Politecnico di Torino

Master of Science in Environmental and Land Engineering
A.Y. 2023/2024
Graduation Session October 2024

# On the water resources availability in the Dora Baltea River: a modelling experiment with the GEOframe system

Supervisor:
Stefania Tamea

Candidate:
Alessia Franco

# ABSTRACT

The *Dora Baltea River* is one of the main tributaries of the *Po River*. Its basin includes the entire *Valle d'Aosta* region, where it originates, and a small part of the *Piemonte* region. The *Dora Baltea* plays a crucial role in providing water to the downstream regions because most of its flow is regulated by many dams that collet and divert the water to meet several needs downstream. In addition, during summer, snow and glacial melting upstream contribute to the river discharge. The flow of the Dora Baltea is vital for various - often conflicting – purposes including agriculture, hydroelectric power production, tourism, sports, and civil uses. Therefore, quantifying the water availability generated in this area is fundamental to assess a sustainable water-use strategy. This thesis contributes to an international project aimed at supporting stakeholders in the development of adaptation strategies to climate change, providing information about the availability of water resources for specific uses in *Piemonte* and in the *Po River* basin.

In this work the semi-distributed hydrological modelling system GEOframe is used to model the *Dora Baltea Basin* and extract the discharge arriving in *Piemonte* in a period of 27 years (1996-2023). The entire model has been built to be run with a daily time resolution in order minimize the computational demand while maintaining good modelling results.

The goal is to evaluate the water reaching the main withdrawal points, such as the onset of *Canale Farini* - an important in-take structure designed to increase and regulate the water flow available in the *Canale Cavour*, made for irrigation purposes of the fertile lands located in *Piemonte.*

Since a gauging station is necessary for model calibration and none exists at the *Canale Farini* confluence, the entire basin has been closed about 5 kms downstream, near *Verolengo*. At the end, to estimate the discharge reaching *Canale Farini*, a post-processing evaluation has been conducted to quantify the amount of water diverted.

Six calibrations have been carried out - four upstream and two downstream - to improve the simulated discharge. Each calibration has been done twice: first with one round of 100 iterations, then with three rounds of 200 iterations. In the end, the values of the objective function – Kling-Gupta efficiency (KGE)- are higher compared to the initial round of calibrations, demonstrating significant improvements through the calibration process. Considering all six calibrations, an average KGE range between 0.60 and 0.80 has been achieved, with the lower values in upstream which are among the most challenging to model considering all the snow melting related processes. Furthermore, the sum of all observed and simulated discharge values is compared, yielding reasonable results, with $\sum Q_{obs}/\sum Q_{sim} = 1.02$ at the basin outlet. The overall result, including all the hydrological processes, are acceptable since, comparing the results with studies that used finer sub-basin discretization, the difference in KGE is about 0.05-0.10. The study demonstrates that it is possible to reduce the number of sub-basins while still achieving good model accuracy and maintaining low computational demand. The model simulates all natural flows, but it could be improved in the future by incorporating anthropogenic changes aimed to flow regulation, such as the dams in the *Valle d'Aosta* region which play a crucial role in the hydrological management of the territory.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1 INTRODUCTION

Nowadays, having a clear and detailed idea of the hydrological cycle and all the related water processes is fundamental to access the water availability. Water is a fundamental source, and it must satisfy all the necessary requests related, mainly, to agriculture, energy production, industrial development, and civil purposes. With changing time and climate, the access to water is becoming critical and it is one of the most important topics in the environmental science. One of the Sustainable Development Goals promoted by the United Nations is related to the clean water and sanitation: it highlights the necessity of a better management of our water in order to be able to have a sustainable use of the source without compromise anything and anyone [46].
In this scenario, developing tools and methodologies useful for a better knowledge and evaluation of the water resources of rivers is becoming increasingly important. In particular, reproducing the natural water flow of rivers could be one of the first step to perform for the drafting of new modern management plans which give the possibility of an equal and sustainable water use. For that reason, hydrological models are considered as efficient instruments able to simulate and forecast the water availability. They can surely help analyse the hydrological problems and the related tasks, and thus provide guidance for the planning and management of water resource (Chow et al., 1988 [47]).

The present project has been developed with the aim of reproducing the discharge timeseries of the *Dora Baltea River*, a tributary of the *Po River*, using the hydrological physically based, semi-distributed model system GEOframe. It has been built with the objective of decrease the computational demand while producing satisfactory results. The study contributes to an international project called LIFE CLIMAX PO (CLIMate Adaptation for the PO river basin district) which "*aims at a climate-smart management of water resources in the Po River basin*" [48].
The model simulates the natural flow in the Dora Baltea catchment during a period of 27 years (1996-2023) in order to have a better understanding of the natural hydrological processes of the territory. The study area uses the water resources for many and – often – competing purposes, going from the hydropower production – since the territory is full of dams – to irrigation – since the river contributes to feed the canal network used for agricultural uses, but it is also important for civil uses, sports and tourism. This basin is an alpine one and it is characterized by high discharges values during Spring due to the snow melting processes whereas precipitation is mostly abundant during Autumn. Moreover, the territory is subject of exceptional floods which have caused historical flooding mainly in the plain area along the main Dora Baltea River course. Indeed, a sustainable and better management of the water resource is needed to preserve the natural flow of the river and also give an equal access to water for all the necessities.

The system GEOframe is an open-source, semi-distributed, component-based hydrological modelling system developed in Java and based on the environmental modelling framework Object Modelling System V3 (OMS3) [6]. It is subdivided into a sequence of components simulating different physical processes of the hydrological cycle

which can be joined together to obtain the discharge values and the hydrological response of the entire basin starting from the available meteorological data. The basin is divided into Hydrological Response Units (HRUs) that are used to identify the meteorological forcing data, the radiation budget, and the evapotranspiration. Any of these components are used to solve the final water balance and connecting all the single units to have the final response of the entire basin.
The water balance analysis is performed in the ultimate step with a calibration of a set of parameters in correspondence of the selected sections equipped with gauging stations. The procedure gives the optimal values to give at each HRU and to obtain the best results by the model. The evaluation of the accuracy of the model is mostly made over the simulated discharge timeseries which are compared with the observed ones. GEOframe reproduces the natural flow of the Dora Baltea basin providing results that can be used to evaluate the water availability in the territory and manage its use better.

The thesis is divided into four main chapters, dedicated to all the step and the results of the modelling procedure. The first one presents the study area, and the first step regarding the geomorphology performed in the GEOframe analysis; the second part is related to meteorological data and their interpolation; the third chapter presents the main part of the GEOframe methodology giving information about the estimation procedure of radiation, evapotranspiration and finally, the discharge obtained using the Embedded Reservoir Model integrated into the software. In conclusion, the last chapter shows the results of all the simulations with the comparison between the observed and simulated discharge timeseries giving a quantitative interpretation of the accuracy of the model by using error metrics. Additionally, a post-processing analysis is presented to quantify the water diverted by the *Canale Farini* – used for irrigation provisioning - located 5 km above the basin outlet.

# 2 DORA BALTEA BASIN

In this chapter the *Dora Baltea Basin* is presented in order to give a clear image of the case study. The section is divided into three parts: the first one describes the basin and gives the geographical context of the *Dora Baltea*, including the basin outlet and "*Canale Farini*" which is a section of interest of the final analysis. The second part describes in details and gives information regarding the inputs of the geomorphological analysis performed by GEOframe as first step. The third and last part explains the structure of the first phase of the modelling: how it is carried out step by step, including the settings used and, at the same time, the results obtained are showed and interpreted.

## 2.1 Geographical context

*Dora Baltea River*, located in North-West Italy, is a left-hand tributary of the *Po River*. It rises in the high *Mont Blanc* after two small rivers start flowing together, namely, the *Dora di Ferret*-being fed by the *Pré de Bar Glacier* in *Val Ferret* - and the *Dora di Veny*-fed instead by the *Miage Glacier* and *Brenva Glacier* in *Val Veny*. Dora Baltea crosses the entire *Valle d'Aosta* region and then it enters into *Piemonte* where it reaches the *Po River* at *Crescentino*, a little downstream from *Chivasso* (metadata reported in *Table 2.1*) [12]. *Figure 2.1* shows the main course of the Dora Baltea.

*Table 2.1, metadata of the confluence of the Dora Baltea [12]*

| UTMX [m] | UTMY [m] | Latitude [°] | Longitude [°] | Elevation (m.a.s.l.) [m] |
|---|---|---|---|---|
| 425270.647 | 5003373.209 | 45.1799 | 8.0489 | 210 |

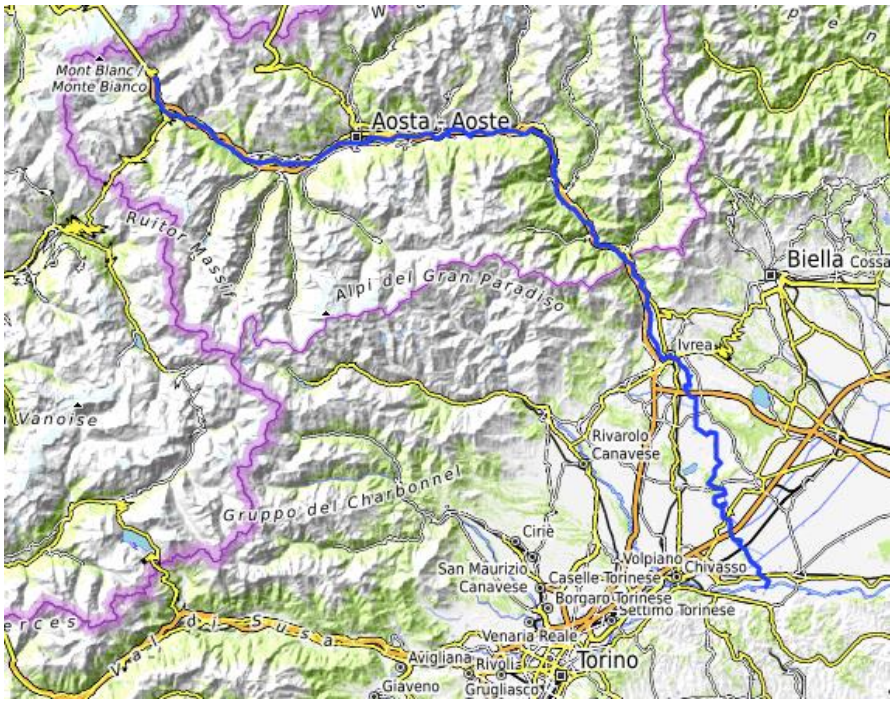



*Figure 2.1, map of the main course of the Dora Baltea. Credit: OpenStreetMap [13]*

In the *Valle d'Aosta* part, *Dora Baltea* counts with many tributaries coming from the right and left side. From the right side all tributaries arrive from the *Gran Paradiso* massif. Instead, from the left part they arrive from *Mont Cervin* and *Mont Rose*. Entering *Piemonte,* the main tributary enters from the right side and is the *Chiusella*. *Figure 2.2* shows the main tributaries of the river. The basin has an area of 3930 km$^2$ of which 90% is in mountainous regions. From the *Mont Blanc* confluence to the *Po River*, it has a length of 152 km [11].

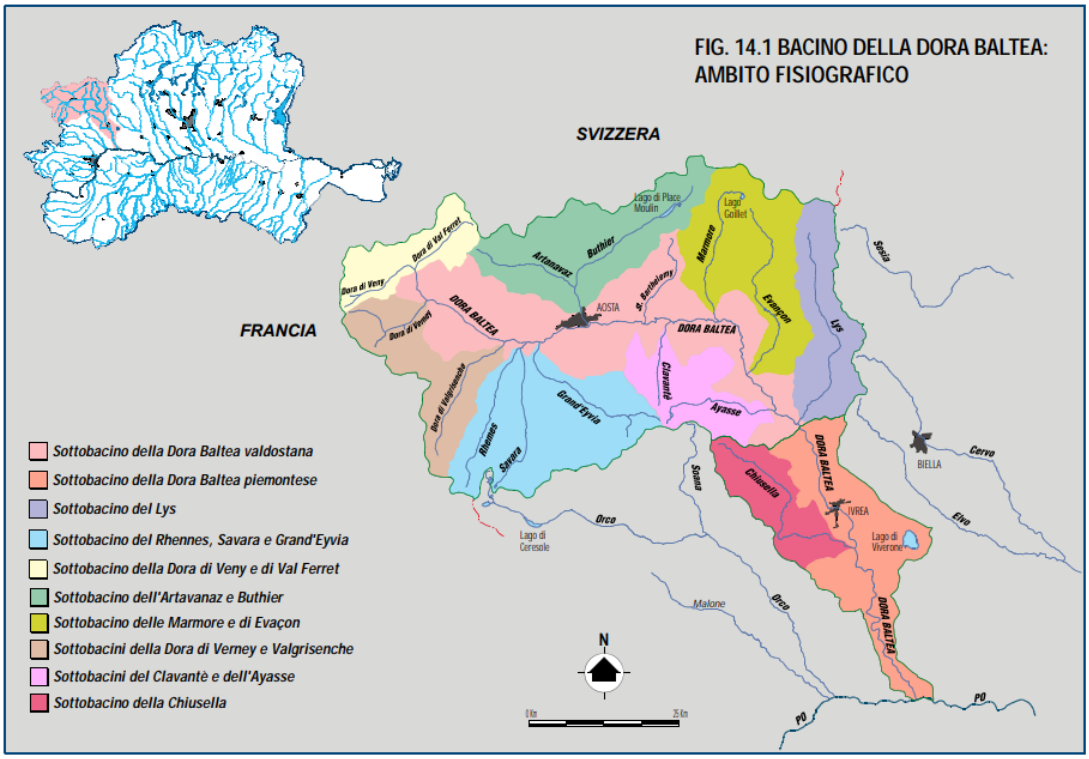



*Figure 2.2, main tributaries of the Dora Baltea Basin. Credit: AdBPo [11]*

The *Dora Baltea Basin* is clearly an alpine basin, the mountain gives protection from the humid air coming from the Atlantic Ocean. For that reason, precipitation is modest in both terms of total values and intensity. The highest flow is from the end of the Spring and the start of the Autumn when the snow precipitation is extremely low. The flat area of the basin has the role of lamination of the flow during this period, the highest values are measured from the gauging station located in the *Chiusella* tributary. [11] *Figure 2.3* shows the elevation histogram obtained from the analysis of the Digital Elevation Model (DEM) of the basin during the geomorphological analysis. The plot is useful to understand the distribution of heights in the basin area. It is clear that the elevation ranges are quite large, going from 139 m.a.s.l. to 4792 m.a.s.l., proving that the *Dora Baltea Basin* is an alpine basin. The most frequent elevation is 208 m.a.s.l. and belongs to the central flat area of Valle d'Aosta, where the main course of the river is located, and to the entire part of Piemonte that is quite flat. Another peak is around 2400 m.a.s.l., underlying the presence of high elevation mountains in all the region considered. The average elevation extracted is 1862 m.a.s.l.

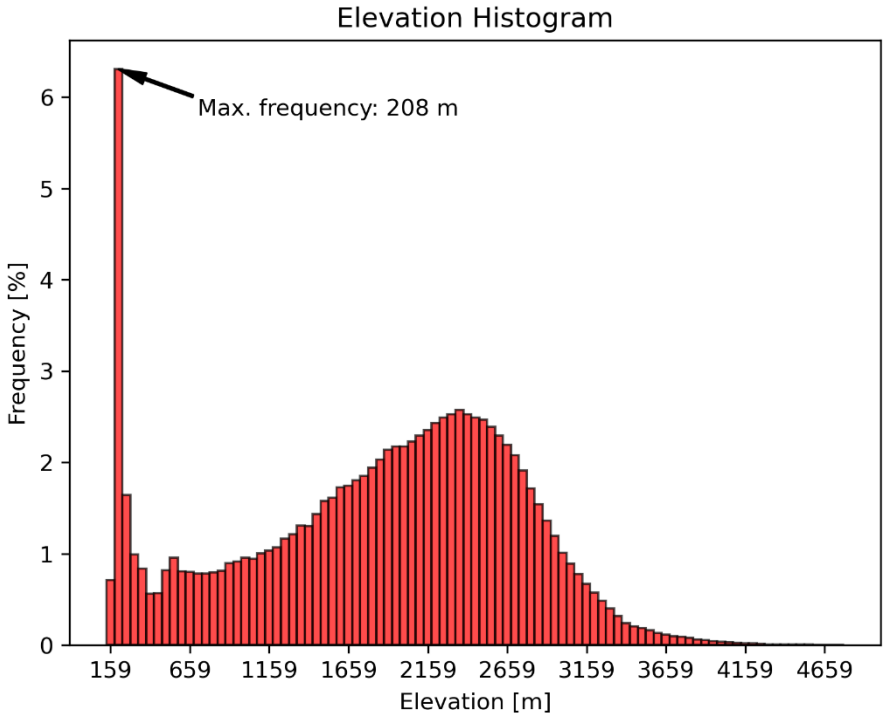



*Figure 2.3, elevation histogram of the Dora Baltea Basin*

In the last part of the catchment, in Piemonte, a lake is located. It is called *Lago di Viverone,* and it is not linked directly with the Dora Baltea River. According to "*Autorità di Bacino del Fiume Po*" [11], the small river "*Roggia Violana*" is a left tributary of the lake, but it is not detected by the maps. Probably, the lake could give some contribution to the groundwater flow, but there are not clear investigations. During the geomorphology analysis the network given as input – extracted by Open Street Map [3] - does not take into consideration the small river and, as consequence, there are no links between the Dora Baltea and the Lago di Viverone. For that reason, the lake has not been considered in the analysis, and it was left for further and future developments.

### 2.1.1 "Canale Farini"

The *Dora Baltea* flow arriving in the flat area of *Piemonte* is largely used as source for different purposes - irrigation for the countryside, hydropower production, tourism, sports, and civil uses. For agricultural purposes, the water is used till the *Sesia River* at East and till the *Po River* at South. In that area of Piemonte there is a vast network of canals made with the purpose of support agriculture by providing water for irrigation [14]. *Figure 2.4* shows a part of the canal network including the part of the case study, briefly described ahead.

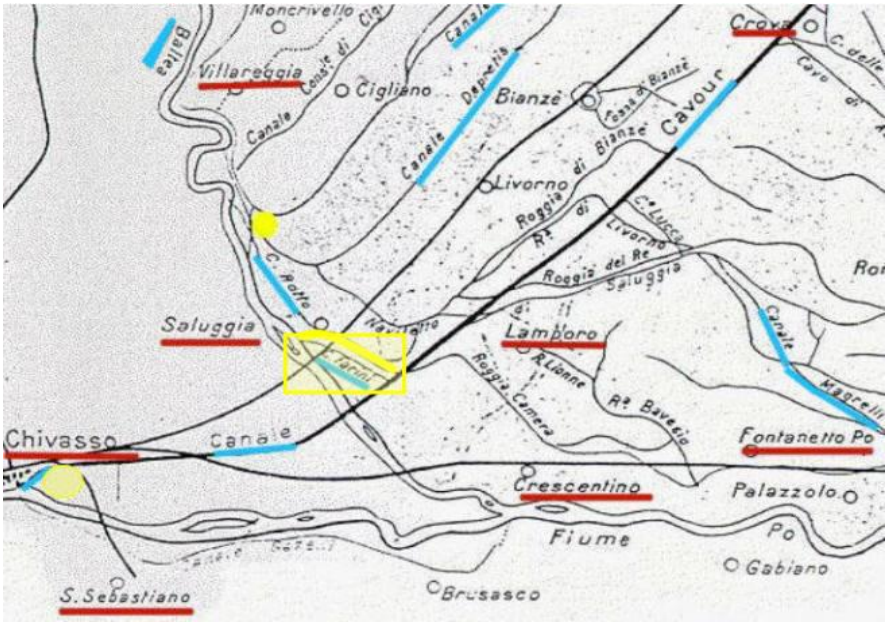



*Figure 2.4, canal network used for irrigation in Northern Italy. Credit: roberto-crosio.net [14]*

The *Canale Farini* is a subsidiary of the *Canale Cavour* which connects the *Po River* with the *Ticino River* and plays a crucial role in distributing water to the nearby agricultural areas. The *Canale Cavour* have a lot of branches forming together a large network that supports the thriving agricultural industry in Northern Italy, particularly rice cultivation. One of these subsidiaries is the *Canale Farini* that connects the Dora Baltea with the *Canale Cavour* to maintain the discharge always constant, even during summer. It has a length of only 3 km, but it has a large section able to carry an average discharge of 70 $m^3/s$ [15]. The canal intake structure is huge as *Figure 2.5* shows. It was made to increase the flow in the Canale Cavour and to solve the problem of an unequal distribution of irrigation water among all farmers during the entire year. All these canals help in the efficient management of water resources for irrigation, ensuring that the fertile lands in the region receive adequate irrigation [14].

*Figure 2.5, Canale Farini intake structure. Credit: Andrea Miola [15]*

At the confluence between the Dora Baltea and the *Canale Farini*, no gauging station has been installed, so the discharge is not measured in a precise way. The nearest hydrometer is on the Dora Baltea just 5 km below the confluence nearby *Verolengo. Figure 2.6* shows the map of the region of interest highlighting the Canale Farini, the Po-Dora Baltea confluence and the only gauging station available nearby.

In addition, measuring the quantity of the flow of *Canale Farini* is quite important to have a clear idea of the availability of water coming from the *Dora Baltea River* and used to irrigate a huge area of cultivations. The hydrological model constructed in the study case gives the possibility to quantify the discharge arriving from the Dora Baltea Basin to the downstream territory. A post-processing evaluation has been used to take into consideration the water diverted by *Canale Farini*.

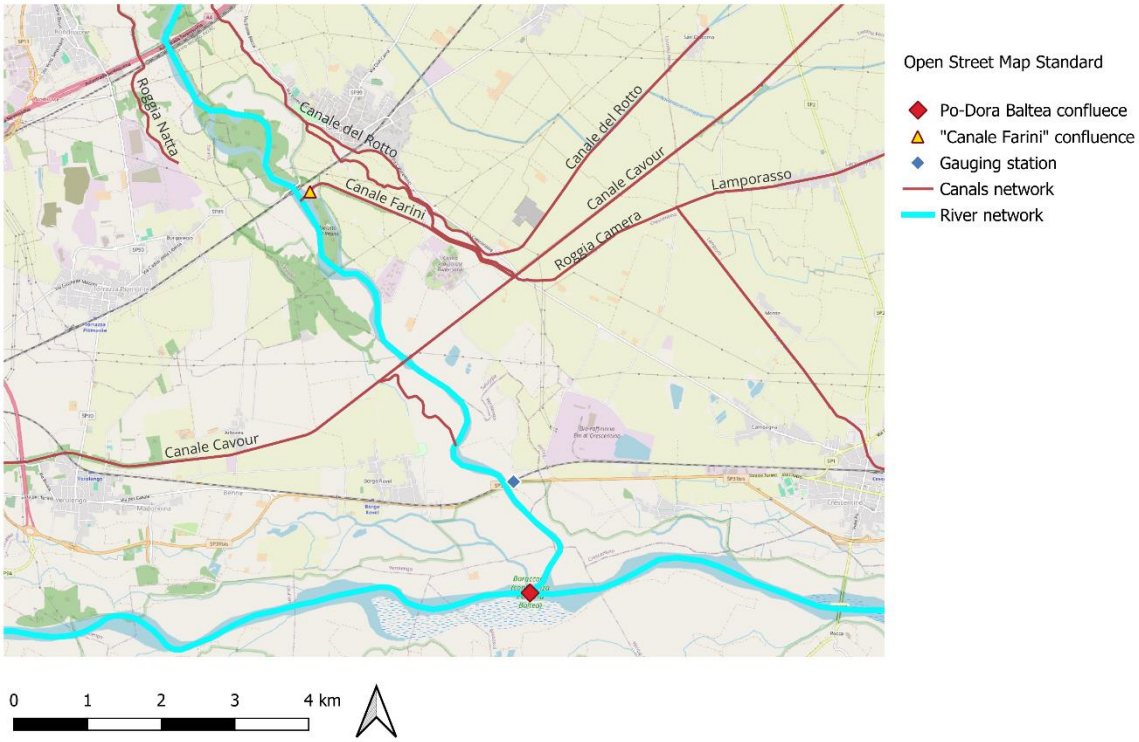



*Figure 2.6, map of the location of the Canale Farini and the gauging station at Verolengo*

## 2.2 GEOframe component: geomorphological analysis

The first step of GEOframe modelling is the geomorphological analysis. It is the start of the entire study, and, for that reason, it is an especially important phase that must be done carefully. All the subsequent steps are dependent on it, so it is important to perform it meticulously. In the two following paragraphs, the inputs, the structure, and the results are presented.

### 2.2.1 Inputs

The first input that is necessary is the Digital Elevation Model (DEM) of the area containing the basin of interest. The Dora Baltea basin (see 2.1) is located in the North-West of Italy, so that part of the whole Italian territory was downloaded by *Tinitaly*, "the dataset is about the elevation of the bare ground, and thus it can be called DTM /Digital Terrain Model"[1] [1]. The DEM was downloaded and then cropped to have only the region of interest of the *Dora Baltea* basin. The DEM used as input to the software includes all the *Valle d'Aosta* region and only a small part of the North-West *Piemonte* as shown in *Figure 2.7*. The resolution of the DEM was reduced from 10 m to 20 m in order to minimize the file size and make the computational demand much more efficient.

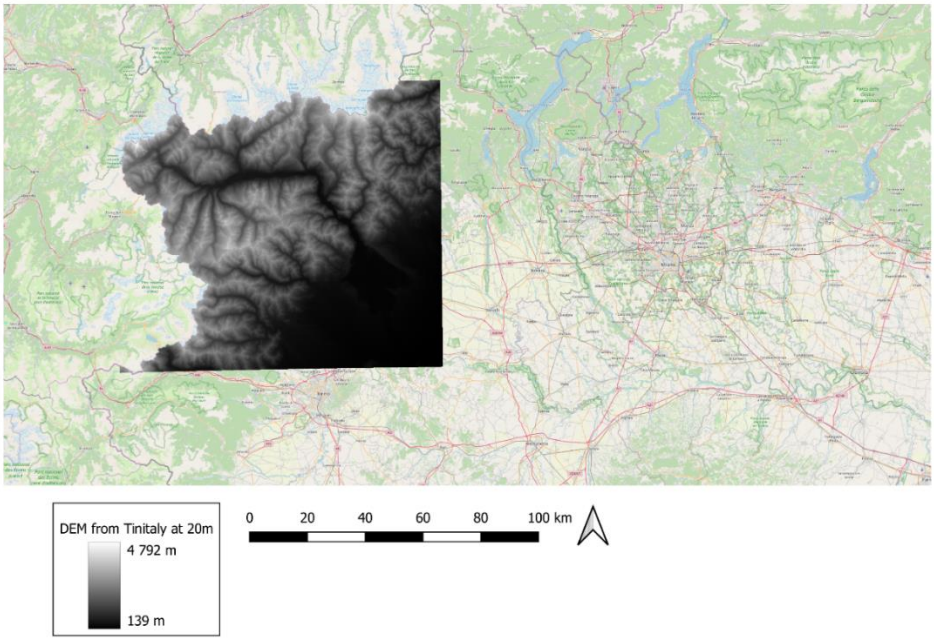



*Figure 2.7, DEM of Valle d'Aosta and a part of Piemonte downloaded from Tinitaly, resolution reduced at 20 m.*

[1] "*It is a seamless digital elevation model (DEM) of the whole Italian territory, presented in 2007. The dataset is about the elevation of the bare ground, and thus it can be called DTM /Digital Terrain Model). This DEM was obtained starting from separate DEMs of single administration region of Italy. The DEM is freely available as a 10m-cell size grid (in GeoTIFF format), in the UTM WGS 84 zone 32 projection system*" [1]

The part of the DEM used includes the basin of interest and a small part of the territory around it to make GEOframe able to extract the basin area in the right way once the basin outlet has been defined. If only the DEM of the basin of interest is used, GEOframe cannot extract the area in the right way. As *Figure 2.8* [2] shows, the DEM covers all the Dora Baltea basin area and the *Orco, Stura di Lanzo and Malone* basin and a small part of the *Sesia, Ceronda, Dora Riparia and Po* basin, but it is only the start input that must be cut during the geomorphological analysis steps.

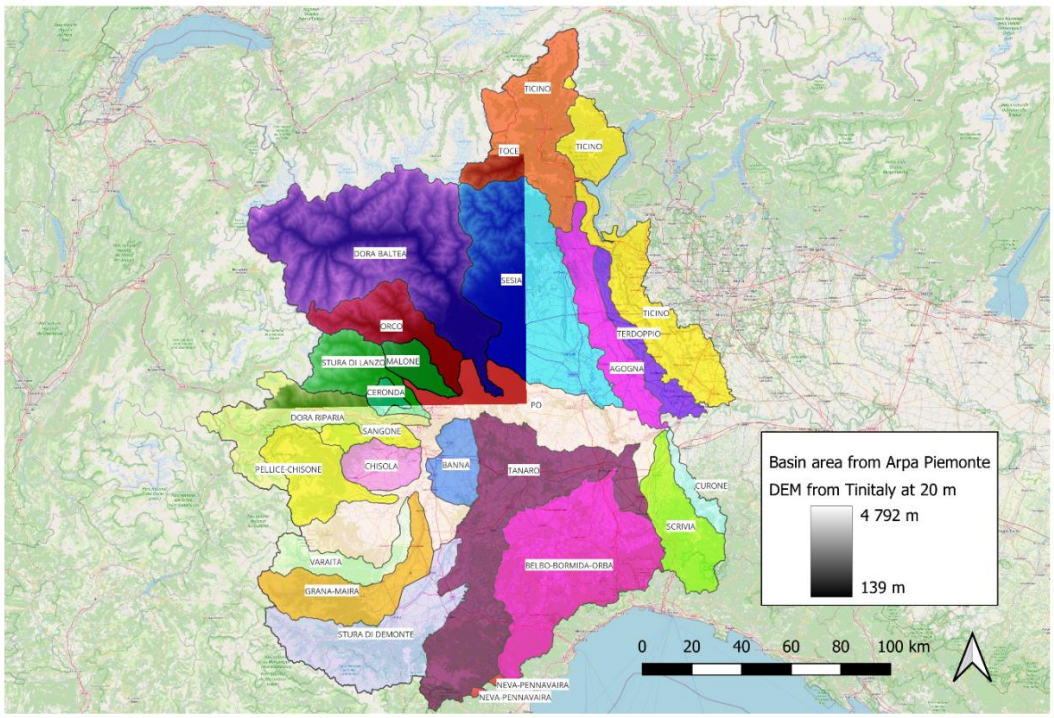



*Figure 2.8, DEM from Tinitaly and Piemonte basins of the most important rivers. Credit: ARPA Piemonte [2]*

The second input of this phase is optional. It is the river network available from the National Authorities that GEOframe uses in the first step of the geomorphological analysis: *carver*. It is not mandatory for the modelling, but it is useful to make the computation more accurate. This step takes the DEM and the shape file of the river network, and it excavates below the river network of a certain amount[2]. This will force the drainage direction to follow the path of the official river network.

The river network was extracted from the *QuickOSM* plugin on the software QGIS (version 3.30.0-'s-Hertogenbosch) [3]. *Figure 2.9* shows the river network used, it is not detailed, and it has very few branches, but it is just a way to give the software a real reference of the river network that has to be extracted in the subsequent steps.

[2] In the study case the "depth lines" parameter referring to the excavation depth has been set at 5.0 as default

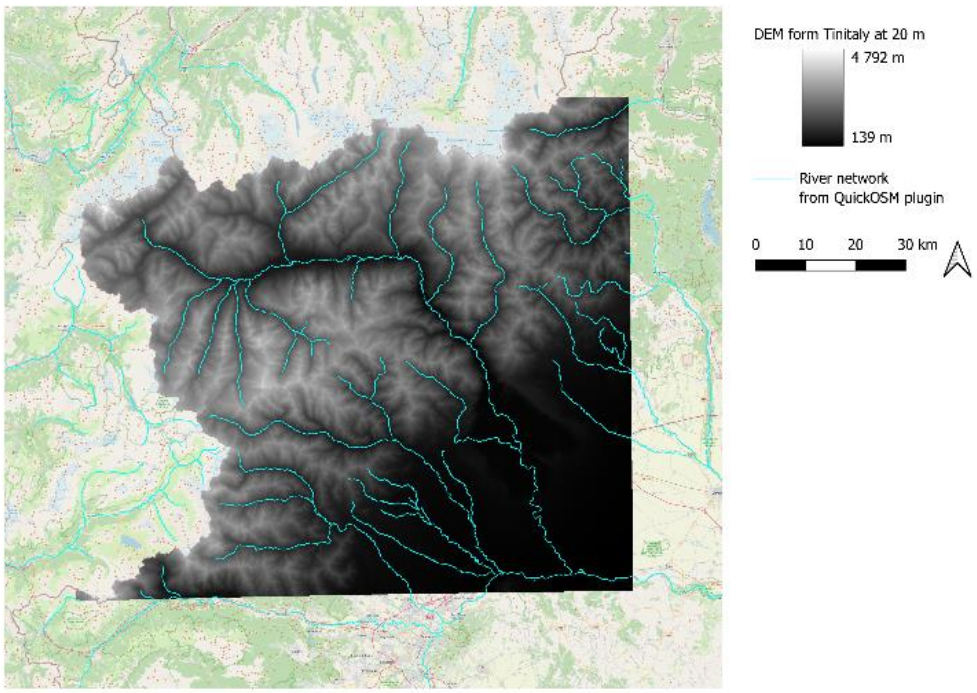



*Figure 2.9, river network extracted from QuickOSM plugin [3]*

The third and last input are the gauging stations data available for the Dora Baltea basin. The station data, including the geographical position of the stations and the timeseries of the discharge, have been downloaded from Regional Environmental Protection Agencies. Concerning the *Valle d'Aosta* region which occupies the largest part of the territory considered, "*Centro Funzionale Regione Autonoma Valle d'Aosta*" has been used [4]. Instead, the data for the *Piemonte* part have been downloaded from "ARPA Piemonte" [5]. Only the stations located inside the Dora Baltea basin area are taken into consideration since the remaining part of the DEM is not part of the study case.

Regarding the geomorphological phase, the only information necessary is the geographical location of the gauging stations. GEOframe requires as input a shape file containing in the attribute table the coordinates, the elevation, and the ID of each point. The coordinates available from ARPA are latitude and longitude expressed with the reference system WGS84, but for completeness the coordinates have been also projected in the UTM zone 32 reference system since it is the one used in the project[3].

In *Figure 2.10* the geographical position of the gauging stations available for the Dora Baltea basin, are shown with the correspondent ID. It is evident that only the stations in the basin area are considered. The basin outlet is remarked with a red star, and it is located in the correspondence of the station n. 3 (located at *Verolengo – Torino*). Furthermore, the "Canale Farini" confluence is shown as the yellow triangle, and it is located only 5 km above the outlet. For the purpose of our analysis the station n. 3 has been considered

[3] WGS84/UTM zone 32 reference system uses the ellipsoidal model WGS84 to define locations and UTM projection of the point (expressed in meters) for precise mapping in the specified geographic zone UTM 32 that covers the Northern Italy, Switzerland, Germany and Denmark [3].

a good compromise to evaluate in post-processing the discharge arriving at “Canale Farini” since no gauging stations are available or installed exactly at the confluence. For that reason, the basin has been closed there.

