Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

Signe Othelie Petrohai Pedersen

Estimation of Van Genuchten Parameters using Ensemble Kalman Filter for Hydrological Modeling

Master's thesis in Civil and Environmental Engineering
Supervisor: Ivan Depina
June 2023

NTNU
Norwegian University of
Science and Technology

Signe Othelie Petrohai Pedersen

Estimation of Van Genuchten Parameters using Ensemble Kalman Filter for Hydrological Modeling

Master's thesis in Civil and Environmental Engineering
Supervisor: Ivan Depina
June 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Norwegian University of
Science and Technology

Preface

This master thesis was written in the spring 2023 as the final product of a 5-year degree in Civil and Environmental Engineering at NTNU in Trondheim. The thesis is conducted for the geotechnical department with Ivan Depina (NTNU) as the main supervisor.

The goal of the thesis was to calibrate hydrological van Genuchten parameters of a soil using the Ensemble Kalman Filter method and sensor data from IoT-sensors installed for the KlimaDigital-project. Automating the calibration of hydrological parameters in soil holds significant potential and could possibly be applicated in problems such as automated early-warning systems for rainfall-induced landslides.

Acknowledgements

First and foremost, I am immensely grateful for my supervisor, Ivan Depina, at NTNU's Geotechnical Department. His guidance, expertise, and unwavering support throughout this thesis journey have been instrumental in shaping the direction of my thesis. Ivan not only defined the topic and idea for this thesis, but also provided me with insight and knowledge to tackle the problem at hand. He assisted me in finding relevant literature, helped in writing the code, and helped me gain a comprehensive understanding of the problem.

I would also like to extend my gratitude to all the professors at the Geotechnical Department at NTNU. Their dedication to teaching and their expertise in geotechnical engineering have played a crucial role in shaping my understanding of the field and inspiring me to pursue this path over the past five years. Their guidance, challenging coursework, and insightful discussions have contributed to my growth as a geotechnical engineer.

Additionally, I would like to express my appreciation to all the other master students at our office at Tower of London at Lerkendalsbygget. The supportive environment we shared have been invaluable during this semester. Their encouragement has made the journey much more enjoyable.

Abstract

Rainfall-induced landslides make out a considerable amount of geohazards in Norway. This is from combination of steep slopes, heavy rainfall during fall, temperature changes, and snow melting in spring and fall. Landslides are dangerous because they can damage critical infrastructure and in the worst case be a threat to human life. Climate change is likely to result in more and intense rainfall events, which is further increasing the landslide risks. Consequently, there is a need to develop strategies to manage landslide risks. One of such strategies are landslide early warning systems that provide a timely that can be used to evacuate people, movable properties, or close roads to reduce consequences in case of a landslide.

In 2018, SINTEF launched the project KlimaDigital with the main objective of creating an early warning system for rainfall induced landslides and debris flow, using monitoring data from sensors installed in landslide prone areas. Sensors were installed on two different slopes in Meråker, Trøndelag to monitor groundwater conditions. The area was selected due to the

high risk for rainfall-induced landslides resulting from the combinations of heavy rainfall and rapid snow-melting events, and topography characterized by steep slopes (Ivan Depina E. O., 2021)

When working with rainfall induced landslide problems, it is crucial to understand groundwater conditions in slopes in response to rainfall infiltration and snow melting. Slopes in the studied area are typically unsaturated and groundwater conditions were monitored with volumetric water content sensors. In this thesis, the volumetric water content sensor data from the Meråker slopes was used to calibrate the hydrological van Genuchten model parameters and the permeability of the soil, using a Plaxis flow-analysis and the Ensemble Kalman Filter method (EnKF). The estimated van Genuchten parameters are defining the Soil Water Characteristics curve (SWCC), that describes relations between suction, permeability, and degree of saturation in the given soil. Calibrated hydrological models provide a basis for more accurate modelling of groundwater conditions (e.g., in response to extreme weather events) and implementation of a reliable landslide early warning system.

Python was used to implement the EnKF algorithm and automate Plaxis analysis of the hydrological model, making it possible to update and run the calculations numerous of times with different hydrological input parameters for each iteration. The Ensemble Kalman filter method works by comparing the Plaxis output values to the real sensor data and updating the parameters based on this difference and the parameter covariance. By iterating through this

process, the unknown parameters were updated until the van Genuchten parameters were estimated with better accuracy and the difference between model predictions and sensor values became relatively small.

The study concludes that the EnKF algorithm shows promise in estimating these parameters; however, the accuracy of the initial conditions and prior parameter knowledge are critical for obtaining more reliable estimations. The research also highlights the potential application of this method in early warning systems for rainfall-induced landslides. However, further refinement and increased certainty in the results are necessary to ensure reliability.

Future work could include incorporation of suction sensors alongside the VWC-sensors to enhance the calibration results. It would also be interesting to implement the results in a safety analysis to investigate how the hydrological conditions influence slope stability.

Norsk Sammendrag

Nedbørsinduserte jordskred er en betydelig fare i Norge på grunn av bratte skråninger, mye nedbør, temperaturendringer og snøsmelting. Klimaendringer øker risikoen ytterligere. Prosjektet KlimaDigital, startet av SINTEF i 2018, har hatt som mål å utvikle et varslingssystem for nedbørsinduserte jordskred ved hjelp av sensorer.

I denne oppgaven er sensorverdier fra KlimaDigital-prosjektet brukt til å kalibrere de hydrologiske van Genuchten parameterne til jorda. Dette er gjort ved å bruke Ensemble Kalman Filter-algoritmen (EnKF). Simuleringene er gjort ved å lage en automatisert Plaxis-modell i Python og EnKF fungerer ved å oppdatere de ukjente inputparameterne basert på forskjellen mellom sensordata og Plaxis-output. Formålet er at Plaxis-outputen skal sammenfalle med sensorverdiene, noe som tyder på en virkelighetsnær modell. I tillegg er hensikten at de ukjente parameterne skal konvergere mot estimater med høyere sikkerhet.

Oppgaven konkluderer med at EnKF-algoritmen viser gode tendenser i estimatet av disse parameterne. I midlertidig er nøyaktigheten av de initielle betingelsene og forhåndskunnskapen om parameterne avgjørende for å oppnå mer pålitelige estimater. Oppgaven peker også på den mulige bruken av denne metoden i varslingssystemer for jordskred. Ytterligere for bedring og økt sikkerhet i resultatene er derimot nødvendig for å sikre pålitelighet.

Fremtidig arbeid vil kunne være å implementere sensorverdier for sug, samt kalibrere parameterne med verdier fra faktiske skredhendelser.

Contents

LIST OF FIGURES

List of tables

1 Introduction

With the changing climate resulting in wetter and warmer weather conditions, geohazards are becoming more significant than ever. In accordance with the UN climate report (United Nations (UN), 2021), wetter and warmer climate will lead to more extreme weather meaning higher intensity rainfall both in Norway and in the word. Heavy rainfall is one of the most common triggers for landslides in Norway and rainfall-induced landslides pose a significant threat to both human lives and critical infrastructure. This threat is making it imperative to develop efficient strategies for managing landslide risks. One of these strategies are landslide early warning systems that reduce consequences by providing timely warning, which allow for the people and movable property to be evacuated before the onset of a landslide.

Working with rainfall-induced landslides requires consideration of the unsaturated soil conditions, partially saturated pores, which results in suction and affects both the permeability and effective stresses of the soil. The relationship between the hydrological soil parameters can be described by a Soil Water Characteristics Curve (SWCC), relating suction to volumetric water content (or saturation) and permeability of the soil. Having a good understanding of the SWCC is essential in being able to model the development of the wetting front and pore pressures accurately and reliably to provide timely warnings in a landslide early warning system.

The hydrological parameters of the soil can be difficult to estimate. Numerous of pedotransfer functions (PTF's) have been developed throughout the last four decades with the purpose of relating the hydrological parameters of the soil to parameters that are easier to determine through laboratory tests, such as the grain size distribution and porosity (Abdelbaki, 2020).

Another option for determining the hydrological parameters in the soil is to implement a sensor system to monitor real life data and calibrate the hydrological parameters with respect to sensor values. In this thesis a calibration like this is done using sensors monitoring volumetric water content (VWC), combined with a hydrological model in Plaxis and the EnKF-algorithm.

The EnKF method has shown great results in a variety of fields where monitoring data and joint parameter estimation play an essential role, such as meteorology and oceanology (Muhammad Mohsan, 2021). In this master's thesis, this algorithm was used to calibrate the unknown van Genuchten parameters, but the method is very general and can be applied to other similar problems in geotechnical engineering where model calibration with data is needed, for

example by using monitoring data of deformations to estimate stiffness parameters. (Muhammad Mohsan, 2021) demonstrated the efficacy of the EnKF method to successfully employing the calibrated stiffness and strength parameters to estimate the Factor of Safety (FOS), yielding favorable results.

1.1 Background

The background for this master thesis is the KlimaDigital project that was launched by SINTEF, in collaboration with NTNU, back in 2021. The objective of the KlimaDigital project is to reduce the risk of infrastructure damage and fatalities caused by geohazards such as landslides and debris flows by creating early warning systems based on data monitoring. Due to climate change causing more heavy rainfall events and extreme weather, the frequency of dangerous landslide events both has, and is expected to further increase. Understanding the state of the soil during heavy rainfall and developing reliable early warning systems is therefore more important than ever (Torun Rise, 2023).

Currently, there is a national early warning system for rainfall-induced landslides operated by NVE. However, this system operates on a regional to national level with the system not being suited to providing warnings on local to regional level, which makes the system less applicable when it comes to issuing warnings on a local and more refined scale.

There is a growing need for reliable and cost-efficient local to regional landslide warning system. To scale down the warning system one needs to be able to obtain more refined local knowledge of the parameters controlling the occurrence of landslides. One of the approaches in obtaining such knowledge is using reliable and coset-efficient monitoring. This approach requires an automated data management system to effectively incorporate sensor data into a model. This master's thesis seeks to address this knowledge need by utilizing sensor values from the KlimaDigital project's monitoring system to calibrate a hydrological model of a slope. A calibrated hydrological model is essential in providing timely and reliable predictions of rainfall-induced landslides.

This thesis's significance lies in its contribution to the development of early warning systems for rainfall-induced landslides, which is an area of research that is still in its early stages. In addition, the thesis shows how the Ensemble Kalman Filter method (and other Kalman Filter methods) can be used to automate calibration of geotechnical models with data. The algorithm is of exceptional value in situations with substantial amounts of real-time data.

1.2 Problem Formulation

The overarching goal of this master thesis was to automate the calibration of hydrological van Genuchten parameters of the soil based on sensor values. The unknown parameters are used to fit the SWCC for the soil, which is an essential part of hydrological analysis of rainfall-induced landslides.

The parameter calibration was to be automated with the Ensemble Kalman Filter algorithm, using an automated hydrological model in Plaxis. The model was to be created using Python scripting in combination with Plaxis to automate the calculations and to easier process the output and input data through the EnKF.

1.3 Objectives

The main objective of the thesis is to automate calibration of a hydrological model of a slope based on sensor data with the EnKF algorithm. To achieve the main objective, the following specific objectives are implemented:

- Analyze and process sensors data in a wetting period for a slope in Meråker.
- Implement and automate a hydraulic model of the slope in Plaxis with Python scripting.
- Study and Implement the Ensemble Kalman Filter algorithm in the programming language Python Couple Plaxis model with the Ensemble Kalman Filter algorithm in Python. Assess the performance of the Ensemble Kalman Filter algorithm and the calibrated hydrological model.

1.4 Limitations

- One significant drawback of this thesis is the inconsistency of the SWCC derived from the van Genuchten parameters between wetting and drying conditions. The selected period chosen for calibrating the SWCC with sensor values was primarily focused on accurately simulating wetting conditions. The wetting period was selected due to the intention of the hydrological model to be used in predicting the groundwater condition in response to rainfall or snow-melting events as a part of the landslide early warning system. However, true wetting conditions are difficult to obtain over many days in real life.

- The implemented Plaxis model not simulating evapotranspiration, snow melting, effects of vegetation, or temperature effects, which all influence the water conditions in the soil, especially in the upper soil layers. Additionally, there are substantial uncertainties in the ground (e.g., variation of soil conditions with depth) and surface conditions (e.g., varying surface cover, geometry) at the site that were not modeled explicitly. However, the sensors indirectly measure these effects, which means that even though the Plaxis model might yield similar results as the sensors after calibrating the parameters, the premises (i.e., parameter estimation) could be wrong.
- The EnKF algorithm updates the unknown parameters based on the difference between Plaxis output and volumetric water content data. Since the soil water characteristic curve, as modeled by the van Genuchten model, is not uniquely defined by the values of volumetric water content, this means that the calibration process can result in multiple parameter combinations producing near identical outcomes.
- A trial pit was excavated at the site during sensor installation. A thin organic layer was identified at the site, but it was not included in the hydrological model to simplify the implementation. This could give wrong premises to the Plaxis model as soil layering and change in soil parameters with depth are not considered.
- Runtime of the simulations were between 5-10 hours, and this often led to the code failing as it lost contact with Plaxis. This made it difficult to test wide ranges of prior knowledge of the unknown parameters, which possibly could have enhanced the results.

1.5 Structure of the thesis

This master thesis is divided into 6 different chapters with the structure as shown below.

In Chapter 2, the fundamental aspects of unsaturated soil theory are presented in detail. The chapter covers various themes such as the classification of soil layers, capillary theory, flow of water in unsaturated soils, the SWCC, soil suction and the van Genuchten model. The EnKF algorithm is also explained.

Chapter 3 presents the data monitoring system on all levels. The chapter outlines a geotechnical overview of the study area where the sensors are installed. In addition, the framework of the

IoT-system is presented and discussed with regards to calibration and expected sensor values. The sensor data used for the calibration is also presented.

Chapter 4 presents a description of the Plaxis model creation process, including the selection of boundary conditions, mesh, and input parameters. In addition, the implementation of the EnKF-method is included.

In chapter 5 the results from the parameter calibration described in chapter 4 are shown. The chapter discusses the results consecutively. An overall discussion of the results and the model is also provided in this chapter.

Chapter 6 provides a final summary of the thesis and the main conclusions.

Chapter 7 presents suggestions for future work.

2 Theory of unsaturated soil

Parts of this section is sited from (Pedersen, 2022).

The theory chapter of this thesis focuses on the characterization of unsaturated soil behavior and its implications for geotechnical engineering. It challenges the traditional assumption of complete saturation or dryness in soil and highlights the significance of the unsaturated zone in flow problems. The chapter discusses concepts such as capillary forces, matric suction, and effective stresses in unsaturated soil. It also explains the flow of water through soil and the coefficient of permeability. The Soil Water Characteristics curve is introduced as a tool to understand the relationship between water content and soil suction. Various equations for the SWCC are discussed, with a focus on the van Genuchten equation.

2.1 Classification of unsaturated soil

In traditional geotechnical practices, soil water content is commonly considered to be either completely saturated or entirely dry, depending on the water table level. This assumption tends to be conservative for stability and safety issues because elevated pore pressure lowers the effective stresses in the soil. However, when investigating water flow through slopes and examining real soil conditions, this assumption appears to be inadequate. The unsaturated area of the soil is of high interest when working with flow of water as the soil water characteristics change drastically in unsaturated conditions.

(Makonto, 2013) noted that the soil layers can be classified into various categories based on the level of saturation throughout the soil.

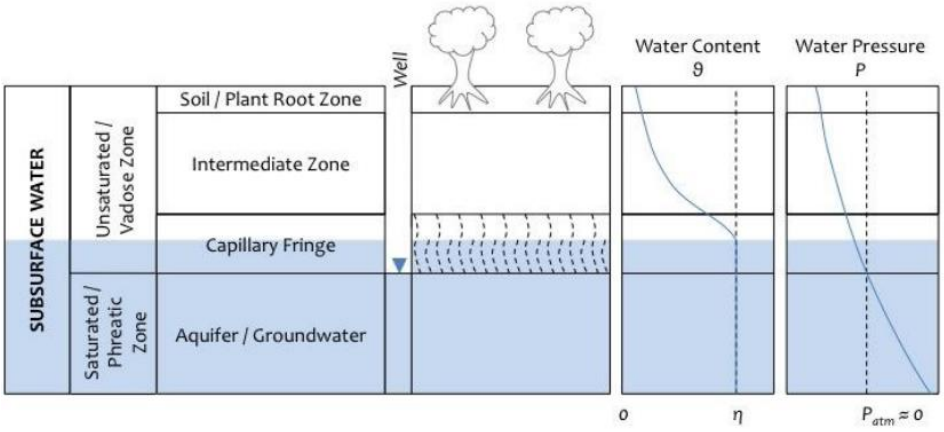


Figure 1: Zone classification of soil for flat ground (Makonto, 2013)

Soil can be categorized into different areas or zones based on the degree of saturation, as illustrated in Figure 1. The portion of the soil that is not fully saturated is called the vadose zone and in this zone the pores between the soil grains are partially filled with water and partially filled with air. The combination of air and water within these spaces affects soil characteristics such as permeability and effective stresses. The interface between air and water in the vadose zone creates suction forces within the soil, which increases the effective stresses. Chapter 2.2 will provide more detailed information on these relations.

The saturated zone beneath the ground water table is called the phreatic zone and here the pores in the soil are fully saturated. Beneath the ground water table, the pore pressures are usually assumed to be hydrostatic and non-negative (no suction). (D. G. Fredlund, 2012)

2.2 Capillary forces and capillary height

In unsaturated soil the pores between the soil grains are partially filled with water and partially filled with air. The boundary between the water surface and the air within the pores of the soil is referred to as the contractile skin. Soil suction appears in these interfaces because of molecules needing to stay in force equilibrium. The water pressure deviates from the air pressure and as a result, tension forces in the contractile skin occur. This is called capillary forces.

