SWB model parameters calibration analysis on ISMN dataset

AUTHOR
SwissDataApps

PUBLISHED
October 28, 2024

Introduction

The Hydromontania project is aimed at providing precise irrigation recommendations to farmers based on soil water balance model (SWB) outputs. Effective water management is essential for agricultural productivity, especially in regions with varied soil and weather conditions like Switzerland. The SWB model, a core component of Hydromontania, utilizes daily precipitation and evapotranspiration data alongside a set of soil and terrain parameters to estimate available soil moisture and water dynamics.

For the model to provide accurate irrigation guidance, these soil and terrain parameters must be calibrated to reflect actual field conditions. This calibration process can be conducted globally or specifically for each location, optimizing the model's accuracy by aligning its output with real-world data. To achieve this, ground truth data—primarily soil moisture measurements—will be gathered and used to fine-tune the SWB model parameters, ensuring that model outputs closely resemble observed soil conditions.

Ground Truth Data

This second report is based on soil moisture data obtained from the International Soil Moisture Network (ISMN), a collaborative initiative aimed at establishing and maintaining a global in situ soil moisture database.

This dataset holds only location inside Switzerland, so we widened the area to all northern Europe, since the soil and weather conditions are similar to an acceptable extent to those of Switzerland.

There are 217 locations, each of them equipped with at least one soil moisture volumetric sensor. Data spans between 2015 and 2022, although it varies between stations.

Soil moisture sensors are layed out at different depths ranging between a few centimeters to almost 2 meters. For the sake of simplicity we narrowed down the data to the first 50 cm and calculated a mean of the different depths in the case there were several in this depth range.

Weather Data

Weather data was obtained from Meteoblue API historic dataset, the same source used in production in Hydromontania project.

We gathered a maximum of 3 years for each of the 217 locations.

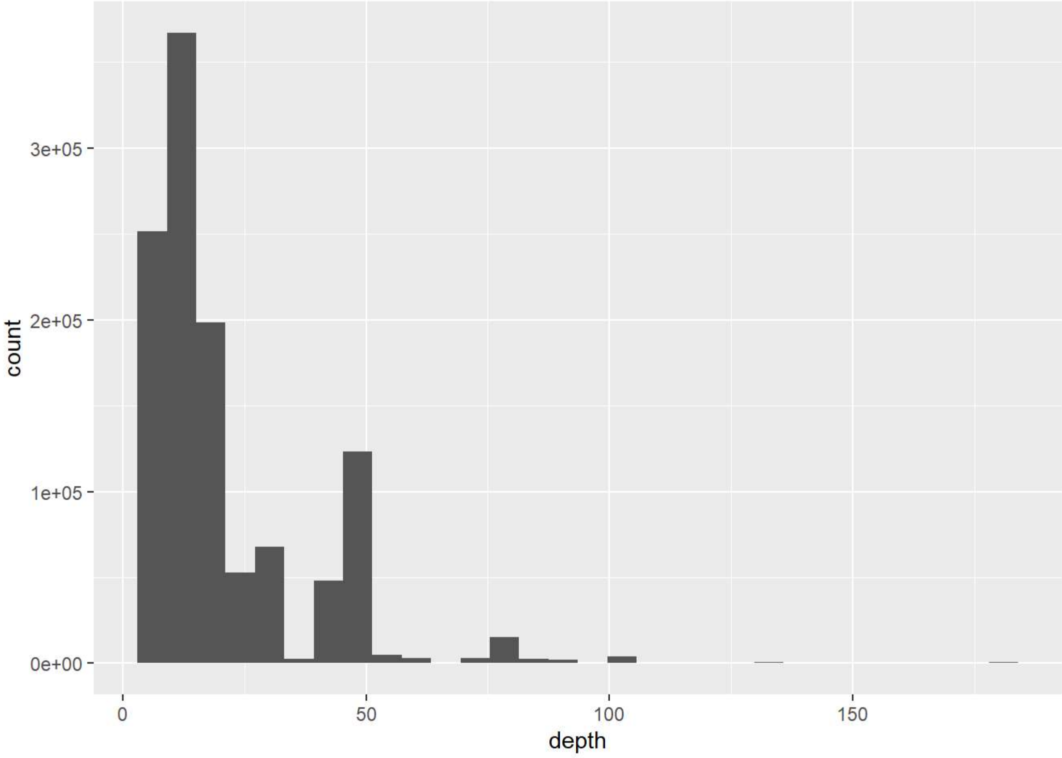
This weather data was used to estimate evapotranspiration with the same variables and algorithm than for the projct in production.

