Turin Polytechnic

Master’s degree in Environmental Engineering
A.Y. 2020/2021
Graduation Session October 2021

Hydrological analysis of a karst system via Ensemble Kalman Filter

Relators:
Prof. Ilaria Butera
Prof. J. Jaime Gómez-Hernández

Candidate:
Alessandro Pansa – s268982

Summary

Thanks

Before going into the discussion of the thesis I want to dedicate these first lines to thank the people who contributed to the work.

A special thanks goes to the supervisor of this thesis, Professor Butera Ilaria, for the continuous and constant availability, the time dedicated, the countless contributions and fundamental ideas for the development of this work.

Thanks also to the co-relator, Professor J. Jaime Gómez-Hernández, for his listening skills, the wise advice, the clever ideas, and his understanding of a shaky English.

It is also necessary to remember and thank Professor Bartolomeo Vigna for making the data of the Bossea Scientific Station available, as well as all the staff who works to maintain the necessary equipment.

I dedicate a huge thanks to my family without whose support today I would not even be here and to my girlfriend for the hours spent listening to me.

At last, a thought goes to my classmates and university colleagues with whom I have shared years and fundamental moments.

Thanks everyone!!

AP

1 Introduction

The thesis work presented below aims to use a tool belonging to the field of data assimilation, the algorithm of Ensemble Kalman Filter, for the estimation of a hydrological response characteristic of the karst system of Prato Nevoso-Bossea, in southern Piedmont. In particular, this method is used to statistically define a set of parameters capable of describing the state of the system through a series of equations typical of traditional hydrology. Those are applied here to an underground system, not entirely accessible and therefore largely not measurable.

This topic was chosen by one party for the personal interest of the student, on the other hand for its innovative vibe, which allows comparison with topics not usually dealt with in degree courses, but nevertheless an integral part of today's scientific landscape, such as the field of data assimilation. Furthermore, the study of aquifer systems takes on an increasingly leading role at a local and global level as the water resource gains greater importance for contemporary society. With this in mind, the work is intended to be part of the modern sensitivity towards this fundamental good.

In carrying out the work, an attempt was made to characterize the response of the karst system in the face of a precipitation event, as an instantaneous unit hydrogram and infiltration coefficient. In addition to all this a study of the snow melting process and its contribution turned out to be necessary. Ultimately, the system state thus defined allows the estimation of an outgoing flow, the Mora stream one, which crosses the stone halls and corridors of the Bossea Cave. The comparison of the estimated flow rate with that measured on site constitutes the basis of the applied method. Weather data are gathered by ARPA Piemonte, in particular at the Borello station in Frabosa Soprana, while the hydraulic and geologic data are gathered by the Bossea Scientific Station, under the aegis of CAI of Cuneo and Turin Polytechnic through the DIATI department.

The paper presents a first part of introduction to the area under study, a description of the karst system and measurement stations. There follows a chapter dedicated to the theoretical foundations from which we start for the work carried out. Finally, the result obtained and the process that led to it in its salient passages are reported as well as some analysis of the results.

All the following algorithms described and utilized are implemented in the MATLAB environment.

2 The Water System of the Bossea Cave

In the following pages there is a summary of the main traits of the area studied, for further information please refer to [1].

2.1 Geographic framework

The Bossea Cave is located in the municipality of Frabosa Soprana, in the Province of Cuneo. The entrance to the cave opens near the hamlet of Fontane, in Val Corsaglia, at an altitude of 836 m above sea level. In Figure 2-1 is shown its location in the territory of the Piedmont Region (IT).

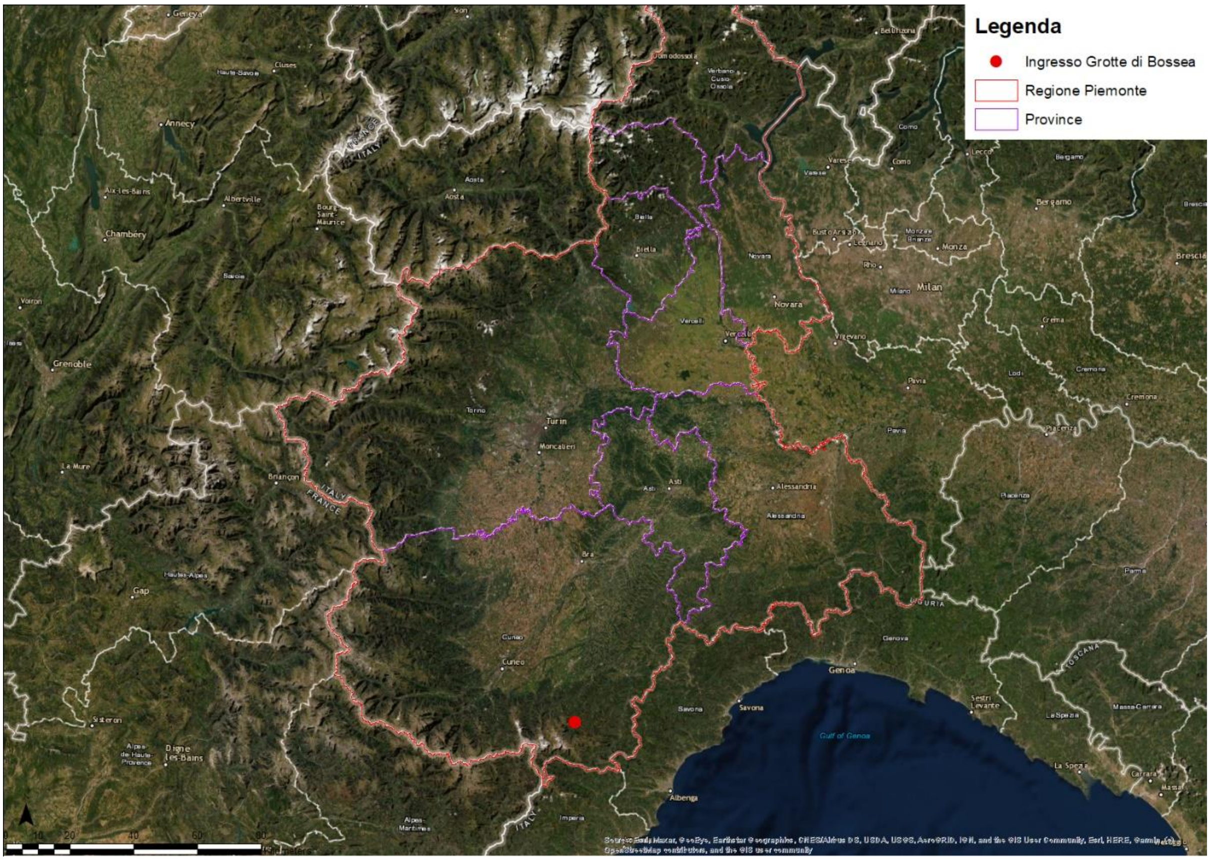


Figure 2-1: Location of the Bossea Cave inside the Piedmont Region.

The halls and galleries of the cave are set in the Mondolé-Artesinera-Bossea karst area, in the municipalities of Frabosa Sottana and Frabosa Soprana. The entire area is approximately included between the Corsaglia stream to the east, the Maudagna stream to the west, Mount Malanotte to the north and Cima Artesinera to the south; with a variable altitude between 800 m asl of the Corsaglia riverbed and 2382 of Monte Mondolé. The entire area is distributed across the watershed between the Maudagna and Corsaglia streams [1].

On the ground can be found an alternance of carbonate, permeable rocks, and impermeable surfaces, above which the water of precipitation streams, to be then absorbed on the limestones. In the summit area, a series of karst basins and plateaus drains the water into the subsoil.

The karst phenomena are typical and extensive in the whole area, with a considerable develop of underground water circulation. The underground flow lines can develop for several kilometres, interspersed with numerous caves. The main water systems in the area are Prato Nevoso-Bossea, Prato Nevoso-Case Bergamino, Artesinera-Stalle Buorch, Dolly-Artesina [1].

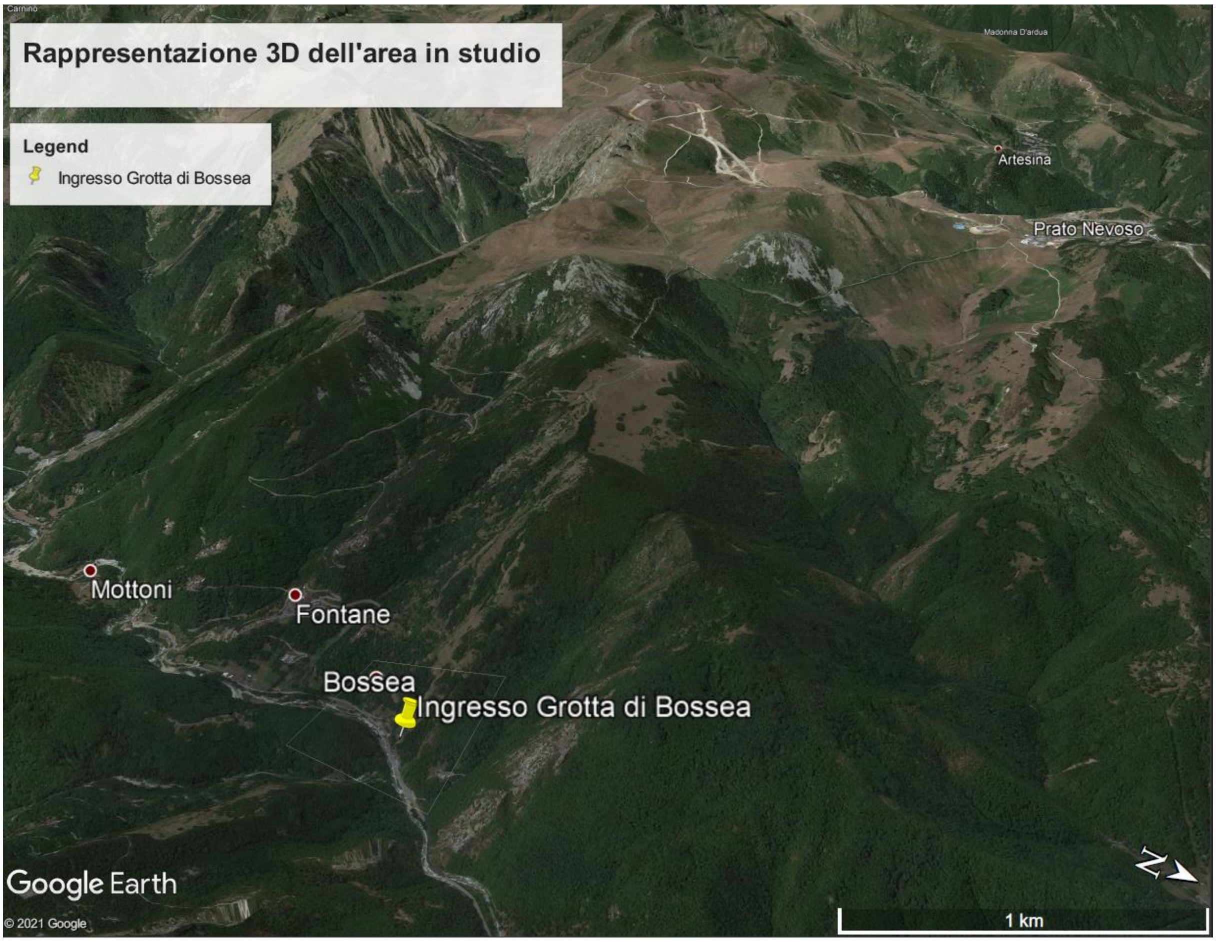
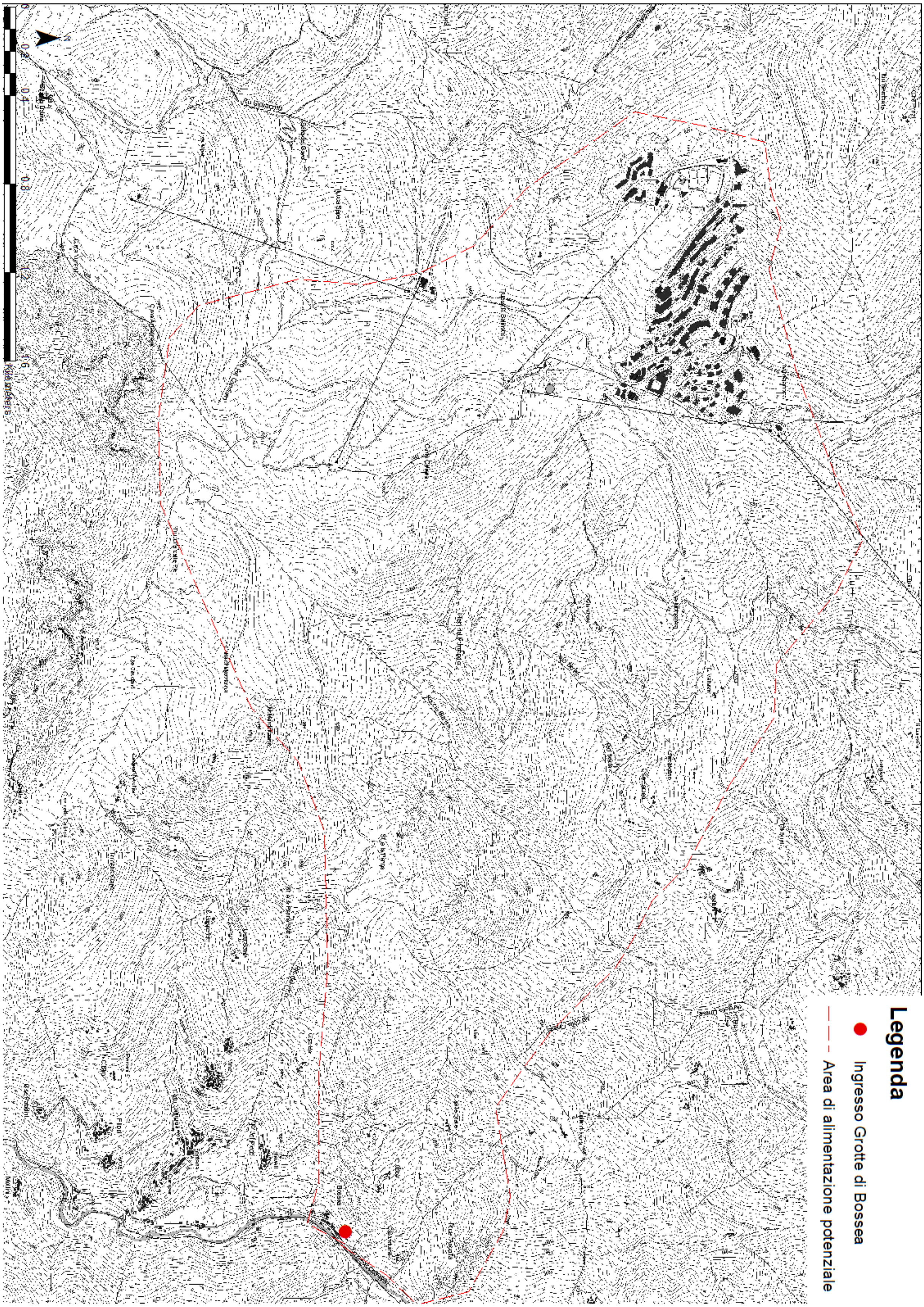


Figure 2-2: 3D model of the studied area, from Google Earth.

In the next page a map is shown displaying the infiltration basin of the karst system, as it has been identified since now, on the Regional Technical Map.


Legenda
Ingresso Grotte di Bossea
Area di alimentazione potenziale

2.2 Climatological framework

The Mediterranean Sea is about 40 km far as the crow flies, and it has an enormous influence on the Ligurian Alps through the influx of air masses of marine origin and full of moisture. These cause copious snowfalls in winter and spring, while in summer and autumn they take form of violent thunderstorms with abundant rainfall, short time extent and sometimes hailstorms. In the summer we usually have the dry season, although rainfalls are still present, therefore normally not droughty. In august the fog can persist in the absorption area for long times. The precipitations follow a typically Mediterranean trend, with peaks in the late summer and autumn (see Figure 2-4).

The water supply of the karst system turns out to be a pluvio-nival one, with contributions caused by the progressive melting of the snow mantle in the spring, to which are added the rainy precipitations, and a totally pluvial income in the fall period.

The mean annual precipitation calculated in the span 2001-2018 (only the years with complete rainfall data are considered) is equal to 1372 mm. this datum presents a strong variability, with standard deviation equal to 381 mm.

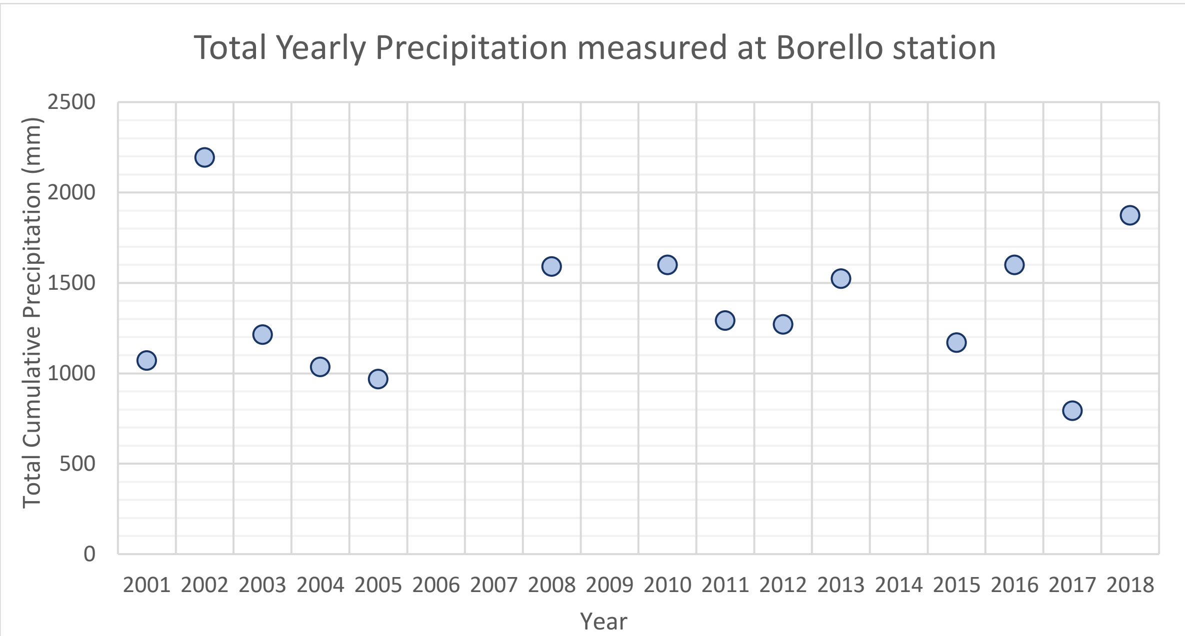


Figure 2-3: Cumulative annual precipitation at the Borello station, only the years whose data are complete are taken into consideration.

In the Figure 2-4 is possible to see how the monthly mean precipitations reach their peaks in the spring months, particularly May, and in autumn, particularly November. The summer is drought, as well as winter, during which, however, the precipitation often takes on a snowy form. Due to the stormy nature assumed by the summer-autumn rainfall, the data for this period shows greater variability than the rest of the year.

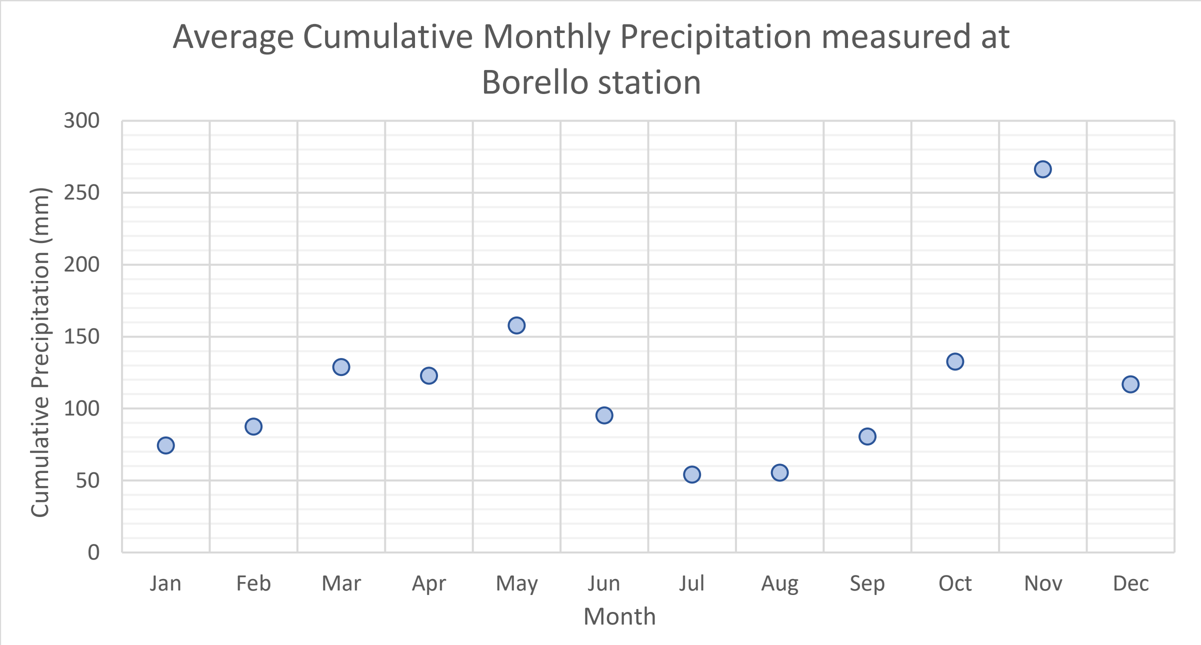


Figure 2-4: Average monthly precipitation at the Borello station in the period 2001-2018.

With regard to snowfall, a maximum height of snow during the season can be found between 50 and 250 cm, usually recorded in the late winter period. There is often a remarkable difference between the lower areas and the higher ones. The snow melting lasts for the entire springtime, starting in March, and continuing usually until May. It proceeds in ascending altitudes, affecting the sunny slopes, usually east and south, and those in the shade, north and west, in a different way.

2.3 Geological and morphological framework

The entrance to the cave opens onto the middle Val Corsaglia, engraved by the stream of the same name. The primitive glacial morphology was almost completely supplanted by the intense erosion of the watercourse. This phenomenon has two main causes: a change in the base level of the Tanaro river, of which the Corsaglia is a tributary, following its capture and a recent uplift of the entire alpine sector in question.

The slopes above the entrance have strong steepness and reduced coverage. The rocks show marked fragmentation linked to gelling. It is also possible to identify some limited sub-horizontal openings, relict galleries, once fully loaded, traces of the ancient base level.

Towards the west, the karst absorption areas open up, consisting of valleys dug by the erosion of temporary waterways. An alternation of waterproof quartzites can be identified, supplanted by limestone where the water infiltrates to reach the watershed. Karst soils are partially covered by insoluble residues on which vegetation grows. The portion of land that finally meets towards Prato Nevoso sees gentler karst forms with a grassy cover, where there are some absorbent sinkholes.

Looking through a geologist's eye, the entire area is included in the Brianzonese-Ligurian series. particularly, the Navonera-Bossea-Prel can be identified [2], limited by important tectonic lines (approximatively in a E-W direction) and characterized by strong compressions, folds and scales, with subvertical faults.

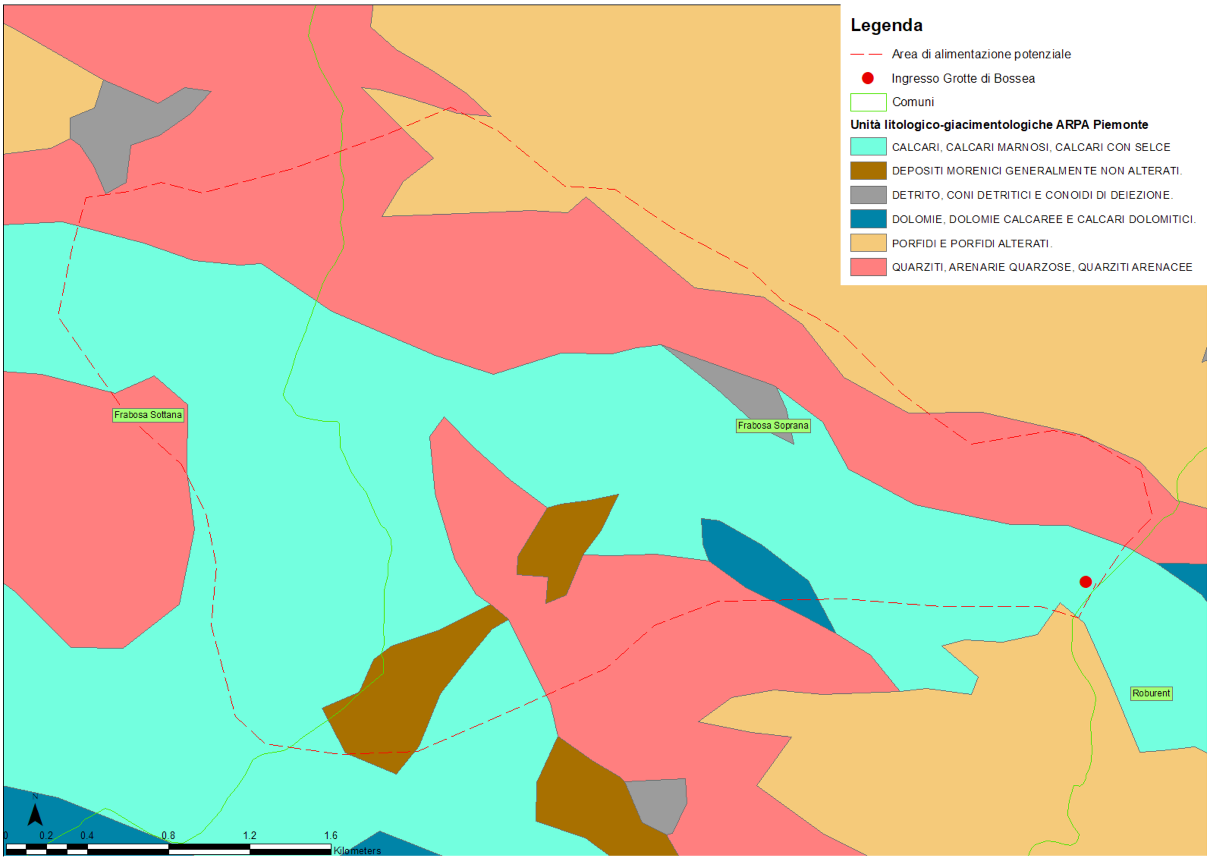


Figure 2-5: Lithologic-giacimentologic map. Font: ARPA Piemonte.

The Brianzonese-Ligurian series here in Bossea is constituted by a volcano-clastic basal succession (porphyroids, quartzites and pelites) ed a limestone-dolomite sequence (dolomites, limestones and schist limestones).

2.4 Basics of hydrogeology

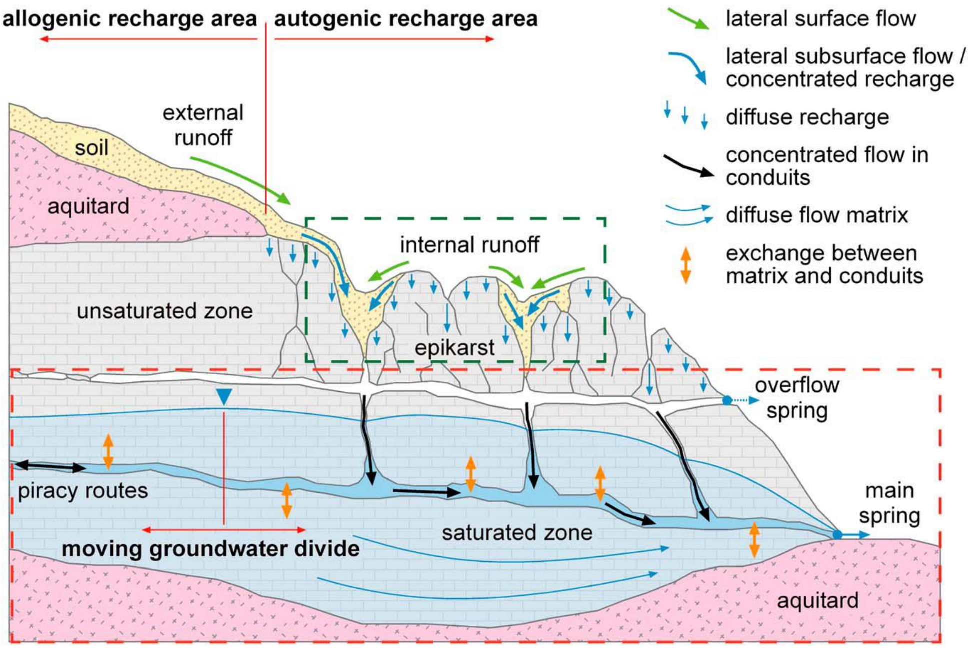


Figure 2-6: Conceptual model of karst inclusive of all characteristic phenomena. In dashed green the area of the epikarst or superficial karst, in dashed red the underground water system [3].

From general point of view, the karst phenomenon involves the erosion of carbonate rocks by the water that flows inside them according to the balance of (2.1).

$$CaCO_3 + H_2O + CO_2 \leftrightarrow Ca^{2+} + 2HCO_3^{2-} \qquad 2.1$$

This phenomenon tends to create a series of drainage paths in the superficial part of the carbonate platform from which there is a widespread infiltration, while in the areas of contact with impermeable soils the runoff is conveyed to points of concentrated infiltration. A transfer occurs when the water is released from a secondary aquifer.

Below the epikarst there is a transfer zone, in which the water flows into fractures or tunnels, generally not fully loaded, and then reaches the saturated zone, where there is a mostly sub-horizontal transfer, up to emerging areas. Fossil tunnels, an indication of ancient transfer areas in conditions different from the current one, can sporadically reactivate during periods of floods.

Returning specifically to the Bossea area, the soils affected by the volcanic succession constitute important permeability thresholds, limiting the carbonate structures. Along the main tectonic lines there can be strongly fractured and cataclasated areas. In these areas, especially in quartzites, non-negligible secondary aquifers can be set up, which feed the karst aquifer through underground transfer.

The karst, housed in the carbonate series, is the main aquifer in the area, with high capacities in terms of ingestion and transport. In particular, a network of fractures with different degrees of karstification allows the transfer of the inflows along a sub-vertical direction until the saturated zone is reached. This seems to be at an altitude of about 940 m asl, while the emerging altitude is about 810 m. The presence of a saturated area so high is explained on the one hand by the evolution of the system and the significant lowering of the local base level and on the other by the existence of some permeability thresholds within the complex.

2.5 The Bossea Cave and the Scientific Station

The discovery of the cave dates back to 1850 and in the following years the first section had been explored. In 1948 the first lighting system was built and in the subsequent period the explorations of the emerged galleries were completed. In 1968 and 1995 part of the underwater tunnels were explored. Access is guaranteed through a partially excavated fossil gallery to facilitate entry.

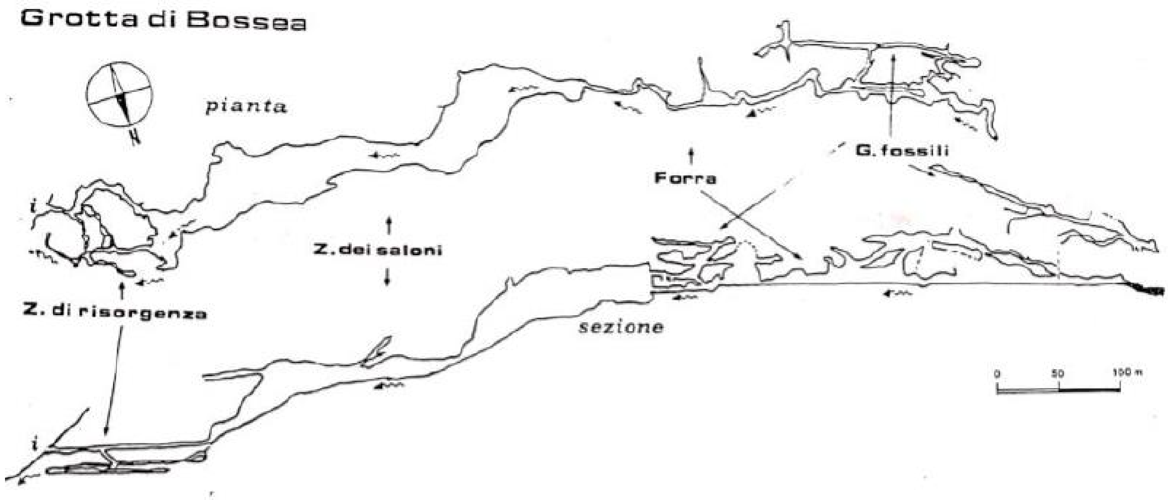


Figure 2-7: Plan and schematic section of the Bossea Cave [1].

The entrance to the cave opens onto the western side of Val Corsaglia at 836 m asl, about 30 m above the torrential bed. It develops in the WNW-ESE direction for a total of 2638 m and a total height difference of +184 m. A perennial stream, the Mora one, runs through it, whose flow rates oscillate between 50 and 1200 l/s.

Along the development of the cavity, five different areas can be distinguished, as indicated in [1]:

- The **resurgence area** characterized by narrow horizontal tunnels with full load morphology and set on three levels. The lower one, crossed by the Mora Stream, the intermediate gallery, active in periods of flooding and the upper gallery through which you have access to the cave and now no longer active. The outflows from this section directly feed the Corsaglia Stream with flows into the riverbed (identified thanks to a thermal camera).
- The **halls' area** is set along a plane of contact between vulcanite, often cataclasated, and limestone. The development is about 550 m on a steep slope. There is a marked collapse morphology with the succession of different halls, the largest of which, the Garelli Hall, extends for 100 meters, is 40 meters high and 60 meters wide. The stream runs deeply recessed on the bottom between cyclopic boulders, high it is possible to notice wrecks of small pipes under pressure. Phases of partial filling of the cavities are evidenced by large clasts partially cemented by calcite on the walls. In the terminal area there is the Ernestina Lake and immediately upstream the 9 meters high waterfall of the same name. In this area there is a shrinkage due to the progressive reduction of porphyroids, whose erosion has conditioned the genesis of the previous salons. Beyond the waterfall, the next section opens up, the gorge.
- • The **gorge** is a tunnel set on a series of vertical fractures following a horizontal course. It extends for 400 m with variable widths between 2 and 4 m and heights up to 30 m. the Mora Stream runs through it in its entirety, with sections in which it has deepened to create small basins. Within this area, a weir was built to measure the flows of the Mora Stream, resulting in a rise in the upstream level and the formation of a single large lake of about 120 m. On the ceiling of the gorge there are relics of large, pressurized ducts, sometimes partially eroded, while on the walls a series of brackets testifies to the subsequent phases of excavation and corrosion.
- • The **fossil tunnels** overlook the gorge and are composed of large fully loaded tunnels where huge concretionary deposits are often present. Different groups are distinguished, such as the Paradise Galleries and the High Branches. In the final part of the cavity, we find the Galleries of Marvels, about 200 m long, which end in the siphon of the Dead Lake at an altitude of +184 m. Beyond they become impassable, occluded by imposing concretions.

• The **submerged area** includes two siphons. A submerged pipeline connects the Dead Lake to a shortly emerged stretch, from here a vertical well leads to the first siphon, consisting of a large hall and a gallery of about 90 m. Then follows the second siphon, only partially explored.

Figure 2-8: Garelli Hall. Photo by A. Morabito.

Inside the Bossea Cave there is a permanent scientific research station, managed by the Bossea Scientific Station of the Cuneo CAI and by the DIATI of the Turin Polytechnic, in collaboration with ARPA Piemonte and ARPA Valle d'Aosta. There are numerous research fields operating in this context, such as biospeleology, geochemistry and hydrochemistry, karst hydrogeology, underground meteorology, natural radioactivity, topographic surveys, hydraulics, and computer processing.

The research station is divided into the *main laboratory*, in the halls area (divided into the Physical-Chemistry and Biospeleology sections) and the *advanced laboratory* in the gorge, fully automated and dedicated to the detection of hydrogeological and meteorological parameters in the less accessible zones. They are followed by a number of peripheral stations and a network of sensors spread over the entire cave.

Thanks to the data measured by the laboratory, it is possible to know the flow rate of the Mora stream, measured at the weir above, thanks to a flow curve (2.2) specially calibrated that allows you to make an indirect measurement of the flow rates by measuring the height of water in the basin upstream of the measuring weir.

$$Q = 0.0647 \cdot y^{1.5} \tag{2.2}$$

- Q flow rate in l/s.
- y water height in mm.

The measurement is carried out automatically, on an hourly basis starting from 2008 and daily in the previous period starting from 1982.

Figure 2-9: Bear Hall and entrance to the scientific station (on the right side of the image). Photo by A. Morabito.

2.6 Borello weather station

In Borello-Case Cane locality, in the municipality of Frabosa Sottana, there is a meteorological station (code 310, altitude 1005 m asl) managed by ARPA Piemonte. The precipitation and air temperature data are measured on a daily scale since 1997. The same data are also available on an hourly scale since 2001. A heated rain gauge is used, which can also measure the equivalent in mm of water of fallen snow.

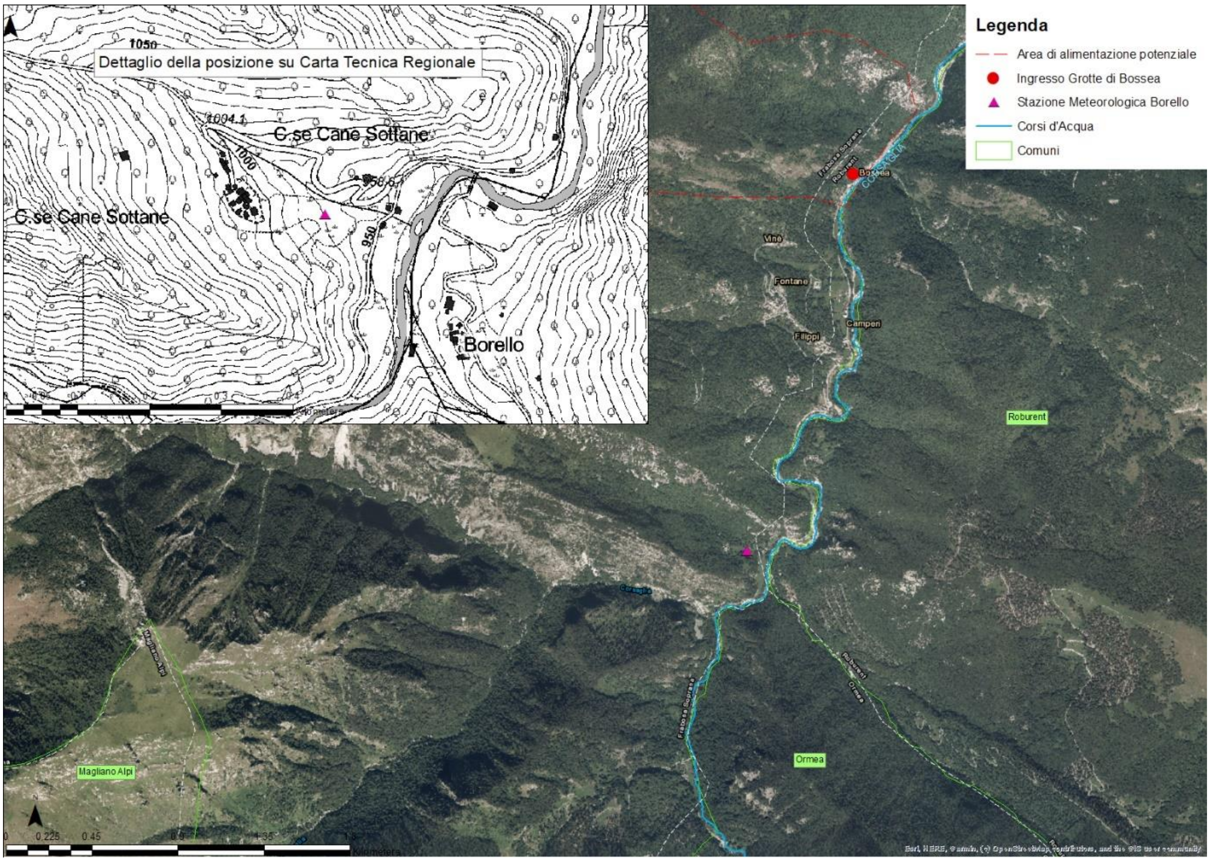


Figure 2-10: Location of the Borello weather station, within the middle Val Corsaglia and in detail.

The distance from the entrance to the Bossea Cave is about 2 km as the crow flies. The precipitation data used later are those recorded at this station, and they are made available by ARPA Piemonte.

In the area there is also the Monte Malanotte station in Prato Nevoso, however the unfortunate positioning led to the non-use of the data.

3 Hydrology Basis

3.1 The inflow-outflow transformation

The purpose of this study is to define a hydrological response of the physical system consisting of the Bossea karst aquifer in the face of a precipitating event. The main source of the following is the [4]. Three fundamental hypotheses are posed in the modelling of the physical behaviour of the system:

1. 1. Uniformity of precipitation on the feeding area
2. Stationarity of physical processes, they do not vary over time
3. Linearity of the response of the system, for which the superimposition of the effects is valid

The first hypothesis is reasonable given the relatively modest extension of the supply basin. Hypotheses 2 and 3 will be discussed later.

The precipitation that falls on the feed area can be divided into two different inflows: the part infiltrated into the subsoil and feeding the aquifer and the part run off on the surface that enters directly into the surface basin without contributing to the flow inside the cave.

The flow rate that feeds the system is therefore:

$$I(t) = \chi \cdot A \cdot \frac{p(t)}{\Delta t} \tag{3.1}$$

Where *p(t)* is the measured precipitation at time *t* and discretized on an interval Δt, *A* is the basin's area (in m^2) and χ the infiltration coefficient, dimensionless.

At this point, a system response is assumed according to the convolution theory of hydrograms [4]:

$$Q(t) = \int_0^t I(\tau) \cdot h(t-\tau) d\tau \tag{3.2}$$

Where *Q(t)* is the outcoming flow rate, *I(t)* is the incoming flow rate and *h(t)* is the transfer function or Instantaneous Unit Hydrogram, IUH. This response is linear (hypothesis 3) and is based on the superposition of the effects, as well as being invariant over time (hypothesis 2).

Furthermore, the continuity equation (the input volume must equal the output volume) also implies:

$$\int_0^\infty h(t)dt = 1 \qquad 3.3$$

The transfer function therefore exists within the set [0,1].

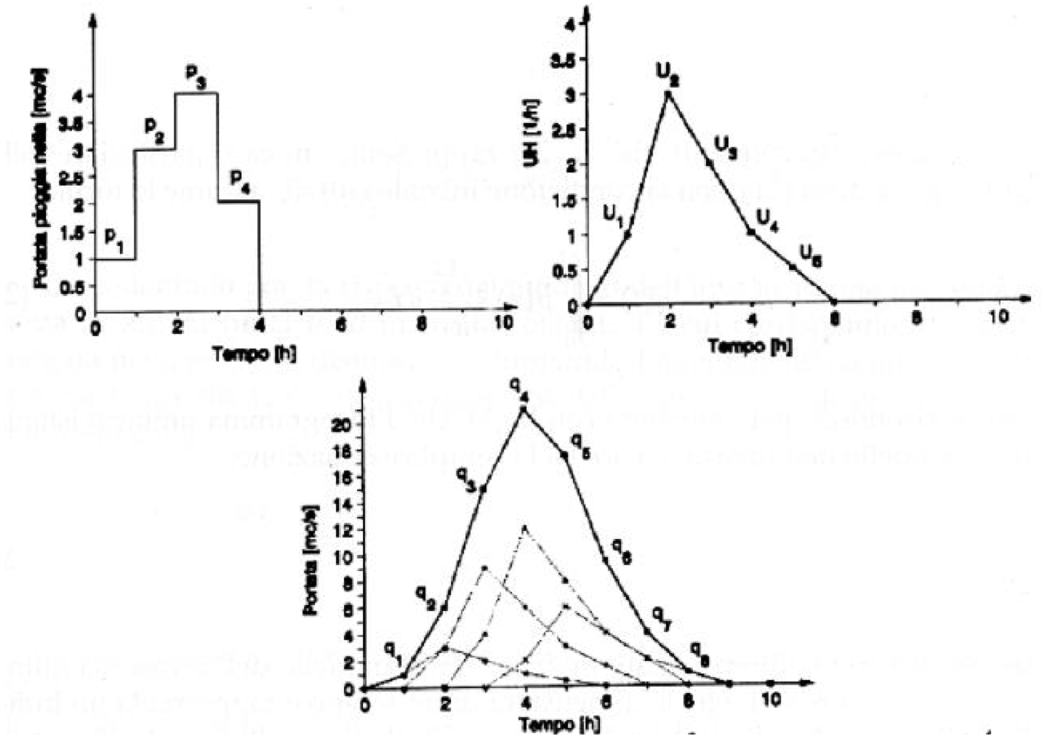


Figure 3-1: Scheme of the hydrological model using the IUH.

The total length of the IUH, T_h, it is comparable to the basin's corrivation time.

In practical terms, the IUH constitutes the measurable flow rate at the closing section in the face of a unitary inflow, from (3.1). It is assumed that each drop falling into a fixed point of the basin takes a fixed time to reach the closing section and the run-off time is equal to the longest travel time. The theoretical duration of the calculated flow rates is equal to the duration of the meteoric event T_p to which the maximum time of the IUH, T_h must be added.

Given the discrete form of the available data, the convolution integral of the (3.2) is substituted with a summation:

$$Q_k = \sum_{j=1}^{k \leq m} I_j \cdot h_{k-j+1} \cdot \Delta t \qquad 3.4$$

where *m* is the number of elements of the IUH and *n* the number of elements of *I*. *Δt* is the discretization interval, in our case it is unitary, because it is equal to a day or an hour.

3.2 Nash-Sutcliffe efficiency coefficient

The Nash-Sutcliffe efficiency coefficient [5], or NSE, aims to evaluate the predictive capacity of a hydrological model, according to the formula:

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_0^t)^2}{\sum_{t=1}^{T}(Q_0^t - \overline{Q_0})^2} \qquad 3.5$$

Where we have the time, t, from 1 to T; the measured flow rates, Q_0^t, of average value $\overline{Q_0}$; the predicted flow rates thanks to the model, Q_m^t.

This index varies between 1 (the model perfectly mimics reality) ands $-\infty$ (the model is unable to predict reality in any way). An NSE value of 0 indicates a predictive capacity of the model equal to that of the average of the observations. In this way we can quantitatively evaluate the goodness of our result by comparing it with real data.

It is also possible to calculate a normalized value, NNSE (Normalized Nash-Sutcliffe Efficiency), according to the formula:

$$NNSE = \frac{1}{2 - NSE} \qquad 3.6$$

Where $NNSE = 1$ is equivalent of $NSE = 1$, $NNSE = 0.5$ to $NSE = 0$ and $NNSE = 0$ to $NSE = -\infty$. This is useful for eliminating the lower limit to $-\infty$.

Often the efficiency is expressed in percentage terms.

4 Data Assimilation

4.1 The Kalman Filter (KF)

The Kalman Filter is deeply explored in [6], while in [7] and [8] the topic is treated as an introduction to the Kalman Filter Ensemble.

The problem consists in the best possible esteem of the values of a dynamic process in the face of a physical estimation model affected by error and a series of observations over time (and space) afflicted by instrumental error.

Is considered a stochastic process governed by the stochastic differential equation, the dynamic model:

$$x_k = Ax_{k-1} + w_k, \quad k > 1 \tag{4.1}$$

Where $x_k \in \mathbb{R}^n$ is the stochastic process and the index k denotes the dependence on the time instant. The initial state is ($k = 1$):

$$x_0 = \xi \sim N(x_0, P_0) \tag{4.2}$$

A measurement system is then added $y_k \in \mathbb{R}^d$, the observational model:

$$y_k = Hx_k + \epsilon_k \tag{4.3}$$

The error terms are:

$$w_k \sim N(0, Q_k) \tag{4.4}$$

$$\epsilon_k \sim N(0, R_k) \tag{4.5}$$

The hypotheses of the gaussianity of the errors are therefore posed (from 4.4 and 4.5) and linearity of both models. Under this hypothesis, the Kalman Filter is the optimal solution to recursively estimate the first and second moments of the stochastic distribution, following a sequential, iterative process that alternates an *a priori* estimation phase of the value x_k to one of data assimilation. Following the notation in [8] we will define *forecast* the first step, with an associated apex *f*, and *analysis* the latter, with apex *a*; $x^{f/a}$ will be the average of the distribution and $P^{f/a}$ the covariance matrix.

Forecast Step

$$x_k^f = Ax_{k-1}^a \tag{4.6}$$

$$P_k^f = AP_{k-1}^a A^T + Q_k \tag{4.7}$$

Analysis Step

$$K_k = P_k^f H^T \left(H P_k^f H^T + R_k\right)^{-1} \tag{4.8}$$

$$x_k^a = x_K^f + K_k (y_k - H x_k^f) \tag{4.9}$$

$$P_k^a = (\boldsymbol{I} - K_k H) P_k^f \tag{4.10}$$

The matrix $K_k \in \mathbb{R}^{m \times d}$ is the *Kalman gain*, whose coefficients represent the linear combination of optimum between the *a priori* estimate and observation. The solution produced by the KF can be defined BLUE, Best Linear Unbiased Estimator. The process stops at a certain time *T*, after a series of cycles.

4.2 The Ensemble Kalman Filter (EnKF)

This variation of the algorithm, introduced by Evensen in 1994 [9], is considered a KF of Monte Carlo type. Shortly, it is proposed to extend the applicability of the filter to a non-linear system (4.11), whose statistical moments are difficult to estimate, through N realizations of the vector x_k representing the state of the system. There are again the observations (4.12).

$$x_k = f(x) + w_k \tag{4.11}$$

$$y = H x_k + \epsilon_k \tag{4.12}$$

An ensemble of realizations is therefore obtained:

$$E^{f,a} = \left[x_1^{f,a}, \dots, x_N^{f,a}\right] \in \mathbb{R}^{m \times N} \tag{4.13}$$

With mean:

$$\bar{x}^{f,a} = \frac{1}{N} \sum_{n=1}^{N} x_n^{f,a} \tag{4.14}$$

And covariance matrix (approximation of the real covariance matrix of the distribution):

$$(P^e)^{f,a} = \frac{1}{N-1} \left[x_1^{f,a} - \bar{x}^{f,a}, \dots, x_N^{f,a} - \bar{x}^{f,a}\right] * \left[x_1^{f,a} - \bar{x}^{f,a}, \dots, x_N^{f,a} - \bar{x}^{f,a}\right]^T \tag{4.15}$$

Furthermore, a vector of observations perturbed by the error is necessary:

$$y_n = y + \epsilon_n, \quad 1 \le n \le N \tag{4.16}$$

$$Y = [y_1, \dots, y_N] \in \mathbb{R}^{d \times N} \tag{4.17}$$

Whose covariance matrix results in:

$$R_e = \frac{1}{N-1}[Y-\bar{Y}] * [Y-\bar{Y}]^T = \frac{1}{N-1}\epsilon * \epsilon^T \qquad 4.18$$

At this point, the algorithm resumes the two basic steps described above:

Forecast Step

$$E^f = f(E) + W, \quad W = (w_i) \qquad 4.19$$

Analysis Step

$$E^a = E^f + (P^e)^f H^T [H(P^e)^f H^T + R_e]^{-1}[Y - HE^f] \qquad 4.20$$

There may be problems in the inversion of the matrix $[H(P^e)^f H^T + R_e]$, covered exhaustively in [10].

4.3 The Ensemble Smoother with Multiple Data Assimilation (ES-MDA)

To obtain a better result, the algorithm of the Ensemble Smoother with Multiple Data Assimilation was finally opted. It allows you to assimilate the same data several times unlike the Smoother Ensemble which uses all the observations in a single step one time. The complete series of measurements will then be assimilated for a certain number of cycles to be defined *a priori*, making corrections gradually more marked each time. For further information see [11].

Given the *smoother* update equation:

$$E^a = E^f + C_{xy}\left(C_{yy} + C_{dd}\right)^{-1}(Y - g(E^f)) \qquad 4.21$$

In which Y represents the observations, E^f the set of realizations of the vector containing the model parameters, $g(E^f)$ the data estimated at the position and time of the observations through the set of parameters, C_{xy} the covariance between the parameter matrix and the estimated data matrix, C_{yy} the self-covariance of the estimated data, C_{dd} the covariance matrix of measurement errors.

The ES-MDA proposes to repeat the *update* step thus structured cyclically, assimilating the same data several times. To avoid problems in the inversion of the matrix $\left(C_{yy} + C_{dd}\right)^{-1}$, which would tend to singularity in this condition, a method of inflation of the covariance matrix of the instrumental error is used, in such a way as to dampen the change made to the model.

We then move from the one major correction made by the *smoother* to several minor corrections made to each cycle of the algorithm.

In this sense it is necessary to introduce a vector of coefficients α_i necessary for this operation, such that:

$$\sum_{i=1}^{N_a} \frac{1}{\alpha_i} = 1 \tag{4.22}$$

Where N_a represents the number of cycles arbitrarily decided.

The *update* step becomes:

$$E_{i+1}^{a} = E_i^{f} + C_{xy,i}\left(C_{yy,i} + \alpha_i C_{dd}\right)^{-1}\left(y_i + \sqrt{\alpha_i}\epsilon - g(E_i^{f})\right) \tag{4.23}$$

In which ϵ is the error covariance matrix and y_i the observations at time i.

The construction of the vector α happens in the following way: you impose a value for α_1(usually equal to 1) and a value of α_{geo}, that controls the change between one step and another. At this point we can calculate:

$$\alpha'_{i+1} = \frac{\alpha'_i}{\alpha_{geo}} \tag{4.24}$$

And then:

$$\alpha_i = \alpha_i \cdot \left(\sum \frac{1}{\alpha'_i}\right) \tag{4.25}$$

So, we can ensure the condition in (4.22).

It is remarkable how the parameter $\alpha_{geo} = 1$ means a uniform change from one cycle to another, while $\alpha_{geo} < 1$ means a decreasing update and on the opposite $\alpha_{geo} > 1$ means an increment of the gain between cycles.

A second approach useful for generating the error's inflation factors α_i is proposed in [12], to which reference should be made for the theoretical discussion, while the main passages will be summarized below.

First of all, the number of cycles N_a must be imposed.

Defined the matrices (n represents the n^{th} iteration):

$$\Delta D^n = \frac{1}{\sqrt{nr-1}}[Y - \bar{Y}] \qquad n \in [0, N_a] \tag{4.26}$$

$$C_e = \boldsymbol{I} \cdot R_k = C_{dd} \tag{4.27}$$

$$A = C_e^{-\frac{1}{2}} \Delta D^0 \tag{4.28}$$

A singular value decomposition of the matrix A is applied:

$$A = U\Sigma V \tag{4.29}$$

with σ_i i^{th} element of the singular values vector Σ.

The following calculation is then carried out:

$$\alpha_1 = \max\{\bar{\sigma}^2, N_a\} \tag{4.30}$$

And a factor $\gamma \in (0,1]$ is defined by solving with the bisection method:

$$f_1(\gamma) = \sum_{k=1}^{N_a} \frac{1}{\gamma^{k-1} \cdot \alpha_1} - 1 = 0 \tag{4.31}$$

Finally, the inflation factors are calculated according to the formula:

$$\alpha_{k+1} = \gamma^k \cdot \alpha_1 \tag{4.32}$$

4.4 Of Inbreeding, Covariance Inflation e Damping Factor

The inbreeding is a phenomenon that occurs in Monte Carlo-type statistical filters when the set representing the stochastic variability of system's state collapses to a unique solution, behaving in a deterministic and non-statistical manner. The problem is known and dealt with in the literature, for example [13], [14], [15]. As indicated in these sources it is possible to measure the validity of the result in terms of inbreeding thanks to RMSE, *Root Mean Square Error*) and ES, *Ensemble Spread*:

$$I = \frac{RMSE}{ES} \tag{4.33}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^{meas} - \langle y^a \rangle_i)^2} \tag{4.34}$$

$$ES = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2} \tag{4.35}$$

where N is the number of elements in the model, y_i^{meas} the measurements, $\langle y^a \rangle_i$ the average of the guessed flow rates; σ_i the variance of the ensemble.

The Covariance Inflation algorithm and the Damping Factor are used to avoid the collapse of the filter in a single solution, as explained in the sources cited above.

The first is applied as in [16], a factor $\lambda > 1$ is defined and it acts on the ensemble matrix:

$$E_{inf}^{f} = \overline{E^f} + \lambda \cdot (E^f - \overline{E^f}) \tag{4.36}$$

While the *damping factor*, let's call it $\beta \in (0,1]$, is applied in the u*pdate* step of the ES-MDA:

$$E_{i+1}^{a} = E_i^f + \beta \cdot C_{xy,i}\left(C_{yy,i} + \alpha_i C_{dd}\right)^{-1}(y_i + \sqrt{\alpha_i}\epsilon - g(E_i^f)) \tag{4.37}$$

The value of the two parameters is adjusted from time to time to ensure correct filter operation.

5 Application of the Ensemble Kalman Filter to the Case Study

In the following chapters we will try to present a summary that traces the salient steps in the development of the defined model proposed in this thesis and its functioning. The goal remains a set of parameters representing the state of the system, capable of receiving rainfall at the input and returning an outgoing flow. It is possible to make assumptions about the physical functioning of the system, while to estimate the value of these state parameters we use a statistical model, the data assimilation algorithms based on the Monte Carlo methods seen above.

5.1 The data

Incoming data are the precipitations, i, and the air temperature, tmp, measured at the Borello weather station and the flow rates, Q, measured at the Bossea Scientific Station, useful for data assimilation.

The complete series available covers the time span between 1-Jan-2001 and 31-Dec-2018. The values are measured on a daily basis. Within the series, the data are not continuous, but some windows are missing for technical reasons.

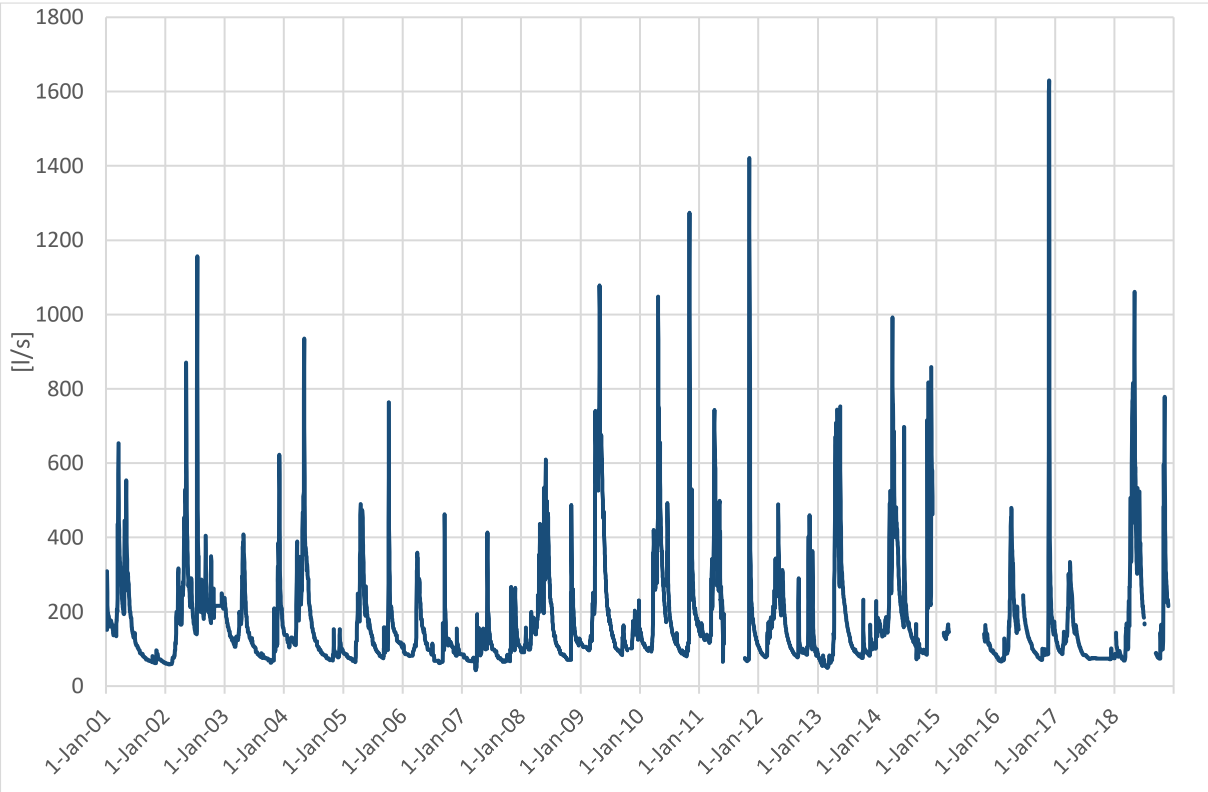


Figure 5-1: Series of flow measurements carried out on the Mora stream by the staff of the Bossea Scientific Station.

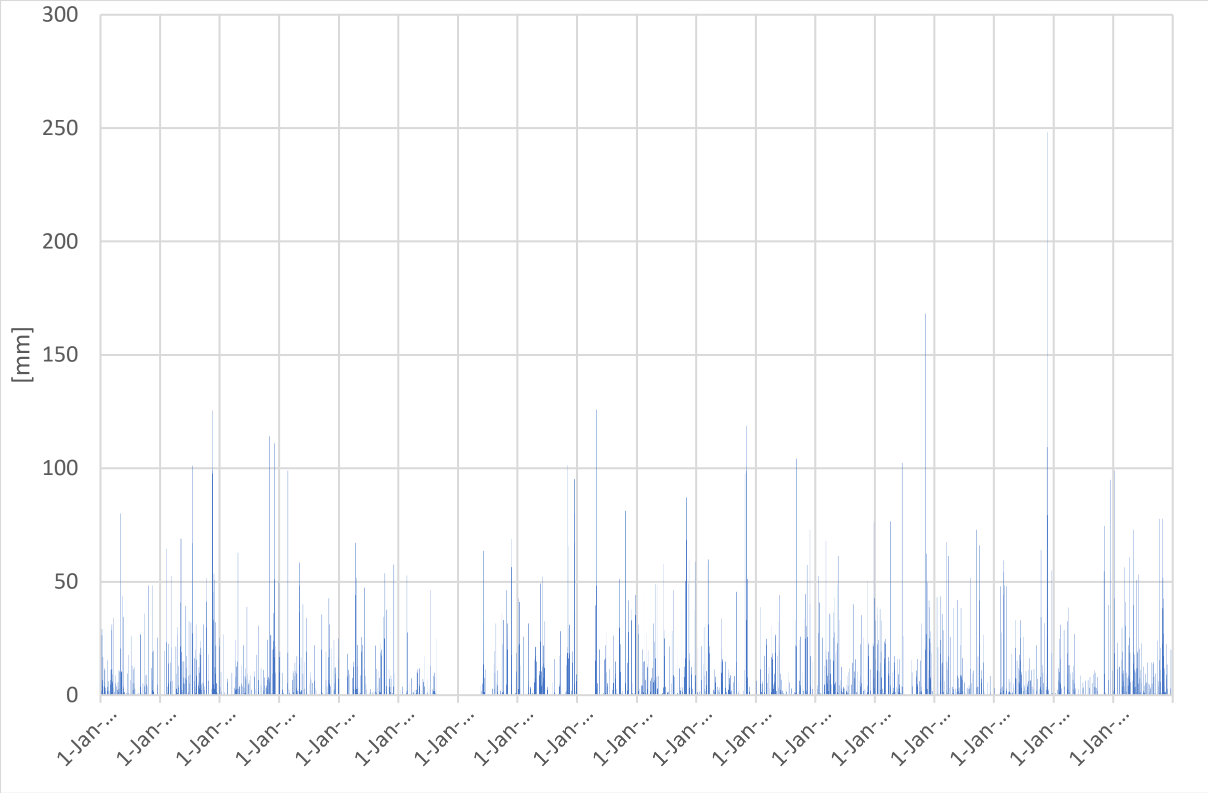


Figure 5-2: Series of precipitation measurements carried out at the Borello station (Frabosa Soprana) by the staff of ARPA Piemonte.

