A Dissertation on

ARTIFICIAL NEURAL NETWORK BASED ASSESSMENT OF GROUNDWATER QUALITY IN DELHI REGION, INDIA

submitted for partial fulfillment of requirements for the degree of

MASTER OF TECHNOLOGY
IN
ENVIRONMENTAL ENGINEERING

SUBMITTED BY:
VISHAL
(2K21/ENE/14)

Under supervision of
PROF. SK SINGH
ASST. PROF. ANUNAY GOUR

DEPARTMENT OF ENVIRONMENTAL ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

MAY 2023

DELHI TECHNOLOGICAL UNIVERSITY

(*FORMERLY* DELHI COLLEGE OF ENGINEERING)

MAIN BAWANA ROAD, NEW DELHI -110042

DEPARTMENT OF ENVIRORNMENTAL ENGINEERING

CERTIFICATE

This is to certify that the research work embodied in this dissertation entitled "**ARTIFICIAL NEURAL NETWORK BASED ASSESSMENT OF GROUNDWATER QUALITY IN DELHI REGION, INDIA"** has been carried out in the Department of Environmental Engineering, Delhi Technological University, New Delhi.

ASST. PROF. ANUNAY GOUR **PROF. S.K. SINGH**

CANDIDATE'S DECLARATION

I, Vishal, Roll No. 2K21/ENE/14 student of MTech (Environmental Engineering), hereby declare that the project dissertation titled " **ARTIFICIAL NEURAL NETWORK BASED ASSESSMENT OF GROUNDWATER QUALITY IN DELHI REGION, INDIA**" which is submitted by me to the Department of Environmental Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology is original and not copied from any source without proper citation. This work has not previouslyformed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

PLACE: VISHAL

DATE:

ACKNOWLEDGEMENT

It gives me immense pleasure to take the opportunity to thank our Head of Department, Environmental Engineering, **Prof A.K. Haritash** for providing the laboratory facilities to carryout the project.

It is of extreme pleasure to express my deep sense of gratitude and indebtedness to my research supervisor **Prof. S.K. SINGH** and **ASST. PROF. ANUNAY GOUR** Department of Environmental Engineering, DelhiTechnological university (Formerly Delhi College of Engineering) for his invaluable guidance, encouragement and patient reviews. Without his helps and guidance, this dissertation would have been never possible. I am also thankful to all other teachers of the department who directly or indirectly helped me in completion of my project successfully.
A special word of thanks to all my friends for their help and support in my work.

VISHAL

INDEX

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ANN	Artificial Neural Network
GQI	Ground Water Quality Index
WQI	Ground Water Quality Index
MAE	Mean Absolute Error
MSE	Mean Square Error
ANFIS	Adaptive Neuro Fuzzy Inference System

SUMMARY

Groundwater is a crucial natural resource that supports both ecosystems and human activity. For the purpose of providing clean drinking water sources and environmental protection, it is essential to evaluate the quality of groundwater. In order to evaluate groundwater quality accurately and effectively, this study suggests a framework called the GQI that uses ANN. By capturing the intricate linkages and nonlinear patterns seen in the dataset, the ANN architecture makes it possible to forecast the quality of groundwater depending on the input parameters. The model goes through a thorough training procedure where the biases and weights are optimised to reduce prediction mistakes. A number of benefits come from using ANN in the GQI technique, including flexibility, adaptability, and the capacity to handle sizable and complicated datasets. The suggested paradigm offers a useful tool for policymakers, managers of water resources, and decision-makers to analyse and manage the quality of groundwater, enabling prompt responses to safeguard both human health and the environment. Once trained, the ANN model may be used to aggregate the normalised values of the water quality parameters and determine the GQI for every given groundwater sample. The GQI offers a thorough and succinct depiction of groundwater quality, making it simple to compare and comprehend the findings.

The study has been carried out with the objective to determine the WQI for duration of 8 years i.e., 2014 to 2022 and model and predict the WQI by ANN approach. The obtained results using the ANN model in the training phase are observed to be substantial, and optimistic with extremely high association having an R-square value of 99.99% - 100% thus exhibiting the suggested program very high efficiency.

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION:

Water is crucial since it is both the most essential element for the existence of life and the element that is found in the greatest quantity in the earth's environment. Understanding the fundamental principles guiding the distribution and flow of water. In Delhi State, groundwater is frequently used to supplement erratic pipe borne water as a reliable supply of drinking water. It has contributed significantly to most societies' social well-being while frequently going unnoticed.

Consuming polluted water contributes to a growing number of health problems throughout the world, most of which increase rates of morbidity and death, particularly in children, and constitute a severe threat to public health in the form of mass epidemics in emerging nations.(Wagh et al., 2018). Additionally, every year 250 million individuals get infected, with 10-20 million dying, largely in developing countries. In recent past years, various experts have demonstrated that the quality of most nations' groundwater has noticeably declined. Nowadays, people use groundwater extensively for drinking, irrigation purposes,and commercial and industrial uses due to a lack of fresh water resources. Due to its natural, concealed existence and lower susceptibility to pollution than surface water, groundwater is generally regarded as a safe and reliable supply of drinking water.(Khalil et al., 2011).

In other instances, it is utilised without treatment since it is thought that the ground contains built-in filtering capabilities and processes for removing gases, dissolved compounds, and particle matter. Because water is a good solvent and has many opportunities to dissolve compounds as it flows through rocks and underlying soil, groundwater is vulnerable to contamination by human activities. The character has an impact on the quality of groundwater supplies.

On the land surface, the timing and amount of stream flow depends on the precipitation, wind, air temperature, topography features, vegetation cover, soil types, and various other factors. The relationship between rainfall, runoff, infiltration, evapotranspiration, and groundwater is very necessary and important but difficult to elaborate. Additionally, uneven warming of the earth's surface results in diurnal phenomena like land and sea wind and redistributes heat.

of groundwater to non-point source (NPS) contamination It is still difficult to accurately and economically model the susceptibility at the regional level. Numerous hydrological studies have utilised geographic information systems (GIS), neural networks (NNs) and fuzzy logic approaches. The benefit of this approach is that it provides a way to deal with imperfect data, making it an attractive alternative for regional and continental environmental modelling when imperfect data predominates.

Water is an important natural resource and second only to fresh air for life on earth. About 60-65% of the human body is made up of water. It exists in all three states of matter such as solid (ice), liquid and gaseous. The liquid form of water is available for use as surface, ground or underground water, but unfortunately the levels of drinking water and available groundwater are decreasing around the world due to the imbalance between the replenishment capacity of aquifers and the rate of water abstraction at the source.

Water scarcity with respect to quantity and water pollution with respect to quality has become major issues in developing as well as developed nations. The current situation in Asian developing countries is highly variable and majorly affected by socioeconomic and physical conditions along with various phases of civilisation and industrialization.(Munoz-arriola, 2020).

Industrialisation, urbanisation, religious activities such as bathing and immersion of idols in sacred rivers and illegal settlements in India have led to the pollution of groundwater and surface waters, increasing the number of waterborne diseases. Groundwater is considered to be the purest and most important source of water available and is used alongside irrigation to meet India's water demand of 50% urban and 80% rural. Groundwater is also used to irrigate about two-fifths of India's total agricultural area. Groundwater quality is deteriorating day by day due to declining water table, inadequate sanitation, introduction of chemical compounds, inefficient or less efficient irrigation practices, poor management of industrial waste and of the leachate mixture produced by the indiscriminate disposal of industrial waste on the ground. The seriousness of groundwater contamination depends on the type and quantity of waste, the methods of disposal, the climatic and hydrological characteristics of the aquifer, the recharge capacity of the area and the speed with which the water is expelled. A quality standard defines acceptable levels of concentrations for contaminants in water used for various purposes, e.g. drinking water purification, irrigation, aquaculture, etc.

Therefore, since groundwater is mainly used to meet domestic needs, it should be adjusted to domestic water quality standards. Groundwater quality should be checked from time to time to provide drinking water. The situation requires immediate attention and we need to start at the smallest possible level to fix the problem. Currently, for drinking purpose, Delhi city suffers due to limited source of supply for drinking, a big portion that is used in Delhi as groundwater source has been removed to satisfy various needs such as household, irrigation needs, agricultural and industrial needs. Groundwater extraction on a big scale has not only upset the demand-supply equilibrium of the groundwater as a resource, but it has also brought up important issues over its quality and quantity. In today era, Delhi is suffering from several acute problems such as bad quality of water contamination and overuse of water. In terms of groundwater development, Delhi has been deemed overexploited(Chatterjee et al., 2009)

WQI is an extremely used technique to ensuring adequate amount and quality of water. It promotes sustainable management practises, which help to provide access to clean, affordable drinking water. Despite efforts to develop a widespread, consistent, and cost-effective method, determining the quality of surface water remains difficult. Plenty ecological informatics frameworks have been developed to meet the demand for dependable data gathering for effective river water quality monitoring through the use of a mix of software and hardware. 4 Despite the fact that the informatics methodology has grown greatly over the course of time, an effective WQI is tool necessary for translation to the obtained data into relevant water quality of river categorization information. Also, WQI is necessary to scaling up the ecological informatics architecture in order to give accurate and timely knowledge to the common public, academics, and water resource management.

The water pollution mitigation (WPM) initiative is based on WQ management. WPM assesses the category of WQ based on the current WQ state and the laws for dumping contaminants into the river. Furthermore, numerous water quality models, such as stochastic and deterministic mathematical models, have been build up in accordance to give the best strategies for preserving water quality. Obtaining an effective and exact water quality model with complicated water resources remains problematic due to variation and complexity in the actual world, ambiguities in the model's framework and parameters, and deviations in field data. As a result, traditional data processing methods are no longer efficient sufficiently for dealing with water quality challenges. Furthermore, efforts are necessary for increasing the uniformity of the model's outcomes. These models, have drawbacks since all of them rely on data sets that demand significant quantity of time for processing in addition to large quantity of input information which is frequently not known.

Deterministic frameworks attempt to characterise all physico-chemical procedures in statistical, utilising inputs derived from historical data, collected through empirically or derived by experience or investigation. In general, differential equations are streamlined in order to discover solutions that are appropriate for model. The solutions for related equation can entail suppositions and simplifications drawn from the model's efficiency, and typically experimental experience is required from the researcher previously to achieving ideal results. Statistical models explore common principles from observed data, which may be accomplished by acquiring field data. Modelling and evaluation by statistical use, necessitate a careful picking of tools for computation, as well as the validation of hypotheses and data. The bulk of these kinds of models are highly complicated and requires a representative quantity of site data to execute the research. Furthermore, statistical-based models of the quality of water that presume a relationship between the simulation and response variation are commonly and linearly distributed. Nonetheless, various characteristics can have an influence on WQ; current data operation approaches are not any more efficient sufficiently to solve this situation, and as a result, parameters have a complicated non-linear relationship to WQ simulation parameters.

Recently, AI (Artificial Intelligence) technique has emerged as a viable alternative way for describing complex nonlinear systems. In general, these frameworks do not include the underlying process and instead create models based on the correlation of inputs and outcomes. AI is currently being utilised extensively for assessing numerous related to water locations (Engineering et al., 2006). Now-a-days, AI has developed strategies for processing optimisation and equipment preference, as well as issue solutions that entail enormous amounts of data that humans cannot comprehend for the determination of law manufacturing. AI technologies are capable of replicating this action and compensating for deficit. Thus, the advancement of parallel computing techniques and increasing computer power have enabled researchers to utilise AI methodologies such as ANN and ANFIS for experimental data modelling solutions (Chang et al., 2005). Neuro-fuzzy approach had been applied successfully in water body engineering domains such as the basin operation model and rainfall-runoff. ANFIS has been shown to improve the accuracy of day-to-day evaporation estimation, reservoir water level forecast, and river flow prediction(Kişi, 2006) .

1.2 Objectives:

This research has many major goals, in consideration of this, the primary goals of this research were to:

1. Prediction of ground water quality index of Delhi region using ANN
2. Comparison of WQI by weighted arithmetic method and ANN

1.3 Organization of thesis

This thesis is provided as detailed below in five sections:

Chapter 1 bring forth the introduction to surface water quality and its indies. ***Chapter 2*** reviewed the literature to provide the work of various scientists in the field of the Water Quality Index. ***Chapter 3*** includes a short overview of the methodology used in this research. ***Chapter 4*** gives the results of surface water quality index and its model utilizing ANN. ***Chapter 5*** provides the study's significant findings as well as future scope.

CHAPTER 2
LITERATURE REVIEW

Since the middle of the 1800s, WQIs have been used to classify WQ, establishing the WQI idea during the past 50 years. In 1965, Horton was the first to classify WQ. Based on 10 water quality parameters deemed significant in most bodies of water. adjusted the preexisting Horton WQI and produced a new, general Water quality index (WQI). The choice for parameters and weight evaluation were given advice by a steering committee comprising 142 WQ professionals. The NSF-WQI is now recognised as the cornerstone for a number of further WQI models.

A steering committee of 142 WQ professionals provided advice on parameter preference and weight assessment, and (Beek, 2015) modified Horton's existing WQI to create a generic WQI in 1970. The NSF-WQI has since become the cornerstone for a number of additional WQI models.

In the middle of the 1990s, the CCME recognised the British Columbia (BC) WQI model. The US NSF-WQI, the Florida Stream WQI (FWQI), the British Columbia WQI (BCWQI), the CCME, and OWQI are WQIs that are regularly utilised. several studies on the quality of river water. (RWQ) have been carried out using the NSF-WQI methodology. Many countries used the NSF-WQI and its variations as the WQI paradigm for monitoring the RWQ.

(Patil et al., 2015)based on thirteen WQPs, constructed a pollution index. Mathematical calculations were used to transform the pollutant concentration values into new units. The extra units were designed such that they were proportional to the polluting units. The index score will be significantly affected by a specific pollutant in this technique even if its polluting effect is greater even after accounting for that parameter's concentration. A simple WQI with eight pollution parameters was called the McDuffie River Pollution Index (RPI). Six of the eight sub-indices were specifically described by as linear operations.

(Tunc Dede et al., 2013) assessed the water quality of the Kirmir basin in Turkey using data of 10 sample stations over the period of 1 year. All water samples were analysed for forty-four WQPs. For the selected WQPs, five distinct WQI models were used. According to the study, the CCMEWQI and OWQI produced the greatest outcomes.(Tunc Dede et al., 2013) compared 1-year WQ data from several sample stations in the Kirmir basin utilising Overall Index of Pollution OIP (nine WQPs) and 4 other WQIs. To determine OIP, a total of forty-four WQPs were analysed. During the application of OIP, the eclipsing problem was detected. (Tunc Dede et al., 2013) used 44 WQPs to assess the water quality of the Kirmir basin in Turkey. On the specified WQPs, five distinct WQIs were used. According to the findings of this study, the CCME WQI and O WQI may utilised to measure quality of water in studied region. (Tunc Dede et al., 2013) compared 10 WQ parameters to determine water quality of the Kirmir basin utilizing Universal Water Quality Index (UWQI) and four additional WQIs. UWQI employed weight factors ranging from 1 to 4. Despite the fact that UWQI had an eclipse problem, its behaviour in analysing WQ data from the research region was judged to be satisfactory.

(Rakhshandehroo et al., 2012) found that for the purpose of predicting groundwater levels in the Shiraz Plain, four different neural network designs were tested, and the outcomes of each were compared using the statistical metrics of mean square error and square of correlation coefficient. The input data used to anticipate groundwater level in the following time step as an output of the networks included effective parameters on groundwater level such as 5-month precipitation and groundwater level histories, temperature or evaporation, and runoff. All networks were trained using data from the years 1993 to 2003, and they were calibrated using data from the months of April 2003 to September 2004. On the basis of ground-water level monitoring made for a further 18 months (from October 2004 to March 2006) in 29 wells situated in the plain, networks were confirmed. The results suggested that artificial neural networks may be used to predict groundwater levels. All wells' groundwater levels were predicted by various networks with acceptable root mean square errors of 0.6–12.17 m. Feed-forward neural networks had the greatest overall performance, and the second-best by neural network Elman.

(Mohanty et al., 2015) deals with the use of an ANN technique to anticipate groundwater levels weekly in a number of wells spread throughout a river basin. Gradient descent with adaptive momentum at 18 locations across the research region, the learning rate backpropagation (GDX) algorithm was used to forecast groundwater levels one week in advance. A proper suitable set of input data for the ANN model was chosen based on the domain expertise and related statistical analysis. Total 18 output nodes and 40 input nodes were taken as a result of the weekly rainfall, pan evaporation, river stage, water level in the surface drain, pumping rates of 18 stations, and groundwater levels of 18 stations. This produced 18 output nodes and 40 input nodes. The ideal number of hidden neurons

for the ANN model during training was discovered to be 40, and the model's performance was deemed adequate (RMSE= NSE = 0.9722, r = 0.9861, and m = 0.2397). The values of the statistical indices RMSE, r, and NSE during model testing were 0.4118 m, 0.9715, and 0.9288, respectively. The created ANN model was then used to anticipate groundwater levels in 18 tubewells for the next 2, 3, and 4 weeks using the same inputs.
(Iqbal et al., 2020) for comparing the constructed Levenberg-Marquardt back-propagation ANN models, Coefficient of Determination(R), Mean Square Error (MSE), Mean Absolute Error (MAE) were used.. The findings (TT-8-24-1 for pre- and post-monsoon) demonstrate that with MAE, MSE, and R values of 0.0338, 0.0023, and 0.97 for the case of pre-monsoon and 0.031, 0.0021, and 0.974 for the case of post-monsoon, respectively, it was determined that an optimistic ANN model with a single hidden layer, 24 neurons, 80% of the data for training, 10% for validation, and 10% for testing and using tangent sigmoid activation function was used.
A research investigation (Kazi et al., 2009) found that when employing 2 rates for same index for evaluating a water body of surface at different stations, such techniques aid in variable categorization. While one indexing approach demonstrated 9 acceptable categorizations for the variable, a different indexing technique demonstrated low quality. As a result, it may be concluded that there is no optimum WQI approach. Usage of WQI is fully dependent on what sources were used, the parameters measured, the weights allocated, their categorization system, and their ultimate analysis of acquired WQI. Bonanno & Giudice (2010) utilized NSF-WQI for Imera Meridionale river surveillance and administration, using 9 water quality parameters (WQPs). The findings established that the NSF WQI was a reliable tool for effectively managing river ecosystems. Barceló-Quintal et al. (2013) used the NSF WQI to assess the water quality of the Lerma River and compared the results to those of the Mexican National Water Commission. The index was created using 7 WQPs obtained from 4 separate sampling stations. The findings indicated that NSFWQI computed almost the similar river quality as the commission of water.
(Daliakopoulos et al., 2005) simulate the groundwater level's declining pattern and make reliable forecasts up to 18 months in the future. Greek island of Crete's Messara Valley was selected for the study because its groundwater resources have been overused for the past fifteen years and the groundwater level has been rapidly dropping. The effectiveness and precision of model prediction are examined for various network architecture types and training procedures. The findings of the various experiments demonstrate that reliable

predictions may be made using a typical feedforward neural network trained using the Levenberg-Marquardt algorithm, with the best outcomes for forecasts of up to 18 months. One such prevalent technique for selecting variables is the Delphi approach, which is based on professional judgement. It has become easier and more accurate to identify characteristics that characterise the varied character of water sources and provide insight into their future WQ while being extremely resilient thanks to the development of new methodologies, such as multi-variation statistical evaluation. The quantity and type of variable 12 must be carefully chosen as a result (Oladipo et al., 2021)[i] .A geometrical technique ultimately replaced the first use of arithmetic means, although this had the unique drawback of not yielding results that could be trusted when the concentration values of the variables chosen were low. The advantage of employing a geometric approach to weights was that results were less affected when extreme scores were used to compare to the mean score generated using an arithmetic method. The stated variables were given weights using the harmonic mean. The suggestions also suggest that the superior to both geometry and arithmetic. As a result, the finalised WQI may contain quite large irregularities in certain variables during water quality monitoring. As computing power utilisation becomes more and more simplified, accumulation has evolved to include new techniques like the use of logarithmic application procedures and the most recent use of fuzzy logic.

