https://doi.org/10.3311/PPci.41202

Periodica Polytechnica Civil Engineering, 69(4), pp. 1311–1320, 2025

Modified Correlations to Predict Shear-wave Velocity Using Cone Penetration Test Data for Hungary

Ahmad Al-Azazmeh[1]*, András Mahler[1]

[1] Department of Engineering Geology and Geotechnics, Faculty of Civil Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
* Corresponding author, e-mail: azazmeh9999@edu.bme.hu

Received: 05 June 2025, Accepted: 03 October 2025, Published online: 29 October 2025

Abstract

This study evaluates and optimizes empirical correlations between shear wave velocity (V_s) and cone penetration test (CPT) data for various soil types in Hungary. A comprehensive database of 914 data pairs was compiled from multiple cities with diverse geological conditions, incorporating SCPTu and MASW measurements. Over 40 existing V_s–CPT correlations were statistically assessed using parameters such as RMSE, RD, *K*, CVK, and RI to determine their accuracy across different soil types and depositional settings. The most promising correlations were further refined using regression analysis, leading to the development of improved models tailored for Hungarian soils. These new correlations were evaluated both graphically and statistically, showing enhanced predictive performance, particularly for coarse-grained soils. The final proposed models demonstrate significant reductions in estimation error, with RMSE improvements exceeding 35%. This work provides geotechnical engineers in Hungary with robust, site-adapted tools for seismic site characterization and supports safer and more reliable subsurface profiling practices.

Keywords

in situ testing, shear wave velocity, cone penetration test, empirical correlation

1 Introduction

Seismic site characterization is an essential aspect of geotechnical investigations, especially for assessing the behavior of soils during earthquakes and other dynamic loading conditions. Shear wave velocity (V_s) plays a crucial role in this process as it provides valuable information about the stiffness and strength properties of soils.

V_s is a measure of the speed at which shear waves propagate through a soil medium. It is typically expressed in meters per second (m/s) or kilometers per second (km/s). The V_s is directly related to the shear modulus (G_{max}) of the soil, which characterizes its resistance to shear deformation at very small strain levels (Eq. (1)) [1]:

$$G_{max} = \rho \cdot V_s^2. \quad (1)$$

The behavior of soils at small strain amplitudes is important because it reflects their response to seismic waves during an earthquake. By knowing the V_s values of soils, engineers and geologists can evaluate their potential for liquefaction, settlement, and amplification of ground motions.

Building codes [2] often include provisions for seismic site classification based on V_s values, particularly for the top 30 m of soil. These codes provide guidelines for designing structures and foundations that can withstand anticipated ground shaking. V_s values are also crucial in micro zonation studies, which aim to assess and map the seismic hazard at urban and regional scale. These studies help in urban planning, land-use management, and risk mitigation strategies.

To obtain V_s measurements, both *in situ* and laboratory tests can be performed. *In situ* methods use specialized equipment to generate and record seismic waves directly in the ground [1]. Common *in situ* tests include the seismic piezocone penetration test (SCPTu), which integrates cone tip resistance (q_c), sleeve friction (f_s), pore pressure (u_2), and shear wave velocity (V_s) measurements within a single sounding [3]. This test is an extension of the conventional cone penetration test (CPT), which measures only q_c and f_s, and the piezocone (CPTu), which also records pore pressure. Another widely used technique is the Multichannel Analysis of Surface Waves (MASW) [4], which involves analyzing surface wave dispersion data recorded by an array of geophones to estimate V_s profiles with depth. MASW is particularly useful for shallow

Cite this article as: Al-Azazmeh, A., Mahler, A. "Modified Correlations to Predict Shear-wave Velocity Using Cone Penetration Test Data for Hungary", Periodica Polytechnica Civil Engineering, 69(4), pp. 1311–1320, 2025. https://doi.org/10.3311/PPci.41202

subsurface investigations, as it is non-invasive, efficient, and effective in characterizing layered soil deposits.

While both primary (P-) and secondary (S-) wave velocities can be measured *in situ*, V_s is preferred for seismic site characterization because it directly relates to the small-strain shear modulus, which governs soil deformation under earthquake loading. P-wave velocities are more influenced by pore fluid compressibility and are therefore less reliable for stiffness characterization in saturated soils [1].

Laboratory tests can also be conducted to determine the V_s of soil samples collected. These tests often involve the use of a torsional shear wave apparatus or bender elements, where the soil sample is subjected to controlled shear stress and the resulting shear wave propagation is measured.

However, laboratory measurements are inherently limited by sample disturbance during retrieval, changes in moisture content, and the inability to replicate *in situ* stress conditions. As a result, laboratory-derived V_s values may differ from field measurements, particularly in soft or sensitive clays.

This study seeks to identify the optimal empirical correlations for estimating shear wave velocity in diverse soil types across Hungary using data from the seismic cone penetration test (SCPTu). Recent local investigations have highlighted the importance of region-specific seismic site characterization: Al-Azazmeh and Mahler [5] developed CPT–V_s correlations for soft Quaternary clays in Szeged, while Al-Azazmeh and Mahler [6] performed nonlinear site-response analyses in District XIII, Budapest, demonstrating the critical role of accurate V_s profiling in urban areas.

The number and spacing of V_s measurements depend on the site's size, complexity, and the level of detail required for the analysis. A comprehensive understanding of the soil's behavior and its variability with depth is crucial for reliable seismic hazard assessment and design of resilient structures.

In Hungary, seismic site classification is guided by Eurocode 8 (MSZ EN 1998-1:2008, [2]), which has been adopted as the national standard. Site classes are determined primarily based on V_{s30} (the average V_s in the top 30 m), in accordance with categories A–E. The Hungarian annex specifies national parameters for design ground acceleration and site amplification factors.

In addition to depositional environment and grain size, geological factors such as over consolidation ratio (OCR), age, and cementation significantly influence V_s values. Older deposits and those with a higher OCR typically display greater stiffness due to diagenetic bonding, cementation, and secondary mineral formation. Cementation, particularly in silty and sandy strata, can increase small-strain stiffness beyond that predicted by density alone. These effects should be accounted for when applying empirical correlations, as they may lead to over- or under-estimation of V_s in soils with atypical geological histories [7].

