SSRG International Journal of Civil Engineering
ISSN: 2348-8352/ https://doi.org/10.14445/23488352/IJCE-V12I4P105

Volume 12 Issue 4, 48-66, April 2025
© 2025 Seventh Sense Research Group®

Review Article

Prediction of Compression Index of Soil - A Perspective Review

Bishwajoty Paul[1], Dipika Devi[2], Santosh Kumar Tamang[3]

[1,2]*Civil Engineering, North Eastern Regional Institute of Science and Technology, Arunachal Pradesh, India.*
[3] *Mechanical Engineering, North Eastern Regional Institute of Science and Technology, Arunachal Pradesh, India.*

[1]*Corresponding Author : bishso.paul@gmail.com*

Received: 03 February 2025 Revised: 05 March 2025 Accepted: 06 April 2025 Published: 30 April 2025

Abstract - *The settlement analysis of structures built over soil masses is integral to design, ensuring both stability and long-term performance. A critical parameter influencing settlement behaviour is the soil's Compression Index (C_c), which provides insights into soil compressibility and potential risk factors essential for informed structural design. Traditionally, the estimation of C_c relies on standardized laboratory procedures (e.g., Bureau of Indian Standards), which, while accurate, are often costly, labour-intensive, and time-consuming. To address these limitations, researchers have explored correlations between C_c and easily measurable index properties of soil, such as Atterberg limits and others. Through exploring these index properties, predictive models based on Linear Regression and Computer-aided Learning algorithms have emerged as efficient alternatives for C_c estimation. This review provides a comprehensive perspective on current methodologies for C_c prediction, highlighting that liquid limit, in-situ void ratio (e_o), and natural moisture content exhibit a significant correlation with C_c estimates across both linear and machine learning models. The findings from this study underscore the potential for data-driven approaches to streamline soil compressibility assessments, offering reliable and time-efficient predictions essential for geotechnical design practices. This paper also shows that predictive models and simple correlations can be developed by using an extended range of index properties obtained from bibliographies, self-generated experimental data or other project sources.*

Keywords - *Compression Index, Computer aided Learning, Index properties, Linear Regression, Machine Learning.*

1. Introduction

1.1. Compression Index

The soil is a naturally formed, complex, porous material consisting of solid particles. The voids or pores within soil may contain water, air, or a combination of both. As per the Indian Standard Classification (ISC) System, soil is classified into three main types, one of which is fine-grained soil, which includes clay and silt and is characterized by low permeability. Due to this low permeability, water is expelled gradually when subjected to load, resulting in a slow settlement process. This gradual decrease in volume caused by the expulsion of water is known as consolidation, and it is an important engineering property that must be thoroughly analyzed for accurate settlement assessment. One key concept in consolidation for determining settlement is the C_c. It is a measure of the soil's compressibility; higher C_c indicates greater potential for settlement. Since this compressibility parameter is used in the calculation of settlement, the soil undergoes upon application of load it has an essential role in designing the foundation of various structures and eventually affecting its cost because the allowable bearing capacity for foundation design is based on two primary criteria: shear failure and settlement. So, an inaccurate determination of C_c will lead to an incorrect settlement estimate, resulting in structural damage. The C_c of soil is equal to the slope as determined from the linear section of the plot between e_o vs logarithmic of effective stress. This plot is made based on the experimental test result accomplished by the oedometer test, which is sometimes referred to as the consolidation test. The oedometer test was developed by Jean Frontard who was a French Civil Engineer, in the year 1910 for soil characteristics in slope failure of an earthen dam.

1.2. Problem Statement

The conventional one-dimensional consolidation test, which is performed as per the procedure laid in BIS code IS:2720 part 15 (1986), is a time-consuming test and generally takes about (7-10) days. The test, in general, involves the application of step-by-step loading in an incremental manner on the soil sample in a period of one day and, thereby, measuring its corresponding settlements and, thereafter, subsequent unloading, which takes more than a week time. In addition to this disadvantage related to time-consumption and cost involvement, another disadvantage in the estimation of C_c is that since its determination involves plotting a graph between void ratio and effective stress in a logarithmic scale,

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

its accuracy is highly dependent upon the persons' experience. Another relatively modern method to determine C_c is the Constant Rate Strain (CRS) test introduced by Hamilton and Crawford in 1959, which is faster than the traditional Oedometer test but has limitations. It is mostly applicable for soft soils and requires a very careful selection of strain rate to avoid misinterpretation in results. This encourages numerous researchers to investigate the possibilities of establishing the correlation between C_c and various physical as well as other properties of soil, which are relatively simple and easy to estimate, which in turn would reduce the necessity of large-scale laboratory testing and will save time as well as resources. The majority of these established correlations were developed using Linear regression, but in recent years, several computer-aided learning techniques have also gained much popularity regarding this aspect and have been used in an attempt to anticipate the C_c of soil. With the usage of various computer-aided learning, several predictive models that are capable of understanding complex relationships among multidimensional data, both linear and nonlinear, are developed that relate C_c to soil index properties. Out of these established equations and predictive models, some are considered to be suitable to all types of soils, while the rest are restricted to certain specific categories of soils. Although these correlation and predictive models are not yet fully reliable enough to eliminate the one-dimensional oedometer test for C_c estimate, still these predictive models can give quick and efficient predictions after training using local datasets and can be of significant help for preliminary estimation of C_c for local conditions. This paper aims to emphasize previously conducted research done for the prediction of C_c using various correlation and computer-aided learning techniques. The paper concludes with critical findings which might be helpful for future investigations in this field.

2. Review of Various Literatures

2.1. Review of Literature on Linear Regression

One of the earliest studies on (C_c) was put forward by Skempton (1944), in which a relationship with liquid limit (w_{LL}) was recommended. Helenelund (1951), Moran et al. (1958), and Koppula (1981) also proposed a relationship with Natural moisture content (w_n). Nishida (1956), Hough (1957), Sowers (1970) and Burland (1990) estimated C_c using the initial void ratio (e_o). Cozzolino (1961) proposed two correlations, one with w_{LL} and the other with e_o. Terzaghi and Peck (1967), Mayne (1980), and Burghignoli et al. (1985) also proposed a correlation equation with w_{LL} for which Mayne's correlation was based on test data from 96 different soil samples which were collected from various published papers. Azzouz et al. (1976) provided two sets of equations, one relating with w_{LL} and the other with w_n. Herrero (1983) provided various equations correlating with w_n, e_o, and G_s for all clays. Koppula (1981), based on 134 cohesive soil test results, used regression techniques such as OLS and Ridge regression for obtaining linear models and a correlation with w_n was proposed. Rendon-Herrero (1983) used data obtained from the Marine Geomechanics Laboratory, University of Rhode Island, to develop correlation equations using different regressions and found that e_o has the most influence.

Nagaraj and Murthy (1983) and Nagaraj et al. (1995) also proposed equations relating to the void ratio, which is obtained at the liquid limit of soil (e_L). Nagaraj et al. (1986) correlated e_L & e with C_c and proposed various equations. Bowels (1989) proposed three empirical equations for Chicago clays, all clays moderately over consolidated and organic silts or clay; for the first two categories of soil was used as the influencing parameter and for the last soil type wn was considered as the influencing parameter. Abdrabbo and Mahmoud (1990) established correlations between w_{LL}, w_n, eo with compressibility parameters of Egyptian clays and concluded that with an increase in w_{LL}, w_n and in-situ void ratio, C_c also increases. Hirata et al. (1990) performed multiple regression analyses to establish the relationship between engineering and index properties of natural and artificial mixed soil. One of the engineering properties in the analysis was C_c, and three different equations were provided correlating C_c with w_{LL} and void ratio. carried out regression analysis to establish the correlation between C_c and Void ratio.

Tsuchida (1991) performed regression analysis over a data set of 200 and 150 consolidated test results of undisturbed clay collected from bay regions of Tokyo and Osaka, Japan, respectively and proposed a correlation between the w_{LL} and C_c. Al-Khafaji and Andersland (1992), based on 72 data points collected from published studies, developed multiple linear regression for the prediction of C_c in which a correlation was established relating C_c with e_o and w_{LL}; it was also found that in-situ void ratio had the most influence for low to medium plasticity soil and for high plasticity soil both had significant influence. Koumoto and Park (1998a, 1998b) suggested correlating C_c with (W_o-W_p) and (e_o-e_p) based on test results for 34 samples of soil. Sridharan & Nagaraj (2000), based on the experimental test result of 10 soil samples using linear regression, proposed various correlations between C_c with w_{LL}, Plasticity Index (I_p), Shrinkage index and concluded that shrinkage index is the most influencing parameter followed by I_p and w_{LL}. Lav & Ansal (2001), using a database of 300 soil sample test results which were taken from different construction sites throughout Turkey, attempted to establish a relationship for consolidation properties and various index properties out of which w_n, w_{LL}, e_o and dry unit weight were considered to yield sufficient good correlation, a set of correlation were also provided for soil subgroup categorized as low plasticity, over-consolidated, normally consolidated and also liquidity index. Yoon et al. (2004), using a data set of 1200 marine clay soil samples from various areas in Korea, namely the south, East and West coasts, performed single and multiple linear regression analyses to establish correlation equations between C_c with other properties of soil such as w_{LL}, w_n, e_o I_p and dry density out of which it was found that e_o and

w_{LL} are having a major effect on the C_c for all the coast. Koumoto and Park (2004) established two correlation equations for C_c, first for remolded clay using a data set of 66 soil samples and second for Ariake clay with a data set of 83 soil samples using porosity (n_0) as the dependent variable for both. Nath and DeDalal (2004) carried out several consolidation tests on mixed soil samples to produce a correlation among the C_c and I_p of various clays. Solanki and Desai (2008) developed a correlation for alluvial deposits of soil collected from 10 zones, out of which 6 zones were from Surat and the remaining from Suda India; based on the results of statistical analysis it came to a conclusion that soil plasticity characteristics have significant influence for determination of C_c.

Vinod and Bindu (2010) proposed a correlation for Clay soils which were collected from 18 sites of Alappuzha, Kottayam and Pathanamthitta districts situated in Kerala, India, and it was found that the Shrinkage index was the most influencing. Al-Kahdaar et al. (2010) used properties of soils collected from 40 boreholes located in Ammarah City, Iraq and performed regression analysis with a single variable using MS-office software to propose an empirical equation for C_c. Slamet Widodo and Abdelazim Ibrahim (2012) correlate physical properties such as e_o, w_n and w_{LL} of Pontianak soil to obtain C_c of the soil using a data set of 20 samples collected from 10 boreholes.