A recent study in North West Delhi (Sarma & Singh, 2023) for investigation of groundwater revealed that bicarbonate was the dominant anion while magnesium was the dominant cation. Multivariate analysis using principal component analysis and Pearson correlation matrix indicated that major ion chemistry in the aquifer under study is primarily due to mineral dissolution, rock-water interactions and anthropogenic factors. Water quality index values showed that only 20% of the samples were acceptable for drinking. Due to high salinity, 54% of the samples were unfit for irrigation purposes. Nitrate and fluoride concentrations ranged from 0.24 to 380.19 mg/l and 0.05 to 7.90 mg/l, respectively due to fertilizer use, wastewater infiltration and geogenic processes.

Behmanesh (2015) used the NSFWQI, BCWQI, and thematic WQI to determine the WQ of the Babolroud River. The above-mentioned WQIs were calculated using various combinations of thirteen WQPs. According to the sensitivity study, the NSFWQI with nine WQPs outperformed the BCWQI and thematic WQI in portraying the RWQ. Effendi et al.(2015) used the NSFWQI to calculate the water quality of the Ciambulawung River according to 9 criteria. The results showed that NSF WQI functioned effectively and assessed the river as excellent. Behmanesh (2015) employed thematic WQI, NSFWQI, and BCWQI. Babolroud River WQ review. Earn 13 WQPs to evaluate RWQ. In accordance to provisional research BC WQI was resulted to be weaker than NSF WQI to represent river WQ.

Groundwater quality data of 54 monitoring wells were collected from the National Water Affairs Regulation Authority of Afghanistan (A. Noori & Singh, 2023), including the dry and wet seasons data and spatial distribution maps and temporal variations were developed to interpret spatial trends in groundwater quality and seasonal variations. Out of the 20 assessments of water quality, 14 during the dry and 15 during the wet seasons revealed concentrations higher than the World Health Organization recommends.

(A. Noori & Singh, 2021) study in Kabul, Afghanistan aimed to comprehensively review the spatiotem-poral groundwater distribution, physicochemical and bacteriological specifications of groundwater, and its consumption finds that the mean annual decrease in groundwater levels between 2008 and 2016 is 1.7m and about 33% of supply wells are not in operation due to declining groundwater table.

CHAPTER 3

MATERIAL AND METHODS

For this study, a systematic flow diagram is shown in figure 3.1, consisting of domain of study, data collection, data standardization, parameters selection, development of WQI, and then use ANN model for water quality modelling.

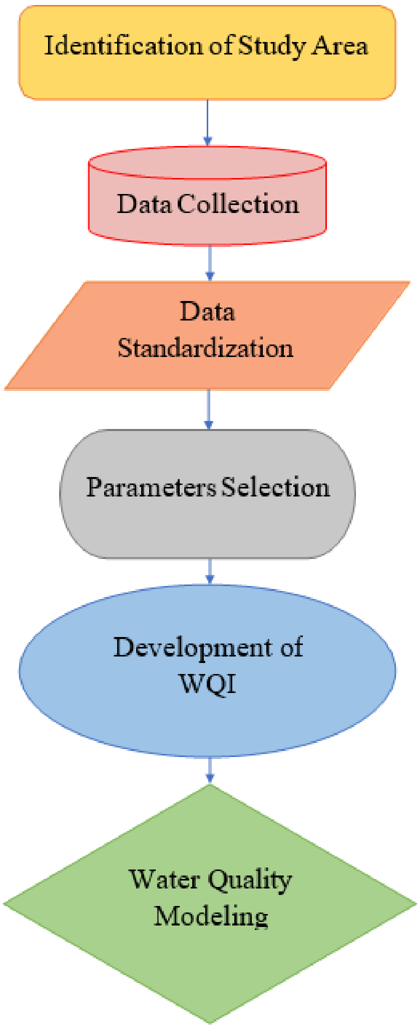

Figure 3.1: Methodology of conducted research work

3.1 Study Area

The National Capital Territory of Delhi located between 28°249170 and 28°539000 N and 77°509240 and 77°209370 E latitudes. It has nine revenue districts and borders two other states: north, west and south to Haryana, north and south to U.P. NCT of Delhi is located on 1483 km^2 of land. The city has a semiarid to humid subtropical climate of , with winters that are typically dry and last from November to January. The monsoon season, which lasts from July to August and is characterised by heavy rains and winds coming from the Arabian Sea, follows the hot and muggy summer months from April to July.

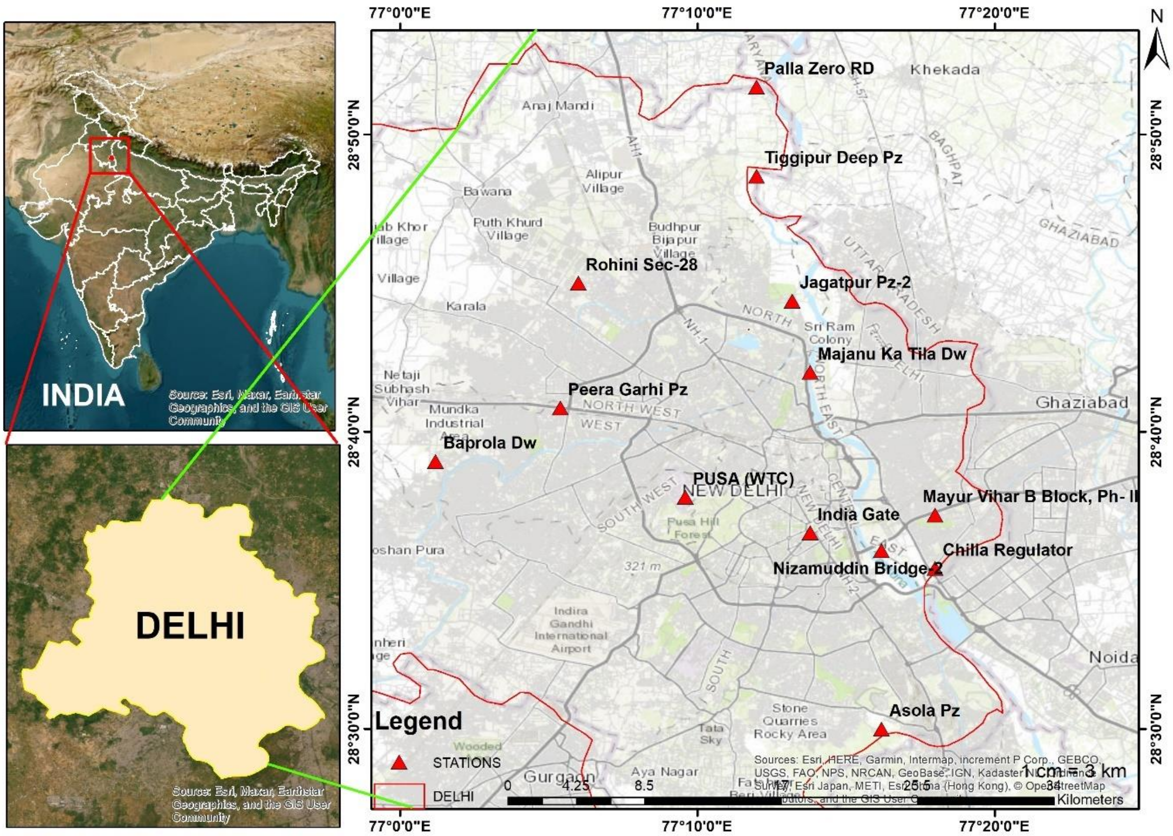

Figure 3.2: Delhi stations considered in present study

3.2 DATA ACQUISITION:

The raw data sets utilised throughout the research investigation have been acquired from the Central Ground Water Board (CGWB) in Delhi, India and consists of 112 samples from 13 distinct location Of Delhi, India. The different locations are:

1. Asola Pz
2. Chilla Regulator
3. Baprola Dw
4. Jagatpur Pz-2
5. Majnu Ka Tila Dw
6. Mayur Vihar B Block Ph11
7. Nizamuddin Bridge -2
8. Palla Zeero Rd
9. Peera Garhi Pz
10. Pusa (Wtc)
11. Rohini Sector 28
12. Tiggipur Deep Pz
13. India Gate

The data was collected for the duration of eight years i.e., from 2014 to 2022. The dataset included 14 important parameters namely are pH, Electrical conductivity (EC), Total Dissolved Solids (TDS), Calcium (Ca^{2+}), Magnesium (Mg^{2+}), Total Hardness, Carbonate (CO_3^{2-}), Bicarbonate (HCO_3^-), Chloride (Cl^-), Nitrate (NO_3^{2-}), Fluoride(F^-), Sulphate (SO_4^{2-}), Sodium (Na^+) and Potassium (K^+).

. However, 14 parameters were considered to show significant values, which are named as: pH, TDS, Alkalinity, Cl^-, NO_3^{2-} , F^- , SO_4^{2-} , Fe, NH_3-N, PO_4^{3-} , BOD, COD, DO and Total Hardness and the created algorithms were assessed using parameters from statistical analysis. The selected parameters are considered on the following basis: Selected parameters represents the overall water quality status, for WQI functionality within the chosen domain, selected parameters should be common in all locations and they should be included in Indian Standards Drinking water specifications (IS 10500: 2012) (BIS, 2012), Central Pollution Control Board (CPCB, 2017)) and World Health Organization (WHO, 2017) drinking WQ standards.

3.3 steps for calculation of WQI:

WQ was measured by employing the WQI, which is derived utilising various characteristics that impact WQ. The suggested system's performance was assessed using a public dataset containing 14 essential water quality metrics. The equation (3.1 – 3.4) that follows was used to determine the WQI were listed step wise below. As tabulated in table 3.2, parameters were assigned a weight (wi) on a scale of one to five based on its importance for drinking acceptability and the well-being of humans. The relative weight (Ri) was calculated for all the 14 parameters. The specified numerical value for Ri is given in Table 3.2.

$$R_i=\frac{W_i}{\Sigma_{i=1}^{n} W_i} \text{ ......} \quad (3.1)$$

For all the factors, a quantity rating scale (S_i) was developed and estimated as:

$$s_i = \left(\frac{Con_i - I_o}{Std_i - I_o}\right) \times 100 \text{ ......} \quad (3.2)$$

The evaluation of sub-index level rating ($SILi$) given below:

$$SIL_i = R_i \times s_i \text{ ......} \quad (3.3)$$

WQI was deliberated as:

$$WQI = \Sigma SIL_i \text{ ......} \quad (3.4)$$

where Con_i is the parameters' concentration in mg/L, Std_i is the standard concentration of each parameter as per BIS, 2012, CPCB, and WHO and n is the number of parameters given tabulated in table 3.2. I_o is the ideal concentration of the parameters in the water. The ideal concentration of each and every parameter is 0 except for pH and DO, for these parameters the ideal value is 7 and 14.8 mg/L respectively. The calculated WQI score was then classified into 5 classes on the basis of table 3.2 shown below:

TABLE 3.1: WQI parameters with assigned weights and acceptable limits used in estimation as per (IS 10500: 2012), as per (CPCB, 2017) and per the (WHO, 2017) drinking water:

S.NO	PARAMETER	W_I	R_I	ACCEPTABLE LIMIT
1	pH	5	0.08	6.5-8.5
2	Total dissolved solids	4	0.07	500
3	Nitrate	5	0.08	45
4	Chloride	4	0.07	250
5	Fluoride	5	0.08	1
6	Sulphate	5	0.08	200
7	Sodium	2	0.08	-
8	Total Hardness	3	0.05	200

All parameters are presented in mg/L except for pH in table 3.2.

TABLE 3.2: SCORE WQI SPECIFIED FOR INDIAN DRINKING STANDARDS

S.NO	WQI SCORE	CLASSIFICATION OF WQI
1.	<50	Excellent Water
2.	50-100	Good Water
3.	100-200	Poor Water
4.	200-300	Very Poor Water
5.	>300	Water unsuitable for drinking purpose

3.4 Artificial Neural Network as WQI Prediction Model:

The ANN is an extremely potent computing approach that may be used to construct a variety of real-world medical applications. ANN models are widely utilised as extremely strong algorithms for machine learning for time series projection in a variety of engineering domains. The ANN model is consisting of 3 layers: input, hidden, and output layer. To manage neurons, every hidden layer includes weights and bias settings. The activation function is implemented to move information from the hidden layer to the output layer. Within neural network (NN) framework, learning methods are utilized to choose weight. The weights were chosen according to minimal performance indicators such as MSE. Figure 3 depicts the FFNN design for the WQC. The ANNs approach were employed for the estimation of WQ in this investigation. There are three major layers in ANNs: input, hidden, and output. For transferring parameters learning from input to output to sigmoid operation, 15 hidden layers were explored.

.

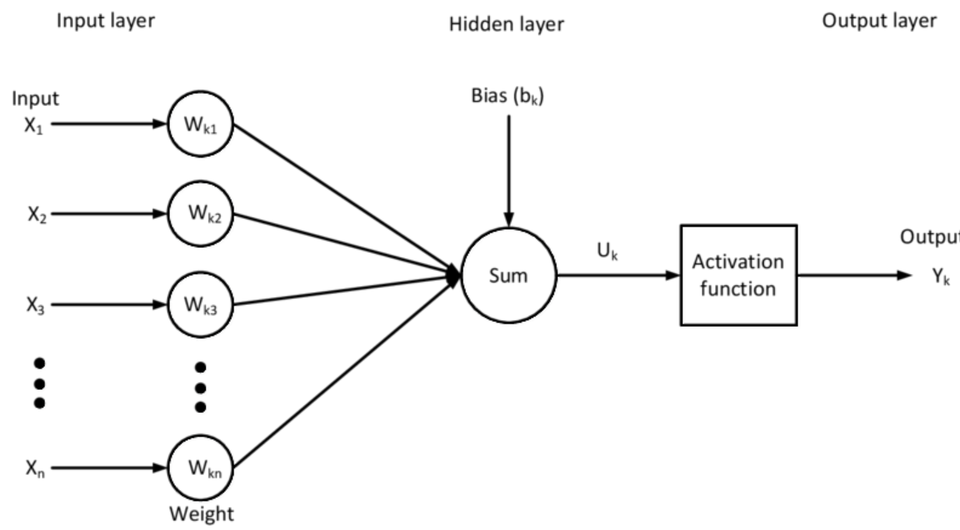

FIGURE 3.3 Architecture of ANN Model for present study

3.4.1.1 *ANN MODEL DEVELOPMENT*:

An innovative and alluring solution to the problem of connecting outputs to input parameters in complex structures (Dawson & Wilby, 2014) is the use of artificial neural networks (ANNs), and estimate is a common justification for doing so. In complicated systems, it could be challenging to link output parameters to input parameters. The use of ANNs offers a creative and alluring solution to this problem. One of the most common applications of this model is prediction.

Establishing appropriate model input, describing network design, preliminary operations and dataset separation, identifying network architecture, expressing model efficiency standards, learning (optimisation of conjunction weighting), and verifying developed model are the main phases in developing ANN models.

3.4.1.2*FEED FORWARD NEURAL NETWORK:*

Artificial neural network with circular node connections is known as a feed-forward neural network. Recurrent neural networks are the exact opposite of feed-forward neural networks, which cycle over various paths. The input is only processed in one direction, making the feed-forward model the most fundamental kind of neural network. Data never moves backwards or in the other direction; it always moves forward. A single layer perceptron is a popular example of a feed-forward neural network in its most basic configuration. A number of inputs are introduced into the layer in this model and multiplied by the weights. The weighted input values are then summed together to get a total. The value generated is frequently 1, and if the total of the values is below the threshold, the output value is -1. The threshold is typically set at zero. In classification problems, the single layer perceptron is a crucial feed forward neural network model. Single layer perceptrons can also contain certain features of artificial intelligence. The neural network may compare the outputs of its nodes with the desired values using a property known as the delta rule, which enables the network to train its weights to create output values that are more accurate. This learning and training procedure results in a gradient descent. Although the process of updating weights in multi-layered perceptrons is almost comparable, it is more formally known as back-propagation. In these circumstances, the network's hidden layers are individually changed in accordance with the output values generated by the top layer.

3.4.2 *NETWORK ARCHITECTURE OPTIMIZATION*:

The current study created a non-linear regression (NLR) framework using WQ monitoring data and a concurrent, feed-forward, fully interconnected multilinear neural network that can be used to accurately mimic the WQ scenario of the Yamuna in the context(Engineering et al., 2006). The two main stages of the development of this model were the selection of the network framework and the specification of the network structure. The Yamuna basin's SWQ condition may be estimated using an NLR technique that is currently being built as a measure of the weighted arithmetic mean. This study used WQ data gathered through monitoring to test an independent, feedforward (FF), fully interconnected multilinear neural network to WQI.

The hidden layers provide the network its ability to generalize. A NN with a hidden layer and a sufficient number of hidden neurons may theoretically anticipate any continuous function, creating a broad and flexible class of general-purpose approximators(Sahoo et al., 2015). The variables in empirical datasets often have a wide range of measurement units, and they are frequently tainted by measurement errors, noise, or disruption. These factors could negatively affect how some ANN training techniques operate.

Then, in this study, an MLP network was built using a NN with settings that included one input layer, fifteen hidden layers, and the WQIs score as the output layer. Datasets based on evidence frequently include variables with a range of measurement units. These components might have a negative impact on how certain ANN training techniques work (B. Singh et al., 2021). To prevent these outcomes, it is necessary to standardize the data at the beginning of the data preparation process, converting the dataset into a 1-D format with a constant range of variation. This prevents any attribute from unnecessarily dominating the NN modeling outputs. In order to pre-process the input WQ data, normalization among the logistic sigmoid function's restrictions, i.e., 0-1 was applied.

The quantity of units for both output and input is typically connected based on the quantity of input forecasters and output parameters(Muthulakshmi et al., 2013). But in ANN modeling, choosing the number of hidden nodes is often a trial-and-error procedure. There are a few broad principles to follow, but there is no magic formula for making decisions. Therefore, the researchers employed the 'Heuristic Search' feature of the Neuro Intelligence ANN program to identify (i) the optimal architectural design for input-WQI simulation; (ii) the most effective method for splitting data. Throughout the investigation, subset preference was random, and the resulting efficiency was one. The degree of variance in the experimental WQI values attributable to the model was the criterion to be used for selecting the best network design because this was an NLR type modeling issue.

