"DATA ASSIMILATION IN A HYDROLOGICAL LANDSLIDE MODEL WITH ENSEMBLE KALMAN FILTER"

Amirahmad Vakilinezhad[1], Signe Othelie Petrohai Pedersen[2], Ivan Depina[1]

KEYWORDS

Data assimilation, Hydraulic parameters estimation, Ensemble Kalman Filter, Uncertainty reduction

ABSTRACT

Landslides pose significant negative impact on people lives, infrastructures, society, and environment. To reduce the consequences of this natural hazard, spatial and temporal estimations of failure are essential. Yet substantial uncertainties are associated with these assessments. This study is motivated by the need to reduce the uncertainty in soil hydraulic parameters estimations for a better landslide susceptibility and hazard assessment. The research highlights the vital role of diminishing uncertainties in soil hydraulic parameters leading to better estimation of hydraulic conductivity, infiltration, pore pressure distribution and soil strength, therefore improved temporal and spatial slope instability forecasting. The study focuses on a case study in the proximity of Meråker in central Norway because of the complex morphology and geology of the area as well as the available instrumented slopes to measure the soil suction and volumetric water content throughout the year. By utilizing the high-resolution rainfall data from the weather stations in the area together with sensor data, a sequential uncertainty reduction of soil hydraulic properties was implemented with the Ensemble Kalman Filter Method (EnKF) and PLAXIS 2D software. A Python script is employed to implement EnKF and automated Plaxis numerical simulations to assimilate data from sensors and calibrate hydraulic parameters. Preliminary findings reveal considerable reduction in soil hydraulic properties uncertainty leading to improved performance of regional slope stability analysis and ultimately contribute to better geohazard management, effective mitigation strategies and fortifying the community resilience.

[1] Norwegian University of Science and Technology (NTNU)
[2] Geovita as

1. INTRODUCTION

Spatial and temporal prediction of water-induced landslides are crucial to protect people's lives and reduce the socioeconomical consequences through improved spatial planning, engineering protection measures, and early warning systems. In this regard understanding the soil hydraulic properties is the key factor to estimate the water infiltration, ground water condition, soil strength reduction and therefore slope stability. Despite of the importance of soil hydraulic parameters, due to the inherent spatial variability and costly and time consuming soil investigation measures it is common that the hydraulic parameters are associated with great uncertainties. In this paper, Ensemble Kalman Filter (EnKF) data assimilation method is employed together with Van Genuchten unsaturated infiltration theory using PLAXIS 2D to calibrate and update the soil hydraulic parameters continuously, leading to more precise estimation of uncertain hydraulic parameters. The EnKF assimilation of data and dynamic updating features make it exceptional tool for integration of real-time sensor data with conventional models such as Van Genuchten model to evaluate the ground water condition in response to rainfall. Moreover, employing novel data assimilation methods such as EnKF is first step to pave the road for further incorporation of innovative techniques in monitoring and early warning systems for water-induced landslides.

2. METHOD

The study utilizes a finite element PLAXIS 2D model combined with EnKF data assimilation algorithm to calibrate the three uncertain soil hydraulic parameters namely Van Genuchten SWCC fitting parameters α, n, and saturated hydraulic conductivity K_s. The sensors in the study area record volumetric water content, suction and temperature. A slope similar to the instrumented slope is modeled in PLAXIS 2D and using the available geotechnical and hydrological data, infiltration analysis is performed by assuming a probability distribution for unknown hydraulic parameters α, n and K_s. The results of the PLAXIS 2D model is compared to the actual sensor data in the field and using the EnKF the initial assumptions are calibrated. After updating the parameters with 10 over the period of 9 days, the goals is to reduce uncertainty associated with initial guesses and obtain more accurate parameters probability distribution. Subsequently these hydraulic parameters can be used to perform a slope stability analysis and enhance the accuracy of the deployed early warning systems in the case study area.

Instrumentation and study area

The instrumented area located along the Stjørdal river in Trøndelag, central Norway. Area covers almost 200 km^2 area and consist of complex geological formation according to Norway National Geological Survey (NGU) shown in

Figure 1. Two locations in the area are chosen to be instrumented with sensors shown in Figure 1.