All this weather data, together with the ground-truth data is stored at the database in case it is needed for future use.

▶ Show Code

▶ Show Code

▶ Show Code



▶ Show Code

▶ Show Code

Soil Water Balance Model Runs

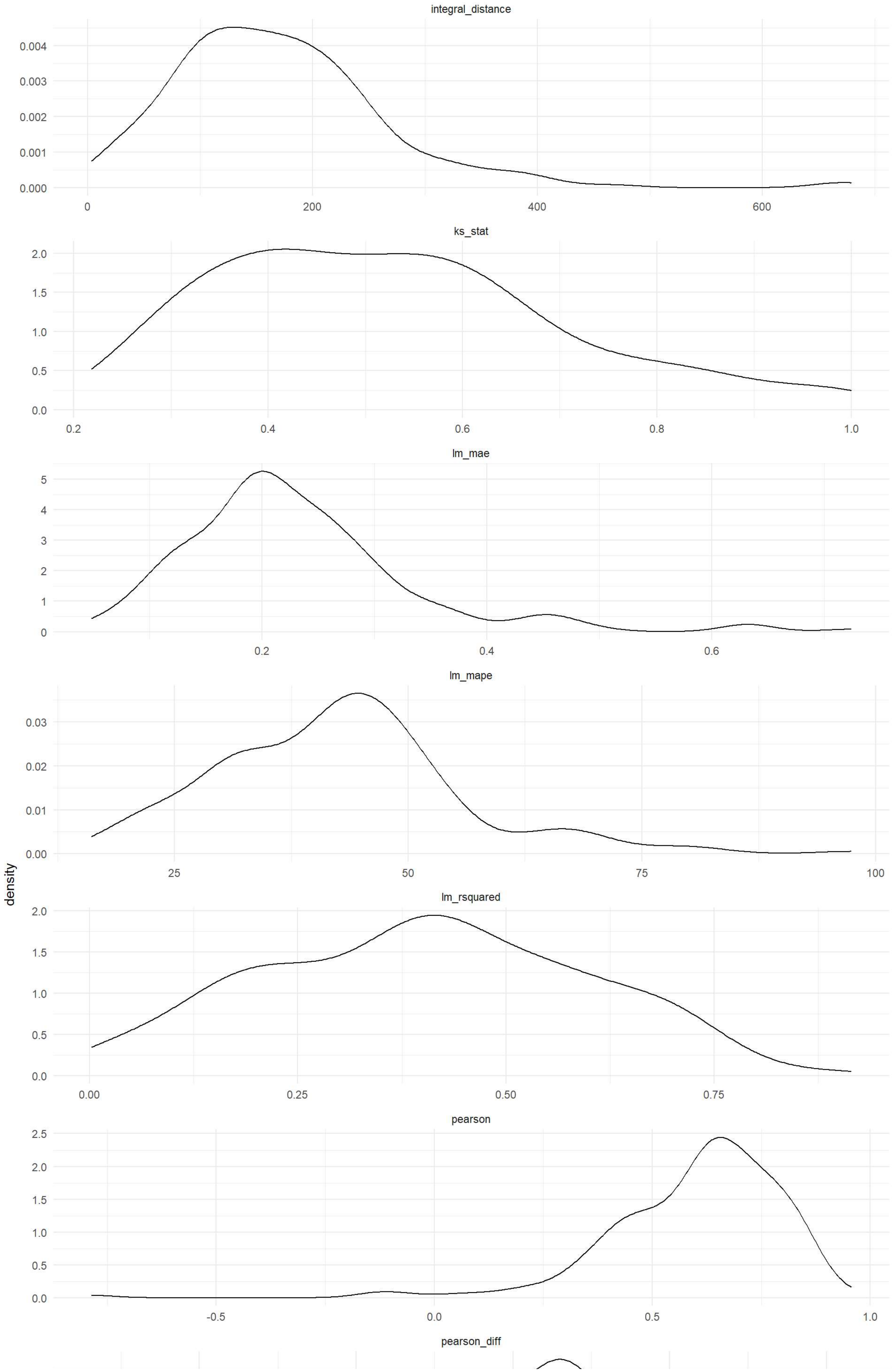
We configured the SWB model to run simulations for each location, using a grid of 20 distinct soil parameter combinations. These parameter sets represent different values of field capacities and wilting points. This grid search approach allowed us to capture a range of soil moisture behaviors and enabled initial testing of which parameter sets best matched the field data.