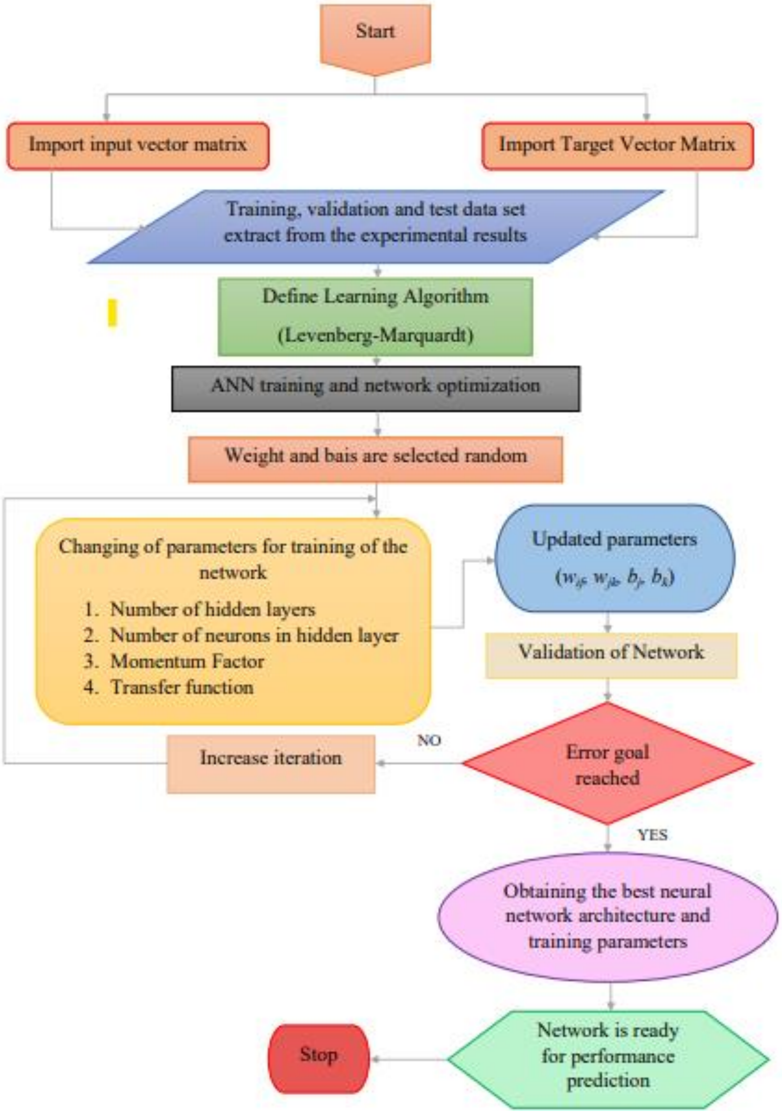

FIGURE 3.4. Flow Diagram of ANN algorithms used in present study

3.4.3 *NETWORK STRUCTURE OPTIMIZATION*:

Testing, training, and cross-validation subsets of the WQ data were created from the optimum network design results. The next stage's main objective was to assess the most effective training method, pace, frequency of iterations, frequency of retrains, as well as training halting conditions. The ideal network structure was found by calibrating each of these elements separately. Until the most precise WQI estimators were established, all model parameters were calibrated by changing one of the preceding parameters while keeping the others constant. The value of the researched parameter was then set to the best score possible and connected, and a new parameter was formed in a similar way until the best scores for each parameter in the model were found. Networking training is a method of the process of continually adjusting a network's synaptic weights and bias degrees by activation from its environment. The link weights in the different networks examined in this study were optimized using the QP approach. QP and the batch propagation (BP) training approach were initially investigated, and their findings were connected.

Initial evaluations in the current study revealed the fact that the QP methodology consistently beat BP technique. The conventional technique manufactured fewer test sets mistakes and stronger R^2 scores more frequently than later. In conclusion, is consistent with actual correlation in the publications, hence, showed that quick prop. learning methodology is one of the quickest and most efficient computations, outperforming the bulk of remaining heuristic variations of the back prop. technique on a wide scope of modelling tasks. As a result, every network underwent training using the conventional QP approach, which has faster stability and convergence and performs fair on analysed dataset compared to BP methodology.

To effectively identify an appropriate training rate, NN are trained as well as validated with dataset's outcomes which are examined. In this procedure, ANN model is taught and continuously optimized in contrast to the cross-validation datasets, allowing the NN structure to be evaluated utilizing the error scores obtained throughout learning on cross-validation dataset. The examine data, on the other hand, is utilized to evaluate ANN model using previously unknown information when model have been constructed. The evaluation of the test set error consistently assesses the network's generalization ability. Learning should be halted when set of test results error achieves the lowest possible level for excellent generalization, and the arrangement of networks to the lowest examine error is called the ideal NN topology. In the account of the conclusions of the optimum network design, the WQ data which consists of 102 datasets. They were separated into a training subset having 88 samples, as well as for cross-validation and testing parts 11 samples were assigned to each. Following that, the network architecture of the current investigation is capable of generating the most precise WQI. The evaluation was carried out by utilizing the strong command of MATLAB 2022b on ANN by using the NN tool. The primary desideratum of this phase was to pick the most appropriate learning techniques, training rate, frequency of iterations, frequency of retraining, and learning halting circumstances. By adjusting these variables one at a time, the optimal network architecture was established. All parameters of the model were validated by altering one of the preceding parameters while holding the others constant until the optimum value of WQI predictions was attained. The investigated parameter's value was therefore set to the optimal point and retained, as well as another parameter was validated in a similar fashion there until the optimal combination of all parameter model’s estimations was discovered.

3.4.4 MODEL SELECTION:

In ANN modelling, one of two basic kinds of choosing models procedures is typically used. The 1st method was according to cross-validation, while the 2nd was in-sample model picking. The given data is divided into three groups in the cross-validation strategy: testing, validation, as well as training. The learning set of data was used for learning the NN, whereas to validate collected dataset is utilized to determine when to halt learning before over-training occurs, with the assumption that a successful model is one that lowers the validity error. The testing set, on the contrary hand, is used for real out of sample assessment, i.e., measuring network efficiency after training is complete, and the actual network error is subsequently determined as the result of the testing error. When using realistic training data, the examining set inaccuracy is an accurate evaluation of the underlying NN performance. As a result, the learning set is utilised to estimate parameters for a variety of distinct neural network specifications. The network that was trained is then tested against the validate set, and the NN model, exhibits the greatest performance on the validate dataset was chosen as the ultimate estimating model. In contrast to the present investigation, the first strategy, which relies on a cross-validation approach, was used. The empirical outcomes were examined using MATLAB 2022b.

3.5 PERFORMANCE MEASUREMENT:

The model's validity, utility, and generalisation performance are assessed on the testing set through the use of an appropriate performance metric like the MSE or the mean absolute error (MAE). One of the finest overall metrics of simulation results is the MSE. It has the benefit of being far less sensitive to outliers (less sensitive to extreme values) than other error metrics. It does not give huge mistakes a high weight, but instead weighs all error amounts uniformly. As a result, the training dataset is utilised to evaluate parameters for a variety of different NN specifications. The trained NN is then tested against the validation dataset, and the NN model along with the greatest performance on the valediction dataset is chosen as the ultimate simulation model (Hanoon et al., 2021). However, in this research, the first method was utilized which is employed on a cross-validation approach. Following that, the model's validity, utility, and generalized performance are assessed on the testing dataset through the use of an appropriate performance metric such as the MSE.

In addition to the previously mentioned error metrics, it is recommended that LR evaluation of simulated dataset and observed dataset may be utilized to objectively and quantitatively estimate the outcomes of validity. The coefficient of correlation (R^2) is a metric of the model's goodness-of-fit to the data that measures the model's simulating power, according to these 26 techniques. As a result, selecting the best of several viable models might be based on the highest value of R2, where the relatively high the R^2 score and the nearer it is to 1.0, the preferable (Yaseen et al., 2019). MSE and LR analysis performance evaluation methodologies were utilized to estimate the proposed model's capacity to predict the WQI in current study MSE and linear regression analysis performance evaluation methodologies were utilized to evaluate the projected model's capacity to simulate WQI in current study. The statistical approaches employed are as:

- Mean Square Error (MSE):

$$MSE = \frac{1}{N} \Sigma_{i=1}^{N} (v_i - \hat{v}_l)^2$$

- Root Mean square error (RMSE)

$$RMSE = \sqrt{\Sigma_{i=1}^{N} \frac{(v_i - \hat{v}_l)^2}{N}}$$

where vi and $\hat{v}i$ are the simulated, the measured values of WQI, respectively, and the total number of parameters represented by N.

- Coorelation Analysis

$$R = \frac{n\Sigma(p \times q) - (\Sigma p)(\Sigma q)}{[n\Sigma(p)^2 - \Sigma(p)^2] \times [n\Sigma(q)^2 - \Sigma(q)^2]} \times 100\%$$

Where, R is the Pearson’s correlation coefficient approach; p is all the input response in the 1st dataset of the training phase; q is all the input response of the 2nd dataset of the training phase; n is known as total response of inputs parameter. As a result, this study distinguished between prospective ANN models established on (i) testing dataset errors and cross-validation (MSE); (ii) co-relation among simulated and reported WQI scores; and (iii) quantity of variation in calculated WQI scores explained by model simulations (R2 scores). Hence, the ANN model presented in study is one with the least testing and cross-validation subset deviations of excellent adjustment, the best co-relation among simulated as well as reported WQI scores, and the greatest R^2 value.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Performance result of ANN

In current study, WQI by employing weighted arithmetic mean method has been calculated. Different weightages were allotted to every parameter in accordance to their effects on human health, severity, and frequency of crossing the limit of acceptance. These weights are assigned on a scale from 1 to 5. Then, relative weight for every parameter was calculated. So that, WQI using weighted arithmetic mean method can be calculated. Following that, a coordinated, fully interconnected ANN with 1 input layer, 15 hidden layer, and the WQI as the output layer was utilized in this investigation. The neurons' responses in the output and input layers were set to 10 (input neurons) and 1 (WQI). The water quality data input was pre-operated by standardization to the scale (0,1). All of this work was done to forecast the WQI of the Delhi region. Furthermore, the power of ANN was used to create additional simulation models for each water quality metric such as pH, TDS, and BOD. To do this, several dataset and parameter combinations are used to create a water quality model for each parameter. Figure 4.1 depicts the error histogram of training state of ANN, having bars of zero error, test, training and validation. It illustrates that the maximum error in observed and simulated results was found to be very close to zero error bar. The extreme error i.e (simulated value differ from original value) of negative value was found to be 0.04899 and positive value was to be 0.022989.

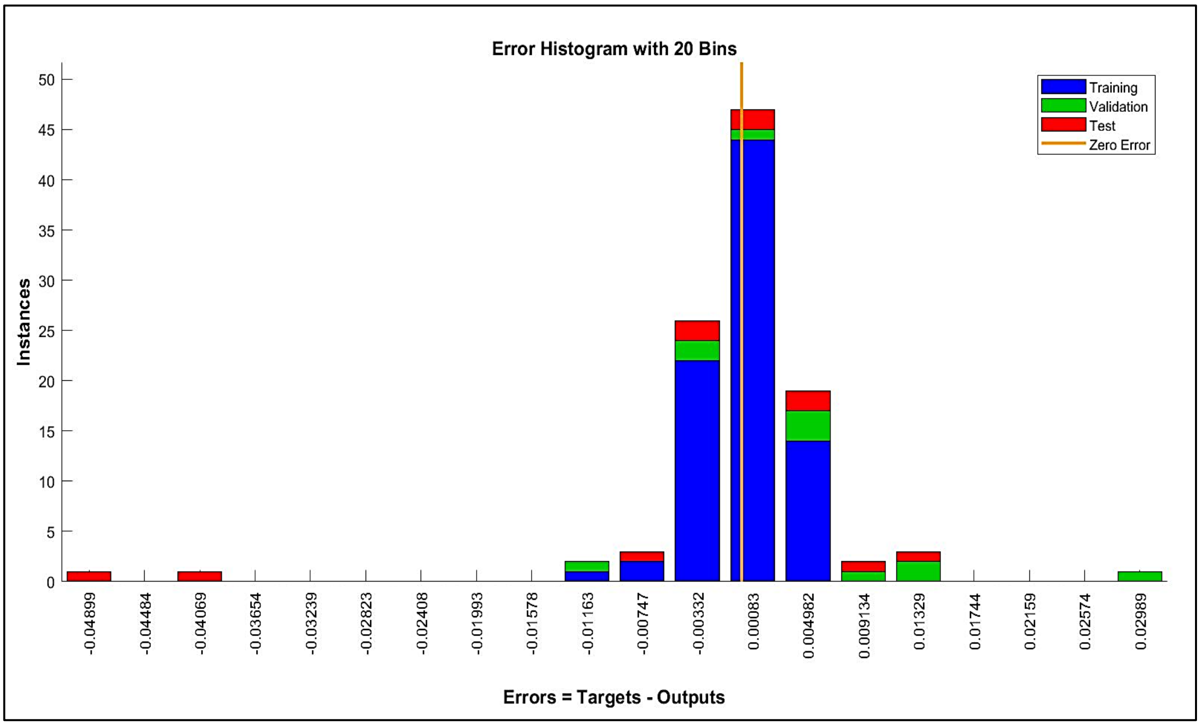

Figure 4.1: Error histogram used in present study

The suggested model technique was validated by training on 80% of total data and implementing ANN model to simulate WQI. The initial learning results demonstrated that ANN model was extremely effective in simulating WQI. Table 3.2 summarises the ANN model's prediction outcomes for the WQI over the testing and learning stages. The simulated results indicated R% = 100%, demonstrating the suggested system's very efficient performance. Based on the correlation regression findings, the simulated outcomes of the ANN model in testing indicated R% = 99.9999%.

TABLE 4.1: Performance of ANN model to simulate WQI in present study

	SAMPLES	MSE	R-Square
TRAINING	83	0.00001334	9.99999e-1
VALIDATION	11	0.00015292	9.99999e-1
TESTING	11	0.00019923	9.99999e-1

Figure 4.2 depicts the ANN model's training state plot. The three graphs are gradient (y-axis) vs epochs (x-axis), mu (y-axis) vs epochs (x-axis), and validation failure (y-axis) vs epochs (x-axis). For system validation, we separated the available data into training and testing sections. The data was split using a random algorithm that selected distinct basic

scores from the total data. The testing phase consists of previously unknown data. It is a procedure for validate the ANN model's competency to forecast water quality. The method picks random numbers from the full dataset to evaluate the model during the testing phase. As per the figure 4.2, validation of data stops at 117 epochs due to 6 consecutive failures after the 113 epochs.

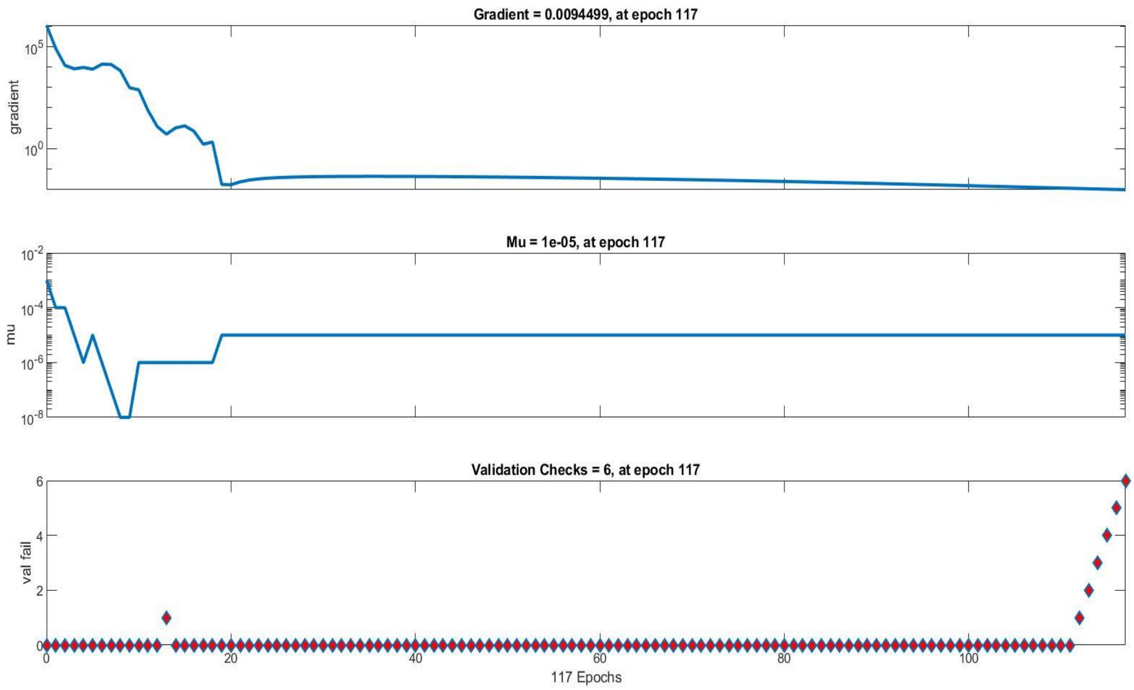

Figure 4.2: Training State Plot of ANN model for WQI assessment in present study

The empirical data given show that the ANN model performs extremely well. The MSE in the WQI network was identified using a performance plot. Figure 4.4 illustrated the effectiveness of ANN model. ANN model's best validation was 0.15371 at epoch 113. The MSE of the ANN model's performance declined dramatically as it learned. The blue, green, and red lines, respectively, show the training process, validation error, and training error. Increased epoch counts suggest that the used training data has minor faults. When the validation fault is eliminated, the training is terminated.