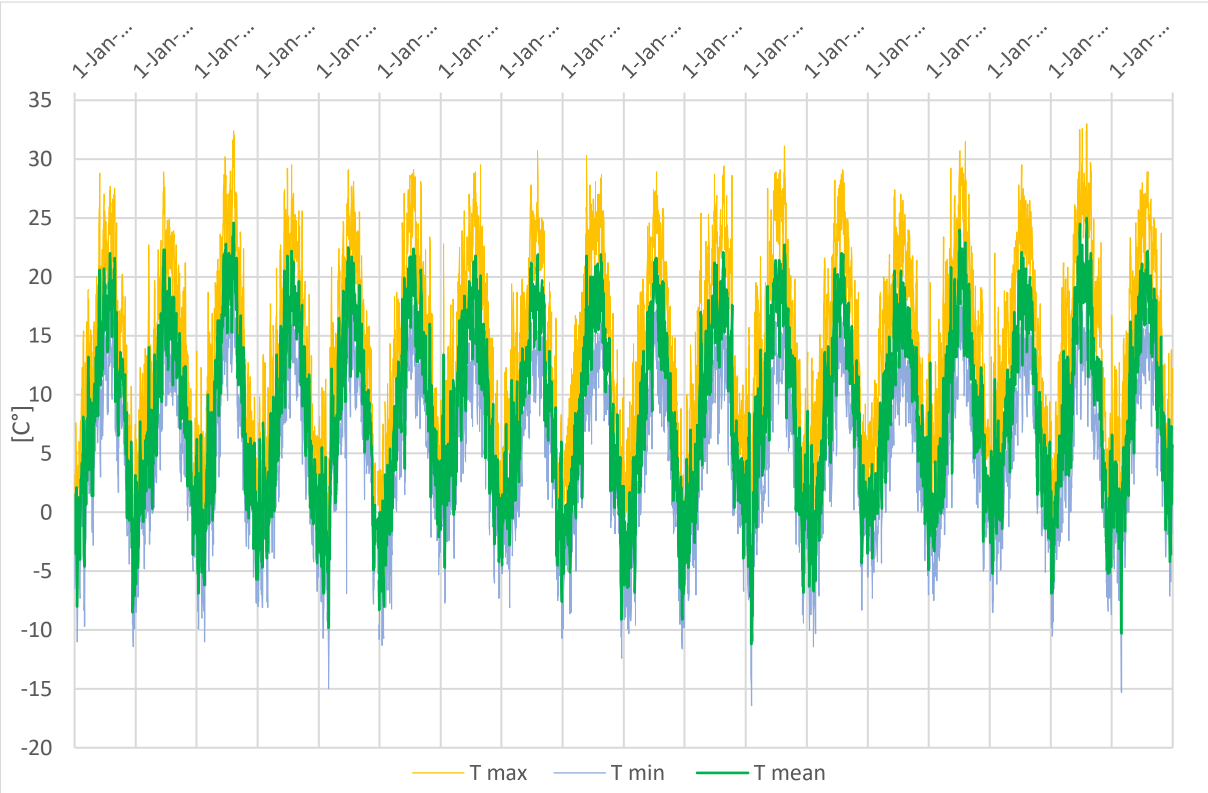


Figure 5-3: Series of air temperature measurements carried out at the Borello station (Frabosa Soprana) by the staff of ARPA Piemonte.

5.2 Measurement error variance

The flow rate measured thanks to (2.2) is affected by an instrumental error. The gauge precision in equal to $\pm 1\ mm$, from which it is possible to estimate the error on the flow rate by inverting the weir outflow law (2.2).

The link between the flow rate and the height of the free surface upstream of the weir is non-linear, so the measurement error will vary as the observed flow rate varies. We can first calculate the height of water measured, called here y, inverting the flow scale. Then we proceed to the construction of a flow vector affected by error by adding to y a millimetre ($Q' = f(y + 1)$). The measurement error will result from the difference between the estimated flow rates and those just calculated, that is:

$$err = Q' - Q \tag{5.1}$$

from which it is possible to calculate the variance of the measurement error useful in the applied algorithms:

$$\sigma^2(err) = \frac{err \cdot err^T}{L - 1} \tag{5.2}$$

With L length of the vector.

5.3 First attempt to estimate the IUH with EnKF

The first approach to solving the problem sees the application of the EnKF for the estimation of the characteristic IUH of the Bossea system.

We start by selecting the flow rate and precipitation data relating to a flood event. Once the suitable period has been identified, we proceed to subtract the minimum flow rate measured in this time interval, generally referred to the first day, in such a way as to consider only the runoff generated by the precipitation of the same period (an example can be seen in the Figure 5-4).

We then proceed with the calculation of the infiltration coefficient, χ, thanks to the relationship between the water flowing in and out according to the (3.1).

In this chapter and the following, reference will be made to a flood event that occurred in the spring of 2003 to illustrate the development of the work, although the code from time to time has been tested on multiple events.

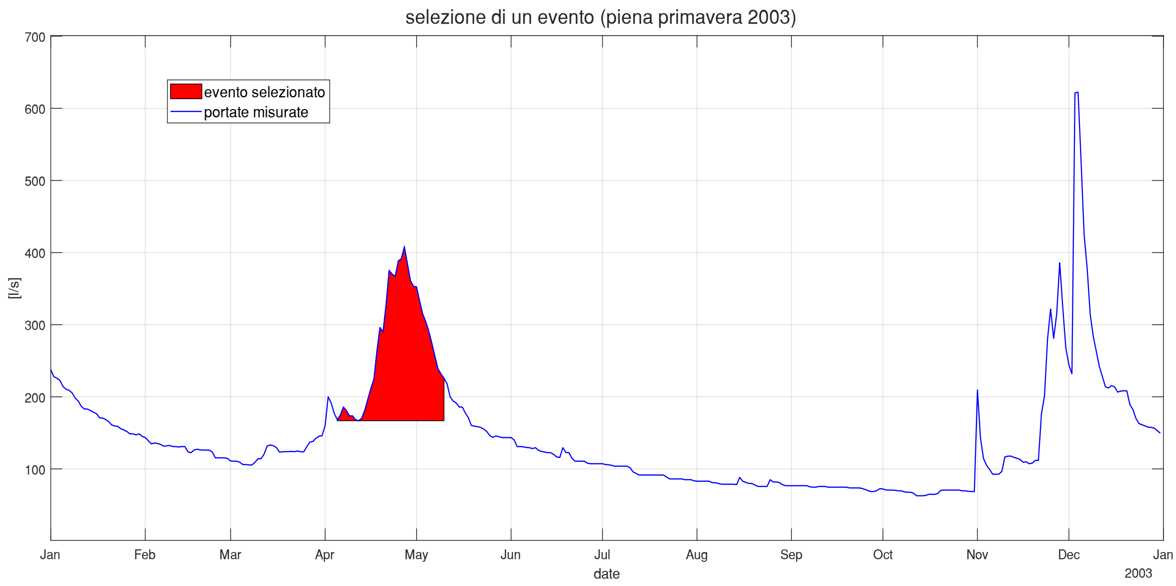


Figure 5-4: Selection of an event that took place in the spring of 2003.

In other words, the infiltration coefficient can be obtained from the calculation:

$$\chi = \left. \sum Q \cdot \frac{86400}{1000} \middle/ \sum \frac{i}{1000} \cdot A \right. \qquad 5.3$$

Where Q the flow rate measured in l/s, i the measured rainfalls in mm/h and A is the basin area, approximately 6 km^2.

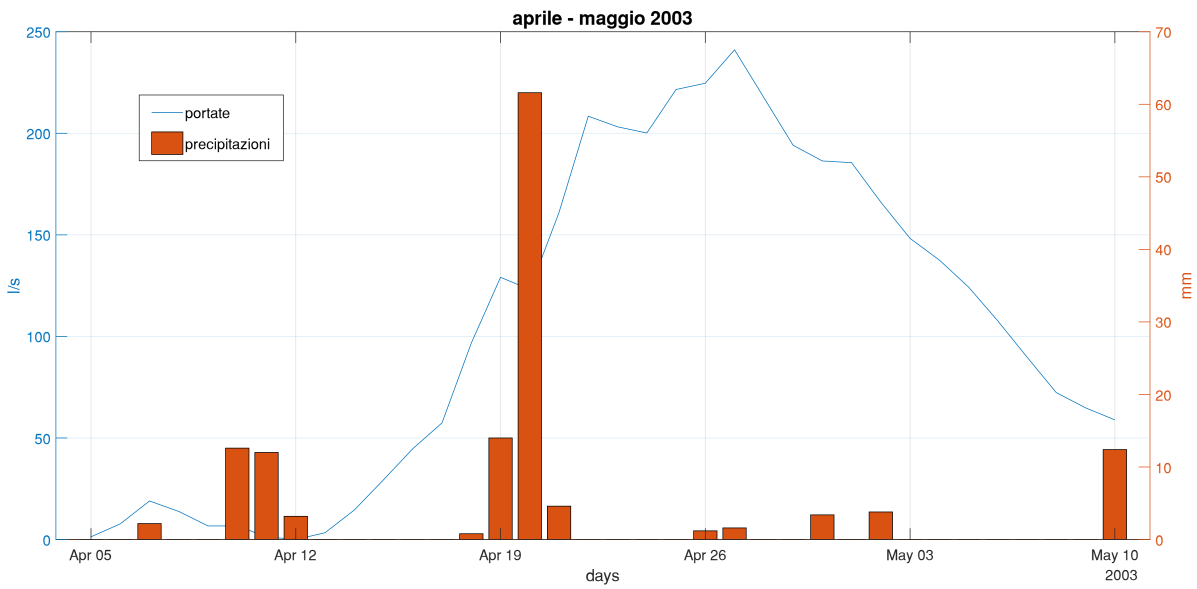


Figure 5-5: Example of data used for the selected event.

At this point it is possible to proceed with the initialization phase of the Ensemble Kalman Filter. It is necessary to define the number of vectors that will make up the stochastic set (in our case 1000) and to define the initial state of the system, a set of those vectors each representing an IUH. Their length is equal to the duration of the event under analysis minus two, because the final values would be unstable and poorly defined by the filter. In fact, if the event in question lasts 35 days, the IUH will have a total length of 33 days. The values constituting the vectors are generated by stochastic extraction from an evenly distributed distribution between 0 and 1. The IUH thus discretized will be called non-parametric IUH, as opposed to a parametric IUH model developed at a later stage, referred to in paragraph 5.6-The IUH. The definition of the initial observational model follows by convolution of the hydrograms as defined in 3.1-The inflow-outflow transformation. Our initial set consisting of the thousand vectors containing the state of the system and the relative observational model is therefore ready.

Then follows the application of the EnKF as seen in paragraph 4.2-The Ensemble Kalman Filter (EnKF). The assimilation of the measures proceeds starting from the first in question and advances along the series in chronological order, using the same value only once, perturbed by the error. This allows the filter to make changes to the whole set for each assimilated data, until it converges to an average value indicative of the solution.

Finally, the forward problem is applied using the average of the filtered set as the value of the IUH, in order to visually compare, at this stage, the goodness of the solution obtained.

In the following figures we see the initial set (Figure 5-6), the final set obtained after applying the filter (Figure 5-7) and the forward problem compared with the measured data (Figure 5-8).

It is possible to see how the solution obtained gives a good result, very similar to the measured values in terms of flow rate. However, the IUH values obtained after applying the filter range well beyond the theoretical boundaries of existence, reaching being negative. The result, therefore, represents an excellent mathematical solution, but it is unacceptable and unjustifiable from a physical point of view. This implies the need to develop mathematical constraints capable of keeping the solution consistent with the hypotheses made about the physical behaviour of the system.

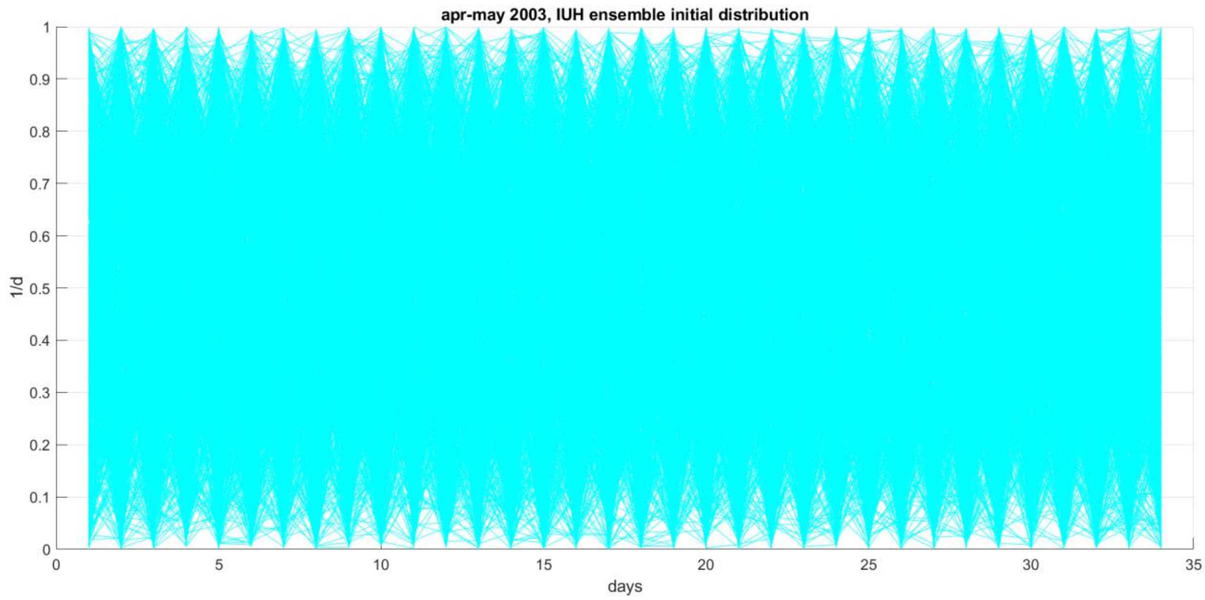


Figure 5-6: Initial set of state vectors (IUH).

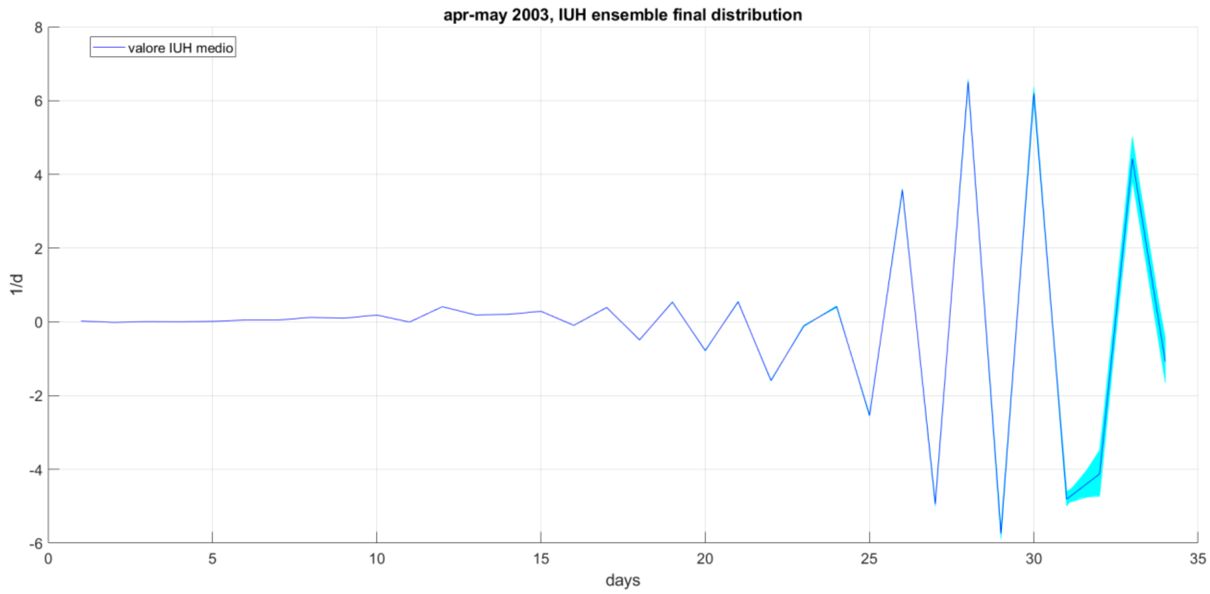


Figure 5-7: Final set of state vectors (IUH) subjected to the filter. Non-parametric IUH.

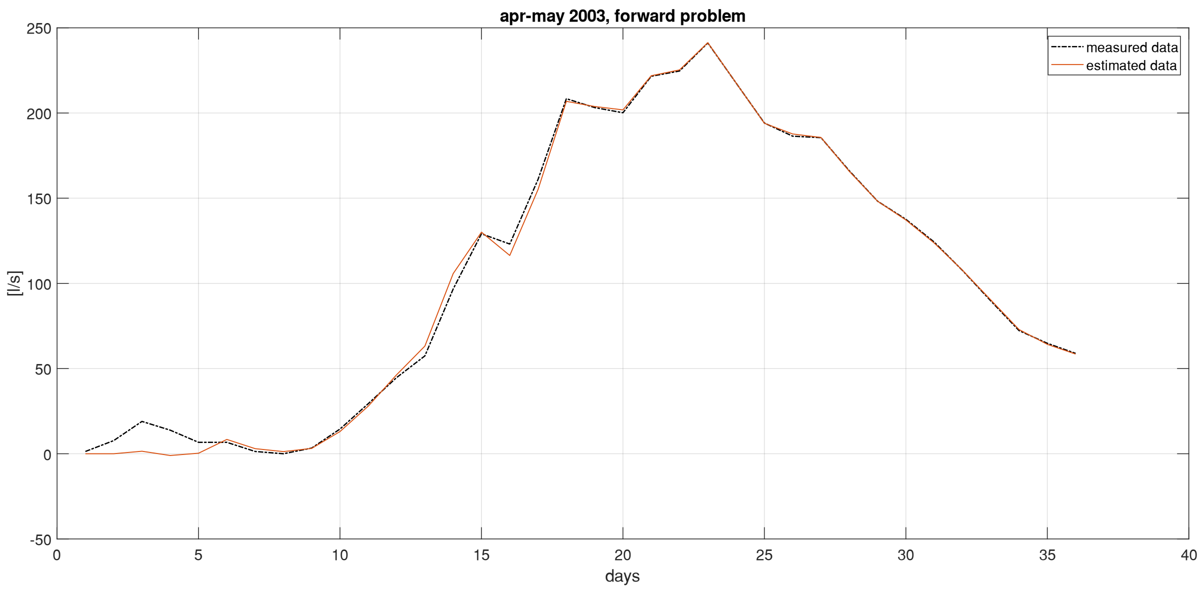


Figure 5-8: Forward problem and comparison with the measured data.

5.4 Adoption of logarithmic values

The problem constituted by the dispersion of the final IUH values beyond the theoretical range of existence can be limited by adopting logarithmic values. In short, a change of variable is performed where the filtered set consists of the natural logarithms of the values generated as described in the previous paragraph, that is to say $E = \ln(E')$ where E' is the generated ensemble and E the ensemble to be used in the filter. In this way the field of existence of the solution is limited to strictly positive values. The results are visible in the following figures (Figure 5-9 and Figure 5-10).

If on the one hand the flow rates estimated by the model in the forward problem are much more approximate than the measured values, on the other hand the average IUH is effectively limited in its field of existence. Nonetheless, a series of further problems arise, first of all the integral value, or rather the summation being discrete values, of the IUH: The average value is 4.5, far from the theoretical one equal to 1. Compliance with this value is necessary to ensure compliance with the mass conservation law (the incoming mass must be equal to the outgoing mass).

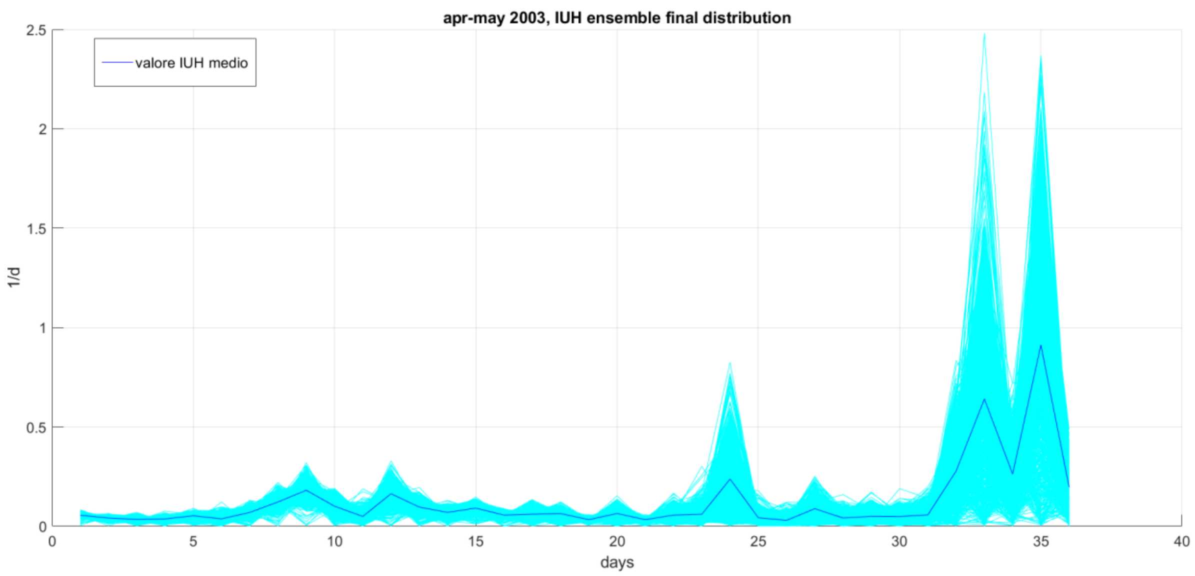


Figure 5-9: Final set of state vectors (IUH) subjected to the filter (after the adoption of logarithms). Non-parametric IUH.

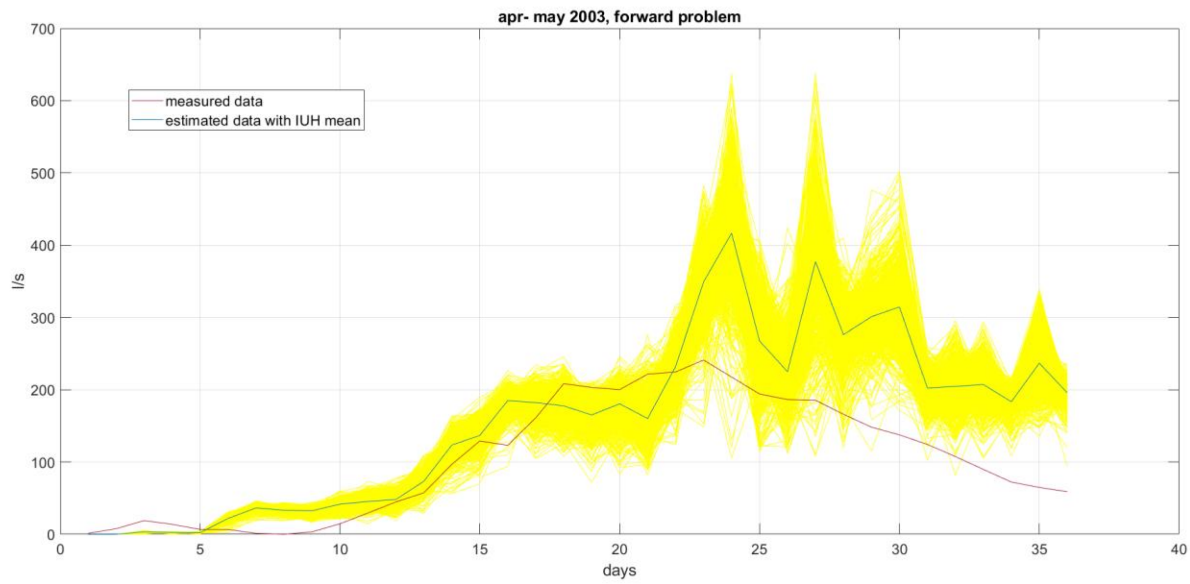


Figure 5-10: Forward problem and comparison with the measured data (after the adoption of logarithms). In yellow the set of solutions obtained and the average solution in blue.

5.5 Adoption of the ES-MDA

At this point, in the development of a solution to the problem under consideration, it was decided to adopt the algorithm of the Ensemble Smoother with multiple data acquisition, as described in the paragraph 4.3-The Ensemble Smoother with Multiple Data Assimilation (ES-MDA).

Once again, the set of system state variables, the IUH, is stochastically extracted from an evenly distributed distribution. The observational model follows the ensemble of system states, the flow rates from the system.

This ensemble is processed through the algorithm, reusing the same measure several times to search for the solution. This is made possible thanks to a decreasing monotone vector of error inflation factors capable of stabilizing the inversion of the filter matrices.

In a first phase, the first and simplest method for the generation of inflation factors described in the appropriate paragraph was adopted, while later we switch in favour of the other method, proposed by Rafiee and Reynolds in [12], again described in the chapter dedicated to the ES-MDA, also reducing the number of cycles required to reach the solution. From the confront in Figure 5-11 it is observed how the RMSE in the first version of the algorithm is more variable, while in the last it remains approximately monotonous decreasing.

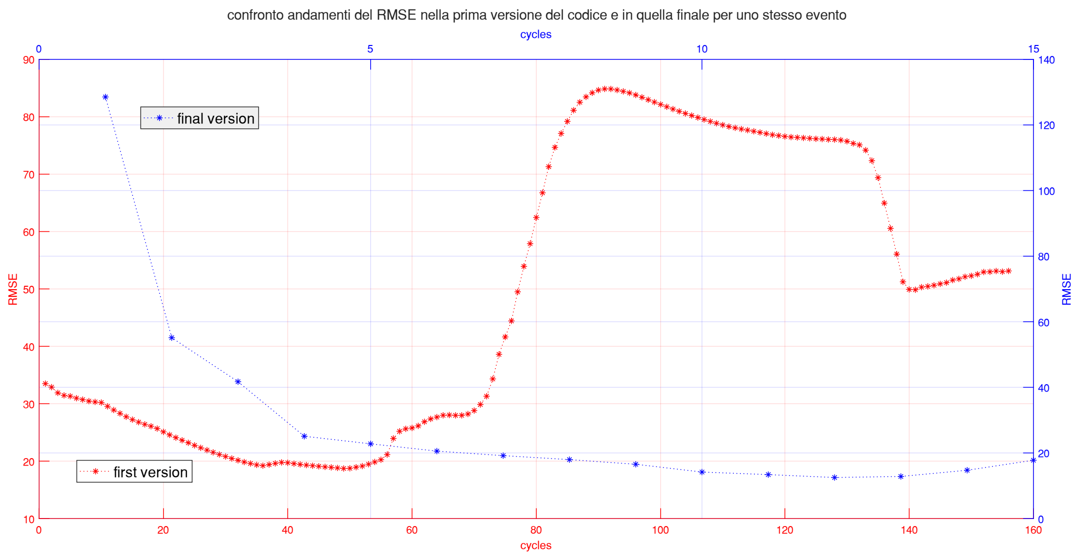


Figure 5-11: Evolution of the RMSE in the first version of the code (red) and in the final one (blue).

Below we can see the results obtained at this point of the process: in Figure 5-9 the final distribution of the system states, in Figure 5-13 the forward problem with all the different possible solutions from the ensemble compared with the measured data.

There is a clear improvement in the result in terms of event prediction and estimated IUH, the integral of the latter is also equal to 1.39, a value closest to the theoretical 1.

The model is however unstable, with difficulty in reaching convergence for some events and little reproducibility of the results, moreover it is applicable to periods not longer than a couple of months, under penalty of singularity of the matrices to be inverted during the process. In the following chapters the remaining difficulties and discrepancies that lead to the final model elaborated will be dealt with, presented below.

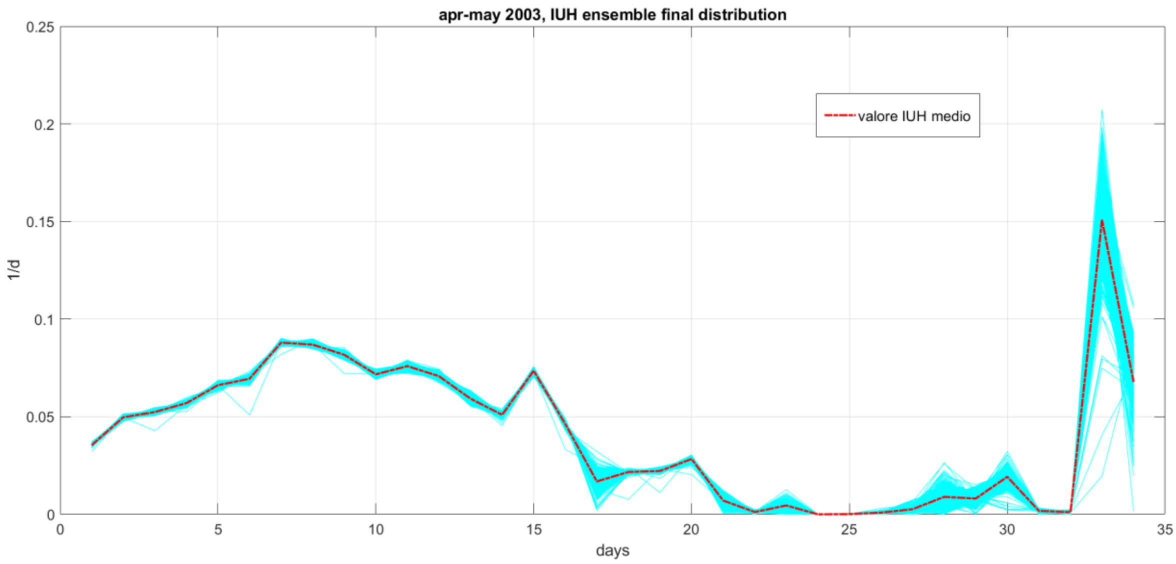


Figure 5-12 Final set of state vectors (IUH) subjected to ES-MDA filtering. Non-parametric IUH.

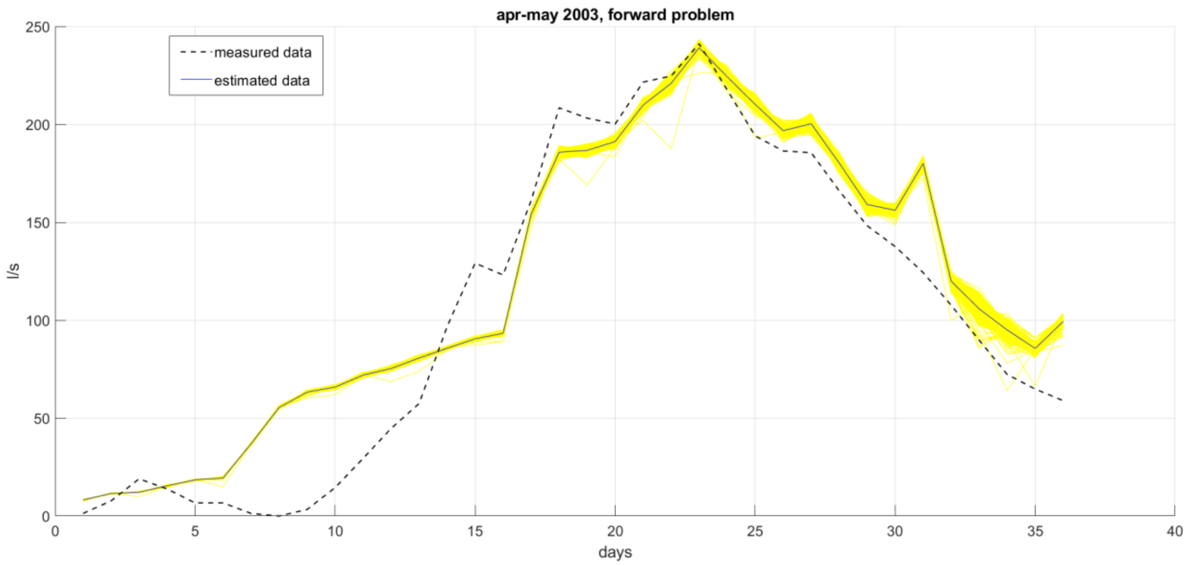


Figure 5-13 Forward problem and comparison with the measured data. In yellow the set of solutions obtained and the average solution in blue.

5.6 The IUH

During the development of the thesis, an influence by the initial set of system states on the final result was observed. By varying the stochastic distribution, different results are obtained.

After multiple attempts, an initial set was opted for consisting of vectors generated through a parametric form of the type:

$$IUH'_n = abs\left(\frac{e^{-\frac{(\ln t-\mu_n)^2}{2\sigma_n^2}}}{\sqrt{2\pi}\sigma_n t} + w_n\right) \qquad 5.4$$

Where the parameters σ_n e μ_n they are stochastically extracted from a continuous uniform distribution between 1 and the maximum time (in days) of the IUH. The term w_n represents instead a normal-distributed white noise $w_n \sim N(0, 0.0001)$. Lastly, the absolute value is necessary to avoid negative values whose logarithm would result in a complex number. The result is visible in Figure 5-14.

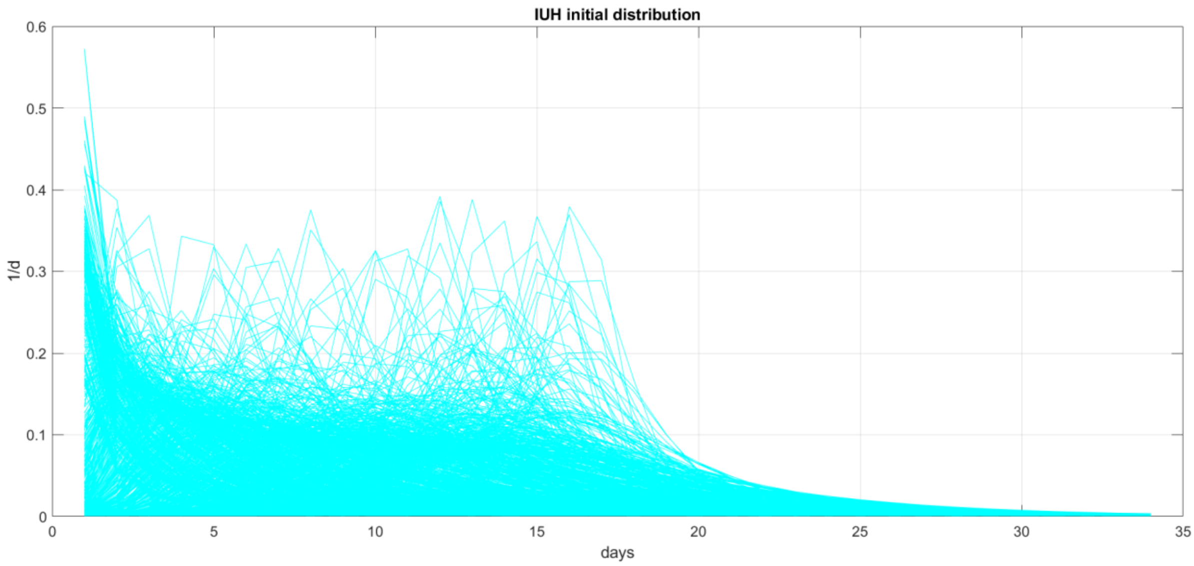


Figure 5-14: Initial set of system states generated according to the above method.

The following images show the results of the algorithm to which the variation just described is added. In Figure 5-15 the ensemble after the action of the filter and in Figure 5-16 the forward problem.

The solution obtained in terms of IUH is more stable with a comparable convergence of the single values, even in the final times. The final result in terms of estimated flow rates is sufficiently similar.

Despite everything, however, this measure alone is not sufficient to make the result stable and reproducible, especially for periods of time longer than about two months.

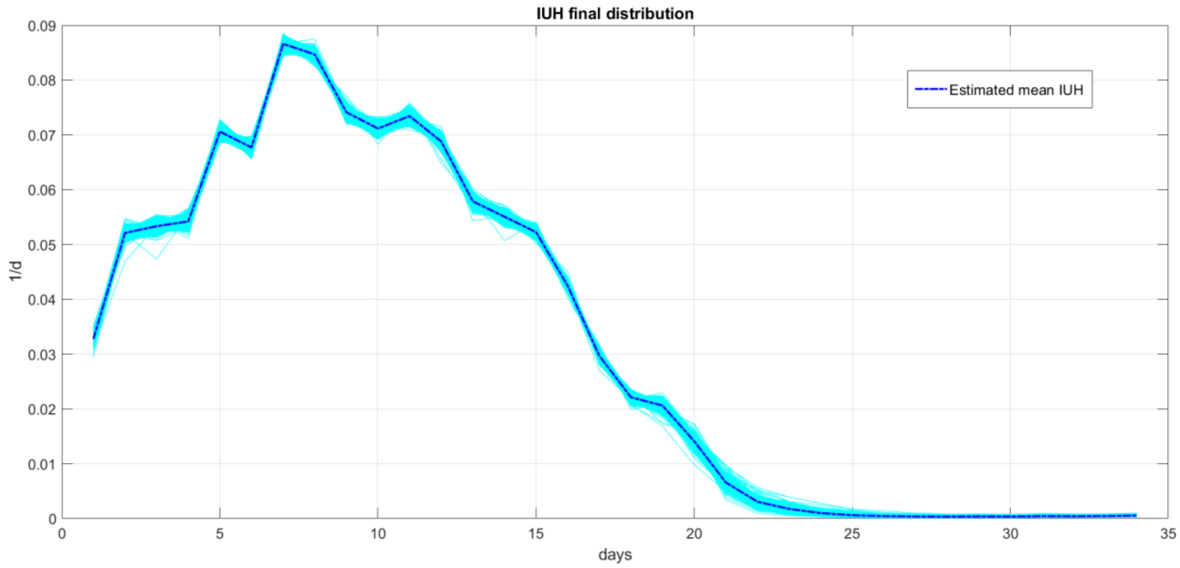


Figure 5-15: Set of IUHs after applying the algorithm. Non-parametric IUH.

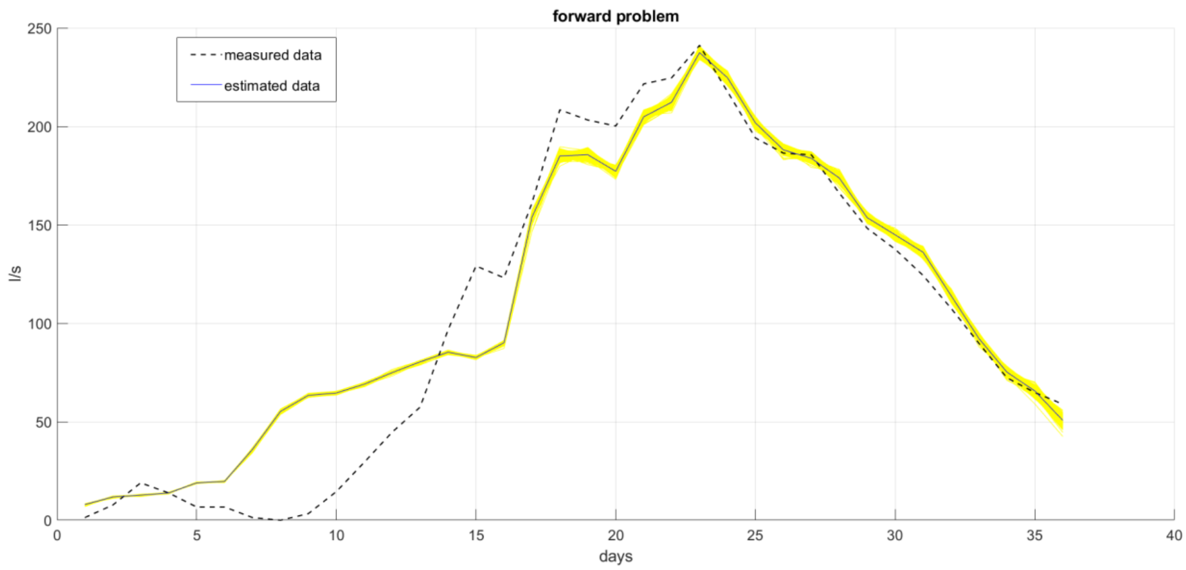


Figure 5-16: Corresponding forward problem.

A second possibility considered later is the adoption of a parametric form for estimating the IUH. In this way, the number of terms to be processed through the algorithm is reduced from one per day considered (generally one per unit of time) to a few parameters. After trying different solutions, comparing them with the non-parametric solution, it was decided to opt for a model of the type described by (5.5).

$$IUH'(t) = \frac{1}{\sqrt{2\pi v_1^2}} \cdot exp\left(-\frac{(t-m_1)^2}{2v_1^2}\right) + \frac{exp\left(-\frac{t}{m_2}\right)}{v_2} \qquad 5.5$$

Only the parameters m_1, v_1, m_2, v_2 are processed by the filter, reducing the computational costs of the process. The initial values are extracted from a uniform continuous distribution respectively in the interval [1,5] for the parameters m_1, m_2 and [5,10] for v_1, v_2.

This solution, together with what will come in the following three chapters, will be taken up again in the discussion about the final model.

Lastly, it was decided to impose the integral value of the IUH equal to 1 by applying:

$$IUH = \frac{IUH'}{\sum IUH'} \qquad 5.6$$

This ensures compliance with the theoretical value; the integral of the curve is always equal to one.

The possibility of developing IUH in parametric and non-parametric form is being carried out in parallel from here on. In this way it is possible to evaluate how strong the approximation due to the adoption of the (5.5).

5.7 The effect of precipitations prior to the period considered

We are now dealing with the problem posed by the flow generated by precipitation prior to the period considered. The outflow measured on the first day of the series under analysis will necessarily be generated by rain or snow prior to it. This contribution will then gradually decrease over the following times. We can call this outflow due to previous events with the name of “flow rate B”.

A first approach consists, as described above, in subtracting the minimum flow rate recorded in the series from the series itself, thus only the excess is considered to be generated in the period. In short, a constant flow rate B equal to the minimum recorded is identified. It goes without saying that it is appropriate to select events with a minimum of measurements on the first day, in order to make the hypothesis more consistent with physical reality.

However, it is reasonable to assume that the flow on day one is entirely generated by previous precipitations, but subsequent ones will gradually be in a smaller percentage, until the influence of these precipitations is overcome. In this sense, it was therefore decided to hypothesize an outflow B with a decreasing exponential trend able to mimic the influence of what happened in the time before the event under analysis on the measured flow. The equation is the (5.7).

$$Q_B(t) = \frac{Q_1}{exp(-q_2)} \cdot exp(-q_2 \cdot t) \tag{5.7}$$

It was decided to use a parametric form that is a function of time and regulated by the parameter q_2, whose value is estimated thanks to the ES-MDA. Q_1 is the flow rate measured on day one.

The value of q_2 it can also be bound to be greater than or equal to a value $q_{2,min}$, to avoid obtaining a result too close to zero and therefore a constant flow rate B over the entire series. With simple arithmetic steps it is calculated:

$$q_{2,min} = \frac{log\left(\frac{Q_1}{Q_n}\right)}{n-1} \tag{5.8}$$

once a minimum value at the time n of the series is identified, reasonably to be found between the firsts measurements. A good indication is to look for a minimum value in the first quarter of the series of observations.

The minimum constraint is then imposed with the change of variable:

$$q_2' = \frac{1}{q_2 - q_{2,min}} - 1 \tag{5.9}$$

Where the parameter q_2' is used as a state variable in the algorithm's ensemble.

The value $q_{2,min}$ represents the limit case in which in the first period, between the first measurement and the minimum, there is no precipitation affecting the system, then the flow rate B will interpolate the measured values. But if there is rain this new contribution will be able to partially justify the outflows and the flow rate B will therefore have to decrease more rapidly.

5.8 The snow melting

The instrument dedicated to the measurement of rainfall, mentioned in the paragraph 2.6-Borello weather station, is a heated rain gauge. This implies the need to distinguish snow from rain. Not only that, but also the snow that falls in the cold months melts with the warm season, hence the need to produce an estimate of the snow melting and its trend.

In general, snowfalls are considered to be those recorded in conjunction with an air temperature of less than one degree Celsius. On the other hand, snow melting occurs when the minimum daily temperatures exceed freezing.

In the work presented here we tried to arrive at a model capable of correctly identifying snowfall (in the winter period they are distinguished because they have no immediate influence on the flow rate) and then redistribute them over time respecting what is the total volume of snow fallen, in equivalent mm of water.

It is necessary to underline how the difference between the measurement altitude and the maximum height of the basin makes the snow measured probably often underestimated and the temperatures detected may be higher than those at altitude, where snow accumulates more easily.

By comparing the results on multiple events, a logic equation was reached that can identify snow events in a manner deemed satisfactory, depending on the measured daily temperature values, the average temperature, tmp_{avg}, the minimum temperature, tmp_{min}, and the maximum temperature, tmp_{max}. There is also a need to consider thermal inertia, considered thanks to a three-day moving average of the measured values, $tmp_{avg}^{mm}, tmp_{min}^{mm}, tmp_{max}^{mm}$.

$$snow = \left(tmp_{min}^{mm} \leq -1 \vee tmp_{avg} \leq 1 \vee tmp_{avg}^{mm} \leq 0\ t_{max} \leq 4 \vee tmp_{min} \leq -1.5\right) \wedge \left(tmp_{min} \leq 0 \vee tmp_{min}^{mm} \leq -1\right) \quad 5.10$$

After the snow we proceed to identify the period in which the snow melt is required. Two conditions have been set in this sense: the first, trivial, is the need for at least one day of snow to have occurred, the second is that the minimum temperature is higher than zero degrees centigrade. Therefore, those following the first snowfall are days of snow melt, during which positive minimum daily temperatures are recorded.

Finally, a model is needed that distributes the volume of snow fallen over the identified snow melt period. To do this, a vector of cumulative sum over time (the time previously identified as liable to melting) of the fallen snow is used, let's call it $SumSnow$, and starting with it we can build the snow melting contribution.

$$Sn'(t) = \frac{0.5 \cdot SumSnow}{\sqrt{2\pi b_1^2}} \cdot exp\left(-\frac{(t-a_1)^2}{2b_1^2}\right) + \\ + \frac{0.5 \cdot SumSnow}{\sqrt{2\pi b_2^2}} \cdot exp\left(-\frac{(t-a_2)^2}{2b_2^2}\right) \quad 5.11$$

All normalized to make the total volume coincide with the volume of fallen snow V_{snow}:

$$Sn(t) = \frac{Sn'(t)}{\sum Sn'(t)} \cdot V_{snow} \quad 5.12$$

The equation (5.11) that describes the progress of the snow melt is a form governed by parameters a_1, b_1, a_2, b_2, whose value is determined by the ES-MDA. They are then added to the set of state variables.

In Figure 5-17 the estimated average result is shown for the snow melt in spring 2003.

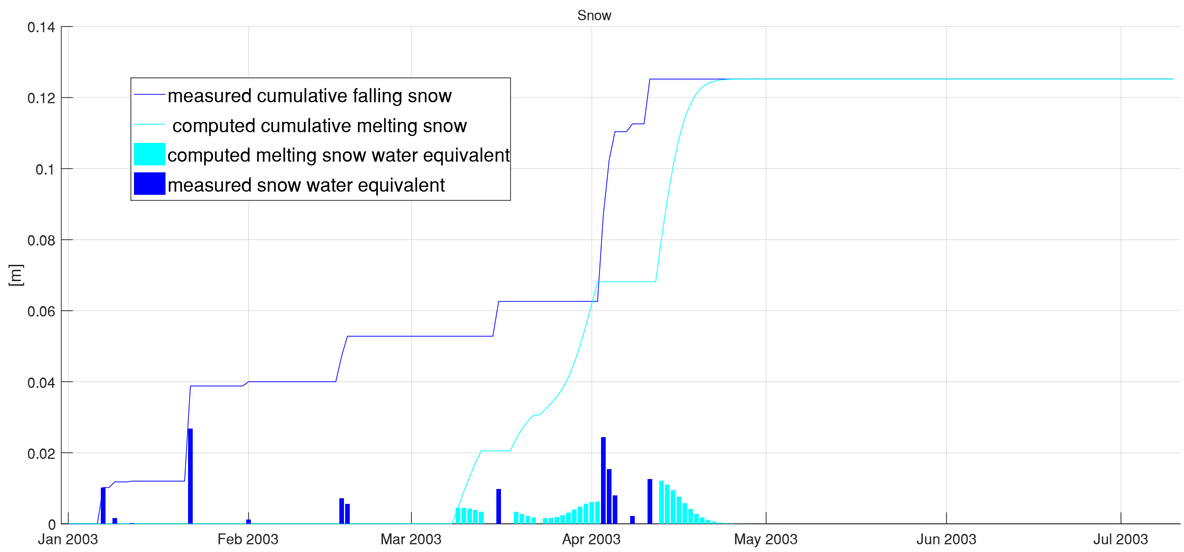


Figure 5-17: Example of snow fusion calculated by the algorithm.

5.9 Estimation of the infiltration coefficient over time

Up to this point the model provided for a constant infiltration coefficient in the selected time interval. It is now proposed to add a vector of infiltration coefficients to the parameters representing the state of the system, in order to evaluate their value day by day thanks to the algorithm. During the works an attempt was made to use a parametric form, but the poor results made it preferable to use a discrete vector of values free to vary.

As mentioned for the IUH, the way in which the values of the initial set are extracted influences the final result. The vectors representing the infiltration coefficients and making up the initial set are extracted thanks to the equation:

$$\chi(t) = \left| c_0 + \frac{c_1}{100} \cdot sin\left(\frac{2\pi}{T_1} \cdot t + \phi_1'\right) + \frac{c_2}{50} \cdot sin\left(\frac{2\pi}{T_2} \cdot t + \phi_2'\right) + \frac{c_3}{10} \cdot sin\left(\frac{2\pi}{T_3} \cdot t + \phi_3'\right) \right| \quad 5.13$$

In which the terms $c_0, c_1, c_2, c_3, \phi_1, \phi_2, \phi_3$ are extracted from an evenly distribute distribution, respectively the first three between 0 and 1, the last three between $-\pi$ and $+\pi$. T_1, T_2, T_3 are randomly drawn from a normal distribution centred at 1, 150, and 365, respectively.

The field of existence of the infiltration coefficient varies between 0 and 1. The logarithmic transformation applied to the set of state variables guarantees us a result included in $\mathbb{R}^+$, but it is still necessary to limit the values above. Hence the change of variable:

$$\chi'(t) = \frac{1 - \chi(t)}{\chi(t)} \quad 5.14$$

This leads to a result in the interval (0,1). The vector composed of the values of χ' calculated in this way, it is actually used in the ES-MDA, only to then apply the inverse of (5.14) to obtain the actual infiltration coefficient when applying the physical model.

In Figure 5-18 an example of infiltration coefficients calculated with the method described above is shown.

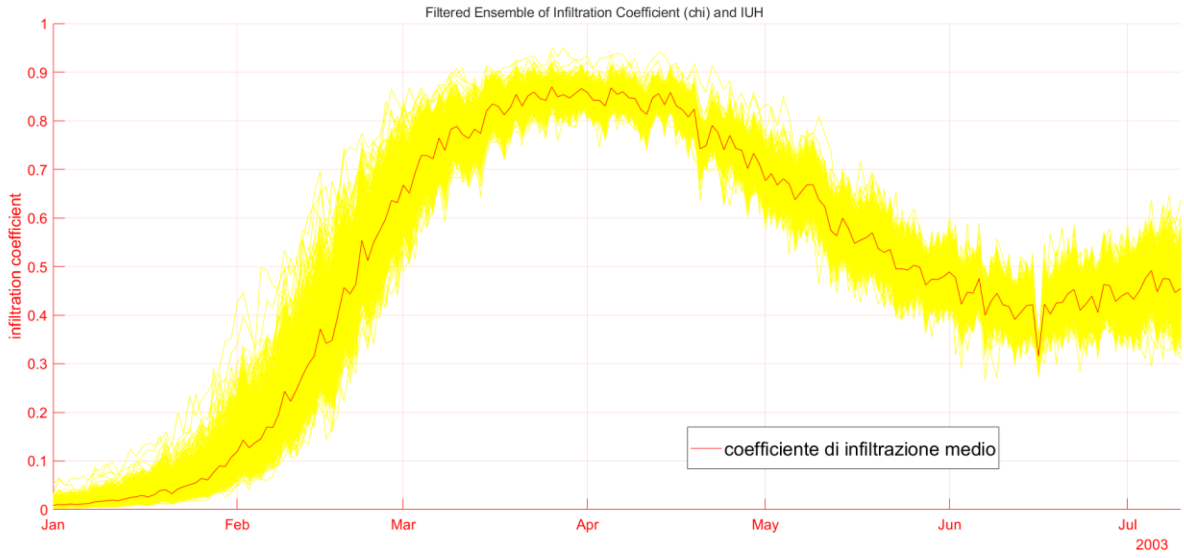


Figure 5-18: Example of infiltration coefficients calculated by the program. The set of values in yellow, its average in red.

5.10 The model

The model obtained by weaving together what has been said in the previous paragraphs is composed of four main elements:

- The instantaneous unit hydrogram
- The infiltration coefficients
- The snow melting as mm of water equivalent
- The outflow due to prior weather events, the flow rate B

The first point, the IUH, sees two possibilities: the generation of a vector then used in the algorithm, a parametric form whose regulatory parameters are used in the algorithm while the IUH is generated from time to time thanks to the chosen equation.

As regards the snow melting, the infiltration coefficients and the flow rate B, the methods described above are used.

The values of the state variables make up a vector x_n which goes to build the initial ensemble $E = [x_1, \dots, x_n, \dots, x_N]$, where N has been decided equal to 1000. The set is then submitted to the ES-MDA algorithm for a number of cycles to be decided, commonly 15. The final set thus obtained converges to the first two statistical moments of the real distribution of the parameters, unknown to us, or at least it should.

The two models are now described separately, different in the formulation of the IUH, while the remaining part is completely similar. In both cases the time t varies between 1 and the max time of the selected series T.

5.10.1 Non-parametric IUH

The state parameter vector will be of the type:

$$x_n = [IUH_1, \dots, IUH_n, \dots, IUH_T, a_1, b_1, a_2, b_2, \chi'_1, \dots, \chi'_n, \dots, \chi'_T, q'_2] \qquad 5.15$$

where T is the maximum time of the input data series.

The physical model provides for the following calculations: the snow melting vector, Sn, following the (5.12) and then (5.11), the flow rate B from (5.9) and (5.7), the tributary flow according to:

$$Q_{in} = (i + Sn) \cdot \chi \cdot \frac{A}{86{,}4} \qquad 5.16$$

where A is the area of the basin and χ the infiltration coefficient following the (5.14). The IUH vector is normalized with the (5.6), then hydrograms convolution is applied (3.2) and the flow rate B is added to obtain the outcoming flow.

5.10.2 Parametric IUH

The state parameter vector will be of the type:

$$x_n = [m_1, v_1, m_2, v_2, a_1, b_1, a_2, b_2, \chi'_1, \dots, \chi'_n, \dots, \chi'_T, q'_2] \qquad 5.17$$

where T is the maximum time of the input data series.

The physical model provides for the following calculations: the IUH vector from (5.5) and them (5.6), snow melting vector, Sn, following the (5.12) and (5.11), the flow rate B thanks to (5.9) and (5.7), the income from:

$$Q_{in} = (i + Sn) \cdot \chi \cdot \frac{A}{86{,}4} \qquad 5.18$$

where A is the basin's area and χ is the infiltration coefficient from (5.14). The convolution of the hydrograms is then applied (3.2) and the flow rate B is added to obtain the outgoing flow.

Below are graphs relating to the analysis of the flood event that occurred in the spring of 2003. From the data in Figure 5-19 you can see how the duration of the selected event is greater, to ensure a good estimate of the snow fallen in winter, moreover, a high initial flow rate is no longer a problem thanks to the estimate of the flow rate B.

The following figures will show part of the initial set (infiltration coefficient and IUH), the result obtained thanks to the algorithm for the same parameters and the forward problem with the various components of the model, compared with the recorded data.

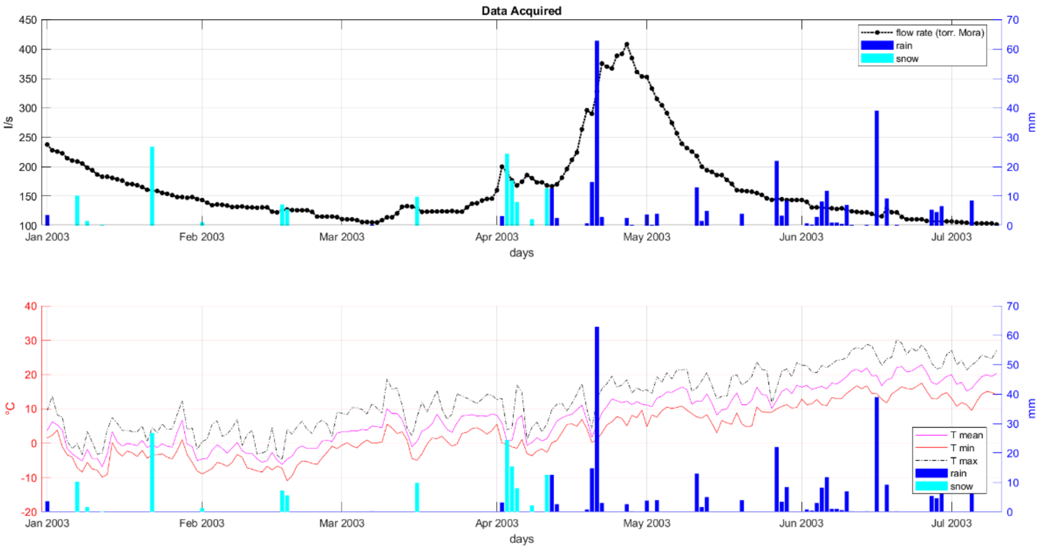


Figure 5-19: Input data, spring 2003.

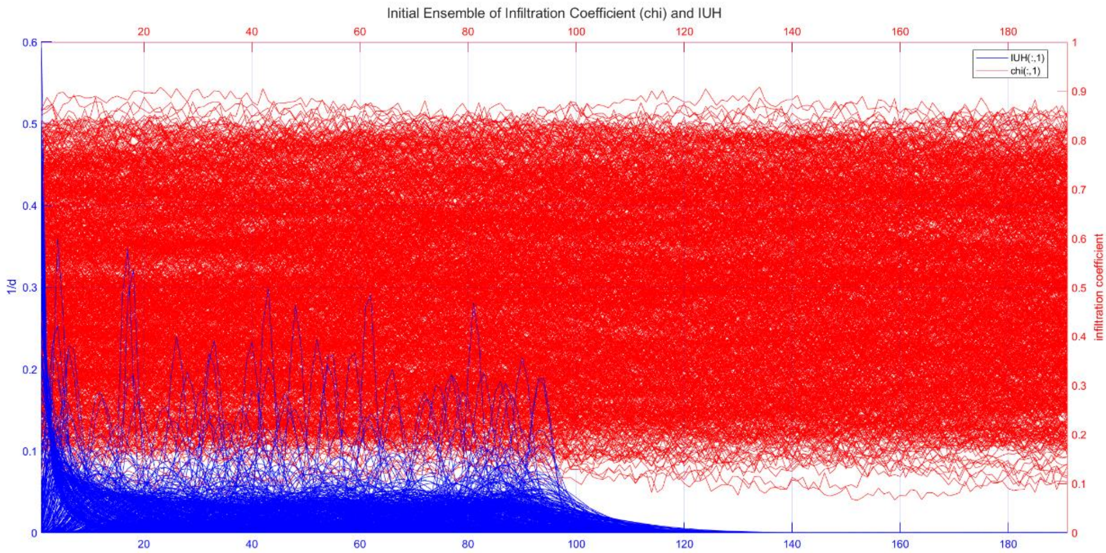


Figure 5-20: Initial set, non-parametric IUH.

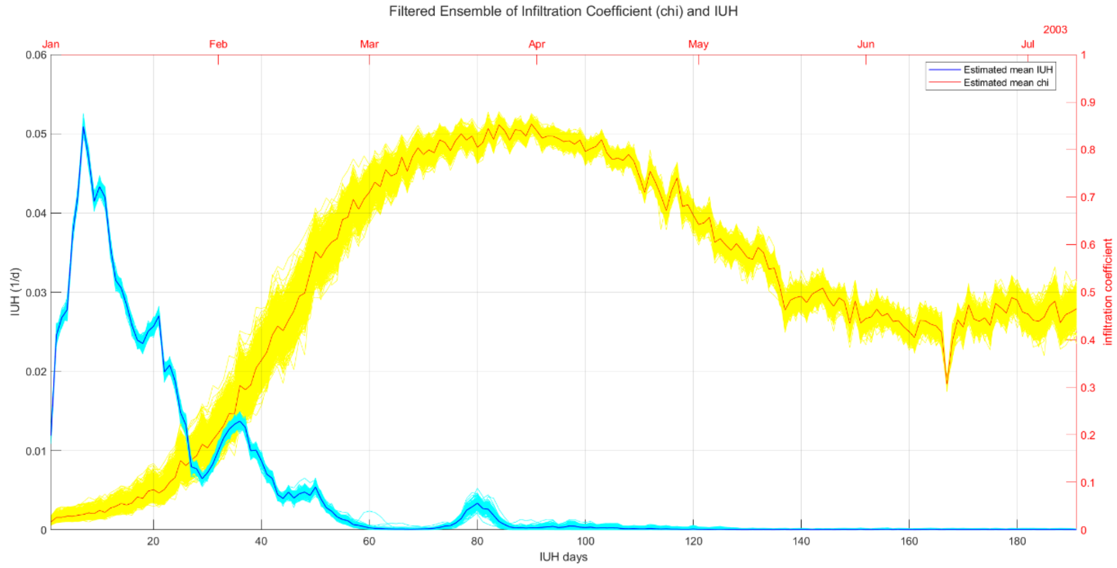


Figure 5-21: Final result in terms of IUH and infiltration coefficient, non-parametric IUH.

