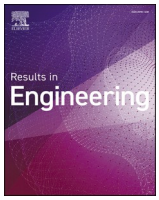
Results in Engineering 30 (2026) 110548

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering



Research paper

A machine learning assisted probabilistic framework for predicting the impact of crack characteristics on heterogeneous deformation of desiccated soils under environmental changes

Ali Mohammadi Kamizji, Hamed Sadeghi *, Milad Jabbarzadeh

Department of Civil Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Keywords:
Soil desiccation cracks
Probabilistic modelling
Reliability analysis
Risk assessment of subsidence and swelling
Environmental impacts

ABSTRACT

Progressive surface deformation of desiccated cracked soils in semi-arid to arid regions in response to seasonal changes has become a challenge to the performance of infrastructures. Hence, a comprehensive analysis of heterogeneous deformation due to changes in crack width was conducted by considering more reliable cumulative distributions for environmental parameters. The probabilistic modelling was performed and validated using data from field reconnaissance and numerical simulations to establish relationships between crack width, spacing, and surface deformation for different crack geometries. Monte Carlo sampling and the first-order second moment method were employed to determine probability distributions and evaluate subsidence risks. Moreover, a machine learning approach was employed to elucidate the relationships between relevant parameters. As a result, the significance of soil-atmosphere interplay for various parameters from the most to the least important was revealed for cracked ground surfaces. Relative humidity, precipitation, and radiation emerged as the most critical factors affecting deformation of cracked soil. The probabilistic estimation of crack widths and the reliability index are used to assess the likelihood that crack sizes will remain within acceptable limits, even under changing climatic conditions. This also accounts for the associated risks of ground subsidence and swelling caused by variations in crack widths.

1. Introduction

1.1. Background and motivation

Soil desiccation cracking is a common geotechnical phenomenon that plays a critical role in soil deformation under environmental actions such as wetting-drying cycles [1–3]. In high-plasticity soils, desiccation cracks disrupt structural uniformity, introduce pronounced heterogeneity, and alter mechanical and hydraulic properties once tensile stresses exceed soil tensile strength [4,5]. Although cracking is a natural process, its irregular spatial distribution leads to non-uniform subsidence and swelling during environmental cycles, posing significant geotechnical challenges [6]. Consequently, probabilistic analysis becomes essential for assessing the long-term reliability of structures founded on desiccation-cracked soils.

Probabilistic methods have gained increasing importance in geotechnical engineering for evaluating structural reliability under uncertainty [7,8]. Natural soil heterogeneity, amplified by environmental forcing such as rainfall, temperature fluctuations, and soil–atmosphere interactions, limits the effectiveness of purely deterministic approaches [9,10]. By explicitly accounting for uncertainties in soil properties and environmental drivers, probabilistic analyses provide a more realistic basis for predicting and mitigating geotechnical risks [11–14], particularly for estimating non-uniform deformation induced by desiccation cracking.

The influence of climatic conditions on soil cracking is well documented [15,16]. Drying and wetting cycles cause volumetric changes associated with moisture loss and gain, while desiccation cracks further intensify these processes by accelerating evaporation and infiltration [17,18]. During drying, cracks enhance moisture gradients and promote heterogeneous shrinkage, whereas during wetting, preferential infiltration along cracks induces uneven swelling [19,10,20]. These coupled processes can result in significant differential ground movement, adversely affecting buildings, pavements, slopes, and embankments [21–23]. Desiccation cracking is inherently stochastic, with random variations in crack aperture, spacing, and spatial distribution,

* Corresponding author.
E-mail addresses: al.mohammad@sharif.edu (A. Mohammadi Kamizji), hsadeghi@sharif.edu (H. Sadeghi), milad.jabbarzadeh@sharif.edu (M. Jabbarzadeh).

https://doi.org/10.1016/j.rineng.2026.110548
Received 27 December 2025; Received in revised form 2 April 2026; Accepted 14 April 2026
Available online 15 April 2026

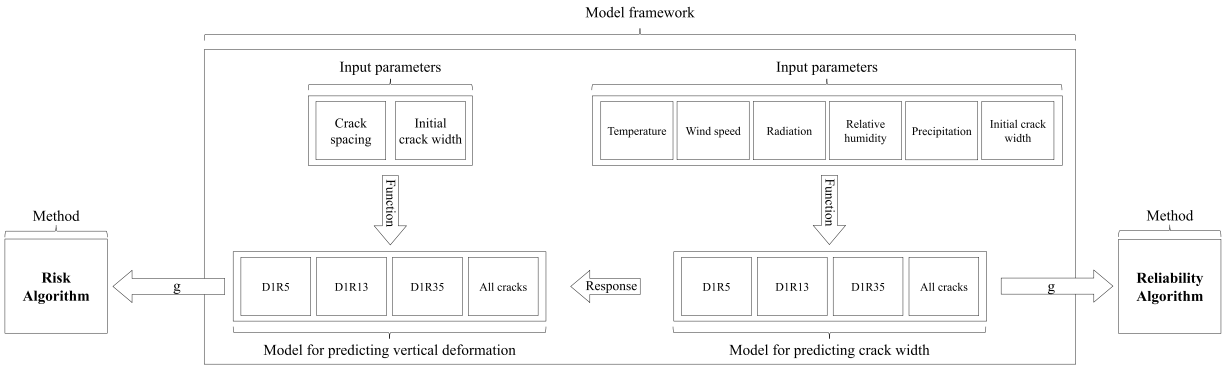


Fig. 1. General layout of the modelling process for reliability assessment of cracked soil deformation.

complicating predictions of deformation magnitude and location. Crack geometry plays a key role in soil response: wider cracks facilitate rapid infiltration during wetting, while narrower cracks increase exposed surface area and sensitivity to atmospheric conditions [24,25]. Accounting for this stochasticity is therefore critical in risk and reliability assessments.

Previous studies have investigated crack formation mechanisms [26], the influence of soil type and moisture content on cracking patterns [27], and climatic controls on crack evolution [28]. Laboratory studies have shown progressive degradation of soil strength and stiffness with repeated hydraulic cycles [29], while numerical simulations revealed increased crack volume and connectivity in expansive soils [30,31]. Field monitoring using distributed fiber optic sensing further identified cyclic crack opening and closure under environmental forcing, a process referred to as soil crack breathing [32]. This hysteretic behavior reflects the continuous interaction between the soil matrix and atmospheric conditions and governs heterogeneous ground deformation.

Despite these advances, existing studies predominantly rely on deterministic or limited probabilistic approaches and rarely address the combined stochastic effects of crack aperture and spacing on subsidence and swelling. This highlights the need for comprehensive probabilistic frameworks capable of representing random crack patterns and their interaction with environmental loading. Such frameworks can be developed using approximation or probabilistic models [33,34]. In this context, machine learning techniques—including artificial neural networks, ensemble trees, and support vector machines—offer promising tools for predicting desiccation cracking behaviour [35–37], although their application to soil cracking remains limited. Employing multiple ML models to assess sensitivity to input variability is therefore crucial for improving predictive robustness [36]. Although recent studies have increasingly applied probabilistic methods to soil cracking problems [34], significant gaps remain in addressing desiccation cracking under stochastic environmental loads. Traditional probabilistic approaches often rely on direct Monte Carlo simulations coupled with computationally expensive finite element models, which limits the feasibility of extensive risk assessment. Furthermore, while recent data-driven studies have successfully utilised machine learning for predicting soil parameters or crack patterns [38], they predominantly focus on deterministic prediction rather than utilising it as surrogate for calculating the reliability index or the probability of failure.

Recent advances have increasingly adopted the concept of Geotechnical Digital Twins (GDTs), in which data-driven or hybrid models provide continuously updatable representations of soil–structure systems for monitoring and decision support. In this context, surrogate models are essential for enabling rapid prediction, uncertainty quantification, and scenario evaluation that are impractical with high-fidelity numerical simulations alone. Although the present study does not aim to develop a complete GDT, the proposed machine-learning–assisted probabilistic framework constitutes a foundational step toward such applications by efficiently translating climatic inputs into crack evolution and deformation risk metrics suitable for integration with near-real-time monitoring data. Recent studies [39,38] have demonstrated GDT applications in deep excavation and landslides, while also highlighting computational cost as a key limitation; the physics-informed surrogate developed herein directly addresses this issue for desiccation-sensitive infrastructure.

1.2. Research objectives and engineering relevance

Despite these advances, a critical gap remains in quantifying how stochastic climatic variability propagates through crack evolution to influence deformation risk in desiccated soils. Most existing studies rely on deterministic analyses or finite-probability approaches applied to prescribed drying-wetting scenarios, which limits their ability to capture the combined effects of climate uncertainty, crack-geometry variability, and serviceability-oriented performance metrics. In particular, the relative influence of climatic drivers across different crack-development stages and their implications for subsidence–swelling risk remain poorly constrained. To address this gap, the present study proposes an integrated probabilistic framework that couples machine-learning–based crack-width prediction with Monte Carlo uncertainty propagation and reliability-based risk assessment, enabling scenario-specific evaluation of climate-driven deformation in cracked soils. This enables engineers to move from qualitative statements to quantified, design-relevant outputs under realistic climate variability. In addition, the framework provides importance ranking of environmental drivers (via First-Order Second-Moment method (FOSM)), supporting prioritisation of monitoring and mitigation in semi-arid to arid regions where seasonal crack breathing governs serviceability risk.

2. Material and methods

In order to model the probabilistic behaviour of variations in crack aperture and soil subsidence or swelling under different climatic conditions, and then calculate the reliability index of crack width and the risk of subsidence, it is important to first identify the influential climatic parameters and establish the relationships between them. The following steps were considered to develop a framework for analysing the variations in crack width under the influence of meteorological data, and then finding the relationship between soil deformation (subsidence in drying conditions and swelling in wetting conditions) and both climatic conditions and crack width and spacing. The general framework of this study is illustrated in Fig. 1, which is discussed in detail in the following steps.

2.1. Soil-atmosphere interaction analysis

The first step is to determine the factors and parameters that affect the changes in crack width in real environmental conditions.

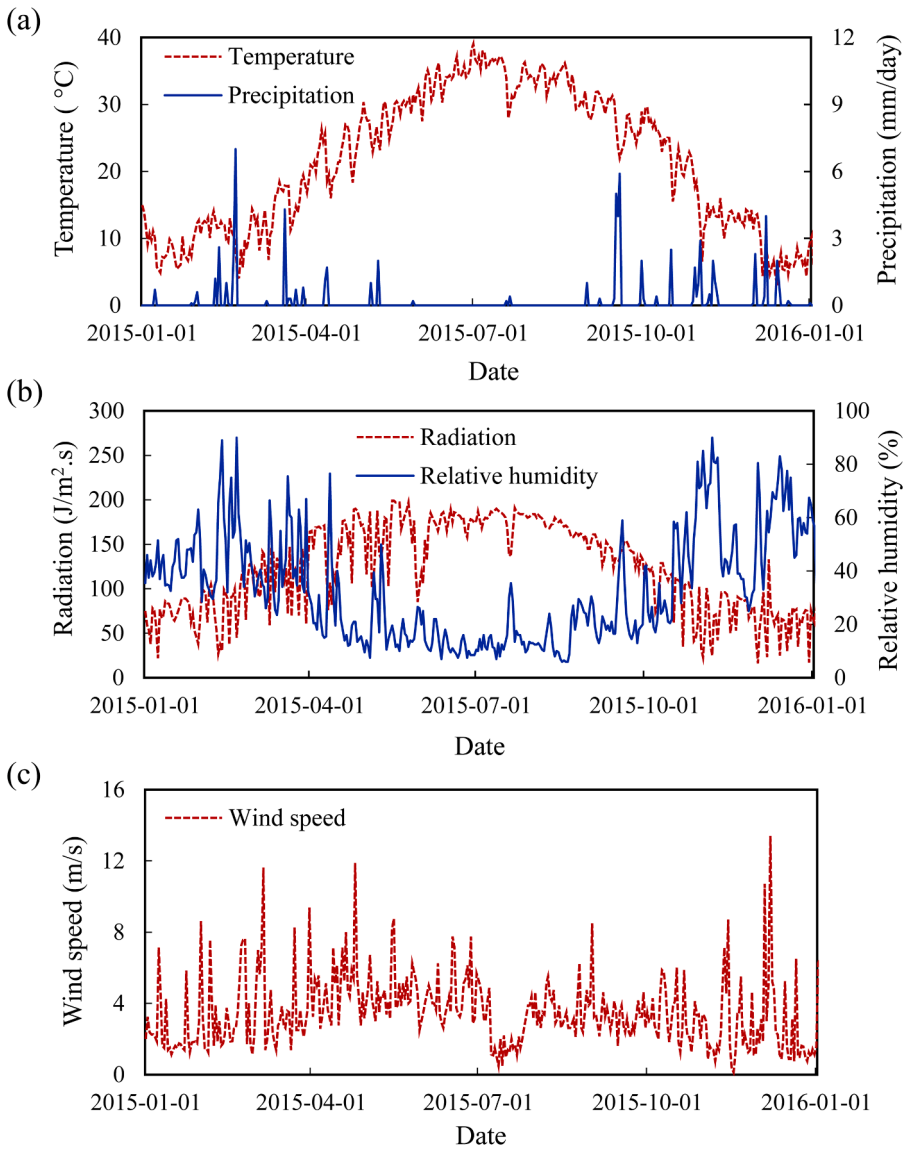


Fig. 2. Environmental data for Qom city, (a) temperature and precipitation, (b) radiation and relative humidity, and (c) wind speed.