2 Site description and data sources

Fig. 1 [8] presents the locations of the investigated sites, which include Sopron, Szeged, Kecskemét, the M8 Highway, and Budapest. For these areas, MASW and SCPTu data were collected and compiled into a comprehensive database. This study evaluates the data, estimates V_s values, and identifies the empirical correlations that best fit each of the studied Hungarian sites.

Budapest, the capital city of Hungary, possesses a diverse and captivating geology that encompasses various ages and formations. Its geological composition comprises a complex assemblage of Paleozoic marine sedimentary rocks, Mesozoic limestones and marls originating from a shallow sea, and Tertiary volcanic formations. The Quaternary period has played a significant role in shaping the region, with glaciations and the formidable Danube River leaving indelible imprints on the landscape.

Gyalog et al. [9] classified the Holocene sediments in the area into three distinct categories: fluvial, alluvial fan, and paludal sediments. Fluvial sediments are found in water-covered brooks and river valleys, while alluvial fan sediments form in the expansive alluvial plain of the Danube River. Paludal sediments dominate the valleys of the Pest Plain. Examination of the Pleistocene-Holocene period reveals a diverse range of sediment types, including aeolic sand, fluvial-aeolic sand, slope sediments, deluvial redeposited loess and clay-silt, fine-grained residual deposits, Upper Pleistocene fluvial sediments, Middle and

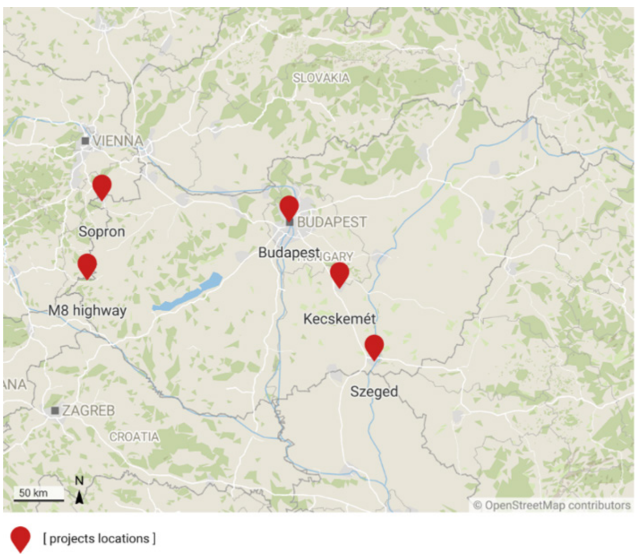

Fig. 1 Locations of sites in Hungary [8]

Lower Pleistocene fluvial sediments, loess, and travertine. These sediments provide invaluable insights into the geological history and composition of the region during the Pleistocene-Holocene period.

Moving to Kecskemét, the South Tisza Basin, a north-south graben-like structure harboring the deepest Quaternary basin in the Alföld region. Quaternary deposits extend over 600 m near the confluence of the Körös and Tisza Rivers. The western border of the basin is defined by an eastward dipping flank situated between the Danube and Tisza Rivers [10]. As we progress westward, the thickness of Quaternary deposits diminishes from 650 m to less than 50 m at the Danube. Fluvial deposits, referred to as the Kecskemét gravel formation, dominate the region. The Quaternary sequence, with a thickness of 200 m, overlays fine-grained Upper Pannonian sediments transported from the northwest. Pliocene-age variegated clays are absent, and the predominant color of the pelitic sediments is gray, both in the Quaternary and Upper Pannonia.

Situated in the Pannonian Basin above the Algyő High, the city of Szeged showcases sedimentary formations from the Neogene and Quaternary periods. The Neogene period [11] left behind marine sediments, while the Quaternary period witnessed the influence of glacial and alluvial deposits that have shaped the landscape. The pre-Cenozoic basement of the region belongs to the Codru nappe system within the Tisza mega-unit. The Algyő High comprises metamorphic rocks, siliciclastic formations, and shallow marine mud rocks and carbonates. In the Szeged Basin, the Lower Triassic period is represented by continental sandstones and variegated shales, while the Anisian and Ladinian periods feature shallow-marine lagoonal dolomites.

Moving westward to Sopron, a city located near the Austria-Hungary border, we encounter a geology profoundly influenced by the presence of the Eastern Alps and the Pannonian Basin. The geological history of Sopron spans various ages, with the oldest rocks dating back to the Paleozoic era and forming part of the Variscan orogeny. These ancient rocks consist of metamorphic formations, including gneisses and schists [11], that have undergone significant tectonic forces and high temperatures. During the Mesozoic era, Sopron and its surroundings were submerged under a shallow sea, resulting in the deposition of sedimentary rocks such as limestones, marls, and sandstones. The Quaternary period played a pivotal role in shaping Sopron, as glaciations carved U-shaped valleys, moraines, and drumlins across the landscape. The region's diverse rocks and formations owe their existence not only to glaciations but also to Tertiary volcanic activity and sedimentation within the neighboring Pannonian Basin.

For each site, we paired the nearest CPTu with the corresponding SCPTu/MASW profile. SCPTu provided ~1 m V_s increments. To make MASW comparable, its coarse layers were refined into finer sublayers so the vertical resolution better matched the CPTu/SCPTu data; the redistribution preserved each original MASW layer's average velocity to avoid bias. This produced higher-resolution MASW V_s suitable for direct depth-by-depth comparison with CPTu-derived and SCPTu values.

Fig. 2 presents a combined analysis of V_s profiles derived from MASW and SCPTu measurements across multiple sites in Hungary. The V_s values span a range of approximately 100–600 m/s, reflecting diverse subsurface conditions, from soft soils to stiff sediments.

