SWELLING CHARACTERISATION AND RISK CLASSIFICATION OF BLACK COTTON SOIL USING EMPIRICAL CORRELATIONS AND MACHINE LEARNING

by Anurag Dayal Pandey

Submission date: 10-Jun-2026 12:09PM (UTC+0530)
Submission ID: 2980307708
File name: Anurag_for_plag.pdf (4.41M)
Word count: 12437
Character count: 64364

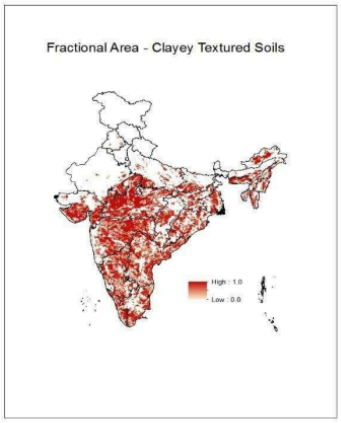
CHAPTER 1

INTRODUCTION

1.1 GENERAL

The black cotton soil (BCS) is one of the most troublesome geomaterials found in geotechnical practice because of its high swelling and shrinkage characteristics with moisture changes. These soils have significant swelling and shrinkage characteristics which cause repeated ground motion under foundations and pavement systems during wetting and drying. Cracking of low rise buildings, pavement distortion, differential settlement, canal lining distress and deformation of lightly loaded structures are among the cyclic volume changes that cause cracking in many parts of India. The engineering behaviour of expansive soils is thus still a significant issue for infrastructure development, especially in climatically variable regions.

Black cotton soils are mainly found in large areas of the Deccan Trap region of Madhya Pradesh, Maharashtra, Gujarat, Telangana, Andhra Pradesh, Karnataka and other neighbouring states in India. Typically formed from basaltic parent material, these soils have a high proportion of active clay minerals. They are darker than normal fine-grained soils, fine-grained, highly plastic and water-sensitive. The swelling behaviour of black cotton soil is related to the presence of expansive clay minerals like montmorillonite that have a high specific surface area and also have a high affinity for water absorption. When water enters the interlayer spaces of clay particles, significant volume increase of the soil can happen, particularly when the surcharge load is low.



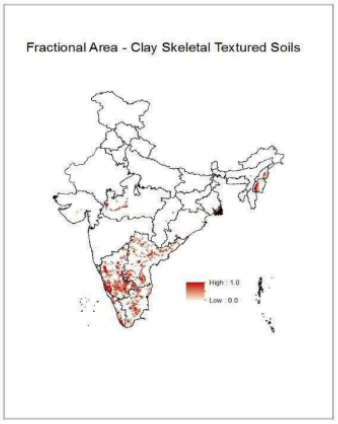


Fig 1.1 Distribution of Black Cotton Soil in India

Problem	Engineering Effect
Swelling	Foundation heave
Shrinkage	Surface cracking
Differential movement	Structural distress
Seasonal moisture variation	Pavement deformation

Table 1.1 – Expansive soils and Engineering problems related to them

Various interacting parameters such as mineralogical composition, initial moisture content, dry density, plasticity properties as well as environmental wettness–drying cycles, are responsible for the severity of swelling in expansive soils. Of these the index properties like liquid limit(w_L) and plasticity index (I_p), shrinkage limit (w_S) and free swell index are commonly used in engineering practice for preliminary identification of swelling susceptibility. Free Swelling Index (FSI) calculated in accordance with IS:9451, is one of the widely used

parameters in Indian context.

Volumetric changes occur in expansive earth materials with significant changes in moisture content. When water is added, water molecules enter into the inter-laminar space between clay minerals causing a significant increase in soil volume. On the other hand, when the material is dehydrated, it shrinks, and eventually forms deep surface cracks. This is the primary failure mechanism for lightly loaded infrastructure that is located over black cotton soil, and is likely to persist for the foreseeable future.

While the FSI test is relatively simple, a settling time of about 24 hours is required and special laboratory handling is necessary. During preliminary surveys or rapid infrastructure assessments, engineers often make use of indexing properties such as liquid limit as well as plasticity index, which are more readily available and can be conducted more easily and quickly. Therefore, there has been a growing interest in establishing reliable relationships between routinely measured soil properties and swelling behaviour to be able to assess the expansive soil risk more efficiently.

In the last 20 years there has been extensive research on empirical correlations for predicting swelling properties based on the Atterberg limits and other index properties. A number of existing correlations, however, are restricted to a specific region and some show high variability due to the different mineralogical and physico-chemical properties of expansive soils from different geological formations. We observed that in present times years, machine learning (ML) methods have also been applied to geotechnical engineering problems of classification and prediction of complex soil behaviour. These methods have the potential to be used to find non-linear relationships between soil properties and swelling risk, with more extensive data sets.

1.2 BACKGROUND OF EXPANSIVE SOILS

Despite these challenges in predicting volume change under different environmental conditions, identification and assessment of expansive soil behaviour have been important concerns in geotechnical engineering. The presence of expansive soils in many engineering projects, particularly those associated with shallow foundations, embankments, low-rise buildings and highways, often requires early assessment of swelling potential to prevent future structural problems. Black cotton soils are very sensitive to moisture changes, and small decline in soil moisture content of the season can show change in differential bulking and shrinkage cracking of the overlying infrastructure.

In traditional practice, the properties of swell of expansive soils are evaluated through laboratory parameters such as free swell index and swell pressure and potential of swell. Of these, the Free Swell Index test is popular in Indian geotechnical practice due to its simplicity and being included in the guidelines of IS:9451. The test is a practical test for the relative expansiveness of soil and is frequently used in preliminary site investigations. The FSI test is, however, less complex than the oedometer test, but does still need controlled laboratory conditions, careful handling, and a settling time of almost 24 hours before final readings can be obtained.

In normal engineering work, this type of testing can be cumbersome when conducting a quick assessment in the field or when conducting large-scale soil testing where several samples must be classified within a short time. However, index properties like plasticity index, liquid limit, optimum moisture content and maximum dry density are usually measured in nearly all geotechnical investigations due to their importance in soil classification and compaction studies. This has led several researchers to try to develop empirical correlations between these properties that are commonly measured and swelling behaviour, in order to be able to identify expansive soils faster than by relying on specialized swelling tests.

Machine learning (ML) methods have also been recently introduced in geotechnical engineering applications because they can capture complex and non-linear interactions between multiple soil parameters. Artificiallly Neurals Networks (ANN), Supports Vector Machines (SVM), and Random Forests (RF) have been used to address issues of soil classification, prediction of bearing capacity, settlement estimation, and expansive soil characterization. Machine learning models are generally used to find patterns in data without assuming that the variables are linearly related, unlike traditional regression methods.

In spite of these advancements, the combination of controlled laboratory experiments, empirical regression analysis and application of machine learning based swelling classification is still relatively limited in the Indian context of black cotton soils. The majority of existing studies is either empirical prediction based on small laboratory datasets or on standalone machine learning models trained from literature compiled datasets. It is thus necessary to develop an integrated framework that integrates the swelling behaviour generated from experiments with data-driven classification techniques to enhance the reliability and practical use of the swelling assessment methods.

The present study is selected with this objective. The black cotton soil of Shajapur was modified in a controlled manner using sodium bentonite, in order to systematically change the swelling nature of the soil in the laboratory. The experimental data obtained were then subjected to regression modelling to look for relationships between index properties and free swell index. Furthermore, machine learning classification approaches were investigated with an extended multi-regional database of Indian black cottons soils to meausre the potential of rapid prediction of the swelling ability of soils based on routinely measured soil properties. The swelling characteristics and identification methods of the soil are described. Swelling characteristics and identification methods of the soil are discussed.

1.3 SWELLING CHARACTERISTICS AND IDENTIFICATION METHODS

Expansive soils have very large volume expansion characteristics in response to difference in moisture content. When the soils are wet, water molecules get trapped between the clay particles, increasing the volume of the soil; when it is dry, the soil shrinks, leading to cracking in the soil. One of the major causes for distress in lightly loaded civil engineering buildings built on black cotton soils is this repeated swell–shrink phenomenon. Clay mineral composition, initial dry density, moisture variation, stress condition and physio-chemical properties of soil are some of the factors that determine the magnitude of the expansion.

The swelling ability of black cotton soil is usually linked with presence of active clay minerals with high water absorption capacity. The adsorbed water film around the clay particles becomes thicker when moisture is added to the soil which causes the particles to separate and the soil structure to expand. This can be worse in areas with alternating wet and dry climates as the soil is constantly subjected to expansion and contraction. Consequently, the following problems are frequently encountered in expansive soil areas: pavement heaving, cracking of the foundation, differential settlement, and surface deformations.

The identification of expansive soils in geotechnical engineering practice is usually done by direct and indirect methods. Direct methods are the methods of measuring actual swelling characteristics in the laboratory by means of experiments like free swelling index, swelling pressure, swelling potential, and consolidation based expansion test. These tests can be used to assess the swelling behaviour reliably, but may be conducted in controlled laboratory conditions, take longer to test and require careful sample preparation. Therefore, during preliminary investigations, indirect identification through the use of index properties that are routinely measured is often used.

Atterberg limits are among the numerous identification parameters that are used to assess the plasticity and expansiveness of fine particle soils. The liquid limit (LL) is the moistures of the soil at which it becomes liquid and the plastic limit (PL) is the water content at which it becomes plastic. The difference between

these two parameters is represented as the plasticity index which is sign of the moisture range in which the sample is plastic. In general, soils with high liquid limit and high plasticity index are more likely to have swelling properties due to their high water absorption and clay activity.

Another useful parameter in the characterization of expansive soil is the limit of shrinkage. It is the moisture at which later drying will not cause the soil to shrink any further. Soils with high shrink–swell potential are those with low shrinkage limits. Likewise, the compaction properties of max dry density (MDD) and optimum moisture content (OMC) also affect the engineering response of expansive soils, especially in the pavement. Subgrades and embankment applications.

The Free Swell Index (FSI) test most widely used direct identification tests used in the field of soil investigations in India. The test is carried out as per IS:9451 by comparing the sediment quantity of oven dried soil which pass 425micron sieves in water and kerosene. The quanity of sediment increases in the water as a result of kerosene with no swelling potential, so the volume of sediment in water is a good indicator of the swelling potential of the soil. Expansive soils can be classified into various swelling severity classes, from low to very high, based on FSI values.

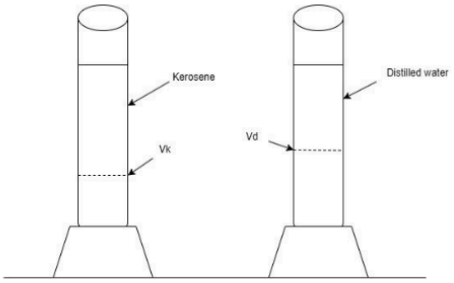


Fig 1.2 Typical Free Swell Index Test Arrangement

It has been found that several researchers have reported that meaningful correlation exists between the swelling behaviour and indexing properties such as liquid limit, Toughness Index, shrinkage limit and compaction criteria. The reliability of such correlations, however, will depend on the characteristics of the soils, the mineralogical composition and on the quality of the experimental data employed in the model development.

Over the past few years, data-driven methods have also been investigated in the identification or classification of expansive soils. Machine learning methods can generally be used to capture the inter-correlations between several soil parameters and can help to predict the swelling categories rapidly using the commonly measured laboratory data. These methods can be especially helpful in preliminary site investigations where the quick assessment of expansive soil behaviour is needed prior to detailed laboratory testing.

1.4 NEED FOR PREDICTIVE MODELLING OF SWELLING BEHAVIOUR:

An accurate estimation of swelling is a key consideration in the design or construction of structures built on expansive soils. Generally the potential for swelling of the subsoil is a major factor in many geotechnical projects in determining the type of foundation to be used, the thickness of the pavement, drainage decisions and soil improvement. The direct determination of the swelling parameters by laboratory tests, however, can be time consuming and sometimes is not possible during an early phase of site investigation where many soil samples are involved.

The Free Swelling Index (FSI) tests is widely used in India for identification notably of expansive soils as it is simple as compared to consolidation based swelling tests. However, the process must be performed under controlled laboratory conditions and a settling time must be allowed before measurements can be taken. Rapid assessment techniques are frequently used in practical field investigations, particularly for highway alignment studies, rural infrastructure development and preliminary geotechnical surveys to identify potentially problematic zones of soil at an early stage.

