Proceedings of The ASME 2023 Conference on Smart Materials, Adaptive Structures and Intelligent Systems
SMASIS2023
September 11-13, 2023, Austin, TX, USA

SMASIS2023-111009

CLASSIFYING SOIL SATURATION LEVELS USING A NETWORK OF UAV-DEPLOYED SMART PENETROMETERS

Puja Chowdhury
Department of
Mechanical Engineering
University of South Carolina
Columbia, South Carolina 29208
Email: pujac@email.sc.edu

Joud N. Satme
Department of
Mechanical Engineering
University of South Carolina
Columbia, South Carolina 29208
Email: jsatme@email.sc.edu

Ryan Yount
Department of
Mechanical Engineering
University of South Carolina
Columbia, South Carolina 29208
Email:RJYOUNT@email.sc.edu

Austin R.J. Downey
Department of
Mechanical Engineering
Department of Civil and
Environmental Engineering
University of South Carolina
Columbia, South Carolina 29208
Email: austindowney@sc.edu

Sadik Khan
Department of Civil and
Environmental Engineering
Jackson state University
Jackson,Mississippi
Email: sadik.khan@jsums.edu

Jasim Imran
Department of Civil and
Environmental Engineering
University of South Carolina
Columbia, South Carolina 29208
Email: IMRAN@sc.edu

Laura Micheli
Department of Civil and
Environmental Engineering
University of South Carolina
Columbia, South Carolina 29208
Email:LMICHELI@mailbox.sc.edu

ABSTRACT

Levees are built to safeguard human lives, essential infrastructure, and farmland. However, failure of levees can have catastrophic impacts due to a fast rate of inundation in areas protected by levees. Earthen levees are prone to failure due to excessive moisture content that reduces the shear strength of the soil. The use of levee monitoring systems has demonstrated the ability to reduce the likelihood of failure by creating maps that depict the saturation levels of the surface of the levee, both in terms of space and time. By utilizing extensive sensor networks to continuously monitor these geo-infrastructure systems, the structural deterioration attributed to changing climate can be studied. Measuring environmental parameters surrounding such structures provides insight into the potential stressors that cause structural failure. Steps can then be taken to mitigate those effects on the levees and maintain structural integrity. However, the massive scale of levees makes it difficult to monitor with conventional wired sensors. This paper presents a preliminary investigation into the development and validation of UAV-deployable smart sensing spikes for soil conductivity levels in levees, which is a measurement modality for determining soil saturation levels. For this work, Gaussian process regression (also known as kriging) is used to model the soil saturation levels between sensing spikes obtaining a continuous moisture map of the levees. The expanded data is then categorized using a clustering-based machine learning approach with conductivity data from sensing spikes as model inputs. The machine learning model output is sorted into three categories: dry, partially saturated, and saturated soil. The findings of a laboratory study are presented, and the implications of the raw and expanded data are discussed. This work will aid in predicting potential levee failure risks and maintenance requirements based on the analysis of the soil conditions using a network of smart sensing spikes.

INTRODUCTION

Infrastructure degradation and failure are influenced by extreme weather and natural catastrophes. Heavy rainfall and earthquakes are shown to have the greatest impact on water-holding infrastructure like dams and levees [1]. Sensors are installed to periodically monitor such structures, send out early alerts for evacuation, and implement preventative measures [2, 3]. Hard-wired large-footprint sensing is pervasively expensive and time-consuming since it requires specialized tools and trained personnel. The introduction of compact, drone-deployable sensors has made a rapid assessment of these structures possible and more cost-effective [4, 5].

When a levee's structural integrity is undermined by water, seepage, earthquakes, or other factors, levee failure occurs [6–8]. Overtopping, a kind of levee collapse brought on by flooding, occurs when the levee's waterside rises higher than the levee's crest, producing external erosion [9]. As the levee seeps, the water flushes out soil particles, creating interior channels, and sand boils. Internal erosion, also known as pipe erosion, is a phenomenon that weakens the structural integrity of levees over time. Subterranean dead roots and wild animals that dig under the levee are considered the main contributors to gaps that water seeps through which eventually leads to failure [7]. Figure 1 shows a detailed layout of a levee structure.