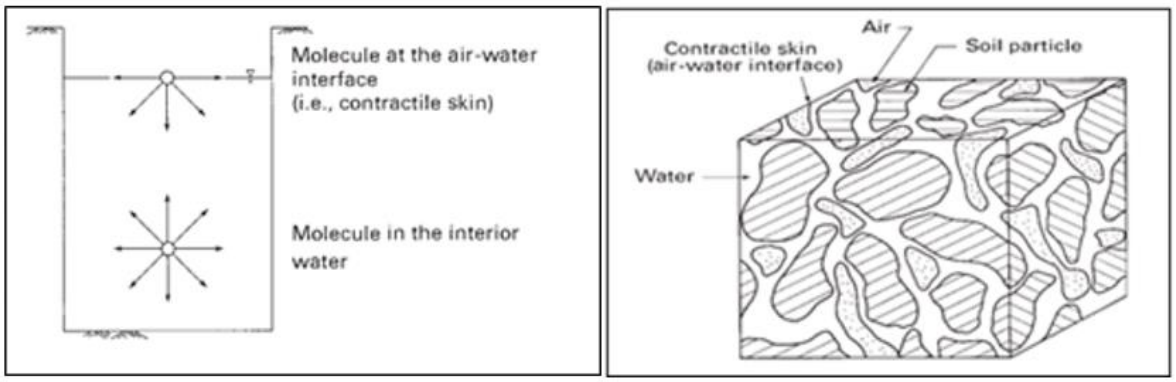


Figure 2: Upscaled element of soil-air-water interfaces and illustration of how tension forces occur in the air-soil interface (D. G. Fredlund, 2012)

The capillary force is often described by the capillary height: the maximum height h_c water can climb in a glass tube. Given by Equation 1:

Equation 1

$$h_c = \frac{2T_s}{\rho_w \, g \, R_s}$$

Where:

T_s= Tension forces

ρ_w = The density of water

g = The gravitational acceleration = 9.81$\frac{m^2}{s}$

R_s = The radius of the tube

When transferring the theory to capillary forces in soil, R_s, is the pore radius of the soil pores and T_s is the tension forces in the soil-air-interface (D. G. Fredlund, 2012).

The tension forces in the air-water interface can be related to the matric suction of the soil as the difference between the pore air pressure and the pore water pressure. Combined with Equation 1, the matric suction can be written as:

Equation 2

$$u_a - u_w = \frac{2T_s}{R_s}$$

(D. G. Fredlund, 2012)

Resulting from the matric suction in the air-water-surface, the effective stresses in the soil increase. The definition of the effective stress in soil was given by (Terazaghi, 1925) and is defined by the total soil pressure, σ subtracted the positive pore pressure, u:

Equation 3

$$\sigma' = \sigma - u$$

However, when working with unsaturated soil, it becomes crucial to determine the stress state within the soil considering both the positive and the negative pore pressures. Bishop's equation, proposed in 1959, is the most widely used formula for this purpose. The equation contains both the total stress/pore pressure-differences and the pore water/pore air-pressure-differences (known as the matric suction).

Equation 4

$$\sigma' = (\sigma - u) + \chi(\mathrm{u_a} - u_w)$$

X is a soil parameter varying based on the degree of saturation in the soil.

Bishop's formula fulfills its purpose of describing the effective stresses with respect to both positive and negative pore pressures. However, the formula contains a soil characteristic parameter, χ , and is therefore not valid as a stress state equation.

When stress dependent failure criterions (such as Mohr Coulomb) are used, the increase in matric suction will lead to a higher degree of mobilization of the undrained shear strength, as the effective stresses increase accordingly.

2.3 Flow of water in soils

In traditional geotechnical practice, the movement of water through soil is explained using Darcy's law and the permeability coefficient, denoted as k_w[m/s]. The coefficient is determined by laboratory tests on saturated soil and is therefore only applicable to soil that is fully saturated. The permeability coefficient of saturated soil is solely dependent on the void ratio. As discussed in earlier sections, assuming that the ground is either completely dry or completely wet is inaccurate and insufficient, particularly when dealing with seepage problems (D. G. Fredlund, 2012).

2.4 Darcy's law

Water flowing through a saturated soil can be described by Darcy's law (1856):

Equation 5

$$v_w = \frac{-k_w(\delta h_w)}{\delta y}$$

Where:

v_w = The velocity of water flowing through a porous medium

k_w= Coefficient of permeability

$\frac{\delta h_w}{\delta y}$ = Hydraulic head gradient

Equation 5 describes the proportional connection between the velocity of water flow in porous media with the hydraulic head gradient, connected by the permeability coefficient.

Experiments done by (Childs, 1950) show that the coefficient of permeability is constant for a given degree of saturation hence the proportionality is valid also for unsaturated soils.

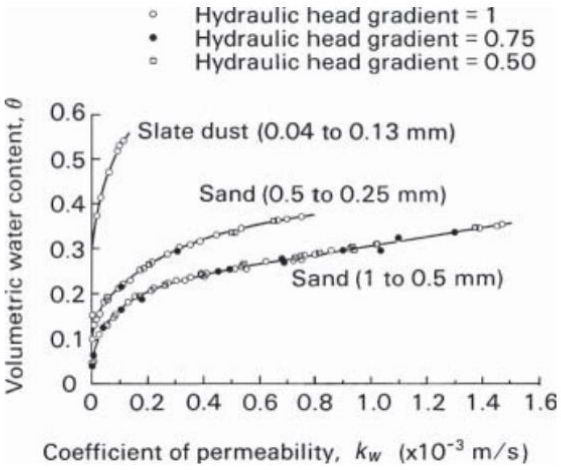


Figure 3: Coefficient of permeability vs. degree of saturation (D. G. Fredlund, 2012)

This means that when working with flow and seepage problems in unsaturated soils, the same principles apply as for saturated soil. The only exception is the coefficient of permeability which for unsaturated soil is a function of more than one variable. This also means that water

will flow from higher hydraulic potential to lower potential gradient, also for unsaturated soil (D. G. Fredlund, 2012).

2.5 The Coefficient of Permeability

The permeability, k, in unsaturated soil theory can be presented as a function of the combination of three different parameters (D. G. Fredlund, 2012):

Equation 6

$$k_w = k_w(S, e)$$

Equation 7

$$k_w = k_w(e, w)$$

Equation 8

$$k_w = k_w(w, S)$$

Where:

S = Degree of saturation

e = Void ratio

w = gravimetric water content

Although permeability in unsaturated soils can be viewed as a function of three parameters, the degree of saturation tends to generate the biggest changes. Therefore, the coefficient of permeability is often viewed as a function of the matric suction only, especially when working with soils with low incompressibility (i.e., Clay) (D. G. Fredlund, 2012).

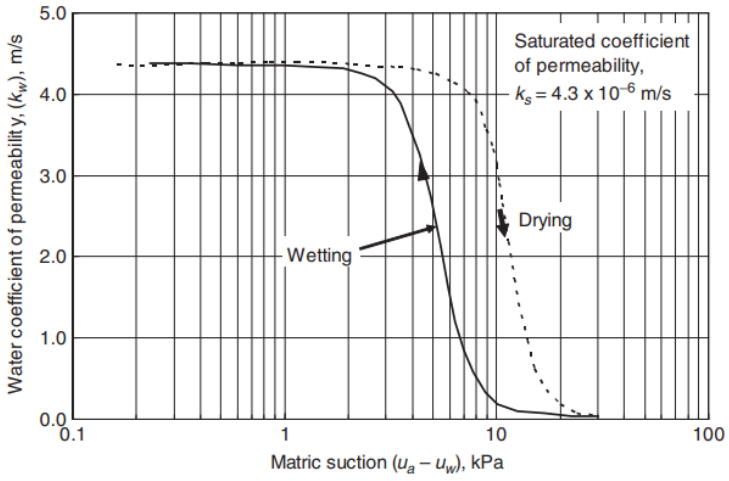


Figure 4: Relation between matric suction and coefficient of permeability for wetting and drying conditions (D. G. Fredlund, 2012)

Water flows though the water filled pores in soils as it gives the least resistance. When the pores are partially filled with air the resistance increases and the velocity of the flow, given by Darcy's law (Equation 5), decreases nonlinearly.

As shown in Figure 4, the coefficient of permeability changes rather quickly as matric suction increases. As desaturation develops and the pores fill up with air, the volume of which water can flow decreases. In addition, the voids shrink because of matric suction and the pathway available for water flow is further reduced. Therefore, the change of permeability in relation to suction is different for wetting and drying. This difference is important to consider when solving unsaturated problems (D. G. Fredlund, 2012).

2.6 Soil Water Characteristic Curve (SWCC)

The correlation between volumetric water content and soil suction can be illustrated using SWCC plots. The SWCC can be determined in a laboratory setting by measuring the difference between air pressure and water pressure using a high-air entry disk, which yields the matric suction. The SWCC can then be plotted along with gravimetric water content to form the curve that characterizes the relationship between the two variables (E.C. Leong, 1997) (W. Scott Sillers, 2001).

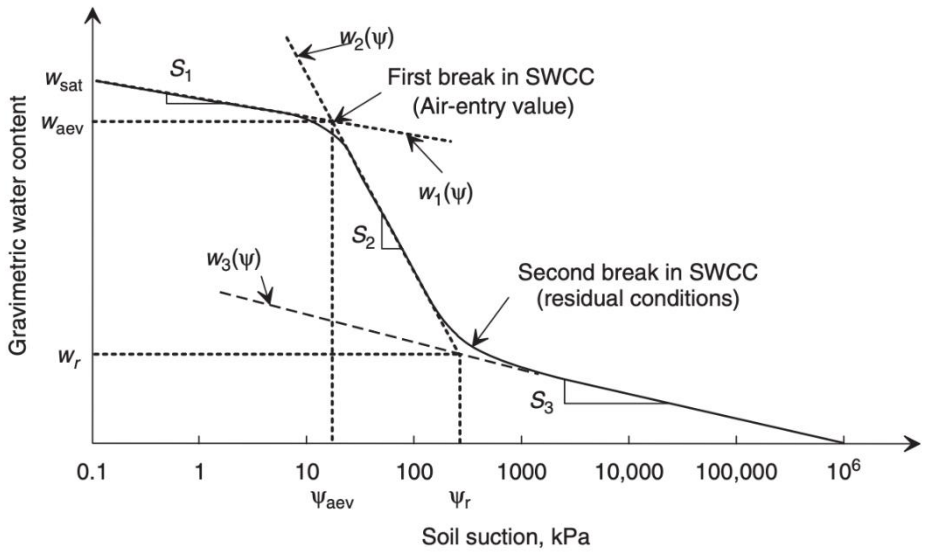


Figure 5: Example of a Soil Water Characteristics Curve with explanations (D. G. Fredlund, 2012)

Figure 5 illustrates an example of a Soil Water Characteristic curve (SWCC) with descriptions of different zones. On the left side of the graph, the soil suction is low, and the gravimetric water content is high, forming the Boundary effect zone. In this zone, the soil pores are predominantly filled with either water or soil (no air) (D. G. Fredlund, 2012).

As the water content decreases, the soil suction increases, and the SWCC transitions to a phase where the pores contain a combination of air and water. The air entry value marks the soil suction where air flows back into the soil pores. In this phase, small changes in volumetric water content leads to significant changes in matric suction because of the pores being less permeable to the water particles (W. Scott Sillers, 2001).

Eventually, the soil reaches residual conditions characterized by high levels of soil suction and low levels of gravimetric water content. In this phase, the pores are mainly filled with air, and water content in minimal, making it nearly impossible for water to flow. The residual water becomes challenging to remove due to its resistance to flow though air-filled pores (D. G. Fredlund, 2012).

2.7 Wetting and drying curves

The SWCC follows a different path for wetting and drying conditions in the soil (D. G. Fredlund, 2012). The difference in wetting and drying curves differs noticeably between soils and comes from factors such as pore layout, contact-angles between water and particles and shrinking/swelling of the soil (T. Mavara, 2018).

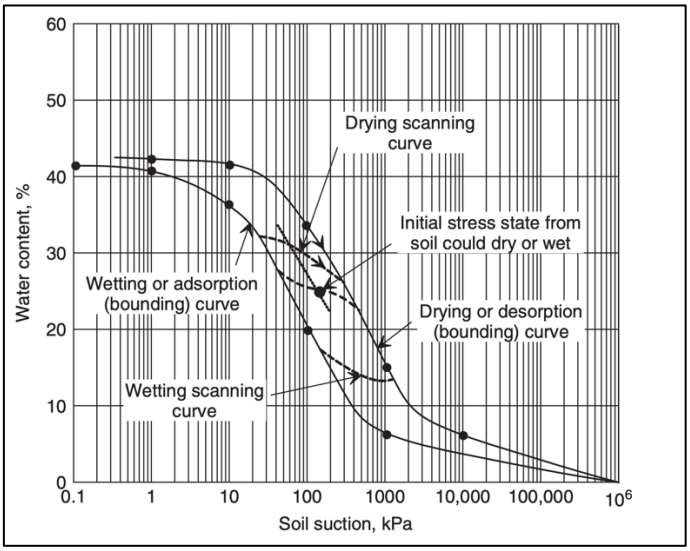


Figure 6: Examples of wetting and drying SWCC (D. G. Fredlund, 2012)

Figure 6 shows examples of the bounding curves for a soil, which are the curves defining the SWCC for the wetting and drying in the initial phases (completely dry and completely saturated

soil). The curves for degrees of saturation between 0% and 100% lay somewhere in between the two bounding curves and form “loops” during wetting and drying cycles. The SWCC can therefore not be uniquely described for a soil, but by defining the bounding curves, a range of soil suction vs water content can be obtained (D. G. Fredlund, 2012) (T. Mavara, 2018).

In this thesis, only the wetting curve will be evaluated as heavy rainfall is the main trigger for water-induced landslides.

2.8 Equations for SWCC

One of the more established variations of the SWCC is the function for the permeability, derived by (R. J. Millington, 1960) and (Mualem, 1976). For this purpose, an equation for the curve has been shown to be beneficial and numerous attempts are made to estimate such an equation. The variety in the derived equations springs from differences in the initial assumptions regarding the pore size distribution of the soil (Robinson, 2019). In the following section, some of these equations and their required inputs are presented, including the van Genuchten equation which is the equation used for flow calculations in this thesis.

2.8.1 Form fitting parameters

To fit the equations, several curve fitting parameters are needed. The form fitting parameters are generally categorized into three different categories (D. G. Fredlund, 2012).

- an a-type parameter related to the air entry value of the soil (i.e., the matric suction of where air begins to enter the pores)
- an n-type parameter related to the rate of desaturation (the “steepness” of the curve)
- an m-type parameter to make the curve more flexible.

Figure 7, Figure 8 and Figure 9 show how the respective parameters affect the (Genuchten, 1980)- and Fredlund and Xing-equations when two of the three parameters are held constant (E.C. Leong, 1997).

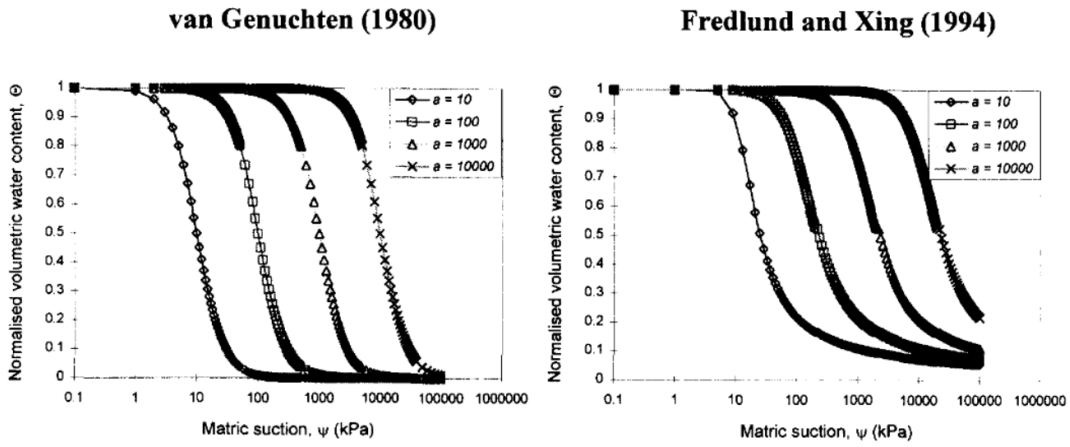


Figure 7: How the a-type parameter affects the SWCC for n = 2 and m = 1

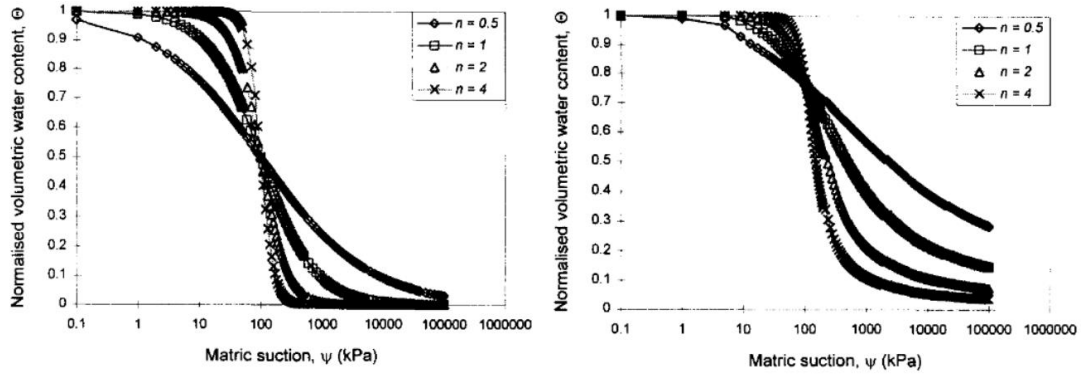


Figure 8: How the n-type parameter affects the SWCC for a = 100 and m = 1

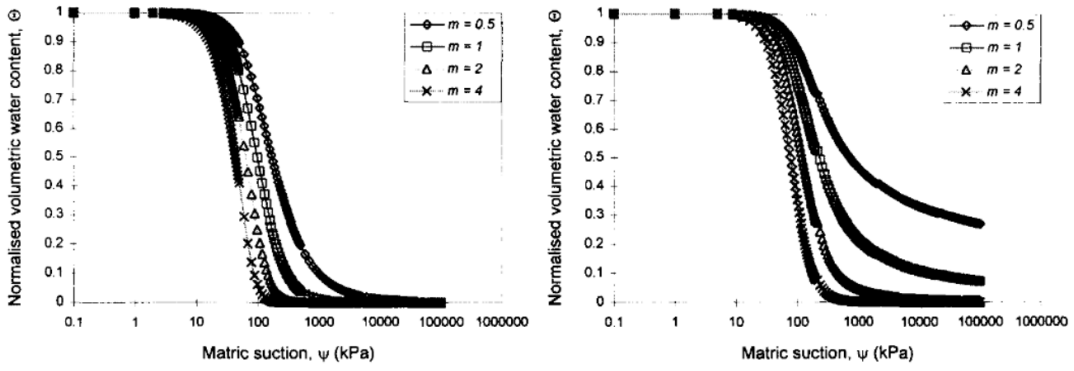


Figure 9: How the m-type parameter affects the SWCC for a = 100 and m = 1

The range of the curve fitting parameters vary greatly. Notably for the van Genuchten equation, the curve is highly sensitive to changes in the n-parameter and less sensitive to changes in the α and m-parameters, compared to the expected range.

