Bulletin of Engineering Geology and the Environment (2024) 83:88
https://doi.org/10.1007/s10064-024-03587-z

ORIGINAL PAPER

Prediction of void ratio and shear wave velocity for soil in quaternary alluvium using cone penetration tests

Priyam Mishra[1] · Pradipta Chakraborty[1]

Received: 6 July 2023 / Accepted: 19 February 2024 / Published online: 28 February 2024

Abstract
Several correlations are available to determine the shear wave velocity using cone penetration test (*CPT*) data. Available correlations are applied for the studied region, which shows the requirement for developing a new correlation for the study area. This study uses *CPT*, standard penetration test (*SPT*), and multichannel analysis of surface wave (*MASW*) data to formulate correlations for predicting void ratio (e) and shear wave velocity (V_S). The estimated void ratio at various depths was taken from the *SPT* bore log available for the site. A regression model has been formulated for predicting e from normalized cone tip resistance (Q_{tn}). In developing the shear wave velocity prediction model, two types of cone data are used: mechanical and electrical. In the prediction model of V_S, various parameters, such as cone tip resistance (q_c), soil behavior type index (I_C), effective stress (σ_0'), e, and depth (z), are considered. The correlation regarding shear wave velocity gives a good prediction with both *CPT* cones. A cone factor (K_C) is introduced in the developed correlation for predicting 'e'. The proposed correlations allow design soil parameters to be easily obtained from cone penetration test data.

Keywords Void ratio · Shear wave velocity · Quaternary alluvium · Cone penetration test · Regression model

Introduction

The shear wave velocity (V_S) is an essential property used in dynamic analysis and is related to the stiffness of the soil. Site characterization (determination of site condition), liquefaction hazard assessment, seismic hazard analyses and ground response analyses use shear wave velocity (V_S) as input parameters. Site classification and liquefaction hazard assessment can be performed for a city (Chakrabortty et al. 2018) or a region (Wang et al. 2017). The expected ground motion estimated from probabilistic or deterministic seismic hazard assessment also required knowledge of shear wave velocity (Al-Ajamee et al. 2022). Therefore, it is imperative to accurately measure or predict shear wave velocity (V_S) for seismic design purposes. Undisturbed samples need to be collected to accurately estimate the soil properties in the laboratory. However, collecting undisturbed samples is often not possible or very difficult. A disturbed soil sample will not give the actual value of V_S, as the soil structure will change, and particles will be oriented in a different configuration. The behavior of altered soil particles will ultimately differ from that of soil deposits, and the properties determined will not depict the actual information. Direct field measurement of velocity should be taken for determining stiffness parameters such as shear modulus or Young's modulus, as it provides convenient and reliable results (Jardine et al. 1986). Various laboratory methods, such as resonant column and bender element tests, as well as field methods such as geophysical techniques, are used to determine shear wave velocity (Gu et al. 2015). In a study by Nilay et al. (2022), three different in situ tests, CPT, SPT, and MASW, were considered for liquefaction hazard mapping. The conclusion drawn was that a CPT-based assessment tends to yield conservative liquefaction potential results for sites within the studied region. Reflecting on the previous discussion, it becomes evident that numerous direct and indirect methods exist for determining V_S. The choice of the appropriate method

Priyam Mishra and Pradipta Chakrabortty contributed equally to this work.

✉ Pradipta Chakrabortty
pradipt@iitp.ac.in

Priyam Mishra
priyam_2121ce10@iitp.ac.in

[1] Department of Civil and Environmental Engineering, Indian Institute of Technology Patna, Bihta, Bihar 801106, India

depends on the specific requirements of the task at hand. Therefore, developing a multivariable nonlinear regression prediction model (Das and Chakrabortty 2022) between V_S and cone test parameters is beneficial, as it gives a reasonably accurate measurement of V_S for both region-wide and site-specific ground response assessments (McGann et al. 2015). Shear wave velocity is affected by soil type, aging conditions, cementation properties, and effective stress (Andrus et al. 2007). The prediction model can consider aging conditions and cementation properties using effective stress and void ratio terms. The value of V_s in soil deposits of the Pleistocene age is greater than that of the Holocene age. This difference influences the researchers to introduce an age scaling factor (*SF*) in the correlation. Many *CPT*-V_S correlations are available worldwide for different sites with different cone parameters. Initially, correlation model was proposed using only two parameters by Baldi et al. (1990). Later on, three parameters were considered by Hegazy and Mayne (1995) and Andrus et al. (2007). Subsequently, models involving four parameters were introduced by Hegazy and Mayne (2006) and Robertson (2009). Some of these correlations applicable to the study area (given in Table 1) have been used to predict *Vs* for the studied soil. Hegazy and Mayne (1995) formulated three correlations (considering sand, clay, and all soil) for *Vs* determination. Hegazy and Mayne (2006) selected a site with relatively complex stratigraphy and proposed a global correlation for V_S determination.

The soil behavior index (I_C) was considered in this correlation. Robertson (2009) gave a global relation for V_S as a function of cone tip q_c, soil behavior index I_C, and effective vertical stress σ'_0. Mousa and Hussein (2022) most recently, provided seven (7) different correlations for shear wave determination using *CPT*. From the literature, it has been inferred that the *Vs-CPT* correlation improves considering *e*. The void ratio articulates the denseness of strata. It is closely related to soil compressibility, permeability, and shear strength and depends upon the particle size and distribution of particle size. As an important parameter, a direct method for estimating '*e*' at a desired depth is not available. Therefore, *CPT* data can be used as an effective way to estimate '*e*'.

A correlation model between *CPT* and '*e*' was developed in this study with available *CPT* and *SPT* data to materialize this concept. A power regression model was fitted with the available data between the factored void ratio (*FVR*) and normalized cone tip resistance. The factored void ratio is defined here as $e^{0.5}$ multiplied by $(I_C)^n$. A correction factor has been proposed to consider the effect of cone type. This proposed model is one of the novelties of the present study. In the next part of this study, two site-specific prediction models have been proposed for estimating *Vs* from *CPT* data. The first *Vs-CPT* model has been presented with four parameters (q_c, I_C, σ'_0 and z). The second model has been proposed with five parameters (q_c, I_C, σ'_0, z and e) based on regression analyses. The need for the proposed site-specific models for estimating shear wave velocity from *CPTs* is explained in Sect. "Shear wave velocity (VS) prediction model".

Study area

Geology

The data used in this study were collected from different soil reports available for the IIT Patna campus (Fig. 1). The study area lies in the alluvium plain of Ganga and its tributaries with the most recent geologic age termed Quaternary alluvium (Sahu et al. 2015). This Quaternary

Table 1 Applicable correlations between CPT and Vs used for the study area

V_S	R^2	RMSE	No of samples	Geologic age	Study
$(10.1\log(q_c) - 11.4)^{1.67}.(\frac{f_s}{q_c} \times 100)^{0.3}$	0.695	–	323	Quaternary	Hegazy and Mayne (1995)
$32.3{q_c}^{0.089}.{f_s}^{0.121}.z^{0.215}$	0.73	–	60	Holocene	Piratheepan (2002)
$0.0831.q_{c1n}.(\frac{\sigma'_0}{P_a})^{0.25}.e^{1.786I_c}$	0.854	–	558	–	Hegazy and Mayne (2006)
$118.8\log(f_s) + 18.5$	0.82	–	161	Quaternary	Mayne (2007)
$2.62{q_t}^{0.395}{I_c}^{0.912}z^{0.124}.SF$	H – 0.73 P – 0.43	–	185	Holocene (H) & Pleistocene (P) age	Andrus et al. (2007)
$10^{(0.55I_c+1.68)}.(q_t - \sigma_v)^{0.25}$	–	–	1035	Quaternary	Robertson (2009)
$18.4{q_c}^{0.144}{f_s}^{0.0832}z^{0.278}$	–	–	513	–	McGann et al. (2015)
$100(1.4 + 1.59f_s + 0.09q_c - 1.33{f_s}^2 - 0.002{q_c}^2 + 0.05f_sq_c)$	–	27.78	37	–	Mola-Abasi et al. (2015)
$10.915{q_t}^{0.317}.{I_c}^{0.210}.z^{0.057}.SF^a$	0.798	13.11	–	–	Zhang and Tong (2017)

