3rd International Conference on Civil Engineering,
Architecture and Urban Management in Iran
August -2020

# A polynomial model for predicting shear wave velocity from cone penetration test data

**Sina L. Afkham*[1], Abolfazl Najafi[2]**

1. Graduate Student, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran; email: s_lahijani@civileng.iust.ac.ir
2. Graduate Student, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran; email: abolfazl_najafi@civileng.iust.ac.ir

## Abstract

Shear wave velocity (Vs) is known to be an essential soil parameter for site characterization purposes. Although shear wave velocity can be directly measured, indirect determination using common in-situ tests, entailing cone penetration test (CPT), are often preferred by engineers. In this paper, the relation between Vs and CPT parameters is reinvestigated and a new polynomial regression form is proposed to correlate shear wave velocity with cone tip resistance, sleeve friction, and vertical effective stress. Using a dataset obtained from several sites in Canada, Japan, Norway, Italy, and China, the performance of the presented correlation is appraised through comparison with the observed data as well as formerly published CPT-Vs equations. The results demonstrate that the suggested functional form provides a more reliable approach for predicting shear wave velocity and thus, can ostensibly be regarded as an improvement to the currently adopted practices.

**Keywords:** shear wave velocity, cone penetration test, CPT-Vs model, small strain shear modulus, empirical correlation, ranking index

## 1. Introduction

Deemed to be an important soil parameter in both earthquake and geotechnical engineering [1], shear wave velocity ($V_s$) is widely utilized to evaluate liquefaction potential, define site categories, implement remedial measures, and estimate foundation settlements [2-5].

$V_s$ is ascertained either by direct approaches, including laboratorial and in-situ tests, or empirical correlations. Although direct measurement is preferred due to its relatively high precision, numerous associated drawbacks restrict the usage of this mechanism. When it comes to laboratorial testing, for instance, the major downside is the requisite of undisturbed sampling, that is neither economical nor feasible [6]. In addition, $G_0$ is susceptible to disturbances which are ineluctable in laboratorial tests [7]. Turning to in-situ testing, this approach can be subsumed under two classifications: borehole-type measurement and surface wave methods. The former, that involves CHT and DHT, not only has high financial and temporal expenses connected with drilling [7], but it also illustrates subsurface conditions at a single level, rather than the average situation along a line. The latter, which entails SASW, although does not need a borehole, requires the solution of an inverse problem, that is predominantly ill-posed. Besides, usage of this method might contribute to missing anomalies as well as inclusions [8, 9].

The aforementioned difficulties associated with the direct measurement of $V_s$ along with the remarkable expenses connected with the provision of specialized equipment culminate in their use only for important projects. Conversely, site investigation techniques, encompassing SPT and CPT, are commonly deployed in a broader scope of sites and thus, empirical equations based on the data obtained by such tests can be utilized to determine shear wave velocity for the purpose of ground response assessment.

The objective of this paper is to propose and validate a polynomial model to predict $V_s$ based on $\sigma_{vo}$ as well as CPT parameters $q_t$, and $f_s$. Firstly, a number of formerly developed correlations will be reviewed. Secondly, the database used in this study will be presented. Finally, a new polynomial model will be suggested and subsequently compared with previously proposed equations.

## 2. Review of the selected correlations between CPT parameters and Vs

Numerous Researchers have investigated relationships between shear wave velocity and CPT parameters, involving cone tip resistance ($q_c$), cone frictional resistance ($f_s$), soil behaviour type index ($I_c$), depth (z), void ratio ($e_0$), and vertical effective stress ($\sigma'_{vo}$). Contingent upon the type of soil, suggested CPT-$V_s$ correlations can be divided into three classifications: (a) equations for cohesive soils; (b) equations for cohesionless soils; and (c) general equations for all soil types. The following paragraphs will summarize a number of previously-published functional forms, which are apposite to this study.