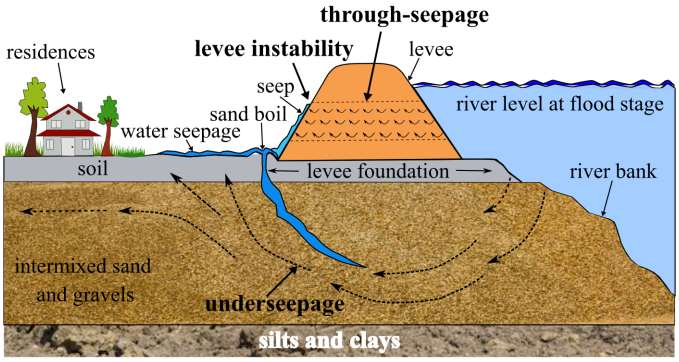

FIGURE 1. Cross-sectional model of an earthen levee with the key layers annotated along with possible failure points.

Information on soil quality is essential for site assessment, resource management, and planning of land use [10]. There are several techniques, including inverse distance weighting (IDW) and ordinary kriging (OK), for interpolating the geographical distribution of variables such as soil quality and moisture content [11] [12]. Regression kriging (RK), a technique that combines regression with spatial interpolation, has been proposed by researchers in recent years to map the distribution of soil attributes spatially [13]. Due to the simplicity of ordinary kriging and availability in many Geographic Information Systems (GIS), it has been frequently employed as an interpolation technique [14].

Finding the cluster structure in a data set with the most similarity within the same cluster and the highest dissimilarity across clusters can be accomplished via clustering [15]. The k-means algorithm is acknowledged as the most established and widely used partitioning technique [16]. Several cluster validity indices, such as the Akaike information criterion (AIC) [17], Davies-Bouldin index (DB) [18], extended Dunn's index (DNg) [19] are approaches presented in the literature for the k-means clustering technique.

As an initial investigation, this work presents the development and validation of UAV deployable smart sensor spikes for soil conductivity levels in levees, a method of detecting soil saturation levels. To obtain continuous conductivity maps of the surface soil in levees for this work, Gaussian process regression, also known as kriging, is utilized to estimate the mapping of soil saturation between sensing spikes. Using a clustering-based machine learning approach and conductivity data from sensing

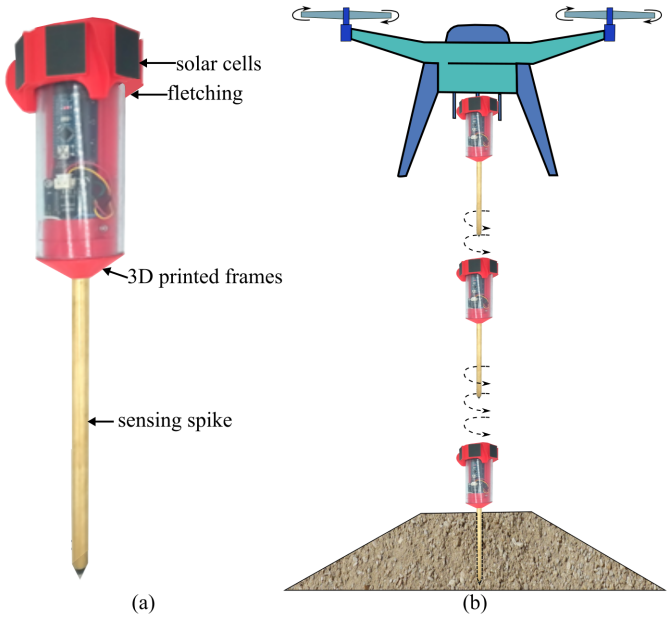

FIGURE 2. A schematic of a UAV deployed smart penetrometer, displaying: (a) the UAV deployable sensor package, and (b) a 2D view of the UAV sensor deployment process via the drone.

spikes as model inputs; the expanded data is then categorized. The output of the machine learning model is divided into three groups: dry, partially saturated, and saturated soil. The results of a lab experiment are reviewed, along with the implications of the raw and interpolated data. The investigation of the soil conditions utilizing a network of smart sensing spikes in this work can help identify levee seepage that has the potential to lead to maintenance or levee failures. The contributions of this work are: 1) The expandment of experimental data using kriging, and 2) The categorization of soil saturation using a network of smart sensing spikes.

