

A MACHINE LEARNING BASED SURROGATE MODEL FOR PREDICTING THE INFLUENCE OF ENVIRONMENTAL CONDITIONS ON DESICCATION CRACK WIDTH

Ali Mohammadi Kamizji

Department of Civil Engineering, Sharif University of Technology, Iran. E-mail: al.mohammad@sharif.edu

Milad Jabbarzadeh

Department of Civil Engineering, Sharif University of Technology, Iran. E-mail: milad.jabbarzadeh@sharif.edu

Hamed Sadeghi

Department of Civil Engineering, Sharif University of Technology, Iran. E-mail: hsadeghi@sharif.edu

Habibollah Sadeghi

Norwegian Geotechnical Institute, Norway. E-mail: habibollah.sadeghi@ngi.no

Climate change is of significant interest in various fields due to its far-reaching impacts. One of the notable consequences of environmental change is the increase in the frequency and severity of soil desiccation cracking, which can lead to a range of serviceability limit state problems. Therefore, the influence of environmental factors on variations in the crack width is explored in this study through machine learning techniques. The research focuses on five primary environmental parameters including temperature, wind speed, radiation, relative humidity, and precipitation. These variables were meticulously recorded over a year, from 2015 to 2016, by local meteorological stations to capture the temporal variations. Various regression methods were explored, including Linear, Bayesian, Polynomial, Dummy, and Gaussian regressions. The dataset, comprising environmental parameters and corresponding crack widths, was randomized to prevent bias. A split of 70% for training and 30% for testing was adopted to validate the model performance. The Mean Absolute Error (MAE) is employed as the metric to evaluate the accuracy of the regression models. Subsequently, a numerical thermo-hydro-mechanical modeling approach was utilized to compute the variability of crack width. The results obtained from the machine learning methods were then compared against the findings from a comprehensive multi-physical modeling. According to the results, linear regression emerged as the most suitable model among other regression techniques, demonstrating the lowest MAE. The findings also reveal that the predictive model demonstrates an excellent match with the numerical simulation results. This finding implies the high dependency of the variations in crack width on the environmental parameters considered. Lastly, a new formulation was derived and proposed that can be used to estimate the crack width under specific environmental conditions.

Keywords: Machine learning, Desiccation crack width, Linear regression, Climate change impact, Geohazard.

1. Introduction

Soil desiccation cracking is a prevalent phenomenon in geotechnical engineering that significantly influences soil deformation processes, especially under fluctuating environmental conditions and natural cycles of wetting and drying (Chen et al. 2023; Ghandilou et al. 2023; Valipour et al. 2024; Tourchi et al. 2024a). The formation of cracks in highly plastic desiccated soils disrupts the uniformity of the soil structure, resulting in considerable heterogeneity in its behavior. These cracks arise when tensile stresses surpass the soil's tensile strength, leading to alterations in both the mechanical and hydraulic properties of the soil (Tang et al. 2020; Gao et al. 2024; Tourchi et al. 2024b; Wei et al. 2024; Valipour et al. 2024b; Mousavi and Jabbarzadeh, 2025). Although desiccation cracks are a natural phenomenon, their irregular distribution across the soil surface generates non-uniform patterns of subsidence and swelling during environmental fluctuations, which present significant challenges in geotechnical contexts (Jabbarzadeh and Sadeghi 2024; Xu et al. 2024; Tourchi et al. 2024c). In addition to the soil surface, the crack itself is exposed to variations in its geometric characteristics under extreme environmental conditions as occurred in parts of Iran plain. Consequently, understanding the relationship between crack geometric factors and environmental parameters is essential for predicting potential subsidence and swelling behaviors in desiccation-cracked soils (Sadeghi et al. 2024). Machine learning (ML) is a reasonable approach to make this correlation since it offers a powerful means to uncover complex patterns that may not be readily apparent through conventional analysis, allowing for more accurate predictions of soil

behavior under varying conditions. Therefore, this research aims to establish a correlation between environmental parameters such as temperature, precipitation, wind speed, relative humidity, and radiation with crack width. The developed model was validated against the recent findings of Sadeghi et al. (2024).

