École Polytechnique fédérale de Lausanne

Master Thesis in Environmental Sciences and Engineering

Investigation of proper boundary conditions for fluid flow in a soil column using model and data

Author:

Mohamed Aymen Lamzouri

Supervisors:

Andrea Rinaldo
Mitra Asadollahi

February 4^{th}, 2022

Contents

List of Figures

List of Tables

Acknowledgments

First of all, I would like to express my gratitude to Professor Andrea Rinaldo for his supervision, advice, and educational content he provided during my studies at EPFL.

I also want to sincerely thank Mitra Asadollahi for her patience, expertise, advice and feedback that allowed me to understand the methods used in this project.

I thank the Laboratory of Ecodydrology (ECHO) and its members for hosting this project and providing me with the experimental data and software packages to be used during this study.

Last but not least, I would like to thank my family, especially my parents, for the unconditional support and guidance they provided me throughout my life.

Abstract

This study aims at investigating the best set of boundary conditions in a lysimeter setup with complex bottom boundary condition, coupled with the calibration of one dimensional homogeneous solute flow transport model parameters. Moreover, this study investigates how significant the impact of evaporation may be in this lysimeter setup, and whether considering air-entrapment in the model allows to increase the accuracy of the simulation results. To achieve these goals, in this work, a set of calibrations on a one dimensional homogeneous solute flow transport model against experimental data (tracer concentrations, pressure, water content and drainage flow) has been carried out under different configurations. Measurement data are collected during a controlled experiment where a single pulse of Bromide (Br) travels through a lysimeter irrigated on a daily basis, while the suction applied on a filter membrane at the bottom controls the drainage flow. The Bayesian inverse modeling MCMC DREAM algorithm is used to determine the soil hydraulic properties and solute transport parameters under multiple configurations for the upper and lower boundary conditions. Identifying the proper boundary condition for this lysimeter setup was particularly challenging because of the presence of a membrane creating a positive pressure overlaid by a suction indirectly applied in the bottom end of the lysimeter, at which direct pressure measurements are not available. The results from calibration, such as posterior parameter distributions and optimal parameters are compared between different configurations and with typical values for the soil under study. To compare the configurations used, the root mean squared error (RMSE) between simulated and measured data is used to quantify the modeling accuracy under a given set of boundary conditions. Additionally, the ability of the model to correctly simulate the trends of the measured observations was assessed to compare the configurations used. Obtained results reveals that combining flux and pressure boundary conditions allows to obtain precise posterior parameters distributions, and simulation with the lowest RMSE, while adequately reproducing measurement data dynamics. Evaporative fluxes were predicted by a cropped model with boundary conditions defined from empirical measurements of pressure and discharge. These predicted evaporative fluxes were used to quantify the impact of evaporation on the system. In this cropped model, simulated total evaporation was 0.90 cm, compared to 98.14 cm of total drainage flux and 92.27 cm of total irrigation flux, while lysimeter was holding on average 20.47 cm of water in its storage on a daily basis. As evaporative fluxes represented less than 5% of the total water mass balance on a daily basis (even with underestimated discharge) it is concluded that evaporation was insignificant and can be neglected. In fact, evaporation is expected to be negligible in such setup as high amount of evaporation occurs in more dried systems (Liu et al., 2002). Air entrapment assumption in the model reduced the error for the cropped configuration, but increased it in the full domain model. Hence, the impact of air-entrapment assumption in the model on the simulation results depend on the boundary conditions setup used. Overall, the findings of this work suggests that the conclusions on optimal boundary conditions for this setup needs to be verified under known boundary conditions or precisely determined model parameters, verifying that optimal boundary conditions obtained with this approach are similar to experimental measurements. Further, investigation of a complete hysteresis assumption for soil hydraulic properties is necessary, and can potentially lead to better inverse parameter calibration results, ideally with a higher number of experimental measurements for unsaturated conditions.

Keywords— *Water flow, Solute transport, HYDRUS-1D, Inverse modeling, Bayesian parameter estimation, Monte Carlo Markov Chain, Boundary conditions*

1 Introduction

Modeling water flow and solute transport in porous media is the key to understand subsurface hydrological and biogeochemical processes and their impact on the environment and human health in multiple fields. For instance, it allows to simulate contaminant transport for groundwater pollution monitoring (e.g., Mirbagheri, 2004). Also, it is of high importance in agriculture where it allows to quantitatively estimate the crop growth, nutrients concentration and moisture content in soil in order to prevent from water shortage or soil salinization (e.g., Chen et al., 2019). To numerically implement one dimensional flow and solute transport model in variably saturated porous media, the Hydrus-1D software package (J. Šimunek et al., 2012) is used to solve the Richards and convection-dispersion (CDE) equations. Nevertheless, solving these governing equations require accurate knowledge of the boundary conditions occurring during the experiment. Soil hydraulic properties (SHPs) and solute transport parameters of the soil are also used in the model, but usually not known a priori due to the high sensitivity and variability of these parameters at different spatial scales (Mertens et al., 2003, Corwin et al., 2006).

Controlled lysimeters experiments are extensively used to accurately measure flux at the domain boundaries, and obtain other measurements such as moisture content, solute concentration and pressure head at various depths. Conducting lysimeter experiments has many aims and practical applications, for example it allows to quantify the biochemical oxygen demand reduction capacity of a soil and the phosphorous and inorganic nitrogen concentrations in soil (Brookman et al., 2005), to measure evapotranspiration and crop coefficients for agriculture products such as rice and sunflower (Tyagi et al., 2000), or study the transport of fuel in it gaseous form and it bio-degradation by microbial activity (Pasteris et al., 2002). Different setups of lysimeters are used to study water flow and solute transport (Pütz et al., 2018), a free draining weighable lysimeter with a surface area of 1 m^2 and length of 2 m situated in an agricultural field with crop rotation was used to collect data in order to develop a lumped parameter approach allowing to estimate the $\delta^{18}O$ transport in soils using a lumped dispersion model (Stumpp et al., 2009). Lysimeters with different soil textures installed in arable land, with suction cups installed at the bottom boundary to regulate the matric potential of the soil similarly to surrounding field, were used to study the evaporation rate using weather data and the FAO Penman-Monteith method combined with water flow modeling based on the Richards equation (Schneider et al., 2021). Another experimental setup, where the bottom of the lysimeter column consisted of a 50 cm gravel filter to make drainage easier and reproduce realistic pressure profile in the soil column, was used to conduct a tracer experiment for studying the closure of the soil water balance (Benettin et al., 2021). To evaluate the impact of redox potential variations caused by water level fluctuations on the greenhouse gases emissions from an agricultural soil, a lysimeter laboratory experiment with a porous nylon membrane plate installed at the bottom boundary to control the water table has been conducted (Wang et al., 2018). The lysimeter setup used in this study has an upper boundary exposed to atmospheric conditions, while the bottom boundary consists of a membrane allowing to build up a water table, overlaid by a suction force indirectly applied via a tube linked to a water bottle. The one dimensional flow and solute transport model parameters are not known a priori, evaporative fluxes has not been determined and measurements of the pressure at the bottom boundary are not available, hence boundary conditions are also not known a priori and this particular setup makes them difficult to estimate.

Lysimeter experiments and measured data can be used for inverse modeling methods to study water flow and solute transport (e.g., Groh et al., 2018; Abdou and Flury, 2004; Ritter et al., 2003; Zhang et al., 2003), and further be scaled to larger spatial resolution applications (Vrugt et al., 2008). Inversion modeling methods provide estimations of soil hydraulic and solute transport parameters by minimizing the residual error between the model outputs and measurements recorded during the experiment (e.g., Vereecken et al., 2007; Schelle et al., 2013). These methods has been widely studied during the last decades for water, heat, and solute transport model parameters calibration. An example of inverse modeling methods is the nonlinear regression technique, which consists of minimizing an objective function based on the least squares approach to fit modeled data with measured data. The implementation of this method revealed it limitations due to correlation among model parameters, non-uniqueness of best performance set of calibrated parameters, and uncertainty in the solution prediction due to the lack of measurement data to increase constraints on parameters (Friedel, 2005). Another algorithm used for global optimization is the shuffled–complex evolution algorithm (SCE-UA, Duan et al., 1992), and was used to estimate flow and transport parameters with supplementary optimization targets using independent knowledge of the stable isotope composition in the soil pore water depth profile (Sprenger et al., 2015).

Recently, statistical simulation algorithms based on Bayesian inference significantly gained in popularity as global optimization schemes for estimating flow and solute transport parameters by minimizing the residual error between measured and modeled data based on an objective function. Bayesian inversion is a statistical method aiming to estimate an unknown set of parameter by posterior sampling from prior information about the probability density function of these parameters. In particular, the Monte Carlo Markov Chain (MCMC) approach is increasingly used for identifying the unknown hydraulic and solute transport parameters. State-of-the-art tools for Bayesian

parameters inference includes the MCMC posterior sampling method based on the Differential Evolution Adaptive Algorithm (*DREAM*) algorithms (Vrugt, 2016), and has led to meaningful estimations of water flow and solute transport model parameters in multiple studies (e.g., Vereecken et al., 2007; Schelle et al., 2013; Moreira et al., 2016; Younes et al., 2017; Younes et al., 2018; Younes et al., 2019). For example, Younes et al., 2019 performed simultaneous estimation of water flow and solute transport by using a one-step Bayesian inversion from measured pressures and outflow concentrations measurements collected during a laboratory lysimeter experiment in a 40 cm length soil column. Additionally, the *DREAM* algorithm is able to maintain ergodicity and detailed balance, and provide uncertainty assessments associated with the obtained parameter estimates, leading to better estimation of the parameters posterior distributions (Rahnamay Naeini et al., 2019). This method has been applied in various contexts, including to model flow and pesticide transport through unsaturated porous media and provided accurate mean estimated parameters values (Younes et al., 2018).

Because accurate knowledge of boundary conditions is considered as a prerequisite for inverse parameter estimation from numerical modeling (Vrugt et al., 2008), little research has been carried out when it comes to simultaneous investigation of proper boundary conditions and inverse parameter calibration to fit with measured data collected from suction controlled lysimeter experiments. Among the challenges encountered when defining boundary conditions for one dimensional water flow and solute transport modeling, the spatial and temporal variability of these boundary conditions makes them difficult to define precisely (Zaradny, 1978). An inverse estimation of the boundary conditions using soil moisture content observations under known hydraulic parameters has been used to estimate upper flux boundary conditions (Li et al., 2021). Also, Schelle et al., 2013 showed that the lysimeter system behavior can be properly predicted under different boundary conditions setups with inverse parameter estimation. In this study, the calibration of one dimensional flow and solute transport model parameters will be performed under various boundary conditions configurations. For this purpose, a suction controlled lysimeter experiment was conducted to obtain precise measurements for moisture content, pressure head, discharge and solute concentration at various depths within the soil column. The Bayesian inversion modeling method based on the Monte Carlo Markov Chain *DREAM* algorithm will be used for soil hydraulic properties and solute transport parameters estimation under multiple configurations for the upper and lower boundary conditions. Calibration results and obtained parameters will be compared between configurations and with reference values from literature. Simulation results will also be compared between configurations to assess the accuracy reached when using a given boundary conditions setup. Calibration of a cropped model under known boundary conditions will be used to simulate the evaporation fluxes and to provide an understanding of the pressure head at different observation points. Finally, the best boundary conditions configuration found for this lysimeter experiment will be tested in a model with air-entrapment assumption to verify if it also provide accurate simulation results and posterior parameters distributions. The benefits of taking into account air-entrapment in the model will be assessed by comparing the simulation results and residual error obtained using the basic model and the model with air-entrapment assumption, under similar best performance boundary conditions configurations.

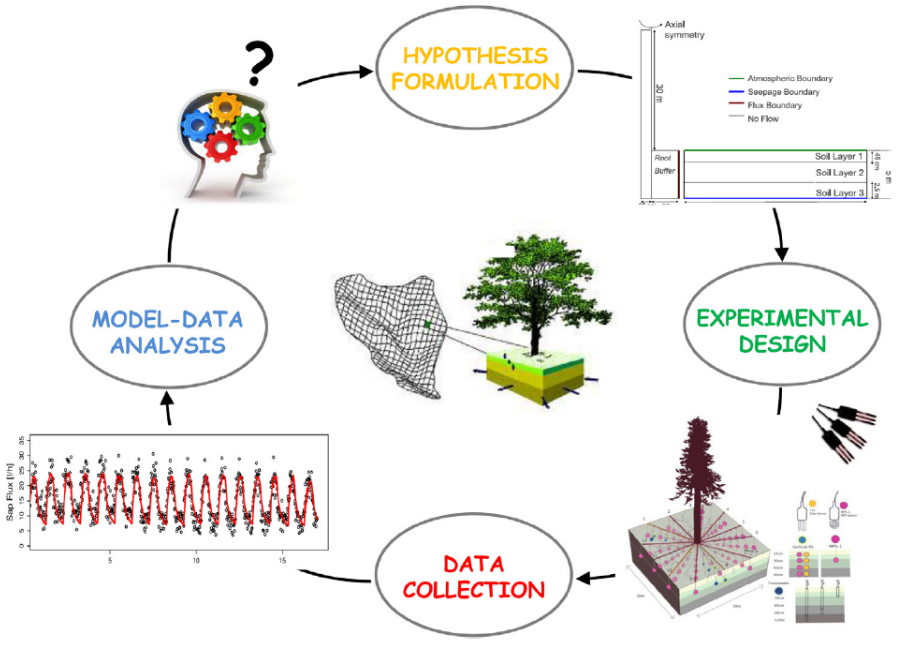


Figure 1: Water-flow and solute transport inverse modeling iterative process (Vrugt, 2016).

2 Methods

2.1 Data description

The experiment was conducted by members of the laboratory during a time period of 80 days, between the 25^{th} May and the 12^{th} August 2021. This is a variably saturated column experiment, in which a shower is used to irrigate the soil with water. The fluid can then exit the lysimeter by an outlet located at the bottom boundary where suction is applied.

The 10^{th} of June (day 17), a conservative tracer, ***Bromide (Br)***, mixed with irrigation water is injected at the upper boundary of the soil column. Moisture probes are installed to measure the moisture content at several observation points (at depths 8, 20, 30 and 42 cm) in the soil domain, a tensiometer is in place to measure the water pressure head in the soil column (at depth 20 cm). In addition, sampling is realized daily from the soil column to measure the concentration of bromide at soil (at 25 cm depth) and at the lower boundary. A filter is installed at the bottom of the soil column which is creating resistance against drainage flow, on top of that there is a suction applied at the bottom boundary. This suction is indirectly applied at the bottom boundary through a tube linked to a water bottle. This suction is not continuous and is applied only when the pressure in the drainage bottle reach a value higher than the aimed suction pressure. Finally, discharge is computed using load cells to calculate the weight variations of the system assuming negligible evaporation. The figure below provide a schematic representation of the experimental setup.

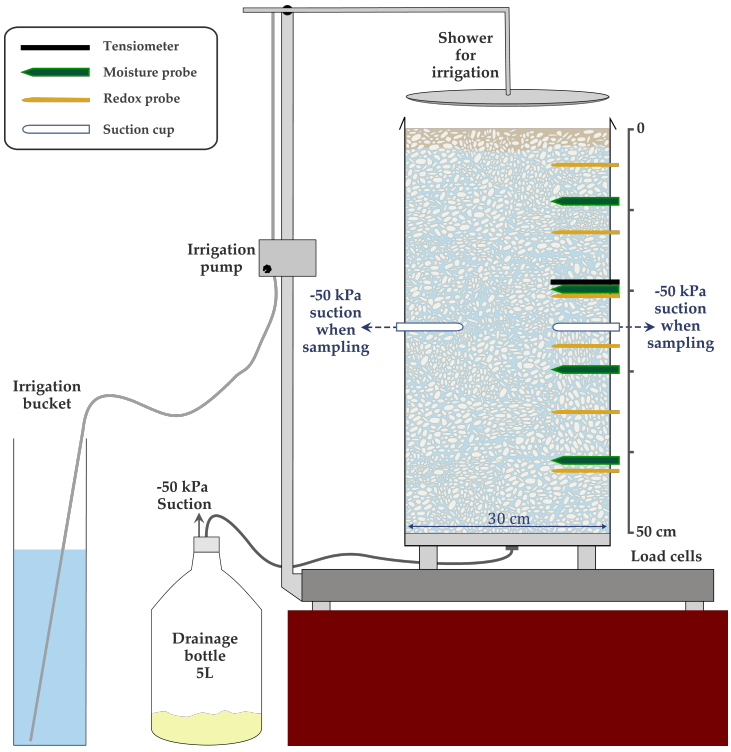


Figure 2: Suction controlled lysimeter setup.

Hence, the data collected and used during the calibration process to estimate longitudinal dispersion and soil hydraulic parameters, are the following: discharge (estimated from weight changes), water content (at depths 8, 20, 30 and 42 cm) and pressure (at depth 20 cm) as well as solute concentration at the middle of the soil column (at depth 25 cm) and bottom of the domain. Irrigation and input concentrations are recorded during the experiment and used as model input.

In order to be used in the solute transport model equations, the measured concentrations will be converted to $[mmol/cm^3]$, according to the following relation:

$$C\ [\frac{mmol}{cm^3}] = \frac{C\ [\frac{mg}{L}]}{M(Br) * 10^3\ [\frac{cm^3}{L}]} \tag{1}$$

With the molar mass of Bromide : $M(Br) = 79.9\ [\frac{mg}{mmol}]$.

2.2 Water flow and solute transport model in HYDRUS-1D

In the experimental setup used for water flow and solute transport, the soil column was irrigated on a daily basis with the volume that equates the drainage volume of the previous day to create a semi-steady state condition. Hence, the flow parameters which are the velocity and pressure depend on both the position and the time: such flow is called **transient**. The domain discretization used in the software is performed according to the **Galerkin finite elements method** application for water flow and solute transport modeling (Neuman, 1975, Perrochet and Bérod, 1993) subdividing the domain into 1 node for each 1 cm of soil.

2.2.1 Governing water flow equations

Richards Equation: The one-dimensional uniform vertical movement of liquid water in a variably saturated rigid porous medium is predicted by a modified form of the Richards equation in HYDRUS-1D (Simunek et al., 2008), as shown below:

$$\frac{\partial \theta(h)}{\partial t} = \frac{\partial}{\partial z}\left[K(h)(\frac{\partial h}{\partial z} + 1)\right] \tag{2}$$

With:

- h : Water pressure head $[cm]$.
- θ : Volumetric water content $[-]$.
- t : Time $[d]$.
- z : Spatial coordinate, here is depth $[cm]$.

Also, K is the hydraulic conductivity expressed by:

$$K(h, z) = K_s(z)K_r(h, z) \tag{3}$$

With:

- K_r: the relative hydraulic conductivity [-].
- K_s: the saturated hydraulic conductivity [cm/d].

Soil Hydraulic Model - Single Porosity Model (Mualem, 1976): The Mualem-Van Genuchten model provide relations for the volumetric water content as following:

$$\theta(h) = \begin{cases} \theta_r + \dfrac{\theta_s - \theta_r}{[1 + |\alpha h|^n]^m} & h < 0 \\ \\ \theta_s & h \geq 0 \end{cases} \tag{4}$$

With:

- θ_s: Saturated volumetric water content.
- θ_r: Residual volumetric water content.
- α, n and m are the Mualem Van Genuchten (MvG) hydraulic shape parameters.

This relation helps to provide an expression for the hydraulic conductivity as a function of the pressure head using the soil water retention curve $\theta(h)$, leading to the following equation:

$$K(h) = K_s S_e^l \left[1 - \left(1 - S_e^{\frac{1}{m}}\right)^m\right]^2 \tag{5}$$

With the effective saturation S_e [-], defined as:

$$S_e(h) = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} \tag{6}$$

And the following relation for parameters m, n:

$$m = 1 - \frac{1}{n}, \quad n > 1 \tag{7}$$

Hence, there is 5 independent parameters defined in the above equations: θ_r, θ_s, α, n and K_s. Concerning the pore-connectivity parameter l in the hydraulic conductivity function, it is generally assumed to be equal to 0.5 in average for many types of soils.

2.2.2 Hysteresis implied by air-entrapment model assumption

Hysteresis in soil represents the difference in the relationship linking the water content of the soil and the water potential computed under wetting and drying processes, as shown in the figure below. This relationship is known as the soil water retention curve. It states that the amount of water being held by the soil when it is wetted is smaller than the amount of water being held by the same soil when it is being dryed.

The wetting is the process leading to an increase of the water content of an unsaturated soil caused by water supply. In the other hand, drying is the decrease of the water content and level of saturation of a soil due to evaporation and drainage.

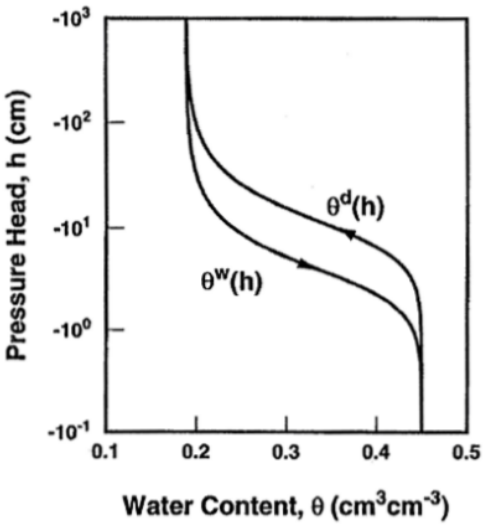


Figure 3: Water retention curve showing hysteresis, with the wetting curve $\theta^w(h)$, and the drying curve $\theta^d(h)$.

Among causes of hysteresis, air-entrapment in the soil pore space is an important factor. It happens when wetting an initially dryed soil, some air get entrapped, hence some space is occupied by the air out of the total pore space, then the volume of water during wetting will be less than during drying. This imply that $\theta_s^w \leq \theta_s^d$ for wetting and drying saturated water contents.

In this work, the impact of considering air-entrapment in the model will be investigated. In the model from Kool et al., 1987 to implement hysteresis, the parameter controlling air-entrapment is θ_s^w, which represents the water content at saturation during the wetting process, differentating it with the saturation water content during drying. The following relations holds for the saturated water contents during wetting and drying processes respectively:

$$\theta(h) = \begin{cases} \theta_r + \dfrac{\theta_s^w - \theta_r}{(1 + |\alpha h|^n]^m} & h < 0 \\ \theta_s^w & h \geq 0 \end{cases} \qquad \theta(h) = \begin{cases} \theta_r + \dfrac{\theta_s^d - \theta_r}{(1 + |\alpha h|^n]^m} & h < 0 \\ \theta_s^d & h \geq 0 \end{cases} \tag{8}$$

2.2.3 Governing solute transport equations

Considering the one-dimensional vertical solute transport in the liquid phase, this complex phenomenon is studied in Hydrus-1D (Simunek et al., 2008) using the Convection-Dispersion Equation (CDE), also called the Advection-Dispersion Equation (ADE), given by:

$$\frac{\partial(\theta C)}{\partial t} = \frac{\partial}{\partial z}\left(\theta D \frac{\partial C}{\partial z}\right) - \frac{\partial(qC)}{\partial z} \tag{9}$$

Where:

- C: Concentration of solute in the liquid phase $[mmol/cm^3]$.
- q: Volumetric flux density $[cm/day]$.
- t: Time $[days]$.
- θ: Soil volumetric water content $[-]$.
- D: Effective dispersion coefficient in the liquid phase $[cm^{-2}d^{-1}]$.

Also, the effective dispersion coefficient in liquid phase is described by the following expression developed by Bear, 1972:

$$\theta D = D_L|q| + \theta D_w \tau_w \tag{10}$$

$$D = \frac{D_L|q|}{\theta} + D_w \tau_w \tag{11}$$

With:

- D_L: Longitudinal dispersivity $[cm]$.
- $|q|$: Absolute value of the darcian fluid (water) flux density $[cm/day]$.
- θ: Soil volumetric water content [-].
- D_w: Molecular diffusion coefficient in liquid water $[cm^2d^{-1}]$.
- τ_w: Tortuosity factor in the liquid phase [-].

2.3 Boundary conditions

In order to solve the previously defined equations, it is needed to define boundary conditions for water flow and for solute transport. In HYDRUS-1D, two different types of boundary conditions are defined (Simunek et al., 2008):

- System-dependent Boundary Conditions: This type of boundary conditions are able to change during the simulation, which means that they depend on the solution found for the previous time step. Hence, boundary conditions might change depending on the saturation conditions, hydraulic properties and/or climate conditions.
- System-independent Boundary Conditions: This type of boundary conditions are known prior the simulation and are not subject to change during the simulation, meaning they are independent of the simulation results.

2.3.1 Water flow boundary conditions

The relevant water flow boundary conditions investigated in this study are described below (J. Šimunek et al., 2012):

- System-independent Boundary Conditions:
  - Variable Pressure Head (VP): This boundary condition allow to enter pressure heads as time series: for each time step a value for the pressure head is imposed at the upper/lower boundary.
  - Variable Flux (VF): This boundary condition allow to enter flux values as time series, for each time step a value for the flux is imposed at the upper/lower boundary.
- System-dependent Boundary Conditions:
  - Seepage Face (SF): This bottom boundary condition imposes a seepage face through which the water exits the saturated boundary of the flow domain. Once the bottom boundary becomes saturated, a pressure head equal to h_{seep} is imposed at the bottom. The value for h_{seep} is selected for trigerring flux across the seepage face, the default value from Hydrus for this parameter is $h_{seep} = 0\ cm_{H_2O}$. It can expressed by the following relations:

$$q(z,t) = 0 \qquad for\ h(z,t) < h_{seep} \qquad z = 50\ cm \tag{12}$$

$$h(z,t) = h_{seep} \qquad for\ q(z,t) < 0 \qquad z = 50\ cm \tag{13}$$

    Then, the flux across the boundary is computed by solving the Richards equation. This boundary condition is set to be used as a bottom boundary condition, and best suits laboratory soil columns experiments, where the water freely drains through a soil surface that is exposed to the exterior atmosphere.
  - Atmospheric Boundary Condition (ATM): Using this condition permit to consider climatic conditions for the upper boundary by specifying irrigation and evaporation fluxes.

For the atmospheric boundary condition, it is possible to specify if water build up on the surface (surface layer) up to a certain height h_{max} or if it runs off (surface runoff) when precipitation exceeds the saturated hydraulic conductivity of the soil surface. In case the specified flux into the soil is higher than the infiltration capacity of the soil, then a pressure head boundary is imposed depending if there is a maximum head specified at soil surface or if excess water exits as surface runoff. Thus, it is a system-dependent boundary condition since the actual flux depend on the soil moisture at each time step of the simulation.

2.3.2 Solute transport boundary conditions

To solve the solute transport equations, boundary conditions must be specified. The following HYDRUS-1D system-independent boundary conditions are used this study (Simunek et al., 2008):

- Concentration Flux Boundary Condition (Upper Boundary): Using this boundary condition imply to specify the concentration of solute in the water infiltrating the soil. The following relation shows the imposed conditions:

$$for\ 0 < t \leq T_{inj} \begin{cases} K\left(\frac{\partial h}{\partial z} - 1\right) = q_{inj} \\ \\ \theta D \frac{\partial C}{\partial z} + qC = q_{inj} C_{inj} \end{cases} \qquad for\ t > T_{inj} \begin{cases} K\left(\frac{\partial h}{\partial z} - 1\right) = 0 \\ \\ C_{inj} = 0 \end{cases} \tag{14}$$

    With q_{inj} the injected flux of water across the upper boundary in $[cm/day]$, and C_{inj} the injected solute concentration in the penetrating flux in $[mmol/cm^3]$ during a period of T_{inj}.
- Zero Concentration Gradient (Bottom Boundary): This boundary condition allow to specify a zero concentration gradient at the outlet of the domain. It can be expressed as follows:

$$\left(\frac{\partial C}{\partial z}\right)_{z=L} = 0 \tag{15}$$

    With the length of the soil column L equal to 50 cm.

2.4 Bayesian parameters calibration

In order to correctly simulate water flow and solute transport under different boundary conditions configurations, the soil hydraulic and solute transport parameters must be determined by calibrating the model. For this purpose, the *DREAM* algorithm (Vrugt, 2016) based on the Monte Carlo Markov Chain method is used. The algorithm explores various set of parameters starting from initial values comprised in the prior distribution of the model parameters. The ranges of the prior distribution are defined in a reasonable interval for the soil under study, and used to infer posterior parameters distribution using the MCMC sampler. Prior distribution type is a continuous uniform distribution because we have poor prior information about parameters values. The objective function to minimize is the log-normal likelihood function defined below. The residual error term in the objective function is normalized by the length of the observation vector, thus the residual error is computed such that no data type dominates the error vector due to a higher number of measurements.

$$\mathscr{L}(\mathbf{x}|\widetilde{\mathbf{y}},\widehat{\sigma}) = -\frac{n}{2}log(2\pi) - \sum_{t=1}^{n} log(\widehat{\sigma}_t) - \frac{1}{2}\sum_{t=1}^{n}\left(\frac{\widetilde{y}_t - y_t(\mathbf{x})}{\widehat{\sigma}_t}\right)^2 \tag{16}$$

With:

- $\mathbf{x}$: vector of unknown parameters.
- $\mathbf{y}$: vector of empirically measured data.
- $\widetilde{\mathbf{y}}$: vector of simulated data.
- $\widehat{\sigma}$: vector of standard deviation of the residual error between measured and simulated data.
- t: time in days.
- n: total number of days.

The standard deviation of the residual error is calibrated along with the water flow and solute transport parameters and takes a constant value from an observation type to another. About the model parameters, the residual and saturated water contents were fixed to the approximated values recorded during the experiment, which are $\theta_r = 0$ and $\theta_s = 0.4$.

An approximation for the parameters depending on the soil composition and bulk density is available in Hydrus. The values are computed based on a Neural Networks approach, which can serve as reference values. Also, available tables from literature providing values for multiple types of soils (Rawls et al., 1983, Carsel and Parrish, 1988) were used to set reference soil hydraulic parameters values for a **Loam** soil.

Using the average soil hydraulic parameters values for various soil types, the prior ranges were defined for the model parameters to be calibrated. A relative wide range for the parameters were chosen to allow searching in a large space. Reference values from literature tables and the hydrus approximations will be used for comparison after the calibration process. All these information are displayed in the table below:

Table 1: Prior ranges, literature references and Hydrus approximation for soil hydraulic parameters.

Parameters	Min	Max	Literature reference	Hydrus approximation
α $[cm^{-1}]$	0.001	0.150	0.063	0.014
n $[-]$	1.05	2.7	1.39	1.45
K_s $[cm/day]$	0.3	40.0	20.64	12.86
D_L $[cm]$	1.0	50.0	-	-

2.5 Configurations definition

Solute transport boundary conditions used were defined in the previous subsection. Additionally, the molecular diffusion is considered negligible as when computing dispersion coefficient in the convection-dispersion equation, the molecular diffusion coefficient is usually very small compared to longitudinal dispersivity in solute transport lysimeter experiments (Selker and Or, 2021, Filipović et al., 2013).

In order to determine the best set of boundary condition for this lysimeter setup having a complex bottom boundary condition, several configurations are defined with different boundary conditions. The assumption of negligible evaporation will be verified using the cropped configurations to simulate evaporative fluxes under known boundary conditions. Also, the impact of air-entrapment assumption in the model will be investigated under best performance configurations. The following table includes the empirically measured observations used for calibration and the parameters calibrated for each configuration:

Table 2: Boundary conditions configurations, with quantities used for calibration and calibrated parameters.

Setup	Observations					Calibrated parameters						
	Q	θ	p	C_Q	C_S	α	n	K_s	D_L	σ	h_{Seep}	θ_s^w
Atm-Seep	+	+	+	+	+	+	+	+	+	+	-	
Atm-SeepC	+	+	+	+	+	+	+	+	+	+	+	
Atm-Flux	+	+	+	+	+	+	+	+	+	+		
Flux-Flux	+	+	+	+	+	+	+	+	+	+		
Pressure-Flux (Cropped)	-	+	-	+	+	+	+	+	+	+		
Flux-Pressure	+	+	+	+	+	+	+	+	+	+		
Pressure-Flux (Cropped HYS)	-	+	-	+	+	+	+	+	+	+		+
Flux-Pressure (HYS)	+	+	+	+	+	+	+	+	+	+		+

- Setup Atm-Seep: atmospheric upper boundary conditions (ATM) coupled with a seepage face bottom boundary condition (SF) using the default seepage pressure parameter $h_{Seep} = 0\ cm$.
- Setup Atm-SeepC: same boundary conditions as in Atm-Seep, except that the seepage pressure parameter h_{seep} is calibrated along with flow-transport parameters.
- Setup Atm-Flux: atmospheric upper boundary conditions (ATM) with a variable flux (VF) bottom boundary condition defined from discharge values.
- Setup Flux-Flux: variable flux (VF) boundary condition at upper and lower boundary.
- Setup Pressure-Flux (Cropped): the defined domain has it upper boundary located at $z = 20\ cm$ depth and it lower boundary at $z = 50\ cm$ depth of the soil column. Variable Pressure (VP) is the top boundary condition, and Variable Flux (VF) at the bottom boundary, inputs are from empirically measured pressure and discharge respectively.
- Setup Flux-Pressure: modeled pressures from the cropped model are used to impose a bottom variable pressure boundary conditions, coupled with flux boundary condition at the upper boundary, assuming continuity conditions for fluid flow within the column.
- Setup Flux-Pressure (Cropped HYS): similar to setup Flux-Pressure (Cropped), but including air-entrapment assumption in the model.
- Setup Flux-Pressure (HYS): similar to setup Flux-Pressure, but investigating air-entrapment assumption in the model.

For the configurations where an upper atmospheric boundary condition is defined, the maximum height of water allowed to be build on surface is $h_{max} = 5\ cm$ according to the experimental layout.

Configurations using a flux boundary condition, without specifying either an upper of bottom pressure head boundary condition, did not converges for days with null discharge. The period was then reduced to 53 days, where discharge is always higher than zero.

3 Results

3.1 Calibrations

3.1.1 Distribution of posterior parameters samples

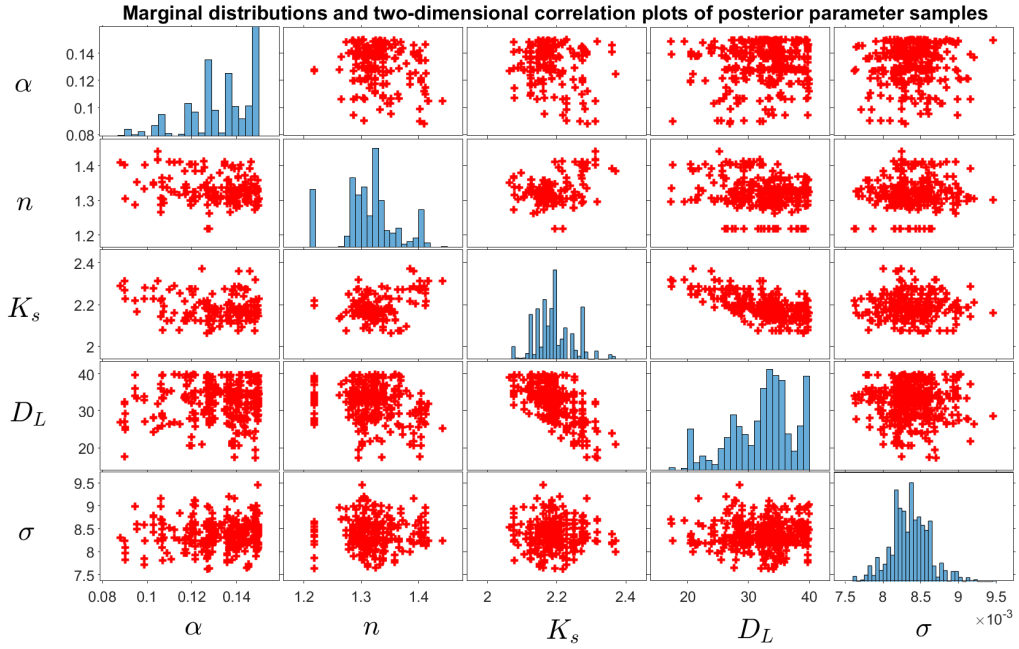


Figure 4: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Atm-Seep setup.

The above figure displays the inferred distributions of the calibrated flow-transport model parameters. The parameters n, K_s and σ are well estimated since the prior range is narrowed in their posterior distributions, and show a visible Gaussian distribution. On the other hand, the parameters α and D_L are not very well estimated since their prior range has not been narrowed in their posterior distributions, and converges towards the upper bound of the prior range. A high pairwise auto-correlation is observed between K_s and D_L, and between K_s and n, which might occur given these parameters are interrelated in the MvG model equation.

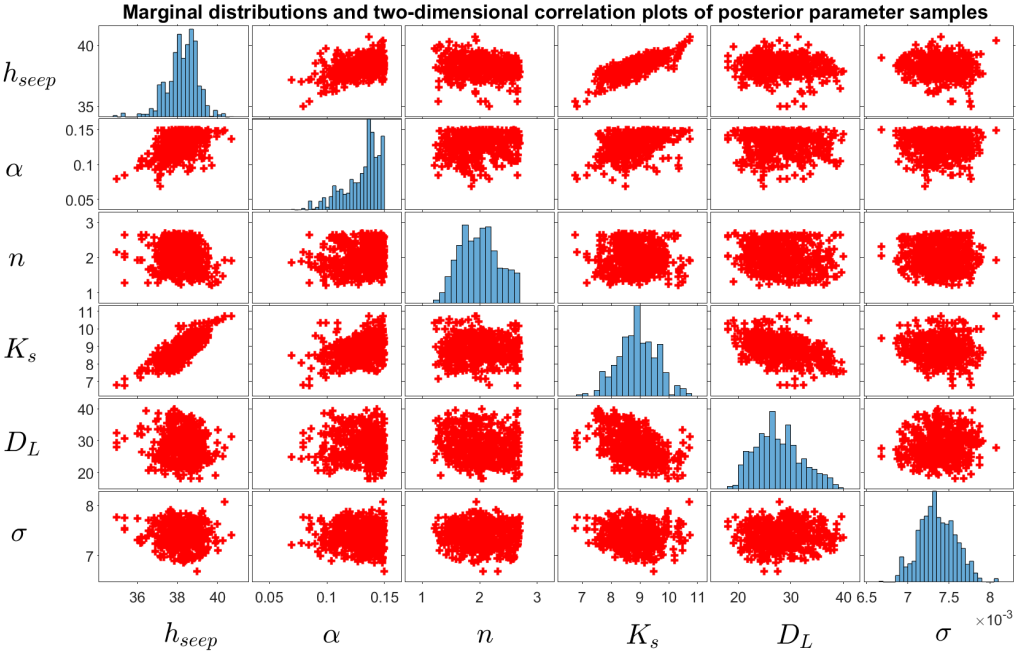


Figure 5: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Atm-SeepC setup.

As showed in calibration results above in Figure 5, the parameters h_{seep}, K_s, D_L and σ are well estimated: the prior range is narrowed in their posterior distributions and show a visible Gaussian distribution and obtained values from posterior distributions clearly converge towards a similar value. On the other hand, the parameters α and n are not very well estimated. Despite the fact that their posterior distribution shows more or less a centered distribution, the prior range of these parameters has not been narrowed in their posterior distribution and do not converge toward a similar value. In particular, α converges towards the upper bound of the prior range. High pairwise correlations between h_{seep} and α and between h_{seep} and K_s are observed, which indicate dependence on this two parameters estimation. Also a visible pairwise correlation between α is reported in the Figure, it can occur since these parameters are interrelated in the MvG model equation.

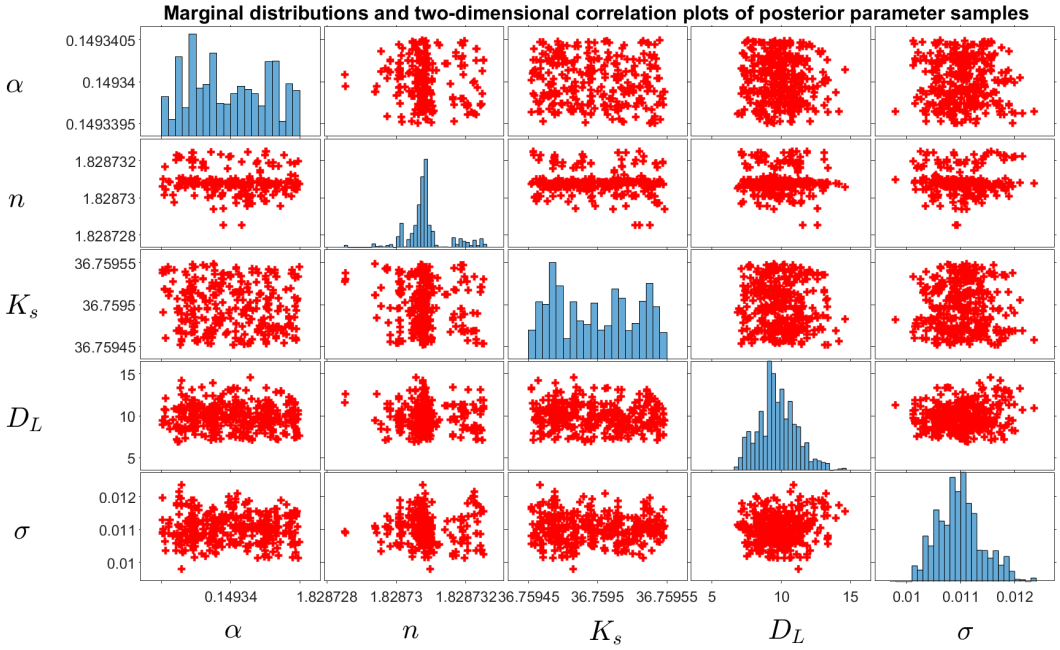


Figure 6: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Atm-Flux setup.

Given the results displayed in Figure 6, one can notice that the parameters n, D_L and σ are well estimated since the prior range is narrowed in their posterior distributions, and show a visible Gaussian distribution, with the posterior samples converging toward a precise value. On the other hand, the parameters α and K_s are not badly estimated since their prior range has also been narrowed in their posterior distributions, still their obtained posterior distributions do not display a Gaussian trend. Globally, no correlation between parameters is remarkable, and all of the calibrated parameters prior ranges has been efficiently narrowed in their posterior distribution. No particular pairwise correlation can be reported, indicating independence between parameters posterior distributions.

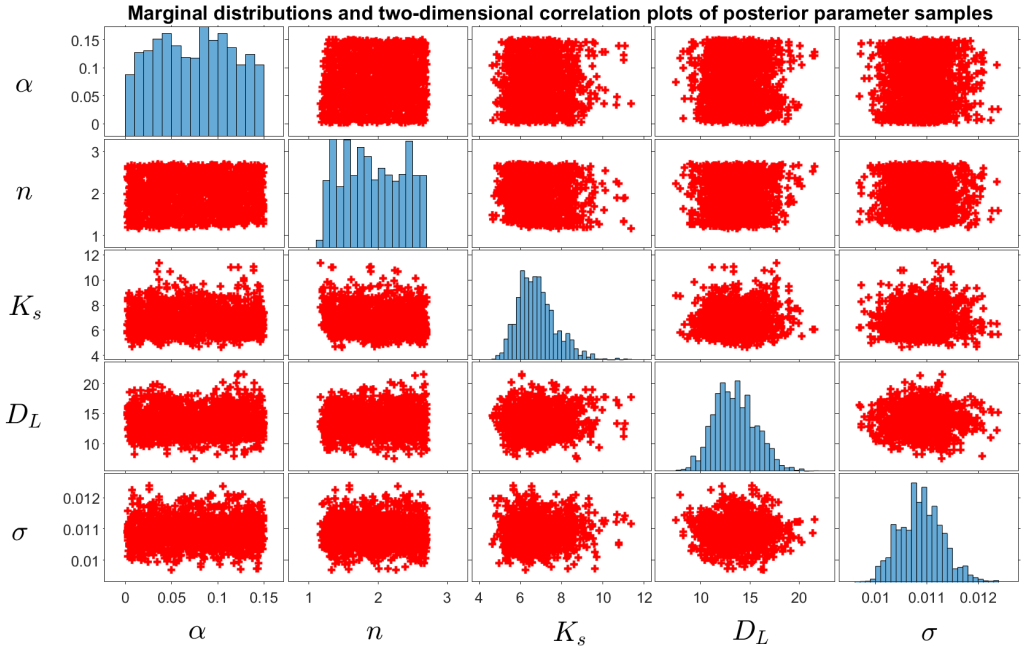


Figure 7: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Flux-Flux setup.