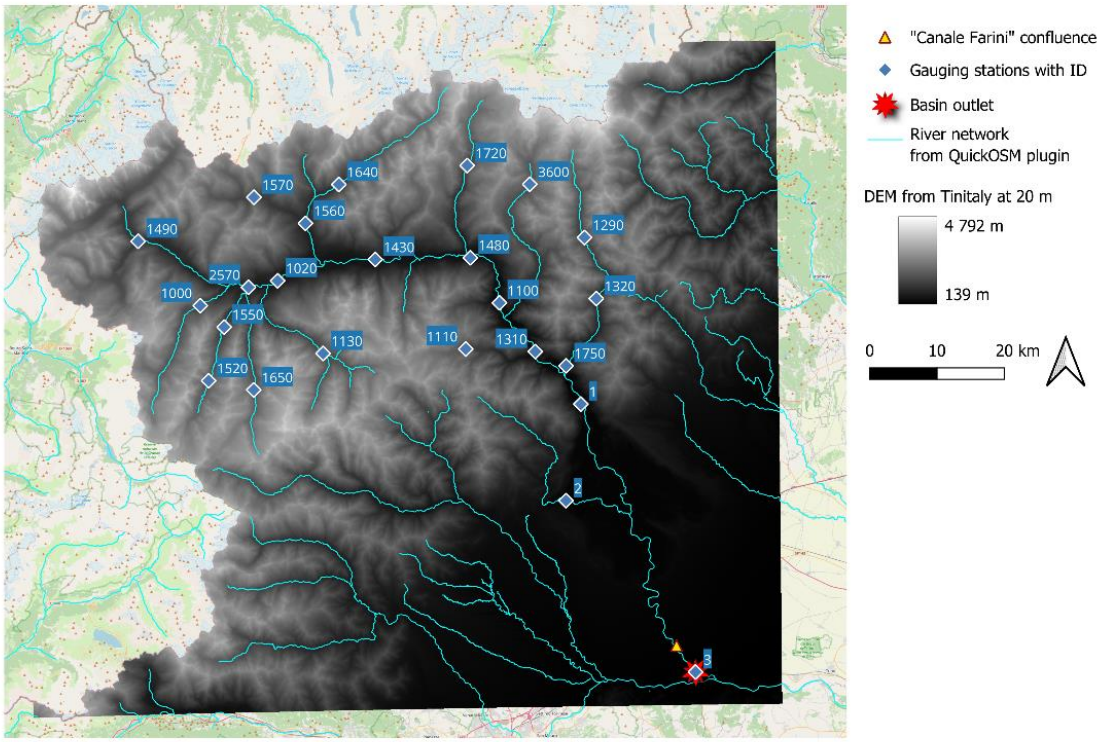



*Figure 2.10, geographical position of the gauging stations with ID label*

In *Table 2.2* all the hydrometer stations metadata are shown, including the coordinates WGS84 and UTM 32, the elevation, the ID, and the complete name.

The ID number seems to have no correlation with the geographical position, but maybe it is correlated to the installation date. The number are in ascending order, but it is important to look at their position on the map as *Figure 2.10* shows.

The station that has the highest elevation (m.a.s.l.) is the n. 1520 - *Rhemes-Notre-Dame – Chanavey* with 1690 m.a.s.l., instead the lowest one is the n. 3 – *Verolengo – Dora Baltea* with 167 m.a.s.l. chosen as the basin outlet.

*Table 2.2, gauging stations metadata*

| Station Name | Station ID | Region | UTMX [m] | UTMY [m] | Latitude [°] | Longitude [°] | Elevation (m.a.s.l.) [m] |
|---|---|---|---|---|---|---|---|
| Arvier - Chamençon | 1000 | VdA | 352607 | 5060047 | 45.6783 | 7.10742 | 1238 |
| Aymavilles - Ponte Dora Baltea | 1020 | VdA | 363990 | 5063401 | 45.7108 | 7.25257 | 618 |
| Champdepraz - Ponte Dora Baltea | 1100 | VdA | 396715 | 5059550 | 45.6818 | 7.67371 | 370 |
| Champorcher - Chardonney | 1110 | VdA | 391539 | 5052836 | 45.6206 | 7.60875 | 1430 |
| Cogne - Crétaz | 1130 | VdA | 370610 | 5052557 | 45.6145 | 7.34048 | 1470 |
| Gressoney-Saint-Jean - Capoluogo | 1290 | VdA | 409219 | 5068878 | 45.7675 | 7.83249 | 1373 |
| Hone - Ponte Dora Baltea | 1310 | VdA | 401759 | 5052289 | 45.6172 | 7.73993 | 340 |
| Issime - Capoluogo | 1320 | VdA | 410936 | 5059941 | 45.6873 | 7.85621 | 960 |
| Nus - Les Iles | 1430 | VdA | 378497 | 5066224 | 45.7389 | 7.43818 | 534 |
| Pontey - Ponte Dora Baltea | 1480 | VdA | 392442 | 5066223 | 45.7412 | 7.61737 | 473 |
| Pré-Saint-Didier - Capoluogo | 1490 | VdA | 343759 | 5069668 | 45.7629 | 6.99078 | 996 |
| Rhemes-Notre-Dame - Chanavey | 1520 | VdA | 353591 | 5048964 | 45.5788 | 7.12338 | 1690 |
| Rhemes-Saint-Georges - Capoluogo | 1550 | VdA | 356051 | 5056788 | 45.6497 | 7.15259 | 1179 |
| Roisan - Moulin | 1560 | VdA | 368333 | 5071722 | 45.7865 | 7.30608 | 745 |
| Saint-Oyen - Moulin | 1570 | VdA | 360850 | 5075709 | 45.8209 | 7.20871 | 1310 |
| Valpelline - Prelé | 1640 | VdA | 373341 | 5077309 | 45.8377 | 7.36902 | 1093 |
| Valsavarenche - Eaux-Rousses | 1650 | VdA | 360190 | 5047479 | 45.5668 | 7.20835 | 1651 |
| Valtournenche - Maen | 1720 | VdA | 392190 | 5079719 | 45.8626 | 7.61112 | 1310 |
| Pont-Sant-Martin - Monte Lys | 1750 | VdA | 406290 | 5050120 | 45.5983 | 7.79844 | 340 |
| Villeneuve - Arbonne | 2570 | VdA | 359750 | 5062539 | 45.7022 | 7.19838 | 660 |
| Ayas - Champoluc | 3600 | VdA | 401233 | 5076702 | 45.8368 | 7.72821 | 1566 |
| Tavagnasco – Dora Baltea | 1 | Piemonte | 408410 | 5044505 | 45.54806 | 7.82667 | 270 |
| Parella Chiusella | 2 | Piemonte | 405963 | 5030311 | 45.42 | 7.79806 | 260 |
| Verolengo Dora Baltea | 3 | Piemonte | 424678 | 5004720 | 45.19194 | 8.04111 | 167 |

In this phase, also the discharge timeseries have been downloaded from the same websites of the Regional Environmental Protection Agencies. At this stage of the project, the data are not an input since only the geomorphological analysis is done. The data are an important source for the last part of the modelling since during the calibration step the discharge data simulated are compared and adapt with the measured data from the hydrometers considered. It is important to have a clear and whole idea of the availability of the data, so *Figure 2.11* shows how the data measured from all the stations are distributed from 1996 to 2023. Almost all stations start to have significative data from 2002 and some of them have some gaps due probably to instrument detection issues (1110 - Champorcher – Chardonney is a clear example). Despite that, simulation period was considered between 1996 and 2023 since the temperature and precipitation data availability is much larger and there are many stations start having data from 1996. Since the modelling period was chosen based on weather stations, the start date was established in 1996.

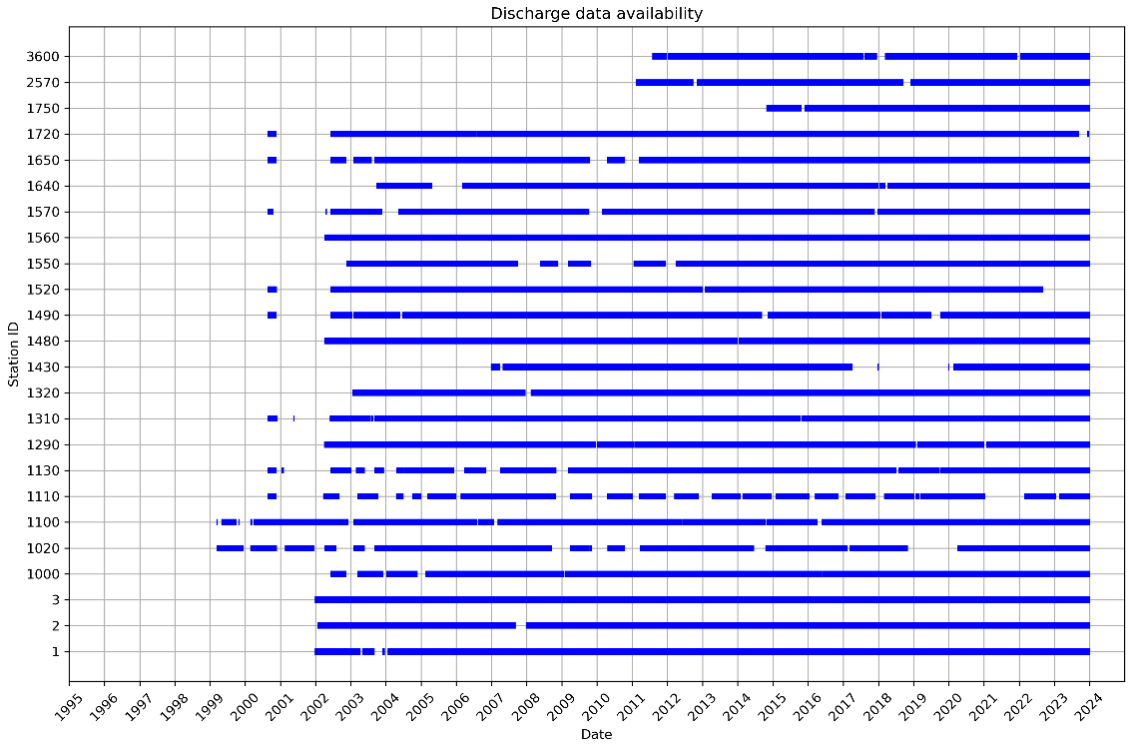



*Figure 2.11, discharge data availability for each station*

## 2.2.2 Structure and results

The aim of the geomorphological analysis is to divide the basin into different sub-basins, which are hydrological response units (HRUs). This part of the modelling is divided into 16 steps able to identify all the topological, hydrological, and spatial information. All the steps are linked together, the outputs are generally *raster* file that must be analysed with a GIS tool[4]. The structure can be divided into two parts: the first part is related to the largest area containing the basin of interest, so the entire DEM downloaded is used; the

[4] The software QGIS (version 3.30.0-'s-Hertogenbosch) has been used [3]

second part is related to the extraction of the area of the basin of interest and the definition of the HRUs.

The first part includes 8 subsequent steps of manipulation of the original DEM given as first input. *Figure 2.12* shows the flowchart of the GIS procedure used to crop the hydrographic basin (adapt from *Bancheri, Rigon, and Manfreda 2020* [6]).

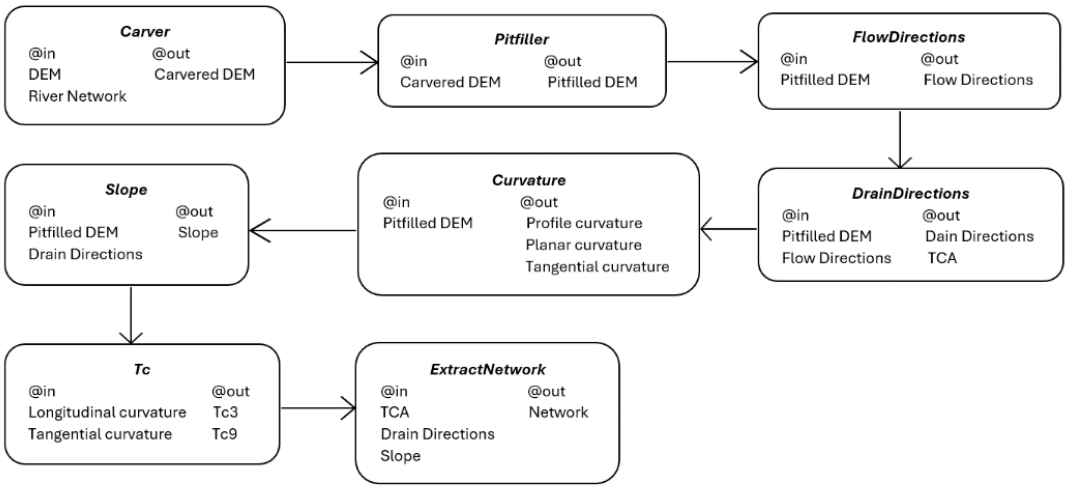



*Figure 2.12, first part of the flowchart of the GIS procedure*

All these steps are briefly described below [7]:

### ***Carver***

It takes the original DEM as input and the shape file of the river network downloaded and it excavates below the river network of a certain amount. The value is set in the code as the *pDepthLines*, for the Dora Basin the default number equal to 5.0 has been used. The output is a raster file containing the DEM with the part excavates below the river path, in this way the drainage is forced to follow the original path of the official river network.

### ***Pitfiller***

It reads the *carvered* DEM and it eliminates all the depressions that can form in the grid. Usually, these sinks are not real, but they are computational errors formed during the creation of the DEM. Using this procedure, all the possible sinks are filled, and, from a mathematical point of view, the derivative exists, and it is continuous in every single point of the DEM. *Figure 2.13* shows the output which is a raster file containing the *pitfilled* DEM showing the elevation distribution of the region. From brown to green, it shows the decreasing elevation. The highest elevation is 4792.4 m.a.s.l. and it corresponds to the *Mont Blanc* in the North-West part of the territory. The lowest elevation is 139.4 m.a.s.l. located in the South-East area in *Piemonte* region.

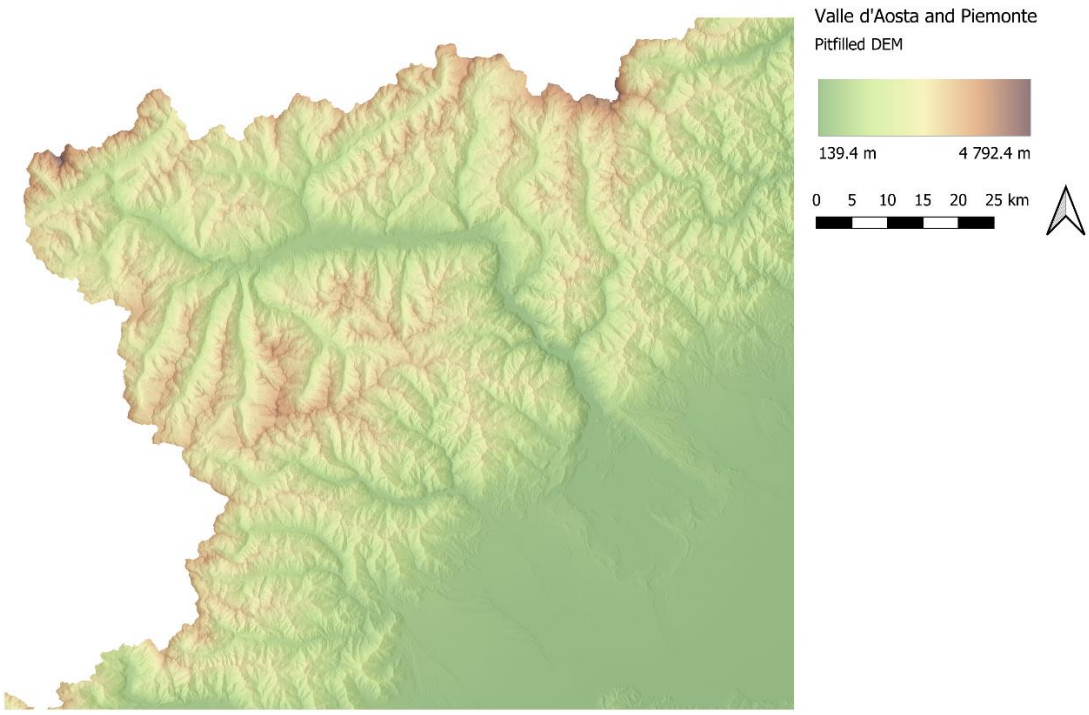



*Figure 2.13, entire pitfilled DEM*

## ***Flow Directions***

It takes the *pitfilled* DEM as input and it gives the direction of the flow according to the topology. The output is a categorical map, where each pixel has a value that goes from 1 to 8 and each number is a codification of the direction of the water flowing. Since the terrain is represented with cells forming together a squared grid, each cell of the DTM can drain just into one of the eight cells available in the surrounding. *Figure 2.14* represents the eight possible directions.

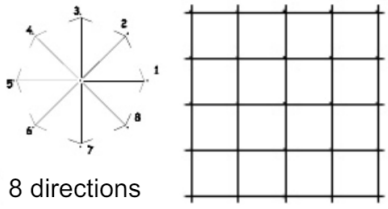



*Figure 2.14, eight possible drainage directions (Formetta and Rigon, GWS 2024 [7])*

## ***Drain Directions***

The method used in the previous step is just a simplification since in the real-world water does not flow only in 8 directions. For that reason, there are some mathematical methods correcting this issue and make the model much more real. An example is the *Orlandini* method used by GEOframe (*Orlandini et al.,2003 method* [9]). The *DrainDirections* code takes the *pitfilled* DEM and the *flow directions* obtained before as input and produces a

new raster file with drain direction corrected. *Figure 2.15* shows the 8 corrected drain directions. So, possible effects of deviation of the flow are minimized. This step gives also another output that is the *Total Contributing Area*. It is a cell map of the contributing area draining into a point of the basin. The raster output file gives the possibility to see the main course of the river network, but this is mostly important for the second part of the GIS procedure where the basin is extracted. The TCA of the Dora Baltea basin is reported *Figure A1.2 in Appendix A1*.

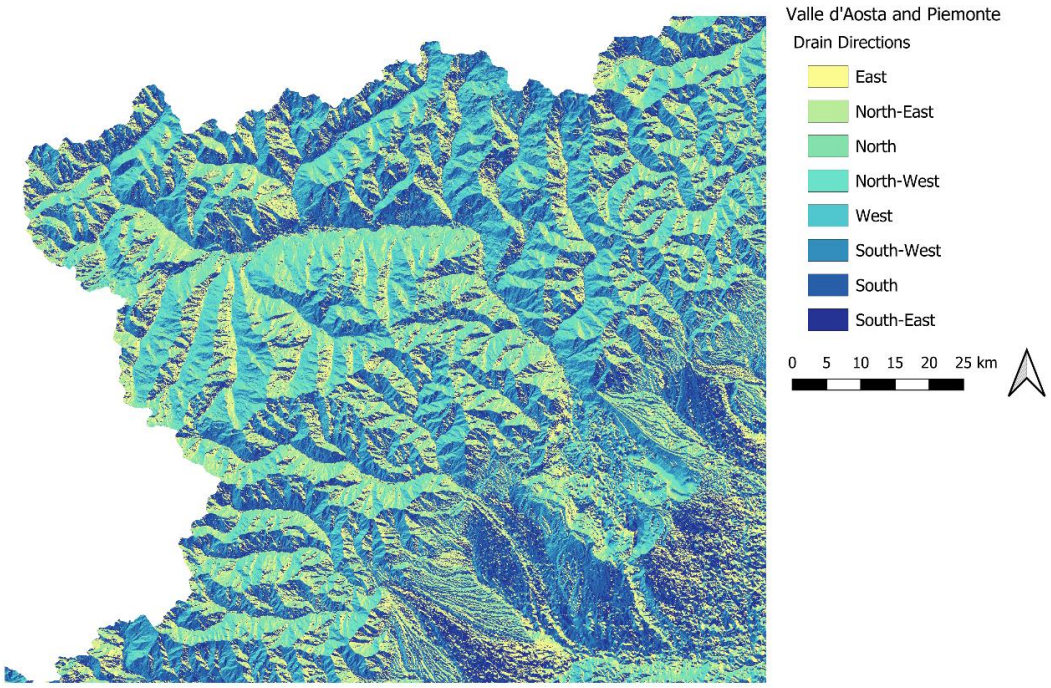



*Figure 2.15, eight drain direction of the entire DEM corrected with the Orlandini method.*

## ***Curvature***

The curvature is defined as "*the inverse of the radius r of the circle tangent to the curve (osculating circle)"* and "*the variation of the tangential vector between two infinitely close positions*" [7]. This step of GEOframe takes only the *pitfilled* DEM as input and gives three different raster files:

1. Planar curvature: "*curvature in a horizontal plane. It is positive for cells with concave contours and negative for cells with convex contour"* [7]. *Figure A1.1 in Appendix A1* reports the result.
2. Profile curvature: "*curvature of the surface in the direction of the steepest slope. It affects the flow velocity of water draining the surface and influences erosion and deposition"* [7]. *Figure 2.16* shows in blue the positive values corresponding to the concave curvatures, thus to the river network and they are in valleys. The yellow area is the flat one.
3. Tangential curvature: "*measures curvature in relation to a vertical plane perpendicular to the gradient direction, or tangential to the contour"* [10].

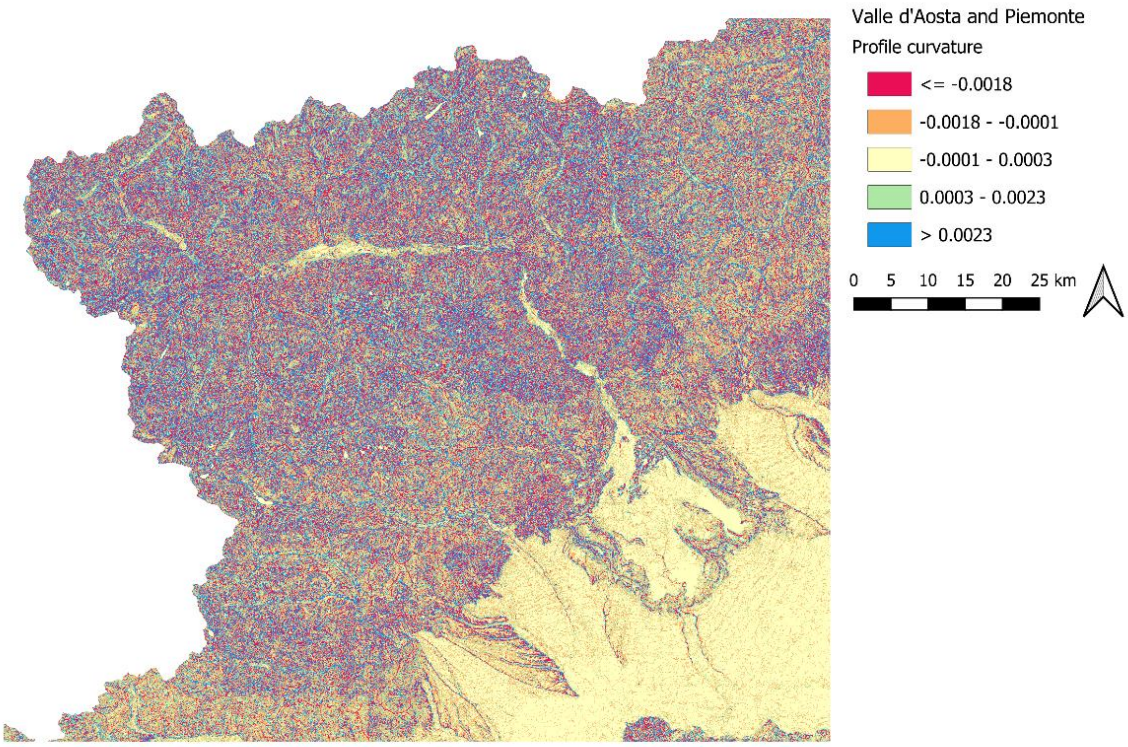



*Figure 2.16. profile curvature of the entire DEM*

***Slope***

This step takes as input the *pitfilled* DEM and the *Drain Directions*. The equation of the slope (2.2) is defined from the equation of elevation and its gradients (2.1):

$$z = f(x, y), f_x = \frac{\partial z}{\partial x}, f_y = \frac{\partial z}{\partial y} \quad (2.1)$$

$$\gamma = arctan\sqrt{f_x^2 + f_y^2} \quad (2.2)$$

The slope is important for the determination of the radiation arriving in each point of the terrain and for the extraction of the river network. Normally, the maximum slope is close to the highest elevation. *Figure 2.17* shows different gradient of slope, the darkest colours indicate the highest and the lowest values. Dark violet is in the higher mountain region where the difference of elevation changes very quickly, instead the dark green is mostly in the flat region in Piemonte where the elevation is quite low.

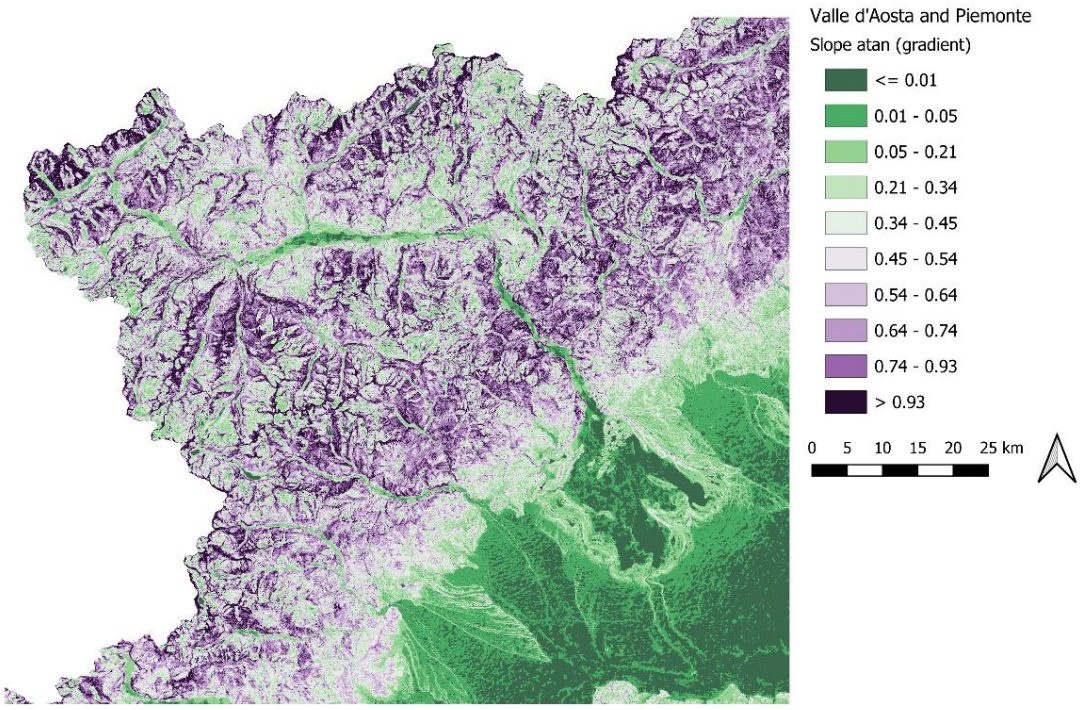



*Figure 2.17, slope expressed in gradient of the entire DEM.*

## ***Tc***

The composition of the curvatures produces 9 main topographic forms, but the most important ones are three: concave, planar and convex. In this step the curvatures files are taken as inputs and two raster files are created: one containing the 9 forms (*Tc9*) and one containing the 3 forms (*Tc3*). *Tc3* has been analysed with QGIS in order to classify the river basin and see better the topographic structure of the terrain. *Figure 2.18* reports the *Tc3* highlighting the concave, planar and convex region. The planar region is exactly where the region becomes flat and, in the *Valle d'Aosta* part, where the main course of the Dora Baltea is located.

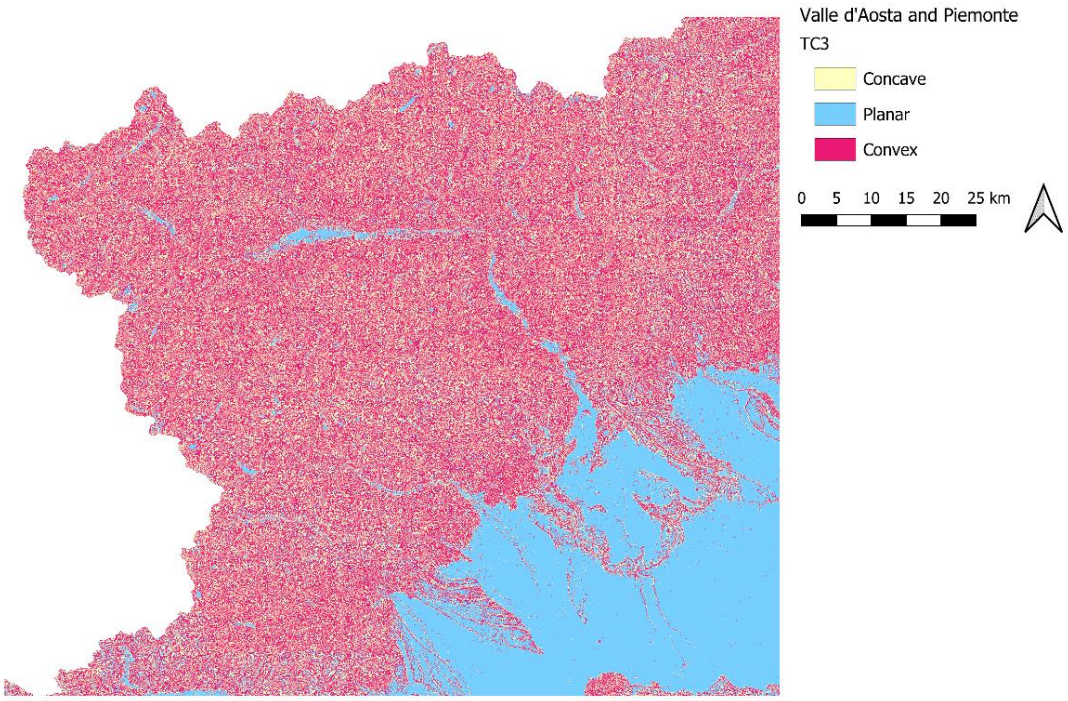



*Figure 2.18, TC3 highlighting the 3 main topographic forms in the entire DEM.*

***Extract Network***

This step takes the *TCA*, *Drain Directions* and *Slope* to extract the river network of the territory. GEOframe uses the method of product of two quantities: TCA and slopes representing stress tangential to the bottom. The gradient is needed, and the tangential stress correspond to the product of the total contributing area over contour length ($\frac{A_T}{b}$) and the slope ($\vec{\nabla}z$) (2.3) [7]:

$$\tau \sim \frac{A_T}{b} |\vec{\nabla}z| \qquad (2.3)$$

The number of branches of the river network can be set in the Java code of this step, a threshold contributing area is defined. In this way, only the points having an area higher than the threshold can form canals. So, the lower this value is the more are the branches. Initially the number was set to 30 to have as much details as possible, but then, looking at the results on QGIS, there were too many branches not corresponding to the reality. So, the value was reduced to 50 to have a more real result. *Figure 2.19* shows the river network extracted over the DEM, it follows the flat part of the region, and the drain directions obtained before. Comparing it with the OMS map, it has a good fit even if there are some parts that do not follow exactly the path of the river. These errors were considered negligible since the main course of the Dora Baltea River is followed almost perfectly.

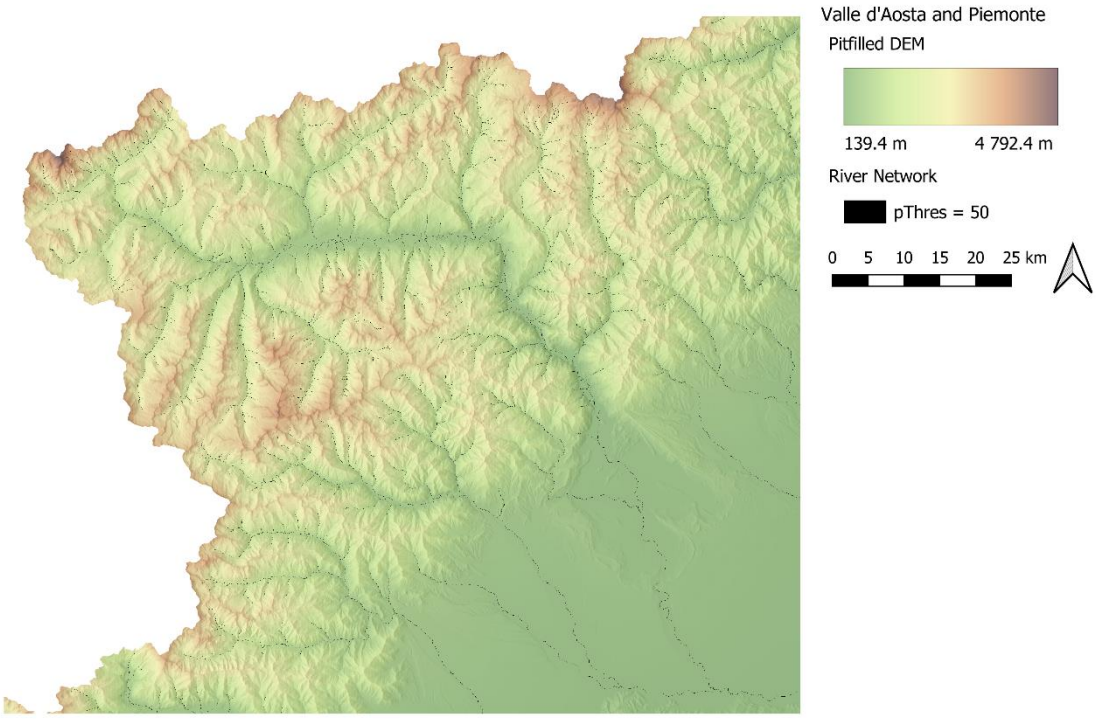



*Figure 2.19, river network of the entire DEM extracted with pThres set at 50.*

The second part of the geomorphological analysis is focused on the extraction of the area of the basin of interest and the creation of the HRUs. As the previous set of phases is composed of 8 subsequent steps: first the raster files obtained before are cropped then the sub-basins are created. *Figure 2.20* shows the flowchart of the final GIS procedure used to obtain the sub-basin division (adapt from *Bancheri, Rigon, and Manfreda 2020* [6]).

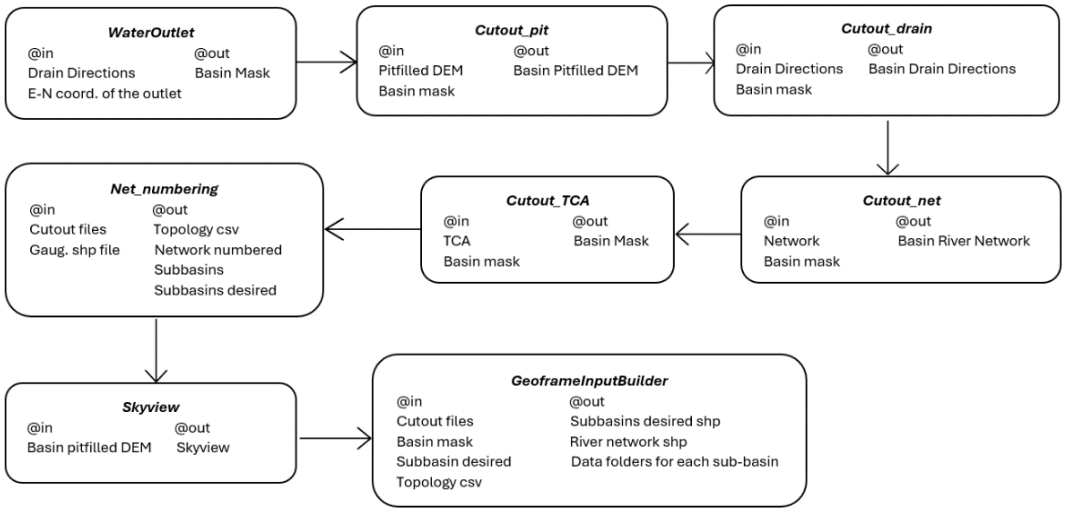



*Figure 2.20, second and last part of the flowchart of the GIS procedure*

All these steps are briefly described below [7]:

***Water Outlet***

The aim of this step is to extract the area of the basin of interest. In order to do that GEOframe requires the *Drain Directions* file and the East and North coordinate of the basin outlet. As explained before, for this case study the Dora Baltea basin is closed at *Verolengo*. *Table 2.3* highlights the coordinates used of the point used:

*Table 2.3, gauging station outlet metadata*

| Station Name | Station ID | Region | UTMX [m] | UTMY [m] | Latitude [°] | Longitude [°] | Elevation (m.s.l.m.) [m] |
|---|---|---|---|---|---|---|---|
| Verolengo Dora Baltea | 3 | Piemonte | 424678 | 5004720 | 45.19194 | 8.04111 | 167 |

It is important that the gauging station is located right above the river network extracted otherwise the software is not able to detect it and extract the basin.
The output is a raster file of the basin mask. *Figure 2.21* shows that the mask is correctly extracted, and it corresponds to the one indicated by "ARPA Piemonte" [2].

