

Cite this article

Muslim AF, Nakao R, Tsukumo T and Hino T
Mixing intensity and spatial strength variability in laboratory deep-mixed columns.
Proceedings of the Institution of Civil Engineers – Ground Improvement,
<https://doi.org/10.1680/jgrim.26.00039>

Review Article

Paper 2600039

Received 29/05/2026; Accepted 26/08/2026

Published with permission by Emerald Publishing Limited under the CC-BY 4.0 license.
(<http://creativecommons.org/licenses/by/4.0/>)

Ground Improvement



Mixing intensity and spatial strength variability in laboratory deep-mixed columns

Afif Fauzan Muslim

Civil Engineering and Architectural Design Course, Graduate School of Science and Engineering, Saga University, Saga, Japan (Orcid:0009-0000-2949-6490)
(corresponding author: afiffauzanmuslim96@gmail.com)

Ryota Nakao

Civil Engineering and Architectural Design Course, Graduate School of Science and Engineering, Saga University, Saga, Japan
(Orcid:0009-0007-2164-9932)

Tadashi Tsukumo

Department of Technical Service, YBM Service Corporation, Saga, Japan

Takenori Hino

Department of Civil Engineering and Architecture, Faculty of Science and Engineering, Saga University, Saga, Japan (Orcid:0009-0009-2688-4741)



Deep mixing method (DMM) column quality is commonly characterised by the coefficient of variation (COV) of unconfined compressive strength, which cannot distinguish depth-wise variability from local horizontal heterogeneity. Internal strength variability was characterised for laboratory deep-mixed columns prepared at four blade revolution counts, T2, T4, T6 and T8 (2000, 4000, 6000 and 8000 rev/m), using a miniature DMM apparatus on kaolin clay. Needle penetration tests at 20 positions on 11–12 slices per column provided spatially resolved q_{uNP} data. Mean q_{uNP} did not increase monotonically with blade revolution count, whereas overall COV decreased from 0.168 at T2 to 0.089 at T8. A two-component variance decomposition revealed distinct spatial patterns: local horizontal heterogeneity dominated at T2 and T4, depth-structured variability dominated at T6, and both components were substantially reduced at T8. Normal and lognormal assumptions were not consistently supported across mixing intensities, with tentative support for a normal approximation found only at T8. Vertical correlation lengths were comparable to or shorter than the 3 cm slice spacing, and horizontal autocorrelation was negligible at the present grid resolution. Overall COV alone is therefore insufficient for DMM quality characterisation; depth-structured columns require denser vertical sampling, whereas locally heterogeneous columns require greater horizontal coverage.

Keywords: autocorrelation/deep mixing method/ground improvement/laboratory test/spatial variability/statistical analysis/UN SDG 9: Industry, innovation and infrastructure/variance decomposition

Notation

A^{*2}	Stephens-modified Anderson–Darling statistic
$C(0)$	sample variance of detrended residuals
$C(\Delta)$	sample spatial covariance at lag Δ
COV_{ss}	within-slice COV
d	separation distance used in the autocorrelation model
d_{ij}	distance between measurement locations i and j
h	horizontal lag distance
$l_p(\theta)$	profile log-likelihood function
M	total number of blade wings
M^*	number of vertical profiles used in MLE
N	total number of measurement points per column
NP	needle penetration gradient (N/mm)
N_d, N_u	rotational speeds during downward and upward movements (rev/min)
n_s	number of measurement points in slice s
p	probability value used in goodness-of-fit tests
q_u	unconfined compressive strength, UCT (kPa)
q_{uNP}	converted UCT from needle penetration test (kPa)
$q_{uNP,s,i}$	converted strength at point i in slice s
$\bar{q}_{uNP}$	overall column mean of q_{uNP}
$\bar{q}_{uNP,s}$	mean q_{uNP} of slice s

$R(\theta)$	exponential correlation matrix
R_{jk}	element of $R(\theta)$
S	number of horizontal slices in a column
T	blade revolution count (rev/m)
V_d, V_u	downward and upward moving speeds of mixing blade (m/min)
$w_{s,i}$	slice-mean-detrended residual at slice s , position i
z_j, z_k	vertical coordinates of slices j and k
Δ	spatial lag distance
Δz	vertical slice spacing
η_h	local horizontal variance ratio
η_z	depth-structured variance ratio
θ	exponential correlation-length parameter (cm)
θ_v	maximum likelihood estimate of vertical correlation length (cm)
$\rho(\Delta)$	empirical autocorrelation function, ACF, at lag Δ
$\rho(d)$	theoretical ACF at separation distance d
σ^2	process variance of exponential covariance model (kPa ²)
σ_h^2	local horizontal, variance of q_{uNP} (kPa ²)
σ_{total}^2	total population variance of q_{uNP} (kPa ²)
$\sigma_{u,s}^2$	within-slice population variance of slice s (kPa ²)

σ_z^2 depth-structured, variance of q_{uNP} (kPa²)
 σ_{ln} lognormal spread parameter

1. Introduction

The deep mixing method (DMM) is an in situ ground improvement technique in which cement slurry is mechanically mixed with soft soil to form stabilised columns. Column quality depends on the uniformity of cement distribution, which controls both the mean and variability of unconfined compressive strength, q_u (Zuo *et al.*, 2023). Strength variability in DMM has been examined probabilistically since Honjo (1982) and later studies have developed statistical frameworks for interpreting core strength data, representative strength and design strength while accounting for variability, sample size and reliability criteria (Liu *et al.*, 2008, 2017). More recent studies have shown that mean strength and standard deviation alone are insufficient because spatial correlation affects column behaviour and reliability (Namikawa, 2019; Namikawa and Koseki, 2013). Construction and mixing parameters also influence binder distribution and strength variability, although their effects on spatial correlation are not always consistent (Chen *et al.*, 2016; Larsson *et al.*, 2005; Zuo *et al.*, 2023).

Existing characterisations commonly treat a DMM column as a single statistical population and report only the overall coefficient of variation (COV). This approach does not distinguish whether q_u variability is organised along the depth axis or scattered locally within horizontal cross-sections. The distinction is important because geotechnical variability is commonly described by both COV and spatial correlation structure, which govern how uncertainty is averaged over space in reliability-based design (Phoon and Kulhawy, 1999). In cement-treated columns, spatial autocorrelation can reduce the overall column strength relative to a uniform column with the same mean strength (Namikawa and Koseki, 2013), and the reliability of DMM column-supported embankments depends on mean strength, COV and spatial correlation length (Wijerathna and Liyanapathirana, 2019). Therefore, two columns with the same overall COV may require different interpretations if one is dominated by depth-structured variability and the other by local horizontal heterogeneity.

The blade revolution count T (rev/m) is the principal mixing-intensity parameter in Japanese DMM practice. The rational mixing number concept of Ota *et al.* (2000) describes a threshold beyond which further increases in T produce diminishing returns. However, how T governs the internal spatial structure of q_u remains unclear. Although Chen *et al.* (2016) linked binder mass-fraction variability to strength variability and showed that improved mixing reduces overall COV, the partitioning of residual variability between depth-structured and local horizontal components has not been systematically examined.

This study forms part of a broader laboratory programme for evaluating DMM blade performance; however, the present paper

focuses on one representative blade design operated at four blade revolution counts: T2, T4, T6 and T8, corresponding to 2000, 4000, 6000 and 8000 rev/m, respectively, with all other mix parameters held constant. Needle penetration tests at 20 positions on 11–12 slices per column provided spatially resolved values of q_{uNP} , the converted unconfined compressive strength. Four complementary analyses are reported: (i) descriptive statistics of q_{uNP} , establishing the mean and COV trends as a function of T ; (ii) decomposition of total q_u variance into depth-structured and within-slice components, an analysis that, to the authors' knowledge, has not previously been applied to DMM strength data; (iii) parametric distribution fitting with formal goodness-of-fit testing; and (iv) spatial autocorrelation analysis with estimation of the vertical correlation length, interpreted within the broader geostatistical framework of scale-of-fluctuation estimation (Lloret-Cabot *et al.*, 2014). Together these provide a more complete spatial characterisation of how mixing intensity governs DMM column quality than is achievable from overall COV alone.