### 2.1. Sykora and Stokoe (1983)

Based on 256 data pairs obtained from 9 sites, Sykora and Stokoe [10] developed an equation between $q_c$ and $V_s$ with the following form:

$$V_s = a{q_c}^b \tag{1}$$

Where a, and b are regression coefficients and $q_c$ is raw tip resistance. The dataset used in this study included $q_c$ and $V_s$ ranging from 1 to 70 MPa and 120 to 500 m/s respectively.

### 2.2. Baldi et al. (1989)

Baldi et al. [11] suggested a CPT-$V_s$ correlation for unaged, uncemented silica sands based on cone penetration tests, laboratory resonant column tests, and dilatometer tests in the calibration chamber. The equation developed by Baldi et al. has the form below:

$$V_s = a{q_c}^b {\sigma'_{vo}}^c \tag{2}$$

In which c is a regression coefficient, and $\sigma'_{vo}$ is vertical effective stress.

### 2.3. Mayne and Rix (1995)

Mayne and Rix [12] proposed a CPT-$V_s$ relation for both intact and fissured clays based on 364 data pairs obtained from 31 sites. The form of the function is stated below:

$$V_s = a{q_c}^b e_0^c \tag{3}$$

Where $e_0$ is the void ratio. The dataset entailed the plasticity index with a range of $8<PI<300$ and the over-consolidation ratio with a range of $1<OCR<100$. In-situ measurement of $V_s$ was made by CHT, DHT, and SASW.

### 2.4. Hegazy and Mayne (1995)

Based on a database of 24 sand sites and by deploying multiple regression, Hegazy and Mayne [13] established the relationship between $V_s$, and CPT parameters $q_t$, and $f_s$. This equation has the following form:

$$V_s = a{q_t}^b f_s^c \tag{4}$$

In which $f_s$ is cone frictional resistance, and $q_t$ is corrected piezocone resistance.

### 2.5. Wair et al. (2012)

Wair el al. [14] advocated the functional form below for CPT-$V_s$ regression analysis:

$$V_s = a{q_t}^b f_s^c {\sigma'}_{vo}^d \tag{5}$$

Where d is a regression coefficient.

### 2.6. Robertson (2009)

Robertson [15] developed a CPT-$V_s$ correlation based on a global set of 1035 datapoints predominately acquired from uncemented soils of Holocene and Pleistocene-age. The database used in this study, moreover, involved normalized cone tip resistance ($Q_{tn}$), friction ratio ($F_r$), and vertical effective stress ($\sigma'_{vo}$) with mean values of 56, 3.13%, and 190 kPa, as well as ranges of 0.67 to 577, 0.15% to 13.13%, and 19kPa to 580 kPa respectively. The equation suggested by Robertson is of the following form:

$$V_s = \left[10^{a+bI_c}\left(\frac{q_t - \sigma_{vo}}{Pa}\right)^{0.5}\right] \tag{6}$$

In which $I_c$ is soil behaviour type index, $\sigma_{vo}$ is vertical total stress, and Pa is atmospheric pressure.

### 2.7. Hegazy and Mayne (2006)

By utilizing a total of 558 data pairs, Hegazy and Mayne [16] proposed a CPT-$V_{s1}$ equation, that has the form below when rearranged to solve for $V_s$:

$$V_s = aQ_{tn}e^{bI_c}\left(\frac{\sigma'_{vo}}{Pa}\right)^{0.25} \tag{7}$$

Where all terms are formerly defined. The CPT data compiled from 73 sites, encompassing 31 cohesive and 42 cohesionless sites. Shear wave velocity was measured by various in-situ techniques including CHT, DHT, SCPT, and SASW.

### 2.8. Mola-Abasi et al. (2015)

Mola-Abasi et al. [17] established a quadratic polynomial relation between $V_s$ and CPT parameters $q_c$, and $f_s$ based on 437 data points obtained from 37 sites in Turkey. The study area was primarily composed of sand and silt with the exiguous existence of gravel and clay deposits. The suggested function has the following form:

$$V_s = a + bq_c + cf_s + d{q_c}^2 + ef_s^2 + fq_cf_s \tag{8}$$

In which e, and f are regression coefficients.

