

SCHOLARLY PUBLICATIONS

*A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT*

@DTU

(40th Edition)

APRIL 2016

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY

(FORMERLY *DELHI COLLEGE OF ENGINEERING*)

GOVT. OF N.C.T. OF DELHI

SHAHBAD DAULATPUR, MAIN BAWANA ROAD

DELHI 110042

PREFACE

This is the **Fortieth** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc to make the bulletin more comprehensive.

This issue contains the information published during **April 2016**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A collaborative recommender system enhanced with particle swarm optimization technique, **3.Rahul Katarya** and **3.Om Prakash Verma**, CSE, DTU
2. An Optimal Fuzzy System for Edge Detection in Color Images using Bacterial Foraging Algorithm, **3.Om Prakash Verma** and **3. Anil Singh Parihar**, ECE, DTU
3. An Optimized and Cost Efficient Realization Of Reversible Braun Multiplier, **3.Neeta Pandey**, **6.Nalin Dadhich** and **6.Mohd. Zubair Talha**, ECE, DTU
4. An Overview - Tribological Aspects at the Interface of Piston Ring and Cylinder Liner for Improving the Performance of IC Engine, **3.R. C. Singh**, R. K. Pandey and **3.S. Maji**, Mechanical Engineering, DTU
5. Atomic Force Microscopy Studies of Homoepitaxial GaN Layers Grown on GaN Template by Laser MBE, B.S.Choudhary, **3. A. Singh**, S.Tanwar, **3.P.K.Tyagi**, M. Senthil Kumar and S.S. Kushvaha, Applied Physics, DTU
6. Blood Generation from Stem Cells: An Overview, **3.Vimal KishorSingh**, **3.Abhishek Saini**, **3.Neeraj Kumar** and **3.Manisha Kalsan**, Biotechnology, DTU
7. CFD Analysis of Angle Axial Annular Diffuser with Both Hub and Casing at Diverging Equal Angles, **3. B.B.Arora**, Mechanical Engineering, DTU
8. Communication and Cyber Security issues in Smart Grid, **3.M.M. Tripathi**, Electrical Engineering, DTU
9. Contingent Demand Fulfilment Decision Problem from Supplier's Perspective, **3.Akhilesh Kumar**, **3.Anjana Gupta** and **3.Aparna Mehra**, Applied Mathematics, DTU
10. Effect of Different Media and Growth Hormones on Shoot Multiplication of In Vitro Grown Centella asiatica Accessions, **3.Arпита Roy**, **3.Koyel Kundu**, **3.Gaurav Saxena**, **3.Lakhan Kumar** and **3.Navneeta Bharadvaja**, Biotechnology, DTU

11. Effect of structured parameters on hot-carrier immunity of transparent gate recessed channel (TGRC) MOSFET, *3.Ajay Kumar, 3.Neha Gupta and 3.Rishu Chauja*, Applied Physics, DTU
12. E Learning on Clouds: A profitable and Efficient Business Model, *3.Rajni Jindal, and 3.Alka Singhal*, DTU
13. Enabling information recovery with ownership using robust multiple watermarks, Vidhi Khanduja, Shampa Chakraverty and *3.Om Prakash Verma*, CSE, DTU
14. Enhanced dielectric properties of multilayered BiFeO₃/BaTiO₃ capacitors deposited by pulsed laser deposition, *3.Savita Sharma*, Monika Tomar, Nitin K. Puri and Vinay Gupta, Applied Physics, DTU
15. Enhancement of power system damping using Fuzzy Power System Stabilizer, *3.J.N.Rai* and Ram Bhagat, Electrical Engineering, DTU
16. Examining the possibility of achieving inclusive growth in India through corporate social responsibility, *3.Upali Arijita Biswas, 3.Suresh Garg* and Archana Singh, DSM, DTU
17. Excitation of Electromagnetic Surface Waves at A Conductor-Plasma Interface by An Electron Beam Research Article, Ved Prakash, Ruby Gupta, Vijayshri, and *3.Suresh C. Sharma*, Applied Physics, DTU
18. Experimental Investigations On Fluid Flow Characteristics of Single Phase Liquid Through Circular Micro channels, *7.1.Ravindra Kumar*, Mohd. Islam and M.M.Hasan, Thermal Engineering, DTU
19. Facile synthesis of higher manganese silicide employing spark plasma assisted reaction sintering with enhanced thermoelectric performance, *3.Saravanan Muthiah, 3.R.C. Singh, 3.B.D. Pathak*, Ajay Dhar, Mechanical Engineering, DTU

20. Forecasting Volatility Using GARCH: A Case Study, **3.Nand Kumar**, Applied Science & Humanitie, **3.Rishabh Verma** and **3.Puneet Gupta**, Mathematics DTU
21. FPGA Accelerated Abandoned Object Detection, Rajesh Rohilla, **3.Aman Raj**, **3.Saransh Kejriwal** and **3.Rajiv Kapoor**, ECE, DTU
22. Human Genome - Underlying Scientific Mysteries, **3.Shivani Kachroo** and Sivakumar JT Gowder, DTU
23. Jain–Durrmeyer operators associated with the inverse Pólya–Eggenberger distribution, **3.Minakshi Dhamija** and **3.Naokant Deo**, Applied Mathematics, DTU
24. MRI Brain Image Classification Using Polynomial Kernel Principal Component Analysis With Neural Network, **3.Mahipal Singh Choudhry**, Shikha Misra, **3.Rajiv Kapoor**, ECE, DTU
25. Optimization of Process Parameters with Effect of Thrust and Torque in Drilling Operation, **3.Ashish Tripathi**, Ranganath M Singari, Vipin Dahiya, Mechanical, Production and Industrial Engineering, DTU
26. Photovoltaic-grid hybrid power fed pump drive operation for curbing the intermittency in PV power generation with grid side limited power conditioning **3.Amritesh Kumar** and **3.Vishal Verma**, Electrical Engineering, DTU
27. Potential application of carbon nanotube core as nanocontainer and nanoreactor for the encapsulated nanomaterial, **3.Pawan K. Tyagi**, **3.Reetu Kumari**, Umananda M. Bhatta, Raghavendra Rao Juluri and Ashutosh Rath, Sanjeev Kumar, P.V. Satyam, Subodh K. Gautam, Fouran Singh, Applied Physics, DTU
28. Prediction of Tool-Life Model for Heavy Machining Using Central Composite Design, **3.Vipin and Priya**, Production and Industrial Engineering, DTU

29. Resonant Ion Beam Interaction with Whistler Waves in A Magnetized Dusty Plasma Research Article, Ruby Gupta, Ved Prakash, **3.Suresh C. Sharma**, Vijayshri and D.N. Gupta, Applied Physics, DTU
30. Statefinder diagnosis for holographic dark energy models in modified $f(R, T)$ gravity, **3.C.P. Singh** and **3.Pankaj Kumar**, Applied Mathematics, DTU
31. Study of Vendors Selection Methods for Efficient Supply Chain Management, **3.Rajiv Chaudhary**, **3.R C Singh** and Reema Chaudhary, Mechanical Engineering, DTU
32. Use of Green Technology (F.S.P.) for processing of 99.9% Copper with Carbon Nano Tubes, **3.R.S. Mishra**, Mechanical Engineering, DTU
33. Utilization of an Industrial Waste Product from Thermal Power Plants in Civil Construction, **3.Alok Verma**, Civil Engineering, DTU
34. Voltage Mode Second Order Notch/All – Pass Filter Realization Using OTRA, Rashika Anurag, **3.Neeta Pandey**, Rohan Chandra and **3.Rajeshwari Pandey**, ECE, DTU

1. Chancellor

2. Pro Vice Chancellor

3. Faculty

4. Teaching-cum-Research Fellow

5. Alumni

6. Research Scholar

7. PG Scholar

8. Undergraduate Student

2.1. Ex Pro Vice Chancellor

3.1. Ex Faculty

6.1. Ex Research Scholar

7.1. Ex PG Scholar

8.1. Ex Undergraduate Student

A collaborative recommender system enhanced with particle swarm optimization technique

Rahul Katarya¹  • Om Prakash Verma¹

Received: 25 September 2015 / Revised: 9 March 2016 / Accepted: 18 March 2016
© Springer Science+Business Media New York 2016

Abstract In a web environment, one of the most evolving application is those with recommendation system (RS). It is a subset of information filtering systems wherein, information about certain products or services or a person are categorized and are recommended for the concerned individual. Most of the authors designed collaborative movie recommendation system by using K-NN and K-means but due to a huge increase in movies and users quantity, the neighbour selection is getting more problematic. We propose a hybrid model based on movie recommender system which utilizes type division method and classified the types of the movie according to users which results reduce computation complexity. K-Means provides initial parameters to particle swarm optimization (PSO) so as to improve its performance. PSO provides initial seed and optimizes fuzzy c-means (FCM), for soft clustering of data items (users), instead of strict clustering behaviour in K-Means. For proposed model, we first adopted type division method to reduce the dense multidimensional data space. We looked up for techniques, which could give better results than K-Means and found FCM as the solution. Genetic algorithm (GA) has the limitation of unguided mutation. Hence, we used PSO. In this article experiment performed on Movielens dataset illustrated that the proposed model may deliver high performance related to veracity, and deliver more predictable and personalized recommendations. When compared to already existing methods and having 0.78 mean absolute error (MAE), our result is 3.503 % better with 0.75 as the MAE, showed that our approach gives improved results.

Keywords Recommender system · Computational intelligence · Fuzzy C-means · Collaborative filtering · Movie

✉ Rahul Katarya
rahulkatarya@dtu.ac.in

Om Prakash Verma
opverma.dce@gmail.com

¹ Department of Computer Science & Engineering, Delhi Technological University, Delhi, India

1 Introduction

In recent times, we have seen numerous applications of recommendation system. The “People you may know” column suggest friends on Facebook and the “recommended channel” column in YouTube recommend the videos according to our interests and past historical searches. These are some of the examples of recommendation used in web network technology. Movie recommendation applications are widely used nowadays which are amalgamated with web and multimedia devices that adopted operating systems such as windows and android. A recommender system is an information filtering technology which is used to give useful recommendations to a group of users for items or services that might interest them [4, 28]. It is based on a utility function that automatically predicts how a user will like an item, depending upon their past behavior, relations with other users, item similarity, context, and various other factors [23, 26]. The recommendation is an instance of data mining where data sets are discovered and arranged in large patterns. Recommendation systems use different techniques, but two of them used mostly, one is content-based (CB) which examine properties of the items recommended [35]. CB based recommender systems deliver the description of the item and knowledge of the user’s preference; another one is collaborative filtering systems that recommend items based on similarity measures among users and items. It builds a model from a user’s past behavior such as items previously purchased; numerical ratings were given to those items [16, 17, 23, 27, 46]. Collaborative filtering (CF) is based on the theory that people who agreed on same decisions in the past will agree with similar decisions in the future, and they will like similar items as they liked in the past [11, 17, 29, 41]. However, it faces various challenges and limitations such as data sparsity [20, 45, 46], whose role is to the evaluation of large itemsets. Another limitation is hard to make predictions based on nearest neighbor algorithms, third is scalability in which numbers of users and numbers of items both increases [18, 19, 25, 41] and last one is the cold start where poor relationship among like-minded people [23, 40, 46]. To address encounters, above-mentioned, a hybrid model based movie recommendation approach is proposed to improve the concerns of both high dimensionality and data sparsity. We moved to other approaches of collaborating filtering, and we landed up on model-based collaborative filtering [20, 29, 33]. It is memory based and uses the entire user-item database to generate prediction and also uses a statistical technique to find nearest neighbors. The clustering technique of model-based collaborative filtering is used because clustering reduces the high dimension of data sparse problem and improves scalability. Clustering group the like-minded people together and thus prevent us from searching the whole separate user space. Then there’s another problem that clustering requires quite large computational time [46]. We use the entire system in two phases: online and offline phase. In offline phase, clustering model results in a relatively small dimensional space and divides active users into different clusters. In online phase, due to predicted ratings of movies, an active user is presented with top-N movie recommendation list. The FCM algorithm is employed to prepare the required clusters because, in FCM a user does not belong to only one cluster but more than one cluster with a different degree of membership, this ensure more accurate recommendation [5, 14, 18, 31]. Besides this, particle swarm optimization (PSO) is adopted, that is an optimization technique to provide better results of FCM. We used PSO instead of GA because it provides a guided change in particles as compared to a random mutation in GA [6–8, 12, 16, 34, 37, 38, 42, 43]. We have used K-means to facilitate basic elements of PSO [8, 14, 15, 24, 30, 36]. We have implemented and perform an experiment on movielens dataset. The results we have obtained have achieved more success than existing

cluster based collaborative filtering method. The rest of this article is organized as follows: Section 2 is dedicated to literature survey which gives details about the related works already done. In section 3 brief overview of PSO, K-Means and FCM are taken into consideration. Fundamentals and methods are also discussed in section 3. In Section 4; we highlighted our proposed framework PSOKM-FCM, which specifies that how the algorithms and techniques behaved in our system. Section 5 specifies the experiment and result performed on movielens dataset, and disused what we have achieved by applying our framework to movielens database. The last section 6 is dedicated to a conclusion and future work with the explanation of achievements and scope of our proposed framework.

2 Related work

Bobadilla et al. [4] performed massive literature survey on recommender systems (RS) and suggested that RS as a valuable tool for the internet that provides relevant information. Nowadays lots of Webs and mobile applications have adopted recommendation approaches for suggesting relevant recommendations. E-commerce and mobile applications now totally depend upon the recommendation techniques and methods. A recommender system is broadly classified into three types of methods: content-based, collaborative based and hybrid based and classified them according to heuristic and model based [1]. Traditional collaborative filtering (TCF) approaches are the most popular approach in a recommendation system research domain. Then, using Pearson correlation coefficient (PCC) measurement, users are grouped such that those who have more similarity are together in a cluster. Next, predicted value of items is calculated and recommended to the user. Then at the end, the top N highest predictive items are selected to recommend the target user [40]. Context based recommender systems are adopting nowadays by various e-commerce industries. In recommendation research authors also applied factorization machines to the context-aware recommender systems [39]. They developed an iterative optimization technique that analytically found the least-square answer for one parameter given the other ones. Factorization Machines (FMs) were applied to model contextual information and deliver context-aware rating predictions. This method resulted in fast context-aware recommendations because the model equation of FMs could be computed in linear time with both the number of context variables and the factorization size. In another research work, a recommender system combined with content-based and collaborative methods was suggested for items of interest to users, and also, exploit item semantics [13]. Authors described a hybrid method in which a user-specific recommendation mechanism was learned that use similarity measures between users and also measures the attributes of items that make them interesting to particular users. A multitask clustering framework was projected for the activity of daily living analysis from visual data gathered from wearable cameras [48]. In their framework, rather than clustering data from different users separately, researchers proposed to look for clustering partitions which were coherent among related tasks. A strategy was suggested that automatically select meaningful semantic concepts for the event detection task based on both the events-kit text descriptions and the concepts high-level feature descriptions [50]. Moreover, authors introduced an event oriented dictionary representation based on the selected semantic concepts. They attempted to learn an event oriented dictionary representation based on the selected semantic concepts. A flexible graph-guided multi-task learning (FEGA-MTL) framework was offered for categorizing the head pose of a person who passages freely in an atmosphere monitored by multiple, large field-of-view surveillance

cameras [47, 49]. The FEGA-MTL framework naturally extended to a weakly supervised setting where the target's walking route is employed as a proxy instead of head orientation. Some researchers also considered the problem of egocentric activity recognition from unlabelled data within a multi-task clustering framework [10]. Two multi-task clustering (MTC) algorithms were proposed and evaluated on first-person vision (FPV) datasets.

3 Material and methods

The collaborative filtering is based on the user preference generated from the Pearson correlation coefficient (PCC), that compute the similarity between n user and other users [28]. Pearson correlation computes the statistical correlation among two item's mutual ratings to determine their similarity. The correlation is calculated as following as shown in Eq. 1.

$$S(u, v) = \frac{\sum_{i \in I_u \cap I_v} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I_u \cap I_v} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in I_u \cap I_v} (r_{v,i} - \bar{r}_v)^2}} \quad (1)$$

Where $S(u, v)$, is the value between -1 and 1 . If a value equals to -1 , there is a perfect negative indication correlation between the two values, vice versa. Here u and v are the users whose similarity has to be calculated. i is the item belonging to a set of items that both the users have rated. $r_{u,i}$ is the rating of the i^{th} item by user u and $r_{v,i}$ is the rating of the i^{th} item by user v . \bar{r}_u is the average rating of user u and \bar{r}_v is the average rating of user v . In [46], it converted the original high dimensional data space into a relatively small dimensional data space for carrying denser information. The main idea is to convert and obtain new coordinate space from the original data, which is denoted by a principal component of data with the largest eigenvalue. Let's assume we have $m \times n$ matrix as the user-rating and in this n -dimension vector specifies user's profile. It turns out then principal components after performing eigenvalue decomposition, and we select the only first d components ($d \leq n$) to keep in the new data space which is based on the value of accumulated proportion of 90 % of the origin alone. According to [2, 9, 15, 21, 24, 30, 36, 44], the simplest clustering algorithm used is k -means, which categorizes items into k clusters. Initially, each k cluster contains random items. Then, for each cluster, a centroid (or center) is computed. The distance of each item from the centroids are then calculated and checked. If an item is discovered to be closer to another cluster, meaning that if the distance between them is less then, it is moved to that cluster. Centroids are calculated, again and again; thus, all item distances are checked. This is done until the stability is reached (that is when no items move in the different cluster during an iteration), and thus algorithm ends. In [46], suggested a common drawback of K -means algorithm is the selection of initial seeds (initial centroid). K -means algorithm aims at reducing an objective function, in this case, a squared error function. The objective function is:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (2)$$

Where $\|x_i^{(j)} - c_j\|^2$ is a selected distance measure among a data point $x_i^{(j)}$ and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centres.

The k-means algorithm is composed of the following steps: a) Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids. b) Assign each object to the group that has the closest centroid. c) When all objects have been assigned, recalculate the positions of the K centroids. d) Repeat Steps b and c until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated. Our experimental outcomes have been compared with the techniques adopted by authors on the same dataset [46]. The initial centroid could affect the final output and can easily lead to produce local optimum. Particle swarm optimization (PSO) was inspired by the social behavior as it is an evolutionary computation technique [22]. In PSO, the possible resolutions called “particles”, fly around in a multi-dimensional search space, to determine an optimal, or sub-optimal, answer by competition as well as support for themselves. The system primarily has a population of random responses. The individual particle is given a random velocity and is drifted through the d-dimensional problem space. The position or location (x_{id}) and velocity (v_{id}) of every particle i in dimension d is revised based on its previous velocity, the previous best particle location (p_{id} or p_{best}), and the previous global best location or position of a particle in the population (p_{gd} or g_{best}). The key approach of PSO lies in accelerating each particle towards its p_{best} and g_{best} locations at each time step. The authorized velocity and position of i th particle are given in Eqs. 3 and 4, respectively [8]:

$$v_{id} = w \times v_{id} + c_1 \times \text{rand}_1() \times (p_{id} - x_{id}) + c_2 \times \text{rand}_2() \times (p_{gd} - x_{id}). \quad (3)$$

$$x_{id} = x_{id} + v_{id} \quad (4)$$

Where w denotes the inertia weight factor and is usually set to a value in the range of 0.5–1.

p_{id} is the location of the particle that understands the best fitness value. p_{gd} is the position of the particles that involve a global best fitness value. c_1 and c_2 are constants recognized as social acceleration coefficients. d signifies the aspect of the problem space. rand_1 , rand_2 are random values in the scope of (0, 1), which ensures wide search through problem space. The inertia weight factor w delivers the essential diversity to the swarm by altering the momentum of particles to escape the stagnation of particles at the local optima. Equation 5 needs each particle to record its existing coordinate x_{id} . The velocity v_{id} specifies the speed of its movement sideways the dimensions in a problem space, and the coordinates p_{id} and p_{gd} . Where the best fitness values were computed [8]:

$$Pi(t+1) = \begin{cases} Pi(t) & f(Xi(t+1)) \leq f(Xi(t)) \\ Xi(t+1) & f(Xi(t+1)) > f(Xi(t)) \end{cases} \quad (5)$$

Where the symbol f represents the fitness function; $P_i(t)$ stands for the finest fitness values and the coordination where the significance was calculated. t denotes the generation step. Finally, PSO algorithm can be outlined as a) Initialize every particle in the population pool within the search space of initial velocities; b) Calculate the fitness value of each of these particles. c) If this value is better than its best fitness value recorded until this time, which is also called as population best (p_{id}), then set the current value as the new p_{id} . d) The global best (p_{gd}) is selected from the candidates, which is the best fitness value of all of them in the pool. This is updated in every iteration after this comparison as the new p_{gd} . e) Calculate the particle velocity according to Eq. 2. f)

Refresh the particle position according to Eq. 3. g) Carry out stages b to e until minimum error criteria are attained. The fuzzy c-means algorithm [3, 5, 14, 31, 32] is a clustering algorithm where each item may belong to two or more clusters, not just one. Here, the degree of membership is considered for each item which is given by a probability distribution over the clusters. Fuzzy c-means (FCM) is a clustering algorithm that is useful when an item's similarity matches with more than one cluster, and we do not want just to consider but various other clusters too [31]. FCM is different in the sense that it does not decide the entire membership of a data point to a given group; instead, it calculates the likelihood (the degree of membership). Since there is no absolute membership in one cluster but the various percentage of membership in different clusters, FCM can be extremely fast. This is because high accurateness can be achieved with a large number of iterations with this clustering method. FCM is grounded on minimization of the following objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m < \infty \quad (6)$$

Where m is some real number larger than 1, u_{ij} is the degree of association of x_i in the cluster j . x_i is the i th of d -dimensional dignified data, c_j is the d -dimension center of the cluster. $\|*\|$ is somewhat norm expressing the similarity between any measured data and the center. Fuzzy partitioning is supported out by an iterative optimization of the objective function mentioned above, with the apprise of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m} \quad (7)$$

This iteration will stop when, $\max_{ij} \{ |u_{ij}^{(k+1)} - u_{ij}^k| \} < \varepsilon$ where ε is a termination criterion between 0 and 1, whereas k is the iteration steps. This process converges to a local minimum or a saddle point of J_m .

4 PSO-KM-FCM based collaborative filtering framework

In this section, our aim is to develop a hybrid clustering model to improve movie prediction accuracy and recommendation to users. We have used some well-known clustering algorithm such as fuzzy c-mean, k-means clustering algorithm and PSO along with an approach named 'type division' method. Initial data contains two parameters 'movie id' and its 'rating' by a user. Type division method distributes the ratings given over the types that particular movie belongs. Then they are used as initials to a k-mean algorithm which outputs the centers for particular algorithm fuzzy c-mean algorithm which need optimal initial centers for better results, the centers produced by k means are further optimised by using them in fuzzy c-mean. After processing fuzzy c-mean algorithm, final results are combined back and used for predicting ratings for futures

movies given by a particular user based on previous data. As shown in Fig. 1, we developed type division method for movies and adopted particle swarm optimization (PSO). In our proposed recommendation system, we have developed a master collaboration of PSO, K-means and fuzzy c-means. We have converted our initial dataset to a new form in which users are divided based on types of movies they watched. Once we have our new dataset we use it on PSO and K-means combination to find initial centres, which are much accurate and precise centres than the assigning of centres randomly, these centres are used by fuzzy c-means for optimization. Type division method is a fundamental approach that divides users according to their interest in different kinds of movies; it uses initial data set generate 19 files, each containing users having an interest in that particular type of movie. A single movie can belong to many types or categories. After generation of files, we use our algorithm on each file to combine users more precisely by their rating on particular movie belonging to that type. By dividing users according to the categories or type of movies, we find particular user interest and their rating for that particular type. The type division method simply segregated the dataset making the recommendation system more accurate for calculation.

Type division method was applied to the database and that had no effect on K-means being applied first except that K-Means had to be applied for each movie kind in the dataset. The initial centres and clusters given by K-Means provided an alternate to the random allocation of input to PSO for optimized center calculation. The final output of K-Means and PSO was a set of centres for each database file. The centres given by K-Means and PSO were used by FCM algorithm to form the final clusters which were directly used for result calculation process. Figure 1 shows that our approach for proposed movie recommendation system. It initially involves breaking the dataset movies into the 19 types given and applying a series of clustering and optimization algorithms explained below.

4.1 Adopting K-means clustering for centres

K-means algorithm is one of the most popular and commonly used clustering algorithms because of factors like flexibility, simplicity and computation efficiency which arises when considering large amounts of data. K-means one by one calculate 'k' cluster centres to assign users to nearest cluster based on the distance calculated and applied by the Type division method in which movies types are categorized according to their users. When there is no more change in centres, the code comes to convergence. For the movie and their users, K-means algorithm aims to partition these users into 'k' groups automatically.

$$J = \sum_{j=1}^k \sum_{i=1}^n \left\| \mathbf{x}_i^{(j)} - \mathbf{C}_j \right\|^2 \quad (8)$$

Where J = objective function, k = number of clusters, n = number of cases, \mathbf{C}_j = centroid for cluster j , $\mathbf{x}_i^{(j)}$ = case i , $\left\| \mathbf{x}_i^{(j)} - \mathbf{C}_j \right\|^2$ = Distance function.

- (a) Clusters the Movie data into k groups.
- (b) Select k points at random, as cluster centres (\mathbf{C}_j where $j = 1, 2..k$).
- (c) Assign objects to their closest cluster centre according to the Euclidean distance function.
- (d) Compute the centroid or mean of all objects in each cluster.
- (e) Repeat steps b, c and d until the same points are assigned to each cluster in consecutive rounds.

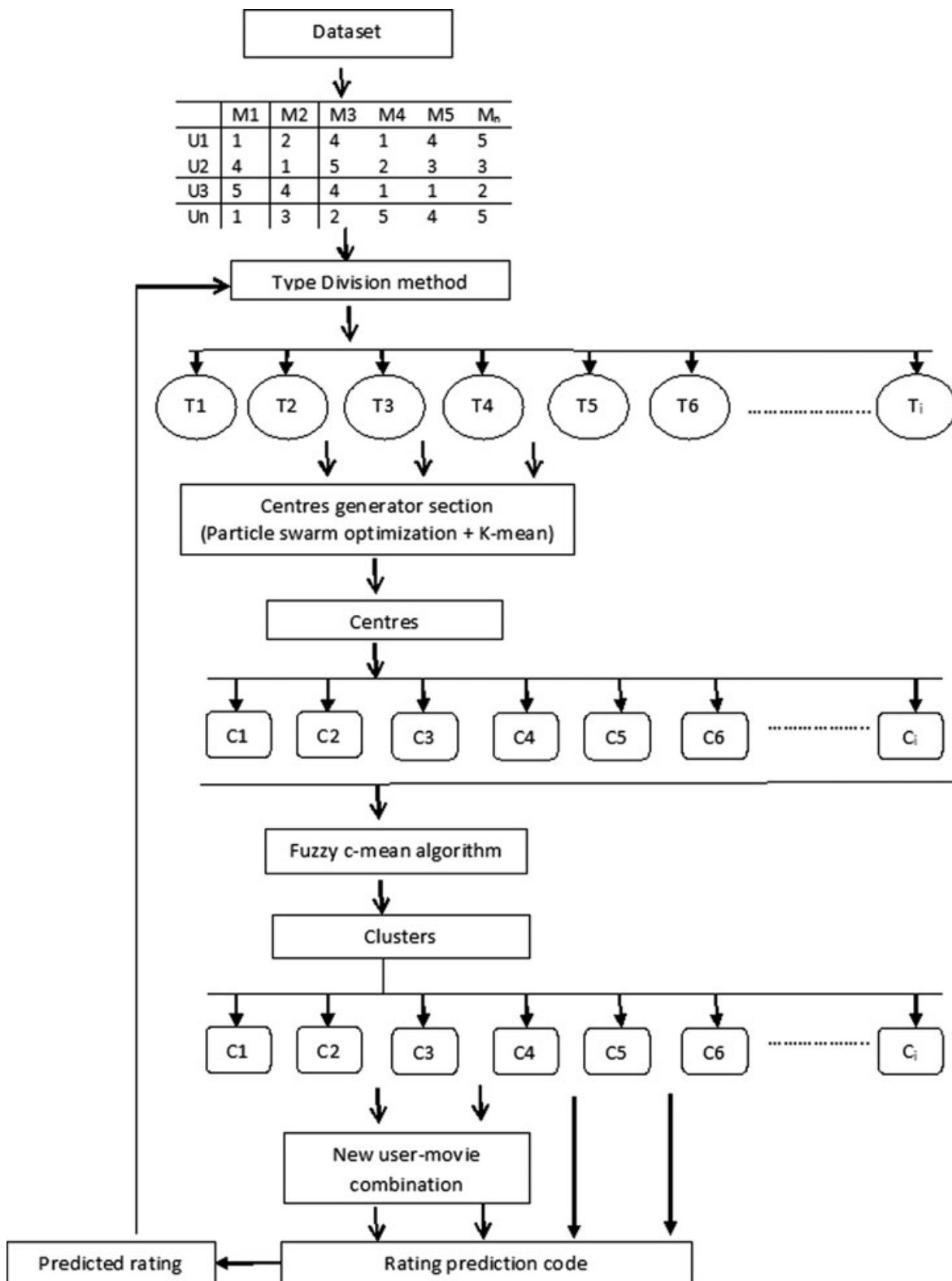


Fig. 1 PSO-KM-FCM based collaborative filtering framework

4.2 Particle swarm optimization application for movie

Let S be the number of users in the swarm, each having a position $x_i \in \mathbb{R}^n$ in the defined search-space and having a velocity $v_i \in \mathbb{R}^n$. Let p_i is the best known position of user i and g is best position of the whole swarm. A PSO algorithm application for movie is shown in Fig. 2.

- For every user $i = 1, \dots, S$ do:
 - Initialize the user's position with a random vector: $\mathbf{x}_i \sim U(\mathbf{b}_l, \mathbf{b}_u)$, where \mathbf{b}_l and \mathbf{b}_u are the lower and upper bound of the search-space.
 - Initialize the user's position to its initial position: $\mathbf{p}_i \leftarrow \mathbf{x}_i$
 - If $(f(\mathbf{p}_i) < f(\mathbf{g}))$ refresh the swarm's finest known position: $\mathbf{g} \leftarrow \mathbf{p}_i$
 - Initialize the user's property: $\mathbf{v}_i \sim U(|\mathbf{b}_u - \mathbf{b}_l|, |\mathbf{b}_u - \mathbf{b}_l|)$
 - Do while terminating criterion is not reached (e.g. max number of times then code run or a solution to satisfying objective function value is found):
 - For every user $i = 1, \dots, S$ do:
 - take random numbers: $r_p, r_g \sim U(0,1)$
 - For every dimension $d = 1, \dots, n$ do:
 - Update the user's property: $\mathbf{v}_{i,d} \leftarrow \omega \mathbf{v}_{i,d} + \varphi_p r_p (\mathbf{p}_{i,d} - \mathbf{x}_{i,d}) + \varphi_g r_g (\mathbf{g}_d - \mathbf{x}_{i,d})$
 - Update the user's position: $\mathbf{x}_i \leftarrow \mathbf{x}_i + \mathbf{v}_i$
 - If $(f(\mathbf{x}_i) < f(\mathbf{p}_i))$ do:
 - Update the user's best-known position: $\mathbf{p}_i \leftarrow \mathbf{x}_i$
 - If $(f(\mathbf{p}_i) < f(\mathbf{g}))$ update the swarm's greatest known position: $\mathbf{g} \leftarrow \mathbf{p}_i$
- Now \mathbf{g} holds the best-found solution.

Fig. 2 Particle swarm optimization application for the movie

The parameters φ_p , ω and φ_g , are nominated by the practitioner. These parameters control the behaviour and efficiency of the PSO algorithm.

4.3 Application of fuzzy c-means and rating prediction

This FCM works by allowing membership to each data point correlating to each cluster center by the distance among the cluster center and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, the summation of membership of each data point should be equal to one. After each iteration membership and cluster, centers are updated according to users and set of centers in Eq. 9, and steps are:

- a) Take cluster centers generated by K-Mean-PSO code.
- b) Calculate the fuzzy membership μ_{ij} using Eq. 9:

$$\mu_{ij} = 1 / \sum_{k=1}^c (d_{ij} / d_{ik})^{\left(\frac{2}{m}-1\right)} \quad (9)$$

Where, μ_{ij} =fuzzy membership, c = number of cluster centre, d_{ij} =Euclidean distance between i^{th} data and j^{th} cluster center, 'm' is the fuzziness index $m \in [1, \infty]$.

- c) Calculate the fuzzy centres v_j with Eq. 10:

$$v_{ij} = \left(1 / \sum_{i=1}^n (\mu_{ij})^m x_i \right) / \left(\sum_{i=1}^n (\mu_{ij})^m \right), \forall j = 1, 2, \dots, C \quad (10)$$

- d) Repeat step b and c until the minimum 'J' value is achieved

$$\|U^{(k+1)} - U^{(k)}\| < \beta. \text{ Where,}$$

k is the iteration step. β is the termination criterion which is between $[0, 1]$.

$U = (\mu_{ij})_{n \times c}$ is the fuzzy membership matrix. J is the objective function.

We divided that movie in their types then finds a user based on the previous rating in corresponding types, and then all this rating are combined to get a final predicted rating.

5 Experiments and results

In this section, we discuss the experimental design and empirically investigate our novel proposed movie recommendation system via K-Means-PSO-FCM technique. We evaluate the performance of the proposed method using Mean Absolute Error. Finally, the results are analyzed and discussed here. We carried out all of our experiments and techniques on Pavilion DV6 2.6 GHz, 6.0 GB RAM and Java using Netbeans IDE 7.3.1 to simulate the model.

5.1 Data set and evaluation criteria

We consider the Movielens dataset to conduct the experiments, which is available online including 100,000 ratings by 943 users on 1682 movies and assigned to a discrete scale of 1–5 (<http://grouplens.org/datasets/movielens/100k/>). Every user has rated at least 20 movies. We divide the dataset into 19 types of movies given for each file in the dataset. We have taken two features for identifying the user preferences as input to recommendation system, and they are ‘movies id’ that tells the categories or types it represents and ‘ratings’ given to the movie. We have 19 different kinds in which all the movies gets categories. Each user rating for a particular movie is distributed among the types that movie belongs to. The distribution task is done by the ‘type division’ method, which goes for individual clustering and gets combined again at the end of final results. We implemented our proposed system on the dataset to make the predictions using the final clusters. To check the quality of recommendation, we used the mean absolute error (MAE) as evaluation measure which has been widely used to compare and measure the performance of recommendation systems. The MAE is a statistical accuracy metric which measures the average of the absolute difference between the predicted ratings according to the technique used and actual ratings on test users as shown in Eq. 10. A lower MAE value corresponds to a more accurate prediction [46].

$$\text{MAE} = \frac{\sum |\tilde{P}_{ij} - r_{ij}|}{M} \quad (11)$$

Where M is the total number of predicted movies, P_{ij} represents the predicted value for user i on item j and r_{ij} is the true rating. To understand whether users are interested in the recommendation movies, we can employ the precision and recall metrics which are widely used in recommender systems to understand intelligence level of recommendations. Precision is the ratio of interesting movies retrieved by a recommendation method to the number of recommendations. Recall gives the ratio of interesting movies retrieved that is considered interesting by the user in the actual system. These two measures clearly conflict in nature because increasing the size of recommended movies N leads to an increase in the recall but decrease the precision. However, these evaluations can only be measured on a real online recommender system. The precision and recall for Top-N (N is the number of predicted movies) recommendations are as follows [4]:-

5.2 Results and discussion

The sparse of user-item rating matrix makes it hard to find real neighbours to form the final recommendation list. In our experiments, we compare the performances and some trends of the existing baseline CF movie recommendation systems with our approach, while the neighbourhood size varies from 5 to 60 in an increment of 5. A detailed result is mentioned in Table 1, as it demonstrates efficient experimental results. The mean absolute error is calculated for K-Means-PSO-FCM on different dataset files. We first calculate the MAE using the above-given method on various dataset files namely u1.base, u2.base, u3.base, u4.base, u5.base, ua.base and ub.base. These files contain a different number of movies. However, we get the mean absolute error for each dataset which is later combined to find the MAE of the entire MovieLens dataset.

Table 1 shows the values of mean absolute error for our approach on the MovieLens dataset.

Figure 3 shows the variation of mean absolute error over different files of the dataset and after that we calculated the Mean, Variation and the standard deviation of the MAE calculated: Mean = 0.7458933, Variance = 3.78615 e-5, Standard Deviation = 6.15317 e-3. Mean absolute error comparison between different techniques. We first try to evaluate the movie recommendation quality with the traditional cluster-based CFs. We compare different methods based on the MAE calculated with a different number of neighbours ranging from 5 to 60.

Figure 4 shows the variations in mean absolute error for a different number of neighbours in clusters for various methods. It shows that all methods try to reach the optimum prediction values where the neighbourhood size varies from 15 to 20, and it becomes relatively stable around 60 nearest neighbours. Without the first step of dimensional reduction, GA-KM and SOM gave very close MAE values, and it seems that GAKM produces slightly better prediction than SOM. When coupling with PCA technique, GAKM shows a distinct improvement in recommendation accuracy compared with SOM. Moreover, the proposed K-Means-PSO-FCM produces the smallest MAE values continually where the neighbour size varies. All K-means clustering CF generate increasing MAE values which indicate the decreasing quality for recommendation due to the sensitiveness of the algorithm. Traditional user-based CF produces relatively worse prediction compared with the basic clustering-based methods. However, our method gave the lowest MAE amongst all methods shown above because of the fuzzy logic used.

Table 2 shows the comparison between the variations in the mean absolute error calculated by different techniques by comparing their mean and standard deviations. Our proposed method has the best MAE among all the methods used and a reasonable variation still keeping the value below all the other methods used. Our work is novel in the sense that, this hybrid of PSO, K-Means and Fuzzy-c-means delivered better

Table 1 MAE of separate files of dataset

Dataset	Mean absolute error
U1.base	0.755745
U2.base	0.745370
U3.base	0.738844
U4.base	0.739792
U5.base	0.744568
Ua.base	0.742731
Ub.base	0.754202

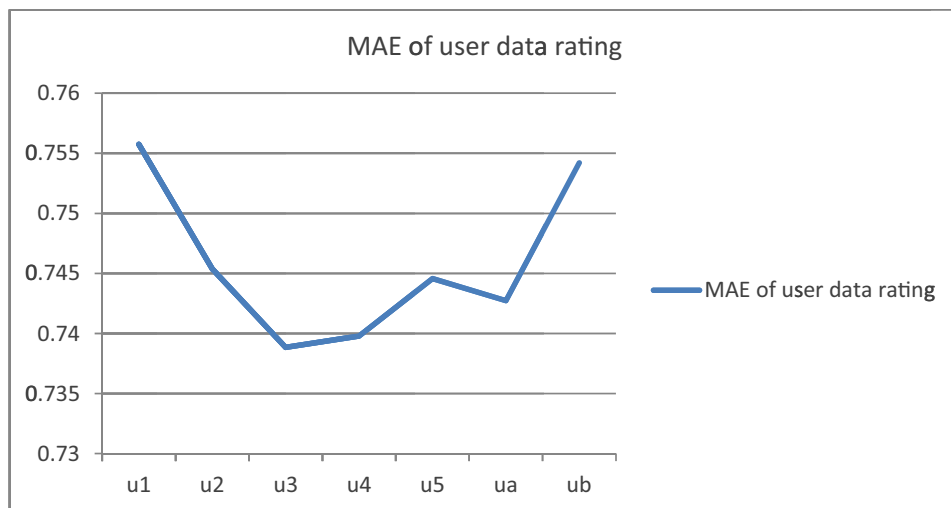


Fig. 3 MAE of separate files of the dataset

results. We initially use FCM, but to make it more efficient, we needed an optimization algorithm. The centers obtained by PSO was used as input for FCM. PSO needs initial clusters and initial centers. For initial clusters, we made random clusters with no similarity factor. For leading centers, we took the average of the no. of movie and rating values and used them as centers. However, the differences between these centers were too small. This produced poor results. So we used K-Means, and the clusters of K-means are used as initial clusters of PSO, and the average of the values is taken as first centers for PSO. We found that the combination of FCM with PSO and K-Means gives lower MAE values than other recommendation systems used before.

6 Conclusion and future work

In this paper, we developed a novel hybrid model based on a collaborative filtering (CF) approach that produces movie recommendations in which we developed type division system for movies and adopted this with particle swarm optimization and clustering algorithms as

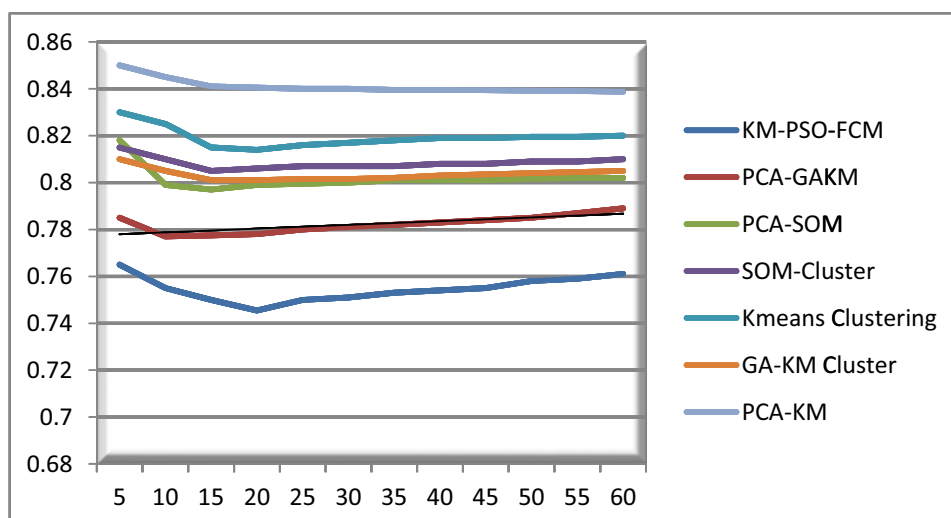


Fig. 4 MAE comparison between different techniques

Table 2 MAE comparison between different techniques

Method	Mean	Standard deviation
Proposed	0.7547	5.067 e-03
PCA-GAKM	0.7821	4.747 e-03
PCA-SOM	0.8018	5.900 e-03
SOM-clustering	0.8078	2.589 e-03
KM-clustering	0.8175	2.244 e-03
GA-KM clustering	0.8040	2.786 e-03
PCA-KM	0.8412	2.845 e-03

fuzzy c-mean and with K-mean algorithm application. In our approach, we divided user-movie according to their movie type, which in return make our dataset more precise, dense, clear and reliable. We used fuzzy c-means (FCM) as the main algorithm for finding a neighborhood for users. To improve the accuracy of FCM we used PSO and K-means algorithm, which give more precise centers for fuzzy c-means and give denser neighborhood selections. We have done the experiment on Movielens dataset; this dataset has many categories. Note that this dataset will only have the ratings of movies rated by users. Users are fully independent and do not have social relationships. Our proposed approach proves that it is capable of generating better prediction accuracy and provide more reliable movie recommendations to users comparing to the present clustering-based collaborative filtering (CCF) approaches. As for future work, we will continue to improve our work approach that can deal with a much larger dataset with higher dimension or attribute. We will add other attributes of users like age, occupation, etc. to give more accurate and reliable rating.

References

1. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans Knowl Data Eng* 17:734–749. doi:[10.1109/TKDE.2005.99](https://doi.org/10.1109/TKDE.2005.99)
2. Ahmad A, Dey L (2007) A k-mean clustering algorithm for mixed numeric and categorical data. *Data Knowl Eng* 63:503–527. doi:[10.1016/j.datak.2007.03.016](https://doi.org/10.1016/j.datak.2007.03.016)
3. Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy *c*-means clustering algorithm. *Comput Geosci* 10:191–203. doi:[10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
4. Bobadilla J, Ortega F, Hernando A, Gutiérrez A (2013) Recommender systems survey. *Knowl Based Syst* 46:109–132. doi:[10.1016/j.knosys.2013.03.012](https://doi.org/10.1016/j.knosys.2013.03.012)
5. Cannon RL, Dave JV, Bezdek JC (1986) Efficient implementation of the fuzzy *c*-means clustering algorithms. *IEEE Trans Pattern Anal Mach Intell* 8:248–255
6. Chen C, Ye F (2004) Particle swarm optimization algorithm and its application to clustering analysis. *IEEE Netw Sens Control* 2:789–794
7. Chen D, Zhao C (2009) Particle swarm optimization with adaptive population size and its application. *Appl Soft Comput* 9:39–48. doi:[10.1016/j.asoc.2008.03.001](https://doi.org/10.1016/j.asoc.2008.03.001)
8. Cui X, Potok TE (2005) Document clustering analysis based on hybrid PSO+K-means algorithm. *Engineering* 185–191
9. Ding C, He X (2004) K -means clustering via principal component analysis. *Twentyfirst Int Conf Mach Learn ICML 04 Cl*:29. doi: [10.1145/1015330.1015408](https://doi.org/10.1145/1015330.1015408)
10. Eccv A (2014) Recognizing daily activities from first-person videos with multi-task clustering. *Accv2014*. doi: [10.1007/978-3-319-16817-3_34](https://doi.org/10.1007/978-3-319-16817-3_34)
11. Ekstrand MD (2010) Collaborative filtering recommender systems. *Found Trends® Hum Comput Inter* 4: 81–173. doi:[10.1561/1100000009](https://doi.org/10.1561/1100000009)

12. Fourie PC (2002) The particle swarm optimization algorithm in size and shape optimization. *Struct Multidiscip Optim* 23:259–267. doi:[10.1007/s00158-002-0188-0](https://doi.org/10.1007/s00158-002-0188-0)
13. Garden M, Dudek G (2006) Mixed collaborative and content-based filtering with user-contributed semantic features. *Proc Natl Conf Artif Intell* 21:1307
14. Ghosh S, Dubey S (2013) Comparative analysis of K-means and fuzzy c-means algorithms. *IJACSA* 4:35–39. doi:[10.14569/IJACSA.2013.040406](https://doi.org/10.14569/IJACSA.2013.040406)
15. Hartigan JA, Wong MA (1979) A K-means clustering algorithm. *J R Stat Soc* 28:100–108
16. Hassan R, Cohanin B (2005) A comparison of particle swarm optimization and the genetic algorithm. 1st AIAA Multidiscip Des Optim Spec Conf 1–13. doi: [10.2514/6.2005-1897](https://doi.org/10.2514/6.2005-1897)
17. Herlocker J, Konstan J (1999) An algorithmic framework for performing collaborative filtering. *Proc 22nd ... 8*. doi: [10.1145/312624.312682](https://doi.org/10.1145/312624.312682)
18. Jafar OAM, Sivakumar R (2013) A comparative study of hard and fuzzy data clustering algorithms with cluster validity indices. *Int Conf Emerg Res Comput Inf Commun Appl*
19. Jones MT (2013) Recommender systems, part 2: introducing open source engines explore software for building a recommendation capability. 1–13
20. Kannan R, Ghinea G, Swaminathan S (2015) What do you wish to see? A summarization system for movies based on user preferences. *Inf Process Manag* 51:286–305. doi:[10.1016/j.ipm.2014.12.001](https://doi.org/10.1016/j.ipm.2014.12.001)
21. Kanungo T, Mount DM, Netanyahu NS et al (2002) An efficient k-means clustering algorithm: analysis and implementation. *IEEE Trans Pattern Anal Mach Intell* 24:881–892. doi:[10.1109/TPAMI.2002.1017616](https://doi.org/10.1109/TPAMI.2002.1017616)
22. Kennedy J, Eberhart R (1995) Particle swarm optimization. *IEEE International Conference on Neural Networks, Proceedings* 4:1942–1948. doi:[10.1109/ICNN.1995.488968](https://doi.org/10.1109/ICNN.1995.488968)
23. Konstan JA, Riedl J (2012) Recommender systems: from algorithms to user experience. *User Model User Adapt Inter* 22:101–123. doi:[10.1007/s11257-011-9112-x](https://doi.org/10.1007/s11257-011-9112-x)
24. Krishna K, Murty MN (1999) Genetic K-means algorithm. 29:433–439
25. Lekakos G, Caravelas P (2008) A hybrid approach for movie recommendation. *Multimed Tools Appl* 36:55–70. doi:[10.1007/s11042-006-0082-7](https://doi.org/10.1007/s11042-006-0082-7)
26. Lü L, Medo M, Yeung CH et al (2012) Recommender systems. 97
27. Man KF, Tang KS, Kwong S (1996) Genetic algorithms: concepts and applications. *IEEE Trans Ind Electron* 43:519–534. doi:[10.1109/41.538609](https://doi.org/10.1109/41.538609)
28. Melville P, Sindhvani V (2010) Recommender systems. *Encycl Mach Learn* 829–837. doi: [10.1162/153244302760200641](https://doi.org/10.1162/153244302760200641)
29. Mukherjee R, Sajja N, Sen S (2003) A movie recommendation system—an application of voting theory in user modeling. *User Model User Adap Inter* 13:5–33. doi:[10.1023/A:1024022819690](https://doi.org/10.1023/A:1024022819690)
30. Nazeer K (2009) Improving the accuracy and efficiency of the K-means clustering algorithm. *Proc World Congr I*:1–5
31. Pal R, Nikhil BCJ (1995) On cluster validity for the fuzzy c-means model.pdf. 370–379
32. Pal NR, Pal K, Keller JM, Bezdek JC (2005) A possibilistic fuzzy c-means clustering algorithm. *IEEE Trans Fuzzy Syst* 13:517–530. doi:[10.1109/TFUZZ.2004.840099](https://doi.org/10.1109/TFUZZ.2004.840099)
33. Park DH, Kim HK, Choi IY, Kim JK (2012) A literature review and classification of recommender systems research. *Expert Syst Appl* 39:10059–10072
34. Parsopoulos KE, Vrahatis MN (2008) Multi-objective particles swarm optimization approaches. *IGI Glob* 20–42. doi: [10.4018/978-1-59904-498-9.ch002](https://doi.org/10.4018/978-1-59904-498-9.ch002)
35. Pera MS, Ng Y-K (2013) A group recommender for movies based on content similarity and popularity. *Inf Process Manag* 49:673–687. doi:[10.1016/j.ipm.2012.07.007](https://doi.org/10.1016/j.ipm.2012.07.007)
36. Pham DT, Dimov SS, Nguyen CD (2005) Selection of K in K-means clustering. *Proc Inst Mech Eng Part C J Mech Eng Sci* 219:103–119. doi:[10.1243/095440605X8298](https://doi.org/10.1243/095440605X8298)
37. Poli R (2008) Analysis of the publications on the applications of particle swarm optimisation. *J Artif Evol Appl* 2008:1–10. doi:[10.1155/2008/685175](https://doi.org/10.1155/2008/685175)
38. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. *Swarm Intell* 1:33–57. doi:[10.1007/s11721-007-0002-0](https://doi.org/10.1007/s11721-007-0002-0)
39. Rendle S, Gantner Z, Freudenthaler C, Schmidt-Thieme L (2011) Fast context-aware recommendations with factorization machines. *Proc 34th Int ACM SIGIR Conf Res Dev Inf* 635–644. doi: [10.1145/2009916.2010002](https://doi.org/10.1145/2009916.2010002)
40. Rosli AN, You T, Ha I et al (2014) Alleviating the cold-start problem by incorporating movies facebook pages. *Clust Comput* 18:187–197. doi:[10.1007/s10586-014-0355-2](https://doi.org/10.1007/s10586-014-0355-2)
41. Shani G, Gunawardana A (2011) Evaluating recommendation systems. *Recomm Syst Handb* 257–297
42. Shi Y, Eberhart R (1999) Empirical study of particle swarm optimization. *Proc Congr Evol Comput* 1945–1950. doi: [10.1109/CEC.1999.785511](https://doi.org/10.1109/CEC.1999.785511)
43. Van Der Merwe DW, Engelbrecht AP (2003) Data clustering using particle swarm optimization. *Congr Evol Comput*. doi:[10.1109/CEC.2003.1299577](https://doi.org/10.1109/CEC.2003.1299577)

44. Wagstaff K, Cardie C, Rogers S, Schroedl S (2001) Constrained K-means clustering with background knowledge. *Int Conf Mach Learn* 577–584. doi: [10.1109/TPAMI.2002.1017616](https://doi.org/10.1109/TPAMI.2002.1017616)
45. Wang L-C, Meng X-W, Zhang Y-J (2012) Context-aware recommender systems. *J Softw* 23:1–20. doi: [10.3724/SPJ.1001.2012.04100](https://doi.org/10.3724/SPJ.1001.2012.04100)
46. Wang Z, Yu X, Feng N, Wang Z (2014) An improved collaborative movie recommendation system using computational intelligence. *J Vis Lang Comput* 25:667–675. doi: [10.1016/j.jvlc.2014.09.011](https://doi.org/10.1016/j.jvlc.2014.09.011)
47. Yan Y, Member S, Ricci E (2007) A multi-task learning framework for head pose estimation under target motion. 6:1–15. doi: [10.1109/TPAMI.2015.2477843](https://doi.org/10.1109/TPAMI.2015.2477843)
48. Yan Y, Ricci E, Liu G et al (2015) Egocentric daily activity recognition via multitask clustering. *Image Process IEEE Trans* 24:2984–2995. doi: [10.1109/TIP.2015.2438540](https://doi.org/10.1109/TIP.2015.2438540)
49. Yan Y, Ricci E, Subramanian R et al (2013) No matter where you are: flexible graph-guided multi-task learning for multi-view head pose classification under target motion. *Proc IEEE Int Conf Comput Vis* 1177–1184. doi: [10.1109/ICCV.2013.150](https://doi.org/10.1109/ICCV.2013.150)
50. Yan Y, Yang Y, Meng D et al (2015) Event oriented dictionary learning for complex event detection. 24: 1867–1878



Rahul Katarya is working as an Assistant Professor in Department of Computer Science & Engineering in Delhi Technological University, Delhi India.



Dr. O. P. Verma is HOD and Professor in Department of Computer Science & Engineering in Delhi Technological University, Delhi India.

An Optimal Fuzzy System for Edge Detection in Color Images using Bacterial Foraging Algorithm

Om Prakash Verma, *Member, IEEE*, Anil Singh Parihar

Abstract—This paper presents a fuzzy system for edge detection, using Smallest Univalued Segment Assimilating Nucleus (SUSAN) principal and Bacterial Foraging Algorithm (BFA). The proposed algorithm fuzzifies the Univalued Segment Assimilating Nucleus (USAN) area obtained from the original image, using a USAN area histogram based Gaussian membership function. A parametric fuzzy intensifier operator (FINT) is proposed to enhance the weak edge information, which results in another fuzzy set. The fuzzy measures: fuzzy edge quality factors and sharpness factor are defined on fuzzy sets. BFA is used to optimize the parameters involved in fuzzy membership function and FINT. The fuzzy edge map is obtained using optimized parameters. The adaptive thresholding is used to de-fuzzify the fuzzy edge map to obtain a binary edge map. The experimental results are analyzed qualitatively and quantitatively. The quantitative measures: Pratt's FOM, Cohen's Kappa, Shannon's Entropy and edge strength similarity based edge quality metric, are used. The quantitative results are statistically analyzed using t-test. The proposed algorithm outperforms many of the traditional and state-of-art edge detectors.

Index Terms— Edge detection, Fuzzy intensifier, Fuzzifier, Edge sharpness, Fuzzy entropy, Bacterial Foraging, USAN area.

I. INTRODUCTION

THE edge is characterized by change in intensity with some direction behavior. The edges represent all non-chromatic information contained within an image. In many applications such as segmentation, object detection, edge detection becomes the essential step. Even in case of noise removal, de-blurring, etc. the edge information plays a significant part. The edges can be divided into the two categories, i.e. strong or prominent edges and weak or fine edges. There are a number of methods available for edge detection, which are able to detect prominent edges. Most of the edge detectors available in the literature, are unable to detect fine or weak edges. In this work, we have proposed an optimal fuzzy system to detect strong and weak edges present within the image.

In literature, many edge detectors have been proposed. Canny [1] proposed an optimal edge detector, it smoothens image using Gaussian convolution and directional gradients to

obtain edge information. Similarly Sobel, Prewitt, Roberts and LoG are also gradients based edge detectors [2]. Prewitt operator uses horizontal and vertical gradients, while Sobel uses the same with more weight to central pixel. Roberts uses a method of computing diagonal gradients, i.e. cross gradients to obtain edge information. The LoG edge detector uses Gaussian convolution and Laplacian to approximate the second derivative for edge detection. It has been noticed that, the edge detector involving Gaussian convolution may suffer from the problems like proper localization of edges, vanishing edges and false edges [3]. Salinas *et al.* [4] proposed regularization to obtain fused edge map and overcome “ill-posed” problem of previous edge detectors such as Canny. Perona *et al.* [5] proposed an anisotropic diffusion based algorithm, which produces the sharp region boundary to get better quality edges as compared to other Gaussian based methods. Molina *et al.* [6] presented an edge detector using finite sample images from Gaussian scale-space. The Sobel method is used to produce the edge map of each image. The final edge map is obtained by using a coarse-to-fine tracking algorithm.

The advancement in fuzzy theory, has contributed towards the growth of many fuzzy based edge detectors. The vagueness present in edge detection, i.e. location and visible clarity of the edges can be dealt with fuzzy reasoning. Bezdek *et al.* [7] used fuzzy logic to model geometric feature based functions for edge detection. The fuzzy reasoning is applied for edge in noisy images [8]. The morphological edge extraction method is integrated with local fuzzy properties and minimization of fuzzy entropy [9-10], are used for edge detection. Jinbo *et al.* [11] proposed a multilevel fuzzy edge detection algorithm. Yishu *et al.* [12] abstracted the image features using multi-scale wavelet and fuzzy C-mean clustering algorithm for edge detection. Bustince *et al.* [13] proposed an edge detector based on interval-valued fuzzy sets (interval type-2 fuzzy). The algorithm uses the difference of maximum and minimum in the neighborhood, to represent the uncertainty measure. It uses interval valued t-norm, t-conorm and entropy to detect the edges. In [14], Barrenechea *et al.* use the similar concept as in [13], to produce fuzzy edge images. The edges are obtained from fuzzy images using binarization and thinning. Molina *et al.* [15] presented an approach to generate fuzzy edge images by fuzzifying the gradient magnitude of image. The algorithm uses various parametric membership functions to capture the uncertainty in the edges. The fuzzy edge images are binarized to obtain edges. Most of the fuzzy based edge detectors uses de-fuzzification of the

Manuscript received on April 15, 2015; first revision on September 16, 2015; second revision on November 15, 2015, accepted on March 10, 2016.

Om Prakash Verma is with the Department of Computer Science and Engineering, Delhi Technological University, Delhi, India (e-mail: opverma.dce@gmail.com).

Anil Singh Parihar is with the Department of Computer Science and Engineering, Delhi Technological University, Delhi, India (e-mail: parihar.anil@gmail.com).

fuzzy edge map to obtain binary edges. Pinto *et al.* [16], proposed a thresholding algorithm based on Atanassov's intuitionistic fuzzy sets and Pagola *et al.* [17] presented a fuzzy thresholding algorithm.

Nature-inspired optimization algorithm has got attention from many researchers in the field of image processing. Passino [18] proposed an algorithm based on foraging behavior of bacteria, named bacterial foraging algorithm (BFA). The BFA has evolved significantly since then and find many applications in image enhancement and edge detection [19-21]. In [19], BFA is used to optimize various parameters for enhancement of the under exposed and over exposed region in the color images. Verma *et al.* [20] proposed a modification to the BFA algorithm for edge detection. The probabilistic derivative approach of ant colony optimization (ACO) [22], is used for determining the direction of movement of a bacterium. The edge map is obtained using the optimal path travelled by the bacteria. This approach results in some discontinuities in edges as well as the double edges. Verma *et al.* [21] proposed an edge detection algorithm in the noisy images using fuzzy derivative and BFA. In this algorithm, the movement of bacteria is governed by fuzzy derivate. This approach fails to detect the weak edges present within the image. Setayesh *et al.* [23] proposed an edge detection technique based on particle swarm optimization (PSO) [24]. The authors proposed two modifications of PSO to optimize the two sets of parameters in smooth region and edge pixels. Etemad and White [25] proposed an ACO inspired edge detection approach. In this approach, the authors have defined type-I pheromone deposited by ant movement and type-II pheromone determined by the Euclidean norm of the image gradient. The optimal path of ants is decided by these two pheromones. Sun *et al.* [26] proposed an edge detection technique based on the theory of universal gravity. In this approach, pixel intensity is considered as gravitational mass. Malina *et al.* [27] proposed another edge detector based on law of universal gravity. The algorithm calculates membership value of the edge fuzzy set using gravitational force between pixels. The effective detection of weak edges is still a challenging task. Most of the methods involve derivatives for edge detection, thus, suffers from the noisy and false edge map. A family of edge detectors uses Gaussian convolution, which results in dislocation of edges. The weak edges get vanished and false edges appear.

In the present work, we have developed an algorithm, which is capable of detecting weak edges along with strong edges. The proposed algorithm uses the USAN area [28] histogram based Gaussian membership function [29] to extract primitive edge information. This algorithm uses the integration of derivatives, which reduces the effect of noise. We propose a parametric fuzzy intensification operator (FINT) to enhance the weak edge information. Fuzzy measures are introduced to assess the quality of weak and strong edges. The sharpness factor and fuzzy entropy are used to define the objective function for optimization. BFA is used to get an optimized fuzzy edge map, which is further de-fuzzified to get binary edge map using adaptive thresholding. Following are the key

contributions of our work:

1. A parametric fuzzy intensifier FINT to enhance weak edge information is proposed.
2. The fuzzy quality measures for strong and weak edges present in the image, are introduced. Further the sharpness factor is defined using fuzzy quality measures.
3. The effective utilization of BFA to obtain optimized fuzzy edge, which resulted in binary edge map after de-fuzzification.

The organization of the paper is as follows: Section II introduces the SUSAN principle to calculate the USAN area. Section II describes the fuzzification of USAN area matrix and FINT operation. In Section IV, the fuzzy measures are defined for sharpening the weak and the strong edges. The optimization using BFA is also described in this section. De-fuzzification process to obtain a binary edge map is described in section V. The experimental results and their analysis are presented in section VI. The conclusion is drawn in section VII.

II. COMPUTATION OF USAN AREA

The USAN (Univalue Segment Assimilating Nucleus) area is calculated using SUSAN (Smallest Univalue Segment Assimilating Nucleus) principle [28], with a mask of 37 pixels. The 37 pixel circular mask approximation in a digital image, is shown Fig. 1. The central pixel is compared with other pixels within the mask [28] using the following equation:

$$C(r, r_0) = \exp \left(- \left(\frac{I(r) - I(r_0)}{t} \right)^\delta \right) \text{ where } r \neq r_0 \quad (1)$$

Where $I(r)$ is the intensity of a pixel at position r , within the mask, $I(r_0)$ is the intensity of the nucleus of mask at position r_0 , and t is brightness threshold. The parameter t is used to yield a proper representation for similar pixels in the neighborhood. The suitable value of parameter t is found experimentally to be 20 for 8-bit images. The value of δ is taken as '6' experimentally [21, 28] as it yields good results for nearly all types of images. The USAN area k at position r_0 is calculated as:

$$k(r_0) = \sum_r C(r, r_0) \quad (2)$$

The value k is approximated to the nearest integer. The USAN area attains a maximum value in smooth region and its value decreases as edges appear within the mask. The variation of USAN area represents the edge information in the image. The USAN area involves integration of differences, therefore the noise effect is reduced. Applying the above process to the each pixel results in a USAN area matrix. The value of USAN area is high for flat regions and low for the edges, but USAN area alone is unable to detect the fine features within the image, especially weak edges. We apply the fuzzy theory to capture the fine features of the image.

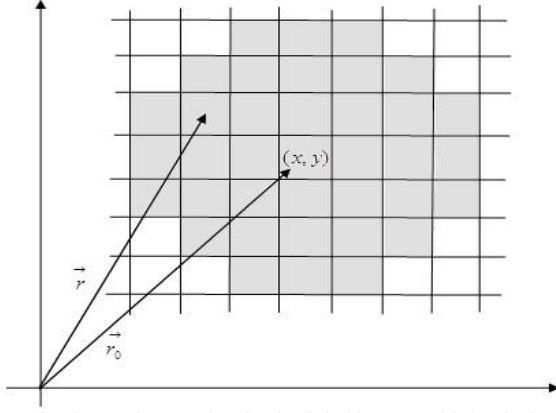


Fig. 1. Circular Mask approximation in digital images, with 37 pixels.

III. FUZZIFICATION

We fuzzify USAN area matrix to obtain a fuzzy property [30], which reflects how much a pixel looks like an edge-pixel. We construct a histogram $h(k)$ of USAN area, which is the total number of pixels having USAN area k . The probability p of having USAN area k is given by:

$$p(k) = \frac{h(k)}{\sum h}; \quad k = 0, 1, \dots, 36. \quad (3)$$

The USAN area histogram satisfies the following condition:

$$\sum_{k=0}^{L-1} p(k) = 1 \quad (4)$$

Where L is the total number of pixels in a mask i.e. 37. The USAN area matrix, obtained from spatial domain is transformed into fuzzy domain using histogram-based Gaussian membership function. This membership function for fuzzification of USAN area matrix is given as:

$$\mu_1(k) = \exp\left(-\left(\frac{(k - k_{\min})^2}{2f_h^2}\right)\right) \quad (5)$$

Where k_{\min} is the minimum USAN area. The fuzzifier parameter f_h is determined as:

$$f_h^2 = \frac{1}{2} \frac{\sum_{k=0}^{L-1} (k - k_{\min})^4 p(k)}{\sum_{k=0}^{L-1} (k - k_{\min})^2 p(k)} \quad (6)$$

The minimum value of k i.e. k_{\min} represents the strong-edge pixels as USAN area is minimum when strong edge is present. Thus, it gives a measure of deviation of the USAN area k from k_{\min} . The use of the ratio (normalized value of moments) rather than the standard variance of k gives robustness against deviations due to non-edge features and noise, but at the same time it may leave weak edges undetected. Since, $\mu_1(k)$ involves only one parameter f_h , it gives little

control to the user. Therefore, to detect the weak edges present in the image, another fuzzy set is obtained using a parametric fuzzy intensification operator (FINT) on fuzzy set generated by $\mu_1(k)$. The membership function for intensifier with

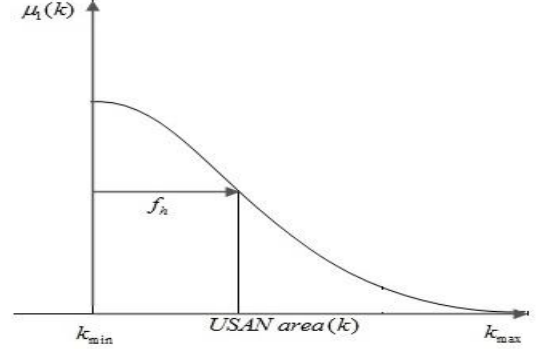


Fig. 2. Histogram-based Gaussian fuzzy membership function.

parameters α , β and γ is given by:

$$\mu_2(k) = \begin{cases} \alpha[\mu_1(k)]^\beta, & 0 \leq \mu_1(k) < 0.5 \\ 1 - \alpha[1 - \mu_1(k)]^\gamma, & 0.5 \leq \mu_1(k) \leq 1 \end{cases} \quad (7)$$

The membership function $\mu_2(k)$ needs to be continuous at $\mu_1(k) = 0.5$, therefore parameters α , β and γ are related in following manner:

$$\alpha = [2^{-\beta} + 2^{-\gamma}]^{-1} \quad (8)$$

Moreover, the parameters α , β and γ needs to be positive to keep the membership value $\mu_2(k)$ between 0 and 1. The strong edges will have a high membership value (in fuzzy set obtained by $\mu_1(k)$ with respect to weak edges as later have higher USAN area. The crossover point, where $\mu_1(k)$ takes the value 0.5 may be considered as a region of division for strong and weak edges. The crossover point will depend on the value of fuzzifier f_h . The application of fuzzy intensification operator (FINT) modifies membership values. The parameters α , β and γ control the shape of the membership function, hence provides more flexibility. The shape of membership function $\mu_2(k)$ becomes straight line for $\alpha = \beta = \gamma = 1$ i.e. no modification. Since, the parameters α , β and γ are related as per Eqn. (8), only two of the three parameters can be varied at a time. The variation in membership function $\mu_2(k)$, while keeping one of the parameters at a fixed value, are shown in Fig. 3. The variations in the value of these parameters alter the membership value of each pixel. The value of fuzzifier f_h , also plays an important role as it decides the crossover point of $\mu_1(k)$, therefore obtaining optimum values of the parameters α , β , γ and f_h becomes imperative.

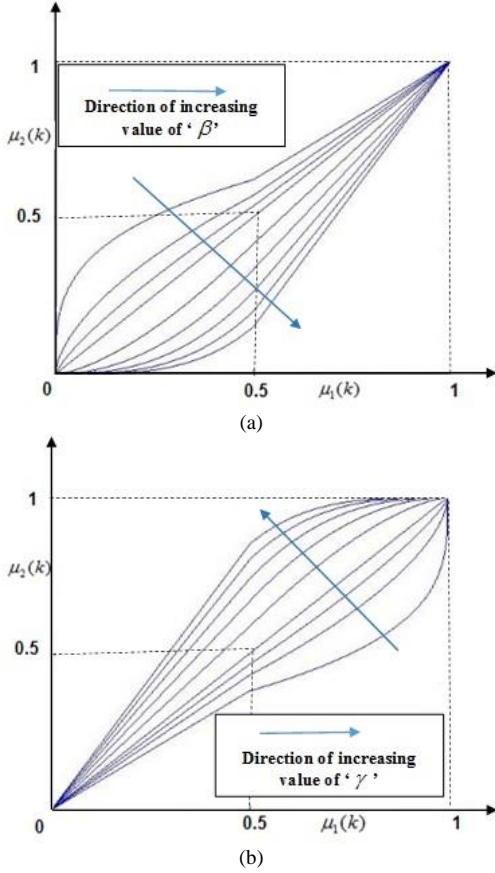


Fig. 3. Fuzzy membership function intensification operator curves (a) with $\gamma=1$, (b) with $\beta=1$.

IV. FUZZY MEASURES AND OPTIMIZATION

To optimize the parameters α, β, γ and f_h , evolutionary optimization technique, namely bacterial foraging algorithm is used in the present work. Other optimization algorithms such as particle swarm, gravitational search, artificial colony, etc. will also work with proposed algorithm. To facilitate the optimization process, the fuzzy measures are defined in the next sections. Further, the objective function for optimization is constructed using these fuzzy measures and fuzzy entropy.

A. Fuzzy Measures

The qualitative assessment of the edge map is required in the optimization process as it tries to achieve best possible edge quality. Assessing the quality of edge map and formulating an ideal measure for it, which can detect all the edges and their exact location, is always a challenge in the edge detection literature. Nearly, all the existing measures leaves some vagueness in one or other situations. The fuzzy theory capture this vagueness; therefore, a fuzzy measure for qualitative assessment becomes instrumental. In the present approach, fuzzy measures are defined, which estimates the edge map quality. The fuzzy edge sharpness measures are defined for both fuzzy sets, i.e. fuzzy sets obtained by $\mu_1(k)$ and $\mu_2(k)$ to measure edge quality of strong and weak

edges. The fuzzy edge sharpness factor for the weak edges of the image is given by:

$$F_w = \sum_{k=0}^{L-1} [\mu_2(k) - c]^2 p(k) \quad (9)$$

Where c is the crossover point of the membership function $\mu_1(k)$. The fuzzy edge sharpness factor indeed depicts, how far the membership values of $\mu_1(k)$ are stretched with respect to the crossover point c after applying the fuzzy intensification operator FINT. Therefore, the desired value of F_w should be obtained by tuning the parameters α, β, γ , and f_h . The average fuzzy edge sharpness factor for the weak edges of the image is given by:

$$F_{avgW} = \sum_{k=0}^{L-1} [\mu_2(k) - c] p(k) \quad (10)$$

Similarly, the fuzzy edge sharpness factor for the strong edges of the image is given by:

$$F_s = \sum_{k=0}^{L-1} [\mu_1(k) - c]^2 p(k) \quad (11)$$

The average fuzzy edge sharpness factor for the strong edges of the image is given by:

$$F_{avgS} = \sum_{k=0}^{L-1} [\mu_1(k) - c] p(k) \quad (12)$$

The fuzzy sharpness factor gives a measure of deviation from the crossover point of fuzzy set and average fuzzy edge sharpness factor gives the overall sharpness of the edges with in the image. The ratio of the absolute average fuzzy edge sharpness factor to the fuzzy edge sharpness factor is defined as the fuzzy edge quality factor. The fuzzy edge quality factor for the weak edges of the image is given by:

$$Q_w = \frac{F_{avgW}}{F_w} \quad (13)$$

The fuzzy edge quality factor for the strong edges of the image is given by:

$$Q_s = \frac{F_{avgS}}{F_s} \quad (14)$$

It can be observed from above definitions neither very low nor very high value of the quality factor represents a good quality edge map. A very low fuzzy edge quality factor results due to very few edges, i.e. poor edge detection. Similarly, a high fuzzy edge quality factor, i.e. membership values have a large deviation from the critical point. Therefore, a high value of the quality factor is due to a large number of false edges in the edge map. Thus the range of values of the fuzzy quality factors for good quality of the edge map exists in between two extreme values. In the present approach, the ratio of these factors is termed as the sharpness factor (section IV.B). The above definitions are coined to capture the extent of the uncertainty present in the edges. These fuzzy measures are used for the computation of the objective function for optimization. The quality of the edges is ascertained by a quantitative analysis based on the fuzzy edge sharpness measures and the fuzzy entropy (section IV.C).

B. Sharpness Factor

The normalized fuzzy edge sharpness factor, called the sharpness factor, is defined to evaluate the fuzzy edge sharpness. The sharpness factor is defined as:

$$S_f = \frac{Q_s}{Q_w} \quad (15)$$

The sharpness factor cannot be used directly as objective function in the optimization process as it will result in parameters (α , β , γ and f_h) pertaining minimum or maximum value of S_f . A very low (high) value of S_f may be due to low (high) value of Q_s and high (low) value of Q_w , but both are undesirable. It is experimentally found that the sharpness factor should lie in between 1.0 and 1.5 for a good quality edge map. The desired value of the desired sharpness factor is denoted as S_{df} .

C. Fuzzy Entropy

In the process of fuzzification, the spatial domain edge information is transformed in the fuzzy domain using USAN area. In the fuzzy domain, the measure of uncertainty or randomness associated with the image is characterized by fuzzy entropy. The fuzzy entropy [31-32] of fuzzy set obtained by $\mu_2(k)$ is defined using Shannon's function as:

$$E = \frac{-1}{L \ln 2} \sum_{k=0}^{L-1} [\mu_2(k) \ln(\mu_2(k)) + (1 - \mu_2(k)) \ln(1 - \mu_2(k))] \quad (16)$$

The fuzzy entropy of an image characterizes the edge information present in the image. The optimization of fuzzy entropy should result in the desired value of the parameters α , β , γ and f_h involve in the membership functions.

D. Objective Function for Optimization

The optimization is performed to obtain suitable values of parameters α , β , γ and f_h such that both strong and weak edges are detected properly. The entropy function and the edge sharpness factor S_f both express the edge quality, hence these must be optimized together. The minimization of fuzzy entropy would reduce the randomness in edge map, but a very low value of fuzzy entropy would result in loss of some actual edges, which is undesirable. Therefore, optimize the entropy function E subject to the constraint $S_{df} = S_f$. An objective function is defined as:

$$J = E + \sigma |S_{df} - S_f| \quad (17)$$

Where $\sigma = 0.5$ is chosen experimentally. The objective function is optimized for four parameters α , β , γ and f_h i.e. four dimensional search space has to be explored for optimum value. The search space needs to satisfy the condition in Eqn. (8) and α , β , γ takes only positive values. The initial value of fuzzier is calculated from image information using Eqn. (6), which controls the shape of the membership function $\mu_1(k)$.

E. Bacterial Foraging algorithm

The bacterial foraging algorithm, introduced by Passino [18] in the year 2002, is one of the nature-inspired evolutionary optimization algorithms, which caught great attention in recent times. The bacterial foraging algorithm is inspired from the foraging strategy of the *E. coli* bacterium. The *E. coli* is made up of plasma cell and contains "cytoplasm" and "nucleoid". At a given temperature and sufficient food bacterium gets longer and divided into daughters (reproduction). In search of food the motion patterns (taxes) of bacteria, govern by chemical attractants (nutrient region) and repellants (noxious region), are called chemotaxis. The *E. coli* bacterium has a control system, which facilitates the search of food and avoids the noxious region by changing the direction of movement. The foraging behavior of the bacteria is mimicked into a mathematical model in the BFA. The major steps of BFA are: Chemotaxis, Swarming, Reproduction, Elimination and Dispersion. An objective function $J(i, j, k, l)$ is defined to represent the fitness (cost) of the i^{th} bacterium at the position $\theta^i(j, k, l)$ in j^{th} chemotaxis, k^{th} reproduction and l^{th} elimination-dispersion step. The positions of bacteria may be initialized randomly or by using some given information.

Chemotaxis: The bacteria search for food, i.e. nutrient rich region and movement of bacteria is governed by nutrient concentration in the search space. The nutrient concentration in search region is represented as a cost function $J(\theta)$ at any position θ . The bacterium moves to attain a better nutrient region, i.e. to reduce cost value $J(\theta)$. Let the step size of movement in random direction is $\lambda(i)$ for the i^{th} bacterium. If direction of movement takes the bacterium in a better nutrient region, it continues to move in the same direction as long as cost function reduces or maximum number of steps N_s reaches. If movement of the bacterium takes in the region of constant nutrients or noxious region, the next step is taken in some random direction, termed as tumble. The random vector $\Delta(i)$ represents the direction of movement, and the movement of a bacterium is given by:

$$\theta^i(j+1, k, l) = \theta^i(j, k, l) + \lambda(i) \frac{\Delta(i)}{\sqrt{\Delta^T(i) \Delta(i)}} \quad (18)$$

Swarming: The bacteria use a cell to cell signaling to swarm together. It signals to attract other bacteria in a nutrient rich region. The bacterium also repels the nearby bacterium, to avoid over consumption of its food by others. These factors are incorporated in BFA by using attractant and repellent parameters.

Reproduction: The health of a bacterium is measured as the value of the objective function at a position. The bacteria are sorted with respect to health, and the healthier bacteria are split into two, in the same position. To make the total number of bacteria same, half of the bacteria are reproduced, and the other half gets eliminated.

Elimination and dispersal: This step redistributes the bacteria population in such a way that bacteria in low nutrient or noxious regions are eliminated and dispersed randomly to explore the regions which might be remain unexplored. The number of eliminations and dispersals kept same in number to

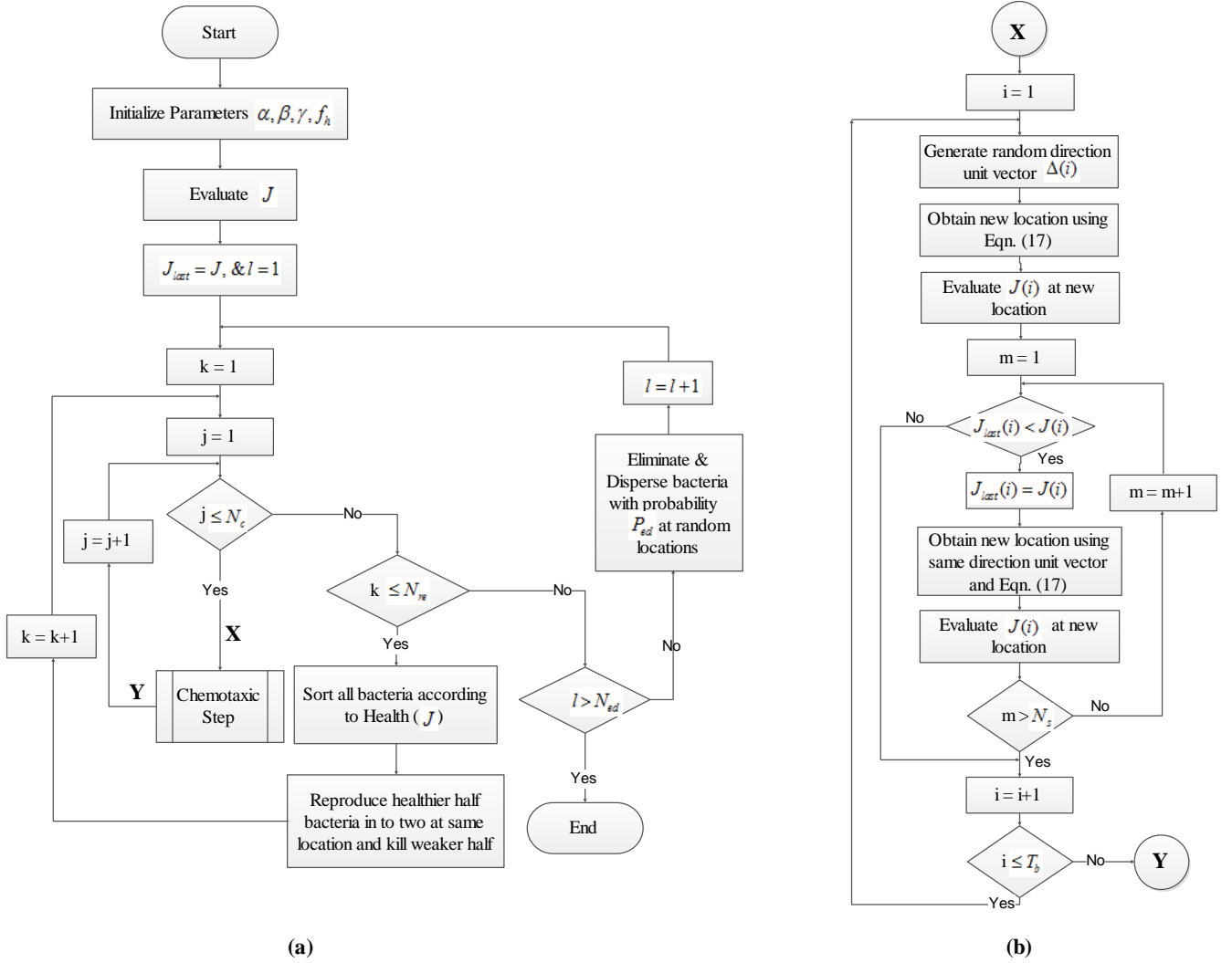


Fig. 4. (a) Flow chart of bacterial foraging algorithm, (b) Flow chart of Chemotaxis step.

make total bacteria count same.

F. Initialization of position of bacteria and the parameters of the BFA

The bacteria are placed randomly in four dimensional search space created by α, β, γ and f_h i.e. position of i^{th} bacterium θ^i is defined by a quartet $(\alpha, \beta, \gamma, f_h)$. The initial positions of all bacteria are taken randomly by taking random values of α, β, γ . Although, the values of parameters α, β, γ are always positive and should follow the relation given by Eqn. (8). The value of f_h is calculated by Eqn. (6), i.e. initially all bacteria are placed in positions having the same value of f_h . The initial values of BFA parameters are set as follows: Number of bacteria $T_b=15$, Swimming length $N_s=5$, Number of iterations in a chemotactic loop $N_c=10$, Number of reproduction steps $N_{re}=6$, Number of elimination and dispersal events $N_{ed}=4$,

probability of elimination/dispersal $p_{ed}=0.26$. Although, the initial position of bacterium has no effect in optimize value of parameters, but a wise selection of positions of bacteria can reduce execution time significantly. The number of bacteria involves another tradeoff between edge quality and complexity. A wide range of values is tried on different images and it is concluded that it needs to keep the optimum number of iterations to execution time reasonable, but at the same time not sacrificing edge map quality. Although, initial values of parameters α, β, γ and f_h does not affect the quality of edge map obtained as these are optimized through BFA.

V. DE-FUZZIFICATION: ADAPTIVE THRESHOLDING

The fuzzification, using optimized parameter results in the edge map with fuzzy membership, termed as the fuzzy edge map. To obtain a well define an edge map, i.e. a classical set having pixels that truly belongs to edges only. Therefore, we need to de-fuzzify the fuzzy set to a classical set.

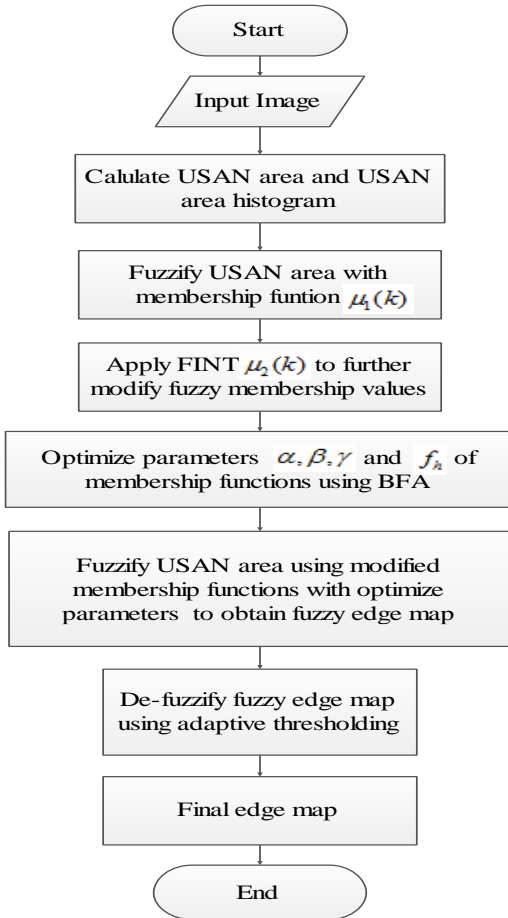


Fig. 5. Block diagram of the proposed approach.

In other words a proper fuzzy set is converted to two-value membership fuzzy set i.e. binarization. This classical set is obtained using adaptive thresholding of the fuzzy edge map. The adaptive thresholding, unlike the global threshold algorithms, allows a user to exact out a higher intensity (membership value) object from the region with varying intensity background in an image. The de-fuzzification/binarization is performed through an adaptive thresholding by analyzing the membership value of each pixel with respect to neighborhood membership values. Therefore, each image pixel is judged adaptively by the localized criteria. The pixels having membership value greater than the local threshold are marked as “object” and the rest of the pixels are considered to be “background”. To achieve a binary edge map, object pixels and background pixels are assigned a value of ‘1’ and ‘0’, respectively. This creates a classical set, i.e. a fuzzy set having membership value ‘0’ and ‘1’ only. In the context of adaptive thresholding, the value of the threshold depends on the particular regions of the fuzzy edge map, hence resulting in adaptive de-fuzzification. A small region within an image is assumed to have approximately uniform intensity in Chow and Kaneko [33] and threshold depends on the local intensity. In an alternative approach, local statistical parameters in the neighborhood of a pixel, are examined to

find the local threshold. The statistical measure of the neighborhood to be considered may depend on the type of image. In simplest measures, it could be *mean* of the local intensity distribution; or the median value; or the median / the mean of the minimum and maximum values, etc. In the present approach the threshold is taken as the average value of mean and median of the local intensity distribution of the fuzzy edge map. The neighborhood size of a pixel need to be sufficient to enclose enough of the foreground and background pixels, but at the same time not too large to lose the advantage of the local statistic. When the variance of the local region is too low the local threshold is replaced with the global threshold to avoid threshold become too low in case of the smooth region. Fig. 5 represents the block diagram of the proposed approach.

VI. EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the algorithm is compared against some well-established edge detection methods, namely Canny, Sobel, Prewitt, Roberts, LoG, and some of the latest edge detection methods described in BFED [20], FEDBF [21], PSODE [23], ACED [25], EDUG [26], GEDT [27], IVFE [13], Canny-gs [15], WFIE [14], M-Sobel [6]. The BFED and FEDBF are based on bacterial foraging algorithm, PSODE is based on particle swarm optimization, ACED is based on ant colony optimization, EDUG and GEDT are based on theory of gravitation. We have also compared our work with some state of art edge detectors based on the fuzzy theory [13-15] and multiscale edge detector [6]. The human judgment is supposed to be the best measure in case of edge detection, but at the same time it can be subjective. Moreover, human visual system (HVS) has its own limitations; therefore, the quantitative measures are also engaged to judge the performance of the proposed approach. The color image edge detection may be performed by applying the proposed algorithm to each component, i.e. red, green, and blue. But, our algorithm uses the HSI color model to represent the color image. In this color model, Hue (H), and Saturation (S) represents chromatic information, whereas intensity (I) represents achromatic information.

TABLE I
INITIAL AND OPTIMAL VALUE OF PARAMETERS FOR LENA IMAGE

Parameter	Initial Value (Random)	Optimized Value for “Lena”	Optimized Value for “Cameraman”
α	5	3.6348	2.9674
β	10	17.4827	15.6352
γ	1.2	0.8732	1.3172
f_h	15	24.0658	21.5961

The edge information is non-chromatic in nature; therefore, only the intensity component is required for edge detection in nearly all practical applications. The use of HSI model over RGB model makes the proposed algorithm universally applicable, i.e. for both gray and color images. The image data set for performance analysis consists of various benchmarks,

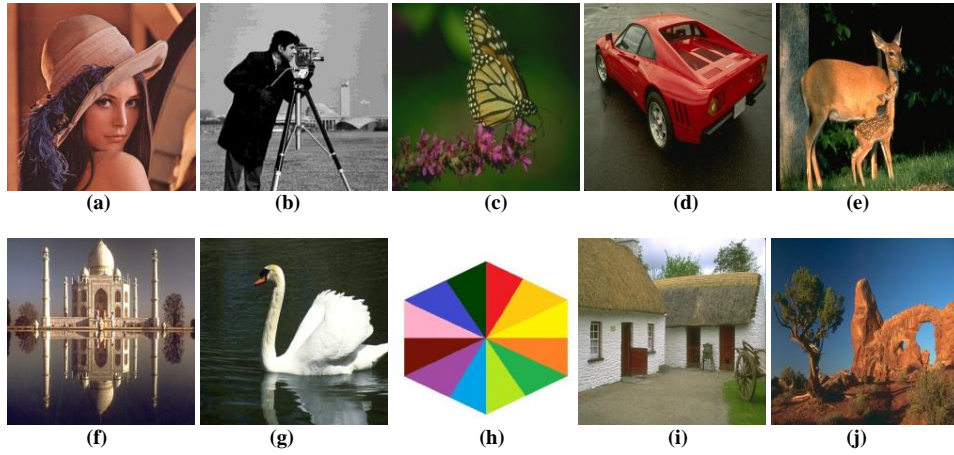


Fig. 6. Test images: (a) Lena (b) Cameraman (c) Butterfly (d) Car (e) Deer (f) Taj (g) Swan (h) Hexagon (i) Hut, (j) Tree

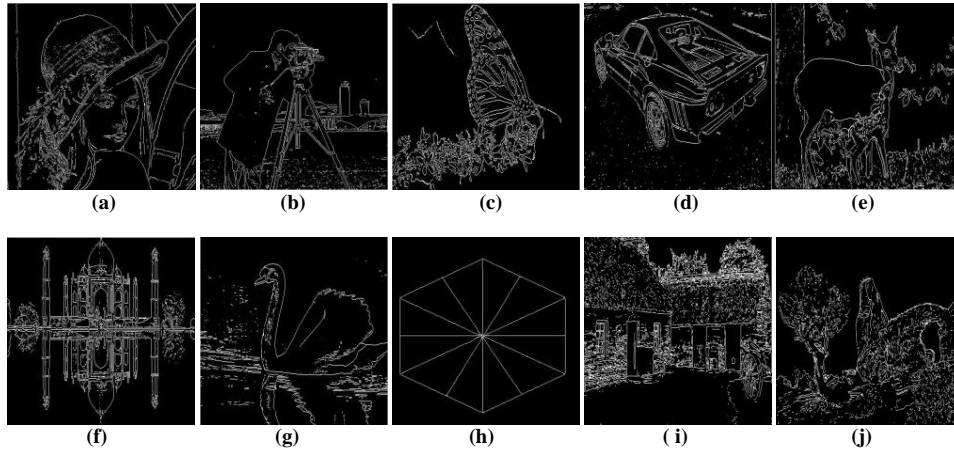


Fig. 7. (a) Lena (b) Cameraman (c) Butterfly (d) Car (e) Deer (f) Taj (g) Swan (h) Hexagon (i) Hut, (j) Tree

natural and synthetic images. We use Berkeley segmentation data set of more than 100 images as a main source of images. Ten such images and corresponding edge maps obtained by proposed algorithm, are shown in Fig. 6 and Fig. 7 respectively. The values of parameters used for Lena and Cameraman images are shown in Table I. The initial values of the parameters is same for all images. This is because, the initial values do not have much effect on the algorithm as BFA optimizes parameters according to image information. Although, if parameters are initialized with proper values, optimization process will take lesser time. The visual analysis of edge maps in Fig. 7 shows, that the proposed filter is able to detect nearly all edges present in images. It may be noted that the proposed approach detects strong and weak edges efficiently. It is evident from the results shown in Fig. 7 that in the region with fine features, only prominent edges are detected. In the highly detailed region, if all the edges are detected, then edge information becomes noisy in nature. Therefore, it is desirable to adjust the parameters in such a way that only prominent edges get detected. In the proposed approach, the parameters are tuned as per nature of the image (i.e. edge information present in the image), also the adaptive

thresholding is used to suppress the unwanted edges. The edge maps by well-established and state-of-art edge detectors are shown in Fig. 8 and Fig. 9 for 'Lena' and 'Cameraman' images. The results of the Canny edge detector, in Fig. 8(a) and 9(a), shows that it gives good edge connectivity, but there are a large number of false edges or unwanted edges. This makes edges map noisy and if the threshold is reduced many prominent edges goes missing. The Sobel edge detector, in Fig. 8(b) and 9(b), on the other hand, gives the lesser number of false edges, but the edge connectivity is poor. Similarity the results for other traditional edges detectors, i.e. Prewitt, Roberts and LoG in Fig 8(c-e) and 9(c-e), shows that either some prominent edges are missing or too many false edges are present in the edge map. Nearly, all the tradition methods shows poor edge connectivity, if the threshold is high and noisy edge map, if the threshold is low. Finding an optimum threshold for each type of image is a critical issue in all these algorithms. The SUSAN edge detector gives thick edges as shown in Fig. 8(f) and 9(f), as the detector works on USAN area alone. The SUSAN edge detector is unable to detect weak edges present in the images. The detector BFED [20] is based on combination of BFA and ACO. Since the algorithm BFED

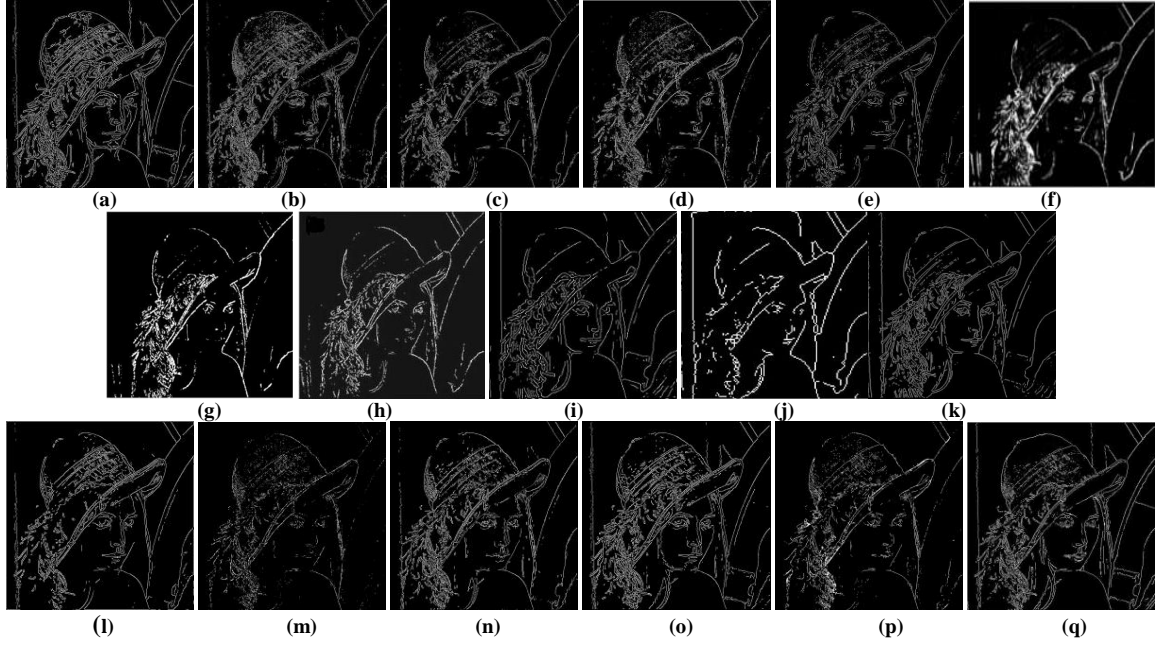


Fig. 8. (a) Canny (b) Sobel (c) Prewitt (d) Roberts (e) LoG (f) SUSAN (g) BFED (h) FEDBF (i) PSODE (j) ACED (k) EDUG (l) GEDT (m) IVFE (n) M-Sobel (o) Canny-gs (p) WFIE (q) Proposed Algorithm

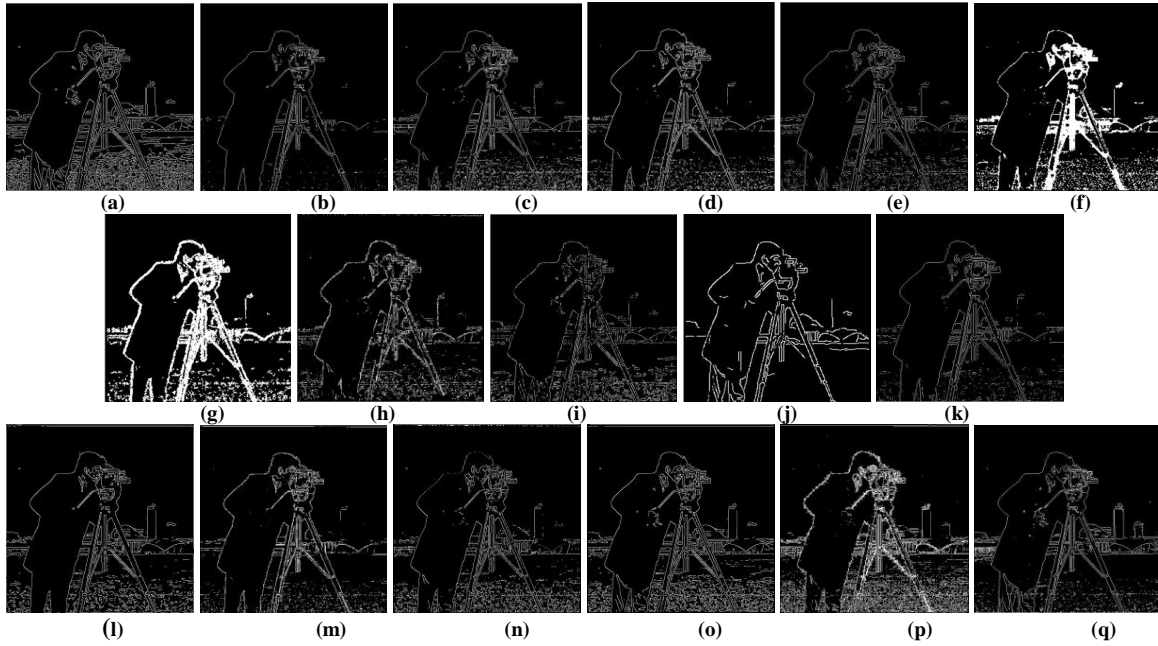


Fig. 9. (a) Canny (b) Sobel (c) Prewitt (d) Roberts (e) LoG (f) SUSAN (g) BFED (h) FEDBF (i) PSODE (j) ACED (k) EDUG (l) GEDT (m) IVFE (n) M-Sobel (o) Canny-gs (p) WFIE (q) Proposed Algorithm

utilizes the ability of BFA to find the optimum path, which is considered as edge map, therefore, mostly strong edges are detected. Also results in Fig. 8(g) and 9(g) shows some disconnected edges. The algorithm in FEDBF [21], also uses BFA to find the edge map, and the fuzzy reasoning is used to find the direction of an optimal path (edge). Moreover, swarming step of BFA is not considered to avoid the convergence of bacteria to only thick edges. The results of

FEDBF algorithm are shown in Fig. 8(h) and 9(h). It is observed that FEDBF gives improved results as compared to other BFA based algorithm BFED, but many of the weak edges are missing. The algorithm in PSODE is based on particle swarm optimization. The results of PSODE in Fig. 8(i) and 9(i), shows that it detects more weak edges, but it has introduced some distortion in the shape of many edges. In ACED [25], the ant colony algorithm based algorithm is

proposed and the results for the same are shown in Fig. 8(j) and 9(j). It is evident from these results that ACED show poor connectivity in the detected edges. The EDUG [26] is an edge detector based on law of universal gravity, and intensity is considered as gravitational mass. The results of EDUG in Fig 8(k) and 9(k), shows that edge connectivity is better, but weak edges are not properly detected. Also the edge detector GEDT [27] considers the intensity of pixel as gravitational mass. A fuzzy set of edges is constructed using gravitation force as membership degree. Canny based edge detector is applied to fuzzy edge image to obtain edges. The authors in [27] has considered various t-norms for implementation, and minimum is best suited for most of the cases. The results for GEDT are shown in Fig. 8(l) and 9(l). These edge maps are comparable to that of canny edge detector. The edge connectivity is good for strong edges, but weak edges are not detected efficiently. The problem in weak edge detection is due to weak gravitational for low intensity pixels. The edges obtained by IVFE [13] are shown in Fig. 8(m) and 9(m). The results shows some missing edges and few noisy edges. The edge detector IVFE divides edge pixels in three category using interval valued fuzzy sets. If pixels with small membership value are included in edge map, it becomes noisy, and if they are excluded, it shows poor edge connectivity. The results for multiscale edge detector M-Sobel [6] are shown in Fig. 8(n) and 9(n). Although, results are improved as compared to Sobel edge detector, yet weak edges are not properly detected. The edge detector Canny-gs [15] fuzzifies gradient magnitudes using different parametric functions. We use sigmoidal membership implementation for comparison purpose. The edge maps for Canny-gs are shown in Fig. 8(o) and 9(o). The edge maps have better connectivity, but some noisy edges are also present. The authors in WFIE [14] presented a method to generate the fuzzy edge map, but their focus is not to develop an edge detector. WFIE converts fuzzy edge map into crisp edges using binarization and thinning. The results of WFIE are shown in Fig. 8(p) and 9(p). WFIE detects strong edges properly, but weak edges shows some missing edges. Although, a proper de-fuzzification may result in good edge map. The visual analysis of all above edge detectors shows that most of the edge detectors fail to detect weak edges. Majority of the edge detectors detects false edges if the threshold is small and leaves missing edges if the threshold is high. The visual perception of Fig. 8 and 9, shows the proposed algorithm performs better or at least comparable to many traditional and state-of-art edge detectors, in weak edge detection. The proposed edge detector is able to detect strong as well as weak edges accurately.

A. Quantitative Analysis

The quantitative analysis is performed using some well-known methods, i.e. Pratt's figure of merit (FOM) [34], Kappa value [35], Shannon's entropy [36], and a recently proposed method edge strength similarity [37]. Since, there is no single universally accepted method for quantitative analysis of edge maps, we have performed the qualitative analysis using all above methods.

1) Pratt's Figure of Merit (FOM)

Pratt's FOM [34] is one of the highly used measure to evaluate the performance of edge detectors in the literature. This measure uses an ideal edge map of an image to compare with the edges obtained by edge detector. The Pratt's FOM is defined as:

$$FOM = \frac{1}{\max\{N_I, N_D\}} \sum_{i=1}^{N_D} \frac{1}{1 + \tau(d_i)^2} \quad (19)$$

Where N_I and N_D are the number of ideal and actually detected edge points respectively, d_i is the distance between i^{th} edge point detected and ideal edge point, and τ is the scaling factor with a value of $1/9$ as in [34]. The expression for FOM in Eqn. (19), suggests that the ideal edge map for given image is required for exact measure. The problem of finding ideal edge map is dealt by taking synthetic images, where exact locations of the edges are known or by taking the best of the well-established methods Canny, Sobel, Prewitt, Roberts

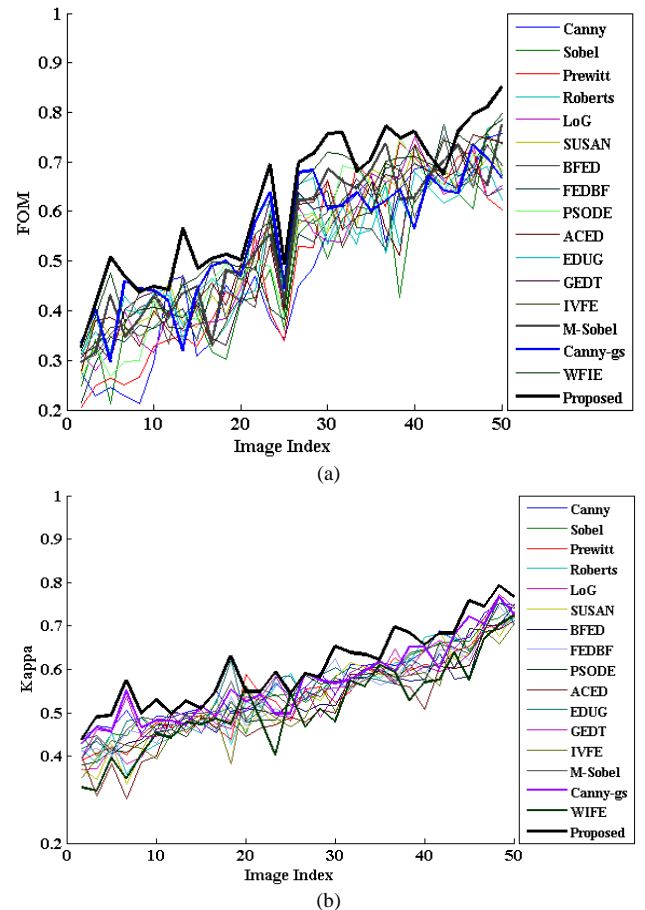


Fig. 10. The quantitative metric curves of the different algorithms: (a) Pratt's FOM, (b) Cohen's Kappa.

and LoG and combine them to give good approximation to ideal edge map.

The FOM values are calculated for the edge maps of images from the Berkeley image database, using various edge detectors. The images are sorted (in the ascending order), and indexed based on the mean value of the FOM metric by different algorithms. In Fig. 10(a), we plot the FOM curves

for different algorithms. Different colors are used to distinguish various algorithm. The FOM curve for the proposed algorithm is highlighted with black color. The proposed algorithm gives highest FOM values for most of the images. However, for few images, FOM values are not much better than other traditional edge detectors. This is because Pratt's FOM method is dependent on the approximation of the ideal edge map, which is difficult to get for some images.

2) Cohen's Kappa

Cohen's Kappa [35] is used for further investigation of experimental results obtained by the proposed algorithm. Cohen's Kappa is defined as:

$$Kappa = \frac{P(a) - P(e)}{1 - P(e)} \quad (20)$$

Where $P(a)$ is observed probability of agreement and $P(e)$ is the probability of random agreement of two approaches. It can be noted from the definition of Cohen's Kappa, that the agreement due to 'chance' i.e. probability of random agreement is factored out. It makes comparison more realistic, i.e. less random. Cohen's Kappa is used for comparison of two edge maps by obtaining pixel to pixel agreement. One of the edge maps is considered as ground truth image, which is a majority image [20]. A majority image/edge map is obtained by using the edge maps of Canny, Sobel, Prewitt, Roberts and LoG. In the majority edge map, a pixel is considered as edge pixel if the majority of the edge maps indicates it to be an edge pixel. In the proposed approach, observed probability of agreement $P(a)$, is calculated as the ratio of number of pixels

declared as edge pixels (by both majority edge map and other edge map in question), to the total number of pixels. The probability of random agreement $P(e)$ is obtained by adding the joint probabilities of labelling a pixel as edge pixel and non-edge pixel by both. The Cohen's Kappa is calculated for the edge maps obtained by various algorithms with respect to majority edge map, and the values are plotted in Fig. 10(b). The Kappa metric curve for the proposed algorithm is highlighted with black color. It may be observed that the proposed method has highest Kappa values for nearly all types of images. However, the Kappa value is slightly less for few images. It is because some of the pixels are declared as non-edge pixels in majority edge map, but they are termed as a weak edge pixel in the proposed algorithm. Therefore, the proposed algorithm is able to detect weak edges better than other traditional filters, which constituted majority edge map.

3) Shannon's Entropy

The quantitative measure Pratt's FOM requires the ideal edge map to compare, which is approximated with the edge maps of the well-established edge detectors. The approximated ideal edge map or majority edge map does not give much information about weak edges. Therefore, Pratt's FOM and Cohen's Kappa are a not true measure for weak edges. Thus, results are further investigated using Shannon's entropy [36], which is defined as:

$$H(I) = - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (21)$$

Where I is the image or edge map, whose entropy is to be

calculated, p_i is probability of a pixel having intensity i , i.e. normalized frequency of occurrence of pixel with intensity i , L represents intensity levels. Shannon entropy represents uncertainty, i.e. information contained within the image. Although, a higher value of entropy does not always reflect the higher edge information. This is because the entropy value may be high due to noisy/ false edges, double edges or broad edges. Also, a very low value of entropy reflects loss of the prominent edges in the edge map. To obtain a fair value of entropy of the ideal edge map, we have used the majority edge map [20] as discussed in section VI.A.2. The majority edge map should have nearly all the prominent edges, but it may not be able to fetch the information about weak

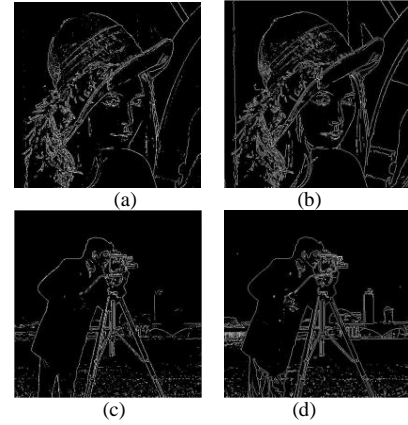


Fig. 11. (a) Majority edges, (b) Proposed algorithm (c) Majority edges (d) Proposed algorithm

edges as shown in Fig. 11. Therefore, the desired entropy should be close to the value of the entropy of the majority edge map.

The Shannon entropy is calculated for edge maps of the images from Berkeley's database, by different edge detectors; and plotted in Fig. 12(a). The entropy for majority edge map is also plotted for the reference. The reference curve and the proposed algorithm curve are highlighted in blue and black colors, respectively. It may be observed that the entropy curve of the proposed algorithm traces closely to that of the majority edge map. However, the proposed algorithm gives little higher values of entropy than the majority edge map. This is because of the weak-edge information contained in edge maps.

4) Edge Strength Similarity based Edge Quality Metric (ESSEM)

Edge strength similarity based edge quality metric (ESSEM) is based a quality metric described in [37]. It compares the quality of the distorted image with the reference/original image. In [37], the edge strength in the diagonal direction of i^{th} pixel is defined as:

$$\chi_i^{j,j+2}(I) = |\partial I_i^j - \partial I_i^{j+2}|^\rho \quad (22)$$

Where ∂I_i^j denotes derivative of image I at i^{th} pixel in the direction j and $\rho = 0.25$. The four directions of the edges are denoted as $j=1,2,3,4$ respectively as shown in Fig. 13. The

total edge strength at i^{th} pixel is defined as:

$$\chi(I, i) = \max \{ \chi_i^{1,3}(I), \chi_i^{2,4}(I) \} \quad (23)$$

The edge strength of distorted image is also calculated in a similar manner. Since, we have to examine the edge quality of the edge map, the edge strength of the original image is compared with the edge map.

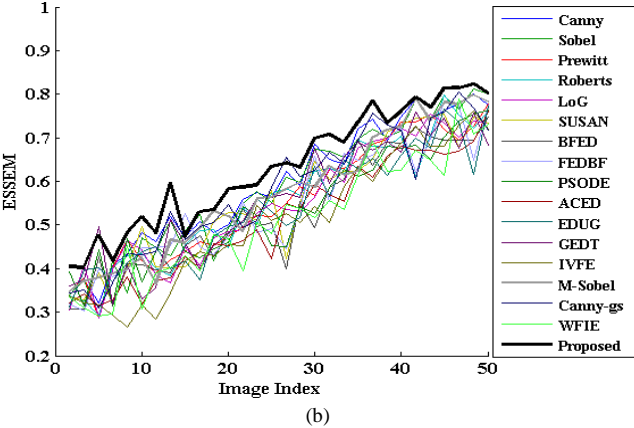
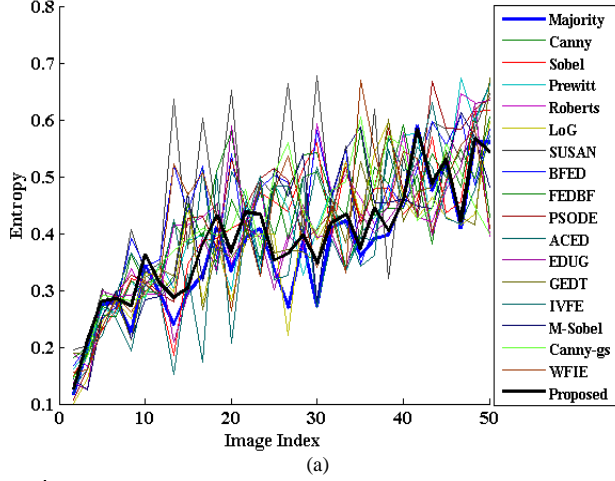


Fig. 12. The quantitative metric curves of the different algorithms: (a) Entropy, (b) ESSEM.

The ESSEM is defined as:

$$ESSEM(I, \xi) = \frac{1}{MN} \sum_{i=1}^{MN} \frac{2\chi(I, i)\chi(\xi, i) + \varepsilon}{(\chi(I, i))^2 + (\chi(\xi, i))^2 + \varepsilon} \quad (24)$$

Where ξ is the edge map, MN is the total number of pixel present in image/edge map, and ε is scaling parameter with value 0.01.

The ESSEM metric for the edge maps obtained by various algorithms is calculated with respect to their original images and values. The ESSEM metric curves for different algorithms

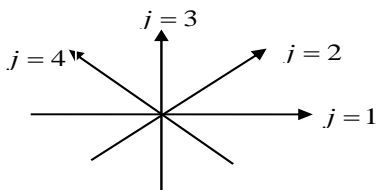


Fig. 13. (a) Directions of edges for directional derivative

are plotted in Fig. 12(b), and the metric curve for the proposed algorithm is highlighted with black color. The ESSEM values do reflect the capability of the edge detector in abstracting the edge information from an image, but the exact value depends upon many factors such as nature, contents, fine details present in the image. It can be observed that the proposed algorithm gives highest values of ESSEM nearly for all the images. Thus, the ESSEM metric supports the fact that our proposed algorithm is able to detect nearly all the prominent and weak edges.

B. Statistical Analysis of Quantitative Results

We perform hypothesis testing using statistical analysis of results obtained in various quantitative measures. The proposed algorithm is compared with other algorithms using the same set of images; thus, a paired t-test analysis [38] is chosen. We propose following null hypothesis (H_0) and alternate hypothesis (H_1):

H_0 : The proposed algorithm does not performs better than the algorithm X.

H_1 : The proposed algorithm performs better than the algorithm X.

Here method X is any algorithm against which the proposed method is being compared. The higher values of the measures FOM, Kappa, and ESSEM, shows better performance. Therefore, one-tail t -test is performed for these measures. However, the entropy value closer to majority edge map reflects better performance. Thus a two-tail t -test is performed between majority edge map and other algorithms for entropy values. The p -values [38] for t -test are shown in Table II. The p -value for the t -test between the majority edge map and the proposed approach is 1.6E-05. The p -values in the Table II shows that the null hypothesis is rejected in each case with at least 0.01 level of significance. The statistical t-test analysis has proved the supremacy of the proposed algorithm over well-established and state-of-art edge detectors.

C. Computational Time Analysis of Proposed Algorithm

We perform computation time analysis by evaluating execution time of the proposed algorithm and other algorithms used for comparison. The execution of algorithms are carried out in a computer with core i3 processor, 2 GB RAM and Windows 7 operating system. The implementation of the algorithms are done in MATLAB R2010b. We have executed each algorithm for 30 times for 50 test images, and an average execution time is recorded in Table III. It is observed that traditional edge detector takes almost negligible time in comparison to nearly all state-of-art edge detectors. This seems to be logical as most of the traditional algorithms involve a convolution with a kernel with few morphological operations. The fuzzy logic based edge detectors: GEDT [27], IVFE [13], Canny-gs [15], M-Sobel [6], and WFIE [14] takes more execution time in comparison to classical edge detectors as most of them are built on either Canny or Sobel edge detectors. All the edge detectors based on evolutionary algorithm take longer execution time.

TABLE II
THE p -VALUES OF THE PAIRED t -TEST BETWEEN AN ALGORITHM X AND THE PROPOSED ALGORITHM

Alg. A→	Canny	[20]	[21]	[23]	[25]	[26]	[27]	[13]	[6]	[15]	[14]
FOM	3.9E-08	6.1E-11	2.4E-09	9.6E-11	6.5E-10	1.5E-08	1.0E-10	5.4E-10	1.5E-07	2.1E-06	1.4E-10
KAPPA	6.3E-06	5.9E-11	4.3E-12	1.5E-10	1.0E-10	2.8E-08	3.4E-08	1.3E-09	1.8E-09	9.6E-08	6.3E-10
ESSEM	5.8E-08	6.8E-12	2.5E-09	1.2E-11	5.3E-15	1.9E-11	1.6E-10	1.7E-13	1.3E-09	6.1E-08	9.2E-14
ENTROP	5.5E-04	2.3E-02	2.5E-03	8.3E-02	3.6E-02	1.3E-01	7.7E-02	6.5E-01	3.2E-04	9.7E-04	3.0E-02

TABLE III
AVERAGE COMPUTATION TIME PER IMAGE (IN SECONDS)

ALGO.	[1]	[20]	[21]	[23]	[25]	[26]	[27]	[13]	[6]	[15]	[14]	PROPOSED
TIME	0.54	25.41	19.37	17.38	26.39	2.31	1.96	2.83	3.74	2.16	1.28	11.46

This is due to the involvement of parameter learning. The gravitational rule based edge detectors (EDUG [26], GEDT [27]) take lesser time in comparison to other evolutionary computation based edge detector. The proposed algorithm takes longer execution time than other fuzzy based edge detectors. This is because we optimize the parameters using BFA. It is evident from the Table III that the proposed algorithm takes the smallest time among all other evolutionary computation based edge detector. As discussed earlier, most of the evolutionary based approaches try to find edges as an optimal path of search-agent (ant, particle, or bacterium). The proposed algorithm optimizes parameter based on fuzzy measures. It requires the lesser number of search agents and hence lesser computation. Although, the proposed algorithm requires longer time than traditional as well as other fuzzy based edge detectors, but there is always a tradeoff between the accuracy and computation time. The proposed algorithm is capable of detecting weak edges with higher accuracy in comparison to other edge detectors.

VII. CONCLUSION

We have presented a fuzzy system for edge detection. The algorithm is designed to detect strong and weak edges efficiently. We have defined a USAN area histogram-based Gaussian membership function, a new parametric fuzzy intensifier FINT, weak and strong fuzzy edge quality, and sharpness factor. The fuzzy entropy and sharpness factor are used to devise the objective function for optimization. The optimized parameters obtained from BFA are used to construct a fuzzy edge map. The fuzzy edge map is defuzzified using adaptive thresholding to result in binary edge map. The results obtained by the proposed algorithm are examined visually and quantitatively. It is compared with well-known traditional methods and state-of-art methods. The proposed algorithm has following key advantages:

1. The proposed algorithm detects edges with more accuracy i.e. better localization of edges, which is supported with visual analysis and quantitative analysis.
2. Improved edges connectivity.
3. The algorithm uses integration of derivatives, which makes our edge detector robust against noisy edges. Therefore, the proposed algorithm detects weak edges with almost negligible noisy edges.

The drawback of the proposed algorithm is that it takes longer execution time as compared to non-evolutionary computation base edge detectors. The computation time can be reduced significantly by using parallel computing. Theoretically in

BFA, bacteria can move simultaneously in search space, but in practical implementation we need to code in sequential manner. A parallel computing hardware/software implementation would make the proposed algorithm comparatively faster. The future scope of the proposed work is to develop parallel computing based implementation to reduce the execution time.

REFERENCES

- [1] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1986, vol. 8, no.1, pp. 679-698, 1986.
- [2] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, 3rd ed., Pearson India.
- [3] M. Basu, "Gaussian Based Edge Detection Methods: A Survey," *IEEE Trans. on systems, man and cybernetics*, vol. 32, no. 3, pp. 252 – 260, 2002.
- [4] R.A. Salinas, C. Richardson, M.A. Abidi, and R.C. Gonzalez, "Data Fusion: Color Edge Detection and Surface Reconstruction through Regularization," *IEEE Trans. on Industrial Electronics*, vol. 43, no. 3, pp. 355-363, 1996.
- [5] P. Perona, J. Malik "Scale-Space and Edge Detection Using Anisotropic Diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.
- [6] C. L. Molina, B. De Baets, H. Bustince, and E. Barrenechea, "Multiscale edge detection based on Gaussian smoothing and edge tracking," *Knowledge-Based Systems*, vol. 44, pp. 101-111, 2013.
- [7] J.C. Bezdek, R. Chandrasekhar, and Attikiouzel, "A geometric approach to edge detection," *IEEE Trans. on Fuzzy Systems*, vol. 6, no. 1, pp. 52-75, 1998.
- [8] F. Russo, "Edge Detection in Noisy Images Using Fuzzy Reasoning," *IEEE Trans. on Instrumentation and Measurement*, vol. 47, no. 5, 1998.
- [9] I. Bloch, "Fuzzy sets in image processing," in *Proc. of ACM Symposium on Applied Computing*, New York, USA, pp. 175 – 179, March 6-8, 1994.
- [10] S. E. El-Khamy, I. Ghaleb, and N. A. El-Yamany, "Fuzzy edge detection with minimum fuzzy entropy criterion," in *Proc. of IEEE Electrotechnical Conference MELECON*, Egypt, pp. 498-503, 7-9 May, 2002.
- [11] Jinbo Wu, Zhouping Yin, "The Fast Multilevel Fuzzy Edge Detection of Blurry Images," *IEEE signal processing letters*, vol. 14, no. 5, pp. 344-347, 2007.
- [12] Yishu Zhai: Xiaoming Liu, "Multiscale Edge Detection Based on Fuzzy C-Means Clustering," in *Proc. of 1st International Symposium on Systems and Control in Aerospace and Astronautics*, Harbin, pp. 1201-1204, June 19-21, 2006.
- [13] H. Bustince, E. Barrenechea, M. Pagola and J. Fernandez, "Interval-valued fuzzy sets constructed from matrices: Application to edge detection," *Fuzzy Sets and Systems*, vol. 160, pp. 1819-1840, 2009.
- [14] E. Barrenechea, H. Bustince, B. De Baets and C. L. Molina, "Construction of Interval-Valued Fuzzy Relations with Application to the Generation of the Fuzzy Edge Images," *IEEE trans. on Fuzzy Systems*, vol. 19, no. 5, pp. 819-830, 2011.
- [15] C. L. Molina, B. De Baets, and H. Bustince, "Generating fuzzy edge images from gradient magnitudes," *Computer Vision and Image Understanding*, vol. 115, pp. 1571-1580, 2011.

- [16] P. M. Pinto, P. Couto, H. Bustince, E. Barrenechea, M. Pagola, and J. Fernandez, "Image segmentation using Atanassov's intuitionistic fuzzy sets," *Expert Systems with Applications*, vol. 40, pp. 15-26, 2013.
- [17] M. Pagola, C. L. Molina, J. Fernandez, E. Barrenechea, and H. Bustince, "Interval Type-2 Fuzzy Sets Constructed From Several Membership Functions: Application to the Fuzzy Thresholding Algorithm," *IEEE trans. on Fuzzy Systems*, vol. 21, no. 2, pp. 230-244, 2013.
- [18] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 3, pp. 52-67, 2002.
- [19] M. Hanmandlu, O. P. Verma, N. K. Kumar, M. Kulkarni, "A Novel Optimal Fuzzy System for Color Image Enhancement Using Bacterial Foraging," *IEEE Trans on Instrumentation and Measurement*, vol. 58, no. 8, pp. 2867 - 2879, 2009.
- [20] O. P. Verma, M. Hanmandlu, P. Kumar, S. Chhabra, A. Jindal, "A novel bacterial foraging technique for edge detection," *Pattern Recognition Letters*, vol. 32, pp. 1187-1196, 2011.
- [21] O. P. Verma, M. Hanmandlu, A. K. Sultania and A. S. Parihar, "A Novel Fuzzy System for Edge Detection in noisy image using Bacterial Foraging," *Multidimensional Systems and Signal Processing*, vol. 24, no. 1, pp 181-198, 2013.
- [22] M. Dorigo, V. Maniezzo, and A. Colomi, "Ant System: Optimization by a Colony of Cooperating Agents," *IEEE Trans. on systems, man, and cybernetics-part b cybernetics*, vol. 26, no. 1, 1996.
- [23] M. Setayesh, M. Zhang, M. Johnston, "A novel particle swarm optimization approach to detecting continuous, thin and smooth edges in noisy images," *Information Sciences*, vol. 246, pp. 28-51, 2013.
- [24] J. Kennedy, R. Eberhart, "Particle Swarm Optimization," in *Proc. of IEEE International Conference on Neural Networks* vol. 4, pp. 1942-1948, 1995.
- [25] S. Ali Etamad, Tony White, "An ant-inspired algorithm for detection of edge features," *Applied Soft Computing*, vol. 11, pp. 4883-4893, 2011.
- [26] G. Sun, Qinhua Liu, Qiang Liu, C. Ji, X. Li, "A novel approach for edge detection based on the theory of universal gravity," *Pattern Recognition*, vol. 40, pp. 2766-2775, 2007.
- [27] C. L. Molina, H. Bustince, J. Fernandez, P. Couto, and B. De Baets, "A gravitational approach to edge detection based on triangular norms," *Pattern Recognition*, vol. 43, pp. 3730-3741, 2010.
- [28] S. M. Smith, and J. M. Brady, "SUSAN: A new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no.1, pp.45-78, 1997.
- [29] M. Hanmandlu, O. P. Verma, P. Gangwar, S. Vasikarla, "Fuzzy Edge and Corner Detector for Color Images," in *Proc. of 6th IEEE International Conference on Information Technology: New Generations*, Las Vegas, NV, pp. 1301 - 1306, April 27-29, 2009.
- [30] S. K. Pal and R. A. King, "Image enhancement using smoothing with fuzzy sets", *IEEE Trans. on System, Man, and Cybernetics*, vol. 11, no. 7, pp. 494 - 501, 1981.
- [31] A. De Luca, S. Termini, "A definition of Nonprobabilistic Entropy in the setting of fuzzy set theory," *Information and Control*, vol. 20, pp. 301-312, 1972.
- [32] H. Cheng, Y. Chen, Y. Sun, "A novel fuzzy entropy approach to image enhancement and thresholding," *Signal Processing*, vol. 75, pp. 277-301, 1999.
- [33] C. K. Chow and T. Kaneko, "Automatic boundary detection of the left ventricle from cineangiograms," *Computers and Biomedical Research*, vol. 5, no. 4, pp.388- 410, 1972.
- [34] I. Abdou and W. Pratt, "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors", in *Proceedings of IEEE* vol. 67, no.5, pp. 756-766, 1979.
- [35] Cohen J., "A coefficient of agreement for nominal Scales," *Educational and Psychological Measurement*, vol. 20 no. 1, pp. 37-46, 1960.
- [36] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, 1948.
- [37] X. Zhang, X. Feng, W. Wang and W. Xue, "Edge Strength Similarity for Image Quality Assessment," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 319-322, 2013.
- [38] D. A. Lind, W. G. Marchal, S. A. Wathen, *Statistical Techniques in Business and Economics*, 13th ed., Tata McGraw Hill Education Pvt. Ltd. India.