Specifically, one-year meteorological data of the Qom city located in Iran with arid to semi-arid environmental condition was collected [40]. Qom is known for its arid climate with severe ground subsidence due to groundwater exploitation and surface soil shrinkage [41,42]. These data include temperature, wind speed, relative humidity, precipitation, and radiation, as shown in Fig. 2.

Using measured environmental data, the probability distribution of these parameters can be found, as shown in Fig. 2. Sadeghi et al. [6] conducted a thermo-hydro-mechanical modelling to investigate heterogeneous subsidence and swelling in a desiccation-cracked soil. Their model was able to solve the soil-atmosphere problem and analyse the desiccation cracks dynamics under environmental change. In this study, the variations in crack width obtained from their numerical models were utilised to establish probabilistic relationships between changes in crack width and the resulting soil deformation through reliability analysis. The data used for surrogate modelling and probabilistic analysis were generated using a coupled thermo–hydro–mechanical framework with well-defined physical and geometric boundary conditions. The groundwater table in the study area was identified at approximately 90 m below ground surface, and a bi-linear pore-water pressure profile was imposed to represent depth-dependent suction conditions in the unsaturated zone. Transient climatic boundary conditions, including temperature, radiation, and precipitation time series measured for the Qom region, were applied at the soil surface and along crack walls. Liquid and gas flow were governed by the generalised Darcy law, vapour diffusion by Fick's law, and heat transfer by Fourier's law. The mechanical behaviour of the soil was simulated using the Barcelona Expansive Model (BExM), which captures suction-dependent strain evolution and is capable of reproducing shrinkage–swelling cycles induced by climatic wetting–drying. These boundary conditions and constraints define the physical domain within which the generated data and subsequent probabilistic analyses are valid.

The next step is to determine the probabilistic distribution of climatic variables. Therefore, the cumulative frequency of the available one-year data was utilised. In Fig. 3, the cumulative frequency at each point was obtained by dividing the number of data points less than or equal to that value by the total number of data points.

This method provided the cumulative distribution chart necessary for the analysis. Consequently, in order to be able to use the cumulative distribution, these charts were replaced with common cumulative distribution charts that have a specific relationship. Then, a graphical method was employed to identify the appropriate standard distribution for each climatic parameter.

Specifically, the cumulative distribution of each climatic parameter

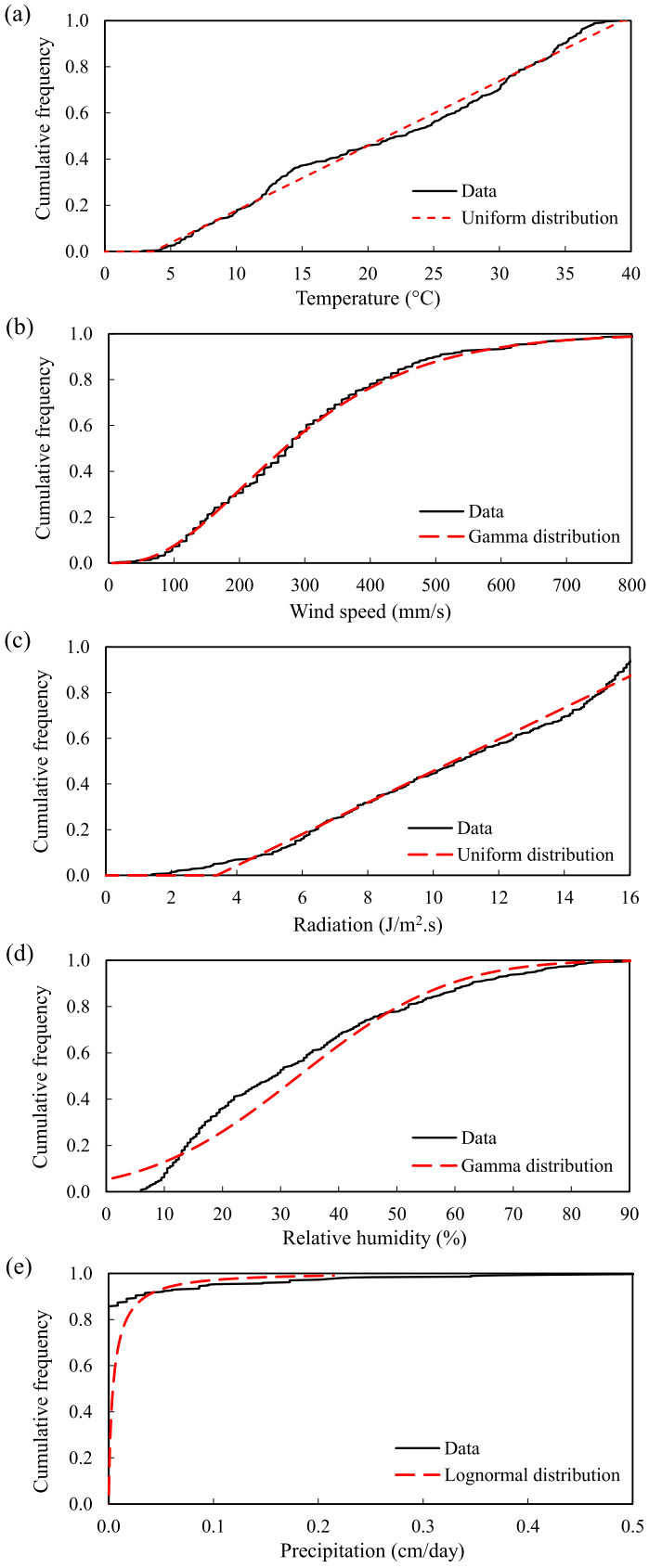


Fig. 3. Comparison of the probabilistic distribution functions for: (a) temperature, (b) wind speed, (c) radiation, (d) relative humidity, and (e) precipitation.

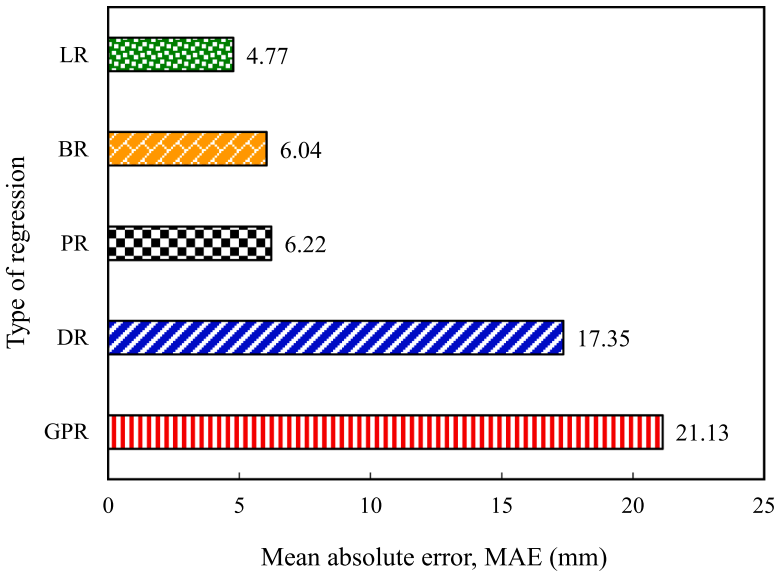


Fig. 4. Evaluation of the best regression technique through comparison of the mean absolute error.

was compared with the cumulative distributions of Normal, Log-normal, Exponential, Uniform, Laplace, and Gamma functions. The distributions that closely matched the cumulative distribution of the data were selected. The mean and standard deviation of the cumulative distribution obtained from the data were then applied to the selected standard distribution. As a result, the standard distributions with the real mean and standard deviation related to the climatic variables were used in place of the original distribution of the climatic parameters, as shown in Fig. 3.

2.2. Machine learning approaches for regression analysis

The variations in crack width obtained from numerical models were correlated with the corresponding environmental parameters to establish a relationship derived from the ML regression models. Statistical analysis was then used to model the probability distribution of crack width as a function of the input environmental variables. The developed predictive models in this study provide a framework for predicting crack-induced soil heterogeneous deformation and associated geohazard risks in desiccated soils. The input features, including temperature, wind speed, precipitation, radiation, and relative humidity, are expressed in heterogeneous physical units and exhibit substantially different numerical scales. To ensure numerical stability during model training and to reduce bias and variance induced by scale imbalance, all input features were standardised prior to model development. Specifically, each feature was transformed using standardisation based on its mean and standard deviation, resulting in zero-mean and unit-variance variables. This preprocessing step ensures equal contribution of all features to the training process, improves convergence behaviour, and enhances overall predictive performance. The ML regression models used in this study include Linear (LR), Bayesian (BR), Polynomial (PR), Dummy (DR), and Gaussian Process regressions (GPR). To train the model, after randomising the data, which includes weather parameters and crack width, 70% of the data was used for training and the remaining 30% was used for testing. To determine which of the regression models are more suitable for finding the relationship and modelling the crack width, the mean absolute error (MAE) was used which can be defined as:

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n} \quad (1)$$

Fig. 4 compares the performance of the examined regression models based on the MAE metric. According to this criterion, linear regression yields the lowest prediction error and is therefore identified as the most suitable surrogate model for crack-width estimation in this study. This behaviour can be attributed to the restricted climatic domain of the dataset and the limited variability of crack-width responses, which exhibit an approximately linear distribution within the investigated range. Under such conditions, linear regression provides improved stability and lower variance. In contrast, higher-complexity approaches, such as BR and GPR, are more sensitive to data sparsity, limited feature space, and constrained output distributions, which may lead to variance error and an increased tendency toward overfitting. The ML framework was implemented using Python (version 3.9) and the Scikit-learn library (version 1.2). For the predictive modelling, a Multiple Linear Regression algorithm based on the ordinary least squares method was employed. Given the linear nature of the correlation between the climatic inputs and crack characteristics, the standard LR estimator from Scikit-learn was utilized. The model was configured with the default hyperparameters. A linear relationship was then established between environmental parameters and crack width.

According to the comprehensive physics-based survey of crack geometry by Sadeghi et al. [6], three initial crack widths were considered based on statistical analysis of available experimental and field studies in previous studies. Since crack geometric properties are highly dependent on the scale of study, a dimensionless parameter, crack ratio (C_R), is proposed, which is the ratio of crack width to crack depth. A dataset was provided for the crack ratio, and analysis found that it follows a log-normal distribution. Based on the probability density function for the crack ratio, three data model values, including model of data 5%, the mean 13%, and the mean plus twice the standard deviation 35%, were chosen. Therefore, given a crack depth of 1 m, the crack width can be simply calculated as 50 mm, 130 mm, and 350 mm. The selected crack widths are associated with C_R values of 5% (C_R5), 13% (C_R13), and 35% (C_R35). In another word, data are firstly normalised to ensure consistency across different spatial scales and measurement techniques. A statistical analysis was then conducted, and representative crack widths were selected from a fitted log-normal distribution to capture low, intermediate, and high crack-development regimes. The spacing between cracks exhibits a random distribution. This model is not restricted to cracks with a specific initial width, offering a more generalised method for predicting soil crack behaviour. In the numerical model, it is assumed that cracks are positioned next to each other, with spacing ranging from 1 m to 10 m.