2. Experimental programme

2.1 Specification of model apparatus

A miniature DMM apparatus was developed to reproduce the slurry-type mechanical mixing process under controlled laboratory conditions. Figure 1(a) shows a representative field-scale DMM machine, while Figure 1(b) shows the laboratory apparatus used in this study. Figure 2 shows the soil tank and experimental configuration. The apparatus comprises a mixing machine, cement-slurry delivery system and soil tank. The soil tank has internal dimensions of 200 mm × 200 mm in plan and 600 mm in height, and was filled with kaolin to a depth of 500 mm. Blade rotation speed can be controlled from 0 to 34 rpm, and vertical penetration and extraction speed from 0 to 1.0 m/min.

The mixing blade assembly is shown in Figure 3. It comprises six wings along a central shaft of 130 mm dia. (1/10 of a representative full-scale diameter of 1300 mm): four upper mixing wings to agitate the soil and distribute cement slurry, two lower cutting wings to assist penetration, and a horizontal anti-co-rotation plate to suppress co-rotation of the soil mass. Cement slurry was stored in a tank adjacent to the apparatus and delivered pneumatically through an internal conduit to the blade tip, where it was ejected into the surrounding soil. Ejection pressure was maintained at approximately 9 kPa. Since no flow meter was installed, slurry delivery was monitored by tracking the tank-level decrease against pre-calculated target values.

During each test, the blade was rotated and penetrated downward through the kaolin at the prescribed speed, with slurry ejected continuously throughout the 450 mm effective improved zone. On reaching the base, the rotation was reversed and the blade extracted upward at the same speed without further slurry ejection. The effective treated column diameter after mixing and curing was approximately 150 mm, slightly larger than the blade diameter owing to lateral soil displacement during rotation.

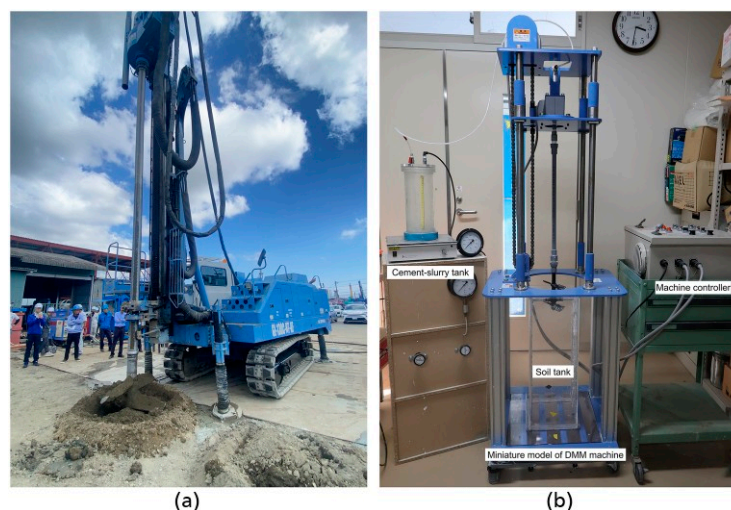


Figure 1. DMM mixing systems: (a) representative field-scale DMM machine and (b) miniature laboratory DMM apparatus

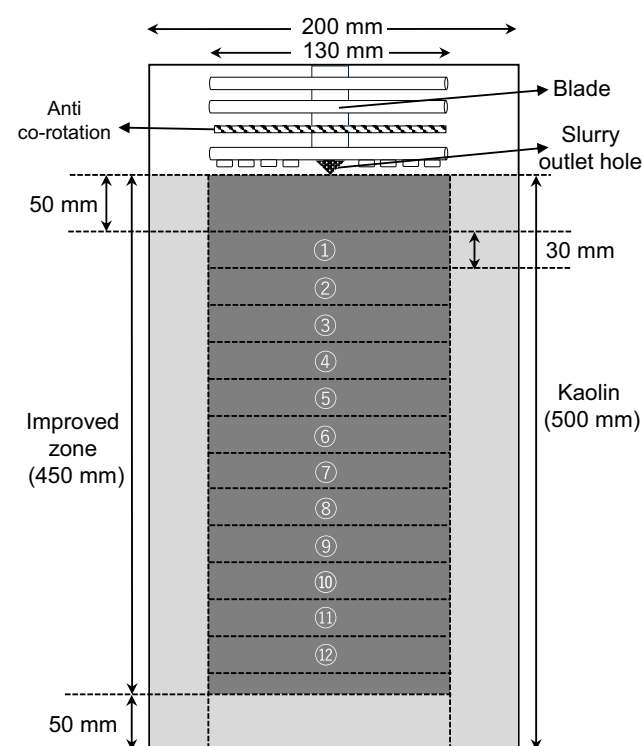


Figure 2. Schematic cross-section of the soil tank and experimental configuration, showing the kaolin layer, improved zone, blade position and slicing arrangement

2.2 Preparation of base soil and cement slurry

The base soil was prepared from powdered kaolin with a liquid limit of 66.6%, a plastic limit of 34.7% and a particle density of 2.61 mg/m^3 . Purified water was added to achieve a water content of 133.2%, equivalent to twice the liquid limit and corresponding

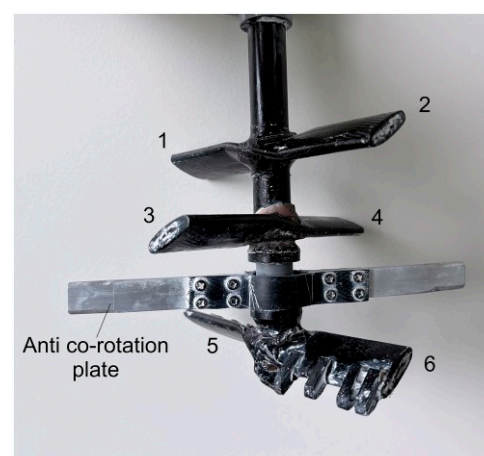


Figure 3. Photograph of the mixing blade assembly showing the six wing components and horizontal anti-co-rotation plate

to a liquidity index of 3.1. This high-fluidity condition was adopted for two reasons: it ensures uniform and reproducible filling of the soil box, minimising the influence of initial heterogeneity on the results; and the torque capacity of the miniature apparatus is limited, such that higher soil stiffness would have prevented adequate blade rotation. Although a water content of approximately 1.0–1.5 times the liquid limit is more representative of natural soft clay, the primary objective is to compare strength variability across mixing conditions rather than to reproduce absolute field behaviour.

The binder was a cement-based stabiliser, Solid Ace #100 (Aso Cement Co.), with a density of 3.13 mg/m^3 . A cement addition rate of 150 kg/m^3 and a water-cement ratio of 1.5 were adopted for all mixing conditions. An admixture at 5% of dry cement

content was included to enhance slurry fluidity and prevent outlet clogging. These proportions were held constant to isolate the effect of mixing intensity.

2.3 Mixing conditions and definition of blade revolution count

Mixing intensity was controlled through the blade revolution count, which is defined as the total number of blade passes per unit length of the improved zone during a single penetration and extraction cycle. Following the standard definition in Japanese DMM (Japanese Cement Association, 2021), T is expressed as:

1.
$$T = \sum M \left(\frac{N_d}{V_d} + \frac{N_u}{V_u} \right)$$

where T is the blade revolution count (rev/m); N_d and N_u are the rotational speeds during downward and upward movements (rev/min), respectively; V_d and V_u are the downward and upward moving speeds of the mixing wings (m/min), respectively; and M denotes the total number of blades composing the mixing wings. In the model tests, $M = 6$ and $V_d = V_u = 0.051$ m/min were fixed for all cases. The four mixing conditions examined in this study are summarised in Table 1. These cases are referred to as T2, T4, T6 and T8 corresponding to $T = 2000, 4000, 6000$ and 8000 rev/m, respectively. The range of T from 2000 to 8000 rev/m was selected to span conditions from clearly under-mixed to very well-mixed states.

Table 1. Mixing conditions adopted in the experimental programme

Case	Number of blade rotations, T : rev/m	Rotational speed, N : rev/min	Penetration speed, V_d : m/min
T2	2000	8.5	0.051
T4	4000	17.0	0.051
T6	6000	25.5	0.051
T8	8000	34.0	0.051

All cases used identical soil preparation, cement type, cement dosage, slurry mix proportion and curing condition, so that differences in the resulting q_u distribution can be attributed solely to the difference in mixing intensity.

