ETH zürich

Uncertainty propagation and sensitivity analysis in hydrology

Riccardo Arrigoni

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Risk, Safety & Uncertainty Quantification

Report Data Sheet

Client:	-
Contract Ref.:	Master thesis Riccardo Arrigoni
Contact:	Dr. Stefano Marelli, Prof. Dr. Bruno Sudret, Prof. Dr. Peter Molnar
Address:	Stefano-Franscini-Platz 5 CH-8093 Zürich Switzerland
ETH Project:	Master thesis Riccardo Arrigoni

Report Ref.	R-MTArrigoni-001
Title:	Uncertainty propagation and sensitivity analysis in hydrology
Authors:	Riccardo Arrigoni
Date:	July 10, 2020
Signature:	

Abstract

This project presents an application of uncertainty quantification techniques, namely polynomial chaos expansions (PCE) and sensitivity analysis (SA), in a hydrology context. More specifically, a Sobol' analysis on a subset of the parameters of the physically based and fully distributed rainfall-runoff model TOPKAPI-ETH (TE) is performed.

The analysis uses a PCE-based approach, and for a relatively low amount of model evaluations is shown to perform reasonably well for all the investigated hydrological quantities. These are a total of 30 scalar values obtained from roughly 1000 TE model runs using a 3-year simulation time span, and describe: magnitude and volume of discharge flood events, 7-day low flows, average snow cover and evapotranspiration averages. For each one a PCE was computed using the (hybrid) LAR algorithm, and afterwards the Sobol' indices were obtained from post-processing of the coefficients of said PCE.

The results obtained can help hydrologists start the calibration process while appropriately considering the importance of all the model parameters.

Keywords: Uncertainty Quantification; Rainfall-runoff models; Polynomial Chaos Expansions; Surrogate modeling; Sensitivity Analysis

Contents

Acknowledgements

I would like to thank Dr. Stefano Marelli and Prof. Dr. Bruno Sudret for giving me the opportunity to write my master thesis at their research group, and together with Prof. Dr. Peter Molnar and the hydrology group, for developing such an interesting project for me to work on.

I would also like to express my gratitude to everyone mentioned above for the expertise they provided, on both uncertainty quantification and hydrology.

A very special thank you goes to my advisor Nora Lüthen for her precious feedback, patience and guidance throughout the thesis, as well as to Sebástian Moraga for his knowledge on the hydrology technical problems.

Lastly, I would also like to thank my family and friends, for their constant support, in school as much as in life.

1 Introduction

In this section an overview of the development of hydrological rainfall-runoff (R-R) models is first presented, up until the starting point of this thesis. Afterwards, a very brief description of the uncertainty quantification software is given, followed by the outline of the thesis.

1.1 Rainfall-runoff models

Computational models are extensively used in hydrology to predict various relevant processes. Some examples are models for stochastic rainfall (or more in general, weather) generation, general circulation models (GCM) and R-R models. This last type of model is going to be the focus of the application presented in this thesis.

Due in part to the limited availability of computational resources, as well as a lesser understanding of the hydrological processes to be analyzed, R-R models were initially designed as conceptual, meaning based on theoretical considerations. The calibration of these models involves a large amount of data and curve fitting, with results not always easy to interpret. The development then continued towards the direction of empirical models. These too require a high amount of data for their calibration, and depend on mathematical expressions valid only for the catchment under observation. Physically based models provide, through solving of differential equations for the continuity of mass and momentum, the most accurate results. These models do not require such a laborious calibration process since they depend on parameters which have a direct physical interpretation. The spatial representation of the catchment to be investigated went from purely lumped models, to semi-distributed and finally to fully distributed models. In this case as well, the evolution was also allowed by an increase in computational power (Paschalis et al., 2014; Ciarapica and Todini, 2002; Devia et al., 2015).

For the case of fully distributed physically based models, an application on a large catchment, although feasible, can still be computationally very demanding due to complex interactions between the various hydrological processes. To circumvent this issue while still keeping the advantages of these kind of models, Todini (1988) suggested to integrate equations in space to obtain a scale-independent representation only based on a limited number of physically interpretable parameters.

The TOPKAPI model is the result of this approach: a model which is parsimonious in its parameters, and therefore in its calibration, and which does not lose physical meaning across spatial scales (Ciarapica and Todini, 2002). It is based on the knowledge and expertise mainly obtained from two earlier models: the ARNO model (Todini, 1996) and the TOPMODEL (Beven and Kirkby, 1979). A substantial improvement is given by the TOPKAPI-ETH (TE) model (Paschalis et al., 2014; Fatichi et al., 2015), used in this project.

Among the fields of application are land use and climate change assessment, extreme flood analysis as well as the possibility of integration with general circulation models (Liu

and Todini, 2002). Thanks to its physically based nature, it is also possible to use it in ungauged catchments.

The process of calibration involves perturbation of the model parameters within an expert-defined space, to minimize the difference between observed and simulated catchment outlet discharge values, or in other words, to maximize indicators such as the Correlation Coefficient R^2 or the Nash-Sutcliffe Efficiency Coefficient NSE (Paschalis et al., 2014).

Though parsimonious in its formulation, the model still remains very complex, and affected by a high amount of uncertainty. Since the model parameters have a direct interpretation in the physical catchment characteristics, they could theoretically be measured through fieldwork. However, this is not always the case in practice, in that some of the parameters represent physical processes either very hard to measure, or highly heterogeneous. This makes the calibration process unavoidable even for this kind of model (Liu and Todini, 2002). A more complete overview of the uncertainties within a hydrological model can be found in 2.5.

To aid in the calibration process and to therefore obtain reliable results from TOPKAPI-ETH, it is helpful to quantify the relevance of its model parameters. This is where sensitivity analysis techniques can be used.

Sensitivity analysis techniques can be divided into two main categories: local and global methods. Local sensitivity analysis can be seen as determining the impact of the input variables around a specific value. The gradient of the model response is therefore computed. Global sensitivity analysis methods quantify the impact of the uncertainty in the input variables, or combinations thereof, on the output variability. Sobol' indices belong to the global methods (Sobol, 1993; Gratiet et al., 2016).

1.2 UQLab

Aside from the TE model, the other software that will be employed during the project is UQLab (Marelli and Sudret, 2014). This is a MATLAB toolbox developed by the Chair of Risk, Safety and Uncertainty Quantification at ETH Zürich. The software finds its theoretical foundation in the uncertainty framework introduced by Sudret (2007). This states that any uncertainty quantification problem may be decomposed as follows:

- Step A: Definition of the computational model
- Step B: Identification and quantification of the sources of uncertainty in the system
- Step C: Propagation of uncertainty through the system
- Step C' (optional): Sensitivity analysis

Within the software is UQLink (Moustapha et al., 2019), which allows to connect to the TE model, as well as the modules to surrogate the TE results using polynomial chaos expansions (PCE) (Marelli and Sudret, 2019), and to perform sensitivity analysis (Marelli et al., 2019).

1.3 Outline

The TE model is described in more detail in Section 2. This includes the input to the model, the input to the uncertainty quantification framework, as well as the chosen output quantities of interest. Afterwards, in Sections 3 and 4, is the explanation of polynomial chaos expansions and sensitivity analysis. The setup used to obtain the results, from running the TE model up to the Sobol' analysis, is described in Section 5. The results are presented in Section 6, then discussed and interpreted in Section 7. This section also includes suggestions for further research on the topic.

2 TOPKAPI-ETH

As briefly mentioned in 1.1, the data is obtained with the TE model. A more detailed description is given in this section.

In the model's fully distributed version, the catchment's area is divided into a grid of cells, or computational elements, over which the differential equations are integrated.

The basin chosen for this thesis is the Kleine Emme, an alpine catchment located in central Switzerland. Its total area amounts to approximately 477 km^2, which for a chosen grid cell size of 100m x 100m, results in $47,707$ computational elements. The elevation ranges from 430 and 2300 m a.s.l. (Peleg, Skinner, Fatichi, and Molnar, Peleg et al.), with a mean value of $1,047$ m a.s.l.. The mean annual precipitation is $1,650$ mm, and the mean slope of approximately $16°$. The catchment is characterized by a high space-time variability of rainfall, is glacier-free and does not show any anthropogenic disturbances such as irrigation withdrawals or hydropower use. Furthermore, data records for both rainfall and streamflow are available. Three raingauges and weather radar coverage (both from MeteoSwiss) provide the necessary meteorological data (Paschalis et al., 2014).

The chosen time span is 1.1.2000-31.12.2002. Historical data for the catchment outlet discharge shows how this time span does not include any flood event with an extremely high return period, such as the one from 2005. The data also indicates a yearly flood event towards the month of July. Additionally, for the chosen time span, an event larger than the yearly flood occurred towards the summer of 2002.

In the following a more in-depth look at the model used is given. In particular, it is possible to distinguish three main different stages: the input to the model, the model itself, as well as the outputs it produces.

2.1 Hydrological input

First is the input representing the climate forcing, in this case made of temporal data for precipitation, air temperature and cloud cover transmissivity for the selected time span. In the context of this project, the entire hydrological input is provided by ground observations.

Also characteristic of both catchment and time span are the initial conditions

2.1.1 Climate forcing

The time series for precipitation can either come from observations, or from stochastic models. If observations are used, such as in this project, a possible source can be measurement stations from the MeteoSwiss network, weather radars, whereas if simulated data is chosen, a wide array of models is available, such as the STREAP model (Paschalis et al., 2013), cluster models or random cascade models.

The air temperature values can be provided by measurement stations (like MeteoSwiss). In order to obtain spatially distributed data, interpolation techniques such as Thiessen

polygons can be employed (Burlando, 2017).

The cloud cover transmissivity was derived from the hourly sunshine duration measurements following the empirical relation proposed by Kasten and Czeplak (1980).

2.1.2 Initial state

The initial conditions were set by running a simulation for an earlier time span finishing just before the very beginning of the one chosen for this project. The state of the simulation at that time, in this case December 31th, 1999, will yield all the necessary information for the successive three years.

Of particular importance are the saturation volume in both the upper and the lower soil layers, and the amount of snow cover.

2.2 Hydrological model

As its main assumptions, TOPKAPI-ETH represents the channel flow, the surface flow and the subsurface flow using the kinematic wave approximation, resulting in three nonlinear reservoirs (Paschalis et al., 2014). For each cell, or computational element, the partial differential equations for the continuity of mass and momentum are approximated through the use of the equations described in 2.3.1. These equations are solved analytically, as shown in Liu and Todini (2002).

The flow paths are computed from the Digital Terrain Model, starting from the so-called source cells and continuing along the downstream cells following the principle of minimum energy. Every active cell moves within a reduced 3×3 window along a tree-shaped network, receiving upstream contributions from up to three cells connected along the edges and being connected with one downstream cell. The source cells are those without any upstream contribution (Ciarapica and Todini, 2002).

From all the available model parameters, it was decided in the context of this thesis to only modify 21. These are scalar values constituting the uncertainty quantification (UQ) input, and influencing 31 variables (see Table 1, note that ETCF refers to evapotranspiration crop factor). The values given to these parameters follow a uniform distribution with ranges a and b, resulting in the notation $x_k \sim \mathcal{U}(a, b)$, with x_k as either p_i or m_j. For a more detailed description of how to generate the parameter samples, see 5.1.

Each set of parameters results in a slightly different characterization of the catchment, and therefore in a different model output.

In order to consider the spatially varying character of the catchment properties, different values must be assigned to different grid cells. However, because of the fine resolution, this would be simply unfeasible as the total amount of information required for all grid cells would soon be too demanding (except for small catchments or coarse resolutions). It is possible however, to identify areas in the catchment having similar characteristics. This results in grid cell IDs for each of the three variable categories: soil, surface and snow. Each variable will have as many entries as IDs for its respective category. For

the soil and surface variables, this produces row vectors, whereas for the snow variables only scalar values (meaning variables valid over the entire catchment). The values of the variables at this stage are then multiplied with the parameter samples to obtain the variables final values, which are written in the files containing the model variables to be changed. The procedure to generate these files according to the different realizations of the parameter set will be explained in 2.4.

The uniform distribution allows one to investigate what effect values different from the initial ones have on the results, while at the same time keeping the values in the entire range equally likely. The boundaries of these ranges were chosen based on expert knowledge.

Note that the parameters can each influence one to two variables, and that each one multiplies the respective entire affected variable(s). The changes therefore always affect the entire grid.

2.3 Output

For each of the chosen outputs, a description of the corresponding physical processes implemented in TOPKAPI-ETH is given below. Although the model offers more accurate settings for some of the outputs, it was concluded to avoid using them, since beside there not being a significant difference between the results with and without the more accurate settings, a higher precision was beyond the project scope.

2.3.1 Catchment outlet hourly discharge peaks

There is a number of reasons why flood events must be considered. For example, it is important to make sure that boundary conditions such as culverts, dikes or bridges are not violated, i.e. that the peak discharge does not exceed a specified design value. In case hydropower production is present on the catchment, a flood event is usually accompanied by a very high amount of sediments, which significantly damages the turbines.

The discharge at the catchment outlet collects the water from the three components mentioned in 2.2.

The subsurface component is a Darcy-based expression which substitutes the full Richard's equation, an assumption justified by the chosen grid cell size (Benning, 1995; Ciarapica and Todini, 2002), whereas the overland and channel flow stem from the Manning-Strickler formula. Using the subscripts s and o for subsurface flow and overland and channel flow respectively, the equations for the specific discharge (discharge per meter width in $m^2\,s^{-1}$) can be defined as follows:

$$q_s = Depth \cdot Ks \cdot s^{Exp} \tag{1}$$

$$q_o = \frac{1}{ManSurf} \cdot tan(\beta)^{\frac{1}{2}} \cdot h_o^{\frac{5}{3}} \tag{2}$$

Table 1: Input parameters

Parameter name	Prob. Input	Meaning	Variables affected
		Soil	
p_1	$\mathcal{U}(0.2, 4)$	Depth of Soil Layers	$Depth$ and $Depthlow$
p_2	$\mathcal{U}(0.3, 3)$	Theta	$Theta\_s$, $Theta\_slow$, $Theta\_r$ and $Theta\_rlow$
p_3	$\mathcal{U}(0.3, 3)$	Permeability	Ksh,Ksv, $Kshlow$ and $Ksvlow$
p_5	$\mathcal{U}(0.3, 3)$	Percolation Parameter	$Exph$,$Expv$, $Exphlow$ and $Expvlow$
		Snow	
p_6	$\mathcal{U}(0.5, 2)$	Shortwave Radiation	SRF
p_7	$\mathcal{U}(0.5, 2)$	Temperature Factor	TF
p_8	$\mathcal{U}(0, 4)$	Threshold Air Temperature	$PrecSF$
p_9	$\mathcal{U}(-1, 4)$	Threshold Snow Melt	Tt_Snow
		Surface	
p_4	$\mathcal{U}(0.2, 4)$	Surface Roughness (Manning)	$MannSurf$
m_1	$\mathcal{U}(0, 5)$	ETCF January	Jan_ETCF
m_2	$\mathcal{U}(0, 5)$	ETCF February	Feb_ETCF
m_3	$\mathcal{U}(0, 5)$	ETCF March	Mar_ETCF
m_4	$\mathcal{U}(0, 5)$	ETCF April	Apr_ETCF
m_5	$\mathcal{U}(0, 5)$	ETCF May	May_ETCF
m_6	$\mathcal{U}(0, 5)$	ETCF June	Jun_ETCF
m_7	$\mathcal{U}(0, 5)$	ETCF July	Jul_ETCF
m_8	$\mathcal{U}(0, 5)$	ETCF August	Aug_ETCF
m_9	$\mathcal{U}(0, 5)$	ETCF September	Sep_ETCF
m_{10}	$\mathcal{U}(0, 5)$	ETCF October	Oct_ETCF
m_{11}	$\mathcal{U}(0, 5)$	ETCF November	Nov_ETCF
m_{12}	$\mathcal{U}(0, 5)$	ETCF December	Dec_ETCF

where Eq. 1 refers to the specific discharge in both soil layers, horizontally and vertically. In this equation, $Depth$ is the soil thickness (in m), Ks the hydraulic conductivity at saturation (in $\mathrm{m\,s^{-1}}$), s the soil moisture (in the $[0, 1]$ range) and Exp is the percolation parameter (dimensionless). In Eq. 2, $MannSurf$ is the Manning surface roughness (in $\mathrm{m^{-\frac{1}{3}}\,s}$), β is the slope angle and h_o the water depth over the soil surface (in m).

From this quantity, it was chosen to consider the extreme values, i.e. the flood events. A number of peak discharges was first selected in a TE standard run, meaning using a set of parameters chosen within the input space, and representative for it. These same events were then identified in each of the other TE runs (the ones with different values of the model parameters). Due to the short time span of the simulation, methods such as the annual maximum flood are not particularly suited. Instead, the sampling technique for the peaks over threshold (POT) method was employed (Lang et al., 1999). This allowed to obtain a longer data series, made of a number of independent flood events.

The POT method specifies a number of criteria to facilitate the choice of a reasonable threshold. For this catchment however, it is known through expert knowledge that the 6 largest events take place during the summer. Adding another event (in this case happening in early spring) provides a seasonal perspective to the analysis. This results in a threshold of 165 m^3/s, or 1.24 mm/h. These units of measurement refer to the discharge collected at the catchment outlet and distributed over the entire catchment area respectively.

The procedure first starts by finding all the peaks above the threshold in the TE standard run. The peaks found are then checked for independence. To do this, Lang et al. (1999) specifies two conditions:

$$\theta > 5\,\text{days} + \ln(A_{sq\,mi}) \tag{3}$$

$$X_{min} < (3/4)\text{min}[Q_1, Q_2] \tag{4}$$

where θ indicates the number of days between the two peaks, $A_{sq\,mi}$ the catchment area in square miles, X_{min} the minimum discharge between the peaks, and Q_1 and Q_2 the two peak discharges.

Should one or both conditions not be satisfied, the peaks belong to the same flood event.

It must be mentioned that Lang et al. (1999) applied these criteria to the discharge on the daily scale, whereas here they refer to hourly values. However, the application to daily values yielded the same results.

Since the initial search for peaks follows the time span (the peaks found are already sorted according to time), a possible issue can arise where an independent peak is found in a secundary peak within the flood event, i.e. too early. Because of this, it is necessary to make sure that the independent peak is actually the maximum value within the event. Therefore, after finding an independent peak and comparing it to the next candidate peak, the following condition can be formulated: should the candidate peak violate Eq. 3 but be within a reasonable time distance (chosen for this project as 2 days) and at the same time be bigger than the independent peak, then this candidate peak becomes the new independent peak. This ensures that the flood event is climbed to the very top while still keeping the same number of independent peaks.

The independent peaks found in the standard run occur at the times shown in Table 2.

Table 2: Peak times in standard TE run

Peak	1	2	3	4	5	6	7
Time	6 Aug '00	21 Mar '01	15 Jul '01	9 Sep '01	7 Jun '02	15 Jul '02	11 Aug '02

At this point, the times at which the independent realizations occur are recorded, and intervals of fixed length are built around them. The duration of each interval is fixed at the minimum duration of an independent flood event as given in Eq. 3. For a catchment area of around 477 square kilometers, this amounts to 10 days.

Each TE run is then searched in all of these intervals to identify the corresponding flood event, where every time the peak discharge, as well as its volume are computed.

The magnitude of a peak is intuitively given by the maximum value in the interval. To compute the volume, the time series is integrated over the respective interval.

2.3.2 Catchment outlet 7-day low flows

The daily discharge can be computed through aggregation of the hourly discharge. The aggregation can be done as shown below:

$$Q_d = \frac{1}{24}\sum_{i=0}^{23} Q_h \tag{5}$$

From this time series, the 7-day low flow is observed. This quantity describes the lowest mean daily discharge in a 7-day period. It defines the flow conditions during dry weather, the availability of habitats along the river network, the irrigation during the warm season, and in case of hydropower production, it regulates the environmental flow release.