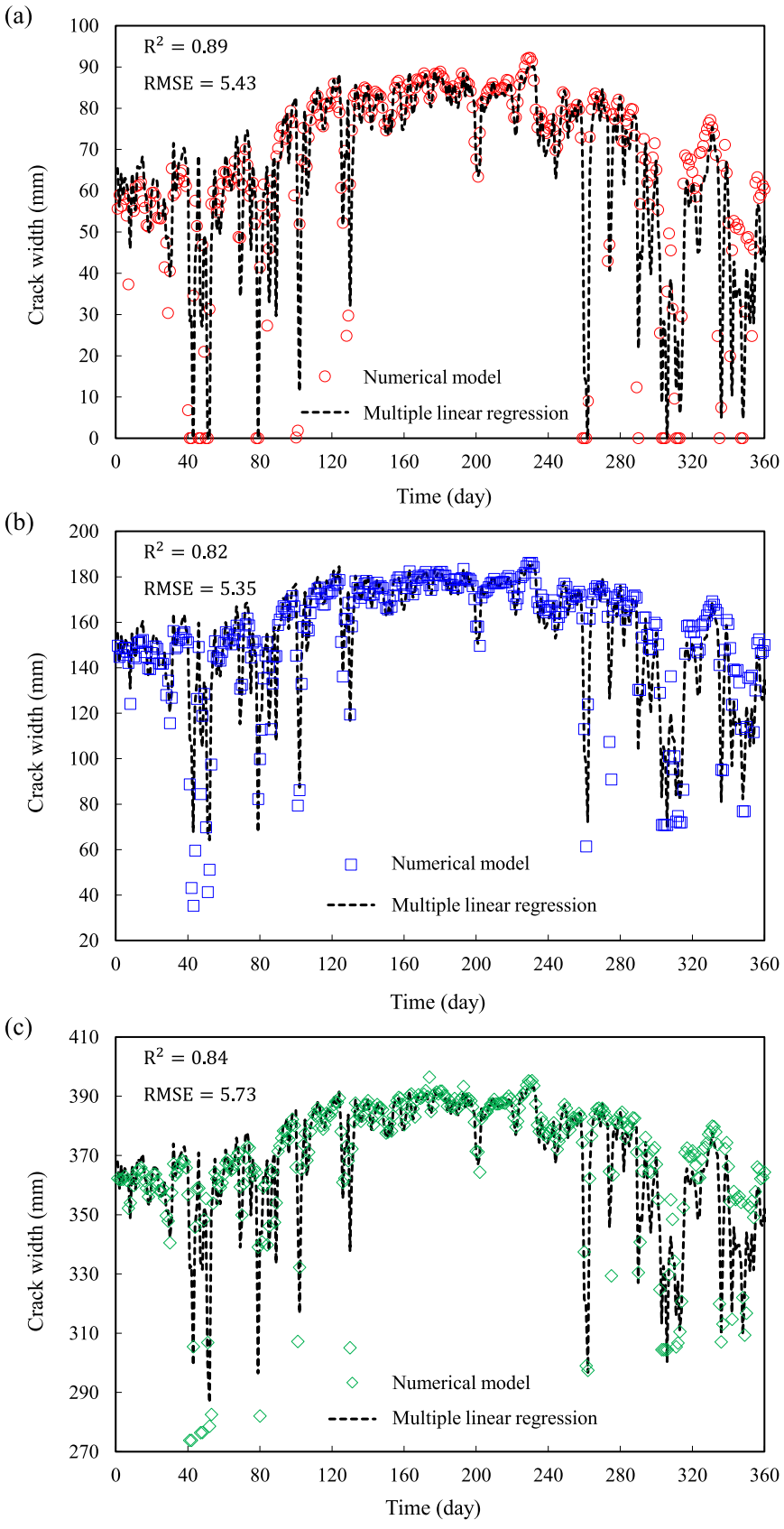


Fig. 5. Comparison of the results of numerical model and ML-based linear regression model for crack width in (a) C_R5, (b) C_R13, and (c) C_R35 crack geometry.

Table 1
The coefficients of the linear regression parameters for relationship between crack width and environmental factors.

Parameter	Modelling scenario		
	C_R5	C_R13	C_R35
A	100.5292	205.3566	400.6014
B	0.0197	-0.2174	0.0045
C	-1.5447×10^{-10}	-1.1433×10^{-9}	-7.4233×10^{-10}
D	-102.2489	-97.7481	-96.5282
E	-0.3831	-0.3092	-0.1775
F	-0.9267	-1.2611	-0.8930

Fig. 5 shows the accuracy of the linear model by comparing it with results of the numerical models, using the R factor and the root mean squared error (RMSE). RMSE values were evaluated with respect to the characteristic magnitude of crack widths for each crack scenario. To facilitate a scale-independent comparison of prediction errors among cracks with different average widths, the normalised root mean squared error (NRMSE) was additionally adopted, which is defined as:

$$NRMSE = \frac{RMSE}{\bar{y}} \tag{2}$$

Where $\bar{y}$ is the mean crack width associated with each crack category. Using this normalisation, the resulting NRMSE values for crack types C_R5, C_R13, and C_R35 are 8.4%, 3.5%, and 1.6%, respectively. The increased scatter observed in Fig. 5 is primarily due to seasonal transitions of environmental conditions. These periods correspond to immediate changes in temperature and moisture conditions, during which small variations in climatic inputs can induce amplified variability in crack-width response. In contrast, mid-season intervals characterised by relatively stable climatic conditions exhibit smoother and more coherent trends. It is observed that the standard deviation of the probabilistic model is lower than that of the numerical results by approximately 2–4 mm. This difference is attributed to the inherent smoothing effect of the linear regression surrogate model used in this study. The regression model approximates the numerical response with a RMSE of approximately 5.5 mm, effectively filtering out high-frequency fluctuations present in the raw data. Consequently, the probabilistic framework captures the primary trends and mean behaviour accurately, while the variance associated with the residual errors is naturally excluded.

Based on the LR method, the variations in crack width with weather conditions can be predicted as:

$$C_W = \mathrm{A} + \mathrm{B}(T) + \mathrm{C}(W) + \mathrm{D}(P) + \mathrm{E}(R) + F(H) \tag{3}$$

where C_W is the crack width, T (°C) is the temperature, W (mm/s) is the wind speed, P (cm/day) is the precipitation, R (J/m^2.s) is the radiation, and H (%) is the relative humidity. In addition, A, B, C, D, E, and F are the coefficients of parameters, which are summarised in Table 1 for each model scenario.

It is also possible to make a relationship between all three scenarios and define a unique equation depending on both climatic factors and initial crack width as:

$$C_w = 58.2411 - 0.0644\,(T) - 1.8516 \times 10^{-10}(W) - 98.8417\,(P) - 0.2899\,(R) - 1.0269\,(H) + 1.0033\,(C_{w0}) \tag{4}$$

where C_W is the crack width, T (°C) is the temperature, W (mm/s) is the wind speed, P (cm/day) is the precipitation, R ($J/m2.s$) is the radiation, H (%) is the relative humidity and the C_{w0} (mm) is the initial crack width. Using this equation, the changes in crack aperture due to environmental changes can be predicted using meteorological data and initial aperture of the crack (C_{w0}).

In addition, the linear regression models were trained independently for each individual crack dataset, with hyperparameters systematically tuned for each configuration. Model generalisation was assessed by comparing the coefficient of determination (R^2) obtained on the training and testing datasets. For all three crack geometries, the testing R^2 values ranged between 0.84 and 0.89, while the corresponding training R^2 values were approximately 0.96. The relatively small gap between training and testing performance, indicates stable generalisation behaviour and the absence of pronounced overfitting. Given this evidence, additional regularisation techniques such as Lasso (L1), Ridge (L2), or Elastic Net were not employed. Under the present conditions, introducing regularisation was considered unnecessary and could potentially bias the regression coefficients without yielding meaningful improvements in predictive accuracy or model robustness.

2.3. Monte Carlo sampling method and model validation

After determining the cumulative distributions of the environmental parameters and also finding the linear relationship between the crack width and environmental parameters, the Monte Carlo Sampling (MCS) method was used to determine the probability distribution of crack width. In this method, using MCS and 200,000 samples, the cumulative distribution from probabilistic modelling for crack width is determined. From this probabilistic cumulative distribution, the exceedance and non-exceedance probability distribution of crack width can be calculated for cracks with different initial crack widths. As a result, it is possible to calculate the mean and standard deviation of the crack width and cumulative distribution of its probabilities using Rtx software [43, 44], which is a powerful software for calculations and performing probabilistic operations. Furthermore, to evaluate the accuracy of the distribution derived from the probabilistic cumulative distribution method, the cumulative distribution of crack width data obtained from numerical modelling is constructed and compared to the probabilistic modelling's results. The mean and standard deviation, as well as the graphical form of the two cumulative distributions obtained from the numerical method and the probabilistic method, were compared. This comparison is shown in Fig. 6 for three cracks with different initial widths. To quantitatively assess the agreement between the distributions obtained from the probabilistic and numerical methods, in addition to visual inspection, the non-parametric Kolmogorov-Smirnov test was employed. By examining the cumulative distribution function (CDF), this test determines whether two datasets are drawn from the same underlying distribution without requiring specific prior assumptions regarding the distribution types. The test statistic, D, is defined as follows:

$$D = \max_x \left| F_{prob}(x) - F_{num}(x) \right| \tag{5}$$

where $F_{prob}(x)$ and $F_{num}(x)$ denote the CDF values of the probabilistic and numerical methods, respectively. Based on the analysis of the plots in Fig. 6, the D statistic was calculated for cracks C_R5, C_R13, and C_R35. The maximum discrepancies observed for these three samples were 0.09, 0.07, and 0.08, respectively (occurring approximately at crack widths of 75, 145, and 375 mm). A comparison of these values indicates that the C_R13 sample, with the lowest D value, exhibits the best fit between the probabilistic and numerical results, followed by C_R35 and C_R5. Given the low values of the D statistic, as well as the proximity of the means obtained from both methods, it can be concluded that the probability distributions derived from both methods for all three cracks are statistically consistent and show no significant discrepancy. The sufficiency of this sample size was verified using the Standard Error of the Mean (SEM) as follows:

$$SEM = \frac{\sigma}{\sqrt{N}} \tag{6}$$

where the σ is the standard deviation and the N is the sample size. Considering the scenario with the highest observed variability

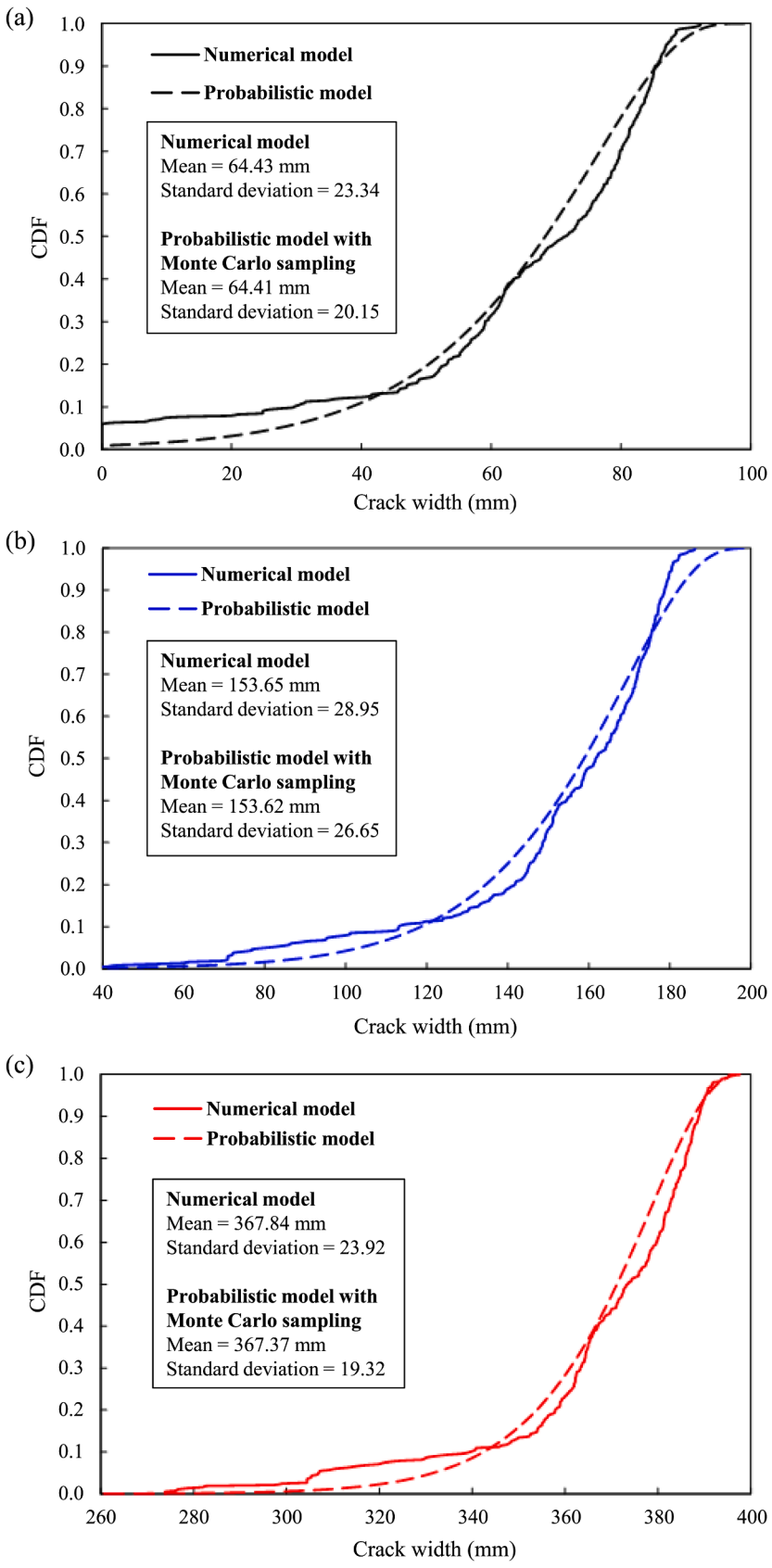


Fig. 6. Comparing the CDF of crack width between the numerical and the probabilistic model for: (a) C_R5, (b) C_R13, and (c) C_R35 crack geometry.