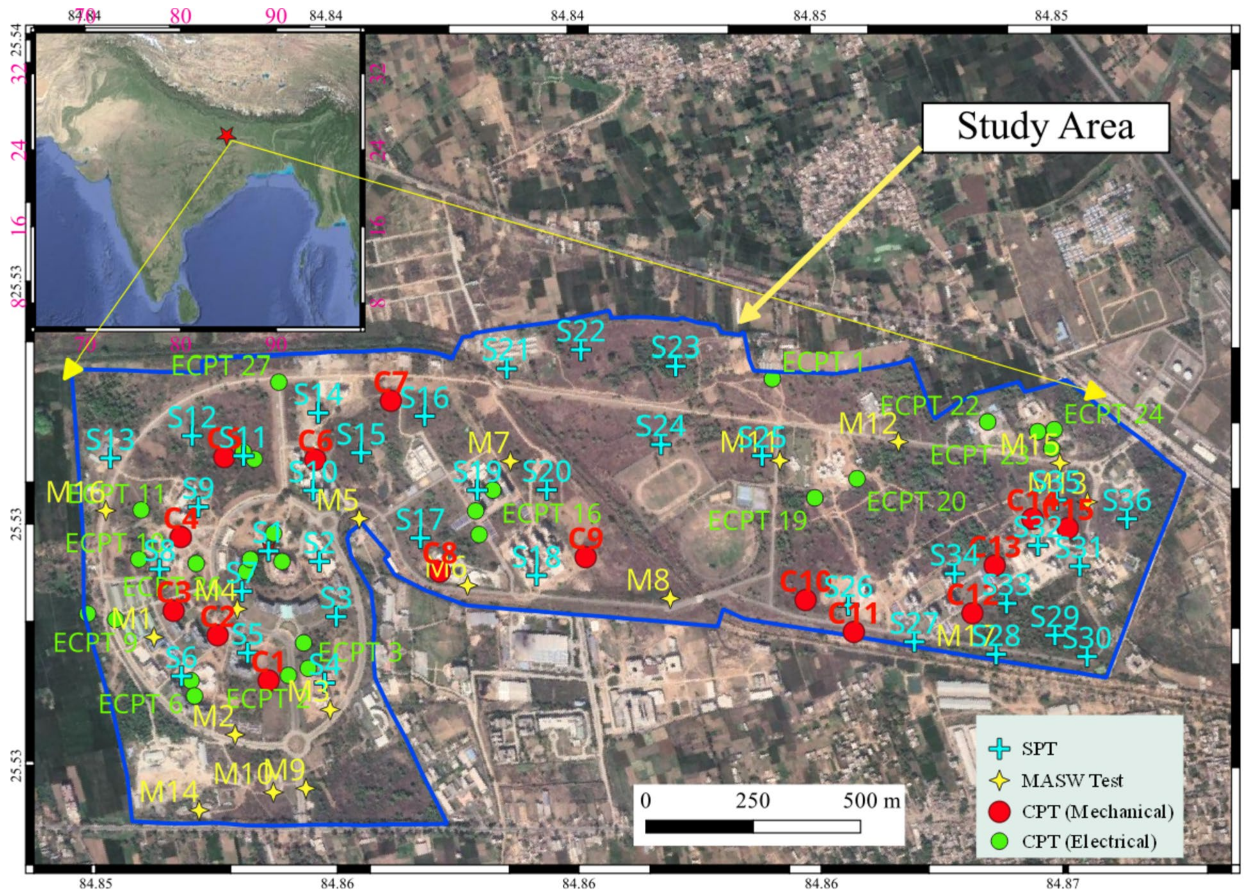


Fig. 1 Locations of CPT, SPT, and MASW tests marked by various symbols on the study area (IIT Patna campus) map prepared by modifying the google map

alluvium refers to sedimentary deposits formed in the most recent geological period through the action of flowing water from the Ganga River and its tributaries, such as the Sone, Gandak, and Koshi Rivers. The entire region lies in the Middle Ganga Plain (*MGP*), which has an almost flat topography. The geology of the studied region is influenced by fine sand particles deposited by the Sone River. This is the only river flowing in this area is dynamic in nature as mentioned by Sahu et al. (2010). The sediment type found near Sone is locally called Sone sand, which contains fine to medium fine-grained sand and gravel with a size range of 0.15 to 1.18 mm. Generally, these sediments can be found in various settings, including floodplains, deltas, alluvial fans, and terraces. The tectonics of the study area lie in an alluvial plain, an active tectonic region underlain by transverse and oblique faults. Two significant faults, namely, the East Patna and West Patna faults, are considered the most active because of the continuous subsidence of the Indian plate into the Eurasian plate. Hence, to thoroughly understand the region's seismicity, proper estimation of dynamic soil properties is essential.

Database for formulating prediction models

The database is formed by collecting data from three different types of testing, namely, i) *CPT*, ii) *SPT*, and iii) *MASW* testing. The collected *CPT* data have two different types of cones, namely, mechanical and electrical cones. From *CPT*, two essential readings are obtained: q_c and sleeve friction (f_s). The collected *CPT* data using a mechanical cone were obtained from existing soil reports containing data at 15 locations on the campus with a penetration depth of 30 m. The collected ECPT data were obtained from soil reports conducted at 27 sites with a maximum penetration depth of approximately 20 m. Continuous readings are available in electrical cone penetration testing (*ECPT*). As velocity measurements taken from *MASW* are available at every 1-m depth until 30 m depth, *CPT* readings are also selected at the same level with 1-m intervals from both the *ECPT* and mechanical cone. Data from *SPT* testing near that of *CPT* and *MASW* testing are considered to determine properties such as unit weight, *e*, etc., along the depth. A comparison of q_c, shear wave velocity (V_S), and *SPT* N-value (N) along depth is shown in Fig. 2. The test results shown in Fig. 2 for a particular location (e.g., C1-M2-S5) are close to

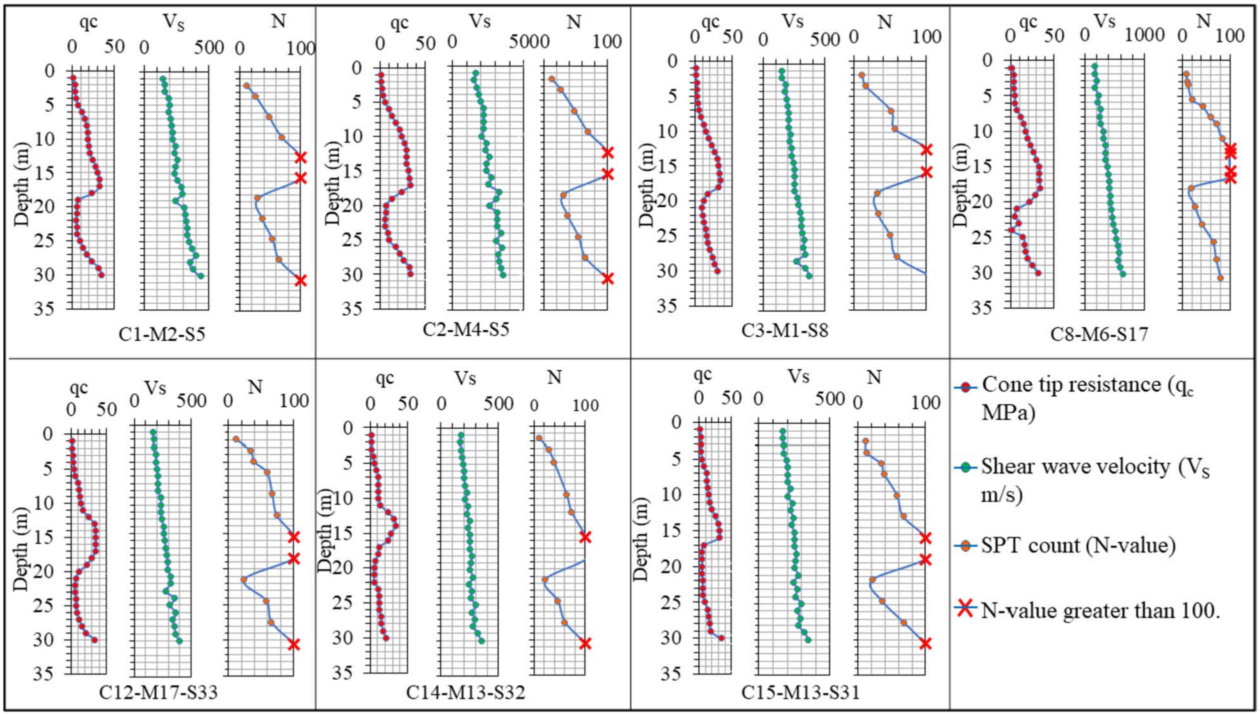


Fig. 2 Comparison between recorded cone tip resistance (q_c), SPT N-value (N), and shear wave velocity (V_S) that are close to each other along with depth at seven locations (e.g., C8, M6 and S17 are tested at nearby locations as shown in Fig. 1)

each other, with a maximum of 150 m apart. In *SPT* profiling, some distinct markers are shown in red, and these red marker values are those with an N-value equal to or greater than 100.

Soil classification

Cone penetration testing has been in use for nearly 40 years. It has a sound theoretical aspect and a simplified testing procedure. In this testing, a cone penetrates into the soil. The resistance offered by the soil to the cone gives essential results, which are called cone parameters. With much advancement in the past years using *CPT*, information regarding soil properties such as soil type, behavior, and strength can be obtained very quickly. Soil stratigraphy and soil type classification are significant applications of *CPT*. Using *CPT*, early soil identification charts were given by Douglas and Olsen (1981). Later, normalized and nonnormalized charts provided by Robertson (1990) and his coworkers gained much popularity. Robertson (1990) proposed the concept of normalization for the cone tip and friction ratio, which is shown in Eq. (1) and (2) as follows:

$$Q_{t1} = [\frac{q_t - \sigma_0}{\sigma'_0}] \quad (1)$$

where, q_t is the cone tip resistance, σ_0 is the total stress and σ_0' is the effective stress at the tested depth.

$$F_r = \left[\frac{f_s}{q_t - \sigma_0}\right] \times 100(\%) \quad (2)$$

where, f_s is the sleeve friction, and F_r is the friction ratio. Robertson and co-workers (Robertson and Wride 1998; Zhang et al. 2002) proposed an modified version of Eq. (1), which introduces a normalized cone tip resistance, expressed as:

$$Q_{tn} = [\frac{q_t - \sigma_0}{p_a}] \times (\frac{p_a}{\sigma'_0})^n \quad (3)$$

where, Q_{tn} is the normalized cone tip resistance; n is the stress exponent; and P_a is atmospheric pressure. Jefferies and Davies (1993) introduced I_c (soil behavior type index) to characterize the soil zone in Qt_1—F_r charts, defining it as the boundary in terms of the radius of concentric circles. Robertson and Wride (1998) provided an equation, Eq. (4) For these concentric circles and the updated Robertson's 1990 chart. These charts are plotted between the normalized cone tip resistance (Q_{tn}) and friction ratio (F_r), and the entire graph is divided into nine different soil zones. Each soil zone provides information about the soil type in that stratum corresponding to a range of I_C values.

$$I_C = [(3.47 - \log Q_{t1})^2 + (\log F_r + 1.22)^2]^{0.5} \quad (4)$$

The stress exponent (n) in Eq. 3 is a function of I_C and overburden pressure and is given by the following equation:

$$n = 0.381 I_C + 0.05\left(\frac{\sigma'_0}{p_a}\right) - 0.15;\ \text{Where } n \le 1. \tag{5}$$

For soil classification using *CPT*, I_C is estimated using the abovementioned equations. It is an iterative process that starts with steps from Eq. 1 to Eq. 5. This iteration will begin by assuming an initial stress exponent 'n' equal to 1. It will stop when the change in two consecutive 'n' values is less than 0.01 ($\Delta n < 0.01$). The Δn is the change observed in 'n' from two successive observations. When the difference in n is below or equal to 0.01, Q_{tn} and I_C at that stage will be termed final values.