In each TE run both the summer and winter 7-day low flows are going to be computed, as they are the result of different physical processes. The centers of the 7-day intervals during which the low flows are taking place are given in Table 3.

Table 3: 7-day low flow times in standard TE run

7-day low flow	summer	winter
Time	26 Aug '01	25 Dec '01

The intervals in which to search for these low flows in the TE runs last 30 days, and are centered around the times in Table 3.

2.3.3 Snow cover

As one of the sources of water to both discharge and evapotranspiration, the snow cover is computed from the results of the enhanced temperature-index (ETI) model (Pellicciotti et al., 2005; Rimkus, 2013). At the grid cell scale, the amount of snow is obtained through a mass balance involving solid precipitation, snow melt (with the aforementioned ETI model) and gravitational snow redistribution from an upstream cell.

For each grid cell, the snow melt M (in mm h^{-1}) is obtained from the following equation:

$$M = \begin{cases} TF \cdot T(t) + SRF \cdot R_{\downarrow}(t) \cdot (1 - \alpha) & T(t) \geq Tt\_Snow \\ 0 & T(t) < Tt\_Snow \end{cases} \tag{6}$$

with TF as the temperature factor (in mm h^{-1} °C^{-1}), T as the air temperature (in °C), SRF as the shortwave radiation (in m^2 mm W^{-1} h^{-1}), $R_{\downarrow}$ as the global irradiance (in W^{-1} m^{-2}), α_i as the snow or ice albedo (dimensionless) and Tt as the threshold temperature for the onset of snow melt (in °C).

From this time series, the mean value of each TE run is taken, resulting in the average snow cover. Thanks to this quantity it is possible to determine which parameters control the snow cover variability.

Since the model returns the number of snow covered grid cells at any given time during the simulation, it is first necessary to divide the values by the total amount of grid cells in the catchment, so as to compute the mean of a time series ranging from 0 to 1.

2.3.4 Total actual evapotranspiration

Evapotranspiration is one of the main components of the hydrological budget as it plays a key role in the soil moisture volume during the warm season. It is composed of two processes, evaporation and transpiration. Assuming an infinite supply of water, it is possible through the Penman-Monteith equation to compute the potential evapotranspiration (Allen et al., 1998; Monteith, 1965). This approach represents the combination of the energy and mass transfer methods (Burlando and Fatichi, 2018).

A simplified version of the Penman-Monteith equation, and also one implemented in the TE model, is given by the Priestley-Taylor equation (Priestley and Taylor, 1972). To obtain this, the model offers two approaches after Makkink (see below).

Another correction can be made to account for the different land-uses on the catchment. This is done by using the crop factors from 2.2:

$$ET = ETCF \cdot ETP \tag{7}$$

where ETP is the potential (meaning not limited by the water supply in the soil) evapotranspiration calculated for a reference crop (often grass). The two approaches after Makkink read as follows (Rimkus, 2013; Jacobs and De Bruin, 1998):

Makkink (A):

$$ETP = ETCF \cdot \frac{s}{s+\gamma} \cdot \left(\frac{c_1 \cdot I_{G\downarrow}}{L} + c_2 \right) \tag{8}$$

where s is the slope of the saturation water vapour-temperature curve at air temperature, γ is the psychrometric constant, $I_{G\downarrow}$ is the global irradiance in $\mathrm{W\,m^{-2}}$, L is the specific heat of vaporization in $\mathrm{W\,m^{-2}\,d\,mm^{-1}}$, and c_1 and c_2 are empirical constants.

A simplification of this expression is given in the next approach.

Makkink (B):

$$ETP = ETCF \cdot \frac{s}{s+\gamma} \cdot \frac{c_1 \cdot I_{G\downarrow}}{L} + c_2 \tag{9}$$

To quantify the impact these factors have on the process, the data is separated into 12 time series per TE run, each containing the evapotranspiration values for the respective month throughout the time span. The time series are then averaged to obtain the monthly means of the evapotranspiration.

It is also important to find out which parameters control the overall variability of the process other than the monthly crop factors. Since evapotranspiration only shows significant values during the warmer months of the year, the average of the months from April to September is going to be analyzed.

This quantity is computed through averaging of the monthly means described above for the aforementioned range.

2.3.5 Further available outputs

Although not the focus of this thesis, the model can also return spatial outputs in the form of maps. These can describe quantities such as the total actual evapotranspiration, the snowmelt or the channel discharge. Other outputs can also be additional time series, for example the infiltration, the heights of the snowpack and of the channel flow, or the albedo.

Furthermore, aside from the time series showing variables aggregated over the entire catchment, it is also possible to save the time series (in the very same representation as in the main output file, but with a different name) at up to 500 grid cells. This allows e.g. to check for local issues (localized scarcity in water supply, despite not being problematic at the catchment scale).

2.4 Running the model

In order to execute properly, the model needs a certain amount of information. Among them are the following:

- The executable file (to be run on Linux): topkapi
- The catchment specific files (named after the catchment itself)

  - The configuration file: KEMME.tpk
  - The spatially distributed catchment properties: KEMME.tes and KEMME.cel
  - The Digital Terrain Model (DTM): KEMME.dem

- The file allowing to recognize on which catchment to run the simulation: fiume.txt
- The time series for precipitation, air temperature and cloud cover transmissivity: files in .csv and .csv.xyz format for each quantity
- The initial state file: 199912310000.stt

The three time series listed above represent the hydrological input to the TE model described in 2.1.1, and are given as fixed for all the simulations, along with the rest of the files in the list. However, each simulation is run with a different set of model parameters according to Table 1, which in turn represent the input to the uncertainty quantification framework.

The model can be run on any Linux machine. For this project, the remote cluster Euler was used, with each simulation requiring one core to compute. Despite this, certain limitations regarding how much RAM and scratch space each core can use make it necessary (for this study) to specify the resource requirements needed for each job. More precisely, 4 GB of both RAM and local scratch space are requested.

In order to run simulations with different parameter values as described in 2.2, the following can be added to the actual execution line, which also allows to change the output location as well as the initial state:
`./topkapi -I=CalibrationDirectory/Calibration.tpk -L=OutputDirectory/ -o=OutputDirectory/ -s=199912310000.stt`.
with `./topkapi` as the Linux executable, and `Calibration.tpk` as the file containing the model variables to be changed. This procedure could theoretically be extended to any of the variables present in KEMME.tpk. In case no `Calibration.tpk` file is provided, the variables values are taken from the configuration file.

2.5 Uncertainties in hydrology

In the case of a hydrological model, uncertainties are due to a wide range of causes, located both in the model, and in the input to it (Burlando, 2017).

Possible sources of uncertainty in the input are due to:

- Lack of data
- Climate variability
- Spatial variability of meteorological forcing
- Measurement errors

On the other hand, model uncertainties can be the result of:

- Poor knowledge of the physical system
- Spatial heterogeneities and anisotropies
- Model approximations
- Model parameters

The result is a model output that is also uncertain. Since it was chosen for this project to fix the meteorological input data, the focus is directed towards the uncertainties due to the model parameters.

3 Polynomial chaos expansions

When dealing with a costly computational model such as the TE model, the quantification of the uncertainties in its parameters can require a high amount of evaluations. To circumvent this issue the use of surrogate modeling techniques such as polynomial chaos expansions can prove to be advantageous.

3.1 General definition and polynomial basis

Defining the PCE of a finite variance computational model can be done with the following expression:

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \cdot \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{10}$$

where $\boldsymbol{X} \in \mathbb{R}^M$ is the input random vector, M the number of input variables, and $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ the matrix of multivariate polynomials, defined as the tensor product of the univariate polynomials:

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \overset{\text{def}}{=} \prod_{i=1}^{M} \psi_{\alpha_i}^{(i)}(x_i) \tag{11}$$

The polynomials must be orthogonal with respect to the input probability distribution:

$$\langle \psi_j^{(i)}(x_i), \psi_k^{(i)}(x_i) \rangle \overset{\text{def}}{=} \int_{\mathcal{D}_{X_i}} \psi_j^{(i)}(x_i) \times \psi_k^{(i)}(x_i) f_{X_i}(x_i) dx_i = \delta_{jk} \tag{12}$$

where:

$f_{X_i} =$ marginal probability density functions

$$\delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

For the case of uniformly distributed variables, this results in Legendre polynomials (Xiu and Karniadakis, 2002). A table showing the polynomial basis of each distribution can be found in Marelli and Sudret (2019).

The other terms in Eq. 10 are the coefficients $y_{\boldsymbol{\alpha}}$, These are used to characterize the surrogate so as to resemble the model response, and will be explained in 3.2.

3.1.1 Truncation

The general definition of a PCE given in Eq. 10 leads to an infinite series. A property is that this series will converge to the true output. However, for practical applications it must be truncated to a manageable form. To do this several schemes are available. One

of the most common is the total degree truncation scheme, where all PCE terms having a total polynomial degree not exceeding p are considered:

$$\mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : |\boldsymbol{\alpha}| \leq p\} \tag{13}$$

with:

$$\text{card}\, \mathcal{A}^{M,p} \equiv P = \binom{M+p}{p} \tag{14}$$

Due to considering every interaction between the input variables, this scheme often results in a large amount of PCE terms. Furthermore, not all PCE terms contribute in the same way to the accuracy of the surrogate, with low order interactions often as the most significant.

To solve this problem while still considering the most important terms, several schemes have been proposed. Among them are the maximum interaction and the hyperbolic truncation schemes.

Chosen for this project is the hyperbolic truncation scheme, which through the use of the q-norm discourages the PCE basis from retaining high-degree terms in high order interactions.

In practice, the basis can be built with the following expression:

$$\mathcal{A}^{M,p,q} = \{\boldsymbol{\alpha} \in \mathcal{A}^{M,p} : \|\boldsymbol{\alpha}\|_q \leq p\} \tag{15}$$

with:

$$\|\boldsymbol{\alpha}\|_q = \left(\sum_{i=1}^{M} \alpha_i^q\right)^{1/q} \tag{16}$$
$$q \in (0, 1]$$

This can be combined with the maximum interaction scheme to further reduce the basis. This scheme sets a maximum number r of non-zero entries in the multi-indices comprising the basis:

$$\mathcal{A}^{M,p,r} = \{\boldsymbol{\alpha} \in \mathcal{A}^{M,p} : \|\boldsymbol{\alpha}\|_0 \leq r\} \tag{17}$$

with:

$$\|\boldsymbol{\alpha}\|_0 = \sum_{i=1}^{M} \mathbf{1}_{\{\alpha_i > 0\}} \tag{18}$$

3.2 PCE Coefficients - Ordinary least-squares regression

The PCE coefficients can be computed through numerous techniques. Two non-intrusive and commonly used methods are projection and ordinary least-squares regression (OLS). In the case of OLS, the coefficients are found by representing the output as the sum of

the truncated PCE and a residual.

$$Y = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \cdot \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) + \varepsilon_P \tag{19}$$

Using a $N \times M$ sample $\boldsymbol{\mathcal{X}}$ (also called experimental design (ED)) of the input random vector, as well as the corresponding N model realizations $\mathcal{Y}$, a best-fit estimate of the coefficients is found so as to minimize that residual. This can be done with the following equation:

$$\hat{\boldsymbol{y}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\mathcal{Y} \tag{20}$$

where $\boldsymbol{A}$ corresponds to the evaluations of the polynomials at the ED points, and $\mathcal{Y} = \{y^{(1)}, ..., y^{(N)}\}^{\mathrm{T}}$ the model realizations on the input $\boldsymbol{\mathcal{X}} = \{\boldsymbol{x}^{(1)}, ..., \boldsymbol{x}^{(N)}\}^{\mathrm{T}}$.

In order to ensure sufficient stability, the size of the ED is often chosen between two to three times the amount of coefficients to be estimated.

The sampling method used to generate the ED has a direct influence on the PCE coefficients, and can be chosen among several alternatives, such as Monte Carlo sampling (MCS), latin hypercube sampling (LHS), Sobol' or Halton Sequence. For this project, LHS was chosen (McKay et al., 1979), since it is both random and space-filling. To guarantee the optimality of the sample, the method is repeated a number of times, with the realization having the maximal minimal distance between any two points being chosen (Fajraoui et al., 2017).

In the context of this project, least squares regression will be used together with the LAR algorithm (see 3.4.1).

3.3 Leave-one-out cross-validation error

To quantify the accuracy and predictive ability of a surrogate model, one commonly used criterion is the generalization error. For the input random vector $\boldsymbol{X}$ and the model response Y, its definition reads:

$$e_{gen} = \mathbb{E}\left[(\mathcal{M}(\boldsymbol{X}) - \hat{\mathcal{M}}(\boldsymbol{X}))^2\right] / \mathrm{Var}[Y] \tag{21}$$

If no validation set is available, one way to estimate e_{gen} is through the leave-one-out cross-validation error e_{LOO}:

$$e_{\mathrm{LOO}} = \frac{\sum_{i=1}^{N} (\mathcal{M}(\boldsymbol{x}_i) - \hat{\mathcal{M}}_{\setminus i}(\boldsymbol{x}_i))^2}{\sum_{i=1}^{N} (\mathcal{M}(\boldsymbol{x}_i) - \hat{\mu}_{\mathcal{Y}})^2} \tag{22}$$

where $x_i \in \boldsymbol{\mathcal{X}}$.

According to the above formula, a metamodel $\hat{\mathcal{M}}_{\setminus i}$ is built on an experimental design missing one point, and a comparison is made between the observed value and the predicted value for that same point. The procedure is then repeated for each point in the ED.

However, if the PCE is built using OLS, the estimation considerably simplifies, as the calculation of N metamodels can be avoided entirely (Marelli and Sudret, 2019):

$$e_{\text{LOO}} = \sum_{i=1}^{N} \left(\frac{\mathcal{M}(\boldsymbol{x}_i) - \hat{\mathcal{M}}(\boldsymbol{x}_i)}{1 - h_i} \right)^2 \Big/ \sum_{i=1}^{N} (\mathcal{M}(\boldsymbol{x}_i) - \hat{\mu}_{\mathcal{Y}})^2 \tag{23}$$

with $\boldsymbol{h}$ as the column vector given by:

$$\boldsymbol{h} = \text{diag}\big(\boldsymbol{A}(\boldsymbol{A}^{\text{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\text{T}}\big) \tag{24}$$

where $\boldsymbol{A}$ refers to the same matrix from Eq. 20.

The UQLab software offers a correction to this error measure to make sure it does not get underestimated. Such a situation can occur especially for the case of small experimental designs. According to Marelli and Sudret (2019), the correction has the following form:

$$e^*_{\text{LOO}} = e_{\text{LOO}} T(P, N) \tag{25}$$

where P indicates the number of regressors with non-zero coefficient, which increases the correction factor, and N the size of the experimental design. For a large N, $T(P, N) \to 1$.

3.4 Least angle regression

Although truncation schemes such as the one chosen and described in 3.1.1 tend to favour sparsity compared to the standard truncation scheme, it is possible to further reduce the basis. Two main categories of algorithms in the PCE context are basis-adaptive PCE and sparse PCE. In the first strategy, starting from a small initial basis, elements are gradually added until the corresponding PCE meets the required generalization error. On the other hand, sparse PCE selects to have non-zero coefficients only a few basis elements causing the greatest impact on the model response. The properties of both the strategies described above can be combined into a single method. One example is the least angle regression (LAR) algorithm (Blatman and Sudret, 2011)(Marelli and Sudret, 2019), where regressors are moved iteratively from a candidate set to an active set based on their correlation with the residual. The algorithm reads as follows:

0. Initialize:
   - $y_{\boldsymbol{\alpha}} = 0$ with $\boldsymbol{\alpha} \in \mathcal{A}^{M,p,q}$
   - Active and candidate sets as $\varnothing$ and $\Psi_{\boldsymbol{\alpha}}$ respectively
   - Residual $r_0 = \mathcal{Y}$
1. Move all coefficients in the current active set towards their least-square value until their regressors are as correlated with the current residual as another regressor $\Psi_{\boldsymbol{\alpha}_j}$ in the candidate set

2. Update coefficients and residual and calculate the leave-one-out error for the current iteration e^j_{LOO}

3. Move $\Psi_{\boldsymbol{\alpha}_j}$ from the candidate to the active set

Steps 1 to 3 are repeated until the active set reaches the size $m = min(P, N-1)$, or using the early-stop criterion if e_{LOO} has increased several times in a row, then the basis resulting in the PCE with the lowest leave-one-out error is selected.

Including basis-adaptivity into this algorithm, it is possible to repeat the steps above for all polynomial degrees $p = 1, ..., p_{max}$. The candidate set of basis elements, is given by the chosen truncation scheme(s). The algorithm stops at the degree resulting in a sparse PCE whose e_{LOO} increases several times in subsequent interations.

3.4.1 Hybrid LAR

Because the LAR algorithm is defined on non-constant regressors only, the leave-one-out error cannot be calculated as in Eq. 23 (see Eq. 22 instead). This issue can be solved by adding the constant regressor to the current active basis after each LAR iteration. In this case, the basis elements are selected using standard LAR, whereas coefficients and e_{LOO} are computed with OLS.

3.5 PCE Moments

The mean value and the variance of a PCE can be calculated directly from its coefficients, using the following equations:

$$\hat{\mu} = \hat{y}_0 \tag{26}$$

$$\hat{\sigma}^2 = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A} \\ \boldsymbol{\alpha}\neq\mathbf{0}}} \hat{y}^2_{\boldsymbol{\alpha}} \tag{27}$$

3.6 Elementary effects

To see the contribution a single input variable is giving to the model response, its elementary effect can be observed. For each input variable, this is the sum of those univariate PCE terms which only involve this variable. In cases of models without significant interactions, the elementary effects allow to obtain a close representation of how each input influences the output.

For each input variable, the corresponding elementary effect EE_i, with $i = 1, \ldots, M$, can be obtained as follows:

$$\mathcal{A}^{M,p,i} = \{\boldsymbol{\alpha} \in \mathcal{A}^{M,p} : \alpha_j = 0,\ j = 1, \ldots, M,\ j \neq i\} \tag{28}$$

$$EE_i = \sum_{\boldsymbol{\alpha} \in \mathcal{A}^{M,p,i}} y_{\boldsymbol{\alpha}} \cdot \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{29}$$

The truncation scheme $\mathcal{A}^{M,p,i}$ is defined for each input variable, and only considers elements of $\boldsymbol{\alpha}$ referring to the input variable for which the elementary effect is to be computed.

3.7 Bootstrap PCE

Local error estimates for a PCE can be obtained using bootstrap resampling. A number of replications B of the experimental design are generated through resampling with substitution, as described in Marelli and Sudret (2019). For each of these replication a PCE is computed, to eventually obtain the desired error estimate.

Should the computation of B PCEs be time consuming, UQLab offers the possibility to use a sparse basis obtained on the original experimental design $\boldsymbol{\mathcal{X}}$, and only compute the PCE coefficients for each bootstrap replication. This is also the UQLab default option when using bootstrap PCE.

4 Sensitivity analysis

To find out how much the input variables (in this case the TE model parameters) influence the variability of the response, sensitivity analysis can be performed. Several ways to do that are present in the literature. For this project it was decided to use the variance decomposition with subsequent computation of the Sobol' indices.

In this technique, the share of the output's total variance due to the uncertainty in the single input variables is represented by the first-order indices, whereas higher-order indices indicate how much of the output variance is due to interactions between the variables.

This section gives a brief overview of the model and variance decomposition as well as of the PCE-based approach to compute the Sobol' indices originally proposed by Sudret (2008).