2.4 Specimen preparation and needle penetration test

After mixing, each column was cured in the soil box at $20 \pm 2^\circ\text{C}$ for 7 days and then extracted. Each extracted column had a total length of approximately 450 mm and was cut into 11 or 12 horizontal slices of approximately 30 mm thickness using wire saws, starting from a depth of 50 mm. Exposed surfaces were smoothed with a straight blade knife to ensure reliable needle contact. Needle penetration tests were performed on each exposed surface in accordance with JGS 3431-2012 using a table-mounted device. The needle was pushed perpendicularly into the surface at a constant rate, and the applied load P (N) and penetration depth L (mm) were recorded until $L = 10$ mm. The needle penetration gradient NP (N/mm) was computed as P/L , and the converted unconfined compressive strength q_{uNP} was obtained from the empirical formula of Uchida *et al.* (2004):

2.
$$q_{uNP} = 41.8 \times NP - 4$$

Figure 4(a) shows a typical exposed slice surface after cutting and smoothing. On each exposed surface, a square measurement grid with 20 mm spacing in both the x - and y -directions was

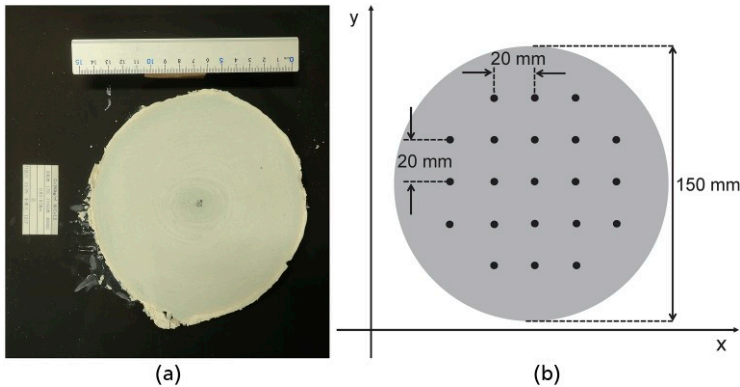


Figure 4. Measurement grid: (a) photograph of a typical exposed slice surface and (b) needle penetration testing plan view showing the 20 measurement positions within the 150 mm effective column diameter

established, as shown in Figure 4(b). The original grid comprised 21 points, but the centre point was excluded because cement tended to accumulate near the blade axis, giving a strength response unrepresentative of the surrounding mixed soil. The remaining 20 points were used at every slice depth and for all mixing conditions, ensuring spatial consistency throughout the dataset.

3. Descriptive statistical analysis

3.1 Overall strength characteristics

Table 2 summarises the overall statistical characteristics of the q_{uNP} for the four mixing conditions. Each of T2, T4 and T6 comprised 240 measurements from 12 slices, whereas T8 comprised 220 measurements from 11 slices because slice 12 was lost during specimen extraction. The grand mean q_{uNP} ranges from 136.5 kPa at T8 to 152.3 kPa at T4, with no monotonic trend, indicating that increasing blade revolution count mainly redistributes and homogenises strength rather than systematically increasing it. Overall COV decreases progressively from 0.168 at T2 to 0.089 at T8, consistent with improved mixing uniformity. Skewness also varies systematically. At T2, the distribution is moderately right-skewed (skewness = 0.68), indicating some locally high-strength zones.

Skewness increases substantially at T4 and T6 (1.37 and 1.42, respectively), reflecting extreme high-strength values and, for T6, distributional distortion arising from the depth-wise strength gradient. At T8, skewness approaches zero (−0.15), consistent with near-uniform mixing.

3.2 Slice-wise strength distribution

Figure 5 shows the vertical distribution of q_{uNP} for the four mixing conditions; Table 3 presents, for each case, the depth-wise range of slice means and the within-slice coefficient of variation COV_{ss} , computed from the 20 measurement points within each horizontal test plane. Two distinct variability patterns emerge. At T2 and T4, COV_{ss} is high and varies substantially between slices. At T2, five of the 12 slices have COV_{ss} exceeding 0.15 (maximum 0.246), reflecting incomplete cement dispersion across the horizontal cross-section. At T4, two slices exceed $COV_{ss} = 0.15$ (maximum 0.234), with individual values reaching 245.8 kPa, the maximum recorded.

At T6, within-slice uniformity improves markedly: all 12 slices show COV_{ss} below 0.10 (range 0.043–0.094, mean 0.070). However, slice means vary substantially with depth, from 121.4 kPa at 15.5 cm to 184.9 kPa at 33.5 cm, a depth-wise range

Table 2. Overall statistical characteristics of q_{uNP} for the four mixing conditions

Case	<i>n</i>	Mean, q_{uNP} : kPa	SD: kPa	COV	Min: kPa	Max: kPa	Skewness
T2	240	143.5	24.0	0.168	92.9	220.7	0.68
T4	240	152.3	22.5	0.148	103.9	245.8	1.37
T6	240	138.8	19.9	0.144	111.8	226.5	1.42
T8	220	136.5	12.1	0.089	104.6	166.4	−0.15

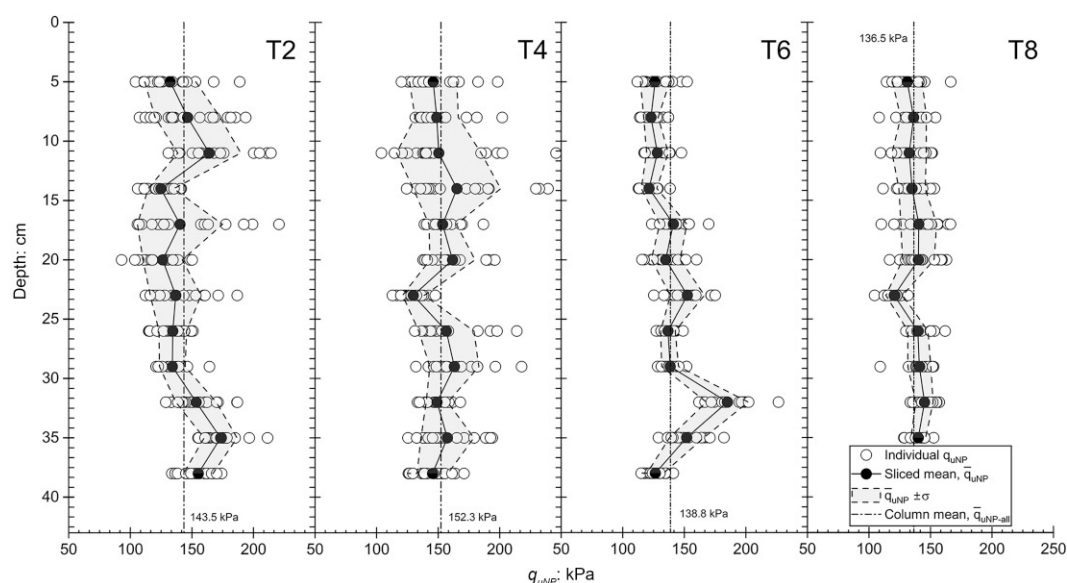


Figure 5. Vertical distribution of q_{uNP} for the four mixing conditions. Open symbols: individual q_{uNP} values; filled circles: slice mean; grey band: $\pm\sigma$; vertical dashed line: overall column mean

Table 3. Slice-wise statistics for the four mixing conditions

Case	Slice-mean range: kPa	Depth-wise range: kPa	COV _{ss} range	Mean COV _{ss}
T2	124.8–173.3	48.5	0.077–0.246	0.127
T4	129.8–165.3	35.5	0.059–0.234	0.125
T6	121.4–184.9	63.5	0.043–0.094	0.070
T8	120.5–145.1	24.6	0.042–0.105	0.074

of 63.5 kPa, the largest across all four cases. This pattern of low within-slice scatter combined with large depth-wise variation in slice means is the defining characteristic of the T6 case and is examined in detail in Section 4. At T8, both aspects of variability are minimised simultaneously. All slices have COV_{ss} below 0.11 (range 0.042–0.105, mean 0.074) and the slice-mean range narrows to only 24.6 kPa, confirming near-uniform cement distribution in both the horizontal and vertical directions.