2.1 Key observations

The compiled SCPTu and MASW datasets from the investigated Hungarian sites were jointly analyzed to identify characteristic patterns and site-specific behaviors in the measured V_s profiles. The following key observations summarize the main geological and geotechnical trends derived from this comparative assessment:

1. Consistency and variability: MASW data reveal a general increase in V_s with depth, suggesting a broadly homogeneous geological trend across the region. However, localized deviations in V_s at greater depths (e.g., between 15–24 m) indicate site-specific

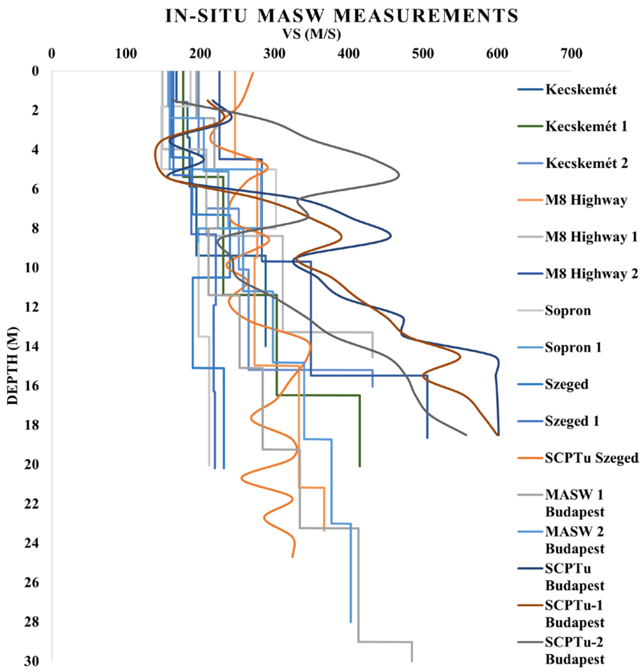

Fig. 2 Summary of all MASW data for sites in Hungary

variations, likely due to differences in soil composition (e.g., clay vs. sand layers) or bedrock topography.

2. Integration of SCPTu data: to complement the spatially averaged MASW results (which represent zones several meters thick), high-resolution SCPTu point measurements (recorded at 1–2 cm intervals) were incorporated. The SCPTu V_s values, typically valid for 0.5 m intervals, were upscaled to match the MASW resolution using a weighted averaging approach, ensuring consistency in depth-specific comparisons.

3 Soil classification

This research study focuses on the classification of soils using measurements from the CPT, specifically cone tip resistance (q_t, MPa), sleeve friction (f_s, kPa), and excess pore water pressure (u_2, measured just behind the cone tip). The main objective is to establish a robust soil classification system using the soil behavior type index (I_c), a dimensionless parameter derived from CPT data that characterizes soil type.

In 1990, Robertson [12] proposed a soil classification framework using the normalized cone resistance (Q_t, dimensionless) to classify soils into nine distinct SBTs. However, Jefferies and Davies [13] introduced the SBT chart in 1991, which provided a more accurate approach to classification.

Expanding upon the work of Jefferies and Davies [14], further refined the classification system in 1993 by introducing the soil index, I_c. Jefferies and Been [15] suggested refining the classification of these loose silts. Consequently, the correlation of I_c became:

$$I_{c(\mathrm{J\&B06})} = \sqrt{\left(3-\log\left[Q_t\left(1-B_q\right)+1\right]\right)^2+\left[1.5+1.3\left(\log F_R\right)\right]^2}, \quad (2)$$

where B_q is the pore pressure ratio, and F_R is the normalized friction ratio used in the calculation of I_c. To assess the sensitivity of I_c values to different formulations, both the Jefferies and Been [15] and Robertson and Wride [16] (most widely used) equations were evaluated for a subset of the data. The differences in computed I_c values were found to be minimal (within ±5%), indicating that either formulation yields similar soil classification in this context.

The soils investigated in this study predominantly exhibit clay-like behavior, characterized by high plasticity, low permeability, and a tendency to undergo volume changes (shrinkage and swelling) with variations in moisture content. They also display high cohesion and relatively low shear strength, making them more susceptible to deformation and stability issues. As shown in Robertson's [17] normalized SBT_n chart (Fig. 3), which refines the original chart proposed in 1990 [12], the data points from our sites are mainly distributed across zones 3 to 9. A significant concentration is observed in zones 3, 4, 5, and 8, indicating that clayey soils are dominant in the study area.

4 Statistical analyses of the best correlations

There are several empirical correlations available for estimating V_s. However, their applicability can vary depending on the specific soil type. Therefore, selecting the most suitable correlation for Hungarian soils is a critical task. The aim of this study is to evaluate the performance of different correlations in predicting V_s across various soil types in Hungary and to provide recommendations on the most reliable ones.

To achieve this, both measured and estimated V_s values from our database were analyzed using a range of statistical parameters. The selected correlations were subjected to rigorous statistical testing to evaluate their reliability and accuracy. Key evaluation metrics included correlation coefficients and goodness-of-fit indicators, which provide insight into how well each correlation model represents the observed data.

Common goodness-of-fit measures include R-squared (coefficient of determination), adjusted R-squared, root mean square error (RMSE), A higher R-squared value and lower RMSE indicate a better fit between the correlation model and the observed data. This analysis provided quantitative measures of the relationships between the variables and allowed for the identification of correlations that demonstrated strong statistical significance and predictive power.

In the assessment of correlations, the RMSE in Eq. (3) is commonly employed to measure the accuracy by comparing predicted values from the correlation with the actual

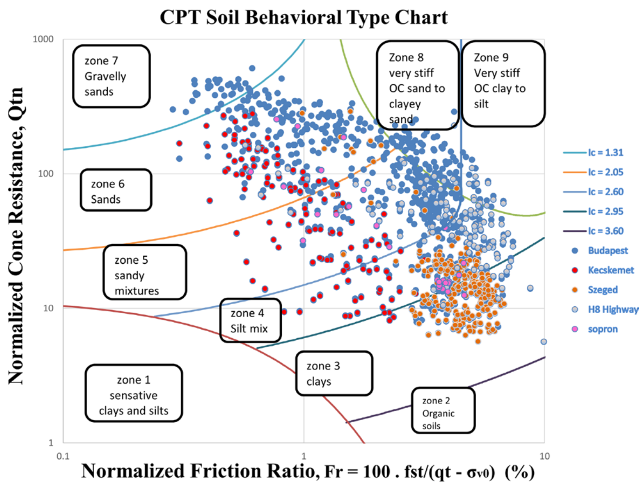

Fig. 3 Scattered data points (914) on the SBT_n chart for normalized cone penetration test (CPT) results

measured values. The correlation exhibiting the lowest RMSE value is considered the most effective in providing accurate predictions. Consequently, RMSE serves as a valuable tool in identifying correlations that offer precise estimations for the studied outcome variable.