Identification of soil strata is completed based on I_C values, in which soil ranges from Zone-3 to Zone-6. Zone 3 belongs to the clayey soil type with an I_C value between 2.95 and 3.6, and Zone 6 belongs to the purely clean sandy type soil (Table 2). From Fig. 3, it can be observed that several datasets lie in the zone of silt mixtures to sand mixtures. The I_C value at all the locations is calculated and plotted along the depth (Fig. 4). It consists of a silt mixture in the first four meters of depth and a sandy mixture in the next eight to ten meters. A sharp change in the I_C value shows a sudden shift in stratigraphy. This abrupt change indicates a layer of different material present at that depth. The measured q_c value is affected by the presence of these thin layers. At the interface, the cone senses these thin layers before entering them from a certain distance. The transition effect induced variation in the q_c value. At the boundary, q_c is affected by both layers, i.e., layer ahead and layer behind. This variation continues up to a certain depth in the next layer. This effect is termed the "thin layer effect".

Table 2 Soil classification based on soil behavior type index (Robertson and Wride 1998)

I_C value	Zone	Soil behavior type
$I_C < 1.31$	7	Gravelly sand to dense sand
$1.31 < I_C < 2.05$	6	Sands: clean sand to silty sand
$2.05 < I_C < 2.60$	5	Sand mixtures: silty sand to sandy silt
$2.60 < I_C < 2.95$	4	Silt mixtures: clayey silt to silty clay
$2.95 < I_C < 3.60$	3	Clays: silty clay to clay
$I_C > 3.60$	2	Organic soils: peats

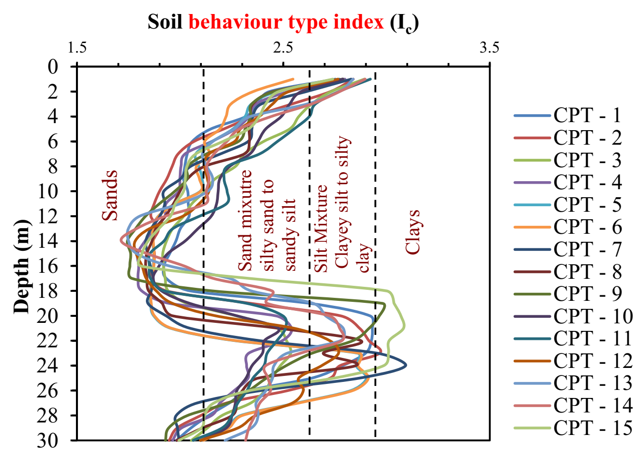


Fig. 4 Variation of soil behavior type index (I_c) along with the depth in the studied area

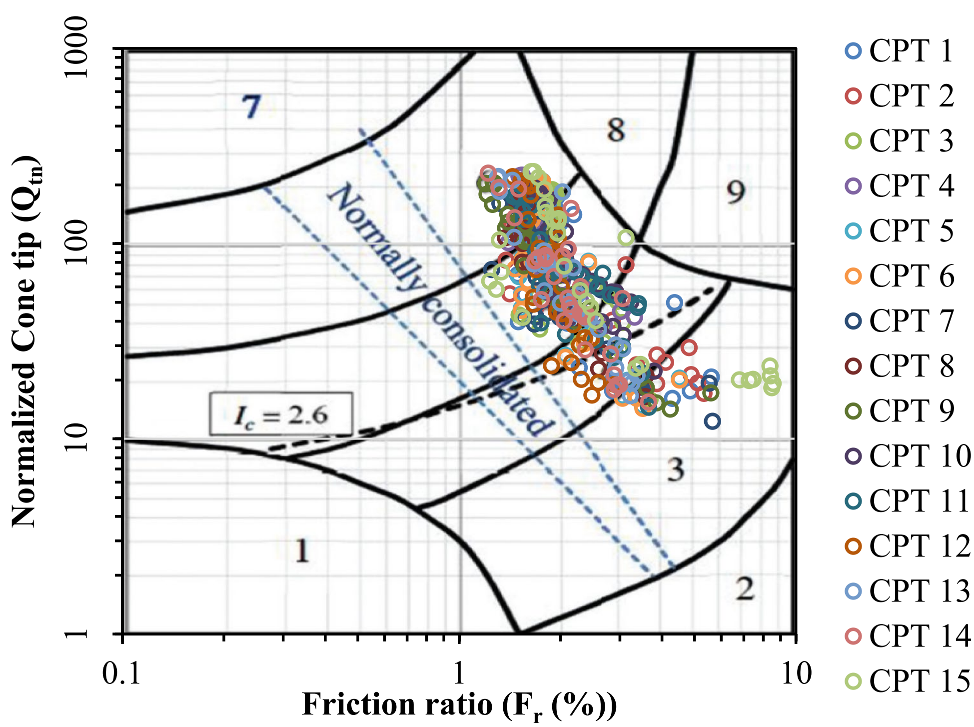


Fig. 3 Soil classification using normalized cone tip resistance (Q_{tn}) and friction ratio (F_r) based on Robertson and Wride (1998)

Void ratio (*e*) Prediction model

Prediction using mechanical cone

The void ratio is usually estimated from laboratory tests of collected soil samples from *SPT*. To eliminate the dependency on *SPT*, a prediction model for '*e*' from *CPT* data was proposed in this study. A rigorous statistical analysis was conducted using the available data, and it was found that a power relation exists between $e^{0.5}I_C^{\ n}$ and Q_{tn}. The term plotted on the Y-axis in Fig. 5b, i.e., $e^{0.5}I_C^{\ n}$, is called the Factored Void Ratio (*FVR*) here. The actual measured void ratio needed for correlation formulation is obtained from the results of *SPT* testing available for the site. It has been assumed that there is no/little change in soil properties within small distances between *SPT* and nearby *CPT* locations (within a distance of 150 m.). From that location, cone tip parameters are chosen at a depth of known void ratios. A total of 194 *CPT* data points are gathered from the mechanical cone, and regression analysis is carried out. The datasets used are shown as bar charts with individual counts and their respective *CPT* locations in Fig. 5a. The trend between the *FVR* and I_C is shown in Fig. 5b. The functional form of the equation obtained between the *FVR* and Q_{tn} is also shown in the figure. The proposed *CPT-e* correlation with an R^2 value of 0.92 is given below:

$$e^{0.5}I_c^{\ n} = 5.2976 \times Q_{tn}^{\ -0.274} \tag{6}$$

Validation of the prediction model for mechanical cone data

In this section, the formulated *CPT-e* correlation given in Eq. 6 is validated with available *CPT* data from different locations on campus, which was not used in developing the relationship. Figure 6a compares the measured and predicted *FVR* against depth; both values were almost identical. Figure 6b corresponds to the predicted *FVR* and *FVR* estimated (based on the estimated '*e*'). Most of the data points are on a 45° degree line or nearby. The data points chosen for validation using mechanical cones agree well with the proposed model. However, electrical cone penetration test (*ECPT*) data, when validated using the proposed model, show a downward shift in the *FVR* compared to the proposed correlation (as shown in Fig. 7). Cone tip resistance (q_c) values from both cones at nearly the same site are plotted against depth. The values of q_c are similar, and not much noticeable change is detected. Additionally, the same vertical soil profile was obtained from the I_C value calculated from two different types of *CPT* testing parameters. After checking all the relevant parameters, a downward shift in the plotted value of *FVR* and Q_{tn} with *ECPT* was present. All other parameters are nearly identical; the only change was that an electrical cone is used in *ECPT*, which is entirely different from the mechanical cone. This observed downward shift may be because of this changed cone type. This necessitates

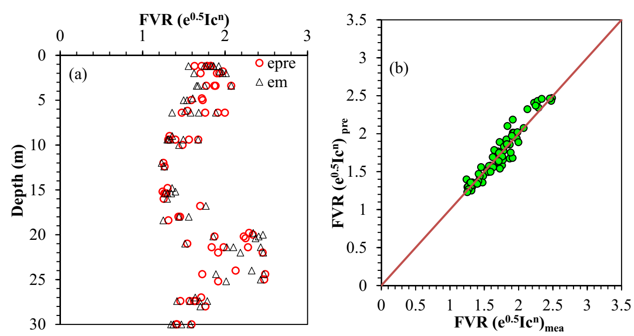


Fig. 6 Validation of void ratio prediction model developed using mechanical cone data (**a**) variation of factored void ratio (FVR) along with depth, (**b**) predicted and estimated FVR plotted along with 45° line

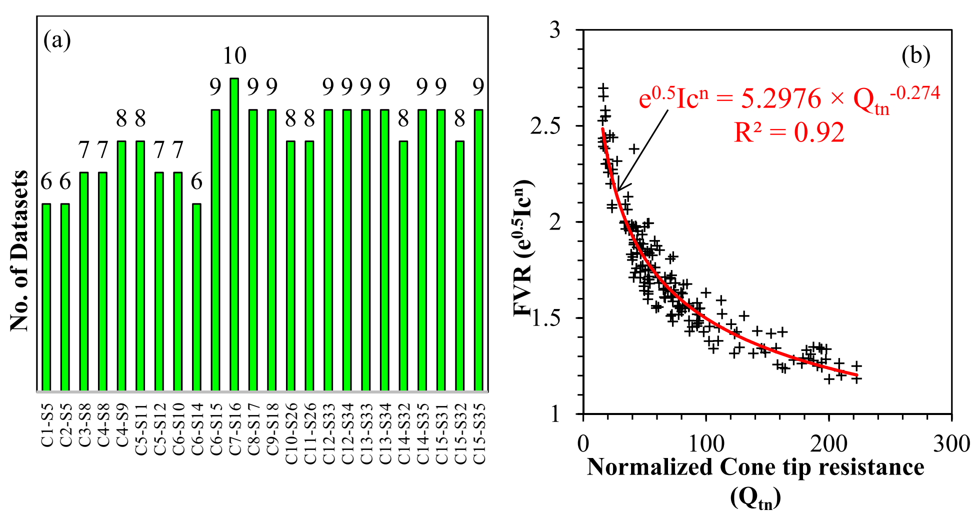


Fig. 5 (**a**) Dataset distribution (total count: 194) at various locations (e.g., C1 is CPT test location 1 and S5 is SPT test location as shown in Fig. 1) used for formulating void ratio prediction model, (**b**) Trend between factored void ratio (FVR) and normalized cone tip resistance (Q_{tn})