Amardeep Singh and Shahid Noor (2012) proposed an empirical equation for the determination of C_c using soil index properties of 23 soil samples collected from different hydropower projects in India. Sari & Firmansyah (2013) derived an empirical formula for the determination of C_c using soil index properties using the results of 466 samples of soil collected from 77 borehole locations spread in the Surabaya area, Indonesia. B. Tiwari and B. Ajmera (2012) proposed various correlations for C_c, using the results of 55 artificially prepared soil samples mixing minerals such as quartz, kaolinite, illite, and montmorillonite at various proportions. Abbasi et al. (2012) proposed an empirical equation for C_c estimate using soil index properties. The soil was collected from various locations in Iran, including Khozestan, Ardabil, Ghazwin, Eshafan and Tehran provinces. Akayuli and Ofosu (2013) established an equation relating Cc and soil index properties using the data of soil investigations conducted which was performed by Research Institute (BRRI) with 90 laboratory results. Bryan et al. (2014) provided a new correlation for the assessment of C_c for Irish soils using a soil data set collected from various published literature. Dway et al. (2014) estimate C_c by forming equations relating soil index properties using samples of soils obtained from 3 locations in Mandalay. Pundreek Dwivedi et al. (2015) based on the data of 23 clay samples of soil, out of which 16 samples were from Bhopal City situated in Madhya Pradesh, India and the remaining were artificial soil mixed with bentonite used Linear regression single and multiple to a developed correlation between C_c and other properties of soil. Güllü et al. (2016) developed correlations for the C_c with a database consisting of 69 data sets from Baghdad City. Kumar K et al. (2016) used a data set of six soil samples, which were gathered from areas near Hyderabad, to develop correlations between Cc, Cv, and C_s. Zaman et al. (2017), using the results of soil samples collected from a major design project of the expressway in Dhaka-Chittagong, proposed a correlation for C_c; a total of 14 undisturbed samples were used. Kootahi and Moradi (2017) proposed correlations for estimation of C_c using data of 1000 soil samples which were collected from 170 different locations spread all over the world. Kok Shien Ng et al. (2018), based on the results of five remolded cohesive soil samples, correlated index properties with C_c and C_v. Salih et al. (2020) proposed new empirical relationships for various geotechnical properties with index properties using a test data set of 170 soil samples collected from Barika and other locations in Sulaimani Governorate, Iraq. An outline of various correlation equations of the past studies is provided in Table 1.

The above-mentioned studies for the prediction of C_c of soil were more or less based on linear regression, out of which most of them were single and few of them multiple. But as from the past research, it is well understood that the C_c of soil is affected by multiple parameters and not only a single parameter, giving a better alternative for accurate prediction of compression index led to the application of multiple regression or even better using machine learning algorithm such as ANN, SVM, RF etc.

2.2. Review of Literature on Machine Learning Algorithm

Kolay, P. K. et al. (2008) developed various predictive models using ANN for the determination of C_c of soil using the experimental results of soil samples, which were gathered from 200 borehole locations spread across various cities in Malaysia. For the model, 13 input parameters were used. The Levenberg-Marquardt algorithm (lm) and BFGS Quasi-Newton (bfg) showed better accuracy than the other algorithm. Hyun Il Park and Seung Rae Lee (2011) used ANN to develop a predictive model based on Feed-Forward Back Propagation Algorithms (FFBPA) using a data set of the soils of the Republic of Korea.

For this purpose, 947 soil data were collected, out of which 852 were used for training the models and the rest, 95 were used to test the models. Based on the result, it was suggested that the model based on the natural water content of soil showed better results in predicting the Cc. P.K. Kolay et al. (2011) used a total of 700 numbers of borehole data of undisturbed soil samples collected from Jabatan Kerja Raya (JKR), Sarawak, Malaysia, for developing an ANN predictive model. The model includes a 3-layer FFBPA. A total of eight training algorithms were used, out of which the Resilient Backpropagation algorithm showed the best prediction. V. Phani Kumar et al. (2011) used MLP-ANN with FFBPA to develop various predictive models using a total of 68 soil sample data from different parts of Chitter district, India. The models used the logsig activation function, and based on the performance results, the model consisting of 8 neurons in the concealed layer was considered to give the best output. Farzaneh Namdarvand et al. (2013) experimented on 100 soil specimens collected from Ahvaz, a city in Iran, to develop various predictive models using MLP-ANN and multiple regression.

Table 1. Summary of Various correlations available for compression index estimate using Linear Regression

Reference	Correlation	Applicability	Input variable
Skempton (1944)	$C_c = 0.007 \times (w_{LL} - 10)$	Remolded cohesive soil	Liquid Limit
Cozzolino (1961)	$C_c = 0.0046 \times (w_{LL} - 9)$	Brazil clays soil	
Terzaghi and Peck (1967)	$C_c = 0.009 \times (w_{LL} - 10)$	NCC soil	
Azzouz et al. (1976)	$C_c = 0.006 \times (w_{LL} - 9)$	Clay soil with w_{LL} < 100 %	
Mayne (1980)	$C_c = (w_{LL} - 13)/109$	All types of clay soil	
Burghignoli et al. (1985)	$C_c = 0.008 \times (w_{LL} - 10)\, w_{LL}$	Italian soft clays	
Bowles (1989)	$C_c = 0.0046 \times (w_{LL} - 9)$	Brazilian clays (Moderately over-consolidated)	
Abdrabbo & Mahmoud (1990)	$C_c = 0.0063 \times (w_{LL} - 10)$	Egyptian clay (10 % < w_{LL} < 110 %)	
Hirata et al. (1990)	$C_c = 0.010 \times w_{LL} + 0.063$	Natural soils and Artificial soil of Osaka and Hyogo (clay content > 20 %)	
Tsuchida (1991)	$C_c = 0.009 \times (w_{LL} - 8)$	Osaka Bay clay	
Sridharan & Nagaraj (2000)	$C_c = 0.008 \times (w_{LL} - 12)$	Remoulded clays (37 % < w_{LL} < 74 %)	
Lav & Ansal (2001)	$C_c = 0.006 \times (w_{LL} + 1)$	All soils	
Yoon et al. (2004)	$C_c = 0.012 \times (w_{LL} + 16.4)$	South Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.011 \times (w_{LL} - 6.36)$	East Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.01 \times (w_{LL} - 10.9)$	West Coast (Korea clay)	
Solanki and Desai (2008)	$C_c = 0.0061 \times (w_{LL} - 0.0024)$	Alluvial deposits, Surat, India	
Vinod and Bindu (2010)	$C_c = 0.0055 \times (w_{LL} - 1.8364)$	Kuttanad clay, Kerela, India (70.8% < w_{LL} < 276.3 %)	
Al-Kahdaar et al. (2010)	$C_c = 0.00556\, w_{LL}$	Al-Ammarah silty clay soil (Iraq)	
Slamet and Abdelazim (2012)	$C_c = 0.01706 \times (w_{LL} - 1.29)$	Pontianak clay soil	
B. Tiwari and B. Ajmera (2012)	$C_c = 0.0075 \times (w_{LL})$	Artificial Soil mix (10 % < w_{LL} < 470 %)	
B. Tiwari and B. Ajmera (2012)	$C_c = 0.012 \times (w_{LL})$	Artificial Soil mix (10 % < w_{LL} < 470 %)	
Abbasi et al. (2012)	$C_c = 0.007\, w_{LL} - 0.043$	Fine-Grained soil of Iran (w_{LL} < 75%)	
Akayuli and Ofosu (2013)	$C_c = 0.004 w_{LL} - 0.03$	Weathered Birimian phyllites samples	
Bryan et al. (2014)	$C_c = 0.0118 \times (w_{LL} - 20.7)$	Irish soft soils	
Dway et al. (2014)	$C_c = 0.0027 \times w_{LL} + 0.1994$	Lean to Fat Clay with Low to High compressibility	
Pundreek Dwivedi et al. (2016)	$C_c = 0.0067 \times (w_{LL}) - 0.0364$	Natural and Artificial soils	
Güllü et al. (2016)	$C_c = 0.00454 w_{LL} - 0.01246$	Clay of Baghdad City	
Kumar K et al. (2016)	$C_c = 0.001 \times (w_{LL}) - 0.013$	Hyderabad clay with High compressibility	

Zaman et al. (2017)	$C_c = 0.01 \times (w_{LL} - 13.61)$	Dhaka-Chittagong clay	
Kootahi and Moradi (2017)	$C_c = -0.096 + 0.012 w_{LL}$	Marine fine-grained soils	
Kok Shien Ng et al. (2018)	$C_c = 0.0062 w_{LL} + 0.0165$	Soil with Low to Intermediate Plasticity (I_p =8 to 18)	
Salih et al. (2020)	$C_c = -0.0022\, w_{LL} + 0.2795$	Fine-grained soil with High and Low Compressibility in Iraq	
Helenelund (1951)	$C_c = 0.85\sqrt{(w_n/100)^3}$	Finnish muds and clays	Moisture content
Moran et al. (1958)	$C_c = 0.0115 \times w_n$	Organic soils, peat, organic silt, and clay	
Azzouz et al. (1976)	$C_c = 0.01 \times (w_n - 5)$	All types of clays	
Koppula (1981)	$C_c = 0.01 w_n$	Chicago and Alberta Clays, (NC soil, S_t < 1.5)	
Herrero (1983)	$C_c = 0.01 \times (w_n - 7.549)$	All types of clay soil	
Abdrabbo & Mahmoud (1990)	$C_c = 0.0066 \times w_n$	Egypt clay soil with 20 % < w_n < 140 %	
Lav & Ansal (2001)	$\ln C_c = 1.235 \ln w_n - 5.65$	All soils	
Yoon et al. (2004)	$C_c = 0.013 \times (w_n - 3.85)$	South Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.01 \times (w_n + 2.83)$	East Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.011 \times (w_n - 11.22)$	West Coast (Korea clay)	
Solanki and Desai (2008)	$C_c = 0.0091 w_n + 0.0522$	Alluvial deposits, Surat, India	
Vinod and Bindu (2010)	$C_c = 0.0072 \times (w_n - 12.625)$	Remoulded Kuttanad clay, Kerela, India	
Al-Kahdaar et al. (2010)	$C_c = 0.0092\, w_n$	Al-Ammarah silty clay soil (Iraq)	
Slamet and Abdelazim (2012)	$C_c = 0.0102\, (w_n + 11.57)$	Pontianak clay soil	
Abbasi et al. (2012)	$C_c = 0.008\, w_n - 0.044$	Fine-grained soil of Iran (w_{LL}< 75%)	
Sari & Firmansyah (2013)	$C_c = 0.0143\, w_n - 0.0165$	Soft clay in Indonesia (w_{LL} = 0 to 100 %)	
Sari & Firmansyah (2013)	$C_c = 0.0179\, w_n - 0.1005$	Soft clay in Indonesia (w_{LL} =30 to 50 %)	
Sari & Firmansyah (2013)	$C_c = 0.0137 \times w_n + 0.0034$	Soft clay in Indonesia (w_{LL} = 50 to 70 %)	
Akayuli and Ofosu (2013)	$C_c = 0.002 \times w_n + 0.14$	Weathered Birimian phyllites samples	
Bryan et al. (2014)	$C_c = 0.014 \times (w_n - 22.7)$	Irish soft soils (35% < w_n < 150%)	
Dway et al. (2014)	$C_c = 0.01 \times w_n + 0.027$	Lean to Fat Clay with Low to High compressibility	
Güllü et al. (2016)	$C_c = 0.00553 \times w_n + 0.05321$	Clay of Baghdad City	
Zaman et al. (2017)	$C_c = 0.0158 \times w_n - 0.179$	Dhaka-Chittagong clay	
Kootahi and Moradi (2017)	$C_c = -0.093 + 0.012\, w_n$	Marine fine-grained soils	
Nishida (1956)	$C_c = 0.54 \times (e_o - 0.35)$	All clays	Void Ratio
Hough (1957)	$C_c = 0.35 \times (e_o - 0.50)$	fine-grained soil, organic silt with little clay	
Hough (1957)	$C_c = 0.29 \times (e_o - 0.27)$	Inorganic and cohesive soil, silty clay	
Cozzolino (1961)	$C_c = 0.43 \times (e_o - 0.25)$	Brazilian clays	
Sowers (1970)	$C_c = 0.75 \times (e_o - 0.50)$	fewer plasticity soils	
Rendon Herrero (1980)	$C_c = 0.3 \times (e_o - 0.27)$	All soil types	