Studies have shown that equations based on three form fitting parameters, rather than two, show a greater compatibility to laboratory test results (D. G. Fredlund, 2012).

2.8.2 Brooks and Corey equation (1964)

(R. H. Brooks, 1964) made an equation for the soil suction as a function of the water content by dividing the equation into zones depending on whether the soil suction is higher or lower than the air entry value. For soil suctions lower than the air entry value, the water content is set to be the saturated water content of the soil (D. G. Fredlund, 2012). The equations can be written:

Equation 9

$$w(\psi) = w_s \qquad \psi < \psi_{aev}$$

Equation 10

$$w(\psi) = w_s \left(\frac{\psi}{a}\right)^{-n} \qquad \psi \geq \psi_{aev}$$

Where:

ψ_{aev}= the air entry value

$w =$ water content

$\psi =$ soil suction

The equation is simple to use and is therefore popular to use. However, as noted by (W. Scott Sillers, 2001), the discontinuous nature of the Brooks and Corey model when matric suction exceeds the air-entry value can lead to instabilities during modeling. Additionally, (D. G. Fredlund, 2012) have suggested that while this model is valid for coarser-grained soils that experience rapid changes in moisture content at low suctions, it may be less applicable to finer-grained soils that display a more gradual slope change in the transitional zone.

2.8.3 The van Genuchten Equation

For this project, the (Genuchten, 1980) formula was chosen as the favorable model for the SWCC, mainly because of the possibility of implementation in the finite element program Plaxis. van Genuchten made a continuous closed form equation for the volumetric water vs soil suction content based on three form fitting parameters a, m, and n. From (D. G. Fredlund, 2012)the van Genuchten equation can be written:

Equation 11

$$w(\psi) = \frac{w_s}{[1 + (a\psi)^n]^m}$$

Or re-arranged for soil suction:

Equation 12

$$\psi = \frac{1}{a}\left[\left(\frac{w_s}{w}\right)^{\frac{1}{m}} - 1\right]^{\frac{1}{n}}$$

Where:

w = water content

ψ = suction

a, n & m = curve fitting parameter

Figure 7, Figure 8 and Figure 9 show how the curve fitting parameters of the van Genuchten model influence the curve. The a-parameter is the same as the inflection point of the curve, marking where the boundary effect zone transitions into the transition zone. The parameter does not affect the shape of the function, in contrast to the n-parameter which is closely related to the pore size distribution index and gives the curve slope (W. Scott Sillers, 2001).

The van Genuchten has been combined with (Burdine, 1953) and (Mualem, 1976) to derive a simplified equation for the hydraulic conductivity, by eliminating the m-parameter with the relations $m = 1 - \frac{1}{n}$ (Mualem, 1976) and $m = 1 - \frac{2}{n}$ (Burdine, 1953). Although this simplifies the equation, (W. Scott Sillers, 2001) specifies that the elimination of the m-parameter reduces the flexibility of the curve and that the original three-parameter van Genuchten equation is favorable.

2.9 Ensemble Kalman Filter

In this thesis, the recursive EnKF method was used to calibrate the van Genuchten parameters. The EnKF-method is a formulation of the Kalman Filter method first described by (Evensen, 1994) with the objective of working with nonlinear oceanography problems. The EnKF-model can easily handle non-linear high-dimension problems, thus making it a well-used and acknowledged data assimilation method for physics-based problems where monitored data can be used to approximate parameters of high uncertainty (Muhammad Mohsan, 2021).

The EnKF-method can easily handle large scale, non-linear problems and is well suited for joint parameter-estimation. The method has showed satisfactory results in various fields such as hydrology, meteorology, and geotechnical engineering (Muhammad Mohsan, 2021).

2.9.1 The EnKF algorithm

By assuming a Probability Density Function (PDF) for the unknown parameters, the EnKF-algorithm yields, as for Monte Carlo simulations, numerous random combinations of input variables to the model. By comparing model output values to known data, the unknown parameters will be updated until the difference is sufficiently small. Using the same notations as (Muhammad Mohsan, 2021) the EnKF can be written:

To define the model output, we define the operator g(**z**):

Equation 13

$$\boldsymbol{y} = g(\boldsymbol{z})$$

Equation 14

$$\boldsymbol{z} = (\boldsymbol{x}\ \boldsymbol{\theta})^{\boldsymbol{T}}$$

where $\boldsymbol{y} \in \mathbb{R}^{N_m}$ is the model output in the measurement space and the $\boldsymbol{z}$ -matrix is a combination of the model state and the model parameters (Muhammad Mohsan, 2021).

The measurements are defined by the measurement matrix, $\boldsymbol{d} \in \mathbb{R}^{N_m}$, which also contains the measurement error $\boldsymbol{e}$.

Equation 15

$$\boldsymbol{d} = \boldsymbol{y} + \boldsymbol{e}$$

The main objective of the EnKF algorithm is to maximize the probability $f(\boldsymbol{z}|\boldsymbol{d})$, which means we will get a close-to-reality estimation of the unknown parameters. Given normal

distributed parameters, it can be shown that this is done by minimizing the cost function given by (Muhammad Mohsan, 2021):

Equation 16

$$J(\boldsymbol{z}) = \left(\boldsymbol{z} - \boldsymbol{z^f}\right)^T \boldsymbol{C_{zz}^{-1}} \left(\boldsymbol{z} - \boldsymbol{z^f}\right) + (g(\boldsymbol{z}) - \boldsymbol{d})^T \boldsymbol{C_{dd}^{-1}} (g(\boldsymbol{z}) - \boldsymbol{d})$$

Where $\boldsymbol{C_{zz}}$ represents the error covariance of $\boldsymbol{z}$ and the $\boldsymbol{z^f}$-matrix symbolizes the previous estimate of the model parameters. $\boldsymbol{C_{dd}}$ is the error covariance of the known data (measurements).

Minimizing the cost function is done by setting the derivative to zero and the solution is a set of equations called the Kalman Filter (Muhammad Mohsan, 2021):

Equation 17

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

Equation 18

$$\boldsymbol{C_{zz}^a} = (\boldsymbol{I} - \boldsymbol{KG})\boldsymbol{C_{zz}}$$

Where:

The notation "a" represents the new estimate while the notation "f" indicates the prior estimate. The K-matrix is called the "Kalman gain" and is defined:

Equation 19

$$\boldsymbol{K} = \boldsymbol{C_{zz}}\boldsymbol{G}(\boldsymbol{G}\boldsymbol{C_{zz}}\boldsymbol{G^T} - \boldsymbol{C_{zz}})^{-1}$$

For each ensemble member in the sample, we can write the equations:

Equation 20

$$z_i^a = z_i^f + \boldsymbol{K^e}(\boldsymbol{d}_i - g(z_i^f))$$

Equation 21

$$\boldsymbol{K^e} = \boldsymbol{C_{zz}^e}\boldsymbol{G}(\boldsymbol{G}\boldsymbol{C_{zz}^e}\boldsymbol{G}^T - \boldsymbol{C}_{dd})^{-1}$$

Where:

$\boldsymbol{C_{zz}^e}$ is the combined covariance matrix for the model state parameters and $\boldsymbol{d}_i = \boldsymbol{d} + \boldsymbol{\epsilon}_i$ is the measurement matrix with added "noise" (Muhammad Mohsan, 2021).

$\boldsymbol{C_{zz}^e}$ is calculated by first finding the ensemble mean for the Ne sample realizations and then the difference between each sample member and the mean value:

Equation 22

$$\overline{\boldsymbol{Z}_t^f} = \boldsymbol{Z}_t^f \boldsymbol{I}_{Ne}$$

Equation 23

$$\boldsymbol{Z}_f' = \boldsymbol{Z}_f^t - \overline{\boldsymbol{Z}_t^f}$$

Where $\boldsymbol{I}_{Ne}$ is a matrix containing only a factor $\frac{1}{Ne}$. The combined state error covariance is then defined:

Equation 24

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

C_{zz} includes the covariances of both the state and parameter errors, along with the cross-covariance that exists between the state variables and parameters.

The measurements are stored in the $\boldsymbol{d}_i$-vector with the added measurement error or "noise" added to each ensemble member. The total measurement matrix is then defined:

Equation 25

$$\boldsymbol{D}_t = (\boldsymbol{d}_{1,t}, \quad \boldsymbol{d}_{2,t}, \quad \boldsymbol{d}_{3,t}, \quad \ldots\ldots \boldsymbol{d}_{Ne,t})$$

Based on these matrices, the main parameter updating function can be written:

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

The function shows how the prior parameters**,** $\boldsymbol{Z}_t^f$ are updated with the Kalman Gain. The last factor of the equation (i.e., the difference between sensor data and Plaxis outputs) weights the parameter update, meaning that if the difference is zero, the parameters are not updated as the model is identical to real life (Muhammad Mohsan, 2021)**.**

3 Data monitoring system

This chapter provides context for this thesis by presenting the data monitoring system installed as a part of the KlimaDigital study. In addition, the study site's geological and geotechnical properties are presented to give a better understanding of the problem and enhance readers' comprehension of the expected results. Furthermore, the chapter outlines the process flow of the data monitoring system and the IoT-sensors, offering insight into the data collection process by including a sensor data overview.

As part of the KlimaDigital project, sensors were strategically installed in various areas within the study site. This thesis focuses on utilizing data from sensors located in location 1-Kvernbekkneset. Consequently, this location is examined in greater detail, particularly concerning its geology and geotechnical properties.

3.1 Study site

The study site for the KlimaDigital project is a 200km^2 area located along the Stjørdal river in Trøndelag. This area has a long history of being at high risk for rainfall induced landslides due to its steep slopes and heavy annual precipitation between 964mm and 1205mm (Leiva, 2019).

The study area was chosen for sensor installation for the KlimaDigital project based on a series of factors, which some are:

- A heatmap of landslide events
- Areas above the marine limit, as glacial till and moraine were favorable soil types.
- Driving distance < 2h from Trondheim
- Thickness of deposits around 2-3m (shallow slopes)
- Steepness of slopes > 25^o
- Easily accessible, not too dense vegetation.

(Ivan Depina E. O., 2021)

Based on these criteria, two areas in Meråker were chosen as suitable for sensor installation, Kvernbekkneset and Kjelberget, situated alongside E14.

Figure 10: Location 1 - Kvernbekkneset

Figure 11: Location 2 – Kjelberget

Figure 12 show a Digital Elevation Model (DEM) made by (Leiva, 2019). The figure shows that the elevation of the region increases rapidly along the Stjørdal river/E14 road from 0-200m in the blue to green areas of the DEM, indication steep slopes. The slope map shows that the steepest slopes are concentrated along the river, with steepness varying from 25-50 degrees.

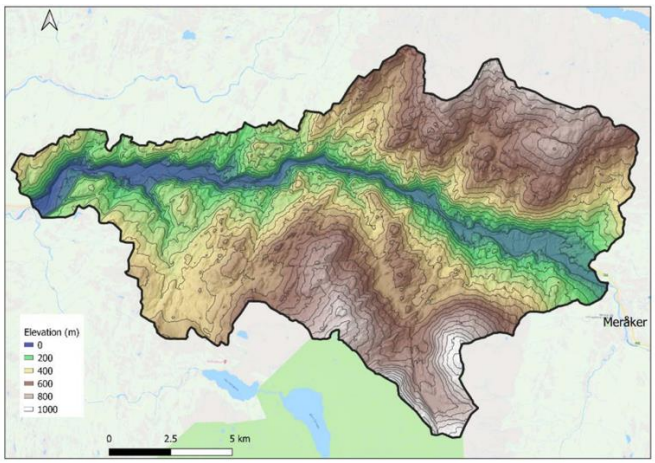


Figure 12: Elevation model for the study site. (Leiva, 2019)

3.2 Geotechnical properties

Figure 13 show the assumed soil deposits of the study area, made by the Norwegian Geological Institute (NGU). As seen from the figure, the area is dominated by weathered rock material (purple) and shallow till/moraine deposits (green), with marine deposits occurring only in small "pockets" in the lower areas. Along the Stjørdal river the ground conditions are primarily fluvial deposits (yellow).

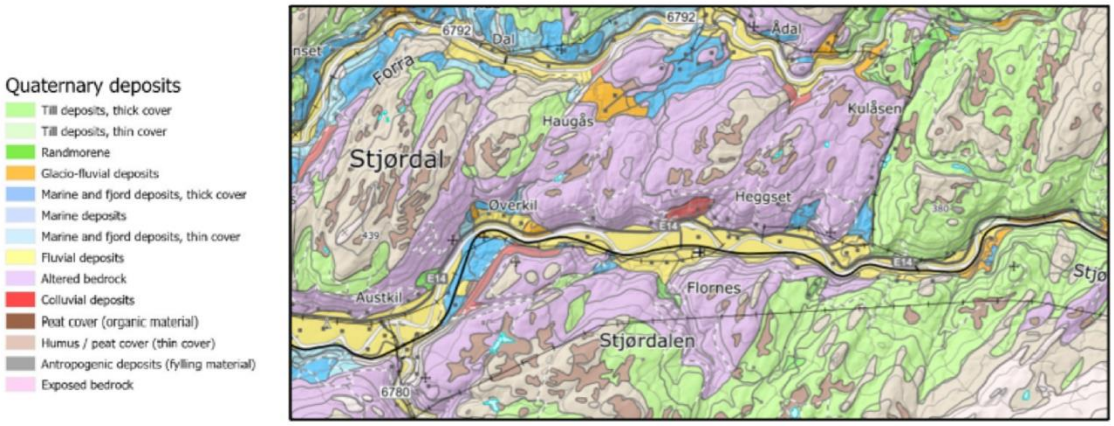


Figure 13: Deposits in the Meråker-area (NGU)

In the eastern sector of the area, the till/moraine occurs mainly as shallow deposits. Consequently, exposed bedrock is common in this area. Moraine material, which is deposited and transported by glaciers, is typically an unsorted and hard packed mixture of different soil materials that can contain everything from clay to sand and small rocks. Because of its hard-packed and unsorted nature, infiltration rates in these deposits are typically low, because the soil leaves little room for water flow *(Ivan Depina E. O., 2021)*.

On the other hand, fluvial deposits found along the Stjørdal rover consists of well-rounded gravel and sand and are often more loosely packed and well sorted because they are transported and deposited by flowing water. As a result, these deposits typically have a higher permeability. These deposits vary in depth from 0,5m to 10m deep *(Ivan Depina E. O., 2021)*.

3.2.1 Geology at sensor location – Kvernbekkneset

The soil in location 2 – Kvernbekkneset mainly consists of shallow moraine deposits (green) with an organic layer on top. The terrain is steep with a visible channel along the Rabb-river (Ivan Depina E. O., 2021).

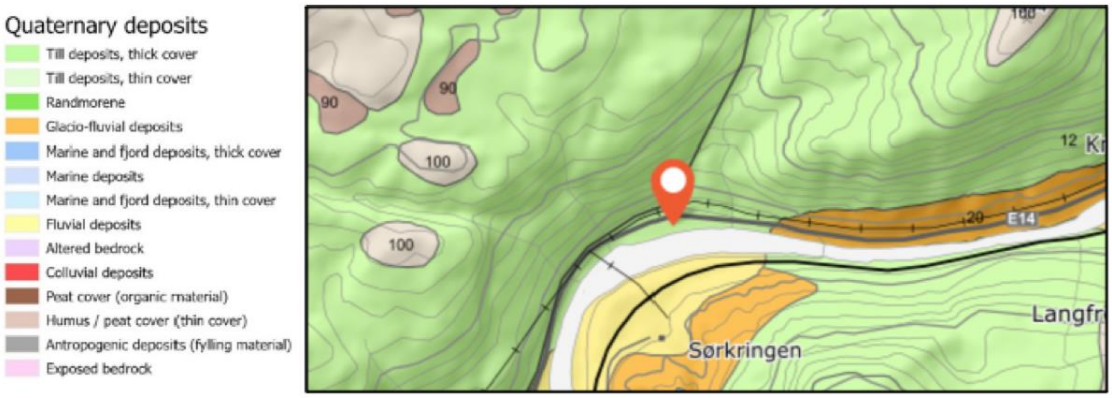


Figure 14: Deposits map location Kvernbekkneset (NGU)

Some laboratory testing was done in this area for the KlimaDigital project. These tests encompassed methods to determine water content and organic content, sieve analysis and hydrometer analysis for soil classification, pycnometer test for determining soil density as well as a large-scale direct shear box test (Ivan Depina E. O., 2021). Figure 15 shows the results from the grain-size distribution tests and the large-scale shear box test:

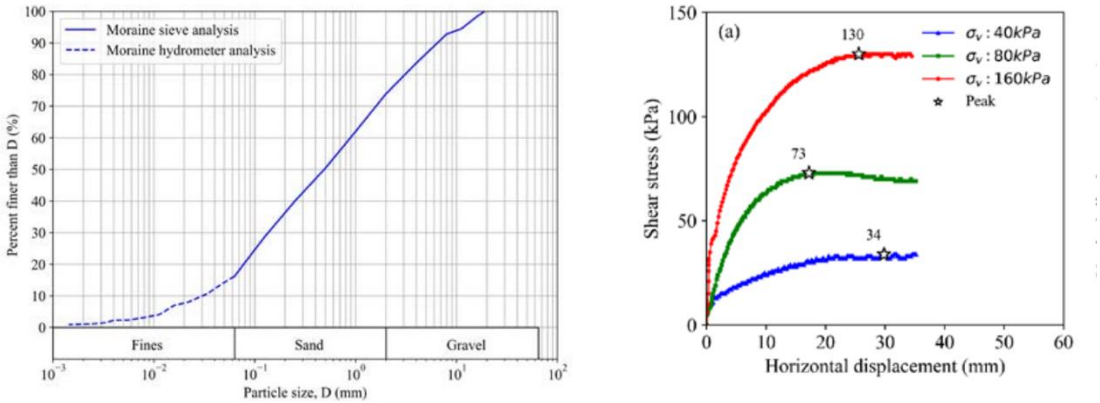


Figure 15: Grain size distribution and results from shear box test at sensor location (Ivan Depina E. O., 2021)

The soil has been classified as silt sand in line with the European Soil Classification System, containing 16% fines, 57.7% sand and 26.1% gravel. The results from the large-scale shear box test correspond to a friction angle 38.2^{o} and cohesion 5.5kPa (ISO, 2017) (Ivan Depina E. O., 2021).