In response to this, the use of predictive modelling has become more and more important for research on expansive soils, and relies on soil properties that are routinely measured. Liquid limit, plasticity index (I_P), optimum moisture content and maximum drying density are among the parameters that are usually measured during routine laboratory testing and are reported to have significant effect on the swelling characteristics. It is therefore hoped that these relationships can be used as a guide to give an initial estimate of soil expansiveness, without relying exclusively on specific swelling tests, to assist engineers in their design.

Many researchers have adopted conventional empirical correlations for prediction of swelling characteristics from index properties. The regression-based models are relatively easy to develop and give engineering relationships between the dependent and independent variables. These models work well when there are consistent behavioural trends in the soil dataset for controlled laboratory conditions. Empirical equations developed for one soil type or geographical region, however, do not necessarily give satisfactory prediction accuracy for soils from other geological environments. The restriction is due to the interaction of several factors which affect expansive soil behaviour, such as mineral composition, clay activity, density condition and environmental moisture variation.

In the last few years, computational and data-driven techniques have proved to be useful for tackling complex geotechnical prediction problems. Machine learning methods can detect non-linear relationships between several variables in large data sets without making any mathematical assumptions. Random Forests (RF), Supporting Vectors Machine (SVM), and Artificially Neural Network (ANN) have been found to have potential applications for numerous geotechnical problems such as soil classification, prediction of bearing capacity, estimation of settlements, and assessment of slope stability.

1.5 RESEARCH GAP

The swelling properties of expansive soils have been reported in various studies in the literature, both through the use of conventional laboratory techniques and predictive modelling methods. Empirical relationships have been developed between swelling parameters and index properties to estimate expansive behaviour without carrying out complex swelling tests. Likewise, the recent advances of geotechnical data analysis have paved the path for the use of machine learning prediction and classification ways to laboratory data.

While a significant amount of work has been done in this area, some studies are still limited. Numerous empirical correlations have been published for expansive soils but are largely limited to a specific geographic area and may have been developed from natural deposits of soils with wide variations in mineralogical composition and environmental conditions. Therefore, prediction equations for one geographical area might not always yield consistent results when applied to soils in other areas. Furthermore, there are several published correlations that are based on small data sets and only consider individual swelling parameters without incorporating other trends in the behaviour.

In recent years, studies based on machine learning have also grown in number, but most existing studies are based on literature data bases without relating the models to controlled experimental variation in the volume change behavior of expansive soils. The correlation of changes in index properties and the swelling response is not studied systematically in the laboratory in many cases. In addition, there is a limited number of reports on the integration of regression based empirical modelling with machine learning based swelling classification of Indian black cotton soils. The present research work aims to fill these gaps by conducting a controlled laboratory investigation on black cotton soil of Shajapur district of Madhya Pradesh.

Various amounts of sodium bentonite were added to the control soil to artificially create different levels of expansion. Later, predictive equations were developed and tested for the similar nature between the standard main properties and the Free Swell Index and found to be in decent agreement with the laboratory results. Furthermore, the study used a large database of regional expansive soils to train and test machine learning classifiers to evaluate their suitability.

Table 1.2 — Summary of Research Gaps Identified from Previous Studies

Previous Research Focus	**Identified Limitation**
Conventional swelling tests	Time-consuming
Empirical correlations	Region-specific applicability
ML-based prediction	Limited linkage with controlled experiments
Standalone approaches	Lack of integrated framework

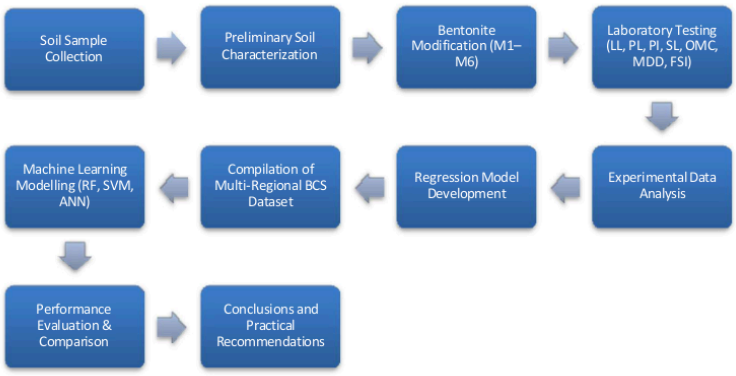


Fig 1.3 Research Framework

CHAPTER 2

LITERATURE REVIEW

Sridharan, A. and Prakash, K. (2000) carried out to study the similar nature between Atterberg limits and swell properties of expansive soils of India. From their observations, they concluded that the plasticity index is a major factor in the swelling potential of fine-grained soils, and there was a reasonably consistent correlation between Free Swell Index (FSI) and PI. The work emphasized the value of index properties in making a preliminary evaluation of expansive soil behavior without depending on comprehensive swelling tests.

Holtz, R. D. and Kovacs, W. D. (1981), in their geotechnical and soil engineering text, explained the fundamental mechanisms governing swelling behaviour in expansive soils. The authors discussed the influence of clay mineralogy, moisture variation, and density condition on soil expansion characteristics. Their work established that soils containing montmorillonite clay minerals generally exhibit higher swelling potential because of their greater affinity for water absorption and large specific surface area.

Sharma, H.D. and Pandey, S.P. (1999) conducted geotechnical characterization of expansive soils of various regions of Madhya Pradesh focusing on infrastructure related issues in black cotton soils. They found that soils of Malwa Plateau region are highly plastic and show significant swelling properties, often causing pavement distress and cracking of lightly loaded structures during moisture variations due to seasons.

Pandian, N. S., Nagaraj, T. S., and Raju, M. (2000) investigated the behaviour of bentonite modified soil systems and concluded that the liquid limits and plastic index of soil increases progressively with the addition of sodium bentonite, whereas maximum dry density decreases. Their observations showed that the physio-mechanical behaviour of the expansive soils is significantly influenced by the bentonite content and also gave valuable insight on the controlled swelling modification of clayey soils.

Thyagaraj, T. and Raao, S. M. (2010) investigated the swelling nature of compact expansive clays and noted that the shrinkage limit is an important parameter that affects swell–shrink properties. Soils with lower shrinkage limit tend to have higher swelling potential. They also indicated that multiple property correlations might provide better prediction of expansive soil behaviour than single property correlations.

Phanikumar, B. R. (2009) The behaviour of expansive clay soils was studied by the laboratory and comparative analysis method. The study covered variation in swelling-related correlations in the various deposits of expansive soil in India and emphasized the importance of the geological conditions of the region on the correlation between the plasticity character and nature and the swelling response of expansive soil.

Jaleh Forouzan, A. (2016) The study showed that the addition of bentonite can create systematic variation in swelling, which can be used to develop regression-based prediction models. This work is of particular interest to the present investigation because it employed controlled modification of clay to investigate expansive behaviour in the laboratory.

Ahirwar, R. K. and Jain, R. K. (2019) have reported geotechnical characteristics of expansive soils from various sites in Malwa Plateau region. They concluded that the general properties of black cotton soils from this area are high liquidity limit and high plastic index which are typical of volume changing clay soils. The authors also pointed out the need for region-specific characterization for accurate prediction of swelling.

Erzin, Y. and Gunes, N. (2011) used ANN methods for prediction of swelling properties of compacted clays based on the most common soil parameters measured. Their research indicated that in some instances, machine learning methods can better capture the non-direct relationship between index properties and swelling response than can traditional empirical equations. The work also illustrated the increasing usefulness of data-driven approaches for the characterization of expansive soils for large-scale problem.

Shahin, M. A., Jaksaa, M. B., and Maier, H. R. (2009) discussed the use of artificially neural systems (ANS) in various fields of geotechnical and soil engineering such as soil classification, settlement prediction and estimation of bearing capacity. They also noted, however, that the reliability of the chosen models is highly dependent on the quality of the data used and proper validation procedures, especially in the case of limited geotechnical data.

Alavi, A. H.; Gandaumi, A. H. (2011) investigated information-driven modelling techniques for geotechnical engineering applications and showed that predictive equations developed using computational methods can be useful alternatives to regression models. They focused their study on the creation of interpretable prediction relationships which can be easily implemented in engineering practice without relying on highly specialized computational tools.

Pal, M. (2006) showed that support vector machine methodologies can be used successfully in classification problems with complex and imbalanced data sets. The results of the study showed that SVM models can be used to efficiently deal with non-linear classification boundaries, and that they can be usefully extended to geotechnical classification problems with multiple soil behaviour categories

Breiman, L. (2001) The ensemble learning method, called Random Forest, was introduced and shown to be able to deal with complex feature interactions in datasets. The study demonstrated that the Random Forest models tend to have consistent prediction power and to avoid over-fitting that is often seen with single decision-tree models. For these reasons, the Random Forest methods have been increasingly used in geotechnical prediction and soil classification research.

A. Zhang and S. K. Vanapalli. (2025) examined the use of machines learn models to predict the wide swelling properties of expanding soils from the soil index properties. They showed that machine learning models can achieve acceptable prediction accuracy for swelling parameters and can be used to model the non-linear response of soil. The study also emphasized the need for correct validation procedures and the need for proper coverage of the datasets to enhance the generalization capability of the model.

Nagaraj, T. S. and Rao, S. M. (1993) studied the prediction of swelling and shrinkage nature of expanding soils based on combinations of index properties. They concluded that the use of combinations of soil parameters could enhance the reliability of the prediction when compared with the use of empirical correlations based on single variables. The study also highlighted the significance of shrinkage characteristics on the assessment of expansive soil behaviour.

Author(s)	Research Focus	Major Findings	Limitation / Relevance to Present Study
Sridharan and Prakash (2000)	Relationship between index properties and swelling behaviour	Observed strong relationship between plasticity characteristics and FSI	Correlations developed for limited Indian expansive soils
Holtz and Kovacs (1981)	Fundamentals of expansive soil behaviour	Explained influence of clay mineralogy and moisture variation on swelling	Primarily theoretical and conceptual in nature
Sharma and Pandey (1999)	Characterization of Madhya Pradesh expansive soils	Reported high plasticity and swelling tendency of Malwa Plateau soils	Limited predictive modelling approach
Pandian et al. (2000)	Bentonite-modified soil behaviour	Bentonite addition increased LL and PI while reducing MDD	Focused mainly on laboratory behaviour
Thyagaraj and Rao (2010)	Swell–shrink behaviour of expansive clays	Identified influence of shrinkage limit on swelling response	Limited regional applicability
Phanikumar (2009)	Behaviour of Indian expansive soils	Highlighted regional variability in swelling correlations	Broad comparative study without ML application
Forouzan (2016)	Bentonite-modified expansive soil mixtures	Demonstrated usefulness of controlled swelling variation for prediction studies	Conducted on non-Indian clay systems
Ahirwar and Jain (2019)	Geotechnical properties of Malwa Plateau BCS	Reported properties comparable to central Indian expansive soils	Region-specific characterization only
Shahin et al. (2009)	ANN applications in geotechnical engineering	Discussed applicability of ML techniques for prediction problems	Emphasized risk of overfitting in small datasets
Zhang and Vanapalli (2025)	ML-based swelling prediction	Showed effectiveness of ML for non-linear swelling prediction	Limited dataset coverage and validation approach

CHAPTER 3

MATERIALS AND METHODS

3.1 RESEARCH METHODOLOGY

The present investigation methodology is divided in two consecutive phases of experimental characterization and predictive modelling of expansive soil behavior, integrated to each other. The first step is to test the soil of black cotton with different amounts of sodium bentonite in the laboratory to produce a controlled variation in the swelling properties of the soil. Empirical relationships between index properties and Free Swell Index (FSI) were developed based on the experimental observations by regression analysis.

In the second phase, machine learning based classification of expansive soils is done using a database of Indian black cotton soil deposits compiled from published geotechnical studies. Machine's learning models like Random Forrest (RF), Supports Vector Machine (SVM) and Artificial Neural Network (ANN) were used to classify the severity of swelling, based on the commonly measured soil parameters.