Herrero (1983)	$C_c = 0.434 \times (e_o - 0.336)$	Alluvial deposits, Surat, India	
Nagaraj and Murthy (1983)	$C_c = 0.2237\ e_L$	Various clays	
Nagaraj and Srinivasa Murthy (1986)	$C_c = 0.2343\ e_L$	NC-saturated uncemented fine-grained soil	
Nagaraj and Srinivasa Murthy (1986)	$C_c = 0.39\ e$	NC-saturated uncemented fine-grained soil	
Bowles (1989)	$C_c = 0.75 \times (e_o – 0.5)$	less plasticity soil; ($St < 5$)	
Bowles (1989)	$C_c = 1.21 + 1.055 \times (e_o - 1.87)$	Clay soils of Sao Paulo; ($S_t > 5$)	
Bowles (1989)	$C_c = 0.208 \times (e_o – 0.083)$	Clay soils of Chicago (Moderately OC)	
Bowles (1989)	$C_c = 0.156\ e_o – 0.0107$	All types of clays (Moderately OC)	
Abdrabbo & Mahmoud (1990)	$C_c = 0.42 \times (e_o – 0.5)$	Clay soils of Egypt ($0.6 < e_o < 2.0$)	
Hirata et al. (1990)	$C_c = 0.633 \times e_o – 0.215$	Natural soils and Artificial soil of Osaka and Hyogo (clay content > 20 %)	
Burland (1990)	$C_c = 0.256\ e_L - 0.04$	Reconstituted clay ($0.6 < e_L < 4.5$)	
Nagaraj et al. (1995)	$C_c = 0.274\ e_L$	Various clays	
Lav & Ansal (2001)	$\ln C_c = 1.272 \ln e_o – 1.282$	All Soils	
Yoon et al. (2004)	$C_c = 0.54 \times (e_o – 0.37)$	South Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.39 \times (e_o – 0.13)$	East Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.37 \times (e_o – 0.28)$	West Coast (Korea clay)	
Solanki and Desai (2008)	$C_c = 0.4066 \times e_o – 0.0415$	Alluvial deposits, Surat, India	
Vinod & Bindu (2010)	$C_c = 0.2875 \times (e_o – 0.5082)$	Remoulded Kuttanad clay, Kerela, India	
Vinod & Bindu (2010)	$C_c = 0.2001 \times (e_L + 0.0755)$	Remoulded Kuttanad clay, Kerela, India	
Slamet and Abdelazim (2012)	$C_c = 0.5217 \times (e_o – 0.2)$	Pontianak clay soil	
B. Tiwari and B. Ajmera (2012)	$C_c = 0.2608 e_o$	Artificial Soil mix ($10\ \% < w_{LL} < 470\ \%$)	
B. Tiwari and B. Ajmera (2012)	$C_c = 0.3921 e_o$	Artificial Soil mix ($10\ \% < w_{LL} < 470\ \%$)	
Abbasi et al. (2012)	$C_c = 0.286\ e_o - 0.054$	Fine-Grained soil of Iran($w_{LL} <$ 75%)	
Sari & Firmansyah (2013)	$C_c = 0.6787 e_o - 0.1933$	Soft clay in Indonesia (w_{LL}=30 to 50 %)	
Sari & Firmansyah (2013)	$C_c = 0.58 e_o - 0.1428$	Soft clay in Indonesia (w_{LL} = 50 to 70 %)	
Dway et al. (2014)	$C_c = 0.196\ e_o + 0.207$	Lean to Fat Clay with Low to High compressibility	
Güllü et al. (2016)	$C_c = 0.08358 e_o + 0.12739$	Clay of Baghdad City	
Zaman et al. (2017)	$C_c = 0.5562 e_o – 0.1453$	Dhaka-Chittagong clay	
Kootahi and Moradi (2017)	$C_c = – 0.167 + 0.510\ e_o$	Marine fine-grained soils	
Salih et al. (2020)	$C_c = 0.2494\ e_o + 0.0045$	Fine-grained soil with High and Low Compressibility in Iraq	
Koppula (1981)	$C_c = 1.325 \times (I_p)$	Remoulded clays	Plasticity Index
Sridharan & Nagaraj (2000)	$C_c = 0.014 \times (I_p + 3.6)$	Remoulded clays ($37\ \% < w_{LL} < 74\ \%$)	
Yoon et al. (2004)	$C_c = 0.165 + 0.014\ I_p$	East Coast (Korea clay)	

Nath & DeDalal (2004)	$C_c = 0.015\ I_p - 0.0198$	Various clays(19% < w_{LL} < 205 %)	
Solanki & Desai (2008)	$C_c = 0.0082 \times (I_p + 0.0915)$	Alluvial deposits, Surat, India	
Vinod & Bindu (2010)	$C_c = 0.0086 \times (I_P + 24.2674)$	Remoulded Kuttanad clay, Kerela, India	
B. Tiwari & B. Ajmera (2012)	$C_c = 0.014 \times (I_p)$	Artificial Soil mix (10 % < w_{LL} < 470 %)	
Akayuli and Ofosu (2013)	$C_c = 0.007 \times (I_p) + 0.01$	Weathered Birimian phyllites samples	
Dway et al. (2014)	$C_c = 0.0038 \times I_p + 0.22$	Lean to Fat Clay with Low to High compressibility	
Zaman et al. (2017)	$C_c = 0.0091\ I_p + 0.128$	Dhaka-Chittagong clay	
Kootahi and Moradi (2017)	$C_c = 0.013 + 0.020\ I_p$	Marine fine-grained soils	
Salih et al. (2020)	$C_c = -\ 0.0049\ I_p + 0.2882$	Fine-grained soil with High and Low Compressibility in Iraq	
Yoon et al. (2004)	$C_c = -1.6\gamma_d + 2.4$	South Coast (Korea clay)	Dry Unit weight
Yoon et al. (2004)	$C_c = -0.66\gamma_d + 1.15$	West Coast (Korea clay)	
Vinod and Bindu (2010)	$C_c = 0.7045\ (\gamma_w/\gamma_d - 0.4711)$	Remoulded Kuttanad clay, Kerela, India	
Abbasi et al. (2012)	$C_c = -0.461\gamma_d + 0.883$	Fine-Grained soil of Iran (w_{LL} < 75%)	
Koppula (1981)	$C_c = 0.009 \times w_n + 0.005 \times w_{LL}$	All types of clay soil	Multiple parameters (Void ratio, moisture content, plastic limit, plasticity index, Specific Gravity)
Herrero (1983)	$C_c = 0.185\ \{G_s \times (l + e_o / G_s)^2 - 0.144\}$	All clays	
Herrero (1983)	$C_c = 0.489\ \{\ln G_s (l + e_o / G_s)^2 + 0.296\}$	All clays	
Herrero (1983)	$C_c = 0.141\ G_s \times (l + e_o / G_s)^{2.382}$	All clays	
Abdrabbo & Mahmoud (1990)	$C_c = (0.095 + 0.00114\ w_n) \times (1+e_o)$	Egyptian clay (20 % < w_n < 120 %)	
Hirata et al. (1990)	$C_c = 0.005 \times w_{LL} + 0.388 \times e_o - 0.245$	Natural soils and Artificial soil of Osaka and Hyogo (clay content > 20 %)	
Al-Khafaji & Andersland (1992)	$C_c = -0.156 + 0.411 \times e_0 + 0.00058 \times w_{LL}$	All clays	
Koumoto and Park (1998a, 1998b)	$C_c = 0.302 \times (e_o - e_p) + 0.064$	Remoulded clay	
Koumoto and Park (1998a, 1998b)	$C_c = S_t^{\ 0.22} \times \{0.009\ (w_o - w_p) + 0.101\}$	Undisturbed clay	
Yoon et al. (2004)	$C_c = -0.0003 w_n + 0.538 e_o + 0.002 w_{LL} - 0.3$	South Coast (Korea clay)	
Yoon et al. (2004)	$C_c = 0.194 \times e_o + 0.0098 \times w_{LL} - 0.0025Ip - 0.256$	The clay soil of the East coast region of Korea	
Yoon et al. (2004)	$C_c = 0.0038 \times w_n + 0.12 \times e_o + 0.0065 \times w_{LL} - 0.248$	The clay soil of the Western coastal region of Korea	
Vinod and Bindu (2010)	$C_c = 0.002 \times (I_P \times G + 110.55)$	Remoulded Kuttanad clay, Kerela, India	
Vinod and Bindu (2010)	$C_c = 0.002 \times (I_s \times G + 65.35)$	Remoulded Kuttanad clay, Kerela, India	
Amardeep Singh and Shahid Noor (2012)	$C_c = 0.002 \times w_{LL} + 0.0025 \times I_p - 0.005$	Clays having low compressibility to Clays with high compressibility	
Abbasi et al. (2012)	$C_c = 0.007 \times w_n + 0.001 \times w_{LL} - 0.077$	Fine-Grained soil of Iran (w_{LL}< 75%)	
Sari & Firmansyah (2013)	$C_c = 1.0941(0.123 \times e_o + 0.01 w_n)$	Soft clay in Indonesia	