HARDWARE DEVELOPMENT

The hardware development phase of this work consists of two steps: 1) developing the sensing spikes, and 2) experimental setup.

SENSING NODE: With the goal of measuring underground moisture while ensuring durability during UAV deployment, the sensing node was developed to be resilient yet lightweight to minimize payload. Field deployments of the sensing spike utilized a helical fetching design 3D printed in PLA, which protected the delicate electronics and provided rotational stability to ensure that the probe points downwards during free fall (see figure 2 (a) [20]. Solar cells were also added to aid in load sharing during sunny conditions, extending battery life. Figure 2 (b) presents a 2D perspective of the UAV sensor deployment procedure via the drone, which is one of the work's goals for the future. For the purpose of simplifying the network of sensing node configuration, this study focused solely on the conductivity measuring aspect of the node. The sensing node has an instrumented spike, depicted in figure 3. The sensing spike has conducting surfaces - an outer tube and an inner rod - separated by an insulating ABS plastic tube. This design allows for the integration of a conductivity module into the tip of the spike, enabling the spike to also function as an underground moisture probe. For the moisture test in this initial study, only a single sort of ion concentration, such as normal water, is used. During the investigations, future studies will evaluate the effect of different water quality and ion concentrations.

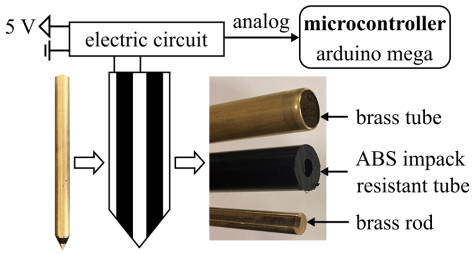

FIGURE 3. The sensing spike construction.

EXPERIMENTAL SETUP: In this work, a simplified configuration of just the sensing spikes is used to demonstrate the capability of the sensing node for monitoring and mapping spatial conductivity. Figure 4 (a) displays the experimental setup used to develop data for this work. For testing, a container of 10.8 x 12.4 inches is used. Sand is then filled to a height of 1.5 inches. Five sensing spikes are placed in the sand as shown in figure 4 (b).

A potential difference of 5 V was provided by a DC power supply to the spikes. Voltage is then measured via an independent analog to digital converters onboard the microcontroller simultaneously to be logged into memory. The coordinate of the five spikes is shown in table 1. A breakdown of the electrical circuit is depicted in figure 5. The spikes of the moisture-sensing network are set in parallel concerning power and ground (shown as F and G respectively). Each spike (R1-R5), modeled as a variable resistor in the schematic, is configured as an independent voltage divider using a constant 3.9 kΩ resistor. Potential points A through E are then measured using an analog to digital converter onboard a microcontroller as $V = [v_1, v_2, \cdots, v_5]$. In this configuration, the voltage drop measured and point A for instance will be directly proportional to the moisture level measured by the spike (R1).

METHODOLOGY: Figure 6 presents a flowchart from experimentation to the moisture classification procedure. Five sensor