3.3 Landslide susceptibility

The study area is situated within the landslide-prone region known as "Trøndelagskysten" (Ivan Depina E. O., 2021), where the occurrence of landslides and debris flows is prevalent. In addition to the ground conditions discussed in the previous chapter, the combination of the area's topography, characterized by natural steep slopes, and its wet climate with high precipitation rated and snowmelt, contributes significantly to this susceptibility.

According to Kjell Hauge from NGI, the chances for rainfall-induced landslides substantially increases when the daily rainfall reaches 6% of the annual precipitation (Rommetveit, 2008). In the case of the study area, this equates to a daily range of 57mm/day to 72mm/day. Figure 16 derived from data provided by NVE, illustrates all recorded landslide events in the area since 1964 in relation to the corresponding daily rainfall. This data, referenced from (Leiva, 2019), demonstrates a discernible connection between daily rainfall levels and the incidence of landslides.

#	Date	Rainfall (mm/day)	Slide type
1	2016-12-05	41.10	Landslide
2	2015-02-09	50.10	Unspecified landslide
3	2012-03-28	63.10	Landslide
4	2012-03-28	29.80	Unspecified landslide
5	2012-03-28	29.80	Landslide
6	2012-03-23	63.10	Mudflow
7	2012-03-23	63.10	Unspecified landslide
8	2012-03-23	63.10	Unspecified landslide
9	2012-03-12	13.00	Unspecified landslide
10	2012-03-12	13.00	Landslide
11	2011-08-16	71.70	Unspecified landslide
12	2011-08-16	71.70	Unspecified landslide
13	2007-03-16	6.10	Unspecified landslide
14	2006-09-26	0.60	Debris flow
15	2003-06-05	0.00	Unspecified landslide
16	2000-08-09	43.40	Unspecified landslide
17	2000-08-08	43.40	Unspecified landslide
18	1992-01-16	39.90	Unspecified landslide
19	1992-01-14	39.90	Unspecified landslide
20	1983-08-28	22.40	Unspecified landslide
21	1983-01-21	30.00	Unspecified landslide
22	1964-02-07	44.00	Unspecified landslide

Figure 16: Registered landslide events by NVE in Meråker, Trøndelag (Leiva, 2019)

A total of 93 mass movements, including rock falls, landslides, snow avalanches, stone slides, debris flows, and clay slides, have been documented in the entire area since 1750. Among these events, approximately 30 were classified as landslides triggered by intense rainfall and snow melting (NVE, 2023) (Ivan Depina E. O., 2021).

Figure 17 show a map of some of the registered events along the Stjørdal river.

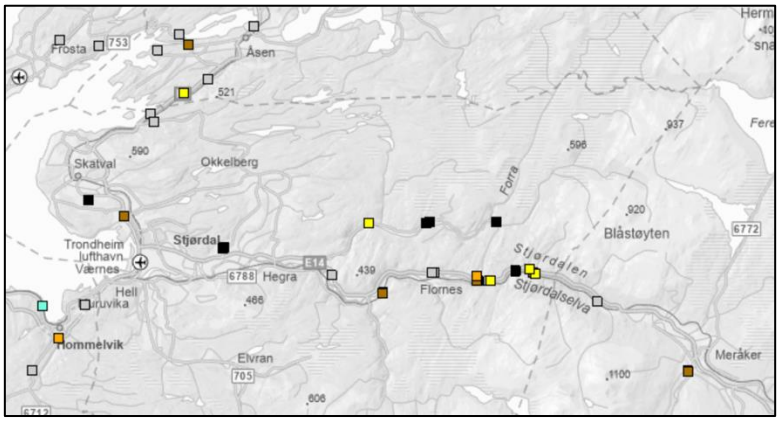


Figure 17: Registered landslide events in the study area

In Figure 17 the yellow dots mark landslides triggered by floods or heavy rainfall. The orange and brown dots mark clay slides and unspecified landslides, respectively.

3.4 Data Monitoring System

As discussed in the previous section, heavy rainfall and snow melting is the main trigger for water-induced landslides. An important part of an early warning system therefore must be analyzing how the inflow influences the hydrological conditions of the soil. Hence: to calibrate hydrological parameters, hydrological sensor data is needed.

For the calibration of van Genuchten parameters in this thesis, sensor data of volumetric water content (VWC) from Kvernbekkneset was used. In this section, the IoT-sensor system is presented, first with and overview and further with more detailed explanation. In addition, the sensor data used for parameter calibration are reviewed and explained.

3.5 Framework

The framework of the system consists of five layers as shown in Figure 18: the device layer, network layer, platform layer, application layer and the user layer.

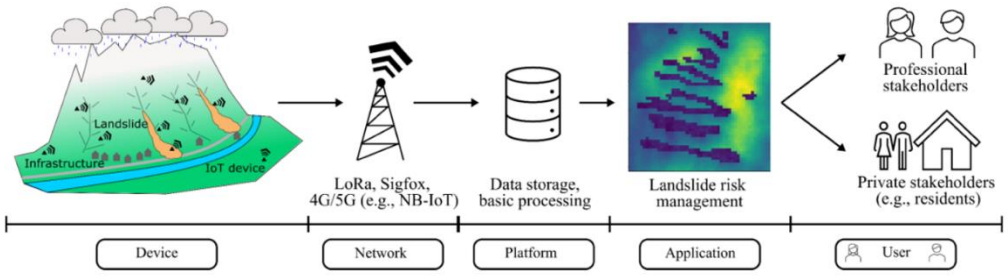


Figure 18: Illustration of the IoT-system (Ivan Depina E. O., 2021)

The device layer forms the foundation of the system and consists of sensors strategically placed across two distinct slopes within the study site. These sensors utilize Narrowband-IoT-technology to accurately measure temperature, pore pressures, volumetric water content, and suction. Additionally, a weather station is deployed in the study area to monitor precipitation. The sensor data is stored in network clouds within the Network layer (i.e., the internet), with the values being updated every fifteen minutes and made accessible though the Platform layer (Emir Ahmet Oguz, 2021).

This master's thesis primarily focuses on the data processing aspect of the system, which corresponds to the Application layer. Within this layer, the sensor data is processed and used to address geotechnical challenges, mainly in relation stability and landslide susceptibility. The User Layer includes all private and non-private stakeholders using the system.

3.6 IoT devices and calibration

For this thesis, data from IoT sensor 5, monitoring VWC was used. The sensors are placed at location Kvernbekkneset with coordinates as shown in Table 1 (D2.1). In this area, three suction sensors and three volumetric water content sensors at depth 0.3m, 0.5m and 0.9m were installed in each point (D2.1 and D2.2). The sensors also monitor ground temperature. A weather station monitoring precipitation was installed in location 1 (Kjelberget) (Emir Ahmet Oguz, 2021).

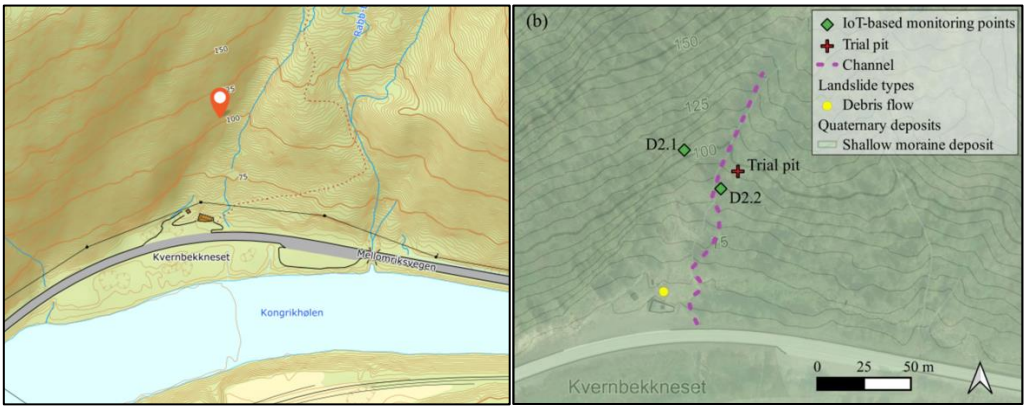


Figure 19: Exact location of IoT-sensors at Kvernbekkneset. (Emir Ahmet Oguz, 2021)

Table 1: Coordinates of IoT-sensors

	X	*Y*
IoT-device 5	63.466612° N	11.461331° E

The monitoring sensors for VWC provide raw data in microvolt (mV) format, which needs to be converted into VWC. Figure 20 shows the equation for the conversion from the manufacturer in addition to a conversion equation found by laboratory testing done for the KlimaDigital project done by (Emir Ahmed Oguz, 2021).

For this project, the calibration equation provided by the sensor manufacturer was used for simplicity.

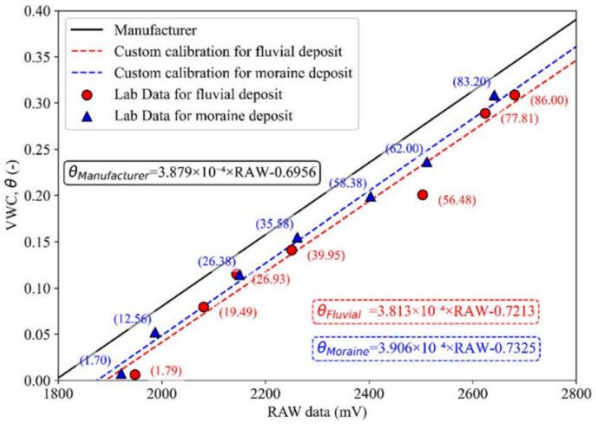


Figure 20: Calibration equations for converting mV to VWC (Emir Ahmet Oguz, 2021)

Soil suction sensors were not utilized for parameter calibration in this thesis. The decision was based on the observation of low suction values recorded by the sensors during the selected time period. The sensor data was specifically collected to simulate wet conditions, characterized by significant precipitation. Due to the intense rainfall during this period, the suction values recorded by the sensors fell below the calibration range, rendering the suction data unreliable and inaccurate.

3.7 Sensor data

Figure 21 shows sensor data of volumetric water content and precipitation from device 5 during a time period from August 2020 to October 2021.

Figure 21: Sensor data for device 5. Graph from (Ivan Depina E. O., 2021)

As we see from the figure, the impact of precipitation is characterized by rapid increase in VWC for all depths. It is interesting to note how the VWC decreases with depth for the spring/summer and fall. This is mainly due to vertical infiltration in the area (Ivan Depina E. O., 2021) and the only deviation from this is during the winter months (January to March) where the upper soil layers are frozen. Is this section we can see that the VWC of depth 0.3m dips below the VWC at depth 0.5m and 0.9m.

For this thesis, data from the last 58 days of this data set was used to calibrate the hydrological parameters of the soil. The data are presented in more detail in Figure 22.

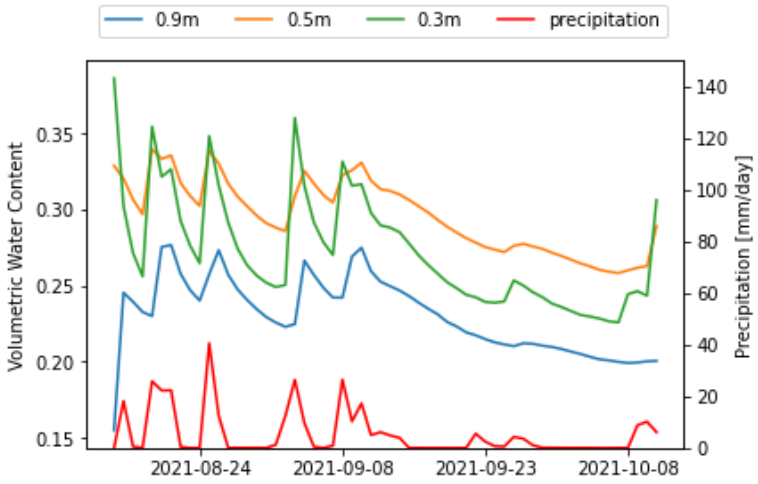


Figure 22: VWC and precipitation plots for IoT-device 5. A selection of this data was used to calibrate the can Genuchten parameters in this thesis.

Figure 22 show a detailed plot of the VWC for depths 0.3m, 0.5m and 0.9m alongside the precipitation. As mentioned above, the data shows a decreasing degree of saturation with depth. All sensors show a rapid response in VWC with regards to the precipitation. The effect is most

prominent in depth 0.3m but sensors at 0.5m and 0.9m show similar tendencies though to a lesser degree.

For the calibration of the van Genuchten parameters, the first 10 days of the data in Figure 22 were chosen. The selection was done based on two main factors:

- This period is characterized by heavy rainfall and is therefore the best option to replicate wetting tendencies in the soil.
- The responses in the VWC (i.e., the "spikes" in the graphs) helps the EnKF-algorithm estimate the unknown parameters as it feeds the model information about how the soil responds to changes in precipitation.

4 Methodology

In the following section, the main idea and method for the project model is presented. The chapter is divided into blocks for each part of the model. First, an overview of the system is presented describing the interactions between the model components: the IoT sensors, the Plaxis model, and the EnKF algorithm. In the rest of the chapter, the process is described in more detail, including all the steps taken to achieve the results. The Plaxis model is presented first with the description of the input parameters and boundary conditions/initial conditions, which is followed by the presentation of the EnKF model and its implementation. Lastly, the choices for number of iterations and prior knowledge distributions for the simulations are presented.

4.1 Model Overview

Figure 23 illustrates an overview of the model calibration with the EnKF algorithm. The algorithm is initiated by specifying a "prior state" for the parameters to be calibrated. The prior state represents any prior knowledge that we may have about the unknown parameters (e.g., literature, prior experience from similar projects). The prior state is specified by assigning a probabilistic distribution to the unknown parameters that represents the likely range of unknown parameter values. The parameters of the distribution are typically selected to allow for a wide range of values (e.g., large variance) due to lack of knowledge of the parameter values. Once the prior distributions are defined, Ne = 20 random samples for each of the unknown parameters are generated from the specified probability distribution for each parameter. In this thesis, the parameters are considered independent. However, in a general case, dependencies between the parameters may be simulated.

Subsequently, Ne Plaxis calculations are performed to generate outputs of effective degree of saturation at three different depths that correspond to the sensor depths in the slope. The Plaxis output is converted by the formula:

Equation 26

$$\theta = S_{eff}(\theta_{sat} - \theta_{res}) + \theta_{res}$$

to match the sensor data where:

$\theta =$ volumetric water content

$S_{eff} =$ effective degree of saturation

$\theta_{sat}, \theta_{res} =$ saturated volumetric content and residual volumetric water content, respectively.

These outputs are then compared to the known VWC values obtained from sensor data using the Ensemble Kalman Filter (EnKF) algorithm. Normally distributed measurement error is added to the sensor data randomly with zero-mean and standard deviation of 0.02.

Based on the covariance and the disparity between the sensor data and Plaxis output, the samples of the unknown parameters are updated. The updated samples are then used as input for a new set of Ne = 20 Plaxis calculations, and the entire process is iterated Nm times, where Nm represents the number of sensor realizations.

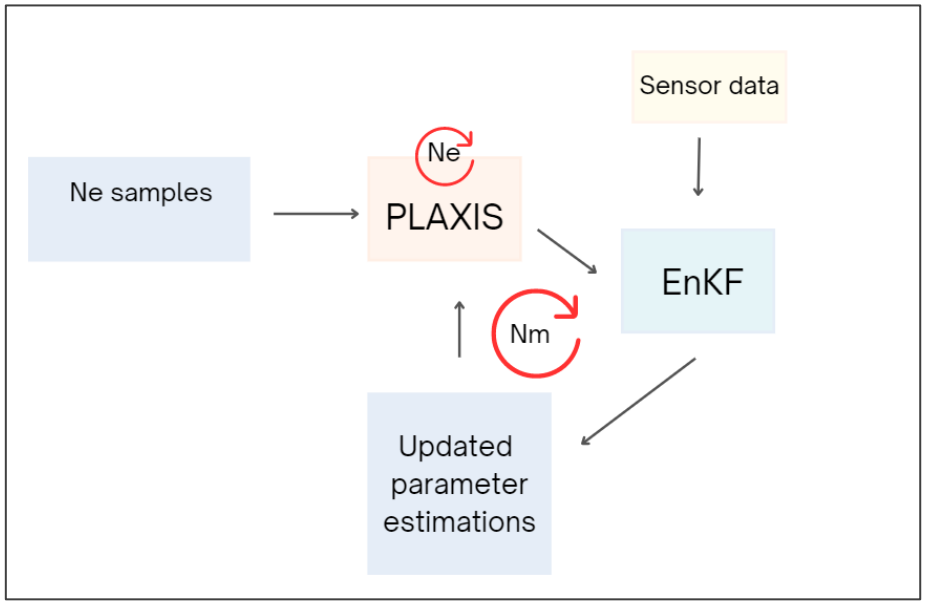


Figure 23: Model Illustration. First, Ne parameter samples are generated, based on the prior distributions of the unknown model parameter. Plaxis performs Ne calculations (one for each sample), and the output is converted to Volumetric Water Content to be comparable to the sensor data. The Plaxis output is then compared to the sensor data in the EnKF algorithm, and the values of samples are updated with the algorithm favoring the values that result in smaller difference between model predictions and sensor measurements. Then Plaxis performs Ne new calculations with the updated parameters. The whole process is repeated Nm times, making the total of calculations (Nm x Ne).

4.2 Plaxis model

The Plxscripting Python-package was used to write the Plaxis program in Python. By doing this, the Plaxis model could be calculated automatically numerous of times.