	- 0.0415	(w_{LL}= 0 to 100 %)	
Sari & Firmansyah (2013)	C_c = 0.4044 (e_o+0.01w_n) - 0.0795	Soft clay in Indonesia (w_{LL} = 0 to 100 %)	
Dway et al. (2014)	C_c = 0.52 - 0.03 x w_{LL} + 0.03 I_p	Lean to Fat Clay with Low to High compressibility	
Güllü et al. (2016)	C_c = 0.004483w_{LL} + 0.028871 e_o - 0.03029	Clay of Baghdad City	
Kootahi and Moradi (2017)	C_c = – 0.117 + 0.009 w_n + 0.004 w_{LL}	Coastal fine-grained soils	
Kootahi and Moradi (2017)	C_c = – 0.151 + 0.364 e_o + 0.007 I_p	Coastal fine-grained soils	
Kootahi and Moradi (2017)	C_c = – 0.150 + 0.361 e_o + 0.006 w_{LL}– 0.006 PL	Coastal fine-grained soils	
Kootahi and Moradi (2017)	C_c = – 0.092 + 0.008 w_n + 0.007 w_{LL} – 0.007 PL	Coastal fine-grained soils	
Kok Shien Ng et al. (2018)	C_c = 0.27 G_S – 0.005w_{LL} -0.26	Soil with Low to Intermediate Plasticity (Ip =8 to 18)	
Koumoto and Park (2004)	C_c = n_o / (371.747 – 4.275 n_o)	Undisturbed Ariake clay (S_t =3.9 – 35.5)	Porosity
Koumoto and Park (2004)	C_c= (0.0109 x C_c+0.0018) x n_o	Remolded clays	
Vinod and Bindu (2010)	C_c/n_o = 0.0108 x Cc + 0.0018	Remoulded Kuttanad clay, Kerela, India	
B. Tiwari and B. Ajmera (2012)	C_c/n_o = 1.0584 n_o + 0.0885	Artificial Soil mix (10 % < w_{LL} < 470 %)	
Sridharan & Nagaraj (2000)	C_c = 0.007 x (I_s + 18)	Remoulded clays (37 < w_{LL}< 74)	Shrinkage Index
Vinod and Bindu (2010)	C_c = 0.0055 x (I_s+ 21.2364)	Remoulded Kuttanad clay, Kerela, India	
Al-Kahdaar et al. (2010)	C_c = 0.24 x LI + 0.21	Al-Ammarah silty clay soil (Iraq)	Liquidity Index

Based on the result, it was found that for the MLR model, dry bulk specific gravity was the only factor that has a significant impact, whereas for the neural network model, the best architecture for the predictive model was proposed to have 8 neurons in the first layer of nodes and 10 neurons in the concealed layer using the Sigmund tangent threshold function. It was concluded that ANN outperforms the MLR model. Xuchao Shi et al. (2013) proposed an optimize predictive model using SVM, which is a supervised computer-aided learning technique based on a structural hazard minimizing postulate, combining it with a Genetic Algorithm (GA-SVM). A total of 49 soil sample datasets were considered for the predictive model. The outcome of the GA-SVM predictive model was also compared with the Backpropagation Neural Network (BPNN), and it was concluded that the GA-SVM model was able to predict better than the BPNN model. Shamshad Alam et al. (2014) developed a predictive model using ANN, which was trained and validated with the records of 391 soil samples. The model was prepared using FFBPA with an activation function as a hyperbolic tangent sigmoid. The outcome of the predictive models was also matched with other regression models and was found to provide better predictability. T. Fikret Kurnaz et al. (2016) used 246 Laboratory data of soil collected from various locations in Turkey to estimate C_c and recompression index using a combined ANN predictive model using FFBPA, meaning instead of a single output, the model was trained to give two outputs. The predictive model having 20 neurons in the concealed layer was able to estimate the most reliable results for C_c, but results for the recompression index were not desirable. Nitish Puri et al. (2017), using a data set of a total of 1053 soil samples collected from various departments in the state of Haryana, proposed various models based on different machine learning algorithms for estimation of geotechnical parameters of soil such as in-place density, C_c, Cohesion and Angle of Shearing resistance. Mohammed Amin Benbourasa et al. (2018) for the estimation of C_c of soil developed models using two machine learning algorithms and Multiple Regression based on 373 sets of soil samples data collected from various laboratories employed in the geotechnical construction projects implemented in Algiers. For the GP model, a total of 9940 generations was used in training the model, whereas using ANN, 210 predictive models were developed with two hidden layers, out of which the best-performing model has 14 nodes in the primary layer and 4 nodes in the secondary layer. Out of the predictive model, the ANN model showed better predictability. Achal Bhardwaj and Vijay Kumar (2020) developed ANN-based predictive models using a total of 266 test data sets of soil samples collected from 72 numbers of boreholes across 18 construction venues in Allahabad, India. In the predictive models, FFBPA was adopted, and the best architect of the model was proposed to have 5 nodes in the first layer of nodes with 4 nodes in the primary concealed layer and 8 nodes in the secondary concealed layer with one output node. Scott Kirts et al. (2018) developed a separate Machine learning Model for fine-grained, coarse-grained and organic peat soil using SVM for prediction of C_c and recompression index of soil using the data set of soil samples which were obtained from the Florida Department of Transportation in the state of Florida. Danial Mohammadzadeh S et al. (2019) developed a predictive model using Gene Expression Programming. 108 data sets of soil specimens were gathered from Mashhad, Iran, for model training and Testing. Based on model performance, it was found that e_o and w_{LL} as input have a better correlation with C_c than PL; also, the Gene Expression Programming model predicted much better than the classical regression model.

Pijush Samui (2019) developed a hybridized ANN model incorporating metaheuristic algorithms such as ABC and LM. For the model, soil sample data sets were collected from real-life urban housing projects in the city of Hai Phong, North Vietnam. Mohammed el Amin Bourouis et al. (2020), using a data set of 203 Soil samples, developed predictive models based on two computer-aided learning techniques, namely Multi-Gene Genetic Programming (MGGP) and a hybrid particle swarm optimization incorporated with Neural Network (NN-PSO). For the MGGP model, the algorithm was set up beforehand and is then translated by the model for optimal results, whereas for the NN-PSO model, two hidden layers along with activation functions as hyperbolic tangent function, which is a ratio between hyperbolic sine and cosine function and the linear saturated function were used. Y. Erzin et al. (2020) developed a predictive model using a popular Machine learning technique known as Robust Optimization (RO) based on a total of 433 oedometer test results collected from the Soil test report of Site investigations at different locations in Mazandaran province of Iran. A total of ten RO models were developed, with three of them using single variable and others multivariable; based on the result, three models, namely RO2, RO6 and RO7, performed very well. Ramachandran Saisubramanian et al. (2021) developed models for the estimation of C_c of coastal clay using Machine learning techniques. The predictive models were evolved with two separate data sets, namely data set-1 and 2. Data set 1 contains a total of 28 soil specimens obtained from eight investigation boreholes, whereas data set 2 contains 200 soil samples collected from bridge and multistorey building projects by government agencies located along the coastline of Puducherry, India. Based on the test results for Data set-1, the best architect model of ANN had 3 nodes in the first layer of nodes, 2 nodes in the concealed layer and 1 node in the output layer as for Data set-2 was 2 nodes in the first layer of nodes as well as concealed layer and 1 nodes in the output layer. Manh Duc Nguyen et al. (2022) proposed a hybrid model for C_c estimate for which the DE algorithm was combined with ANFIS to enhance the accuracy of prediction.

The model was developed using a test data set of 817 samples of soil which were collected from various project works in Hong River Delta located in Vietnam. The outcome of the model was also compared with the outcome of other models like REPTree and Dstump, and it was found that the hybrid model outperforms the regular model. Worku Firomsa Kabeta et al. (2022) proposed a predictive model and Correlation for Jimma Clay soil using Linear Regression and ANN using the test results of 24 sets of Soil samples. Based on the analysis, it was found that the ANN model outperforms the Regression model. Long Tsang et al. (2023) developed two different predictive models using a tree-structured machine learning algorithm with 391 soil sample data sets collected from 125 site locations in North Iran; the test result of soil samples was collected from published literature. In both predictive models, a five-fold cross-validation algorithm was used to avoid inaccurate predictions. Yu Huat.Chia et al. (2023) performed two different analyses, namely 1 and 2, based on the data set of 116 and 137 numbers of the soil of Alluvium formation of Malaysia and Al-Nasiriya city collected from various published literature developed predictive models for estimating the C_c using tree-structured computer-aided learning techniques, namely Random Forest (RF) and Gradient Boosting Tree (GBT). Based on analysis, although both models showed good predictability, GBT displayed stronger predictive power than RF.

Huifen Liu et al. (2023) proposed three computer-aided learning predictive models- ANN, RF model, and SVM for the determination of C_c of soft soils collected from a cluster of cities in China referred to as GBA. A dataset consisting of 743 records of measured parameters of soils of the said area was utilized for teaching and verifying the predictive models. Although all three models performed well in predicting, the ANN predictive model outperformed the other two. R Akshaya et al. (2024), utilizing computer-aided learning such as SVM and k-Nearest Neighbors (kNN) based on the results of 359 total data acquired from various studies, proposed a predictive model for C_c. The predictive model was developed with four input variables. Based on the analysis, the best prediction SVM model is the model with linear kernel, while the best kNN model is the model with Chebyshev distance. Also, the SVM model outperforms the kNN model. Sungyeol Lee et al. (2024) used Statistical Analysis and computer-aided learning techniques such as RF, XGB, and Light GBM for the development of models which are capable of predicting C_c of soft, fine-grained soil along the southern coast of Korea for which a data set of 4868 soil samples were utilized. Out of all the models, the one developed with the RF classifier showed the most accurate result, and the significance of the soil index properties was in the order of w_n, w_{LL}, PL. Mintae Kim et al. (2024), based on the results comprising 915 Soil samples data collected from various literatures, developed a predictive model using computer-aided learning techniques, namely RF-Regressor (RFR), GB-Regressor (GBR) and AdaBoost-Regressor (ABR) for C_c of fine-Grained soil. Based on the outcome of the analysis, it was concluded that the computer-aided learning algorithm, particularly GBR and RFR, showed significant capability in predicting the C_c. Qi Ge et al. (2024) utilized a comprehensive data set of 1080 soil samples from eight distinctive geographic locations with three numbers from Asia, three from Europe and two from Africa for the development of predictive models for the compression index of soil.

In the study, computer-aided learning techniques, namely RF, GB-Decision Trees (GBDT), XGB and a Stacking model, were applied for model training and validation. Among all of these predictive models, the stacking model was the most accurate, proving fewer prediction errors and powerful generalization abilities. Ali Ulvi Uzer (2024) developed an ANN predictive model using a database of 560 soil samples with low to high plasticity collected from the soil testing Laboratory of Istanbul Technical University, Turkey. For the training and validation of the predictive model, a k-fold cross-validation technique was utilized, indicating that 80% of the total data was used for the model teaching stage, and the remaining 20 % was used for the verifying stage. A brief summary of various machine learning algorithm techniques used in predicting C_c is presented in Table 2.

3. Methodology

The Schematic illustration of this review paper is given in Figure 1. This paper consists of a wide range of literature relating to the prediction of C_c of soil. Efforts have been made in this study to investigate the past work done by various researchers and summarize it.

A total of 75 literature have been reviewed and were categorized into two groups, namely literature on linear regression and literature on machine learning algorithms. The published literature chosen was from various timelines, past as well as recent. The advantages and limitations of both correlations developed by regression and predictive machine learning models were also discussed for better understanding.