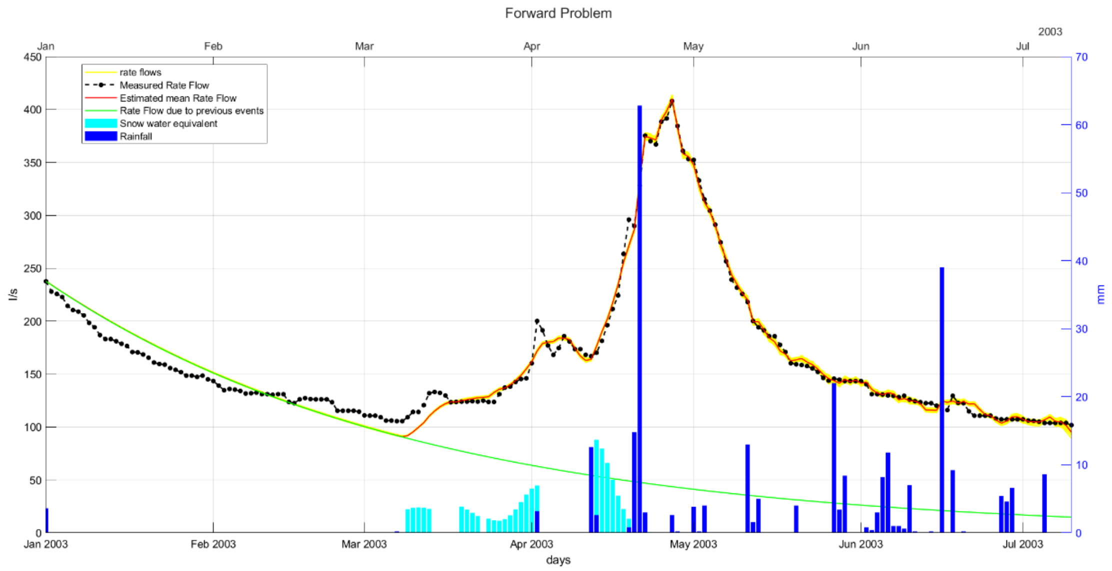


Figure 5-22: Forward problem, non-parametric IUH. In green the outflow B, in blue the contribution of the snow melt.

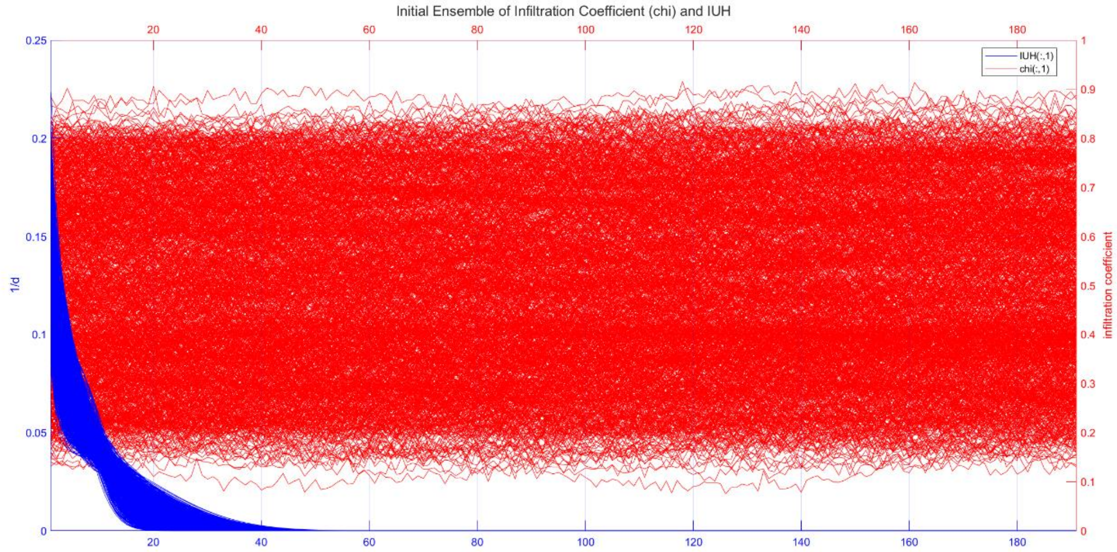


Figure 5-23: Initial set, parametric IUH (the curves representing the different IUH are generated thanks to the parameters part of the system state vector).

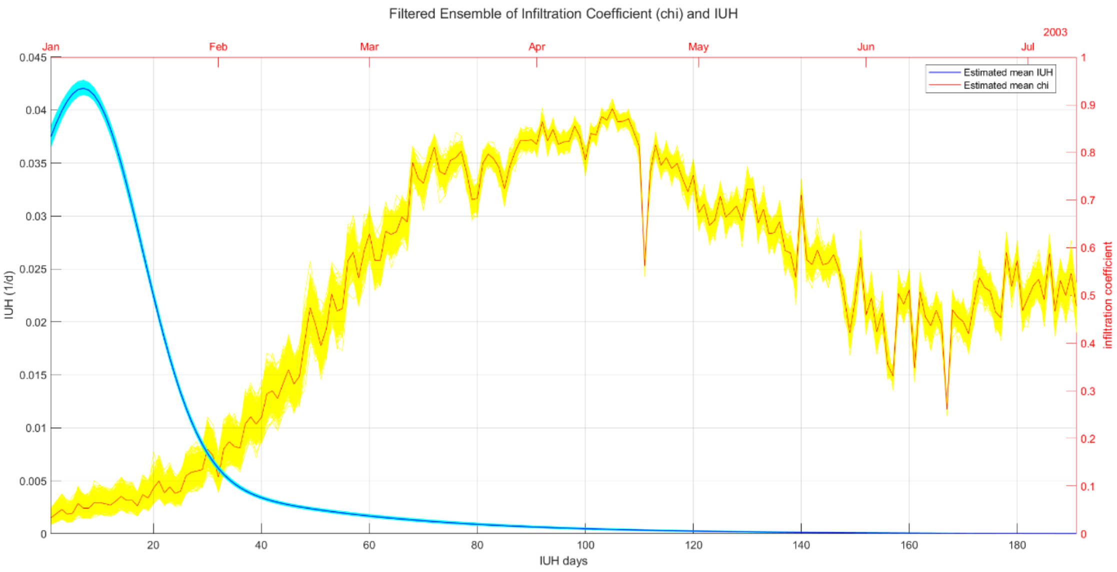


Figure 5-24: Results obtained thanks to the algorithm, parametric IUH. Again, the IUH curves are generated thanks to the corresponding parameters.

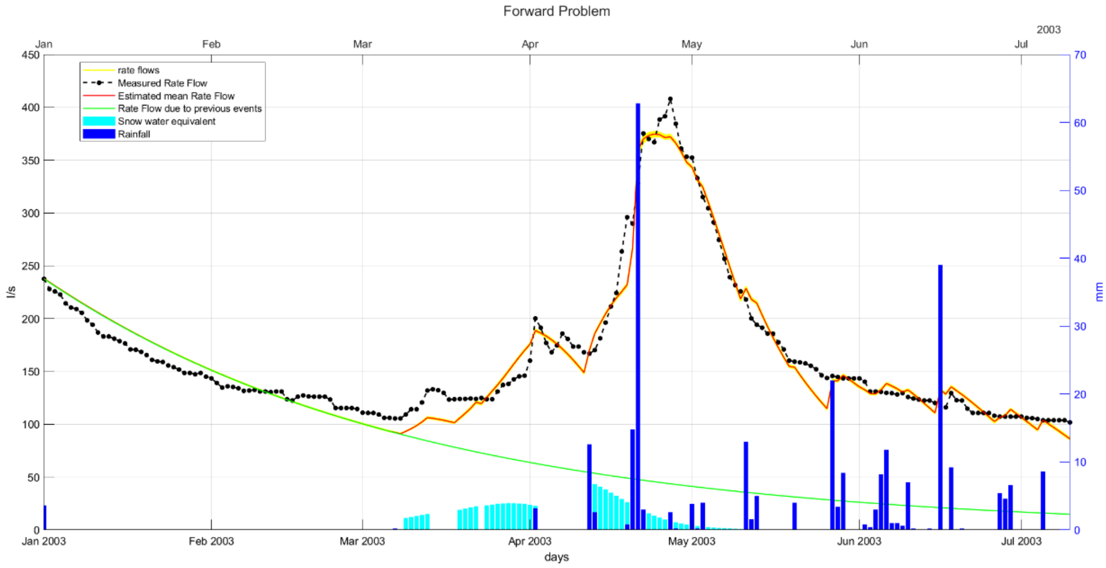


Figure 5-25: Forward problem, parametric IUH. In green the outflow B, in blue the contribution of the snow melt.

5.11 The inbreeding problem

Now, to continue with the development of the model, we proceed to an evaluation of the results obtained so far in terms of efficiency of Nash-Sutcliffe, NSE, and of inbreeding.

Let's keep with example the realization and the event analysed in the previous paragraph (spring 2003) and let's see how the results behave in terms of NSE: with non-parametric IUH a value of 0.98 is recorded, while with parametric IUH of 0.96. In these terms, the model is therefore satisfactory.

With regard to inbreeding, the indicator calculated according to (4.33) provides the following results: 12.12 for non-parametric IUH and 39.20 for parametric IUH. From this point of view, the results are everything but acceptable.

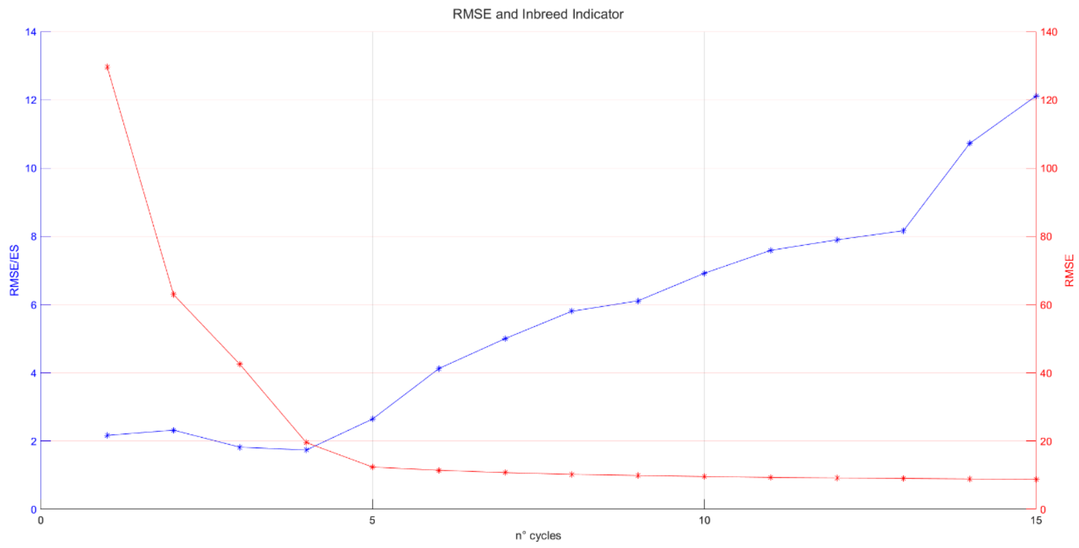


Figure 5-26: RMSE and RMSE/ES trend for the spring 2003 event, non-parametric IUH.

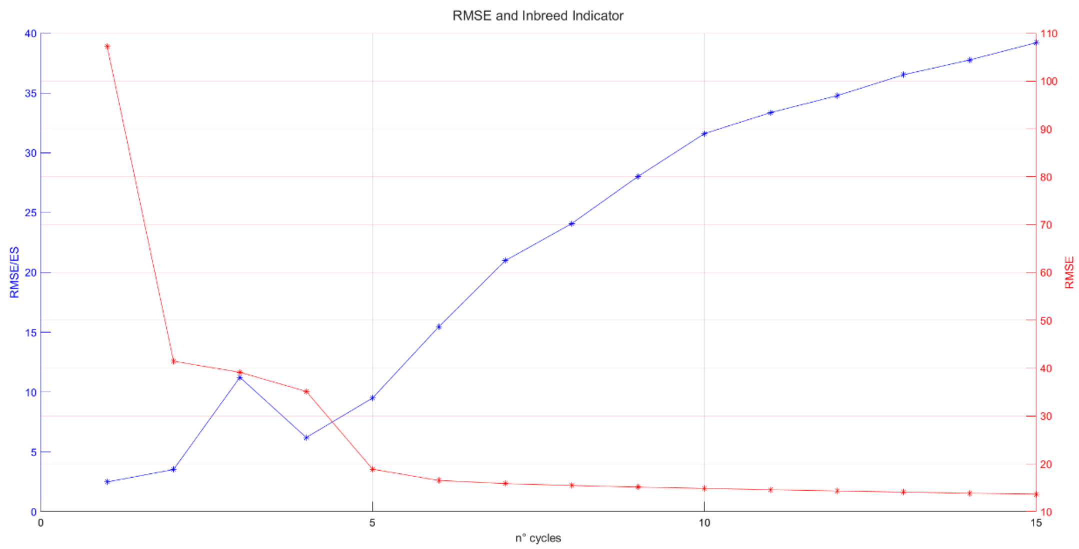


Figure 5-27: RMSE and RMSE / ES trend for the spring 2003 event, parametric IUH.

In Figure 5-26 and Figure 5-27 are shown the complete trends of the RMSE and the inbreeding indicator.

Yet in Figure 5-22 and Figure 5-25 the problem is to some extent visible; the set of solutions is poorly distributed and always very close to the average.

At this point the problem arises of containing the collapse of the filter without compromising the hydrological validity of the final result. During the works it was seen that none of the solutions proposed in the paragraph 4.4-Of Inbreeding, Covariance Inflation e Damping Factor is able, on its own, to solve the problem. On the other hand, when both *damping factor* and *covariance inflation* are used, better results are obtained. Added to this is the possibility of increasing the variance of the error (5.2) to allow greater variability of the result. In this way the error on the observations is able to also incorporate the error of the model. It was decided to use a multiplier factor, let's call it *error multiplier*, to achieve this result.

It was not possible to define unique values of damping factor (DF), error multiplier (EM) and covariance inflation coefficient *(λ)* able to work for any selected event, therefore each time it is advisable to proceed with the selection of these coefficients by trial and error.

Below are the graphs of the results obtained for the same event analysed above, but they are applied:

- Non-parametric IUH: $DF = 0.8$, $\lambda = 1.1$, $EM = 2$;
- Parametric IUH: $DF = 0.4$, $\lambda = 1.15$, $EM = 3$.

In the first case (non-parametric IUH) we have an NSE value of 98% and an inbreeding index of 3.95; in the second case (parametric IUH) we have an NSE value of 96% and an inbreeding index of 3.05.

We can also see in Figure 5-30 and Figure 5-33 how the trend of the RMSE remains predominantly monotonous and decreasing, while the RMSE / ES ratio remains low and does not skyrocket as the cycles advance, as was the case previously.

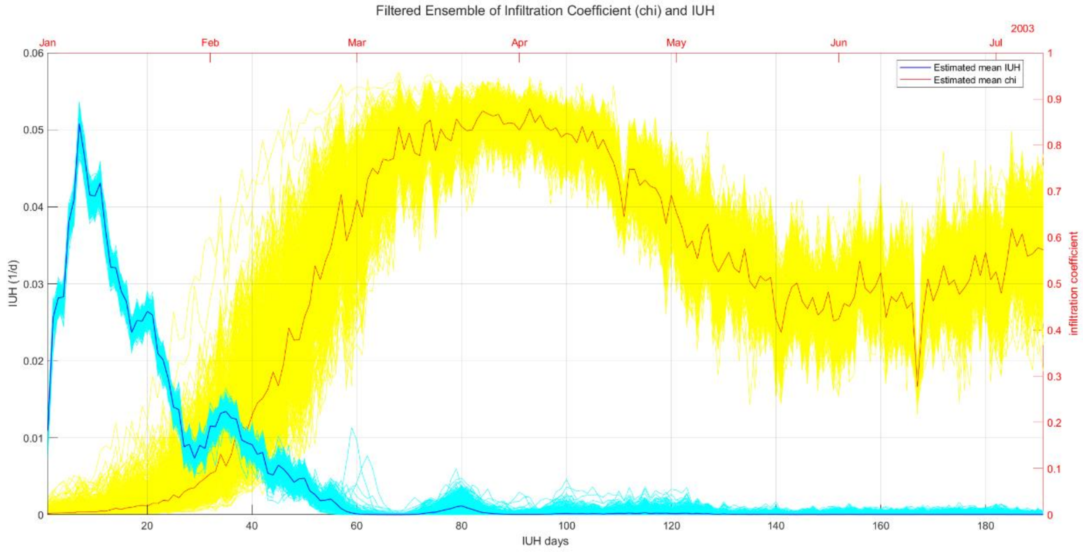


Figure 5-28: IUH and infiltration coefficient obtained with the definitive algorithm. Non-parametric IUH.

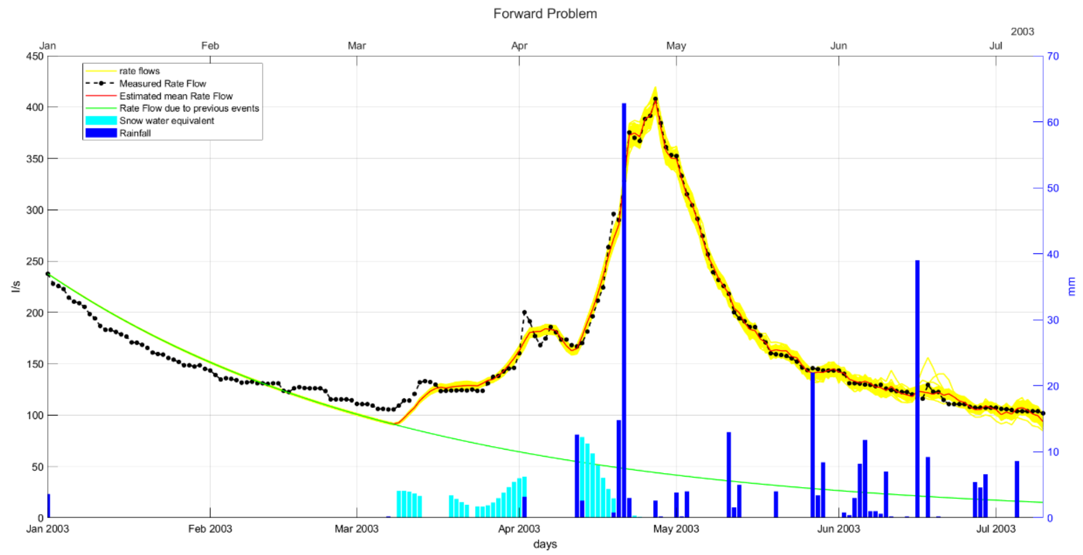


Figure 5-29: Forward problem, non-parametric IUH.

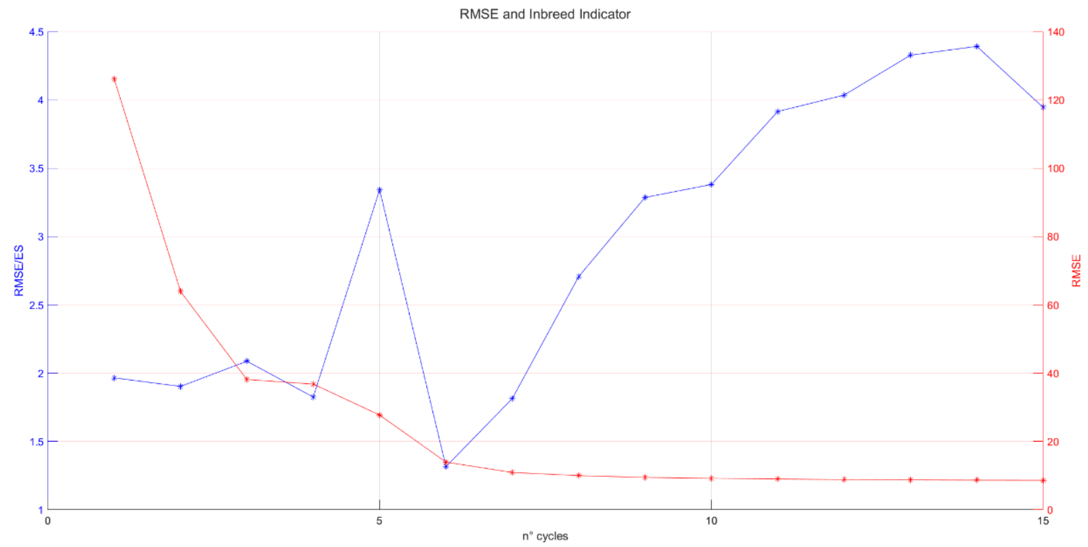


Figure 5-30: RMSE and RMSE / ES trend. Non-parametric IUH.

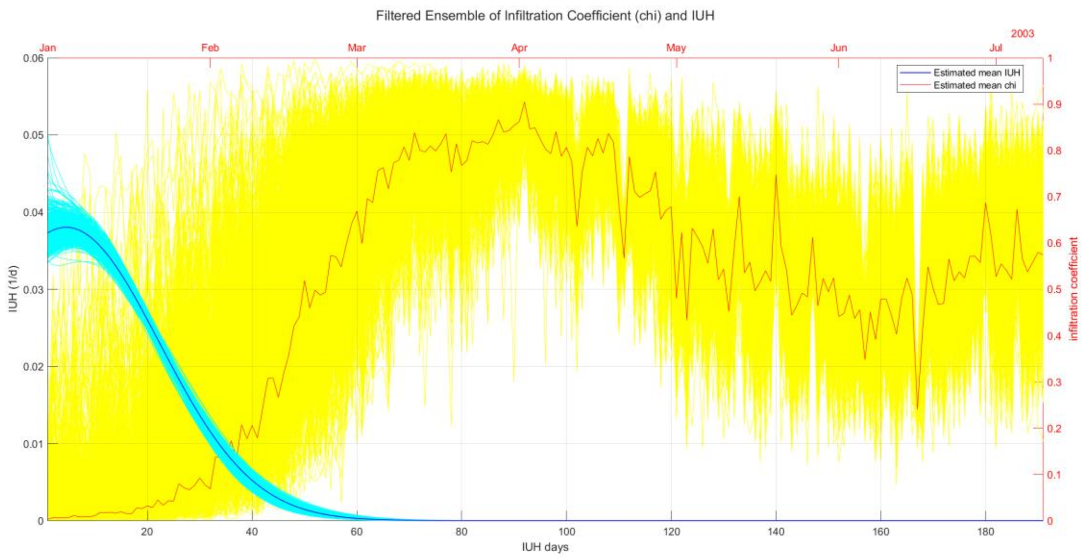


Figure 5-31: Final result in terms of IUH and infiltration coefficient, parametric IUH.

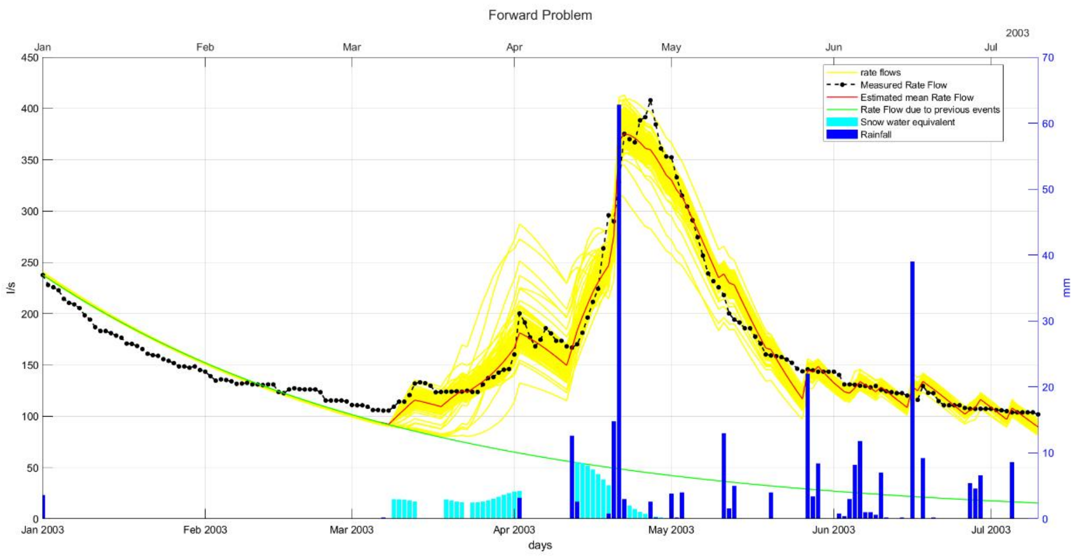


Figure 5-32: Forward problem. Parametric IUH.

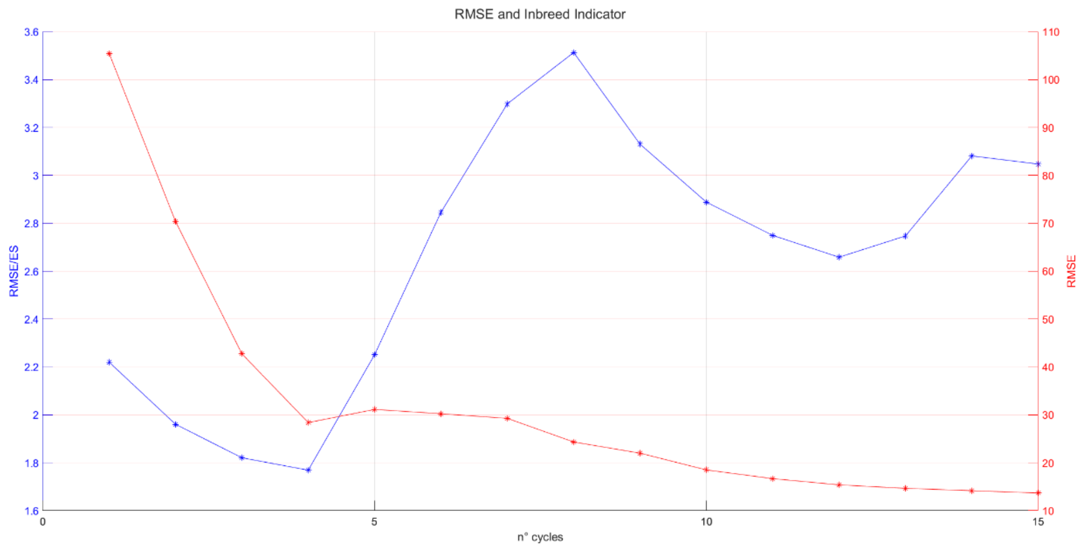


Figure 5-33: RMSE and RMSE / ES trend. Parametric IUH.

5.12 The final model

In Appendix A-Analysed Events the events analysed by the model are reported as graphs and useful data. The following conclusions have been drawn from the observation of these results.

Ultimately, the model thus developed provides for the estimation of a series of data:

- A flow rate due to previous events (the flow rate B), mostly useful for correctly estimating the other parameters that most directly affect the selected event
- The IUH, the response of the system to a unitary influence, in a parametric or non-parametric form
- The infiltration coefficient, in a discrete form as a vector in which each element represents the infiltration coefficient of the corresponding day
- The snow melting, that is the input contribution that undergoes a translation over time; the snow that fell in winter does not turn into an income until spring, hence the need for an estimate of the melting over time

The water system under analysis is complex, with fractured parts, conducted under full load, siphons, and transfers from detrital aquifers. This implies a variability in terms of response, IUH, and infiltration coefficient.

The IUH in parametric form displays a rather acceptable result, but on the one hand the difficulty in maintaining variability in the ensemble and avoiding its collapse and on the other hand the differences, in some cases marked, with the non-parametric result suggest that this approximation is very strong and not always acceptable. However, the concomitant estimate of the infiltration coefficient is able to absorb part of the irregularity of the real system.

On the other hand, the infiltration coefficient presents a variability that differs considerably from the first hypothesis of a constant value over the period. In general, recurring seasonal variations are observed, especially in summer involving values that are lower than the average. This somewhat stochastic trend can be explained thanks to several concomitant factors: rainfall, assumed to be constant on the supply basin, can actually have a spatially discontinuous character due to their sometimes-stormy nature; the aquifer system reacts differently according to the stored water volume; temperatures can vary significantly in short periods, as well as soil moisture and vegetation behaviour, consequently evapotranspiration can change dramatically and quickly.

6 Results Interpretation

Once the model described in the previous chapter was developed, the work continued with the analysis of a series of flood events identified in the 2001-2018 period, the data of which are reported in Appendix A-Analysed Events. These data were then processed, and the results are presented below.

6.1 Of IUH

The average values of the instantaneous unit hydrogram calculated for the different events were recorded and compared with each other. It was decided to group them according to seasonal criteria, in two groups: autumn events (including summer ones) and spring events.

The results obtained with the non-parametric model are shown below. From the images it is clear that this solution involves a distinctly irregular response with high variability between different events. The average response, in red in the figures, was calculated as the daily average of the reported values, in order to identify a general trace of the seasonal behaviour of the aquifer. The confrontation between Figure 6-1 and Figure 6-2 allows us to observe the differences in the two seasonal responses: where spring distributes the influx over a longer period and gradually, going to zero in about 100 days, autumn is characterized by impulsive behaviour, with an initial peak and a smaller exhaustion queue in terms of volumes and times.

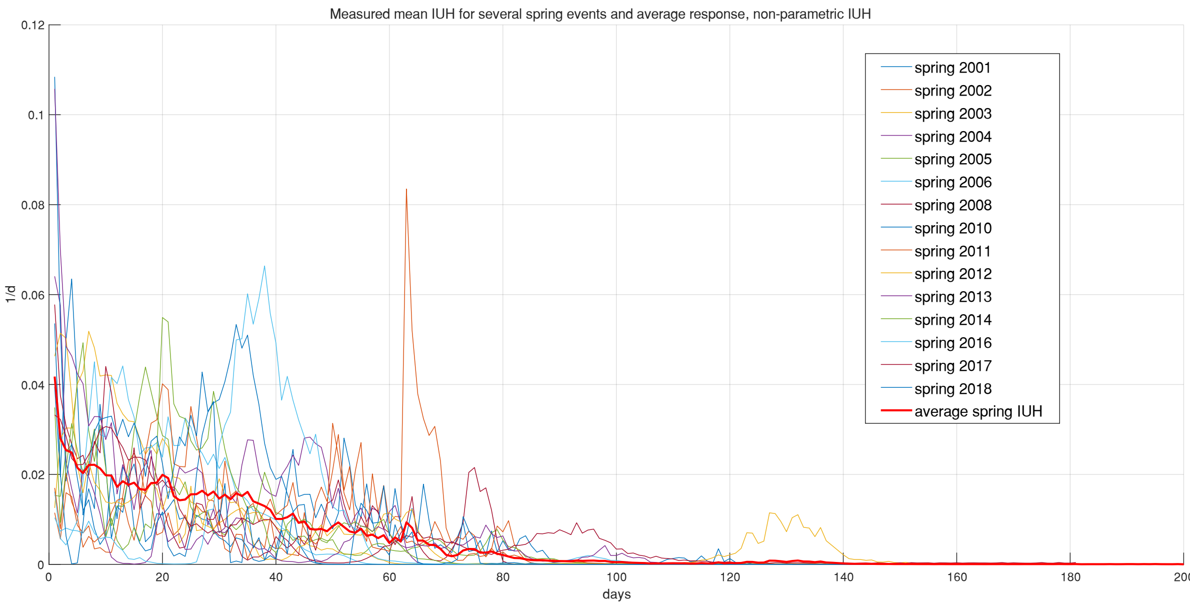


Figure 6-1: Non-parametric IUH recorded for spring events and daily average of values.

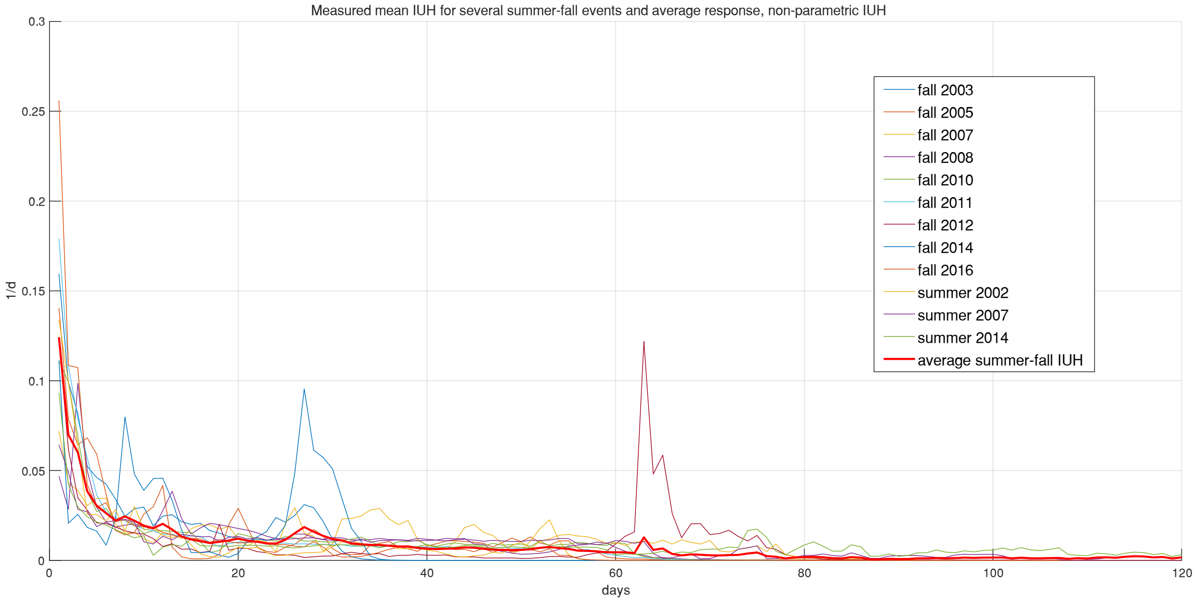


Figure 6-2: Non-parametric IUH recorded for autumn events and daily average of values

.

At this point the responses estimated with parametric IUH are going to be analysed. The adoption of a parametric shape, it goes without saying, eliminates the irregular trend of the curve. The seasonal mean response this time is calculated using the average of the recorded parameters. In addition, there is a second possibility for estimating an average response: after calculating the daily average of the responses obtained during the various events, we proceed to estimate the parameters capable of fitting (with Trust-Region and Bisquare weights algorithms) better this average response. The values obtained are presented below, the interpolation operations return in both cases (spring and autumn) a value of R^2 higher than 0.99.

The average values of the parameters recorded in the spring season are: $m_1 = 7.54, v_1 = 25.64, m_2 = 9.96, v_2 = 58.54$, while the interpolation parameters for the average daily spring response: $m_1 = 10.57, v_1 = 17.91, m_2 = 4.054, v_2 = 15.84$. For the summer-autumn season the average values of the parameters are: $m_1 = 1.13, v_1 = 25.41, m_2 = 1.94, v_2 = 5.91$, while the fitting parameters: $m_1 = 0.072, v_1 = 2.313, m_2 = 21.75, v_2 = 32.82$. The average response estimated by interpolation does not fully respect the integral equal to 1, with a spring value of 0.94 and an autumn value of 1.07.

Again, the observations made for the previous model are valid: the spring responses tend towards volumes distributed over long times, while the autumn ones behave in a more impulsive manner.

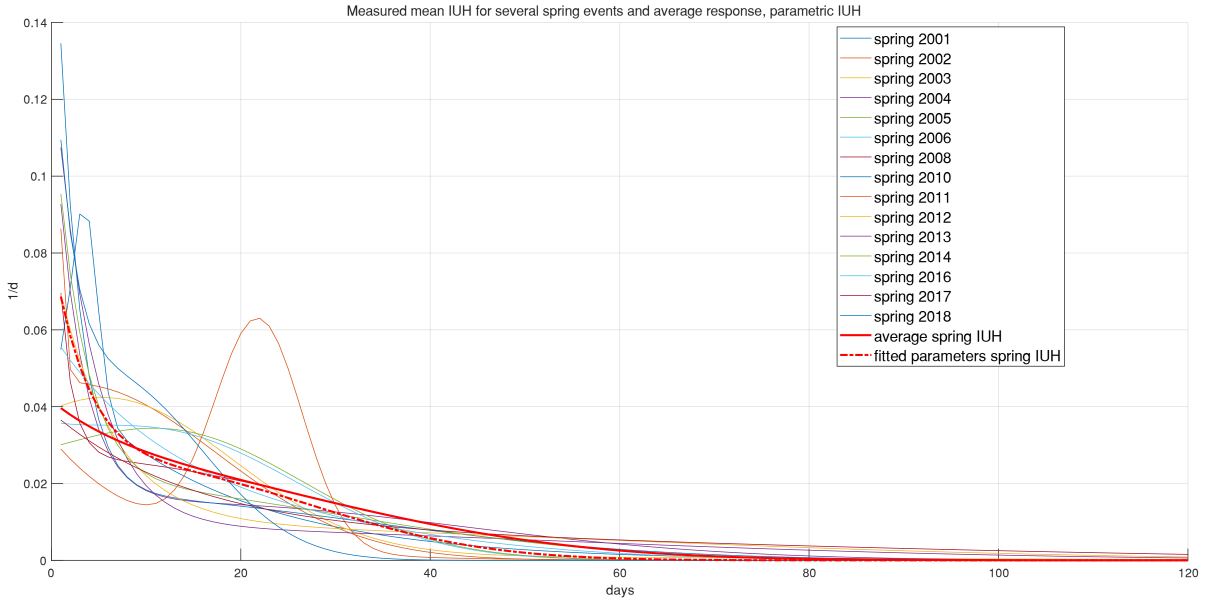


Figure 6-3: Parametric IUH recorded for spring events and response with average parameters.

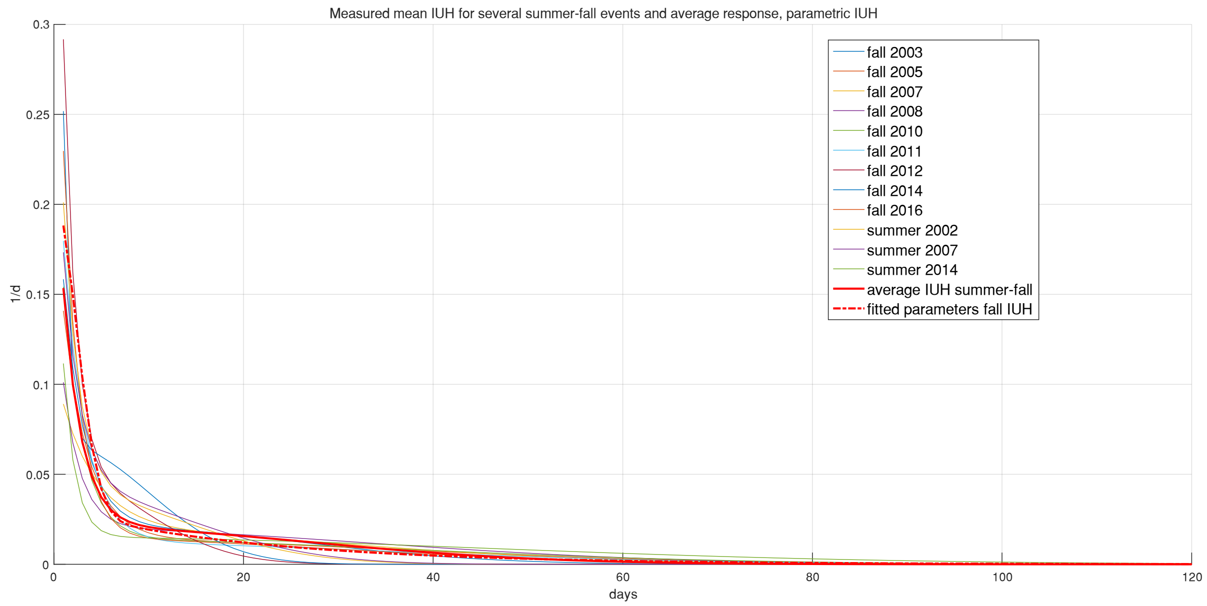


Figure 6-4: Parametric IUH calculated for fall events and response with mean parameters

.

The response of the two models is similar in the trend. A decrease in the time in which the IUH goes to zero can be observed in the parametric model, present in both the spring and autumn seasons. The adoption of IUH in parametric form reduces the variability of seasonal responses, making them more uniform among themselves. Is interesting to observe how the area subtended by those average calculated IUH is very close to 1 in all the four cases.

Part of the difference in the behaviour recorded in the two periods of the year can be explained by the complexity of the aquifer system and the different feeding. Spring involves snow melting and widespread rains, these phenomena affect the groundwater in the fractured with a delay in response times; summer and autumn are the seasons of thunderstorms with short and intense precipitations, able to affect the karst pipes with full load through the sinkholes and to activate siphons, in this way a large volume is immediately conveyed to the measurement section, while the runoff coefficient grows in the face of high volumes not able to widely affect the fracture in the short time of a stormy downpour.

Now, we can make a comparison between those average responses and the ones of a specific event. In other words, we put into a confrontation the response obtained through the ES-MDA and a similar one, in which the IUH has been switched with the average one of the same season and type (parametric or non-parametric). To do so a couple of flood events has been selected, whose IUH doffers enough to the average one to be a good yardstick. Particularly we are talking about the events registered in the 2006 spring and in 2014 autumn, chosen by chance before the test. In the following table the NSE of those average IUH responses are reported, then we can see the charted flow rates. A more extensive analysis suggest placing the selected events between the worsts in term of efficiency and error of the mean IUH response, with only a couple that performed in an even less favourable way. This analysis displays how the average IUH has its predictive power, and it is consistently higher the average flow rate. Moreover, the parametric type of IUH seems to lead to a better average response with a better behaviour when used instead of the specific one.

Table 6.1: NSE of the events selected to compare average IUH with specific ones.

spring 2006	non-parametric average IUH	0.60	parametric average IUH	0.78
fall 2014	non-parametric average IUH	0.65	parametric average IUH	0.74

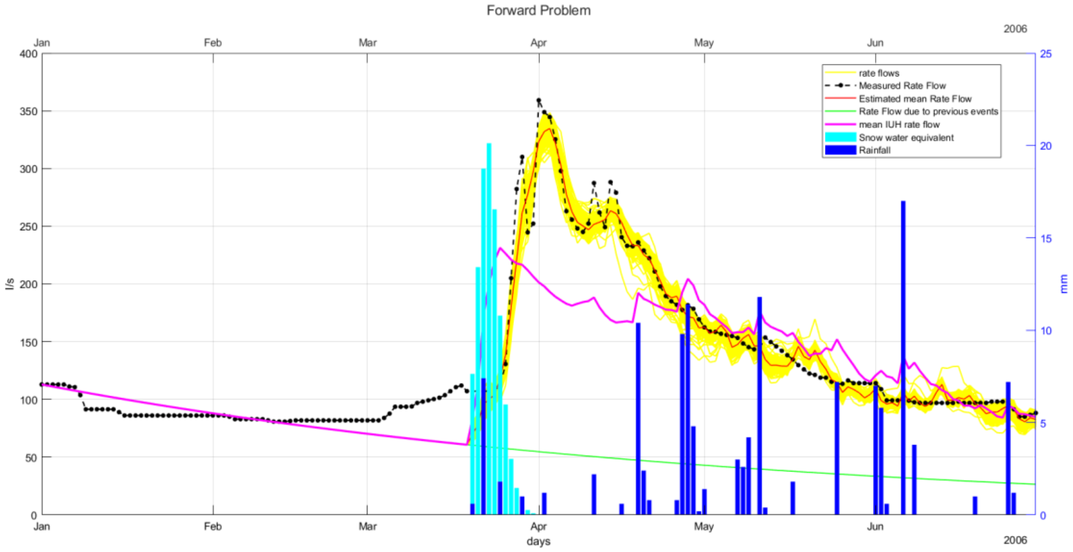


Figure 6-5: Confront between the event specifical IUH response and the average IUH response, non-parametric IUH, spring 2006.

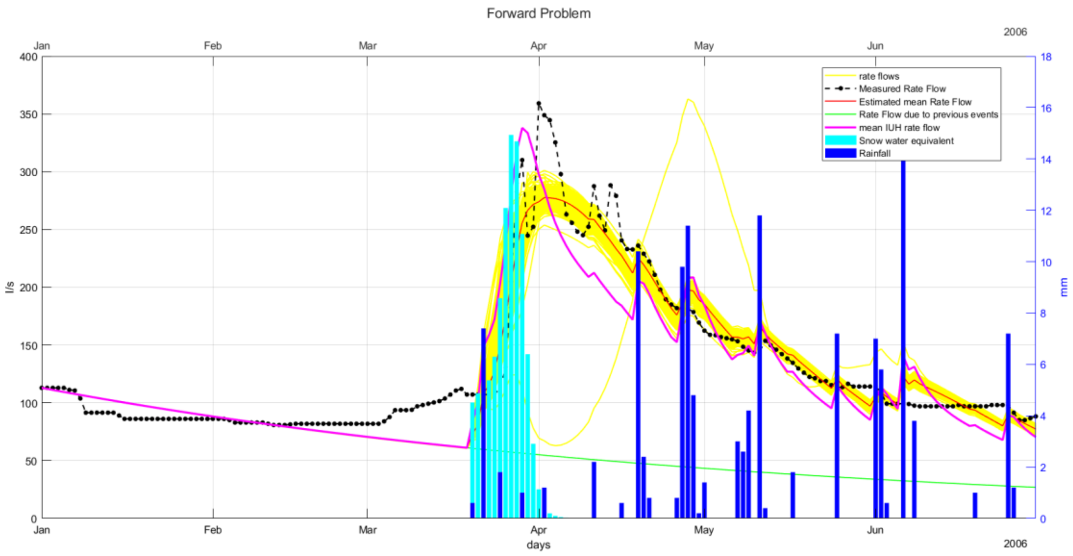


Figure 6-6: Confront between the event specifical IUH response and the average IUH response, parametric IUH, spring 2006.

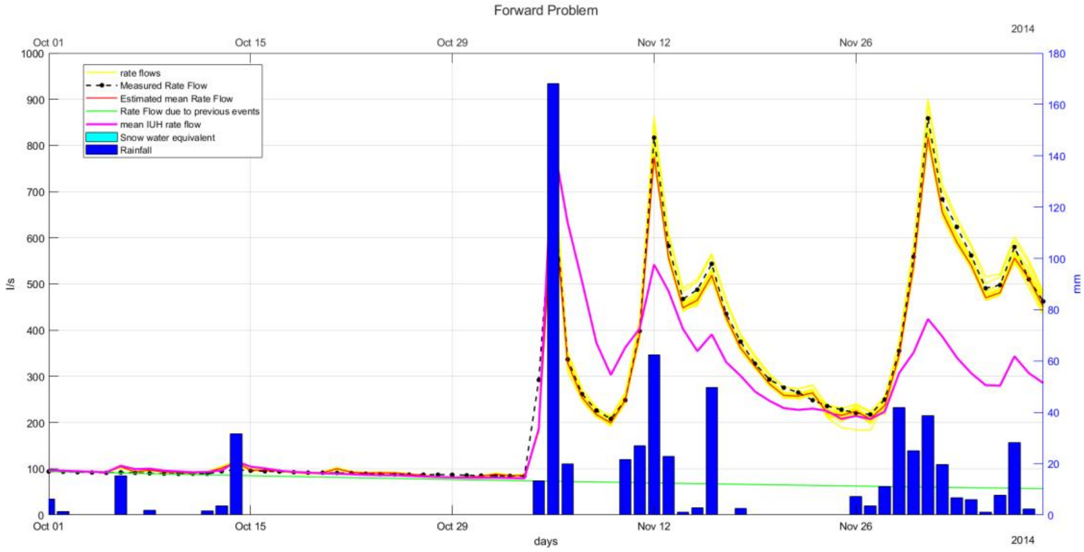


Figure 6-7: Confront between the event specifical IUH response and the average IUH response, non-parametric IUH, fall 2014.

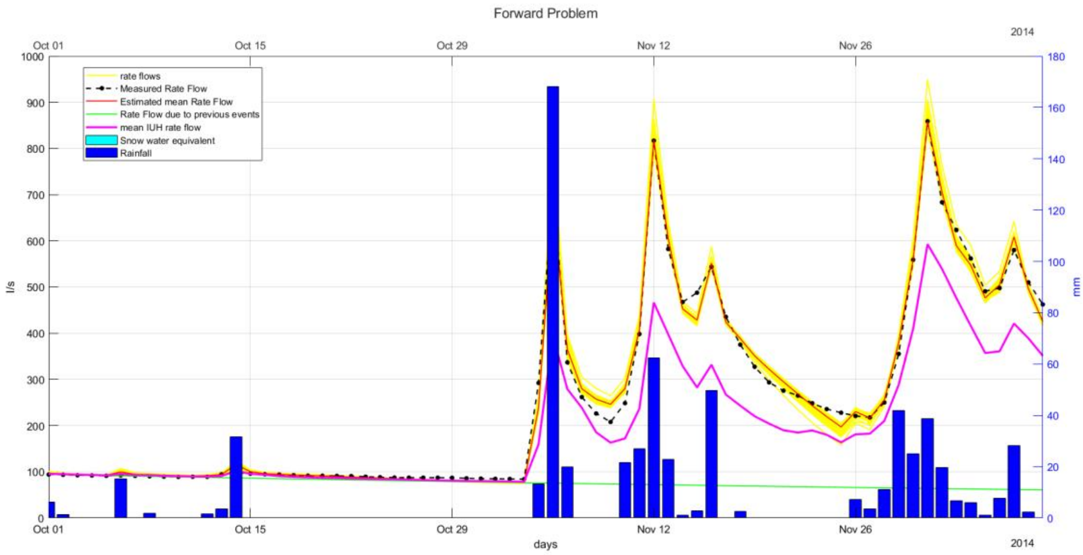


Figure 6-8: Confront between the event specifical IUH response and the average IUH response, non-parametric IUH, fall 2014.

6.2 Of infiltration coefficient

On the front of the infiltration coefficient, in the first instance, the correlations of the estimated data with the rains and its distribution over time were analysed.

The diagrams depicting the ratio between the daily infiltration coefficient and cumulative precipitation in the previous three days are shown below (Figure 6-9) and in ten days before (Figure 6-10). The results from this point of view are completely comparable both when obtained with parametric IUH and with non-parametric IUH. As can be seen from the graphs, no trend is observed in the distribution of data.

When the data is arranged in chronological order, as in Figure 6-11 and Figure 6-12, we can notice a certain repetitiveness in the behaviour of the infiltration coefficient. For this reason, it was decided to continue the analysis following a temporal criterion.

Specifically, the vectors of average values calculated were compared on the basis of the corresponding day of the year. In this way it was possible to estimate an average infiltrative year from January 1st to December 31st. The development of a solid statistic is hampered by the low number of data available, generally around ten values per day. In particular, the summer months, little affected by flood events and therefore not very represented, suffer from a dramatically reduced sample population, around 3-4 values per day.

An initial evaluation of the statistical distribution of daily values using Pearson's fit test, where the sample size allows it, suggests the approximation with normal distribution truncated in the interval [0,1] as acceptable.

In Figure 6-13 the values referring to the model with non-parametric IUH are shown, in Figure 6-14 with parametric IUH. The first shows a strong discontinuity of the average values in early August, where the data are reduced in number, as already mentioned. For the rest there is a comparable trend between the two graphs, with a period of high infiltration between March and May, with higher maximum values when referring to the model with non-parametric IUH. Subsequently, the coefficient gradually decreases in value, until it reaches its minimum around October and then rises again.

Although it may seem counterintuitive a reduction in infiltration during a period typically characterized by flood events, it is possible to make hypotheses to justify this behaviour. As already mentioned, the high intensity of storm events typical of the season leads the influx to move mainly through sinkholes and karst tunnels, while the fractured portion of the aquifer is only marginally affected, for this reason large portions of the territory are subject to strong runoff.

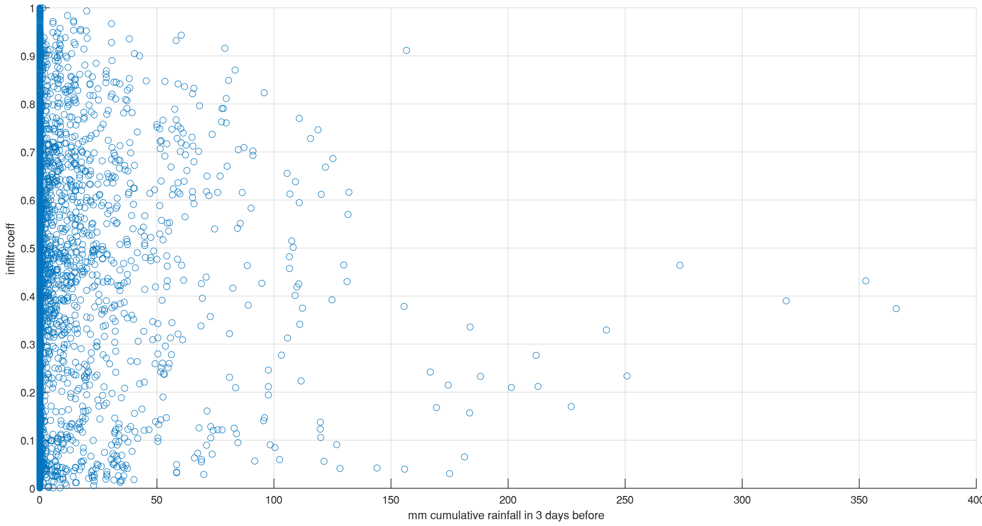


Figure 6-9: Coefficient of infiltration in relation to the accumulated precipitation in the previous three days. Values estimated with parametric IUH. The result is completely comparable with that obtained thanks to the non-parametric IUH.

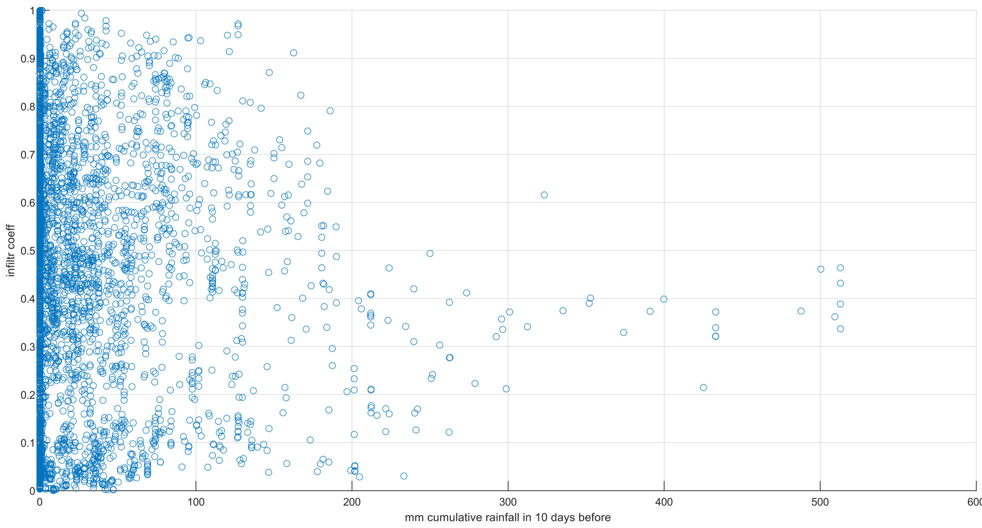


Figure 6-10: Coefficient of infiltration in relation to the accumulated precipitation in the previous ten days. Values estimated with parametric IUH. The result is completely comparable with that obtained thanks to the non-parametric IUH.

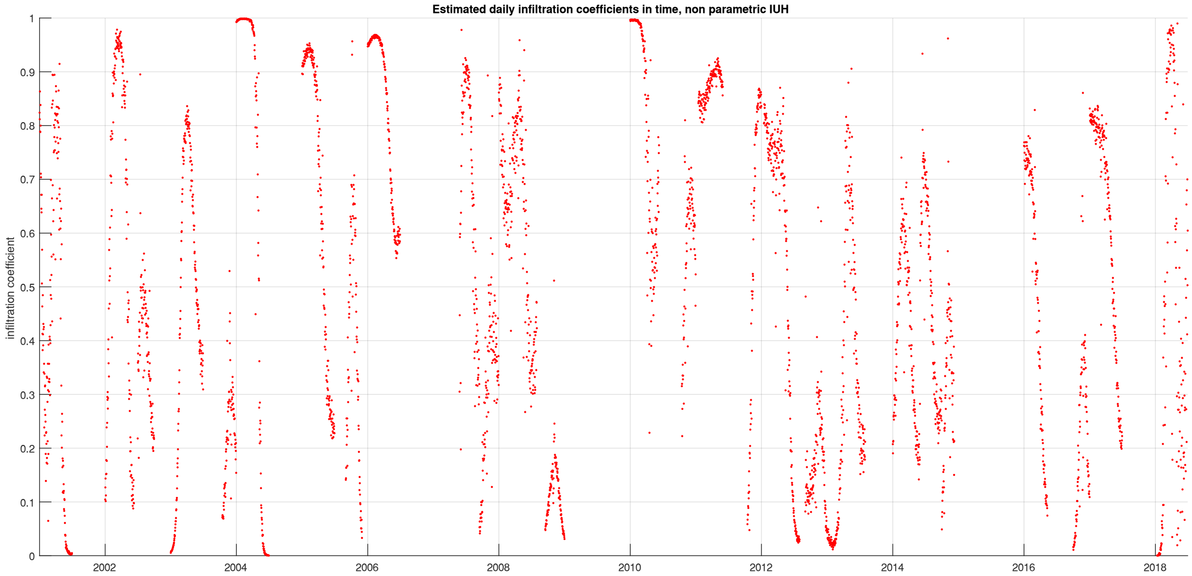


Figure 6-11: Distribution of daily infiltration coefficients over time, model with non-parametric IUH.

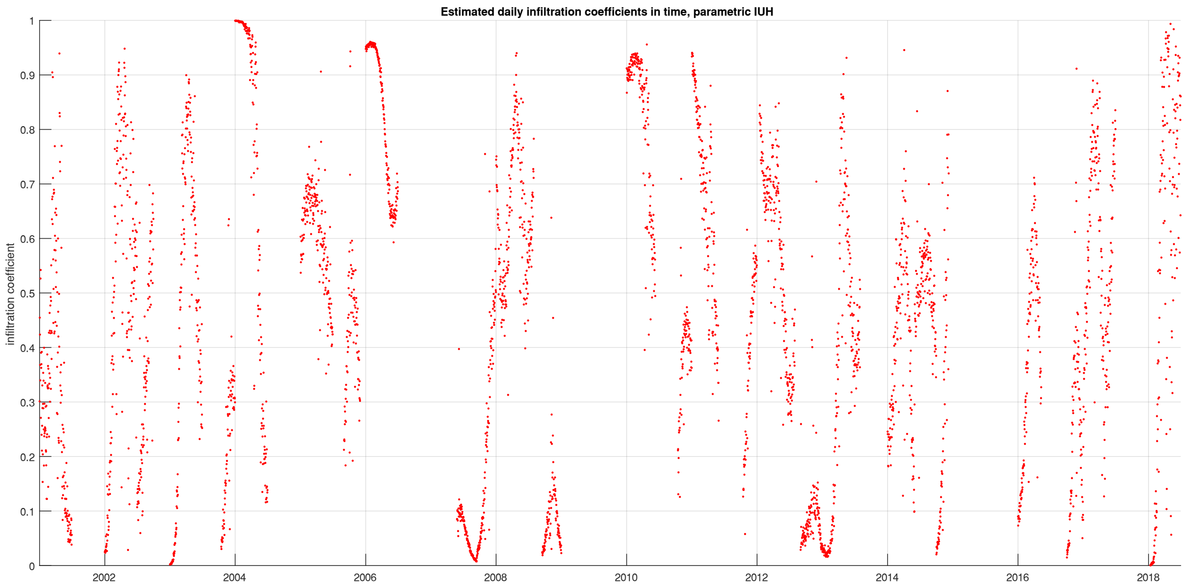


Figure 6-12: Distribution of daily infiltration coefficients over time, model with parametric IUH.

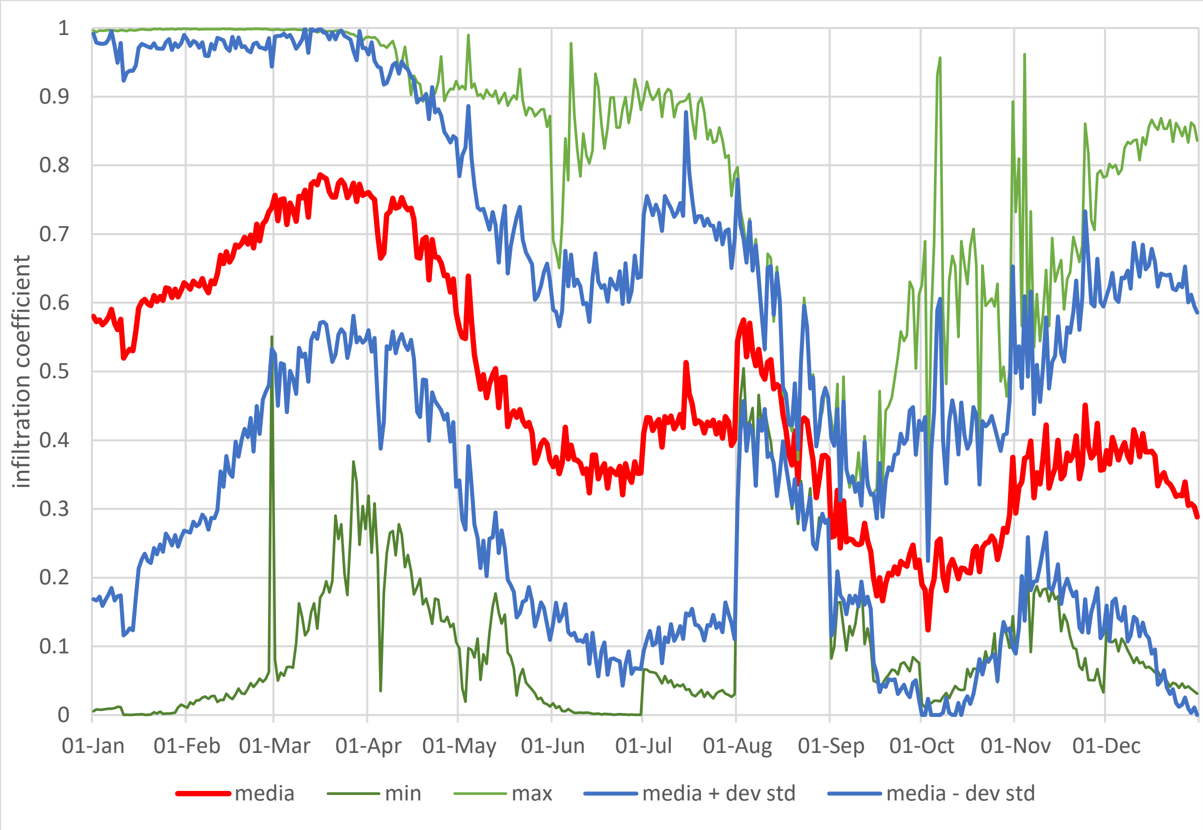


Figure 6-13: Average yearly infiltration calculated with a non-parametric IUH model.