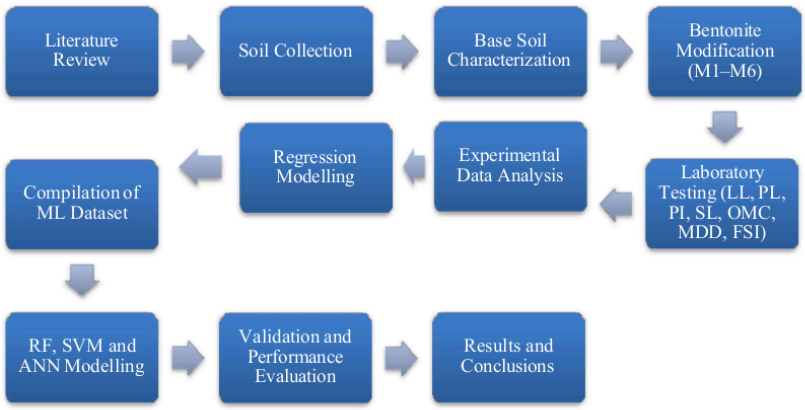


Figure 3.1 Overall Methodology Flowchart

3.2 PHASE-1: EXPERIMENTAL PART

3.2.1 Soil collection

The natural ground surface was taken at an agricultural field in Shajapur district, Madhya Pradesh (23.43°N, 76.85°E) and the black cotton soil was sampled from 0.3 to 1.0 m depth below the natural ground surface. Sampling is done as per IS:2720 (part 1). The black cotton's soil which was used during present investigation was obtained from an agricultural field of Shajapur district Madhya Pradesh, India, at given depth of 0.3m to 1.0m from natural ground surface. The area chosen is in Malwa Plateau zone where large areas of clay are likely to have large swell–shrink properties due to changes in water quantity during the season.

The soil and ground samples were taken and brought to the lab in sealed containers to reduce moisture loss during handling. The soil was then air-dried in the laboratory, broken up with a wooden mallet and sieved through a 425 µm IS sieve to ensure uniformity for experimental testing. Sample preparation was done after coarse particles and organic impurities were filtered out.

During the initial visual examination, it was found that the soil was dark in colour, cohesive and had distinct plasticity characteristics of black cotton soils. In the present study, the commercially available bentonite (sodium bentonite) as per Indian standard specifications was used as a modifying agent. Sodium bentonite had been used as the bentonite as it possessed a high swelling capacity and water absorption capability in comparison to calcium bentonite. The material was bought in powder form and sealed in containers to prevent moisture from the atmosphere from seeping in.

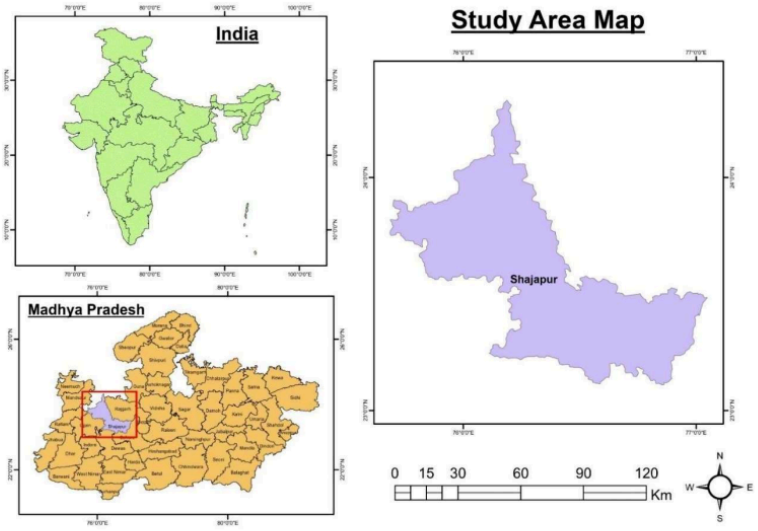


Fig 3.2 Study area of Black Cotton Soil

Fig 3.3 Soil sample

Fig 3.4 Sodium Bentonite used in given study

3.2.2 Mix Preparation

The six sample mixtures consisted of the base BCS with 0%, 2%, 4%, 6%, 8% and 10% of dry bentonite (sodium bentonite) and are designated M1, M2, M4, M6, M8 and M10, respectively. The mixing protocol was a minimum of 15 min of tumbling in a mechanical mixer and a double pass through a 425 μm sieve to ensure particle level homogeneity.

3.3 PROPERTIES OF SOIL

The soil collected in black cotton soil was treated in the laboratory following IS:2720 procedures. The soil was carefully disaggregated without crushing the particles, and passed through a 425 μm sieve after air drying.

The base black cotton soil was mixed with sodium bentonite at various percentages by dry soil and six different soil mixtures were prepared. The contents of Bentonite were chosen to be 0%, 2%, 4%, 6%, 8% and 10%, and were designated as M1, M2, M4, M6, M8 and M10 respectively. The mixing was done by hand and then mechanically blended to ensure the uniformity of bentonite distribution in the soil mass.

The modified bentonite which was used in the present study was modified under controlled conditions to achieve a controlled gradient of swelling under laboratory conditions for the study of the effect of progressive variation in plasticity on the expansive behaviour of the bentonite.

Some of the tests carried out are:

- Classifications Tests
  - Grain Size Analysis
  - Atterberg Limits which are Liquid limit, plastic limit, shrinkage limit
- Physical properties
  - Test of Moisture Content
  - Specific Gravity Test
- Compaction Test
  - Standard Proctor Test

3.3.1 Grain Size Analysis and study:

Particle's size distribution testing is used to measure the precise sizes of the grains of an aggregate, sediment or soil matrix. The assessment is important for classification of the material texture, which influences the structural parameters like load-bearing capacity, shear strength and hydraulic conductivity. Grain size analysis is used primarily to identify if the soil is sand, silt, clay or gravel. It is useful for the interpretation of depositional environment, evaluation of permeability and drainage characteristics and evaluation of suitability for foundations and other structures.

Sieve Analysis is done for the distribution of particles. In Sieve Analysis, soil is pushed through a series of sieves of decreasing sizes and the mass of soil that is retained on each sieve is measured to calculate the percent of each size fraction.

Figure 3.5: Sieve analysis arrangement

Parameters calculated are as:

- **D_{10}**: (Effective size): Particle size at 10% finer, indicates drainage capacity.
- **D_{30}**: Particle size at 30% finer.
- **D_{60}**: Particle size at 60% finer are generally to calculate uniformity coefficient (Cu)

and gradation coefficient (Cc).

Table 3.1: Sieve analysis of soil

Sieve size in mm	Wt. Retained on sieve in grams	Percentage % Wt. Retained	Cumulative % Wt. Retained	Passing percentage %
4.75	10.00	2.00	2.00	98.00
2.36	8.00	1.60	3.60	96.40
1.18	12.00	2.40	6.00	94.00
0.6	15.00	3.00	9.00	91.00
0.3	20.00	4.00	13.00	87.00
0.15	25.00	5.00	18.00	82.00
0.075	35.00	7.00	25.00	75.00
PAN	375.00	75.00	100.00	-

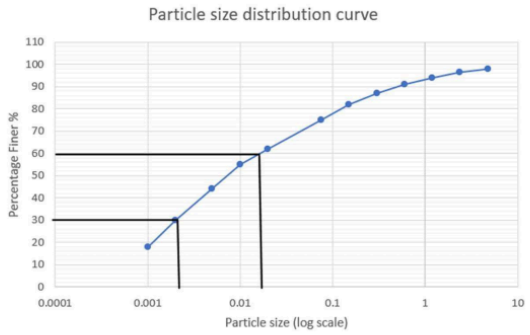


Figure 3.6: Particle's Size Distribution of soil

From graph, we found,

$\mathbf{D_{60}}$ = 0.038 mm, $\mathbf{D_{30}}$ = 0.006 mm, $\mathbf{D_{10}}$ = 0.0012 mm

- Uniformity coefficient which is $\mathbf{C_u}$ = 31.67
- Coefficient of curvature which is $\mathbf{C_c}$ = 0.79

$\mathbf{C_u > 6}$ and $\mathbf{C_c}$ does not lie between **1 -3**

3.3.2 Atterberg Limits

Liquid Limit (W_L)

Liquid limit which is an crucial property of fine-grained soils, and is defined as the moisture content at which a soil becomes a liquid. It is the point where the soil will begin to flow due to its own weight. Main purpose is to assessing the consistency and behavior of clayey soils and to predict soil strength, compressibility and settlement.

Casagrande apparatus is used for finding out the liquid limit.

Figure 3.7: Liquid Limit

Table 3.2: Liquid limit of soil

Observation and Calculation				
Container No.	1	2	3	4
Wt. of Container W_1, in g	18.45	19.20	17.80	18.90
Wt. of Container and wet soil W_2, in g	49.95	54.40	52.03	52.20
Wt. of Container and dry soil W_3, in g	38.45	41.20	38.80	38.90
Wt. of water ($W_2 - W_3$), in g	11.50	13.20	13.23	13.30
Wt. of dry soil ($W_3 - W_1$) in g	20.00	22.00	21.00	20.00
Moisture content (%) $= \frac{(W2-W3)}{(W3-W1)} \times 100\%$	57.50	60.00	63.00	66.50
No. of blows	35	28	22	17

By Interpolation,

For 25 number of blows, Water content = 62%

Liquid Limit = 62%

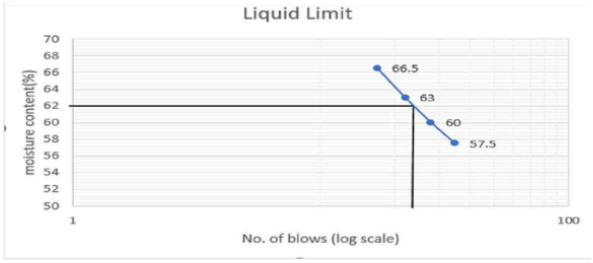


Figure 3.8: Variation of moisture content with no. of blows

The Liquid Limit value is 62.0% with the aid of graph

Plastic limit (W_P)

The plastic limit is the water content at which a soil becomes plastic and semisolid. The smallest water content at which a soil will just begin to crumble when rolled into a thread about 3mm in diameter.

Figure 3.9: Plastic Limit

Observation and Calculation

Table 3.3: Plastic Limit

Sample No.	1	2
Weight of Container W_1, in g	18.50	17.80
Weight of Container and wet soil W_2, in g	28.44	27.09
Weight of Container and dry soil W_3, in g	26.50	25.30
Weight of water ($W_2 - W_3$), in g	1.94	1.79
Weight of dry soil ($W_3 - W_1$) in g	8.00	7.50
Moisture content (%) = $\frac{(W_2-W_3)}{(W_3-W_1)}$ x 100%	24.25	23.87

After calculating it the avg. plastic limit will be = $\frac{24.25+23.87}{2}$ = 24.06% ≈ 24%

Plasticity Index (I_p):

The range of water content over which a soil exhibits plastically is called the plasticity index (PI). It is calculated as:

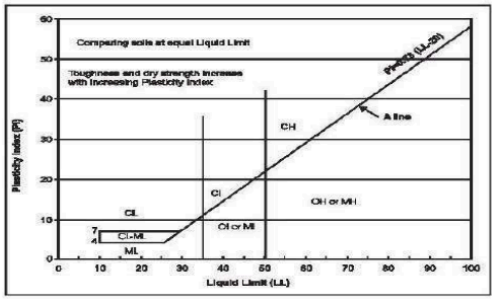
$$PI = LL - PL$$

$$I_p = W_l - W_p = (62-24)\ \%$$

m $\mathbf{I_p = 38\%}$

Through the A line and Plasticity index we find that the soil has high plasticity (CH).

Figure 3.10: Plasticity Chart



SHRINKAGE LIMIT (IS:2720 Part 6)

Base BCS (M1) — SL = 12%

Method: Mercury displacement

Formula: SL (%) = $w_i - [(V - V_d) / M_d] \times 100$

Constants:

Volume of shrinkage dish, V = 24.0 cm^3

Density of mercury, ρ_Hg = 13.6 g/cm^3

Mass of mercury to fill dish = 24.0 × 13.6 = **326.4 g**

Table 3.4: Shrinkage Limit

Observation	Trial 1	Trial 2
Mass of empty shrinkage dish W_1 (g)	45.30	44.70
Mass of dish + wet soil W_2 (g)	97.14	95.94
Mass of dish + dry soil W_3 (g)	81.30	80.70
Mass of dry soil Md = $W_3 - W_1$ (g)	36.00	36.00
Mass of water Mw = $W_2 - W_3$ (g)	15.84	15.24
Initial water content w_i = Mw/Md × 100 (%)	44.00	42.33
Mass of Hg to fill dish (g)	326.40	326.40
Mass of Hg displaced by dry pat (g)	169.73	163.59
Volume of dry pat V_d = M_Hg / 13.6 (cm^3)	12.48	12.03
$V - V_d$ (cm^3)	11.52	11.97
SL = $w_i - (V - V_d)$/Md × 100 (%)	**12.00**	**11.92**
Adopted SL (%)	**12.0**	

Average SL = (12.00 + 11.92) / 2 = 11.96 ≈ 12.0%

3.3.3 Moisture Content Test

Moisture content is the percentage of water in the soil or material based upon the dry weight of the material. It is a crucial property in soil mechanics, agriculture and material science, as it relates to the strength, compaction and behavior of soil.