2. Materials and Methods

2.1. Numerical Model and Soil-atmosphere Interaction Problem

To determine the factors that affect fluctuations in crack width under real environmental conditions, thermo-hydro-mechanical numerical models were developed to simulate the interaction between cracked soil and the atmosphere (Sadeghi et al., 2024). Meteorological data for Qom city in Iran was collected, as this area is experiencing land subsidence due to its arid climate and extensive deep desiccation cracks (Sadeghi et al., 2023). The weather dataset, reported by Jabbarzadeh et al. (2024), includes temperature, wind speed, relative humidity, precipitation, and radiation. The results from the numerical model are utilized to develop a predictive model for variations in desiccation crack width in response to changes in environmental conditions. In the numerical model, a dimensionless parameter known as the crack ratio (C_R), defined as the ratio of crack width to crack depth, is utilized. Three scenarios for the C_R were considered: 5% (C_{R5}), 13% (C_{R13}), and 35% (C_{R35}). It is also assumed that the crack depth is 1 m, based on field observations from the Qom plain. These scenarios are also employed in this study to validate the results of the ML-based regression analysis.

2.2. Machine Learning Approaches for Regression Analysis

To predict variations in crack width due to changes in climatic parameters, supervised machine learning algorithms were utilized, including Linear Regression (LR), Bayesian Regression (BR), Polynomial Regression (PR), Dummy Regression (DR), and Gaussian Process Regression (GPR). Considering the initial uncertainty about whether the relationship between the target data (soil crack width) and the climatic parameters was linear or non-linear, these algorithms were selected to investigate various potential relationships. To train the models, the data encompassing both environmental variables and crack width was randomized, with 70% allocated for training and 30% for testing. The Mean Absolute Error (MAE) metric was used to evaluate and compare model performance in capturing crack width relationships, defined as:

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

Fig.1 presents each regression method alongside the MAE when compared with numerical modeling data. Based on the MAE results, LR emerged as the most accurate model with the lowest error rate. Consequently, a LR model was utilized to establish a relationship between climatic parameters and crack width.

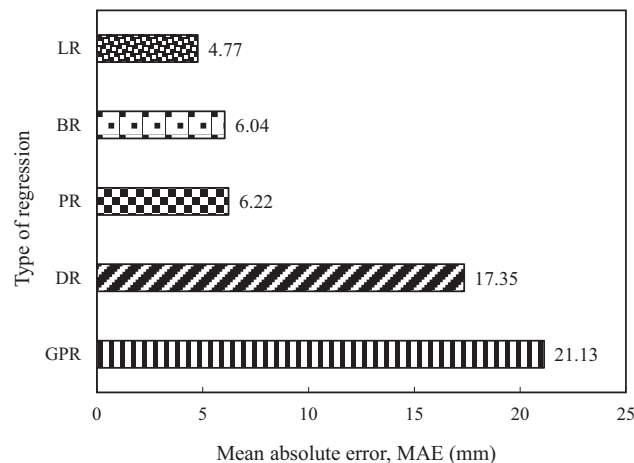


Fig. 1. Assessment of the five regression techniques adopted in this study based on the MAE metric.

3. Interpretation of the Results

Based on LR method, the variations in crack width with environmental conditions can be estimated with the following general function:

$$C_W = A + B(T) + C(W) + D(P) + E(R) + F(H) \quad (2)$$

where C_w is the crack width, $T(^{\circ}\text{C})$ is the temperature, $W(\text{mm/s})$ is the wind speed, $P(\text{cm/day})$ is the precipitation, $R(\text{J}/(\text{m}^2.\text{s}))$ is the radiation, and $H(\%)$ is the relative humidity. In addition, A, B, C, D, E, and F are the coefficients of model parameters that can be obtained by ML model. These constants were summarized in Table 1 for each scenario.

Table 1. The coefficients of the environmental parameters.

Parameter	Modeling scenario		
	C_{R5}	C_{R13}	C_{R35}
A	100.5	205.3	400.6
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

To further improve the proposed model, a relationship between all three scenarios was made; hence a unique equation correlating both environmental factors and initial crack width was defined as:

$$C_w = 58.2411 - 0.0644(T) - 1.8516 \times 10^{-10}(W) - 98.8417(P) - 0.2899(R) - 1.0269(H) + 1.0033(C_{wi}) \quad (3)$$

Using this equation, the changes in crack aperture due to environmental changes can be predicted using meteorological data and the initial aperture of the crack (C_{wi}). Fig. 2 compares the variations in crack width for both the numerical model and the ML-based model, indicating an excellent match.