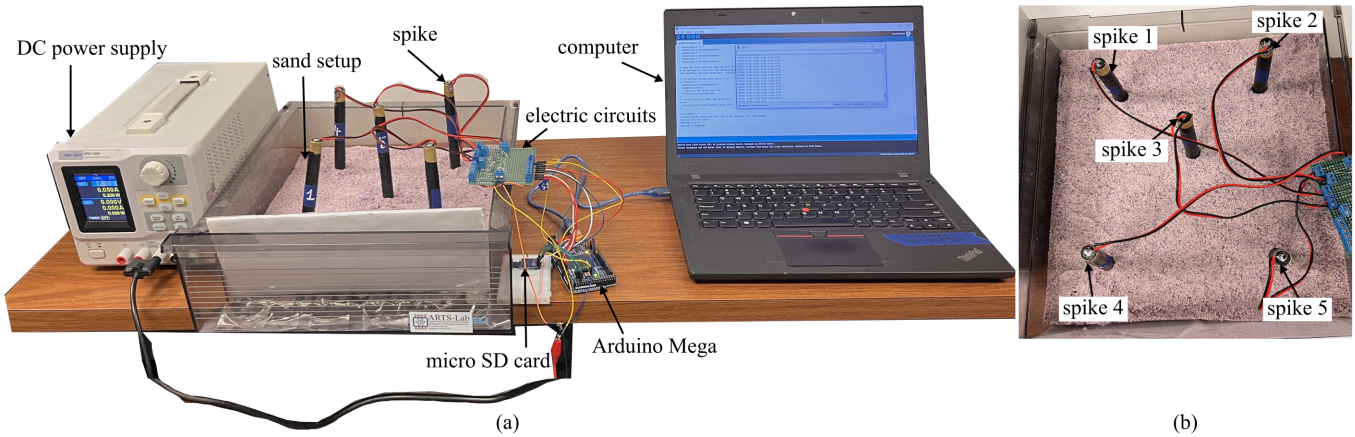

FIGURE 4. Experimental setup of the soil classification moisture test using a network of sensing spikes with key components annotated, showing: (a) table-top test, and; (b) layout of the five sensing spikes used in this work.

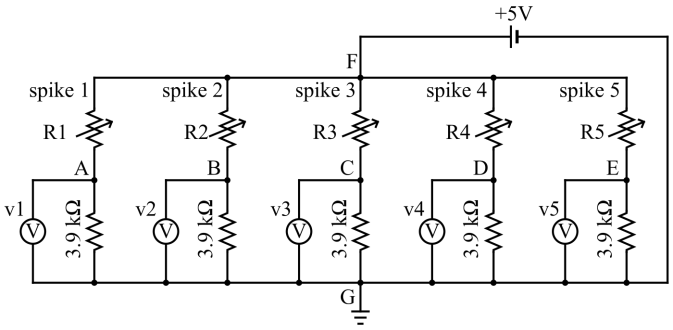

FIGURE 5. Electrical circuit of the network of sensing spikes.

TABLE 1. Five-spike moisture test position coordinates.

spikes	notation of fixed resistor with respect to the spike	x-coordinate (inch)	y-coordinate (inch)	z-coordinate (inch)
1	v_1	2.5	9.3	1.69
2	v_2	8.3	9.3	1.69
3	v_3	5.4	6.4	1.69
4	v_4	2.5	2.5	1.69
5	v_5	8.3	2.5	1.69

spikes are inserted in the sand-filled box during the experiment. Voltage measurements are taken before introducing water flow to acquire a baseline measurement of the dry sand. Water is then added to the corner of the container, and moisture spreads throughout the sand. After gathering data from sensor spikes, the dataset is fed into a mapping algorithm as five simultaneous measurements, where the data is interpolated to cover the entire test area using the kriging model. By applying k-means clustering, the data is then classified into the three categories of moisture

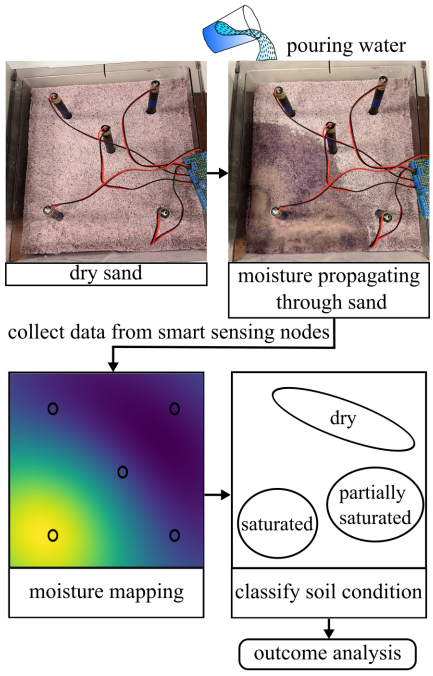

FIGURE 6. A flowchart from the experiment and the soil condition analysis process.

levels.