4.1 Hoeffding-Sobol' decomposition

4.1.1 Model decomposition

By using the Hoeffding-Sobol' decomposition, it is possible to divide a computational model into summands of increasing dimension. A model with M input parameters is defined:

$$Y = \mathcal{M}(\boldsymbol{X}) \tag{30}$$

To simplify, it is assumed that:

$$X_i \sim \mathcal{U}(0,1), i = 1, \ldots, M \tag{31}$$

The Sobol' decomposition is carried out as follows:

$$\mathcal{M}(x_1 \ldots x_M) = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(x_i) + \sum_{1 \le i < j \le M} \mathcal{M}_{ij}(x_i, x_j) + \ldots + \mathcal{M}_{12\ldots M}(x_1, \ldots, x_M) \tag{32}$$

In the above equation the term $\mathcal{M}_0$ is the mean of $\mathcal{M}(\boldsymbol{X})$, whereas the summands represent every possible contribution given by the input variables to the model response and are orthogonal to each other. The integrals of these summands with respect to their variables, a subset $i_1, \ldots, i_s$ of all the input variables, vanish, leading to terms with zero mean:

$$\int_0^1 \mathcal{M}_{i_1 \ldots i_s}(x_{i_1}, \ldots, x_{i_s}) dx_k = 0 \tag{33}$$

$$\text{for all} \quad k \in i_1, ..., i_s \tag{34}$$

It is shown in Sudret (2008) how the decomposition in Eq. 32 is unique, provided $\mathcal{M}(\boldsymbol{X})$ is square-integrable on $\mathcal{D}_{\boldsymbol{X}}$.

4.1.2 Variance decomposition

The model decomposition can be applied to obtain an analogous expression for the variance of the model response.

By defining the total variance of Y as:

$$D = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}^2(\boldsymbol{x}) d\boldsymbol{x} - \mathcal{M}_0^2 \tag{35}$$

and thanks to Eq. 32 as well as to the orthogonality of the summands, it is possible to decompose the total variance as follows:

$$D = \sum_{i=1}^{M} D_i + \sum_{1 \leq i < j \leq M} D_{ij} + \ldots + D_{12\ldots M} \tag{36}$$

It is also possible to define the partial variances using the following equation:

$$D_{i_1 \ldots i_s} = \int_{[0,1]^s} \mathcal{M}^2_{i_1 \ldots i_s}(x_{i_1}, \ldots, x_{i_s}) dx_{i_1} \ldots dx_{i_s} \tag{37}$$

4.2 Sobol' indices

The Sobol' indices can be computed as the ratios of partial and total variances:

$$S_{i_1 \ldots i_s} = \frac{D_{i_1 \ldots i_s}}{D} \tag{38}$$

For the case of single input variables, the first-order indices can be expressed as simply as:

$$S_i = \frac{D_i}{D} \tag{39}$$

The total contribution of an input variable to the output variance is given by the total Sobol' indices. These are computed through the sum of all the indices involving that variable:

$$S_i^T = \sum_{\{i_1 \ldots i_s\} \supset i} S_{i_1 \ldots i_s} \tag{40}$$

Obtaining the relative importance of each input variable with respect to the model response also allows to determine which variables can be considered as unimportant. After running the Sobol' analysis, this allows to fix those variables to deterministic values.

To obtain the values of the partial and total variances, various methods can be employed. One of them is to realize that a PCE is an approximation to the Hoeffding-Sobol decomposition of a model in terms of polynomials.

4.3 PCE-based estimation of Sobol' indices

By expressing the total variance in terms of the PCE coefficients, the partial variances can also be formulated using a similar notation as for the total variance:

$$\sigma(\mathcal{M}_{\mathbf{v}}(\boldsymbol{x}_{\mathbf{v}}))^2 = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A}_{\mathbf{v}} \\ \alpha\neq 0}} \hat{y}_{\boldsymbol{\alpha}}^2 \tag{41}$$

Where $\mathbf{v}$ denotes the index set of the PCE expansion terms depending only on the subset of input variables $\boldsymbol{x}_{\mathbf{v}}$.

$$\mathbf{v} \stackrel{\text{def}}{=} \{i_1, \ldots, i_k\} \subset \{1, \ldots, M\} \tag{42}$$

From the ratio between Eqs. 41 and 27 it is possible to obtain the Sobol' indices directly from post-processing of the PCE coefficients, allowing to obtain all indices basically for free.

Since the first-order indices allocate a share of the total variance to each single model parameter, their values will range from 0 to 1. The total indices will always be greater than or equal to the first-order indices, since they also contain the interaction terms, when any are present. These can be repeated among different total indices (interaction between two parameters increases value of two total indices). The sums will therefore be:

$$\sum_{i=1}^{M} S_i \leq 1 \leq \sum_{i=1}^{M} S_i^T \tag{43}$$

4.3.1 Bootstrap-based Sobol' indices confidence intervals

Through bootstrap resampling it is possible to obtain the 95%-confidence intervals of the Sobol' indices. To do this, the set of coefficients of each of the B bootstrap PCE replications is used to compute the indices, resulting in B realizations of each index. From this, the empirical quantiles can be calculated to obtain the confidence intervals.

5 Workflow and setup of the numerical experiments

As mentioned in 2.4, the simulations are run on ETH's high performance computing cluster Euler. One of the many advantages this choice offers is the possibility to run parallel computations. This section gives a general overview of how it is possible to connect the TE model to MATLAB using UQLink, and how the results in Section 6 are obtained. This also includes the settings of both the PCEs and the Sobol' analysis. A graphical representation of the procedure is given in the following scheme:

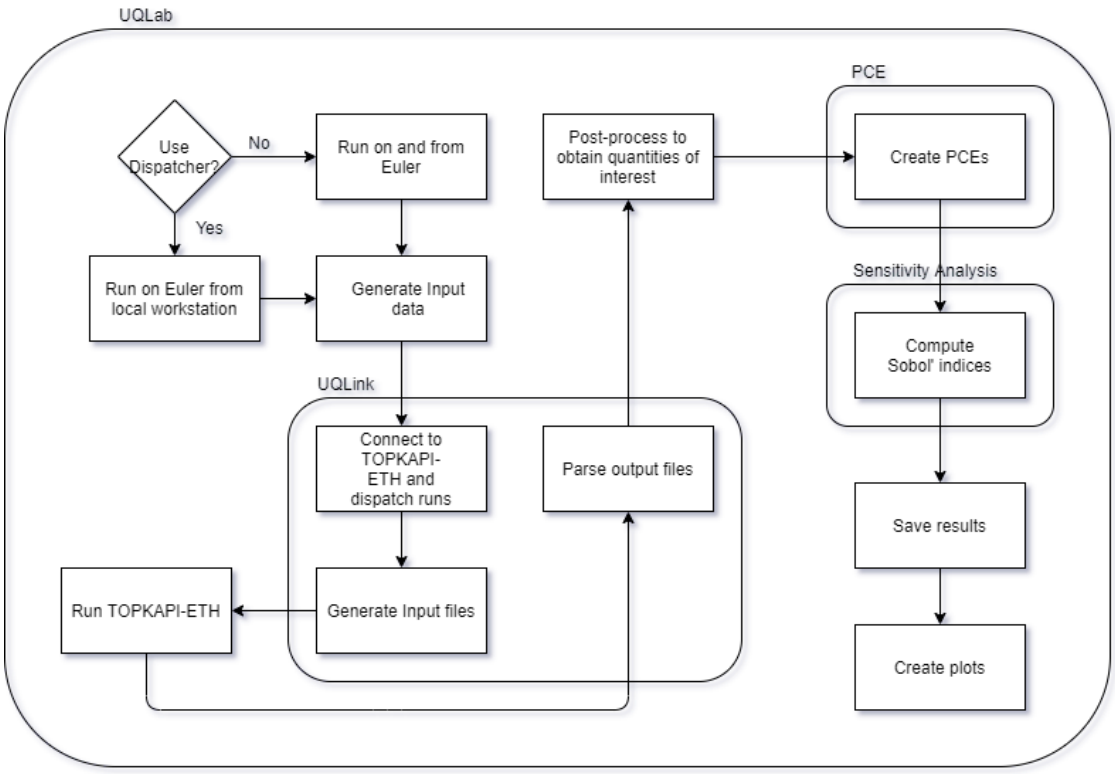


Figure 1: Setup of SA study using the computing cluster Euler and several specialized UQLab modules

As Fig. 1 shows, each step of the setup happens within a UQLab session. In case the default dispatcher is used, the main MATLAB script can be run directly from the local workstation. Else, it is necessary to manually connect to the cluster. Thanks to the Euler batch system the MATLAB script must only be executed once, with all runs within the same job.

Independently from how the cluster is reached, once there the input data consisting of the different parameter sets is generated. At this point, the module UQLink takes care of connecting to the third-party software, in this case TOPKAPI-ETH, and to dispatch the requested model runs. To do this, it creates a text file with the input values at the start of each run. The UQLink module also takes care of submitting the job and feeding the input file to the third-party software. As soon as the job is completed the UQLink parser looks for the specified output file and converts its content (more specifically only the content

of interest to the user) in MATLAB-readable variables.

This sequence is carried out for all runs. In order to keep track of each run separately, UQLink automatically appends a numeric counter at the end of the input file name, right before the extension. In case the output file has the same name (with different file extension) as the input file, the same counter will be appended. However, this is not the case of the TE model, which produces a file named 1.catchment.ts. It is still possible to assign a unique name to each of these files. This is done by copying and renaming them in the same way as the input files (once again with different file extension). UQLink will then append a counter on this renamed file.

Once all the simulations have run, it is possible to obtain the quantities of interest from the model outputs. These quantities are the subject of the next two phases, starting with the PCE module, where a polynomial chaos expansion is created for each quantity of interest.

After the PCEs have been obtained, the computation of the Sobol' indices can finally be carried out. This is possible thanks to the Sensitivity Analysis module.

5.1 Creating the PCE

The PCEs are created for a total of 30 quantities of interest, with each one having the same settings.

The initial basis is provided by the hyperbolic truncation and the maximum interaction schemes, and is then reduced using the LAR algorithm. In the truncation schemes, the q-norm was set at 0.7, whereas a maximum interaction of 2 was allowed. The maximum polynomial degree ranges from 2 to 7, thanks to the degree-adaptivity offered by UQLab. For each quantity of interest, the selected degree is the one resulting in the metamodel with the smallest e^*_{LOO}. The amount of bootstrap replications was set at 500.

To check if the PCEs do in fact converge with an increase in the ED size, the original LHS design is enriched with additional batches of points. The enrichments still follow the LHS method, so that the result is a nested LHS experimental design. The size of the initial LHS design and of each enrichment amounts to 5 times the number of PCE dimensions. The PCE convergence is then checked by sequentially adding an enrichment to the current ED.

It happened that due to some particular combination of the input parameters, the TE model failed to produce an output. This was the case for 3 runs, which were omitted from the ED. The final sizes of the experimental design steps are given in the array below:

$$N_{ED} = \{105; 209; 314; 419; 524; 628; 733; 837; 942; 1,047\}$$

The results shown in the next section were obtained using the largest ED, consisting of $1,047$ points, except for the convergence plots in Appendix C.

Since the discharge peaks and 7-day low flows, as well as the evapotranspiration monthly averages show rather low values, the corresponding metamodels returned, to a

varying extent, negative values for the lower end of the data set. Because the Sobol' analysis is a SA method based on the variance of a quantity, small fluctuations into negative values, although not physically valid, do not impact the quality of its results. This is also the reason why possible solutions such as surrogating the logarithm of the data may lead to inaccurate results: a metamodel built after the log-transform fits small values much better than larger values, which affects the variance significantly. On the other hand, a metamodel built on the actual data fits all values equally well.

In any case, since the PCEs are not used for prediction purposes, the validity of the subsequent Sobol' analysis is not impacted.

5.2 Computing Sobol' indices

The Sobol' indices are computed using the PCE-based approach discussed in 4.3 for each of the quantities of interest. From each Sobol' analysis the indices up to order 2 are obtained, as well as the total indices. The maximum order was chosen based on the limit given by the maximum interaction scheme. The confidence intervals are obtained by evaluating 500 bootstrap replications.

6 Results

6.1 Catchment outlet discharge

6.1.1 Hourly peaks

In order to give the analysis of the hourly discharge peaks described in 2.3.1 a seasonal perspective, and since except for one event the peaks always take place during the summer, the spring peak as well as the maximum summer peak, where maximum refers to the independent peak magnitudes in the standard TE run, were analyzed in more detail. These are the peaks number 2 and 6 respectively. It must also be mentioned how peak 6 is much higher than the rest of the summer peaks, suggesting a flood event with a higher return period.

Two quantities which can help characterizing the shape of a hydrograph are the peak magnitude, as well as the peak volume. For this reason, these are also the quantities shown in this peak analysis.

The following figures show the first-order indices and the difference to the total indices for the peak magnitudes and volumes. For each parameter, this difference denotes the importance of interaction terms involving that parameter. Also shown are the confidence intervals for the first-order indices, as well as (for some of the quantities shown here, refer to Appendix D for the rest) the elementary effects for the most relevant parameters. To be considered as relevant, a parameter must have a first-order index higher than 0.07. This threshold is entirely subjective, and mainly chosen so as to include as many indices as possible. Note that the value of this threshold is the same for all the examined quantities (not just the hourly peaks).

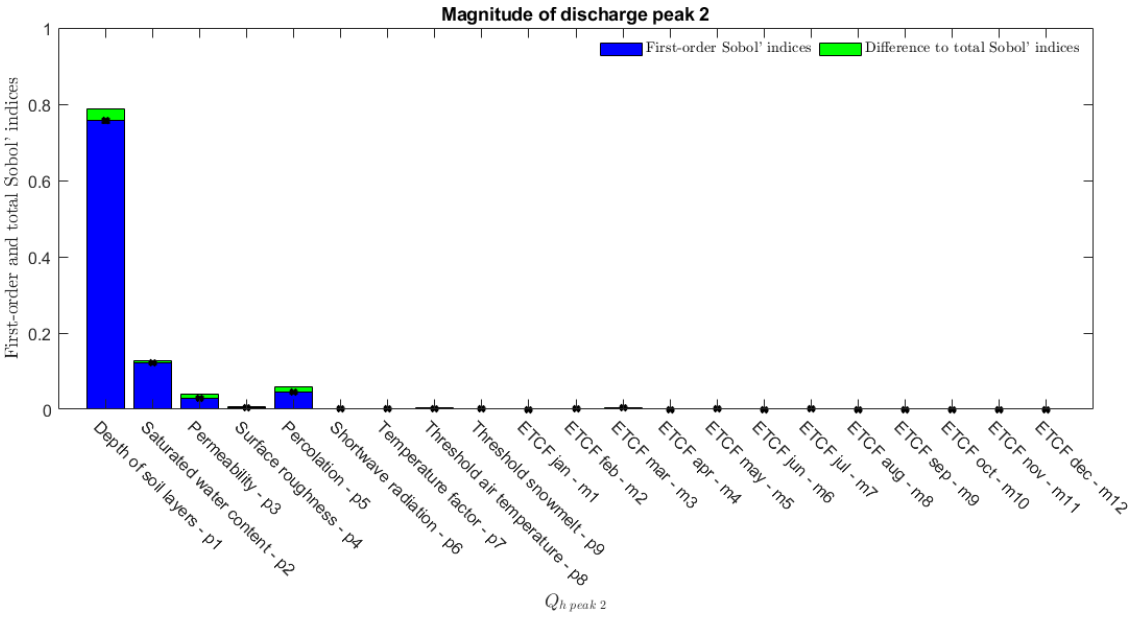


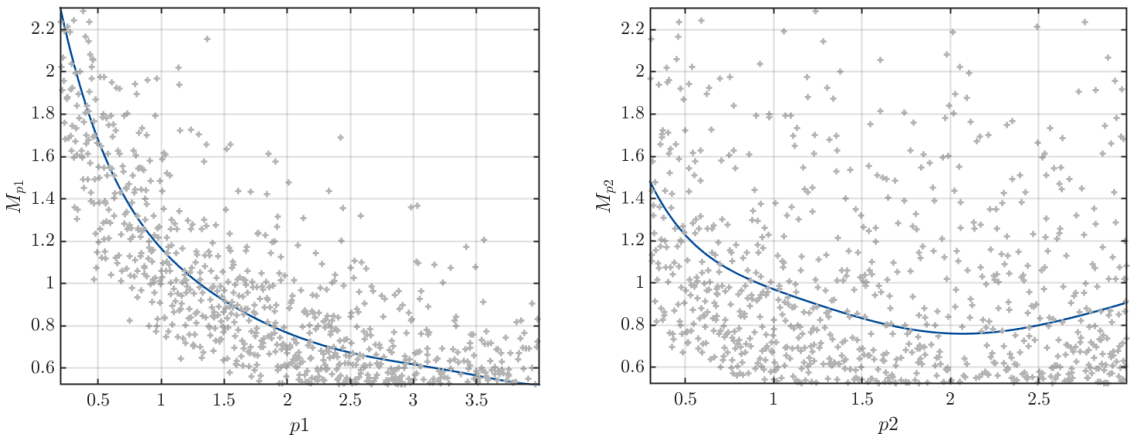
Figure 2: First-order and total indices for the *magnitude* of hourly peak 2 (spring peak)

The first-order indices for the spring peak magnitude in Fig. 2 show a strong depen-

dence on $p1$, the depth of the soil layers, with approximately 75% of the total variance. Another relevant parameter appears to be the saturated water content $p2$, although the difference in importance remains evident. A small contribution to the total variance is also given by both permeability and percolation.

The confidence intervals bounds for this peak magnitude are very narrow, indicating the obtained indices are very close to the actual true values of the PCE approximation, and the underlying PCE has essentially converged.

The second order contributions for this quantity are practically negligible.



(a) Elementary effect for $p1$

(b) Elementary effect for $p2$

Figure 3: Elementary effects for most relevant parameters of *magnitude* of hourly peak 2 (spring peak)

The elementary effects for the two most significant parameters are given in Fig. 3. It can be observed how generally for both parameters the lowest values result in the highest discharge magnitudes. For $p1$ a low value means a reduced capacity of the soil to hold water, resulting in a faster saturation, and therefore in more water in the overland and channel flow networks, which provokes a faster catchment response. A similar observation can be made for $p2$, although it must also be noted how for high values the discharge has a minimum at $p2$ around 2, and then starts to slightly increase again.

From the summer peak results in Fig. 4 it can be seen how $p1$ and $p2$ are the most relevant parameters for this peak as well, although in a slightly different proportion. A very small contribution is also given by the July evapotranspiration crop factor $m7$, which is reasonable given how peak 6 happens in that same month.

The confidence intervals bounds for the magnitude of this flood event are still quite narrow, though wider than for peak 2.

The total indices for peak 6 indicate significant interaction terms involving parameters $p1$ and $p2$, hinting at second order indices mostly involving these 2 parameters.

The elementary effects for $p1$ and $p2$ in Fig. 5 show a similar behavior as in Fig. 3, indicating how despite the seasonality varies the proportion of the indices for these two parameters, they remain the most relevant ones throughout the year.