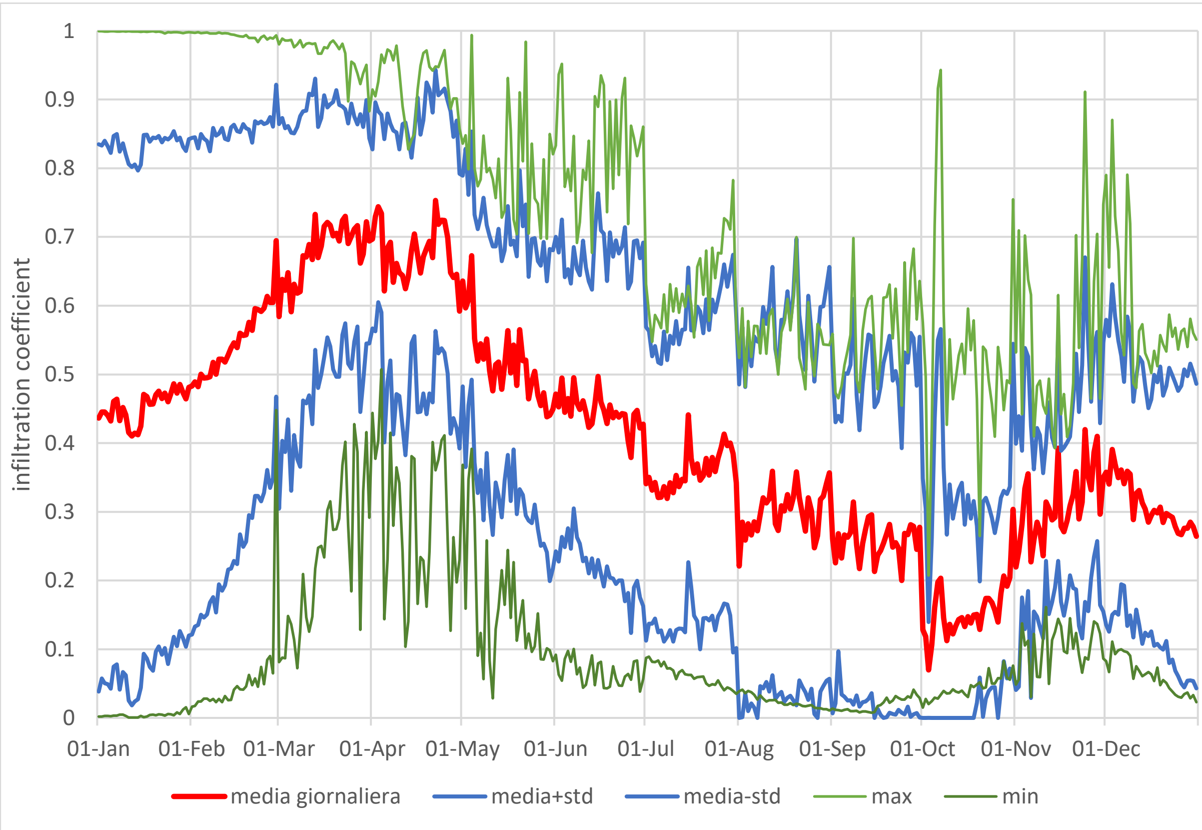


Figure 6-14: Average yearly infiltration calculated with a parametric IUH model.

6.3 Of flow rate B

As regards the flow rate B, the outflow generated by precipitation events prior to the period considered, we have taken into account the values of the regulating parameter of the (5.7), q_2, to evaluate its statistical distribution. Again, the Pearson test is adopted for the purpose. The analysis of the results obtained with non-parametric IUH considers to be valid a log-normal distribution with parameters $\mu = -2.55$ $\sigma = 1.42$, and probability of exceeding equal to 0.09, or a gamma distribution with parameters $a = 0.70$ $b = 0.26$, with probability of exceeding equal to 0.36. The results are reported in Figure 6-15 in the form of cumulative distribution functions. The results obtained with parametric IUH are interpreted well through a log-normal distribution of parameters $\mu = -4.22$ $\sigma = 1.01$, with probability of exceeding equal to 0.50. Again, the empirical and theoretical cumulative distribution functions are reported in Figure 6-16.

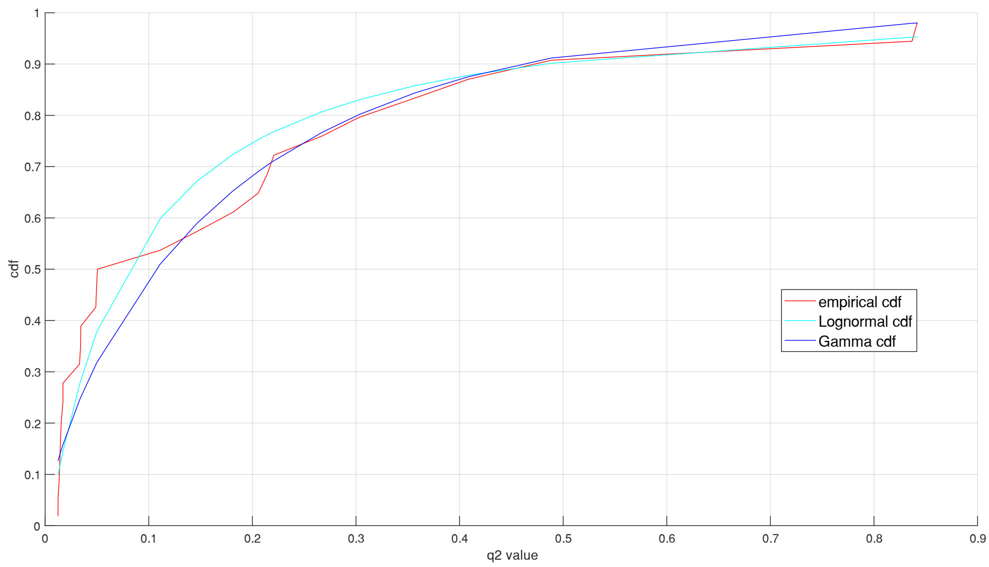


Figure 6-15: Cumulative distribution functions, empirical and theoretical, for the parameter q_2 obtained from the model with non-parametric IUH

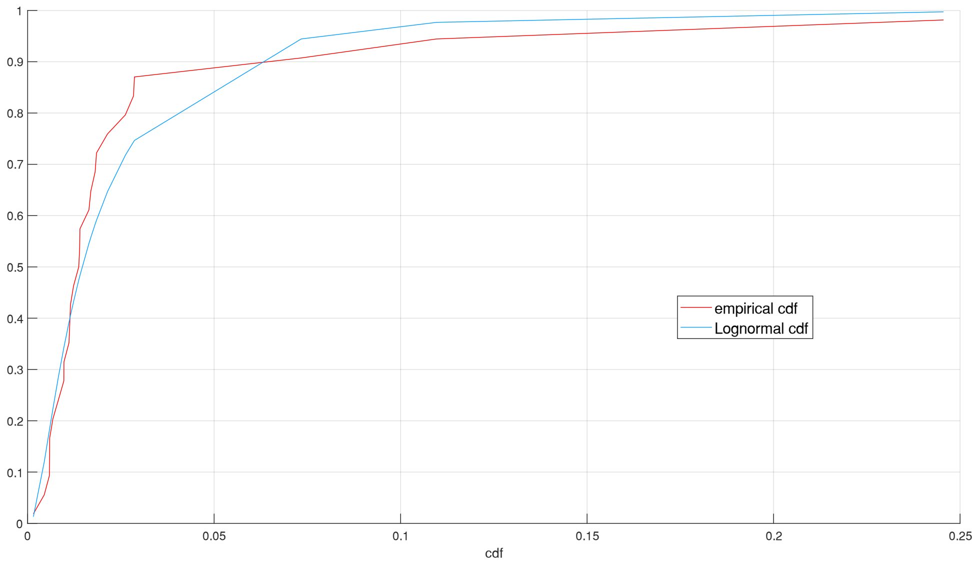


Figure 6-16: Cumulative distribution functions, empirical and theoretical, for the parameter q_2 obtained from the model with parametric IUH.

Furthermore, the relationship between the regulator parameter q_2 and the flow rate on the first day of the series, to which the flow rate B is constrained by construction, was evaluated. However, the data, even separated on a seasonal basis, do not have a reciprocal dependence, as can be seen in the images below.

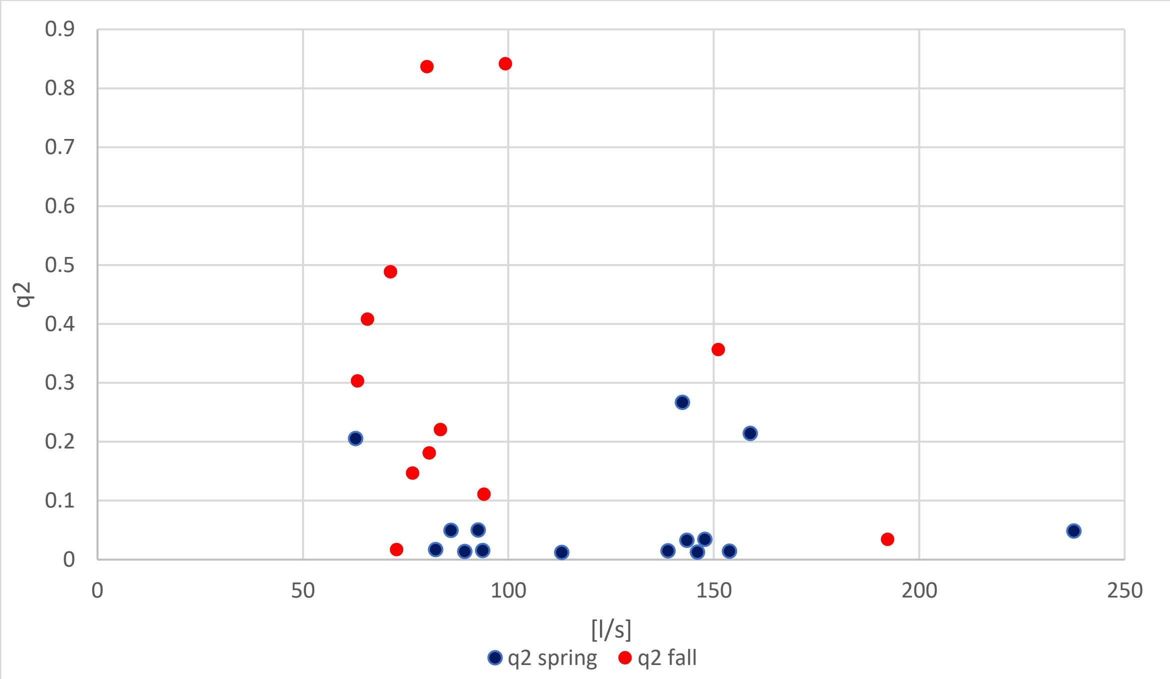


Figure 6-17: Scatter plot of parameter q_2 and flow rate on the first day of the series. Non-parametric IUH.

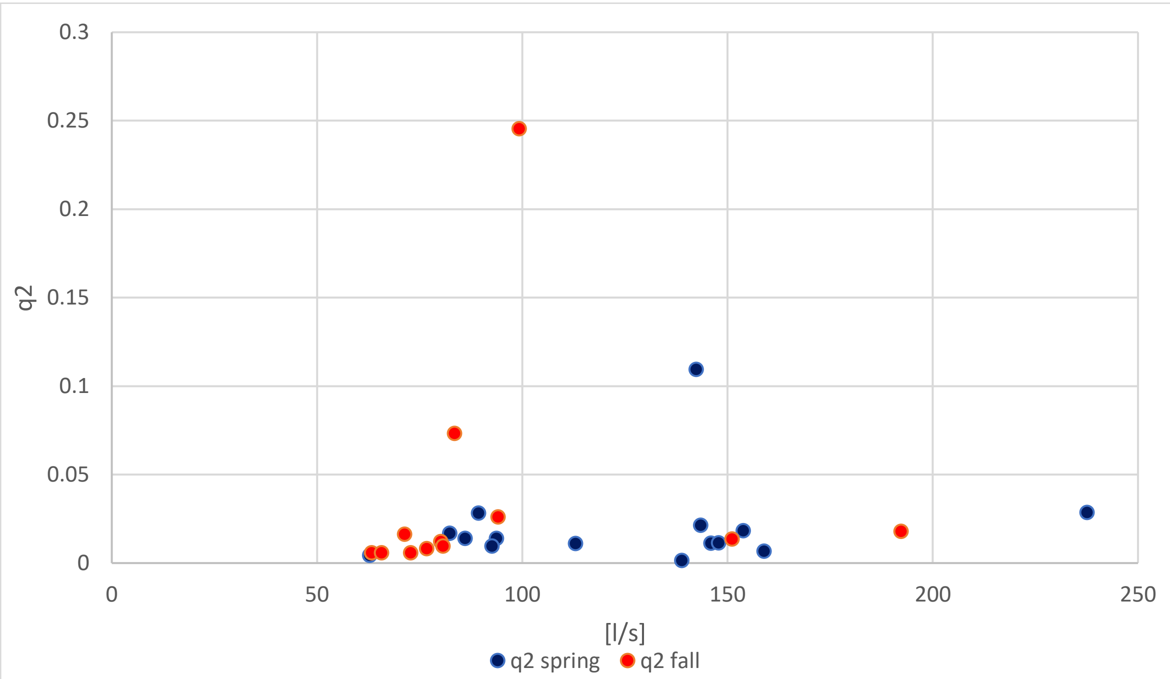


Figure 6-18: Scatter plot of parameter q_2 and flow rate on the first day of the series. Parametric IUH.

6.4 Of snow melting

The data collected on the equation (5.11), used to describe the progress of the snow merger, are quite variable. An analysis performed using the χ^2 test reveals that the average parameters of the events estimated thanks to the model with parametric IUH can be approximated to random variables belonging to a gamma or log-normal distribution, each of suitable parameters. On the other hand, what is obtained from the non-parametric model is adequately described by gamma distributions, again of suitable parameters. In Table 6.1 the values of these statistical distributions are summarized.

Table 6.1: Parameters of the gamma distributions able to describe the parameters of the snow melting equation.

non-parametric IUH – gamma distribution parameters	a_1	b_1	a_2	b_2
a	1.84	0.87	1.62	1.53
b	4.45	11.49	6.64	4.61
parametric IUH – gamma distribution parameters	a_1	b_1	a_2	b_2
a	1.51	1.38	1.18	1.56
b	10.14	6.83	9.59	5.19

In general, the snow melting is closely linked to the temperature and the volume of snow fallen, this makes it a phenomenon difficult to describe in detail. This strong variability is reflected in the diversity of the values assumed by the parameters in the various events studied, as well as in the diversity of days involved. The impossibility of accurately predicting daily temperatures makes it impossible to establish in advance when the snow melt will occur, if not in a general way.

6.5 *Is it possible to identify a system signature?*

In other words, to conclude the analysis of the results hitherto carried on, we wonder if it is possible to identify a set of parameters capable of describing the behaviour of the system with a good approximation in the face of precipitation, an average model with some predictive skills.

Now, from the analyses carried out, it is necessary to extrapolate this average, possibly representative model. The choices made to do so are then described:

- The IUH was selected differently for the two developed models: the non-parametric IUH is the daily average of the mean values recorded for the different events, divided into spring and summer-autumn; the parametric IUH was selected as the interpolation model of the average daily response (which was discussed in the paragraph 6.1)
- The outflow B depends on the parameter called q_2 and the flow rate at the beginning of the series. This can be considered known (the last measured value available before the forecasting process), while for the parameter q_2 it was decided to use the median (more stable than the average) of the estimated values thanks to the analysis of the events, again following a seasonal classification between spring and summer-autumn
- The infiltration coefficient was calculated as the average daily coefficient (see Figure 6-13 and Figure 6-14). In addition, the average values to which the daily standard deviation is subtracted and the average values to which the daily standard deviation is added are considered as borderline cases
- The snow melting is calculated by assuming the median of the regulatory parameters. As for the days in which the phenomenon occurs, the daily temperatures are considered to be known. A temporal and quantitative abstraction of the phenomenon requires accurate analysis not yet carried on at this stage

Ultimately there are two average models depending on the choice of IUH made.

Once an identification response has been defined in this way, flood events are required to test this "system signature". It was therefore decided to pre-select four events, two spring and two autumns, drawn by lot: Spring 2004, Spring 2013, Autumn 2010, and Autumn 2007. For each event, the efficiency of Nash-Sutcliffe and the volume error are measured each in three different values representing respectively the model with mean infiltration coefficient minus standard deviation, mean infiltration coefficient and mean infiltration coefficient plus standard deviation. At the end of the paragraph, the summary tables of the different efficiencies measured and the comparison graphs between the calculated responses and the data actually measured are shown.

From what can be observed thanks to the NSE values, the models behave better than the average, in general terms. On the other hand, the considerable variability of the responses makes even large errors possible. The phenomenon of snow melting is described in a satisfactory manner, but the impossibility of knowing the temperatures in large advance prevents it from being used for long-term forecasts. The model struggles to predict the long-term depletion curves of the aquifer, both as regards the tail of the IUH and on the side of the flow rate B.

The predictive model was evaluated in terms of overall outcoming volumes as well, compared with the real values in terms of percentage error. We can say that the adoption of the average infiltration coefficient reduces the deviation from reality, although the mistakes made are still quite remarkable in some cases. It is also observed how the confidence interval of the infiltration coefficient leads to extreme errors in both directions: the extreme values often lead to a noticeable underestimation or overestimation. In addition, between the model developed thanks to parametric IUH and the counterpart there does not seem to be any evident differences for the autumn season, while in spring there is an average better result in the first case, parametric IUH, whereas the model with non-parametric IUH tends to underestimate the outgoing volumes.

The discussion regarding the signature of the aquifer will be resumed in the next chapter.

Spring 2004 forecast	$\chi - std$	χ	$\chi - std$	
NSE	-0.41	0.28	0.58	Non-parametric IUH
Volume error (%)	-53.75	-33.20	-12.65	
NSE	0.23	0.69	0.81	Parametric IUH
Volume error (%)	-37.96	-20.41	-2.87	

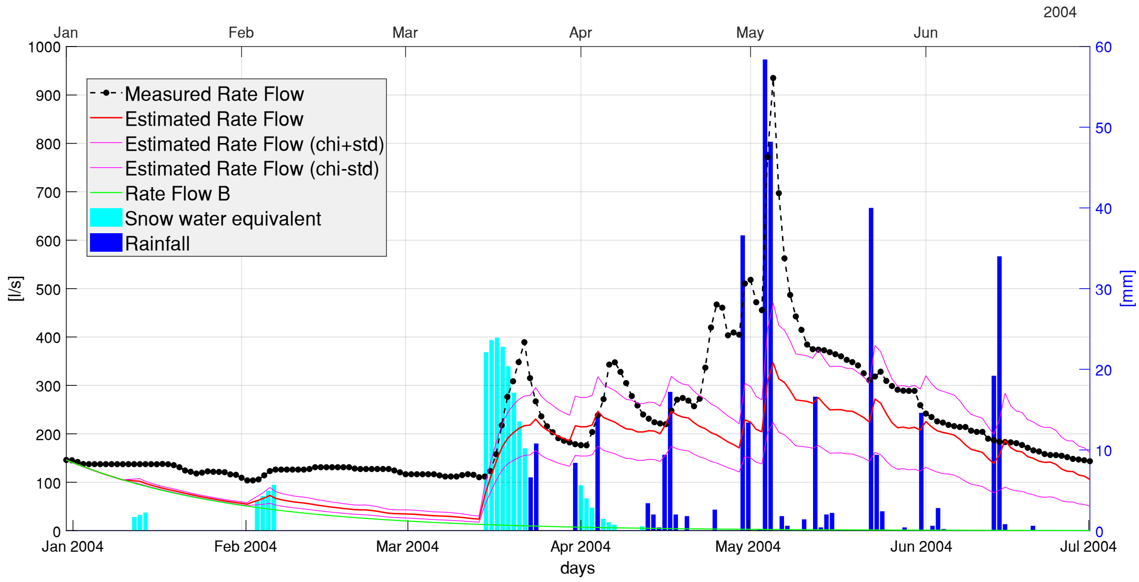


Figure 6-19: Spring 2004, non-parametric IUH.

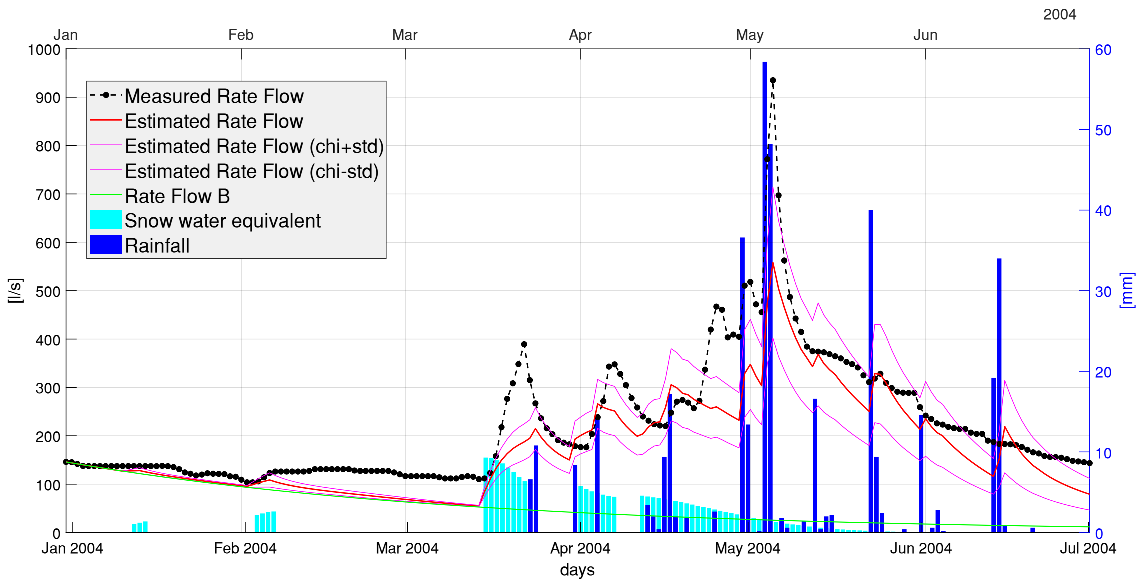


Figure 6-20: Spring 2004, parametric IUH.

Fall 2007 forecast	$\chi - std$	χ	$\chi - std$	
NSE	-3.48	-0.49	-1.70	Non-parametric IUH
Volume error (%)	-80.89	-36.14	10.48	
NSE	-1.39	-1.42	-8.07	Parametric IUH
Volume error (%)	-51.82	-4.33	47.12	

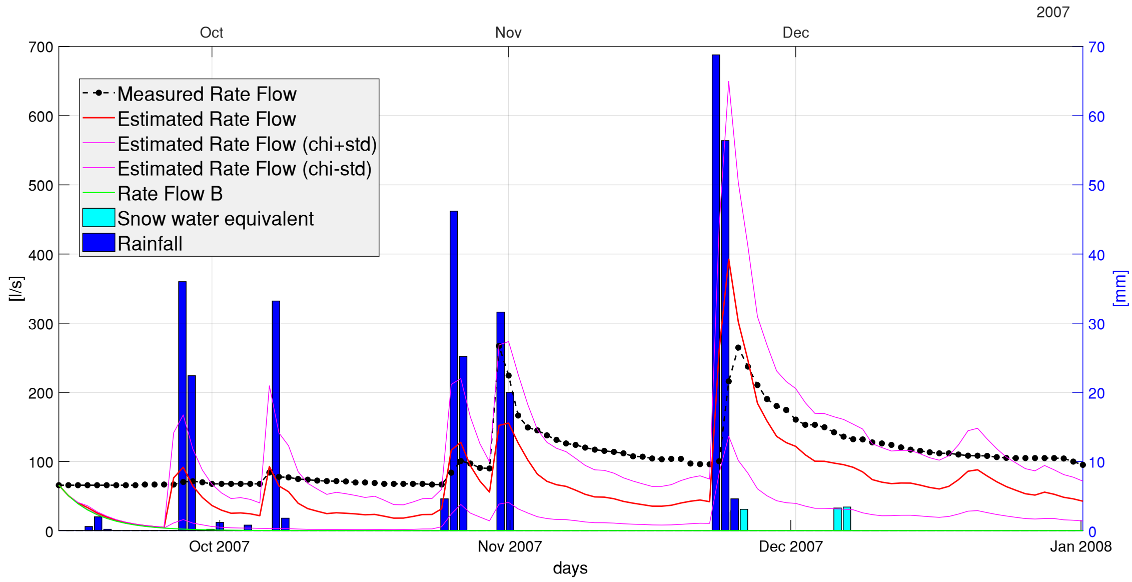


Figure 6-21: Fall 2007, non-parametric IUH.

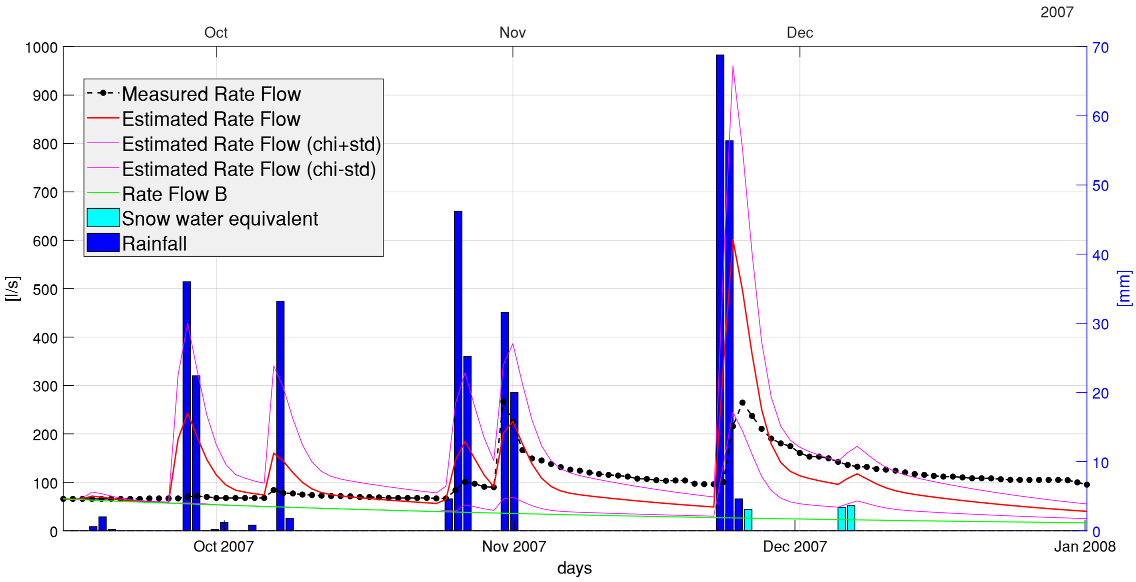


Figure 6-22: Fall 2007, parametric IUH.

Spring 2013 forecast	$\chi - std$	χ	$\chi - std$	
NSE	0.59	0.83	0.54	Non-parametric IUH
Volume error (%)	-34.69	0.27	35.23	
NSE	0.60	0.72	0.41	Parametric IUH
Volume error (%)	-27.98	1.84	31.65	

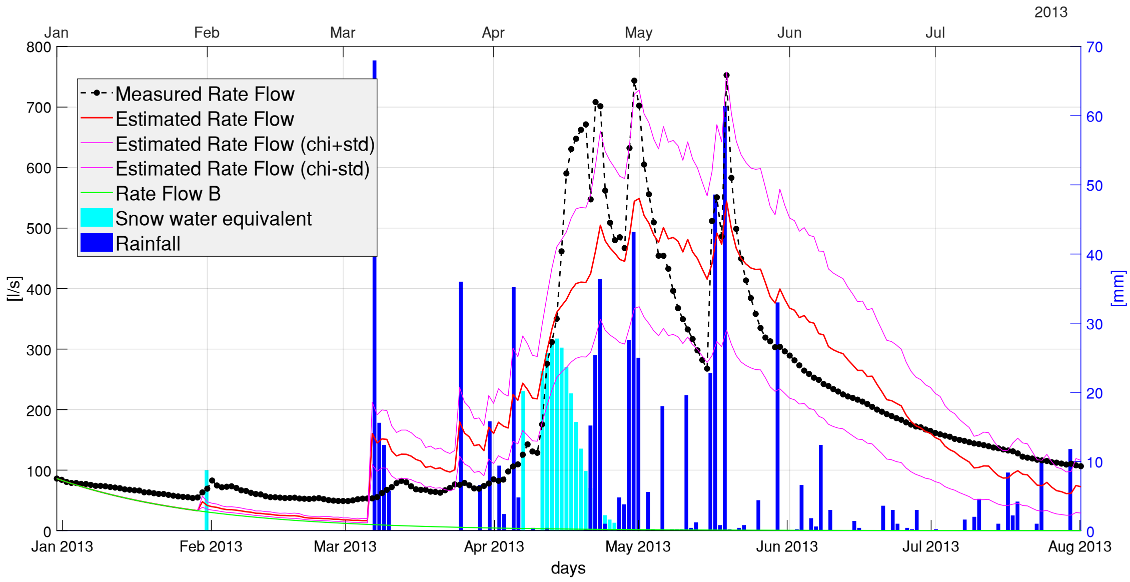


Figure 6-23: Spring 2013, non-parametric IUH.

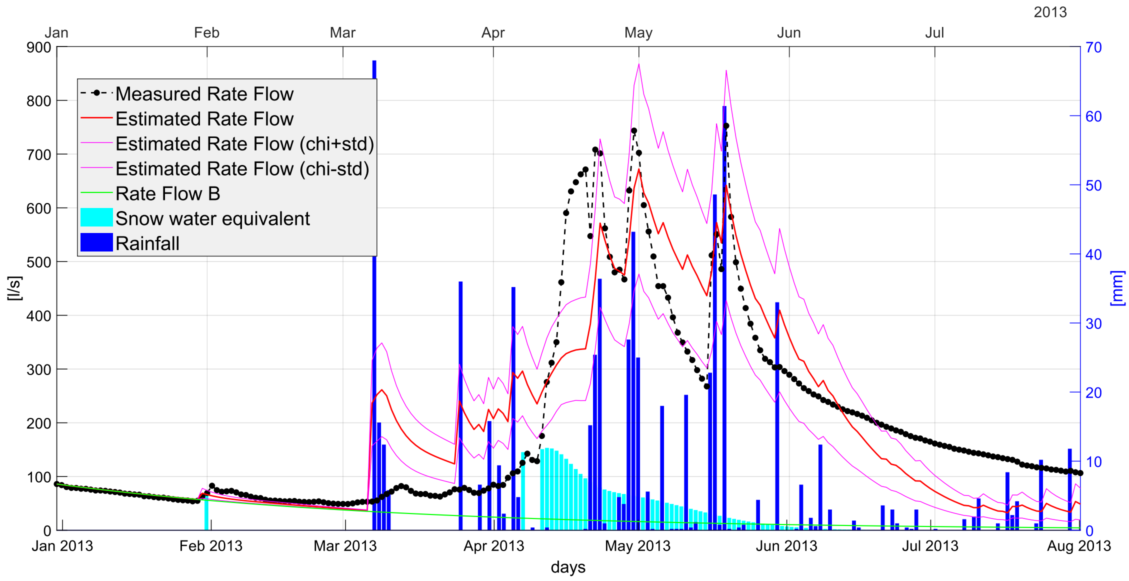


Figure 6-24: Spring 2013, parametric IUH.

Fall 2016 forecast	$\chi - std$	χ	$\chi - std$	
NSE	0.06	0.74	0.61	Non-parametric IUH
Volume error (%)	-71.24	-15.49	40.25	
NSE	0.55	0.84	0.16	Parametric IUH
Volume error (%)	-36.16	16.11	68.88	

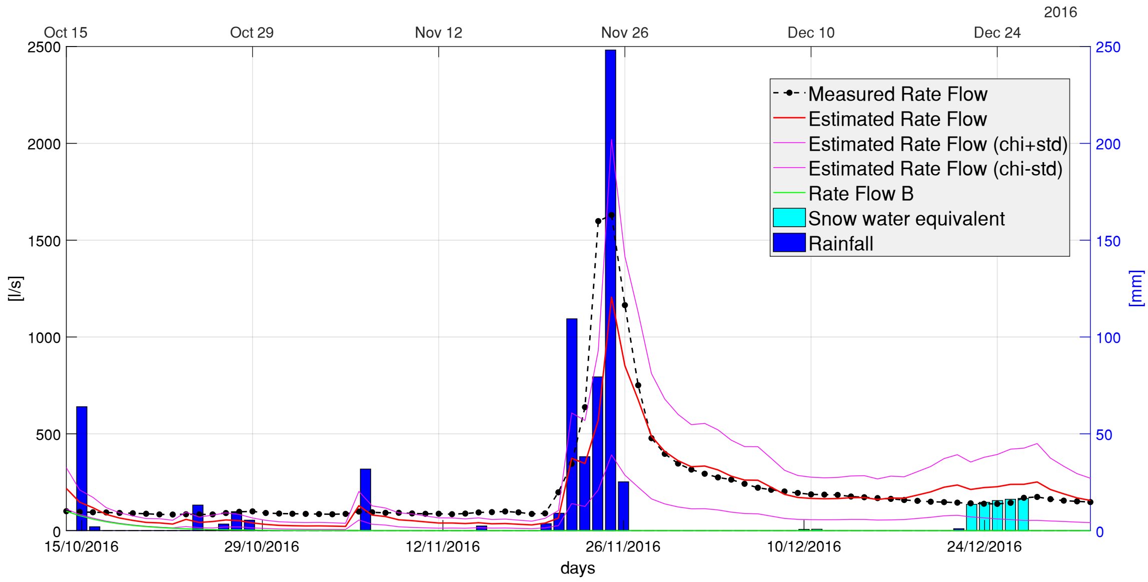


Figure 6-25: Fall 2016, non-parametric IUH.

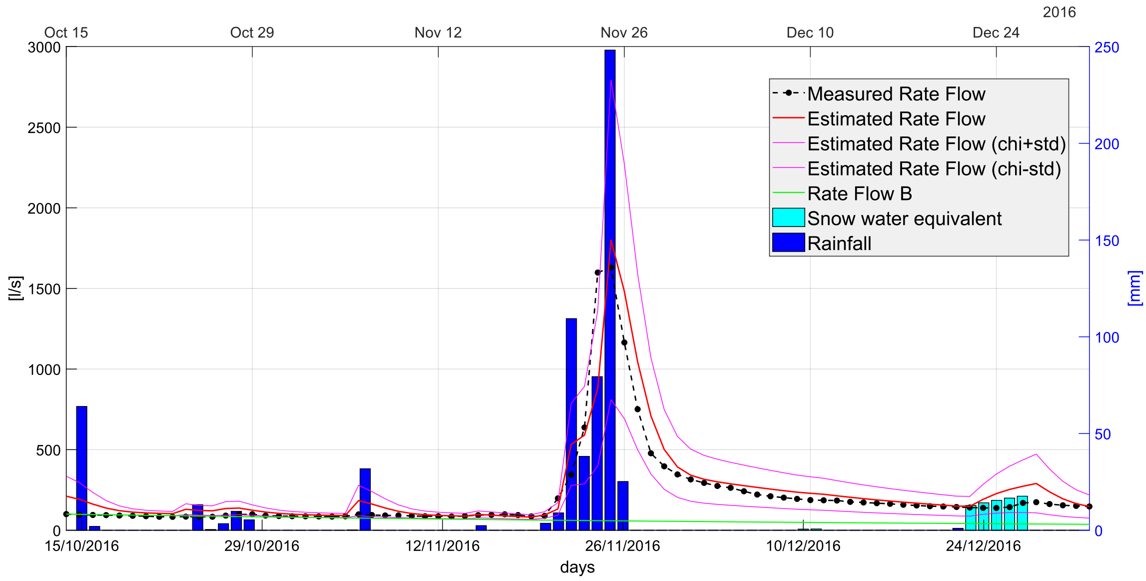


Figure 6-26: spring 2016, parametric IUH.

7 Conclusions

Since now we have seen the model chosen to describe our karst system and the methods used to concretely define it as well as some analysis performed on the data consequently obtained. We now try to discuss its effectiveness.

The data collected thanks to the analysis of flood events in the period 2001-2018, in Appendix A, suggest a good ability to estimate the response of the system *a posteriori*, that is, when the extent data of the event are available, and it is therefore possible to proceed with methods of data assimilation. This system response also varies considerably from one event to another. This reflects the complexity of the aquifer, consisting of fully loaded pipelines, fractured limestones and affected by transfers from secondary detrital aquifers.

In general, the use of a non-parametric IUH allows for a more efficient solution, thanks to the ability to better describe the irregularity of the response. On the other hand, the parametric IUH reduces the accuracy of the response on a single event, but, as seen in paragraph 6.1, the different events, grouped according to seasonal criteria, have a more homogeneous behaviour, making it easier to identify a characteristic trend.

The infiltration coefficient is characterized by considerable random behaviour. There are remarkable differences between chronologically contiguous values, and this has made it impossible to adopt some sort of parametric equation to describe their progress. Instead, we have chosen to use discrete daily values. This approach proved to be better than a single constant infiltration coefficient over the selected period. This calls into question the first hypothesis posed in paragraph 3.1, the uniformity of precipitation on the basin, because part of the high variability of the parameter in question can be explained by uneven precipitation on the feeding area, the intensity of which varies in space and with the altitude, thus affecting different portions of the aquifer in different ways. In any case, the infiltration coefficient is deeply linked to many environmental factors, such as temperature, soil and air moisture, vegetation behaviour, and the distribution of precipitation, therefore its irregularity is necessary to ensure sufficient variability for our approximate model. In this sense, this parameter conceals, summarizes, and compresses part of the complexity typical of natural phenomena that we have tried to describe up to now.

The snowfall and snow melting forecast model, although developed by observing the flow rate trend and therefore still to be compared with snow data collected in the field, was found to be fundamental for the development of an efficient model. The influence of snow melting in spring is remarkable, just as in autumn the contribution of snowfall has no (immediate) influence on the watershed. In this way we were able to correct the contributions and redistribute them in the most appropriate period. This improved the quality of the results and allowed a temporal extension for the single event analysed without deteriorating the final response.

The flow rate B, designed to reduce the influence of previous rainfall, underlines a weakness of the model: the difficulty in predicting long-term exhaustion queues. The response through IUH, especially in the parametric form, finds it difficult to correctly describe the gradually decreasing flows in the absence of precipitation for prolonged times, becoming dependent on the rate flow B to guarantee this contribution. In general, given the hypothesis of stationarity of physical processes, the tail of the instantaneous unit hydrogram and the flow rate B should have a comparable trend, while this is not the situation in most cases. This suggests different behaviour between different parts of the system, described however through a single answer (the IUH). This approximation is therefore incomplete and unable to fully describe the system in all its complexity.

The previous chapter saw an attempt to identify a unique system signature. Now, we can say that this is not possible on all state parameters. The autumn IUH of the different events are similar, so the approximation to an average response is acceptable, especially with the use of the parametric form. The spring response is more varied, but its close link with snow melting means that overall, it is possible to obtain a typical seasonal behaviour described by the sum of these two components and able to predict events with a good approximation. It should be remembered again that the temperature-dependent snow melting process is difficult to predict before time in its temporal development. The infiltration coefficient has a repetitive trend on an annual basis, however significant variations with respect to the average are possible and probable. Furthermore, we can affirm that the single events tend to present a moment of maximum infiltration and then the values usually decrease after this maximum. Finally, the flow rate B is the most difficult parameter to approximate to a univocal response. Its contribution to the outflow is often significant and this involves a non-negligible level of error. However, this part of the model, being dependent on the antecedent data, can be guessed by those, as an interpolation of flow rate data. The behaviour expected thanks to the model with parametric IUH on this parameter seems to be better than that obtained with non-parametric IUH, with a more distributed trend over time.

Ultimately, if an approximate answer in terms of seasonal IUH seems possible, the trend of the other parameters is difficult to predict accurately, leaving room for statistical evaluations. Generally speaking, an answer estimated in the ways seen above seems to behave better than the average of the flow rates, but it is possible to make even considerable errors. The system signature predicted through the model with parametric IUH seems to have greater predictive capabilities than the counterpart, although the model is more imprecise on individual events.

7.1 Possible developments

At the end of the conclusions, some future developments of the work described so far are hoped for. First of all, the application to different aquifer systems would allow a more complete study of the algorithm and possible improvements. A further study of the dependencies of the infiltration coefficient on other environmental parameters, such as soil moisture or air temperature, could allow to restrict the variability of the parameter. The model used to define the snow melting requires in-depth analysis to define its validity and applicability, perhaps by applying the model in locations where snow data are available. It is also possible to evaluate the adoption of other parametric forms than those used, where provided for by the model. A further development could concern an application on a yearly basis data and the comparison with seasonal responses. Finally, the analysis of sufficiently extensive data over time could help identify variations in the behaviour of the aquifer over time and therefore help define future behaviours and possible evolutions.

Figure 7-1: Detail of the Mora stream inside the Bossea Cave. Photo by A. Morabito.

8 Bibliography

[1] M. Civita, F. Gregoretti, A. Morisi, G. Olivero, G. Peano, B. Vigna, E. Villavecchia and F. Vittone, Atti della stazione scientifica di della Grotta di Bossea, Savigliano: Gruppo Speleologico Alpi Marittime C.A.I. Cuneo - Dipartimento Georisorse e Territorio del Politecnico di Torino, 1990.

[2] M. Vanossi, "Analisi stratigrafico-strutturale della zona tra le alte valli del Casotto e dell'Ellero," *Atti SOcietà Geologica Universita di Pavia,* vol. 24, pp. 72-120, 1974.

[3] A. Hartmann, N. Goldscheider, T. Wagener, J. Lange and M. Weiler, "Karst water resources in a changing world: Review of hydrological modeling approaches," *Reviews of Geophysics,* no. 52, pp. 2018-242, 2015.

[4] V. T. Chow, D. R. Maidment and L. W. Mays, Applied Hydrology - second edition, Mcgraw-Hill, 2013.

[5] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I — A discussion of principles," *Journal of Hydrology,* vol. 10, no. 3, pp. 282-290, 1970.

[6] R. G. Brown and P. Y. Hwang, Introduction to Random Signals and Applied Kalman Filtrering, United States of America: John Wiley & Sons, Inc., 2012.

[7] A. Almendral-Vazquez and A. Randi Syversveen, "The Ensemble Kalman Filter - theory and applications in oil industry," Norsk Regnesentral - Norwegian Computing Center, Oslo, 2006.

[8] A. Carrassi, M. Bocquet, L. Bertino and G. Evensen, "Data Assimilation in the Geosciences, An overview on methods, issues and perspectives," 2018.

[9] G. Evensen, "Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *JOURNAL OF GEOPHYSICAL RESEARCH,* vol. 99, pp. 10,143-10,162, 1994.

[10] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics,* no. 53, pp. 343 - 367, 2003.

[11] A. A. Emerick and A. C. Reynolds, “Ensemble smoother with multiple data assimilation,” *Computers & Geosciences,* vol. 55, pp. 3-15, 2013.

[12] J. Rafiee and A. C. Reynolds, “Theoretical and efficient practical procedures for the generation of inflation factors for ES-MDA,” *Inverse Problems,* vol. 33, no. 11, 2017.

[13] Z. Chen, T. Xu, J. J. Gómez-Hernández and A. Zanini, “Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter,” *Mathematical Geosciences,* 2021.

[14] X. Liang, X. Zheng, S. Zhang, G. Wu, Y. Dai and Y. Li, “Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble Kalman filter assimilation,” *Quarterly Journal of the Royal Meteorological Society,* no. 138, pp. 263-273, 2012 A.

[15] T. Xu, J. J. Gómez-Hernández, H. Zhou and L. Li, “The Power of Transient Piezometric Head Data in Inverse Modeling: An Application of the Localized Normal-score EnKF with Covariance Inflation in a Heterogenous Bimodal Hydraulic Conductivity Field,” *Advances in Water Resources,* 2013.

[16] J. L. Anderson and S. L. Anderson, “A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts,” *Monthly Weather Review,* vol. 127, pp. 2741-2758, 1999.

[17] M. E. P. Raddusa, F. Boano, A. Viglione and B. Vigna, *Tesi di Laurea Magistrale - Analisi di portate in un sistema acquifero carsico,* Torino, 2020.

[18] “Sito Web ARPA Piemonte,” [Online]. Available: http://www.arpa.piemonte.it/.

[19] G. Evensen, “Sampling strategies and square root analysis schemes for the EnKF,” *Ocean Dynamics,* no. 54, pp. 539-560, 2004.

[20] “openspeleo,” [Online]. Available: www.openspeleo.org.

[21] “La Grotta di Bossea,” [Online]. Available: www.grottadibossea.com.

[22] T. Xu and J. J. Gómez-Hernández, "Joint identification of contaminant source location, initial release time and initial solute concentration in an aquifer via Enseble Kalman Filtering," *WATER RESOURCES RESEARCH,* 2016.

[23] H. H. Bauser, D. Berg, O. Klein and K. Roth, “Inflation method for ensemble Kalman filter in soil hydrology,” *Hydrol. Earth Syst. Sci.,* vol. 22, pp. 4921-4934, 2018.

[24] J. Gómez-Hernández, *How to implement the ensemble Kalman filter,* 2013.

[25] “Gruppo Speleologico Apli Marittime C.A.I. Cuneo,” [Online]. Available: http://speleocuneo.blogspot.com/p/bossea.html.

[26] V. Todaro, M. D'Oria, M. G. Tanda and J. J. Gómez-Hernández, “Ensemble smoother with multiple data assimilation for reverse flow routing,” *Computers & Geosciences,* no. 131, pp. 32-40, 2019.

[27] S. J. Fletcher, Data Assimilation for the Geosciences, Elsevier, 2017.

[28] I. Butera, J. J. Gómez-Hernández and S. Nicotra, “Contaminant-Source Detection in a Water Distribution System Using the Ensemble Kalman Filter,” *Journal of Water Resources Planning and Management,* vol. 147, no. 7, 2021.

[29] W. B. White, “Conceptual models for karstic aquifers,” *Speleogenesis,* vol. 1(1), pp. 1-6, 2003.

[30] P. J. Van Leeuwen, “Comment on ‘‘Data Assimilation Using an Ensemble Kalman Filter Technique’’,” *Monthly Weather Review,* vol. 127, pp. 1374-1377, 1998.

[31] G. Evensen, “Analysis of iterative ensemble smoothers for solving inverse problems,” *Computational Geosciences,* no. 22, pp. 885 - 908, 2018.

[32] A. A. Emerick, “Analysis of geometric selection of the data-error covariance inflation for ES-MDA,” *Journal of Petroleum Science and Engineering,* vol. 182, 2019.

[33] G. Evensen and P. J. van Leeuwen, “An Ensemble Kalman Smoother for Nonlinear Dynamics,” *M O N T H L Y W E A T H E R R E V I E W,* vol. 128, pp. 1852-1867, 2000.

[34] J. L. Anderson, “An adaptive covariance inflation error correction algorithm for ensemble filters,” *Tellus A: Dynamic Meteorology and Oceanography,* vol. 59, no. 2, pp. 210-224, 2007.

[35] "Ambiente carsico e umano in Val Corsaglia," in *Atti dell'incontro di Bossea 14-15 settembre 1991*, Bossea, 1991.

[36] G. Evensen, "Accounting for model errors in iterative ensemble smoothers," *Computational Geosciences,* no. 23, pp. 761-775, 2019.

Appendix A Analysed Events

The following are the 27 flood events analysed with the developed code and the results are reported in the form of diagrams and tables. The results are organized in chronological order, the model with non-parametric IUH precedes the one with parametric IUH. In the time interval of available data, all flood events whose data were complete were selected and analysed, selecting a suitable time window. It is generally more extensive for spring events in order to better consider the snow contribution of winter.

The analysis of the events is summarized as follows: first the starting data are shown (image with the measured flows of the Mora stream, black line, and the meteorological events divided into rains, blue, and snow, blue and image with temperatures and rainfall), the result with non-parametric IUH follows as a table of salient values, final result in terms of IUH, blue, and infiltration coefficient, red, and forward problem (flow rates measured in black and estimated in red, rainfall measured in blue, snow melting estimated in blue), then result with parametric IUH in the same exposure sequence.

The tables summarize some values considered summary of the final model. The bias is the average error of the model, the Mean Absolute Error is the average of the error in absolute value, NSE represents the Nash-Sutcliffe efficiency referred to in paragraph 3.2, the value ${}^{V_{in}}/_{V_{out}}$ indicates the ratio between the input and output volumes estimated by the model (whose value should be equal to 1 in theory), while the remaining parameters concern inbreeding and is dealt with extensively in the paragraphs 4.4 and 5.11.

I. Spring 2001

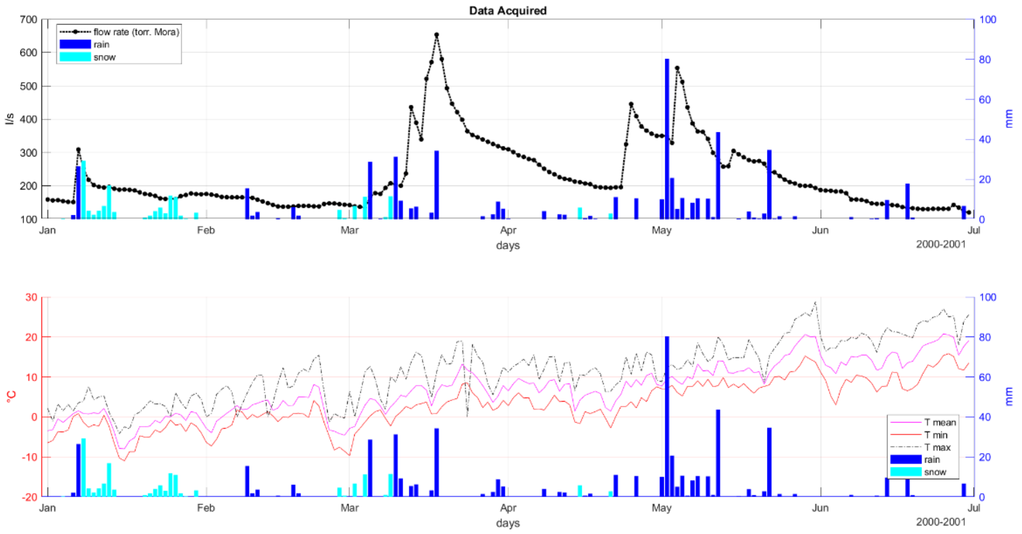
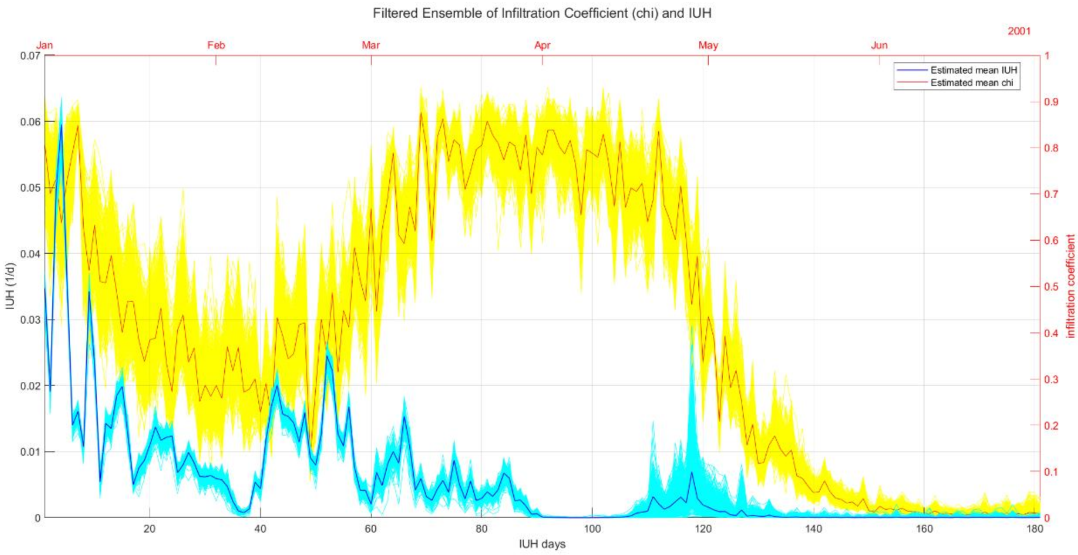
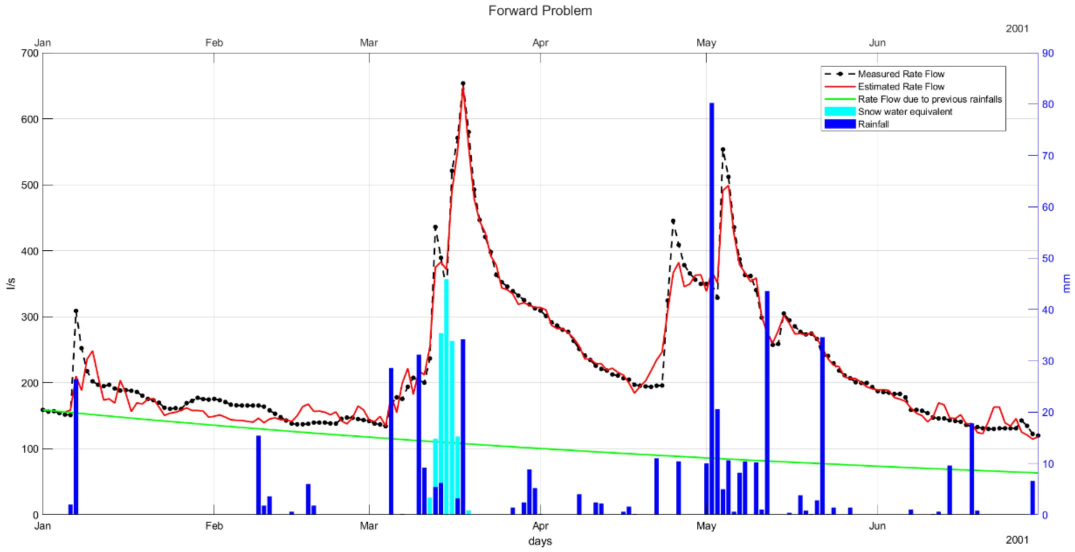


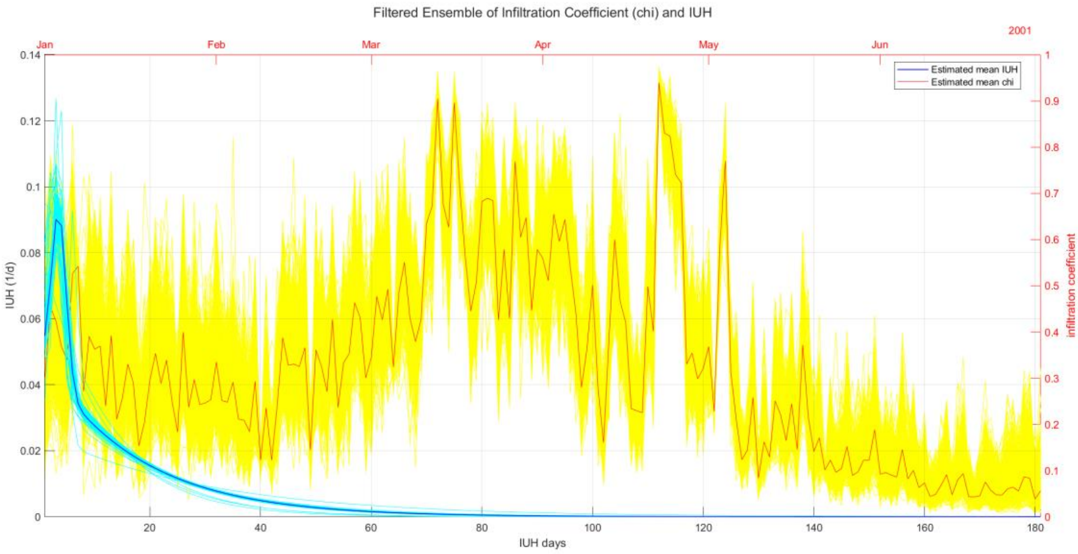
Figure 8-1: data used.

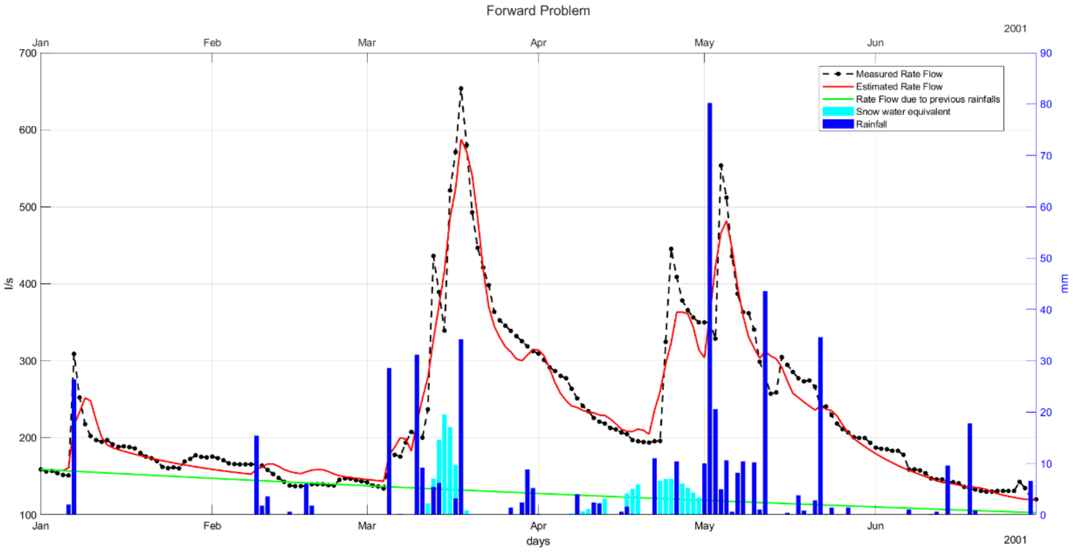
non-parametric IUH					
Bias [l/s]	-3.54	V_{in}/V_{out}	1.08	λ	1.1
Mean Absolute Error [l/s]	13.47	RMSE	20.41	DF	0.7
NSE	0.96	RMSE/ES	5.29	EM	2





parametric IUH					
Bias [l/s]	-2.77	V_{in}/V_{out}	1.02	λ	1.13
Mean Absolute Error [l/s]	16.92	RMSE	27.38	DF	0.4
NSE	0.94	RMSE/ES	5.3	EM	2




Forward Problem
2001
Jan
Feb
Mar
Apr
May
Jun
700
600
500
400
300
200
100
l/s
90
80
70
60
50
40
30
20
10
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days

II. Spring 2002

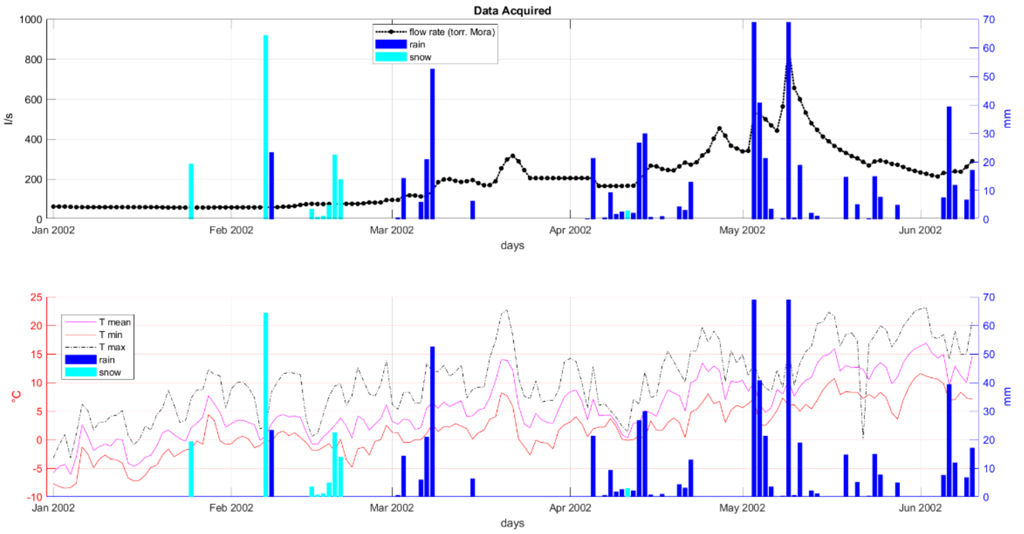
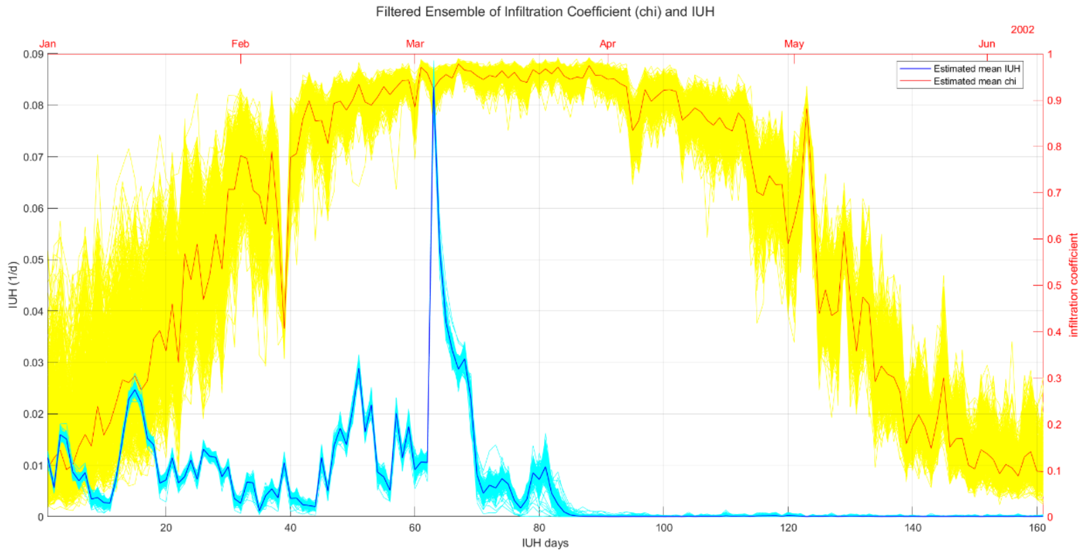
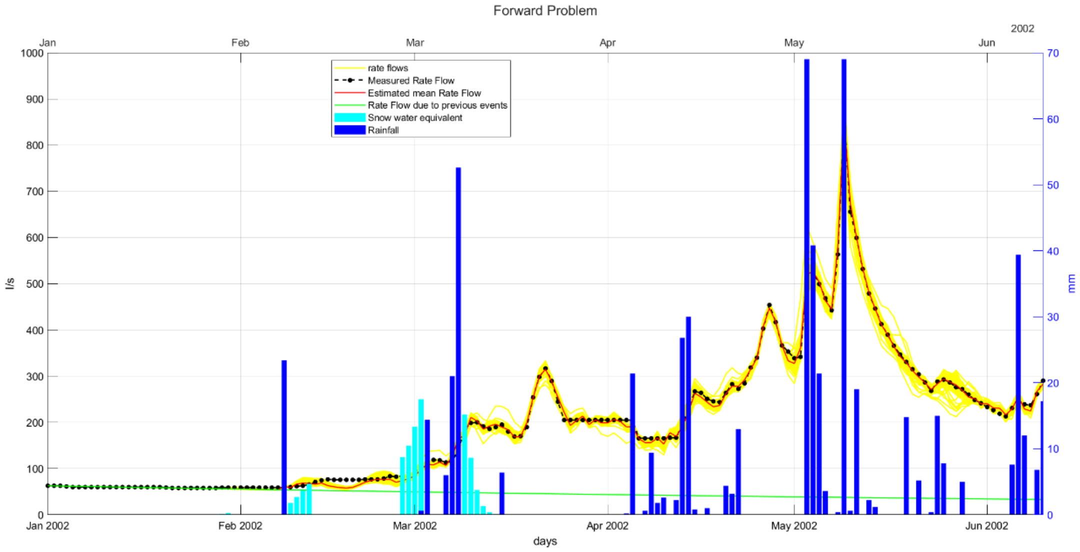


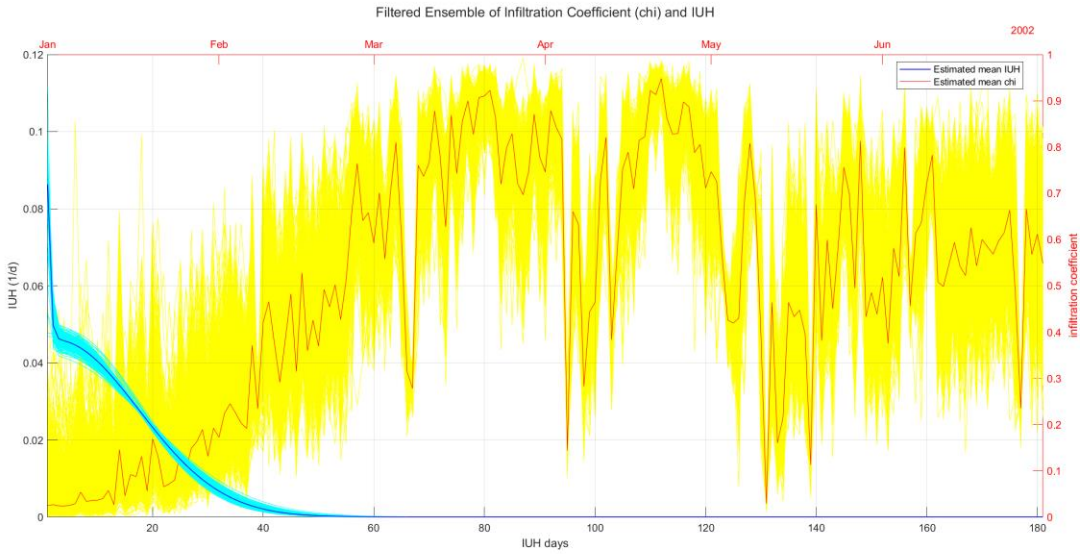
Figure 8-2: data used.