Overall, COV decreases monotonically with T but cannot distinguish a column dominated by local horizontal heterogeneity (T2, T4) from one whose cross-sections are individually uniform but differ systematically with depth (T6). Section 4 decomposes total variance into depth-structured and local horizontal components to quantify this distinction.

4. Variance decomposition

4.1 Two-component partitioning of total q_{uNP} variance

Total q_{uNP} variance within each column was decomposed into a depth-structured component, reflecting systematic differences in mean strength between depth levels, and a local horizontal component, reflecting point-to-point scatter within each horizontal test plane. Denoting total population variance as σ_{total}^2 , the decomposition is:

$$3. \quad \sigma_{total}^2 = \sigma_z^2 + \sigma_h^2$$

where σ_z^2 is the depth-structured (between-slice) variance and σ_h^2 is the local horizontal (within-slice) variance, defined, respectively, as:

$$4. \quad \sigma_z^2 = \frac{1}{N} \sum_{s=1}^S n_s (\bar{q}_{uNP,s} - \bar{q}_{uNP})^2$$

$$5. \quad \sigma_h^2 = \frac{1}{N} \sum_{s=1}^S \sum_{i=1}^{n_s} (q_{uNP,s,i} - \bar{q}_{uNP,s})^2$$

Here, S is the number of slices, n_s is the number of measurement points in slice s , N is the total number of measurement points, $q_{uNP,s,i}$ is the converted strength at point i in slice s , $\bar{q}_{uNP,s}$ is the mean strength of slice s , and $\bar{q}_{uNP}$ is the overall column mean. The variance ratios are defined as:

$$6. \quad \eta_z = \frac{\sigma_z^2}{\sigma_{total}^2}; \eta_h = \frac{\sigma_h^2}{\sigma_{total}^2}; \eta_z + \eta_h = 1$$

where η_z quantifies the proportion of total q_{uNP} variance attributable to depth-level mean differences, whereas η_h quantifies the proportion attributable to local horizontal heterogeneity within each test plane. Both ratios are expressed as percentages to compare cases with different absolute variances.

Since each mixing condition is represented by a single physical column, the point estimates of η_z and η_h describe the individual model columns rather than the population of all DMM columns produced under each T setting. Hierarchical bootstrap 95% confidence intervals are reported alongside the decomposition results in Table 4, computed by resampling slices with replacement and, within each resampled slice, resampling measurement points with replacement (2000 replications). The intervals are wide owing to the small number of slices per column, but the rank ordering of cases by η_z is preserved across the full bootstrap distribution.

Figure 6 shows the variance ratios as a stacked bar chart; Figures 7(a) and 7(b) show depth profiles of slice-mean q_{uNP} and within-slice COV_{ss}, respectively. Three distinct regimes are identified. At T2, local horizontal heterogeneity dominates ($\eta_h = 63.2\%$),

Table 4. Variance decomposition of q_{uNP} into depth-structured and local horizontal components for the four mixing conditions, including hierarchical-bootstrap 95% confidence intervals

Case	σ_{total}^2 : kPa ²	σ_z^2 : kPa ²	σ_h^2 : kPa ²	η_z : %	95% CI	η_h : %	95% CI
T2	577.2	212.3	364.9	36.8	[16,59]	63.2	[41,84]
T4	507.9	86.7	421.2	17.1	[7,38]	82.9	[62,94]
T6	397.5	291.0	106.5	73.2	[40,84]	26.8	[16,60]
T8	145.8	40.5	105.3	27.8	[8,51]	72.2	[49,92]

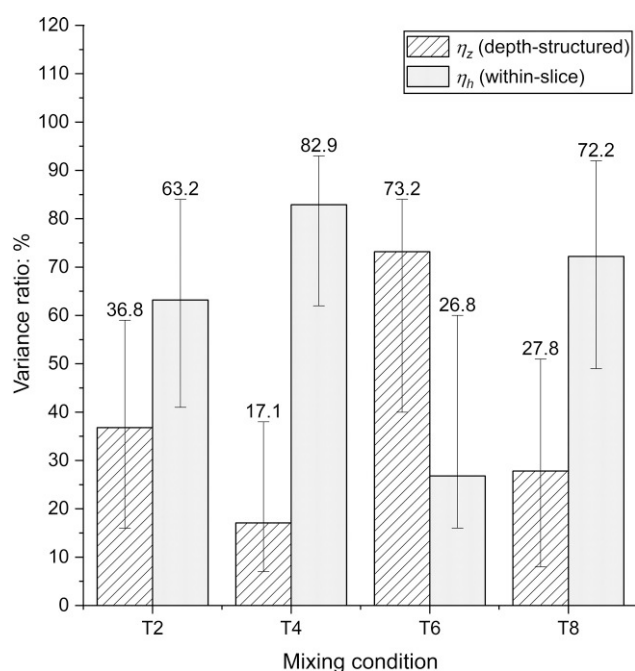


Figure 6. Variance decomposition of q_{UNP} into depth-structured (η_z) and local horizontal (η_h) components for the four mixing conditions. Error bars indicate hierarchical-bootstrap 95% confidence intervals

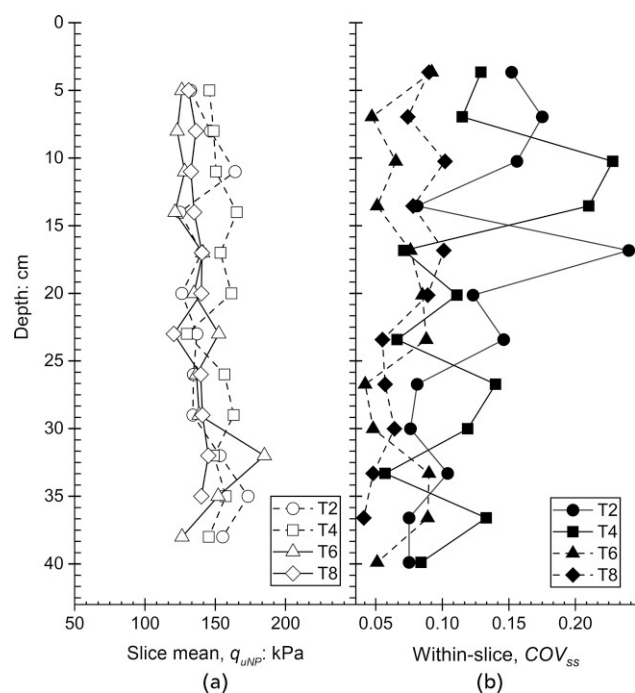


Figure 7. Depth profiles of (a) slice-mean q_{UNP} and (b) within-slice q_{UNP} for the four mixing conditions

with COV_{ss} spanning 0.077–0.246, reflecting insufficient mixing to homogenise cement distribution within each cross-section. At T4, local dominance intensifies ($\eta_h = 82.9\%$), making T4 the most

locally heterogeneous case despite its moderate overall COV of 0.148.

The most striking result occurs at T6, where η_z reaches 73.2%, the highest depth-structured ratio across all cases, while COV_{ss} remains consistently tight (0.043–0.094). Slice means nevertheless vary substantially with depth (121.4–184.9 kPa, range 63.5 kPa), demonstrating that within-slice homogeneity does not guarantee depth-wise uniformity. The mechanism responsible for this depth-wise gradient at T6 is identified as a subject for further investigation.

At T8, total variance falls to 145.8 kPa², less than one quarter of the T2 value, with both components substantially reduced and the slice-mean range narrowing to 24.6 kPa. The non-monotonic behaviour of η_z (36.8% → 17.1% → 73.2% → 27.8%) indicates that within-slice homogeneity is largely achieved near T6, while depth-wise uniformity is not fully established until T8, broadly consistent with the rational mixing number concept of Ota *et al.* (2000), but adds a spatial interpretation to the mixing threshold. Reporting η_z and η_h alongside overall COV therefore provides a more complete description of column quality.