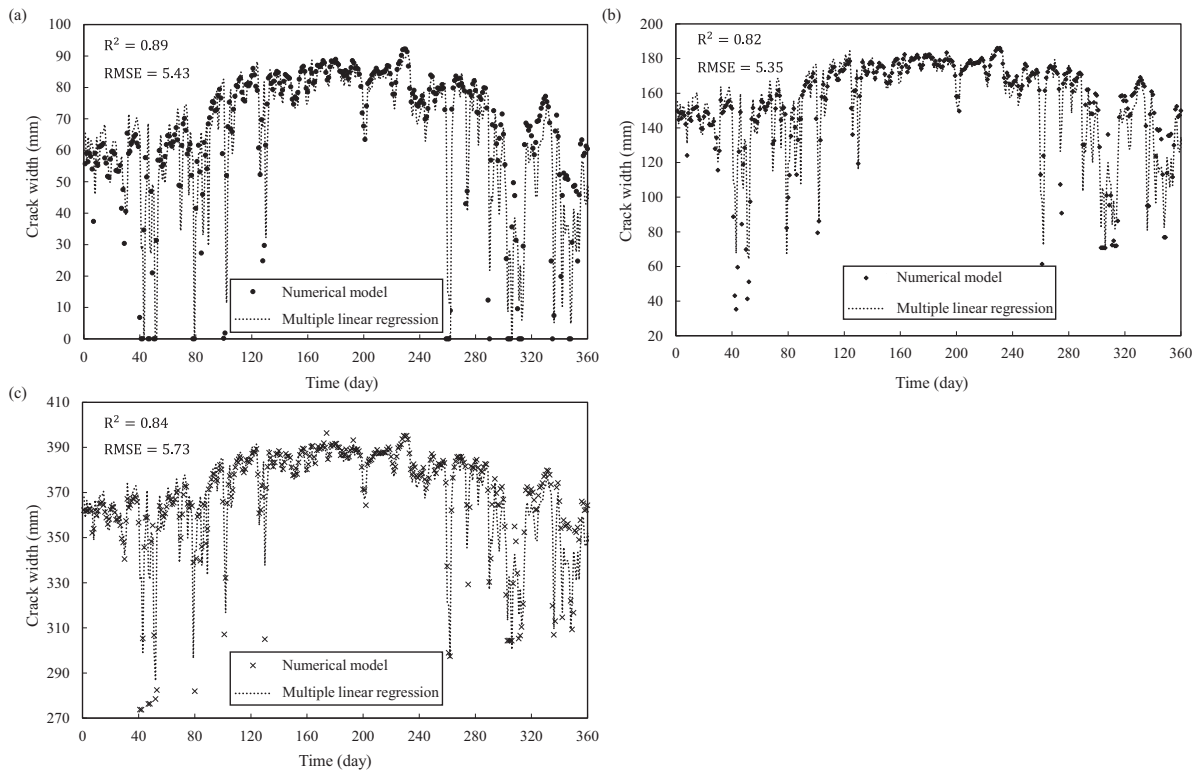


Fig. 2. Comparison between numerical modeling results and linear regression model for (a) C_{R5} , (b) C_{R13} , and (c) C_{R35}

As shown in Fig. 2, the width of the crack is continuously changing due to changes in environmental conditions. This fluctuation leads to alterations in the existing crack geometry, which in turn intensifies the deformation of the soil surface. The results presented in Fig. 2 also indicate that a ML-based model can predict the seasonal changes in crack width under environmental change, similar to the complex numerical model. Consequently, Eq. (3) serves as a simpler alternative for this purpose. While the linear regression model offers a computationally efficient and data-lean approach for predicting desiccation crack width, it is best viewed as a first screening tool, capable of quickly identifying areas where climate change suggests a high risk of desiccation cracking. Subsequent, more detailed investigations employing comprehensive thermo-hydro-

mechanical numerical modeling, are recommended for a more in-depth understanding and precise prediction of soil behavior in these high-risk zones. Moreover, the developed model is applicable to arid climates characterized by soils with a high susceptibility to desiccation cracking.

4. Conclusions

This study provides a machine learning-driven linear regression model to establish a correlation between changes in desiccation crack width and environmental parameters. The model includes the significant impact of climatic factors such as temperature, wind speed, precipitation, radiation, and relative humidity on crack width in highly plastic desiccated soils. By employing different machine learning techniques, it is found that linear regression method has the lowest Mean Absolute Error (MAE). The predictive model was then validated by numerical modeling results. The model demonstrated a strong fit with numerical data, as evidenced by low MAE and high R^2 values. Further investigations into the geometric characteristics of cracks and their interaction with environmental changes could provide deeper insights into subsidence and swelling behaviors in desiccated soils.

Acknowledgments

The financial support provided by the Research Grant Office at Sharif University Technology by way of grants G4010902 and QB020105 is gratefully acknowledged.