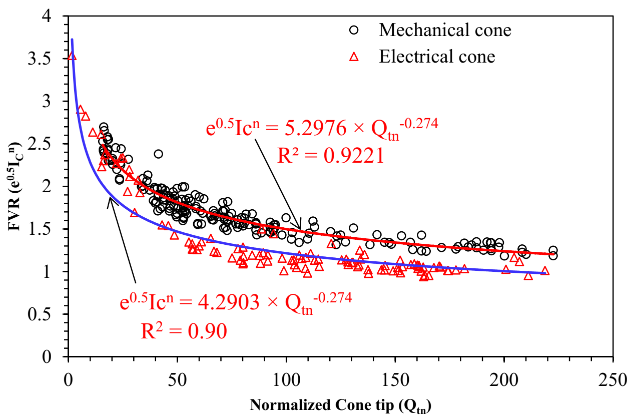


Fig. 7 Estimation of cone type effect on power relation between factored void ratio (FVR) and normalized cone tip resistance (Q_{tn})

introducing a particular factor to the formulated correlation that will take care of this cone effect, known as the cone factor (K_C). The trend line is plotted across the data between *FVR* and Q_{tn} from *ECPT* to calculate the cone factor. The *ECPT* data, collected from nearly the exact location of mechanical testing, are used, and *FVR*-Q_{tn} is plotted, which gives the following equation:

$$e^{0.5}I_c^{\ n} = 4.2903 \times Q_{tn}^{\ -0.274} \quad (7)$$

By comparing Eq. 6 and Eq. 7, the cone factor for the electrical cone is proposed as follows:

$$K_c = \frac{4.2903 \times Q_{tn}^{\ -0.274}}{5.2976 \times Q_{tn}^{\ -0.274}} = 0.809 \quad (8)$$

Therefore, the proposed mechanical and electrical cone factors are 1 and 0.809, respectively. This factor can be multiplied by Eq. 6 to predict '*e*' from various cones. The modified predicted model takes the following form:

$$e^{0.5}I_c^{\ n} = 5.2976 \times Q_{tn}^{\ -0.274} \times (K_C) \quad (9)$$

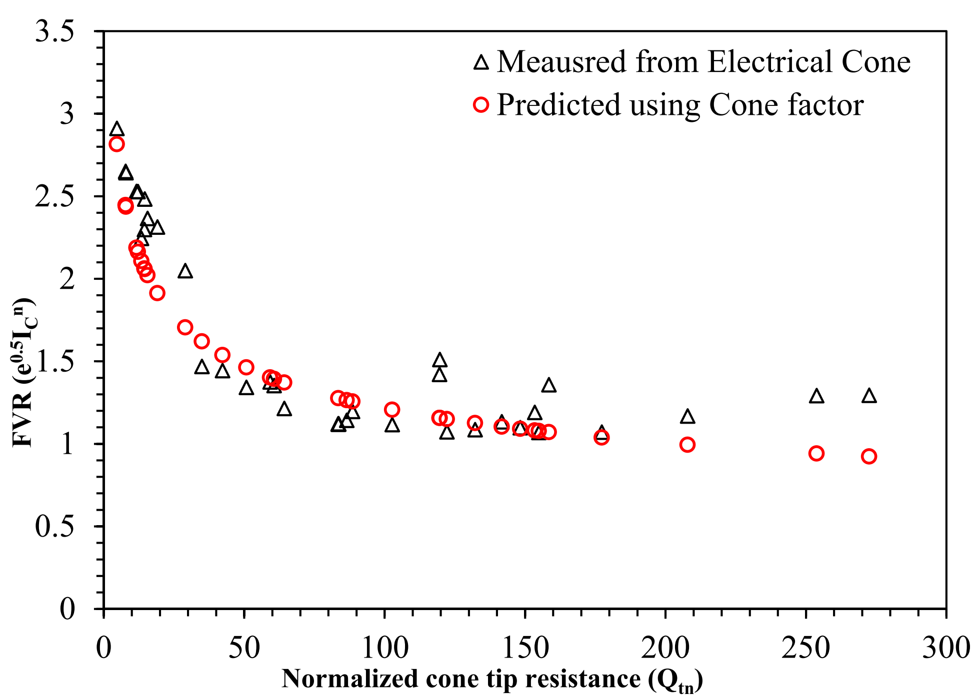
Validation of the prediction model for *ECPT* data

To validate the proposed model, values obtained from electrical cones at different testing locations are used in Eq. 9 considering the cone factor (K_C). The predicted and estimated *FVR* from *ECPT* data is shown in Fig. 8. In Fig. 9a, *FVR* predicted and measured values lie on or near the 45-degree line, with a comparison of *FVR* along depth in Fig. 9b. Both figures show the excellent predictability of *e* from the proposed model.

Shear wave velocity (*VS*) prediction model

The correlation between cone penetration testing (CPT) and shear wave velocity (V_S) is commonly represented in literature through various forms: linear, as demonstrated by Sykora and Stokoe (1983), nonlinear with a single

Fig. 8 Validation of void ratio prediction model using the ECPT data with consideration of cone factor



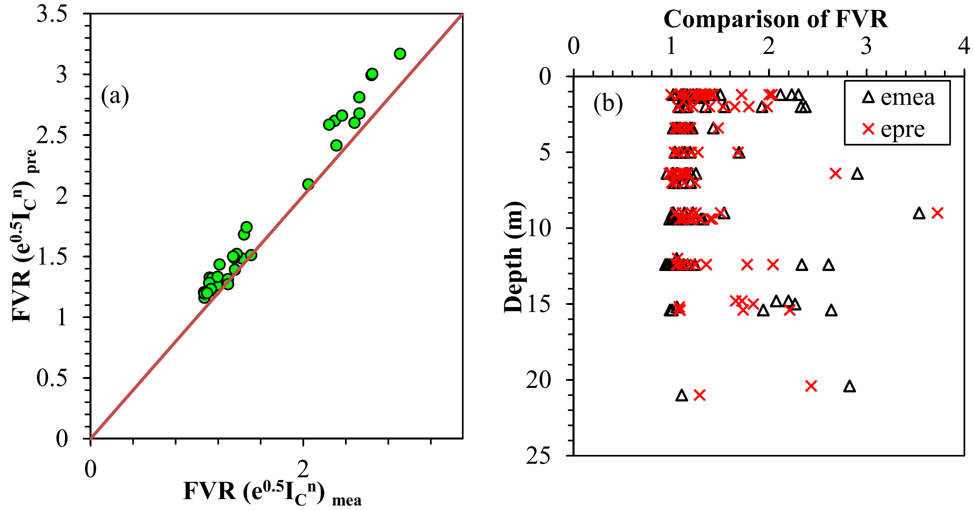


Fig. 9 (**a**) Efficiency of model showing predicted and measured FVR, (**b**) Validation of FVR along depth using electrical Cone

parameter proposed by Jaime and Romo (1988), or nonlinear with multiple parameters as explored by Robertson (2009). Some existing relations, such as the one by Andrus et al. (2007), adopt a power law equation, while others, like those presented by Hegazy and Mayne (1995, 2006) rely on logarithmic relationships. However, this simple power law equation with a single parameter does not work well, as previous studies show that V_S depends upon many factors other than soil type and testing conditions. For this reason, a nonlinear model with multiple variables gives better efficiency in prediction. These multiple variables are direct (e.g., q_c) or indirect cone parameters (e.g., I_c) or in situ soil properties (e.g., total or effective stresses). A series of correlation models available for predicting *Vs* based on q_c data for the Quaternary alluvial deposit available in the studied region. Some of these applicable models, as presented in Table 1, have been used to predict *Vs*. The predicted *Vs* (V_{pre}) and measured *Vs* (V_{mea}) from *MASW* for locations C2-M5 are shown in Fig. 10. The result indicates that none works well and cannot accurately provide the V_S. The difference between V_{pre} and V_{mea} along with the depth shows the requirement of developing a site-specific model for the studied region. Using the functional form of the correlation, a simple *CPT*-V_S correlation has been proposed by Mishra et al. (2023) that is applicable to the same study region. However, in that study, a limited number (90 datasets from 3 different locations) of available datasets was used. Therefore, to improve the prediction model, large datasets (453 pairs) from both testing types (mechanical and electrical) are combined, forming an updated *CPT*-V_S correlation.

Existing models have been considered for the selection of the functional form of the model and parameters to be incorporated in the model. The *CPT* and *Vs* relationship has been investigated by various researchers since the 1980s (Robertson and Campanella 1983; Robertson et al.

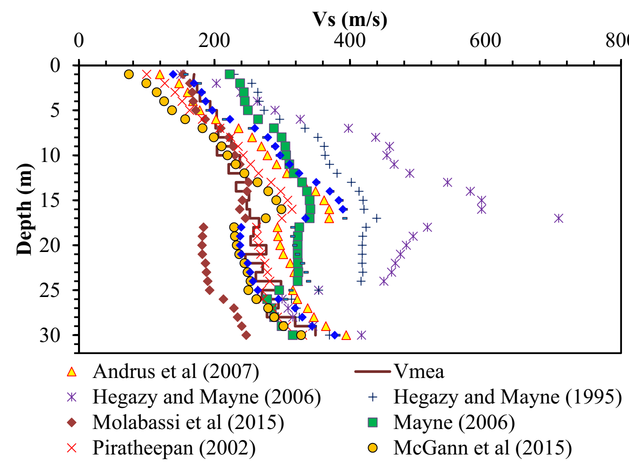


Fig. 10 Variation of shear wave velocity along depth estimated using different available CPT-V_S models and experimentally measured for the studied site

1986; Hegazy and Mayne 1995; Mayne and Rix 1995). One of the limitations of these earlier-developed models is that most of these relations are valid for either sand or clays. Later, these have been addressed by including parameters such as *Ic* and *e,* which relate the soil type with predicted *Vs* values (Piratheepan 2002; Andrus et al. 2007; Robertson 2009; Long and Donohue 2010; Gadeikis et al. 2013; Cai et al. 2014; Sara 2014; Ahmad et al. 2015; Mola-Abasi et al. 2015; McGann et al. 2015; Abbaszadeh Shahri and Naderi 2016; Mohamed Ahmed and Ahmed 2017; Zhang and Tong 2017; Tun and Ayday 2018; Fayed and Mousa 2020; Yang et al. 2022; Mousa and Hussein 2022; Khan et al. 2022; Mishra et al. 2023). Two models have been proposed in the following subsections, one without considering *e* (correlation model 1) and another considering *e* (correlation model 2).