## 3. Compiled database

The database comprises 15 undisturbed samples of sands, which were obtained from 6 sites in Canada, 4 sites in Japan, and one site each in Italy, Norway, and China. Moreover, the samples from Canada, Japan, and Italy were acquired by in-situ ground freezing, while the ones from Norway and China were taken with tube and Mazier core barrel respectively. It is also worth noting that hydraulic fills, mine tailings, and natural deposits (mainly alluvial) were included in soil samples. This field database was presented by Mayne [18] and subsequently deployed by numerous other researchers. Table 1 demonstrates the CPT data as well as relevant geotechnical properties of the samples.

| Sample No. | $\gamma$ ($kN/m^3$) | $\sigma_{vo}$ (kPa) | $\sigma'_{vo}$ (kPa) | $e_0$ | $q_t$ (MPa) | $f_s$ (kPa) | $F_r$ (%) | $I_c$ | $V_s$ (m/s) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 20.53 | 270 | 180 | 0.492 | 10.2 | 183 | 1.84 | 2.237 | 175.2 |
| 2 | 18.65 | 123 | 102 | 0.820 | 19.9 | 188 | 0.95 | 1.681 | 183.8 |
| 3 | 19.11 | 144 | 123 | 0.720 | 12.8 | 130 | 1.03 | 1.885 | 204.6 |
| 4 | 18.74 | 164 | 143 | 0.790 | 13.9 | 122 | 0.89 | 1.847 | 198.7 |
| 5 | 18.53 | 108 | 87 | 0.857 | 19.7 | 59 | 0.30 | 1.347 | 205.2 |
| 6 | 18.27 | 98 | 84 | 0.947 | 13.1 | 31 | 0.24 | 1.453 | 180.7 |
| 7 | 17.88 | 72 | 51 | 1.043 | 8.0 | 58 | 0.73 | 1.776 | 137.9 |
| 8 | 19.02 | 726 | 516 | 0.768 | 17.2 | 121 | 0.73 | 2.213 | 227.0 |
| 9 | 18.18 | 135.4 | 120 | 0.970 | 6.1 | 24 | 0.40 | 1.932 | 172.0 |
| 10 | 18.33 | 175 | 160 | 0.981 | 8.6 | 31 | 0.37 | 1.852 | 195.7 |
| 11 | 18.83 | 60 | 55 | 0.762 | 1.8 | 15 | 0.87 | 2.345 | 101.5 |
| 12 | 18.62 | 121 | 100 | 0.849 | 4.0 | 16 | 0.41 | 2.059 | 146.6 |
| 13 | 18.73 | 178 | 138 | 0.825 | 5.0 | 18 | 0.38 | 2.044 | 145.2 |
| 14 | 19.54 | 120 | 110 | 0.724 | 3.4 | 14 | 0.42 | 2.162 | 135.4 |
| 15 | 20.30 | 57 | 42 | 0.589 | 11.8 | 30 | 0.26 | 1.400 | 167.0 |

**Table 1: Database presented by Mayne [18]**

## 4. Development of a new polynomial model

In this study, the preliminary premise, utilized to establish the relationship between CPT parameters and $V_s$, is that inputs can be related to the output through a polynomial equation. Based on this assumption, the subsequent task would be to obtain a polynomial function, for which the sum of squared errors between measured and calculated outputs is minimum. As an attempt to address this conundrum, The connection between input and output variables can be represented by a disparate analogue of Volterra functional series, known as Kolmogorov-Gabor polynomial [19]:

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \cdots \quad (9)$$

Where A= ($a_0$, $a_1$,…, $a_m$), and X= ($x_0$, $x_1$,…, $x_m$) are the vectors of weights, and input variables respectively. Should only the first three terms be used, equation (9) can be written in the following form for three input variables:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 {x_1}^2 + a_5 {x_2}^2 + a_6 {x_3}^2 + a_7 x_1 x_2 + a_8 x_1 x_3 + a_9 x_2 x_3 \quad (10)$$