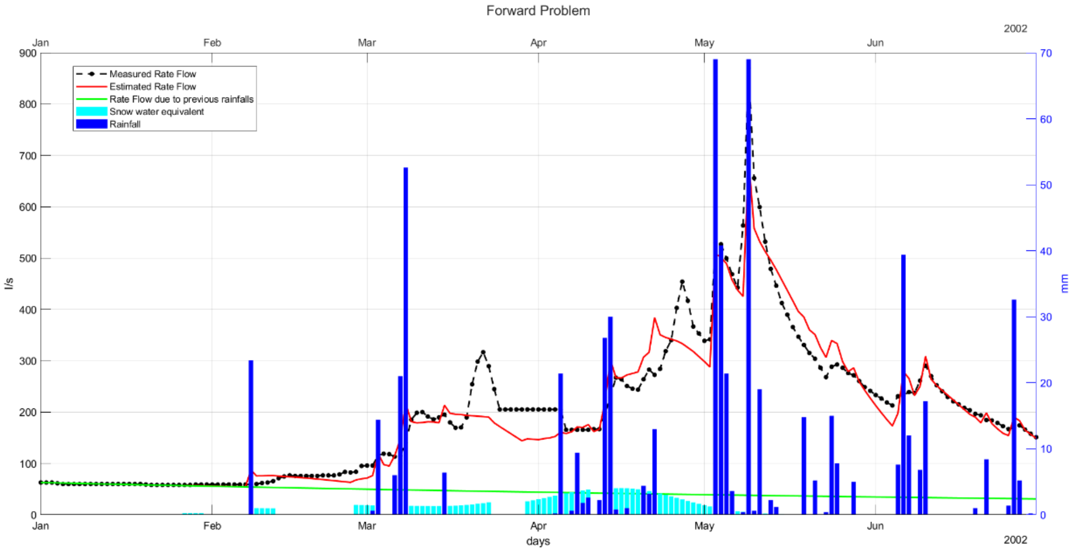
non-parametric IUH					
Bias [l/s]	-1.81	V_{in}/V_{out}	1.43	λ	1.12
Mean Absolute Error [l/s]	5.57	RMSE	8.72	DF	0.8
NSE	1.00	RMSE/ES	1.87	EM	2





parametric IUH					
Bias [l/s]	-5.73	V_{in}/V_{out}	1.05	λ	1.15
Mean Absolute Error [l/s]	23.71	RMSE	38.86	DF	0.4
NSE	0.92	RMSE/ES	6.52	EM	4





Forward Problem
2002
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days

III. Summer 2002

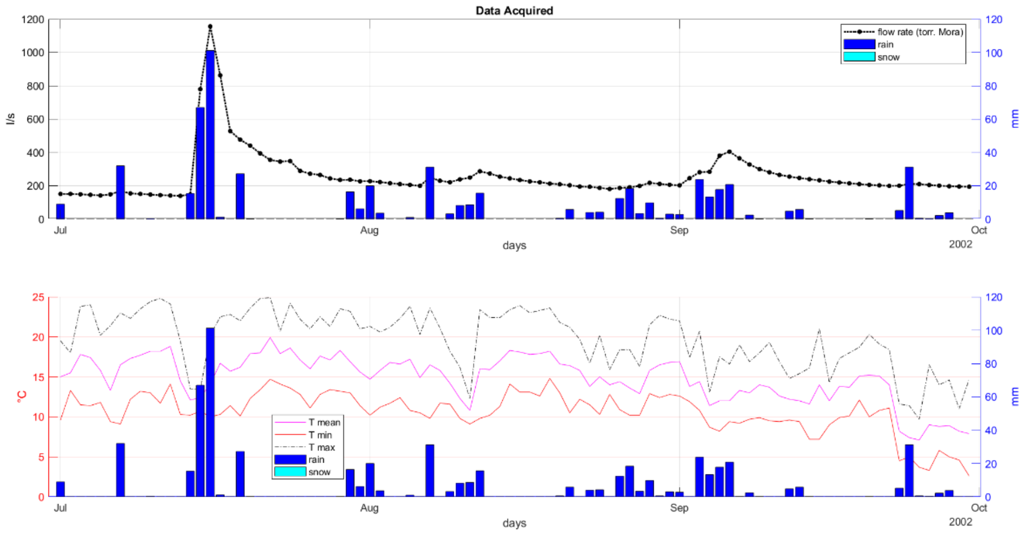
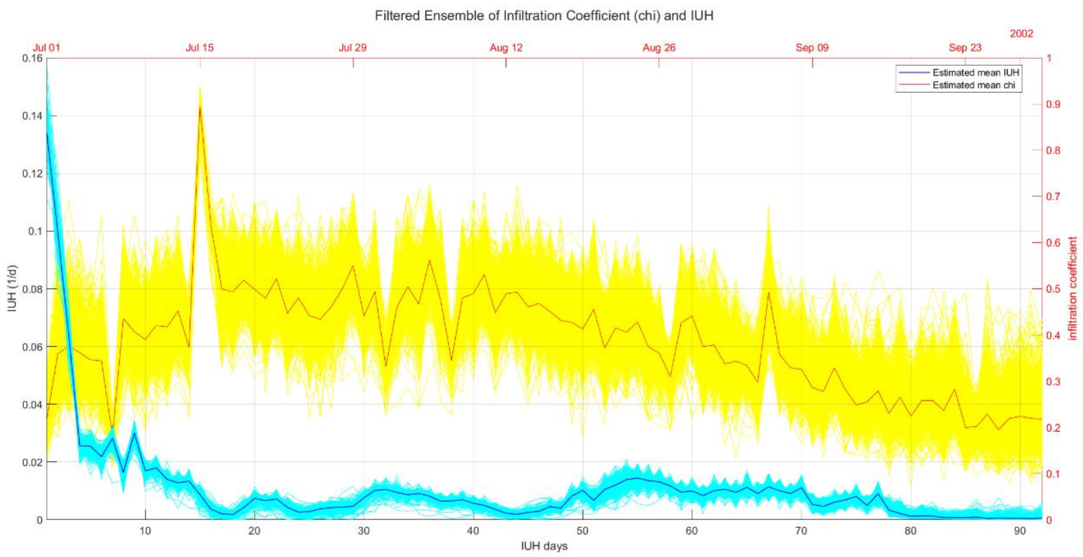
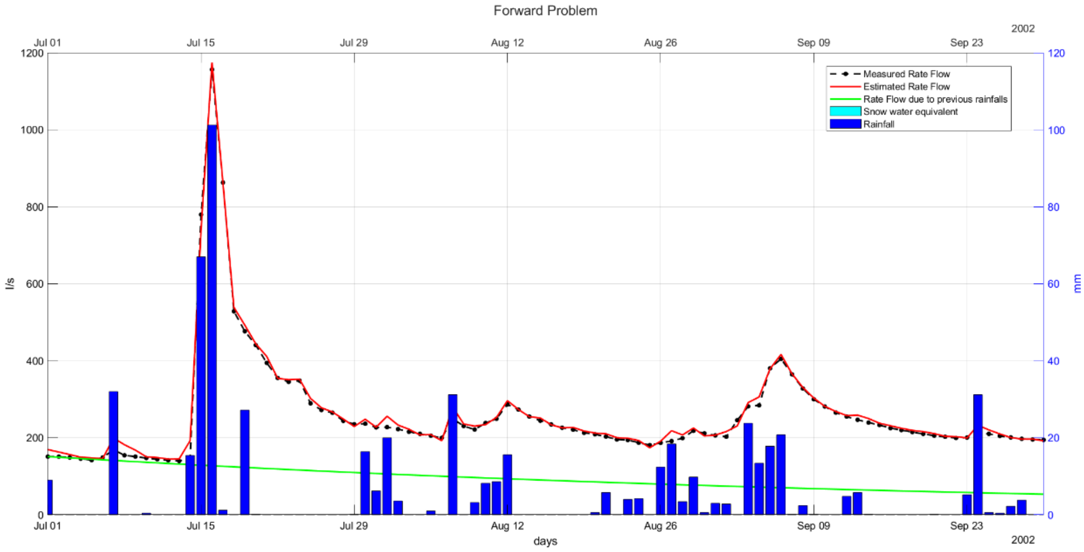


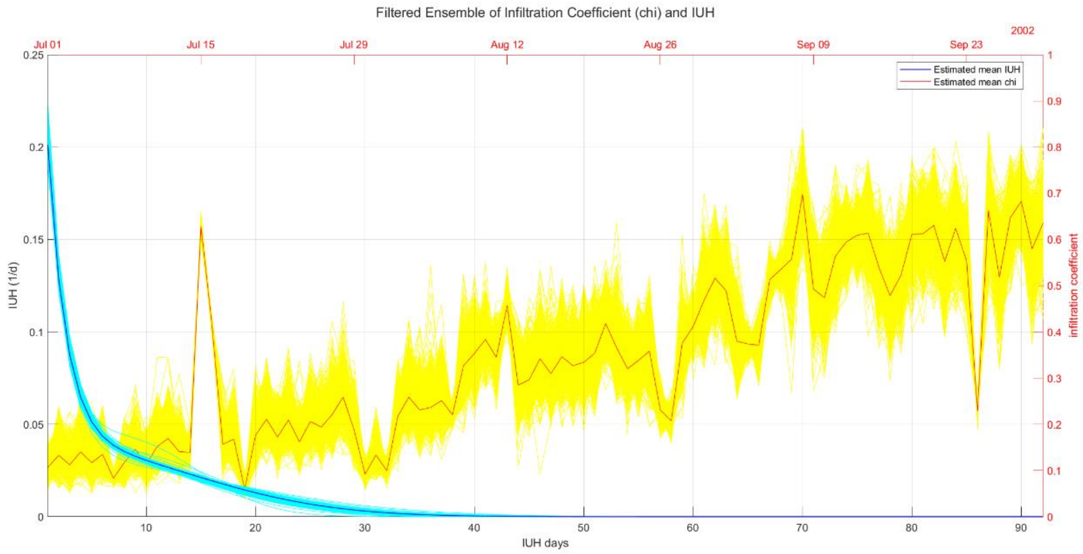
Figure 8-3:data used.

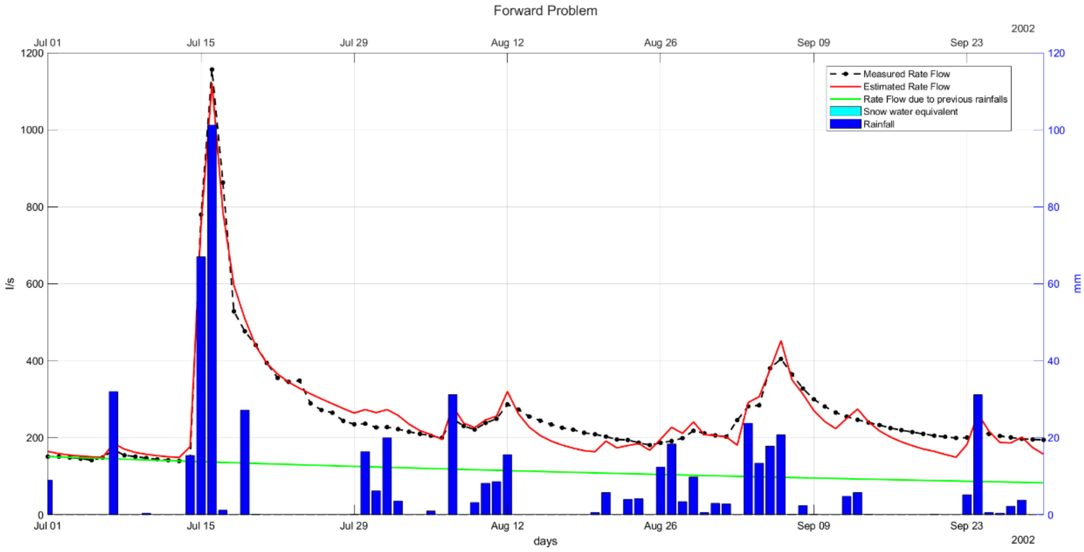
Non-parametric IUH					
Bias [l/s]	4.66	V_{in}/V_{out}	1.18	λ	1.1
Mean Absolute Error [l/s]	7.60	RMSE	13.75	DF	0.4
NSE	0.99	RMSE/ES	2.16	EM	4





Parametric IUH					
Bias [l/s]	-2.06	V_{in}/V_{out}	1.04	λ	1.1
Mean Absolute Error [l/s]	22.58	RMSE	29.61	DF	0.4
NSE	0.96	RMSE/ES	4.45	EM	4





Forward Problem
2002
Jul 01
Jul 15
Jul 29
Aug 12
Aug 26
Sep 09
Sep 23
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days

IV. Spring 2003

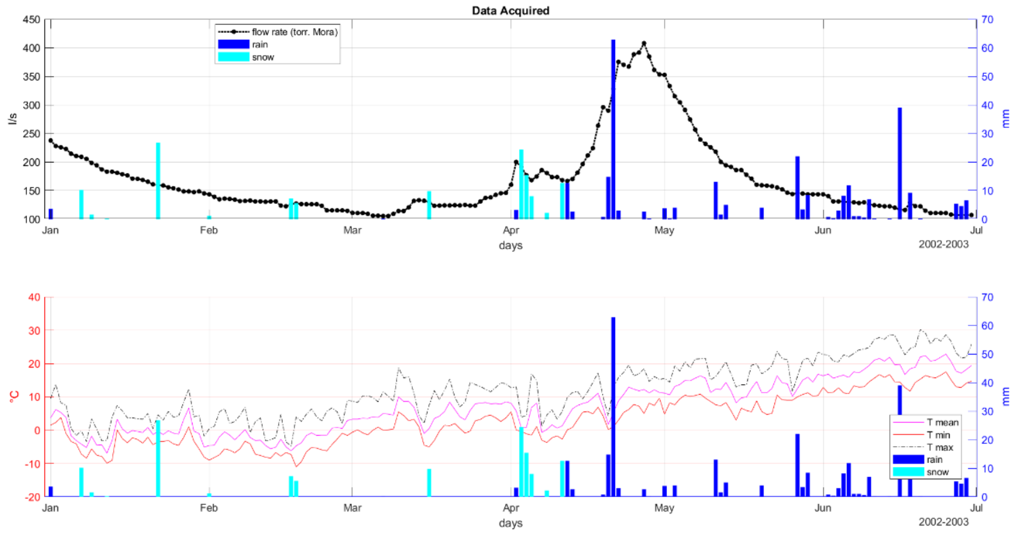
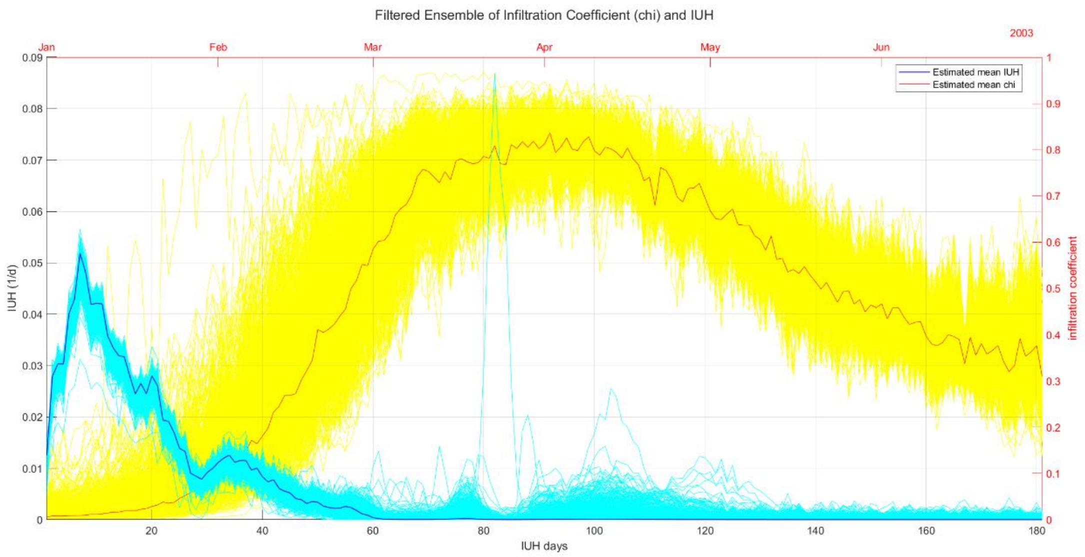
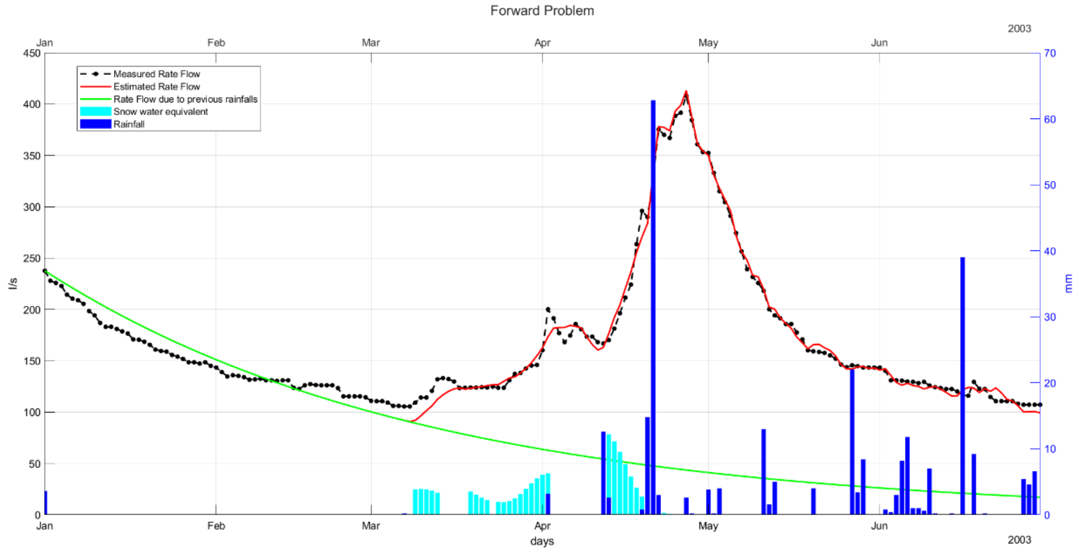


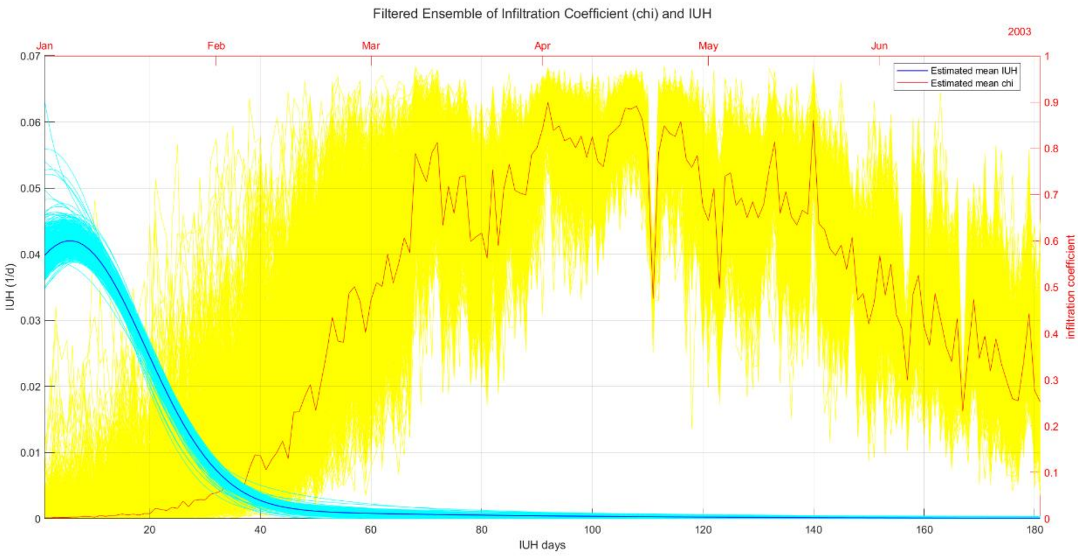
Figure 8-4:data used.

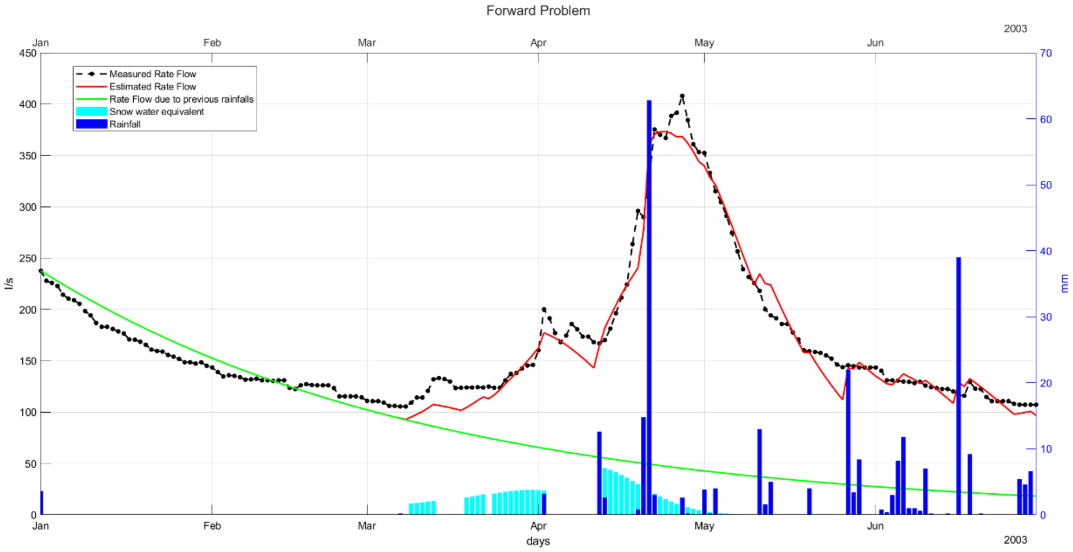
Non-parametric IUH					
Bias [l/s]	-1.07	V_{in}/V_{out}	1.09	λ	1.1
Mean Absolute Error [l/s]	7.96	RMSE	10.42	DF	0.6
NSE	0.98	RMSE/ES	2.12	EM	4





Parametric IUH					
Bias [l/s]	-1.93	V_{in}/V_{out}	1.09	λ	1.15
Mean Absolute Error [l/s]	11.07	RMSE	15.07	DF	0.75
NSE	0.96	RMSE/ES	4.24	EM	4




Forward Problem
2003
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days

V. Fall 2003

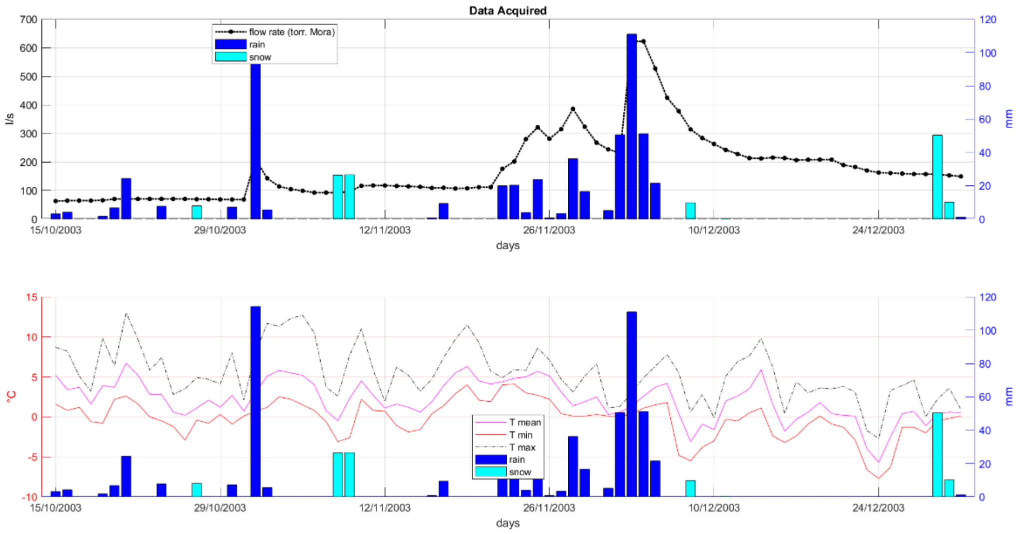
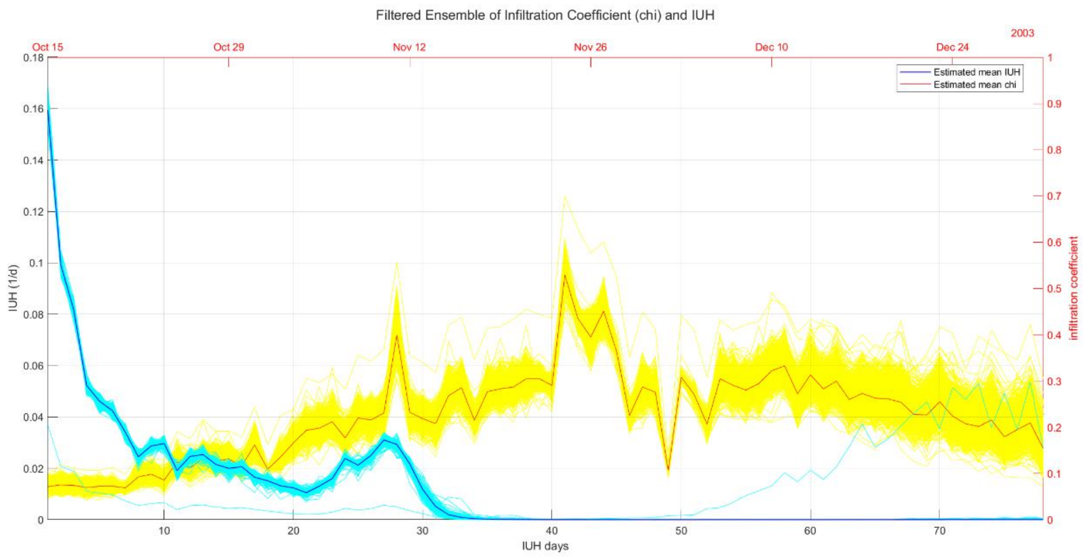
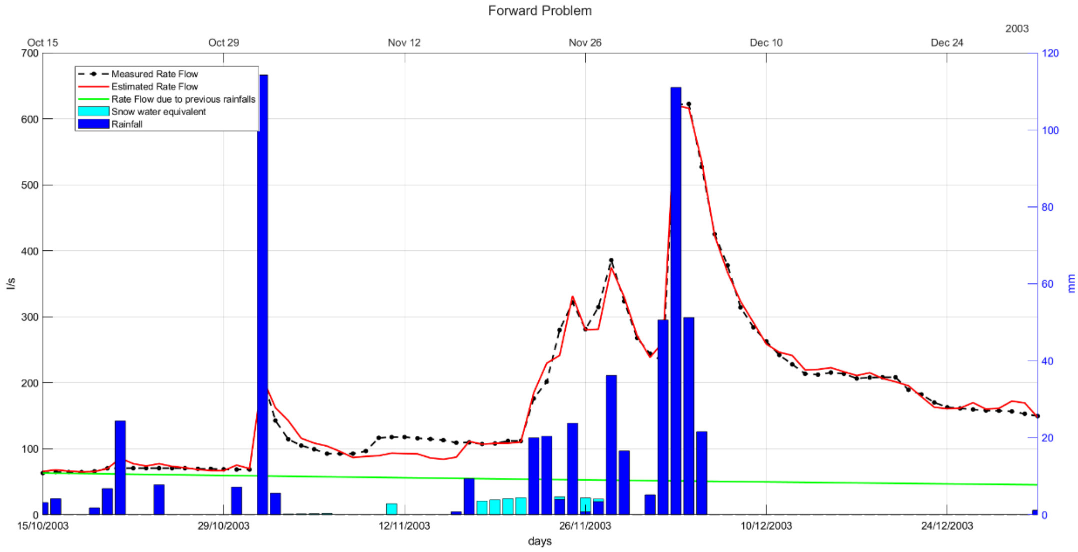


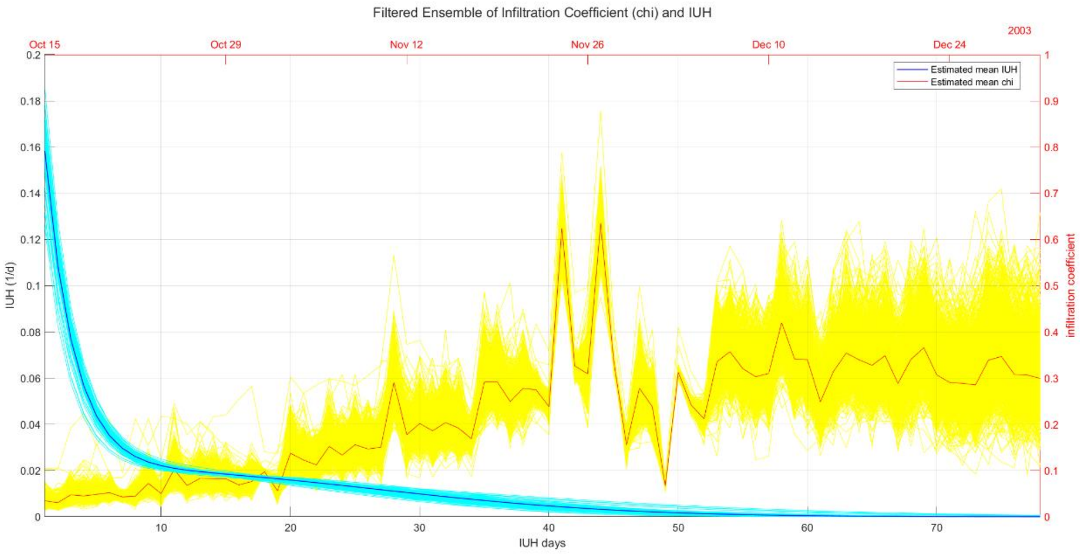
Figure 8-5:data used.

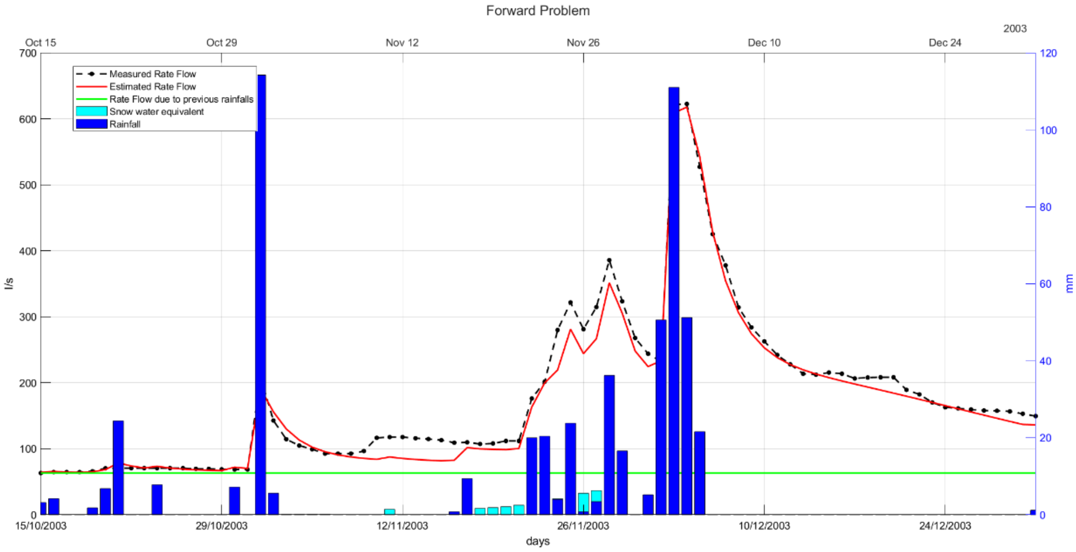
Non-parametric IUH					
Bias [l/s]	0.01	V_{in}/V_{out}	1.01	λ	1.05
Mean Absolute Error [l/s]	8.98	RMSE	14.45	DF	0.7
NSE	0.99	RMSE/ES	3.03	EM	3





Parametric IUH					
Bias [l/s]	-9.45	V_{in}/V_{out}	1.09	λ	1.1
Mean Absolute Error [l/s]	11.80	RMSE	16.8	DF	0.7
NSE	0.98	RMSE/ES	2.89	EM	3





Forward Problem
2003
Oct 15
Oct 29
Nov 12
Nov 26
Dec 10
Dec 24
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
15/10/2003
29/10/2003
12/11/2003
26/11/2003
10/12/2003
24/12/2003

VI. Spring 2004

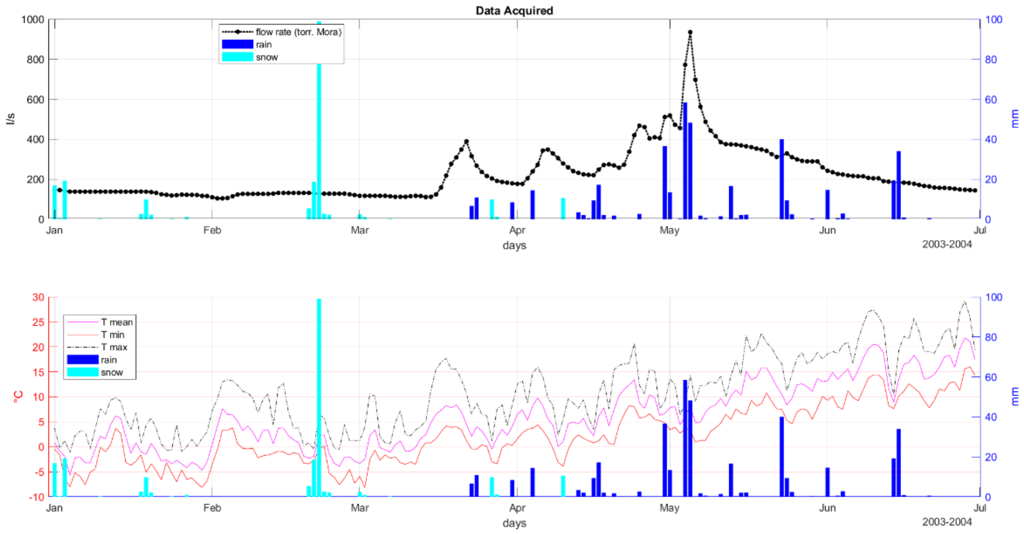
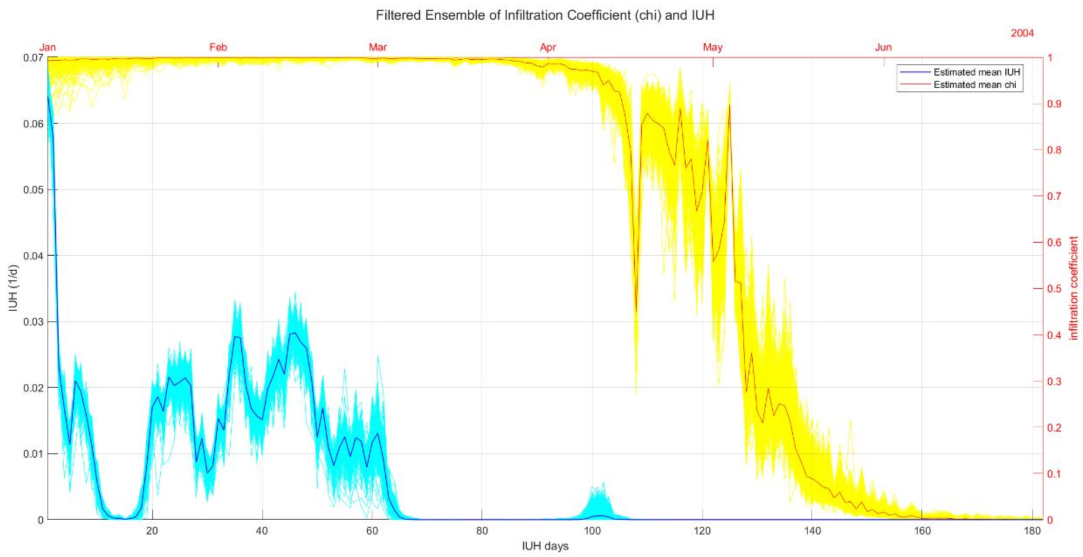
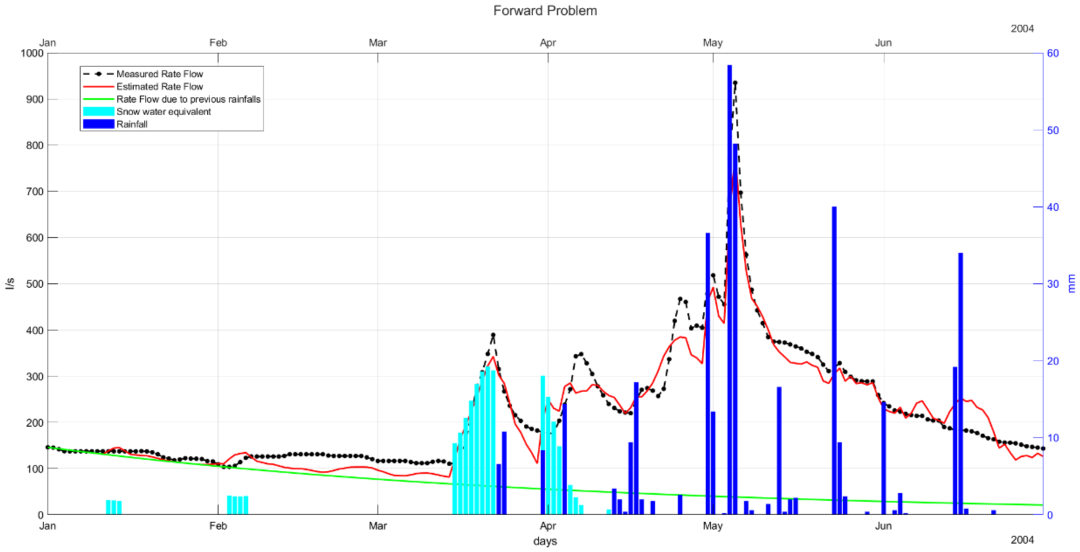


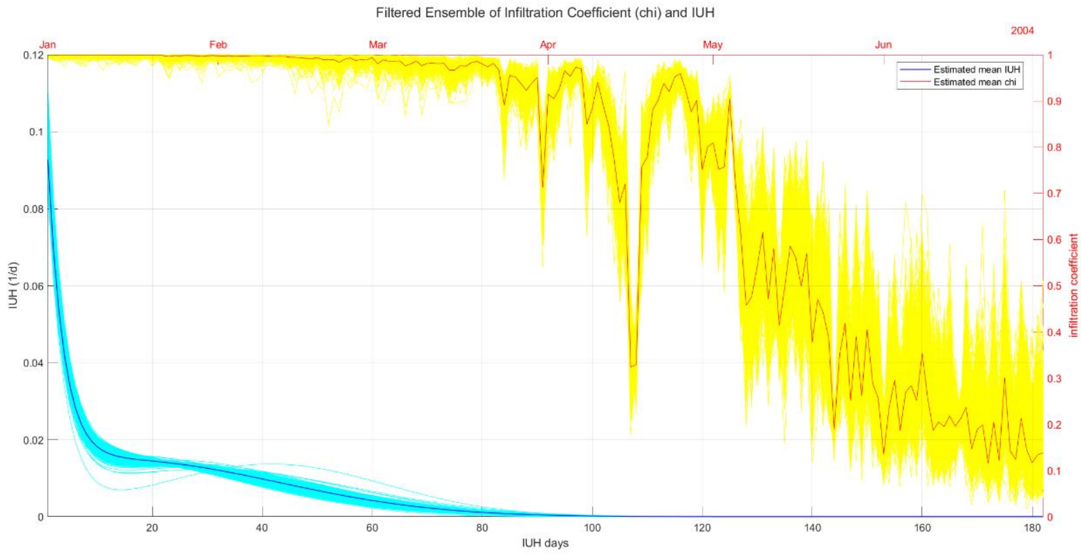
Figure 8-6:data used.

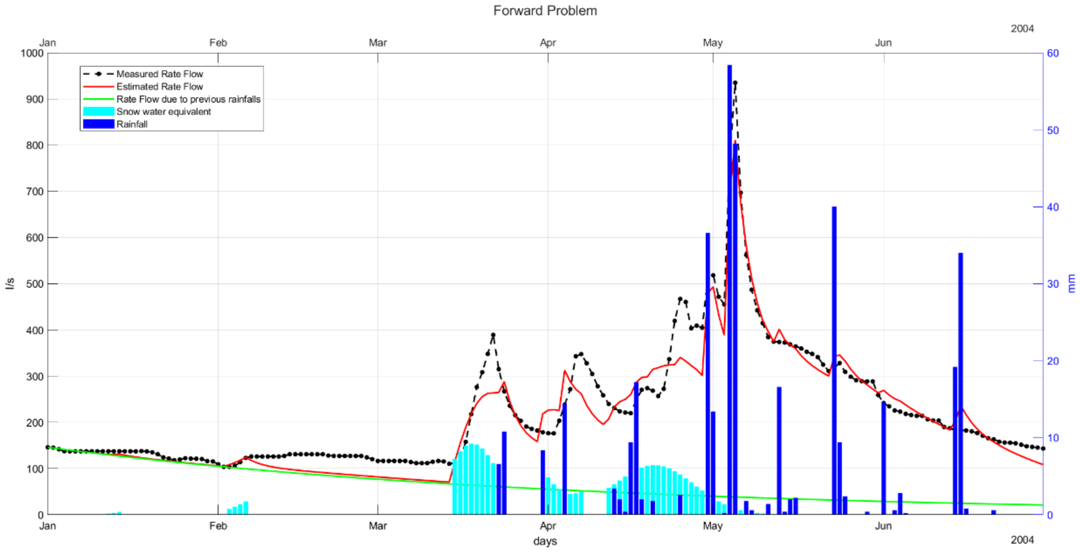
Non-parametric IUH					
Bias [l/s]	-10.78	V_{in}/V_{out}	1.02	λ	1.12
Mean Absolute Error [l/s]	25.02	RMSE	36.37	DF	0.5
NSE	0.93	RMSE/ES	7.81	EM	4





Parametric IUH					
Bias [l/s]	-14.06	V_{in}/V_{out}	1.06	λ	1.13
Mean Absolute Error [l/s]	27.69	RMSE	37.88	DF	0.35
NSE	0.91	RMSE/ES	7.95	EM	3





Forward Problem
2004
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days

VII. Spring 2005

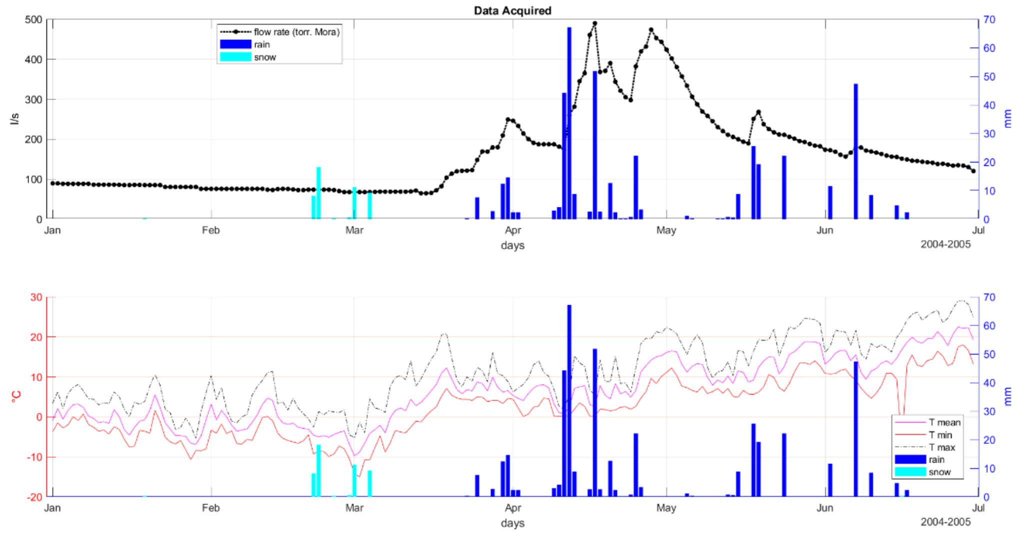
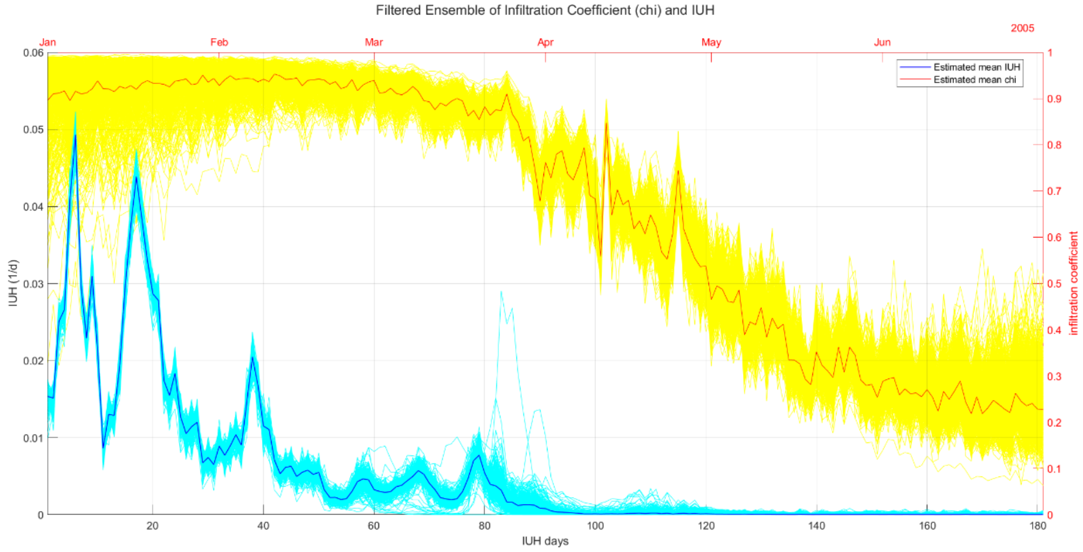
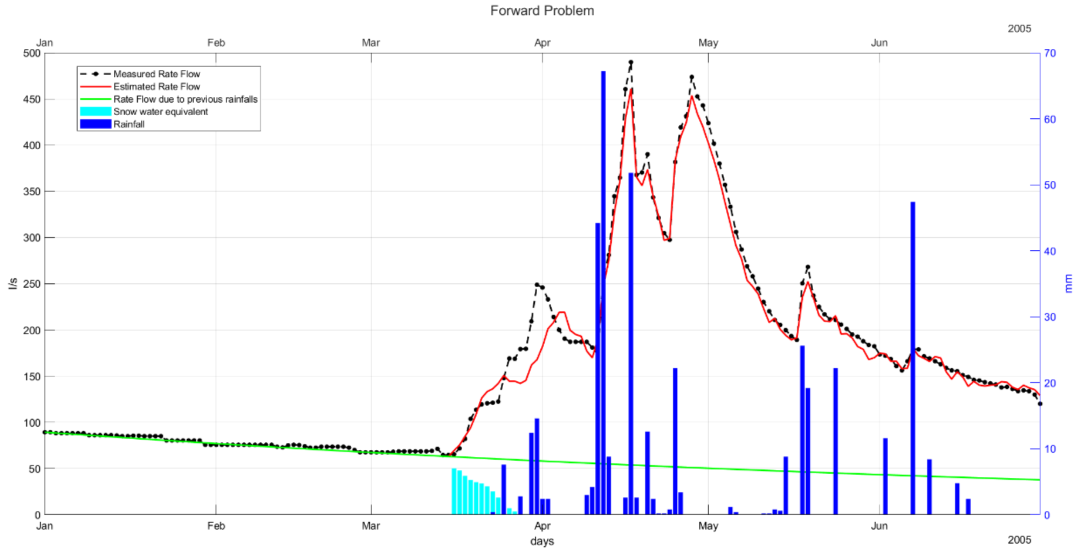


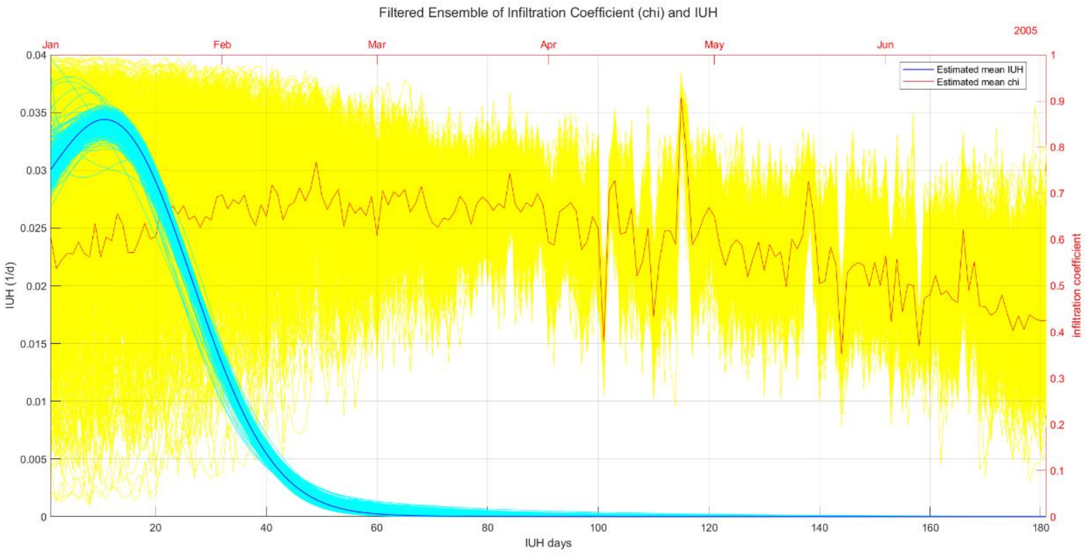
Figure 8-7:data used.

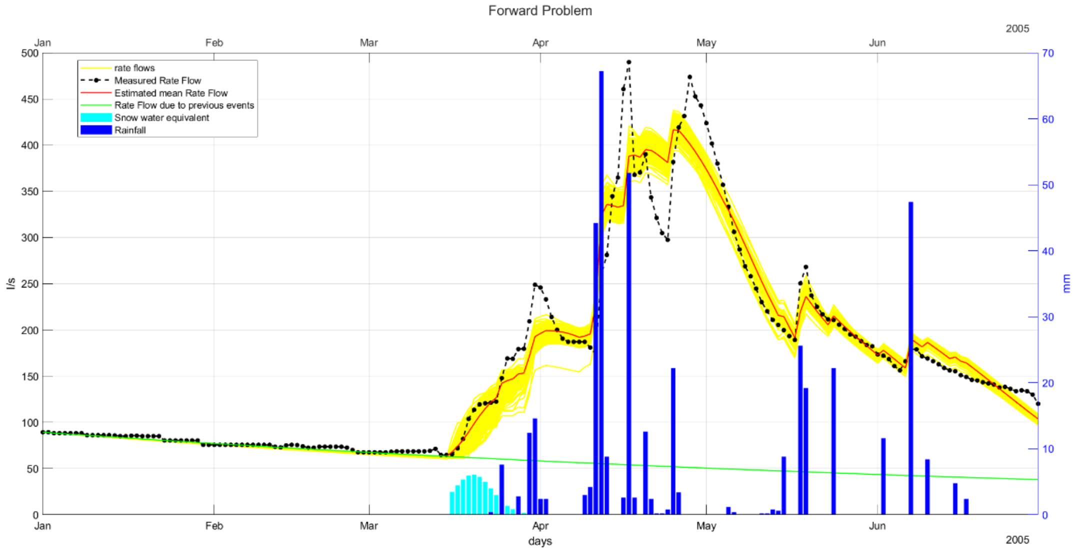
Non-parametric IUH					
Bias [l/s]	-4.94	V_{in}/V_{out}	1.08	λ	1.08
Mean Absolute Error [l/s]	7.47	RMSE	12.96	DF	0.5
NSE	0.98	RMSE/ES	3.36	EM	2





Parametric IUH					
$Bias$ [l/s]	-2.06	V_{in}/V_{out}	1.04	λ	1.12
Mean Absolute Error [l/s]	12.88	RMSE	23.88	DF	0.4
NSE	0.95	RMSE/ES	7.49	EM	4




Forward Problem
2005
Jan
Feb
Mar
Apr
May
Jun
rate flows
Measured Rate Flow
Estimated mean Rate Flow
Rate Flow due to previous events
Snow water equivalent
Rainfall
l/s
mm
days
2005

VIII. Fall 2005

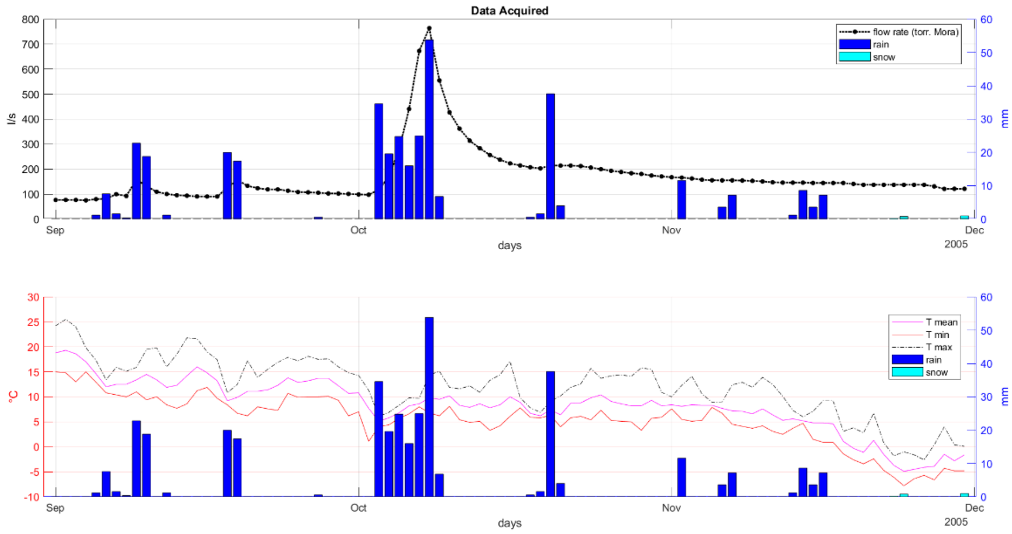
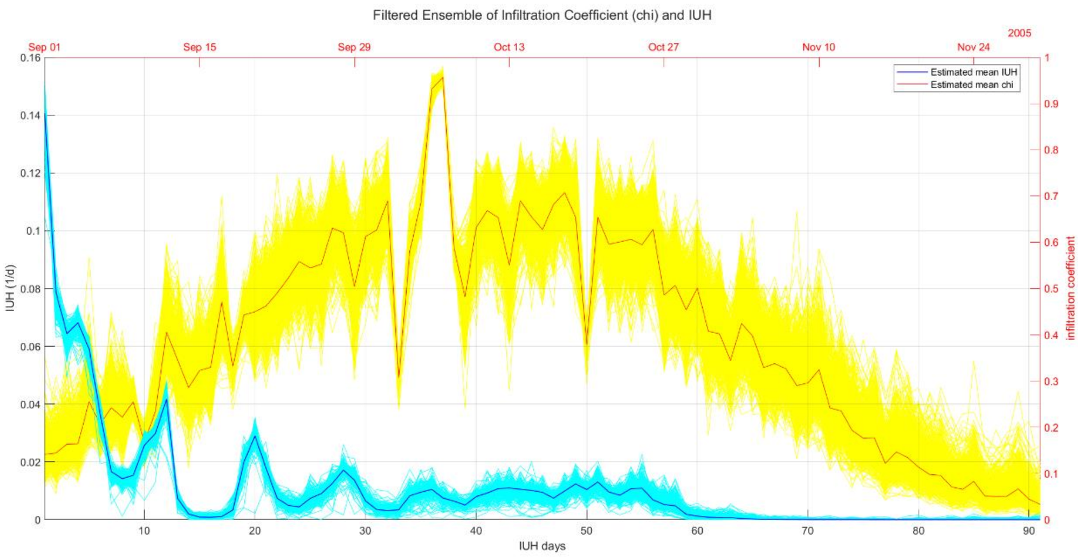
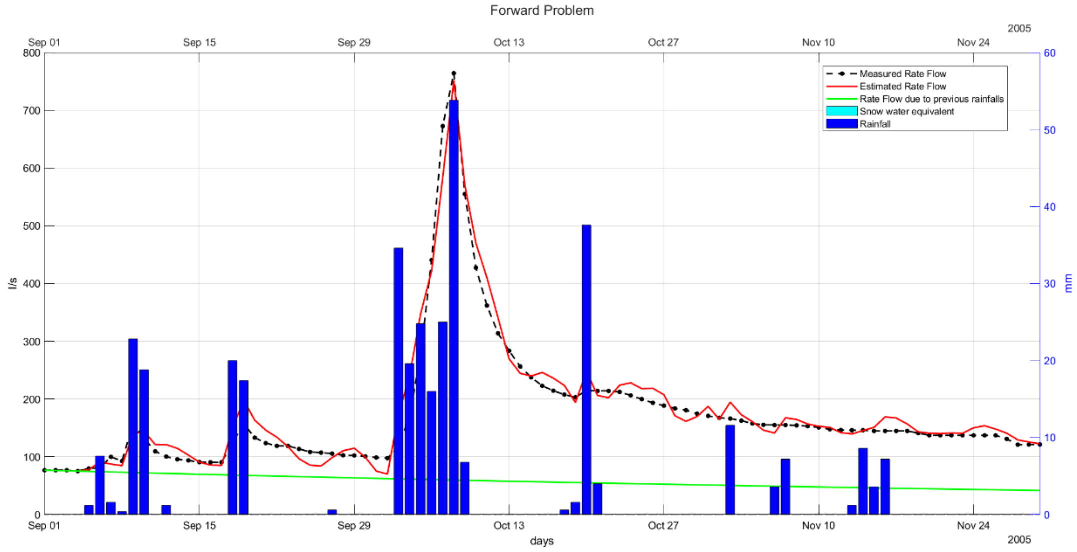


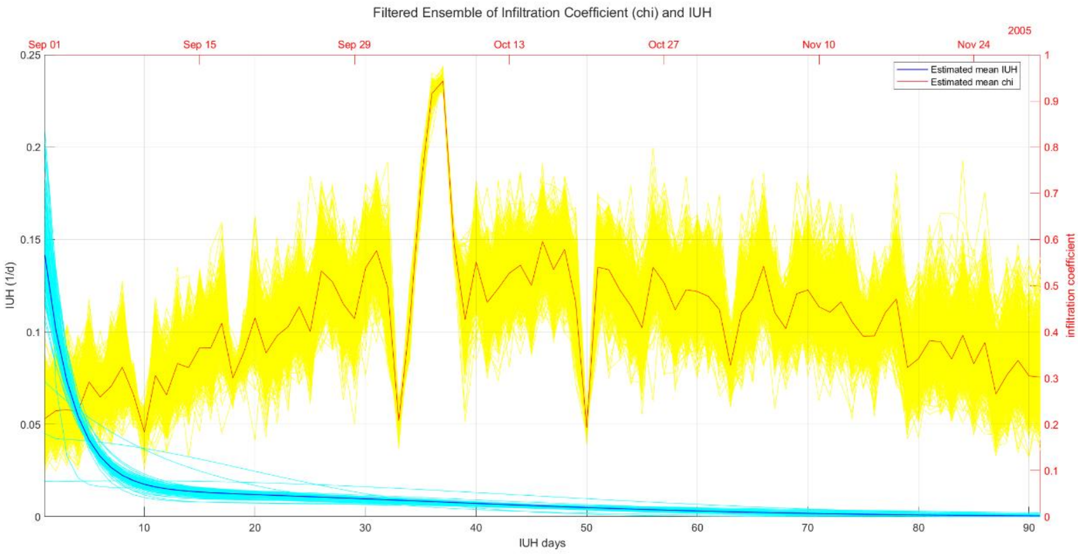
Figure 8-8:data used.