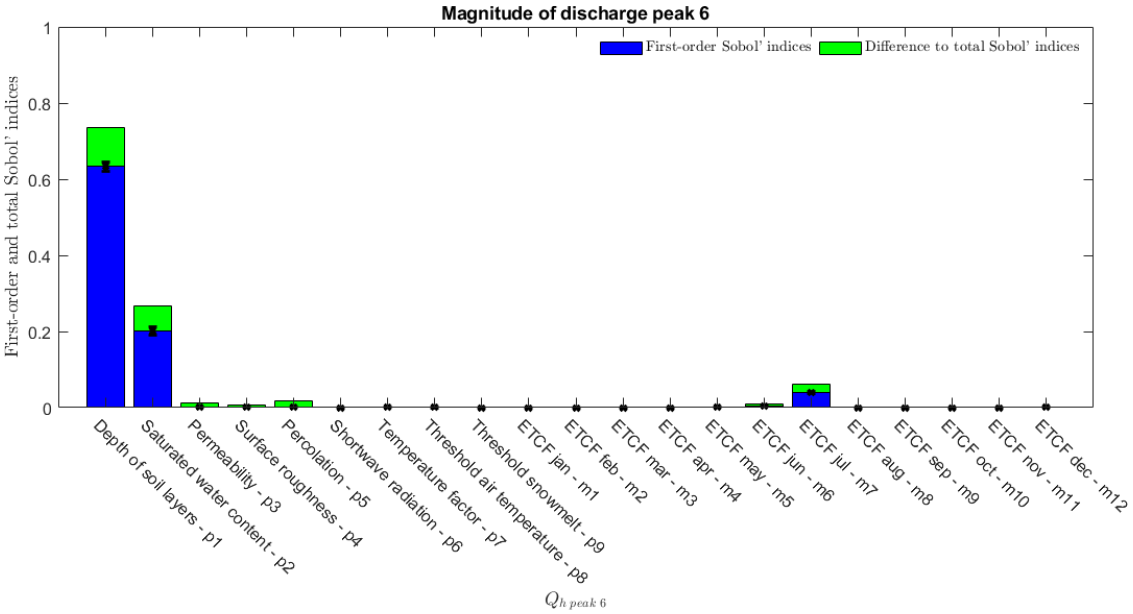
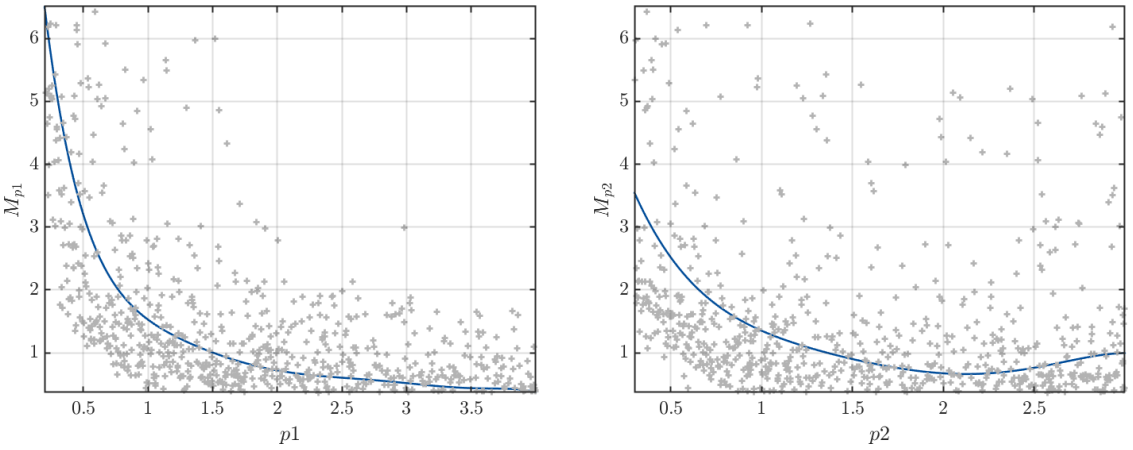


Figure 4: First-order and total indices for the *magnitude* of hourly peak 6 (summer peak)



(a) Elementary effect for $p1$

(b) Elementary effect for $p2$

Figure 5: Elementary effects for most relevant parameters of *magnitude* of hourly peak 6 (summer peak)

The first-order indices for the discharge volume on the spring peak in Fig. 6 show how, unlike the magnitude for this same peak, there is not one key parameter explaining most of the variance. Instead, 4 parameters are the most relevant, with a relative importance ranging from 15% to 30% of the total variance, and with very narrow confidence intervals. These 4 parameters include $m3$, which can once again be justified seeing how it concerns the month in which the peak is taking place, as well as $p1$, $p2$ and $p3$. The threshold temperature for precipitation state transition $p8$ also appears in the plot, although with minor importance. However, this is plausible, given how the air temperature at the time of the spring peak begins to rise again after the colder months. During the days preceding the spring peak, the air temperature is still relatively low, possibly very similar to $p8$.

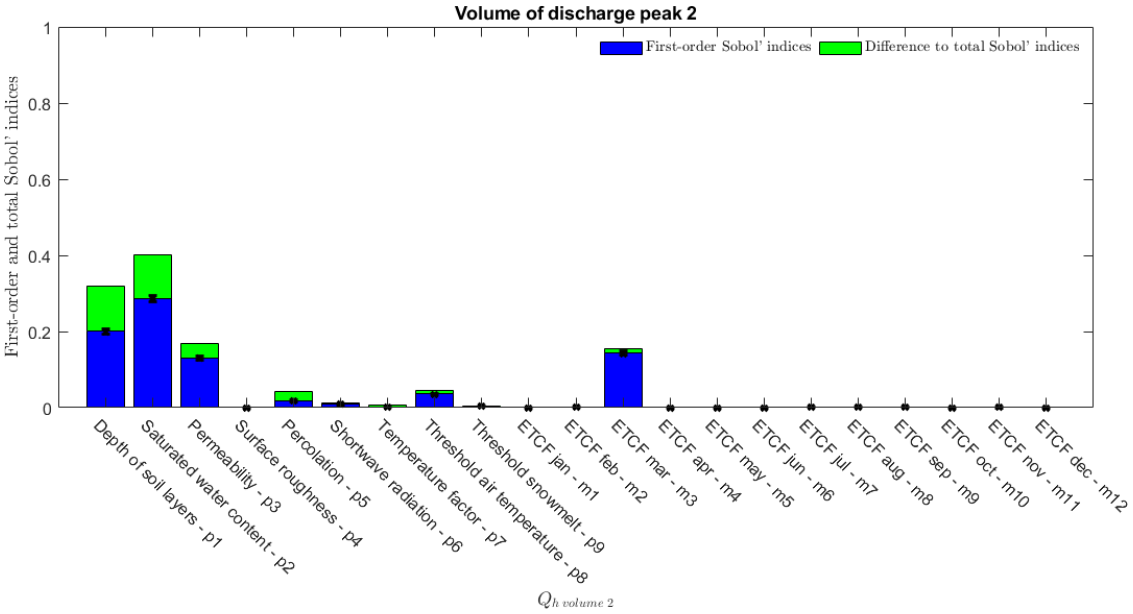


Figure 6: First-order and total indices for the *volume* of hourly peak 2 (spring peak)

The volume total indices in Fig. 6 indicate mostly a $p1p2$ term, accounting for approximately 10% of the total variance, as well as less significant interactions of these two parameters with $p3$ and $p5$.

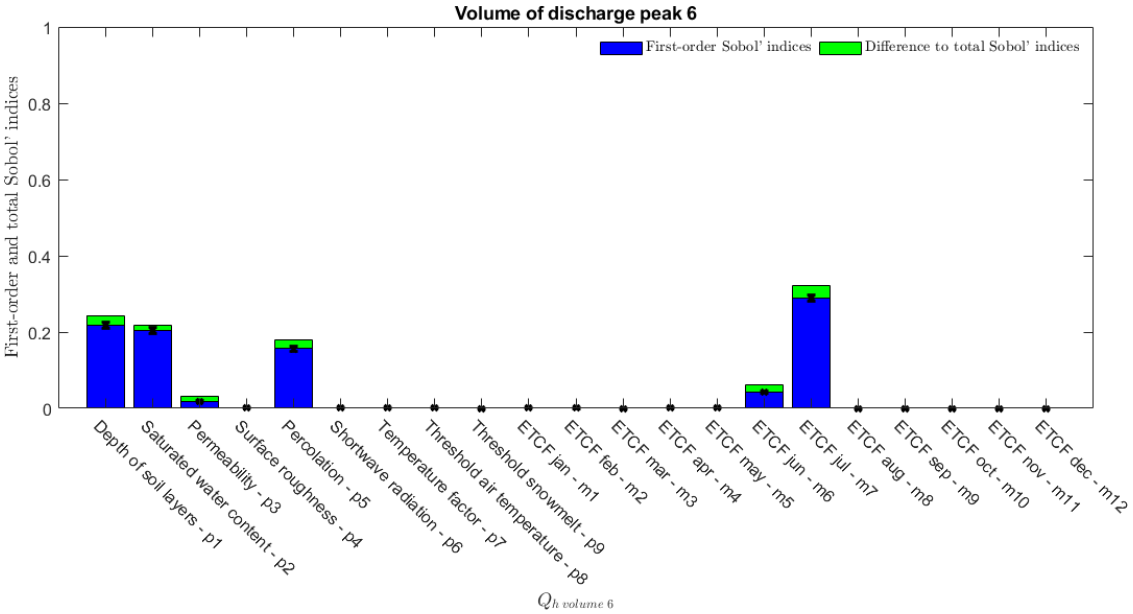


Figure 7: First-order and total indices for the *volume* of hourly peak 6 (summer peak)

Looking at the summer peak (Fig. 7), aside from the usual $p1$ and $p2$, the percolation $p5$ accounts for approximately 15% of the total variance, which was not seen in Fig. 6. Another difference is the much heavier importance of the evapotranspiration crop factor $m7$.

Higher order interactions for this quantity were found to be spread across several

parameters, namely all those having non-zero first-order indices.

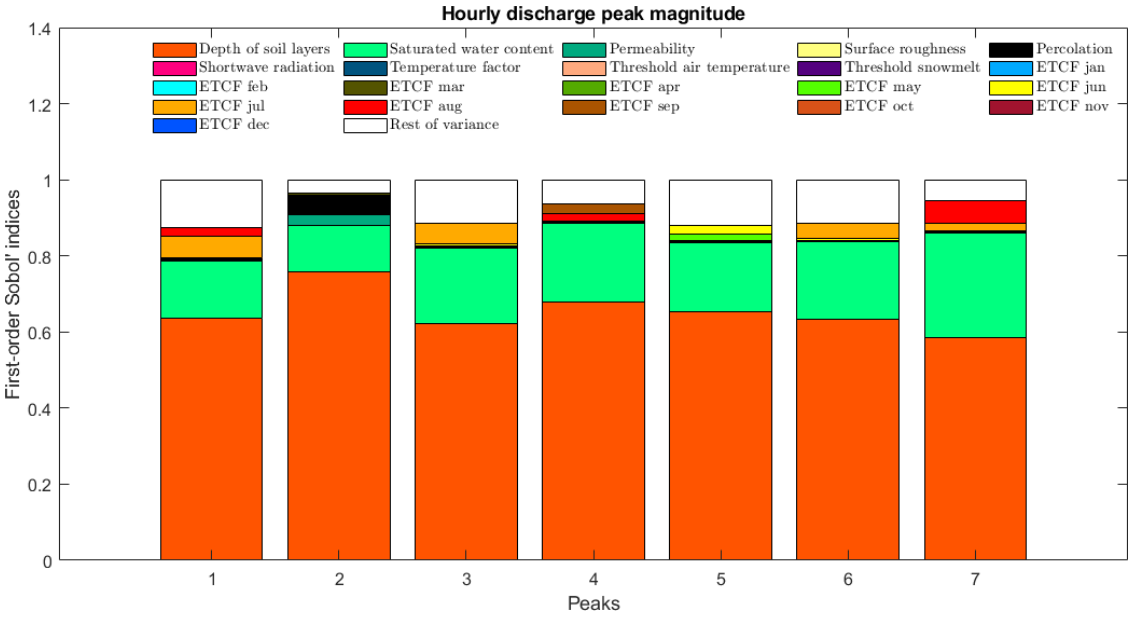


Figure 8: Stacked first-order indices for the *magnitudes* of the hourly peaks

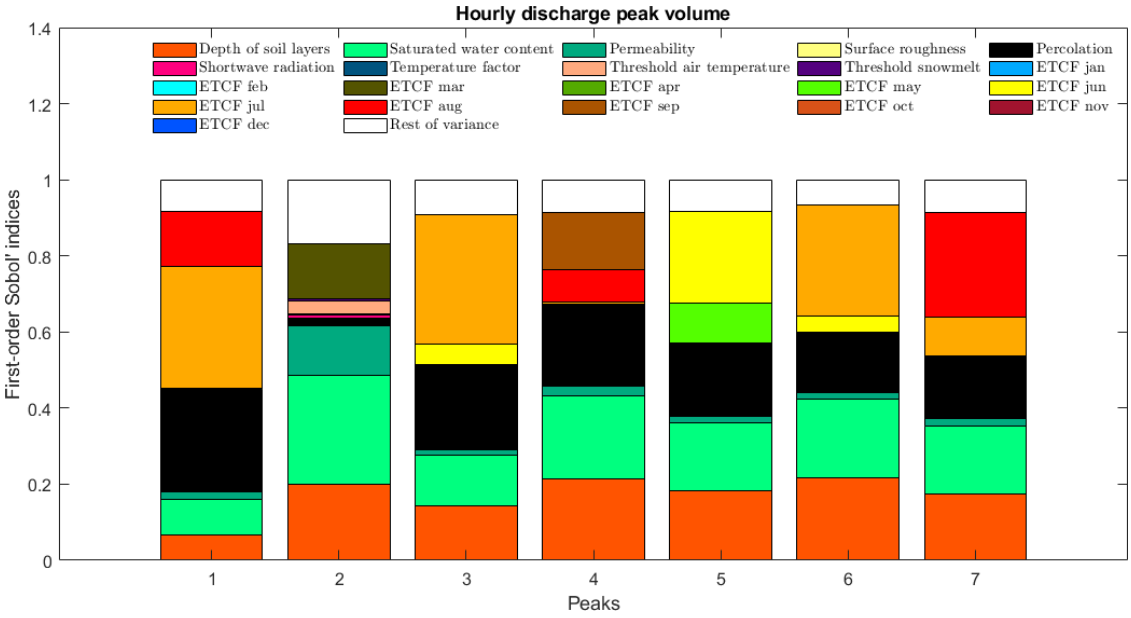


Figure 9: Stacked first-order indices for the *volumes* of the hourly peaks

Next, the first-order indices for all seven peaks in the form of a stacked bar plot are displayed, where different colors denote different model parameters. Fig. 8 confirms the paramount role of $p1$ and $p2$, which for all the peaks, although in different proportions, account combined to $80 - 90\%$ of the output variability. The influence of these two parameters is much more pronounced in peak 2 than in any other peak, making it quite easy to recognize the spring flood event. What the figure also shows is the similarity of peaks 1, 3 and 6. This is not surprising given the strong similarity in their timing

within their respective year. Peak 7 also takes place at a similar timing, but shows slightly different results. This could be due to the fact that peak 6, given the higher return period, has altered the initial conditions in that a lot of water has already been discharged from the catchment.

Comparing with the stacked first-order indices for the peak volumes displayed in Fig. 9, it is possible to notice how the percolation parameter $p5$ exhibits a rather low importance in the summer peak magnitudes, but represents a significant fraction of the total variance for the summer peak volumes. At the same time, an almost opposite behavior can be observed for the spring peak. This suggests a seasonal dependence for this parameter.

The evapotranspiration crop factors in the peak volumes from Fig. 9 shows a pattern where significant importance is given to the crop factor relative to the peak timing, along with a small carryover from the previous month's crop factor. The exceptions to this are the spring peak, where only a dependence on the corresponding crop factor is identified, and peak 1, where another proportion is observed. This could suggest how peak 1 is actually part of the yearly July flood event, only having a delayed response.

The peak volumes in Fig. 9 also show how the contribution to the total variance of the permeability is much more pronounced in the spring peak than in the summer peaks.

6.1.2 7-day low flows

Following are the bar plots for the 7-day low flow for both a summer and a winter period.

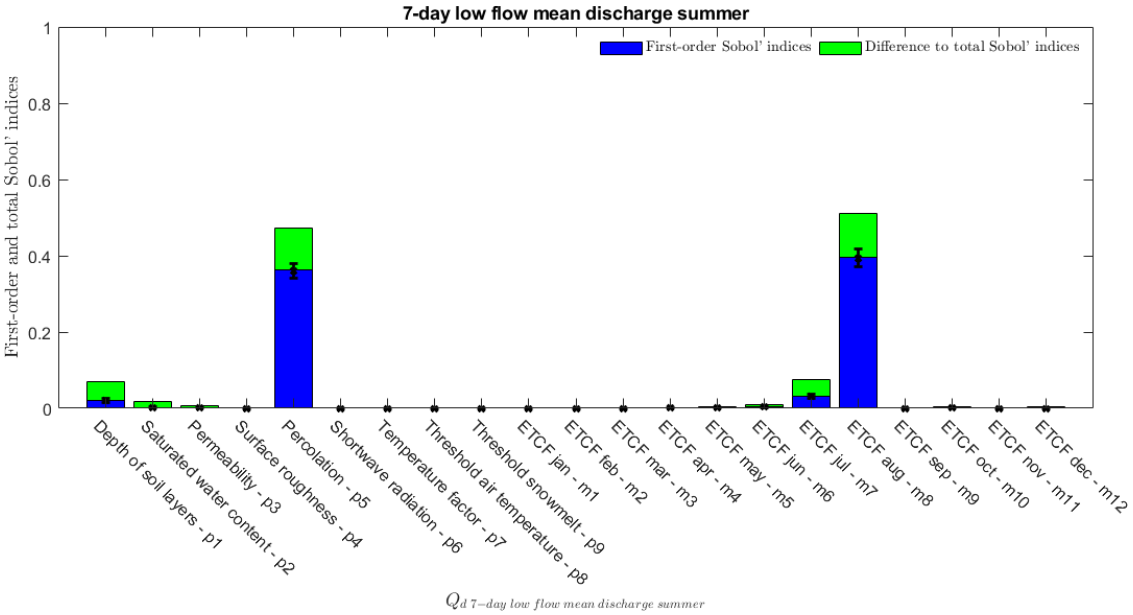


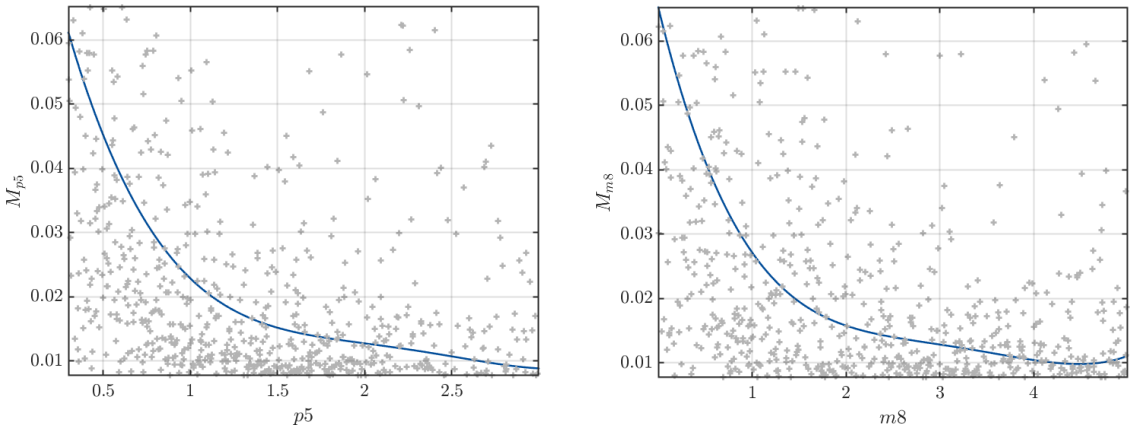
Figure 10: First-order and total indices for the summer 7-day low flow

In the first-order Sobol' indices for the summer 7-day low flow shown in Fig. 10, the two main parameters responsible for approximately 80% of the total variance are the percolation $p5$ and the August evapotranspiration crop factor $m8$, the same month in

which the low flow is taking place.

Though still at reasonably low width, the bounds of the confidence intervals are the highest out of the quantities presented in this section. This quantity is also one that exhibited a very slow convergence of the first-order indices, with little improvement after the first few enrichments. For the convergence plots, see Appendix C.

Among the second-order indices, the most significant term represents the interaction between the two main parameters $p5$ and $m8$.