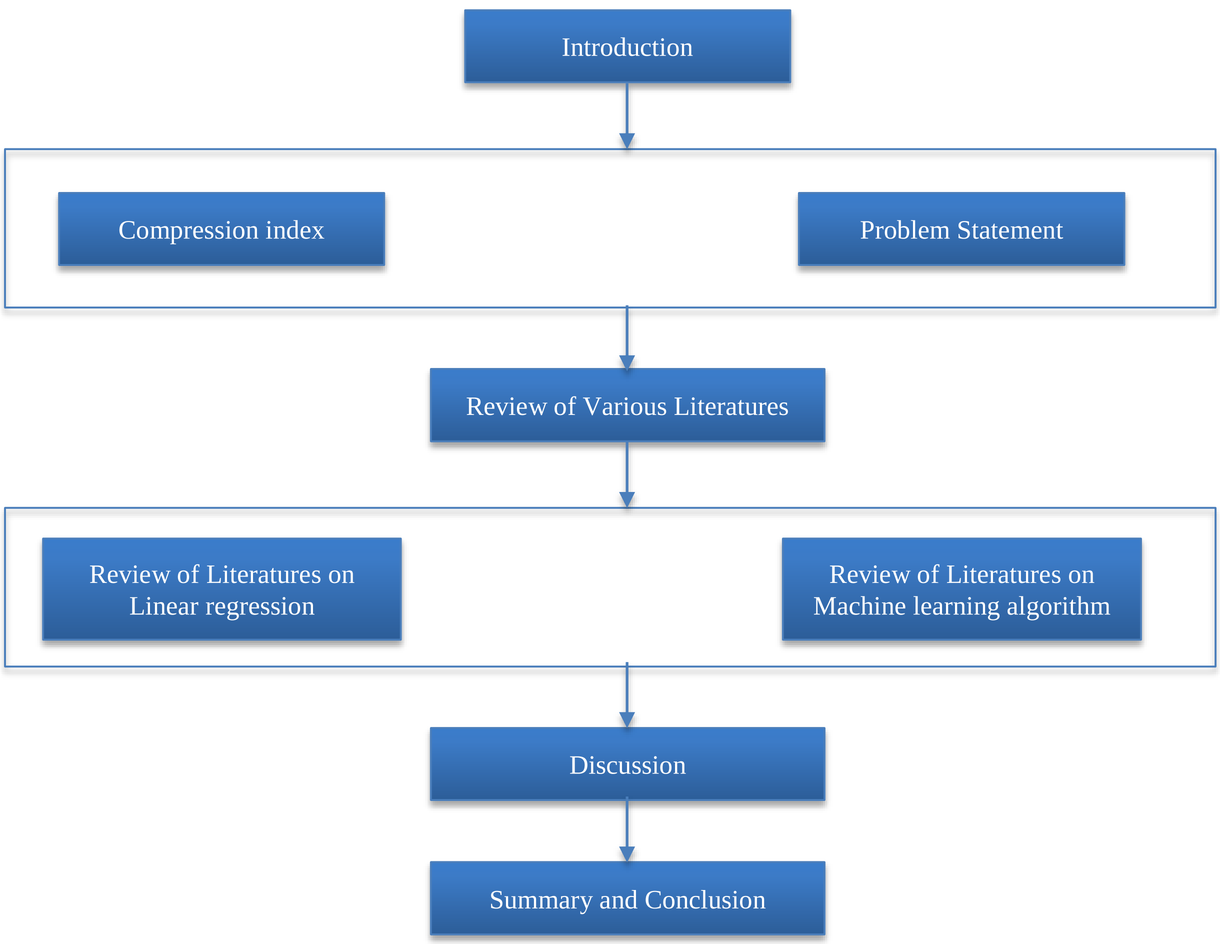

Fig. 1 Flow chart of review

Table 2. Summary of various input variables used in predictive model development

Reference	Input variable	No of data	Techniques used (Statistical Indices)
Hyun Il Park and Seung Rae Lee (2011)	w_n, e_o, w_{LL}, I_p	947	ANN (R^2 =0.885)
P.K. Kolay et al. (2011)	w_n, Bulk density, Dry density, e_o, Specific gravity w_{LL}, I_p, Gravel (%), Sand (%), silt (%), Clay (%), Pre-consolidation pressure	700	ANN (R=0.0756 to 0.7541)
V. Phani Kumar et al. (2011)	Fine Fraction w_{LL}, I_p, Maximum Dry Density, Optimum Moisture Content	68	ANN (R^2 =0.974)
Farzaneh Namdarvand et. al. (2013)	% clay, % silt, % sand, Wet bulk density, Dry bulk density, Friction coefficient, Viscosity coefficient, PL	100	ANN (R = 0.63) MLR (R = 0.47)
Xuchao Shi et al. (2013)	I_p, w_n, e_o, Density	49	GA-SVM (R = 0.972)
Shamshad Alam et al. (2014)	w_n, w_{LL}, e_o, I_p	391	ANN (R =0.852)
T. Fikret Kurnaz et. al. (2016)	w_n, e_o, w_{LL}, I_p.	246	ANN (R^2 =0.897)
Nitish Puri et al. (2018)	w_{LL} and e_o	1053	LR (R^2 = 0.92) SVM (R^2 = 0.92) ANN (R^2 = 0.92) RF (R^2 = 0.94) M5P (R^2 = 0.95)
Mohammed Amin Benbourasa et al. (2018)	Wet density, Water content, e_o, w_{LL}, I_p, Fine content	373	MRA (R = 0.64) GP (R = 0.048) ANN (R = 0.75)
Achal Bhardwaj and Vijay Kumar (2018)	SPT-N value w_{LL}, I_p, G_s, Dry unit weight	266	ANN (R^2 = 0.9706)
Scott Kirts et al. (2018)	w_n, e_o, Dry unit weight, Moist unit weight, Automatic hammer SPT blow count, Overburden stress, Fines content (−200)	619	SVM (R^2 = 0.77 - 0.91)
Danial Mohammadzadeh S et. al. (2019)	w_{LL}, PL, e_o	108	GEP (R^2 = 0.832)
Pijush Samui (2019)	Sample depth, Sand (%), Loam (%), Clay (%), w_n Wet density, Dry Density, e_o, w_{LL}, PL, I_p, LI	441	ABC-LM-ANN (R^2 = 0.864)
Mohammed el Amin Bourouis et.al (2020)	w_n, e_o, Vertical stress	203	MGGP (R=0.9983) NN-PSO (R =0.999)
Y. Erzin et al. (2020)	e_o, w_{LL}, w_n, I_p, G_s	433	RO (R^2 =0.9226)
Ramachandiran Saisubramanian et al. (2021)	Data set 1: PL, w_{LL} & w_n Data set 2: PL, w_{LL}, I_p, G_s, Swell Percentage, N- value, PL/w_{LL}	28	ANN (Data set 1: R^2 =0.994, Data set 2: R^2 =0.994) MLR (Data set 1: R^2 =0.99, Data set 2: R^2 =0.999)
Manh Duc Nguyen et al. (2021)	w_n, e_o, w_{LL}, G_s, PL, Clay content, Sample depth	817	ANFIS-DE (R =0.825) REPTree (R =0.7802) Dstump (R =0.7325)
Worku Firomsa Kabeta et al.	w_{LL}, PL, I_p	24	ANN (R^2=0.939) MLR (R^2=

(2022)			0.841)
Long Tsang et.al (2023)	e_o, w_n, w_{LL}, I_p, G_s	391	RF (R^2 =0.818) XGB (R^2 =0.833)
Yu Huat.Chia et al. (2023)	For analysis 1: e_o, w_{LL}, PL, w_n, For analysis 2: e_o. w_{LL}, PL	For analysis 1: 116 For analysis 2: 137	RF (Analysis 1: R^2=0.71, Analysis 2: R^2 =0.86) GBT (Analysis 1: R^2=0.63, Analysis 2: R^2 =0.89)
Huifen Liu et al. (2023)	Moisture content, Density, Void ratio	743	ANN (R^2 =0.827) RF (R^2 =0.769) SVM (R^2 =0.689)
R Akshaya et al. (2024)	w_{LL}, I_p, w_n, and e_o .	359	SVM (R^2 =0.64) kNN (R^2 =0.60)
Sungyeol Lee et al. (2024)	w_n, w_{LL}, I_p, PL, e_o	4868	LR (R^2 =0.687) RF (R^2 =0.72) XGB (R^2 =0.70) LGBM(R^2 =0.71)
Mintae Kim et al. (2024)	w_{LL}, PL, I_p, w_n, e_o, G_s	915	RFR (R^2 =0.926) GBR (R^2 =0.930) ABR (R^2 =0.921)
Qi Ge et al. (2024)	w_{LL}, I_p, e_o, w_n	1080	RF (R^2 =0.843) GBDT (R^2 =0.84) XGB (R^2 =0.839) Stacking Model (R^2 =0.848)
Ali Ulvi Uzer (2024)	w_n, w_{LL}, PL, I_p, e_o	560	ANN (R^2 =0.81)

Notations:

C_c - Compression Index; RF – Random Forest; RFP – Random Forest Regressor; w_o – Optimum Moisture content; M5P - M5 Tree; kNN - K-Nearest Neighbors; GP – Genetic Programming; ABR - AdaBoost Regressor; e_p - void ratio at PL; w_{LL}- liquid limit; SVM - Support Vector Machine; n_o – porosity; e_L – void ratio at w_{LL}; RO - Robust Optimization; e_o - initial or in-situ void ratio; w_n- Natural water content; I_s – Shrinkage Index; C_v = Coefficient of consolidation; G_s- Specific Gravity of soil; GBR - Gradient Boosting Regressor; GBDT - Gradient Boosting Decision Trees; NCC – Normally Consolidated clay; LI – Liquidity Index; γ_w - unit weight of water; MRA- Multiple Regression Analysis; LR – Linear Regression; GBT - Gradient Boosting Tree; GEP - Gene Expression Programming; PL – Plastic Limit; I_p – Plasticity Index; MLR- Multiple Linear regression; MGGP - Multi-Gene Genetic Programming; LGBM – Light Gradient Boosting Method; XGB – Extreme Gradient Boosting; GA-SVM-Genetic Algorithm and Support Vector Machine; γ_d - dry unit weight; NN- PSO - Neural network and Particle Swarm Optimization ANN- Artificial Neural Network; FFBPA – Feed forward Back propagation Algorithm; S_t - sensitivity of the clay; ABC-LM-ANN - Artificial bee colony Levenberg Marquardt Artificial Neural Network; ANFIS-DE – Adaptive Network based Fuzzy Inference System with Differential Evolution; REPTree - Reduced Error Pruning Trees; Dstump - Decision Stump; R^2- Coefficient of determination; R - Coefficient of correlation.

4. Discussion

The main aim of this research is to investigate the past work in predicting or estimating the Cc of soil and summarize it. Based on the findings, it is possible to state that most of the past studies related to the prediction of Cc have been focused on fine-grained soils because these soils, such as clay and silt, undergo settlement over an extended period due to their low permeability. Findings also revealed that linear regression as a predictive technique has been used extensively in the past, whereas machine learning techniques such as ANN, XGB, etc., along with various hybrid techniques, have been gaining much popularity nowadays. This is because, with advancements in technology in recent times, many people are able to use various computer learning algorithms to analyze complex non-linear behavior.