From posterior parameter samples above, one can observe that the prior distribution of the parameters K_s, D_L and σ are narrowed in their posterior distributions and have a Gaussian distribution. Concerning the parameters α and n, the posterior samples of these parameters are more or less uniformly distributed across their prior distribution, their prior ranges are not narrowed and their posterior distribution is not Gaussian. Globally, no remarkable correlation is observed between posterior samples of parameters.

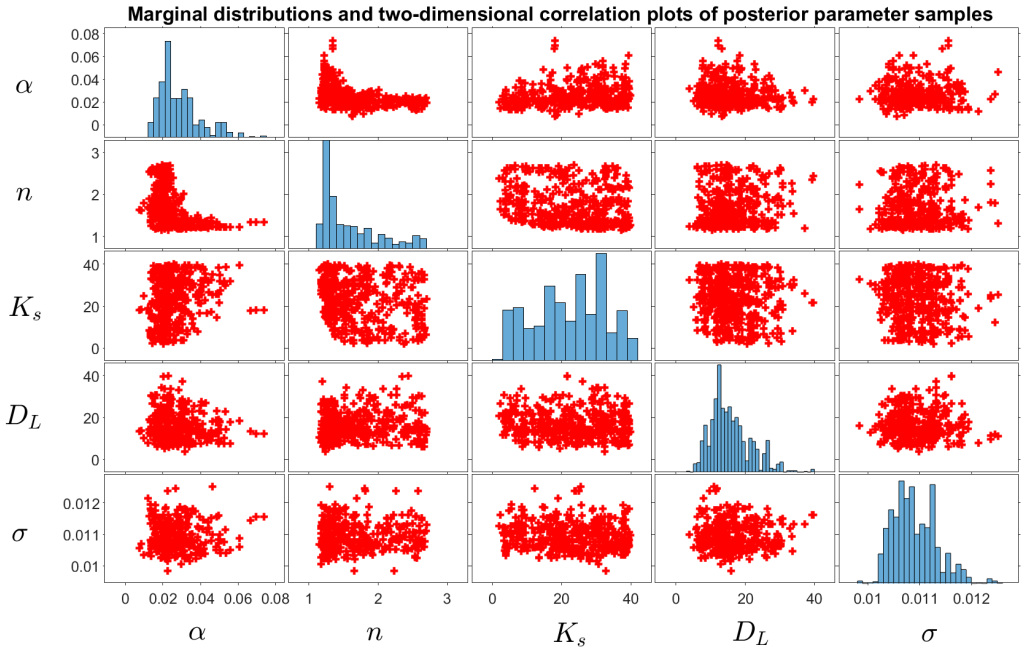


Figure 8: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Flux-Pressure (Cropped) setup.

In the figure above, one can observe that the prior range of the calibrated parameters were narrowed in their posterior distributions. Especially, the posterior distributions for α, D_L and σ display a visible Gaussian distribution shape. For n and K_s, their distribution converges toward a particular value within the prior interval, but the distribution do not seem to be a Gaussian distribution. Concerning the correlation between parameters, the posterior pairwise distributions do not display remarkable correlation between parameters.

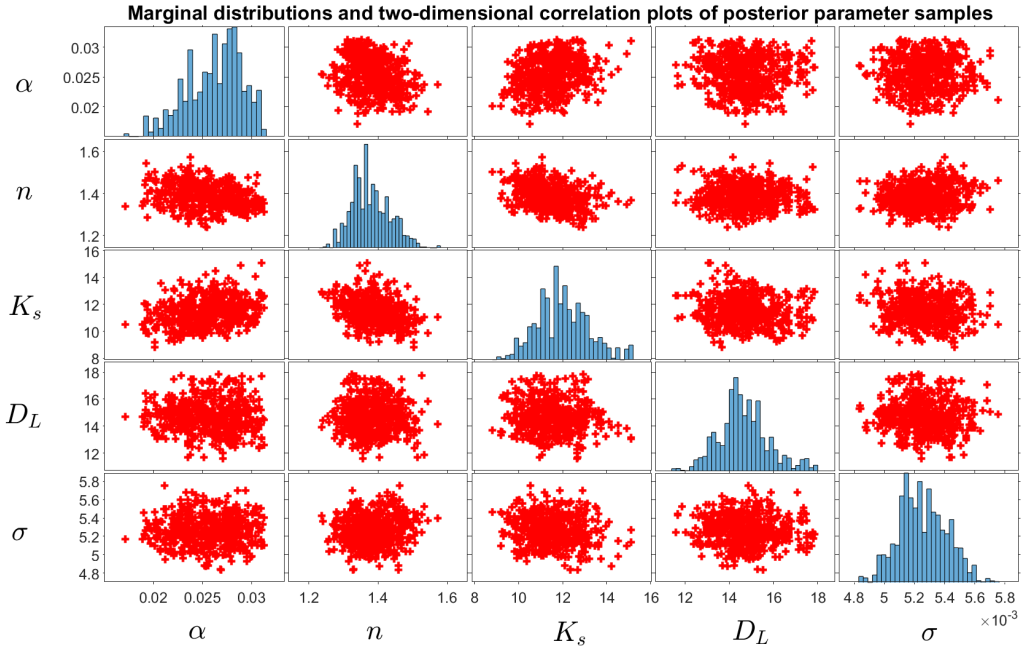


Figure 9: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Flux-Pressure setup.

From Figure 9, the posterior distributions of the calibrated parameters are displayed. For this configuration, the prior distribution of all of the calibrated parameters were narrowed in their posterior distributions, which are visible Gaussian distributions. Posterior samples converge toward a similar value within a small range. Correlations between α, n and K_s are observed, which might occur since these parameters are interrelated in the MvG model equation.

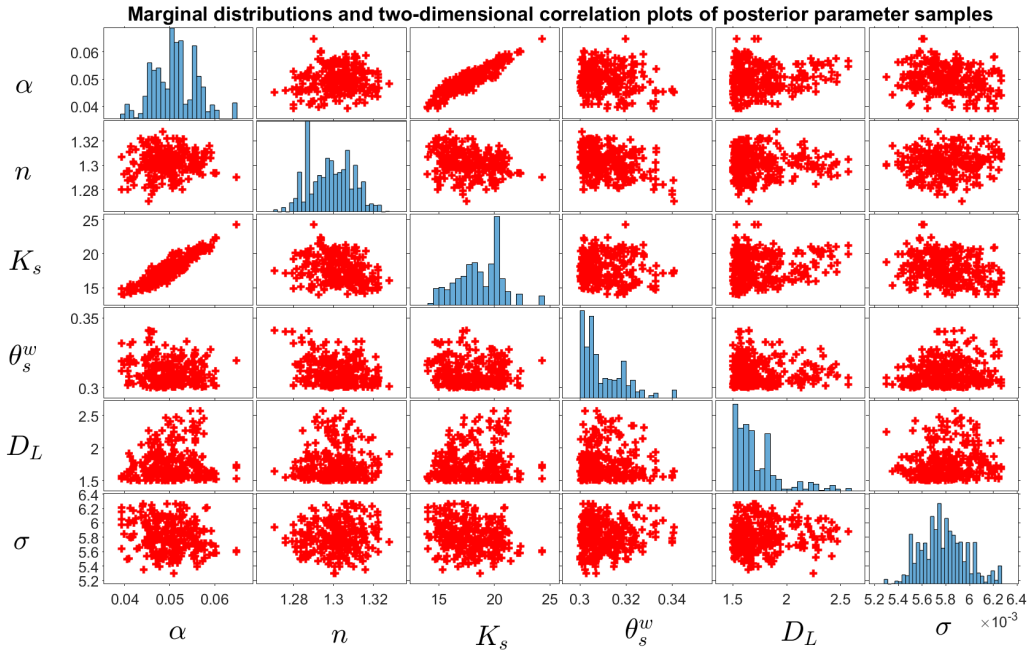


Figure 10: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Flux-Pressure (Cropped HYS) setup.

In Figure 10, it can be seen that the prior range of all of the calibrated parameters were narrowed in their posterior distributions. Especially, posterior distributions for α, n, K_s and σ show a visible Gaussian distribution. For θ_s^w and D_L, the posterior distributions converges towards one bound of the interval and do not look like a Gaussian distribution. Most importantly, for all the parameters, the posterior distributions converges toward a similar value, that is comprised in a small interval. High pairwise correlation are observed between α and n and between α and

K_s, which is possible because these parameters are interrelated in the MvG model equation.

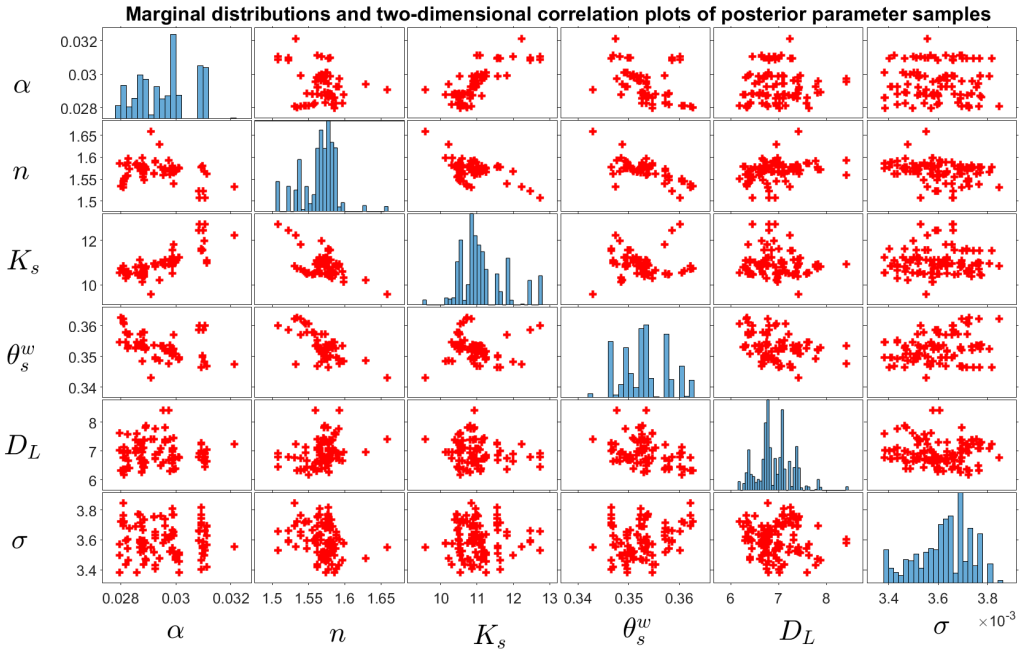


Figure 11: Posterior distributions (on-diagonal) and pairwise correlations (off-diagonal) of posterior parameters samples - Flux-Pressure (HYS) setup.

From the posterior samples results in Figure 11, the posterior distributions for all parameters has been efficiently narrowed toward a small interval with limited possible values. In addition, posterior distributions seems to be a Gaussian distributions, centered in a precise value. Pairwise correlations are observed between posterior samples of the α, n, K_s and θ_s^w parameters, which is might occur given these parameters are interrelated in the MvG model equation.

3.1.2 Obtained parameters

Table 3: Best performance parameters from posterior distributions for each configuration.

Setup	α $[cm^{-1}]$	n [-]	K_s $[cm/day]$	D_L $[cm]$	σ [-]	h_{seep} $[cm]$	θ_s^w [-]
Atm-Seep	0.137	1.305	2.210	29.687	0.008	-	-
Atm-SeepC	0.147	2.352	9.234	24.222	0.007	38.351	-
Atm-Flux	0.149	1.829	36.760	9.379	0.011	-	-
Flux-Flux	0.130	1.311	6.751	13.080	0.011	-	-
Pressure-Flux (Cropped)	0.035	1.248	16.033	12.299	0.011	-	-
Flux-Pressure	0.028	1.356	13.772	14.376	0.005	-	-
Pressure-Flux (Cropped HYS)	0.053	1.288	20.130	1.510	0.006	-	0.303
Flux-Pressure (HYS)	0.032	1.531	12.191	6.682	0.004	-	0.348
Literature reference	0.063	1.390	20.640	-	-	-	-
Hydrus approximation	0.014	1.450	12.860	-	-	-	-

The obtained parameters displayed in the above table results from the MCMC algorithm that finds a set of parameters allowing to reduce the residual between modeled and observed data under a given set of boundary conditions.

From the Mualem Van Genuchten single porosity model, increasing α leads to decreasing the value for the moisture content for a same pressure head value. The best performance values for α are very high for configurations Atm-

Seep/Atm-SeepC and Atm-Flux/Flux-Flux. Thus, the fact that calibrated α is towards the upper boundary means that it allows the model under this boundary conditions setup to reduce the moisture content as much as possible to match the measured data. Also, the order of magnitude for α in these configurations is way larger than the literature values for a loam soil, while for the other configurations - in which pressure and flux boundary conditions were used together - the calibrated value for α is much lower and matches with the order of magnitude of the reference values.

Concerning the n parameter, an increasing value of n also leads to reduce the moisture content for a same pressure head value, according to the model equations for water flow. For the Atm-SeepC and Atm-Flux configurations, it is observed that the calibration predicts a high value for α and for n, meaning that it was needed to additionally reduce simulated moisture contents values to match with the measured data. While for the other configurations, a value of n between 1.25 to 1.45 is usually obtained, and matches with the reference values for this parameter.

For the saturated hydraulic conductivity, this parameter controls the ability of the soil to transmit water under a hydraulic gradient. Now recalling the Richards equation, a high value for K_s would lead to an increasing hydraulic conductivity, and a higher moisture content gradient over time. The reference value is roughly around 12 to 20 cm/day according to literature and Hydrus approximation. A low value were predicted from model calibration in the configurations Atm-Seep/Atm-SeepC and Flux-Flux. A very high value is obtained under Atm-Flux configuration. And values of K_s in the same range as reference values are obtained in the Flux-Pressure/Flux-Pressure(HYS).

The longitudinal dispersivity D_L is used in the CDE equation, increasing the value of this parameter leads to an augmentation of the concentration gradient per unit of time, and the model predicts the solute to expand more rapidly when travelling the soil domain. Thus a higher value for D_L leads to a smoother concentration curve and reduced peak concentration magnitude. A high value for D_L is obtained from calibration with the Atm-Seep/Atm-SeepC, while values comprised between 6.68 and 14.38 are obtained using all the other scenarios.

The seepage face parameter h_{seep} determines the pressure head reached at the bottom node of the domain that allows water to leave the soil column when using a seepage face boundary condition. This parameter were calibrated in the Atm-SeepC scenario, and the obtained value is $h_{seep} = 38\ cm$, meaning that pressure should reach this value at bottom node before flux allowed, allowing to minimize the residual error using this boundary conditions setup.

The θ_s^w parameter controls the saturated moisture content when the soil undergoes a wetting process under air-entrapment assumption in the water flow model. The value obtained for the cropped hysteresis model is $\theta_s^w = 0.3031\ cm$, which is lower than the one obtained for the full domain model under Flux-Pressure (HYS) boundary conditions with hysteresis.

Finally, the σ parameter represent the standard deviation of the residual error between modeled and measured data, thus a small value for this parameter means that the residual error value is roughly equal from one observation to another. Globally, the obtained value for this parameter is in the same order of magnitude for all configurations. Smaller value for σ are obtained in the Flux-Pressure, Cropped (HYS) and Flux-Pressure (HYS) configurations.

3.1.3 Water retention curves

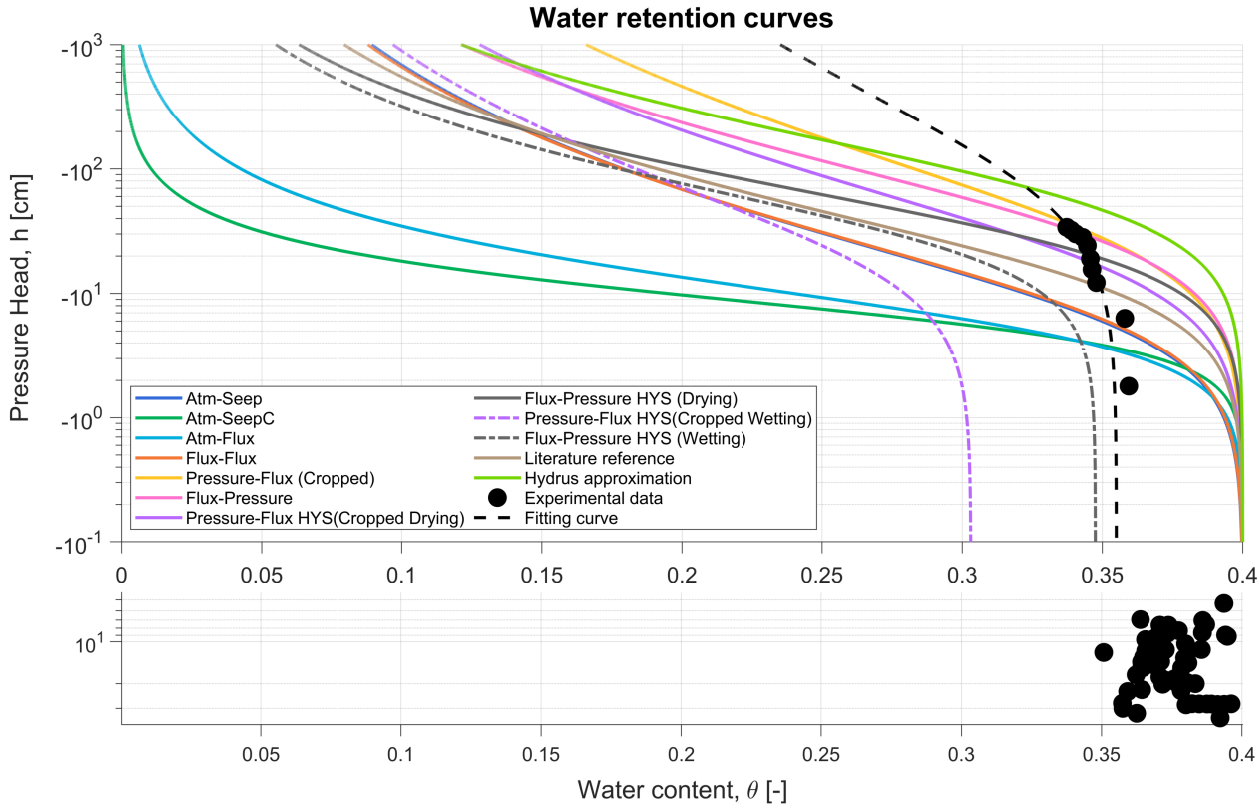


Figure 12: Soil water retention curves for all boundary conditions configurations and experimental data.

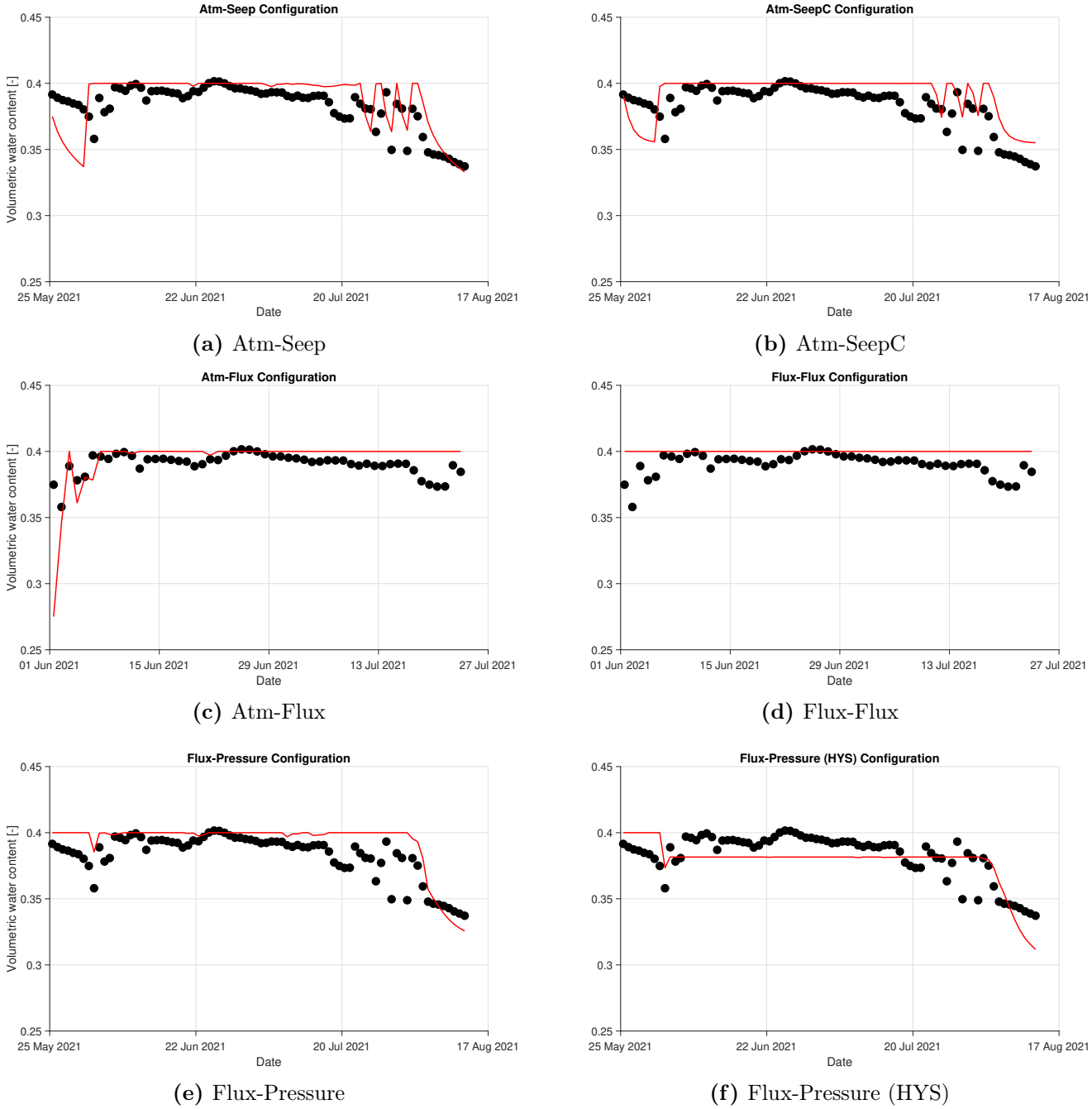
The soil water retention curve provide information about the water movement, the water availability and holding capacity of the soil under study. Comparing soil water retention curves allows to determine if the predicted set of parameters under a given set of boundary conditions have similar water flow characteristics than the predicted parameters under another configuration, or than the reference parameters water flow characteristics. For the configurations with hysteresis model assumption, the drying curve is displayed. Since the measured water content during the experiment is comprised between 0.3 and 0.4, the focus will be on this interval when comparing between configurations.

From the above figure, one can see that the Atm-SeepC and Atm-Flux setups water retention curves have very different characteristics than other configurations and reference parameters. The water retention curves of the Atm-Seep and Flux-Flux configurations are almost identical. The Pressure-Flux (Cropped) and Flux-Pressure have similar water retention curves, with trends similar to the literature and hydrus reference curves. The Pressure-Flux HYS (Cropped) and Flux-Pressure HYS configurations have similar water retention curves, and have also the same trends as the literature and hydrus reference curves.

Nevertheless, none of the retention curves is perfectly matching the one displayed using pressure and moisture content measurements at $z = 20\ cm$ depth collected during the experiment. The fitted curve was adjusted to the measured data by manually tuning the soil hydraulic properties, finding values equal to $n = 1.15$, $K_s = 30\ cm/day$, $\alpha = 0.015\ cm^{-1}$, $\theta_r = 0$ and $\theta_s = 0.355$. From the plot, this curve suggests that two different curves for wetting and drying would maybe allow to fit with the retention curve of the measured data. With the model under air-entrapment assumption with the Flux-Pressure configuration, the predicted saturation water content for wetting ($\theta_s^w = 0.3451$) was lower than the saturation content of the Fitting curve ($\theta_s = 0.355$), also the Flux-Pressure HYS (Drying) curve and the Fitting curve are significantly different, air-entrapment assumption alone is not allowing to fit with measured moisture-pressure relationship.

Still, as visible in the subplot below, most of the data were collected under saturation condition with positive pressure values, the available information under unsaturated conditions used for inverse calibration is not sufficient, more measurements under unsaturated conditions are needed to obtain a precise fitting curve.

3.2 Simulations



(a) Atm-Seep

(b) Atm-SeepC

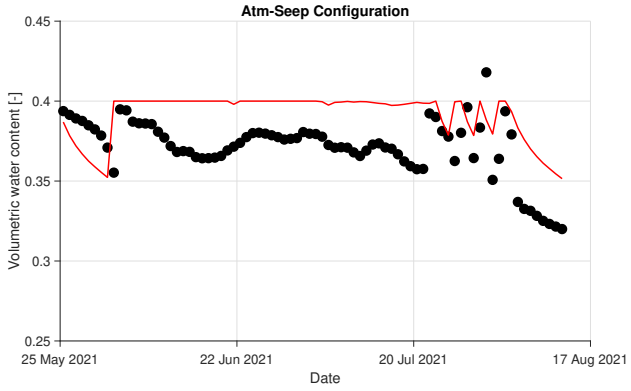
(c) Atm-Flux

(d) Flux-Flux

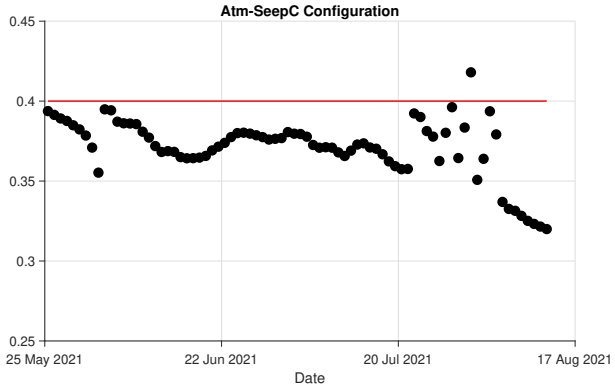
(e) Flux-Pressure

(f) Flux-Pressure (HYS)

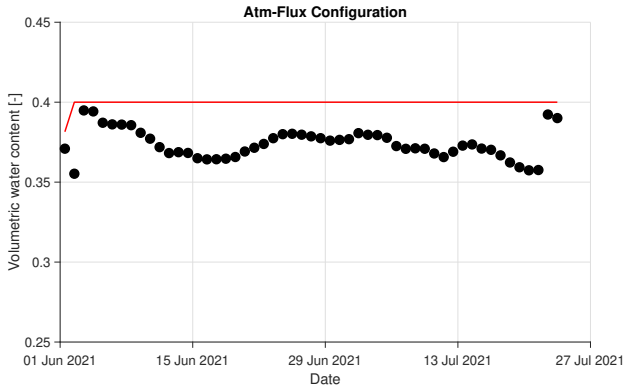
Figure 13: Moisture content at $z = 8\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).



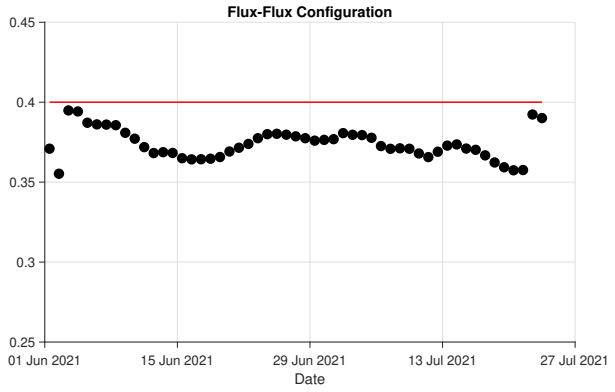
(a) Atm-Seep



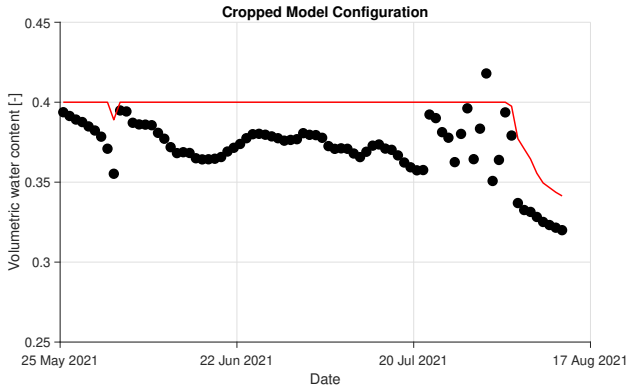
(b) Atm-SeepC



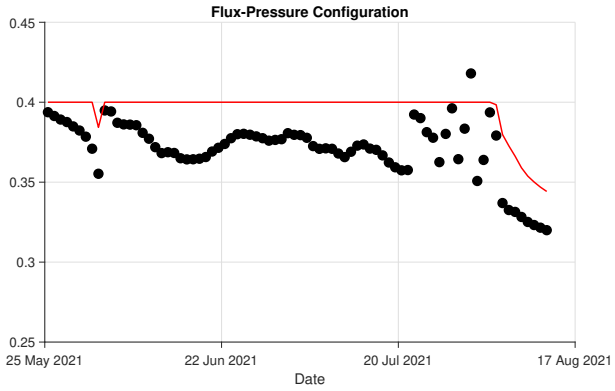
(c) Atm-Flux



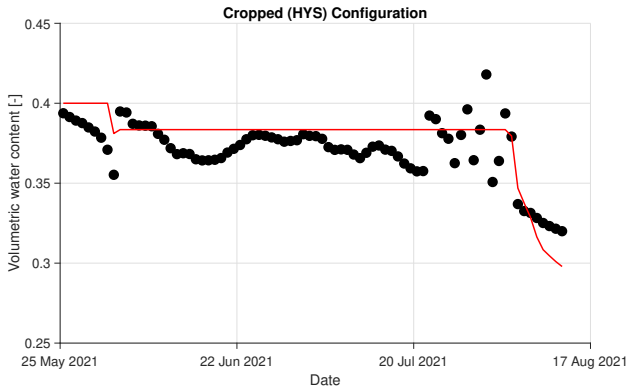
(d) Flux-Flux



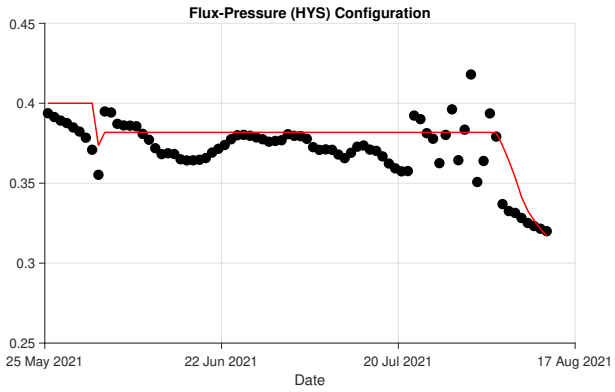
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure

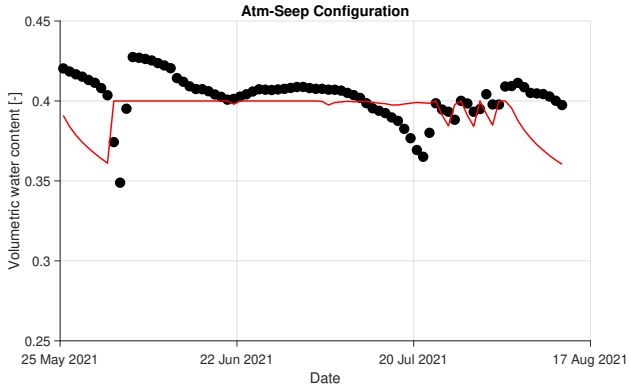


(g) Pressure-Flux (Cropped HYS)

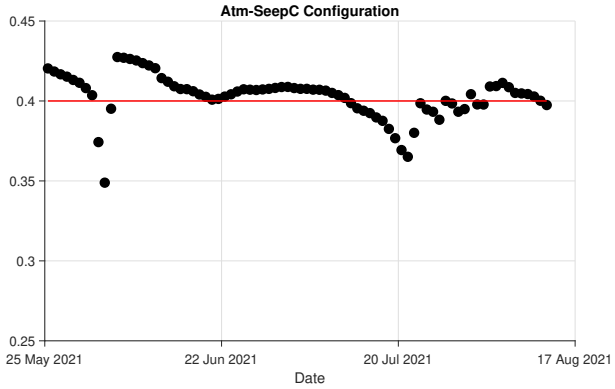


(h) Flux-Pressure (HYS)

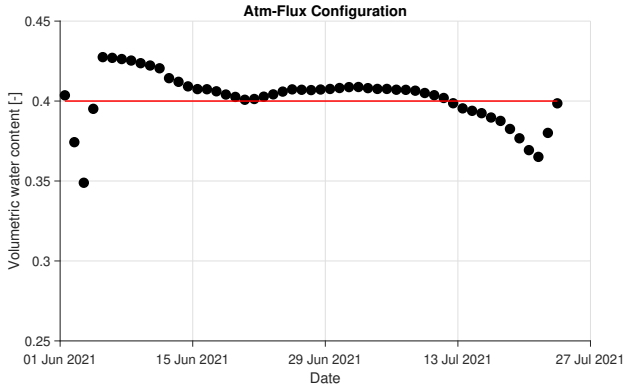
Figure 14: Moisture content at $z = 20\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).



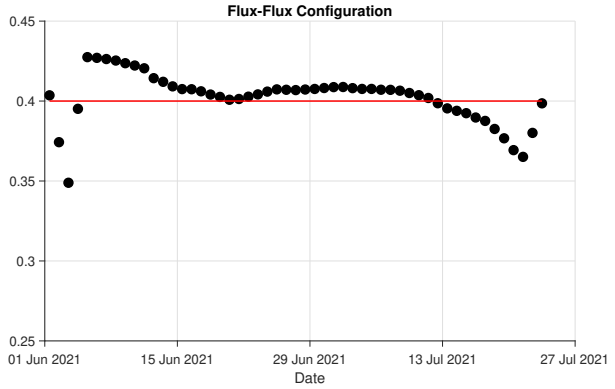
(a) Atm-Seep



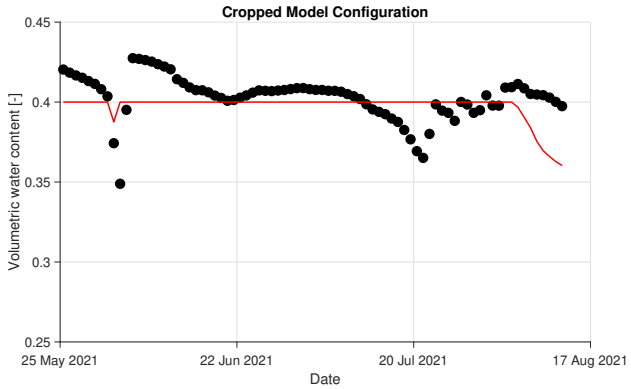
(b) Atm-SeepC



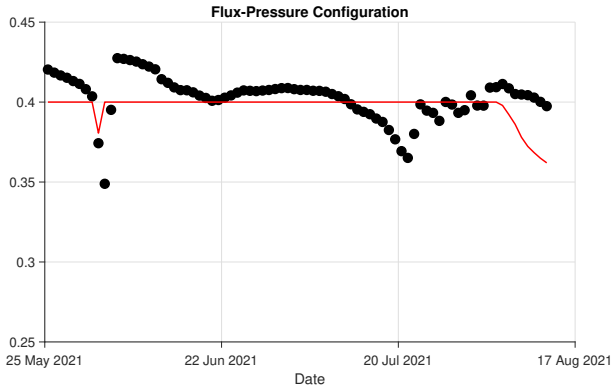
(c) Atm-Flux



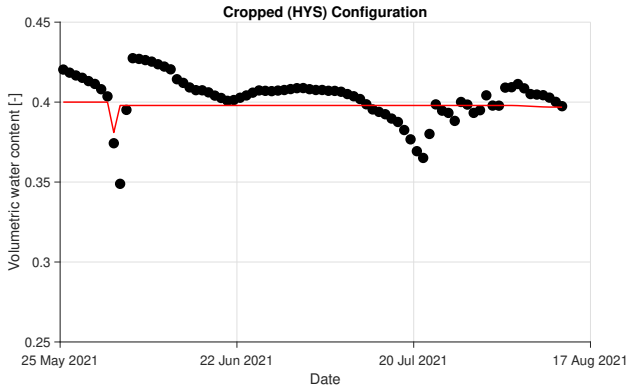
(d) Flux-Flux



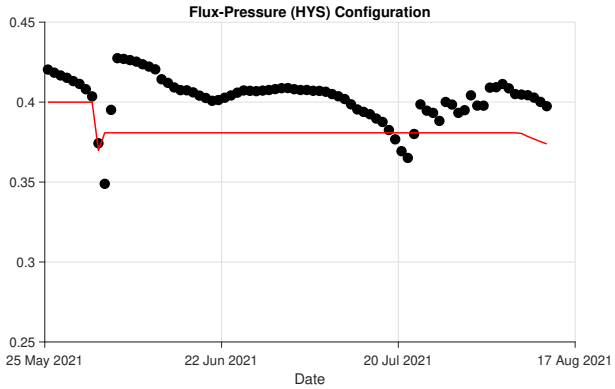
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure

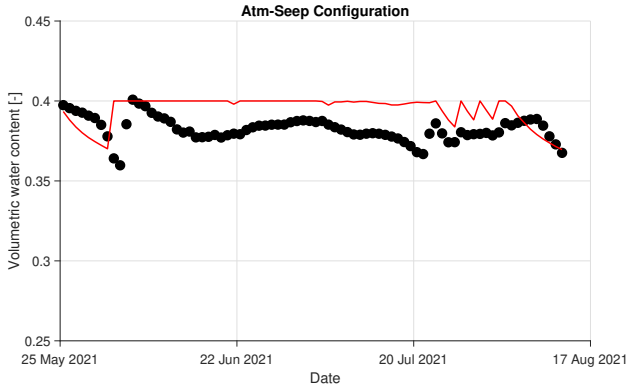


(g) Pressure-Flux (Cropped HYS)

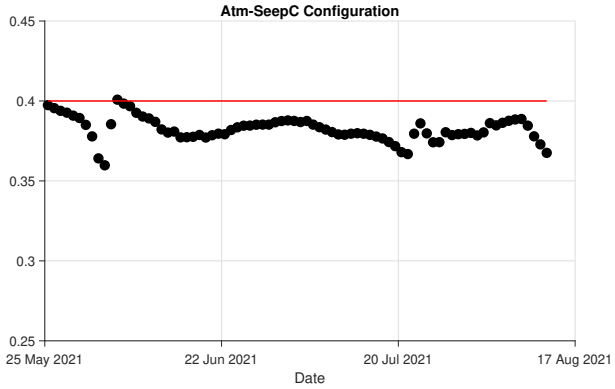


(h) Flux-Pressure (HYS)

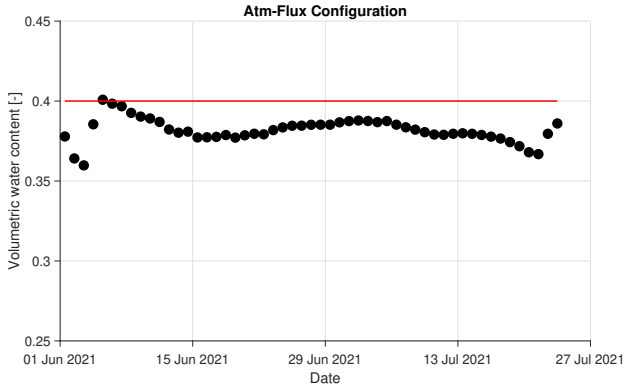
Figure 15: Moisture content at $z = 30\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).