(a) Elementary effect for $p5$

(b) Elementary effect for $m8$

Figure 11: Elementary effects for most relevant parameters of summer 7-day low flow

The summer 7-day low flow elementary effects show how for both parameters, low values result in higher discharges. Low values of these parameters also provoke the most drastic changes in the flow value.

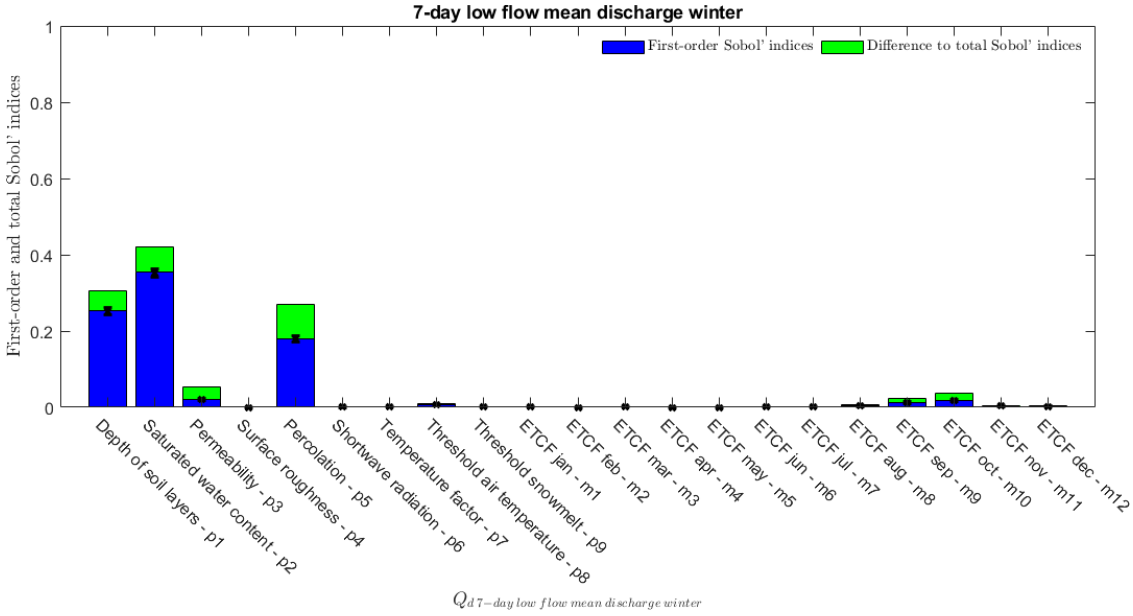


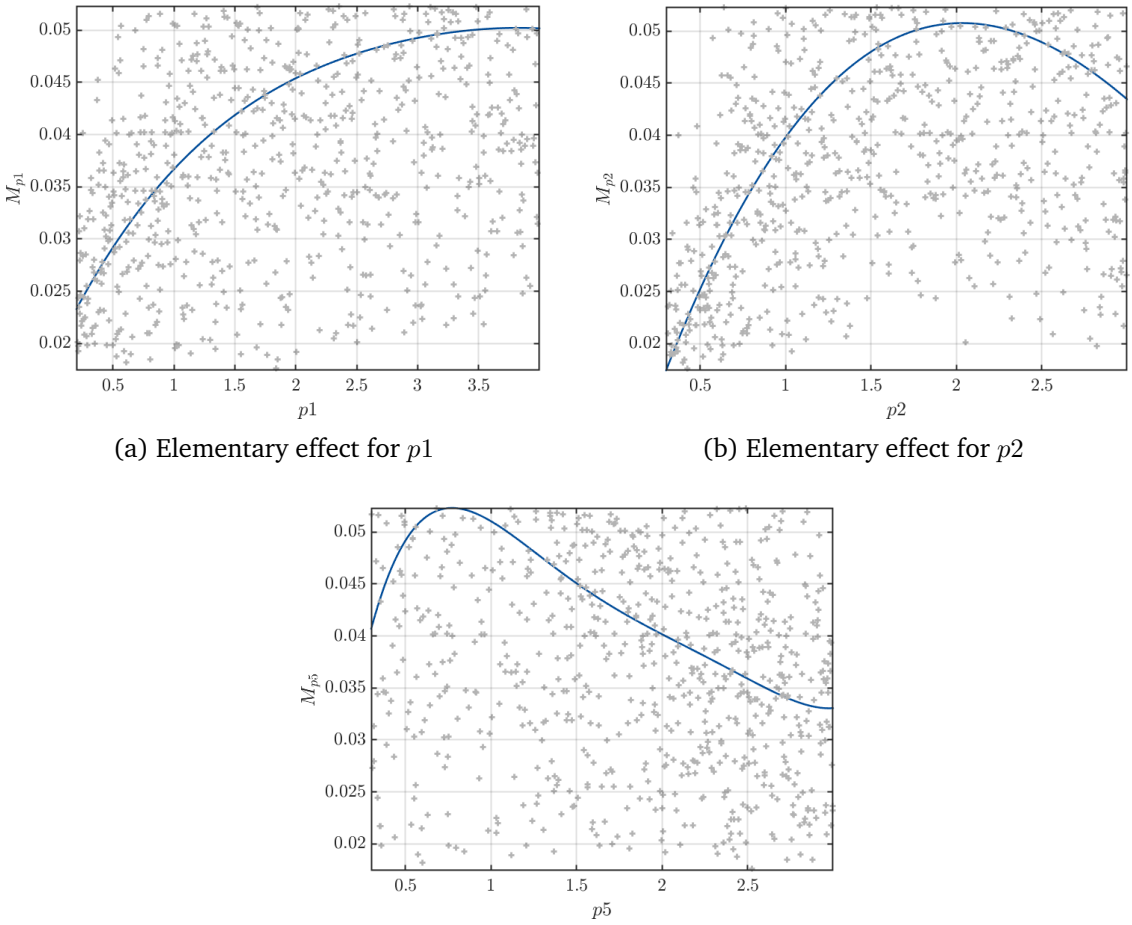
Figure 12: First-order and total indices for the winter 7-day low flow

Fig. 12 illustrates quite clearly how the winter low flow is a result of different pro-

cesses than the summer low flow. The three most important indices indicate soil parameters, whereas the evapotranspiration coefficients play a negligible role.

The reaction to the enrichments is also different than for the summer low flow, with the final ED size resulting in narrower confidence intervals.

Second-order indices indicate mostly a $p2p5$ term, along with a few other less significant interactions.



(a) Elementary effect for $p1$

(b) Elementary effect for $p2$

(c) Elementary effect for $p5$

Figure 13: Elementary effects for most relevant parameters of winter 7-day low flow

The plots of the elementary effects in Fig. 13 show an almost opposite behavior for $p1$ and $p2$ compared to Fig. 3, where low values resulted in higher discharges. In this case, this 7-day low flow takes place during winter and during a period of low precipitation. This means that the river network will be mostly fed by the subsurface component, which makes for a discharge both slower and more spread-out over time. Lower values for $p1$ and $p2$ result in a higher influence of the subsurface flow component, and thus in lower discharges. The percolation has a maximum at around 0.8.

The total indices show how significant interaction is present between almost all parameters. However, the most relevant interaction only concerns $p2$ and $p5$. In fact, the total index for $p1$ contains smaller, less significant interactions, even though their sum

may appear similar to the one for $p2$ and $p5$.

6.2 Average snow cover

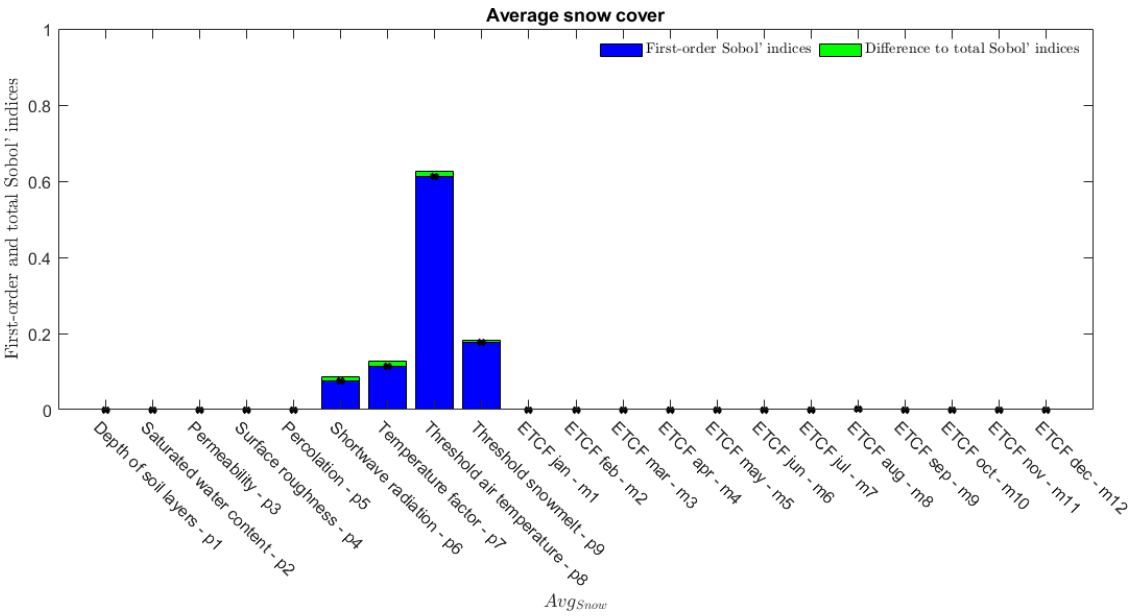


Figure 14: First-order and total indices for the average snow cover

The average snow cover first-order indices in Fig. 14 show a dependency only on the snow parameters, namely from $p6$ to $p9$. These are also the parameters of the ETI model from 2.3.3. The reason for this could be that the snow cover dictates the availability of water for the discharge, but not the other way around. In other words, even if a parameter combination should provoke a high discharge volume during spring, it will only do so up to a certain point if there is not enough snowmelt available. It is also possible to see from the figure how almost all of the variance is explained by first-order indices, and how much more pronounced the influence of the threshold temperature for precipitation state transition $p8$ is with respect to the other snow parameters.

6.3 Evapotranspiration

6.3.1 Monthly averages

Below are the results for the monthly evapotranspiration averages. A more detailed view is given for the February and June averages, so as to identify the most different seasonal patterns.

The February evapotranspiration average in Fig. 15 is influenced mostly by the February crop factor. The percolation also plays a role, although much smaller. This is a parameter that affects the flow of water through the soil, therefore influencing the water supply for evapotranspiration. Some rather minor influence is also given by the snow

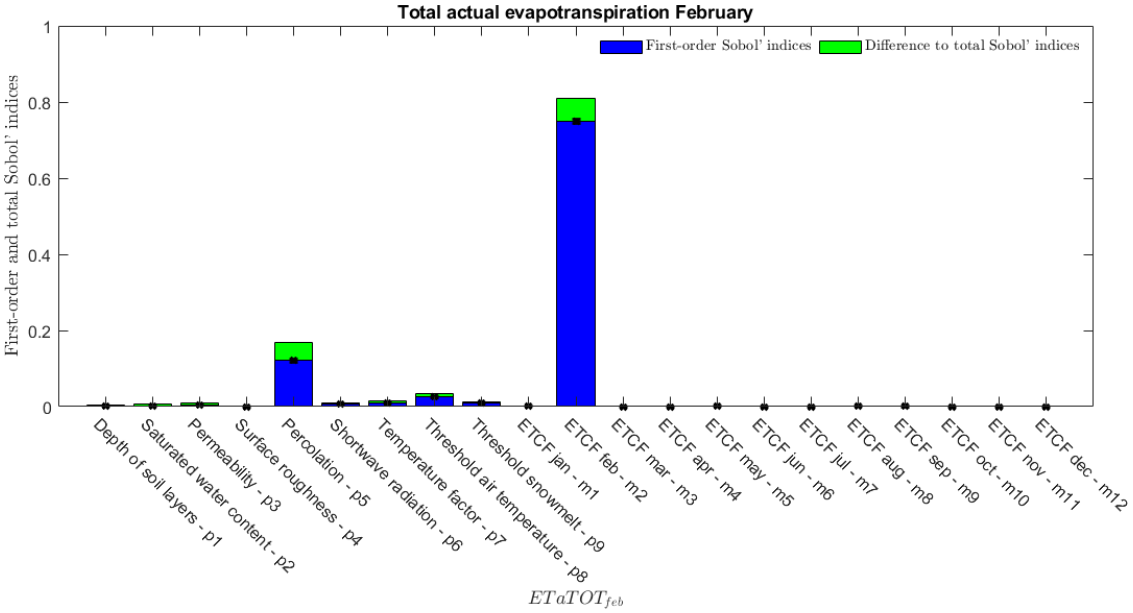


Figure 15: First-order and total indices for the February evapotranspiration

parameters, indicating how the snow cover does have some control over the availability of water.

It must also be noted however, that the values for the evapotranspiration in the winter months are relatively small, and how even though the results for this quantity do show some differences to the ones in Fig. 16 (evapotranspiration during a summer month), it does not have a major influence on the total water budget during that period.

This quantity shows very small confidence intervals bounds. The total indices show interactions mostly between the two parameters mentioned above, $p5$ and $m2$.

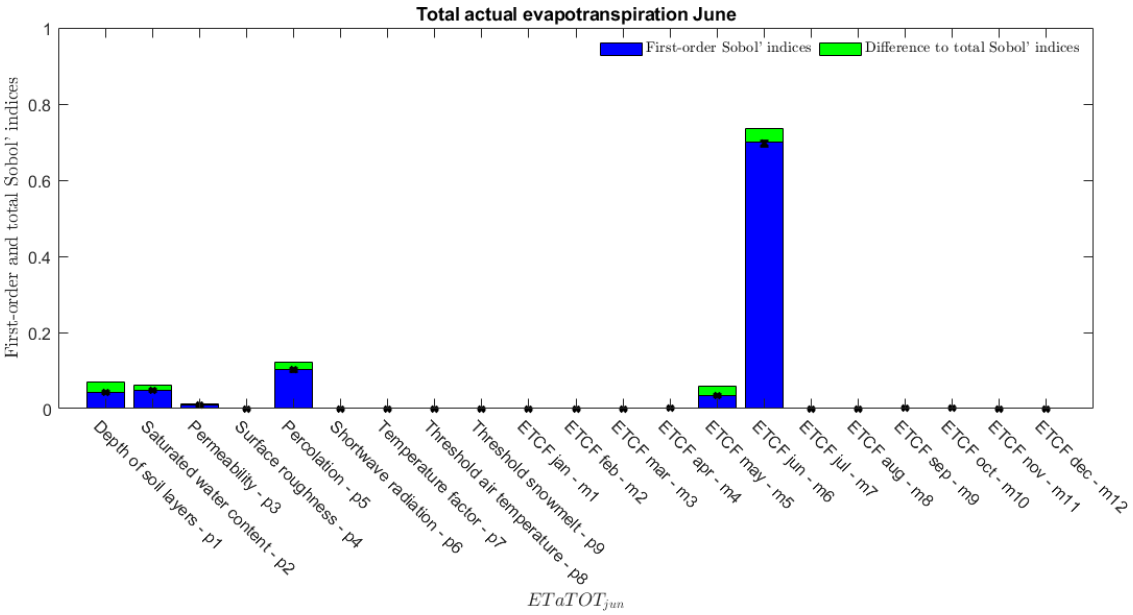


Figure 16: First-order and total indices for the June evapotranspiration

The first-order indices in Fig. 16 show similarities with the ones for the February evapotranspiration average, with percolation and June crop factor explaining most of the variance. However, the influence from the snow parameters, although already rather small, disappears altogether for this month, with some contribution given by $p1$, $p2$ and $p3$. The first two of these parameters were also the most relevant ones for the peak discharge magnitudes from Fig. 8, where they showed how important their role is in controlling the soil water storage capacity. During a month of high evapotranspiration, such as June, water availability can become an issue, therefore explaining the presence of these parameters. Also present in the figure is the crop factor for the previous month, a pattern that was already noticed in the peak discharge volumes from 9.

The second-order indices show an interaction between $p1$ and $m6$, along with a number of less important interaction terms.

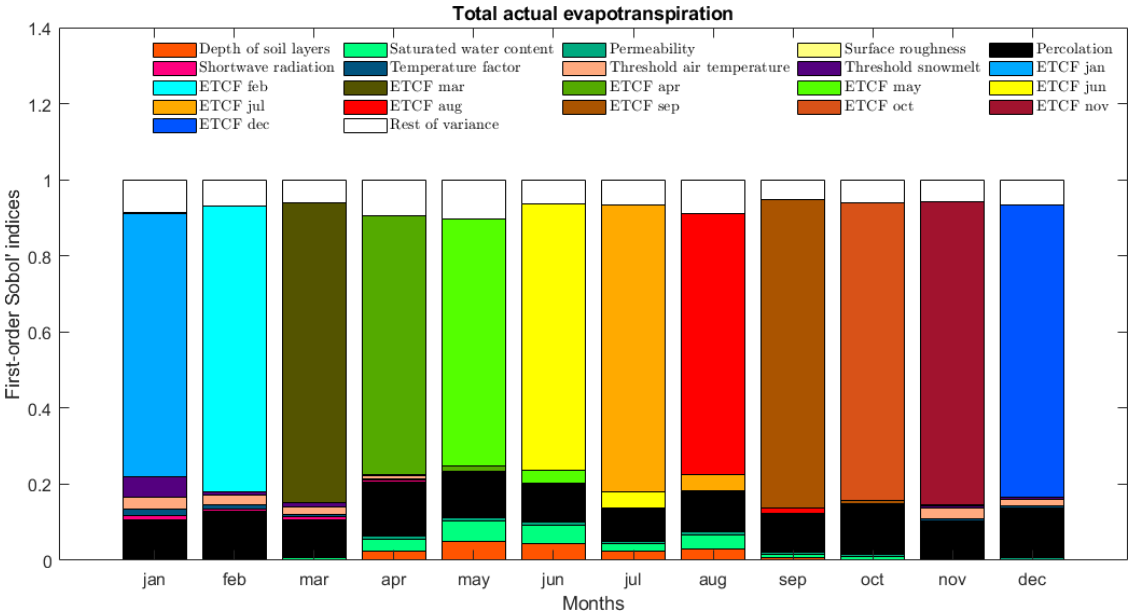


Figure 17: Stacked first-order indices for the monthly evapotranspiration averages

In Fig. 17 the stacked first-order indices for the monthly evapotranspiration averages are displayed. From this, a number of observations can be made.

First of all, around $90 - 95\%$ of the variance is explained by the first-order indices already. Throughout the year, the percolation $p5$ accounts for a significant part of the total variance. As mentioned in 6.3.1, this parameter acts on the water flow through the soil, where the water available for evapotranspiration is stored. Also, aside from the dominant role of each crop factor in their respective month, there seems to be a small carryover during the months from April to September of the previous month's crop factor. These are also the months where the evapotranspiration reaches its highest values. This carryover was noticed previously in the hourly and daily discharges taking place during these same months, and confirms a link between these two processes (higher evapotranspiration causes lower discharges).

During almost the same period, specifically from April to August, a small influence of $p1$ and $p2$ can be seen, whereas the monthly averages from November to March show a small dependency on the snow parameters instead.