4.2 Inter-column replication

Two additional columns were prepared under the T4 and T6 conditions to assess repeatability; the replicates are denoted T4-R and T6-R. Numerical comparisons are summarised in Table 5, and depth profiles and variance-decomposition results are shown in Figure 8. Mean q_{UNP} is reproduced closely for T4 (152.3 vs. 152.7 kPa) but less so for T6 (138.8 vs. 152.4 kPa). Overall COV differs by only 0.008 between T4 replicates and 0.024 between T6 replicates, indicating good reproducibility of global variability magnitude.

In contrast, the variance partitioning shows greater between-column variability. For T4, η_z increases from 17.1% in T4 to 38.1% in T4-R; for T6, η_z decreases from 73.2% in T6 to 62.3% in T6-R. Despite these differences, the rank ordering is preserved: both T6 columns exhibit higher η_z than both T4 columns, indicating that the qualitative distinction between locally heterogeneous behaviour at T4 and more depth-structured behaviour at T6 is reproducible.

The hierarchical-bootstrap 95% confidence intervals in Table 5 partly overlap, so the replicate comparison does not constitute a formal statistical separation. Nevertheless, the consistent rank ordering supports the interpretation that mixing intensity governs the spatial character of variability, not only its magnitude. The T4

Table 5. Repeatability comparison for T4 and T6, including hierarchical-bootstrap 95% confidence intervals

Quantity	T4	T4-R	T6	T6-R
Mean q_{UNP} : kPa	152.3	152.7	138.8	152.4
Overall COV	0.148	0.156	0.144	0.119
η_z (%) [95% CI]	17 [7, 38]	38 [24, 51]	73 [40, 84]	62 [37, 78]
η_h (%) [95% CI]	83 [62, 94]	62 [49, 76]	27 [16, 60]	38 [22, 63]

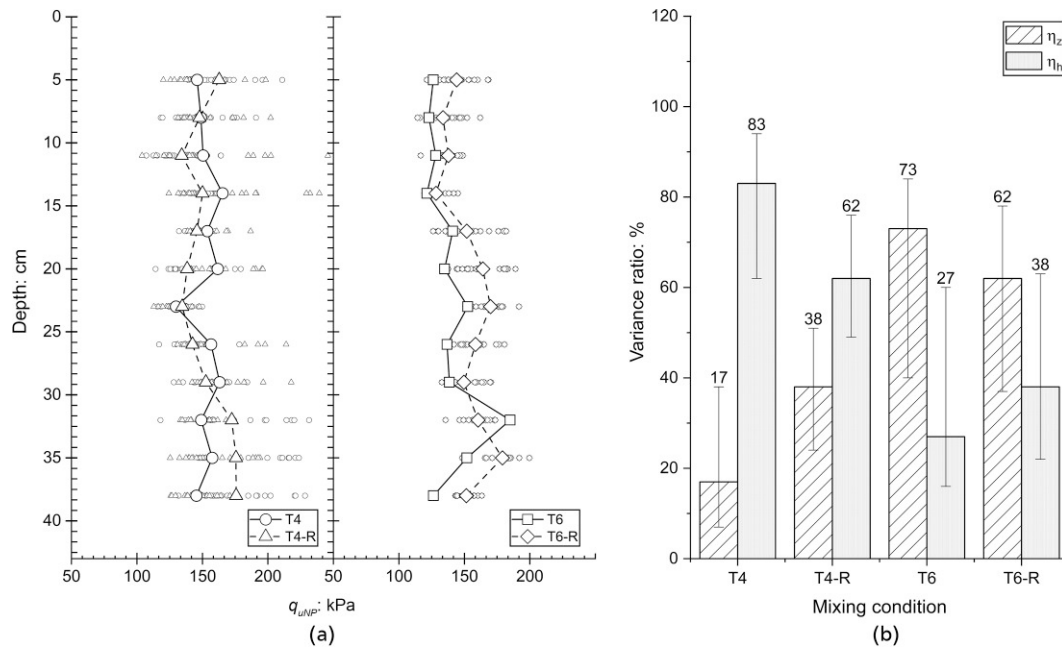


Figure 8. Repeatability comparison for the original and replicate T4 and T6 columns: (a) depth profiles of q_{uNP} and (b) variance partitioning into η_z and η_h with hierarchical-bootstrap 95% confidence intervals

vs. T4-R contrast suggests that variance partitioning is sensitive to factors not fully controlled by the present apparatus, potentially including local fluctuations in slurry-delivery pressure. This does not invalidate the qualitative findings but indicates that single-column estimates of η_z and η_h may carry inter-column uncertainty of approximately 10–20 percentage points. Replicate columns are therefore recommended in future studies.

5. Probability distribution analysis

5.1 Analytical approach

The variance decomposition in Section 4 showed that mixing intensity changes not only the magnitude of q_{uNP} variability but also its spatial organisation. This section examines whether these changes are also reflected in the probability distribution of q_{uNP} . This is important because DMM design and quality-control practice recognise substantial scatter in in situ treated-soil strength and require correction factors to account for this variability (CDIT, 2002), while probabilistic assessments of representative/design strength depend on the adopted statistical model and lower-tail criteria (Liu *et al.*, 2017). In addition, uncertainty in the estimated strength parameters depends on sample size and spatial correlation (Namikawa, 2019).

Normal and Lognormal distributions were fitted by maximum likelihood. These distributions were selected because both have been examined in previous statistical studies of cement-treated soil strength, although neither should be assumed universally applicable without goodness-of-fit checking (Liu *et al.*, 2017; Namikawa and Koseki, 2013). The normal distribution was tested as a symmetric benchmark that may become appropriate under

relatively uniform mixing conditions, whereas the Lognormal distribution was tested as a positive-valued, right-skewed alternative for heterogeneous strength data. Goodness of fit was assessed using the Lilliefors-corrected Kolmogorov–Smirnov test (Lilliefors, 1967) and the Stephens-modified Anderson–Darling test (Stephens, 1974, 1976) at $\alpha = 0.05$. A distribution was treated as acceptable only when it passed both corrected tests simultaneously.

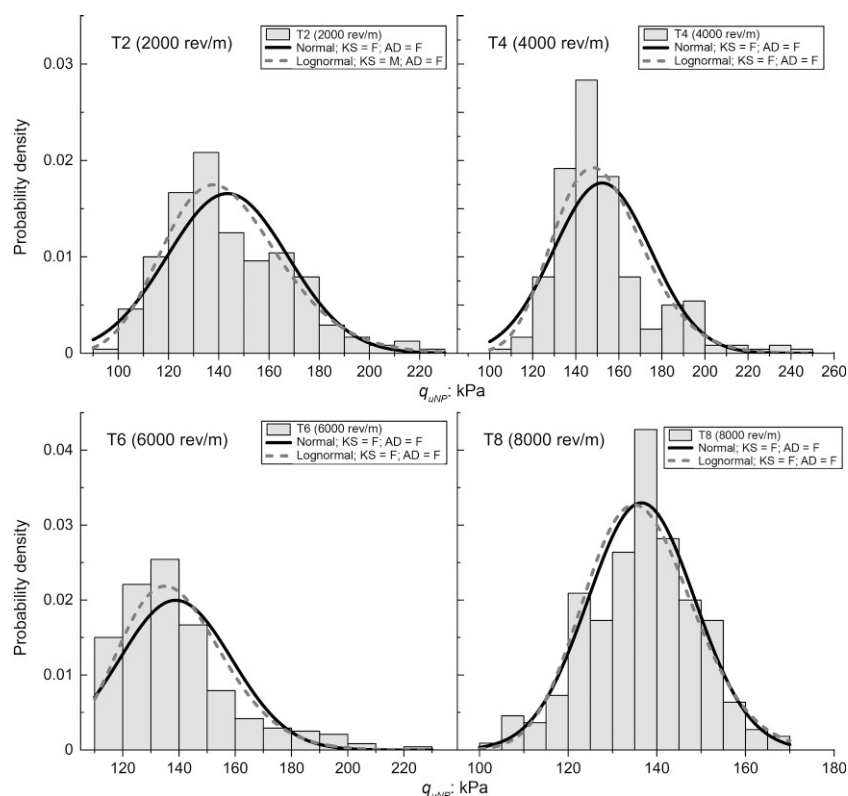
Distributions were also fitted to slice-mean-detrended residuals, $w_{s,i} = q_{uNP,s,i} - \bar{q}_{uNP,s}$, to determine whether apparent non-normality originates from the depth-trend structure or from the local residual field. Because these residuals are spatially nested within slices rather than strictly independent, the detrended p -values should be interpreted as descriptive screening metrics, not exact significance tests.