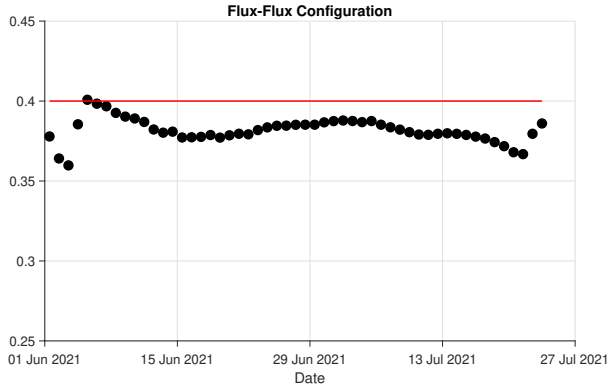
(a) Atm-Seep



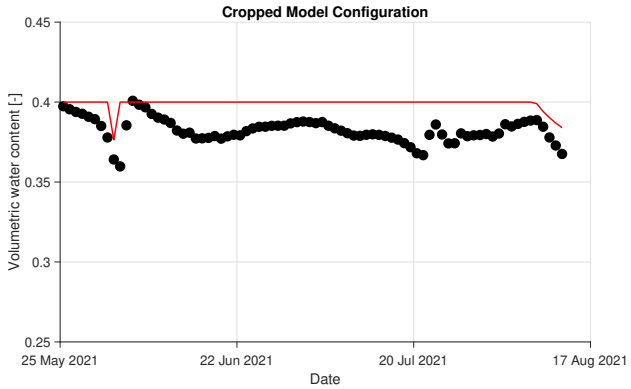
(b) Atm-SeepC



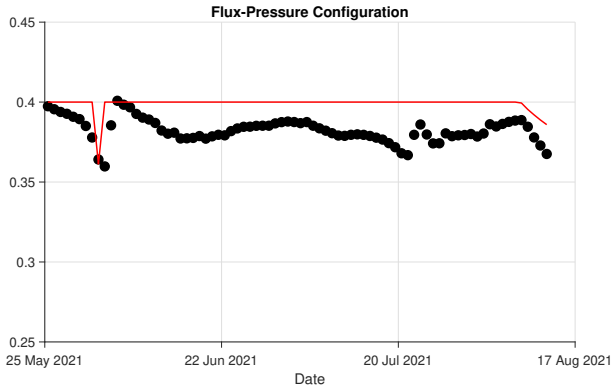
(c) Atm-Flux



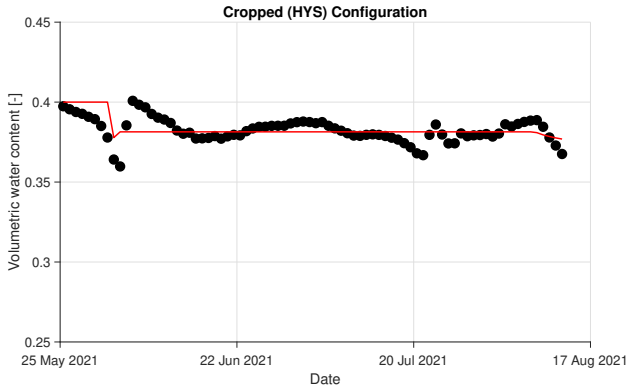
(d) Flux-Flux



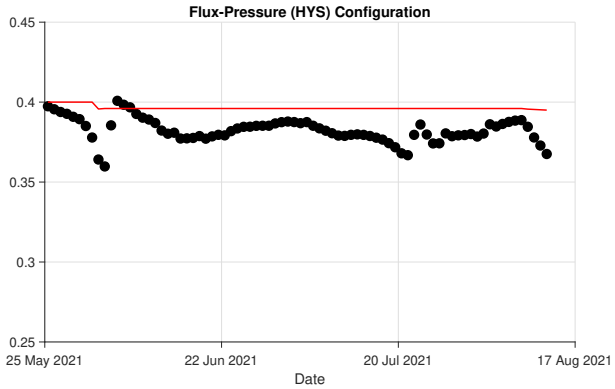
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure

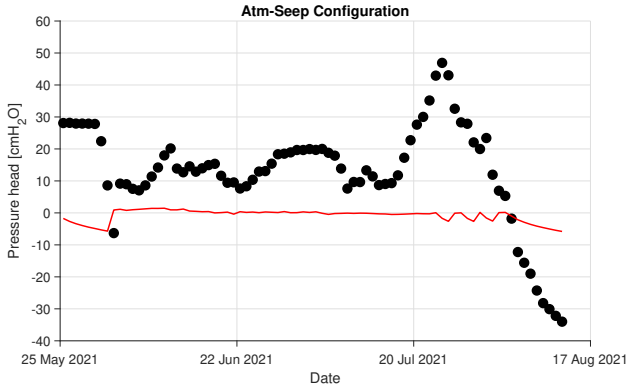


(g) Pressure-Flux (Cropped HYS)

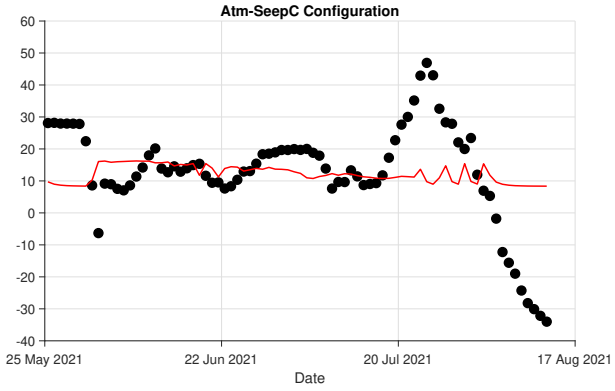


(h) Flux-Pressure (HYS)

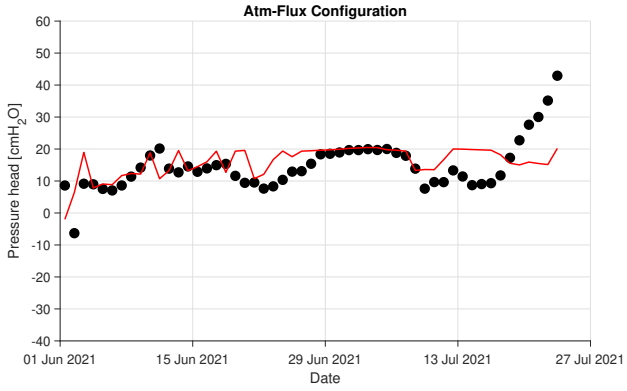
Figure 16: Moisture content at $z = 42\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).



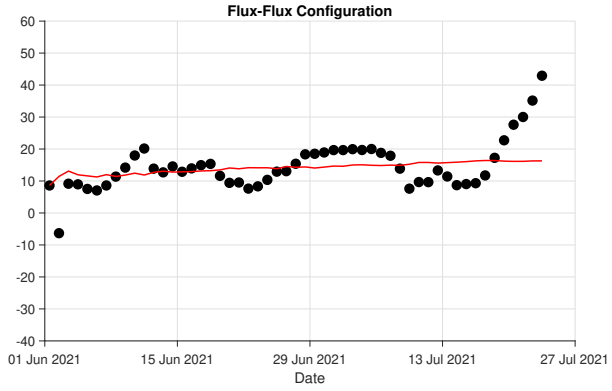
(a) Atm-Seep



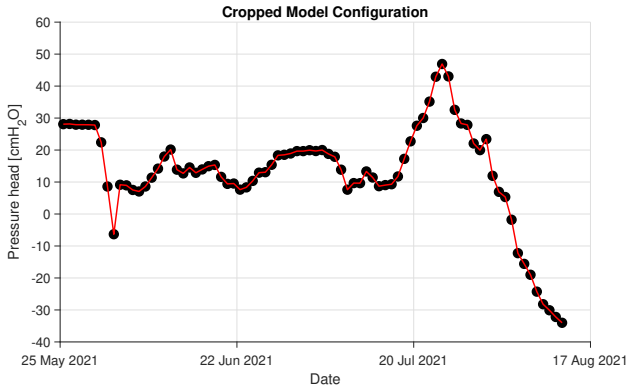
(b) Atm-SeepC



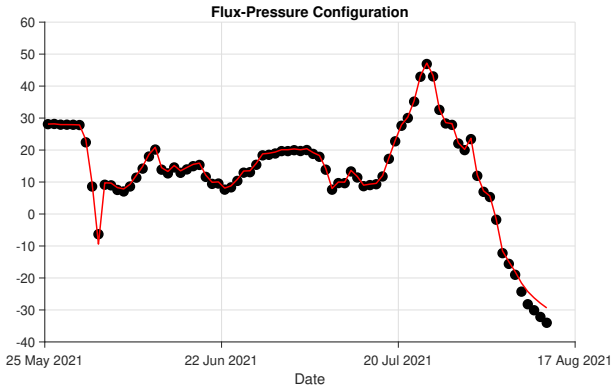
(c) Atm-Flux



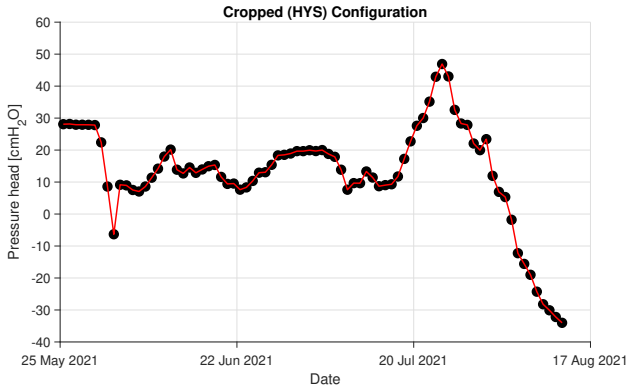
(d) Flux-Flux



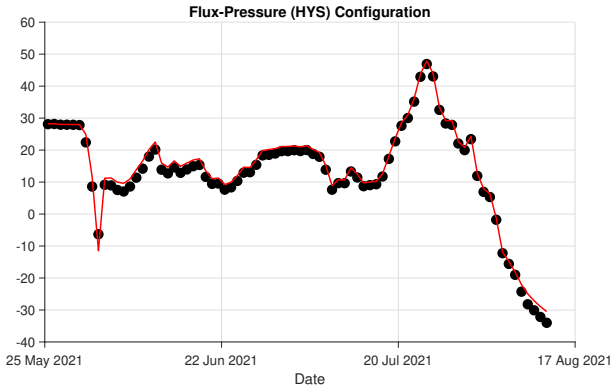
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure

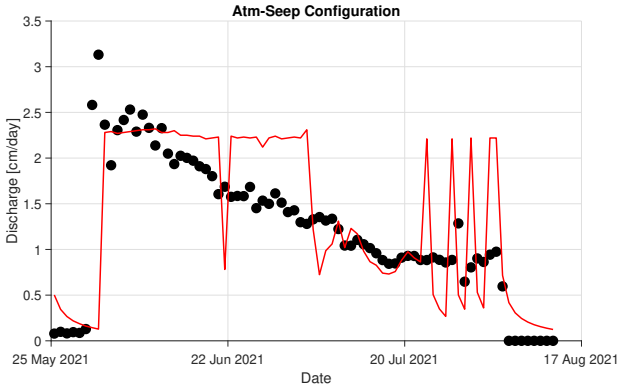


(g) Pressure-Flux (Cropped HYS)

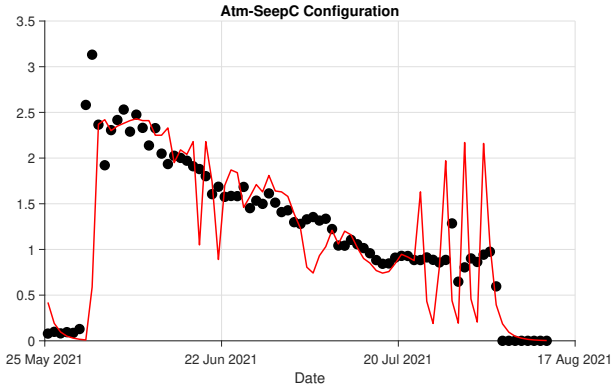


(h) Flux-Pressure (HYS)

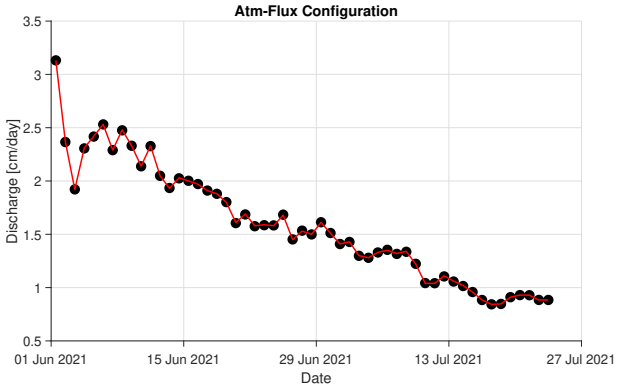
Figure 17: Pressure head at $z = 20\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).



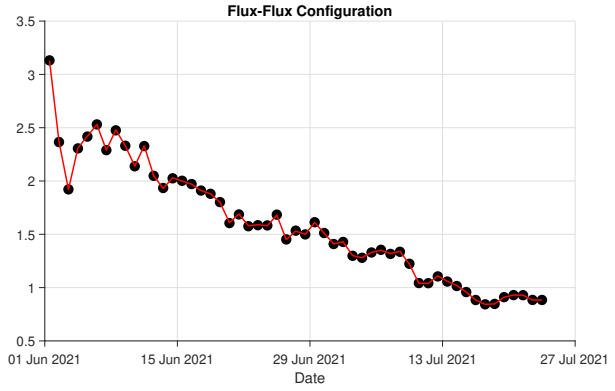
(a) Atm-Seep



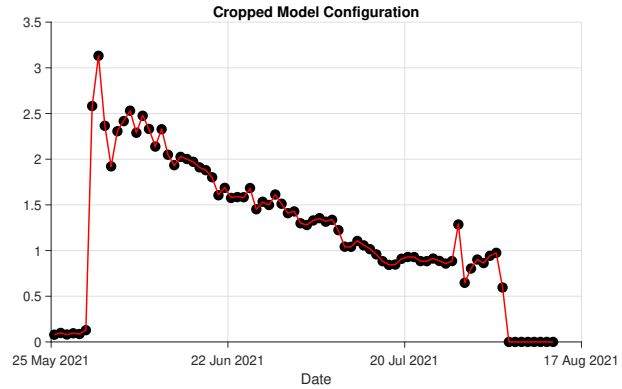
(b) Atm-SeepC



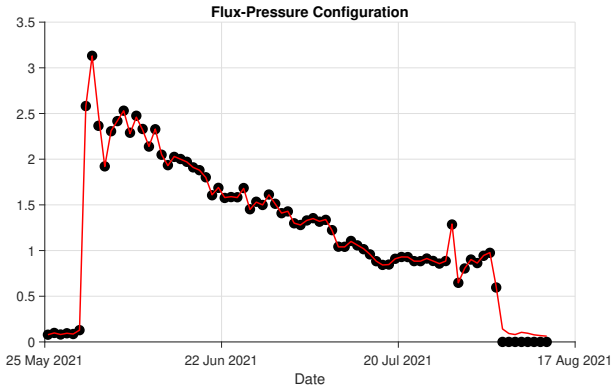
(c) Atm-Flux



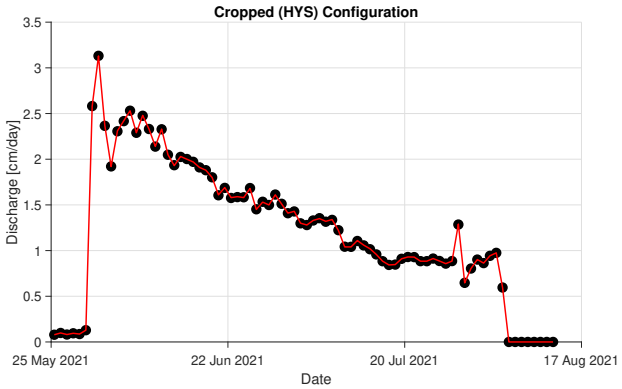
(d) Flux-Flux



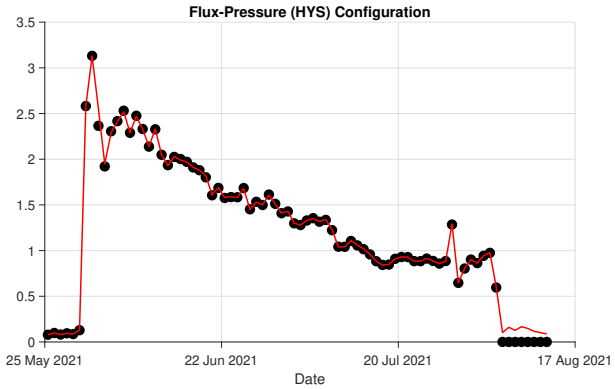
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure



(g) Pressure-Flux (Cropped HYS)



(h) Flux-Pressure (HYS)

Figure 18: Discharge at the outlet $z = 50\ cm$ (black dots are measured values from the experiment, and red curve is simulated values).

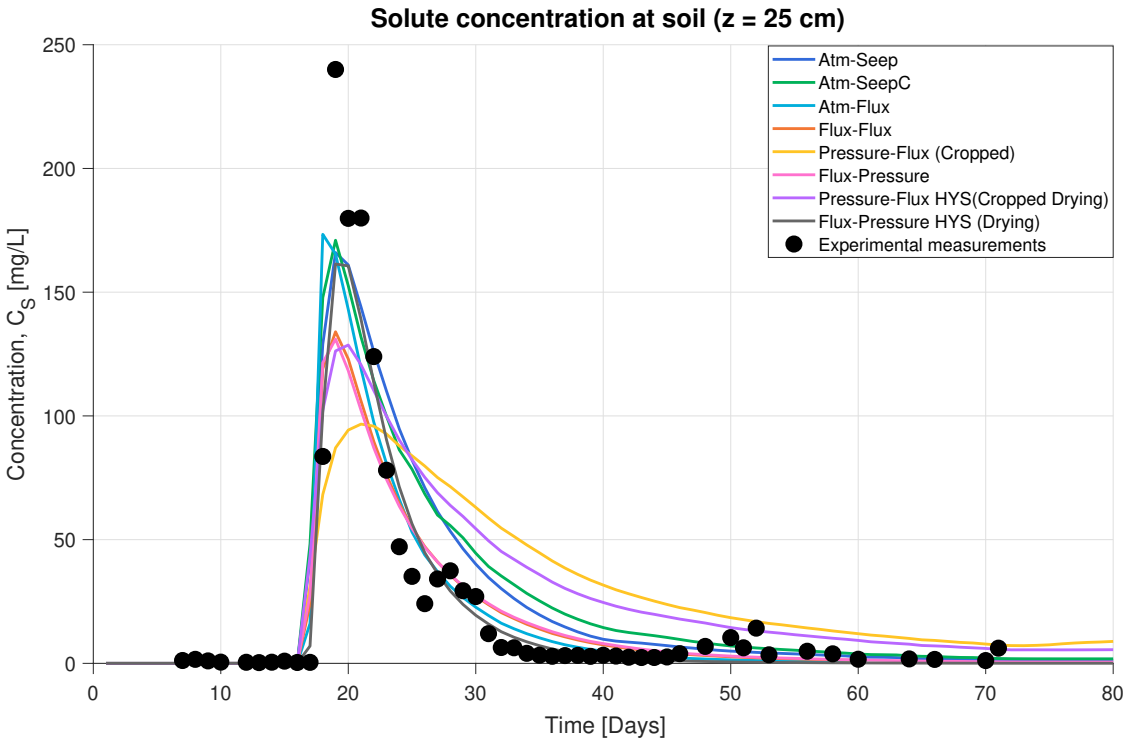


Figure 19: Concentration at soil $z = 25$ cm for all configurations.

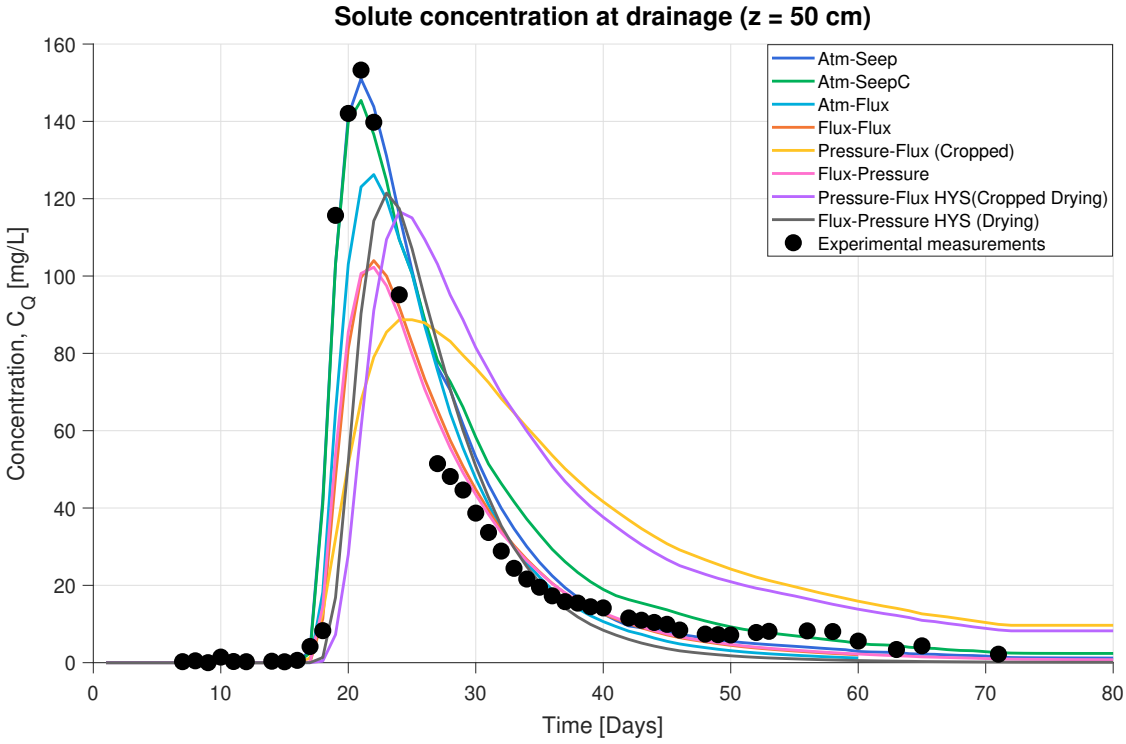


Figure 20: Concentration at the outlet $z = 50$ cm for all configurations.

3.2.1 Moisture content

Figures 13 to 16 provide the results for moisture content simulations at the various observation nodes, recalling the first observation node (at $z = 8$ cm) is not included in the cropped model scenarios.

Concerning the first observation node (Figure 13), as one can see, the Atm-Seep and Atm-SeepC configurations allow to properly catch the dynamics of measured moisture content at this depth. The Atm-Flux setup predicts a very low moisture content at the beginning, and tends to predicts saturation conditions during the entire duration of the experiment similarly to Flux-Flux configurations. Finally, the Flux-Pressure and Flux-Pressure(HYS) correctly predicts the trends observed during the experiment, especially few days after the beginning and near the end of the experiment duration.

For the other observation points (at $z = 20,\ 30$ and $42\ cm$ respectively), one can observe similar results from one observation point to another. The Atm-Seep is properly catching the trends visible in the measured data, while the Atm-SeepC do now predicts saturation conditions during the entire duration of the experiment similarly to Atm-Flux and Flux-Flux configurations. The cropped model and the Flux-Pressure configuration correctly predicts the moisture during the time period at this observation point. Finally, the Cropped (HYS) and Flux-Pressure(HYS) also accurately simulates the trends observed during the experiment.

3.2.2 Pressure head

For the Atm-Seep setup, the pressure head is very badly simulated and the model is not able to catch the highly variable pressure head occurring at the soil column. When the seepage pressure is calibrated (Atm-Seep), the pressure head is more precisely simulated, even though the modeled pressures do not display some of the observed trends in the measured data, the magnitude of the simulated pressure heads are now much closer to the magnitude of measured pressures. For the Atm-Flux and Flux-Flux setups, the trends observed for measured data are also present in the simulated pressures, still, the magnitude of the pressure head do not reach the peak pressure measured near the end of the considered period, and do not exceed 20 cm_{H_2O}. Concerning the Cropped Model and the Flux-Pressure configurations, with and without hysteresis, the simulated pressure is almost exactly equal to the measured pressure, and all the variations are properly simulated.

3.2.3 Discharge

For the Atm-Seep configuration, the modeled discharge do more or less follow the global trend visible in the measurement data, still some modeled peaks are very sharp near the end of the experiment in comparison with the magnitude of the variations of the measured discharge at this period, residuals can reach values up to $1cm/day$. When the seepage face parameter is calibrated in the Atm-SeepC, there is less disparities between modeled and measured discharge, the measured data dynamics are better reproduced, still very high variations of the modeled discharge are visible near the end of the experiment. For the Atm-Flux and Flux-Flux boundary conditions configurations, the flux bottom boundary condition imposes the daily discharge observed at the bottom boundary, which allow to perfectly simulate flux at the lower end of the soil column. Also, for Pressure-Flux (Cropped Model) and Flux-Pressure configurations, with and without air-entrapment model assumption, the discharge is very accurately simulated and perfectly reproduces the measured data dynamics.

3.2.4 Concentrations

For the Atm-Seep/Atm-SeepC configurations, the concentrations are well modeled, the residual error between measured and modeled peaks are not very high given the magnitude order of the concentrations. In particular, the modeled peak is very close to the measured peak concentrations, which highlights that this setup seems to be adequate to catch the solute transport dynamics in the soil column.

For the Atm-Flux configuration, the modeled peak soil concentration is reached a few days before the one observed in the measured data, with a lower magnitude. Also the other hand, for the concentration at the outlet, the magnitude of the modeled peak is closer to the magnitude of the measured peak, but reached a few days later than in the measurements. Similar results are obtained under the Flux-Flux configuration, with a higher difference between modeled and measured peak magnitudes both at soil and at the outlet.

For the Pressure-Flux (cropped model) configuration, it can be seen that the simulated concentrations displays a smoother curve than the measured concentrations. Indeed, the concentration at soil, the modeled peak concentration has a much lower magnitude than the peak visible in the measured data. In addition, the drop in concentration from the peak takes more time in the modeled data than in the measured data. Similarly, in the concentrations at the outlet, the modeled concentration curve display a smoother curve, the peak concentration is observed a few days later in comparison with the peak in measured data.

For the Flux-Pressure setup, the modeled peak concentration at soil occurs at the same time than the measured peak but has an order of magnitude smaller than the one observed in the measured data. Similar findings are

observed for the concentration at the outlet, still the magnitude of the concentration peak is more close to the magnitude of the concentration peak visible in the measured data.

For the Pressure-Flux (cropped model) with hysteresis (HYS), the modeled concentration at soil curve is efficiently centered around the peak, and the drop towards a zero concentration is achieved at the end of the experiment. Still the magnitude of the modeled peak is lower than the measured peak. Similar results are observed for the concentration at the outlet, still the modeled peak at discharge is closer to the measured concentration peak, but occurs a few days later than observed in the experiment.

Finally, for the Flux-Pressure (HYS) setup, the resulting modeled concentrations match properly with the measured concentrations. At soil, the predicted peak concentration occurs at the same day than the one observed in the measured data, with a lower magnitude (difference equal to around 80 mg/L) because the model average the maximum value with the previous and next time steps. For the concentration at the outlet, the peak concentration modeled occurs a few days after the measured peak, there is a difference of about (30 mg/L) which is less than for the modeled concentration at soil.

3.2.5 Errors for observations

In order to evaluate the performance and the residuals magnitude of the simulation for each scenario, the root mean square error (RMSE) for each data type is computed according to the following formula:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{17}$$

The RMSE can be interpreted as the Euclidean distance between the vectors of modeled and measured observations, thus it provides a robust measure to evaluate the accuracy of the model and it ability to reproduce the experiment dynamics under a given set of boundary conditions. Also the total root mean squared error $RMSE_T$ is computed with all the observations used for calibration. Hence, results for each scenario are available in the table below:

Table 4: $RMSE$ for each observation type and total root mean squared error $RMSE_T$ for each configuration using best performance parameters.

Setup	$RMSE_\theta$ [-]	$RMSE_Q$ [cm/day]	$RMSE_p$ [cm]	$RMSE_{C_S}$ $[mmol/cm^3]$	$RMSE_{C_Q}$ $[mmol/cm^3]$	$\mathbf{RMSE_T}$ [-]
Atm-Seep	2.08×10^{-2}	0.67	20.52	2.71×10^{-4}	1.17×10^{-4}	7.63
Atm-SeepC	2.22×10^{-2}	0.55	15.58	2.88×10^{-4}	1.37×10^{-4}	5.81
Atm-Flux	2.05×10^{-2}	2.86×10^{-3}	7.42	2.71×10^{-4}	1.62×10^{-4}	2.68
Flux-Flux	1.97×10^{-2}	2.86×10^{-3}	7.01	2.96×10^{-4}	2.21×10^{-4}	2.53
Cropped	2.09×10^{-2}	2.40×10^{-3}	3.04×10^{-3}	4.72×10^{-4}	4.07×10^{-4}	0.02
Flux-Pressure	1.97×10^{-2}	3.29×10^{-3}	1.14	2.92×10^{-4}	2.00×10^{-4}	0.43
Cropped(HYS)	1.24×10^{-2}	2.40×10^{-3}	3.04×10^{-3}	3.69×10^{-4}	4.51×10^{-4}	0.01
Flux-Pressure(HYS)	1.39×10^{-2}	4.61×10^{-2}	1.73	1.84×10^{-4}	2.96×10^{-4}	0.65

First, what can be realized is that the seepage face scenarios (Atm-Seep/Atm-SeepC), in contrast with other scenarios, both produces very high residual errors when simulating pressure head at a depth of $z = 20$ cm. Also, the produced residuals for discharge are also important, around ten times larger than in most of the other scenarios. When it comes to moisture content, the residuals are lower and has the same order of magnitude as other scenarios. Finally, simulated concentrations are satisfying, in particular for the simulated concentration at discharge, which has the lowest RMSE of all the scenarios. When the seepage face parameter were calibrated (Atm-SeepC), the RMSE were lower than in the default configuration for seepage face (Atm-Seep), but most importantly, this boundary condition setup do not permit to simulate the high peaks of pressures visible in the measured data, as described in the results section. Overall, the $RMSE_T$ is lower when calibrating seepage face parameter.

Concerning the configurations Atm-Flux/Flux-Flux, the residual error for pressure head are half lower than the one observed in the seepage face configurations: using a bottom flux boundary condition increases the quality of simulated pressures, but the magnitude remains higher than when this boundary condition were coupled with a pressure boundary condition at one boundary of the domain. Most importantly, this boundary conditions setup allows to fix the simulated discharge to the one measured, leading to a very low RMSE for discharge. When it comes to moisture content, the simulated values has a slight lower RMSE than in the seepage face configurations, but globally in the same order of magnitude than other setups. For concentrations the RMSE is also in the same order than in other configurations, with a relatively low error for the outlet concentration. Generally, the overall RMSE is lower than in seepage faces setups.

Now, for the Pressure-Flux (cropped model) setup, the simulation produces very low errors for all of the simulated data. Especially for pressure, the obtained error is a lot lower than for all of the model where all the soil domain is considered: using a variable pressure boundary condition is significantly improving the quality of the simulated pressures. Nevertheless, using a cropped model with approximated simulated concentrations as upper concentration flux boundary condition increases significantly the RMSE for concentration (2 to 3 times larger than other setups). Still, given the very low RMSE for θ, Q and pressure p, this model provide a good approximation of the pressure head occurring at the observation points of the column, and specially at the bottom boundary. In this case, the overall RMSE is the one of the lowest ones comparing with other configurations.

When used to put in place a pressure boundary condition in the Flux-Pressure setup, the RMSE were considerably reduced compared with all of the other full column setups (Atm-Seep/Atm-SeepC and Atm-Flux/Flux-Flux). In fact, the RMSE of pressure is 7 times lower than in Atm-Flux/Flux-Flux setups and 15-20 times lower than in Atm-Seep/Atm-SeepC setups. For volumetric water content, the RMSE is also lower than in other configurations. When it comes to concentrations, the RMSE is in the same order as in previous setups, revealing changing flow boundary conditions do not significantly impact the simulated concentrations. Also, the $RMSE_T$ for the best performance parameters considering all the observations used during calibration is very low in comparison with the other setups. Still, overall the error is higher than in the Cropped model configuration for pressure, but one should keep in mind that in this configuration the pressure at the boundary are not taken from measurements but from Cropped model simulation, which increase the uncertainty when modeling. The $RMSE_T$ is higher than in the Cropped model setup, but drastically lower than in the other scenarios.

Then, for the Pressure-Flux (cropped model) (HYS), the $RMSE_T$ is half of the one obtained from the standard Cropped model setup. In particular, the RMSE for θ is also lowered when considering hysteresis in the cropped model, but RMSE for Q and pressure p are equal to the ones from Cropped Model setup. Also, the RMSE for concentration are in the same order as in Cropped Model configuration, with a lower RMSE for C_S and a higher one for C_Q.

Finally, for the complete soil domain Flux-Pressure (HYS) configuration with hysteresis in the soil hydraulic properties, the $RMSE_T$ is higher than in the standard model Flux-Pressure setup, knowing they have similar boundary conditions setups. The main improvement from adding hysteresis in the model is that it reduces the RMSE for moisture content, but produces a higher error for discharge. Concerning the concentrations, the RMSE is lower for C_S and higher for C_Q when comparing with Flux-Pressure configuration.

4 Discussion

4.1 Simulation and calibration

From the simulation results, it has been deduced that the boundary conditions achieving the lowest root mean square errors are the ones where flux and pressure boundary conditions were implemented in the model, both with and without hysteresis. Globally, using this configuration allow to properly simulate the flow dynamics during the experiment: moisture content, discharge and pressure head dynamics are efficiently modeled, clearly better that under Atm-Seep/Atm-SeepC configurations. Still, the performance of the modeled concentration is not the best using these configurations, especially for the cropped models when input modeled concentrations were used at the upper solute transport boundary conditions. Indeed, configurations such as Atm-Seep/Atm-SeepC showed a better robustness and modeled peak concentration much closer to what was observed during the experiment. However, these two setups do not properly simulate the flow which is problematic since solute transport modeling also relies on flux modeling. Alternatively, other configurations such as Atm-Flux and Flux-Flux provide more or less adequate results in terms of water flow and efficient solute transport modeling, but hardly reproduce the trends for pressure head at soil. Additionally, using a flux bottom boundary condition do not permit to simulate the days where a null-discharge was occurring, leading the model to diverge during these time periods.

Now, given that the MCMC method is able to properly find the best parameter set in a wide prior range research space, it is essential to verify that simulation results are obtained from a meaningful and reasonable set of parameters, obtained from Gaussian distributions on a narrowed interval, and with values matching expectations from literature.

From the calibration results, it has been showed that the Atm-Seep/Atm-SeepC and Atm-Flux/Flux-Flux were not able to properly narrow the prior range of parameters in their posterior distributions, and additionally did not displayed Gaussian distributions for all the parameters. In particular, the obtained value for α is highly suspicious and unusually high in comparison with literature and hydrus reference values for the type of soil studied in the experiment, this affirm that using this boundary conditions configurations do not lead to reliable results for modeling water flow, even with this extreme value detected as an optimum. Oppositely, the Flux-Pressure and Pressure-Flux (cropped model) configurations successfully achieved narrowing the prior range in the posterior samples of parameters and very apparent Gaussian curve were visible. This confirms that these sets of boundary conditions are the most reliable to properly simulate the water flow, obtain meaningful and precise set of parameters from posterior distributions, even if the solute transport modeling achieves a lower performance comparing with the one obtained under other setups. Also, from the Richards equation, one can intuitively deduce that using pressure combined with flux boundary conditions allow to obtain a unique solution for the equation, which is not the case if only flux is used as a boundary condition, because it can have multiple pressure values corresponding to a same flux value, explaining the divergence for null discharge time periods. However, if neither flux or pressure measurements are available, the Atm-SeepC can be used, but it is not expected that this configuration allow to properly simulate the pressure dynamics occurring during the experiment.

4.2 Cropped pressure dynamics

To understand the pressure dynamics occurring in the soil column, the cropped model with Pressure-Flux setup is used to simulate the pressure head at the lower part of the domain under known boundary conditions.

The simulated pressure heads in multiple observation points of the cropped model using the best performance parameters are displayed in Figure 21. As shown in the figure, the pressure head at the bottom boundary ($z = 50\ cm$) is initially very high as the column is close to a saturation state, then a negative peak occurs due to high discharge, and at the end of the experiment the pressure head becomes negative as the volumetric water contents measured are lower than in the beginning. Globally, the pressure head is highly variable at every observation points of the cropped model.

Figure 22 displays the simulated pressure head at the bottom boundary and the empirically measured discharge exiting the soil column. One can see that the simulated pressure at the bottom of the soil column is highly variable and do not show any particular trend jointly with the discharge values, in fact there is no particular pressure head value for which the flux would become possible at the bottom boundary. However, for a seepage face condition, it is needed to define a pressure head starting from which water can freely exits the soil column from the bottom boundary. The default seepage pressure $h_{seep} = 0\ cm$ and the calibrated seepage pressure $h_{seep} = 38\ cm$ are displayed. Hence, the calibrated seepage pressure is optimized such that it fits with the average pressure head occurring at the bottom boundary to reduce the error, but flux is occurring at pressure values below h_{seep} so this bottom boundary condition fails at reproducing the flux and pressure dynamics observed in the data, as observed

in the simulation results for bottom seepage face configurations.

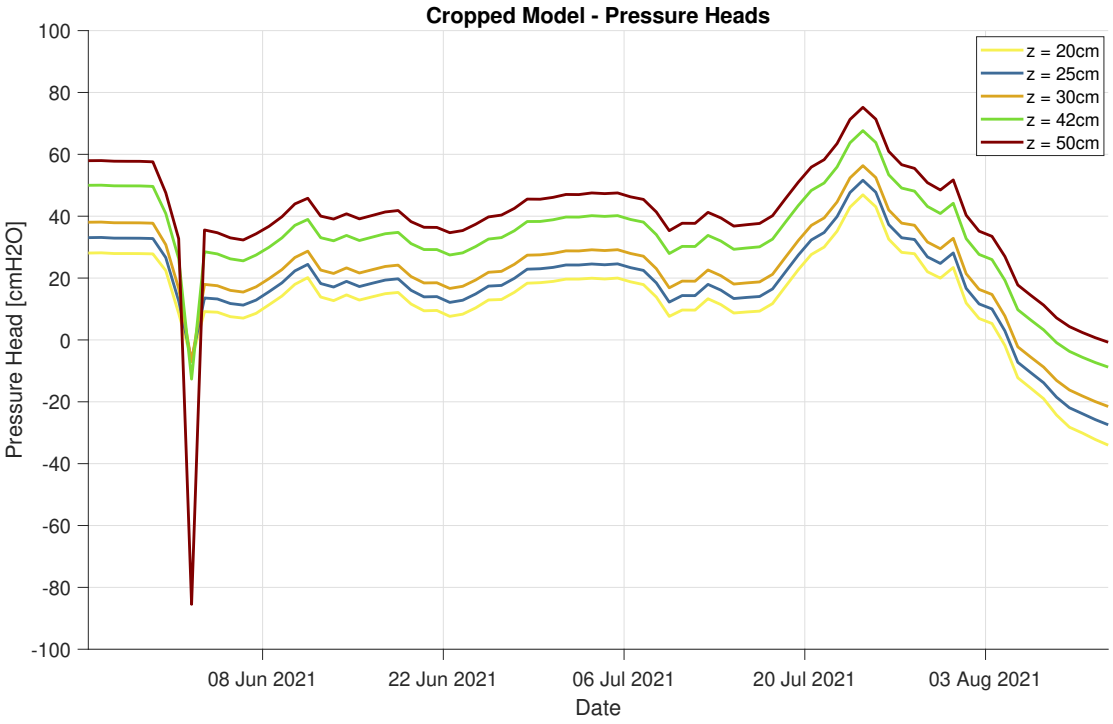


Figure 21: Pressure heads simulated at bottom observation nodes - Pressure-Flux (Cropped) Setup.

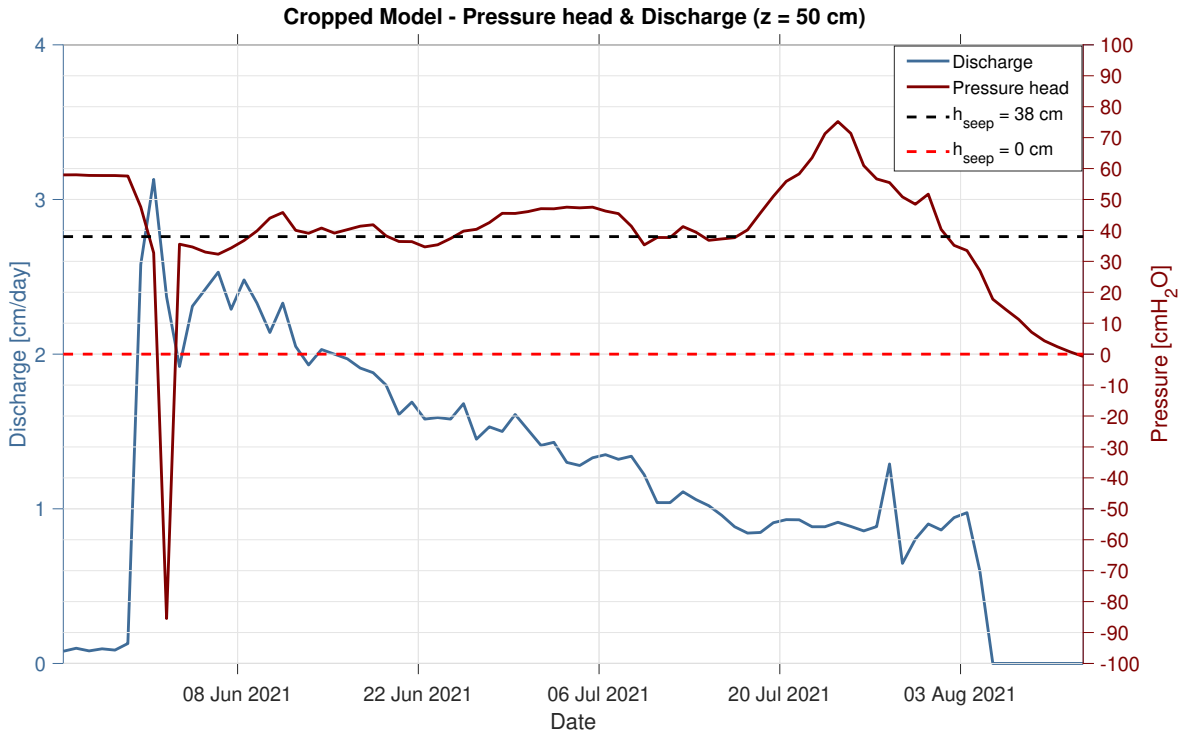


Figure 22: Pressure heads and discharge simulated - Pressure-Flux (Cropped) Setup.

4.3 Evaporation assumption under atmospheric boundary conditions

In order to verify the evaporation assumption previously implemented under atmospheric boundary conditions, the cropped model configuration is useful because it allows to simulate evaporation under known boundary conditions, with variable pressure head from measurements as upper boundary condition and variable flux from measured drainage as bottom boundary condition. In fact, not using an upper atmospheric boundary condition allows to estimate the occurring evaporation from the model simulation.

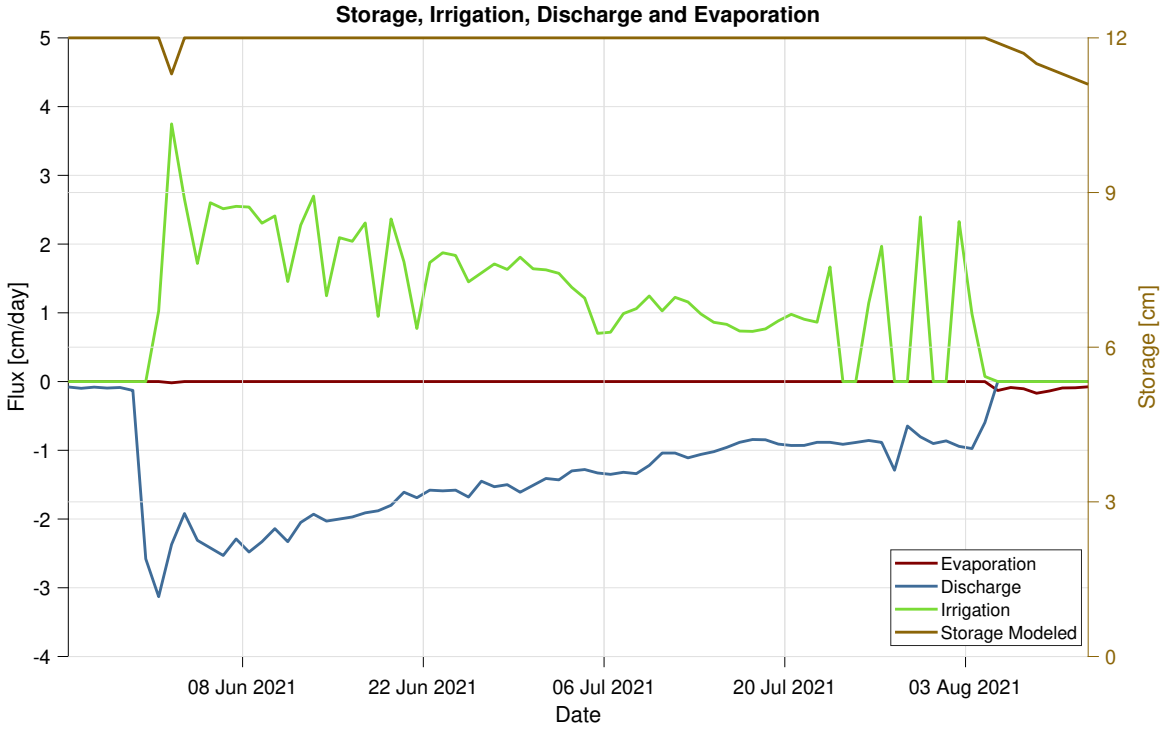


Figure 23: Irrigation, Simulated Storage, Discharge and Evaporation - Pressure-Flux (Cropped) Setup.

The above figure displays the simulated fluxes at different observation points in the cropped model. From the figure, the evaporation simulated by the model is close to zero during the entire duration of the experiment, except at the end where the column is no longer close to saturation conditions (unsaturated conditions) and there is no flux across the bottom boundary.