References

- Chen, R., Lindqwister, W., Wu, F., Mielniczuk, B., Hueckel, T. and Veveakis, M., (2023). The physics of desiccation cracks 2: Modeling and prediction of the crack patterns. *Geomechanics for Energy and the Environment*, 35, p.100489.
- Gao, H., An, R., Zhang, X., Wang, G., Liu, X. and Xu, Y. (2024). Dynamic evolution of desiccation cracks and their relationship with the hydraulic properties of expansive soil. *International Journal of Geomechanics*, 24(3), p.04023299.
- Ghandilou, M.J., Tourchi, S. and Sadeghi, H. (2023). Numerical investigation of cyclic wetting and drying of Boom clay based on the Barcelona Expansive Model. In 84th EAGE Annual Conference & Exhibition, Vol. 2023, No. 1, pp. 1-5. European Association of Geoscientists & Engineers.
- Jabbarzadeh, M. and Sadeghi, H. (2024). A novel analytical-statistical-numerical framework for investigating the thermo-hydro-mechanical behavior of heterogeneous subsidence phenomena, moisture and temperature variations in cracked soils under soil-atmosphere interaction. *Sharif Journal of Mechanical Engineering*.
- Jabbarzadeh, M., Sadeghi, H., Tourchi, S. and Darzi, A.G., (2024). Thermo-hydraulic analysis of desiccation cracked soil strata considering ground temperature and moisture dynamics under the influence of soil-atmosphere interactions. *Geomechanics for Energy and the Environment*, 38, p.100558.
- Mousavi, Z. and Jabbarzadeh, M., 2025. 3D DEM-based analysis of cylindrical rock specimen failure and micro-fracturing: impact of stiffness and tensile strength. *Discover Geoscience*, 3(1), pp.1-21.
- Sadeghi, H., Darzi, A. G., Voosoghi, B., Garakani, A. A., Ghorbani, Z. and Mojtahedi, S. F. F. (2023). Assessing the vulnerability of Iran to subsidence hazard using a hierarchical FUCOM-GIS framework. *Remote Sensing Applications: Society and Environment*, 31, 100989.
- Sadeghi, H., Jabbarzadeh, M. and Tourchi, S., (2024). Thermo-hydro-mechanical modelling of the heterogeneous subsidence and swelling in the desiccation cracked clayey strata. *Engineering Geology*. p.107798.
- Tang, C.S., Cheng, Q., Leng, T., Shi, B., Zeng, H. and Inyang, H.I. (2020). Effects of wetting-drying cycles and desiccation cracks on mechanical behavior of an unsaturated soil. *Catena*, 194, p.104721.
- Tourchi, S., Jabbarzadeh, M. and Sadeghi, H., (2024). Thermo-hydro-mechanical analysis of soil strata suffered from desiccation cracking. In *Geotechnical Engineering Challenges to Meet Current and Emerging Needs of Society* (pp. 1362-1367). CRC Press.
- Tourchi, S., Jabbarzadeh, M., Lavasan, A., Sadeghi, H. and Racek, O., (2024). Thermo-Hydro-Mechanical Dynamics of a Rock Slope: Integrated Field and Numerical Analysis. *Journal of Rock Mechanics and Geotechnical Engineering*.
- Tourchi, S., Jabbarzadeh, M., Sadeghi, H. and Lavasan, A.A. (2024). Thermo-hydro-mechanical modelling of rock cliff-atmosphere interaction: the case of the Pozary test site in Czechia. In No. EGU24-20407), *Copernicus Meetings*.
- Valipour, M., Jabbarzadeh, M. and Sadeghi, H. (2024). The influence of different hydraulic conductivities on two-phase flow characteristics of gas-injected sand-bentonite soil considering embedded fractures. In 14th National Congress on Civil Engineering.
- Valipour, M., Jabbarzadeh, M., Sadeghi, H. and Roghangar, K., (2024). Volumetric behavior of a soil column subjected to freeze-thaw cycles under coupled thermo-hydro-mechanical conditions. *GeoMontréal 2024*, 1-8.
- Wei, Y., Tang, C.S., Lu, Y., Zhang, M., Cheng, Q., Li, L. and Shi, B. (2024). Evaporation and desiccation cracking of soils: Experiment evidence and insight on the freeze-thaw cycle dependence. *Engineering Geology*, 329, p.107399.
- Xu, J.J., Tang, C.S., Yang, Y., Li, L., Zhang, H., Cheng, Q., Zhang, X.Y., Liu, B. and Shi, B. (2024). Breathing phenomenon of soil desiccation cracking: Insights from novel geophysical observations. *Journal of Geophysical Research: Earth Surface*, 129(1), p.e2023JF007318.