Sand Stiffness Variability Induced by Stochastic Distributions of Calcite Precipitates: A Monte Carlo-DEM Study

Sun, M.[1]; Liu, P.F.[2]; Chen, Y.X.[2]; Bate, B.[2]*; Xue, F.[3]

1 Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, SAR, China; formerly, Institute of Geotechnical Engineering, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China, 310058

2 Institute of Geotechnical Engineering, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China, 310058

3 Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, SAR, China

ABSTRACT

The inclusion of calcite precipitates ($CaCO_3$) in soft soil can improve the mechanical properties. Understanding the variability in sand stiffness due to heterogeneous precipitates is crucial for stiffness evaluation and prediction. A novel Discrete Element-Monte Carlo (DE-MC) method was proposed to quantify the sand stiffness variability induced by stochastic distributions of calcite precipitates, specifically focusing on shear wave velocity (V_s) as an indicator of soil stiffness. A total of 1972 samples were constructed to simulate stochastic spatial distributions of calcite precipitates. Through joint stochastic analysis, the preferential paths formed by calcite clusters were identified as significant contributors to V_s variability. The normalized connectivity per unity distance contact weight ($C_{d,n}$) exhibited the most correlated relation with V_s. Two weight selection methods were applicable for using $C_{d,n}$ to characterize and predict V_s. The results suggest that the DE-MC method has the potential to assess the variability in sand stiffness quantitatively.

Keywords

Monte Carlo simulation; discrete element method; variability; shear wave velocity; calcite precipitation

1 Introduction

Due to the demand for infrastructure and engineering construction, the need for various types of buildings and construction land has increased sharply. However, construction land resources are relatively limited, highlighting the significance of reinforcing and solidifying soft soil on the construction site [1, 2]. On the other hand, Sustainable Development Goal (SDG) 11, titled "Sustainable Cities and Communities", advocates for an environmentally-friendly approach to

addressing existing soft ground soil, aiming to reach the target of carbon neutrality [3, 4].

The inclusion of calcite precipitates induced by either microbes or enzymes in a porous medium represents an innovative eco-friendly method for cementing and strengthening soft soil. This approach typically results in increases in stiffness and strength, as well as reductions in hydraulic conductivity [5-8]. The shear stiffness property of cemented soil can vary significantly in different affiliation forms between soil particles and calcite precipitates [9], which can be calculated by

$$G_0 = \rho V_s^2, \tag{1}$$

where G_0 is the small strain shear modulus, ρ is the overall sample density, and V_s is the shear wave velocity. Uniform surface coating assumption of the calcite precipitates on coarse-grained soils prevails decades ago [10, 11]. Precipitations at particle contacts and pore filling are also considered alternative affiliation forms on coarse-grained soils [6, 12-15]. The effect of spatial variability on the evolution of V_s is further revealed through both column tests and microstructural experiments. Bate et al. [16] employed spectral-induced polarization (SIP) to quantify both the content and the size of calcite clusters. Scanning electron microscopy (SEM) and energy-dispersive X-ray spectroscopy (EDS) revealed the preferential precipitation of calcite at the particle contacting area [16-18], while the pore throat area and pore void were occupied by calcite, especially at high calcite content [6, 15]. In field applications, spatial variability prevails and still poses challenges for engineering applications [19-22]. Due to the natural spatial variability of calcite inclusion, quantification of the geotechnical properties has not been sufficiently investigated.

Recently, the discrete element method (DEM) has gained popularity in predicting the geotechnical properties of soil enhanced using microbially induced carbonate precipitation (MICP) [12, 23-26]. With the advancement in the DEM technique, spatial variability of calcite inclusion was investigated. For example, Wu et al. [12] proposed a novel DEM modeling scheme to reproduce various precipitation patterns of calcite and quantitatively investigated the microstructural characteristics and their evolution upon external loading. They classified the initial interparticle bonds into effective, partially effective, surface coating, and pore filling. Sun et al. [23] demonstrated that the V_s increased with the connectivity of the calcite aggregation cluster. However, DEM simulation of the spatial variability of calcite inclusion is very limited but highly desired.

The Monte Carlo method, known as random sampling or statistical testing method, unveils fundamental mechanisms via statistical analysis of the many instances or realizations of a stochastic process or state. By setting random numbers to simulate various physical processes, population properties can be inferred from samples. Zhou et al. [27] used the Monte Carlo method to address parameter uncertainty problems and produce a spatial slope failure probability distribution map. Saifuddin et al. [28] examined the applicability of the Markov-

chain Monte Carlo method in inversions of surface-wave phase velocity, obtaining a variation of S-wave velocity profiles from the inversion. Ma et al. [29] combined the Monte Carlo method with the fictitious domain method to improve the accuracy of the hydrodynamic force acting on the particle in a viscous fluid. In this paper, we used a combined DEM and Monte Carlo method to randomize the spatial distribution of calcite precipitation based on the authors' prior research [23].

The following tasks were conducted to quantify the sand stiffness variability induced by stochastic distributions of calcite precipitates: (1) A total of 1972 DEM models were established to account for the spatial variability of the calcite precipitates based on a proposed Discrete Element-Monte Carlo (DE-MC) method. (2) Interval prediction and confidence interval estimation were performed for the estimator V_s, and statistical analysis was carried out. (3) The correlation analysis between several contact-related features and V_s was conducted. (4) A quantified formula between $C_{d,n}$ and V_s was fitted based on the results of DE-MC simulation.

2 The proposed Discrete Element-Monte Carlo (DE-MC) method

2.1 Overview of the DE-MC method

Given the variability in the spatial distribution of calcite cementation within the sand matrix, variations in shear wave velocity (V_s) are to be expected but have not been quantified. In this case, the Monte-Carlo method could be applied in the quantitative prediction and interval estimation of V_s. In this paper, a Discrete Element-Monte Carlo (DE-MC) method was proposed, which consists of DEM simulation, Monte-Carlo process, and joint stochastic analysis (Fig. 1). A total of 116 groups of randomized calcite distribution were established, with 17 calcite content samples in each group.

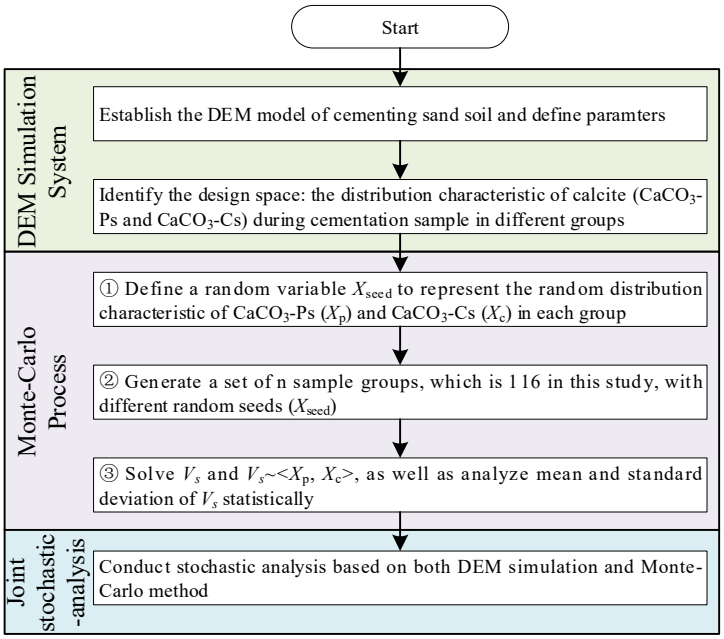


Fig. 1 Procedures of Monte-Carlo simulation in this study [30, 31]