As described in the experimental layout, the discharge is calculated using weight difference, and assumes null evaporation fluxes by default. To tackle this issue, another simulation is done with a discharge reduced by 10% to verify if the model still predicts negligible evaporation. The table below summarizes the total volume of irrigation, discharge and evaporation simulated using the cropped model, with empirically measured discharge and reduced discharge. Also, the ratio of total evaporation by the average water storage in the column (equal to 20.47 cm) is computed. From the table, the total evaporation modeled represent less than 5% of the average water storage of the column, even with reduced discharge. From these results, it is fairly correct to maintain the negligible evaporation assumption.

Table 5: Irrigation, Discharge and Evaporation total volumes - Pressure-Flux (Cropped) Setup.

Flux	Pressure-Flux	Pressure-Flux (reduced discharge)
Irrigation [cm]	92.27	92.27
Discharge [cm]	98.14	88.33
Evaporation [cm]	0.90	0.92
Evaporation/Storage [%]	**4.4**	**4.5**

4.4 Air-entrapment model assumption efficiency for simulation

As observed in the simulation residual errors results (Table 4), running the full domain model under air-entrapment assumption with the most efficient boundary conditions configuration do not achieve better results in term of total root mean squared error. Even though it allows to slightly better model the moisture content dynamics and reduce $RMSE_\theta$, it do not allow to better model the other water flow and solute transport variables. For the cropped model, it was indeed contributing to modestly reduce the $RMSE_T$, but for the entire domain the standard model under Flux-Pressure boundary conditions achieve better results than the one with air-entrapment assumption. The obtained parameters from calibration in common for the two models are similar, either using the cropped model or the full model under flux-pressure boundary conditions. The efficiency of including air-entrapment in the model globally reduces the accuracy of the simulation for the full domain model, but increases it for the cropped model. Hence, including air-entrapment in the model leads to different results depending on the boundary conditions configurations used.

It was observed that under air-entrapment assumption, with best boundary conditions configuration, the peak concentration modeled was closer to the measured peak (both at soil and at drainage) than the peak simulated with same boundary condition but without the air-entrapment assumption. If tracking the peak concentration is the main focus of modeling, then using the air-entrapment assumption is in fact useful. According to the water retention curves, the fitting curve to the measured data could be obtained with a model under hysteresis, but under air-entrapment assumption only the obtained drying curve was not fitting the data.

4.5 Next steps

The investigated approach to determine the fluid flow boundary conditions implemented in this study needs to be applied to a system with precisely measured boundary conditions, or with the water flow and solute transport parameters known with precision. Then the approach would be used to calibrate parameters under different boundary conditions configurations, or to try different configurations if parameters are known. This is to verify if the setup leading to lowest residual error is similar to boundary conditions setup measured during the experiment. If the best performance boundary conditions setup found is matching the ones measured, and the set of parameters obtained is on average equal to the one obtained from calibration under known boundary conditions (or from known parameters), then the method could potentially be applied for other types of soils and experiments with complex boundary conditions to determine a configurations allowing to properly model the experiment.

Concerning the evaporation assumption, the validity of the assumption was verified by simulating evaporation in a cropped model under known boundary conditions. However, the bottom flux boundary condition was set using the discharge values computed based on weight difference, but this method is neglecting evaporation by default. Thus, it is needed to simulate the cropped model using discharge values from direct flux measurement at the lower end of the lysimeter, ideally under precisely measured evaporation, to check if the approach for simulating evaporation using this cropped model is correct, or leads to results different from what is measured during the experiment.

In this study, only the air-entrapment assumption in the model was investigated, and only θ_s^w was calibrated as a parameter differentiating the wetting and drying processes in the hysteresis model. Further, it should be necessary to calibrate all of the parameters of the Mualem Van Genuchten single porosity model, to evaluate if accounting for hysteresis in the model leads to better results in terms of root mean squared error and posterior parameters distribution. It could be possible that under calibration of all the soil hydraulic properties in a complete hysteresis model assumption for all soil hydraulic parameters, the soil water retention curve fits better with the measurement pressure-moisture content relationship. However, more data points for unsaturated conditions are needed to provide more information for calibrating the soil hydraulic properties.

Depending on the scale of the experiment, neglecting the diffusion process could not be always valid, and one should find proper prior range for this parameter to be calibrated together with model parameters. Also, if a non-conservative (i.e reactive) solute is injected, the solute transport model equations will need to consider decay depending on the undergoing chemical equations in place. Finally, it could be necessary to account for the temperature dependence of the soil hydraulic properties if the experiment is conducted in variable temperature conditions.

5 Conclusion

The present study focused on testing an approach to determine the best water flow boundary conditions configuration to accurately model a suction controlled lysimeter experiment conducted under laboratory conditions. During the experiment, several data were collected from probes installed on the lysimeter, including the moisture content, the discharge, the pressure head, the concentrations at soil and drainage. The water flow model were modeled based on the combination of the Richards equation and Mualem Van Genuchten single porosity model, with and without air-entrapment model assumption. The solute transport model was relying on the convection-dispersion equation, where the impact of diffusion was considered to be negligible. The calibration of the water flow and solute transport unknown parameters were performed using the Bayesian inverse modeling MCMC Dream algorithm based on measurement data. It allows to estimate posterior parameters distributions from the prior ranges of parameters defined in a wide interval covering reasonable values recorded for different types of soil.

The calibration process was performed under multiple configurations for boundary conditions, and quality of the posterior distribution samples was assessed while the values of the parameters were compared with reference values from litterature. Also, the observations errors for each simulation using calibrated best performance parameters under a given configuration were computed and compared with other configurations. Using measurements collected during the experiment, the evaporation fluxes were simulated using a cropped model to verify the validity of the negligible evaporation assumption. Finally, the impact of considering air-entrapment in the model is assessed under best performance boundary conditions configuration found previously. The interpreted results are outlined below:

- From posterior parameters distributions and simulation errors assessments, the best performance boundary conditions setup for this experiment is the combination of an upper flux boundary condition with a bottom pressure bottom boundary conditions, which is also the most used setup in several inverse Bayesian modeling for water flow and solute transport in various lysimeter experiments (e.g. Younes et al., 2017; Younes et al., 2018; Younes et al., 2019.)
- Other boundary conditions setups do not allow to properly reproduce measured data dynamics, especially the pressure head that is highly variable at all depths within the soil column as suggested by the cropped model simulation. Specifying a flux boundary condition without combining it with pressure boundary condition diverges for days with null-discharge, and do not allow to model pressure head variations measured during the experiment.
- If neither precise flux or pressure measurements are available at the boundary, it is still possible to use upper atmospheric conditions with a bottom seepage face boundary condition with a calibration of h_{seep}. However, for this experiment, it do not allow to properly simulate the pressure head fluctuations observed in the measured data.
- Evaporation fluxes simulated in the cropped model has very small magnitudes in comparison with the average water storage, and could then be neglected when considering an upper atmospheric boundary conditions. Still the bottom boundary condition was set using discharge indirectly measured from weights changes. The approach needs to be verified using direct flux measurements and ideally measured evaporation.
- Air-entrapment assumption in the model for the entire soil domain do not lead to reduce the residual error between modeled and measurement data, nor to improve the quality of posterior distribution obtained from Bayesian inverse modeling, under similar best performance boundary conditions. On the contrary, it allows to slightly reduce the total root mean squared error in the cropped model. Efficiency of considering air-entrapment in the model depends on the set of boundary conditions used. Calibrating all of the hysteresis MvG parameters is needed to verify if the hysteresis assumption increases the simulation accuracy, and allow to fit a drying retention curve to the measured pressure-moisture content relationship.

Last but not least, even if the results from this study are quite encouraging for obtaining an adequate set of boundary conditions with a precise posterior parameters estimation, they should be confirmed by applying the investigated approach under a system with precisely known boundary conditions, or with the water flow and solute transport parameters known with precision, to check if findings are similar to experimental results. Afterwards, one could generalize this approach to other lysimeter experiments with complex boundary conditions, similarly to the one studied in this work.

References

Abdou, H. M., & Flury, M. (2004). Simulation of water flow and solute transport in free-drainage lysimeters and field soils with heterogeneous structures. *European Journal of Soil Science*, *55*, 229–241. https://doi.org/10.1046/j.1365-2389.2004.00592.x

Bear, J. (1972). *Dynamics of Fluids in Porous Media*. American Elsevier Publishing Company.

Benettin, P., Nehemy, M. F., Asadollahi, M., Pratt, D., Bensimon, M., McDonnell, J. J., & Rinaldo, A. (2021). Tracing and Closing the Water Balance in a Vegetated Lysimeter. *Water Resources Research*, *57*(4). https://doi.org/10.1029/2020WR029049

Brookman, S. K. E., Chadwick, n., & Retter, A. R. (2005). Lysimeter experiments to determine the ability of soil to reduce concentrations of BOD, available P and inorganic N in dirty water. *Environmental Technology*, *26*(11), 1189–1204. https://doi.org/10.1080/09593332608618593

Carsel, R. F., & Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, *24*(5), 755–769. https://doi.org/10.1029/WR024i005p00755

Chen, S., Mao, X., Barry, D. A., & Yang, J. (2019). Model of crop growth, water flow, and solute transport in layered soil. *Agricultural Water Management*, *221*, 160–174. https://doi.org/10.1016/j.agwat.2019.04.031

Corwin, D. L., Hopmans, J., & de Rooij, G. H. (2006). From Field- to Landscape-Scale Vadose Zone Processes: Scale Issues, Modeling, and Monitoring. *Vadose Zone Journal*, *5*(1), 129–139. https://doi.org/10.2136/vzj2006.0004

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, *28*(4), 1015–1031. https://doi.org/10.1029/91WR02985

Filipović, V., Kodešová, R., & Petošić, D. (2013). Experimental and mathematical modeling of water regime and nitrate dynamics on zero tension plate lysimeters in soil influenced by high groundwater table. *Nutrient Cycling in Agroecosystems*, *95*. https://doi.org/10.1007/s10705-012-9546-5

Friedel, M. J. (2005). Coupled inverse modeling of vadose zone water, heat, and solute transport: Calibration constraints, parameter nonuniqueness, and predictive uncertainty. *Journal of Hydrology*, *312*(1), 148–175. https://doi.org/10.1016/j.jhydrol.2005.02.013

Groh, J., Stumpp, C., Lücke, A., Pütz, T., Vanderborght, J., & Vereecken, H. (2018). Inverse Estimation of Soil Hydraulic and Transport Parameters of Layered Soils from Water Stable Isotope and Lysimeter Data. *Vadose Zone Journal*, *17*. https://doi.org/10.2136/vzj2017.09.0168

J. Šimunek, M. Th. van Genuchten, & M. Šejna. (2012). HYDRUS: Model Use, Calibration, and Validation. *Transactions of the ASABE*, *55*(4), 1263–1276. https://doi.org/10.13031/2013.42239

Kool, J. B., Parker, J. C., & van Genuchten, M. T. (1987). Parameter estimation for unsaturated flow and transport models — A review. *Journal of Hydrology*, *91*(3), 255–293. https://doi.org/10.1016/0022-1694(87)90207-1

Li, N., Yue, X. Y., Wang, W. K., & Ma, Z. T. (2021). Inverse Estimation of Spatiotemporal Flux Boundary Conditions in Unsaturated Water Flow Modeling. *Water Resources Research*, *57*(1). https://doi.org/10.1029/2020WR028030

Liu, C., Zhang, X., & Zhang, Y. (2002). Determination of daily evaporation and evapotranspiration of winter wheat and maize by large-scale weighing lysimeter and micro-lysimeter. *Agricultural and Forest Meteorology*, *111*(2), 109–120. https://doi.org/10.1016/S0168-1923(02)00015-1

Mertens, J., Madsen, H., Kristensen, M., Jacques, D., & Feyen, J. (2003). Sensitivity of soil parameters in unsaturated zone modeling and the relation between effective, laboratory and in-situ estimates, 818.

Mirbagheri, S. A. (2004). Modeling contaminant transport in soil column and ground water pollution control. *International Journal of Environmental Science & Technology*, *1*(2), 141–150. https://doi.org/10.1007/BF03325827

Moreira, P. H. S., Genuchten, M. T. v., Orlande, H. R. B., & Cotta, R. M. (2016). Bayesian estimation of the hydraulic and solute transport properties of a small-scale unsaturated soil column. *Journal of Hydrology and Hydromechanics*, *64*(1), 30–44. https://doi.org/10.1515/johh-2016-0002

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Neuman, S. P. (1975). Galerkin approach to saturated-unsaturated flow in porous media. *Finite Elements in Fluids, Viscous Flow and Hydrodynamics Chapter 10*, *1*, 201–217.

Pasteris, G., Werner, D., Kaufmann, K., & Höhener, P. (2002). Vapor Phase Transport and Biodegradation of Volatile Fuel Compounds in the Unsaturated Zone: A Large Scale Lysimeter Experiment. *Environmental Science & Technology*, *36*(1), 30–39. https://doi.org/10.1021/es0100423

Perrochet, P., & Bérod, D. (1993). Stability of the standard Crank–Nicolson–Galerkin scheme applied to the diffusion-convection equation: Some new insights. *Water Ressources Research*, *29*(9), 3291–3298.

Pütz, T., Fank, J., & Flury, M. (2018). Lysimeters in Vadose Zone Research. *Vadose Zone Journal*, *17*(1), 180035. https://doi.org/10.2136/vzj2018.02.0035

Rahnamay Naeini, M., Analui, B., Gupta, H. V., Duan, Q., & Sorooshian, S. (2019). Three decades of the Shuffled Complex Evolution (SCE-UA) optimization algorithm: Review and applications. *Scientia Iranica*, *26*(4), 2015–2031. https://doi.org/10.24200/sci.2019.21500

Rawls, W. J., Brakensiek, D. L., & Miller, N. (1983). Green-ampt Infiltration Parameters from Soils Data. *Journal of Hydraulic Engineering*, *109*(1), 62–70. https://doi.org/10.1061/(ASCE)0733-9429(1983)109:1(62)

Ritter, A., Hupet, F., Muñoz-Carpena, R., Lambot, S., & Vanclooster, M. (2003). Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods. *Agricultural Water Management*, *59*(2), 77–96. https://doi.org/10.1016/S0378-3774(02)00160-9

Schelle, H., Durner, W., Iden, S. C., & Fank, J. (2013). Simultaneous Estimation of Soil Hydraulic and Root Distribution Parameters from Lysimeter Data by Inverse Modeling. *Procedia Environmental Sciences*, *19*, 564–573. https://doi.org/10.1016/j.proenv.2013.06.064

Schneider, J., Groh, J., Pütz, T., Helmig, R., Rothfuss, Y., Vereecken, H., & Vanderborght, J. (2021). Prediction of soil evaporation measured with weighable lysimeters using the FAO Penman–Monteith method in combination with Richards' equation. *Vadose Zone Journal*, *20*(1), e20102. https://doi.org/10.1002/vzj2.20102

Selker, J., & Or, D. (2021). *Soil Hydrology and Biophysics: Solute Transport in Soils and Vadose Materials*. Oregon State University.

Simunek, J., Jirka, Saito, H., Sakai, M., & Van Genuchten, M. (2008). *The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media*.

Sprenger, M., Volkmann, T. H. M., Blume, T., & Weiler, M. (2015). Estimating flow and transport parameters in the unsaturated zone with pore water stable isotopes. *Hydrology and Earth System Sciences*, *19*(6), 2617–2635. https://doi.org/10.5194/hess-19-2617-2015

Stumpp, C., Stichler, W., & Maloszewski, P. (2009). Application of the environmental isotope 18O to study water flow in unsaturated soils planted with different crops: Case study of a weighable lysimeter from the research field in Neuherberg, Germany. *Journal of Hydrology*, *368*(1), 68–78. https://doi.org/10.1016/j.jhydrol.2009.01.027

Tyagi, N. K., Sharma, D. K., & Luthra, S. K. (2000). Determination of evapotranspiration and crop coefficients of rice and sunflower with lysimeter. *Agricultural Water Management*, *45*(1), 41–54. https://doi.org/10.1016/S0378-3774(99)00071-2

Vereecken, H., Kasteel, R., Vanderborght, J., & Harter, T. (2007). Upscaling Hydraulic Properties and Soil Water Flow Processes in Heterogeneous Soils A Review. *Vadose Zone J.*, *6*, 1–28. https://doi.org/10.2136/vzj2006.0055

Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, *75*, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A., & Vesselinov, V. V. (2008). Inverse Modeling of Subsurface Flow and Transport Properties: A Review with New Developments. *Vadose Zone Journal*, *7*(2), 843–864. https://doi.org/10.2136/vzj2007.0078

Wang, J., Bogena, H. R., Vereecken, H., & Brüggemann, N. (2018). Characterizing Redox Potential Effects on Greenhouse Gas Emissions Induced by Water-Level Changes. *Vadose Zone Journal*, *17*(1), 170152. https://doi.org/10.2136/vzj2017.08.0152

Younes, A., Mara, T. A., Voltz, M., Guellouz, L., Baalousha, H. M., & Fahs, M. (2018). A new efficient Bayesian parameter inference strategy: Application to flow and pesticide transport through unsaturated porous media. [Publisher: Elsevier Ltd]. *Journal of Hydrology (Amsterdam)*, *563*, 887–899. https://doi.org/10.1016/j.jhydrol.2018.06.043

Younes, A., Mara, T., Fahs, M., Olivier, G., & Ackerer, P. (2017). Hydraulic and transport parameter assessment using column infiltration experiments. *Hydrology and Earth System Sciences*, *21*, 2263–2275. https://doi.org/10.5194/hess-21-2263-2017

Younes, A., Zaouali, J., S., K., Lehmann, F., & Fahs, M. (2019). Bayesian Simultaneous Estimation of Unsaturated Flow and Solute Transport Parameters from a Laboratory Infiltration Experiment. *Water*, *11*, 1660. https://doi.org/10.3390/w11081660

Zaradny, H. (1978). Boundary condition in modeling water flow in unsaturated soils. *Soil Science*, *125*(2), 75–82.

Zhang, Z. F., Ward, A. L., & Gee, G. W. (2003). Estimating Soil Hydraulic Parameters of a Field Drainage Experiment Using Inverse Techniques. *Vadose Zone Journal*, *2*(2), 201–211. https://doi.org/10.2113/2.2.201

A Chains convergence

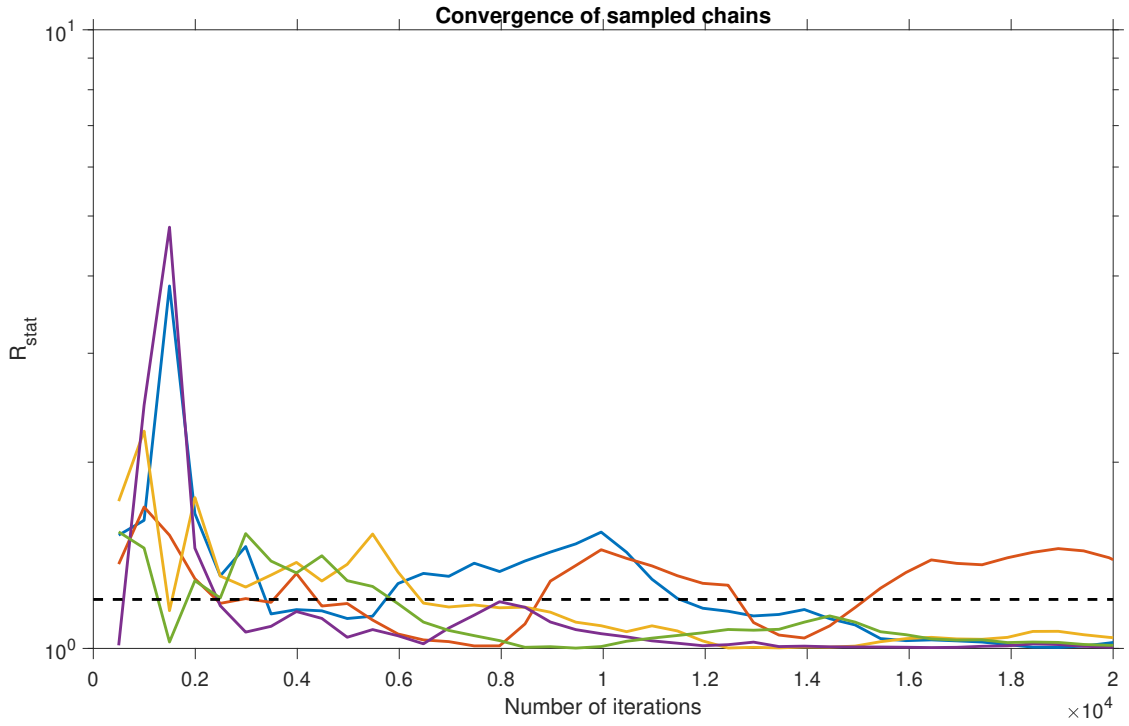


Figure 24: Atm-Seep configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters)

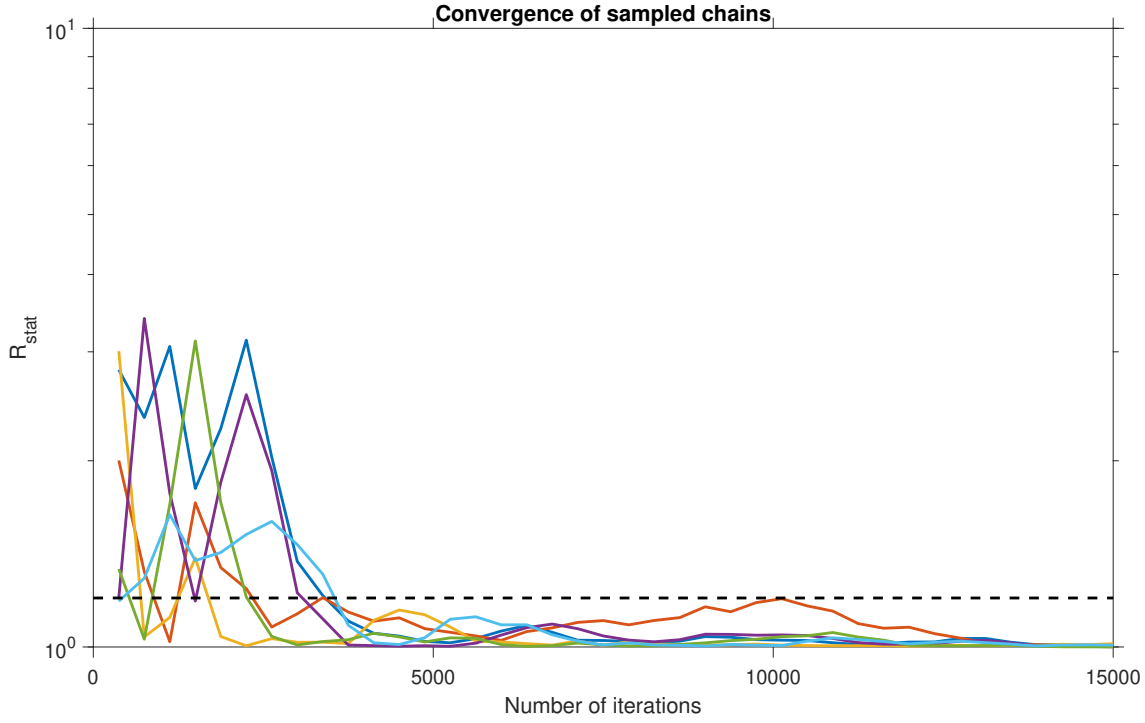


Figure 25: Atm-SeepC configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

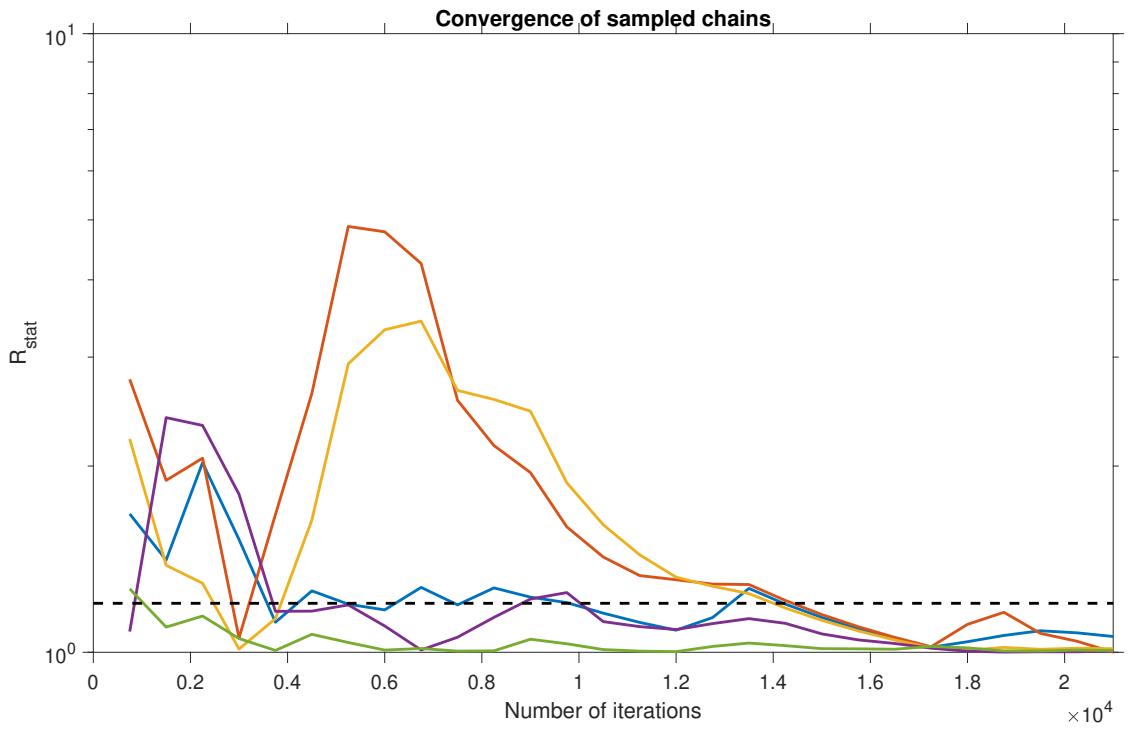


Figure 26: Atm-Flux configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

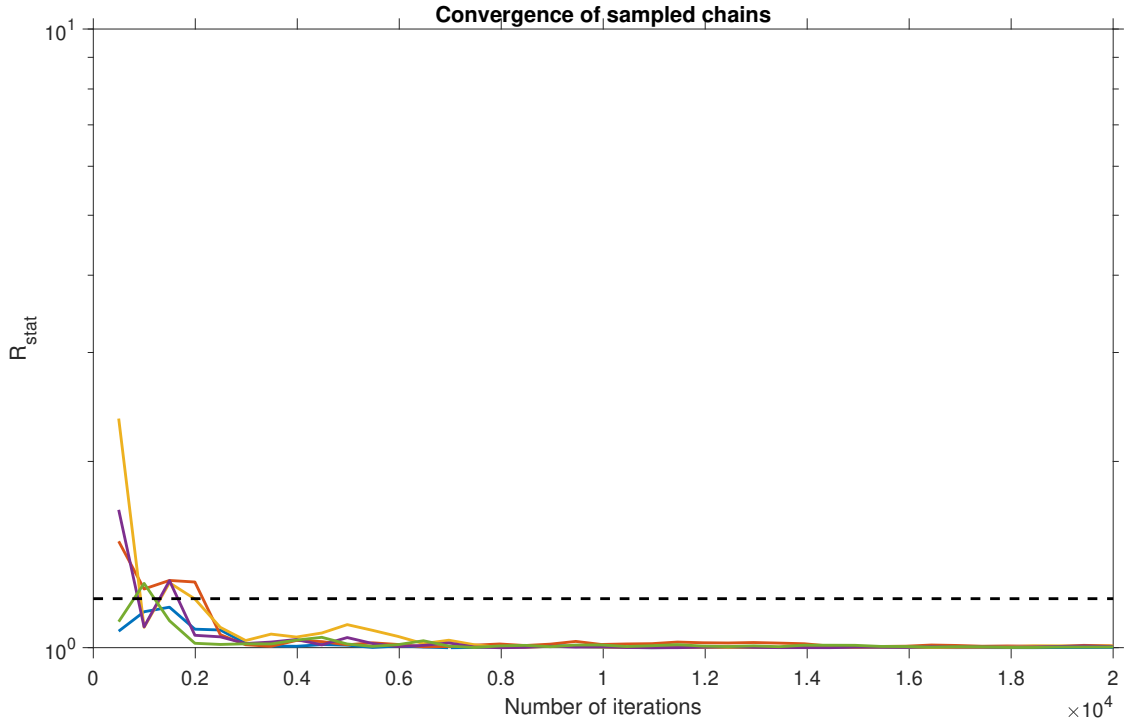


Figure 27: Flux-Flux configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

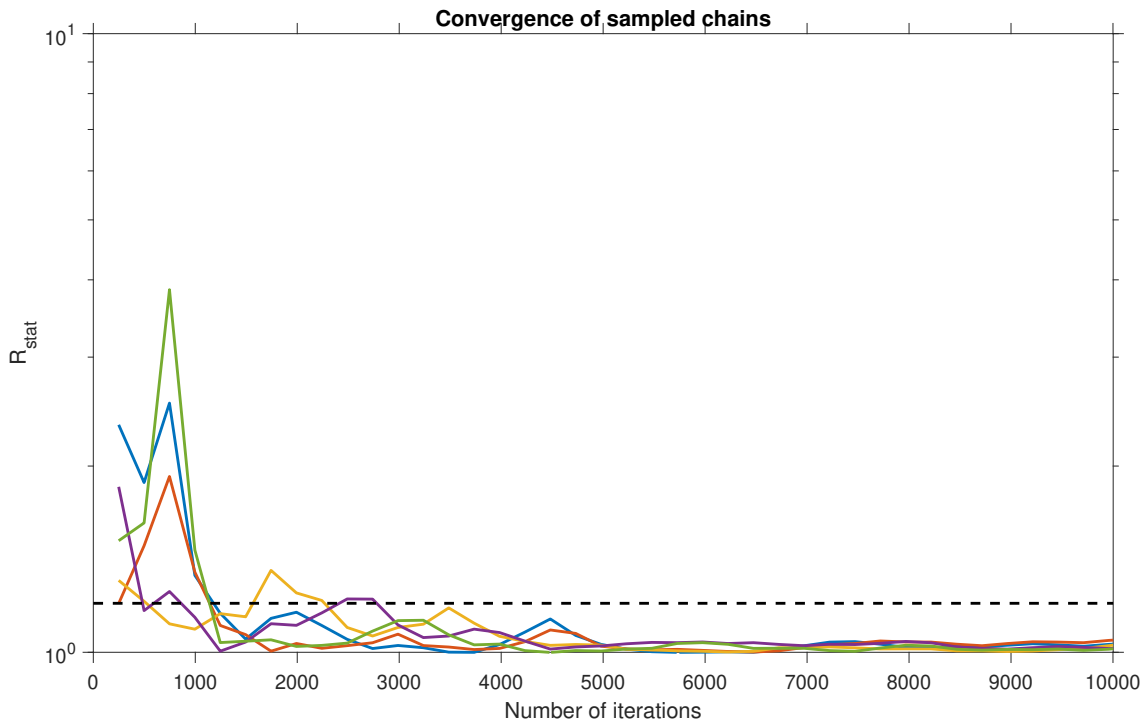


Figure 28: Cropped model configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

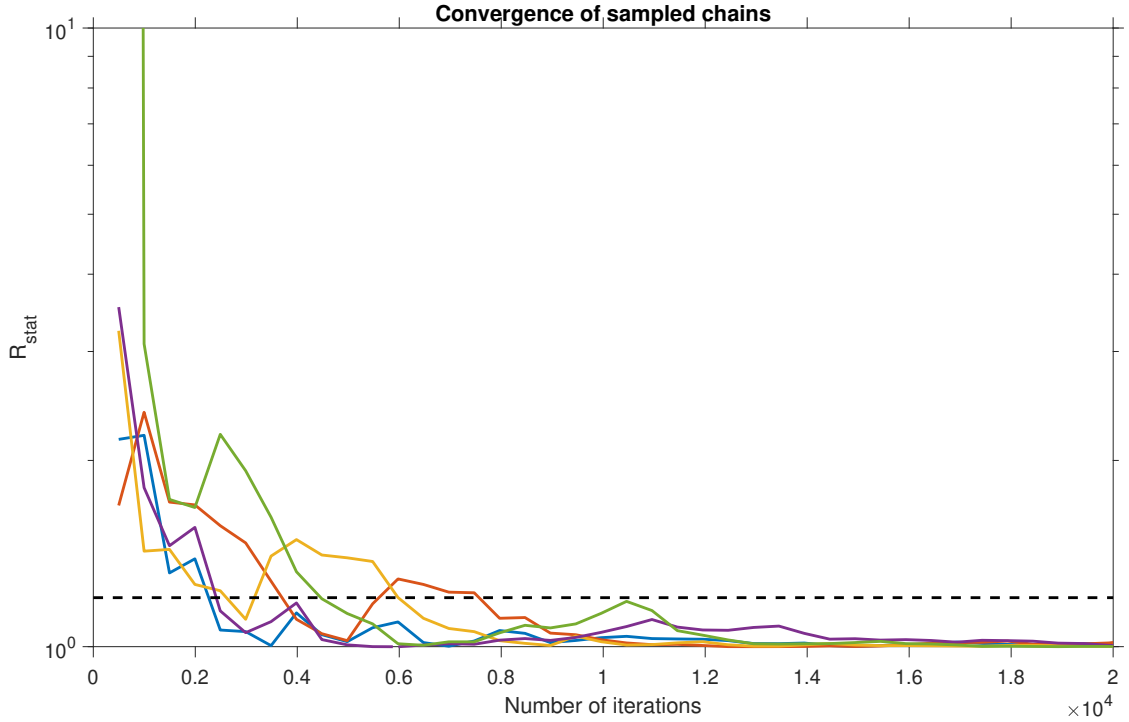


Figure 29: Flux-Pressure configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

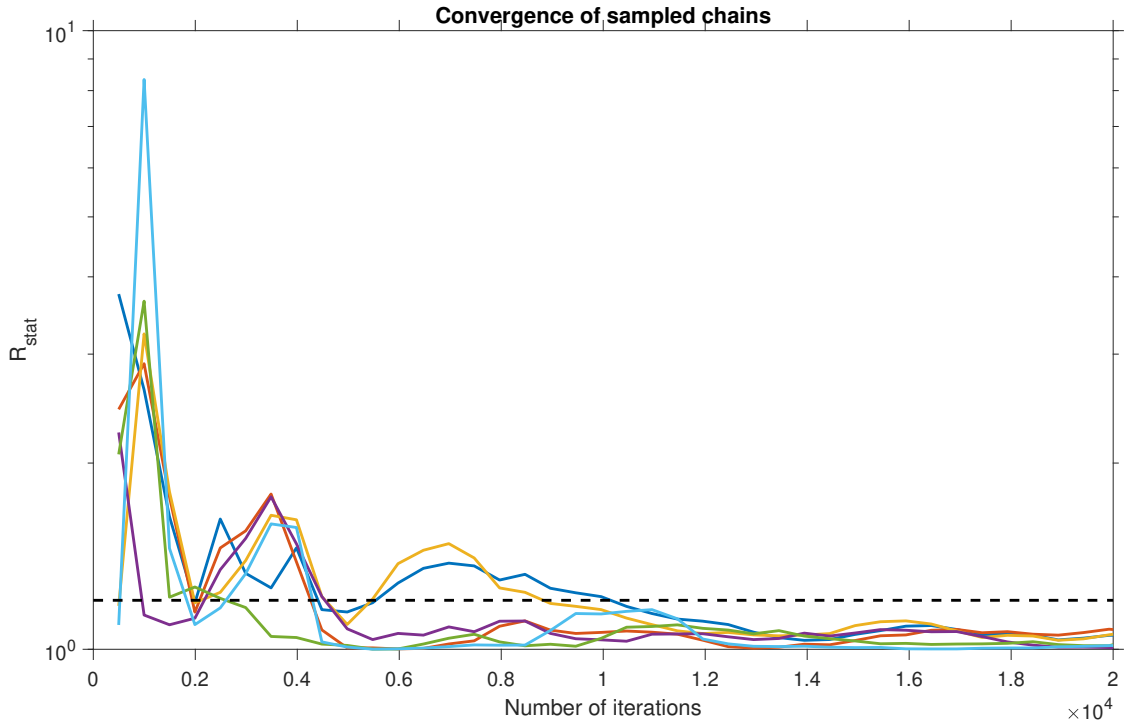


Figure 30: Cropped Model (HYS) configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

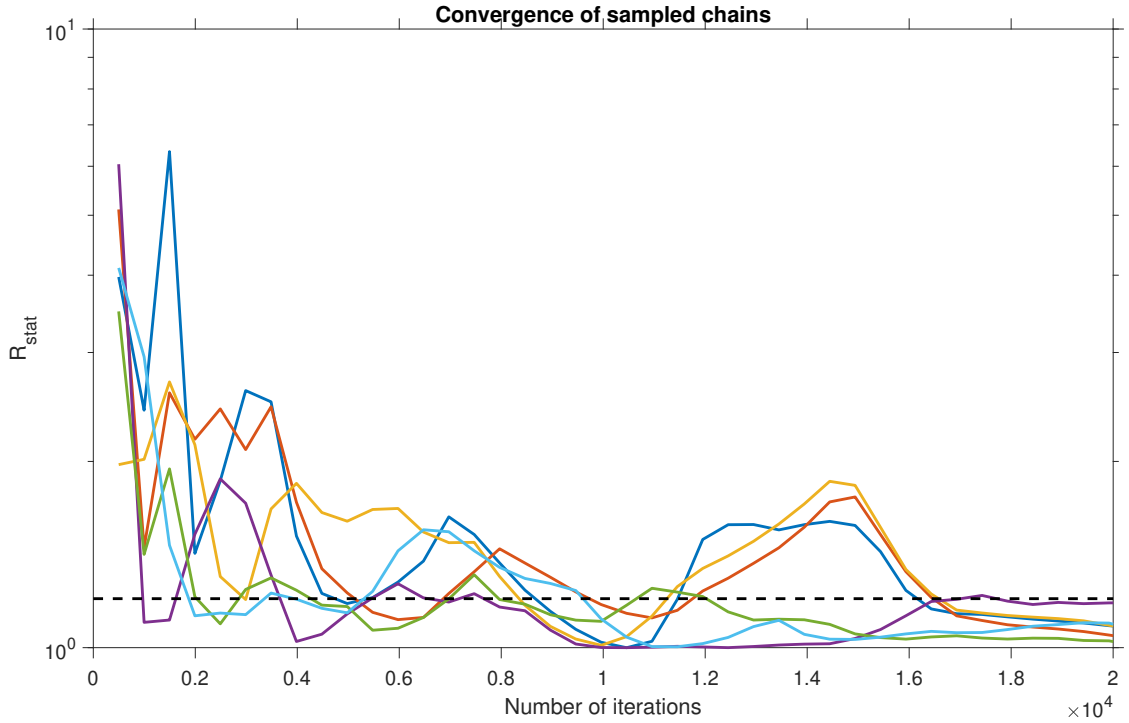


Figure 31: Flux-Pressure (HYS) configuration - Chains convergence (curves should reach R_{stat} or below to achieve convergence for all parameters).

B Moisture content

B.1 First observation node $z = 8\ cm$

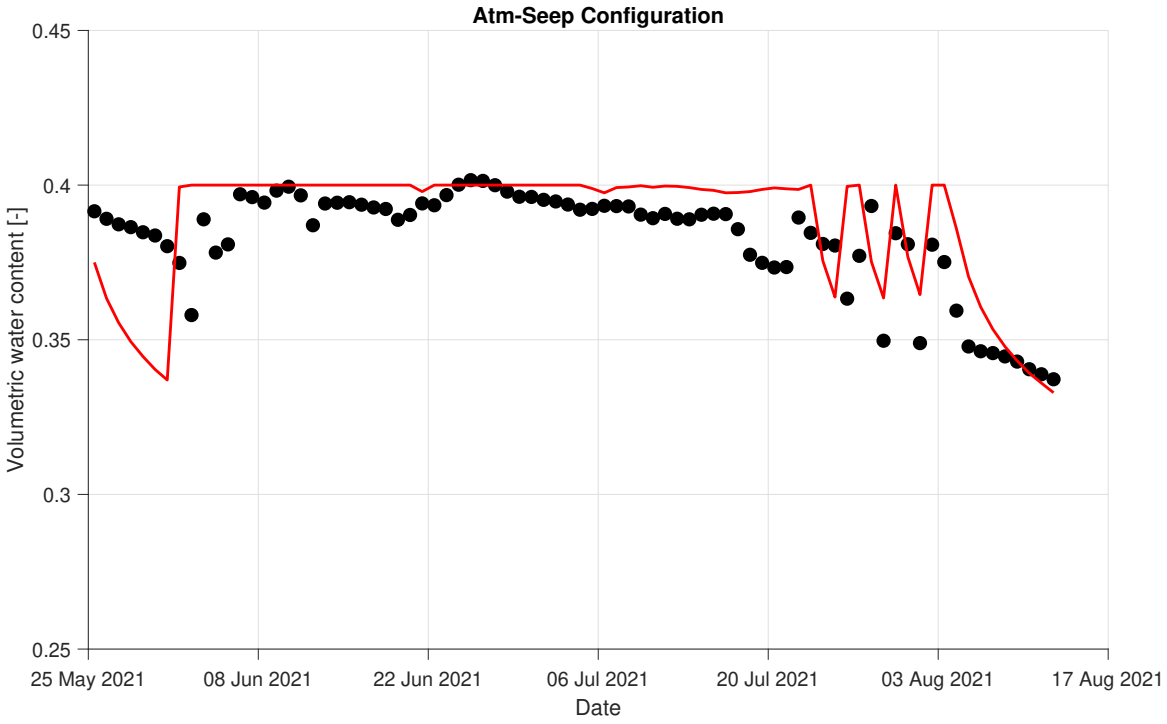


Figure 32: Moisture content at $z = 8\ cm$ versus time (black dots are measured values and red curve is simulated values).

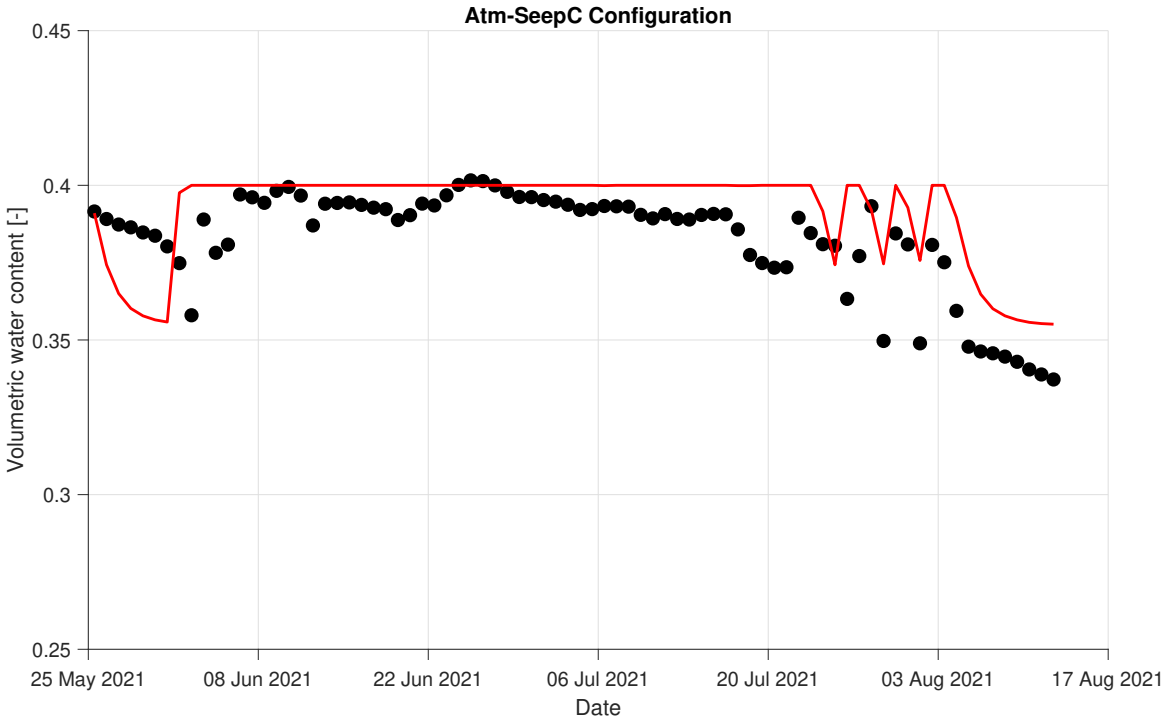


Figure 33: Moisture content at $z = 8\ cm$ versus time (black dots are measured values and red curve is simulated values).