Om Prakash Verma (M'76) received his B.E. degree in Electronics and communication engineering from Malaviya National Institute of Technology, Jaipur, India, M. Tech. degree in Communication and Radar Engineering from Indian Institute of Technology (IIT), Delhi, India and PhD in the area of applications of soft and evolutionary computing in image processing from University of Delhi, Delhi, India. From 1992 to 1998 he was assistant professor in Department of ECE at Malaviya National Institute of Technology, Jaipur, India. He joined Department of Electronics & Communication Engineering, Delhi Technological University (formerly Delhi College of Engineering) as Associate Professor in 1998. Currently, he is Head & Professor in Department of Computer Science and Engineering, Delhi Technological University, Delhi, India. His present research interest includes: Applied Soft Computing, Nature Inspired Algorithms, Swarm Intelligent, Evolutionary Computing, and Image Processing.



Anil Singh Parihar received his B. Tech Degree in Electronics and Communication Engineering in 2005 from Priyadarshini College of India and the Master of Engineering degree in Electronics and Communication Engineering from Delhi College of Engineering, New Delhi, India, in 2008. He joined the Department of Information Technology at Delhi Technological University, Delhi, India as Assistant Professor in 2010. Currently, he is Assistant Professor in the Department of Computer Science & Engineering at Delhi Technological University, Delhi, India. His research interest includes Image Processing, Biometrics, Soft Computing and Evolutionary Algorithms.

AN OPTIMIZED AND COST EFFICIENT REALIZATION OF REVERSIBLE BRAUN MULTIPLIER

By

NEETA PANDEY *

NALIN DADHICH **

MOHD. ZUBAIR TALHA ***

* Associate Professor, Department of Electronics and Communication Engineering, Delhi Technological University, India.

** _*** UG Scholar, Department of Electronics and Communication Engineering, Delhi Technological University, India.

ABSTRACT

In CMOS logic, there is a steady increase in power dissipation which appears in the form of heat to the surrounding environment and affects the reliability. The research efforts are made towards looking into alternatives that go beyond the traditional CMOS technologies, and reversible logic has emerged as a promising choice. In this paper, an optimized and cost efficient realization of reversible Braun multiplier is presented. The design of a 4x4 bit multiplier is developed, designed and presented in this paper as an illustration. The architecture is iterative and hence this can easily be extended to the generalized multiplier of any order. The proposed design of a 4x4 reversible Braun multiplier uses three types of reversible gates namely, PG, HNG and TG gates. The proposed design is compared with an already presented reversible multiplier design showing that the proposed multiplier design is more efficient in terms of quantum cost, constant inputs, garbage outputs and the number of elementary reversible gates.

Keywords: Reversible Multiplier, Reversible Circuit, Reversible Gate, Braun Multiplier.

INTRODUCTION

The CMOS technology has constantly moved to take up a central position in electronic system design. It finds the use in most of the cutting edge commercial applications. The processes have shrunk and the design automation tools are in common place. This has accelerated the design of complex chips operating at higher frequencies than a decade ago. Although, the basic CMOS design principle remains unaltered, there is a significant change in the emphasis and approach of the design. There is a steady increase in power dissipation in the form of heat to the surrounding environment which affects the reliability. The research efforts are made toward looking into alternatives that go beyond the traditional CMOS technologies and reversible logic wherein operations performed in an invertible manner have emerged as a promising choice. New technologies are introduced to reduce heat generation but Landauer (Landauer, 1961) pointed out that, the gates always dissipate energy regardless of underlying technology. There is loss of information in classical computing, which is a dissipative process. The entropy of system is reduced by $\Delta S = \kappa \ln 2$ for every bit of information that is lost during computation process, and

hence $H = T\Delta S = \kappa T \ln 2$ Joules of heat is dissipated where κ and T are Boltzmann's constant and absolute temperature respectively. In 1973, Bennett (Bennett, 1973) of IBM research showed that, a classical computing can be broken down into a series of steps which can be performed reversibly, it is emphasized that, a reversible computer can run forward till the completion of computation to get the final result, print the result, and then reverse back all of its steps to return to its initial step and thus resulting in no heat generation. The literature survey on reversible logic shows that a progress has been made in the direction of exploring new reversible gates and applying the concepts to the areas such as DNA computing (Sarker, 2011), cellular automata (Morita, 2008), cryptography and encoding/decoding methods (McGregor, John Patrick & Lee, 2001).

Multiplication is an essential operation used in almost all the modern digital systems including digital signal processors, bioinformatics, complex image processing and microcontrollers. The performance of the digital system depends heavily on the multiplier due to its extensive use. It is therefore, highly desirable that, these are fast and consume low power without performance degradation. In an effort to apply reversible logic to multiplier, the objective

An Overview- Tribological Aspects at the Interface of Piston Ring and Cylinder Liner for Improving the Performance of IC Engine

R. C. Singh¹, R. K. Pandey² and S. Maji³

^{1,3}(Department of Mechanical Engineering, Delhi Technological University, Delhi-110042, India)

²(Department of Mechanical Engineering, I.I.T. Delhi, New Delhi- 110 016, India)

¹rcsingh68@hotmail.com

Abstract : The scientists are reluctantly working on tribo-pairs to determine the tribological effects between them under different condition of lubrication. I. C. Engine is one of the major inventions which effects human life of civilization. There are many parts in I.C. engine which have relative motion among them. The most important part is piston ring and cylinder liner. Continuous dynamic loading and motion plays at the interface of them. The intrinsic international norms forced researchers to concentrate on power producing units and alternative source of energy. An overview pertaining to tribological aspects for improving the performance of IC engine has been carried in this paper.

Keywords : IC engine, Chemical and mechanical energy, Lubrication, Piston ring pack Face profiles, complex motion.

I. INTRODUCTION

Internal combustion engine (IC) as shown schematically in Fig. 1 is the most important mechanical invention done by human beings, which has played great role in the industrialization of the globe after world war-II. But, now due to fast depletion of conventional fuel resources and increasing environmental issues, there is worldwide relentless pressure on the researchers to develop ever more fuel efficient and compact IC engines having reduced environmental issues. In the last couple of decades, there have been many studies on the frictional evaluations at the various interfaces in IC engines in order to identify the crucial interfaces of the engine components for minimizing the interfacial frictional losses associated with it.

An Overview- The major portion of fuel energy (i.e. chemical energy) goes as waste in the form of heat. Even, significant portion of chemical energy liberated during the combustion of precious fuel is consumed in the frictional resistance present at the various interfaces of the moving engine components. Fig. 2a illustrates percentage of chemical energy taken away by various modes in a typical IC engine. It can be seen that the portion of chemical energy consumed in friction during mechanical motions in a typical IC engine is considerably large. Moreover, Fig. 2b also provides break ups of frictional losses in a typical IC engine. Due to considerably large frictional losses in piston assembly, this system contributes in more fuel and lubricating oil consumptions and in this way it happens to be a potential source of hydrocarbons and particulate emissions.

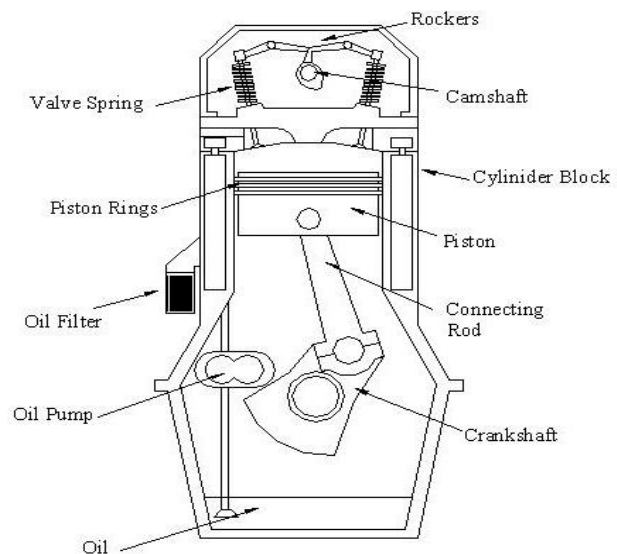
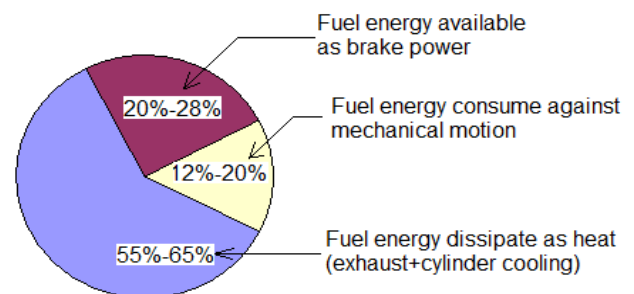
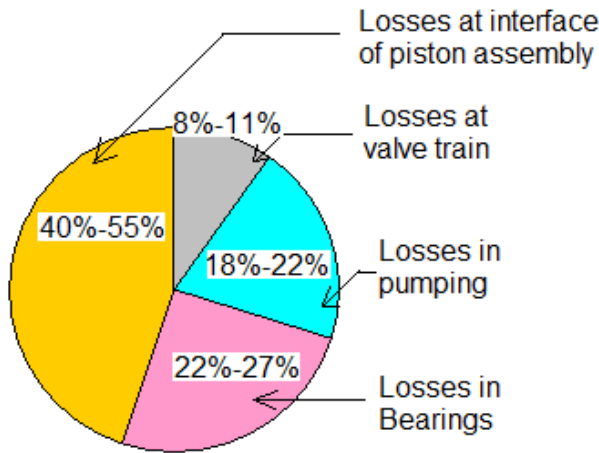


Fig. 1 Vital components of a typical IC engine [1]



(a) Break-ups of fuel's chemical energy



(a) Break-ups of mechanical energy losses

Fig. 2 Chemical and mechanical energy losses in a typical IC engine [1]

Reductions of fuel consumption and emission in an internal combustion engine are largely a function of improved lubrication. Therefore, advanced concepts are now being explored to reduce the friction at various interfaces in an IC engine. The development of smart IC engines and their proper use are of great importance for the national economy, individual and environment. Energy efficient IC engines can save billions of dollars in the case of an industrialized nation. It is worth to mention here that improvements in the tribological performance of interfaces in IC engines can lead to the following benefits:

- Reduced fuel consumption
- Increased engine power output
- Reduced oil consumption
- A reduction in harmful exhaust emissions
- Improved durability, reliability and engine life
- Reduced maintenance requirements and longer service intervals

With large numbers of IC engines in use across the globe, even a fraction of improvements in engine efficiency and emission level can have a major influence on the world fuel economy and the environment in a long term.

Piston and Piston Rings - It will be appropriate to say that heart of the reciprocating internal combustion engine is the piston assembly. Figure 1.3 presents photographic view of piston assembly. Piston and piston rings form a critical unit in transforming the fuel energy into useful kinetic energy. The piston ring pack includes the piston rings, which is a series of compression rings and oil ring. The primary role of the compression piston ring is to maintain an effective gas seal between the combustion chamber and the crankcase. The piston rings of the piston assembly, which form a labyrinth seal, achieve this function by closely conforming to their grooves in the piston and interfacing to the cylinder wall. The additional role of the piston rings is to transfer heat from the piston into the cylinder wall and limit the amount of lubricating oil that is transported from the crankcase to the combustion chamber. This flow path is perhaps the largest contributor to engine oil consumption and leads to increase in

harmful exhaust gas emissions as the lubricating oil mixes and reacts with the other contents present in combustion chamber.



Fig. 3 Photographic view of piston assembly [2]

Figs. 4(a) and 4(b) illustrate face profiles and functions of piston rings in an IC engine. Two top piston rings are compression rings. The pressure generated during the combustion pushes the piston rings radially outward, which causes engagement of the entire piston ring face with the cylinder wall. This process helps in gas sealing. The second compression ring, which is known as a scraper ring is designed to assist in the limiting of the upward oil flow in addition to providing a secondary gas sealing. Figure 4 (b) illustrates surplus oil scraping from the cylinder wall by the second compression piston ring. For scraping function, the second compression ring is provided a tapered-faced profile. The bottom piston ring in the piston assembly is known as oil control ring, which has two running faces (or lands) and a spring element to enhance the radial load. The role of this ring is to limit the amount of oil transported from the crankcase to the combustion chamber.



(a) Face profiles of piston rings

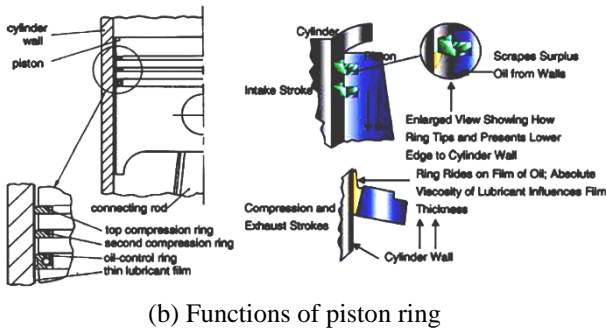


Fig. 4 Face profiles and functions of piston rings [1,3]. The piston ring is the most complicated tribological component in the internal combustion engine to analyze because of large variations of load, speed, temperature and lubricant availability. In one single stroke of the piston, the piston ring interface with the cylinder liner wall may experience boundary, mixed and full fluid film lubrication. During the engine cycle the piston itself exhibits a complex motion i.e. transverse motion, axial motion, and secondary motion; as illustrated in Figs. 5(a), 5(b), and 5(c). Such motions result in hydrodynamic/mixed lubrication at the various interfaces in piston assembly. In presence of poor lubrication at the piston skirt and cylinder wall, noise generation may take place due to piston slap.

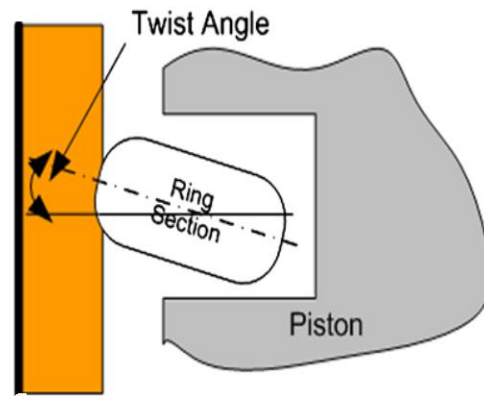
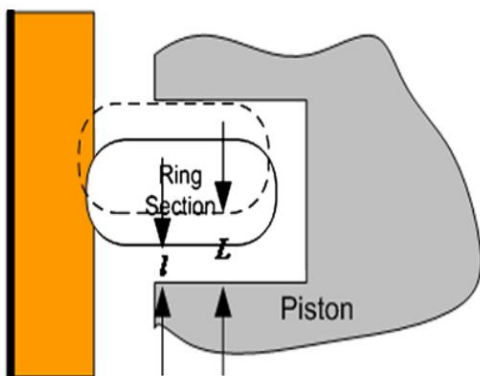
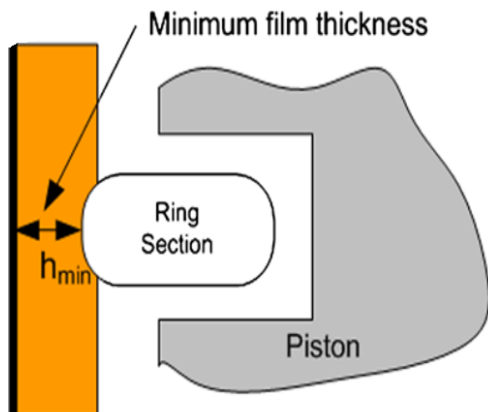


Fig. 5 Movements of piston rings during operation [4]

II. CONCLUSION

It is widely understood that surface profiling and surface topography of mating solids have remarkable influence on tribological behaviors in both dry and lubricated sliding conditions. Though the influences of the micro and nano scale topography are more complex but, at the same time it offers interesting possibilities for friction reduction. The most promising results related to friction reduction may come out with some novel surface profiles in presence and absence of surface texturing. Therefore, in light of stringent federal legislations pertaining to better fuel economy and reduced emissions, there are needs to revisit lubrication and tribological contact design aspects in IC engine for friction reduction. Improving performance of piston assembly from tribological aspects (reducing friction) will have immense influence on fuel saving and controlling the emissions.

REFERENCES

- [1] S. C. Tung, M. McMillan, "Automotive Tribology Overview of Current Advances and Challenges for the Future" Tribology International, Vol. 37, 2004, PP. 517-536.
- [2] <http://www.aalcar.com> dated 01/08/2012
- [3] <http://www.sacskyranch.com> dated 01/08/2012
- [4] K. Wannatong, S. Chanchaona, S. Sanitjai, "Simulation Algorithm for Piston Ring Dynamics," Simulation Modeling Practice and Theory, Vol. 16, 2008, PP. 127-146.
- [5] P. Jost, J. Schofield, "Energy Savings through Tribology: A Techno-economic Study" Proceedings of Institution of Mechanical Engineers, Vol. 195, 1981, PP. 151-173.
- [6] C. M. Taylor, "Automobile Engine Tribology – Design Considerations for Efficiency and Durability" Wear, Vol. 221, 1998, PP. 1-8.
- [7] R C Singh, "Thermohydrodynamic Analysis of Lubricated Piston Rings of Internal Combustion Engine" Ph. D. Thesis, University of Delhi, 2013.



Atomic force microscopy studies of homoepitaxial GaN layers grown on GaN template by laser MBE

B. S. Choudhary, A. Singh, S. Tanwar, P. K. Tyagi, M. Senthil Kumar, and S. S. Kushvaha

Citation: *AIP Conference Proceedings* **1724**, 020120 (2016); doi: 10.1063/1.4945240

View online: <http://dx.doi.org/10.1063/1.4945240>

View Table of Contents: <http://scitation.aip.org/content/aip/proceeding/aipcp/1724?ver=pdfcov>

Published by the *AIP Publishing*

Articles you may be interested in

[Stress distribution of GaN layer grown on micro-pillar patterned GaN templates](#)

Appl. Phys. Lett. **103**, 012102 (2013); 10.1063/1.4813077

[Optical and structural studies of homoepitaxially grown m-plane GaN](#)

Appl. Phys. Lett. **100**, 172108 (2012); 10.1063/1.4706258

[Homoepitaxial N-polar GaN layers and HEMT structures grown by rf-plasma assisted molecular beam epitaxy](#)

J. Vac. Sci. Technol. B **30**, 02B113 (2012); 10.1116/1.3676175

[Atomic force microscopy analysis of cleaved facets in III-nitride laser diodes grown on free-standing GaN substrates](#)

Appl. Phys. Lett. **88**, 041910 (2006); 10.1063/1.2167400

[GaN homoepitaxial layers grown by metalorganic chemical vapor deposition](#)

Appl. Phys. Lett. **75**, 1276 (1999); 10.1063/1.124666

Atomic Force Microscopy Studies of Homoepitaxial GaN Layers Grown on GaN Template by Laser MBE

B.S. Choudhary^{1, 2}, A. Singh³, S. Tanwar², P.K. Tyagi³, M. Senthil Kumar¹ and S.S. Kushvaha^{1, a)}

¹ CSIR- National Physical Laboratory, Dr K.S. Krishnan Road, New Delhi 110012, India.

² Rajasthan Technical University, Rawatbhata Road, Kota 324010, India.

³ Department of Applied Physics, Delhi Technological University, Delhi 110042, India.

^{a)}kushvahas@nplindia.org

Abstract. We have grown homoepitaxial GaN films on metal organic chemical vapor deposition (MOCVD) grown 3.5 μm thick GaN on sapphire (0001) substrate (GaN template) using an ultra-high vacuum (UHV) laser assisted molecular beam epitaxy (LMBE) system. The GaN films were grown by laser ablating a polycrystalline solid GaN target in the presence of active r.f. nitrogen plasma. The influence of laser repetition rates (10-30 Hz) on the surface morphology of homoepitaxial GaN layers have been studied using atomic force microscopy. It was found that GaN layer grown at 10 Hz shows a smooth surface with uniform grain size compared to the rough surface with irregular shape grains obtained at 30 Hz. The variation of surface roughness of the homoepitaxial GaN layer with and without wet chemical etching has been also studied and it was observed that the roughness of the film decreased after wet etching due to the curved structure/rough surface.

Keywords: Gallium Nitride, GaN template, Laser molecular beam epitaxy, Homo-epitaxial growth, Atomic force microscopy, Wet chemical etching.

INTRODUCTION

The III-nitride semiconductors and their alloys constitute a diverse and fascinating class of materials which have wide application in the field of laser diodes, light-emitting diodes, high electron mobility transistors and ultra- violet detectors due to their remarkable properties such as wide direct band gaps, good thermal conductivity and high breakdown field [1]. For the growth of epitaxial GaN, conventionally hydride vapor phase epitaxy (HVPE) [2], molecular beam epitaxy (MBE) [3] and metal organic chemical vapor deposition (MOCVD) [4] growth techniques are used. There is a high probability of formation of interfacial compounds on chemically vulnerable substrates for GaN film grown by HVPE and MOCVD due to their high growth temperature. Laser MBE (LMBE) is originated as an alternative technique to grow GaN at a relatively lower temperature because in it, laser ablation produces high kinetic energy film precursors [5-8]. Here, we report the effect of laser repetition rate on the surface morphology of homoepitaxial GaN films grown on GaN template using LMBE.

It has been well known that wurtzite GaN has two different polarities along the c-axis, e.g. Ga-polar (0001) and N-polar (000-1). Polarities of GaN surface greatly affect the growth [9] and device performances [10, 11] so it is of key importance to determine the polarity of GaN surface. A simple method of determining the polarity uses an aqueous solution of KOH that etches selectively N-polar surface of the GaN smooth layer while there is almost no effect of KOH on Ga-polar surface [12]. Thus, it is of great importance to know that how this KOH etching affects the surface morphology of LMBE grown GaN layers, especially when the GaN surface is rough. Here, we report the etching behavior in KOH solution of GaN layers prepared by LMBE at different laser repetition rates. It is found that KOH etches the surfaces and decrement in surface roughness has been observed. Maximum change in rms roughness has been observed for rougher GaN layer grown at large repetition rate (i.e.30 Hz).

EXPERIMENTAL DETAILS

The growth of epitaxial GaN films on MOCVD grown GaN template was carried out in an UHV-LMBE system. The LMBE system is equipped with reflection high energy electron diffraction (RHEED) and a radio-frequency (RF) nitrogen plasma. The base pressure in the main growth chamber was better than 2×10^{-10} Torr. A KrF excimer laser (248-nm wavelength, 25 ns pulse) was used to ablate a high purity polycrystalline HVPE grown GaN target. We grew GaN films on GaN templates for 2 hours at different laser repetition rates (10-30 Hz) at a growth temperature of 700 °C, keeping r.f. nitrogen plasma constant. The surface morphologies of these samples were characterized by atomic force microscopy (AFM) in tapping mode using silicon tips of curvature radii less than 10 nm. In order to determine the etching behavior, these homoepitaxially grown samples were placed in KOH solution (KOH:H₂O = 1:5 weight ration) for 30 min at room temperature. The root-mean-square (rms) surface roughness of as grown LMBE GaN layers and etched GaN layers was determined from AFM images of scan area $2 \mu\text{m} \times 2 \mu\text{m}$.

RESULTS AND DISCUSSION

Before the growth of GaN on GaN template, the surface morphology of GaN template was studied by AFM. The AFM image of GaN template revealed a quite smooth surface, consisting of array of straight narrow (width 50-120 nm) terraces with step height of $\sim 3 \text{ \AA}$ which agrees well with one GaN (0001) monolayer (2.6 \AA) (not shown here) [3, 13,14]. It is well known that in MBE growth of GaN, the Ga/N flux ratio is a crucial parameter which determines the mode of GaN growth, i.e., Ga-rich or Ga/N ~ 1 condition normally promotes a two dimensional (2D) smooth GaN growth [3] while N-rich condition results in 3D growth [8]. In our case, we employed laser ablation of a HVPE grown GaN bulk target under the RF activated nitrogen plasma ambient to maintain the stoichiometry as GaN plume may also contain GaN_{1-x} species [5]. Previously in-situ RHEED monitoring of the GaN layers grown on GaN templates showed a spotty RHEED pattern of GaN layer which indicates the 3D epitaxial growth due to the N-rich growth condition [15].

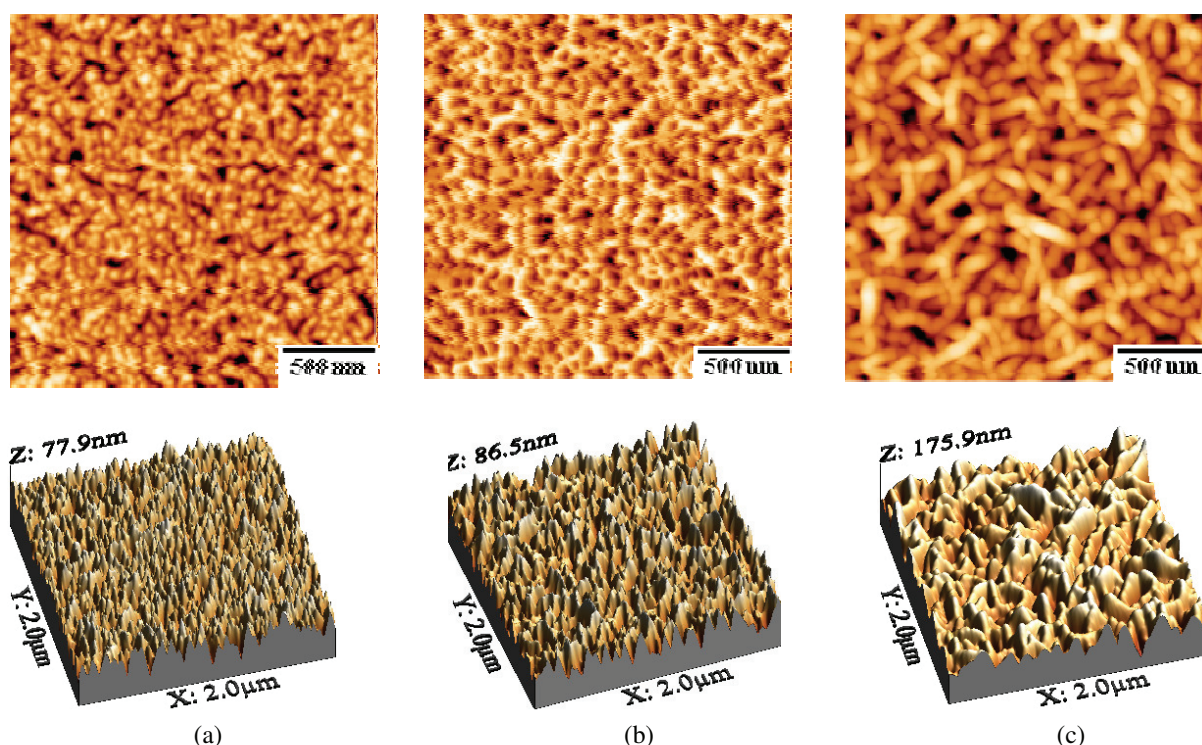


FIGURE 1. Typical 2-D and 3-D AFM images of homoepitaxial GaN layers grown on GaN template at growth temperature of 700°C with laser repetition rates of (a) 10 Hz, (b) 20 Hz and (c) 30 Hz. Scan area: $2 \mu\text{m} \times 2 \mu\text{m}$.

Surface morphologies of the LMBE grown GaN films on GaN template at growth temperature of 700°C with different laser repetition rates of 10 to 30 Hz were investigated using AFM and shown in Figure 1. The AFM images also revealed that the GaN films have grown in 3D island growth mode. For the GaN film grown on GaN template at 10 Hz, mostly granular film with a surface rms roughness of ~7.6 nm has been observed [Fig.1(a)], but the surface roughness increases to ~12.1 nm when the laser repetition rate increased to 20 Hz [Fig.1(b)]. On further increasing the laser repetition rate to 30 Hz, the surface roughness increases to ~19.6 nm and nanoworm/curved like surface structures have been obtained [Fig.1(c)]. The field-emission scanning electron microscopy (FE-SEM) image of the 30 Hz grown GaN film showed the formation of prominent GaN nanowall network [15]. This increment in rms roughness with increasing the laser repetition rate is attributed to the increment in variation of lateral width of islands as shown by the histograms after statistical analysis of AFM images. In [Fig.2(a)], histogram of sample grown at 10 Hz laser repetition rate showed that lateral size of GaN islands is quite uniform (70-80 nm). This may be due to the better surface diffusion of Ga/N adatoms on growth front at the low repetition rate of 10 Hz (low growth rate). On increasing the laser repetition rate to 20 Hz, islands with lateral size of ~90-120 nm have been observed as shown by histogram [Fig.2 (b)]. On further increasing laser repetition rate to 30 Hz, histogram in Fig.2(c) showed increment in lateral size of island ~120-160 nm because at higher laser repetition rate, the number of GaN adatoms arriving onto the surface per unit time is higher which restricts the adatom surface mobility, leading to formation of bigger islands (more favorable 3D growth) with an enhanced surface roughness.

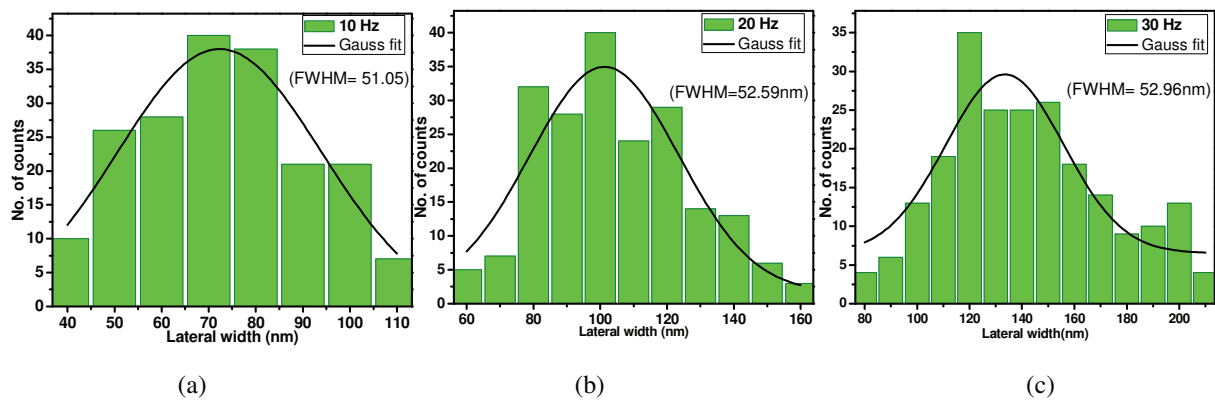


FIGURE 2. Histograms representing the distribution of grain size width of GaN layers grown at (a) 10 Hz, (b) 20 Hz and (c) 30 Hz.

It has been also observed that on increasing the laser repetition rate from 10 Hz to 30 Hz, variation in island lateral size also increases as shown by the increment in full width at half maxima (FWHM) values of Gaussian fitted histograms (51.05 nm for 10 Hz, 52.59 nm for 20 Hz, 52.96 nm for 30 Hz). The adjacent GaN islands can more likely merge together due to higher flux rates at 20 and 30 Hz

Table 1. Grain size and rms roughness of homoepitaxial GaN layers grown at different laser repetition rates

Laser repetition rate (Hz)	Grain size (nm)	Rms roughness (nm)
10 Hz	70-80	7.6
20 Hz	90-120	12.1
30 Hz	120-160	19.6

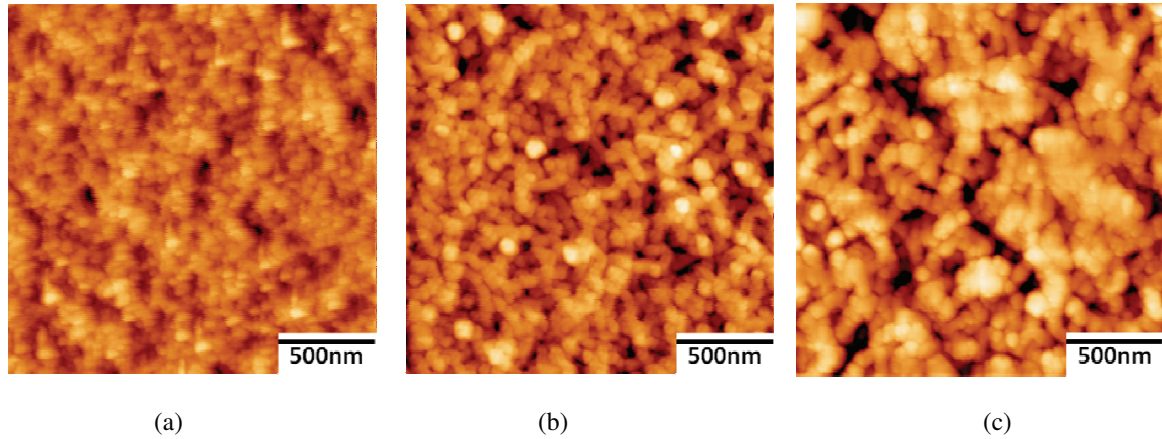


FIGURE 3. Typical 2-D AFM images after etching in KOH solution for GaN layers grown at laser repetition rates of (a) 10 Hz, (b) 20 Hz and (c) 30 Hz. Scan area: $2\ \mu\text{m} \times 2\ \mu\text{m}$.

Fig.3 shows the AFM images of the GaN films after wet chemical etching with 1:5 weight ratio KOH solution for 30 min. For the GaN film grown at 10 and 20 Hz, there is no abrupt change in the morphology. However, the sample grown at 30 Hz, shows a clear change in surface morphology after etching in KOH solution. Since the sample grown at 30 Hz has a lot of curved shape structures, it is likely that the etching rate is fast for such kind of structures. The previous first-principle theoretical analysis showed that the etching reaction take place predominantly at kink site of Ga-terminated GaN (0001) surfaces compared to the step-terrace structures [16].

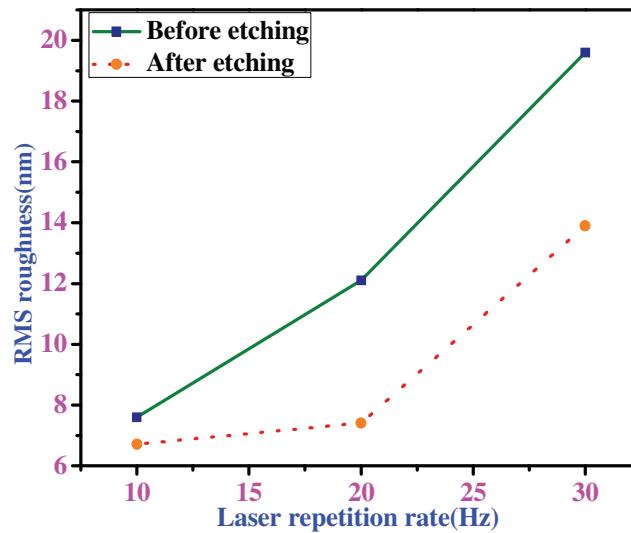


FIGURE 4. Variation of surface rms roughness before and after wet etching of GaN layers grown at different laser repetition rates.

A decrement in rms surface roughness has been observed for all the samples after etching as shown in Fig.4. It has been reported that long time (> 30 sec) etching with KOH solution of the N-polar GaN surface promote the rougher surface [17]. Since we have used the Ga-polar GaN template, it is likely that the LMBE grown GaN film follow the template polarity [18]. The decrement in rms roughness as well as abrupt change in surface morphology after etching is more for sample grown at 30 Hz repetition rate suggests that the etching rate also depends on the surface features i.e. whether surface is smooth or rough. It has been reported that there is no change in the surface morphology measured by the optical microscope after etching of Ga-polar surface using KOH solution for 30 min at room temperature [12]. In contrast to previous report [12], we want to emphasize that the spatial resolution is limited

by the optical microscope compared to AFM and the change in surface rms roughness is likely possible before and after etching especially for rough/curved surfaces when the surface morphology studied by the high resolution imaging techniques.

CONCLUSION

The GaN homoepitaxial layers were grown on GaN (0001) templates using a LMBE system at various laser repetition rates. We found that the surface morphology of the GaN layer critically depends on the laser repetition rates during ablation of GaN target. Mostly 3D growth of GaN has been observed due to the N-rich growth condition. The wet chemical etching of LMBE grown GaN film showed the smoother film compared to the as-grown film for the nano-structured film. This observation indicates that wet chemical etching critically depends on the surface structures apart from the polarity, and several etch conditions should be used when evaluating different nano-structured GaN films. The obtained results suggest that the LMBE technique could be an alternative for growth of high quality homo-epitaxial nano-structured GaN films for future III-nitride based devices.

ACKNOWLEDGMENTS

The authors are grateful to Director, NPL for the constant encouragement and support. This work was funded by Council of Scientific and Industrial Research (CSIR) under project PSC-0109. BSC would like to acknowledge the financial support by RTU Kota.

REFERENCES

- [1] S. Adachi, *Properties of Group-IV, III-V and II-VI Semiconductors* (Wiley, England, 2005) pp. 1-368.
- [2] L. Zhang, J. Yu, X. Hao, Y. Wu, Y. Dai, Y. Shao, H. Zhang and Y. Tian, *Scientific Reports* **4**, 4179 (2014).
- [3] S. S. Kushvaha, P. Pal, A. K. Shukla, A. G. Joshi, G. Gupta, M. Kumar, S. Singh, B. K. Gupta, and D. Haranath, *AIP Advances* **4**, 027114 (2014).
- [4] T.H. Seo, A.H. Park, S. Park, Y.H. Kim, G.H. Lee, M.J. Kim, M.S. Jeong, Y.H. Lee, Y.B. Hahn and E.K. Suh, *Scientific Reports* **5**, 7747 (2015).
- [5] R. D. Vispute, V. Talyansky, R. P. Sharma, S. Choopun, M. Downes, T. Venkatesan, K. A. Jones, A. A. Iliadis, M. A. Khan and J. W. Yang, *Appl. Phys. Lett.* **71**, 102 (1997).
- [6] Y. Kawaguchi, J. Ohta, A. Kobayashi and H. Fujioka, *Appl. Phys. Lett.* **87**, 221907 (2005).
- [7] S. S. Kushvaha, M. Senthil Kumar, K. K. Maurya, M. K. Dalai and N. D. Sharma, *AIP Adv.* **3**, 092109 (2013).
- [8] M. Senthil Kumar, S. S. Kushvaha and K. K. Maurya, *Sci. Adv. Mater.* **6**, 1215 (2014).
- [9] L. K. Li, M. J. Jurkovic, W. I. Wang, J. M. Van Hove, and P. P. Chow, *Appl. Phys. Lett.* **76**, 1740 (2000).
- [10] J. A. Majewski, G. Zandler and P. Vogl, *Phys. Stat. Sol. (a)* **179**, 285 (2000).
- [11] H. Morkoç, R. Cingolani, B. Gil, *Solid State Electronics* **43**, 1909 (1999).
- [12] D. Li, M. Sumiya, S. Fuke, D. Yang, D. Que, Y. Suzuki, and Y. Fukuda, *J. Appl. Phys.* **90**, 4219 (2001).
- [13] V. Darakchieva, B. Monemar, and A. Usui, *Appl. Phys. Lett.* **91**, 031911 (2007).
- [14] E. J. Tarsa, B. Heying, X. H. Wu, P. Fini, S. P. DenBaars, and J. S. Speck, *J. Appl. Phys.* **82**, 5472 (1997).
- [15] S. S. Kushvaha, M. Senthil Kumar, A. K. Shukla, B. S. Yadav, Dilip K. Singh, M. Jewariya, S. R. Ragam and K. K. Maurya, *RSC Adv.* **5**, 87818 (2015).
- [16] M. Oue, K. Inagaki, K. Yamauchi and Y. Morikawa, *Nanoscale Research letters* **8**, 232 (2013).
- [17] D.C. Du, J.C. Zhang, X.X. Ou, H. Wang, K. Chen, J.S. Xue, S.R. Xu and Y. Hao, *Chin. Phys. B* **20**, 037805 (2011).
- [18] P.R. Hageman, V. Kirilyuk, W.H.M. Corbeek, J.L. Weyher, B. Lucznik, M. Bockowski, S. Porowski and S. Muller, *J. Cryst. Growth* **255**, 241 (2003).

Blood Generation from Stem Cells: An Overview

Vimal Kishor Singh^{1*}, Abhishek Saini¹, Neeraj Kumar¹, Manisha Kalsan¹

¹Stem Cell Research Laboratory, Department of Biotechnology, Delhi Technological University, Delhi

²Correspondence: Dr. Vimal Kishor Singh O/I Stem Cell Research Laboratory, Department of Biotechnology, Delhi Technological University, Delhi, India

Abstract: *Blood transfusions are the mainstay in modern clinical modalities. The worldwide spread blood banks secure the supply of various blood components which solely depend upon voluntary paid/unpaid donor. According to the WHO reports present collections and supply of safe/adequate blood is inadequately distributed among economically developed and developing countries. Further, blood transfusions are prone to a number of risks/ adverse reactions that may be imposed due to inefficient blood screening and processing methods and put extreme supply constraints, especially in the underdeveloped countries. In past few years stem cells from different sources are demonstrated to give rise to the significant amount of functionally mature erythrocytes, platelets and neutrophils by using a cocktail of growth factors under standard culture conditions. Therefore, a logistic approach to ex vivo manufacturing of erythrocytes (& other blood components) through modern tissue culture techniques has surfaced in the past few years. Ex vivo expansion of erythrocytes to generate transfusable units of clinical grade RBCs from a single dose of starting material (e.g. umbilical cord blood) has been well established. More and more information is coming from the various research groups showing the use of other potential sources (e.g. embryonic stem cells, induced pluripotent stem cells, and stress induced erythroid progenitors), different cocktails of growth factors, methodologies to get optimal yields of the functional RBCs/ blood cells. One can imagine that with the advent of more defined methods for large-scale production of blood cells the existing system of donor derived blood supplies could be replaced in future. However, various regulating factors should essentially control the fate of large-scale production of RBCs including issues related to the quantity/ quality and per unit cost of these products. These issues would be solved with the invention of more automated culture techniques such as the use of bioreactors and identification of more direct molecular agents regulating all phases of their ex vivo generation such as stem cell proliferation, differentiation, and maturation of final products. Recent studies have shown the possibilities of generating several hundred units of erythrocytes by using continuously feeding hollow fiber bioreactors along with other types of bioreactors. Even more, the functionality and acceptability of ex vivo cultured RBCs is demonstrated by their in-human transfusion which provides essential motivation to develop more sophisticated methods/culturing techniques to pave the way for commercial production of blood components at a feasible cost.*

Keywords: Ex vivo RBCs expansion, Blood Pharming, iPSCs, RBCs, Manufacturing blood, CD34⁺ HSCs.

1. Introduction

Modern clinical modalities revolve around blood bank system which is an inevitable part of almost all the surgical interventions and support all kinds of blood transfusion needs [1]. According to the WHO reports about 103 million blood units are collected across the globe every year [2]. However, huge discrepancies occur among economically developed and developing countries as their accessibility to safe and adequate blood supplies are concerned. The people living in developing countries have limited supplies of safe blood as compared to the economical developed countries. Developed countries have a small population (15% of the global population) but have access to more than 50% of the total global blood collection/supplies per annum. WHO reports states a higher rate of blood donations scored (>30,000 annual blood donations/per blood center) by ~8000 blood centers scattered in 159 high-income countries [3]. This indicates that safe and sufficient blood units are comparatively easily available in the developed countries and the blood collection seems to be sufficient for the time being at least. While, the situation is not so comfortable in economically developing countries where only 3700 blood collections per center are reported. Further, almost~ 80% of the world population resides in these countries which reduces the per person availability of the adequate blood supplies up to a significant level. This is reported by WHO showing that there are only 10 donations per 1000 people in the 82 low

income and middle income countries [4]. The situation is worsened due to inefficient screening facilities available in most of the underdeveloped countries. It is reported that ~39 countries lack efficient screening facilities for most severe transfusion-transmissible infections (TTIs) viz. HIV, hepatitis B, hepatitis C, and syphilis [4]. Further, only 106 countries are equipped with national blood transfusion guidelines and only 13% of the low-income countries have established a national haemovigilance system to monitor the quality of the transfusion processes. An important fact is that the proportion of elderly population is rapidly growing most developed countries (>60 years old) and in the absence of young donors it would be difficult to meet the demands for blood transfusions even in the developed countries by the year of 2050 [5].

In the absence of sufficient blood inventories, it would be hard to find exact match for >30 blood group system with 308 recognized antigens including ABO & Rh antigens [6]. In addition, there are demands for rare phenotypic blood, patients with various hemoglobinopathies, polytransfusion patients and polyimmunization which put more constraints on the present donor derived blood supply system across the globe [4]. Although, there are few reports showing procedures to develop universal blood group by using antigen masking and/or enzymatic cleavage techniques but these methods are not yet developed properly [7]. Apart from these supplies related problems, the association of various types of adverse reactions with the donor derived blood

supplies system has been challenging the clinicians to find more suitable alternatives to avoid these problems. Although, a number of testing, preclinical precautionary measures are taken to avoid any infectious agent [e.g., HIV-1, HIV-2, HTLV-1, HTLV-2, Hepatitis B, Hepatitis C, Syphilis (*T pallidum*), Chagas disease (*T cruzi*), and West Nile Virus, Cytomegalovirus (CMV)], but blood transfusions are often subjected to cause several complicated situations [1, 3-4]. These adverse reactions/complications are reported to causes delayed cure rate with prolonged hospitalization periods and may further raise the overall cost of treatment [3]. For the last few years, researchers have been trying to find various non-donor derived sources such as artificial blood. However, the majority of artificial blood molecules such as hemoglobin-based oxygen carriers or perfluorocarbon solutions remains inefficient oxygen carriers and thus have limited functional application [8-12].

Another more promising approaches are the ex vivo expansion of erythrocytes from stem cells. As discussed in the later sections of the article a number of protocols have been derived showing a significant amount of functionally matured RBCs by using different types of stem cells.

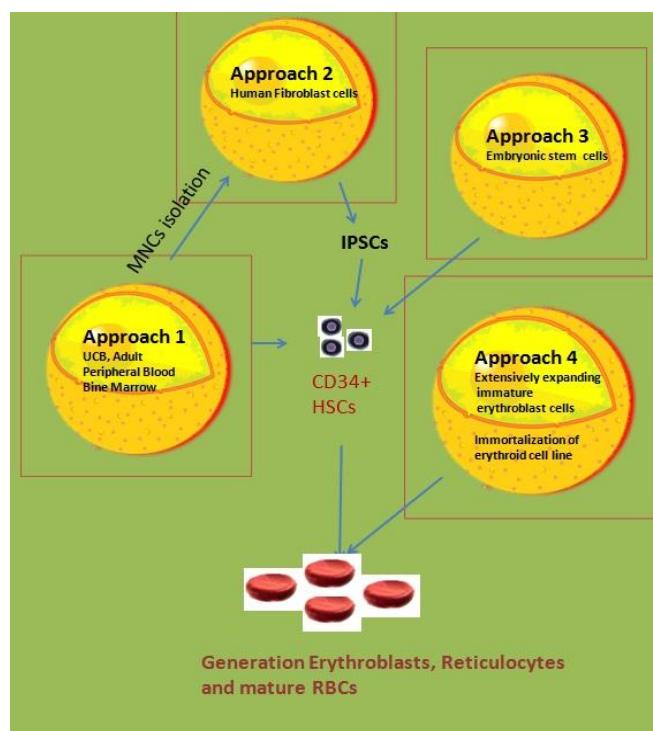


Figure 1: Various approaches in use for Ex vivo RBCs expansion: Ex vivo generation of transfusable units of RBCs can be done through different source materials and methodologies as shown above. CD34⁺ HSPCs (may be isolated from CB, PB, and BM), ESCs/IPSCs derived erythroid progenitors, and highly expanding erythroid progenitors due to stress erythropoiesis are equally potent but the overall yield, degree of maturation, methodologies used are differently developed by various research groups Ex Vivo Erythropoiesis and Different Approaches

The first report came from Nakamura and his colleague who demonstrated survival of lethally irradiated mice through

transfusion of erythroid progenitors/ cells which was generated by ex vivo differentiation of Embryonic Stem cell line (ESCs) [13].

These findings tempted many researchers who eventually developed various protocols to derive functionally matured erythrocytes from different types of stem cells viz. CD34⁺ HSCs, Embryonic stem cells, and iPSCs based methods (Figure 1). The clinical utility of cultured RBCs was established through proof-of-concept study by Giarratana et al, who shown first in-human blood transfusion of ex-vivo generated RBCs [14]. These studies indicated that expansion of erythrocytes from stem cells could become a powerful tool to solve the problems of blood supply shortages and other associated issues. With the advent of more suitable efficient culturing protocols a regular mass scale production could be feasible in near future. However, a number of relevant issues need to be defined to ensure the sufficient yields of transfusable RBCs units through these protocols. Fortunately, people have demonstrated possibilities of numerical expansion of erythrocyte/ precursors by using similar techniques. Umbilical cord blood-derived CD34⁺ cells can be expanded ~ 2 million fold to give rise to enucleated erythroid cells that are equivalent to ~0.5 transfusion units of donor derived blood [15]. Similarly, UCB derived CD34⁺ are reported to generate ~10 transfusion unit equivalents of RBCs by a technically complex protocol [16]. Although, these primitive methods are not best suitable for the large scale commercial production, but their potential up gradation to scale-up the overall production would become feasible in near future. Another lacuna is the sub-optimal enucleation efficiency of RBCs population generated by these protocols indicating their incomplete functional maturation. However, it is believed that these nucleated RBCs precursors, cellular intermediates/nucleated erythroid precursors generated through ex vivo expansion methods may also be of use if transfused. It is proposed that these maturing nucleated cells along with enucleated erythrocytes can support the cellular content at the time of transfusion and may also give rise to functionally matured cell by proliferating into 4–64 enucleated RBCs in circulation. This is supported by reports showing functional maturation of cellular intermediates/nucleated erythroid precursors in vivo on post-transfusion through their interaction with internal factors regulating their maturation. There are further studies both in human and animal models showing similar reports [15, 17]. This is also supported by several clinical reports showing occasional use of compatible Cord Blood (40-80 ml/4–7 ×10¹⁰RBCs and 2–8 ×10⁹ erythroblast cells) for transfusion in emergency situations [17].

These reports support the feasibilities of regular ex vivo RBCs expansion to overcome different clinical short-comes that exist because of supply constraints. However existing technological barriers need development of more automated and efficient ex vivo expansion methods for large scale production. In order to enhance the present scale of RBCs production a more cost-effective and timely automated ex vivo culture approaches are required. This is possible if we

have an in-depth knowledge of the various regulatory elements and their mode of action playing axial role in almost all of these methods demonstrated so far.

The present article focuses on the various approaches developed for ex vivo erythrocyte expansion and other important regulatory elements to discuss the fate of large-scale RBCs production.

During past few decades considerable progress has occurred in the understanding and therefore developing methods for biological control over the expansion of erythrocytes from hematopoietic stem cells. This information is very well exploited to generate terminally differentiated, fully functional RBCs [18]. There are different types of stem cells such as CD34+ HSPCs, Embryonic stem cells/induced pluripotent stem cells, and immortalized erythroid precursors which can be differentiated into erythroid lineages by using almost similar/overlapping protocols (Figure2). The overall ex vivo expansion process may be categorized into three broad phases Such as phase I: -commitment, phase II: expansion, and phase III: maturation or enucleation. Recent studies has demonstrated transdifferentiation (directly differentiating of human fibroblast cells into erythroblasts) to bypass the HSPC state (Figure3) [19]. These reports are promising and a detailed technical evaluation of their potential to generate large-scale RBCs would be helpful in establishing the essential guidelines.

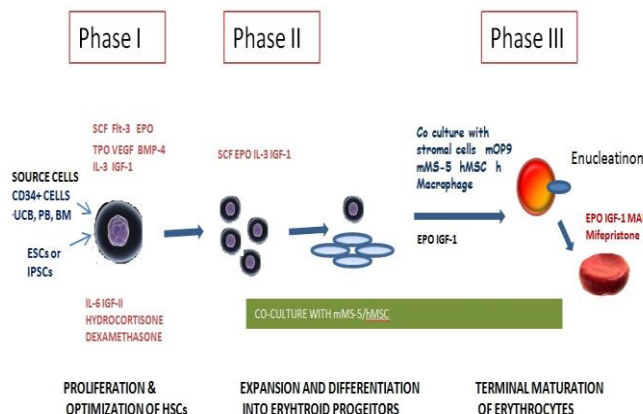


Figure 2: An overview of the methodologies/protocols: Ex vivo RBCs expansion/ manufacturing blood may involve various phases such as depicted above. Phase I deals with the initial source material collection from a variety of source material(s).these cells grown are on the basis of their availability, suitability, and expansion potential in medium generally supplemented with growth factors to enhance HSCs proliferation: subsequently, phase II entails culture of these cells in the presence of Erythropoietin to induce their differentiation and maintenance into erythropoietic progenitor stage. The final Phase has been described to support their maturation. Here, Erythropoietic progenitors may be co cultured with murine/human stromal cell line support to induce their maturation and enucleation resulting into mature RBCs. These RBCs are evaluated for their biochemical properties and various antigenic profiles to ensure their nativeness.

2. The Fate of Blood Production is Regulated by Various Factors

Efforts are being made develop optimal production methods by controlling different variables of ex vivo RBCs expansion procedures. There are a number of important factors which may be summarized as follow:

- 1) Issues related to the initial source material
- 2) Quantitative issues
- 3) Growth factors and their optimal uses
- 4) Biophysical and biochemical parameters
- 5) Issues related to the antigenic profiling
- 6) Ethnicity-related issues
- 7) Issues related to the enucleation efficiency of cultured RBCs
- 8) Production cost

The most important factor is the initial source material for various methods that are being explored for large scale RBCs production. An ideal source material would be a discarded material (reducing production cost), available in abundance, possess high potential to give rise to final product i.e. matured RBCs. and non-immunogenic. Most of these criteria fits well with umbilical CB which is easily available and have been used to produce sufficient amount of transfusable RBCs.

Another important factor is the extremely huge quantity of any blood component that would be required to produce. For example, RBCs are present in 2×10^{12} cells/unit (200ml of blood) with an expected cell density of 3×10^6 cells/ml. In order to support global population a mammoth amount of blood units (e.g. >103 million units) would be an essential requirement. Considering the present ex vivo expansion technologies an extremely large amount of ingredients and adequately skilled manpower would be an inevitable requirement. This is hypothesized that a cell density of 5×10^7 cells/ml should be produced regularly to support this much demand. It is estimated that production of single unit of blood through present static culture methods will require 660 liters of culturing medium and 9500 lab scale 175cm² culturing flasks (70ml medium/flask) [20-22]. This indicates the need for more automated 3D culture techniques to ensure such high production efficiency.

Fortunately, use of bioreactors may be helpful in reducing both the culturing assets and associated labor. This could help researchers achieve required cell density with limited efforts. It is reflected from the reports showing production of large density of RBCs in a wave bioreactor [22].

Similar results are reported (cell density of 10^7 cells/ml) by using a small-scale stirred bioreactor [23]. Further, a massive yield of 2×10^8 cells/ml was produced in a hollow fiber bioreactor [24]. The point is that with the advent of the technology these 3D bioreactors may be scaled up with to produce desired units of RBCs in 3–4 weeks of time. Thus it seems feasible to produce large amounts of RBCs, but before that many factors such as degree of cell maturation and

various cultural parameters should be defined in detail. Next, there are few reports showing the preliminary data on direct transformation of stem cells into the erythropoietic lineages which will dramatically reduce both the time and cost of production. For example, generation of erythropoietic cell lines from ESCs/IPSCs through reprogramming of any somatic cells directly into erythroblasts using suitable genes is reported by some researchers. Cheng's group has reported in humans that the differentiated erythroid cells obtained from human embryo might be used as a potential initial source for RBCs production [25-26].

In brief, it is conceivable that limited production of RBCs may take place in near future. The upcoming advancements in the existing tissue/ cell culture technology and identification of most efficient source materials would ensure sufficient ex vivo expansion of RBCs/ other blood cells in a regular fashion. The quality issues of these products may be resolved with the development of highly stringent, comprehensive quality control, analytical measures. As discussed earlier, all types of stem cells (e.g. CD34⁺ HSCs, hESCs and, iPSCs) have enough potential to support the production of clinical grade blood cells in huge quantities and invention of more efficient source materials like erythropoietic hESC cell lines which can be directly transformed/induced to generate erythrocytes, would further ensure the unlimited supply these products. Since a number of growth factors are used in these protocols and the molecular mechanisms involved in various stages of ex vivo erythropoiesis governed through these growth factors is not precisely defined. The exploration of these fundamental processes would be helpful in developing pharmacological alternates for the various stages of erythropoietic development.

Next, use of new pharmacological alternates can further simplify the production processes for clinical grade blood by removing the need for feeder cells and thus may help in reducing the unexpected risks associated with their use due to undefined immunogenic and infectious exposures. The omission of feeder cells from these methods is not only essential for the generation of GMP products but is also required to simplify the commercial scale production methods. Use of straightforward and simple procedures may also be helpful in reducing the cost of production.

3. Acknowledgement

We are thankful to the Honorable Chairman and the Honorable Vice-Chancellor of the Delhi Technological University, Shahbad Daultapur, Bawana road, Delhi-42, for support. Dr. Vimal K. Singh particularly thanks the Department of Science and Technology and Indian National Science Academy (INSA), INDIA, for the research grant.

4. Conflict of Interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation

in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

- [1] Alter, H. J., and Klein, H. G. (2008). The hazards of blood transfusion in historical perspective. *Blood* 112, 2617–2626. doi: 10.1182/blood-2008-07-077370
- [2] World Health Organization. (2011). Global Database on Blood Safety: Summary Report 2011. Available online at: http://www.who.int/bloodsafety/global_database/GDBS_Summary_Report_2011.pdf (Accessed May 4, 2014).
- [3] Department of Health and Human Services. (2010). The 2009 National Blood Collection and Utilization Survey Report. Washington, DC: DHHS. Available online at: <http://www.aabb.org/programs/biovigilance/nbcus/Documents/09-nbcus-report.pdf> (Accessed May 4, 2014).
- [4] Department of Health and Human Services. (2013). The 2011 National Blood Collection and Utilization Survey Report. Washington, DC: DHHS. Available online at: <http://www.aabb.org/programs/biovigilance/nbcus/Documents/11-nbcus-report.pdf> (Accessed May 4, 2014).
- [5] U.S. Census Bureau. (2004). Global Population Composition. Available online at: <http://www.census.gov/population/international/files/wp02/wp-02004.pdf> (Accessed September 14, 2012).
- [6] Daniels, G., Castilho, L., Flegel, W. A., Fletcher, A., Garratty, G., Levene, C., et al. (2009). International society of blood transfusion committee on terminology for red blood cell surface antigens: macao report. *Vox Sang.* 96, 153–156. doi: 10.1111/j.1423-0410.2008.01133.x
- [7] Bagnis, C., Chiaroni, J., and Bailly, P. (2011). Elimination of blood group antigens: Hope and reality. *Br. J. Haematol.* 152, 392–400. doi: 10.1111/j.13652141.2010.08561.x
- [8] Winslow, R. M. (2006). Current status of oxygen carriers ("blood substitutes"): 2006. *Vox Sang.* 91, 102–110. doi: 10.1111/j.1423-0410.2006.00789.x
- [9] Henkel-Honke, T., and Oleck, M. (2007). Artificial oxygen carriers: a current review. *AANA. J.* 75, 205–211.
- [10] Natanson, C., Kern, S. J., Lurie, P., Banks, S. M., and Wolfe, S. M. (2008). Cellfree hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. *J. Am. Med. Assoc.* 299, 2304–2312. doi: 10.1001/jama.299.19.jrv80007.
- [11] Silverman, T. A., and Weiskopf, R. B. (2009). Hemoglobinbased oxygen carriers: Current status and future directions. *Transfusion* 49, 2495–2515. doi: 10.1111/j.1537-2995.2009.02356.x.
- [12] Castro, C. I., and Briceno, J. C. (2010). Perfluorocarbonbased oxygen carriers: Review of products and trials. *Artif. Organs* 34, 622–634. doi: 10.1111/j.15251594.2009.00944.x
- [13] Hiroyama, T., Miharada, K., Sudo, K., Danjo, I., Aoki, N., and Nakamura, Y.

- (2008). Establishment of mouse embryonic stem cell-derived erythroid progenitor cell lines able to produce functional red blood cells. *PLoS ONE* 6:3. doi: 10.1371/journal.pone.0001544.
- [14] Giarratana, M. C., Rouard, H., Dumont, A., Kiger, L., Safeukui, I., Le Pennec, P. Y., et al. (2011). Proof of principle for transfusion of in vitro-generated red blood cells. *Blood* 118, 5071–5079. doi: 10.1182/blood-2011-06-362038
- [15] Neildez Nguyen, T. M., Wajcman, H., Marden, M. C., Bensidhoum, M., Moncollin, V., Giarratana, M. C., et al. (2002). Human erythroid cells produced ex vivo at large scale differentiate in to red blood cells in vivo. *Nat. Biotechnol.* 20, 467–472. doi: 10.1038/nbt0502-467.
- [16] Fujimi, A., Matsunaga, T., Kobune, M., Kawano, Y., Nagaya, T., Tanaka, I., et al. (2008). Ex vivo large-scale generation of human red blood cells from cord blood CD34+ cells by co-culturing with macrophages. *Int. J. Hematol.* 87, 339–350. doi: 10.1007/s12185-008-0062-y
- [17] Ende, M., and Ende, N. (1972). Hematopoietic transplantation by means of fetal (cord) blood. A new method. *Va. Med. Mon.* 99, 276–280.
- [18] Migliaccio, A. R., Masselli, E., Varricchio, L., and Whitsett, C. (2012). Ex-vivo expansion of red blood cells: how real for transfusion in humans? *Blood Rev.* 26, 81–95. doi: 10.1016/j.blre.2011.11.002
- [19] Szabo, E., Rampalli, S., Risueno, R. M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., et al. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* 468, 521–526. doi: 10.1038/nature09591
- [20] Timmins, N. E., and Nielsen, L. K., (2009). Blood cell manufacture: current methods and future challenges. *Trends Biotechnol.* 27, 415–422. doi: 10.1016/j.tibtech.2009.03.008
- [21] Zeuner, A., Martelli, F., Vaglio, S., Federici, G., Whitsett, C., and Migliaccio, A. R. (2012). Concise review: stem cell-derived erythrocytes as upcoming players in blood trans- fusion. *Stem Cells* 30, 1587–1596. doi: 10.1002/stem.1136
- [22] Timmins, N. E., Athanasas, S., Günther, M., Buntine, P., and Nielsen, L. K. (2011). Ultrahigh-yield manufacture of red blood cells from hematopoietic stem cells. *Tissue Eng. Part. C Methods* 17, 1131–1137. doi: 10.1089/ten.tec.2011.0207
- [23] Ratcliffe, E. I., Glen, K. E., Workman, V. L., Stacey, A. J., and Thomas, R. J. (2012). A novel automated bioreactor for scalable process optimisation of haematopoietic stem cell culture. *J. Biotechnol.* 161, 387–390. doi: 10.1016/j.jbiotec.2012. 06.025.
- [24] Housler, G. J., Miki, T., Schmelzer, E., Pekor, C., Zhang, X., Kang, L., et al. (2012). Compartmental hollow fiber capillary membrane-based bioreactor technology for in vitro studies on red blood cell lineage direction of hematopoietic stem cells. *Tissue Eng. Part. C Methods* 18, 133–142. doi: 10.1089/ten.tec.2011.0305
- [25] Huang, X., Shah, S., Wang, J., Ye, Z., Dowey, S. N., Tsang, K. M., et al. (2013). Extensive ex vivo expansion of functional human erythroid precursors established from umbilical cord blood cells by defined factors. *Mol. Ther.* 22, 451–463. doi: 10.1038/mt.2013.201
- [26] Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., et al. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS ONE* 8:3 doi: 10.1371/journal.pone.0059890

Author Profile



Dr. Vimal Kishor Singh received the M.Sc. in Bio-Medical Sciences from University of Delhi. PhD from Institute of Nuclear Medicine and Allied Sciences (DRDO)/University of Delhi Founder and O/I of **Stem Cell Research Laboratory**, Dept. of Biotechnology, Delhi Technological University



Abhishek Saini received the B.Sc. Microbiology (H) degree from Delhi University and M.Sc. in Biotechnology from TERI University. At present pursuing Ph.D. from Delhi Technological University.



Neeraj Kumar received the B.Sc. degree in Chemistry from Delhi University and M.Sc. degree in Biochemistry from Jamia Millia Islamia University and M.Tech degree in Bioinformatics from Delhi Technological University. Presently, pursuing his Ph.D. from Delhi University.



Manisha Kalsan received a B. Tech degree in Biotechnology from D. C. R. University of Science and Technology, Murthal and an M. Tech degree in Bioinformatics from Delhi Technological University and is presently pursuing her Ph.D. from Delhi University.

CFD Analysis of Angle Axial Annular Diffuser with Both Hub and Casing at Diverging Equal Angles

B.B.ARORA¹

¹(Deptt. of Mech. Engg., Delhi Technological University, Delhi, India)

¹prof_bbar@yahoo.com

Abstract : The development of flow inside the axial annular diffuser whose hub and casing walls are diverging at equal angles have been predicted with the help of computational fluid dynamics (CFD). The equivalent cone angle for axial annular diffuser has been taken as 20° and area ratio 2 and 3 for the present study has been taken as for The analysis has been carried out for flow regime with various experimentally obtained inlet velocity profiles for area ratio 2 with or without swirl. The flow behaviour inside the annular diffuser was analyzed for the same experimental geometry with CFD and pressure recovery coefficient was computed which agrees reasonably well with the experimental/available data. Further prediction of expected flow behaviour inside the annular diffuser with area ratio 3 has been carried out. Pressure recovery coefficient on the mass average basis has also been computed.

Keywords: Annular diffusers; pressure recovery coefficient; swirl, equivalent cone angle

I. INTRODUCTION

The function of a diffuser is the efficient conversion of kinetic energy into pressure. Many fluid-dynamical systems involve diffusion in annular passages. Need for annular diffusers may arise from a necessity to provide a central core to allow a coaxial shaft in a given situation. Annular diffusers are widely used in engineering, in particular, as outlet devices of pumps and turbines often located downstream of turbo machinery in a number of applications. In aircraft application, annular diffusers often operate downstream of compressors. Such diffusers handle flows having substantial amount of swirl and unsteadiness made up of turbulence and periodic flow components introduced by the turbo machinery. The swirling component of velocity may arise either from the presence of inlet guide vanes or any other components preceding the diffuser, e.g., a compressor, or from rotation of the central shaft through the diffuser. The introduction of presence of swirl alters the flow field considerably and this affects the overall performance of a system.

Swirling flows through annular diffusers have been investigated by many researchers such as Sovran and Klomp [3], Shrinath^[4] Hoadley^[7], Colodipietro et al. [8], Shaalan et al. [9], Kumar^[10], Lohmann et al. [11] and Sapre et al. [13], Agrawal et al.^[14], Singh et al.^[16, 23], Kochevsky^[18], Mohan et al. [19], Japikse, D^[20], Kochevsky, A. N^[21] and Yeung et al.^[22], Manoj et al.^[266, 83], Arora et al.^[24, 25, 29] These investigators found improved diffuser performance with swirl till a point after that it deteriorated. The performance of an annular diffuser apart from swirl is dependent on a large number of geometrical and dynamical parameters. The effectiveness of annular axial diffusers worsens with flow separation. The separation of the flow can be suppressed or shifted from one location to another with the help of swirl. The efforts have been made to design

an annular diffuser for no flow separation^[2,5,6], however little success has been achieved.

Literature on annular diffusers reveals that earlier studies have been carried out either with parallel hub diverging casing and both hub and casing diverging. The experimental/ analytical data on the pressure recovery coefficient or coefficient of energy losses^[1, 12] for a wide range of geometrical parameters and swirl intensities are scant. Experimental studies on annular diffuser^[17] require sophisticated instrumentation and complicated time-consuming procedures which is not economically viable and thus has limited the research activity in the field of annular diffusers^[12].

The present study is therefore carried out to examine with the help of Computational Fluid Dynamics (CFD), the detailed flow behavior of axial annular diffusers with parallel hub diverging casing for same equivalent cone angle of 20° and area ratio of 2 and 3. Experimental velocity profiles were obtained with the axial annular diffuser having hub parallel and casing diverging and area ratio of the experimental diffuser was 2.01, and equivalent cone angle of 20.02°. CFD analysis of the diffuser with same configuration and dynamic parameters was carried out with different turbulence models. The model which predicted the results more closely with the experimental results was chosen for further investigations. CFD Study has been carried out to predict the effect of experimentally obtained inlet velocity profiles without swirl (0°) and with inlet swirl angles of 7.5°, 12°, 17° and 25° on the performance of annular diffusers.

II. EXPERIMENT SETUP

Figure 1(a) shows the actual experimental setup used for the present investigation. The test rig consists of a single stage centrifugal blower which delivers 1.5m³/s at 1m water gauge

pressure. It draws air from the atmosphere through a very fine mesh filter and delivers it to a settling chamber through a well-designed conical divergence. A symmetrical damper placed at the blower intake controlled and kept the flow rate constant through the system. A piece of heavy fabric serving as flexible coupling was used to seal the gap between the blower and settling chamber in order to prevent the vibrations reaching to settling chamber from the blower. The settling chamber is provided with fine mesh screens and a honey comb section. The purpose of the screens is multifold in serving as flow steadying and straightening, reducing the level of turbulence and losses. It is further connected to a constant-area annular duct made up of two commercial pipes; through a smooth converging section. Smooth transition from the annulus to the conical casing of the diffuser was ensured by inserting suitable metal shims between flanges and the inside was finished off with plasticine. Diffuser hub was made from cast aluminum and machined smooth whereas the casing of the annular test-diffusers was made of transparent Plexiglas. This was done to permit flow visualization inside the annulus so formed. The air flowing through the diffuser was finally exhausted into the atmosphere.

The measurements of static pressure and yaw angle were made manually with the help of Cobra probe, Traversing Mechanism and Manometers.



Fig. 1. (a) Actual Diffuser test setup

Figure 2 shows Annular diffuser Geometrical Parameters of the half section as the diffuser has been taken as axially symmetrical.

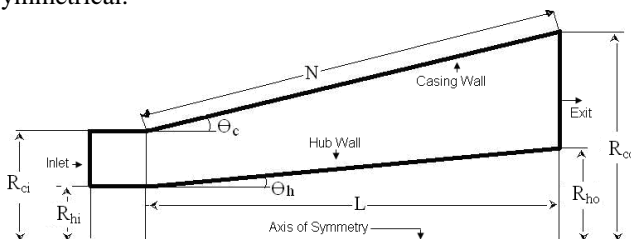


Fig. 2. Annular Diffuser Geometrical Parameters

III. CFD MODELLING

Annular diffuser geometry was sketched with proper meshing scheme with the help of GAMBIT software. In the pre-study $k-\epsilon$ turbulence models such as standard, RNG and realizable were tried on the geometries for which experimental data were available. The results obtained were validated with the available experimental results. The boundary conditions fed at the inlet is the same velocity profile as experimentally obtained

with turbulence specification of 3% turbulence and hydraulic diameter as calculated from the geometry of the diffuser inlet. The outlet boundary condition is pressure normal to the pressure outlet with turbulence specification of 3% turbulence and hydraulic diameter as calculated from the geometry of the diffuser outlet. The solution controls for momentum, swirl velocity, turbulence kinetic energy and turbulence dissipation rate are second order up winding. The convergence criteria for residuals are 10^{-6} for various parameters involved in the present study such as continuity, velocity components v_x , v_r , and v_z , swirl, k and ϵ ; the results were found to be stable.

The modeling was repeated for various mesh sizes varying from 50000 to 500000 mesh cells to attain the grid independence. It was found that the model which approached more closely to the experimental results was 2D axisymmetric RNG "renormalization group" $k-\epsilon$ turbulence model with moderate mesh size of 0.07cm. The RNG-based $k-\epsilon$ turbulence model [15] is derived from the instantaneous Navier-Stokes equations, using a mathematical technique called "renormalization group" (RNG) methods. The same model has been used for carrying out the analysis for other geometries considered for the present study.

The governing equations for 2D axisymmetric geometries used are the same as used by Arora[]

Grid independence Tests were carried out on number of geometries with regards to meshing and turbulence modeling. Various turbulence models and meshes were tested with inlet velocity profile which was obtained experimentally for a fully developed flow on an annular diffuser having an equivalent cone angle of 20° whose both casing and hub were diverging and area ratio of 2. The turbulence model which is in close proximity with the experimental results as shown in Fig 3 is Renormalization-group (RNG) $k-\epsilon$ model.

$k-\epsilon$ RNG model with mesh size of 0.07 was considered for the present CFD analysis to reduce the computational time without foregoing the desired accuracy.

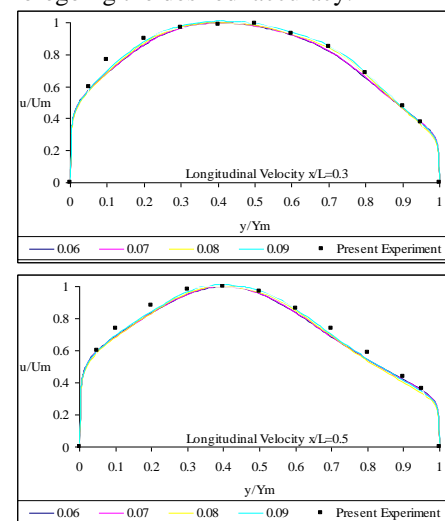


Fig. 3. Grid Independence Test of Longitudinal Velocity (0°) for 20° equivalent cone angle diffuser

Pressure recovery Coefficient

The static pressure rise non-dimensionalised with respect to the diffuser inlet dynamic head is defined as the static pressure recovery [29].

For one-dimensional flow of perfect gas without any energy loss, the ideal pressure rise for given diffuser can be computed by considering energy conservation.

The same turbulence model was used to predict the performance of axial annular diffusers with various geometries and one such model was validated with experimental results as shown in Fig. 4 and Fig. 5

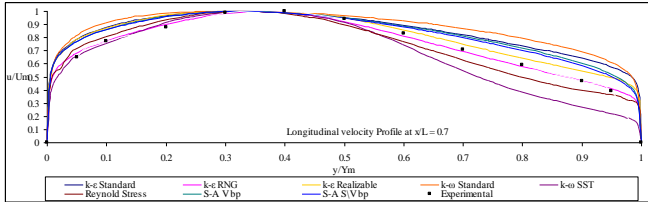


Fig. 4 Validation of Turbulence Model with Experimental Results at $x/L = 0.7$

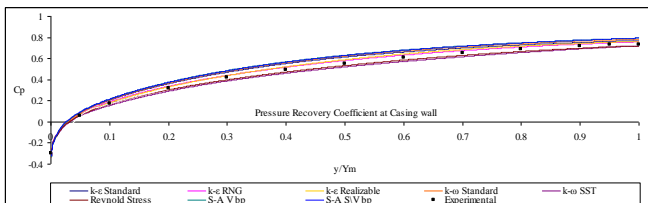


Fig. 5 Validation of Pressure Recovery Coefficient with experimental Results at the casing wall

IV. RESULTS & ANALYSIS

Velocity Profile

Figure 6-8 show the Longitudinal and swirl velocity profiles for the area ratios 2 to 4 respectively. The velocity profiles are shown for inlet swirl angles 0° , 7.5° , 12° , 17° and 25° . Fig.6 (a) to 8(a) represents the Longitudinal whereas 6 (b) to 8(b) represents the swirl velocity profiles. All the velocity profiles are shown in terms of non-dimensional velocity as local longitudinal or swirl velocity divided by local maximum velocity of the transverse where velocity is required. The non-dimensional velocity has been shown as a function of non-dimensional diffuser passage height of the particular traverse (y/Y_m). The $y/Y_m = 0$ is the hub position and for $y/Y_m = 1$, it is the casing position of the traverse. The graphs are shown at various traverses of the diffuser passage at $x/L = 0.1, 0.3, 0.5, 0.7$ and 0.9 for all the area ratios and inlet swirl angles.

Fig.6 (a) to 8(a) illustrate that the flow is hub generated for no swirl condition and the shift increases to larger extent towards hub with the increase in the area ratio for same inlet velocity profile in this case too. The peak of the velocity at $x/L = 0.9$ in the order of area ratio 2,3 and 4 occurs at y/Y_m at 0.45, 0.41 and 0.32 for 10° equivalent cone angle and for 20° equivalent cone angle it happens at $y/Y_m = 0.4, 0.34$ and 0.29 . Moreover the velocity at the casing recedes at faster rate. The flow is pushed towards the casing with the introduction of swirl.

Fig.6 (a) for area ratio 2 illustrates that the reversal of flow takes place at the casing wall for 7.5° inlet swirl at $x/L = 0.9$ from casing to $y/Y_m = 0.98$, whereas for inlet swirl of 12° and above it takes place at the hub wall. The reversal takes place at $x/L = 0.7$ for 12° inlet swirl and at $x/L = 0.5$ and 0.3 for 17° and 25° respectively. For 12° inlet swirl the flow is reversed up to $y/Y_m = 0.05$ and 0.13 for $x/L = 0.7$ and 0.9 respectively. For 17° inlet swirl the flow is reversed up to $y/Y_m = 0.05, 0.15$ and 0.23 for $x/L = 0.5, 0.7$ and 0.9 respectively. Whereas for 25°

inlet swirl the flow is reversed up to $y/Y_m = 0.004, 0.09, 0.18$ and 0.26 for $x/L = 0.3, 0.5, 0.7$ and 0.9 respectively. Fig.7 (a) for area ratio 3 illustrates that the reversal of flow takes place at the casing wall for no swirl (0°) and 7.5° inlet swirl at $x/L = 0.7$ and 0.5 respectively. For no swirl condition it takes place from casing to $y/Y_m = 0.98$ and 0.94 at $x/L = 0.7$ and 0.9 respectively, whereas for 7.5° inlet swirl it takes place from casing to $y/Y_m = 0.98, 0.91$ and 0.87 at $x/L = 0.5, 0.7$ and 0.9 respectively. For inlet swirl of 12° and above it takes place at the hub wall. The reversal takes place at $x/L = 0.5$ for 12° inlet swirl and at $x/L = 0.3$ for 17° and 25° . For 12° inlet swirl the flow is reversed up to $y/Y_m = 0.13, 0.25$ and 0.34 for $x/L = 0.5, 0.7$ and 0.9 respectively. For 17° inlet swirl the flow is reversed up to $y/Y_m = 0.07, 0.23, 0.33$ and 0.38 for $x/L = 0.3, 0.5, 0.7$ and 0.9 respectively. Whereas for 25° inlet swirl the flow is reversed up to $y/Y_m = 0.18, 0.32, 0.39$ and 0.44 for $x/L = 0.3, 0.5, 0.7$ and 0.9 respectively.

The graphs demonstrate that the velocity recedes at a faster rate near the hub as well as casing wall as the flow progresses along the diffuser passage due to the growth of boundary layer. The non-dimensional velocity near the hub wall tends to decrease with the introduction of swirl and it increases near the casing wall.

It is also evident from the graphs of non swirling flow (0°), that the location of maximum non dimensional velocity shifts towards the hub for downstream of the diffuser passage. The shift increases to larger extent with the increase in the area ratio for same inlet velocity profile. This is due to the fact that the stall increases at the casing wall with increase in the area ratio for same equivalent cone angle diffusers. The stall tends to shift from casing to the hub wall with the introduction of swirl as observed by examining the Figures 6 and 7. The shift is stronger with the increase in the inlet swirl.

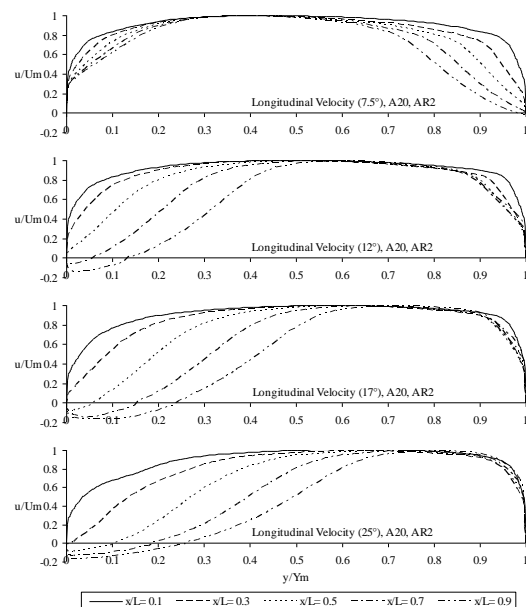


Fig. 6(a) Non-dimensional Longitudinal Velocity versus diffuser passage height for Area ratio 2 for inlet swirl angles ($\alpha = 0^\circ$ to 25°) at various traverses at $x/L = 0.1$ to 0.9 .

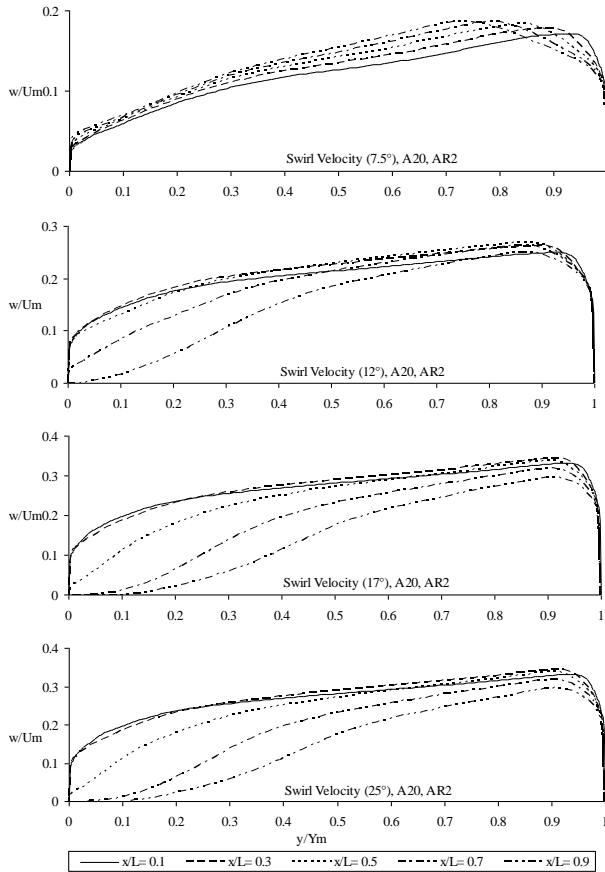


Fig. 6(b) Non-dimensional Swirl Velocity versus diffuser passage height for Area ratio 2 for inlet swirl angles ($\alpha = 0^\circ$ to 25°) at various transverses at $x/L = 0.1$ to 0.9 .

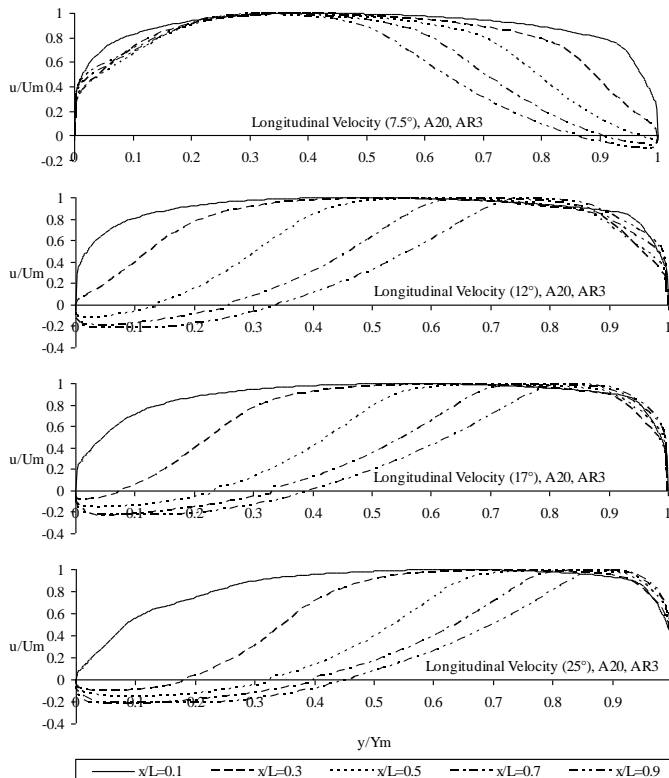


Fig. 7(a) Non-dimensional Longitudinal Velocity versus

diffuser passage height for Area ratio 3 for inlet swirl angles ($\alpha = 0^\circ$ to 25°) at various transverses at $x/L = 0.1$ to 0.9 .

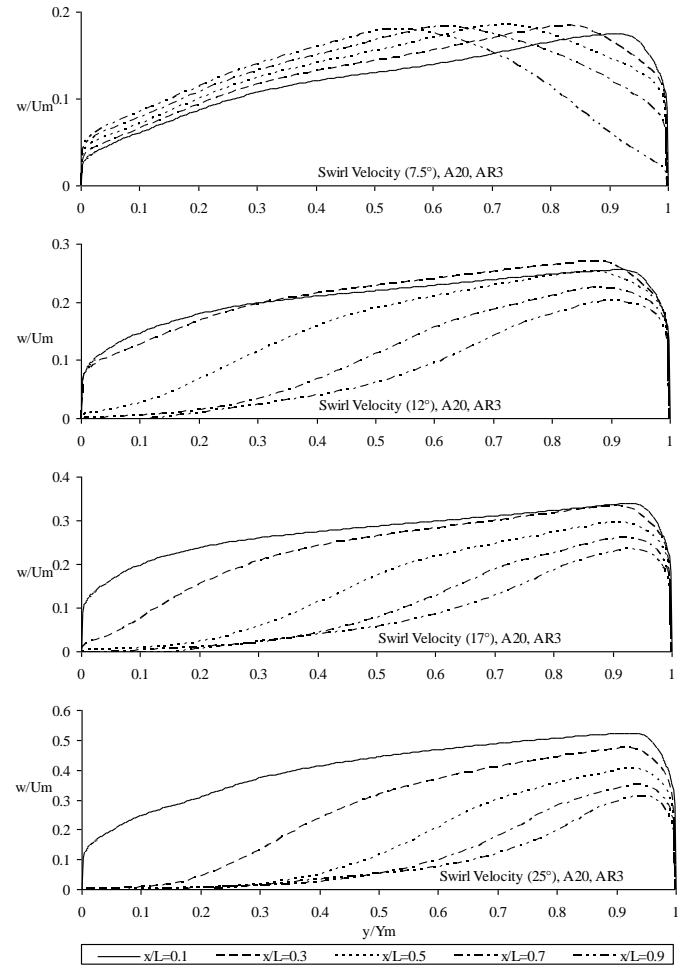


Fig. 7(b) Non-dimensional Swirl Velocity versus diffuser passage height for Area ratio 3 for inlet swirl angles ($\alpha = 0^\circ$ to 25°) at various transverses at $x/L = 0.1$ to 0.9 .

Pressure Recovery Coefficient

Figure 8 indicate pressure recovery coefficient at casing wall (C_p) for diffuser for area ratios 2 to 4 as a function of non-dimensional diffuser passage x/L for various inlet swirl angles 0° , 7.5° , 12° , 17° and 25° . C_p increases with the diffuser passage in each case. The marginal increase in C_p is sharp in the beginning of the diffuser passage and later on it decreases with the diffuser passage.

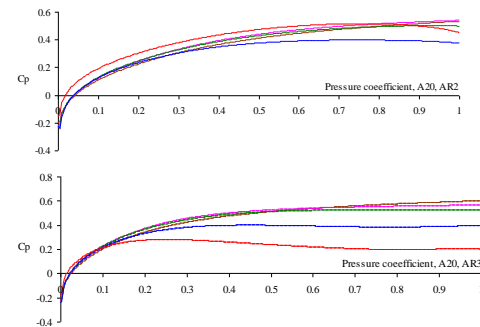


Fig. 8. Pressure recovery Coefficient predicted by RNG $k-\epsilon$ model

The pressure recovery coefficient (C_p) increases with the increasing diffuser passage length for flow without swirl. C_p increases with the increasing swirl as compared to the non swirling flow in the beginning of diffuser passage length. However for 7.5° , 12° , 17° and 25° inlet swirl C_p is lower than the flow without swirl beyond $x/L = 0.82$, 0.95 , 0.78 and 0.47 respectively in case of Area ratio 2. C_p is highest up to diffuser passage length of 0.36 for 25° inlet swirl. From $x/L = 0.36$ to 0.56 it is maximum for 17° inlet swirl, then from 0.56 to 0.95, it is for 12° inlet swirl and in the last for flow without swirl.

In diffuser with Area ratio 3, the C_p increases as compared to no swirl condition with increasing swirl in the beginning of diffuser passage length and the marginal increase is more than diffuser with Area ratio 2. However for 7.5° , 12° , 17° and 25° inlet swirl C_p is lower than the flow without swirl beyond $x/L = 0.67$, 0.97 , 0.60 and 0.25 respectively. C_p is highest up to diffuser passage length of 0.18 for 25° inlet swirl. From $x/L = 0.18$ to 0.28 it is maximum for 17° inlet swirl, then from 0.28 to 0.97, it is for 12° inlet swirl and in the last for flow without swirl. The reason for above phenomenon is that the swirl decays as we move downstream of diffuser passage.

V. CONCLUSIONS

Validated CFD RNG $k-\epsilon$ model has been employed to predict axial annular diffuser performance. Following inferences have been drawn from the predicted computational results for area ratios 2 and 3 for various inlet swirl angles.

1. The longitudinal velocity decreases continuously as the flow proceeds downstream, for both swirling and non swirling inlet flow
2. Due to the development of boundary layer shapes of velocity profiles are dissimilar at different locations of the flow passage
3. At any diffuser transverse the maxima of velocity profile is not at the middle rather it is towards the hub for non swirling flow. It shifts towards the casing side with the increase of swirl.
4. With the introduction of swirl, the flow is forced towards casing wall which strengthens the flow towards casing wall than the hub wall.
5. Pressure recovery coefficient increases with the diffuser passage. However the marginal recovery decreases with the diffuser passage.
6. With the introduction of swirl the recovery is sooner towards the casing wall. The effect of swirl emerges to decay gradually as the flow proceeds downstream of the flow. Further the recovery is negligible or nil towards the diffuser exit.
7. Stall is not observed on the hub wall for non swirling flow, however it is observed on casing wall for area ratio 3
8. CFD analysis in the pre study is reasonably in good agreement with the experimental data.
9. The RNG $k-\epsilon$ turbulence model used in the present work can be used to predict the flow behaviour in advance and the pressure recovery coefficient and diffuser effectiveness can be computed for a particular type of diffuser.

REFERENCES

- [1] Howard, J.H.G., Thorriton Trurnp A.B. and Henseler, H.J., "Performance and Flow Regimes for Annular Diffusers", ASME Paper No. 67 WA/FE 21, 1967.
- [2] Sovran Ed. G. Diffusers with Rectangular, Conical or Annular Cross Section", Fluid Mechanics of Internal Flow, Elsevier Amsterdam, pp.270-319, 1967.
- [3] Srinath.T., "An investigation of the effects of swirl on flow regimes and performances of annular diffuser with inner and outer cone angles." M.A.Sc. thesis, University of Waterloo, Canada 1968.
- [4] Stevens, S.J and Markland, E., "The Effect of Inlet Conditions on the Performance of Two Annular Diffusers", ASME Paper No.68 WA/FE 38.
- [5] Stevens, S.J., "The Performance of Annular Diffusers". Proc. Instn. Mech. Engrs., Vol. 182, Part 3D, pp.58-70, 1968.
- [6] Hoadley, D., "Three Dimensional Turbulent Boundary Layers in an Annular Diffuser", Ph.D. Thesis, Department of Engineering, University of Cambridge, London, 1970.
- [7] Coladipietro, R., Schneider, J.H and Shridhar, K., "Effects of Inlet Flow Conditions on the Performance of Equi Angular Annular Diffusers", CSME Paper No. 73 84, 1974.
- [8] Sovran, G and Klomp, E.D., "Experimentally Determined Optimum Geometries for Rectilinear
- [9] Shaalan, M.R.A and Shabaka, I.M.M., "An Experimental Investigation of the Swirling Flow Performance of an Annular Diffuser at Low Speed", ASME Paper No. 75 WA/FE 17, 1975
- [10] Kumar D.S., "Effect of swirl on flow through annular diffusers", Ph.D. Thesis, I.I.T. Delhi, 1977.
- [11] Lohmann, R.P., Markowski, S.J and Prookman, E.T., "Swirling Flow Through Annular Diffusers with Conical Walls", Journal of Fluids Engineering, Trans. ASME, Vol.101, pp.224-229, 1979.
- [12] Chithambaran, V.K., Aswatha, Narayana P.A., Chandrashekhra Swamy, N.V. "Mean velocity characteristics of plane diffuser flows with inlet velocity distortion." Journal of Indian Institute of Science, 65(A):79-93, 1984.
- [13] Sapre, R.N., Singh, S.N., Agrawal, D.P and Malhotra, R.C., "Flow Through Equiangular Wide Angle Annular Diffusers", 15th NCFMFP, Srinagar, July, 1987.
- [14] Agrawal, D.P., Singh, S.N., Sapre, R.N and Malhotra, R.C., "Effect of Hub Rotation on the Mean Flow of Wide Angle Annular Diffusers", HydroTurbo 1989, Czechoslovakia, 1989.
- [15] Choudhury D., "Introduction to the Renormalization Group Method and Turbulence Modeling." Fluent Inc. Technical Memorandum TM-107, 1993.
- [16] Singh S.N., Agarwal D.P., Sapre R.N. and Malhotra R.C., "Effect of inlet swirl on the performance of wide angled diffusers", Indian journal of Engineering & Materials Sciences, Vol.1, pp 63-69, 1994.
- [17] Buice, C. U., and Eaton, J. K., "Experimental Investigation of Flow Through an Asymmetric Plane Diffuser", Report No. TSD-107, Thermosciences Division, Department of Mechanical Engineering, Stanford University, Stanford, CA, USA, 1997.
- [18] Klomp, E.D., "Performance of Straight-Walled Annular diffuser with Swirling Flow", The Aeronautical Journal, Vol. 101, No. 1010, pp. 467-480, 1997.
- [19] Mohan, R., Singh, S.N., and Agrawal, D.P., "Optimum Inlet Swirl for Annular Diffuser Performance Using CFD", Indian Journal of Engineering and Materials Sciences, Vol. 5, pp. 15-21, 1998.
- [20] Japikse, D., "Correlation of Annular Diffuser performance with Geometry, Swirl, and Blockage", Proceedings of the 11th Thermal and Fluids Analysis Workshop (TFAWS), Cleveland, Ohio, August 21-25, 2000, pp. 107-118.
- [21] Kochevsky, A. N. "Numerical Investigation of Swirling Flow in Annular Diffusers With a Rotating Hub Installed at the Exit of

- Hydraulic Machines”, *Journal of Fluids Engineering*, Trans. ASME, Vol. 123, pp.484-489, 2001.
- [22] Yeung, W. W. H. and Parkinson, G. V., “Analysis and Modeling of Pressure Recovery for Separated Reattaching Flows,” *ASME Journal of Fluids Engineering*, Vol. 126, No. 3, pp. 355–361, 2004.
- [23] Singh, S. N., Seshadri, V., Saha, K., Vempati, K. K., and Bharani, S., “Effect of Inlet Swirl on the Performance of Annular Diffusers Having the Same Equivalent Cone Angle”, *Proceedings of the Institution of Mechanical Engineers, Part G, Journal of Aerospace Engineering*, Vol. 220, pp. 129-143, 2006.
- [24] Arora, B. B., Pathak, B. D. “Effect of Geometry on the Performance of Annular Diffuser,” *International Journal of Applied Engineering Research*, Vol.5, No.20. (2009) pp. 2639-2652, ISSN 0973-4562.
- [25] Arora, B. B., Manoj Kumar, Subhashis Maji, “ Study of Inlet conditions on Diffuser Performance,” *International Journal of Theoretical and applied Mechanics*, Vol.5, No.2. (2010) pp. 201-221, ISSN 09736085.
- [26] Manoj Kumar, Arora, B. B., Subhashis Maji, “ Analysis of flow separation in wide angle annular diffusers,” *International Journal of Applied Engineering Research*, Vol.5, No.20. (2010) pp. 3419-3428, ISSN 0973-4562
- [27] Arora, B. B., Pathak, B. D. “CFD analysis of axial annular diffuser with both hub and casing diverging at unequal angles,” *International Journal of dynamics of fluid* Vol.7, No.1. (2011) pp. 109-121, ISSN 0973-1784.
- [28] Manoj Kumar, Arora, B. B., Maji. Subhashis and Maji, S. “Effect of inlet swirl on the flow behavior inside annular diffuser,” *International Journal of dynamics of fluid* Volume 7, Number 2 (2011), pp. 181-188. ISSN 0973-1784
- [29] Arora, B. B. 2014, “Performance analysis of parallel hub diverging casing axial annular diffuser with 20° equivalent cone angle”, *Australian Journal of Mechanical Engineering*, Vol. 12, No. 2, June, pp. 179-194, <http://dx.doi.org/10.7158/M11-823.2014.12.2>.

Communication and Cyber Security issues in Smart Grid

M.M. Tripathi

Department of Electrical Engineering, Delhi Technological University, New Delhi, India

Abstract— Smart Grid is an Information and Communication Technology (ICT) enabled Power grid. It is efficient, secure, reliable and self-healing power grid. Integration of micro grids, electric vehicles and other utilities make it more interesting. The deregulation of electricity sector has necessitated the use of many advanced software and embedded technologies to handle the size and complexity of power network. Smart grid needs to be supported by efficient and secure communication architecture design and implementation. At the same time it is necessary to ensure the security and privacy of data and information moving or stored in the smart grid system to have near 100% uptime of the power grid. This paper presents a comprehensive analysis of the various communication and cyber security issues involved with the successful operation of Smart Grid.

Keywords— Smart grid, Cyber Security, Communication, Generation, Transmission, Distribution, Advance metering, restructuring.

I. INTRODUCTION

Smart Grid often called Intelligent Grid is a digitized electricity network which treats electricity unconventionally; not as a commodity but as a value added service. Its important features are deregulation, distributed generation, enhanced participation of consumers, generation and storage options, power quality, optimized asset utilization with high operational efficiency, self-healing and resiliency against attack and natural disaster. Integration of micro grids, electric vehicles and other utilities make it inspiring. The restructuring and deregulation of electricity market has necessitated the movement and analysis of large chunk of data in power network. The Smart grid is a new age transmission grid which is reliable, secure, affordable and efficient with improved power quality and low operational and maintenance cost. With the Advanced metering infrastructure, supported by bidirectional data flow, Demand side management with load and price forecasting and online bidding Smart Grid becomes most crucial to today's power system. The elements of the Smart Grid are (i) integrated communications (ii) advanced sensing and instrumentation technologies (iii) smart components (iv) advanced control strategies (v) improved interfaces.

Some of important features of Smart grid are:

- Distributed generation/ Micro-grids
- Enhanced customer participation
- Generation and storage options
- Better quality of power
- Bi-directional data flow
- Optimized utilization of Infrastructure with high operational system efficiency
- Self-healing and resiliency against cyber-attack and natural disaster
- Restructuring of Power system

A Smart grid is consisting of Power plus Energy Layers supported by the Communication Layer. Smart grid has better energy management, efficient and reliable network that integrates all energy supply and demand sources as well as smart devices. It also emphasizes the need of reliable and secure communications which pose biggest challenge of designing and deployment of appropriate communication architecture for smart grid which is interoperable, scalable, self-organizing, and secure against cyber-attacks.

The large scale introduction of distributed generation and non-conventional energy sources supports the institution of regional micro-grids. Making use of techniques like virtual power plants or virtual power systems they try for aggregating and autonomously deciding their own generation and demand side resources to balance production and consumption in as compact as feasible entities. Interaction with near-by distribution network and with the connected production, storage and consumption appliances would be the key to the economical and efficient management of such grids. Economical energy management in buildings needs in depth use of communication network infrastructure to and in buildings moreover because the provision of the mandatory interfaces to existing appliances, existing distributed generation and energy and service providers. The massive scale introduction of electrical vehicles would have an effect on the energy infrastructure. Providing sufficient charging points needs interaction between the energy infrastructures, the transport infrastructure, the vehicle data and information systems and also the communication network infrastructure so as to gather, process and deliver the required data.

The major software component in smart grid are the databases that collects demand orders from the customers

and keeps the record of the consumption and the decision making tool that determines generation and distribution scheme of the network. Artificial Intelligence tools may be utilized for forecasting of demand, price and other electricity parameters for better decision making, planning and operation. Various optimization techniques will be used for formulation of best optimized bids. Web technologies would be used for creating online systems for bidding, monitoring and management of smart grid, electricity market and consumer services. Smart meters provide communication (TCP/IP) between customer and supplier. Many smart devices using embedded technologies would become part of smart grid [1]. In the new system utility firms once consolidating the customer's orders will decide the optimum generation and distribution scheme to satisfy customer's demands. A virtual energy provisioning system might improve the stability and reliability of the electrical grid that is progressively integrated with distributed energy resources. The new system could permit customers to directly order the electricity via the web in terms of power needs and time of usage.

Smart grid provides numerous opportunities for generators, market operators and customers to work and manage the sophisticated energy infrastructure. Customers will get cash discounts, marketers would be able to create a reliable dispatch scheme; and generators would maximize their asset utilization and profit. Customer would get opportunity to select energy supplier as and when he/she desires via web [2]. Power would be transmitted on the same distribution line for all service suppliers and calculation of power flow and revenue share are going to be determined based on software system applications. Massive amount of information from generators, distributors, customers and retailers, would need it to be stored and moved safely. Thus security and as privacy of information/ data is of utmost importance.

II. SMART GRID WORLD WIDE AND INDIA

Smart grid has been implemented in many countries worldwide and their status in different countries are shown below in table 1 [3], [4], [5], [6] and [7]. In 2003 the Electricity Act 2003 was passed in India. After this the traditional Indian Power System is transformed to restructured Power System paving the way for the automation of Power Grid. Indian power system is operated as five Regional grids viz. Southern Regional grid (SR), Western Regional grid (WR), Northern Regional grid (NR), Eastern Regional grid (ER), and North Eastern Regional grid (NER). In addition to this India have 31 state grids and more than 100 area-wise grids. To control the grid there is a hierarchical system as shown in fig. 1.

Market operation is coordinated by Market operators along with RLDCs [1]. Use of information, electronics &

communication technology (ICT) paved the way for conceptualization of "Indian Smart Grid" by integration of large scale generation, distributed energy resources, transmission and distribution system to improve the operation and control of power grid [8]. Smart grid implementation in India includes followings [9]:

- Advance metering infrastructure (AMI)
- Smart meters and sensors
- Metered data transmission network system
- Metered data management
- SCADA applications and data transfer network.
- Intelligent electronic devices (IEDs) applications
- Peak demand reduction in distribution systems
- Voltage regulation system
- Demand response from Industry and home users

TABLE 1 SUMMARIZED R STATUS OF SMART GRID IMPLEMENTATION IN DIFFERENT COUNTRIES

Country	Smart grid status
Italy	Italy was the first country to adopt the smart grid technology. Most of the electricity customers are provided with smart metering system.
Denmark	Denmark has adopted smart metering by most of the electricity distribution system operators.
Finland	Finland has 100 percent installation of hourly metering and settlement system in the country.
France	France has already rolled out its nationwide implementation of smart meters.
Spain	In line with new regulations, Spain had planned to deploy 10 million smart meters in the country.
Ireland	A nationwide program has started for the deployment of smart meters from 2013.
The Netherland	A two-year trial period to roll out the smart metering was initiated as per the legislation adopted by the Dutch parliament in July 2008. The Country has already installed over 100000 residential smart meters.
UK	The roll out of dual-fuel smart metering for 27 million households started in October 2008 which is likely to be completed by 2021.
Germany	The German power utility announced in February 2008 that it will install approx. one million smart electricity meters.
China	The China has announced to develop a national smart grid by the year 2020.

Indian power sector is one of the biggest in the world, with a capacity of 156,092.23 MW. It has 4th place in installed

capacity and 6th place in energy uses, globally. Energy demand for India has reached to 157,107 MW in the year 2012. India's rising population and growing economy need a rise in energy demand from 800,000 MW to 950,000 MW by the year 2030. India's huge energy supply-demand gap and its high losses have sought attention of the Government and other stakeholders in the smart grid. Indian stakeholders are required to take active interest in design and implementation of smart grid; they must acquire the required skills and invest in development of human resources in this area. At last but not least, Indian smart grid also requires standards and specifications for the smart grid. The priority areas should be:

- Advanced metering infrastructure
- Communication network Infrastructure
- Efficient electricity transportation system
- Energy storage systems
- Demand response
- Energy efficiency
- Distributed and hierarchical management of smart grid
- Setting up of cyber security policies, infrastructure and management system

electricity very efficiently. In smart grid several electrical devices, sensors, smart meters etc. would be connected to one another in a complex way, reporting back their status of power consumption, health condition and other signals via appropriate communication protocols. There would be many challenges in the design of the smart grid communications network where electrical devices and smart meters would be exchanging information related to continuously changing power requirements [10]. Various communication media which may be used in smart grids are:

- Optical Fiber network
- Wireless Radio network
- Zig-Bee network
- Cellular network (2G, 3G, 4G, WiMAX)
- Satellite network (GSM/ GPRS)
- Power line Communication (PLC)
- Advanced Digital Subscriber Lines (ADSLs)

The smart grid communication systems can be abstracted to four layers such as field layer, network layer, transport layer, and application layer. Field layer includes remote terminal unit, intelligent electronic device, information collection terminal, smart meters, and smart devices. Network layer would be consisting of wired or wireless network using low range media and switching devices used in LAN. Transport layer would include the wired or wireless network of long range communication such as optical fiber, power line communication, and mobile network and routing devices used in wide area network. Control and information system of master station of smart grid would be the application layer [11].

In smart grid communication network the transmission and distribution substations and control centers can be connected to one another in a meshed network, which can be built over optical fiber technology. The communication architecture for the lower distribution network may be divided into a number of networks arranged in hierarchical network architecture. Smart meters can be deployed in such architecture which is part of advanced metering infrastructure (AMI) for enabling automated two-way communication between various layers of network. These Smart meters would be equipped with power reading capability and would also work as communication gateway using Internet Protocol (IP) based communication [10].

A huge amount of raw data is collected by smart meters, sensors, phasor measurement units (PMUs) from different part of the smart grid and being send continuously to the central computation system. Subsequently, this considerably big amount of unstructured data must be processed, analyzed and stored in a cost effective ways. An enormous amount of computing, storage and software services must be provided to compute this vast amount of data and produce some valuable information in least possible time. Many technologies such as distributed and parallel computing,

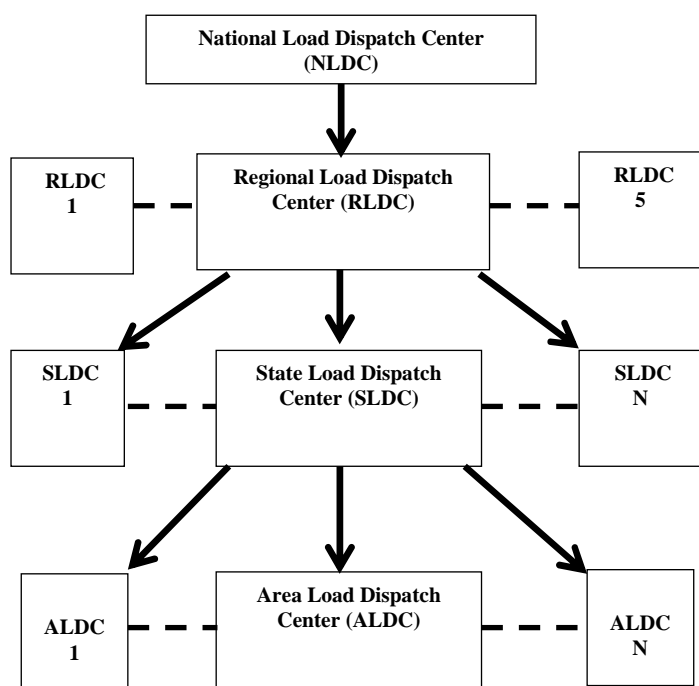


Fig.1: Hierarchical Grid Control System in India

III. COMMUNICATION ISSUES IN SMART GRID

Smart grid would allow the power producers, distributors and consumers to have real-time awareness of operating requirements and capabilities of one another hence smart grid would be able to generate, distribute, and use the

Grid computing or cloud computing may be utilized for this purpose. Cloud computing and big data analysis looks very promising for the future of smart grid [12]. This allows distributed power generation such as local solar and wind generation and also promises higher efficiency in the power distribution system. Scalability and fast communication is crucial for practical deployment of smart grids into home and building automation. Communication in smart grid must be private and secure as many autonomic functions that would run over it might be critical to grid operations. It is very important to design and deploy the future communication architecture for smart grid which is interoperable, scalable, self-organizing, and secure against cyber-attacks [10], [13], [14], [15] and [16].

IV. CYBER SECURITY ISSUES IN SMART GRID

Increasing interconnection of smart grids additionally will increase the exposure of grid to potential attackers and/or unintentional errors. Networks that connect more often to totally different networks introduce most common vulnerabilities which will span multiple smart Grids and increase the possibilities of failures. Different variety of interconnections invites additional denial of service attacks, malicious code, compromised system and intrusions. Because the variety of nodes will increase within the network, the amount of entry points increase which can be used by potential adversaries for exploitation. Extensive information & data collection their flows might broaden the potential for compromises of information and data and breach of confidentiality and client privacy.

Components of good grid security includes physical elements and management applications, cyber infrastructures needed to support necessary designing, operational and market functions, cyber-attacks and its impact on the system, actions and measures to mitigate risks from cyber threats. Risk assessment to understand the potential of unwanted outcome ensuing from internal or external factors, as determined from the probability of occurrences and also the associated consequences is important to attenuate unacceptable risk levels by applying risk mitigation solutions. This might be performed through the preparation of a lot of sturdy supporting cyber security infrastructure or applications.

A. Cyber Security issues in Generation

Power generation management primarily involve managing the generator power output and terminal voltage by applying local automatic transformer (AVR) and governor control (GC) schemes. AVR and governor control don't rely on the supervisory control and data acquisition (SCADA) transmission infrastructure for its operations as each the terminal voltage and rotor speed are detected locally. Albeit these applications are susceptible to malware that would enter the station local area network (LAN) through different

entry points like USB keys. Additionally associates working in company may compromise plant cyber security mechanisms to gain an entry point into the native local network and may disrupt normal operation by corrupting the logic or settings within the digital controllers. The automatic generation control (AGC) is a secondary frequency management that's involved with fine controlling the system frequency to its nominal value [17] and [18]. The AGC depends on tie-line and frequency measurements provided by the SCADA mensuration system. An attack on AGC may have direct impacts on system frequency, stability and operation. Denial of service (DoS) style of attacks might create a big impact on AGC operation once supplemented with another attack that needs AGC operation.

B. Cyber Security issues in Transmission

Power system state estimation [19] is used as a technique by that estimates of system variables like voltage magnitude, phase angle (state variables) and power flows on totally different sections are calculated based on likely faulty measurements from field devices. The control center application performs computations by making use of large no. of measurements it receives via wide-area network. False information injection attacks, that escape detection by existing faulty measurement identification algorithms, could also be dangerous.

The reactive (VAR) compensation with FACTS devices [17] is the method of controlling reactive power injection or absorption in a power grid to boost the performance of the transmission network. FACTS devices interact with each other to exchange operational data via communication link. It can face Denial of cooperative operation that may be a DoS attack. During such attack, the communication to some or all the FACTS devices might be stopped by flooding the network with unwanted packets. This may end in the loss of vital data and therefore have an effect on semi-permanent and dynamic management capabilities. De-synchronization (timing-based attacks) might disrupt steady operation of Cooperative FACTS devices (CFDs).

False data injection attacks [20] may well be utilized to send incorrect operational information like status and control data. This might lead to uncalled-for reactive power compensation and in unstable operational conditions. The phase angles of voltage phasors measured by PMUs directly facilitate the computation of real power flows within the network, and will therefore assist in decision making at the central control unit. Any attack on WAN backbone is going to be a danger to PMU.

C. Cyber security and distribution system

Modern relays are internet protocol (IP) capable and support various communication protocols [21]. Cyber-attack on the communication set up or malicious amendment to the control logic might end in unexpected tripping of distribution feeders, resulting in load segments not served.

Advance metering Infrastructure primarily depends on the deployment of good smart meters to supply real-time meter readings. Smart meters offer utilities with the flexibility to implement load management & control (LMC) to disable control devices once demand spikes. The capability to remotely disable smart meters through load management shift provides potential threats from attackers.

Embedded systems are used heavily in the grid to support observance and management functions. Intelligent electronic devices (IEDs) are placed to control relays throughout the grid [22]. Recent events have shown that IEDs are often maliciously reprogrammed to halt intended management functions. Deployments of embedded devices at large scale in smart grid additionally incentivize the employment of cheaper hardware leaving very little computing capability to support varied security functions like malware or intrusion observance. The deployment of secure computation within embedded platforms provides a key challenge to cyber security.

D. Cyber security in communication network of Smart grid

Cyber security [17] inside the communication network is known as protection of data and systems from unauthorized access, disclosure, modification, destruction or disruption. The objectives of Cyber security are confidentiality, integrity and availability. These 3 objectives need to be ensured altogether in various stages of data processing (i) storage state in storage media (ii) processing state in RAM and (iii) transmission state in communication media.

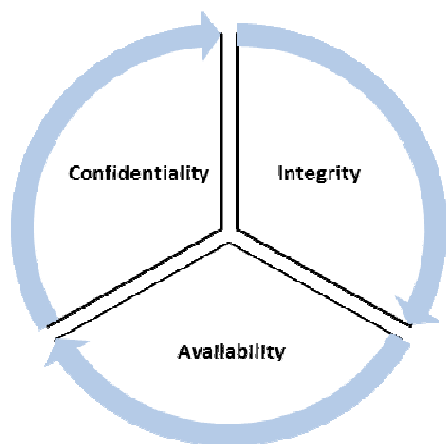


Fig. 2: Three Objectives of Cyber-security in Smart Grid

Confidentiality is outlined in literature as protection of data from unauthorized access or revelation. The authorized users only ought to get access to data and un-authorized ones ought to be prevented from doing this. Integrity is defined as protection of data from unauthorized modification or destruction. It's necessary to make sure that the data and information and system containing information is correct, non-corrupted and complete. Availability refers to the protection of data and information systems from unauthorized disruption. It is important to make sure the timely and reliable access to and use of data and

information systems. The graphical representation of the 3 objectives of the cyber-security in smart grid is shown in fig. 2 above.

Other cyber-security concerns are authentication, identification authorization, access control, non-repudiation and privacy of data and user's information. Other than this auditability and reliability of data and information is also very important. Some of the important cyber-security features are shown graphically below in fig. 3. It is must to implement the features listed in fig. 3 in the smart grid for secure and reliable operation of the grid.

There are several threats to the ICT infrastructure of smart grid and it regularly faces many types of attacks. Some of the Security threats and attacks to ICT infrastructure of Smart grid is shown in fig. 4 and fig. 5 respectively. Other than the tools and techniques mentioned in table 2, many management, audit, measurement, monitoring, and detection tools are used for better management of cyber-security in the communication network of smart grid as shown in the fig. 6 below. It is difficult to respond these attacks as the ICT system is designed for trust and without authentication. Also it is evolving as it is being employed for uses beyond design. Some cyber security features and tools/ techniques for their implementation as listed below in table 2 can be practiced based on the risk assessment of the system [23].

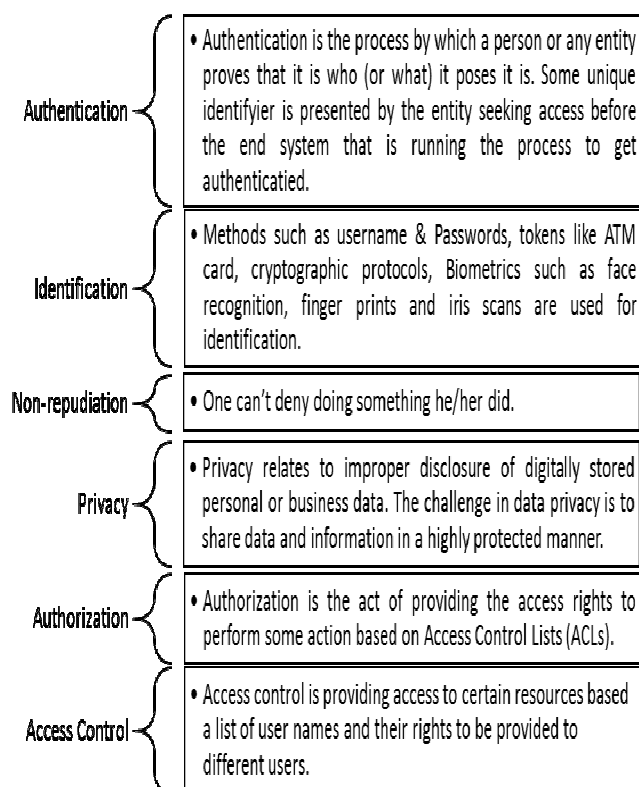


Fig. 3: Features of secure Smart grid

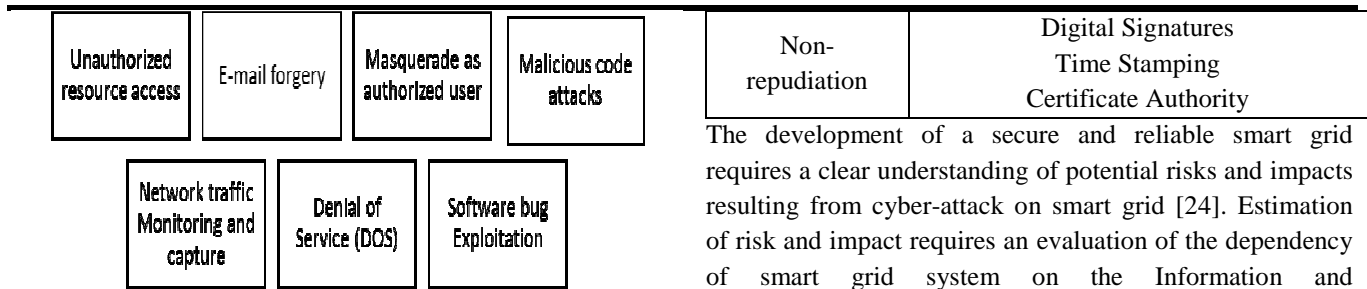


Fig. 4: Security threats to ICT Infrastructure

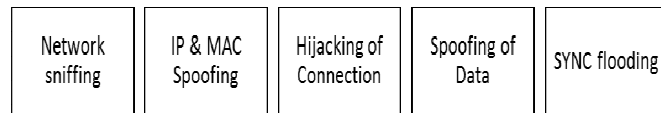


Fig. 5: Cyber-attacks on ICT infrastructure

TABLE II CYBER SECURITY FEATURES AND METHOD OF IMPLEMENTATION IN SMART GRID

Cyber-security feature	Cyber-security techniques/ tools
Confidentiality	Encryption Data Separation Symmetric Key Encryption Public Key Encryption Virtual Private Networks (VPNs)
Integrity	Digital Signatures Message Integrity Safeguards Time Stamping
Availability	Protection from attack Protection from unauthorized users Resistance to routine failures
Identification	User ID and passwords PIN number Biometric
Authentication	Secure Tokens Smart Cards Single Sign-on Password Authentication Challenge/Response Authentication Physical/Token Authentication Smart Card Authentication Biometric Authentication Location-Based Authentication Device-to-Device Authentication
Authorization	Certificates Attribute use Role-Based Authorization Tools
Access Control	Role-based Access Control Passwords Network Firewalls Host-based Firewalls Virtual Networks Physical Protection Personnel Security

The development of a secure and reliable smart grid requires a clear understanding of potential risks and impacts resulting from cyber-attack on smart grid [24]. Estimation of risk and impact requires an evaluation of the dependency of smart grid system on the Information and Communication Technology (ICT) infrastructure and its ability to continue the operations in presence of potential cyber-attacks [25]. Regular evaluation of the cyber-physical relationships of smart grid with its ICT infrastructure and specific review of possible attack vectors is important to determine the effectiveness and reliability of the cyber security system in place.

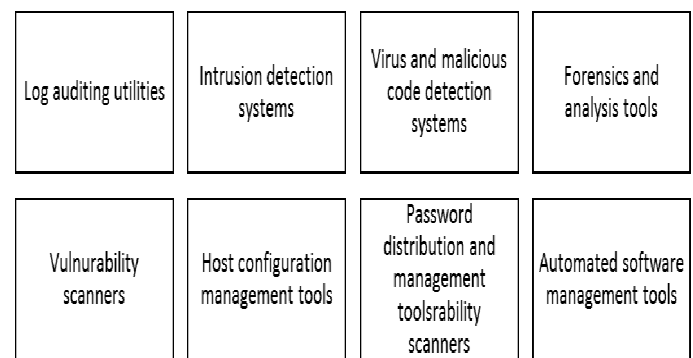


Fig. 6: Cyber-security management tools in Smart grid

V. CONCLUSION

Many countries in the world have moved towards making their power grid smarter and others are in the process. With increasing integration of many regional grids across the globe would make it possible to form power cloud where power can be drawn from the cloud in any amount at any time and place. This would necessitate storage and movement of large chunk of data, high speed network connectivity of various systems in power grid, use of smart devices and high speed computing techniques to handle the complexity and size of the smart grid.

Many new scalable communication architecture, protocols and software need to be designed and developed to handle the real time requirements of operation and control of smart grid. It would also require that policy, procedures and technologies are identified and implemented to operate the smart grid in a secure and reliable manner.

Estimation of potential risk and its impact assessment on the business continuity of smart grid system would decide the appropriate tools and technologies to be put in place for mitigation of potential cyber-attacks on the smart grid. By making correct choices it is possible to run the future smart grid in an efficient, secure and reliable manner as it is important for sustainable growth.

REFERENCES

- [1] Xiang Lu, Zhuo Lu, Wenye Wang and Jianfeng Ma," On Network Performance Evaluation toward the Smart Grid: A Case Study of DNP3 over TCP/IP", 978-1-4244-9268-8/11/\$26.00 ©2011 IEEE
- [2] Tongdan Jin, Mechehoul M., "Ordering Electricity via Internet and its Potentials for Smart Grid Systems, " Smart Grid, IEEE Transactions on, vol.1, no.3, pp.302-310, Dec. 2010
- [3] Jingjing Lu, Da Xie, and Qian Ai," Research on Smart Grid in China ",IEEE T&D Asia 2009
- [4] V S K Murthy Balijepalli, S A Khaparde, R P Gupta," Towards Indian Smart Grid", 978-1-4244-4547-9/09/\$26.00 ©2009 IEEE
- [5] Arup Sinha, S.Neogi , R.N.Lahiri, S.Chowdhury, S.P.Chowdhury , N.Chakraborty, "Smart Grid Initiative for Power Distribution Utility in India ",978-1-4577-1002-5/11/\$26.00 ©2011 IEEE
- [6] Mohsen Fadaee Nejad, Amin Mohammad Saberian, Hashim Hizam, Mohd Amran Mohd Radzi, Mohd Zainal Abidin Ab Kadir , "Application of Smart Power Grid in Developing Countries", 978-1-4673-5074-7/13/\$31.00 ©2013 IEEE
- [7] The Smart Grid Vision for India's Power Sector: A White Paper
- [8] Yong Wang, Da Raun, Dawn Gu, Jason Gao, Daming Liu, Jianping Xu and Fang Chen," Analysis Of Smart Grid Security Standards", 978-1-4244-8728/11/\$26.00©2011 IEEE
- [9] Arup Sinha, S.Neogi, R.N.Lahiri, S.Chowdhury, S.P.Chowdhury, N. Chakraborty, "Smart Grid Initiative for Power Distribution Utility in India ",978-1-4577-1002-5/11/\$26.00 ©2011 IEEE
- [10] Zubair Md. Fadlullah, Mostafa M. Fouda, and Nei Kato, Akira Takeuchi, Noboru Iwasaki, and Yousuke Nozaki,"Toward Intelligent Machine-to-Machine Communications in Smart Grid" IEEE Communications Magazine, April 2011, 0163-6804/11/\$25.00 © 2011 IEEE
- [11] International Telecommunication Union, "ITU Internet Reports 2005: The Internet of Things", 2005
- [12] Berthold Bitzer, Enyew Sileshi Gebretsadik, "Cloud Computing Framework for Smart Grid Applications", 978-1-4799-3254-2/13/\$31.00 ©IEEE
- [13] Xudong Wang and Ping Yi,"Security Framework for Wireless Communications in Smart Distribution Grid", IEEE TRANSACTIONS ON SMART GRID, VOL. 2, NO. 4, DECEMBER 2011
- [14] Vehbi C. Güngör, Dilan Sahin, Taskin Kocak, Salih Ergüt, Concettina Buccella, Carlo Cecati and Gerhard P. Hancke," Smart Grid Technologies: Communication Technologies and Standards", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 7, NO. 4, NOVEMBER 2011 529
- [15] Fariba Aalamifar, Hossam S. Hassanein and Glen Takahara," Viability of Powerline Communication for the Smart Grid," 978-1-4673-1114-4/12/\$31.00 ©2012 IEEE
- [16] Yi Xu and Wenye Wang "Wireless Mesh Network in Smart Grid: Modeling and Analysis for Time Critical Communications", IEEE Transaction on Wireless Communications, Vol. 12, No. 7, July
- [17] Siddharth Sridhar, Adam Hahn and Manimaran Govindarasu, "Cyber-Physical System Security for the Electric Power Grid" Proceedings of the IEEE | Vol. 100, No. 1, January 2012
- [18] Siddharth Sridhar, and Manimaran Govindarasu," Model-Based Attack Detection and Mitigation for Automatic Generation Control" IEEE Transaction on Smart Grid, Vol. 5, No. 2, March 2014
- [19] Saman Zonouz, Katherine M. Rogers, Robin Berthier, Rakesh B. Bobba, William H. Sanders, Thomas J. Overbye" SCPSE: Security-Oriented Cyber-Physical State Estimation For Power Grid Critical Infrastructure" IEEE Transactions on Smart Grid
- [20] Lanchao Liu, Mohammad Esmalifalak, Qifeng Ding, Valentine A. Emesih, and Zhu Han, "Detecting False Data Injection Attacks on Power Grid by Sparse Optimization" IEEE Transaction on Smart Grid, Vol. 5, No. 2, March 2014
- [21] Higgins, N.; Vyatkin, V.; Nair, N.-K.C.; Schwarz, K.; "Distributed Power System Automation with IEC 61850, IEC 61499, and Intelligent Control," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on , vol.41, no.1, pp.81-92, Jan. 2011
- [22] A. M. Gaouda, Ahmed Abd-Rabou and Abdul Rahman Dahir, "Developing Educational Smart Grid Laboratory", 978-1-4673-6355-6/13/\$31.00©2013 IEEE
- [23] Vehbi C. Güngör, Dilan Sahin, Taskin Kocak, Salih Ergüt, Concettina Buccella, Carlo Cecati and Gerhard P. Hancke, , " Smart Grid Technologies: Communication Technologies and Standards", IEEE Transactions on Industrial Informatics, Vol. 7, No. 4, November 2011, pp 529-539
- [24] Tosin Daniel Oyetoyan, Reidar Conradi and Kjell Sand," Initial Survey of Smart Grid Activities in the Norwegian Energy Sector – Use Cases, Industrial Challenges and Implications for Research", 978-1-4673-1864-8/12/\$31.00c 2012 IEEE
- [25] Boban Panajotovic, Milan Jankovic, Borislav Odadzic," ICT and Smart Grid", 978-1-4577-2019-2/11/\$26.00 ©2011 IEEE.