Table 2
The comparison of the mean and standard deviation for 200,000 and 500,000 Monte Carlo samples.

Crack scenario	Mean value		Standard deviation	
	200,000 sampling	500,000 sampling	200,000 sampling	500,000 sampling
C_R5	64.41	64.42	20.15	20.16
C_R13	153.62	153.61	26.65	26.62
C_R35	367.37	367.56	19.32	19.32

($\sigma \approx 28.95$ mm), the SEM is calculated as 0.065mm. This negligible margin of error indicates that the chosen sample size is sufficient to achieve stable statistical moments and minimises sampling uncertainty. However, to rigorously ensure that the probabilistic models have achieved global convergence and to validate the stability of the results, we conducted a secondary verification test. In this phase, the Monte Carlo simulation was repeated with a significantly larger dataset consisting of 500,000 samples for the critical crack scenarios: C_R5, C_R13, and C_R35. Table 2 presents a comparative analysis of the mean and standard deviation values obtained from both sample sizes (200,000 and 500,000). As indicated in the table, despite increasing the sample size by a factor of 2.5, the variations in the statistical outputs are minimal and practically insignificant. This consistency confirms that the simulation results had already converged at 200,000 samples.

It is important to note that Rtx functioned as the primary stochastic computational engine in this study. Distinguished by its intuitive framework for reliability analysis, Rtx offers superior versatility compared to similar platforms, particularly in handling user-defined limit state functions through both analytical methods (e.g., FOSM, FORM) and simulation-based approaches. Specifically, the explicit linear transfer function derived from the ML regression (Eq. 2) was seamlessly embedded within the Rtx environment. While the software facilitates various reliability algorithms, it was configured here to perform Monte Carlo simulations by randomly sampling the input PDFs, fitted to field meteorological and geometric data, and propagating them through the embedded equation to generate the output CDFs for serviceability risk.

2.4. FOSM method and model validation

In addition to the MCS method, there is another technique called FOSM method, which is used to obtain the mean and standard deviation of a distribution based on its relationship. Although FOSM is an approximate method, it is highly accurate when the relationship used is linear, as in linear regression. One of the advantages of the FOSM method is its ability to identify the significance of probability parameters in probability modelling. This helps determine which climatic parameters have a greater impact on the probability distribution of crack width. The FOSM method is used when the function relating weather parameters to crack width has a linear mean and standard deviation, allowing the method to effectively handle this linear relationship:

$$g(X) \approx g(M_x) + \nabla g^T(M_x)(x - M_x) \tag{7}$$

The g is a function of the vector X which is the vector of random variables and M_x is the vector of their means. Hence, if the variance of the relationship is considered, it will be:

$$\mathrm{Var}[g(X)] = \nabla g^T(M_x) \sum_{XX} \nabla g(M_x) \tag{8}$$

If the expression for the variance calculation in the FOSM method is expanded, it will be:

$$\mathrm{Var}[g(X)] = (\nabla g_1 \sigma_1)^2 + (\nabla g_2 \sigma_2)^2 + (\nabla g_3 \sigma_3)^2 + \ldots + (\nabla g_n \sigma_n)^2 + \sum_{i=1}^{n} \times \sum_{j=1}^{n} \nabla g_i g_j \sigma_i \sigma_j \rho_{ij} \tag{9}$$

The direct participation or contribution of the i-th random variable (i.e. the i-th environmental parameter) to this variance can be expressed as:

$$x_i = (\nabla g_i \sigma_i)^2 \tag{10}$$

Additionally, the indirect participation of the i-th variable X is achieved through correlation coefficients. Therefore, the total importance of a variable is determined by both its direct participation based on σ_i^2 and its indirect participation based on σ_i. A criterion defined in this methodology is the importance measurement, which is the absolute value of the gradient of the function g with respect to σ_i, or $|\nabla g_i\ \sigma_i|$. If this importance measurement is larger for a given parameter, it indicates that the parameter and its associated variable have greater importance in the overall variance. In the FOSM method, evaluating this importance measure is key to pinpointing the most critical environmental variables for refining crack width probabilistic modelling.

$$\boldsymbol{\omega} = \nabla g^T \boldsymbol{D_x} \tag{11}$$

where ω is the importance measure of the input variables and $\boldsymbol{D_x}$ is the covariance matrix of the input variables X.

2.5. Risk assessment based on the probabilistic distribution

To model the probability of soil subsidence or swelling, the probabilistic distribution of the input parameters is first established. A key parameter is the variance of crack width, which was previously obtained through probabilistic modelling using linear regression and the cumulative distribution of weather factors. Additional input variables for modelling soil subsidence or swelling include the distance between successive cracks (C_S) and the initial width of the crack (C_{W0}). The distances from the first to the thirteenth crack are specified. First, based on the cumulative distribution of these values for each crack with a specified initial width, the closest common cumulative distribution is identified from among common probabilistic distributions, including Normal, Log-normal, Exponential, Uniform, Laplace, and Gamma. The mean and standard deviation from the cumulative distribution of distances was then used to graphically fit the closest matching common distribution. In the numerical model, the crack spacing values were selected randomly since cracks occur irregularly on the ground surface. Based on field observation data and initial simulations, the crack spacing range was chosen to be between 1 m and 10 m. Although the observed crack spacings varied broadly, the dataset size was insufficient to reliably fit a specific higher-order parametric distribution (like Normal or Lognormal). Under such conditions of data scarcity, assuming a preferential concentration around a mean without ample evidence introduces epistemic bias. Therefore, a uniform distribution over the observed range was adopted as the most unbiased. For the distance between cracks, the uniform distribution with a mean of 5150 mm and a standard deviation of 2222.8 was found to be the most suitable.

To complete the probabilistic modelling of soil subsidence and swelling, a relationship between changes in crack width, the distance between two cracks, and the initial crack width is required. This relationship is determined using linear regression through ML for various initial crack widths. Using these relationships and the probabilistic distributions of each input parameter, the Monte Carlo method is applied to generate the cumulative distribution of subsidence under various scenarios. This allows the mean, standard deviation, and overall distribution to be determined. By utilising the cumulative distribution of

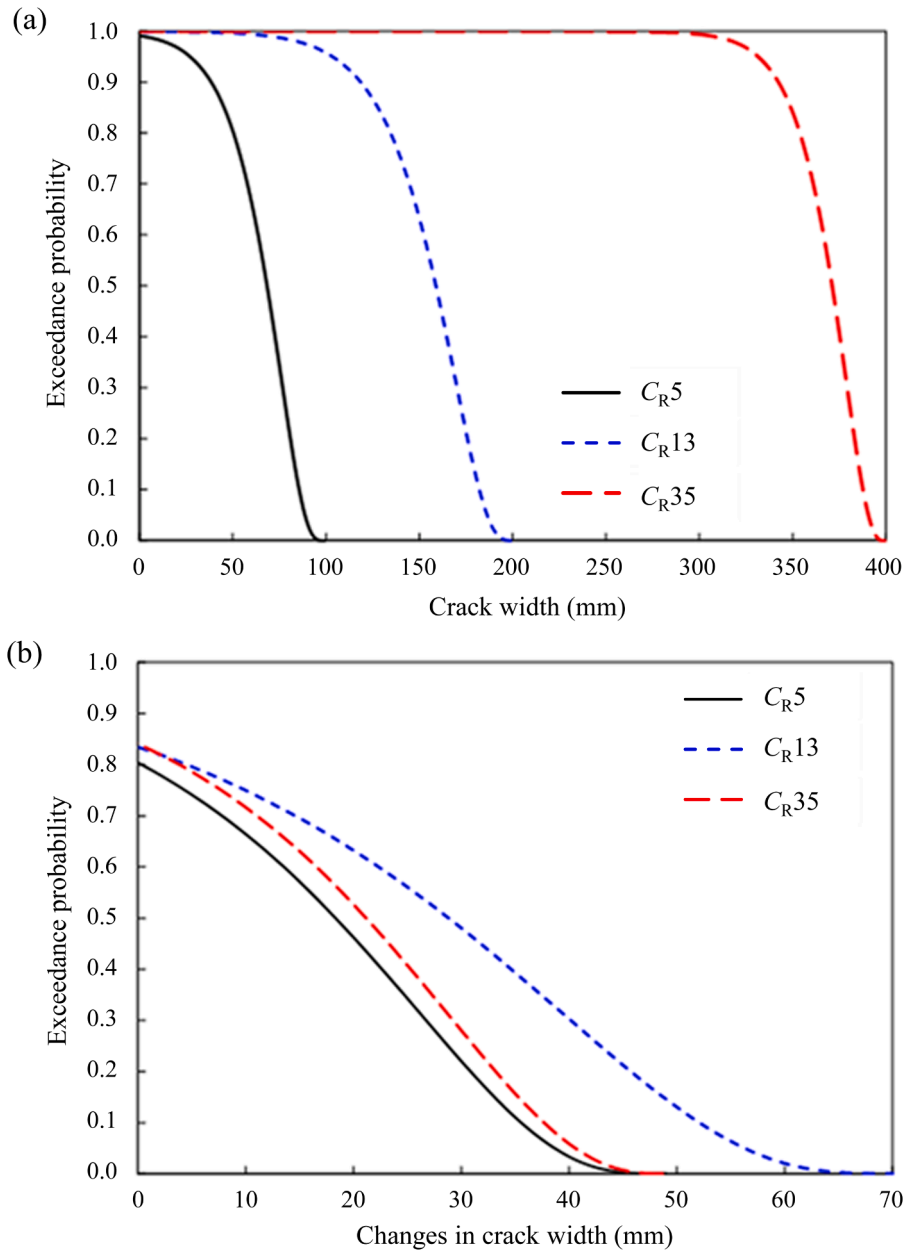


Fig. 7. Exceedance probability curves for: (a) crack width and (b) Changes in the crack width (mm), reflecting the probability distributions of influential parameters.

subsidence obtained through the Monte Carlo method, the risk of subsidence and swelling can be further analysed, considering that subsidence is a consequence of changes in crack width.

3. Result and discussion

3.1. Reliability analysis based on exceedance probability distribution

The close alignment between the mean and standard deviation values obtained from both numerical and probabilistic modelling, along with the high similarity in the probability distributions of crack width produced by the two methods, confirms the accuracy of the probabilistic modelling results. After confirming the results obtained from probabilistic modelling, it is possible to analyse the reliability index derived from the probability distribution of crack widths for all three types of cracks with initial widths of 5, 13, and 35 cm. This analysis is represented by the exceedance probability distribution diagram, as shown in Fig. 7(a). From the cumulative distribution diagram obtained through probabilistic modelling, the exceedance probability diagram for crack width can be derived as its complement. This relationship is defined as:

$$\mathrm{CCDF} = 1 - \mathrm{CDF} \tag{12}$$

where CDF is cumulative distribution function which represents the probability that a given variable will be less than or equal to a specific value as shown in Fig. 3. Complementary Cumulative Distribution Function (CCDF) is exceedance probability, or the probability that a given variable (such as crack width) will exceed a specific value.

Each point in this diagram shows a specific crack width value on the horizontal axis and the corresponding probability on the vertical axis, indicating the likelihood that the crack width will exceed the specified value at that point. For example, the probability of the crack width exceeding 7.5 cm is about 35% for C_R5, while it is 100% for both C_R13 and C_R35, which aligns with their initial crack sizes. Additionally, the probability of the crack width exceeding 17.5 cm is approximately 18% for C_R13 and 100% for C_R35, which is expected given that the initial size of C_R35 is greater than 17.5 cm. For C_R5, the probability is zero,

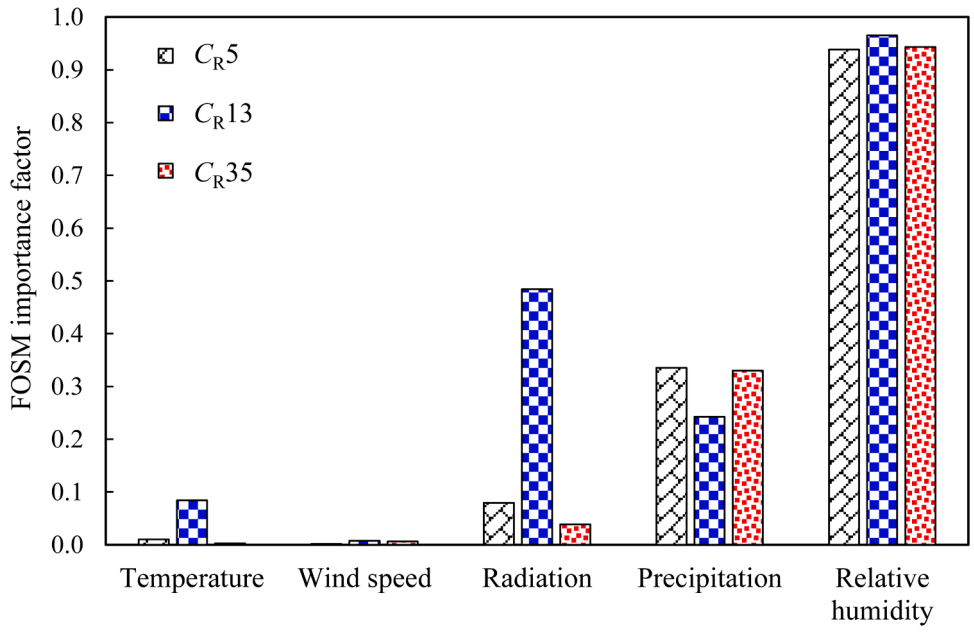


Fig. 8. The relative importance of five environmental parameters affecting the crack width based on FOSM method.

indicating that it is not possible for the crack size to reach this value. Given that the initial width of the C_R35 crack is greater than that of the other two cracks, it is expected that wider cracks are more likely to occur, as shown in the diagram. Similarly, it was anticipated that the C_R13 would exhibit larger widths compared to the C_R5 crack, due to its greater initial width.