In the current study, input variables involve corrected piezocone resistance, cone frictional resistance, and vertical effective stress. Thus, by means of equation (10), shear wave velocity can be predicted as follows:

$$V_s = a_0 + a_1 q_t + a_2 f_s + a_3 \sigma'_{vo} + a_4 {q_t}^2 + a_5 {f_s}^2 + a_6 {\sigma'_{vo}}^2 + a_7 q_t f_s + a_8 q_t \sigma'_{vo} + a_9 f_s \sigma'_{vo} \quad (11)$$

Deploying multiple regression analysis, the coefficients $a_i$ are calculated, so that the sum of squared subtractions between observed and predicted $V_s$ is minimum. Finally, $V_s$ can be expressed in the following proposed correlation:

$$V_s = 59.34 + 9.74 q_t - 20.21 f_s + 513.84 \sigma'_{vo} - 0.27 {q_t}^2 - 1871.45 {f_s}^2 - 1049.20 {\sigma'_{vo}}^2 + 10.79 q_t f_s + 7.03 q_t \sigma'_{vo} + 623.23 f_s \sigma'_{vo} \quad (12)$$

In which $q_t$, $f_s$, and $\sigma'_{vo}$ are all in units of MPa.

## 5. Comparison between the proposed model and previously developed correlations

Fig. 1 depicts a comparison between the measured $V_s$ values and the predicted $V_s$ values using each of the formerly presented functional forms. The plots are associated with the coefficient of determination as well as root mean square error. These two statistical indices reveal how well the database can be replicated by each of the regression forms and can be calculated as follows:

$$R^2 = 1 - \left[ \frac{\sum_{i=1}^{n} \left( Y_{i(Actual)} - Y_{i(model)} \right)^2}{\sum_{i=1}^{n} \left( Y_{i(Actual)} \right)^2} \right] \quad (13)$$

$$RMSE = \left[ \frac{\sum_{i=1}^{n} \left( Y_{i(Actual)} - Y_{i(model)} \right)^2}{n} \right]^{0.5} \quad (14)$$

With regard to the scatter plots in Fig. 1, it appears that Eq. (7), proposed by Hegazy and Mayne (2006), is least applicable to the current dataset, followed by Eq. (6), developed by Robertson (2006), with $R^2$ and RMSE values of 0.47 and 23.34 respectively. Conversely, the proposed regression form ostensibly provides the best performance, followed by Eqs. (2) and (5) which give analogous representations of the observed data.

In order to evaluate and juxtapose the accuracy of the aforementioned correlations, the mean value of the ratio of the estimated shear wave velocity to the measured shear wave velocity can be deployed [20-22]. This ratio is denoted by K and determined using the formula below:

$$K = \frac{V_{sc}}{V_{sm}} \quad (15)$$

Where $V_{sc}$, and $V_{sm}$ are estimated and measured shear wave velocity respectively.