And below are the statistical parameters used for the assessment of the correlations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(V_{s,measured} - V_{s,calculated}\right)^2}, \quad (3)$$

$$K = V_{s,calculated} / V_{s,measured}, \quad (4)$$

$$\Delta = V_{s,measured} - V_{s,calculated}, \quad (5)$$

$$\mathrm{RI} = \left|\mu_{\ln K}\right| + \sigma_{\ln K}, \quad (6)$$

$$C_{VK} = \mu_K / \sigma_K, \quad (7)$$

$$\mathrm{RD} = \sqrt{\left[1-\mu_K\right]^2 + \left[\sigma_K\right]^2}. \quad (8)$$

In Eqs. (3) to (8), $V_{s,measured}$ and $V_{s,calculated}$ represent the measured and predicted shear wave velocities, respectively; K is their ratio, μ_K and σ_K are the mean and standard deviation of K, and $\mu_{\ln K}$ and $\sigma_{\ln K}$ denote the mean and standard deviation of the natural logarithm of K.

Lower values of the indexes – RI, RMSE, RE, K, and RD – indicate a better correlation. Accuracy is determined by the proximity of the mean K to 1 or the mean Δ to 0.

Over 40 V_s-CPT correlations were assessed for this study Table 1 [17–23] displays the primary correlations that were evaluated in this study. In Andrus et al. [18] and Stuyts et al. [24], the focus has predominantly been on establishing direct correlations between V_s and parameters such as cone tip resistance (q_c) or cone total resistance (q_t). These correlations have proven to be simpler to apply and have yielded satisfactory results.

Over the years, various correlations have been proposed to estimate V_s using data obtained from CPT or piezocone penetration testing (CPTu) for different soil types:

1. For sands and cohesionless soils, correlations suggested by Hegazy and Mayne [21], Fear and Robertson [25], Andrus et al. [18], Piratheepan [23], and others were considered.
2. For clays and cohesive soils, correlations proposed by Piratheepan [23], Cai et al. [20], and others were examined.

Table 1 Previously suggested correlations between V_s and CPT or CPTu data for various soil classification

Authors	Correlations	Location/site	Soil	Notes
Andrus et al. [18]	$V_{s1} = 16.5 \cdot q_{t1N}^{0.411} \cdot I_c^{0.97}$	California, South Carolina and Japan	Holocene-Age Soils	
Andrus et al. [18]	$V_s = 2.62 \cdot q_t^{0.395} \cdot I_c^{0.912} \cdot z^{0.124} \cdot \mathrm{SF}^a$	California, South Carolina and Japan	Combined Holocene–Pleistocene soils	Scaling factor (SF) is a parameter that depends on soil deposit age, where a = 0.88–0.92 for Holocene-age sites and a = 1.11–1.12 for Pleistocene sites
Wolf [19]	$V_s = 11.97 \cdot q_t^{0.262} \cdot I_c^{0.709} \cdot z^{0.107}$	Hungary	All soils	
Cai et al. [20]	$V_s = 7.954 \cdot q_t^{0.403}$	Jiangsu, China	Soft clays	R^2 value of 0.631
Cai et al. [20]	$V_s = 38 \cdot Q_{tn}^{0.61}$	Jiangsu, China	Soft clays	Purely statistical analysis of the Jiangsu soft clay database
Hegazy and Mayne [21] (all soils)	$V_s = \left[\left(10.1 \log q_t - 11.4\right)\right]^{1.67} \cdot \left(100 \dfrac{f_s}{q_t}\right)^{0.3}$	Worldwide	Clays, sands, intermediate soils, and mine tailings	q_c, σ'_{V0} in kPa, V_s in m s^{-1}
Mayne [22]	$V_s = 118.8 \log f_s + 18.5$	Worldwide	Sands, clays and organic plastic clays	f_s in kPa
Piratheepan [23]	$V_s = 25.3 \cdot q_c^{0.163} \cdot f_s^{0.029} \cdot D^{0.155}$	California, Japan, and Canada	Sands	q_c, f_s in kPa, D in m
Piratheepan [23]	$V_s = 11.9 \cdot q_c^{0.269} \cdot f_s^{0.108} \cdot D^{0.127}$	California, Japan, and Canada	Clays	q_c, f_s in kPa, D in m
Robertson [17] (all soils)	$V_s = \left(10^{0.55 I_c + 1.68} \cdot \dfrac{q_t - \sigma'_V}{P_a}\right)^{0.5}$	Worldwide	Predominantly Holocene and Pleistocene-age general soils	q_c, σ'_{V0} in kPa, V_s in m s^{-1}
Robertson [17] (all soils) (q_{c1N})	$V_{s1} = \left(10^{0.55 I_c + 1.68} \cdot q_{c1N}\right)^{0.5}$	Worldwide	All soils	q_{c1N} is dimensionless

3. Correlations applicable to all soil types were reviewed, including those suggested by Hegazy and Mayne [19], Piratheepan [21], Hegazy and Mayne [24], Mayne [20], Andrus et al. [16], and Tonni and Simonini [25].

These correlations were evaluated to determine their suitability for estimating V_s in the current study.

Table 2 [7, 17–19, 21–23, 26–28] presents a summary of the analysis results for the parameters K, Coefficient of Variation of K (CVK), Relative Deviation (RD), Reliability Index (RI), and Root Mean Square Error (RMSE) for all soils. Similar tables were also prepared for Quaternary and Tertiary soils; however, these are not included in the paper. The correlations were evaluated based on whether they overestimated or underestimated the V_s. This classification was determined by comparing the mean K values or the mean difference between the measured and calculated shear wave velocities of the datasets.

If the average K value (K_{mean}) is greater than 1 or the mean difference value is negative, the correlation is classified as overestimated. The analysis revealed that most of the correlations had a K value larger than 1 or a negative Δ value, indicating an underestimation tendency for V_s.