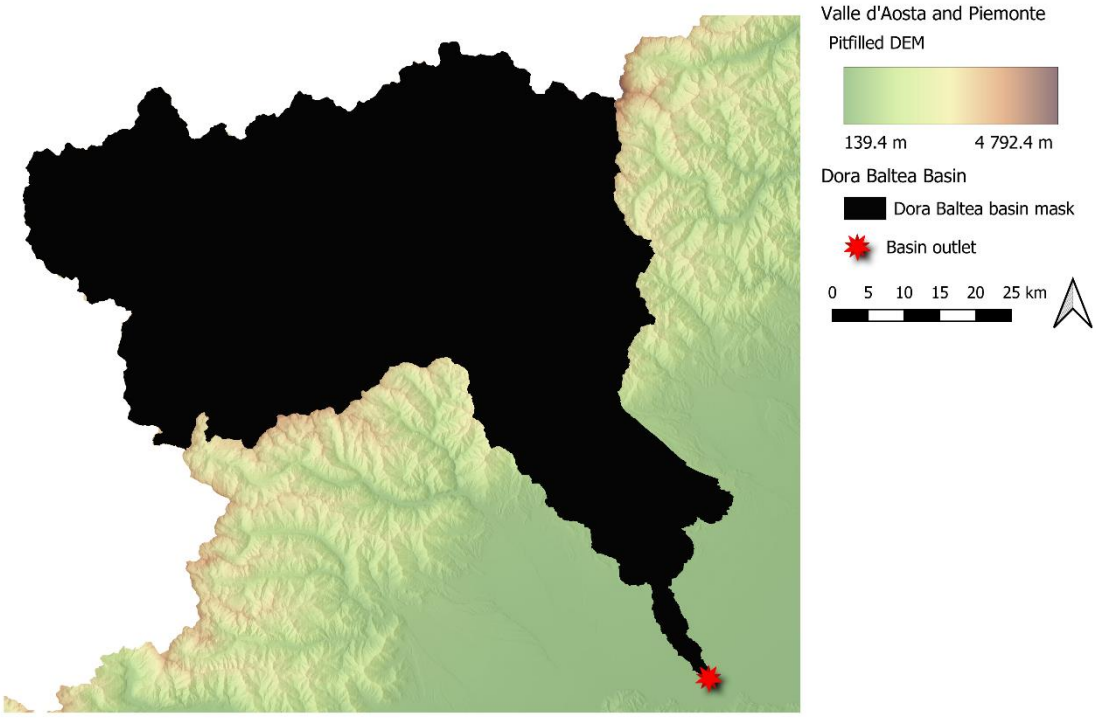



*Figure 2.21, Dora Baltea basin mask*

***Cutouts***

All these 4 cascade steps have the role of crop the *pitfilled DEM*, the D*rain Directions*, the *River Network,* and the *TCA*. They take the mask and the raster file to cut as input and they give the raster basin file. *Figure 2.22* shows the *pitfilled DEM* and the *River Network*

together highlighting the basin outlet. *Figure 2.23* reports the drain direction of the area and finally *Figure A1.2 in Appendix A1* shows the TCA that highlights the main course of the Dora Baltea.

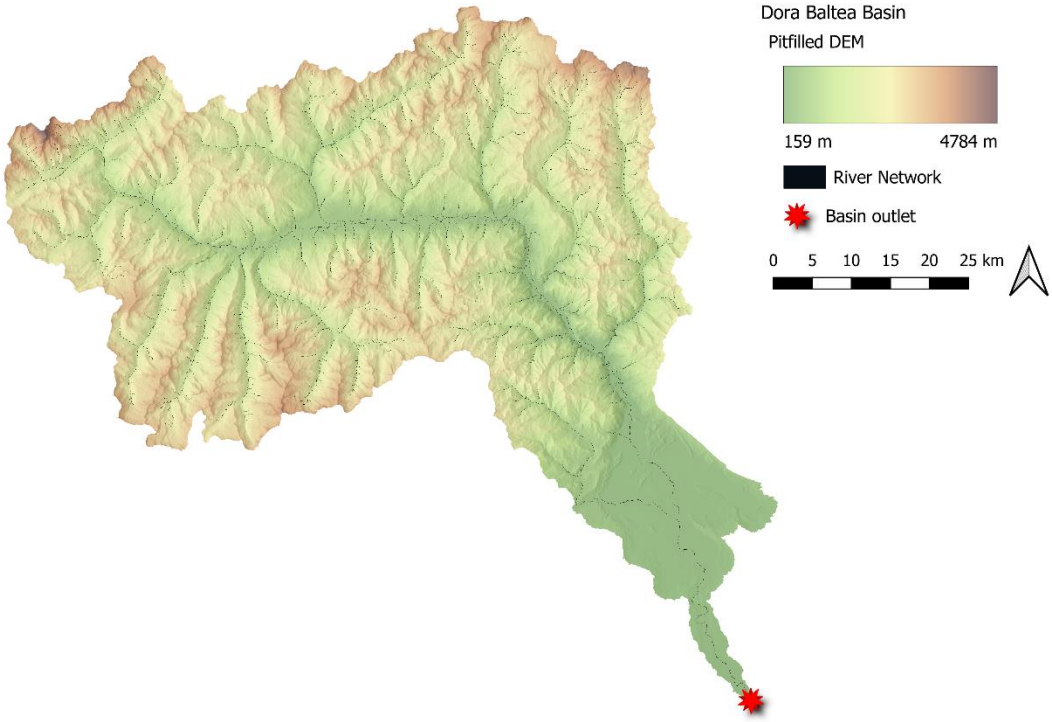



*Figure 2.22, Dora Baltea basin DEM and river network*

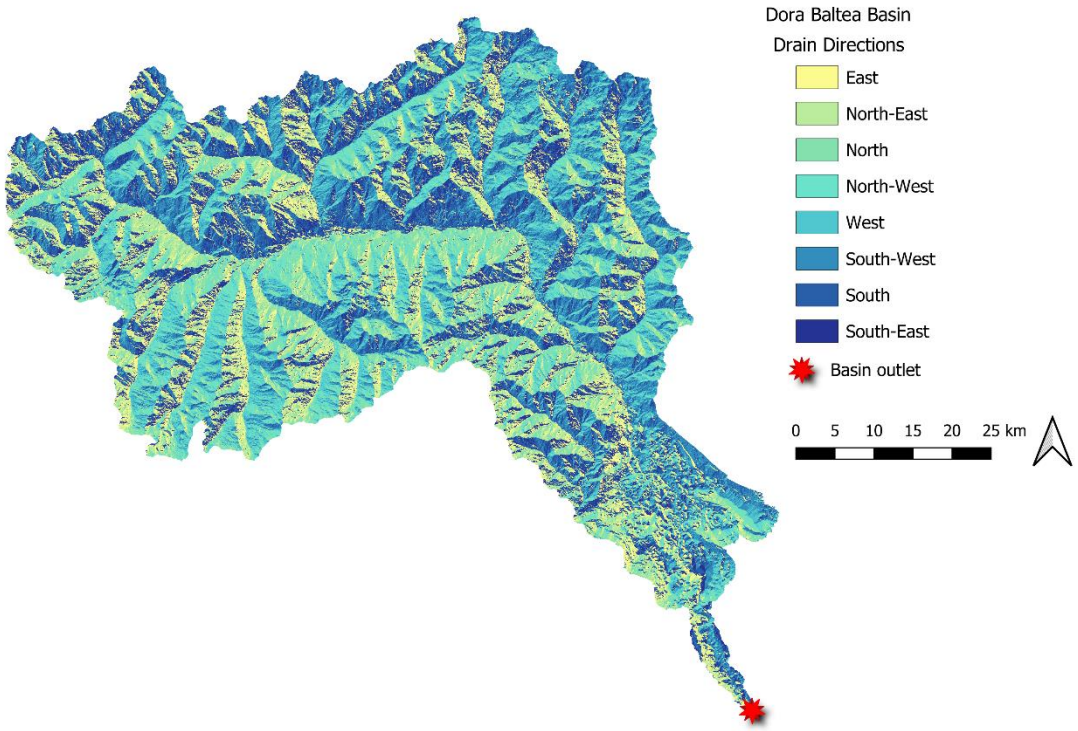



*Figure 2.23, Dora Baltea drains directions.*

***Net numbering***

This is the crucial phase where the sub-basins are created. It takes as input all the 4 cutout files and the shape file of the monitoring points containing the East and North coordinate of each station available inside the basin. Furthermore, it is defined a threshold up to which the basin we are extracting can be aggregated. It indicates the minimum area in $m^2$ that each sub-basin can have. The default value of the software is 5 $km^2$ and the raster file is automatically created by the code. However, in this case study the 50 $km^2$ was used as threshold to have an easy manageable number of sub-basins and, at the same time, aggregate the mountain regions and the flat one in an effective way. The position of the available gauging stations was also considered and since the available monitoring points at high elevation are very few, set a low value would not have sense since the final calibration cannot be done in a so finer way. It was chosen to aggregate a maximum of 3 upstream sub-basins to calibrate them together with the only gauging station available nearby. *Valle d'Aosta* areas with lower elevation were divided with a finer discretization and a lower sub-basin area has been obtained since more hydrometers were available there. In the *Piemonte* part the sub-basin's area are larger since the monitoring points are less. A finer discretization probably would lead to a better modelling of the mountain area that have a rapid variation of the slope and they are strongly affected by snow accumulation and melting processes that are very variable processes. Nevertheless, to make it possible, a less sparse gauging-station network is for sure required. It is also clear that creating more sub-basins would certainly require a huge computational demand and looking at the hydrometers network of the Dora Baltea basin the 50 $km^2$ threshold was

set as the empirical optimal value to have a good size to model the different hydrological processes and a finer discretization was excluded to reduce the computational cost.

The outputs of this step are the raster file of the sub-basin's delimitation, and the river network numbered according to the number of each HRU. *Figure 2.24* displays the sub-division and the correspondent ID. The sub-basin limits are exactly in correspondence of the monitoring points to perform the calibration at the end of the modelling, so compare the simulated discharge with the measured one. The final output is a csv file of the topology of the sub-basins, which contains the routing structure: all the connections between the HRUs are reported in order to know which sub-basin drain into another.

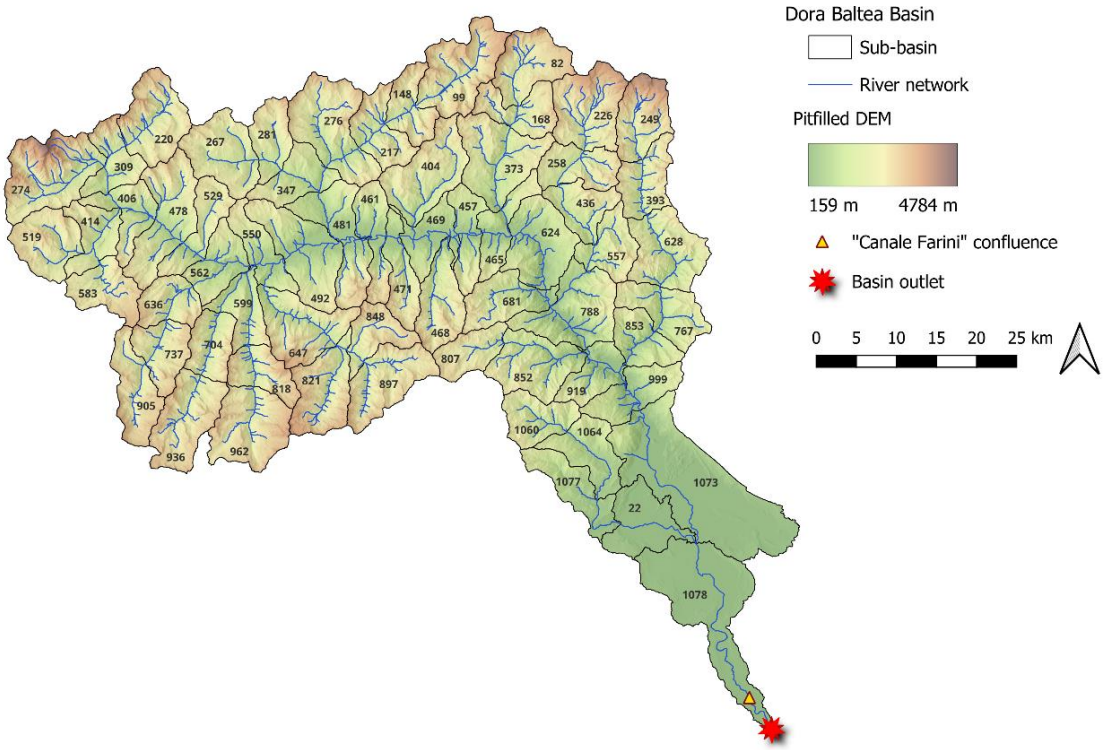



*Figure 2.24, sub-basin delimitation and label with ID numbers of the Dora Baltea basin*

## ***Skyview***

The *skyview* tells the amount of sky visible from each point, it is a multiplier that indicates how much portion of the sky we can see from a point. It goes from 0 to 1, it is high if we are in a flat area, and it is low if we are in mountain area. *Figure 2.25* shows that we have higher numbers going downstream approaching the *Piemonte* part of the basin and we have lower values in correspondence of the mountains and glaciers.

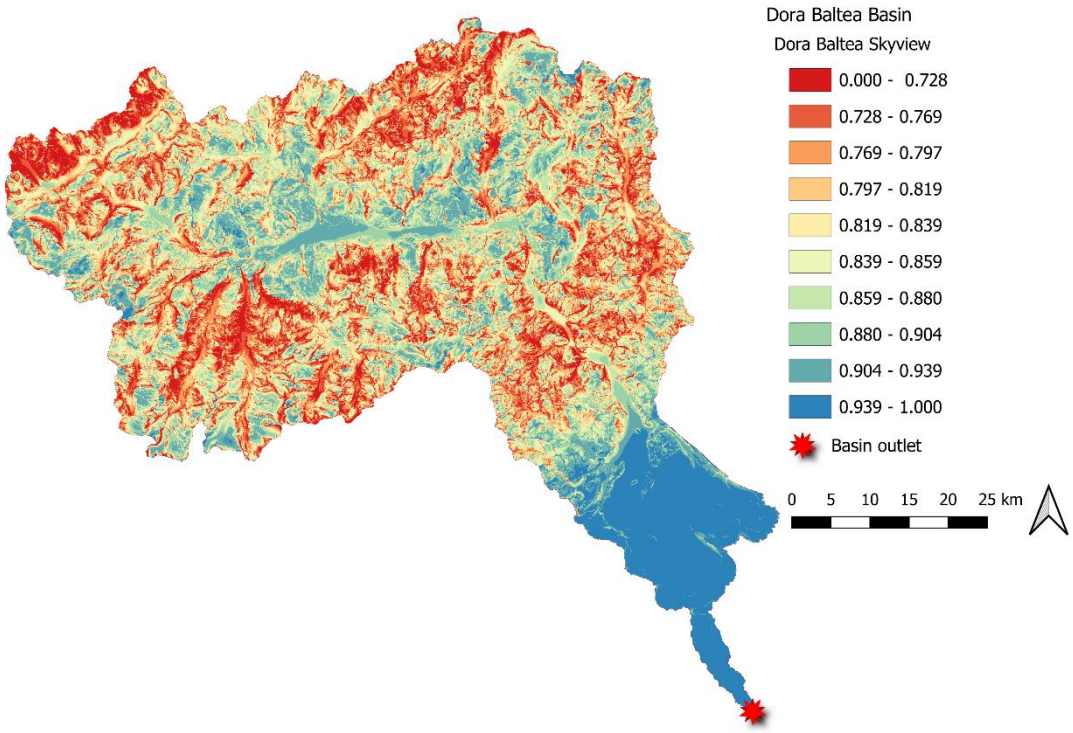



*Figure 2.25, Dora Baltea sky-view*

***GEOframe Input Builder***

In this last phase the input data structure for applying the GEOframe components is created. Taken together the cutouts file, the basin mask, the sub-basins, and the topology file, two shapefiles are created: one for the sub-basins and one for the river network containing all the information. Furthermore, the data structure is created in order to have several folders for each sub-basin to apply GEOframe components in parallel.

The total area of the Dora Baltea basin is 3930 km$^2$ according to "*Autorità di Bacino del Fiume Po*" [11], the basin extracted has a total area of 3877 km$^2$. The value is a bit lower because the basin has been closed at *Verolengo* and not at the real confluence with the Po River which is few km downstream nearby *Crescentino*.
The chosen threshold area of the sub-basin and the gauging-network leads to the creation of 64 sub-basins with areas between 39 km$^2$ and 200 km$^2$. The threshold was set at 50 km$^2$ but the software takes also in consideration the gauging station in order to give the possibility to calibrate at each hydrometer. For that reason, the lowest value is below the threshold, but the mean area is 61 km$^2$. Each HRU is characterised by an area, coordinates x and y in UTM of the centroid, the elevation in correspondence of the centroid, the mean elevation of the entire sub-basin, the length of the network and the *skyview* centroid value (detailed reported in *Table A1.1 in Appendix A1*). The centroid elevation ranges from 230 to 2746 m.a.sl. and the mean elevation goes from 237 to 2831 m.a.s.l. *Table 2.4* reports the average, max. and min. value of the elevation considering the centroid and the mean over the entire sub-basin. To the max. and the min. value also the correspondent sub-basin is reported.

*Table 2.4, comparison of elevation between the mean and the centroids values*

| | Centroid | Sub-basin ID | Mean | Sub-basin ID |
|---|---|---|---|---|
| **Avg. elevation [m]** | 1538 | - | 1941 | - |
| **Max elevation [m]** | 2746 | 519 | 2831 | 249 |
| **Min elevation [m]** | 230 | 22 | 237 | 1078 |

Looking at the centroid elevation the mean one is 1538 m, the maximum one is 2746 m corresponding to the sub-basin 519 and the minimum one is 230 m corresponding to the sub-basin 22. If we look instead at the mean elevation the mean is 1941 m, the max. is 2831 m (sub-basin 249), and the min. is 237 m (sub-basin 1078). The elevation difference is on average 403 m between the two different values. The mean is higher than the centroids and they identify different sub-basins. In the following step of the modelling where the temperatures and precipitation are interpolated, the mean one was used since it was considered much more representative. There are sub-basins that have the elevation difference higher than 800 m (see *Figure 3.8*) and it surely underestimates the temperature ranges and so, the snow processes.

*Figure 2.26* shows the final result of the geomorphology analysis, highlighting the gauging-stations and the basin outlet.

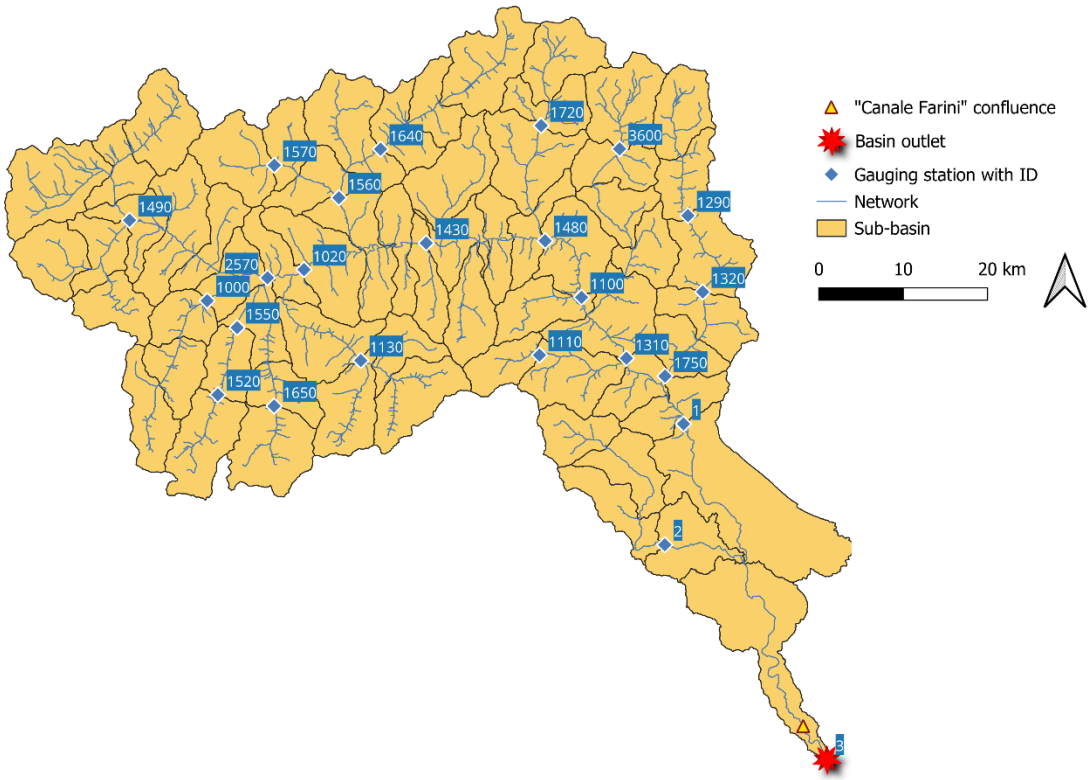



*Figure 2.26, final result of the geomorphological analysis of the Dora Baltea basin*

# 3 METEOROLOGICAL DATA

In this chapter the meteorological data of the Dora Baltea Basin are presented and analysed. After that, the spatial interpolation of temperature and precipitation data performed by the model GEOframe component is showed. The section is divided into four parts: the first one gives information about the data availability and data distribution for the Dora Baltea catchment. The second part is about the GEOframe interpolation method: Kriging. Thus, a briefly part about the theory and how GEOframe applies it is proposed. The second- and third-part shows, in order, the results of the temperature and precipitation interpolation of the region of interest.

## 3.1 Temperature and precipitation data

The meteorological data available from the *Regional Environmental Protection Agencies* are the necessary inputs to perform the spatial interpolation during the modelling. Thus, the timeseries of the data and the shape files containing the coordinates of each station have been retrieved and analysed.

The meteorological station data, including the geographical position and timeseries of temperature and precipitation, were downloaded online from "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [4] concerning the *Valle d'Aosta* part. Instead, the data for the *Piemonte* part have been downloaded from "*ARPA Piemonte*" [5]. A daily scale was chosen in order to model a large number of years and make the computation more efficient. Indeed, having only one value for each day decrease the computational demand. However, it is clear that considering the daily average the precision of the model is affected. Nevertheless, this time resolution was considered sufficient to analyse the hydrological status of the basin without compromising the computational capabilities of the compiler used.

The temperature data downloaded are obtained using a thermometer able to measure the air temperature in Celsius degree. Considering a daily scale, the values are averaged over 24 hours. Instead, the precipitation data downloaded are measured using tipping bucket rain gauge as instrument that provides values expressed in millimetres. The daily value indicates the height of the fallen rain and the water equivalent of solid precipitation correspondent to the cumulative value over 24 hours, calculated from 00:00 to 24:00 UTC. Almost all the rain gauges used are unheated, so during the winter season, they can record precipitation values outside of the event due to the melting of snow or having no data when there is ice inside.

The data need to be pre-processed to remove incoherences, organize them and create a shape file containing all the stations with their coordinates. This step is fundamental to have an idea of the data availability and so, based on that, decide the range of years in which perform the modelling. The pre-processing was done with Excel and QGIS (version 3.30.0-'s-Hertogenbosch) [3]. Finally, the data were processed with a *jupyter* code [53] available from GEOframe developers, in order to make the data readable by the model.

The *Dora Baltea Basin* extracted in the geomorphology step and showed in *Figure 2.26* contains 92 meteorological stations represented by red dots in *Figure 3.1*.

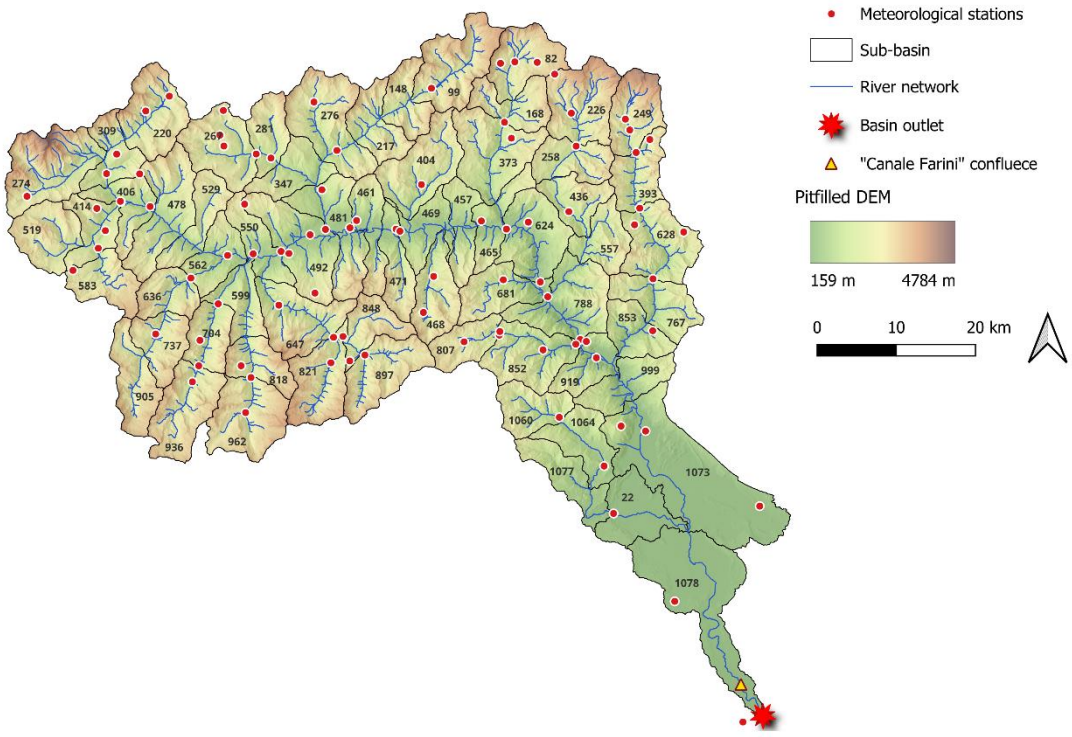



*Figure 3.1, Dora Baltea and all the meteorological stations*

These 92 stations are sub-divided into 91 temperature stations – with a density of 2.35/100 km$^2$ - and 82 precipitation stations – with a density of 2.12/100 km$^2$ (look at *Table 3.1* further ahead).

It is clear that not all the stations have data for both quantities downloaded: not all the temperature stations are also precipitation stations. In the *Piemonte* part only 8 stations are available for the Dora Baltea Basin, and they all have both parameters available. In the *Valle d'Aosta* part, only one station has only precipitation data and no temperature data. *Figure 3.2* and *Figure 3.3* shows the temperature stations as red dots and the precipitation stations as green dots. Notice that station n. 11 in *Piemonte* is outside the basin area to let the interpolation perform better since few stations are available in that region.

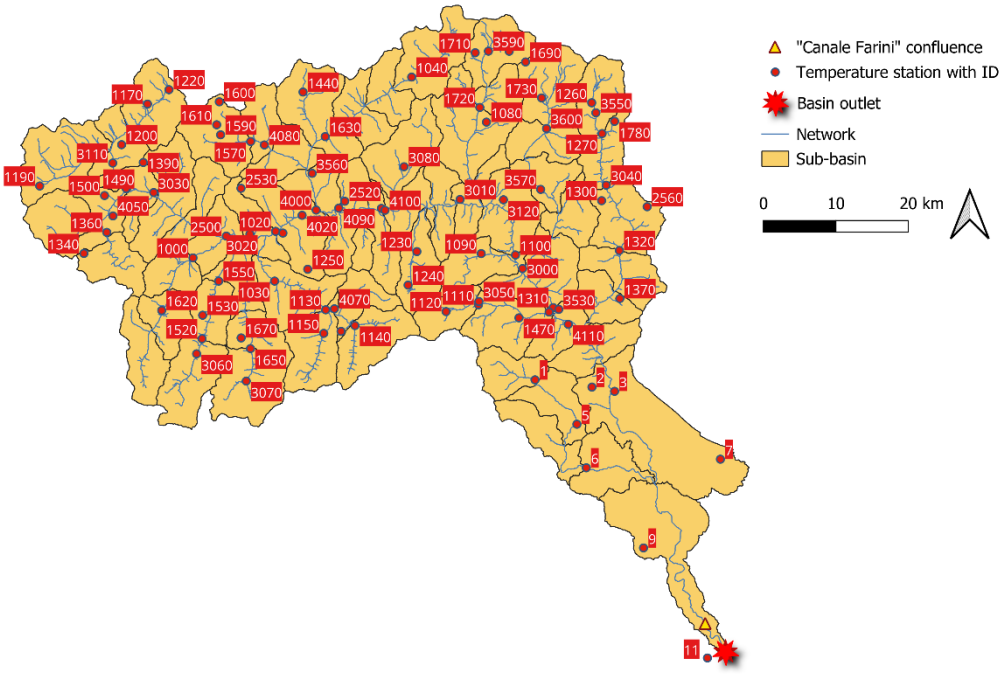



*Figure 3.2, Dora Baltea basin and temperature station with ID labels*

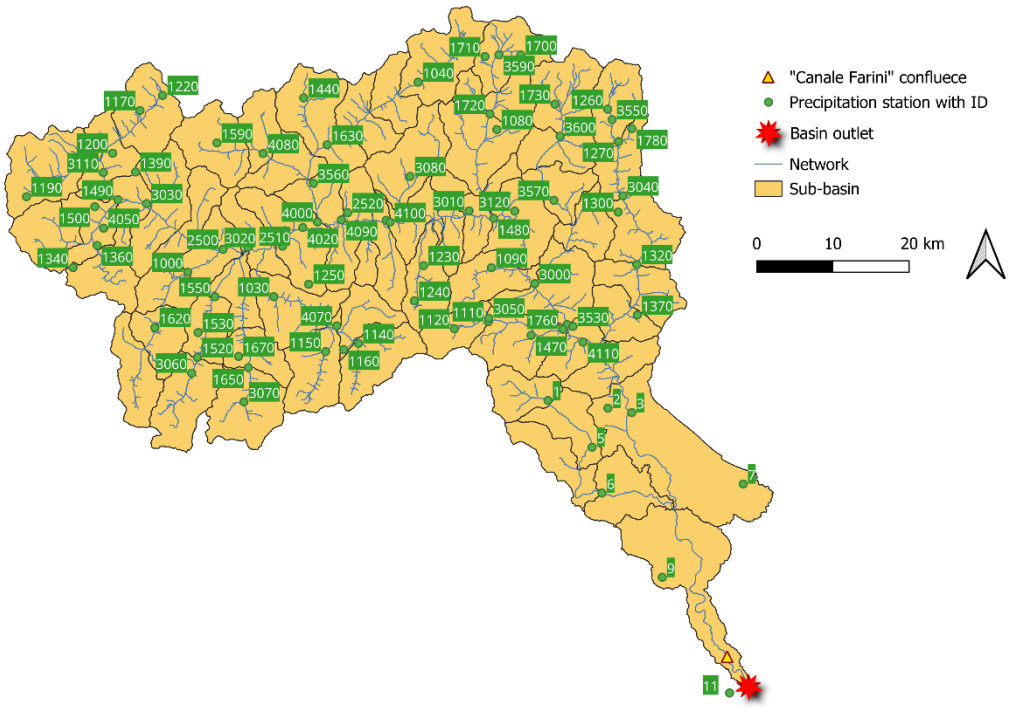



*Figure 3.3, Dora Baltea basin and all precipitation stations with ID labels*

GEOframe needs to have different shapefiles if the temperature and precipitation stations are not the same. If a single shapefile is used for both parameters, the software does not have the ability to recognize which station has temperature and/or precipitation data. *Table A1.2 in Appendix A1* reports all the metadata of the stations included in the shapefiles used as input. As it was done for the gauging stations, the geographical position available online are expressed with latitude and longitude using the reference system WGS84. For completeness, the coordinates have been also projected in the UTM zone 32 reference system since it is the one used in the project.

During this phase, it is also important to look how the data measured are distributed. *Figure 3.4* shows the data availability for temperature stations. All the white gaps represent data missing due probably to instrument detection issues. There are 10 stations that have been collecting data since 1996, confirmed also by the metadata written by "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [22] available online. During the initial phase in which the software was tested to understand its functionality, it was found out that GEOframe can perform interpolation provided that at least three stations with complete data are present. So, it was considered a viable choice to set the start date at 01/01/1996.

Together with the start date is also defined the time in HH:MM format. In the current case study, the hour of each value measured and so interpolated is set at 12:00. Since it is a daily analysis, the time was set at noon in order to guarantee a correct estimation of the mean radiation in the subsequent step (see 4.1).

*Piemonte*'s stations are the last 8 in the graph (*Figure 3.4*) and almost all of them have been functioning since 1996. From the graph is also clear that there is a substantial number of stations start having data from 2000 on, and 2 new stations have been installed recently in 2023.

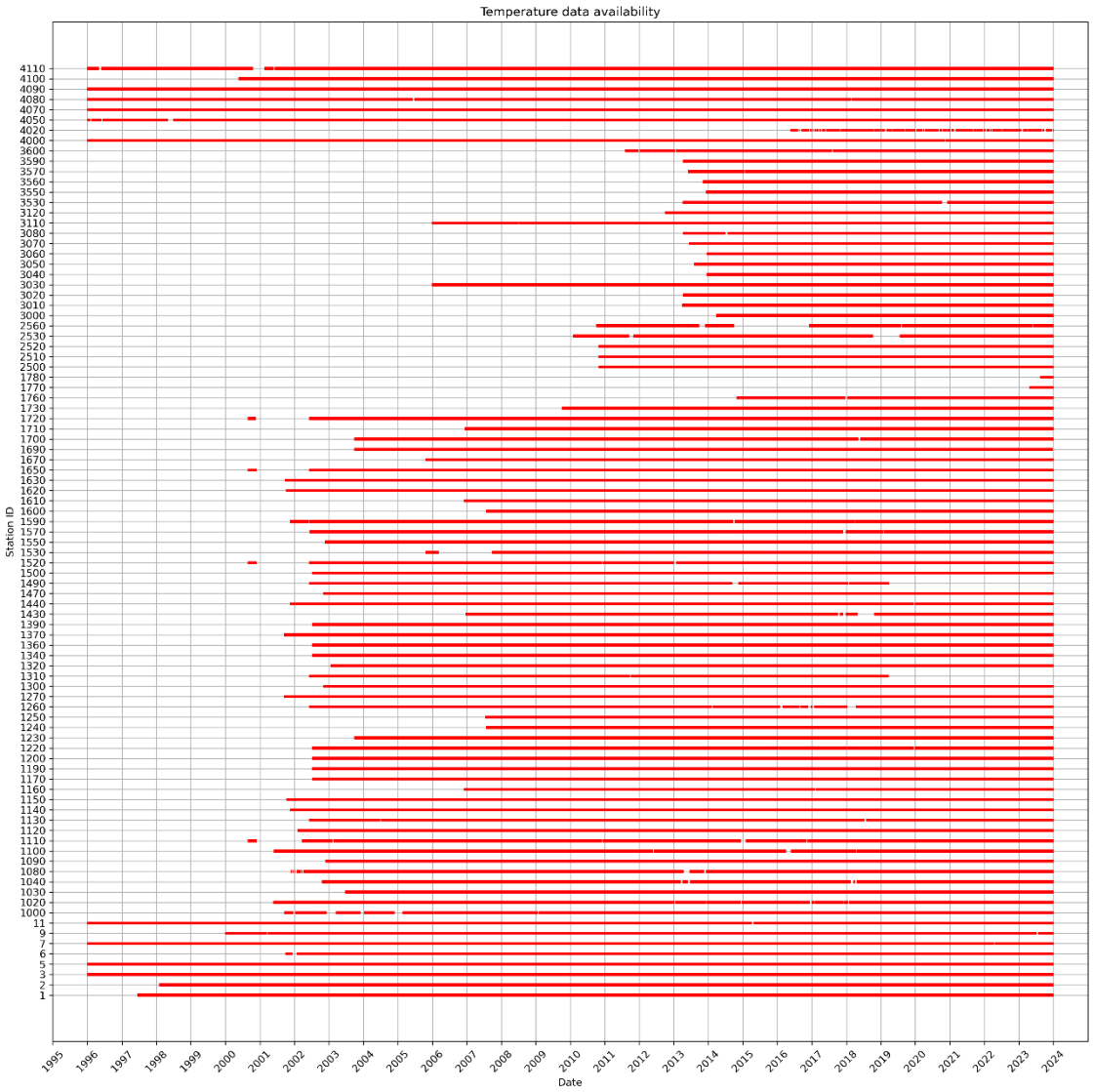



*Figure 3.4, temperature data availability for the Dora Baltea basin*

Finally, *Figure 3.5* shows the data availability for precipitation stations. Even though almost all the stations started having data in the same period of the temperature ones, it is clear that there are much more gaps. As the histogram in *Figure 3.6* shows, there are 39 out of 82 precipitation stations that are located above 1500 m.a.s.l. Since, almost all stations are unheated, it is highly probable that the gaps in the measurements are caused by the formation of ice inside, making the instrument unable to function until the ice has melted. Indeed, looking at the timeseries, the gaps seem to be very cyclical, and they appear in the coldest period during winter. The station with the highest elevation is the n. 1390 located in "*Morgex – Lavancher"* with an elevation of 2842 m.a.s.l. *Figure A1.3 in Appendix A1* clearly shows the pattern described above: no data available for every year in the winter period from December to April.