The code begins with opening and creating a Plaxis model-file and creating a slope by using polygons and lines. The slope length of 20m was initially determined considering the characteristics of the study area. However, it should be noted that while the study area does have long slopes, the actual water contribution area is dependent on the topography, and it may differ from the model geometry The water flow from higher up the slope might be channelized in ravines, and it may not significantly affect the groundwater conditions in the slope. Therefore, it might not be realistic to incorporate an excessively long slope in the model.

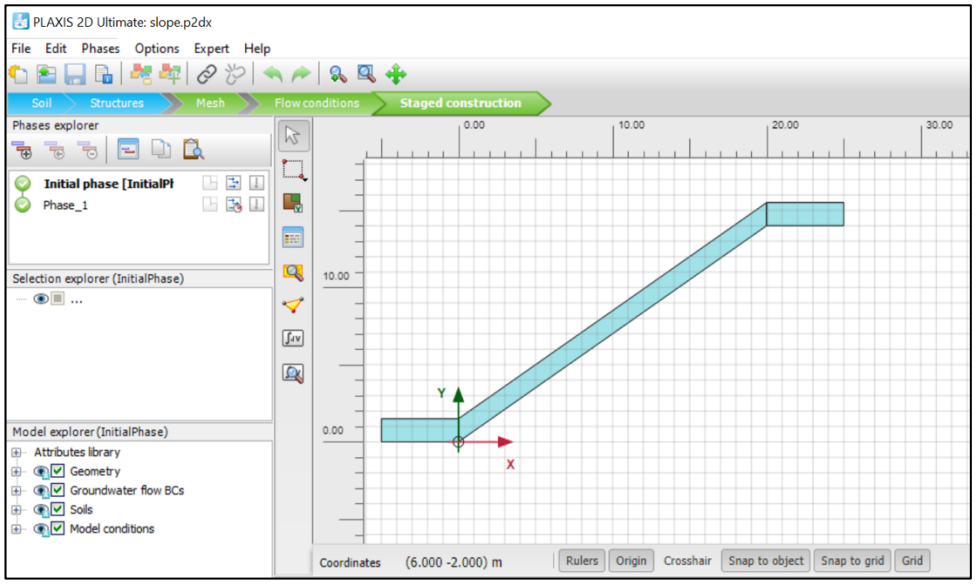


Figure 24: Slope model in Plaxis

Two areas at the top and bottom of the slope are modelled to prevent substantial outflow of water on the left and right boundaries and ensure slope-parallel flow from the top of the slope and to the bottom of the slope.

Both the depth to bedrock and slope angle was scripted as variables, making it easy to adjust these parameters later if needed. Based on test runs of the model and the elevation map constructed by (Leiva, 2019) (Figure 12), the slope angle was set to be 35 degrees for all the simulations. The depth to bedrock was set to be 1.5m based on trial pits that were dug during the installations of sensors.

4.2.1 Material definition

For simplicity, one material was defined for the whole depth of the slope. Adding soil layers would double the unknown parameters in the algorithm, which would increase the runtime and complexity significantly.

The code continues by defining a “make material”-function, assigning hydrological properties to the soil, taking in the unknown van Genuchten parameters and saturated permeability as inputs. For van Genuchten implementation into Plaixs, the parameters were converted as described on Bentley Communities by (Khan, 2022):

- ga is defined by dividing the parameter a by the unit weight of water $ga\ [\frac{1}{m}] = \frac{a}{10}$

- $gn = n$
- $gc = m = 1 - \frac{1}{gn}$ $(automatically\ implemented\ in\ Plaxis)$
- $gl = 1.250$

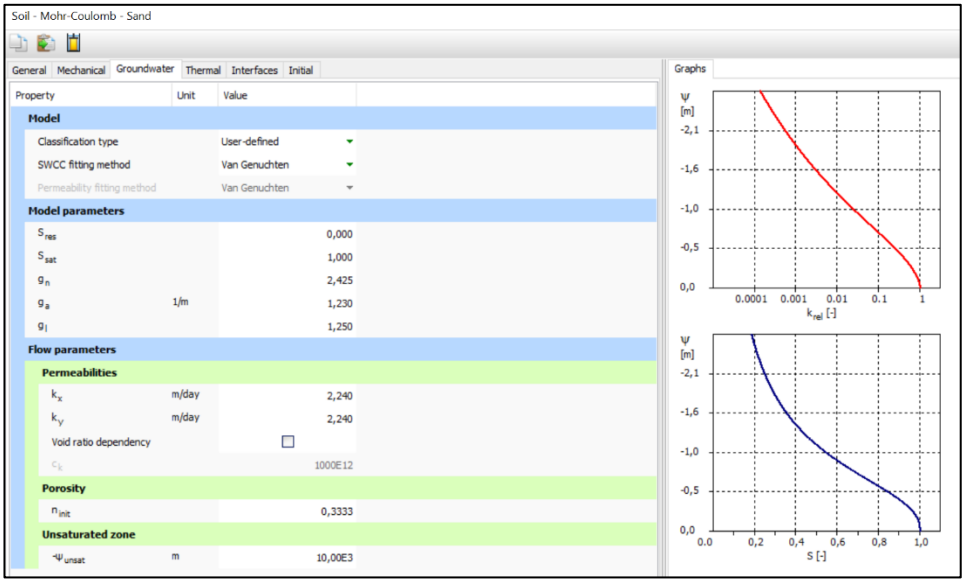


Figure 25: Material Input example

Plaxis utilizes the van Genuchten parameters and the saturated permeability to display the SWCC. While the saturated permeability k_x is not directly inputted into the van Genuchten equation, Plaxis leverages this parameter to estimate a relative permeability k_{rel}, thus establishing the unsaturated permeability as a function of the saturated permeability and suction.

The gl parameter did not impact the shape of the SWCC significantly in this study and was therefore set to default value 1.250 for the sake of simplicity in the implementation of the algorithm.

4.2.2 Mesh and selected nodes

Based on trial calculations, a fine mesh was required to obtain a good solution to the infiltration process in the model. If the mesh is too coarse, Plaxis struggles with calculating smooth transitions between zones of different saturations in the slope.

The model mesh is set to 0.01, generating 2074 elements and 17313 nodes.

To simulate the IoT sensor placement, three nodes were chosen at depths of 0.3m, 0.5m and 0.9m from the top of the soil in the middle of the slope.

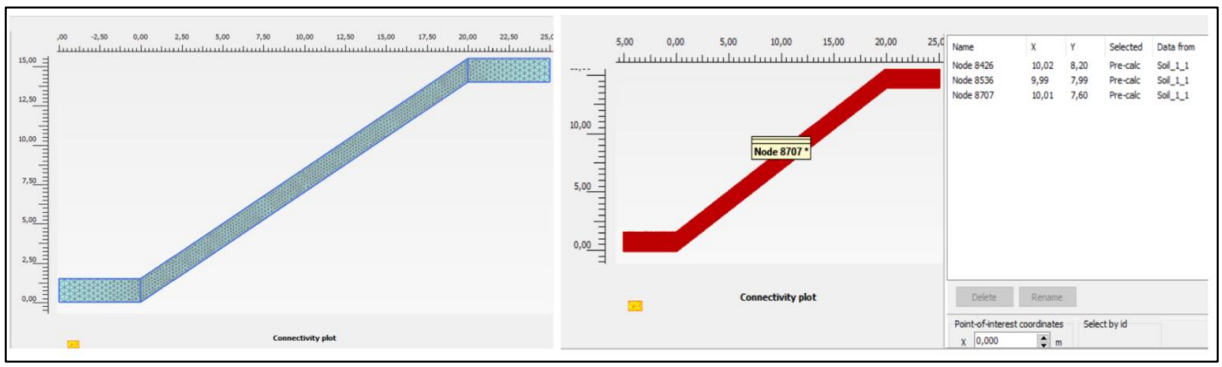


Figure 26: Mesh and node points from Plaxis

4.2.3 Infiltration and boundary conditions

The lower boundary of the model is assumed to be bedrock and it is specified as a closed boundary for the whole slope. The left and right sides of the slope are defined as “seepage” boundaries, while the upper boundary is specified as “infiltration”. To estimate the initial groundwater condition, a steady-state analysis was conducted with the infiltration on the upper boundary defined by the average annual rainfall of around 1200 mm/year.

The initial steady-state analysis is followed by a transient analysis. In the transient analysis the infiltration on the upper boundary of the model is specified to replicate the actual rainfall during the considered time-period. This was implemented with a Plaxis discharge function, which was defined using real precipitation data from the monitoring system. The discharge function is defined as a table with daily rainfall rates from day 0 to 58. Because of wetting and drying hysteresis in the SWCC, a period with heavy rainfall over several days was preferable. When looking at the precipitation graph in Figure 27, the first 12 days of the period of 58 days has several days with rainfall > 20mm/day and this period was chosen based on this.

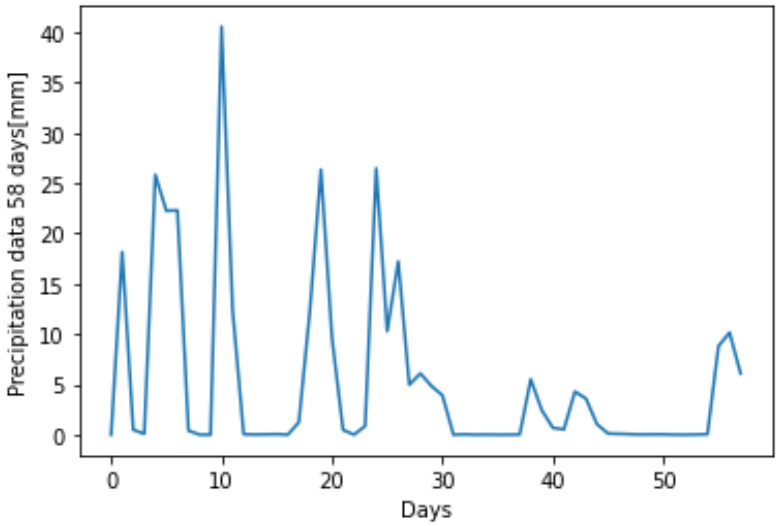


Figure 27: Precipitation from sensor data day 0-58

The initial phase is constructed to define the initial groundwater conditions in the slope. In this phase, the precipitation is sat to be 0.005 m/day with a steady-state groundwater flow analysis. In phase 1, a transient groundwater flow analysis is done with the time interval as an input value from the Python script. During the first iteration $Nm = 0$, the time interval is $t = Nm + 1$ and subsequently $t = 10$ in the last iteration. In combination with the discharge function, which is detailing the precipitation during the 10 days, this ensures realistic infiltration conditions for the model automatically when running the code.

4.3 Ensemble Kalman Filter implementation

The EnKF-algorithm begins with estimating Ne = 20 realizations or samples for each of the unknown parameters, n, α and k_{sat} based on a lognormal distribution with a given expected value μ and standard deviation σ. A measurement error σ_m= 0.02 is also defined in the initial stage.

The expected value and standard deviation are defined as the system's "prior knowledge" and was varied for each simulation to investigate how the results were affected. A lognormal distribution was used for the parameters to avoid negative numbers in the samples, which could lead to Plaxis-errors and numerical fails in the simulation. The parameters were transformed using e^x before Plaxis input.

Plaxis then runs one calculation for each value in the sample, in total Ne = 20 calculations in one iteration.

The Plaxis model outputs a matrix of effective saturation (which is converted to volumetric water content promptly by using Equation 26) in the three node points (0.3m, 0.5m and 0.9m). These results are the model output $\boldsymbol{y}$ of the system, $\boldsymbol{y} = g(\boldsymbol{z})$ where the $g(\boldsymbol{z})$-function represents the Plaxis calculations. The $\boldsymbol{z}$-matrix is the system state of Plaxis defined by model state $\boldsymbol{\theta}$ (i.e., output of interest) and unknown parameters $\boldsymbol{x}$ (i.e., van Genuchten parameters). In the code, the parameter samples $\boldsymbol{x}$ were noted $\boldsymbol{z_f}$ for prior parameter estimates and $\boldsymbol{z_a}$ for updated estimations.

In the code, a $\boldsymbol{Y}$-matrix was defined to contain all unknown parameters and Plaxis outputs:

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{\alpha_i} \\ \boldsymbol{k_{sat}} \\ \boldsymbol{n} \\ \boldsymbol{\theta_1} \\ \boldsymbol{\theta_2} \\ \boldsymbol{\theta_3} \end{bmatrix}$$

Where the three first elements are the parameter vectors, and the three last elements are the Plaxis output vectors at each depth. The elements of the $\boldsymbol{Y}$-matrix are then copied over to the $\boldsymbol{z}^f$-matrix, which only contains prior parameter estimations.

Further, the EnKF algorithm works by calculating the sample covariance $\boldsymbol{C_{zz}} = \frac{\boldsymbol{z_d}(\boldsymbol{z_d})^T}{Ne-1}$ of the Ne parameter samples, where $\boldsymbol{z_d}$ is the difference between the sample and the sample mean.

The measurement matrix $\boldsymbol{D}$ contains the sensor data, and the measurement error σ_m= 0.02 is added to each element in the matrix to account for any noise in the sensor data. The parameter updating process starts by first calculating the difference between measurements and Plaxis output:

$$difference = \boldsymbol{D} - \boldsymbol{Gz}$$

And then updates the parameters based on the difference:

$$\boldsymbol{z}^a = \boldsymbol{z}^f + \boldsymbol{K} * difference$$

Where K is the Kalman gain, defined by:

$$\boldsymbol{K}^e = \boldsymbol{C}_{zz}^{\boldsymbol{e}}\boldsymbol{G}(\boldsymbol{G}\boldsymbol{C}_{zz}^{\boldsymbol{e}}\boldsymbol{G}^T - \boldsymbol{C}_{dd})^{-\mathbf{1}}$$

As presented in chapter 2.9.1.

If there is no difference between the measurements and model predictions, the samples will not be updated as it means that we have good model performance. If there is a difference between the model predictions and sensor measurements, the samples of the unknown parameters are updated based on the calculated Kalman gain. The size of the Kalman gain and its relation to predictions or observations depends on the relative size of the parameter covariance matrix and the measurement error covariance matrix. If the measurement error covariance is small relative to the parameter covariance, that means that our measurements are very accurate, and we tend to trust them more than predictions. In case the parameter covariance is small, relative to the measurement error covariance, that means that we have very little variability in model predictions and that the model is likely to be accurate and that we can trust it more than the measurements. Finally, towards the end of the calibration the size of the parameter covariance matrix is likely to become very small as we are converging to certain values of the unknown parameters, while the error covariance matrix will remain constant. This means that uncertainty in the estimated parameter values cannot be fully eliminated as there will always remain some uncertainty originating from the measurement error.

4.4 Model input

The number of simulations and iterations was determined as follows:

Ne = number of sample guesses

Nm = Iterations or number of measurements. In this case one measurement for each day for 10 days was used.

After conducting multiple tests runs of the simulations, it was observed that using $Ne = 20$ parameter guesses yielded satisfactory results without compromising the runtime and computer capacity. While some simulations were attempted with $Ne = 30$, it did not sufficiently improve parameter convergence and significantly increased the likelihood of Plaxis not converging and crashing.

$Nm = 10$ days was chosen based on the precipitation graph in Figure 27, which displayed the highest variability within this timeframe. Extending the simulations beyond this period was deemed unnecessary, although one simulation with Nm = 12 is presented in the results for comparison reasons. Moreover, the test simulations demonstrated rapid convergence of the

unknown parameters, and prolonging the simulation period would only add more time to the process, without yielding better outcomes.

This resulted in a total of $Ne\ x\ Nm$ = 20 x 10 = 200 calculations for each simulation. The runtime for each simulation varied between 5 - 10 hours.

The initial probability density functions (pdf's) of the unknown parameters were given as input for each simulation, all other parameters and conditions were kept the same. The unknown parameters were modelled as lognormally distributed to avoid negative values of the parameters. The pdfs were chosen based on expected values from the literature and the results from the test simulations.

5 Results and discussion

In this chapter, the outcomes and analysis of the research are presented. The results were obtained by varying prior knowledge of the parameters to evaluate changes in convergence, posterior estimates in the calibration process, and the stability of the algorithm. The volumetric water contents from the Plaxis analysis over Nm iterations (days) are shown in comparison to the to the sensor data for the representative depths. The compliance between sensor data and Plaxis results are the main indicator of good parameter estimations as the goal is to establish a model as close to real-life as possible.

The results from each simulation are discussed consecutively as they are presented. Lastly, a more collective discussion is presented to further compare the results and give an understanding of how the model performed as a whole.

5.1 Results overview

The overview of the prior estimate vs the posterior estimate for the unknown parameters is presented in Table 2. The table show the expected values and standard deviation of the initial guesses (black numbers) compared to the expected values and standard deviation obtained from the data assimilation process (blue numbers) after Nm = 10 iterations.

Table 2: Overview of results. The table show the prior distribution estimate of the unknown van Genuchten parameters a, n and k_{sat} in comparison to the posterior estimates obtained by the data assimilation process.

Simulation	μ_a Prior	μ_a Posterior	σ_a Prior	σ_a Posterior	μ_n Prior	μ_n Posterior	σ_n Prior	σ_n Posterior	$\mu_{k_{sat}}$ Prior	$\mu_{k_{sat}}$ Posterior	$\sigma_{k_{sat}}$ Prior	$\sigma_{k_{sat}}$ Posterior
i	6.0	6.73	1.0	0.885	2.0	1.887	0.5	0.108	4.0	5.64	2.0	1.358
ii	6.0	2.58	1.0	0.36	3.5	2.325	0.5	0.137	4.0	2.68	2.0	0.365
iii	6.0	5.465	1.0	0.358	3.5	2.378	0.5	0.156	5.5	2.525	2.0	0.368
iv	2.0	1.561	1.0	0.121	3.5	3.034	0.5	0.185	4.0	3.724	2.0	0.771

5.2 Results simulation i)

Figure 28 (Plot i.A) - i.C)) exhibit Plaxis ensemble results compared to the corresponding sensor data for simulation i). The sensor data is represented by the blue line, while the x-marks depict the ensemble Plaxis calculation results.