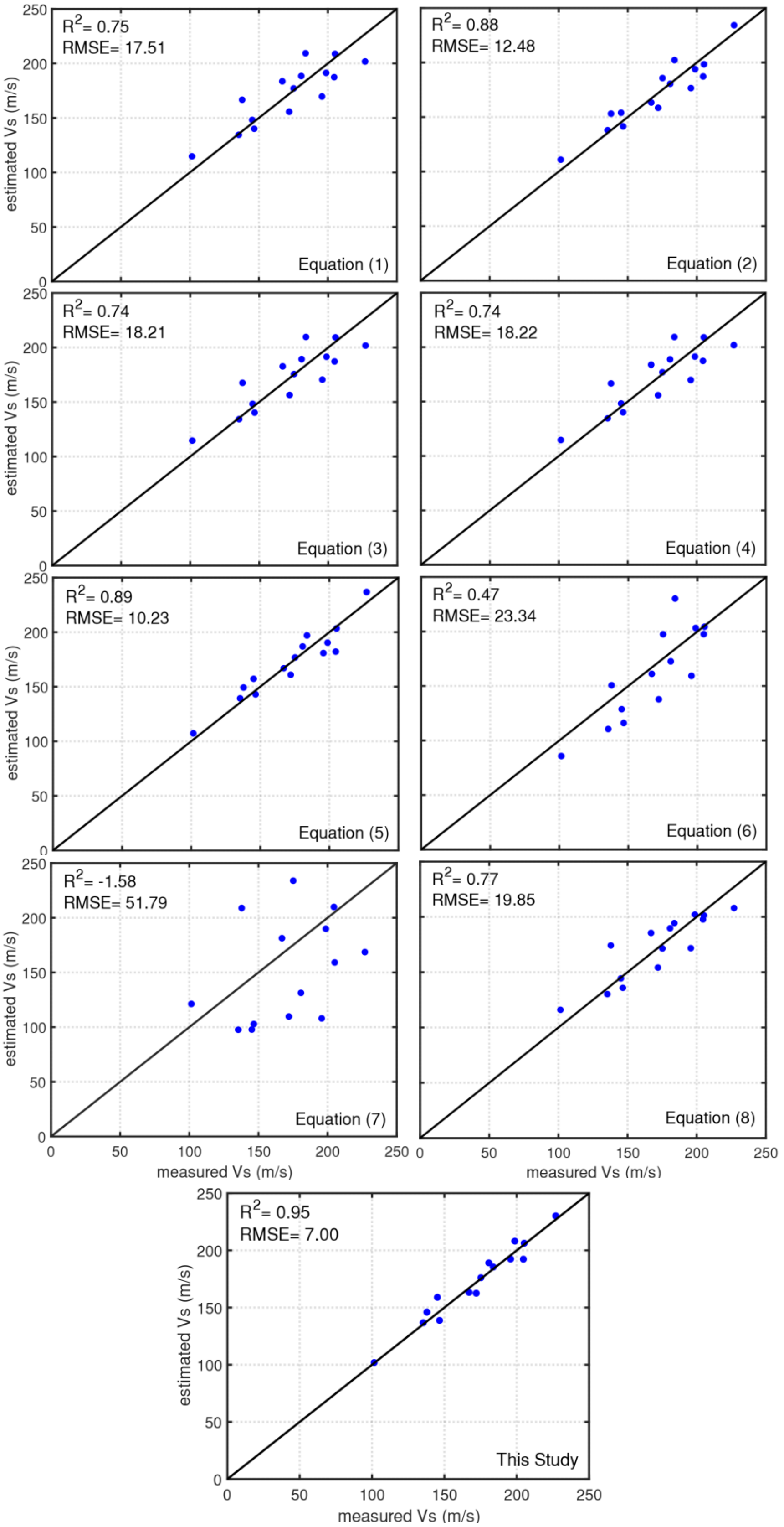


**Fig.1. Comparison of measured and estimated $V_s$ for selected functional forms.**

Table 2 demonstrates the mean value of K ($\mu_k$) for each of the functional forms. Based on these results, it is evident that Eq. (5), developed by Wair et al. (2012), Eq. (2), suggested by Baldi et al. (1989), and the correlation proposed in this study provide the highest accuracy (closeness of the mean of K to 1). Nevertheless, Eq. (7), developed by Hegazy and Mayne (2006), apparently gives the least accurate prediction for the current dataset at 0.937, followed by Eq. (6), suggested by Robertson (2009), with a mean of K value of 0.963. It is also worth noting that all the equations generally overestimate the measured data, bar the Eqs. (6) and (7), for which $\mu_k$ is below 1.