Contingent Demand Fulfilment Decision Problem from Supplier's Perspective

Akhilesh Kumar

Department of Applied Mathematics
Delhi Technological University
Delhi, India
akhilesh.maths@gmail.com

Anjana Gupta

Department of Applied Mathematics
Delhi Technological University
Delhi, India
anjanagupta@dce.ac.in

Aparna Mehra

Department of Mathematics
Indian Institute of Technology, Delhi
Delhi, India
apmehra@iitd.ac.in

Abstract— The planning of production and other operations in a firm is done based on demand forecasts. But production decisions get affected by actual realization of the demand. These decisions are taken keeping into consideration production capacities. Sometimes sudden and unforeseen demands from customers arise. This kind of situation generally arises when customers of a supplier firm get some new contracts or there is a short term but a high fluctuation in their demands. When this sudden or contingent demand arises on customers' side, they create a demand for corresponding spare parts to their suppliers. In this situation, the vendor's decision to satisfy this demand fully or partially is restricted by its own available resources, as processes like production capacity expansion cannot be carried out in such a short response time. The supplier has to decide on the extent of fulfilment of demands of various customers. The decision making in this situation is difficult, as here the supplier has to consider not just the profit from the deal but also some other criteria of business sustainability like business relations and long sightedness in terms of future business prospects which are qualitative in nature.

In this paper, we consider the problem of a supplier firm which manufactures products that are used by its customer firms as spare parts to manufacture their specific products. We attempt to address the decision making on the extent of satisfaction of contingent demand of each customer. While deciding on the extent of demands of its customers, we try to address the concern of the supplier firm to concentrate not just on the profit, but also on the future business and relationship with its client firms. For this, we try to quantize these linguistically expressible aspects, so as to use them in our optimization problem.

Keywords: *Multiobjective optimization; contingent demand; linguistic; membership function; AHP (Analytical Hierarchy Process)*

I. INTRODUCTION

The planning of production and other operations in a firm is done based on demand forecasts. But production decisions get affected by actual realization of the demand. These decisions are taken keeping into consideration production capacities. Sometimes sudden and unforeseen demand can arise from customers. This kind of situation is fairly common for the supplier firm producing those products which are frequently purchased by other firms for further use in their own production processes. The unforeseen demands at customer end can be either due to some new contracts

received by these firms or a high fluctuation in their product demand for a short duration.

When this sudden or contingent demand arises on customers' side, they place a demand for corresponding spare parts to the vendor. The supplier's decision to satisfy this demand, fully or partially, is restricted by his own available resources since the long term strategic actions like production capacity expansion cannot be carried out in such a short response time. In this kind of situation, a supplier firm finds itself in a fix to decide on the demand fulfilment of multiple customers for various products under consideration. The reason behind this problem is that the firm needs to consider not just a single objective of obtaining maximum profit from the demand fulfilment but also some other important aspects like present business relations with its customers and future business prospects with them. As these aspects are qualitative in nature and cannot be measured quantitatively thereby making it difficult for the decision maker to compare these with the profit.

In this paper, we consider the problem of a supplier firm which manufactures products that are used by its customer firms as spare parts to manufacture their particular products. We address the decision making on the extent of satisfaction of contingent demand of each customer. For this decision making problem, on the extent of demands of its customers, we present to quantitatively measure the factors of future business and present relations with customers. A multi-objective optimization is proposed which is converted into a single objective optimization problem using the classical weighted sum approach and the weights (or priority) of various objective functions is obtained using experts knowledge and analytical hierarchy process. A survey is conducted with managers of a manufacturing firm with similar profile and a similar situation generally faced by them. The problem is solved for a sample data provided by the firm.

II. MOTIVATION AND RESEARCH GAP

Plethora of research has been done on the strategy building and development of decision support systems (DSS) of firms for the supplier selection and order allocation [10]. Some of recent developments in the field are [3], [11]. Looking through the other direction of the supply chain, there is again ample development of DSS for production firms demand fulfillment decisions. Some of well known developments are [1], [4], [7] and [8]. The literature suggests strategy building depending on the known or estimated demand. But, so far no research has

addressed the issue of unforeseen demand from customers, especially for a supplier having its customers as business firms.

The decision makers of any business organization, while taking decisions on any deal or processes of supply chain, consider many strategy oriented factors directing towards long term gains and sustainable business development. Experts from industry opinion that major factors kept into consideration during business to business (B2B) deals are generally business relations with the organization, future business with the organization and future of the product segment under negotiation with the organization. These factors are qualitative in nature, in fact some aspects involved in the assessment can be judged in a linguistic manner. So, no work has tried to quantify these and therefore these are not incorporated into quantitative DSS.

III. QUANTIFICATION OF QUALITATIVE FACTORS

Various components kept into consideration by strategic decision makers were identified through experts' opinion from industry for the above mentioned qualitative factors. We develop the corresponding quantitative measures in context of our problem.

A. Developing Quantitative Measure for Business Relation

For developing quantitative measure for the aspect of business relation with a customer firm, following factors are identified to differentiate a customer firm from the other.

- (1) *Fraction of the total demand of the customer which it allocates to the decision maker supplier firm.* This gives a measurable value of the proportion of the business for which the supplier is a partner of the customer.
- (2) *The profit percentage (proportion) earned from the customer firm out of all the customers taken together.*
- (3) *Satisfaction or ease of doing business with the customer.* This can be measures in terms of a function of number transactions for payments which are categorized as unsuccessful, satisfactory and excellent. Unsuccessful transactions are termed as those which have failed due to some error or miscommunication. Satisfactory are those which are successful but delayed from promised due dates. The excellent one are termed as those transactions which have been performed and that too within the promised time frame.
- (4) *Payment lead-time or grace period.* This is measured as the time span between the delivery of the placed order and the receipt of its payment by the customer. Sometimes this is considered as the time span promised by the buyer for the payment. The lead-time for the payment is generally negotiated by the customers in advance. The lesser the payment lead-time, the better is the customer considered to do business with.

Once quantized, let us denote these factors as L_f ($f = 1, 2, 3, 4$), respectively. Among the above four factors, first two are quantitative in such a manner that they are directly

proportional to the extent business relationship. So, we try to quantize the other two factors.

For quantizing the satisfaction or ease of doing business with the customer using the above mentioned three types of transactions, we propose to use a membership function defined as follows.

$$\mu_{low} = \frac{\text{unsuccessful transactions}}{\text{total transactions}} \quad (1)$$

$$\mu_{medium} = \frac{\text{satisfactory transactions}}{\text{total transactions}} \quad (2)$$

$$\mu_{high} = \frac{\text{excellent transactions}}{\text{total transactions}} \quad (3)$$

Then using the weights with these functions (all three weights being non-negative with their sum equals 1), which in practice can be subjectively defined by the decision maker, the satisfaction factor may be calculated as following.

$$L_3 = w_{unsuccessful} \mu_{low} + w_{satisfactory} \mu_{medium} + w_{high} \mu_{excellent} \quad (4)$$

Now for making the payment lead-time or grace period to reflect in the measure of business relations, one needs to consider the subjectivity in the consideration of these time spans as small, medium or high. For this we propose to use the three point membership function approach. As these lead-times are defined in terms of number of days so demonstrate these membership functions to be defined linearly as follows.

$$\mu_{small} = \begin{cases} 1 & \text{if } 0 \leq t_i < 10 \\ (30 - t_i)/20 & \text{if } 10 \leq t_i < 30, \\ 0 & \text{if } t_i \geq 30 \end{cases} \quad (5)$$

$$\mu_{medium} = \begin{cases} 0 & \text{if } 0 \leq t_i < 10 \\ (t_i - 10)/20 & \text{if } 10 \leq t_i < 30 \\ (60 - t_i)/30 & \text{if } 30 \leq t_i < 60 \\ 0 & \text{if } t_i \geq 60 \end{cases} \quad (6)$$

$$\mu_{high} = \begin{cases} 0 & \text{if } t_i \leq 30 \\ (t_i - 30)/30 & \text{if } 30 \leq t_i < 60 \\ 1 & \text{if } t_i \geq 60 \end{cases} \quad (7)$$

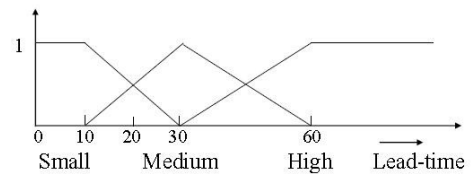


Fig. 1. Membership functions of lead-time

We then propose to use the weights of these membership functions to describe the preference or acceptance of each category of lead-times as w_{small} , w_{medium} and w_{high} . Where, w_{small} , w_{medium} , $w_{high} \in [0, 1]$ such that $w_{small} + w_{medium} + w_{high} = 1$.

$$L_4 = w_{small} \mu_{small} + w_{medium} \mu_{medium} + w_{high} \mu_{high} \quad (8)$$

To prefer shorter lead-times one should consider

$$w_{small} > w_{medium} > w_{high}.$$

After quantizing all the four factors as described above, we now propose to aggregate these as a single factor as follows.

$$L = \sum_{f=1}^4 wtl_f L_f, \quad (9)$$

where, wtl_f is the weight assigned to each quantized factor L_f for their preference or importance in the decision making process.

B. Developing Quantitative Measure for Future Business Prospects

This factor is considered as an initiative with the vision of promoting the business with those customers and for those products which are envisaged as key role players in future business. So, here we should quantize this factor corresponding to each product and each buyer. In our problem, as the buyer is a firm who manufactures some products using the products from the supplier as spare parts, so accordingly we identify the influencing aspects as following.

- (1) Number of years of business collaboration
(since when buyer purchasing that product from supplier)
- (2) Reputation of buyer as an enterprise in its market
 - (a) financial stability
 - (b) market value
 - (c) recent news/ results
- (3) Customer satisfaction
- (4) Friendly managerial policy
- (5) Market performance of product of the buyer which uses i^{th} product as spare part/ raw material

Once quantized, for a product k ($k = 1, 2, \dots, K$), let us denote these factors as E_{kf} ($f = 1, \dots, 5$), respectively. Out of the above listed factors only the first and last are product specific in addition to being buyer specific but the other three are buyer specific only.

For quantizing and scaling the years of business collaboration, we propose to define three point membership functions for partnership in its *initial*, *intermediate* and *high* stages for purchase of each product under consideration. These are illustrated in equations (20) – (22). Further, the weights of these membership functions to describe preference of the durations of collaboration be defined as $w_{initial}$, $w_{intermediate}$ and w_{high} . Where, $w_{initial}$, $w_{intermediate}$, $w_{high} \in [0, 1]$ such that $w_{initial} + w_{intermediate} + w_{high} = 1$. The corresponding component can then be expressed as for the product k ($k = 1, 2, \dots, K$).

$$E_{k(f=1)} = w_{initial} \mu_{k initial} + w_{intermediate} \mu_{k intermediate} + w_{high} \mu_{k high}. \quad (10)$$

The second factor of reputation of a business firm can only be described in linguistic terms and can be detailed by the market analyst of the supplier company as *financial stability*, *market value* and *recent news/ results*. All of these three elements can only be graded linguistically by a market expert. We incorporate the grading done by expert for each of these three as ratings of the buyer on a 4 point scale $\{0, 1, 2, 3\}$ as follows.

0 (poor) 1 (satisfactory) 2 (good) 3 (excellent)

For aggregating these three elements to represent the reputation factor as a number on the scale of $[0, 1]$, we first quantize each rating on the scale of $[0, 1]$ by dividing them by 3 and then take their arithmetic mean [2]. For example, if a buyer is rated as 1 for financial stability, 1 for market status and 0 for recent news/ results, then their scores are obtained as 0.33, 0.33 and 0. Then by obtaining the arithmetic mean of these scores we obtain the aggregated score for reputation of the buyer as 0.22.

The third factor considered as customer satisfaction is taken from the scorecards ratings given by each buyer to the supplier. These ratings are given in terms of percentages so we can convert these percentage scores on the scale of $[0, 1]$.

The fourth factor of friendly managerial policy of buyer is again a linguistic term. A buyer can be judged for his policy by the market analyst and given a grade on a 5 point scale as shown below. The friendlier the managerial policy of the buyer, the better the future business collaboration can be expected with the buyer. This grading can be quantized on the scale of $[0, 1]$ using Harrera's approach [5].

Feedback	Grading	Score
Extremely poor	1	0.2
Poor	2	0.4
Satisfactory	3	0.6
Good	4	0.8
Extremely Good	5	1

The fifth factor considered as the market performance of the product of the buyer, which uses i^{th} product as spare part/ raw material, is again such a factor that can only be judged and graded on a 5 point scale. So, this also can be quantized on similar line as described above.

After quantizing all these five factors, as described above, we propose to aggregate these into a single factor as follows.

$$E_k = \sum_{f=1}^5 wte_f E_{kf}, \quad (11)$$

where, wte_f is the weight assigned to each quantized factor E_{kf} for their preference or importance in the decision making process.

Once above described aspects of future business and present business relations are quantized, these can be incorporated as parameters in our decision making problem.

IV. PROBLEM DESCRIPTION

The proposed model considers the problem of a supplier, which is a firm manufacturing ' K ' types of products. This supplier firm is having a set of ' I ' customers who purchase these products for using these in their own production. The supplier firm receives a sudden additional demand for its products from its customers. The supplier has to decide on the proportion of the respective demands to be satisfied. In case, the supplier is capable of satisfying these additional demands through its present resources then the problem is trivial. So, let us assume the case when the supplier has limited resources (which cannot be expanded all of a sudden) to meet all such demands. The supplier wants to decide on the demand

fulfilment proportions based on criteria of maximum possible profit to be obtained, present business relations with each customer firm, and future business prospects. The scenario is depicted in Figure 1.

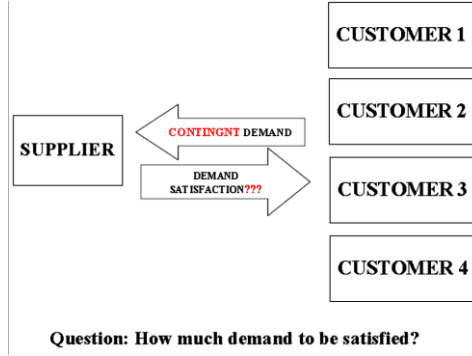


Fig. 2. Depiction of problem structure

We attempt to measure the present business relations and future business aspects quantitatively from the linguistic feedbacks on various relevant components of experts of the firm. This quantitative measurement is explained in the case studied in section V. For understanding the decision model we, for the time being, assume these to be measured and known parameters.

V. PROBLEM FORMULATION

A. Notations for the Proposed Model

TABLE I. NOTATIONS

Indices

I	number of customers
K	number of different types of products manufactured by supplier
i	index for customers ($i = 1, 2, \dots, I$)
k	index for products ($k = 1, 2, \dots, K$)

Parameters

D_{ik}	Demand of k^{th} product from i^{th} customer/buyer
C	Total Budget/Money in hand
C_{ik}	Cost of producing and delivering k^{th} product to i^{th} buyer
P_{ik}	Per unit price of k^{th} product paid by i^{th} buyer
e_{ik}	Production resource consumption
S_k	Space occupied by k^{th} product
S	Total storage available space
M	Total production capacity
L_i	Business Relationship factor with i^{th} buyer
E_{ik}	Factor corresponding to future demand of k^{th} product by i^{th} buyer

Variables

X_{ik}	Number of units of k^{th} product to be supplied to i^{th} buyer
----------	--

B. Model

Our assumption, that the vendor under its present capacity cannot satisfy the contingent demands of all its customers for various products, can be formulated as follows.

$$\sum_i \sum_k X_{ik} < \sum_i \sum_k D_{ik} \quad (12)$$

The multi-objective mathematical model representing our problem is described as follows.

$$(P) \quad \max z_1 = \sum_i \sum_k (P_{ik} - C_{ik}) * X_{ik} \quad (13)$$

$$\max z_2 = \sum_i \sum_k L_i * X_{ik} \quad (14)$$

$$\max z_3 = \sum_i \sum_k E_{ik} * X_{ik} \quad (15)$$

subject to constraints

$$\sum_i \sum_k e_{ik} * X_{ik} \leq M \quad (16)$$

$$\sum_i \sum_k S_k * X_{ik} \leq S \quad (17)$$

$$\sum_i \sum_k C_{ik} * X_{ik} \leq C \quad (18)$$

$$0 \leq X_{ik} \leq D_{ik} \quad \forall i, k. \quad (19)$$

The objective (13) is aimed to maximize the total profit to be earned by satisfying the demands of various products from customers fully or partially. The objective (14) aims to maximize the business relationship of the vendor with all the customers. The objective (15) aims to give more weight to fulfilling the demand of those products whose demand may grow in future. Constraint (16) gives a bound on the total of actually fulfilled demand because of production capacity constraints. Constraint (17) gives a bound on the same because of available storage space. Constraint (18) gives a bound on the same because of total available budget. Constraint (19) limits the number of units of any product to be supplied to a customer to not more than those actually demanded and not less than zero.

C. Solution Algorithm

Step 1. The objectives are first *normalized* by solving three optimization problems each with a single objective to obtain the “Positive Ideal Solution” (PIS) and the “Negative Ideal Solution” (NIS) following the approach proposed by Lai and Hwang [6].

Step 2. The *weights* for the objectives are calculated using *analytical hierarchical process* (AHP), a popular multi-criteria decision making method introduced by Saaty [9] in 80's through numerous research articles and books.

Step 3. *Weighted sum approach* is used to solve the multi-objective decision making problem (P) by converting it into single objective optimization problem.

VI. IMPERICAL ILLUSTRATION

We considered a case of a supplier who faces a contingent demand of two of its products from four buyers. Case oriented assumed data is obtained from an industry expert of a manufacturing firm. The data is tabulated below.

TABLE II. Cost (per unit) of producing and supplying k^{th} product to i^{th} buyer (including logistic costs) (in Rs.): C_{ik}

C_{ik}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Product 1	2654	2521	2654	2654
Product 2	1600	1520	1600	1600

TABLE III. P_{ik} : price paid by i^{th} buyer for k^{th} product (in Rs.)

P_{ik}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Product 1	3800	3728	4000	4000
Product 2	2280	2232	2400	2400

TABLE IV. Contingent demand of k^{th} product from i^{th} buyer: D_{ik}

D_{ik}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Product 1	50	110	35	20
Product 2	65	110	35	20

TABLE V. Production resource consumption e_{ik} Space occupied by k^{th} product (in cubic inch): S_k

	Product 1	Product 2
e_{ik} (machine-hours)	39.81	28.80
S_k (cubic inches)	735	506

TABLE VI. Bounds of the model M , S and C

M (Available machine hours)	S (Total storage space available – in cubic inches)	C (Total Budget available)
18070	259688	Rs. 9,03,500

Now we demonstrate the quantification of the factors of business relation L_i and future business prospects E_{ik} .

A. Quantification of Business Relation

1. The fraction of the regular demand of i^{th} customer which he purchases from the supplier.

L_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
$L_{i(f=1)}$	0.5	0.85	0.5	0.2

2. Profit percentage earned from i^{th} customer.

L_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
%age	25%	50%	20%	5%
$L_{i(f=2)}$	0.25	0.5	0.2	0.05

3. Satisfaction or ease of doing business with the customer

Transactions type / Total transactions	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Unsuccessful	5/100	2/80	5/50	0/10
Satisfactory	30/100	8/80	15/30	3/10
Excellent	65/100	70/80	30/50	7/10

Computing the three point membership values as following.

	Buyer 1	Buyer 2	Buyer 3	Buyer 4
$\mu_{i \text{ low}}$	0.05	0.025	0.1	0
$\mu_{i \text{ medium}}$	0.3	0.1	0.5	0.3
$\mu_{i \text{ high}}$	0.65	0.875	0.6	0.7

The weights for the three types of transacions prescribed by the expert are as follows.

weights	Buyer 1	Buyer 2	Buyer 3	Buyer 4
$w_{\text{unsuccessful}}$	0	0	0	0
$w_{\text{satisfactory}}$	0.3	0.4	0.4	0.2
$w_{\text{excellent}}$	0.7	0.6	0.6	0.8

The weighted avarage score thus obtained are as following.

L_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
$L_{i(f=3)}$	0.545	0.565	0.56	0.62

4. Payment lead-time or grace period

Using the membership functions as defined in equations (5), (6) and (7) and weighted aggregate as defined in equation (8) for payment lead-times.

		w_{small} 0.5	w_{medium} 0.35	w_{high} 0.15	
Buyer	Average lead-time	μ_{small}	μ_{medium}	μ_{high}	$L_{i(f=4)}$
Buyer 1	75	0	0	1	0.15
Buyer 2	35	0	0.833	0.167	0.3166
Buyer 3	20	0.5	0.5	0.425	0.425
Buyer 4	5	1	0	0	0.5

Aggregating the above scores; the weights for the preferences of these four factors given by the industry expert are listed below in the following table showing the calculation of L_i .

L_{if}	L_{i1}	L_{i2}	L_{i3}	L_{i4}	
weights (w_{if})	0.2	0.5	0.2	0.1	L_i
Buyer 1	0.5	0.25	0.55	0.15	0.35
Buyer 2	0.85	0.5	0.565	0.3166	0.56466
Buyer 3	0.5	0.2	0.48	0.425	0.3385
Buyer 4	0.2	0.05	0.62	0.5	0.239

B. Quantification of Future Business Prospects

1. Years of Business Collaboration

The three point membership functions for categorizing the years of business collaborations as initial, intermediate and high have been defined as follows.

$$\mu_{ik \text{ initial}} = \begin{cases} 1 & \text{if } 0 \leq P_{ik} < 5 \text{ years} \\ (8 - P_{ik})/3 & \text{if } 5 \leq P_{ik} < 8 \text{ years} \\ 0 & \text{if } P_{ik} > 8 \text{ years} \end{cases} \quad (20)$$

$$\mu_{ik \text{ intermediate}} = \begin{cases} 0 & \text{if } 0 \leq P_{ik} < 5 \text{ years} \\ (P_{ik} - 5)/3 & \text{if } 5 \leq P_{ik} < 8 \text{ years} \\ (15 - P_{ik})/7 & \text{if } 8 \leq P_{ik} < 15 \text{ years} \\ 0 & \text{if } P_{ik} > 15 \text{ years} \end{cases} \quad (21)$$

$$\mu_{ik \text{ high}} = \begin{cases} 0 & \text{if } P_{ik} \leq 8 \text{ years} \\ (t_i - 8)/7 & \text{if } 8 \leq P_{ik} < 15 \text{ years} \\ 1 & \text{if } P_{ik} > 15 \text{ years} \end{cases} \quad (22)$$

The data of pertaining to years of collaboration of the four buyers for each of the two products is as follows.

Years	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Product 1	20	10	3	1
Product 2	10	4	1	1

Using the prescribed weights as $w_{\text{initial}} = 0.02$, $w_{\text{intermediate}} = 0.02$, $w_{\text{high}} = 0.02$, the first factor is calculated as follows.

$E_{ik(f=1)}$	Buyer 1	Buyer 2	Buyer 3	Buyer 4
Product 1	0.833	0.343	0.06	0.02
Product 2	0.343	0.08	0.02	0.02

2. Reputation of buyer as an enterprise in its market

E_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>Financial Stability</i>	3	2	1	1
<i>Market Value</i>	3	2	1	1
<i>Recent Results</i>	3	2	1	0

The scores for the reputation factor are thus obtained as follows.

E_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
$E_{i(f=2)}$	1	0.67	0.33	0.2

3. Customer satisfaction factor from scorecard ratings are taken as follows.

E_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>%age</i>	90%	95%	80%	50%
$E_{i(f=3)}$	0.9	0.95	0.8	0.5

4. Friendly Managerial Policy

Linguistically characterized managerial policy of the each of the buyers is quantized as following.

E_{if}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>Points</i>	2	5	4	3
$E_{i(f=4)}$	0.4	1.0	0.8	0.6

5. Market Performance of product of k^{th} buyer which uses i^{th} product as spare part/ raw material.

Similarly, the market performances are quantized through qualitative grading as follows.

<i>Market Performance</i>	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>Product 1</i>	5	4	3	2
<i>Product 2</i>	4	3	3	4

$E_{ik(f=5)}$	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>Product 1</i>	1	0.8	0.6	0.4
<i>Product 2</i>	0.8	0.6	0.6	0.8

Aggregating the scores; the weights for the preferences of these four factors given by the industry expert are listed below, followed by the tabulation of calculation of scores for the factor E_{ik} .

Weights	Business Partnership in Market	Reputation Satisfactor	Customer Managerial Policy	Market Performance
(w_{te_f})	0.35	0.2	0.15	0.15

		E_{ik1}	E_{ik2}	E_{ik3}	E_{ik4}	E_{ik5}	
	w_{te_f} :	0.35	0.2	0.15	0.15	0.15	E_{ik}
Buyer 1	<i>Product 1</i>	0.5	1	0.9	0.4	1	0.72
	<i>Product 2</i>	0.5	1	0.9	0.4	0.8	0.69
Buyer 2	<i>Product 1</i>	0.85	0.67	0.95	1	0.8	0.844
	<i>Product 2</i>	0.85	0.67	0.95	1	0.6	0.814
Buyer 3	<i>Product 1</i>	0.5	0.33	0.8	0.8	0.6	0.571
	<i>Product 2</i>	0.5	0.33	0.8	0.8	0.6	0.571
Buyer 4	<i>Product 1</i>	0.2	0.2	0.5	0.6	0.4	0.335
	<i>Product 2</i>	0.2	0.2	0.5	0.6	0.8	0.395

C. Obtaining Weights of Multi-objectives problem using AHP

Using 25 responses from the management personals of the firm, the weights corresponding to the three objectives of our model have been obtained using AHP as

$$w_1 = 0.46, w_2 = 0.31 \text{ and } w_3 = 0.23.$$

Then the single objective integer programming problem was solved using the MATLAB 2014a.

VII. RESULTS AND ANALYSIS

The solution of the single objective problem obtained after normalizing the multi-objective problem is obtained as follows.

TABLE VII. Demand Fulfilment Decision

X_{ik}	Buyer 1	Buyer 2	Buyer 3	Buyer 4
<i>Product 1</i>	50	110	35	7*
<i>Product 2</i>	65	110	35	20

The optimal decision on fulfillment of demand of the four buyers for the two products is found to satisfy all the demands completely except that of Buyer 4 for the Product 1.

Let us try to analyse and verify these results with the actual situation and managerial decision making practices.

In our case, let the Buyer 1 is a buyer having good business relationship with the supplier and is having significantly long collaboration for 25 years. Buyer 2 is signified with its very high demand and over that this buyer purchases an 85% of his demand from this supplier. So, this is also one of the buyers having good business relations and future business prospects with the supplier. Among the rest of two buyers, Buyer 3 is having comparatively higher demand, purchasing more demand proportions and contributes to more profit than the Buyer 4.

Comparing these features, the demands of first three buyers have been fulfilled which is within the suppliers capacity of its available resources. Lastly, falling short of resources, the supplier has decided on the demand of Product 2 fully but that of Product 1 partially. This is due to the fact that the market performance of the product of this buyer using Product 1 is good but that of the Product 2 is poor. So, looking through his own future business aspect this is justified.

VIII. CONCLUSIONS

In this paper we propose a multi-objective optimization model to decide on satisfying contingent demands of various buyers of various products for a supplier with a vision on present business relations and future business aspect along with the profit making objectives. For this we have demonstrated the quantification of linguistic terms featuring many factors which are intended to be considered in the decision making process. Using weighted sum approach, the problem is converted into a single objective optimization problem which can then be solved using the standard optimization techniques.

Acknowledgment

The first author is thankful to the University Grant Commission, India, for research grant.

References

- [1] Y. Amer, L. Luong, S-H. Lee, "Case study: Optimizing order fulfillment in a global retail supply chain, *Int. J. Production Economics*, vol. 127, 2010, p. 278–291.
- [2] C.T. Chen, W.S. Tai, "Measuring the intellectual capital performance based on 2-tuple fuzzy linguistic information", in: *Proceedings of the 10th Annual Meeting of Asia Pacific Region of Decision Sciences Institute, APDSI, Taiwan, 2005*.
- [3] L. X. Cui, K.L. Mak, and S.T. Newman, "Optimal supplier selection and order allocation for multi-product manufacturing featuring customer flexibility" *International Journal of Computer Integrated Manufacturing*, vol. 28 (7), 2015 p. 729 – 744.
- [4] A. Gupta, C. D. Maranasa, and C. M. McDonald , "Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management", *Computers & Chemical Engineering*, vol. 24 (12), December 2000, p. 2613-2621.
- [5] F. Herrera, and E. Herrera-Viedma, "Linguistic decision analysis: steps for solving decision problems under linguistic information", *Fuzzy Sets and Systems*, vol. 115, 2000, p. 67 – 82.
- [6] Y. J. Lai, C. L. Hwang, *Fuzzy Multiple Objective Decision Making: Methods and Applications, Lecture Notes in Economics and Mathematical Systems*, Springer, 1994.
- [7] J. Li, "Customer demand satisfaction in production systems: a due-time performance approach", *Robotics and Automation, IEEE Transactions*, vol 17 (4), 2002, p. 472 – 482.
- [8] A. Nagurney, "Optimal supply chain network design and redesign at minimal total cost and with demand satisfaction", *International Journal of Production Economics*, vol 128 (1), November 2010, p. 200–208.
- [9] T. L. Saaty, *The Analytic Hierarchy Process*, New York: McGraw Hill. International, Revised editions, Paperback (1996, 2000), Pittsburgh: RWS Publications, 1980.
- [10] M. Setak, S. Sharifi and A. Alimohammadian, "Supplier Selection and Order Allocation Models in Supply Chain Management: A Review", *World Applied Sciences Journal*, vol. 18 (1), 2012, p. 55-72.
- [11] S.A. Torabi, M. Baghersad, S.A. Mansouri, "Resilient supplier selection and order allocation under operational and disruption risks, *Transportation Research Part E: Logistics and Transportation Review*, vol. 79, July 2015, p. 22–48.

Effect of Different Media and Growth Hormones on Shoot Multiplication of *In Vitro* Grown *Centella asiatica* Accessions

Arpita Roy, Koyel Kundu, Gaurav Saxena, Lakhan Kumar and Navneeta Bharadvaja*

Plant Biotechnology Laboratory, Department of Biotechnology, Delhi Technological University, New Delhi-110042, India

Abstract

Centella asiatica also known as Gotu Kola is an important medicinal plant which contains several bioactive compounds such as triterpenoid saponins including asiaticoside, madecassoside, centelloside, asiatic acid, etc. In addition, *Centella sp.* contains other components including flavonoids, phytosterols, tannins, amino acids, sugars, etc. Due to its medicinal importance, this plant is being overexploited and it is essential to conserve this plant. In the present investigation, comparative study of different accessions of *Centella asiatica* for shoot multiplication in different media and growth hormones was performed. Shoot multiplication and phytochemicals production from different accessions is required to be assessed to choose an accession giving optimum production of phytochemicals. For this, different media containing different combinations of macro and micro nutrients have to be tested that influence the growth of the plant. Assessment of best culture media and concentrations of plant hormone for shoot culture are some of the critical culture conditions to achieve. So there is a need to optimise such conditions that will enhance the growth of the plant. For shoot culture, explant of different accession of *Centella asiatica* were inoculated in different media i.e. Murashige and Skoog (MS), Gamborg's B5 and Nitsch medium which were supplemented with standard concentrations of plant growth hormones. The cultures were incubated at $25 \pm 2^\circ\text{C}$ with photoperiod of 16 hours. After six week of incubation period, highest growth was found in MS media in all accessions. Further MS medium was supplemented with different combinations of growth hormones. After six week of incubation, MS medium supplemented with 1mg/l BAP showed the highest growth of the plant followed by 2mg/l BAP+0.5 NAA.

Keywords: *Centella asiatica*; Murashige and skoog (MS); Gamborg's B5; Nitsch; 6-Benzylaminopurine (BAP); Naphthaleneacetic acid (NAA)

Introduction

Medicinal plants are the traditional source of many pharmaceutically important compounds. In recent times, they are utilized by the pharmaceutical companies for the preparation of several formulations. *Centella asiatica* is one of the important traditional medicinal plant belonging to family Apiaceae and commonly known as Gotu kola in India. It is an important perennial medicinal herb found in the tropical and subtropical countries like India, Sri Lanka and Bangladesh. *C. asiatica* contains several triterpene, saponins like asiaticoside, asiatic acid, sapogenins, madecassic acid, vellarin, adecassoside, glycosides and centelloside [1]. Leaves contains high amount of triterpenoids [2]. It possesses several important properties like antileprotic, anti-stress, anti-feedent, anti-tuberculosis activities, wound-healing properties [3,4], antibacterial, atherosclerosis and fungicidal activity [5]. It is used in the treatment of leprosy, wound, cancer, fever, allergies [6], abscesses, asthma, catarrh, convulsions, dysentery, eczema, gonorrhea, hypertension, bronchitis, headache, jaundice, pleuritis, rheumatism, ulcers, spasms, tuberculosis, urethritis, etc. [7]. Leaves of this plant are rich in Vitamin B, C, and minerals such as magnesium, potassium, calcium, phosphorus and aluminium [8,9]. It is also used as brain tonic and blood purifier [10]. *C. asiatica* contains various flavonoids which include quercetin and kaempferol, rutin and naringin [11]. Roots are rich in amino acids like aspartic, glutamic, serine, alanine, threonine, histidine and lysine [12]. Underground parts contain many polyacetylenic compounds [12]. Due to its medicinal importance, this plant is overexploited and there is a decline in the population of the *Centella asiatica*. International Union for Conservation of Nature and Natural resources (IUCN) listed it as threatened plant and endangered species [13]. Tissue culture techniques can play an important role in the clonal multiplication of elite clones of this plant as well as conservation of its germplasm. For this study, different accessions of *Centella asiatica*

have been used. An accession refers to the collection of plant material from a single species which is collected at one time from a specific geographical location. Each accession is an attempt to capture the diversity present in a given plant population. Accession number is given a unique identifier, and it is used to maintain associated information in the database. It exhibits significant variations in morphological parameters like growth of leaf, flowering, stomatal frequency, etc. The purpose of using different accession was to choose the best accession for the phytochemicals production. To the best of our knowledge there is no study which describes the effect of different media on the shoot multiplication of this plant and it provides an opportunity to explore the role of media and growth hormone on the enhancement of *in vitro* culture of different accessions of *C. asiatica*. In this study, effects of different media like MS, Gamborg's B5 and Nitsch and the concentrations of plant growth hormones on the growth of three *Centella asiatica* accessions were observed and reported.

Materials and Method

Plant material

Three different accessions of *in vitro* grown plantlets of *Centella asiatica* were collected from the NBPGR (National Board of Plant

*Corresponding author: Navneeta Bharadvaja, Plant Biotechnology Laboratory, Department of Biotechnology, Delhi Technological University, New Delhi-110042, India, E-mail: navneetab@dtu.co.in

Received April 05, 2016; Accepted April 12, 2016; Published April 19, 2016

Citation: Roy A, Kundu K, Saxena G, Kumar L, Bharadvaja N (2016) Effect of Different Media and Growth Hormones on Shoot Multiplication of *In Vitro* Grown *Centella asiatica* Accessions. Adv Tech Biol Med 4: 172. doi: [10.4172/2379-1764.1000172](http://dx.doi.org/10.4172/2379-1764.1000172)

Copyright: © 2016 Roy A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Genome Research), New Delhi. Nodes of *C. asiatica* were utilized as explant for this study.

Culture medium and conditions

Effect of different media on shoot multiplication and shoot length:

Three different plant tissue culture medium i.e. Murashige and Skoog (MS) (1962) [14] with 3% (w/v) sucrose and 0.8% (w/v) agar, Gamborg's B5 (HiMedia Laboratories Pvt. Ltd., India) [15] and Nitsch (HiMedia Laboratories Pvt. Ltd., India) [16] media were used. The composition of each medium is mentioned in Table 1. These media were supplemented with standard concentration of plant growth hormone i.e. 4.0 mg/L BAP+0.4 mg/L NAA. The pH of media was adjusted to 5.8 with 1N NaOH or 1N HCl and media were autoclaved at 121°C for 20 minutes at 15 psi. The explants were then inoculated in the medium under aseptic conditions and incubated at 25 ± 2°C with photoperiod of 16 hours under cool-white fluorescent tubes for six weeks.

Effect of growth hormones on shoot multiplication and shoot length:

MS medium was augmented with different concentrations of plant growth hormones i.e. 1.0 mg/l BAP, 2.0 mg/l BAP, 1.0 mg/l BAP+0.5 NAA and 2.0 mg/l BAP+0.5 mg/l NAA. Explants were inoculated and

cultures were incubated at 25 ± 2°C with photoperiod of 16 hours under cool-white fluorescent tubes for six weeks. Regenerated shoots were subcultured every three weeks in the same media (Table 1).

Data analysis

Visual observations were recorded in terms of number of shoots per explant and the length of each shoot. Experiments were done in triplicates and means of each experiment was carried out to detect the significant differences.

Results and Discussion

Effect of different media on shoot multiplication and shoot length

This study was an attempt to correlate the effect of different media and plant hormone concentration on the shoot multiplication of *Centella asiatica* accessions. To initiate the study, nodal explant were taken from *in vitro* grown plants. Shoot multiplication of *Centella asiatica* nodal explants cultured on MS, Gamborg's B5 and Nitsch media supplemented with 4.0 mg/L BAP+0.4 mg/L NAA. After two weeks of incubation explants showed the growth response in different media. After six weeks of incubation period it was found that MS media showed highest shoot multiplication as compared to Gamborg's B5 and Nitsch media in all the three accessions. In case of MS media, the highest shoot multiplication was observed as follow, 5.5 ± 0.22 in accession no.-342109, 5.0 ± 0.22 in accession no.-347492 and 4.5 ± 0.22 in accession no.-331514. In case of Nitsch media highest shoot multiplication was observed as follow 4.3 ± 0.37 in accession no.-342109, 4 ± 0.33 in accession no.-331514 and 2 ± 0.31 in accession no.-347492. Whereas Gamborg's B5 media showed lowest shoot multiplication in all the three accessions. Similar results were also reported where MS media was supplemented with 4.0 mg/l BAP+0.1mg/l NAA [17]. The details of each accession are mentioned in Table 2 (Figure 1).

Effect of growth hormones on shoot multiplication and shoot length

Nodal explants were cultured in MS medium supplemented with different concentrations and combination of auxin (NAA) and cytokinin (BAP) to assess their effect on shoot multiplication of *C. asiatica* accessions. The highest number of shoots as well as length of

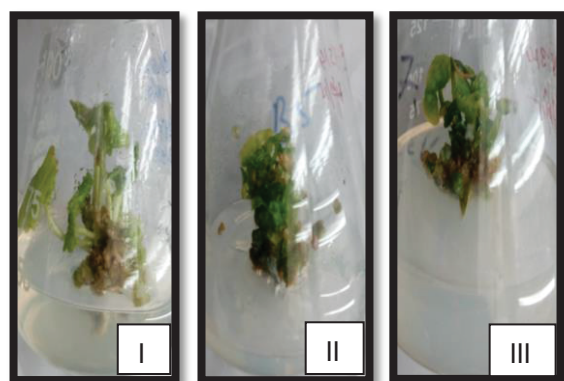


Figure 1: *In vitro* culture of *Centella asiatica* after six weeks of inoculation. I) MS, II) B5 and III) Nitsch Medium.

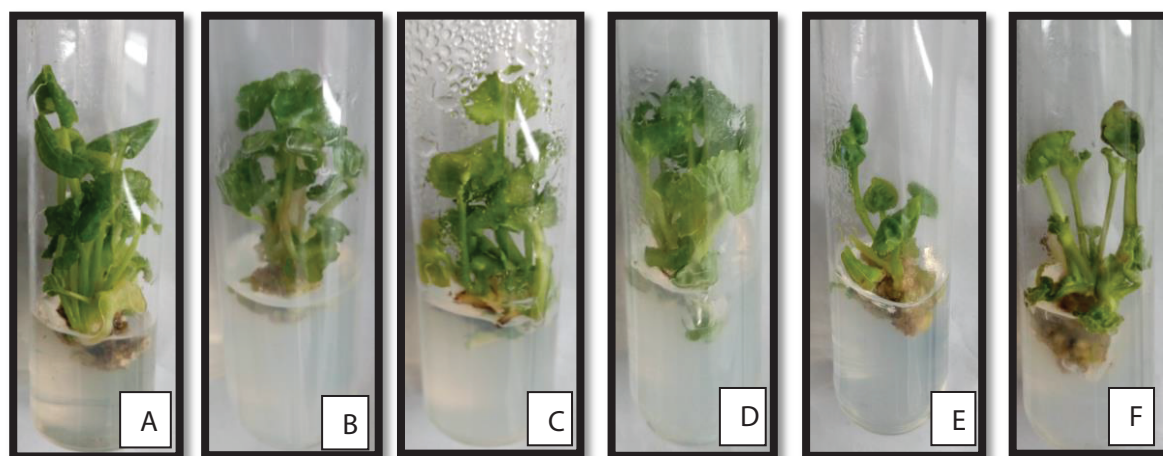


Figure 2: Figure showing effect of MS+1mg/l BAP on the growth in accession no-342109 (A), accession no-347492 (B) accession no-331514 (C) and also effect of MS+2mg/l BAP+0.5NAA on the growth in accession no-342109 (D), accession no-347492 (E) accession no-331514 (F).

MS Medium [14]		B5 Medium [15]		Nitsch Medium [16]	
Components	mg/L	Components	mg/L	Components	mg/L
Macronutrients		Macronutrients		Macronutrients	
Ammonium nitrate	1,650.00	Potassium nitrate	2500	Potassium nitrate	950
Potassium nitrate	1,900.00	Ammonium Sulphate	134	Ammonium nitrate	720
Calcium chloride (anhydrous)	332.2	Calcium chloride.2H ₂ O	150	Magnesium sulphate anhydrous	90.34
Magnesium sulfate (anhydrous)	180.7	Magnesium sulphate	122.1	Potassium phosphate monobasic	68
Potassium phosphate monobasic	170	Sodium phosphate monobasic	130.4	Micronutrients	
Micronutrients		Micronutrients		Manganese sulphate.H ₂ O	18.94
Manganese sulfate monohydrate	16.9	Manganese sulphate. H ₂ O	10	Boric acid	10
Ferrous sulfate heptahydrate	27.8	Boric acid	3	Molybdic acid (sodium salt).2H ₂ O	0.25
Zinc sulfate heptahydrate	8.6	Potassium iodide	0.75	Zinc sulphate.7H ₂ O	10
Boric acid	6.2	Molybdic acid (sodium salt).2H ₂ O	0.25	Copper sulphate.5H ₂ O	0.025
Potassium iodide	0.83	Zinc sulphate.7H ₂ O	2	Ferrous sulphate.7H ₂ O	27.85
Sodium molybdate dehydrate	0.25	Copper sulphate.5H ₂ O	0.025	EDTA disodium salt.2H ₂ O	37.25
Cobalt chloride hexahydrate	0.025	Cobalt chloride.6H ₂ O	0.025	Vitamins	
Cupric sulfate pentahydrate	0.025	Ferrous sulphate.7H ₂ O	27.8	Myo-Inositol	100
Disodium EDTA dehydrate	37.26	EDTA disodium salt.2H ₂ O	37.3	Thiamine hydrochloride	0.5
Vitamins		Vitamins		Pyridoxine hydrochloride	0.5
myo-Inositol	100	Myo-Inositol	100	Nicotinic acid	5
Nicotinic acid	0.5	Thiamine hydrochloride	10	Folic acid	0.5
Pyridoxine hydrochloride	0.5	Pyridoxine hydrochloride	1	Biotin	0.05
Thiamine hydrochloride	0.1	Nicotinic acid	1	Glycine	2
Sugar		Sugar		Sugar	
Sucrose	30000	Sucrose	20000	Sucrose	20000

Table 1: Composition of MS, B5 and Nitsch medium investigated in this study.

Different media with standard hormones	Accessions					
	342109		347492		331514	
	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)
MS + 4 mg/l BAP + 0.4 mg/l NAA	5.5 ± 0.22	2.6 ± 0.11	5.0 ± 0.22	2.9 ± 0.31	4.5 ± 0.22	2.1 ± 0.22
B5 + 4 mg/l BAP + 0.4 mg/l NAA	1.0 ± 0.38	1.4 ± 0.22	1.0 ± 0.33	0.6 ± 0.22	2.0 ± 0.38	0.35 ± 0.24
Nitsch + 4 mg/l BAP + 0.4 mg/l NAA	4.3 ± 0.37	1.25 ± 0.31	2.0 ± 0.31	0.8 ± 0.33	4.0 ± 0.33	0.48 ± 0.21

Table 2: Effect of different media on the number and length of shoots of three different accessions of *Centella asiatica* after six week of inoculation. Values are expressed as mean ± Standard Error (M ± SE). MS: Murashige and Skoog medium; B5: Gamborg's B5 medium; BAP: 6-Benzyl amino purine; NAA: Naphthaleneacetic acid.

MS Media with different hormonal concentrations	342109		347492		331514	
	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)	Number of Shoots (M ± SE)	Length of Shoots (in cm) (M ± SE)
MS+1 mg/l BAP	12.5 ± 0.22	3.5 ± 0.32	7.0 ± 0.22	2.7 ± 0.31	11.5 ± 0.22	2.9 ± 0.24
MS+2 mg/l BAP	3.0 ± 0.22	2.5 ± 0.22	3.0 ± 0.37	1.6 ± 0.24	4.0 ± 0.4	1.7 ± 0.31
MS+1 mg/l BAP+0.5 NAA	2.7 ± 0.37	1.0 ± 0.31	2.0 ± 0.22	1.3 ± 0.32	3.0 ± 0.38	1.2 ± 0.21
MS+2 mg/l BAP+0.5 NAA	5.3 ± 0.38	2.0 ± 0.24	6.2 ± 0.22	2.7 ± 0.24	8.2 ± 0.33	1.5 ± 0.22

Table 3: Effect of different hormone concentrations on number and length of shoots of three accessions of *Centella asiatica* after six week of inoculation. Values are expressed as mean ± Standard Error (M ± SE). MS: Murashige and Skoog medium; BAP: 6-Benzyl amino purine; NAA: Naphthaleneacetic acid.

shoots was observed in MS medium with hormonal concentration of 1mg/l BAP followed by 2 mg/l BAP+0.5 mg/l NAA. Multiplication of shoots was observed after 14 days. After six weeks of incubation MS medium supplemented with 1mg/l BAP showed shoot multiplication as follow, 12.5 ± 0.22 in accession no.- 342109, 7 ± 0.22 in accession

no.- 347492 and 11.5 ± 0.22 in accession no.- 331514. While MS medium supplemented with 2 mg/l BAP+0.5 mg/l NAA showed shoot multiplication of 5.3 ± 0.38 in accession no.- 342109, 6.2 ± 0.22 in accession no.- 347492 and 8.2 ± 0.33 in accession no.- 331514. Similar results were reported where BAP alone showed the good shoot

induction. In general, BAP is the most efficient growth hormone for the shoot proliferation [18]. It mimics as an inhibitor agent and function against apical dominance of shoot induction and shoot bud formation [19]. Several studies reported that media supplemented with BAP and NAA have also useful for the shoot multiplication. One article reported that MS media supplemented with 22.2 μ M BA+2.68 μ M NAA showed highest growth [20] where as another article reported that maximum shoot multiplication was observed at 2 mg/l BAP [21]. Results of this study indicate that large scale propagation of this plant is feasible and several plantlets can be regenerated from a single nodal explant. Details of this experiment are mentioned in Table 3 (Figure 2).

Conclusion

In this investigation, it was concluded that the MS medium with concentration of BAP 1 mg/l supports the maximum shoot multiplication and length of shoots for all the three accessions of *Centella asiatica*. It was also observed that the maximum number of shoots and length of shoots obtained for Accession-342109 in comparison to the other two accessions. Further analysis of phytochemicals will be done with all these cultures. Findings in this investigation have proven an efficient media and plant growth hormone concentration for the mass propagation of this plant. These findings would be useful in conservation and micropropagation of this plant. Future efforts are in progress to evaluate the phytochemical present in this plant.

Acknowledgement

Sincere thanks to the Department of Biotechnology, Delhi Technological University for making necessary facilities during this study and National Bureau of Plant Genetic Resources for providing the plant materials.

References

- Glasby JS (1991) Dictionary of Plants Containing Secondary Metabolites. Taylor and Francis, London.
- Zainol NA, Voo SC, Sarmidi HR, Aziz RA (2008) Profiling of *Centella asiatica* (L.) Urban Extract. The Malaysian Journal of Analytical Sciences 12: 322-327.
- Chakraborty T, Sinha Babu SP and Sukul NC (1996) Preliminary evidence of antifilarial effect of *Centella asiatica* on canine dirofilariasis. Fitoterapia 67: 110-112.
- Srivastava R, Shukla YN, Kumar S (1997) Chemistry and pharmacology of *Centella asiatica*: a review. J Medi Arom Plant Sci 19: 1049-1056.
- Oyediji OA, Afolayan AJ (2005) Chemical composition and antibacterial activity of the essential oil of *Centella asiatica* growing in South Africa (J). Pharmaceutical Biology 43: 249-252.
- Inamdar PK, Yeole RD, Ghogare AB, De Souza NJ (1996) Determination of biologically active constituents in *Centella asiatica*. J. Chromatography 742: 127-130.
- Hausen BM (1993) *Centella asiatica* (Indian pennywort), an effective therapeutic but a weak sensitizer. Contact Dermatitis 29: 175-179.
- Herbert D, Paramasivan CN, Prabhakar R, Swaminathan G (1994) In vitro experiments with *Centella asiatica*: investigation to elucidate the effect of an indigenously prepared powder of this plant on the acid-fastness and viability of *M. tuberculosis*. Indian J Lepr 66: 65-68.
- Brinkhaus B, Lindner M, Schuppan D, Hahn EG (2000) Chemical, pharmacological and clinical profile of the East Asian medical plant *Centella asiatica*. Phytomedicine 7: 427-448.
- Jorge OA, Jorge AD (2005) Hepatotoxicity associated with the ingestion of *Centella asiatica*. Rev Esp Enferm Dig 97: 115-124.
- Zainol MK, Abd-Hamid A, Yusuf S (2003) Antioxidative activity and total phenolic compounds of leaf, root and petiole of four accessions of *Centella asiatica* (L.) Urban (J). Food Chemistry 81: 575-581.
- Arora D, Kumar M, Dubey SD (2002) *Centella asiatica*-a review of its medicinal uses and pharmacological effects. Journal of Natural Remedies 2: 143-149.
- Singh S, Gautam A, Sharma A, Batra A (2010) *Centella asiatica* L. a plant with immense potential but threatened. International Journal of Pharm. Sci Review and Research 4: 9-17.
- Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiology Plant 15: 473-497.
- Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res 50: 151-158.
- Nitsch JP, Nitsch C (1969) Haploid plants from pollen grains. Science 163: 85-87.
- Das R, Hasan MF, Hossain MS, Rahman M (2008) Micropropagation of *Centella asiatica* L. an important medicinal herb. Progress. Agri 19: 51-56.
- George S, Remashree AB, Sebastian D, Hariharan M (2004) Micropropagation of *Centella asiatica* L. through axillary bud multiplication. Phytomorphology 54: 31-34.
- Wang PJ, Charle A (1991) Micropropagation through meristem culture in biotechnology in agriculture and forestry. Springer Verlag 17: 41-44.
- Tiwari KN, Sharma NC, Tiwari V, Singh BD (2000) Micropropagation of *Centella asiatica* (L.), a valuable medicinal herb. Plant cell, Tissue and Organ Culture 63: 179-185.
- Karthikeyan K, Chandran C, Klothungan S (2009) Rapid clonal multiplication through *in vitro* axillary shoot proliferation of *Centella asiatica* L. Indian Journal of Biotechnology 8: 232-235.

Citation: Roy A, Kundu K, Saxena G, Kumar L, Bharadvaja N (2016) Effect of Different Media and Growth Hormones on Shoot Multiplication of *In Vitro* Grown *Centella asiatica* Accessions. Adv Tech Biol Med 4: 172. doi: [10.4172/2379-1764.1000172](https://doi.org/10.4172/2379-1764.1000172)

OMICS International: Publication Benefits & Features

Unique features:

- Increased global visibility of articles through worldwide distribution and indexing
- Showcasing recent research output in a timely and updated manner
- Special issues on the current trends of scientific research

Special features:

- 700+ Open Access Journals
- 50,000+ editorial team
- Rapid review process
- Quality and quick editorial, review and publication processing
- Indexing at major indexing services
- Sharing Option: Social Networking Enabled
- Authors, Reviewers and Editors rewarded with online Scientific Credits
- Better discount for your subsequent articles

Submit your manuscript at: <http://www.omicsonline.org/submission>

Effect of structured parameters on the hot-carrier immunity of transparent gate recessed channel (TGRC) MOSFET

Ajay Kumar¹ · Neha Gupta¹ · Rishu Chaujar¹

Received: 18 August 2015 / Accepted: 20 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract In this work, the impact of parameter variation on hot-carrier effect immunity in transparent gate recessed channel (TGRC)—MOSFET based on the hydrodynamic energy transport model have been studied. The parameters of TGRC-MOSFET investigated include the oxide thickness, negative junction depth, and substrate doping. TCAD analysis shows the performance of TGRC-MOSFET in terms of transfer characteristics, transconductance, electric field, electron velocity, electron mobility and electron temperature. The simulation results indicate the improved hot-carrier immunity for TGRC-MOSFET in 30 nm device.

1 Introduction

The size of MOSFET's has been constantly scaled down to achieve advanced speed and packing density in VLSI technology. Apart from the intricacy of designing devices at these nanometer scales, high power dissipation and an increase in device temperature are prime concerns in semiconductor device design. The hot-carrier deprivation becomes the severe limitation to the reliability of nanometer devices and the packing density of VLSI (Ren and Hao 2002). Hot-carrier effect can be suppressed by reducing the supply voltage and gate oxide thickness (Fiegna et al. 1994),

but the supply voltage and gate oxide thickness are becoming inherent limits in nanometer regime. Under the influence of high lateral fields in short-channel MOSFETs, carriers in the channel and pinch-off regions of the transistor reach non-equilibrium energy distributions. The generation of these hot-carriers is the primary source of several reliability problems (Gupta et al. 2014). Hot-carriers can acquire sufficient energy to surmount the energy barrier at the Si–SiO₂ interface or tunnel into the oxide (Kumar et al. 2014b).

In MOSFETs, the substrate's hot carriers have been attracted towards the source and the drain electrode by the strong channel electric field before they get to interface when the device is scaled down to 100 nm regime. Therefore, sharp decrement in substrate hot-carrier is observed and then the influence of channel hot-carrier plays a dominant role in 100 nm device degradation (Ren and Hao 2002). These kind of mobile carriers in device create some physical damage which might change the device characteristics. Numerous device engineering schemes and novel device structures have been reported in the literature to overcome the problems arising from downscaling of device dimensions. Further scaling has faced serious limitations with respect to fabrication technology and device performances. Many solutions are proposed to overcome these limitations. Some of the solutions include such as SOI MOSFET (An et al. 2004), multi-gate MOSFET (Biswal et al. 2015), Recessed Channel (RC) MOSFET (Chaujar et al. 2008b), GAA MOSFET (Moreno et al. 2015), Silicon Nanowire MOSFET (Gupta et al. 2015a, b) etc. Conventional Recessed Channel MOSFET (CRC) MOSFET improves the short channel effects (SCEs) and HCEs appreciably because of the presence of groove which splits the source and drain regions.

In TGRC-MOSFET (Kumar et al. 2015, 2016a, b), a transparent conducting oxide is used instead of a metallic gate as it

✉ Rishu Chaujar
rishu.phy@dce.edu

Ajay Kumar
ajaykumar@dtu.ac.in

Neha Gupta
nehagupta@dtu.ac.in

¹ Department of Engineering Physics,
Delhi Technology University, Delhi 110042, India

has a relatively high concentration of free electrons in its conduction band. The free electrons arise either from defects in material or from extrinsic dopants, the impurity levels of which lie near the conduction band edge. Transparent conducting oxides open a broad range of power saving optoelectrical circuits. Electrical conductivity and optical transparency of indium tin oxide (ITO) is quite high and the ease with which it can be deposited as a thin film makes it a desirable option. The circuit can be improved by using ITO since it decreases the device power consumption (Pammi et al. 2011).

The metal gate is replaced by a transparent conducting oxide (Edwards et al. 2004; Kawazoe and Ueda 1999) called ITO. ITO is a solid solution with doped impurity of indium oxide (In_2O_3) and tin oxide (SnO_2). Commercially available ITO has a resistivity in the order of $1 \times 10^{-5} \Omega\text{-cm}$ (Minami 2005). It is transparent and colorless in thin layers. ITO is one of the most widely used transparent conducting oxides because of its two chief properties, its electrical conductivity and optical transparency as well as the ease with which it can be deposited as a thin film (Chaujar et al. 2008a, b).

2 Device structure and its description

A schematic representation of the TGRC-MOSFET is shown in Fig. 1. ITO is a transparent conducting material and shows high transparency as well as good conducting material when the workfunction is 4.7 eV. Total gate length is 30 nm. Oxide thickness is varying from 1 nm to 3 nm and negative junction depth is varying from 5 nm to 15 nm. Here substrate doping is p-type with concentration of $1 \times 10^{16}\text{cm}^{-3}$ and source/drain are n-type with uniform doping $1 \times 10^{19}\text{cm}^{-3}$. All the simulation have been done using ATLAS device simulator. Figure 1b shows the meshing in different regions of the device.

3 Simulation methodology and calibration

The simulations are based on the inversion layer Lombardi capacitance–voltage–temperature (CVT) mobility model with the Auger recombination model and Shockley–Read–Hall (SRH) for minority carrier recombination. Further, the hydrodynamic energy transport model comprising of the continuity equations, energy balance equations of the carriers, momentum transport equations and the Poisson equation have been used (ATLAS User's Manual 2014).

Calibration of model parameters used in simulation has been performed according to the published results obtained (Granzner et al. 2006) using above mentioned models and the results are shown in Fig. 2. The data has been extended (Granzner et al. 2006) and then plotted.

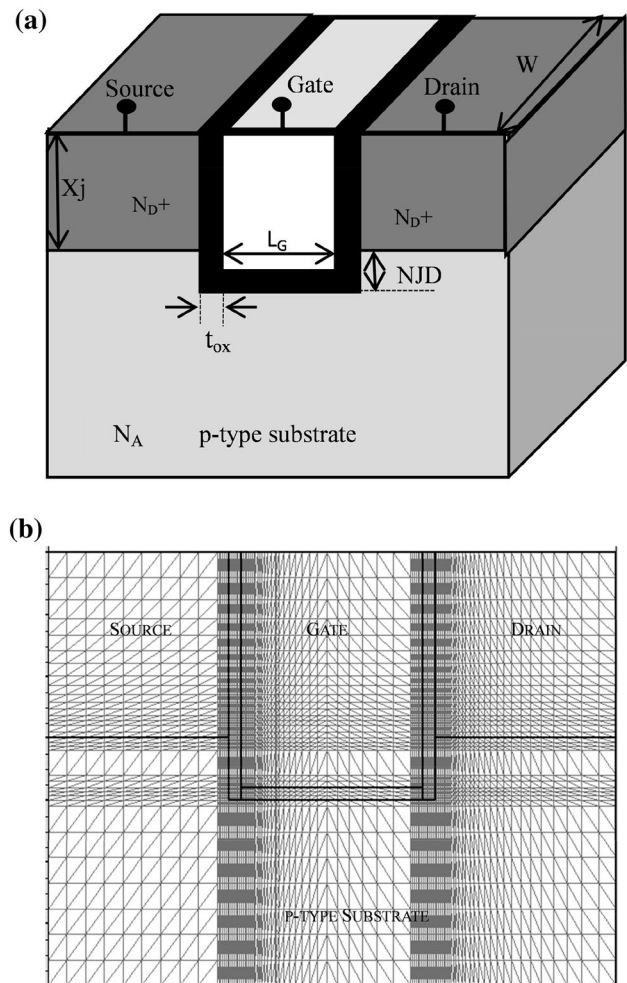


Fig. 1 a Schematic structure of transparent gate recessed channel MOSFET. b 2-D meshed structure of transparent gate recessed channel MOSFET

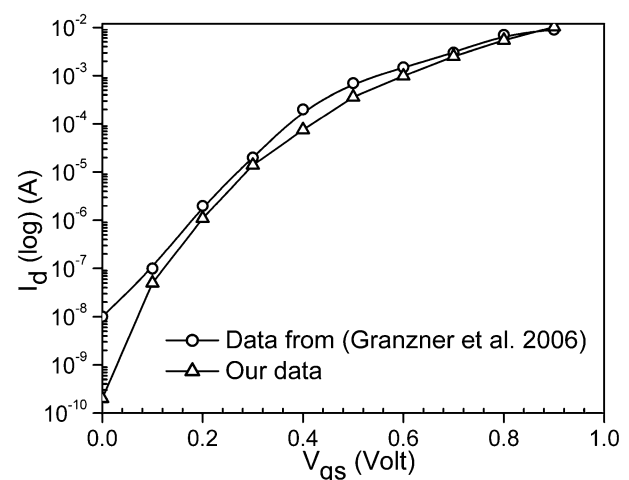


Fig. 2 Calibration with the published result, I_{ds} – V_{gs} characteristics of 40 nm gate length TGRC-MOSFET

4 Result and discussion

4.1 Effect of oxide thickness

If gate length is scaled down to 25 nm with 1.6 nm oxide thickness then, I_{OFF} is 70 nA while if gate length is reduced below 20 nm (18 nm) with oxide thickness below 1.5 nm (1.3 nm), then I_{OFF} increases above 100 nA due to quantum effects and tunneling of electron and causes SCEs (mainly increase in I_{OFF}) (Hu 2009; Iwai 2009) as mentioned in ITRS. Thus, there is need to optimize the value of oxide thickness. Therefore in this section, hot-carrier reliability of TGRC is studied in terms of oxide thickness.

In this section, the effects of oxide thickness variation have been observed. Figure 3 reflects the variation of the drain current (I_{ds}) and Transconductance (g_m) along the channel in TGRC-MOSFET at different oxide thickness. When oxide thickness is reduced from 3 to 1 nm, I_{ds} and g_m are enhanced because of more free charge carriers are generated in the channel due to the incorporation of ITO in gate region. These free charge carriers are contributing to enhancing the drain current. Thus, there are very less number of charge carriers which increases the hot-carriers in the channel and therefore hot carriers are reduced with the reduction of oxide thickness.

The hot-carrier effect is very sensitive to the electric field in the channel region (Ren and Hao 2002). Figure 4 shows variation of the electric field along the channel in TGRC-MOSFET at different oxide thickness. It is evident from Fig. 4 that electric field is very high on the source side as compared to drain side when the oxide

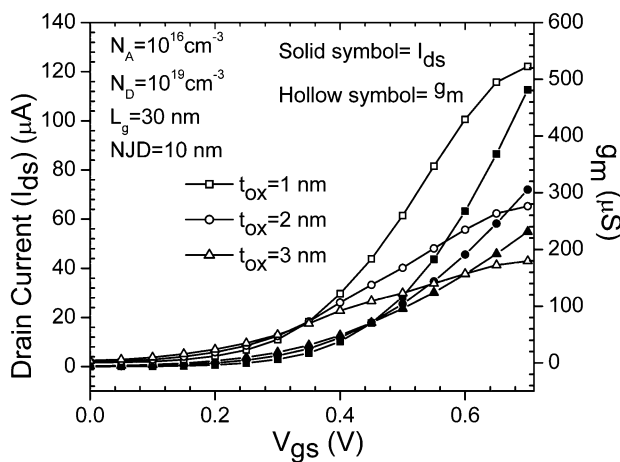


Fig. 3 Variation of drain current (I_{ds}) and transconductance (g_m) w.r.t. gate to source voltage (V_{gs}) in TGRC-MOSFET at different oxide thickness

thickness is 1 nm. If the oxide thickness is decreased from 3 to 2 nm, the electric field increases 29 % at the source side and increases 82 % when oxide thickness reduces from 2 to 1 nm, thus reflecting that if electric field is less at the drain side as compared to source side then the hot carrier generation will drop (Chaujar et al. 2008c).

Variation of electron temperature along the channel for TGRC-MOSFET at different oxide thickness is shown in Fig. 5. For improved device performance, electron temperature should be low (Kumar et al. 2014b). When oxide thickness is reduced from 3 to 2 nm and then to 1 nm, the electron temperature reduces by 4.97 %. It is clearly evident that hot-carrier-effect is less pronounced when oxide

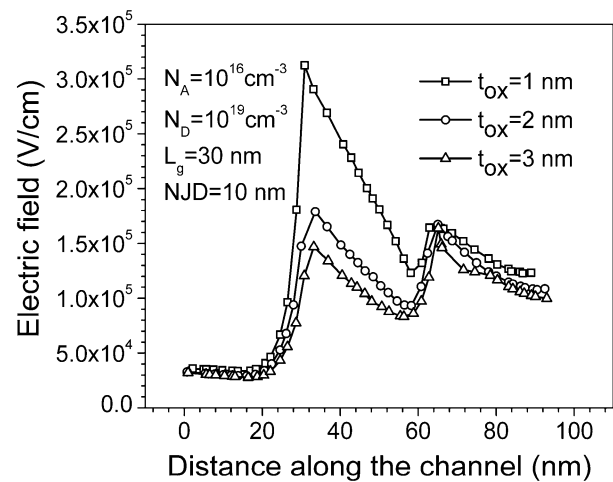


Fig. 4 Variation of electric field along the channel in TGRC-MOSFET at various oxide thicknesses

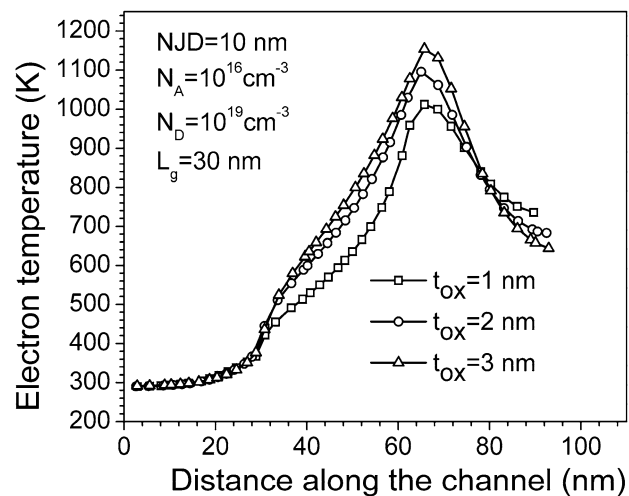


Fig. 5 Variation of electron temperature along the channel for TGRC-MOSFET at various oxide thicknesses

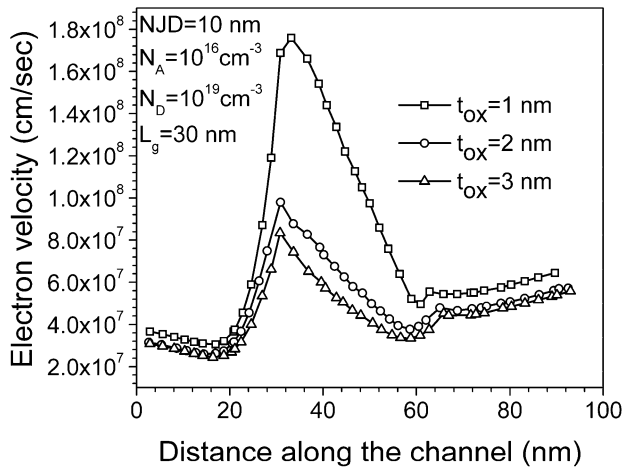


Fig. 6 Variation of electron velocity along the channel in TGRC-MOSFET at various oxide thicknesses

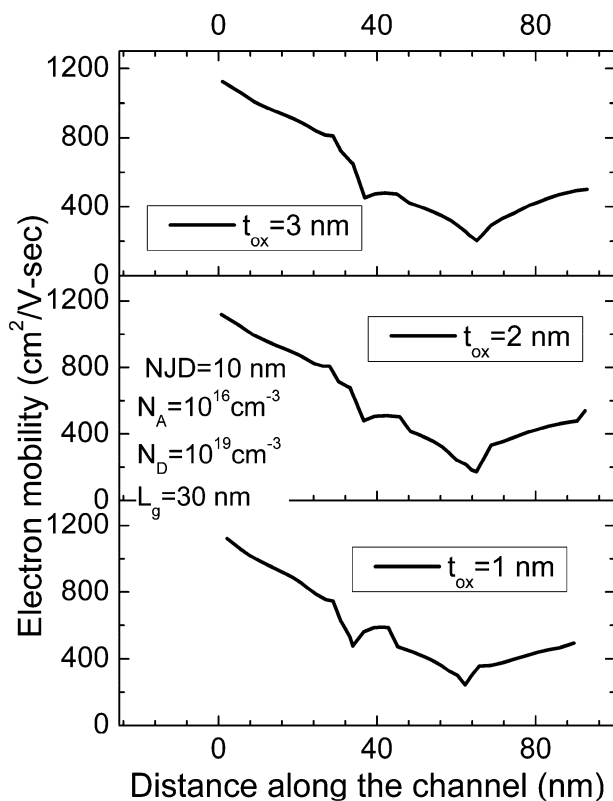


Fig. 7 Variation of electron mobility along the channel in TGRC-MOSFET at various oxide thicknesses

thickness is reduced. In the same manner, electron velocity increases at the source side nearly 19 % if oxide thickness decreases from 3 to 2 nm and increases 95 % when oxide thickness reduces from 2 to 1 nm as shown in Fig. 6.

Figure 7 shows the variation of electron mobility along the channel for TGRC-MOSFET at various oxide

thicknesses. Electron mobility is high at source side in TGRC-MOSFET when oxide thickness is 1 nm and it is further decreases for 2 and 3 nm respectively. Electron mobility is lower at the drain side of 1 nm oxide thickness and increases when t_{ox} increases from 1 to 2 nm and further from 2 to 3 nm. Four scattering mechanisms largely stimulate the mobility of charge carriers in ITO thin films: lattice scattering, ionized impurity scattering, neutral impurity scattering, and grain boundary scattering (Thilakan and Kumar 1997).

4.2 Effect of negative junction depth (NJD)

Further, different parameters of TGRC-MOSFET have been evaluated by the variation of negative junction depth (NJD). When NJD is reduced from 15 to 10 nm and from

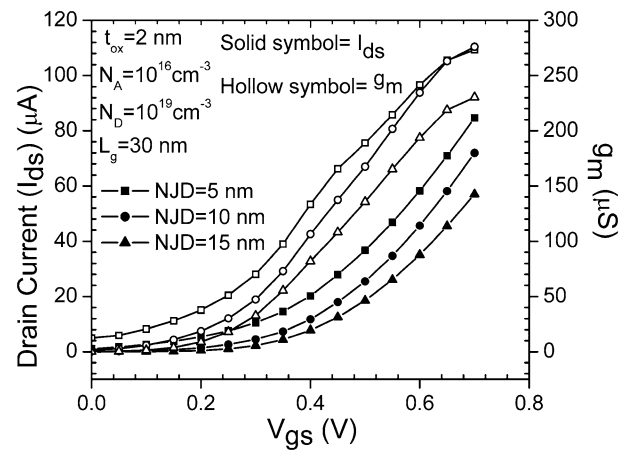


Fig. 8 Variation of drain current (I_{ds}) and transconductance (g_m) w.r.t. gate to source voltage (V_{gs}) in TGRC-MOSFET at different NJDs

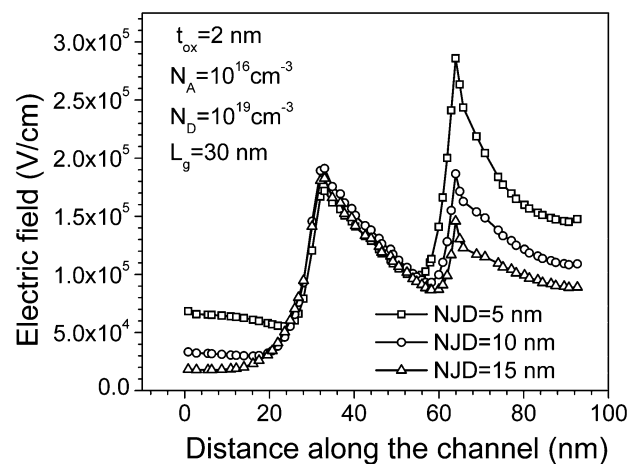


Fig. 9 Electric field variation along the channel for TGRC-MOSFET at different NJDs

10 to 5 nm, drain current is enhanced by 26 and 17.5 % while transconductance is enhanced by 20 and 2.9 % respectively as decreasing the NJD causes more number of charge carriers to surpass from drain to source and hence, there is less impact of corner effect in TGRC-MOSFET. Thus hot-carriers are reduced causing an increase in drain current and transconductance as shown in Fig. 8.

Figure 9 shows electric field variation along the channel for TGRC-MOSFET at different NJDs. When NJD is reduced from 15 to 10 nm and further from 10 nm to 5 nm, it is observed that electric field is increased at source side by 6.8 and 18.8 % respectively while it decreases at the drain side. So we can say that with a rise in NJD, the reduction in the electric field at the drain end can be interpreted as a

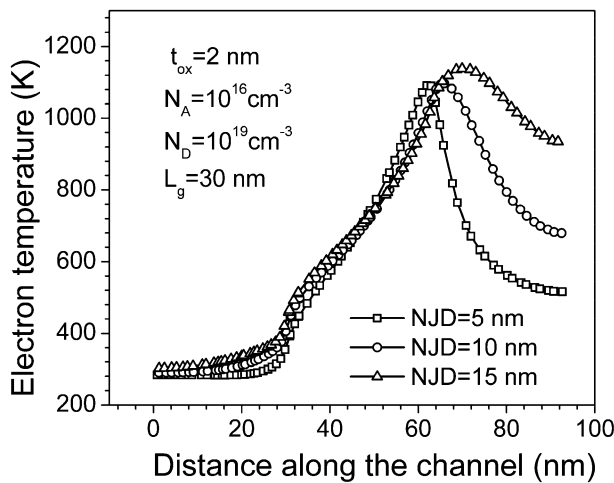


Fig. 10 Electron temperature variation along the channel for TGRC-MOSFET at different NJDs

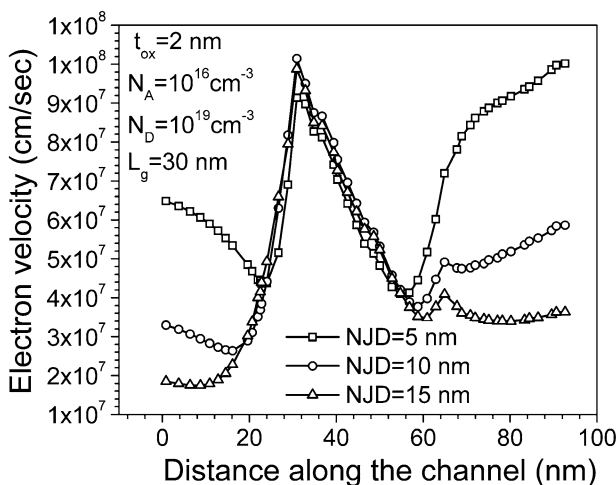


Fig. 11 Electron velocity variation along the channel for TGRC-MOSFET at different NJDs

reduction in hot carrier effects, higher device breakdown voltage and lower impact ionization (Kumar et al. 2014b).

Moreover, the variation of electron temperature has been observed at different NJDs. When NJD reduces from 15 to 10 nm and from 10 to 5 nm, then electron temperature is reduced by 27.3 and 24.1 % respectively at the drain side due to the incorporation of ITO (Kumar et al. 2015) as shown in Fig. 10. Figure 11 shows electron velocity variation along the channel for TGRC-MOSFET at different NJD. It is clearly evident from Fig. 11 that electron velocity increases on the source side as compared to drain side when NJD varies from 15 to 10 nm and further, 10 to 5 nm by 61.4 and 70.6 % respectively which reduced the hot carrier effects.

Electron mobility at the source is higher and lower at drain side when NJD is 15 nm and it reduces if the NJD reduces from 15 to 10 nm and further from 10 to 5 nm then the electron mobility enhanced by 0.7 and 2.6 % respectively as shown in Fig. 12. The main purpose of NJD is to eliminate short channel effects in the device. If NJD is tuned, then the probability of short channel effects will be less, but we need the reduction of short channel effects as well as higher mobility at source side for higher performance of the device. Here, optimized the device performance in terms of reduced SCEs and enhanced electron mobility. Higher electron mobility leads to a device for high switching applications such as logic gates.

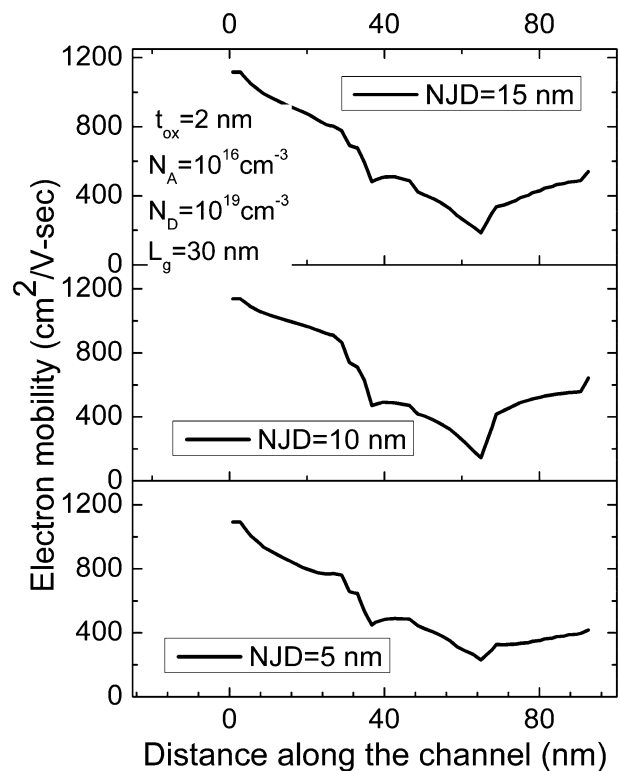


Fig. 12 Electron mobility variation along the channel for TGRC-MOSFET at different NJDs

4.3 Effect of substrate doping

Yadav and Rana (2012) reported that if the substrate (channel) doping reduces, threshold voltage also reduces due to which driving current (I_{ON}) and thus electron mobility enhances. Due to this reason, there is a need to study the effect of substrate doping on the performance of our proposed device. In this section, the efficacy of substrate doping on hot-carrier effects of TGRC-MOSFET has been studied.

Figure 13 reflects the variation of drain current and transconductance in accordance with substrate doping. Both drain current and transconductance are enhanced with the reduction of substrate doping owing to reduction in threshold voltage of device due to more recombination of charge carriers and thus impact ionization of charge carriers are reduced causing the drain current and transconductance to enhance by 57.4 and 18 % respectively when substrate doping is reduced from $1e17$ to $1e15 \text{ cm}^{-3}$ in TGRC-MOSFET. Due to this, the gate controllability is improved and hot-carrier effects are reduced.

Further, the effect of substrate doping on the electric field has been observed. Figure 14 reflects the variation of the electric field along the channel, and it is observed that when substrate doping is reduced from $1e17$ to $1e15 \text{ cm}^{-3}$, the electric field is enhanced by 2.3 % on the source side and reduced by 3.7 % at the drain side. Hence, the less electric field at the drain side reduces hot-carriers in the TGRC-MOSFET. Thus, reduction in hot-carrier effects leads to TGRC-MOSFET for high switching applications.

Figure 15 shows the variation of electron temperature along the channel length with a variation of substrate doping and it is clearly observed that the electron temperature is reduced by 6.67 % when substrate doping is reduced.

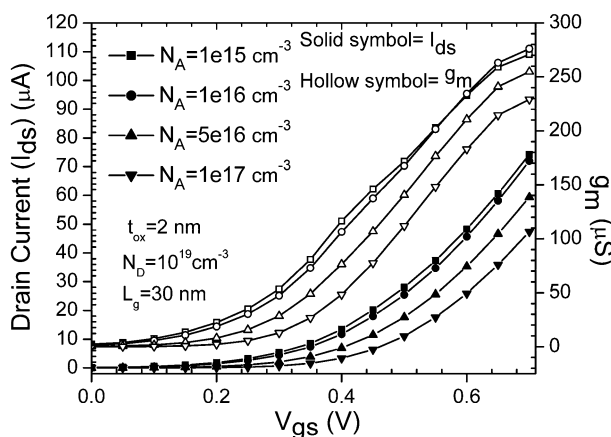


Fig. 13 Variation of drain Current (I_{ds}) and transconductance (g_m) w.r.t. gate to source voltage (V_{gs}) in TGRC-MOSFET at different substrate doping

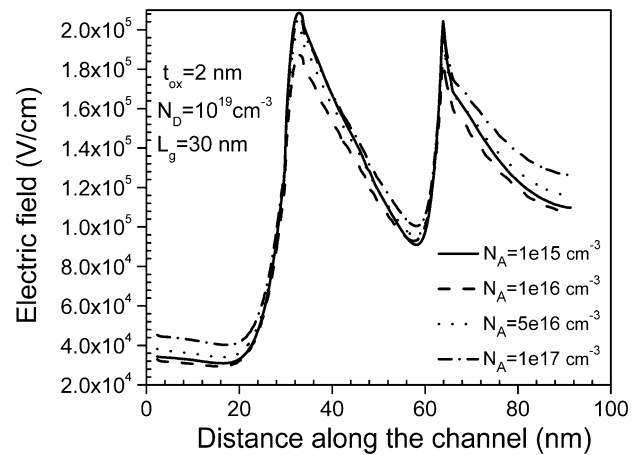


Fig. 14 Variation of electric field along the channel in TGRC-MOSFET at various substrate doping

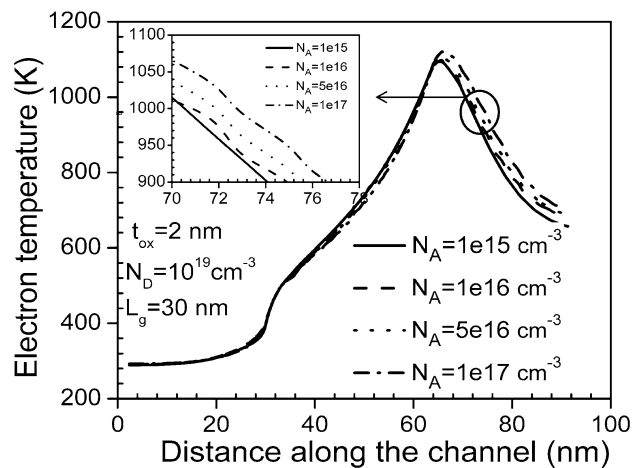


Fig. 15 Variation of electron temperature along the channel in TGRC-MOSFET at different substrate doping

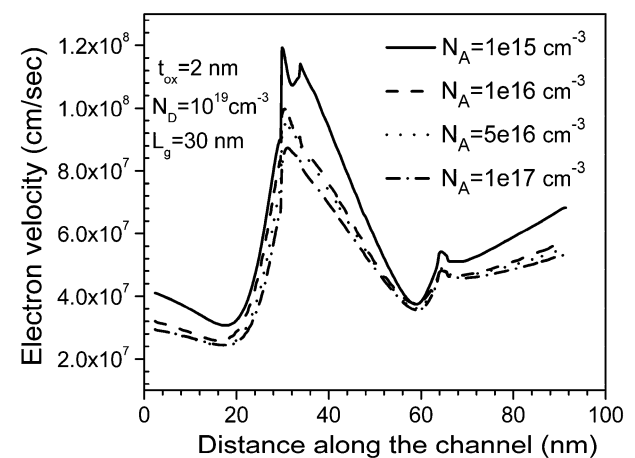


Fig. 16 Electron velocity along the channel in TGRC-MOSFET at different substrate doping

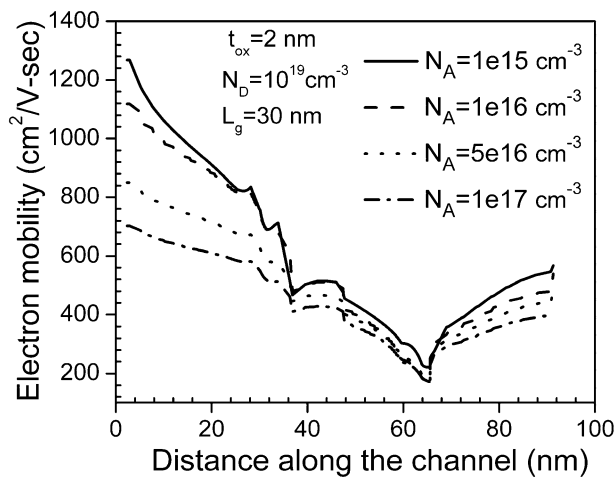


Fig. 17 Electron mobility along the channel in TGRC-MOSFET at different substrate doping

Thus, more hot-carrier immunity is observed when substrate doping reduces from $1e17$ to $1e15$ cm^{-3} . Similar kind of work has been reported by Yadav and Rana (2012). Further, more hot-carrier degradation is observed in terms of electron velocity along the channel with the variation of substrate doping. Figure 16 shows that electron velocity is 28 % higher at source side as compared to drain side and it is increased when substrate doping is reduced from $1e17$ to $1e15$ cm^{-3} . When substrate doping reduces then the velocity of charge carriers is more at the source end as compared to drain end.

Moreover, the efficacy of substrate doping variation on hot-carrier effects is observed in terms of electron mobility. Figure 17 depicts that electron mobility is higher (35.5 %) on the source side as compared to drain side and it is further enhanced when substrate doping is reduced from $1e17$ to $1e15$ cm^{-3} . Electron mobility is inversely proportional to temperature. Thus, mobility is enhanced with the reduction of substrate doping; hence, the electron temperature is decreased in TGRC-MOSFET, which leads to TGRC-MOSFET for high-temperature applications.

5 Conclusion

Effect of structured parameters, such as NJD, oxide thickness and substrate doping are investigated to study the device's hot-carrier immunity. This study proves that the TGRC-MOSFET has a high hot-carrier immunity which is influenced by geometrical structure intensely. By reducing the oxide thickness and enhancing NJD, the capability of hot-carrier-immunity can be further increased. Thus, TGRC-MOSFET is a reliable device for high-performance applications in low power CMOS technology.

Acknowledgments The authors would like to express gratitude to Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University and Department of Science and Technology (DST) to carry out this work.

References

- An X, Huang R, Zhao B, Zhang X, Wang Y (2004) Design guideline of an ultra-thin body SOI MOSFET for low-power and high-performance applications. *Semicond Sci Technol* (IOP) 19:347–350. doi:[10.1088/0268-1242/19/3/009](https://doi.org/10.1088/0268-1242/19/3/009)
- ATLAS User's Manual (2014) SILVACO Int., Santa Clara, CA
- Biswal SM, Baral B, De D, Sarkar A (2015) Analytical subthreshold modeling of dual material gate engineered nano-scale junctionless surrounding gate MOSFET considering ECPE. *Superlattices Microstruct* 82:103–112. doi:[10.1016/j.spmi.2015.02.018](https://doi.org/10.1016/j.spmi.2015.02.018)
- Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008a) TCAD assessment of gate electrode workfunction engineered recessed channel (GEWE-RC) MOSFET and its multilayered gate architecture—part I: hot-carrier-reliability evaluation. *IEEE Trans Electron Devices* 55(10):2602–2612. doi:[10.1109/TED.2008.2003085](https://doi.org/10.1109/TED.2008.2003085)
- Chaujar R, Kaur R, Saxena M, Gupta M and Gupta RS (2008b) Solution to CMOS Technology for high Performance Analog Applications: GEWE-RCMOSFET. *Advanced Optoelectronic Materials and Devices*, pp 201–205
- Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008c) Investigation of multi-layered-gate electrode Workfunction engineered recessed channel (MLGEWE-RC) sub-50 nm MOSFET: a novel design. *Int J Numer Model Electron Networks Devices Fields* 22:259–278. doi:[10.1002/jnm.699](https://doi.org/10.1002/jnm.699)
- Edwards PP, Porch A, Jones MO, Morgan DV, Perks RM (2004) Basic materials physics of transparent conducting oxides. *Dalton Trans* 19:2995–3002
- Fiegna C, Iwai H, Wada T, Saito M, Sangiorgi E, Ricco B (1994) Scaling the MOS transistor below 0.1 μm : methodology, device structure, and technology requirements. *IEEE Trans Electron Device* 41(6):941–949
- Granzner R, Polyakov VM, Schwierz F, Kittler M, Luyken RJ, Rosner W, Stadel M (2006) Simulation of nanoscale MOSFETs using modified drift-diffusion and hydrodynamic models and comparison with Monte Carlo results. *Microelectron Eng* 83:241–246. doi:[10.1016/j.mee.2005.08.003](https://doi.org/10.1016/j.mee.2005.08.003)
- Gupta N, Kumar A, Chaujar R (2014) Simulation analysis of Gate Electrode Workfunction Engineered (GEWE) Silicon Nanowire MOSFET for hot carrier reliability. *IEEE 1st International conference on Microelectronics, Circuits and Systems*, Kolkata, pp 150–153
- Gupta N, Kumar A, Chaujar R (2015a) Impact of Device Parameter Variation on RF performance of Gate Electrode Workfunction Engineered (GEWE)-Silicon Nanowire (SiNW) MOSFET. *J Comput Elect Springer* 14(3):798–810. doi:[10.1007/s10825-015-0715-z](https://doi.org/10.1007/s10825-015-0715-z)
- Gupta N, Kumar A, Chaujar R (2015b) Oxide bound impact on hot-carrier degradation for gate electrode workfunction engineered (GEWE) silicon nanowire MOSFET. *Microsyst Technol Springer*. doi:[10.1007/s00542-015-2557-9](https://doi.org/10.1007/s00542-015-2557-9)
- Hu C (2009) *Modern semiconductor device for integrated circuit*, Prentice Hall, pp 261–274
- Iwai H (2009) Roadmap for 22 nm and beyond. *Microelectron Eng* 86:1520–1528. doi:[10.1016/j.mee.2009.03.129](https://doi.org/10.1016/j.mee.2009.03.129)
- Kawazoe H, Ueda K (1999) Transparent conducting oxides based on the spinel structure. *J Am Ceram Soc* 82:3330–3336
- Kumar A, Gupta N, Chaujar R (2014) Novel design: transparent gate recessed channel (TGRC) MOSFET for improved reliability

- applications. IEEE 1st International conference on Microelectronics, Circuits and Systems", Kolkata, pp 1–5
- Kumar A, Chaujar R, Monica (2014) Thermal behavior of novel transparent gate recessed channel (TGRC) MOSFET: TCAD Analysis. Tech-Connect World Innovation Conference and Expo, Washington DC
- Kumar A, Gupta N, Chaujar R (2015) Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behaviour. *Microsyst Technol*. doi:[10.1007/s00542-015-2554-z](https://doi.org/10.1007/s00542-015-2554-z)
- Kumar A, Gupta N, Chaujar R (2016a) Power gain assessment of ITO based transparent gate recessed channel (TGRC) MOSFET for RF/wireless applications. *Superlattices Microstruct* 91:290–301. doi:[10.1016/j.spmi.2016.01.027](https://doi.org/10.1016/j.spmi.2016.01.027)
- Kumar A, Gupta N, Chaujar R (2016b) TCAD RF performance investigation of transparent gate recessed channel MOSFET. *Microelectron J* 49:36–42. doi:[10.1016/j.mejo.2015.12.007](https://doi.org/10.1016/j.mejo.2015.12.007)
- Minami T (2005) Transparent conducting oxide semiconductors for transparent electrodes. *Semicond Sci Technol* 20:S35–S44. doi:[10.1088/0268-1242/20/4/004](https://doi.org/10.1088/0268-1242/20/4/004)
- Moreno E, Villada MP, Ruiz FG, Roldán JB, Marin EG (2015) A new explicit and analytical model for square Gate-All-Around MOS-FETs with rounded corners. *Solid-State Electronics* 111:180–187. doi:[10.1016/j.sse.2015.06.004](https://doi.org/10.1016/j.sse.2015.06.004)
- Pammi SVN, Jung HJ, Yoon SG (2011) Low-temperature nanocluster deposition (NCD) for improvement of the structural, electrical, and optical properties of ITO thin films. *IEEE Trans Nanotechnol* 10:1059–1065. doi:[10.1109/TNANO.2010.2103568](https://doi.org/10.1109/TNANO.2010.2103568)
- Ren H, Hao Y (2002) The influence of geometric structure on the hot-carrier-effect immunity for deep-sub-micron grooved gate PMOSFET. *Solid State Electron* 46:665–673. doi:[10.1016/S0038-1101\(01\)00338-0](https://doi.org/10.1016/S0038-1101(01)00338-0)
- Thilakan L, Kumar J (1997) Studies on the preferred orientation changes and its influenced properties on ITO thin films. *Vacuum* 48:463. doi:[10.1016/S0042-207X\(96\)00309-0](https://doi.org/10.1016/S0042-207X(96)00309-0)
- Yadav VK, Rana AK (2012) Impact of channel doping on DG-MOS-FET parameters in nano regime-TCAD simulation. *Int J Comp Appl* 37:36–41. ISSN: 0975–8887

E-learning on Clouds: A profitable and Efficient Business Model

Rajni Jindal, Alka Singhal,
Delhi Technological University

Abstract

The paper proposes that E-learning on clouds can be a profitable business Model as cloud eliminates the upfront investment required to setup involving purchase of software and hardware, it also eliminates the licensing and renewal cost. In a nutshell, provider does not require any initial infrastructure but has to pay for the services they use and all the hard work is done by the third party. The paper also compares the Net Profit Value (NPV) of both the cost models (On Premise Vs On Cloud) by taking case studies to show that adopting clouds can be profitable.

Keywords

E-learning, Cloud Computing, Cost Model, Net Profit value

1. Introduction

“An Investment in knowledge pays the best return”- Benjamin Franklin

E-learning solutions facilitate the delivery of the right information and skills to the right people at the right time.

The globalization of the economy, shortage of the skilled worker, free agent mentality, new flexible work situation and numerous factors have created problems not easily solved by traditional education, spurring initial growth of E-learning industry [1].

1.1 Evolution of E-learning technology

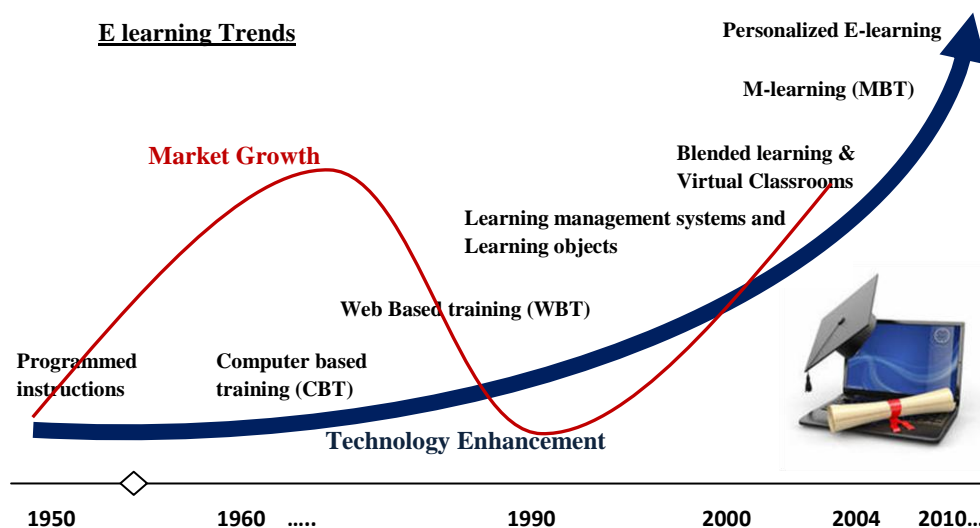


Figure 1: Evolution of E-learning systems

I. E-learning evolved in the form of computer based training (CBT) which considerably relied on CD-ROM. It did not face the bandwidth problem and was able to deliver rich content without clogging up a network while downloading audio and video files. It had large storage capacity but was static in nature. If it was distributed, it was difficult to update it[1].

II. In 1990s, E-learning stepped into the web and it was called Web based training (WBT). Internet provided Educators and trainers with a platform which provided the facility of distance learning and online training. The ability of the Internet to serve as an efficient, economical and ubiquitous delivery system for instruction has facilitated the development of tens of thousands of online courses and hundreds of online degree and certificate programs. Studies have confirmed that learning outcomes of online students are equivalent and in some cases may be superior to those of face-to-face students. In terms of technology, online courses were comprised of hyperlinked web pages through which learner advance by clicking. But collaboration of learner with the instructor was limited to E-mails[2].

Technology	Era (start)	Content Quality	Semantics	Update	Content Distribution
Computer Based training	1960	Rich but static in nature	No	Very Difficult	Easy and no network clogging
Web Based training	1990	Less multimedia content and Dynamic in nature	No	Fast	Ubiquitous but network clogging
Collaborative Learning (LMS)	2000	Rich and Dynamic in nature	No	Fast	Ubiquitous and improved network performance
Blended Learning	2002	Rich and Dynamic in nature	Yes	Fast	High speed internet
Mobile Learning	2007	Rich and Dynamic in nature	No	Fast	Mobile internet
Personalized Learning	2010	Rich and Dynamic in nature	Yes	Fast	High speed and mobile internet

Figure 2: Comparative Study between Different Era of E-learning system

III. In 2000, it was time of Web collaborating tools and learning Management system. LMS were designed and developed to track, manage and tie together entire enterprise program of learning. A number of vendors have begun to bundle all of these tools together into a package to create integrated learning system that can be either hosted by the customer or the vendor to address organizational learning needs. May vendors like SCORM, Moodle provides standard API enclosing all the facilities of LMS including content management, delivery etc[2].

IV. Finally it was technology which took E-learning to a new Era. E-learning was now blended learning, mobile learning and then became personalized learning. Blended learning uses combination of various methodologies to included face to face and live learning (Valiathan 2002). Mobile learning promotes delivery of content on mobile devices increases capabilities of learner to access the content anytime and everywhere (Wagner 2007). Personalized learning facilitates and supports individual learning and was inspired by learner's requirement [1].

2. Traditional E-learning and E-learning on clouds

In Traditional computing one has to buy the assets and pay for them while in Cloud computing one buys services and pay for use. Traditional computing requires technical setup and expanded staff. It is single tenant and requires lengthy deployment whereas in cloud computing, it is more administrative work required and less deployment time required and fast ROI.

E-learning system requires a big setup including many hardware and software resources. There are many educational institutions that cannot afford such investments and so cloud computing is the best solution.

There are many challenges standing in front of E-learning to hinder its progress. Some of them are high concurrency demand, optimizing resource allocation, security, sustainable services and reliable infrastructure [3][4].

Cloud computing can prove to withstand all of these challenges as it provides multitenant model which can provision elastic, scalable services.

3. Comparative study of Cost Model for on-Premise and cloud based E-learning system

3.1 Cost Model for on-Premise E-learning system

To compare the cost of on-Premise Vs on cloud, we first calculate the net profit value (NPV) in the case of purchase i.e. On-premise E-learning system.

As On-premise E-learning system includes cost of setting up data centre and running it. So as follows is an estimate value of Net profit value of setting up own data centre as shown fig3.

For calculating the Net Profit Value we use the following formulae stated by R.W.Johnson and W.G.Lewellen[5].

They have proposed that the decision of lease and buy can be analyzed by comparing the net profit value of the object with a span of time and the higher the NPV is the rational choice [5]. As setting up own data centre involves buying all the assets, therefore we use the formula of buying in case of On-Premise cost.

The proposed formula referred [5] was as below:

$$\text{Net Profit Value (purchase)} = \frac{\sum_{T=1}^Y (Pt - Lt) - Tc(Pt - Lt - Dt)}{(1+k)^T} + \frac{S - Tg(S-B)}{(1+k)^T} - \text{Cost}_{\text{purchase}}$$

Pt=Cash revenue expected from the use of the asset at year T

Lt=Pre-tax cash required to operate the asset at year T

Tc=Corporate Tax

Dt=Depreciation charge for year T

k=cost of capital

Cost_{purchase} =Cash purchase price of the asset

S=expected salvage value of the asset at the end of its life

Tg =tax rate for gain or loss on disposal of the asset

B=expected book value of the asset at the end of useful life

We moderate the formula for E-learning system because there is no salvage value after its retirement because only small market exists for used CPU equipment [6]. In the referred paper [6], they have assumed no expected cost revenue and no depreciation charges but in our case we have to consider the cost revenue paid by the learners in response to the E-learning services provided to them. Therefore, it is required that the cash revenue obtained by delivering the E-learning services should be considered and also the tax paid on the basis of net profit. And depreciation cost also plays an important role in making cloud a strong candidate of business because a major regular maintenance cost is spent over updation of the technology (hardware and software) in case of On-premise setup. In addition, Cost of operating data centre cannot be ignored as it involves Cost_{power} and Cost_{cooling}.

$$(\text{Modified}) \text{ Net Profit Value (NPV)} = \frac{\sum_{T=1}^Y (Pt - Lt) - Tc(Pt - Lt - Dt)}{(1+k)^T} - \text{Cost}_{\text{purchase}} - \text{Cost}_{\text{operation}}$$

Pt=Cash revenue expected from the use of the asset at year T

Lt=Pre-tax cash required to operate the asset at year T (Licensing cost)

Tc=Corporate Tax

D_t =Depreciation charge for year T

k =cost of capital

$Cost_{purchase}$ =Cost price of the asset (will be discussed in detail in section 3.3)

$Cost_{operation}$ = Cost of operating data centre as well as delivering services (will be discussed in detail in section 3.3)

3.2 Cost Model for cloud based E-learning system

Cost of using cloud services depends on the type of services being utilized and further storage, communication resources it will provision and consume.

Cloud providers charge applications for the use of their resources according to a fee structure. In addition to support on-demand computing, clouds uses virtualization technologies, which enable applications to setup and deploy a custom virtual environment suitable for them as shown in fig3.

Net profit value can be easily calculated by the net profit from cash revenue by providing the E-learning service after removing the cost of service to the cloud provider [6].

$$\text{Net Profit Value (NPV)} = \frac{\sum_{T=1}^Y \frac{P_t - (T_c * P_t)}{(1+k)^T}} - Cost_{service}$$

P_t =Cash revenue expected from the use of the asset at year T

T_c =Corporate Tax

k =cost of capital

$Cost_{service}$ =Cost of service of the cloud provider.

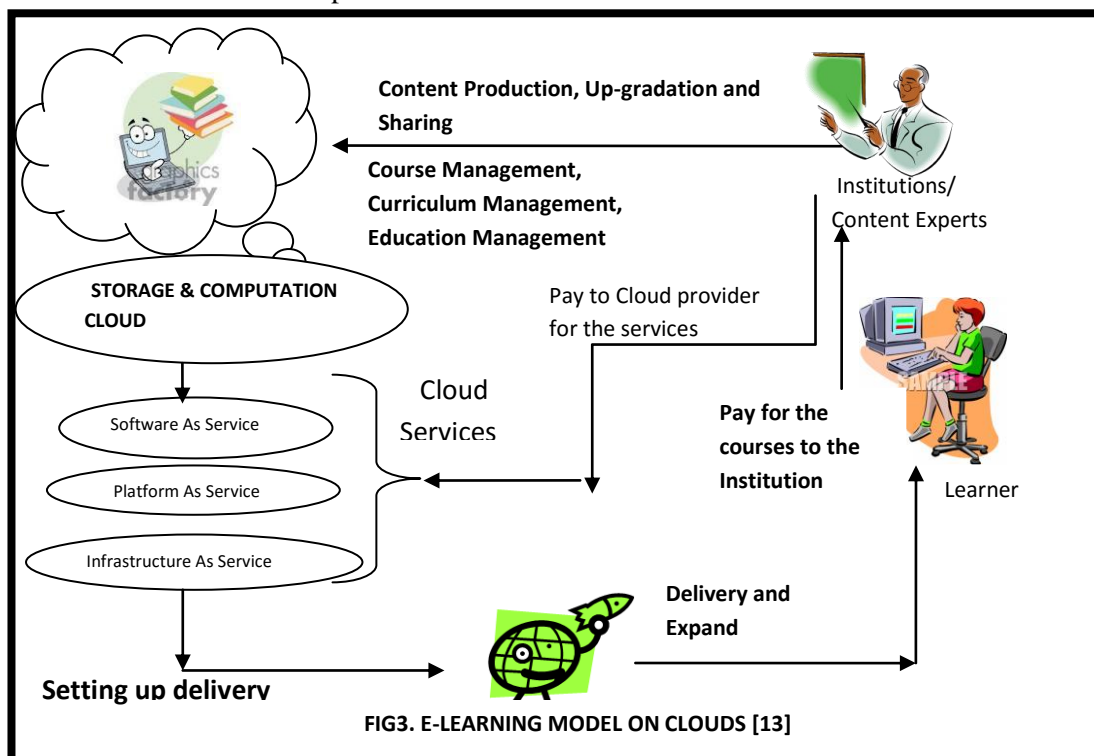


FIG3. E-LEARNING MODEL ON CLOUDS [13]

3.3 Comparing the Net profit Value of On-Premise versus Cloud Based Service

If we compare the formula, each equation is composed of two parts, First part calculates cash revenue earning by providing services excluding corporate tax and second part is the cash investment for the asset and its operation or the service.

As First term is approximately same in both the cases, therefore the comparison lies in the second part which deals with investment of the provider.

In On-premise setup, second part is composed of $Cost_{purchase}$ and $Cost_{operation}$. [10]

$Cost_{purchase}$ includes the cash investment on the asset. It is composed of Cost of Server, Cost of Software, Cost of Network and Cost of space [9]. The parameters on which these costs are measured are as follows [9]

Parameters for Cost of Server are

Parameter	
N_s	Number of physical servers in resource pool
VI_{ps}	Cost per physical server of same configuration

Parameters for Cost of Software are

Parameter	
V_{is}	Unit Price software
S_s	Subscription factor
N_{slic}	Number of software license

Parameters for Cost of Network are

Parameter	
N_{switch}	Number of new network switches per year
$SNIC$	Number of NIC per virtualized server
$PNIC$	Number of ports per NIC
P_s	Price per switch
N_{port}	Port number of a network switch

Parameters for Cost of Space are

Parameter	
A_p	Cost per square foot to build Data centre
RSF	Square feet per rack
$RSPACE$	Percent of space taken by racks in all(<1)
W_{Server}	Weight of a physical server
W_{rack}	Weight of a rack

$Cost_{operation}$ includes the cash investment in operating the data centre. It is composed of Cost of power, Cost of Cooling and Cost of administration. The parameters for measuring these costs are as follows:

Parameters for Power cost

Parameter	
Srp	Sum of the power rating of working servers
Es	Price per hour of 1kw of electricity
Nrack	Number of racks in working

Parameters for Cooling cost

Parameter	
L	Cooling Load Factor - amount of power consumed by the cooling equipment for 1W of heat dissipated
P	Airflow Redundancy Constant – redundant airflow required to cool data center

In Cloud based services, second part is composed of $Cost_{service}$ i.e. Cost of service of the cloud provider. It can also be customized according to the requirements. It is scalable, dynamic in nature.

Though cash investment on the asset is one time spend but the depreciation cost and operating cost is regular in nature whereas in using cloud based services there is no operating and depreciation cost and also no initial investment on licensing and purchase is required.

To calculate the cost of Buying (A), a case study of real-world customer environment was taken [11]. These companies were using a myriad of integration solutions to manage integration communities that used EDI and many other formats. This cost analysis was based on 100 connections (trading partners or customers) and 10 million transactions. Total Cost of On-Premise Integration: Nearly \$1 million per year. In comparison, to calculate the cost of service of the cloud provider, we have taken a rough estimate of 2000 learners taking 6 courses from SCORM cloud. Using link <http://scorm.com/scorm-solved/scorm-cloud-pricing> [12] and the estimated cost was \$1900 per connection and so annually \$22800

Therefore it can be estimated that cost of Cloud based service is better than cost of On-premise service. Cloud can make E-learning a profitable Business Model.

Conclusion and Future Work

Therefore, the paper states that E-learning on clouds can be a profitable business Model as cloud eliminates the upfront investment required to setup involving purchase of software and Hardware it also eliminates the licensing and renewal cost. In a nutshell, provider does not require any initial infrastructure but has to pay for the services they use and all the hard work is done by the third party.

But still there are lots of compromises one has to make to adopt cloud services like security, lack of control on data, customization and redundancy issues.

So the best way can be devised to find a balance between both delivery models and adapt according to the business model which can be obtained using Hybrid cloud.

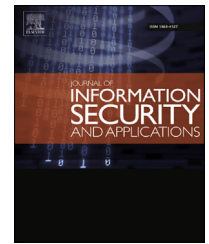


References

- [1] Brian W. Rittenburg, Ginger C. Spickler, "E-learning book- the engine of the knowledge economy" July 2000.
- [2] Michael H. Breitner, Gabriela Hoppe. "A Glimpse at Business Models and Evaluation Approaches for E-Learning" Springer, E-Learning 2005, pp 179-193
- [3] Laisheng, Wang Zhengxia, "Cloud Computing: a New Business Paradigm for E-learning", Third International Conference on Measuring Technology and Mechatronics Automation 2011
- [4] Laisheng, Wang Zhengxia, "Cloud Computing: a New Business Paradigm for E-learning", presented at Third International Conference on Measuring Technology and Mechatronics Automation 2010
- [5] R.W. Johnson, W.G. Lewellen, "Analysis of Lease or Buy Decision", J Finance, vol 27, no4, 1972 pp 815-823
- [6] Edward Walker, "The Real Cost of a CPU hour", IEEE Computer Society, 2009
- [7] Ewa Deelman, Gurmeet Singh, Miron Livny, Bruce Berriman, John Good, "The Cost of Doing Science on the cloud: The Montage Example", Proceeding of Super Computing 2008, Austin, Texas
- [8] "On-premise Vs Cloud based Solutions", GFI white paper 2010.
- [9] Xinhui Li, Ying Li, Tiancheng Liu, Jie Qiu, Fengchun Wang, "The Method and Tool of Cost Analysis for Cloud Computing" published in IEEE International Conference on Cloud Computing 2009
- [10] Chandrakant D. Patel, Amip J. Shah1, "Cost Model for Planning, Development and Operation of a Data Center" Internet Systems and Storage Laboratory HP Laboratories Palo Alto HPL-2005-107(R.1) June 9, 2005
- [11] <http://scorm.com/scorm-solved/scorm-cloud-pricing>
- [12] <http://liaison.com/resource-center/white-papers/companies-look-to-the-cloud-for-data-integration>
- [13] Rajni Jindal, Alka Singhal "Social Networking based E-Learning System on Clouds" International Journal of Computer Applications, March 2012, pp.45-48

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jisa

Enabling information recovery with ownership using robust multiple watermarks

Vidhi Khanduja ^{a,*}, Shampa Chakraverty ^a, Om Prakash Verma ^b

^a Department of Computer Engineering, Netaji Subhas Institute of Technology, Delhi, India

^b Department of Computer Science and Engineering, Delhi Technological University, India

ARTICLE INFO

Article history:
Available online

Keywords:
Information recovery
Database watermarking
Right protection
Robustness
Tamper detection

ABSTRACT

With the increasing use of databases, there is an abundant opportunity to investigate new watermarking techniques that cater to the requirements for emerging applications. A major challenge that needs to be tackled is to recover crucial information that may be lost accidentally or due to malicious attacks on a database that represents asset and needs protection. In this paper, we elucidate a scheme for robust watermarking with multiple watermarks that resolve the twin issues of ownership and recovery of information in case of data loss. To resolve ownership conflicts watermark is prepared securely and then embedded into the secretly selected positions of the database. Other watermark encapsulates granular information on user-specified crucial attributes in a manner such that the perturbed or lost data can be regenerated conveniently later. Theoretical analysis proves that the probability of identifying target locations, False hit rate and False miss rate is negligible. We have experimentally verified that the proposed technique is robust enough to extract the watermark accurately even after 100% tuple addition or alteration and after 98% tuple deletion. Experiments on information recovery reveal that successful regeneration of tampered/lost data improves with the increase in the number of candidate attributes for embedding the watermark.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

With the growth of Information and Communication Technology in the progress of mankind, the information has taken center-stage. A major proportion of the internet content is dynamically generated from databases. This has driven the capabilities, sizes and performance of databases to grow in exponential magnitudes. In this scenario where end users are

demanding more and more information to be available on the net, data providers are burdened to supply accurate data and at the same time ensure its security. A major threat faced is that unauthorized and illicit copies of true data can be easily generated and distributed using the same enabling technologies. To counter such attacks, legal as well as technological solutions are being devised to assert data ownership. Watermarking is one such widely accepted technological measure to protect the data (Kerr et al., 2003).

* Corresponding author. Department of Computer Engineering, Netaji Subhas Institute of Technology, Delhi, India. Tel.: +919999695797. E-mail address: vidhikhanduja9@gmail.com (V. Khanduja).

<http://dx.doi.org/10.1016/j.jisa.2016.03.005>

2214-2126/© 2016 Elsevier Ltd. All rights reserved.

High risk databases such as databases from medicine, weather, transportation, automobile, military, etc., contain critical attributes. The information conveyed by critical attributes needs to be preserved. Additionally, these high-risk databases are subject to various malicious attacks resulting in threat to their ownership (Agrawal et al., 2003).

This motivated us to propose a technique that not only recovers information irrespective of a number of alterations occurred but also resolves ownership issues. We follow an information-centric data recovery technique to preserve the information. The proposed robust watermarking technique prepares dual watermarks and securely conceals them within the database. One watermark contains the ownership details while the second one holds information to be preserved.

This paper proceeds with related work in Section 2. Section 3 elucidates the proposed watermarking technique. Section 4 presents the Security Analysis of the proposed scheme. Section 5 presents experimental results and analysis to verify robustness and information recovery followed by comparative evaluation of our work with a prior work. We conclude the proposed technique in Section 6.

2. Related work

Literature is rife with interesting works in the domain of robust watermarking for digital databases. R. Agrawal, Peter J. Haas, J. Kiernan identified the need of database watermarking in 2002 (Agrawal et al., 2003). They proposed that database relations can be watermarked in some algorithmically selected attributes out of several candidates attributes in a tuple. The technique suffered from certain drawbacks. The technique is primary key dependent and therefore is not suitable for databases that does not contain primary key. Also, technique does not provide mechanism for multi-bit watermark. Their technique generates limited potential locations that can be used to hide watermark bits without being subjected to removal or destruction. Several techniques proposed in literature enhanced the work of Agrawal et al. (2003) by embedding multi-bit watermark in selected LSBs (Ali and Mahdi, 2011; Cui et al., 2007; Farfoura et al., 2012).

Xinchun, Xiaolin and Gang have proposed weighted algorithm (Cui et al., 2007). The technique assigns the weight to attributes according to their significance. This result in the increase in the chances of selection of high rank significant attributes as compared to less important ones. The mark bit-string comes from meaningful character of a database owner such as names. Robustness against subset deletion, addition and alteration attacks is discussed. However, primary key dependent embedding process is based on work proposed in Agrawal et al. (2003). In Ali and Mahdi (2011), Dr. Yossra H. Ali and Bashar Saadoon Mahdi proposed a technique that uses threshold generator based on simple combination of odd number of register. However, this technique has same limitations as in Agrawal et al. (2003).

In Farfoura et al. (2012), they proposed time-stamping based protocol to resolve additive attacks. The embedding

process was reversible so that original values can be regained after extraction. Reversibility allows one to recover the original data completely from the watermarked database after authenticating with a time-stamp protocol. However, technique can be applied where data authentication and original content recovery were required at the same time. Once the original data is recovered, the database loses ownership protection. Recently, semi-blind reversible technique has been proposed in Ifthikar et al. (2015). In this technique, the knowledge of mutual information for every candidate feature is employed to create watermark. Genetic algorithm is used to resolve the constraint optimized problem of optimal watermark encoding to ensure high data quality.

Most of the prior works on robust watermarking assume the watermark to be a random bit stream, which is otherwise meaningless (Agrawal et al., 2003; Ali and Mahdi, 2011; Cui et al., 2007; Shehab et al., 2008; Sion et al., 2004). Little other notorious work based on fragile watermarking schemes exists in literature (Camara et al., 2014; Guo et al., 2006; Guo, 2011; Khan and Husain, 2013; Li et al., 2004), resolving tamper detection and localization of perturbation issues. However, all these works do not take into account the Data/Information recovery. The application of watermarking in data recovery has received scant attention so far.

To our knowledge, the only work that reports a watermarking scheme for recovering data is by Khataeimaragheh and Rashidi (2010). In this work, authors have proposed a fragile watermarking technique that can detect and correct distortions in RDBs by embedding watermarks created from each attribute value, thereby recovering true data. There are certain serious shortcomings in their approach. Firstly, it can only be used to detect and recover altered data and preclude proper recovery from deletions. Secondly, the probability of accurately detecting, localizing and hence rectifying errors reduces drastically when the number of errors exceeds two.

In contrast to prior work, our proposed approach deciphers data in terms of information it represents and helps recover it from altered as well as deleted data. We utilize the watermark as information carrier to address the need to recover lost information due to frequent attacks. We propose a scheme for robust watermarking with multiple watermarks that resolve the twin issues of ownership along with recovery of information.

3. Architecture of the proposed scheme

Information is framed in the form of a relational database \mathcal{R} . The database \mathcal{R} is built upon a relational model. We define relational schema as $\mathcal{R}(\mathcal{K}_p, \mathcal{A})$ comprising N_t tuples, i.e. $N_t = |\mathcal{R}|$. \mathcal{K}_p is the primary key attribute and \mathcal{A} is a set of attributes such that $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_{N_a}\}$, $|\mathcal{A}| = N_a$. \mathcal{A} represents a set of attributes in \mathcal{R} excluding \mathcal{K}_p .

Fig. 1 depicts block diagram of the proposed watermarking scheme showing the main components. The proposed scheme consists of four phases: Watermark Preparation, Watermark Insertion, Watermark Extraction, and Decision

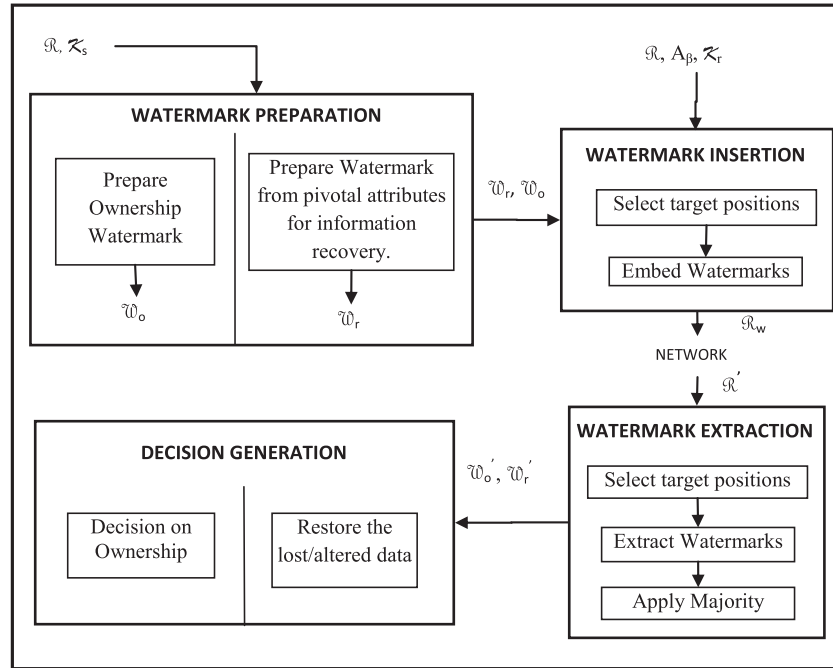


Fig. 1 – Block diagram of the proposed watermarking model.

Generation. In the watermark preparation phase, we prepare dual watermarks; one watermark for ownership proof and another for information recovery.

The ownership watermark \mathcal{W}_o is prepared by taking hash of owner's identity and \mathcal{K}_s . The information recovery watermark \mathcal{W}_r is prepared by extracting the information in form of clusters from the pivotal attributes that need to be preserved. The watermark insertion phase embeds the watermarks \mathcal{W}_o and \mathcal{W}_r into securely selected target positions using secret parameters \mathcal{K}_s , \mathcal{K}_r , and A_β and converts a data set \mathcal{R} into watermarked data set \mathcal{R}_w .

In case of any dispute over ownership or data loss during transmission, watermark extraction phase extracts watermark bits from the respective marked locations and recovers the watermarks \mathcal{W}'_o and \mathcal{W}'_r after applying majority voting. To resolve conflicts over ownership, the regenerated ownership watermark \mathcal{W}_o and extracted ownership watermark \mathcal{W}'_o are compared in Decision Generation phase. Decision generation phase takes another decision whether to recover the data losses during transmission based on comparison of information recovery watermark extracted \mathcal{W}'_r with regenerated watermark \mathcal{W}_r . We now discuss each of these phases in detail.

3.1. Watermark preparation

This step prepares watermarks that are secretly concealed within a database \mathcal{R} . We create two different watermarks — one for ownership proof \mathcal{W}_o and other for information

recovery \mathcal{W}_r . We now describe each one of them subsequently.

3.1.1. Ownership watermark \mathcal{W}_o

\mathcal{W}_o is prepared using identity of the owner. Firstly, the watermark (w) is selected by owner of the database representing copyright and/or the database information. Final watermark \mathcal{W}_o is prepared by applying a hash function to the selected watermarking text (w) concatenated with the key \mathcal{K}_s . Following Kerckhoff's principle (Kerckhoffs, 1883) that says a cryptosystem must be secure even if everything about the system is publically known except the key, we use \mathcal{K}_s . In such cases, if direct watermark, i.e. the owner's identity is taken as watermark, then an attacker, say Mallory, can guess a watermark knowing the owner of a database. Concatenating w with \mathcal{K}_s makes the watermark secure against malicious attacks. \mathcal{W}_o is given in Eq. (1) as

$$\mathcal{W}_o = \text{Hash}(w \parallel \mathcal{K}_s) \quad (1)$$

where \mathcal{K}_s is the secret key selected by the owner of a database, \parallel is concatenation operator, and $\text{Hash}(\cdot)$ is a cryptographic hash function (Schneier, 1996). Hash provides the security against Mallory as she cannot find the watermark unless she has knowledge of w and \mathcal{K}_s .

Furthermore, using hash provides owner with the liberty to select his own watermark. Owner can select large files such as an image, audio or video as a watermark. Keeping in mind the restriction on the length of a watermark to be embedded

(as watermarks are embedded repeatedly for robustness), we consider the hash to represent the shortened reference to original selected watermark.

Many cryptographic hash algorithms exist in literature, e.g. RIPE-MD, MD5, SHA-2, SNEFRU, etc. We employ SHA-2 algorithm as a cryptographic hash function that yields 256-bit (64-hex characters) hash value owing to its improved resilience against attacks (Wikipedia, 2015). Final watermark is a set of N_w bits $\mathcal{W} = b_{N_w} \dots b_1$.

3.1.2. Information recovery watermark \mathcal{W}_i

\mathcal{W}_i represents the vital information in \mathcal{R} that is protected against tamperedness or deletion. For larger databases, sometimes we need to protect the relevant information conveyed by it instead of an entire database. To achieve this, we propose an information-centric data recovery watermarking technique. Fig. 2 depicts the steps followed to extract the information from a database. We now explain each of these steps subsequently.

A. Select pivotal attributes for clustering: The owner first decides a set of pivotal attributes for information preservation. We define a set of pivotal attributes A_p as a set of attributes whose information is preserved. The granular information of these crucial attributes is captured and encapsulated into a watermark.

$$A_p = \{A_p \subset A \mid A_p \neq \emptyset\} \quad (2)$$

B. Perform k-means clustering: We aim at extracting the relevant information that consumes less storage space. The information conveyed by all the tuples under an attribute is the degree of discernability in the values. Thus, we implement k-means clustering algorithm to categorize them into clusters.

K-means clustering classifies n observations (x_1, x_2, \dots, x_n) into k clusters/partitions $S = \{S_1, S_2, \dots, S_k\}$, such that each observation belongs to a cluster with nearest centroid (Hartigan and Wong, 1979). Each element of S is a set containing observations having same centroid. A set $C = \{C_1, C_2, \dots, C_k\}$ holds the centroids of k clusters formed.

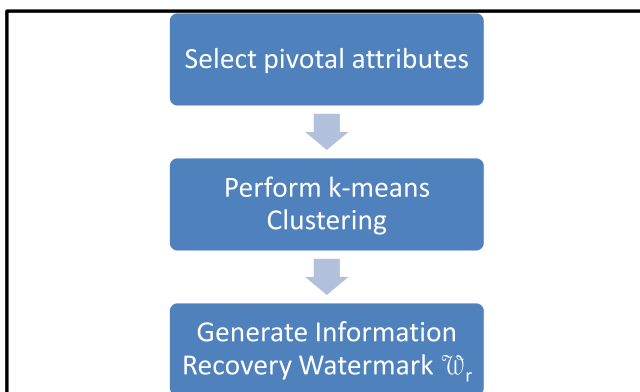


Fig. 2 – Block diagram showing preparation of Information recovery Watermark.

The onus of specifying the number of clusters rests on the owner. Thus, she has complete control over the granularity of the information that is encoded into the watermark. The optimal choice of k will strike a balance between the degree of compression and the accuracy of recovered data. Fig. 3 shows the cluster formation on the two selected pivotal attributes of a database (National Geochemical Survey Database of the US). We chose the pH-value of water content of the sample and Grain size of soil sample as the pivotal attributes from the geochemical database (National Geochemical Survey Database of the US). Details of the database used are given in Section 5 of experimental results and analyses. By applying k means clustering on these two attributes, we are able to extract relevant information with two, four, six and eight clusters. This shows that as the cluster size increases, the information represented by each cluster becomes more accurate. We have shown the results for two attributes; the same principles apply to any number of attributes.

C. Generate information recovery watermark $\mathcal{W}_i(t)$: The cluster set S is mapped to a smaller set of cluster-Ids, which are concealed as information recovery watermarks $\mathcal{W}_i(t)$ (Khanduja et al., 2014). Each set S_j is mapped to unique Cluster-Id I_j using mapping function

$$\psi: S \rightarrow I \quad (3)$$

where S is a set of clusters and I is the designated cluster-Ids. Thus, for every tuple, recovery information of its pivotal attribute is its designated cluster-Id, referred as $\mathcal{W}_i(t)$. The cluster-Ids, their representative data points C , i.e. centroid of each cluster, and minimum and maximum difference between the pivotal attributes that belong to same cluster and that cluster's centroid are saved with the owner or with a trusted entity. This information is restored later to detect tampering and recover the information.

3.2. Watermark insertion

After creating the watermarks, we embed them into the fractional parts of securely selected numeric attributes to minimize the distortions. The pseudo-code for embedding the watermarks is inscribed in the Pseudo 1 *Embed_Watermarks()*. The owner specifies the attribute set A_β , comprising of attributes of type float that can tolerate small amount of perturbations without compromising usability constraints. Usability constraints are pre-defined for every attribute by owner of a database depending upon its application.