All these measurement gaps certainly influence the interpolation; but since they are the only meteorological available data in the region of interest and they are associated with a denser temperature data availability, they were considered sufficient for the modelling.

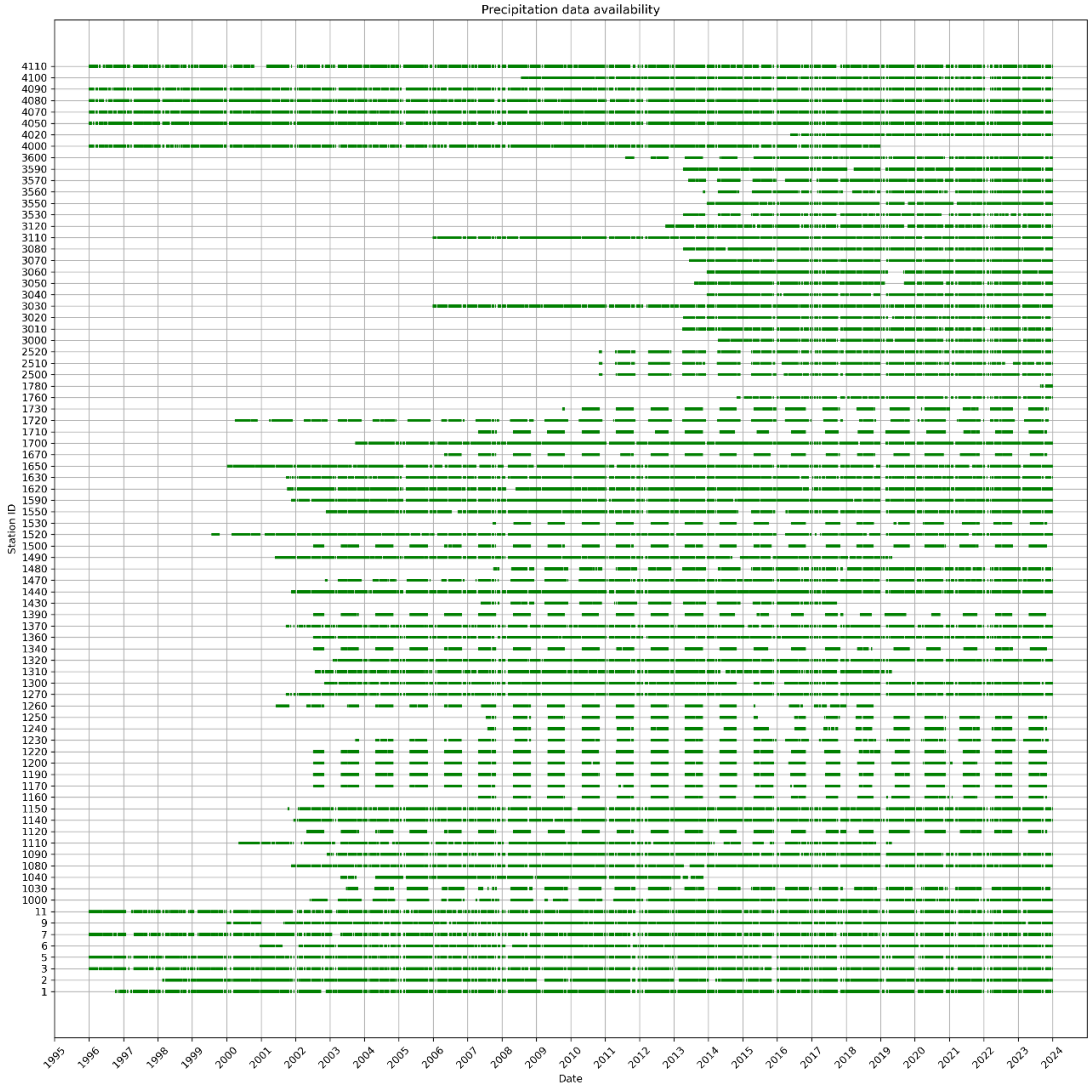



*Figure 3.5, precipitation data availability for the Dora Baltea basin*

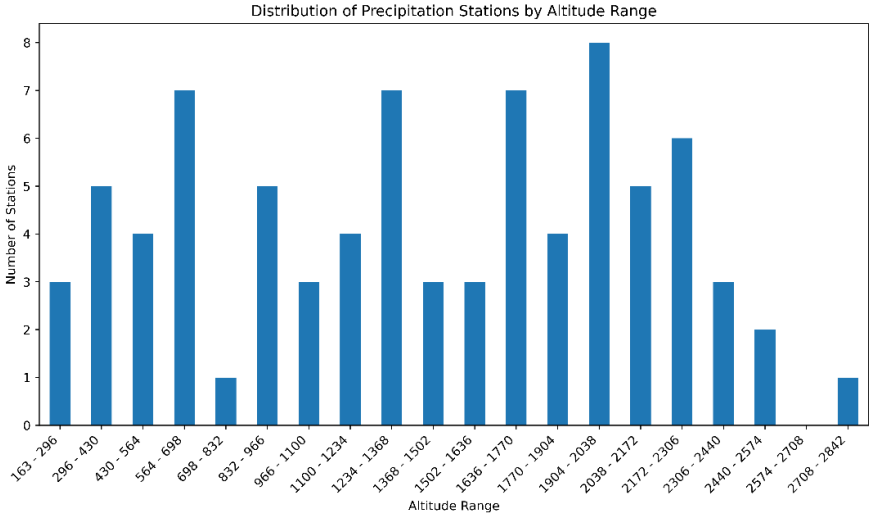



*Figure 3.6, distribution of precipitation station by altitude range*

*Table 3.1* summarizes the numbers of station available for each measured quantity and the simulation period.

*Table 3.1, final properties of the Dora Baltea basin*

| Area[km$^2$] | Nr. Sub-basins | Nr. Precipitation Stations | Nr. Temperature Stations | Nr. Total stations | Simulation Period |
|---|---|---|---|---|---|
| 3877 | 64 | 82 | 91 | 92 | 1996 - 2023 |

## 3.2 Kriging theory

Temperature and precipitation data are spatially interpolated by GEOframe using Kriging. "*It is essentially an interpolation technique based upon linear least squares estimation extended to ReV (regionalized variable). A ReV is a deterministic function which generally varies such a rate that it is best considered in terms of its probabilistic interpretation*" *(Virdee and Kottegoda 1984)* [16]. Basically, it is a way to know the meteorological variable values in all the other points in which we do not have measurements. Kriging is mostly used when the set of data that must be interpolated have high natural variability, in our case temperature and precipitation. It is based on the concept that spatially close data points are more likely to have similar values than those farther apart. So, the variable dispersion decreases with distance and weights are obtained to decide in each point which measure can be trusted more [17].

Kriging uses an instrument to correlate observations with distance: the variogram. As *Equation (3.1)* shows, it is the semi variance between two points. It is essentially the measure of dissimilarity, so how much the two values differ each other at distance *h* [19].

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} (z(u) - z(u+h))^2 \qquad (3.1)$$

where N(h) is the number of couples at distance h from each other, z(u) is the head point and z(u+h) is the tail point. The plot obtained is first an empirical variogram where the available realization values are used. "*In order to be extended to any distance the variogram need to be fitted to a theoretical variogram model* (explained below)*. The fitting to the theoretical variogram is also necessary to find the values of the weights*" *(Bancheri et al. 2018)* [20]. At the end, the plot compares the values of the variogram with different distances h, the higher the distance is the lower is the correlation. The variogram contains three different parameters [18]:

- *Sill*: "*total variability inherent in the data*" *(Formetta and Rigon, 2024)* [18]. It is the upper limit of the variogram,
- *Range*: "*separation distance at which the variability reaches the sill. When samples are separated by distances beyond the range of continuity, they have no spatial correlation*" *(Formetta and Rigon, 2024)* [18],

- *Nugget effect*: "*expected difference between samples when the separation distance is almost zero*" *(Formetta and Rigon, 2024)* [18]. It is a high variation at a scale smaller than the sampling distance and it can be due to measurement errors.

The graphical representation of the variogram plot is showed in *Figure 3.7*.

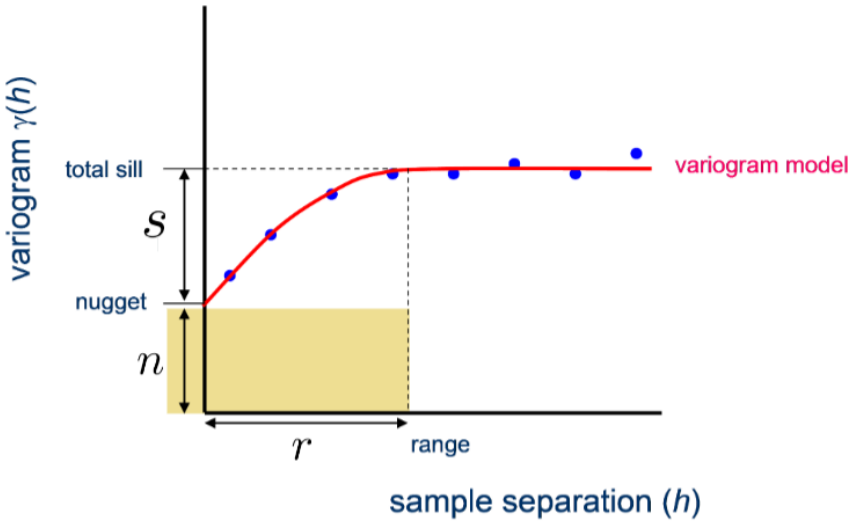



*Figure 3.7, variogram plot highlighting the 3 different parameters. Credit: geoframe.blogspot.com [18]*

The variogram have different models [21], GEOframe gives the possibility to choose from the main three:

- Exponential: it grows rapidly in the first part and then it reaches the constant value of the sill. "*The spatial dependence decreases exponentially with increasing distance and disappears completely at an infinite distance" (Mohd Aziz et al. 2019)* [20],
- Linear: it grows linearly with distance, and it never reaches the stable value of the sill. It is good for phenomenon that has a variability that continuously increases with the distance,
- Spherical: it grows rapidly at first then it slows down and it reaches the sill. It is good for phenomenon with a limited variability till a certain distance.

Finally, we distinguish between three types of kriging:

- Simple Kriging, "*which considers the mean to be known and constant throughout the study area*" *(Bancheri et al. 2018)* [20],
- Ordinary Kriging, "*which accounts for local fluctuations of the mean, limiting the stationarity to the local neighbourhood*" (in this case the mean throughout the study case is unknown but constant) *(Bancheri et al. 2018)* [20],
- Detrended Kriging, "*which considers that the local mean varies within the local neighbourhood" (Bancheri et al. 2018)* [20]. The procedure removes first the trend and then the ordinary kriging is applied. At the end, the sum of the interpolated values and the trend gives the final result.

The workflow, used by GEOframe, for solving the interpolation problem with Kriging is summarized in the following steps [19]:

1. Get the data from gauges,
2. Build the empirical variogram among existing data, defining the type of Kriging to use,
3. Fit a theoretical model to the variogram, to be chosen from the three possibilities reported above,
4. Use the theoretical model for solving Kriging system, getting the *sill*, *range,* and *nugget* values of the variogram,
5. Produce continuous surface maps or pointwise timeseries of the quantity desired in any point of the domain,
6. Calculate estimation errors.

From a practical point of view in GEOframe, the data from gauges are downloaded and pre-processed, then point n. 2, 3 and 4 are done together with a Java code that gives as results the parameters values of *sill*, *range,* and *nugget*. Finally, simulation files are created for each sub-basin and, after the run is done, the timeseries of temperature and precipitation are created.

## 3.3 Kriged temperature and precipitation

In the following two paragraphs the results of the Kriging interpolation for temperature and precipitation are reported.

### 3.3.1 Temperature interpolation results

The temperatures are interpolated from the meteorological stations to the sub-basin centroids. The Java code (reported in *Figure A1.4 in Appendix A1*) takes as inputs the start date set on 01/01/1996, the time step, the elevation of each station and all the stations' timeseries. Furthermore, the method of Kriging interpolation is defined. In this case the detrended Kriging was selected since temperatures are quite affected by elevation and, so the height dependencies are taken into consideration. Finally, the type of theoretical variogram to use is defined. The exponential variogram was chosen since the spatial dependency decreases slowly till it disappears completely, so it was considered the good one to interpolate temperatures quantities. The outputs of this step are the values of *sill*, *range* and *nugget* reported in *Table A1.3 in Appendix A1*. The codes gives both values with and without the trend. The values obtained with the trend are higher than the others because a global scale is considered, and the systematic variations are much larger than the ones at a local scale. The trend gives a better result in a large scale, and it seems to capture the global spatial structure of the phenomenon.

The values are then used to create the simulation files able to generate the timeseries of each sub-basin. In this step, the elevation of the each HRU is given as input. GEOframe uses as default the centroid elevation, but as *Figure 3.8* shows the difference between the average elevation and the centroids elevation in a lot of sub-basins is too hight. Clearly, almost every HRU has an elevation in the centroids that is on average 403 m less than the

mean one. Sub-basin n. 461 has an elevation difference of almost 1200 m, having 1707 m.a.s.l. as average elevation and 538 m.a.s.l. as centroids elevation. It is clear that considering the centroids elevation seems to be not representative of the topographical setting of each sub-basin. Furthermore, as it was reported before in *2.3*, considering the centroids elevation leads to underestimate the low temperatures and so the snow processes.

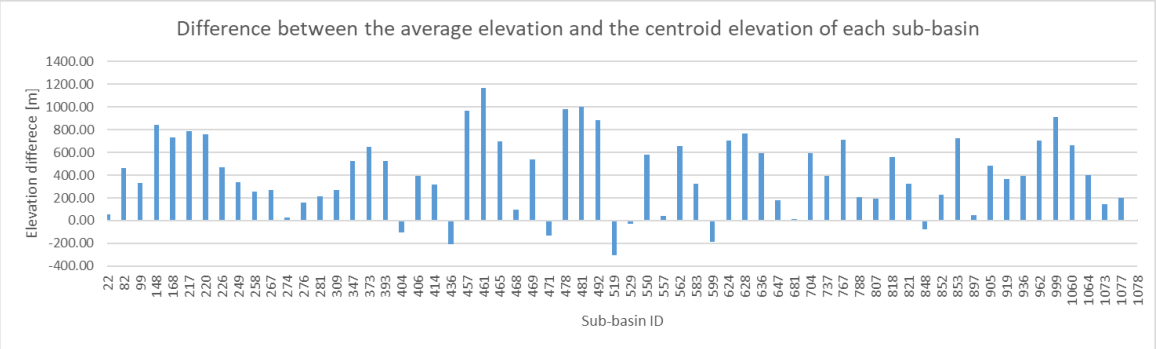



*Figure 3.8, plot of the difference between the average elevation and the centroid elevation of each HRU*

Finally, the daily temperature timeseries are obtained for each sub-basin running the last Java code with all the parameters set before. *Figure 3.9* shows the entire timeseries for the sub-basin 1078 having the lowest elevation and the sub-basin 249 having the highest elevation.

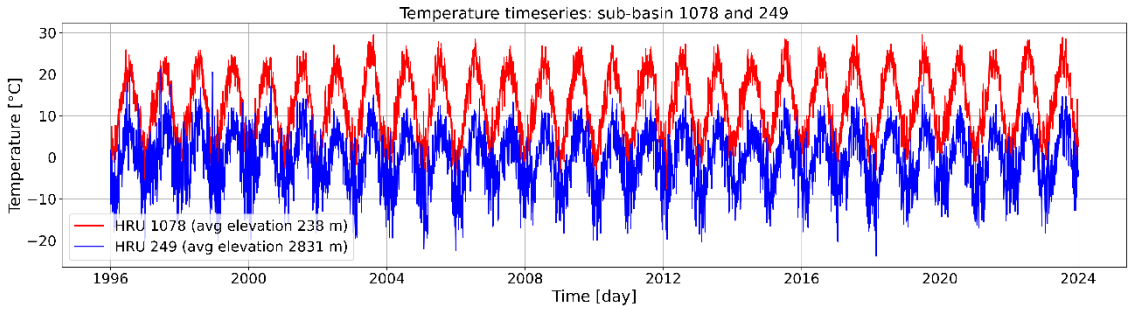



*Figure 3.9, plot of the timeseries of sub-basin 1078 and sub-basin 249*

As expected, the temperature values of the sub-basin with the highest average elevation are about 10 units lower than the ones in sub-basin with the lowest elevation. This path is much more evident during the coldest periods.

It is important to notice that GEOframe evaluates during Kriging the closest station from each sub-basin centroid. Then, temperatures are evaluated based on the identified closest station and the average elevation of the sub-basin. Indeed, the temperature are interpolated considering also the average elevation of the sub-basin. *Figure 3.10* and *Figure 3.11* clearly shows this consideration in the year 2011, selected as an example: in the first plot where the elevation difference between the closest station and the sub-basin is incredibly low, the timeseries are remarkably similar. Whereas, when the elevation difference is higher, as in *Figure 3.11*, the timeseries have lower values even though they follow the same path.

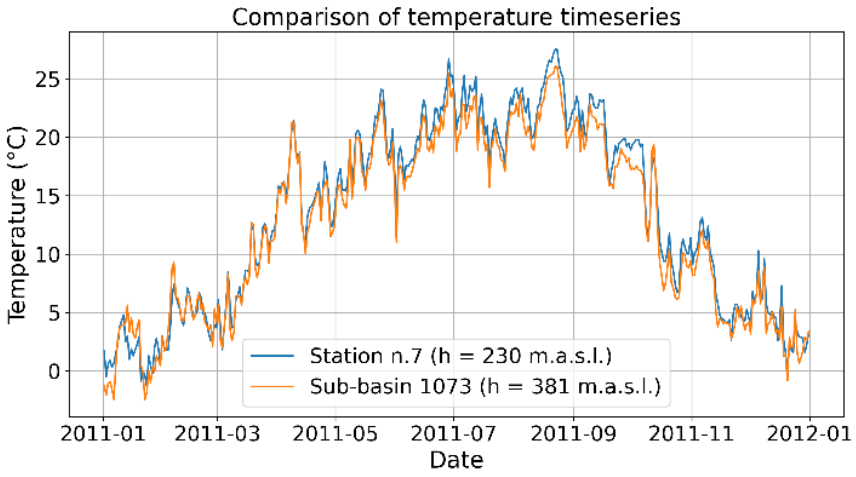



*Figure 3.10, comparison of the timeseries for station n. 7 and sub-basin 1073*

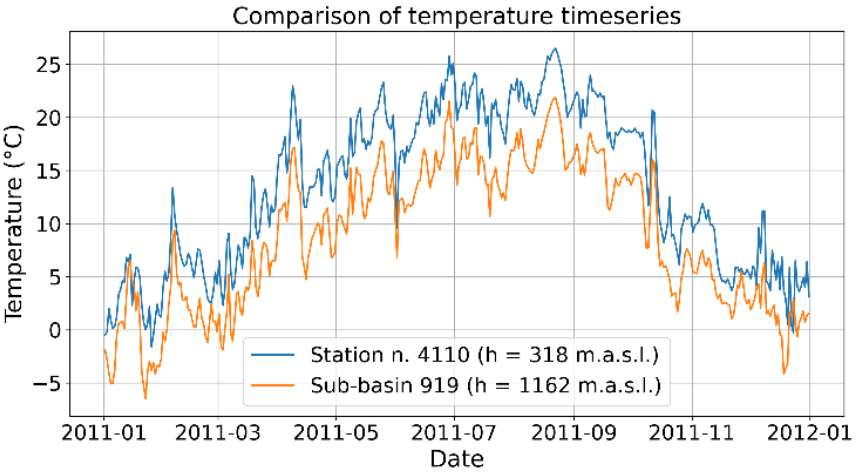



*Figure 3.11, comparison of the timeseries for station n. 4110 and sub-basin 919*

The average temperature of each sub-basin is calculated over the entire period chosen for the modelling. *Figure 3.12* shows the map of the Dora Baltea Basin with distinct colours based on the average yearly temperature obtained between 1996 and 2023.

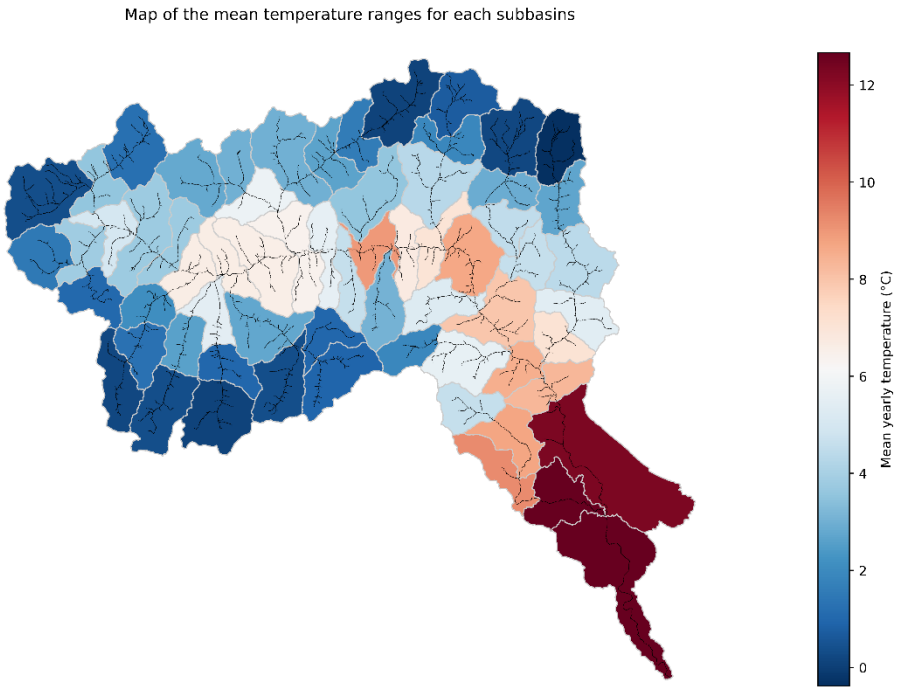



*Figure 3.12, map of the average yearly temperature between 1996 and 2023 in the Dora Baltea basin*

As can be observed, the mean annual temperatures vary from -0.4 to 12.7 °C. It is clear that the lowest mean temperatures are located where the elevation is higher. In the central part of *Valle d'Aosta* - where the main course of the Dora Baltea River is located and where the territory becomes flatter - mean temperatures increase. Reaching the basin outlet and entering in *Piemonte*, the temperatures are higher since the mean elevation is around 350 m.a.s.l. and the territory have a foothill topology.

The created temperature map has been compared with the same map of the entire Italian territory available from ISPRA (*Istituto Superiore per la Protezione e la Ricerca Ambientale*) [23]. The map of the mean yearly temperature calculated from 1951 to 2019 reported in *Figure 3.13* shows a remarkably similar pattern in the same studied territory. The period used by ISPRA is much larger than the one used by the model, but it was considered a reliable source to compare the values obtained. Over the mountains in *Valle d'Aosta* the temperatures reported are a little bit lower than the one simulated, but when the territory becomes flatter the values seems to be similar.

In addition, how GEOframe use the elevation influence over the temperature is also clear looking at *Figure 3.14* where the linear interpolation of the mean yearly temperature, considering the elevation, is showed. As it was said before the highest value is in the sub-basin with the lowest elevation (n. 1078) and the lowest value is in the sub-basin with the highest elevation (n. 249). The temperature lapse rate is of approximately 0.52 °C/mm. The same trend is reported by "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [24].

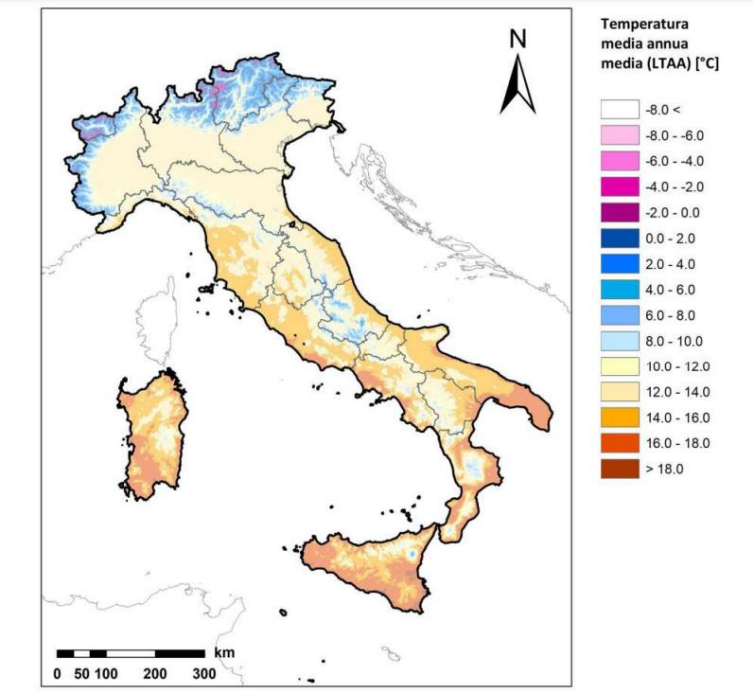



*Figure 3.13, mean yearly temperature over Italy from 1951 to 2019. Credit: isprambiente.it [23]*

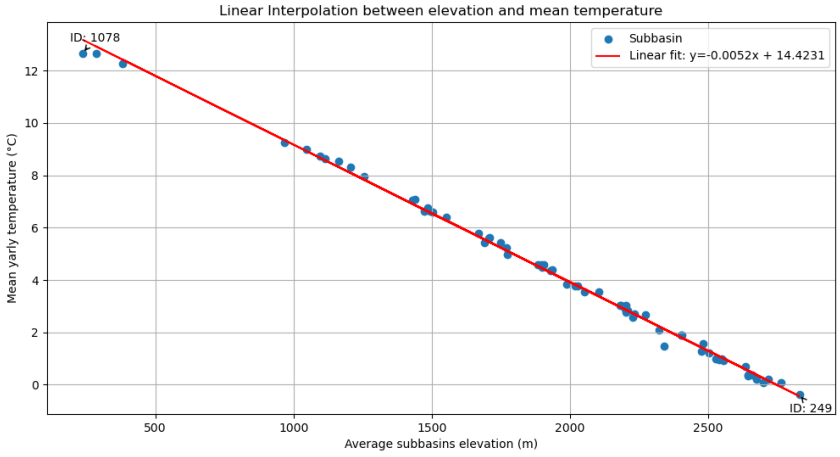



*Figure 3.14, linear interpolation between elevation and mean yearly temperature modelled of each sub-basin.*

The mean yearly temperature of the sub-basins within the Dora Baltea catchment can be seen in *Figure 3.15* where each blue line represents an HRU. It is clear that all sub-basins follow the same pattern. The three upper lines are the sub-basins located in the flattest part of the basin in *Piemonte* and the other ones have similar trend with a decreasing value approaching the highest elevation. The coldest year seems to be 2010, whereas the hottest years are 2022 and 2023. It is also evident the increasing temperature over the period due to global warming.

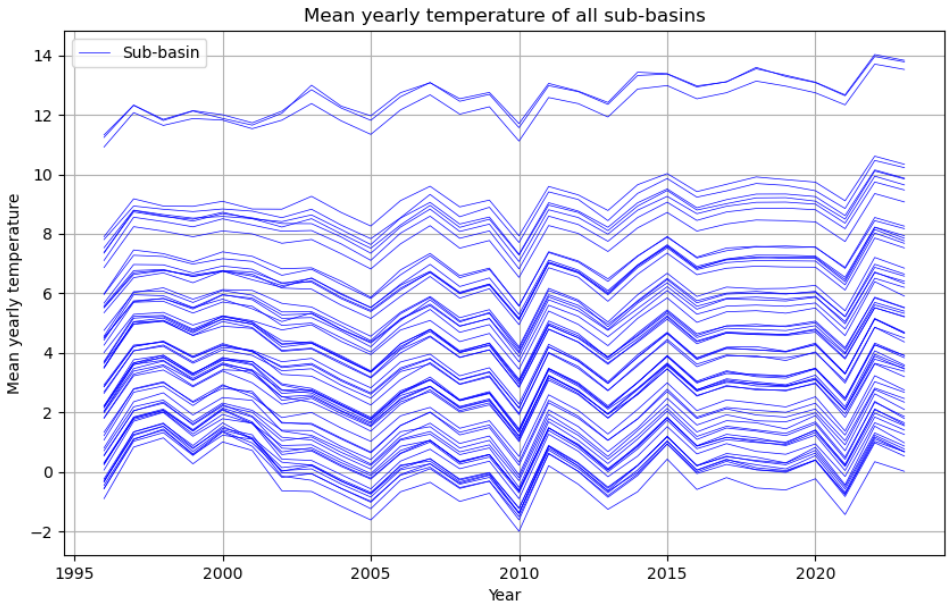



*Figure 3.15, mean yearly temperature of all sub-basins in the Dora Baltea catchment.*

Considering the timeseries obtained, another consideration about the increasing temperature trend was done. Looking at *Figure A1.5 in Appendix A1* where the temperature timeseries of the sub-basin n. 1078 is reported with the trend line, we can notice that the slope is of 0.0002 °C per day that correspond to 0.073 °C per year. Over a span of 27 years, which is the period of modelling, this would amount to a total increase of approximately 1.971 °C, which is considerable and would have significant impacts on local ecosystems, snow cover and glacial melting. For comparison, studies on temperature trends in the Western Alps, including the *Valle d'Aosta*, have shown the same noticeable increases in temperature over the past few decades. For instance, the ISPRA [23] reports show that temperature increases in the Alps are typically higher than the global average, underscoring the region's vulnerability to climate change, with potential consequences for biodiversity, hydrology, and human activities.

It is also remarkable that the exact same trend is reported in *Figure A1.6 in Appendix A1* where the same timeseries with measured data in station n. 11 near the sub-basin n. 1078 is shown. It is again clear that GEOframe used that as reference to interpolate.

Furthermore, the mean yearly temperature map in 2023 was compared with the same map of only the *Valle d'Aosta* region available from "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [24]. As it can be seen in *Figure A1.7 in Appendix A1*, the temperature ranges are terribly similar even though there is an underestimation in the highest mountain regions where the temperatures interpolated seem to be around 0.5°C higher. Probably a denser meteorological network, especially in the mountain region, would have certainly increased the overall result of the entire interpolation.

### 3.3.2 Precipitation interpolation results

The precipitation values were estimated with the same procedure of the temperature values. The setting of the code to estimate the variogram parameters are remarkably similar. The type of Kriging used, even in this case, is the detrended one. The values of *sill*, *range* and *nugget* are reported in *Table A1.4 in Appendix A1*. The global values without the trend considered are high. This is because precipitation phenomena are quite variable with the distance and the spatial distribution is less uniform than the one of the temperatures. Adding the trend the values decrease a lot explaining the high variation over a larger scale. Looking at the extremely low value of the nugget effect with the trend it is clear that the model can explain in an effective way the spatial variability, only small local variations can have an estimation with low accuracy.

As it was showed before, the stations available with precipitation data are less than temperature ones and they have much more gaps in the measurements. These issues affected certainly the interpolation since the approach of the Kriging interpolation is the same: finding the minimum sub-basin centroid-station distance and interpolate from that data. Thus, having less stations with not uniform data decrease the accuracy of interpolation. Also, in this case the average elevation of the sub-basin is used to perform the interpolation.

At the end of the process, the precipitation timeseries for each sub-basin are produced. *Figure 3.16* shows the precipitation timeseries over the entire modelling period, from 1996 to 2023, of the sub-basin where the outlet is located: n. 1078. The highest values of precipitation are during autumn as expected for the climate of the catchment.

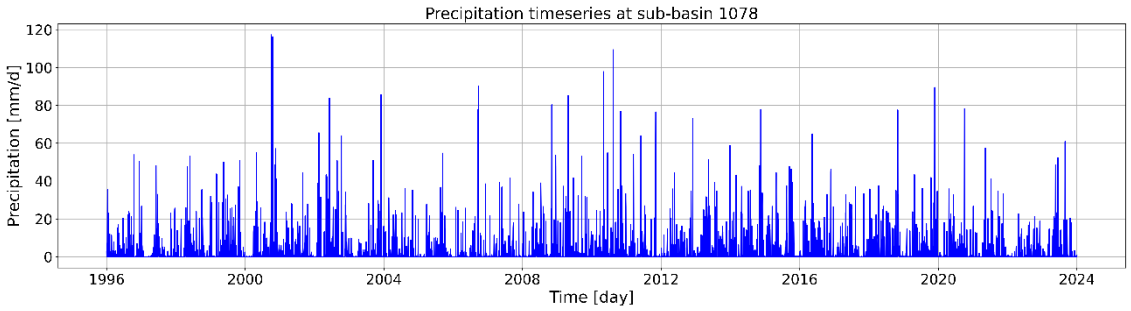



*Figure 3.16, precipitation timeseries simulated at sub-basin 1078 from 1996 to 2023.*

The trend reported is similar to the closest station since the elevation difference between the station and the sub-basin is basically the same. Indeed, the same comparison done for the temperatures was carried out even with precipitation data. *Figure 3.17* shows the comparison of the cumulative precipitation in 2023 between a couple of station-sub-basin that have a low elevation difference. *Figure 3.18* shows the same graph in the same year but of a couple station-sub-basin that have a high elevation difference. The precipitation is expected to be higher approaching higher elevation and both graphs show this trend. Indeed, the total value at the end of the year is lower in *Figure 3.17* where the elevations are lower. *Figure 3.18* has higher values but even though the elevation difference between the station and the sub-basin is high, the precipitation values does not differ so much. It

is clear that the dependency from elevation is much more pronounced in the temperature values than in the precipitation ones. Lower temperature creates atmospheric conditions that are often good for precipitation, especially when combined with other factors like humidity and atmospheric movements. Thus, having more precipitation approaching higher elevation, and so lower temperature, is a general trend, but there are many aspects to take into consideration. Furthermore, *Figure 3.17* shows also the gaps present in the observed data making the interpolation more difficult and less comparable.

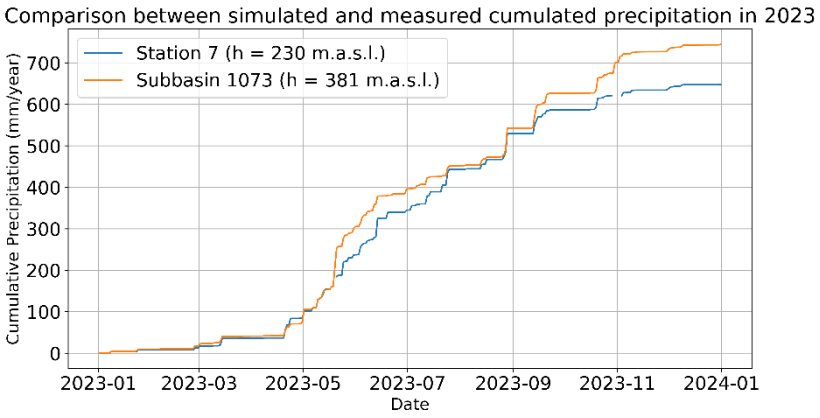



*Figure 3.17, comparison between simulated and measured cumulative precipitation in 2023 at station n. 7 and sub-basin 1073.*

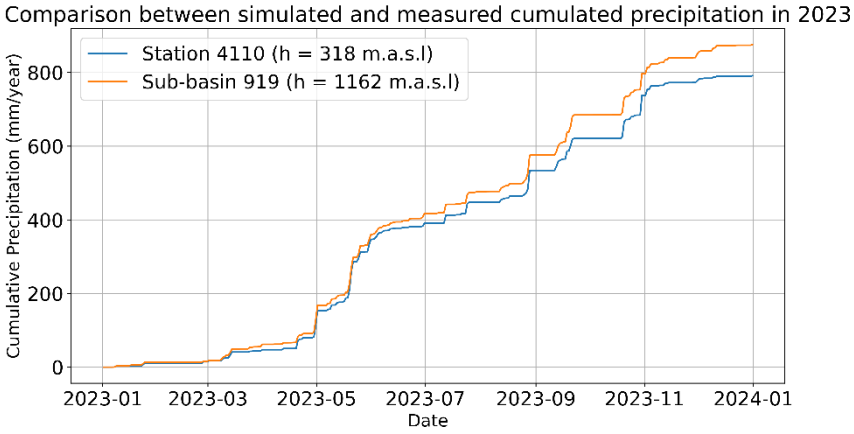



*Figure 3.18, comparison between simulated and measured cumulative precipitation in 2023 at station n. 4110 and sub-basin 919.*

The average cumulative precipitation of each sub-basin is calculated over the entire period chosen for the modelling. *Figure 3.19* shows the map of the *Dora Baltea Basin* with distinct colours based on the average yearly precipitation obtained between 1996 and 2023.