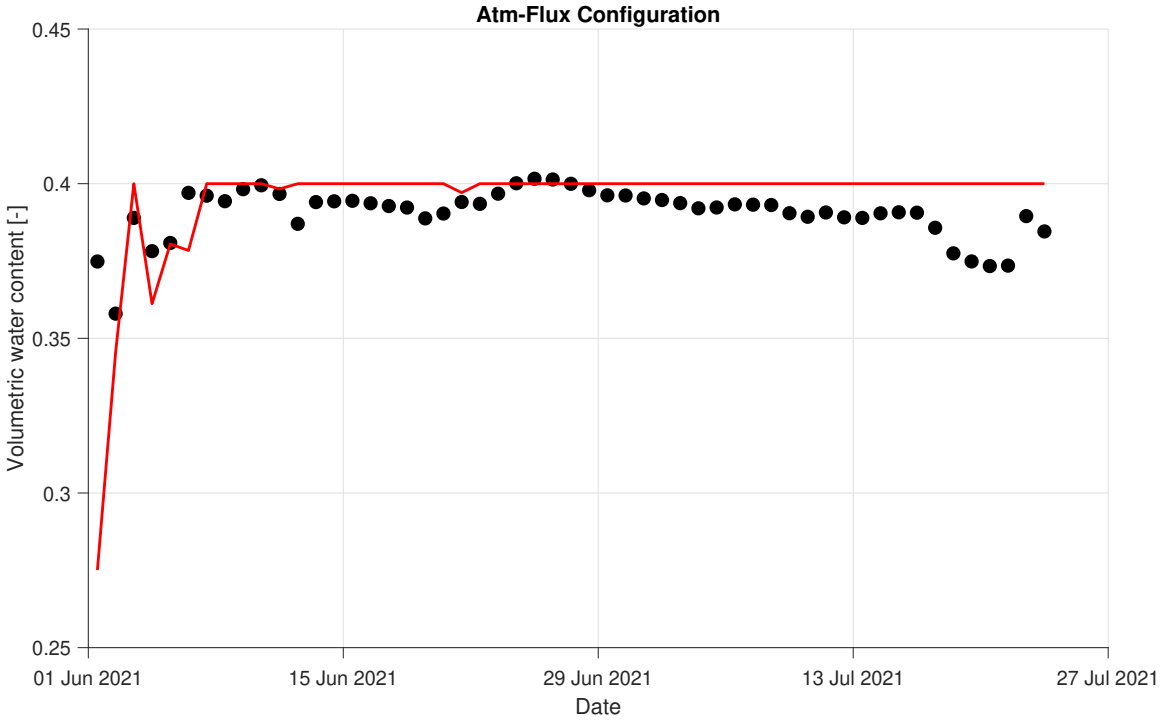


Figure 34: Moisture content at $z = 8\ cm$ versus time (black dots are measured values and red curve is simulated values).

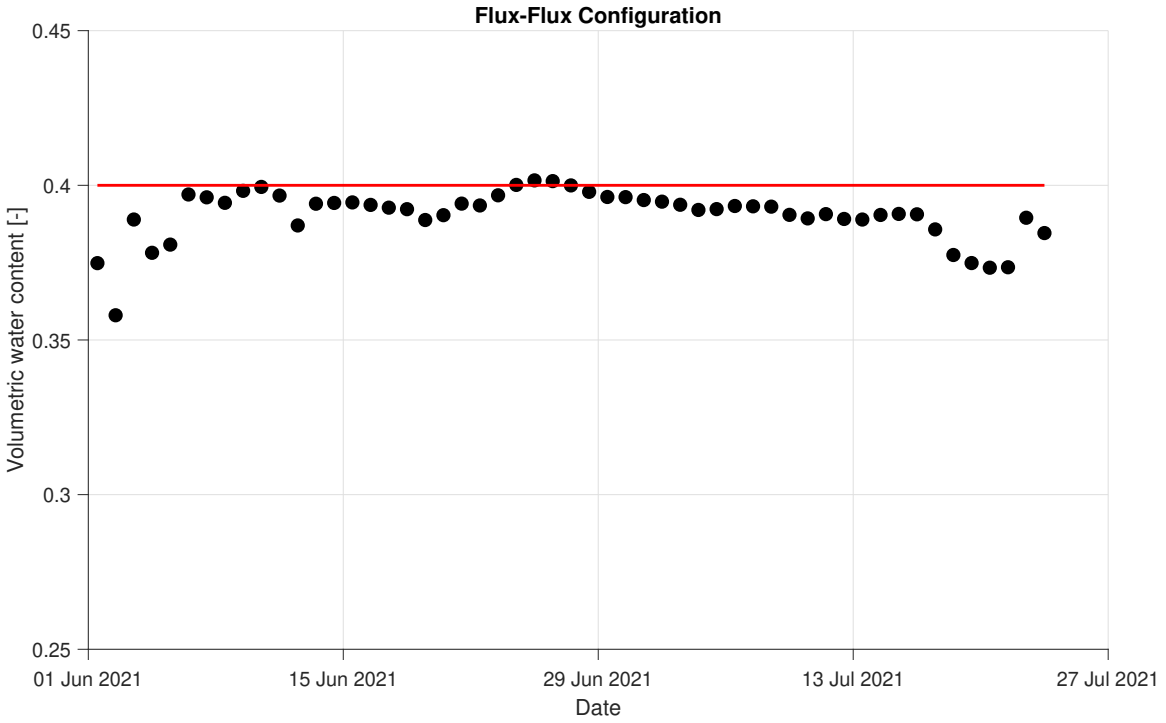


Figure 35: Moisture content at $z = 8$ cm versus time (black dots are measured values and red curve is simulated values).

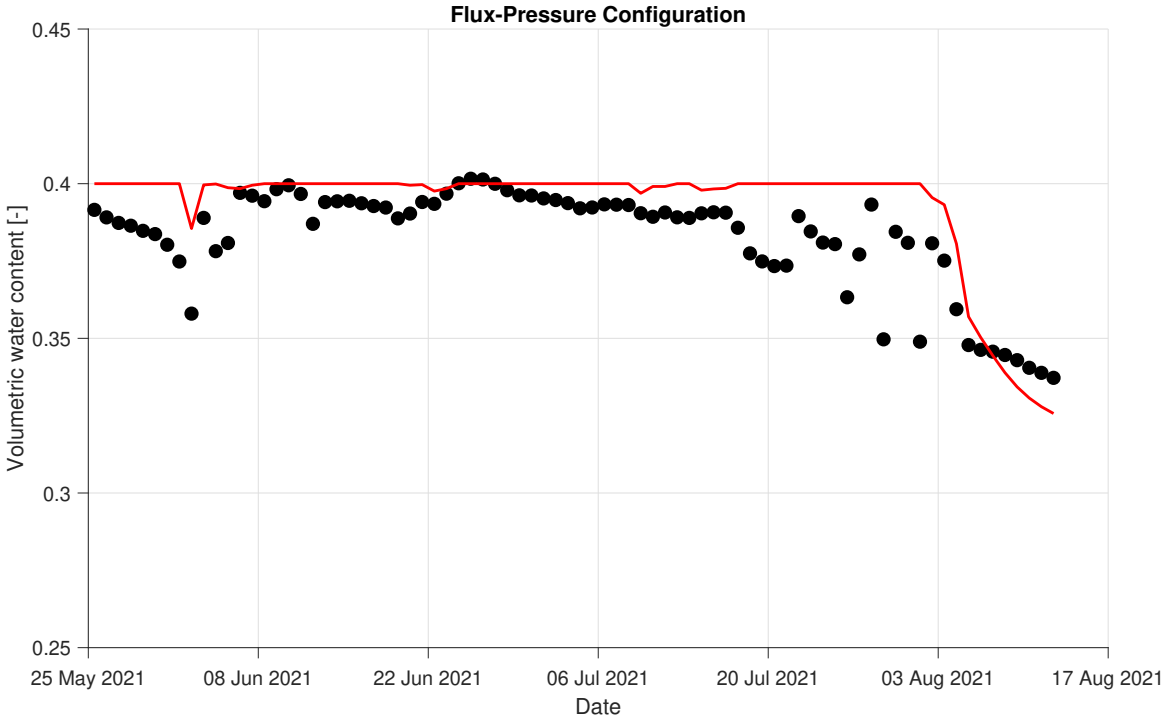


Figure 36: Moisture content at $z = 8$ cm versus time (black dots are measured values and red curve is simulated values).

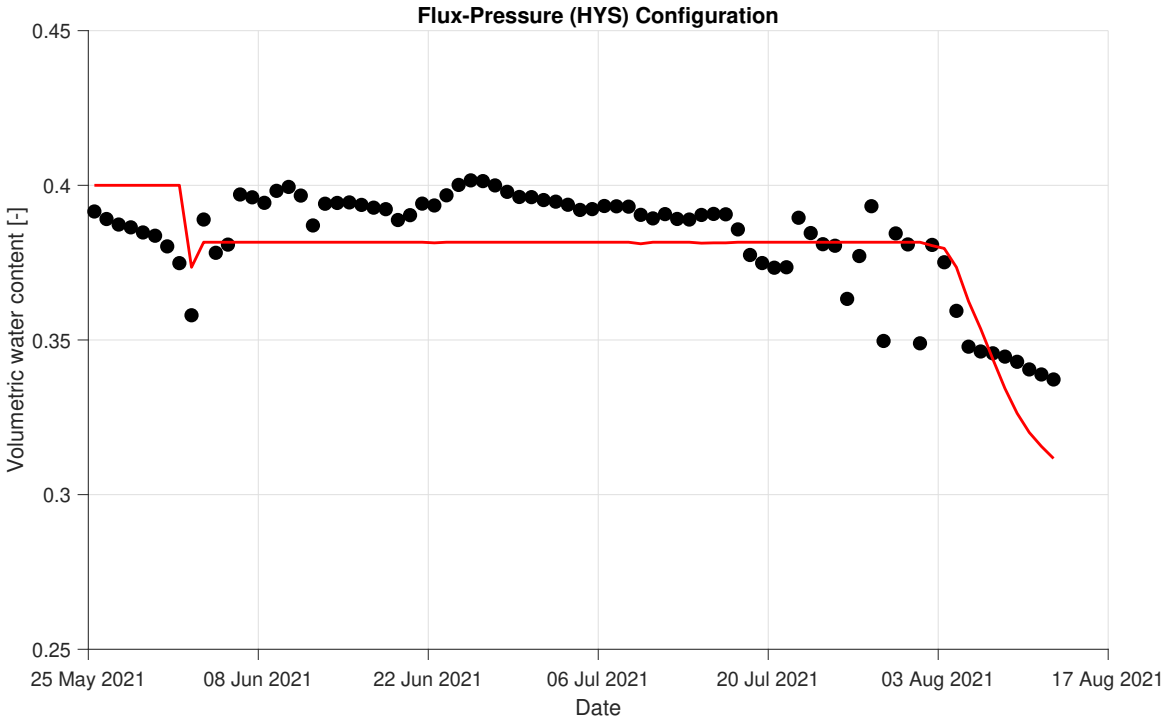


Figure 37: Moisture content at $z = 8\ cm$ versus time (black dots are measured values and red curve is simulated values).

B.2 Second observation node $z = 20\ cm$

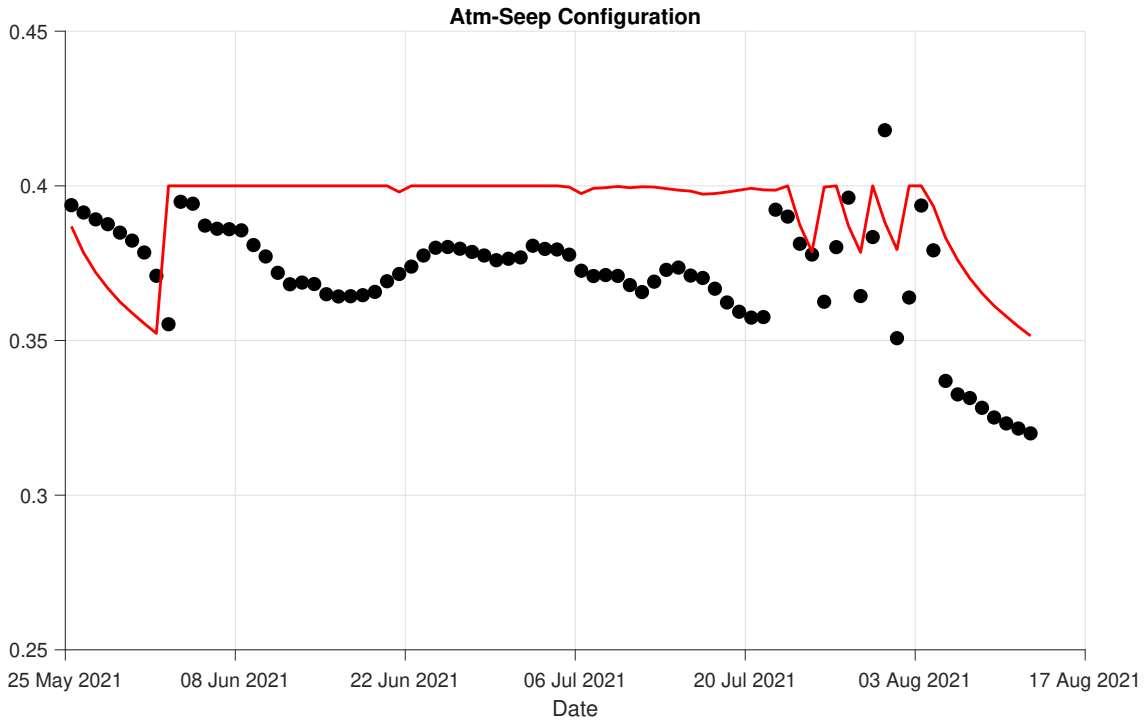


Figure 38: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

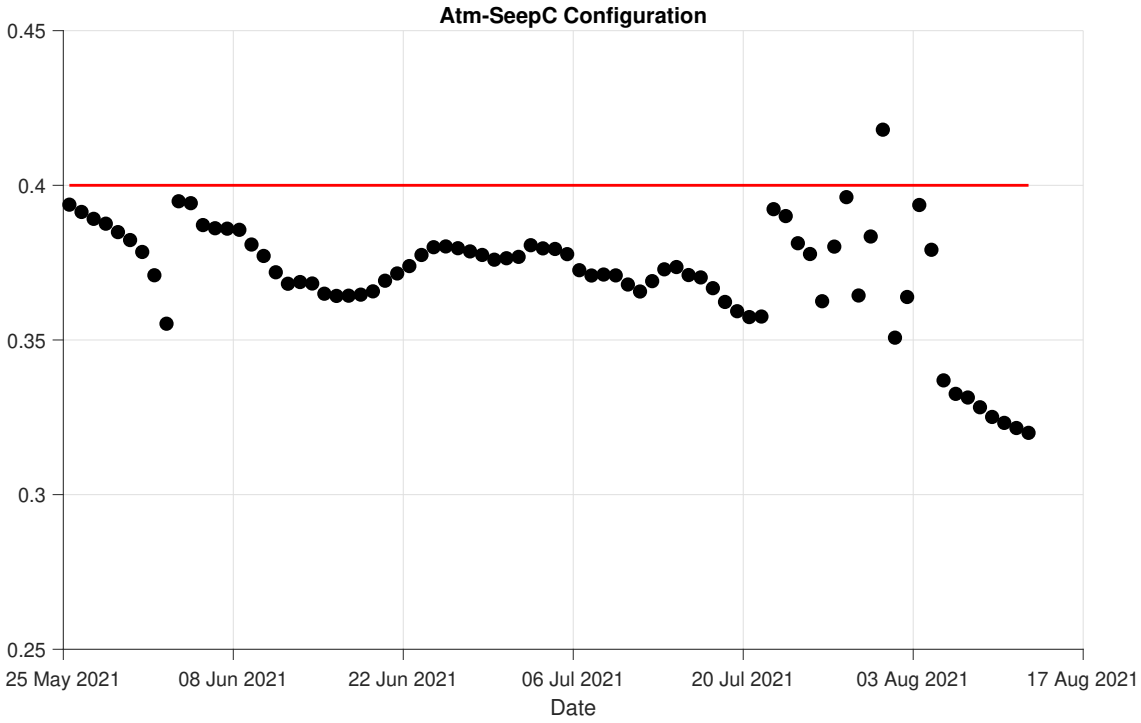


Figure 39: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

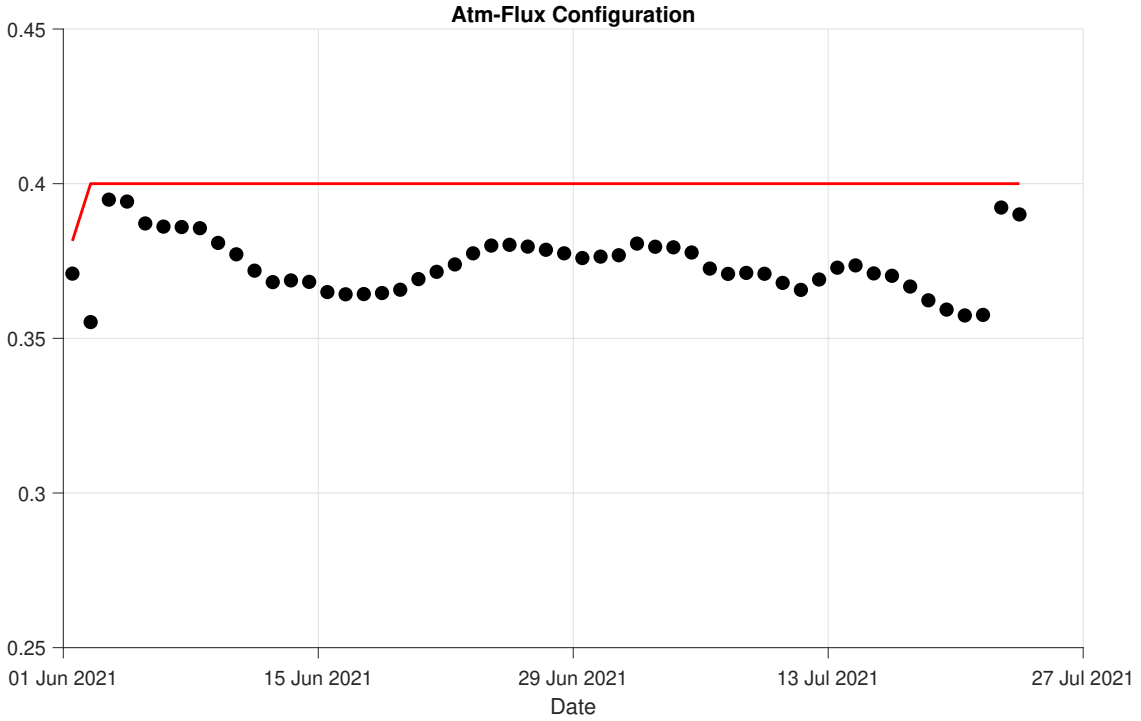


Figure 40: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

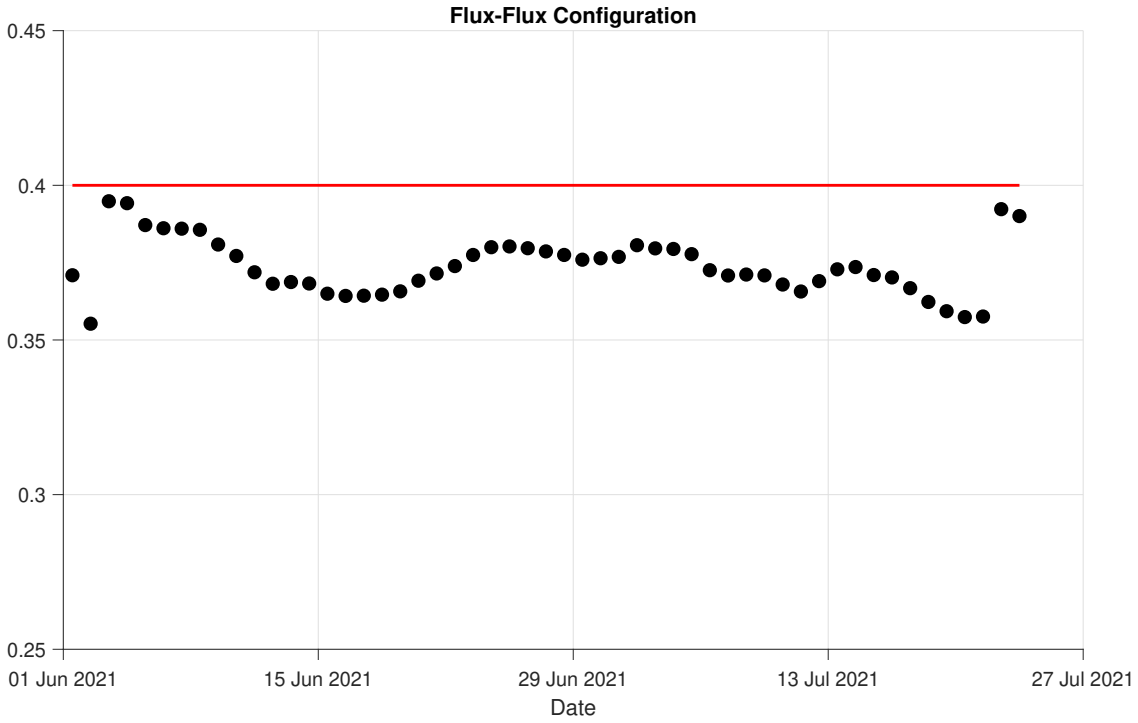


Figure 41: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

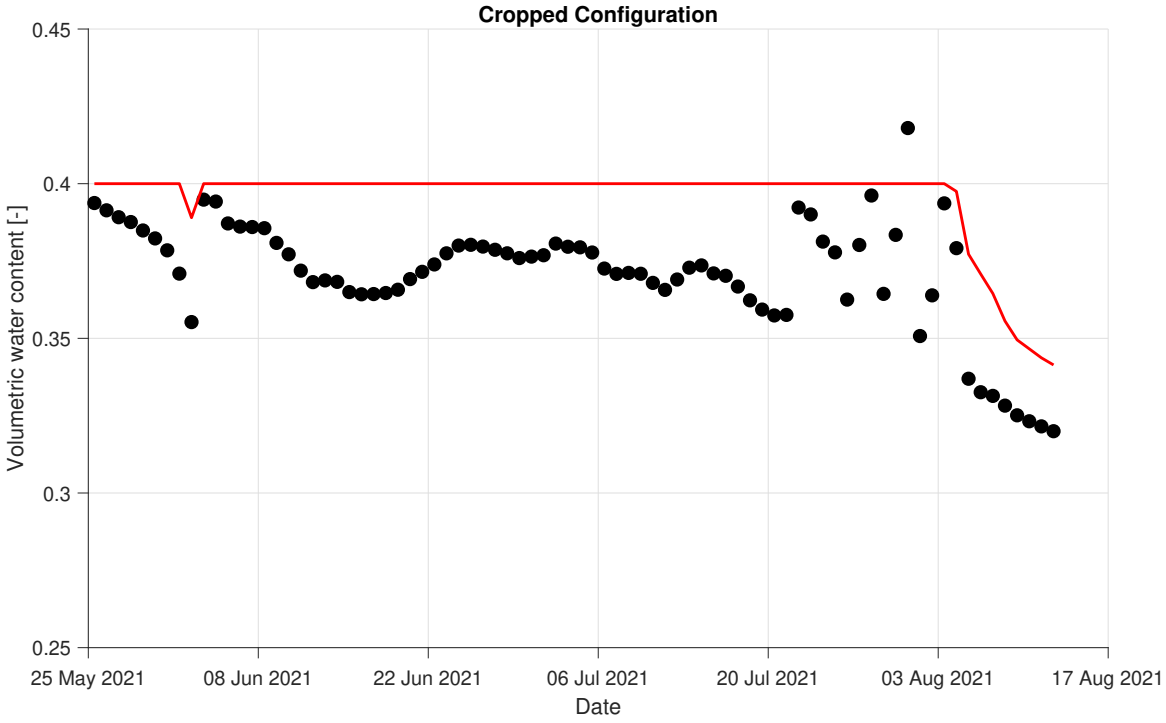


Figure 42: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

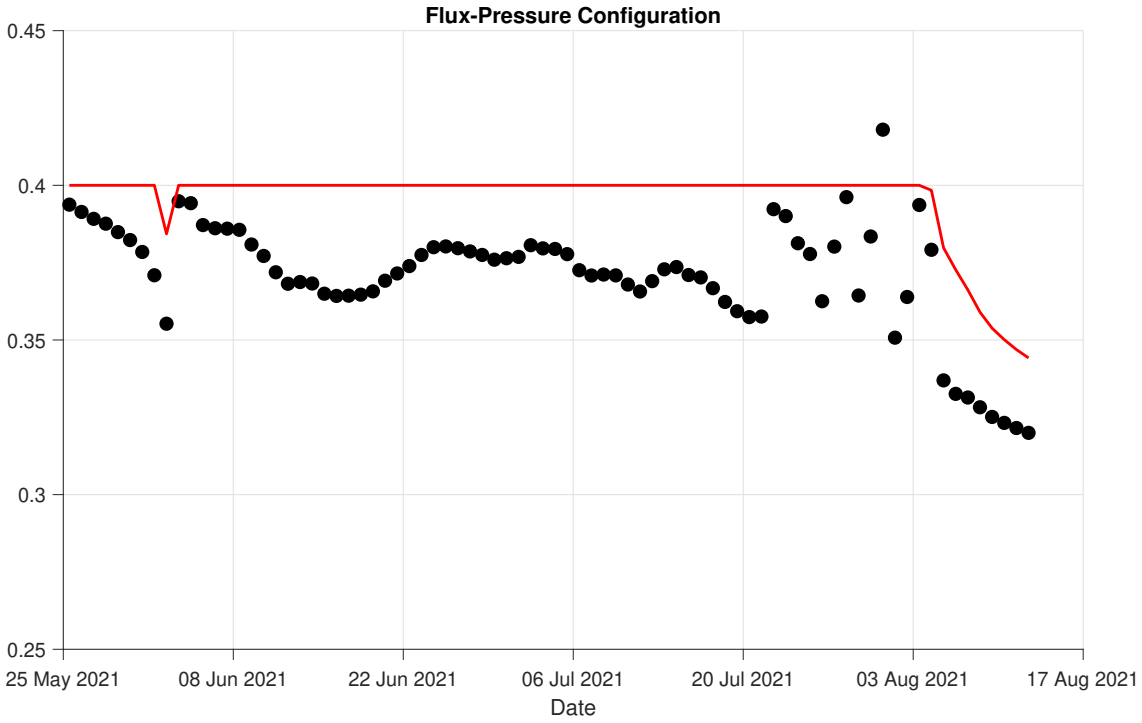


Figure 43: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

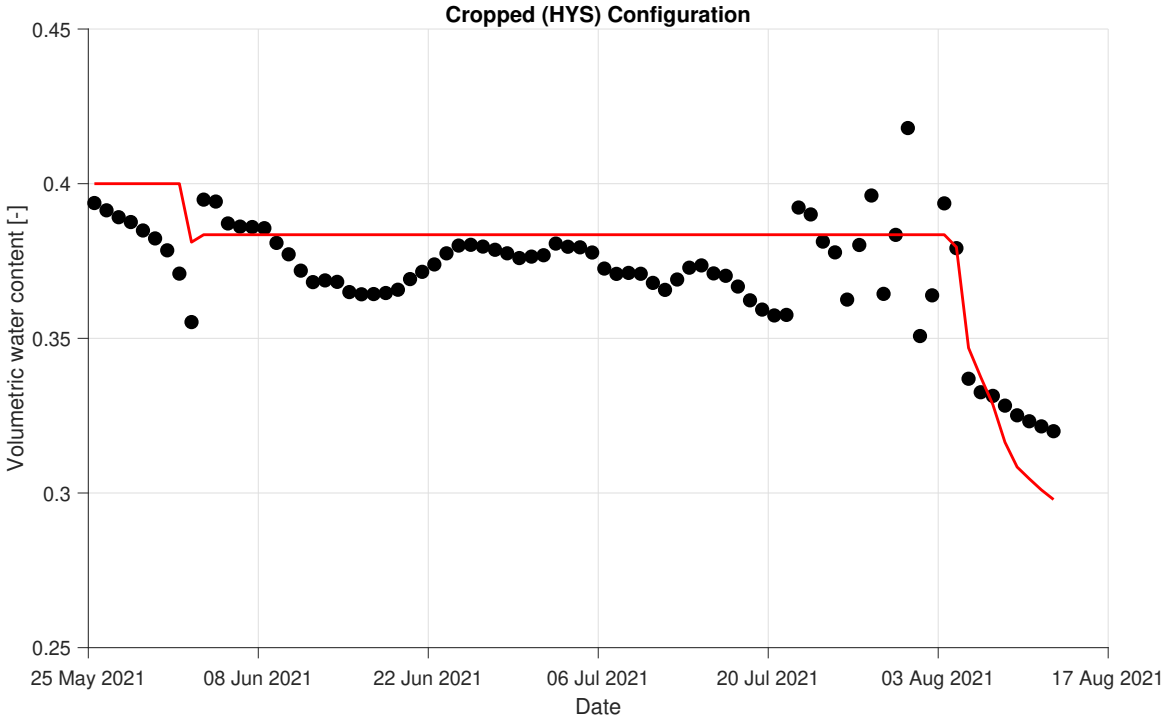


Figure 44: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

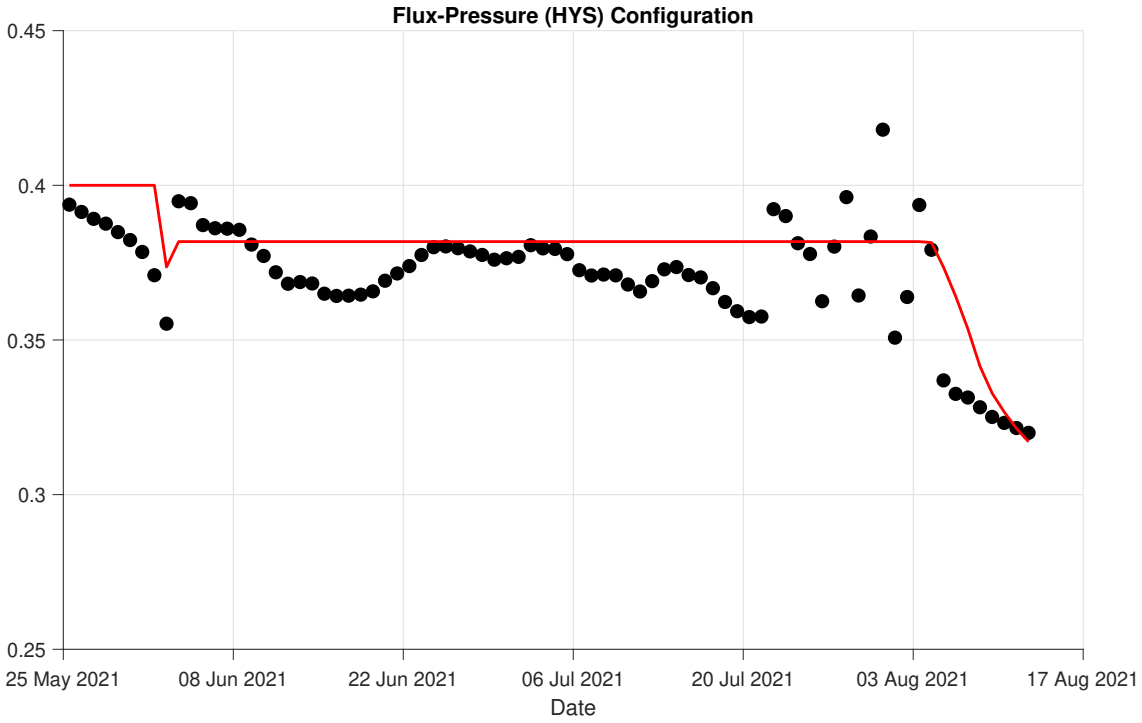


Figure 45: Moisture content at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

B.3 Third observation node $z = 30\ cm$

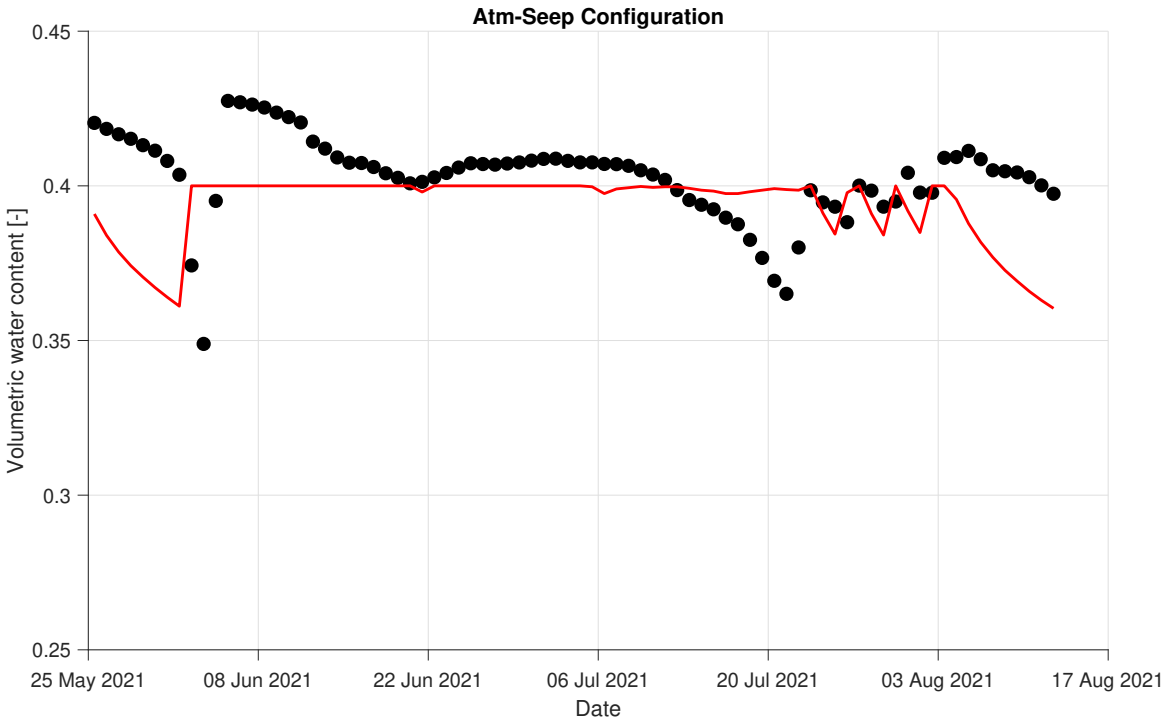


Figure 46: Moisture content at $z = 30\ cm$ versus time (black dots are measured values and red curve is simulated values).