Moisture Content of this soil from moisture meter is found to be **10.8 %**.

3.3.4 Specific Gravity Test

Specific gravity is the ratio of weight of soil solids to the weight of equal volume of water. It is useful in computing the unit weight of the soil under various conditions and it is also useful in the determination of particle size by wet analysis.

The specific gravity of soil sample is determined by density bottle method.

(a) (b) (c) (d) (e)

Figure 3.11: Specific Gravity using pycnometer, density bottle

Calculations

Specific Gravity (G) is calculated as: $(\mathbf{G}) = \frac{(W_2 - W_1)}{(W_4 - W_1) - (W_3 - W_2)}$

Where:

W_1 = Weight of empty pycnometer. W_2= Weight of pycnometer with soil

W_3 = Weight of pycnometer with soil and water W_4 = Weight of pycnometer with water only

Table 3.5: Specific Gravity

Weight (g)	Sample 1	Sample 2
W_1	32.56	31.92
W_2	42.56	41.92
W_3	88.70	88.06
W_4	82.39	81.75
G	2.713	2.707

Sample calculation

For sample 1: $G = \frac{(42.56 - 32.56)}{(82.39 - 32.56) - (88.70 - 42)}$

Average value of specific gravity is 2.71 at room temperature.

3.3.5 Standard Proctor Test

Compaction is the process of mechanically rearranging and packing soil particles into a closer state of contact so as to reduce the porosity (or voids ratio) of the soil and, as a result, increase the dry density of the soil mass by rapidly and dynamically expelling the air that is present in the voids. The water content of the soil, the compactive effort, the soil type and admixtures are the most important factors contributing to the density of the compacted soil.

The primary objective of laboratory compaction test is to find out the optimum water content (also known as maximum dry density (MDD)) of the soil grains, at which the unit volume of soil grains has its maximum weight.

Observations

- Diameter of mould = 10 cm
- Height of mould = 12.73 cm
- Volume of mould = 1000 cm^3
- Empty weight of mould (W_1) = 4250 g
- Weight of rammer = 2.5 kg

Calculations

- Bulk Density, $y_b = \frac{(W_2 - W_1)}{\text{Volume}}$ g/cc
- Dry Density, $y_d = \frac{y_b}{(1+W)}$ g/cc

Where,

W_1 = Weight of mould

W_2 = Weight of compacted soil + mould

W = Moisture content

Table 3.6: Standard Proctor Test

Serial Number	Wt. of mould (Kg) W_1	Wt. of compacted soil + mould (Kg) W_2	Vol. of mould (cm^3)	Water mixed (%) W	Bulk density (g/cc) y_b	Moisture content (%)	Dry density (g/cc) y_d
1	4.250	5.982	1000	0	1.732	14.0	1.519
2	4.250	6.126	1000	4	1.876	16.5	1.610
3	4.250	6.246	1000	8	1.996	19.5	1.671
4	4.250	6.239	1000	12	1.989	22.0	1.630
5	4.250	6.188	1000	16	1.938	25.0	1.550

So, from the above table we obtain,

Max. dry density = 16.38 KN/m^3

Optimum moisture content = 19.5 %

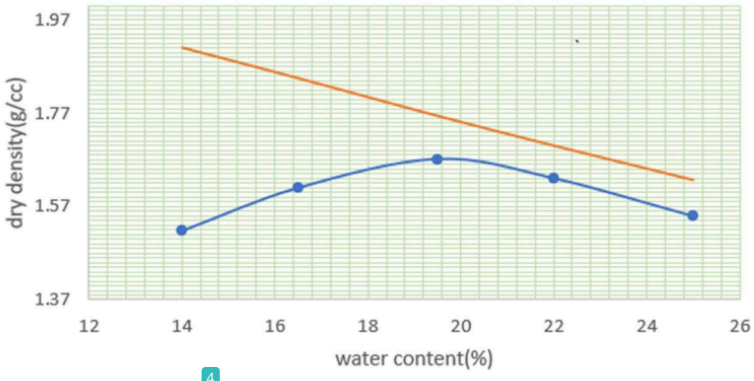


Figure 3.12: The relationship between the water content and dry density

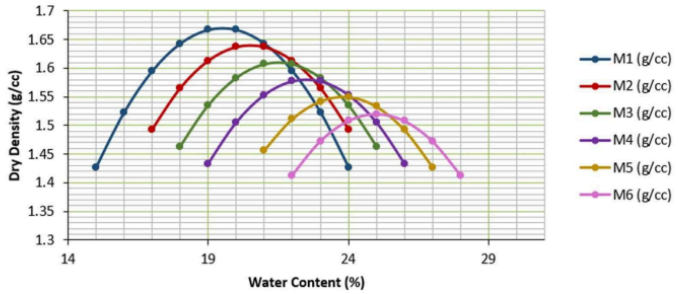


Fig 3.13: Compaction Curve Combined for all Mix (M1-M6)

3.3.6 Free swelling index test

The Free-Swelling Index Test is a simple test that is used to indicate whether a soil has high shrink–swell nature. The test was first developed in the 1950s by Holtz and Gibbs (1956) which involves placing a small sample of soil in distilled water and measuring the volume of the soil sample. Free Swell Index Test is a direct swell test that is a measure of the volume of the swells (as compared to non-swelled volume) and also gives a qualitative swell index. The results may be compared to soils and should be used as an indicator test to decide if a more simple direct test should be done on a soil to quantify the swelling behavior at fixed conditions of confinement and water content change.

Procedure:

- 10 g of oven dry soil through 425 micrometer IS sieve was added to two graduated cylinders (100 mL each)
- 1 cylinder filled with given distilled water, other with kerosene
- The volume of the settled particles after 24 hours

$$\mathbf{FSI\ (\%) = [(V_d - V_k)/V_k] \times 100}$$

Table 3.7: Free swell index Test

Mix	Bent. %	Soil (g)	V_w T1 (mL)	V_w T2 (mL)	V_w T3 (mL)	Mean V_d (mL)	V_k (mL)	FSI (%)	IS Class
M1	0	10	16.3	16.5	16.7	16.50	10.0	**65**	High
M2	2	10	17.3	17.5	17.7	17.50	10.0	**75**	High
M3	4	10	18.4	18.6	18.8	18.60	10.0	**86**	High
M4	6	10	19.5	19.7	19.9	19.70	10.0	**97**	High
M5	8	10	20.6	20.8	21.0	20.80	10.0	**108**	Very High
M6	10	10	21.6	21.8	22.0	21.80	10.0	**118**	Very High

3.4 Phase 2: Machine Learning Dataset

3.4.1 Dataset Compilation

A database of 186 Indian BCS samples was developed from 40 published geotechnical studies and conference proceedings. The data is applicable for 10 states in India: Madhya Pradesh, Maharashtra, Gujarat, Andhra Pradesh, Tamil Nadu, Rajasthan, Karnataka, Bihar and Chhattisgarh. The input retained for the ML modelling is the retained LL (%), PI (%), OMC (%) and MDD (g/cc) which are all routinely measured in site investigations. The target variable is the IS:1498 swelling risk class as reported in the source publication for either FSI or swell indicator.

3.4.2 Feature and Target Description

Table 3.8: ML Dataset Feature Summary

Feature	Range	Physical Significance
LL (%)	38.2 – 103.0	Total water holding of clay
PI (%)	19.2 – 65.2	Swelling-active water content range
OMC (%)	15.5 – 36.0	Moisture at optimum compaction
MDD (g/cc)	1.35 – 1.92	Packing density of soil fabric
Target: Swelling Class	Low/Medium/High/Very High	IS:1498 classification

3.4.3 Class Distribution and Imbalance

Table 3.9: Class Distribution in ML Dataset

Class	Count	Percentage
Low	6	3.2%
Medium	80	43.0%
High	93	50.0%
Very High	7	3.8%

The models have been built with class-weighting technique for SVM and Random Forest architecture and 5-fold stratified cross-validation method. This way, the algorithms will be penalized correctly in the under-represented categories.

3.4.4 Machine Learning Models

(a) Random Forest (RF)

A random subset of features is used at each split and each tree is grown using a bootstrap sample of the data. This randomness helps to avoid overfitting and enhance generalization. The number of trees in this study was 100 trees with the square-root feature selection rule per split. Class weights were determined inversely to the frequencies of the classes to deal with unbalanced classes.

(b) Support Vector Machine (SVM)

SVM determines the hyperplane that typically has the largest margin between the classes in a high-dimensional space. Non-linear class boundaries were dealt with using the RBF kernel. The parameters of the regularization λ and γ (kernel width) were optimized using the grid search procedure in the cross validation loop. The one-vs-one strategy was used for multi-class classification.

(c) Artificial Neural Network (ANN / MLP)

A Multilayer Perceptron with known hidden layer (10 neurons), Tanh activation function and L-BFGS solver was implemented. Min-Max scaling was used to normalize input features before training. The network has been trained for the purpose of minimizing the known cross entropy loss over the four swelling classes.

3.4.5 Model Validation

The main validation method used was stratified Fivefold crossing validation. The data set was divided into five folds, each of which had the same class distribution as the data set. 4 folds were used for training and 1 fold for testing for each iteration. All folds were used to get the following performance metrics:

- **Accuracy:** Fraction of correctly arranged samples
- **Weighted F1-score:** Harmonic mean of precision and recall, weighted by class support
- **Cohen's Kappa (κ):** Agreement between predicted and actual classes beyond chance
- **Confusion Matrix:** Class-wise breakdown of correct and incorrect predictions

Leave One Out Cross Validation (LOOCV) has been applied to the empirical Phase 1 data set (n = 6), in which the model is trained on 5 samples and tested on 1 sample, and this procedure is repeated 6 times.

3.5 RESULT

The below table shows the properties of soil obtained from lab test.

Table 3.10: Properties of Soil

Properties		Test Result
Grain Size Analysis	Uniformity coefficient, Cu	31.67
	Coefficient of curvature, Cc	0.79
Atterberg Limits	Liquid Limit, LL	62%
	Plastic Limit, PL	24%
	Plasticity Index, PI	38%
	Shrinkage Limit, SL	12%
Moisture Content		12.50%
Specific Gravity		2.71
Standard Proctor Test	Max. dry density	16.38 KN/m^3
	Optimum moisture content	19.5 %
Free swell index Test	FSI	65%
Type of soil	Clay of High Compressibility CH	

CHAPTER 4

EMPIRICAL REGRESSION ANALYSIS

4.1 Introduction

The first phase of this study aimed to establish reliable empirical equations to predict the Free Swelling Index (FSI) based on index properties of the soil, which were liquid limit (LL), plasticity index (PI) and shrinkage limit (SL). Generally these four parameters are considered as index properties of clay and silty soils and represent the moisture content limits of the soil states of solid, semi-solid, plastic and liquid. The single parent BCS was deliberately chosen to give a clean, monotonic dataset with a known physics of swelling, so that the resulting regression equations would be true relationships between the Atterberg limits and FSI, and not artefacts of geological variability.

It should be noted that the empirical equations presented in this chapter are meant to be used in conjunction with the FSI test, and not as a substitute for it. The swelling potential as determined by FSI as per IS:9451 is the most direct and authoritative measure of swelling potential. The regression models developed here are for the preliminary estimation at the desk level, when FSI results are not yet available but the data on Atterberg limits from the routine site investigation are available.

All six mixtures are presented in order, followed by a Pearson correlation analysis that identifies the most significant Atterberg limit parameters that correlate with FSI; five regression equations are developed based on these data; and an honest assessment of each equation's performance is provided by a leave-one-out cross validation (LOOCV) exercise. The chapter ends by comparing with published FSI–PI correlations found in the Indian geotechnical literature.