DATA INTERPOLATION: To interpolate the data for all the spatial points, ordinary kriging is adopted. Where the spikes' locations are $S = [s_1, s_2, \cdots, s_5]$ with their coordinates $[X, Y] = [(x_1, y_1), (x_2, y_2), \cdots, (x_5, y_5)]$. The voltage measurements are given as $V = [v_1, v_2, \cdots, v_5]$. Provided the five observations at the discretized locations, the kriging model attempts to accurately

map a continuous v_k at all possible s_k where $s_k = (x_k, y_k)$. The desired prediction is given in the form,

$$v_k = \mu + \varepsilon(s_k) \tag{1}$$

where μ is the true mean of the entire dataset. As the true mean value μ is undetermined, the estimation is performed by ordinary kriging and $\varepsilon(\cdot)$ is the error (small scale variation) at s. The estimation $\hat{v_k}$ can be presented as:

$$\hat{v_k} = \sum_{i=1}^{n} \lambda_i v_i \tag{2}$$

where λ is the interpolation weight. Here assuming $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1$ to make the unbiased result. For ordinary kriging, three conditions need to be satisfied [21]. 1) Linearity ($\hat{v_k} = \sum_{i=1}^{n} \lambda_i v_i$) 2) Unbiasedness ($\sum_{i=1}^{n} \lambda_i - 1$), and 3) Minimized error: the selection of the most appropriate values for the coefficients λ_n and the Lagrange multiplicator $2m$. E is the estimated function. Therefore, the loss function for the problem becomes:

$$L_{krging} = E\left(v_k - \sum_{i=1}^{n} \lambda_i v_i\right) - 2m\left(\sum_{i=1}^{n} \lambda_i - 1\right) \tag{3}$$

PyKrige library is utilized to do the ordinary kriging model. The [X, Y, V] is used to train the Gaussian variogram models. After that, the trained model is used to estimate the values for the overall setup (10 x 10) with a resolution of 0.01 inches on each axis.

CLUSTERING

This work classifies moisture levels in earthen levees into three clusters, namely 'dry', 'partially saturated', and 'saturated'. The data obtained by the interpolation process is fed into a k-means clustering algorithm to be classified into one of the three categories. To measure the similarity between different points, for instance, s_p and s_q, the squared Euclidean distance is used with Voltage(v) being the sole feature considered. So, the similarity is shown as:

$$||s_p - s_q||_2^2 = (v_p - v_q)^2 \tag{4}$$

Based on the Euclidean distance metric the k-means algorithm can be considered a simple optimization problem. The iterative approach is followed to minimize the within-cluster sum of squared error (SSE) or cluster inertia. The objective would be:

$$L_{SSE} = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{(i,j)} ||v_i - c_j||_2^2 \tag{5}$$

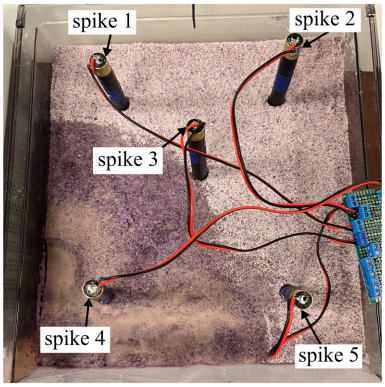

FIGURE 7. The experimental setup showing the positions of the five spikes during moisture propagating.

TABLE 2. Five spikes voltages for a single time stamp.

spikes	spike 1	spike 2	spike 3	spikes 4	spike 5
voltage (V)	0.014	0.014	0.024	1.549	0.078

where, c_j is the centroid for cluster j and $w^{(i,j)} = 1$ if the sample v_i is in cluster j or 0 otherwise. As three clusters are being used here, m=3.

RESULTS AND DISCUSSION

This study's findings are divided into two categories: kriging and clustering.

During kriging, only a single timestamp of measurements is considered for moisture mapping of the whole experimental area. From table 2, spike 4 voltage measurement is shown to be higher compared to the other four spikes while spikes 1 and 2 show the lowest voltage reading. So the area surrounding spike 4 is considered to have the highest level of moisture. Figure 7 shows the experimental setup after introducing water flow to the corner of the test area, which correlates to the moisture map produced by the kriging algorithm shown in Figure 8 (a).

Spatial kriging outcomes from the moisture test are shown in figure 8 (a). From moisture mapping, the voltage range shows around 0 to 1.5 V. Moisture mapping by spatial kriging shows similar outcomes to the experimental outcomes as figure 7. The bottom left corner has the largest voltage meaning more moisture compared to the top right corner of the mapping which is around 1.5 V.