Robertson [17] correlation had used the normalized net corrected cone resistance Q_{tn}. In most moderately dense sandy soils, where $q_t \sim q_c$ and $(q_t - \sigma_v) \sim q_c$, Q_{tn} is essentially equal to q_{c1N} as shown in Fig. 4. Replacing Q_{tn} with q_{c1N} in the relationship by Robertson [17] was tested and it was observed that this substitution (i.e., Robertson [17] q_{c1N}) demonstrated a slight improvement to the V_s estimation by giving a slightly lower RMSE value.

Fig. 4 presents a comparison between estimated and measured V_s for Quaternary soils in Budapest based on 687 data pairs, using selected empirical correlations: Wolf [19], Mayne [22], Robertson [17], and Robertson [17] q_{c1N}.

A total of 687 data pairs were analyzed for Quaternary soils, with measured V_s values ranging from 120 to 550 m/s and an average of 255 m/s. The strongest agreement with measured data was observed for Wolf [19] and Robertson [17] q_{c1N}, while all correlations tended to underestimate V_s values above 400 m/s. Linear trends and clustering below the 1:1 line indicate systematic underprediction, especially at greater depths where V_s increases.

Fig. 5 shows the comparison between estimated and measured V_s for Tertiary soils at the study sites, based on 227 data pairs. Measured V_s values ranged from 250 to 600 m/s, with an average of 360 m/s, primarily between depths of 7 and 30 m. The selected empirical correlations, Wolf [19], Andrus et al. [18], Robertson [17], and Robertson [17] q_{c1N}, show improved performance when applied specifically to Tertiary deposits. Among them, Andrus et al. [18] and Robertson [17] provided the most accurate predictions, with RMSE values of 60.1 and 58.3, respectively. While all correlations generally follow a linear trend, they tend to underestimate V_s values exceeding 400 m/s, particularly in deeper, denser soil layers. The Wolf [19] Pleistocene formula, although originally developed for Quaternary soils, demonstrated strong predictive performance for the

Table 2 Results of the statistical analysis for all soils data group (as an example, similar tables were calculated for other soil groups)

Correlation	Δ_{mean}	$\lvert\Delta_{mean}\rvert$	K_{mean}	CVK	RD	RI	RMSE	Estimation
Hegazy and Mayne [26] (all soils)	−33.16	33.16	1.15	3.65	0.35	0.40	81.84	Overestimated
Hegazy and Mayne [21] (all soils)	−22.08	22.08	1.13	4.90	0.27	0.31	58.10	Overestimated
Mayne [22] (all soils)	−3.59	3.59	1.06	4.98	0.22	0.25	52.44	Overestimated
Mayne and Rix [7] (intact clays)	53.73	53.73	0.78	3.11	0.33	0.57	80.68	Underestimated
Robertson [17] (all soils)	47.22	47.22	0.81	6.49	0.23	0.38	61.11	Underestimated
Robertson [17] q_{c1N} (all soils)	43.21	43.21	0.83	6.47	0.21	0.35	58.88	Underestimated
Piratheepan [23] (all soils)	62.80	62.80	0.74	5.94	0.29	0.48	75.40	Underestimated
Andrus et al. [18] (all soils)	47.18	47.18	0.81	6.19	0.23	0.39	60.25	Underestimated
Wolf [19] (Holocene fluvial sediments)	35.79	35.79	0.86	5.24	0.22	0.37	56.56	Underestimated
Wolf [19] (tertiary deposits)	−5.63	5.63	1.04	3.22	0.33	0.36	72.59	Overestimated
Wolf [19] (Pleistocene fluvial sediments)	46.76	46.76	0.82	6.29	0.22	0.37	61.67	Underestimated
Wolf [19] (Pleistocene eolian sediments)	−33.06	33.06	1.18	5.19	0.29	0.35	62.89	Overestimated
Wolf [19] (Holocene-Pleistocene)	6.09	6.09	1.00	6.19	0.16	0.18	44.14	Underestimated
Wolf [19] (all soils)	0.31	0.31	1.03	6.34	0.16	0.18	42.28	Overestimated
Karray et al. [28]	44.50	44.50	0.86	3.06	0.31	0.49	82.55	Underestimated
Tonni and Simonini [27] (sands and silts)	17.73	17.73	0.99	3.26	0.30	0.35	73.97	Underestimated