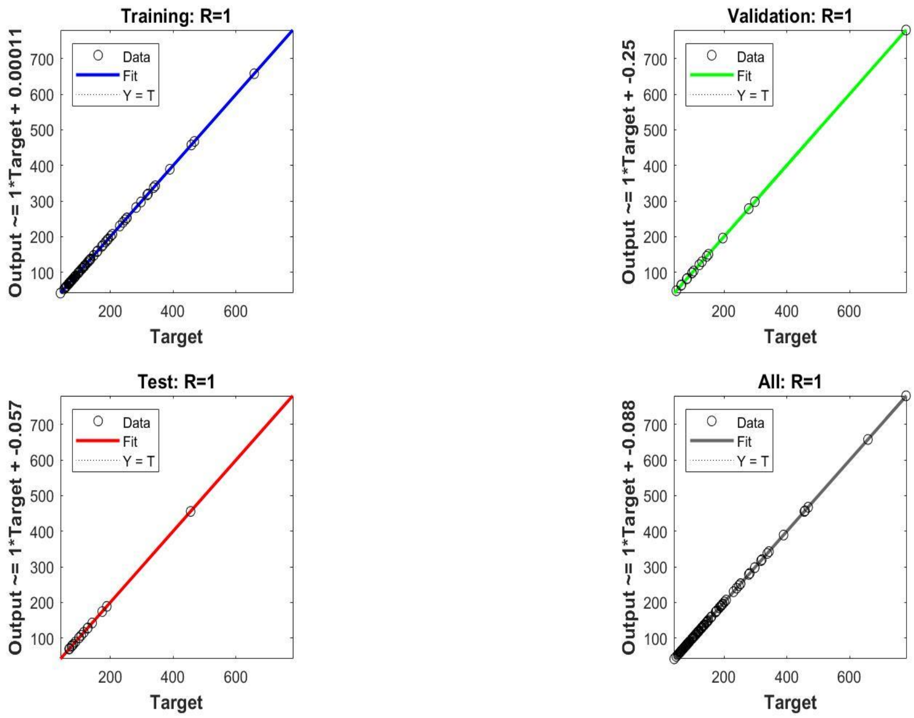

Figure 4.3: Regression plot of the ANN model to simulate WQI in present study

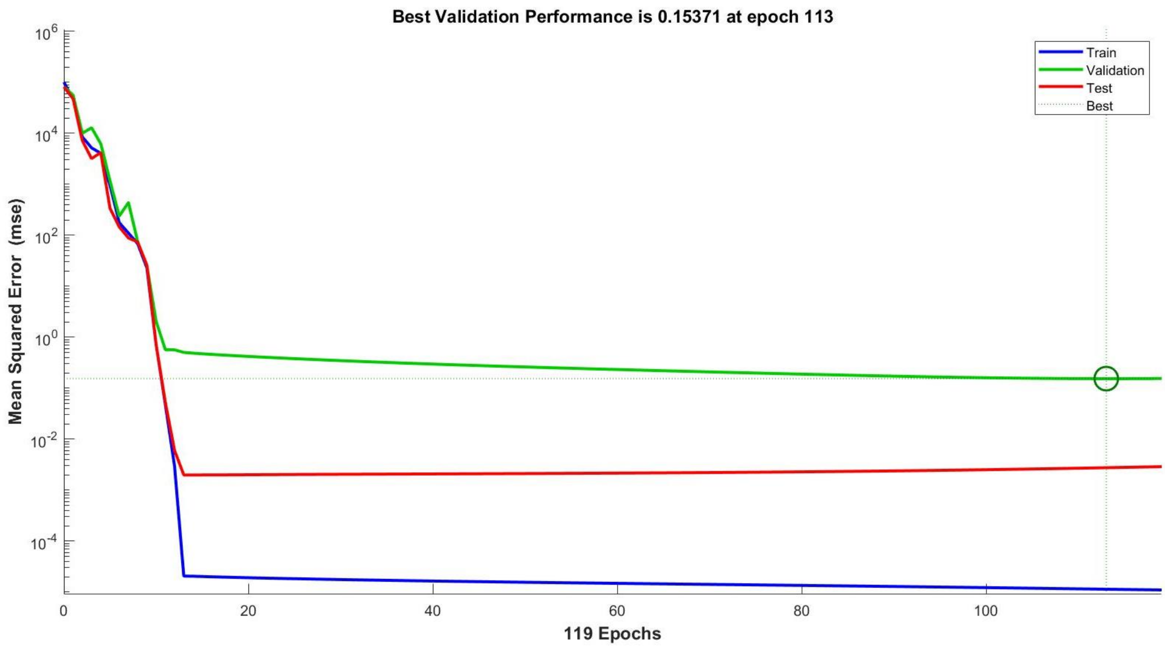

Figure 4.4: Performance of training phase WQ data utilizing the ANN model

Figure shows how the testing and training phases work. The regression plot was used to analyze the relationship between the simulated and actual WQI scores using Pearson's correlation. The output (y-axis) results are the simulated WQI scores produced using the ANN model, and the target (x-axis) scores are the observed WQ index data. The expected results are more in line with the desired results. The strong correlation between the observed and anticipated scores led to the construction of a successful model. The co-relation study demonstrated that the created model performed quite well.

By using an ANN model to mimic WQI and training on 80% of the entire data, the recommended model method was proven to be effective. The preliminary learning outcomes showed that the ANN model was quite good at imitating WQI. The prediction results for the WQI over the testing and learning stages. The simulated outcomes showed that R% = 100%, illustrating the system's extremely effective performance. The results of the ANN model's testing simulation suggested that R% = 100% based on the findings of the correlation regression analysis. The simulated results show the expected results during the testing stage. According to the estimate metrics (MSE and R), the simulated simulation scores were quite close to the experimental ones. The ANN model's regression curve for simulating WQI.

4.2 Piper diagram of stations for year 2015-2022:

Piper trilinear classification over groundwater chart software of US geological survey suggested that most areas of New Delhi and its adjoining places are affected by salinity hazard due to excessive presence of Na^+, Ca^+, Mg^+, Cl^- ions. Interpretation of surface and subsurface geological and hydrological data indicate that integration of lithological, geomorphic and tectonic factors has led to restricting the circulation of surface and sub-surface water in a graben like structure causing rise of water level. During summer, high evaporation causes capillary rise of shallow groundwater and subsequent precipitation of salts in soil.

The ground water availability in Delhi is controlled by the hydrogeological formations characterized by the presence of alluvial formation and quartzites hard rocks. The rock formation is widely varied with variation in land formation like ridge areas. It is traversing across the city and is quite significant to control the occurrence and movement of groundwater. In shallow aquifers, the groundwater occurs under phreatic confined condition. Contrarily, it is in semiconfined to confined conditions in deep aquifers. The shallow aquifers contain saline water and the depth varies from 5 to 10 m. Piper diagram was created for Delhi Region using analytical data obtained from groundwater quality parameters from Central Pollution Control Board. From the given data, locations have been extracted on basis of continuous availability of data and further helps in plotting the coordinates in Arc-GIS. From the piper plot from 2013-2019, from the piper plot, we classify the sample into 10 locations. 60% of sample shows the sodium chloride type. Rest 20% are of mixed type and further 20%are of magnesium bicarbonate types.

Evaluation from the piper plot suggested that there is clear indication of Na^{+} and Mg^{2+} species exceeded Ca^{2+} and Mg^{2+} as well there are various weak acids like CO_3^{2-} HCO_3^{-} exceeded the strong acids of Cl^{-} and SO_4^{2-}. In India Gate (28.61,77.23), the piper plot evaluates that from 2013 to 2016, it shows sodium chloride type then move towards calcium magnesium chloride type in 2017 to 2019. In Jagatpur (28.74,77.22), initially it shows magnesium bicarbonate type physiochemistry then from 2015 onwards, it shows calcium chloride type, further followed by calcium magnesium chloride type. In Majnu ka Tila (28.7,77.23), it shows calcium magnesium chloride mixed type in 2013, then slightly moves towards sodium chloride type of physiochemistry in later years of 2018 and 2019. In Mayur Vihar (28.62,77.3), piper plot suggests this area is dominated by sodium chloride from 2013 to 2019 except initially it shows the possibility of magnesium bicarbonate in 2013. In Nizamuddin (28.6,77.27), piper plot suggests that moving from 2013 it shows magnesium bicarbonate type but in later years, it shows fluctuations between sodium chloride and calcium sodium bicarbonate mixed type of physiochemistry. In Palla (28.86,77.2), piper plot suggested its shows sodium chloride type throughout the years of evaluation. In Peeragarhi (28.68,77.09), piper plot suggested that initial year of 2013 it shows calcium chloride type of physiochemistry and then moving towards the calcium magnesium mixed type of it in the later years of evaluation. In Asola (28.5,77.27), the piper plot assessed that in 2013,2014, 2015,2016, 2018, 2019; it shows Ca-Mg-Cl^{-} type except in 2017 in which it is more inclined towards $MgHCO_3^{-}$. In Baprola (28.65,77.02), the piper plot evaluate that it moved from NaCl type in 2013 to $MgHCO_3^{-}$ IN 2015 then again shows $NaCl^{-}$ type of physiochemistry. In CHILIA (28.59,77.3), the piper plot evaluates that initially it shows $MgHCO_3$ then moved to Ca-Cl^{-} type till 2018. But in later stages it shows Ca-Mg-Cl^{-} type of hydrochemistry.

Lithology is another geogenic factor that plays a dominant role in the hydro chemical groundwater process in Delhi. The lithology of the region is known to control the major-ion chemistry of the groundwater system (Kumar & Kumar, 2015).The weathering of carbonate concretions in parts of older alluvial plains, for example, is known to affect the evolution of hydro chemical facies of groundwater (Kumar & Kumar, 2015). Hydro chemical facies of groundwater in the Delhi region reveal them to be predominantly controlled by both changing geomorphology and anthropogenic influences. Facies variation in the alluvial plains clearly indicates that in moving laterally away from the Yamuna River, hydro chemical facies also evolve from bicarbonate-dominant facies in active floodplains into chloride-type facies in older alluvial plains. Hydro chemical facies in active floodplains are known to be bicarbonate in shallower aquifers (Sarkar and Shekhar, 2013). It has also been observed that general hydro chemical facies in the Delhi Ridge area and its elevated sandy surface were of Ca–Mg–HCO_3 type, which could account for continuous recharge in parts of these areas. The presence of high concentrations of sulfate and chloride ions in the groundwater system could be attributed to bed seepage from drains and wastewater (Subramanian and Saxena, 1983). In areas of low salinity, a Ca–Mg–HCO_3-type hydro chemical facies have been reported that is affected by local recharge. General assessment of the groundwater quality for all districts of Delhi reveals the presence of groundwater contaminants such as fluoride, nitrate, and lead in some localities. The groundwater quality of the western district seems to be more problematic than that in other districts of Delhi.

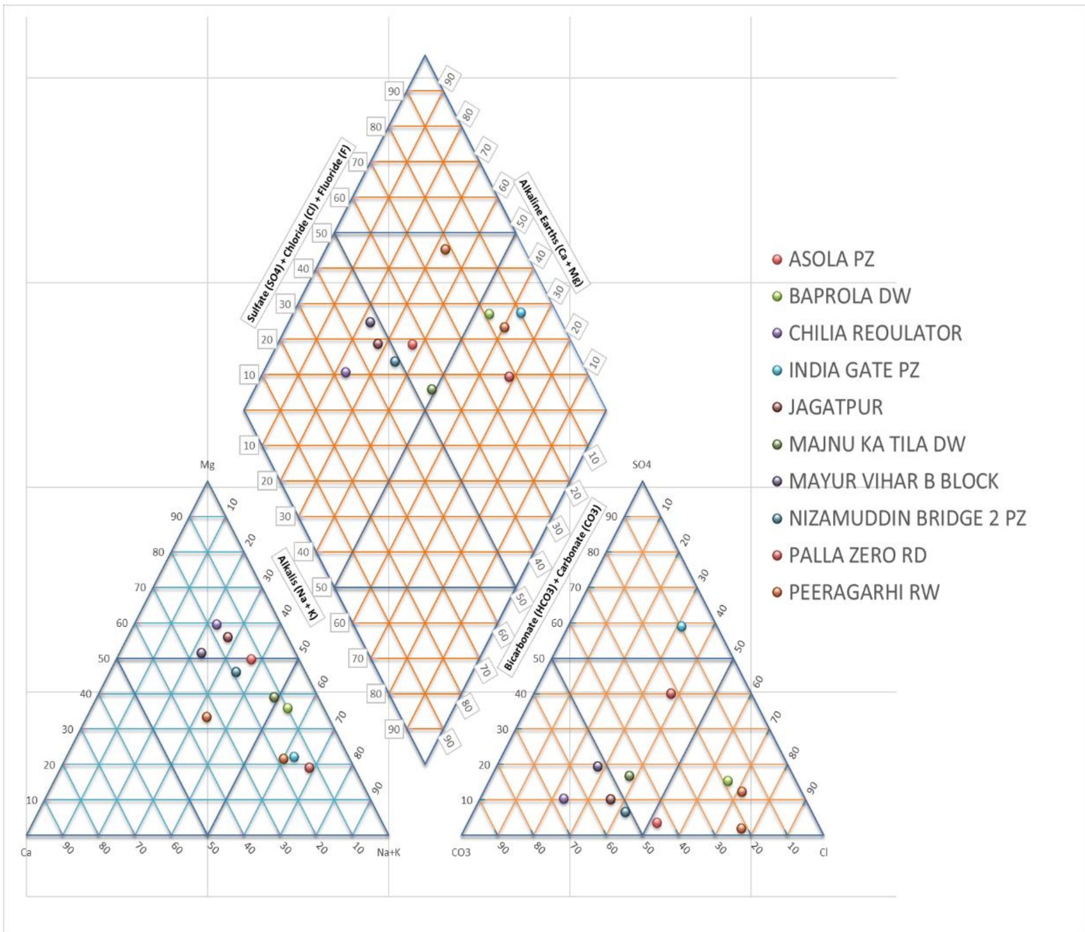

Figure 4.5: Piper diagram of water quality at 10 stations in current study for the year 2015

In 2015; Asola, Majnu ka Tila are in domain of mixed type while Baprola, India Gate, Peeragarhi and Palla road is in domain of Sodium Chloride (NaCl) and Chilia Regulator, Jagatpur, Mayur Vihar B Block and Nizamuddin is in domain of Magnesium Bicarbonate.

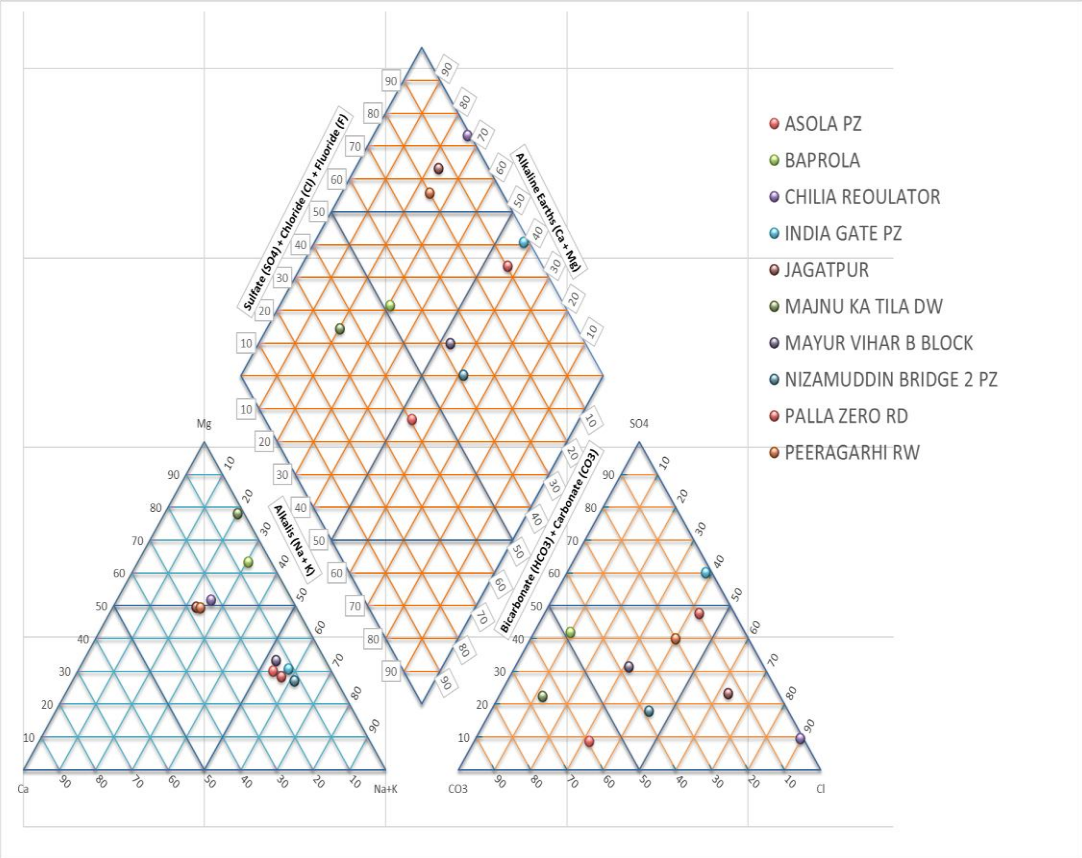

Figure 4.6: Piper diagram of water quality at 10 station for the year 2016

In 2016, Asola, Majnu ka Tila, Mayur Vihar B Block and Nizamuddin is in mixed type domain while India Gate PZ is in sodium chloride type domain. Baprola is in magnesium bicarbonate domain. India Gate and Jagatpur is in calcium chloride type domain.

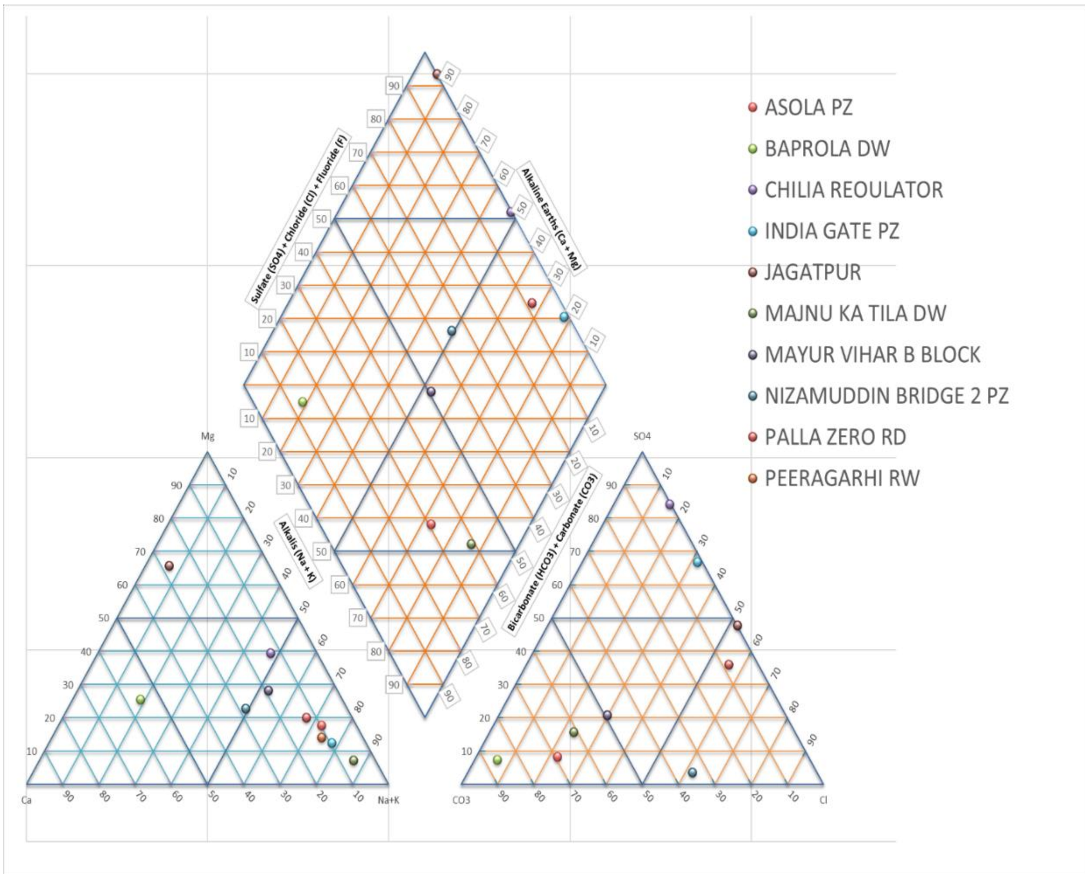

Figure 4.7: Piper diagram of water quality at 10 station for the year 2017

In 2017, India Gate and Peeragarhi lies in sodium chloride domain while Nizamuddin is in mixed type domain. Baprola lies in Magnesium Bicarbonate domain. Mayur Vihar, Palla zero RD, lies in mixed type domain.