At each spot, sensors are located at two proximate locations and three depths in Figure 2. This study focuses on sensor data from "Location 2" in Figure 1, where the soil Volumetric Water Content (VWC) has been measured at depths of 0.3, 0.5 and 0.9 m from the

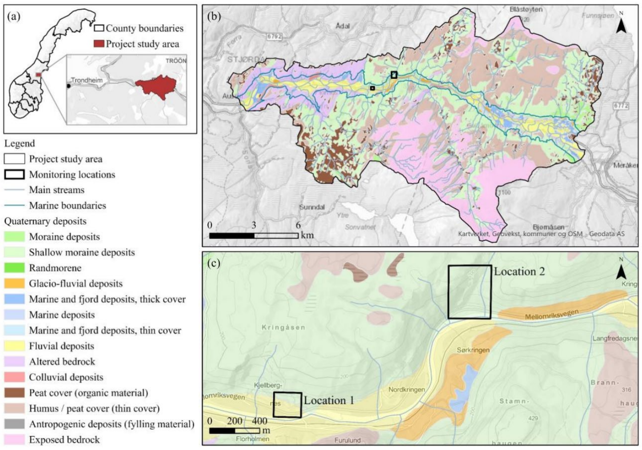


Figure 1. (a) Study area, (b) Geological formation of the area, (c) Two selected monitoring locations. [1]

ground surface, and is accessible in 15-minute resolution. A weather station to measure the daily rainfall also planted at the area.

Annual rainfall of 964 mm to 1205 mm [2] has been registered in the case study area. Soil conditions are characteristic for steep areas with inclination greater than 25 degrees and shallow dept to bedrock with soil thickness varying between 0.5 m to 10 m above the bedrock [1]. Geotechnical soil properties are presented in [1] after literature review and sets of laboratory tests from the soil samples. Soil friction angle and cohesion are measured as 38° and 5.5 kPa respectively and the soil is classified as silty sand. This characteristic for moraine, which is a well-graded soil type consisting of silt, sand, and gravel.

Ensemble Kalman Filter (EnKF)

The main objective of data assimilation is to successively adjust the state of knowledge about unknown parameter distribution based on the initial assumed distribution, model predictions, and incoming measured state or parameter. Kalman Filter (KF) is one of the well-known data assimilation systems, optimizing the accuracy of the system by processing and adjusting the unknown parameters [3]. Ensemble Kalman Filter (EnKF) is the extension of KF, applicable in processing nonlinear systems by employing the set of samples of unknown parameters known as ensembles to approximate the state or parameter's distribution and therefore optimizing the accuracy of the model [4].

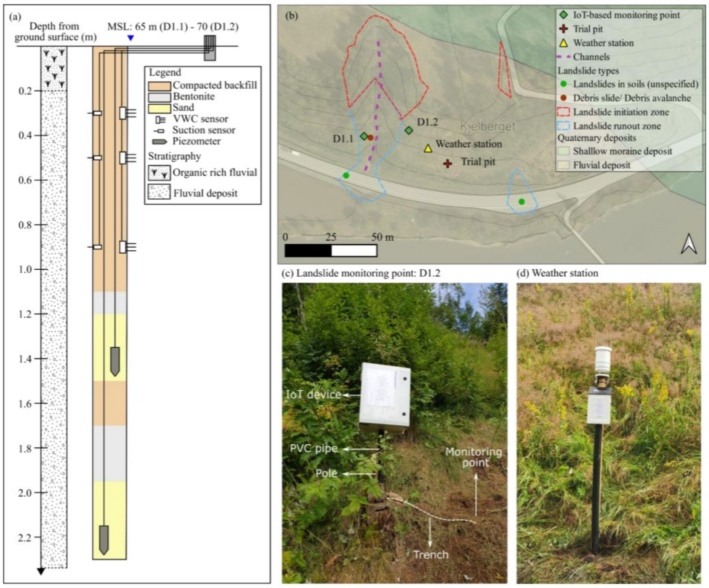


Figure 2. Study area [1] *(a) Soil layering and sensor locations, (b) Two sensors, weather station and trial pit locations, (c) Monitoring point and (d) weather station*

Comprehensive notation and formulation of the EnKF is provided in [5]. Following the same notation, EnKF is formulated and incorporated in this study.