Correlation model 1

First, a multiparameter regression model was formulated to predict *Vs* based on the *CPT* data. Both mechanical and electrical cone test data have been used to formulate *CPT*-V_S correlation for the study area. The shear wave velocity (V_S) has been estimated from *MASW* tests conducted at different locations and reported in the literature (Nilay et al. 2022). A total of 453 data pairs from 33 *CPT* and 16 *MASW* sites were considered. The site locations and the number of datasets at each location are shown in Fig. 11. The data pairs are very close, having a maximum distance of 100 m. Statistical regression analysis has been performed on these datasets, and a nonlinear multivariable equation has been proposed to predict V_S. The proposed equation is as follows:

$$V_s=156.885q_c^{0.033}I_c^{0.120}\sigma'^{-0.169}z^{0.366} \quad (10)$$

where, q_c is cone tip resistance, I_C is soil behavior type index, $\sigma_0\prime$ is effective stress at particular depth, and z is the depth. The predictive equation for shear wave velocity (V_S) is derived through non-linear regression incorporating multiple variables. This choice is informed by an evaluation of existing models such as Hegazy and Mayne (1995), and Andrus et al. (2007). Additionally, insights from prior research highlight the significance of cone tip resistance (q_c) as a pivotal parameter associated with the undisturbed shear strength of the soil, as demonstrated by Hegazy and Mayne (2006). Their findings indicate that cone tip resistance (q_c) exhibits superior variability in predicting V_S compared to sleeve friction (f_s). In the present study, q_c (containing a correlation coefficient (r) of 0.34 with *Vs*) is considered, alongside the soil behavior type index (I_C) in the correlation equation. As soil samples are not extracted during CPT, the inclusion of the I_C (with a correlation coefficient of 0.26 with *Vs*) incorporated valuable insights into soil type and its behavior. The correlation equation also considers other parameters, namely effective stress ($\sigma_0\prime$) and depth (z). The inclusion of these parameters is justified by the high correlation observed between shear wave velocity (*Vs*) with depth ($r=0.95$), as well as effective stress ($r=0.90$). The selection of these variables was made based on correlation coefficient analyses independently. The observed lower *r*-values for (q_c) and (I_C) indicate a non-linear relationship with *Vs*, while effective stress and depth demonstrate a robust linear relationship with *Vs*. A visual examination of the relationship between *Vs* and depth confirms a positive correlation, indicating an increase in *Vs* along depth. A similar relationship is observed with $\sigma_0\prime$. Therefore, the inclusion of all these parameters in the CPT-VS correlation is deemed beneficial.

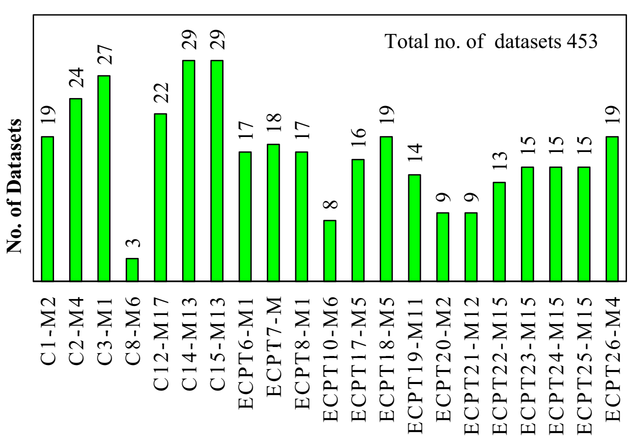


Fig. 11 Number of data pairs at various locations (e.g., C1 is CPT test location 1 and M2 is MASW test location as shown in Fig. 1) used for formulating CPT-Vs prediction model

The above correlation Eq. (10) has a coefficient of determination of 0.858. ANOVA was used to estimate the significance of the model. The proposed model has also been validated for the site using the datasets not included in the model formulation and discussed in subSect. "Validation of the proposed models".

Correlation model 2

As mentioned earlier, including *e* in the prediction model increases the model's efficiency. In some studies, '*e*' is considered an input parameter for the computation of V_S, and the *CPT*-V_S correlation model is significantly improved. Therefore, another prediction model has been proposed in this section considering *e* in the regression model. The uniqueness of this proposed model is that all the parameters used in the regression model can be estimated from the *CPT* tests including void ratio '*e*' presented in Sect. "Void ratio (e) Prediction Model". Therefore, the dependency of the model on other test methods has been eliminated in this proposed model. After incorporating *e* in the equation, the following *CPT*-V_S model has been proposed:

$$V_s=151.859q_c^{0.044}I_c^{0.165}\sigma'^{-0.202}z^{0.397}e^{0.035} \quad (11)$$

The above correlation equation has a coefficient of determination of 0.849. ANOVA was used to estimate the significance of the model.

Validation of the proposed models

The validation of the *CPT*-V_S correlation is essential to ensure the accuracy and reliability of the predictions. Several methods are available to validate *CPT*-V_S correlations, including laboratory and field testing. One standard method is the comparison of V_s measurements obtained from different techniques or instruments. Field testing involves conducting *CPT* and geophysical tests at the same site to compare

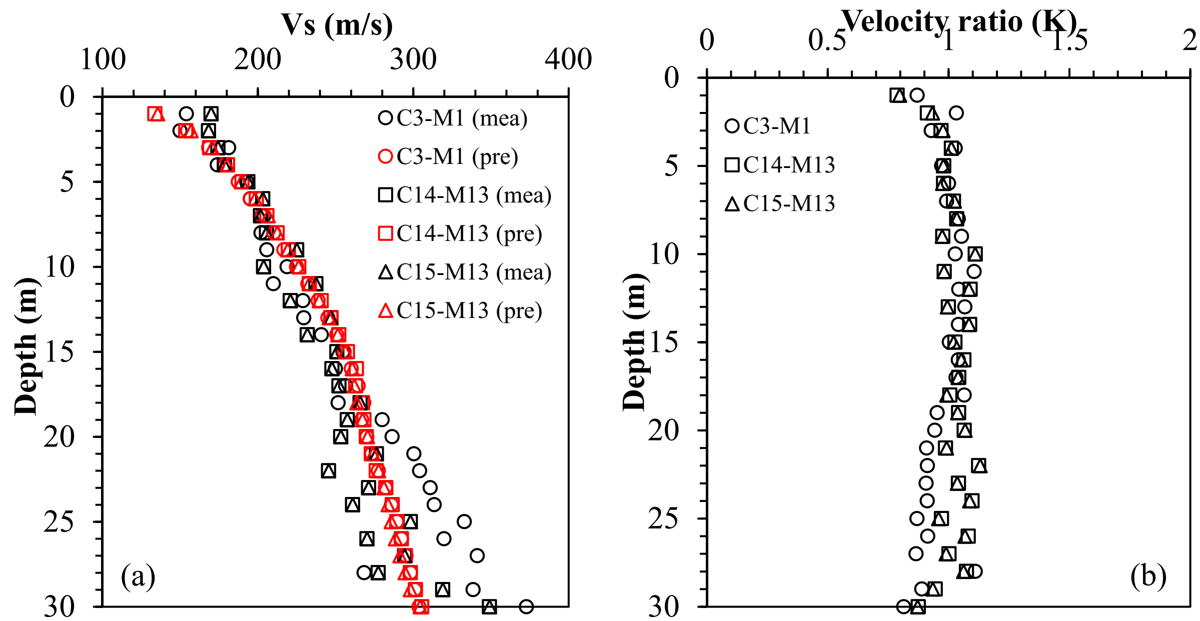


Fig. 12 Validation of prediction model for various locations (C3-M1, C14-M13, C15-M13) inside the studied site using mechanical cone (**a**) shear wave velocity profile, (**b**) velocity ratio (K-value)

the *Vs* directly. The developed *CPT-V*$_S$ correlations have been compared in this section with the *MASW* data to validate the correlations. In this section, the correlations have been validated using the data from some selected locations inside the campus, as shown in Figs. 12a and 13a for correlation model 1. From both types of cones used, the measured

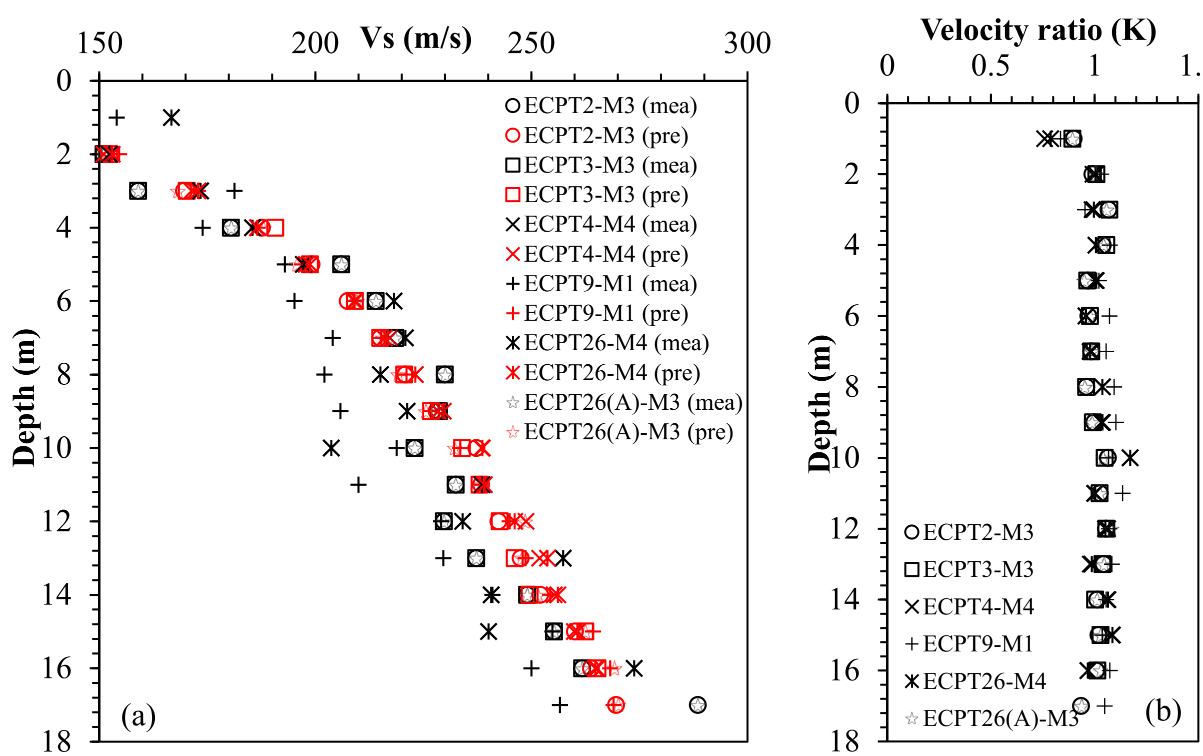


Fig. 13 Validation of prediction model with ECPT data sites (ECPT2-M3, ECPT3-M3, ECPT4-M4, ECPT9-M1, ECPT26-M4, ECPT26(A)-M3) (**a**) in terms of shear wave velocity profile, (**b**) velocity ratio (K-value)