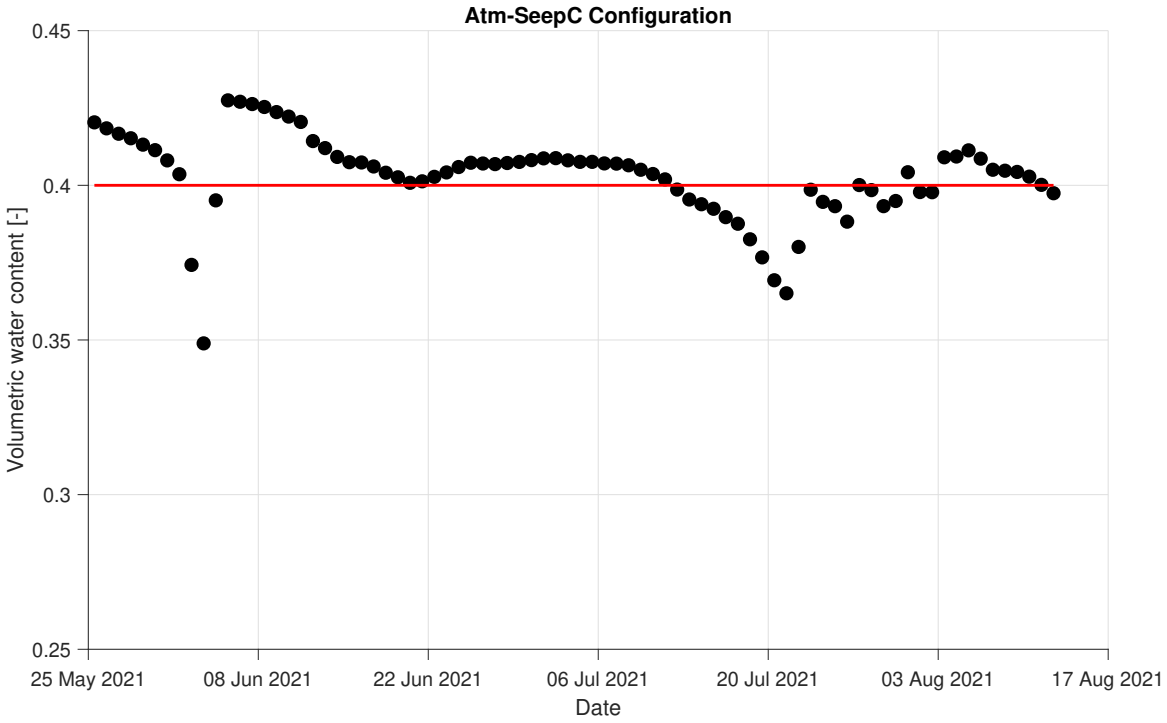


Figure 47: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

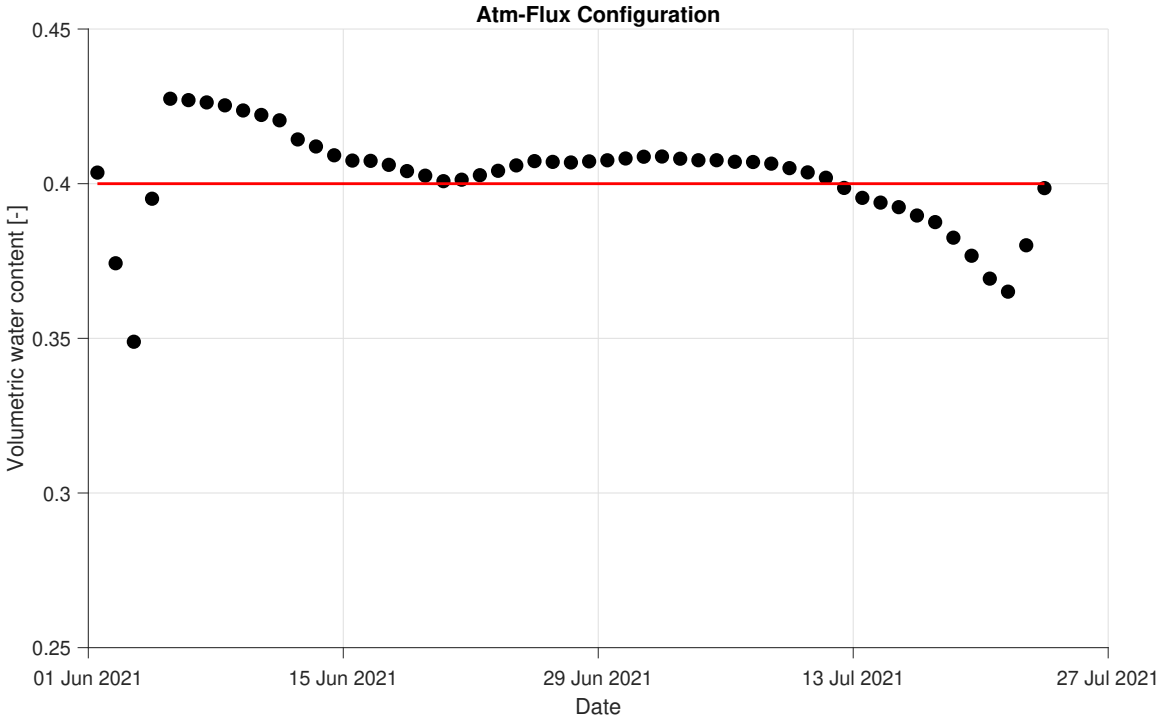


Figure 48: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

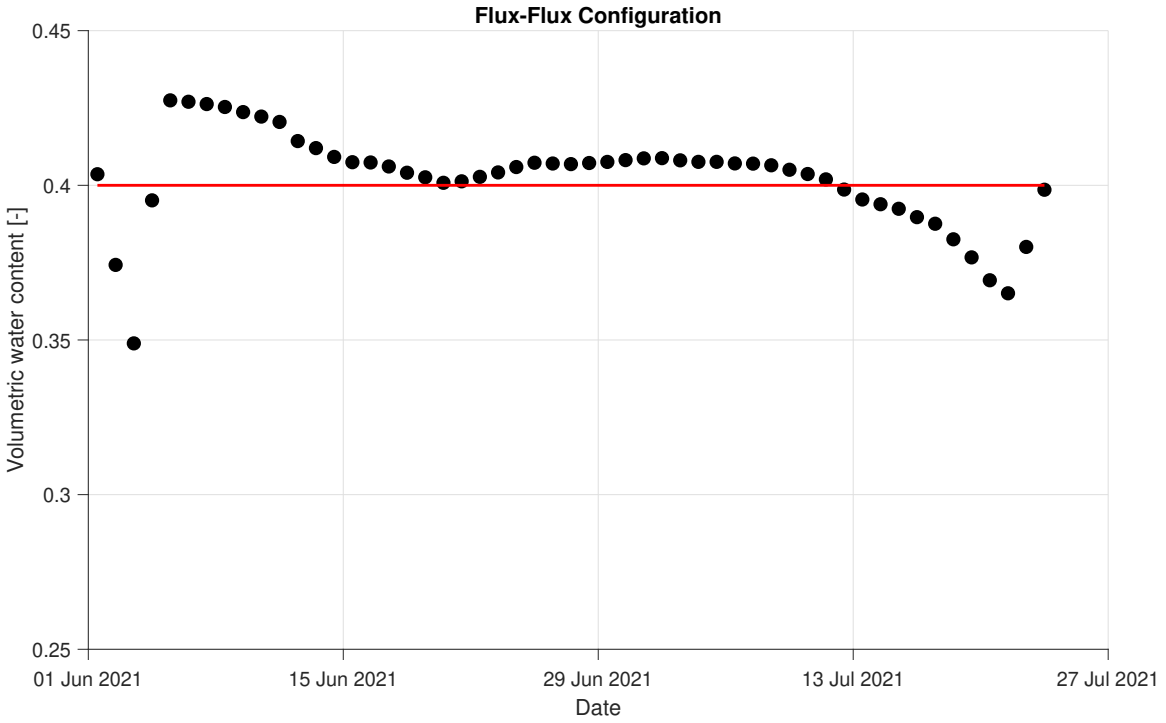


Figure 49: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

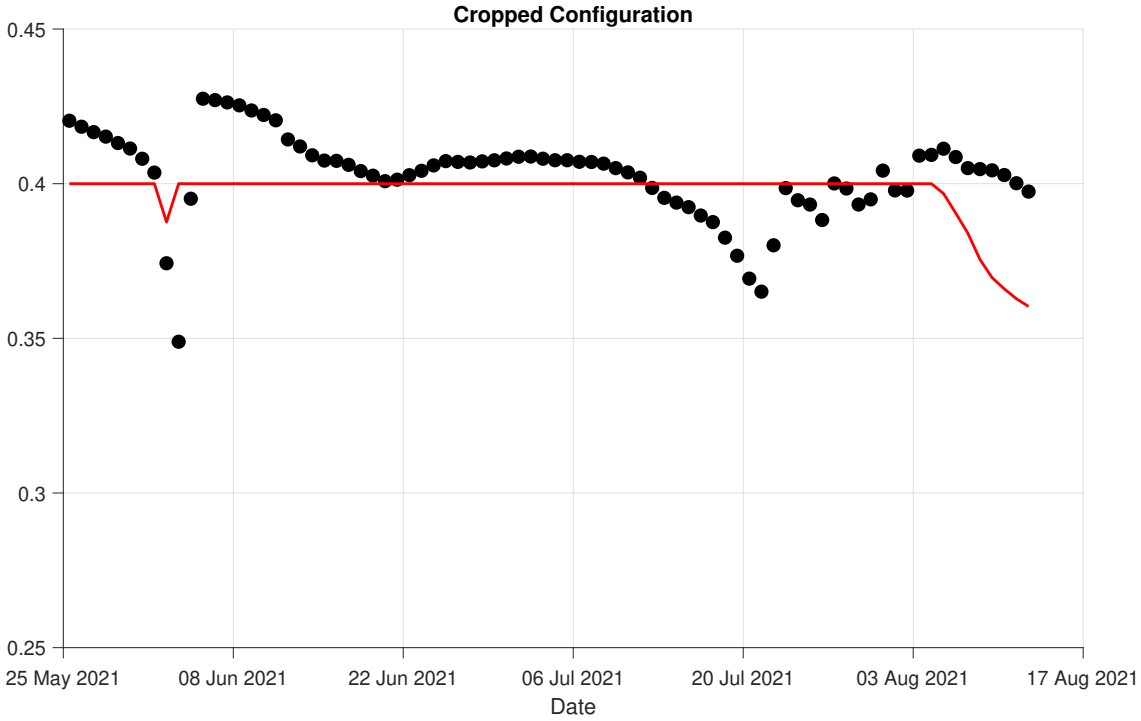


Figure 50: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

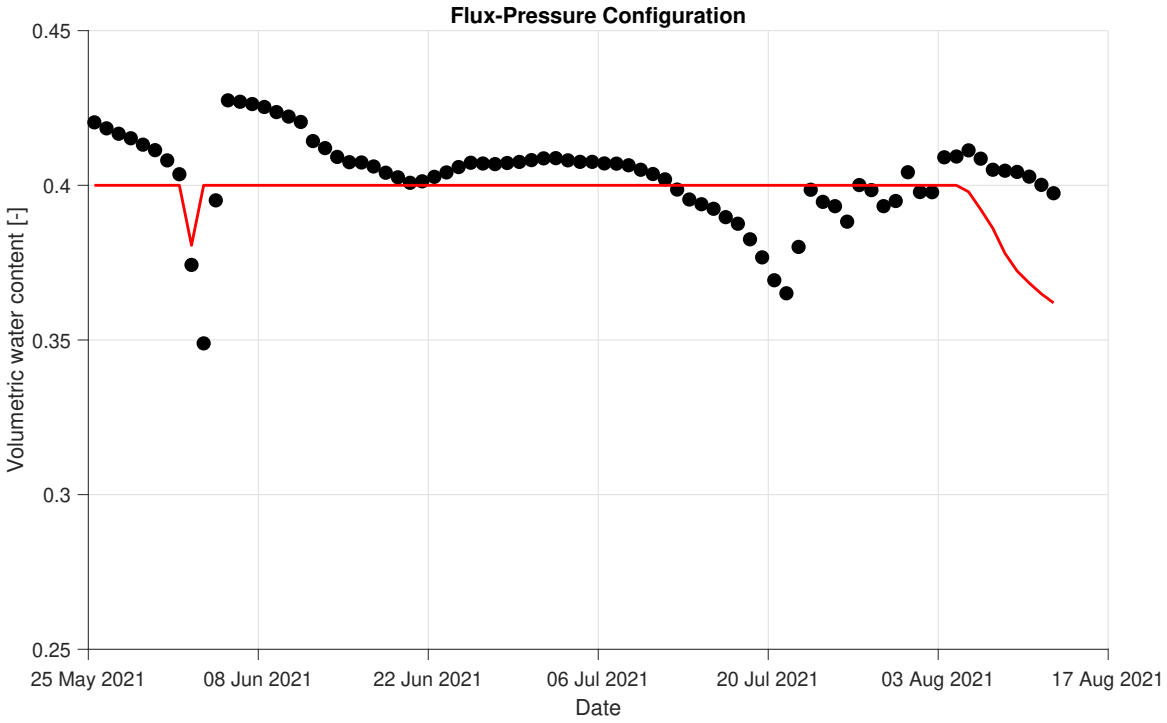


Figure 51: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

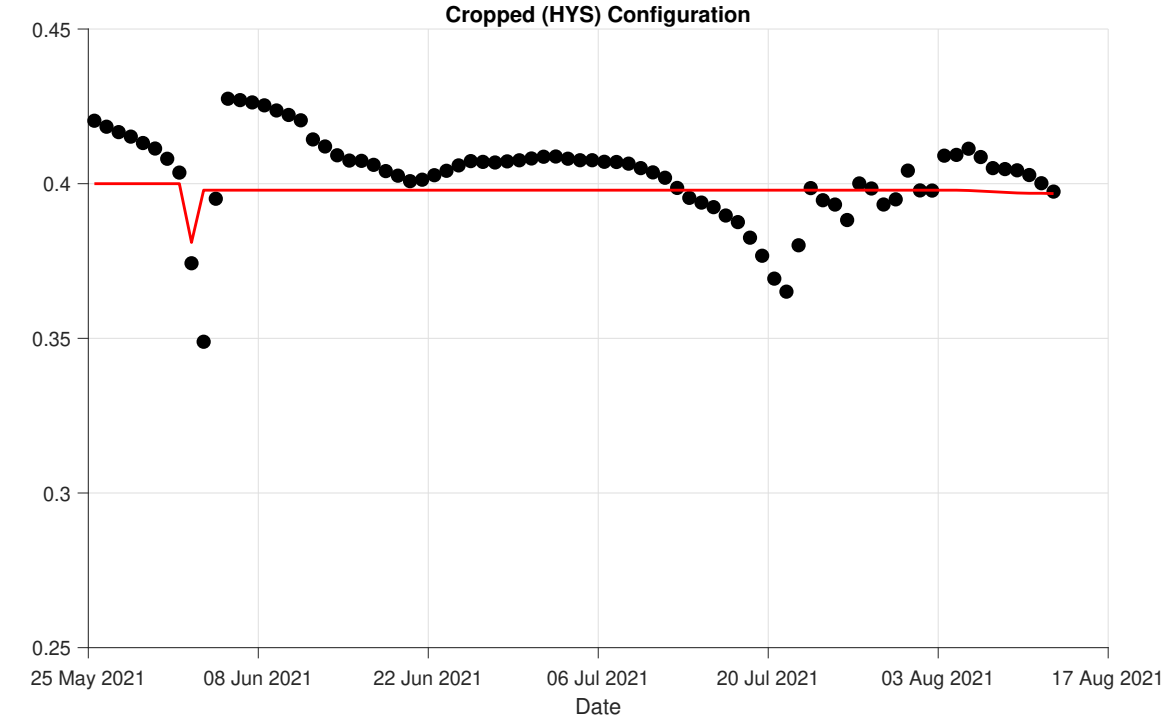


Figure 52: Moisture content at $z = 30$ cm versus time (black dots are measured values and red curve is simulated values).

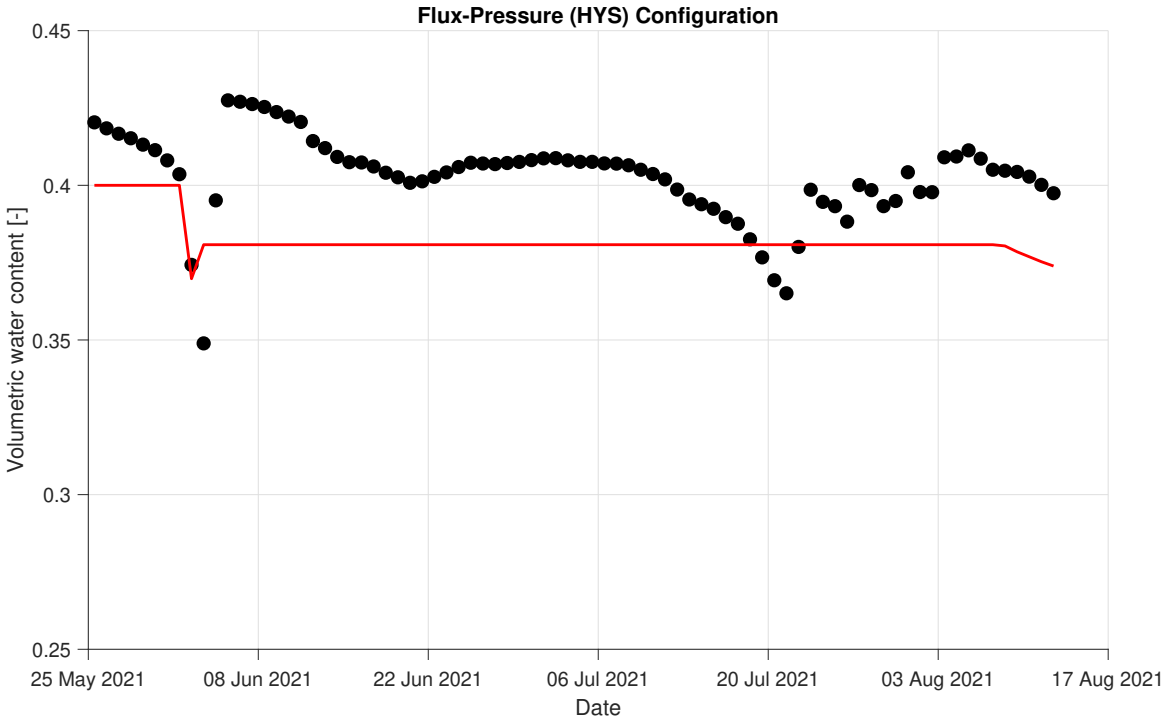


Figure 53: Moisture content at $z = 30\ cm$ versus time (black dots are measured values and red curve is simulated values).

B.4 Fourth observation node $z = 42\ cm$

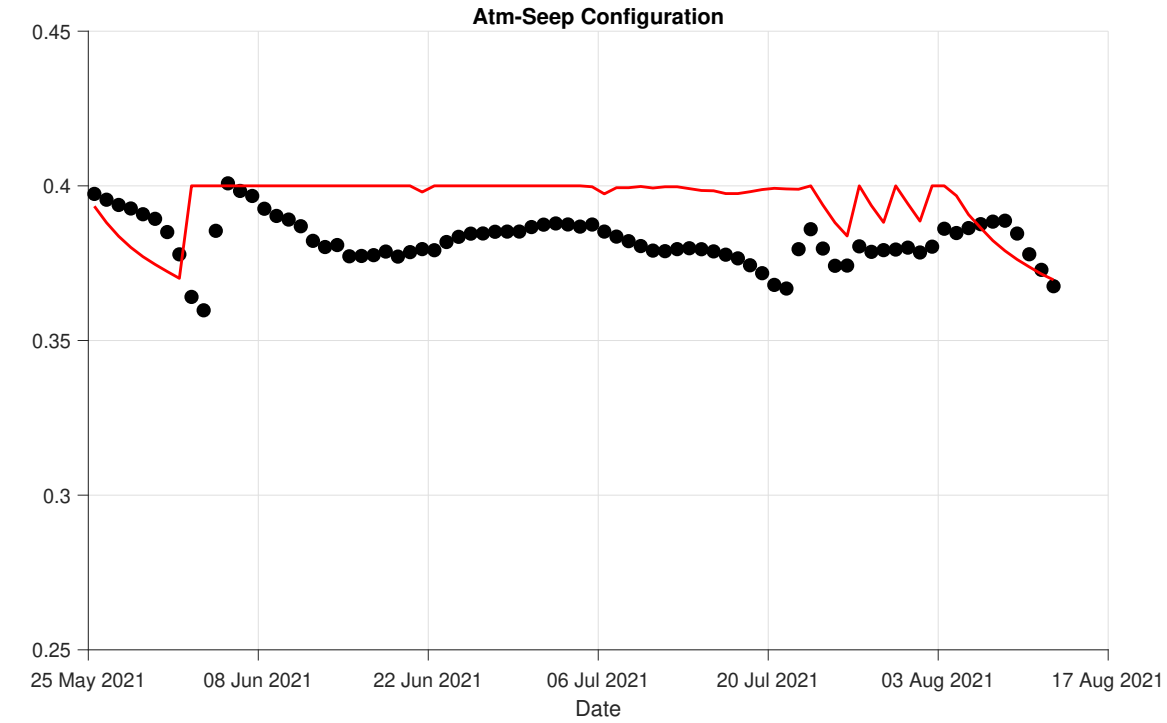


Figure 54: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

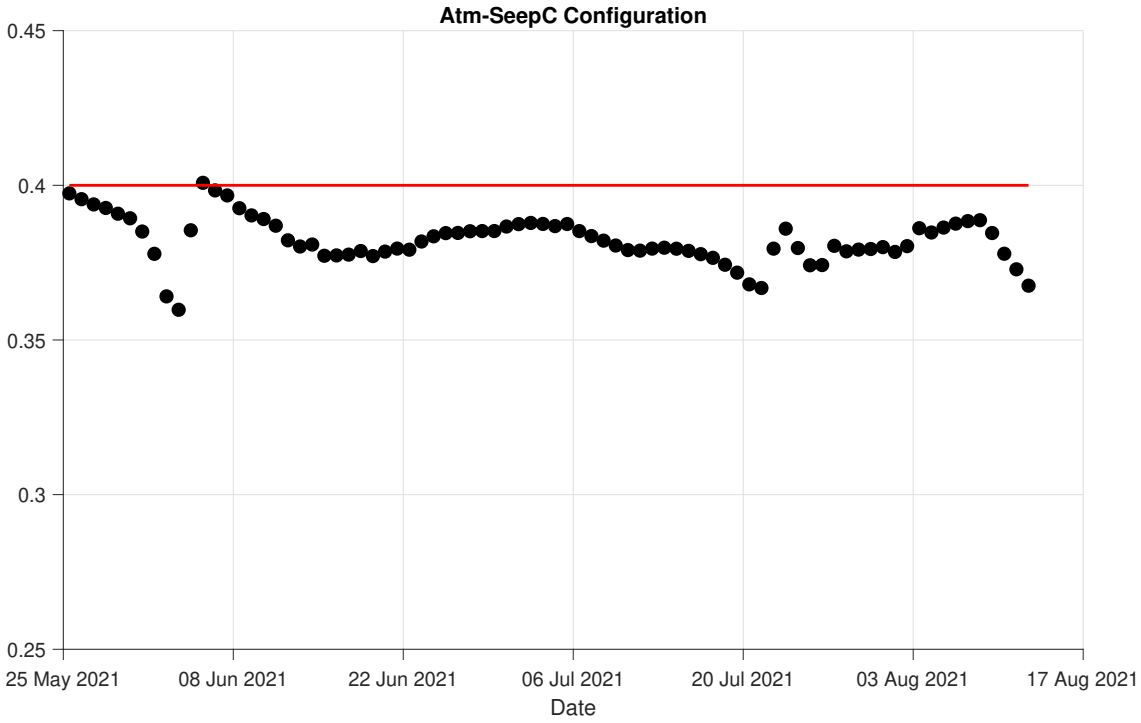


Figure 55: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

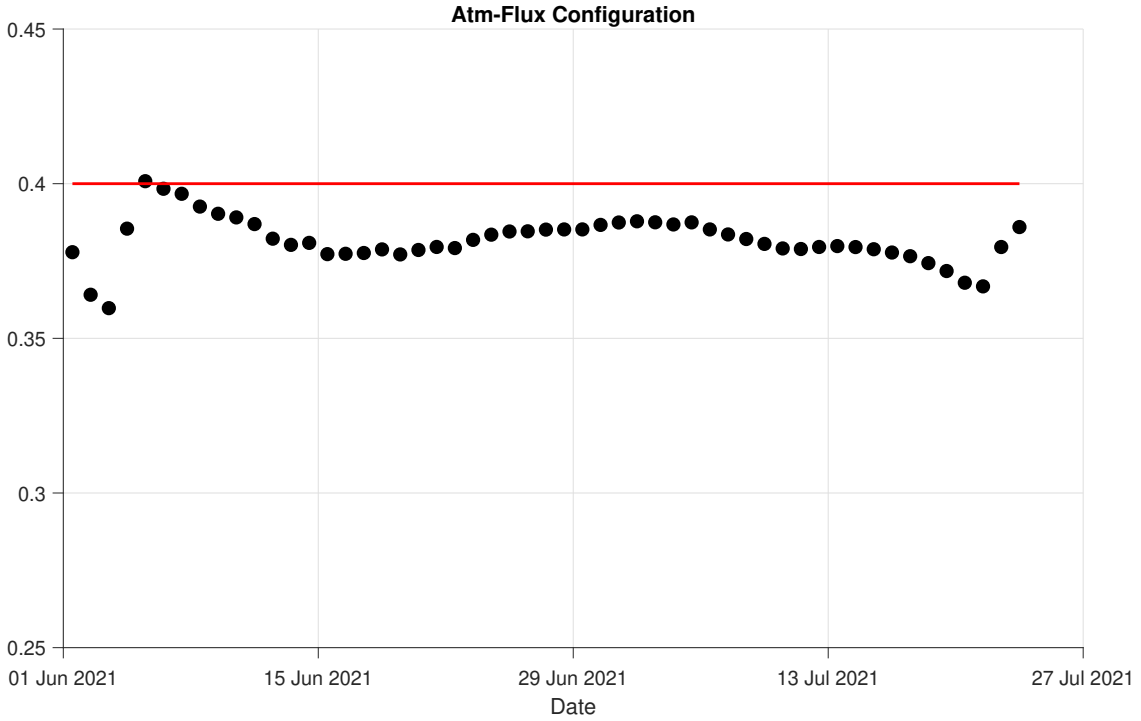


Figure 56: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

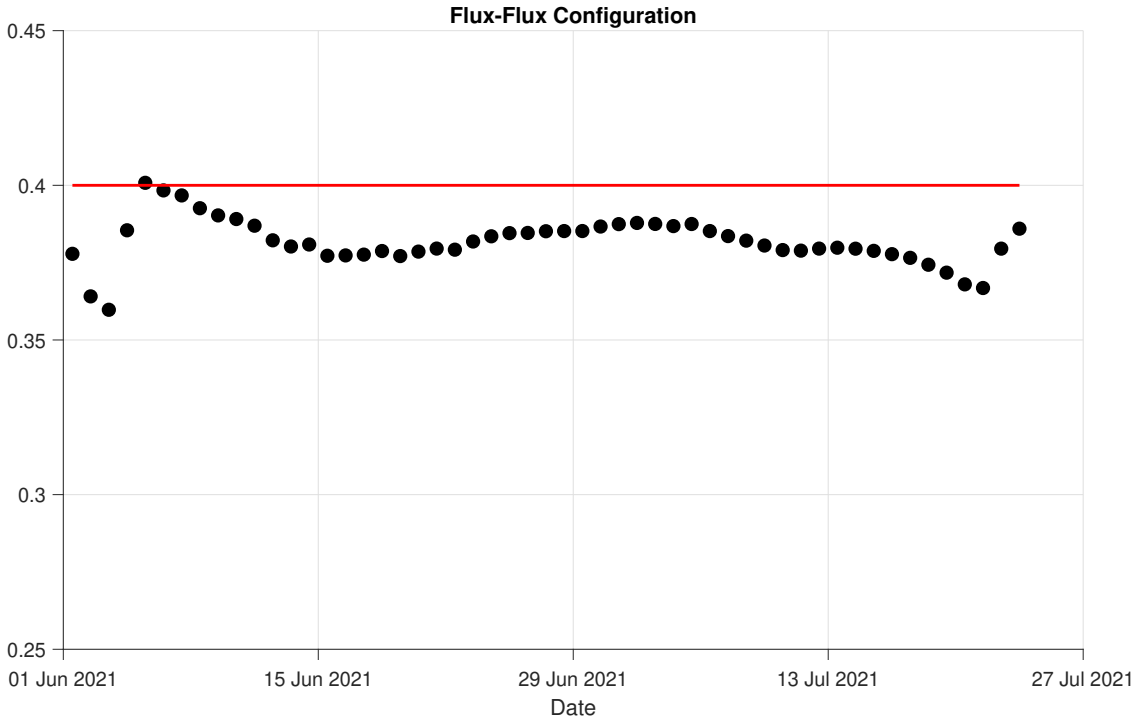


Figure 57: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

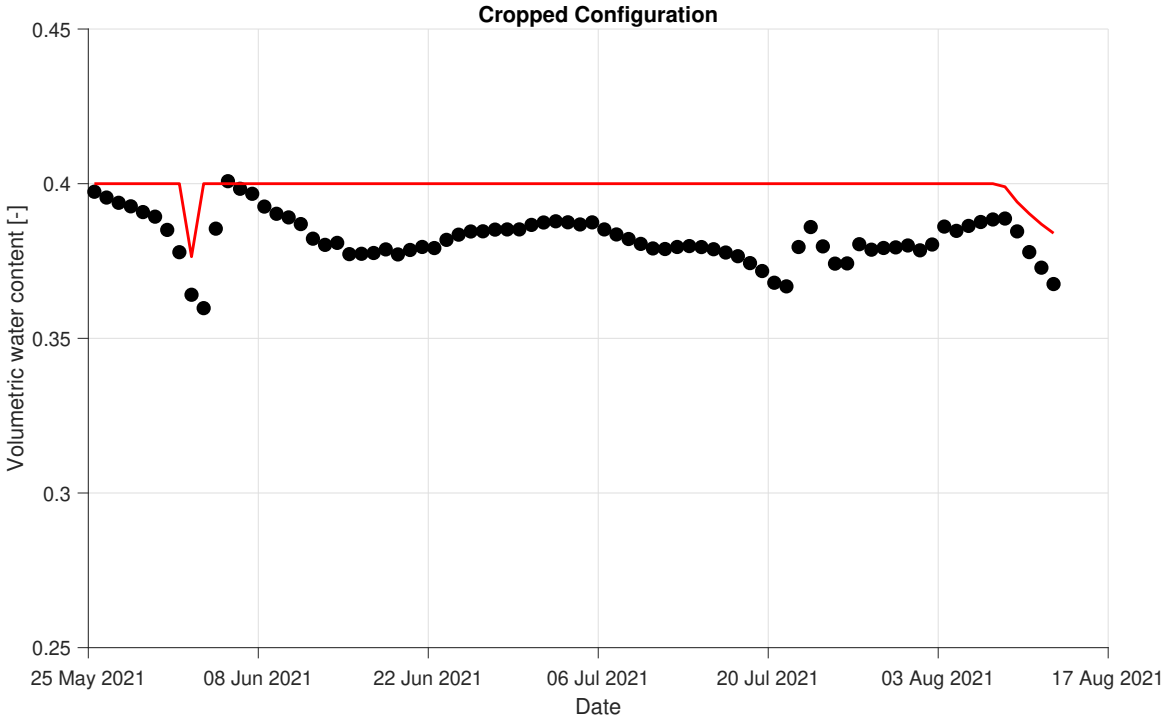


Figure 58: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

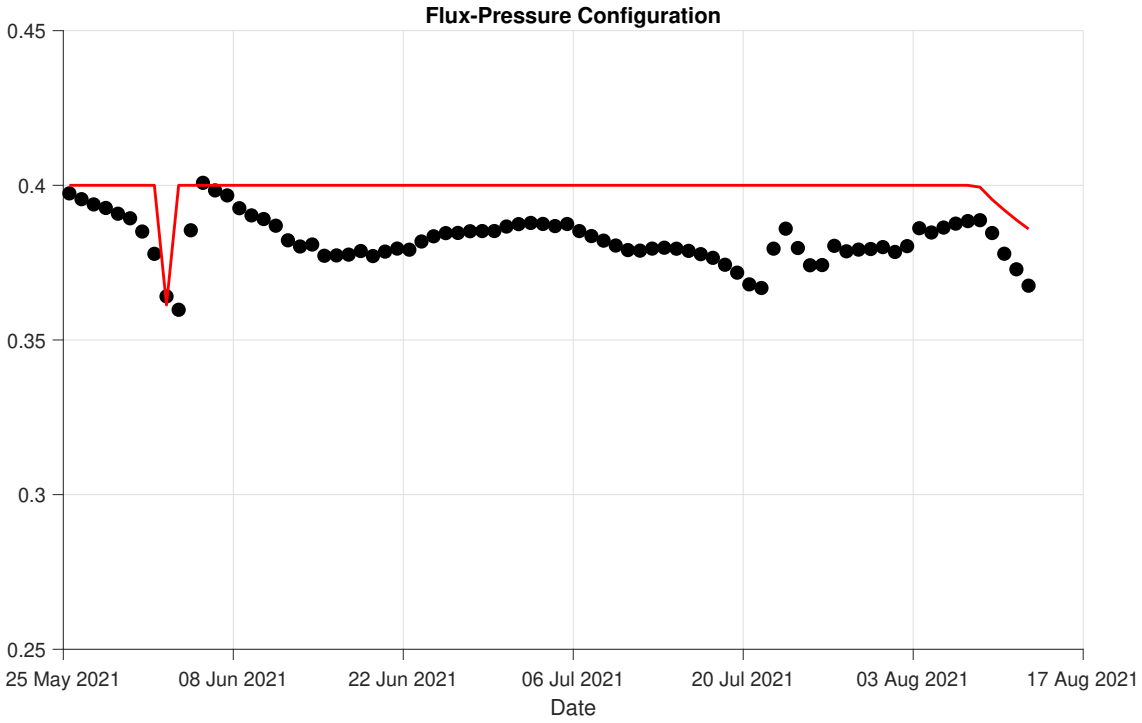


Figure 59: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

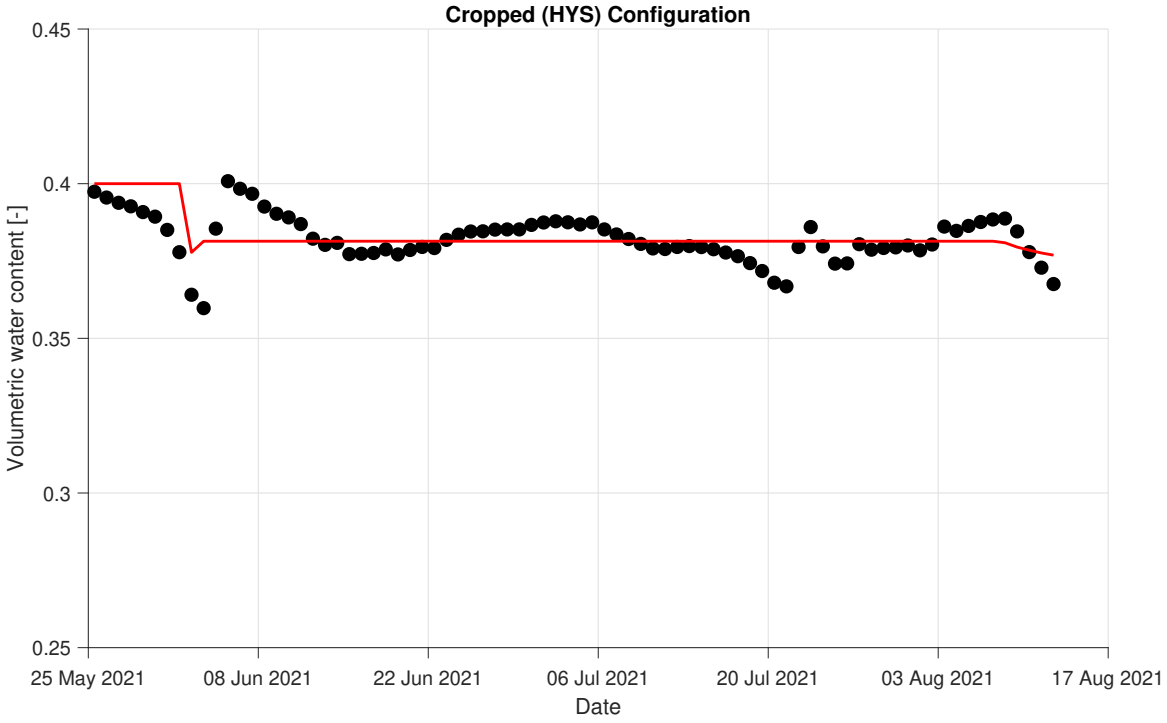


Figure 60: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

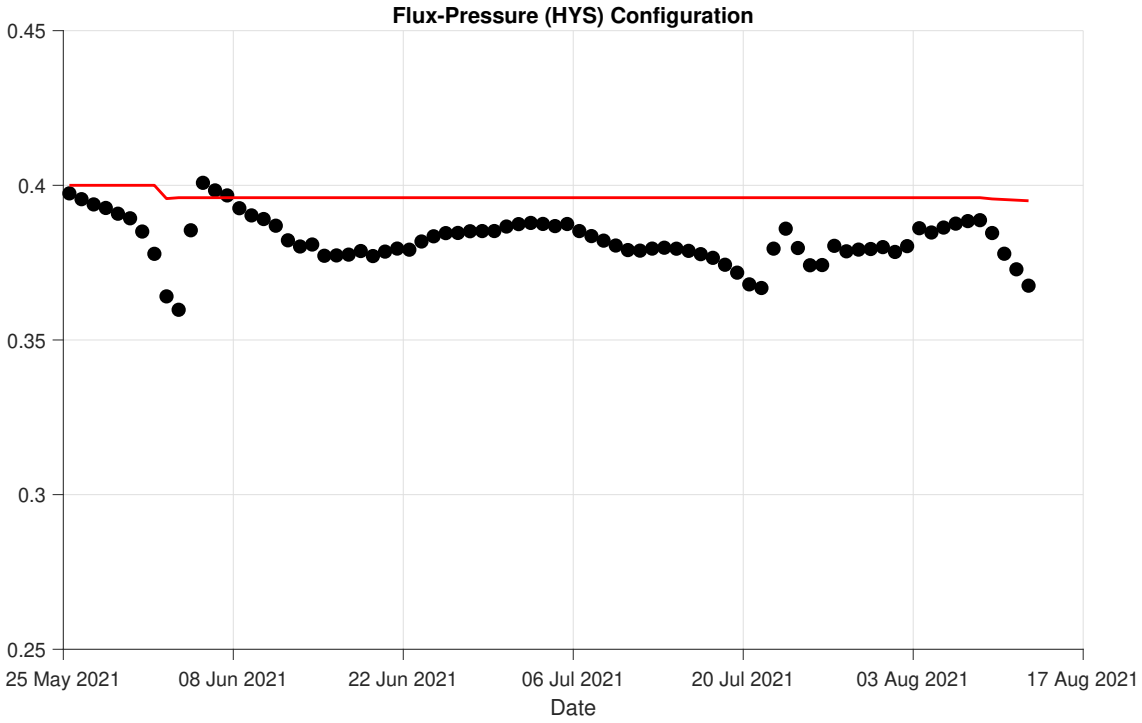


Figure 61: Moisture content at $z = 42\ cm$ versus time (black dots are measured values and red curve is simulated values).

C Pressure heads

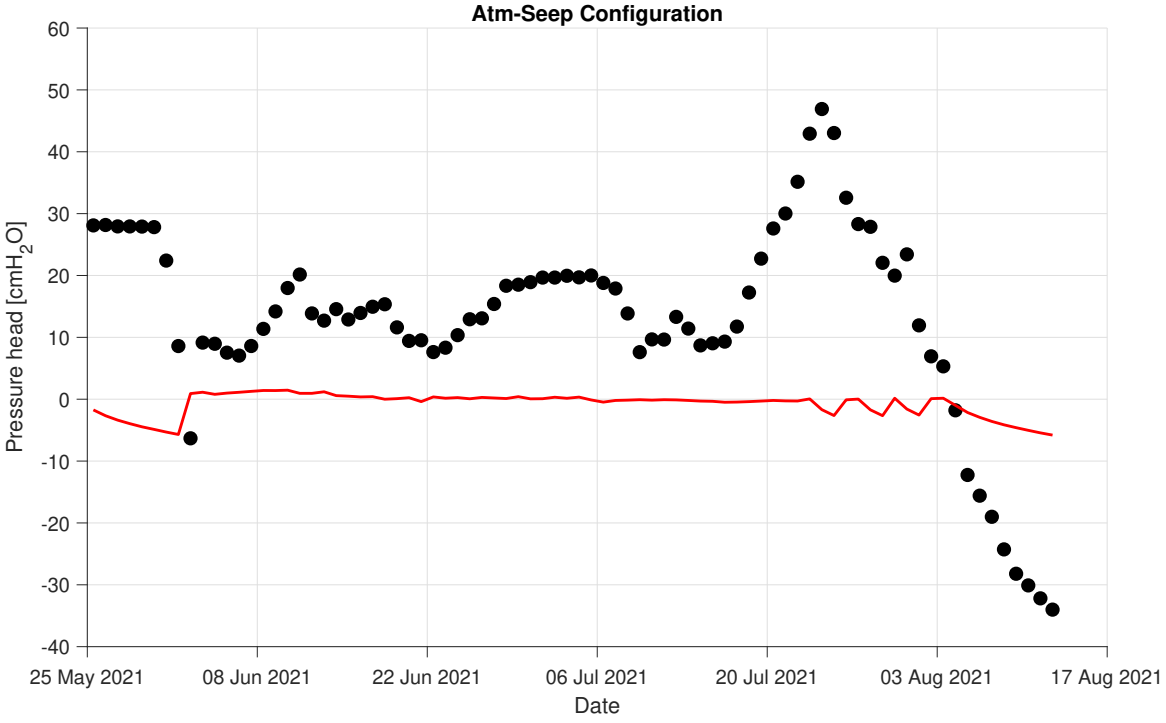


Figure 62: Pressure head at $z = 20\ cm$ versus time (black dots are measured values and red curve is simulated values).