In addition to accuracy, the precision of empirical equations is of consequence and needs to be appraised. Deemed to refer to scatter around the mean, precision can be quantified by the standard deviation of K ($\sigma_k$) [20-22]. Based on the results presented in Table 2, it appears that the CPT-$V_s$ correlation proposed in this study and Eq. (5), suggested by Wair et al. (2012), respectively provide the highest and next highest precision. With a $\sigma_k$ of 0.309, Eq. (7), developed by Hegazy and Mayne (2006), conversely gives the least precise estimation for the current database.

Derived from the fact that K can theoretically range from zero to infinity, the distribution of K around the mean is nonsymmetric and hence, equal weights are not given to underestimation and overestimation. To mitigate this problem, Briaud and Tucker [20] proposed ranking index (RI), by which the overall performance of a correlation can be expressed. RI is calculated as follows:

$$RI = \left|\mu_{\ln(K)}\right| + \sigma_{\ln(K)} \tag{16}$$

Table 2 illustrates RI values for each of the regression forms. Based on these results, it is evident that the equation suggested in the current study gives the best prediction of the observed data with RI of 0.049. It should also be considered that Eqs. (1), (3), (4), and (8), all standing at 0.104, represent the dataset alike. RI values of discrete functional forms are plotted in Fig. 2 as well. Regarding this line graph, The RI curve pursues the profile of the RMSE curve and is graphed above it [23, 24].

In order to further investigate the capability of the proposed approach, scaled relative error ($E_r$) is deployed as a means of illustrating the discrepancy between the observed and predicted values compared to the size of measurements:

$$E_r = \frac{V_{sc} - V_{sm}}{V_{sm}} \tag{17}$$

Where both terms are formerly defined. Fig. 3 shows the percentage of cumulative frequency versus the percentage of scaled relative error for each of the regression forms presented in this study. Based on this graph, it appears that using Eqs. (7), and (6), about 20%, and 40% of $V_s$ values are predicted within a 10% error margin respectively. Conversely, using the correlation proposed in the current study, 100% of the $V_s$ values are predicted within a 10% error, followed by Eq. (5) at about 90%.

In the final analysis, the appraisement of a number of previously developed equations revealed that they are predominantly inapplicable to the database used in this study and that the proposed functional form provides a comparatively superior prediction of the shear wave velocity. These differences might be underlain by numerous factors encompassing geotechnical

| Reference | $\mu_k$ | $\sigma_k$ | RI |
|---|---|---|---|
| Equation (1) | 1.010 | 0.100 | 0.104 |
| Equation (2) | 1.006 | 0.068 | 0.072 |
| Equation (3) | 1.011 | 0.100 | 0.104 |
| Equation (4) | 1.011 | 0.100 | 0.104 |
| Equation (5) | 1.005 | 0.060 | 0.064 |
| Equation (6) | 0.963 | 0.139 | 0.190 |
| Equation (7) | 0.937 | 0.309 | 0.432 |
| Equation (8) | 1.009 | 0.103 | 0.104 |
| This Study | 1.006 | 0.044 | 0.049 |