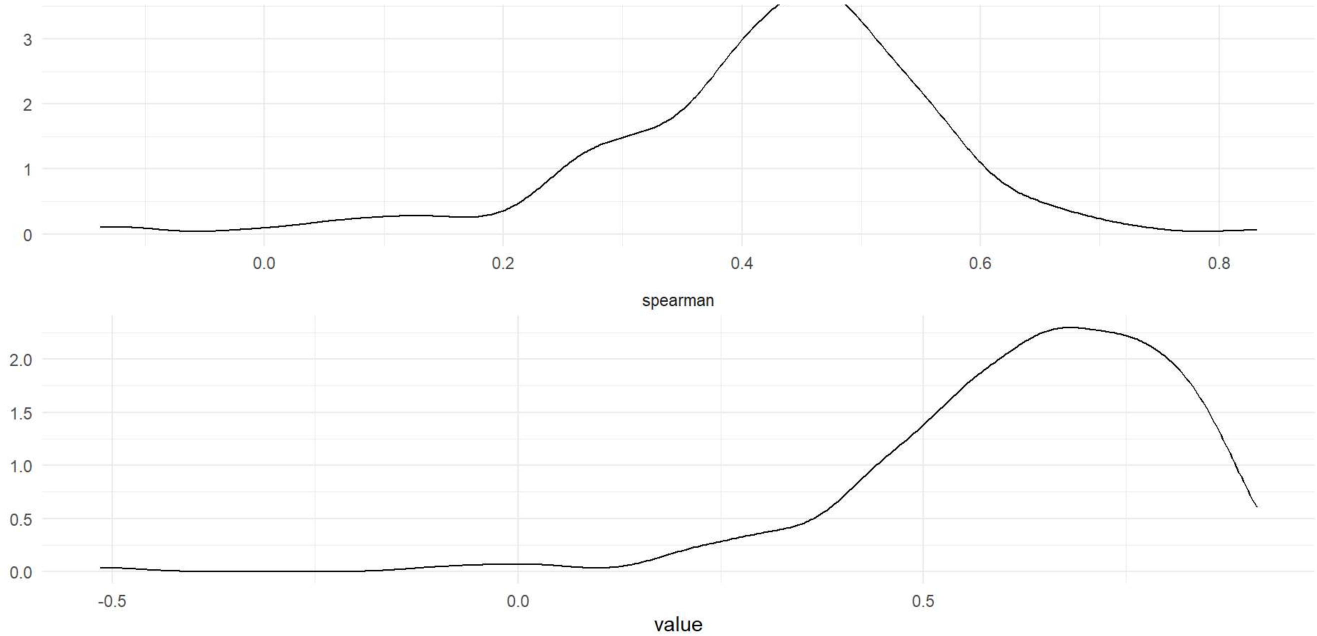
SWB model returns several variables, the most important being the amount of water available in the soil expressed in milimiters (AVAIL). To make the analysis simpler and more clear we calculated a new variable based on AVAIL, which si simply the ratio of AVAIL and the total water holding capacity assigned for the soil. This variable is called hydro_state and ranges between 0 (totally dry) and 1 (totally wet).

Model fitness analysis

We calculated several metrics for each combination of location (loc_id) and sets of soil parameters (soil_params_id) and used all of these to define which of the parameters sets was associated with a better fit between the model's hydro_state and ground-truth data.

The following plots show the distribution of the metrics obtained for the 217 analyzed locations. This first analysis shows promising values of performance (error) considering the data we are dealing with and the previous analysis results.