Fig. 7(b) compares the probability of changes in crack width among the three scenarios. The results indicate that the probability of changes is higher for the C_R13 crack compared to the others. Additionally, the probability of crack width changes in C_R35 is higher than in C_R5, relative to their initial widths. For example, the probabilities that the variation in crack width exceeds 2.5 cm for the C_R5, C_R13, and C_R35 cracks are 34%, 56%, and 40%, respectively. This indicates that the probability of crack deformation under environmental change is highest for the C_R13 crack. Contrary to the previous graph, this result shows that even though the initial width of the C_R13 crack is less than that of C_R35, and the probability of larger crack widths is lower for C_R13 compared to C_R35, the probability of changes in crack width is greater for C_R13 than for C_R35. This suggests that larger cracks reach equilibrium faster during cyclic wetting and drying caused by rainfall and dry seasons.

The reliability index is a useful tool for assessing the likelihood that cracks will remain within specified limits, even in the face of environmental variations. In fact, the reliability index (β) serves as a measure of a system's reliability in not exceeding a specified threshold. It can be considered a metric indicating the distance from a critical point where failure may occur. The reliability index is defined as:

$$\mathrm{EP} = \phi\ (\beta) \tag{13}$$

where EP is exceedance probability and a point in the curves of Fig. 7, and β is a dimensionless number that indicates how many standard deviations away the threshold is from the mean of the distribution. The CDF of the standard normal distribution, denoted as ϕ, is used to represent the probability that a standard normal random variable is less than or equal to a specified value. A higher β indicates that the system is more reliable and safer, similar to having a larger safety margin.

In this analysis, a threshold of 30 mm for crack width variation was adopted as the limit state. As shown in Fig. 7 (b), the exceedance probabilities for cracks C_R5, C_R13 and C_R35 are 21%, 48%, and 29%, corresponding to β values of 0.81, 0.05, and 0.55, respectively. To provide a physical interpretation, reliability metrics correspond directly to the mean variations observed in the secondary cracks, which are 14.41 mm for C_R5, 17.37 mm for C_R35, and 23.63 mm for C_R13. Although the mean variation for C_R13 (23.63 mm) is technically below the 30 mm threshold, it is significantly closer to the limit compared to the other cracks. This proximity explains the extremely low reliability index ($\beta = 0.05$) and indicates that the infrastructure is operating with a negligible safety buffer. Therefore, unlike C_R5 which maintains a safe distance from the threshold, C_R13 represents a near-critical condition where the serviceability of the infrastructure is compromised, necessitating immediate monitoring or maintenance actions to prevent functional failure. A practical example of this risk is evident in the study region of Qom, where extensive networks of deep desiccation cracks have propagated into the immediate vicinity of high-voltage transmission towers [19]. These critical infrastructure assets are highly sensitive to differential settlement. The near-critical condition of C_R13 implies that even minor climatic fluctuations could push ground deformation beyond the serviceability limit, potentially destabilising the tower foundations and compromising the daily electricity supply.

The notably high probability of exceedance for the C_R13 scenario is attributed to two governing factors. First, the specific climatic forcing of the Qom region (arid to semi-arid with the maximum temperature of 39.1°C and the minimum relative humidity of 6%) imposes significantly more severe hydraulic gradients and evaporation rates than typical reference studies, driving aggressive soil desiccation. Second, the C_R13 geometry appears to function as a critical aspect ratio for instability; it possesses sufficient aperture to facilitate rapid deep-soil moisture exchange, unlike the confined C_R5, yet retains a highly reactive soil structure that has not yet reached the hydro-mechanical equilibrium observed in the wider C_R35 cracks.

3.2. FOSM analysis of parameters affecting the crack width

Given the linear relationship between crack width and weather factors, the FOSM analysis is highly efficient. This method determines the significance of each random variable in modelling crack width using the FOSM importance factor. It is important to note that the probabilistic analysis presented herein, including the FOSM importance factors, assumes statistical independence among the climatic input parameters. While these parameters are physically dependent on soil-atmosphere interactions, this study adopts the independence assumption to

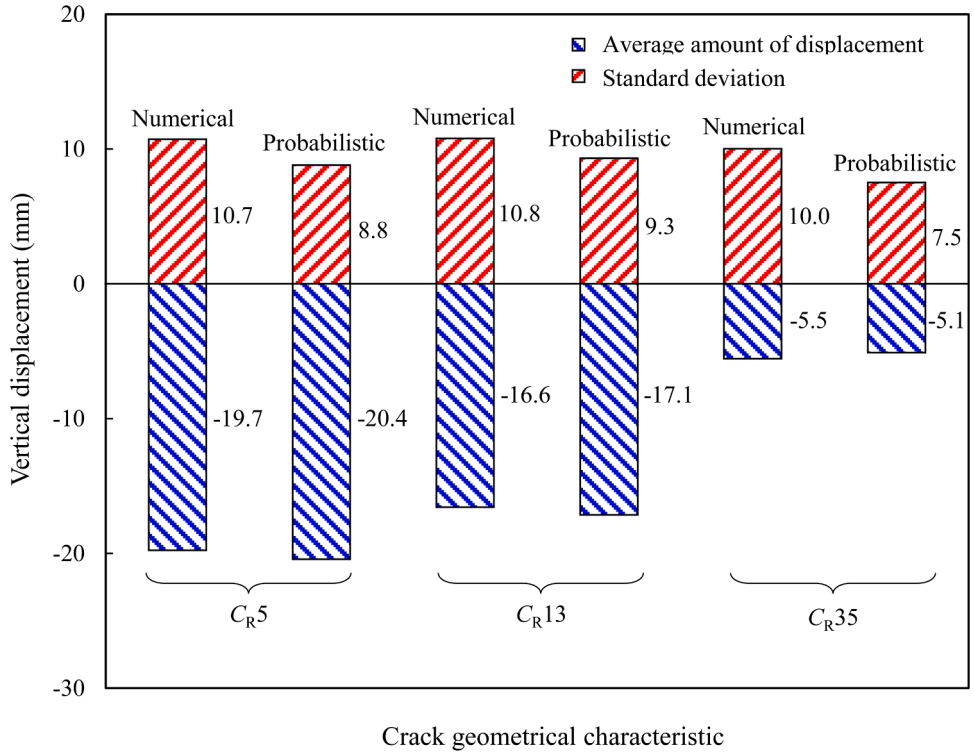


Fig. 9. Comparing the results of numerical and probabilistic models in terms of the mean and standard deviation of vertical displacement for all scenarios.

simplify the derivation of the surrogate model and to avoid the complexities associated with constructing joint probability density functions. Therefore, the dominant role of relative humidity is identified within the context of uncorrelated inputs. The parameter ω is defined as follows:

$$\boldsymbol{\omega} = -\frac{\nabla g^T \boldsymbol{D_x}}{\| \nabla g^T \boldsymbol{D_x} \|} \tag{14}$$

This standardises the vector size, enabling the comparison of different elements and providing a clear sense of their magnitude ratios. According to the FOSM analysis for importance factor, Fig. 8 shows the most influential parameter for all three cracks is the moisture percentage, which has the highest importance value across the cracks. The analysis also reveals that the temperature parameter has minimal impact on C_R5 and C_R35, but is significant for C_R13. Conversely, the radiation parameter shows low importance for C_R5 and C_R35, while it is more significant for C_R13. In contrast, the precipitation parameter has a greater impact on the crack width for C_R5 and C_R35 compared to C_R13. The results show that in the wetting and drying paths the relative humidity and radiation are the most influential climatic parameters in soil swelling and shrinkage, respectively. The C_R13 geometry acts as an optimal evaporative channel; its aperture is sufficiently wide to minimise aerodynamic resistance and maximise radiative heat capture, yet the surrounding soil matrix retains sufficient saturation to drive active shrinkage. In contrast, the wider C_R35 cracks represent a soil mass that has approached its residual moisture content corresponding to the soil shrinkage limit. Consequently, it becomes insensitive to further thermal forcing by temperature or radiation, i.e., no further deformation physically occurs along the drying path, but remains highly responsive to precipitation (wetting path), as the large aperture facilitates rapid infiltration and subsequent swelling [22]. This confirms that intermediate crack geometries (C_R13) are governed by thermodynamic drivers.

The dominance of relative humidity observed in the FOSM results (Fig. 8) can be explained by its direct control on soil-atmosphere moisture exchange. Relative humidity governs the vapour pressure deficit between the soil surface (including exposed crack walls) and the surrounding air, thereby directly regulating evaporation rate, matric suction development, and associated shrinkage strains. In contrast to temperature and radiation, which influence evaporation indirectly through energy availability, relative humidity acts as a primary driver of the moisture flux gradient. In desiccated and cracked soils, this effect is further amplified because crack networks increase the effective evaporating surface area and shorten vapour diffusion paths, making crack-width evolution highly sensitive to relative humidity fluctuations. As a result, variations in relative humidity exert a dominant influence on crack development and deformation risk, consistent with the FOSM sensitivity trends shown in Fig. 8. This statistical finding aligns with the experimental observations of Zeng et al. [28], who demonstrated that fluctuations in environmental humidity are the governing trigger for the initiation and propagation rate of desiccation cracks in clayey soils.

3.3. Risk analysis of the variations in crack width

As mentioned in Section 2.5, a linear relationship can be developed to predict heterogeneous deformation in cracked soils. The dependency of soil vertical deformation on environmental changes can be predicted using the parameters of crack width and spacing as follows:

$$C_R5: \ \delta = 0.3141 - 0.4164(C_W) + 0.001177\ (C_S) \tag{15}$$

$$C_R13: \delta = 29.9244 - 0.3393(C_W) + 0.00098(C_S) \tag{16}$$

$$C_R35: \delta = 138.1255 - 0.3839(C_W\ ) - 0.00042\ (C_S) \tag{17}$$

In these predictive equations, C_S has a known value can be measured and C_W can be calculated using Eq. (2) or Eq. (3) based on measured environmental parameters. Considering these scenarios all together, a unique equation for soil deformation can be defined based on initial crack width, variable crack width, and crack spacing as:

$$\delta = 0.4165(C_{W0}) - 0.3674(C_W\ ) + 0.00057(C_S\ ) - 19.1177 \tag{18}$$

where δ (mm) is the vertical displacement of soil, $C_{W0}(mm)$ is the initial crack width and $C_S(mm)$ is the spacing between the cracks. To verify the