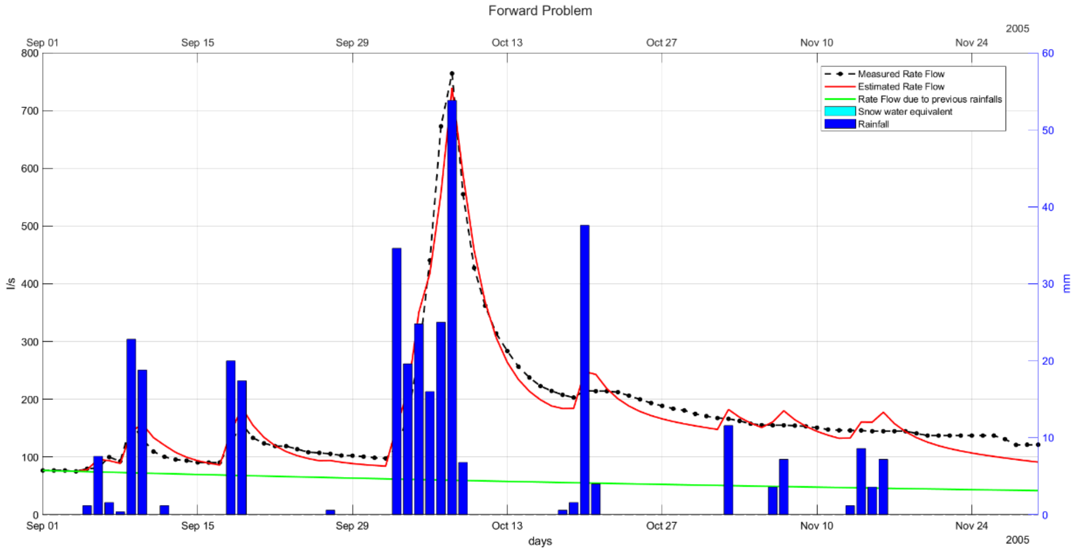
Non-parametric IUH					
Bias [l/s]	2.54	$^{V_{in}}/_{V_{out}}$	1.05	λ	1.1
Mean Absolute Error [l/s]	14.81	RMSE	23.17	DF	0.4
NSE	0.96	RMSE/ES	4.31	EM	3





Parametric IUH					
$Bias$ [l/s]	-4.23	V_{in}/V_{out}	1.09	λ	1.1
Mean Absolute Error [l/s]	18.15	RMSE	24.71	DF	0.3
NSE	0.96	RMSE/ES	3.1	EM	3




Forward Problem
2005
Sep 01
Sep 15
Sep 29
Oct 13
Oct 27
Nov 10
Nov 24
800
700
600
500
400
300
200
100
0
l/s
60
50
40
30
20
10
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days
2005

IX. Spring 2006

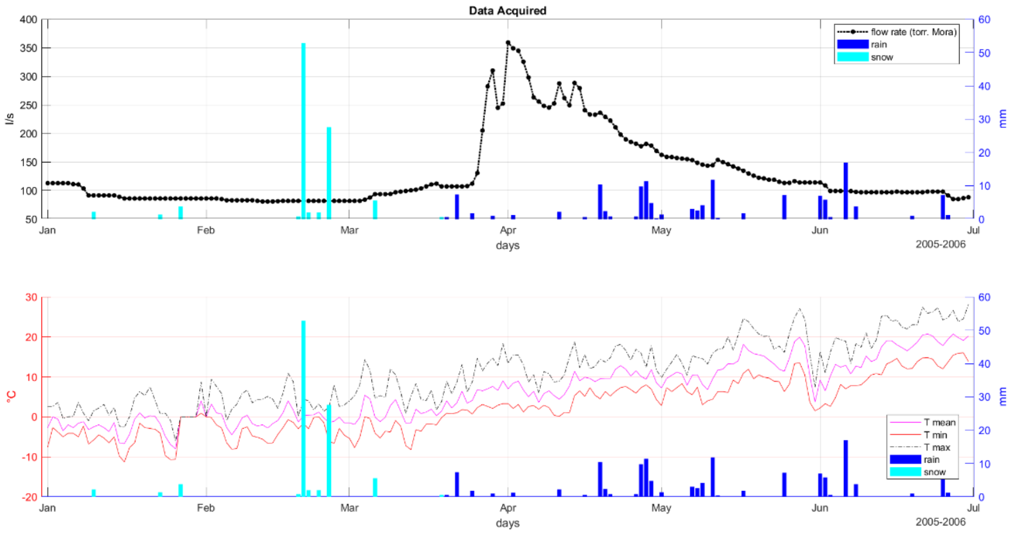
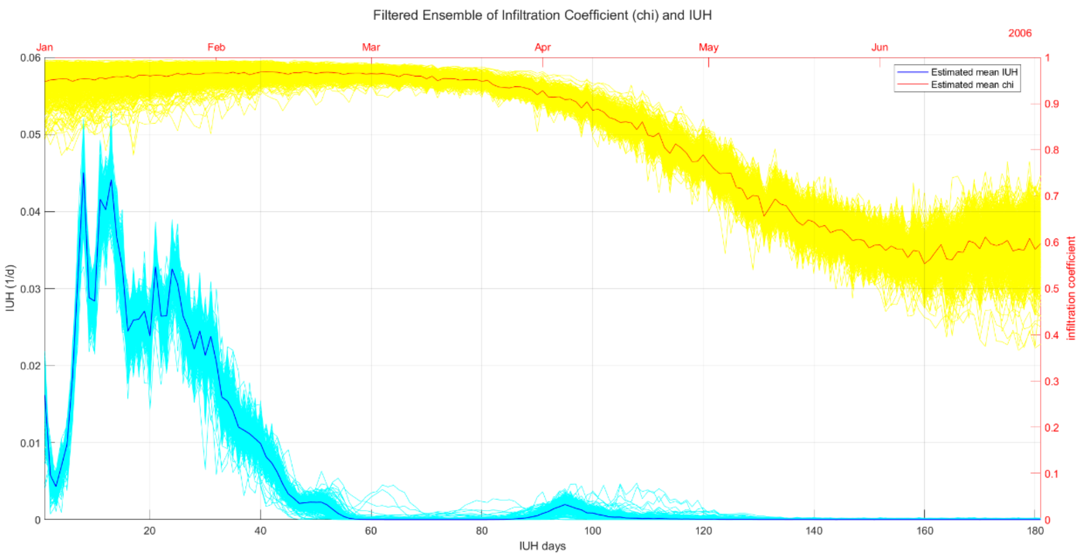
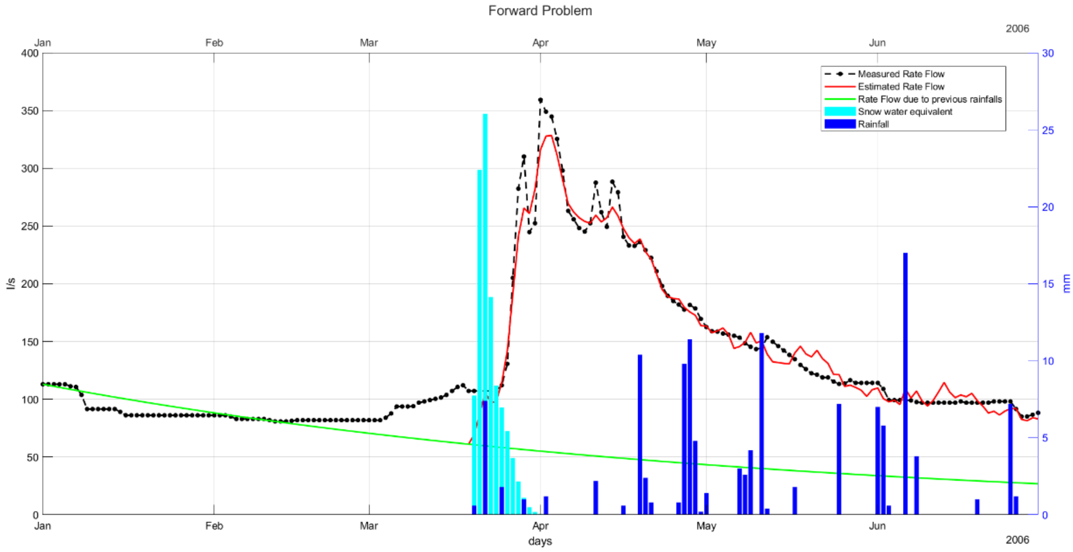


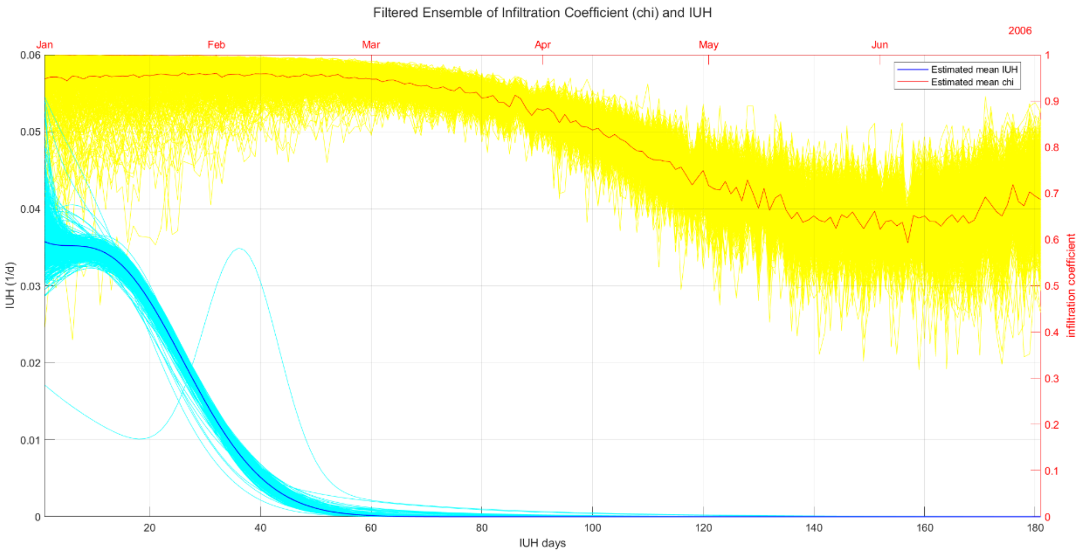
Figure 8-9:data used.

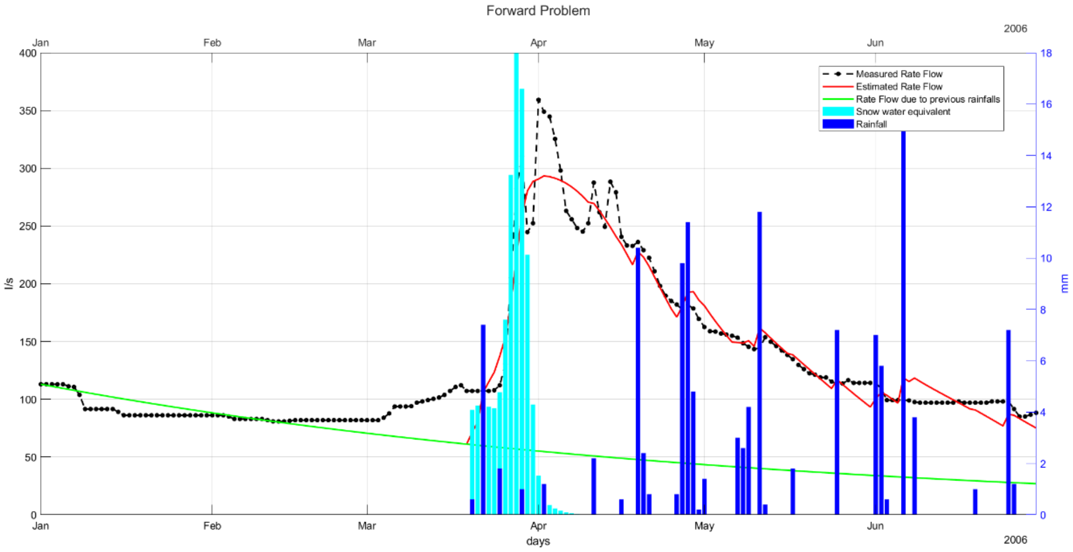
Non-parametric IUH					
Bias [l/s]	-4.17	V_{in}/V_{out}	1.08	λ	1.05
Mean Absolute Error [l/s]	10.42	RMSE	15.66	DF	0.4
NSE	0.94	RMSE/ES	4.50	EM	3





Parametric IUH					
Bias [l/s]	-4.09	V_{in}/V_{out}	1.05	λ	1.1
Mean Absolute Error [l/s]	12.93	RMSE	18.93	DF	0.35
NSE	0.92	RMSE/ES	5.19	EM	4




Forward Problem
2006
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
0
50
100
150
200
250
300
350
400
2
4
6
8
10
12
14
16
18

X. Summer 2007

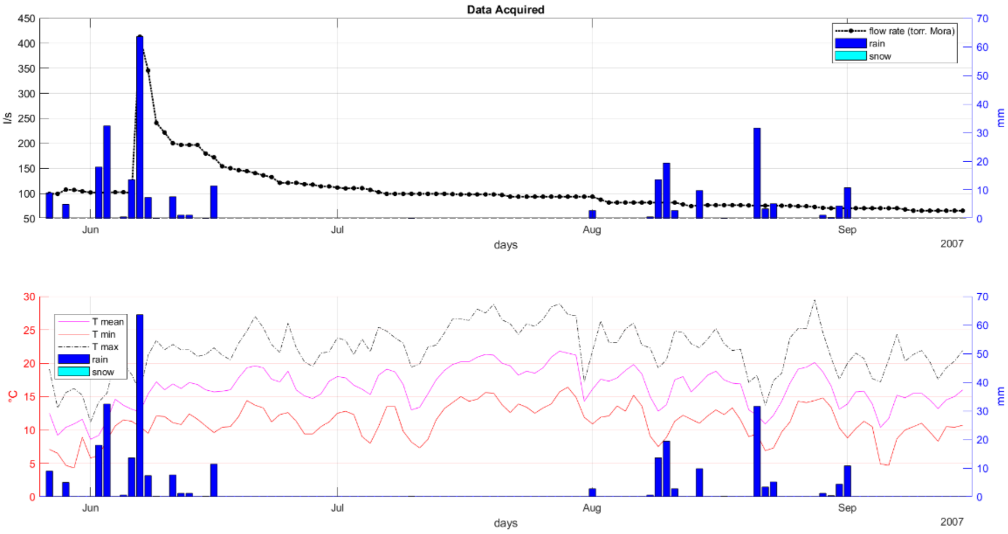
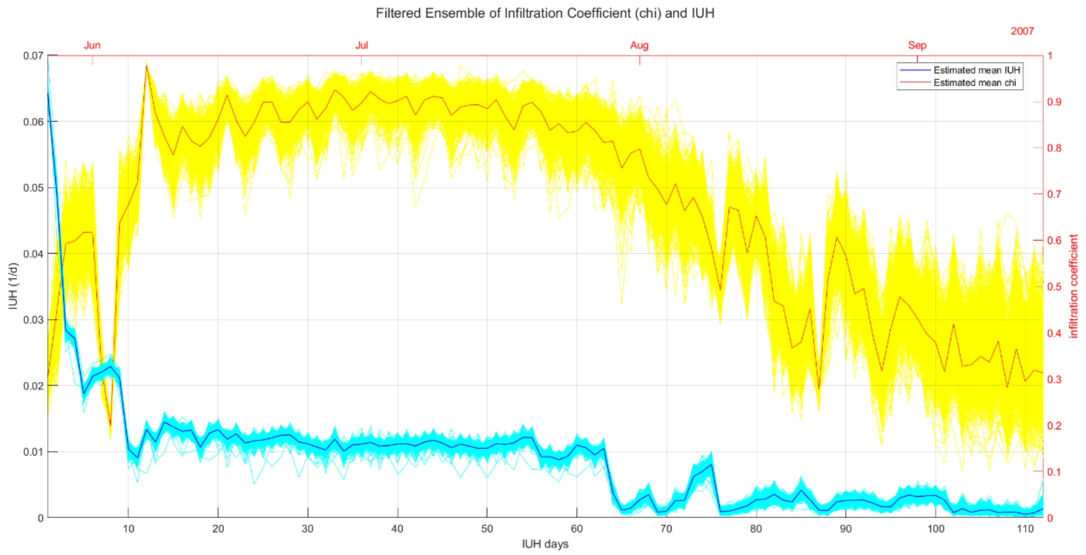
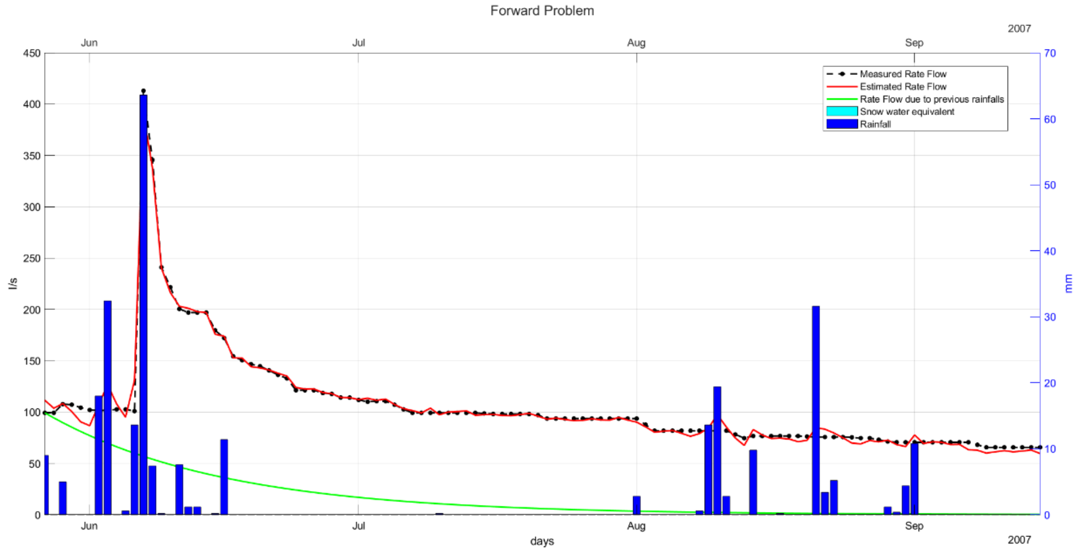


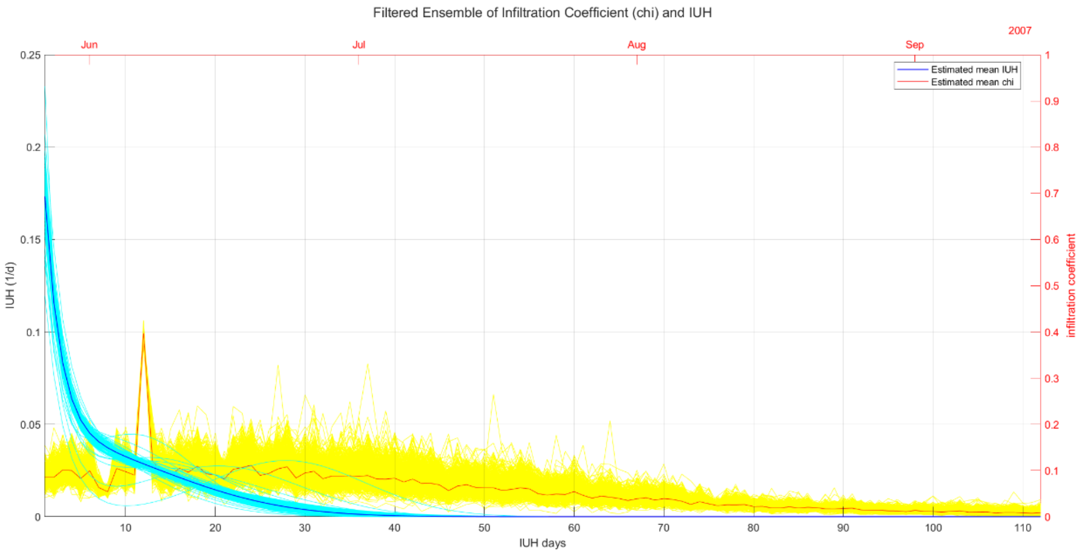
Figure 8-10:data used.

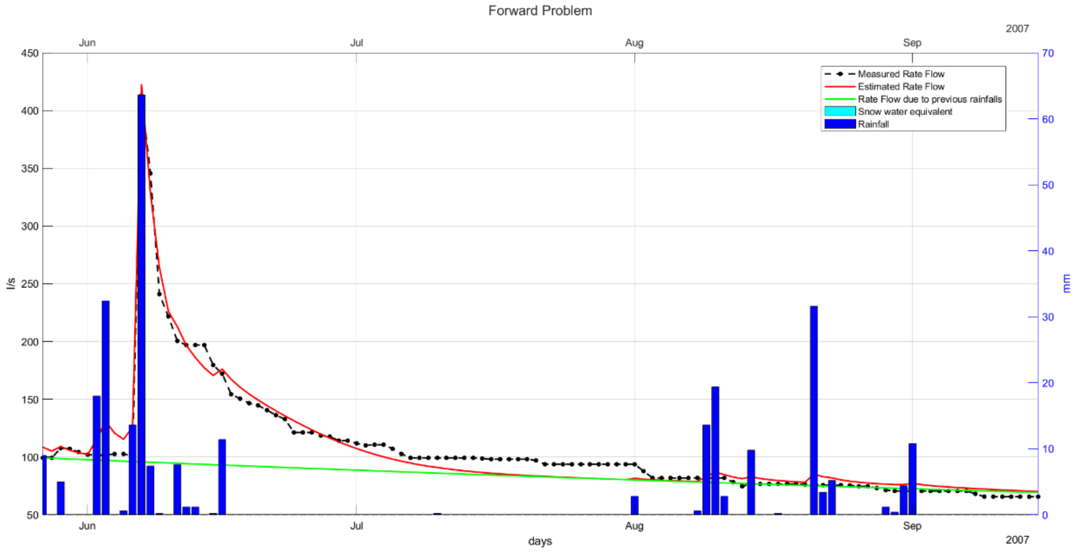
Non-parametric IUH					
Bias [l/s]	-0.24	V_{in}/V_{out}	1.17	λ	1.1
Mean Absolute Error [l/s]	3.65	RMSE	7.09	DF	0.7
NSE	0.99	RMSE/ES	2.94	EM	2





Parametric IUH					
Bias [l/s]	0.04	V_{in}/V_{out}	1.00	λ	1.1
Mean Absolute Error [l/s]	7.50	RMSE	9.22	DF	0.5
NSE	0.97	RMSE/ES	3.29	EM	4





Forward Problem
2007
Jun
Jul
Aug
Sep
450
400
350
300
250
200
150
100
50
l/s
70
60
50
40
30
20
10
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days

XI. Fall 2007

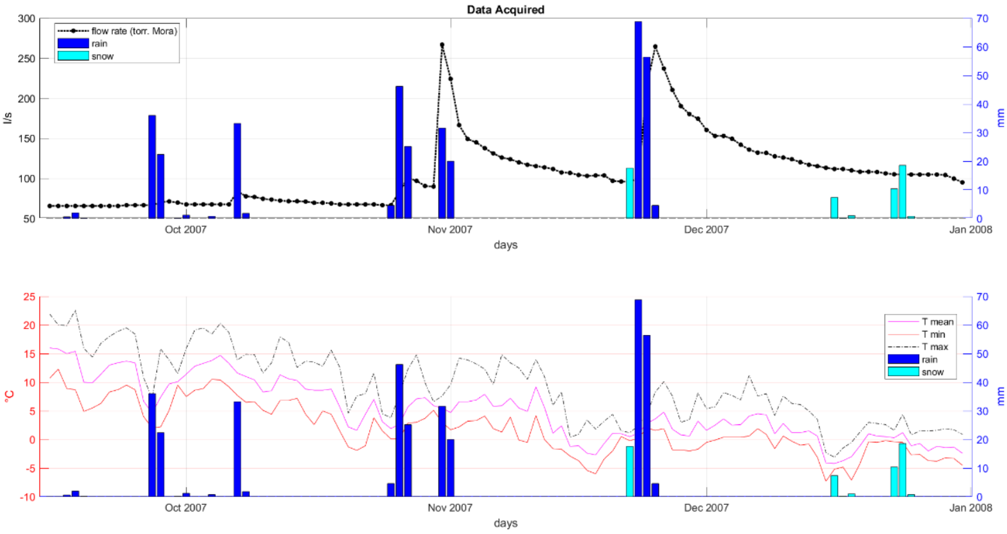
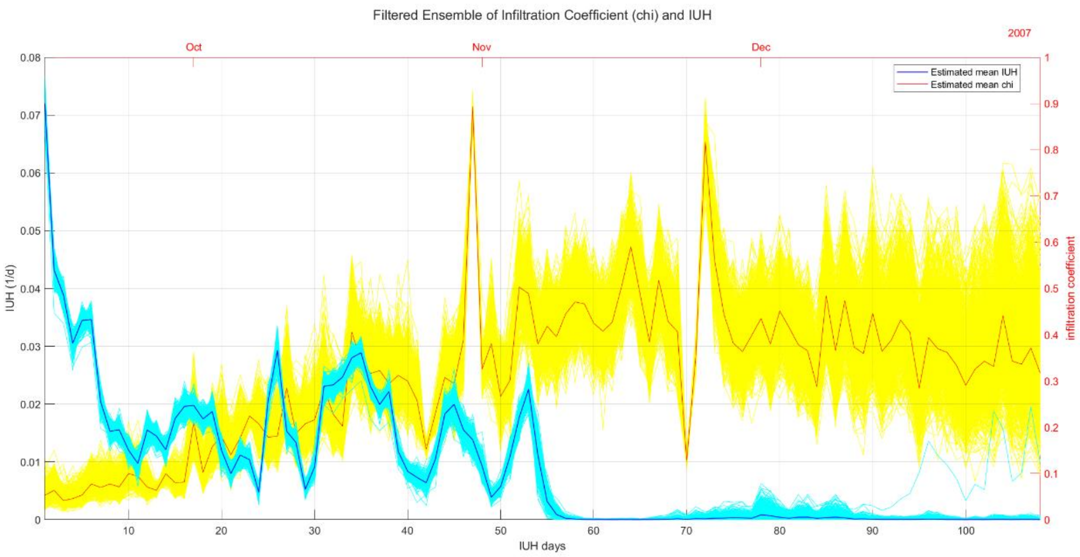
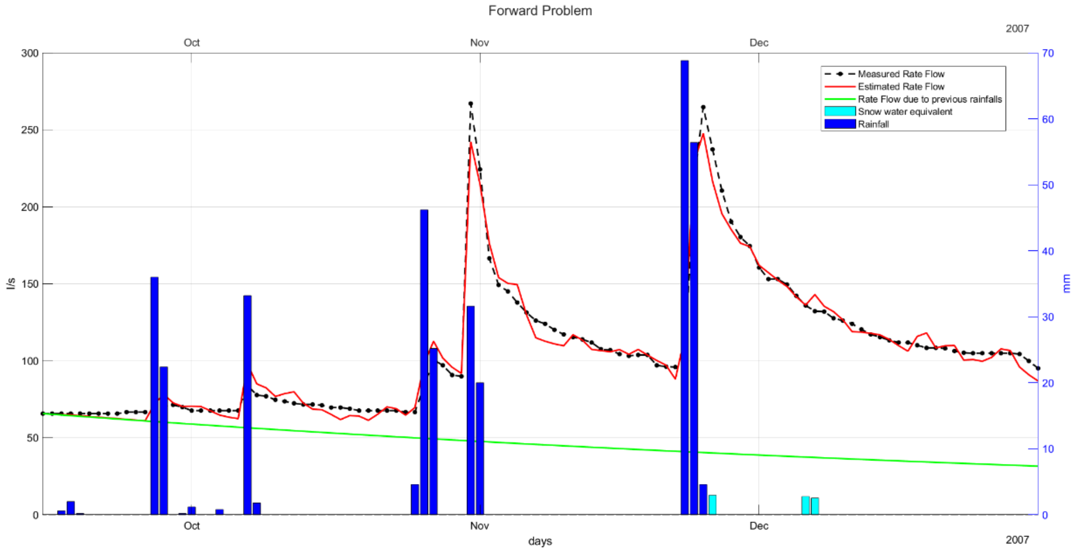


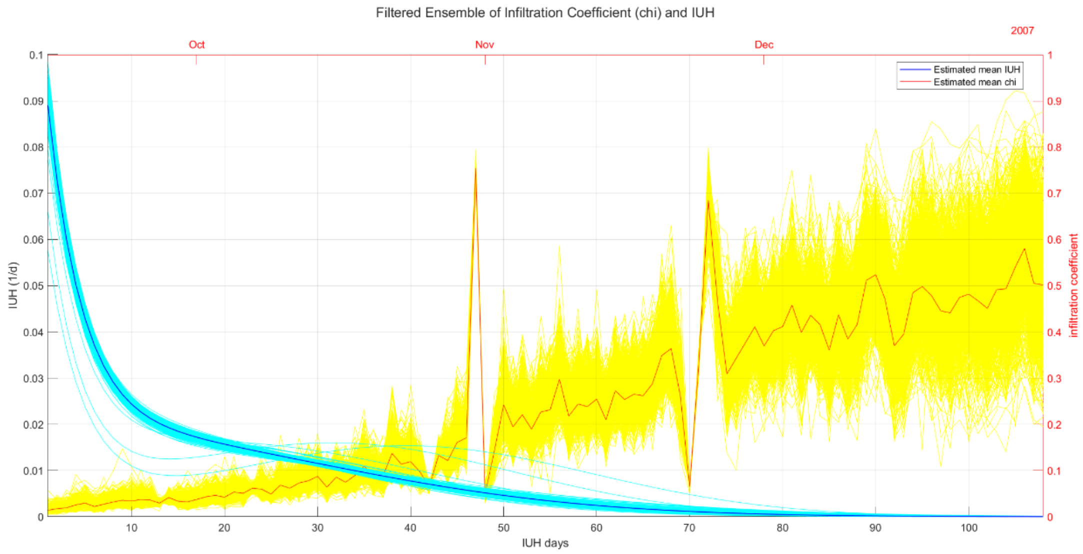
Figure 8-11:data used.

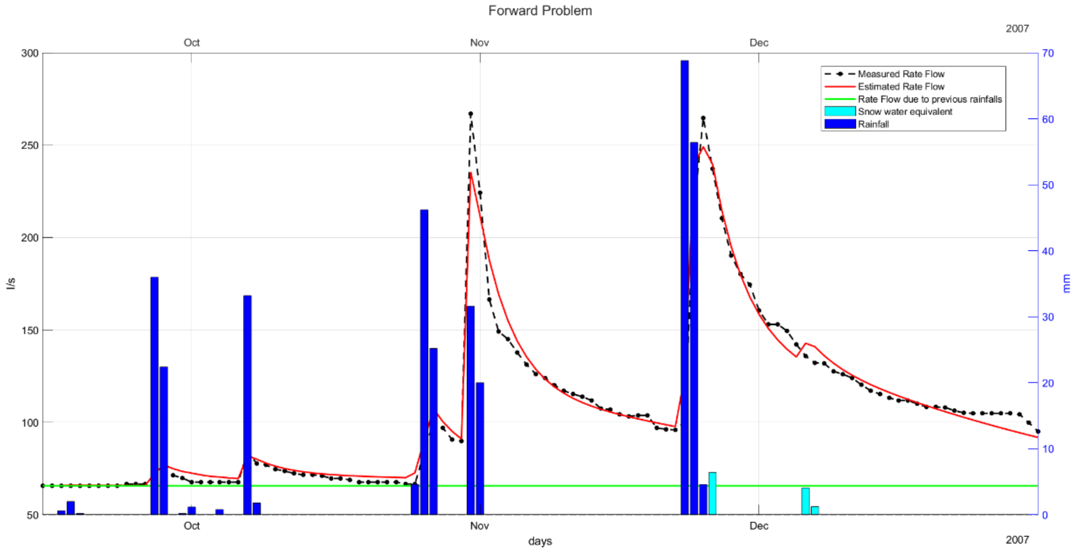
Non-parametric IUH					
Bias [l/s]	-0.53	V_{in}/V_{out}	1.10	λ	1.1
Mean Absolute Error [l/s]	4.83	RMSE	7.55	DF	1.00
NSE	0.98	RMSE/ES	3.43	EM	2





Parametric IUH					
Bias [l/s]	0.63	V_{in}/V_{out}	1.09	λ	1.1
Mean Absolute Error [l/s]	4.19	RMSE	7.36	DF	0.6
NSE	0.98	RMSE/ES	3.13	EM	2





Forward Problem
2007
Oct
Nov
Dec
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
300
250
200
150
100
50
70
60
50
40
30
20
10
0

XII. Spring 2008

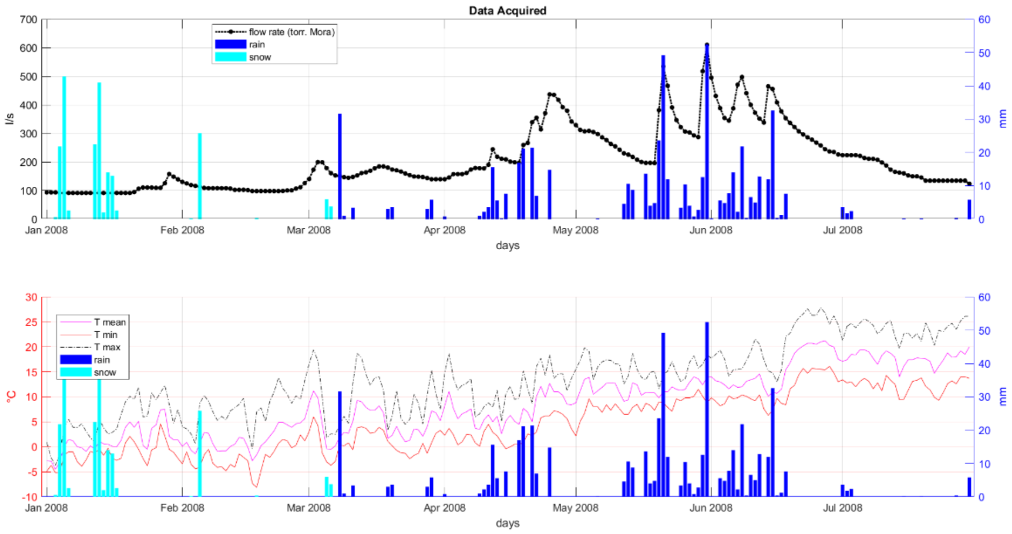
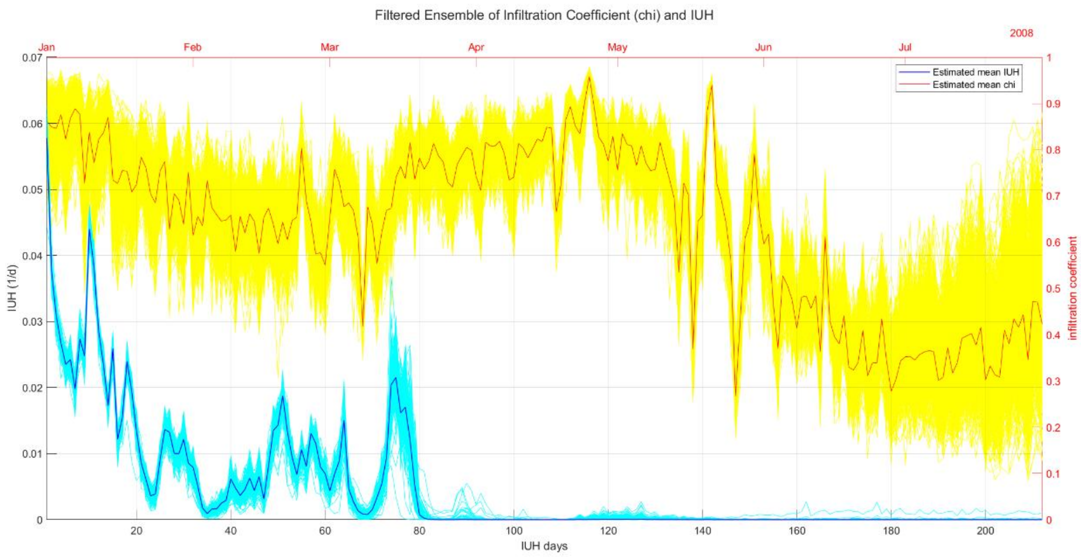
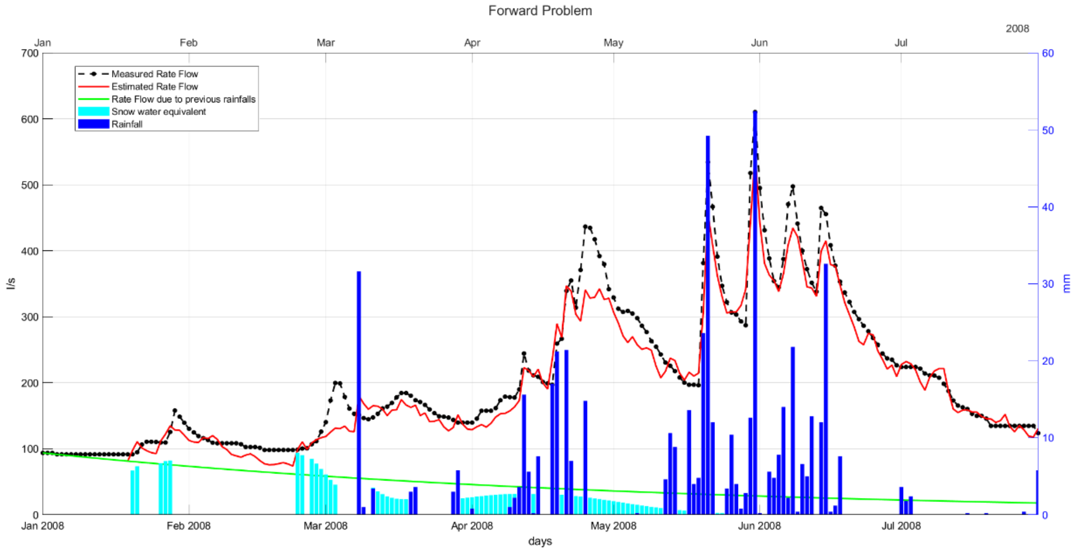


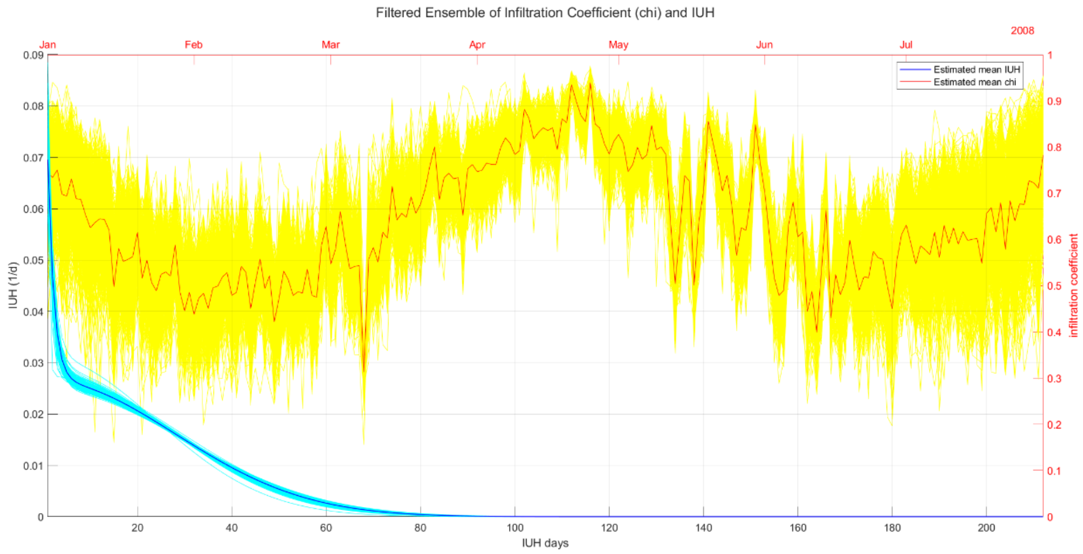
Figure 8-12:data used.

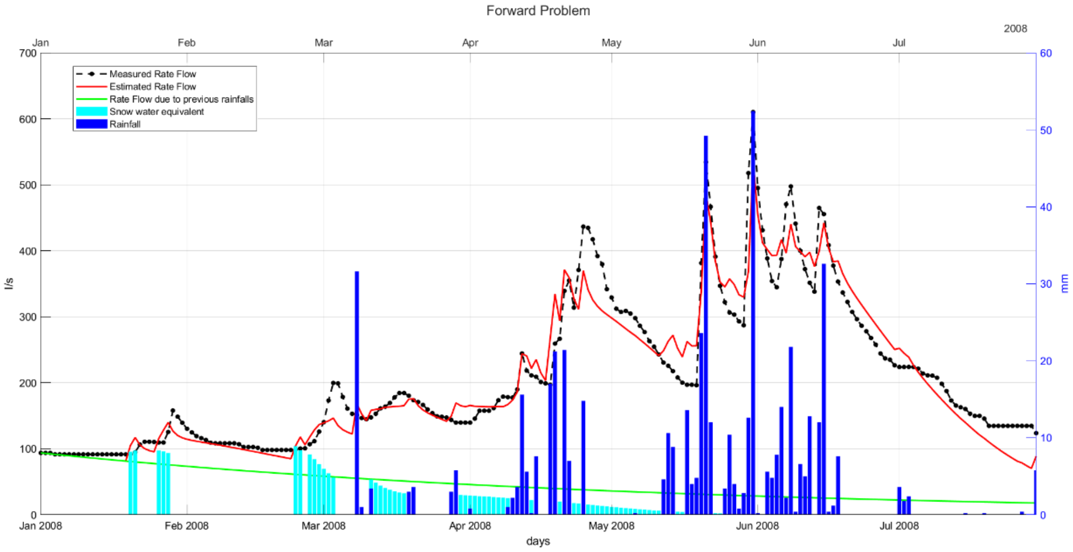
Non-parametric IUH					
Bias [l/s]	-13.25	V_{in}/V_{out}	1.08	λ	1.1
Mean Absolute Error [l/s]	18.72	RMSE	23.32	DF	0.6
NSE	0.94	RMSE/ES	3.6	EM	4





Parametric IUH					
Bias [l/s]	-4.44	V_{in}/V_{out}	1.03	λ	1.1
Mean Absolute Error [l/s]	22.82	RMSE	32.15	DF	0.3
NSE	0.92	RMSE/ES	7.81	EM	4




Forward Problem
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
Jan
Feb
Mar
Apr
May
Jun
Jul
2008
Jan 2008
Feb 2008
Mar 2008
Apr 2008
May 2008
Jun 2008
Jul 2008

XIII. Fall 2008

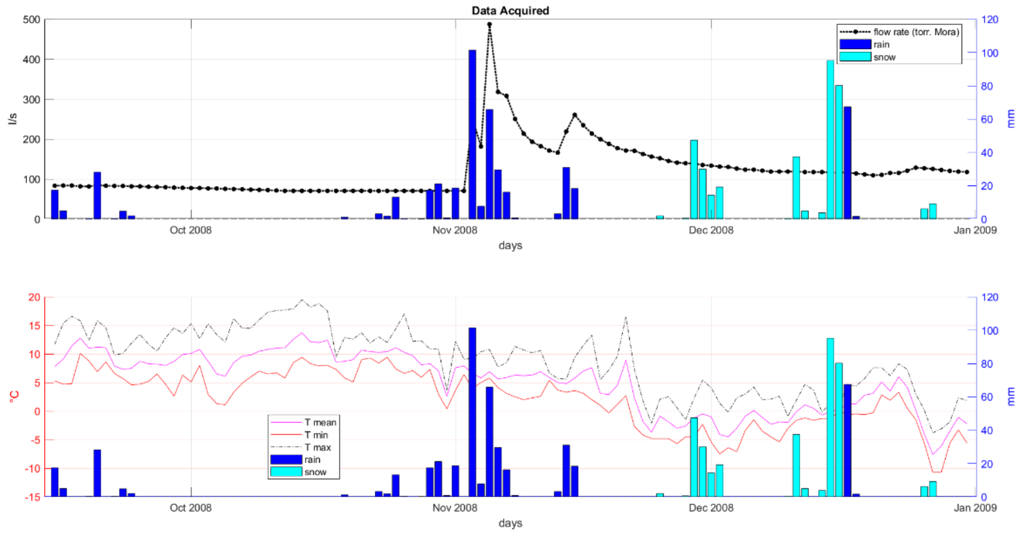
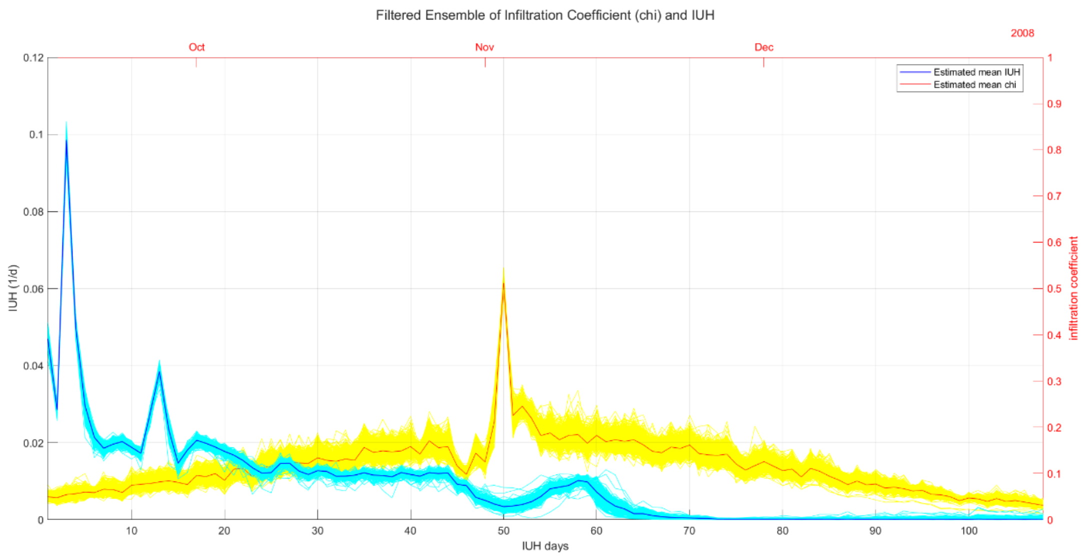
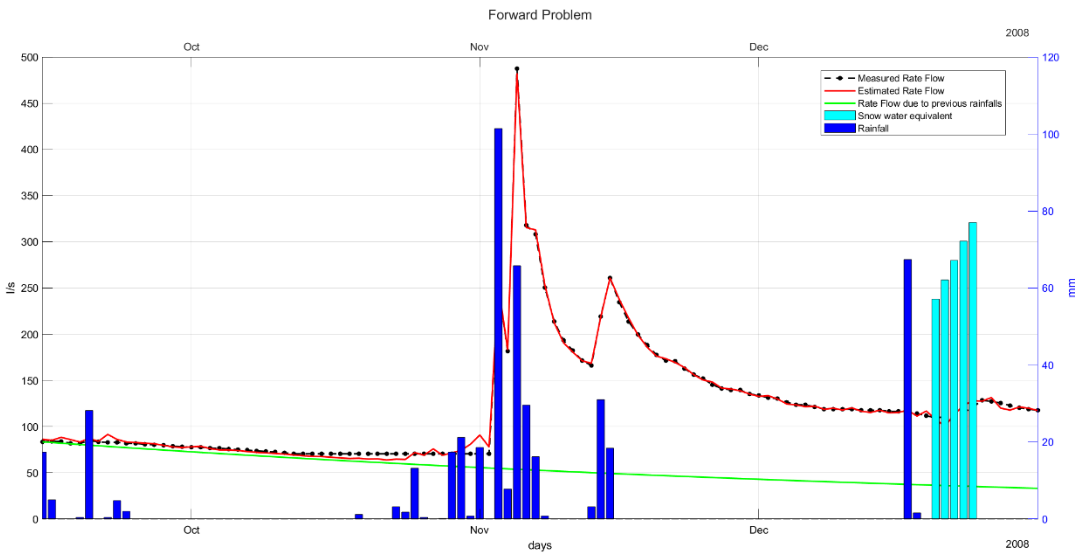


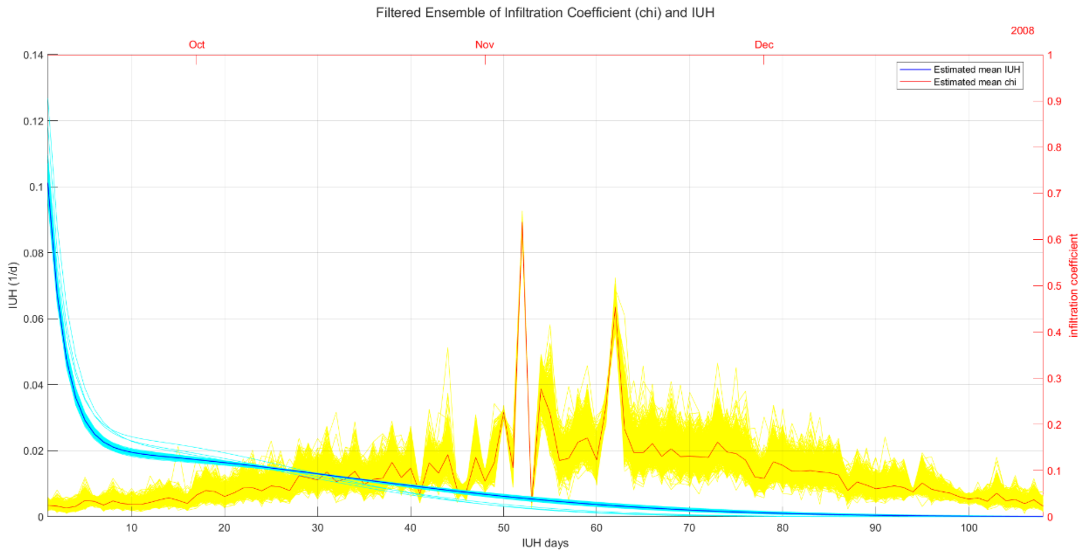
Figure 8-13:data used.

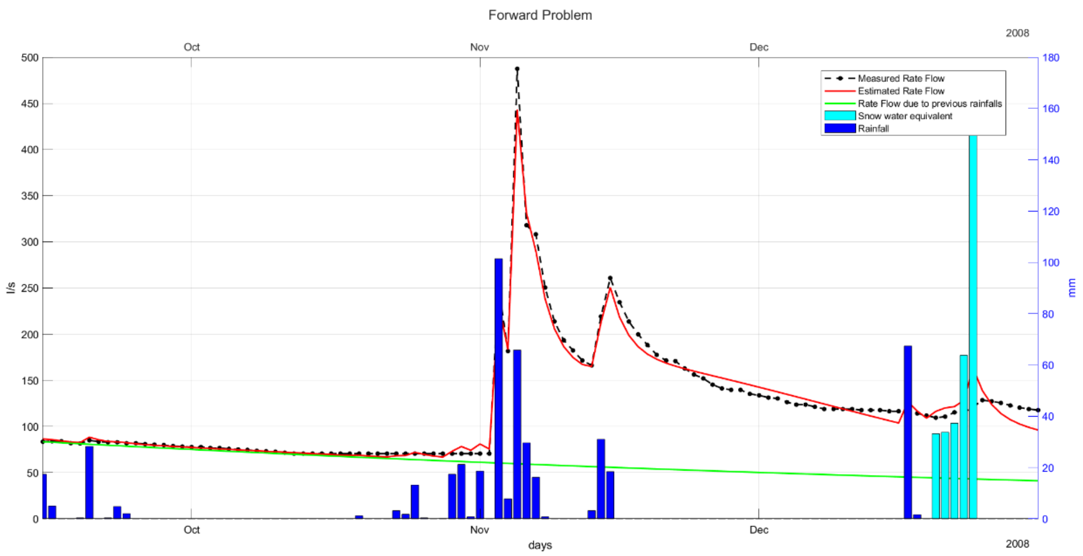
Non-parametric IUH					
Bias [l/s]	0.05	V_{in}/V_{out}	1.16	λ	1.05
Mean Absolute Error [l/s]	2.62	RMSE	4.67	DF	0.7
NSE	1.00	RMSE/ES	2.38	EM	2





Parametric IUH					
Bias [l/s]	-0.55	V_{in}/V_{out}	1.18	λ	1.1
Mean Absolute Error [l/s]	6.36	RMSE	9.40	DF	0.6
NSE	0.98	RMSE/ES	2.62	EM	2




Forward Problem
2008
Oct
Nov
Dec
l/s
mm
days
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall

XIV. Spring 2010

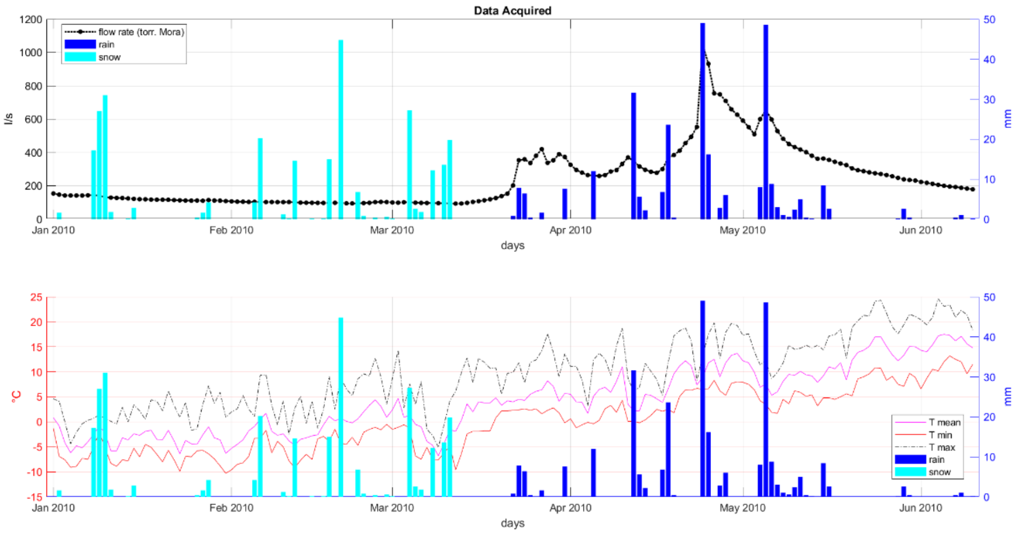
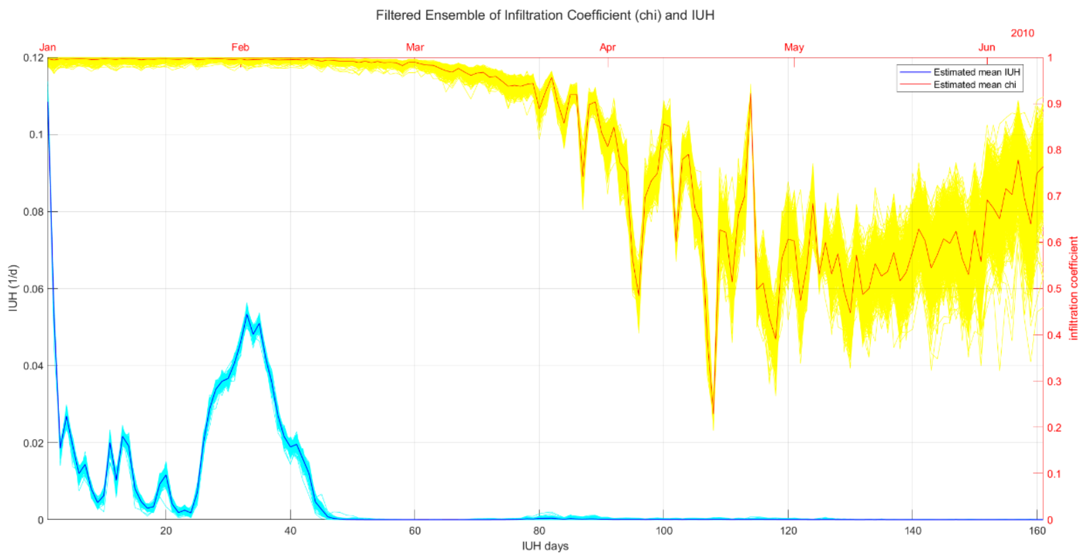
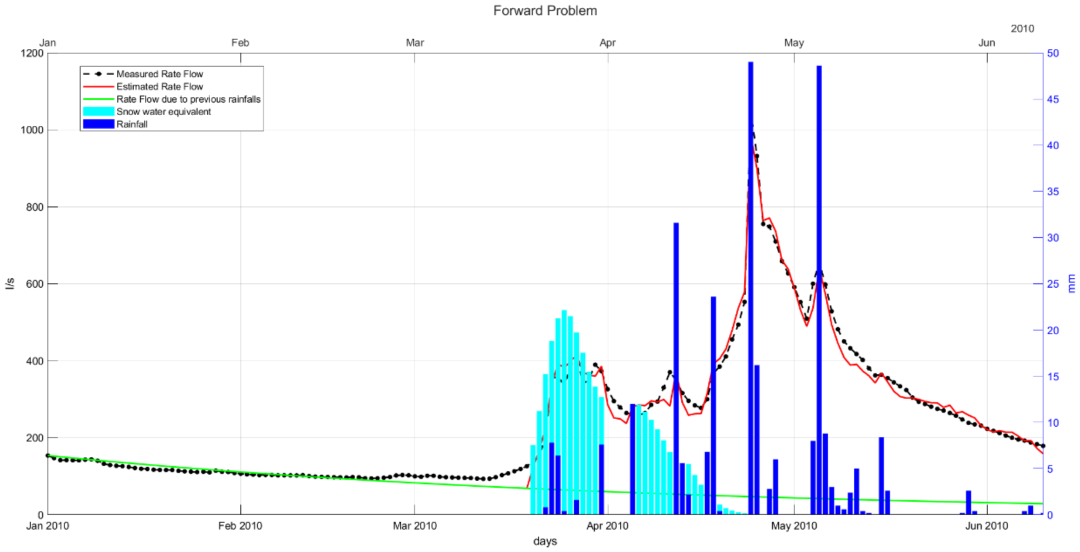


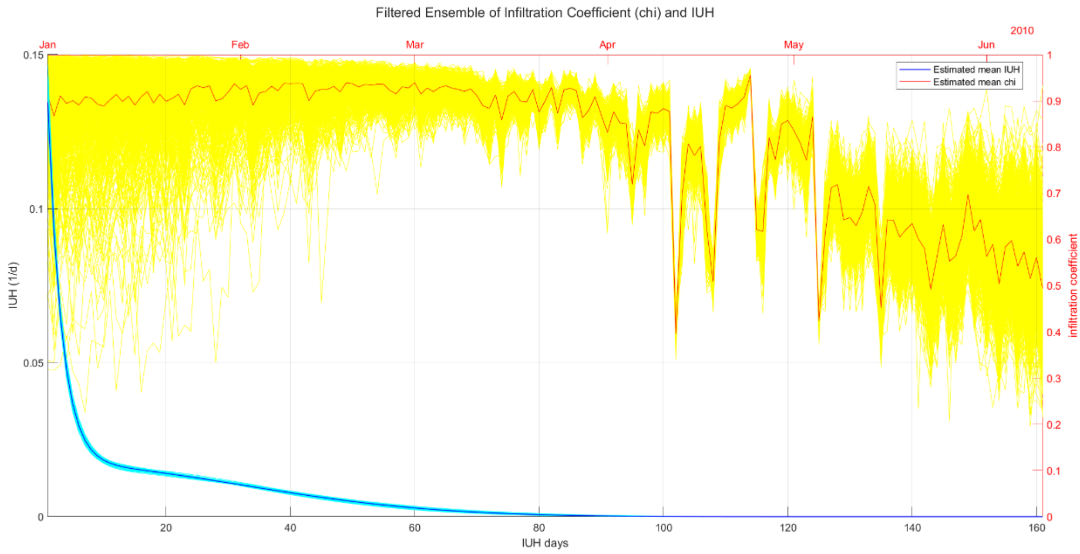
Figure 8-14:data used.