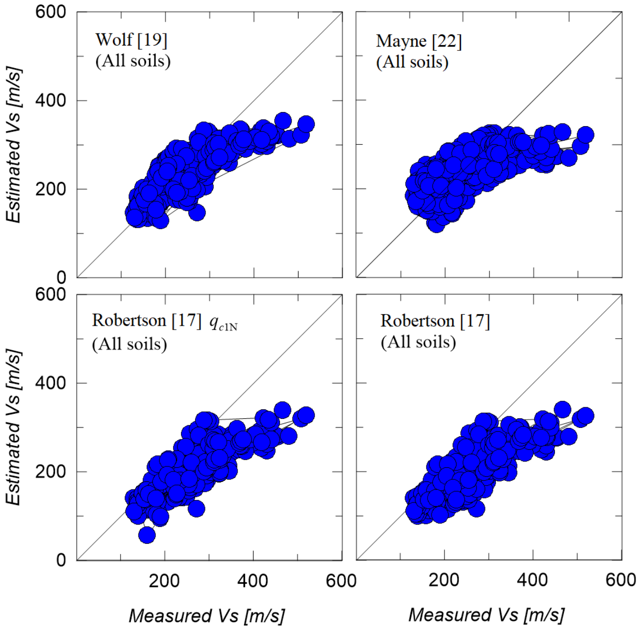

Fig. 4 Estimated V_s against measured V_s of Quaternary age deposits

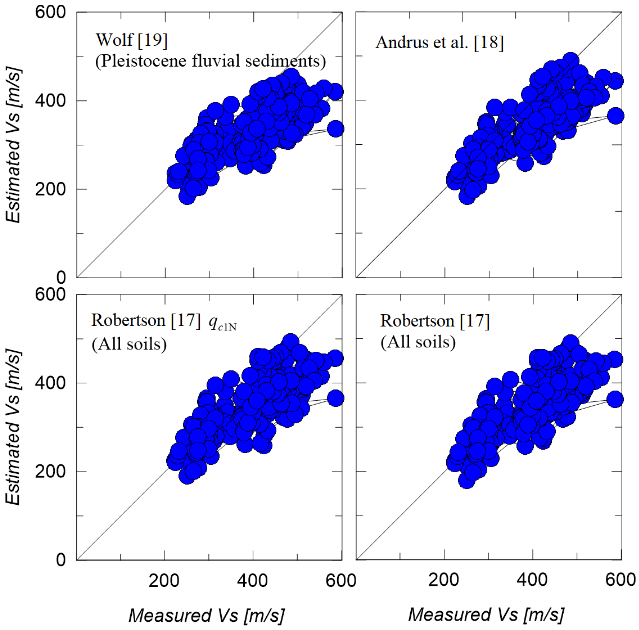

Fig. 5 Estimated V_s against measured V_s of Tertiary age deposits

Tertiary dataset as well; therefore, it was included in Fig. 5 for comparative purposes.

Fig. 6 presents a comparison between estimated and measured V_s for all soil types using selected empirical correlations, based on 914 data points. The analysis includes a broad range of soils with V_s values extending up to 800 m/s. Among the tested models, those by Wolf [19], particularly the Holocene-Pleistocene and all-soils variants and Robertson [17] q_{c1N} demonstrated strong linear trends and close agreement with measured data. RMSE values were lowest for Wolf [19] (56.4) and Robertson [17] q_{c1N} (56.2), confirming their robustness across diverse soil conditions.

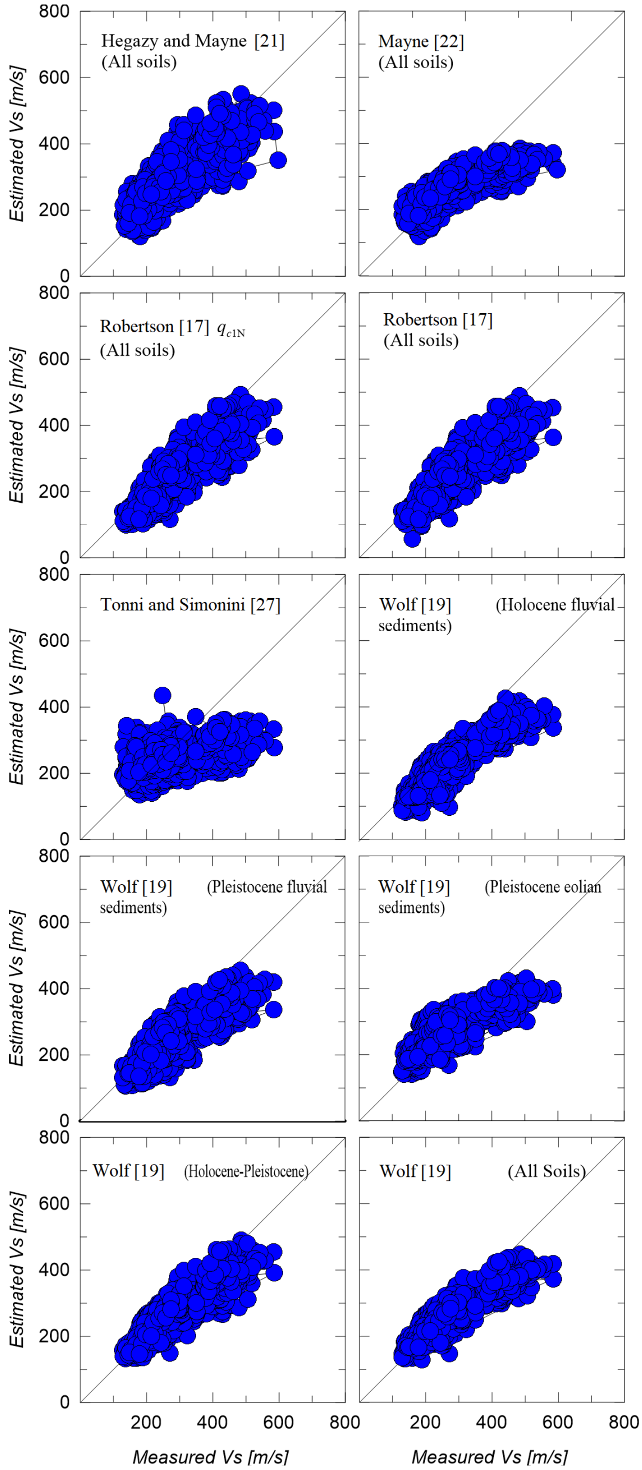

Fig. 6 Estimated V_s against measured V_s of all soil's deposits

5 New modified correlations

To achieve this optimization, the RMSE was minimized by adjusting the multipliers (factors) of the input parameters for each correlation tested. Robertson [17] (with replacing Q_{tn} with q_{c1N}) correlation was selected as the model equation to generate the proposed correlations:

$$V_s = \left(\sigma'_V / P_a\right)^{0.25} \cdot \left(10^{aI_c+b} \cdot q_{c1N}\right)^c . \quad (9)$$

Truly normalized (i.e., dimensionless) cone penetration resistance corrected for overburden stress (q_{c1N}) and I_c parameters in these correlations contribute to a better prediction of V_s and help reflect the geological site characteristics in reality.

Through the optimization process, notable improvements were obtained in the model equation for the studied categories.

All soils: based on the (914) data pairs the recommended best-fit equation for predicting V_s in m/s is:

$$V_s = \left(\sigma'_V / P_a\right)^{0.25} \cdot \left(10^{0.398 I_c + 3.65} \cdot q_{c1N}\right)^{0.387}. \quad (10)$$

Fig. 7 shows the proposed correlation (Eq. (10)) for all soils compared to the original one (before regression), most of the predicted values of V_s are within 20% of the measured V_s, and 36.6% improvement in RMSE.