2.2 DEM simulation

2.2.1 Virtual entity of generated calcite in the DEM model

In the EICP experiment, newly generated calcite cementing sand soil was transferred from a sample with a single uncemented contact model to a sand-calcite mixture [13, 16, 32]. The generated sand-calcite mixture sample, which contained two calcite precipitate forms: $CaCO_3$-P and $CaCO_3$-C [23], had the following characteristics: a. the calcite appeared as a small non-spherical entity; b. the formation space of calcite cementation was limited; c. the ratios of sand and calcite content were extremely uneven (Fig. 2a,b). Thus, an innovative DEM sample generation method, the virtual entity method, was adopted rather than traditional direct generation methods. In the DEM simulation, to be specific, Hertz-Mindlin model was used for sand-sand contacts due to the incompressible and unbonded nature of sand; linear parallel bond model (Linearpbond) was used to characterize the single calcite precipitate ($CaCO_3$-P, Fig. 2c) and calcite aggregation ($CaCO_3$-C, Fig. 2d), considering the model is appropriate for simulating finite-sized pieces of cement-like materials [33, 34]. The DEM model generated calcite by changing the original contact model between sand particles, and a toroidal model was used to calculate the mass and volume. Since sand particles were simulated as elastic spheres in the DEM, the toroidal model was applied with the cylindrical assumption of calcite precipitates at particle contacts (contact cementation association pattern) in prior research (Fig. 3, Eq. 2), which is elaborated by EDS image and theoretical calculation [13, 23, 35, 36]. The total volume of calcite precipitation at particle contacts was calculated using the toroidal assumption, as elaborated in a prior study [23] by

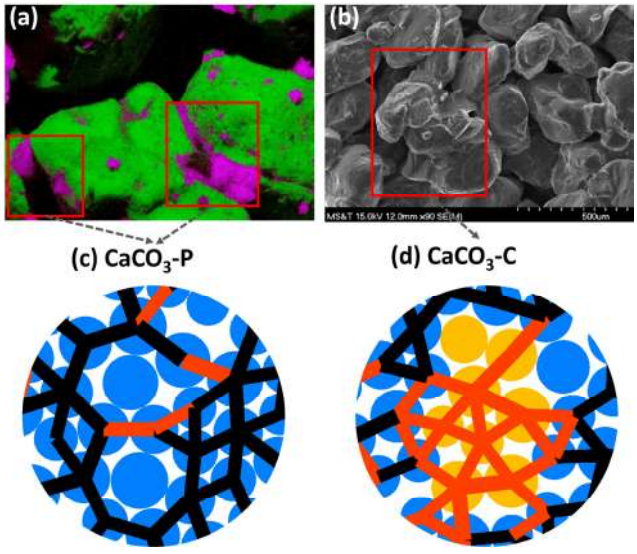


Fig. 2 (a) EDS image: Si-green color and Ca-purple color, (b)SEM image: example of sand-calcite clusters, (c)(d) precipitate forms in DEM model, blue denotes pure sand, orange denotes sand-calcite clusters, black denotes Hertz-Mindlin contact and red denotes Linearpbond contact.

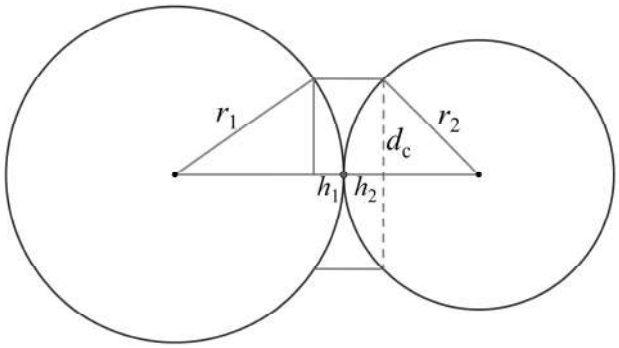


Fig. 3 Calculation of particle cementation volume.

$$h_i = \frac{d_i}{2}\left(1 - \cos\left(\arcsin\frac{d_c}{d_i}\right)\right),$$

$$v_{cntct} = d_c \sum h_i - \sum\left(\frac{d_i^2}{4}\arcsin\frac{d_c}{d_i} - \frac{d_c}{2}\left(\frac{d_i}{2} - h_i\right)\right), \quad (2)$$

where h_i = [h_1, h_2], which denotes the height of the toroidal area of each disk; d_i = [d_1, d_2], which denotes the diametral sizes of the two disks; d_c is the cementation diameter; v_{cntct} is total cementation volume. By changing the original contact model between sand particles, the virtual entity method is able to break through the limitation of basic physical properties of cementation such as mass, volume, and shape, so that it can be applied with high degrees of freedom in DEM simulation where only spherical particles can be generated. The validity of objective DEM modeling relies on the rationality of condition setting and parameter selection. SEM images from previous studies provide distribution and morphological characteristics of cementation [16, 23]; Bender element measurement provides data on V_s for calibrating microscopic parameters; SIP provides size information of cementation. These prerequisites enable cementation to exist as a virtual entity and yield nature-like experimental results.

2.2.2 Sample preparation

A 2D rectangular stripe (height × width = 0.3 m × 0.1 m) with rigid, frictionless walls was created in PFC^{2D}, within which particles (expressed as disks) were then generated by a five-step expansion method, reaching the two-dimensional porosity of 0.192 [23]. A vertical stress of 5 kPa under the plane strain condition was then applied using a rigid servo system. Following this, a high-damping boundary (α=1.0) was set in this model to simulate the continuous medium, enabling the dynamic effect to conform to the actual propagation law [37]. The parameters used in the model are listed in Table 1. For the measurement of V_s, a pair of virtual "walls" with a coefficient of friction of 0.7, called the transmitter and the receiver, were set in the samples at a distance of 0.1m (Fig. 4) [38]. A sinusoidal velocity pulse ($v_0 = 5(1.0 - \cos(2\pi ft))$ m/s) with the excitation frequency (f) close to the natural frequency of the sample was applied to the transmitter in the y-direction, aiming to maximize the received vibration signal and to minimize the errors caused by noises and P-waves [39, 40]. The received vibration

at the receiver end was recorded for V_s calculation.

Table 1 Parameters of the sample in the DEM model

Parameter	Sand	CaCO3	Sample
Diameter (mm)	2.12-3	-	-
Shear modulus (Pa)	6.5×10^{11}	-	-
Poisson ratio	0.25	-	-
Density (kg/m^3)	2650	2600	-
Friction coefficient	0.5	0.5	-
Tensile strength (Pa)	-	5.0×10^{9}	-
Cohesion	-	2.0×10^{10}	-
Effective modulus (Pa)	-	8.0×10^{10}	-
Viscous damping factor	0	-	-
Local damping factor	0.3	-	-
Relative density	-	-	0.85
Porosity Lab	-	-	0.37
Porosity 2D	-	-	0.192
Confining pressure (kPa)	-	-	5
Soil density (kg/m^3)	-	-	1659

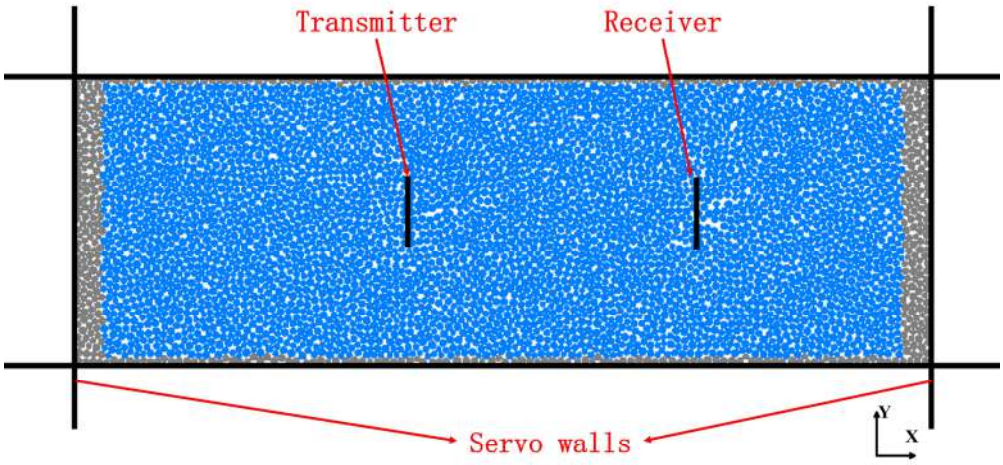


Fig. 4 Shear wave measurement in the DEM model

$CaCO_3$-P (represents individual calcite precipitation) was assigned to contacts when calcite content was below 1%, beyond which $CaCO_3$-C (represents calcite aggregation) was assigned. The quantity of $CaCO_3$-P precipitation at particle contacts increases with calcite content increment, while $CaCO_3$-C only forms when calcite content exceeds 1%. In this study, the formation of calcite precipitates was achieved by randomly adjusting the contact model at local contacts with the precipitates [23] (Fig. 5). Based on an experimental study, the simplified size of $CaCO_3$-P and $CaCO_3$-C in the DEM model are 1.06mm/1.33mm and 12mm respectively [23].