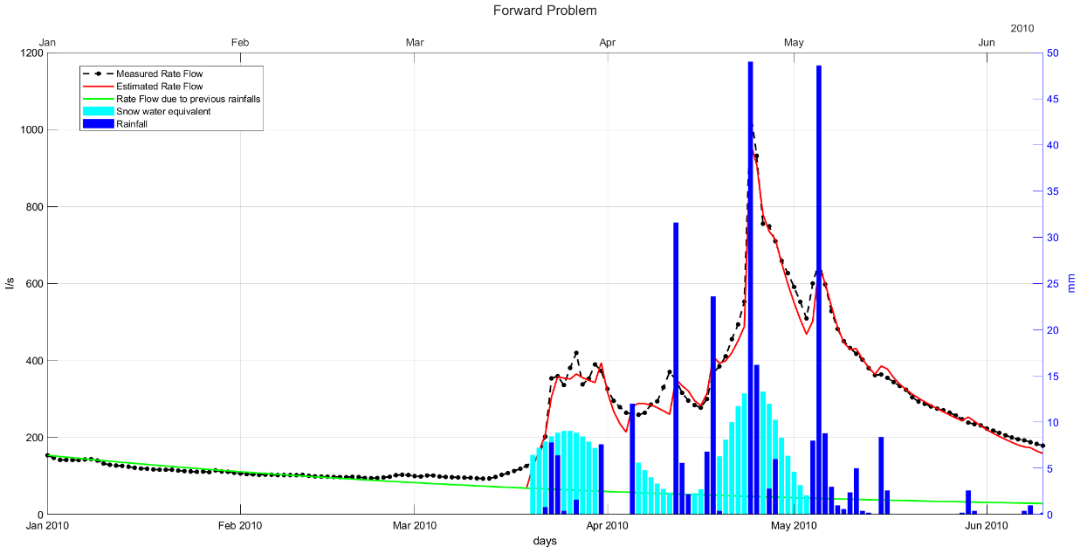
Non-parametric IUH					
Bias [l/s]	-4.50	V_{in}/V_{out}	1.04	λ	1.07
Mean Absolute Error [l/s]	16.28	RMSE	26.30	DF	0.55
NSE	0.99	RMSE/ES	4.61	EM	2





Parametric IUH					
Bias [l/s]	-7.17	V_{in}/V_{out}	1.08	λ	1.1
Mean Absolute Error [l/s]	15.74	RMSE	24.18	DF	0.5
NSE	0.98	RMSE/ES	3.85	EM	4




Forward Problem
2010
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
Jan 2010
Feb 2010
Mar 2010
Apr 2010
May 2010
Jun 2010

XV. Fall 2010

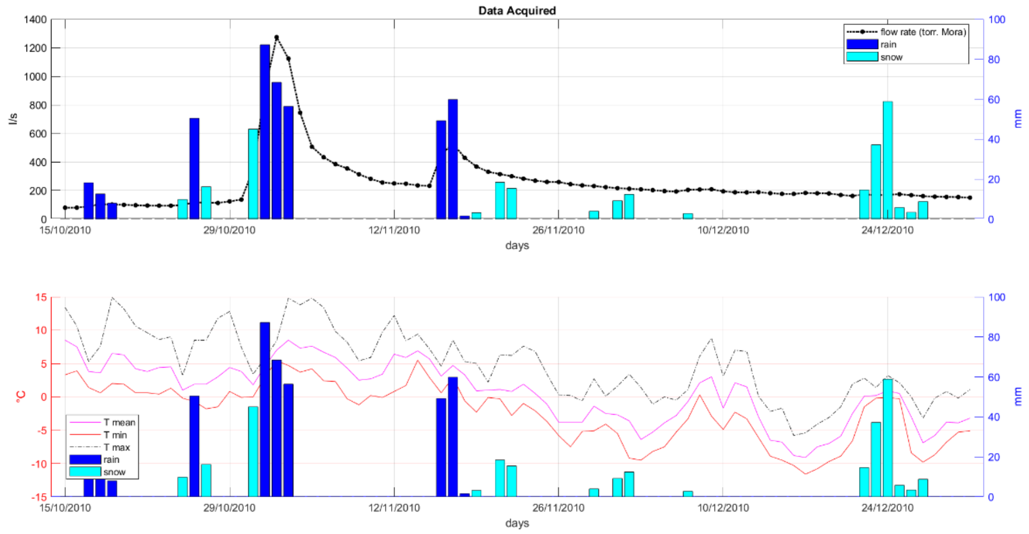
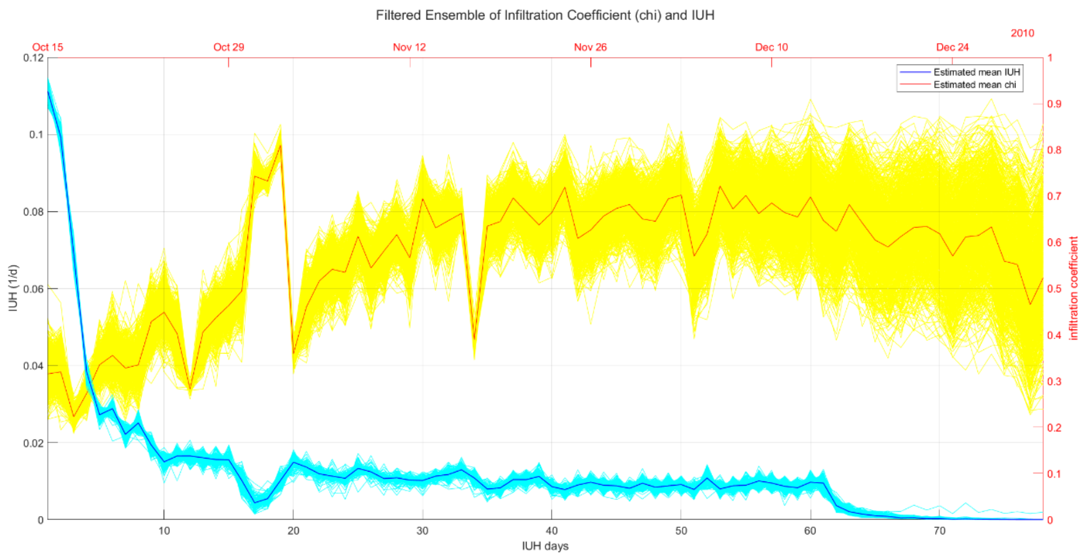
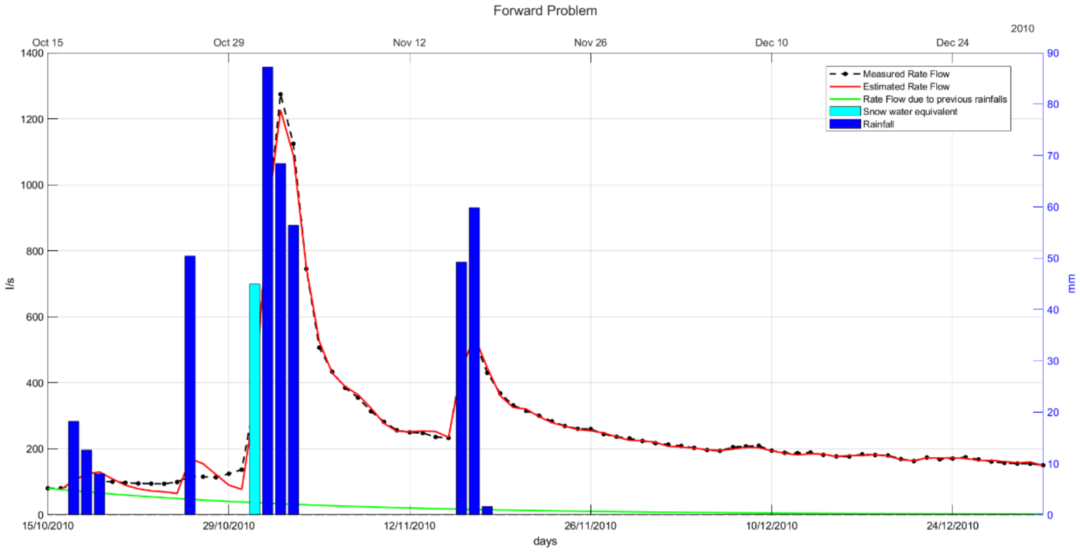


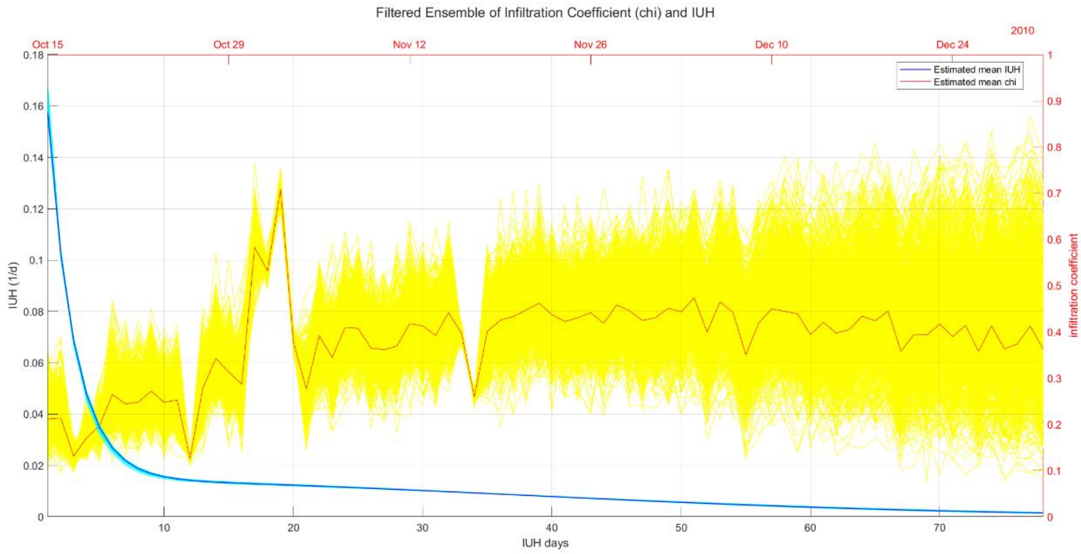
Figure 8-15:data used.

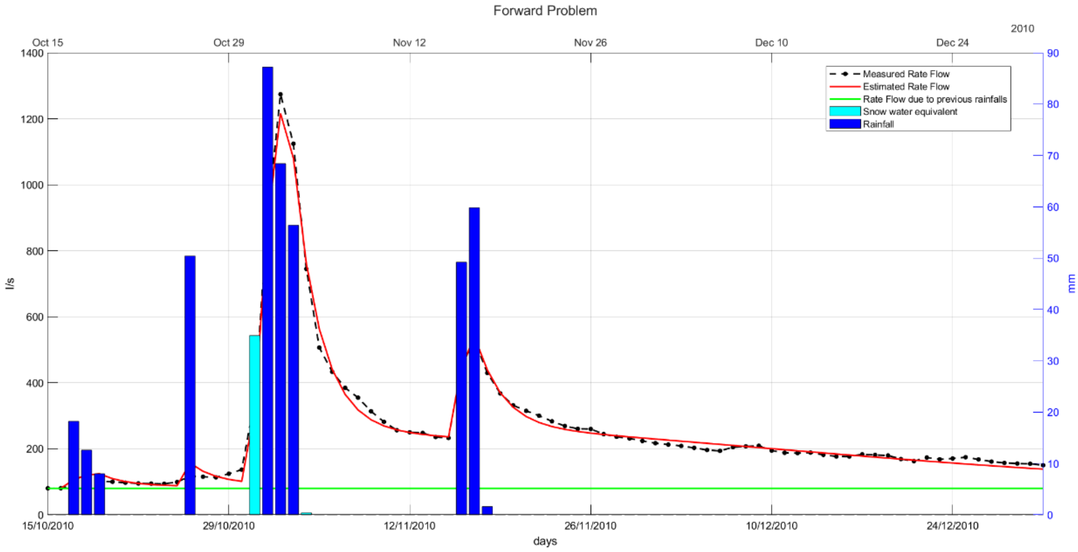
Non-parametric IUH					
Bias [l/s]	-1.39	$^{V_{in}}/_{V_{out}}$	1.04	λ	1.07
Mean Absolute Error [l/s]	9.26	RMSE	18.29	DF	0.55
NSE	0.99	RMSE/ES	4.44	EM	2





Parametric IUH					
Bias [l/s]	-2.65	V_{in}/V_{out}	1.06	λ	1.1
Mean Absolute Error [l/s]	12.74	RMSE	17.43	DF	0.5
NSE	0.99	RMSE/ES	2.92	EM	3





Forward Problem
2010
Oct 15
Oct 29
Nov 12
Nov 26
Dec 10
Dec 24
l/s
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
15/10/2010
29/10/2010
12/11/2010
26/11/2010
10/12/2010
24/12/2010
days

XVI. Spring 2011

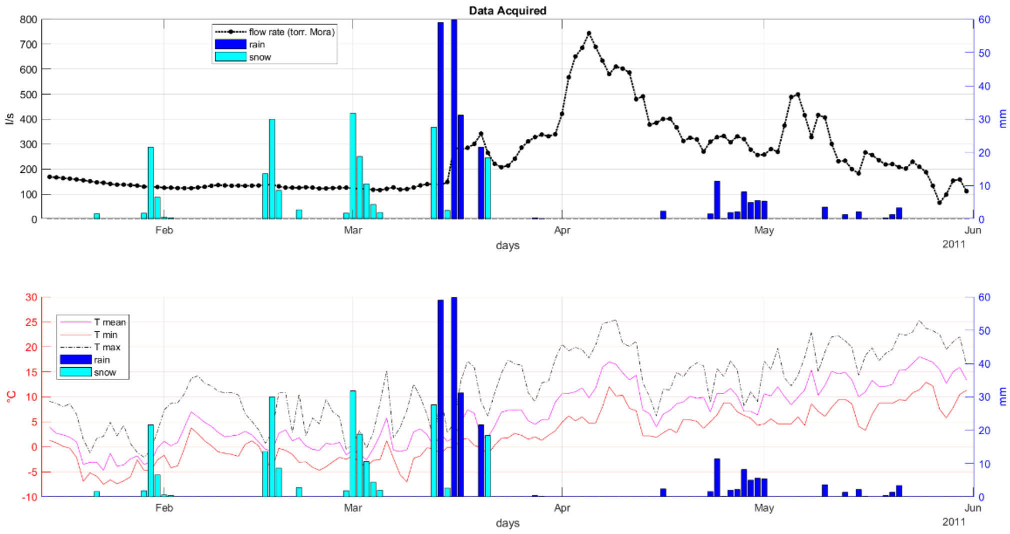
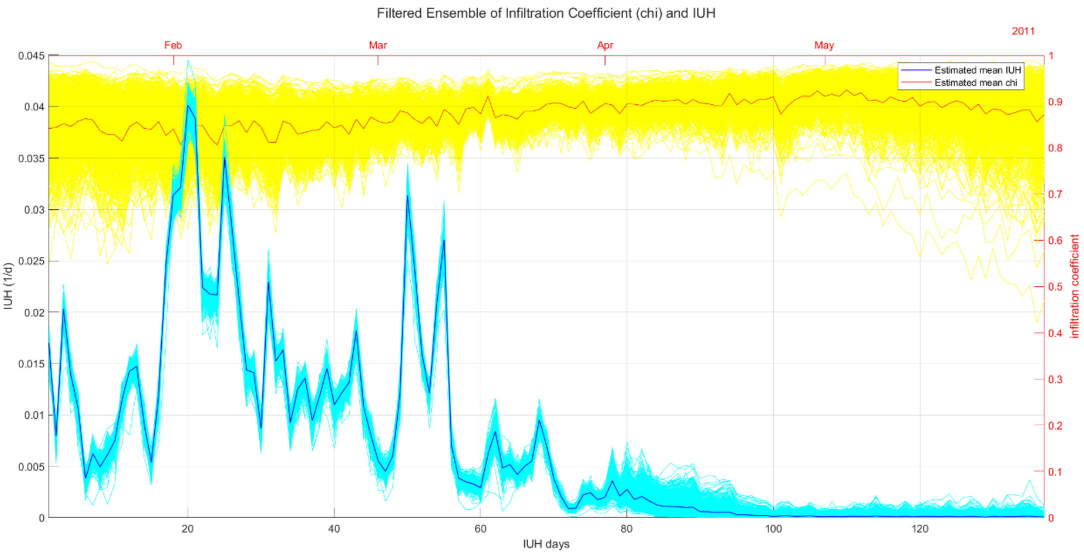
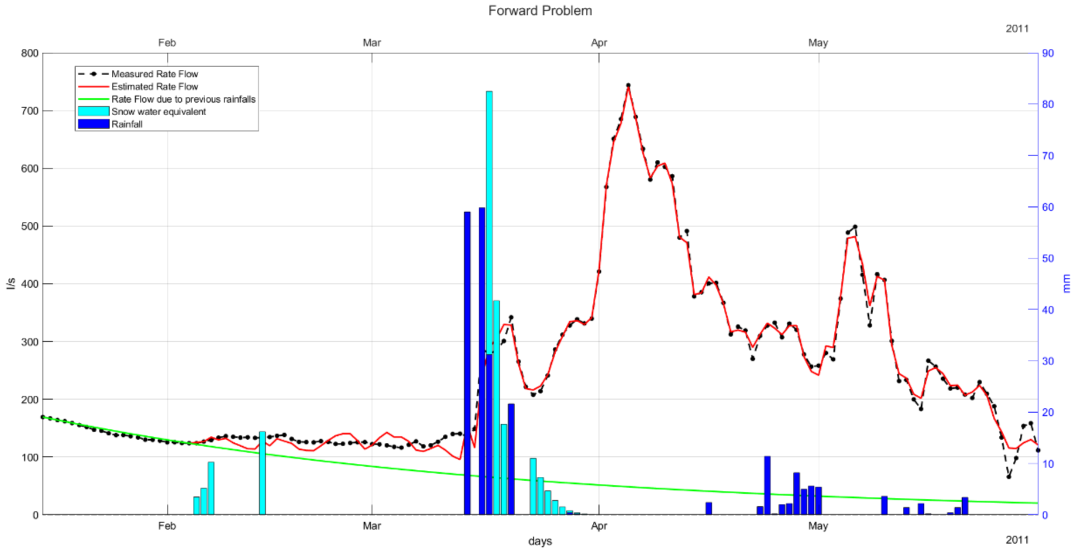


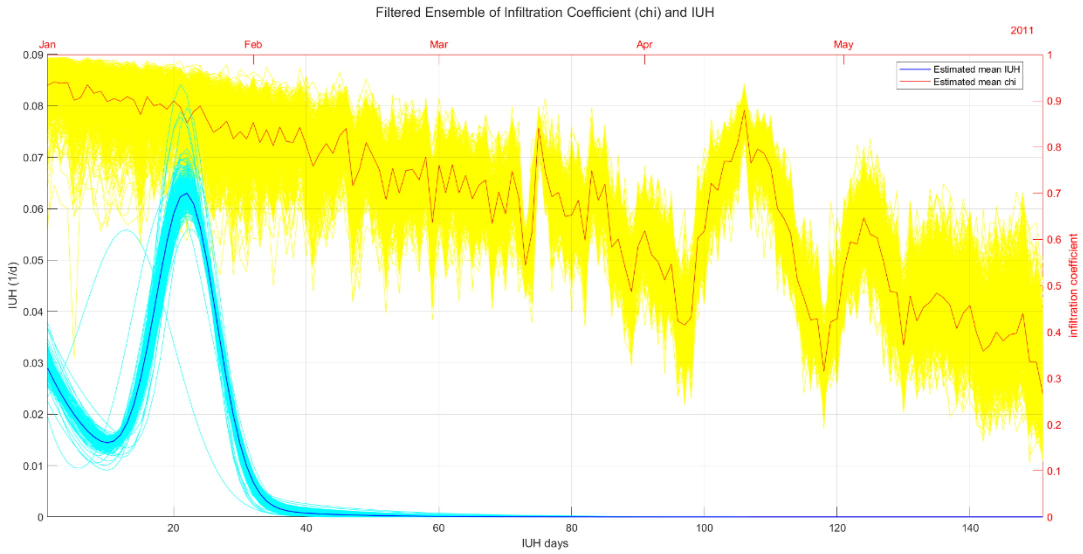
Figure 8-16:data used.

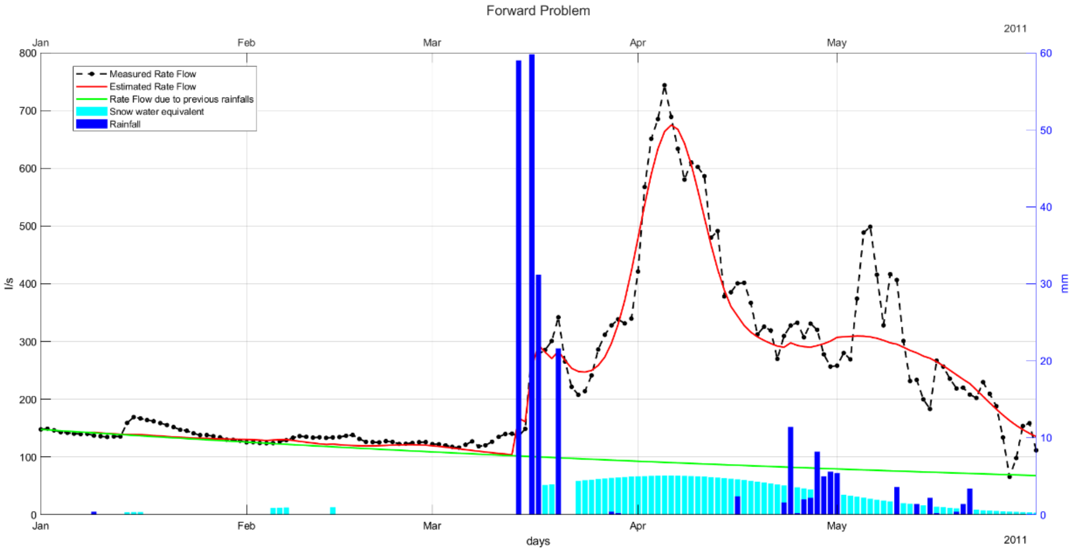
Non-parametric IUH					
Bias [l/s]	-2.55	V_{in}/V_{out}	1.10	λ	1.07
Mean Absolute Error [l/s]	10.00	RMSE	14.86	DF	0.6
NSE	0.99	RMSE/ES	3.93	EM	3





Parametric IUH					
Bias [l/s]	-8.40	V_{in}/V_{out}	1.04	λ	1.1
Mean Absolute Error [l/s]	25.77	RMSE	41.05	DF	0.4
NSE	0.92	RMSE/ES	5.0	EM	3





Forward Problem
2011
Jan
Feb
Mar
Apr
May
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
0
100
200
300
400
500
600
700
800
10
20
30
40
50
60

XVII. Fall 2011

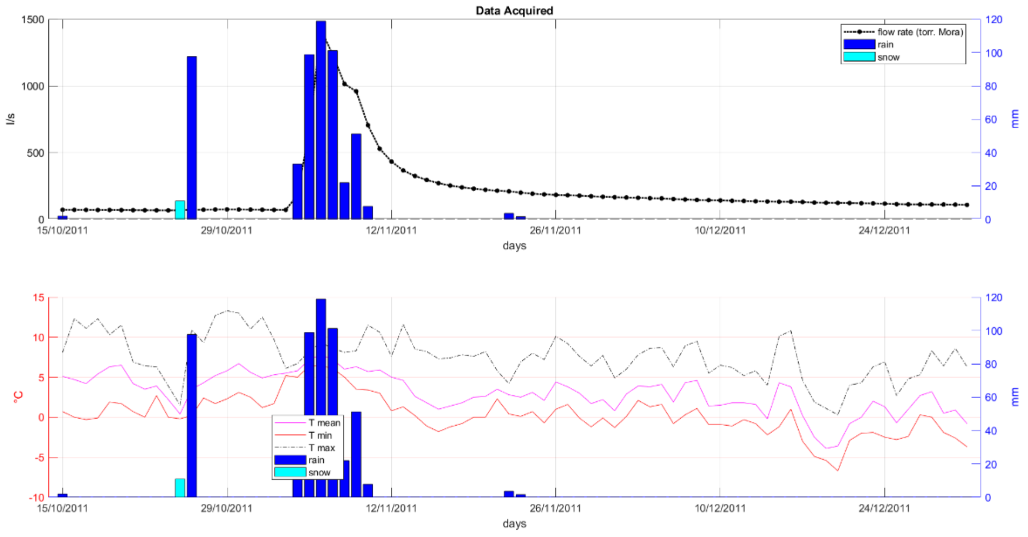
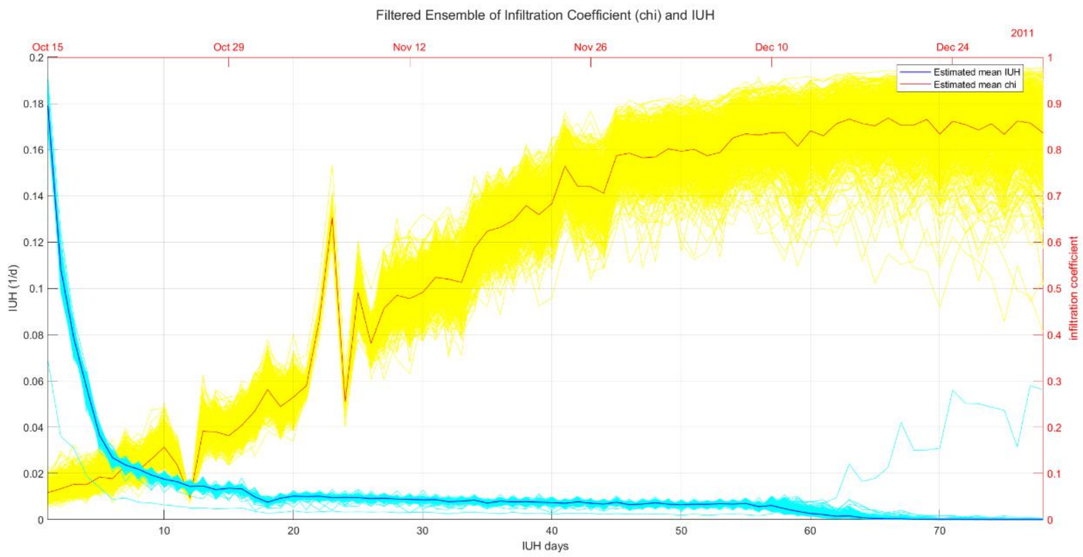
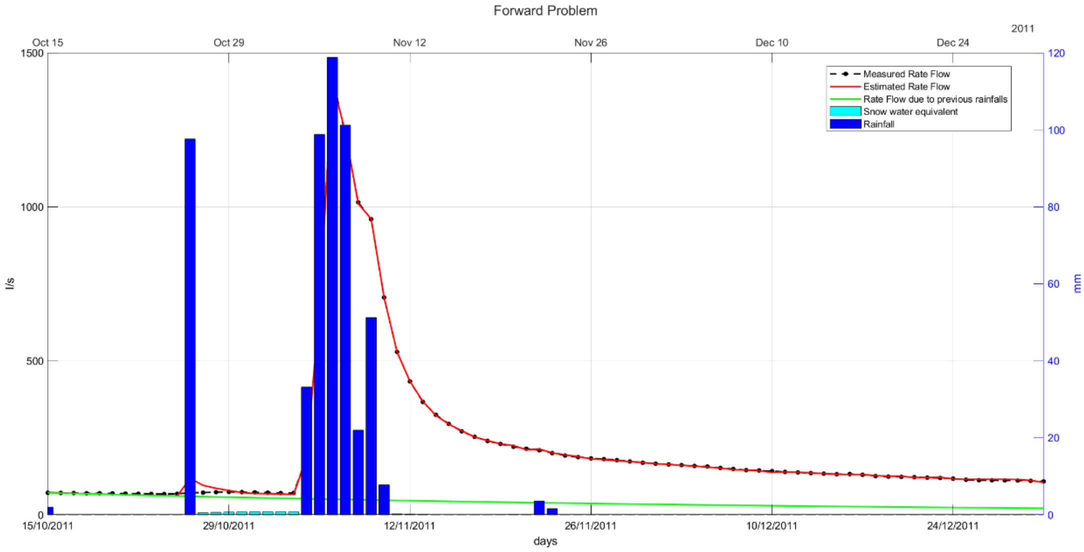


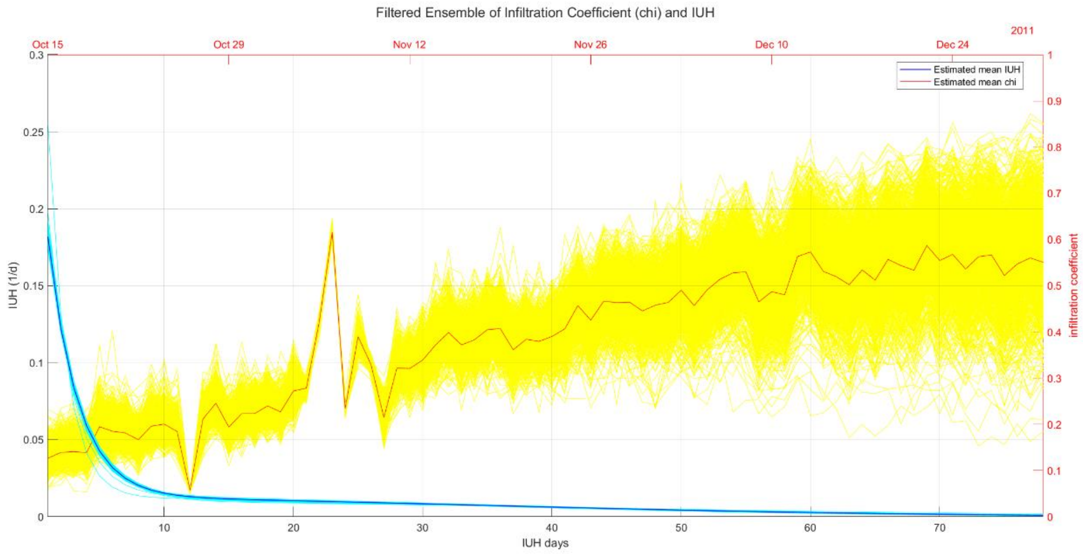
Figure 8-17:data used.

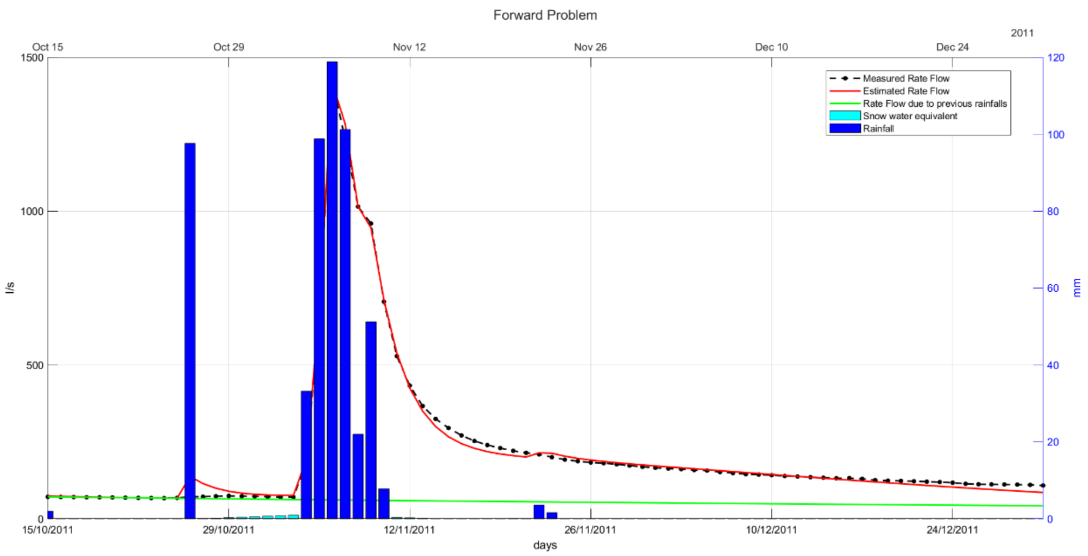
Non-parametric IUH					
Bias [l/s]	-0.08	$^{V_{in}}/_{V_{out}}$	1.03	λ	1.07
Mean Absolute Error [l/s]	3.21	RMSE	8.36	DF	0.6
NSE	1.00	RMSE/ES	1.74	EM	3





Parametric IUH					
Bias [l/s]	-0.70	V_{in}/V_{out}	1.05	λ	1.07
Mean Absolute Error [l/s]	10.88	RMSE	17.75	DF	0.3
NSE	1.00	RMSE/ES	3.22	EM	3




Forward Problem
2011
Oct 15
Oct 29
Nov 12
Nov 26
Dec 10
Dec 24
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
15/10/2011
29/10/2011
12/11/2011
26/11/2011
10/12/2011
24/12/2011

XVIII. Spring 2012

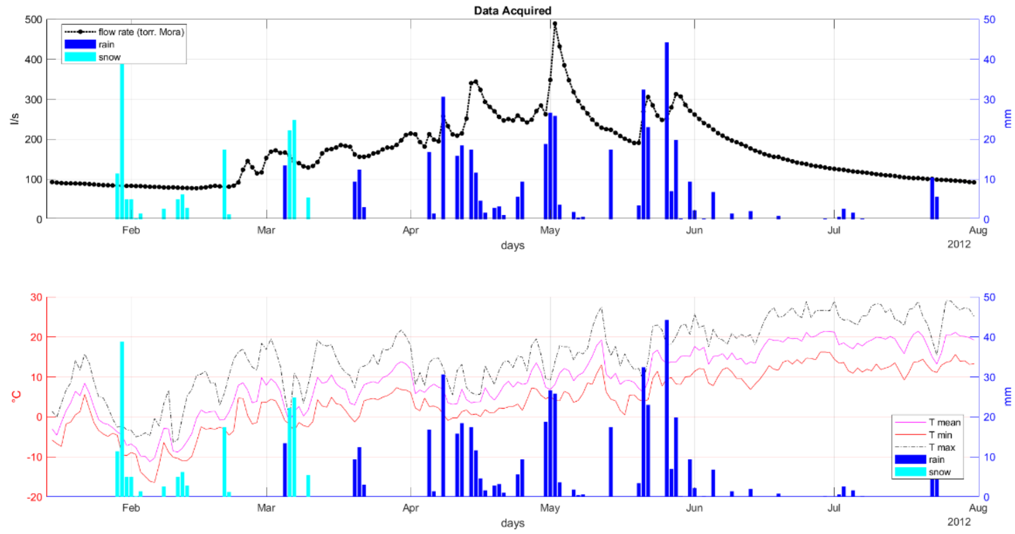
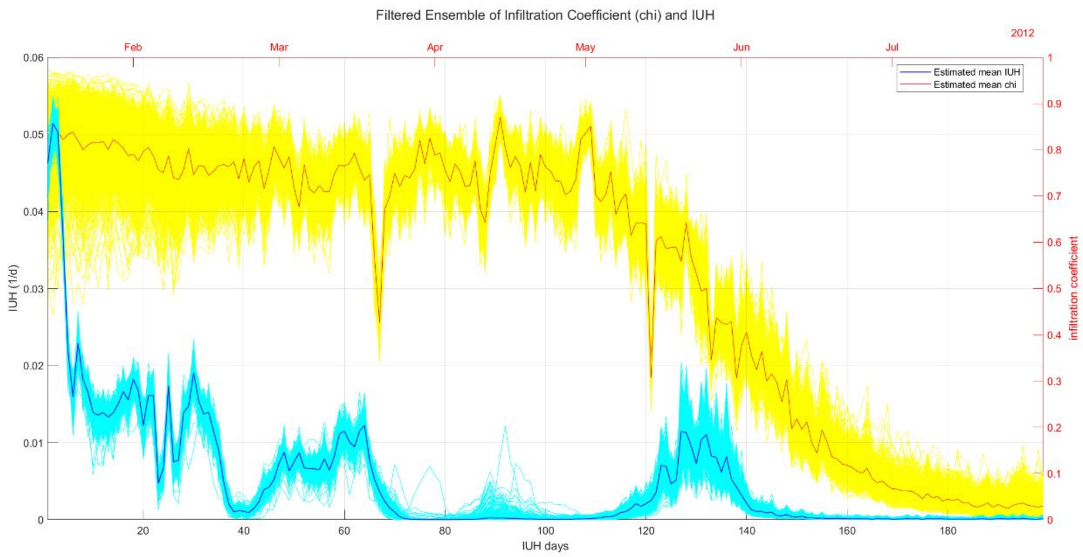
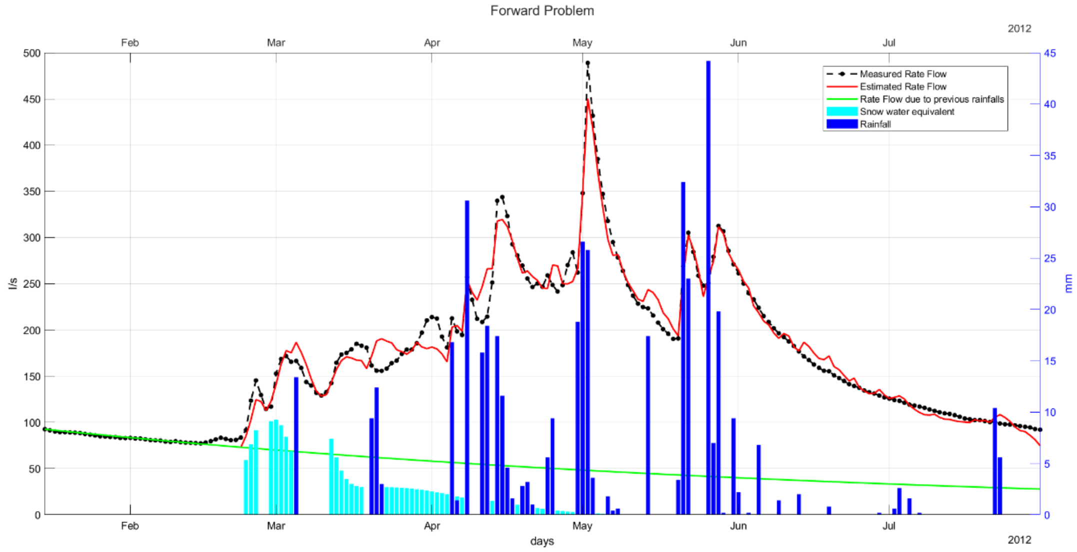


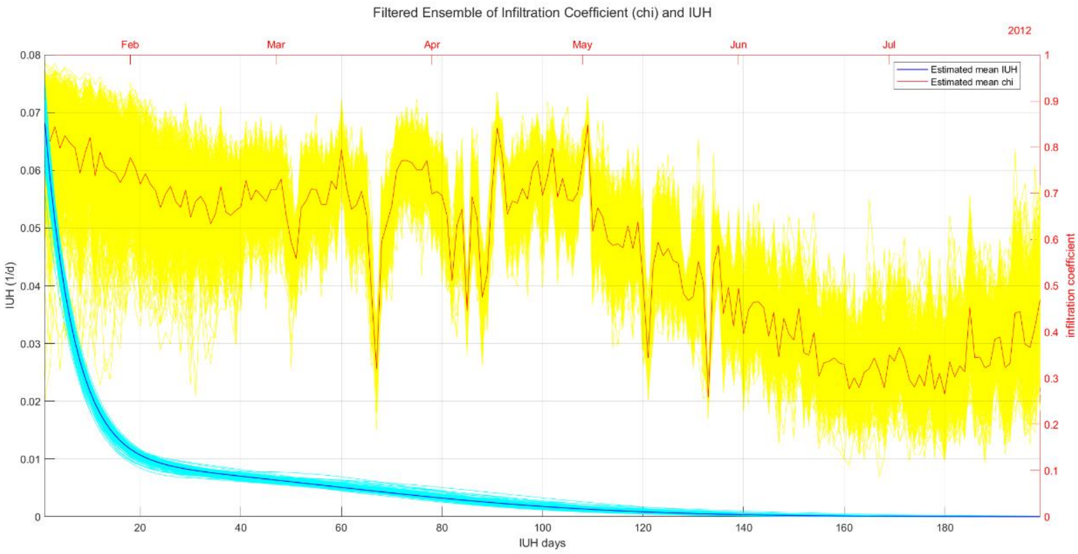
Figure 8-18:data used.

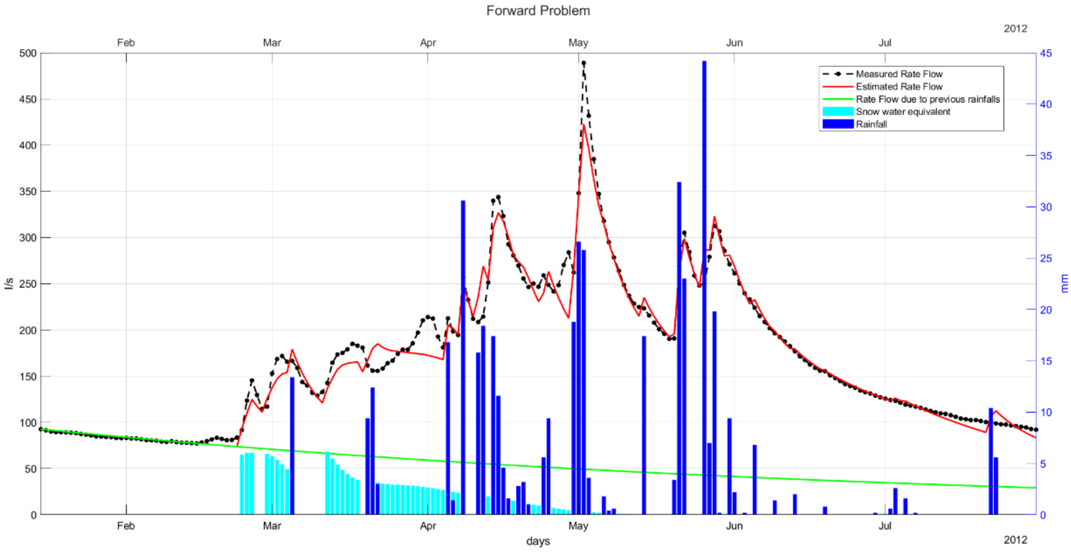
Non-parametric IUH					
Bias [l/s]	-0.95	V_{in}/V_{out}	1.14	λ	1.08
Mean Absolute Error [l/s]	8.30	RMSE	12.90	DF	0.5
NSE	0.98	RMSE/ES	4.79	EM	3





Parametric IUH					
Bias [l/s]	-2.36	V_{in}/V_{out}	1.07	λ	1.10
Mean Absolute Error [l/s]	8.12	RMSE	14.10	DF	0.4
NSE	0.97	RMSE/ES	4.87	EM	3





Forward Problem
2012
Feb
Mar
Apr
May
Jun
Jul
l/s
mm
days
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall

XIX. Fall 2012

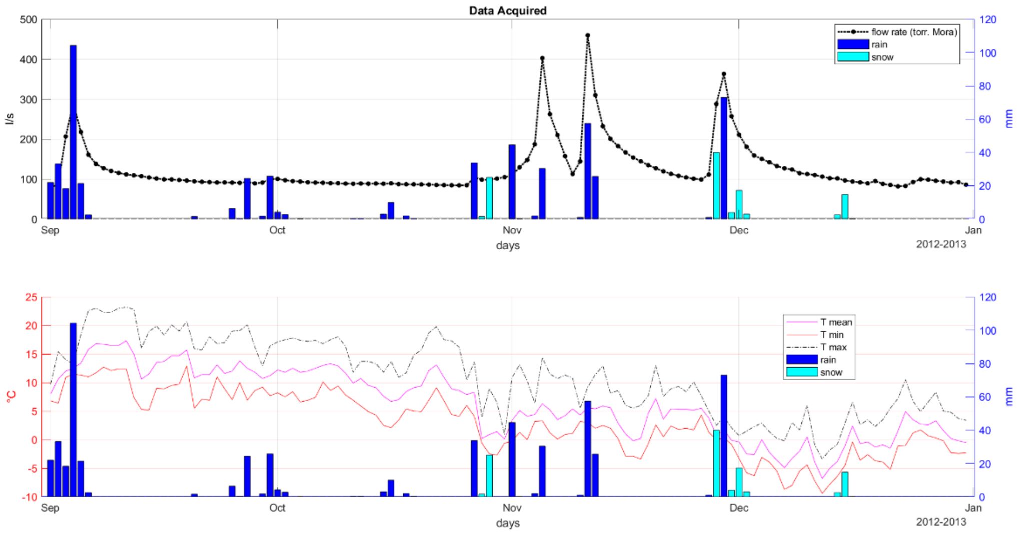
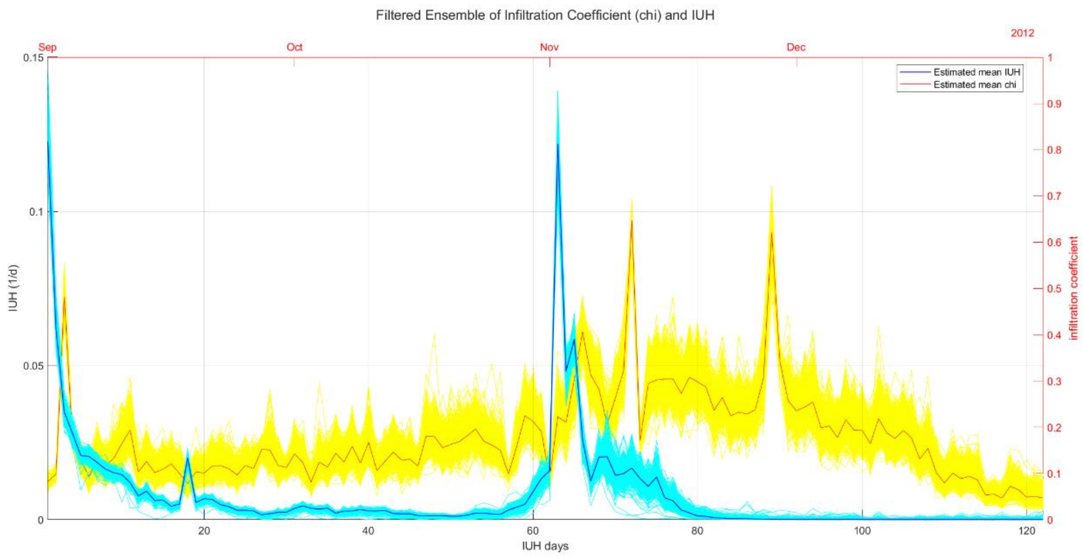
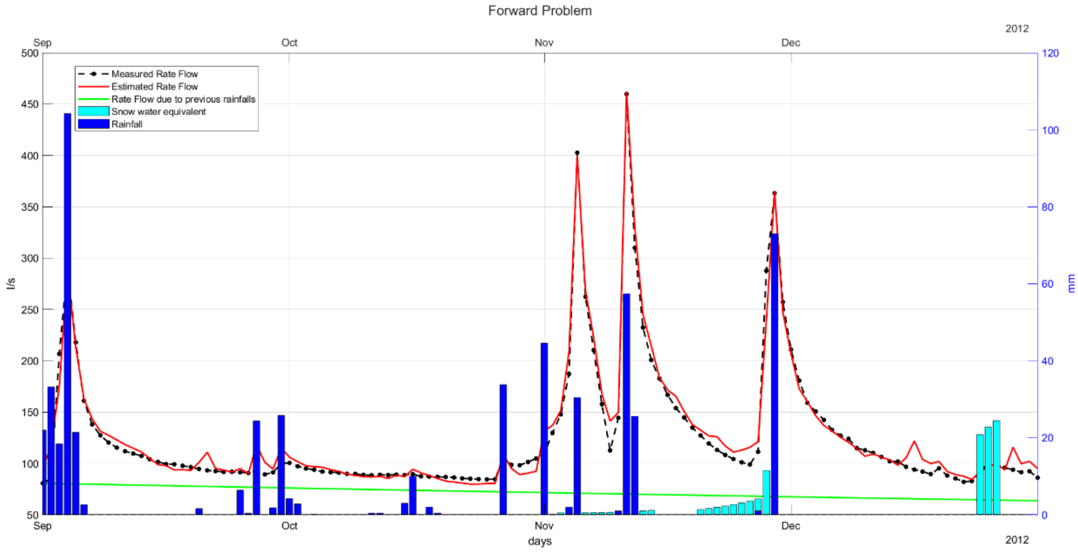


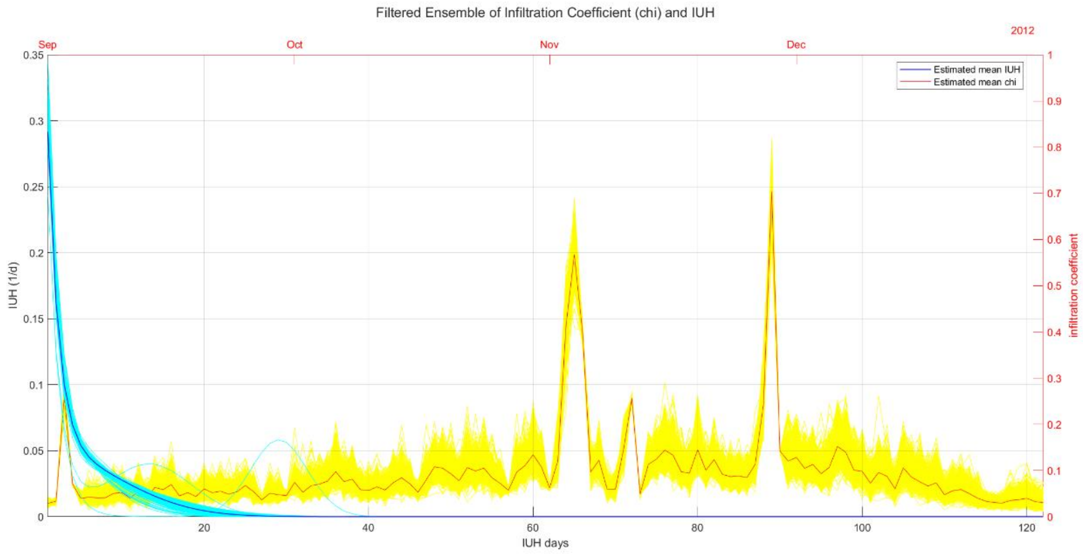
Figure 8-19:data used.

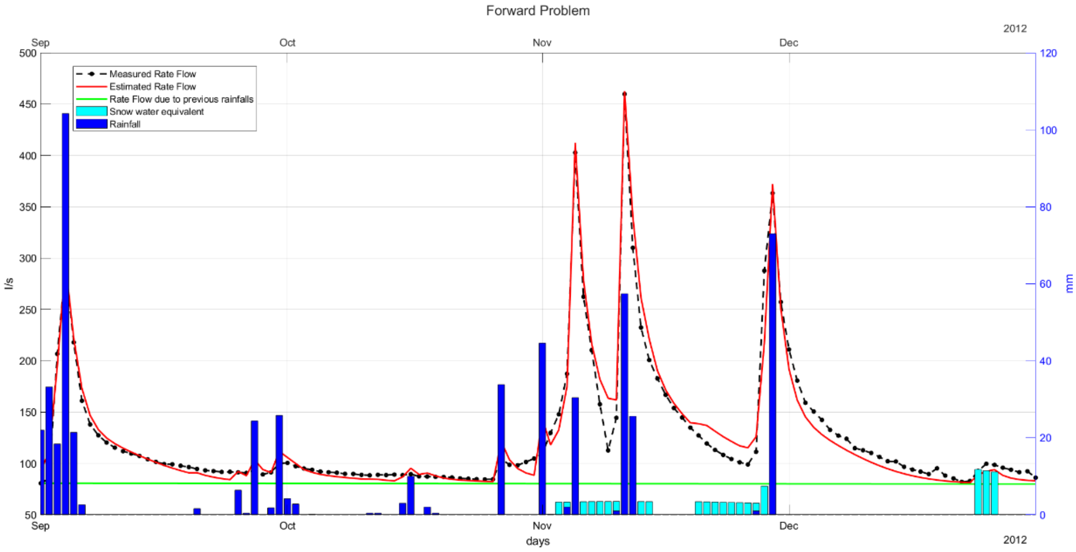
Non-parametric IUH					
Bias [l/s]	2.90	V_{in}/V_{out}	1.51	λ	1.08
Mean Absolute Error [l/s]	7.76	RMSE	12.14	DF	0.6
NSE	0.97	RMSE/ES	3.17	EM	3





Parametric IUH					
Bias [l/s]	0.44	V_{in}/V_{out}	1.00	λ	1.10
Mean Absolute Error [l/s]	9.00	RMSE	17.72	DF	0.5
NSE	0.96	RMSE/ES	3.87	EM	3





Forward Problem
2012
Sep
Oct
Nov
Dec
500
450
400
350
300
250
200
150
100
50
l/s
120
100
80
60
40
20
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days

XX. Spring 2013

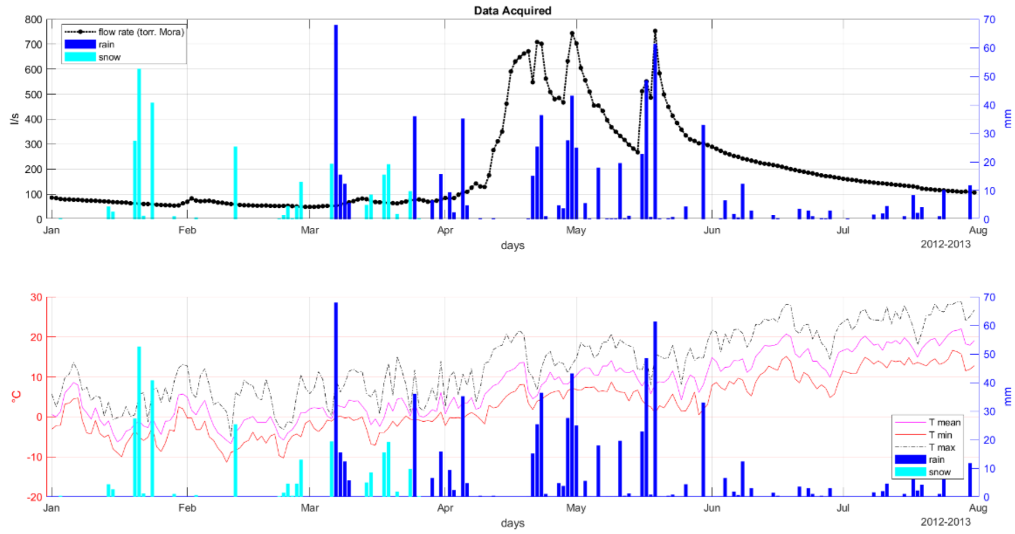
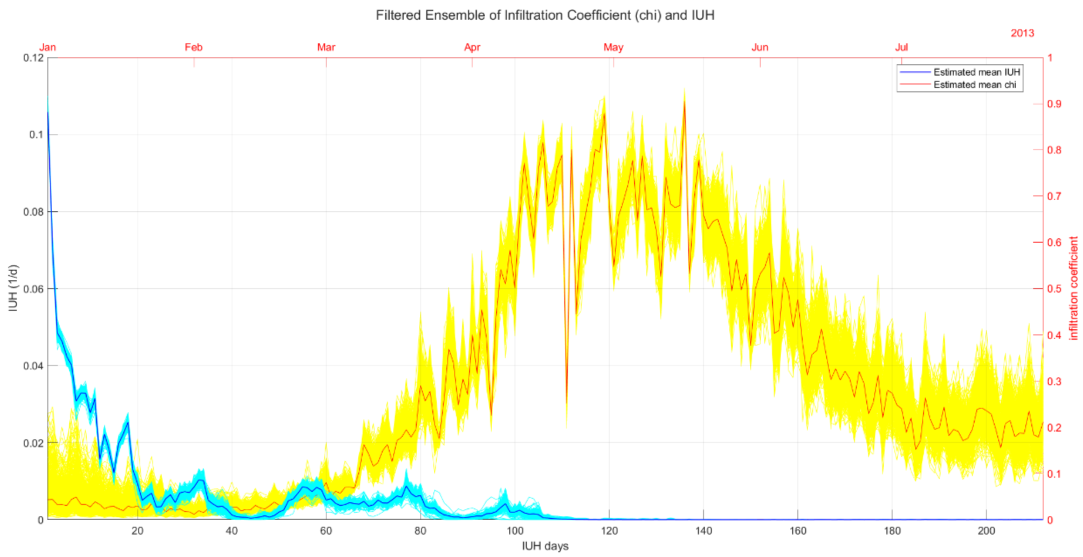
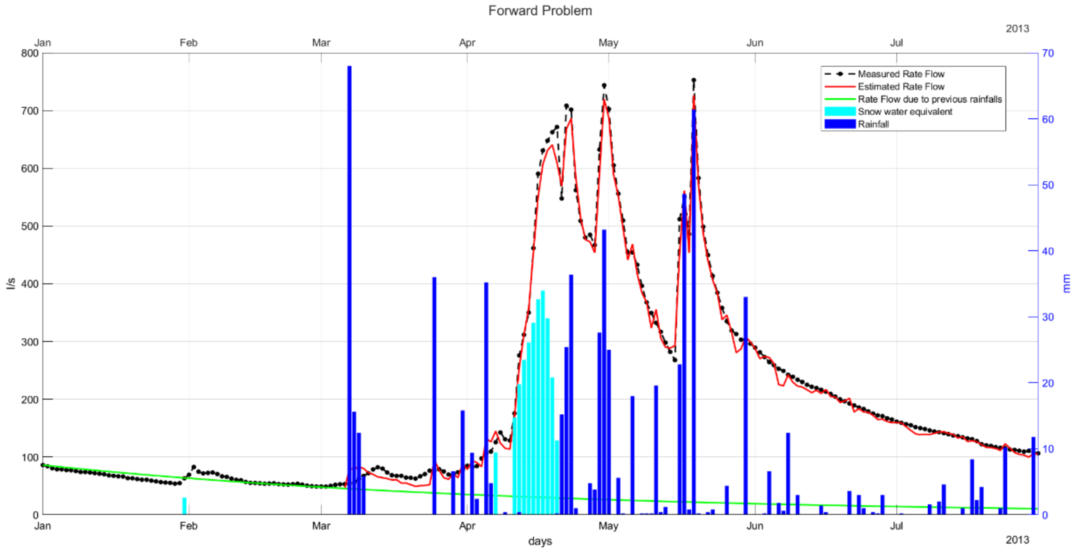


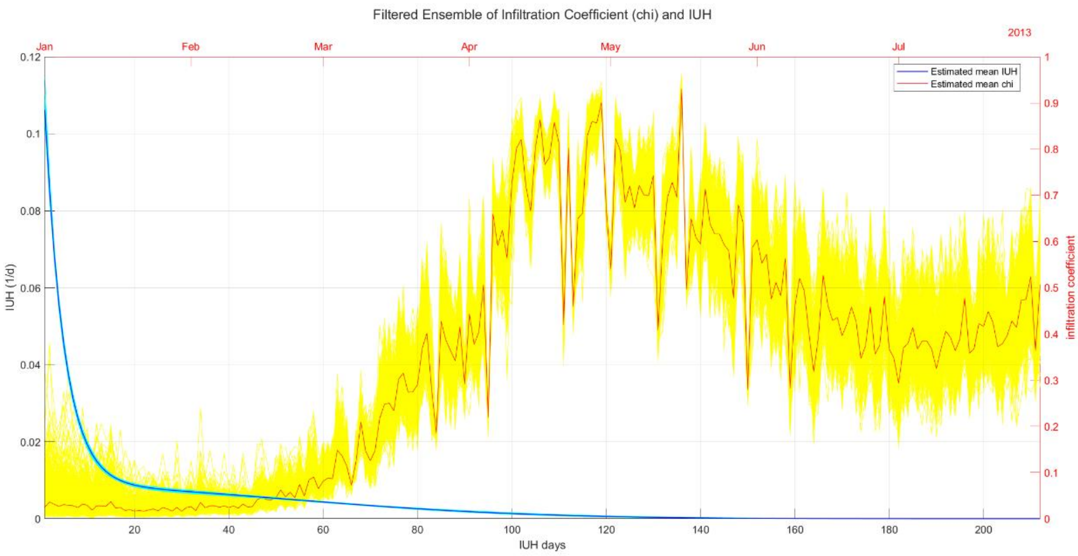
Figure 8-20:data used.

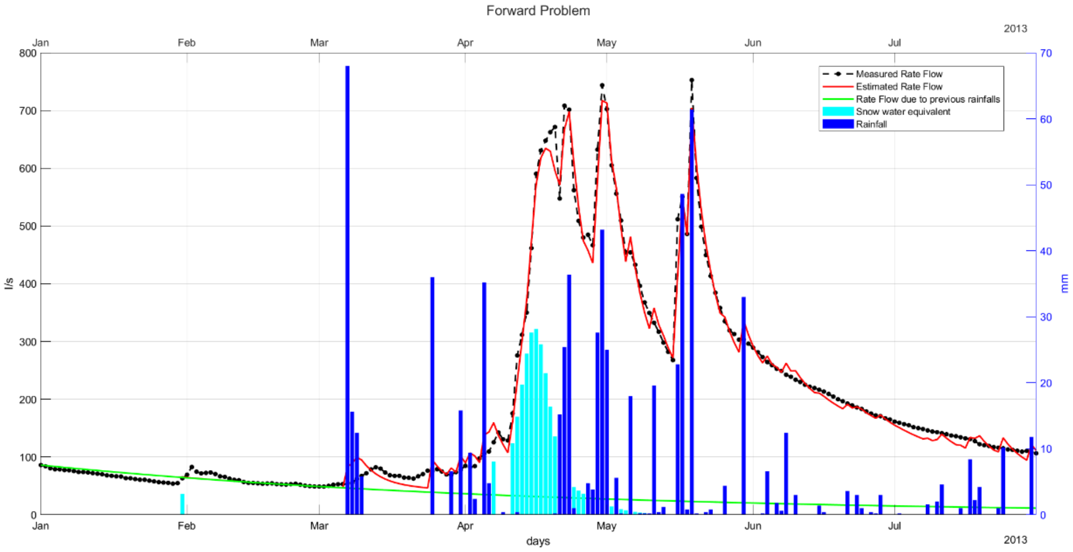
Non-parametric IUH					
Bias [l/s]	-3.87	V_{in}/V_{out}	1.05	λ	1.08
Mean Absolute Error [l/s]	9.53	RMSE	13.96	DF	0.6
NSE	0.99	RMSE/ES	3.87	EM	3





Parametric IUH					
Bias [l/s]	-1.24	V_{in}/V_{out}	1.07	λ	1.10
Mean Absolute Error [l/s]	11.78	RMSE	17.95	DF	0.4
NSE	0.99	RMSE/ES	6.30	EM	3

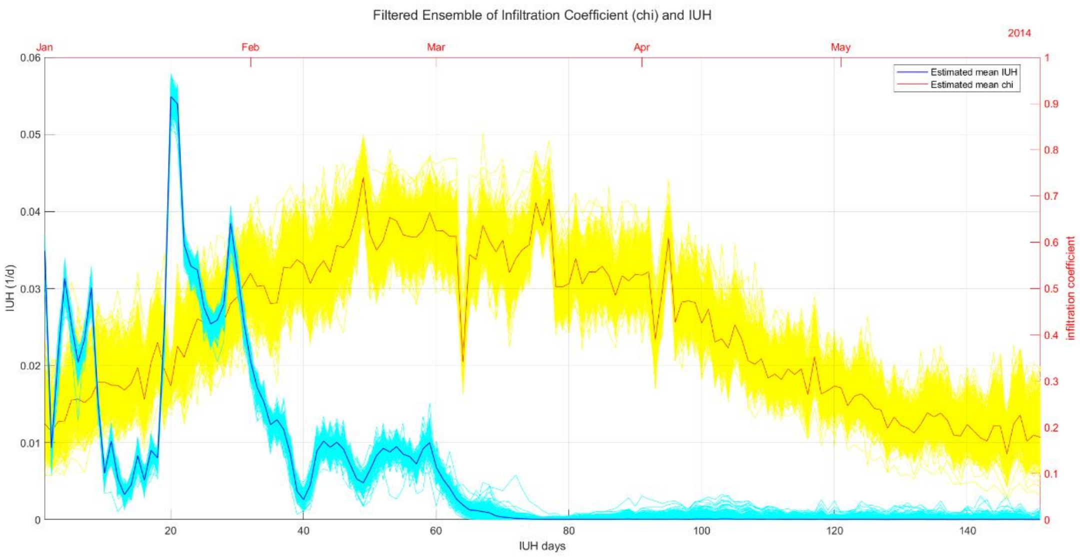


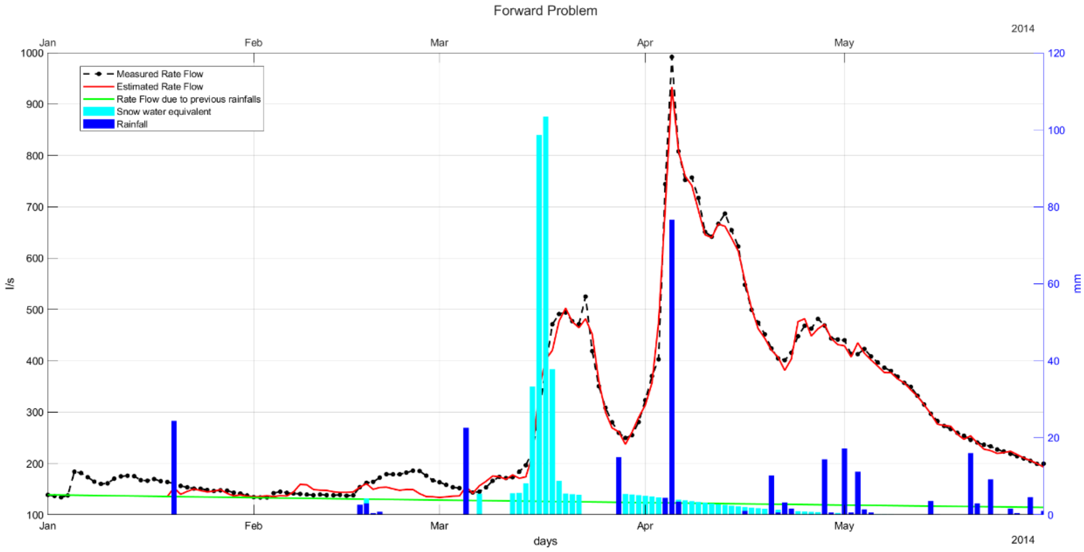

Forward Problem
2013
Jan
Feb
Mar
Apr
May
Jun
Jul
800
700
600
500
400
300
200
100
0
l/s
70
60
50
40
30
20
10
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days
2013

XXI. Spring 2014

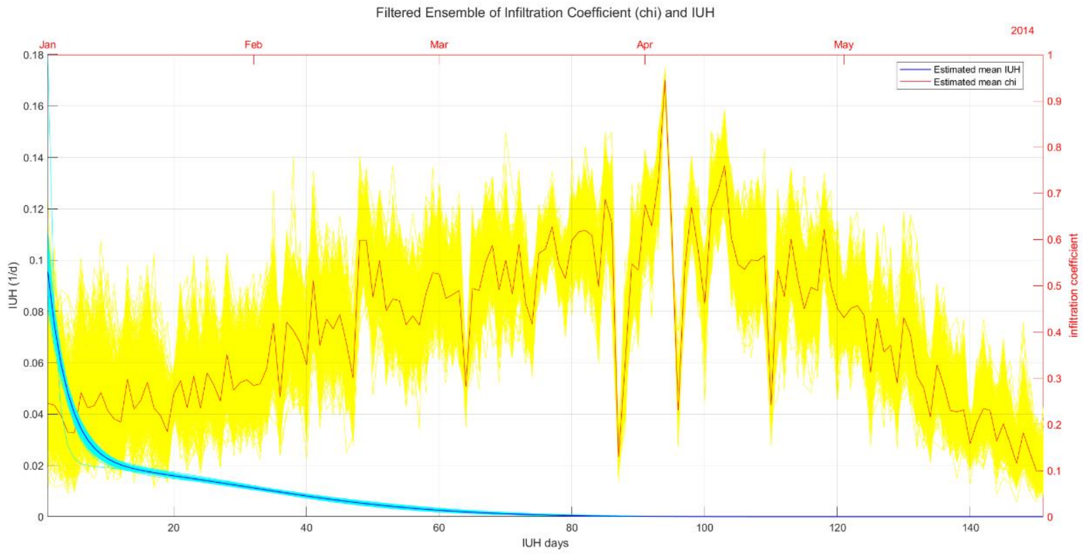
Figure 8-21:data used.