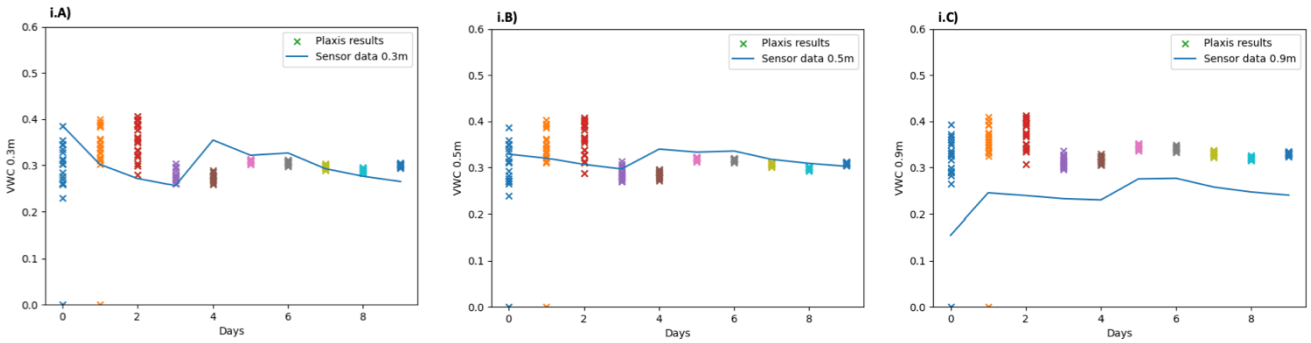


Figure 28: Results from simulation i). Plaxis data of VWC (x) plotted against sensor data (blue lines) for all depths.

Plots i.A) and i.B) corresponding to depths 0.3m and 0.5m, respectively, demonstrate a notable overall alignment between the sensor data and Plaxis results, with the Plaxis outcomes converging to nearly identical values as the sensor data in both cases. However, it appears that Plaxis exhibits a slower response to increases in VWC, as is seen by the observed gaps between sensor-plots and Plaxis-plots in i.A) and i.B) between day 4 - 5.

A reason for this could be that EnKF over-estimates the k_{sat} parameter which leads to the water draining too fast for day 2-4. As is seen in Figure 29 (graph i.E) and i.H)), which show the evolution of the estimation of k_{sat} during the 10-day simulation period, the permeability estimations increase during these days before decreasing noticeably afterwards, resulting in the Plaxis output not keeping up with the rapid increase in VWC at day 3-4, which leads to deviations between Plaxis and the sensor data.

Plot i.C), representing a depth of 0.9m, shows less conformity with the sensor data, as is seen by the Plaxis outputs consistently deviate from the sensor data with around 0.05%. There are several possible reasons for this deviation:

As the sensor at depth 0.9m consistently shows lower VWC than sensors at 0.5m and 0.3m for the whole period, there is reason to believe a low-permeable soil layer is starting somewhere between depth 0.5m and 0.9m. (Ivan Depina E. O., 2021) also suggest that the low values of VWC at depth 0.9m is partially due to vertical infiltration in the slope. A combination of these

theories is supported by the geotechnical investigations, concluding with the area being dominated by steep slopes and hard packed moraine with poor infiltration (chapter 3.2.1).

However, a permeability-paradox occurs when analyzing unsaturated soil in Plaxis. As the water drains from a soil, the permeability decreases as soil suction increases. If the permeability is highly sensitive to changes in soil suction (i.e., high values of n), this could lead to a rapid decrease in permeability caused by the high saturated permeability.

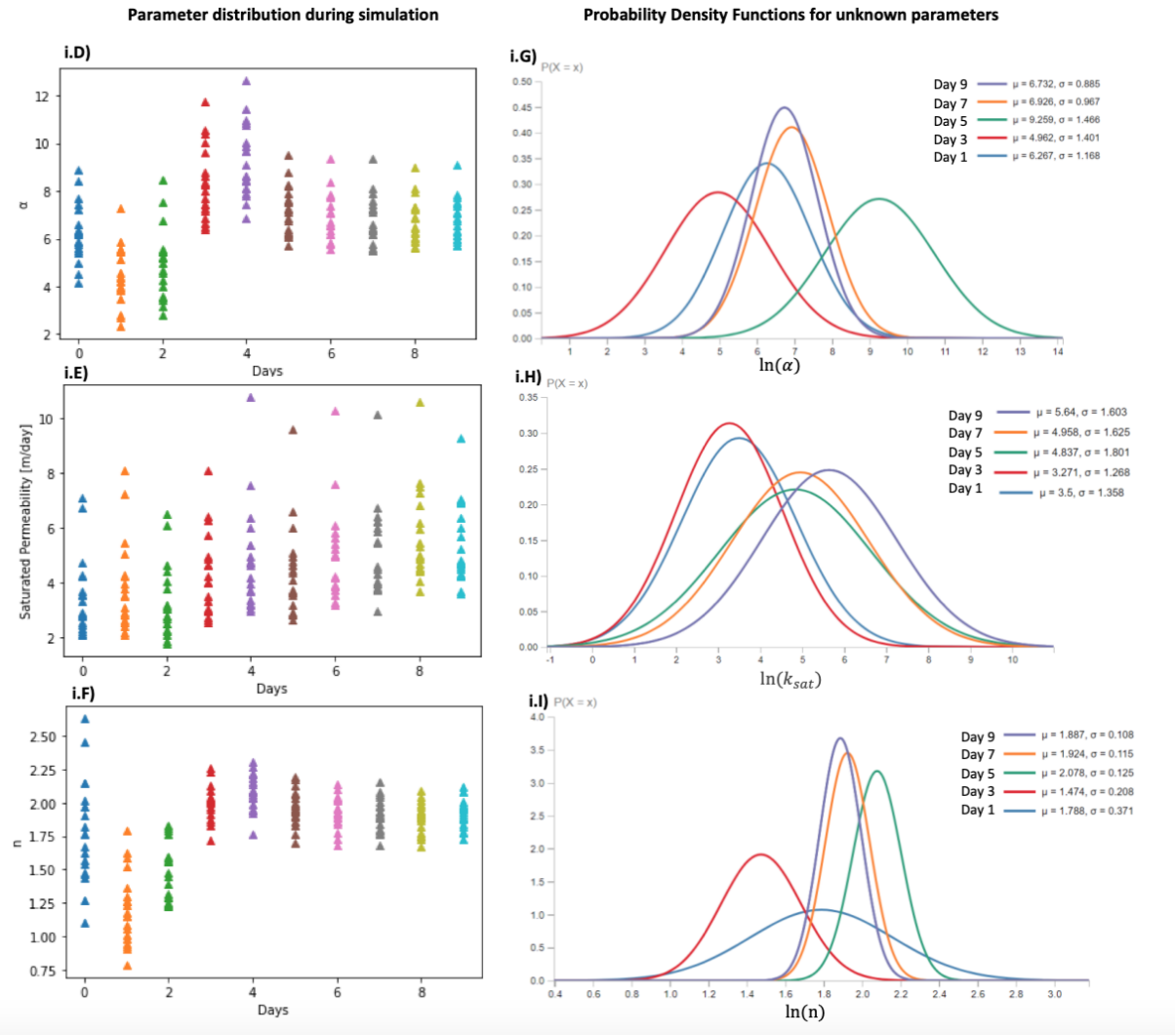


Figure 29: Results from parameter estimation (i). Prior estimate: μ_α= 6.0, σ_α = 1.0, $\mu_{k_{sat}}$ = 4.0, $\sigma_{k_{sat}}$ = 2.0, μ_n = 2.0, σ_n = 0.5

Figure 1 (Graph i.D) – i.F)) shows the parameter estimations during the 10-day simulation and graphs i.G) – i.I) show the probability distribution of the parameters from every other iteration during the simulation where day 1 indicate the "prior knowledge" of each parameter.

The α-parameter shows a slight convergence tendency from prior state, seen by the standard deviation changing from $\sigma_{\alpha,prior}$ = 1.0 to $\sigma_{\alpha,posterior}$ = 0.885, indicating a more accurate estimation. The n-parameter converges significantly and is therefore the most accurate estimate in this simulation with $\sigma_{n,prior}$ = 0.5 to $\sigma_{t=posterior}$ = 0.108.

The results for the saturated permeability k_{sat} don't demonstrate any satisfactory convergence as the standard deviation only changes from $\sigma_{k_{sat},prior}$ = 2.0 to $\sigma_{k_{sat},posterior}$= 1.358. However, graph i.A) – i.C) show a significant convergence of Plaxis outputs. There are a few reasons this could be:

The model is relatively insensitive to the change in permeability, as the results doesn't seem to change much although the permeability is widely spread between values 3 – 9m/day.

However, when we investigate the SWCC obtained from Plaxis using the expected values for the posterior estimates of the van Genuchten parameters, we can see that the permeability in fact changes substantially due to soil suction.

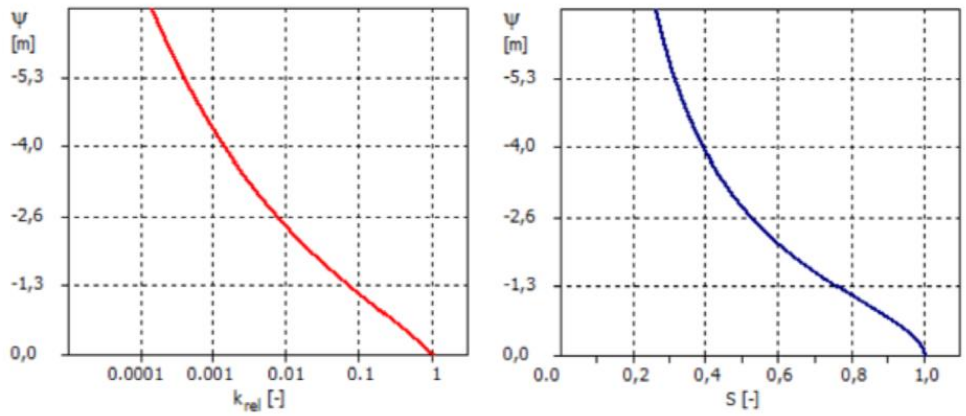


Figure 30: SWCC from Plaxis using the estimated van Genuchten parameters. Ga = a/10 = 0.673, gn = 1.887, gl = 1.125, k_{sat} = 5.64 m/day

Here $k_{rel} = \frac{k}{k_{sat}}$ is defined as a percentage of the saturated permeability and ψ is the soil suction. Looking at graphs i.A) – i.C) we see that the VWC from Plaxis outputs lays in the range of 0.3%, which, by using Equation 26, translates to a degree of saturation S ≈ 0.7%.

Reading off the $S - \psi$ –plot in Figure 30 we see that this corresponds to ψ ≈ -1.5 [m] → $\frac{k}{k_{sat}}$ = 0.055 → $0.055 * k_{sat} =$ 0.31 m/day, which is a significant percentage loss from k_{sat}.

Compared to the SWCC from the other simulations, the SWCC from simulation i) shows a slower response in k_{sat} to changes in soil suction, due to the n-parameter being in the lower range.

Another reason for the high variability in k_{sat} could also be the relatively high variability shown in the estimation of the two other parameters α and n. Although both parameters seem to converge to high certainty estimates the SWCC is, as discussed in chapter 2.8.3, highly sensitive to changes in the n-parameter. Therefore, the reason could be that the highest estimates of k_{sat} (≈9m/day) were paired with high values of n in the simulation, resulting in k_{rel} changing quickly for small changes of soil suction and vice versa.

To conclude, the parameter estimations obtained from simulation i) are not particularly reliable. Although the sensor data and Plaxis outputs have good compliance for depth 0.5 and somehow good compliance in depth 0.3m, the posterior distributions of the parameters are not sufficient to describe the hydrological conditions of the soil in a unique and satisfying way.

5.3 Results simulation ii)

Graph ii.A – ii.C) show Plaxis outputs (x-marks) plotted against sensor data (blue line) for simulation ii).

For depth 0.5m (ii.B)), the Plaxis outputs are almost identical to the sensor data which is a good indicator of reliable parameter estimations at this depth.

For depth 0.3m, the Plaxis data show an overall good compliance to the sensor data, yet there is a significant deviation in Plaxis' response to increase in VWC, especially during day 3-5. This deviation is likely a result of a higher permeability layer in the top layer of the soil, which makes EnKF slightly underestimate the saturated permeability k_{sat} to obtain better results at depth 0.5m and 0.9m, which show signs of lower permeability as the responses to increased VWC are slower.

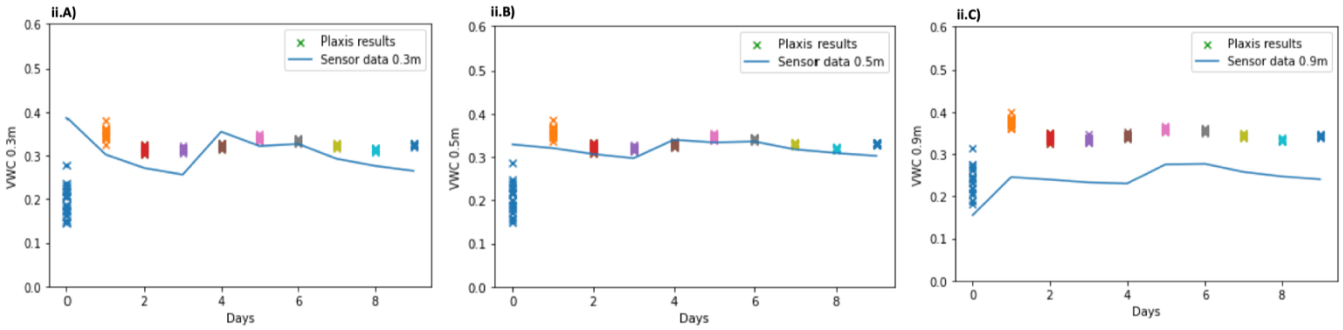


Figure 31: Results from simulation ii). Plaxis data of VWC (x) plotted against sensor data (blue lines) for all depths.

As for the problem in simulation i), the VWC from Plaxis is overestimated for depth 0.9. The same explanations as for simulation i) are valid for this case.

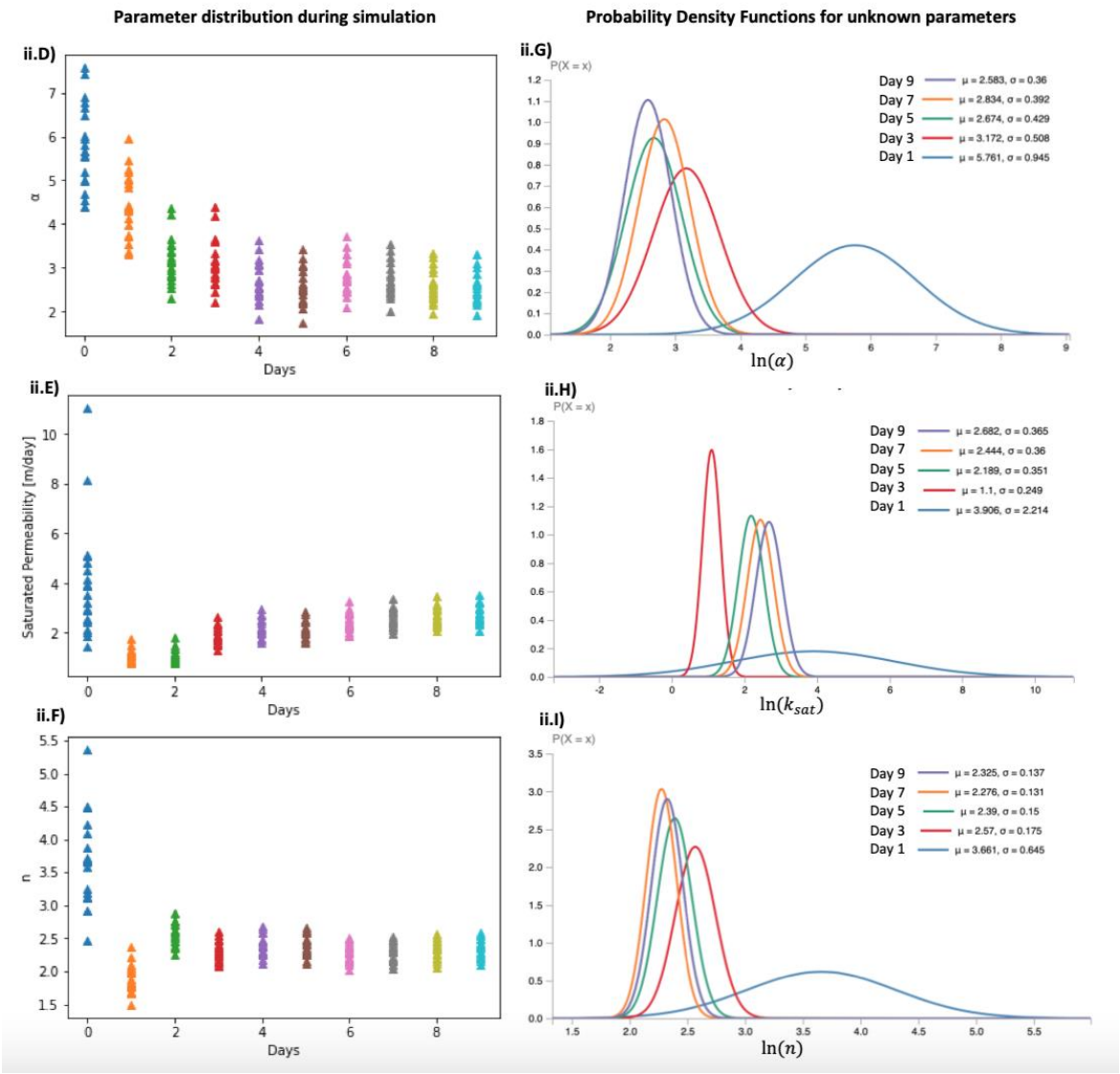


Figure 32: Results from parameter estimation (ii). Prior estimate: μ_α = 6.0, σ_α = 1.0, $\mu_{k_{sat}}$ = 4.0, $\sigma_{k_{sat}}$ = 2.0, μ_n = 3.5, σ_n = 0.5

Figure 32 shows the results from the parameter estimations for simulation ii) (graph ii.D) – ii.I)).