Let us define the output of a model that aims to capture the considered phenomenon (e.g., rainfall infiltration) as g(**z**). The input to the model, **z**, is a matrix that combines model state (estimation), $\boldsymbol{x}$, and the unknown model parameters (forward calculation input), $\boldsymbol{\theta}$

$$\boldsymbol{y} = g(\boldsymbol{z}) \qquad Eq(1)$$

$$\boldsymbol{z} = [\mathbf{x}, \boldsymbol{\theta}]^T \qquad Eq(2)$$

Measurement matrix **d** is defined as observation of the model output **y** with the addition of a measurement error **e**.

$$\boldsymbol{d} = \boldsymbol{y} + \boldsymbol{e} \qquad Eq(3)$$

In order to get the close to reality estimations the posterior likelihood of the model parameters given the observations, f (**z** | **d**), is maximized, and this can be done by minimizing the cost function [5].

$$J(\boldsymbol{z}) = \left(\boldsymbol{z} - \boldsymbol{z^f}\right)^T \boldsymbol{C}_{\boldsymbol{zz}}^{-1}\left(\boldsymbol{z} - \boldsymbol{z^f}\right) + (g(\boldsymbol{z}) - \boldsymbol{d})^T \boldsymbol{C}_{\boldsymbol{dd}}^{-1}(g(\boldsymbol{z}) - \boldsymbol{d}) \qquad Eq(4)$$

$\mathbf{C}_{\mathbf{zz}}$, $\mathbf{z}^{\mathbf{f}}$ and $\mathbf{C}_{\mathbf{dd}}$ represent error covariance of **z**, previous estimation of model parameters and error covariance of measurements respectively. The state-parameter vector is updated by maximizing the posterior likelihood [5] as follows:

$$\boldsymbol{z^a} = \boldsymbol{z^f} + \boldsymbol{K}\big(\boldsymbol{d} - g(\boldsymbol{z})\big) \qquad Eq(5)$$

$$\boldsymbol{C}_{zz}^{a} = (\boldsymbol{I} - \boldsymbol{KG})\boldsymbol{C}_{zz} \qquad Eq(6)$$

The “a” superscript represents the new estimation. New parameter estimation and error covariance estimation can be obtained through equation 5 and 6 using previous estimation and **K** matrix which is known as “Kalman gain”.

$$\boldsymbol{K} = \boldsymbol{C}_{zz}\boldsymbol{G}\left(\boldsymbol{G}\boldsymbol{C}_{zz}\boldsymbol{G}^{T} - \boldsymbol{C}_{zz}\right)^{-1} \qquad Eq(7)$$

For each ensemble the equations can be written:

$$z_i^a = z_i^f + \boldsymbol{K}^{e}\left(\boldsymbol{d}_i - g\left(z_i^f\right)\right) \qquad Eq(8)$$

$$\boldsymbol{K}^{e} = \boldsymbol{C}_{zz}^{e}\boldsymbol{G}\left(\boldsymbol{G}\boldsymbol{C}_{zz}^{e}\boldsymbol{G}^{T} - \boldsymbol{C}_{dd}\right)^{-1} \qquad Eq(9)$$

$\mathbf{d}_i = \mathbf{d} + \boldsymbol{\varepsilon}_i$ and $\boldsymbol{C}_{zz}^{e}$ are the measurement matrix with noise and combined co-variance matrix respectively. $\boldsymbol{C}_{zz}^{e}$ is calculated using equations 10, 11 and 12 by first finding ensemble members mean and differentiating each ensemble member from the mean value.

$$\overline{\boldsymbol{Z}_t^f} = \boldsymbol{Z}_t^f \boldsymbol{I}_{Ne} \qquad Eq(10)$$

$$\boldsymbol{Z}_f' = \boldsymbol{Z}_f^{T} - \overline{\boldsymbol{Z}_t^f} \qquad Eq(11)$$

$$\boldsymbol{C}_{zz}^{e} = \frac{\boldsymbol{Z}_f'\left(\boldsymbol{Z}_f'\right)^{T}}{Ne - 1} \qquad Eq(12)$$

The measurement matrix is defined as **D** and it is integrated in the main updating function as shown in equation 14.