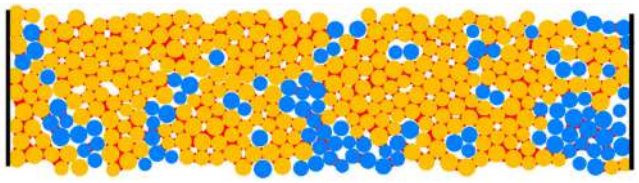


Fig. 5 Cemented sample (Blue denotes pure sand, orange denotes sand coated with $CaCO_3$-C, and red denotes $CaCO_3$ in the model figures)

2.3 Monte Carlo algorithm process

The detailed Monte-Carlo simulation process is as follows:

(1) Construct probability process. Unlike the single uncertain variable of most Monte-Carlo simulations, double variables in this study—the distribution characteristic of $CaCO_3$-Ps (X_p) and $CaCO_3$-Cs (X_c)—had a complex interrelation spatially and quantitatively under a certain calcite content [30, 41]. To address this challenge, a fundamental tool for Monte Carlo simulation, the random-seed variable (X_{seed}), was employed in the FISH Function to generate the X_p and X_c. The FISH Function is the only object that can be executed in the FISH (an embedded programming language within PFC) to manipulate DEM samples. In this way, the simple X_{seed} representing the overall random distribution of calcite cementation, can be executed instead of the X_p and X_c representing the random distribution characteristic of the two calcite forms in each sample group (Eq. 3). The methodology adopted in this study is outlined with reference to a prior study [23]. This redefinition simplified the generation of randomized 116 sample groups.

(2) Sample from a known probability distribution. Given that the calcite generation was evenly distributed within the region, X_p and X_c, represented by X_{seed}, also followed a uniform distribution. The sample matrix was set to represent the relationship among samples, variables, and estimators (Eq. 4). In Matrix 1, total $m*n$ samples were generated, where the number of columns (n) represents the number of groups of samples, and the number of rows (m) represents the sample numbering in each group. Matrix 2 indicated mathematically that each sample was determined by two variables -- X_{seed} for n and $CaCO_3$% for m. Through modeling each sample in Matrix 2, the Vs estimator matrix can be achieved as shown in Matrix 3. In this study, n=116 and 116 X_{seed} were allocated to 116 DEM sample groups to simulate the V_s.

(3) Solve and analyze estimators V_s. V_s was measured as the target estimator in the process, and the corresponding confidence interval was obtained. Statistically, the relationship between estimators V_s and uncertain variables (X_p, X_c), as well as the mean and standard deviation of V_s, ought to be analyzed.

Table 2 Calcite contents and excitation frequencies of 116 groups

m*	$CaCO_3$%	Frequency（kHz）	m*	$CaCO_3$%	Frequency（kHz）
1	0	4	10	2.55	8

2	0.11	4	11	3.17	8
3	0.23	5	12	4	10
4	0.34	7	13	5	15
5	0.45	8	14	6	20
6	0.55	8	15	7	25
7	0.66	8	16	8	25
8	1.24	8	17	9	25
9	1.88	8			

The m refers to the sample numbering.

$$S_{n\text{-}m}=F(X_{\text{seed}_n},\text{CaCO}_3\%)=F\left(X_\text{c},X_\text{p}\right)=\begin{matrix}p_\text{p}X_\text{p},\ \text{CaCO}_3\%\le 0.66\%\\ p_\text{c}X_\text{c},\ \text{CaCO}_3\%>0.66\%\end{matrix}, \tag{3}$$

$S_{n\text{-}m}$ refers to the sample with sample number m in the nth group; X_p and X_c are the distribution characteristic of $CaCO_3$-Ps and $CaCO_3$-Cs, respectively; p_p and p_c are the variation coefficients of $CaCO_3$-Ps and $CaCO_3$-Cs, respectively.

$$\begin{pmatrix} S_{1\text{-}1} & S_{2\text{-}1} & \cdots & S_{n\text{-}1}\\ \vdots & \vdots & \ddots & \vdots\\ S_{1\text{-}m} & S_{2\text{-}m} & \cdots & S_{n\text{-}m}\\ \vdots & \vdots & \ddots & \vdots\\ S_{1\text{-}17} & S_{2\text{-}17} & \cdots & S_{n\text{-}17}\end{pmatrix} = \begin{pmatrix} F(X_{\text{seed}_1},0) & F(X_{\text{seed}_2},0) & \cdots & F(X_{\text{seed}_n},0)\\ \vdots & \vdots & \ddots & \vdots\\ F(X_{\text{seed}_1},\text{CaCO}_3\%) & F(X_{\text{seed}_2},\text{CaCO}_3\%) & \cdots & F(X_{\text{seed}_n},\text{CaCO}_3\%)\\ \vdots & \vdots & \ddots & \vdots\\ F(X_{\text{seed}_1},9\%) & F(X_{\text{seed}_2},9\%) & \cdots & F(X_{\text{seed}_n},9\%)\end{pmatrix} \sim \begin{pmatrix} Vs_{1\text{-}1} & Vs_{2\text{-}1} & \cdots & Vs_{n\text{-}1}\\ \vdots & \vdots & \ddots & \vdots\\ Vs_{1\text{-}m} & Vs_{2\text{-}m} & \cdots & Vs_{n\text{-}m}\\ \vdots & \vdots & \ddots & \vdots\\ Vs_{1\text{-}17} & Vs_{2\text{-}17} & \cdots & Vs_{n\text{-}17}\end{pmatrix} \tag{4}$$

Matrix 1: sample number Matrix 2: model expression Matrix 3: the estimator

The particle packing of an uncemented sand matrix was selected for all 1972 samples to eliminate variations in V_s caused by contact arrangements. In addition, the post-processing and data calculation could be simplified by using the same initial particle packing in the preceding DEM simulation. Shear wave measurements were performed on these samples after reaching equilibrium.

2.4 Joint stochastic-analysis

In the joint stochastic analysis, the random distribution of macro-level V_s and the variation of micro-spatial cementation are described to estimate the evolution interval of V_s and interpret the relation between shear wave velocity and spatial variability.

A Jarque-Bera test was used to test the normal goodness of fit for the V_s distribution, which was based on the sample skewness and kurtosis [42]. The test statistics were defined as:

$$JB=\frac{S^2}{6/n}+\frac{(K\text{-}3)^2}{24/n}, \tag{5}$$

$$S=\frac{\widehat{\mu_3}}{\widehat{\sigma^3}}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i\text{-}\bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i\text{-}\bar{x})^2\right)^{3/2}}, \tag{6}$$

$$K=\frac{\widehat{\mu_4}}{\widehat{\sigma^4}}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{4/2}}, \tag{7}$$

where n is the number of samples, S is the sample skewness, and K is the sample kurtosis. The skewness and kurtosis of a standard normal distribution are 0 and 3, respectively. The hypothesis was that the V_s of different calcite content obeyed a normal distribution with unknown mean and variance, which were tested using the sample skewness and kurtosis. Moreover, the V_s results were considered not to follow a normal distribution if the random value of the 2-freedom chi-square distribution (P-value) for the test statistics was less than 5%.