To embed two different watermarks into the database, we divide A_β into two subsets: A_{β_0} represents candidate attributes for \mathcal{W}_0 , and A_{β_r} represents a set of candidate attributes for embedding \mathcal{W}_i such that $A_\beta = A_{\beta_0} \cup A_{\beta_r}$ and $A_{\beta_0} \cap A_{\beta_r} = \emptyset$. We next calculate derived primary key $\mathcal{K}_d(t)$ for all the tuples using Eq. (4), which aids in selection of target locations for watermarking (Khanduja et al., 2015)

$$\mathcal{K}_d(t) = \text{MSB}(N(t \cdot \mathcal{A}_1)) \parallel \text{MSB}(N(t \cdot \mathcal{A}_2)) \quad (4)$$

where $N(t \cdot \mathcal{A}_1)$ and $N(t \cdot \mathcal{A}_2)$ represents the normalized attribute values and \parallel is the concatenation operator. These attributes

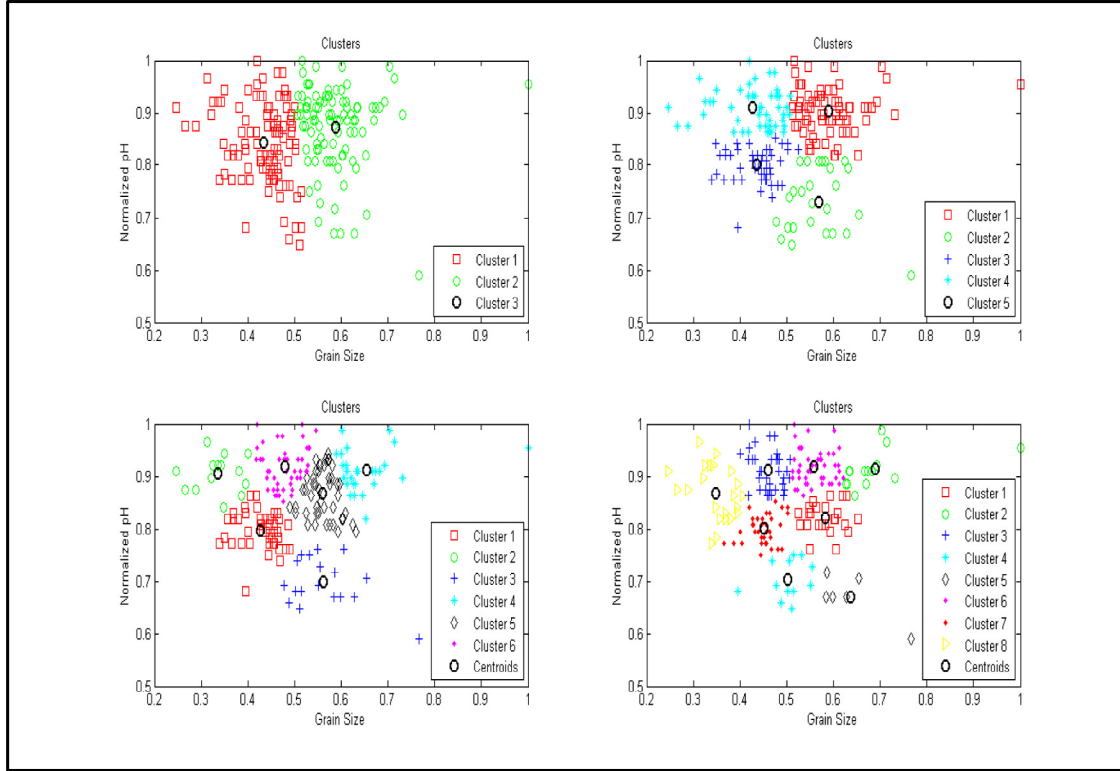


Fig. 3 – Formation of clusters by taking two pivotal attributes pH value and grain size.

are selected by the owner such that they contain crucial information that cannot be deleted from a database. $\mathcal{K}_d(t)$ replaces the primary key, thereby making the technique applicable to databases that does not have a primary key.

Normalization is carried out using a common divider that is applied to all items in an attribute. This helps to defeat linear transformation attack. In this attack, numeric values are linearly transformed, e.g. attacker may convert the data to different unit of measurement (like Fahrenheit to Celsius). In such a case, entire attribute get changed to different value, and hence watermark is removed. Thus, normalizing the attribute values and then using them defeats this attack.

$MSB(N(t \cdot A_1))$ and $MSB(N(t \cdot A_2))$ are the most significant bits (MSBs) of two of the crucial attributes A_1 and A_2 in relation \mathcal{R} . The significance of using MSBs is that in case of any perturbations in these attributes, only LSBs will change, abstaining MSBs. MSB space is assumed to be a domain where minor changes on the data items have a minimal impact on the MSB labels. Any further attempt to change MSBs will affect the usefulness of the data.

For each tuple, we calculate its tuple hash $\mathcal{H}(t)$ using Eq. (5), which identifies the watermarking bit index, the target

attribute and target bit position for concealing \mathcal{W}_t (Khanduja et al., 2015)

$$\mathcal{H}(t) = \text{Hash}(\mathcal{K}_s \parallel \mathcal{K}_d(t) \parallel \mathcal{K}_s) \quad (5)$$

where \mathcal{K}_s is a secret key selected by the owner, $\mathcal{K}_d(t)$ is the derived key of the tuple t and \parallel is the concatenation operator. Presence of a secret key and a hash function makes the watermark inherently robust against various malicious attacks. Furthermore, the secret key \mathcal{K}_s is concatenated with \mathcal{K}_d to enhance the security. Since owner knows \mathcal{K}_s , she can reproduce $\mathcal{H}(t)$ but Mallory, an attacker, does not know \mathcal{K}_s . This makes it more difficult for an attacker to compute $\mathcal{H}(t)$.

For the information recovery, we calculate tuple hash $\mathcal{H}_r(t)$ as

$$\mathcal{H}_r(t) = \text{Hash}(\mathcal{K}_r \parallel \mathcal{K}_d(t) \parallel \mathcal{K}_r) \quad (6)$$

where \mathcal{K}_r is the secret key decided by owner for information recovery process, and $\mathcal{K}_d(t)$ is the primary key attribute of tuple t . This recovery tuple hash $\mathcal{H}_r(t)$ is used to calculate target locations for embedding \mathcal{W}_t .

Embed_Watermarks(.)*Input:* $\mathcal{R}, A_\beta, A_{\beta_0}, \beta, \alpha, \mathcal{K}_d, \mathcal{K}_s, N_H$ *Output:* Watermarked Database \mathcal{R}_w

1. **For each** tuple $t \in \mathcal{R}$
2. **For each** candidate attribute $A_i \in A_\beta$
3. **If** $A_i \in A_{\beta_0}$
4. Calculate $\mathcal{H}(t)$ using equation (5)
5. **If** $((\text{mod}(\mathcal{H}(t), \beta_0) == i))$
6. Calculate $x = \text{mod}(\mathcal{H}(t), \alpha)$
7. Calculate $p = \text{mod}(\mathcal{H}(t), N_H)$
8. Embed hextobin($\mathcal{W}_o(p)$) in fractional part of $A_i(t)$ starting at x^{th} position
9. **End If** (line5)
10. **Else**
11. Calculate $\mathcal{H}_r(t)$ using equation (6)
12. Calculate $x = \text{mod}(\mathcal{H}_r(t), \alpha)$
13. Embed dectobin($\mathcal{W}_r(t)$) in fractional part of $A_i(t)$ starting at x^{th} position
14. **End If** (line 3)
15. **End for** (line 2) (line 1)

Pseudo 1 Pseudo-code for embedding watermarks in the Relational Database \mathcal{R} .

We divide the watermark insertion phase into following two sub-phases:

A. Select target positions: The proposed algorithm decides the attributes that are earmarked for embedding $\mathcal{W}_i(t)$ and \mathcal{W}_o (Lines 2 to 5). Only that candidate attribute $A_i \in A_{\beta_0}$ that satisfies $(\text{mod}(\mathcal{H}(t), \beta_0) == i)$ is selected for embedding \mathcal{W}_o . $\mathcal{W}_i(t)$ is embedded repeatedly in all the attributes present in A_{β_r} . Algorithm computes the secret positions within the fractional part where the watermarks are embedded (Lines 6, 8, 12 and 13).

B. Embed Watermark: The selected hex character of \mathcal{W}_o is converted into binary first and then embedded. Each hex character will take four binary places that are concealed at $x^{\text{th}}, (x+1), (x+2)$ and $(x+3)$ fractional positions in A_i . Similarly, information recovery watermark $\mathcal{W}_r(t)$ is converted to binary form and then embedded in the fractional part of A_i starting from selected x^{th} position. We consider, five-bit recovery watermark that is embedded into the candidate attributes repeatedly within a tuple.

We need minimum one member in set A_{β_r} to embed \mathcal{W}_i . However, we recommend $|A_{\beta_r}| \geq 3$, which embeds \mathcal{W}_i thrice to allow triple redundancy based majority voting. As the number of candidate attributes increases, probability of successful information recovery also increases due to majority voting applied in extraction of watermark process.

3.3. Watermark extraction

In case of any dispute over ownership or Alice suspects that her database is illegally copied or tampered with, Alice extracts the watermark to verify the ownership from suspected database \mathcal{R}' . In case of applications, where information recovery is done at receiver end, the secret key

\mathcal{K}_r along with A_{β_r}, α and C is securely transferred to receiver. Receiver then extracts \mathcal{W}_i to recover the required information. Receiver cannot extract \mathcal{W}_o as he is not having $\mathcal{K}_s, \beta, A_{\beta_0}$. Thus, the ownership of the original sender/owner is still preserved within a database.

The pseudo-code of watermark extraction process from the suspected watermarked database \mathcal{R}' is inscribed in Pseudo 2. Lines 1 to 8 are same steps of selecting target locations as done in *Embed_Watermarks(.)*. For each tuple t , first embedded watermark characters are extracted then majority voting is applied on extracted redundant bits. If conditions on line 4 and 6 are true then single hex character of \mathcal{W}_o per tuple is extracted and stored in matrix M_o (Line 9). If the candidate attribute belongs to A_{β_r} , $\mathcal{W}_i(t)$ is extracted from multiple attributes per tuple and stored in matrix M_i , which is cleared to zero initially for each tuple (Line 14). Some of the $\mathcal{W}_i(t)$ may be obliterated due to attacks. Hence majority voting on all the values of M_i is performed per tuple to get the correct cluster-id for every tuple (line 17). Finally, the recovery watermark of all the tuples are combined to get \mathcal{W}_r' (line 20).

When all tuples are processed, majority voting on $M_o[p][1:N_r + 1]$ is performed and the extracted hex character is stored in $\mathcal{W}_o'(p)$ (Line 19). For each position p , hex character of \mathcal{W}_o may be embedded N_r or $N_r + 1$ times assuming N_r as number of times each watermark character is embedded repeatedly into a database. This is due to use of lower bound while calculating N_r in Eq. (7)

$$N_r = \left\lfloor \frac{N_t}{N_H} \right\rfloor \quad (7)$$

where N_t is the number of tuples and N_H is the number of hex characters in watermark.

Extract_Watermarks(.)

Input: $\mathcal{R}', A_\beta, A_{\beta o}, \beta_o, \alpha, \mathcal{Z}_d, \mathcal{Z}_s, N_H, N_r$

Output: Information recovery watermark \mathcal{W}_r' , Ownership watermark \mathcal{W}_o'

1. **For each** tuple $t \in \mathcal{R}'$
2. **For each** candidate attribute $A_i \in A_\beta$
3. Clear matrix $M_r[1:\beta - 1]$.
4. **If** $A_i \in A_{\beta o}$
5. Calculate $\mathcal{Z}(t)$ using equation (5)
6. **If** $((\text{mod}(\mathcal{Z}(t), \beta_o) == i)$
7. Calculate $x = \text{mod}(\mathcal{Z}(t), \alpha)$
8. Calculate $p = \text{mod}(\mathcal{Z}(t), N_H)$
9. Extract the values present at x^{th} to $(x + 3)^{th}$ position in the fractional part of $A_i(t)$, convert them to hex character and store in matrix $M_o[p][1:N_r + 1]$
10. **End If** (Line 6)
11. **Else**
12. Calculate $\mathcal{Z}_r(t)$ using equation (6).
13. Calculate $x = \text{mod}(\mathcal{Z}_r(t), \alpha)$
14. Extract the values present at the x^{th} to $(x + 4)^{th}$ position in the fractional part of $A_i(t)$, convert them to decimal form and store in the matrix $M_r[i]$.
15. **End If** (line 4)
16. **End For** (line 2)
17. Perform majority voting on $M_r[1:\beta_r]$ to obtain cluster-id of t and store the value in $\mathcal{W}_r'(t)$.
18. **End for** (line 1)
19. Perform majority voting on $M_o[p][1:N_r + 1]$ for each watermark position p and save in $\mathcal{W}_o'(p)$.
20. The cluster-ids of all tuples are combined to form \mathcal{W}_r' .

Pseudo 2 Pseudo-code for extracting watermarks from suspected watermarked database \mathcal{R}'

Table 1 – Change in an attribute values after embedding \mathcal{W}_i .

Values of attributes before embedding	Values of attributes after embedding \mathcal{W}_i	Parameters
48.1234	48.1226	$\mathcal{W}_i = 10$, BPos = 5
98.4596	98.4574	$\mathcal{W}_i = 15$, BPos = 6
567.1009	567.1008	$\mathcal{W}_i = 16$, BPos = 5
42.3333	42.3368	$\mathcal{W}_i = 20$, BPos = 6

Table 1 show the values of some of the watermarked attributes under various parameters when \mathcal{W}_i is embedded. We now illustrate the process of embedding and extracting back information recovery watermark in detail.

We trace the steps mentioned in pseudo 1 to embed \mathcal{W}_i in selected attribute. Let, Attribute $A_i = 48.1234$ is chosen for watermarking. Extracting fractional part of A_i yields 1234. Next, we convert it to binary form and get fractional part as $(10011010010)_2$. Considering BPos = 5 and $\mathcal{W}_i = 10$, we take binary

of watermark as 01010 and embed it at 5th LSB of fractional part. Now, watermarked fractional part becomes $(10011001010)_2$. So, watermarked attribute $A_i^w = 48.1226$.

While extracting a watermark from the watermarked attribute we follow steps of pseudo 2. Let us take the same watermarked attribute $A_i^w = 48.1226$ to extract the embedded watermark. Converting the fractional part of the selected attribute to binary will give $=10011001010$. BPos is calculated and watermark bits are extracted. Say, we get BPos = 5, therefore 01010 is extracted as \mathcal{W}_i , which is = 10 in decimal.

3.4. Decision generation

Mallory might try to destroy the watermark embedded into the database and then claim ownership over the pirated copy. Such disputes are handled in this phase. The watermark bits extracted \mathcal{W}_k from a suspected database are matched with the corresponding bits of the original ownership watermark \mathcal{W}_o . The original watermark is either stored or can be re-generated in watermark preparation phase. If the number of matches

N_m is very large, we suspect piracy. To assess such quantities, a threshold τ_w is taken (Lipschutz and Schiller, 2005). We define τ_w as the threshold value that decides the database piracy. It tells the permissible fraction of matches of extracted watermark bits with original embedded watermark bits to claim ownership over disputed database. This value is decided by the owner depending on the sensitivity of the data.

If $(N_m \geq \tau_w * N_w)$, then database is pirated. (8)

For tamper detection of pivotal information, we check whether the suspected data lies within the range of cluster it belongs to. This is achieved by comparing the stored minimum and maximum deviation from centroid with the difference of suspected value and its centroid. The centroid of the suspected value is obtained using extracted watermark \mathcal{W}' and stored centroid set. If the suspected data are found to be in different cluster, \mathcal{R}' is tampered and the value is replaced with its centroid. Thus, stored information C along with \mathcal{W}' will regenerate the tampered information. In case of deletion of pivotal attribute, i.e. data are completely lost, regenerated information is added. The entire data is not recovered, but the information conveyed is recovered.

4. Security analysis

Suppose Alice owns a database \mathcal{R} which is freely available on the internet. Alice has embedded watermarks \mathcal{W}_0 into the database to generate a watermarked data set \mathcal{R}_w . Mallory, an attacker, does not have any knowledge of the original database \mathcal{R} and the secret parameters β, α and K_s . Mallory may try to destroy these watermarks to claim the database to be hers and/or delete/alter important information from the database. Under such circumstances, we justify the security of our model w.r.t. following parameters:

4.1. Probability to correctly identify potential locations

The technique embeds ownership watermark \mathcal{W}_0 into all the tuples. For each tuple different attribute is selected out of β_0 attributes. Without the knowledge of A_{β_0} , Mallory will select A_f as a set of all the attributes of type float such that $A_{\beta_0} \subseteq A_f$ and α_f as number of LSBs where watermark can be embedded such that $\alpha \leq \alpha_f$.

Mallory will select β_f candidate attributes out of which any one attribute is selected per tuple. Thus, probability of correctly identifying watermarked positions of all the N_t tuples is given as

$$P_1 = \left(\frac{1}{\beta_f * \alpha_f C} \right)^{N_t} \quad (9)$$

Taking $\alpha_f = 10, \beta_f = 15, N_t = 6000$ in Eq. (9), we get $P_1 = 1.37 * e^{-20990}$, which is negligible. Hence, attribute and watermark bit positions selection is a secure process.

4.2. False hit rate

We define P_{FH} as the probability of extracting a valid watermark from a non-watermarked database. We calculate P_{FH} and

prove that it is low in our technique. Each extracted character $\mathcal{W}_0^*(p)$ from a non-watermarked relation has same probability of 1/2 to match or not match the original embedded hex character of the watermark. Thus, we define P_{hex} as the probability that at-least one more than half (majority threshold) of N_r hex characters can be detected from non-watermarked relation by sheer chance. We calculate P_{hex} by using binomial distribution considering N_r independent trials (Khunduja et al., 2015).

$$P_{hex} = \sum_{j=N_r/2+1}^{N_r} b\left(j; N_r, \frac{1}{2}\right) = B((0.5 * N_r + 1); N_r, 0.5) \quad (10)$$

where $b(j; n, p)$ is the probability of obtaining j success in n Bernoulli trials with probability p for success and $1-p$ for failure, and $B(j; n, p)$ is referred to as cumulative binomial probability representing the probability of obtaining at-least j success from n Bernoulli's trials (Hypergeometric distribution). For a watermark of length N_H , the false hit rate P_{FH} is given by

$$P_{FH} = B(\tau_w * N_H; N_H, P_{hex}) \quad (11)$$

where τ_w is the watermark threshold on entire database, N_H is the length of watermark in hexadecimal. Fig. 4 shows the variation of P_{FH} with τ_w by keeping other parameters constant. We took $N_t = 6000, N_H = 64$ and plotted the graph by varying τ_w . The figure illustrates that false hit rate is monotonic decreasing with increase in threshold τ_w . We also observed the effect of N_t . It is found that as N_t increases, P_{FH} increases but slowly. False hit rate P_{FH} is low for the nominal value of $\tau_w = 0.6$, which means that our method is robust and secure.

4.3. False miss rate

We define False Miss rate P_{FM} as the probability of not detecting a valid watermark from a watermarked relation. Mallory will try to destroy the watermark \mathcal{W}_0 through various attacks. If Mallory successfully deletes $(\tau_w * N_H + 1)$ hex characters of \mathcal{W}_0 , a watermark is destroyed. We define P_{FM} as given in Eq. (12).

$$P_{FM} = \prod_{i=1}^{\tau_w * N_H + 1} Prob(i) \quad (12)$$

where $Prob(i)$ is the probability of successfully destroying i th watermark character. We now analyze the false miss rate against Subset addition, alteration and deletion attacks subsequently.

A. In subset addition attack: Mallory tries to demolish the watermark \mathcal{W}_0 by adding spurious tuples in the database. In this way, she hopes to overshadow the watermark embedded in the database. In the proposed technique, the watermark character is embedded repeatedly N_r times into the database and then a majority voting mechanism is applied to extract the correct watermark. This means Mallory will add minimum of (N_r) additional spurious tuples to make watermark character corrupted in majority voting test. Probability that i th watermark character is completely destroyed with random addition of tuples is calculated using hyper-geometric distribution (Khunduja et al., 2013) as:

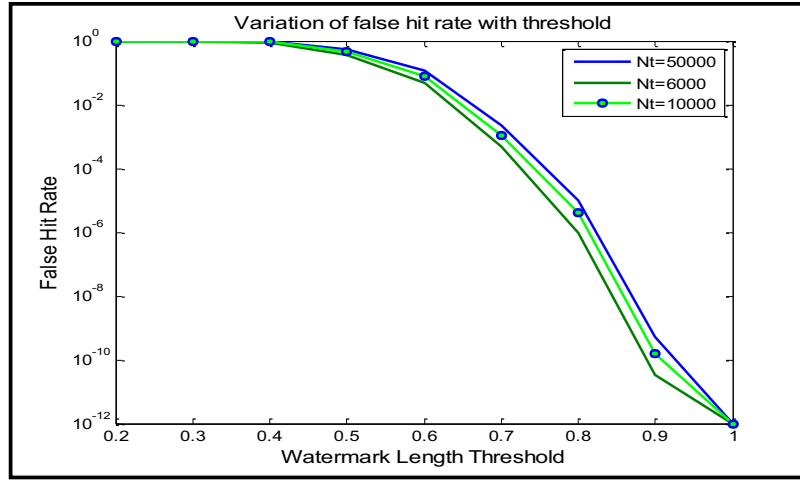


Fig. 4 – Variation of false hit rate with watermark threshold τ_w .

$$Prob(i) = \left(\frac{N_t C * \frac{N_t - N_r}{\phi - N_r} C}{N_t C} \right) \quad (13)$$

Taking $\phi = 1000$, $N_t = 6000$, $\tau_w = 0.6$, $N_w = 50$ we calculated $P_{FM}(\text{Subset Addition}) = 3.2e-413$ by substituting Eq. (13) in Eq. (12). Thus P_{FM} is negligible and close to zero for subset addition attack.

B. In subset alteration attack: Mallory tries to modify ϕ tuples randomly to alter the embedded watermark \mathcal{W}_0 . Let us consider the worst-case scenario where altering any tuple will result in change in watermark. In order to completely obliterate a hex character of a watermark, Mallory has to alter one more than half of tuples corresponding to i^{th} character of a watermark, i.e. $((N_r/2) + 1)$. Probability that i^{th} character of a watermark is completely destroyed with random alterations of tuples is given by hypergeometric distribution (Khanduja et al., 2013) as:

$$Prob(i) = \left(\frac{N_t C * \frac{N_t - N_r}{\phi - N_r/2 - 1} C}{N_t C} \right) \quad (14)$$

Taking $\phi = 1000$, $N_t = 6000$, $\tau_w = 0.6$, $N_w = 50$, we calculated $P_{FM}(\text{Subset Alteration}) = 1.2e-88$ by substituting Eq. (14) in Eq. (12), thus P_{FM} is negligible. Hence, proposed technique is robust against this attack.

C. In subset deletion attack: Mallory tries to delete random tuples from the database with an intention of destroying the watermark. In each tuple, we embedded one hex character of watermark \mathcal{W}_0 ; hence to delete the entire watermark, she must destroy a significant portion of the database. To destroy i^{th} character of a watermark, random deletions must include N_r tuples containing i^{th} character of a watermark. The probability that i^{th} hex character of a watermark is completely destroyed with random deletions using Eq. (13). It is found to be negligible.

5. Experimental results and analysis

The experiments are validated on an Intel Core™ i7 2.30 GHz system in MATLAB 7.8.0. The database (National Geochemical

Survey Database of the US) of a national geochemical analysis of stream sediments and soils conducted in the United States (US) is used. These data comprise a complete, national-scale geochemical coverage of the US and enable construction of geochemical maps, refine estimates of baseline concentrations of chemical elements in the sampled media, and provide context for a wide variety of studies in the geological and environmental sciences. In the database, the stream sediments and soils in the US from existing data are analyzed. The database contains information about soil samples at different geological locations. Database contains 287 attributes of text, numeric and categorical data types. Pivotal attributes pH, grain size, orgn_pct (estimated % organics) are considered as elements of A_p and is being categorized into 31 categories. The candidate attributes A_β chosen are Al_ICP40, Ca_ICP40, Fe_ICP40, K_ICP40, Na_ICP40, P_ICP40, Ti_ICP40, Ag_ICP40, As_ICP40, Mg_ICP40, Co_ICP40, Cu_ICP40, which gave the amount of aluminum, calcium, iron, potassium, sodium, phosphorous, titanium, silver, arsenic, magnesium, cobalt and copper respectively present in the soil samples. We used subset comprising 16,000 tuples with $\beta = 12$ and $\alpha = 10$. We have implemented SHA-2 algorithm, thus $N_H = 64$ (Wikipedia, 2015).

5.1. Robustness analysis

In Section 4, we have theoretically proven the resilience of our technique. We now experimentally verify that the watermark \mathcal{W}_0 embedded is robust and can withstand several attacks. To prove the robustness of the proposed technique, we compare our results with one of the latest work in literature in our domain. In Farfoura et al. (2012), the authors have proposed robust watermarking technique targeting float type attributes.

A. Subset addition attack: We performed the experiment by varying percentage of added tuples from 10 to 100. Fig. 5 shows the comparison of our technique with that of Farfoura et al. (2012). We found that the percentage of watermark extracted accurately is 100 even when 100% of extra tuples are added, verifying the robustness of our technique against this attack. The work in Farfoura et al. (2012) demonstrates that the watermark is substantially destroyed on 60% addition of new tuples.

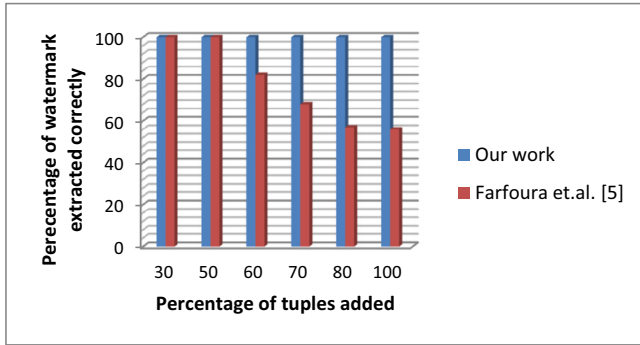


Fig. 5 – Watermark degradation by Subset Addition attack of our work and work in Farfoura et al. (2012).

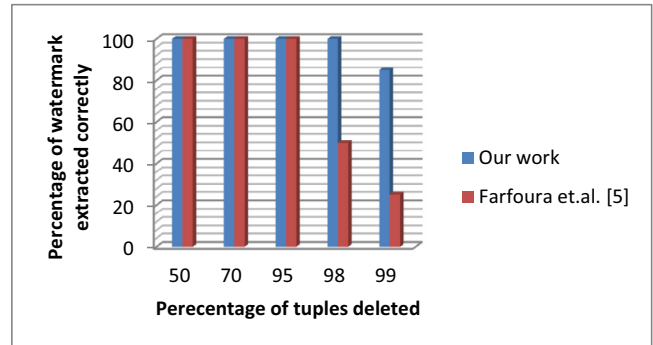


Fig. 7 – Watermark degradation by Subset deletion Attack.

B. Subset alteration attack: For altering the numeric attributes, we have randomly chosen any position from the fractional part of the number and replace it by a random digit. Fig. 6 shows the comparison of our technique with that of Farfoura et al. (2012). The work in Farfoura et al. (2012) demonstrates that the watermark is substantially destroyed when more than 80% of tuples are altered. This compromise on robustness is overcome in our technique, since the experimental results have proven that on 100% tuple alteration, we are successfully able to recover the complete watermark.

C. Subset deletion attack: Fig. 7 shows that even on 98% deletion, we are able to extract the entire watermark. The database would be rendered useless if more than 98% of the tuples are deleted. The results of work done in Farfoura et al. (2012) shows that the extracted watermark bits starts degrading after 95% deletion.

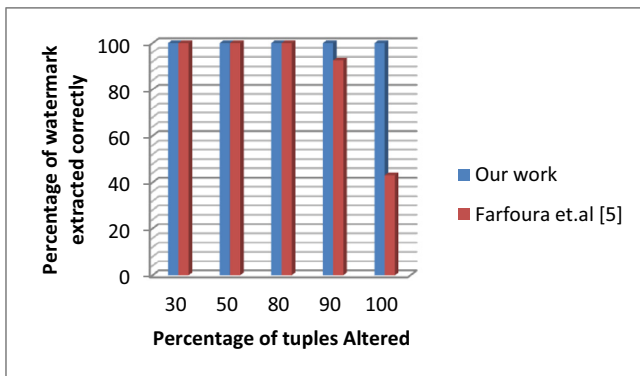


Fig. 6 – Watermark degradation by Subset Alteration attack.

5.2. Information recovery analysis

The information concealed in a database along with the stored centroid values are used to recover the lost/altered data. The category of the perturbed data is extracted from the candidate attributes of that tuple. Using the category, we get to know the cluster Id associated with the perturbed data. We then replace the altered/lost value with its centroid. We statistically measure the amount of distortion produced by replacing the original data with its centroid. The experiment is performed by deliberately introducing errors in the database, i.e. by altering the original attribute values in a manner such that it results in a change of their cluster Ids. Table 2 shows the change in standard deviation of two sample clusters (i) before adding distortion and (ii) after adding distortion and recovering the same by replacing with their original cluster Ids. We varied the amount of distortion introduced to examine their effect.

Results show that there is a decrease in standard deviation of the elements after information recovery as compared to standard deviation of the original clusters. This is due to the replacement of altered attribute values with their respective centroids. Furthermore, there is a decrease in standard deviation with the increase in number of attributes altered as shown in Table 2.

We next performed the experiments to check the accuracy of information extracted correctly from candidate attributes against the following attacks:

A. Subset alteration attack: Fig. 8 shows a plot of percentage of correct information extracted against percentage of tuples altered. This is observed by varying the number of perturbed candidate attributes per tuple (β , from 1 to 10). It is observed that as more candidate attributes are altered, percentage of correctly extracting \mathcal{W}_i decreases slowly. Only when all the candidate attributes are altered, percentage drops significantly with a number of tuples altered.

Table 2 – Statistical measure of distortion introduced in pivotal attribute after information recovery.

Change in original clusters	Cluster 1			Cluster 2		
	σ before distortion	σ after recovery	Difference	σ before distortion	σ after recovery	Difference
20%	0.31781	0.29289	0.1492	0.13058	0.11546	0.01512
40%	0.31781	0.13171	0.18610	0.13058	0.11176	0.01882
60%	0.31781	0.08202	0.23579	0.13058	0.06114	0.06944

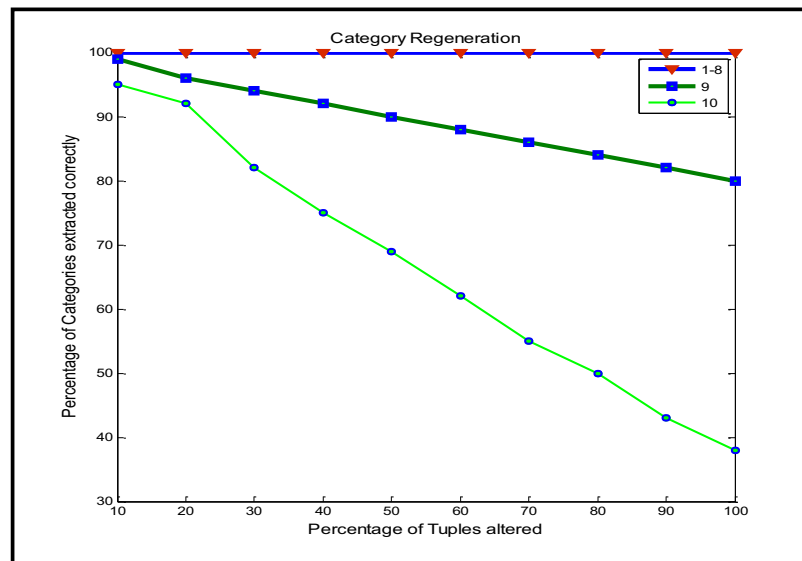


Fig. 8 – Graph showing percentage of information recovered in subset alteration attack.

It may be noted that our technique does not provide any mechanism to recover the data in case an entire tuple is deleted. We consider that deleting data value completely is against usability constraints. However, important information can still be retrieved by modifying the proposed scheme. We can store the partition/cluster set S along with C . The S contains information about data that belong to a particular cluster. This enables us to trace the missing data, and we can recover the information accordingly at the cost of additional storage space. Thus, proposed technique provides complete information recovery against alteration and deletion attacks.

B. Rounding attack: We studied the effect of rounding attacks on candidate attributes of type float. The results are plotted in Fig. 9 showing the percentage of tuples from which recovery watermark is correctly extracted, when an attacker rounds-off the fraction at tenths or hundredths places with an intention to remove a watermark.

Results show that even when half of the values of the candidate attributes are altered by rounding attack; recovery information from 80% to 90% of total tuples in a database is successfully extracted. Even if all the candidate attributes are

rounded off to tenths place, the recovery information is correctly extracted from 69.8% of the total tuples. We have shown rounding attack till hundredths place. Altering the values in the database beyond this i.e. thousandths position and above can violate the usability constraints of the database. The resilience to this attack is due to (i) the repeated embedding of recovery information in multiple attributes, and (ii) embedded the watermark within the selected positions among α LSBs of the fractional part of the targeted attribute.

5.3. Imperceptibility

We use mean, standard deviation and Kullback-Leibler distance (KL) as the statistical metrics to measure the quantitative alterations occurred while embedding watermarks in a database. Table 3 shows the result of this experiment after taking modulus of the difference in mean, standard deviation and KL distance of attributes before and after embedding watermark in a database. We calculated the KL distance between the distributions before and after embedding watermarks in candidate

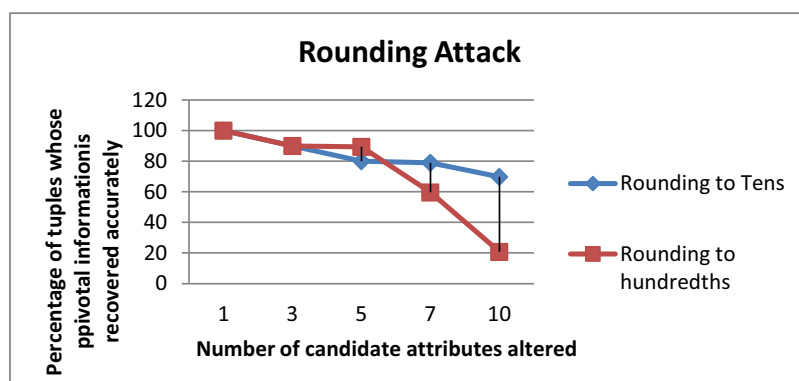


Fig. 9 – Graph showing percentage of tuples whose pivotal information is correctly recovered after rounding attack.

Table 3 – Change in mean and standard deviation of attributes during watermarking.

Attributes	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉
Mean	0.014	0.188	0.086	0.164	0.23	0.062	0.196	0.19	0.11
Standard deviation	1.0942	0.4480	0.0176	0.3245	0.1172	0.0511	0.0063	0.4913	0.4329
KL	0.0177	0.0006	0.1067	0.016	0.0018	0.0436	0.0156	0.0002	0.0035

attributes. Negligible changes in these statistics prove the imperceptibility of our algorithm.

5.4. Comparative evaluation

The prior work proposed in literature discusses a fragile watermarking technique that can detect and correct distortions in relational databases by embedding watermarks created from each attribute value, thereby recovering true data (Khataeimaragheh and Rashidi, 2010). However, the technique only recovers the tampered data; in case of deletion of an attribute value, the recovery technique completely fails. Furthermore, if two or more than two values of different tuples and attributes are altered, then this technique will fail to localize the exact positions where the alterations might have occurred. Hence, recovery cannot be done in such a scenario. We found that the probability of accurately detecting, localizing and hence rectifying errors reduces drastically when the number of errors exceeds two. Table 4 enumerates the key differences between our technique and of Khataeimaragheh and Rashidi (2010).

Furthermore, we compared the robustness results of our technique for ownership proof with work of Farfoura et al. (2012). The graphs shown in Figs. 5, 6 and 7 proves superiority of our technique w.r.t. robustness.

6. Conclusions

In this paper we discussed a novel robust watermarking technique that regenerates information of pivotal attributes lost during database transmission. The proposed scheme regenerates crucial information encoded in the data in the event of both alterations as well as deletion of pivotal attributes in a database. Our proposed dual watermarking technique is well-

suited for web databases that necessitate both information and ownership protections. The major contributions are summarized below:

1. Information and ownership protection of RDBs is achieved by embedding dual watermarks.
 - i) Watermark can be utilized as an information carrier to detect information loss and its recovery.
 - ii) Watermark can contain owner specific information to claim ownership.
2. Unsupervised learning such as k-means gives a sound basis in discovering salient information contained in the data and facilitates the recovery of lost information due to rampant attacks in the databases.

REFERENCES

- Agrawal R, Haas PJ, Kiernan J. Watermarking relational data: framework, algorithms and analysis. *VLDB J* 2003;12(2):157–69.
- Ali YH, Mahdi BH. Watermarking for relational database by using threshold generator. *Eng & Tech J* 2011;29:33–43.
- Camara L, Li J, Li R, Xie W. Distortion-free watermarking approach for relational database integrity checking. *Math Probl Eng* 2014;Article ID 697165. <<http://dx.doi.org/10.1155/2014/697165>>.
- Cui X, Qin X, Sheng G. A weighted algorithm for watermarking relational databases. *Wuhan Univ J Nat Sci* 2007;12(1):79–82.
- Farfoura ME, Horng S-J, Lai J-L, Run R-S, Chen R-J, Khan MK. A blind reversible method for watermarking relational databases based on a time-stamping protocol. *Expert Syst Appl* 2012;39(3):3185–96.
- Guo H, Li Y, Lui A, Jajodia S. Fragile watermarking scheme for detecting malicious modifications of database. *Inf Sci (Ny)* 2006;176:1350–78.
- Guo J. Fragile watermarking scheme for tamper detection of relational database. In: *International conference on computer and management*, 2011. 2011. p. 1–4.

Table 4 – Comparison of proposed technique with prior work.

	Our technique	Khataeimaragheh (Khataeimaragheh and Rashidi, 2010)
Security issues handled	We proposed a robust watermarking technique that resolves twin issues of ownership proof and information recovery.	A fragile watermarking technique for tamper detection with data recovery is proposed.
Objective	Focus on the information contained in the databases and recovers them from losses.	No focus on information recovery. It recovers the data if there is an alteration in single bit attribute.
Recovery in case of alteration in a database	Alterations in information contained in any of the pivotal attributes are detected and recovered. Entire information is recovered for altered pivotal attributes.	Localization fails for two or more than two alterations in different tuples/attributes, and hence the recovery fails. Alterations in LSBs are not recovered as a watermark is formed by truncating signature of database.
Recovery in case of deletion of attribute/s	Technique recovers the information lost in case of deletion of pivotal attributes.	Recovery fails in case of deletion of attribute value.
Primary key independence	Technique is primary key independent.	Technique is primary key dependent.

- Hartigan JA, Wong MA. A k-means clustering algorithm. *J R Stat Soc Ser C Appl Stat* 1979;100-8.
- Hypergeometric distribution. <<http://stattrek.com/probability-distributions/hypergeometric.aspx>>.
- Ifthikar S, Kamran M, Anwar Z. RRW-A robust and reversible watermarking technique for relational Data. *IEEE Trans Knowl Data Eng* 2015;27(4):1131-45.
- Kerckhoffs A. La cryptographie militaire. *J Sci Mil* 1883;9(1):5-83.
- Kerr IR, Maurushat A, Tacit CS. Technical protection measures: tilting at copyright's windmill. *Ottawa Law Rev* 2003;34:6-82.
- Khan A, Husain SA. A fragile zero watermarking scheme to detect and characterize malicious modifications in database relations. *ScientificWorldJournal* 2013;Article ID 796726.
- Khanduja V, Verma OP, Chakraverty S. Watermarking relational databases using Bacterial Foraging Algorithm. *Multimed Tools Appl* 2013;74(3):813-39.
- Khanduja V, Chakraverty S, Verma OP, Tandon R, Goel S. A robust multiple watermarking technique for information recovery. In: *IEEE international advance computing conference*. 2014. p. 250-5.
- Khanduja V, Chakraverty S, Verma OP. Watermarking categorical data: robustness and algorithm analysis. *Def Sci J* 2015;65(3):226-32.
- Khataeimaragheh H, Rashidi H. A novel watermarking scheme for detecting and recovering distortions in database tables. *Int J Database Manag Syst* 2010;2:1-11.
- Li Y, Guo H, Jajodia S. Tamper detection and localization for categorical data using fragile watermarks. In: *Proceeding of ACM workshop on digital rights management*. 2004.
- Lipschutz S, Schiller J. *Schaum's outline of theory and problems of Introduction to probability and statistics*. India: McGraw Hill Education; 2005.
- National Geochemical Survey Database of the US. <http://mrdata.usgs.gov/geochem/select.php> [accessed 05.02.15].
- Schneier B. *Applied cryptography protocols, algorithms and source in C*. 2nd ed. India: Wiley; 1996.
- Shehab M, Bertino E, Ghafoor A. Watermarking relational databases using optimization-based techniques. *IEEE Trans Knowl Data Eng* 2008;20(1):116-29.
- Sion R, Atallah M, Prabhakar S. Rights protection for relational data. *IEEE Trans Knowl Data Eng* 2004;16(12):1509-25.
- Wikipedia. SHA-2 [Online]. 2015. <<https://en.wikipedia.org/wiki/SHA-2>>.



Enhanced dielectric properties of multilayered BiFeO₃/BaTiO₃ capacitors deposited by pulsed laser deposition

Savita Sharma, Monika Tomar, Nitin K. Puri, and Vinay Gupta

Citation: [AIP Conference Proceedings](#) **1724**, 020098 (2016); doi: 10.1063/1.4945218

View online: <http://dx.doi.org/10.1063/1.4945218>

View Table of Contents: <http://scitation.aip.org/content/aip/proceeding/aipcp/1724?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Stress induced enhanced polarization in multilayer BiFeO₃/BaTiO₃ structure with improved energy storage properties](#)

[AIP Advances](#) **5**, 107216 (2015); 10.1063/1.4934578

[Enhanced ferroelectric photovoltaic response of BiFeO₃/BaTiO₃ multilayered structure](#)

[J. Appl. Phys.](#) **118**, 074103 (2015); 10.1063/1.4928964

[Magnetic spin structure and magnetoelectric coupling in BiFeO₃-BaTiO₃ multilayer](#)

[Appl. Phys. Lett.](#) **106**, 082904 (2015); 10.1063/1.4913444

[Formation of artificial BaTiO₃/SrTiO₃ superlattices using pulsed laser deposition and their dielectric properties](#)

[Appl. Phys. Lett.](#) **65**, 1970 (1994); 10.1063/1.112837

[Magnetic properties of BiFeO₃-BaTiO₃ and BiFeO₃-PbTi\(Zr\)O₃ glassy sputtered films](#)

[J. Appl. Phys.](#) **64**, 5434 (1988); 10.1063/1.342338

Enhanced dielectric properties of multilayered BiFeO₃/BaTiO₃ capacitors deposited by pulsed laser deposition

Savita Sharma^{1,2}, Monika Tomar³, Nitin K. Puri² and Vinay Gupta^{1*}

¹*Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India*

²*Department of Applied Physics, Delhi Technological University, Delhi-110042, India*

³*Miranda House, Department of Physics, University of Delhi, Delhi 110007, India*

^{a)} Corresponding author's e-mail: drvguptavinay@gmail.com, vgupta@physics.du.ac.in

Abstract. We report on dielectric studies of BiFeO₃(BFO)/BaTiO₃(BTO) multilayer structure fabricated by pulsed laser deposition technique. Multilayered capacitors were prepared by increasing number of alternating individual layers from 2 to 7 while maintaining the total thickness of the layered structure as 350 nm. The dielectric constant of the BFO/BTO multilayer structure was significantly increased to 772 (at 1 kHz) and the dielectric loss decreased to 1.08 (at 1 kHz) in comparison to that of bare BFO and BTO thin films of 350 nm thickness. Further the dielectric constant increased with increasing number of individual layers. The increase in dielectric constant is related to the enhancement of ferroelectricity and reduced leakage current which is due to the induced stress at the interface of BFO and BTO layer in the multilayer capacitor. A high value of ferroelectric polarization ($\sim 99.80 \mu\text{C}/\text{cm}^2$) was obtained for 6-layer BFO/BTO structure. The multilayer structure exhibited superior dielectric properties and can be undeniably used as the dielectric layer in silicon-based capacitors and tunable microwave device applications.

INTRODUCTION

Ferroelectric thin films have drawn enough consideration by displaying a wide range of dielectric, pyroelectric, piezoelectric and electro-optic properties leading to a variety of applications in dynamic random access memories (DRAM) capacitors, optoelectronics, infrared sensors, non volatile memories etc [1-4]. The crucial features for the quality and reliability of such devices include 1) low leakage current/low loss, 2) fast dielectric response, 3) long shelf life [5-6]. Wide ranging studies have been executed to enhance the dielectric properties of ferroelectric thin films. One of the encouraging methods is to use thin films in the form of multilayers or superlattices. The conventional physical properties such as, low leakage current, strong poling properties and charge effects can be modified with the presence of strain at interfaces [7-8]. Enhanced multiferroic or ferroelectric properties including improved dielectric constant can be obtained with the blend of different ferroelectric, ferromagnetic and/or multiferroic materials. Zhao et al. [8] prepared BTO/BST/STO multilayer structure, and reported that the dielectric constant was greatly enhanced. Peng et al and Xu et al. fabricated BST multilayers with varying concentration of Ba, Sr and Ti and concluded that multilayer structure was favorable towards reducing the dielectric loss while improving the tenability [9-10]. Zhou et al. [11] identified that the BTO/STO multilayered structure had an enlarged dielectric constant. Pontes et al. examined the dielectric properties of STO/BTO bilayer thin films prepared by chemical route

[12]. Mostly, multilayer thin films are studied for binary or ternary systems, while no systematic study have been executed with effect of variation of individual number of layers (i.e. number of interfaces) on dielectric properties.

Barium titanate, BaTiO_3 (BTO) is one of the most studied perovskite type ferroelectric material with tetragonal symmetry, and carry a relatively high dielectric constant, used most extensively in microelectronic industry, especially as pyroelectric detectors, dielectric material in multilayer ceramic capacitors (MLCC) etc [13-16]. Bismuth ferrite, BiFeO_3 (BFO), a room temperature multiferroic material, demonstrates some promising applications for ferroelectric random access memory and magnetoelectric devices due to its multiferroicity, large remnant polarization and high ferroelectric curie temperature [14,17-18]. However, one of the major problems of BFO in above mentioned applications is its high leakage current. The low leakage current density in BFO can be obtained by A or B site substitution with different ions [19] or by mixing BFO with other perovskites such as PbTiO_3 and BaTiO_3 [20-21]. Detailed literature survey reveals that there is hardly any systematic study on the BiFeO_3 with BaTiO_3 multilayer system.

In this study, BFO/BTO multilayered structures were prepared on Pt/Ti/SiO₂/Si substrates by using multi-target pulsed laser deposition technique. The surface morphology and dielectric properties have been studied using Atomic force microscopy (AFM) and Keithley semiconductor characterization unit (4200). The influence of number of interfaces on the dielectric properties of BFO/BTO multilayer structure has also been examined.

EXPERIMENTAL DETAILS

Multilayered BFO/BTO structures with 2 to 7 alternating layers of BFO and BTO were deposited by pulsed laser deposition using individual BTO and BFO ceramic targets. The multilayered structure was prepared on conducting (bottom electrode) platinized silicon Pt/Ti/SiO₂/Si substrate. The BTO and BFO ceramic targets of one inch diameter were prepared by conventional solid-state route as described in our earlier publication [22]. The 4th harmonic of Nd:YAG pulsed laser ($\lambda = 266$ nm) was used to ablate the respective targets to deposit the multilayer structure. During the deposition the substrate was maintained at 750 °C at an oxygen pressure of 200 mT. Surface profiler (Dektak 150) was used to measure the thickness of (BFO/BTO) multilayered structure. The total thickness of all the multilayered structure was fixed at ~350 nm. The pure single layer BFO and BTO films of 350 nm were also prepared under similar conditions by PLD for comparison. The ferroelectric and dielectric measurements were performed in metal-insulator-metal (MIM) capacitor configuration with Pt/Ti as bottom electrode, while Au is used as top electrode. The 600 μm diameter and 40 nm thin Au circular dots were deposited by thermal evaporation technique through a shadow mask on the surface of multilayered structure. Room temperature ferroelectric polarization hysteresis of the prepared samples was measured in MIM capacitor configuration at a frequency of 1 kHz and at an applied bias of 10 V using a Radiant Technology Precision ferroelectric workstation. The current-voltage characteristics (I-V) and capacitance-frequency (C-f) characteristics were measured using Keithley (4200) semiconductor characterization system. Surface morphology and topography were studied by Atomic force microscopy (AFM) technique using Bruker Dimension icon.

RESULTS AND DISCUSSION

Figure. 1 represents the schematic of the typical multilayer BFO/BTO (6-layers) capacitor structure with gold (Au) top electrode and platinum (Pt) being the bottom electrode. Thickness of the total multilayers was kept constant as ~ 350 nm in all the multilayer structure.

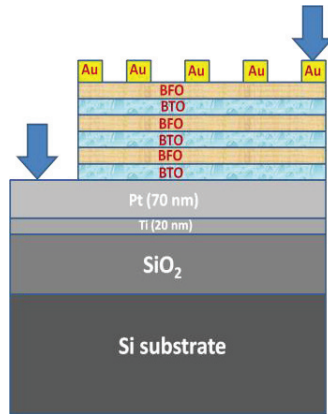


FIGURE 1. Schematic of the BFO/BTO multilayer capacitor structure

Figure. 2 shows the XRD pattern of the prepared multilayer BFO/BTO/Pt/Si structure having three alternate layers. It may be clearly seen from Fig. 2 that the prepared sample exhibits XRD peaks corresponding to either BFO or BTO without formation of any secondary phase indicating the growth of polycrystalline sample having both rhombohedral distorted structure of BFO and tetragonal structure of BTO.

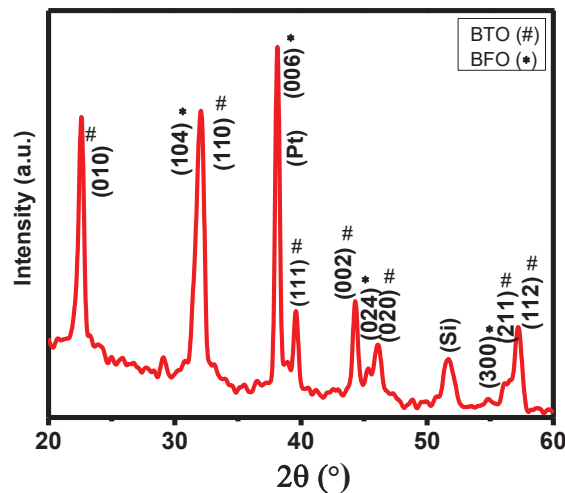


FIGURE 2. XRD pattern of BFO/BTO multilayer structure

The surface morphology of multilayers have been studied by AFM. The three-dimensional AFM micrographs of bare BFO and BTO thin films along with multilayer structure having four and seven alternate layers are shown in Fig. 3(a) to (d). Fig. 3 indicates the fine morphology of multilayer structure having homogenous distribution of grains. The grain size of bare BTO and BFO films are about 50 nm. The grain size decreases significantly in the

multilayer structure. The root mean square roughness (Ra) of the bare BTO and BFO thin films is in the range of 65 nm and 80 nm respectively which increased slightly in the case of BFO/BTO multilayer structures.

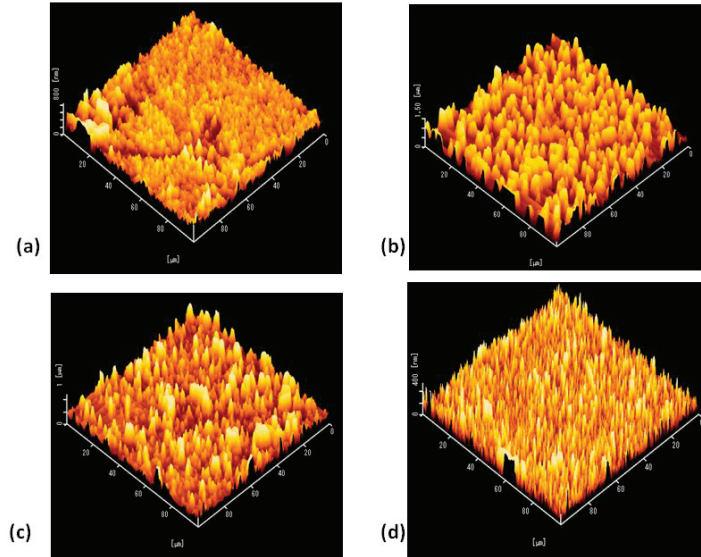


FIGURE 3. AFM images of the samples : (a) BFO thin film (b) BTO thin film (c) 4-layer BFO/BTO structure (d) 7-layer BFO/BTO structure

Figure. 4 shows the room temperature PE-hysteresis curve obtained at 100 kHz for BFO/BTO multilayered structure having six alternate layers. The variation in PE hysteresis curve with varying number of individual layers in multilayer structure have been reported elsewhere. The ferroelectric properties increased and leakage current decreased with the increase in number of alternate layers upto six in BFO/BTO multilayer structure. The saturation polarization was found to be maximum ($99.80 \mu\text{C}/\text{cm}^2$) for BFO/BTO multilayer structure having six alternate layers. The ferroelectric properties degraded with further increase in number of BFO/BTO alternating layers to 7 and may be due to formation of poor interface between the individual BTO and BFO layers.

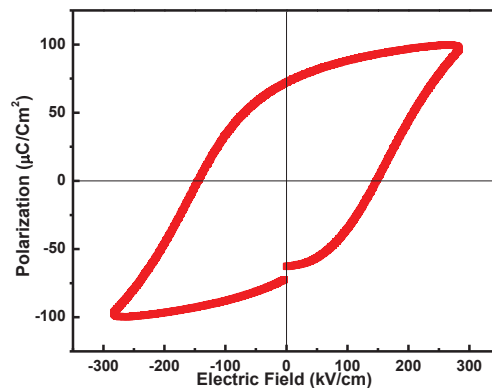


FIGURE 4. Room temperature electric hysteresis loop of multilayer BFO/BTO structure with six alternate layers

Figures. 5 and 6 represent the dielectric constant (ϵ_r) and dielectric loss ($\tan \delta$) of BFO/BTO multilayers as a function of frequency at room temperature. The decreasing trend of dielectric constant with increasing frequency

for all the prepared structures can be seen from fig.5. The corresponding results obtained for bare BFO and BFO thin films are also included in fig. 5, and found to follow similar behavior but with much lower values. It may be noted that ϵ_r increases and $\tan \delta$ decreases continuously with increase in number of alternate layers in multilayer structure upto six, and thereafter decreases for a multilayer structure having seven layers. The ϵ_r of BFO/BTO multilayer structure with six alternate layers is maximum 772 (at 1 kHz) and $\tan \delta$ is lowest 1.08 (at 1 kHz). It undoubtedly points that the dielectric properties were enhanced significantly with multilayer structure. A space charge layer has been brought in this micro-region at the interface between heterogeneous BTO and BFO layers. It may also be noted from fig. 6 that the $\tan \delta$ decreases with increasing frequency in the range 1 kHz to 100 kHz, then tends to attain the constant value after 100 kHz. Furthermore, it is important to note that the loss exhibited by individual BFO or BTO films is higher comparison to those obtained for the multilayered structure. It is proclaimed that the stress at the interface region between different layers plays a key role towards the improvement of the dielectric properties of the multilayer thin film structures [23]. To analyze this further, microstructure and interface of BFO/BTO multilayer thin films will be studied in detail in future. Enhancement in dielectric property in present work is obtained by increasing the stacking periodicity of alternating layers from 2 to 6. The multilayer capacitors having thin alternating layers (~ 60 nm) but with higher periodicities (six in present case) exhibits superior dielectric and ferroelectric property. Similar results have also been reported for sol-gel derived multilayer structure by Shen et al. and Zhao et al. [24, 10]. However this trend is further restrained when the film have more number of periodicities (seven layers) and the average thickness of individual layer is about 50 nm. In the seven layered system, the disorderness increases, the interfaces between the individual layers are tainted and the designed multilayers leans to transform into an homogenous mixing of phases owing to decrease in dielectric constant.

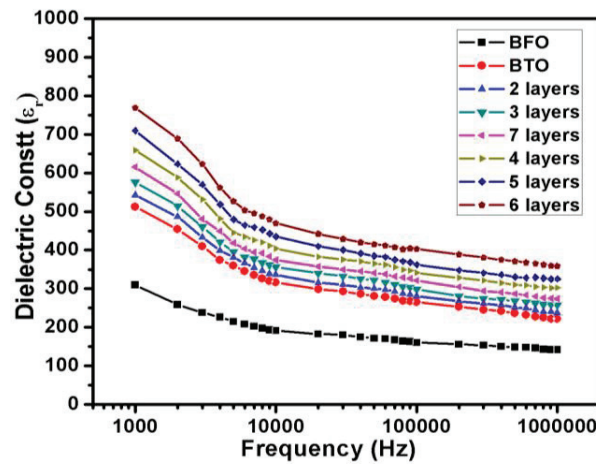


FIGURE 5. Frequency dependence of dielectric constant of multilayered structures along with bare BFO, BTO thin films

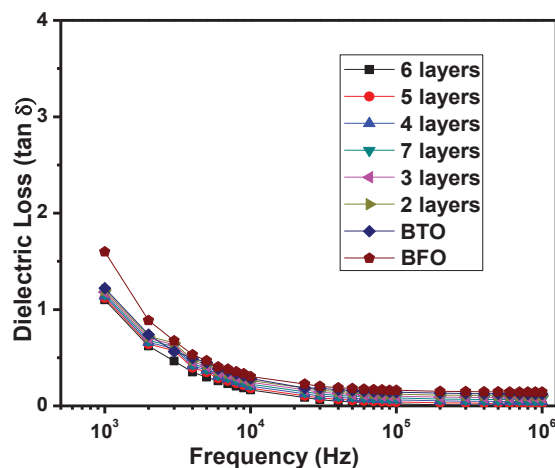


FIGURE 6. Frequency dependence of dielectric loss for multilayered structures along with bare BFO, BTO thin films

CONCLUSION

BFO/BTO multilayer structures were prepared on Pt/Ti/SiO₂/Si substrates by using pulsed laser deposition technique. The dielectric constant of the multilayer films was significantly increased and their $\tan \delta$ loss was found to decrease with increase in number of individual component layers upto six but above that the value of dielectric constant decreased and loss increased. We have shown that the dielectric losses are gradually reduced when the number of interfaces of the component layers were increased. This enhancement in dielectric properties may be attributed to stress at the interface region and further degradation in dielectric properties with increase in individual layers above six may be due to transformation of individual layers into an homogenous mix phase of BFO and BTO. The enhanced dielectric constant and low tangent loss suggest that BFO/BTO multilayer structures have potential application for tunable microwave device applications as well as dielectric layer in silicon-based capacitors.

ACKNOWLEDGMENTS

The authors gratefully thank Dr. Ashok Kumar of NPL, India, for (P-E) measurements and acknowledge the financial support from University of Delhi, India, to carry out the work. One of the authors (SS) is thankful to the Delhi Technological University (DTU) for the teaching assistantship.

REFERENCES

1. M.A. Mohiddon, A. Kumar, and K.L. Yadav, *Physica B* 395, 1 (2007).
2. J. Wang, T. J. Zhang, J. H. Xiang, and B. S. Zhang, *Mater. Chem. Phys.* 108, 445 (2008).
3. M. N. Kamalasanan, S. Chanra, P. C. Joshi, and A. Mansingh, *Appl. Phys. Lett.* 59, 3547 (1991).
4. B. Bihari, J. Kumar, G. T. Stauff, P. C. Van Buskirk, and C. S. Hwang, *J. Appl. Phys.* 76, 1169 (1994).
5. G.W. Dietz, W. Antpohler, M. Klee, and R. Waser, *J. Appl. Phys.* 78, 6113 (1995).
6. H. H. Huang, M. C. Wang, H. J. Lin, M. H. Hon, F. Y. Hsiao, N. C. Wu and C. S. His, *Ceramics International* 37, 3167 (2011).
7. T. X. Li, M. Zhang, Z. Hu, K. S. Li, D. B. Yu and H. Yan, *Solid State Communications* 151, 1659 (2011).
8. N. Zhao, L. Wan, L. Cao, D. Yu, S. Yu, R. Sun, *Mater. Lett.* 65, 3574 (2011).
9. D. W. Peng, J. R. Cheng, and Z. Y. Meng, *Technology of Electroceramics* 21, 668 (2008).
10. Z. Xu, D. Yan, D. Xiao, P. Yu and J. Zhu, *Ceramics International* 39, 1639 (2013).
11. Y. Zhou, *Solid State Communications* 150, 1382 (2010).

12. F. M. Pontes, E. R. Leite, E. J. H. Lee, E. Longo and J. A. Varela, [Thin Solid Films](#) 385, 260 (2001).
13. D. Z. Hu, M. R. Shen, and W. W. Cao., [Microelectron Eng.](#) 83, 553 (2006).
14. W. Cai, C. Fu, J. Gao, Q. Guo, X. Deng and C. Zhang, [Physica B](#) 406, 3583, (2011).
15. X. Ren, [Nature Mater.](#) 3, 91 (2004).
16. G.H. Haertling, [J. American Ceramic Society](#) 82, 797 (1999).
17. M. Sarkar, S. Balakumar, P. Saravanan and S. N. Jaisankar, [Mater. Res. Bull.](#) 48, 2878 (2013).
18. S. Hong, T. Choi, J. H. Jeon, Y. Kim, H. Lee, H. Y. Joo, I. Hwang, J. S. Kim, S. O. Kang, S. V. Kalinin and B. H. Park, [Adv. Mater.](#) 25, 2339 (2013).
19. S. Gupta, M. Tomar, A. R. James and V. Gupta [Ferroelectrics](#) 454, 41 (2013).
20. M. A. Khan, T. P. Comyn and A. J. Bell, [Appl. Phys. Lett.](#) 91, 032901 (2007).
21. M. M. Kumar, A. Srinivas and S. Suryanarayana, [J. Appl. Phys.](#) 87, 855 (2000).
22. S. Sharma, M. Tomar, A. Kumar, N. K. Puri and V. Gupta, [Physica B](#) 448, 125 (2014).
23. O. Nakagawara, T. Shimuta, T. Makino, S. Arai, H. Tabata and T. Kawai, [Appl. Phys. Lett.](#) 77, 3257 (2000).
24. M. R. Shen, S. B. Ge, and W. W. Cao, [J. Phys. D : Appl. Phys.](#) 34, 2935 (2001).

Enhancement of power system damping using Fuzzy Power System Stabilizer

J. N. Rai¹, Ram Bhagat²

^{1, 2}(Electrical Engineering, Delhi Technological University, Delhi, India)

Email: jnrai.phd@gmail.com

Abstract : The use of Power System Stabilizers (PSS) has become very common in operation of large electric power systems. However, it is very difficult to design a stabilizer that could present better performance in all operating points of power systems. In an attempt to cover a wide range of operating conditions, FLC has been suggested as a possible solution to overcome this problem.

This paper presents a study of Fuzzy Power System Stabilizer (FPSS) for stability enhancement of a single machine infinite bus system. In order to accomplish the stability enhancement, speed deviation and acceleration of the rotor synchronous generator are taken as the inputs to the fuzzy logic controller. These variables take significant effects on damping the generator shaft mechanical oscillations. The stabilizing signals were computed using the fuzzy membership function depending on these variables. Simulink Block Design and Matlab-8.1 is utilized in implementing the study. The performance of FPSS is compared with the conventional power system stabilizer and without power system stabilizer.

Keywords: FLC Fuzzy Logic Control, FPSS: Fuzzy Power System Stabilizer, AVR: Automatic Voltage Regulator

I. INTRODUCTION

In early days, many power generating plants were equipped with continuously acting automatic voltage regulators. As the power generated increase and high response exciters come into picture with the use automatic voltage regulators grew it became apparent that the high performance of these voltage regulators had a destabilizing effect on the power system. Power oscillations of small magnitude and low frequency often persisted for long periods of time[1]. In some cases, this presented a limitation on the amount of power able to be transmitted within the system. Power system stabilizers were developed to aid in damping of these power oscillations by modulating the excitation supplied to the synchronous machine.

The power system stabilizer normally consists of a phase-lead compensation blocks, a signal washout block, and a gain block. The input signal to the stabilizer is the equivalent rotor speed deviation. But due to constant stabilizer gain and the complexity of system modeling under different operating conditions calls for new technology to be introduced in damping of small signal oscillation of the system giving origin new type of linguistic based power system stabilizer called FPSS. The FPSS removes most of the shortcomings of the conventional power system stabilizer before this fuzzy logic stabilizer auto tuned power system stabilizer is also consider under this thesis to explain some of the generalized aspects of power system stabilizer[2].

II. SYSTEM MODELING

Thevenin's equivalent circuit of a general system configuration for the synchronous machine connected to the large system is shown in figure 1. This general system is used for the study of small signal stability study.

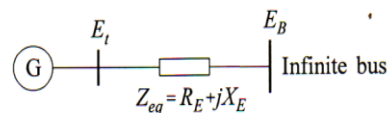


Figure 1: Equivalent circuit General configuration of single machine connected to large system through transmission line

The Block diagram representation shown in figure 2 can be used to describe the small-signal performance.

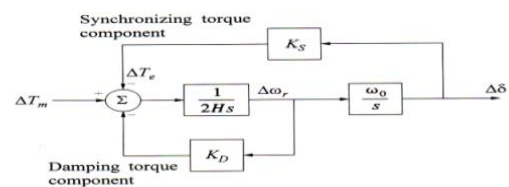


Figure 2: Block diagram of a single-machine infinite bus system with classical generator model

Solving the block diagram we get the characteristic equation:

$$s^2 + \frac{K_D}{2H}s + \frac{K_s \omega_o}{2H} = 0$$

III EFFECT OF EXCITATION

Under the effect of excitation system dynamics, the block diagram developed is extended to include the excitation system[4]. The terminal voltage error signal, which forms the input to the voltage transducer block, is given by

$$\Delta E_t = K_5 \Delta \delta + K_6 \Delta \psi_{fd}$$

Where

$$K_5 = \frac{e_{d0}}{E_{to}} [-R_a m_1 + L_1 n_1 + L_{aq} n_1] + \frac{e_{q0}}{E_{to}} [-R_a n_1 + L_1 m_1 + L'_{ad} m_1]$$

$$K_6 = \frac{e_{d0}}{E_{to}} [-R_a m_2 + L_1 n_2 + L_{aq} n_2] + \frac{e_{q0}}{E_{to}} [-R_a n_2 + L_1 m_2 + L'_{ad} (\frac{1}{L_{fd}} - m_2)]$$

The influence on small-signal stability is examined by considering the excitation system model shown in Figure 3

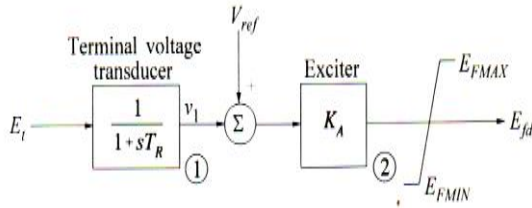


Figure 3 Thyristor excitation system with AVR

The only nonlinearity associated with the model is that due to the ceiling on the exciter output voltage represented by E_{FMAX} and E_{FMIN} . For small-disturbance studies, these limits are ignored as we are interested in a linearized model about an operating point such that E_{fd} is within the limits. Limiters and protective circuits are not modeled as they do not affect small-signal stability. The block diagram representation with exciter and AVR is shown in fig4

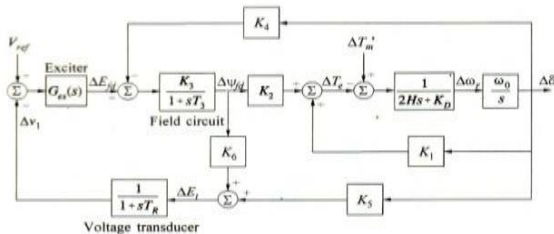


Figure 4:Block diagram representation with exciter & AVR

Change in the time response of the system for the 5% change in mechanical input depicts that the system is unstable.

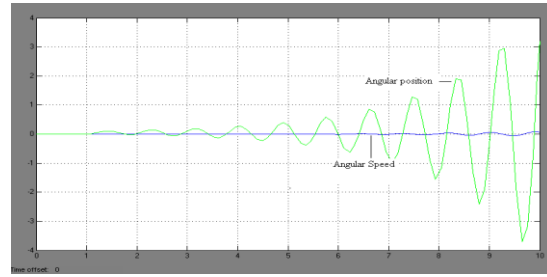


Figure 5: System Response for 5% change in input torque

With AVR, constant K_5 may have either negative or positive values as shown in figure 6, which influences the damping and synchronizing torque coefficient shown in Fig7(a) & 7(b).

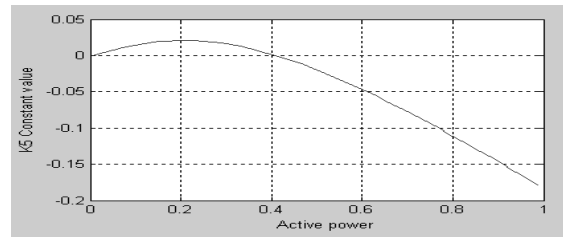


Figure 6 Variation of K_5 with per unit active power

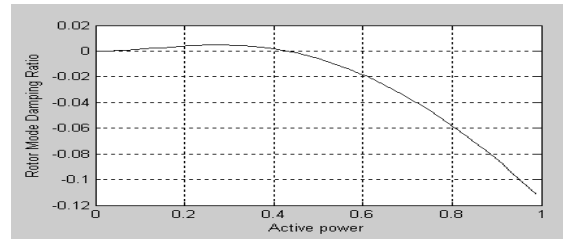


Figure 7(a)Variation of Damping Torque Coefficient with per unit power

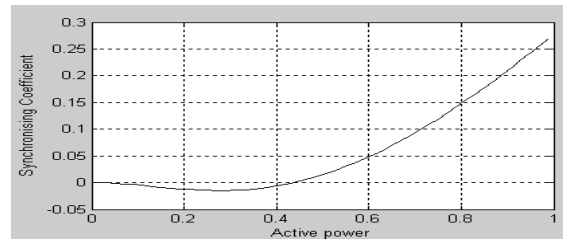


Figure 7(b) Variation of Synchronizing Coefficient with per unit power

Now the effect of variation in excitation on synchronization coefficient and damping torque coefficient is shown in fig 8(a) & 8(b).

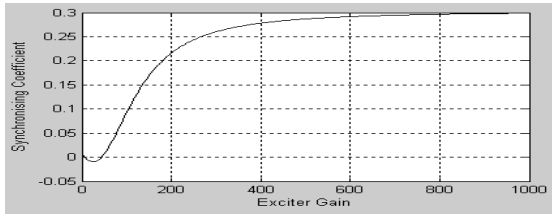


Figure 8(a) Variation of Synchronizing coefficient K_5 With K_A

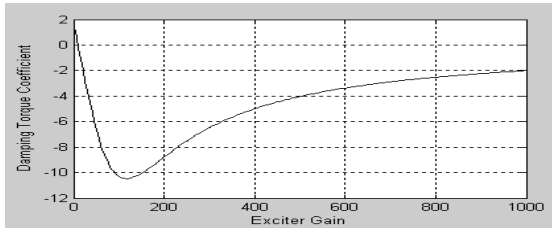


Figure 8(b) Variation of damping torque coefficient with K_A

From the above analysis, we concluded that the effect of AVR on damping and synchronizing torque component is thus primarily influence by constant K_5 and exciter gain K_A . With K_5 negative, the AVR action introduces a positive synchronizing torque component and negative damping torque component. This effect is more pronounced as exciter response increases. The main cause of instability of the system is negative damping coefficient; this adverse affect of low damping should be removed by adding damping to the system.

IV POWER SYSTEM STABILIZER

The basic function of a power system stabilizer (PSS) is to add damping to the generator rotor oscillations by controlling its excitation using auxiliary stabilizing signal(s). To provide damping, the stabilizer must produce a component of electrical torque in phase with the rotor speed deviations.

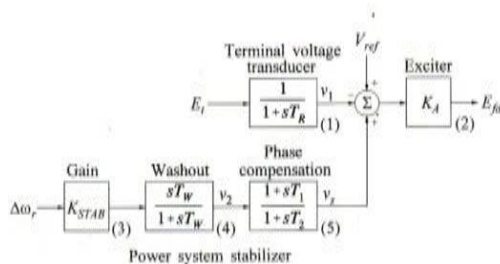


Figure 9 Thyristor excitation system with AVR and PSS

Model used in Simulink/ Matlab to examine the effect of power system stabilizer with automatic voltage regulator on single machine infinite bus system is shown below in figure 10

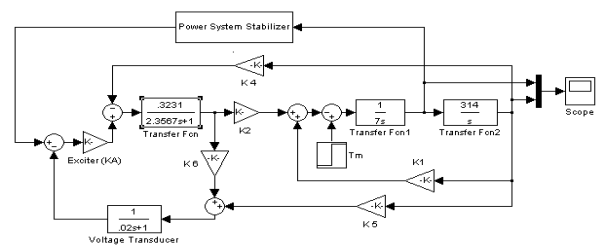


Figure 10 Simulink Model with AVR and PSS

With above constants we will analyze the variation of angular speed and angular position with time this is shown in figure 11(a) and 11(b).

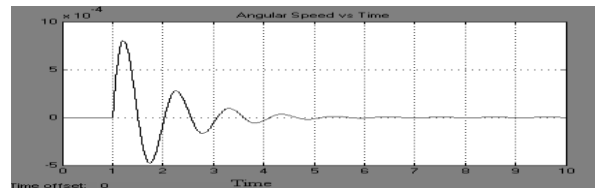


Figure 11(a) Variation of Angular Speed with Time

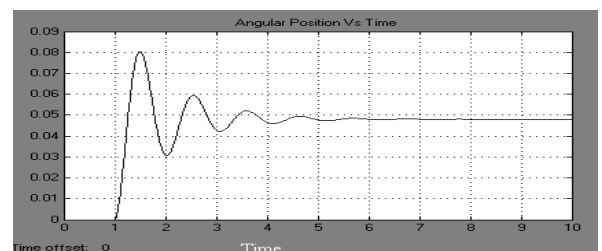


Figure 11(b) Variation of Angular position with Time

With Power system stabilizer the rotor mode damping ratio and damping coefficient increases with increase in exciter gain, This is as evident from the figure 12(a) & 12(b)

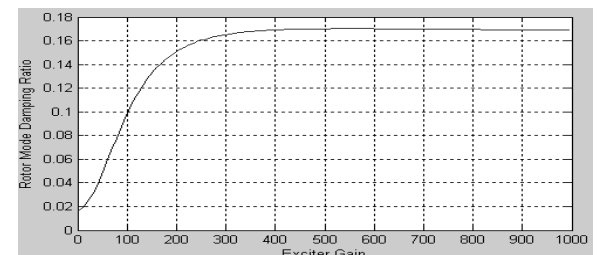


Figure 12(a) Rotor Mode Damping Ratio with Exciter Gain

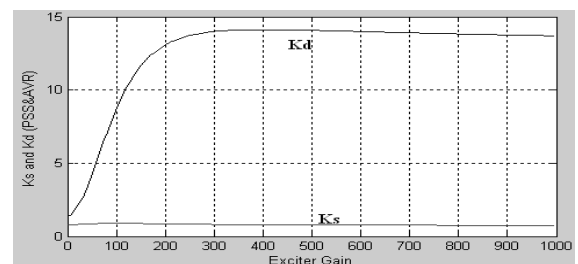


Figure 12(b) Damping and Synchronizing Torque coefficient Vs Exciter Gain for PSS

These power system stabilizers suffer from a limitation that these are not much efficient for damping small signal oscillations over wide range for operating conditions. Also it requires a deep understanding of a system, exact equations and precise numeric values.

To overcome this problem, a FPSS was developed without real-time model identification.

V. FUZZY LOGIC POWER SYSTEM STABILIZER

A fuzzy power system stabilizer (FPSS) is developed using the concept of fuzzy basis functions. The linguistic rules, regarding the dependence of the plant output on the controlling signal, are used to build the FPSS. The FPSS is designed for cogeneration, but simulation studies are based on a one machine-infinite bus model[3,6].

Selection of input and out put variables

Define input and control variables, that is, determine which states of the process should be observed and which control actions are to be considered. For FPSS design, generator speed deviation ($\Delta\omega$) and acceleration ($\Delta\dot{\omega}$) can be observed and have been chosen as the input signal of the FPSS. The dynamic performance of the system could be evaluated by examining the response curve of these two variables.

In practice, only shaft speed ($\Delta\omega$) is readily available. The acceleration signal ($\Delta\dot{\omega}$) can be derived from the speed signals measured at two successive sampling instants using the following equation:

$$\Delta\dot{\omega}(k) = \frac{((\Delta\omega(k) - \Delta\omega(k-1)))}{\Delta T}$$

The control variable is the output from the fuzzy logic controller[4].

Membership function

First input to the fuzzy controller is acceleration i.e. ($\Delta\dot{\omega}$) it is normalized by multiplying with some factor X_1 so that its value lies between -1 and 1. The membership function for acceleration is shown in figure 13(a).

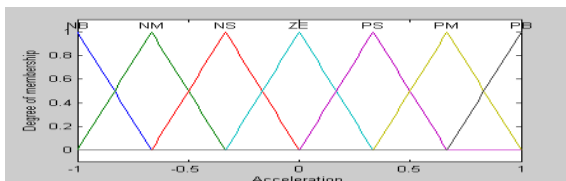


Figure 13(a) Membership Function of Acceleration

Membership function of second input i.e. Speed deviation ($\Delta\omega$) also has same membership function that of acceleration and with 50% overlap between adjacent fuzzy subsets, the function is shown in figure 13(b).

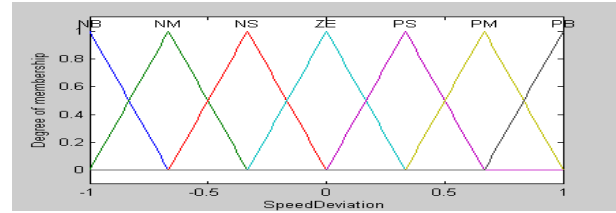


Figure 13(b) Membership Function of Speed Deviation

There is only one output from the fuzzy logic controller i.e. voltage signal, its membership function also contains seven subsets with 50% overlap the only difference is that we have to renormalized the output by multiplying the output from fuzzy controller with a constant gain

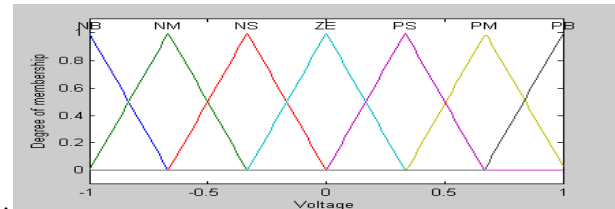


Figure 13(c) Membership function of output Voltage signal.

Defuzzification

There are five methods of defuzzification namely centroid method, bisector method, middle of maximum method, largest of maximum method and smallest of maximum method. Here in fuzzy logic controller centroid method of defuzzification was opted. The output of the controller is given by following expression:

$$u_k = \frac{\sum_{i=1}^P y_i w_i}{\sum_{i=1}^P w_i}$$

Fuzzy rule base

A set of rules which define the relation between the input and output of fuzzy controller can be found using the available knowledge in the area of designing FPSS is shown in table 1[5].

VI SIMULATION RESULTS

Model used in Simulink / MATLAB to analyze the effect of fuzzy logic controller in damping small signal oscillations when implemented on single machine infinite bus system is shown below in figure 14(a)

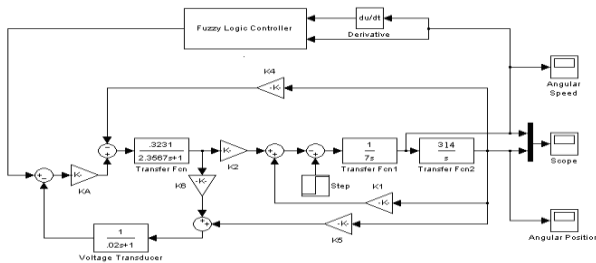


Figure 14(a) Simulink Model with Fuzzy Logic Controller

Table1: Decision Table for fuzzy logic controller

$\Delta\dot{\omega}$ $\Delta\omega$	NB	NM	NS	ZE	PS	PM	PB
NB	NB	NB	NB	NB	NB	NB	NB
NM	NB	NM	NM	NM	NM	NS	ZE
NS	NM	NM	NS	NS	NS	ZE	PS
ZE	NM	NS	NS	ZE	ZE	PS	PM
PS	NS	ZE	ZE	PS	PS	PM	PM
PM	PS	PS	PS	PM	PM	PM	PB
PB	PM	PM	PM	PB	PB	PB	PB

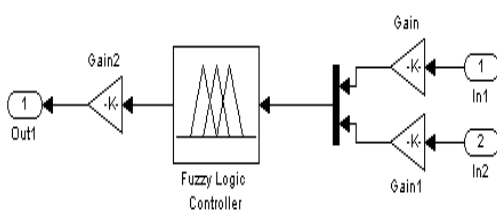


Figure14(b) Expanded form of Fuzzy logic controller block

The fuzzy logic controller block consists of fuzzy logic block and two type of scaling factors. One is input scaling factors,

these are two in number one for each input as the other is output scaling factor which determine the extent to which controlling effect is produced by the controller

Using fuzzy logic power system stabilizer it can be inferred that it does not require any complex mathematical calculations and the response with fuzzy logic is much improved than with conventional power system stabilizer. It is illustrated using the plots of angular speed and angular acceleration shown in figure 15(a) and 15(b).

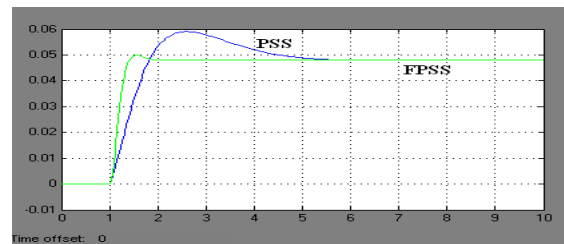


Figure 15(a) Variation of Angular position with time for PSS and FPSS

Figure15(a) shows the relative plots for variation of angular position with time for PSS and FPSS. These results are for 5% change in mechanical torque. From figure it can be perceived that with the application of fuzzy logic the rise time and the settling time of the system decreases also there is significant decrease in the peak overshoot of the system. The system reaches its steady state value much earlier with fuzzy logic power system stabilizer compared to conventional power sytem stabilizer.

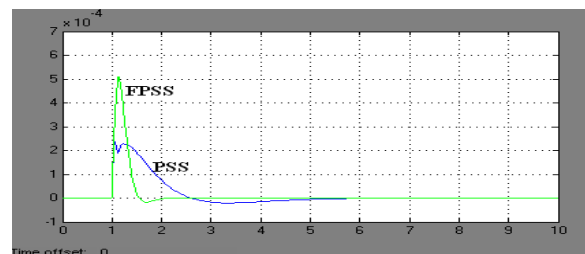


Figure 15(b) Variation of Angular speed with time for PSS and FPSS

From relative plot in fig 15(b) it can be retrieved that oscillations in angular speed reduces much faster with fuzzy logic power system stabilizer than with conventional power system stabilizer. As shown in figure with fuzzy logic the variation in angular speed reduces to zero in about 2 seconds, but with conventional power system stabilizer it takes about 6 seconds to reach to final steady state value.

VI CONCLUSIONS

In this paper FPSS shows the better control performance than power system stabilizer in terms of settling time and damping effect. The proposed FPSS produces better damping effect than PSS. To further increase the stability of the power

system a technique was introduced to tune parameters of fuzzy logic controller.

System considered

System operating condition in per unit for 4×555 MVA, 24 KV, 50 Hz generating unit on a common 2220MVA, 24 KV base are:

$$Q=P/3; \quad E_t=1.0$$

Thyristor exciter gain:

$$K_A=200; \quad T_R=0.02;$$

Frequency of oscillation is taken as 10 rad/sec.

REFERENCES

[1] Michael J. Basler and Richard C. Schaefer "Understanding Power System Stabilizer," Basler Electric Company IEEE Trans., pp. 46-67, August 2005.

[2] K.Tomsovic, "Fuzzy systems applications to power systems", IEEE PES Tutorial, July 2000.

[3] John Yen and Reza Langari "Fuzzy Logic Intelligent Control and Information" Pearson Education Publishing Company.

[4] Recommended Practice for Excitation System Models for Power System Stability Studies," IEEE Standard 421.5-1992, August 1992.

[5] Timothy J. Rose, "fuzzy logic with engineering applications", Mc – Graw Hill, Inc, New York, 1997 [6] J N Rai, Vijay Priyadarshi "Dynamic Stability Enhancement of Power System Using Fuzzy Logic based Power System Stabilizer" Proceeding of emerging trends in Electrical and Electronics Engineering, ISBN: 978-93-84869-25-0, vol I, PP 137-144, Feb. 2015.

Examining the possibility of achieving inclusive growth in India through corporate social responsibility

Upali Arijita Biswas¹ · Suresh Garg¹ ·
Archana Singh¹

© Springer Science+Business Media Dordrecht 2016

Abstract The paper attempts to evaluate whether corporate social responsibility (CSR) can contribute towards achievement of inclusive growth in India. Using content analysis of annual reports belonging to 42 non-financial companies across 8 years (2004–2011), the study identifies company disclosures under areas that fall within the purview of both inclusive growth, as conceptualised by the Indian government and themes of CSR subject indices mentioned in corporate social disclosure literature. The study is unique in its attempt to address the relation between CSR and inclusiveness through a longitudinal analysis of 322 annual reports, in a timeframe when governmental intervention regarding CSR was negligible and the disclosures were primarily voluntary. The findings revealed that company CSR efforts are in congruence with some requirements for inclusive growth, like CSR activities in areas such as health, education, infrastructure, environment and energy and can contribute to the welfare of special interest groups such as women, children and backward sections. Apart from contributing to CSR literature in India, the paper will be of interest to corporations, government and other stakeholders of the CSR process.

Keywords Corporate social responsibility · Inclusive growth · Annual reports · Content analysis · India · Corporate social disclosure · Longitudinal study

✉ Upali Arijita Biswas
upali29@yahoo.co.in

Suresh Garg
skgarg63@yahoo.co.in

Archana Singh
sarchana03@yahoo.co.in

¹ Delhi School of Management, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 42, India

Introduction

“...inclusive growth allows people to contribute to and benefit from economic growth”

(Ianchovichina et al. 2009)

In recent years, India has emerged as one of the fastest growing economies. However, the growth the country has witnessed has not permeated to every section in society. Problems of gender inequality adversely affect women, while groups such as Scheduled Castes (SC), Scheduled Tribes (ST), etc. have often not been included in growth (Planning Commission 2008, 2011a). The lack of inclusiveness can be witnessed in the slow progress of human development indicators such as literacy and education, maternal and infant mortality rates, etc., slow growth in sectors such as agriculture, slow rate of decline in poverty despite accelerated economic growth, dearth of basic amenities to a segment of the population, etc. (Planning Commission 2008). In the Eleventh Five-Year Plan (2007–08 to 2011–12), the Government of India (GoI) had focused on achieving faster and more inclusive growth (Planning Commission 2011a). In its attempt to do so, the government had initiated a number of programmes as well. However, the burgeoning challenges of urban development, such as poverty, malnourishment, etc., and global challenges, such as climate change that jeopardise the long-term sustainability of human kind, warrant a more collaborative solution, with the government, the business sector, civil society, organisations and common citizens as partners (MCA 2009). Through different initiatives, policies and programmes, across different ministries, especially the Ministry of Corporate Affairs (MCA), the government has tried to both acknowledge the contribution and further encourage the corporate sector to play a supporting role in alleviating the new social and economic problems that have emerged in the latter half of the 21st century. In 2009, the MCA released the first set of voluntary guidelines for corporate social responsibility (CSR) (MCA 2009) and initiated ‘Corporate Week’ to highlight the contribution of the corporate sector towards inclusive growth. Subsequent years have seen the Indian government introducing both voluntary guidelines (see MCA 2011) as well as regulations mandating CSR (see MCA 2013; SEBI 2011) to give impetus to CSR activities. For the increasing focus on CSR to be truly valid, it is necessary to evaluate whether CSR can be treated as a means to an end. One way to do so is to ascertain the congruence of CSR activities with areas of focus under inclusive growth. The aim of the study is to, therefore, explore how far CSR activities undertaken by companies can be instrumental in achieving inclusive growth. The research will be of relevance to the government, as it analyses the complementary nature of the governmental priority areas of CSR and inclusive growth. Apart from adding to academic literature, the study also provides an insight to corporations and their CSR partners such as NGOs, about how they can contribute to inclusive growth through their CSR efforts. The paper is structured as follows. The methodology adopted follows the literature review. The next section details the findings of the study. The paper concludes with a brief review of the findings.

Literature review

Corporate social disclosure

Corporate social disclosures (CSDs) refers to both financial and non-financial information provided in a company's annual report or other social reports about the organisation's commitment to issues such as employee welfare, environment and other social issues that affect people, (Guthrie and Mathews 1985; Haron et al. 2006). CSD has been on the research agenda for more than 35 years (Orij 2010), and over the years, three methods have largely been used to measure corporate social activities—social accounting, reputational scales and content analysis of corporate publications (Abbott and Monsen 1979). Ernst and Ernst's unpublished annual summary, in 1971, about the presence of content on specific social involvement categories such as environmental, equal opportunity, personnel, community involvement, products and their corresponding sub-categories, in the annual reports of Fortune 500 companies, is one of the earliest instances of evaluation of CSD using content analysis (Abbott and Monsen 1979; Ullmann 1985). CSD studies such as Guthrie and Parker (1989) and Gray et al. (1995a, b) continued the trend of classifying CSR in categories such as environment, customers, energy, community, charitable and political donations, employee-related categories and general CSR statements and using content analysis as the method of evaluating CSD. A similar CSR category index is visible in Hackston and Milne's (1996) work, which was based on prior works such as Ernst and Ernst in 1978 and Ng in 1985 (Aras et al. 2010). This index has found relevance in CSD studies conducted in different national contexts over the years (see Aras et al. 2010; Cowan and Deegan 2011; Tilt 2001). Mainstream CSR literature can thus be seen to broadly allude to four major themes of CSR categories: natural environment; employee; community and customers (Gray et al. 1995b), and most CSD authors have studied organisational disclosures in light of slightly different adaptations of these categories. Some authors have restricted their study to a detailed analysis of environmental disclosures as a representative of CSR information (see Montabon et al. 2007; Tilt 2001). The unobtrusive method of content analysis has also gained favour with a number of CSD researchers (see Aras et al. 2010; Cowan and Deegan 2011; Freundlieb and Teuteberg 2013; Gautam and Singh 2010; Mahadeo et al. 2011; Rahman et al. 2011).

In India, there is growing awareness and activism amongst both corporations and their stakeholders regarding the corporations' responsibilities, which have been driven both internally by corporate will and externally by increased governmental and public expectations (Mohan 2001). A large number of corporations such as Tata Group, Birla Group, Bajaj, ITC, Mahindra and Mahindra, Hero Honda, etc. are increasingly contributing in the fields of education, healthcare services, rural infrastructure, community welfare, environment protection, relief and emergency assistance, preserving art, heritage, etc. and are supported and encouraged by the Chambers of Commerce and other industry associations in India (Das Gupta 2007). With organisations successfully incorporating CSR initiatives, the concept of CSR has grown into an industry in itself, with full-time managers, websites, newsletters and with several pages devoted to the company's social goals and good works undertaken in annual reports of many major organisations (Agarwal 2008). With increasing amount of data being made available about CSR in India, extant CSR literature has explored CSR communication and

reporting in India from different perspectives (see Alon et al. 2010; Dhanesh 2014; Gautam and Singh 2010; Jammulamadaka 2013; Narwal and Singh 2013; Planken et al. 2010). However, CSR has seldom been explored in the context of inclusive growth and this study aims to bridge this gap.

Inclusive growth

In the Eleventh Five-Year Plan, the Indian government adopted multi-dimensional economic and social objectives of inclusive growth that could be classified in six major categories of income and poverty, education, health, women and children, infrastructure and environment, where results could be easily monitored and were directly in the central government's purview (Planning Commission 2008). Other areas such as agriculture, sex ratio, etc. were considered as state-monitored subjects. The approach paper to the twelfth five-year plan (2012–2017), while discussing the progress in the direction of inclusive growth in the eleventh plan and the possibilities in the next plan, further noted that (Planning Commission 2011a):

- Poverty reduction was imperative for growth and it could be achieved by improving incomes and creating adequate livelihood opportunities for a growing labour force.
- Apart from governmental policy and programmes, institutional and attitudinal changes are also required to successfully bring about inclusive growth.
- To ensure education to all, as proposed by the right to education (RTE), more educational facilities would be required.
- The suggestions by the expert group on low carbon strategy for inclusive growth were effective initiatives in the fight against climate change. The committee provided insights about curtailing environmental problems through a host of issues from the perspective of demand and supply, such as in the power sector. Other considerations involved the investment in renewable technologies, particularly solar, wind, second generation biofuels and hydropower, providing sector-specific emission recommendations, highlighting National Missions for integrated management of water resources, sustainable habitat including the introduction of sustainable transport systems, enhanced energy efficiency, strategic knowledge for climate change, etc. (Planning Commission 2011b)

The approach paper also noted that for inclusive growth, there needed to be substantial improvements in the provision of basic amenities such as water, electricity, roads, sanitation and housing, and the needs of SC, ST and other backward caste (OBC) population, women, children and minorities required particular attention (Planning Commission 2011a).

Methodology

To fulfil the objectives of the study, the paper uses content analysis of annual reports to analyse the type of CSR activities undertaken by companies in India. The broad concept of inclusive growth has not been assessed directly as a part of CSR disclosure research. A review of government policy documents helped identify the areas

considered, by the government, as a part of inclusive growth and extant CSD studies provided the guidelines to enable the study of whether, and the extent to which, these areas were represented as a part of CSR disclosures in the annual reports of the companies under consideration. The sample used in this research and the method of content analysis is discussed in the following sub-sections.

Content analysis

The technique of content analysis can be used to make replicable and valid inferences (Krippendorff 1981) about an organisation's social and environmental disclosures (Gray et al. 1995b; Milne and Adler 1999) by codifying qualitative and quantitative information into pre-defined categories (Guthrie and Abeysekera 2006). The technique is unobtrusive and can cope with large volumes of data (Krippendorff 1981). According to Raman (2006), content analysis requires the researcher to identify an appropriate document for analysis, a unit of analysis and themes, in which, to classify the content. For this study, annual reports of select organisations were chosen as the document for analysis, since they have served as a source of raw data for previous social and environmental reporting studies (see Aras et al. 2010; Gao 2011; Gray et al. 1995b; Guthrie and Abeysekera 2006; Hackston and Milne 1996; Haron et al. 2006; Raman 2006). Further, annual reports provide the organisation's construction of its own social imagery (Gray et al. 1995b), in which an organisation consciously presents and omits content to be presented to its stakeholders (Guthrie and Abeysekera 2006).

Weber (1990) proposed six possible classifications of unit of text—word, word sense, sentence, theme, paragraph and whole text—to be used in content analysis. Of these, sentences are used as the unit of analysis for the study since they are most reliable and have been most commonly used in content analyses of CSD (Milne and Adler 1999). Sentences overcome the unreliability that can occur with page measurements not accounting for different page sizes, font sizes or margins (Ng, 1985 in Aras et al. 2010; Raman 2006) and words losing meaning when considered out of context (Raman 2006).

Some areas of focus regarding inclusive growth were obtained from the review of complementary documents to the 5-year plans and have been discussed in the literature review. The major themes of classification used in CSD research, such as that of Hackston and Milne (1996), include employee health and safety, employee others, product, environment, energy, community and general CSR policies, with a number of sub-categories listed under each theme. The themes for analysis in the study (Table 1) are based on congruence of ideas in Hackston and Milne's (1996) CSR subject index and areas of inclusive growth identified by the Planning Commission (Planning Commission 2008, 2011a).

Agriculture was not a part of Hackston and Milne's (1996) work but was included due to the important role it plays in the economy and its propensity, along with employment and livelihood generation, to contribute to poverty reduction, a major factor for inclusive growth (see Planning Commission 2008). It notes special initiatives for farmers, animal husbandry, etc. Sub-categories of classification for themes such as energy, environment and employment and livelihood and the decision rules for determining CSD were adapted from Hackston and Milne (1996). Sub-categories corresponded well with discussions on inclusiveness in government reports (see

Table 1 Selection of themes for analysis in the study

Theme for analysis in the study	Similar theme listed under		
	Hackston and Milne (1996)	Planning commission (2008)	Planning commission (2011a)
1. Employment and livelihood	Employee others (opportunities for women/minorities), community	Employment and livelihood	Employment and livelihood
2. Agriculture	No specific mention	Agriculture ^a	Agriculture ^a
3. Health	Employee health and safety, community	Health	Health
4. Education	Community	Education	Education
5. Infrastructure development	Community	Infrastructure development	Infrastructure development
6. Energy	Energy	Energy	The energy challenge
7. Environment	Environment	Environment	Natural resource management—water, land and forests

^a State subject. The rest fall within the purview of the central government

Planning Commission 2008, 2011a, b). Owing to increased governmental focus, initiatives pertaining to renewable energy were also included along with sub-categories covered in Hackston and Milne (1996), in the energy category. Disclosure in any of the categories mentioned above or their sub-categories, by a company in its annual report, was noted by the presence of sentences divulging the company's CSR activities in that area. The content analysis also noted whether CSR activities were carried out for specific beneficiaries who are of primary interest to the government, such as women, children, elderly, physically challenged (PC), backward sections (BSs) or the neighbouring community (NC). NC refers to areas adjacent to company operations where the CSR activity has the propensity to affect all beneficiary groups living in the area, and the company has not distinguished any beneficiary group for their CSR programme. BS refers to socially and economically backward sections such as SC, ST and OBC, etc. Activities undertaken specifically for women, children, elderly, PC and BS, whether in the NC or the rest of the country, were catalogued separately. The study briefly reviews activities initiated for employees, under the themes of classification identified, since a number of employee-related themes are indexed as a part of the CSR index proposed by Hackston and Milne (1996). Activities undertaken for the board of directors were not considered. The high degree of congruence of themes between CSD studies and inclusive growth, in itself, highlights the possibility of CSR programmes contributing towards greater inclusiveness in the country.

Sample for content analysis

Annual reports of 42 non-financial companies listed on the Nifty index of the National Stock Exchange (NSE) in India, as on March 31, 2011, were selected for evaluation in

the timeframe 2004–2011. The companies chosen are traded regularly and have a large market capitalisation, with operations often spanning across the country, and therefore likely to indulge in CSR activities to retain a positive outlook from the public and government. Companies belonging to industries such as automobile, cement and cement products, construction, consumer goods, energy, industrial manufacturing (IM), IT, metals, pharmaceuticals (pharma) and telecom were considered in the study. Eight financial companies were not considered owing to slightly different reporting norms. The choice of the 8-year period (2004–2011) was based on the availability of annual reports as well as on the lack of any attempt to institutionalise CSR prior to 2004. Further, the timeframe covered the era leading up to the introduction of mandatory CSR directive for Central Public Sector Enterprises (CPSE) by the Department of Public Enterprises (DPE) (DPE 2010) and for the top 100 listed companies by the Securities and Exchange Board of India (SEBI) (SEBI 2011) and consequently covers the companies' idea of CSR, as implemented voluntarily. Post 2012, the pattern of reporting has been changing owing to directives in quick succession, which encourage companies to better align their CSR programmes with growth objectives. Companies are likely to build on their preliminary work to ensure compliance with emerging requirements and parameters. In totality, 322 annual reports were analysed since not all the companies existed or were listed on any stock exchange in certain years. Hence, their annual reports, corresponding to these years, were unavailable on both, company websites and report databases.

The standard reliability statistic for content analysis Krippendorff's α (Hayes and Krippendorff 2007), which has been used in previous studies (see Aras et al. 2010; Cowan and Deegan 2011; Hackston and Milne 1996), was calculated to ensure inter-coder reliability. Three coders performed content analysis on a report on the basis of rules devised by Hackston and Milne (1996), to analytically ensure reliability of what could be described as CSD. The value of Krippendorff's α (0.946) displayed acceptable level of reliability. The following section discusses the findings of the study.

Findings

The analysis of CSR content mentioned in annual reports reveals that despite the lack of mandatory requirements from stakeholders such as the government, a number of initiatives have been undertaken by companies that can contribute significantly to inclusiveness. A company was considered to have contributed towards the welfare of a group if it was found to undertake at least one activity (education, health, etc.) for the group. Table 2 identifies BS, NC, women, children, elderly and PC as the beneficiary groups, for which companies have directed certain CSR programmes. These groups serve as special interest groups for the study. Of these special interest groups, NC seemed to have received the most attention from the companies over the years (Table 2). Activities mentioned with respect to NC are likely to benefit all categories of special interest groups. CSR activities directed towards children and BS have also gained prominence over the years, while limited activities were undertaken by the companies for the elderly and PC. The following tables further explore the CSR initiatives undertaken by companies for these special groups as well as activities undertaken under each theme of analysis discussed in the methodology.

Table 2 Beneficiaries explicitly mentioned in CSR disclosures

Year	NC	BS	Women	Children	Aged	PC
2004	66.67	38.89	47.22	47.22	25.00	11.11
2005	66.67	56.41	43.59	61.54	10.26	10.26
2006	70.00	45.00	40.00	57.50	12.50	5.00
2007	82.93	60.98	65.85	75.61	14.63	2.44
2008	82.50	55.00	52.50	70.00	22.50	2.50
2009	83.33	47.62	50.00	78.57	26.19	7.14
2010	92.86	73.81	54.76	73.81	11.90	11.90
2011	95.24	78.57	57.14	78.57	11.90	9.52

Employment and livelihood

Mandatory allocation of percentage of jobs for BS has been a source of contention for private companies in India. While that debate is beyond the purview of the paper, it takes note on whether companies have mentioned BS employment and initiated programmes for them as a part of their CSR activities. Table 3 discusses the role of companies in providing special provision for the employment opportunities for special interest groups, such as BS, women and PC, within the organisations as well as allocating CSR resources to train and provide help to generate livelihood opportunities to empower them. Post 2006, there is some consistency in reporting about recruitment of BS and PC, but not in employment of women. According to the Presidential Directives and other guidelines issued from ministries such as the Ministry of Social Justice and Empowerment and the DPE, it is compulsory for public sector units (PSUs) to report on reservations in recruitment of SC, ST and OBC and special interest groups such as those with physical disabilities. All PSUs seemed to abide by these rules but recruitment of women received lesser attention. Some private enterprises mentioned being equal opportunity employers, fewer still explicitly mentioned recruitment opportunities provided to BS or women and

Table 3 Percentage of companies reporting about activities to encourage employment and livelihood opportunities for different sections of the population

	Recruitment by company			Providing training/helping set up livelihoods	
	BS	Women	PC	Total reported	Specifically for women
2004	13.89	5.56	16.67	41.67	16.67
2005	12.82	5.13	10.26	46.15	10.26
2006	15.00	5.00	12.50	47.50	15.00
2007	24.39	17.07	12.20	53.66	24.39
2008	22.50	25.00	15.00	65.00	27.50
2009	16.67	28.57	9.52	69.05	30.95
2010	19.05	7.14	9.52	64.29	7.14
2011	19.05	9.52	11.90	71.43	7.14

recruitment of PC received negligible mention. The number of private sector enterprises reporting about employment of BS and women, however, saw an increase after 2006, though most did not furnish any statistics supporting their claims.

Despite companies showing less enthusiasm in formulating special provisions for recruitment of the interest groups, an increasing number (71 % in 2011) of companies provided either training and/or provided infrastructural support such as vocational training, repair tool kits, etc. to youth. Comparatively, fewer opportunities are made available to women, either through employment or support such as training, women development centres or empowering women self-help groups, to enable them earn their livelihoods (Table 3). Livelihood opportunities were discussed primarily in the context of youth and women. Some companies mentioned training PC persons and youth belonging to BS from around their areas of operations and then absorbing them in the organisation. These cases were considered as recruitment by company in Table 3. While companies from most industries reported on providing training and livelihood opportunities as a part of their CSR, IM, cement and consumer goods, industries appeared to give it greater attention.

Agriculture

The agricultural sector has not witnessed much progress in India, although a large section of the population is involved in it. Table 4 briefs about the nature of support received by farmers as a part of company CSR programmes. Corporate support was described in the form of monetary data such as financial aid given to farmers and in a few cases, the income gained by farmers because of this support. Such financial details were inconsistently reported (Table 4). Numerical information about number of farm equipment and farm animals provided and number of beneficiaries has been increasing over the years. Declarative information refers to information that is neither in monetary terms nor provides numerical information. Declarative information pertained to issues such as providing assistance to famers about crops and developing watershed and irrigation projects. Such information is also on the rise. Companies from industries like consumer goods and cement reported supporting agriculture-related issues prominently over the timeframe of the study, while some telecom, metal and automobile companies reported their initiatives only in later years.

Healthcare

Tables 5 and 6 show the two-pronged approach adopted by companies of providing a healthy workplace and medical facilities for employees and their families and

Table 4 Percentage of companies reporting about activities to assist agriculture in the country

Type of information about agricultural support:	2004	2005	2006	2007	2008	2009	2010	2011
Crops/watershed/irrigation	25.00	10.26	20.00	21.95	20.00	26.19	35.71	33.33
Numerical information	19.44	10.26	15.00	21.95	22.50	26.19	35.71	30.95
Financial assistance	0.00	0.00	2.50	4.88	5.00	7.14	11.90	2.38

Table 5 Percentage of companies reporting about activities to provide healthcare to communities

Year	Total percentage reporting on public health	Of the total number of companies reporting on public health, percentage of companies reporting about:							Total percentage reporting on medical research
		Financial support	BS	NC	Women	Children	Elderly	PC	
2004	50.00	2.78	19.44	55.56	27.78	25.00	22.22	11.11	8.33
2005	48.72	2.56	23.08	46.15	20.51	20.51	7.69	10.26	5.13
2006	52.50	7.5	17.50	60.00	25.00	30.00	12.50	17.50	7.50
2007	68.29	9.76	34.15	68.29	26.83	29.27	14.63	14.63	4.88
2008	72.50	5	30.00	70.00	25.00	35.00	20.00	22.50	7.50
2009	73.81	4.76	23.81	73.81	23.81	35.71	19.05	28.57	4.76
2010	76.19	2.38	40.48	80.95	26.19	38.10	9.52	23.81	7.14
2011	66.67	11.90	42.86	80.95	42.86	30.95	9.52	23.81	7.14

supporting healthcare initiatives for special interest groups, which can make a significant difference in making healthcare available to a broad spectrum of people. Table 5 displays the total percentage of companies undertaking public health initiatives and supporting medical research as a part of their CSR activities and also identifies the percentage of companies working for different interest groups. Companies across different sectors reported sponsoring different healthcare initiatives as a part of their CSR activities. Companies mentioned providing healthcare facilities mostly for NC through initiatives such as health check-up camps and donating mobile medical units to NGOs to expand availability of health services to the underprivileged in both rural and urban areas (Table 5). Maternal and infant care also received impetus in the form of sanitation programmes and health camps. Apart from sponsoring medical facilities for children in hospitals, companies contributed towards providing a balanced meal to children, particularly through Aanganwadis and the Mid-Day Meal (MDM) scheme (for school-going children), to combat malnutrition amongst children. Healthcare for BS received mention but increased substantially in 2010 and 2011. Initiatives for the

Table 6 Percentage of companies reporting about activities to safeguard health of employees

Year	Reducing hazards, pollutants at the workplace	Promoting physical/mental health of employees	Health and safety standards implemented	Safety committee / safety training	Low-cost healthcare for employees
2004	47.22	8.33	38.89	47.22	13.89
2005	48.72	7.69	35.90	35.90	15.38
2006	55.00	12.50	47.50	45.00	20.00
2007	51.22	12.20	56.10	58.54	19.51
2008	60.00	17.50	50.00	57.50	22.50
2009	71.43	28.57	54.76	57.14	19.05
2010	64.29	23.81	57.14	47.62	19.05
2011	66.67	23.81	64.29	57.14	30.95

elderly were least reported and mostly restricted to clinics for cataract and ocular implants. Similarly, initiatives for PC were related to providing them with implants and prosthetics, although few companies have consistently supported PC. Support for mental disabilities was seldom mentioned. Table 6 also investigates some organisational health and safety initiatives undertaken for employees. Companies across industries have consistently mentioned about making the workplace safe by removing various types of work hazards and through the formulation of safety committees and imparting safety training. To make the work environment safe, companies are increasing efforts to control hazardous substances at work and providing safety gears to employees. Low-cost healthcare facilities for workers, their families and retired staff and disclosures of company policy and intent to promote employee physical health or mental health through counselling also received mention but to a lesser extent. In general, Table 6 indicates that employee health and safety is receiving greater attention from employers.

Education

Table 7 represents the initiatives undertaken by companies to educate employees as well as the special interest groups considered in the study. Most companies claimed to impart job-related training to their employees and some assisted their employees to attain higher education at different institutes within the country and abroad. Barring 2004, more than 50 % of the companies extended their support to educational activities, such as supporting schools, libraries, donating books, computer centres for children, providing financial assistance, offering scholarships to children to continue education, providing training to qualify for higher educational intuitions such as IITs, etc., as a part of their CSR. Initiatives spanned across the country though a lot of these activities were centred in communities around company operations. Education has received significant support from companies across industries and in particular, consistent support from the IT and metal industries. Education for children, particularly in rural areas, has been a

Table 7 Percentage of companies reporting about activities to encourage education amongst employees and communities

Year	For employees		Total percentage reporting on Education	Of the total number of companies reporting on education, percentage of companies reporting about:						
	In-house training	Higher education		Financial support	BS	NC	Children	Women and female child	Adult	PC
2004	72.22	33.33	41.67	13.89	19.44	25.00	41.67	25.00	11.11	11.11
2005	71.79	25.64	51.28	2.56	33.33	17.95	58.97	20.51	10.26	0.00
2006	70.00	25.00	57.50	12.50	25.00	35.00	50.00	12.50	5.00	7.50
2007	80.49	24.39	68.29	17.07	31.71	36.59	70.73	39.02	2.44	9.76
2008	77.50	35.00	60.00	5.00	37.50	45.00	62.50	35.00	2.50	20.00
2009	71.43	35.71	64.29	14.29	38.10	50.00	71.43	28.57	7.14	16.67
2010	83.33	52.38	69.05	9.52	64.29	52.38	64.29	42.86	11.90	11.90
2011	83.33	42.86	59.52	16.67	69.05	66.67	61.90	35.71	9.52	16.67

priority area (Table 7). Education for the daughter and young women was also encouraged mostly through scholarships and funding schools for them. Initiatives for adult education received scant mention, while special initiatives for BS, particularly in areas dominated by them saw a significant increase in 2010 and 2011. Education and healthcare for children saw significant support in the form of contribution to the development of new infrastructure such as hospitals and schools. This is reflected in the column dedicated to infrastructure for children in Table 8.

Community infrastructure

Community development has been detailed through information about infrastructure in an overwhelming number of annual reports as indicated by Table 8. Over the years, a lot of this effort has been focused on neighbourhood infrastructure. This commonly included construction and repair of community structures such as halls, wells and tube wells for drinking water, roads, etc. Making such basic amenities available to the BS and people in remote rural India has also been given some consideration. This is significant since BS have been denied the opportunity to avail many facilities available to other members of society, owing to dogmas and taboos that India is trying to overcome, and they have remained poor. Directing funds for community infrastructure, schools or hospitals in villages with a majority BS population enables them to avail the advantages of development. Support to old-age homes has received little attention and PC barely fair better. Not much is mentioned about making areas handicap accessible even for employees or supporting institutions set up specifically to cater to the special needs of PC. In general, community infrastructure has featured prominently in the CSR efforts across different industries.

Energy

Table 9 shows that most companies have revealed steps taken to conserve energy, essentially under the influence of Section 217(1)(e) of the Companies Act, 1956 that mandates the filing of the details regarding energy conservation and technology

Table 8 Percentage of companies reporting about activities to develop community infrastructure

Year	Total percentage reporting on community infrastructure	Of the total number of companies reporting on community infrastructure, percentage of companies reporting about:					
		Financial support	BS	NC	Children	Elderly	PC
2004	63.89	66.67	19.44	47.22	13.89	2.78	2.78
2005	71.79	64.10	0.23	0.56	0.26	0.03	0.03
2006	67.50	62.50	25.00	57.50	20.00	5.00	2.50
2007	73.17	56.10	39.02	63.41	31.71	0.00	7.32
2008	77.50	57.50	27.50	55.00	25.00	2.50	5.00
2009	78.57	61.90	26.19	64.29	28.57	7.14	9.52
2010	90.48	71.43	47.62	66.67	28.57	2.38	9.52
2011	90.48	66.67	38.10	71.43	40.48	2.38	4.76

Table 9 Percentage of companies reporting on issues related to energy

Activities	2004	2005	2006	2007	2008	2009	2010	2011
Energy conservation	83.33	71.79	85.00	82.93	85.00	83.33	90.48	88.10
Energy from waste	11.11	15.38	20.00	29.27	22.50	28.57	26.19	28.57
Energy efficiency of products	8.33	7.69	2.50	7.32	25.00	16.67	26.19	21.43
Renewable energy initiatives	22.22	20.51	22.50	31.71	42.50	59.52	38.10	57.14

absorption, in all annual reports. Some service sector firms, such as IT firms, noted that due to the nature of their activities, which did not consume much power, the reporting directives did not apply to them. IT companies, however, stated providing green and energy efficient solutions to their customers. Company disclosures, in the area of energy, are predominantly narrative and inform about projects undertaken. The amount of energy savings or number of units involved were also disclosed but to a lesser extent. The government has advocated the fuelling of energy needs by using low carbon emission technology and by encouraging renewable energy through programmes such as the National Solar Mission (Planning Commission 2011a, b). Table 9 reflects the increasing focus on renewable energy sources, particularly since 2008. Adoption and propagation of non-conventional sources of energy, particularly solar, hydro-electric and wind power, was reported by companies across sectors such as energy, consumer goods, construction, cement and IM. Energy from waste has received maximum impetus from cement industries while receiving sporadic mention from consumer goods and metals industries. IM companies mostly reported information about energy efficiency of products.

Environment

Table 10 shows how the companies have addressed environmental concerns. Pollution control has been an important issue for all types of industries, and the reporting in this

Table 10 Percentage of companies reporting on activities undertaken to protect the environment

Initiatives affecting:	Percentage of companies reporting in a year							
	2004	2005	2006	2007	2008	2009	2010	2011
Pollution control	72.22	74.36	77.50	87.80	85.00	88.10	85.71	90.48
Environment monitoring	19.44	20.51	30.00	14.63	20.00	28.57	28.57	42.86
Resource reduction—recycling water, etc.	52.78	56.41	55.00	60.98	65.00	69.05	66.67	61.90
Prevention of environmental damage/ reforestation	33.33	33.33	42.50	36.59	52.50	45.24	64.29	61.90
Wild life conservation	0.00	7.69	5.00	4.88	2.50	7.14	9.52	11.90
Support to historical structures/parks	2.78	12.82	5.00	7.32	15.00	16.67	16.67	16.67
Waste management	30.56	28.21	35.00	41.46	55.00	47.62	23.81	16.67
Sustainable buildings	27.78	15.38	17.50	26.83	37.50	40.48	42.86	30.95

regard has increased over the years. There is an increasing focus on environment monitoring, which includes companies monitoring the environmental changes due to company activity, emission, pollution, etc. by setting up monitoring centres at different places and particularly in areas where the impact of company activity is likely to be high. Water conservation was an important aspect of conservation. Since 2007, more than 60 % of the companies, considered in this study, have made provisions for recycling water and other liquids and invested in processes which use water efficiently. The area has received attention from all industries, though sporadically from the IT, telecom and pharma sectors. Few companies also mentioned about working towards becoming water positive. The number of companies participating in reforestation, through plantation over closed mines, reclamation of used territory or development of a green belt around operations, has gradually increased (Table 10). Wildlife protection was propagated, mostly by the metal industry, through awareness generation and financial assistance, though the amount is rarely mentioned. Some companies support the conservation of historical structures and parks around them. Cement companies predominantly reported about waste management initiatives such as using fly ash generated during production into bricks, etc. that could be used elsewhere. Metals, automobile and consumer goods companies sporadically reported on the issue and their efforts included reducing waste, recycling metal, developing alternate solutions to landfills, etc.

Support to government schemes

The Eleventh Five-Year Plan (2007–2012) gave special impetus to several programmes aimed at building rural and urban infrastructure and providing basic services with the objective of increasing inclusiveness and reducing poverty. The CPSE guidelines encourage the PSUs to synergise with initiatives of state governments, district administrations, local administration as well as central government departments or agencies, self-help groups, etc. (DPE 2010, 2013). The Indian government has identified 13 flagship programmes, including the MDM scheme, to facilitate inclusive growth in the country (Planning Commission 2011a). The analysis of the annual reports (Table 11) yielded that apart from participating in various community development programmes around the regions they operate in, through either cash, products or employee hours, both PSUs and private sector organisations also fulfil their social obligations by supporting some governmental flagship programmes for inclusiveness. Supporting the MDM scheme for children is a popular choice of private and public companies. Apart from the central government schemes, companies worked with state governments and panchayats at the village level. Both public and private sector companies also contributed to the Prime Minister's Relief fund or State fund, particularly at times of great natural or human tragedy. Table 11 indicates that barring 2010, the percentage of public and private sector companies, working with governmental schemes at different levels, is quite similar. The reason for the discrepancy between the percentage of public (87.5 %) and private (55.88 %) sector companies, supporting governmental schemes in 2010, cannot be attributed to CPSE guidelines since the effects, if any, should have been visible in the annual report of 2011 though that is not the case. The increase in reporting may be considered to arise from the anticipation of forthcoming governmental guidelines or attributed to sporadic reporting by companies.

Table 11 Percentage of public and private sector companies supporting government schemes for inclusive growth (central programmes for inclusiveness and poverty reduction)

Year	Company type	Companies supporting government sponsored initiatives	
		As percentage of PSUs and private companies	Total percentage of companies
2004	PSU	28.57	27.78
	Private	27.59	
2005	PSU	12.5	12.82
	Private	12.9	
2006	PSU	25	22.5
	Private	21.88	
2007	PSU	37.5	36.59
	Private	36.36	
2008	PSU	50	37.5
	Private	34.38	
2009	PSU	37.5	45.24
	Private	47.06	
2010	PSU	87.5	61.9
	Private	55.88	
2011	PSU	50	45.24
	Private	44.12	

All the PSUs considered in the study have operations in economically backward rural areas where inhabitants cannot access basic amenities. Some PSUs, such as the Power Grid Corporation of India and NTPC, by virtue of their operations being related to power distribution, have been contributing to governmental projects such as Rajiv Gandhi Gramin Viduyati Karan Yojana (RGVKY) and Accelerated Power Development and Reforms Programme (APDRP) by facilitating electrification of rural areas. Some PSUs also expressed support for the Mahatma Gandhi National Rural Employment Guarantee scheme (MGNREGA) which aims to provide employment to the rural poor. PSUs and private sector organisations, involved in industries such as metals and mining, mentioned providing rehabilitation and relocation to villages that might have been affected by their operations. Some companies claimed they were helping villagers develop the new areas by also providing support for housing, community infrastructure, health, etc. The CSR reporting of the PSUs indicated that towards 2010 and 2011, more and more activities included the interests of weaker and underprivileged sections of society.

Discussion and conclusion

Inclusive growth demands a focus on productive employment rather than increasing incomes for excluded groups through means such as income redistribution (Ianchovichina et al. 2009) and therefore makes sustainable livelihoods a necessity

for inclusiveness. The study revealed that companies were attempting to provide livelihood opportunities either by making special provisions to employ BS, like in the case of PSUs, or imparting vocational training and providing infrastructural support, particularly to youth and women, to help gain useful employment. By discussing initiatives for employees, the study also identified some company activities for BS, PC, women employees, which in turn can provide an impetus to improving the quality of life for these special interest groups. Company support extended to areas such as education, health and community infrastructure that have benefited women, children and BS, not only in NC but also across the nation. Companies have helped farmers acquire information about crops and have provided equipment, farm animals and much needed irrigation facilities. Energy and the environment have also received attention, with increasing number of companies making an effort to embrace energy conservation, renewable energy sources, reduce pollution and reduce resource usage. While identifying the CSR initiatives that could contribute to inclusiveness, the study found that not all companies had CSR programmes that contributed to the different aspects of inclusive growth considered in the study (education, health, etc.). Neither did all companies cater to the interests of all the special interest groups, for example, companies showed much less enthusiasm for activities for elderly and PC. Despite such shortcomings, the study successfully highlights the congruence of CSR activities with the goals of inclusive growth and that the nature of company CSR efforts, though limited in its current form, has the potential to yield substantial results. With regulatory, instrumental and mimetic forces such as governmental regulations mandating CSR spending and reporting, growing awareness about the advantages of being proactive about CSR and competitors implementing CSR programmes, it is likely that the CSR programmes, facilitating inclusiveness, will become prominent. However, it is also imperative to discuss some of the factors, the content analysis of 322 annual reports revealed, as hindering CSR from playing a more significant role in the area of inclusive growth. The patterns of reporting suggested the following factors:

- Lack of impact monitoring mechanisms—Companies did not mention having an impact assessment mechanism, which reduces the chance of improvement and monitoring progress. This in turn give the impression that CSR programmes are implemented without much thought or enthusiasm and with a low sense of ownership that might impede any prospects for development through CSR activities.
- Limited input from local population or knowledgeable groups—Some companies mentioned implementing their CSR programmes with NGO partners, but the exact nature of association remained ambiguous. Limited input from proposed beneficiaries about the nature of their problems can render programmes ineffective since companies may direct focus to less significant matters.
- Necessity to link CSR and inclusive growth—The government has advocated CSR as well as inclusive growth, but has failed to provide a clear picture of how CSR could serve as a means for inclusive growth. Few companies have acknowledged that they could contribute to mitigating social problems hindering inclusiveness and

provided some insight about the CSR initiatives undertaken in this regard. Greater awareness and a more consolidated effort will be constructive.

- **Support to generic activities**—Focus on generic activities (Porter and Kramer 2006), which companies have no particular expertise in, cause these activities to be treated as one-off or philanthropic ventures. Such initiatives do not utilise organisational core competencies in helping communities, which is a waste of resources. Most companies made generic statements about supporting numerous different areas. Inclusive growth can only be achieved through a long-term perspective (Ianchovichina et al. 2009); however, few initiatives appear strategic and most programmes lacked long-term planning and structure.
- **Ambiguous reporting**—Some companies provided the same information over a number of years. Monetary and other statistical details about most activities were not disclosed. The lack of transparency also impedes the analysis of CSR activities as well as judging the kind of impact these activities have had.
- **Extending CSR reforms to partners in the value chain**—Companies seldom mentioned extending CSR expectations to their vendors, retailers or suppliers. A concerted and consolidated effort across industries and value chains may be helpful.
- **Government, industry and society partnership** needs to be developed further for better results. The impact of new rules and possible changes in attitude need to be monitored.

Limitations and future research

The current study restricted itself to the analysis of company annual reports due to paucity of time and resources and because all companies, under consideration, did not provide any standard CSR report such as a GRI report, etc. With new regulations regarding CSR in place, annual reports since 2012 have a Business Responsibility Report section, to report CSR activities, and an increasing number of companies are providing stakeholders with CSR information through a separate CSR or GRI report. Future research can consider other documents as a source of information for CSR activities. The pattern of reporting and focus is likely to become clearer in the next few years when a contrast between work prior and post implementation of the mandatory regulations for CSR could be evaluated. The authors assumed the CSR activities reported by the companies to be accurate. Studies can replicate the methodology adopted and investigate other aspects of inclusive growth, extend the study to micro, small and medium sector enterprises or explore the potential of CSR for inclusive growth in the context of different countries. The paper does not address the degree of inclusiveness possible through CSR, rather provides an insight that CSR activities of top organisations, even without external compulsions like those imposed by Companies Act, 2013, are, at the least, capable of increasing inclusiveness. This study shows that the potential of CSR aiding inclusive growth is far from truly harnessed and requires greater collaboration between all stakeholders concerned.

Appendix 1

Table 12 List of companies considered for the study

Sr. no.	Company name	Industry type (base on NSE)
1.	ACC Ltd.	Cement and cement products
2.	Ambuja Cements Ltd.	Cement and cement products
3.	Bajaj Auto Ltd.	Automobile
4.	Bharat Heavy Electricals Ltd. ^a	Industrial manufacturing
5.	Bharat Petroleum Corporation Ltd. ^a	Energy ^b
6.	Bharti Airtel Ltd.	Telecom ^b
7.	Cairn India Ltd.	Energy ^b
8.	Cipla Ltd.	Pharmaceuticals
9.	Coal India Ltd. ^a	Metals
10.	DLF Ltd.	Construction
11.	Dr. Reddy's Laboratories Ltd.	Pharmaceuticals
12.	GAIL (India) Ltd. ^a	Energy ^b
13.	Grasim Industries Ltd.	Cement and cement products
14.	HCL Technologies Ltd.	Information Technology ^b
15.	Hero MotoCorp Ltd.	Automobile
16.	Hindalco Industries Ltd.	Metals
17.	Hindustan Unilever Ltd.	Consumer goods
18.	I T C Ltd.	Consumer goods
19.	Infosys Ltd.	Information Technology ^b
20.	Jaiprakash Associates Ltd.	Construction
21.	Jindal Steel and Power Ltd.	Metals
22.	Larsen & Toubro Ltd.	Construction
23.	Mahindra and Mahindra Ltd.	Automobile
24.	Maruti Suzuki India Ltd.	Automobile
25.	NTPC Ltd. ^a	Energy ^b
26.	Oil & Natural Gas Corporation Ltd. ^a	Energy ^b
27.	Power Grid Corporation of India Ltd. ^a	Energy ^b
28.	Ranbaxy Laboratories Ltd.	Pharmaceuticals
29.	Reliance Communications Ltd.	Telecom ^b
30.	Reliance Industries Ltd.	Energy ^b
31.	Reliance Infrastructure Ltd.	Energy ^b
32.	Reliance Power Ltd.	Energy ^b
33.	Sesa Goa Ltd.	Metals
34.	Siemens Ltd.	Industrial manufacturing
35.	Steel Authority of India Ltd. ^a	Metals
36.	Sterlite Industries (India) Ltd.	Metals
37.	Sun Pharmaceutical Industries Ltd.	Pharmaceuticals
38.	Tata Consultancy Services Ltd.	Information technology ^b
39.	Tata Motors Ltd.	Automobile

Table 12 (continued)

Sr. no.	Company name	Industry type (base on NSE)
40.	Tata Power Co. Ltd.	Energy ^b
41.	Tata Steel Ltd.	Metals
42.	Wipro Ltd.	Information technology ^b

^a PSUs. The rest belong to the private sector

^b Service industries. The rest are manufacturing firms

References

- Abbott, W. F., & Monsen, R. J. (1979). On the measurement of corporate social responsibility: self-reported disclosures as a method of measuring corporate social involvement. *The Academy of Management Journal*, 22(3), 501–515.
- Agarwal, S. K. (2008). *Corporate social responsibility in India*. Los Angeles: Response Books, Sage Publications.
- Alon, I., Lattemann, C., Fetscherin, M., Li, S., & Schneider, A. (2010). Usage of public corporate communications of social responsibility in Brazil, Russia, India and China (BRIC). *International Journal of Emerging Markets*, 5(1), 6–22.
- Aras, G., Aybars, A., & Kutlu, O. (2010). Managing corporate performance: investigating the relationship between corporate social responsibility and financial performance in emerging markets. *International Journal of Productivity and Performance Management*, 59(3), 229–254.
- Cowan, S., & Deegan, C. (2011). Corporate disclosure reactions to Australia's first national emission reporting scheme. *Accounting and Finance*, 51(2), 409–436.
- Das Gupta, A. (2007). Social responsibility in India towards global compact approach. *International Journal of Social Economics*, 34(9), 637–663.
- Dhanesh, G. S. (2014). CSR as organization-employee relationship management strategy: a case study of socially responsible information technology companies in India. *Management Communication Quarterly*, 28(1), 130–149.
- DPE (2010). *Guidelines on corporate social responsibility for central public sector enterprises*. DPE, GoI. <http://dpe.nic.in/newgl/glch1223.pdf>. Accessed 15 Jan 2012.
- DPE (2013). *Guidelines on corporate social responsibility and sustainability for central public sector enterprises*. DPE, GoI. http://dpe.nic.in/sites/upload_files/dpe/files/glch1231.pdf. Accessed 15 Jan 2015.
- Freundlieb, M., & Teuteberg, F. (2013). Corporate social responsibility reporting—a transnational analysis of online corporate social responsibility reports by market-listed companies: contents and their evolution. *International Journal of Innovation and Sustainable Development*, 7(1), 1–26.
- Gao, Y. (2011). CSR in an emerging country: a content analysis of CSR reports of listed companies. *Baltic Journal of Management*, 6(2), 263–291.
- Gautam, R., & Singh, A. (2010). Corporate social responsibility practices in India: a study of top 500 companies. *Global Business and Management Research: An International Journal*, 2(1), 41–56.
- Gray, R., Kouhy, R., & Lavers, S. (1995a). Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure. *Accounting, Auditing & Accountability Journal*, 8(2), 47–77.
- Gray, R., Kouhy, R., & Lavers, S. (1995b). Constructing a research database of social and environmental reporting by UK companies. *Accounting, Auditing & Accountability Journal*, 8(2), 78–101.
- Guthrie, J., & Abeysekera, I. (2006). Content analysis of social, environmental reporting: what is new? *Journal of Human Resource Costing & Accounting*, 10(2), 114–126.
- Guthrie, J., & Mathews, M. R. (1985). Corporate social accounting in Australasia. In L. E. Preston (Ed.), *Research in corporate social performance and policy* (pp. 251–277). New York: JAI.