Quaternary-age soils: based on the (687) quaternary data pairs, the recommended best-fit equation for predicting V_s in m/s is:

$$V_s = \left(\sigma'_V / P_a\right)^{0.25} \cdot \left(10^{0.372 I_c + 4.08} \cdot q_{c1N}\right)^{0.368}. \quad (11)$$

Tertiary-age soils: for predicting V_s in m/s, the recommended best-fit equation based on (227) data pairs is:

$$V_s = \left(\sigma'_V / P_a\right)^{0.25} \cdot \left(10^{0.545 I_c + 2.1} \cdot q_{c1N}\right)^{0.47}. \quad (12)$$

Figs. 7 and 8 illustrate the performance of the proposed Eq. (10) for estimating V_s across 914 soil samples. In Fig. 7, the "original correlation" refers to the initial output of Eq. (10) prior to any regression adjustment. This baseline correlation, which reflects the general trend observed before model calibration, yielded a RMSE of 58.8 m/s.

After applying regression analysis to improve the model's fit, the RMSE was significantly reduced to 37.27 m/s, demonstrating enhanced predictive accuracy. The improvement is evident by the tighter clustering of data points around the 1:1 reference line in the scatter plot.

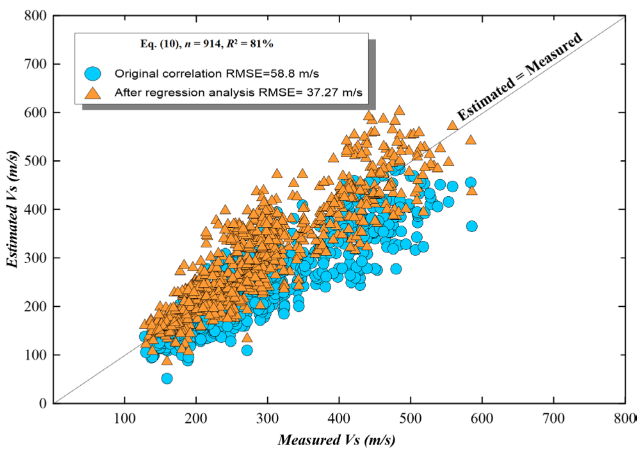

Fig. 7 The proposed correlation (Eq. (10)) for all soils compared to the original one (before and after regression)

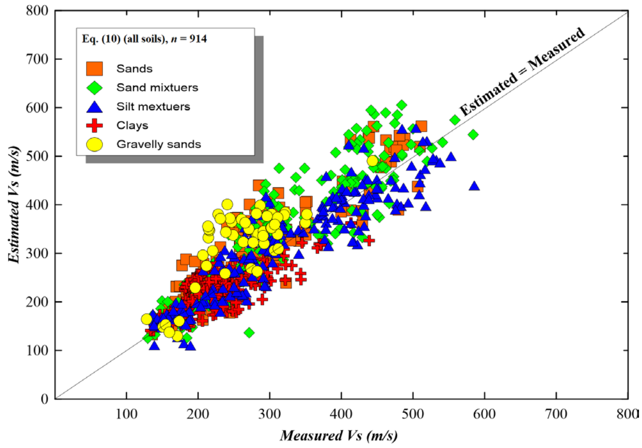

Fig. 8 Estimated V_s against measured V_s categorized by soil type, including sands, mixtures, clays, and gravelly sands. This representation highlights the model's varying accuracy across different soil classes7 conclusion and recommendation

In Fig. 8, the data are categorized by soil type rather than soil age as used in previous representations. This shift allows for a more meaningful interpretation of the model's performance, as soil type directly influences geotechnical behavior and V_s propagation. The results show that the model performs better for coarse-grained soils such as sands and gravelly sands, with data points closely aligned with the 1:1 line. In contrast, finer-grained soils like clays exhibit greater dispersion, reflecting the natural variability and complexity in predicting V_s in cohesive materials. This classification provides important insight into the model's sensitivity to soil composition and supports its application across diverse geological contexts.

6 Conclusions

This study evaluated and improved empirical correlations for estimating V_s from CPT data, focusing on Hungarian soils. A comprehensive database of 914 V_s–CPT pairs from five cities with varied geology was compiled. Over 40 existing correlations were assessed using statistical indicators such as RMSE, K, RD, RI, and CVK. Results showed that many widely used models over- or underestimate V_s when applied to local conditions, especially in fine-grained soils.

To enhance prediction accuracy, regression-based modifications were applied to the best-performing models, resulting in significantly improved correlations. The proposed equation for all soils (Eq. (10)) reduced RMSE from 58.8 m/s to 37.27 m/s, showing a 36.6% improvement and tighter clustering of estimated V_s around the 1:1 line. Further analysis by soil type revealed better model performance in coarse-grained soils (e.g., sands, gravelly sands) and more variability in cohesive soils like clays, emphasizing the influence of soil composition on V_s prediction.

These findings highlight the need for region-specific and soil-type-sensitive models in seismic site characterization. The modified correlations provide a reliable tool for Hungarian geotechnical practice, improving V_s estimates is essential for site classification, liquefaction assessment, and seismic design. Future research should consider advanced modeling techniques and expanded datasets to further enhance predictive capabilities across complex subsurface conditions.

References

[1] Kramer, S. L. "Geotechnical Earthquake Engineering", Prentice-Hall, 1996. ISBN 978-0-13-374943-4

[2] MSZT "MSZ EN 1998-1:2008 Eurocode 8: Tartószerkezetek tervezése földrengésre. 1. rész: Általános szabályok, szeizmikus hatások és az épületekre vonatkozó szabályok" (MSZ EN 1998-1:2008 Eurocode 8: Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings), Hungarian Standards Institute (MSZT), Budapest, Hungary, 2008. (in Hungarian)

[3] Robertson, P. K., Cabal, K. L. "Guide to Cone Penetration Testing for Geotechnical Engineering", [pdf] Gregg Drilling & Testing, Inc., Signal Hill, CA, USA, 2015. Available at: https://www.g-i.co.nz/wp-content/uploads/greggs-cpt-guide-6th-2015.pdf [Accessed: 17 August 2025]

[4] Park, C. B., Miller, R. D., Xia, J. "Multichannel analysis of surface waves", Geophysics, 64(3), pp. 800–808, 1999. https://doi.org/10.1190/1.1444590