Quantifying spatial distribution variability in V_s typically requires thousands of Monte Carlo samples, which is computationally intensive, time-consuming and lacks theoretical background. An indicator, the normalized connectivity per unity distance($C_{d,n}$), was introduced to represent the connectivity of the sample [23]. The DEM model can be seen as a weighted undirected graph by graph theory with weighting systems on Linearpbond contacts and Hertz contacts, in which vertexes and edges symbolize particles and contacts, respectively. The shortest path (also known as preferential propagation path in the DEM model) and weighted shortest distance (also known as shortest propagation distance in the DEM model) from the transmitter to the receiver can be calculated by Dijkstra's algorithm in reference to the prior study [23]. A preliminary indicator, C_d, was used to represent the weighted shortest propagation distance. However, C_d represents absolute distance, which cannot allow quantitative comparison of samples with different sizes. The $C_{d,n}$, normalizing the shortest propagation distance to represent the connectivity of all samples, was thereby introduced and defined as [23]:

$$\begin{cases}\omega_h = \dfrac{1}{C^{2/3}} \\ \omega_{LPB} = \dfrac{1}{C}\end{cases}, \tag{8}$$

$$\begin{cases} C_d = \omega_h N_h + \omega_{LPB} N_{LPB} \\ C_{d,n} = \dfrac{C_d - \omega_{LPB}\left(\frac{L}{d_{50}} - 1\right)}{\omega_h\left(\frac{L}{d_{50}} - 1\right) - \omega_{LPB}\left(\frac{L}{d_{50}} - 1\right)}\end{cases}, \tag{9}$$

where C is the coordination number in the sample; ω_h and ω_{LPB} are weights of Hertz contacts and Linearpbond contacts; N_h and N_{LPB} are the numbers of the Hertz contact model and the Linearpbond model, respectively; L is the distance between the transmitter and the receiver; C_d is the connectivity, represented as the weighted shortest path; $C_{d,n}$ is the normalized connectivity per unity distance, represented as the normalized weighted shortest path. A lower value of $C_{d,n}$ indicates a shorter preferential propagation path and greater connectivity.

3 Results

3.1 Sensitivity analysis

3.1.1 Peak-to-peak time determination

Shear wave velocity (V_s) was determined by dividing the propagation distance (0.1m) by the time. A comparison between the peak-to-peak and first-arrival time methods (Fig. 6) revealed that the V_s calculated by both two methods were steady at low frequency (<8kHz). However, the determination of shear wave propagation time using the first arrival method is subjective due to the influence of near-field effects (i.e., compression waves and reflected shear waves) [43]. Although both the two methods exhibited some dispersion at high frequency (>16kHz), it was particularly pronounced in the first arrival method. Therefore, the peak-to-peak method was selected in the subsequent analysis.

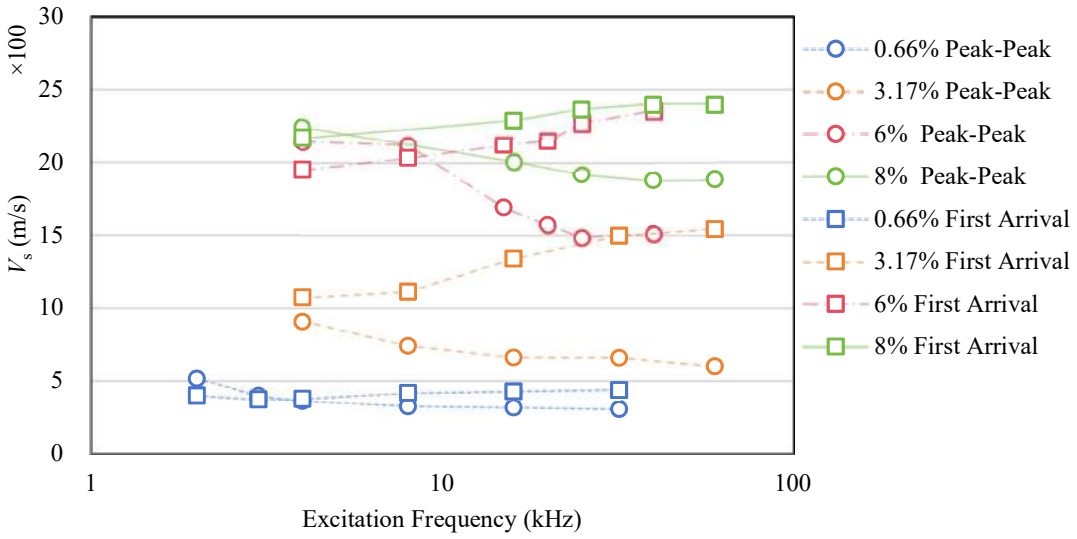


Fig. 6 V_s measured by two methods under four typical calcite content (0.66%, 3.17%, 6%, 8%)

3.1.2 Damping coefficients

In DEM simulation, damping contains two components, the viscous damping coefficient and local damping coefficient. The former is not applicable to dry, non-viscous materials, so it was set to 0 to avoid a reduction in output frequency. The range of the latter of [0.1, 0.7] was tested in DEM simulations (Fig. 7a). Using experimentally measured shear wave velocity ($V_{s,exp}$) from a prior study [16], an optimal damping coefficient was determined as follows. Compared with the experimentally measured arrival time (t_{ar}, calculated by $V_{s,exp}$), local damping coefficients exceeding 0.3 delayed wave arrival (Fig. 7a). However, high-frequency noise appeared at α<0.2 (Fig. 7b). In view of the above results, the local damping factor (α) was set to 0.3, and the damping ratio was $D = \alpha/\pi$=0.095, which was within the small-strain damping ratios of cemented sands of [0.02, 0.15] (Fig. 8) [44, 45].

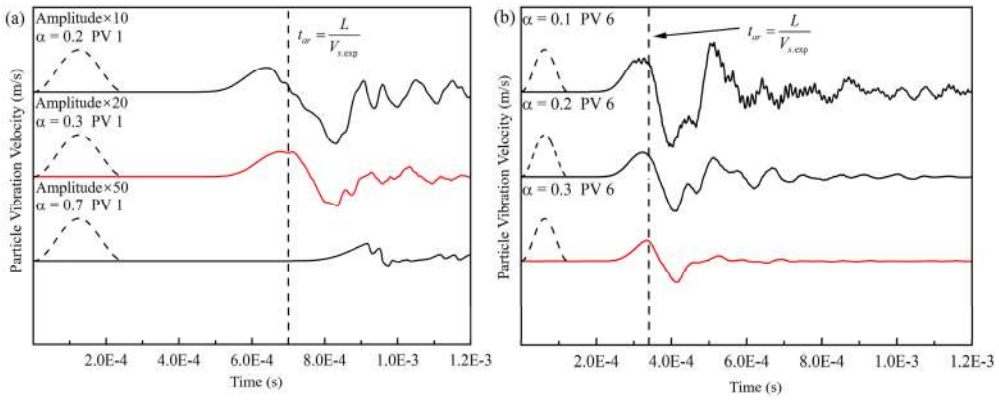


Fig. 7 Signals of the receiver under $CaCO_3$% = 0.11% (a) and 0.66% (b) with local damping factors (α) of [0.1, 0.7]; t_{ar}, L, $V_{s,\,exp}$ are the wave propagation time, propagation distance and V_s from experiments [16].

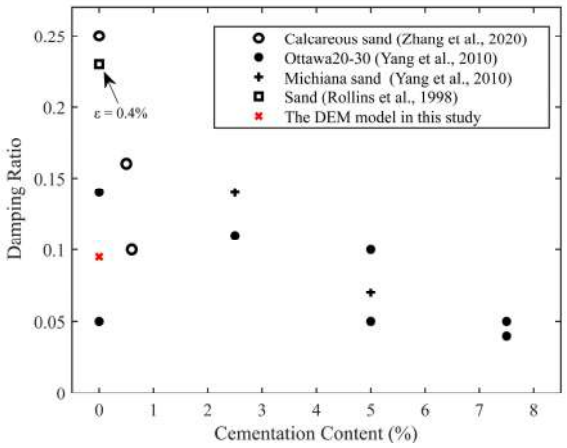


Fig. 8 Damping ratio and cementation content relationships [44-46]

3.1.3 Excitation frequency

It was observed that V_s was greatly affected by different excitation frequencies in both experiments and simulations [47-49]. Sets of samples with four calcite contents were tested for sensitivity analysis of excitation frequency (Fig. 9). For samples of the same calcite content, varying excitation frequencies could cause differences in the calculated V_s, with a maximum variation of 358m/s at 8% $CaCO_3$. This variation, attributed to wave dispersion, was particularly obvious when the cementation content was high, significantly impacting both experimental and simulation results. Therefore, characterizing the relationship properly between input and output frequencies is essential to accurately determine V_s [48, 50]. The input frequency should fall within the range of its linear relationship with the output frequency, and thus this threshold can be considered a reasonable input frequency [48]. For the same sample of four different calcite contents, it was found that the input and output frequencies could be well matched when the excitation frequencies of the samples were selected as 8kHz, 8kHz, 20kHz, and 25kHz, and their frequencies fell within the reasonable threshold (Fig. 9), which meant that the cementation content is positively correlated with the inherent frequency of the sample. Excitation frequencies near the inherent frequencies of each calcite content group were used (Table 2). The propagation distance (L) and excitation frequency used in the shear wave

simulation ought to be both satisfied by $\lambda/d_{50} < 10$ and $L/\lambda > 2$, where λ is the wavelength and d_{50} is the median particle size [38, 48]