- Guthrie, J., & Parker, L. D. (1989). Corporate social reporting: a rebuttal of legitimacy theory. *Accounting and Business Research*, 19(76), 343–352.
- Hackston, D., & Milne, M. J. (1996). Some determinants of social and environmental disclosures in New Zealand companies. *Auditing & Accountability Journal*, 9(1), 77–108.
- Haron, H., Yahya, S., Manasseh, S., & Ismail, I. (2006). Level of corporate social disclosure in Malaysia. *Malaysian Accounting Review*, 5(1), 159–184.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.
- Ianchovichina, E., Lundstrom, S., & Garrido, L. (2009). What is inclusive growth?. PRMED Knowledge Brief. World Bank. <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTDEBTDEPT/0,,contentMDK:21870580~pagePK:64166689~piPK:64166646~theSitePK:469043,00.html>. Accessed 25 June 2015.
- Jammulamadaka, N. (2013). The responsibility of corporate social responsibility in SMEs. *International Journal of Organizational Analysis*, 21(3), 385–395.
- Krippendorff, K. (1981). *Content analysis* (1st ed.). California: Sage.
- Mahadeo, J. D., Oogarah-Hanuman, V., & Soobaroyen, T. (2011). A longitudinal study of corporate social disclosures in a developing economy. *Journal of Business Ethics*, 104(4), 545–558.
- MCA (2009). *Corporate social responsibility voluntary guidelines 2009*. MCA, GoI. http://www.mca.gov.in/Ministry/latestnews/CSR_Voluntary_Guidelines_24dec2009.pdf. Accessed 12 Jan 2012.
- MCA (2011). *National voluntary guidelines on social, environmental and economic responsibilities of business*. MCA, GoI. http://www.mca.gov.in/Ministry/latestnews/National_Voluntary_Guidelines_2011_12jul2011.pdf. Accessed 12 Jan 2012.
- MCA (2013). *The companies act, 2013*. MCA, GoI. <http://www.mca.gov.in/Ministry/pdf/CompaniesAct2013.pdf>. Accessed 22 Jan 2012.
- Milne, M. J., & Adler, R. W. (1999). Exploring the reliability of social and environmental disclosures content analysis. *Accounting, Auditing & Accountability Journal*, 12(2), 237–256.
- Mohan, A. (2001). Corporate citizenship: perspectives from India. *Journal of Corporate Citizenship*, 2, 107–117.
- Montabon, F., Sroufe, R., & Narasimhan, R. (2007). An examination of corporate reporting, environmental management practices and firm performance. *Journal of Operations Management*, 25(5), 998–1014.
- Narwal, M., & Singh, R. (2013). Corporate social responsibility practices in India: a comparative study of MNCs and Indian companies. *Social Responsibility Journal*, 9(3), 465–478.
- Orij, R. (2010). Corporate social disclosures in the context of national cultures and stakeholder theory. *Accounting, Auditing & Accountability Journal*, 23(7), 868–889.
- Planken, B., Sahu, S., & Nickerson, C. (2010). Corporate social responsibility communication in the Indian context. *Journal of Indian Business Research*, 2(1), 10–22.
- Planning Commission (2008). *Inclusive growth vision and strategy*. Planning commission. http://planningcommission.nic.in/plans/planrel/fiveyr/11th/11_v1/11v1_ch1.pdf. Accessed 20 Jan 2012.
- Planning Commission (2011a). *Faster, sustainable and more inclusive growth: An approach to the Twelfth Five Year Plan (2012–17)*. Planning Commission. http://planningcommission.nic.in/plans/planrel/12appdrft/approach_12plan.pdf. Accessed 20 Jan 2012.
- Planning Commission (2011b). *Low carbon strategies for inclusive growth: An interim report*. Planning Commission. http://planningcommission.nic.in/reports/genrep/Inter_Exp.pdf. Accessed 20 Jan 2012.
- Porter, M. E., & Kramer, M. R. (2006). Strategy and society: the link between competitive advantage and corporate social responsibility. *Harvard Business Review*, 84(12), 78–92.
- Rahman, N. H. W. A., Zain, M. M., & Yahaya Al-Haj, N. H. Y. (2011). CSR disclosures and its determinants: evidence from Malaysian government link companies. *Social Responsibility Journal*, 7(2), 181–201.
- Raman, S. R. (2006). Corporate social reporting in India: a view from the top. *Global Business Review*, 7(2), 313–324.
- SEBI (2011). *SEBI Board meeting*. Press release (PR No. 145/2011 November 24). SEBI. <http://www.sebi.gov.in/sebiweb/home/list/4/23/0/0/Press-Releases>. Accessed 24 Nov 2011.
- Tilt, C. A. (2001). The content and disclosure of Australian corporate environmental policies. *Accounting, Auditing & Accountability Journal*, 14(2), 190–212.
- Ullmann, A. A. (1985). Data in search of a theory: a critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. *The Academy of Management Review*, 10(3), 540–557.
- Weber, R. P. (1990). *Basic content analysis* (1st ed.). Newbury Park: Sage Publications.



Excitation of Electromagnetic Surface Waves at A Conductor-Plasma Interface by An Electron Beam

Research Article

Ved Prakash^{1,*}, Ruby Gupta^{2,†}, Vijayshri¹ and Suresh C. Sharma³

¹ School of Sciences, Indira Gandhi National Open University, Maidan Garhi, New Delhi 110 068, India

² Department of Physics, Swami Shraddhanand College, University of Delhi, Alipur, Delhi 110 036, India

³ Department of Applied Physics, Delhi Technological University, Shahbad Daultpur, Bawana Road, Delhi 110 042, India

[†] Corresponding author: rubyssndu@gmail.com

Abstract. Electromagnetic surface waves are driven to instability on a conductor plasma interface via Cerenkov and fast cyclotron interaction by an electron beam. A dispersion relation and the growth rate of the instability for this process has been derived. Numerical calculations of the growth rate and unstable mode frequencies have been carried out for the typical parameters of the surface plasma waves. The plasma and beam responses are obtained using fluid treatment and the growth rate is obtained using the first-order perturbation theory. The growth rate increases with the beam density and scales as one-third power of the beam density in Cerenkov interaction and is proportional to the square root of beam density in fast cyclotron interaction. In addition, the real frequency of the unstable wave increases with the beam energy and scales as almost one half power of the beam energy. The effect of the plasma parameters and the strength of the external magnetic field on unstable frequencies and growth rates are analyzed.

Keywords. Surface plasma wave; Electron beam; Instability; Dispersion relation; Growth rate

PACS. 52

Received: March 1, 2015

Accepted: November 11, 2015

Copyright © 2016 Ved Prakash, Ruby Gupta, Vijayshri and Suresh C. Sharma. *This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

*Current address: India Meteorological Department, Ministry of Earth Science, Lodhi Road, New Delhi 110003, India

1. Introduction

Propagation of surface waves on the interface between two media derives attention due to its technological applications as well as its relation to astrophysical problems [1]. A surface plasma wave is an electromagnetic wave that propagates at the boundary between two media with different conductivities and dielectric properties such as a conductor-plasma boundary, only if the permittivity of one of the media is negative or has a nonzero negative part. Trivelpiece and Gould [2] first reported the experimental observations of surface plasma waves using a cylindrical plasma column enclosed in a glass tube that was coaxial with a circular metallic waveguide. These waves in plasmas have already been the subject of many theoretical, numerical and experimental investigations due to their spatial frequency spectrum [3, 4]. The problem of transferring energy from a beam of particles into electromagnetic wave energy has been given considerable attention in various fields of physics. Beam-energy extraction requires that phase matching between waves and particles is maintained for as long as possible. It is well known that Cerenkov and cyclotron resonances have this desirable property of maintaining synchronization. The SPW amplitude falls off rapidly as one move away from the interface. Therefore, SPW excitation by laser or electron beam injection has been observed and studied extensively by several investigators [5–11]. Denton et al. [5] have studied the process of SPW excitation over a metal surface by charged particles. Liu and Tripathi [6] have developed a theory of the excitation, reflection and scattering of a SPW over a metal surface by modeling a localized surface ripple as an electron density perturbation. Khankina et al. [7] studied the excitation of surface waves in magnetoactive plasma by a moving charged particle along the spiral line relative to the constant magnetic field. SPW can also be excited by an ion beam [12], by *attenuated total reflection* (ATR) configuration [13], by ripples of suitable wave number over the metallic interface [14], by light via prism coupling [15] or by laser [16–18]. The SPW excitation by charged particle on the metal vacuum interface can provide important information on the structure of the energy spectrum of the electron Fermi fluid in metals. Later, Liu and Tripathi [14] observed that a laser incident on a metal film excites a surface plasma wave at the metal-free space interface or it can also be excited by a relativistic electron beam. Macchi et al. [19] studied the parametric excitation of electron surface wave in the interaction of intense laser pulse with an over-dense plasma. Shokri and Jazi [20] showed that non reciprocal electromagnetic surface waves can be excited in semi bounded magnetized plasmas by an electron beam flowing on the plasma surface. Borisov and Nielsen [21] studied the excitation of plasma waves by unstable photoelectron and thermal electron populations on closed magnetic field in the Martian ionosphere. Kumar and Tripathi [22] studied the excitation of a surface plasma wave over a plasma cylinder by a relativistic electron beam propagating in a plasma cylinder and an annular beam propagating outside the plasma cylinder.

In the present paper, we study the excitation of SPWs by an electron beam propagating across an external magnetic field parallel to the conductor-plasma interface. In Section 2, we study the plasma and beam electron responses to the SPW perturbation. In Section 3, we have derived the dispersion relation and growth rate of SPWs. The variation of the growth rate of the unstable mode as a function of electron plasma density n_{e0} for different propagating distances

from the interface has been discussed in Cerenkov and fast cyclotron interactions. In Section 4, we discuss our results.

2. Plasma and Electron Beam Response

Consider a conductor-plasma interface at $x = 0$, with conductor in region $x < 0$ characterized by effective relative permittivity ϵ_c and dielectric constant ϵ_L , while plasma in region $x > 0$ with dielectric constant ϵ_p (cf. Figure 1).

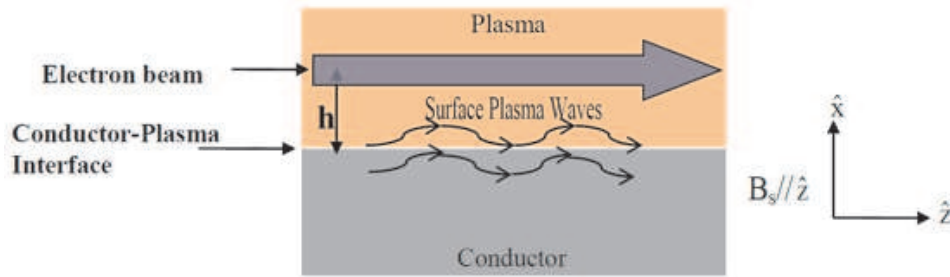


Figure 1. Schematic of a conductor-plasma interface with an electron beam propagating above the interface at a distance h .

The equilibrium electron and ion densities are n_{e0} and n_{i0} , respectively, immersed in a static magnetic field B in the z -direction. We assume the t, z variations of fields as $\mathbf{E}, \mathbf{B} \sim \exp[-i(\omega t - k_z z)]$ and consider E field to be polarized in the x - z plane. An electron beam is considered propagating along z -axis at a height ' h ' above the conductor-plasma interface, with density n_{b0} and equilibrium beam velocity $v_{b0}\hat{z}$. To obtain the response of plasma electrons to the fields of the surface plasma waves, we solve the equation of motion

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{e}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right), \quad (1)$$

where $\mathbf{v} = v_{b0}\hat{z} + \mathbf{v}_{b1}$, \mathbf{v}_{b1} refers to perturbed velocity.

The magnetic field of the wave is

$$B_w = \left(\frac{c}{\omega} \right) (\mathbf{k} \times \mathbf{E}) = \left(\frac{c}{\omega} \right) (k_z E_x - k_x E_z) \hat{y}.$$

Here $E_x = \left(\frac{ik_z}{k_2} \right) E_z$ for $\nabla \cdot \mathbf{E} = 0$ SPWs.

The decay rate of wave amplitude in plasma is given by $k_1 = \frac{\frac{\omega \epsilon_p}{c}}{[-(\epsilon_p + \epsilon_c)]^{\frac{1}{2}}}$, and the decay rate of wave amplitude in conductor is $k_2 = \frac{\frac{\omega}{c}}{\left[-\frac{\epsilon_p + \epsilon_c}{\epsilon_c^2} \right]^{\frac{1}{2}}}$, where $\epsilon_c = \epsilon_L - \frac{\omega_{pe}^2}{\omega^2}$, $\omega_{pe}^2 = \frac{4\pi n_{e0} e^2}{m}$.

For k_1 and k_2 to be real, we must have $\epsilon_p + \epsilon_c < 0$, which gives

$$\omega < \omega_{sc}, \quad (2)$$

where $\omega_{sc} = \frac{\omega_p}{(\epsilon_p + \epsilon_L)^{\frac{1}{2}}}$ is the surface wave cut-off frequency.

Eq. (2) is essential for the existence of surface plasma wave.

The perturbed electron velocity in x-, y- and z-directions, after linearization are obtained as

$$v_{b1x} = \frac{\left[e \left(\frac{ik_z}{k_2} \right) - e \left(\frac{v_{b0}}{\omega} \right) \left\{ \left(\frac{ik_z^2}{k_2} \right) - k_x \right\} \right] (\omega - k_z v_{b0}) E_z}{im [(\omega - k_z v_{b0})^2 - \omega_{ce}^2]}, \quad (3)$$

$$v_{b1y} = \frac{\left[e \left(\frac{ik_z}{k_2} \right) - e \left(\frac{v_{b0}}{\omega} \right) \left\{ \left(\frac{ik_z^2}{k_2} \right) - k_x \right\} \right] \omega_{ce} E_z}{m [(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} \quad (4)$$

and

$$v_{b1z} = \frac{e E_z}{im (\omega - k_z v_{b0})} \quad (5)$$

where $\omega_{ce} = eB/mc$ is the electron cyclotron frequency.

From equation of continuity $\frac{\partial n}{\partial t} + \nabla \cdot (nv) = 0$, where $n = n_{b0} + n_{b1}$, we obtain the perturbed beam density n_{b1} and perturbed electron density n_{e1} as

$$n_{b1} = \frac{en_{b0} \left[ik_z - \left(\frac{v_{b0}}{\omega} \right) (ik_z^2 - k_2 k_x) \right] E_z}{m [(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} + \frac{en_{b0} k_z E_z}{im (\omega - k_z v_{b0})^2} \quad (6)$$

and

$$n_{e1} = -\frac{n_{e0} e k_z E_z}{im} \left[\frac{\omega_{ce}^2}{\omega^2 (\omega^2 - \omega_{ce}^2)} \right]. \quad (7)$$

3. Dispersion Relation and Growth Rate

The perturbed current density will be

$$\mathbf{J}_1 = -e(n_{b0} \mathbf{v}_{b1} + n_{b1} v_{b0} \hat{z}) \delta(x-h).$$

By retaining only those terms which go as $(\omega - k_z v_{b0})^{-2}$, the z-component of current density is obtained as

$$J_{1z} = \left[-\frac{e^2 v_{b0} \left[ik_z - \left(\frac{v_{b0}}{\omega} \right) (ik_z^2 - k_2 k_x) \right] E_z n_{b0}}{m [(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} - \frac{e^2 v_{b0} k_z E_z n_{b0}}{im (\omega - k_z v_{b0})^2} \right] \delta(x-h)$$

or

$$J_{1z} = -\frac{e^2 v_{b0} n_{b0} E_z}{m} \left[\frac{\left[ik_z - \left(\frac{v_{b0}}{\omega} \right) (ik_z^2 - k_2 k_x) \right]}{[(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} + \frac{k_z}{i(\omega - k_z v_{b0})^2} \right] \delta(x-h). \quad (8)$$

Using Eq. (8) in the z-component of wave equation $\nabla^2 E - \nabla(\nabla \cdot E) + \frac{\omega^2}{c^2} \epsilon_p E = -\frac{4\pi i \omega}{c^2} J$ we get

$$\left(k_z^2 - k_z^2 + \frac{\omega^2}{c^2} \epsilon_p\right) E_z = -\frac{\omega v_{b0}}{c^2} \omega_{pb}^2 \left[\frac{\left[k_z - \left(\frac{v_{b0}}{\omega}\right)(k_z^2 + i k_2 k_x)\right]}{[(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} - \frac{k_z}{(\omega - k_z v_{b0})^2} \right] E_z \delta(x-h) \quad (9)$$

where $\omega_{pb}^2 = \frac{4\pi n_{b0} e^2}{m}$.

Multiplying by E_z^* and integrating from $x = 0$ to ∞ , taking $\int_0^\infty E_z^* E_z dx = 1$ and $\int_0^\infty E_z^* E_z \delta(x-h) dx = \exp(-2k_2 h)$, Eq. (9) can be rewritten as

$$\omega^2 - k_z^2 c^2 \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) = -\omega v_{b0} \omega_{pb}^2 \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) \left[\frac{\left[k_z - \left(\frac{v_{b0}}{\omega}\right)(k_z^2 + i k_2 k_x)\right]}{[(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} - \frac{k_z}{(\omega - k_z v_{b0})^2} \right] e^{-2k_2 h} \quad (10)$$

or

$$(\omega - \omega_{spw})(\omega + \omega_{spw}) = -\omega v_{b0} \omega_{pb}^2 \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) \left[\frac{\left[k_z - \left(\frac{v_{b0}}{\omega}\right)(k_z^2 + i k_2 k_x)\right]}{[(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} - \frac{k_z}{(\omega - k_z v_{b0})^2} \right] e^{-2k_2 h} \quad (11)$$

where ω_{spw} is the root of ω , given as

$$\omega_{spw}^2 = k_z^2 c^2 \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right). \quad (12)$$

Eq. (12) represents the standard dispersion relation of SPW [23].

In Cerenkov interaction, $(\omega - k_z v_{b0})^2 \approx 0 \Rightarrow \omega \sim k_z v_{b0}$, therefore neglect the first term on RHS in Eq. (11) and assume perturbed quantities $\omega = \omega_{spw} + \delta$ and $\omega = k_z v_{b0} + \delta$, where δ is the small frequency mismatch.

The growth rate is obtained as

$$\gamma = \text{Im}(\delta) = \frac{\sqrt{3}}{2} \left[\frac{\omega_{pb}^2 k_z v_{b0}}{2} \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) e^{-2k_2 h} \right]^{\frac{1}{3}}. \quad (13)$$

The real part of frequency ω_r is obtained from the real part of δ as

$$\omega_r = k_z \left(\frac{2eV_b}{m}\right)^{\frac{1}{2}} - \frac{1}{2} \left[\frac{\omega_{pb}^2 k_z v_{b0}}{2} \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) e^{-2k_2 h} \right]^{\frac{1}{3}} \quad (14)$$

where V_b is the beam potential.

In cyclotron interaction $(\omega - k_z v_{b0})^2 \gamma_o^2 \approx \omega_{ce}^2$, therefore neglecting the second term on RHS in Eq. (11), we get

$$(\omega - \omega_{spw})(\omega + \omega_{spw}) = -\omega v_{b0} \omega_{pb}^2 \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c}\right) \left[\frac{\left[k_z - \left(\frac{v_{b0}}{\omega}\right)(k_z^2 + i k_2 k_x)\right]}{[(\omega - k_z v_{b0})^2 - \omega_{ce}^2]} \right] e^{-2k_2 h} \quad (15)$$

where $(\omega - k_z v_{b0})\gamma_o + \omega_{ce}$ corresponds to slow cyclotron interaction, and $(\omega - k_z v_{b0})\gamma_o - \omega_{ce}$ corresponds to fast cyclotron interaction.

Considering slow cyclotron interaction, and assuming perturbed quantities $\omega = \omega_{spw} + \delta$ and $\omega = k_z v_{b0} - \omega_{ce} + \delta$, the growth rate is obtained as

$$\gamma = 0. \quad (16)$$

The phase velocity of the unstable mode is obtained from the real part of ω as

$$\gamma = \left[\frac{\omega_{pb}^2}{4} \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c} \right) \left(1 - \frac{\omega_{ce}}{\omega} \right) e^{-2k_z h} \right]^{\frac{1}{2}}. \quad (17)$$

That is, in case of slow cyclotron interaction there is no growing mode as the phase velocity exceeds the beam velocity.

Now, considering fast cyclotron interaction and assuming perturbed quantities $\omega = \omega_{spw} + \delta$ and $\omega = k_z v_{b0} + \omega_{ce} + \delta$, the growth rate is obtained as

$$\gamma = \left[\frac{\omega_{pb}^2}{4} \left(\frac{\epsilon_p + \epsilon_c}{\epsilon_p \epsilon_c} \right) \left(1 - \frac{\omega_{ce}}{\omega} \right) e^{-2k_z h} \right]^{\frac{1}{2}}. \quad (18)$$

4. Results and Discussion

Typical parameters of the SPWs used for numerical calculations are: electron plasma density $n_{e0} = 10^{10} \text{ cm}^{-3}$ and 10^{11} cm^{-3} , mass of electron $m_e = 9.1 \times 10^{-28} \text{ g}$, charge of electron $e = 4.8 \times 10^{-10} \text{ ergs}$, magnetic field $B = 30 \text{ G}$, dielectric constant of plasma $\epsilon_p = 1.2$, dielectric constant of conductor $\epsilon_L = 4$ and beam velocity $v_{bo} = 2 \times 10^{10} \text{ cm/s}$. The electron beam is assumed to travel at a distance of 2 cm, 4 cm and 6 cm from the conductor-plasma interface with beam density $n_{bo} = 2 \times 10^9 \text{ cm}^{-3}$.

We have plotted the dispersion curves of surface plasma waves on a conductor-plasma interface for the two values of electron plasma densities using Eq. (12) in Figure 2.

We have also plotted the beam modes via Cerenkov interaction and fast cyclotron interaction with SPWs, where the beam is assumed to travel at a distance 2 cm from the interface. The frequencies and the corresponding wave numbers of the unstable mode are obtained by the points of intersection between the beam modes and the plasma modes. We can say that the unstable wave frequencies and the axial wave vector k_z (cm^{-1}) of the SPWs increase with an increase in plasma density, in Cerenkov as well as cyclotron interactions. The variation in the value of applied magnetic field does not affect the Cerenkov interaction, but the fast cyclotron beam mode with $B > 45 \text{ G}$ could not interact with SPW for plasma density of 10^{10} cm^{-3} and the beam mode with $B > 135 \text{ G}$ do not intersect with SPW dispersion curve. It implies that the value of magnetic field should be less for a less dense plasma for the existence of cyclotron interaction between electron beam and SPW mode.

Using Eq. (13), we have plotted in Figure 3 the growth rate γ (rad./sec) of the surface plasma waves as a function of unstable frequency ω in Cerenkov interaction with a beam travelling at different distances from the interface, and for two values of plasma electron densities.

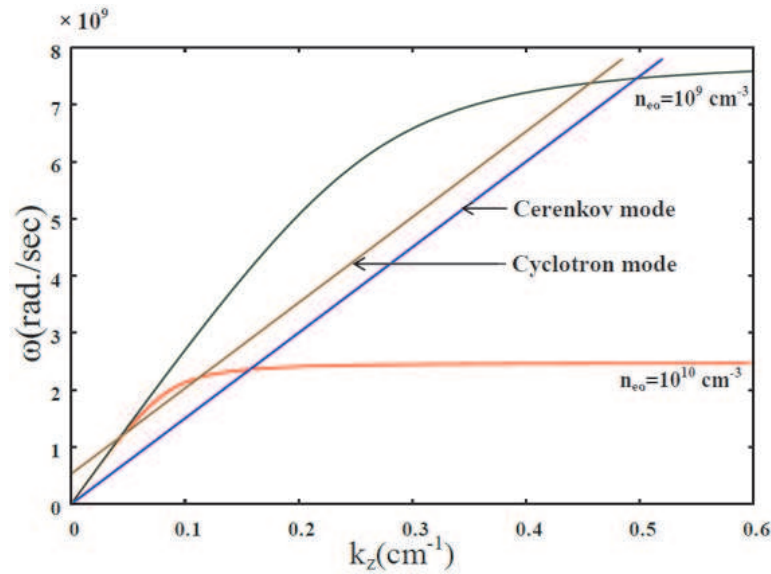


Figure 2. Dispersion curves of SPWs over a magnetized plasma for different values of n_{e0} and beam modes. The parameters are given in the text.

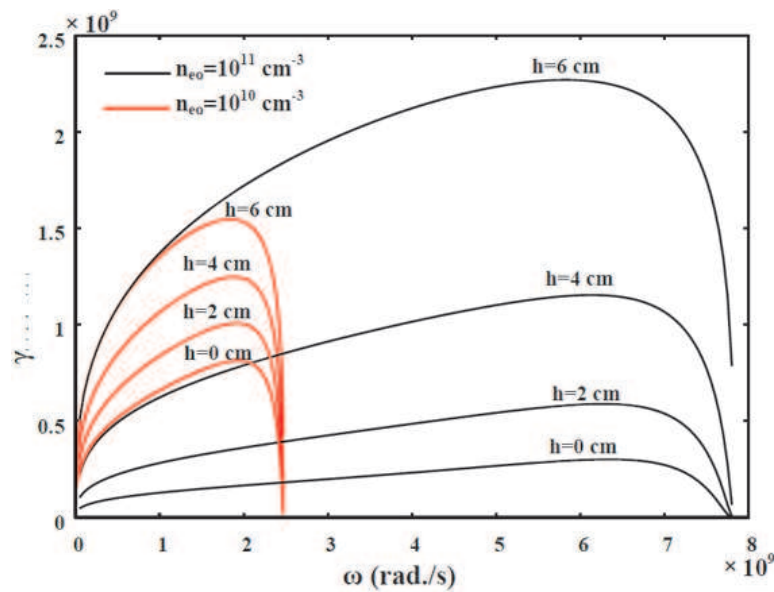


Figure 3. Growth rate γ of the unstable mode as a function of wave frequency for a beam propagating at different heights from the interface via Cerenkov interaction, with two different values of n_{e0} .

From Figure 3, it can be seen that the growth rate of the unstable mode first increases and then shows a maxima in all the curves. When $h = 0$ cm, i.e. when the beam is propagating along the interface, the growth rate is more for a denser plasma. However, as the value of h increases i.e. as the beam moves away from the interface, the growth rate decreases rapidly in case of a denser plasma as compared to the rate of decrease of growth rate in less dense plasma. The growth rate γ (in rad./sec) of the unstable wave decreases from 1.55 to 0.54 for plasma density of 10^{10} cm^{-3} , and from 2.27 to 0.09 for plasma density of 10^{11} cm^{-3} , when h increases from

0 cm to 6 cm. The growth rate of the unstable mode also increases with the beam density and scales as the one-third power of the beam density [cf. Eq. (13)].

In Figure 4, we have plotted the variation of the growth rate γ (in rad./sec) of SPW as a function of unstable frequency via fast cyclotron interaction and the beam mode for same beam velocity and static magnetic field, for the two values of plasma densities and for electron beam travelling at different distances from the interface.

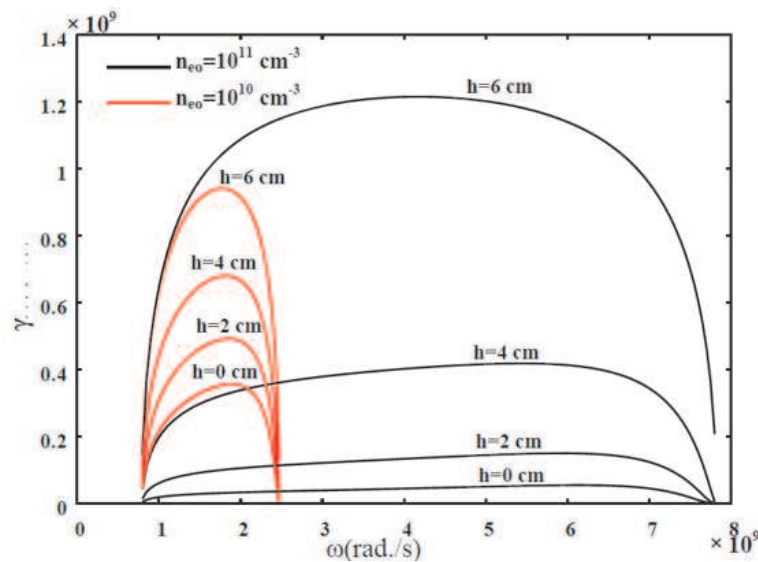


Figure 4. Growth rate γ of the unstable mode as a function of wave frequency for a beam propagating at different heights from the interface via Cerenkov interaction, with two different values of n_{e0} .

The behavior of the curves is almost same as in the case of Cerenkov interaction. The values of growth rate are comparatively smaller in cyclotron interaction, with a maximum of 0.95 for $n_{e0} = 10^{10} \text{ cm}^{-3}$ and a maximum of 1.21 for $n_{e0} = 10^{11} \text{ cm}^{-3}$. The growth rate also increases with the beam density and scales as the square root of the beam density. In case of slow cyclotron interaction there is no growing mode as the phase velocity exceeds the beam velocity [cf. Eq. (17)].

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

- [1] B. Buti, *Advances in Space Plasma Physics*, World Scientific, Singapore (1985).
- [2] A.W. Trivelpice and R.W. Gould, *J. Appl. Phys.* **30** (1959), 1784.
- [3] D.J. Cooperberg, *Phys. Plasmas* **5** (1998), 853.

- [4] K.L. Kartwright, P.J. Christenson, J.P. Verboncoeur and C.K. Birdsall, *Phys. Plasmas* **7** (2000), 1740.
- [5] C. Denton, J.L. Gervasoni, R.O. Barrachina and N.R. Arista, *Phys. Rev. A* **57** (1998), 4498.
- [6] C.S. Liu and V.K. Tripathi, *IEEE J. of Quant. Electron.* **34** (1998), 1503.
- [7] S.I. Khankina, V.M. Yakovenko and I.V. Yakovenko, *Telecommun. Radio Eng. (Engl. Transl.)* **59** (2003), 140.
- [8] V. Prakash and S.C. Sharma, *Phys. Plasmas* **16** (2009), 0937031.
- [9] R. Gupta, S.C. Sharma and V. Prakash, *Phys. Plasmas* **17** (2010), 122105.
- [10] R. Gupta, S.C. Sharma and V. Prakash, *Phys. Plasmas* **17** (2010), 112701.
- [11] R. Gupta, S.C. Sharma and V. Prakash, *Phys. Plasmas* **18** (2011), 053704.
- [12] I.O. Girka, V.O. Girka and I.V. Pavlenko, *Progress in Electromagnetics Research M* **21** (2011), 267,.
- [13] E. Kretschmann, *Z. Phys.* **241** (1971), 313.
- [14] C.S. Liu and V.K. Tripathi, *IEEE Transactions of Plasma Science* **28** (2) (2000), 353.
- [15] N.C. Chen, W.C. Lien, C.R. Liu, Y.L. Huang, Y.R. Liu, C. Chou, S.Y. Chang and C.W. Ho, *J. Appl. Phy.* **109** (2011), 0431041.
- [16] V.N. Anisimov, V.Y. Baranov, O.N. Derkach, A.M. Dykhne, D.D. Malyuta, V.D. Pismennyi, B.P. Rysev and A.Y. Sebrant, *IEEE J. Quantum Electron* **24** (1988), 675.
- [17] J. Parashar, H.D. Pandey and V.K. Tripathi, *J. Plasma Phys.* **59** (1998), 97.
- [18] A. Bouhelier, F. Ignatovich, A. Bruyant, C. Huang, Franks, D.G. Colas, J.C. Weeber, A. Dereux, G.P. Wiederrecht and L. Novotny, *Optics Letters* **32** (2007), 2535.
- [19] A. Macchi, F. Cornolti and F. Pegoraro, *Phys. Plasmas* **9** (2002), 1704.
- [20] B. Shokri and B. Jazi, *Phys. Plasmas* **10**(12) (2003), 4622.
- [21] N. Borisov and E. Neilsen, *Annales Geophysicae* **23** (2005), 1249.
- [22] G. Kumar and V.K. Tripathi, *Phys. Plasmas* **15** (2008), 073504.
- [23] C. S. Liu and V. K. Tripathi, *Electromagnetic Theory for Telecommunications*, Cambridge University Press, Delhi (2007).

ABSTRACT

Microchannels which are light in weight, smaller in size, handle low amount of coolants and provide higher surface area to volume ratio give higher heat transfer coefficients are efficient cooling devices for various modern electronic systems where higher heat fluxes are encountered. In this connection an experimental investigation has been carried out on the laminar flow and convective heat transfer of deionized water in circular microchannels. To explore the validity of classical correlations of friction factor based on conventional sized channels for predicting the fluid behavior of single phase liquid flow of water through circular microchannels experiments were conducted with deionized water with Reynolds number ranging from 71 to 1495. The copper microchannels used in the experiment have the hydraulic diameter of 279 μm and 45 mm length. Pressure drop and flow rates were measured to analyze the flow characteristics. The results show good agreement between the classical correlations of friction factor and the experimentally measured data.

Keywords: Microchannels, Reynolds numbers; Friction factor

INTRODUCTION

Miniaturization of Microfluidic devices has led to development of high heat fluxes during operation of these devices. We are in the era of developing Micro and Nano devices leading tremendous increase in power density which arises significant thermal management problems. One of the solutions proposed is to enhance heat transfer by using microchannels. In a microchannel a fluid is used to carry away heat from the small hot surface by forcing it through passages having hydraulic diameters ranging from 10 μm to 200 μm [1]. As a microchannel has higher heat transfer surface area to fluid volume ratio, so it provides high heat transfer coefficient for convective heat transfer. However this small channel experiences a considerable pressure drop.

The pioneer work in the field of heat transfer using microchannel heat sink for electronic cooling was first time demonstrated by Tuckerman and Pease [2] by achieving high heat flux removal capacity of up to 800 W/cm² with microchannels in single-phase and two-phase flows. They noted that as the hydraulic diameter of the channel decreases, the heat transfer coefficient increases. This landmark work paved the door for further research in the area of microchannel heat transfer. However it has been reported by various researchers that phenomena in microgeometry may differ from those in macroscopic counterparts and great differences still exist between the available measured data of Nusselt number and friction factor. One of the important factor for variations in reported data may be differences in channel roughness [3]. The review conducted by Steinke and Kandlikar [4] on the results of single-phase liquid friction factors in microchannels concluded that conventional Stokes and Poiseuille flow theories could apply for single-phase liquid flow in microchannel flows. Mala and Li [5] conducted experiments on flow of water through microtubes with inner diameters from (50 μm to 254 μm) made of fused silica and stainless steel. For higher values of Reynolds number, they found a noticeable increase in friction factor, compared to the

predictions based on conventional Poiseuille flow theory. They also observed the material dependence of flow behavior, for similar flow and hydraulic diameter; fused silica microtubes exhibited higher friction factor than stainless- steel microtubes. But for larger microtubes, as well as for lower Reynolds number flows, the experimental data were in agreement with the conventional theory. The most salient feature of their work was the attribution of these deviations of microscale flow characteristics to the effects of surface roughness of the microtubes. Choi et al. [6] conducted experiments in circular silica microtubes with inside diameters 3, 7, 10, 53 and 81.2 μm having length 24 to 52 mm with nitrogen as working fluid for evaluating heat transfer and pressure drop characteristics. They observed that for both laminar and turbulent flows Nusselt number depends on Reynolds number in a different manner as compared to macroscale theory. Based on the measurements they found the critical Reynolds number 2000, which is concurrent with microtubes. Qu et al. [7] conducted experiments to investigate the flow behavior of water through trapezoidal silicon microchannels ($51 \mu\text{m} \leq D \leq 169 \mu\text{m}$) and apply the roughness viscosity model to explain the experimentally obtained higher values of friction factor. Pfund et al. [8] experimentally investigated the water flow characteristics through high-aspect-ratio smooth and rough rectangular microchannels ($128 \mu\text{m} \leq D \leq 521 \mu\text{m}$) for Reynolds number ranging 60 to 3450, and concluded that the laminar flow friction was considerably higher than the classical value for the rougher microchannels. Wu and Cheng [9] in their experimental study found that for channels having similar geometry, the friction factor obtained was higher for the channel having the greater surface roughness and it was in accordance with the majority of the works highlighting the effects of surface roughness on microflow characteristics.

Li et al. [10-11] in their experimental study for laminar flow of deionized water through glass, silicon and stainless steel microtubes ($79.9 \mu\text{m} \leq D \leq 166.3 \mu\text{m}$) and obtained the Poiseuille number relatively higher for rougher stainless steel microtubes than for smoother glass and silica microtubes. Further they found that for rough stainless steel microtubes ($373 \mu\text{m} \leq D \leq 1570 \mu\text{m}$) the friction factor was not only higher than the predicted by conventional theory but increased further as relative roughness increased. However experimental values of friction factor found in agreement with the values predicted by conventional theory for water flows through relatively smooth fused silica microtubes ($50 \mu\text{m} \leq D \leq 100 \mu\text{m}$). They also concluded that classical theory successfully addressed the frictional characteristics for flows in microtubes with relative surface roughness less than about 1.5 %. Most of the experimental studies conducted have revealed deviations in the values of friction factor from classical conventional theory the reasons asserted are by various researchers are experimental uncertainty, surface roughness.

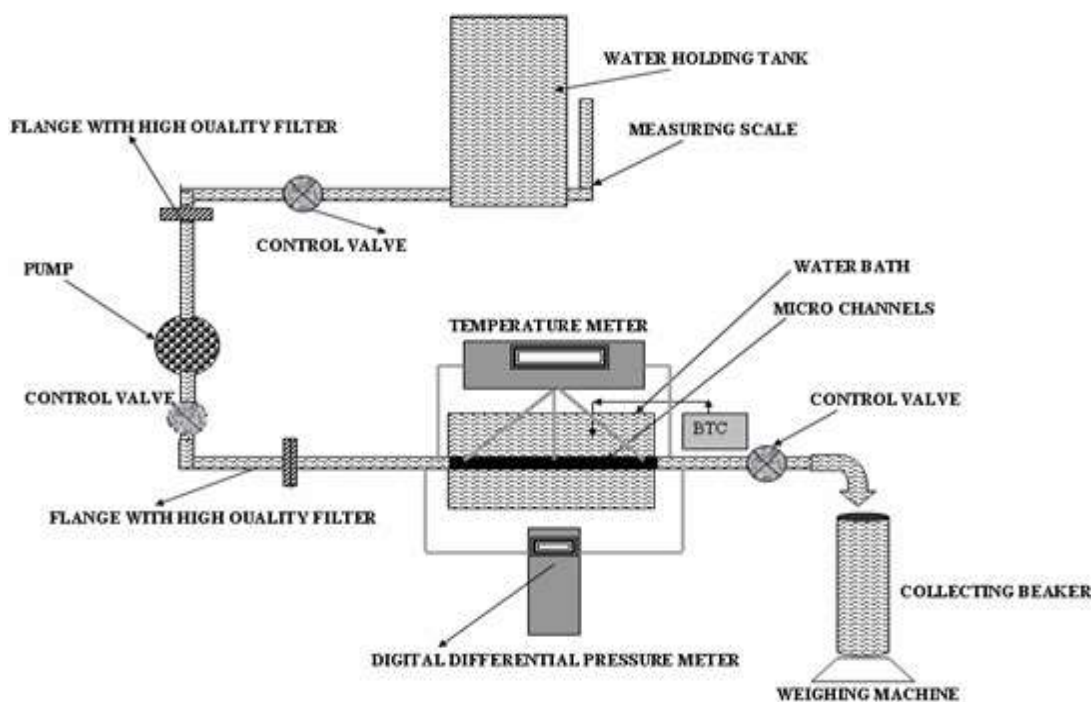
From the review of literature it has been observed that there is a large scatter in the data measured for friction factor and reasons suggested by various studies are contradictory. Meanwhile recent papers have shown that conventional theory is applicable by careful evaluation of geometry or by considering other effects. In this present work an attempt is made to investigate friction factor experimentally and then comparing with theory.

EXPERIMENTAL SETUP

The Experimental setup used in this investigation for measurements of pressure and temperature difference at inlet and outlet of the test section is shown in Figure 1. A flow pump is used to drive the Deionized water from a water holding tank; the flow pump provides smooth and steady flow over a wide range of flow rates that corresponds to a Reynolds number ranging from 50 to 1500. In order to avoid blockage of the microchannels, a submicron filter of $0.1 \mu\text{m}$ was installed between outlet of pump and inlet of test section. The test section was enveloped by the heated oil in the oil bath. Details of the microchannel test section are shown in Figure 2. The test section consists of total seventy nine stainless steel tubes having the inner diameter 279 and 45 mm long arranged in the circumferential manner. Three copper-constantan (T-Type) thermocouples were used to measure the temperatures at the inlet and outlet of the test section as well as of the oil of the oil bath and hand operated digital RS232 manometer (has a pressure range of 0 to 100 Psi with an accuracy of $\pm 0.3\%$ of its full scale at 25°C) was used to measure the differential pressure between the inlet and outlet of the test section as shown in Figure 2. HJ-123 heater of 500 W was placed in the oil bath to heat the oil and the connection of this heater through the blind temperature controller (BTC) which cut the electric supply of the heater when the temperature of the oil reaches the desired temperature. The first step in conducting the experiment was to fill the water tank with Deionized water and note down the initial as well as final reading of measuring scale of water tank before and after filling. This gives the

volume of water contained in the water tank. Once the water was filled, heater is turned on the heater and wait until the temperature of the oil reaches the desired temperature. Once the oil in the oil bath beaker attains this temperature, open the valve of water tank and motor is switched on. This allowed water to flow through the Microchannel Heat Exchanger. A set flow rate was established with the help of controlled valves by monitoring the digital manometer. After a steady state was reached, temperatures of water at inlet and outlet of the test section were recorded from the monitor of temperature indicator keeping in mind that the temperature of the oil in the oil bath remained constant with time during measurement. Once the temperature measurements were completed, water is collected in the volumetric beaker from the exit section for predetermined period of time and measured the flow rate. This procedure was repeated for several times for various differential pressure readings. The experiments cover the Reynolds numbers from 71 to 1495.

Figure: 1



Schematic diagram of Experimental setup

In conventional channels, internal laminar flow and heat transfer the surface irregularities are considered negligible for laminar flow, and the Nusselt number for thermally fully developed flow is dependent on the cross sectional shape of the channel.

Data Reduction

In this Experimental study, Reynolds number is defined by

$$Re = \frac{\rho V D}{\mu} \quad (1)$$

Where ρ is density of the fluid flowing through microchannel and V is the mean velocity of the flowing fluid, μ is the viscosity of fluid and D is the hydraulic diameter defined by

$$D = \frac{4A}{P} \quad (2)$$

Where A is the area of cross section and P is the wetted perimeter of the microchannel.

The Experimental friction factor based on the pressure drop measurements in microchannel flow are represented by Darcy friction factor which is given as

$$f_e = \frac{2\Delta p D}{\rho V^2 L} \quad (3)$$

Where Δp is the pressure drop across the length of the microchannel, L is the length of the microchannel, D is the Hydraulic diameter of the channel, ρ is the density of fluid flowing through microchannel and V is the mean velocity of the flowing fluid. The experimental friction factor is compared with theoretical friction factor for fully developed laminar flow in circular pipes based on classical macroscale theory given as

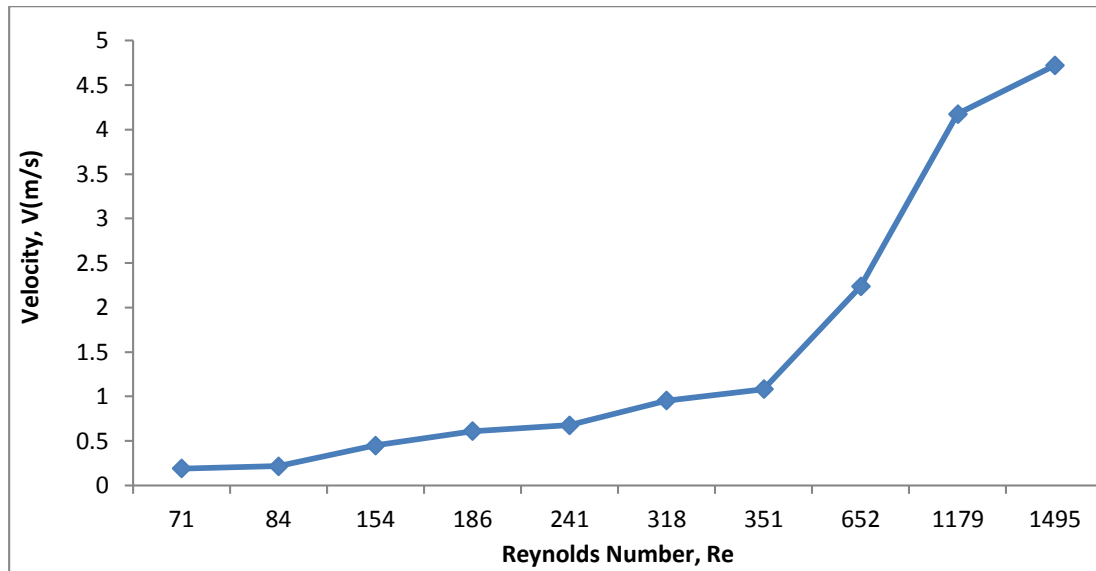
$$f_{th} = 64/Re \quad (4)$$

In microchannel flow analysis, hydrodynamically developed flow is very important because of short length of microchannel [4]. Kays and Crawford [12] suggested a good approximation for development flow length as given by relation ($x^+ = x/ReD$), and fully developed flow is assumed valid above the value of $x^+ = 0.05$ [12].

RESULTS AND DISCUSSIONS

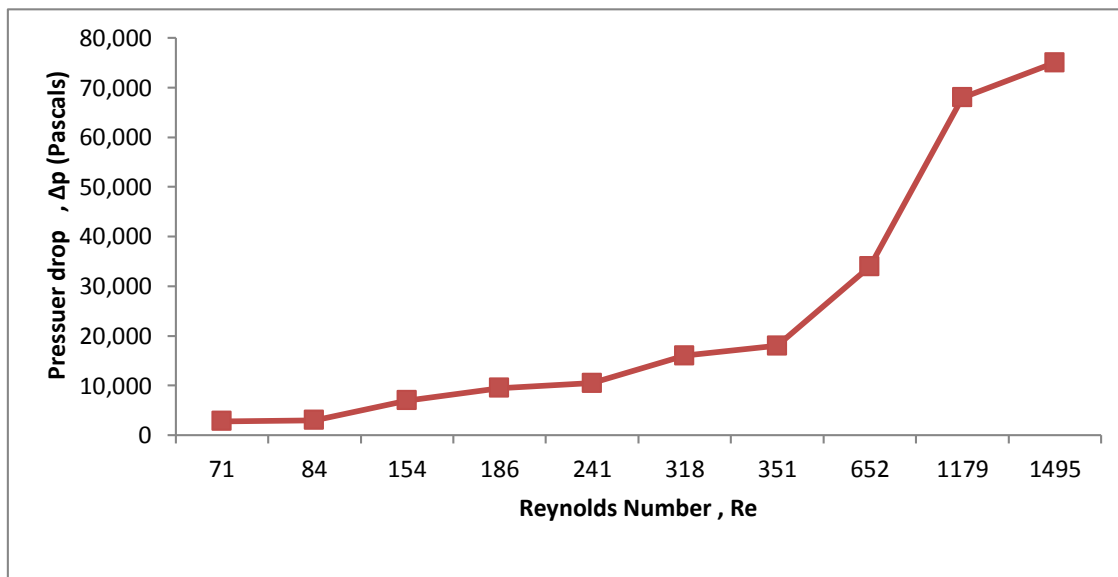
In this experimental study the inlet and outlet plenums were incorporated to induce a uniform flow distribution in microchannels and total pressure drop between the inlet and outlet plenums is measured. The flow rate was varied from 54 mL/min to 1367 mL/min and steady-state laminar flow was assumed, the flow velocity in microchannels varied from 0.19 m/s for lowest Reynolds number 71 to 4.72 m/s for highest Reynolds number 1495 as shown in Fig.2. It can be seen from Fig.3 that the pressure drop is linearly proportional to Reynolds number for a given microchannel dimension and the fluid flow is laminar without any transient effects as predicted by classical theory of fluid dynamics. The values of friction factors obtained from experimental pressure drop data are shown in Fig. 4. Each curve shows that experimental values are very close and consistent with the values predicted by the theory of fully developed flow as shown in Fig.5. Steinke and Kandlikar [4] compared their experimental results with previous researchers data and suggested that the entrance effects caused by developing flows was very important and should be taken into consideration in microchannel flow. Our experimental results show the flow in microchannels to be fully-developed as all of the x^+ values (0.11-2.26) obtained are bigger than 0.05 [12] and our experimental results are in good agreement with classical theory. Further the values of f_{exp}/f_{th} obtained by assuming fully-developed flow were between 0.79 and 1.23 except one value of 1.4 having (average 1.03) which shows that the flow in microchannels can be well predicted by applying conventional friction factor theory and the flow inside the microchannels is fully-developed.

Figure: 2



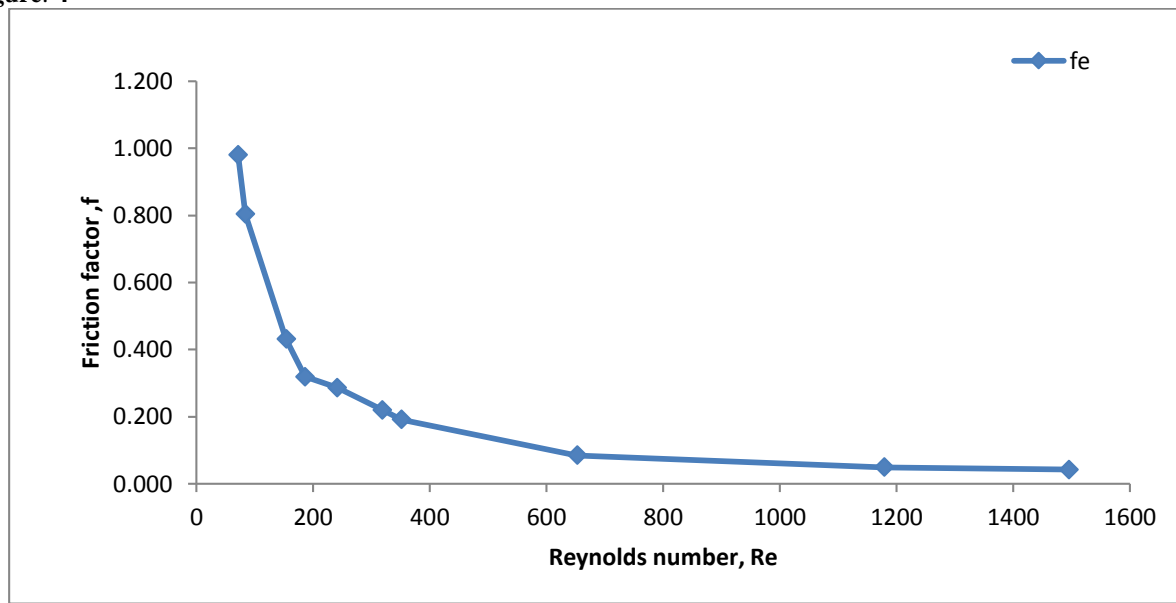
Fluid velocity variation with Reynolds number

Figure: 3



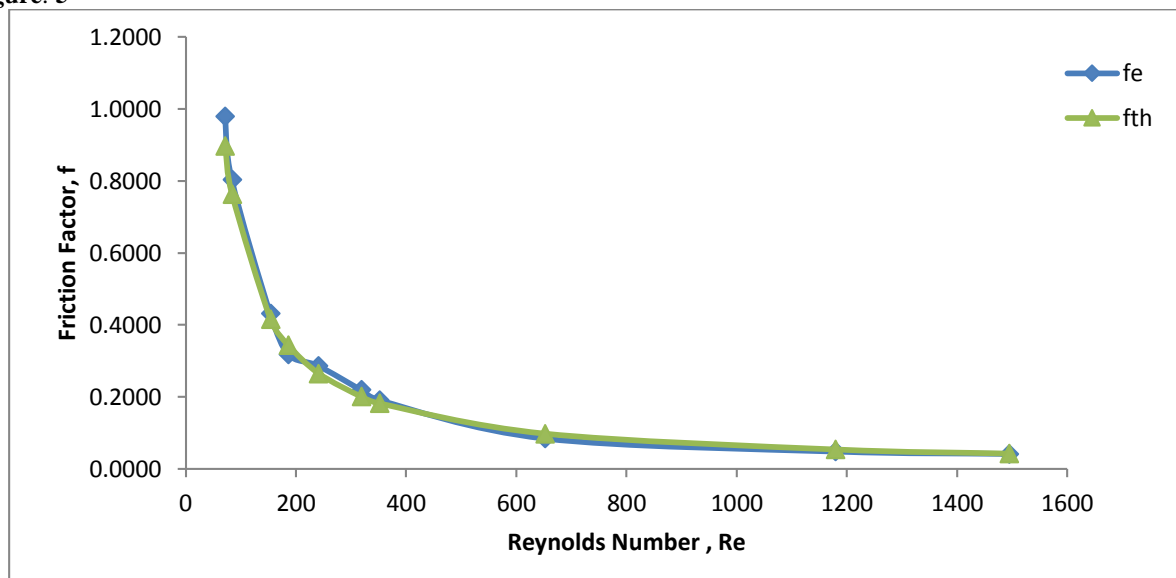
Total pressure drop variation with Reynolds number

Figure: 4



Friction factor variation with Reynolds number

Figure: 5



Friction factor (Experimental & Theoretical) variation with Reynolds number

CONCLUSIONS

Experimental investigations have been carried out for friction factors in single phase liquid flow inside circular microchannels. The experimental results obtained showed good agreement with conventional hydraulic theory and also suggested that the flow inside the microchannels was fully developed laminar flow in the range of experiments

conducted ($71 < Re < 1493$) inside the circular microchannels having hydraulic diameter of $279 \mu\text{m}$ and 45 mm length. The flow in microchannels can be well predicted by applying conventional friction factor theory and the flow inside the microchannels is fully-developed.




NOMENCLATURE

Re	Reynolds Number	Δp	Experimental pressure drop
Nu	Nusselt Number	L	Length of microchannel
D	Microchannel hydraulic diameter	x^+	Hydraulic entrance length
MEMS	Micro-Electro Mechanical Systems	ρ	Density of fluid
f	Friction factor	μ	Viscosity of fluid
V	Mean flow velocity	f_{th}	Theoretical Friction factor
A	Microchannel cross section area	f_e	Experimental friction factor

REFERENCES

- [1] Kandlikar, S. G.; History, Advances, and Challenges in Liquid Flow and Flow Boiling Heat Transfer in Microchannels: A Critical Review, *ASME J. Heat Transfer*, Vol.134/034001-15. March,2012
- [2] Tuckerman DB, Pease RFW ; High-performance heat sinking for VLSI. *IEEE Electron Device Letters* ,5:126–9,,1981
- [3] Zhang Chengbin, Chen youngping, Shi Mingheng ,Effect of roughness elements on laminar flow and heat transfer in microchannels, *Chemical Engineering and Processing* 49,1188-1192,2010
- [4] Steinke, M.E., and Kandlikar, S.G., Single-Phase Liquid Friction Factors in Microchannels, *International Journal of Thermal Sciences*, Vol. 45 pp. 1073-1083, 2006
- [5] Mala, G. M., and Li, D., Flow Characteristics of Water in Microtubes, *International Journal of Heat and Fluid Flow*, vol. 20, no. 2, pp. 142-148, 1999.
- [6] Choi,S.B., Barron, R.F., Warrington, R.O., Liquid flow and heat transfer in microtubes, *ASME Micromechanical sensors, Actuators and Systems*, Vol. 32, pp. 123-128,1991
- [7] Qu, W., Mala, G. M., and Li, D., Pressure-Driven Water Flows in Trapezoidal Silicon Microchannels, *International Journal of Heat and Mass Transfer*, vol. 43, no. 3, pp. 353-364, 2000.
- [8] Pfund, D., Rector, D., Shekarriz, A., Popescu, A., and Welty, J., Pressure Drop Measurements in a Microchannel, *American Institute of Chemical Engineers Journal*, vol. 46, pp. 1496-1507, 2000.
- [9] Wu, H. Y., and Cheng, P., An Experimental Study of Convective Heat Transfer in Silicon Microchannels with Different Surface Conditions, *International Journal of Heat and Mass Transfer*, vol. 46, no. 14, pp. 2547-2556,2003.
- [10] Li, Z. X., Du, D. X., and Guo, Z. Y., Experimental Study on Flow Characteristics of Liquid in Circular Micro-tubes, *Microscale Thermophysical Engineering*, vol. 7, pp. 253-265, 2003.
- [11] Li, Z., He, Y-L., Tang, G-H., and Tao, W-Q., Experimental and Numerical Studies of Liquid Flow and Heat Transfer in Microtubes, *International Journal of Heat and Mass Transfer*, vol. 50, no. 17-18, pp. 3447-3460, 2007.
- [12] Kays, W.M., and Crawford, M.E., *Convective Heat and Mass Transfer*, third ed., McGraw-Hill, 1993

AUTHOR BIBLIOGRAPHY

	<p>Ravindra Kumar is a Senior Lecturer in Department of Mechanical Engineering, G.B.Pant Institute of Technology, Okhla , New Delhi, India. He is M.E. in Thermal Engineering from Delhi College of Engineering and currently pursuing Ph.D. in Department of Mechanical Engineering, Jamia Millia Islamia, New Delhi, India. He is having 15 years teaching and 4 years industrial experience.</p>
	<p>Mohammad Islam is a senior professor in Department of Mechanical Engineering, Jamia Millia Islamia, New Delhi, India. He is engaged in teaching and research activities for last 25 years. Prof. M. Islam has published 72 papers in various national and internal journals and guided several M.Tech and Ph.D. thesis in Thermal areas. He is expert member of several reputed government organizations. He is also member of various selection committees of government and affiliated colleges. He is life member of deferent technical societies</p>
	<p>M. M. Hasan is a senior professor in Department of Mechanical Engineering, Jamia Millia Islamia, New Delhi, India. He is engaged in teaching and research activities for last 28 years. Prof. M.M.Hasan has published 65 papers in various national and internal journals and guided several M.Tech and Ph.D. thesis in Thermal areas. He is life member of various technical societies i.e. ISTE,ISME,Combustion Institute,SAE</p>



Regular Article

Facile synthesis of higher manganese silicide employing spark plasma assisted reaction sintering with enhanced thermoelectric performance

Saravanan Muthiah^{a,b}, R.C. Singh^b, B.D. Pathak^b, Ajay Dhar^{a,*}^a CSIR-Network of Institutes for Solar Energy, CSIR-National Physical Laboratory, Dr. K. S. Krishnan Marg, New Delhi 110012, India^b Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Article history:

Received 18 February 2016

Received in revised form 18 March 2016

Accepted 27 March 2016

Available online xxxx

Keywords:

Spark plasma sintering

Transmission electron microscopy

Thermoelectric materials

Microstructure

Figure-of-merit

ABSTRACT

This work reports a facile synthesis of aluminium doped higher manganese silicides using a single-step spark plasma assisted *in-situ* reaction sintering process, which takes only a few minutes as opposed to few hours employing conventional processing methods. Despite employing this cost-effective material processing technique, a figure of merit ~ 0.65 at 873 K was realized, which is comparable to the best value reported thus far. The phase purity, elemental distribution, morphology and structure was investigated employing X-ray diffraction, scanning, transmission electron microscopy and energy dispersive spectroscopy, based on which the thermoelectric performance has been discussed.

© 2016 Elsevier B.V. All rights reserved.

Power generation using thermoelectric devices is gaining momentum worldwide, as a green source of energy primarily owing to the recent environmental concerns and depleting fossil fuels reserves [1–3]. However, it is well recognized that in order for thermoelectric technology to compete with other established green energy generation technologies, like photovoltaics, the material and processing costs of thermoelectric modules, have to be drastically reduced [3–10]. Apart from this, as most of the thermoelectric efficient materials contain toxic elements adds another dimension to this problem from an environmental point of view. Thus the thrust of thermoelectric research is currently focussed on the development of efficient materials that are earth-abundant for mass production, non-toxic and also chemically stable at the device operating temperatures.

Higher manganese silicides (HMS) are a unique family of thermoelectric materials that satisfy these criteria, due to which they are being seriously explored as p-type materials in thermoelectric devices [11–25]. These have proven to be one of the best materials for cost-effective and stable thermoelectric power generation employing non-toxic constituent element for the mid-temperature thermoelectric device applications [2,11,24,25]. The added advantage of HMS is that they are the most promising p-type compatible counterpart to the existing n-type $\text{Mg}_2(\text{Si}, \text{Sn})$ thereby leading to a thermoelectric device which consists of earth-abundant & non-toxic elements [24]. However, the synthesis of p-type HMS is rather difficult due to its narrow composition range in the Mn–Si phase diagram [4]. A small variation in

composition can results in several similar HMS phases, such as, $\text{Mn}_{11}\text{Si}_{19}$ ($\text{MnSi}_{1.72}$), $\text{Mn}_{15}\text{Si}_{26}$ ($\text{MnSi}_{1.73}$), $\text{Mn}_{27}\text{Si}_{47}$ ($\text{MnSi}_{1.74}$) and Mn_4Si_7 ($\text{MnSi}_{1.75}$) [2–5,11–23]. All these are Nowotny Chimney-ladder phases which possess tetragonal crystal structure with nearly the same a-axis and an unusually elongated c-axis [4,5] and share similar properties thus are very hard to differentiate.

There are several reported studies on HMS, synthesized using different material processing routes, such as, liquid metallurgy, including arc melting [12,19,20,23], powder metallurgy employing mechanical alloying [11,13–16,21,22], and melt spinning [17,18], followed by consolidation of the synthesized HMS phases either using hot pressing [16,19,23] or spark plasma sintering (SPS) [11,13–15,17,18,21,22]. However, experimentally it is very difficult to control the precise composition of HMS phase due to the low vapour pressure of Mn and loss of Si during processing [13,14] coupled with a large difference in the melting points of its constituent elements. Added to this is the fact that HMS is not the eutectic composition in the phase diagram of Mn–Si and it is thus rather tedious to experimentally obtain the exact HMS composition, employing melting processes. In most of the reported studies, it has been observed that during the synthesis of HMS employing liquid [18–20,23] or powder metallurgy processes [13–16, 21,22,25], the high temperature melting, extended hours of ball milling and long durations of sintering lead to a different post-processing stoichiometry compared to its starting HMS composition. Thus the compositional control in HMS to a large extent dictates its thermoelectric properties, which perhaps accounts for the large scatter in their reported thermoelectric figure-of-merit (ZT) in the range of 0.28 to 0.65 [13–19,21–23,25].

* Corresponding author.

E-mail address: adhar@nplindia.org (A. Dhar).

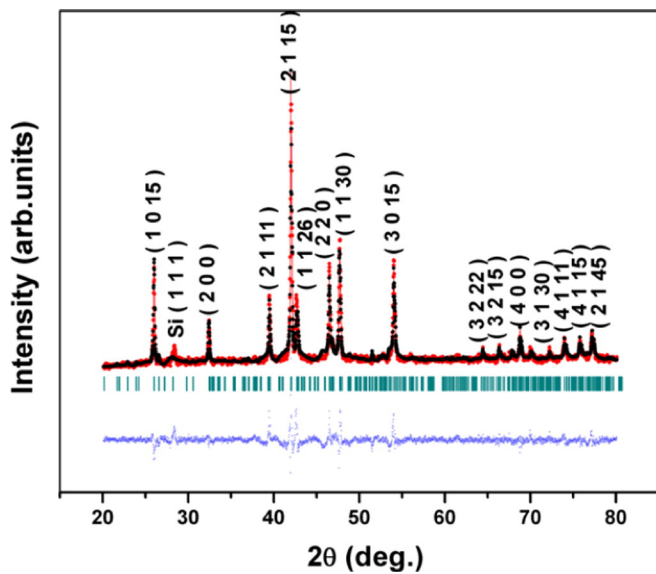


Fig. 1. XRD pattern of 5 at.% Al-doped $\text{MnSi}_{1.73}$ sample with Rietveld refinement. The red and black lines represent the observed and fitted data, respectively. The blue line represents the difference between the observed and fitted data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In order to enhance their ZT, HMS has also been doped with several elements, such as, Al [18], Ge [19], Cr [21], CNTs [22], SiGe [23] and double/triple doping (Ge, Mo, W, Al) [25]. Among these, the highest ZT ~ 0.65 was reported by Luo et al. [17] for Al-doped $\text{MnSi}_{1.73}$ phase, synthesized using melt-spinning followed by SPS. This enhanced ZT was attributed to the nanoscale crystallite size features due to rapid solidification employing SPS, which reduces the thermal conductivity owing to enhanced phonon scattering.

SPS is a pressure-assisted reaction sintering process in which high amperage pulsed direct current is employed to sinter the samples

under pressure and thus high sintering rates over large volumes can be realized. The kinetics and enhanced diffusion coefficients during SPS process assist in synthesizing compounds, which otherwise cannot be synthesized employing conventional processing methods. The SPS process makes it possible to sinter at lower temperatures and in shorter times because the particles are heated rapid by the spark plasma, which has a surface activating and cleaning effect on the particles being sintered. The different aspects of the SPS process and its mechanism are reported by Risbud and Han viewpoint set [26].

In the present study, we report a facile synthesis of HMS phase, 5 at.% Al-doped $\text{MnSi}_{1.73}$, with a ZT ~ 0.65 , which is comparable to the state-of-the-art reported thus far [18], employing a cost-effective and single-step process by *in-situ* reaction sintering of blended elemental powders employing SPS process. This good thermoelectric performance arises from a very low thermal conductivity (~ 1.7 W/mK) of the synthesized HMS samples, which originates from the low crystallite size (120–150 nm) of the HMS phase, obtained using the *in-situ* spark plasma assisted reaction sintering process. In this process, the *in-situ* reaction between the elemental powders and their subsequent consolidation takes place in a single step and thus the synthesis takes only a few minutes as opposed to few of hours employing conventional material processing techniques [11–23]. The synthesized HMS samples were characterized for phase purity, composition and microstructure using X-ray diffraction (XRD), field emission scanning electron microscopy (FESEM), high resolution transmission electron microscopy (HRTEM) and energy dispersive spectroscopy (EDS).

High purity Mn (99.95%) powders, Si (99.95%) powders, and Al (99.95%) powders (all from M/s Alpha Aesar), were weighed and blended in proper chemical stoichiometric proportion. The blended powders were handled in glove box (MB20, Mbraun, Germany) to avoid oxidation and other atmospheric contamination. These powders were then sintered at a heating rate of $100^\circ\text{C}/\text{min}$ using spark plasma sintering (SPS-725, Japan) at 1173 K for 10 min at 60 MPa in a graphite die & punch (12.7 mm diameter) under vacuum. The density of the sintered HMS samples was measured by the Archimedes principle and was found to be $>98.8\%$ of the theoretical density. X-ray diffraction analysis was carried out using monochromatic $\text{CuK}\alpha$ radiation employing

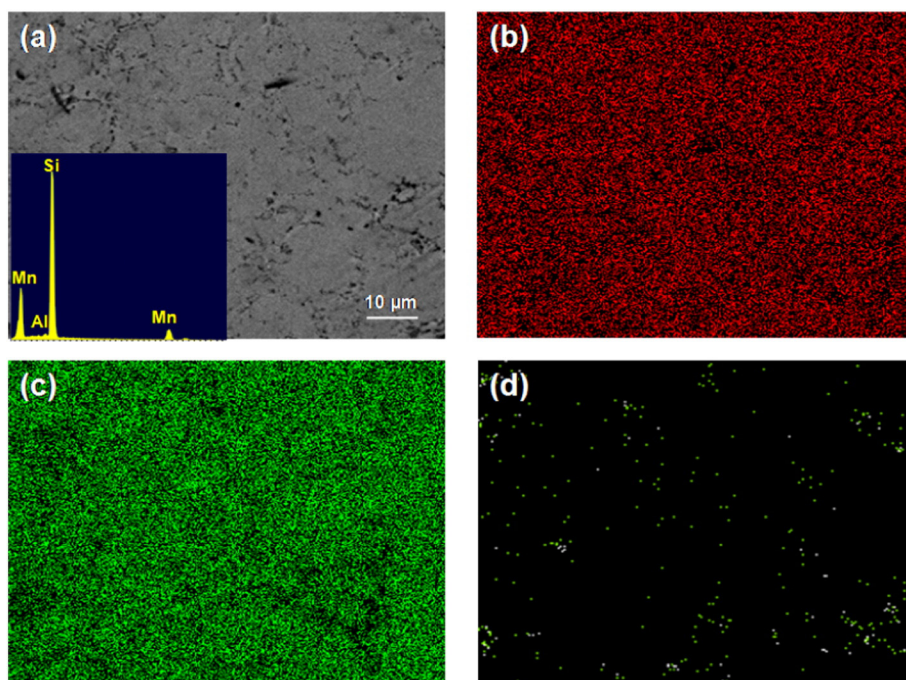


Fig. 2. SEM images of (a) as-synthesized HMS ($\text{MnSi}_{1.73}$) sample, and EDS spectrum (inset: A) showing the presence of Mn, Si and Al, and (b), (c) & (d) Individual Elemental maps of Mn, Si & Al, respectively.

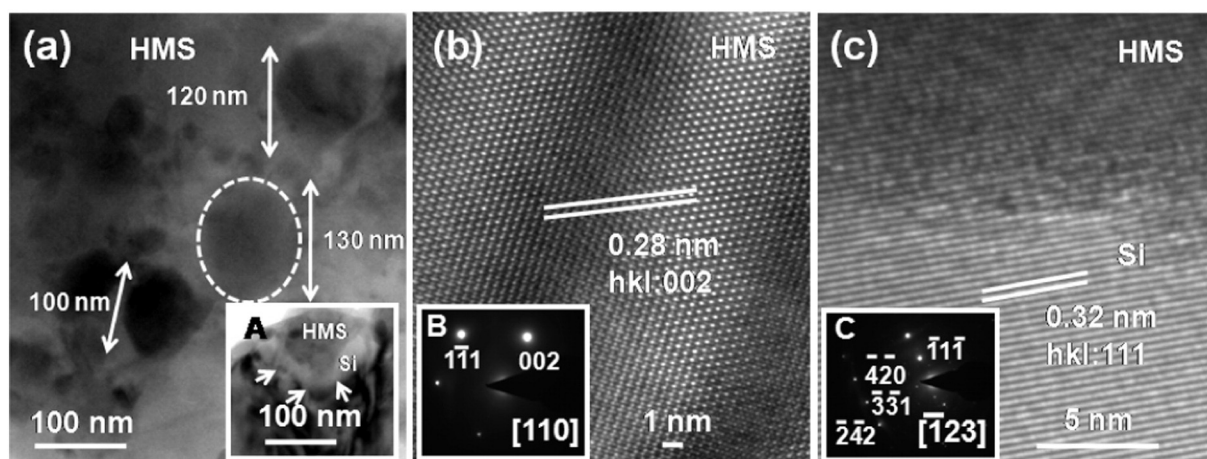


Fig. 3. HRTEM images of HMS Al-doped $\text{MnSi}_{1.73}$ sample showing (a) distribution of grains in the microstructure, (b) atomic scale image of a $\text{MnSi}_{1.73}$ grain, and (c) atomic scale image of $\text{MnSi}_{1.73}$ with the presence of Si. Insets: (A) $\text{MnSi}_{1.73}$ grain surrounded with Si, (B) electron diffraction pattern of $\text{MnSi}_{1.73}$, and (C) electron diffraction pattern of Si.

RigaKu X-ray diffractometer at a scan speed of 1° min^{-1} . The microstructure of the as-synthesized HMS samples was studied using FESEM (Zeiss, SUPRA V40, Germany) with an attached EDS. In order to elucidate the microstructural features of the HMS samples, a HRTEM (Tecnai G2F30 STWIN with field emission electron gun source, operated at 300 kV electron accelerating voltage) was employed.

Rectangular specimens ($3 \times 3 \times 12 \text{ mm}$) were cut from the centre of the HMS 12.7 mm diameter samples for the evaluation of Seebeck coefficient and electrical conductivity, which were measured by four-probe DC method in helium atmosphere using Ulvac-ZEM3. A circular disc sample of 12.7 mm diameter was used for thermal diffusivity measurement employing laser flash analyser (Linesis, LFA 1000). The specific heat capacity was measured by differential scanning calorimeter (Netzsch DSC 404 F1). Finally, thermal conductivity (κ) was calculated from the relation, diffusivity (α) \times specific heat (C_p) \times density (ρ). The accuracies in thermoelectric property measurements are: $\pm 3\%$ for thermal diffusivity, $\pm 10\%$ for electrical conductivity, $\pm 6\%$ for the Seebeck coefficient, $\pm 8\%$ for specific heat and $\pm 0.5\%$ for density.

The X-ray diffraction pattern of the Al-doped $\text{MnSi}_{1.73}$ sample is shown in Fig. 1. The Rietveld refinement of the XRD data suggests the formation of $\text{MnSi}_{1.73}$ phase and the calculated lattice parameters ($a = 0.55$, $b = 0.55$ and $c = 6.55 \text{ nm}$) are in good agreement with the corresponding JCPDS data file (no. 00-020-0724), however, no apparent change in the lattice parameters was observed due to Al addition. The XRD pattern although suggests the presence of an additional Si-phase, but it is devoid of the cubic-MnSi phase, which generally is an inevitable by-product in the synthesis of HMS [14–16] and is known to deteriorate the thermoelectric properties. The elemental mapping of $\text{MnSi}_{1.73}$ sample also confirms a nearly uniform dispersion of Mn, Si & Al (Fig. 2) and the inset of Fig. 2(a) shows the EDS spectrum suggests that the composition of the synthesized HMS sample is very close to its initial stoichiometry.

HRTEM studies were performed on the synthesized $\text{MnSi}_{1.73}$ sample, in real and reciprocal space, which led to several important features pertaining to the microstructure and phase formation. Fig. 3(a) exhibits a typical TEM micrograph which indicated the grains in the range ~ 120 – 150 nm . It was noted that the individual grains reveal good crystallinity, which is apparent from the high resolution TEM micrographs (Fig. 3(b) & (c)). Fig. 3(b), which shows the lattice scale image of a typical 130 nm grain (region encircled with white dotted line in Fig. 3(a)), clearly reveals a set of atomic planes along $hkl:002$ with interatomic spacing of 0.28 nm of the tetragonal phase of $\text{MnSi}_{1.73}$ ($\text{Mn}_{15}\text{Si}_{26}$, tetragonal, space group: $I\bar{4}2d$, lattice parameters: $a = 0.55$, $b = 0.55$, $c =$

6.56 nm , JCPDS no.: 00-020-0724). The inset in Fig. 3(b) shows a corresponding selected area electron diffraction pattern (SAEDP) recorded along $[110]$ zone axis of a tetragonal crystal structure of $\text{MnSi}_{1.73}$ sample. A set of important planes, $hkl:002$ and $1\bar{1}\bar{1}$ are also marked on the SAEDP in the reciprocal space. The lattice scale images as shown in Fig. 3(c) clearly show the presence of excess Si, surrounding the $\text{MnSi}_{1.73}$ grains, although as a minor phase. The inset in Fig. 3(a) clearly delineates such a microstructure of a $\text{MnSi}_{1.73}$ grain with the presence Si (as a light grey contrast around the periphery marked with white dotted lines). A corresponding atomic scale image (Fig. 3c) clearly shows a distinct set of Si-atomic planes with the interplanar spacing of 0.32 nm ($hkl:111$, fcc crystal structure of Si) stacked along at the interface of $\text{MnSi}_{1.73}$. The SAEDP recorded from a Si rich region along $[\bar{1}23]$ zone axis exhibits a cubic crystal structure of Si with important diffracted planes, $hkl:\bar{1}\bar{1}\bar{1}$, $\bar{2}4\bar{2}$, $\bar{3}\bar{3}\bar{1}$, $\bar{4}\bar{2}0$, marked on an inset C in Fig. 3(c). The XRD and HRTEM analysis clearly confirm the formation $\text{MnSi}_{1.73}$ phase (120 – 150 nm) with some excess Si, although devoid of MnSi phase.

Fig. 4(a–c) show the temperature dependence of the electrical and thermal transport properties of the as-synthesized $\text{MnSi}_{1.73}$. The Seebeck coefficient (Fig. 4(a)) initially increases with increasing temperature from $\sim 128 \mu\text{V/K}$ at 323 K to a peak value of $\sim 211 \mu\text{V/K}$ at 773 K , although a small decrease is observed with further increase in temperature, which can be attributed to the thermal excitation of electrons across the band gap. This observed temperature dependence is similar both in behaviour and magnitude to the other reported studies on similar HMS phases [14–19,21,23]. The positive Seebeck coefficient over the entire measured temperature range confirms the p-type conduction in this HMS phase. The electrical conductivity, as shown in Fig. 3(a), suggests a continuous decrease with increasing temperature with a near-saturation at higher temperatures ($\sim 3 \times 10^4 \text{ S/m}$). This behaviour is typical of a degenerate semiconductor and close to that reported earlier for Al [18], Ge [19] and CNT [22] doped HMS phase. Fig. 4(c), which shows the temperature dependence of the calculated power factor ($\alpha^2\sigma$) of the synthesized HMS sample, suggests the magnitude of the power factor $\sim 1.28 \times 10^{-3} \text{ W/mK}^2$ at 873 K , which is comparable to those reported earlier [14,19].

The temperature dependence of the total thermal conductivity, shown in Fig. 4(b), suggests a very low value thermal conductivity $\sim 1.7 \text{ W/mK}$ at the highest temperature of 873 K , which is the lowest reported thus far for HMS containing no additional MnSi phase, although much lower values of thermal conductivity have been reported for HMS containing precipitated MnSi phase [14–16,18,19,21,22,25]. However, it

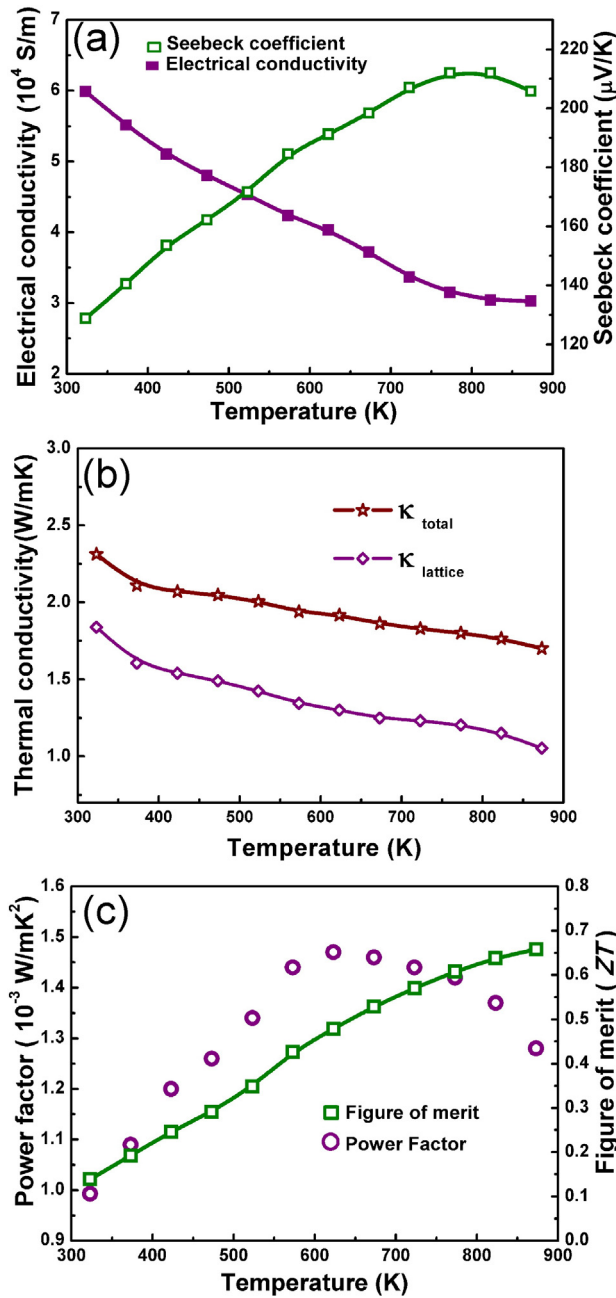


Fig. 4. Temperature dependence of the thermoelectric properties of synthesized MnSi_{1.73} sample (a) Seebeck coefficient and electrical conductivity, (b) total and lattice thermal conductivity, (c) power factor and thermoelectric figure-of-merit.

has been well documented by several researchers that although MnSi is a stable intermetallic compound but it deteriorates the transport properties of the HMS phase, despite favourably decreasing the thermal conductivity, thus leading to an overall deterioration in the ZT [16,19,21,22]. The temperature dependence of lattice thermal conductivity, as shown in Fig. 3(b) was obtained by subtracting the electronic contribution, (calculated from the electrical conductivity, using the Wiedemann–Franz law) from the total thermal conductivity considering the Lorentz number $2.45 \times 10^{-8} \text{ W}\Omega\text{K}^{-2}$. It is apparent from this figure that the total thermal conductivity has a major contribution from its lattice counterpart, which generally can be reduced by several structural modifications in the thermoelectric material, including engineering its crystallite size. The heat-carrying phonons, in general, have a large spectrum

of wavelengths with different wavelengths contributing different amounts to the effective heat conduction. It is well known that in several thermoelectric materials, such as, PbTe [27] PbSeTe [28] and Si [29] most of the heat is transported by mid-to-long wavelength phonons [30–33] thus the meso-scale features are expected to play a dominant role in reducing the lattice conductivity, owing to phonon scattering [9,34]. In our synthesized HMS MnSi_{1.73} samples, the crystallite size was found to be in the range of 120–150 nm, (as evidenced from XRD and TEM studies). Thus, the reduction in thermal conductivity in our synthesized HMS samples can primarily be attributed to the mid-to-long wavelength phonons scattering by the meso-scale crystallites, which introduce a high density of grain boundaries leading to abundant grain boundaries participating in phonon scattering process.

The meso-scale crystallite size (~120–150 nm) in our MnSi_{1.73} samples can be attributed to their synthesis employing spark plasma reaction sintering, which is known to yield alloys with meso-scale crystallite sizes [34–36]. During SPS, a rapid sintering rate is applied to the powdered green compact under pressure so that the sample spends little time in the low-temperature regime, where non-densifying mechanisms, such as, surface diffusion are active, and proceeds quickly to higher temperatures where densifying mechanisms, such as, grain boundary or volume diffusion are dominant. Since grain growth occurs in both these regimes, it is believed that the use of rapid sintering rates can greatly impede the grain growth while simultaneously improving the final sintered density of the product [9]. There are large number of reports in the literature on the synthesis of several alloys using spark plasma reactive sintering wherein meso-scale crystallite size ~100–200 nm has been observed [9,34–36].

The calculated temperature dependence of ZT ($\alpha^2\sigma/\kappa$) for the synthesized MnAl_{1.73} sample, as shown in Fig 4(c) exhibits a ZT ~ 0.65 at 873 K, which is the best reported value thus far for HMS. Although, Luo et al. [18] have reported a similar ZT value, but their material processing involved melt-spinning followed by consolidation employing spark plasma sintering process, which is both time-consuming as well as expensive. On the contrary, we report a similar value of ZT in HMS albeit employing a single-step spark plasma assisted reaction-sintering process, which takes a few minutes compared to several hours required using other material processing techniques employed for synthesizing thermoelectric HMS [13–23]. The good thermoelectric performance in the synthesized MnAl_{1.73} sample, despite a moderate value of power factor, is primarily due to its very low thermal conductivity (~1.7 W/mK), which arises due to the meso-scale crystallites (120–150 nm) in the HMS, synthesized employing plasma assisted reaction-sintering.

In summary, we report the synthesis of Al-doped MnSi_{1.73} devoid of any additional MnSi phase, employing spark plasma sintering, wherein the *in-situ* reaction between the elemental powders and their subsequent consolidation takes place in a single step and in a few minutes as opposed to few hours required for processing the same material employing conventional methods. The phase purity, composition, elemental distribution, morphology and structure of the Al-doped MnSi_{1.73} was confirmed employing X-ray diffraction, FESEM, HRTEM and EDS. Using this simple and cost-effective material processing technique, a ZT ~ 0.65 at 873 K was realized in 5 at.% Al-doped MnSi_{1.73}, which is comparable to the best value reported in the literature. Higher manganese silicides are one of the most stable thermoelectric materials, which are cost-effective as well as environmental friendly owing to which they are currently being seriously explored as a p-type counterpart to the well-known n-Mg₂(Si,Sn).

This work was supported by the CSIR-TAPSUN (Network Project NWP-54) programme entitled “Novel approaches for solar energy conversion under technologies and products for solar energy utilization through networking”. The authors would also like to thank Dr. A.K. Srivastava for TEM analysis and Dr. Bhasker Gahtori for XRD analysis. The authors are also grateful to Radhey Shyam and N. K. Upadhyay for their technical support.

References

- [1] W. Liu, Q. Jie, H.S. Kim, Z. Ren, *Acta Mater.* 87 (2015) 357–376.
- [2] K. Koumoto, R. Funahashi, E. Guilmeau, Y. Miyazaki, A. Weidenkaff, Y. Wang, C. Wan, *J. Am. Ceram. Soc.* 96 (1) (2013) 1–23.
- [3] C. Dames, *Scripta. Mater.* 111 (2016) 16–22.
- [4] L.D. Ivanova, A.A. Baikov, *J. Thermoelectricity* 3 (2009) 60.
- [5] Y. Miyazaki, D. Igarashi, K. Hayashi, T. Kajitani, K. Yubuta, *Phys. Rev. B.* 78 (2008) 214104.
- [6] J.J. Pulikkotil, D.J. Singh, S. Auluck, M. Saravanan, D.K. Misra, A. Dhar, R.C. Budhani, *Phys. Rev. B.* 86 (2012) 155204.
- [7] S. Muthiah, J. Pulikkotil, A.K. Srivastava, A. Kumar, B.D. Pathak, A. Dhar, R.C. Budhani, *Appl. Phys. Lett.* 103 (2013) 053901.
- [8] S. Muthiah, B. Sivaiah, B. Gahtori, K. Tyagi, A.K. Srivastava, B.D. Pathak, A. Dhar, R.C. Budhani, *J. Electron. Mater.* 43 (2014) 2035–2039.
- [9] B. Gahtori, S. Bathula, K. Tyagi, M. Jayasimhadri, A.K. Srivastava, S. Singh, R.C. Budhani, A. Dhar, *Nano Energy* 13 (2015) 36–46.
- [10] K. Tyagi, B. Gahtori, S. Bathula, V. Toutam, S. Sharma, N.K. Singh, A. Dhar, *Appl. Phys. Lett.* 105 (2014) 261902.
- [11] G. Gurjar, P.S. Khare, S. Muthiah, B.D. Pathak, Ajay Dhar, *Int. J. Electron. Electr. Eng.* 7 (2014) 287–292.
- [12] A. Allam, P. Boulet, C.A. Nunes, J. Sopousek, P. Broz, M.C. Record, *J. Alloys. Comp.* 551 (2013) 30–36.
- [13] Y. Sadia, Y. Gelbstein, *J. Electron. Mater.* 41 (2012) 1504–1508.
- [14] T. Itoh, M. Yamada, *J. Electron. Mater.* 38 (2009) 925–929.
- [15] A. Famengo, S. Battiston, M. Saleemi, S. Boldrini, S. Fiameni, F. Agresti, M.S. Toprak, S. Barison, M. Fabrizio, *J. Electron. Mater.* 42 (2013) 2020–2024.
- [16] D.K. Shin, K.W. Jang, S.C. Ur, I.H. Kim, *J. Electron. Mater.* 42 (2013) 1756–1761.
- [17] W. Luo, H. Li, Y. Yan, Z. Lin, X. Tang, Q. Zhang, C. Uher, *Intermetallics* 9 (2011) 404–408.
- [18] W. Luo, H. Li, F. Fu, W. Hao, X. Tang, *J. Electron. Mater.* 40 (2011) 1233–1237.
- [19] A.J. Zhou, T.J. Zhu, X.B. Zhao, S.H. Yang, T. Dasgupta, C. Stiewe, R. Hassdorf, E. Mueller, *J. Electron. Mater.* 39 (2009) 2002–2007.
- [20] V. Ponnambalam, D.T. Morelli, *J. Electron. Mater.* 41 (2011) 1389–1394.
- [21] G. Liu, Q. Lu, X. Zhang, J. Zhang, Y. Shi, *J. Electron. Mater.* 41 (2012) 1450–1455.
- [22] D.Y. Nhi Truong, H. Kleinke, F. Gascoin, *Dalton Trans.* 43 (2014) 15092–15097.
- [23] A.J. Zhou, X.B. Zhao, T.J. Zhu, S.H. Yang, T. Dasgupta, C. Stiewe, R. Hassdorf, E. Mueller, *Mater. Chem. Phys.* 124 (2010) 1001–1005.
- [24] Y. Thimont, Q. Lognone, C. Goupil, F. Gascoin, E. Guilmeau, *J. Electron. Mater.* 43 (2014) 2023–2028.
- [25] D.Y. Nhi Truong, D. Berthebaud, F. Gascoin, H. Kleinke, *J. Electron. Mater.* 44 (2015) 3603–3611.
- [26] S.H. Risbud, Y.-H. Han, *Scripta. Mater.* 69 (2013) 105–206.
- [27] Z.G. Chen, G. Han, L. Yang, L. Cheng, J. Zou, *Prog. Nat. Sci.: Mater. Int.* 22 (2012) 535–549.
- [28] Q. Hao, G. Zhu, G. Joshi, X. Wang, A. Minnich, Z. Ren, G. Chen, *Appl. Phys. Lett.* 97 (2010) 063109.
- [29] G. Zhu, H. Lee, Y. Lan, X. Wang, G. Joshi, D. Wang, J. Yang, D. Vashae, H. Guilbert, A. Pillitteri, *Phys. Rev. Lett.* 102 (2009) 196803.
- [30] G.J. Snyder, E.S. Toberer, *Nat. Mater.* 7 (2008) 105–114.
- [31] D.M. Rowe, *CRC Handbook of Thermoelectrics*, CRC Press, 1995.
- [32] G. Chen, *Int. J. Therm. Sci.* 39 (2000) 471–480.
- [33] W. Kim, J. Zide, A. Gossard, D. Klenov, S. Stemmer, A. Shakouri, A. Majumdar, *Phys. Rev. Lett.* 96 (2006) 045901.
- [34] S. Bathula, M. Jayasimhadri, B. Gahtori, N.K. Singh, K. Tyagi, A.K. Srivastava, Ajay Dhar, *Nanoscale* 7 (2015) 12474–12483.
- [35] D.V. Dudina, A.K. Mukherjee, *J. Nanomater.* (2013) 625218.
- [36] G. Cabouro, S. Chevalier, E. Gaffet, Yu Grin, F. Bernard, *J. Alloys. Comp.* 465 (2008) 344–355.

Forecasting Volatility Using GARCH: A Case Study

Nand Kumar¹, Rishabh Verma², Puneet Gupta³

¹ (Department of Applied Science & Humanitie ,Delhi Technological University, India)

^{2,3} (Department of Mathematics ,Delhi Technological University, India)

¹nand.dce@gmail.com

Abstract : Forecasting volatility is fundamental to the risk management process in order to price derivatives, devise hedging strategies and estimate the financial risk of a firm's portfolio of positions. In recent years, Autoregressive Conditional Heteroscedasticity (ARCH) type models have become popular as a means of capturing observed characteristics of financial returns like thick tails and volatility clustering. These models use time series data on returns to model conditional variance Our study shows that GARCH volatility captures the most information of future volatility The implied volatility calculated from the study subsumes only 46% of realized volatility whereas GARCH Volatility subsumes 70% of realized volatility, therefore, Garch volatility is a better measure of volatility in option pricing.

Keywords (11Bold): Geometric Brownian motion, Black- Scholes model, Implied Volatility, ARCH, GARCH Volatility.

I. INTRODUCTION

Since the introduction of the Black-Scholes model (1973), researchers have studied the empirical performance of the model. The Black-Scholes model explains that the price of heavily traded assets follow a geometric Brownian motion that looks like a smile or smirk with constant drift and volatility. When applied to a stock option, the model incorporates the constant price variation of the stock, the time value of money, the option's strike price and the time to the option's expiry. According to the geometric Brownian motion model, the returns on a certain stock in successive, equal periods of time are independent and normally distributed. Thus they form a Markov process. The assumptions on which this model is based meet the financial market laws and rules imposed by the Market Efficiency Hypothesis. These rules and laws suppose that only present information about a stock is sufficient to determine the future price of that stock. So theoretically the geometric Brownian motion seems to be a good way to model future stock.

The next major step of this paper is to determine estimates of volatility. In recent years, Autoregressive Conditional Heteroscedasticity (ARCH) type models have become popular as a means of capturing observed characteristics of financial returns like thick tails and volatility clustering. These models use time series data on returns to model conditional variance. An alternative way to estimate future volatility is to use options prices, which reflect the market's expectation of volatility. Day and Lewis (1993) compare the relative information content and predictive power of implied volatility and ARCH-type forecasts for crude oil futures. A similar study by Xu and Taylor (1996) examines the informational efficiency of the PHLX currency options market in predicting volatility. Duffie and Gray (1995) compare the forecasting accuracy of ARCH type models, Markov switching models,

and implied volatilities for crude oil, heating oil and natural gas markets.

Our study attempt to test the hypothesis that GARCH volatilities subsume information contain in returns and provides the best month-ahead volatility forecasts for SBI equity for which we have taken closing price data of SBI equity from 23rd October 2007 to 30th October 2013. However for finding implied volatility, data is considered from 26th July, 2013 to 30th October, 2013 because the options in Indian Capital Markets have been introduced only recently and sufficient time-series options data is not available. Also we have divided our data in In-Sample and Out-of-Sample data and performed regression analysis on both. The Out-of-Sample data we have considered for OLS regression is from 8th October 2013 to 30th October 2013.

II. DATA AND METHODOLOGY

Geometric Brownian Motion- To simulate the stock price using Geometric Brownian Motion we performed Monte Carlo Simulation on the SBI stock price from 2nd August 2013 to 8th October 2013. We take the closing price of the stock on 2nd August 2013(Rs. 1680.60) as time 0. Referring the above mentioned price as S_0 we find one of the outcomes of day 1 by using the log normal property.

$$\ln S_1 \sim \emptyset [\ln S_0 + (\mu - \sigma^2/2)T, \sigma^2 T]$$

Table 1: Calculated Parameters for Monte Carlo Simulation

μ	-0.068%
σ	2.804%
Price	1680
annual volatility	44.518%

The average (μ) and standard deviation(σ) of log return is calculated. Thus the constant annualized volatility comes out to be 44.518% shown in the above table. Then stimulation is done by Monte Carlo Simulation using μ and σ as parameters. Various movements of stock prices observed are shown in figure 1.a,b,c.

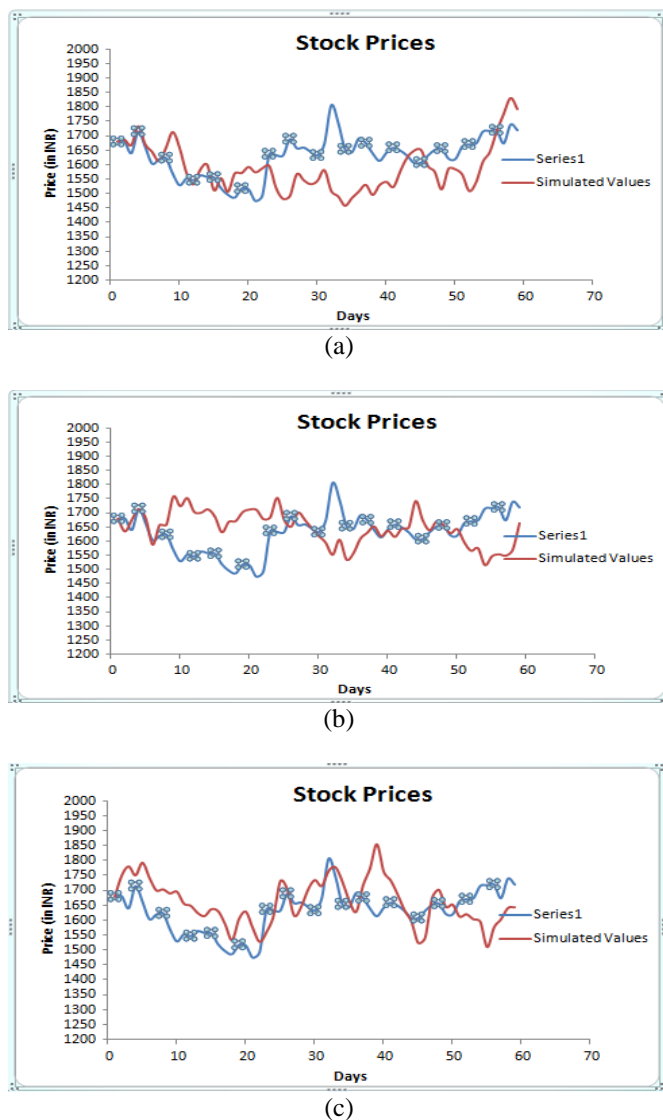


Figure 1: a,b,c Various movements of stock prices observed

Relaxing the assumption of constant volatility in Black Scholes Model, we have considered various stochastic

volatility such as Implied volatility, Historical volatility, EWMA volatility and GARCH volatility in our study.

Implied Volatility- Implied volatility of an option contract is that value of the volatility of the underlying instrument which when feeded as an input in an option pricing model (such as Black-Scholes) will return a theoretical value equal to the

current market price of the option. Thus $\sigma_{c,t}$ and $\sigma_{p,t}$ was calculated for each day from 26th July ,2013 to 8th October,2013 (for In-Sample comparison) by using the stock price , strike price, rate of interest (4 percent), option market

price, time of expiration. Here $\sigma_{c,t}$ and $\sigma_{p,t}$ denote the implied volatility of call and put options respectively at time t. And weighted average of both the volatilities is considered as follows:

$$\sigma_t = (\sigma_{c,t} + \sigma_{p,t}) / 2$$

Such a weighted average is simple to implement and it avoids the noise , created by single implied volatility , that might impact an observation .

Historical volatility- Merton (1980) has shown that the accuracy of an estimate of volatility using past volatility increases with the sampling frequency within a given overall observation period. We thus choose to use daily data and preferred to omit the usual estimator of the mean to avoid excessive noise, and use the following formula:

$$\sigma_{h,t} = \sqrt{\frac{252}{\tau_t} \sum_{i=t-\tau_t}^t R_i^2}$$

where R_i denotes the log-return on day i and is calculated as

$$R_i = \ln(S_i/S_{i-1}),$$

where S_i is the index level on the same day i .

Exponentially-weighted average volatility

Assuming that volatility varies with time, the EW version compensates to some extent for one of the shortcomings of simple historical volatility by giving greater weight to the recent observation.

For each closing price observation at time t, we also measure the volatility using an exponentially-weighted average of past daily volatility, including the day of the observation, with a decay factor of 0.94 . We use the formula:

$$\sigma_{f,t} = \sqrt{252 * 0.06 * \sum_{i=0}^t 0.94^n (R_i^2)}$$

where 0.94 is the decay factor, 0.06 the sum of the weights, and where R denotes the log-return on day i .

We calculated all volatilities from 23rd October, 2007 to 8th October, 2013 on all trading days

GARCH

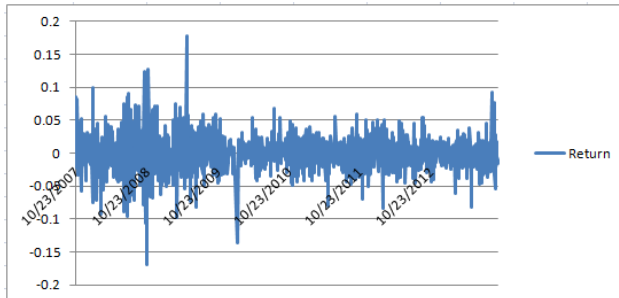


Figure 2 Return series from 23rd October 2007 to 8th October 2013

Returns series are preferred over prices in analysis of financial time series because they have attractive statistical properties like stationarity.

But here the actual distribution of return series has fatter tails compared to fitted normal distribution. Thick tails can be modeled by assuming a “conditional” normal distribution for returns; where conditional normality implies that returns are normally distributed on each day, but that parameters of the distribution change from day to day. Returns are thus not identically distributed with mean 0 and variance σ^2 at each point in time. Instead, it is fair to say that σ^2 changes with time t . The persistence of volatility in option market is an indication of autocorrelation in variances.

Testing for Stationarity

Ljung box test is used for testing the stationarity character of return series and return square series.

The null hypothesis for this test is that the first m autocorrelations ($\rho_1, \rho_2, \dots, \rho_m$) are jointly zero.

$$H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0$$

Using matlab the p value for return series is .0612. Thus our null hypothesis is accepted i.e., ($h=0$).

But in case of return square series, p value is approximately zero. Thus our null hypothesis is rejected which concludes that return square series is not stationary.

Testing for ARCH effect

Using Matlab the null hypothesis is rejected ($h = 1, p = 0$) in favour of the ARCH(2) alternative. The F statistic for the test is 44.53, much larger than the critical value from the χ^2 distribution with two degrees of freedom, 5.9915. The test concludes that there is significant volatility clustering in the residual series.

Also from correlogram diagram as shown below we can say that there is no significant autocorrelation in return series (Fig 3). But there is significant autocorrelation in return square series

(Fig 4) suggesting volatility dependent on past volatilities.

ACF

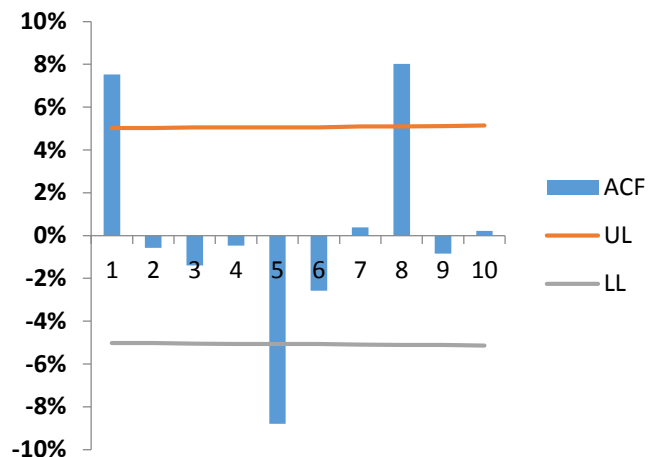


Figure 3. Correlogram showing autocorrelation of return series at various lags

ACF

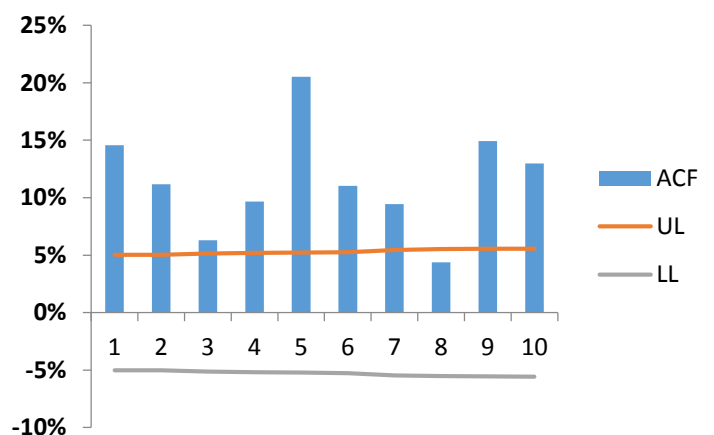


Figure 4. Correlogram showing autocorrelation of return square series at various lags

Time varying volatility is modeled statistically by estimating a conditional variance equation in addition to the returns generating process. In practice, the GARCH (1, 1) model comprising only three parameters in the conditional variance equation is sufficient to capture the volatility clustering in the data. The conditional variance equation of GARCH (1,1) model is:

$$\sigma_t^2 = a_0 + a_1(r_{t-1})^2 + \beta(\sigma_{t-1})^2$$

Return series is considered from 23rd October, 2007 to 8th October, 2013 (for In-Sample comparison).

We have calculated the parameters (a_0, a_1, β) using maximum log likelihood method in excel.

Table 2: Long Run Volatility

Likelihood	3556.3255
ω	7.26144E-06
α_1	0.05
β_1	0.94
α_o	0.01
Long run volatility	
σ	2.69%
Annualized σ	42.78%

Long run annual volatility came out to be 42.78 percent. Using GARCH(1,1) model and estimated parameters, we calculated volatilities from 23rd October 2007 to 8th October 2013.

Next we predict volatilities by different models to perform the OLS regression against realized volatilities.

Realized Volatility

For each stock price observation at time t , we measure the Realized volatility by the sample standard deviation of the daily index returns over the remaining life R_t of the option. Again, we deliberately omit the estimator of the mean, which would have been too noise-sensitive. We use the following formula:

$$\sigma_{f,t} = \sqrt{\frac{252}{\tau_t} \sum_{i=t+1}^{t+\tau_t} R_i^2}$$

where R_i denotes the log-return on day i .

The study attempts to report the results of OLS regression of the realized volatility on In-Sample values and forecast values given by the various models and estimators. Also in order to compare bias and efficiency of various estimators and models for estimation, we have calculated the following errors such as

Bias = $E(\sigma_{rst} - \sigma_t)$,

Mean square error MSE) = $E[(\sigma_{rst} - \sigma_t)^2]$,

Relative bias = $E[(\sigma_{rst} - \sigma_t)/\sigma_t]$,

Mean absolute error (MAE) = $E[Abs(\sigma_{rst} - \sigma_t)]$

RESULTS AND DISCUSSIONS

Stock Prices of SBI equity were simulated by Geometric Brownian motion. The behavior of simulated prices was found to be similar to the path followed by closing prices of the given equity as shown earlier.

Different volatilities were calculated for all trading days from 23rd October 2007 to 8th October 2013 and each volatility was plotted against time as below:

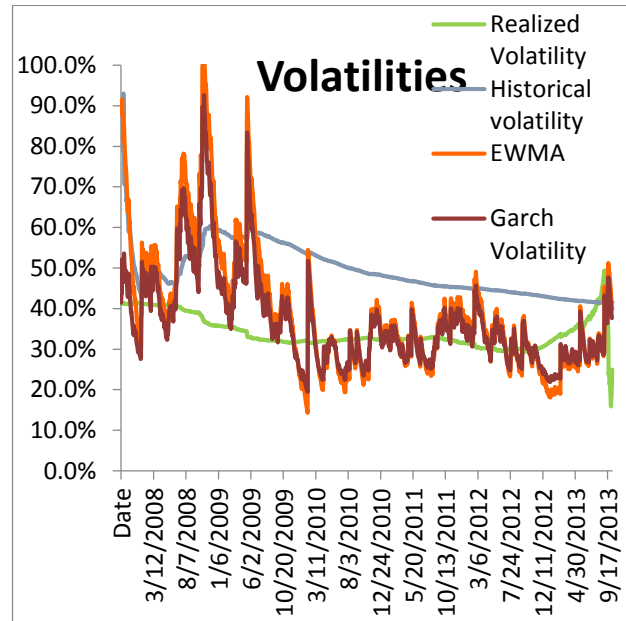


Figure 5. showing the relationship between realized volatility, historical, implied and GARCH volatilities.

Errors were calculated for each type of volatility as shown in table 1a(appendix). Historical Volatility was rejected as it has maximum error. The Bias, Mean square error and Mean absolute error for EWMA was greater as compared to these errors found in implied volatility. The comparison between GARCH and Implied Volatility cannot be conclusive based on these errors.

Therefore to get a clearer picture, we regress all the four volatilities with the realized volatility.

Performing regression on In-Sample data between Historical and Realized Volatility (Table 2b), the R-square comes out to be 0.027501 and a negligible p-value. In the case of EWMA (Table 2c), the R-square comes out to be 0.164429 and the p-value is again negligible. In the case of GARCH (Table 2d), the R-square is 0.204015 and p-value is negligible. For the regression between Implied and Realized Volatility (Table 2a), R-square comes out to be 0.265774 and a negligible p-value. The study suggest that in case of regression on In-Sample data no volatilities is explaining the realized volatility to the satisfactory level.

Similarly performing regression on Out-of-Sample data between Historical and Realized Volatility (Table 2f), the R-square comes out to be **0.271066** and a negligible p-value. In the case of EWMA (Table 2g), the R-square comes out to be **0.575202** and the p-value is again negligible. In the case of Implied (Table 2e), the R-square is **0.467892** and a negligible p-value. For the regression between GARCH and Realized Volatilities (Table 2h), R-square comes out to be **0.723662** and a negligible p-value. The study suggest that GARCH Volatility outperforms all the other three volatilities as it best explains the realized volatility in the case of Out-of-Sample data.

V. CONCLUSIONS

In this paper, firstly we have validated the performance of geometric Brownian motion on SBI stock price by seeing the path followed by simulated prices and closing prices. Then we have compared different volatilities to find out the correct measure of volatility which can be used as an input in Black Scholes formula. The results suggest that errors for GARCH volatility are found to be less as compared to other volatilities. The OLS regression of forward looking implied volatility is less than 50% which implies that implied volatility subsumes less part of future volatility whereas the OLS regression for forecasted GARCH volatilities against realized volatility, is more than 70% which suggest that GARCH volatility subsumes maximum part of future volatility. Thus GARCH forecasted volatility comes out to be an effective measure of volatility and may be used by traders and hedgers in indian derivatives market.

REFERENCES

- [1] Abdelmoula Dmouj (2006,Nov). Stock price modeling: Theory and Practice. Faculty of Sciences, Amsterdam:BMI Paper.
- [2] Ahoniemi (2005). Modeling and Forecasting implied volatility – an econometric analysis of the VIX index, Helsinki Centre of Economic Research.
- [3] Ajay Pandey (2005,June). Volatility Models and their Performance in Indian Capital Markets, Vikalpa (Vol. 30) , 11
- [4] Akgiray, V (1989). Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts, *Journal of Buisness*, 62(1)
- [5] Bollerslev t. (1986) ,Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31 (1986) . North-Holland
- [6] Carol Alexander (1998), Risk Management and Analysis: Measuring and Modelling Financial Risk, Wileys
- [7] Christensen, B.J & Prabhala, N.R (1978). The relation between implied and realized volatility, *Journal of Financial Economics*
- [8] Christensen,B.J.& Hansen, C.S (2002) . Stock market volatility and the information content of stock index options, *Journal of Financial Economics*.
- [9] Day , T.E & Lewis, C.M(1992). Stock market volatility and the information content of stock index options, *Journal of Financial Economics*.
- [10]Engle, RF and Patton, AJ(2001). What Good is a Volatility Model? .*Quantitative Finance*,1(2).
- [11]Fama, EF(1965). The Behaviour of Stock-Market Prices. *Journal of Business*,38(1)
- [12]Hamao,Y,Masulis, RW and Ng, VK(1990). Correlations in Price Changes and Volatility across International Stock Markets, *Review of Financial Studies*, 3(2).
- [13]Huang Kun (2011), Modeling Volatility of S&P 500 Index Daily Returns: A comparison between model based forecasts and implied volatility.
- [14]Jaykumar Dagha(2006-08).Derivatives: Forwards,Futures & Options, The Best Way to Mitigate Volatility. Alliance Buisness School
- [15]John C. Hull, Options, Futures and Other Derivatives . 7th Ed. Pearson Education :Prentice Hall Publishers(2008), 4-5
- [16]Namita Sharma (1998). Forecasting Oil Price Volatility, Faculty of the Virginia Polytechnic Institute and State University, 2-3
- [17]Peter Gross (2006). Parameter Estimation for Black-Scholes Equation, URA
- [18]S.C. Gupta and V.K. Kapoor (1970). Fundamentals of Mathematical Statistics. 11th ed. Sultan Chand & Sons
- [19]Schwert, G W (1989). Why Does Stock Market Volatility Change over Time?. *Journal of Finance*,44(5)
- [20]Sheldon M. Ross (2010), Introduction to Probability and Statistics for Engineers and Scientists . 4th ed. Academic Press an Imprint of Elsevier
- [21]Vladimir M. Ionesco (2011). The performance of implied volatility in forecasting future volatility: an analysis of three major equity indices from 2004 to 2010. MIT Sloan School of Management, 6-8

Appendix:

Table 1a: Bias, Mean Square Error, Relative Bias, Mean Absolute error and OLS regression on In-Sample data and OLS regression on Out- of- Sample data

Volatility	Bias	Mean Square Error	Relative Bias	Mean Absolute error	Regression square(R ²) (In-Sample)	Regression Square(R ²) (Out-of-Sample)
Historical Volatility	.152979	.0282405	.464096	.154111	.027501	.271066
EWMA	.052231	.0238237	.14816	.10714	.164429	.575202
Implied Volatility	.041756	.01018	.16825	.08125	.265774	.467892
GARCH Volatility	.032462	.01294	.10525	.077155	.204015	.723662

Table 2a: OLS regression between Implied and Realized Volatility (In-Sample data)

SUMMARY OUTPUT							
<i>Regression Statistics</i>							
Multiple R	0.515533						
R Square	0.265774						
Adjusted R Square	0.252377						
Standard Error	0.050105						
Observations	56						
ANOVA							
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>		
Regression	1	0.049072	0.049072	19.54682	4.78E-05		
Residual	54	0.135566	0.00251				
Total	55	0.184638					

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.600918	0.05776	10.40368	1.65E-14	0.485116	0.71672	0.485116	0.71672
X Variable 1	-0.62915	0.142303	4.42318	4.78E-05	-0.91444	-0.34385	-0.91444	-0.34385

Table 2b: OLS regression between Historical and Realized Volatility (In-Sample data)

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.165834							
R Square	0.027501							
Adjusted R Square	0.026857							
Standard Error	0.039017							
Observations	1512							
ANOVA								
	<i>Df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.065003	0.065003	42.70091	8.7E-11			
Residual	1510	2.29866	0.001522					
Total	1511	2.363663						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.289969	0.007781	37.26498	4.3E-236	0.274706	0.305233	0.274706	0.305233
X Variable 1	0.102201	0.01564	6.534594	8.7E-11	0.071523	0.132879	0.071523	0.132879

Table 2c: OLS regression between EWMA and Realized Volatility (In-Sample data)

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.405498							
R Square	0.164429							

Adjusted R Square	0.163875							
Standard Error	0.036166							
Observations	1512							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.388654	0.388654	297.147	6.32E-61			
Residual	1510	1.975009	0.001308					
Total	1511	2.363663						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.300223	0.002509	119.6515	0	0.2953	0.305143	0.2953	0.305143
X Variable 1	0.102314	0.005935	17.23795	6.32E-61	0.090672	0.113957	0.090672	0.113957

Table 2d: OLS regression between GARCH and Realized Volatility (In-Sample)

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.45168							
R Square	0.204015							
Adjusted R Square	0.203483							
Standard Error	0.031737							
Observations	1498							
ANOVA								
	<i>Df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.386237	0.386237	383.4318	3.43E-76			
Residual	1496	1.506867	0.001007					

Total	1497	1.893084						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.286727	0.002678	107.07	0	0.281474	0.29198	0.281474	0.29198
X Variable 1	0.13603	0.006947	19.581	3.43E-76	0.122403	0.149657	0.122403	0.149657

Table 2e: OLS regression between Implied and Realized Volatility(Out- of- Sample data)

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.684026							
R Square	0.467892							
Adjusted R Square	0.414681							
Standard Error	0.07432							
Observations	12							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.048569	0.0485	8.7931	0.014158			
Residual	10	0.055234	0.0055					
Total	11	0.103803						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	-0.13075	0.172306	0.7597	0.4649	-0.51423	0.252722	-0.51423	0.252722
X Variable 1	1.1757	0.396482	2.9653	0.0141	0.292283	2.059116	0.292283	2.059116

Table 2f: OLS regression between Historical and Realized Volatility(Out-of-Sample data)

SUMMARY OUTPUT								

Regression Statistics								
Multiple R	0.52064							
R Square	0.271066							
Adjusted R Square	0.198172							
Standard Error	0.086986							
Observations	12							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.028137	0.028137	3.718654	0.082655			
Residual	10	0.075665	0.0075665					
Total	11	0.103803						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.229024	0.080058	2.86072	0.016936	0.050643	0.407405	0.050643	0.407405
X Variable 1	0.859824	0.445879	1.928381	0.082655	-0.13366	1.853303	-0.13366	1.853303

Table 2g: OLS regression between EWMA and Realized Volatility(Out-of-Sample data)

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.758423							
R Square	0.575202							
Adjusted R Square	0.532722							
Standard Error	0.066404							
Observations	12							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			

Regression	1	0.059708	0.059708	13.54061	0.004248			
Residual	10	0.044095	0.00441					
Total	11	0.103803						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.261411	0.036479	7.166128	3.05E-05	0.180132	0.342691	0.180132	0.342691
X Variable 1	2017.684	548.3201	3.679756	0.004248	795.951	3239.418	795.951	3239.418

Table 2h: OLS regression between GARCH and Realized Volatility(Out-of-Sample data)

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.849507							
R Square	0.723662							
Adjusted R Square	0.696359							
Standard Error	0.052948							
Observations	13							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.079956	0.079956	28.52036	0.000237			
Residual	11	0.030838	0.002803					
Total	12	0.110795						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
			-					
Intercept	-31.7633	6.016778	5.27911	0.000261	-45.0061	-18.5204	-45.0061	-18.5204
X Variable 1	79.25606	14.84072	5.34046	0.000237	46.59186	111.9203	46.59186	111.9203

FPGA Accelerated Abandoned Object Detection

Rajesh Rohilla*, Aman Raj, Saransh Kejriwal, Rajiv Kapoor
Department of Electronics & Communication Engineering
Delhi Technological University, Delhi, India

Abstract—In this paper, a method to address the problem of detecting abandoned object(s) in a live video feed has been presented. The proposed technique utilizes a static background modeling algorithm and identifies any object lying abandoned for a given period of time. Our approach uses a FIFO queue for each pixel in a video frame as an essential element and applies a processing logic. Since the serial processing of such pixel queues on a conventional computing platform is relatively a slow process, the proposed algorithm is synthesized on an FPGA with the aim of making a custom chip that can be embedded in a development board for abandon object detection applications. Such hardware implementation also speeds up the execution of the algorithm by exploiting the parallel nature of algorithm and makes the computation real-time. The proposed technique successfully detects abandoned objects of different sizes present in the video.

Keywords—Abandoned object detection, Video surveillance, FPGA

I. INTRODUCTION

Abandoned objects are a common sight in any modern day crowded area such as a marketplace, railway station, school, and public transport. However, these abandoned objects can contain dangerous explosives planted by people with malevolent intentions to cause heavy loss of life and property. Such explosives are often disguised as commonplace objects, which are easily ignored by a passer-by and left unattended at a location, whilst keeping the bomber at a safe distance to trigger the explosive.

In an urban area, we are constantly monitored by Closed-Circuit Television (CCTV) cameras, to locate and identify any such object and trace its origins. However, the current surveillance system is largely man-powered, where a small group of security personnel is required to constantly gaze at scores of cameras, spread throughout the area to find anything suspicious. It is easy to imagine that this is not an easy task to continue for hours, and it is likely that the aforementioned personnel may miss an object here or there. Also, due to low manpower, an area cannot be monitored by an individual constantly, leaving abundant room for error.

Several tactics are deployed to address this problem. High-resolution overhead cameras are used for surveillance in significant locations such as national monuments or private establishments, to detect such unattended objects of even small size. However, these cameras are expensive, and even with their resolution, there is no guarantee that the security personnel will spot the object. Also, in some cases security officers often patrol the area under surveillance, to maintain vigilance over any such object. This is commonly seen in railway stations and areas of high-security zones. However, this tactic requires excessive manpower, especially if the area

under surveillance is huge. Also, if there happens to be a bomb concealed in an abandoned object, this raises the risk of grievous physical harm to the officers who would examine this object.

Also, the public is often encouraged to raise an alarm in case they find any such object, once they have established that it does not belong to anyone nearby. Television advertisements and loudspeaker announcements are used to motivate people to survey the area by themselves to confirm its safety. However, most people do not follow such instructions out of ignorance, or the inhibition of raising a false alarm. Even in a genuine case, a common person can find it difficult to send the message to the appropriate authority. None of the above methods uses any sort of autonomous solution. This creates a need to develop an automated system which intelligently detects such situations and inform the concerned people.

Our approach takes input from the camera a video stream and models a reference static frame and other objects which are ordinarily present in it. It then updates the background to incorporate any static object that has been abandoned or has been removed from the original background. Finally, it compares the new background with the original background and generates their difference frame which only contains the abandoned objects enclosed inside ovals, informing the security personnel who is monitoring the area in the control room. By implementing this approach in the form of a custom hardware design on FPGA, we hope to make a chip in future that can be easily embedded on a development board for such applications.

II. RELATED WORK

Many methods have been previously proposed to automate the detection of abandoned objects. Jing Chang [2] utilizes selective tracking to determine whether the owner of the so-detected abandoned luggage is in proximity by detecting “skin color information” and body contours. However, it may be difficult to determine in crowded situations where people are rapidly moving towards and away from said luggage. Also, detecting such attributes of a person is subject to lighting changes and occlusions due to an ambient crowd. Most of the proposed methods, such as the one proposed by M. Bhargava [1] search for an owner of the luggage in the midst of the crowd, which requires the system to have a lot of memory at its disposal. Background Subtraction [7] is also commonly utilized, but it is a practical only if the background so acquired updates itself with lighting changes over time, to prevent noise contours. Gaussian Mixture Model (GMM) is commonly utilized to achieve background updating [12] [3] [10]. Although GMM is an efficient method for background modeling but it is memory intensive. We present a memory

*Corresponding author, E-mail: rajesh@dce.ac.in

efficient algorithm to achieve a result that is comparable to that of GMM, also, we can explicitly control the amount of memory utilized for storage.

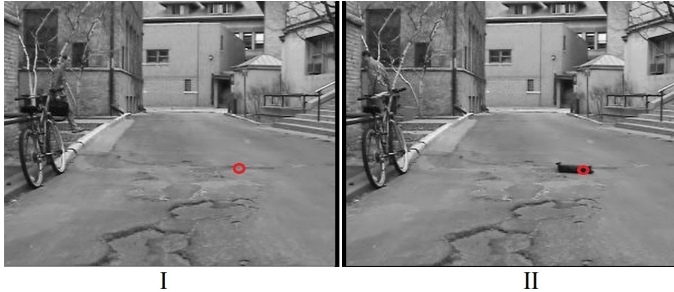


Fig. 1: A static frame- **I** without object and **II** with object

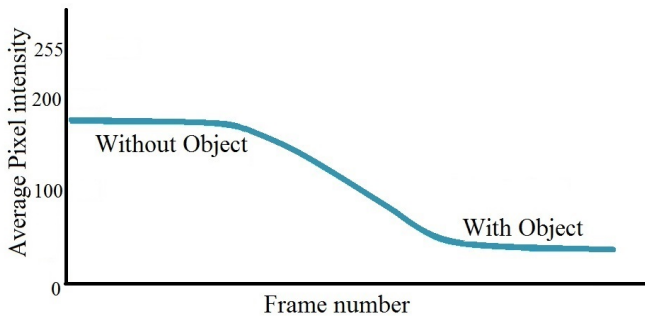


Fig. 2: Variation of pixel intensity at encircled position in Fig.1(I)-(II) with incoming frames. Averaging is performed over all the coming inputs video frames, highlighting effect of introduction of an object that is black in this case.

III. ALGORITHM

A. Obtaining Reference Static Frame

The proposed algorithm considers an assumption during the initialization period. It assumes that there are no abandoned objects in the first ' n ' frames of the video. We model an initial static background from the video feed which contains only those objects, which are at rest. Such objects form a general composition of the background, which can be trees, poles, signboards, buildings, roads and more. This is computed by forming an image using the average value of each pixel in the first ' n ' frames. This is done as per the designed algorithm flow which takes only a few seconds. This computed initial static background is stored in memory for further use.

B. Updating Current Static Frame

After initializing the static background containing solely the stationary objects that are expected to be present in the video feed, we constantly update by doing computations as proposed in III-A and models an updated static background. This is done after the desired interval to incorporate incoming of new static objects into the video. Thus, if a new object comes into the cameras field of view and it has not been moved

for a certain period of time, its presence will be visible in the updated static background. This happens even if the object is blurred due to the low resolution of surveillance video. Since the effect of rapidly moving objects while updating the static background will be very less, hence, this updating is unaffected by object occlusion due to crowds. This helps the proposed algorithm to detect such objects in the presence of spatial and temporal occlusion successfully, experimental results of the same is shown in the result section.

C. Current Frame Comparison

In case an object is left unattended in the video feed for an extended period of time, its presence will be seen in the updated static frame. This updated static frame is compared with the previous reference static frame without the object, such comparison will reveal the object as a highlighted blob. This comparison is performed by calculating a difference between both frames.

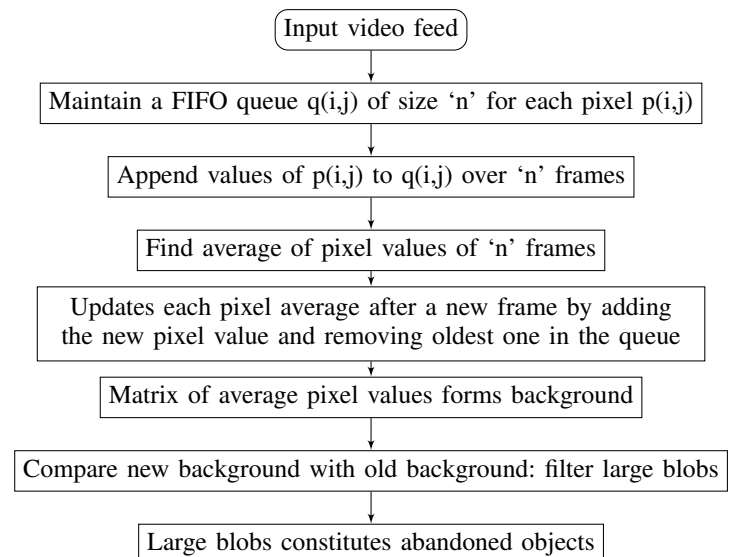
D. Blob Detection

Blob detection (also known as connected component labeling) is applied on the grayscale image produced by comparison in the previous section and attempts to find the bright areas. In order to remove the effect of intermittent movement in a video feed that results in the presence of many smaller blobs during frame comparison, we merge such blobs that are close together into one larger blob.