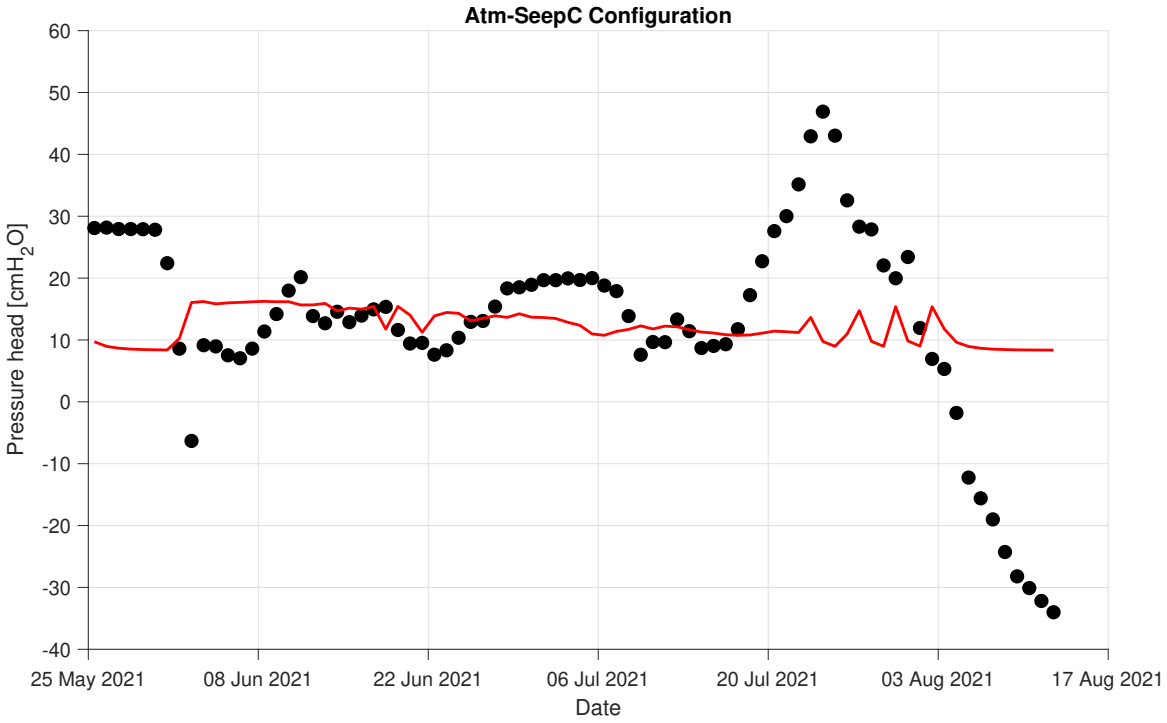


Figure 63: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

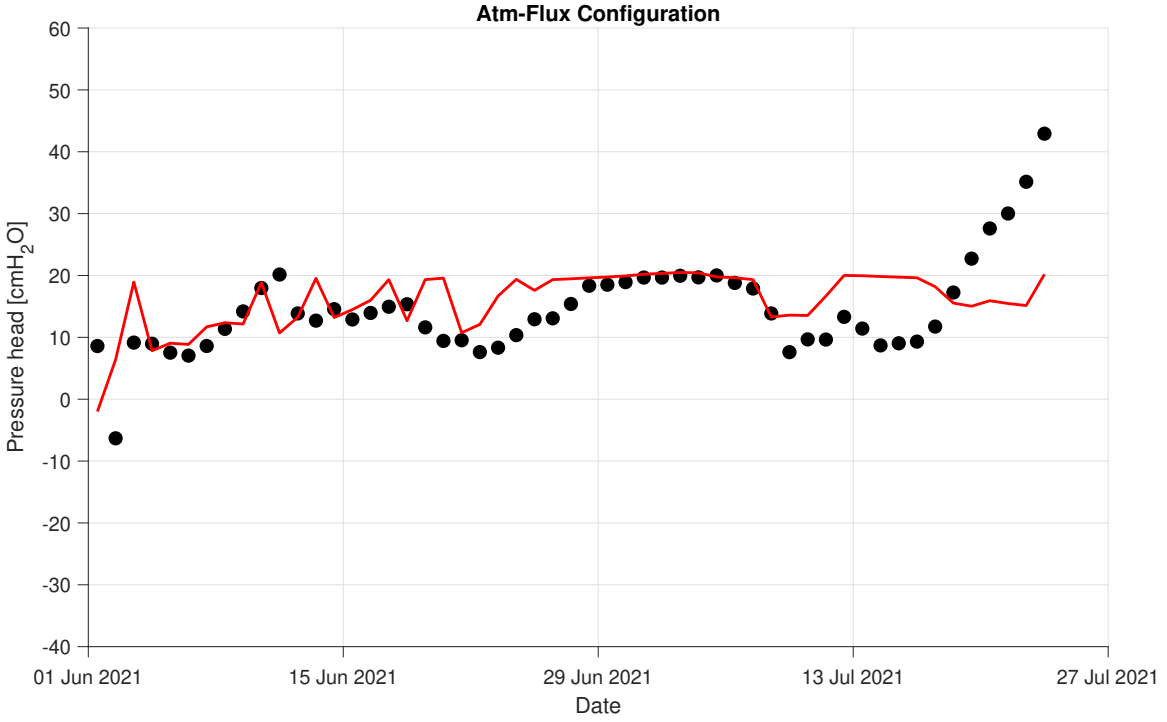


Figure 64: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

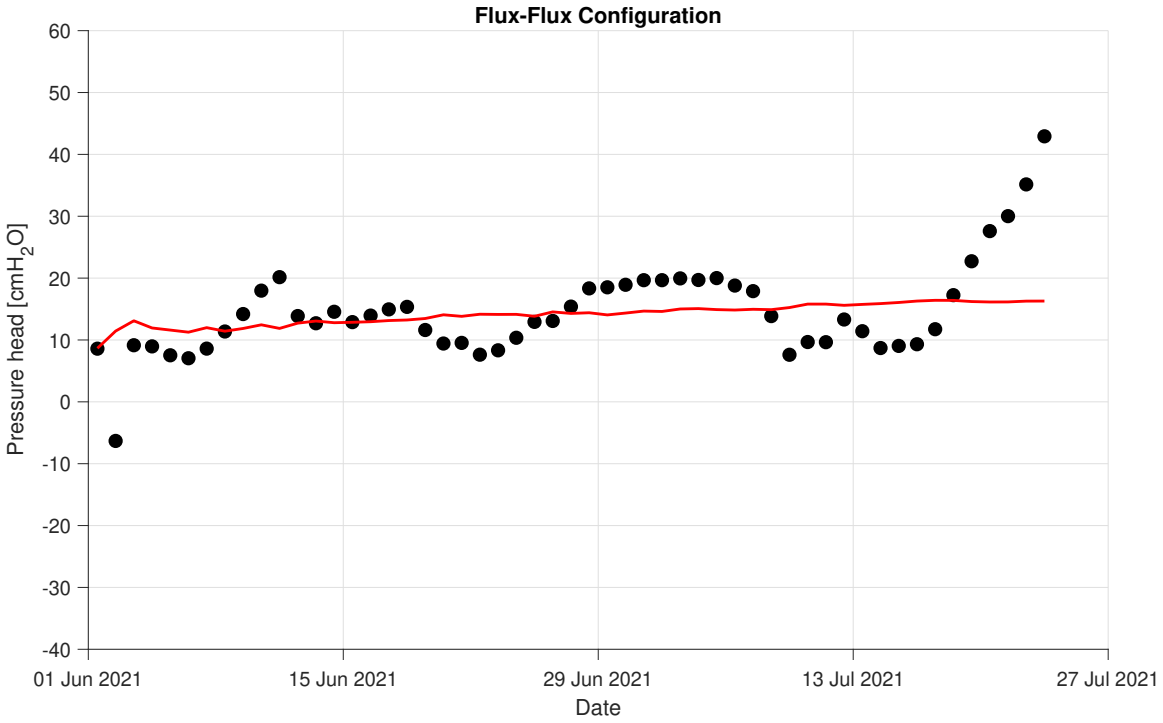


Figure 65: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

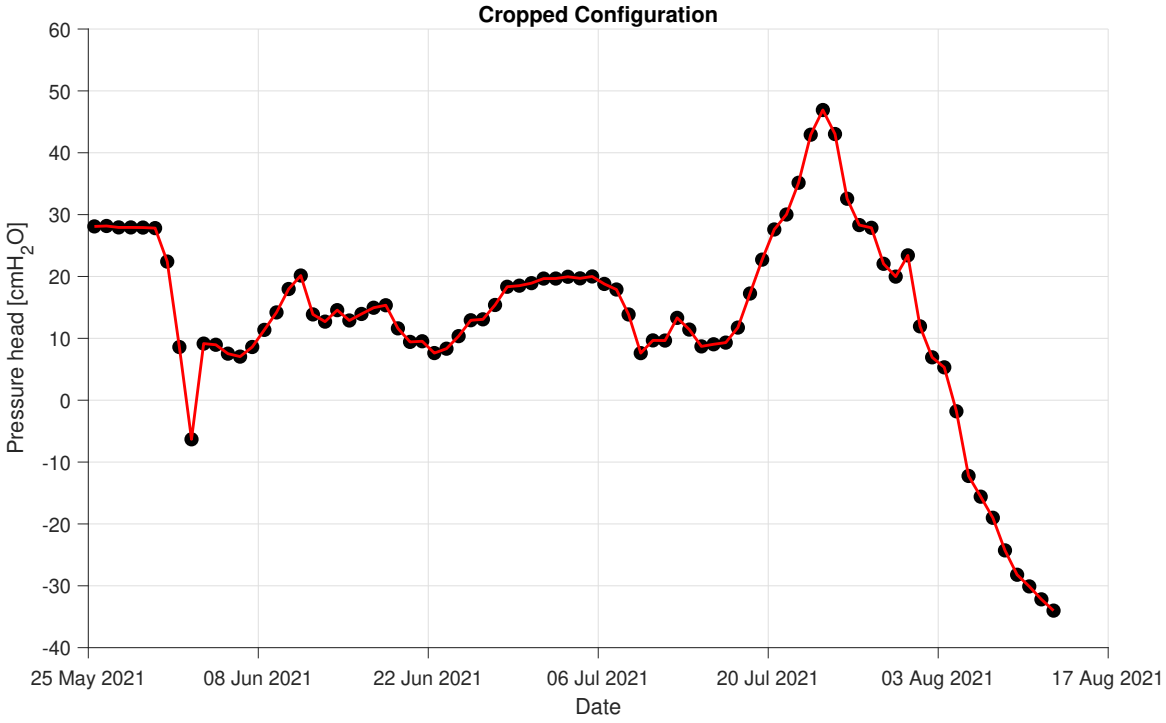


Figure 66: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

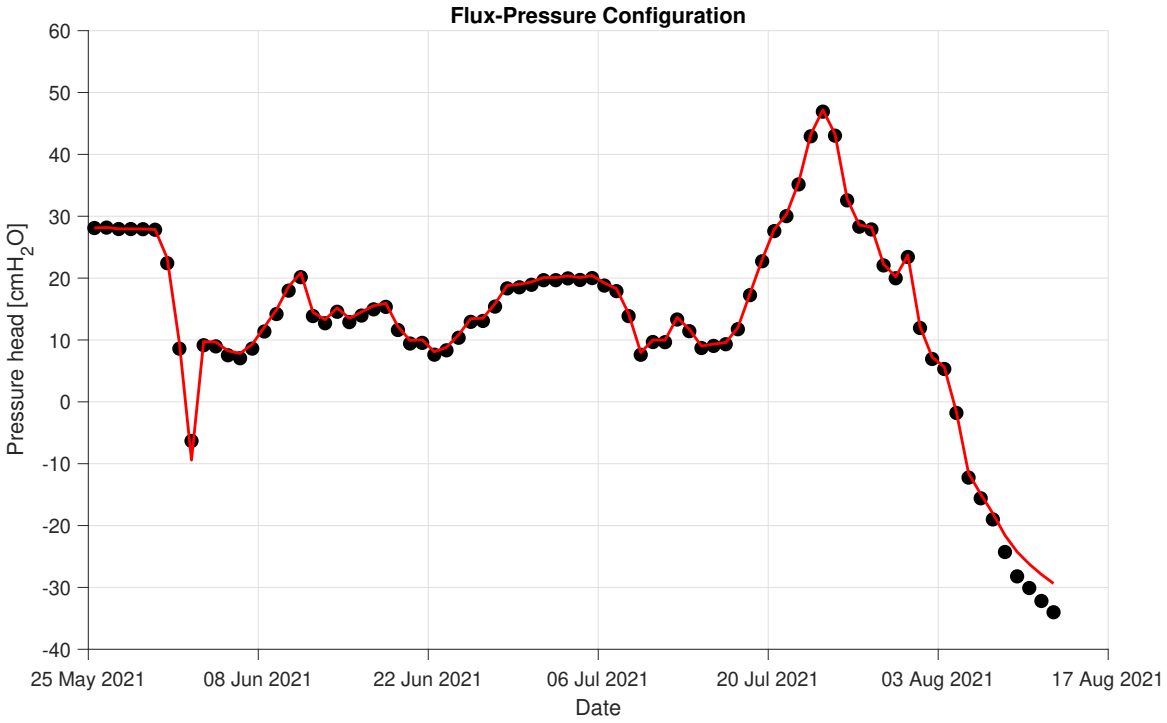


Figure 67: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

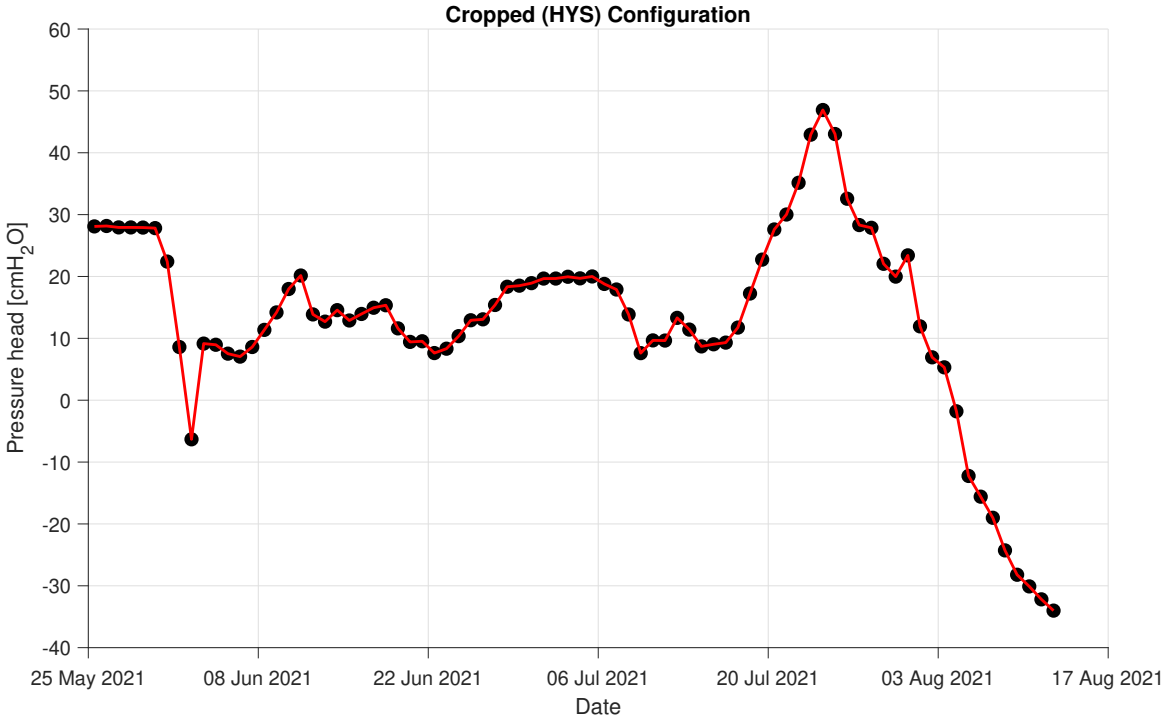


Figure 68: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

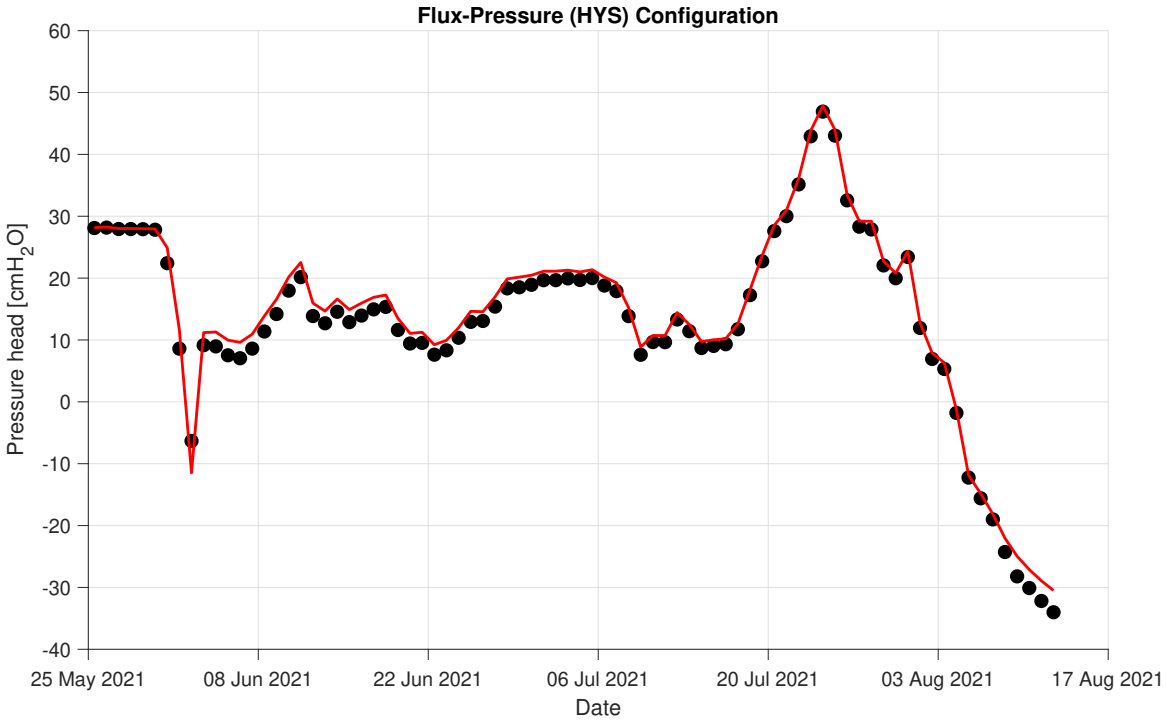


Figure 69: Pressure head at $z = 20$ cm versus time (black dots are measured values and red curve is simulated values).

D Discharge

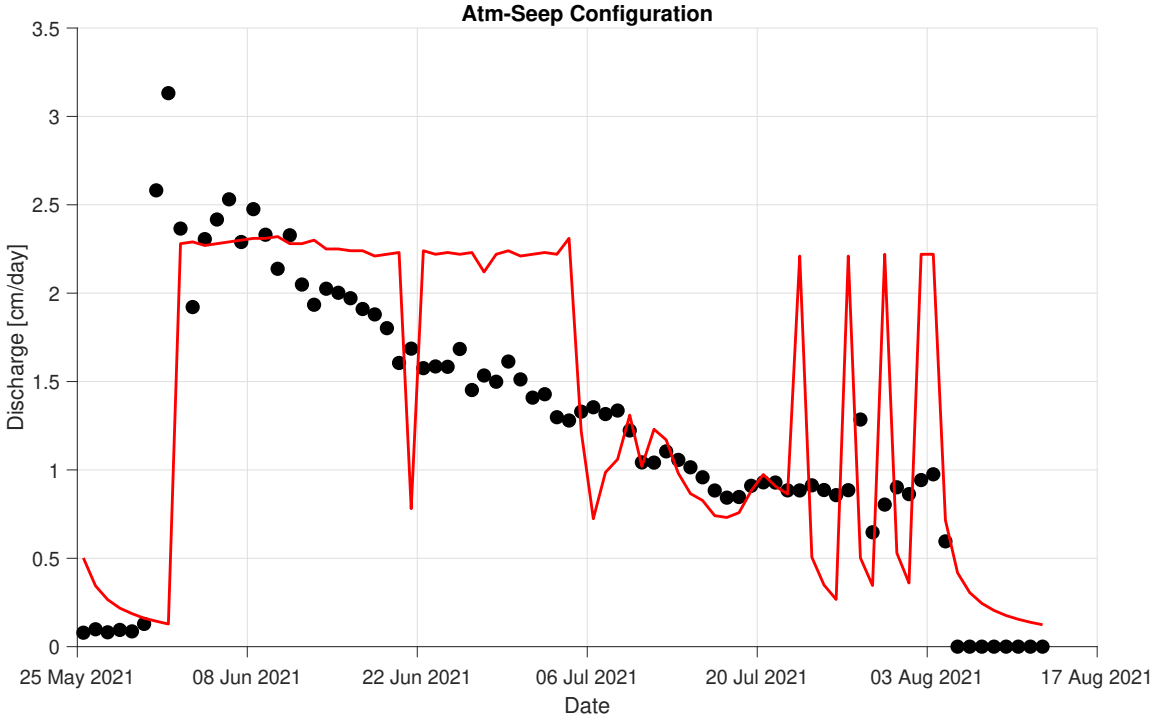


Figure 70: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

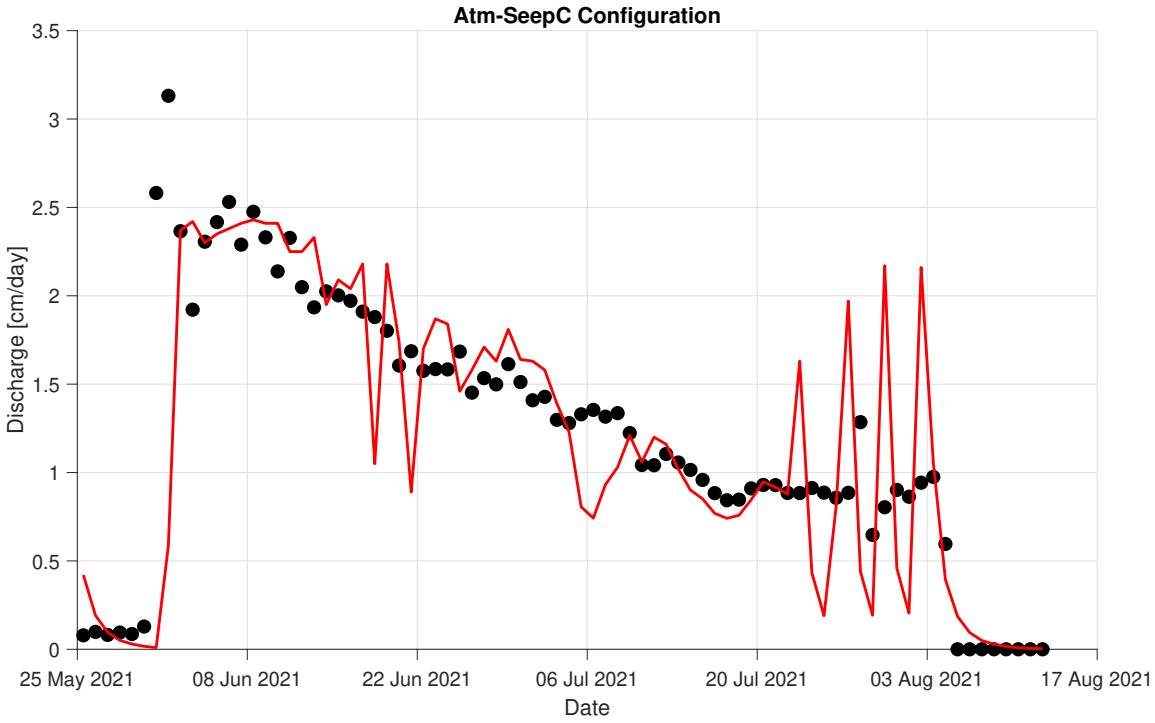


Figure 71: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

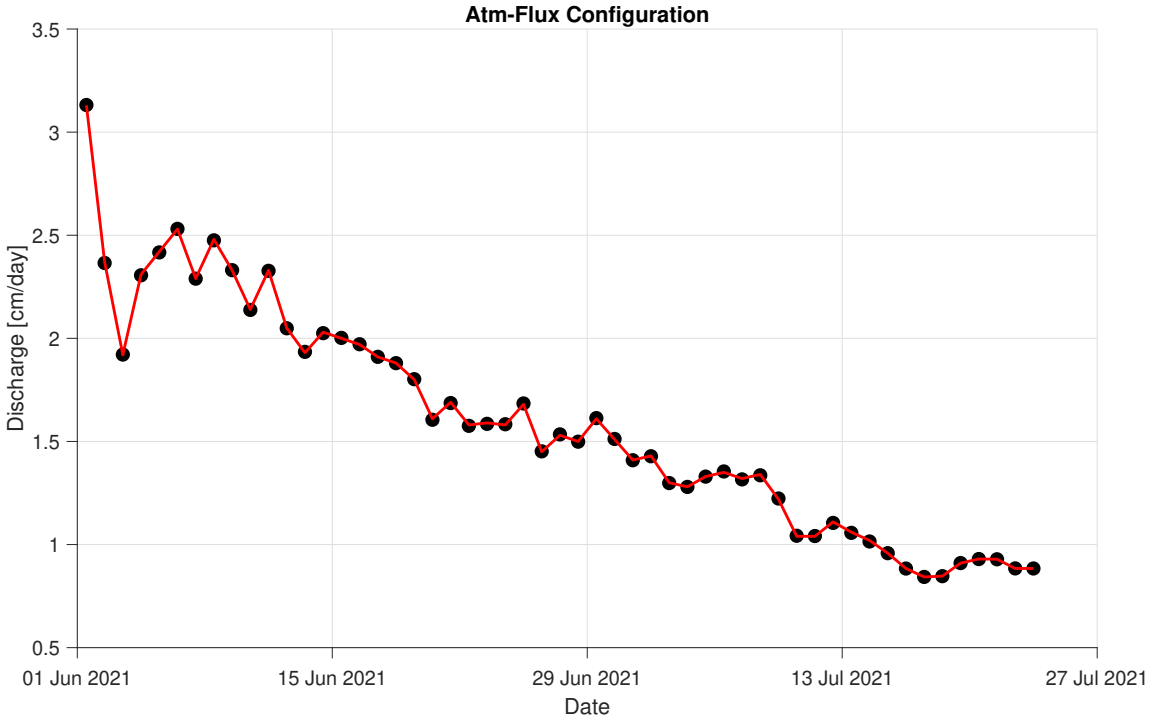


Figure 72: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

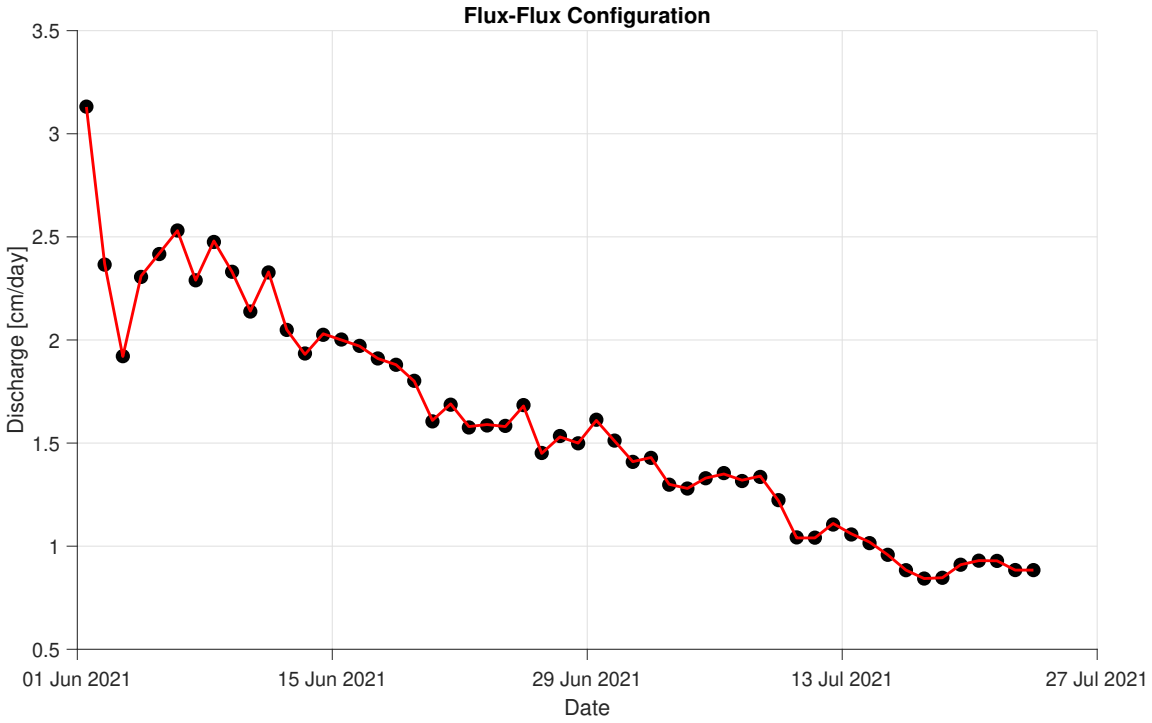


Figure 73: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

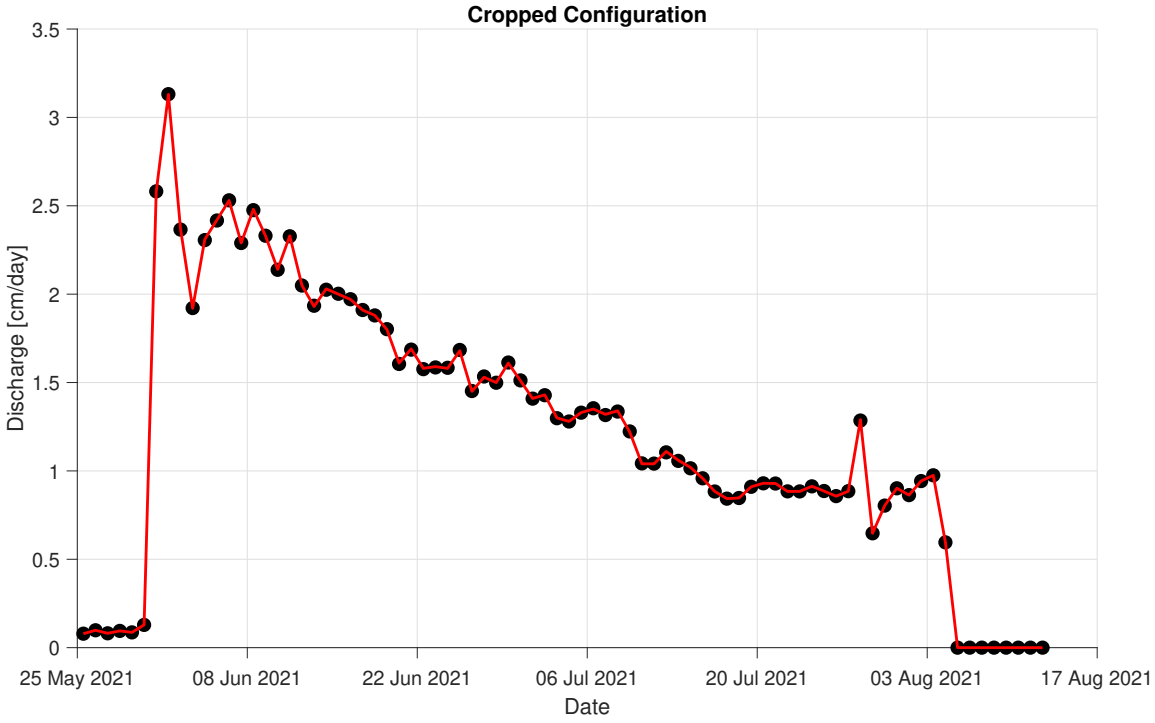


Figure 74: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

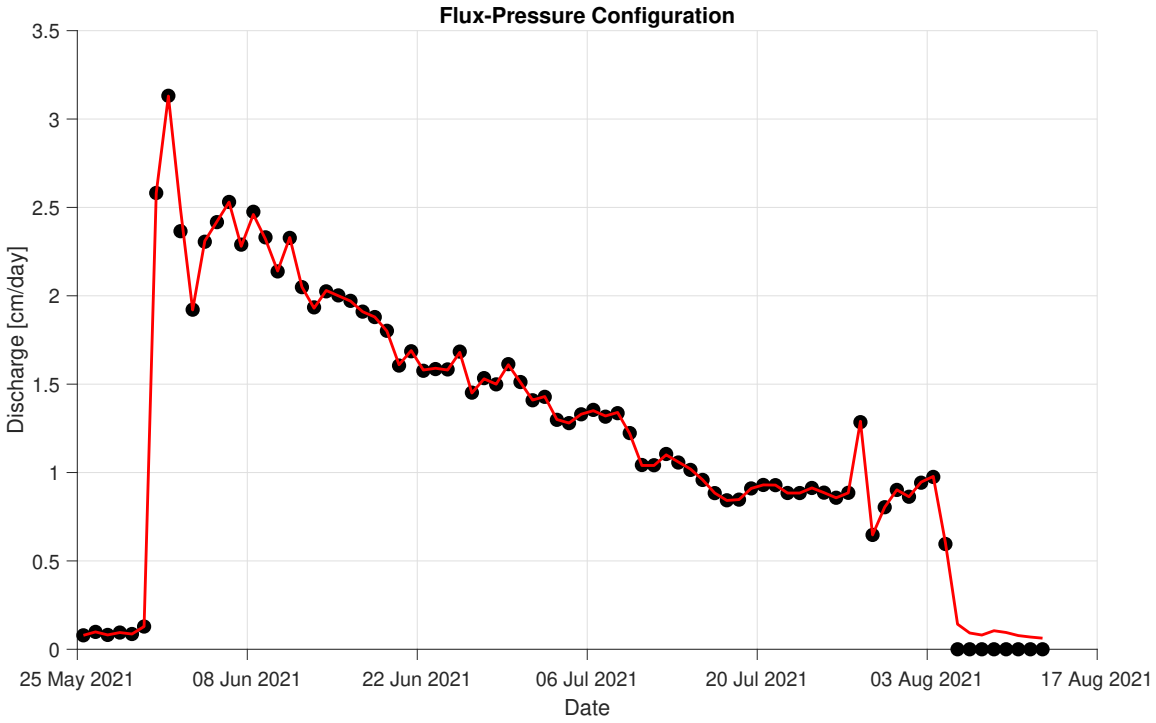


Figure 75: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

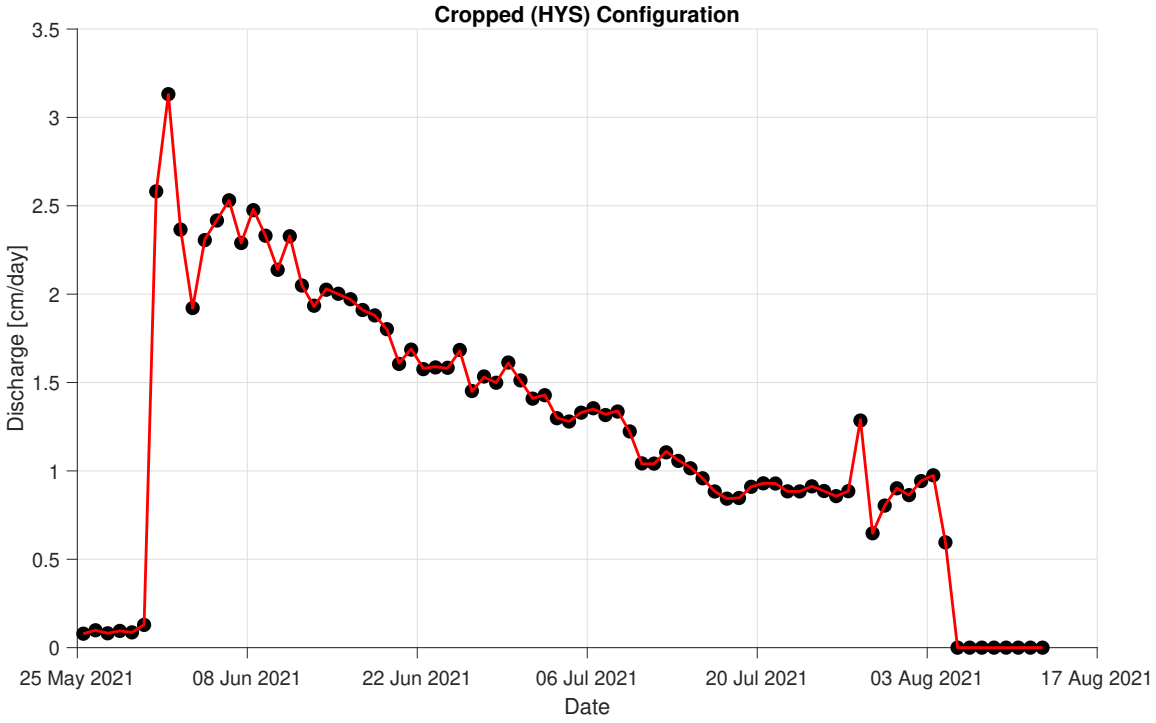


Figure 76: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

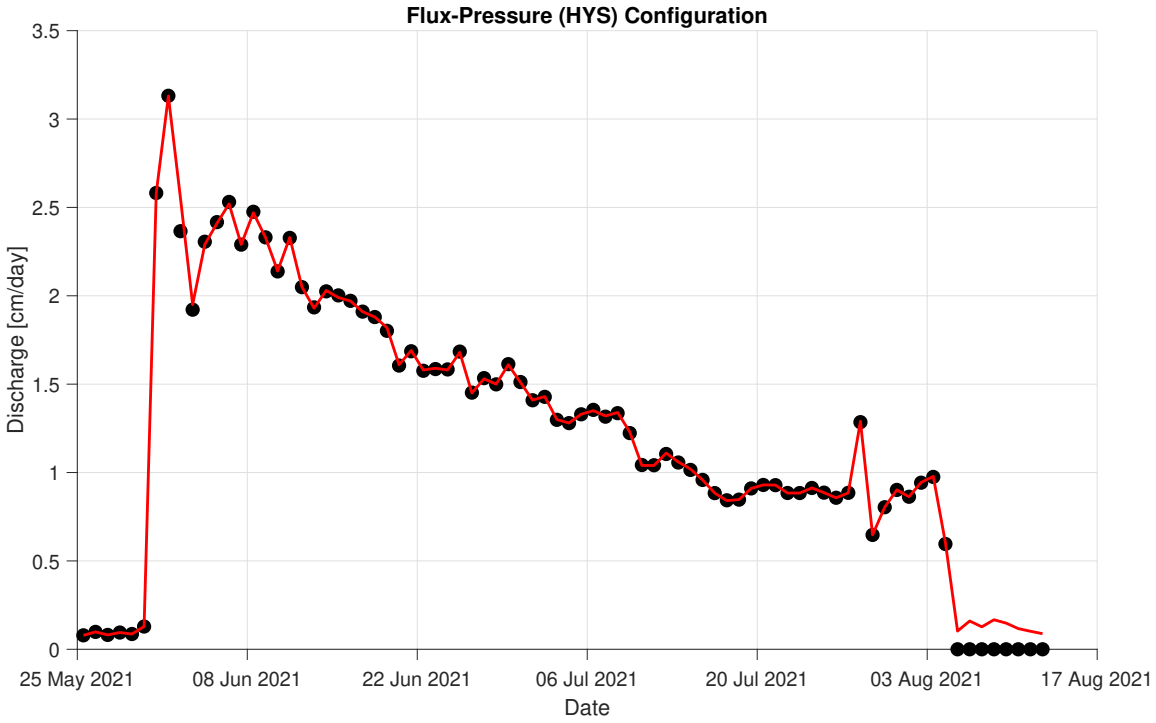
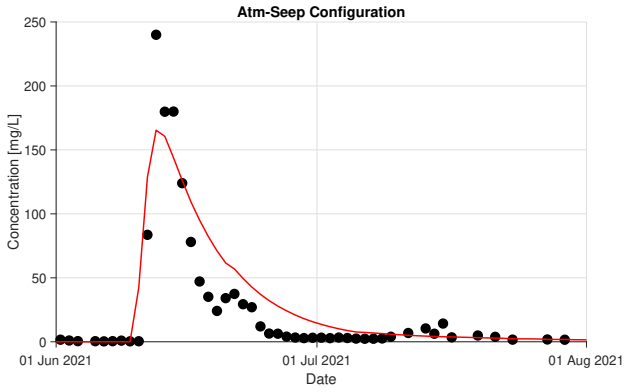
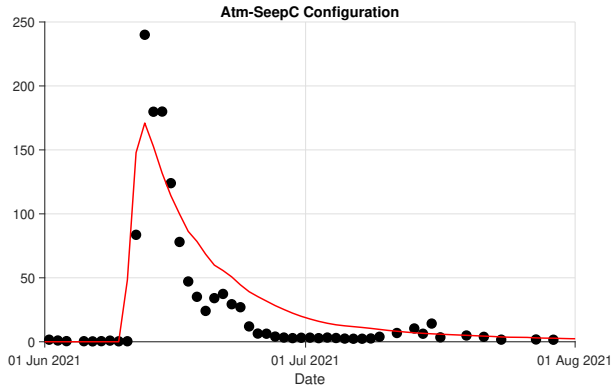


Figure 77: Discharge at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

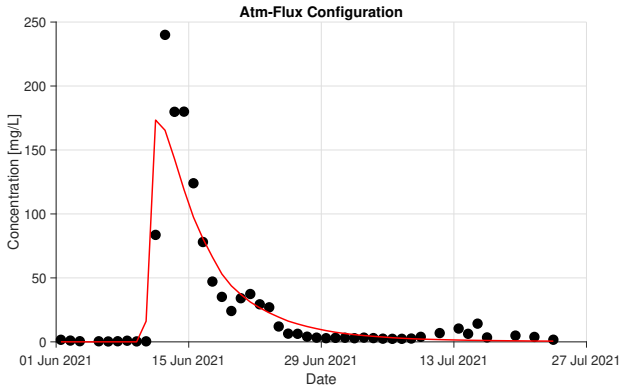
E Concentration at soil



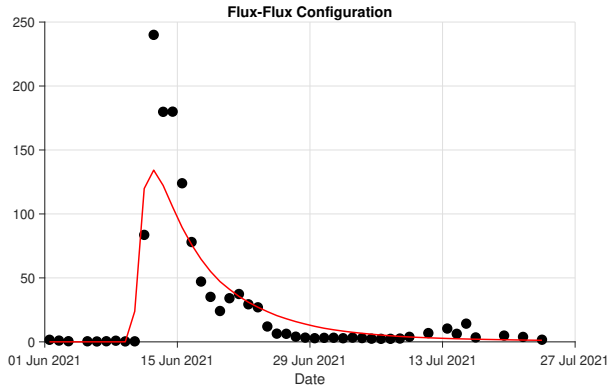
(a) Atm-Seep



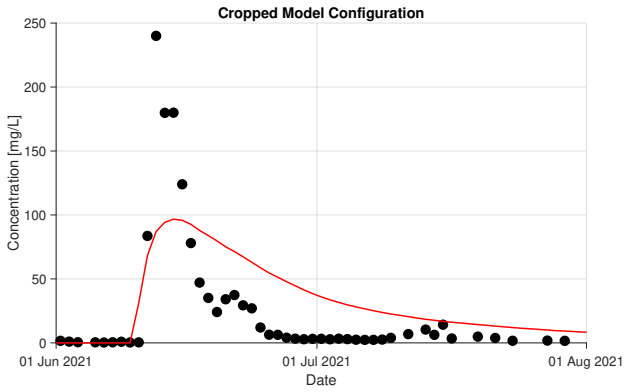
(b) Atm-SeepC



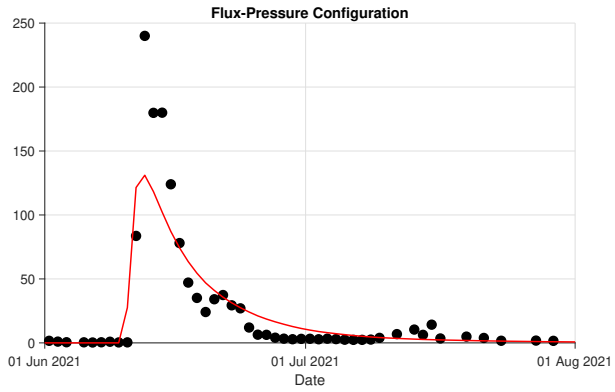
(c) Atm-Flux



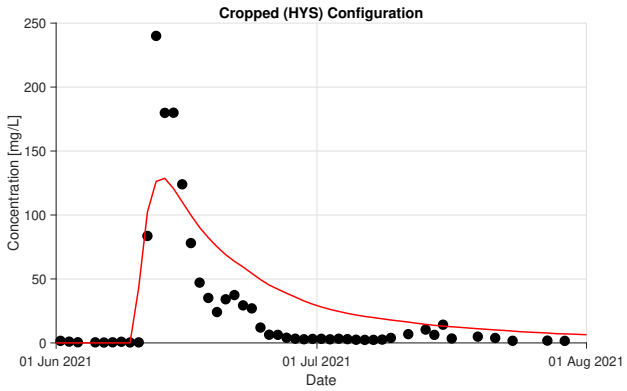
(d) Flux-Flux



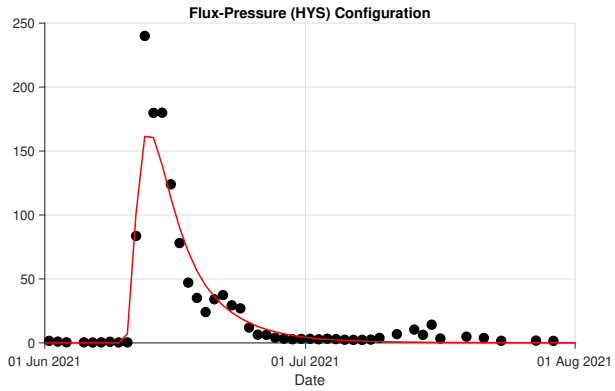
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure



(g) Pressure-Flux (Cropped HYS)



(h) Flux-Pressure (HYS)

Figure 78: Concentration at soil $z = 25\ cm$ (black dots are measured values from the experiment, and ~~red curve is simulated values~~).

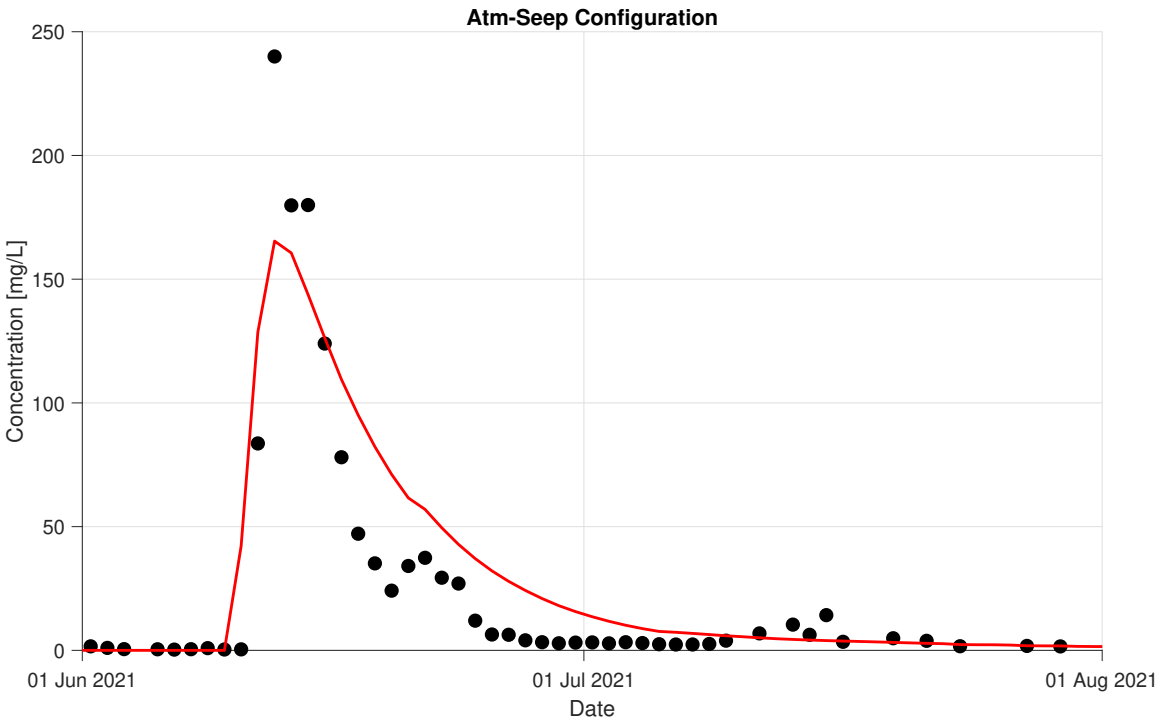


Figure 79: Concentration at $z = 25$ cm versus time (black dots are measured values and red curve is simulated values).

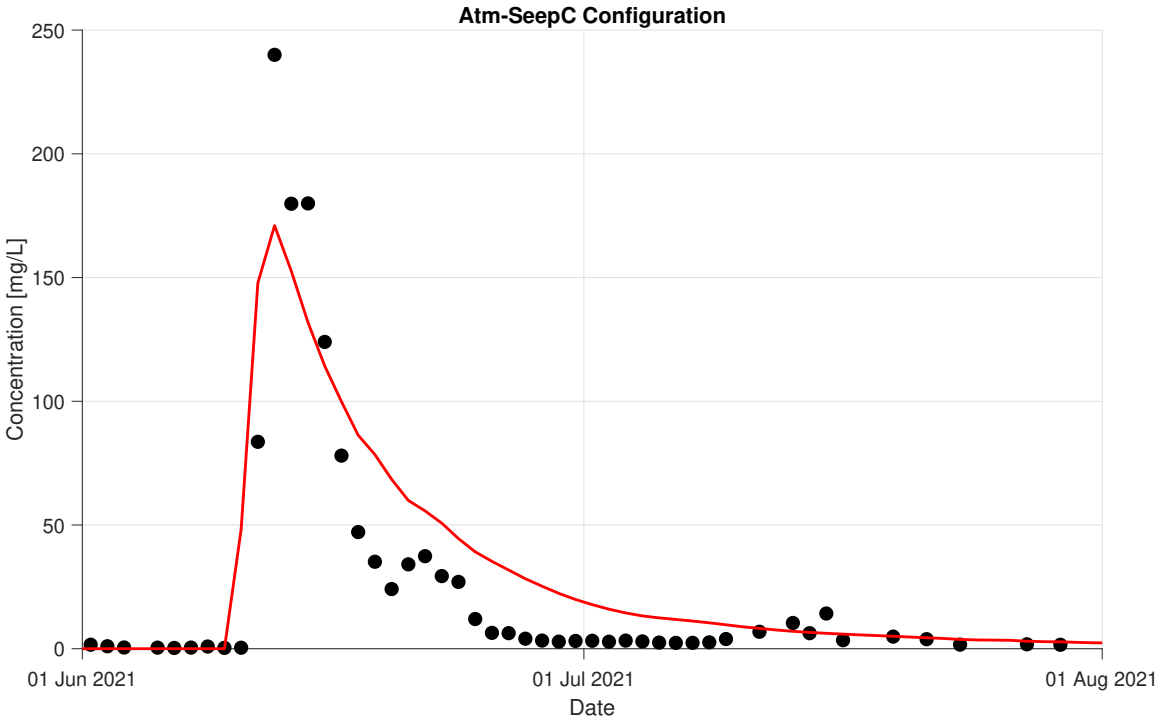


Figure 80: Concentration at $z = 25$ cm versus time (black dots are measured values and red curve is simulated values).