5.2 Goodness-of-fit results and distribution shape

The test results are summarised in Table 6. Histograms with fitted probability density functions are shown in Figure 9. At T2, the normal distribution is rejected by both tests $KS\ p = 0.0010$; $AD\ A^* = 2.211$, $p < 0.01$. The Lognormal distribution gives a marginal fit, passing the KS test ($p = 0.059$) and falling near the AD acceptance threshold ($A^* = 0.768$, $p = 0.05$ – 0.10). However, this should not be interpreted as strong evidence of Lognormal behaviour because the distribution remains right-skewed and tail-sensitive. This is consistent with $\eta_h = 63.2\%$, where local within-slice heterogeneity produces isolated high-strength zones that generate right-skewed behaviour (skewness = 0.68). Slice-mean detrending does not restore

Table 6. Goodness-of-fit test results for normal and lognormal fits to raw and slice-mean-detrended q_{UNP} data

Case	Distribution	Lilliefors $KS\ p$	Stephens $AD\ A^{*2}$	$AD\ p$ bracket	Result
T2 raw	Normal	0.0010	2.211	<0.01	Fail
T2 raw	Lognormal	0.0590	0.768	0.05–0.10	Marginal
T4 raw	Normal	0.0010	7.686	<0.01	Fail
T4 raw	Lognormal	0.0010	4.476	<0.01	Fail
T6 raw	Normal	0.0010	7.148	<0.01	Fail
T6 raw	Lognormal	0.0010	4.112	<0.01	Fail
T8 raw	Normal	0.0681	0.527	>0.15	Pass
T8 raw	Lognormal	0.0034	1.064	0.01–0.025	Fail
T2 detrended	Normal	0.0075	2.020	<0.01	Fail
T4 detrended	Normal	0.0010	6.629	<0.01	Fail
T6 detrended	Normal	0.0589	1.048	0.01–0.025	Fail/KS marginal pass
T8 detrended	Normal	0.1330	0.873	0.025–0.05	Fail/KS pass

**Figure 9.** Probability density histograms of raw q_{UNP} for the four mixing conditions, with fitted normal and lognormal probability density functions

normality ($KS\ p = 0.0075$; $AD\ p < 0.01$), suggesting that the local residual field remains non-Gaussian at T2.

At T4, both distributions are rejected for the raw data. Strong right skew (skewness = 1.37) and extreme values reaching 245.8 kPa produce a heavy-tailed distribution that neither model adequately represents. The high $\eta_h = 82.9\%$ suggests that isolated high-strength zones remain within individual cross-sections, generating outliers that distort the distribution. Slice-mean detrending also fails to produce normal residuals ($KS\ p = 0.0010$; $AD\ p < 0.01$), indicating that the local residual field remains strongly non-Gaussian.

At T6, both distributions are rejected for the raw data. The histogram does not show a simple unimodal distribution and instead exhibits a pronounced right tail associated with large differences in slice-mean strength. This behaviour is consistent with pooling slices whose means span 121.4–184.9 kPa and with the high depth-structured variance ratio ($\eta_z = 73.2\%$) identified in Section 4. After slice-mean detrending, the departure from a simple normal form is reduced and the KS test no longer rejects normality. This indicates that the poor fit of the raw T6 distribution is mainly associated with depth-wise variation in slice means rather than the local residual field alone. The AD test, however,

continues to reject the detrended residuals, indicating that residual tail mismatch remains.

At T8, the raw dataset is the only case for which the normal distribution is not rejected by either test at $\alpha = 0.05$. The near-zero skewness and low COV are consistent with near-uniform mixing. The T8 dataset is slightly smaller ($n = 220$), which marginally reduces test power; therefore, the acceptance of normality should be treated as suggestive rather than definitive. For the slice-mean-detrended T8 residuals, the KS test does not reject normality ($p = 0.1330$), but the AD test remains marginally sensitive to tail mismatch ($A^* = 0.873$, $p = 0.025-0.05$); therefore, the detrended T8 residuals are treated as a KS pass but not a full normal fit.

5.3 Implications for distribution characterisation

The lognormal spread parameter decreases monotonically with increasing T , from $\sigma_{ln} = 0.164$ at T2 to 0.145 at T4, 0.142 at T6, and 0.090 at T8, confirming that distributional spread narrows with increasing T . However, because the lognormal distribution is rejected or only marginally supported for T2–T6, this parameter should be interpreted as a descriptive measure of log-scale spread rather than a structural distributional parameter. Overall, the conventional lognormal assumption is not consistently supported in the present miniature-column dataset. Only T8 provides tentative support for a normal approximation. More importantly, distribution shape reflects the spatial organisation of strength: the distorted T6 distribution is mainly associated with depth-wise mean variation, whereas the near-normal T8 distribution reflects the simultaneous reduction of both spatial variance components. Overall COV alone is therefore insufficient to characterise DMM column quality.

6. Spatial correlation analysis

6.1 Empirical autocorrelation and maximum likelihood estimation of vertical correlation length

The spatial autocorrelation structure of q_{uNP} was examined in the horizontal and vertical directions using slice-mean-detrended residuals. This detrending was applied to remove the depth-structured component identified in Section 4, so that the ACF analysis represents the residual spatial structure of the local q_{uNP} field. This follows the geostatistical treatment in which spatial variability is separated into a deterministic trend and a residual component characterised by variance and autocorrelation structure (Baecher and Christian, 2003; Vanmarcke, 1977). For each measurement point at slice s and horizontal position i , the detrended residual is defined as:

$$7. \quad w_{s,i} = q_{uNP(s,i)} - \bar{q}_{uNP(s)}$$

where $\bar{q}_{uNP(s)}$ is the mean of the 20 measurement points within slice s . The empirical autocorrelation function (ACF) was computed as:

$$8. \quad \hat{\rho}(\Delta) = \frac{C(\Delta)}{C(0)}$$

where $C(\Delta)$ is the sample covariance between all pairs of residuals separated by lag Δ , and $C(0)$ is the sample variance. Horizontal pairs were formed from measurement points within the same slice, whereas vertical pairs were formed from corresponding horizontal positions across different depth levels and pooled across the 20 vertical profiles. For the vertical direction, the one-dimensional exponential covariance model was adopted following Vanmarcke (1977):

$$9. \quad C(h) = \sigma^2 \exp\left(-\frac{h}{\theta}\right)$$

or equivalently in correlation form, following Honjo (1982):

$$10. \quad \rho(h) = \exp\left(-\frac{h}{\theta}\right)$$

where h is the vertical lag distance, σ^2 is the process variance and θ is the exponential correlation-length parameter. The exponential form was selected because it provides a simple single-parameter representation of spatial correlation and has been widely used in geotechnical random-field modelling (Baecher and Christian, 2003; Vanmarcke, 1977), including cement-treated soil strength data (Honjo, 1982; Namikawa, 2019; Namikawa and Koseki, 2013).