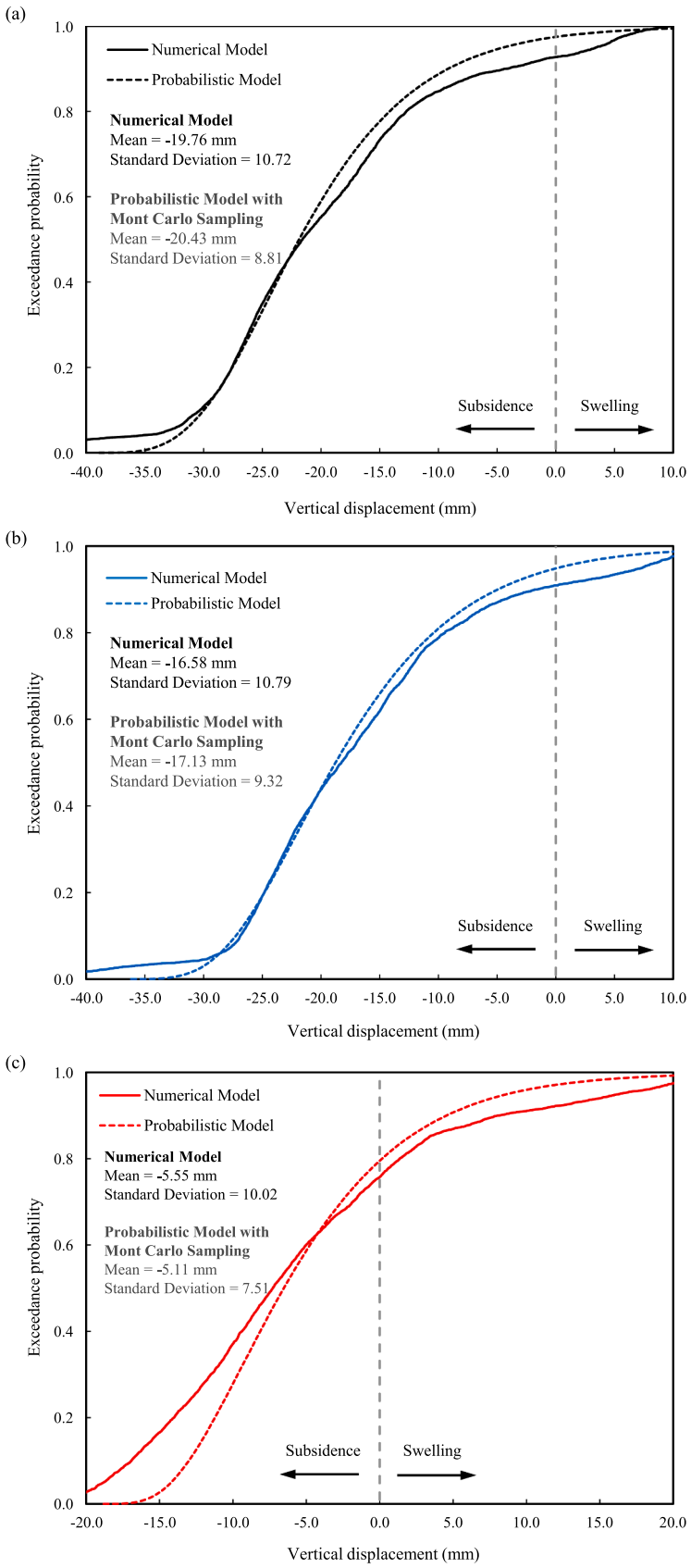


Fig. 10. Comparison of numerical and probabilistic models in terms of the cumulative distribution function (CDF) of vertical displacement for (a) C_R5, (b) C_R13, and (c) C_R35.

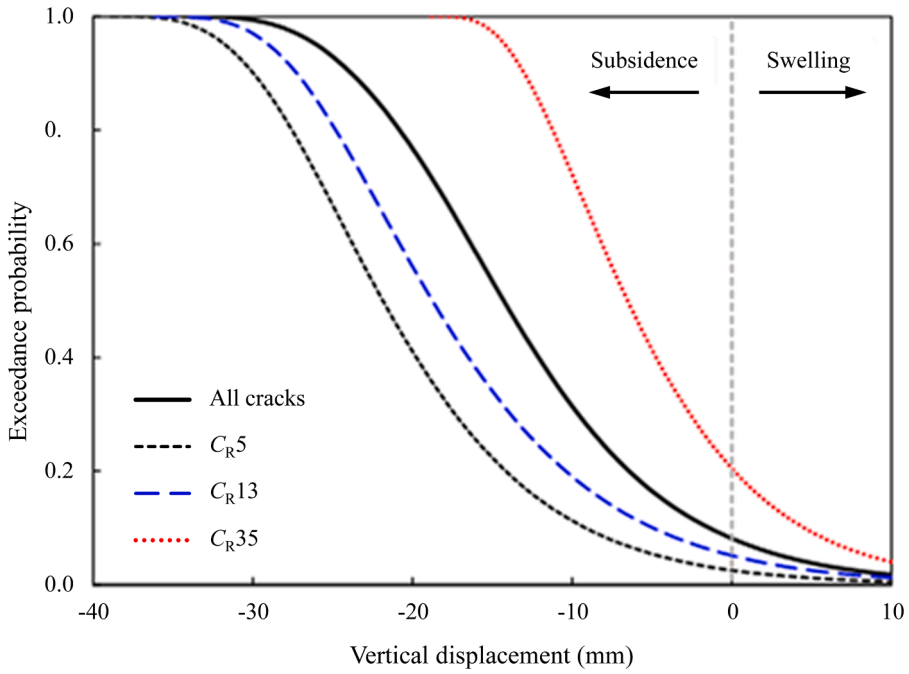


Fig. 11. Risk curves of vertical displacement in cracked soils reflecting the probability distributions of influential parameters.

accuracy of the probabilistic modelling results, the cumulative distribution from this method was compared with that obtained from numerical modelling. The average values and standard deviations from both methods are illustrated in Fig. 9, and the cumulative distributions obtained from both methods are shown in Fig. 10. As shown in Fig. 9, the mean and standard deviation obtained from the probabilistic and numerical modelling methods are very close to each other, indicating the accuracy and validity of the probabilistic modelling approach.

The risk analysis of changes in crack width, which assesses subsidence and swelling, involves calculating risk by multiplying the probability of occurrence by the effect, i.e., the extent of vertical soil deformation. As soil deformation is continuous, a graph illustrating subsidence and swelling is used for risk assessment. This analysis is derived from the CCDF shown in Fig. 10, which represents the CDF of ground subsidence and swelling. The diagram in Fig. 11 includes a vertical line that separates subsidence from swelling, helping to identify the probability for each condition. Furthermore, the results illustrate the general transition between subsidence and swelling, with the initial crack width treated as a probabilistic variable.

Each point in the probability diagram in Fig. 11 represents the likelihood that the amount of subsidence is equal to or less than the value indicated on the horizontal axis. For example, in Fig. 11, the probability that soil subsidence is 15 mm or less is 22% for C_R5, 34% for C_R13, and 97% for C_R35. Overall, the general probability for soil with existing cracks is 53%. The findings suggest that the probability and risk of subsidence generally decrease with the initial crack width. As observed, the probability that soil subsidence is 25 mm or less is 100% for C_R35, while it is <90% for C_R5 and C_R13. For instance, the value corresponding to a 50% probability — meaning there is a 50% chance that the amount of subsidence will be less than or equal to this value — is 22 mm for C_R5, 19 mm for C_R13, and 7 mm for C_R35. This indicates that the risk of subsidence decreases as the initial crack width increases.

As the probability of swelling increases, the probability of subsidence decreases. The results indicate that the risk of swelling is higher for C_R35 compared to the other two cracked soils. This suggests that increasing the initial crack width heightens the potential for swelling, as wider cracks allow more infiltration during rainfall, leading to greater expansion. Specifically, the probability of swelling is 2% for C_R5, 4% for C_R13, and approximately 20% for C_R35. In this research area, which is characterised by arid environmental conditions, the likelihood of subsidence is significantly greater than that of swelling due to arid environment and deep groundwater conditions. This pronounced dominance of subsidence over swelling is driven by the coupling of arid climatic forcing and the intrinsic hysteretic behaviour of the soil material. Climatically, the meteorological conditions in Qom result in prolonged high-evaporation periods that drive continuous shrinkage, whereas precipitation events are often too short or infrequent to induce full saturation and volume recovery. Physically, this behaviour is governed by the fundamental disparity between the soil's stiffness during loading and unloading phases. According to Terzaghi's consolidation theory, the swelling index (C_S) for fine-grained clays is typically only 0.2 to 0.25 of the compression index (C_C). Consequently, the volumetric strain recovery during the wetting path (swelling) is mechanically limited to a fraction of the strain accumulated during the drying path (subsidence). This physical characteristic ensures that net settlement accumulates over time, even under cyclic suction changes. This finding aligns with recent regional studies by Garakani et al. [41] and Sadeghi et al. [42], who identified cumulative land subsidence—rather than transient cyclic heave—as the governing infrastructure risk in this locale. It is critical to note that subsidence driven by desiccation cracking and surface shrinkage is distinct from the subsidence mechanisms observed in humid climates. In regions such as Shanghai, land subsidence is predominantly governed by the consolidation of saturated layers due to groundwater drawdown in the absence of surface cracks [45]. Consequently, the probabilistic framework presented here is specifically developed for arid to semi-arid environments, where the serviceability risk is controlled by the suction variations in the unsaturated zone rather than saturated consolidation.

Additionally, the overall probability of swelling for soils with any initial crack width is approximately 7%, which tends to decrease as the crack width increases. The results demonstrate that while the risk of subsidence decreases with an increase in the initial crack width, the risk of swelling increases. This is evident from the higher probability of swelling in C_R35 compared to C_R5 and C_R13. Although subsidence remains the predominant hazard in this arid environment, the 20% swelling probability observed for the C_R35 scenario represents a non-negligible engineering risk. Cracks with higher width facilitate rapid deep-soil saturation by acting as a preferential hydraulic pathway

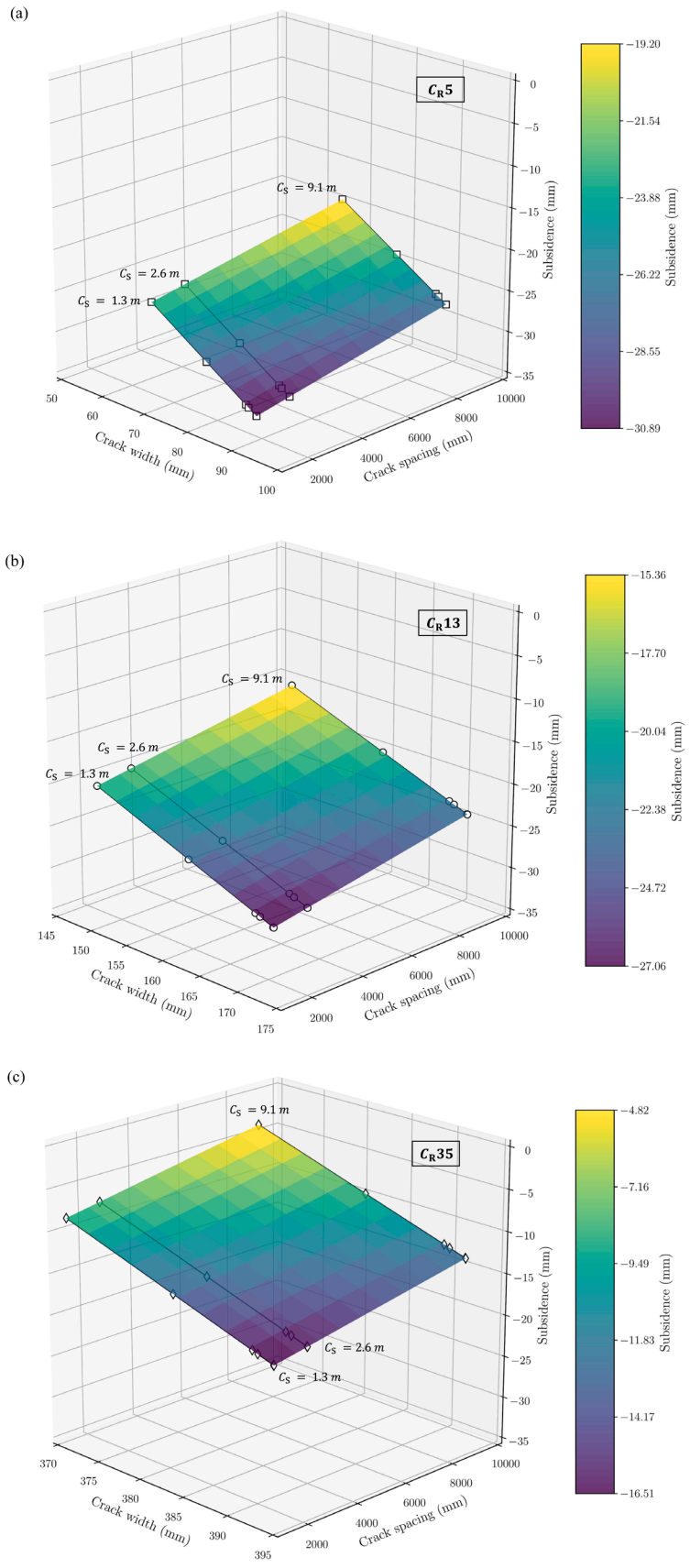


Fig. 12. Three-dimensional illustration of soil subsidence as a function of crack width and spacing along a drying path.

Table 3
Comparison of the present study with recent investigations on desiccation cracking.