Utilizing data from moisture mapping to classify the soil conditions. Figure 8 (b) displays three clusters labeled dry, partially saturated, and saturated based on k-mean clustering. The green-colored cluster is the saturated area which matches the ground truth shown in Figure 7. The orange color cluster is the partially saturated area and finally, the cyan color cluster indi-

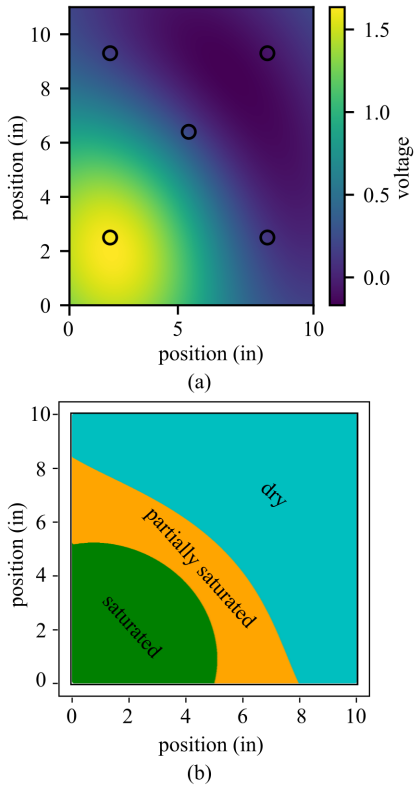

FIGURE 8. Outcomes from the moisture test showing: (a) spatial kriging outcomes; and (b) categorizing soil conditions by clustering.

TABLE 3. Threshold and centroid values for *k*-mean clustering.

minimum v (V)	maximum v (V)	centroid v (V)	categories
-0.168	0.324	0.019	dry
0.324	1.000	0.629	partially saturated
1.000	1.635	1.372	saturated

cates dry sand.

Table 3, shows the threshold of each cluster in voltage and centroid value for each cluster. This voltage is directly proportional to the soil moisture. The highest voltage of the centroid is 1.372 V at the threshold between 1.0 to 1.635 V. This centroid value is close to the spike 4 value 1.549 V values from table 2 and considered this cluster as a saturated one. Moisture at spike 4 is greater than at other spikes.

The centroid value of 0.629 V is considered partially saturated as this threshold 0.324 to 1.000 V is close to spike 5, 3 values from table 2. The lowest value of the centroid is 0.019 V and ranges between -0.168 to 0.324 V whose location is the right side of the experimental setup figure 7. Ordinary kriging produces a negative value as a measurement. When nearby data eliminate outlying data, ordinary kriging (OK) encounters negative weights. This cluster is categorized as dry.

Dry, partially saturated, and saturated soil is properly categorized as such in the soil condition analysis from the moisture propagation stage from experimental figure 7, moisture mapping figure 8 (a), and finally clustering figure 8 (b).

CONCLUSION

This paper presents a preliminary experiment on the development and validation of UAV-deployable smart sensing spikes for soil conductivity levels in levees, a measurement method for estimating soil saturation levels. For this investigation, Gaussian process regression, widely known as kriging, is utilized to estimate the mapping of soil saturation between sensing spikes to get continuous conductivity maps of the surface soil in levees. The expanded data is then classified using a clustering-based machine learning approach, with conductivity data from sensor spikes contributing as model inputs. The measured voltage is exactly proportional to the soil moisture as it is inversely related to soil resistance. The threshold for the saturated condition is chosen by clustering and ranges from 1.00 to 1.635 V with a centroid value of 1.372 V. The centroid value for the dry state is 0.019 V, with a threshold of -0.168 to 0.324 V. Using a threshold of 0.324 to 1.000 V and a centroid of 0.629 V, the partially saturated state that lies between dry and saturated conditions is also characterized. To identify possible levee failure concerns and maintenance needs, this work evaluates soil conditions utilizing a network of smart sensing sensor spikes.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation Grant number 2152896. The support of the National Science Foundation is gratefully acknowledged. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Tsai, C.-C., Yang, Z.-X., Chung, M.-H., and Hsu, S.-Y., 2022. "Case study of large-scale levee failures induced by cyclic softening of clay during the 2016 meinong earthquake". *Engineering Geology,* ***297***, p. 106518.

[2] Pengel, B., Krzhizhanovskaya, V., Melnikova, N., Shirshov, G., Koelewijn, A., Pyayt, A., Mokhov, I., et al., 2013. "Flood early warning system: sensors and internet". *IAHS Red Book,* ***357***, pp. 445–453.