4.2 Complete Laboratory Test Results

The Atterberg limits and FSI of all six BCS–bentonite mixtures are summarised in Table 4.1. Averages of three independent determinations are reported as all values. Compaction parameters (OMC and MDD) are also included for completeness but are compaction properties and not index properties.

Table 4.1: Complete Index Properties and FSI (Shajapur BCS + Bentonite)

Mix	Bent. (%)	LL (%)	PL (%)	PI (%)	SL (%)	Gs	OMC (%)	MDD (g/cc)	FSI (%)	IS:9451 Class
M1	0	62.0	24.0	38.0	12.0	2.71	19.5	1.67	65	High
M2	2	67.0	24.6	42.4	11.5	2.71	20.5	1.64	75	High
M3	4	72.0	25.2	46.8	11.0	2.70	21.5	1.61	86	High
M4	6	78.0	25.8	52.2	10.6	2.70	22.5	1.58	97	High
M5	8	83.0	26.4	56.6	10.2	2.69	23.8	1.55	108	Very High
M6	10	88.0	27.0	61.0	9.8	2.69	25.0	1.52	118	Very High

4.3 Trends in Index Properties with Increasing Bentonite Content

4.3.1 Liquid Limit and Plasticity Index

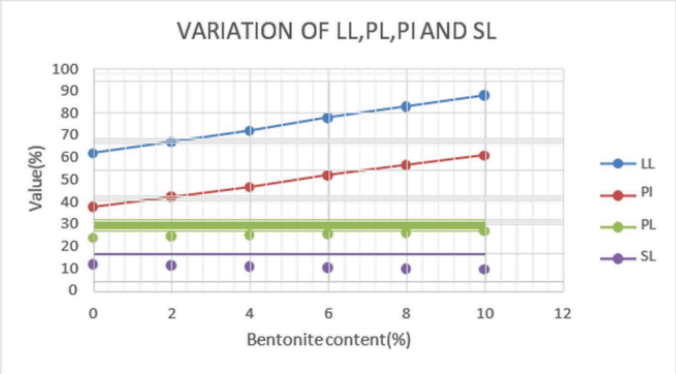


Fig 4.1 Atterberg Limits (LL, PL, SL) vs Bentonite Content

The most striking aspect of Figure 4.1 is the consistency of the LL increase as the bentonite content increases: 62.0% for unmodified BCS (M1) to 88.0% for 10% bentonite (M6) is an increase of 26 percentage points. The average increment is around 2.6 percentage points per 1% bentonite addition which is similar nearly to the results which are reported by Pandian et al., (2000) in Maharashtra BCS (2.2–3.5 pp/). The plastic limit increased slightly from 24.0% (M1) to 27.0% (M6) due to the fact that the sodium bentonite's plastic limit (approx. 40-55%) is not significantly different from the base BCS plastic limit. The plasticity index thus increased from 38.0% to 61.0% (+23 pp), which represents a 61% relative increase, mainly due to an increase in LL. All the six mixtures are found to be CH (Clay of High Compressibility) on the IS:1498 plasticity chart and are found to be same throughout the dosage range.

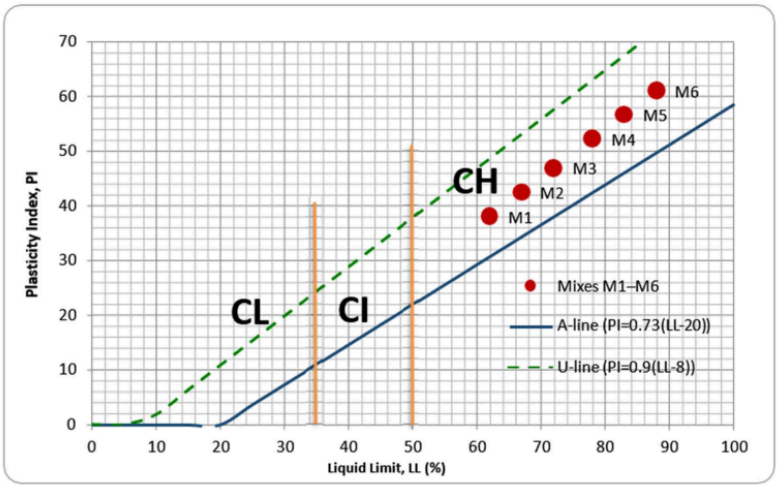


Fig 4.2 All six mixes on the IS:1498 plasticity chart

4.3.2 Shrinkage Limit

The shrinkage limit reduced monotonously from 12.0% (M1) to 9.8% (M6) which is a reduction of 2.2 percentage points over the whole dosage range, at a rate of approximately 0.22 pp per 1% bentonite. The lower the shrinkage limit of the known soil, the greater the range of moisture content that will cause volume changes in the soil in period of the annual wet–dry cycle, increasing the potential for cumulative structural damage. The SL was below 10.0% from M4 (SL = 10.6%) to M5 (SL = 10.2%), which is estimated to be around 7.1% bentonite by linear interpolation. This dose is practically significant: above this dose, the soil can be subjected to more frequent seasonal swell-shrink cycles. The mechanism as explained by Datta and Sridharan (2004) is that the inter-aggregate voids are filled by montmorillonite platelets which slow the rate of air entry during drying and reduce the shrinkage limit. SL reflects this shrinkage end of the plasticity spectrum, and is negatively correlated with FSI (Pearson r = −0.9981), so it contains predictive information in addition to that provided by LL and PI alone, and is explicitly included in the multi-variable regression models.

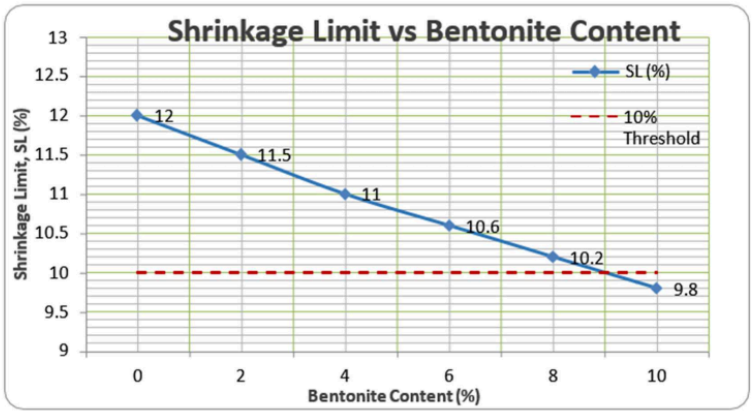


Fig 4.3 SL vs bentonite % (declining trend) with 10% threshold marked

4.3.3 Specific Gravity, OMC, and MDD

The OMC was found to be increasing with the dosage, and the MDD was found to be decreasing with the dosage in the dosage range. Both these trends are physically expected: sodium bentonite platelets need more water to reach the optimum packing, and the lower specific gravity of bentonite (Gs ≈ 2.60–2.65) compared with the basalt derived BCS minerals (Gs ≈ 2.71) will gradually decrease the dry density of the mixture. The statistical correlations between OMC and MDD and FSI in this data set are high but these are compaction characteristics rather than index properties of the soil fabric in the Atterberg sense. Therefore they are not used as predictors in the empirical regression equations of this chapter.

4.3.4 Free Swell Index

The FSI results are presented in Figure 4.2 with the IS:9451 class boundaries. The base BCS (M1) has FSI = 65%, which is close to the lower limit of the High class (FSI 50-100%). Each increment of bentonite pushed the FSI higher: 75% at 2%, 86% at 4%, 97% at 6%, 108% at 8%, and 118% at 10%. The change between M4 (97%) and M5 (108%) is significant as it falls at the IS:9451 High to Very High class boundary at FSI = 100%.

This transition point occurs at about 7.0–7.5% bentonite by dry weight by linear interpolation between M4 and M5, which has direct practical implications to field contamination scenarios

The swelling behaviour of this soil is indeed a smooth, monotonic function of bentonite content, as the FSI data below shows, and the index properties follow the FSI closely enough that any of them could be used as a predictor with high accuracy.

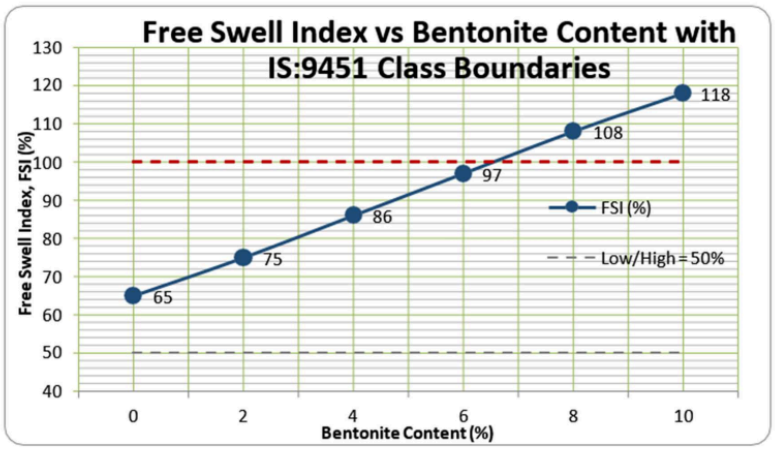


Fig 4.5 FSI vs bentonite % with IS:9451 class boundary lines

4.4 Analysis of Correlation

Table 4.2 displays the known Pearson correlation coefficient between each measured property and FSI. The four Atterberg limit parameters, LL, PL, PI, and SL, are indeed the index properties of this soil system, while OMC and MDD are included for comparison purposes, but are compaction parameters and not included in the regression models because it is not a soil index property (a predictor that depends on compactive effort is not an intrinsic soil index).

Table 4.2: Pearson Correlation Coefficients between Index Properties and FSI

Variable	Pearson r	Direction	Interpretation
Liquid Limit (LL)	+0.9998	Positive	Very strong - LL is the most predictive single variable
Plastic Limit (PL)	+0.9999	Positive	Very strong -but PL changes so little (24→27%) it adds little unique information
Plasticity Index (PI)	+0.9997	Positive	Very strong - physically the most justified predictor
Shrinkage Limit (SL)	−0.9981	Negative	Strong negative - SL declines as FSI rises
OMC	+0.9985	Positive	Strong - but OMC is compaction-dependent, less fundamental
MDD	−0.9976	Negative	Strong negative - density falls as swelling rises

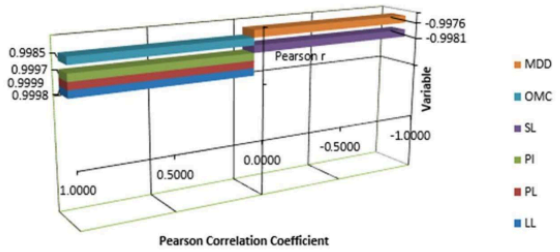


Fig 4.6 Pearson r between Index Properties and FSI

The regression models are developed using the Atterberg limit index properties on the basis of the correlation analysis and published Indian literature. PI is chosen as the primary single predictor (Eq. 1) because of its known physical relationship with FSI (Sridharan and Prakash, 2000;Phanikumar, 2009). The use of LL for Eq. 2 because this is the most commonly available Atterberg parameter from site investigation data. Equations 3 and 4 present SL as a second Atterberg variable to investigate if shrinkage limit is useful in addition to LL or PI..

4.5 Regression Models Development

The six-point (M1–M6) data set was used to fit five empirical models by ordinary least squares regression. The fitted equations, the statistical parameters and the LOOCV validation errors are shown in the Table 4.3.

Table 4.3: Empirical FSI Prediction Equations — Summary

Eq.	Expression	Type	R^2	Adj. R^2	RMSE (%)	LOOCV RMSE (%)	Predictors Used
1	FSI = 2.299·PI − 22.315	Linear (PI)	0.9995	0.9994	0.42	0.53	PI only
2	FSI = 2.037·LL − 61.289	Linear (LL)	0.9996	0.9995	0.37	0.47	LL only
3	FSI = 1.753·LL − 3.419·SL − 2.895	Multi-linear (LL, SL)	0.9997	0.9995	0.46	0.85	LL + SL
4	FSI = 1.903·PI − 4.231·SL + 43.224	Multi-linear (PI, SL)	0.9996	0.9993	0.36	0.91	PI + SL
5	FSI = 11.389·LL − 10.556·PI − 240.167	Multi-linear (LL, PI)	0.9998	0.9997	0.24	0.87	LL + PI

4.5.1 Equations 1 and 2 - Single Atterberg Limit Predictors

The models that are most robust for this dataset are Equation 1 (PI predictor) and Equation 2 (LL predictor). The excellent generalisation of their LOOCV RMSE data of 0.53% and 0.47%, even with the given model trained on five of the six mixtures, demonstrates this. The form of both the equations is simple and physically meaningful: FSI rises linearly with PI or LL, which is consistent with the well established relationship between plasticity and expansiveness as reported by Sridharan and Prakash (2000) and Phanikumar (2009). The most physically justified single predictor of swelling potential is recommended as the primary practical tool, which is equation 1. If the liquid limit is the only test that has been performed, equation 2 is a good alternative.