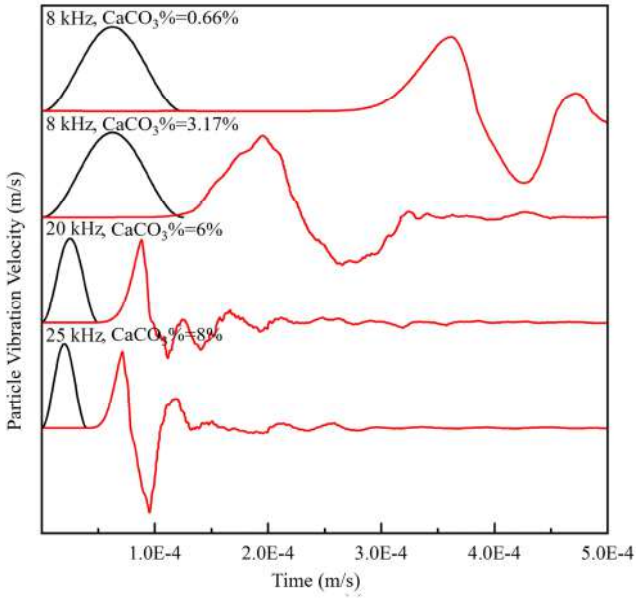


Fig. 9 Excitation frequency characterization of samples with different calcium carbonate content

3.2 Evidence from DE-MC

3.2.1 Macro-level results of DE-MC: Typical V_s vs $CaCO_3$% relationship

The stochastic analysis results of V_s in the DE-MC are shown in Table 3, Fig. 10, and Fig. 11. The growth of shear wave velocity can be divided into four stages (Sun et al., 2022): (1) In Stage I ($CaCO_3 \leq 0.66\%$, $CaCO_3$-P only), the mean V_s increased monotonically with the calcite content at a rate of 308.98m/s, as evidenced by the slope. The V_s frequency histograms displayed a normal distribution with a variance of [9.76, 35.49]. Samples in Stage I demonstrated positive skewness, implying a higher dispersion above the mean V_s and aggregation in the low-speed region. (2) In Stage II ($0.66\% \leq CaCO_3 \leq 2.55\%$, $CaCO_3$-C appeared), the V_s increased slower than Stage I, at a rate of 164.62m/s. The variance was [62.38, 171.20]. The kurtosis exceeded twice that of the standard normal distribution, indicating a deviation from normality. It suggested that, despite the generation of cemented clusters, the spatial variability was slight, and V_s remained relatively unaffected. Furthermore, the histograms had a positive (left) skew distribution, suggesting that a small number of samples exhibited high-speed regions, likely due to the formation of calcite clusters that facilitated wave propagation. (3) In Stage III ($2.55\% \leq CaCO_3 \leq 6\%$), the increasing rate of V_s was 262.18m/s, which was the highest of the four stages. The V_s frequency histograms displayed a normal distribution with a variance of [220.15, 260.74]. (4) In Stage IV ($6\% \leq CaCO_3 \leq 9\%$), the increasing rate was reduced to 189.81m/s. The histograms displayed a normal distribution, and the variance was [137.91, 203.39]. Samples in Stage IV demonstrated negative skewness, implying a higher dispersion below the mean V_s and aggregation in the high-speed region. In summary, the actual V_s for calcium carbonate reinforcement tended to be lower than the

325 predicted velocity in Stages I and II, and slightly higher than that in Stage IV.

Table 3 V_s distribution of samples with different calcite content

Stage	$CaCO_3$%	$\overline{V_s}$ (m/s)	Standard deviation(m/s)	Skewness	Kurtosis	Normal distribution
I	0.11	210.40	9.76	-0.13	2.77	○
	0.23	227.78	14.22	0.16	4.15	○
	0.34	248.97	18.29	0.67	3.93	○
	0.45	282.13	22.98	0.91	6.49	○
	0.55	322.39	24.89	0.68	3.75	○
	0.66	380.34	35.49	0.38	3.24	○
II	1.24	447.77	62.38	1.77	11.97	×
	1.88	525.58	92.21	1.48	6.36	×
	2.55	640.11	124.69	1.20	5.68	×
	3.17	765.48	171.20	1.44	6.75	×
III	4.00	958.21	220.15	0.64	3.09	○
	5.00	1211.31	260.74	0.36	2.60	○
	6.00	1482.56	238.12	-0.15	2.49	○
IV	7.00	1756.64	203.39	-0.19	3.11	○
	8.00	1974.69	154.59	-0.34	2.89	○
	9.00	2136.25	137.91	-0.70	3.47	○

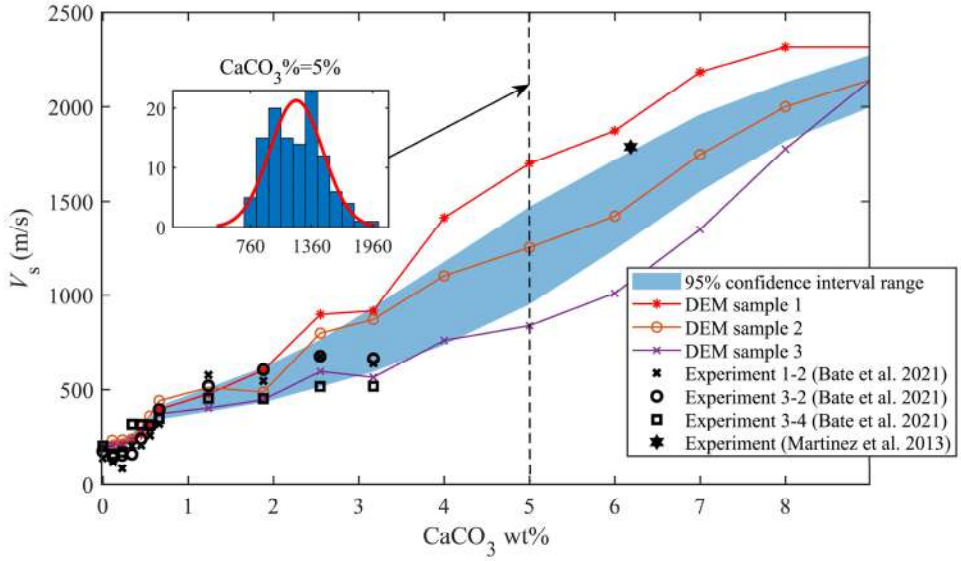


Fig. 10 Variation range of V_s [16, 51]

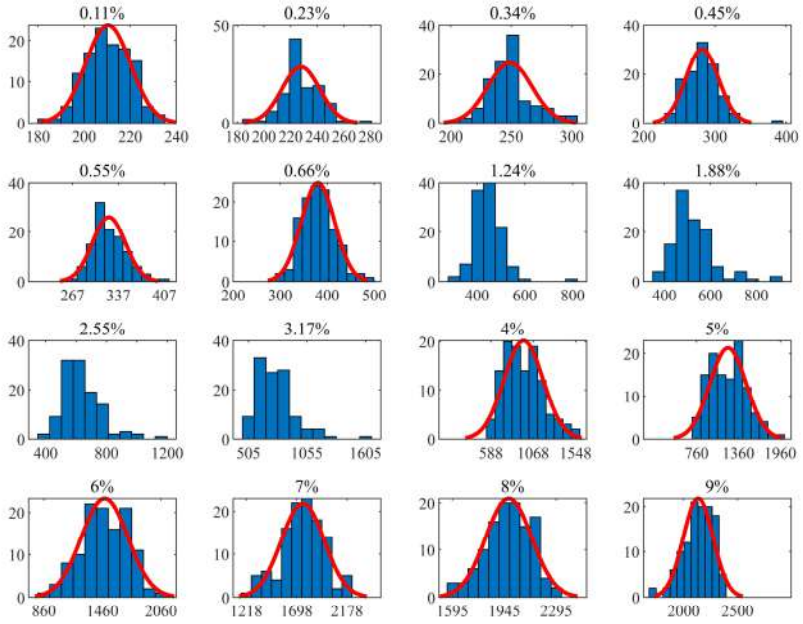


Fig. 11 V_s frequency histogram with different calcium carbonate content. The red line is the normal distribution fitting curve