4.1. Discussion on Advantages

Since C_c is normally determined using time-consuming and labour-intensive laboratory testing. It is desirable to have predictive models and empirical correlations that are capable of estimating Cc using simple, measurable soil properties. This will help in a quick and easy determination of $C_{c,}$ which in return reduces the cost of the project and saves time and effort. Empirical models based on regression analysis and machine learning models, such as standalone models or hybrid models, have shown great potential in C_c prediction. Empirical models have the advantage that they are simple, quick and easy to implement but have limited flexibility. Machine learning models, on the other hand, are designed to handle complex patterns by automatically capturing nonlinear relationships between soil properties and compression characteristics. This is due to the reason that when any machine learning model is developed, the first step is to train the model using large data sets using featured selection and optimizing techniques, which helps the model learn the complex interconnection between input and output parameters, which is then followed by the second step of testing the trained model with new data sets and validating the results with statistical indices.

This makes the model capable of predicting the outcome with very high accuracy if it is a generalized one.

4.2. Discussion on Limitations

Although both techniques use some similarities in the prediction of outcomes, such as recognizing data patterns, regression analysis is mainly focused on finding the best fit for the data set used, due to which, upon testing this correlation with new sets of data, they can display results with deviation up to 30 % (Spagnoli and Shimobe 2020), whereas machine learning seeks to find the best generalization model to give better performance against future data set by tackling the effect of overfitting and underfitting exhibiting better results compared to classical regression analysis.

Regardless of how well a machine learning model shows evidence of improved outcomes, certain limitations related to machine learning models shall also be addressed before their application in real-life situations, such as the availability of high-quality data, generalization, interpretability of model, etc.

Also, due to the structural non-homogeneity of soil, the use of any empirical correlation established or machine learning models trained on a limited data set to give reliable results out of the scope of geographic location is very difficult. Another difficulty with Machine learning models is the selection of input variables many a time, and a specific variable is not considered if measurements of that variable are not shown reliably as part of the dataset or even vary a lot or are not measured well, as in such cases the model may not perform well with those variables as inputs and might result in either underfitting or overfitting of the model.

Table 3. List of commonly used input parameters for compression index estimate

Input Parameters	Frequency	Reference
Liquid Limit	52	[1, 4, 5, 9, 14, 16-22, 27, 28, 31-34, 36-50, 53, 54, 56-59, 62-66, 68-74]
Void Ratio	50	[3, 5, 6, 8, 11-13, 15, 17-19, 21-25, 27, 28, 31-33, 35, 37, 40, 41, 44-46, 48-50, 52-55, 57-62, 64, 65, 67-74]
Natural Moisture content	40	[2, 7, 9, 10, 12, 18, 27, 28, 31-33, 35-37, 40, 42, 45-50, 53, 55, 56, 58, 60-63, 65-74]
Plasticity index	31	[10, 26, 28, 30-32, 34, 40-42, 44, 45, 48-50, 53-56, 58, 59, 62-65, 68, 69-73]
Specific Gravity	10	[12, 32, 38, 53, 59, 62, 63, 65, 70, 74]
Plastic Limit	11	[48, 51, 52, 63, 64, 66, 69, 70, 72, 73, 74]
Dry Density	7	[28, 32, 46, 51, 53, 54, 73]
Bulk Density	6	[51, 53, 55, 58, 67, 73]
Porosity	3	[29, 32, 41]
Fine content (%)	3	[54, 58, 60]
SPT-N value	2	[59, 63]
% sand	3	[51, 53, 73]
Shrinkage index	2	[26, 32]
% clay	3	[51, 53, 73]
% silt	2	[51, 53]
Dry Unit Weight	2	[59, 60]
OMC	1	[54]
Friction Coefficient	1	[51]
Overburden stress	1	[60]
Vertical stress	1	[61]

% swell	1	[63]
% gravel	1	[53]
PL/w_{LL}	1	[63]
Pre-consolidation pressure	1	[53]
Liquidity Index	2	[47, 73]
Automatic hammer SPT blows	1	[60]
Viscosity coefficient	1	[51]
Moist unit weight	1	[60]
Clay Content	1	[74]
Sample Depth	2	[73, 74]
Loam (%)	1	[73]

Table 4. Most commonly used techniques for compression index estimate

Technique used	Frequency	Reference
LR	55	[1-48, 51, 57, 63, 64, 69]
ANN	13	[49, 50, 51, 53, 54, 56-59, 63, 64, 67, 72]
, RF	6	[57, 65-67, 69, 71]
XGB	3	[65, 70, 71]
GEP	1	[52]
GP	1	[58]
RO	1	[62]
GBT	1	[66]
SVM	4	[57, 60, 67, 68]
LGBM	1	[69]
GA-SVM	1	[55]
MGGP	1	[61]
NN-PSO	1	[61]
KNN	1	[68]
GBR	1	[70]
Staking model	1	[71]
GBDT	1	[71]
RFR	1	[70]
ABR	1	[70]
ABC-LM-ANN	1	[73]
ANFIS-DE	1	[74]
REPTree	1	[74]
Dstump	1	[74]

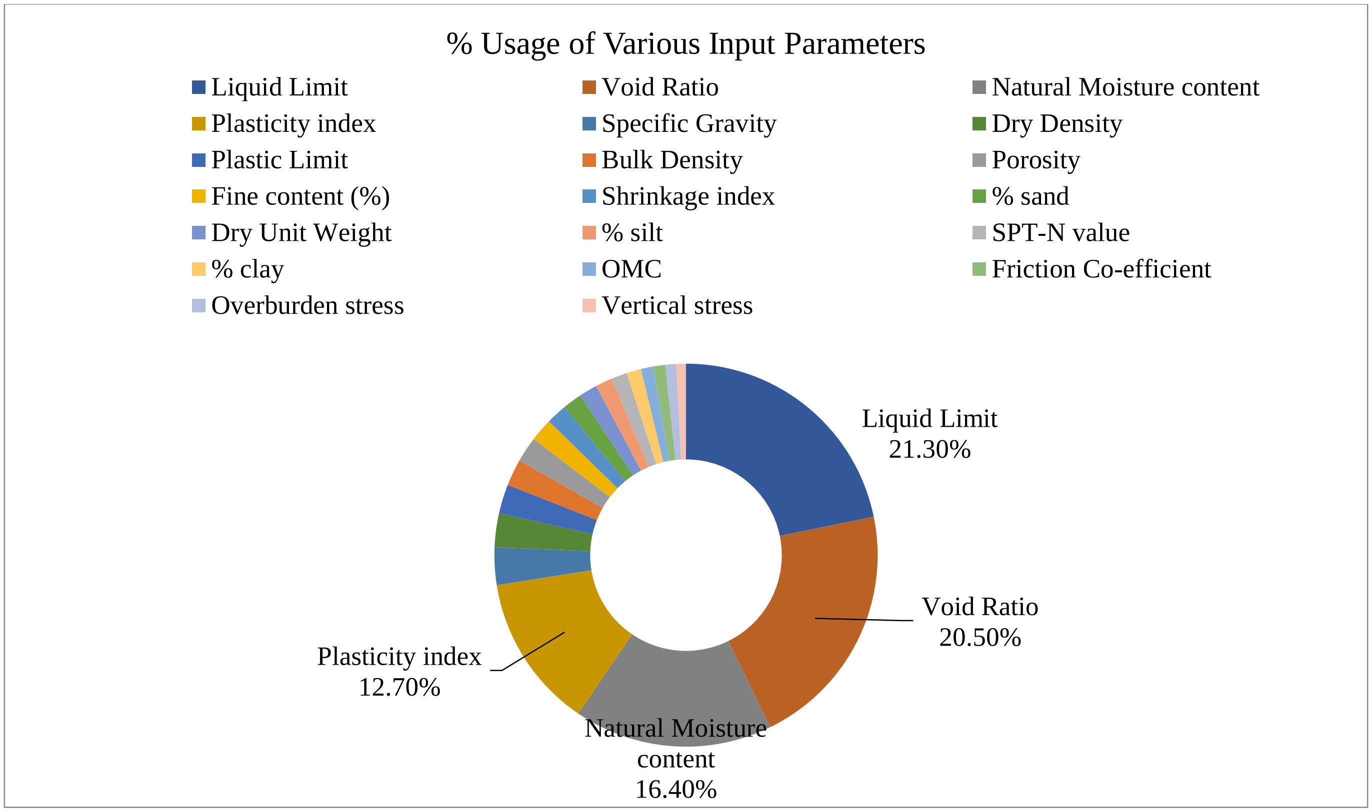

Fig. 2 Percentage usage of various input parameters

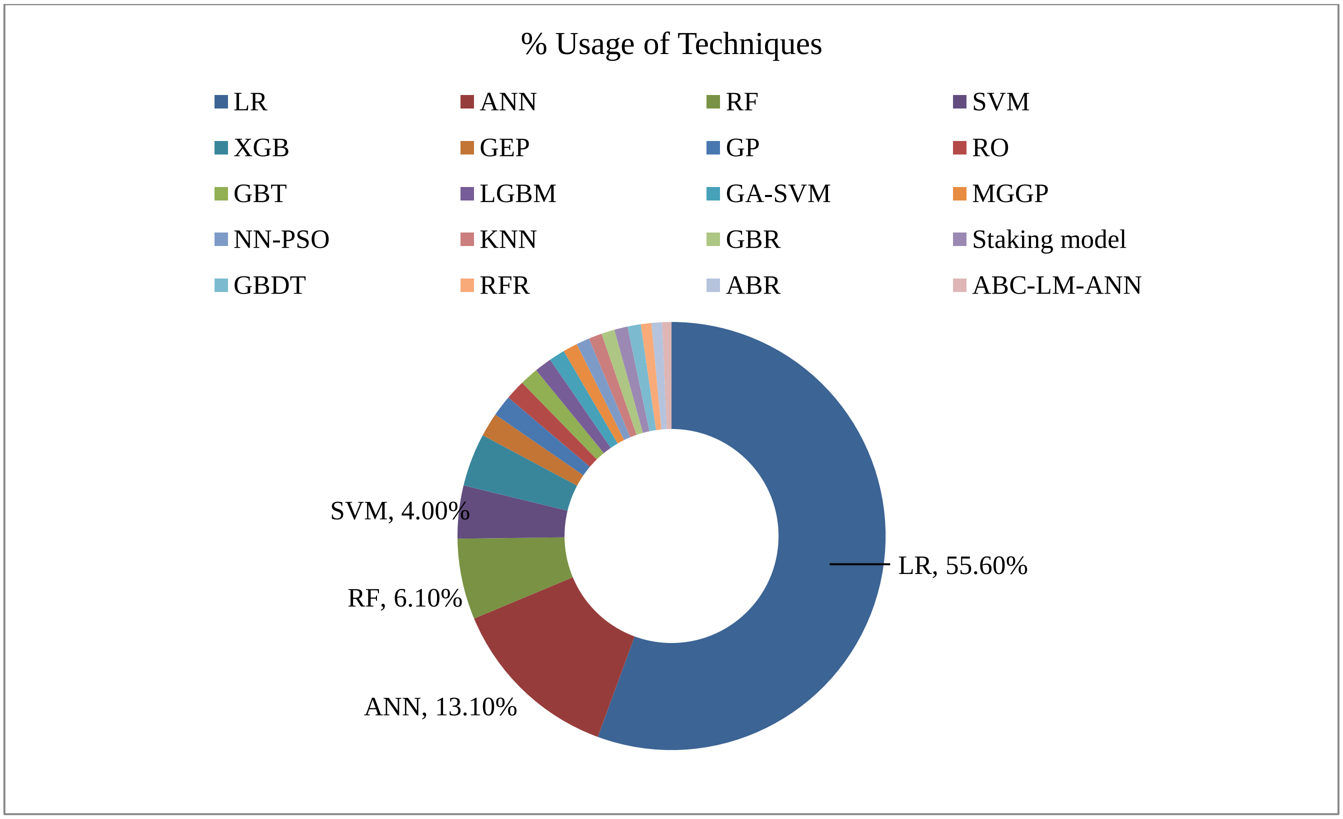

Fig. 3 Percentage usage of various techniques for compression index estimate

5. Summary and Conclusion

The present paper comprises a catalogue of empirical equations and predictive models for estimating (C_c) of soil, which will be useful for geotechnical engineers for quick identification of empirical equations and predictive models as per their regional suitability for easy determination of compression index and predict the amount of settlement the structure might undergo due to its loading.