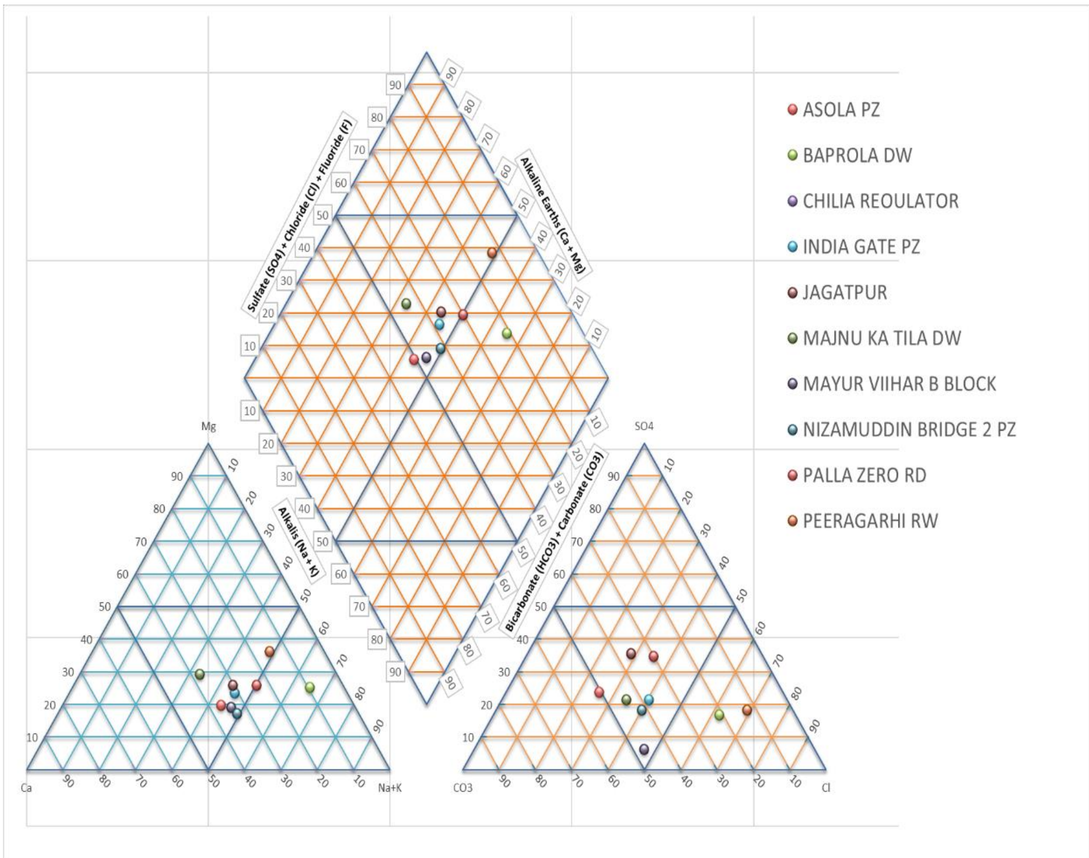

Figure 4.8: Piper diagram of water quality at 10 station for the year 2018

In 2018, majorly all stations except Baprola DW lies in mixed type domain whereas Baprola lies in sodium chloride domain. This may be due to there equal contribution of rainfall as well as seepage of polluted water in groundwater at all stations.

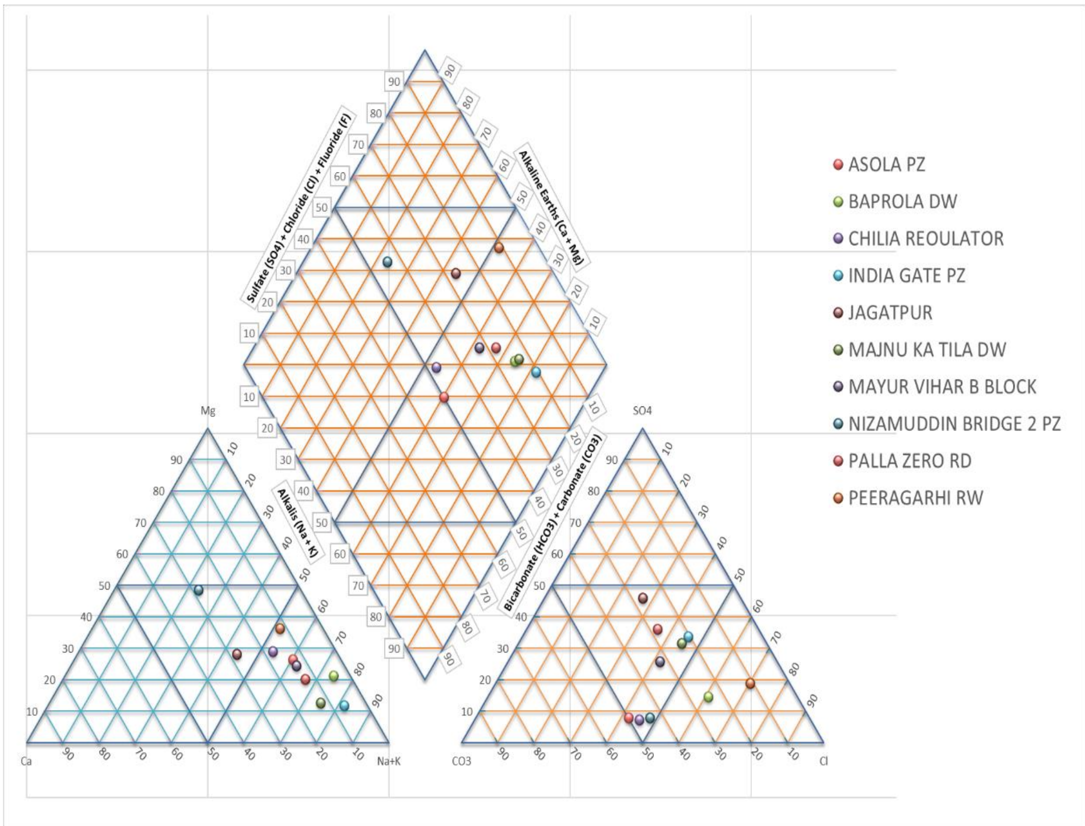

Figure 4.9: Piper diagram of water quality at 10 station for the year 2019

In 2019, except Nizamuddin Bridge PZ, Palla Zero RD, Peeragarhi RW all stations lies in mixed type domain while remaining stations lies in sodium chloride type domain.

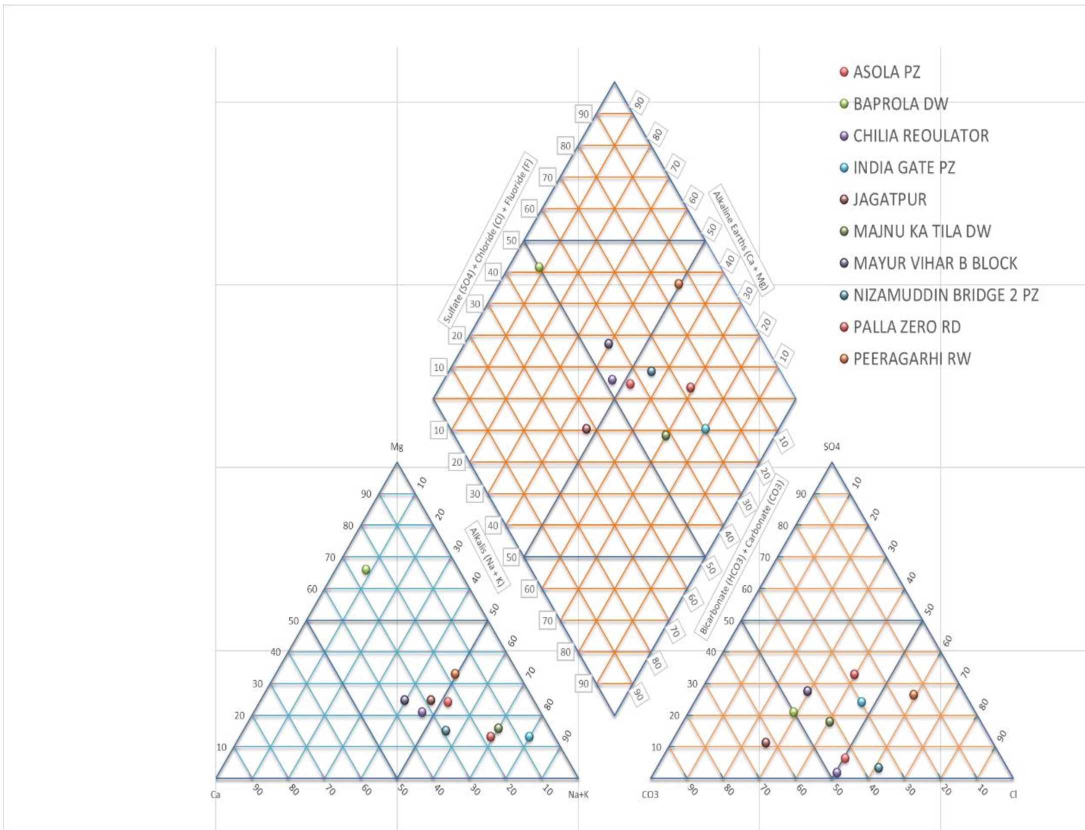

Figure 4.10: Piper diagram of water quality at 10 station for the year 2020

In 2020, India gate PZ, Majnu ka Tila, Palla Zero RD and Peeragarhi RW lies in sodium chloride domain while majnu ka tila lies in magnesium bicarbonate domain. However, remaining stations lies in mixed type domain.

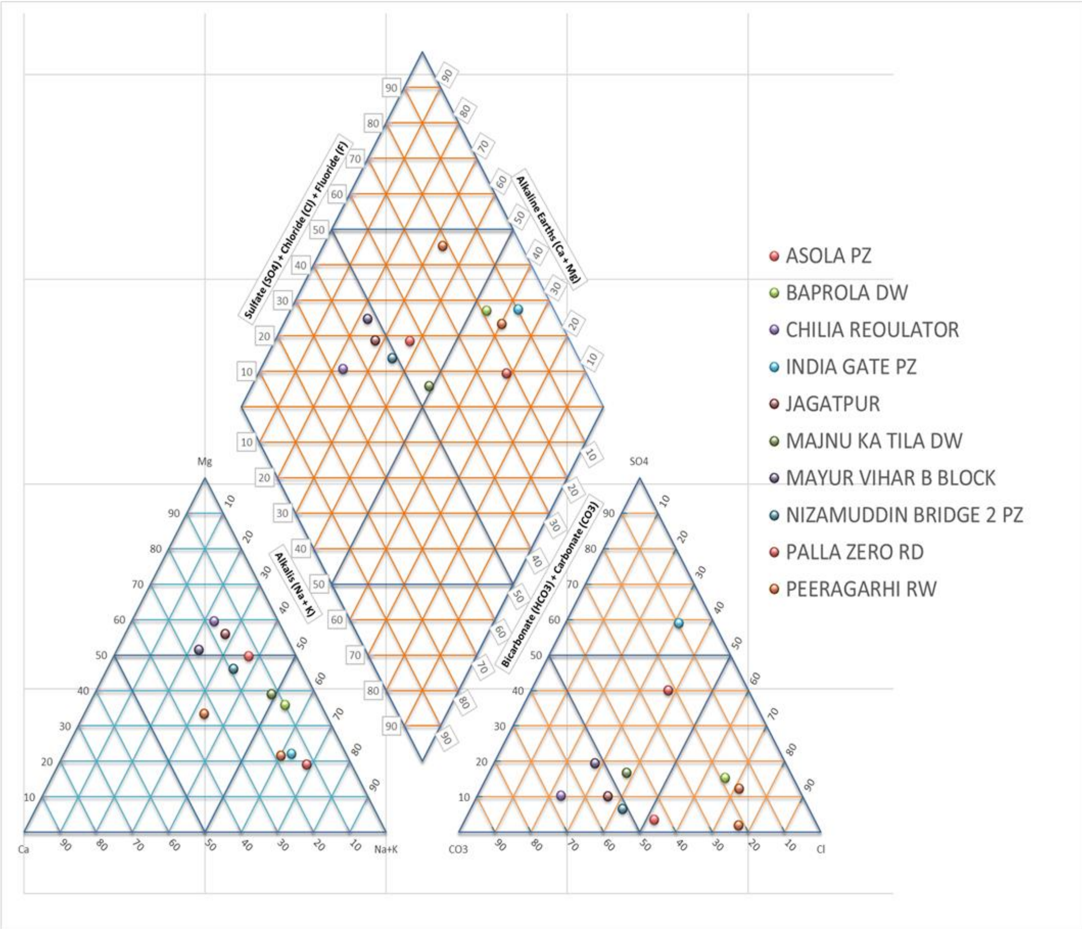

Figure 4.11: Piper diagram of water quality at 10 station for the year 2021

In 2021, India Gate PZ, Baprola DW, Palla road zero and Peeragarhi RW lies in sodium chloride domain while Majnu ka tila and Asola PZ lies in mixed type domain while Chilia Regulator, Mayur Vihar B Block and Nizamuddin bridge 2 PZ lies in magnesium bicarbonate region.

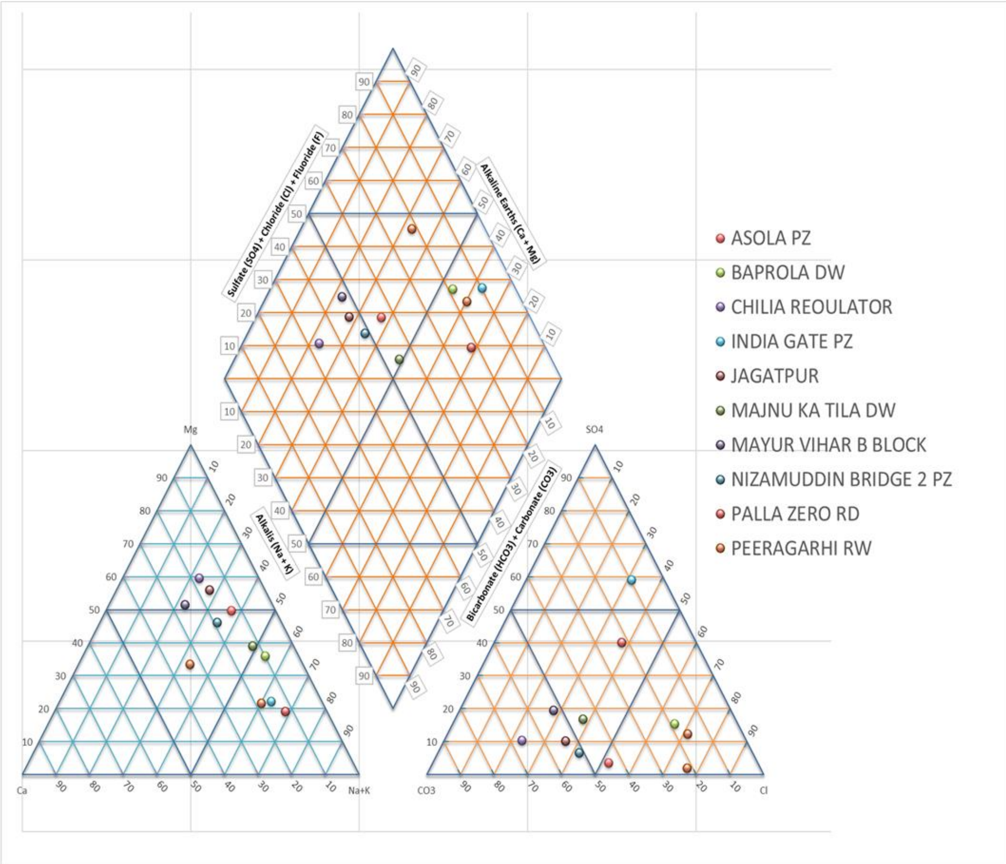

Figure 4.12: Piper diagram of water quality at 10 station for the year 2022

In 2022, Baprola lies in mixed type domain, while India Gate PZ, Majnu ka Tila , Peeragarhi RW lies in sodium chloride type domain .However, Jagatpur lies in magnesium bicarbonate domain and Majnu ka Tila and Chilia lies in mixed type domain.

4.3 Variation OF WQI by weighted arithmetic WQI and ANN model:

Yearly comparison i.e., from 2014 to 2022 of WQI of 13 stations by weighted arithmetic method and ANN has been shown below. From representation of comparison of WQI, it clearly shows that there is negligible difference between WQI outcoming from both tools. Hence our ANN approach is considered of best accuracy because of its results as compared with ANN. In year 2014 it was found that the water quality index score of station Jagatpur is least which is less than 50 hence, excellent WQ. While India Gate Pz station showed the maximum value of WQI of 389 which is higher than the 300 resulting in water unsatiable for drinking purposes. Stations having WQI score less than 100 were Asola, Chilia Regulator, Nizamuddin Bridge 2 Pz and Tiggipur.

For year 2015 it was found that there is the slight increase in the WQI of station India Gate, whereas Jagatpur showed no major difference in WQI scores. But the value of other stations showed reasonable fair increase in the WQI scores. In year 2016 it was found that there was an abrupt dip in WQI value of station India Gate Pz and rise in Peeragarhi DW. While the other all stations showed that WQI scores are less than 100 except for Palla zero RD, Rohini Sector 28 and Tiggipur. Out these three stations Rohini Sector 28 showed the maximum value of WQI of 250. It is found that in year 2017 the WQI of region Baprola showed an abrupt rise, causing in very bad ground water quality. The WQ at Peeragarhi DW station was some what improved but still it was more than 300 i.e., not fit for drinking. While all other stations showed the WQI lower than 100 except for Mayur Vihar B block, Palla Zero RD and PUSA (NRL). But no station showed the best WQ i.e., less than 50.

In 2018, there was a fair increase in overall WQI scores all over the Delhi region signifying the deterioration of WQ of all region of Delhi. Out of 13 stations only 3 stations showed WQI score less than 100. Peeragarhi DW showed the most polluted ground water00 having WQI score greater than 300 followed by Mayur Vihar B block having WQI score between 200 – 250. In 2019 the ground water quality was improved drastically, it was observed that, WQI of station Asola Pz, Chilla Regulator, Baprola DW, Jagatpur, Mayur Vihar B Block, Nizamuddin Bridge -2 and Tiggipur was observed to be less than 100. Whereas the WQI of Majnu Ka Tila Dw, Palla Zero RD and PUSA (NRL) was slightly more than 100. But the WQ of station Peeragarhi DW still remains same i.e., WQI > 300. In 2020 WQI of station Baprola DW increase drastically which is found to be more than 350. However, the WQI of other stations compared to 2019 was increased too. In 2020, the WQI of Baprola is close to 450 followed by peeragarhi DW of 350. In

2021, the maximum WQI is found to be of 550 followed by Rohini Sector 28. In case of 2022, WQI of peeragarhi DW is found to be maximum of 800 followed by India Gate PZ.