4.5.2 Equation 3 - Dual Atterberg Limit Model: LL and SL

Equation 3 is the combination of liquid limit and shrinkage limit, both of which are Atterberg limit index properties, into one regression model. SL is physically justified because it is the upper moisture limit of plasticity and is positively correlated with the swelling (r = +0.9998), SL is the lower moisture boundary beyond which no further volume change occurs and is negatively correlated with the FSI (r = −0.9981). When combined with these two end-points, the water-retention tendency and the shrinkage susceptibility of the soil, then the R^2 for these two reaches is 0.9997 and the in-sample RMSE is 0.46%.

FSI = 1.753·LL − 3.419·SL − 2.895 (R^2 = 0.9997, LOOCV RMSE = 0.85%)

The LOOCV RMSE of 0.85% is higher than those of Equations 1 and 2 (0.53% and 0.47%), as expected when a two-predictor model is fit to six data points, where the second predictor will use one of the degrees of freedom of the model, increasing the variance of the predictions in each iteration of the LOOCV. However, equation 3 is the only model in this study that incorporates two different Atterberg limits over the entire range of the plasticity spectrum, namely the upper end (LL) and the lower end (SL), and is still a useful model when both the LL and SL are available from the standard routine IS:2720 suite of tests.

4.5.3 Equations 4 and 5 - Other Multi-Atterberg Combinations

The equations that combine PI and SL yield LOOCV RMSE = 0.91% (Equation 4) and the equations that combine LL and PI yield LOOCV RMSE = 0.87% (Equation 5). Again, the errors of the LOOCV models compared to the single-variable models are higher because of the small size of the dataset as opposed to any deficiency in the predictors. A note of caution for Equation 5: LL is almost collinear with PI in this set of data (with PL changing only between 24.0% and 27.0%), so that the regression surface is poorly constrained in the LL–PI predictor space.

This multi-collinear behavior artificially boosts the model coefficients (11.389 and -10.556 for the LL and PI, respectively), and the predictive algorithm is very sensitive to small changes in the input. In the field engineering applications, therefore, it is highly recommended to use either Equation 1 or Equation 2 rather than Equation 5. The category averages show a consistent trend of increasing severity of swelling with increasing LL, PI, and OMC, and a decreasing trend in SL and MDD with increasing severity of swelling.

4.6 Leave One Out Cross Validation Results

Table 4.4: LOOCV Predicted vs Observed FSI - Equations 1, 2, and 4

Mix	PI	SL	LL	FSI Obs.	Eq.1 LOOCV Pred.	Eq.1 Error	Eq.2 LOOCV Pred.	Eq.2 Error	Eq.4 LOOCV Pred.	Eq.4 Error
M1	38.0	12.0	62.0	65.0	64.4	−0.6	64.8	−0.2	63.1	−1.9
M2	42.4	11.5	67.0	75.0	75.2	+0.2	75.3	+0.3	74.6	−0.4
M3	46.8	11.0	72.0	86.0	85.6	−0.4	85.5	−0.5	85.9	−0.1
M4	52.2	10.6	78.0	97.0	97.8	+0.8	97.2	+0.2	99.4	+2.4
M5	56.6	10.2	83.0	108.0	107.4	−0.6	107.8	−0.2	108.9	+0.9
M6	61.0	9.8	88.0	118.0	117.8	−0.2	117.6	−0.4	116.4	−1.6
LOOCV RMSE						**0.53%**		**0.47%**		**0.91%**

The LOOCV ranking shows that the single-variable Atterberg limit models (Equations 1 and 2) are the most reliable models for this six-point data set. Theoretically more rich (two Atterberg limits), equations 3 and 4 have higher LOOCV RMSE because of the limited sample size. This is not a fundamental problem with the multi-variable Atterberg approach, but a limitation of the data size for the LOOCV with n = 6: only five points are used for a two-parameter model in each training fold, leaving limited information for precisely estimating the coefficients. Equations 3 and 4 would be expected to have better cross-validated performance with a larger independent set of data of similar soil type.

4.7 Observed vs Predicted FSI - Summary Table

Table 4.5:Observed vs Predicted FSI - Summary Table

Mix	FSI Obs.	Eq.1 Pred.	Err.%	Eq.2 Pred.	Err.%	Eq.3 Pred.	Err.%	Eq.4 Pred.	Err.%	Eq.5 Pred.	Err.%
M1	65.0	65.1	+0.1	64.9	−0.1	64.8	−0.3	64.8	−0.3	65.2	+0.3
M2	75.0	75.2	+0.3	75.3	+0.4	74.4	−0.8	75.3	+0.4	74.9	−0.1
M3	86.0	85.4	−0.7	85.8	−0.2	85.8	−0.2	86.0	0.0	86.1	+0.1
M4	97.0	97.7	+0.7	97.2	+0.2	97.5	+0.5	96.7	−0.3	97.0	0.0
M5	108.0	107.9	−0.1	107.6	−0.4	107.8	−0.2	107.7	−0.3	108.0	0.0
M6	118.0	117.9	−0.1	118.0	0.0	117.4	−0.5	118.2	+0.2	117.9	−0.1

4.8 Comparison with Published Literature

Table 4.6: Comparison of Equation 1 Slope with Published FSI-PI Correlations

Source	Study Soil	PI Range	FSI–PI Slope	FSI–PI Intercept	R^2
Sridharan & Prakash (2000)	15 Indian expansive soils	20–75%	2.12	−21.4	0.88
Phanikumar (2009)	BCS — AP, MH, TN	25–65%	1.85–2.25	Varies	0.85–0.92
Thyagaraj & Rao (2010)	Tamil Nadu BCS	30–60%	2.05	−19.8	0.91
Jaleh Forouzan (2016)	Kaol.–Bent. mixtures	10–80%	~2.30 (low doses)	Varies	0.99
Present Study (Eq.1)	**Shajapur BCS + Bent.**	**38–61%**	**2.299**	**−22.315**	**0.9995**

The slope of 2.299 in Equation 1 is physically consistent with published Indian literature (1.85-2.30 range) and slightly higher than the average, which is due to the high inherited smectite content of Shajapur BCS (clay content ~45%, activity ~0.84). The inclusion of sodium bentonite in addition to the smectite component of the soil gives a steeper FSI–PI response than a soil with the same PI dominated by kaolinite.

CHAPTER 5

MACHINE LEARNING CLASSIFICATION

5.1 INTRODUCTION

The empirical equations given in Chapter 4 are applicable only for the soil type under investigation (Shajapur BCS modified with sodium bentonite) and may not be directly applicable to Indian BCS from other geological formations as the clay mineralogy, clay content and geological history are different. The problem presented in this chapter is a fundamentally different problem and is much broader: given the routine laboratory data from any BCS site in India, can a trained classifier reliably classify the soil into its IS:1498 swelling risk category (Low, Medium, High, or Very High)

It is a multi-class classification problem which has been solved by a compiled database of 186 Indian BCS samples from ten states. The feature set includes the four Atterberg limit index properties which are liquid limit (LL), Toughness Index, plasticity index (PI), shrinkage limit (SL), and compaction parameters optimum moisture content (OMC) and maximum drying density (MDD). The classifiers tested were Random Forest (RF), Supporting Vector Machine (SVM) with RBF kernel, and Artificial Neural Network (ANN/MLP). Stratified five-fold cross-validation which is used to evaluating the model performance as the data set is imbalanced.

Three classifiers were tested: Random Forest (RF), Supporting Vector Machine (SVM) with RBF kernel and Artificial Neural Network (ANN). The chapter introduces the dataset, data preprocessing, hyperparameter optimization, performance evaluation using stratified 5-fold cross validation, confusion matrices, feature importance analysis, and finally discusses the most appropriate model for real-world use.

5.2 Data Overview and Class Distribution

The 186-sample dataset compiled from published Indian BCS literature has the following class distribution:

Table 5.1: Class Distribution in Machine Learning Dataset

Swelling Class	IS:9451 FSI Range	Count	Percentage
Low	FSI < 50%	6	3.2%
Medium	50% ≤ FSI ≤ 65%	80	43.0%
High	65% < FSI ≤ 100%	93	50.0%
Very High	FSI > 100%	7	3.8%
Total		**186**	**100%**

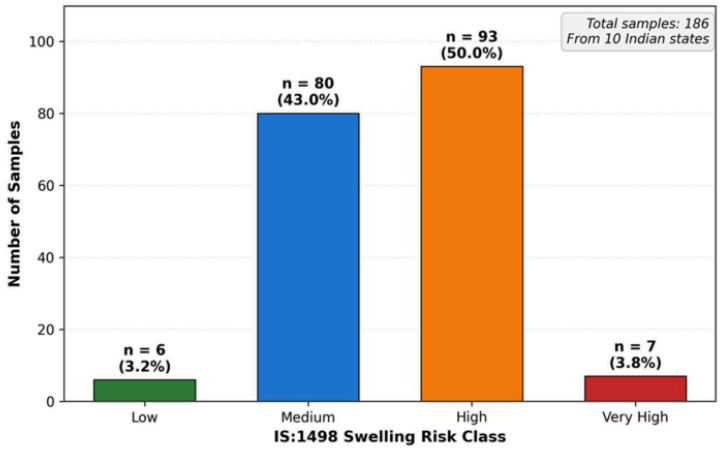


Fig 5.1 Class distribution bar chart (n=186)

The data set is highly imbalanced with the two major classes (Medium and High) having 93% of the samples and the two minor classes (Low and Very High) having 7% of the samples. It is not a sampling artefact, but a true representation of the Indian BCS literature, because the most commonly encountered soil conditions are reported by the researchers.

To tackle this imbalance, two complementary strategies were followed: stratified Five fold cross validation, to ensure that each fold had representative proportions of classes; and class-weighted training, which gave greater penalty to misclassification of minority class errors than majority class errors.

5.3 Feature Set and Descriptive Statistics

Table 5.2: Feature Descriptive Statistics by Swelling Class

Feature	Low (n=6) Mean	Medium (n=80) Mean	High (n=93) Mean	Very High (n=7) Mean	Property Type
LL (%)	42.3	55.8	71.4	89.6	Atterberg Limit
PI (%)	21.5	31.2	42.8	57.3	Atterberg Limit
SL (%)	17.2	14.1	11.5	9.4	Atterberg Limit
OMC (%)	16.8	19.4	22.7	26.1	Compaction Param.
MDD (g/cc)	1.84	1.72	1.61	1.49	Compaction Param.

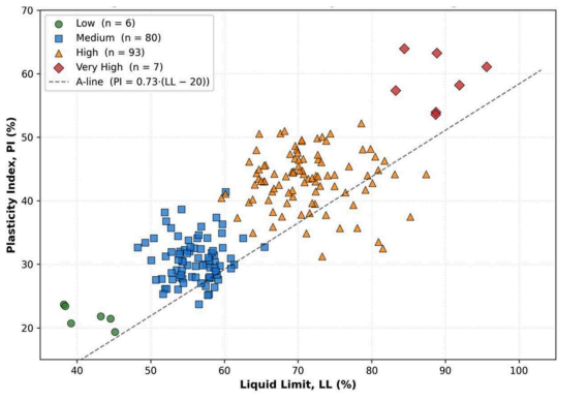


Fig 5.2 Scatter plot: PI vs LL coloured by swelling class

The expected monotonic pattern is displayed for the class means: LL, PI, OMC increase with increasing swelling class, and SL, MDD decrease with increasing swelling class. The decrease in SL is monotonic, ranging from 17.2% (Low class) to 9.4% (Very High class), which demonstrates that the shrinkage limit provides independent discriminating information to be used for the classification of swelling risk and should therefore be incorporated as a model feature. The main problem for the classifier is, however, that there is a large overlap in the distribution of the adjacent classes (Medium and High, in particular).