The mean annual cumulative precipitation varies from 578 to 1192 mm/year. The general trend, as reported before, is to have more precipitation approaching higher values. Indeed, over the mountain regions, the cumulative precipitation seems to be higher due probably also to the snow contribute. The highest value is in the left upper part of the basin, where

the *Mont Blac* is located. The lower values are located in the central part of the *Valle d'Aosta* where the main course of the *Dora Baltea River* is located and where water is much more drained into the river.

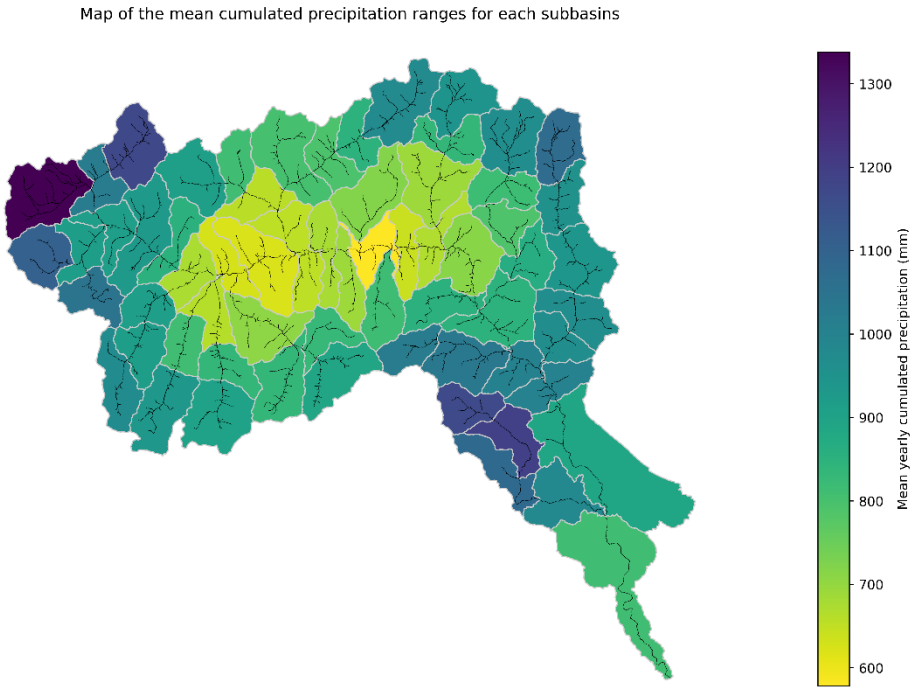



*Figure 3.19, map of average yearly cumulative precipitation between 1996 and 2023 in the Dora Baltea basin*

The created precipitation map has been compared with the same map of the entire Italian territory available from ISPRA (*Istituto Superiore per la Protezione e la Ricerca Ambientale*) [25]. The map of the mean yearly cumulative precipitation calculated from 1951 to 2019 reported in *Figure 3.20* shows a remarkably similar pattern in the same studied territory. The lower and higher values are reported in the same part of the territory. The comparison is the same done with the temperature values.

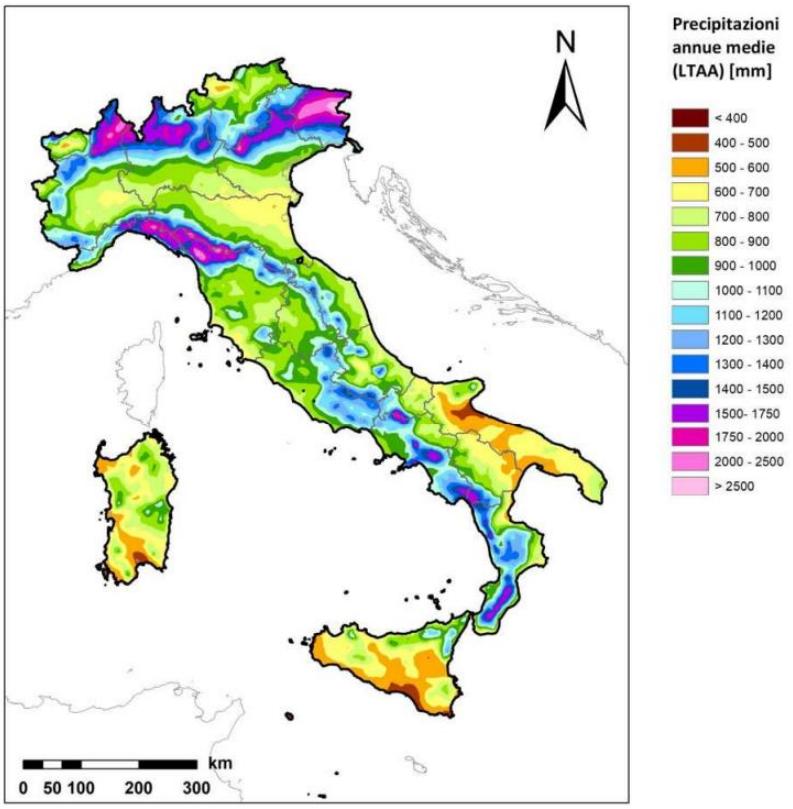



*Figure 3.20, mean yearly cumulative precipitation over Italy from 1951 to 2019. Credit: isprambient.it [25]*

Furthermore, the mean yearly cumulative precipitation map in 2023 has been compared with the same map in the same year of only the *Valle d'Aosta* region available from "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [26]. In *Figure A1.8 in Appendix A1* the comparison can be retrieved. The two maps seem to be terribly similar in both mountain and flat regions. The precipitation interpolation is considered robust.

Finally, the mean yearly cumulative precipitation of the sub-basins within the Dora Baltea catchment can be seen in *Figure 3.21* where each blue line represents an HRU. In the first period of interpolation all sub-basins seem to have a similar trend. From 2005 to 2015 there are much more fluctuations and not all basins follow the same trajectory. This is due probably to the measured data used as input that have, as it was reported before, a lot of gaps, thus the interpolation was harder and with a lower accuracy. The wettest year with 2000 mm/year is certainly in 2000 when a significant flood happened, whereas the driest years are 2005 and 2022 with 400 mm/year. The interannual variability is high since the mean yearly precipitation of the entire basin can vary almost 900 mm/year.

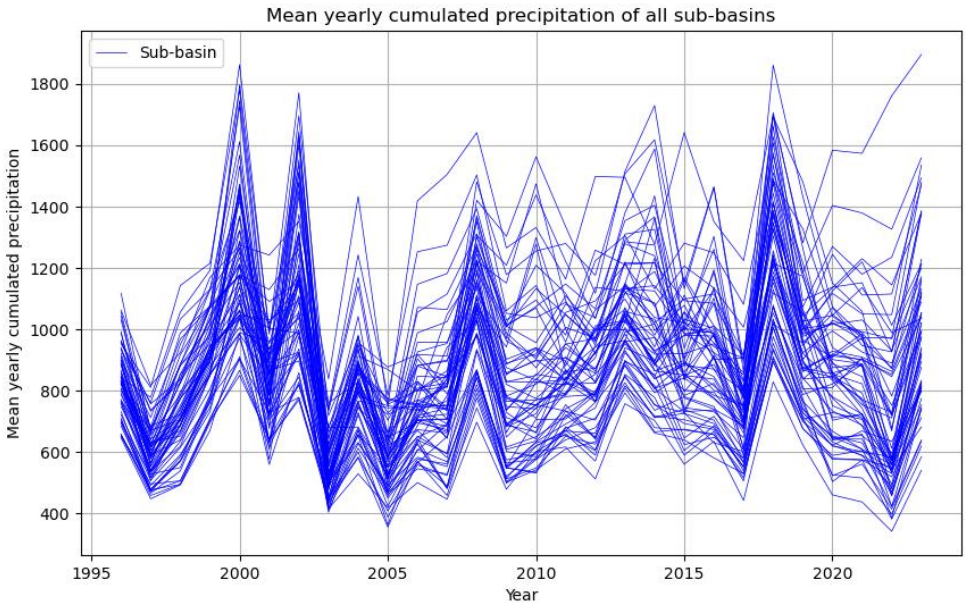



*Figure 3.21, yearly cumulative precipitation of all sub-basins in the Dora Baltea catchment*

The overall result was considered good. It is clear that a dense pluviometer network would have improved the interpolation capturing the spatial variability of the precipitation in mountainous areas.

Anyway, it is important to note that the interpolation was done based on the data available online from the Regional Authorities. All these datasets are not validated, and they all clearly report on their websites that some data can have some errors and outsiders. Looking the results, the values obtained seem to be acceptable for the modelling, but they are clearly an estimation.

# 4 GEOframe ESTIMATION COMPONENTS

In this chapter the GEOframe estimation components are reported. The section is divided into three main parts. Starting from the radiation values estimation for the Dora Baltea Basin: the first part describes briefly the theory applied by the software and the second part contains the interpretation of the results obtained. Following, the evapotranspiration estimation is reported explaining briefly the theory and the results obtained. Finally, the Embedded Reservoir Model used by the software is described in the last part together with the initial results in terms of discharge values.

## 4.1 Radiation estimation in GEOframe

Radiation data in GEOframe are computed using the three different components of solar radiation: shortwave radiation, longwave radiation, and net balance radiation. The net radiation is the balance between the two parts of the longwave radiation and the shortwave radiation. It is the final product used to evaluate also the evapotranspiration in the subsequent step of the modelling.

The three components are briefly described from the software theoretical point of view, in the following sections.

***Longwave radiation component***

The longwave radiation refers to the infrared energy emitted by the Earth. It "*is an important component of the radiation balance on earth and it affects many phenomena, such as evapotranspiration, snowmelt (Plüss and Ohmura, 1997), glacier evolution (MacDonell et al., 2013), vegetation dynamics (Rotenberg et al., 1998), plant respiration, and primary productivity (Leigh Jr., 1999)"* (Formetta et al. 2016) [27]. These types of waves are emitted by the Earth, and they all reach the atmospheric layer where are absorbed due to greenhouse effect, clouds, and other components. Then, they are re-emitted in all directions. For that reason, we distinguish between:

- *Upwelling* $LW_{\uparrow}$, part of the longwave component emitted by the Earth and capable of crossing the atmospheric layer and reaching the space,
- *Downwelling* $LW_{\downarrow}$, part of the longwave component emitted by the Earth, stopped by the atmosphere and re-emitted back toward the Earth's surface. They contribute to heat the Earth surface.

Several simplified models (SMs) were developed over the years to model the two different components by using easily available meteorological observation such as air temperature, relative humidity, incoming solar radiation, and cloud cover. SMs formulation for $LW_{\uparrow}$ and $LW_{\downarrow}$ are based on the Stefan-Boltzmann equation (Formetta et al. 2016) [27]:

$$LW\uparrow = \varepsilon_s * \sigma * T_s^4 \tag{4. 4}$$

$$LW\downarrow = \varepsilon_{all-sky} * \sigma * T_a^4$$

*(4. 2)*

where:

- $\sigma = 5.670 * 10^{-8}$ is the Stephan-Boltzmann constant,
- $\varepsilon_{all\text{-}sky}$ [-] is the atmosphere effective emissivity,
- $\varepsilon_s$ [-] is the soil emissivity,
- $T_a$ [K] is the near-surface air temperature,
- $T_s$ [K] is the surface soil temperature.

In order to solve the first *Equation (4.1)*, for simplicity the surface soil temperature $T_s$ is substitute with the air temperature. In our case, the temperature timeseries interpolated with Kriging are used for each sub-basin. Concerning the soil emissivity $\varepsilon_s$, it is normally a property of the surface, and it varies based on the type of soil. *Figure 4.1* reports six distinct types of surfaces and the corresponding value of soil emissivity (Formetta et al. 2016) [27].

| Nature of surface | Emissivity |
|---|---|
| Bare soil (mineral) | $0.95 - 0.97$ |
| Bare soil (organic) | $0.97 - 0.98$ |
| Grassy vegetation | $0.97 - 0.98$ |
| Tree vegetation | $0.96 - 0.97$ |
| Snow (old) | $0.97$ |
| Snow (fresh) | $0.99$ |

*Figure 4.1, soil emissivity for each nature of surface (Brutsaert, 2005). Credit: Formetta et al. 2016 [27]*

For simplicity in all sub-basins the value of the grassy vegetation correspondent to 0.98 was used. It is also the default value suggested by the software developers.

Looking at *Equation (4.2),* the atmosphere effective emissivity $\varepsilon_{all\text{-}sky}$ accounts for the increase of the overall value in cloud cover conditions. It is formulated according to *Equation (4.3):*

$$\varepsilon_{all-sky} = \varepsilon_{clear} * (1 + a * c^b)$$

*(4. 3)*

where: c is the clearness index [-] quantifying the presence of clouds in the sky, a and b are two calibration coefficients.

$\varepsilon_{clear}$ is the emissivity considered in the absence of clouds. It is evaluated in ten separate ways according to the literature formulation of the ten simplified models. The complete list of parametrizations implemented are presented in *Figure 4.2* (Formetta et al., 2016) showing the formulations and in *Figure 4.3* (Formetta et al., 2016) showing the different parameters X, Y and Z used in the formula.

| # | Component Name | Formulation | Reference |
|---|---|---|---|
| 1 | Angstrom | $\epsilon_{clear} = X - Y \cdot 10^{Z \cdot e}$ | Angstrom [1918] (2) |
| 2 | Brunt's | $\epsilon_{clear} = X + Y \cdot e^{0.5}$ | Brunt's [1932](3) |
| 3 | Swinbank | $\epsilon_{clear} = X \cdot 10^{-13} \cdot T_a^6$ | Swinbank [1963](4) |
| 4 | Idso and Jackson | $\epsilon_{clear} = 1 - X \cdot exp(-Y \cdot 10^{-4} \cdot (273 - T_a)^2)$ | Idso and Jackson [1969](5) |
| 5 | Brutsaert | $\epsilon_{clear} = X \cdot (e/T_a)^{1/7}$ | Brutsaert [1975] (6) |
| 6 | Idso | $\epsilon_{clear} = X + Y \cdot 10^{-4} \cdot e \cdot exp(1500/T_a)$ | Idso [1981] (7) |
| 7 | Monteith and Unsworth | $\epsilon_{clear} = X + Y \cdot \sigma \cdot T_a^4$ | Monteith and Unsworth [1990] (8) |
| 8 | Konzelmann | $\epsilon_{clear} = X + Y \cdot (e/T_a)^{1/8}$ | Konzelmann et al [1994] (9) |
| 9 | Prata | $\epsilon_{clear} = [1 - (X + w) \cdot exp(-(Y + Z \cdot w)^{1/2})]$ | Prata [1996](10) |
| 10 | Dilley and O'brien | $\epsilon_{clear} = X + Y \cdot (T_a/273.16)^6 + Z \cdot (w/25)^{1/2}$ | Dilley and O'brien [1998] (11) |

**Table 1 Clear sky emissivity formulations: $T_a$ is the air temperatue [K], w $[kg/m^2]$ is precipitable water = 4650 $[e_0/T_a]$ and e [kPa] screen-level water-vapour pressure.**

*Figure 4.2, clear sky emissivity formulations available in the software implementation. Credit: Formetta et al. 2017 [27]*

| Component Name | X | Y | Z |
|---|---|---|---|
| Angstrom | 0.83 | 0.18 | −0.07 |
| Brunt?s | 0.52 | 0.21 | [−] |
| Swinbank | 5.31 | [−] | [−] |
| Idso and Jackson | 0.26 | −7.77 | [−] |
| Brutsaert | 1.72 | 7 | [−] |
| Idso | 0.70 | 5.95 | [−] |
| Monteith and Unsworth | −119.00 | 1.06 | [−] |
| Konzelmann et al | 0.23 | 0.48 | [−] |
| Prata | 1.00 | 1.20 | 3.00 |
| Dilley and O'brien | 59.38 | 113.70 | 96.96 |

**Table 2 Models parameters values as presented in their literature formulation.**

*Figure 4.3, models parameters values for the models available in the software implementation. Credit: Formetta et al. 2016 [27]*

Since there are no elements to evaluate the clearness index and all the parameters, the *Idso model* (Idso, 1981 [28]) (n. 6 in the table) was used. It is the default model suggested by the software developers.

***Shortwave radiation component***

The shortwave radiation refers to the visible and ultraviolet electromagnetic waves emitted from the Sun and arriving to the Earth. *"In the absence of clouds, solar radiation arrives at Earth's ground surface in two classes*" (Formetta et al. 2013) [29]:

- *Direct radiation* ($S_{\downarrow}$), is that part of the solar beam that arrives at the surface without any interaction with Earth's atmosphere (Formetta et al. 2013) [29],
- *Diffuse radiation* (d $_{\downarrow}$), is shortwave radiation scattered downwards back to Earth's surface after hitting molecules of the atmospheric gases and aerosols (Formetta et al. 2013) [29].

GEOframe uses the formulation given by Javier Gonzales Corripio in 2002 (Corripio, 2002 [30]) and in 2003 (Corripio, 2003 [31]). Under cloudless sky conditions the incident shortwave $S_{\downarrow}$ is given by the *Equation (4.4)* (Formetta et al. 2013):

$$S_{\downarrow} = C_1 * I_{sc} * E_0 * cos(\theta_s) * (T_s + \beta_s) * \psi \quad (4.4)$$

where:

- $C_1 = 0.9751$ is the fraction of solar radiation that is included between 0.3 and 3.0 μm wavelengths,

- $E_0$ [–] is a correction factor related to Earth's orbit eccentricity computed according to Spencer (1971),
- $I_{sc}$ [W * $m^2$] is the sola constant ~ 1367 [$Wm{-}2$Wm−2],
- $T_s$ [-] is the product of the atmospheric transmittance,
- $B_s$ [m] is a correction factor for the increased transmittance with elevation z[m] defined according to Corripio (2002),
- θs [rad] is the angle between the Sun vector and the surface plane,
- ψ is the shadows index those accounts for the sun or shadow of the point under analysis.

Concerning the modelling of the diffusive part $d_{\downarrow}$, it is defined by Iqbal (1983) [32], reported in *Equation (4.5)*.

$$d_{\downarrow} = (\, d_{\downarrow r} + \, d_{\downarrow a} + \, d_{\downarrow m}) * V_s$$

*(4. 5)*

where $d_{\downarrow r}$, $d_{\downarrow a}$ and $d_{\downarrow m}$ are the diffusive irradiance components after the first pass through the atmosphere due to the Rayleigh scattering, the aerosol scattering and multiple reflection, respectively. Finally, $V_s$ is the sky view factor.

At the end, the shortwave radiation component is given by the sum of the direct and diffusive part.

***Net radiation balance component***

The net radiation is defined as the balance between the incoming and the outgoing radiation. In the GEOframe component it is defined as the local difference between downwelling radiation and upwelling radiation (Bancheri 2016) [33]. It is the final balance of the shortwave and longwave components evaluated previously as reported above. The formulation implemented in the software is reported in *Equation (4.6)* (Bancheri 2016).

$$R_n = (\, 1 - \alpha) * (R_s + d) + \, RI_d - \, RI_u$$

*(4. 6)*

where:

- $R_n$ [W * $m^2$] is the net radiation,
- α [-] is the albedo,
- $R_s$ [W * $m^2$] is the direct shortwave radiation,
- d [W * $m^2$] is the diffuse shortwave radiation,
- $RI_d$ [W * $m^2$] is the longwave radiation released by the atmosphere,
- $RI_u$ [W * $m^2$] is the longwave radiation released by the soil.

The net radiation timeseries are produced for each sub-basin by the software and they are the input files used to estimate the evapotranspiration.

### 4.1.1 Radiation component implementation and results

Since the radiation data availability for the *Dora Baltea Basin* were too few to make a good estimation, the components of the software were used to evaluate the radiation timeseries for each sub-basin.

The theory reported above of the radiation values estimations is applied in the GEOframe software using only a simulation. Firstly, the Java codes for each sub-basins are generated using a *Jupyter* notebook [53] provided by the developers, in which all the inputs are defined and the parameters of the model chosen are set (see *Figure A1.9 in Appendix A1*). Then, using the provided code able to run in cascade all the sub-basins simulation, the radiation timeseries are provided. For each sub-basin 7 timeseries are created as output: four for the shortwave components (diffuse part, direct part, the sum of the two components and the values at the top of the atmosphere), two for the longwave component (upwelling and downwelling), and, finally, one for the net component.

The inputs required for each sub-basin, are only the kriged temperature timeseries, the sky view raster file obtained during the geomorphology analysis (*see 2.2.2*), the *pitfilled* DEM and the position of the centroid where the radiation is evaluated.

Concerning the longwave radiation component, the *Idso model* (Idso, 1981 [28]), as reported in 4.1.1, was used. The software developers recommend it, and the parameters set are reported in *Table 4.1*.

*Table 4.1, parameters of the Idso model used in the GEOframe software.*

| Parameter | Value |
|---|---|
| **X** | 0.7 |
| **Y** | 5.95 |
| **$\varepsilon_s$** | 0.98 |
| **a** | 0 |
| **b** | 1 |

The coefficient a and b in the *Equation (4.3)* takes into consideration the presence of clouds. For simplicity, the coefficients were set to 0 and 1 to consider only the emissivity under cloudless sky.

Instead, concerning the shortwave radiation component, in the code (see *Figure A.9 in Appendix A1*) three parameters are set to evaluate the timeseries: ozone layer, albedo and amount of km visible. Even in this case, the parameters were set as default, and they are reported in *Table 4.2*.

*Table 4.2, parameters of the shortwave radiation formulation used in the software.*

| Parameter | Value |
|---|---|
| **Ozone layer (pCmO3) [cm]** | 0.6 |
| **Albedo (pAlphag) [-]** | 0.9 |
| **Visibility (pVisibility) [km]** | 80 |

The shortwave radiation component takes also in consideration the sky view factor to know which parts are much exposed to the solar radiation.

In *Figure 4.4* the timeseries of the total shortwave radiation component for the entire period of modelling is showed for sub-basin 1078. The sub-basin taken into consideration has a centroid's elevation of 231 m.a.s.l. and a high sky view factor value. In *Figure 4.5*, the same plot is showed for sub-basin 82 which has a centroid's elevation of 2174 m.a.s.l. and a low sky view factor value. Clearly, since we are modelling daily radiation at noon, the overall estimated values are high. However, the seasonal pattern is complied with since we have the lowest values during the winter period and the highest values during the summer period in both cases. Furthermore, since sub-basin 82 is located at higher elevation, it shows lower values due to its position lowering the type of exposition to the solar radiation.

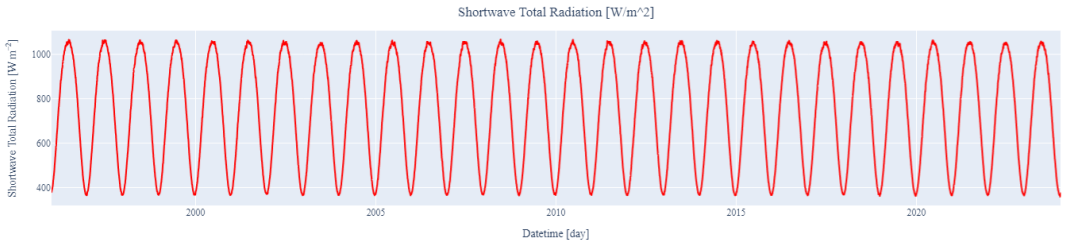



*Figure 4.4, total shortwave radiation timeseries for the sub-basin 1078*

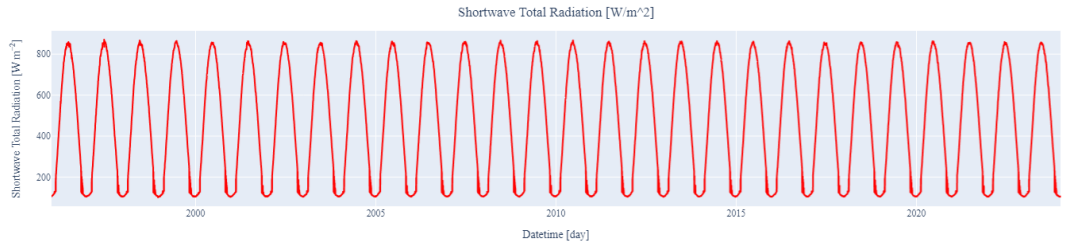



*Figure 4.5, total shortwave radiation timeseries for the sub-basin 82*

It is clear that the radiation values in the real world have a huge variation during the day, especially they change a lot during night and day. However, since the time step of the modelling of the *Dora Baltea Basin* was set as daily, the radiation values were computed at noon. Compared to hourly measured data, values at noon correspond to one of the highest net radiation values of the day.

In the *Dora Baltea Basin* are available radiometers able to detect the hourly total shortwave radiation. To evaluate the accuracy of the data, four stations were taken as reference and a comparison with the modelled data was done.

The first example of comparison is reported below whereas the second one can be found in *Figure A1.10 and Table A1.5 in Appendix A1.*

Sub-basin 481 is in the central part of *Valle d'Aosta* and its centroid is located near the station 4090 (*Saint-Christophe – Aeroporto*) where a radiometer is located. *Figure 4.6* shows the locations. Sub-basin 481 has a centroid's elevation of 1966 m.a.s.l. instead the station 4090 is located at 545 m.a.s.l.

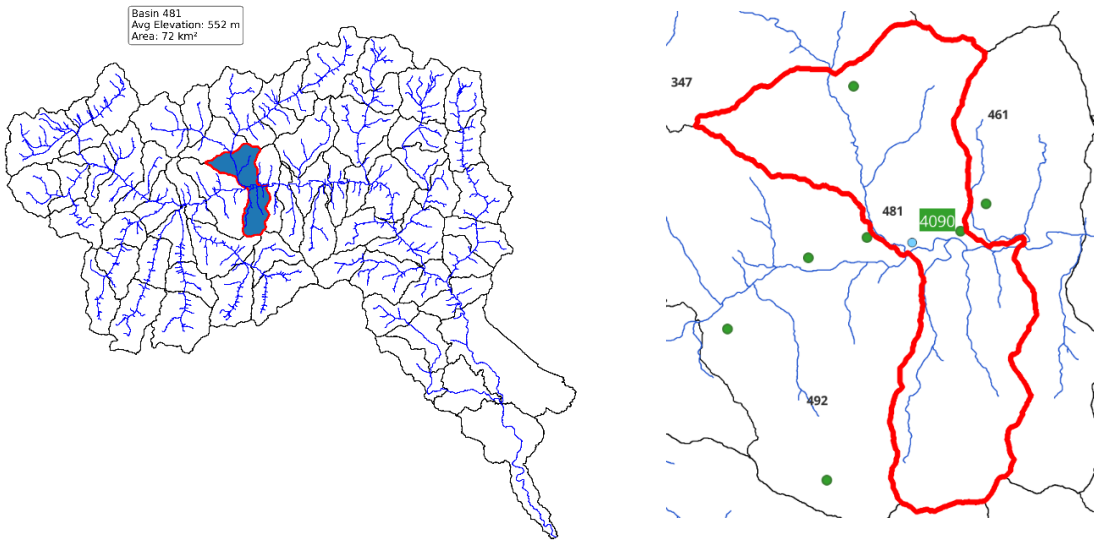



*Figure 4.6, on the left map of the entire Dora Baltea basin with the sub-basin 481 highlighted in blue. On the right the zoom on the sub-basin 481 with the location of its centroid (blue dot) and the meteorological station n. 4090 (green dot with ID label)*

The station and the centroid's sub-basin taken into consideration are in the flat part of the territory and even they are not in the same position, the sky view factor, which plays a key role in the total shortwave radiation estimation, is basically the same.

Looking at the timeseries of the total shortwave radiation balance produced by the software in sub-basin 481, the average values can be evaluated. The average value over the year 2022 and 2023 is of 739.40 W/m$^2$.

The hourly dataset for the year 2022 and 2023 of the radiation values were downloaded from "*Centro Funzionale Regione Autonoma Valle d'Aosta*" [4]. *Figure 4.7* shows the histogram of the mean values recorded in 2022 and 2023 divided by hours. The value at noon is highlighted.

The simulated total shortwave radiation is 183 W/m$^2$ higher than the one measured at the same hour. This difference could be explained considering that the values were estimated in a cloudless condition sky. So, the effects of the clouds were not taken into consideration. Clouds has a key role in absorbing the radiation and so decrease the amount

of shortwave component arriving to the Earth surface. For that reason, the estimated values were considered acceptable. Furthermore, as it was reported above, the centroid's sub-basin and the station are not in the same position, so this also influence a bit the estimation.

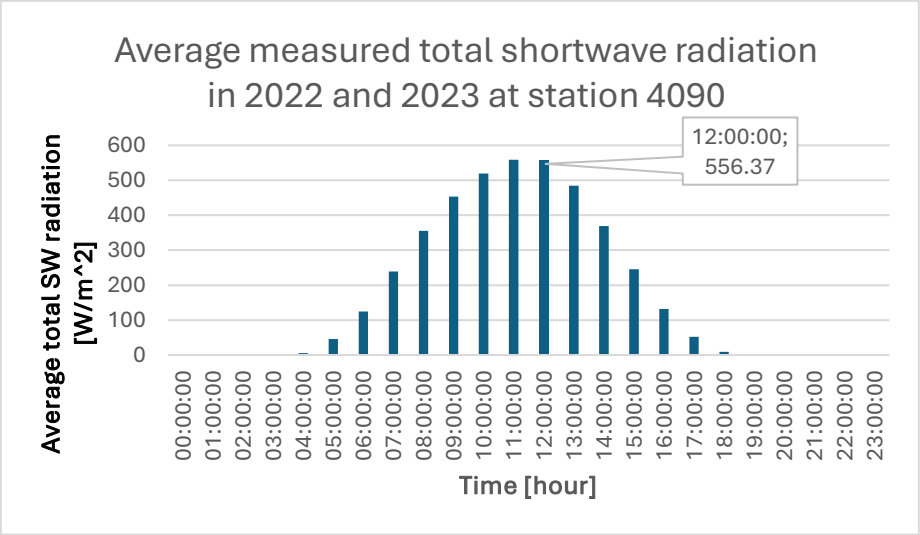



*Figure 4.7, average measured total shortwave radiation in 2022 and 2023 at station n. 4090*

Nevertheless, snowmelt and evapotranspiration processes are highly dependent on the time of the day. The values obtained are clearly an estimation since a daily scale was used. For a better modelling, an hourly temporal resolution could be used in order to capture the diurnal patterns. Another option is also to use smaller sub-basins to give the possibility to capture better the difference in the exposition and therefore capture better the radiation values. It is clear that both options require a greater computational demand, and the aim of the entire case study was to reduce it but maintaining satisfactory results.

## 4.2 Evapotranspiration estimation in GEOframe

In this paragraph, the evapotranspiration estimation is reported. In the first part a brief introduction and description of the methods used by the software GEOframe is made, in particular with focus on the method used for the current study case. In the second part the results and their interpretation are showed.

### 4.2.1 Evapotranspiration component implementation: the Penman-Monteith FAO method

GEOframe offers three distinct types of models to simulate the evapotranspiration values: Priestley-Taylor, Penman-Monteith FAO and Penman-Monteith with different vegetation covers for each sub-basin. All of them are implemented in the component and the choice is given to the user. In the *Dora Baltea Basin*, the Penman-Monteith with different vegetation covers has been used since is the only one that give the possibility to make more heterogeneous the vegetation parameters, whereas improving the overall result.

Penman-Monteith method is born to "*the need to standardize one method to compute reference evapotranspiration ($ET_0$) from meteorological data*" (Allen et al. 1998) [34].

The method gives good results, and it is recommended by FAO (Food and Agriculture organization of the United Nations) as the "*sole standard method*" [34].

The original Penman-Monteith formulation was uploaded in 1998 by FAO by integrating the original form with the equations of the aerodynamic and surface resistance. The final method to estimate $ET_c$ given by FAO can be used to evaluate also the crop evapotranspiration which is the value for a crop in standard conditions with healthy growth and no pests. It is obtained by multiplying $ET_0$ by the crop coefficient $K_c$. The final formulation can be derived in *Equation (5.1)* (Allen et al. 1998) [34].

$$ET_c = K_c \frac{0.408 * \Delta * (R_n - G) + \gamma * \frac{C_p}{T + 273} u_2 * (e_s - e_a)}{\Delta + \gamma * (1 + C_d * u_2)}$$

with:

$$G = f(x) = \begin{cases} G_{morn} * R_n, & daylight \\ G_{night} * R_n, & nightime \end{cases}$$

*(5.1)*

where:

- $ET_0$ is the reference evapotranspiration [mm * day$^{-1}$],
- $K_c$ is the crop coefficient [-],
- $R_n$ is the Net radiation at the crop surface [MJ * m$^{-2}$ * day$^{-1}$],
- G is the soil heat flux density [MJ * m$^{-2}$ * day$^{-1}$],
- T is the mean daily air temperature at 2 m height [°C],
- $u_2$ is the wind speed at 2 m height [m * s$^{-1}$],
- $e_s$ is the saturation vapour pressure [kPa],
- $e_a$ is the actual vapour pressure [kPa],
- $e_s - e_a$ is the saturation vapour pressure deficit [kPa],
- $\Delta$ is the slope vapour pressure deficit [kPa * °C$^{-1}$],
- $\gamma$ is the psychrometric constant [kPa * °C$^{-1}$],
- $C_d$ is the coefficient equal to 0.34 and $C_p$ is a coefficient equal to 900 in the case of a daily time step and equal to 37 in the case of an hourly time step.

The formula reported above is implemented in the GEOframe ET component and gives the possibility to determine the $ET_c$ and heat-flux values for each sub-basin. Some details about the inputs used are reported below.

The net radiation is given in the software component in timeseries of [W/m$^2$] values, but since the formula reported above requires values in [MJ * m$^2$], the conversion is necessary, and it is automatically done by the software.

The soil heat flux density G is split between night and day since the contributing values change during the day. However, if the time resolution is set at daily, this distinction is not done.
The timeseries required by the formulation for the wind velocity, the relative humidity, the soil heat flux, and the atmospheric pressure are not always easily available. For that reason, GEOframe gives the possibility to give as input null timeseries from the start date of the modelling period till the end date. The software can evaluate them using the elevation of each sub-basin or, in the case of the wind velocity, it is set a constat value. As done for the kriging interpolation, the elevation can be chosen between the one evaluated in the centroid or the average one.
Instead, the crop coefficient is given as input as a constant value for every sub-basin if the second model with homogeneous vegetation cover is chosen or with a csv file in which the crop height and the $K_c$ is specified for each sub-basin, making a heterogeneous distinction. FAO has also tabled different values of crop coefficient in the three different stage of growth – initial, middle and end - of different crops [35].

### 4.2.2 Dora Baltea Basin evapotranspiration results

Regarding the Dora Baltea Basin, Penman-Monteith FAO with different crop coefficient has been applied. Since the time step of modelling was set at daily, the distinction between day and night of the soil heat flux density G is not done.

Wind speed, atmospheric pressure, relative humidity, and the soil heat flux have been given as null timeseries inputs since no dense network was available to retrieve this data sets. The elevation used for each sub-basin was the average one for the same reason of the Kriging interpolation (*see 3.2*).

The crop file given as inputs to make the estimation less homogeneous was constructed based on the land cover map available for the territory. Regarding the *Valle d'Aosta* region, the data have been extracted from the "Geoportale SCT – Sistema delle Conoscenze Territoriali" [36] where different land classifications are available from remote sensing data. The same procedure was done for the remaining part of the Piemonte region where the data have been extracted from "Geoportale Regione Piemonte" [37].

Looking at the land covers maps, all sub-basins are quite heterogenous, and they occupy an area between 50 and 80 $km^2$ each. For that reason, a rough estimation of the crop height and coefficient is needed. Based on the FAO table [35], the main distinction was made between the mountain regions where glaciers and permanent snow are present and flat regions where there are much more forests and agricultural cultivations.

In almost all the alpine sub-basins the only vegetation that can be found has been classified by the land cover as alpine meadows. Indicatively, the crop coefficient values are around 0.3 and 0.6 depending on the stage. It reflects the reduced evapotranspiration due to the low temperature during all the year. The remaining part in that regions are covered by snow, ice and rocks which lowering down the crop coefficient. A mean value of 0.3 and 0.5 has been used in all the alpine sub-basins. The crop coefficient has been increased approaching lower elevation and so having more vegetation.

In the flatter regions, low altitude meadows and forests of conifers or wooden tree are the most present. Approaching the urbanized zones the vegetation is lower but still present. In the earlier and middle stage of growth of these cultivations the crop coefficient is between 0.8 and 1.2 according to the FAO estimation values [35]. Even in this case, a rough average was done looking at the land cover map and a range between 0.5 and 1.2 was used. The crop coefficient has been increased approaching the lowest elevation, especially in the Piemonte region where agricultural cultivations are much more present.