and predicted V_S profiles show good agreement. Not much discrepancy is observed in the predicted and measured data. The absolute percentage difference in the expected value is less than 18% of the calculated value. Another term defined by Zhang and Tong (2017) as the velocity ratio (K), which is the ratio of the predicted to measured velocity, is shown in Figs. 12b and 13b for the mechanical and electrical cones, respectively. The value of K shows the variation in predicted velocity and estimated velocity. The value of K for most data closer to 1 offers the best predictability power of the correlation model. The figure shows that at every level of depth, 'K' values are within approximately 25% of the measured data.

Comparing the two correlation models given in Eq. (10) and (11), no significant improvement is observed in the predicted V_S values with the introduction of *e,* as shown in Fig. 14. There may be two probable reasons for such a minor variation. The first is the consideration of I_C values in both correlations. The *void ratio* expresses the grain compactness. The I_C can also incorporate the effects of soil type and grain compactness. The second reason is the incorporation of σ_0' and q_c in both models. These two parameters already consider the stiffness of the soil. Therefore, including *e* does not increase the accuracy of the model.

Additionally, ANOVA was used on the collected and predicted data to determine the significance of Eq. (10) and (11). The degrees of freedom for the numerator and denominator were suitably determined based on the number of groups and the sample size. The significance level was set at 0.05. The obtained F values in this study are 0.0064 & 0.00068; the F-critical values are 3.854 & 3.854; and the resulting p values are 0.98 & 0.98, respectively. ANOVA for both correlations gives nearly the same F value, F_{cri} value, and p values. Therefore, it is clear that the acquired F value is much smaller when compared with the critical F value (3.854). Therefore, it can be concluded that the observed differences in the means are not statistically significant, or it is likely that the variation in the data can be attributable to random chance or elements unrelated to the treatments under comparison. ANOVA with a p value of 0.98 indicated no statistically significant difference between the observed variation across groups. In other words, the null hypothesis, which states

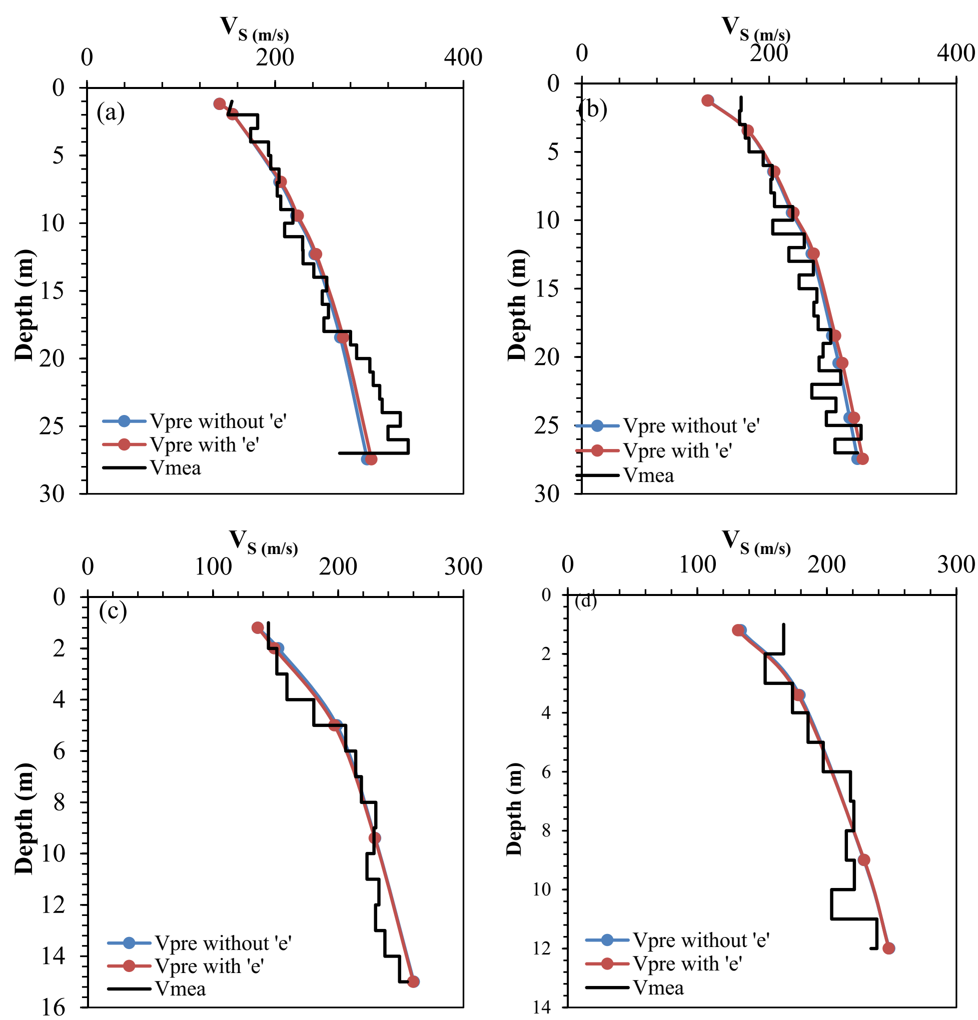


Fig. 14 Comparison of Vs Correlation Models 1 and 2 at different locations (**a**) C3-M1 (**b**) C15-M13 (**c**) ECPT2-M3 (**d**) ECPT4-M4

no significant differences between the groups being compared, is strongly supported.

Another tool for evaluating different regression equations is the computation of residuals for the fitted regression models. Therefore, for this reason, residuals (ε) are computed using the following equation:

$$\varepsilon = \frac{\ln\left(V_{mea}\right) - \ln\left(V_{pre}\right)}{S_{V|X}} \tag{12}$$

where $S_{V|X}$ is an estimate of the conditional standard deviation (Ang and Tang 2007) and defined as:

$$\mathrm{S}_{\mathrm{V|X}} = \sqrt{\frac{\sum\left(\ln\left(\mathrm{V}_{\mathrm{mea}}\right) - \ln\left(\mathrm{V}_{\mathrm{pre}}\right)\right)^2}{\mathrm{n}-4}} \tag{13}$$

where n is the number of datasets included in the regression. In Fig. 15, ε calculated from Eq. (12) is shown along the depth for a few typical sites (blue circle markers) considered for validation. Figure 15-A (a, b, c… etc.) displays ε for the Correlation Model 1, i.e., without consideration of void ratio, while in the right-side Fig. 15-B (i, ii, iii…etc.) are the computed ε for Correlation Model 2. The continuous black line is the moving average with ± σ. At locations C15-M3, ECPT9-M1, and ECPT4-M4, nearly zero residuals are observed, which means that the developed correlation perfectly predicts the V_S value. Locations C3-M1 and C14-M13 show concentrated biases at depths greater than 20 m and 25 m for the latter case, which results in a slight underestimation of the V_S value at the mentioned level of depth. Collectively, observing all the residues for the formulated correlation is consistent with the considered validation sites.