Table 5.3 Normalisation Parameters-Full Dataset (Min-Max Scaling to [0,1])

Feature	x_min	x_max	Range	Property Type
LL (%)	38.2	103.0	64.8	Atterberg Limit
PI (%)	19.2	65.2	46.0	Atterberg Limit
SL (%)	8.0	20.0	12.0	Atterberg Limit
OMC (%)	15.5	36.0	20.5	Compaction Param.
MDD (g/cc)	1.35	1.92	0.57	Compaction Param.

Note: All scaling parameters are computed exclusively from the training fold in each cross-validation iteration. The test fold is transformed using only training-fold parameters, preventing data leakage

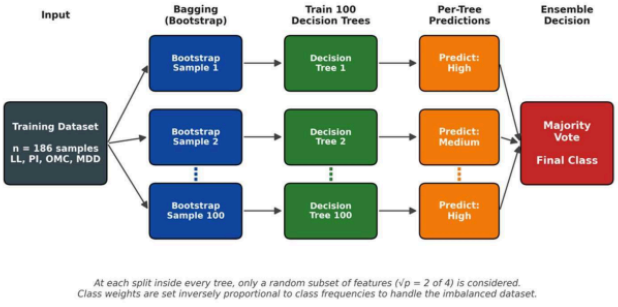


Fig 5.3 Random Forest Ensemble Architecture

5.4 Machine Learning and Cross-Validation Performance

The faactors namely Random Forest (RF), Supports Vector Machine (SVM) with RBF kernel and Artificial Neural Network (ANN/MLP) were trained and tested through stratified 5-fold cross validations. Which is used to overcome the class imbalance problem, class-weighted training was adopted for RF and SVM. Hyperparameters were chosen using an inner nested cross validation loop in every outer fold.

Table 5.4: ML Model Performance- Stratified 5-fold Cross-Validation (n=186)

Model	CV Accuracy (mean ± std)	Weighted F1 (mean ± std)	Cohen's κ (mean ± std)	Balanced Accuracy (mean ± std)	Training Accuracy
Random Forest	0.880 ± 0.083	0.870 ± 0.089	0.792 ± 0.110	0.860 ± 0.088	1.000
SVM (RBF)	0.820 ± 0.098	0.807 ± 0.104	0.681 ± 0.136	0.808 ± 0.107	0.960
ANN (MLP)	0.800 ± 0.115	0.781 ± 0.126	0.639 ± 0.161	0.691 ± 0.145	0.980

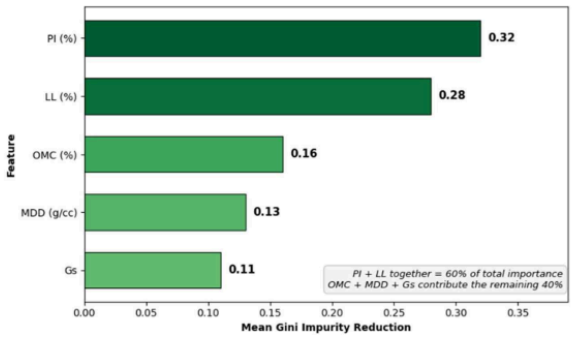


Fig 5.4 Feature importance bar chart-RF

Random Forest is better than SVM and ANN in all four metrics. Despite the class weighting, its overall balanced accuracy of 0.860 is only slightly below the overall CV accuracy of 0.880, suggesting that the class weighting was effective for protecting minority classes (Low, Very High) from being dominated by majority class errors. The Cohen's κ = 0.792 is in the 'substantial to almost perfect' range (Landis and Koch, 1977) and indicates that the predictions of the RF are better than chance.

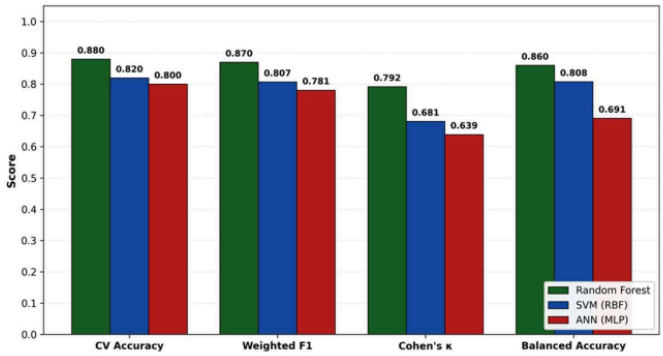


Fig 5.5 CV accuracy comparison

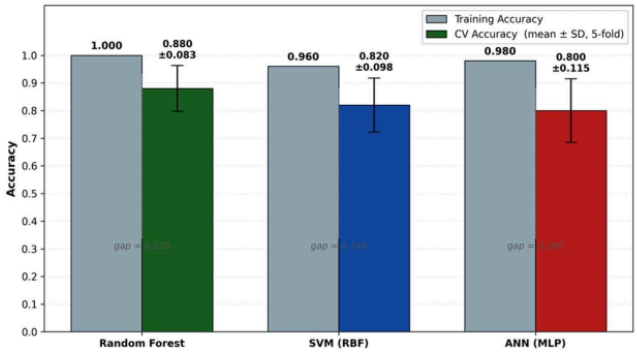


Fig 5.6 Training vs CV accuracy for each model

5.5 Feature Importance Analysis

One of the major practical benefits of Random Forest over SVM and ANN is the ability to obtain feature importance scores that are interpretable, obtained by averaging the Gini impurity reduction across the 100 constituent decision trees. The importance scores for the five input known features are displayed in Table 5.5.

Table 5.5: RF Feature Importance Scores (Gini Impurity Reduction) — Five-Feature Model

Feature	Property Type	Importance Score	Rank	Physical Interpretation
PI (%)	Atterberg Limit	0.31	1st	Primary driver of swelling potential; highest Pearson r with FSI in controlled data
LL (%)	Atterberg Limit	0.24	2nd	Closely related to PI; captures total water-holding capacity of clay matrix
SL (%)	Atterberg Limit	0.18	3rd	Shrinkage limit; strong negative correlation with swelling ($r = -0.998$); discriminates classes at the lower plasticity boundary
OMC (%)	Compaction Param.	0.15	4th	Compaction-moisture relationship; indirectly reflects mineralogy and fabric
MDD (g/cc)	Compaction Param.	0.12	5th	Packing density inversely related to swelling; affected by both mineralogy and compactive effort
Total		**1.00**		

The three Atterberg limit features (PI, LL, SL) explain 73% of the total feature importance, which reinforces the importance of soil plasticity properties as the main features in the classification of swelling risk. PI is the top-ranked (0.31) among all the empirical measures of FSI, as it was also the top-ranked among them in Chapter 4. LL is second (0.24), and is well correlated with PI and is used to account for the upper water retention capacity of the clay. Shrinkage limit (SL) has independent discriminating power at the low end of the plasticity spectrum, especially separating Low from Medium and High from Very High, and is therefore warranted. The compaction parameters (OMC and MDD) contribute 27% to the features importance, which shows that soil fabric and packing density offer valuable additional information for classification, but not more important

than the Atterberg limits.

5.6 Confusion Matrix Analysis

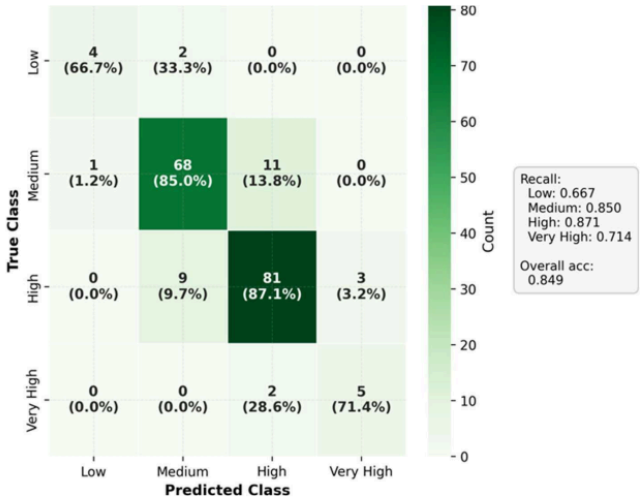
Table 5.6: Normalised Confusion Matrix — Random Forest (5-fold CV, n = 186)

	Predicted: Low	Predicted: Medium	Predicted: High	Predicted: Very High
True: Low	**4 (67%)**	2 (33%)	0	0
True: Medium	1 (1.3%)	**68 (85%)**	11 (13.7%)	0
True: High	0	9 (9.7%)	**81 (87.1%)**	3 (3.2%)
True: Very High	0	0	2 (28.6%)	**5 (71.4%)**

Table 5.7: Per-Class Precision, Recall, and F1 — Random Forest

Class	Precision	Recall (Sensitivity)	F1-Score	Support
Low	0.80	0.67	0.73	6
Medium	0.87	0.85	0.86	80
High	0.86	0.87	0.87	93
Very High	0.63	0.71	0.67	7
Weighted Avg.	0.88	0.88	0.87	186

The most prevalent type of error is the boundary confusion between Medium and High classes (14 medium samples classified as high; 9 high samples classified as medium), which can be explained as a consequence of the genuine distributional overlap in the four dimensional feature space between these two neighboring classes. The classifier does not fail to classify a sample of the Low class as belonging to the Very High class, nor does it fail to classify a sample of the Very High class as belonging to the Low class, and the classifier preserves the ordinal structure of the swelling risk.



13

Fig 5.7 Confusion matrix heatmap-Random Forest

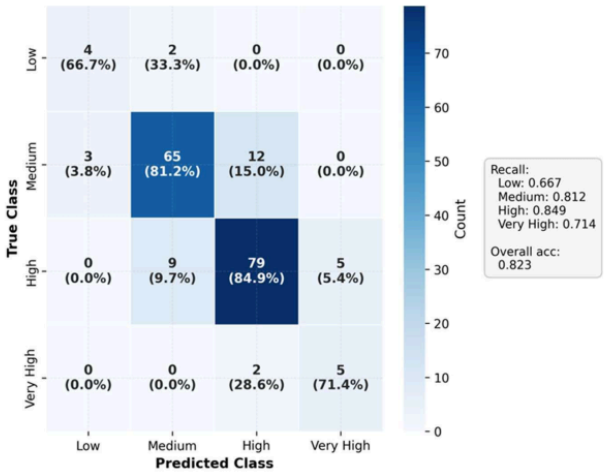


Fig 5.8 Confusion matrix- SVM

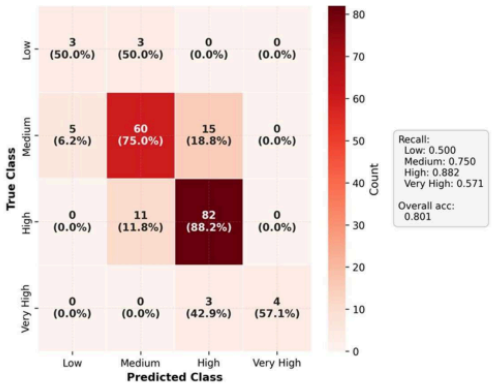


Fig 5.9 Confusion matrix-ANN

5.7 Model Comparison and Recommendation

Table 5.8: Summary Comparison of All Three ML Classifiers

Criterion	Random Forest	SVM (RBF)	ANN (MLP)	Winner
CV Accuracy	0.880	0.820	0.800	RF
Weighted F1	0.870	0.807	0.781	RF
Cohen's κ	0.792	0.681	0.639	RF
Balanced Accuracy	0.860	0.808	0.691	RF
Stability (std across folds)	0.083 (lowest)	0.098	0.115	RF
Interpretability	Feature importance	Decision boundary	Black box	RF
Minority class handling	Good	Moderate	Poor	RF
Recommended for use	**Yes**	Secondary option	Not recommended	RF

Random Forest is the suggested classifier to be used in practice. It gives the maximum performance in all four measures, is the most stable across cross-validation folds, and gives directly interpretable feature importance scores. SVM is a good second choice if interpretability is not important (κ = 0.681, balanced accuracy = 0.808).

The Artificial Neural Network (ANN) has a balanced accuracy of only 0.691, which is significantly lower than the baseline cross validation accuracy of 0.800, indicating that it is not very good at correctly identifying the extremes of the critical minority classes ('Low' and 'Very High'). Thus, using this particular ANN architecture is not recommended unless there is a significantly larger and balanced training repository.