Reference	Methodology	Objective	Key finding	Comparison with current study
Cordero et al. [2]	Large-scale field test	Crack pattern evolution under natural evaporation	Confirmed that crack formation is non-uniform and boundary-dependent	**Agreement:** Validates assumption of heterogeneous deformation. **Advance:** This study also quantifies the heterogeneity probabilistically.
Zeng et al. [28]	Laboratory tests (Controlled chamber)	Influence of relative humidity (RH) on cracking	Identified RH as a dominant factor in crack initiation and propagation	**Agreement:** Aligns with FOSM result (Fig. 8) ranking RH as the most critical parameter. **Advance:** This study extends the idea to cyclic RH conditions rather than static controlled values.
Tang et al. [5]	Laboratory wetting-drying cycles	Mechanical behaviour (shear strength) degradation	Cracks induce irreversible structural degradation and strength loss.	**Complementary:** While Tang et al. focus on strength, our study focuses on the deformation risk (subsidence) resulting from this degradation.
Gao et al. [30]	Numerical seepage simulation	Internal crack propagation and hydraulic conductivity	Cracks significantly increase permeability and infiltration depth.	**Agreement:** Supports the finding that wide cracks (C_R35) pose higher swelling risks due to rapid infiltration. **Advance:** We translate this hydraulic behaviour into a reliability index for infrastructure serviceability.

during precipitation events. Such swelling episodes can contribute to cyclic distress, particularly when followed by subsequent drying that reopens cracks and amplifies damage accumulation. For infrastructure systems such as pavements, shallow foundations, and embankments, these cyclic shrink–swell responses may lead to differential movements, loss of serviceability, and increased maintenance demand. Therefore, swelling-related risks should be considered alongside subsidence when assessing long-term performance under variable climatic conditions.

The analysis reveals a distinct hydraulic regime shift governing the deformation risk. As shown in Fig. 11, the probability of subsidence decreases markedly as the initial crack width increases. This inverse relationship is physically driven by the transition from an evaporation-dominated regime to an infiltration-dominated regime. Narrower cracks (C_R5, C_R13) primarily enhance evaporation and shrinkage (leading to subsidence) while limiting rapid water ingress [24,25]. Conversely, wide cracks (C_R35) function as open channels that facilitate high-velocity preferential flow during precipitation events, allowing water to saturate deep soil layers rapidly [30,46]. This rapid saturation of deep layers compared with intact surficial soil layers, i.e. in the absence of cracks, triggers immediate swelling (leading to heave), which hydro-mechanically counteracts the subsidence potential. Therefore, the reduced subsidence risk for wide cracks is not an indication of stability, but rather a signal of increased susceptibility to reaching the serviceability limit state due to swelling.

3.4. Quantified visualisation and comparative analysis of subsidence and swelling

A three-dimensional diagram can be constructed based on the linear relationship between subsidence and swelling, considering the influence of initial crack width, spacing, and changes in crack width. Due to this linear relationship, the subsidence-C_W-C_S diagram forms a plane in three-dimensional space, as shown in Fig. 12.

These graphs were selected based on a five-day warming sequence to illustrate the trend of soil drying under fluctuating seasonal weather conditions. The results show that subsidence increases as crack width and spacing decrease during warm seasons. This happens because soil cracking creates preferential pathways for moisture and heat transfer, increasing evaporation during warm periods and infiltration during rainy seasons. As a result, the cracks not only accelerate ground deformation but also, due to their varying and random distances from one another, cause heterogeneous deformation, leading to differential deformations and damage to both surface and underground structures.

Comparing the results across all three scenarios reveals that as crack width increases, the plane shifts upwards, leading to lower subsidence. This suggests that soil with narrower cracks is more susceptible to subsidence under similar climatic conditions, while soil with wider cracks is more prone to swelling. In addition to the vertical movement of the plane within the three-dimensional subsidence-crack width-crack spacing framework, the angle of the plane also changes, indicating a steeper slope when the crack ratio is at 5%. This implies that soil with deep, narrow cracks is more sensitive to variations in crack width and spacing. As the initial crack width increases, the sensitivity of the cracked soil to geometric changes decreases, resulting in a lower deformation rate.

To contextualise the present probabilistic findings, Table 3 contrasts key outcomes with recent laboratory, field-monitoring, and modelling studies on desiccation cracking under climatic forcing. The comparison highlights consistent evidence for (i) cyclic opening-closure of crack networks under hydro-thermal variability, (ii) the strong role of surface energy and evaporative demand in sustaining crack development, and (iii) the feedback of crack geometry on preferential flow and moisture redistribution. This synthesis clarifies the domain of applicability of the proposed framework and demonstrates qualitative consistency with reported mechanisms across multiple settings.

To synthesise the probabilistic findings with physical processes, Fig. 13 illustrates the mechanistic feedback loop governing the soil-atmosphere interaction. The cycle begins with atmospheric forcing, where thermodynamic drivers (mainly temperature, radiation) and hydraulic drivers (precipitation, relative humidity) act on the soil surface. These inputs regulate the interface mechanisms: during drying, intermediate cracks (C_R13) maximise the evaporative flux by permitting deep radiative penetration and turbulent airflow, while during wetting, wide cracks (C_R35) act as preferential flow paths for rapid deep-soil saturation. This forcing alters the soil response state variables—specifically suction and effective stress—leading to volumetric deformation (subsidence or swelling). Crucially, this deformation creates a feedback loop: as the soil shrinks, cracks widen, further enhancing the evaporative surface area and infiltration capacity; conversely, swelling closes cracks, self-limiting hydraulic conductivity. This cyclic interaction explains the study's observation that C_R13 is governed by energy drivers due to active evaporation, whereas C_R35 is governed by hydraulic drivers due to rapid re-saturation.

4. Conclusions

The reliability of crack width evolution and the associated risks of soil subsidence and swelling during cyclic climatic conditions was investigated in this study using probabilistic modelling techniques,

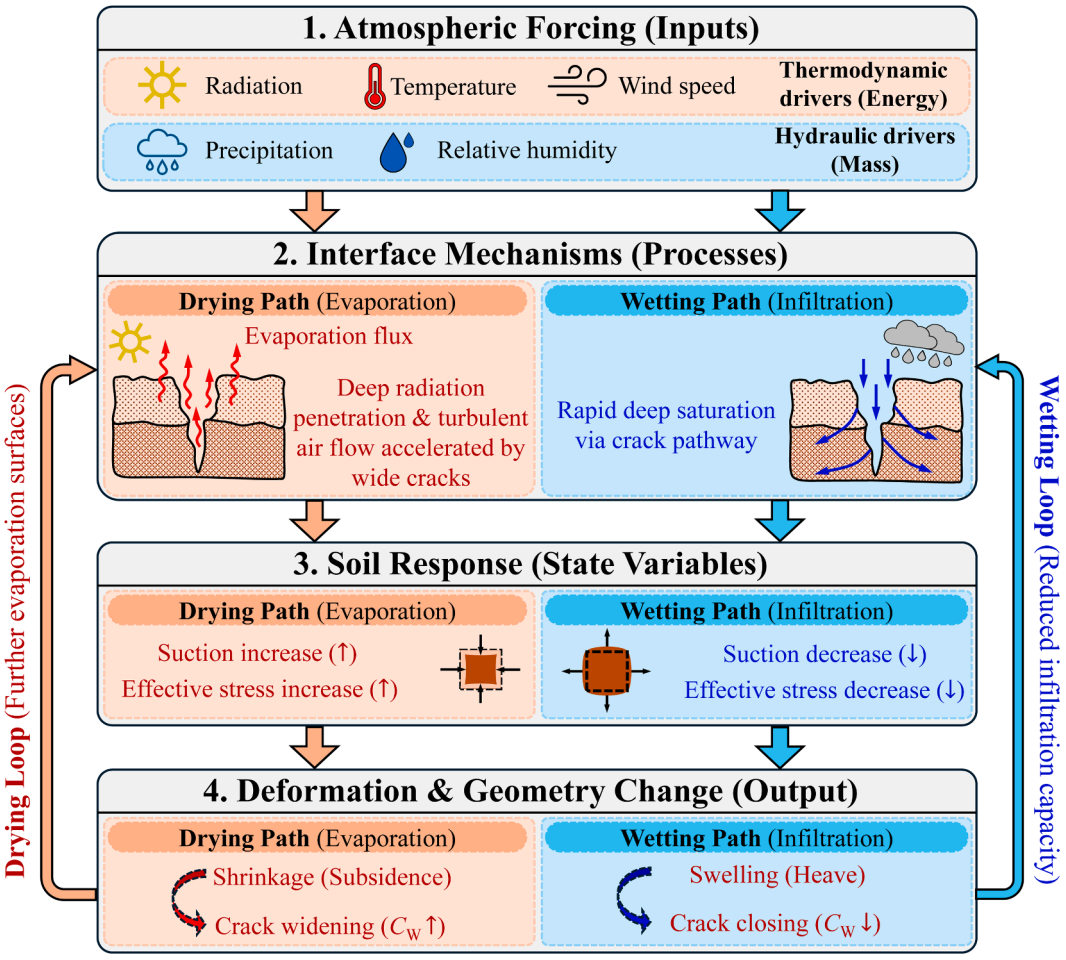


Fig. 13. Schematic layout of mechanistic framework of explaining the soil-atmosphere interaction loop. Atmospheric forcing drives changes in soil suction through evaporation and infiltration processes. These suction changes induce volumetric deformation (subsidence or swelling), which alters the crack geometry.

including Monte Carlo sampling, the FOSM method, and ML tools. The key findings of this study can be summarised as follows:

- On the basis of MAE, the linear regression was identified as the most suitable model for cracks C_R5, C_R13, and C_R35, with an MAE of 4.77, outperforming other regression models. The model demonstrated a strong linear relationship between crack width and environmental parameters, as indicated by favorable R^2 and RMSE values. The model C_R5 demonstrates the accuracy of the results, with an R^2 of 0.89 and an RMSE of 5.43.
- The probabilistic cumulative distribution of crack widths, obtained through Monte Carlo sampling and linear regression, demonstrated a strong correlation with numerical modelling. Both methods indicated similar mean values for C_R13, with numerical modelling yielding 153.65 mm and probabilistic modelling showing 153.62 mm, while their variances differed slightly. C_R13 exhibited the greatest potential for variation in crack width at a 50% probability. In contrast, C_R5 displayed greater reliability across different crack width changes, suggesting less variation and enhanced stability.
- The FOSM method proved effective in calculating the mean and variance of the cumulative crack width distributions and identifying the most influential parameters, with relative humidity emerging as the dominant factor across all three scenarios, characterised by a FOSM importance factor close to 1. Precipitation ranked second in significance, followed by radiation, temperature, and wind speed. This is due to the high potential of expansive soil for swelling even with a little water infiltration.
- The risk analysis of soil subsidence and swelling, driven by crack width variations, highlighted the dominance of subsidence over swelling in long-term soil behaviour. Subsidence was found to be far more prevalent, with overall probabilities of 92% for subsidence and only 8% for swelling. However, the analysis also revealed the significance of swelling as a contributing factor to ground deformation. C_R5, C_R13, and C_R35 exhibited subsidence probabilities of 98%, 96%, and 80%, respectively.
- This study highlights the critical shift from traditional deterministic approaches, which often overlook inherent uncertainties, toward probabilistic frameworks in engineering design. Unlike conventional deterministic methods that perform binary assessments against a fixed threshold (ignoring the variability of input parameters), the reliability index (β) provides a quantitative measure of safety. For instance, considering a limiting threshold such as a crack width of 30 mm, β does not merely indicate compliance but quantifies the standard deviation distance between the mean performance and the failure limit. A higher β value directly correlates with a lower probability of failure, thereby confirming the suitability of the soil or infrastructure conditions with greater confidence. Consequently, instead of relying solely on deterministic verification, the methodology adopted in this research allows for controlling the probability of failure. This approach enables the definition of appropriate target reliability levels for various thresholds, ensuring that the risk of exceeding critical limits is statistically managed.
- Finally, a linear relationship was proposed between subsidence and swelling magnitudes, crack width, and the distance between successive cracks. A plane defined by the initial widths of the three

cracks corresponds with the risk analysis results from the Monte Carlo method, showing that the probability of subsidence decreases as the initial crack width increases.

While the proposed models demonstrate promising performance, several limitations should be acknowledged to contextualise the findings as follows:

- To reduce computational complexity, the study simplified the interactions between influential factors, treating input parameters as independent variables. Future iterations could employ probabilistic graphical models to account for the conditional dependencies between parameters.
- The data utilised is derived from a specific geographical location with a distinct climate. As a result, the generalisation of the proposed models to other climatic zones or geological settings is limited.
- The machine learning models used in this study perform interpolation well but are restricted to the distribution of the training labels. While Deep Learning (DL) architectures could theoretically enhance generalisation capabilities, they were excluded from this study. Given the limited size of the available dataset, the application of DL would have introduced a high risk of overfitting, potentially degrading the model's performance on unseen data.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Ali Mohammadi Kamizji: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Hamed Sadeghi:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Milad Jabbarzadeh:** Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Chen, W. Lindqwister, F. Wu, B. Mielniczuk, T. Hueckel, M. Veveakis, The physics of desiccation cracks 2: modeling and prediction of the crack patterns, Geomech. Energy Environ. 35 (2023) 100489, https://doi.org/10.1016/j.gete.2023.100489.

[2] J.A. Cordero, P.C. Prat, A. Ledesma, Experimental analysis of desiccation cracks on a clayey silt from a large-scale test in natural conditions, Eng. Geol. 292 (2021) 106256, https://doi.org/10.1016/j.enggeo.2021.106256.