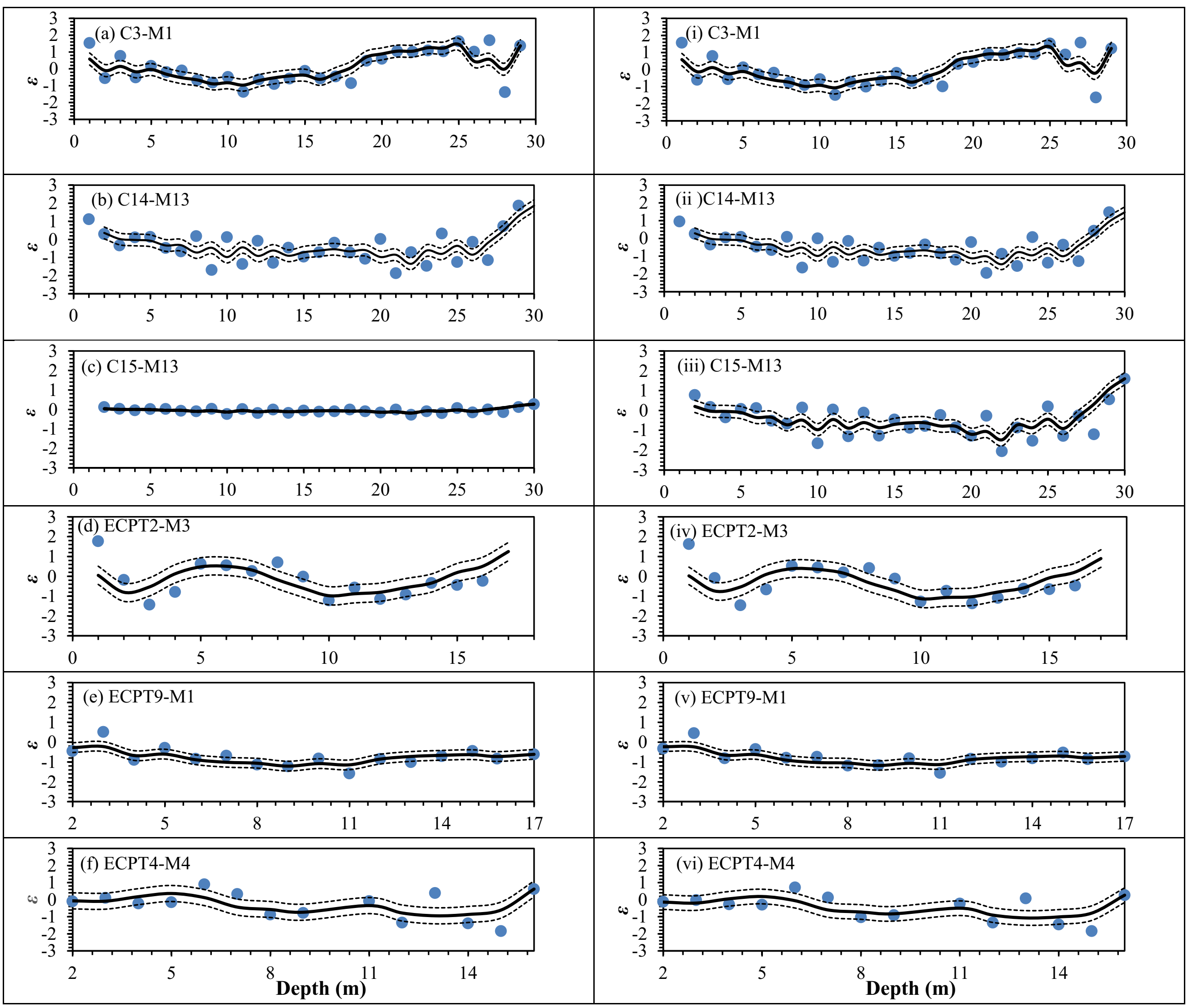


Fig. 15 Residue plot along depth for different locations (**a–f**) shows residue for Correlation Model 1; (i-vi) shows residue for Correlation Model 2

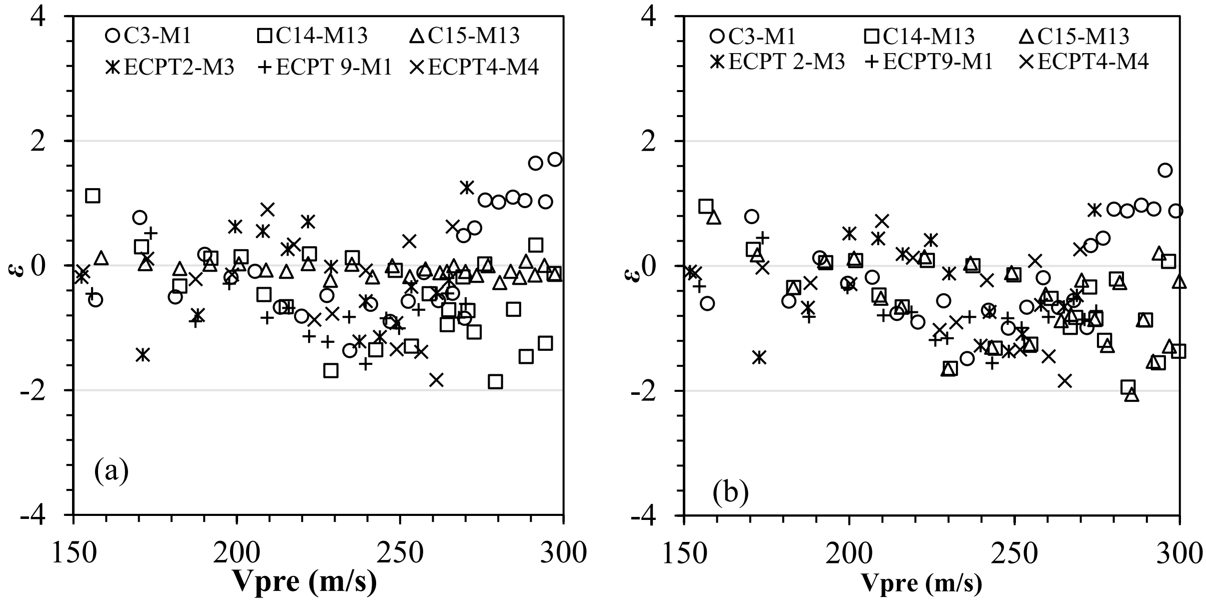


Fig. 16 Residue plot as a function of V_{pre} for (**a**) Correlation Model 1 (**b**) Correlation Model 2

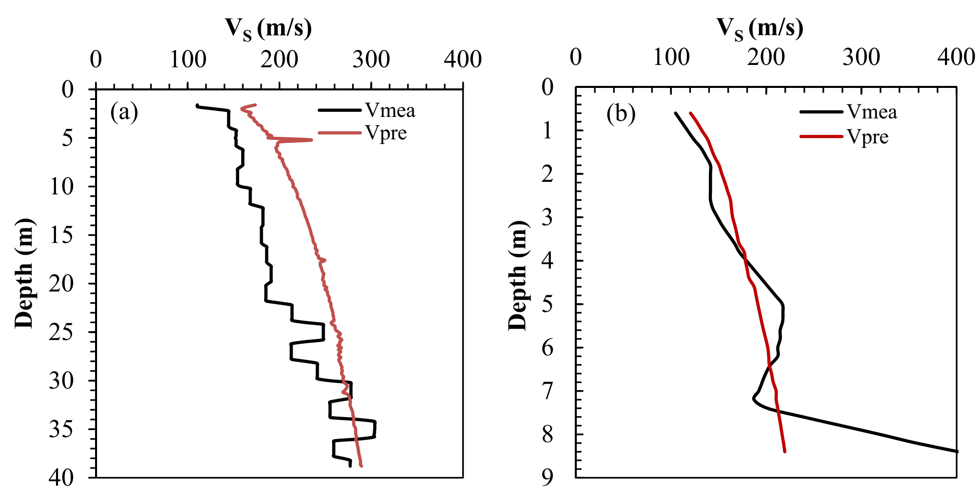


Fig. 17 Validation of CPT-MASW relation for a similar site outside the campus worldwide (**a**) Korean peninsula, Korea (**b**) Eskisehir, Turkey (V_{mea} and V_{pre} are measured and predicted value of V_S)

For correlation model 2 (Fig. 15-B), the same trend for ε is seen, i.e., residual estimates for the same selected areas are practically negligible. In Fig. 16, the X-axis represents the V_{pre} values obtained from Eq. 10 and 11. In contrast, the Y-axis represents the residuals given by Eq. 12. The figure shows that the ε value with a random scattering of residuals points around the zero line suggests that the regression model captures the actual measured value of V_S. Additionally, comparing correlation models 1 and 2 in Fig. 16b, no significant change has been observed in the residue plot.

Table 3 Proposed equations for void ratio (e) and shear wave velocity (V_s) estimation for soil in Quaternary alluvial

Equation No.	Equation	Remarks
Equation (6)	$e^{0.5}I_c^{\,n} = 5.2976 \times Q_{tn}^{-0.274}$	Mechanical cone
Equation (9)	$e^{0.5}I_c^{\,n} = 5.2976 \times Q_{tn}^{-0.274} \times (K_C)$	Electrical Cone
Equation (10)	$V_s=156.885q_c^{0.033}I_c^{0.120}\sigma'^{-0.169}z^{0.366}$	Correlation model 1
Equation (11)	$V_s=151.859q_c^{0.044}I_c^{0.165}\sigma'^{-0.202}z^{0.397}e^{0.035}$	Correlation model 2

Validation of *VS* at a similar site worldwide

Two external sites are selected from the literature, similar to the study areas. Here, explicitly mentioning site similarity means the site of the same geologic age and soil type formation. The literature shows *CPT* parameters along with the measured shear wave velocity. The V_{pre} value from the developed correlation is compared to the V_{mea} for validation. Figure 17a and b show the validation of V_S data for sites in the Korean peninsula region (Sun et al. 2013) and Eskisehir, Turkey, soil deposits (Mola-Abasi et al. 2015), respectively. The results show decent agreement between the predicted and measured results.

Conclusions

This research underway with the collection of *CPT* test data from two different types of cones and V_{mea} from MASW tests from the IIT Patna campus. Previous studies available in the literature describe that many correlations are available to estimate shear wave velocity from CPT data. However, shear wave velocity predicted (V_{pre}) and measured (V_{mea}) show a significant difference, necessitating the development of a new correlation for the studied site. While selecting the parameters for correlation, previous researchers indicated that consideration of void ratio would be beneficial for the developed model. Whereas, very limited correlations are available for predicting void ratio from the CPT data. Therefore, a prediction model has been proposed for estimating the void ratio based on the *CPT* results, and then *CPT*-V_S correlations are formulated. From this study, the following conclusions are drawn:

- One of the novelties of this study is the proposed Eq. (6) and given in Table 3 for predicting the void ratio from the CPT data. The data analysis of formulated model shows a power relation between the factored void ratio (*FVR*) ($e^{0.5}I_C{}^n$) and normalized cone tip resistance (Q_{tn}). Formulated *FVR*-Q_{tn} relation for *void ratio* computation when used with *ECPT* data, a downward shift has been observed, which necessitates the introduction of a cone correction factor (K_C) for the electrical cone. The value for *Kc* has been estimated and proposed as 0.809 for the electrical cone. A similar study can be performed using other cones to propose the value for *Kc* for other cones.
- The second novelty of this study is the models (shown in Table 3) for prediction of shear wave velocity. During the estimation of $V_{S,}$ cone parameters such as q_c, I_C, $\sigma\prime_{0,}$ and *z* are considered in the first correlation model (Eq. 10). Moreover, the void ratio effect in V_S prediction is checked by adding the *e* term in the model. The correlation model (discussed in Sect. "Void ratio (e) Prediction Model") for *e* can be used to estimate it from *CPT* data.
- No significant improvement was observed after the introduction of *e* in correlation model 2 (Eq. 11). This is because I_C is already included in both models. The I_C can integrate the effects of soil type and grain compactness. Therefore, it is recommended to use this I_C index if a V_S prediction model is to be formed using *CPT* parameters.
- The proposed equations are validated inside and outside the study area with similar soil conditions. The site conditions, geologic similarity, etc., must be considered before using this correlation. If the direct measurement of V_S is possible, then preference should be given to all those methods.