The α-parameter (graph ii.D) and ii.G)) show a considerable amount of improvement in standard deviation, with $\sigma_{\alpha,prior}$= 0.945 and $\sigma_{\alpha,posterior}$= 0.36, which is a decrease by 1/3. Also k_{sat} show a great improvement in standard deviation during the simulation with $\sigma_{k_{sat},prior}$ = 2.214 and $\sigma_{k_{sat},posterior}$ = 0.365. Lastly, the n - parameter converges to expected value μ_n = 2.325 with a decrease in standard deviation from $\sigma_{n,prior}$ = 0.645 to $\sigma_{n,posterior}$ = 0.137.

Compared to the overall results from simulation i), simulation ii) show a higher degree of certainty in the posterior parameter estimations. This is mainly due to the factor of convergence in the estimations. Decreasing standard deviations over time means that the expected area for the parameter estimation is shrinking, resulting in increased certainty.

It is especially important to note the difference in estimated k_{sat} for simulation i) and ii) as it varies considerably between the two simulations ($k_{sat,i}$ = 5.64 m/day vs $k_{sat,ii}$ = 2.682 m/day). The effect of this is seen in the VWC-plots (ii.A – ii.C and i.A – i.C) as the Plaxis output shows a slower response to increase in water content for case ii) vs i).

The difference in the α and n – parameter for the two simulations is also interesting to mark as α_i = 6.732 and α_{ii} = 2.583 while n_i = 1.887 and n_{ii} = 2.325. As discussed in chapter 2.8.1, the effect of the α – parameter in the van Genuchten equation is a shift in the air-entry value of the function, describing for what matric suction value air can start to fill the pores.

This means that for the same degree of saturation (S), the SWCC for simulation ii) in Figure 33 yields higher values of matric suction than in simulation i) (Figure 30).

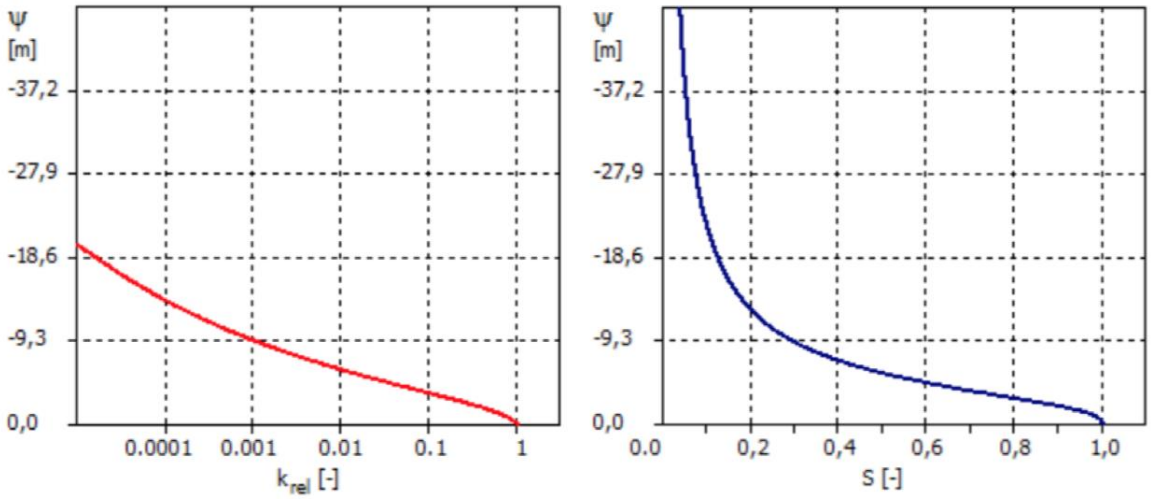


Figure 33: SWCC from Plaxis using the estimated van Genuchten parameters. Ga = a/10 = 0.258, gn = 2.325, gl = 1.125, k_{sat} = 1.68 m/day.

The difference in the α-parameter also show notable differences in the relationship between k_{sat} and $k_{non-saturated}$. Doing the same rough calculation as for simulation i), the VWC obtained from Plaxis lays in the range of 0.3%, which translates to degree of saturation at ≈ 0.7% (Equation 26). For the SWCC in Figure 33, the soil suction corresponding to S = 0.7%

is approximately ψ = 4 which again translates to $k_{rel} = \frac{k}{k_{sat}} \approx 0.1$. With $k_{sat,ii}$ = 1.68 m/day as estimated from the simulation. This means that the *average* non-saturated permeability for the soil k = 0.1 x k_{sat} = 0.168 m/day. Compared to simulation i), where the average non-saturated permeability was calculated to be $k = 0.31$m/day, the difference in non-saturated permeability for the two cases is significantly smaller than for the saturated permeability. This is important to note as it highlights how the differences in numerical results between simulations may appear bigger than the actual effect.

Simulation ii) show overall somewhat satisfying results, with good correspondence between sensor data and Plaxis as well as convergence for all the unknown parameters. When compared to simulation i) the results however show how the SWCC is not uniquely described and that a variety of parameter combinations can yield similar model results.

5.4 Results simulation iii) and iv)

The inclusion of these additional results serves the purpose of demonstrating the impact of prior knowledge and iteration time on the outcomes. By comparing all the results, we can examine whether the parameters converge to consistent values across different starting points and for different iteration numbers.

The results from simulation iii) and iv) in Figure 34 and Figure 35 show similar tendencies as i) and ii), respectively, with regards to Plaxis output and sensor data compliance. These relations are described in the prior section and are therefore not analyzed in detail in this section.

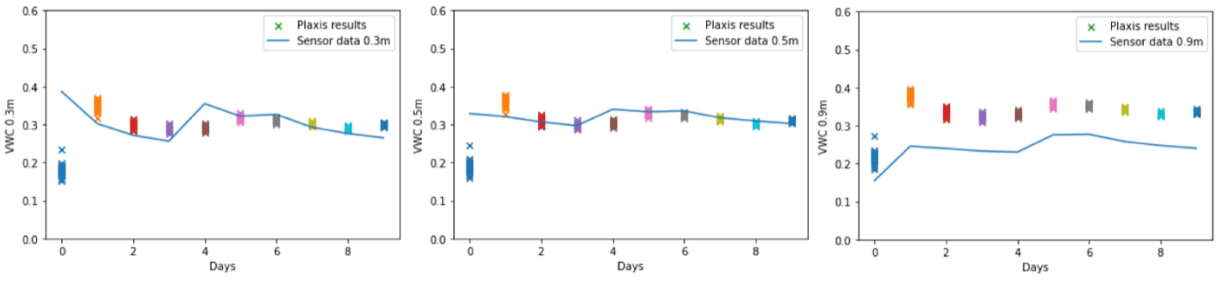


Figure 34: Plaxis results vs sensor data for simulation iii)

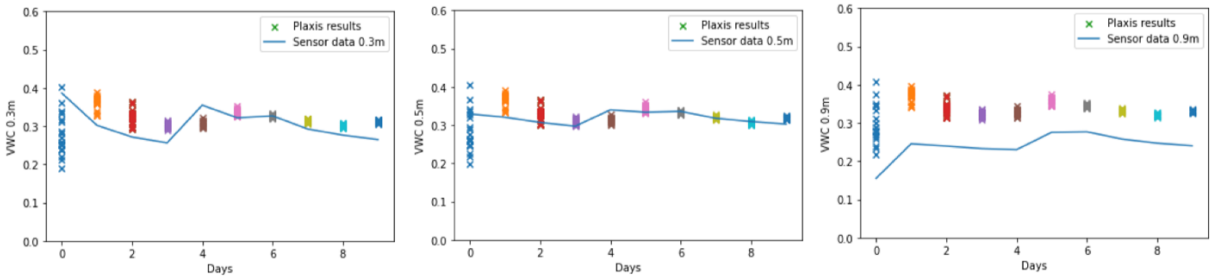


Figure 35: Plaxis results vs sensor data for simulation iv)

Parameter estimations for simulation iii) and iv) are shown in Figure 36. Like results i) and ii), we see variable amounts of convergence in the estimated parameters. In simulation iii), both the saturated permeability k_{sat} and n-parameter converge to significantly more reliable estimates which is seen by the standard deviation for k_{sat} changing from $\sigma_{k_{sat},prior}$ = 2.0 to $\sigma_{k_{sat},posterior}$ = 0.368, and from $\sigma_{n,prior}$ = 0.5 to $\sigma_{n,posterior}$ = 0.156 for n. The α-parameter doesn't have such a strong tendency seen from the plot, however the data shows that the standard deviation for α changes from $\sigma_{\alpha,prior}$ = 1.0 to $\sigma_{\alpha,posterior}$ = 0.358 which is a noteworthy change.

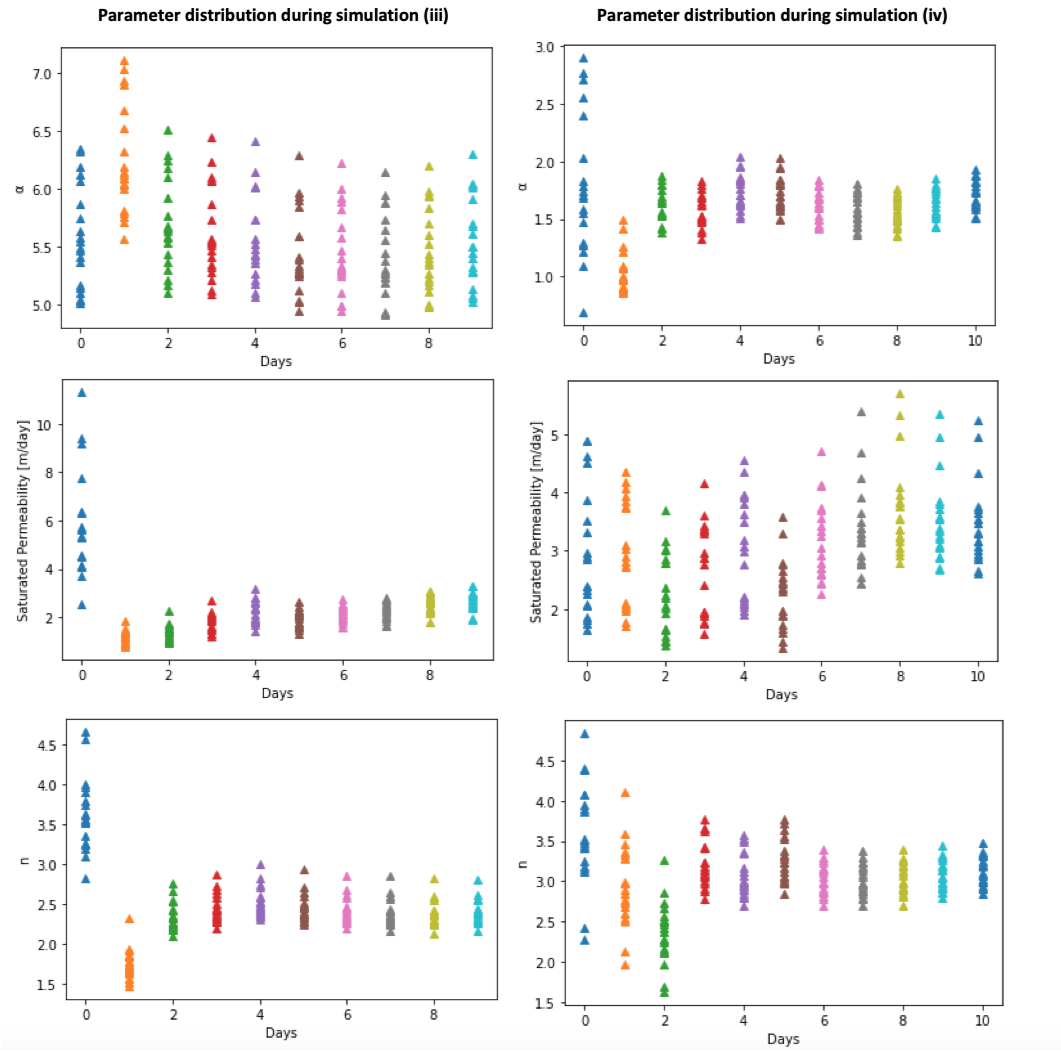


Figure 36:) Results from parameter estimation (iii) and iv)). Prior estimates (iii): $\mu_\alpha = 6.0$, $\sigma_\alpha = 1.0$, $\mu_{k_{sat}} = 5.5$, $\sigma_{k_{sat}} = 2.0$, $\mu_n = 3.5$, $\sigma_n = 0.5$. Prior estimates (iv): $\mu_\alpha = 2.0$, $\sigma_\alpha = 1.0$, $\mu_{k_{sat}} = 4.0$, $\sigma_{k_{sat}} = 2.0$, $\mu_n = 4.0$, $\sigma_n = 0.5$

Parameter estimations for simulation iv) show a great improvement in standard deviation for α, with $\sigma_{\alpha,prior}$ = 1.0 and $\sigma_{\alpha,posterior}$ = 0.121, and for n with $\sigma_{n,prior}$ = 0.5 and $\sigma_{n,posterior}$ = 0.185. In contrast to iii), the saturated permeability shows less convergence than iii) with $\sigma_{k_{sat},prior}$ = 2.0 and $\sigma_{k_{sat},posterior}$ = 0.771.

When analyzing the results in Table 2, it is evident that there are similarities in the results from simulation ii) and iii) as both the k_{sat}-parameter and n-parameter converge to almost identical values. While these parameters demonstrate a high degree of similarity, it is important to note that the α-parameter exhibit a noticeable deviation. The α-parameter states the air-entry value for the soil, marking for what soil suction the permeability (and water content) starts to decrease in the SWCC.

This is further substantiated by analyzing the Plaxis output vs sensor data-plots for both cases (Figure 31 and Figure 34). In simulation ii), the permeability seems to be lower than for simulation iii) as the Plaxis outputs show a faster response to VWC in iii), even though k_{sat} converges to almost identical values. This is likely caused by the lower value for the α-parameter for simulation ii), meaning that the unsaturated permeability starts to decrease for lower soil suctions and vice versa for simulation iii).

Despite very similar values for k_{sat} and n in ii) and iii), the results are not sufficiently consistent to confidently assert that the parameter estimates are correct.

Another example of this issue is the similarities observed between simulation ii) and iv). These results demonstrate very similar Plaxis outputs, even though the van Genuchten parameters exhibit significantly different values. These findings underscore the importance and difficulty in analyzing hydrological parameters, especially for SWCC's. As mentioned in the limitations section, different parameter combinations may yield similar results, which makes it challenging to make high certainty conclusions.

5.4.1 Effect of iterations

As seen from Figure 36 and the results for simulation iv), which was done using Nm = 12 days, the degree of convergence does not seem to be affected by the increase in number of iterations. Based on the observed trend in the results, it appears that when the parameters exhibit significant convergence, this tends to happen rapidly within the first few iterations (days).

However, it is important to note that this pattern may vary depending on the data set used for parameter calibration. If the initial data set consists of limited instances of significant fluctuations or extreme values, introducing more iterations could potentially yield positive effects by providing additional information about the measurement state.

5.4.2 Effect of prior knowledge

The effect of prior knowledge on parameter estimation is a significant factor to consider. Based on the results in this thesis, it appears that the α- and n-parameter respond more to changes in the initial prior knowledge than k_{sat}, although this tendency is not very clear and could very likely originate from the small changes done in the initial conditions for k_{sat}.

However, by narrowing down the area of initial uncertainty, it becomes possible to enhance the reliability of the parameter estimations as it will lead the algorithm closer to the correct results.

5.5 Overall discussion

The results of the parameter calibration are positive, showing that the Plaxis outputs align well with the measured data. The updated parameters are also getting closer to a more accurate estimation within the expected range. However, it is important to emphasize the fact that it is challenging to draw definite conclusions from these simulations because of the nature of unsaturated soil theory itself. This difficulty doesn't necessarily mean there's anything wrong with the model or algorithm used. The problem lies in the fact that unsaturated soil problems are highly non-linear and lacks unique solutions for limited information. This limitation is clearly demonstrated in this thesis, where similar Plaxis outputs are obtained despite variations in the estimated van Genuchten parameters, and similar van Genuchten parameters result in significant varieties in for example soil suction.

To obtain more reliable conclusions, we need additional information about the problem. This could involve using measurements of soil suction along with VWC-data. Another approach could be calibrating the parameters based on hydrological data from actual landslide-events.

In the context of early warning systems, it is crucial to have more certainty about the hydrological conditions of the soil as the unsaturated conditions can lead to shallow shear circles as the waterfront changes. These critical shear circles may not be considered if the unsaturated conditions are not analyzed. In addition, changes in the α-parameter affect the soil suction which again influences the effective stresses and slope stability. More accurate information about the hydrological state of the soil is essential to create a reliable early-warning system.

Despite these challenges, using the EnKF-method for parameter calibration could be effective. An alternative method would be manually updating the parameters and evaluating the results manually, which would be time consuming and difficult. The automated nature of the EnKF approach is advantageous in this regard and the method has great potential when it comes to automatic calibration for hydrological parameters in soil.

6 Summary and main conclusions

This master's thesis focused on the parameter calibration of the van Genuchten parameters (α, n, k_{sat}) using the Ensemble Kalman Filter algorithm and sensor data for volumetric water content. The work involved setting up an automated water flow model using Plaxis-scripting and incorporating the EnKF algorithm into the Python code. The main objectives of this thesis were:

- Analyze and process sensors data in a wetting period for a slope in Meråker.
- Implement and automate a hydraulic model of the slope in Plaxis with Python scripting.
- Study and Implement the Ensemble Kalman Filter algorithm in the programming language Python Couple Plaxis model with the Ensemble Kalman Filter algorithm in Python. Assess the performance of the Ensemble Kalman Filter algorithm and the calibrated hydrological model.

The simulations began with specifying a prior state for the unknown parameters based on prior knowledge and assigning probabilistic distributions to represent the likely range of parameter values. Random samples were generated from these distributions, and Plaxis calculations were performed to obtain outputs of effective degree of saturation at different depths corresponding to sensor placements.

A comparison was then made between the Plaxis output and sensor data using the EnKF algorithm, with normally distributed measurement errors added to the sensor data. The samples of the unknown parameters were updated based on the covariance and disparity between the sensor data and Plaxis output. The updated samples were used as input for new Plaxis calculations, and this process was iterated multiple times.