6.3.2 Evapotranspiration April to September

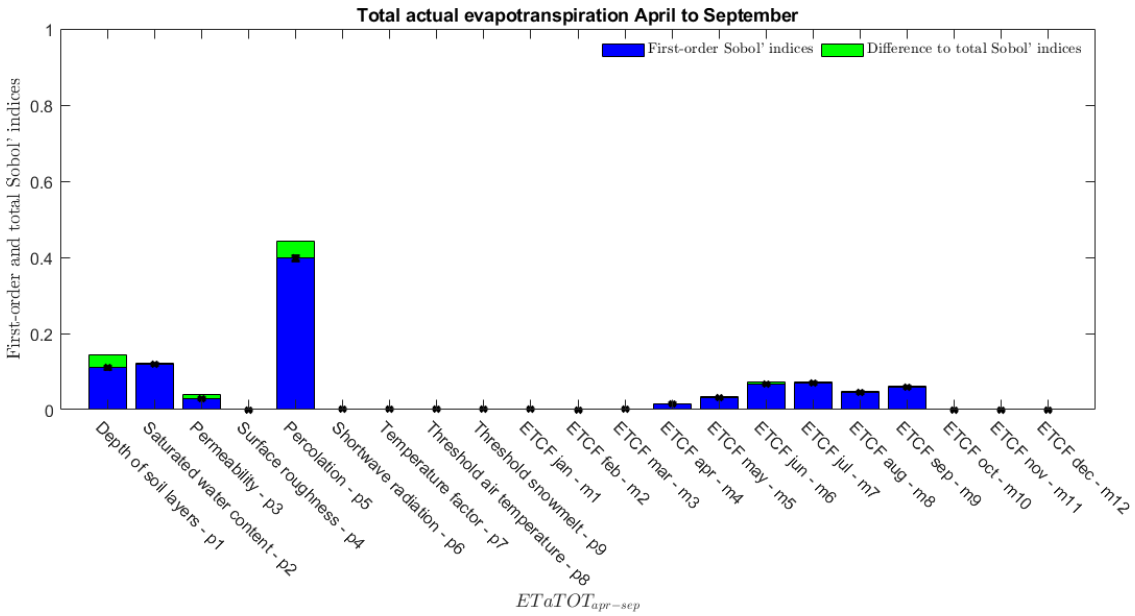


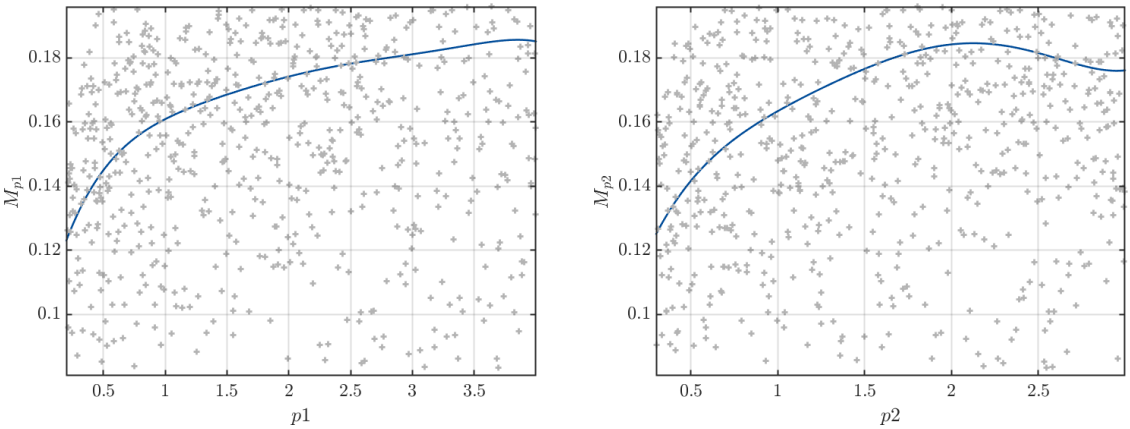
Figure 18: First-order and total indices for the evapotranspiration from April to September

By averaging the April to September monthly evapotranspiration averages, it can be observed from Fig. 18 the predominant role of the percolation $p5$ during these months, although it must be noted how this parameter is actually relevant throughout the entire year. The monthly crop factors belonging to the aggregated averages add up to approximately 30% of the output variability, with neither one standing out. This is expected, since each crop factor is relevant mostly in its respective monthly average.

The figure also shows the recurring influence of the parameters $p1$ and $p2$ during this period.

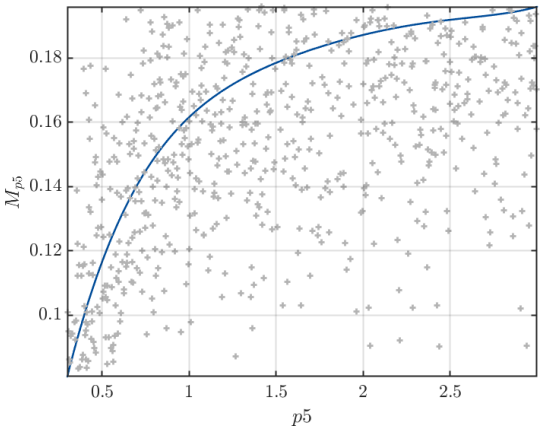
The sum of the first-order indices amounts to approximately 95% of the total variance, with the rest mainly explained by a $p1p5$ interaction term.

The elementary effects in Fig. 19 show rather similar behaviors, where higher values result in a more pronounced evapotranspiration (more water can be stored in the soil, and is therefore available for evapotranspiration). These curves follow an almost opposite behavior with respect to the elementary effects (for these same three parameters) of the volume of the summer discharge peak (see Appendix D).



(a) Elementary effect for $p1$

(b) Elementary effect for $p2$



(c) Elementary effect for $p5$

Figure 19: Elementary effects for most relevant parameters of the evapotranspiration from April to September

7 Discussion and conclusions

Summary and overall performance Within this thesis, an application of uncertainty quantification techniques, mainly sensitivity analysis, in a hydrology setting has been presented. After running $1,047$ TOPKAPI-ETH model evaluations, several outputs were selected and surrogated using polynomial chaos expansions. This made it possible to carry out the sensitivity analysis on those quantities based on the results of the PCEs, therefore allowing to limit the computational costs to the evaluation of the PCE coefficients.

Each quantity is surrogated separately. This is different from what is done in hydrological models, where the results of one quantity influence the others. Still, this is not an issue since each PCE is fed the same input data. For example, a combination of model parameters which causes a low amount of snowmelt will still result in a low spring discharge peak volume, even without the results of the PCE for the average snow cover.

The sensitivity analysis conducted on $1,047$ model evaluations performed quite well for all quantities of interest, therefore confirming the validity of the application in a hydrology context of UQ techniques, both SA and indirectly PCEs. It was possible to identify which model parameters had the most influence on the total variance for all the chosen quantities. Additionally, smaller, although still relevant indices could be obtained. For the observed quantities, these are results both useful and easy to interpret. They reflect the insights an expert would obtain or expect, with the further advantage of being reasonably well and objectively quantified (indices values), which allows hydrologists to start the calibration process properly considering the actual importance of each of the involved model parameters.

From a total of 30 quantities of interest, the 10 that were examined in more detail in Section 6 are briefly summarized in Table 6 in Appendix F in terms of first-order indices of the most relevant parameters. It can be seen how for neither one more than 4 parameters are necessary to explain most of the total variance.

Although the model chosen for the application is composed of parameters rather easy to interpret due to its physically based nature, obtaining the same insights as the ones from this project by using a trial-and-error approach would be a more laborious process.

The analysis also confirmed how the Manning surface roughness $p4$ has no influence or close to none on all the analyzed quantities. This was also observed on other research work, where the effects on flood volume and attenuation due to changes in channel and overland roughness for small-to-mid size catchments were found to be rather small.

Seasonality A seasonal effect could be observed in the Sobol indices, with different parameters governing the output variance in different seasons. This is evident for example in the evapotranspiration monthly averages. In this case, the seasonality also controls the output variance (see Appendix E). Seasonality was also observed in the discharge peak magnitudes, where the permeability only influenced the variance of the spring peak, and the percolation showed a stronger importance than in any of the summer peaks. With regard to the discharge peak volumes, only the spring peak proved to be influenced by

the snow parameters, although in small measure. This peak volume was also the one showing a high importance of the permeability and a low importance of the percolation, a behavior which was reversed in the summer peak volumes.

Importance of interaction terms A high percentage of the total variance could be explained by first-order indices for most of the quantities in the analysis. Only for the case of the discharge peak volumes the seasonality was found to have an effect, where the spring peak discharge volume had more second-order interaction terms than the summer peaks.

For those quantities whose amount of interaction terms is relatively low, it is insightful to look at the plots of the elementary effects. These are plotted in the results section, as well as in Appendix D for the most relevant parameters.

For all the analyzed quantities, the Sobol' indices converged faster than the e_{LOO}. This is expected, given how the indices are computed from the PCE coefficients.

Discharge timings During the peak analysis, in addition to peak volumes as well as peak and 7-day low flow discharge magnitudes, the discharge timings were also investigated. They were not included in the results in Section 6 since issues were noticed, with values bifurcating into several events, and often truncated at the start or end of an interval. However, this does not impact the validity of the discharge magnitudes, and in the case of the hourly peaks, the discharge volume. Furthermore, it does not impact the validity of the UQ techniques involved in the project either, since the issue lies in finding an efficient strategy to obtain those values in the first place. Possible solutions include further restricting the intervals in which to search (like it is already being done in practice by hydrologists). A possible problem is that this yields the maximum (or minimum) value within an interval, rather than finding the exact same event reliably. On the other hand, the observed values for the peak timing did not show a high variability, which would mitigate this problem.

An alternative to this would be substituting the standard TE run with default values for another combination of the model parameters. However, since the input space for these parameters is rather large, it is very difficult to find a parameter combination representative of the entire space. This could perhaps be solved by using multiple sources (multiple "standard runs" on which to look for independent events).

A third possibility could be to identify the independent flood events and low flows on each of the TE model runs, eliminating the dependence on the standard run altogether. While slower, the difference in the required computational time would be in the order of minutes at most. This alternative would call for a new definition of the peak discharge volume, which could be the discharge amount while above the baseflow.

7.1 Outlook

There remains a wide margin for both improvement on the current topic and application on further aspects of hydrology. One example is sensitivity analysis of the likelihood. This

would involve analyzing the two coefficients R^2 and NSE mentioned in 1.1.

The procedure presented in this thesis can easily be applied to another catchment, as well as to quantities of interest obtained from other output time series, and depending on other model parameters. These could be discharge quantities outside of the flood events, such as the baseflow briefly mentioned earlier, or the (monthly) mean infiltration. The peak discharges could also be looked at from another perspective, meaning selecting independent events using different conditions than the ones from Lang et al. (1999), for example focusing more directly on the soil water content. It is also possible to consider local quantities within the catchment. Some examples are water withdrawals for irrigation or hydropower use, and particularly the sources of water for the flood events: spring peaks must not necessarily be fed by snowmelt. Further, because the application only depends on the TE model and its parameters, the feasibility of the procedure extends to the entire range of application of the hydrological model, including for example ungauged catchments.

Should the application presented in this thesis be applied to a longer simulation time span, the extreme flood events to be analyzed could be chosen based on considerations different from the POT sampling technique, such as for example the return period. In fact, as already mentioned in 6.1.1, of the 7 peaks considered, peak 6 showed the highest value overall by a significant margin, indicating it belongs to a much lower probability of occurrence than the usual July-August yearly flood events.

Additionally, should there be a need to investigate an entire time series, its entries could be compressed to principal components. The Sobol' indices could be calculated on those components, which would then be expanded back to the original time series. Although on a different time scale (storm event vs years), the feasibility of this was proven successfully in Nagel et al. (2017).

Doing sensitivity analysis only on the important parameters would drastically reduce the PCE basis, therefore allowing higher degrees and q-norm. This can be done after running the Sobol' analysis on all 21 parameters (like it was done in this project). From its results, the important parameters for each quantity of interest can be defined as those having non-zero values of the total-indices for that same quantity. The same TE runs could be used for this additional sensitivity analysis. The Sobol' indices would converge faster, making better use of the available TE model runs. This would be helpful especially when the ED is not as large as in this project, yet contains enough points for an accurate identification of important and unimportant parameters.

References

Allen, R. G., L. S. Pereira, D. Raes, M. Smith, et al. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *Fao, Rome 300*(9), D05109.

Benning, R. (1995). Towards a new lumped parameterization at catchment scale. Technical report, Landbouwuniversiteit Wageningen.

Beven, K. J. and M. J. Kirkby (1979). A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal 24*(1), 43–69.

Blatman, G. and B. Sudret (2011). Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics 230*(6), 2345–2367.

Burlando, P. (2017, September). Lecture notes in hydrology i.

Burlando, P. and S. Fatichi (2018, September). Lecture notes in hydrology ii.

Ciarapica, L. and E. Todini (2002). Topkapi: A model for the representation of the rainfall-runoff process at different scales. *Hydrological Processes 16*(2), 207–229.

Devia, G. K., B. Ganasri, and G. Dwarakish (2015). A review on hydrological models. *Aquatic Procedia 4*(1), 1001–1007.

Fajraoui, N., S. Marelli, and B. Sudret (2017). On optimal experimental designs for sparse polynomial chaos expansions.

Fatichi, S., S. Rimkus, P. Burlando, R. Bordoy, and P. Molnar (2015). High-resolution distributed analysis of climate and anthropogenic changes on the hydrology of an alpine catchment. *Journal of Hydrology 525*, 362–382.

Gratiet, L. L., S. Marelli, and B. Sudret (2016). Metamodel-based sensitivity analysis: polynomial chaos expansions and gaussian processes. *arXiv preprint arXiv:1606.04273*.

Jacobs, A. and H. De Bruin (1998). Makkink's equation for evapotranspiration applied to unstressed maize. *Hydrological processes 12*(7), 1063–1066.

Kasten, F. and G. Czeplak (1980). Solar and terrestrial radiation dependent on the amount and type of cloud. *Solar energy 24*(2), 177–189.

Lang, M., T. Ouarda, and B. Bobée (1999). Towards operational guidelines for over-threshold modeling. *Journal of hydrology 225*(3-4), 103–117.

Liu, Z. and E. Todini (2002). Towards a comprehensive physically-based rainfall-runoff model.

Marelli, S., C. Lamas, K. Konakli, C. Mylonas, P. Wiederkehr, and B. Sudret (2019). UQLab user manual – Sensitivity analysis. Technical report, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland. Report # UQLab-V1.3-106.

Marelli, S. and B. Sudret (2014). Uqlab: A framework for uncertainty quantification in matlab. In *Vulnerability, uncertainty, and risk: quantification, mitigation, and management*, pp. 2554–2563.

Marelli, S. and B. Sudret (2019). UQLab user manual – Polynomial chaos expansions. Technical report, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland. Report # UQLab-V1.3-104.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics 21*(2), 239–245.

Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology 19*, 205–34.

Moustapha, M., S. Marelli, and B. Sudret (2019). UQLab user manual – The UQLink module. Technical report, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland. Report # UQLab-V1.3-110.

Nagel, J. B., J. Rieckermann, and B. Sudret (2017). Uncertainty quantification in urban drainage simulation: fast surrogates for sensitivity analysis and model calibration. *arXiv preprint arXiv:1709.03283*.

Paschalis, A., S. Fatichi, P. Molnar, S. Rimkus, and P. Burlando (2014). On the effects of small scale space–time variability of rainfall on basin flood response. *Journal of Hydrology 514*, 313–327.

Paschalis, A., P. Molnar, S. Fatichi, and P. Burlando (2013). A stochastic model for high-resolution space-time precipitation simulation. *Water Resources Research 49*(12), 8400–8417.

Peleg, N., C. Skinner, S. Fatichi, and P. Molnar. Temperature effects on heavy rainfall modify catchment hydro-morphological response.

Pellicciotti, F., B. Brock, U. Strasser, P. Burlando, M. Funk, and J. Corripio (2005). An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for haut glacier d'arolla, switzerland. *Journal of Glaciology 51*(175), 573–587.

Priestley, C. H. B. and R. Taylor (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review 100*(2), 81–92.

Rimkus, S. (2013). Documentation and user guide to the hydrological model topkapi-eth.

Sobol, I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical modelling and computational experiment 1*, 407–414.

Sudret, B. (2007). Uncertainty propagation and sensitivity analysis in mechanical models–contributions to structural reliability and stochastic spectral methods. *Habilitationa diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France 147*.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety 93*(7), 964–979.

Todini, E. (1988). Rainfall-runoff modeling—past, present and future. *Journal of Hydrology 100*(1-3), 341–352.

Todini, E. (1996). The arno rainfall—runoff model. *Journal of hydrology 175*(1-4), 339–382.

Xiu, D. and G. E. Karniadakis (2002). The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing 24*(2), 619–644.

A PCE results

Although not the main focus of the thesis, it is still worth it to consider the quality of the PCEs, as they are the foundation of the Sobol' analysis when using the PCE-based approach.

As it can be seen from Table 5, the PCE degree chosen by UQLab (i.e. with the smallest LOO error) was found to range between 6 and 7, whereas the modified leave-one-out error stays in the 1×10^{-3} to 1×10^{-2} range, with the exception of the average snow cover, which attains even lower values.

The size of the full PCE basis using the hyperbolic truncation and maximum interaction schemes is given in Table 4 for every maximum polynomial degree tried in the analysis.

Table 4: Size of full basis per PCE maximum polynomial degree

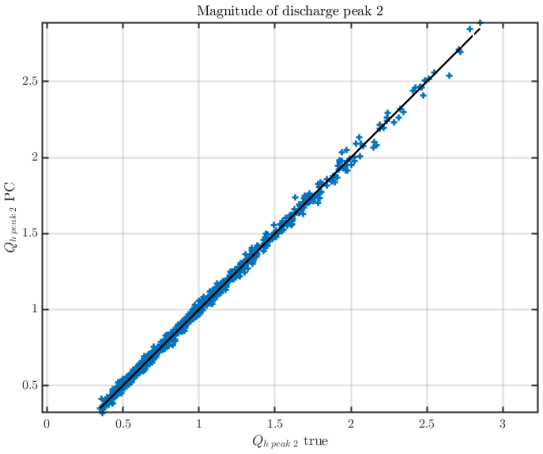
Max. polynomial degree	2	3	4	5	6	7
Size of full basis	43	274	715	736	1,387	2,248

Table 5: Quantities of interest (1047 runs).