[3] Melnikova, N., Jordan, D., and Krzhizhanovskaya, V., 2015. "Experience of using fem for real-time flood early

warning systems: Monitoring and modeling boston levee instability". *Journal of Computational Science,* ***10***, pp. 13–25.

[4] Zwęgliński, T., 2020. "The use of drones in disaster aerial needs reconnaissance and damage assessment – three-dimensional modeling and orthophoto map study". *Sustainability,* ***12***(15), jul, p. 6080.

[5] Carroll, S., Satme, J., Alkharusi, S., Vitzilaios, N., Downey, A., and Rizos, D., 2021. "Drone-based vibration monitoring and assessment of structures". *Applied Sciences,* ***11***(18), sep, p. 8560.

[6] Orlandini, S., Moretti, G., and Albertson, J. D., 2015. "Evidence of an emerging levee failure mechanism causing disastrous floods in italy". *Water Resources Research,* ***51***(10), oct, pp. 7995–8011.

[7] Camici, S., Barbetta, S., and Moramarco, T., 2015. "Levee body vulnerability to seepage: the case study of the levee failure along the foenna stream on 1 january 2006 (central italy)". *Journal of Flood Risk Management,* ***10***(3), feb, pp. 314–325.

[8] Rapti, I., Lopez-Caballero, F., Modaressi-Farahmand-Razavi, A., Foucault, A., and Voldoire, F., 2018. "Liquefaction analysis and damage evaluation of embankment-type structures". *Acta Geotechnica,* ***13***(5), feb, pp. 1041–1059.

[9] Balistrocchi, M., Moretti, G., Orlandini, S., and Ranzi, R., 2019. "Copula-based modeling of earthen levee breach due to overtopping". *Advances in Water Resources,* ***134***, dec, p. 103433.

[10] Lee, E., and Griffiths, J., 1987. "The importance of pedological soil survey in land use planning, resource assessment and site investigation". *Geological Society, London, Engineering Geology Special Publications,* ***4***(1), pp. 453–466.

[11] Uygur, V., Irvem, A., Karanlik, S., and Akis, R., 2010. "Mapping of total nitrogen, available phosphorous and potassium in amik plain, turkey". *Environmental Earth Sciences,* ***59***, pp. 1129–1138.

[12] Conforti, M., Matteucci, G., and Buttafuoco, G., 2017. "Organic carbon and total nitrogen topsoil stocks, biogenetic natural reserve 'marchesale'(calabria region, southern italy)". *Journal of Maps,* ***13***(2), pp. 91–99.

[13] Gia Pham, T., Kappas, M., Van Huynh, C., and Hoang Khanh Nguyen, L., 2019. "Application of ordinary kriging and regression kriging method for soil properties mapping in hilly region of central vietnam". *ISPRS International Journal of Geo-Information,* ***8***(3), p. 147.

[14] Wackernagel, H., 2003. *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.

[15] Sinaga, K. P., and Yang, M.-S., 2020. "Unsupervised k-means clustering algorithm". *IEEE access,* ***8***, pp. 80716–80727.

[16] Jain, A. K., and Dubes, R. C., 1988. *Algorithms for clustering data*. Prentice-Hall, Inc.

[17] Bozdogan, H., 1987. "Model selection and akaike's information criterion (aic): The general theory and its analytical extensions". *Psychometrika,* ***52***(3), pp. 345–370.

[18] Davies, D. L., and Bouldin, D. W., 1979. "A cluster separation measure". *IEEE transactions on pattern analysis and machine intelligence*(2), pp. 224–227.

[19] Pal, N. R., and Biswas, J., 1997. "Cluster validation using graph theoretic concepts". *Pattern Recognition,* ***30***(6), pp. 847–857.

[20] Chowdhury, P., Satme, J. N., Flemming, M., Downey, A. R., Elkholy, M., Imran, J., and Khan, M. S., 2023. "Stand-alone geophone monitoring system for earthen levees". In Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2023, Vol. 12486, SPIE, pp. 168–174.

[21] Malviæ, T., and Baliæ, D., 2009. "Linearity and lagrange linear multiplicator in the equations of ordinary kriging". *Nafta,* ***59***(1), pp. 31–37.