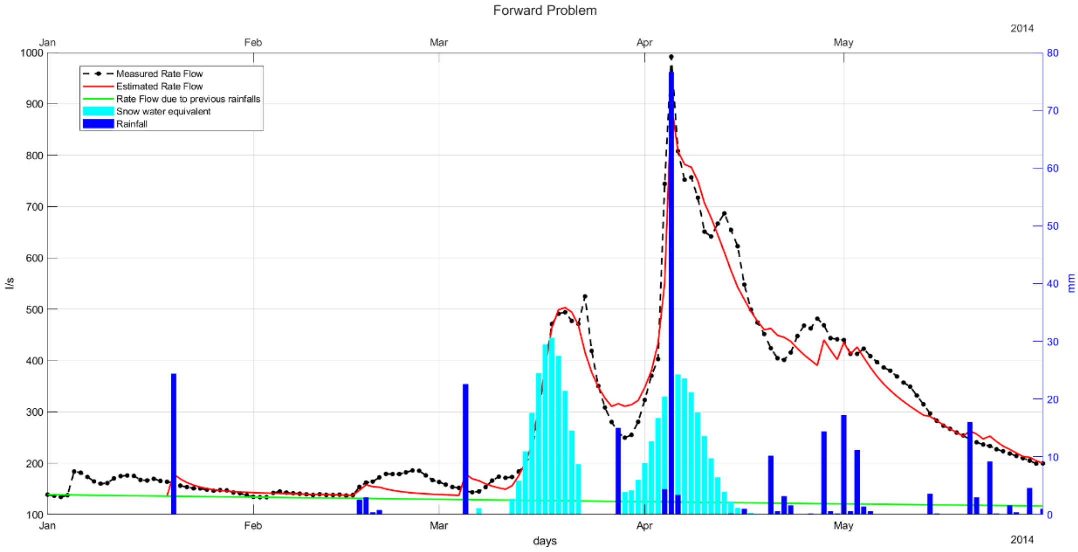
Non-parametric IUH					
Bias [l/s]	-8.14	V_{in}/V_{out}	1.05	λ	1.07
Mean Absolute Error [l/s]	14.00	RMSE	20.94	DF	0.5
NSE	0.99	RMSE/ES	4.84	EM	2





Parametric IUH					
Bias [l/s]	-8.24	V_{in}/V_{out}	1.05	λ	1.10
Mean Absolute Error [l/s]	22.65	RMSE	34.76	DF	0.35
NSE	0.96	RMSE/ES	5.72	EM	3




Forward Problem
Jan
Feb
Mar
Apr
May
2014
1000
900
800
700
600
500
400
300
200
100
l/s
80
70
60
50
40
30
20
10
0
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days

XXII. Summer 2014

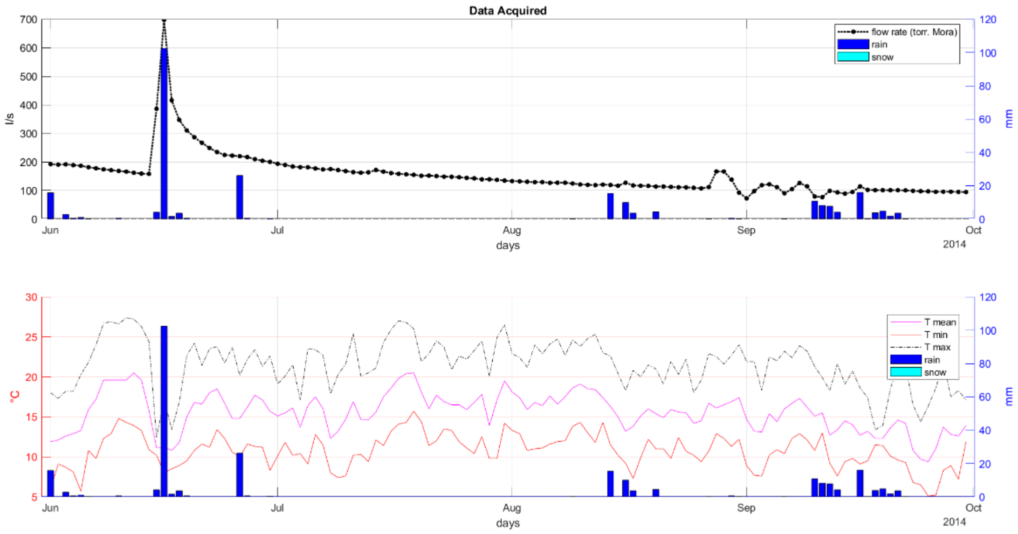
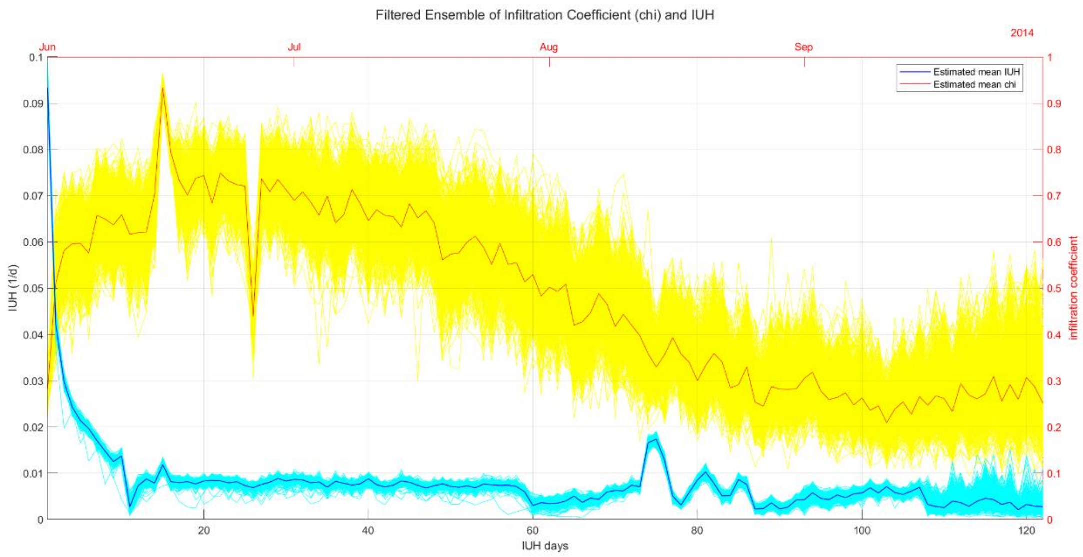
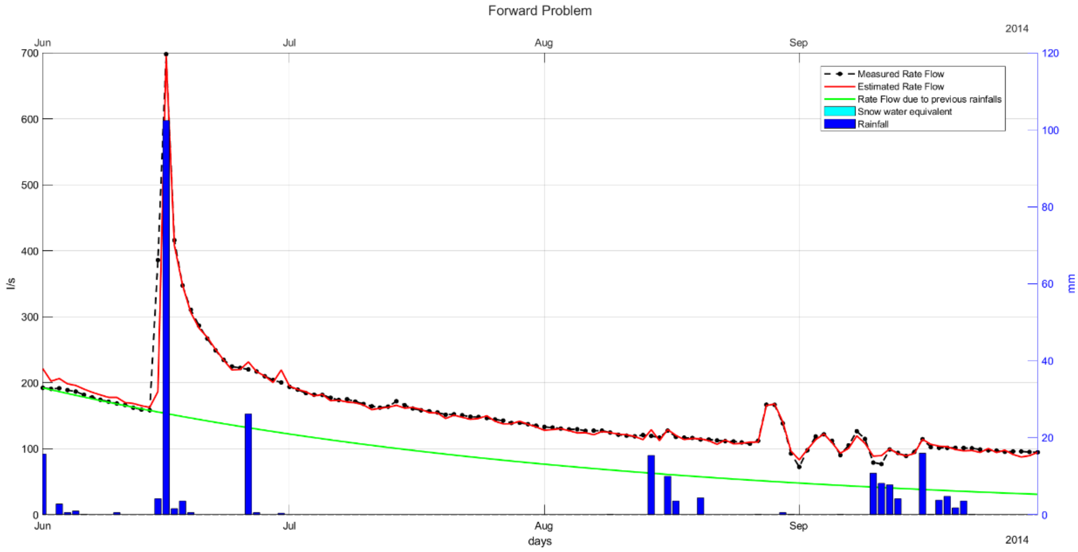


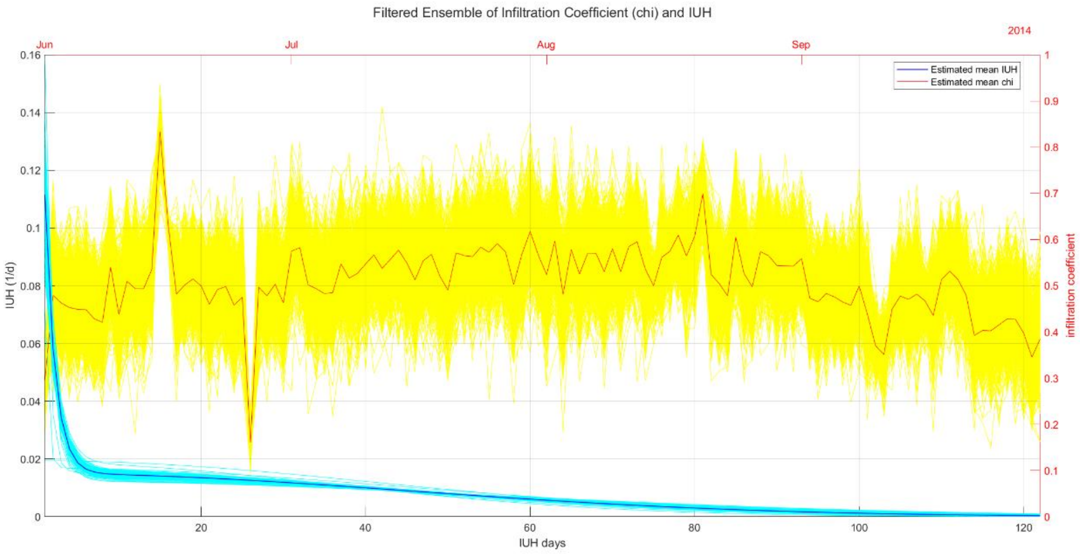
Figure 8-22:data used.

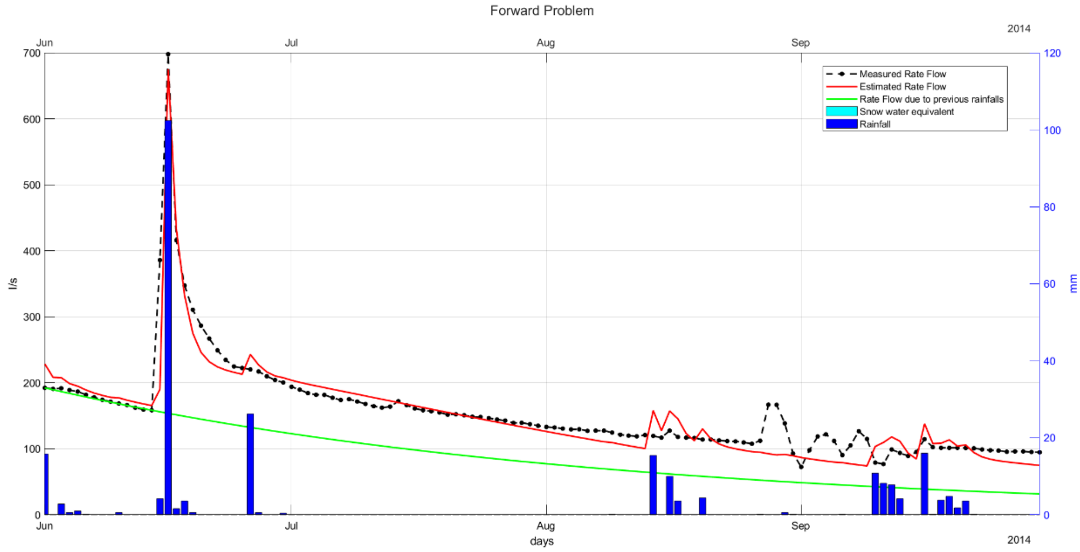
Non-parametric IUH					
Bias [l/s]	-1.34	V_{in}/V_{out}	1.18	λ	1.10
Mean Absolute Error [l/s]	5.62	RMSE	19.11	DF	0.6
NSE	0.94	RMSE/ES	5.45	EM	3





Parametric IUH					
Bias [l/s]	-4.80	V_{in}/V_{out}	1.20	λ	1.10
Mean Absolute Error [l/s]	16.30	RMSE	26.87	DF	0.30
NSE	0.88	RMSE/ES	6.68	EM	3




Forward Problem
2014
Jun
Jul
Aug
Sep
l/s
mm
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
days

XXIII. Fall 2014

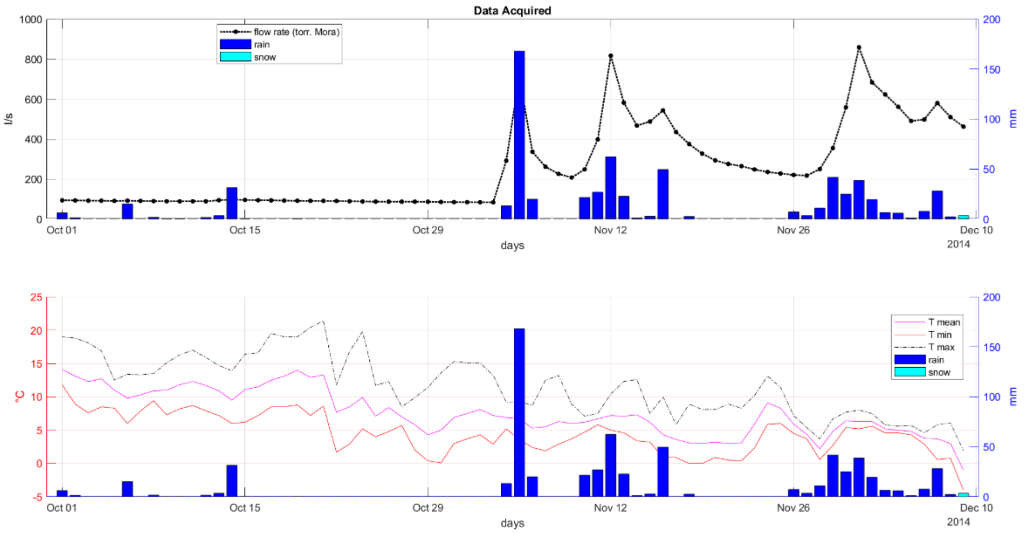
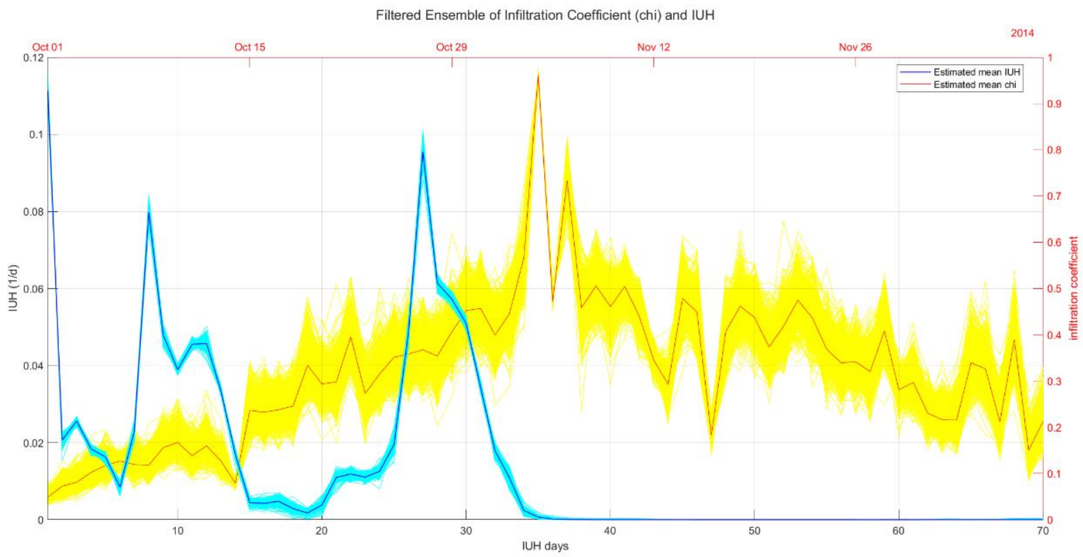
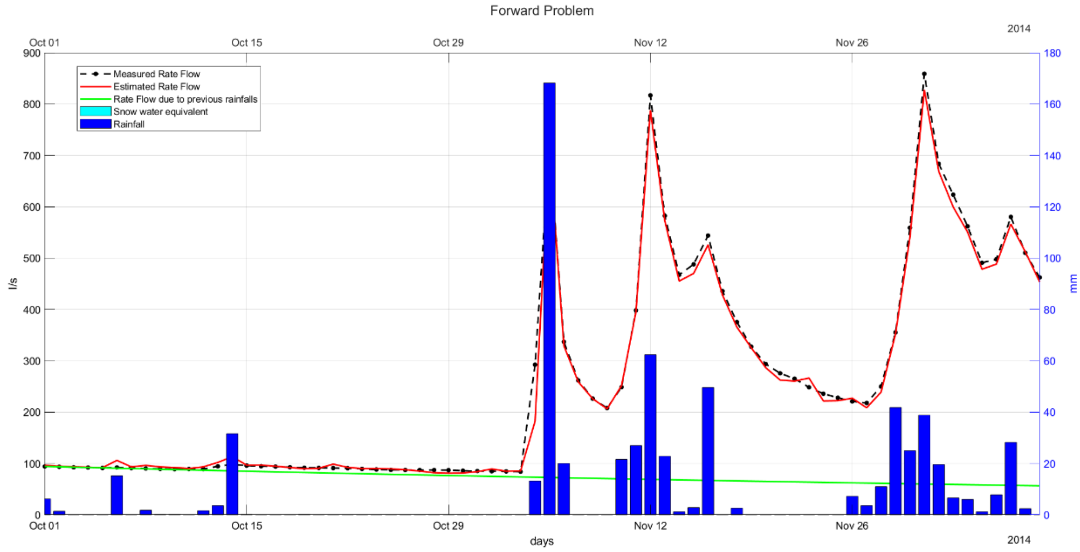


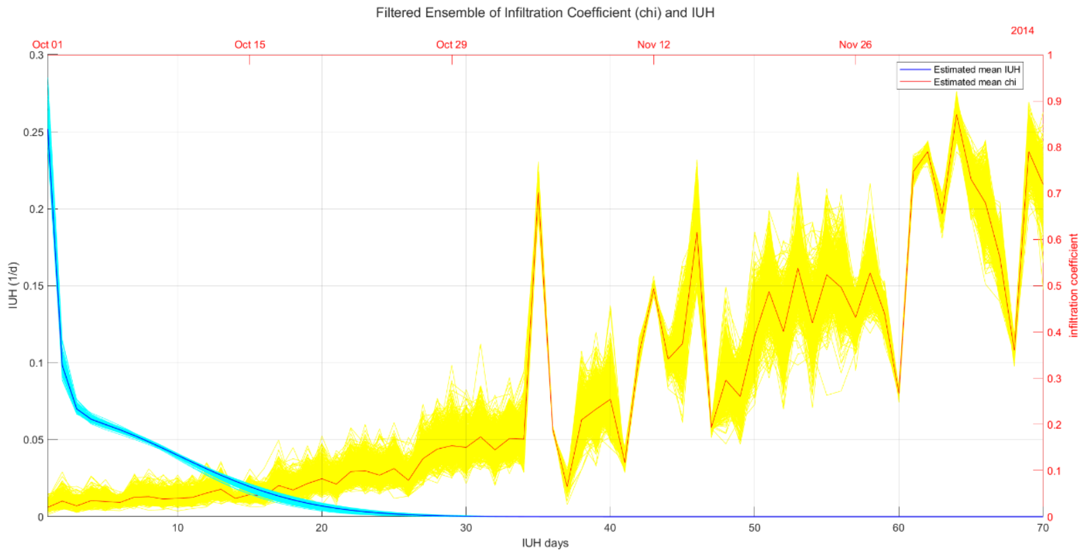
Figure 8-23:data used.

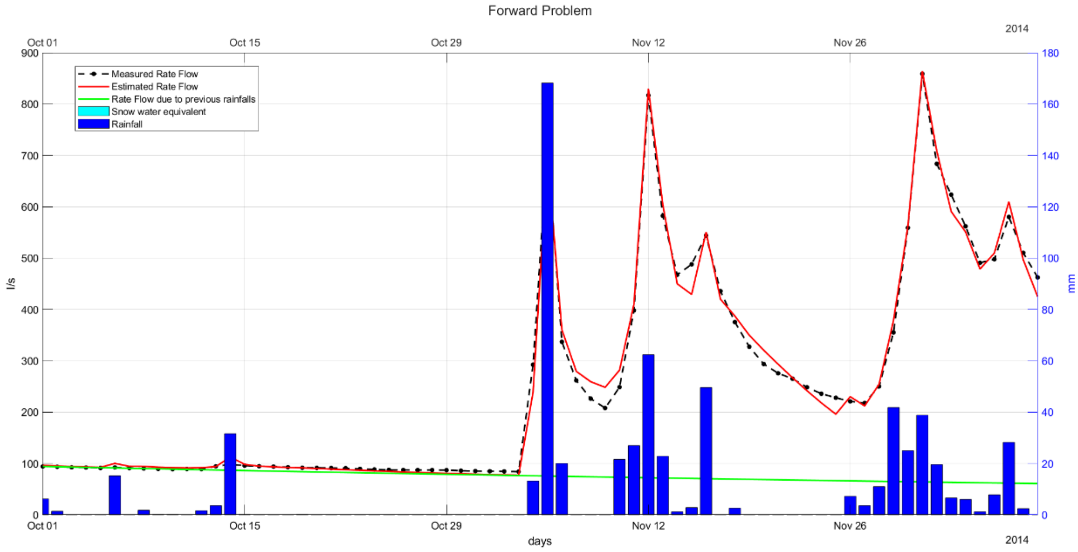
Non-parametric IUH					
Bias [l/s]	-4.80	V_{in}/V_{out}	1.27	λ	1.08
Mean Absolute Error [l/s]	8.19	RMSE	16.05	DF	0.8
NSE	0.99	RMSE/ES	6.57	EM	2





Parametric IUH					
Bias [l/s]	0.33	V_{in}/V_{out}	1.18	λ	1.10
Mean Absolute Error [l/s]	12.19	RMSE	18.97	DF	0.5
NSE	0.99	RMSE/ES	4.96	EM	3





Forward Problem
Oct 01
Oct 15
Oct 29
Nov 12
Nov 26
2014
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days

XXIV. Spring 2016

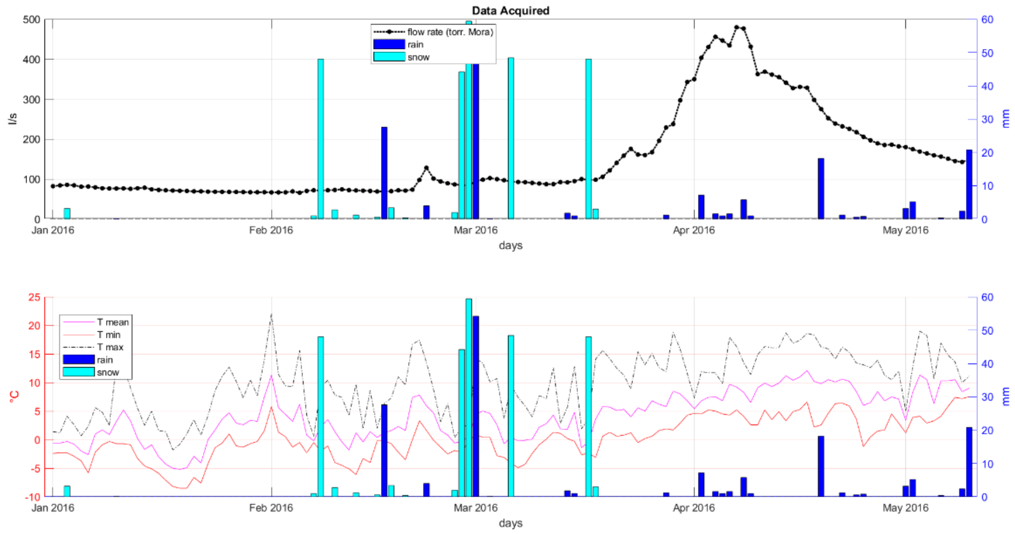
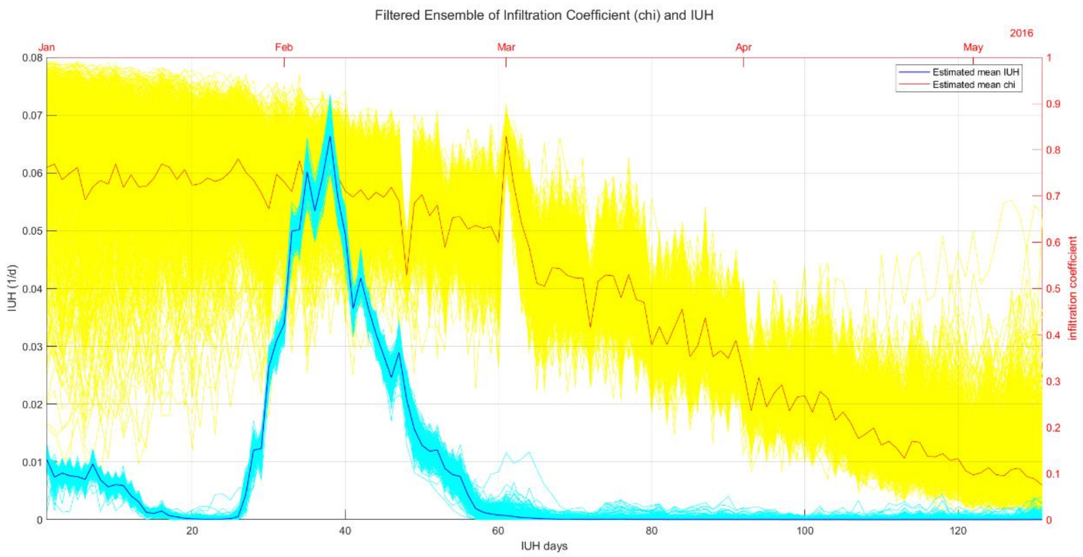
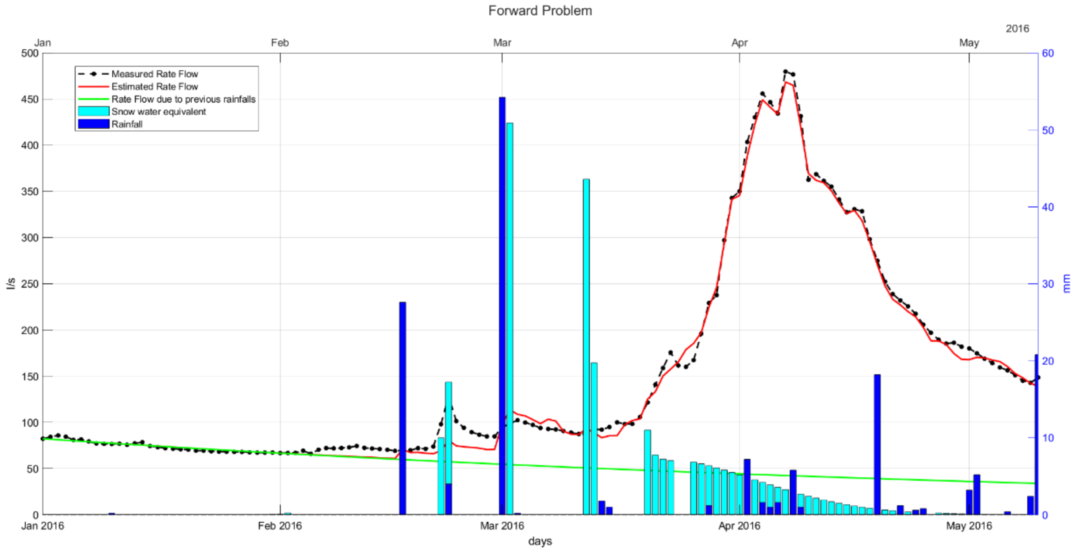


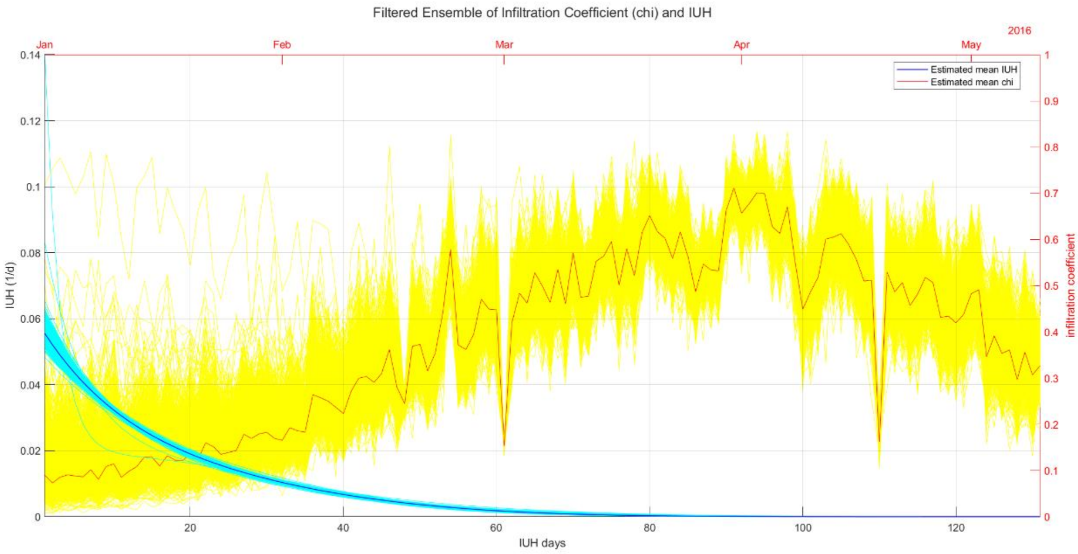
Figure 8-24:data used.

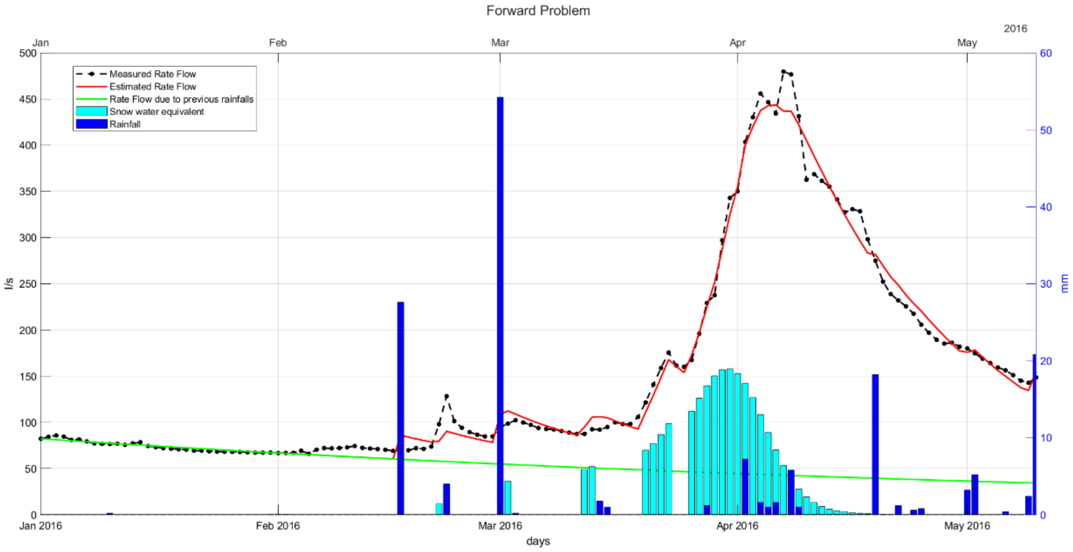
Non-parametric IUH					
Bias [l/s]	-3.23	V_{in}/V_{out}	1.12	λ	1.12
Mean Absolute Error [l/s]	6.11	RMSE	10.15	DF	0.8
NSE	0.99	RMSE/ES	2.03	EM	3





Parametric IUH					
Bias [l/s]	-0.94	V_{in}/V_{out}	1.15	λ	1.10
Mean Absolute Error [l/s]	7.65	RMSE	12.00	DF	0.6
NSE	0.99	RMSE/ES	4.34	EM	3





Forward Problem
2016
Jan
Feb
Mar
Apr
May
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
Jan 2016
Feb 2016
Mar 2016
Apr 2016
May 2016

XXV. Fall 2016

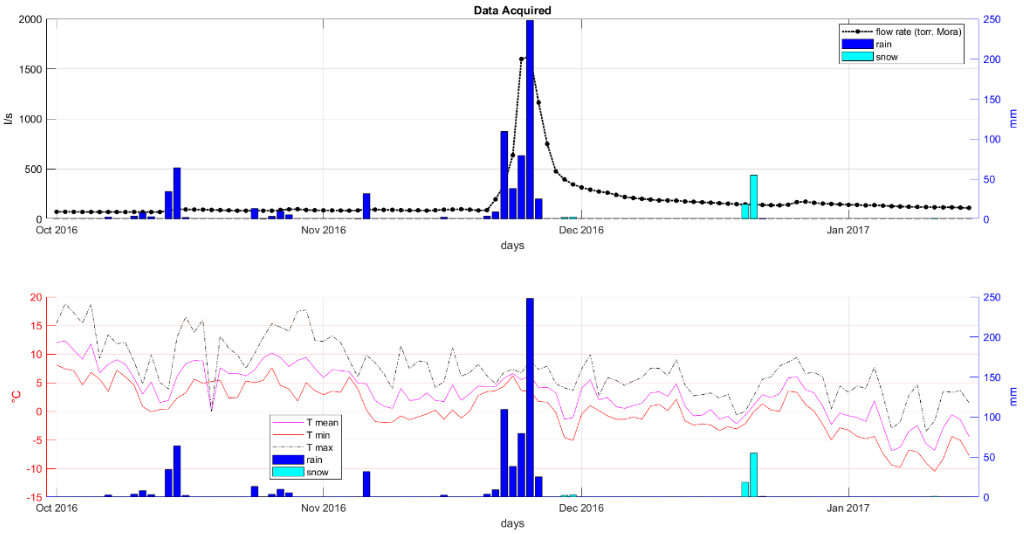
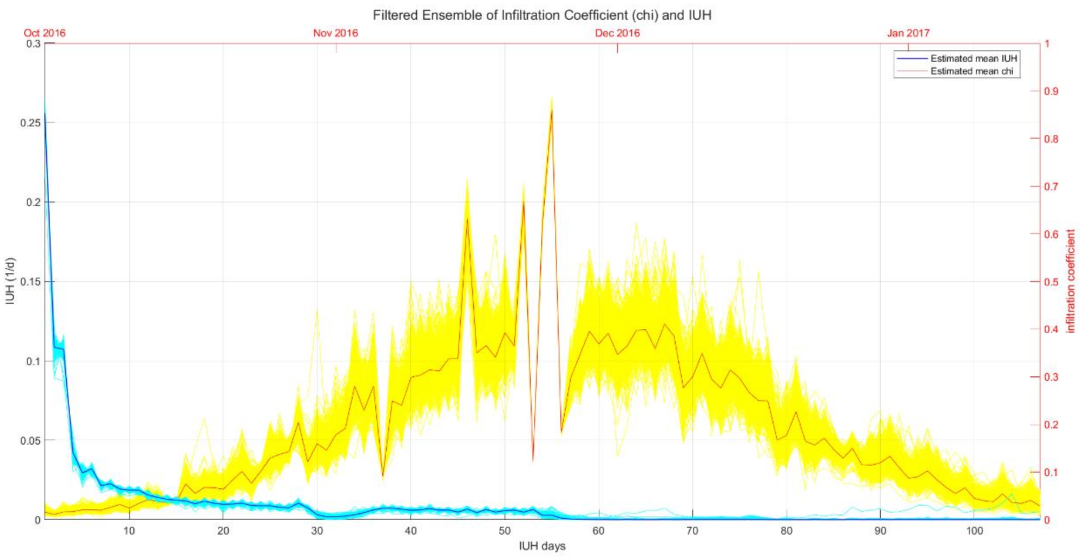
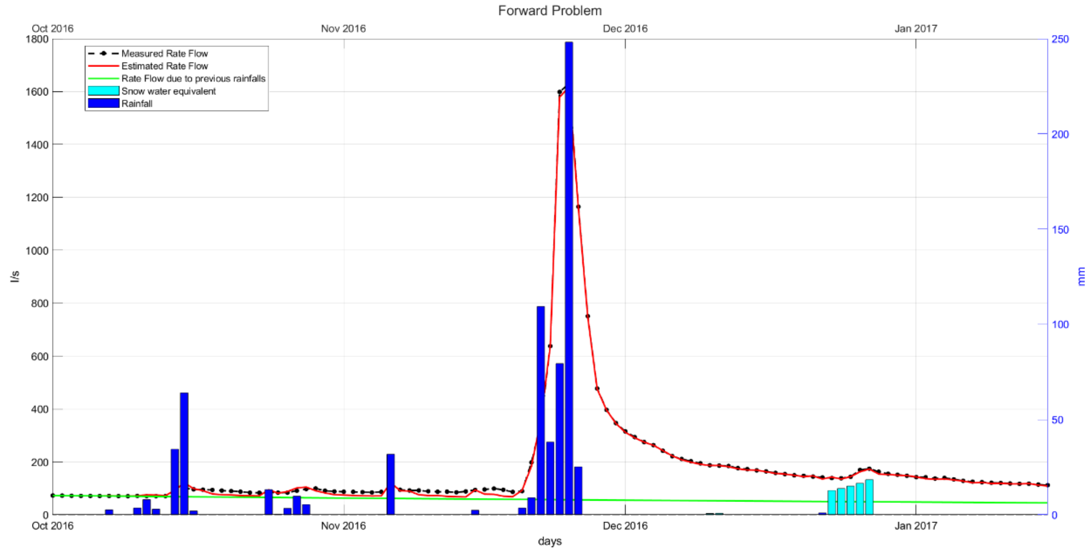


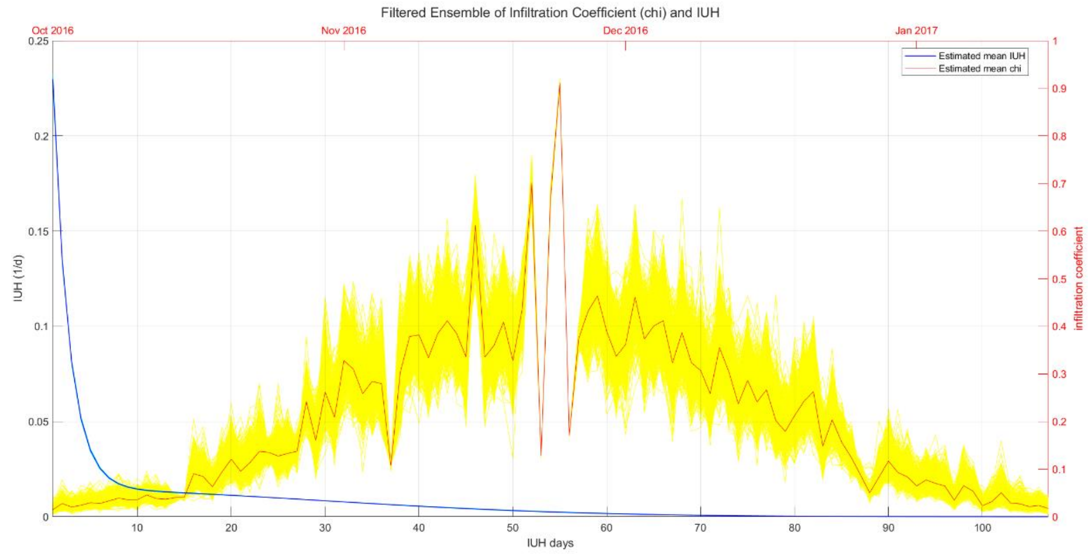
Figure 8-25:data used.

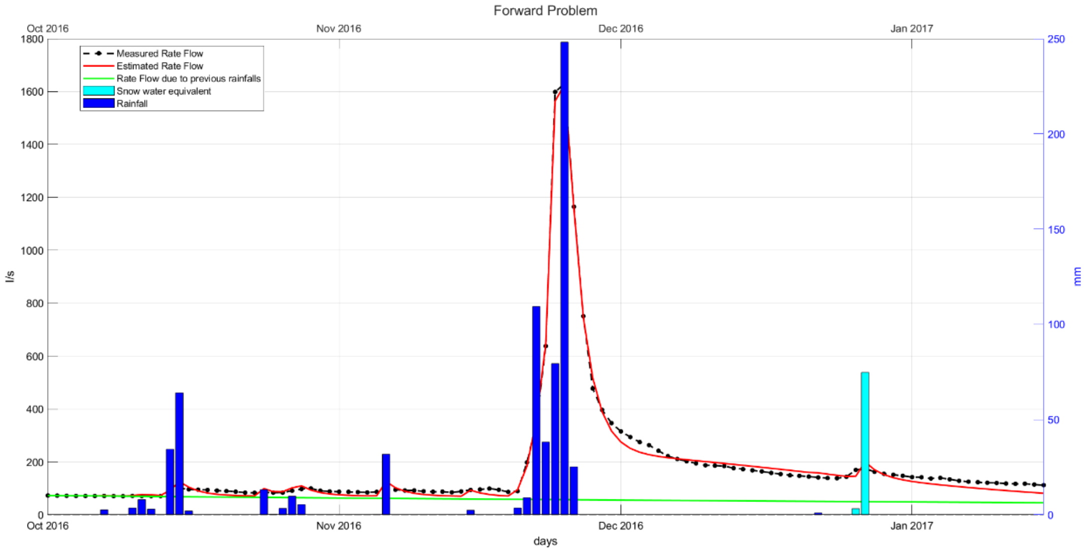
Non-parametric IUH					
Bias [l/s]	-4.03	V_{in}/V_{out}	1.02	λ	1.09
Mean Absolute Error [l/s]	6.49	RMSE	10.70	DF	0.7
NSE	1.00	RMSE/ES	4.09	EM	2





<table>
<tr><td colspan="6">Parametric IUH</td></tr>
<tr><td>$Bias$ [l/s]</td><td>-4.19</td><td>V_{in}/V_{out}</td><td>1.04</td><td>λ</td><td>1.10</td></tr>
<tr><td>Mean Absolute Error [l/s]</td><td>14.28</td><td>RMSE</td><td>17.87</td><td>DF</td><td>0.7</td></tr>
<tr><td>NSE</td><td>1.00</td><td>RMSE/ES</td><td>3.76</td><td>EM</td><td>3</td></tr>
</table>




Forward Problem
Oct 2016
Nov 2016
Dec 2016
Jan 2017
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
1800
1600
1400
1200
1000
800
600
400
200
0
250
200
150
100
50
0

XXVI. Spring 2017

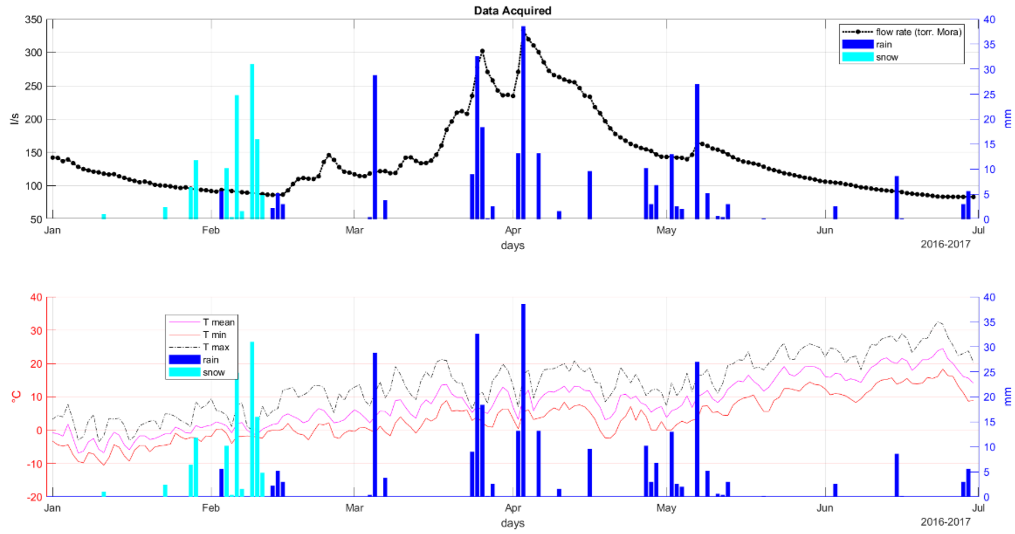
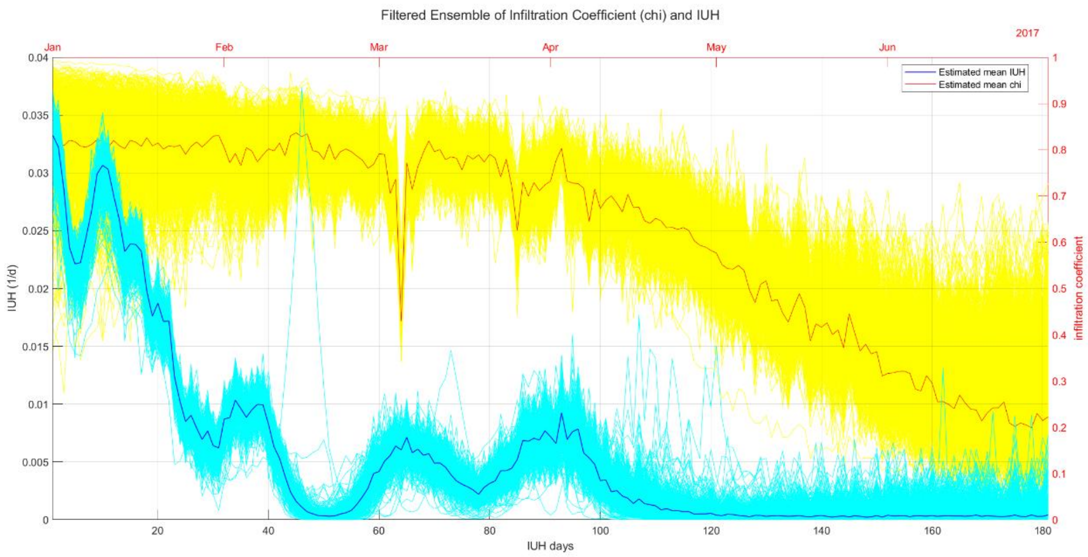
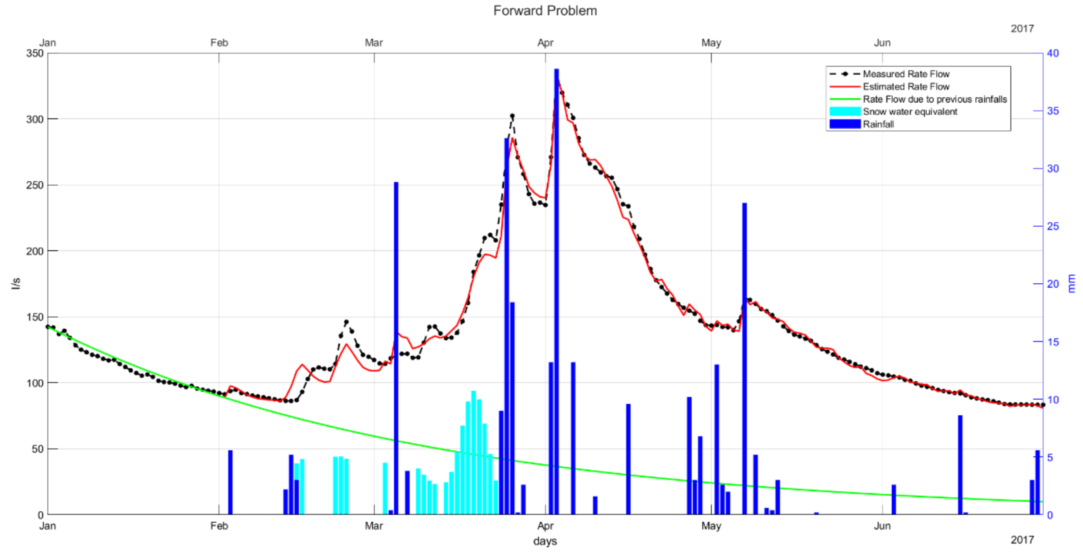


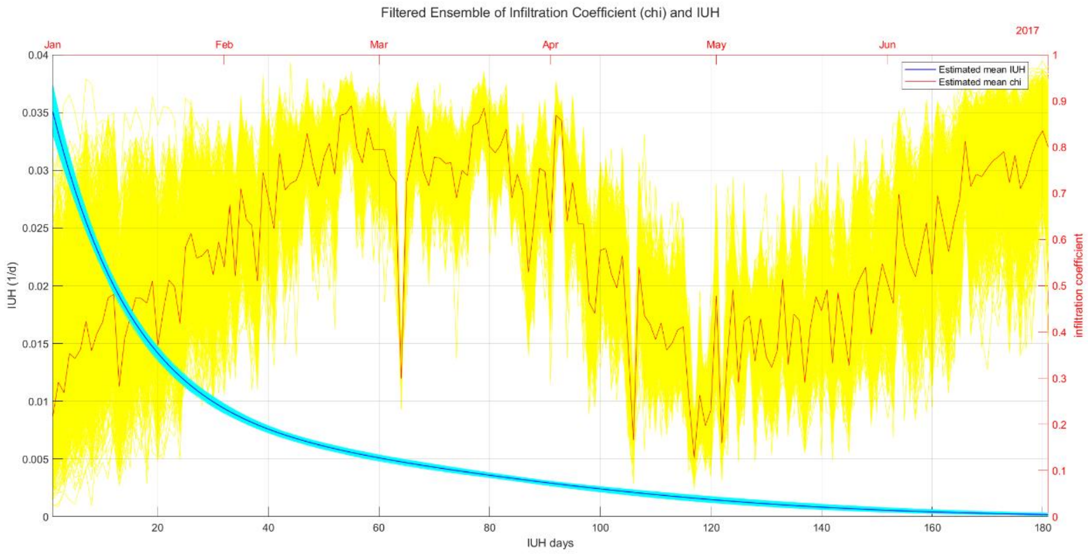
Figure 8-26:data used.

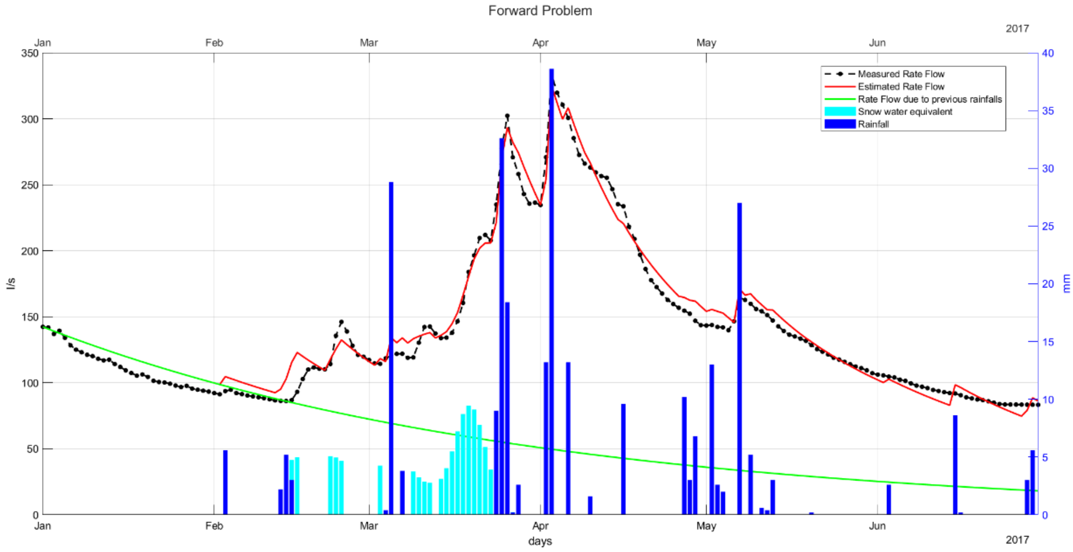
Non-parametric IUH					
Bias [l/s]	-2.03	V_{in}/V_{out}	1.11	λ	1.10
Mean Absolute Error [l/s]	4.68	RMSE	8.79	DF	0.7
NSE	0.99	RMSE/ES	2.46	EM	3





Parametric IUH					
Bias [l/s]	0.11	V_{in}/V_{out}	1.15	λ	1.12
Mean Absolute Error [l/s]	6.75	RMSE	8.48	DF	0.9
NSE	0.98	RMSE/ES	6.24	EM	3





Forward Problem
2017
Jan
Feb
Mar
Apr
May
Jun
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
l/s
mm
days
0
50
100
150
200
250
300
350
0
5
10
15
20
25
30
35
40

XXVII. Spring 2018

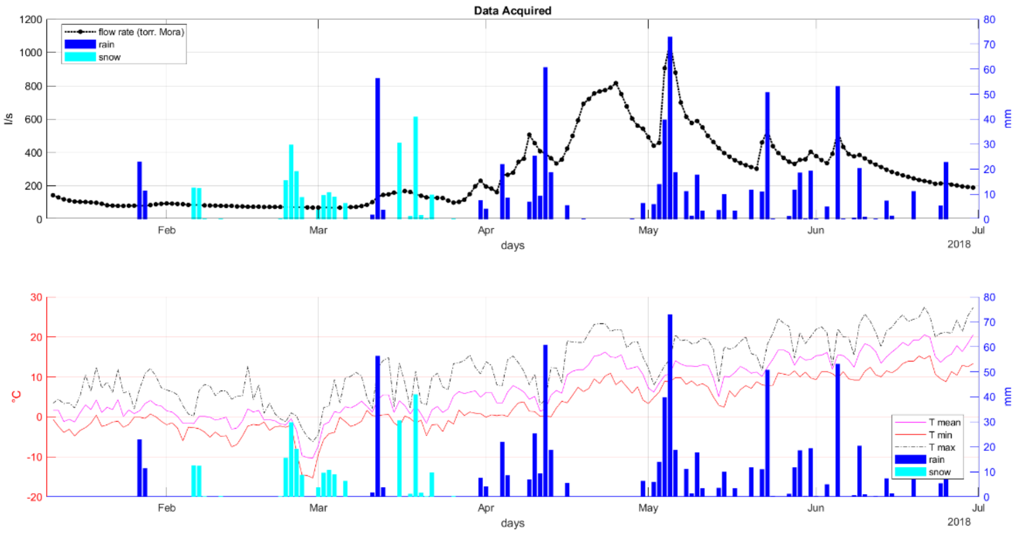
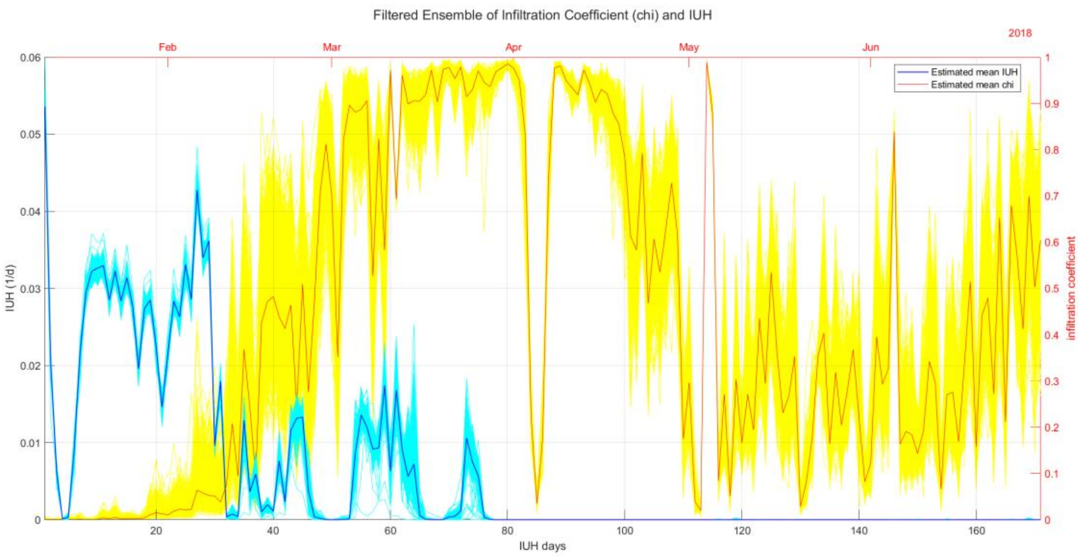
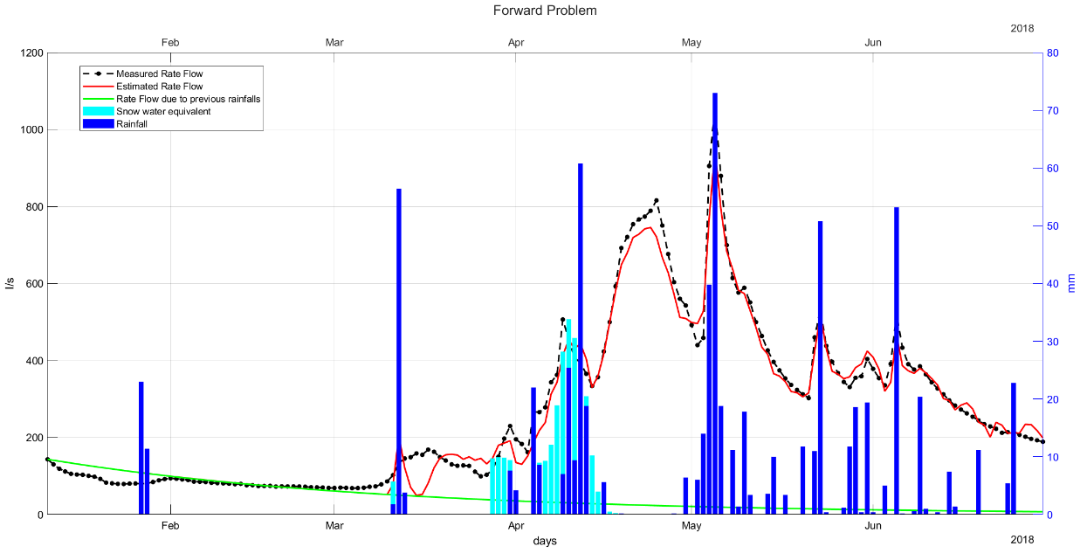


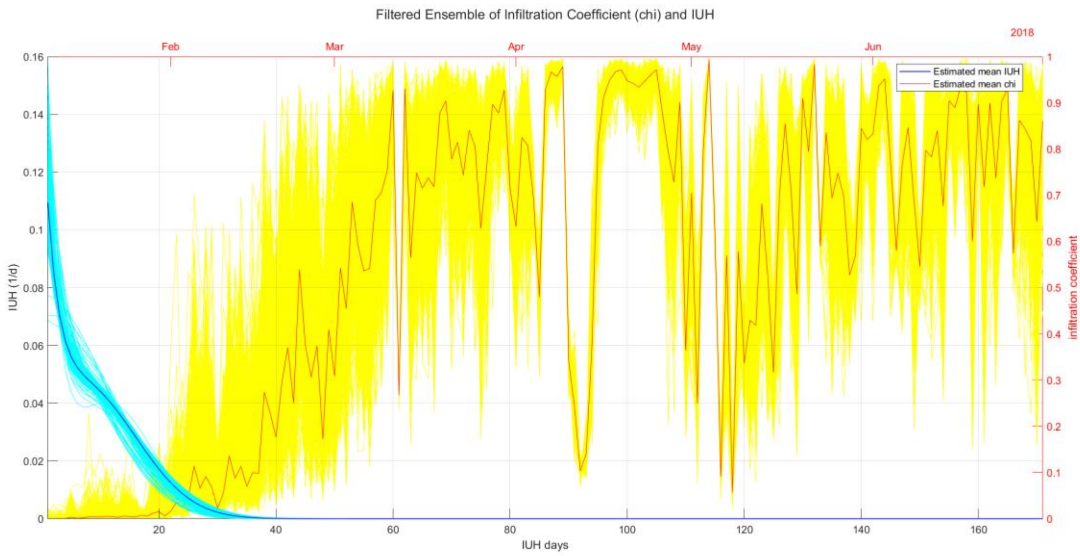
Figure 8-27:data used.

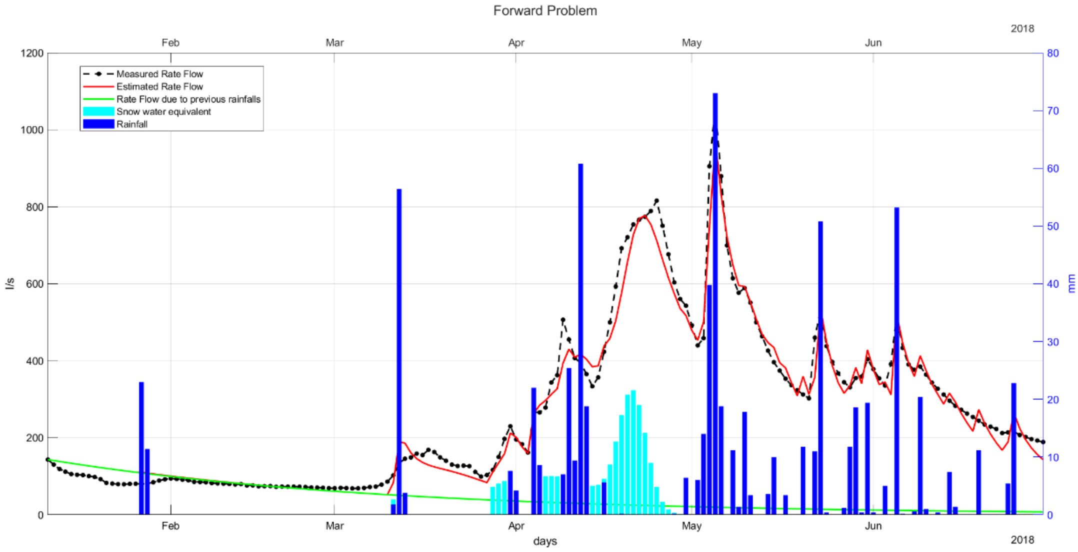
Non-parametric IUH					
Bias [l/s]	-5.70	$^{V_{in}}/_{V_{out}}$	1.10	λ	1.15
Mean Absolute Error [l/s]	25.17	RMSE	39.39	DF	0.8
NSE	0.97	RMSE/ES	6.68	EM	3





Parametric IUH					
Bias [l/s]	-4.95	V_{in}/V_{out}	1.03	λ	1.18
Mean Absolute Error [l/s]	23.59	RMSE	39.97	DF	0.5
NSE	0.98	RMSE/ES	4.62	EM	3





Forward Problem
2018
Feb
Mar
Apr
May
Jun
1200
1000
800
600
400
200
0
l/s
Measured Rate Flow
Estimated Rate Flow
Rate Flow due to previous rainfalls
Snow water equivalent
Rainfall
80
70
60
50
40
30
20
10
0
mm
days