3.2.2 Micro-scope spatial distribution of calcite-sand clusters

A close look at the shear wave propagation path could reveal influencing factors at the particle contact level. Three samples at 5% calcite content showed three preferential paths through which the shear wave arrived at the receiving end first (Fig. 10, Fig. 12). The optimal path of Sample 1 passed through the most calcite clusters and had the longest consecutive Linearpbond contact length, with the measured V_s of 1700m/s; although the optimal path of Sample 2 passed through most calcite clusters, there was a node remained uncemented, resulting in a continuous Linearpbond contact length shorter than that of Sample 1, with the measured V_s of 1255m/s; the optimal path of Sample 3 passes through the least number of calcite clusters, and there were four nodes remained uncemented, resulting in the shortest continuous Linearpbond contact length, with the measured V_s of 838m/s. Through the comparison of the above three paths, it can be seen that the larger the proportion of calcite clusters (Linearpbond contact) and the continuous Linearpbond contact length, the faster the transmission of the shear wave.

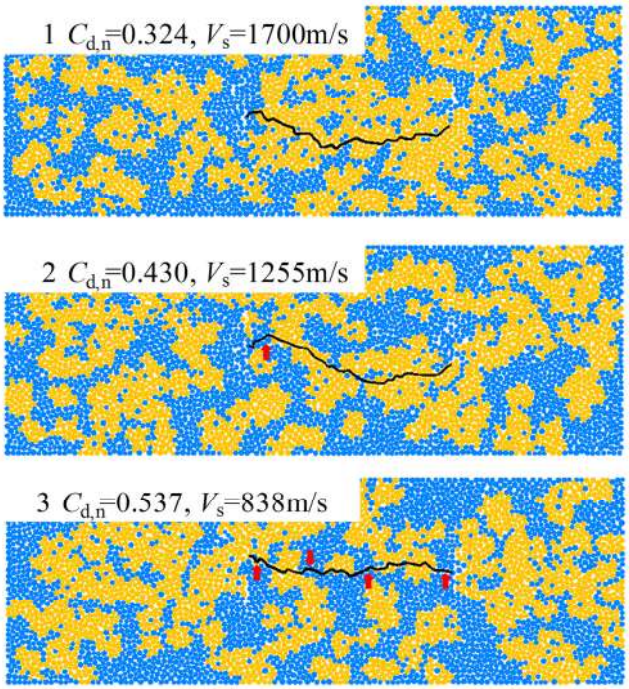


Fig. 12 Distribution of calcite-sand clusters in Samples 1, 2, and 3 at 5% calcite content. Black lines indicate the optimal wave propagation paths; red arrows indicate uncemented particles; $C_{d,n}$ refers to the normalized coordination number along the shear wave propagation path [23].

4 Discussion

4.1 Significance and implications

The Monte-Carlo application verified the rationality of the DEM simulation of calcite-cemented samples. Plenty of sample data also revealed evidence at the macro and micro levels. Based on the constructed sample dataset and valid sample generation method, features of quantifying spatial variability of calcite precipitates and relations between macro-level V_s and micro-level $C_{d,n}$ can be further discussed.

4.2 Feature correlation analysis

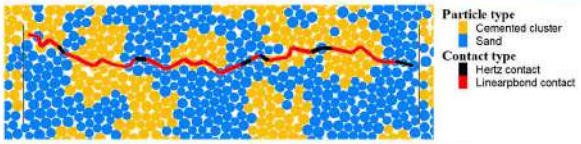
The validity of $C_{d,n}$ characterizing V_s can be evaluated based on the dataset constructed from 1972 samples. The lists of transmitters and receivers of 1972 samples were traversed, and Dijkstra's algorithm was used to calculate the optimal path of each sample. A Pearson correlation analysis was carried out of the eight contact-related features and the target value V_s (Fig. 13). The features pertaining to the Linearpbond contact were all strongly positively correlated with V_s (>0.8), with QLLC exhibiting the highest correlation. This indicates that continuous cementation contributed more to the increase in V_s than dispersed cementation. However, it is worth noting that QLLC was an absolute variable and may not be applicable when parameters such as wave propagation distance and particle size change. Therefore, characterization indicators should be chosen from relative variables, including LP, PLLC and $C_{d,n}$, rather than absolute features. Among these relative values, $C_{d,n}$ exhibited the strongest correlation with V_s, suggesting that the contact weight, based on physical significance, is a

reasonable means of characterizing V_s.

(a)

	C_d	QHC	QLC	QC	NLLC	LP	PLLC	$C_{d,n}$	V_s
C_d	1	0.98	-0.97	-0.02	-0.84	-0.98	-0.73	1	-0.87
QHC	0.98	1	-1	-0.19	-0.82	-1	-0.69	0.98	-0.83
QLC	-0.97	-1	1	0.26	0.8	1	0.67	-0.97	0.8
QC	-0.02	-0.19	0.26	1	-0.05	0.22	-0.23	-0.02	-0.16
NLLC	-0.84	-0.82	0.8	-0.05	1	0.81	0.97	-0.84	0.9
LP	-0.98	-1	1	0.22	0.81	1	0.69	-0.98	0.82
PLLC	-0.73	-0.69	0.67	-0.23	0.97	0.69	1	-0.73	0.84
$C_{d,n}$	1	0.98	-0.97	-0.02	-0.84	-0.98	-0.73	1	-0.87
V_s	-0.87	-0.83	0.8	-0.16	0.9	0.82	0.84	-0.87	1

(b)



(c)

	Feature	Definition	Value in (b)
Absolute variable	C_d	Weighted shortest distance from the transmitter to the receiver	15.28
	QHC	The quantity of hertz contacts on the shortest path	8
	QLC	The quantity of Linearpbond contacts on the shortest path	35
	QC	The quantity of all contacts on the shortest path	43
	QLLC	The quantity of Linearpbond contacts within the longest consecutive Linearpbond contacts on the shortest path	10
Relative variable	LP	The proportion of the number of Linearpbond contacts to all the contacts on the shortest path	0.814
	PLLC	The proportion of thenumber of Linearpbond contacts within the longest contsecutive Linearpbond contacts to the total number of contacts on the shortest path	0.286
	$C_{d,n}$	The normalized connectivity per unity distance	0.537
	V_s	Shear wave velocity measured	838.44

Fig. 13 Correlation analysis between features and V_s. (a) Correlation matrix; (b) An example sample shortest propagation path; (c) The feature definitions and examples

4.3 Quantitative characterization of $C_{d,n}$-V_s

Since $C_{d,n}$ represented the characteristic of the shortest distance in the microscopic scope, the empirical relationship of $C_{d,n}$-Vs can quantify the contribution of calcite cementation, predict Vs, and analyze sample variability.

The first step to calculate $C_{d,n}$ was selecting the weights of contacts. The above weights result was calculated by the coordination number, which was based on physical meaning (Eq. 8). More samples provided statistics information to select weights based on the Monte-Carlo method. The optimal values and appropriate intervals for the contact model weight can be calculated using the Monte-Carlo method, and the $C_{d,n}$-V_s correlation was found to be the most negative when the weight ratio is 0.83 (Fig. 14).

The quantified $C_{d,n}$ relationships based on physical meanings and the Monte-Carlo method were compared in Table 4. Both methods could characterize V_s without considering calcite content, but each had advantages and disadvantages. Weight selection based on physical meanings resulted in greater aggregation of wave velocity points, less error with the fitted curve, and more accurate V_s predictions, but the weights need to be recalculated in different simulations. The weights selected based on the Monte-Carlo method can be applied as constants, but the V_s of different samples were more discrete and less accurately predicted. Both weight selection results are applicable for using $C_{d,n}$ to characterize V_s, and making reasonable predictions.