**Table 2: Statistical information for presented regression forms.**

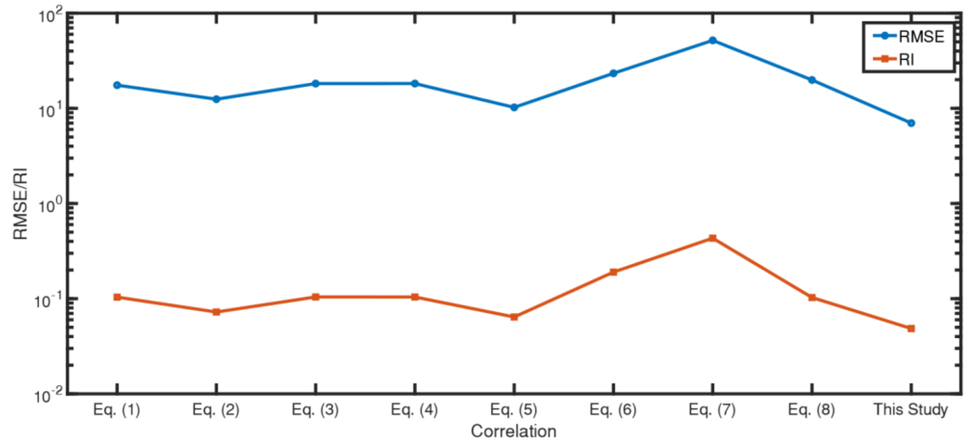


**Fig. 2: Results of RMSE, and RI**

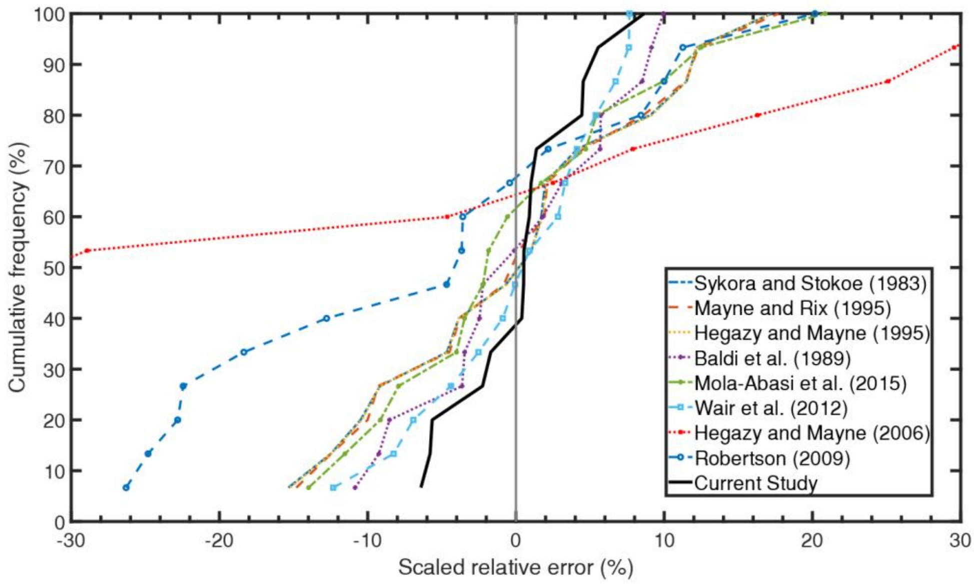


**Fig. 3: Scaled relative errors of predicted $V_s$**

conditions of the study region, water table fluctuations, and geological age remarkably affecting the correlations.

## 6. Conclusion

In the current study, a new polynomial regression form was proposed for the prediction of $V_s$ based on CPT parameters $q_t$, and $f_s$ as well as $\sigma '_{vo}$. A dataset, consisting of 15 undisturbed samples of sand, was used to assess the performance of the suggested polynomial correlation in terms of $R^2$, RMSE, the ratio of the estimated to the measured shear wave velocity (K), ranking index (RI), and scaled relative error between the observed and predicted $V_s$. Based on these results, it appeared that the proposed equation provides the best prediction of $V_s$ thanks to the preponderance of both its accuracy and precision. The findings of this study also corroborate the conclusion reached by an array of antecedent researches that an empirical CPT-$V_s$ correlation should merely be used in a site-specific sense.

## References

[1] Ohta, Y. and N. Goto, Empirical shear wave velocity equations in terms of characteristic soil indexes. Earthquake engineering & structural dynamics, 1978. 6(2): p. 167-187.

[2] Borcherdt, R.D., Estimates of site-dependent response spectra for design (methodology and justification). Earthquake spectra, 1994. 10(4): p. 617-653.

[3] Dobry, R., et al., New site coefficients and site classification system used in recent building seismic code provisions. Earthquake spectra, 2000. 16(1): p. 41-67.

[4] Andrus, R.D. and K.H. Stokoe II, Liquefaction resistance of soils from shear-wave velocity. Journal of geotechnical and geoenvironmental engineering, 2000. 126(11): p. 1015-1025.