Acknowledgements The author(s) greatly acknowledge IIT Patna and the Department of Higher Education (Govt. of India) for funding the present research work to carry out the doctoral research study of the first author for which no specific Grant number has been allotted.

Data availability The data that supports the findings of this research are available from the corresponding author, upon a reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

References

AbbaszadehShahri A, Naderi S (2016) Modified correlations to predict the shear wave velocity using piezocone penetration test data and geotechnical parameters: a case study in the southwest of Sweden. Innov Infrastruct Solut 1:1–9. https://doi.org/10.1007/s41062-016-0014-y

Ahmad I, Waseem M, Abbas M, Ayub U (2015) Evaluation of shear wave velocity correlations and development of new correlation using Cross-hole data. Int J Georesources Environ - IJGE (formerly Int'l J Geohazards Environ) 1 1(1):42–51. https://doi.org/10.15273/ijge.2015.01.005

Al-Ajamee M, Baboo A, Kolathayar S (2022) Deterministic Seismic Hazard Analysis of Grand Ethiopia Renaissance Dam (GERD). In: Conference on Performance-based Design in Earthquake, Geotechnical Engineering, Geotechnical, Geological and Earthquake Engineering book series. Springer International Publishing, pp 1903–1913

Andrus RD, Mohanan NP, Piratheepan P et al (2007) Predicting shear-wave velocity from cone penetration resistance. In: 4th International conference on earthquake geotechnical engineering

Ang AH-S, Tang WH (2007) Probability concepts in engineering: emphasis on applications in civil and environmental engineering. Wiley, New York

Baldi G, Bellotti R, Ghionna VN et al (1990) Modulus of sands from CPT's and DMT's. In: Proc 12th Int Conf soil Mech Found Eng Rio Janeiro, 1989 Vol 1 28:165–170. https://doi.org/10.1016/0148-9062(91)93491-n

Cai G, Puppala AJ, Liu S (2014) Characterization on the correlation between shear wave velocity and piezocone tip resistance of Jiangsu clays. Eng Geol 171:96–103. https://doi.org/10.1016/j.enggeo.2013.12.012

Chakrabortty P, Kumar U, Puri V (2018) Seismic site classification and liquefaction hazard assessment of Jaipur city, India. Indian Geotech J 48:768–779. https://doi.org/10.1007/s40098-017-0287-x

Das A, Chakrabortty P (2022) Large strain dynamic characteristics of quaternary alluvium sand with emphasis on empirical pore water pressure generation model. Eur J Environ Civ Eng 26:5729–5752. https://doi.org/10.1080/19648189.2021.1916605

Douglas BJ, Olsen RS (1981) Soil classification using electric cone penetrometer. In: Proceedings of conference on cone penetration testing and experience, St. Louis. pp 209–227

Fayed AL, Mousa AA (2020) Shear wave velocity in the East Nile delta clay: correlations with static CPT measurements. Geotech Geol Eng 38:2303–2315. https://doi.org/10.1007/s10706-019-01089-4

Gadeikis S, Dundulis K, Žaržojus G et al (2013) Correlation between shear wave velocity and cone resistance of Quaternary glacial clayey soils defined by seismic cone penetration test (SCPT), Lithuania. J Vibroengineering 15(2):992–998

Gu X, Yang J, Huang M, Gao G (2015) Bender element tests in dry and saturated sand: Signal interpretation and result comparison. Soils Found 55:951–962. https://doi.org/10.1016/j.sandf.2015.09.002

Hegazy YA, Mayne PW (1995) Statistical correlations between Vs and cone penetration data for different soil types. In: Proceedings of the International symposium on cone penetration testing, CPT, pp 173–178

Hegazy YA, Mayne PW (2006) A global statistical correlation between shear wave velocity and cone penetration data. In: Proceeding of GeoShanghai, American Society of Civil Engineers (ASCE) 243–248. https://doi.org/10.1061/40861(193)31

Jaime A, Romo MP (1988) The Mexico earthquake of september 19, 1985—Correlations between dynamic and static properties of mexico city clay. Earthq Spectra 4:787–804. https://doi.org/10.1193/1.1585502

Jardine RJ, Potts DM, Burland JB, Fourie AB (1986) Studies of the influence of non-linear stress–strain characteristics in soil–structure interaction. Geotechnique 36:377–396. https://doi.org/10.1680/geot.1986.36.3.377

Jefferies M, Davies M (1993) Use of CPTu to estimate equivalent SPT. Geotech Test J 16:458–468. https://doi.org/10.1520/gtj10286j

Khan Z, Yamin M, Attom M, Al Hai N (2022) Correlations between SPT, CPT, and Vs for reclaimed lands near Dubai. Geotech Geol Eng 40:4109–4120. https://doi.org/10.1007/s10706-022-02143-4

Long M, Donohue S (2010) Characterization of Norwegian marine clays with combined shear wave velocity and piezocone cone penetration test (CPTU) data. Can Geotech J 47:709–718. https://doi.org/10.1139/t09-133

Mayne PW, Rix GJ (1995) Correlations between shear wave velocity and cone tip resistance in natural clays. Soils Found 35:107–110. https://doi.org/10.3208/sandf1972.35.2_107

Mayne PW (2007) NCHRP Synthesis 368: Cone penetration testing. Transportation Research Board, Washington, DC, 118

McGann CR, Bradley BA, Taylor ML et al (2015) Development of an empirical correlation for predicting shear wave velocity of Christchurch soils from cone penetration test data. Soil Dyn Earthq Eng 75:66–75. https://doi.org/10.1016/j.soildyn.2015.03.023

Mishra P, Paul A, Chakrabortty P (2023) Correlation between cone tip resistance and shear wave velocity for quaternary alluvium. Lect Notes Civ Eng 332 LNCE:75–86. https://doi.org/10.1007/978-981-99-1459-3_7

Mohamed Ahmed S, Ahmed SM (2017) Correlating the shear wave velocity with the cone penetration test. In: World Congress on Civil, Structural, and Environmental Engineering. pp 2371–5294

Mola-Abasi H, Dikmen U, Shooshpasha I (2015) Prediction of shear-wave velocity from CPT data at Eskisehir (Turkey), using a polynomial model. Near Surf Geophy 13(2):155–168

Mousa A, Hussein M (2022) Prediction of shear wave velocity in fine-grained soils from cone penetration test data: Toward a global approach. 2676:565–582. https://doi.org/10.1177/03611981221075627

Nilay N, Chakrabortty P, Popescu R (2022) Liquefaction hazard mapping using various types of field test data. Indian Geotech J 52:280–300. https://doi.org/10.1007/s40098-021-00570-3

Piratheepan P (2002) Estimating shear-wave velocity from SPT and CPT data. Thesis, Clemson University

Robertson PK (1990) Soil classification using the cone penetration test. Can Geotech J 27(1):151–158. https://doi.org/10.1139/t90-014

Robertson PK (2009) Interpretation of cone penetration tests - A unified approach. Can Geotech J 46:1337–1355. https://doi.org/10.1139/t09-065

Robertson PK, Campanella RG (1983) Interpretation of cone penetration tests. Part I: Sand. Can Geotech J 20(4):718–733. https://doi.org/10.1139/t83-078

Robertson PK, Wride CE (1998) Evaluating cyclic liquefaction potential using the cone penetration test. Can Geotech J 35(3):442–459. https://doi.org/10.1139/t98-017

Robertson PK, Campanella RG, Gillespie D, Rice A (1986) Seismic cpt to measure in situ shear wave velocity. J Geotech Eng 112:791–803. https://doi.org/10.1061/(ASCE)0733-9410(1986)112:8(791)

Sahu S, Raju NJ, Saha D (2010) Active tectonics and geomorphology in the Sone-Ganga alluvial tract in mid-Ganga Basin, India. Quat Int 227:116–126. https://doi.org/10.1016/j.quaint.2010.05.023

Sahu S, Saha D, Dayal S (2015) Sone megafan: A non-Himalayan megafan of craton origin on the southern margin of the middle Ganga Basin, India. Geomorphology 250:349–369. https://doi.org/10.1016/j.geomorph.2015.09.017

Sara A (2014) Prediction of the shear wave velocity VS from CPT and DMT at research sites. Front Struct Civ Eng 8:83–92. https://doi.org/10.1007/s11709-013-0234-6

Sun CG, Cho CS, Son M, Shin JS (2013) Correlations between shear wave velocity and in-situ penetration test results for Korean soil deposits. Pure Appl Geophys 170:271–281. https://doi.org/10.1007/s00024-012-0516-2

Sykora DE, Stokoe KH (1983) Correlations of in-situ measurements in sands of shear wave velocity. Soil Dyn Earthq Eng 2:125–136

Tun M, Ayday C (2018) Investigation of correlations between shear wave velocities and CPT data: a case study at Eskisehir in Turkey. Bull Eng Geol Environ 77:225–236. https://doi.org/10.1007/s10064-016-0987-y

Wang C, Chen Q, Shen M, Juang CH (2017) On the spatial variability of CPT-based geotechnical parameters for regional liquefaction evaluation. Soil Dyn Earthq Eng 95:153–166. https://doi.org/10.1016/j.soildyn.2017.02.001

Yang Z, Cui Y, Guo L et al (2022) Semi-empirical correlation of shear wave velocity prediction in the Yellow river delta based on CPT. Mar Georesources Geotechnol 40:487–499. https://doi.org/10.1080/1064119x.2021.1913458

Zhang M, Tong L (2017) New statistical and graphical assessment of CPT-based empirical correlations for the shear wave velocity of soils. Eng Geol 226:184–191. https://doi.org/10.1016/j.enggeo.2017.06.007

Zhang G, Robertson PK, Brachman RWI (2002) Estimating liquefaction-induced ground settlements from CPT for level ground. Can Geotech J 39(5):1168–1180. https://doi.org/10.1139/t02-047