[3] T. Wang, C.S. Tang, L. Lin, T.T. Chen, Z.L. Cai, Soil desiccation cracking under a non-uniform temperature field: experimental investigation and DEM modeling, J. Rock Mech. Geotech. Eng. (2025), https://doi.org/10.1016/j.jrmge.2025.08.015.

[4] M. Jabbarzadeh, H. Sadeghi, S. Tourchi, A.G. Darzi, Thermo-hydraulic analysis of desiccation cracked soil strata considering ground temperature and moisture dynamics under the influence of soil-atmosphere interactions, Geomech. Energy Environ. 38 (2024) 100558, https://doi.org/10.1016/j.gete.2024.100558.

[5] C.S. Tang, Q. Cheng, T. Leng, B. Shi, H. Zeng, H.I. Inyang, Effects of wetting-drying cycles and desiccation cracks on mechanical behavior of an unsaturated soil, Catena 194 (2020) 104721, https://doi.org/10.1016/j.catena.2020.104721.

[6] H. Sadeghi, M. Jabbarzadeh, S. Tourchi, Thermo-hydro-mechanical modelling of the heterogeneous subsidence and swelling in the desiccation cracked clayey strata, Eng. Geol. 343 (2024) 107798, https://doi.org/10.1016/j.enggeo.2024.107798.

[7] J.T. Christian, G.B. Baecher, Unresolved problems in geotechnical risk and reliability. Geo-Risk 2011: Risk Assessment and Management, 2011, pp. 50–63, https://doi.org/10.1061/41183(418)3.

[8] K.K. Phoon, J. Ching (Eds.), Risk and reliability in geotechnical engineering, Risk and reliability in geotechnical engineering, 651, CRC Press, Boca Raton, FL, USA, 2015.

[9] Z. Feng, L. Wang, Q. Peng, J. Li, T. Liang, Effect of environmental factors on soil properties under different land use types in a typical basin of the North China Plain, J. Clean. Prod. 344 (2022) 131084, https://doi.org/10.1016/j.jclepro.2022.131084.

[10] S. Tourchi, M. Jabbarzadeh, A.A. Lavasan, H. Sadeghi, O. Racek, Thermo-hydro-mechanical dynamics of a rock slope: integrated field and numerical analysis at the Požáry test site in the Czech Republic, J. Rock Mech. Geotech. Eng. (2024), https://doi.org/10.1016/j.jrmge.2024.09.052.

[11] R. Bea, Reliability and human factors in geotechnical engineering, J. Geotech. Geoenvironmental Eng. 132 (5) (2006) 631–643, https://doi.org/10.1061/(ASCE)1090-0241(2006)132:5(631).

[12] J. Cao, T. Wang, G. Zhou, X. Feng, C. Zhu, Parameter estimation of grouting pressure and surface subsidence on the reliability of shield tunnel excavation under incomplete probability information, Comput. Geotech. 173 (2024) 106530, https://doi.org/10.1016/j.compgeo.2024.106530.

[13] Q.F. Gao, X.Y. Wu, L. Zeng, H.C. Yu, X.K. Shi, Reliability analysis of soil slope stability considering spatial variability of desiccation cracks, Comput. Geotech. 179 (2025) 106977, https://doi.org/10.1016/j.compgeo.2024.106977.

[14] M. Pan, S.H. Jiang, X. Liu, G.Q. Song, J. Huang, Sequential probabilistic back analyses of spatially varying soil parameters and slope reliability prediction under rainfall, Eng. Geol. 328 (2024) 107372, https://doi.org/10.1016/j.enggeo.2023.107372.

[15] J.T. Ritchie, J.E. Adams, Field measurement of evaporation from soil shrinkage cracks, Soil Sci. Soc. Am. J. 38 (1) (1974) 131–134, https://doi.org/10.2136/sssaj1974.03615995003800010040x.

[16] W.K. Song, Y.J. Cui, Modelling of water evaporation from cracked clayey soil, Eng. geol. 266 (2020) 105465, https://doi.org/10.1016/j.enggeo.2019.105465.

[17] M.J. Ghandilou, S. Tourchi, H. Sadeghi, Numerical investigation of cyclic wetting and drying of boom clay based on the Barcelona Expansive Model, in: 84th EAGE Annual Conference & Exhibition, 2023, pp. 1–5, https://doi.org/10.3997/2214-4609.202310400.

[18] H. Zeng, C.S. Tang, A. Fraccica, C. Zhu, B.G. Tian, B. Shi, Multi-scale investigation on dynamic characteristics of clayey soil evaporation and cracking, Comput. Geotech. 171 (2024) 106312, https://doi.org/10.1016/j.compgeo.2024.106312.

[19] M. Jabbarzadeh, H. Sadeghi, Multi-physical modeling of climate-driven elasto-plastic deformation, stress redistribution, and water potential in desiccation-cracked soils of arid regions, J. Rock Mech. Geotech. Eng. (2025), https://doi.org/10.1016/j.jrmge.2025.05.018.

[20] L. Zhao, P. Ma, Q. Mu, Z. Jia, J. Zhao, Z. He, B. Huo, Z. Li, J. Peng, Influences of freeze-thaw cycles on desiccation cracking of an intact loess, Eng. Geol. 344 (2025) 107864, https://doi.org/10.1016/j.enggeo.2024.107864.

[21] M. Jabbarzadeh, Z. Mousavi, Coupled multi-physics modeling of spatio-temporal behavior of embankment subjected to freeze-thaw cycles considering layered construction phases, Cold Reg. Sci. Technol. (2025) 104652, https://doi.org/10.1016/j.coldregions.2025.104652.

[22] H. Sadeghi, A.G. Darzi, A. Mirpanji, A.A. Garakani, A hybrid saturated-unsaturated framework for modeling land subsidence under the combined effects of water exploitation and environmental conditions, Eng. Geol. (2025) 108322, https://doi.org/10.1016/j.enggeo.2025.108322.

[23] F. Yazdani, H. Sadeghi, P. AliPanahi, M. Gholami, A.K. Leung, Evaluation of plant growth and spacing effects on bioengineered slopes subjected to rainfall, Biogeotechnics 2 (2) (2024) 100080, https://doi.org/10.1016/j.bgtech.2024.100080.

[24] J.E. Adams, R.J. Hanks, Evaporation from soil shrinkage cracks, Soil Sci. Soc. Am. J. 28 (2) (1964) 281–284, https://doi.org/10.2136/sssaj1964.03615995002800020043x.

[25] T.G. Poulsen, Predicting evaporation from moist, cracked soil, based on near-surface wind speed, crack width and crack distance, Eur. J. Soil Sci. 73 (1) (2022) e13215, https://doi.org/10.1111/ejss.13215.

[26] J.H. Li, L.M. Zhang, Study of desiccation crack initiation and development at ground surface, Eng. Geol. 123 (4) (2011) 347–358, https://doi.org/10.1016/j.enggeo.2011.09.015.

[27] Q. Yao, T. Chen, C. Tang, M. Sedighi, S. Wang, Q. Huang, Influence of moisture on crack propagation in coal and its failure modes, Eng. Geol. 258 (2019) 105156, https://doi.org/10.1016/j.enggeo.2019.105156.

[28] H. Zeng, C.S. Tang, C. Zhu, F. Vahedifard, Q. Cheng, B. Shi, Desiccation cracking of soil subjected to different environmental relative humidity conditions, Eng. Geol. 297 (2022) 106536, https://doi.org/10.1016/j.enggeo.2022.106536.

[29] A. Kumar, A. Azizi, D.G. Toll, Deterioration of a compacted soil due to suction loss and desiccation cracking, Can. Geotech. J. (2024), https://doi.org/10.1139/cgj-2023-0453.

[30] H. Gao, R. An, X. Zhang, G. Wang, X. Liu, Y. Xu, Dynamic evolution of desiccation cracks and their relationship with the hydraulic properties of expansive soil, Int. J. Geomech. 24 (3) (2024) 04023299, https://doi.org/10.1061/IJGNAI.GMENG-9299.

[31] M. Jabbarzadeh, H. Sadeghi, Exploring the inter-relationship between climatic conditions and heat-water transfer in deformable cracked surface strata, InterPore J. (2026), https://doi.org/10.69631/rk5nyd57.

[32] J.J. Xu, C.S. Tang, Y. Yang, L. Li, H. Zhang, Q. Cheng, X.Y. Zhang, B. Liu, B. Shi, Breathing phenomenon of soil desiccation cracking: insights from novel geophysical observations, J. Geophys. Res.: Earth Surf. 129 (1) (2024) e2023JF007318, https://doi.org/10.1029/2023JF007318.

[33] J.A. Hanson, B.O. Hardin, K. Mahboub, Fracture toughness of compacted cohesive soils using ring test, J. geotech. eng. 120 (5) (1994) 872–891, https://doi.org/10.1061/(ASCE)0733-9410(1994)120:5(872).
[34] B. Jamhiri, Y. Xu, M. Shadabfar, S. Costa, Probabilistic machine learning for predicting desiccation cracks in clayey soils, Bull. Eng. Geol. Environ. 82 (9) (2023) 355, https://doi.org/10.1007/s10064-023-03366-2.
[35] D.J. Armaghani, M. Koopialipoor, M. Bahri, M. Hasanipanah, M.M. Tahir, A SVR-GWO technique to minimize flyrock distance resulting from blasting, Bull. Eng. Geol. Environ. 79 (8) (2020) 4369–4385, https://doi.org/10.1007/s10064-020-01834-7.
[36] B.T. Pham, M.D. Nguyen, D. Van Dao, I. Prakash, H.B. Ly, T.T. Le, L.S. Ho, K. T. Nguyen, T.Q. Ngo, V. Hoang, H.T.T. Ngo, Development of artificial intelligence models for the prediction of Compression Coefficient of soil: an application of Monte Carlo sensitivity analysis, Sci. Total Environ. 679 (2019) 172–184, https://doi.org/10.1016/j.scitotenv.2019.05.061.
[37] Y. Wang, X. Gao, P. Jiang, X. Guo, R. Wang, Z. Guan, L. Chen, C. Xu, An extreme gradient boosting technique to estimate TBM penetration rate and prediction platform, Bull. Eng. Geol. Environ. 81 (2022) 1–19, https://doi.org/10.1007/s10064-021-02527-5.
[38] X. Wang, Y. Pan, J.J. Chen, Digital twin with uncertainty-informed deep learning for prompt quantitative risk assessment of deep excavation, Comput.-Aided Civ. Infrastruct. Eng. 40 (25) (2025) 4226–4252, https://doi.org/10.1111/mice.70049.
[39] L.Y. Ju, T. Xiao, J. He, W.F. Xu, S.H. Xiao, L.M. Zhang, A simulation-enabled slope digital twin for real-time assessment of rain-induced landslides, Eng. Geol. (2025) 108116, https://doi.org/10.1016/j.enggeo.2025.108116.
[40] M. Jabbarzadeh, H. Sadeghi, A novel analytical-statistical-numerical hybrid framework for investigating the thermo-hydro-mechanical behavior of heterogeneous subsidence phenomena, moisture and temperature variations in cracked soils under soil-atmosphere interaction, Sharif J. Mech. Eng. (2024), https://doi.org/10.24200/j40.2024.65047.1718.
[41] A.A. Garakani, S.T. Banafshehvaragh, S. Saheb, H. Sadeghi, Assessing the vulnerability of power transmission towers to land subsidence and forecasting future trends using multi-source datasets: insights from Moein Abad, Iran. Innov. Infrastruct. Solut. 10 (4) (2025) 1–23, https://doi.org/10.1007/s41062-025-01958-3.
[42] H. Sadeghi, A.G. Darzi, B. Voosoghi, A.A. Garakani, Z. Ghorbani, S.F.F. Mojtahedi, Assessing the vulnerability of Iran to subsidence hazard using a hierarchical FUCOM-GIS framework, Remote Sens. Appl.: Soc. Environ. 31 (2023) 100989, https://doi.org/10.1016/j.rsase.2023.100989.
[43] M. Mahsuli, T. Haukaas, Computer program for multimodel reliability and optimization analysis, J. comput. civ. eng. 27 (1) (2013) 87–98, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000204.
[44] H. Nasrazadani, M. Mahsuli, Probabilistic framework for evaluating community resilience: integration of risk models and agent-based simulation, J. Struct. Eng. 146 (11) (2020) 04020250, https://doi.org/10.1061/(ASCE)ST.1943-541X.0002810.
[45] J.C. Chai, S.L. Shen, H.H. Zhu, X.L. Zhang, Land subsidence due to groundwater drawdown in Shanghai, Geotechnique 54 (2) (2004) 143–147, https://doi.org/10.1680/geot.2004.54.2.143.
[46] S. Tourchi, M. Jabbarzadeh, H. Sadeghi, Thermo-hydro-mechanical analysis of soil strata suffered from desiccation cracking, Geotechnical Engineering Challenges to Meet Current and Emerging Needs of Society, CRC Press, pp. 1362–1367. https://doi.org/10.1201/9781003431749-249.