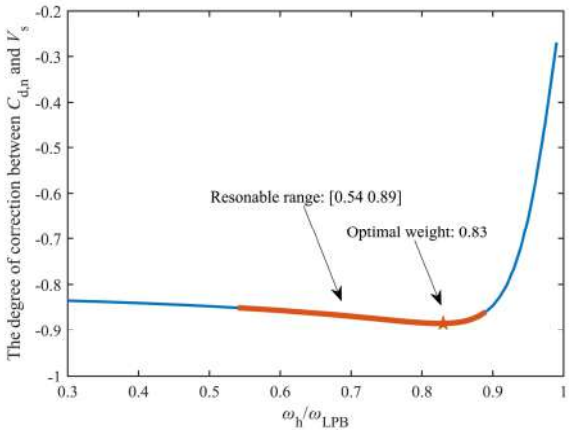
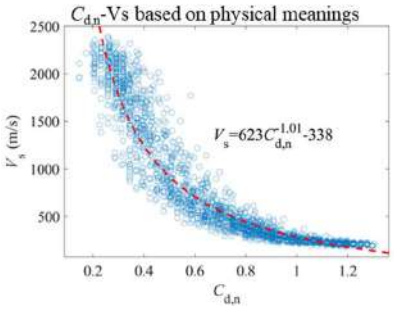
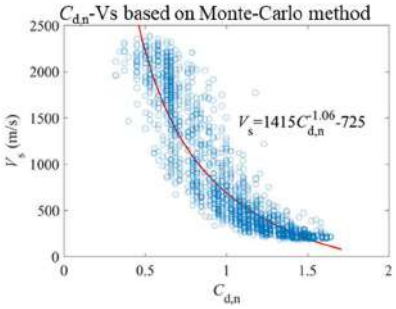


Fig. 14 The degree of correction between $C_{d,n}$ and V_s under different ω_{LPB}/ω_h

Table 4 Comparison of two methods of weight selection for contact model

	Based on Physical meanings	Based on Monte-Carlo method
ω_h	0.475	1
ω_{LPB}	0.328	0.83
$C_{d,n}$ range	[0.145, 1.303]	[0.318, 1.643]
V_s-$C_{d,n}$ fitting equation	V_s=623$C_{d,n}^{1.01}$-338	V_s=1415$C_{d,n}^{1.06}$-725
V_s-$C_{d,n}$ fitting figure		

5 Limitations and future work

Though the DE-MC method provides significant evidence for studying sand stiffness variability induced by stochastic distributions of calcite precipitates, it also has several limitations in quantitative analysis. Firstly, all samples were kept in the same K_0 stress state (i.e., the stress state was not considered); other microscopic features such as contact force, coordination number change, and force chain distribution were neglected. A recommended future direction was to expand the quantification of indicators like $C_{d,n}$ to consider different stress states and incorporate the aforementioned microstructural features. Secondly, the sizes of samples and particles were fixed and the diversity was not enough. Therefore, the generalization ability of the $C_{d,n}$-Vs derived formula needs further verification. Further study can focus on building a more complete physical-informed database based on existing simulation methods, to solve the huge computational problem of traditional numerical methods.

6 Conclusions

In this study, an innovative Discrete Element-Monte Carlo (DE-MC) method was applied in samples of calcite-cemented soil to quantify the sand stiffness variability induced by stochastic distributions of calcite precipitates. During the Monte Carlo process, a random variable X_{seed} represented the random distribution characteristic of $CaCO_3$-Ps (X_p) and $CaCO_3$-Cs (X_c) in different samples. A total of 1972 DEM samples were established. Cementation spatial distribution and V_s in the simulation were presented, and microscopic features were further constructed and analyzed. The following conclusions are drawn:

(1) The V_s frequency histograms and 95% confidence intervals revealed the difference of V_s in different samples, which was attributed to the preferential paths caused by calcite-sand clusters. Typically, the spatial distribution variability of calcite-sand clusters started to stand out in Stage II, and was the most obvious in Stage III, thereby leading to the significant distribution difference of V_s.

(2) Contact-related features of the sample dominated preferential paths, which can predict V_s. The correlation between eight selected contact-related features and V_s was calculated. The normalized connectivity per unity distance contact weight $C_{d,n}$ was the most correlated among relative variables.

(3) The availability of $C_{d,n}$ was evaluated by weighting contacts to quantify the cementation contribution. A lower $C_{d,n}$ value indicated greater connectivity and higher V_s. Two methods of weight selection, physical meanings and Monte-Carlo, were provided and compared, both showing advantages in practical prediction. The fitted formulas are $V_s=623C_{d,n}^{-1.01}-338$ and $V_s=1415C_{d,n}^{-1.06}-725$, respectively.

Acknowledgments

The work presented in this paper was supported by the National Natural Science Foundation of China (Award No.: 42177118), the Basic Science Center Program for Multiphase Evolution in Hypergravity of the National Natural Science Foundation of China (Award No.: 51988101), and Hong Kong Research Grants Council (RGC) (No. 17200123). The authors would also like to acknowledge the MOE Key Laboratory of Soft Soils and Geo-environmental Engineering.

References

[1] D.-S. Cho, J. Kim, Stability of the front wall and the horizontal behavior of composite reinforced-earth retaining walls, Acta Geotechnica (2024).

[2] J. Wang, H. Duan, K. Chen, I.Y.S. Chan, F. Xue, N. Zhang, X. Chen, J. Zuo, Role of Urban Underground-Space Development in Achieving Carbon Neutrality: A National-Level Analysis in China, Engineering (2024).
[3] J. Geng, J. Wang, J. Huang, D. Zhou, J. Bai, J. Wang, H. Zhang, H. Duan, W. Zhang, Quantification of the carbon emission of urban residential buildings: The case of the Greater Bay Area cities in China, Environmental Impact Assessment Review 95 (2022) 106775.
[4] J. Wang, Y. Huang, Y. Teng, B. Yu, J. Wang, H. Zhang, H. Duan, Can buildings sector achieve the carbon mitigation ambitious goal: Case study for a low-carbon demonstration city in China?, Environmental Impact Assessment Review 90 (2021) 106633.
[5] M. Nemati, Modification of porous media permeability, using calcium carbonate produced enzymatically in situ, Enzyme and Microbial Technology 33(5) (2003) 635-642.
[6] X. Pan, J. Chu, Y. Yang, L. Cheng, A new biogrouting method for fine to coarse sand, Acta Geotechnica 15(1) (2019) 1-16.
[7] M.-J. Cui, J. Chu, H.-J. Lai, Optimization of one-phase-low-pH enzyme-induced carbonate precipitation method for soil improvement, Acta Geotechnica (2024).
[8] I. Ahenkorah, M.M. Rahman, M.R. Karim, S. Beecham, Cyclic liquefaction resistance of MICP- and EICP-treated sand in simple shear conditions: a benchmarking with the critical state of untreated sand, Acta Geotechnica (2024).
[9] A. Nafisi, Q. Liu, B.M. Montoya, Effect of stress path on the shear response of bio-cemented sands, Acta Geotechnica 16(10) (2021) 3239-3251.
[10] J.C. Santamarina, A. Klein, M.A. Fam, Soils and waves:Particulate materials behavior, characterization and process monitoring, Journal of Soils and Sediments 1(2) (2001) 130-130.
[11] H. Lin, M.T. Suleiman, D.G. Brown, Investigation of pore-scale CaCO3 distributions and their effects on stiffness and permeability of sands treated by microbially induced carbonate precipitation (MICP), Soils and Foundations 60(4) (2020) 944-961.
[12] H. Wu, W. Wu, W. Liang, F. Dai, H. Liu, Y. Xiao, 3D DEM modeling of biocemented sand with fines as cementing agents, International Journal for Numerical and Analytical Methods in Geomechanics 47(2) (2022) 212-240.
[13] B. Bate, J. Cao, C. Zhang, N. Hao, S. Wang, Monitoring lime and cement improvement using spectral induced polarization and bender element techniques, Journal of Rock Mechanics and Geotechnical Engineering 13(1) (2021) 202-211.
[14] F. Tagliaferri, J. Waller, E. Andò, S.A. Hall, G. Viggiani, P. Bésuelle, J.T. DeJong, Observing strain localisation processes in bio-cemented sand using x-ray imaging, Granular Matter 13(3) (2011) 247-250.
[15] M.-J. Cui, J.-J. Zheng, R.-J. Zhang, H.-J. Lai, J. Zhang, Influence of cementation level on the strength behaviour of bio-cemented sand, Acta Geotechnica 12(5) (2017) 971-986.
[16] B. Bate, J. Cao, C. Zhang, N. Hao, Spectral induced polarization study on enzyme induced carbonate precipitations: influences of size and content on stiffness of a fine sand, Acta Geotechnica 16(3) (2020) 841-857.