$$\boldsymbol{D}_t = \left(\boldsymbol{d}_{1,t}, \boldsymbol{d}_{2,t}, \boldsymbol{d}_{3,t}, \dots, \boldsymbol{d}_{Ne,t}\right) \qquad Eq(13)$$

$$\boldsymbol{Z}_t^a = \boldsymbol{Z}_t^f + \boldsymbol{C}_{zz}^{e}\boldsymbol{G}^{T}\left(\boldsymbol{G}\boldsymbol{C}_{zz}^{e}\boldsymbol{G}^{T} + \boldsymbol{C}_{dd}\right)^{-1}\left(\boldsymbol{D}_t - \boldsymbol{G}\boldsymbol{Z}_t^f\right) \qquad Eq(14)$$

In this study, 3 soil hydraulic parameters α, n and K_s, initially characterized by broad probability distribution are selected to be calibrated. Ne = 20 random samples from the distributions are attained and 20 analyses are performed to obtain ensemble of result, which consists of soil VWC predictions at measurement locations. The estimations are updated by incorporating the measurements from sensors data with 0.02 error. This process is repeated for N_m = 9 days and at each day the analysis started with updated parameter probability distribution. Ideally, outputs of analyses should be converging the real sensor data from the field after iterations.

FINITE ELEMENT PLAXIS 2D MODEL

Python scripting is used to automate the PLAXIS 2D analysis. With 20 random estimations for three parameters the analysis should be performed 20 times for each day. 9 days of data are processed to be used in this study, resulting in total

180 PLAXIS 2D analyses, which would be time consuming to perform manually.

The FEM model begins with defining the geometry and boundary conditions. The geometry of the model is illustrated in Figure 3 where a 35° slope is created with 1.5 m depth to impermeable layer following the available information from trial pits in the field. Soil material is specified following the values presented by [1]. Three unknown parameters α, n and K_s are defined as a function to be updated at each iteration in the model.

Similar to the sensor locations at 0.3, 0.5 and 0.9 m from the surface, three closest nodes on the mesh at the same depths are chosen to obtain the volumetric water content after each analysis and compared to sensor values. The probability distribution function is updated and subsequently 20 new random samples of unknown parameters are generated o run the model.

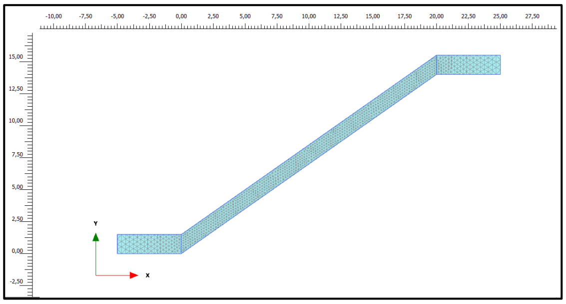
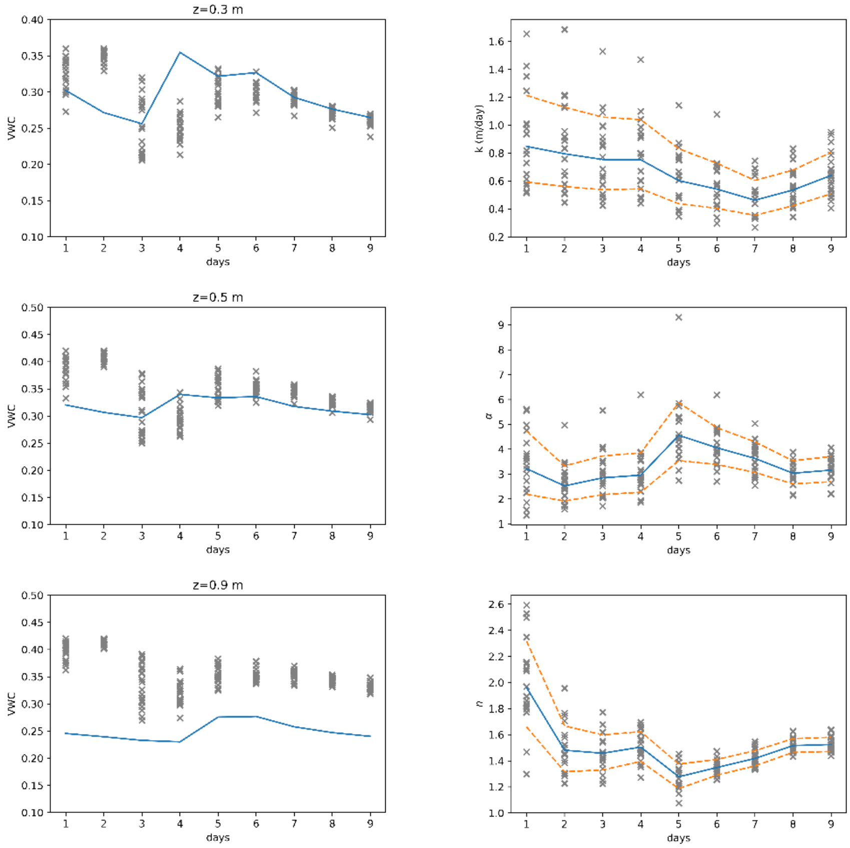