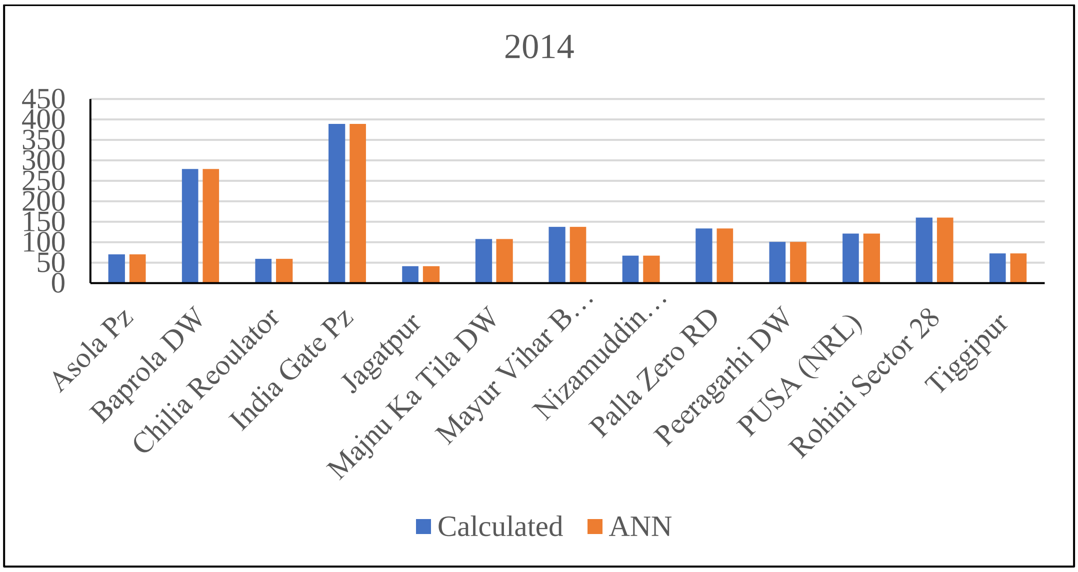

Figure 4.13: Comparison of Observed and simulated values of WQI in 2014

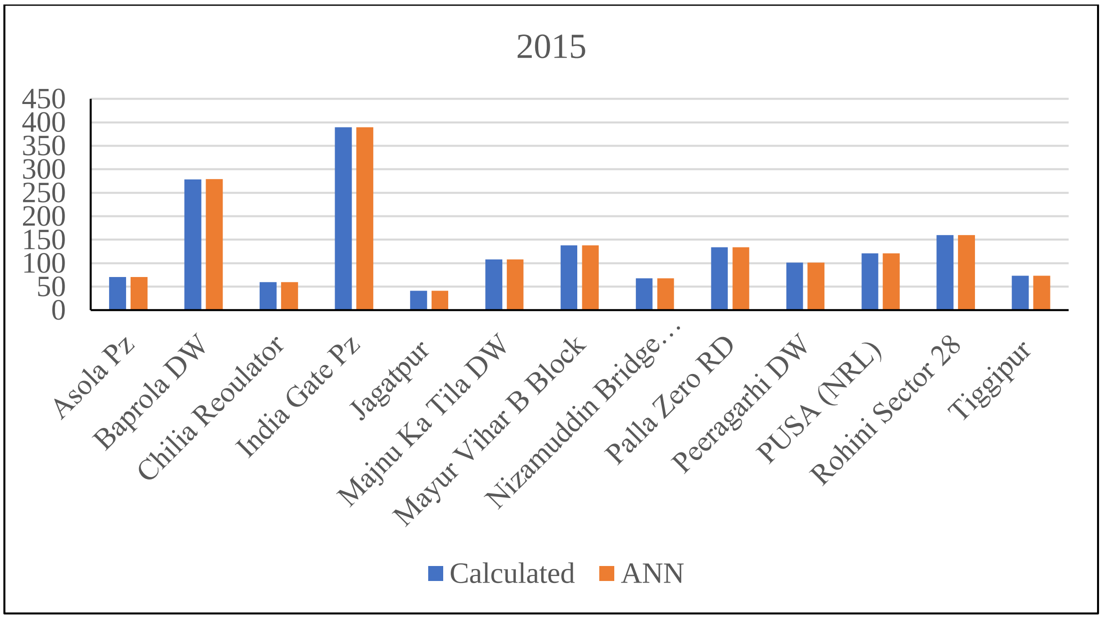

Figure 4.14: Comparison of observed and simulated values of WQI in 2015

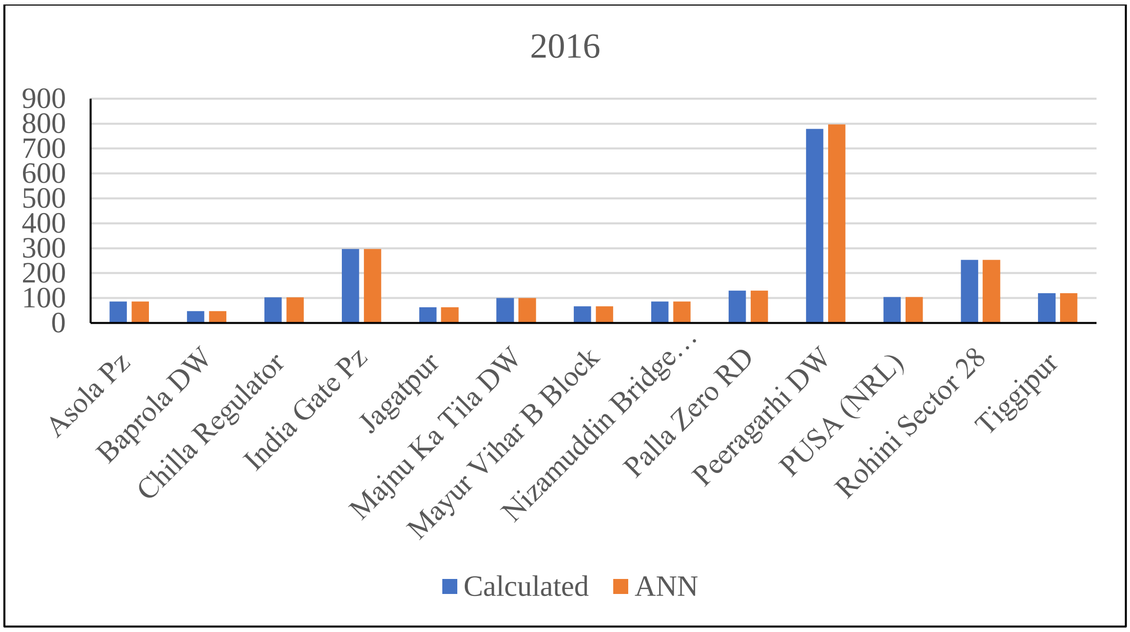

Figure 4.15: Comparison of Observed and simulated values of WQI in 2016

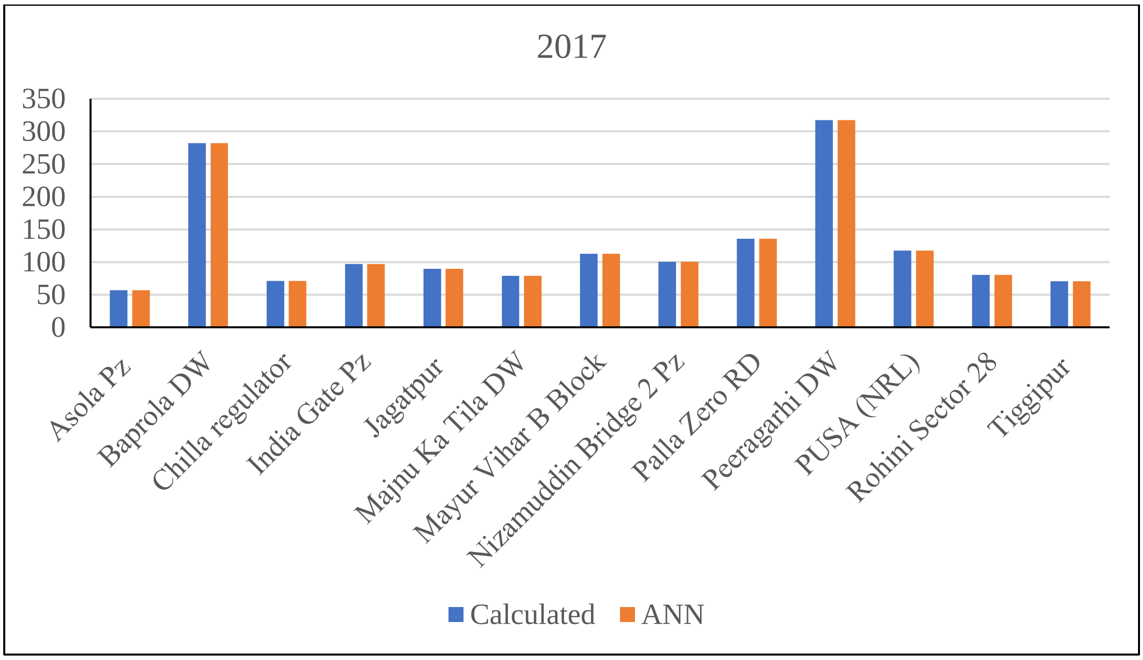

Figure 4.16: Comparison of Observed and simulated values of WQI in 2017

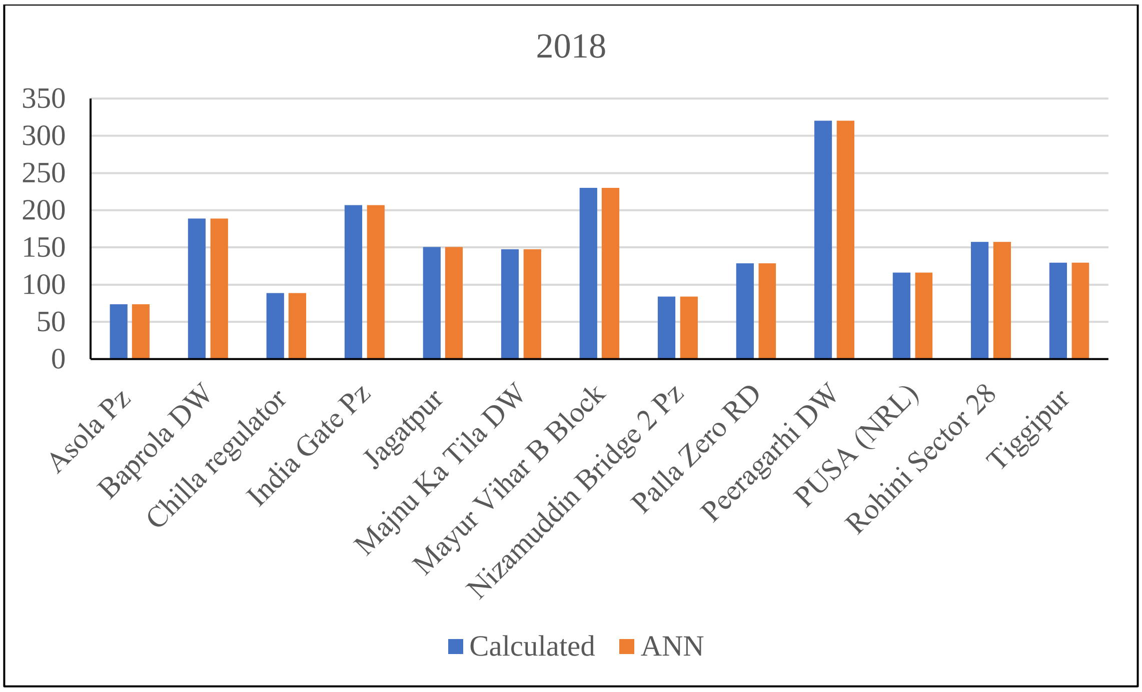

Figure 4.17: Comparison of Observed and simulated values of WQI in 2018

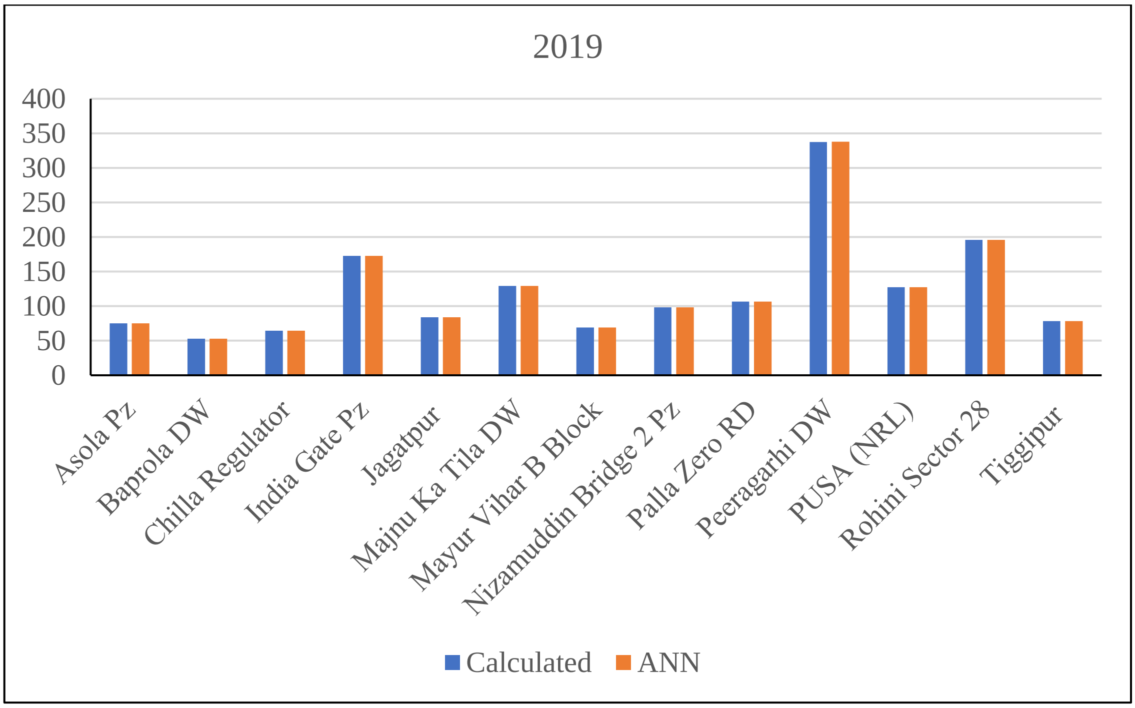

Figure 4.18: Comparison of Observed and simulated values of WQI in 2019

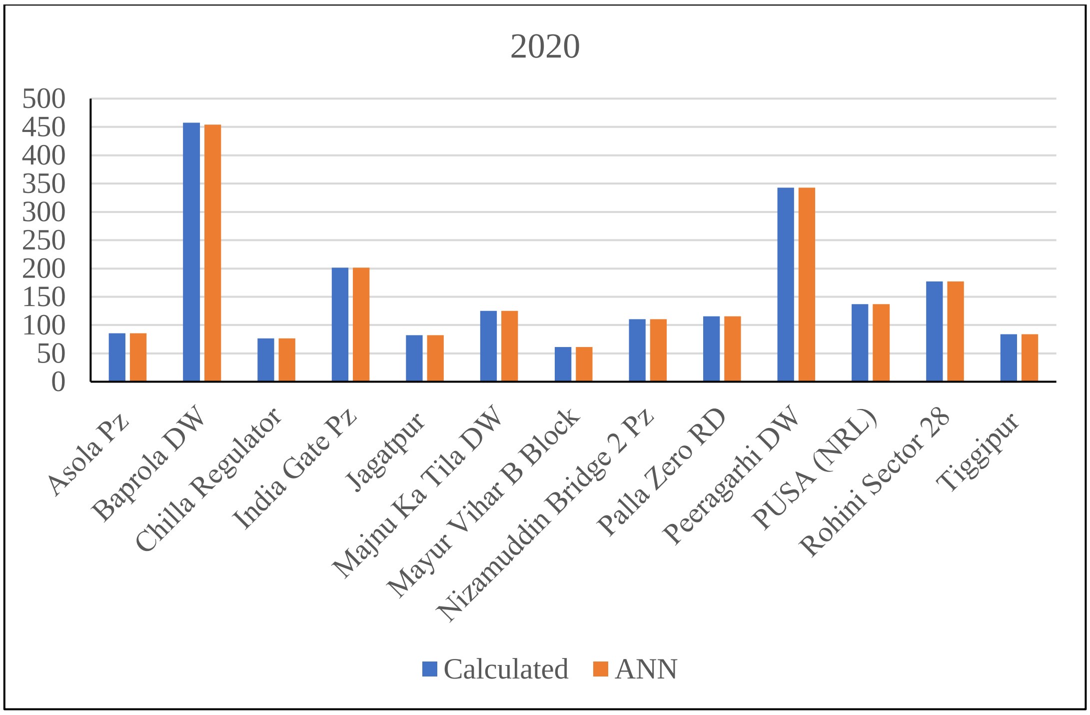

Figure 4.19: Comparison of Observed and simulated values of WQI in 2020

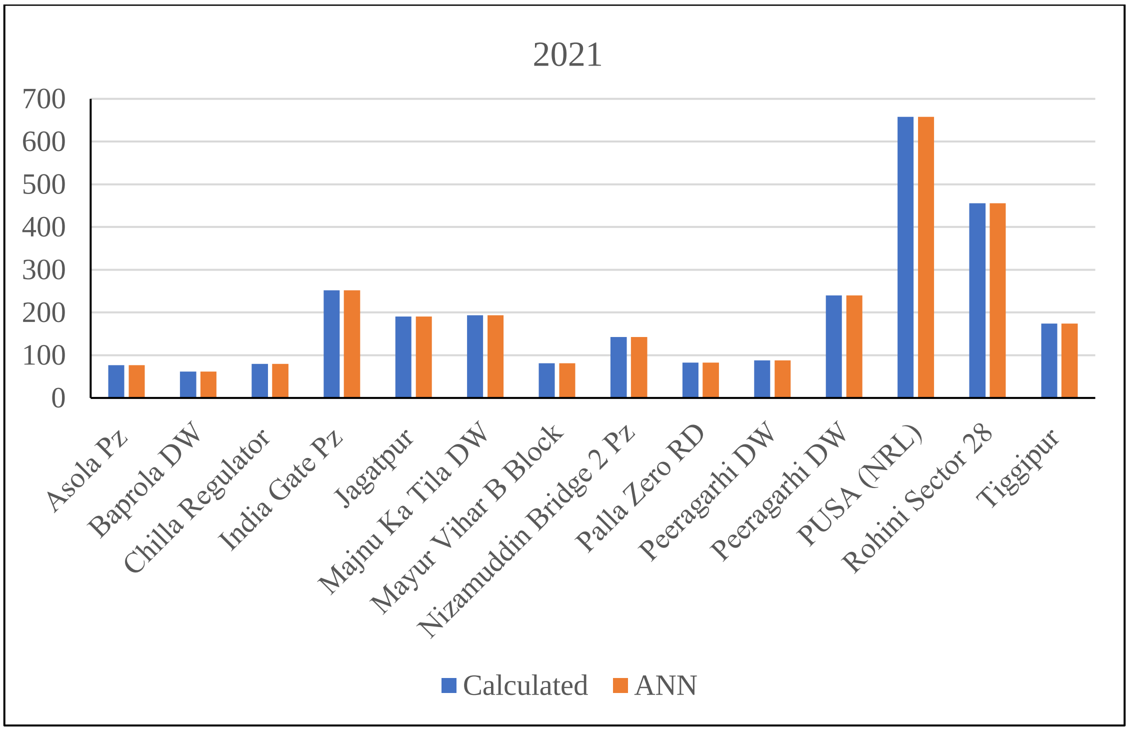

Figure 4.20: Comparison of Observed and simulated values of WQI in 2021

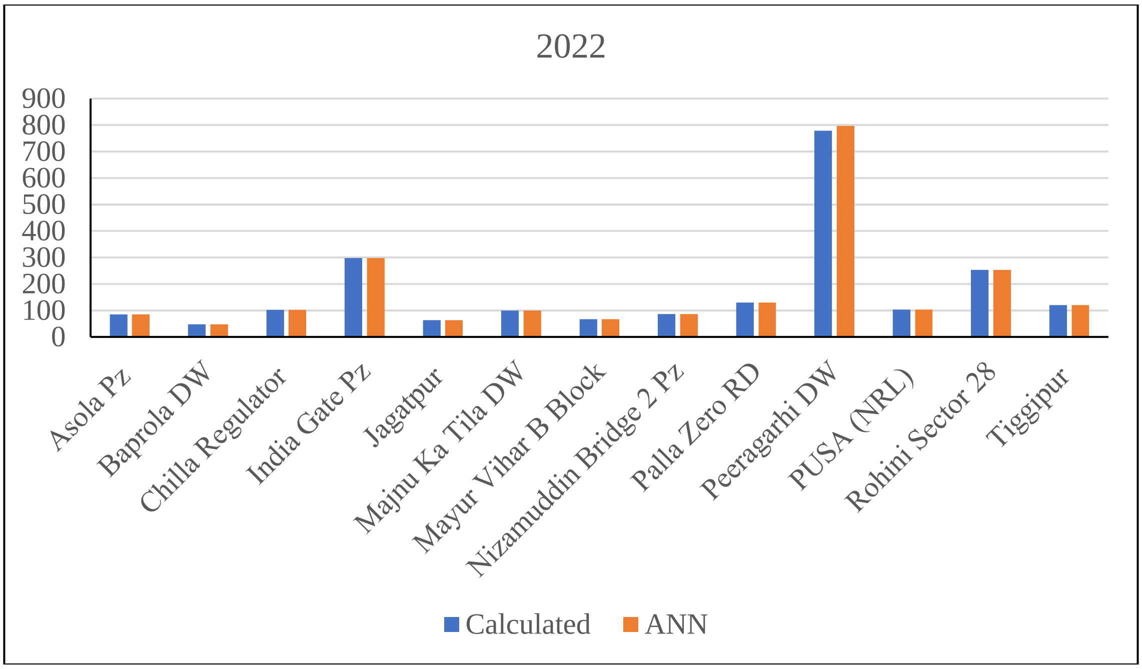

Figure 4.21: Comparison of Observed and simulated values of WQI in 2022

WQI REPRESENTATION BY IDW TOOL (INVERSE DISTANCE WEIGHTED):