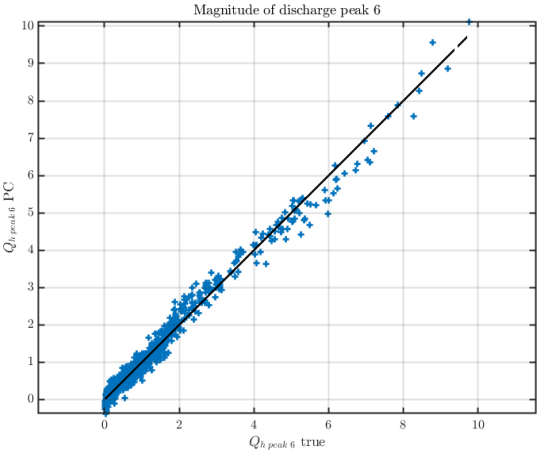
Variable name	e^*_{LOO}	PCE degree	Sparse PCE basis	CV [%]
$Q_{h\,peak\,1}$	2.41×10^{-2}	7	151	124.291
$Q_{h\,volume\,1}$	2.54×10^{-2}	6	133	67.055
$Q_{h\,peak\,2}$	4.47×10^{-3}	7	148	52.268
$Q_{h\,volume\,2}$	1.92×10^{-2}	7	202	12.035
$Q_{h\,peak\,3}$	3.00×10^{-2}	7	164	134.395
$Q_{h\,volume\,3}$	2.99×10^{-2}	7	191	78.177
$Q_{h\,peak\,4}$	1.74×10^{-2}	7	114	108.049
$Q_{h\,volume\,4}$	2.84×10^{-2}	7	142	69.369
$Q_{h\,peak\,5}$	1.98×10^{-2}	7	160	135.801
$Q_{h\,volume\,5}$	2.43×10^{-2}	7	206	68.586
$Q_{h\,peak\,6}$	2.96×10^{-2}	7	124	123.605
$Q_{h\,volume\,6}$	2.65×10^{-2}	7	191	72.916
$Q_{h\,peak\,7}$	1.54×10^{-2}	7	212	86.320
$Q_{h\,volume\,7}$	2.33×10^{-2}	7	164	59.160
$Q_{d\,7-day\,low\,flow\,mean\,discharge\,summer}$	7.00×10^{-2}	7	139	104.935
$Q_{d\,7-day\,low\,flow\,mean\,discharge\,winter}$	3.69×10^{-2}	7	204	34.644
Avg_{Snow}	6.13×10^{-4}	7	131	14.604
$ETaTOT_{jan}$	6.81×10^{-3}	7	128	68.089
$ETaTOT_{feb}$	5.99×10^{-3}	7	137	64.952
$ETaTOT_{mar}$	6.99×10^{-3}	7	112	62.850
$ETaTOT_{apr}$	1.04×10^{-2}	7	182	56.733
$ETaTOT_{may}$	1.76×10^{-2}	7	154	48.874
$ETaTOT_{jun}$	1.76×10^{-2}	7	166	43.548
$ETaTOT_{jul}$	1.51×10^{-2}	7	121	49.057
$ETaTOT_{aug}$	2.09×10^{-2}	7	184	47.003
$ETaTOT_{sep}$	1.26×10^{-2}	7	105	52.269
$ETaTOT_{oct}$	6.00×10^{-3}	7	205	59.219
$ETaTOT_{nov}$	4.57×10^{-3}	7	158	63.728
$ETaTOT_{dec}$	4.15×10^{-3}	7	147	65.013
$ETaTOT_{apr-sep}$	9.66×10^{-3}	6	209	25.524

B Y-Y plots

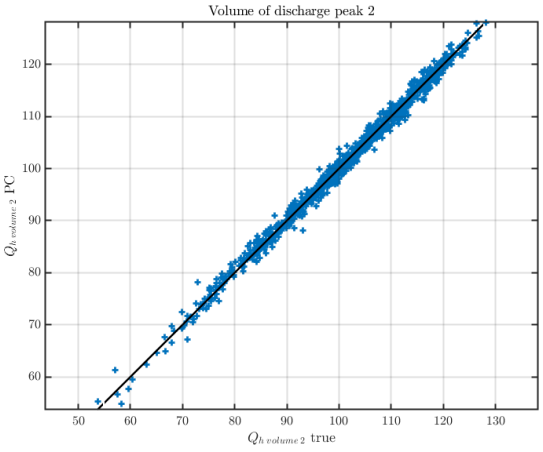
In the following, the Y-Y plots for the quantities selected in section 6 are shown. It can be seen how for many of them the metamodel returns a negative response for small values. However, as already explained in 5.1, this is not an issue as the PCEs are not used for prediction, but only for sensitivity analysis.



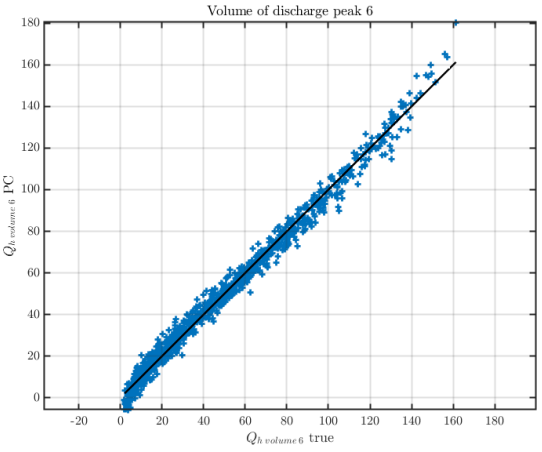
(a) Discharge *magnitude* of hourly peak 2



(b) Discharge *magnitude* of hourly peak 6

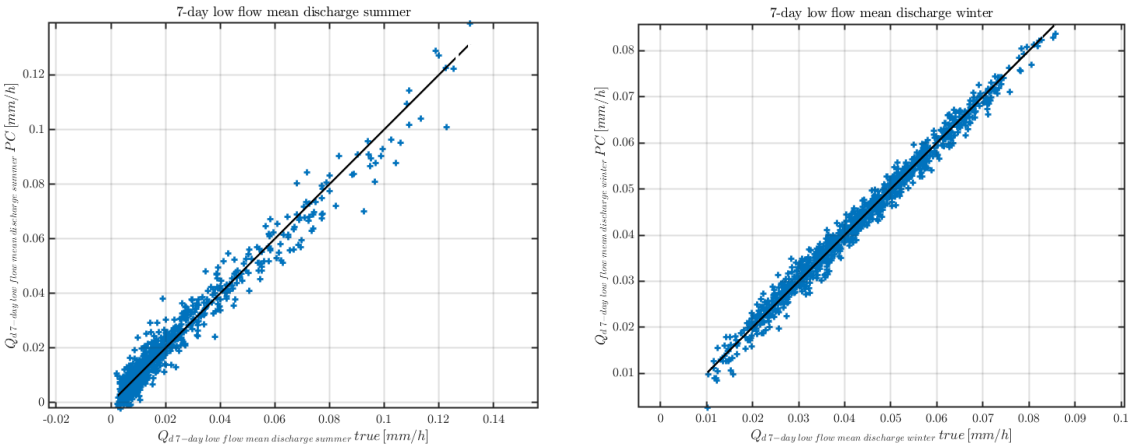


(c) Discharge *volume* of hourly peak 2



(d) Discharge *volume* of hourly peak 6

Figure 20: *Magnitude* and *volume* of hourly discharge peaks 2 and 6



(a) Summer 7-day low flow (b) Winter 7-day low flow

Figure 21: Summer and winter 7-day low flow

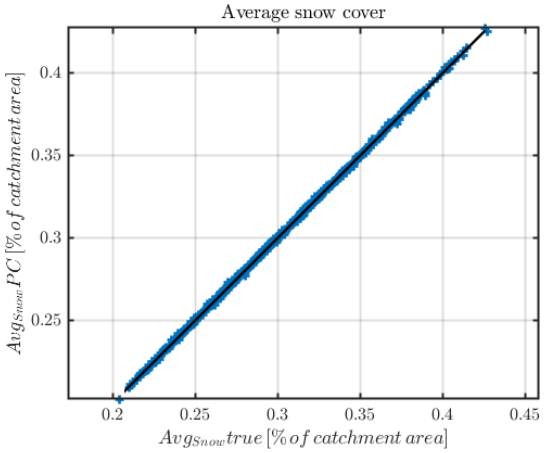
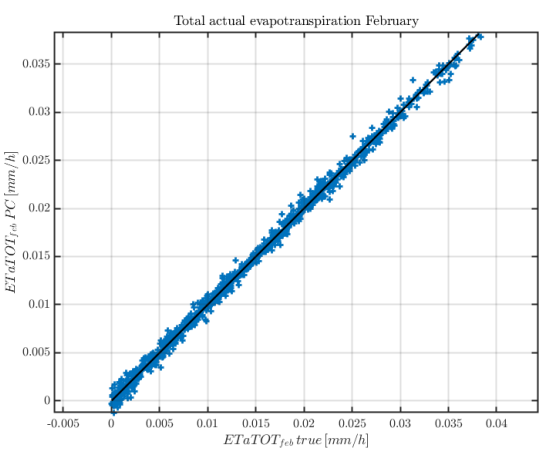
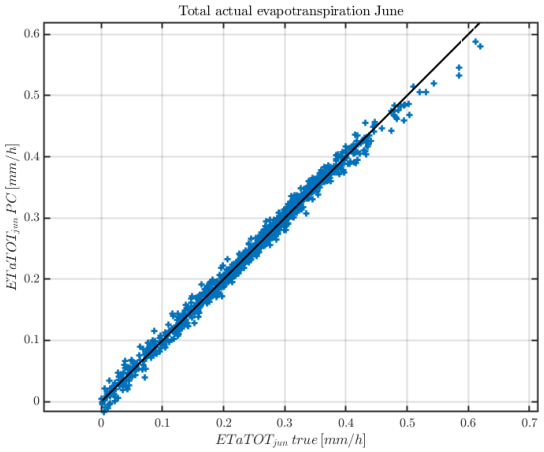


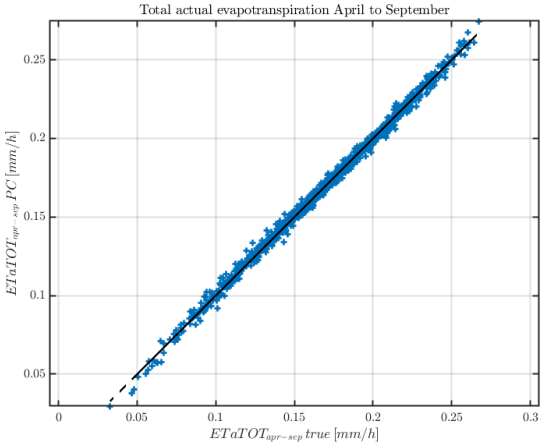
Figure 22: Average snow cover



(a) February evapotranspiration



(b) June evapotranspiration



(c) Evapotranspiration April to September

Figure 23: Monthly evapotranspiration averages for February and June, as well as April-September aggregated averages

C Convergence plots

The following plots show the convergence of the most relevant first-order indices as well as that of the modified leave-one-out error for a few selected quantities of interest. More specifically, one quantity converging later than the others in its category (discharge volume of peak 5, Fig. 24, later convergence than other peak volumes), one quantity whose PCE showed a quite high leave-one-out error (summer 7-day low flow, Fig. 25), and one converging very fast (average snow cover, Fig. 26).

The values for each Sobol' index, as well as the bounds of its confidence interval, are plotted as a percentage of its respective final value (value at final ED size).

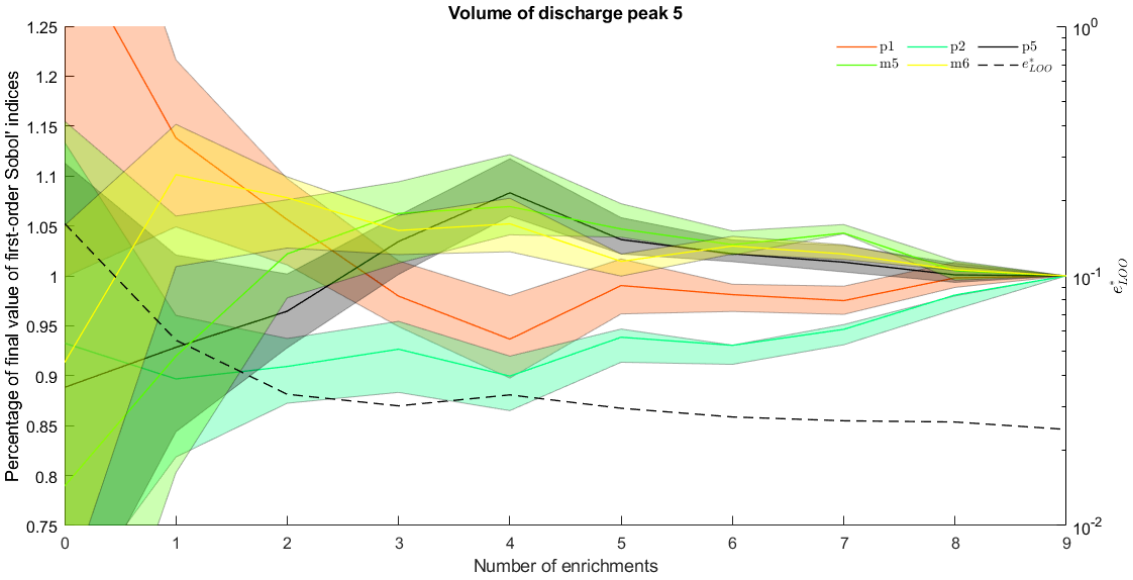


Figure 24: *Volume* of hourly discharge peak 5 convergence

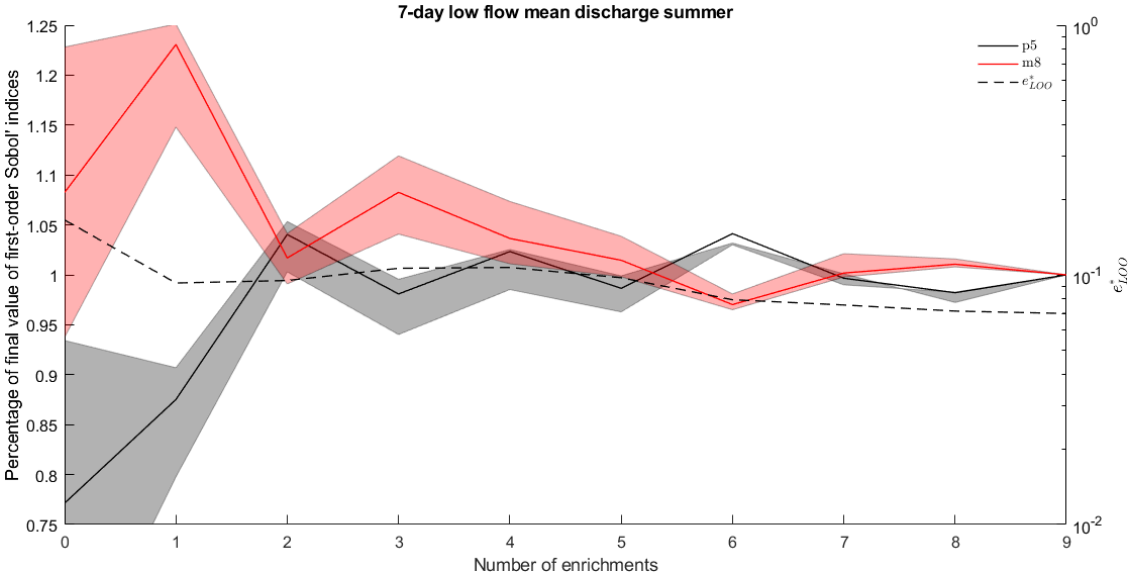


Figure 25: Summer 7-day low flow convergence

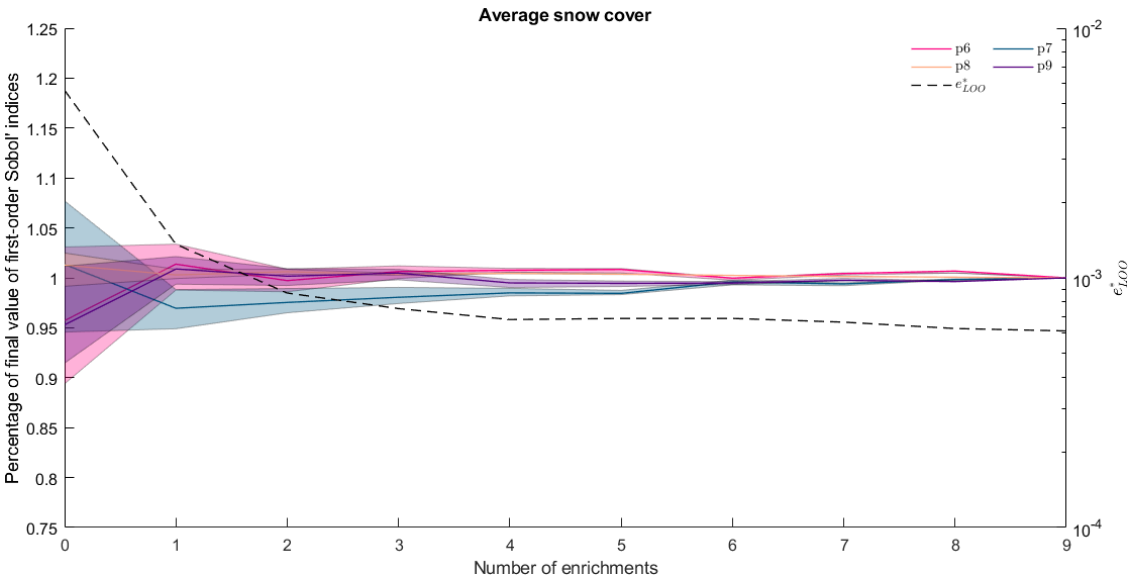
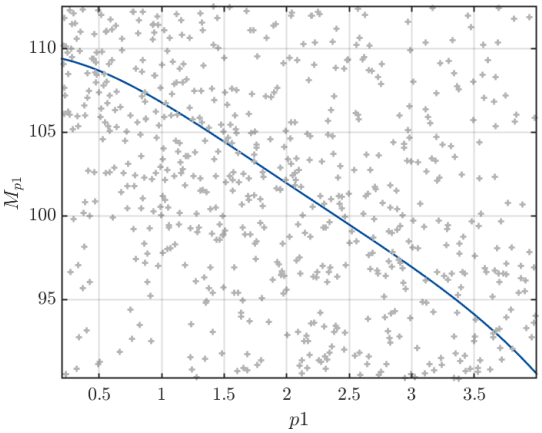


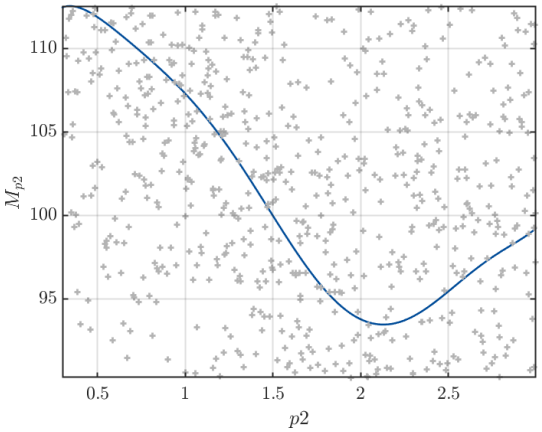
Figure 26: Average snow cover convergence

D Elementary effects

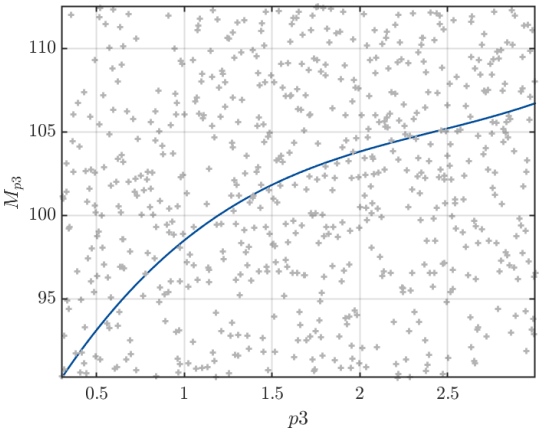
For some of the quantities of interest selected and analyzed in section 6 (out of a total of 30) the elementary effects of the most relevant parameters were shown. The relevant elementary effects for the rest of those quantities are shown here.