[5] Kayen, R., et al., Shear-wave velocity–based probabilistic and deterministic assessment of seismic soil liquefaction potential. Journal of Geotechnical and Geoenvironmental Engineering, 2013. 139(3): p. 407-419.

[6] Ahmed, S.M. Correlating the shear wave velocity with the cone penetration test. in Conference: 2nd international conference on geotechnical research and engineering (ICGRE’17), Barcelona, Spain. 2017.

[7] Stokoe, K.H. and J.C. Santamarina. Seismic-wave-based testing in geotechnical engineering. in ISRM International Symposium. 2000. International Society for Rock Mechanics and Rock Engineering.

[8] McGann, C.R., et al., Development of an empirical correlation for predicting shear wave velocity of Christchurch soils from cone penetration test data. Soil Dynamics and Earthquake Engineering, 2015. 75: p. 66-75.

[9] Tarantola, A., Inverse problem theory and methods for model parameter estimation. Vol. 89. 2005: siam.

[10] Sykora, D.W. and K.H. Stokoe, Correlations of in situ measurements in sands of shear wave velocity. Soil Dynamics and Earthquake Engineering, 1983. 20: p. 125-136.

[11] Baldi, G., et al. Modulus of sands from CPT's and DMT's. in Congrès international de mécanique des sols et des travaux de fondations. 12. 1989.

[12] Mayne, P.W. and G.J. Rix, Correlations between shear wave velocity and cone tip resistance in natural clays. Soils and foundations, 1995. 35(2): p. 107-110.

[13] Hegazy, Y. and P. Mayne. Statistical correlations between Vs and cone penetration data for different soil types. in Proceedings of the international symposium on cone penetration testing, CPT. 1995.

[14] Wair, B.R., J.T. DeJong, and T. Shantz, Pacific Earthquake Engineering Research Center. Disponible en Interner: http://peer. berkeley. edu/publications/peer_reports/reports_2012/webPEER-2012-08-DeJong. pdf, consultado el, 2012. 1.

[15] Robertson, P., Interpretation of cone penetration tests—a unified approach. Canadian geotechnical journal, 2009. 46(11): p. 1337-1355.
[16] Hegazy, Y.A. and P.W. Mayne, A global statistical correlation between shear wave velocity and cone penetration data, in Site and Geomaterial Characterization. 2006. p. 243-248.
[17] Mola-Abasi, H., U. Dikmen, and I. Shooshpasha, Prediction of shear-wave velocity from CPT data at Eskisehir (Turkey), using a polynomial model. Near Surface Geophysics, 2015. 13(2): p. 155-168.
[18] Mayne, P., The Second James K. Mitchell Lecture Undisturbed sand strength from seismic cone tests. Geomechanics and Geoengineering: An International Journal, 2006. 1(4): p. 239-257.
[19] Ivakhnenko, A.G., Polynomial theory of complex systems. IEEE transactions on Systems, Man, and Cybernetics, 1971(4): p. 364-378.
[20] Briaud, J.-L. and L.M. Tucker, Measured and predicted axial response of 98 piles. Journal of Geotechnical Engineering, 1988. 114(9): p. 984-1001.
[21] Giasi, C., C. Cherubini, and F. Paccapelo, Evaluation of compression index of remoulded clays by means of Atterberg limits. Bulletin of Engineering Geology and the Environment, 2003. 62(4): p. 333-340.
[22] Cherubini, C. and T. Orr, A rational procedure for comparing measured and calculated values in geotechnics. Coastal Geotechnical Engineering in Practice, Yokohama, 2000. 1: p. 261-265.
[23] Zhang, M. and L. Tong, New statistical and graphical assessment of CPT-based empirical correlations for the shear wave velocity of soils. Engineering Geology, 2017. 226: p. 184-191.
[24] Onyejekwe, S., X. Kang, and L. Ge, Assessment of empirical equations for the compression index of fine-grained soils in Missouri. Bulletin of Engineering Geology and the Environment, 2015. 74(3): p. 705-716.