Figure 3. PLAXIS 2D model geometry

3. RESULTS AND DISCUSSIONS

Prior and posterior parameters estimations are presented in Table 1. After 9 iterations, the posterior parameters are obtained. It is expected that the output of the FEM analysis resulted from posterior distributions would be getting closer to the sensor data if the model is able to capture the rainfall infiltration process. In Figure 4(a) the VWC recorded from sensors and simulation are shown by the blue line and black crosses respectively. At all depth the initial estimations result in wide range of VWC that converges after a couple of iterations although at the beginning they are far from the real values. At 0.3 m the simulation results are capturing the VWC change after six days. At the 0.5 m depth similarly the simulation results are getting closer to sensor data after six days iteration. However, at 0.9 m depth the predicted VWC is not in good agreement with the measured values. This can be due to the bias in the model, resulting from soil hydrological properties variability at deeper points and inaccurate implementation of initial and boundary conditions in our model.

Figure 4(b) shows the hydraulic parameters estimations. The mean value is shown by blue line and orange dash lines show the one standard deviation from the mean. Saturated hydraulic conductivity shows acceptable convergence and reach to an almost steady number with a small deviation from the mean in 9 days. α parameter distribution hasn't changes significantly after 9 iterations, however, the model is not dramatically



(a) (b)

Figure 4. (a) Sensor data and simulation results of VWC at 0.3, 0.5 and 0.9 m. (b) initial k_{sat}, α and n parameter estimation and distributions in 9 days

Table 1. Mean and standard deviation of the unknown parameters before and after EnKF

Prior Values						Posterior Values					
μ_α	σ_α	μ_n	σ_n	μ_{Ks}	σ_{Ks}	μ_α	σ_α	μ_n	σ_n	μ_{Ks}	σ_{Ks}
4.0	2.0	2.0	0.5	1.0	0.5	3.19	0.49	1.52	0.05	0.65	0.15

sensitive to α parameter since at shallower depth, the model shows acceptable performance and captures the actual VWC values although the α parameter wasn't estimated well. On the other hand, the n parameter deviation from the mean is getting smaller after each iteration and shows acceptable convergence in 9 days.

4. CONCLUSIONS

Concluding the study, it is evident that the EnKF can be employed to automate calibration of numerical models with real-time data. In all scenarios the EnKF effectively reduced the uncertainty in initial estimated parameters. Moreover, efficiency of EnKF is showed at 0.3 m and 0.5 m depth where estimations and sensor values converged precisely. Although the results at 0.9 m depth are not as close as our expectations, these issues can be addressed by improving the model capability such as modelling layered soil that can capture the variation in hydraulic parameters, and improve modelling of boundary conditions, accounting for evapotranspiration, and increasing the number or frequency of sensor data.

REFERENCES

[1] E. A. Oguz, I. Depina, B. Myhre, G. Devoli, H. Rustad, and V. Thakur, "IoT-based hydrological monitoring of water-induced landslides: a case study in central Norway," *Bull. Eng. Geol. Environ.*, vol. 81, no. 5, 2022, doi: 10.1007/s10064-022-02721-z.

[2] C. G. Leiva and S. Mechanics, "Report Susceptibility assessment of rainfall," 2019.

[3] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960, doi: 10.1115/1.3662552.

[4] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. C5. 1994. doi: 10.1029/94jc00572.

[5] M. Mohsan, P. J. Vardon, and F. C. Vossepoel, "On the use of different constitutive models in data assimilation for slope stability," *Comput. Geotech.*, vol. 138, no. March, p. 104332, 2021, doi: 10.1016/j.compgeo.2021.104332.