2014

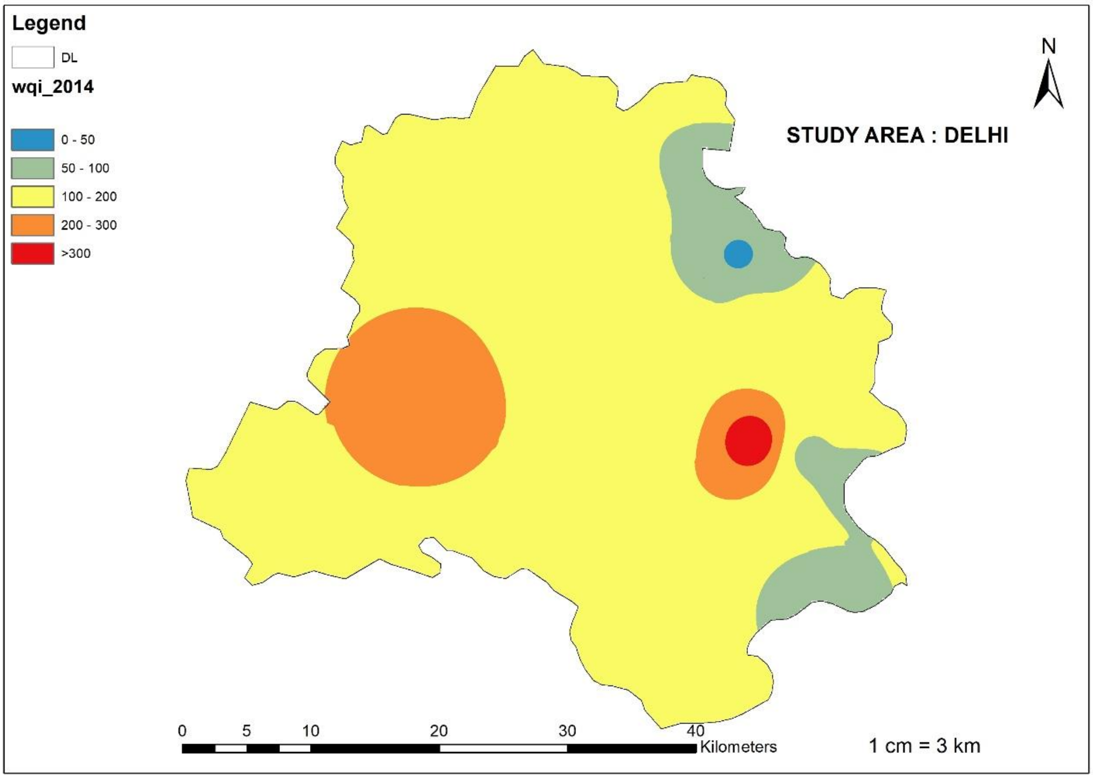

Figure 4.22: WQI representation of Delhi Region Of 2014

In 2014, WQI of Northern Delhi is worst i.e in the range of 450-650. Some parts of Western portion and South eastern portion is of good quality i.e in the range of 60 to 100. Eastern and western portion and some parts of middle Delhi is of poor quality i.e in the range of 150 to 200.

2015

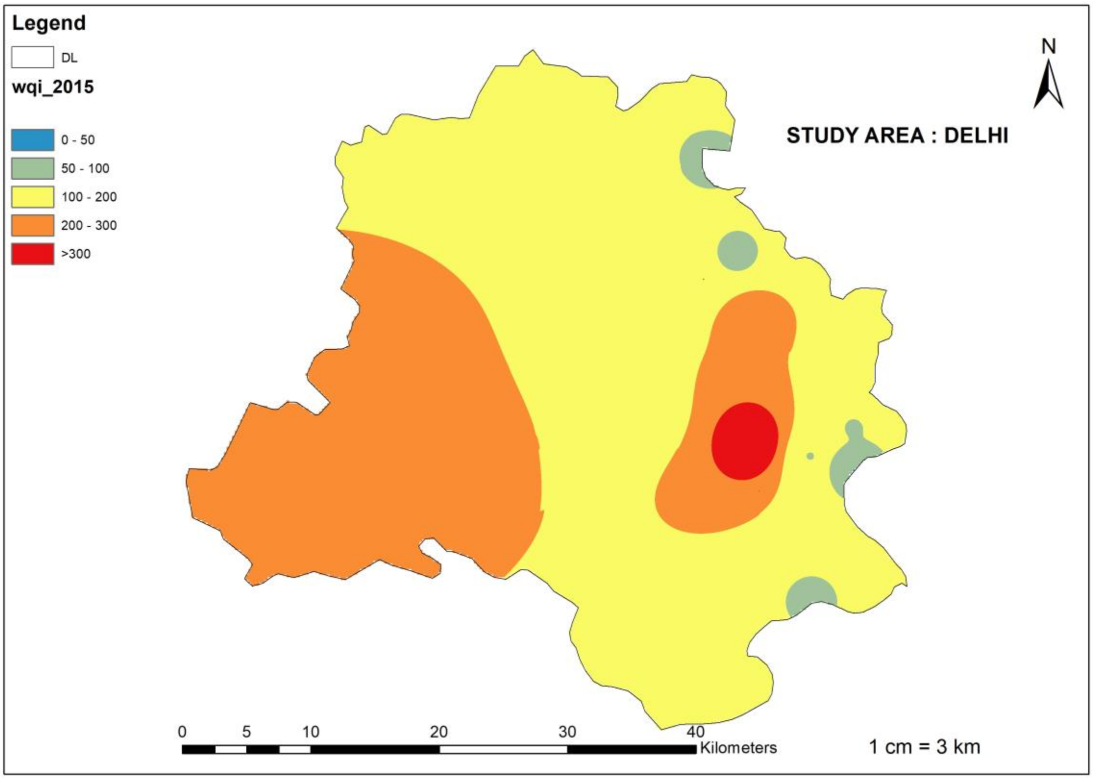

Figure 4.23: WQI representation of Delhi Region of 2015

In 2015, western Delhi is of poor water quality i.e in the range of 200 to 250 while northern, southern and some portion of eastern Delhi is of good quality i.e in the range of 70 to 150. Central Delhi has WQI ranges between 70 to 130.

2016

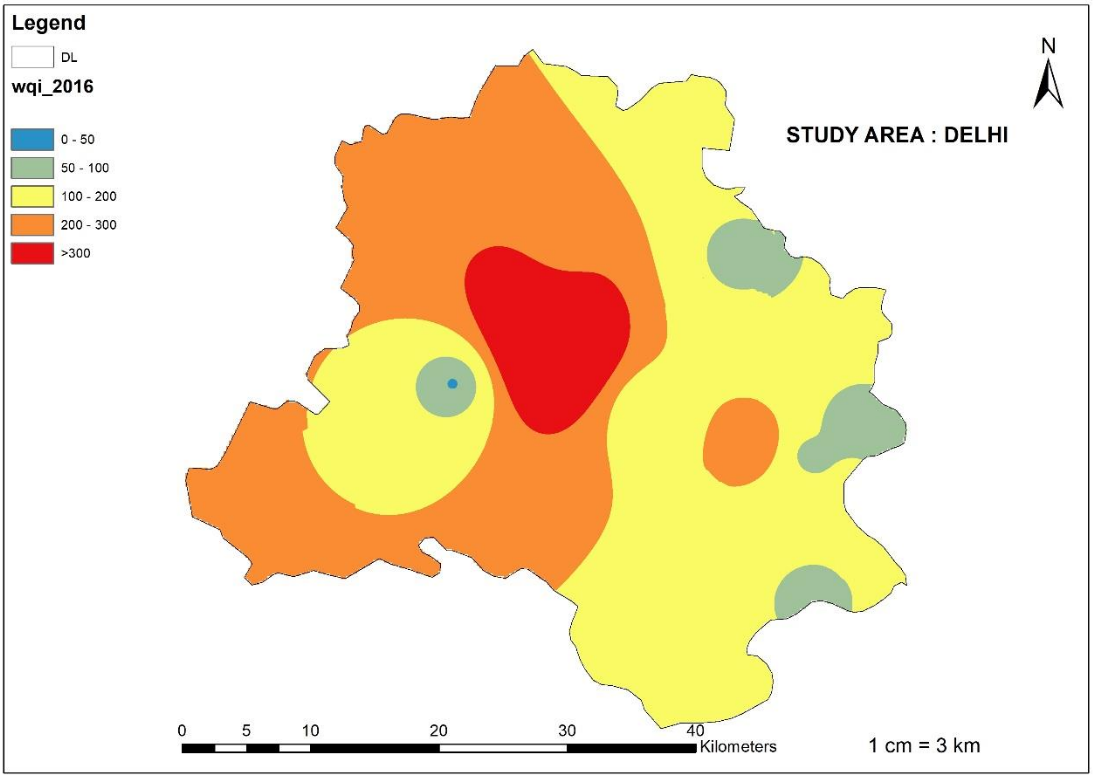

Figure 4.24: WQI Representation of Delhi Region of 2016

In 2016, the eastern and north eastern Delhi has WQI of 100 to 150 while western and middle portion of Delhi has poor water quality of 200 to 250. Northern portion of Delhi has WQI ranges between 110 to 120. Central Delhi has WQI ranges between 300-400.

2017

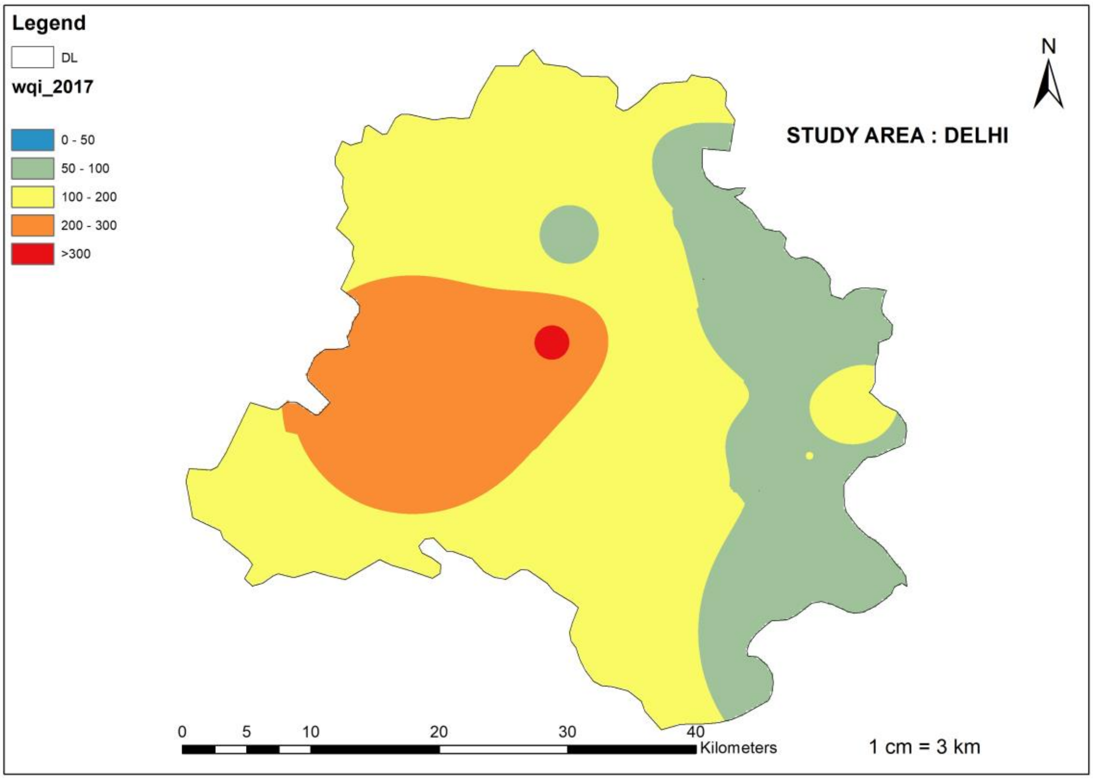

Figure 4.25: WQI Representation of Delhi Region of 2017

In 2017, western Delhi has WQI ranges from 180 to 320 i.e of severe quality while middle part has WQI of 110 to 130. However, the eastern and north eastern portion has WQI ranges between 60 to 80.

2018

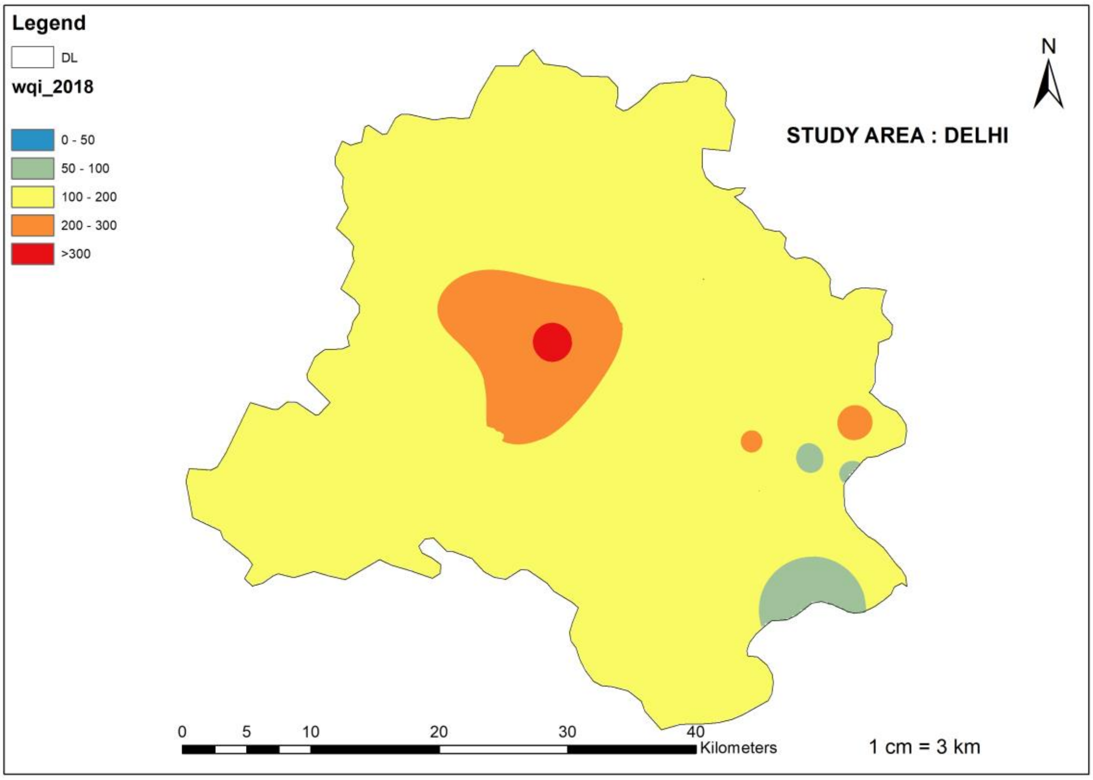

Figure 4.26: WQI Representation of Delhi Region of 2018

In 2018, the western portion has WQI ranges between 175 to 200 while eastern portion of Delhi has WQI in good range i.e between 75 to 120. Some part of northern and southern Delhi has WQI ranges between 75 to 120 i,e of good quality.

2019

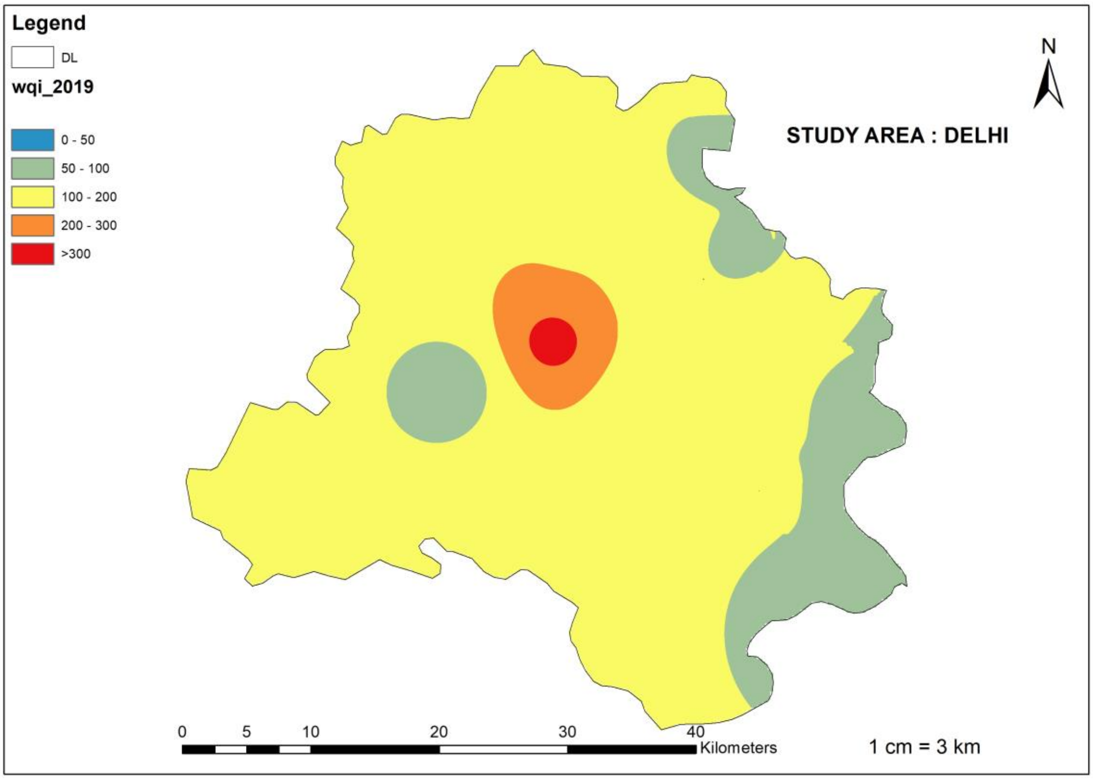

Figure 4.27: WQI Representation of Delhi Region of 2019

In 2019, some portion of western and eastern portion of Delhi ranges between 50 to 100. Although some west part ranges between 120 to 150 i.e of moderate quality. However, middle portion has severe quality ranges between 250 to 300. North eastern to south eastern portion of Delhi has WQI ranges between 50 to 100 i.e of good quality.