[17] H. Lin, M.T. Suleiman, D.G. Brown, E. Kavazanjian, Mechanical Behavior of Sands Treated by Microbially Induced Carbonate Precipitation, Journal of Geotechnical and Geoenvironmental Engineering 142(2) (2016).

[18] M. Sarkis, A. Naillon, F. Emeriault, C. Geindreau, Tensile strength measurement of the calcite bond between bio-cemented sand grains, Acta Geotechnica 19(3) (2024) 1555-1570.

[19] J. He, J. Chu, S.-f. Wu, J. Peng, Mitigation of soil liquefaction using microbially induced desaturation, Journal of Zhejiang University-SCIENCE A 17(7) (2016) 577-588.

[20] C. Zeng, Y. Veenis, C.A. Hall, E.S. Young, W.R.L. van der Star, J.-j. Zheng, L.A. van Paassen, Experimental and Numerical Analysis of a Field Trial Application of Microbially Induced Calcite Precipitation for Ground Stabilization, Journal of Geotechnical and Geoenvironmental Engineering 147(7) (2021).

[21] H.-l. Kou, C. Wu, B.-A. Jang, D. Wang, Spatial Distribution of CaCO3 in Biocemented Sandy Slope Using Surface Percolation, Journal of Materials in Civil Engineering 33(6) (2021) 06021004.

[22] L.A.v. Paassen, R. Ghose, T.J.M.v.d. Linden, W.R.L.v.d. Star, M.C.M.v. Loosdrecht, Quantifying Biomediated Ground Improvement by Ureolysis: Large-Scale Biogrout Experiment, Journal of Geotechnical and Geoenvironmental Engineering 136(12) (2010) 1721-1728.

[23] M. Sun, J. Cao, J. Cao, S. Zhang, Y. Chen, B. Bate, Discrete element modeling of shear wave propagation in carbonate precipitate–cemented particles, Acta Geotechnica 17(7) (2022) 2633-2649.

[24] P. Yang, E. Kavazanjian, N. Neithalath, Particle-Scale Mechanisms in Undrained Triaxial Compression of Biocemented Sands: Insights from 3D DEM Simulations with Flexible Boundary, International Journal of Geomechanics 19(4) (2019) 04019009.

[25] K. Feng, B.M. Montoya, T.M. Evans, Discrete element method simulations of bio-cemented sands, Computers and Geotechnics 85 (2017) 139-150.

[26] E. Kashizadeh, A. Mukherjee, A. Tordesillas, Experimental and numerical investigation on heap formation of granular soil sparsely cemented by bacterial calcification, Powder Technology 360 (2020) 253-263.

[27] G. Zhou, T. Esaki, Y. Mitani, M. Xie, J. Mori, Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach, Engineering Geology 68(3) (2003) 373-386.

[28] Saifuddin, H. Yamanaka, K. Chimoto, Variability of shallow soil amplification from surface-wave inversion using the Markov-chain Monte Carlo method, Soil Dynamics and Earthquake Engineering 107 (2018) 141-151.

[29] S. Ma, Z. Wei, X. Chen, CFD-DEM combined the fictitious domain method with monte carlo method for studying particle sediment in fluid, Particulate Science and Technology 36(8) (2017) 920-933.

[30] F. Kalateh, M. Kheiry, A Review of Stochastic Analysis of the Seepage Through Earth Dams with a Focus on the Application of Monte Carlo Simulation, Archives of Computational Methods in Engineering 31(1) (2023) 47-72.

[31] X. Peng, D.-Q. Li, Z.-J. Cao, W. Gong, C.H. Juang, Reliability-based robust geotechnical design using Monte Carlo simulation, Bulletin of Engineering Geology and the Environment 76(3) (2017) 1217-1227.

[32] E. Kavazanjian, E. Iglesias, I. Karatas, Biopolymer soil stabilization for wind erosion control, 17th International Conference on Soil Mechanics and Geotechnical Engineering, ICSMGE 2009, 2009, pp. 881-884.

[33] R.D. Mindlin, H. Deresiewicz, Elastic Spheres in Contact Under Varying Oblique Forces, Journal of Applied Mechanics 20(3) (1953) 327-344.

[34] D.O. Potyondy, P.A. Cundall, A bonded-particle model for rock, International Journal of Rock Mechanics and Mining Sciences 41(8) (2004) 1329-1364.

[35] A.L. Fernandez, J.C. Santamarina, Effect of cementation on the small-strain parameters of sands, Canadian Geotechnical Journal 38(1) (2001) 191-199.

[36] M.H. Sadd, G. Adhikari, F. Cardoso, DEM simulation of wave propagation in granular materials, Powder Technology 109(1-3) (2000) 222-233.

[37] X. Zhou, Q. Sheng, Z. Cui, Dynamic boundary setting for discrete element method considering the seismic problems of rock masses, Granular Matter 21(3) (2019) 66.

[38] X.M. Xu, D.S. Ling, Y.P. Cheng, Y.M. Chen, Correlation between liquefaction resistance and shear wave velocity of granular soils: a micromechanical perspective, Géotechnique 65(5) (2015) 337-348.

[39] J. Ahn, G. Biscontin, J.M. Roesset, Wave propagation in nonlinear one-dimensional soil model, International Journal for Numerical and Analytical Methods in Geomechanics 33(4) (2008) 487-509.

[40] A.V. da Fonseca, C. Ferreira, M. Fahey, A Framework Interpreting Bender Element Tests, Combining Time-Domain and Frequency-Domain Methods, GEOTECHNICAL TESTING JOURNAL 32(2) (2009) 91-107.

[41] L. Wang, T. Xiao, S. Liu, W. Zhang, B. Yang, L. Chen, Quantification of model uncertainty and variability for landslide displacement prediction based on Monte Carlo simulation, Gondwana Research 123 (2023) 27-40.

[42] C.M. Jarque, Jarque-Bera Test, in: M. Lovric (Ed.), International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 701-702.

[43] Q.-y. DONG, Y.-q. CAI, C.-j. XU, J. WANG, H.-l. SUN, C. GU, Measurement of small-strain shear modulus Gmax of dry and saturated sands by bender element and resonant column tests, Chinese Journal of Geotechnical Engineering 35(12) (2013) 2283-2289.

[44] K.M. Rollins, M.D. Evans, N.B. Diehl, W.D.D. Iii, Shear Modulus and Damping Relationships for Gravels, Journal of geotechnical and geoenvironmental engineering (5) (1998) 124.

[45] X. Zhang, Y. Chen, H. Liu, Z. Zhang, X. Ding, Performance evaluation of a MICP-treated calcareous sandy foundation using shake table tests, Soil Dynamics and Earthquake Engineering 129 (2020) 105959.

[46] L. Yang, L. Salvati, Small Strain Properties of Sands with Different Cement Types, (2010).

[47] J. O'Donovan, E. Ibraim, C. O'Sullivan, S. Hamlin, D. Muir Wood, G. Marketos, Micromechanics of seismic wave propagation in granular materials, Granular Matter 18(3) (2016) 56.
[48] X. Gu, X. Liang, Y. Shan, X. Huang, A. Tessari, Discrete element modeling of shear wave propagation using bender elements in confined granular materials of different grain sizes, Computers and Geotechnics 125 (2020).
[49] Z. Ning, A. Khoubani, T.M. Evans, Shear wave propagation in granular assemblies, Computers and Geotechnics 69 (2015) 615-626.
[50] X.-m. XU, D.-s. LING, B. HUANG, Y.-m. CHEN, Determination of shear wave velocity in granular materials by shear vibration within discrete element simulation, Chinese Journal of Geotechnical Engineering 33(09) (2011) 1462-1468.
[51] B.C. Martinez, J.T. DeJong, T.R. Ginn, B.M. Montoya, T.H. Barkouki, C. Hunt, B. Tanyu, D. Major, Experimental Optimization of Microbial-Induced Carbonate Precipitation for Soil Improvement, Journal of Geotechnical and Geoenvironmental Engineering 139(4) (2013) 587-598.