[5] Al-Azazmeh, A., Mahler, A. "Development of CPT–V_s correlations for soft quaternary clays of Szeged city Hungary", In: Proceedings of the XVIII European Conference on Soil Mechanics and Geotechnical Engineering (ECSMGE 24), Lisbon, Portugal, 2024, pp. 282–285. ISBN 978-1-032-54816-6 https://doi.org/10.1201/9781003431749-25

[6] Al-Azazmeh, A., Mahler, A. "Nonlinear Site Response Analysis in District XIII, Budapest", Slovak Journal of Civil Engineering, 33(3), pp. 54–61, 2025. https://doi.org/10.2478/sjce-2025-0020

[7] Mayne, P. W., Rix, G. J. "Correlations Between Shear Wave Velocity and Cone Tip Resistance in Natural Clays", Soils and Foundations, 35(2), pp. 107–110, 1995. https://doi.org/10.3208/sandf1972.35.2_107

[8] Datawrapper "Create a map", [online] Available at: https://app.datawrapper.de/select/map [Accessed: 17 August 2025].

[9] Gyalog, L., Maros, G., Pelikán, P. (eds.) "Budapest geokalauza" (Budapest geoguide), Hungarian Institute of Geology and Geophysics, 2016. ISBN 9789636713041 (in Hungarian)

[10] Csiky, G. "Lithological and geological study of the Pliocene and Quaternary deposits in the Danube–Tisza interfluve region", Mineralogica et Petrographica Acta, 17(2), pp. 131–142, 1963.

[11] Benkó, Z., Bugledits, É., Récsi, A. "Ductile and brittle deformations in the Sopron Gneiss Formation (Sopron Mts., W-Hungary) and their implication on the formation of the Pannonian Basin", Zeitschrift für Geomorphologie, 56(Suppl. 2), pp. 3–19, 2012. https://doi.org/10.1127/0372-8854/2012/S-00082

[12] Robertson, P. K. "Soil classification using the cone penetration test", Canadian Geotechnical Journal, 27(1), pp. 151–158, 1990. https://doi.org/10.1139/t90-014

[13] Jefferies, M. G., Davies, M. P. "Soil classification by the cone penetration test: Discussion", Canadian Geotechnical Journal, 28(1), pp. 173–176, 1991. https://doi.org/10.1139/t91-023

[14] Jefferies, M. G., Davies, M. P. "Use of CPTU to Estimate Equivalent SPT N60", Geotechnical Testing Journal, 16(4), pp. 458–468, 1993. https://doi.org/10.1520/GTJ10286J

[15] Jefferies, M., Been, K. "Soil Liquefaction: A Critical State Approach", CRC Press, 2006. ISBN 9780203301968

[16] Robertson, P. K., Wride, C. E. "Evaluating cyclic liquefaction potential using the cone penetration test", Canadian Geotechnical Journal, 35(3), pp. 442–459, 1998. https://doi.org/10.1139/t98-017

[17] Robertson, P. K. "Interpretation of cone penetration tests – a unified approach", Canadian Geotechnical Journal, 46(11), pp. 1337–1355, 2009. https://doi.org/10.1139/T09-065

[18] Andrus, R. D., Mohanan, N. P., Piratheepan, P., Ellis, B. S., Holzer, T. L. "Predicting shear-wave velocity from cone penetration resistance", presented at the 4th International Conference on Earthquake Geotechnical Engineering (4ICEGE), Thessaloniki, Greece, June 25–28, 2007.

[19] Wolf, Á. "Cölöpalapok viselkedése szeizmikus terhelés hatására szemcsés talajkörnyezetben" (Behavior of pile foundations under seismic loading in granular soil environment), PhD Thesis, Széchenyi István University, 2018. (in Hungarian) https://doi.org/10.15477/SZE.MMTDI.2019.004

[20] Cai, G., Puppala, A. J., Liu, S. "Characterization on the correlation between shear wave velocity and piezocone tip resistance of Jiangsu clays", Engineering Geology, 171, pp. 96–103, 2014. https://doi.org/10.1016/j.enggeo.2013.12.012

[21] Hegazy, Y. A., Mayne, P. W. "Statistical correlations between V_s and cone penetration data for different soil types", In: Proceedings of the International Symposium on Cone Penetration Testing (CPT'95), Linköping, Sweden, 1995, pp. 173–178.

[22] Mayne, P. W. "In-Situ Test Calibrations for Evaluating Soil Parameters", In: Proceedings of the Second International Workshop on Characterisation and Engineering Properties of Natural Soils, Singapore, Singapore, 2006, pp. 1602–1652. ISBN 9780415426916 https://doi.org/10.1201/NOE0415426916.ch2

[23] Piratheepan, P. "Estimating Shear-Wave Velocity from SPT and CPT Data", MSc Thesis, Clemson University, 2002.

[24] Stuyts, B., Weijtjens, W., Jurado, C. S., Devriendt, C., Kheffache, A. "A Critical Review of Cone Penetration Test-Based Correlations for Estimating Small-Strain Shear Modulus in North Sea Soils", Geotechnics, 4(2), pp. 604–635, 2024. https://doi.org/10.3390/geotechnics4020033

[25] Fear, C. E., Robertson, P. K. "Estimating the undrained strength of sand: a theoretical framework", Canadian Geotechnical Journal, 32(5), pp. 859–870, 1995. https://doi.org/10.1139/t95-082

[26] Hegazy, Y. A., Mayne, P. W. "A Global Statistical Correlation between Shear Wave Velocity and Cone Penetration Data", In: Proceedings of GeoShanghai 2006, Shanghai, China, 2006, pp. 243–248. ISBN 978-1-7138-5198-1 https://doi.org/10.1061/40861(193)31

[27] Tonni, L., Simonini, P. "Shear wave velocity as function of cone penetration test measurements in sand and silt mixtures", Engineering Geology, 163, pp. 55–67, 2013. https://doi.org/10.1016/j.enggeo.2013.06.005

[28] Karray, M., Lefebvre, G., Ethier, Y., Bigras, A. "Influence of particle size on the correlation between shear wave velocity and cone tip resistance", Canadian Geotechnical Journal, 48(4), pp. 599–615, 2011. https://doi.org/10.1139/t10-092