The crop coefficient values were kept low since an average of the entire year is considered, thus without using the three distinct stages of the crop growth.

Regarding the crop height which is also defined in the same file for each sub-basin, the values were set constant at 1.2 m. Evaluating the difference between all the crop height was considered not easy since a unique value for the entire year and sub-basin is used. As result, for simplicity, a homogeneous value was used.

Initially, the values of $K_c$ were set too low, and the results were not reliable comparing them with data available by ISPRA (*Istituto Superiore per la Protezione e la Ricerca Ambientale*) report [38]. *Figure A1.11 in Appendix A1* shows the mean yearly $ET_c$ for each sub-basin obtained. The ranges obtained during the first trial were 100 -700 mm/year.

In order to have better and comparable results, the $K_c$ values were increased of 0.2 and the overall result improved. *Figure 4.8* shows the map of the mean yearly evapotranspiration ranges for each sub-basin and *Figure 4.9* reports the same map of the ISPRA report where the mean was done on a larger temporal scale (from 1951 to 2019).

The values obtained during the second evaluation ranges between 255 mm/year in the alpine sub-basins and 1184 mm/year in sub-basin with lowest elevation. The map given by ISPRA reports the mean yearly $ET_0$ values of the entire territory of Italy and in the Dora Baltea Basin they range from 200 mm/year to 800 mm/year. Since the crop evapotranspiration is given by the multiplication of the $ET_0$ by the crop coefficient used, the $ET_c$ ranges are higher in the flatter zones since the $K_c$ is around 1. Even though the comparison uses different estimation period and $ET_0$ instead of $ET_c$, it has been considered a good reference to compare the one obtained.

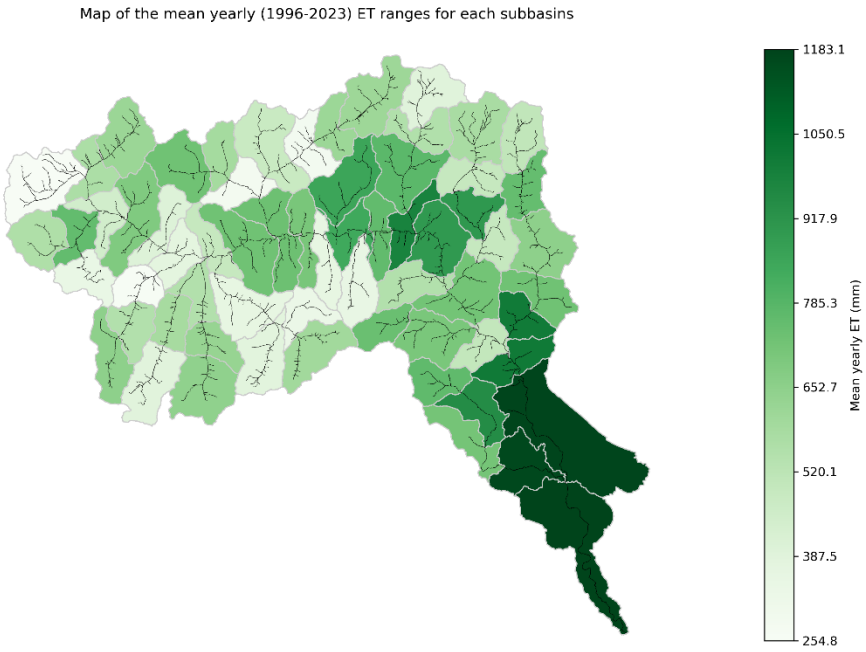



*Figure 4.8, map of the mean yearly evapotranspiration for each sub-basin*

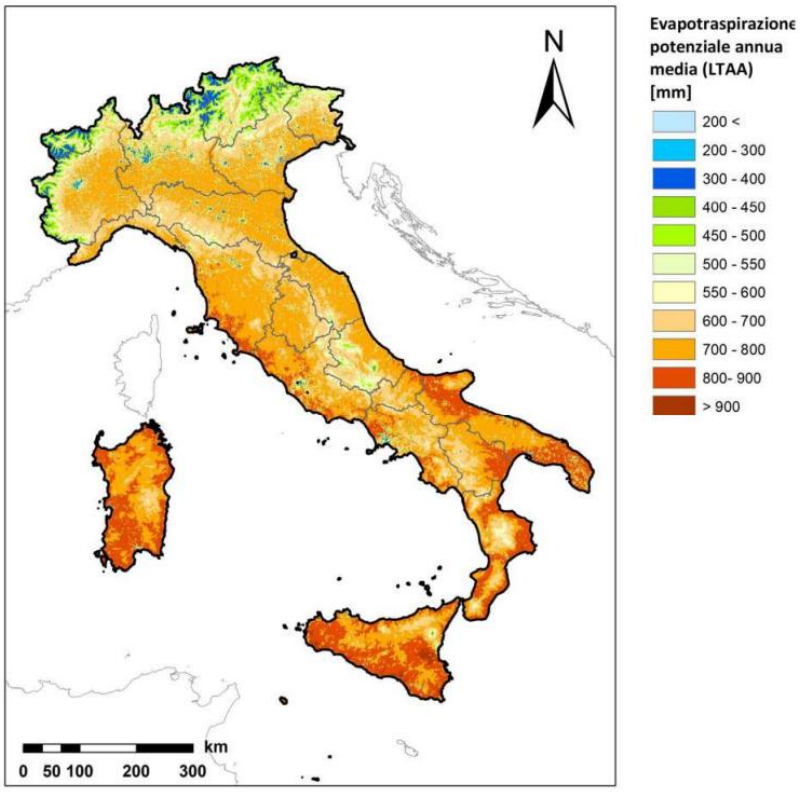



*Figure 4.9, mean yearly ET over Italy from 1951 to 2019. Credit: isprambiente.it [38]*

Another comparison was made using the ERA5 Copernicus Climate Change Service (C3S) [39], where the $ET_0$ was evaluated during March 2022. The entire Valle d'Aosta region has a mean value between 25 and 50 mm/month, as *Figure 4.10* shows.

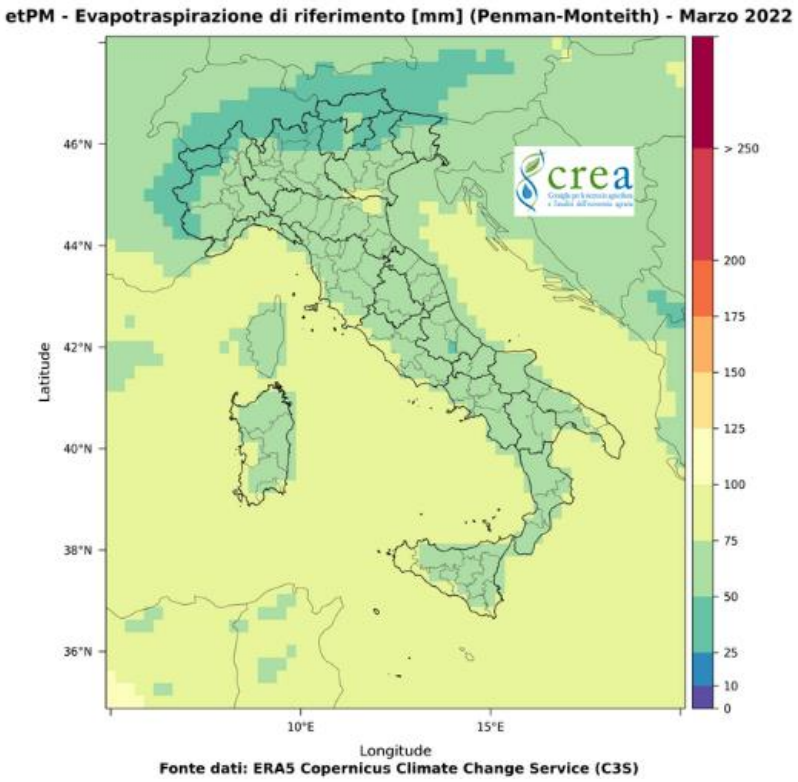



*Figure 4.10, reference evapotranspiration evaluated by ERA5 Copernicus. Credit: pianetapsr.it [39]*

To compare this value the sub-basin 481 located in the central part of the Valle d'Aosta region was taken as example. *Figure 4.11* shows is location, its area, and the average elevation.

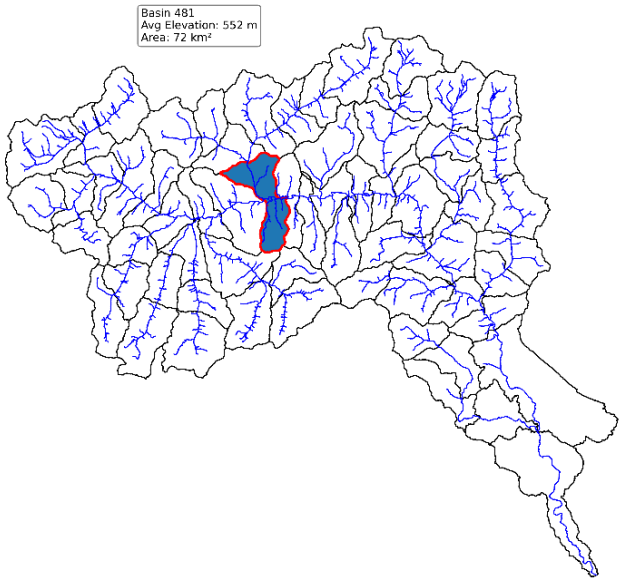



*Figure 4.11, map of the entire Dora Baltea basin with the sub-basin 481 highlighted in blue*

*Figure 4.12* reports the timeseries of the ET values simulated in the sub-basin 481. The graph shows the general path of the ET having lower values during the winter period and higher values during the spring when the climate conditions make the vegetation grow

and so higher evapotranspiration values. Looking at the monthly $ET_c$ simulated in March 2022, the cumulated value over the entire month is of 51.5 mm/month. Considering that it is the crop evapotranspiration and not the reference evapotranspiration as the ERA5 Copernicus evaluated, the simulated value is a little bit above the range indicated and it was considered a good result.

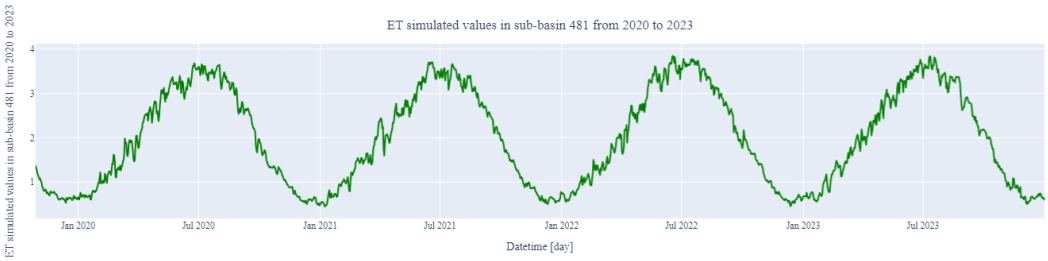



*Figure 4.12, simulated evapotranspiration timeseries for sub-basin 481 from January 2020 to July 2023*

*Figure 4.13* shows the mean annual cumulated ET for each sub-basin. It clearly follows the same yearly trend of the temperature interpolated (*see Figure 3.15*). The graph shows an evident correlation between the evapotranspiration and temperatures, having higher values of ET during the drier years like the 2003 – when a heatwave hit the whole Italy. It is also clear that precipitation is related to ET since the water is needed to make the process happens. However, it was found out that "*during the 2003 heatwave, evapotranspiration in large area over the Alps was above average despite low precipitation*" (Mastrotheodoros et al. 2020) [40]. Indeed, the results of the analysis seem to follow this trend, having less precipitation but high ET during the drier year.

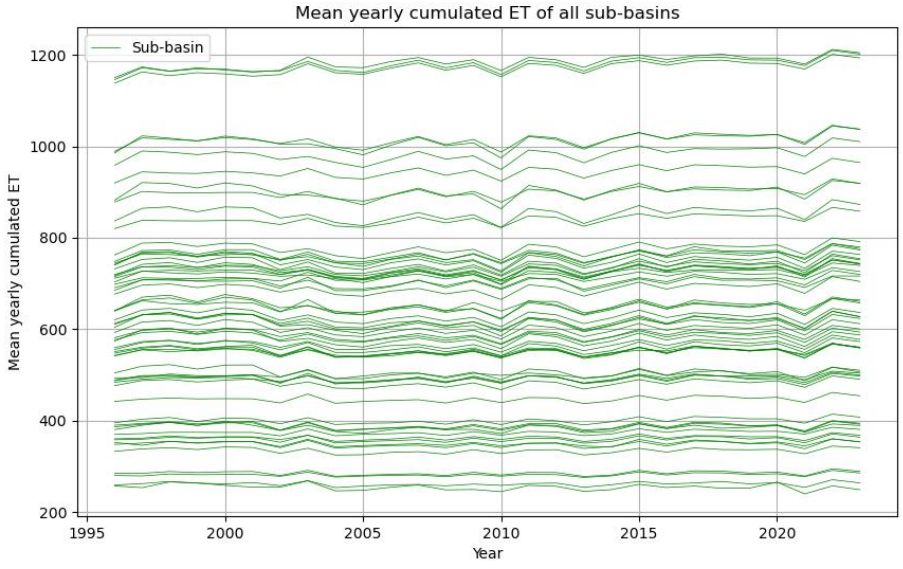



*Figure 4.13, yearly cumulative ET of all sub-basins in the Dora Baltea catchment*

The overall result of the ET estimation seems to be good looking at the comparison. It is clear that the values have been obtained based on radiation estimated in the previous step which are also an estimation since a cloudless condition has been used. Furthermore, a daily time scale was used increasing the assumptions and having average results. During the first trial of estimation the values obtained were underestimated due to low crop coefficients. However, increasing them of 0.2 gave much better results. Evapotranspiration plays a significant role in the water balance, and it is also one of the calibrated parameters in the last step of the modelling, having the possibility to get better results. The analysis of the contribute to the water balance and the water sink of ET is reported further ahead during the calibration results interpretation (see 5.2).

## 4.3 Embedded reservoir model (ERM) component

In this last section dedicated to the GEOframe estimation components the embedded reservoir model implemented inside the software is presented. In the first paragraph the functioning of the system is briefly described whereas in the second paragraph the first discharge timeseries obtained are reported.

### 4.3.1 Discharge calculation through ERM

The last step of the GEOframe model is the runoff calculation carried out with a discretization of all the HRUs using a reservoir system: the embedded reservoir model (ERM) [41]. ERM "*schematizes each HRU as a group of storages (reservoirs) and solves the water budget for each one*" (Bancheri, Rigon, and Manfreda 2020) [6]. Each HRU is defined as a series of reservoir accumulating water arriving from the precipitations and successively releasing discharge. All the discharges are released when the threshold capacity – set with a series of parameters - of each reservoir is achieved. The mathematical equations describe in a simplified way all the parts of the hydrological cycle. The ERM structure is represented in *Figure 4.14* [42]. All the different hydrological processes are organized in five integrated reservoirs: snow (white reservoir), canopy (green reservoir), rootzone (orange reservoir), groundwater (blue reservoir) and runoff (red reservoir).

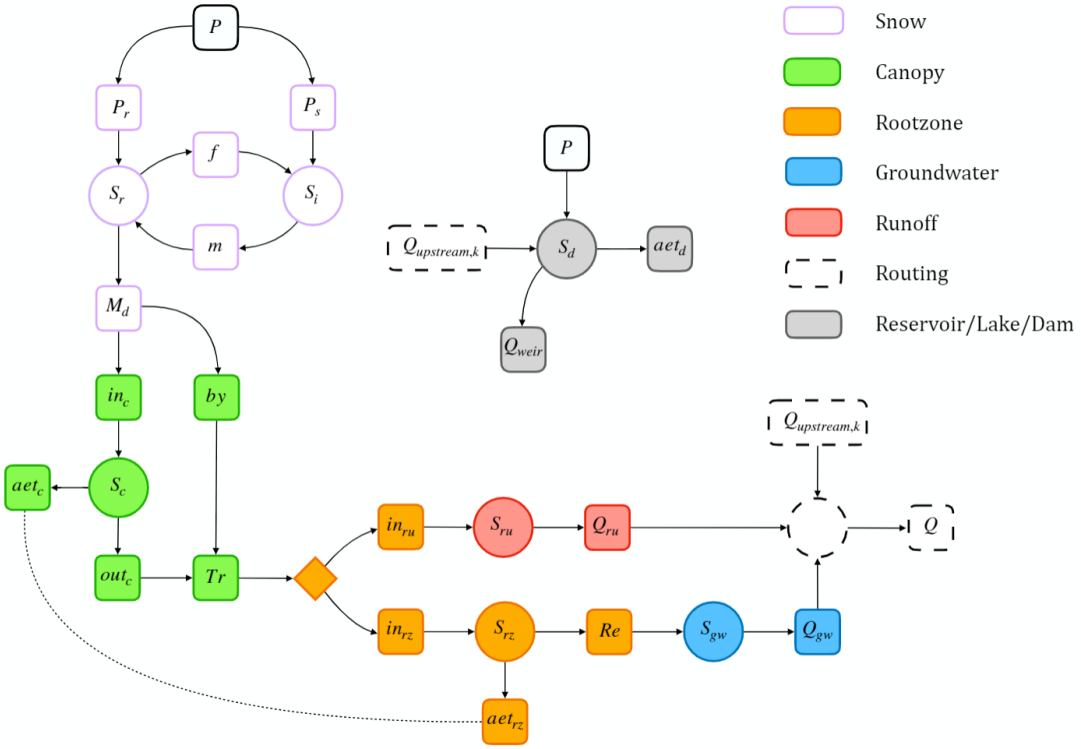



*Figure 4.14, structure of the runoff generation in the ERM in GEOframe system. Credit: Busti, Andreis, Formetta and Rigon 2021 [42]*

Each layer represented in *Figure 4.14* is like a node crossed by incoming and outgoing fluxes. For each reservoir, the ruling equations are established to manage the water budget, and they are available in details in *Appendix A2* [42]. Instead, a schematic description is reported below [6].

Firstly, precipitation is divided into the liquid component ($P_r$) and solid component ($P_s$) thought a rain-snow separator. Then, the rainfall ($S_r$) and snow storage ($S_s$) are evaluated, providing as output the melting discharge which is the input/source of water. The melting discharge ($M_d$) is firstly available for the canopy, which has a storage ($S_c$). The water coming in can rather enter the storage or bypass it reaching the ground (Tr). From the storage the actual canopy evapotranspiration ($aet_c$) is evaluated together with the canopy reservoir outflow ($out_c$) which reaches the ground adding to the by-passed water. As output the throughfall can rather remain on the surface contributing to the runoff storage ($S_{ru}$) and then to the runoff discharge ($Q_{ru}$) or percolate into the rootzone ($in_{rz}$). The water entering the rootzone storage ($S_{rz}$) contributes to the actual rootzone evapotranspiration ($aet_{rz}$) and to the groundwater recharge (Re). The groundwater maximum storage ($S_{gw,max}$) is recharged and produces the deep groundwater discharge ($Q_{gw}$) as outflow. At the end, the two contributions of the runoff, $Q_{ru}$ and $Q_{gw}$, gives the sub-basin out discharge (Q) as also reported in *Equation (6.1)*. In this step the contribution of the upstream discharge given by lakes or dams can be also considered.

$$Q = Q_{ru} + Q_{gw} \tag{6. 5}$$

In *Table 4.3* are reported all the parameters of the five layers set as input values in the ERM calculation and calibrated in the following step.

In the ERM the "*graph-based structure called NET3 is employed for the management of process simulations. NET3 is designed using a river network/graph structure analogy, where each HRU is a node of the graph, and the channel links are the connections between the nodes*" (Bancheri, Rigon, and Manfreda 2020) [6]. Indeed, the sub-basins are divided into upstream and downstream. In this way, the discharge is computed in parallel for all the sub-basins considered as "leaves," so the one that do not have any sub-basin upstream. Whereas all the sub-basins with tributaries are computed in consecutive steps since the water from upstream is the input necessary to evaluate the discharge of the sub-basin downstream.

*Table 4.3, set of parameters of the ERM component.*

| Layer | Parameter Symbol | Parameter Name | Unit |
|---|---|---|---|
| Snow | $\alpha_r$ | Rainfall measurements errors coefficient | [-] |
| Snow | $\alpha_s$ | Snow measurements errors coefficient | [-] |
| Snow | $T_M$ | Melting temperature | [$\Theta^1$] |
| Snow | $\alpha_m$ | Combined melting factor | [$L^1\ \Theta^{-1}\ T^{-1}$] |
| Snow | $\alpha_f$ | Freezing factor | [$L^1\ \Theta^{-1}\ T^{-1}$] |
| Snow | $\alpha_l$ | Water retention capacity coefficient | [-] |
| Canopy | $k_c$ | Canopy storage coefficient | [-] |
| Canopy | p | Free throughfall coefficient | [-] |
| Rootzone | $S_{RZ,max}$ | Rootzone maximum storage | [$L^1$] |
| Rootzone | g | Groundwater recharge coefficient | [$L^1\ T^{-1}$] |
| Rootzone | h | Groundwater recharge exponent | [-] |
| Rootzone | $p_B$ | Spatial variability of the soil moisture capacity | [-] |
| Runoff | $S_{RU,max}$ | Runoff maximum storage | [$L^1$] |
| Runoff | c | Runoff discharge coefficient | [$L^1\ T^{-1}$] |
| Runoff | d | Runoff discharge exponent | [-] |
| Groundwater | $S_{GW,max}$ | Groundwater maximum storage | [$L^1$] |
| Groundwater | e | Deep groundwater discharge coefficient | [$L^1\ T^{-1}$] |
| Groundwater | f | Deep groundwater discharge exponent | [-] |

### 4.3.2 Dora Baltea ERM initial results

As first round the ERM component has been run to have the initial results of all the layers according to the equations reported in *Appendix A2*. The outputs are 25 timeseries for each sub-basin containing different values for each layer considered in the ERM structure. The most important one is the total discharge at the outlet of each sub-basin.

Each sub-basin has its own simulation code with all the parameters including the area of the HRU. An example part of the code made for each sub-basin is reported in *Appendix B1*. However, as it has been done for the other steps, all the simulations are run in cascade using a unique file.

In this step the 18 parameters reported in *Table 4.3* are required as input, but since the optimal and basin related values are set in the calibration step, the default values were used. They can be retrieved in *Table 4.4,* and they are the one used by the developers for the open-source example basin [43].

*Table 4.4, set of default parameters of the ERM component.*

| Parameter Symbol | Parameter Name | Unit | Default values |
|---|---|---|---|
| $\boldsymbol{\alpha_r}$ | Rainfall measurements errors coefficient | [-] | 0.678 |
| $\boldsymbol{\alpha_s}$ | Snow measurements errors coefficient | [-] | 0.802 |
| $\boldsymbol{T_M}$ | Melting temperature | [$\Theta^1$] | 0.821 |
| $\boldsymbol{\alpha_m}$ | Combined melting factor | [$L^1\ \Theta^{-1}\ T^{-1}$] | 1.47 |
| $\boldsymbol{\alpha_f}$ | Freezing factor | [$L^1\ \Theta^{-1}\ T^{-1}$] | 0.132 |
| $\boldsymbol{\alpha_l}$ | Water retention capacity coefficient | [-] | 0.0932 |
| $\boldsymbol{k_c}$ | Canopy storage coefficient | [-] | 0.701 |
| p | Free throughfall coefficient | [-] | 0.509 |
| $\boldsymbol{S_{RZ,max}}$ | Rootzone maximum storage | [$L^1$] | 100 |
| g | Groundwater recharge coefficient | [$L^1\ T^{-1}$] | 1.22 |
| h | Groundwater recharge exponent | [-] | 1.99 |
| $\boldsymbol{p_B}$ | Spatial variability of the soil moisture capacity | [-] | 7.07 |
| $\boldsymbol{S_{RU,max}}$ | Runoff maximum storage | [$L^1$] | 31.2 |
| **c** | Runoff discharge coefficient | [$L^1\ T^{-1}$] | 1.74 |
| d | Runoff discharge exponent | [-] | 1.26 |
| $\boldsymbol{S_{GW,max}}$ | Groundwater maximum storage | [$L^1$] | 538 |
| **e** | Deep groundwater discharge coefficient | [$L^1\ T^{-1}$] | 1.27 |
| f | Deep groundwater discharge exponent | [-] | 1.20 |

Since the Dora Baltea Basin is a territory which does not have a great contribution of groundwater due to its topography, it is evident that the default groundwater parameters are set too high. Another evidence appears from *Figure B1.1 in Appendix B1* which shows the yearly water balance from all sub-basins and weighted by the area performed before calibration. An excess of water in the groundwater layer is explicit. Indeed, during the calibration step, a significant improvement has been conducted in that layer.

As result the basin outlet discharge has been analysed to look at the accuracy of the results. The simulated discharge timeseries have been compared with the measured one (n.3 – *Verolengo*). As reported in *Appendix B2*, the Kling-Gupta efficiency (KGE) [44] is the objective function (OF) used for the calibration process and it evaluates the adherence between the simulated and the measured discharge timeseries. *Figure 4.15* shows the comparison between the simulated and measured timeseries of the discharge at basin outlet over the entire modelling period (1996-2023). The correspondent KGE is also reported.

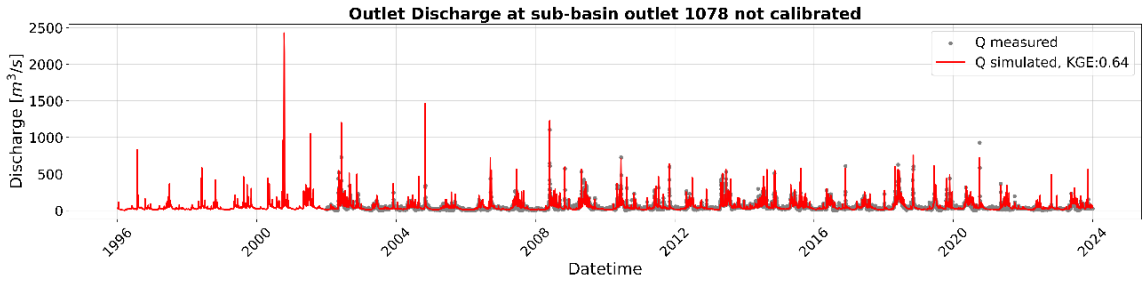



*Figure 4.15, comparison of the simulated and measured timeseries of the sub-basin outlet 1078 with KGE value*

A KGE of 0.64 has been obtained and since the parameters are not calibrated, the result can be considered good. Between 1996 and 2002 no data are available at station n.3, so the comparison cannot be done. From 2002 on, the peaks seem to be respected although some are overestimated like in 2005 and some underestimated like in 2021. To have a better view a zoom on the year 2023 has been done and reported in *Figure 4.16*.

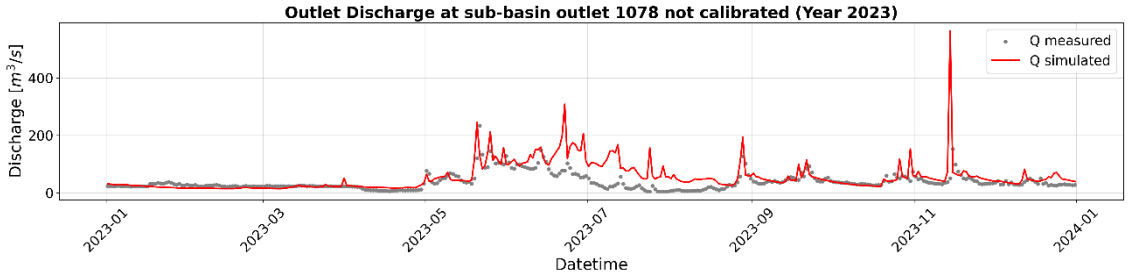



*Figure 4.16, comparison of the simulated and measured timeseries of the sub-basin outlet 1078*

The mean low value seems to be modelled in an effective way as the peaks, despite the peak in November which is clearly overestimated. During the summer period a general overestimation is evident due probably to the drawn water by *Canale Farini* which is not taken into consideration.

The subsequent step of calibration was carried out to fine-tune the accuracy of the outputs by determining the most appropriate set of parameters relevant to the topography of the study area. The final aim is to get a better match between the simulated and the observed discharge timeseries. A post-processing operation was then carried out on the results to obtain the timeseries of the discharge at the basin outlet, net of the volume taken away by Canale Farini.

# 5 MAIN RESULTS AND DISCUSSION

In this chapter the final main results obtained are showed together with their analysis and discussion. This section is divided into three main parts: the first one is about the explanation of the functioning of the calibration procedure and the set-up for the final step of the *Dora Baltea Basin* is showed together with the obtained results; the second section includes the analysis of the water balance made to access the accuracy of the model developed. Finally, the post processing of the *Canale Farini* discharge is presented.

## 5.1 Calibration procedure

The calibration procedure implemented in GEOframe is called *LUCA (Let Us CAlibrate)*. It is a "*multiple-objective, stepwise, automated procedure for hydrologic model calibration*" (Hay and Umemoto 2007) [45]. It is a research global algorithm and it "*assures that intermediate and final stated of the model are simulated consistently with measured values*" (Hay and Umemoto 2007) [45].

Within the GEOframe software, the calibrator is implemented on the OMS console where a simulation file named *LUCA* is used. The set of 18 parameters to be calibrated (*see Table 4.3*) are defined by an upper and a lower limit. Then, the algorithm generates random points inside the range while the Objective Function (OS) – that is KGE - is used to access an evaluation criterion for each of those points. This process is iterated for a number of steps – set of parameters - and for a number of rounds – execution of one or more steps - defined by the user, as the schematic structure shows in *Figure 5.1* (Bancheri 2017) [41]. At the end of each round the set of parameters with the highest KGE value is printed in a txt file. In the event that the value of a round produces a lower KGE value than the previous round, then the set of parameters printed is the same as the precious round since it has the highest KGE. *Figure B1.5 in Appendix B1* shows the structure of the code implemented where in the first 3 lines are defined: the output folder in which are printed the resulting parameters, the topology file to be used and the ERM simulations files created before for each sub-basin. Then, the calibration start date is reported together with the start date and end date of the entire modelling. Here, the number of rounds is defined, then the upper and lower limits of each parameter are set. The last part defines the objective function KGE, together with the files containing the discharge timeseries simulated in the basin outlet and the discharge timeseries observed in the reference gauging station. The last line defines the number of steps for each round.

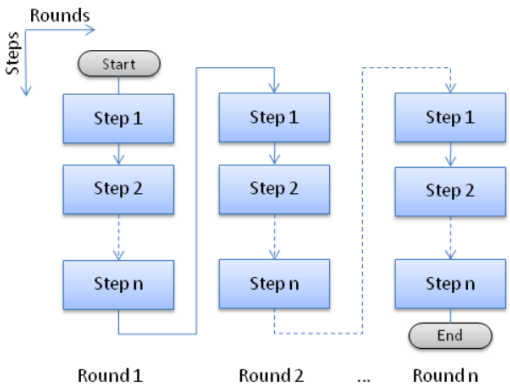



*Figure 5.1, schematic representation of LUCA's rounds and steps. Credit: Bancheri 2017 [41]*

The immediate subsequent step is the validation phase where the optimal parameters values of the calibration step are used to simulate the discharge timeseries in the modelling period not calibrated. This phase is fundamental to access the forecasting ability of the model.

### 5.1.1 Dora Baltea calibration set up and main results

For the Dora Baltea Basin six calibrations have been performed: four upstream and two downstream. In this way, after the upstream calibrations the two downstream calibrations have been run using the parameters fixed in the higher calibrated sub-basins. *Figure 5.2* shows the set-up of the calibrations done on the entire basin area.

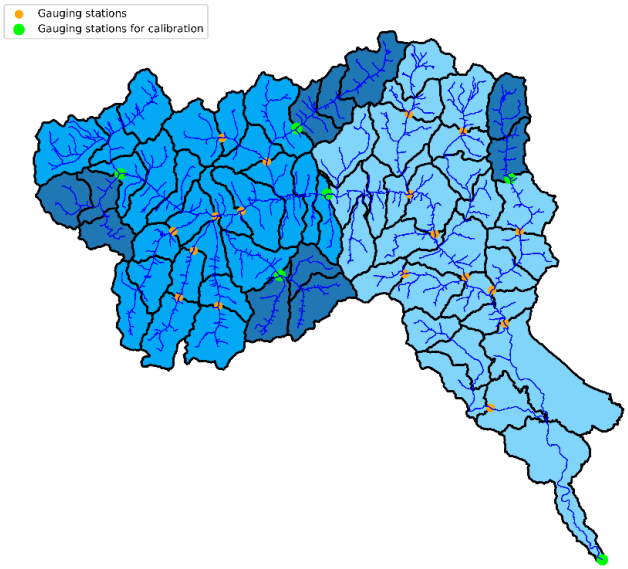



*Figure 5.2, calibrations set-up of the Dora Baltea basin*

The sub-basins highlighted in dark blue identify the upstream calibrated HRUs. They were chosen based on the available gauging stations and with the purpose to have distinct types of calibration in the mountainous regions - where glaciers and snow accumulation play a huge role in the hydrological processes. The four sets of sub-basins are located in the highest regions of the *Valle d'Aosta* territory. Whereas the two downstream

calibrations have been performed over the main course of the river: one located in the middle of the *Valle d'Aosta* territory – highlighted in light blue- and one at the outlet basin – highlighted in lighter blue. *Table 5.1* summarizes the metadata of all the six calibration stations used. In addition, in *Table B1.1 in Appendix B1* all the basin IDs and the correspondent calibration station are reported.

The calibration period was selected looking at the data availability for each station considering also the gaps in the measurements (see *Figure 2.11*). In every particular case the start date has been set on October 1st since it is considered the start of the hydrological year, and the calibration should perform better.

*Table 5.1, calibration stations metadata*

| Station ID | Station Name | Period of data availability | Period used for calibration | Period used for validation | Correspondent sub-basin outlet |
|---|---|---|---|---|---|
| **1490** | Pré Saint Didier Capoluogo | 29/08/2000 31/12/2023 | 01/10/2004 31/12/2023 | 01/01/1996 30/09/2004 | 414 |
| **1640** | Valpelline – Prelé | 01/10/2003 31/12/2023 | 01/10/2005 31/12/2023 | 01/01/1996 30/09/2005 | 217 |
| **1130** | Cogne – Crétaz | 29/08/2000 31/12/2023 | 01/10/2004 31/12/2023 | 01/01/1996 30/09/2004 | 821 |
| **1290** | Gressoney – Saint-Jean – Capoluogo | 22/03/2002 31/12/2023 | 01/10/2004 31/12/2023 | 01/01/1996 30/09/2004 | 393 |
| **1430** | Nus – Les Iles | 01/01/2007 31/12/2023 | 01/10/2009 31/12/2023 | 01/01/1996 30/09/2009 | 471 |
| **3** | Verolengo | 01/01/2002 31/12/2023 | 01/10/2007 31/12/2023 | 01/01/1996 30/09/2007 | 1078 |

Ideally, each available gauging station could be used as calibration station, but the time required to perform them would be extremely high. In order to reduce the computational demand and the time required, six calibrations have been considered as a good compromise.

All the six calibrations have been performed twice: first with only 100 steps and 1 round and then with 200 steps and 3 rounds. This approach has been used to give the possibility to the calibrator algorithm to investigate a major number of set of parameters and, finally, obtain improved results. In *Table 5.2* are summarized the results of the calibration and validation procedures of each case. It evidences the capacity of the calibrator to improve the final results while increasing the number of sets and rounds. Indeed, during the first attempt – which was carried out with a lower number of steps and rounds – the obtained KGE values are lower.

*Table 5.2, comparison of the KGE values for calibration and validation period in the two attempts of calibration*

| Station ID used for calibration | KGE [-] cal. 1st attempt | KGE [-] val. 1st attempt | KGE [-] cal. 2nd attempt | KGE [-] val. 1st attempt |
|---|---|---|---|---|
| **1490** | 0.47 | 0.23 | 0.57 | 0.42 |
| **1640** | 0.51 | 0.46 | 0.60 | 0.17 |
| **1130** | 0.57 | 0.40 | 0.60 | 0.44 |
| **1290** | 0.61 | 0.42 | 0.62 | 0.28 |
| **1430** | 0.76 | 0.71 | 0.79 | 0.81 |
| **3** | 0.67 | 0.59 | 0.68 | 0.67 |

The improvement is particularly evident in the first upstream calibration where the KGE improved of 0.10 (from 0.47 to 0.57). Conversely, the improvement downstream between the first and second calibration is not that strong. For the last calibration including the basin outlet, there was a strong increase during the validation period, from 0.59 to 0.67. However, in both tests, the best performance is for the calibration at station n. 1430, carried out along the main course of the river Dora Baltea. This is also explained by the greater reliability of the data coming from this station because it is placed in a plain zone and the model is able to make corrections that account for all inputs given from the mountain areas nearby. Besides, it becomes even more marked from the first to the second attempt in the validation process, where it increases from 0.71 to 0.81. On the other hand, the values in the upstream KGE during the validation processes are low because of the absence of data from 1996 to 2000 and a larger period used for calibration. For a better interpretation of the results, *Table 5.3,* and *Table 5.4* reports – for both attempts - the model evaluation metrics (see *Appendix B2*) between the observed and simulated discharges for the calibration period. In the last three lines the comparison between the observed and simulated mean, maximum and minimum discharge are reported.