E. Decision making

We use a very efficient approach of comparing sizes of various blobs in the frame. If any blob is too small to be a significant object, we discard it as noise. We display the larger blobs in a separate window, making it easy for the security personnel to locate such objects and take necessary action.

IV. FLOWCHART



V. MATHEMATICAL FORMULATION

Consider the n^{th} frame of the video as I_n , such that $I_n(i, j)$ represents the value of pixel present at i^{th} row and j^{th} column of n^{th} frame. For each such pixel value a queue $Q(i, j)$ of size N , a sum of pixel values $S(i, j)$ and average of pixel values $A(i, j)$ is maintained over the incoming frames. The total number of frames N to be considered for modeling a static background can be selected by user as per the requirement. It is directly related to the measure of time after which an object will be declared abandoned. Starting from the first frame, $n = 1$:

- If $n < N$

$$Q(i, j) = I_1(i, j), I_2(i, j), \dots, I_{n-1}(i, j), I_n(i, j)$$

Hence sum of elements of this queue is given by:

$$S_n(i, j) = \sum_{k=1}^n I_k(i, j) \quad (1)$$

Average of elements of this queue is given by:

$$A_n(i, j) = 0 \quad (2)$$

As we are calculate the average value of each pixel in pixel queue only when first N frames has been observed by system and hence the background frame is blank here.

- If $n \geq N$

$$Q(i, j) = I_{n-N+1}(i, j), I_{n-N+2}(i, j), \dots, I_{n-1}(i, j), I_n(i, j)$$

Now sum of elements of this queue is given by:

$$S_n(i, j) = S_{n-1}(i, j) + I_n(i, j) - Q(i, j).front \quad (3)$$

where $I_n(i, j)$ is pixel value in the latest frame and $Q(i, j).front$ is the oldest pixel value in the queue $Q(i, j)$

Average of elements of this queue is given by:

$$A_n(i, j) = S_n(i, j)/N \quad (4)$$

We model a background now such that:

$$\left[B_n(i, j) \right]_{a \times b} = \left[A_n(i, j) \right]_{a \times b} \quad (5)$$

where $a \times b$ is the dimension of the frame and $B_n(i, j)$ is the pixel value in the static background frame. This ensures that we are always computing the average of the latest N number of frames coming in the video feed. Hence for an appropriate value of N , the proposed algorithm automatically updates background. We simultaneously maintain a separate video frame $W_n(i, j)$ that contains the abandoned objects detected by the algorithm. If there are no such objects, this frame remains empty.

VI. FPGA IMPLEMENTATION

The proposed algorithm is implemented on Xilinx Zynq-7020 all programmable system on chip (SoC) FPGA board. It has Processing System (PS) that contains Dual ARM Cortex-A9 MPCore processor with CoreSight that provides high-speed sequential logic implementations and a powerful Programmable Logic (PL) that contains Artix-7 FPGA with 85,000 logic cells and other resources that give flexibility to

designers to implement complex parallel operations on it with sequential processes. The graphical view of this SoC is shown in Fig-3. We use an RGB camera to capture live video feed which is pre-processed by the PS part before giving input to PL part where the FPGA design of our proposed algorithm for abandoned object detection lies.

The PL section is programmed using Vivado High-Level Synthesis (HLS) library provided by Xilinx. The data transfer is done using the AXI-Stream bus which is highly efficient and fast for real-time high-bandwidth data transfer. Every new frame sent to PL part is converted into a FIFO matrix having a FIFO queue for each pixel as outline in the previous section. This FIFO queue updated with each new incoming video frame. The sum and averages of pixel values in FIFO queue is calculated after a N number of latest frames is observed that can be set while running the algorithm. The design gives output the frame W_n in the form of FIFO buffer that contains the abandoned objects blobs detected by the algorithm. This is sent to PS part where it is converted into video format frame and is projected on a monitor.

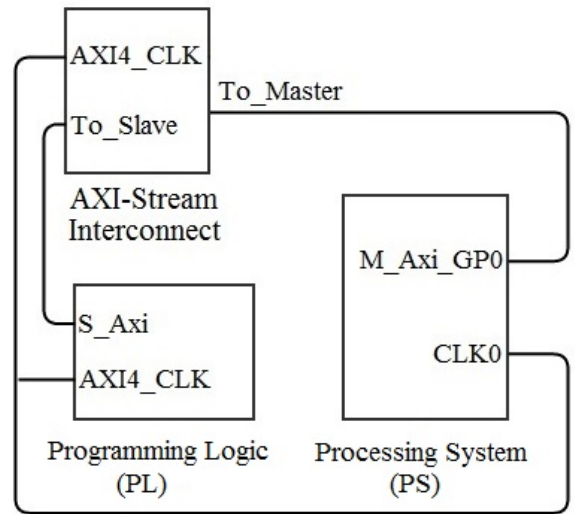


Fig. 3: Block diagram of our FPGA system

RESULTS

We tested our proposed approach on the AVSS2007 dataset. This dataset contains videos from different scenarios, such as abandoned objects and parked vehicles. Since the abandoned object scene fit our problem, so we tested our algorithm on the sequences - AB-Easy, AB-Medium, and AB-Hard that contains abandoned luggage placed on a platform. We also compared our method with current state of the art studies of [9], [4], [6], [11], [5], [8]. Precision, Recall and F-measure values for this comparison are shown in TABLE-I. The luggage left is easily detected in the case of AB-Easy as shown in Fig. 4 due to no crowd and large size of luggage, on the other hand for AB-Medium and AB-Hard this is challenging due to the small size of the abandoned object and crowd on the platform.

Noteworthy, our method localizes the abandoned objects in all three sequences.



Fig. 4: Detection results of the sequence AB-Easy of AVSS2007



Fig. 5: Detection results of the sequence AB-Medium of AVSS2007

	[9]	[4]	[6]	[11]	[5]	[8]	Ours
Precision	0.05	0.21	0.40	0.35	0.97	1.0	1.0
Recall	1.0	1.0	0.67	1.0	1.0	1.0	1.0
F-measure	0.09	0.35	0.50	0.52	0.98	1.0	1.0

TABLE I: Comparison of different methods on AVSS2007 video dataset

We also tested our algorithm on our own dataset for two types of camera positions having different environment situations:

- (i) A horizontally placed camera on the table top in minimally crowded place, for example, our lab.
- (ii) An overhead surveillance camera in very crowded place.

Since occlusion of the object of interest in object detection algorithms is a major concern, so we created various temporal and spatial occlusion situations while testing the performance of our algorithm. Also, the situations when there are objects of various dimensions and it's variable distance from the camera were created. The results as shown in TABLE-II reflects the state of the art performance on such tasks.

REFERENCES

- [1] M. Bhargava, C.-C. Chen, M. S. Ryoo, and J. K. Aggarwal. Detection of object abandonment using temporal logic. *Machine Vision and Applications*, 20(5):271–281, 2009.
- [2] J.-Y. Chang, H.-H. Liao, and L.-G. Chen. Localized detection of abandoned luggage. *EURASIP Journal on Advances in Signal Processing*, 2010(1):1–9, 2010.
- [3] Y. Dedeoğlu. Human action recognition using gaussian mixture model based background segmentation.
- [4] R. H. Evangelio, T. Senst, and T. Sikora. Detection of static objects for the task of video surveillance. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 534–540. IEEE, 2011.
- [5] Q. Fan, P. Gabbur, and S. Pankanti. Relative attributes for large-scale abandoned object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2736–2743, 2013.
- [6] Q. Fan and S. Pankanti. Modeling of temporarily static objects for robust abandoned object detection in urban surveillance. In *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*, pages 36–41. IEEE, 2011.
- [7] H. Lee, S. Hong, and E. Kim. Probabilistic background subtraction in a video-based recognition system. *KSII Transactions on Internet & Information Systems*, 5(4), 2011.
- [8] H.-H. Liao, J.-Y. Chang, and L.-G. Chen. A localized approach to abandoned luggage detection with foreground-mask sampling. In *Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on*, pages 132–139. IEEE, 2008.
- [9] F. Porikli, Y. Ivanov, and T. Haga. Robust abandoned object detection using dual foregrounds. *EURASIP Journal on Advances in Signal Processing*, 2008(1):1–11, 2007.
- [10] C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*, volume 2. IEEE, 1999.
- [11] Y. Tian, R. S. Feris, H. Liu, A. Hampapur, and M.-T. Sun. Robust detection of abandoned and removed objects in complex surveillance videos. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 41(5):565–576, 2011.
- [12] Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 2, pages 28–31. IEEE, 2004.

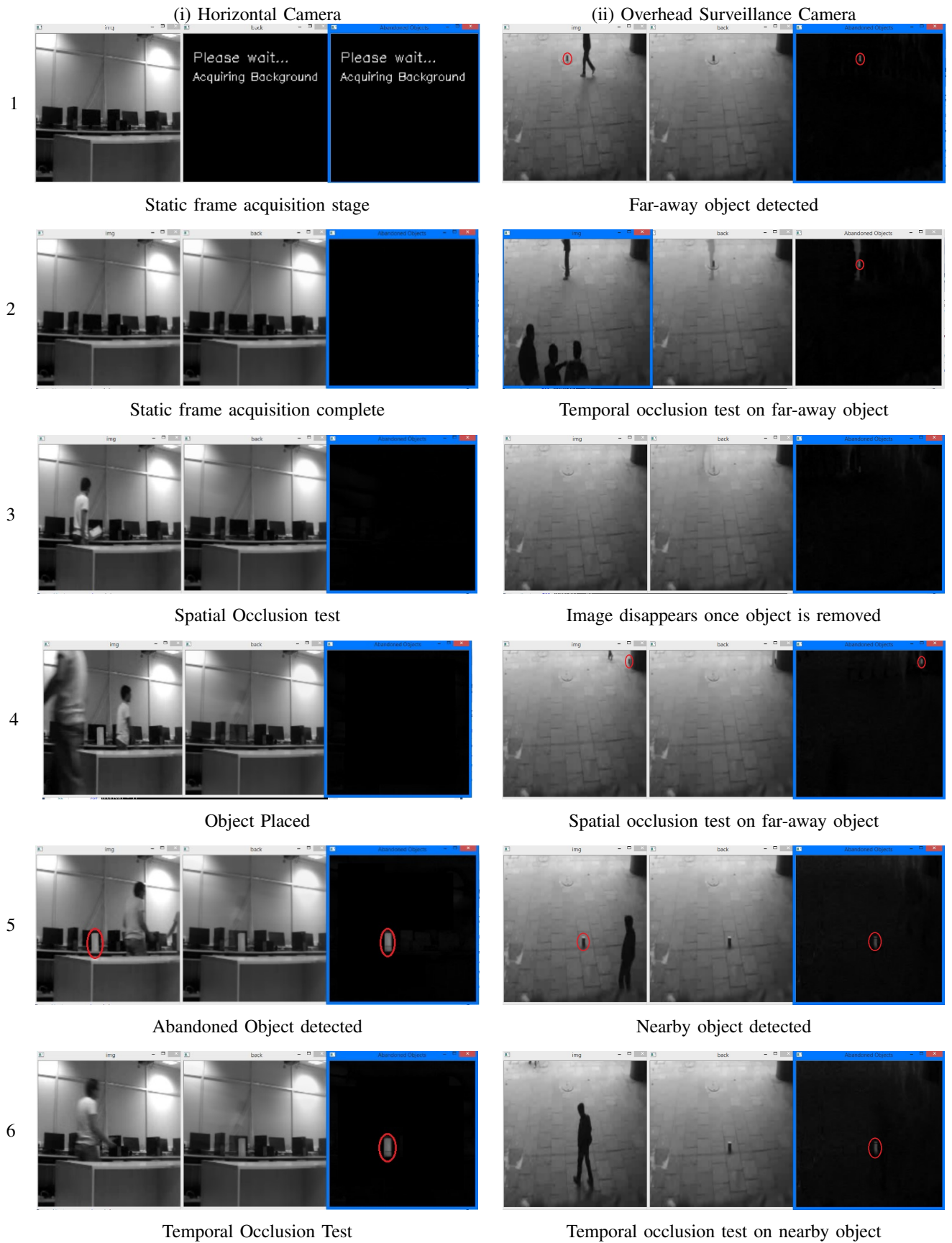


TABLE II: Results of proposed algorithm on video feed in different conditions

Human Genome - Underlying Scientific Mysteries

Shivani Kachroo¹ and Sivakumar JT Gowder^{2*}

¹Delhi Technological University, New Delhi, India

²Qassim University, College of Applied Medical Sciences, Buraidah, Kingdom of Saudi Arabia

*Corresponding author: Sivakumar JT Gowder, College of Applied Medical Sciences, Qassim University, Buraidah, Kingdom of Saudi Arabia, Tel: +966666873969; Fax: +966663802268; E-mail: sivakumargowder@yahoo.com

Rec date: Mar 02, 2016; Acc date: Mar 11, 2016; Pub date: Mar 16, 2016

Copyright: © 2016 Kachroo S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

A gene refers to the functional unit of heredity. Genes are made up of DNA (deoxyribonucleic acid). In humans, the size of genes varies from a few hundred DNA bases to more than 2 million bases. The genes in human beings (*Homo sapiens*) are referred as human genome. It is made up of 23 chromosome pairs with a total of about 3 billion DNA base pairs. The gene sequences to make a protein occupy only about 1.1% of the chromosomes. The rest is well known as junk DNA; that is useful in genetic fingerprinting.

Many scientists now believe that the junk DNA has a role in evolution. The human genome is more than 95% worthless, but 5% of the 3.2 billion nucleotides make the whole human genome [1,2]. A critical implication of the human genome is that as long as we move up in the complexity of the organisms, it surely doesn't correspond to the increase in genes, but is proportional to the number of controls and regulations on these genes. The scientists have used software and programs to go through billions of DNA letters and then locate these gene sequences. These programs advanced with time, but there is still a lot of scopes as a lot is going on in the protein manufacturing including splicing, transposable elements, assembling and post-translational reframing. As a contrast, a recent finding shows that a Japanese fish that is also known as "Fugu" has no junk DNA at all. This discovery helps us to understand more deeply the cause and consequence of development from coding and non-coding sequences [1].

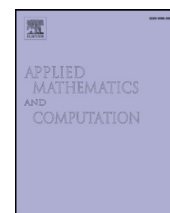
Another contribution is in explaining the variations in humans by the human genome by the common changes in the nucleotides of genetic code called the single nucleotide polymorphisms (SNPs). However, differences between any two individuals in the same population could be greater than that between two different groups. As published lately in a magazine, nobody knows the identities of the people who donated their DNA to the project, but Dr. Venter

knowingly selected DNA from five individuals, one African-American, one Asian-American, one Hispanic-American and two others and found no way of telling which was which [2-4].

It has been predicted by the genome scientists that by 2020, the information will be sufficient enough to diagnose most of the diseases before the symptoms appear. Till now 1,778 genes identified with illness and only 483 prospective candidate genes targets for therapy are known in the human body. The upcoming research will speedily increase the number of genes [2,5]. Some other questions need to be answered before the large scale application of the data, including epigenetic status of the genes, vulnerability to variations and evolution. The more information regarding these could also open up a new era of personalized medicine and therapeutics. Personalized medicine has a control over individualizing medicine both in terms of treatment and diagnosis [6]. But we should also be prepared for the more challenging ethical questions which will arise with the advancements in this knowledge bank [7].

References

1. News this Week (2001) Science 291: 1177-1207.
2. Gee H (2016) The Guardian Special report: the ethics of genetics.
3. Radford T (2001) The Guardian. Special report: the ethics of genetics.
4. Kember S (2003) Book: Cyber feminism and Artificial Life. Taylor & Francis, Routledge, London, UK.
5. Tina HS (2016) The Wired Magazine. The Science News, Science.
6. Greek R, Gowder S (2013) Animal Models in Drug Development, New Insights into Toxicity and Drug Testing, ISBN: 978-953-51-0946-4, InTech.
7. Beyond the Genome (2015) Studies of the epigenomic signatures of many healthy and diseased human tissues could provide crucial information to link genetic variation and disease. Nature 518: 187-206.



Jain–Durrmeyer operators associated with the inverse Pólya–Eggenberger distribution



Minakshi Dhamija, Naokant Deo*

Delhi Technological University, Formerly Delhi College of Engineering Department of Applied Mathematics, Bawana Road, 110042 Delhi, India

ARTICLE INFO

MSC:

41A25

41A36

Keywords:

Baskakov operators

Jain operators

Pólya–Eggenberger distribution

ABSTRACT

The present paper deals with the Jain–Durrmeyer operators based on inverse Pólya–Eggenberger distribution. First, we give the moments with the help of Vandermonde convolution formula and then study approximation properties of these operators which include uniform convergence and degree of approximation.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

The Pólya–Eggenberger (P-E) distribution was introduced by Eggenberger and Pólya [11] in the year 1923. The P-E distribution with parameters (n, A, B, S) is defined as:

$$P(X = k) = \binom{n}{k} \frac{\prod_{i=0}^{k-1} (A + iS) \prod_{i=0}^{n-k-1} (B + iS)}{\prod_{i=0}^{n-1} (A + B + iS)}, \quad k = 0, 1, \dots, n. \quad (1.1)$$

This gives the probability of getting k white balls out of n drawings from an urn contains A white and B black balls, if each time one ball is drawn at random and then replaced together with S balls of the same color.

In literatures, the inverse Pólya–Eggenberger(I-P-E) distribution is defined as:

$$P(X = k) = \binom{n+k-1}{k} \frac{\prod_{i=0}^{k-1} (A + iS) \prod_{i=0}^{n-1} (B + iS)}{\prod_{i=0}^{n+k-1} (A + B + iS)}, \quad k = 0, 1, \dots, n, \quad (1.2)$$

gives the probability that k white balls are drawn preceding the n -th black ball. The details have been given about these two distributions (1.1) and (1.2) in [21].

Based on I-P-E (1.2), Stancu [28] studied a generalization of the Baskakov operators for a real-valued function bounded on $[0, \infty)$, defined as:

$$V_n^{(\alpha)}(f; x) = \sum_{k=0}^{\infty} v_{n,k}^{(\alpha)}(x) f\left(\frac{k}{n}\right) = \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{1^{[n, -\alpha]} x^{[k, -\alpha]}}{(1+x)^{[n+k, -\alpha]}} f\left(\frac{k}{n}\right) \quad (1.3)$$

where α is a non-negative parameter which may depend only on the natural number n and $t^{[n, h]} = t(t-h)(t-2h) \dots (t - (n-1)h)$, $t^{[0, h]} = 1$ represents the factorial power of t with increment h . In the case when $\alpha = 0$, operators (1.3) reduce

* Corresponding author. Tel.: ++919868981676.

E-mail addresses: minakshidhamija11@gmail.com (M. Dhamija), dr_naokant_deo@yahoo.com, naokant.deo@gmail.com (N. Deo).

to the following classical Baskakov operators [6].

$$V_n(f, x) = \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{x^k}{(1+x)^{n+k}} f\left(\frac{k}{n}\right). \quad (1.4)$$

In the year 1972, Jain [20] introduced a new class of linear operators (known as Jain operators) as:

$$J_n^{(\beta)}(f; x) = \sum_{k=0}^{\infty} b_{n,k}^{(\beta)}(x) f\left(\frac{k}{n}\right), \quad x \geq 0, \quad (1.5)$$

where $0 \leq \beta < 1$ and

$$b_{n,k}^{(\beta)}(x) = \frac{nx(nx+k\beta)^{k-1} e^{-(nx+k\beta)}}{k!}.$$

These operators reduce to Szász–Mirakyan operators for special case $\beta = 0$. Several mathematicians studied these Jain operators and their integral variant (See [3,4,12,15,25,29]). Very recently, Gupta and Greubel [16] introduced the Durrmeyer variant of these operators (1.5) in a very different manner as:

$$\begin{aligned} D_n^{(\beta)}(f; x) &= \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\beta)}(t) dt \right)^{-1} b_{n,k}^{(\beta)}(x) \int_0^{\infty} b_{n,k}^{(\beta)}(t) f(t) dt \\ &= \sum_{k=0}^{\infty} \frac{\langle b_{n,k}^{(\beta)}(t), f(t) \rangle}{\langle b_{n,k}^{(\beta)}(t), 1 \rangle} b_{n,k}^{(\beta)}(x), \end{aligned} \quad (1.6)$$

where $\langle f, g \rangle = \int_0^{\infty} f(t)g(t)dt$.

These operators also reduce to Szász–Mirakyan Durrmeyer operators for $\beta = 0$. Moreover, the estimation of moments for (1.6) has been calculated by using Stirling numbers of first kind and confluent hypergeometric function. Durrmeyer form of different operators have been studied by many mathematicians and researchers (See [7,8,22]). Motivated from [16], now we propose Hybrid type operators as:

$$\begin{aligned} L_n^{(\alpha)}(f; x) &= \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) f(t) dt \\ &= \sum_{k=0}^{\infty} \frac{\langle b_{n,k}^{(\alpha)}(t), f(t) \rangle}{\langle b_{n,k}^{(\alpha)}(t), 1 \rangle} v_{n,k}^{(\alpha)}(x). \end{aligned} \quad (1.7)$$

For $\alpha = 0$, these operators include the hybrid operators (Baskakov–Szász–Durrmeyer operators) presented in [1,18] etc.

These days several mathematicians have worked on P-E based positive linear operators [2,13,17,24,26,27] and few mathematicians have discussed I-P-E based positive linear operators [14,19,28]. Very recently Deo et al. [9] defined I-P-E based Baskakov–Kantorovich operators and studied its properties, asymptotic formula as well as degree of approximation.

Now in the present paper, we obtain the moments of I-P-E distribution based operators (1.3) by using Vandermonde convolution formula, however these moments have already been calculated by Stancu in [28] by using hypergeometric series. Then we consider Jain–Durrmeyer operators associated with the inverse Pólya–Eggenberger distribution (1.7) and study uniform convergence and degree of approximation for these operators.

2. Auxiliary results

Let \mathbb{N} be the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $k \in \mathbb{N}_0$, let $e_k(x) = x^k$ be the test functions. The computation of the images of test functions by Stancu–Baskakov operators (1.3) was done in [28]. Now we find those images with the help of Vandermonde convolution formula and following relation:

$$t^{[i+j,h]} = t^{[i,h]}(t - ih)^{[j,h]} \quad (2.1)$$

for $i, j \in \mathbb{N}$ and $h \neq 0$.

Lemma 2.1. For the Stancu–Baskakov operators (1.3), the following results hold:

$$V_n^{(\alpha)}(e_0; x) = 1, \quad V_n^{(\alpha)}(e_1; x) = \frac{x}{1-\alpha},$$

and

$$V_n^{(\alpha)}(e_2; x) = \frac{1}{(1-\alpha)(1-2\alpha)} \left[x^2 + \frac{x(x+1)}{n} + \alpha \left(1 - \frac{1}{n} \right) x \right].$$

Proof. To achieve the results of this lemma we use the relation (2.1), we have

$$\begin{aligned} V_n^{(\alpha)}(e_0; x) &= \sum_{k=0}^{\infty} v_{n,k}^{(\alpha)}(x) = \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{x^{[k, -\alpha]} 1^{[n, -\alpha]}}{(1+x)^{[n+k, -\alpha]}} \\ &= \frac{1^{[n, -\alpha]}}{1^{[n, -\alpha]}} = 1. \end{aligned}$$

From (1.3) we have

$$\begin{aligned} V_n^{(\alpha)}(e_1; x) &= \sum_{k=0}^{\infty} v_{n,k}^{(\alpha)}(x) \frac{k}{n} = \sum_{k=1}^{\infty} \binom{n+k-1}{n} \frac{x^{[k, -\alpha]} 1^{[n, -\alpha]}}{(1+x)^{[n+k, -\alpha]}} \\ &= \sum_{k=1}^{\infty} \binom{n+k-1}{n} \frac{x^{[1, -\alpha]} (x+\alpha)^{[k-1, -\alpha]} 1^{[1, -\alpha]} (1+\alpha)^{[n-1, -\alpha]}}{(1+x)^{[n+k, -\alpha]}} \\ &= x \frac{(1+\alpha)^{[n-1, -\alpha]}}{(1-\alpha)^{[n+1, -\alpha]}} = \frac{x}{1-\alpha}, \end{aligned}$$

and

$$\begin{aligned} V_n^{(\alpha)}(e_2; x) &= \sum_{k=0}^{\infty} v_{n,k}^{(\alpha)}(x) \frac{k^2}{n^2} = \sum_{k=0}^{\infty} \frac{k(k-1) + k}{n^2} v_{n,k}^{(\alpha)}(x) = \frac{1}{n} \left[V_n^{(\alpha)}(e_1; x) \right. \\ &\quad \left. + (n+1) \sum_{k=2}^{\infty} \binom{n+k-1}{n+1} \frac{x^{[2, -\alpha]} (x+2\alpha)^{[k-2, -\alpha]} 1^{[2, -\alpha]} (1+2\alpha)^{[n-2, -\alpha]}}{(1+x)^{[n+k, -\alpha]}} \right] \\ &= \frac{x}{n(1-\alpha)} + \left(\frac{n+1}{n} \right) \frac{x(x+\alpha)(1+\alpha)(1+2\alpha)^{[n-2, -\alpha]}}{(1-2\alpha)^{[n+2, -\alpha]}} \\ &= \frac{1}{(1-\alpha)(1-2\alpha)} \left[x^2 + \frac{x(x+1)}{n} + \alpha \left(1 - \frac{1}{n} \right) x \right]. \end{aligned}$$

□

Lemma 2.2 [16]. For $0 \leq \alpha < 1$, we have

$$\frac{\langle b_{n,k}^{(\alpha)}(t), t^r \rangle}{\langle b_{n,k}^{(\alpha)}(t), 1 \rangle} = P_r(k; \alpha),$$

where $P_r(k; \alpha)$ is a polynomial of order r in variable k and $\langle f, g \rangle = \int_0^\infty f(t)g(t)dt$.

In particular

$$\begin{aligned} P_0(k; \alpha) &= 1, \\ P_1(k; \alpha) &= \frac{1}{n} \left[(1-\alpha)k + \frac{1}{1-\alpha} \right], \\ P_2(k; \alpha) &= \frac{1}{n^2} \left[(1-\alpha)^2 k^2 + 3k + \frac{2!}{1-\alpha} \right], \\ P_3(k; \alpha) &= \frac{1}{n^3} \left[(1-\alpha)^3 k^3 + 6(1-\alpha)k^2 + \frac{(11-8\alpha)k}{1-\alpha} + \frac{3!}{1-\alpha} \right], \\ P_4(k; \alpha) &= \frac{1}{n^4} \left[(1-\alpha)^4 k^4 + 10(1-\alpha)^2 k^3 + 5(7-4\alpha)k^2 + \frac{10(5-3\alpha)k}{1-\alpha} + \frac{4!}{1-\alpha} \right], \\ P_5(k; \alpha) &= \frac{1}{n^5} \left[(1-\alpha)^5 k^5 + 15(1-\alpha)^3 k^4 + 5(1-\alpha)(17-8\alpha)k^3 + \frac{15(15-20\alpha+6\alpha^2)k^2}{1-\alpha} \right. \\ &\quad \left. + \frac{(274-144\alpha)k}{1-\alpha} + \frac{5!}{1-\alpha} \right]. \end{aligned}$$

Lemma 2.3. For the operators $L_n^{(\alpha)}$ given by (1.7), the moments up to second order are given by

$$L_n^{(\alpha)}(e_0; x) = 1, \quad L_n^{(\alpha)}(e_1; x) = x + \frac{1}{n(1-\alpha)},$$

and

$$L_n^{(\alpha)}(e_2; x) = \frac{1-\alpha}{1-2\alpha} \left[x^2 + \frac{x(x+1)}{n} + \alpha \left(1 - \frac{1}{n} \right) x \right] + \frac{3x}{n(1-\alpha)} + \frac{2}{n^2(1-\alpha)}.$$

Proof. Starting from r -th order moment and making use of Lemma 2.2, we get

$$L_n^{(\alpha)}(e_r; x) = \sum_{k=0}^{\infty} \frac{\langle b_{n,k}^{(\alpha)}(t), t^r \rangle}{\langle b_{n,k}^{(\alpha)}(t), 1 \rangle} v_{n,k}^{(\alpha)}(x) = \sum_{k=0}^{\infty} P_r(k, \alpha) v_{n,k}^{(\alpha)}(x). \quad (2.2)$$

With the help of Lemma 2.1 and statement (2.2) we shall find the first three moments

$$L_n^{(\alpha)}(e_0; x) = \sum_{k=0}^{\infty} P_0(k, \alpha) v_{n,k}^{(\alpha)}(x) = 1,$$

$$\begin{aligned} L_n^{(\alpha)}(e_1; x) &= \sum_{k=0}^{\infty} P_1(k, \alpha) v_{n,k}^{(\alpha)}(x) = \sum_{k=0}^{\infty} \frac{1}{n} \left[(1-\alpha)k + \frac{1}{1-\alpha} \right] v_{n,k}^{(\alpha)}(x) \\ &= (1-\alpha) V_n^{(\alpha)}(e_1; x) + \frac{1}{n(1-\alpha)} V_n^{(\alpha)}(e_0; x) = x + \frac{1}{n(1-\alpha)}, \end{aligned}$$

and

$$\begin{aligned} L_n^{(\alpha)}(e_2; x) &= \sum_{k=0}^{\infty} P_2(k, \alpha) v_{n,k}^{(\alpha)}(x) \\ &= \sum_{k=0}^{\infty} \frac{1}{n^2} \left[(1-\alpha)^2 k^2 + 3k + \frac{2!}{1-\alpha} \right] v_{n,k}^{(\alpha)}(x) \\ &= (1-\alpha)^2 V_n^{(\alpha)}(e_2; x) + \frac{3}{n} V_n^{(\alpha)}(e_1; x) + \frac{2}{n^2(1-\alpha)} V_n^{(\alpha)}(e_0; x) \\ &= \frac{1-\alpha}{1-2\alpha} \left[x^2 + \frac{x(x+1)}{n} + \alpha \left(1 - \frac{1}{n} \right) x \right] + \frac{3x}{n(1-\alpha)} + \frac{2}{n^2(1-\alpha)}. \end{aligned}$$

□

Remark 2.1. Taking Lemma 2.3 into the account, we get the following central moments:

$$L_n^{(\alpha)}(e_1 - x; x) = \frac{1}{n(1-\alpha)}, \quad (2.3)$$

and

$$\begin{aligned} L_n^{(\alpha)}((e_1 - x)^2; x) &= \frac{1-\alpha}{1-2\alpha} \left[\frac{\alpha}{1-\alpha} x^2 + \frac{x(x+1)}{n} + \alpha \left(1 - \frac{1}{n} \right) x \right] \\ &\quad + \frac{x}{n(1-\alpha)} + \frac{2}{n^2(1-\alpha)}. \end{aligned} \quad (2.4)$$

Lemma 2.4. For the operators $L_n^{(\alpha)}$, we have

$$|L_n^{(\alpha)}(f; x)| \leq \|f\|,$$

where $f \in C[0, \infty)$ and $\|f\| = \sup_{x \in [0, \infty)} |f(x)|$.

Proof. The definition of Stancu–Jain operators (1.7) and Lemma 2.3 yield

$$\begin{aligned} |L_n^{(\alpha)}(f; x)| &\leq \left| \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) f(t) dt \right| \\ &\leq \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) |f(t)| dt \leq \|f\|. \end{aligned}$$

□

3. Direct results

Using the well-known Bohman–Korovkin–Popoviciu theorem (see [23]) we get the uniform convergence of the Stancu–Jain operators (1.7).

Theorem 3.1. Let $f \in C[0, \infty) \cap E$ and $\alpha(n)$ be such that $\alpha \rightarrow 0$ as $n \rightarrow \infty$, then we have

$$\lim_{n \rightarrow \infty} L_n^{(\alpha)}(f; x) = f(x)$$

uniformly on each compact subset of $[0, \infty)$, where $C[0, \infty)$ is the space of all real-valued continuous functions on $[0, \infty)$ and

$$E := \left\{ f : x \in [0, \infty), \frac{f(x)}{1+x^2} \text{ is convergent as } x \rightarrow \infty \right\}$$

Proof. Taking Lemma 2.3 into the account and the fact that $\alpha \rightarrow 0$ as $n \rightarrow \infty$, it is clear that

$$\lim_{n \rightarrow \infty} L_n^{(\alpha)}(e_i; x) = x^i, \quad i = 0, 1, 2$$

uniformly on each compact subset of $[0, \infty)$. Hence, applying the well-known Korovkin-type theorem [5] regarding the convergence of a sequence of positive linear operators, we get the desired result. \square

Modulus of continuity is the main tool to measure the degree of approximation of linear positive operators towards the identity operators.

Definition 3.1. Let $f \in C_B[0, \infty)$ be given and $\delta > 0$. The modulus of continuity of the function f is defined by

$$\omega(f, \delta) := \sup\{|f(x) - f(y)| : x, y \in [0, \infty), |x - y| \leq \delta\}, \quad (3.1)$$

where $C_B[0, \infty)$ is the space of all real-valued continuous and bounded functions on $[0, \infty)$.

Definition 3.2. For $f \in C[0, \infty)$ and $\delta > 0$

$$\omega_2(f, \delta) := \sup\{|f(x+h) - 2f(x) + f(x-h)| : x, x \pm h \in [0, \infty), 0 \leq h \leq \delta\} \quad (3.2)$$

is the modulus of smoothness of second order.

Definition 3.3. Let f be from the space $C_B[0, \infty)$ endowed with the norm

$$\|f\| = \sup_{x \in [0, \infty)} |f(x)| \text{ and let us consider Peetre's } K\text{-functional}$$

$$K_2(f, \delta) = \inf_{g \in W_\infty^2} \{\|f - g\| + \delta \|g''\|\}, \quad (3.3)$$

where $\delta > 0$ and $W_\infty^2 = \{g \in C_B[0, \infty) : g', g'' \in C_B[0, \infty)\}$. Also, from ([10], p. 177, Theorem 2.4), there exists an absolute constant $A > 0$ such that

$$K_2(f, \delta) \leq A \omega_2(f, \sqrt{\delta}), \quad (3.4)$$

Theorem 3.2. If $f \in C_B[0, \infty)$, then for $x \in [0, \infty)$, it follows:

$$|L_n^{(\alpha)}(f; x) - f(x)| \leq 2 \cdot \omega(f, \delta), \quad \text{with } \delta = (L_n^{(\alpha)}((e_1 - x)^2; x))^{\frac{1}{2}}.$$

Proof. From (1.7), Lemma 2.3 and property of modulus of continuity, we have

$$\begin{aligned} & |L_n^{(\alpha)}(f; x) - f(x)| \\ & \leq \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) |f(t) - f(x)| dt \\ & \leq \left[1 + \frac{1}{\delta} \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) |t - x| dt \right] \omega(f, \delta). \end{aligned}$$

Applying Cauchy–Schwarz inequality for integration, we get

$$\begin{aligned} |L_n^{(\alpha)}(f; x) - f(x)| & \leq \left[1 + \frac{1}{\delta} \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{1/2} \right. \\ & \quad \times \left. \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) (t - x)^2 dt \right)^{1/2} \right] \omega(f, \delta). \end{aligned}$$

Again using Cauchy–Schwarz inequality for sum, we have

$$\begin{aligned} & \left| L_n^{(\alpha)}(f; x) - f(x) \right| \\ & \leq \left[1 + \frac{1}{\delta} \left\{ \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right\}^{1/2} \right. \\ & \quad \times \left. \left\{ \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) (t-x)^2 dt \right\}^{1/2} \right] \omega(f, \delta) \\ & = \left[1 + \frac{1}{\delta} \left\{ L_n^{(\alpha)}(e_0; x) \right\}^{1/2} \left\{ L_n^{(\alpha)}((e_1 - x)^2; x) \right\}^{1/2} \right] \omega(f, \delta). \end{aligned}$$

If we choose $\delta = \left(L_n^{(\alpha)}((e_1 - x)^2; x) \right)^{1/2}$, then we obtain

$$\left| L_n^{(\alpha)}(f; x) - f(x) \right| \leq 2\omega(f, \delta).$$

□

Theorem 3.3. Let be $f \in C[0, \infty)$, then for any $x \in [0, \infty)$ gives

$$\left| L_n^{(\alpha)}(f; x) - f(x) \right| \leq A\omega_2\left(f, \frac{1}{2}\delta_n(x, \alpha)\right) + \omega(f, \delta_\omega),$$

where A is an absolute constant, $\delta_\omega = L_n^{(\alpha)}(e_1 - x; x)$ and

$$\delta_n(x, \alpha) = \left(L_n^{(\alpha)}((e_1 - x)^2; x) + \left(L_n^{(\alpha)}(e_1 - x; x) \right)^2 \right)^{\frac{1}{2}}.$$

Proof. For $x \in [0, \infty)$ consider the following operators:

$$\tilde{L}_n^{(\alpha)}(f; x) = L_n^{(\alpha)}(f; x) - f\left(x + \frac{1}{n(1-\alpha)}\right) + f(x). \quad (3.5)$$

From Lemma 2.3, we get that

$$\tilde{L}_n^{(\alpha)}(e_0; x) = 1$$

and

$$\tilde{L}_n^{(\alpha)}(e_1; x) = x$$

i.e. the operators $\tilde{L}_n^{(\alpha)}$ preserve constants and linear functions. Therefore

$$\tilde{L}_n^{(\alpha)}(e_1 - x; x) = 0. \quad (3.6)$$

Let $g \in W_\infty^2$ and $t \in [0, \infty)$. By Taylor's expansion, we have

$$g(t) = g(x) + (t-x)g'(x) + \int_x^t (t-u)g''(u)du.$$

Applying $\tilde{L}_n^{(\alpha)}$ to above expression and using (3.6), we get

$$\begin{aligned} \tilde{L}_n^{(\alpha)}(g; x) - g(x) &= g'(x) \cdot \tilde{L}_n^{(\alpha)}(e_1 - x; x) + \tilde{L}_n^{(\alpha)}\left(\int_x^t (t-u)g''(u)du; x\right) \\ &= L_n^{(\alpha)}\left(\int_x^t (t-u)g''(u)du; x\right) - \int_x^{x+\frac{1}{n(1-\alpha)}} \left(x + \frac{1}{n(1-\alpha)} - u\right)g''(u)du. \end{aligned}$$

Also, we consider following:

$$\left| \int_x^t (t-u)g''(u)du \right| \leq (t-x)^2 \|g''\|,$$

then

$$\left| \tilde{L}_n^{(\alpha)}(g; x) - g(x) \right| \leq \left(L_n^{(\alpha)}((e_1 - x)^2; x) + \left(L_n^{(\alpha)}(e_1 - x; x) \right)^2 \right) \|g''\|.$$

Again using the expression (3.5) for the operators $\tilde{L}_n^{(\alpha)}$ and Lemma 2.4, it follows:

$$\begin{aligned}
|L_n^{(\alpha)}(f; x) - f(x)| &\leq |\tilde{L}_n^{(\alpha)}(f - g; x)| + |\tilde{L}_n^{(\alpha)}(g; x) - g(x)| \\
&\quad + |g(x) - f(x)| + \left| f\left(x + \frac{1}{n(1-\alpha)}\right) - f(x) \right| \\
&\leq 4\|f - g\| + \delta_n^2(x, \alpha) \|g''\| + \omega(f, \delta_\omega),
\end{aligned}$$

with $\delta_n^2(x, \alpha) = L_n^{(\alpha)}((e_1 - x)^2; x) + (L_n^{(\alpha)}(e_1 - x; x))^2$ and $\delta_\omega = L_n^{(\alpha)}(e_1 - x; x)$.

Now, taking infimum on the right-hand side over all $g \in W_\infty^2$ and using the relation (3.4), we get

$$\begin{aligned}
|L_n^{(\alpha)}(f; x) - f(x)| &\leq 4K_2 \left(f, \frac{\delta_n^2(x, \alpha)}{4} \right) + \omega(f, \delta_\omega) \\
&\leq A\omega_2\left(f, \frac{1}{2}\delta_n(x, \alpha)\right) + \omega(f, \delta_\omega).
\end{aligned}$$

□

Now we present ordinary approximation in terms of Lipschitz constant defined by

$$Lip_M(\beta) = \left\{ f \in C_B(0, \infty) : |f(t) - f(x)| \leq M \frac{|t - x|^\beta}{(t + x)^{\beta/2}} \right\}, \quad (3.7)$$

where M is a positive constant and $0 < \beta \leq 1$.

Theorem 3.4. Let be $f \in C_B[0, \infty)$, then for any $x \in [0, \infty)$, the following inequality holds:

$$|L_n^{(\alpha)}(f; x) - f(x)| \leq M \left(\frac{\lambda_n(x, \alpha)}{x} \right)^{\beta/2},$$

where

$$\lambda_n(x, \alpha) = L_n^{(\alpha)}((t - x)^2; x).$$

Proof. By using the fact that $f \in Lip_M(\beta)$, $0 < \beta \leq 1$ it follows:

$$\begin{aligned}
|L_n^{(\alpha)}(f; x) - f(x)| &\leq \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) |f(t) - f(x)| dt \\
&\leq M \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) \frac{|t - x|^\beta}{(t + x)^{\beta/2}} dt.
\end{aligned}$$

Applying Hölder inequality for integration and sum respectively by taking $p = \frac{2}{\beta}$ and $q = \frac{2}{2-\beta}$,

$$\begin{aligned}
|L_n^{(\alpha)}(f; x) - f(x)| &\leq M \left\{ \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right) \right\}^{(2-\beta)/2} \\
&\quad \times \left\{ \sum_{k=0}^{\infty} \left(\int_0^{\infty} b_{n,k}^{(\alpha)}(t) dt \right)^{-1} v_{n,k}^{(\alpha)}(x) \int_0^{\infty} b_{n,k}^{(\alpha)}(t) \frac{(t - x)^2}{t + x} dt \right\}^{\beta/2}.
\end{aligned}$$

Using the fact that $\frac{1}{t+x} \leq \frac{1}{x}$, we finally get

$$|L_n^{(\alpha)}(f; x) - f(x)| \leq M \left(\frac{\lambda_n(x, \alpha)}{x} \right)^{\beta/2}.$$

where

$$\lambda_n(x, \alpha) = L_n^{(\alpha)}((t - x)^2; x).$$

□

Acknowledgment

The authors are extremely thankful to the referee for his valuable comments and suggestions, leading to a better presentation of the paper.

References

- [1] U. Abel, V. Gupta, M. Ivan, The complete asymptotic expansion for Baskakov–Szász Durrmeyer operators, *Ann. Tiberiu Popoviciu Itinerant Semin. Funct. Equ. Approx. Convexity* 1 (2003) 3–15.
- [2] J.A. Adell, J. De la Cal, On a Bernstein type operator associated with the inverse Polya–Eggenberger distribution, *Rend. Circolo Matem. Palermo: Ser. II* 33 (1993) 143–154.
- [3] O. Agratini, Approximation properties of a class of linear operators, *Math. Methods Appl. Sci.* 36 (2013) 2353–2358.
- [4] O. Agratini, On an approximation process of integral type, *Appl. Math. Comput.* 236 (1) (2014) 195–201.
- [5] F. Altomare, M. Campiti, Korovkin-type approximation theory and its application, *de Gruyter studies in Mathematics*, 17, Walter de Gruyter & Co., Berlin, 1994.
- [6] V.A. Baskakov, A sequence of linear positive operators in the space of continuous functions, *Dokl. Acad. Nauk. SSSR* 113 (1957) 249–251.
- [7] N. Deo, N. Bhardwaj, Some approximation results for Durrmeyer operators, *Appl. Math. Comput.* 217 (2011) 5531–5536.
- [8] N. Deo, Faster rate of convergence on Srivastava–Gupta operators, *Appl. Math. Comput.* 218 (21) (2012) 10486–10491.
- [9] N. Deo, M. Dhamija, D. Miclăuş, Stancu–Kantorovich operators based on inverse Pólya–Eggenberger distribution, *Appl. Math. Comput.* 273 (1) (2016) 281–289.
- [10] R.A. DeVore, G.G. Lorentz, *Constructive Approximation*, Springer, Berlin, 1993.
- [11] F. Eggenberger, G. Pólya, Über die statistik verkehrter vorgänge, *Z. Angew. Math. Mech.* 1 (1923) 279–289.
- [12] A. Farcaş, An asymptotic formula for Jain's operators, *Stud. Univ. Babes-Bolyai: Math.* 57 (4) (2012) 511–517.
- [13] Z. Finta, Quantitative estimates for some linear and positive operators, *Stud. Univ. Babes-Bolyai: Math.* XLVII (3) (2002) 71–83.
- [14] Z. Finta, On approximation properties of Stancu operators, *Stud. Univ. Babes-Bolyai, Math.* XLVII (4) (2002) 47–55.
- [15] V. Gupta, R.P. Agarwal, *Rate of Convergence in Simultaneous Approximation*, Springer, Switzerland, 2014.
- [16] V. Gupta, G.C. Greubel, Moment estimations of new Szász–Mirakyan–Durrmeyer operators, *Appl. Math. Comput.* 271 (15) (2015) 540–547.
- [17] V. Gupta, T.M. Rassias, Lupaş–Durrmeyer operators based on Polya distribution, *Banach J. Math. Anal.* 8 (2) (2014) 145–155.
- [18] V. Gupta, G.S. Srivastava, Simultaneous approximation by Baskakov–Szász type operators, *Bull. Math. Soc. Sci. Math. Roum.* 37 (85) (1993) 73–85.
- [19] E. Iulia, S. Laze, On the Stancu type linear positive operators of approximation constructed by using the beta and the gamma functions, *Stud. Univ. Babes-Bolyai, Math.* LIV (2) (2009) 117–126.
- [20] G.C. Jain, Approximation of functions by a new class of linear operators, *J. Aust. Math. Soc.* 13 (3) (1972) 271–276.
- [21] N.L. Johnson, S. Kotz, *Discrete Distributions*, Houghton-Mifflin, Boston, 1969.
- [22] H.S. Jung, N. Deo, M. Dhamija, Pointwise approximation by Bernstein type operators in mobile interval, *Appl. Math. Comput.* 214 (1) (2014) 683694.
- [23] P.P. Korovkin, Convergence of linear positive operators in the spaces of continuous functions (russian), *Doklady Akad. Nauk. SSSR* 90 (1953) 961–964.
- [24] D. Miclăuş, On the monotonicity property for the sequence of Stancu type polynomials, *Analele Ştiinţifice Ale Universităţii “Al.I.Cuza” Din Iaşi (S.N.)*, 2014.
- [25] P.K. Patel, V.N. Mishra, Jain–Baskakov operators and its different generalization, *Acta Math. Vietnam.* 10 (2014) 1–19.
- [26] Q. Razi, Approximation of a function by Kantorovich type operators, *Mat. Vesnik.* 41 (1989) 183–192.
- [27] D.D. Stancu, Approximation of functions by a new class of linear polynomial operators, *Rev. Roum. Math. Pures Appl.* 13 (1968) 1173–1194.
- [28] D.D. Stancu, Two classes of positive linear operators, *Anal. Univ. Timişoara, Ser. Ştin. Matem.* 8 (1970) 213–220.
- [29] S. Tarabie, On Jain–Beta linear operators, *Appl. Math. Inf. Sci.* 6 (2) (2012) 213–216.

MRI BRAIN IMAGE CLASSIFICATION USING POLYNOMIAL KERNEL PRINCIPAL COMPONENT ANALYSIS WITH NEURAL NETWORK

Mahipal Singh Choudhry¹, Shikha Misra², Rajiv Kapoor³

^{1,3}ECE Department, Delhi Technological University, (India)

²ECE Department, Sachdeva Institute of Technology, (India)

ABSTRACT

Magnetic Resonance (MR) Imaging has come up as widely accepted and revolutionary innovation in field of medical science and brain imaging especially. A new method is proposed here for MRI brain image classification using Polynomial Kernel Principle Component Analysis (KPCA) with Neural Network. In this paper, we are having various stages namely pre-processing, feature extraction, feature reduction and classification of MRI brain images. Here for improving the MRI image quality, imadjust function is used as pre-processing stage of MRI image. In second stage, features are reduced by Polynomial Kernel Principle Component Analysis. In last stage, MRI images are classified as normal or abnormal image by Artificial Neural Network. Different feature reduction methods like PCA, LDA, SVD, Gaussian KPCA and Polynomial KPCA with P (power of kernel) = 2, 3, 4 and 5 are used. The results show that classification rate of 99.8 % is achieved for $p = 4$ of KPCA.

Keywords: ANN, DWT, FCM, k-NN, KPCA, LDA, MRI, SOM, SVD, SVM.

I. INTRODUCTION

MRI is a powerful medical imaging technique to provide detail and reliable information about human brain. MRI is preferred over other imaging methods for human brain imaging because it does not involve any ionizing radiation and it is used in non-invasively form [6]. In proposed method, firstly preprocessing of the MR images are employed using imadjust MATLAB function and then feature extraction is done by discrete Wavelet Transform (DWT). In Wavelet Transform, different frequencies are examined with different resolutions [9] and DWT coefficients are used as feature vectors of image. The wavelet coefficients are extracted from MR images by DWT in form of localized frequency information about the functions of image that is used for classification [4].

There are different types of Wavelet Transform functions with their strength and limitations [1]. Haar wavelet functions are used in proposed method due to suitability for brain MR images. To increase the processing speed, feature reduction method is used by transforming high dimensional input feature space into a lower-dimensional feature space. There are various techniques like Principal Component Analysis [10], Kernel PCA [18], Linear Discriminant Analysis [11] and Singular value Decomposition [14] are used for feature reduction. There are

different types of Kernel PCA. In this method, Kernel PCA with variable power is specially used as feature reduction method because it does not involve any nonlinear optimization [12] [6].

MR image classification techniques have been broadly classified in two classes of supervised techniques and unsupervised techniques. Artificial Neural Networks (ANN) [2] [13], Support Vector Machine (SVM) [6], k-Nearest Neighbor (k-NN) [13] and feed forward back propagation [17] are important supervised techniques and Self-Organization Map (SOM) [6] and Fuzzy C-Means (FCM) [8] are typical unsupervised techniques. Out of these classifiers, ANN, k-NN, SVM and FCM are commonly used to obtain higher classification rate for detection of abnormality of brain. Neural networks have many advantages like requirement of less number of training sets, having large numbers of training algorithms and capability of detecting nonlinear relation between independent and independent parameters [19] etc. Due to these advantages, Neural Network is used in proposed method for MRI brain image classification.

This paper is structured in following parts. The part 2 presents the methodology of paper. The part 3 presents KPCA as feature reduction method. For classification of MRI brain image, artificial neural network is used in part 4. The experimental results are shown in part 5 and in last part conclusion is described.

II. METHODOLOGY

The suggested methodology has different stages like MRI data acquisition, pre-processing of MRI images, feature extraction from images, dimension reduction and MRI images classification as normal or abnormal. The complete methodology is shown in Fig.1.

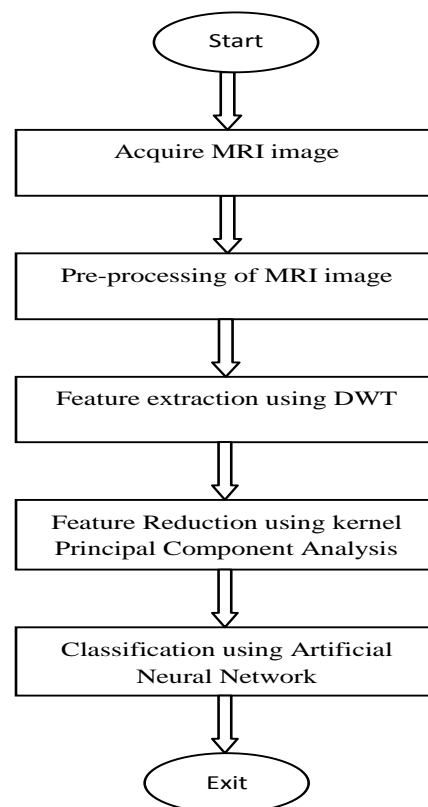


Fig. 1 Methodology

2.1 Data Acquisition

In this paper, we implemented diagnosis method on real MRI data of human brain with high resolution axial images of size 256×256 pixel. Total 120 MR brain images from input dataset are used for training and testing purpose. Out of 120 MR images, we are taken 50 as training images and 70 as testing images. For training five sets of ten images are used. One set of training data has normal MR images. The four sets have abnormal or brain tumor infected MR images in which tumor location are different.

2.2 Pre-Processing of MRI Images

Some MRI images are darker than others due to data acquisition scanner problems. For improving the quality of image, image enhancement techniques are employed. Intensity adjustment is one of the important methods for image enhancement technique, in which the intensity value of original image are mapped into a new range of intensity value. In this paper, gamma correction [7], which is implemented in MATLAB using `imadjust` function, is used to improve MRI images quality.

2.3 Feature Extraction using DWT

For feature extraction from MR brain images Short Time Fourier Transform (STFT) and Wavelet Transform can be used. In STFT, both time and frequency are represented in limited precision. Precision of STFT is determined by window size. Once specific window size is chosen for time window, it will be same for all frequency. Wavelets are short time localized waves with zero average value and possibility of time shifting, flexible etc. DWT is widely used for feature extraction because in case of Wavelet Transform, different frequencies are examined with different resolutions [4]. In proposed method, DWT are used for features extraction from MR images. The features are represented in form of DWT coefficients [18]. MR brain image is applied through series of half band LPF and HPF. The output of HPF is detail coefficients and output of LPF is approximation coefficients [1]. The image can be represented with different resolution levels by wavelet transform. Two level decomposition using DWT is explained in Fig.2.

A_2	D_2^V	D_1^V
D_2^H	D_2^D	
D_1^H		D_1^D

Fig.2 Two Levels Decomposition using DWT

Where, D_2^V , D_2^H , D_2^D are detail of image of input image I_j when resolution is j .

III. KERNEL PRINCIPAL COMPONENT ANALYSIS

To increase the processing speed, feature reduction method is used. The features reduction techniques are applied on features which are extracted by DWT in the form of wavelet coefficients. There are different methods for feature reductions like PCA, KPCA, LDA and SVD. Complex structure of data, which is used for classification, may not be preserved in LDA and SVD. If discriminatory information is contained in the variance of data instead of mean, then LDA will not work properly. In this case PCA is preferred. If the distributions are

non-Gaussian, PCA may not find the best direction for discriminating between two classes. It is difficult to visualize data in dimensions greater than three in PCA. In this case KPCA is preferred over PCA. KPCA applies similar theory as the PCA and it projects data points on low dimensional subspace that captures the highest possible amount of variance in the data. In PCA, data points are separated linearly in original space (referred to as R_d) but in KPCA, data points are separated in a high dimensional space which is known as feature space (referred to as F), by using a mapping function " Ψ " as shown in Fig. 3.

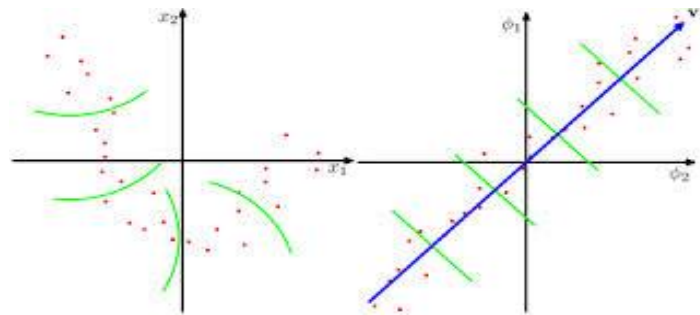


Fig. 3 (a) Nonlinearly distributed data before applying KPCA (b) Data into linearly distributed feature space after applying KPCA

Due to limitation of above methods, KPCA is used in this research paper. Advantages of KPCA are:

1. It does not require nonlinear optimization, knowledge of the network architecture or the number of dimensions.
2. A new point can be quickly projected onto a pre-computed basis.
3. It gets eigenvector with higher variance (principal component) than PCA.

In PCA, covariance matrix is constructed for feature reduction. It is difficult to construct covariance matrix in feature space due to nonlinear projection. Due to this reason kernel trick is used to develop the kernel matrix $Ke(x, x') = \Psi(x) \cdot \Psi(x')$ without explicitly doing the mapping [16] [18]. The data in the feature space is projected onto a low-dimensional subspace, spanned by the eigenvectors that capture most of the variance. One important fact is that without knowing the mapping " ϕ " or the feature space " F ", KPCA is applied to data input. Instead computations are performed on the inner product of pairs of points which are stored in a kernel matrix. The procedure of working with the data in feature space without knowing the mapping " Ψ " is known as "the kernel trick" and is a central part of the kernel PCA method. There are normally two types of kernels used in KPCA: the polynomial kernel and Gaussian kernel.

3.1 Polynomial Kernel

$$Ke(x_j, x_k) = (x_j^T x_k)^p \quad .(1)$$

Or

$$Ke(x_j, x_k) = (x_j^T x_k + c)^p \quad .(2)$$

Where, p denotes polynomial's order and c (constant) > 0

3.2 Gaussian Kernel

$$Ke(x_j, x_k) = \exp(-|x_j - x_k|^2 / 2\sigma^2) \quad .(3)$$

Or

$$K_e(x_j, x_k) = \exp(-\sigma |x_j - x_k|^2) \quad (4)$$

To kernalize the procedure, $x_i \rightarrow \Psi(x_i)$ and

$$K_{e_{jk}} = \Psi(x_j) \Psi(x_k) \quad (5)$$

If data are non-centered set, then inner product computation is applied for centering data.

$$(\tilde{K}_e)_{jk} = (\Psi(x_j) - \bar{\Psi})(\Psi(x_k) - \bar{\Psi}) \quad (6)$$

Where, \tilde{K}_e denotes the centered kernel matrix

$\bar{\Psi}$ represents the center point in the feature space which is written by

$$\bar{\Psi} = \frac{1}{n} \sum_{i=1}^n \Psi(x_i) \quad (7)$$

The eigenvectors of \tilde{K} can be obtained by solving the given equation

$$\tilde{K} \tilde{e} V = \lambda V \quad (8)$$

Where normalized eigenvectors are stored in the column of 'V' and " λ " is a matrix in which the non-diagonal values are zero and diagonal consist of the corresponding Eigen values. The eigenvectors in V are arranged in descending order according to the size of their corresponding Eigen value. Also the Eigen values in λ are sorted in descending order [12].

Using a similar calculation, this can be expressed easily in terms of $K_{e_{ij}}$.

$$K_{e1} = K_e - \mathbf{1}_N K_e - K_e \mathbf{1}_N + \mathbf{1}_N K_e \mathbf{1}_N \quad (9)$$

Where, $\mathbf{1}_N$ is NXN matrix in which all elements equal to $1/N$. In this paper, polynomial kernels, with varying order ($p = 2, 3, 4$) are used to achieve good classification efficiency.

IV. ARTIFICIAL NEURAL NETWORK (ANN)

ANN is highly interconnected numbers of neurons which are arranged processing into different layers [2]. ANN consists of many layers and each layer contains one or more neuron. A single neuron has many inputs; these inputs are multiply with weights and weighted inputs are summed up by summer and then activation function is applied. The biological neuron can be modeled as an artificial neuron [19]. This is explained in Fig. 4 with following steps.

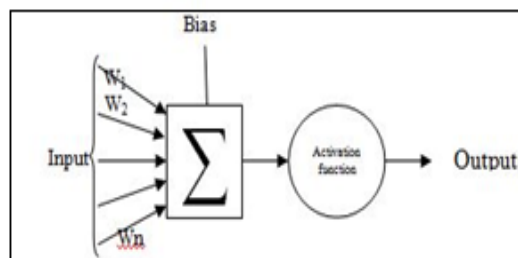


Fig. 4 Single Neuron Model

1. The synapses works as weight element in neuron. When a signal is applied at the input of synapse, the input signal is multiplied by the synaptic weight w_k . The weight w_k is positive or negative depending upon type of synapse.
2. A summer output is the weighted sum of input signal.

3. The output of summer is applied to an activation function which is used to limit the amplitude of the neuron's output. The commonly used normalized amplitude range are $[0, 1]$ or $[-1, 1]$. A bias is used for altering the input of the activation function.

Two fundamental kinds of networks are occurred in neural networks i.e. using feedback and without using feedback. If the output is calculated for each input value, then this is called the network without feedback. In networks with feedback, the output values are applied back to input for tracking the input value. If the information flows in forward direction, then it is called feed forward network [17]. It is shown in Fig.5.

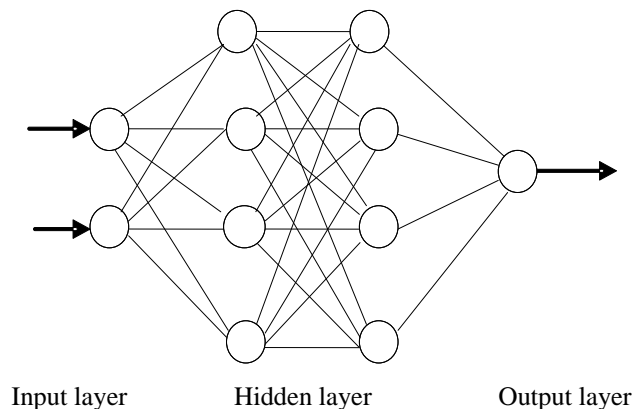


Fig. 5 Feed Forward Neural Networks

There are different types of feed forward network like multilayer perceptions (MLP) and radial basis function (RBF) etc. In MLPs, classes are separated via hyper-planes and in RBFs via hyper-spheres. MLPs have input layer, one or more hidden layers and an output layer. In radial basis function network, a single hidden layer is used which has radial basis activation function for hidden neuron. MLPs have one or more hidden layers but RBFs have only one hidden layer. RBFs require more hidden neurons which lead to curse of dimensionality. The perception learning rule is used for single layer neuron but for more than one layer neuron networks, this training algorithm is not suitable because the output of hidden layer is not available for calculating the output and updating the weight so for multilayer perceptions, perception learning algorithm is not used. For this, error back propagation learning is employed, in which training is done in a supervised manner.

In this paper, multilayer perceptions network with error back propagation learning is used. Back propagation learning rule is based on the error-correction learning rule [13]. In this learning rule, firstly input signal (vector) are applied into input layer and this signal is flowed in forward direction layer by layer through hidden layer to output layer. Then output vector is generated for applied input vector which is known as actual response of the neural network. This is called forward pass. In forward pass, weights of neuron are fixed. Now actual output is compared with desire output or target output. If the difference between desire output and actual output is not equal to zero, then error signal will be generated. This error signal is propagated in backward direction so this is called as "error back propagation". Now according to perceptions error correction rule, weights are updated for minimizing the error signal. In backward pass, weights of neurons are modified so that actual response tends to desire output [5] [13]. Forward NN have three layers like input, hidden and output layer. In this paper, seven different intensity based features after reduced by KPCA are used so seven input elements are used in first layer.

V. RESULTS

For testing seven sets of ten images are used. Three set of training data contains normal MR images & rest four sets contain abnormal or brain tumor infected MR images. The result in terms of classification accuracy for ANN with different feature reduction method is given in Table 1 for MR brain images.

Table 1: Comparison of polynomial KPCA with other feature reduction techniques

Classifier(ANN) with different feature reduction methods	Classification Accuracy (%)
PCA	86.7 %
LDA	93.3%
SVD	80.2%
KPCA (Gaussian)	86.7 %
KPCA(polynomial p=2)	93.7%
KPCA(polynomial p=3)	96.7%
KPCA(polynomial p=4)	99.8 %
KPCA(polynomial p=5)	86.7 %

From above table, it is clear that KPCA with ANN for power $p = 4$ gives maximum classification accuracy (99.8 %).

VI. CONCLUSION

This work proposed the efficient method for detection of normal or abnormal brain images. Various classification rates are obtained using different power of applied kernel. Confusion matrix obtained by classification gives, 93.7% accuracy with $p=2$, 96.7% accuracy with $p=3$, 99.8% accuracy with $p=4$, 86.7% accuracy with $p=5$. From above result, it can be concluded that KPCA with kernel's power = 4 gives best classification result. Here, in this paper, discrete wavelet transform (DWT) is used as feature extraction technique from brain MRI images. There are various others transforms and techniques available for extracting image features so in future work, different feature extraction methods can be used to achieve higher classification accuracy for large MRI data.

VII. ACKNOWLEDGEMENTS

Special thank to Institute of Nuclear Medicine & Allied Science (INMAS), Defence Research & Development Organization (DRDO), Govt of India, for providing real data set of brain MR images.

REFERENCES

- [1] Stephane Mallat, "A WAVELET TOUR OF SIGNAL PROCESSING" Second Edition, Ecole Polytechnique, Paris Courant Institute, New York University, 1999, Elsevier (USA)
- [2] S.Haykin, Neural Networks, A Comprehensive Foundation, Prentice Hall, 1999.
- [3] L.M. Fletcher-Heath, L.O. Hall, D.B. Goldgof, and F.R. Murtagh, "Automatic segmentation of non-enhancing brain tumors in magnetic resonance images", Artif. Intell. Med 21, 2001, 43-63.
- [4] K. Karibasappa, S. Patnaik, "Face recognition by ANN using wavelet transform coeff," IE (India) Journal of Computer Eng., vol . 85, pp. 17-23, 2004.

- [5] M. O'Farrell, E. Lewis, C. Flanagan, and N. Jackman, "Comparison of k-NN and neural network methods in the classification of spectral data from an optical fibre-based, sensor system used for quality control in the food industry", *Sens. Actuators B: Chemical* 111–112C (2005) 354–362.
- [6] S.Chaplot, L.M.Patnaik and N.R.Jagannathan "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network", *Biomed Signal Process. Control*, 1 (2006) 86–92.
- [7] Gerard Blanchet, and Maurice Charbit, "Digital Signal and Image Processing using MATLAB", TK5102.9.B545 2006 ,621.382'2--dc22
- [8] M.Maitra, A.Chatterjee "Hybrid multiresolution Slantlet transform and fuzzy c-means clustering approach for normal-pathological brain MR image segregation", *Med. Engg. Phys.* (2007), doi: 10.1016/j.medengphy .2007 .06.009.
- [9] S. Kara, and F. Dirgenali, "A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks", *Expert Syst.Appl*, 32 (2007).
- [10] Noramalina Abdullah, Lee Wee Chuen, and Umi Kalthum Ngah Khairul Azman, "Improvement of MRI Brain Classification using Principal Component Analysis", 2011 IEEE International Conference on Control System, Computing and Engg.
- [11] Quanquan Gu, Zhenhui Li, and Jiawei Han, "Linear Discriminant Dimensionality Reduction", Department of Computer Science, University of Illinois at Urbana- Champaign Urbana, IL 61801, US.
- [12] Daniel Olsson, Master's Thesis Report Applications and Implementation of Kernel Principal Component Analysis to Specific Data Sets, January 28, 2011
- [13] N.Hema Rajini, and R.Bhavani, "Classification of MRI Brain Images using k- Nearest Neighbor and Artificial Neural Network", IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011, MIT, Anna University,
- [14] Davi Marco Lyra-Leite, Joao Paulo Carvalho Lustosa da Costa, and Joao Luiz Azevedo de Carvalho, "Improved MRI reconstruction and denoising using SVD-based low-rank approximation", Digital Signal Processing Group Department of Electrical Engineering University of Brasilia Brasilia-DF, Brazil.
- [15] Harvard Medical School, Web, data available at <http://med.harvard.edu/AANLIB/>.
- [15] Quan Wang, "Kernel Principal Component Analysis and its Applications in Face Recognition and Active Shape Models" arXiv: 1207.3538v1 [cs.CV] 15 Jul 2012
- [17] S.N.Deepa, and B.Aruna Devi, Artificial Neural Networks design for Classification of Brain Tumour" 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, 2012, Coimbatore, India
- [18] Devvi Sarwinda, and Anjani M Arymurthy, "Feature Selection Using Kernel PCA for Alzheimer's Disease Detection with 3D MR Images of Brain", ICACIS 2013 ISBN: 978-979-1421-19-5
- [19] Walaa Hussein Ibrahim, Ahmed Abdel Rhman Ahmed Osman and Yusra Ibrahim Mohamed, "MRI Brain Image Classification Using Neural Networks", 2013 International Conference On Computing, Electrical And Electronics Engineering (Iccee), 978-1-4673-6232, 2013.
- [20] Jagdish H Pujar Pallavi S. Gurjal, Shambhavi D.S, Kiran S. Kunnur. "Medical Image Segmentation based on Vigorous Smoothing and Edge Detection Ideology", *International Journal of Electrical and Computer Engineering* 5:2 2010

Optimization of Process Parameters with Effect of Thrust and Torque in Drilling Operation

Ashish Tripathi¹, Ranganath M Singari², Vipin Dahiya³

¹(Mechanical, Production and Industrial Engineering,
Delhi Technological University, Delhi, India)

Email: tripathiashishme90@gmail.com

Abstract : The measurement of torque and thrust in drilling enable to find the amplitude of vibration produced that will decide the type of setup needed for the work. The objective of this research paper is to find thrust and torque in mild steel and aluminum 6061 and to find optimal values of drilling parameters so that minimum value of torque and thrust is obtained. After finding the values of thrust and torque in mild steel and aluminum 6061 the comparison of values for both the materials has been presented. The lower value of thrust and torque will result into higher tool life and lower vibration in machine tool structure.

Keywords : Drilling, Torque, Thrust, Taguchi, Optimization

I. INTRODUCTION

Drilling is a method to produce hole in which different investigation and studies has been completed over a compass of time. Drilling is used for assembling industry like watch assembling industry, Aerospace industry, Automobile industry, semiconductors and restorative commercial ventures. Drilling can be described as a process where a multi-point tool is used to remove unwanted materials to produce a desired hole. Hole making had long been recognized as the most prominent machining process, requiring specialized techniques to achieve optimum cutting condition. It is therefore, essential to optimize quality and productivity simultaneously.

Commercial enterprises in assembling area have depended on the utilization of handbook based upon the data and the administrator's experience. To add to a surface unpleasantness model and to enhance, it is fundamental to comprehend the present status of work here. Customary practice prompts poor surface complete and lessening in the ease of use and efficiency alongside usefulness because of inadequately streamlined utilization of machining capacity. This causes high assembling cost and low item quality. In the furniture industry, for instance, large quantities of holes have to be drilled due to the use of connections, handles and hinges. A considerable part of this field is still being devoted to major drilling-optimization issues such as the appropriate cutting parameters or tool geometries. Controllable procedure parameters incorporate cutting pace, device geometry, nourish and instrument setup. Different components, for example, device, work piece, apparatus wear, machine vibration, corruption, work piece and instrument material variability can never be controlled effectively.

The vital cutting parameters considered for exchange here are profundity of cut, sustain and cutting speed. It is observed that in the majority of the cases surface harshness diminishes with expansion in pace of cutting and decline in feed and profundity of cut. Since these cutting parameters will choose the kind of chips, which are gotten at the season of machining of a solitary steady material in this manner, which need to examine them so that no such manufactured up edge chips development.

Drilling is presumably the most vital customary mechanical procedure related with chipboard preparing. In the industry of furniture, chip drilling requires entirely different process parameters for the optimization process: in the earlier process, the smoothness or surface roughness of the surface processed and tool wear are given equal importance. In chip drilling, the earlier parameter is prioritized over the new given the complexity to drill.

A suitable model can help in the focused recruitment of the most apt feed rates, spindle speeds and geometrical cutting tool shapes.

The measurement of torque and thrust in drilling enable to find the amplitude of vibration produced that will decide the type of setup needed for that work. Furthermore, can determine the work piece material suitable for different type of drill material. In addition, the works that has been previously done on the subject were done manually there was less work done by the better design of experiment technique Taguchi. The material also is not an old material but a relatively new and is having better and good properties than the previously available material.

The objective of this research paper is to find thrust and torque in mild steel and aluminum 6061 and to find optimal values of drilling parameters so that minimum value of torque and thrust is obtained. After finding the values of

thrust and torque in mild steel and aluminum6061 the comparison of values for both the materials has been presented. The lower value of thrust and torque will result into higher tool life and lower vibration in machine tool structure.

II. LITERATURE REVIEW

S.R. Karnik et al. [1] explained the delamination behaviour as a function of drilling process parameters at the entrance of the CFRP plates. The analysis in high speed drilling is performed by developing an artificial neural network (ANN) model with speed, feed and point angle as the affecting parameters. A multilayer feed forward ANN architecture, trained using error-back propagation-training algorithm (EBPTA) is employed for this purpose. Drilling experiments were conducted as per full factorial design using cemented carbide (grade K20) twist drills that serve as input– output patterns for ANN training. The validated ANN model is then used to generate the direct and interaction effect plots to analyze the behaviour. The simulation results explain the effectiveness of the ANN models to analyze the effects of drilling process parameters on delamination factor. The analysis also demonstrates the advantage of employing higher speed in controlling the delamination during drilling. Erkan Bahçe and Cihan Ozel [2] investigated, BUEs arisen on the cutting edges and the effect of drilling parameters (rotation speed, feed rate, drill diameter and point angle) on the surface roughness of the work piece experimentally in the drilling process of Al 5005 alloy on CNC milling machine. The objective of the paper is to obtain an optimal setting of the drilling process parameter (cutting speed, feed) resulting in an optimizing axial force in drilling on mild steel. The effect of drilling process parameter on axial force using Taguchi's parameter design approach. Results indicates that the selected process parameter affect the machining characteristics. The result indicate that the selected process parameter affect the mean and the variance of the optimal force. The percentage contribution of the parameters reveal that the influence of the feed (64.72%) in controlling both mean and variation of axial force is significantly larger than that of cutting speed(33.38%).The predicted optimum axial force is (720N).The result have been validated by confirmation experiments. C.C. Tsao and H. Hocheng [3] predicted and evaluated thrust force and surface roughness in drilling of composite material using candle stick drill. The approach was based on Taguchi method and the artificial neural network. The experimental results showed that the feed and the drill diameter are the most significant factors affecting the thrust force, while the feed rate and speed contributed the most to the surface roughness. In the study, the objective was to establish a correlation between the feed, speed and drill diameter with the induced thrust force and surface roughness in drilling composite laminate. The correlations were obtained by multi-variable regression analysis and radial basis function network (RBFN) and compared with the experimental results. The results indicated the RBFN is more effective than multivariable regression analysis. Reddy Sreenivasalu and Ch.Srinivasa Rao[4] ,investigated the effect of drilling parameters on surface roughness and roundness error in drilling of Al6061 alloy

with HSS twist drill. Optimal control factors for hole quality were determined using Taguchi grey relational analysis. Cutting speed, feed rate, drill diameter, point angle and cutting fluid mixture ratio were considered, as the control factors.L18 orthogonal array was determined for experimental trials. Grey relational analysis was employed to minimize surface roughness and roundness error. Yogendra Tyagi et al. [5] discussed the feasibility of machining Mild Steel by CNC drilling machine with a HSS Tool by applying taguchi methodology. The signal-to-noise ratio applied to find optimum process parameter for CNC drilling machining. A L9 orthogonal array and analysis of variance (ANOVA) were used to study the performance characteristics of machining parameter (spindle speed, feed, and depth) with consideration of good surface finish as well as high material removal rate (MRR). Multiple regression equations were formulated for estimating predicted value for surface roughness and material removal rate.

Dinesh Kumar et al. [6] described the Taguchi technique for optimization of surface roughness in drilling process. In their investigation, Taguchi technique was used as one of the method for minimizing the surface roughness in drilling mild steel. The objective of this study was to investigate the influence of parameters, such as cutting speed and feed rate, and point angle on surface roughness produced when drilling Mild steel. B. Shivapragash et al. [7] focused on multiple response optimization of drilling process for composite Al-TiBr2. The study provided to minimize the damage events occurring during drilling process for composite material. A statistical approach used to analyze experiment data. Taguchi method with grey relational analysis was used to optimize the machining parameters with multiple performance characteristics in drilling of MMC Al-TiBr2. The results shows that the maximum feed rate, low spindle speed are the most significant factors, which affect the drilling process, and the performance in the drilling process can be effectively improved by using this approach. R. S. Jadoun et al. [8] showed the effect of parameters on production accuracy obtained through ultrasonic drilling of holes in alumina-based ceramics with using silicon carbide abrasive. Production accuracy in ultrasonic drilling involved both dimensional accuracy and formed accuracy. The parameters considered were work piece material, tool material, grit size of the abrasive, power rating and slurry concentration. Taguchi's optimization approach was used to obtain the optimal parameters. The significant parameters were also identified and their effect on oversize, out-of-roundness and conicity were studied. The results obtained were validated by conducting the confirmation experiments.

J. Mathew et al. [9] studied that thrust is a major factor responsible for delamination and it mainly depends on tool geometry and feed rate. Trepanning tools, which were used in this study, were found to give reduced thrust while making holes on thin laminated composites. In this work the peculiarities of trepanning over drilling of unidirectional composites has been emphasized. The models for prediction of critical thrust and critical feed rate at the onset of delamination during trepanning of uni-directional composites based on fracture mechanics and plate theory also have been presented. Mathematical models correlating thrust and torque

with tool diameter and feed rate have been developed through statistically designed experiments and effect of various parameters on them have been discussed. It is observed that sub-laminate thickness is the most decisive parameter from the viewpoint of critical feed rates. L-B Zhang et al. [10] studied the analysis for multidirectional composite laminates is based on linear elastic fracture mechanics (LEFM), classical bending plate theory and the mechanics of composites. This paper presents a general closed-form mechanical model for predicting the critical thrust force at which delamination is initiated at different ply locations. Good correlation is observed between the model and the experimental results.

Mustafa Kurt et al. [11] investigated the effect of cutting parameters and changes in the diameter of the drill on drilling temperature, cutting forces and surface roughness in dry drilling of Al 2024 alloy using diamond like carbon-coated drills. The cutting temperature was measured by using a thermocouple. Kistler dynamometer was used for the measurements of the cutting forces during drilling process. The surface roughness of drilled hole surface was determined by using surface roughness meter. Finally, the Taguchi technique was employed for the optimization of drilling process.

A Çiçek et al. [12] investigated the effects of deep cryogenic treatment and drilling parameters on surface roughness and roundness error were in drilling of AISI 316 austenitic stainless steel with M35 HSS twist drills. In addition, optimal control factors for the hole quality were determined by using Taguchi technique. Two cutting tools, cutting speeds and feed rates were considered as control factors, and L8 (23) orthogonal array was determined for experimental trials. Multiple regression analysis was also employed and derived the predictive equations of the surface roughness and roundness error achieved via experimental design. Minimum surface roughness and roundness error were obtained. Shravan Kumar Singh [13], the objective of the paper was to obtain an optimal setting of the drilling process parameter (cutting speed, feed) resulting in an optimizing axial force in drilling on mild steel. The effect of drilling process parameter on axial force using Taguchi's parameter design approach. Results indicate that the selected process parameter affect the machining characteristics. The result indicate that the selected process parameter affect the mean and the variance of the optimal force. The percent contribution of parameters as quantified in the S/N pooled ANOVA. The percentage contribution of the parameters reveal that the influence of the feed (64.72%) in controlling both mean and variation of axial force was significantly larger than that of cutting speed (33.38%). U.A. Khashaba et al. [14] studied the effects of the drilling parameters, speed, and feed, on the required cutting forces and torques in drilling chopped composites with different fibre volume fractions. Three speeds, five feeds, and five fiber volume fractions are used in this study. The results show that feeds and fibre volumes have direct effects on thrust forces and torques. On the other hand, increasing the cutting speed reduces the associated thrust force and torque, especially at high feed values. Using multivariable linear regression analysis, empirical formulas that correlate favourably with the obtained results have been developed. These formulas would be useful in drilling

chopped composites. The influence of cutting parameters on peel-up and push-out delaminations that occurs at drill entrance and drill exit respectively the specimen surfaces have been investigated. No clear effect of the cutting speed on the delamination size is observed, while the delamination size decreases with decreasing the feed. Delamination-free in drilling chopped composites with high fibre volume fraction remains as a problem to be further investigated. Irfan Khan et al. [15] presented the use of Taguchi method to achieve the minimum hole size expansion in drilling of acrylic sheet. The purpose of this paper was to investigate the effect of cutting parameters, such as cutting speed and feed rate, and point angle on hole size (without considering the thermal effect). The analysis of experiments showed that Taguchi method could successfully verify the optimum cutting parameters. The level of the best of the cutting parameters on the hole expansion was determined. It was also found that cutting speed was the major factor on which hole size expansion depends. It was lower with higher cutting speed. Feed was another important factor which causes hole expansion but its dependence was lesser as compared to the speed. Tool angle was the most complex factor. It was suggested that for achieving minimum hole expansion on the Acrylic always higher feed rates and higher cutting speeds were preferred with non-standard angles.

N.S. Mohan et al. [16] studied the drilling parameters and specimen parameters evaluated were speed, feed rate, and drill size and specimen thickness on FRP. A signal-to-noise ratio is employed to analyze the influence of various parameters on peel up and push down delamination factor in drilling of glass fibre reinforced plastic (GFRP) composite laminates. The main objective of this study was to determine factors and combination of factors that influence the delamination using Taguchi and response surface methodology and to achieve the optimization machining conditions that would result in minimum delamination. From the analysis, it is evident that among the all-significant parameters, specimen thickness and cutting speed have significant influence on peel up delamination and the specimen thickness and feed have more significant influence on push down delamination. In practice, during study of drilling generally empirical equations are used to calculate thrust force and torque. These equations are approximate, because they do not consider all the cutting parameters. They often use only the feed speed and the diameter of the drill. Tsao reported that the feed rate and the drill diameter are recognized as the most significant factors affecting the thrust force. Depending upon the variables various researchers used various methods to find out the force and thrust produced in the drilling process. E. S. Lee [17] studied the machinability of GFRP by means of tools made of various materials and geometries were investigated experimentally. By proper selection of cutting tool material and geometry, excellent machining of the work piece is achieved. The surface quality relates closely to the feed rate and cutting tool. When using glass fibre reinforced plastics (GFRP) it is often necessary to cut the material, but the cutting of GFRP is often made difficult by the delamination of the composite and the short tool life. E. Ugo. Enemuoh et al. [18] studied new comprehensive approach to select cutting parameters for

damage-free drilling in carbon fiber reinforced epoxy composite material is presented. The approach is based on a combination of Taguchi's experimental analysis technique and a multi objective optimization criterion. The optimization objective includes the contributing effects of the drilling performance measures: delamination, damage width, surface roughness, and drilling thrust force. A hybrid process model based on a database of experimental results together with numerical methods for data interpolation are used to relate drilling parameters to the drilling performance measures. Case studies are presented to demonstrate the application of this method in the determination of optimum drilling conditions for damage-free drilling in BMS 8-256 composite laminate. Hocheng and C.C. Tsao [19] studied the critical thrust force at the onset of delamination, and compare the effects of these different drill bits. The results confirm the analytical findings and are consistent with the industrial experience. Ultrasonic scanning is used to evaluate the extent of drilling-induced delamination in fiber-reinforced materials. The advantage of these special drills is illustrated mathematically as well as experimentally, that their thrust force is distributed toward the drill periphery instead of being concentrated at the center. The allowable feed rate without causing delamination is also increased. Ranganath M S. et al. [20] conducted a series of experiments on an aluminium material to find out the optimization process on them. These experiments were classified in taguchi or the RSM process to decide which is better. The process is done by the design of experiments and the RSM technique was found to better and predicted better results and good surface finish. A lot of work has been done on single material drilling and its analysis. However, no work in which taguchi is being applied to compare the torque and thrust values in two different materials. The comparison of torque and thrust value will enable us to find the dependence of torque and thrust in strength of material.

III. EXPERIMENTAL SETUP

Radial Drilling machine is type of drilling machine which has more number of degree of freedom than conventional drilling machine and this machine is named so because the drilling tool can be move radial in radial arm fixed to machine.



Figure 1: Radial Drilling machine

Aluminium-6061 is precipitation hardening aluminium alloy. It is easily available in pre-tempered grades such as 6061-O which means it is annealed and tempered grades such as

6061-T6 which means it is solution zed and artificially aged and 6061-T651 which means it is solution zed, stress-relieved stretched and artificially aged. The first digit of the four digits in the designation implies the alloy group in terms of the major alloying elements of which it is constituted.

Table 1: Composition of Al-6061

ELEMENT	PERCENT PRESENT
Magnesium(Mg)	0.80-1.20
Silicon(Si)	0.40-0.80
Iron(Fe)	0.0-0.70
Copper(Cu)	0.15-0.40
Chromium(Cr)	0.04-0.35
Zinc(Zn)	0.0-0.25
Titanium(Ti)	0.0-0.15
Manganese(Mn)	0.0-0.15
Others(Total)	0.0-0.15
Other(each)	0.0-0.15

Table 2: Properties of Aluminium 6061

PROPERTY	VALUE
Density	2.7 g/cc
Brinell Hardness Number	95
Ultimate Tensile Strength	310 MPa
Tensile Yield Strength	276 MPa
Modulus of Elasticity	68.9 GPa
Poisson's ratio	0.33
Fatigue Strength	96.5 MPa
Machinability	50%(comparing with other aluminium alloys)
Shear Modulus	26 GPa
Shear Strength	207 MPa
Thermal Conductivity	167 W/m-k
Melting Point	582-652°C

6061 is mostly used for the following:

- Constructing structure of aircraft, such as wings and fuselages. 6061 is more generally utilised as a part of homebuilt aircraft than business or military aircraft. Aluminium alloy 2024 has more strength, but 6061 aluminium alloy is easily workable and it also behaves resistant to corrosion even when the material surface is peeled off. Whereas, 2024 is corroded easily, hence it is used with a thin coating of Al-clad to make it corrosion resistive.
- Yacht construction, including small utility boats.
- wheel spacers used as a automotive parts
- Packaging of foodstuffs and beverages using aluminium cans.
- Bicycle components and frames.

- The acclaimed Pioneer plaque had been made of this specific amalgam of aluminium.
- The auxiliary chambers and baffle systems in firearm sound silencer (essentially gun silencer for decreasing weight and enhancing mechanical operation). Lower and upper receivers of majority of the AR-15 rifle variants.

Most of the aluminium the docks and the gangways are built with 6061-T6 extrusions, and welded into spot.

- aluminium 6061 is likewise utilized in some ultra-high vacuum (UHV) chambers

Parts for remote control model aircraft, for most of helicopter rotor segments

Table 3: Chemical composition of mild steel

Elements	Maximum Weight %
Carbon(c)	0.45
Sulphur(s)	0.60
Manganese(Mn)	1.00
Phosphorous(p)	0.40
Silicon(s)	0.35

Table 4: Mechanical and Physical properties of Mild Steel

PROPERTIES	VALUES
Density	7085 kg m-3
Thermal conductivity	48 Jm-1K-1S-1
Thermal expansion	11.3×10-6 K
young's modulus	210 GNm-2
Tensile strength	600 MNm-2

Dynamometer

A dynamometer is a device used to find the torque and thrust in the machine in either drilling or any other machining process. Dynamometer is nothing but a simple electrical machine, which is used to measure the value of the force and torque exactly in the process.



Figure 2: dynamometer in Metal Cutting Lab, DTU

Taguchi Method

Designs of Experiment Techniques are those techniques, which will provide us a complete set of optimal parameters in

few numbers of experiments. These methods consume less time and are cost effective. These methods are optimization methods. The most widely used optimization methods are response surface methodology and taguchi method

Taguchi has built up a technique for the utilization of outlined analyses, including an expert's handbook. This strategy has taken the outline of investigations from the selective universe of the analyst and brought it even more completely into the universe of assembling. His commitments have additionally made the professional work easier by bolstering the utilization of less trial plans, and giving a clearer comprehension of the variation nature and the monetary outcomes of value building in the realm of assembling. Taguchi presents his methodology, utilizing test configuration for:

- planning items/forms in order to be strong to natural conditions;
- planning and creating items/forms in order to be strong to part variety;
- minimizing variety around an objective worth

The goal of the parameter configuration is to advance the settings of the procedure parameter values for enhancing execution attributes and to distinguish the item parameter values under the ideal procedure parameter values. Furthermore, it is normal that the ideal procedure parameter qualities got from the parameter configuration are unfeeling to the variety of natural conditions and other commotion variables. In this way, the parameter configuration is the key stride in the Taguchi technique to accomplishing high caliber without expanding expense.

Taguchi Design and Optimization

In this analysis, three machining variables were chosen as control components, and every parameter was intended to have three levels, signified 1, 2, and 3 (Table 5). The trial outline was by L27 (3¹³) cluster in view of Taguchi system, while utilizing the Taguchi orthogonal exhibit would extraordinarily lessen the quantity of tests. An arrangement of examinations planned utilizing the Taguchi system was directed to explore the connection between the procedure parameters and delamination element.

Table 5: Different Factors

d(drill bit diameter)(mm)	S(speed)(mm/min)	f(feed)(mm/rev)
8	150	0.12
10	300	0.20
12	440	0.30

Taguchi Orthogonal Array Design

L27 (3³3), Factors: 3, Runs: 27

Columns of L 27(3¹³) Array

d= drill diameter (mm)

s=speed of rotation (rpm)

f=feed (mm)

T1=Torque (Kgf-m), T2=Thrust (Kgf)

FOR AL-6061

Table 6: Experimental Details

Exp.No.	A	B	C	T1	T2
1	1	1	1	0.2	72
2	1	1	1	0.2	76
3	1	1	1	0.2	69
4	1	1	2	0.2	85
5	1	1	2	0.2	87
6	1	1	2	0.3	81
7	1	1	3	0.4	145
8	1	1	3	0.6	168
9	1	1	3	0.4	152
10	2	1	2	0.4	121
11	2	1	2	0.5	132
12	2	1	2	0.7	136
13	2	2	3	0.7	148
14	2	2	3	0.7	158
15	2	2	3	0.7	153
16	2	2	1	0.4	101
17	2	2	1	0.3	92
18	2	2	1	0.3	97
19	3	2	3	1.0	255
20	3	2	3	1.2	262
21	3	2	3	1.2	259
22	3	3	1	0.6	157
23	3	3	1	0.6	158
24	3	3	1	0.6	158
25	3	3	2	0.9	212
26	3	3	2	0.9	203
27	3	3	2	0.9	205

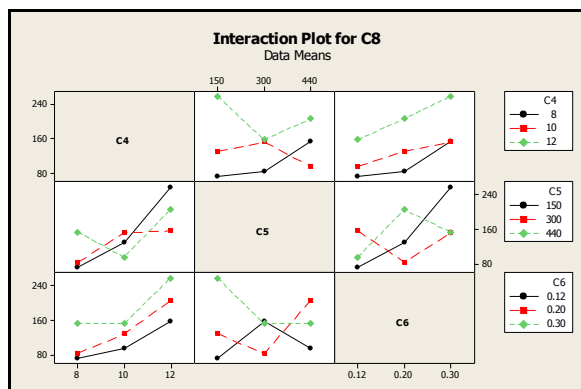


Figure 3: Interaction Plot for Thrust

In first row: for minimum thrust
If s=150 then d=8; if f=0.12 then d=8
If s=300 then d=8; if f=0.20 then d=8
If s=440 then d=10; if f=0.32 then d=8

In second row: for minimum thrust
If d= 8 then s=150; if f=0.12 then s=150
If d=10 then s=440; if f=0.20 then s=300
If d=12 then s=300; if f=0.30 then s=440

In third row: for minimum thrust
If d=8 then f=0.12; if s=150 then f=0.12
If d=10 then f=0.12 ;if s=300 then f=0.20
If d=12 then f=0.12 ; if s=440 then f=0.12

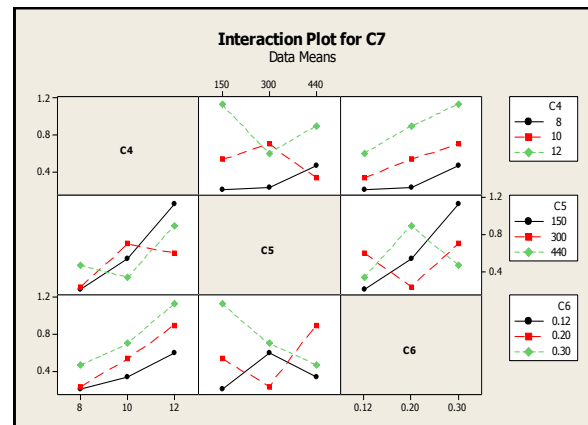


Figure 4: Interaction Plot for Torque

In first row: for minimum torque
If s=150 then d=8; if f=0.12 then d=8
If s=300 then d=8; if f=0.20 then d=8
If s=440 then d=10; if f=0.30 then d=8

In second row: for minimum torque
If d= 8 then s=150; if f=0.12 then s=150
If d=10 then s=440; if f=0.20 then s=300
If d=12 then s=300; if f=0.30 then s=440

In third row: for minimum torque
If d=8 then f=0.12; if s=150 then f=0.12
If d=10 then f=0.12; if s=300 then f=0.20
If d=12 then f=0.12; if s=440 then f=0.12

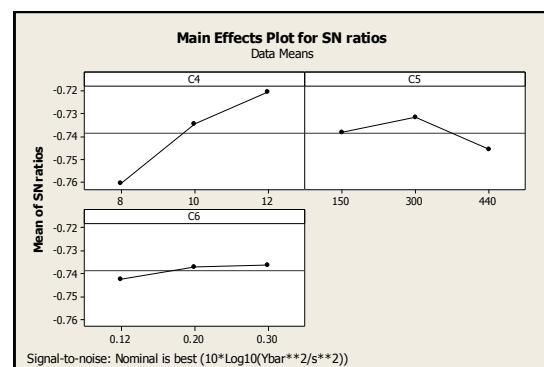


Figure 5: Main Effect Plot for S-N Ratio

Nominal is best ($10 \cdot \log_{10} (\bar{Y} \cdot 2 / s^2)$)

Table 7 Response Table for Signal to Noise Ratios

Level	C4	C5	C6
1	-0.7607	-0.7385	-0.7426
2	-0.7348	-0.7319	-0.7373
3	-0.7206	-0.7456	-0.7361
Delta	0.0401	0.0137	0.0065
Rank	1	2	3

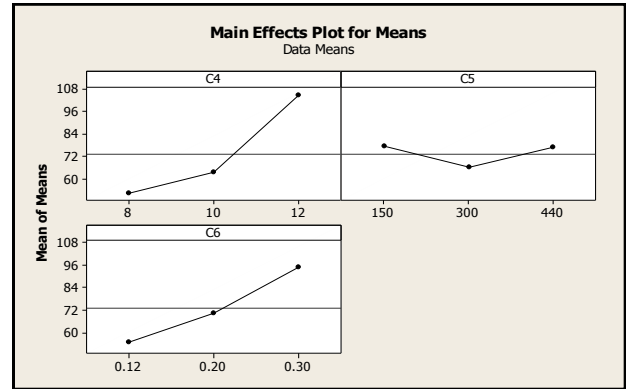


Figure 7: Main Effects Plot for Thrust

Table 8: Response Table for Means

Level	C4	C5	C6
1	52.09	77.09	54.63
2	63.48	66.09	70.39
3	104.27	76.67	94.83
Delta	52.18	11.00	40.19
Rank	1	3	2

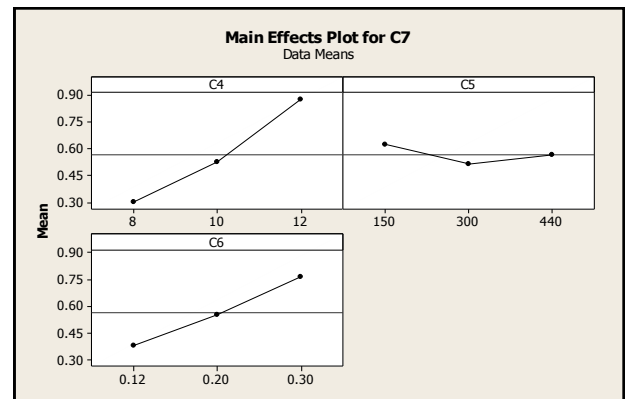


Figure 8: Main Effects plot for Torque

Table 9: Response Table for Standard Deviations

Level	C4	C5	C6
1	56.88	83.85	59.48
2	69.07	71.87	76.58
3	113.28	83.52	103.17
Delta	56.40	11.98	43.69
Rank	1	3	2

From the graph shown in figure 7 the thrust increases with increases in drill diameter and feed but it decreases with increases in speed of rotation of drill.

Figure 8 indicates that the value of torque increases as the value of drill diameter increases and the relation is linear. The torque decreases as the value of speed of rotation of drill increases. While the torque and feed are nearly directly proportion to each other.

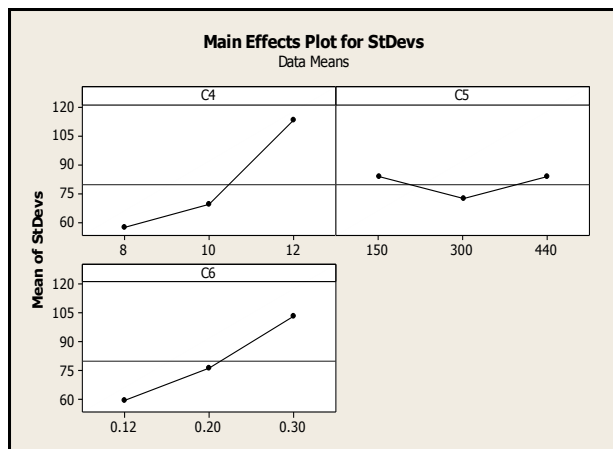


Figure 6: Main Effects Plot for StDevs

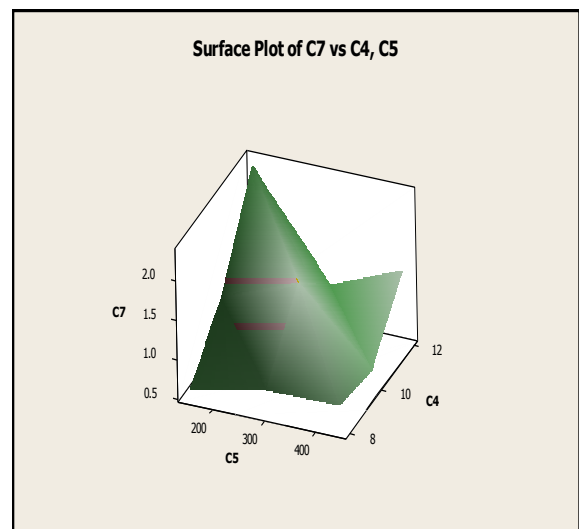


Figure 9: Thrust vs. Drill diameter and speed of Drill

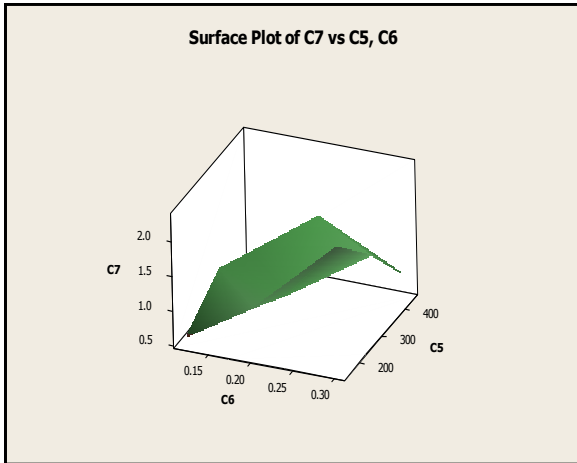


Figure 10: Torque VS feed and speed of Drill

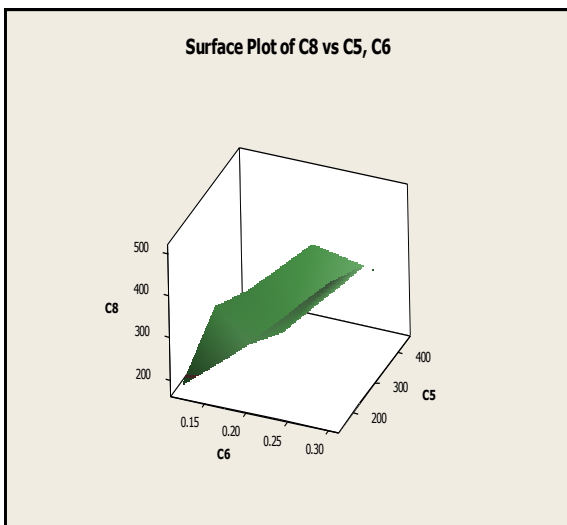


Figure 11: Thrust VS feed and speed of Drill

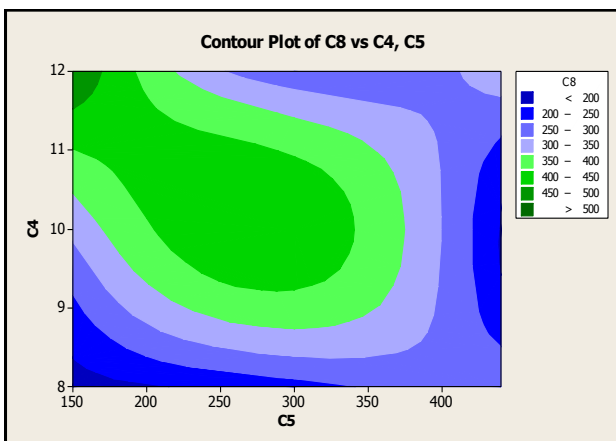


Figure 12: Contour Plot of Thrust vs. Drill diameter and Speed of Drill

A contour plot is a graphical technique for representing a 3-dimensional surface by plotting constant z slices called contours on a 2-dimensional format.

Figure 12 clearly indicates that 1-for low value of thrust speed should be high and and drill should be low.

2-for low speed and high drill diameter the value of thrust is high.

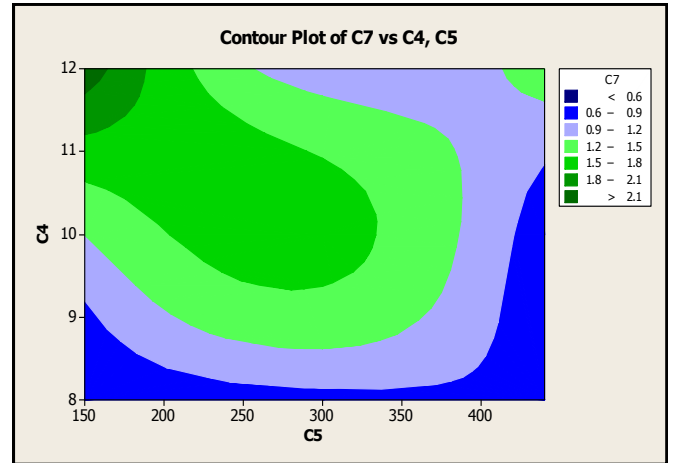


Figure 13: Contour Plot of Torque vs. Drill diameter and Speed of Drill

Figure 13 indicates that 1-for low value of torque speed should be high and and drill should be low.

2-for low speed and high drill diameter the value of thrust is high

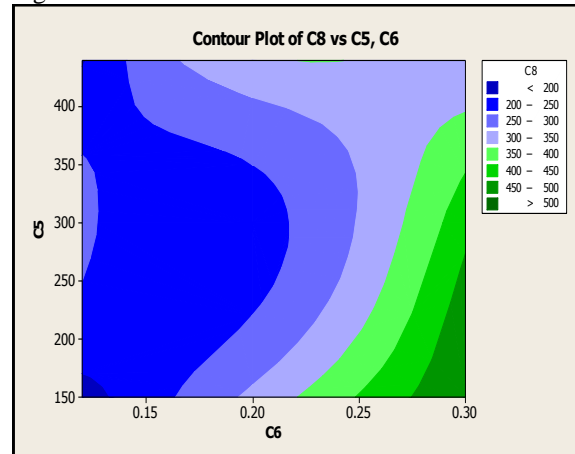


Figure 14: Contour Plot of Thrust vs. feed and Speed of Drill

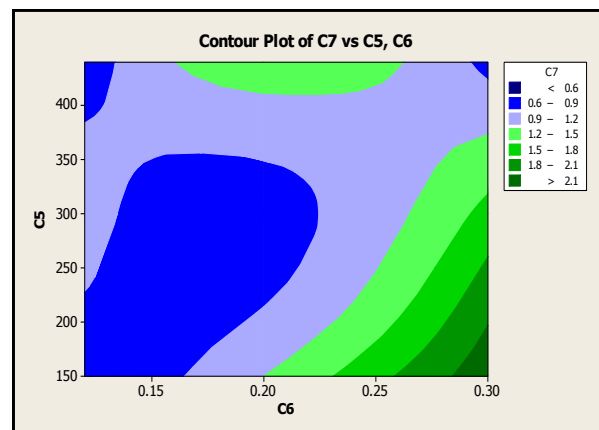


Figure 15: Contour Plot of Torque vs. feed and Speed of Drill

From Figure 14 we can say that the value of thrust is between 200Kg to 250Kgfor feed between 0.15mm/rev to 0.20mm/rev for all speeds and the value of thrust is higher for higher feed and lower speed of rotation of drill.

Figure 15 clearly indicates that

1-For lower torque feed should be low.

2-The value of torque is between 0.6Kg-N to 0.9Kg-N for Medium Feed And Speed.

FOR MILD STEEL

Table 10: Response Table for Signal to Noise Ratios

Level	C4	C5	C6
1	-0.7371	-0.7236	-0.7295
2	-0.7274	-0.7282	-0.7257
3	-0.7191	-0.7317	-0.7283
Delta	0.0180	0.0081	0.0037
Rank	1	2	3

Table 11: Experimental Details

S.NO	A	B	C	T1	T2
1	1	1	1	0.6	180
2	1	1	1	0.6	180
3	1	1	1	0.6	180
4	1	2	2	0.8	232
5	1	2	2	0.8	232
6	1	2	2	0.8	232
7	1	3	3	0.8	300
8	1	3	3	0.8	300
9	1	3	3	0.8	300
10	2	1	2	1.2	312
11	2	1	2	1.2	312
12	2	1	2	1.2	312
13	2	2	3	1.6	433
14	2	2	3	1.6	433
15	2	2	3	1.6	433
16	2	3	1	0.7	195
17	2	3	1	0.7	195
18	2	3	1	0.7	195
19	3	1	3	2.3	500
20	3	1	3	2.3	500
21	3	1	3	2.3	500
22	3	2	1	1	260
23	3	2	1	1	260
24	3	2	1	1	260
25	3	3	2	1.4	340
26	3	3	2	1.4	340
27	3	3	2	1.4	340

Table 12: Response Table for Means

Level	C4	C5	C6
1	119	166	106.2
2	157.3	154.7	147.9
3	184.1	139.6	206.3
Delta	65.1	26.4	100.1
Rank	2	3	1

Table 13: Response Table for Standard Deviations

Level	C4	C5	C6
1	129.6	180.4	115.5
2	171	168.3	160.8
3	200	151.9	224.3
Delta	70.4	28.4	108.7
Rank	2	3	1

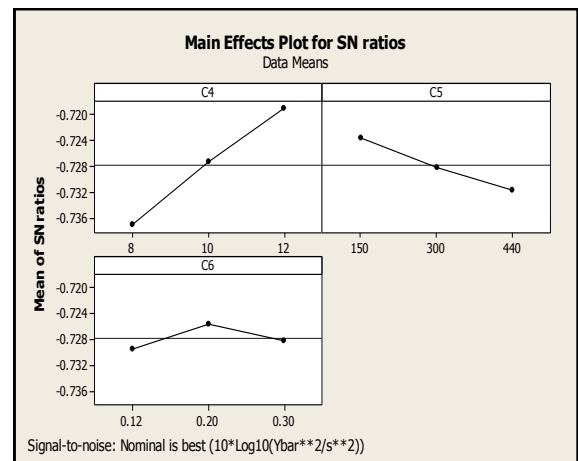


Figure 16: Main Effect Plot for S-N Ratio

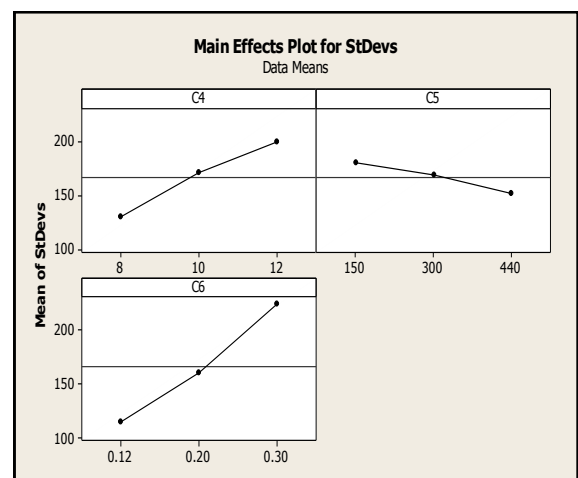


Figure 17: Main Effect Plot for StDevs

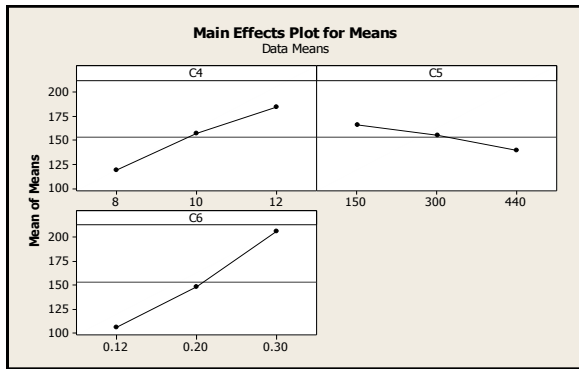


Figure 18: Main Effect Plot for Means

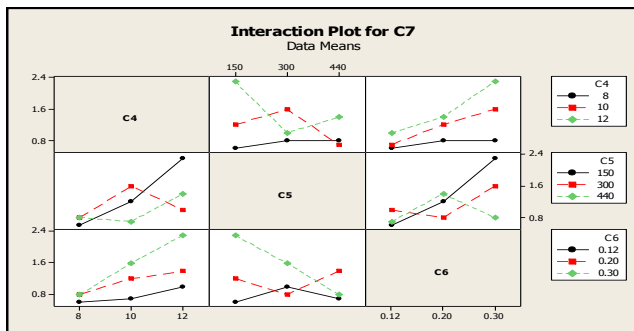


Figure 19: Interaction Plot for Torque

In first row: for minimum torque
If s=150 then d=8; if f=0.12 then d=8
If s=300 then d=8; if f=0.20 then d=8
If s=440 then d=10; if f=0.32 then d=8

In second row: for minimum thrust
If d= 8 then s=150; if f=0.12 then s=150
If d=10 then s=440; if f=0.20 then s=300
If d=12 then s=300; if f=0.30 then s=440

In third row: for minimum thrust
If d=8 then f=0.12; if s=150 then f=0.12
If d=10 then f=0.12 ;if s=300 then f=0.20
If d=12 then f=0.12 ; if s=440 then f=0.12

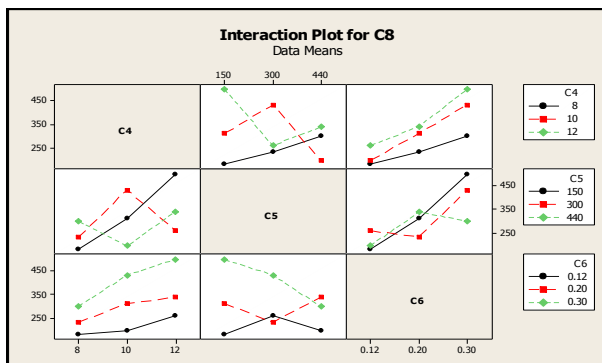


Figure 20: Interaction Plot for Thrust

In first row: for minimum thrust
If s=150 then d=8; if f=0.12 then d=8
If s=300 then d=8; if f=0.20 then d=8
If s=440 then d=10; if f=0.32 then d=8

In second row: for minimum thrust
If d= 8 then s=150; if f=0.12 then s=150
If d=10 then s=440; if f=0.20 then s=300
If d=12 then s=300; if f=0.30 then s=440

In third row: for minimum thrust
If d=8 then f=0.12; if s=150 then f=0.12
If d=10 then f=0.12 ;if s=300 then f=0.20
If d=12 then f=0.12 ; if s=440 then f=0.12

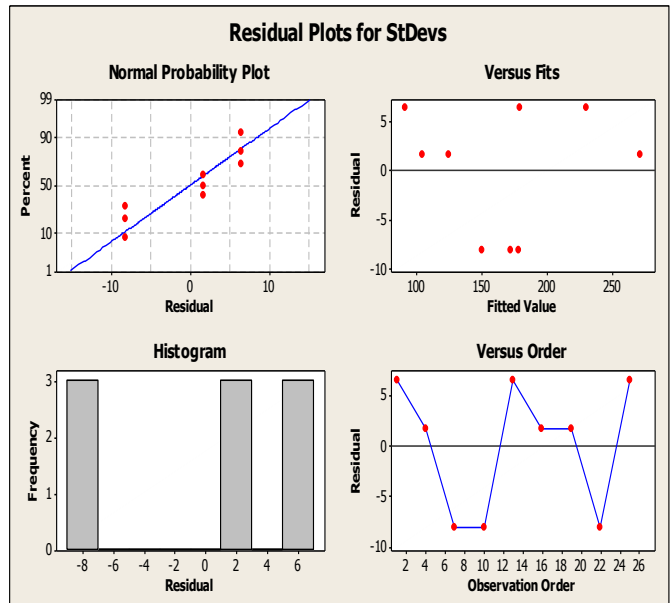


Figure 21: Residuals plot for StDevs

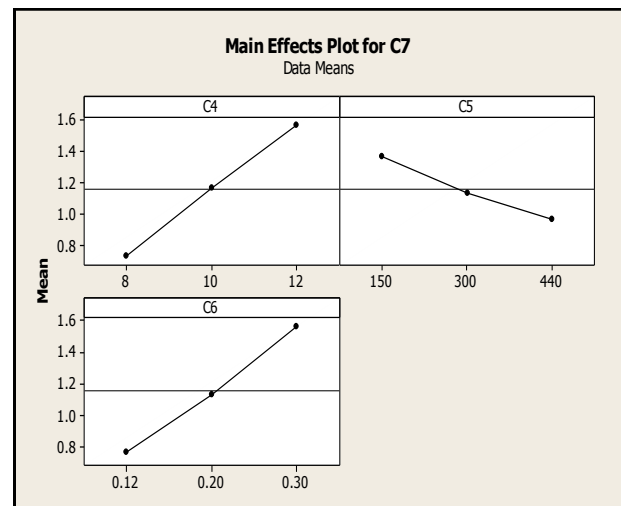


Fig 22: Main effect plot for torque

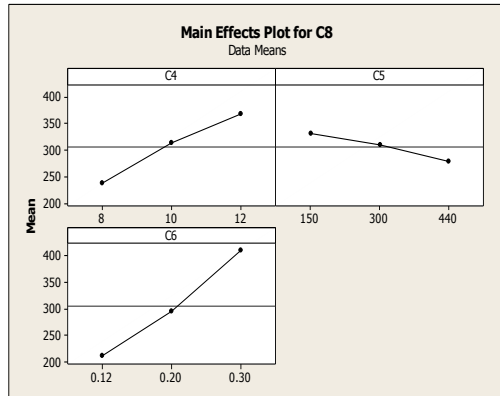


Fig 23 : Main effects plot fir thrust

From the Figure 22 the value of torque increases as the value of drill diameter increases and the relation is linear. The torque decreases as the value of speed of rotation of drill increases. While the torque and feed are nearly directly proportion to each other.

From the Figure 23 the thrust increases with increases in drill diameter and feed but it decreases with increases in speed of rotation of drill.

IV. RESULTS AND DISCUSSIONS

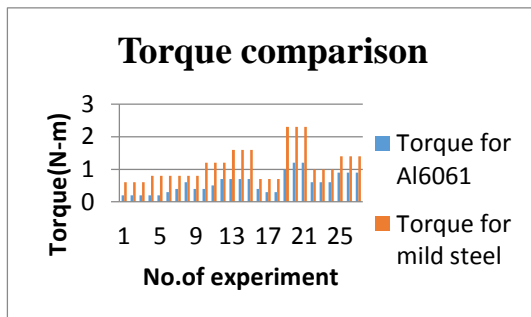


Figure 24: Torque comparison between Al-6061 and Mild steel

We can see from Figure 24 that the value of torque is different for the different set of condition but the value torque is lowest for condition when,

d=8mm
S=150rpm
f=0.12mm

For both the material. But the value of torque is more for mild steel due to having more strength of mild steel.

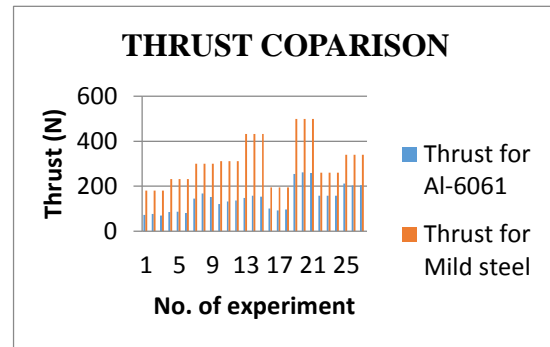


Figure 25: Thrust comparison between Al-6061 and Mild steel

We can see from Figure 25 that the value of torque is different for the different set of condition but the value torque is lowest for condition when

d=8mm
S=150rpm
f=0.12mm

For both the material. But the value of torque is more for mild steel due to having more strength of mild steel.

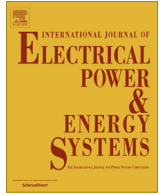
V. CONCLUSIONS

The following are the conclusion drawn based on the experimental investigation conducted at three levels by employing taguchi and anova technique to determine optimum level of process parameters.

- From the main effects plot for aluminium the value of thrust and torque increases with drill diameter and feed but it increases at faster rate with drill diameter. But with increases in drill rotational speed the thrust and torque value decreases and then increases after certain speed.
- From the main effects plot for mild steel the value of thrust and torque increases with drill diameter and feed but it increases at faster rate with feed. But with increases in drill rotational speed the thrust and torque value decreases.
- The Anova revealed that drill diameter and feed are dominant factors followed by drill speed for torque and thrust for aluminium and mild steel as work piece material.
- The optimal combination process parameters for minimum torque and thrust are obtained at 8mm, 150r.p.m and 0.12 mm/rev.
- Taguchi gives systematic simple approach and efficient method for the optimum operating condition.

REFERENCES

- [1] A.Cicek, T. Kivak and G. Samtas, Application of Taguchi Method for Surface Roughness and Roundness Error in Drilling of AISI 316 Stainless Steel, *Journal of Mechanical Engineering*, Vol. 58, 2012, pp. 122-129.
- [2] B. Shivapragash, K. Chandrasekharan, C. Parthasarathy and M. Samuel, Multiple Response Optimizations in Drilling Using Taguchi and Grey Relational Analysis, *International Journal of Modern Engineering Research*, Vol. 3, 2013, pp.765-768.
- [3] C. Tsao and H. Hocheng, Evaluation of thrust force and surface roughness in drilling composite material using Taguchi analysis and neural network, *Journal of Materials Processing Technology*, Vol. 203, 2008, pp. 342-348.
- [4] Ch.Deva Raj Reddy Sreenivasulu, Dr. Ch.Srinivasa Rao, Optimization of operating parameters to minimize burr size in drilling using taguchi method & grey relational analysis for Al 6061 ,*International Conference on Challenges and Opportunities in Mechanical Engineering and Management Studies (ICCOMIM)*, Bonfring/829-833.
- [5] D. Kumar, L.P. Singh and G. Singh, , Operational Modeling For Optimizing Surface roughness in Mild Steel Drilling Using Taguchi Technique, *International Journal of Research in Management*, Vol. 2, 2012, pp. 66-77.
- [6] E. S. Lee, Precision Machining of Glass Fibre Reinforced Plastics with Respect to Tool Characteristics, *The International Journal of Advanced Manufacturing Technology*. Vol. 17, 2011, pp. 791-798.
- [7] E. Ugo. Enemuoh, A. Sherif El-Gizawy, A.Chukwujekwu Okafor “An approach for development of damage-free drilling of carbon fiber reinforced thermosets”, *International Journal of Machine Tools & Manufacture*. Vol. 41, 2001, pp.1795–1814.
- [8] Erkan Bahçe and Cihan Ozel (2013). Experimental Investigation of the Effect of Machining Parameters on the Surface Roughness and the Formation of Built Up Edge (BUE) in the Drilling of Al 5005, *Tribology in Engineering*, Dr. Hasim Pihitli(Ed.), ISBN: 978-953-51-1126-9, InTech, DOI: 10.5772/56027.
- [9] Gupta A, Singari RM (2016). Experimental Analysis of Tool Wear in Drilling of EN-31 Using Artificial Neural Network. *International journal of advanced production and industrial engineering 1(1):1-6*. (9)
- [10] H. Hocheng, C.C. Tsao, Comprehensive analysis of delamination in drilling of composite materials with various drill bits, *Journal of Materials Processing Technology*. Vol.140, 2003, pp. 335–339.
- [11] Irfan Khan, Dheeraj Luthra, Mukesh Verma, Sarabjot Singh, Selection of optimum Drilling Parameter in Drilling of commercial Acrylic sheet to achieve minimum Hole Expansion by using Taguchi Approach, *International Journal of Engineering Science and Technology (IJEST)*, Vol. 4, 2012, pp.1256-1261.
- [12] Jadoun R S, Pradeep Kumar, Mishra B K, Mehta R C S ,Parametric optimization for out-of-roundness of holes in ultrasonic drilling of engineering ceramics using the Taguchi method, *INCARF 2003*, IIT Delhi.
- [13] Jose Mathew, N Ramakrishnan, N K Naik, Investigations into the Effect of Geometry of a Trepanning Tool on Thrust and Torque During Drilling of GFRP Composites, Vol. 91, 1999, pp.1-11.
- [14] Karnik, S.R., V.N. Gaitonde, J.C. Rubio, A.E. Correia, A.M. Abrao and J.P. Davim, 2008. Delamination analysis in high speed drilling of carbon fiber reinforced plastics (CFRP) using artificial neural network model. *Mater. Des.*, 29: 17681776.DOI:10.1016/j.matdes.2008.03.014.
- [15] Kurt M., Kaynak Y. , Bakır B., Köklü U., Atakök G., Kutlu L., 2009, "Experimental Investigation And Taguchi Optimization For The Effect Of Cutting Parameters On The Drilling Of Al 2024-T4 Alloy With Diamond Like Carbon (DLC) Coated Drills", 5. Sixth International Advanced Technologies Symposium (IATS'09), Karabük, Turkey.
- [16] L.B. Zhang, L.J. Wang, X.Y. Liu, H.W. Zhao, X. Wang,H.Y. Lou Mechanical model for predicting thrust and torque in vibration drilling fibre-reinforced composite materials, *International Journal of Machine Tools and Manufacture* Vol. 41 ,2001 ,pp. 641–657.
- [17] Mohan, N.S., S.M. Kulkarni and A. Ramachandra, 2007. Delamination analysis in drilling process of GFRP composite materials. *J. Mater. Process. Technol.*, 186: 265-271. DOI: 10.1016/j.jmatprotec.2006.12.043.
- [18] Rajiv Chaudhary, Ranganath M S, R. C. Singh, Vipin, “Experimental Investigations and Taguchi Analysis with Drilling Operation: A Review” *International Journal Of Innovation and Scientific Research*. Vol. 13(1), 2015.
- [19] Ranganath M. S, Applications of TAGUCHI Techniques in Turning (AKN Learning, Delhi: Delhi, 2015).
- [20] Shravan Kumar Singh , Optimization Of Axial Force In Drilling Process On Mild Steel By Taguchi's Approach *International E-journal of Mathematics and Engineering*. Vol.162, 2012, pp.1502 – 1508.
- [21] U.A. Khashaba, M.A. Seif and M.A. Elhamid, Drilling analysis of chopped composites, *Composites Part A, Applied Science and Manufacturing*. Vol. 38, 2007, pp. 61-70.
- [22] Y. Tyagi, V. Chaturvedi, and J. Vimal, Parametric Optimization of Drilling Machining Process using Taguchi Design and ANOVA Approach, *Journal of Emerging Technology and Advanced Engineering*, Vol. 2, 2012, pp. 339-347



Photovoltaic-grid hybrid power fed pump drive operation for curbing the intermittency in PV power generation with grid side limited power conditioning



Amritesh Kumar*, Vishal Verma

Department of Electrical Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:

Received 20 August 2015

Received in revised form 11 March 2016

Accepted 12 March 2016

Keywords:

Grid connected PV system

Power conditioning

Photovoltaic (PV)

Maximum power point tracking

Synchronous reference frame (SRF)

Voltage source converter (VSC)

ABSTRACT

Existing pump drive systems are fed from the grid. In case grid fails, it takes time to restart the pump drive, and unnecessarily power is wasted in restarting. Moreover, to reduce the burden on the grid and also to curb the intermittency a hybrid configuration is proposed in which PV system is integrated with grid to cater the pump load operation. Even if the grid fails, the PV system works in islanded mode to cater the pump load at the permissible limit depending on the generation by PV system alone in the absence of storage systems. To establish a bidirectional power flow a voltage source converter (VSC) based grid coupling is considered which also provide limited power conditioning to utilize the capacity of VSC at maximum to solve Grid power quality issues like unbalance current, harmonics, etc. The synchronous reference frame (SRF) method is used for current decomposition for determining the limited compensation to be provided by the grid coupled VSC, based on its unutilized capacity. MATLAB based simulation and experimental results show the efficient working of hybrid configuration in both grid connected mode and islanded mode with limited power conditioning.

© 2016 Elsevier Ltd. All rights reserved.

Introduction

The larger part of the electrical energy is utilized by continuous duty pumps used in industrial, agricultural, and household sectors. In the majority of these pumps are fed by power drawn from the distribution systems, and for isolated cases, these are fed by diesel generation sets. To reduce the dependence on the grid renewable energy source needs to be proposed [1]. Out of the all available renewable sources, solar Photovoltaic (PV) energy is suitable for all aforementioned applications [2].

PV panels, exhibits a typical non-linear I - V and P - V characteristics. This varies with varying irradiance and temperature condition. With the high cost of PV panels, there is a need to extract maximum available power from PV panels to avoid heating of the panel due to circulating current and in turn, enhance the total conversion efficiency. In literature different approaches have been proposed for maximum power point tracking (MPPT) such as the lookup table method, perturb and observe method (P&O), and incremental conductance method (IC) [3–5]. For fast tracking and efficient control author has also suggested fuzzy logic base MPPT controller [6].

Amongst different algorithm P&O is very simple to implement, but suffers from oscillations at the MPP point, and the same is overcome by more advanced algorithm such as an incremental conductance method (IC).

Most of the earlier reported literature advocated the use PV pumping system in standalone mode [7]. In literature DC motor driven pump load control based on artificial neural network have been advocated [8] Though PMDC motors are operationally most suitable due to their direct control with charge controller, but due to the ruggedness and absence of brushes in induction motors (IM) are more widely accepted even with additional DC-AC converter. In literature mainly V/f control and slip optimization techniques are reported for induction motor used in PV based pumping application to realize the increase in efficiency. With the advancement in embedded systems and power electronics the sensorless field oriented control (FOC) is considered suitable most for an IM drive for pump application especially with PV applications [9,10].

In standalone operation DC link capacitor voltage experience variation with change in isolation levels, which puts the IM drive out of FOC, if the change is significant. Thus, for smooth operation of the IM drive in FOC the MPPT must operate with some storage system. Such systems are devoid of expansion of PV infrastructure and cost high due to battery capacity storage system. Since the

* Corresponding author. Mobile: +91 9818209336.