The simulations were conducted with specific iteration times and prior knowledge of the unknown parameters, and the results indicated a reasonable correspondence between sensor data and Plaxis output. The correspondence was best at depth 0.5m with Plaxis outputs being almost identical to sensor values for some simulations. Bigger variations between sensor data and Plaxis output was observed at depth 0.3m and especially at depth 0.9m, assumably due to factors like vertical infiltration and lower permeable layers which were not considered in the Plaxis model.

The estimation of van Genuchten parameters showed convergence in most simulations, indicating higher certainty estimates of the parameters. However, the inherent nonlinearity of unsaturated soil problems renders them ill-posed. Multiple parameter combinations can produce almost identical results, as evidenced by similar Plaxis responses obtained with different estimations of the van Genuchten parameters. This lack of uniqueness makes it impossible to determine the correct results without additional information.

The findings suggest that the EnKF algorithm can determine hydrological parameters, but improved accuracy in initial conditions and prior parameter knowledge is necessary for more certain estimation. The potential application of this method in early warning systems for rainfall-induced landslides is present, but further refinement and certainty in the result is needed for sufficiently reliable results. Incorporating calibration with suction sensors as well as VWC-sensors or calibrate the parameters with data from landslide-events could be advantageous to achieve higher certainty.

7 Future work

Recommendations for future work with parameter calibration using EnKF and Plaxis scripting include:

- Expand the simulations and consider potential soil layers.
- Calibrate the parameters using suction-sensors as well as VWC-sensors. This would improve the reliability of the results drastically.
- Test the calibration on a variety of different data sets or over a significantly longer time period to compare the results.
- Implement the results in stability analysis to examine how the hydrological condition of the soil influences the stability. This would be crucial before implementing the calibration results in early-warning systems.
- EnKF can be used to estimate stiffness parameters of the soil by using monitoring data of deformations, as done by (Muhammad Mohsan, 2021).

8 Bibliography

Abdelbaki, A. (2020). *Assessing the best performing pedotransfer functions for predicting the soil-water characteristic curve according to soil texture classes and matric potentials.* European Journal of Soil Sience.

Andrey Guber, Y. P. (2010). *Multimodeling with Pedotransfer Functions. Documentation and User Manual for PTF Calculator (CalcPTF).* USDA.

Bishop, A. (1959). *The Principle of Effective Stress.* Teknisk Ukeblad.

Burdine, N. (1953). *Relative permeability calculations from pore size distribution data.* Journalof Petroleum Technology.

Carsel, R. F. (1988). *Developing Joint Probability DIstributions of Soil Water Retention Characteristics.* Athens: U.S. Environmental Protection Aoenc.

Childs, E. C.-G. (1950). *The permeability of porous materials.* London: Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences.

D. G. Fredlund, H. R. (2012). *Unsaturated Soil Mechanics in Engineering Practice* (1 ed.). Hoboken, New Jersey: John Wiley & Sons.

E.C. Leong, H. R. (1997). *Review of Soil Water Characteristic Curve Equations.* Journal of Geotechnical and Geoenvironmental Engineering.

Emir Ahmed Oguz, I. D. (2021). *A Case Study on IoT-based Hydrological Monitoring of Water-Induced Landslides in Central Norwar.* Trondheim: NTNU.

Emir Ahmet Oguz, I. D. (2021). *A Case Study on IoT-based Hydrological Monitoring of Water-Induced Landslides in Central Norway.* Trondheim: NTNU.

Evensen, G. (1994). *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods for forecast error statistics.* Bergen: Journal of Geophysical Research.

Evensen, G. (1994). *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics.* Bergen: Nansen Environmental and Remote Sensing Cente.

Genuchten, M. T. (1980). *A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils.* Soil Science Society of America Journal.

ISO. (2017). *ISO. (2017). ISO 14688-2:2017 Geotechnical investigation and testing — Identification and classification of soil — Part 2: Principles for a classification. In.* ISO.

Ivan Depina, E. O. (2021). *Dissemination of sensor data and landslide predictions through an early warning system.* Trondheim: Sintef.

Ivan Depina, E. O. (2021). *KlimaDigital Case Study.* Trondheim: Sintef.

K. Liu, P. V. (2018). *Sequential reduction of slope stability uncertainty based on temporal hydraulic measurements via the ensemble Kalman filter.* Elsevier.

Khan, F. (2022). *Bentley Communities*. Retrieved April 3, 2023, from https://communities.bentley.com/products/geotech-analysis/w/wiki/63021/hydraulic-

conductivity-input-in-plaxis-using-van-genuchten-functional-forms-in-a-groundwater-flow-analysis

Leiva, C. G. (2019). *Susceptibility assessment of rainfall Induced Landslides. Case study: Hegra-Meråker.* Trondheim: SINTEF.

Makonto, O. T. (2013). *Vadose zone classification and aquifer vulnerability of the.* Pretoria: University of Pretoria.

Mualem, Y. (1976). *A New Model for Predicting the Hydraulic Conductivity of Unsaturated Porous Media.* Haifa: Water Resources Research .

Muhammad Mohsan, P. J. (2021). *On the use of different constitutive models in data assimilation for.* Delft: Elsevier.

NGU. (n.d.). *Løsmassekart*. Retrieved March 12, 2023, from https://geo.ngu.no/kart/losmasse_mobil/

NVE. (2023). *Skredregistrering*. Retrieved May 1, 2023, from https://www.skredregistrering.no/#Forsiden

Pedersen, S. P. (2022). *Flow of water in unsaturated soil in Plaxis.* Trondheim : NTNU.

R. H. Brooks, A. T. (1964). *Hydraulic properties of porous media.* Fort Collins, Colorado: Colorado State University.

R. J. Millington, J. P. (1960). *Permeability of porous solids.* Adelaide: Departments of Agronomy and Agricultural Chemistry, Waite Institute, University of Adelaide.

Robinson, K. (2019). *Experimental and Numerical Investigations of Infiltration into Unsaturated Soil.* Trondheim: NTNU.

Rommetveit, A. (2008). *Yr*. Retrieved Mars 5, 2023, from https://www.yr.no/artikkel/sa-mye-regn-gir-skredfare-1.6273561

Seboong Oh, N. L. (2014). *Uniqueness of the Suction Stress Characteristic Curve under Different Confining Stress Conditions.* Vadoze Zone Journal .

T. Mavara, R. B. (2018). *Hysteresis in Soil Water Characteristic Curve of Unsaturated Soil and its Influence on Slope Stability.* Toronto, Ontario, Canada: Department of Civil Engineering – York University.

T.Mengistu, Z. (2021). *SoilFlow : The Soil Water and Energy Balance Model.* Oslo: NVE.

Terazaghi, K. v. (1925). *Principles of soil mechanics. IV. Settlement and consolidation of clay.* Engineering News-Record.

Torun Rise, B. M. (2023, April 2). *KlimaDigital*. Retrieved from SINTEF: https://www.sintef.no/projectweb/klimadigital/

United Nations (UN). (2021). *Climate Change 2021: The Physical Science Basis.* United Nations.

W. Scott Sillers, D. G. (2001). *Mathematical attributes of some soil—water characteristic curve models.* Springer, Dordrecht.

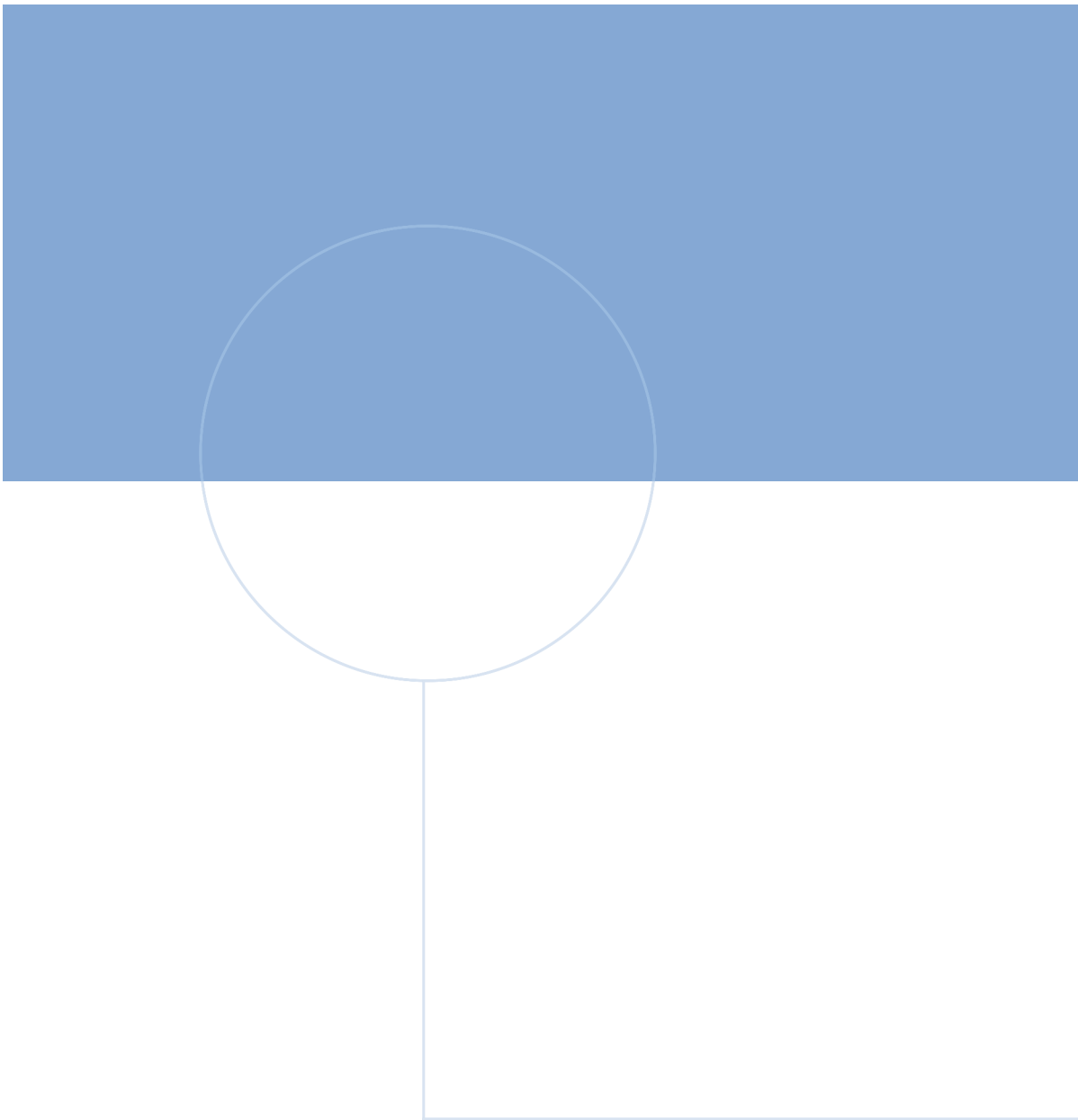
Yanwei Fan, J. G. (2018). *Application of Philip infiltration model to film hole.*

9 Appendix A
Python Code

```python
#Key elements of the code

#Importing packages
from plxscripting.easy import*
import numpy as np
import requests
import matplotlib.pyplot as plt
from datetime import datetime
import math


#Begin making the Plaxis-model

def plaxis_model(permeabilitet, S_res, S_sat, ga, gn, time):

    degrees = 35.0                      #Steepness of slope
    model_length = 20.0                 #Length of slope
    depth_to_bedrock = 1.5              #Depth to bedrock


    s_i.new()                           #Creating a server for Plaxis-scripting
    g_i.gotostructures()                #Go to structures in Plaxis



    #Defining a function for making the material

    def lag_materiale(permeabilitet, S_res, S_sat, ga, gn):


        #theta_r     Residual water content
        #theta_s     Saturated water content
        #alpha       Inverse of the air entry value
        #n           Pore size distribution index

        #Making the material
        sand = g_i.soilmat('Identification', 'Sand', 'SoilModel', 'Mohr-Coulomb', 'ERef',
                           20000, 'cRef', 5, 'phi', 30, 'cInc', 1,
                    'GroundwaterClassificationType', 'User-defined',
                    'SWCCFittingMethod', 'Van Genuchten', 'SaturationSaturated', S_sat,
                    'ClayFraction', 0, 'SiltFraction', 10, 'SaturationResidual', S_res,
                    'PermHorizontalPrimary', perm,
                    'PermVertical', perm, 'GenuchtenGn', gn, 'GenuchtenGa', ga/10)

        return sand


    sand = lag_materiale(permeabilitet,  S_res, S_sat, ga, gn)

    #Defining start points and end points for the polygons.
    .
    .

    #Making the slope-polygons and setting the material to the slope.
    skraning.setmaterial(sand)
    start_element.setmaterial(sand)
    slutt_element.setmaterial(sand)



    #Making the mesh
    g_i.gotomesh()
    g_i.mesh(0.01)
    g_i.selectmeshpoints()


    #Defining the stress points at depth 0.3m, 0.5m and 0.9m


    g_o.addcurvepoint("node", (model_length/2, stress_point_03))
    g_o.addcurvepoint("node", (model_length/2, stress_point_05))
    g_o.addcurvepoint("node", (model_length/2, stress_point_09))
    g_o.update()


    #Defining flow conditions
    g_i.gotoflow()
    .
    .
    .
    #Defining the boundary conditions and infiltrarion in phase 0
   .
   .
   .
    #Defining infiltration in phase 1 and setting a dishacrge function
    .
    .
    .
    #Setting the discharge function
    g_i.set(g_i.DischargeFunction_1.Signal, "Table")
    g_i.set(g_i.DischargeFunction_1.Table, table)


    g_i.gotostages() #Staged construction
    .
```

```
    .
    # Set up the InitialPhase to flow only
    .
    .
    # Increase the tolerated error
    .
    .
    # Set up phase 1 to transient groundwater flow
    .
    .
    #Defining time interval for calculation
    g_i.set(g_i.Phase_1.TimeInterval, time)


    #Calculations
    def calculations():

        list_of_sat_results = []

        g_i.calculate(True)

        #Checking if analysis was successfull
        if(strInPh.split()[1] == 'OK' and strPh_1.split()[1] == 'OK'):

            print('Flow analysis successful.')


            #Collecting the Saturation results in a list
            for k in range(3):


                values_sat = g_o.getcurveresults(g_o.CurvePoints.Nodes[k], phase1_o, g_o.ResultTypes.Soil.EffSaturationAsRatio)
                list_of_sat_results.append(float(values_sat))


        return list_of_sat_results

# Returning the results
    return calculations()



#Sensor data is defined in lists:

    # VWC03 Volumetric water content depth 0.3
    # VWC05 Volumetric water content depth 0.5
    # VWC09 Volumetric water content depth 0.9


# Defining the number of iterations Nm

Nm = len(VWC_0_3) #Number of sensor values

# Measurement error
sigma_m = 0.01


# Defining number of realizations/ensemble members
Ne = 20


# Defining prior knowledge of unknown parameters and

#mu = expected value
#std = Standard deviation

#Making the samples log-normal
std_ln = np.sqrt(np.log(1 + std**2/mu**2))
mu_ln = np.log(mu - 0.5*std_ln**2)


#Making the Y-matrix
Y = np.zeros((6, Ne, Nm))

#Making connection to Plaxis
p = 's<V/+?84wSZG8TT4'
s_i, g_i = new_server('localhost', 10000, password = p)
s_o, g_o = new_server('localhost', 10001, password = p)



# ITERATIONS BEGIN

for i in range(Nm):

    #Making initial parameter samples based on prior knowledge (iteration 0)

    if(i==0):

        Y[0,:,i] = np.random.normal(loc = mu_ln_perm, scale = sig_ln_perm, size = Ne)

        Y[1,:,i] = np.random.normal(loc = mu_ln_ga, scale = sig_ln_ga, size = Ne)

        Y[2,:,i] = np.random.normal(loc = mu_ln_gn, scale = sig_ln_gn, size = Ne)
```

```
        # Save values in files
        file1 = open()



    # Updating the Y-matrix with updated parameters
    else:

        Y[0,:,i] = Z_a[0,:]

        Y[1,:,i] = Z_a[1,:]

        Y[2,:,i] = Z_a[2,:]

        .
        .
        .
        # Writing the values to textfiles and saving them
        .
        .

    #Running the Plaxis model Ne times, one for each parameter sample.
    for j in range(Ne):

        #Calling the Plaxis-model function
        list_of_results = plaxis_model(permeabilitet = perm_input , S_res = S_res, S_sat = S_sat,
                                       ga = np.exp(Y[1,j,i]), gn = np.exp(Y[2,j,i]), time = i + 1)


        # Adding the Plaxis-results to the Y-matrix
        if list_of_results:

            Y[3,j,i] = list_of_results[0]*(theta_s-theta_r) + theta_r
            Y[4,j,i] = list_of_results[1]*(theta_s-theta_r) + theta_r
            Y[5,j,i] = list_of_results[2]*(theta_s-theta_r) + theta_r

            # Close output
            g_o.close()
            # Start new output
            s_o, g_o = new_server('localhost', 10001, password = p)


#Ensemble Kalman Filter Algorithm


    # Defining the G matrix
    G = np.zeros((3,6))
    G[0,4] = 1
    G[1,5] = 1

    # Copy values
    Z_f = Y[:,:,i]


    # Calculating Mean values of sample parameters
    Z_f_mean = np.matmul(Z_f, np.ones((Ne,Ne))/Ne)


    # Difference between parameter samples and mean
    Z_d = Z_f - Z_f_mean

    # Covariance matrix
    Czz = 1/(Ne - 1)*np.matmul(Z_d, Z_d.T)


    # Adding noise to the measurements
    D =  np.matmul(np.reshape(np.array([VWC_0_3[i], VWC_0_5[i], VWC_0_9[i]]), (3,1)),
                   np.ones(shape = (1,Ne))) + np.random.normal(size = (3, Ne))*sigma_m

    # Difference between measurements and Plaxis predictions
    obs_dif = D - np.matmul(G, Z_f)

    #Defining the Cdd-matrix
    Cdd = np.eye(3)*sigma_m**2

    #Matrix multiplications
    cov_inv = np.linalg.inv(np.matmul(G,np.matmul(Czz,G.T)) + Cdd)

    # Calculating the change that will be added to the parameter samples

    chng = np.matmul(Czz, np.matmul(G.T,np.matmul(cov_inv, obs_dif)))


#Updating the prior parameters by adding the change

    Z_a = Z_f + chng
```



NTNU

Norwegian University of
Science and Technology