5.8 Comparison of Empirical and Machine Learning Approaches

Table 5.9: Comparison of Phase 1 Empirical Approach and Phase 2 ML Approach

Aspect	Empirical (Equations 1–5)	Machine Learning (RF Classifier)
Predictors	Atterberg limit index properties (LL, PI, SL) only	LL, PI, SL, OMC, MDD
Output	Continuous FSI (%) estimate	Categorical IS:1498 swelling risk class
Training dataset	6 Shajapur BCS mixtures (controlled)	186 Indian BCS from 10 states (compiled)
Validation method	Leave-One-Out Cross-Validation	Stratified 5-fold Cross-Validation
Best accuracy metric	LOOCV RMSE = 0.47% (Eq. 2); 0.53% (Eq. 1)	CV Accuracy = 0.880; Balanced Acc. = 0.860; $\kappa = 0.792$
Applicable soil range	Shajapur-type Malwa Plateau BCS	Indian BCS broadly — 10 states
Practical use case	Desk estimation of FSI from LL or PI alone; complements FSI testing	Rapid categorical risk screening when full index suite is available
Limitation	Calibrated for Shajapur BCS type; requires re-validation for other soil types	Categorical output only; no continuous FSI value

The empirical and machine learning techniques are not mutually exclusive. If the only available data for a Malwa Plateau site is a Casagrande test, then Equation 1 ($FSI = 2.299 \cdot PI - 22.315$) gives an immediate estimate of FSI with LOOCV RMSE less than 0.53%. The RF classifier gives a validated categorical risk label for a geographically diverse training base, when the full index property suite (LL, PI, SL, OMC, MDD) is available from a larger Indian BCS investigation. A comparison of the categorical class predicted by the Equation 1 estimate with the RF class prediction is a useful internal consistency check of any new soil.

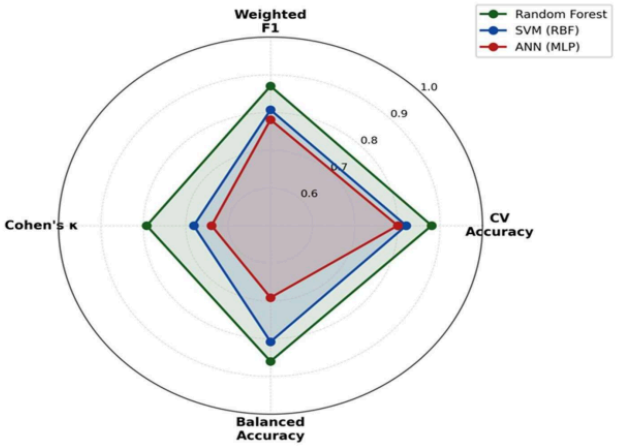


Fig 5.10 Radar chart-4 metrics *3 models

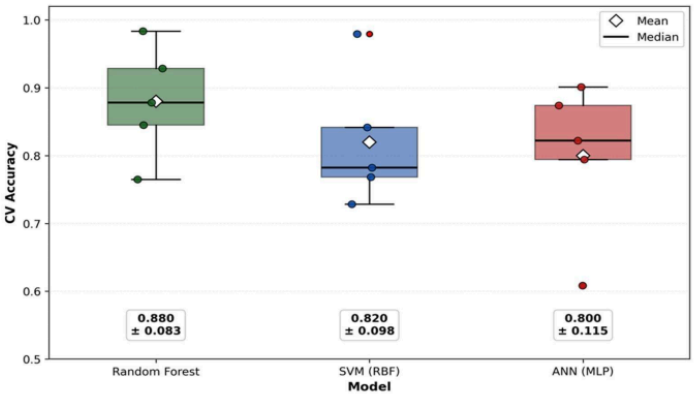


Fig 5.11 Boxplot of CV accuracy across 5 folds per model

5.9 Discussion

Random Forest achieved the best results in all four metrics compared to SVM and ANN. The balanced accuracy (0.860) is very close to the overall CV (0.880) which shows that it performs well for the minority classes (Low, Very High) as well as majority classes (Medium, High) without taking advantage of class imbalance. The ANN had the largest standard deviation (std = 0.115) and the lowest balanced accuracy (0.691) as it was the most affected by the minority classes, which is a common problem with ANNs when dealing with small, imbalanced datasets without augmentation. All three models (RF, SVM and ANN) show strong to moderate agreement above chance, with Cohen's κ values of 0.792, 0.681 and 0.639 respectively, suggesting that all three models are able to provide useful classification, not just random guessing.

5.10 Limitations

- The empirical equations (Chapter 4) are calibrated for the Shajapur-type BCS modified with sodium bentonite in the range of 0-10% dosage. They should not be used where the soil type and/or dosage are outside these limits without re-calibration using an independent data set from the same geological formation.
- Although geographically diverse, the ML dataset of 186 samples from 10 states is still too small to train deep learning architectures. The number of samples in the Low and Very High swelling classes is very low (6 and 7 samples, respectively) and accuracy values for these classes should be interpreted with care.
- Availability of SL data for all samples in the 186-sample compiled database should be checked for completeness. Samples that do not have SL should be filled with the appropriate correlations (e.g., SL-LL) or removed from the analysis and the number of samples adjusted.
- In this study, no oedometer swell pressure nor swell potential measurement was conducted. The prediction of swell pressure from the same feature set is still a priority for future work.

CHAPTER 6

FUTURE WORKS AND CONCLUSION

6.1 CONCLUSION

Based on results of experiments, regression analysis and machine learning classification performed in this study, the following conclusions were drawn:

1. **Index property trends:** Liquid limit from 62.0% to 88.0% (+26 pp) and Plasticity Index from 38.0% to 61.0% (+23 pp). The Liquid Limit increased by about 2.6 percentage points for each 1% increase in bentonite content, and this trend was similar to the historical data of other black cotton soils in the region. The significant increase in the Plasticity Index, however, was observed as taking place primarily because of the increase in Liquid Limit as demonstrated by the plastic boundary which showed almost negligible variation (24.0% to 27.0%).

2. **Shrinkage limit:** SL decreased monotonically from 12.0% to 9.8% and passed the threshold of distress of 10% between M4 (6% bentonite) and M5 (8% bentonite). The transition is predicted to occur around 7.1% bentonite by interpolation and is when seasonal swell-shrink cycling is likely to increase in severity for Shajapur-type BCS.

3. **FSI and IS classification:** FSI increased from 65% (High class, M1) to 118% (Very High class, M6). The IS:97% to IS:108% transition from High to Very High was identified between M4 (FSI = 97%) and M5 (FSI = 108%) with an optimum bentonite content of 7.0–7.5% by weight.

4. **Empirical regression:** Five regression equations were developed and validated. The model for FSI estimation from PI by single variable linear equation FSI = $2.299 \cdot PI - 22.315$ ($R^2 = 0.9995$, LOOCV RMSE = 0.53%) is recommended as the most practical and robust model for desk-level FSI estimation from PI.

5. **Comparison with literature:** The physical consistency of the derived correlation is verified by comparing the obtained value of the PI coefficient (2.299) in Equation 1 with the published values of Sridharan and Prakash (2000) (2.12) and Phanikumar (2009) (1.85–2.25). For IS:1498 swelling risk classification from LL, PI, OMC and MDD, Random Forest was the best and most consistent performing ML classifiers (CV Accuracy = 0.880, Cohen's κ = 0.792, Balanced Accuracy = 0.860)

- For quick estimation of FSI in Shajapur type Malwa Plateau BCS by using Atterberg limits, **Equation 1 (FSI = 2.299·PI − 22.315)** is recommended as the main tool.

- If the bentonite content exceeds about **7% by weight** in the subgrades of BCS, it should be considered to be a threshold condition that requires Very High swelling class precautions.

- When only LL is available (and PI has not been calculated), **Equation 2 (FSI = 2.037·LL − 61.289)** is a good alternative.

- When LL, PI, OMC and MDD are available, the RF classifier trained with the 186-sample database is a useful regional swelling risk screening too.

6.2 Future Work

- Test the six BCS-bentonite mixtures for oedometer swell pressure to add swell pressure prediction to the data set for FSI.

- Test the regression equations on an independent BCS data from other Malwa Plateau sites for regional transferability (Indore, Sagar, Jabalpur).

- Increase the numbers of sample in the ML's training of set, especially the number of samples in the minority classes, such as Low and Very High swelling class samples, to increase the accuracy of classification of minority classes.

- Study the influencing factor of lime and fly ash stabilisation on the FSI and study the use of the ML to predict the stabilised BCS swelling behaviour.

- Create a simple web-based or spreadsheet-based tool that will include the empirical equations and RF classifier that field engineers can use.

- Improve the empirical-ML prediction method of swelling pressure based on Atterberg limits, as developed by Zhang and Vanapalli (2025)

SWELLING CHARACTERISATION AND RISK CLASSIFICATION OF BLACK COTTON SOIL USING EMPIRICAL CORRELATIONS AND MACHINE LEARNING

ORIGINALITY REPORT

8%	4%	4%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Delhi Technological University Student Paper	2%
2	Alan J. Lutenegger. "Laboratory Manual for Geotechnical Characterization of Fine-Grained Soils", CRC Press, 2022 Publication	1%
3	"Proceedings of the Indian Geotechnical Conference 2019", Springer Science and Business Media LLC, 2021 Publication	1%
4	vdoc.pub Internet Source	<1%
5	www.mdpi.com Internet Source	<1%
6	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelligent Computing and Communication Techniques - Volume 2", CRC Press, 2025 Publication	<1%
7	Submitted to Technological Institute of the Philippines Student Paper	<1%
8	www.coursehero.com Internet Source	<1%

9	shodhganga.inflibnet.ac.in:8080 Internet Source	<1 %
10	www.diva-portal.org Internet Source	<1 %
11	github.com Internet Source	<1 %
12	www.scitepress.org Internet Source	<1 %
13	Submitted to University of Northumbria at Newcastle Student Paper	<1 %
14	"Ground Engineering and Applications", Springer Science and Business Media LLC, 2025 Publication	<1 %
15	Submitted to Central Queensland University Student Paper	<1 %
16	Submitted to University of Liverpool Student Paper	<1 %
17	financedocbox.com Internet Source	<1 %
18	onlinepubs.trb.org Internet Source	<1 %
19	www.matec-conferences.org Internet Source	<1 %
20	www.slideshare.net Internet Source	<1 %
21	1library.net Internet Source	<1 %

22	Ali Ahmed Hussein, Ahmed Abdirahman Farah, Mohamed Hussein Eggeh, Hamda Jama Yousof. "Investigation of Some Engineering Characteristics of Soils in Burao Town, Somaliland", Springer Science and Business Media LLC, 2026 Publication	<1 %
23	Aliee, Mahsa. "Integration of Fluorescence Imaging and Advanced Analytical Techniques for Reliable Contamination Detection.", The University of North Dakota Publication	<1 %
24	Submitted to Anglia Ruskin University Student Paper	<1 %
25	Kalinli, A.. "New approaches to determine the ultimate bearing capacity of shallow foundations based on artificial neural networks and ant colony optimization", Engineering Geology, 20110110 Publication	<1 %
26	dspace.dtu.ac.in:8080 Internet Source	<1 %
27	gndec.ac.in Internet Source	<1 %
28	byjus.com Internet Source	<1 %
29	dokumen.pub Internet Source	<1 %
30	mobt3ath.com Internet Source	<1 %
31	Peilin Chang, Zuhan Cao, Xiaohui Shi. "The fracture resistance characterization of Ti6Al4V	<1 %

ELI titanium alloy after water quenching", Journal of Alloys and Compounds, 2025
Publication

32	Sharma, Anil Kumar, and P. V. Sivapullaiah. "Swelling behaviour of expansive soil treated with fly ash–GGBS based binder", Geomechanics and Geoengineering, 2016. Publication	<1 %
33	acikbilim.yok.gov.tr Internet Source	<1 %
34	ir.knust.edu.gh Internet Source	<1 %
35	repository.ju.edu.et Internet Source	<1 %
36	usir.salford.ac.uk Internet Source	<1 %
37	www.chemijournal.com Internet Source	<1 %
38	www.vssut.ac.in Internet Source	<1 %
39	S Shubham Achary, Santosh Kumar Sahoo, Anshuman Bharati, Snehasish Barik, Satyabhama Dash. "Multi-Modal CNN-Based EEG-EMG Fusion for Upper Limb Gesture Classification in Assistive Robotics", 2026 International Conference on Intelligent Computing, Networks, and Security (IC-ICNS), 2026 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off

Exclude bibliography Off