E-mail addresses: amritesh@dce.ac.in (A. Kumar), vishalverma@dce.ac.in (V. Verma).

major part of the industrial, agricultural, and household sectors are connected with the grid the storage can be grid based. In grid connected mode DC link voltage can be maintained easily without any battery source. Excess power from PV unit can be transferred to the grid and similarly deficit may be drawn from the grid. Such PV system is connected to the grid through transformer for providing isolation for voltage source converter (VSC) [11]. In literature many control algorithms for PV grid integration have been proposed [12]. Control derives using instantaneous reference theory has been proposed by author [13], but the algorithm lacks in the estimation of current reference generation due to corrupt PCC point. Modifications and improvements are suggested based on power balance or $p-q$ theory using unit template [14–16]. Even these methods, though are very easy to implement, but prone to malfunction under noisy operation at PCC and there exists no decoupling of real and reactive components [17]. To have decoupled control in literature SRF based control have been proposed. In the current proposed system SRF based Indirect current control of grid coupling is advocated, as it gives better control and dynamics as compare to voltage control and other control discussed above. The control apart from transacting the excess/deficit power to/from the grid, utilizes the remaining available capacity of VSC for compensating power quality (PQ) problems on priority such as reactive power compensation, negative sequence compensation, and harmonic compensation [18,19]. In literature many controls have been discussed for alleviating harmonics problems and power quality problems [20,21]. For estimating different PQ problems in load current, synchronous reference frame (SRF) is used for decomposing load current into positive sequence fundamental real reactive power current, negative sequence fundamental current, and harmonic current [22]. VSC's are operated such that its first priority is the transaction of real power and the remaining capacity can be used for selective compensation of PQ problems based on priority and severity of the problems at the load side [23–25].

The paper presents the technique which evacuates maximum available power from solar PV panels, maintains a near constant DC bus voltage to ease the FOC of IM drive for pump applications, transact excess/deficit power required by the pump load and provides limited compensation to power quality problems (PQ) on the distribution grid. Apart from utilizing the VSC resource through PQ compensation, the proposed system also helps in reducing the total cost of operation as pump load operates in continuous duty cycle without need to shut down the pump even if the grid fails due to hybrid configuration. In the proposed scheme indirect vector control scheme is used for IM drive in grid connected mode. The effectiveness of the proposed system is evaluated through simulation under grid connected mode and varying insolation conditions. Moreover, to track the fast changing insolation condition an effective incremental conductance (IC) MPPT method is used for maximum evacuation of power from the PV panel.

The proposed hybrid photovoltaic pumping system with power conditioning capability (HPVPSPCC) is simulated under the MATLAB simulink environment. The contents are dealt in the following sections: (II) System architecture (III) PV array modeling, (IV) IC method based MPPT techniques, (V) Control, (VI) MATLAB based simulation, (VII) Performance evaluation (VIII) Hardware implementation.

System configuration

Fig. 1 shows the block diagram of HPVPSPCC comprising of indirect vector controlled induction motor (IM) pumping system connected across the tank capacitor (C_r) which is fed by both PV generation and grid in grid tied mode and PV generation in off grid mode. The grid based loads are varying types and in the present case, these are balanced RL load, unbalanced R load, thyristor converter type load, etc. PV system consists of PV panels, DC-DC boost converter with MPPT operation. A large tank capacitor (C_r) is

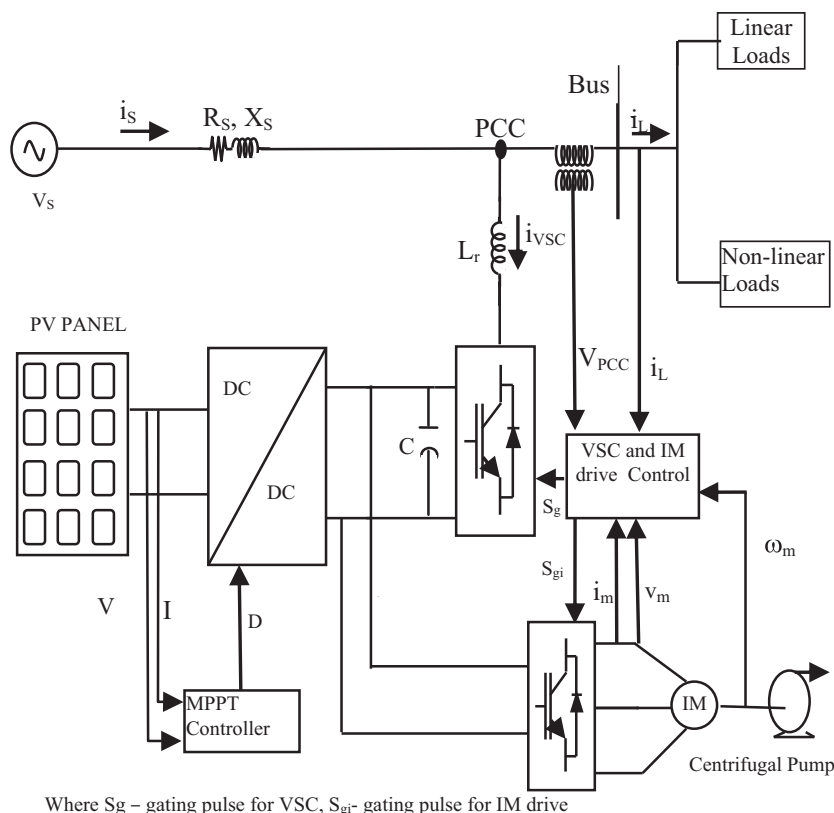


Fig. 1. Block diagram for system configuration of HPVPSPCC.

connected across the DC-link to make its stiff so as to accommodate small variation in solar insolation without changing operating condition. Such stiff condition will make the IM pumping system works efficiently and uninterruptedly with limited demand from the grid. Indirect Sensorless vector control technique is being employed for IM drive operation by taking feedback of speed to generate the required reference electromagnetic torque (T_e) for lifting the load demanded from centrifugal pump. Hysteresis current control based strategy is used here to force the reference current generated by torque reference (I_q^*) and flux (I_d^*) current components.

The proposed SRF based reference current generation based on load current decomposition is integrated with prioritized and selective compensation of PQ problems based on the available capacity of the VSC not utilized for real power transfer. The indirect control strategy is being used with VSC to shape up the source current according to reference current generated through SRF by passing it through hysteresis controller. To filter out higher order harmonics LCL filter is used at the AC side terminal.

Modeling of PV arrays

Basic equations that govern the I - V characteristics of an ideal PV cell is represented in literature [26]:

$$I = I_{pv,cell} - I_{o,cell} \left[\exp \left(\frac{qV}{akT} \right) - 1 \right] \quad (1)$$

where $I_{pv,cell}$ is the current generated by the incident light (changing with varying isolation), I_d represents current in the Shockley diode, $I_{o,cell}$ is the reverse saturation current of the diode, k is the Boltzmann constant, T is the temperature of the p - n junction in Kelvin and ' a ' is the ideality factor of the diode. But in actual scenario many such single units are connected in series and parallel to form the panels. These panels are then arranged again in series and parallel to form the array which suits the power required. The interconnection of panels encounters contact resistance (R_s), across which there will be drop in voltage as shown in Fig. 2. Thus the typical equation considering all the losses is given by:

$$I = I_{pv} - I_0 \left[\exp \left(\frac{V + R_s I}{V_t a} \right) - 1 \right] - \frac{V + R_s I}{R_p} \quad (2)$$

where $V_t = \frac{N_s k T}{q}$, N_s – Number of cells connected in series; R_p = Equivalent parallel Resistance; R_s = Equivalent contact series Resistance.

Commercially available Kyocera make (KC200GT) PV panels are considered for the modeling of PV array to get the output of power 46 kW with V_{mp} (array) 564 V and I_{mp} (array) 81.5 A at 1000 W/m² by connecting 11 strings in parallel having 21 panels in series in a string. The parameters of the panel are shown in Table 1. The P - V characteristic of the aforesaid array is depicted in Fig. 3 at irradiance level of 1000 W/m², 800 W/m², 600 W/m² which represents the maximum power at the corresponding voltage level.

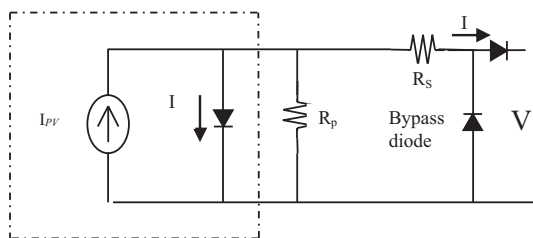


Fig. 2. PV panel model.

Table 1

Parameters of the KYOCERA PV panel (KC200GT) at nominal operating conditions.

Number of parallel connected PV string	11
Number of series connected PV module	21
Number of series connected cell (N_s)	54
Open circuit voltage of module (VOC)	32.9 V
Short Circuit Current (I_{sc})	8.21 A
Maximum Power (P_{max})	200.13 W
Maximum Power Voltage (V_{mpp})	26.3 V
Maximum Power Current (I_{mpp})	7.61 A
Temperature Coefficient of V_{oc}	-0.1230 V/K
Temperature Coefficient of I_{sc}	0.0032 A/K
R_p	415.405 Ω
R_s	0.221 Ω

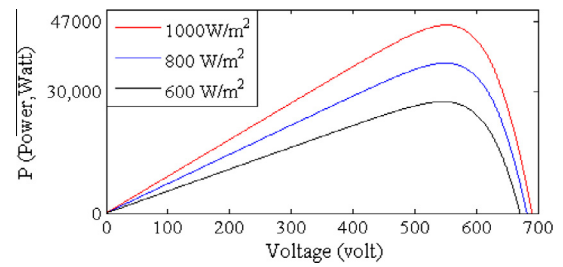


Fig. 3. P - V characteristic.

Control theory

A. MPPT controller based on boost converter

The MPPT controller based on Incremental Conductance (IC) operates to draw maximum power from PV array. The said algorithm may be implemented either by voltage or duty cycle perturbation. But duty cycle based MPPT controller gives better dynamics and fast control as compare to voltage perturbation control. At maximum power point $\frac{dP}{dV} = 0$, this can be expressed as:

$$\frac{dP}{dV} = \frac{d(VI)}{dV} = I + V \frac{dI}{dV} \quad (a)$$

where V is the PV array voltage, I is the PV array output current as shown in Fig. 1.

$$\text{At maximum power point } \frac{dP}{dV} = 0; \quad \text{i.e. } \frac{dI}{dV} = -\frac{I}{V} \quad (b)$$

To arrive at the MPPT according to Eq. (b), voltage and current is sensed at the output terminal of solar array and data is processed through IC based controller to give output command in the form of duty cycle which is being used to give gate pulse to boost converter as shown in Fig. 4 converter to make the operation of the solar array at the MPP point.

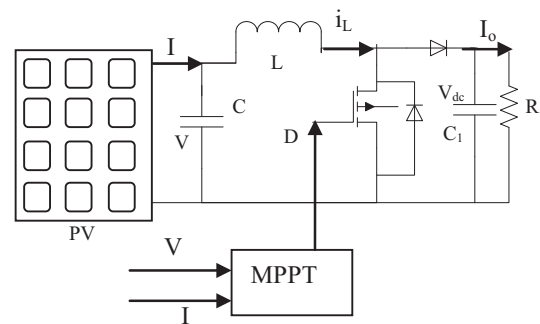


Fig. 4. Boost converter governing through MPPT controller.

$$V_L = V - V_{dc} = L \frac{di_L}{dt}, \quad (c)$$

where V_L is the voltage across the inductor

$$\frac{di_L}{dt} = \frac{V - V_{dc}}{L} \quad (d)$$

Since the output voltage of the converter (V_{dc}) is maintained through grid side so output current may derived as:

$$I_o = I(1 - D), \quad (e)$$

where I_o is the boost converter output current and I is the PV output current.

For arriving at continuous operation of the converter the inductor (L) value is kept well above L_{min} as:

$$L_{min} = \frac{DV_{dc}(1 - D)^2}{2f_o} \quad (f)$$

where V_{dc} is the dc link voltage maintained by the grid (800 V), f is the switching frequency, D is the duty cycle, I_o is the output current.

Eqs. (a)–(f) are used to design the L and C_1 of the DC-DC converter as shown in Fig. 4. For calculating the value of L_{min} , assumption is made that insolation level changes between 1000 W/m² to 600 W/m² only. With this assumption at full insolation level I_o is 57.5 A (as V_{dc} is maintained constant at 800 V by grid side VSC) and I is 81.5 A. Using Eqs. (e) and (f), D and L_{min} is calculated as 0.3 and 0.22 mH (switching frequency 5 kHz). Similarly for 600 W/m², D and L_{min} is calculated as 0.4 and 0.34 mH. Since for both extreme of insolation level current has to be continuous, the higher value of L_{min} is considered for designing the inductor. To have small ripple in inductor current inductor value is chosen 1.5 times that of L_{min} (0.34 mH). So L is design for 0.5 mH.

As the DC link voltage is maintained by the grid side VSC, so the DC link capacitor (C_1) of VSC is designed on the basis of allowable change on the dc bus voltage, and change in supply current (i_{VSC}). Hence, is

$$C_1 = \frac{3L_r(i_1^2 - i_2^2)}{2\alpha V_{dc}^2},$$

where L_r is the interface inductor, “ α ” represents the maximum ripple allowed in dc bus voltage. With allowable voltage ripple as 1%, at 800 V and current change from 25 A to 60 A, the computed value of capacitor is 2091, which is taken as 2200 μ F due to practical availability.

The IC based MPPT controller assign the duty cycle to the DC-DC boost converter so that at every instance power should be at MPP tracked point. Since the operation of the pumping system is in conjunction with grid, the excess/deficit power not being utilized by the IM pumping system will be transacted with the grid through VSC converter. Voltage level of DC bus is maintained in a way that proper vector controlled drive operation is assured along with adequate power conditioning offered to the grid connected load. Indirect vector control technique is used in the proposed IM pumping system, as it provides better flexibility in the controlling by decoupling both torque and flux component separately [27].

The tracking of MPP by IC method yields a typical duty ratio, which extracts the power from the PV panel by acting as derived current source. When the DC bus is kept constant across the tank, it pertains to transfer of all the energy extracted from the PV array to mains; any displacement from equilibrium position would reflect the movement away from MPP position. It is therefore any undesired oscillations are avoided on DC bus, by carefully selecting the tank capacitor and control. Further an algorithm capable of decomposing the current into various components as aforesaid is realized using SRF theory, based on compensation

looking at the available unused capacity of VSC and curbing the disturbances/oscillations on the DC bus. The following section deals with the algorithm for current decomposition, calculation of safe unused capacity of rating of VSC, scaling of compensation gain based on priority and DC bus control. Using SRF theory, load current is decomposed into positive sequence fundamental frequency active and reactive power current, harmonic current and fundamental frequency negative sequence component using technique [25]. To get positive fundamental part load current reference frame is being rotated at fundamental frequency ω_1 . This decomposition gives DC quantities for fundamental frequency component (I_d, I_q) and also harmonic components rotating at frequency ω_1 less than actual harmonic frequency. This decomposed current is passed through low pass filter (LPF) to get current corresponding to fundamental component only. Similarly for negative sequence load current reference is being rotated at fundamental frequency $-\omega_1$. This gives negative sequence real and reactive component after passing through LPF (I_{dn}, I_{qn}). Following mathematical transformation is used for transforming α - β to d - q component at fundamental frequency ω_1 :

$$\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} \cos(\omega_1 t) & \sin(\omega_1 t) \\ -\sin(\omega_1 t) & \cos(\omega_1 t) \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} \quad (3)$$

Similarly for extracting DC component of negative sequence current the SRF axis is rotated in negative direction at fundamental frequency using equation:

$$\begin{bmatrix} i_{d-} \\ i_{q-} \end{bmatrix} = \begin{bmatrix} \cos(\omega_1 t) & -\sin(\omega_1 t) \\ -\sin(\omega_1 t) & \cos(\omega_1 t) \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} \quad (4)$$

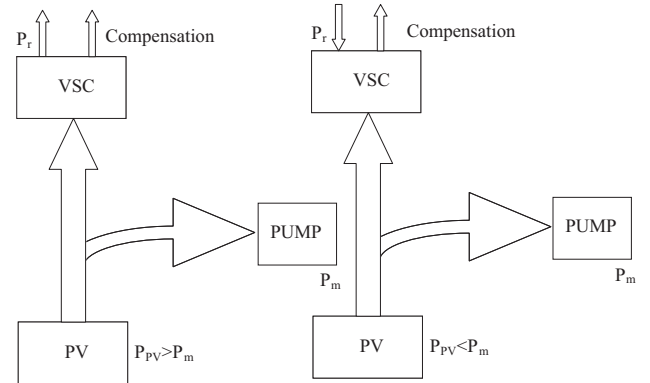


Fig. 5. (a) Power flow diagram with excess power generated. (b) Power flow diagram with deficient power generation.

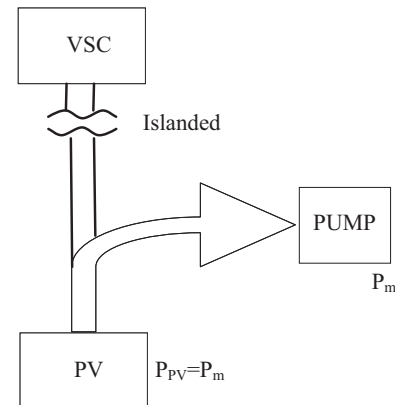


Fig. 6. Power flow diagram in islanded mode.

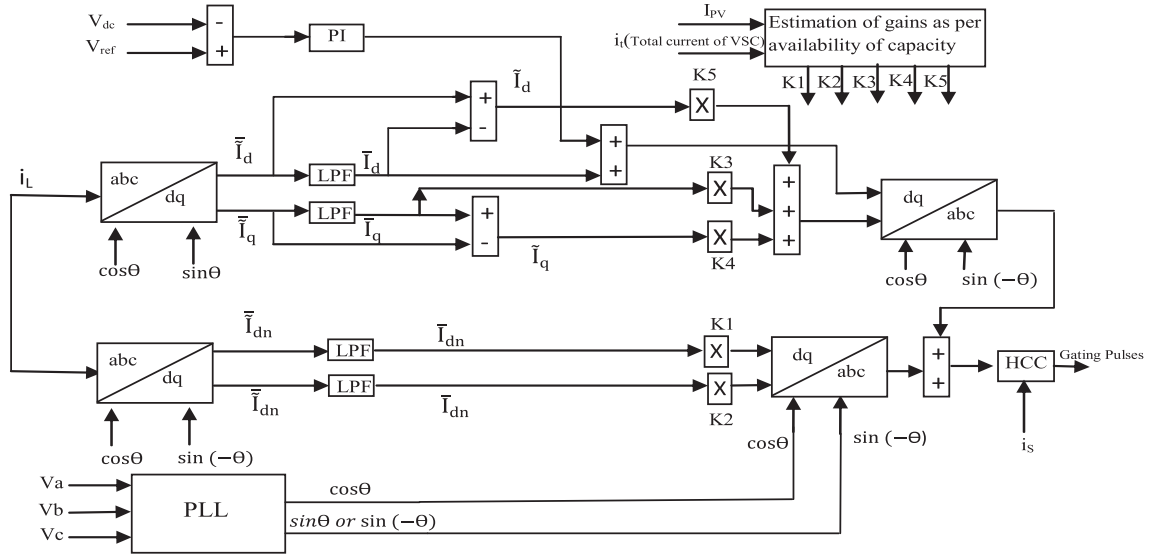


Fig. 7. Main blocks for the proposed control scheme.

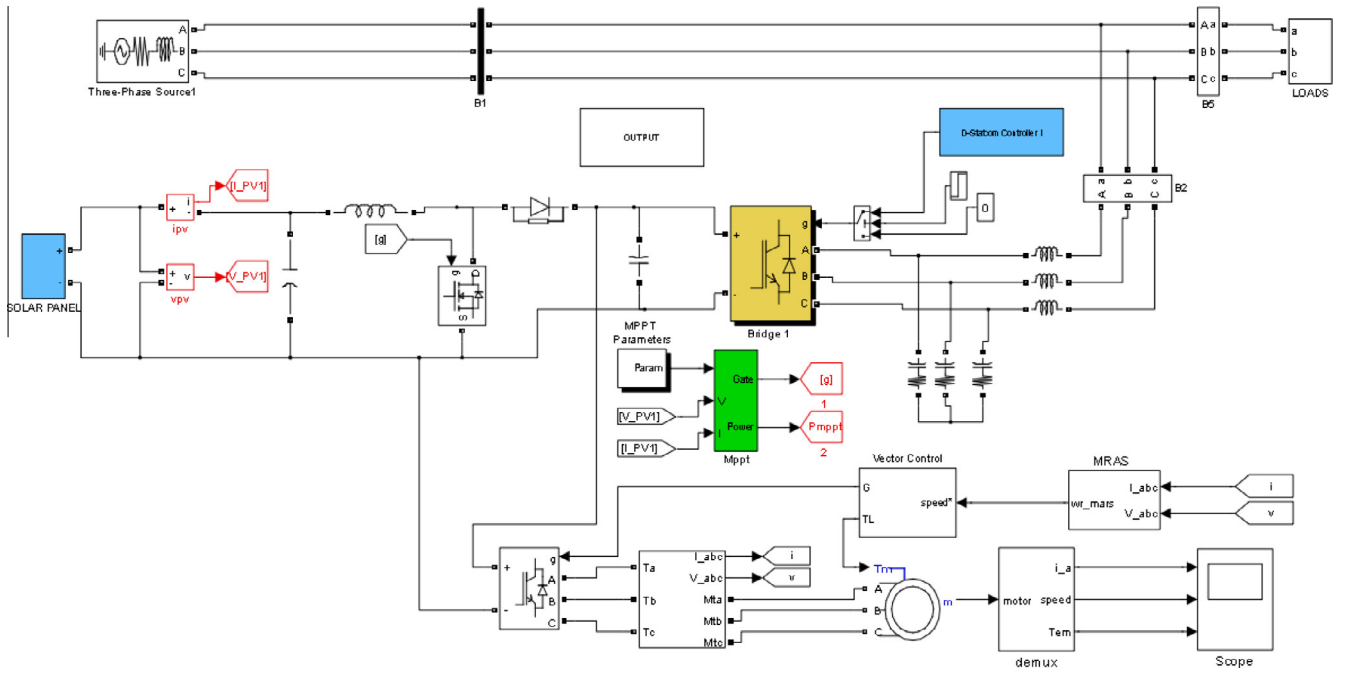


Fig. 8. Matlab simulation diagram.

The reference currents for the indirect current control of VSC are derived by computing the compensation gain of decomposed current components other than positive sequence fundamental frequency component corresponding to real power. These gains are adjusted on the basis of priority and unused capacity of VSC. This priority based limited compensation of PQ disturbances circumvent the overloading of power Converter and avoid oscillation on DC bus and ensure evacuation of maximum power from the PV array. Priority assignment is situation based and multi objective hence is not being included in present scope. However, in order to demonstrate clearly the limited and selective compensation priorities have been scaled based on the type of PQ disturbances. Harmonics, negative sequence and reactive power compensation have been assigned in priority in decreasing order. The PQ disturbances which gets full compensation assign the gain of '0' and the one which

reward uncompensated assumes the gain of '1'. Based on the limited capacity compensation of VSC gains are adjusted to compensate the PQ disturbances in the order of priority.

The compensation provided must assume that;

$$I - I_{PV} - (I_h + I_r + I_n) * k \geq 0 \quad (5)$$

where $I_h = \sqrt{(I_{dh}^2 + I_{qh}^2)}$ - Harmonic current, $I_n = \sqrt{(I_{dn}^2 + I_{qn}^2)}$ - negative sequence current and I_r - reactive current component, I - Full rated capacity of VSC.

DC bus voltage is controlled through reference current generated after passing the error in voltage at the DC bus through PI controller. SRF based theory is used to decompose the load current into positive fundamental real and reactive current component, negative fundamental component and harmonic current. Indirect

control scheme is incorporated to generate the reference current for source based on selective and priority based compensation. Selection for the particular PQ problem is based on the available capacity of the VSC converter and severity of the PQ problem. In this paper harmonic compensation, unbalanced compensation and reactive power compensation has been prioritize in the decreasing order. After estimating the available capacity of the VSC various gains (K_1, K_2, K_3) are estimated to generate the reference source current through gain resolver block as shown in Fig. 7.

PWM gating pulses are generated by comparing reference currents $i_{sa}^*, i_{sb}^*, i_{sc}^*$ and source current in the hysteresis controller block (HCC). HCC band is selected depending on the IGBT switching capability used in the VSC. The generated gating pulses are passed through a D flip flop to limit the frequency more than that a critical value to avoid short through and system entering limit cycles. The complete control scheme is shown in the Fig. 7.

B. The control HPVPSPCC unit under 3 different cases as:

Case I. When the PV power generation is more than the demanded pump load as shown in Fig. 5(a). In this case excess power ($P_{pv}-P_m$) is transferred to the grid through grid VSC which also functions to eliminate PQ problems according to the capacity of the VSC. The excess power is being transferred to the grid by sensing the DC link voltage. In this case because of excess power generation DC link voltage swells, that is detected by the SRF based control algorithm in which I_d^* reference is generated by passing the error in voltage through PI controller.

Case II. With demand load power more than generation, the deficient power ($P_{pv}-P_m$) as shown in Fig. 5(b) is taken from the grid source to make the drive operation uninterruptable. This deficiency in power is detected by sag in voltage of the DC link.

Case III. In this case suppose grid fails, then HPVPSPCC unit will start its operation in islanded mode after disconnecting itself from the main grid through proper island detection as shown in Fig. 6. Now pump load is fed through only PV source, so load torque command has to be given according to the PV power generation to avoid swelling of the DC bus till its maximum torque. If the PV generation is more than what maximum load torque that motor can handle additional battery storage may be required. As it is pump load torque command has to be set according to following equations:

$$T_L = K\omega_m^2; \text{ keeping Full load torque} = 239 \text{ N/m}^2, \omega_m = 1480 \text{ RPM}, K = 0.009989, \text{ So } P_{\text{shaft}} = T\omega_m = \frac{T^3/2}{K^{1/2}}.$$

Matlab-based simulation

Induction motor drive with pump load, PV array with MPPT based DC-DC converter, composite load on the grid (Linear and non-linear load), grid source and VSC based power conditioning units are simulated under the MATLAB/Simulink environment and output waveforms are collected in the block named 'OUTPUT' for evaluation of performance as shown in Fig. 8. PV panels are connected in series and parallel to deliver maximum power of 46 kW through IC based MPPT charge controller. The command code for the same is written in embedded MATLAB function. The PV pumping system response has been tested under fast changing insolation conditions, varying load torque conditions in grid tied and the off grid modes. Timer block is used to vary the load torque

Table 2

Utility grid parameters.

Description	Parameters
Main supply voltage	415 V
DC link voltage	800 V
Line frequency	50 Hz
DC bus capacitance (μF)	470 μF
Line impedance (Z_s)	$R_s = 0.4 \Omega, X_s = 1 \text{ mH}$

Table 3

Induction motor parameters.

Description	Parameters
Rated motor voltage	400 V
Rated power	37 kW
Rated speed	1480 RPM
Magnetization inductance (L_m)	0.02711 mH
Rotor inductance (L_r)	0.0278 mH
Rotor resistance (R_r)	0.05 Ω
Stator resistance (R_s)	0.082 Ω
Stator inductance (L_s)	0.0278 mH

command from rated 240 N/m² to 200 N/m² at different instant of time. The pump load reference for the torque indirectly generates the reference for the speed command. Stator input of the IM is fed through inverter. The detailed parameters of utility grid and induction motor are shown in Tables 2 and 3 respectively.

This paper also covers the aspect of effective utilization of grid tied VSC to compensate for PQ problems with an available capacity of VSC, raised by composite load, viz. reactive power, and harmonic and unbalance currents. Harmonics are generated by thyristor converter fed load ($R = 5 \Omega, L = 15 \text{ mH}$ on DC side), unbalanced is created through unbalance load ($R_a = 5 \Omega, L_a = 15 \text{ mH}; R_b = 10 \Omega, L_b = 15 \text{ mH}; R_c = 10 \Omega, L_c = 15 \text{ mH}$). Grid source is assumed to be infinite and line parameters are modeled by series impedance equivalent to short circuit impedance. The proposed HPVPSPCC is simulated under varying insolation, and varying torque conditions under both grid tied and off grid mode with limited VSC capacity.

Performance evaluation

The complete HPVPSPCC model with photovoltaic and grid as hybrid source and pump load as aforesaid is simulated under MATLAB simulink environment for grid tied and in off grid mode. First, effectiveness of MPPT algorithm is evaluated under fast changing insolation conditions by connecting the output of MPPT charge controller to a resistive load separately. The output waveforms are shown in Fig. 9. To evaluate the performance of the proposed system both Grid tied VSC and PV system is started to operate to make the DC link capacitor voltage to 800 V, and is kept constant by pushing the surplus energy generated by the PV system to the grid. The evaluation of the proposed system is considered in three stages so as to present the result in lucid way. In the first case torque is kept constant and solar power is varied. Besides this the system is shown to utilize the available capacity for VSC for transfer of power whereas in second case insolation is kept same and demanded load torque is varied to demonstrate the effectiveness of the converter to manage the deficit power from the mains. The third case deals with islanded condition, where the available power from the PV system decides the speed, thereby the torque afford by the pump.

Constant torque operation

At $t = 0.2 \text{ s}$ as shown in Fig. 9(i-j), IM drive is started at no load and the speed command is kept at zero to facilitate the development

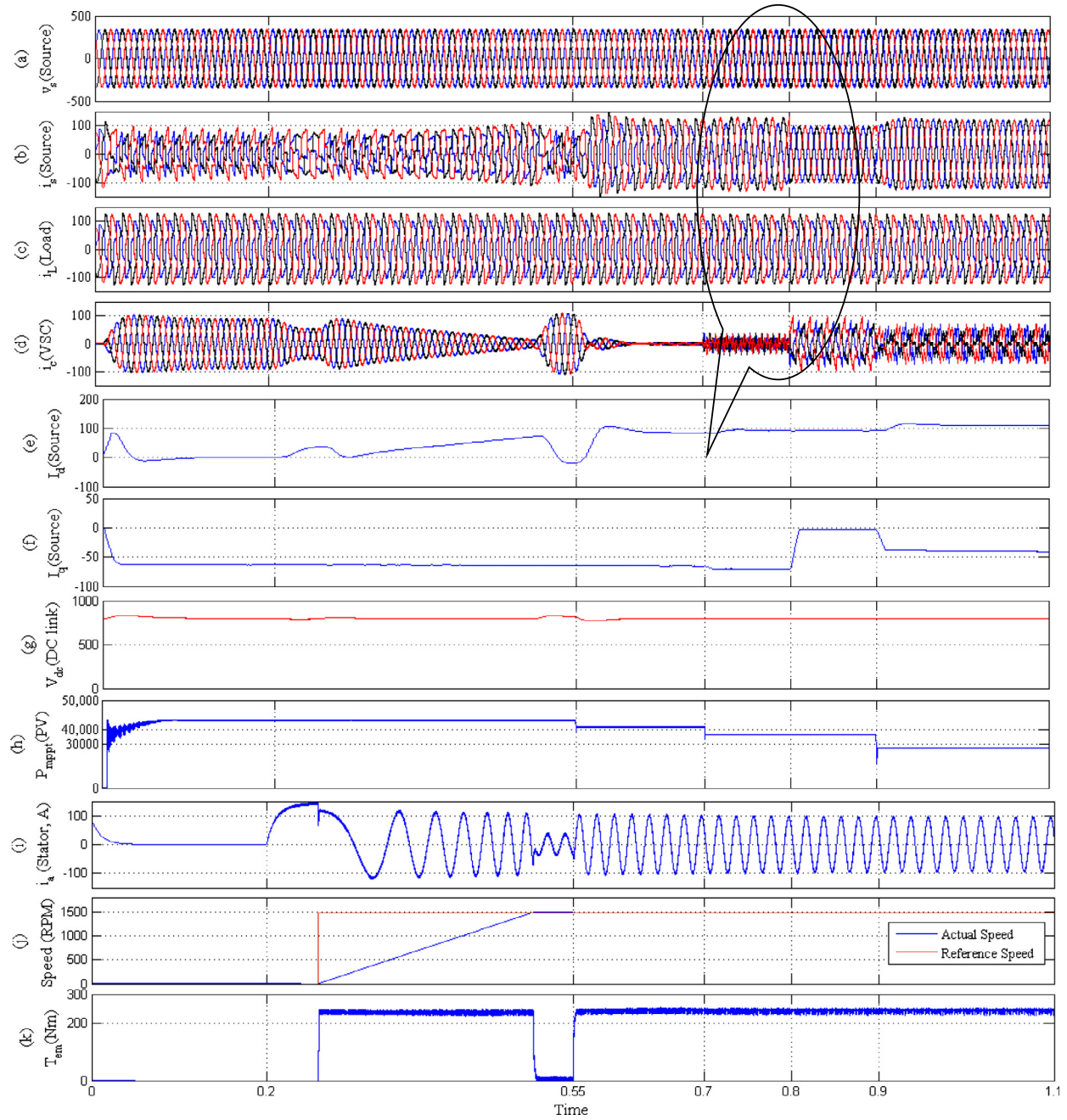


Fig. 9. Dynamic response of HPVSPCC under constant torque mode. IM comes in operation from $t = 0.2$ s onwards. From $t = 0.7$ onwards varying compensation is being provided depending on availability and capacity of VSC.

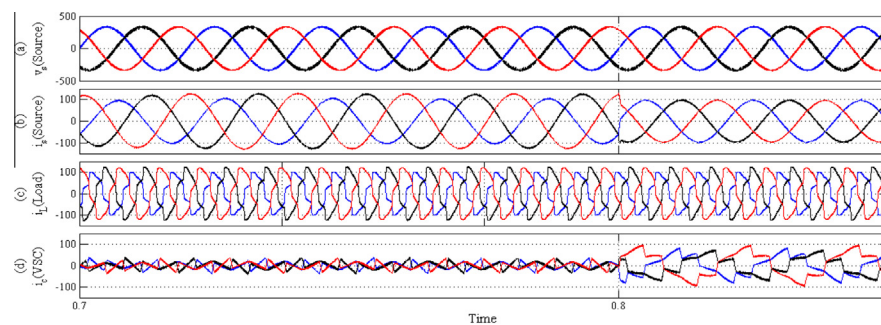


Fig. 10. Zoomed waveform of HPVSPCC under constant torque mode.

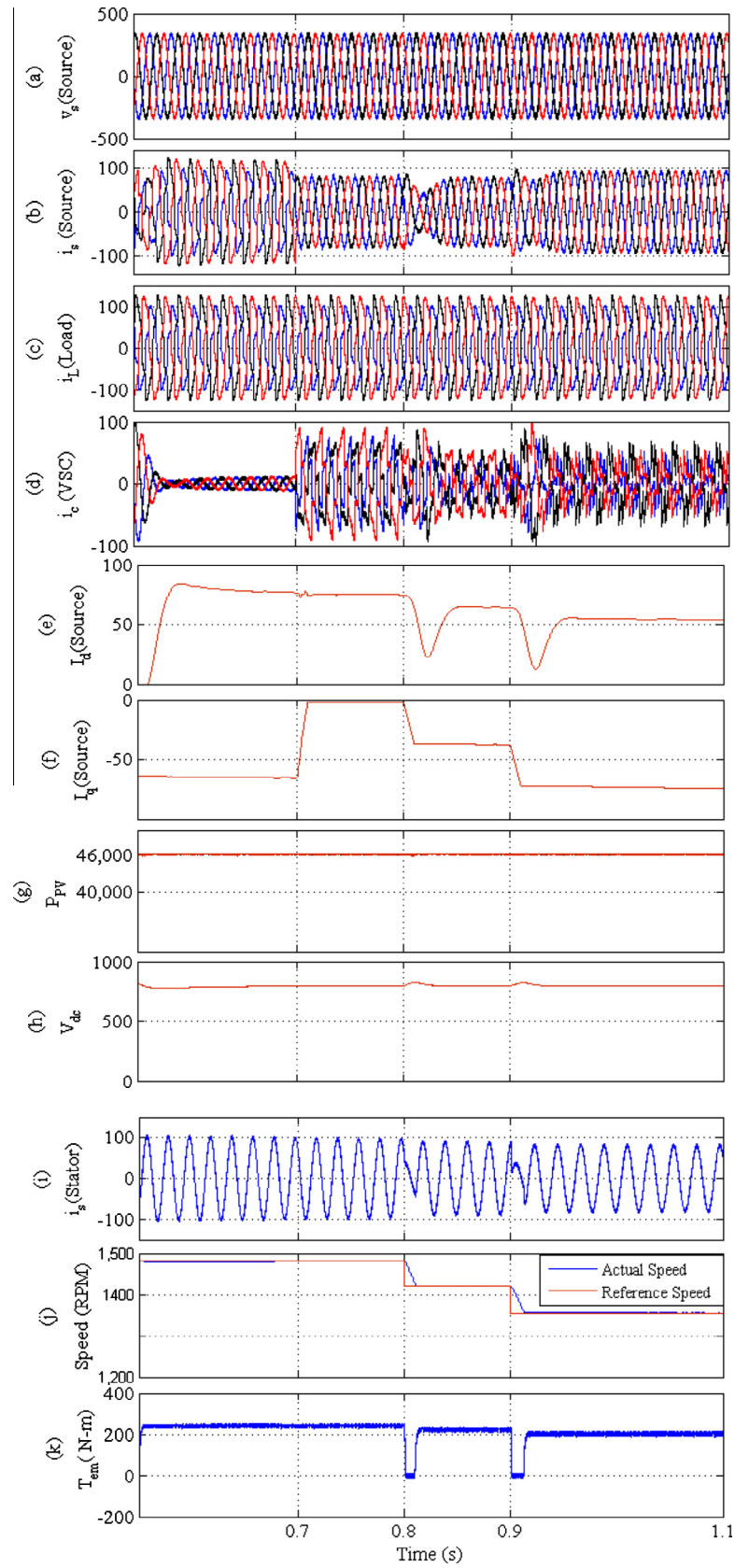


Fig. 11. Dynamic response of HPVPSPCC under constant insolation mode.

of requisite flux. Once a level of flux (0.7 Wb) is reached the vector controller is started. It may be observed from Fig. 9(d), that due to drawl of very low frequency current by stator to enable development of flux, the power returned to the grid reduces from 63 A to 35 A toward $t = 0.26$ s. Further it may be seen that there is marked increase in real current and as a result the current signature changes. The moment the vector controller commences its operation at $t = 0.26$ s, the speed of the pump motor religiously follows the reference presented by the controller as shown in Fig. 9(j). Further Fig. 9(i) reveals that during $t = 0.26$ s to $t = 0.45$ s when the drive is accelerating in vector control mode the frequency of the current waveform gradually increases from low frequency to 50 Hz. Toward the $t = 0.45$ s, where it attains the rated speed. After reaching the rated speed T_{em} comes to the zero level, this results into decrease in stator current shown in Fig. 9(i) and increase in the current supplied through VSC as shown in Fig. 9(d) equal in magnitude as during $t = 0.2$ s to $t = 0.26$ s. It may also be observed from Fig. 9(k) that the drive accelerated at constant peak torque of 300 N m for fastest acceleration. During this period since T_{em} remained constant, the rise in speed corroborates to rise in power linearly; the current supplied to grid accordingly droops as shown in Fig. 9(d). As a result current drawn from the grid rises with same rate to cater to the demand of the connected load, and when rated speed is reached and surplus power is supplied by the VSC, the real power component in the source current drops drastically during $t = 0.45$ s to 0.55 s (refer Fig. 9(b)). At $t = 0.55$ s when full pump load torque of 239 N m is applied and insolation level is decreased to 900 W/m² which is typically sufficiently enough to cater to the need of the load. It may be observed from Fig. 9(i) that the stator current is increased to 70 A. Further the source current continued to maintain the same signature throughout the duration of $t = 0.55$ s to $t = 0.7$ s, during this time PQ compensation is deliberately not shown to observe clarity in results. At 0.7 s insolation level has further gone down to 800 W/m² still giving room to the VSC for power conditioning. It may be observed that source current became pure sinusoidal but with some unbalancing when harmonic compensation is provided by the VSC as evident from Fig. 9(b). At $t = 0.8$ s source current becomes pure sinusoidal with unity power factor when comprehensive compensation viz. harmonic, reactive and negative sequence component is provided by VSC as shown in zoomed Fig. 10. It is also evident that reactive component of source current, I_q (Source) became 0 at $t = 0.8$ s. It may be noted that there are no traces of any unbalance or harmonic in current waveform (refer Fig. 9(b)). At 0.9 s when insolation again drop down to 600 W/m², generating only 26 kW for PV system as against demanded 37 kW, the remaining power is supplied by the grid marked by increase in source current as evident in Fig. 9(b).

Constant insolation mode

Transient operation remains the same as aforesaid till 0.55 s, whole operation is been shown in Fig. 11(i–k) from $t = 0.55$ s to $t = 1.1$ s. At $t = 0.55$ s when full pump load torque of 239 N m is applied, it may be observed from Fig. 11(i) that the stator current is increased to 63 A. At $t = 0.7$ s with load torque remains same, capacity of VSC is utilized for full compensation resulting into sinusoidal unity power factor (UPF) operation of source current. UPF operation is also evident from Fig. 11(f) where i_q becomes 0 during $t = 0.7$ s to $t = 0.8$ s. At $t = 0.8$ s load torque has decreased to 200 N m which results into decrease in stator current to 62 A as observed from Fig. 11(i), moreover there is also decrease in the I_d (Source) current as more power now fed to the grid due to less demand from the pump load. At 0.8 s comprehensive compensation, viz. harmonics and negative sequence compensation command is given to VSC except half of the reactive power till 0.9 s

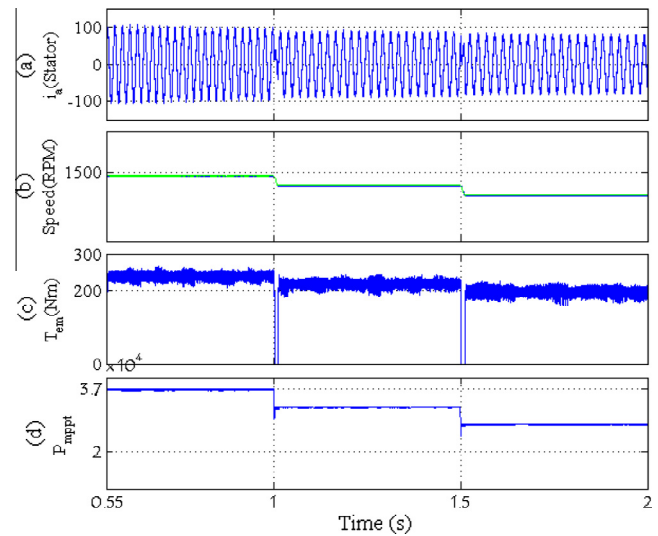


Fig. 12. Dynamic response of HPVPSPCC under islanded mode.

due to increase in real power transaction with grid. At $t = 0.9$ s there is further decrease in the torque from 200 N m to 180 N m, resulting into decrease in the speed in also as evident from Fig. 11 (j). At the same time reactive power compensation control has been stopped resulting into further increase in source current as evident from Fig. 11(b).

Islanded mode

Assuming that grid fails at 0.55 s and the PV output power is just sufficient to cater the full load pump load. In this mode of operation as discussed in control scheme torque command is being set by properly estimating the PV power generation to avoid any swelling of DC link, so that there should be perfect match between power generation and power demand. At 1 s PV power has decreased 32 kW as shown in Fig. 12(d) resulting into decrease in T_{em} from 239 N m to 210 N m and same follows with speed. This

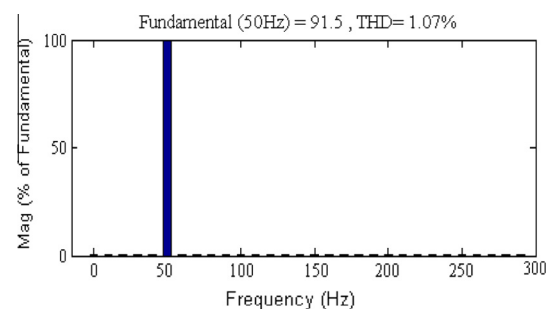


Fig. 13. Screen shot of THD of source current harmonic.

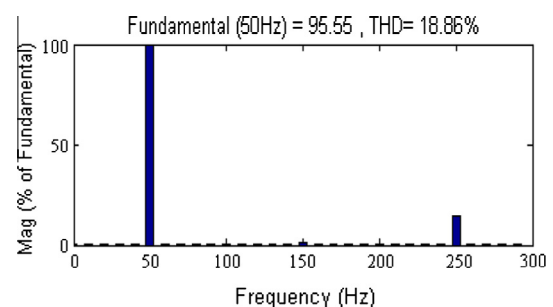


Fig. 14. Screen shot of THD of load current harmonics.

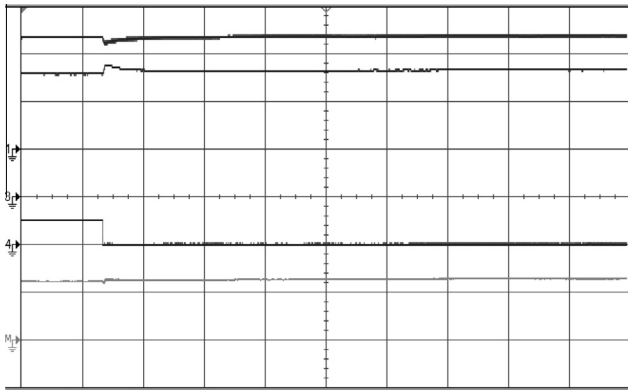


Fig. 15. Experimental result showing PV current (I_{PV} , 3 A/div), and DC link voltage ($V_{DC-link}$, 115 V/div), net drive loading on DC link, PV power (P_{PV} , 870 W/div).

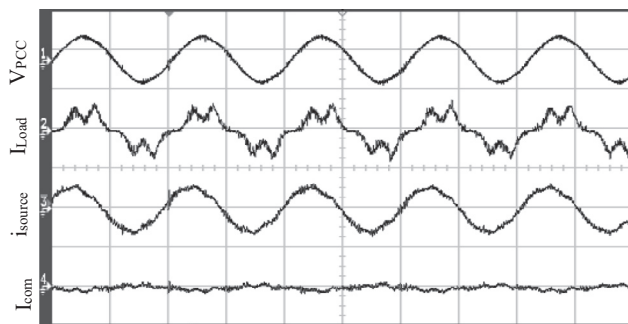


Fig. 16. Measured source voltage (V_{PCC} , 190 V/div), load current (I_{Load}), source current (I_{source} , 10 A/div) and compensating current (I_{comp}).

mode of operation helps the pump to work on continuous duty cycle without even need to shut down fully in case of grid failure.

Further Figs. 13 and 14 show the THD of source current and load current. Despite being 16% THD in load current, the THD of source current remains at 1.07% as shown in Fig. 13.

Hardware implementation

The scaled down hardware prototype of HPVPSPCC is developed for evaluation of the control algorithms for meeting the outlined objectives of the present work is shown in Fig. 19. The development of hardware prototype includes fabrication of power circuits, sensor interfaces and synchronizing switching for the inverters. The LEM LV25P and ABB EL 25 P1, hall effect voltage and current sensors are employed for sensing the voltage at PCC, DC link of inverter, source and load side current of grid. Three phase VSC

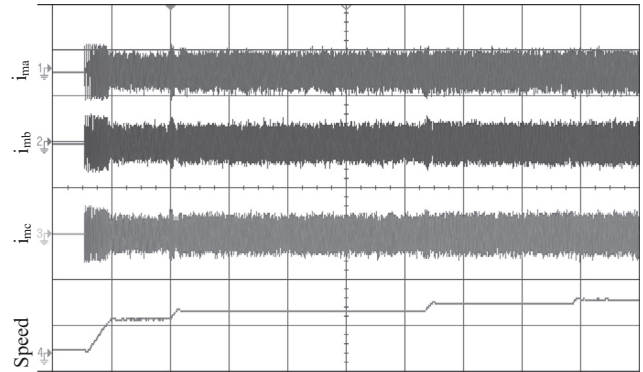


Fig. 17. Experimental results showing motor current i_{ma} (5 A/div), i_{mb} , i_{mc} and speed (1000 rpm/div).

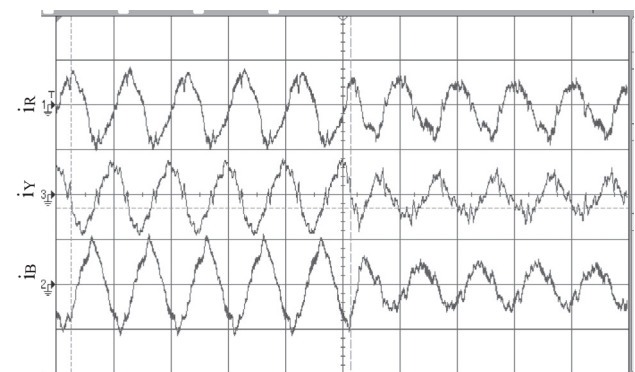


Fig. 18. Measured VSC current (i_R , i_Y , i_B) with change in insolation.

converter is formed by connecting 3 no. of IGBT modules in parallel (each module contains 2 no. IGBT connected in series as high and low side switch). Hitachi make 3 phase 200 V, 1 HP, motor is used for centrifugal pump. dSPACE 1104 real time controller is used for knitting the control of both grid side inverter and the vector controlled induction motor drive. Gating pulses are generated both for grid side inverter and for induction motor drive using hysteresis current control. 5 PV panels each of maximum power point (MPP) voltage of 30 V, and current of 8 A is connected in series to form as a input to the boost converter for extracting maximum power.

Fig. 15 shows the maximum PV panels output current, power, and net loading at DC bus with respect to Common DC link voltage which is maintained at 300 V through grid side VSC. With sudden removal of drive load DC bus gets slightly increased before settling to 300 V of steady state using grid side VSC under proper control.

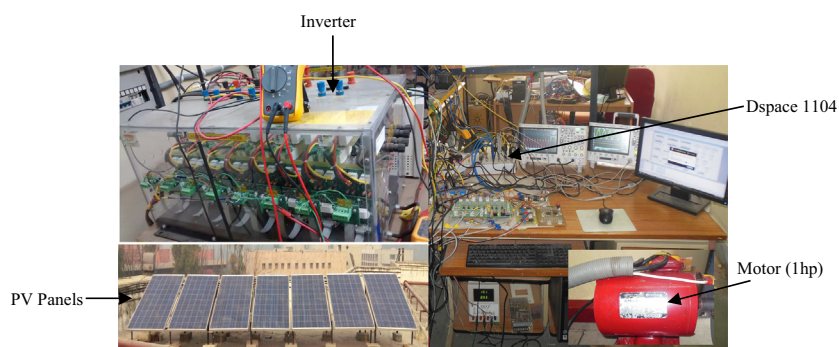


Fig. 19. Hardware setup.

This confirms the working of proper control algorithm as DC link remains constant even when drive load is disconnected. That means whole PV power is being fed to the grid. Under high insolation level i.e. generation more than that of required by the constant torque pump drive operation, the remaining power is sent to the grid via VSC as shown in Fig. 18. With sudden intentional decrease in insolation, demand some active power to be drawn from VSC to support the drive at its constant torque mode. Changing insolation has resulted into phase reversal of each phase current (i_R , i_Y , i_B) as evident from Fig. 18. In both the cases the VSC is being utilized for active power exchange only. For validating the algorithm for PQ issues only diode rectifier type of non-linear load is being considered. With insolation level just enough to cater the drive load, the grid side VSC capacity is being utilized only for full non-linear compensation as demonstrated in Fig. 16. From Fig. 16 it is clear that while load is demanding non-linear current, source current (i_s) remains sinusoidal. Full compensation is supported through grid side VSC. Fig. 17 shows the each phase motor current with respect to motor speed. To check effectiveness of the control varying speed reference command is being given resulted into varying torque mode operation.

Conclusion

The simulated and experimental results clearly demonstrate the ability of the proposed control scheme to evacuate MPP tracked power from the PV array and support the induction motor pumping system with grid connected mode and islanded mode with priority and selective based power conditioning of PQ problems posed by the connected non linear loads with varying reactive power and unbalance. The MPPT used in the control tracks the power very fast, even under step change of insolation, and the indirect current control of source current along with regulation of DC bus shows the uninterruptable continuous duty cycle operation of pumping system even under changing insolation conditions. Sensorless speed estimation of speed provides rugged system with cost effective solution. The proposed SRF based decomposition enables the control for providing limited compensation of PQ problems depending on available capacity of VSC. The implemented scheme derives the advantage of simplicity and is capable of delivering under varying insolation conditions effectively. Such technique is envisaged to benefit the PV pumping system and grid/microgrid by the limited compensation, thereby effectively utilizing the connected hardware.

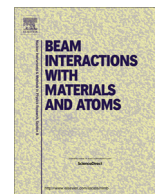
References

- [1] Singh M, Khadkikar V, Chandra A, Varma RK. Grid interconnection of renewable energy sources at the distribution level with power-quality improvement features. *IEEE Trans Power Delivery* 2010;26:307–15.
- [2] Kjaer Soeren Baekhoej, Pedersen John K, Blaabjerg Frede. A review of single-phase grid-connected inverters for photovoltaic module. *IEEE Trans Ind Appl* 2005;41:1292–306.
- [3] Elgendy MA, Zahawi B, Atkinson DJ. Assessment of perturb and observe MPPT algorithm implementation techniques for PV pumping applications. *IEEE Trans Sustain Energy* Jan 2012;3(1):21–33.
- [4] Carannante G, Fraddanno C, Pagano M, Piegari LPN. Experimental performance of MPPT algorithm for photovoltaic sources subject to inhomogeneous insolation. *IEEE Trans Ind Electron* 2009;56:4374–80.
- [5] Kadri R, Gaubert J-P, Champenois G. An improved maximum power point tracking for photovoltaic grid-connected inverter based on voltage-oriented control. *IEEE Trans Ind Electron* 2011;58:66–75.
- [6] Algazar Mohamed M, Al-monier Hamdy, EL-halim Hamdy Abd, Salem Mohamed Ezzat El Kotb. Maximum power point tracking using fuzzy logic control. *Int J Electr Power Energy Syst* 2012;39:21–8.
- [7] Correa T, Nevesy F, Seleme SI, Silva SR. Minimum motor losses point tracking for a stand-alone photovoltaic pumping system. In: *Proc. IEEE ISIE'08 Conf.*, June 30–July 2 2008. p. 1501–6.
- [8] Kassem Ahmed M. MPPT control design and performance improvements of a PV generator powered DC motor-pump system based on artificial neural networks. *Int J Electr Power Energy Syst* 2012;43:90–8.
- [9] Arribas JR, Gonzalez CMV. Optimal vector control of pumping and ventilation induction motor drives. *IEEE Trans Ind Electron* 2002;49(4):889–95.
- [10] Vitorino MA, de Rossiter Correa MB, Jacobina CB, Lima AMN. An effective induction motor control for photovoltaic pumping. *IEEE Trans Ind Electron* 2010;58(4):1162–70.
- [11] Abdelsalam AK, Massoud AM, Ahmed S, Enjeti PN. High-performance adaptive perturb and observe MPPT technique for photovoltaic-based microgrids. *IEEE Trans Power Electron* 2011;26:1010–21.
- [12] Sait H Habeebullah, Daniel S Arul. New control paradigm for integration of photovoltaic energy sources with utility network. *Int J Electr Power Energy Syst* 2011;33:86–93.
- [13] Singh BN, Rastgoufard P, Singh B, Chandra A, Haddad KA. Design, simulation and implementation of three pole/four pole topologies for active filters. *Inst Electr Eng Proc Electr Power Appl* 2004;151(4):467–76.
- [14] Singh BN, Singh B, Chandra A, Haddad KA. Design, digital implementation of active filters with power balance theory. *Inst Electr Eng Proc Electr Power Appl* 2005;152(5):1149–60 [power].
- [15] Depenbrock M, Staudt V, Wrede H. Concerning instantaneous power compensation in three-phase systems by using p–q–r theory. *IEEE Trans Power Electron* 2004;19(4):1151–2.
- [16] Aredes M, Akagi H, Watanabe EH, Salgado EV, Encarnação LF. Comparisons between the p–q and p–q–r theories in three-phase four-wire systems. *IEEE Trans Power Electron* 2009;24(4):924–33.
- [17] Singh B, Verma V. Selective compensation of power-quality problems through active power filter by current decomposition. *IEEE Trans Power Delivery* 2008;23(2).
- [18] Hong Chih-Ming, Chen Chiung-Hsing. Intelligent control of a grid-connected wind-photovoltaic hybrid power systems. *Int J Electr Power Energy Syst* 2014;55:554–61.
- [19] Kamatchi Kannan V, Rengarajan N. Photovoltaic based distribution static compensator for power quality improvement. *Int J Electr Power Energy Syst* 2012;42:685–92.
- [20] Rahmani Bijan, Li Weixing, Liu Guihua. An advanced universal power quality conditioning system and MPPT method for grid integration of photovoltaic system. *Int J Electr Power Energy Syst* 2012;42:661–71.
- [21] Zaveri N, Chudasama A. Control strategies for harmonic mitigation and power factor correction using shunt active filter under various source voltage conditions. *Int J Electr Power Energy Syst* 2012;42:661–71.
- [22] Singh B, Verma V, Solanki J. Neural network-based selective compensation of current quality problems in distribution systems. *IEEE Trans Ind Electron* 2007;54(1):53–60.
- [23] Miret J, Castilla M, Matas J, Guerrero JM, Vasquez JC. Selective harmonic-compensation control for single-phase active power filter with high harmonic rejection. *IEEE Trans Ind Electron* 2009;56:3117–27.
- [24] Liu Bin, Wu Jia Ju, Li Jun, Dai Ji Yang. A novel PFC controller and selective harmonics suppression. *Electr Power Energy Syst* 2013;44:680–7.
- [25] Verma Vishal, Kumar Amritesh. Grid coupled maximum power point tracked photovoltaic system with selective power conditioning capability. In: *IEEE international conference on power and energy (PECon)*; 2012. p. 886–91.
- [26] Villalava MG, Gazoli JR, Ruppert E. Modelling and circuit-based simulation of photovoltaic arrays. *Braz J Power Electron* 2009;14(4):35–45.
- [27] Bose BK. *Modern power electronics and AC drives*. fourth ed. Pearson Education; 2004.



Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research B

journal homepage: www.elsevier.com/locate/nimb

Potential application of carbon nanotube core as nanocontainer and nanoreactor for the encapsulated nanomaterial

Pawan K. Tyagi^{a,*}, Reetu Kumari^a, Umananda M. Bhatta^b, Raghavendra Rao Juluri^c, Ashutosh Rath^c, Sanjeev Kumar^d, P.V. Satyam^c, Subodh K. Gautam^e, Fouran Singh^e

^a Department of Applied Physics, Delhi Technological University, Delhi 110042, India

^b Centre for Emerging Technologies, Jain University, Jakkasandra, Kanakapura Taluk, Ramanagaram Dist, Karnataka 562 112, India

^c Institute of Physics, Sachivalaya Marg, Bhubaneswar 751005, India

^d Department of Physics, Rajiv Gandhi University, Itanagar, Arunachal Pradesh 791112, India

^e Inter University Accelerator Centre, Aruna Asaf Ali Marg, New Delhi 110067, India

ARTICLE INFO

Article history:

Received 16 November 2015

Received in revised form 20 February 2016

Accepted 9 March 2016

Available online xxx

Keywords:

Filled-carbon nanotube

Electron-irradiation

Nano-engineering

Shape-alteration

In-situ transmission electron microscopy

ABSTRACT

Fe₃C nanorod filled inside carbon nanotube has been irradiated inside transmission electron microscope at both room and high temperature. *In-situ* response of Fe₃C nanorod as well as CNT walls has been studied. It has been found that when electron irradiation is performed at room temperature (RT), nanorod first bends and then tip makes at the end whereas at high temperature (~490 °C) nanorod slides along the tube axis and then transforms into a faceting particle. Extrusion of solid particle filled in the core of CNT has also been demonstrated. It is suggested that these morphological changes in nanorod may have happened due to the compression which was generated either by shrinkage of tube or by local electron beam heating. Presented results demonstrate that CNT core could be used as nano-container or reactor.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Concave geometry of hollow core of carbon nanotube (CNT) with confined space offers the tremendous possibility to generate the nanomaterial of superior physical, chemical or electronic properties [1]. In previous reports [2–15], CNT filled with ferromagnetic material has been proposed as a novel material and their myriad potential applications such as magnetic recording media [16], magnetic force microscopy (MFM) [17], biomedicine [18,19] and spintronics [20] have been demonstrated. This has been attributed to the fact that in filled-CNT, tube-walls not only protect the filled nanomagnets against harsh environment but also prohibit coalescence. Filling of CNT core with different materials have facilitated the fabrication of nanostructures in a controlled fashion and has introduced the interesting applications of CNT such as containers, conduits, pipettes, and coaxial cables. In order to further expand the applications of filled-CNT key aspect will be post-growth modification techniques which can modify shape as well as position of the filled materials in a controlled manner. In recent reports, elec-

tron irradiation performed inside high-resolution transmission electron microscopy (HRTEM) demonstrated that when post-growth modification is performed in a controlled manner, it is possible to control the shape, size and properties of the CNTs as well as encapsulated metal by producing defects, at both room [21,22] and high [23] temperatures. Under the energetic electron beam, extrusion of filled material has been demonstrated [24,25]. This evidenced that electron beam might be a tool which has post-synthesis alteration capability on CNT [23,26–30]. Hence, we envision that CNTs either fully or partially filled have many prospective applications. Some of them have been already demonstrated where CNT was used as nanopipette [31], high pressure nanoextruder [24], diagnostic tool [18] and preservative nanocell [18].

Another interesting applications of ferromagnetic nanoparticles, and nanorod of sharp tip filled inside CNT has been reported in the area of high-density magnetic recording media and medicinal imaging [32–34]. Hence, being motivated with the approach adopted in references, [24,25] we performed electron irradiation on Fe₃C-filled CNT at both room and high temperature. *In-situ* response of Fe₃C nanorod as well as CNT was monitored and irradiation induced structure related phenomena such as: shrinkage, thinning, sharpening of tip and shape transformation of nanorod were studied.

* Corresponding author.

E-mail addresses: pawan.phy@dce.edu, pawankumartyagi@gmail.com (P.K. Tyagi).

2. Experimental details

Fe₃C-filled MWCNTs were grown by modified thermal chemical vapor deposition (Thermal CVD) technique. More details are described previously [35]. The as-grown CNTs were scratched from the quartz tube and a very small amount of the sample was dispersed in isopropyl alcohol. Few drops of the sample were casted on Cu-substrate and glass substrate for XRD and Raman measurement, respectively. As-grown filled-CNTs sample was characterized by using Bruker D8 Advance X-ray diffractometer, Hitachi S3100 Scanning Electron Microscope and Raman spectrometer (Renishaw, model Invia) with 514 nm Ar ion laser operated at power 5 mW.

For TEM measurement dispersed sample was sonicated for 10 min and then one drop of the dispersion was casted onto uncoated copper grid of 1000 mesh for high resolution TEM imaging (JEOL JEM 2010 operated at 200 kV) for carrying out *in-situ* electron irradiation experiments as well as imaging. The grid was

then transferred to the specimen holder for electron irradiation and microscopy measurement. Electron beam of energy 200 keV was used. Irradiation at both room and high temperature ($\sim 490^\circ\text{C}$) was carried out at current density of $100\text{--}300\text{ A/cm}^2$ which corresponds to electron concentration $6 \times 10^{21}\text{--}19 \times 10^{21}$ per cm^2 per second.

3. Results and discussion

SEM image of as-grown MWCNTs is shown in Fig. 1(a). In order to confirm that wire type structures shown in Fig. 1(a) are filled-MWCNT, the sample was further characterized by using XRD measurement as well as Raman measurement. The observed diffraction pattern of filled-MWCNT is shown in Fig. 1(b). The diffraction peak observed at $2\theta = 26.4^\circ$ is indexed as a characteristic peak of MWCNT, representing the inter-wall spacing. The observed value of average d_0 is $3.3529 \pm 0.0005\text{ \AA}$ (PCPD file no: 89-8487). The diffraction peaks at $2\theta = 36.82^\circ$, 43.38° and 61.45° assigned to (112), (121) and (222) planes of orthorhombic crystal structure (cementite phase) of Fe₃C. These reported values are in accordance with 37.6° , 43.7° and 61.3° (PCPD file no: 89-2722) for same planes. Remaining peaks are identified and found to correspond to Cu substrate.

Raman spectrum of as-grown sample is shown in Fig. 1(c). Main Raman features have observed at position around 1345 cm^{-1} , 1573 cm^{-1} and 1606 cm^{-1} . D peak at 1347 cm^{-1} assigned to K-point phonons of A_{1g} symmetry associated with breathing vibration of a 6-fold aromatic ring. This mode activates by disorder in graphitic structures. G peak position at 1575 cm^{-1} assigned to zone center phonons of E_{2g} symmetry or sp² stretch vibration in benzene or condensed benzene rings. D' peak at 1611 cm^{-1} originates when phonons in defected graphite with a small q activates.

Fig. 2(a) is the TEM image of cementite nanorod encapsulated inside the CNT. The diameter of the metal nanorod is 10.2 nm. Lattice fringes (Fig. 2b) of cementite crystal encapsulated inside the core of CNT reveal that cementite is a single crystal. The image in Fig. 2(b) demonstrates that the lattice fringe with a periodicity of

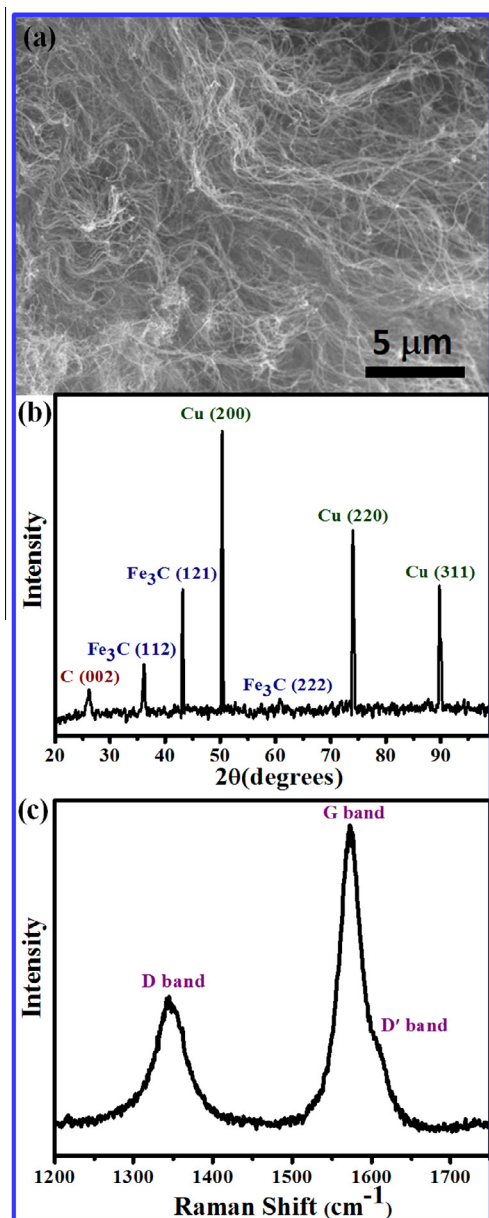


Fig. 1. (a) SEM micrograph, (b) XRD pattern and (c) Raman spectrum of Fe₃C-filled MWCNTs.

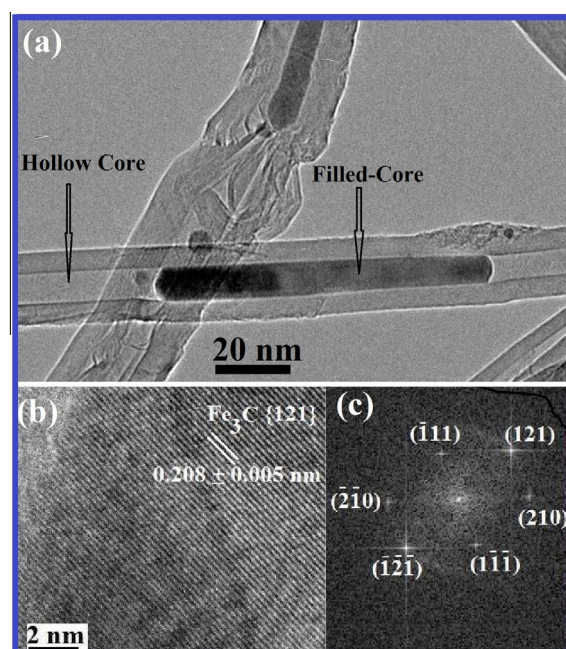


Fig. 2. Panel (a) TEM image of nanorods encapsulated inside the MWCNT. Panel (b) and (c) is the lattice image of nanorod and their corresponding Fast Fourier Transforms (FFT) pattern, respectively. The encapsulated crystal was identified as iron carbide (cementite, Fe₃C).

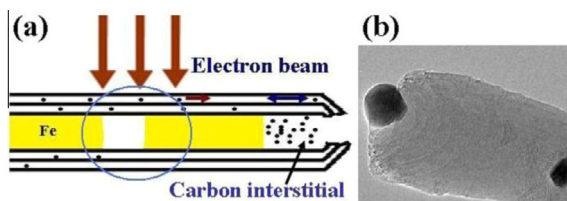


Fig. 3. (a) Illustration of the electron-irradiation process as well as path of carbon interstitials migration. (b) When subjected to continuous electron-irradiation for 240 s, Fe_3C nanorod eventually transformed into a nanoparticle.

0.208 ± 0.005 nm which correspond to the $\{121\}$ planes of Fe_3C (orthorhombic structure, space group Pbnm , and the unit cell parameters are $a = 4.523$ Å, $b = 5.089$ Å, $c = 6.7428$ Å). Furthermore, Fourier transform (Fig. 2c) of lattice image reveal that filled-crystal is iron carbide (cementite, Fe_3C). Since, no diffraction spots splitting were observed in the Fourier transform. It can be concluded that the core of CNT can be used as nano-container for the growth of single crystalline Fe_3C nanorod. As Stamatin et al. [36] reported that CNT morphology dependent on iron particle and their forms (Fe or Fe_3C) resulted from ferrocene decomposition. In present case, nucleation of CNT starts from Fe and then Fe_3C -filled CNT forms. More details of growth mechanism as well as experimental conditions can be found in Ref [35].

Irradiation process has been illustrated in schematic diagram, shown in Fig. 3(a). A time-resolved picture series of the electron micrographs of Fe_3C nanorod which was imaged during the sustained electron irradiation is shown in Fig. 4. After 240 s of continuous irradiation, nanorod was found to be transformed into a faceted nanoparticle as shown in Fig. 4(f). Axial sliding of nanorod towards the open-end of CNT was observed before this transformation. As evident from Fig. 4(f), Fe_3C nanoparticle was in spherical shape at open end of CNT, whereas faceting at the rear side. This shape transformation has also been observed inside the core of the CNT and depicted in Fig. 5. Here we presumed that particle (in a spherical form) has uniform thickness and that the intensity of electron beam varies with particle thickness. Also, FWHM of peak in intensity profile, measured along diameter by using Image J software is supposed to roughly represent the thickness. Curve (thickness vs. irradiation time) was plotted and shown in Fig. 6 (a). This revealed that continuous electron irradiation increased the particle thickness. In bright field TEM image, amplitude of contrast varied with thickness and thicker area appears darker. Quantitatively mass–thickness contrast (C) was calculated by using relation: $C = \frac{\Delta I}{I} = 1 - e^{-Q\Delta t}$, where Q and Δt are total elastic scattering cross section and change in thickness, respectively [37]. Mass–thickness-contrast was then plotted in Fig. 6(b) and found to be nonlinear. This nonlinearity may have arisen due to significant contribution of diffraction-contrast. It was also observed as

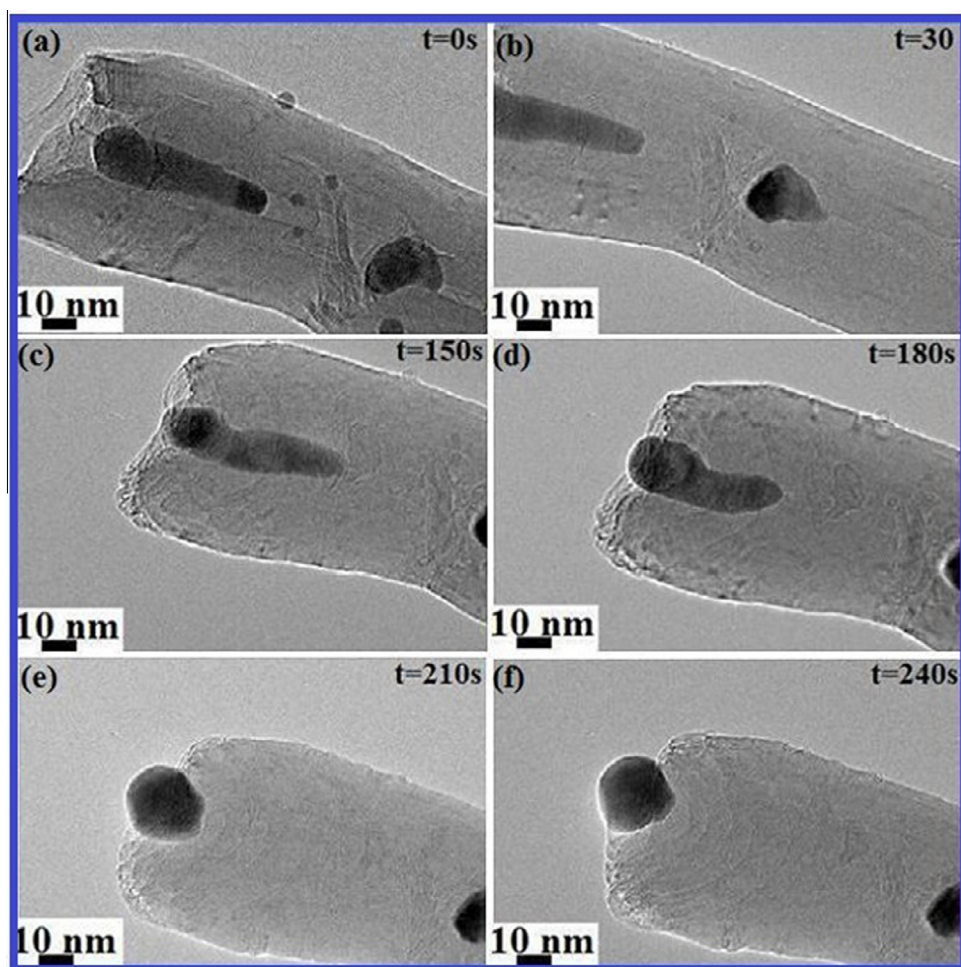


Fig. 4. Effect of electron-irradiation performed at 490°C on Fe_3C nanorod filled near the open-end of the CNT. Panels (b–f) depict the deformation, extrusion and shape transformation of nanorod due to the contraction of graphitic shell. Finally, sustained electron irradiation transformed the rod into a particle, faceting on the front side where as spherical on the rear side panel (f).

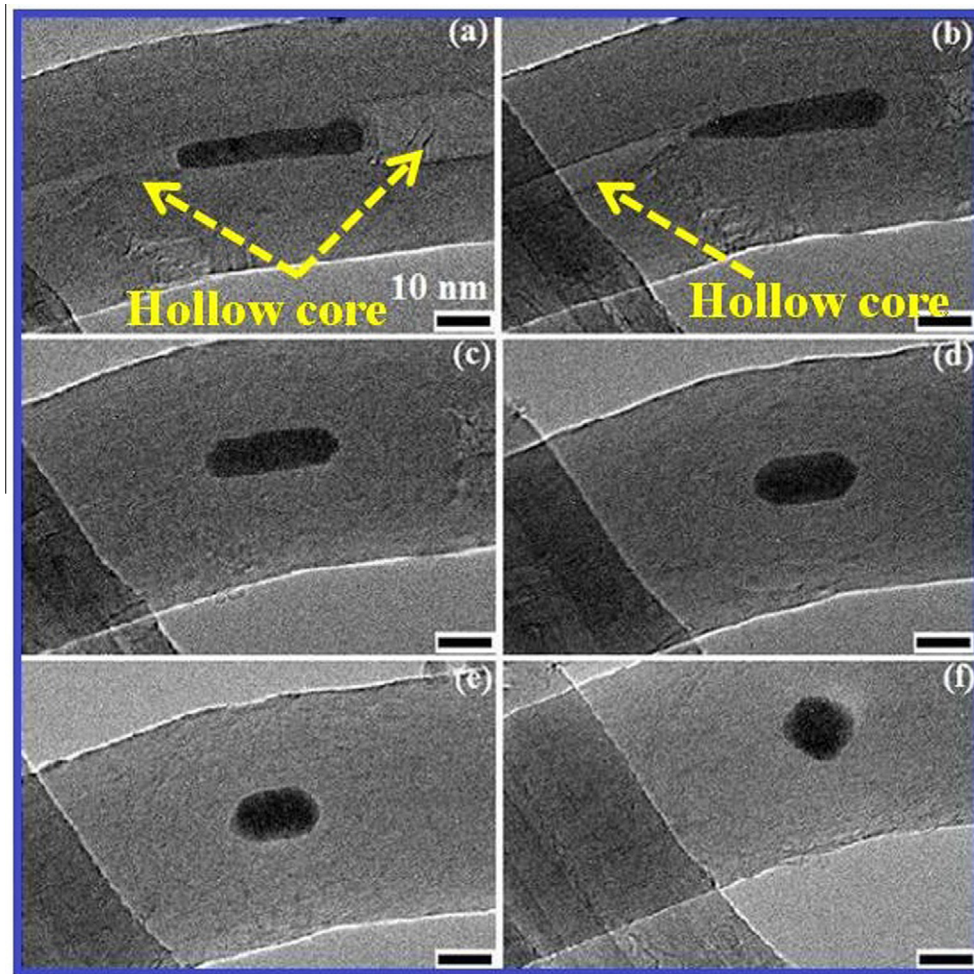


Fig. 5. Effect of electron-irradiation performed at 490 °C on Fe_3C nanorod filled in interior of the CNT. CNT before (a) and (b) after 1 min of irradiation both ends of nanorod are veiled by graphitic layer and then nanorod transformed into faceted-nanoparticle (e and f). Scale bar is 10 nm.

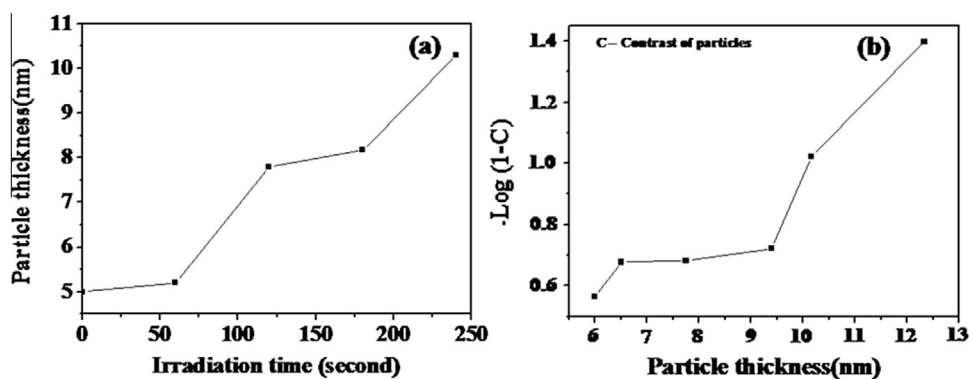


Fig. 6. (a) Thickness of particle estimated through FWHM of peak in intensity profile of bright field TEM image with irradiation time, (b) A typical mass-thickness-contrast profile with increasing particle thickness where C is contrast of particle.

electron irradiation proceeds both ends of the nanorod were veiled by thick graphite flake and then shape transformation occurred as shown in Fig. 5(c–f).

It was observed that shape transformation and extrusion of the Fe_3C nanorod occurred only if electron-irradiation was performed at high temperature. This might have happened due to contraction of graphitic shell as evident in Fig. 7. However at room temperature shape deformation and tip formation in nanorod were seen as evi-

dent from Fig. 8. In Fig. 8, sequential high resolution TEM images of Fe_3C -filled CNT depicted the bending and formation of sharp-tip at various time intervals. These images were recorded at room temperature under sustained electron-irradiation for duration of 10 min.

An explicit interpretation of irradiation-induced changes is not possible without the precise knowledge of production of atomic defects. Hence, we need to understand how defects are produced

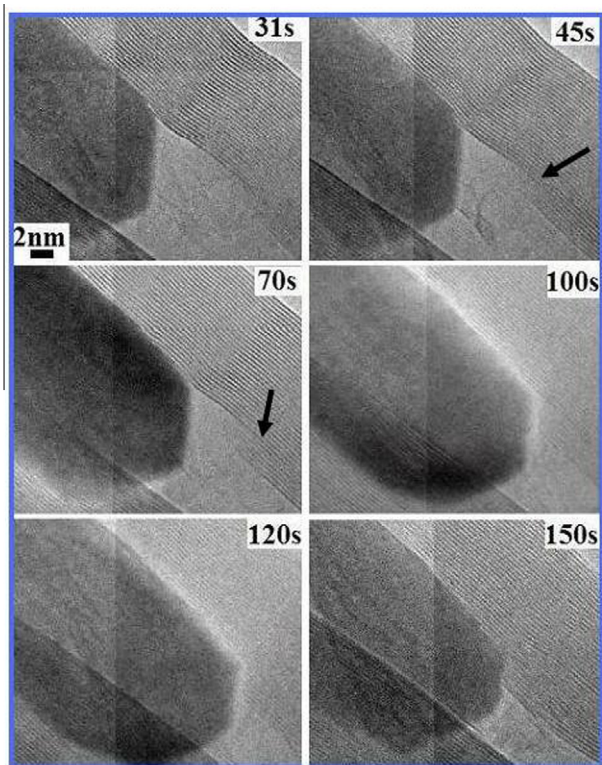


Fig. 7. High resolution images were captured at various time intervals from the real-time movie which was recorded during sustained electron-irradiation at high temperature ($\sim 490^\circ\text{C}$). Irradiation period is mentioned in panels. Cementite nanorod extrudes towards the terminating CNT walls and after 150 s, eventually lead to formation of cementite-carbon contact. Position of terminating CNT walls is indicated by arrow. All frames are taken at same magnification.

in both CNT as well as in nanorod lattice. Previous reports [21–23,38,39] evidenced that number of complex defects such as the vacancies; Frankel pair (interstitial-vacancy pair), interstitials, Wigner defects and intershell covalent bonding in the lattice of CNT were formed when CNT was subjected under energetic electron beam. It was stated that this happened due to the knock-on displacement of carbon atoms by the energetic electrons. In many cases irradiation induced changes were governed by annealing and diffusion of vacancies and interstitials [38,39]. Recently in work done by Anshika et al. [40] it has been seen that under electron irradiation at room temperature vacancies are produced in Fe_3C nanorod lattice. If electron irradiation was performed at 490°C inverse Ostwald ripening of nano cluster was observed in nanorod lattice. Vacancy diffusion-mediated mass/adatom transfer in nanoparticle has been reported by Sun et al. [25]. They reported that vacancies are formed in the part of nanoparticle with low pressure, and then migrate to the high-pressure part. In the present work (Figs. 4 and 5), it was observed that as the irradiation proceeded, hollow core was filled with carbonaceous structure. These structures are supposed to be formed of carbon atoms either sputtered from inner or outer shell of CNT or displaced from its initial position as interstitials [40]. These carbon atoms were reported to migrate faster in hollow core of CNT than in open space between the adjacent CNT shells [26]. Carbon atoms were reported to migrate from high pressure part to low pressure part [38,41,42]. This pressure in CNT core was generated due to reconstruction of CNT lattice to heal vacant site and subsequent shrinkage of CNT [38,41,42]. This pressure was not uniform in CNT core as irradiated area of CNT was seen to be always under high compression [24]. As concentration of vacant site i.e. vacancies in CNT increase then bending, shrinkage and finally breakage occurs [43]. It can be suggested that migration of carbon atoms in hollow core finally led to the filling of CNT core and growth of other graphitic structures as

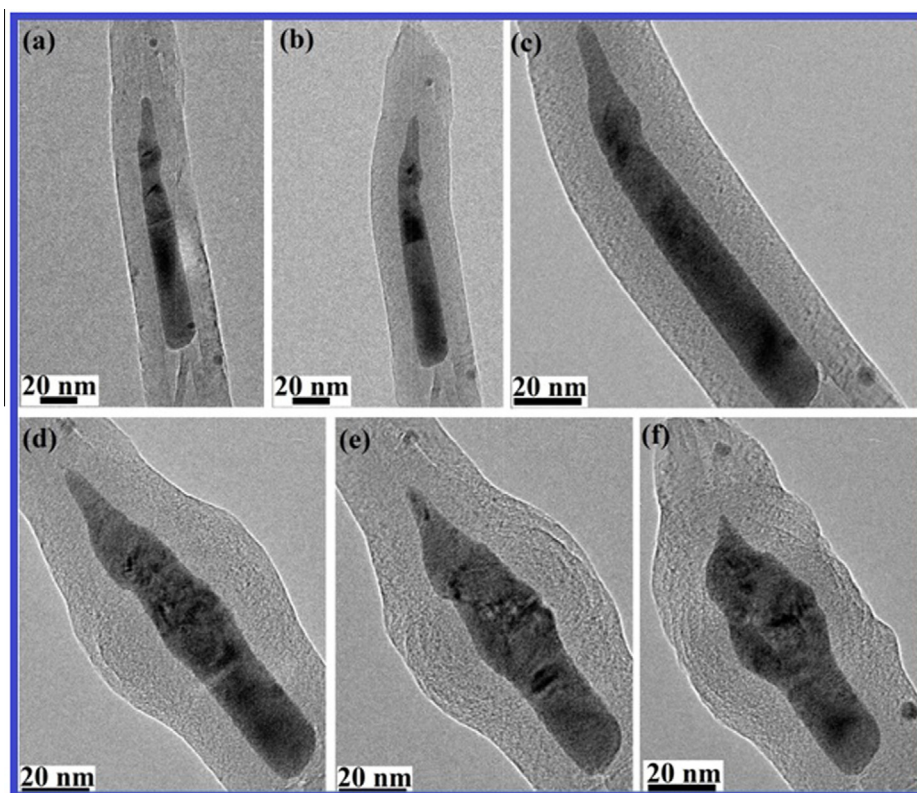


Fig. 8. Sequential high resolution TEM images of Fe_3C -filled MWCNTs depicted the bending and formation of sharp-tip at various time intervals. The images have been recorded at room temperature after sustained electron-irradiation for duration of 10 min.

evident in Figs. 4 and 5. If irradiation further proceeded, interstitials aggregation or clustering led to the amorphization of CNT lattice i.e. complete blockage of interstitials migration in both CNT core and in open space between the adjacent CNT shells, as shown in Figs. 4 and 5. In this case, non-isotropic pressure is exerted on nanorod from all direction. Degree and uniformity of this exerted pressure was seen to depend on degree of disorder in surrounding graphitic structures. CNT walls seem to exert high and more uniform pressure. It can be suggested that vacancy-mediated mass transfer facilitated the shape transformation in these nanocrystals. This process is not probable one because other processes such as surface energy minimization as well as electron beam heating also influenced the shape transformation. As evident from Fig. 4(f), nanocrystal is in spherical shape at open end of CNT. This reflects that surface energy minimization also contributed if pressure is not exerted. Third possibility of local electron beam heating cannot be ruled out. Under e-irradiation, significant electron–electron interactions generated the electronic excitations which may be localized due to the reduction of conduction electrons [41,42,44] and transferred the energy to the nanorod lattice in picoseconds [41,42,44]. This might have resulted either in strong heating of the lattice or locally melting of nanorod. If filled-CNT is irradiated at room temperature, pressure could not be generated due to the shrinkage of tube as CNT lattice gets quickly damaged. However, bending, tip formation and shape deformation of nanorod were observed when filled-CNT was irradiated at room temperature as evidenced from Fig. 8. These deformations occurred due to electron beam heating. Therefore, it can be concluded that controlled faceted shape transformation occurs only due to pressure induced elastic displacement process.

4. Conclusion

In the reported work, response of Fe₃C-filled carbon nanotube under electron irradiation has been studied at both room and high temperature. Extrusion of nanocrystal has been demonstrated. Present study reveals that the compression present in nanorod which was exerted by CNT walls controlled the shape transformation of nanorod. Bending, shape transformation and tip formation has occurred due to the localized heating in nanorod. It can be envisioned that the present studies will facilitate the post synthesis alteration in filled-CNTs and pave way for numerous applications. Carbon nanotube could thus be used as nanoreactor for the shape transformation of encapsulated crystal.

References

- [1] Q. Fu, G. Weinberg, D.S. Su, Selective filling of carbon nanotubes with metals by selective washing, *New Carbon Mater.* 23 (2008) 17.
- [2] J.P. Cheng, B. Zhang, G.F. Yi, Y. Ye, M.S. Xia, Preparation and magnetic properties of iron oxide and carbide nanoparticles in carbon nanotube matrix, *J. Alloys Compd.* 455 (2008) 5.
- [3] S. Karmakar, S.M. Sharma, M.D. Mukadam, S.M. Yusuf, A.K. Sood, Magnetic behavior of iron-filled multiwalled carbon nanotubes, *J. Appl. Phys.* 97 (2005) 054306.
- [4] L. Jian-hua, H. Ruo-yu, L. Gua-hua, W. Dong-guang, An easy approach to encapsulating Fe₃O₄ nanoparticles in multiwalled carbon nanotubes, *New Carbon Mater.* 25 (3) (2010) 192.
- [5] F. Wolny, T. Mühl, U. Weissker, K. Lipert, J. Schumann, A. Leonhardt, B. Büchner, Iron filled carbon nanotubes as novel monopole-like sensors for quantitative magnetic force microscopy, *Nanotechnology* 21 (2010) 435501.
- [6] U. Weissker, S. Hampel, A. Leonhardt, B. Büchner, Carbon nanotubes filled with ferromagnetic materials, *Materials* 3 (2010) 4387.
- [7] C.N.R. Rao, A. Govindraj, R. Sen, B.S. Satishkumar, Synthesis of multi-walled and single-walled nanotubes, aligned-nanotube bundles and nanorods by employing organometallic precursors, *Mater. Res. Innovation* 2 (1998) 128.
- [8] A. Leonhardt, S. Hampel, C. Müller, I. Mönch, R. Koseva, M. Ritschel, et al., Synthesis, properties, and applications of ferromagnetic-filled carbon nanotubes, *Chem. Vap. Deposition* 12 (2006) 380.
- [9] L.J. Li, A.N. Khlobystov, J.G. Wiltshire, G.A.D. Briggs, R.J. Nicholas, Diameter-selective encapsulation of metallocenes in single-walled carbon nanotubes, *Nat. Mater.* 4 (2005) 481.
- [10] E.P. Sajitha, V. Prasad, S.V. Subramanyam, A.K. Mishra, S. Sarkar, C. Bansal, Size-dependent magnetic properties of iron carbide nanoparticles embedded in a carbon matrix, *J. Phys.: Condens. Matter* 19 (2007) 046214.
- [11] R.S. Iskhakov, S.V. Komogortsev, A.D. Balaev, A.V. Okotrub, A.G. Kudashov, Fe nanowires in carbon nanotubes as an example of a one-dimensional system of exchange-coupled ferromagnetic nanoparticles, *JETP Lett.* 78 (2003) 236.
- [12] M. Kumar, Y. Ando, Chemical vapor deposition of carbon nanotubes, *J. Nanosci. Nanotech.* 10 (2010) 3739.
- [13] M.A. Correa-Duarte, M. Grzelczak, V. Salgueirino-Maceiro, M. Giersig, L.M. Liz-Marzan, M. Farle, K. Sieradzki, R. Diaz, Alignment of carbon nanotubes under low magnetic fields through attachment of magnetic nanoparticles, *J. Phys. Chem. Lett.* 109 (2005) 19060.
- [14] R. Bhatiya, V. Prasad, Synthesis of multiwall carbon nanotubes by chemical vapor deposition of ferrocene alone, *Solid State Commun.* 150 (2010) 311.
- [15] A. Leonhardt, M. Ritschel, R. Kozhuharova, A. Graff, T. Mühl, R. Huhle, et al., Synthesis and properties of filled carbon nanotubes, *Diamond Related Mater.* 12 (2003) 790.
- [16] Y.C. Stephen, S.W. Mark, R.K. Peter, P.B. Fischer, Single-domain magnetic pillar array of 35 nm diameter and 65 Gbits/in² density for ultrahigh density quantum magnetic storage, *J. Appl. Phys.* 76 (1994) 6673.
- [17] F. Wolny, U. Weissker, T. Mühl, M.U. Lutz, C. Müller, A. Leonhardt, B. Büchner, Stable magnetization of iron filled carbon nanotube MFM probes in external magnetic fields, *J. Phys. Conf. Ser.* 200 (2010) 112011.
- [18] I. Mönch, A. Leonhardt, A. Meye, S. Hampel, R.K. Kozhuharova, D. Elefant, et al., Synthesis and characteristics of Fe-filled multi-walled carbon nanotubes for biomedical application, *J. Phys.: Conf. Ser.* 61 (2007) 820.
- [19] M. Martincic, G. Tobias, Filled carbon nanotubes in biomedical imaging and drug delivery, *Expert Opin. Drug Deliv.* 12 (2015) 563.
- [20] F. Rossella, C. Soldano, P. Onorato, V. Bellani, Tuning electronic transport in cobalt-filled carbon nanotubes using magnetic fields, *Nanoscale* 21 (2014) 788.
- [21] K. Suenaga, H. Wakabayashi, M. Koshino, Y. Sato, K. Urita, S. Iijima, Imaging active topological defects in carbon nanotubes, *Nat. Nanotech.* 2 (2007) 358.
- [22] A. Hashimoto, K. Suenaga, A. Gloter, K. Urita, S. Iijima, Direct evidence for atomic defects in graphene layers, *Nature* 430 (2004) 870.
- [23] K. Urita, K. Suenaga, T. Sugai, H. Shinohara, S. Iijima, *In-situ* observation of thermal relaxation of interstitial-vacancy pair defects in a graphite gap, *Phys. Rev. Lett.* 94 (2005) 155502.
- [24] L. Sun, F. Banhart, A.V. Krasheninnikov, J.A. Rodriguez-Manzo, M. Terrones, P. M. Ajayan, Carbon nanotubes as high-pressure cylinders and nanoextruders, *Science* 312 (2006) 1202.
- [25] Litao Sun, Arkady V. Krasheninnikov, Tommy Ahlgren, Kai Nordlund, Florian Banhart, Plastic deformation of single nanometer-sized crystals, *Phys. Rev. Lett.* 101 (2008) 156101.
- [26] Y. Gan, J. Kotakoski, A.V. Krasheninnikov, K. Nordlund, F. Banhart, The diffusion of carbon atoms inside carbon nanotube, *New J. Phys.* 10 (2008) 023022.
- [27] A. Kis, G. Csanyi, J.-P. Salvetat, T.-N. Lee, E. Coutreau, A.J. Kulik, W. Benoit, J. Brugger, L. Forro, Reinforcement of single-walled carbon nanotube bundles by intertube bridging, *Nat. Mater.* 3 (2004) 153.
- [28] C. Gomez-Navarro, P.J. De Pablo, J. Gomez-Herrero, B. Biel, F.J. Garcia-Vidal, A. Rubio, F. Flores, Tuning the conductance of single-walled carbon nanotubes by ion irradiation in the Anderson localization regime, *Nat. Mater.* 4 (2005) 534.
- [29] M. Terrones, H. Terrones, F. Banhart, J.-C. Charlier, P.M. Ajayan, Coalescence of single-walled carbon nanotubes, *Science* 288 (2000) 1226.
- [30] W. Mickelson, S. Aloni, W.Q. Han, J. Cummings, A. Zettl, Packing C60 in boron nitride nanotubes, *Science* 300 (2003) 467.
- [31] H.J. Hwang, K.R. Byun, J.W. Kang, Carbon nanotubes as nanopipette: modeling and simulations, *Physica E* 23 (2004) 208.
- [32] G. Reiss, A. Hutten, Magnetic nanoparticles – applications beyond data storage, *Nat. Mater.* 4 (2005) 725.
- [33] Q.A. Pankhurst, J. Connolly, S.K. Jones, J. Dobson, Applications of magnetic nanoparticles in biomedicine, *J. Phys. D* 36 (2003) 167.
- [34] M.H. Kryder, Magnetic recording beyond the superparamagnetic limit magnetics, *Conference INTERMAG 2000 Digest of Technical Papers, IEEE International*, 2000, p. 575.
- [35] R. Kumari, L. Krishnia, V. Kumar, S. Singh, H.K. Singh, R.K. Kotnala, R.R. Juluri, U. M. Bhatta, P.V. Satyam, B.S. Yadav, Z. Naqvi, P.K. Tyagi, Fe₃C-filled carbon nanotubes: permanent cylindrical nanomagnets possess exotic magnetic properties, *Nanoscale* 8 (2016) 4299.
- [36] I. Stamatina, A. Morozan, A. Dumitru, V. Ciupina, G. Prodan, J. Niewolski, H. Figiel, The synthesis of multi-walled carbon nanotubes (MWNTs) by catalytic pyrolysis of the phenol-formaldehyde resins, *Physica E* 37 (2007) 44.
- [37] David B. Williams, C. Barry Carter, *Transmission Electron Microscopy: A Textbook for Materials Science*, Springer, p. 371.
- [38] F. Banhart, Irradiation effects in carbon nanostructures, *Rep. Progr. Phys.* 62 (1999) 1181.
- [39] A. Krasheninnikov, F. Banhart, Engineering of nanostructured carbon materials with electron or ion beams, *Nat. Mater.* 6 (2007) 723.
- [40] Anshika Singh, Reetu Kumari, Vinay Kumar, Lucky Krishnia, Zainab Naqvi, Amrith K. Panwar, Umananda M. Bhatta, Arnab Ghosh, P.V. Satyam, Pawan K. Tyagi, Electron irradiation induced buckling, morphological transformation, and inverse Ostwald ripening in nanorod filled inside carbon nanotube, *Appl. Surf. Sci.* 360 (2016) 1003.

- [41] A.V. Krashennnikov, K. Nordlund, Ion and electron irradiation-induced effects in nanostructured materials, *J. Appl. Phys.* 107 (2010) 071301.
- [42] A. Krashennnikov, F. Banhart, Engineering of nanostructured carbon materials with electron or ion beams, *Nat. Mater.* 6 (2007) 723.
- [43] J.X. Li, F. Banhart, The engineering of hot carbon nanotubes with a focused electron beam, *Nanoletters* 4 (2004) 1143.
- [44] A. Misra, P.K. Tyagi, P. Rai, D.R. Mahapatra, J. Ghatak, P.V. Satyam, D.K. Avasthi, D.S. Misra, Axial buckling and compressive behavior of nickel-encapsulated multiwalled carbon nanotubes, *Phys. Rev. B* 76 (2007) 014108.

Prediction of Tool-Life Model for Heavy Machining Using Central Composite Design

Vipin¹, Priya²

¹(Prod. & Ind. Engineering, Delhi Technological University, Delhi, India)

²(Mahavir Sawami Institute of Technology, Sonipat, Haryana, India)

vipin2867@yahoo.co.in

Abstract : The paper presents a study for tool-life models for turning cast steel material with uncoated cemented carbide tool. In this heavy machining operation under dry conditions with varying depth of cut between 5-10 mm is used. The tool-life models have been developed from the data obtained experimentally in terms of cutting speed, feed rate and depth of cut using factorial design method and regression analysis.

Keywords: Tool-life, cutting speed, feed, depth of cut, metal removal rate.

I. INTRODUCTION

For the design of cutting tools; ability to predict the tool life during machining is necessary. This involves determination of tool changing strategies and optimum cutting conditions. The extensive research in this area has contributed greatly for understanding the problem. There is no machining theory provide adequate relationships between work and tool material properties, tool-life and cutting conditions and, tool geometrical parameters. The major difficulties are the lack of suitable data and complexity of the machining process. Tool life depends on a number of variables i.e. machine tool, tool material and geometry, microstructure of work material being cut, the desired cubic-inch removal rate, cutting conditions and whether or not cutting fluids are used, etc. No machining theory is available to predict the tool life for a practical machining situation.

Researchers Hassan and Suliman [1990], Chua *et al* [1993], Lin and Chen [1995], Choudhury and El-Baradie [1998], Santos *et al* [1999], Ozler *et al* [2001], Lima *et al* [2005], Vipin *et al* [2009] worked on predicting the tool-life. However the models developed by these Researchers are mostly empirical relations between cutting speeds, feeds and varying depth of cuts up to 5mm. These models are silent with regards to heavy machining process for depth of cut above 5mm. Ranganath M. S. *et al* [2014], Ranganath M. S. *et al* [2014] worked on Response Surface Methodology.

The present work therefore emphasis to develop mathematical models to predict tool life equation for medium and heavy machining operation for depth of cut between 5 to 10 mm. The depth of cut, feed rate and cutting speed were obtained from double column heavy vertical boring and turning machine for turning operations on work of 3.6m diameter, EN6 cast steel material. The data so obtained was used to develop Tool life equation with the help of Central composite Design.

First- and second-order models are developed with 95% confidence level by using response surface methodology and

2³ factorial design of experiment. 3D plots of tool-life, and dual-response contours of metal removal rate and tool life for different cutting conditions are developed by the model equations.

II. TOOL LIFE MODEL

The proposed relationship between the machining response (tool life) and machining independent variables can be represented by the following:

$$T = C(V^\alpha f^\beta d^\gamma)\zeta \quad (1)$$

where T is the tool life in minutes, V is cutting speeds (m/min),

f is feed rates (mm/rev), d is depth of cut (mm)

C, α, β, γ are constants, ζ is a random error.

Equation (1) can be rewritten in the following logarithmic form:

$$\log T = \log C + \alpha \log V + \beta \log f + \gamma \log d + \log \zeta \quad (2)$$

The linear model of Eq. (2) is:

$$y = a_0x_0 + a_1x_1 + a_2x_2 + a_3x_3 + \psi \quad (3)$$

where y is the measured tool life for a logarithmic scale, $x_0 = 1$ (dummy variable),

and $x_1 = \log V, \quad x_2 = \log f, \quad x_3 = \log d,$

$\psi = \log \zeta$, where ψ is random error with zero mean and constant variance,

$a_0 = \log C,$ and $a_1, a_2,$ and a_3 are the model parameters.

The estimated response can be written as:

$$y_1 = y - \psi = b_0x_0 + b_1x_1 + b_2x_2 + b_3x_3 \quad (4)$$

where y_1 is the estimated response,

and b_0, b_1, b_2 , and b_3 are estimates of a_0, a_1, a_2 , and a_3 respectively.

The Quadratic model for machining response (tool life) is expressed as:

$$y_1 = y - \psi = b_0x_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \quad \text{-----}(5)$$

Equation (5) is useful when the two-way interactions between V, f , and d are significant for second-order effects.

The parameters of Equations (4) and (5) have been estimated by the regression analysis as per the scheme given below in Fig. 1.

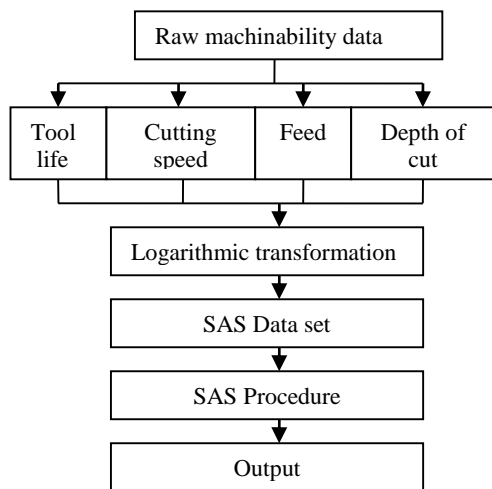


Fig. 1. Program for processing data vector in regression analysis [Vipin and Kumar, 2009]

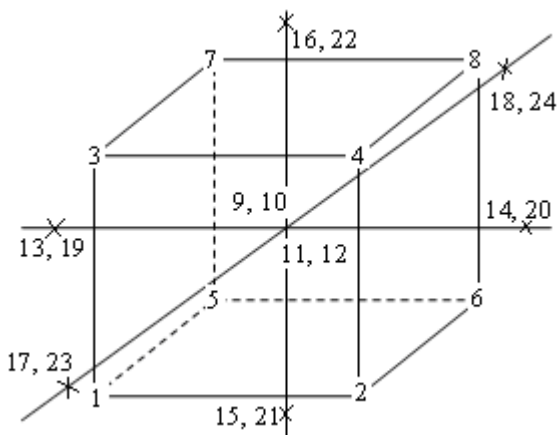


Fig.2. Representation of 23 Central design

III. EXPERIMENTAL SETUP

The first-order model was developed by an experimental design consisting of 12 experiments. Twelve experiments constitute Eight experiments (2^3 factorial designs) and Four experiments (an added centre point repeated four times) as shown in Fig.2. This was done to predict the 'b' parameters as

used in the Equation (4). The blocks provide the confidence interval of the parameters and help in the analysis of variance. A second-order model is developed by adding six augment points to the factorial design. Depending on the capacity of the machine, an augment length of ± 1 was chosen. The augment points consist of three levels for each of the independent variables denoted by -1, 0, 1. These six experiments were repeated twice to develop the second-order model. The resulting 12 or 24 experiments form the central composite design Fig. 2 shows 2^3 designs. (Vipin *et. al.* [2009])

Double column heavy vertical boring and turning machine was used to perform the experiment with a 125 hp motor shown in Fig. 3. The machine has the capability to accommodate up to 4m diameter and height up to 6m length. The present investigation was carried out on the work piece of 3.6m diameter and height of 5m. The cutting tool used was a P-40 uncoated cemented carbide insert (SNMA 25 07 24) with tool holder (PSBN 50, 50 T 25). The work piece material was cast steel (EN 6).

The experimentation was performed using cutting tool having fresh cutting edge each time under dry cutting. As the cutting edge of the insert had worn out, machine was stopped. The time of wear of cutting tool was recorded accordingly. The time between two successive tool failures and of each cut was recorded for further investigation.

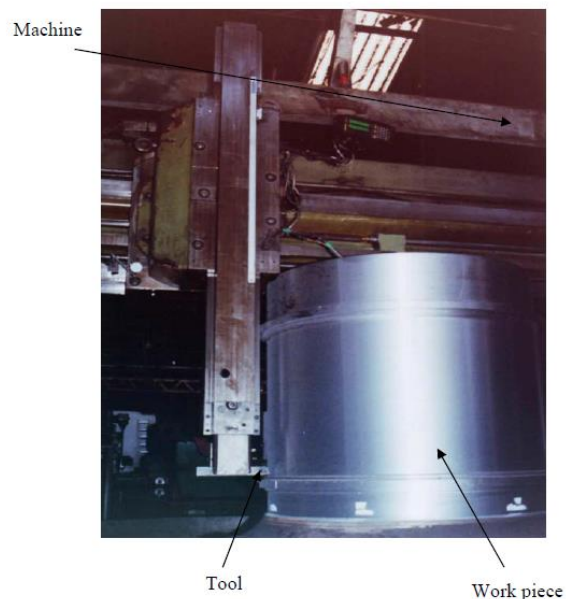


Fig.3. Double Column heavy vertical Boring and Turning Machine

Table 1. Levels of independent variables

Level	Low	Medium	High
Coding	-1	0	1
Cutting speed (m/min) V	75	90	105
Feed (mm/rev) f	0.6	0.8	1.0
Depth of cut (mm) d	6	8	10

III. RESULTS & ANALYSIS

The levels of independent variables and coding are presented in Table 1. The experimental cutting conditions together with measured tool life values are shown in Table 2.

Table 2. Experimental conditions and results

Trial No.	Coding			Tool life (min) T
	x ₁	x ₂	x ₃	
1	-1	-1	-1	42
2	1	-1	-1	21
3	-1	1	-1	33
4	1	1	-1	19
5	-1	-1	1	32
6	1	-1	1	16
7	-1	1	1	25
8	1	1	1	14
9	0	0	0	21
10	0	0	0	20
11	0	0	0	19
12	0	0	0	22
13	-1	0	0	32
14	1	0	0	17
15	0	-1	0	23
16	0	1	0	19
17	0	0	-1	26
18	0	0	1	20
19	-1	0	0	31
20	1	0	0	18
21	0	-1	0	22
22	0	1	0	21
23	0	0	-1	27
24	0	0	1	19

First-order model

The predicted tool life model for the first 12 block obtained in coded form is:

$$y' = 5.65x_0 - 1.96x_1 - 0.36x_2 - 0.56x_3 \quad (6)$$

Table 3. Analysis of variance for 12 tests

Source	DF	SS	MS	F	P
Regression	3	0.194209	0.064736	42.25	0.000
Residual Error	8	0.012257	0.001532		
Lack of Fit	5	0.010257	0.002051	3.08	0.192
Pure Error	3	0.002000	0.000667		
Total	11	0.206467			

Table 3 shows the analysis of variance with the ratio of lack of fit to pure error i.e. F-statistics is 3.08, whilst the P-statistics is 0.192. The model is adequate for determination of tool life prediction. Equation (6) has been transformed in terms of V, f and d for tool life as

$$T = 446683.592 V^{-1.96} f^{-0.36} d^{-0.56} \quad (7)$$

Metal removal rate Q (cm³/min) is given as:

$$Q = V \cdot f \cdot d \quad (8)$$

Tool life decreases with the increase of cutting speed, feed and depth of cut. Cutting speed is the dominant factor, followed by depth of cut and feed because of the respective values of exponents.

Equation 7 is utilized to develop 3D tool-life plots shown in Fig. 4(a)-(c) for different cutting speed, feed and depths of cuts. These 3D plots help to predict the tool life at any zone of the experimental domain.

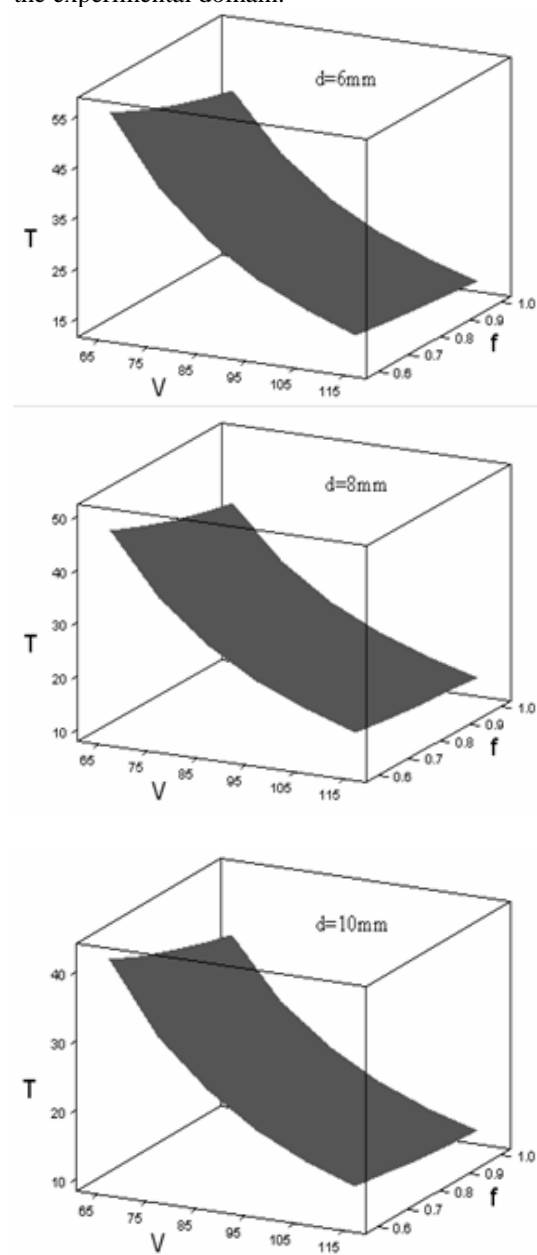


Fig. 4. 3D plots of first-order tool life model at different cutting speed and feed

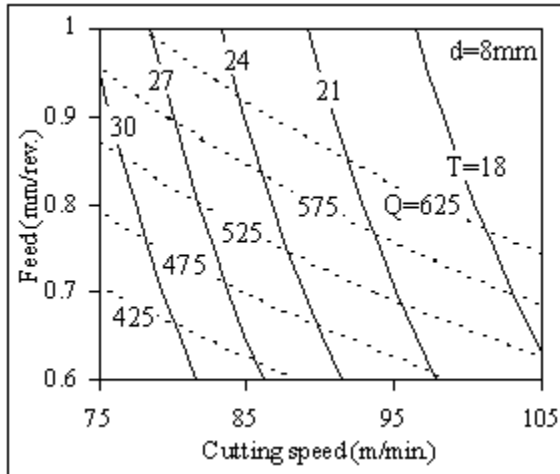


Fig. 5. First-order cutting speed- feed plane for 8.0mm depth of cut

For develop tool-life and metal removal rate contours respectively in the cutting speed–feed plane for depth of cut 8.0mm equation 7 and 8 used as shown in Fig. 5. These contours are used to predict beforehand the tool life and metal removal rate at any zone of the experimental domain. One can also use these contours for comparing various Tool-lives to predict the cutting parameters for predetermined metal removal rate.

Second-order model

First-order model was found to be adequate; the second-order model was postulated to extend the variables range in obtaining the relationship between the tool life and the machining independent variables. The model was based on the central composite design with added augment points to the nucleus of the design. The distance of the augment points was 1 unit.

The model equation is given by:

$$y' = 0.19 + 9.28x_1 - 3.54x_2 - 14.7x_3 - 2.79x_1^2 + 0.018x_2^2 + 7.52x_3^2 + 1.73x_1x_2 + 0.382x_1x_3 - 0.182x_2x_3 \quad (9)$$

Tables 4 indicates the analysis of variance for the ratio of lack of fit to pure error i.e. F-statistics is 1.28, whilst the P-statistics is 0.353. The second-order tool life was explained by the model with factors V, f and d.

Table 4. Analysis of variance for second-order model

Source	DF	SS	MS	F	P
Regression	9	0.290672	0.032297	65.31	0.000
Residual Error	14	0.006923	0.000495		
Lack of Fit	5	0.002873	0.000575	1.28	0.353
Pure Error	9	0.004050	0.000450		
Total	23	0.297596			

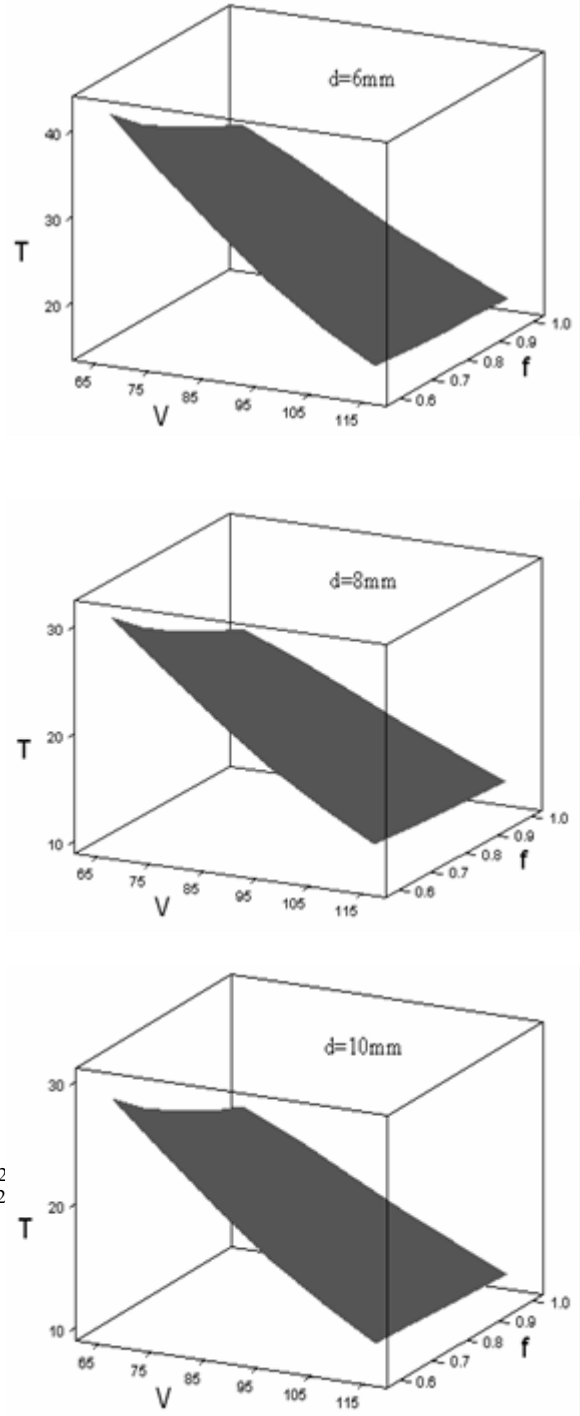


Fig. 6. 3D plots of second-order tool life model at different cutting speed and feed

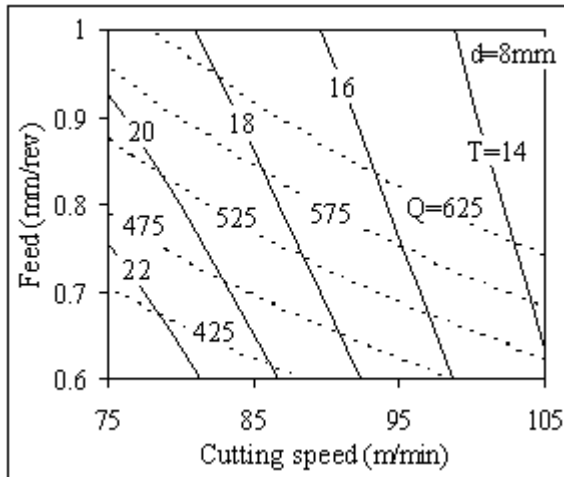


Fig. 7. Second-order cutting speed- feed plane for 8.0mm depth of cut

The model Eq. (9) is plotted in the speed-feed plane for three selected levels of depth of cut in Fig. 6(a)–(c). The contours do not show any sign of non-linearity and thereby confirm that the first-order model is adequate. The contours do not show any sign of non-linearity and thereby confirm that the first-order model is adequate for the present context. Figure 7 is a plot of the dual response of metal removal rate and tool life.

V. CONCLUSIONS

- Regression analysis has been successfully used to develop the predict tool life.
- The tool-life equation shows that the cutting speed is the main influencing factor on the tool wear, followed by depth of cut and feed in the operation model. Increasing any of these three cutting variables reduces the tool life.
- The tool-life contours/results are useful in determining the optimum cutting conditions for a given tool life.
- Dual-response contours provide useful information about the maximum attainable tool life for a given metal removal rate as a function of all three independent cutting variables.

REFERENCES

- [1] Choudhury, I.A. and El-Baradie, M.A., 1998, Tool-life prediction model by design of experiments for turning high strength steel (290 BHN), *Journal of Materials Processing Technology*, vol. 77, pp. 319–326.
- [2] Chua, M.S., Rahman, M., Wong, Y.S. and Loh, H.T., 1993, Determination of optimal cutting conditions using design of experiments and optimization techniques, *International Journal of Machine Tools & Manufacture*, vol. 33, pp. 297–305.
- [3] Hassan, G.A. and Suliman, S.M.A., 1990, Determination of Optimal Machinability Data for Steel Turning, *Journal of Institution of Engineers (India)*, vol. 71, pp. 77–82.
- [4] Lima, J.G., Abrao, A.M., Faustino, M. and Davim, J.P., 2005, Hard turning: AISI 4340 high strength low alloy steel and AISI D2 cold work tool steel, *Journal of Materials Processing Technology*, vol. 169, pp. 388–395.
- [5] Lin, Z.C. and Chan, D.Y., 1995, A study of cutting with CBN tool, *Journal of Materials Processing Technology*, vol. 49, pp. 149–164.
- [6] Ozler, L., Inan, A. and Ozel, A.I.C., 2001, Theoretical and experimental determination of tool life in hot machining of austenitic manganese steel, *International Journal of Machine Tools & Manufacture*, vol. 41, pp. 163–172.
- [7] Ranganath M. S., Vipin and Harshit, “Optimization of Process Parameters in Turning Operation Using Response Surface Methodology: A Review”, *International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459, vol. 4, iss. 10, 2014, pp. 351–360.
- [8] Ranganath M. S., Vipin, Nand Kumar and R. Srivastava, “Surface Finish Monitoring in CNC Turning Using RSM and Taguchi Techniques”, *International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459, vol. 4, iss. 9, 2014, pp. 171–179.
- [9] Santos, A.L.B.D., Duarte, M.A.V., Abrao, A.M., Machado, A.R., 1999, An optimization procedure to determine the coefficients of the extended Taylor’s equation in machining, *International Journal of Machine Tools & Manufacture*, vol. 39, pp. 17–31.
- [10] Vipin, B. B. Arora and R. S. Mishra, “Tool-Life Prediction Model for Heavy Machining of Cast Steel Material (EN 6) With Cemented Carbide Tool (P-40)”, *International Journal of Engineering Studies*, ISSN 0975- 6469, vol. 1, no. 4, 2009, pp. 269–277.
- [11] Vipin and H. Kumar, “Surface Roughness Prediction Model by Design of Experiments for Turning Lead Gun Metal”, *International Journal of Applied Engineering Research*, ISSN 0973-4562, vol. 4, no.12, 2009, pp. 2621–2628.



Resonant Ion Beam Interaction with Whistler Waves in A Magnetized Dusty Plasma

Research Article

Ruby Gupta^{1,†}, Ved Prakash^{2,*}, Suresh C. Sharma³, Vijayshri² and D.N. Gupta⁴

¹Department of Physics, Swami Shraddhanand College, University of Delhi, Alipur, Delhi 110 036, India

²School of Sciences, Indira Gandhi National Open University, Maidan Garhi, New Delhi 110 068, India

³Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110 042, India

⁴Department of Physics and Astrophysics, University of Delhi, Delhi 110 007, India

[†]Corresponding author: rubyssndu@gmail.com

Abstract. The theory of whistler wave interaction with an ion beam injected parallel to the magnetic field in an unbounded plasma is considered. The excited whistler waves propagate parallel to the beam direction and their phase velocity is a characteristic of beam-whistler resonant cyclotron coupling. The frequency and the growth rate of the unstable wave increase with the relative density of negatively charged dust grains. The ion beam velocity responsible for maximum growth rate increases as the charge density carried by dust increases. The maximum value of growth rate increases with the beam density and is proportional to the square root of beam density. These results should shed light on mechanisms of whistler wave excitation in space plasmas by artificial beams injected from spacecraft in the ionosphere and the magnetosphere.

Keywords. Whistler wave; Ion beam; Dispersion relation; Threshold velocity; Growth rate.

PACS. 52

Received: March 1, 2015

Accepted: November 26, 2015

Copyright © 2016 Ruby Gupta, Ved Prakash, Suresh C. Sharma, Vijayshri and D.N. Gupta. *This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

*Current address: India Meteorological Department, Ministry of Earth Science, Lodhi Road, New Delhi 110003, India

1. Introduction

A whistler wave is generated by a thunderstorm or lightning, and is rich in low frequency components. It propagates through the ionosphere along the Earth's magnetic field, and gets dispersed in such a way that the higher frequencies move faster than the lower ones. At various locations on the Earth there are stations that continuously record sonograms of whistler activity and provide a spectrum of the wave frequency versus time of arrival. These sonograms are used as an effective diagnostic tool for studying the ionospheric conditions. The whistler waves can be driven by electron temperature anisotropy or by charged particle beams [1–10]. A large number of papers have appeared regarding the theory of whistler excitation by a pulsed or modulated thin beam injected parallel to the magnetic field in an unbounded homogeneous magnetoplasma. Experiments performed with charged particle beams, modulated or unmodulated, have exhibited Cerenkov and cyclotron emission of whistler waves [11–13]. The whistler instability driven by an ion beam or by a ring beam is a consequence of the interaction between the solar wind and newborn ions of planetary, interstellar or cometary origin [14]. Krafft et al. [4] have studied emission of whistler waves by a density-modulated electron beam in a laboratory plasma and results have been compared to the excitation by loop antenna. Krafft et al. [12] have studied whistler wave excitation in a magnetized laboratory plasma by a density-modulated electron beam for frequency modulation below but in the range of electron cyclotron frequency. In this case, the maximum emission of the whistler waves occurred when the phase velocity of the whistler wave was equal to the beam velocity.

There has been a great deal of interest in studying waves and instabilities in dusty plasma. Dusty plasmas are found in space environments such as the lower ionosphere of the Earth, planetary atmospheres, asteroid zones, nebulae, and tails as well as in a variety of low temperature plasma devices. The presence of dust grains modifies the properties of waves in plasma [15–20]. Barkan et al. [15] have reported experimental results on the current driven electrostatic ion-cyclotron instability in a dusty plasma, where they found that the presence of negatively charged dust grains enhanced the growth rate of the instability. D'Angelo [16] has investigated the dispersion relation for low-frequency electrostatic waves in a magnetized dusty plasma. In the presence of negatively charged dust grains, he has found that the mode frequencies increased as the density ratio of negatively charged dust grains to positive ions is increased. Merlino et al. [17] have presented theoretical and experimental results on low frequency electrostatic waves in a plasma containing negatively charged dust grains and found that the presence of negatively charged dust grains modifies the properties of current driven electrostatic ion cyclotron instability through the quasineutrality condition even though the dust grains do not participate in the wave dynamics.

In this paper, we present the results of whistler wave interaction with an ion beam injected parallel to the external static magnetic field in a dusty magnetized plasma. In Section 2, we study the plasma electron, beam ion and dust grain response to whistler wave perturbation. We obtain the dispersion relation and growth rate of excited whistlers using first order perturbation theory. The variation of the growth rate of the unstable mode as a function of relative density $\delta \left(= \frac{n_{i0}}{n_{e0}} \right)$, (where n_{i0} is the ion plasma density and n_{e0} is the electron plasma density) of negatively charged dust grains has been discussed in slow cyclotron interaction. Results and discussions are given in Section 3 and conclusions are given in Section 4.

2. Instability Analysis

Consider a plasma with equilibrium electron, ion and dust particle densities as n_{e0} , n_{i0} and n_{d0} , respectively, immersed in a static magnetic field B_s in the z -direction. An ion beam with velocity $v_{b0}\hat{z}$ and density n_{b0} propagates through the plasma. In equilibrium, there is overall charge neutrality, i.e.,

$$en_{i0} + en_{b0} = en_{e0} + Q_{d0}n_{d0}.$$

We assume the t, z variations of whistler fields as $\mathbf{E}, \mathbf{B} \sim \exp[-i(\omega t - \mathbf{k} \cdot \mathbf{r})]$ and consider \mathbf{E} field to be polarized in x - z plane,

where $\mathbf{k} = k_x\hat{x} + k_z\hat{z}$ and $\omega_{ci} \ll \omega \ll \omega_{ce}$, ω_{ce} and ω_{ci} are the electron and ion cyclotron frequencies respectively.

The magnetic field of the wave is $\mathbf{B} = \frac{c}{\omega} \mathbf{k} \times \mathbf{E}$.

The equation of motion for the perturbed beam electrons is

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \frac{e}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B}_s + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right), \quad (2.1)$$

where $\mathbf{v} = v_{b0}\hat{z} + \mathbf{v}_{b1}$, \mathbf{v}_{b1} refers to perturbed velocity.