*Table 5.3, model evaluation metrics for the Dora Baltea basin in the first calibration attempt*

| | **1490 cal** | | **1640 cal** | | **1130 cal** | | **1290 cal** | | **1430 cal** | | **3 cal** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Calibration period** | 2004-2023 | | 2005-2023 | | 2004-2023 | | 2004-2023 | | 2009-2023 | | 2007-2023 | |
| **$\sum Q_{obs}$ \ $\sum Q_{sim}$ [-]** | 1.66 | | 1.05 | | 1.36 | | 1.31 | | 0.75 | | 1.05 | |
| **Pearson Correlation [-]** | 0.71 | | 0.64 | | 0.74 | | 0.78 | | 0.76 | | 0.69 | |
| **MAE [$m^3/s$]** | 4.19 | | 0.79 | | 3.48 | | 2.04 | | 9.46 | | 26.2 | |
| **RMSE [$m^3/s$]** | 6.23 | | 1.57 | | 5.41 | | 2.81 | | 17.0 | | 48.4 | |
| **% bias** | -42.5 | | -8.49 | | -29.8 | | -24.6 | | +3.78 | | -4.83 | |
| | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. |
| **Mean Q [$m^3/s$]** | 6.66 | 3.82 | 1.89 | 1.69 | 6.03 | 4.09 | 4.11 | 2.99 | 19.8 | 21.1 | 55.9 | 53.2 |
| **Max. Q [$m^3/s$]** | 59.1 | 130 | 24.2 | 29.6 | 76.3 | 100 | 38.5 | 35.1 | 296 | 294 | 1104 | 812 |
| **Min. Q [$m^3/s$]** | 1.5 | 0.13 | 0.5 | 0.19 | 0.0 | 0.23 | 0.1 | 0.13 | 2.7 | 3.19 | 0.65 | 11.14 |

*Table 5.4, model evaluation metric for the Dora Baltea basin in the second calibration attempt*

| | 1490 cal | | 1640 cal | | 1130 cal | | 1290 cal | | 1430 cal | | 3 cal | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Calibration period** | 2004-2023 | | 2005-2023 | | 2004-2023 | | 2004-2023 | | 2009-2023 | | 2007-2023 | |
| **$\sum Q_{obs} \setminus \sum Q_{sim}$ [-]** | 1.37 | | 1.05 | | 1.40 | | 1.21 | | 0.74 | | 1.02 | |
| **Pearson Correlation [-]** | 0.73 | | 0.61 | | 0.76 | | 0.75 | | 0.80 | | 0.72 | |
| **MAE [$m^3/s$]** | 3.90 | | 0.72 | | 3.31 | | 1.96 | | 9.04 | | 25.3 | |
| **RMSE [$m^3/s$]** | 5.82 | | 1.38 | | 5.02 | | 2.94 | | 16.0 | | 45.3 | |
| **% bias** | -30.2 | | -8.37 | | -31.7 | | -18.5 | | +5.67 | | -2.24 | |
| | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. | Obs. | Sim. |
| **Mean Q [$m^3/s$]** | 6.66 | 4.64 | 1.89 | 1.70 | 6.03 | 3.97 | 4.11 | 3.23 | 19.8 | 21.4 | 55.9 | 54.6 |
| **Max. Q [$m^3/s$]** | 59.1 | 107 | 24.2 | 27.3 | 76.3 | 82.0 | 38.5 | 46.4 | 296 | 267 | 1104 | 781 |
| **Min. Q [$m^3/s$]** | 1.5 | 0.11 | 0.5 | 0.33 | 0.0 | 0.17 | 0.1 | 0.1 | 2.7 | 1.31 | 0.65 | 10.60 |

The difference between the two attempts is noticeable, showing a general improvement. Looking at the water balance, which compares the sum of the observed and simulated discharge values, the calibration performed at station n. 1640, 1430 and 3 have results closest to 1 – with the highest reached during the second attempt in the last calibration having 1.02 as result. In the second try, the values have been improved, passing from 1.66 to 1.37 in the first calibration (n. 1490). The Pearson Correlation indicates, with a number close to 1, a good positive correlation; in the second attempt it ranges between 0.61 and 0.80, demonstrating an acceptable goodness of fit. Instead, the MAE estimates the average of absolute errors, thus a measure that conveys how much forecasts deviate on average from observed values. Looking at the results, the values are on average around 3, but the highest number is reached in the last calibration due to an overestimation during the summer period since the water diverted from *Canale Farini* is not taken into account. The RMSE provides a measure of the average error between the observed and simulated values; it penalizes larger errors more than MAE. Indeed, the higher value is observed, even in this case, in the last calibration. Finally, the bias percentage is calculated, which indicates the average trend of the model to overestimate or underestimate the values observed. Except for the calibration at the station n. 1430 – which has a positive value – all the others give low negative values, which denotes that there is an average trend of underestimation. In the second attempt, the highest underestimation is of 31.7 % corresponding to one of the poorest recorded upstream calibration performances. There has, however, been a reduction from the first attempt to the second attempt in the percentage of bias – especially in the last calibration where an underestimation of only 2.24 has been reached. This demonstrates once again the power of the calibration procedure.

The mean values of the discharge in the simulated and observed data series are of a similar magnitude; maximum values are greater than the observed values, apart from n. 1430; minimum values of simulated discharges are almost identical, indicating good modelling of the low flows. Maximum and minimum extreme values show a large variability, suggesting a seasonal cycle of discharge related with snow and ice melt processes—a result that was also expected from the alpine basin of the Dora Baltea.

For completeness, a comparison with an already implemented model with GEOframe over the Valle d'Aosta region is made. The model has been implemented by "*Autorità di Bacino del Fiume Po"* to "*update the existing numerical modelling for water resource management*" (Roati et al. 2022 [52]) starting from the Valle d'Aosta region "*since is the most upstream part of the district*" (Roati et al. 2022 [52]). The basin outlet - concerning the Valle d'Aosta part - of the citated model is at *Tavagnasco* which is the hydrometer n. 1 in the current model. The resulting KGE value obtained at the outlet (n. 1) is of 0.76 [51], but this value cannot be compared since no calibration have been performed at the same hydrometer. Nevertheless, the KGE value obtained at *Verolengo* – which is the outlet basin of the current model, and it is located around 40 km below *Tavagnasco* – is 0.08 lower than the reference. The difference is noticeable, but the model proposed used a higher discretization of the basin and a higher number of calibrated sections with more rounds and steps. This surely increases the overall result and – with these remarks – the current model could be considered good and comparable. Obviously, it is a rough comparison, but it gives evidence for future improvements by increase, for example, the number of sections calibrated, and the iterations used - which can surely improve the overall result in terms of KGE as demonstrated by the improvement obtained between the first and second attempt.

Results of time series obtained after the second attempt are shown in the following two paragraphs for each calibration divided into upstream and downstream. In addition, goodness-of-fit scatter plots have been added in order to give a better visualization. The same results for the first attempt of calibration can be retrieved in *Appendix B1*.

### 5.1.2 Upstream calibrations

All the four upstream calibrations have been run separately in order to access at the optimum set of parameters for all the mountainous regions taken into account. *Figure 5.3* shows in details the 11 sub-basins calibrated in four different gauging stations. Only one - the calibration at station n. 1290 - considers just 2 sub-basins, while the others calibrate 3 subbasins each.


*Figure 5.3, upstream calibration set-up for the Dora Baltea basin*

In *Table B1.2 in Appendix B1*, all resulting parameters values of each calibration are reported for the second attempt. The values of the combined melting factor, the melting temperature and the groundwater parameters have the same order of magnitude for every calibration highlighting a similar path in the Dora Baltea Basin territory. Instead, the runoff and rootzone maximum storage vary a bit between the various parts of the territory.

In the following pages, the comparison of the uncalibrated and calibrated timeseries, during the second attempt, are reported for each calibration: firstly, the entire period of modelling – including the calibration and validation period - is used to have a clear overview; then a zoom on the year 2023 is done to get more details. Following, the scatterplots for the calibration and for the validation period are showed to assess the goodness of fit. Furthermore, in *Appendix B1* the resulting timeseries of the first calibration attempt can be retrieved.

## ***Results of calibration at station 1490 - Pré Saint Didier Capoluogo***

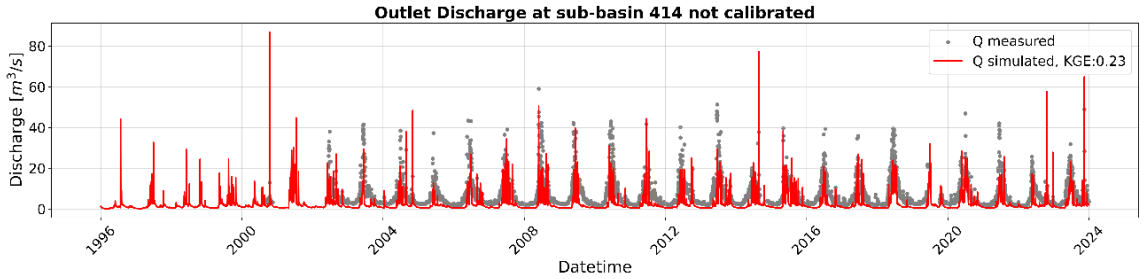



*Figure 5.4, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 1490*

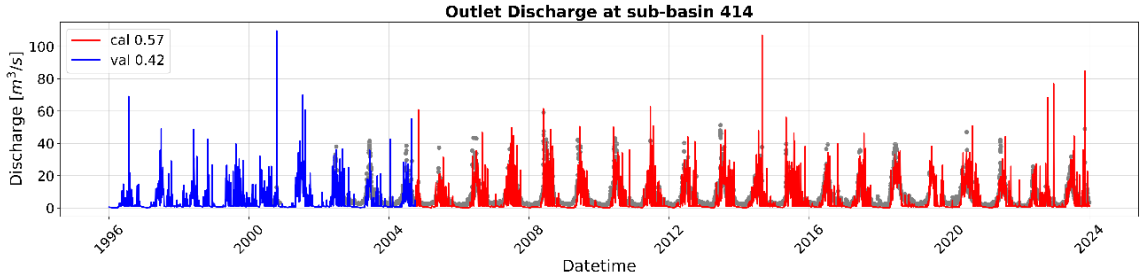



*Figure 5.5, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 1490*

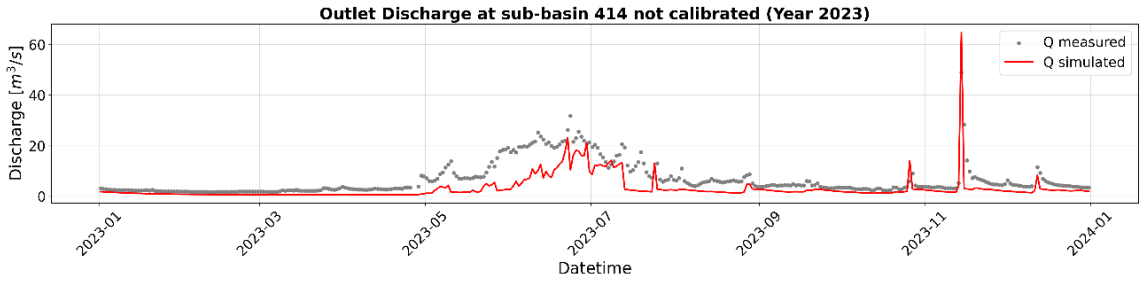



*Figure 5.6, comparison between observed and simulated discharge timeseries in 2023, without calibration at station 1490*

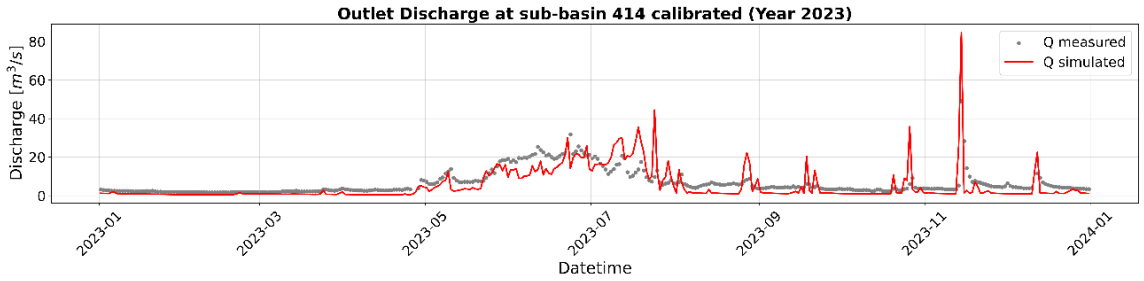



*Figure 5.7, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 1490*

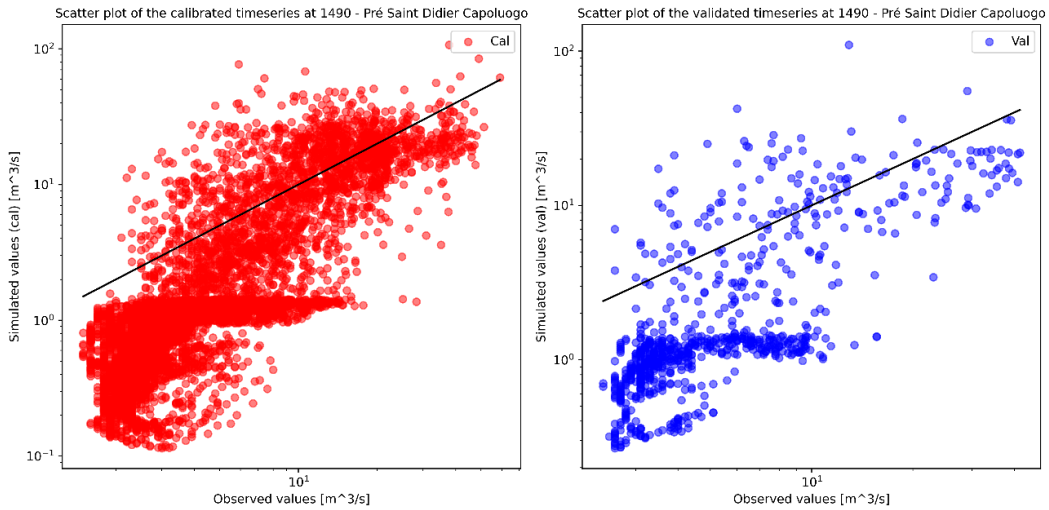



*Figure 5.8, scatterplot of the calibrated timeseries (red) and validated timeseries (blue) for the station 1490*

The comparison of the timeseries before and after calibration shows a significant improvement in the OF: passing from a KGE value of 0.23 to 0.57. Analysis of the 2023 zoom shows that the low flow conditions are well represented, and the peak flows show a similar pattern but slightly overestimated. This calibration managed to improve the values during the summer season since the underestimations considerably decreased as the melting process was better represented. The scatterplot in *Figure 5.8* indicates an overall underestimation in the calibration period as the percentage of bias of – 30.2 % highlighted.

## ***Results of calibration at station 1640 - Valpelline – Prelé***

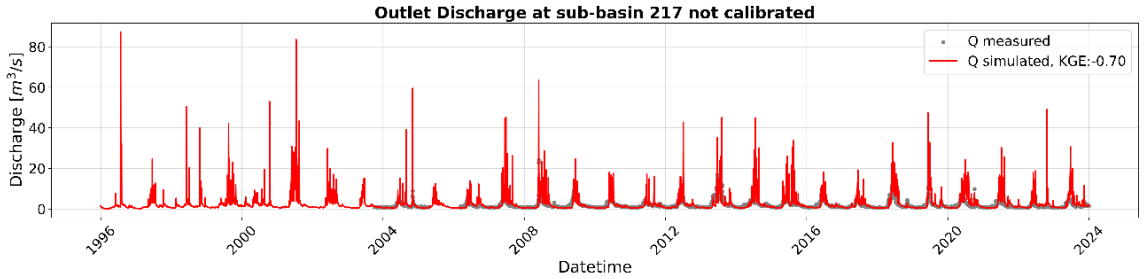



*Figure 5.9, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 1640*

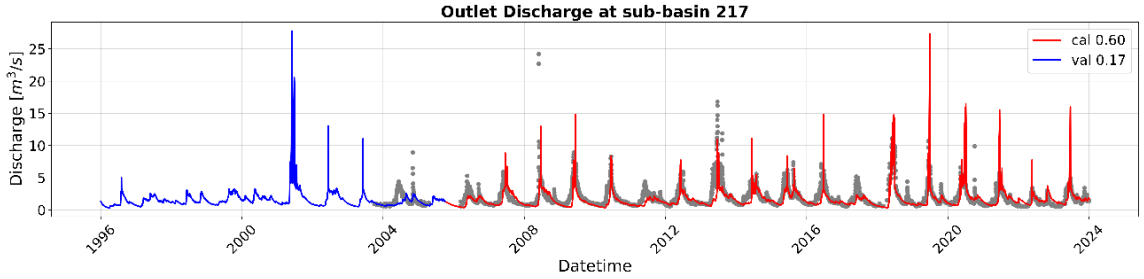



*Figure 5.10, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 1640*

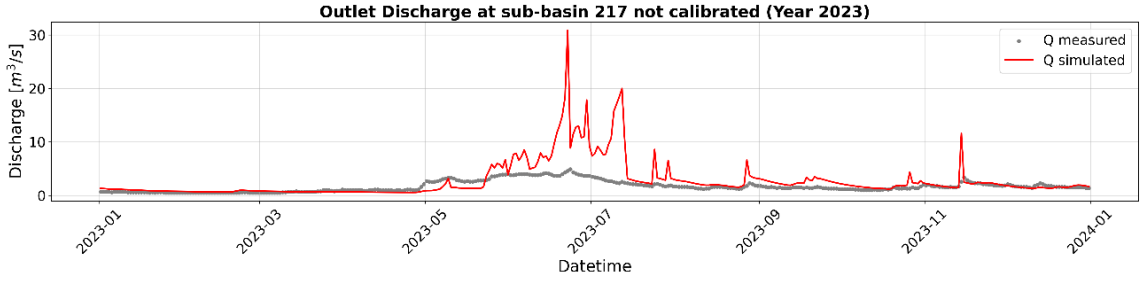



*Figure 5.11, comparison between observed and simulated discharge timeseries in 2023, without calibration at station 1640*

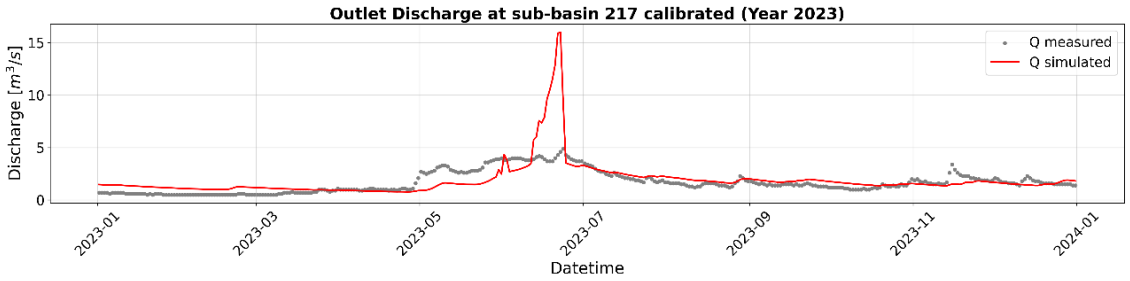



*Figure 5.12, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 1640*

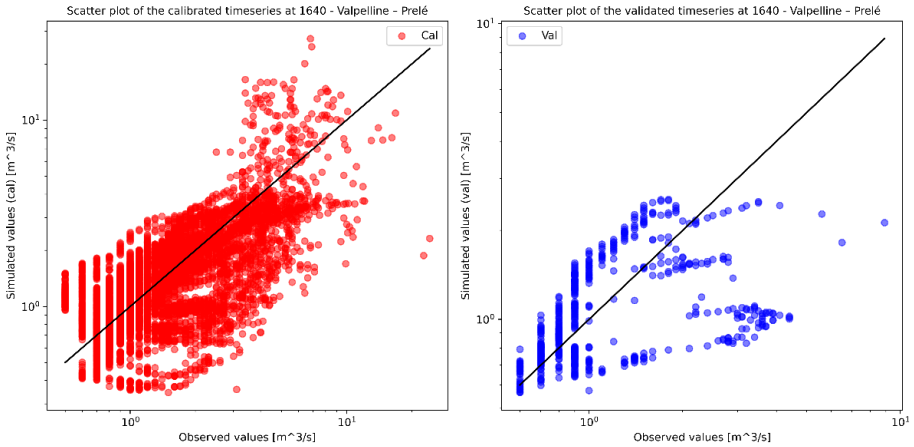



*Figure 5.13, scatterplot of the calibrated timeseries (red) and validated timeseries (blue) for the station 1640*

Compared to the other increases during the whole modelling process, the one obtained in the OF value during this calibration phase is the biggest - it goes from a negative KGE value of -0.70 to a positive one of 0.60. The low flow is well captured also in this case, while peaks are highly overestimated, especially during the summer season. Such a feature reflects the fact that there is a major dam called *Place Moulin* within that vicinity – which has a volume capacity of 9.3 million $m^3$ [50] - which strongly affects the hydrological layering and water accumulation within the area. Since the model does not consider the presence of the dam, it simulates a major quantity during the summer without considering the regulation of the discharge given by the reservoir. A better analysis could be done by looking at cumulative volumes that are much more significant given the presence of the dam. Indeed, looking at *Figure 5.14*, during the coldest and more rainy seasons the dam collects water to release it later and thus the model tends to overestimate the cumulative volumes. Whereas, during the hottest period in summer the dam release water to respond to the water request downstream. For that reason, during the summer period the simulated cumulative volumes in 2023 are remarkably similar to the observed ones since the dam release water. Whereas during the rest of the year there is a systematic overestimation due to the fact that the dam holds back the water.

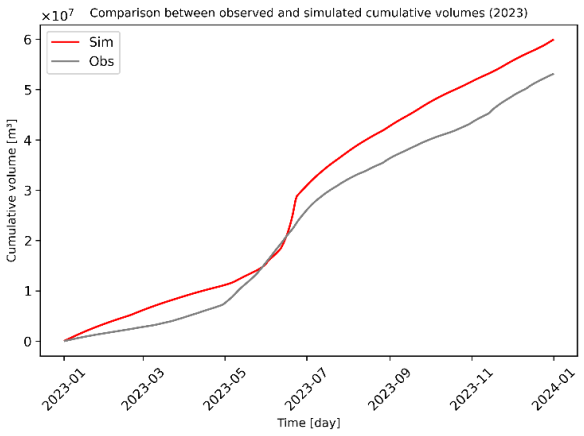



*Figure 5.14, comparison between simulated and observed cumulative volumes in 2023 at sub-basin 217*

## ***Results of calibration at station 1130 – Cogne – Crétaz***

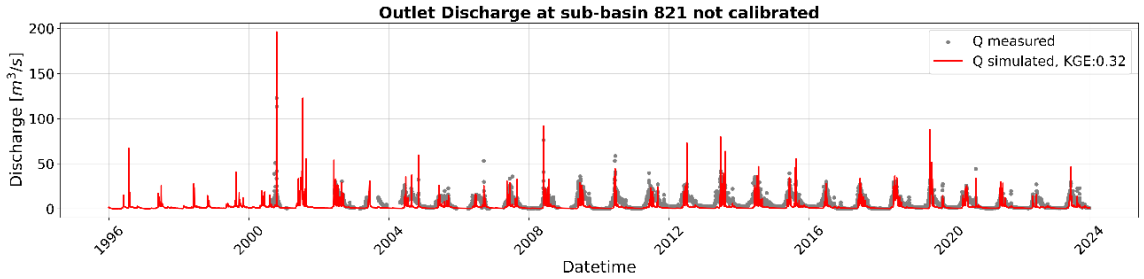



*Figure 5.15, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 1130*

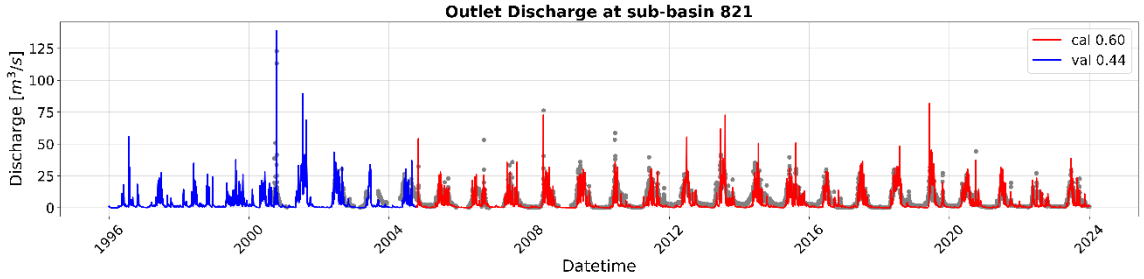



*Figure 5.16, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 1130*

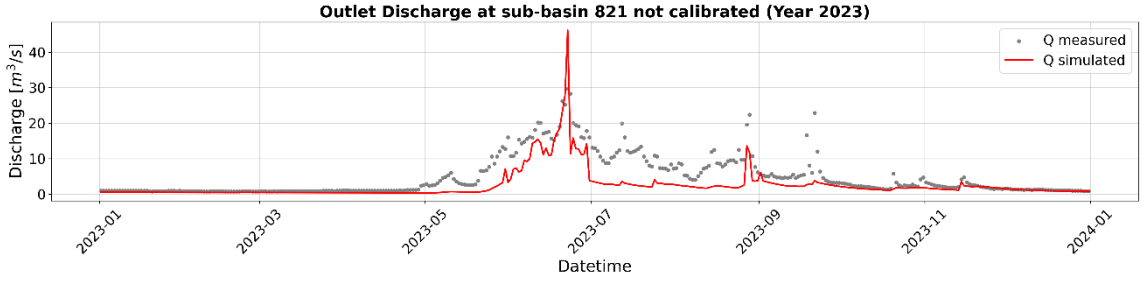



*Figure 5.17, comparison between observed and simulated discharge timeseries in 2023, without calibration at station 1130*

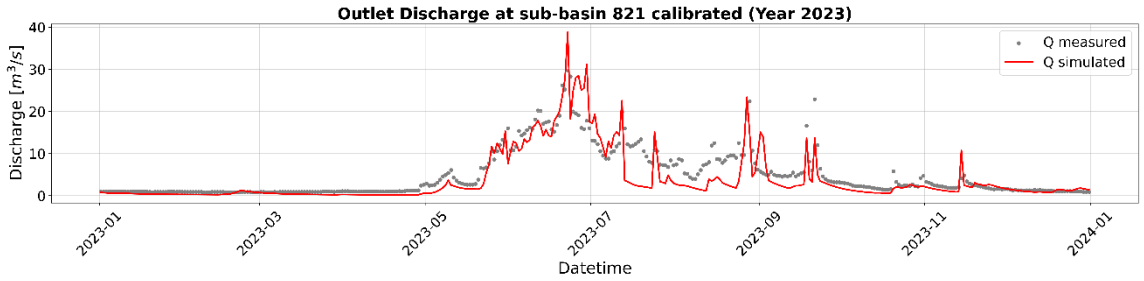



*Figure 5.18, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 1130*

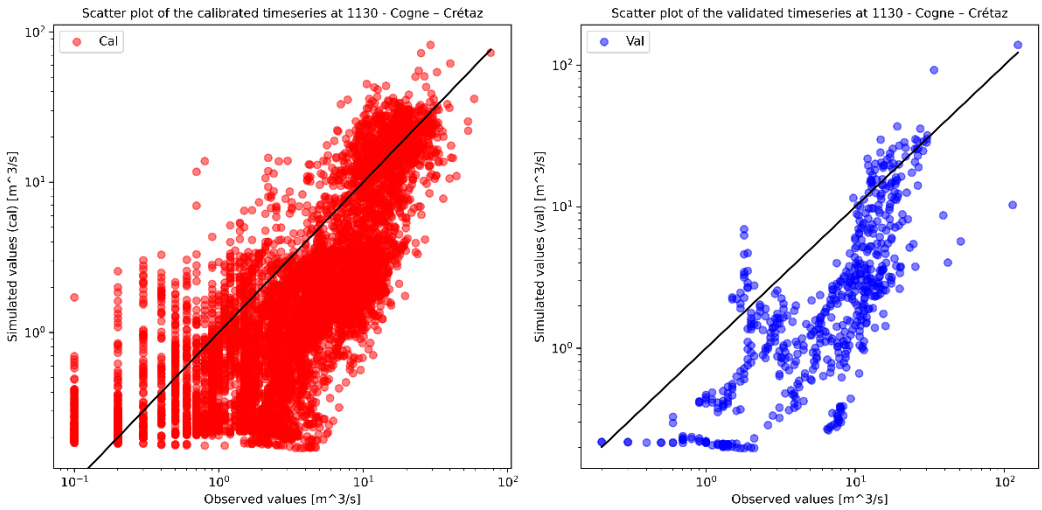



*Figure 5.19, scatterplot of the calibrated timeseries (red) and validation timeseries (blue) for the station 1130*

The improvement of the OF is evident passing from a KGE value of 0.32 to 0.60. The low flows in this case are almost perfectly followed. Whereas the peaks during the spring and summer period are well simulated without overestimating too much. In the concluding phase of summer, there exists a tendency to underestimate conditions, likely attributable to an excessively low melting temperature coupled with a freezing factor that enhances the formation of ice and snow and, as consequence, lowering down the discharge. The scatterplot illustrates the trajectory of this underestimation, as evidenced by the observed bias percentage of -31.7%.

## ***Results of calibration at station 1290 - Gressoney – Saint-Jean – Capoluogo***

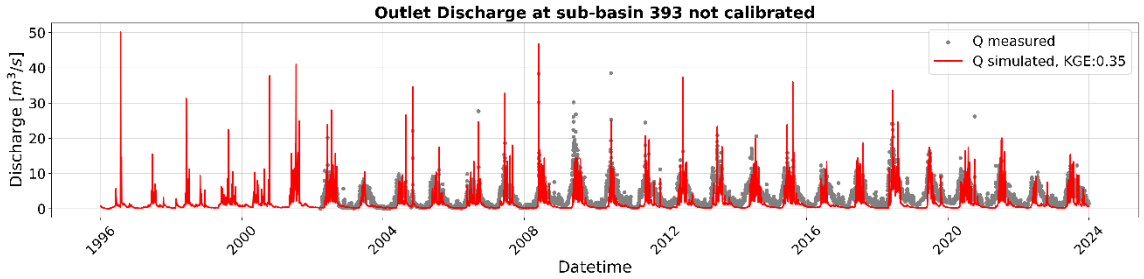



*Figure 5.20, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 1290*

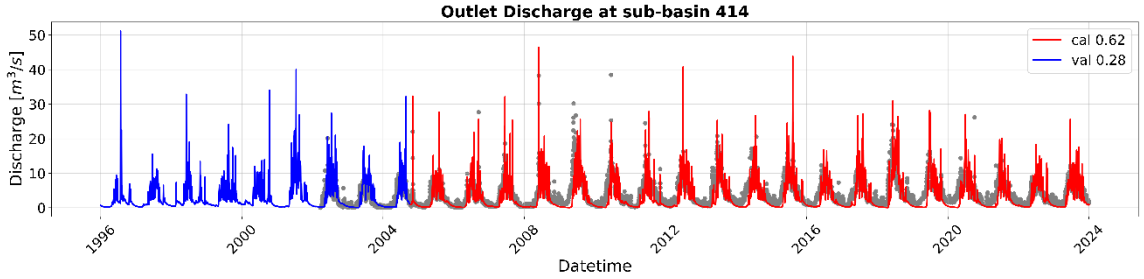



*Figure 5.21, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 1290*

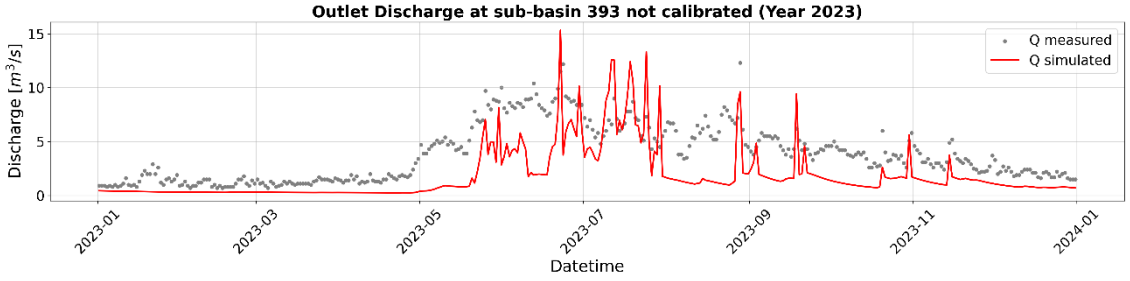



*Figure 5.22, comparison between observed and simulated timeseries in 2023, without calibration at station 1290*

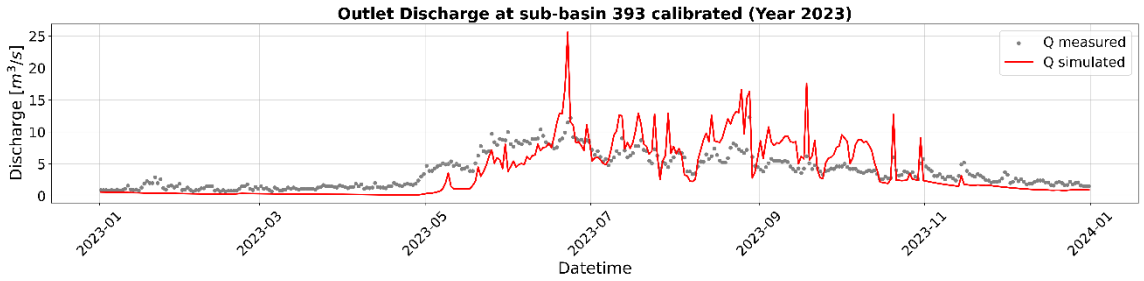



*Figure 5.23, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 1290*

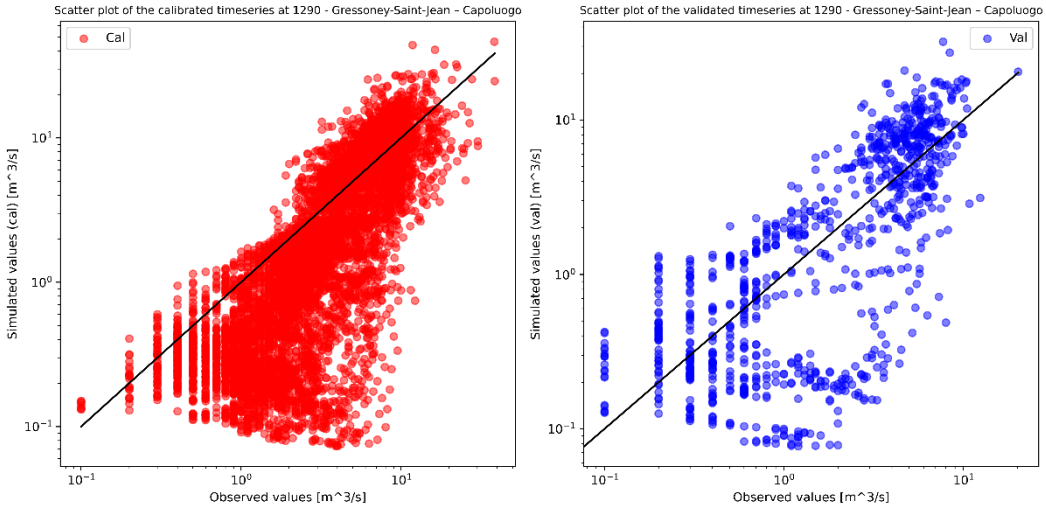



*Figure 5.24, scatterplot of the calibrated timeseries (red) and validated timeseries (blue) for the station 1290*

In this last upstream calibration, the KGE value improved by 0.23, going from 0.35 to 0.62. The great adherence between the observed and simulated discharge values is evident in the year 2023, where peaks, although slightly overestimated in the summer period, follow the same trend. The low flow measurements are somewhat underestimated, but the calibrated timeseries is quite different from the uncalibrated timeseries, suggesting a significant improvement—and compared to the first attempt at calibration, where the calibrated timeseries seems to have a clear different trend than the observed one. Also in this case, the presence of a dam affects the quality of the fit since it was not considered in the analysis – even though less evidently compared to the second upstream calibration since the dam in this case has a volume of 4.17 million $m^3$ [50] compared with the 9.3 million $m^3$ [50] in the previous case. As in the previous four calibrations, the scatterplot in *Figure 5.24* shows a slight underestimation of 18.5%, as highlighted by the % of bias.