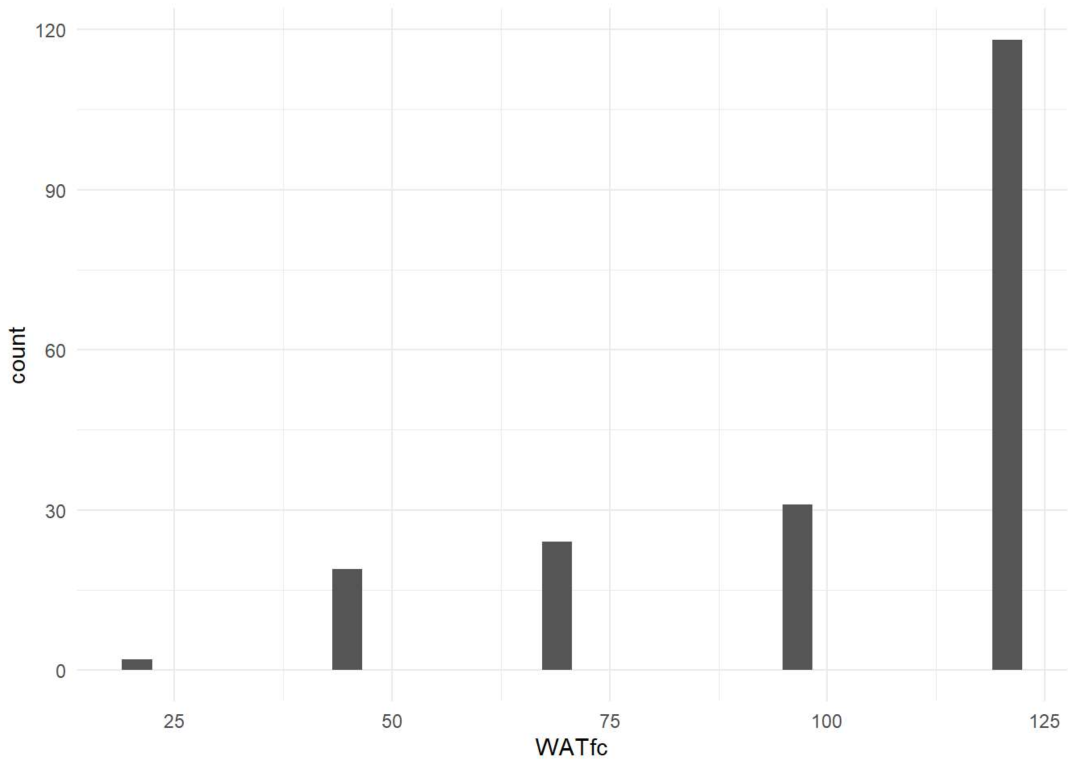

integral_distance
0.004
0.003
0.002
0.001
0.000
0
200
400
600
ks_stat
2.0
1.5
1.0
0.5
0.0
0.2
0.4
0.6
0.8
1.0
lm_mae
5
4
3
2
1
0
0.2
0.4
0.6
lm_mape
0.03
0.02
0.01
0.00
25
50
75
100
density
lm_rsquared
2.0
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
pearson
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
0.0
0.5
1.0
pearson_diff



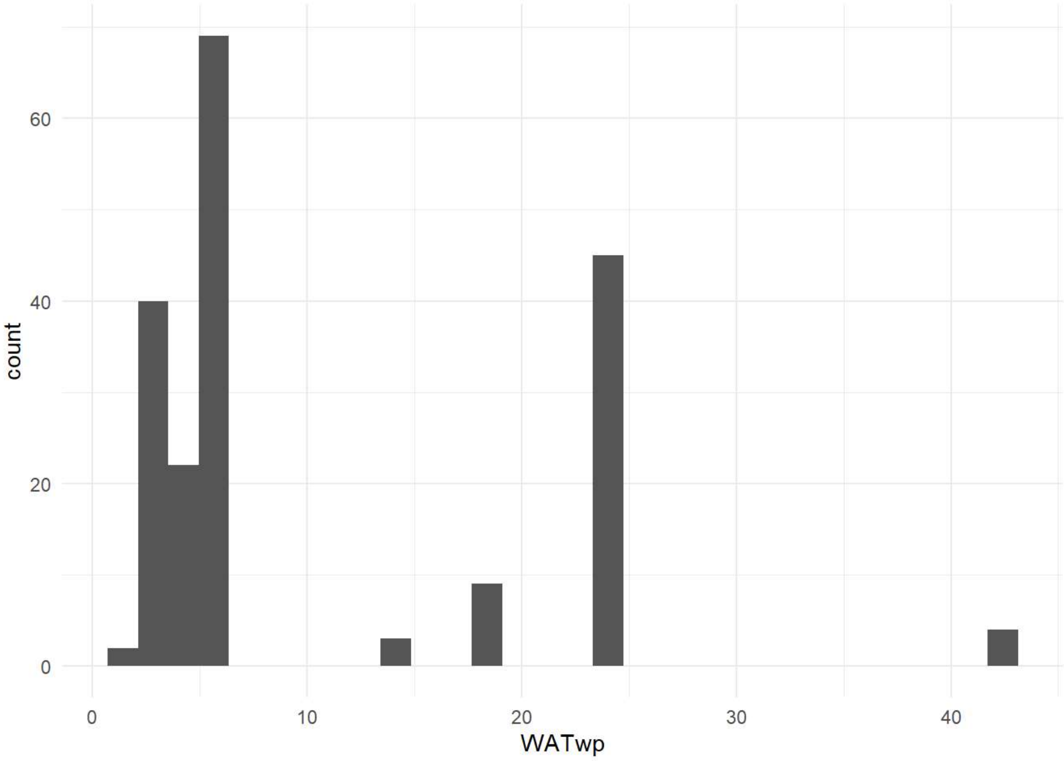
Next follows a visualization of the distribution of the best parameters for the different locations.