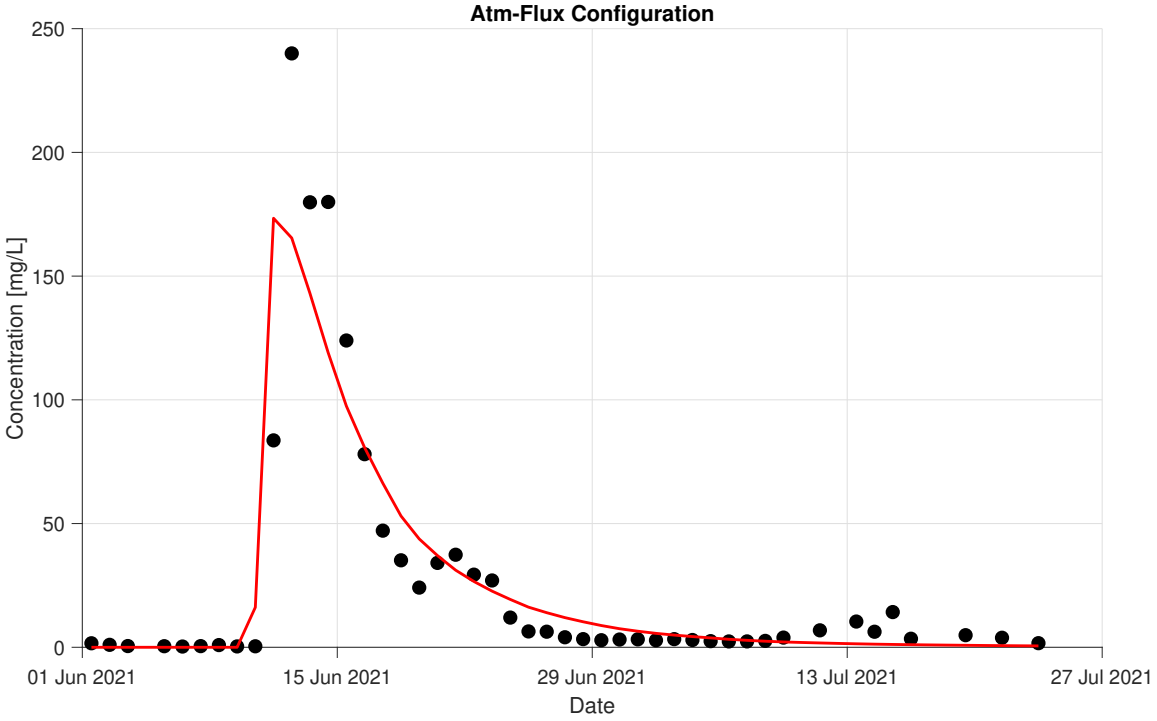


Figure 81: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

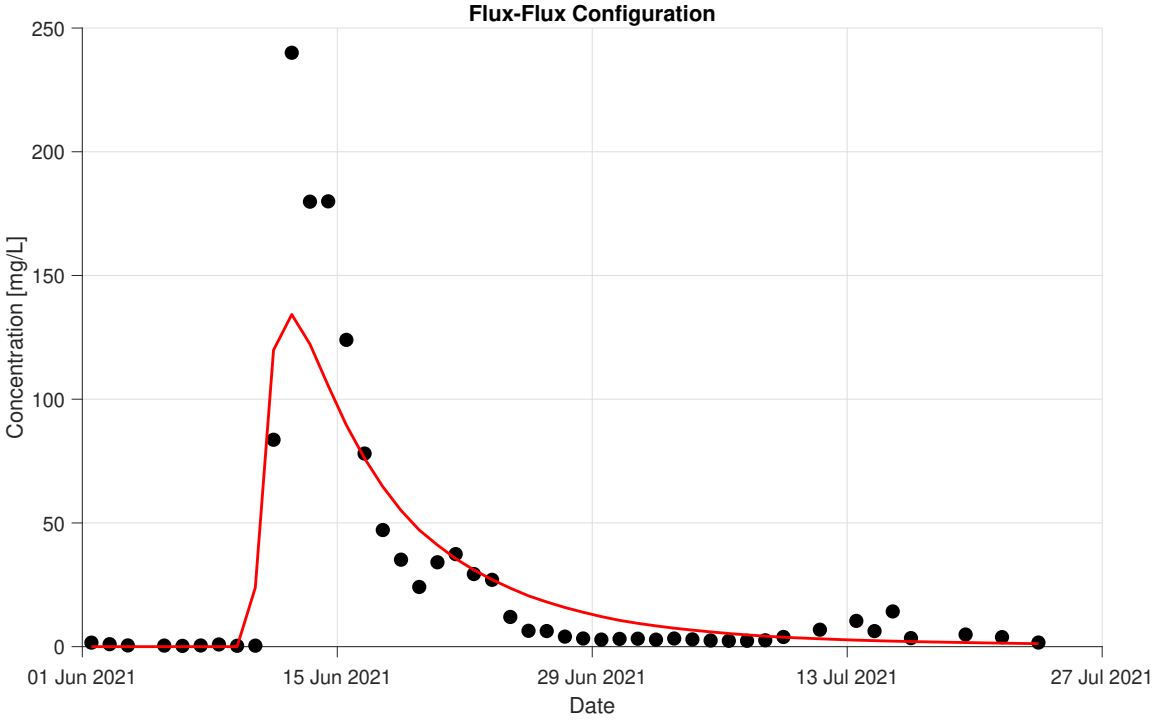


Figure 82: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

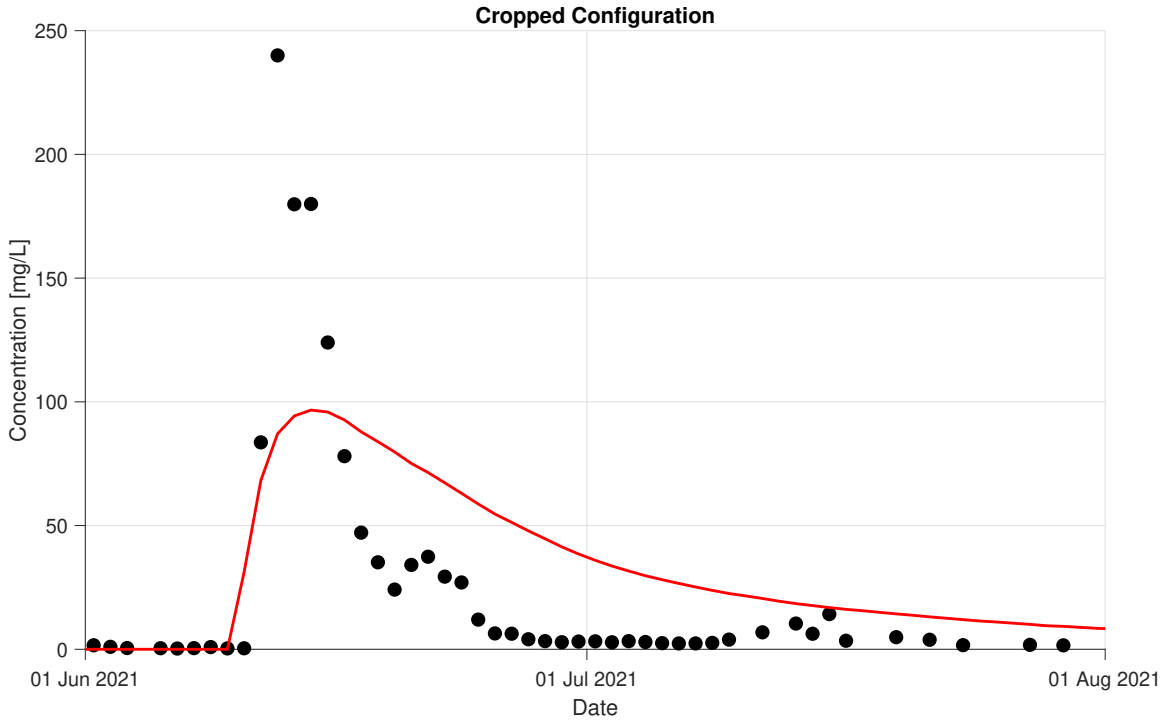


Figure 83: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

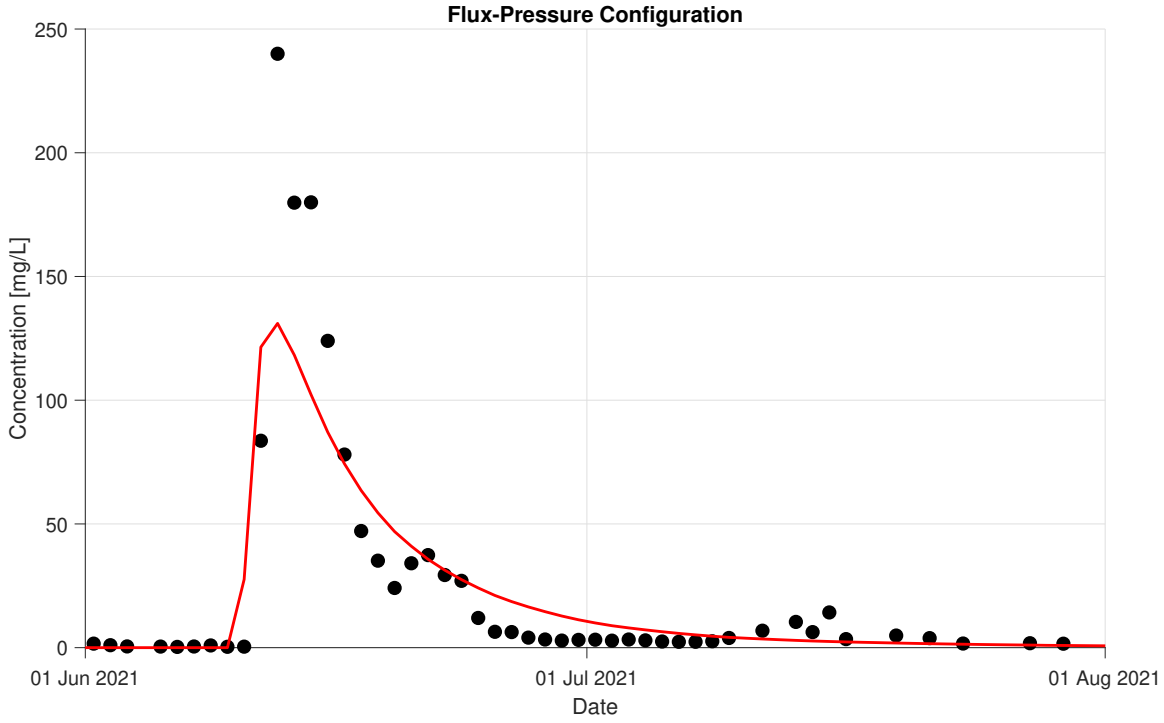


Figure 84: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

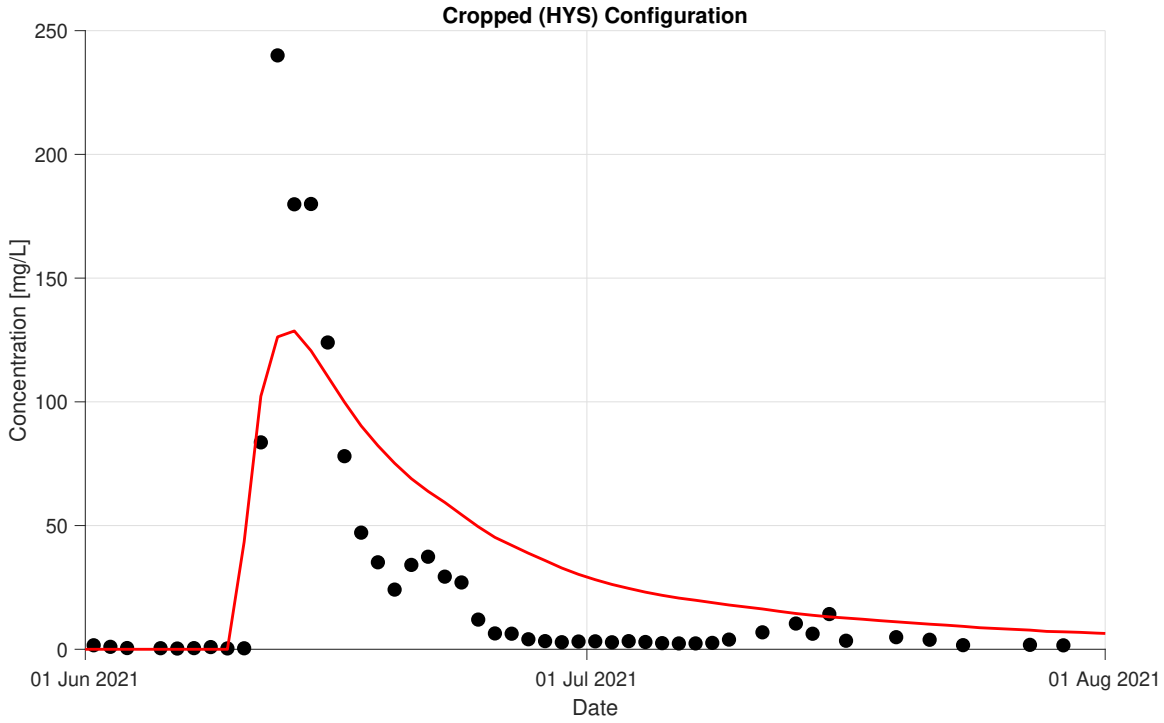


Figure 85: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

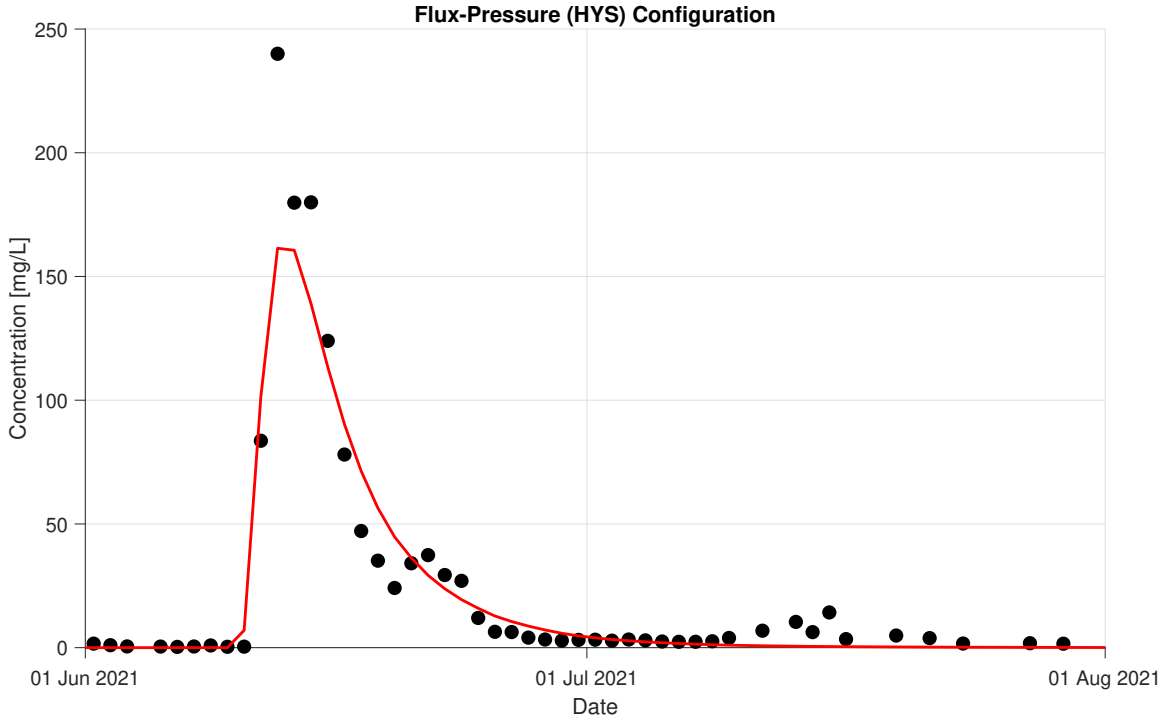
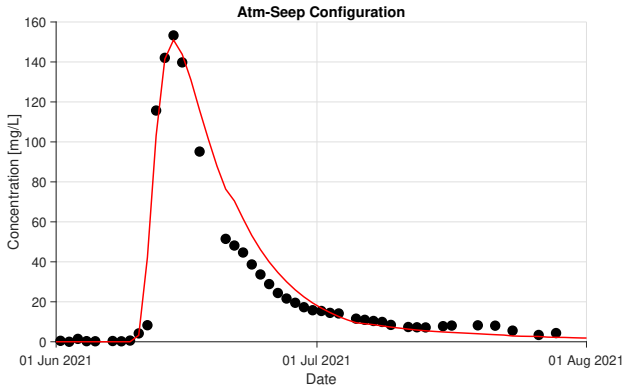
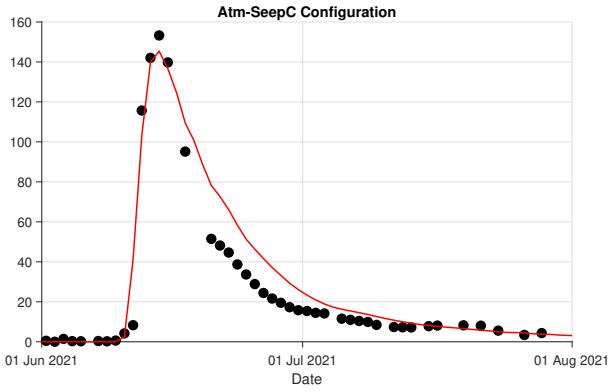


Figure 86: Concentration at $z = 25\ cm$ versus time (black dots are measured values and red curve is simulated values).

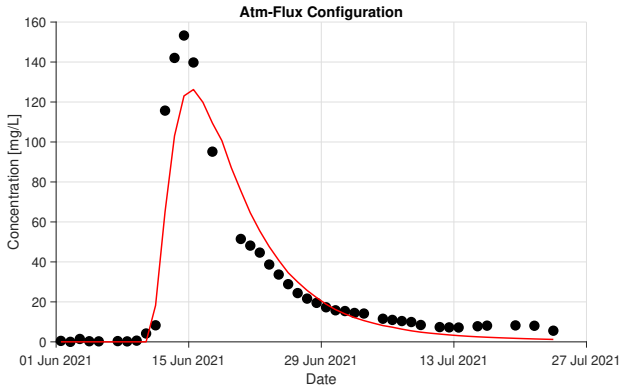
F Concentration at discharge



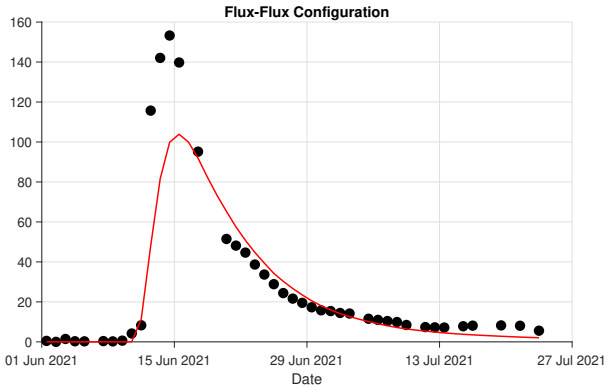
(a) Atm-Seep



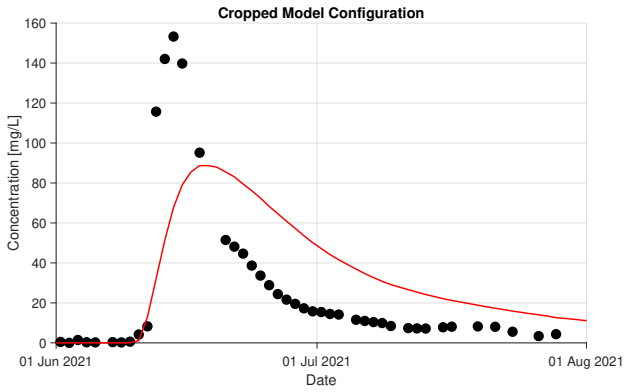
(b) Atm-SeepC



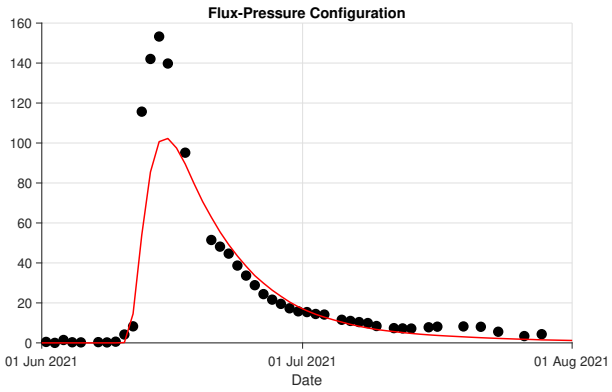
(c) Atm-Flux



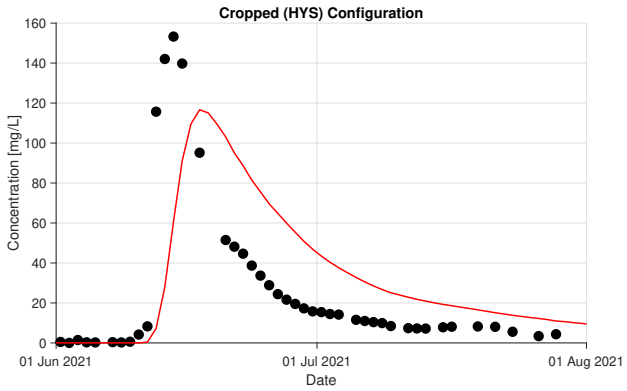
(d) Flux-Flux



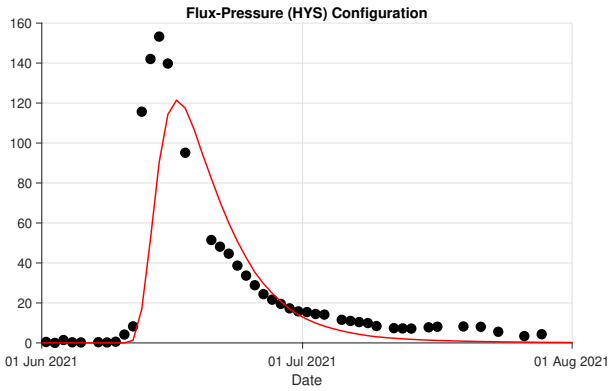
(e) Pressure-Flux (Cropped)



(f) Flux-Pressure



(g) Pressure-Flux (Cropped HYS)



(h) Flux-Pressure (HYS)

Figure 87: Concentration at the outlet $z = 50\ cm$ (black dots are measured values from the experiment, ~~and red curve is simulated values~~).

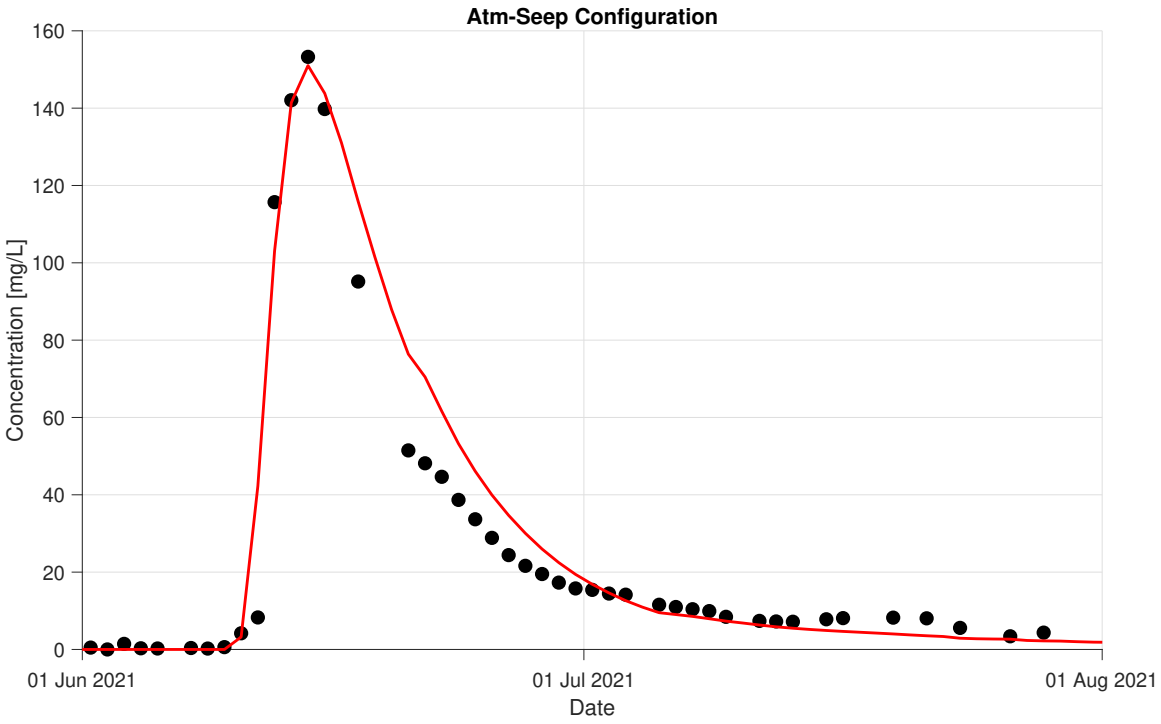


Figure 88: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).

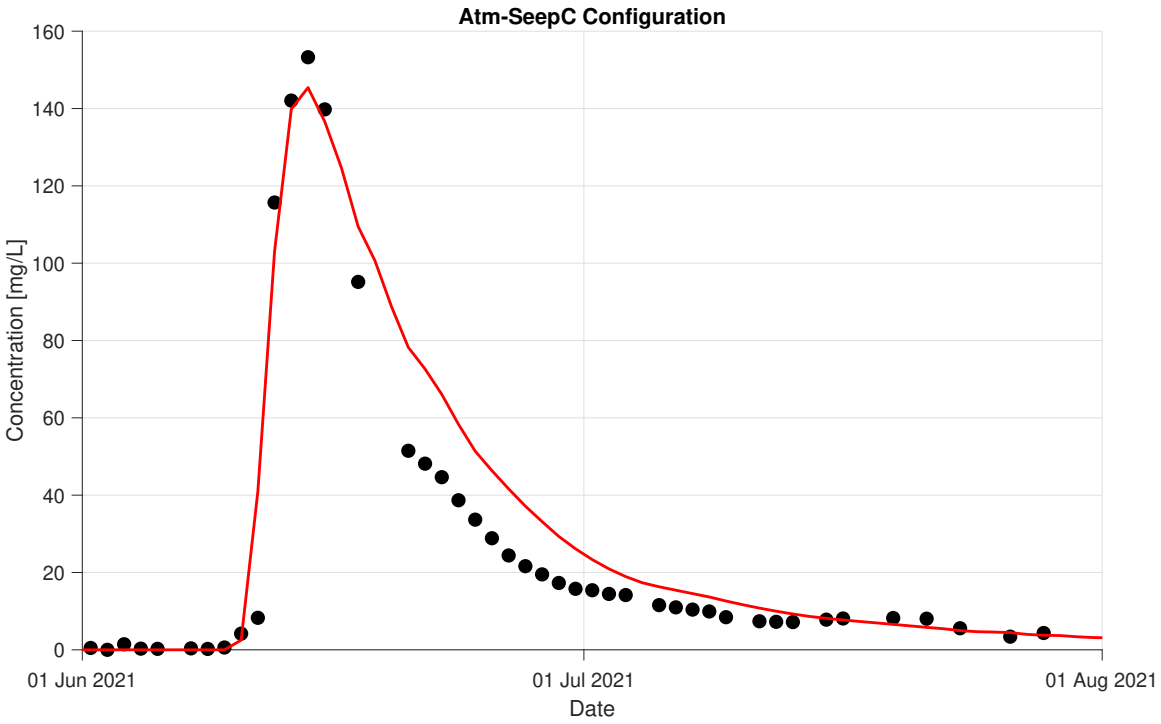


Figure 89: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).

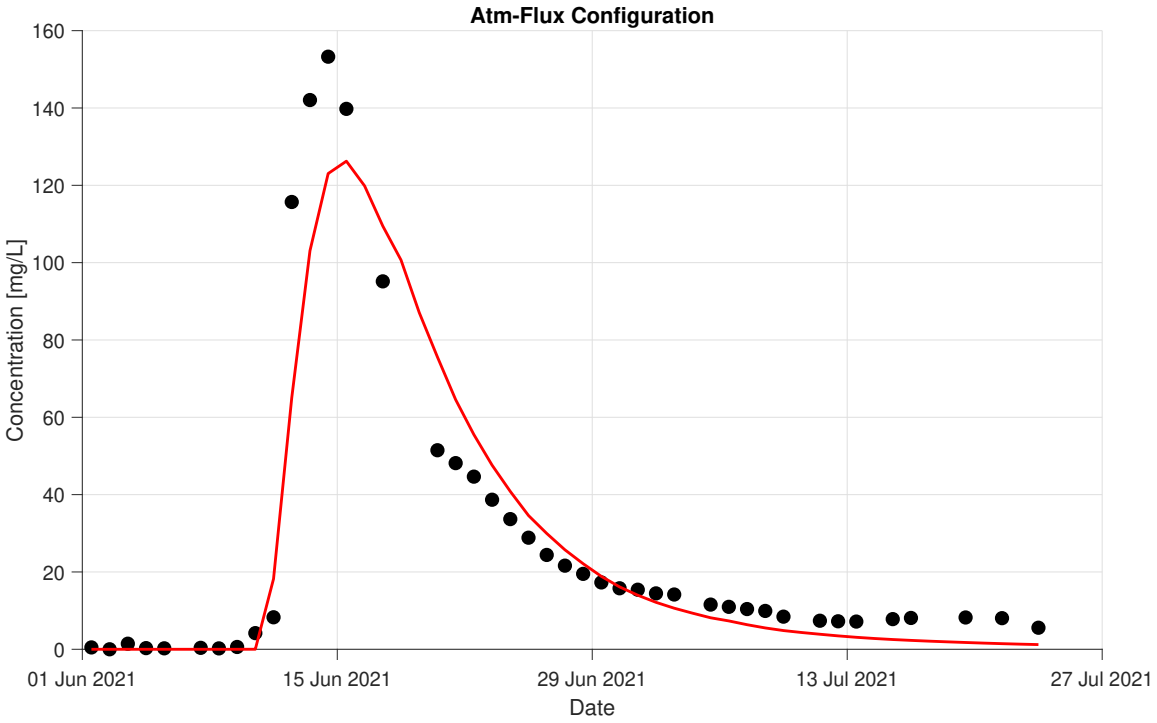


Figure 90: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).

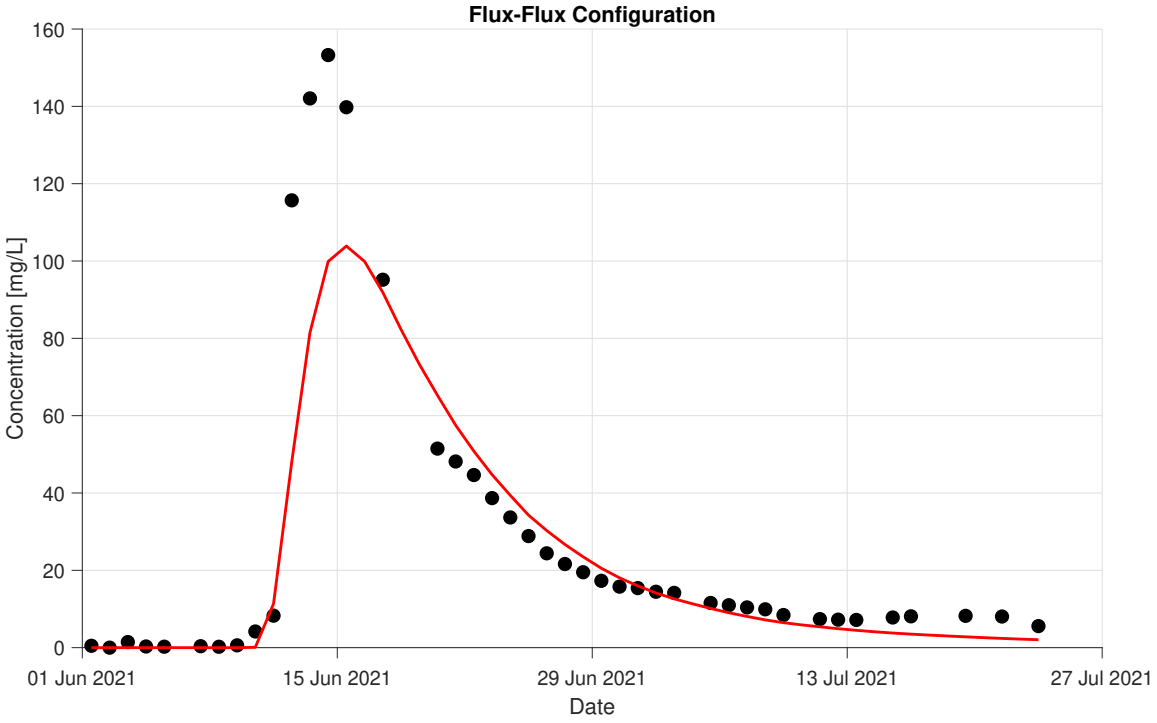


Figure 91: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).

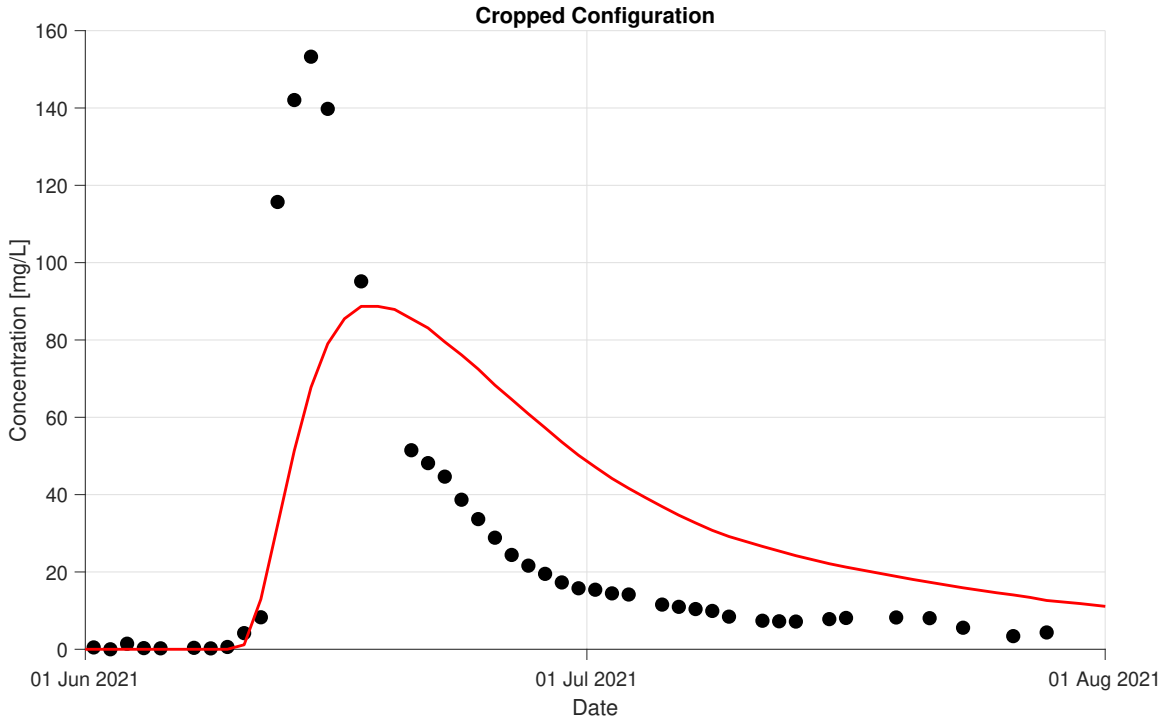


Figure 92: Concentration at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

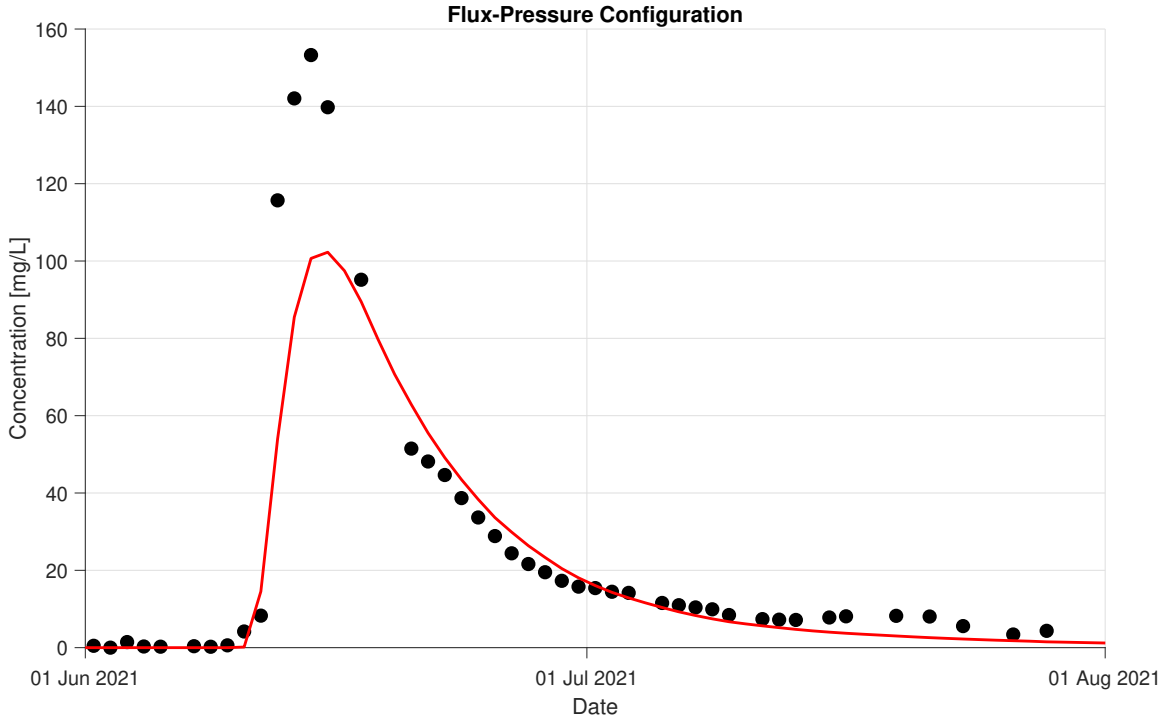


Figure 93: Concentration at $z = 50$ cm versus time (black dots are measured values and red curve is simulated values).

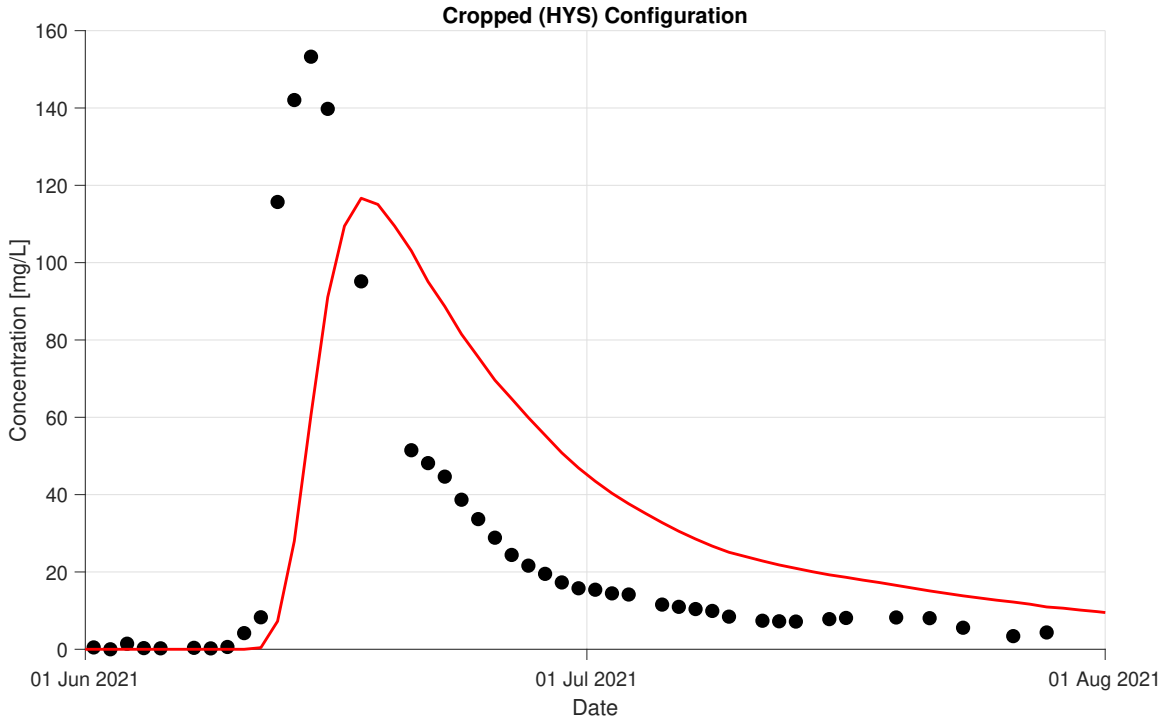


Figure 94: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).

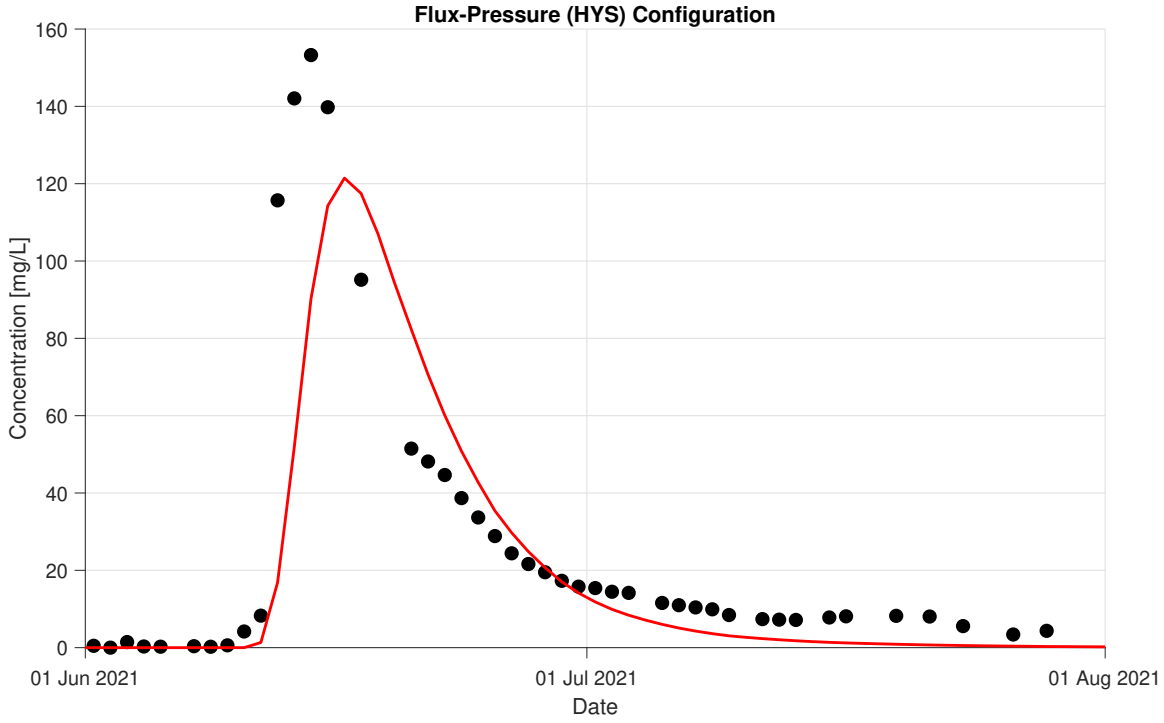


Figure 95: Concentration at $z = 50\ cm$ versus time (black dots are measured values and red curve is simulated values).