Based on the analysis of the literature review focusing on compression characteristics (Compression index of soil), some key aspects or points are drawn:

- The applicability of established correlations and Predictive models is a robust way for a preliminary estimate of the compression index of soil. However, since the soil is a very complex material, and its composition varies depending upon its geological location, these correlation and predictive models give more or less accurate results if applied according to their suitability.
- It was observed that for establishing correlation and predictive models for the C_c estimate of soil, various input or dependent parameters, most of which are index properties were utilized. A list of input or dependent parameters, along with their usage frequency, is given in Table 3.
- It was also observed that although a vast majority of correlations were established using Linear regression such as single or multiple, recent studies have resorted to the use of machine learning algorithms such as ANN, Random, RF, and XGB as a potential alternative technique for prediction of the compression characteristics of soil which offers competence in nonlinear modeling. A list of techniques, along with their frequency, is given in Table 4.
- The performance and validation of the correlations and the predictive models were done based on various statistical indices' results, importance, and relevance. The most commonly used statistical indices include R and R^2. A higher value of these indices shows greater strength and models' reliability for future prediction of data with confidence.
- From Figure 2, it can be interpreted that w_{LL} e_o w_n and I_p are the most frequently used input variables, indicating a strong correlation with C_c.
- From Figure 3, it can be seen that Linear Regression, ANN, RF and SVM are the most frequently used techniques for C_c estimate.
- Different predictive models have a different emphasis on the aspects of the system. Hence, if there is a strong statistical correlation between them and the outcome being predicted, the accuracy of the model increases. Because of these, certain variables such as liquid limit or void ratio are effective in specific predictive models, whereas some other models could experience difficulties with such variables since the relationships between the variable and the outcome may not be linear or other interactions interfere with the model's ability to predict the outcome.
- Since the studies mentioned in this paper were mainly focused only on the development of empirical correlations and predictive models based on the limited data sets, details regarding their practical application in real-world scenarios with case studies are not mentioned here.

5.1. Future Research Directions

Future research could focus on the enhancement of machine learning models by implementing a hybrid approach and using a multiscale deep learning model for better-incorporating soil heterogeneity.

Acknowledgments

The authors would like to acknowledge and sincerely thank the organizing committee of the International Conference on Computer-Aided Modeling for the Sustainable Development of Smart Cities (CAMSSC), sponsored by the Anusandhan National Research Foundation (ANRF), held at the Department of Civil Engineering, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh, India, during November 27–30, 2024, for allowing us to present the paper and sponsoring the paper for publication.

References

[1] Alec Westley Skempton, "Notes on the Compressibility of Clays," *Quarterly Journal of the Geological Society*, vol. 100, no. 1-4, pp. 119- 135, 1944. [CrossRef] [Google Scholar] [Publisher Link]

[2] K.V. Helenelund, "On Consolidation and Settlement of Loaded Soil Layers," *Soil Science*, vol. 73, no. 2, 1952. [Google Scholar] [Publisher Link]

[3] Yoshichika Nishida, "A Brief Note on Compression Index of Soil," *ASCE Journal of the Soil Mechanics and Foundations Division*, vol. 82, no. 3, pp. 1-14, 1956. [CrossRef] [Google Scholar] [Publisher Link]

[4] Karl Terzaghi, Ralph Brazelton Peck, and Gholamreza Mesri, *Soil Mechanics in Engineering Practice*, Wiley, pp. 1-549, 1996. [Google Scholar] [Publisher Link]

[5] E.V.M. Cozzolino, "Statistical Forecasting of Compression Index," *Proceedings of the 5th International Conference on Soil Mechanics and Foundation Engineering*, Paris, vol. 1, pp. 51-53, 1961. [Google Scholar] [Publisher Link]

[6] Benjamin K. Hough, *Basic Soils Engineering*, Ronald Press Company, pp. 1-513, 1957. [Google Scholar] [Publisher Link]

[7] Proctor Moran, Mueser, and Rutledge, *Study of Deep Soil Stabilization by Vertical Sand Drains*, Bureau of Yards and Docks, Department of the Navy, 1958. [Google Scholar] [Publisher Link]

[8] George F. Sowers, *Introductory Soil Mechanics and Foundations: Geotechnical Engineering*, Macmillan, pp. 1-621, 1979. [Google Scholar] [Publisher Link]

[9] Amr S. Azzouz, Raymond J. Krizek, and Ross B. Corotis, "Regression Analysis of Soil Compressibility," *Soils and Foundations*, vol. 16, no. 2, pp. 19-29, 1976. [CrossRef] [Google Scholar] [Publisher Link]

[10] S.D. Koppula, "Statistical Estimation of Compression Index," *Geotechnical Testing Journal*, vol. 4, no. 2, pp. 68-73, 1981. [CrossRef] [Google Scholar] [Publisher Link]

[11] Oswald Rendon-Herrero, "Universal Compression Index Equation," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 106, no. 11, pp. 1179-1200,1981. [CrossRef] [Google Scholar] [Publisher Link]

[12] Oswald Rendon-Herrero, "Closure to "Universal Compression Index Equation"," *Journal of Geotechnical Engineering*, vol. 109, no. 5, pp. 755-761, 1983. [CrossRef] [Google Scholar] [Publisher Link]

[13] T.S. Nagaraj, and B.R. Srinivasa Murthy, "Rationalization of Skempton's Compressibility Equation," *Géotechnique*, vol. 33, no. 4, pp. 433-443, 1983. [CrossRef] [Google Scholar] [Publisher Link]

[14] Paul W. Mayne, "Cam-Clay Predictions of Undrained Strength," *Journal of the Geotechnical Engineering Division*, vol. 106, no. 11, pp. 1219-1242, 1980. [CrossRef] [Google Scholar] [Publisher Link]

[15] T.S. Nagaraj, and B.R. Srinivasa Murthy, "A Critical Reappraisal of Compression Index Equations," *Geotechnique*, vol. 36, no. 1, pp. 27-32, 1986. [CrossRef] [Google Scholar] [Publisher Link]

[16] Giuseppe Scarpelli, "Properties of Italian Clay Soils Geotechnical Engineering in Italy," *AGI, ISSMFE Golden Jubilee*, pp. 239-249, 1985. [Google Scholar] [Publisher Link]

[17] Joseph E. Bowles, *Physical and Geotechnical Properties of Soils*, McGraw-Hill, pp. 1-578, 1984. [Google Scholar] [Publisher Link]

[18] F.M. Abdrabbo, and M.A. Mahmoud, "Correlations between Index Tests and Compressibility of Egyptian Clays," *Soils and Foundations*, vol. 30, no. 2, pp. 128-132, 1990. [CrossRef] [Google Scholar] [Publisher Link]

[19] Shigeyoshi Hirata, Shintaro Yao, and Kazuhiko Nishida, "Multiple Regression Analysis between the Mechanical and Physical Properties of Cohesive Soils," *Soils and Foundations*, vol. 30, no. 3, pp. 91-108, 1990. [CrossRef] [Google Scholar] [Publisher Link]

[20] Tsuchida Takashi, "A New Concept of E-Log p Relationship for Clays," *Proceedings of the 9th Asian Regional Conference on Soil Mechanics and Foundation Engineering*, pp. 71-74, 1991. [Google Scholar]

[21] A.W. Al-Khafaji, and O.B. Andersland, "Equations for Compression Index Approximation," *Journal of Geotechnical Engineering*, vol. 118, no. 1, pp. 148-153, 1992. [CrossRef] [Google Scholar] [Publisher Link]

[22] J.B. Burland, "On the Compressibility and Shear Strength of Natural Clays," *Géotechnique*, vol. 40, no. 3, pp. 329-378, 1990. [CrossRef] [Google Scholar] [Publisher Link]

[23] T.S. Nagaraj, and B.R. Srinivasa Murthy, "Prediction of the Pre-Consolidation Pressure and Recompression Index of Soils," *Geotechnical Testing Journal*, vol. 8, no. 4, pp. 199-202, 1985. [CrossRef] [Google Scholar] [Publisher Link]

[24] T. Koumoto, and J.H. Park, "Compression Index Equation for Remolded Clays," *Transactions-Japanese Society of Irrigation Drainage and Reclamation Engineering*, vol. 193, pp. 81-86, 1998. [Google Scholar]

[25] T. Koumoto, and J.H. Park, "Compression Index Equation for Undisturbed Clays," *Transactions-Japanese Society of Irrigation Drainage and Reclamation Engineering*, vol. 194, pp. 59-63, 1998. [Google Scholar] [Publisher Link]

[26] A. Sridharan, and H.B. Nagaraj, "Compressibility Behaviour of Remoulded, Fine-Grained Soils and Correlation with Index Properties," *Canadian Geotechnical Journal*, vol. 37, no. 3, pp. 712-722, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[27] M. Ayşen Lav, and Atilla M. Ansal, "Regression Analysis of Soil Compressibility," *Turkish Journal of Engineering and Environmental Sciences*, vol. 25, no. 2, pp. 101-109, 2001. [Google Scholar] [Publisher Link]

[28] Gil Lim Yoon, Byung Tak Kim, and Sang Soo Jeon, "Empirical Correlations of Compression Index for Marine Clay From Regression Analysis," *Canadian Geotechnical Journal*, vol. 41, no. 6, pp. 1213-1221, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[29] J.H. Park, and T. Koumoto, "New Compression Index Equation," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 130, no. 2, pp. 223-226, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[30] A. Nath, and S.S. DeDalal, The Role of Plasticity Index in Predicting Compression Behaviour of Clays, Materials Science, vol. 9, no. 1, 2024. [Online]. Available: https://www.researchgate.net/publication/290629787_The_role_of_plasticity_index_in_predicting_compression_behaviour_of_clays

[31] C.H. Solanki, and J.A. Desai, "Significance of Atterberg Limits on Compressibility Parameters of Alluvial Deposits-New Correlations," *Proceedings of Indian Geotechnical Conference*, Bangalore, India, pp. 17-19, 2008. [Google Scholar]