The vertical correlation length θ_v was estimated by maximum likelihood, following the multivariate-normal covariance formulation of DeGroot and Baecher (1993) and Namikawa and Koseki (2013). The 20 vertical profiles at distinct horizontal positions were treated as approximately exchangeable realisations of the residual vertical random field. This is a simplifying assumption because all profiles were obtained from a single laboratory column. However, a one-way analysis of variance of mean q_{uNP} against radial distance from the grid centre showed no statistically significant radial trend for any mixing condition ($p > 0.05$), supporting this approximation. For each vertical profile i , the vector of n_s detrended residuals was modelled as a multivariate normal vector with zero mean and covariance $\sigma^2 \mathbf{R}(\theta)$, where $\mathbf{R}(\theta)$ is the $n_s \times n_s$ correlation matrix with elements:

$$11. \quad R_{jk}(\theta) = \exp\left(-\frac{|z_j - z_k|}{\theta}\right)$$

where z_j and z_k are the depth centre coordinates of slices j and k , respectively. The matrix $\mathbf{R}(\theta)$ was decomposed using the Cholesky factorisation $\mathbf{R} = \mathbf{L}\mathbf{L}^T$ for numerical stability, avoiding direct matrix inversion. Treating the 20 profiles as independent, the total profile log-likelihood obtained by analytically eliminating σ^2 is:

$$12. \quad l_p(\theta) = -\frac{M^* \cdot n_s}{2} (1 + \ln(2\pi\hat{\sigma}^2(\theta))) - \frac{M^*}{2} \ln|\mathbf{R}(\theta)|$$

where $M^* = 20$ is the number of vertical profiles and n_s is the number of slices in each profile. For any given θ , the profile estimate of the variance is:

$$13. \quad \hat{\sigma}^2(\theta) = \frac{1}{M^* \cdot n_s} \sum_{i=1}^{M^*} w_i^T \mathbf{R}(\theta)^{-1} w_i$$

The MLE of the vertical correlation length is then:

$$14. \quad \hat{\theta}_v = \arg \max_{\theta > 0} l_p(\theta)$$

obtained by evaluating $l_p(\theta)$ over a fine grid of candidate values. The 95% confidence interval for θ_v was computed using the profile likelihood ratio criterion. All values of θ_v satisfying

$$15. \quad l_p(\hat{\theta}_v) - l_p(\theta) \leq 1.921$$

are included in the interval, where $1.921 = \chi_{1,0.95}^2/2$ corresponds to half of the 5% upper-tail critical value of the chi-squared distribution with one degree of freedom.

6.2 Estimated correlation lengths

The empirical ACF profiles are shown in Figure 10, and the MLE results for the vertical correlation length are summarised in Table 7. After slice-mean detrending, the horizontal ACF values are effectively zero for all mixing conditions. The maximum correlation across all cases and horizontal lags is only 0.049, which is within the range of sampling noise. Thus, no meaningful horizontal spatial correlation is detectable at the present measurement scale. Because the horizontal grid spans only 8 cm and the

Table 7. Maximum likelihood estimates of vertical correlation length θ_v for the four mixing conditions

Case	θ_v : cm	95% CI: cm	Slice spacing: cm
T2	2.07	[1.39, 2.95]	3.0
T4	3.10	[2.32, 4.24]	3.0
T6	1.72	[0.98, 2.48]	3.0
T8	2.40	[1.64, 3.39]	3.0

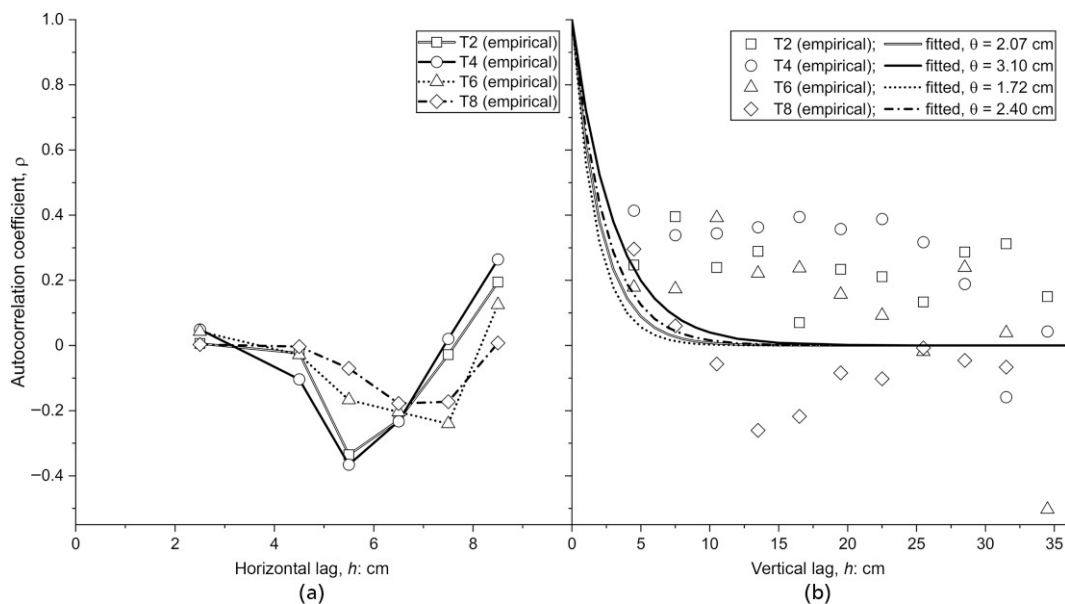


Figure 10. Empirical autocorrelation functions of slice-mean-detrended q_{UNP} residuals for the four mixing conditions: (a) horizontal direction and (b) vertical direction with fitted exponential models

inter-point spacing is 2 cm, this result should be interpreted as a resolution-limited finding, not as evidence that horizontal correlation is absent in general.

The detrended vertical ACF profiles show positive but irregular short-lag correlations. The fitted vertical correlation lengths range from 1.72 to 3.10 cm. T4 gives the longest value ($\theta_v = 3.10$ cm), indicating the greatest residual vertical continuity after detrending. This is consistent with its high within-slice heterogeneity, because unresolved local strength zones may extend slightly across the 3 cm slice boundaries. T6 gives the shortest value ($\theta_v = 1.72$ cm), consistent with its more homogenised within-slice structure and weak residual continuity beyond the nearest measurement depth.

All estimated θ_v values are comparable to, or smaller than, the 3.0 cm slice spacing. The MLE estimates are therefore weakly constrained, and the confidence intervals are correspondingly wide (Lloret-Cabot *et al.*, 2014). Because autocorrelation-distance estimates may be biased under limited-data conditions and are sensitive to the assumed correlation function (Cami *et al.*, 2020; Qi and Liu, 2019), the values in Table 7 should be regarded as approximate order-of-magnitude estimates. Finer vertical sampling would be required for more reliable θ_v estimation.

The estimated laboratory values are much smaller than reported field-scale vertical correlation parameters. Honjo (1982) reported vertical autocorrelation distances of approximately 0.4–4.0 m for cement-stabilised soil. Al-Naqshabandy *et al.* (2012) reported vertical scales of fluctuation of 0.2–0.7 m for lime-cement columns, corresponding to exponential correlation lengths of approximately 0.1–0.35 m. Pittaro and Mace (2024) reported vertical scales of fluctuation of approximately 0.5–3.5 m for DMM in the Kallang formation, corresponding to approximate exponential correlation lengths of 0.25–1.75 m under the same convention. These values are several times to two orders larger than the present estimates of $\theta_v = 1.72$ –3.10 cm. Direct transfer of θ_v from the miniature tests to field conditions is therefore inappropriate.

Overall, the dominant spatial structure is the depth-wise mean gradient identified by the variance decomposition, rather than a residual random field with substantial correlation length. Once the depth trend is removed, the residual q_{uNP} field approaches spatially uncorrelated behaviour at the scale of the present measurement grid.

Although these numerical values are specific to the present tests, the framework itself is not inherently tied to the specimen scale or base material, provided that spatially referenced strength data are available. In the field, the needle-penetration values used here could be replaced by strength data from cored samples or suitable in situ measurements such as cone penetration tests, which have been used to characterise strength variability in lime-cement columns (Al-Naqshabandy *et al.*, 2012). Measurements at multiple positions across the column diameter over several

depths would provide the within-depth variability used to estimate the local horizontal component σ_h^2 , whereas differences among the corresponding depth-level mean strengths would provide the depth-structured component, σ_z^2 . Applying the variance decomposition of Section 4 would therefore indicate whether field-column variability is predominantly depth-structured or locally horizontal, with the former warranting denser vertical sampling and the latter greater cross-sectional coverage. The numerical values and appropriate sampling intervals would nevertheless need to be established from field-scale data and site conditions.

7. Conclusions

The internal strength variability of laboratory deep-mixed columns was characterised across four blade revolution counts, T2–T8, using descriptive statistics, variance decomposition, distribution fitting and spatial autocorrelation analysis. The main findings are summarised below.