On linearizing Eq. (2.1), we obtain

$$\frac{\partial \mathbf{v}_{b1}}{\partial t} = \frac{e\mathbf{E}}{m} + \mathbf{v}_{b1} \times \hat{z}\omega_{ci}. \quad (2.2)$$

Writing x, y, z -components of Eq. (2.2), we obtain the perturbed beam electron velocities as

$$v_{b1x} = \frac{ie}{m} \frac{\bar{\omega}^2}{\omega(\bar{\omega}^2 - \omega_{ci}^2)} E_x + \frac{e}{m} \frac{\bar{\omega}\omega_{ci}}{\omega(\bar{\omega}^2 - \omega_{ci}^2)} E_y, \quad (2.3)$$

$$v_{b1y} = -\frac{e}{m} \frac{\bar{\omega}\omega_{ci}}{\omega(\bar{\omega}^2 - \omega_{ci}^2)} E_x + \frac{ie}{m} \frac{\bar{\omega}^2}{\omega(\bar{\omega}^2 - \omega_{ci}^2)} E_y, \quad (2.4)$$

$$v_{b1z} = \frac{ie}{m\bar{\omega}} E_z, \quad (2.5)$$

where $\bar{\omega} = \omega - k_z v_{b0}$.

Substituting the perturbed beam velocities from Eqs. (2.3)-(2.5) in the equation of continuity, we obtain the perturbed ion beam density as

$$n_{b1} = -\frac{n_{b0}ek_z}{im\bar{\omega}^2} E_z. \quad (2.6)$$

The response of the plasma electrons can be obtained from Eqs. (2.3)-(2.6) by replacing $e, m, \omega_{ci}, n_{b0}$ by $-e, m_e, -\omega_{ce}, n_{e0}$ respectively and putting $v_{b0} = 0$:

$$v_{e1x} = -\frac{e(i\omega E_x + \omega_{ce} E_y)}{m_e(\omega^2 - \omega_{ce}^2)}, \quad (2.7)$$

$$v_{e1y} = \frac{e(\omega_{ce} E_x - i\omega E_y)}{m_e(\omega^2 - \omega_{ce}^2)}, \quad (2.8)$$

$$v_{e1z} = \frac{eE_z}{m_e i\omega}. \quad (2.9)$$

The perturbed electron current density

$$J_{e1} = -n_{e0} e v_{e1}. \quad (2.10)$$

Substituting Eqs. (2.7), (2.8) and (2.9) in Eq. (2.10), we get

$$J_{e1x} = n_{e0} \frac{e^2}{m_e} \frac{(i\omega E_x + \omega_{ce} E_y)}{(\omega^2 - \omega_{ce}^2)}, \quad (2.11)$$

$$J_{e1y} = -n_{e0} \frac{e^2}{m_e} \frac{(\omega_{ce} E_x - i\omega E_y)}{(\omega^2 - \omega_{ce}^2)}, \quad (2.12)$$

$$J_{e1z} = -n_{e0} \frac{e^2 E_z}{m_e i\omega}. \quad (2.13)$$

The response of the dust grains can also be obtained from Eqs. (2.11)-(2.13) by replacing e , m_e , ω_{ce} , n_{e0} by Q_{d0} , m_d , ω_{cd} , n_{d0} respectively and putting $v_{bo} = 0$:

$$J_{d1x} = n_{d0} \frac{Q_{d0}^2}{m_d} \frac{(i\omega E_x + \omega_{cd} E_y)}{(\omega^2 - \omega_{cd}^2)}, \quad (2.14)$$

$$J_{d1y} = -n_{d0} \frac{Q_{d0}^2}{m_d} \frac{(\omega_{cd} E_x - i\omega E_y)}{(\omega^2 - \omega_{cd}^2)}, \quad (2.15)$$

$$J_{d1z} = -n_{d0} \frac{Q_{d0}^2 E_z}{m_d i\omega}. \quad (2.16)$$

The perturbed beam current density is given as

$$J_{b1} = -en_{b0} v_{b1} - en_{b1} v_{b0} \hat{z}. \quad (2.17)$$

Writing the x, y and z components of Eq. (2.17) and using Eqs. (2.3)-(2.6), we obtain

$$J_{b1x} = \frac{ie^2 n_{bo}}{m\omega} \frac{\bar{\omega}^2}{(\bar{\omega}^2 - \omega_{ci}^2)} E_x + \frac{e^2 n_{bo}}{m\omega} \frac{\bar{\omega} \omega_{ci}}{(\bar{\omega}^2 - \omega_{ci}^2)} E_y, \quad (2.18)$$

$$J_{b1y} = -\frac{e^2 n_{bo}}{m\omega} \frac{\bar{\omega} \omega_{ci}}{(\bar{\omega}^2 - \omega_{ci}^2)} E_x + \frac{ie^2 n_{bo}}{m\omega} \frac{\bar{\omega}^2}{(\bar{\omega}^2 - \omega_{ci}^2)} E_y, \quad (2.19)$$

$$J_{b1z} = \frac{ie^2 n_{bo}}{m\bar{\omega}^2} \omega E_z. \quad (2.20)$$

The wave equation which governs the mode structure of low frequency whistler waves is given as

$$\nabla^2 E - \nabla(\nabla \cdot E) + \left(\frac{\omega^2}{c^2}\right) E = -\frac{4\pi i\omega}{c^2} J_1. \quad (2.21)$$

Writing the x, y and z components of Eq. (2.18), we get

$$AE_x + iBE_y = 0, \quad (2.22)$$

$$AE_y - iBE_x = 0, \quad (2.23)$$

$$CE_z = 0, \quad (2.24)$$

where

$$A = -k_z^2 c^2 + \omega^2 - \frac{\omega_{pe}^2 \omega^2}{(\omega^2 - \omega_{ce}^2)} - \frac{\omega_{pd}^2 \omega^2}{(\omega^2 - \omega_{cd}^2)} - \frac{\omega_{pb}^2 \bar{\omega}^2}{(\bar{\omega}^2 - \omega_{ci}^2)},$$

$$B = \frac{\omega \omega_{ce} \omega_{pe}^2}{(\omega^2 - \omega_{ce}^2)} + \frac{\omega \omega_{cd} \omega_{pd}^2}{(\omega^2 - \omega_{cd}^2)} + \frac{\bar{\omega} \omega_{ce} \omega_{pb}^2}{(\bar{\omega}^2 - \omega_{ci}^2)}, \quad C = 1 - \frac{\omega_{pe}^2}{\omega^2} - \frac{\omega_{pd}^2}{\omega^2} - \frac{\omega_{pb}^2}{\bar{\omega}^2}$$

$$\omega_{pb}^2 = \frac{4\pi n_{b0} e^2}{m}, \quad \omega_{pe}^2 = \frac{4\pi n_{e0} e^2}{m_e} \text{ and } \omega_{pd} = \left(\frac{4\pi n_{d0} Q_{d0}^2}{m_d} \right)^{\frac{1}{2}}.$$

A non-trivial solution of Eqs. (2.22)-(2.24) demands that the determinant of coefficients of E_x and E_y must vanish, i.e., $A^2 - B^2 = 0$.

It gives two distinct modes of wave propagation. We examine here right-hand polarized electromagnetic whistler waves where the ions' contributions have been neglected due to the range of frequencies considered, obtained from $A - B = 0$, or

$$k_z^2 - \frac{\omega^2}{c^2} \varepsilon = -\frac{\omega_{pb}^2 \bar{\omega}}{c^2 (\bar{\omega} + \omega_{ci})}, \quad (2.25)$$

$$\text{where } \varepsilon = 1 - \frac{\omega_{pe}^2}{\omega(\omega - \omega_{ce})} - \frac{\omega_{pd}^2}{\omega(\omega - \omega_{cd})}.$$

Eq. (2.25) can be written as

$$\left(k_z - \frac{\omega}{c} \sqrt{\varepsilon} \right) \left(k_z + \frac{\omega}{c} \sqrt{\varepsilon} \right) (\bar{\omega} + \omega_{ci}) = -\frac{\omega_{pb}^2 \bar{\omega}}{c^2}, \quad \text{or}$$

$$\left(k_z - \frac{\omega}{c} \sqrt{\varepsilon} \right) \left(k_z + \frac{\omega}{c} \sqrt{\varepsilon} \right) \left(k_z - \frac{\omega + \omega_{ci}}{v_{bo}} \right) = -\frac{\omega_{pb}^2}{c^2} \left(k_z - \frac{\omega}{v_{bo}} \right). \quad (2.26)$$

Eq. (2.26) gives the dispersion relation of parallel whistler waves in a dusty plasma in the absence of beam given as

$$k_z = \frac{\omega}{c} \left[1 - \frac{\omega_{pe}^2}{\omega(\omega - \omega_{ce})} - \frac{\omega_{pd}^2}{\omega(\omega - \omega_{cd})} \right]^{\frac{1}{2}}. \quad (2.27)$$

Assuming perturbed quantities $k_z = \frac{\omega}{c} \sqrt{\varepsilon} + \Delta$ and $k_z = \frac{\omega + \omega_{ci}}{v_{bo}} + \Delta$ in Eq. (2.26), we get

$$\Delta^2 = -\frac{\omega_{pb}^2 \omega_{ci}}{2c\omega \sqrt{\varepsilon} v_{bo}}.$$

Therefore, Growth rate $\gamma = \text{Im}(\Delta)$ or

$$\gamma = \left[\frac{\omega_{pb}^2}{2c\omega \sqrt{\varepsilon}} \frac{\omega_{ci}}{v_{bo}} \right]^{\frac{1}{2}}. \quad (2.28)$$

It may be noted from Eq. (2.25), that Cerenkov or fast cyclotron interaction between parallel beam and parallel whistlers is not possible, as there is only slow cyclotron interaction term on right hand side of Eq. (2.25). The growth rate depends on the external static magnetic field, unstable whistler frequency, plasma electron density, dust grain density, beam velocity and beam density. It can be inferred from Eq. (2.26) that the whistlers exhibit a growth rate only if the beam velocity is greater than their phase velocity.

3. Results and Discussion

The parameters used in the present calculations are: ion plasma density $n_{i0} = 10^{12} \text{ cm}^{-3}$, electron plasma density $n_{e0} = 10^{12} - 0.2 \times 10^{12} \text{ cm}^{-3}$, magnetic field $B_S = 300 \text{ G}$, mass of ion $m_i = 39 \times 1836 m_e$ (potassium-plasma), mass of dust grain $m_d = 10^{12} m_i$, dust grain density $n_{d0} = 10^8 \text{ cm}^{-3}$, beam density $n_{b0} = 10^9 \text{ cm}^{-3}$ and beam velocity $v_{b0} = 1.2 \times 10^9 \text{ cm/s}$. The relative density of negatively charged dust grains $\delta \left(= \frac{n_{i0}}{n_{e0}} \right)$ has been varied from 1 to 5. In Figure 1, we have plotted the dispersion curves of whistler waves in a dusty plasma, for different values of δ . We have also plotted the beam mode via slow cyclotron interaction for an ion beam travelling inside the dusty plasma. The velocity of the beam is chosen in such a way so that they intersect the dispersion curves of whistler waves in the required frequency range $\omega_{ci} \ll \omega \ll \omega_{ce}$. The frequencies and the corresponding wave numbers of the unstable mode obtained from the points of intersection between the beam mode and the electron whistler modes are given as Table 1. The unstable frequencies of the whistler waves in presence of dust grains and the parallel wave vector k_z decrease with increase in relative density of negatively charged dust grains δ .

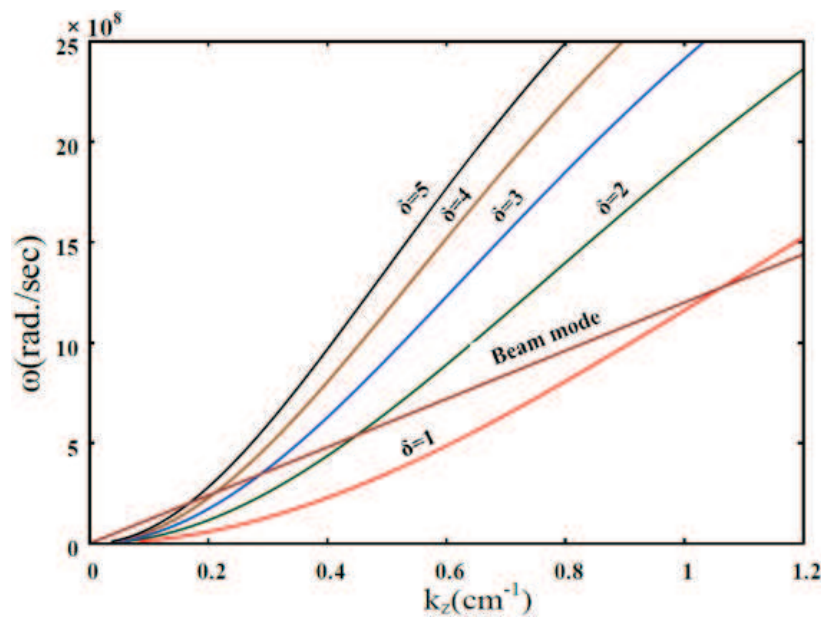
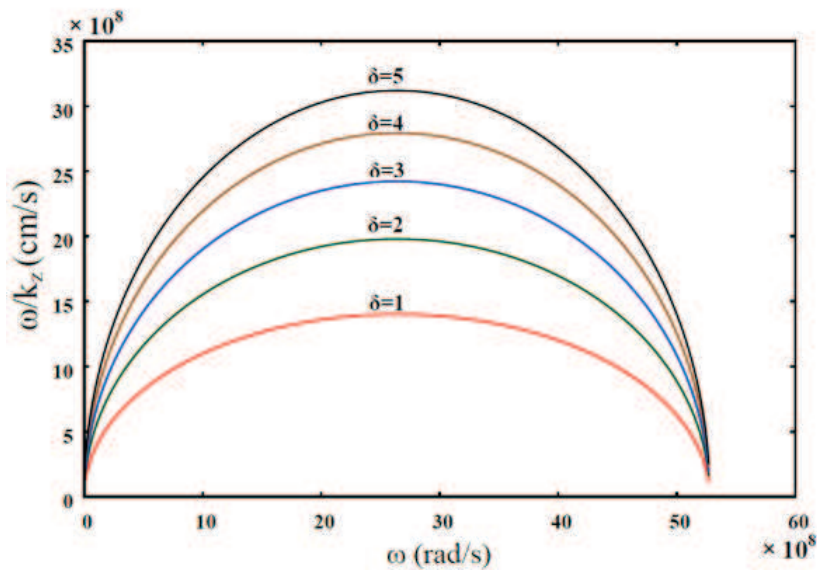


Figure 1. Dispersion curves of whistler waves for different values of $\delta \left(= \frac{n_{i0}}{n_{e0}} \right)$ and a beam mode via slow cyclotron interaction.

Table 1. Unstable wave frequencies ω (rad./s) and axial wave numbers k_z (cm^{-1}) for different values of δ from Figure 1.

δ	k_z (cm^{-1})	ω (rad./sec) $\times 10^8$
1	1.06	12.79
2	0.45	5.40
3	0.29	3.53
4	0.21	2.61
5	0.17	2.09

The phase velocity of the whistler waves has been plotted as a function of parallel whistler frequency in Figure 2. It can be seen that, as the whistler frequency increases, the phase velocity first increases then remains constant for certain frequency range and then decreases. The phase velocity of the whistlers is found to increase from 14.1×10^8 cm/s to 31.4×10^8 cm/s as δ increases from 1 to 5, indicating that the addition of dust in plasma effects the propagation of whistler waves.

**Figure 2.** Phase velocity of whistler waves as a function of unstable whistler frequency for different values of δ .

When the frequency is much smaller than the electron cyclotron frequency, the phase velocity increases i.e., the higher frequencies move faster than the lower ones and ascending frequency whistlers are observed, which produce an ascending pitch whistle. When the frequency is near (but smaller than) the electron cyclotron frequency, the phase velocity starts decreasing with whistler frequency and the low frequency whistlers reach the Earth's surface before the high frequency ones, giving out a descending pitch tone. The frequency regime in which the ascending tone turns into a descending tone decreases with an increase in the value of δ . The phase velocity for all the values of δ shows maxima at a whistler frequency of 26.3×10^8 rad/s.

The growth rate of the whistler waves has been plotted using Eq. (2.28), as a function of δ in Figure 3. The unstable wave frequencies and wave numbers of the whistler waves for plotting the growth rate of Figure 3, have been used from Table 1. From Figure 3, it can be seen that the growth rate of the unstable mode increases with δ . The growth rate of the unstable mode also increases with the beam density and is proportional to the square root of beam density. The beam velocity needed to excite the whistler instability increases with an increase in the value of external static magnetic field. It also depends on the plasma density, and a higher beam velocity is required for interaction with whistler waves in denser plasma.

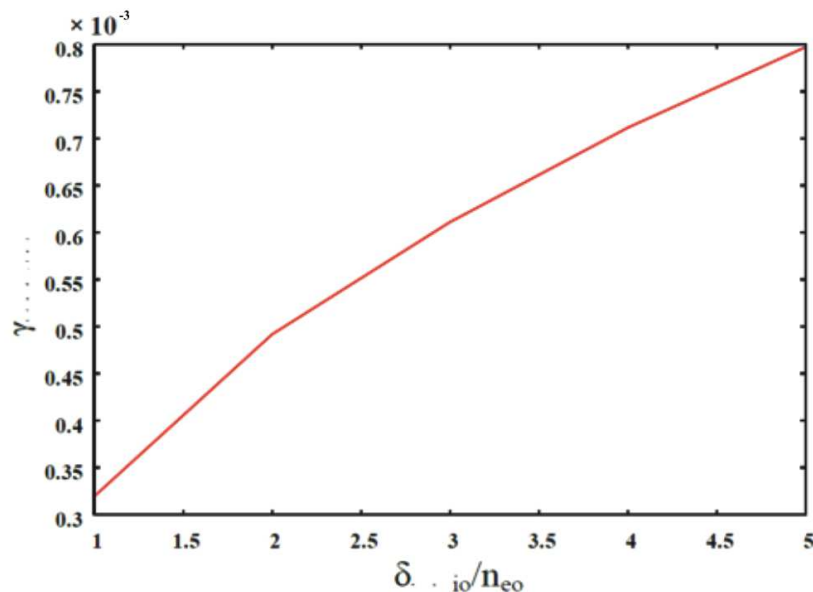


Figure 3. Growth rate δ (rad./sec) of the unstable whistler mode in a dusty plasma via slow cyclotron interaction as a function of $\delta \left(= \frac{n_{i0}}{n_{e0}} \right)$.

4. Conclusion

In this paper, we study the excitation of whistler waves via slow cyclotron interaction with an ion beam in magnetized plasma. The dispersion relation of the whistler modes has been derived for beam-whistler interaction and the standard relation is retrieved in the absence of beam. Cerenkov and fast cyclotron interaction between ion beam and parallel whistlers has not been observed analytically. The waves show a spatiotemporal growth rate, which depends on the beam number density, beam velocity, plasma number density and external static magnetic field.

The main physical process occurring during beam-wave interaction is that the beam ions bunch along the magnetic field, which are continuously accelerated or decelerated while keeping resonance with the emitted wave. The bunches are the main cause which support the wave emission whereas the nonresonant beam ions practically do not exchange energy with the wave. All the loss of the resonant beam particles' energy is transformed into emitted wave energy and the wave grows. The unstable frequency and the growth rate are sensitive to the beam velocity. The results presented here should be useful for the understanding of wave activities in space and laboratory plasmas.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed equally and significantly in writing this article. All the authors read and approved the final manuscript.

References

- [1] H.A. Shah and V.K. Jain, *J. Plasma Phys.* **31** (1984), 22,.
- [2] L. James, L. Jassal and V.K. Tripathi, *J. Plasma Phys.* **54** (1995), 119.
- [3] R.C. Borcia, G. Matthieussent, E.L. Bel, F. Simonet and J. Solomon, *Phys. Plasmas* **7** (2000), 359.
- [4] C. Krafft, G. Matthieussent, P. Thevenet and S. Bresson, *Phys. Plasmas* **1** (1994), 2163.
- [5] I. Talukdar, V.K. Tripathi and V.K. Jain, *J. Plasma Phys.* **41** (1989), 231.
- [6] A. Volokitin, C. Krafft and G. Matthieussent, *Phys. Plasmas* **2** (1995), 4297.
- [7] K. Sauer and R.D. Sydora, *Ann. Geophys.* **28** (2010), 1317.
- [8] J.W. Cipolla, K.I. Golden, and M.B. Silevitch, *Phys. Fluids* **20** (1977), 282.
- [9] M.A. Dorf, I.D. Kaganovich, E.A. Startsev and R.C. Davidson, *Phys. Plasmas* **17** (2010), 23103.
- [10] C. Krafft, and A. Volokitin, *Phys. Plasmas* **5** (1998), 4243.
- [11] N. Baranets, Y. Ruzhin, N. Erokhin, V. Afonin, J. Vojta, J. Smilauer, K. Kudela, J. Matisin and M. Ciobanu, *Advances in Space Res.* **49** (2012), 859.
- [12] C. Krafft, P. Thevenet, G. Matthieussent, B. Lundin, G. Belmont, B. Lembege, J. Solomon, Lavergnat and T. Lehner, *Phy. Rev. Lett.* **72** (1994), 649.
- [13] M. Starodubtsev, C. Krafft, B. Lundin and P. Thevenet, *Phys. Plasmas* **6** (1999), 2862.
- [14] K. Akimoto, and D. Winske, *J. of Geophys. Res.* **94** (1989), 17259.
- [15] R. Barkan, N. D'Angelo and R.L. Merlino, *Planet. Space Sci.* **43** (1995), 905.
- [16] N. D'Angelo, *Planet. Space Sci.* **38** (1990), 1143.
- [17] R.L. Merlino, A. Barkan, C. Thompson and N. D'Angelo, *Phys. Plasmas* **5** (1998), 1607.
- [18] V. Prakash and S.C. Sharma, *Phys. Plasmas* **16** (2009), 93703.
- [19] V. Prakash, S.C. Sharma, Vijayshri and R. Gupta, *Phys. Plasmas* **21** (2014), 33701.
- [20] V. Prakash, S.C. Sharma, Vijayshri and R. Gupta, *Laser Part. Beams* **31** (2013), 411.

Statefinder diagnosis for holographic dark energy models in modified $f(R, T)$ gravity

C.P. Singh¹ · Pankaj Kumar¹

Received: 18 January 2016 / Accepted: 31 March 2016
© Springer Science+Business Media Dordrecht 2016

Abstract In this paper we consider the non-viscous and viscous holographic dark energy models in modified $f(R, T)$ gravity in which the infra-red cutoff is set by the Hubble horizon. We find power-law and exponential form of scale factor for non-viscous and viscous models, respectively. It is shown that the Hubble horizon as an infra-red cut-off is suitable for both the models to explain the recent accelerated expansion. In non-viscous model, we find that there is no phase transition. However, viscous model explains the phase transition from decelerated phase to accelerated phase. The cosmological parameters like deceleration parameter and statefinder parameters are discussed to analyze the dynamics of evolution of the Universe for both the models. The trajectories for viscous model are plotted in r - s and r - q planes to discriminate our model with the existing dark energy models which show the quintessence like behavior.

Keywords Cosmology · Dark energy · Modified gravity theory

1 Introduction

It is strongly believed that the Universe has entered a phase of the accelerated expansion which has been confirmed by the recent observations like supernovae Ia (Riess et al. 2007; Suzuki et al. 2012), cosmic microwave background radiation (Bennett et al. 2003; Komatsu et al. 2011), baryon acoustic

oscillation (Percival et al. 2010) and Planck data (Ade et al. 2014). Within the framework of general relativity (GR), the cause of the acceleration can be attributed to the existence of a mysterious component of the Universe dubbed as “dark energy” (DE), which makes up ~ 70 % of the total cosmic energy in the Universe. The Λ -cold dark matter (Λ CDM) model presents the simplest and most successful description of the recent accelerated expansion scenario and accommodates the observations very well. Despite of many attractive features, it has some theoretical problems like fine-tuning and cosmic coincidence problems (Carroll 2001; Peebles and Ratra 2003; Padmanabhan 2003). To overcome from these problems, a number of dynamical dark energy models such as scalar field (quintessence, phantom, k -essence, etc.) models (Chiba et al. 2000; Caldwell 2002; Padmanabhan 2002; Copeland et al. 2006), chaplygin gas models (Bento et al. 2002), holographic dark energy (HDE) models (Setare 2007; Xu 2009; Sheykhi and Jamil 2011), etc. have been explored in the literature.

In the recent years, the HDE models have been emerged as a viable candidates to explain the problems of modern cosmology. The HDE models explain the recent accelerated expansion as well as the coincidence problem of the Universe (Li 2004; Pavón and Zimdahl 2005). The concept of HDE is based on the holographic principle proposed by 't Hooft (1993) and found it's roots in the quantum field theory. Cohen et al. (1999) have shown that in the quantum field theory, the formation of black hole sets a limit which relates ultra-violet (UV) cut-off length γ to infra-red (IR) cut-off length L . According to the authors, the quantum zero-point energy $\rho_h = \gamma^4$ of a system of size L should not exceed the mass of a black hole of the same size, i.e., $L^3 \rho_h \leq LM_p^2$, where $M_p = (8\pi G)^{-1/2}$ is the reduced Planck mass. In a paper, Li (2004) has taken the largest allowed L to saturate this inequality and thus obtained dark energy density of the

✉ C.P. Singh
cpsphd@rediffmail.com

P. Kumar
pankaj.11dtu@gmail.com

¹ Department of Applied Mathematics, Delhi Technological University (Formerly Delhi College of Engineering), Bawana Road, Delhi 110 042, India

Universe $\rho_h = 3c^2 M_p^2 L^{-2}$, known as HDE density. In the formalism of HDE, the Hubble horizon is a most natural choice for the IR cut-off, but it leads to a wrong equation of state (EoS) of dark energy (Hsu 2004). However, Pavón and Zimdahl (2005), and Banerjee and Pavón (2007) have shown that the viable EoS of dark energy could be achieved by taking the interaction between HDE and dark matter (DM).

On the other hand, the modified theories of gravity such as $f(R)$ gravity (Bamba et al. 2008; Nojiri and Odintsov 2011), $f(G)$ gravity (Nojiri and Odintsov 2005; De Felice and Tsujikawa 2009), $f(R, G)$ gravity (Bamba et al. 2010; Elizalde et al. 2010), etc. have also been proposed to explain the recent accelerated expansion of the Universe. The $f(R)$ gravity is one of the simplest and successful modified theories of GR, which fits with the observations very well. Recently, Harko et al. (2011) have proposed a new modified gravity theory known as $f(R, T)$ gravity, where R as usual stands for the Ricci scalar and T denotes the trace of energy-momentum tensor. This modified theory presents a maximal coupling between geometry and matter. A number of authors (Sharif and Zubair 2012; Chakraborty 2013; Harko 2014; Singh and Singh 2014; Baffou et al. 2015; Shabani and Farhoudi 2014) have discussed the modified $f(R, T)$ gravity in different context to explain the early and late time acceleration of the Universe. In a recent paper (Singh and Kumar 2014), the authors have discussed the viscous cosmology in this theory which shows the recent phase transition of the Universe. The HDE models have not been yet discussed in detail in the framework of $f(R, T)$ gravity. In some papers (Houndjo and Piattella 2012; Fayaj et al. 2014), reconstruction of $f(R, T)$ gravity from HDE and anisotropic model of HDE have been discussed.

Hsu (2004) in GR and Xu et al. (2009) in Brans-Dicke theory have shown that the Hubble horizon as an IR cut-off is not a suitable candidate to explain the recent accelerated expansion. However, Pavón and Zimdahl (2005) in GR, and Banerjee and Pavón (2007) in Brans-Dicke theory have shown that the interaction between HDE and DM can change the scenario and Hubble horizon as an IR cut-off may explain the recent accelerated expansion. In this paper our interest is to study the HDE model with Hubble horizon as an IR cut-off in $f(R, T)$ gravity without considering the interaction between HDE and DM. As it is known, the $f(R, T)$ gravity has coupling between geometry and matter, therefore, it will be interesting to discuss HDE with Hubble horizon as an IR cut-off in this modified theory. We show that Hubble horizon as an IR cut-off is suitable to explain the accelerated expansion in this theory without interaction between HDE and DM.

The Hubble parameter H and the deceleration parameter q are well known cosmological parameters which explain the evolution of the Universe. However, these two parameters can not discriminate among various DE models.

In this context, Sahni et al. (2003) and Alam et al. (2003) have introduced a new geometrical diagnostic pair $\{r, s\}$, known as statefinder parameters, which is constructed from the scale factor and its derivatives up to the third order. The statefinder pair $\{r, s\}$ is geometrical in the nature as it is constructed from the space-time metric directly. Therefore, the statefinder parameters are more universal parameters to study the DE models than any other physical parameters. In a flat Λ CDM model, the statefinder pair has a fixed point value $\{r, s\} = \{1, 0\}$. One can plot the trajectories in r - s and r - q planes to discriminate various DE models. We discuss the statefinder diagnostic and obtain the fixed point values of statefinder pair $\{r, s\} = \{1, 0\}$ as in the case of Λ CDM model.

To be more realistic, the perfect fluid Universe is just an approximation of the viscous Universe. The dissipative processes in the relativistic fluid may be modeled as bulk viscosity. Misner (1968) was the first to use the viscosity concept in cosmology. The origin of the bulk viscosity in a physical system can be traced to deviations from the local thermodynamic equilibrium (see Maartens 1996; Brevik and Grøn 2013). In a cosmological fluid, the bulk viscosity arises any time when a fluid expands (or contracts) too rapidly so that the system does not have enough time to restore the local thermodynamic equilibrium (Avelino and Nucamendi 2009). The bulk viscosity, therefore, is a measure of the pressure required to restore equilibrium. When the fluid reaches again the thermal equilibrium then the bulk viscous pressure ceases (Xinzhang and Spiegel 2001). Therefore, sufficient large bulk viscous pressure could make the effective pressure negative. Thermodynamic states with negative pressure are metastable and cannot be excluded by any law of nature. These states are connected with phase transitions. A phase transition in viscous early universe was discussed by Tawfik and Harko (2012).

In Weinberg (1972), the theoretical concept of the bulk viscosity in cosmology has been discussed which provides the insight into the nature of the bulk viscosity. Within the context of the early inflation, it has been known since long time ago that an imperfect fluid with bulk viscosity can produce an accelerated expansion without the need of a cosmological constant or some other inflationary scalar field (Heller et al. 1973; Zimdahl 1996). An accelerating Universe can be achieved for the right viscosity coefficient. At the late time, since we don't know the nature of the Universe's contents (dark matter and dark energy components) very clearly, the bulk viscosity is reasonable and can play a role as a dark energy candidate. Therefore, it is natural to consider the bulk viscosity in an accelerating Universe. It has been shown that inflation and recent acceleration can be explained using the viscous behavior of the Universe, and plays an important role in the phase transition of the Universe (Murphy 1973; Padmanabhan and Chitre 1987;

Brevik and Gorbunova 2005; Hu and Meng 2006; Singh et al. 2007; Wilson et al. 2007; Kumar and Singh 2015; Sasidharan and Mathew 2015).

The concept of viscous DE has been discussed extensively in the literature (Cataldo et al. 2005; Sebastiani 2010; Setare and Sheykhi 2010). Feng and Li (2009) have shown that the age problem of the Ricci dark energy can be alleviated using the bulk viscosity. Motivated by the above works, we extend our analysis to viscous HDE with the same IR cut-off which shows the recent phase transition of the Universe. We obtain the statefinder parameters for viscous HDE which achieve the value of Λ CDM model and show the quintessence like behavior.

The paper is organized as follows. In the next section we discuss the formalism of $f(R, T)$ gravity theory and present its field equations. In Sect. 3 we discuss the non-viscous HDE model and find the exact power-law solution of the scale factor which avoids the big bang singularity. We also find the cosmological parameters like deceleration parameter and statefinder parameters and discuss their behaviors. Section 4 describes the viscous HDE model and its solution. Section 5 presents the summary of our findings.

2 The formalism of modified $f(R, T)$ gravity theory

The general form of the Einstein-Hilbert action for the modified $f(R, T)$ gravity in the unit $8\pi G = 1$ is as follows (Harko et al. 2011; Singh and Kumar 2014):

$$S = \frac{1}{2} \int d^4x \sqrt{-g} [f(R, T) + 2\mathcal{L}_m], \quad (1)$$

where g stands for the determinant of the metric tensor $g_{\mu\nu}$, R is the Ricci scalar and T represents the trace of the energy-momentum tensor, i.e., $T = T^\mu_\mu$, while \mathcal{L}_m denotes the matter Lagrangian density. The speed of light is taken to be unity. As usual the energy-momentum tensor, $T_{\mu\nu}$ of matter is defined as

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta(\sqrt{-g}\mathcal{L}_m)}{\delta g^{\mu\nu}}. \quad (2)$$

In fact, this modified theory of gravity is the generalization of $f(R)$ gravity and is based on the coupling between geometry and matter. The corresponding field equations have been derived in metric formalism for the various forms of $f(R, T)$.

Varying the action (1) with respect to the metric tensor $g_{\mu\nu}$ for a simple form of $f(R, T) = R + f(T)$, i.e., the usual Einstein-Hilbert term plus an $f(T)$ correction (Harko et al. 2011; Singh and Kumar 2014) which modifies the general relativity and represents a coupling with geometry of the

Universe, we get the following field equations.

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = T_{\mu\nu} - (T_{\mu\nu} + \ominus_{\mu\nu})f'(T) + \frac{1}{2}f(T)g_{\mu\nu}, \quad (3)$$

where a prime denotes derivative with respect to the argument. The tensor $\ominus_{\mu\nu}$ in (3) is given by

$$\ominus_{\mu\nu} = -2T_{\mu\nu} + g_{\mu\nu}\mathcal{L}_m - 2g^{\alpha\beta} \frac{\partial^2 \mathcal{L}_m}{\partial g^{\mu\nu} \partial g^{\alpha\beta}}. \quad (4)$$

The matter Lagrangian \mathcal{L}_m may be chosen as $\mathcal{L}_m = -p$ (Harko et al. 2011), where p is the thermodynamical pressure of matter content of the Universe. Now, Eq. (4) gives $\ominus_{\mu\nu} = -2T_{\mu\nu} - pg_{\mu\nu}$. Using this result, Eq. (3) reduces to

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = T_{\mu\nu} + (T_{\mu\nu} + pg_{\mu\nu})f'(T) + \frac{1}{2}f(T)g_{\mu\nu}, \quad (5)$$

which are the field equations of the modified $f(R, T)$ gravity theory.

Here, we are interested to study the behavior of HDE in this modified theory for a spatially homogeneous and isotropic flat Friedmann-Robertson-Walker (FRW) space-time, which is expressed in comoving coordinates by the line element,

$$ds^2 = dt^2 - a^2(t)(dx^2 + dy^2 + dz^2), \quad (6)$$

where $a(t)$ stands for the cosmic scale factor. In what follows we study the non-viscous and viscous HDE models with deceleration parameter and statefinder parameters in $f(R, T)$ gravity theory to describe the recent acceleration.

3 Non-viscous holographic dark energy cosmology

In this model, let us consider the Universe filled with HDE plus pressureless DM (excluding baryonic matter), i.e.,

$$T_{\mu\nu} = T_{\mu\nu}^h + T_{\mu\nu}^m, \quad (7)$$

where $T_{\mu\nu}^h$ and $T_{\mu\nu}^m$ represent the energy-momentum tensors of HDE and DM, respectively. Many authors have described the recent accelerated expansion by assuming the interaction between HDE and DM in the different theories of gravity. In this paper, we consider a non-interacting matter field in $f(R, T)$ gravity. The generalized Einstein Eqs. (5) yield

$$3H^2 = \rho_m + \rho_h + (\rho_m + \rho_h + p_h)f'(T) + \frac{1}{2}f(T), \quad (8)$$

$$2\dot{H} + 3H^2 = -p_h + \frac{1}{2}f(T), \quad (9)$$

where ρ_m , ρ_h and p_h denote the energy density of DM, the energy density of HDE and the pressure of HDE, respectively. An overdot denotes the derivative with respect to cosmic time t . As the field equations (8) and (9) are highly non-linear, therefore, we assume $f(T) = \alpha T$ (see, Harko et al. 2011), where α is a coupling parameter. Now, the field equations (8) and (9) reduce as

$$3H^2 = \rho_m + \rho_h + \alpha(\rho_m + \rho_h + p_h) + \frac{1}{2}\alpha T, \quad (10)$$

$$2\dot{H} + 3H^2 = -p_h + \frac{1}{2}\alpha T. \quad (11)$$

The equation of state (EoS) of HDE is given by $p_h = w_h \rho_h$ and the trace of energy-momentum tensor is given by $T = \rho_m + \rho_h - 3p_h$. Now, from (10) and (11), a combined evolution equation for H can be written as

$$2\dot{H} + (1 + \alpha)[(1 + w_h)\rho_h + \rho_m] = 0. \quad (12)$$

In the literature, various forms of HDE (the general form is $\rho_h = 3c^2 M_p^2 L^{-2}$, where c^2 is a dimensionless constant, M_p stands for the reduced Planck mass and L denotes the IR cut-off radius) have been discussed depending on the choices of IR cut-off such as Hubble horizon, future event horizon, apparent horizon, Granda-Oliveros cut-off, etc. In this work, we consider the Hubble horizon ($L = H^{-1}$) as an IR cut-off to describe the recent acceleration. The corresponding energy density ρ_h is given by

$$\rho_h = 3c^2 H^2. \quad (13)$$

Form (10) and (13), the energy density of DM can be obtained as

$$\rho_m = \frac{3(\alpha c^2 w_h - 3\alpha c^2 - 2c^2 + 2)}{(3\alpha + 2)} H^2. \quad (14)$$

Using (13) and (14) into (12), we finally get

$$\dot{H} + \frac{3(\alpha + 1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha + 2} H^2 = 0, \quad (15)$$

which, on solving gives (where $\alpha \neq -1$, $\alpha \neq -2/3$)

$$H = \frac{1}{c_0 + \frac{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} t}, \quad (16)$$

where c_0 is a positive constant of integration. Equation (16) can be rewritten as

$$H = \frac{H_0}{1 + \frac{3H_0(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} (t - t_0)}, \quad (17)$$

where H_0 is the present value of the Hubble parameter at the cosmic time $t = t_0$, the time where the HDE starts to

dominate. Using the relation $H = \frac{\dot{a}}{a}$, the cosmic scale factor a is given by

$$a = c_1 \left[1 + \frac{3H_0(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} \times (t - t_0) \right]^{\frac{3\alpha+2}{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}}, \quad (18)$$

where c_1 is an another positive constant of integration. One can rewrite a as follows

$$a = a_0 \left[1 + \frac{3H_0(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} \times (t - t_0) \right]^{\frac{3\alpha+2}{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}}, \quad (19)$$

where a_0 is the present value of the scale factor at the cosmic time $t = t_0$. We obtain the power-law evolution of the Universe which avoids the big-bang singularity.

The deceleration parameter q , which is defined as $q = -a\ddot{a}/\dot{a}^2$, is a geometric parameter which describes the acceleration or deceleration of the Universe depending on its negative or positive value. In this case, the deceleration parameter is given by

$$q = \frac{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} - 1. \quad (20)$$

Here, we obtain a constant deceleration parameter as expected due to the power-law scale factor. The accelerated expansion may be obtained if the parameters satisfy the constraint $\frac{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} < 1$. In a paper, Li et al. (2013) have studied the Planck constraints on HDE and obtained the tightest and self-consistent value of constant c from Planck + WP + BAO + HST + lensing as $c = 0.495 \pm 0.039$. Therefore, let us consider here and thereafter $c = 0.5$ for further discussion which is lying in this observed range. Now, for example let's take $\alpha = 1$ and $w_h = -0.5$, we obtain $q = -0.25$ which shows the accelerated expansion. Moreover, one may obtain the accelerated expansion even when the HDE does not violate the strong energy condition $\rho_h + 3p_h > 0$ for suitable values of coupling parameter α . Thus, the HDE with Hubble horizon as an IR cut-off can successfully explain the accelerated expansion in the framework of $f(R, T)$ gravity without assuming the interaction between HDE and DM in contrast to the works done in GR (Hsu 2004) and in Brans-Dicke theory (Xu et al. 2009). It is to be noted that this model does not show the phase transition as the deceleration parameter is constant.

In order to get a robust analysis to discriminate among DE models Sahni et al. (2003) and Alam et al. (2003) have introduced a new geometrical diagnostic pair $\{r, s\}$, known as statefinder parameters, which is constructed from the

scale factor and its derivatives up to the third order. The statefinder pair $\{r, s\}$ is geometrical in nature as it is constructed from the space-time metric directly. The statefinder pair $\{r, s\}$ provides a very comprehensive description of the dynamics of the Universe and consequently the nature of the DE. It is defined as

$$r = \frac{\ddot{a}}{aH^3}, \quad s = \frac{r-1}{3(q-1/2)}. \quad (21)$$

In this model, we obtain the statefinder parameters r and s as

$$r = \frac{18(\alpha+1)^2(2\alpha c^2 w_h + c^2 w_h + 1)^2}{(3\alpha+2)^2} - \frac{9(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{3\alpha+2} + 1, \quad (22)$$

$$s = \frac{(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1)}{(3\alpha+2)[(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1) - \frac{1}{2}]} \times [2(\alpha+1)(2\alpha c^2 w_h + c^2 w_h + 1) - (3\alpha+2)]. \quad (23)$$

We observe that the statefinder parameters $\{r, s\}$ are constant and the values of these parameters depend on the coupling parameter α , constant c and EoS parameter w_h of HDE. In the papers (Sahni et al. 2003; Alam et al. 2003), it has been observed that Λ CDM model and Λ CDM model have fixed point values of statefinder pair $\{r, s\} = \{1, 1\}$ and $\{r, s\} = \{1, 0\}$, respectively. In our work, it is observed that fixed point of Λ CDM model can be achieved for $\alpha = -\frac{1+c^2 w_h}{2c^2 w_h}$ as a particular case of this model. Thus, for a suitable value of w_h , which may be obtained by observations, we can find the coupling parameter α for which $\{r, s\} = \{1, 0\}$ and viceversa.

4 Viscous holographic dark energy cosmology

In the non-viscous HDE model, we have obtained the constant value of deceleration parameter which does not describe the phase transition. But, the astronomical observations show that the phase transition is an integral part of the evolution of the Universe. Therefore, in this section, it would be of interest to investigate whether a viscous HDE with the Hubble horizon as an IR cut-off could be helpful to explain the phase transition, i.e., time-dependent deceleration parameter with signature flip from positive to negative in order to elucidate the observed phase transition of the Universe.

In an accelerating Universe, it may be natural to assume that the expansion process is actually a collection of state out of thermal equilibrium in a small fraction of time due to the existence of possible dissipative mechanisms. In an isotropic and homogeneous FRW model, the dissipative process may

be treated via the relativistic theory of bulk viscosity proposed by Eckart (1940) and later on pursued by Landau and Lifshitz (1987). It has been found that only the bulk viscous fluid remains compatible with the assumption of large scale homogeneity and isotropy. The other processes, like shear and heat conduction, are directional mechanisms and they decay as the Universe expands. In a cosmological context the inclusion of viscosity broadens the applicability of the theory considerably. Bulk viscosity can produce an accelerated expansion even without dark energy matter due to the presence of an effective negative pressure. A number of papers have appeared on viscous cosmology (for review, see Grøn 1990; Brevik and Grøn 2013). Recently, the present authors (Singh and Kumar 2014; Kumar and Singh 2015) have studied the effect of viscous fluid in $f(R, T)$ gravity and discussed the recent phase transition of the Universe.

Using the Eckart formalism for dissipative fluids, we can assume that the effective pressure of HDE is a sum of the thermodynamical pressure (p_h) and the bulk viscous pressure (Π), i.e.,

$$P_{eff} = p_h + \Pi = p_h - 3\zeta H, \quad (24)$$

where ζ is the positive coefficient of the bulk viscosity. Now, the matter Lagrangian is taken as $\mathcal{L}_m = -P_{eff}$ for which Eq. (4) gives $\ominus_{\mu\nu} = -2T_{\mu\nu} - P_{eff}g_{\mu\nu}$. In this model we follow the same concept as discussed in Sect. 3 to analyze the behavior of the Universe. Using $f(T) = \alpha T$, Eq. (3) yields the field equations for viscous HDE in the framework of $f(R, T)$ gravity as

$$3H^2 = \rho_m + \rho_h + \alpha(\rho_m + \rho_h + p_h - 3\zeta H) + \frac{1}{2}\alpha T, \quad (25)$$

$$2\dot{H} + 3H^2 = -p_h + 3\zeta H + \frac{1}{2}\alpha T. \quad (26)$$

In this case, the trace of energy-momentum tensor is $T = \rho_m + \rho_h - 3(p_h - 3\zeta H)$. Using this value of T into (25) and (26), a single evolution equation of H is given by

$$2\dot{H} + (\alpha+1)(\rho_m + \rho_h + p_h) - 3(\alpha+1)\zeta H = 0. \quad (27)$$

From (25), we get

$$\rho_m = \frac{3}{(3\alpha+2)} H [(\alpha c^2 w_h - 3\alpha c^2 - 2c^2 + 2)H - \alpha\zeta]. \quad (28)$$

Now, Using (13) and (28) into (27), we get

$$\dot{H} + \frac{3(\alpha+1)}{3\alpha+2} (2\alpha c^2 w_h + c^2 w_h + 1) H^2 - \frac{3(\alpha+1)(2\alpha+1)\zeta}{(3\alpha+2)} H = 0. \quad (29)$$

Equation (29) is solvable for H if the coefficient of bulk viscosity ζ is known. Many authors have studied the cosmological models by assuming the various forms of the bulk viscous coefficient (for review, see Maartens 1995). Ren and Meng (2006a, 2006b) have taken a general form $\zeta = \zeta_0 + \zeta_1 H + \zeta_2 H^2$ of bulk viscous coefficient. However, in our case it is too difficult to solve Eq. (29) with this general form of ζ . Therefore, we assume the bulk viscous coefficient of the form $\zeta = \zeta_0 + \zeta_1 H$ (Avelino and Nucamendi 2010; Singh and Kumar 2014) by taking $\zeta_2 = 0$. With this form of ζ , Eq. (29) reduces to

$$\dot{H} + \frac{3(\alpha+1)}{3\alpha+2} (2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H^2 - \frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} H = 0. \quad (30)$$

The solution of (30) is given by

$$H = \frac{e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} t}}{c_2 + \frac{(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1)}{(2\alpha+1)\zeta_0} e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} t}}, \quad (31)$$

where c_2 is a constant of integration. The scale factor a in the terms of cosmic time t is

$$a = c_3 \left[c_2 + \frac{(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1)}{(2\alpha+1)\zeta_0} \times e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} t} \right]^{\frac{3\alpha+2}{3(\alpha+1)(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1)}}, \quad (32)$$

where $c_3 > 0$ is another constant of integration. Equation (32) can be rewritten as

$$a = a_0 \left[1 + \frac{(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H_0}{(2\alpha+1)\zeta_0} \times \left(e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} (t-t_0)} - 1 \right) \right]^{\frac{(3\alpha+2)/(\alpha+1)}{3(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1)}}. \quad (33)$$

One can observe that the model avoids the big-bang singularity. In this case, the deceleration parameter is obtained as

$$q = \frac{3(\alpha+1)}{(3\alpha+2)H_0} \left[(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H_0 - (2\alpha+1)\zeta_0 \right] e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} (t_0-t)} - 1. \quad (34)$$

It is observed that the value of q is time-dependent which comes due to the introduction of bulk viscous term in HDE. The phase transition of the Universe can be explained using this value of deceleration parameter. The deceleration parameter must change its sign from positive to negative to explain the recent phase transition (deceleration to acceleration) of the Universe. In fact, q must change the sign at the

time $t = t_0$ because we have assumed t_0 is the time where the viscous HDE begins to dominate. In other words, the Universe must decelerate for $t < t_0$ (matter dominated epoch) and accelerate for $t > t_0$ (HDE dominated epoch). We observe that the Universe shows the transition from decelerated to accelerated phase at cosmic time t_0 if $3(\alpha+1)[(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H_0 - (2\alpha+1)\zeta_0] = (3\alpha+2)H_0$. Therefore, the value of coupling parameter α can be obtained for a given value of w_h , which may be obtained from the observations, or vice-versa to get the phase transition. Thus, we have shown that the bulk viscous HDE with Hubble horizon as an IR cut-off can explain the recent phase transition of the Universe in the framework of $f(R, T)$ gravity.

Next, we discuss the another geometrical parameters, i.e., statefinder parameters. In this case, the statefinder parameter r is obtained as

$$r = \frac{9(\alpha+1)^2}{(3\alpha+2)^2 H_0} \left[(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H_0 - (2\alpha+1)\zeta_0 \right] \times \left[\frac{(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) H_0 - (2\alpha+1)\zeta_0}{H_0 e^{\frac{6(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} (t-t_0)}} + \frac{(2\alpha c^2 w_h + c^2 w_h - 2\alpha \zeta_1 - \zeta_1 + 1) - \frac{3\alpha+2}{\alpha+1}}{e^{\frac{3(\alpha+1)(2\alpha+1)\zeta_0}{(3\alpha+2)} (t-t_0)}} \right] + 1. \quad (35)$$

The second statefinder parameter s is not given here due to complexity but one can find it by using the values of q and r from (34) and (35) in $s = \frac{r-1}{3(q-1/2)}$. Our model reproduces the fixed point value $\{r, s\} = \{1, 0\}$ of Λ CDM model when the parameter α satisfies the condition $\alpha = -\frac{1}{2} \left(1 + \frac{H_0}{c^2 w_h H_0 - \zeta_1 H_0 - \zeta_0} \right)$. For this value of α , the statefinder pair is independent of time and remains fixed throughout the evolution as in Λ CDM model. Indeed, we have obtained time-dependent statefinder pair which means that a general study of the behavior of this pair is needed.

We plot the trajectories in r - s plane for some particular values of parameters α and w_h to discriminate our model with existing models of DE. Here, we have taken $c = 0.5$, $H_0 = 1$, $\zeta_0 = \zeta_1 = 0.05$ and $t_0 = 1$. In Figs. 1 and 2, the fixed points $\{r, s\} = \{1, 1\}$ and $\{r, s\} = \{1, 0\}$ have been shown as SCDM model and Λ CDM model, respectively. It is obvious from both the figures that for any values of α and w_h , which are consistent with the model, the viscous HDE model always approaches to the Λ CDM model, i.e., $\{r, s\} = \{1, 0\}$ in the late time evolution. In the early time of the evolution our model can approach in the vicinity of SCDM model for some values of α as can be seen in Figs. 1 and 2. It is interesting to note that for larger negative values of α the trajectories may start from Λ CDM in the early time and approach to the same Λ CDM model during the late time evolution.

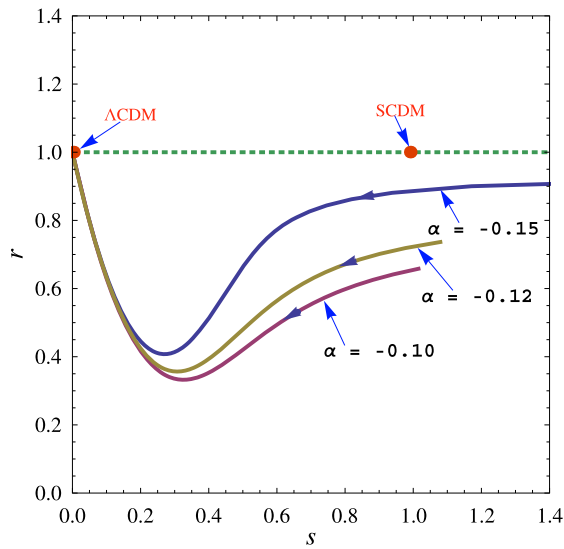


Fig. 1 The trajectories in r - s plane are plotted for $w_h = -0.5$ and different values of α

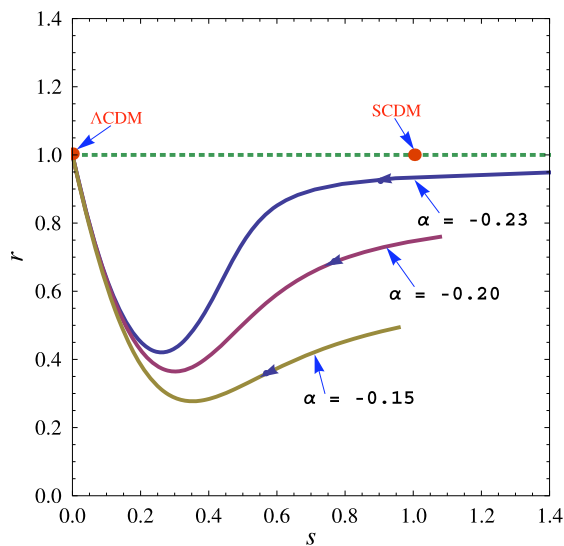


Fig. 2 The trajectories in r - s plane are plotted for $w_h = -1$ and different values of α

In the quiescence model with constant EoS (Q_1 -model) (Sahni et al. 2003; Alam et al. 2003) and the Ricci dark energy (RDE) model (Feng 2008), it has been shown that the trajectories in r - s plane are vertically straight lines. In the both models, s is constant throughout the evolution of the Universe, while r increases in RDE model and decreases in Q_1 -model starting from the initial point $r = 1$. It has also been shown (Sahni et al. 2003; Alam et al. 2003) that the trajectories for the quintessence scalar field model (Q_2 -model), where the scalar potential $V(\phi)$ varies as $V(\phi) \propto \phi^{-\beta}$, $\beta \geq 1$, and chaplygin gas model approach asymptotically to the Λ CDM model in the late time. Comparing this viscous HDE model with Q_1 -model and RDE model, we find

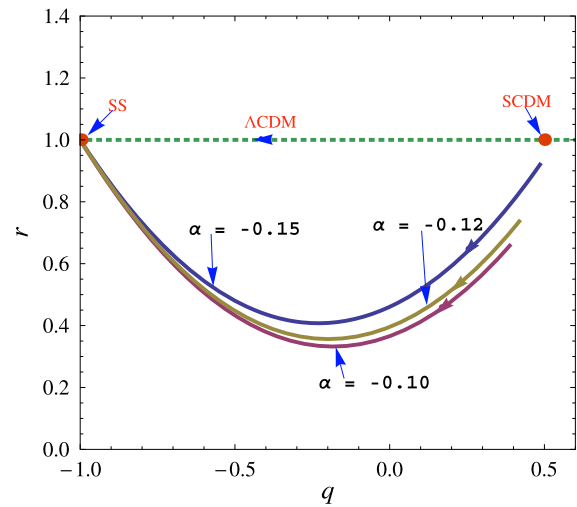


Fig. 3 The trajectories in r - q plane are plotted for $w_h = -0.5$ and different values of α

that our viscous HDE model produces the curved trajectories which approach to Λ CDM model in the late time. Further, we observe that our model almost shows the similar trajectories like Q_2 -model for some values of α and w_h in r - s plane. For $\alpha = -0.15$, $w_h = -0.5$ and $\alpha = -0.23$, $w_h = -1$ as shown in Figs. 1 and 2, respectively, the trajectories show almost similar behavior as Q_2 -model for $\beta = 2$.

It may be observed from Figs. 1 and 2 that for a fixed value of α , say $\alpha = -0.15$, the trajectories show more deviation from the Q_2 -model as we increase the negative value of w_h . Further, for small negative values or positive values of α , the trajectories deviate more and more from the Q_2 -model for fixed value of w_h (see, Sahni et al. 2003; Alam et al. 2003). We may fix α , and vary ζ_0 and ζ_1 , the same behavior of trajectories can be observed for suitable values of ζ_0 and ζ_1 .

Figures 3 and 4 plot the trajectories in r - q plane. Here, we have taken $c = 0.5$, $H_0 = 1$, $\zeta_0 = \zeta_1 = 0.05$, and $t_0 = 1$. The SCDM model and SS model (steady-state cosmology) have been shown by the fixed points $\{r, q\} = \{1, 0.5\}$ and $\{r, q\} = \{1, -1\}$, respectively. We observe the signature flip in the value of q from positive to negative which explain the recent phase transition successfully. The Λ CDM model starting from the fixed point of SCDM model evolves along the dotted line and ends at the fixed point of SS model. It can be observed from both the figures that for any values of α and w_h , the viscous HDE model always approaches to the SS model, i.e., $\{r, q\} = \{1, -1\}$ as Λ CDM, Q_2 and chaplygin gas models approach in the late time evolution. However, in the early time of evolution the model starts close to SCDM model for some values of α as shown in Figs. 3 and 4. Moreover, it may start exactly from the fixed point $\{r, q\} = \{1, 0.5\}$ of SCDM for suitable value of α . In r - q plane also the trajectories corresponding to our model show the Q_2 -model like behavior. Again, comparing our viscous

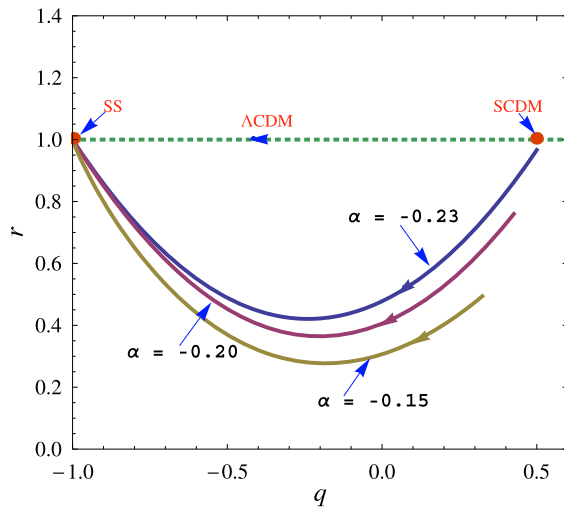


Fig. 4 The trajectories in r - q plane are plotted for $w_h = -1$ and different values of α

HDE model with Q_1 -model, RDE model, Q_2 -model and chaplygin gas model, we find that viscous HDE model is compatible with Q_2 -model.

5 Conclusion

In GR and Brans-Dicke theory, some authors (Hsu 2004; Xu et al. 2009) have found that the Hubble horizon is not a viable candidate to explain the accelerated expansion of the Universe. However, Pavón and Zimdahl (2005), and Banerjee and Pavón (2007) have shown that the interaction between HDE and DM can describe the accelerated expansion. Therefore, it is clear that one can observe accelerated expansion if the interaction between the different matter contents is considered. In this work, we have studied non-viscous and viscous HDE models with Hubble horizon as an IR cut-off in the frame work of modified $f(R, T)$ gravity. The $f(R, T)$ gravity theory presents a maximal coupling between geometry and matter. Therefore, we have explored the consequences of the coupling of matter with the geometry of the Universe instead of taking the interaction between HDE and DM as many authors have studied. We have investigated the possibility whether the Hubble horizon as an IR cut-off could explain an accelerated expansion in $f(R, T)$ gravity. We have shown that the non-viscous and viscous HDE models with Hubble horizon as an IR cut-off may explain the accelerated expansion in the frame work of this modified theory. Further, we have investigated statefinder pair $\{r, s\}$ to discriminate our non-viscous and viscous HDE models with other existing DE models. We summarize the results of these two models as follows: in non-viscous HDE model, we have found an accelerated expansion under the constraint of parameters. In this case, we have obtained constant deceleration and statefinder parameters. Due to constant q , it is not

possible to analyze the phase transition of the Universe. We have found the fixed point $\{r, s\} = \{1, 0\}$ of Λ CDM model as a particular case of this model. Thus, non-viscous HDE model is consistent with Λ CDM model.

In viscous HDE model, we have obtained the recent phase transition of the Universe as the deceleration parameter comes out to be time-dependent and shows signature flip from positive to negative. In this model, the statefinder parameters also are the function of cosmic time t . These time-dependent parameters are possible due to the inclusion of bulk viscous fluid in HDE model which could explain the recent phase transition in a better way. It is interesting to note that the viscous HDE model gives the Λ CDM model fixed point $\{r, s\} = \{1, 0\}$ and remains fixed in Λ CDM model throughout the evolution for a specific value of α as discussed in Sect. 4. The statefinder diagnostic have been discussed through the trajectories of r - s and r - q planes as shown in Figs. 1–4 to discriminate our model with the existing DE models. In Figs. 1 and 2, it has been observed that some of the trajectories pass through the vicinity of SCDM during early time but ultimately all approach to Λ CDM model in the late time. In Figs. 3 and 4, it can be seen that the trajectories may start from SCDM model for a suitable value of α in early time but all the trajectories approach to SS model in the late time evolution. It has been noticed that for some values of α the trajectories of the viscous HDE model are similar to the trajectories of Q_2 -model (Sahni et al. 2003; Alam et al. 2003). Therefore, the viscous HDE in the framework of $f(R, T)$ gravity gives more general results in comparison to Λ CDM and Q_2 -model at least at the level of statefinder diagnostic as we are able to achieve the behavior of both the models.

Acknowledgements The authors are thankful to the referee for his constructive comments and suggestions to improve the manuscript. One of the author PK expresses his sincere thank to University Grants Commission (UGC), India, for providing Senior Research Fellowship (SRF).

References

- Ade, P.A.R., et al.: *Astron. Astrophys.* **571**, A16 (2014)
- Alam, U., et al.: *Mon. Not. R. Astron. Soc.* **344**, 1057 (2003)
- Avelino, A., Nucamendi, U.: *J. Cosmol. Astropart. Phys.* **04**, 06 (2009)
- Avelino, A., Nucamendi, U.: *J. Cosmol. Astropart. Phys.* **08**, 009 (2010)
- Baffou, E.H., et al.: *Astrophys. Space Sci.* **356**, 173 (2015)
- Bamba, K., Nojiri, S., Odintsov, S.D.: *J. Cosmol. Astropart. Phys.* **0180**, 045 (2008)
- Bamba, K., et al.: *Eur. Phys. J. C* **67**, 295 (2010)
- Banerjee, N., Pavón, D.: *Phys. Lett. B* **647**, 477 (2007)
- Bennett, C.L., et al.: *Astrophys. J.* **148**, 1 (2003)
- Bento, M.C., Bertolami, O., Sen, A.A.: *Phys. Rev. D* **66**, 043507 (2002)
- Brevik, I., Gorbunova, O.: *Gen. Relativ. Gravit.* **37**, 2039 (2005)
- Brevik, I., Grøn, Ø.: *Relativistic Universe Models: Recent Advances in Cosmology*, p. 97. Nova Scientific Publication, New York (2013)
- Caldwell, R.R.: *Phys. Lett. B* **545**, 23 (2002)
- Carroll, S.M.: *Living Rev. Relativ.* **4**, 1 (2001)

- Cataldo, M., Cruz, N., Lepe, S.: Phys. Lett. B **619**, 5 (2005)
- Chakraborty, S.: Gen. Relativ. Gravit. **45**, 2039 (2013)
- Chiba, T., Okaba, T., Yamaguchi, M.: Phys. Rev. D **62**, 023511 (2000)
- Cohen, A.G., Kaplan, D.B., Nelson, A.E.: Phys. Rev. Lett. **82**, 4971 (1999)
- Copeland, E.J., Sami, M., Tsujikawa, S.: Int. J. Mod. Phys. D **15**, 1753 (2006)
- De Felice, A., Tsujikawa, S.: Phys. Rev. D **80**, 063516 (2009)
- Eckart, C.: Phys. Rev. **58**, 919 (1940)
- Elizalde, E., et al.: Class. Quantum Gravity **27**, 095007 (2010)
- Fayaj, V., et al.: Astrophys. Space Sci. **353**, 301 (2014)
- Feng, C.J.: Phys. Lett. B **670**, 231 (2008)
- Feng, C.-J., Li, X.-Z.: Phys. Lett. B **680**, 355 (2009)
- Grøn, Ø.: Astrophys. Space Sci. **173**, 191 (1990)
- Harko, T.: Phys. Rev. D **90**, 044067 (2014)
- Harko, T., et al.: Phys. Rev. D **8**, 024020 (2011)
- Heller, H.M., et al.: Astrophys. Space Sci. **20**, 205 (1973)
- Houndjo, M.J.S., Piattella, O.F.: Int. J. Mod. Phys. D **21**, 1250024 (2012)
- Hsu, S.D.H.: Phys. Lett. B **594**, 13 (2004)
- Hu, M.-G., Meng, X.-H.: Phys. Lett. B **635**, 186 (2006)
- Komatsu, E., et al.: Astrophys. J. Suppl. **192**, 18 (2011)
- Kumar, P., Singh, C.P.: Astrophys. Space Sci. **357**, 120 (2015)
- Landau, L.D., Lifshitz, E.M.: Fluid Mechanics. Butterworth-Heinemann, Oxford (1987)
- Li, M.: Phys. Lett. B **603**, 1 (2004)
- Li, M., et al.: J. Cosmol. Astropart. Phys. **09**, 021 (2013)
- Maartens, R.: Class. Quantum Gravity **12**, 1455 (1995)
- Maartens, R.: [arXiv:astro-ph/9609119](https://arxiv.org/abs/astro-ph/9609119) (1996)
- Misner, C.W.: Astrophys. J. **151**, 431 (1968)
- Murphy, G.L.: Phys. Rev. D **8**, 4231 (1973)
- Nojiri, S., Odintsov, S.D.: Phys. Lett. B **631**, 1 (2005)
- Nojiri, S., Odintsov, S.D.: Phys. Rep. **505**, 59 (2011)
- Padmanabhan, T.: Phys. Rev. D **66**, 021301 (2002)
- Padmanabhan, T.: Phys. Rep. **380**, 235 (2003)
- Padmanabhan, T., Chitre, S.M.: Phys. Lett. A **120**, 443 (1987)
- Pavón, D., Zimdahl, W.: Phys. Lett. B **628**, 206 (2005)
- Peebles, P.J.E., Ratra, B.: Rev. Mod. Phys. **75**, 559 (2003)
- Percival, W.J., et al.: Mon. Not. R. Astron. Soc. **401**, 2148 (2010)
- Ren, J., Meng, X.H.: Phys. Lett. B **633**, 1 (2006a)
- Ren, J., Meng, X.H.: Phys. Lett. B **636**, 2 (2006b)
- Riess, A.G., et al.: Astrophys. J. **659**, 98 (2007)
- Sahni, V., et al.: JETP Lett. **77**, 201 (2003)
- Sasidharan, A., Mathew, T.K.: Eur. Phys. J. C **75**, 348 (2015)
- Sebastiani, L.: Eur. Phys. J. C **69**, 547 (2010)
- Setare, M.R.: Phys. Lett. B **644**, 99 (2007)
- Setare, M.R., Sheykhi, A.: Int. J. Mod. Phys. D **19**, 1205 (2010)
- Shabani, H., Farhoudi, M.: Phys. Rev. D **90**, 044031 (2014)
- Sharif, M., Zubair, J.: J. Cosmol. Astropart. Phys. **03**, 28 (2012)
- Sheykhi, A., Jamil, M.: Phys. Lett. B **694**, 284 (2011)
- Singh, C.P., Kumar, P.: Eur. Phys. J. C **74**, 3070 (2014)
- Singh, C.P., Singh, V.: Gen. Relativ. Gravit. **46**, 1696 (2014)
- Singh, C.P., Kumar, S., Pradhan, A.: Class. Quantum Gravity **24**, 455 (2007)
- Suzuki, N., et al.: Astrophys. J. **746**, 85 (2012)
- 't Hooft, G.: [arXiv:gr-qc/9310026](https://arxiv.org/abs/gr-qc/9310026) (1993)
- Tawfik, A., Harko, T.: Phys. Rev. D **85**, 084032 (2012)
- Weinberg, S.: Gravitation and Cosmology. Wiley, New York (1972)
- Wilson, J.R., Mathews, J.G., Fuller, G.M.: Phys. Rev. D **75**, 043521 (2007)
- Xinzhang, C., Spiegel, E.A.: Mon. Not. R. Astron. Soc. **323**, 865 (2001)
- Xu, L.: J. Cosmol. Astropart. Phys. **09**, 016 (2009)
- Xu, L., Li, W., Lu, J.: Eur. Phys. J. C **60**, 135 (2009)
- Zimdahl, W.: Phys. Rev. D **53**, 5433 (1996)

Study of Vendors Selection Methods for Efficient Supply Chain Management

Rajiv Chaudhary¹, R C Singh², Reema Chaudhary³,

^{1,2}(Department of Mechanical Engineering, Delhi Technological University, Delhi, India)

³(Bhagwan Parashuram Institute of Technology (BPIT), Delhi, India)

Email: rch_dce@rediffmail.com

Abstract : *The ambition of the customers has raised suddenly, by getting newer variety products regularly, which has changed the market dynamics drastically. Experts, dealing with the related issues are looking for new solutions to new challenges. Incorporating flexibility in Supply Chain Management has become indispensable for any production system today, which has to highly influenced by the suppliers. Obviously, the supplier selection method has to be dynamic, to respond the variation in organization's requirement, supplier's capacity, quality, lead time, unit cost and transportation cost varying with time. Also, the supplier may not be repeated next time, even for the same set of items. Several theories have been proposed for this critical process, which include various quantitative/qualitative methods and models. Several other issues considering, supplier relationship choice, supplier evaluation and selection taking firm's requirements into account, long-term vendor selection, and negotiating and evaluating suppliers etc also has been undertaken. This paper reviews various aspects of these supplier selection methods.*

Keywords : Vendor Selection, Quantitative, Qualitative, Mathematical Programming

I. INTRODUCTION

Optimizing the business operations, can not be a matter of one's choice in today's competitive environment. Moreover, meeting out the increasing and fluctuating customer's demand within due time has also become another challenge in today's continuously changing environment. Out of various factors desired for the steady and smooth business, the role of supply chain is very important, due to which a good chain of supply or supplier needs to be ensured. Hence, the selection of good suppliers becomes a key parameter to ensure success of any business organization which also helps in making the organization competitive. Any organization concerned in this direction would be following the procedure of vendor selection which is actually identifying, evaluating, and contracting with suppliers.

II. VENDOR SELECTION METHODS

A good number of literature is available on supplier selection problems [1,2,3,4&5]. Most of the methodologies proposed for the supplier selection problem, discusses about the cases of TSSP [Traditional Supplier Selection Problem], in which the suppliers are given certain relative ranking and the supplier with top ranking is supposed to be same throughout the phase unless ranked or assessed again.

Also the problems of supplier selection have been dealt by broadly placing the problems into two categories of Quantitative and Qualitative models.

III. QUANTITATIVE METHODS

Quantitative methods use several Mathematical programming techniques for the vendor selection as well as order quantity

decisions simultaneously. These mathematical programming models are intended to minimize or maximize some objective function, subject to both vendor and buyer constraints and ultimately selecting several vendors in order.

Various quantitative models used are linear programming (LP), mixed-integer linear program (MILP), mixed-integer non-linear (MINLP), fuzzy goal programming, data envelopment analysis (DEA), dynamic programming and multi-objective programming etc [7]. In some of the work, researches have also applied meta-heuristic such as genetic algorithm (GA) [8-10]. In one another variation the vendor selection problem has been also formulated applying quality, delivery, and capacity constraints involving price-break regimes [11].

IV. QUALITATIVE MODELS

There are many qualitative models such as AHP [13 14], Fuzzy-AHP and weighted point method, matrix approach, vendor performance matrix approach, vendor profile analysis, Analytical Network Process, TOPSIS and Fuzzy-TOPSIS have been proposed by various researchers in the past to solve TSSP.

One another model was developed for TSSP applying AHP and quality management system principles. For enhancing the performance of the vendor selection in comparison to the traditional approaches one case based reasoning approach has also been applied for the efficient supplier selection. Analytical Hierarchy Process (AHP), is one of the most used qualitative technique for supplier selection. It was developed by Saaty in 1970 for handling with qualitative factors, is used in multi-criteria decision making problem. Based on the

analysis some ranking of the supplier is prepared. By analyzing these ranking the position can be interpreted, in terms of being the most-preferred, least preferred or absolutely preferred suppliers while implementing this flexible supplier selection model. Under the condition of competitive pressures, organizations resorted to restructuring of the operations by focusing more on their core activities which are strategically important and other peripheral operations were outsourced.

According to an estimate made by Accenture, a majority of these companies are spending about half of their total budget on outsourcing. Also, the practice of outsourcing is further expanding and exploring new prospects. As a common perception, conventionally outsourcing used to be involving non-core activities only with the basic purpose of reducing costs and concentrating on their core competency. But as per the emerging trend and practices outsourcing is accessing to almost every domain of business, be it engineering, research & development, product development or even marketing. Making vendor selection considering both tangible and intangible factors has been also devised by integration of qualitative methods and linear programming.

V. CONCLUSION

Adopting some scientific practice for proper vendor selection has become an essential condition for ensuring successful supply chain in the organization, which helps the firms in achieving their goal comfortably. Also the supplier selection helps in dealing with market uncertainties, along with reducing the risks associated with the performance. Through various Qualitative as well as Quantitative methods we follow a structured approach in modeling the decision making problem of vendor selection, which simplifies the complexity of the situation and suggests some solution in that situation. The models proposed have been found applicable successfully in various situations, which makes us now capable of dealing with the complex problems of vendor selection based on certain criteria and now decision makers may be facilitated in selecting the best suited supplier among the numerous suppliers.

REFERENCES

- [1] Weber, C. A., Current, J. R., & Benton, W. C. (1991). Vendor selection criteria and methods. *European Journal of Operational Research*, 50, 2–18.
- [2] Boer, L. D., Wegen, L. V. D., & Telgen, J. (1998). Outranking methods in support of supplier selection. *European Journal of Purchasing & Supply Management*, 4, 109–118.
- [3] Wadhwa, V., & Ravindran, A. R. (2007). Vendor selection in outsourcing. *Computers & Operations Research*, 34, 3725–3737.
- [4] Chan, F. T. S., & Chan, H. K. (2004). Development of the supplier selection model a case study in the advanced technology industry. *Proceedings of the Institution of Mechanical Engineers*, 218, 1807–1824.
- [5] Ware, N. R., Singh, S. P., & Banwet, D. K. (2012). Supplier selection problem: A state-of-the-art review. *Management Science Letters*, 2(5), 1465–1490.
- [6] Amin, S. H., Razmi, J., & Zhang, G. (2011). Supplier selection and order allocation based on fuzzy SWOT analysis and fuzzy linear programming. *Expert Systems with Applications*, 38, 334–342.
- [7] Narasimhan, R. (1983). An analytical approach to supplier selection. *Journal of Purchasing and Materials Management*, 19(4), 27–32.
- [8] Ding, H., Benyoucef, L., & Xie, X. (2005). A simulation optimization methodology for supplier selection problem. *International Journal of Computer Integrated Manufacturing*, 18(2–3), 210–224.
- [9] Che, Z. H. (2010). A genetic algorithm-based model for solving multi-period supplier selection problem with assembly sequence. *International Journal of Production Research*, 48(15), 4355–4377.
- [10] Yeh, W. C., & Chuang, M. C. (2011). Using multi-objective genetic algorithm for partner selection in green supply chain problems. *Expert Systems with Applications*, 38, 4244–4253.
- [11] Chaudhry, S.S., Forst, F.G. and Zydiak, J.L.. Vendor Selection with Price Breaks. *European Journal of Operational Research*, 1993, 70: 52-66.
- [12] Ghodsypour, S. H., & O’Brine, C. (1998). A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *International Journal Production Economics*, 56–57, 199–212.
- [13] Saaty, T. L. (1980). *The analytical hierarchy process: Planning, priority setting, resource allocation*. New York, NY: McGraw-Hill.
- [14] Saaty, T. L. (1990). How to make decision: The analytical decision process. *European Journal of Operation Research*, 48, 9–26.



ORIGINAL ARTICLE

Use of Green Technology (F.S.P.) for processing of 99.9% Copper with Carbon Nano Tubes

R.S. Mishra

Mechanical Engineering Dept. DTU Delhi-110042 INDIA

Email: rsmishra1651956@yahoo.co.in

ABSTRACT

The states of development of green technology (FSP) for processing of Copper with carbon nanotubes is presented in this paper because Friction stir processing (FSP), is a green technology process in which solid-state microstructural modification technique is viewed in terms of severe plastic deformation technique due to high processing strain involved. It is also unique green technology from the viewpoint of its applicability to a localized region. To investigate the various parameters affecting the friction stir processed copper with carbon nano tubes and enhancement of the microstructure, hardness and tensile properties of the composite material the behaviour of copper with carbon nano tubes has been studied in detail with single pass and multi passes (i.e. double passes and triple passes) and SEM results are presented for each case in the present investigation.

Key words: Microstructure; Hardness, Tensile properties; Friction stir processing of Copper, Green Technology

Received: 30th Nov. 2015, Revised: 27th Dec. 2015, Accepted: 29th Dec. 2015

©2016 Council of Research & Sustainable Development, India

How to cite this article:

Mishra R.S. (2016): Use of Green Technology (F.S.P.) for processing of 99.9% Copper with Carbon Nano Tubes. AJMECS, Vol. 1[1]: January, 2016: 49-59.

INTRODUCTION

The term "technology" refers to the application of knowledge for the betterment of society. Green Technology encompasses a continuously evolving group of methods and materials for generating energy to non-toxic products. Creativity and innovation in *green technology as per society requirement* is also known as *clean technology*. *Clean technology* can communicate with those in the manufacturing industry who are developing and distributing *green technologies*. The major expectations will bring innovation and changes in daily life of similar magnitude to the information technology explosion over the last two decades. In early stages, it is impossible to predict what green technology may eventually encompass. The goals of green technology is to inform the developments in this rapidly growing fields including meeting the needs of society in ways that can continue indefinitely into the future without damaging or depleting natural resources in terms of sustainability. Also ending the "cradle to grave" cycle of manufactured products, by creating products that can be fully reclaimed or re-used, reducing waste and pollution by changing patterns of production and consumption, developing alternatives to technologies:

1. whether fossil fuel or chemical intensive agriculture
2. demonstrated to damage health and the environment and creating a center of economic activity around technologies
3. Products that benefit the environment, speeding their implementation and creating new careers that truly protect the planet.

This study focuses on green technology to contribute towards sustainability i.e. meeting present needs without compromising the ability of future generations to meet their needs.

Friction stir processing (FSP) is a green technology process in which relatively new solid-state micro-structural modification technique can also viewed as a severe plastic deformation technique because of the high processing strain involved. It is also unique from the viewpoint of its applicability to a localized region. FSP was developed for micro-structural modification of metallic materials. In FSP, a rotating tool is plunged into a material and high plastic deformation is produced. FSP is used to enhance ductility, induces super plasticity and improve corrosion resistance properties. Dynamic recrystallization of the deformed zone forms an ultrafine-grained structure. FSP has been successfully applied to various cast aluminium and magnesium and copper alloys to eliminate casting defects and there by improve their mechanical properties [24, 25]. Two modes of metal transfer during friction stir processing have been discussed [15]. The first mode of metal transfer is generated between the tool shoulder and the plate and takes place as layer-by-layer deposition of metal one over the other. The second mode of metal transfer is generated by the extrusion of metal around the tool pin, when it reaches a state of sufficient plasticity. Metal transfer, generated between the tool shoulder and the plate, plays an important role in influencing the mechanical properties during friction stir process [29]. Modes of metal transfer are clearly visible in the microstructure characteristics, but they are not too distinct in macrostructure of most processed samples. Friction stir processing can be applied as a single-pass for processing a small area. For large engineering components in which the contact areas are relatively large, single pass FSP may not be adequate. Multi-pass FSP with a certain level of overlap between the successive passes is required for large contact areas. For both single and multi-pass processes, it is important to assess the microstructural evolution and its influence on the mechanical properties [2, 7, 17, 23].

Effect of single and multiple-pass friction Stir processing on microstructure, hardness and tensile properties of a 99.99% Cu with carbon nano tubes:

Copper is a mostly used industrial and functional metal for various thermal, electrical and electronic applications, i.e. electronic packaging, electrical contacts and resistance welding electrodes. This is because of good thermal and electrical conductivity, high plasticity and excellent resistance to corrosion and oxidation. Nevertheless, low mechanical strength and poor wear resistance limit its applications [3, 10, 16, 26, and 27]. Carbon nano tubes are unique tubular structures of nanometer diameter and large length/diameter ratio. The nanotubes may consist of one up to tens and hundreds of concentric shells of carbons with adjacent shells separation of 0.34 nm. The carbon network of the shells is closely related to the honeycomb arrangement of the carbon atoms in the graphite sheets. The amazing mechanical and electronic properties of the nanotubes stem in their quasi-one dimensional (1D) structure and the graphite-like arrangement of the carbon atoms in the shells. Thus, the nanotubes have high Young's modulus and tensile strength, which makes them preferable for composite materials with improved mechanical properties. The nanotubes can be metallic or semiconducting depending on their structural parameters. This opens the ways for application of the nanotubes as central elements in electronic devices including field-effect transistors (FET), single-electron transistors and rectifying diodes. Possibilities for using of the nanotubes as high-capacity hydrogen storage media were also considered for experiment purpose [26].

Here the states of development of FSP for processing of Copper with carbon nano tubes are addressed. This paper investigates the parameters affecting the friction stir processed copper with carbon nanotubes and enhancement of the microstructure, hardness and tensile properties of the composite material. The behaviour of Copper with carbon nano tubes has been studied with single pass, double passes and triple passes The SEM results are also presented for each case.

LITERATURE REVIEW

Bahram A. Khiyavi, *et al* [1] produced copper reinforced metal matrix composite (MMC) using micron sized chromium particles via friction stir processing (FSP) in order to study effects of adding Cr particles to copper based matrix by FSP. Microstructures, micro hardness and wear properties were studied in order to evaluate the microstructures and mechanical properties of fabricated composites. H.R. Akramifard *et al* [4] in their investigation, pure Cu sheets were reinforced with 25 μ m SiC particles to fabricate a composite surface layer by friction stir processing (FSP). In order to improve distribution of reinforcing SiC particles, a net of holes were designed by drill on the surface of pure Cu sheet. H. Sarmadi, *et al* [5] focused on friction stir processing (FSP) used to produce copper-graphite surface composites. Five tools with different pin profile were employed in order to achieve a comprehensive dispersion. Results showed that the tool with rectangular pin give rise to a better dispersion of graphite particles. Furthermore, four copper-graphite composites containing different graphite content were prepared using rectangular tool through repeating the process passes. Friction and wear performance of the composites were also studied using a pin-on-disc tribometer. It was indicated that the friction coefficients of composites were lower than pure annealed copper and decreased with increase in graphite content.

Galvao, *et al.* [6] used 1 and 3 mm-thick copper-DHP plates were processed with the aim of simulating surface (SFSP) and bulk (VFSP) processing. The influence of the processing conditions on the microstructure and mechanical properties of the processed materials was analyzed. It was found that the tool geometry, which has a close relation with the plastic deformation and dynamic recrystallization kinetics inside the stirred volume, the processing parameters and the heat exchange conditions, which determine the extent of dynamic recrystallization and annealing phenomenon, are determinant in FSP. Kudzanayi Chiteka [8] studied making a choice in selection of friction stir welding/processing (FSW/P) tool material which has become an important task in determining the quality of the weld produced. The tool material selection depends on the operational characteristics such as temperature, wear resistance and fracture toughness that determine the type of materials to be joined. Soft materials can be easily welded using tool steels while harder materials need harder tool materials such as carbide based materials and polycrystalline cubic boron nitride (PCBN). K. Surekha, A. El-Sayed [9] the objective of their study was to obtain a high strength, high conductivity copper by friction stir processing. Three millimeter thick pure copper plate was friction stir processed to a depth of 2.8 mm at low-heat input conditions by varying the travel speed from 50 to 250 mm/min at a constant rotation speed (300 rpm) to obtain fine grains. Mohsen Barmouz, *et al* [11] studied multi-pass friction stir processing (MFSP) which was used for improvement of microstructural and mechanical properties of in situ Cu/SiC composites. Field emission scanning electron microscopy and optical microscopy images indicate that multi-pass FSP notably enhances the separation and dispersion of SiC particles and also reduces the grain size in the composite matrix, SiC particles size and porosity contents. Mohsen Barmouz, *et al* [12] their study was to produce copper reinforced metal matrix composite (MMC) layers using micron sized SiC particles via friction stir processing (FSP) in order to enhance surface mechanical properties. Microstructural evaluation using optical microscopy (OM) and scanning electron microscopy (SEM) indicated that an increase in traverse speed and a decrease in rotational speed cause a reduction in the grain size of stir zone (SZ) for the specimens friction stir processed without SiC particles. With the aim of determining the optimum processing parameters, the effect of traverse speed as the main processing variable on microstructure and microhardness of MMC layers was investigated. P. Xue, *et al* [13] worked with large-area bulk ultrafine grained (UFG) pure Cu which was successfully prepared by multiple-pass overlapping friction stir processing (FSP) under additional rapid cooling. It was observed that overlapping FSP does not exert a significant effect on the microstructure and mechanical properties of the FSP UFG Cu. Similar average grain size was achieved in the transitional zone (TZ) of the multiple-pass

FSP sample compared to that in the nugget zone of the single-pass FSP sample, and the TZ exhibited a strong {111} (112) type A fiber shear texture. Q. Zhang, et al [14] investigated situ Al₃Ti/Al-5.5Cu composites fabricated by powder metallurgy and subsequent forging which were subjected to multiple pass friction stir processing (FSP) with and without active cooling.

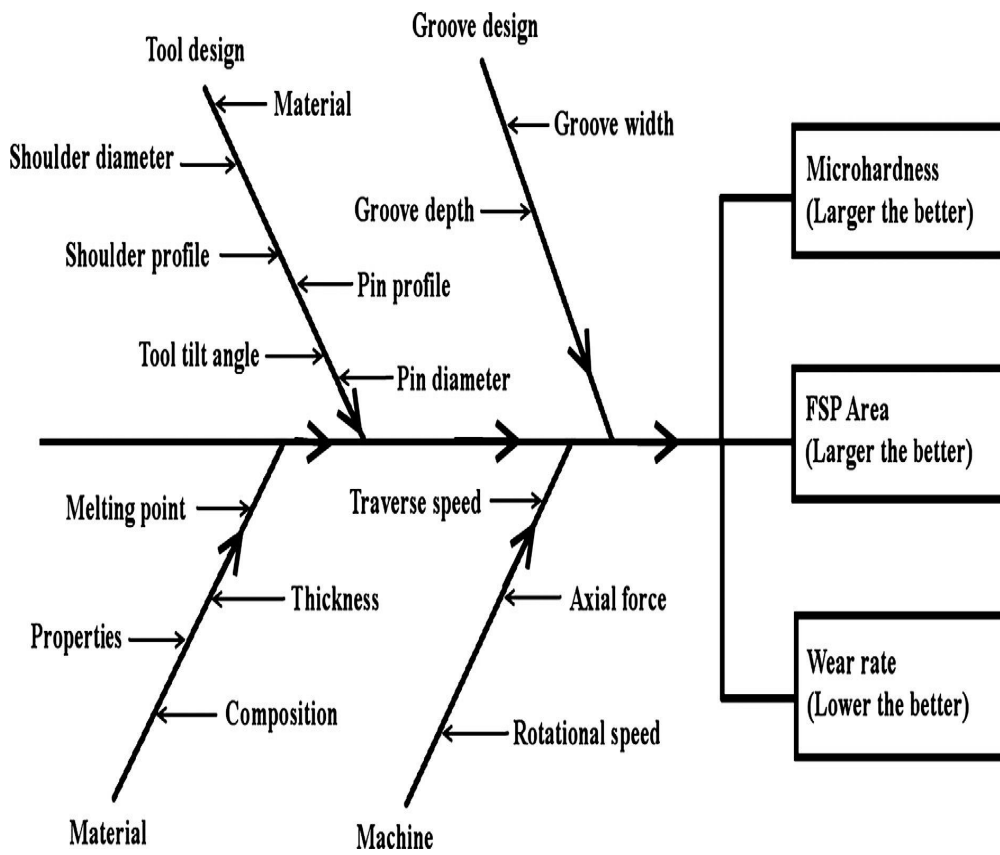
Valentin N. Popov [26] studied Carbon nanotubes: properties and application. V Jegannathan et al. [25] investigated the parameters affecting the friction stir processed copper and enhancement of the mechanical properties of the composite material. The results showed that the grain size of fabricated composite reduce, also it is indicated that in comparison to base metal (copper) micro hardness of friction stir processed composites in stir zone (SZ) increase significantly. The results obtained also indicated that the selected FSP parameters significantly influence the area of surface composite by the distribution of material particles. Higher tool rotational speed and lower processing speed produce an excellent distribution of material particles and higher area of surface composite due to higher frictional heat, increased stirring and material transportation. Ranganath M S et al [21] conducted the experimental studies for assessing the tribological performance of aluminium at the sliding contacts with mild steel plate, using a pin-on-disk tribometer as per ASTM-G 99.

The study has been done in order to explore the friction and wear behaviors at the interface of tribo-pair. The tribological properties as coefficient of friction and specific wear rate of aluminium 6061 are investigated. The Tribological tests are carried out at 500, 1000, 1500 rpm for 1000 meters in dry condition based on Response Surface Methodology. Track diameter, rotating speed and normal load are considered as the design parameters. Using central composite design, the problem is converted into single response optimization problem and the optimum combination of design parameters are found as 50mm track diameter, 500 rpm of rotating speed and 0.5 kg of normal load. The ANOVA result shows that the rotational speed is the most significant factor, followed by load and Track diameter for co-efficient of friction.

Whereas the Track diameter is the most significant factor, followed by rotating speed and normal load for specific wear rate. Finally, microscopic images are investigated to identify the wear mechanism. Sathiskumar, R. et al [18] applied Friction stir processing (FSP) to fabricate boron carbide (B₄C) particulate reinforced copper surface composites. The effect of FSP parameters such as tool rotational speed, processing speed and groove width on microstructure and microhardness was investigated. A groove was contrived on the 6mm thick copper plates and packed with B₄C particles. FSP was carried out using five various tool rotational speeds, processing speeds and groove widths. Sathiskumar R., et al [19] applied the friction stir processing technique to fabricate boron carbide particulate reinforced copper surface composites and investigated the effect of B₄C particles and its volume fraction on microstructure and sliding wear behavior of the same.

A groove was prepared on 6 mm thick copper plates and packed with B₄C particles. The dimensions of the groove was varied to result in five different volume fractions of B₄C particles (0, 6, 12, 18 and 24 vol.%). Sathiskumar R., et al [20] In this work FSP technique was applied to prepare copper surface composites reinforced with variety of ceramic particles such as SiC, TiC, B₄C, WC and Al₂O₃. Empirical relationships were developed to predict the effect of FSP parameters on the properties of copper surface composites such as the area of the surface composite, microhardness and wear rate. The FSP parameters which influence the properties of surface composite are shown in Fig. 1. Salar Salahi and Vahid Rezazadeh [22] studied Fracture mechanism in metals using Friction Stir Processing (FSP) which was a challenging investigation and made by means of a rotating tool inserted in a work piece providing heat transfer and plastic deformation. In their paper, improving ductility during FSP was determined as a purpose and the microstructure and fracture mechanism of samples were investigated during Friction Stir Processing (FSP) of pure copper.

Fig. 1: FSP parameters influencing the properties of surface composite



EXPERIMENTAL SETUP

The machine used for Friction stir processing was a special FSW machine is shown below.

Fig. 2: Friction stir welding / processing machine (Central workshop, Delhi Technological University)



Fig. 3: Tools used during Friction stir processing

The specimen size of the copper plate that is used for processing 200 mm x 74 mm x 5 mm. One groove of 1mm width and 2 mm deep was made on the 99.99% pure Copper Plates using horizontal milling machine with a 1mm (width) saw cutter was cut in the middle of the specimen plate for processing. The tool Material used is H13 steel with shoulder diameter 15mm, threaded pin diameter 8 mm, pin length 2.5 mm with tool rotational speed 960 rpm, tool angle 2° and table traverse speed 25 m / min.

Initially without carbon nanotubes one specimen plate was processed with single pass. Then second specimen plate was processed after filling carbon nanotubes in the groove cut (single pass). Then third specimen plate was processed after filling carbon nanotubes in the groove cut (double pass). Finally the fourth specimen plate was processed after filling carbon nanotubes in the groove cut (triple pass). The processed pieces were taken for the following tests: Tensile Strength Test, Brinell hardness test, Microstructure test.

TENSILE STRENGTH TEST

After the FSP the specimen for tensile testing were cut from the job. The specimens are as shown in figure. The specification of specimen is as following:

1. Length of specimen- 101.6 mm
2. Gauge Length-25.4 mm
3. Gauge width- 6.00mm
4. Gauge thickness- 5.00mm

The specimens were chosen marked with marker on their ends. Care was taken to ensure that the specimens did not have any notching or cracks from manufacturing or any surface defects that would adversely affect the tensile tests. Before loading the specimens in the Instron machine, the computer system connected to the machine was given inputs such as gauge length and width of the specimen. The computer system was then prepared to record data and output necessary load-deflection graphs. The specimens were loaded into the Instron machine, and a tensile test was performed. The data was recorded electronically in text files and the load-deflection curve was shown on the computer screen as a visual representation. The average of different values of 3 specimens each from 2 jobs and parallel material was taken as the final values. The stress strain graphs and load elongations graph are shown for the ultimate strength of the specimen. The

ultimate tensile strength of the processed material comes out to be lesser than the parent material.

Table1a: Specimens details and output results during tensile tests

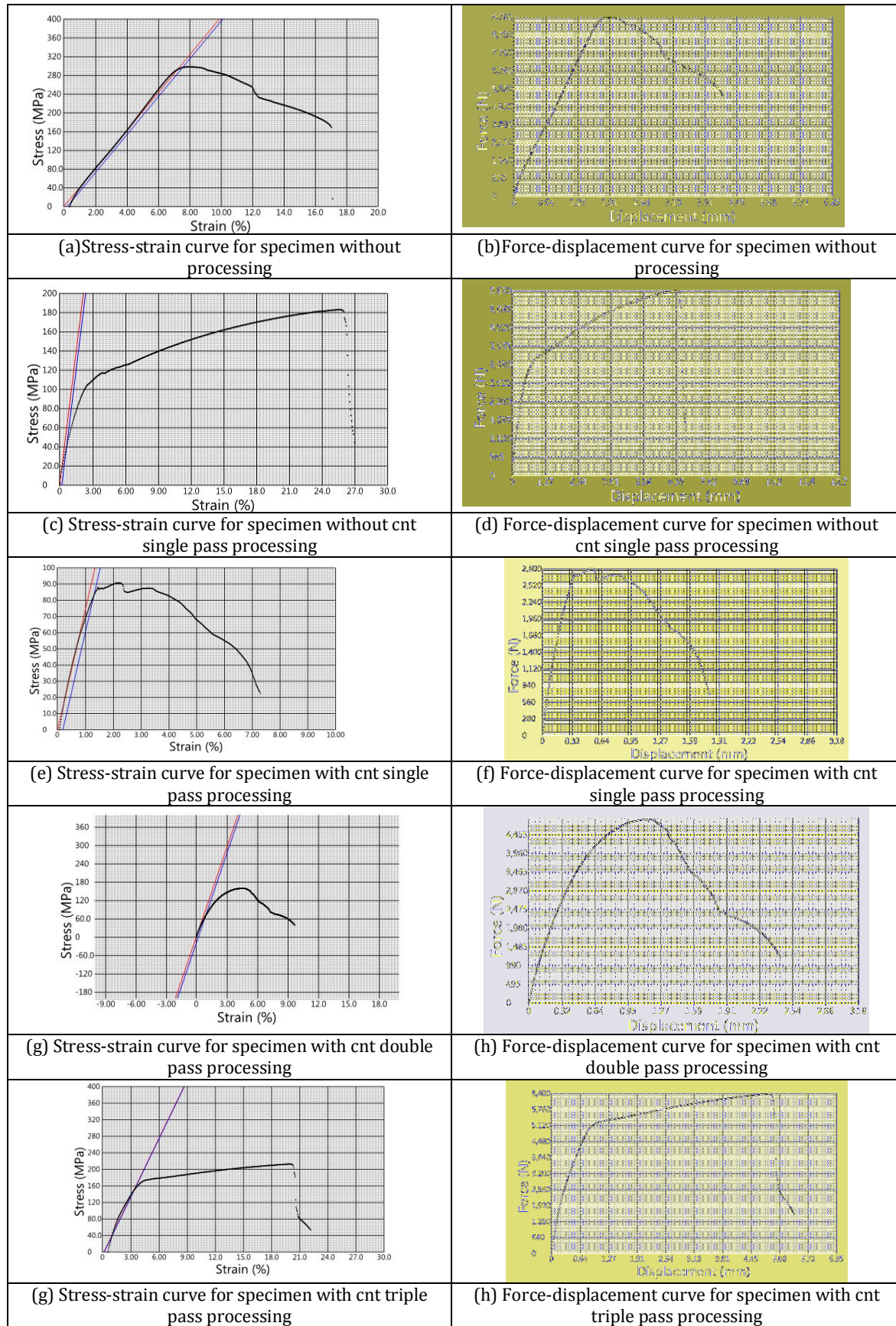
Out Put (Generic metals tensile from position)	Without processing	Single Pass Without CNT
Width (mm)	6.00	6.00
Thickness (mm)	5.40	5.30
Gauge Length (Initial) (mm)	25.0	25.0
Gauge Length (Final) (mm)	28.7	31.8
Area (mm ²)	32.4	30.8
Ultimate Force (N)	9670	5650
Ultimate Stress (MPa)	298	183
Offset @ 0.2% (N)	9620	1240
Offset @ 0.2% (MPa)	297	40.2
TE (Auto) (%)	17.1	26.3

Table1b: Specimens details and output results during tensile tests

Out Put (Generic metals tensile from position)	Single Pass With CNT	Two Pass With CNT	Three Pass With CNT
Width (mm)	6.00	6.00	6.00
Thickness (mm)	5.14	5.05	5.00
Gauge Length (Initial) (mm)	25.0	25.0	25.0
Gauge Length (Final) (mm)	26.8	27.4	30.4
Area (mm ²)	30.8	30.3	30.0
Ultimate Force (N)	2800	4860	6380
Ultimate Stress (MPa)	90.7	160	213
Offset @ 0.2% (N)	2550	2120	4540
Offset @ 0.2% (MPa)	82.7	70.0	151
TE (Auto) (%)	7.28	9.72	22.2

Table2: Specimens details and output results during tensile tests

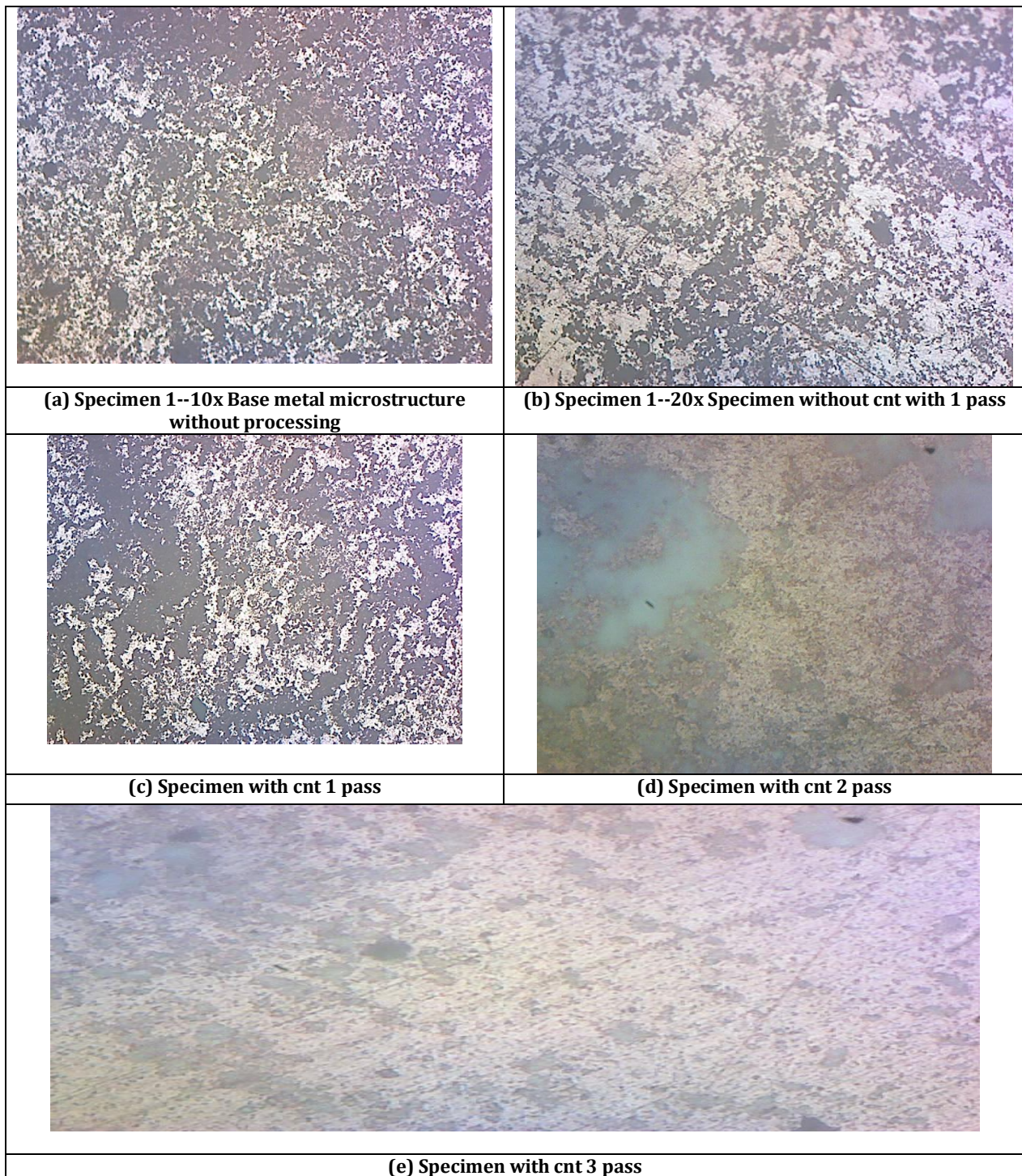
Expts	Description	Depth of indentation	Brinell hardness BHN
1	Without carbon nanotube and without processing	2.8 mm	11.14
2	single pass processing Without carbon nanotube	3.09	8.93
3	Single pass With carbon nanotube	3.47	6.823
4	Double passes With carbon nanotube	3.40	7.162
5	Triple passes With carbon nanotube	3.00	9.55

Fig. 4: Output results during tensile tests after Friction stir processing

BRINELL HARDNESS TEST

The Brinell scale characterizes the indentation hardness of materials through the scale of penetration of an indenter, loaded on a material test-piece. It is one of several definitions of hardness in materials science. The brinell hardness was conducted on all specimens. As we can see the brinell hardness of the processed specimen increases as the number of passes are increased due to the more compact microstructure. The brinell hardness number comes out to be highest for third pass in comparing with single and double passes.

Fig. 5 (a-e): Microscopic results of all specimens after Friction stir processing



MICROSTRUCTURE ANALYSIS

The optical microscope is the principal tool used to characterize the internal grain structure of steel. Traditionally, the structure revealed by the microscope is called the microstructure. The mechanical properties of given steel are strongly influenced by its microstructure. An optical microscope uses reflected light to generate an image. A beam of light is directed down onto the surface and the image is generated either on film or the eye by light reflected along the same direction.

This is performed on specimens either cut to size or mounted in a resin mold. The samples are polished to a fine finish, normally one micron diamond paste, and usually etched in an appropriate chemical solution prior to examination on a metallurgical microscope.

For microstructure test we had 5 specimens. There are two examination methods in metallography such as Macroscopy and Microscopy. In macroscopy the examination of the structural characteristics or chemical characteristics of a metal or an alloy is done by the unaided eye or with the aid of a low-power microscope or binocular, usually under 10x. In microscopy similar examination is done with the prepared metal specimens, employing magnifications with the optical microscope of from 100x to as high as 2000x. Figure 5 represents the microscopic results of all 5 specimens after processing.

CONCLUSION

The effect of friction stir processing (FSP) parameters on copper material with carbon nano tubes as composite material has been studied. Tensile strength, microstructure and micro hardness tests were conducted on the specimens prepared at the central workshop of Delhi Technological University, India. The following conclusions have been made from present investigation.

1. Fabrication of copper-carbon nano tube composites by friction stir processing (FSP) is possible.
2. The microstructure for different specimens showed that after single, double and triple passes the microstructure gets more compact and observed no defects after processing.
3. The brinell hardness number comes out to be highest for third pass in comparing with single and double pass.

REFERENCES

1. Bahram A. Kheyavi, Abdolhossein Jalali Aghchai, Mohammadreza Arbabtafti, Mohammad Kazem Besharati Givi, Jalal Jafari (2014): Effect of friction stir processing on mechanical properties of surface composite of Cu reinforced with Cr particles, *Advanced Materials Research* Vol. 829: 851-856
2. B.C. Liechty, B.W. Webb, (2008): Flow field characterization of friction stir processing using a particle-grid method, *Journal of materials processing technology* 208: 431-443
3. Guangyu Chai, Ying Sun, Jianren 'Jenny' Sun and Quanfang Chen (2008): Mechanical properties of carbon nanotube-copper nanocomposites, *IOP Science, Journal of Micromechanics and Microengineering* Volume 18 Number 3, Guangyu Chai *et al* 2008 *J. Micromech. Microeng.* 18 035013
4. H.R. Akramifard, M. Shamanian, M. Sabbaghian, M. Esmailzadeh, H.R. Akramifard, M. Shamanian, M. Sabbaghian, M. Esmailzadeh (2008): Microstructure and mechanical properties of Cu/SiC metal matrix composite fabricated via friction stir processing *Materials and Design*, 2013,
5. H. Sarmadi, A.H. Kokabi, S.M. Seyed Reihani, Friction and wear performance of copper-graphite surface composites fabricated by friction stir processing (FSP), *Wear* 304 (2013) 1-12
6. Galvao, A. Loureiro and D. M. Rodrigues, Influence of process parameters on the mechanical enhancement of copper-DHP by FSP, *Advanced Materials Research* Vol. 445, 2012, pp 631-636
7. Jian-Qing Su, TW Nelson, TR McNelley, R S Mishra, Development of nanocrystalline structure in Cu during friction stir processing (FSP), *Materials Science and Engineering: A*, Elsevier, 2011/6/25, 5458-5464
8. Kudzanayi Chiteka, Friction Stir Welding/Processing Tool Materials and Selection, *International Journal of Engineering Research & Technology*, Vol. 2 Issue 11, 2013, pp 8-18, ISSN: 2278-0181
9. K. Surekha, A. Els-Botes, Development of high strength, high conductivity copper by friction stir processing, *Materials and Design* 32 (2011) 911-916
10. Lucie B Johannes, Leonard L Yowell, Edward Sosa, Sivaram Arepalli and Rajiv S Mishra, Survivability of single-walled carbon nanotubes during friction stir processing, *IOP Science, Nanotechnology*, Lucie B

- Johannes *et al* 2006 *Nanotechnology* 17 3081. doi:10.1088/0957-4484/17/12/044, Received 6 February 2006, in final form 20 April 2006. Published 2 June 2006. IOP Publishing Ltd
11. Mohsen Barmouz , Mohammad Kazem Besharati Givi, Fabrication of in situ Cu/SiC composites using multi-pass friction stir processing: Evaluation of microstructural, porosity, mechanical and electrical behaviour, *Composites: Part A* 42, 2011, pp 1445–1453
 12. Mohsen Barmouz, Mohammad Kazem Besharati Givi, Javad Seyfi, On the role of processing parameters in producing Cu/SiC metal matrix composites via friction stir processing: Investigating microstructure, microhardness, wear and tensile behaviour, *Materials Characterization* 62, 2011, pp 108 – 117
 13. P. Xue, B.L. Xiao, Z.Y. Ma, Achieving Large-area Bulk Ultrafine Grained Cu via Submerged Multiple-pass Friction Stir Processing, *J. Mater. Sci. Technol.*, 2013, 29(12), 1111–1115
 14. Q. Zhang, B.L. Xiao, P. Xue, Z.Y. Ma, Microstructural evolution and mechanical properties of ultrafine grained Al₃Ti/Al–5.5Cu composites produced via hot pressing and subsequent friction stir processing, *Materials Chemistry and Physics* 134, 2012, pp 294– 301
 15. Rajiv S. Mishra, Murray W. Mahoney, *Friction Stir Welding and Processing*, ASM International, 2007
 16. Rajiv S Mishra , Bulk nanomaterials from friction stir processing: Features and properties, *Book: Bulk Nanostructured Materials*, Wiley-VCH Verlag GmbH & Co KGaA Weinheim, 255-272
 17. RS Mishra, H Jones, GW Greenwood ,An empirical correlation for the grain-boundary diffusion of impurities in copper, *Journal of materials science letters*, 1988/7/1, 728-730
 18. R Sathiskumar, N Murugan, I Dinaharan and S J Vijay, Role of friction stir processing parameters on microstructure and microhardness of boron carbide particulate reinforced copper surface composites, *Indian Academy of Sciences, Sadhana Vol. 38, Part 6*, 2013, pp. 1433–1450
 19. R. Sathiskumar, N. Murugan, I. Dinaharan, S.J. Vijay, Characterization of boron carbide particulate reinforced in situ copper surface composites synthesized using friction stir processing, *Materials Characterization* 84, 2013, pp 16 – 27
 20. R. Sathiskumar , N. Murugan , I. Dinaharan , S.J. Vijay, Prediction of mechanical and wear properties of copper surface composites fabricated using friction stir processing, *Materials and Design* 55, 2014, 224–234
 21. Ranganath. M. S, R. C. Singh , Rajiv Chaudhary, R. K. Pandey, Experimental Investigation of Friction and Wear Behavior at the Interface of Aluminium and Mild Steel, *International Journal of Advance Research and Innovation*, Volume 2, Issue 4 (2014) 775-780
 22. Salar Salahi and Vahid Rezazadeh, Fracture Mechanism in Friction Stir Processed Annealed Pure Copper Samples, *World Applied Sciences Journal* 23 (12): pp 54-58, 2013, ISSN 1818-4952
 23. Scott F. Miller, New friction stir techniques for dissimilar materials processing, *Manufacturing Letters* 1 (2013) 21–24
 24. V. Jeganathan Arulmoni, Ranganath M S , R S Mishra, Effect of Process Parameters on Friction Stir Processed Copper and Enhancement of Mechanical Properties of the Composite Material: A Review on Green Process Technology, *International Research Journal Of Sustainable Science & Engineering*, IRJSSE, Volume 2, Issue: 4, (2014) ISSN: 2347-6176
 25. V. Jeganathan Arulmoni, R. S. Mishra, Ranganath M. S, Experimental Investigations on Friction Stir Processed Copper and Enhancement of Mechanical Properties of the Composite Material, *International Journal of Advance Research and Innovation*, ISSN 2347 – 3258, Volume 2, Issue 3 (2014) 557-563
 26. Valentin N. Popov, Carbon nanotubes: properties and application, *Materials Science and Engineering*, R 43 (2004) 61–102
 27. Yan-Hui Li, William Housten, Yimin Zhao and Yan Qiu Zhu, Cu/single-walled carbon nanotube laminate composites fabricated by cold rolling and annealing, *IOP Science, Nanotechnology*, Yan-HuiLi *et al* 2007 *Nanotechnology* 18 205607. Published 23 April 2007. IOP Publishing Ltd
 28. Yong X. Gan, Daniel Solomon 2 and Michael Reinbolt, Friction Stir Processing of Particle Reinforced Composite Materials, *Materials* 2010, 3, 329-350.

Utilization of an Industrial Waste Product from Thermal Power Plants in Civil Construction

Alok Verma¹

¹(Civil Engineering, Delhi Technological University, Delhi, India)

Email: alokverma@gmail.com

Abstract : Use of fly ash, obtained primarily from coal combustion for power generation, in concrete is important from the point of view of checking environmental pollution, cost effectiveness. The pozzolanic properties of the ash, including its lime binding capacity, makes it useful for the manufacture of cement, building materials concrete and concrete-admixed products. Depending on their physical and chemical properties and the amount of replacements with cement, fly ashes may provide an economical production possibility in construction industry. This paper reviews considerations and state of art of utilization of fly ash in the civil construction sphere and presents results of its effects on performance of cement mortar in aggressive environments.

Keywords: Fly ash, industrial waste, construction, aggressive

I. INTRODUCTION

Fly ash is an environmentally polluting waste product collected at various thermal power plants in great quantities. It has been used for many purposes such as in brick making, in land filling and in cement based composite materials such as cement mortar and cement concrete as a part replacement of cement or in addition to cement. The content of fly ash as a part replacement of cement in fly ash cement concrete mixes, excluding high volume fly ash concrete, has generally been in the range of 0% to 30% [1, 2, 3, 4; 5]. Effect of fly ash as a partial replacement of cement as well as sand on strength and sorption behaviour of concrete has been investigated [6]. The partial replacement of sand by fly ash is seen to be more effective in case of mixes of low cement content [6]. Ramyar and Inan [7] have presented equations to estimate sulphate expansion depending on mineral admixture substitution level among other parameters provided the admixture has satisfactory composition, fineness and pozzolanic activity.

Effects of utilization of fly ash on various properties of cement based composites have been investigated by many researchers. The pulverized fly ash makes a big contribution in concrete because of its pozzolanic properties and it is reflected in lower heat evolution [8]. Effects of use of fly ash on concrete workability have been investigated [2, 9, 10, 11, 12]. Compressive strength of fly ash concrete has been widely investigated [10, 13-20]. Modulus of elasticity of concrete made with the use of ash has been reported to depend on solution to ash ratio and paste to aggregate ratio [14, 21]. Drying shrinkage aspects connected with use of fly ash in concrete have been studied and it has been reported that the use of fly ash improves the performance of concrete

[10, 14,16, 20]. Freeze-thaw behaviour, thermal expansion properties of cement based materials and self healing of shrinkage cracks have been investigated and beneficial effects of use of fly ash have been reported [9, 19, 21] Effect of fly ash on electrical resistivity of concrete has been investigated and it has compared well with the use of rice husk ash in concrete [22]. Fly ash concretes with 35% and 50% replacement of cement have been found to provide better resistance compared to normal concrete in marine environment [23].

II EXPERIMENTAL PROGRAMME

It was intended to observe the performance of fine fly ash blended cement mortar mixes in different concentrations of acidic environments and compare the performances. The experimental programme included the consideration of the effect of sulphuric acid environments of varying degrees on plain and fly ash blended cement mortar specimens. In the present investigation the water cement ratio was kept constant equal to 0.50. The ratios of cement and sand in mortar were 1:4 and 1:6 by weight. Abbreviations of some typical mortar mix designations are explained below in Table 1.

Mortar cube samples of 70 mm size were prepared as per standard procedure and immersed in two tanks of sulphuric acid of 0.1 N and 0.2 N concentrations respectively. The compressive strength and variation in weight of 5 samples were observed after 28, 90 and 180 days of curing. Variations of compressive strength and mass in specimens and microstructural changes were noted with time.

Table 1 – Explanation Of Some Mortar Mix Designations

Typical mix designation	Explanation
CM4W	Cement Mortar 1:4 Water cured
CM4N1	Cement Mortar 1:4 Normal 0.1 acid environment cured
CM6F30N2	Cement Mortar 1:6 blended with Fly ash 30% Normal 0.2 acid environment cured

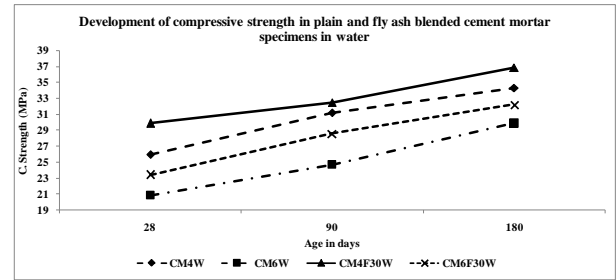


Figure 1 - Compressive strength of plain and fly ash blended cement mortar specimens cured in water

III RESULTS AND DISCUSSION

Compressive strength of plain and blended mortar specimens is shown in Table 2. Generally, the effect of sulphuric acid to degrade the mortar specimens with time and the degradation depends on many factors such as concentration of sulphuric acid, time of exposure to aggressive environment etc. It is observed in Table 2 that compressive strength of any mix increases with age. It is seen that the compressive strength of blended mixes with respect to their respective plain counterparts is more. The trend may be due to the presence of fly ash and the pozzolanic reactions in these mixes.

Table 2 – Compressive Strength Of Plain And Fly Ash Blended Cement Mortar Specimens

Mix Designation	Compressive strength (MPa) at the age			Mix Designation	Compressive strength (MPa) at the age		
	28 days	90 days	180 days		28 days	90 days	180 days
CM4W	26.00	31.20	34.32	CM4F30W	29.90	32.50	36.92
CM6W	20.80	24.70	29.90	CM6F30W	23.40	28.60	32.24
CM4N1	22.75	25.48	27.04	CM4F30N1	26.00	26.91	29.77
CM6N1	18.07	19.24	20.93	CM6F30N1	20.41	22.62	23.27
CM4N2	20.80	23.40	24.31	CM4F30N2	24.05	25.09	26.91
CM6N2	15.73	18.20	20.28	CM6F30N2	18.20	21.71	22.62

In the case of comparison of fly ash blended mixes with their plain cement mortar counterparts, it is seen in Figure 1 that percentage difference of their compressive strengths from those of plain mixes come down with age. It is indicated by the values of 15.00%, 4.17% and 7.58% at the ages of 28 days, 90 days and 180 days respectively for the mix CM4F30W. It indicates the beneficial effect of fly ash in acid aggressive environment due to pozzolanic reactions.

Incremental percentage changes in compressive strengths in various durations for plain and blended mortar specimens in water and acid environments with respect to respective control mixes have been presented in Figure 2.

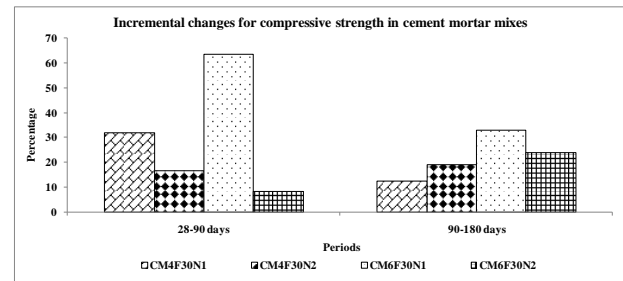


Figure 2 -Incremental changes for compressive strength in cement mortar mixes

It is seen that incremental changes in compressive strengths in the periods of 28-90 days and 90-180 days, as shown in Table 2, are encouraging in the case of blended mixes when exposed to aggressive environments. The incremental changes in both of these periods determine the rate of growth between growths of any mix compared to control mix at different stages. The incremental rates of growth for fly ash blended mixes in acidic environments were all positive indicating the beneficial effect of fly ash in cement mortar. It may also be suggested that incremental growth rates may provide a valuable insight in the analysis of the effect of a blend in a cement based mix.

IV CONCLUSIONS

Following conclusions may be drawn from this review study.

1. Appropriate remedies enable fly ash to be used in cement based materials for the satisfaction of particular requirements concerning strength and durability. It is suggested that causes in variability in characteristics of ash and its effects on the quality of the blended cement based on the variability of calcium and loss on ignition values should be investigated
2. Cement–superplasticizers compatibility can be altered by the physical and chemical characteristics of the mineral additions such as fly ash. It is to be

seen that if the ash contains high chloride content and high loss on ignition values, the ash should be suitably treated

3. Setting behaviour of cement mortar mixes is influenced with the blending of fly ash. This affects the microstructure of the cement mortar paste even after hardening of the paste has started.
4. Use of fly ash in cement mortar mixes has been reported to be beneficial. This is all the more beneficial in the case of exposure to aggressive environments.
5. Many aspects such as effect of variability of fly ash characteristics, such as fineness of fly ash, material characteristics etc should be further investigated.

REFERENCES

1. Singhal Anupam, V.K. Tewari, Satya Prakash, 'Utilization of treated spent liquor sludge with fly ash in cement and concrete', *Building and Environment*, Volume 43, Issue 6, June 2008, Pages 991-998
2. Ozcan F., Atis C.D., Karahan O., Uncuoglu E. and Tanyildizi H., 'Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume concrete', *Advances in Engineering Software*, 40 (2009) 856-863
3. Nath P., P. Sarker, *Effect of Fly Ash on the Durability Properties of High Strength Concrete*, Procedia Engineering, Volume 14, 2011, Pages 1149-1156
4. Nili Mahmoud, A.M. Salehi, 'Assessing the effectiveness of pozzolans in massive high-strength concrete', *Construction and Building Materials*, Volume 24, Issue 11, November 2010, Pages 2108-2116
5. Skaropoulou A., S. Tsvivilis, G. Kakali J.H. Sharp, R.N. Swamy, 'Long term behavior of Portland limestone cement mortars exposed to magnesium sulfate attack', *Cement & Concrete Composites* 31 (2009) 628-636
6. Joseph Glory, Ramamurthy K., 'Influence of fly ash on strength and sorption characteristics of cold-bonded fly ash aggregate concrete', *Construction and Building Materials*, Volume 23, Issue 5, May 2009, Pages 1862-1870
7. Aydın Serdar, Hilmi A. Aytaç, Kambiz Ramyar, 'Effects of fineness of cement on polynaphthalene sulfonate based superplasticizer – cement interaction', *Construction and Building Materials*, 23 (2009) 2402 – 2408
8. Pedersen K.H., Jensen A.D., Rasmussen M.S. Skjoth-, Johansen K. Dam, 'A review of the interference of carbon containing fly ash with air entrainment in concrete', *Progress in Energy and Combustion Science*, Volume 34, Issue 2, April 2008, Pages 135-154
9. Khatib J.M., 'Performance of self-compacting concrete containing fly ash', *Construction and Building Materials*, Volume 22, Issue 9, September 2008, Pages 1963-1971
10. Sahmaran M., Yaman I.O., 'Hybrid fibre reinforced self-compacting concrete with a high-volume coarse fly ash', *Construction and Building Materials*, 21 (2007) 150-156
11. Bakharev T., Sanjayan J. G. and Y. B. Cheng, 'Sulfate attack on alkali - activated slag concrete', *Cement and Concrete Research* 32 (2002) 211 – 216
12. Bensted J., 'A discussion of the paper – Studies about a sulfate resistant cement: influence of admixtures', *Cement and Concrete Research*, 25 (5) 1995 1129-1130
13. Bensted John, 'A discussion of the review paper "Sulphate attack research—whither now?" by M. Santhanam, M.D. Cohen, and J. Olek', *Cement and Concrete Research* 32 (2002) 995-1000
14. Inan G., A.B. Go ktepe, K. Ramyar, A. Sezer, 'Prediction of sulfate expansion of PC mortar using adaptive neuro-fuzzy methodology', *Building and Environment* 42 (2007) 1264-1269
15. Irassar E.F., 'Sulfate attack on cementitious materials containing limestone filler — A review', *Cement and Concrete Research* 39 (2009) 241-254
16. Verma A., Shukla M. and Sahu A. K., 'Application of fuzzy logic to visual examination in the assessment of sulphate attack on cement based materials', *International Journal of Fuzzy Logic Systems (IJFLS)* Vol.3, No2, April 2013
17. Verma A., Shukla M. and Sahu A. K., 'Sulphate attack considerations of use of superplasticizers in concrete', *National seminar on Emerging and Sustainable Technique in Civil Engineering, CTE, Udaipur, October 18-19, 2010*
18. Verma A., Shukla M. and Sahu A. K., 'Influence of Aggressive Chemical Environment on High Volume Fly Ash Concrete', *Concrete Research Letters*, Vol. 4 (2013) 550-556
19. Verma A., Shukla M. and Sahu A. K., 'Performance of Fly Ash and Stone Dust Blended Concrete in Acidic Environment', *Concrete Research Letters*, Vol. 4 (2013) 570-579
20. Wang J. G., 'Sulfate attack on hardened cement paste', *Cement and Concrete Research*, Vol. 24, No. 4, 1994, pp. 735-742
21. Wang Xiao-Yong, Han-Seung Lee, 'A model for predicting the carbonation depth of concrete containing low-calcium fly ash', *Construction and Building Materials*, Volume 23, Issue 2, February 2009, Pages 725-733
22. Yazıcı Halit, 'The effect of silica fume and high-volume Class C fly ash on mechanical properties, chloride penetration and freeze-thaw resistance of self-compacting concrete', *Construction and Building Materials*, Volume 22, Issue 4, April 2008, Pages 456-462
23. Zingg Anatol, Frank Winnefeld, Lorenz Holzer, Joachim Pakusch, Stefan Becker, Renato Figi, Ludwig Gauckler, 'Interaction of polycarboxylate-based superplasticizers with cements containing different C3A amounts', *Cement and Concrete Composites*, Volume 31, Issue 3, March 2009, Pages 153-162

VOLTAGE MODE SECOND ORDER NOTCH/ALL - PASS FILTER REALIZATION USING OTRA

By

RASHIKA ANURAG *

NEETA PANDEY **

ROHAN CHANDRA ***

RAJESHWARI PANDEY ****

* Associate Professor, Department of Electronics and Communication Engineering, JSS Academy of Technical Education, NOIDA, India.

** **** Associate Professor, Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.

ABSTRACT

This paper presents a Second Order Notch/All pass filter based on Operational Transresistance Amplifier (OTRA). It uses two OTRA's and five resistances. To add electronic tunability, the filter uses a capacitor array that can be controlled by switches. The switches are also used to provide inverting and non inverting Notch/All pass as response. This adds flexibility in phase response of all pass filter. The notch/all pass configuration is a modified extension of the Delyiannis-Friend circuit. Through the addition of a second active block which basically acts as a summer for the input voltage and the output of the first OTRA block. The functionality of the proposed filters is verified through SPICE simulations using CMOS based implementation of OTRA. The power supply for the implementation is 1.5V and is based on 0.5 submicron technology.

Keywords: Notch Filter, All Pass Filter, Delay Equalizer, OTRA.

INTRODUCTION

The OTRA (Cam, Kacar, Cicekoglu, H. Kuntman and A. Kuntman, (2003); Chen, Tsao, Liu and Chiu, (1995); Chien, (2014); Lo, Chien and Chiu, (2009); Mostafa and Soliman, (2006); R. Pandey, N. Pandey, Sriram and Paul, (2012); Salama and Soliman, (1999, 2000)) has emerged as an alternate analog building block since it inherits all the advantages offered by current mode techniques. The OTRA is a high gain current input voltage output device. Due to low impedance input and output terminals, limitations on the response by the time constants of the capacitors are reduced. The device is unaffected by stray capacitances due to virtually grounded inputs (Chien, 2014). Ideally, the transresistance gain of OTRA approaches infinity and external negative feedback must be used which forces the input currents to be equal.

The continuous time filters are imperative in signal processing applications. The second order filters being basic building block in realisation of higher order filters, are important class of circuit. The notch filters are used to suppress the magnitude of a particular frequency while passing other frequencies un-attenuated. The all pass filter also known as delay equaliser is widely used for linearizing the phase. The literature survey on OTRA based

second order notch/all pass filters (Cakir, Cam and Cicekoglu, (2005); Chang, Ko, Guo, Hou and Horng, (2015); Chen, Tsao and Liu, (2001); Chen et al., (1995); Gokcen and Cam, (2009); Gokcen, Kilinc and Cam, (2011); Kilinc and Cam, (2005); Kilinc, Keskin and Cam, (2007); R. Pandey et al., (2012); Salama and Soliman, (1999) is comprehended in Table 1 which shows that the structures,

- use excessive numbers of active components (Chen et al., (1995); Gokcen et al., (2011); R. Pandey et al., (2012));
- large number of capacitances (Chang et al., (2015); Chen et al., (2001); Chen et al., (1995); Gokcen and Cam, (2009); Kilinc and Cam, (2005); Kilinc et al., (2007); Salama and Soliman, (1999)) and resistances (Gokcen et al., (2011); R. Pandey et al., (2012));
- requires negative resistance (Salama and Soliman, (1999) and hence will require additional active components;
- cannot provide filter responses with equal capacitance values (Cakir et al., (2005); Gokcen and Cam, (2009); Kilinc and Cam, (2005); Kilinc et al., (2007))
- can provide only notch response (Chang et al., 2015)