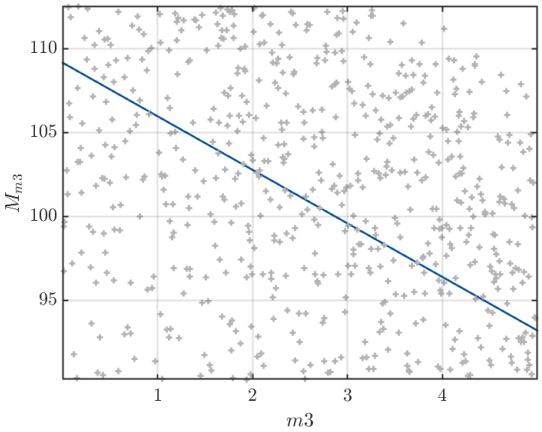
(a) Elementary effect for $p1$



(b) Elementary effect for $p2$

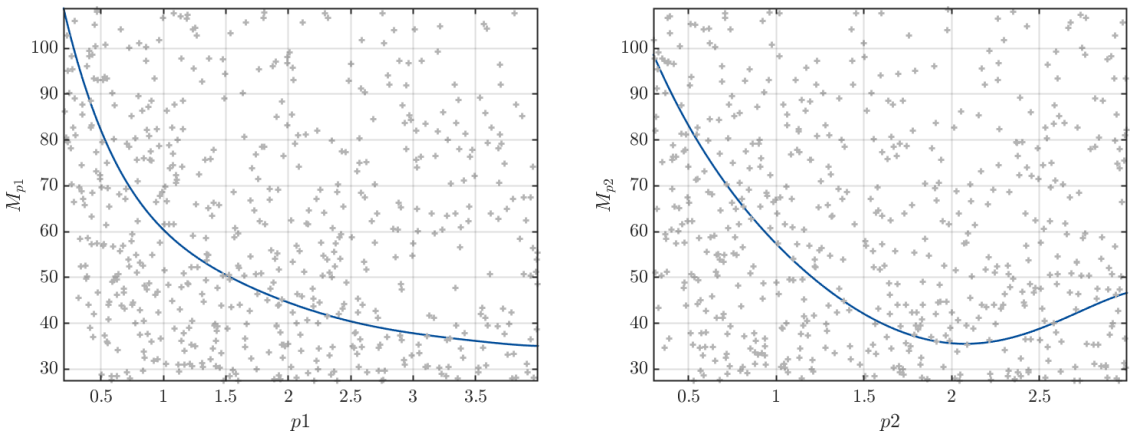


(c) Elementary effect for $p3$



(d) Elementary effect for $m3$

Figure 27: Elementary effects for most relevant parameters of the *volume* of hourly peak 2



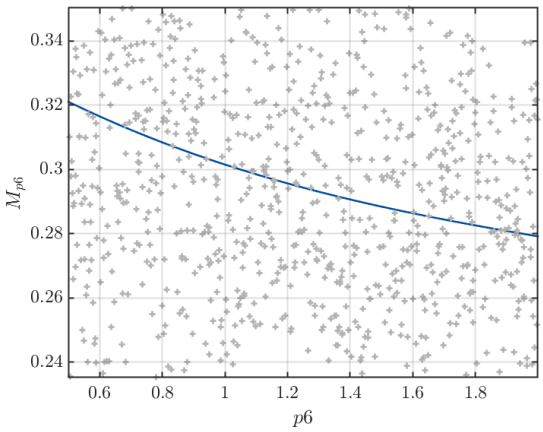
(a) Elementary effect for $p1$

(b) Elementary effect for $p2$

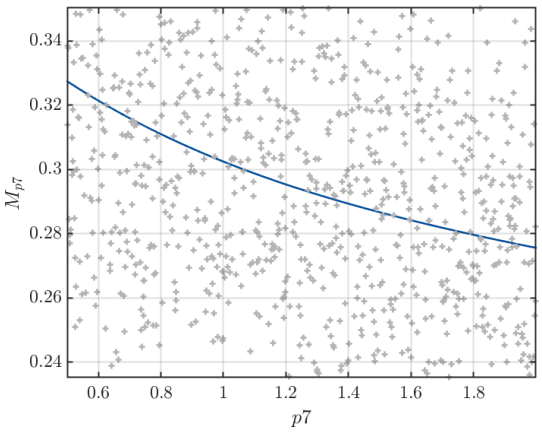
(c) Elementary effect for $p5$

(d) Elementary effect for $m7$

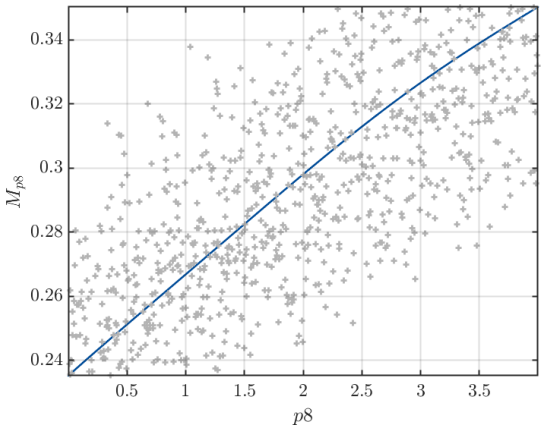
Figure 28: Elementary effects for most relevant parameters of the *volume* of hourly peak 6



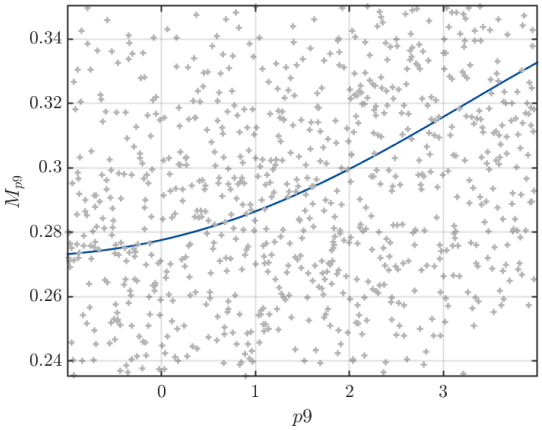
(a) Elementary effect for $p6$



(b) Elementary effect for $p7$

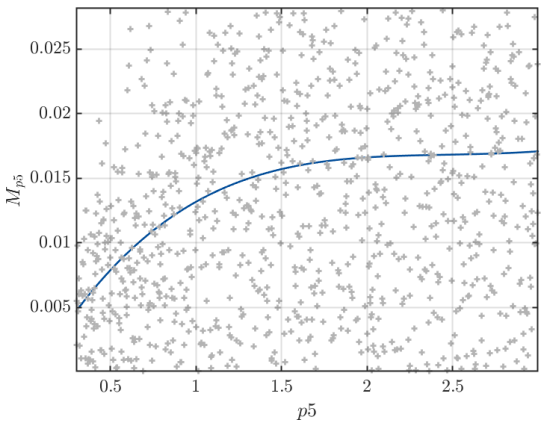


(c) Elementary effect for $p8$

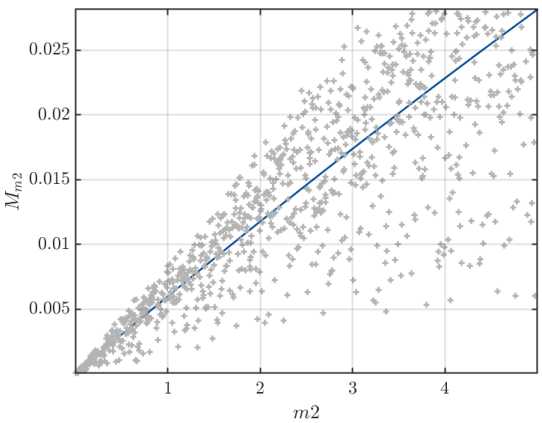


(d) Elementary effect for $p9$

Figure 29: Elementary effects for most relevant parameters of the average snow cover

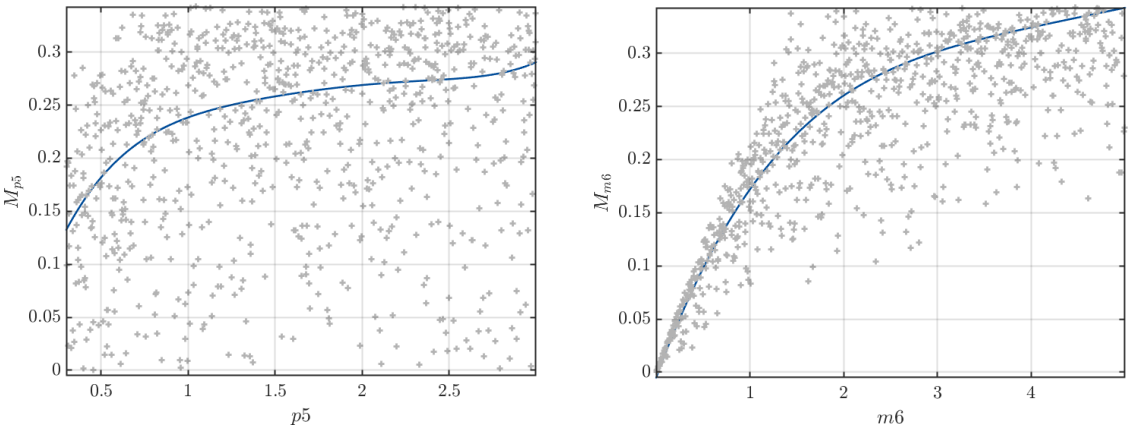


(a) Elementary effect for $p5$



(b) Elementary effect for $m2$

Figure 30: Elementary effects for most relevant parameters of the February evapotranspiration



(a) Elementary effect for $p5$

(b) Elementary effect for $m6$

Figure 31: Elementary effects for most relevant parameters of the June evapotranspiration

E Variance-scaled plots of first-order indices

The stacked bar plots showing the first-order indices for all quantities of interest from Section 6 are repeated here, only scaled with the total variance. It can be seen how the maximum summer peak (peak 6) shows a higher variance for both magnitude and volume. The spring peak (peak 2) on the other hand has the lowest variance overall (again for both magnitude and volume). For both 7-day low flows and monthly evapotranspiration averages, a higher variance resulted in a PCE with higher e^*_{LOO} (see Table 5). Additionally, the monthly evapotranspiration averages showed a strong seasonality, with the variance of the summer months on a different order of magnitude than the one for the winter months.

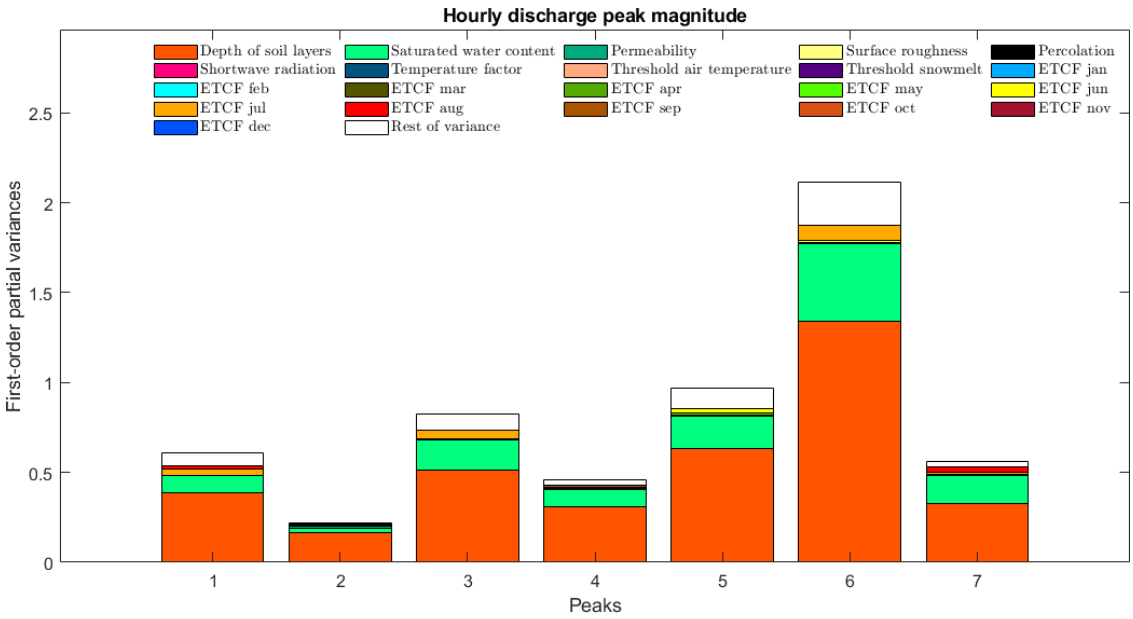


Figure 32: Stacked first-order indices of *magnitudes* of hourly discharge peaks

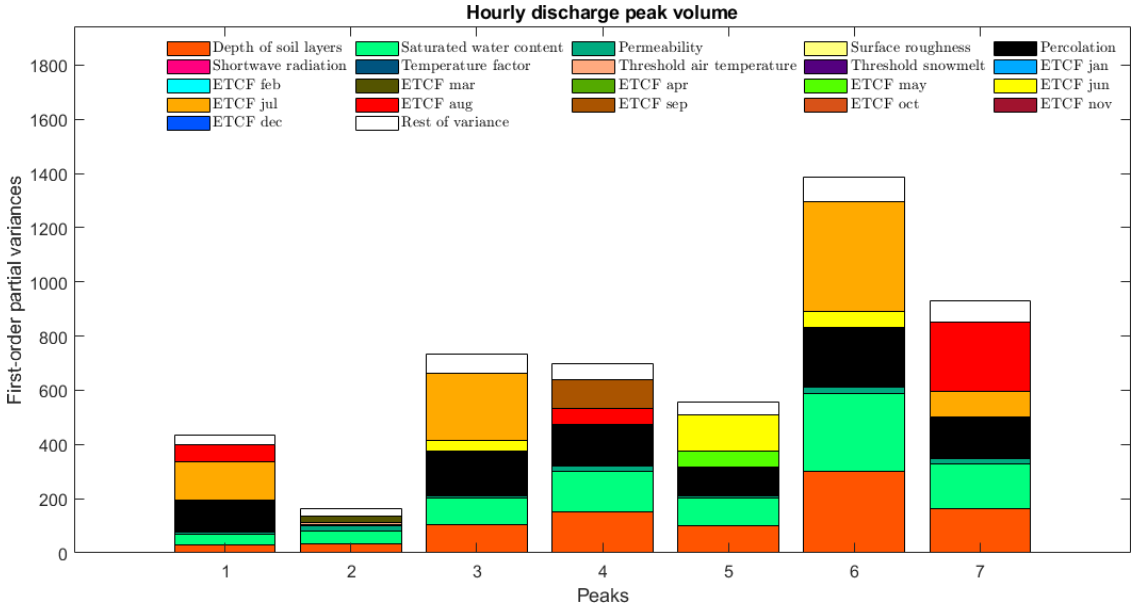


Figure 33: Stacked first-order indices of *volume* of hourly discharge peaks

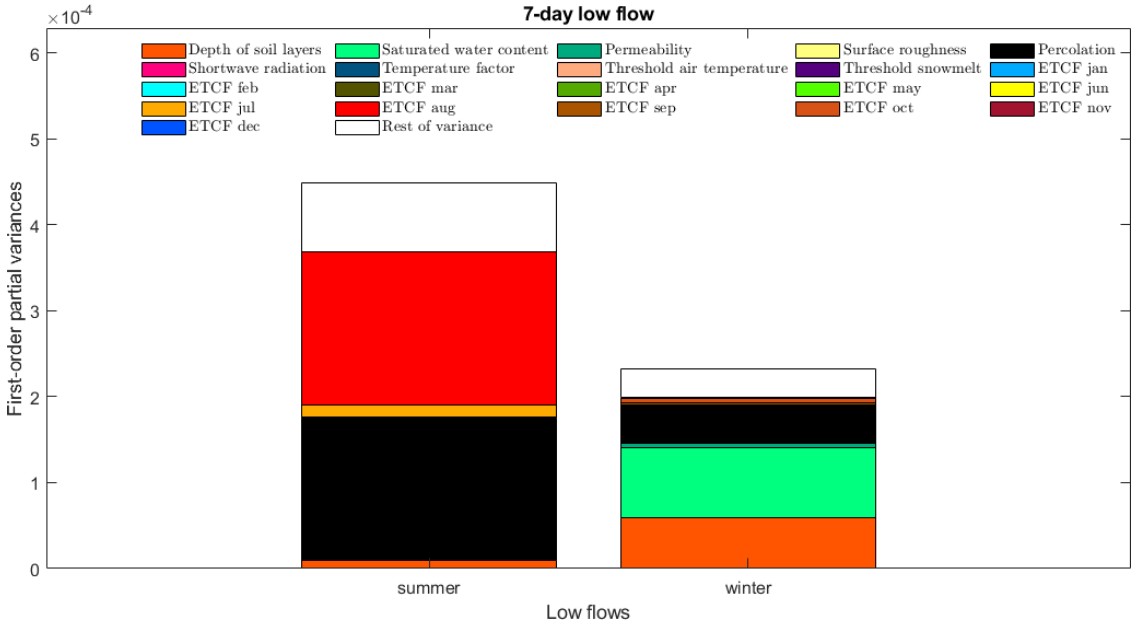


Figure 34: Stacked first-order indices of 7-day low flows

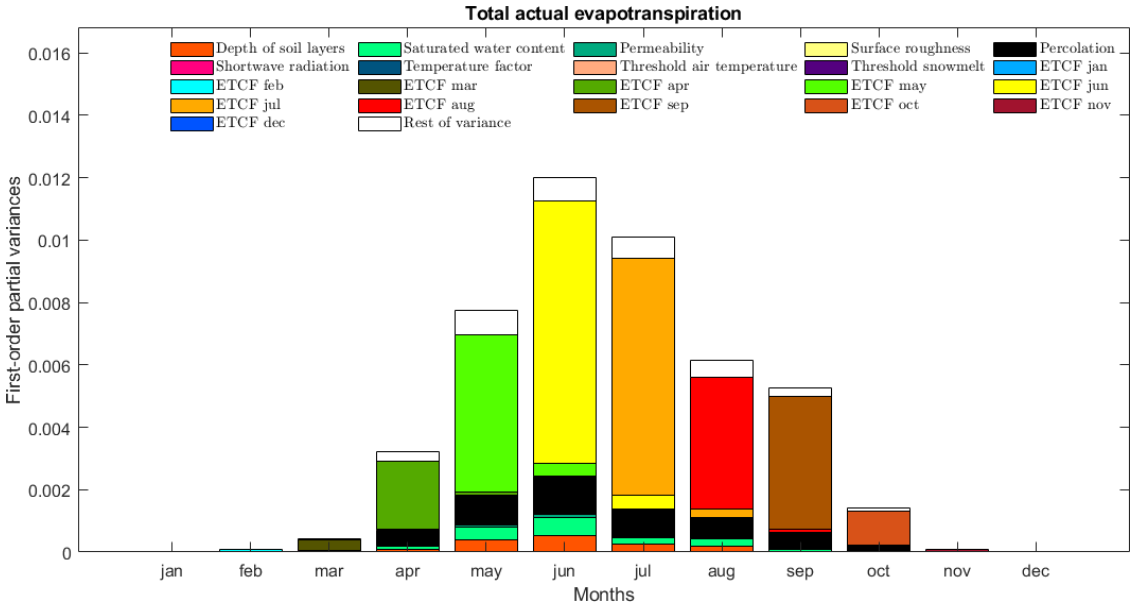


Figure 35: Stacked first-order indices of monthly evapotranspiration averages

F Summary

For the 10 out of 30 quantities analyzed in more detail in Section 6, the most relevant parameters together with the value of their first-order Sobol' index, are shown here.

Table 6: Relevant indices for quantities of interest selected in Section 6

Variable name	Parameter / Index value									
$Q_{h\,peak\,2}$	$p1$	0.7571	$p2$	0.1219						
$Q_{h\,peak\,6}$	$p1$	0.6345	$p2$	0.2015						
$Q_{h\,volume\,2}$	$p1$	0.2003	$p2$	0.2868	$p3$	0.1300	$m3$	0.1432		
$Q_{h\,volume\,6}$	$p1$	0.2182	$p2$	0.2050	$p5$	0.1566	$m7$	0.2903		
$Q_{d\,7-day\,low\,flow\,mean\,discharge\,summer}$	$p5$	0.3643	$m8$	0.3972						
$Q_{d\,7-day\,low\,flow\,mean\,discharge\,winter}$	$p1$	0.2531	$p2$	0.3538	$p5$	0.1804				
Avg_{Snow}	$p6$	0.0757	$p7$	0.1135	$p8$	0.6130	$p9$	0.1780		
$ETaTOT_{feb}$	$p5$	0.1212	$m2$	0.7505						
$ETaTOT_{jun}$	$p5$	0.1020	$m6$	0.6987						
$ETaTOT_{apr-sep}$	$p1$	0.1099	$p2$	0.1198	$p5$	0.3980				