2020

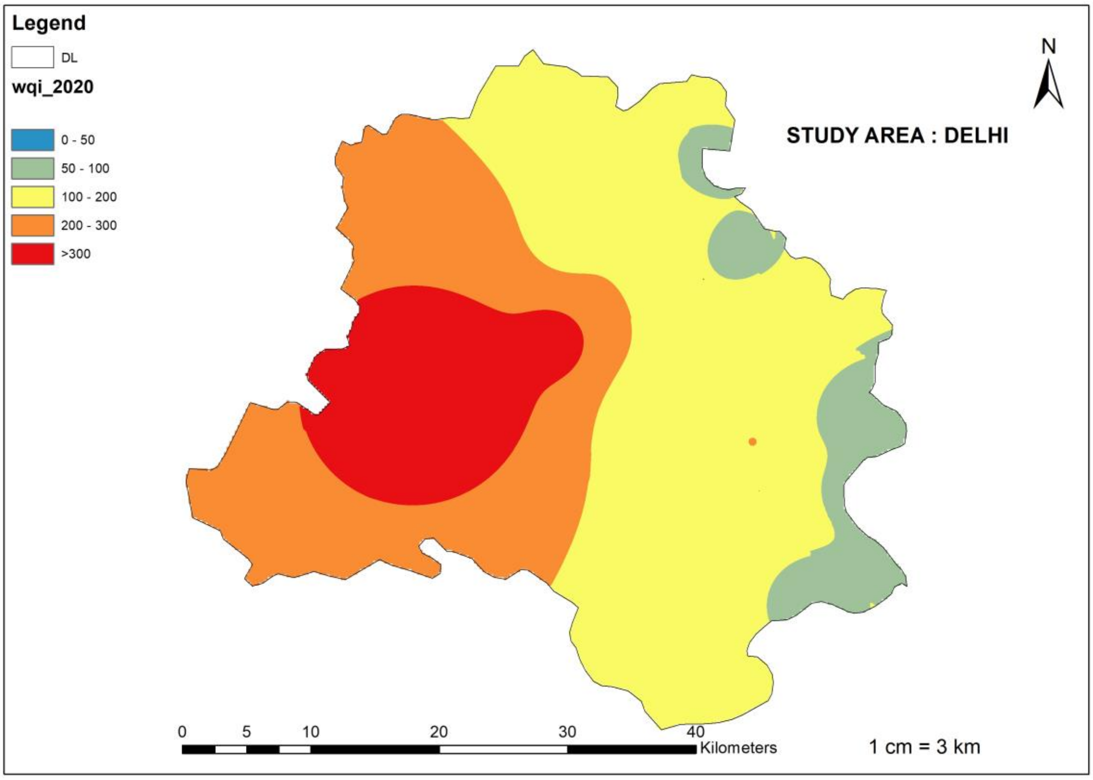

Figure 4.28: WQI Representation of Delhi Region of 2020

In 2020,the western portion of Delhi has severe WQI ranges between 350 to 450 while middle portion has WQI ranges between 175 to 200 while eastern portion has WQI ranges between 61 to 100.

2021

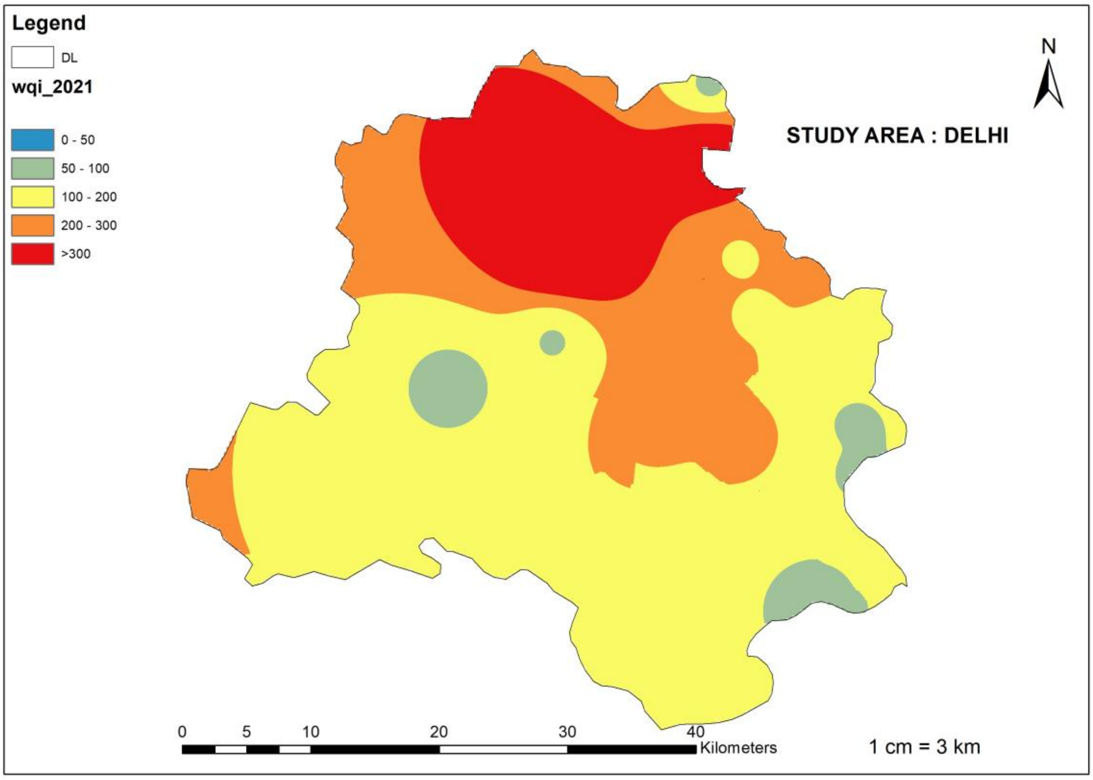

Figure 4.29: WQI representation of Delhi Region of 2021

In 2021, Northern Delhi has severe WQI of approx 450 while western and southern portion has WQI between 61 to 100. South eastern and some part of western Delhi has WQI ranges between 60 to 100. Eastern and western Delhi and some part of central Delhi has WQI ranges between 160 to 190.

CHAPTER 5
CONCLUSION

The mean WQI of India Gate is found to be maximum amongst all stations of Delhi Region. That shows there is drastic ground water pollution due to availability of various infusion of industrial contaminants directly into the ground water. Secondly a possibility of landfill leachate intrusion. Thirdly, hydrogeochemical properties of the soil present.

The percentage error between WQI that is assessed by arithmetic mean method and feed forward method of artificial neural network is very less which shows that our model is accurate. Hence it can be used for future prediction of WQI among the Delhi Region. For future prediction, various tools are available, in this study we used IDW of ARCGIS to get the values of WQI in the adjoining Region of the Delhi.

It explored universal issues associated with the construction of ANN models, with applicability to the Delhi ground water, and illustrated the usefulness of the ANN technique particularly in the aforementioned sector. From the results, it can be concluded that ANN performed very well with $R^2 = 1$ in modelling of WQI. It is concluded from outcomes, that ANN out-perform in modelling of WQI in comparison to conventional model. Furthermore, selected parameters are modelled utilizing the ANN tool.

The study's outcomes emphasizes that the ANN facilitates straightforward modelling of the WQI and enables for assessment of the relative relevance and effectiveness of inputs to model simulation. As a result, this study emphasizes that ANN is a useful technique for assessing groundwater WQ since it facilitates compute of the WQI and liberate significant work and time by optimizing the evaluation. Hence, method is widely applicable and may be used effectively to every aquatic system on the planet. As a result, empirical data analysis approaches such as ANN are recommended for long-term environmental surveillance record analysis.

Future Scope:

Although an exhaustive attempt has been made in this study to address groundwater water quality and its WQI, there are some murky areas that warrant further investigation. One of them might be by engaging additional WQPs such as heavy metals, coliforms, silica oxide, and so on, the second by incorporating more territories and states, and the third by examining all WQPs using conventional methodologies. Also attempts can also made to predict future WQI by using ANN. Future prediction can be possible by using a large data sample in current study. WQI prediction by ANN can also be validated by using different methods of WQI such as NSF method of WQI which depicts the accuracy of ANN results.

REFERENCES:

1. Aenab, A. M. (2013). *Evaluating Water Quality of Ganga River Within Uttar Pradesh State by Water Quality Index Analysis Using C ++ Program*. *3*(1).
2. Aenab, A. M., Singh, S. K., & Al-rubaye, A. A. M. (2012). *Evaluation of Tigris River by Water Quality Index Analysis Using C ++ Program*. *2012*(July), 523–527.
3. Aenab, A. M., Water, S., & Singh, S. K. (2015). *Critical Assessment of River Water Quality and Wastewater Treatment Plant (WWTP) Critical Assessment of River Water Quality and Wastewater Treatment Plant (WWTP). January*. https://doi.org/10.9790/2402-08244448
4. Ali, S. K. S., & Noori, R. (2022). Groundwater quality assessment and modeling utilizing water quality index and GIS in Kabul Basin , Afghanistan. *Environmental Monitoring and Assessment*. https://doi.org/10.1007/s10661-022-10340-0
5. Beek, V. (2015). *Chapter – 1 introduction 1.1*.
6. Chang, Y., Chang, L., & Chang, F. (2005). *Intelligent control for modeling of real-time reservoir operation , part II : artificial neural network with operating rule curves*. *1444*(December 2004), 1431–1444. https://doi.org/10.1002/hyp.5582
7. Chatterjee, R., Gupta, B. K., Mohiddin, S. K., Singh, P. N., Shekhar, S., & Purohit, R. (2009). Dynamic groundwater resources of national capital territory, Delhi: Assessment, development and management options. *Environmental Earth Sciences*, *59*(3), 669–686. https://doi.org/10.1007/s12665-009-0064-y
8. Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, *309*(1–4), 229–240. https://doi.org/10.1016/j.jhydrol.2004.12.001
9. Dawson, C. W., & Wilby, R. L. (2014). *Progress in Physical Geography Hydrological modelling using*. https://doi.org/10.1177/030913330102500104
10. Engineering, S., Hom, H., & Kong, H. (2006). *Neural network and genetic programming for modelling coastal algal blooms Nitin Muttil and Kwok-Wing Chau **. *28*, 223–238.
11. Goyal, D., Haritash, A. K., & Singh, S. K. (2021). A comprehensive review of groundwater vulnerability assessment using index-based , modelling , and coupling methods. *Journal of Environmental Management*, *296*(March), 113161. https://doi.org/10.1016/j.jenvman.2021.113161
12. Iqbal, M., Ali Naeem, U., Ahmad, A., Rehman, H. ur, Ghani, U., & Farid, T. (2020). Relating groundwater levels with meteorological parameters using ANN technique. *Measurement: Journal of the International Measurement Confederation*, *166*, 108163. https://doi.org/10.1016/j.measurement.2020.108163
13. Karamveer, A. K. H., Priyanka, M., & Singh, S. S. K. (2017). Hydrochemical characterization and suitability assessment of groundwater in Baga – Calangute stretch of Goa , India. *Environmental Earth Sciences*, *76*(9), 1–10. https://doi.org/10.1007/s12665-017-6679-5
14. Kazi, T. G. Ã., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A., & Shah, A. Q. (2009). *Ecotoxicology and Environmental Safety Assessment of water quality of polluted lake using multivariate statistical techniques : A case study*. *72*, 301–309. https://doi.org/10.1016/j.ecoenv.2008.02.024
15. Khalil, B., Ouarda, T. B. M. J., & St-Hilaire, A. (2011). Estimation of water

quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis. *Journal of Hydrology*, *405*(3–4), 277–287. https://doi.org/10.1016/j.jhydrol.2011.05.024

16. Khan, S., & Singh, S. K. (2019). *Assessment of the Impacts of Point Load on River Yamuna at Delhi Stretch , by DO-BOD Modeling of River , Using MATLAB Programming*. *December 2014*.
17. Kişi, Ö. (2006). Daily pan evaporation modelling using a neuro-fuzzy computing technique. *Journal of Hydrology*, *329*(3–4), 636–646. https://doi.org/10.1016/j.jhydrol.2006.03.015
18. Kumar, A., & Kumar, C. (2015). Characterization of Hydrogeochemical Processes and Fluoride Enrichment in Groundwater of South-Western Punjab. *Water Quality, Exposure and Health*, *2008*(October 2008), 373–387. https://doi.org/10.1007/s12403-015-0157-7
19. Lake, D. (2019). *Determining the Water Quality Index of an Urban Water Determining the Water Quality Index of an Urban Water Body*. *December 2014*.
20. Malik, V. S., Singh, R. K., & Singh, S. K. (2010). *IMPACT OF URBANIZATION ON GROUNDWATER OF GURGAON DISTRICT , IMPACT OF URBANIZATION ON GROUNDWATER OF GURGAON DISTRICT , HARYANA , INDIA*. *June 2014*.
21. Mohanty, S., Jha, M. K., Raul, S. K., Panda, R. K., & Sudheer, K. P. (2015). Using Artificial Neural Network Approach for Simultaneous Forecasting of Weekly Groundwater Levels at Multiple Sites. *Water Resources Management*, *29*(15), 5521–5532. https://doi.org/10.1007/s11269-015-1132-6
22. Munoz-arriola, F. (2020). Resilience, Response, and Risk in Water Systems: Shifting Management and Natural Forcings Paradigms. *Resilience, Response, and Risk in Water Systems*, *July*, 399. https://doi.org/10.1007/978-981-15-4668-6
23. Muthulakshmi, L., Ramu, A., Kannan, N., & Murugan, A. (2013). Application of correlation and regression analysis in assessing ground water quality. *International Journal of ChemTech Research*, *5*(1), 353–361.
24. Nehra, V. (2018). *ASSESSMENT OF WATER QUALITY OF NAJAFGARH DRAIN AND ITS IMPACT ON RIVER YAMUNA Vaishali Nehra DEPARTMENT OF ENVIRONMENTAL ENGINEERING*.
25. Noori, A. R., & Singh, S. K. (2021). Spatial and temporal trend analysis of groundwater levels and regional groundwater drought assessment of Kabul , Afghanistan. *Environmental Earth Sciences*, *80*(20), 1–16. https://doi.org/10.1007/s12665-021-10005-0
26. Noori, A., & Singh, S. (2021). Status of groundwater resource potential and its quality at Kabul, Afghanistan: a review. *Environmental Earth Sciences*, *80*. https://doi.org/10.1007/s12665-021-09954-3
27. Noori, A., & Singh, S. (2023). Assessment of seasonal groundwater quality variation employing GIS and statistical approaches in Kabul basin, Afghanistan. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-022-02876-5
28. Oladipo, J. O., Akinwumiju, A. S., Aboyeji, O. S., & Adelodun, A. A. (2021). Comparison between fuzzy logic and water quality index methods: A case of water quality assessment in Ikare community, Southwestern Nigeria. *Environmental Challenges*, *3*(January), 100038. https://doi.org/10.1016/j.envc.2021.100038
29. Patil, P. N., Sawant, D. V, & Deshmukh, R. N. (2015). Physico-chemical

parameters for testing of water, a review paper. *International Journal of Chemical Studies*, *3*(4), 24–28.
30. Raghav, A., & Singh, S. K. (2021). *Drinking water quality evaluation for groundwater of Delhi , India using GIS techniques*. *27*(1).
31. Rakhshandehroo, G. R., Vaghefi, M., & Aghbolaghi, M. A. (2012). Forecasting Groundwater Level in Shiraz Plain Using Artificial Neural Networks. *Arabian Journal for Science and Engineering*, *37*(7), 1871–1883. https://doi.org/10.1007/s13369-012-0291-5
32. Sahoo, M. M., Patra, K. C., & Khatua, K. K. (2015). Inference of Water Quality Index Using ANFIA and PCA. *Aquatic Procedia*, *4*(Icwrcoe), 1099–1106. https://doi.org/10.1016/j.aqpro.2015.02.139
33. Sarma, R., & Singh, S. (2023). Assessment of groundwater quality and human health risks of nitrate and fluoride contamination in a rapidly urbanizing region of India. *Environmental Science and Pollution Research*, *30*, 1–18. https://doi.org/10.1007/s11356-023-26204-0
34. Sarma, R., & Singh, S. K. (2021). *Simulating contaminant transport in unsaturated and saturated groundwater zones*. *2018*. https://doi.org/10.1002/wer.1555
35. Shan, V., & Haritash, S. K. S. A. K. (2021). Evaluation of water quality and potential metal contamination in ecologically important Bhindawas bird sanctuary , India. *Applied Water Science*, *11*(1), 1–9. https://doi.org/10.1007/s13201-020-01334-9
36. Singh, B., Sihag, P., Singh, V. P., Sepahvand, A., & Singh, K. (2021). Soft computing technique-based prediction of water quality index. *Water Supply*, *21*(8), 4015–4029. https://doi.org/10.2166/ws.2021.157
37. Singh, S. K. (2015). *Removal of chloride from ground water by Bio adsorption*. *May*.
38. Singh, S. K. (2019a). *Characterisation Study and Correlation Analysis For Water Quality of Dal Lake , India*. *April 2015*.
39. Singh, S. K. (2019b). *Limnological Study of Water Quality Parameters of Dal Lake , India*. *February 2014*. https://doi.org/10.15680/IJIRSET.2015.0402078
40. Singh, S. K. (2019c). *Modeling of arsenic transport in groundwater using MODFLOW : A case study Modeling of arsenic transport in groundwater using MODFLOW : A case study*. *October 2016*.
41. Singh, S. S., & Singh, S. K. (2019). *Evaluating Water Quality of River Yamuna in Delhi by Regression Analysis Evaluating Water Quality of River Yamuna in Delhi by Regression Analysis*. *June 2015*.
42. Tunc Dede, O., Telci, I. T., & Aral, M. M. (2013). The Use of Water Quality Index Models for the Evaluation of Surface Water Quality: A Case Study for Kirmir Basin, Ankara, Turkey. *Water Quality, Exposure and Health*, *5*(1), 41–56. https://doi.org/10.1007/s12403-013-0085-3
43. Wagh, V., Panaskar, D., Aamalawar, M., Lolage, Y., Mukate, S., & Adimall, N. (2018). Hydrochemical Characterisation and Groundwater Suitability for Drinking and Irrigation uses in Semiarid Region of Nashik, Maharashtra, India. *Hydrospatial Analysis*, *2*(1), 43–60. https://doi.org/10.21523/gcj3.18020104