[32] P. Vinod, and J. Bindu, "Compression Index of Highly Plastic Clays–An Empirical Correlation," *Indian Geotechnical Journal*, vol. 40, no. 3, pp. 174-180, 2010. [Google Scholar]

[33] Slamet Widodo, and Abdelazim Ibrahim, “Estimation of Primary Compression Index (Cc) Using Physical Properties of Pontianak Soft Clay,” *International Journal of Engineering Research and Applications*, vol. 2, no. 5, pp. 2232-2236, 2012. [Google Scholar] [Publisher Link]

[34] Amardeep Singh, and Shahid Noor, “Soil Compression Index Prediction Model for Fine Grained Soils,” *International Journal of Innovations in Engineering and Technology*, vol. 1, no. 4, pp. 34-37, 2012. [Google Scholar] [Publisher Link]

[35] Putu Tantri Kumala Sari, and Yerry Kahaditu Firmansyah, “The Empirical Correlation Using Linear Regression of Compression Index for Surabaya Soft Soil,” *The 2013 World Congress on Advances in Structural Engineering and Mechanics*, pp. 3008-3019, 2013. [Google Scholar]

[36] Bryan A. McCabe et al., “Empirical Correlations for the Compression Index of Irish Soft Soils,” *Proceedings of ICE Geotechnical Engineering*, vol. 167, no. 6, pp. 510-517, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[37] Hamza Güllü, Hanifi Canakci, and Ali Alhashemy, “Development of Correlations for Compression Index,” *Afyon Kocatepe University Journal of Sciences and Engineering*, vol. 16, no. 2, pp. 344-355, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[38] Kok Shien Ng, Yee Ming Chew, and Nur Izzati Ahmad Lazim, “Prediction of Consolidation Characteristics from Index Properties,” *E3S Web of Conferences: International Conference on Civil and Environmental Engineering*, vol. 65, pp. 1-5, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[39] K. Shiva Prashanth Kumar, and N. Darga Kumar, “Evaluation of Coefficient of Consolidation in CH Soils,” *Jordan Journal of Civil Engineering*, vol. 10, no. 4, pp. 515-528, 2016. [Google Scholar] [Publisher Link]

[40] Wasif Zaman, Rezwan Hossain, and Hossain Shahin, “Correlation Studies between Consolidation Properties and Some Index Properties for Dhaka-Chittagong Highway Soil,” *Proceedings of 1st International Conference on Engineering Research and Practice*, Dhaka, Bangladesh, vol. 1, 2017. [Publisher Link]

[41] Binod Tiwari, and Beena Ajmera, “New Correlation Equations for Compression İndex of Remoulded Clays,” *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 138, no. 6, pp. 757-762, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[42] C.F.A. Akayuli, and Bernard Ofosu, “Empirical Model for Estimating Compression İndex from Physical Properties of Weathered Birimian Phyllites,” *Electronic Journal of Geotechnical Engineering*, vol. 18, pp. 6135-6144, 2013. [Google Scholar]

[43] Pundreek Dwivedi, Rakesh Kumar, and P.K. Jain, “Prediction of Compression Index (Cc) of Fine-Grained Remolded Soils from Basic Soil Properties,” *International Journal of Applied Engineering Research*, vol. 11, no. 1, pp. 592-598, 2016. [Google Scholar] [Publisher Link]

[44] Nihad Bahaaldeen Salih, “Geotechnical Characteristics Correlations for Fine-Grained Soils,” *IOP Conference Series: Materials Science and Engineering: 4th International Conference on Buildings, Construction and Environmental Engineering*, Istanbul, Turkey, vol. 737, no. 1, pp. 1-12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[45] Seinn Moh Moh Dway, and Daw Aye Aye Thant, “Soil Compression Index Prediction Model for Clayey Soils,” *International Journal of Scientific Engineering and Technology Research*, vol. 3, no. 11, pp. 2458-2462, 2014. [Google Scholar]

[46] Nader Abbasi, Akbar A Javadi, and Reza Bahramloo, “Prediction of Compression Behaviour of Normally Consolidated Fine-Grained Soils,” *World Applied Sciences Journal*, vol. 18, no. 1, pp. 6-14, 2012. [Google Scholar] [Publisher Link]

[47] Rana Mohammed Al-Kahdaar, and Abbas Fadhil Ibrahim Al-Ameri, “Correlations between Physical and Mechanical Properties of Al-Ammarah Soil in Messan Governorate,” *Journal of Engineering*, vol. 16, no. 4, pp. 5946-5957, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[48] Karim Kootahi, and Gholam Moradi, “Evaluation of Compression Index of Marine Fine-Grained Soils by the Use of Index Tests,” *Marine Georesources & Geotechnology*, vol. 35, no. 4, pp. 548-570, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[49] Hyun Il Park, and Seung Rae Lee, “Evaluation of the Compression Index of Soils Using an Artificial Neural Network,” *Computers and Geotechnics*, vol. 38, no. 4, pp. 472-481, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[50] T. Fikret Kurnaz et al., “Prediction of Compressibility Parameters of the Soils Using Artificial Neural Network,” *SpringerPlus*, vol. 5, pp. 1-11, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[51] Farzaneh Namdarvand, Alireza Jafarnejadi, and Gholamabbas Sayyad, “Estimation of Soil Compression Coefficient Using Artificial Neural Network and Multiple Regressions,” *International Research Journal of Applied and Basic Sciences*, vol. 4, no. 10, pp. 3232-3236, 2013. [Google Scholar]

[52] S. Danial Mohammadzadeh et al., “Prediction of Compression Index of Fine-Grained Soils Using a Gene Expression Programming Model,” *Infrastructures*, vol. 4, no. 2, pp. 1-12, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[53] Prabir K. Koyal, A.B. Rosmina, and Yap Shirley, “Prediction of Compression Index for Tropical Soil by using Artificial Neural Network (ANN),” *Proceedings of 13th International Conference of the IACMAG*, pp. 542-547, 2011. [Google Scholar]

[54] V. Phani Kumar, and Ch. Sudha Rani, “Prediction of Compression Index of Soils Using Artificial Neural Networks (ANNs),” *International Journal of Engineering Research and Applications*, vol. 1, no. 4, pp. 1554-1558, 2011. [Google Scholar] [Publisher Link]

[55] Xu Chao Shi, and Ying Fei Gao, “Application of Genetic Arithmetic and Support Vector Machine in Prediction of Compression Index of Clay,” *Applied Mechanics and Materials*, vol. 438-439, pp 1167-1170, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[56] Shamshad Alam, Sunil Khuntia, and Chittaranjan Patra, "Prediction of Compression Index of Clay Using Artificial Neural Network," *Proceedings of International Conference on Industrial Engineering Science and Applications*, pp. 387-390, 2014. [Google Scholar]

[57] Nitish Puri, Harsh Deep Prasad, and Ashwani Jain, "Prediction of Geotechnical Parameters Using Machine Learning Techniques," *Procedia Computer Science*, vol. 125, pp. 509-517, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[58] Mohammed Amin Benbouras et al., "A New Approach to Predict The Compression Index Using Artificial Intelligence Methods," *Marine Georesources & Geotechnology*, vol. 37, no. 6, pp. 704-720, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[59] Achal Bhardwaj, and Vijay Kumar, "Prediction of Coefficient of Compression of Soil Using Artificial Neural Network," *Geotechnical Characterization and Modeling*, vol. 85, pp. 917-926, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[60] Scott Kirts et al., "Soil-Compressibility Prediction Models Using Machine Learning," *Journal of Computing in Civil Engineering*, vol. 32, no. 1, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[61] Mohammed el Amin Bourouis, Abdeldjalil Zadjaoui, and Abdelkader Djedid, "Contribution of Two Artificial Intelligence Techniques in Predicting the Secondary Compression Index of Fine-Grained Soils," *Innovative Infrastructure Solutions*, vol. 5, no. 96, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[62] Yusuf Erzin et al., "Prediction of Compression Index of Saturated Clays Using Robust Optimization Model," *Journal of Soft Computing in Civil Engineering*, vol. 4, no. 3, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[63] Ramachandiran Saisubramanian, and V. Murugaiyan, "Prediction of Compression Index of Marine Clay Using Artificial Neural Network and Multilinear Regression Models," *Journal of Soft Computing in Civil Engineering*, vol. 5, no. 4, pp. 114-124, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[64] Worku Firomsa Kabeta, Fekadu Fufa Feyessa, and Yerosan Feyissa Keneni, "Numerical Modelling for Prediction of Compression Index from Soil Index Properties in Jimma Town, Ethiopia," *U. Porto Journal of Engineering*, vol. 8, no. 6, pp. 102-120, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[65] Tsang Long et al., "Tree-Based Techniques for Predicting the Compression Index of Clayey Soils," *Journal of Soft Computing in Civil Engineering*, vol. 7, no. 3, pp. 52-67, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[66] Yu Huat Chia, Danial Jahed Armaghani, and Sai Hin Lai, "Predicting Soil Compression Index Using Random Forest and Gradient Boosting Tree," *Proceedings of the Chinese Institute of Engineers*, Taipei, Taiwan, vol. 1, pp. 31-40, 2023. [Google Scholar] [Publisher Link]

[67] Huifen Liu, Peiyuan Lin, and Jianqiang Wang, "Machine Learning Approaches to Estimation of the Compressibility of Soft Soils," *Frontiers in Earth Science*, vol. 11, no. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[68] R. Akshaya, and K. Premalatha, "Prediction of Soil Compression Index Using SVM and kNN," *Proceedings of IOP Conference Series: Earth and Environmental Science, International Conference on Creative and Innovative Solutions in Civil Engineering*, Jaipur, India, vol. 1326, no. 1, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[69] Sungyeol Lee et al., "A Study on Developing a Model for Predicting the Compression Index of the South Coast Clay of Korea Using Statistical Analysis and Machine Learning Techniques," *Applied Science*, vol. 14, no. 3, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[70] Mintae Kim, Muharrem A. Senturk, and Liang Li, "Compression Index Regression of Fine-Grained Soils with Machine Learning Algorithms," *Applied Science*, vol. 14, no. 19, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[71] Qi Ge et al., "Explainable Ensemble Learning Approaches for Predicting the Compression Index of Clays," *Journal of Marine Science and Engineering*, vol. 12, no. 10, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[72] Ali Ulvi Uzer, "Accurate Prediction of Compression Index of Normally Consolidated Soils Using Artificial Neural Networks," *Buildings*, vol. 14, no. 9, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[73] Pijush Samui et al., "A New Approach of Hybrid Bee Colony Optimized Neural Computing to Estimate the Soil Compression Coefficient for a Housing Construction Project," *Applied Science*, vol. 9, no. 22, pp. 1-18, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[74] Manh Duc Nguyen et al., "Hybridization of Differential Evolution and Adaptive-Network- Based Fuzzy Inference System in Estimation of Compression Coefficient of Plastic Clay Soil," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 149-146, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[75] Giovanni Spagnoli, and Satoru Shimobe, "Statistical Analysis of Some Correlations between Compression Index and Atterberg Limits," *Environmental Earth Sciences*, vol. 79, 2020. [CrossRef] [Google Scholar] [Publisher Link]