- Mean q_{uNP} did not vary monotonically with blade revolution count, whereas overall COV decreased from 0.168 at T2 to 0.089 at T8. Under the present laboratory conditions, increasing mixing intensity mainly reduced strength variability rather than increasing mean strength.
- Variance decomposition identified distinct spatial patterns. T2 and T4 were dominated by local horizontal heterogeneity ($\eta_h = 63\%$ –83%), whereas T6 was dominated by depth-structured variability ($\eta_z = 73\%$). At T8, total variance fell to less than one quarter of the T2 value and the slice-mean range narrowed substantially. Thus, within-slice homogeneity and depth-wise uniformity did not develop simultaneously with increasing mixing intensity.
- Normal and lognormal fits were not consistently supported. Raw q_{uNP} distributions at T4 and T6 were rejected by both models, T2 showed only marginal lognormal support, and the poor T6 fit mainly reflected depth-wise variation in slice means. Only T8 provided tentative support for a normal approximation.
- Detrended horizontal autocorrelation was effectively zero at the present grid resolution. Detrended vertical correlation lengths were 1.72–3.10 cm, comparable to the 3.0 cm slice spacing, indicating that the dominant spatial structure was the depth-wise mean variation rather than a strongly correlated residual field.

Overall, COV alone is an incomplete descriptor of DMM column quality. Columns with high η_z require denser vertical sampling, whereas those with high η_h require greater horizontal coverage. Reporting η_z and η_h alongside COV provides a more informative characterisation of strength variability. It may also support future comparison of alternative blade geometries in terms of depth-wise uniformity and local horizontal heterogeneity.

Acknowledgements

The authors gratefully acknowledge the GI Column Association for providing the miniature DMM apparatus and funding support for the experimental programme. The authors also thank Mr. Shota Sato for assisting with the laboratory experiments.

REFERENCES

- Al-Naqshabandy MS, Bergman N and Larsson S (2012) Strength variability in lime-cement columns based on cone penetration test data. *Proceedings of the Institution of Civil Engineers – Ground Improvement* **165**(1): 15–30. [10.1680/grim.2012.165.1.15](https://doi.org/10.1680/grim.2012.165.1.15).
- Baecher GB and Christian JT (2003) *Reliability and Statistics in Geotechnical Engineering*. John Wiley & Sons.
- Cami B, Javankhoshdel S, Phoon K-K and Ching J (2020) Scale of fluctuation for spatially varying soils: estimation methods and values. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **6**(4): 03120002. [10.1061/AJRUA6.0001083](https://doi.org/10.1061/AJRUA6.0001083).
- CDIT (Coastal Development Institute of Technology) (2002) The deep mixing method: principle, design and construction. A.A Balkema Publishers, Lisse.
- Chen EJ, Liu Y and Lee F-H (2016) A statistical model for the unconfined compressive strength of deep-mixed columns. *Géotechnique* **66**(5): 351–365. [10.1680/jgeot.14.P.162](https://doi.org/10.1680/jgeot.14.P.162).
- DeGroot DJ and Baecher GB (1993) Estimating autocovariance of in-situ soil properties. *Journal of Geotechnical Engineering* **119**(1): 147–166. [10.1061/\(ASCE\)0733-9410\(1993\)119:1\(147\)](https://doi.org/10.1061/(ASCE)0733-9410(1993)119:1(147)).
- Honjo Y (1982) A probabilistic approach to evaluate shear strength of heterogeneous stabilized ground by deep mixing method. *Soils and Foundations* **22**(1): 23–38. [10.3208/sandf1972.22.23](https://doi.org/10.3208/sandf1972.22.23).
- Japanese Cement Association (2021) *Manual for Ground Improvement Using Cement-Based Solidification Materials* 5th ed, Japan Cement Association, Tokyo.
- Larsson S, Stille H and Olsson L (2005) On horizontal variability in lime-cement columns in deep mixing. *Géotechnique* **55**(1): 33–44. [10.1680/geot.2005.55.1.33](https://doi.org/10.1680/geot.2005.55.1.33).
- Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association, Taylor & Francis* **62**(318): 399–402. [10.1080/01621459.1967.10482916](https://doi.org/10.1080/01621459.1967.10482916).
- Liu Y, Zheng J and Guo J (2008) Statistical evaluation for strength of pile by deep mixing method. In *Geotechnical Engineering for Disaster Mitigation and Rehabilitation* (Liu H, Deng A, and Chu J (eds)), Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 195–200. [10.1007/978-3-540-79846-0_15](https://doi.org/10.1007/978-3-540-79846-0_15).
- Liu Y, Jiang YJ, Xiao H and Lee FH (2017) Determination of representative strength of deep cement-mixed clay from core strength data. *Géotechnique* **67**(4): 350–364. [10.1680/jgeot.16.P.105](https://doi.org/10.1680/jgeot.16.P.105).
- Lloret-Cabot M, Fenton GA and Hicks MA (2014) On the estimation of scale of fluctuation in geostatistics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* **8**(2): 129–140. [10.1080/17499518.2013.871189](https://doi.org/10.1080/17499518.2013.871189).
- Namikawa T and Koseki J (2013) Effects of spatial correlation on the compression behavior of a cement-treated column. *Journal of Geotechnical and Geoenvironmental Engineering* **139**(8): 1346–1359. [10.1061/\(ASCE\)GT.1943-5606.0000850](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000850).
- Namikawa T (2019) Evaluation of statistical uncertainty of cement-treated soil strength using Bayesian approach. *Soils and Foundations* **59**(5): 1228–1240. [10.1016/j.sandf.2019.04.010](https://doi.org/10.1016/j.sandf.2019.04.010).
- Ota K, Watanabe K, Kawamura M, Fujii M and Hibino S (2000) Experiment for evaluating mixing performance in the deep mixing method, part 1: research objectives and experimental method. In *Proceedings of the 35th Japan National Conference on Geotechnical Engineering, Presented at the 35th Japan National Conference on Geotechnical Engineering*. pp. 1203–1204. (In Japanese).
- Phoon KK and Kulhawy FH (1999) Characterization of geotechnical variability. *Canadian Geotechnical Journal* **36**(4): 612–624. [10.1139/t99-038](https://doi.org/10.1139/t99-038).
- Pittaro GA and Mace N (2024) Spatial variability for deep cement mixing in Kallang formation in Singapore. Review of scale of fluctuation and its influence in design. In *Geotechnical Aspects of Underground Construction in Soft Ground* 1st ed. CRC Press, London, pp. 480–487. [10.1201/9781003413790-57](https://doi.org/10.1201/9781003413790-57).
- Qi XH and Liu HX (2019) Estimation of autocorrelation distances for in-situ geotechnical properties using limited data. *Structural Safety* **79**: 26–38. [10.1016/j.strusafe.2019.02.003](https://doi.org/10.1016/j.strusafe.2019.02.003).
- Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association, Taylor & Francis* **69**(347): 730–737. [10.1080/01621459.1974.10480196](https://doi.org/10.1080/01621459.1974.10480196).
- Stephens MA (1976) Asymptotic results for goodness-of-fit statistics with unknown parameters. *The Annals of Statistics* **4**(2): 357–369.
- Uchida N, Eto Y, Ono H and Miura N (2004) Strength evaluation of deep mixing soil-cement by needle penetration test. *Journal of the Japanese Society of Soil Mechanics and Foundation Engineering* **52**(7): 23–25. (In Japanese).
- Vanmarcke EH (1977) Probabilistic modeling of soil profiles. *Journal of the Geotechnical Engineering Division* **103**(11): 1227–1246.
- Wijerathna M and Liyanapathirana DS (2019) Significance of spatial variability of deep cement mixed columns on reliability of column-supported embankments. *International Journal of Geomechanics* **19**(8): 04019087. [10.1061/\(ASCE\)GM.1943-5622.0001473](https://doi.org/10.1061/(ASCE)GM.1943-5622.0001473).
- Zuo J, Wang B, Li W et al. (2023) Quality assessment and quality control of deep soil mixing columns based on a cement-content controlled method. *Scientific Reports* **13**(1): 4813. [10.1038/s41598-023-31931-y](https://doi.org/10.1038/s41598-023-31931-y).

How can you contribute?

To discuss this paper, please email up to 500 words to the editor at support@emerald.com. Your contribution will be forwarded to the author(s) for a reply and, if considered appropriate by the editorial board, it will be published as discussion in a future issue of the journal.

Proceedings journals rely entirely on contributions from the civil engineering profession (and allied disciplines). Information about how to submit your paper online is available at www.emeraldgroupublishing.com/journal/jgrim, where you will also find detailed author guidelines.