Wilting point is expressed in two different ways, as an absolute quantity (mm), and in relation the field capacity (percentage).

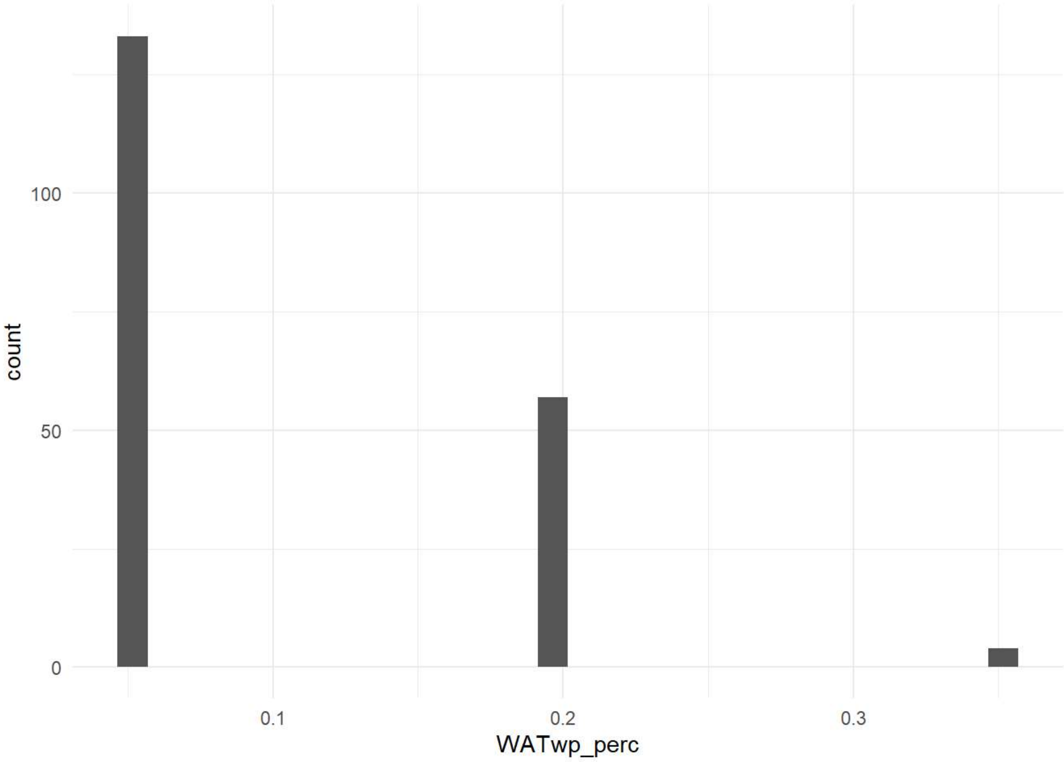
▶ Show Code



▶ Show Code



▶ Show Code



▶ Show Code

Finally for this first report version, follows a simple dashboard to verify visually and individually by location the matching of the modeled hydro_state (blue) for the best set of parameters with the ground-truth (red).

In the bottom-right corner there is a button to expand the dashboard to improve readability. The sidebar allows to resize the plots grid size and to filter the locations shown by performance metrics.