### 5.1.3 Downstream calibrations

In the flat region of the territory, two consecutive calibrations were performed, each maintaining the previously calibrated parameters fixed. In the case of the first calibration, carried out at station n. 1430, the three upstream sub-basins groups were processed with the three different sets of fixed values, as *Figure 5.25* shows – highlighting in green the sub-basins already calibrated. In the second calibration, all the five sets of parameters obtained before were kept fixed, as illustrated in *Figure 5.26*. Compared to earlier calibrations, these two elaborates a significantly larger number of HRUs: 25 sub-basins were calibrated using gauging station n. 1430 and 28 sub-basins using gauging station n. 3.

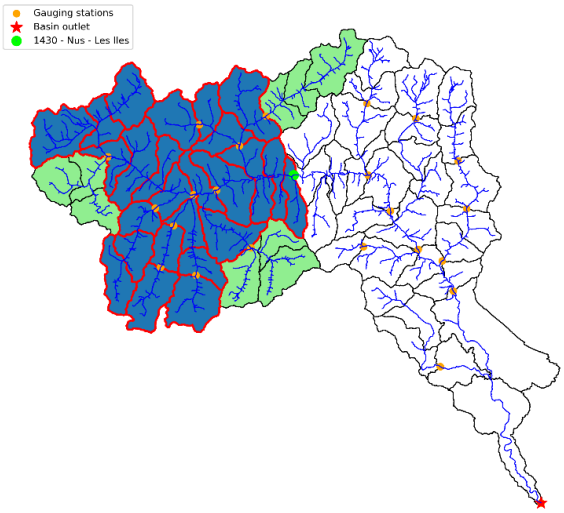



*Figure 5.25, calibrated sub-basins at station n. 1430 (blue) and sub-basins with parameters kept fixed (green)*

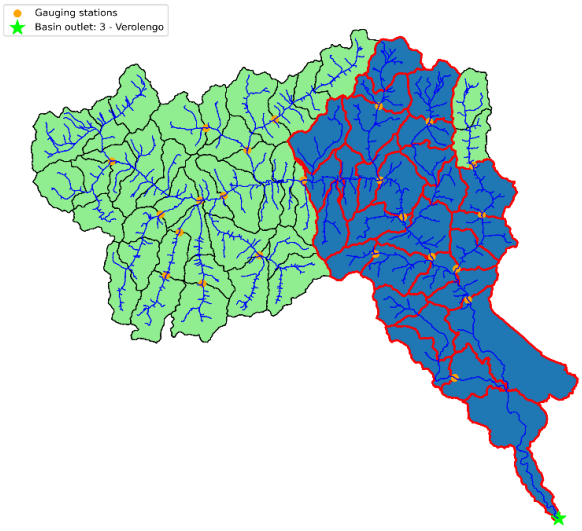



*Figure 5.26, calibrated sub-basins at station n. 3 (blue) and sub-basin with parameters kept fixed (green)*

In *Table B1.3 in Appendix B1*, as for the upstream calibrations, the set of parameters values obtained during the second attempt are reported. In this case, the parameters are not so similar because a larger and more heterogeneous zone of the territory is taken into consideration. In particular, in both calibrations, a mix of mountainous and plain areas are included due to the considerable number of HRUs involved. This may reduce the accuracy of the research of the optimal parameters, producing an overall average. However, melting temperature, freezing factor, and water retention capacity have comparable magnitudes in the two scenarios. On the other hand, runoff and groundwater parameters deviate slightly. Overall, the values tend to be somewhat higher compared to the upstream calibration, suggesting an appropriate acknowledgment of the existence of sub-basins at lower elevations.

The time series results at the sub-basin outlet for each calibration are shown in the following sections. Also in this case, calibrated versus un-calibrated data are compared and a scatterplot assess the modelling performance. Only for the final calibration result, a more detailed analysis is made.

## ***Results of calibration at station 1430 – Nus – Les Iles***

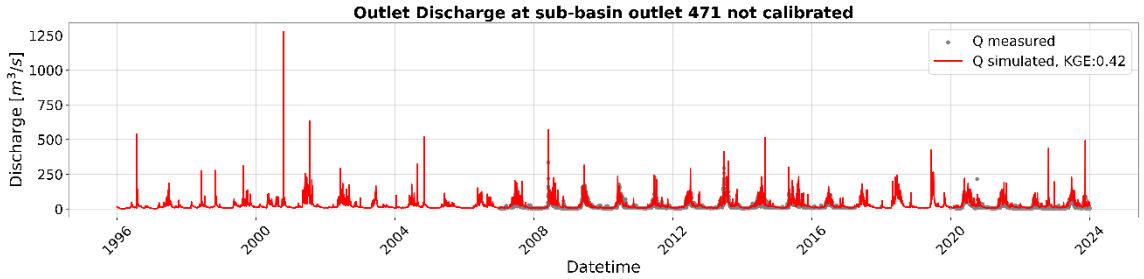



*Figure 5.27, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 1430*

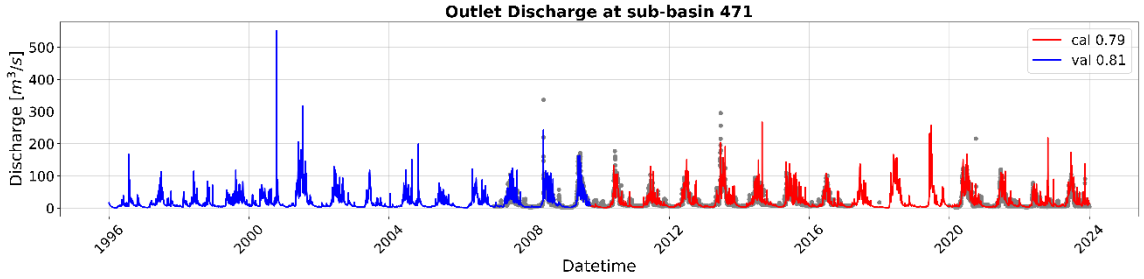



*Figure 5.28, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 1430*

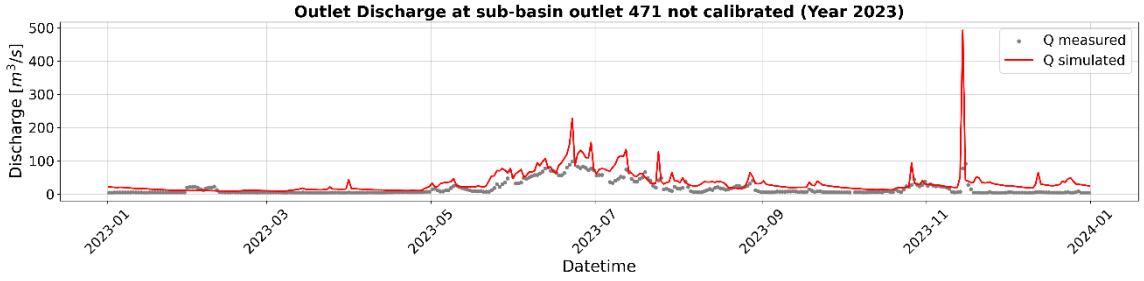



*Figure 5.29, comparison between observed and simulated timeseries in 2023, without calibration at station 1430*

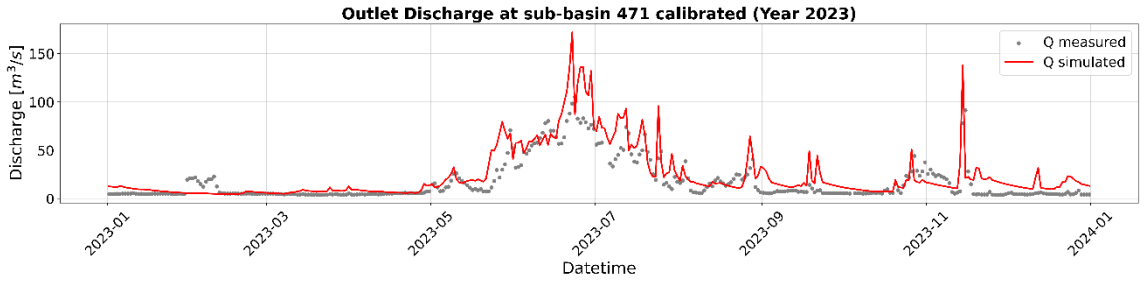



*Figure 5.30, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 1430*

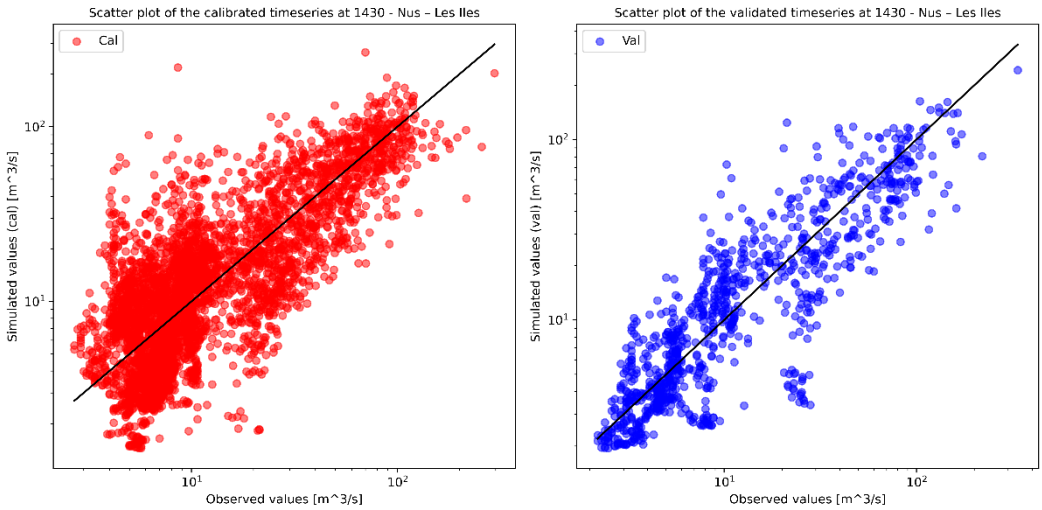



*Figure 5.31, scatterplot of the calibrated timeseries (red) and validated timeseries (blue) for the station 1430*

The calibration performed at station n. 1430 reveals a particularly good improvement of the OF, giving the highest result in terms of KGE value. The start value of 0.42 has been improved to 0.79 in the calibration period and to 0.81 in the validation period. The peaks are clearly lowered down and they seem to be well interpolated since there is only a slightly overestimation. The low flows are well captured despite a low underestimation, but the trend is clearly followed rather closely with a proficient level of resolution. The scatterplot in *Figure 5.31* shows a commendable level of accuracy for both periods with a minimal overestimation, as evidenced from the +5.67 % bias.

## ***Results of calibration at station 3 – Verolengo***

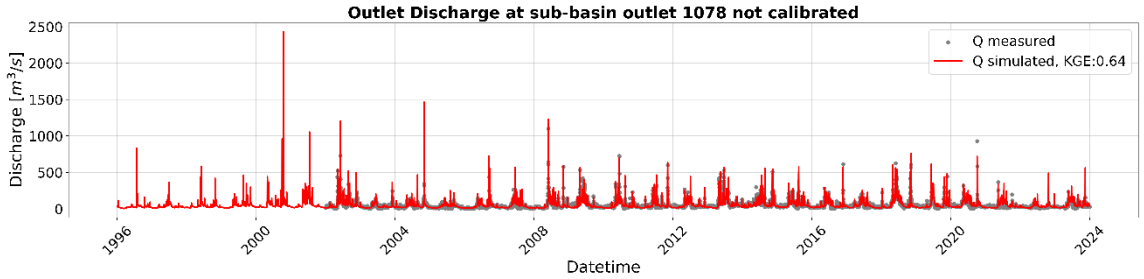



*Figure 5.32, comparison between observed and simulated discharge timeseries between 1996 and 2023, without calibration at station 3*

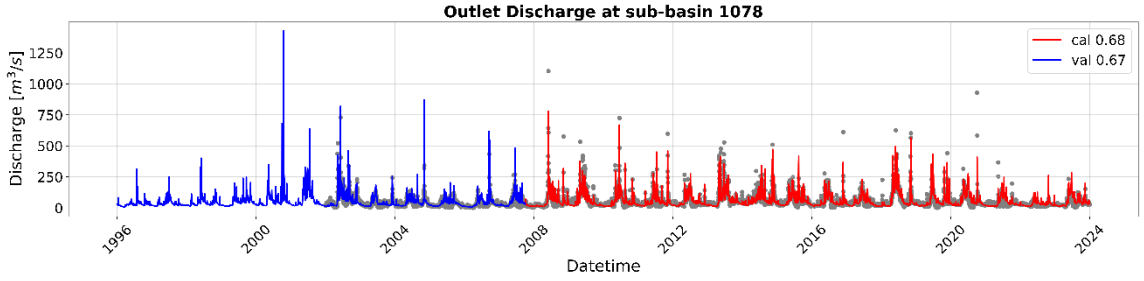



*Figure 5.33, comparison between observed and simulated discharge timeseries between 1996 and 2023, with calibration at station 3*

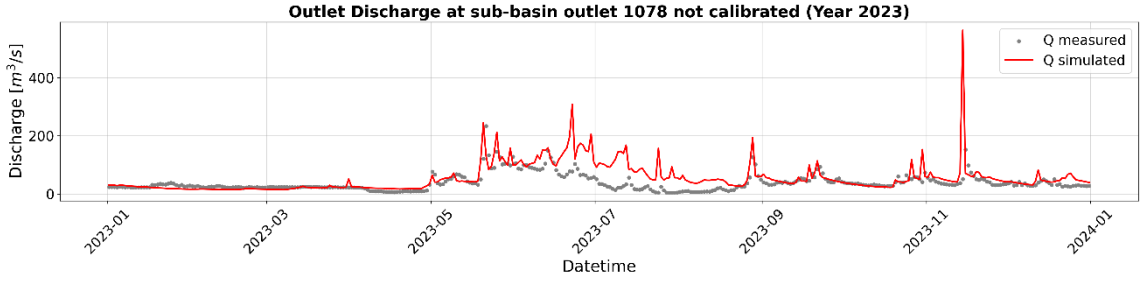



*Figure 5.34, comparison between observed and simulated discharge timeseries in 2023, without calibration at station 3*

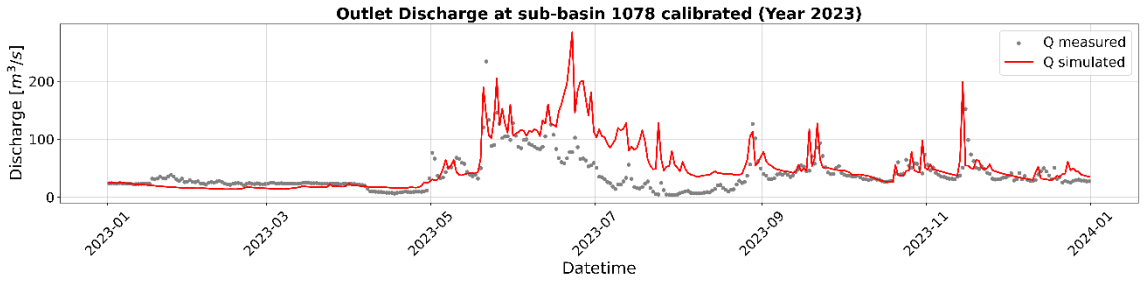



*Figure 5.35, comparison between observed and simulated discharge timeseries in 2023, with calibration at station 3*

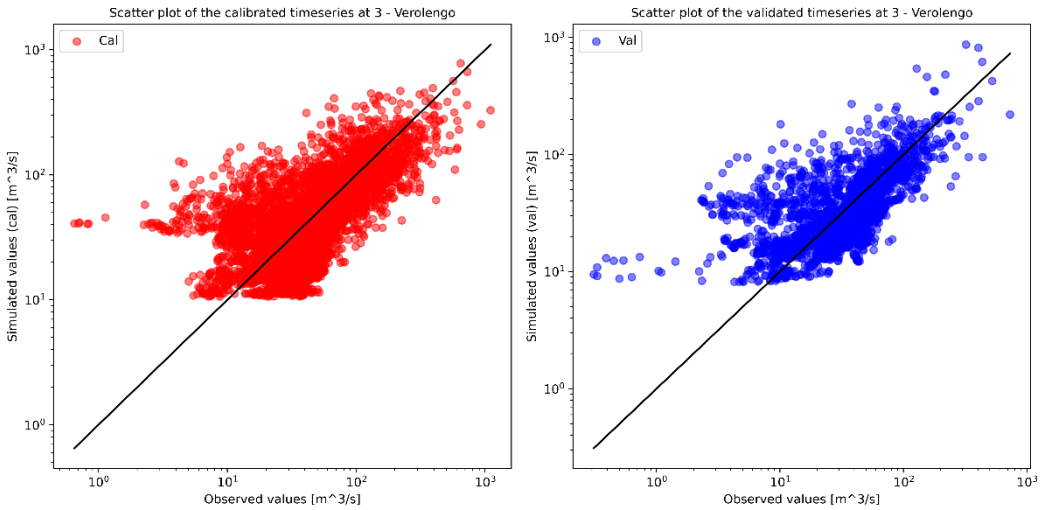



*Figure 5.36, scatterplot of the calibrated timeseries (red) and validated timeseries (blue) for the station 3*

As it was already noticed before, the improvement in the last calibration between the uncalibrated and calibrated timeseries is not so hight. This could be related to the fact that the water drawn by the *Canale Farini* is not taken into consideration. However, the final value of KGE is of 0.68 for the calibration period and 0.67 for the validation period, which is considered good. The modelled discharge is in line with the observation in terms of magnitude and timing. The peaks are much better modelled with the calibration without overestimating. As well as the low flows that despite the first part of the year in 2023 (*Figure 5.35*) are followed very well. The only overestimation is evident during the summer period related to the water diverted 5 km above the basin outlet by the canal. The scatterplot (*Figure 5.36*) shows a particularly good adherence between the observed and simulated data, confirmed by the low value of the percentage of bias (-2.24%).

Cumulative distribution functions (CDFs) of the simulated and observed discharge of the *Dora Baltea Basin* outlet during the calibration period (2007-2023) are shown in *Figure 5.37*. Both observed (blue) and simulated (red) curves exhibit similar trends. They both have starting points with cumulative probability values of 0 at extremely small discharges and thereafter both tend to increase till 1. For flows in the range of 0 - 80 $m^3/s$ there is a similarity between these two curves which suggest that a proper calibration procedure has been employed for low and moderate values. However, for discharges greater than 100 $m^3/s$ the two lines appear to be somewhat distinct. When compared to the blue curve (observed), the red curve (simulated) increases slightly faster above 100 $m^3/s$ crossing over till they end up having the same behaviour. This clearly highlights the slightly overestimation of the intermediate fluxes as the timeseries plots already showed. Globally, the model is capable to represented in an effective way the higher values and both curves reaches the cumulative probability of 1. The simulated and observed timeseries cover the whole discharge range available in the dataset, both low and high flows. The CDFs were cut at 400 $m^3/s$ to see better the difference of the two. Nonetheless, the entire CDF is provided in *Figure B1.2 in Appendix B1* indicating higher peak

discharge reached by the observed values, as it can also be observed though the entire timeseries in *Figure 5.33* where few extremely high peaks are not captured by the model.

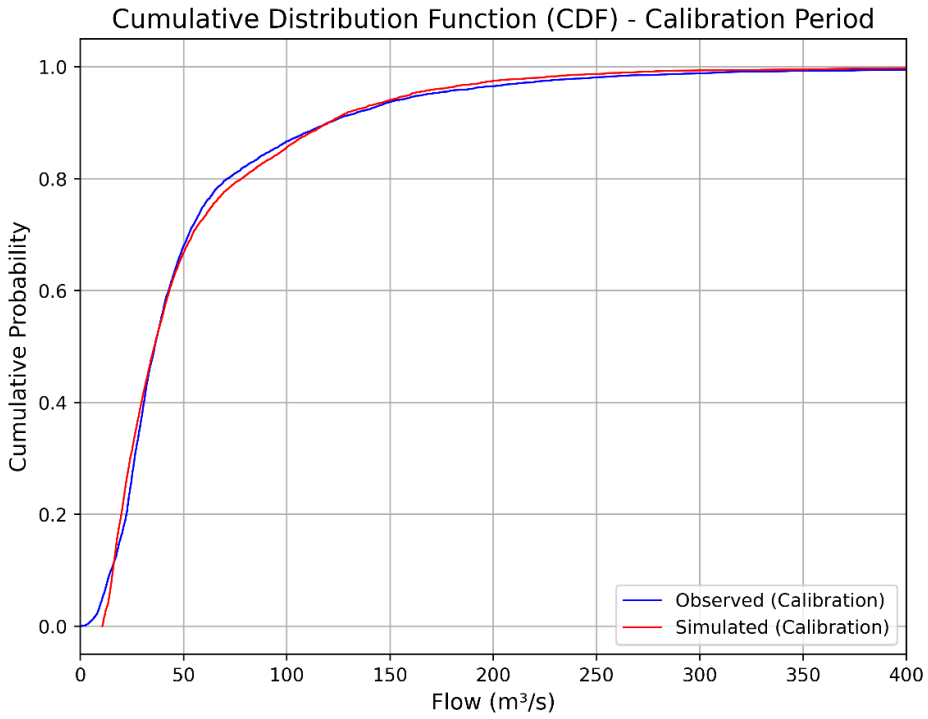



*Figure 5.37, cumulative distribution function (CDF) of observed and simulated discharge at the basin outlet cut at Qmax = 400 $m^3/s$, calibration period 2007-2023*

## 5.2 Water balances

At the end of the calibration procedure, the water balance over the entire period of modelling have been analysed. In *Figure 5.38* the yearly water balance on the Dora Baltea catchment is shown. The mass balance has been carried out considering snow and precipitation as positive contributions -identified with sources - while groundwater and surface runoff together with the canopy and the rootzone evapotranspiration as negative contributions – identified with sinks. Looking at the red dots which give the result of the entire balance for each year, the surpluses or the deficits can be observed. All the yearly values have been weighted by the respective area of each sub-basin. Since the average was made on a yearly time scale the storages were neglected because they give an exceptionally low contribution over the total yearly sinks.

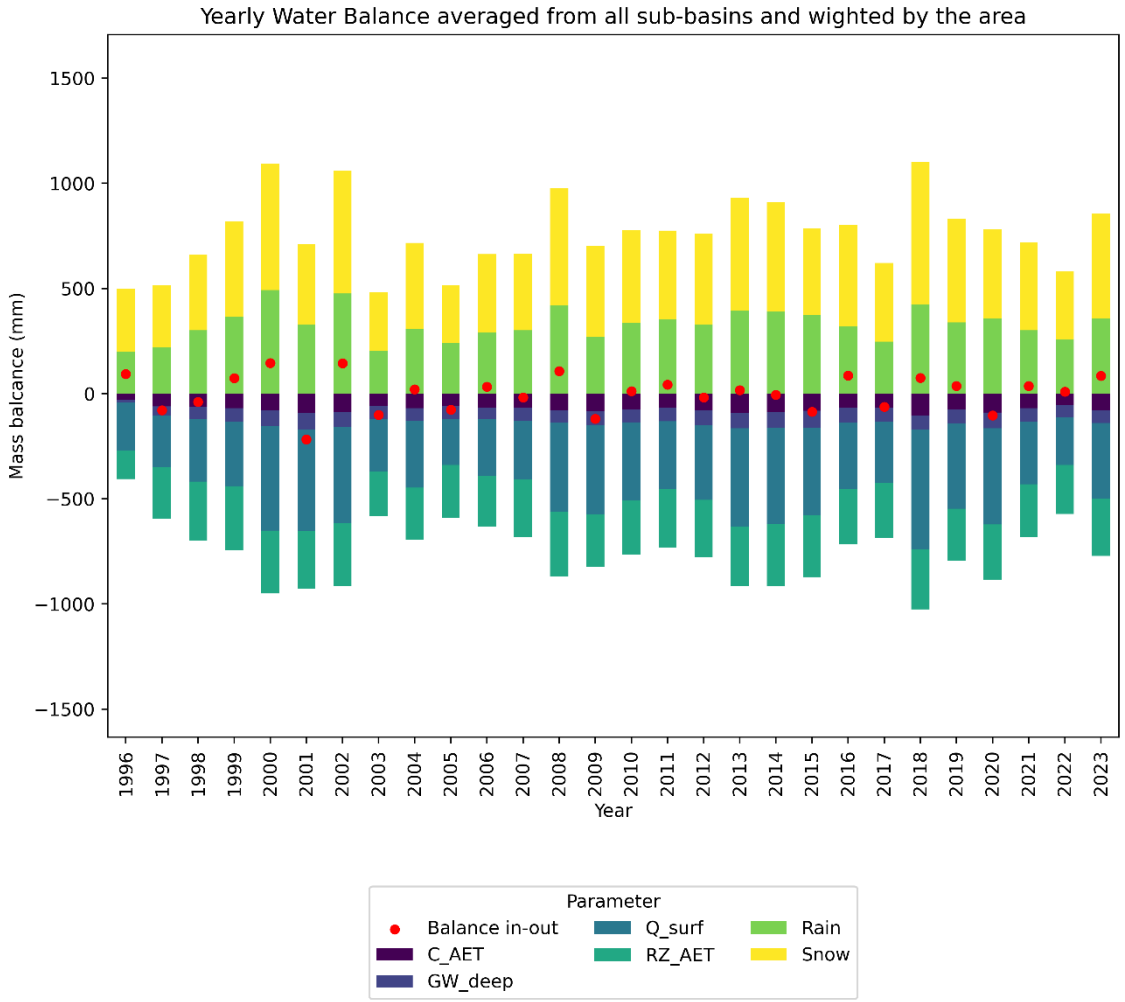



*Figure 5.38, yearly water balance of the calibrated Dora Baltea basin averaged from all-sub-basins and weighted by the area.*

As it can be seen the snow gives a huge contribution in accordance with the fact that we are analysing an alpine basin. The output is dominated by the surface runoff and the groundwater contribution has been lowered down in comparison with the uncalibrated water balance, available in *Figure B1.1 in Appendix B1*. The rootzone ET is much higher than the canopy ET but they follow the same trend. The water balance is nearly closed for the calibration period (2004 – 2023) with some deviations toward the positive values due probably to the precipitation interpolation which suffers in accuracy due to lacks in measured data. In the validation period (1996 – 2003) the deviation reaches higher values especially during the extreme events like the flood in 2000. This could be related to the fact that no gauging stations have data for that period and the calibration was not performed. The model capacity on a back casting prevision is extremely limited since the entire model has been calibrated with more recent data and the previous condition could not be efficiently identified.

The surface runoff accounts on average for 47% of the sinks and it varies a lot during the years with a maximum of 57% in 2018 when the cumulative precipitations seem to be

higher compared to the other years (see *Figure 3.21*). The yearly groundwater discharge, as expected, gives only between 3% and 10% of sinks contribution.

On average ET accounts for 45% of the sinks, ranging from 39% to 52% according to the water availability. Interestingly, as it was already showed in the evapotranspiration analysis, the higher ET values are related to the drier years like 2003 and 2022. This shows that drier years do have less runoff due to less precipitation but also due to an ET which is proportional larger, as some studies demonstrate (Mastrotheodoros et al. 2020 [40]).

Another recurrent situation is clear looking at the wettest years, like 2000 and 2002, where a surplus of water is evident, and the territory seems to not be able to manage an excess of water by having higher but not sufficient runoff and ET leading to flood events.

To conclude the water balance analysis, some simulated hydrological components were plotted together to look if the model represents all the processes with a good accuracy. In *Figure 5.39* are represented the snow water equivalent (SWE) and the different fluxes related – snowmelt, snow, and rain precipitations – for the winters 2022-2023 in sub-basin 414 with an average elevation of 1986 m.a.s.l. The hydrological year is represented starting from 01/10/2022 and ending to 30/09/2023.

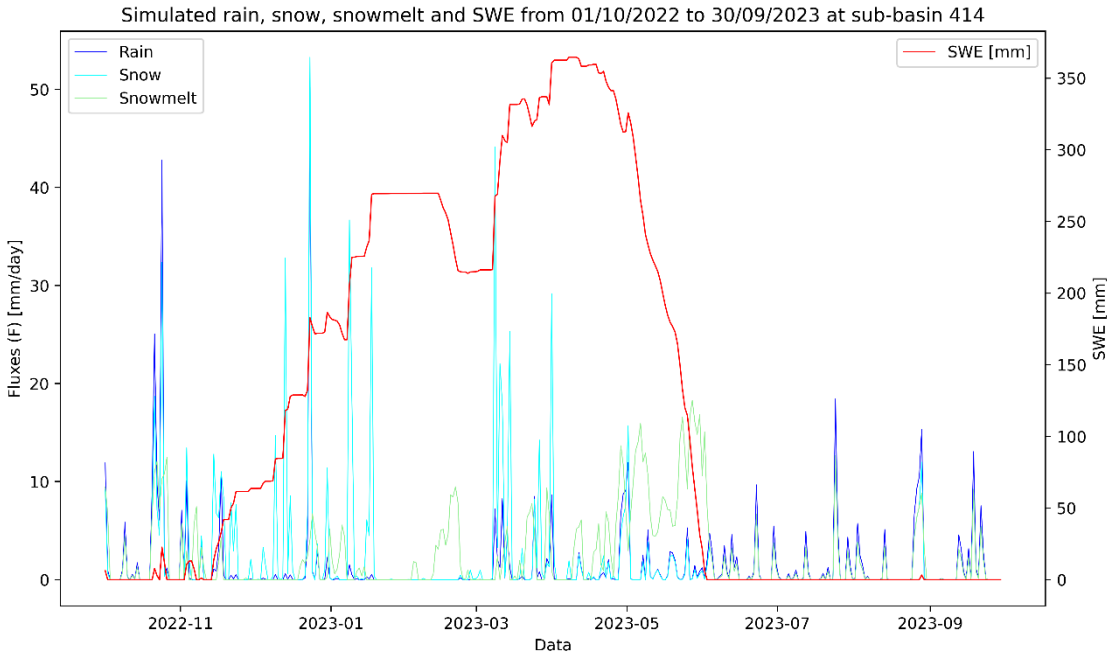



*Figure 5.39, simulated rain, snow, snowmelt, and SWE for sub-basin 414 (average elevation 1986 m.a.s.l.) during the hydrological year 2022-2023*

Snow, rain, and snowmelt are represented as daily values whereas SWE is represented as yearly cumulative magnitude. The increase in snow cover during the winter months is consistent with colder conditions when most precipitation would fall as snow. Heavy snowmelt starts around March 2023, reaches its peak in late April and May, and coincides with melting accumulated snow at rising temperatures. It should be observed that SWE

increases from November 2022 up to the peak in March-May 2023, indicating snowpack gained during winter months. After May 2023, SWE rapidly declines, indicating fast melting of the snowpack. This further coincides with the high values of snowmelt during the period. By June 2023, SWE has reached zero, showing that the complete melting of the snowpack has occurred. This is also confirmed by the near station - available by the "*Centro Funzionale Regione Autonoma Valle d'Aosta*" - which registered data of the snow cover, and they indicate absence of snow during the summer period.

This graph gave a perfect picture of the hydrological cycle in an alpine sub-basin of the Dora Baltea catchment, dominated by the accumulation of snow in winter, followed by its heavy melting in late Spring and early Summer, and then rain takes over in warmer months. This evidences the fact that choosing to use the average elevation of the sub-basin instead of the centroid's elevation was a viable choice. Indeed, considering a more reliable elevation of the area leads to better accuracy in the snow processes simulation.

## 5.3 Post-processing of the Canale Farini discharge

The presence of the Canale Farini 5 km above the basin outlet certainly gives a contribution in the estimation of the discharge since it diverts water from the river for irrigation purposes (*see 2.1.1*). The model does not consider the in-take structure but the high difference during the summer period between the observed and simulated discharge values (see *Figure 5.35*) at the outlet suggests a high amount of water drawn by the canal.

Accessing to the data of water used by the canal is not easy since no gauging station are installed at the confluence. For that reason, it is important to implement hydrological models in order to have access to this information useful for the quantification of the water available for agriculture in the nearby territory. The discharge values of the canal were modelled from 2019 to 2023 by *"Coutenza Canali Cavour"* [49]. The modelled timeseries have some gaps as showed in *Figure B1.3 in Appendix B1*, especially during the year 2023 which has no data during the summer period. However, they have been used to try to evaluate the quantity of water coming from the Dora Baltea River using the results obtained by the model. *Figure 5.40* reports the timeseries of the discharge of the Canale Farini in 2021. During February, some data are missing, but the graphs clearly show the general trend of the water flow of the canal: the discharge increases from April to June where the peaks are located and then it decreases till the end of September. Thus, during summer the canal diverts the highest amount of water from the river to provide the source to the cultivations. For the rest of the year the water diverted seems to be much lower due also to the recharge given by precipitation since the canal is an open air one.

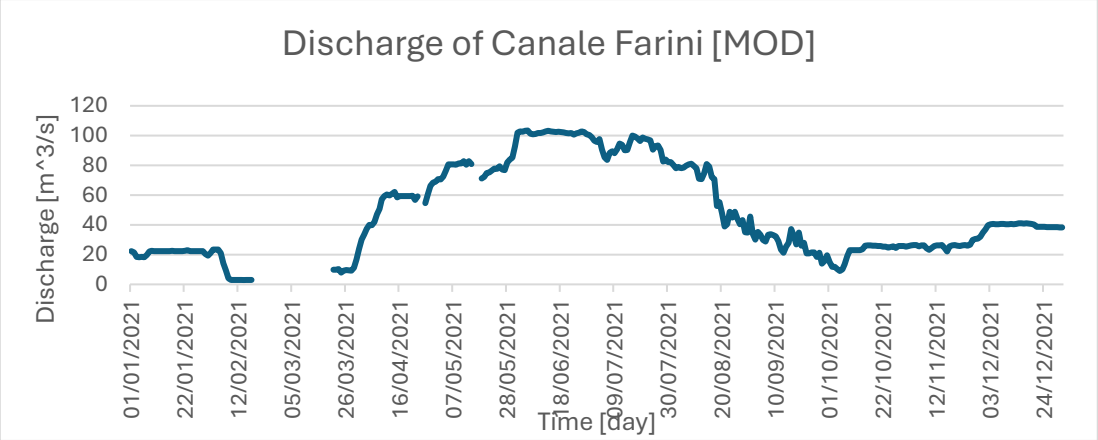



*Figure 5.40, discharge of Canale Farini in 2021 [49]*

Looking at the model results in 5.1.3, it is evident that some discrepancies between the simulated and observed discharge values at the outlet basin are present. In order to access the distribution over time of this dissimilarities, the difference between the two values has been evaluated. This was carried out with the purpose to identify when the two values are much more different and see if it matches the period when the water in the Canale Farin is higher. The analysis was conducted firstly by making the difference between the simulated and the observed discharge values during the period 2019-2023 and, then by dividing it by season. After that, the mean seasonal value for each year has been

evaluated. To compare it with the modelled data of the canal, the same procedure has been done for the discharge of the Canale Farini considering that there is 9.26 % of missing data during the period, and thus the accuracy is lowered down.

The histogram in *Figure 5.41* and the related table (*Table 5.5*) shows the comparison between the mean seasonal value of the observed and simulated difference and the Canale Farini discharge. It is clearly showed that most of the discrepancies between the observed and simulated discharge values are during the summer season, having an average of 49.38 $m^3/s$ of difference. They also have comparable values within the 5-year period, going from a minimum of 33.2 to a maximum of 64.1 $m^3/s$. In the exact same period, the mean season discharge values of the Canale Farini are the highest registered. The trend suggests that during summer the water diverted is much higher than during the rest of the year when the water required by cultivations is lower and the canal can be just fed by precipitation events. Whereas in all the other seasons, as *Table 5.5* reports in details, the overall trend is to have a low underestimation of the simulated discharge compared with the observed one, in accordance with the error metrics results of the model – indicating an overall underestimation of -2.24%. In addition, *Table 5.5* reports also the seasonal cumulative precipitation for each year simulated by the model in the sub-basin of interest: the highest values are during the Autumn and Spring season whereas the lowest values are during Winter. Thus, the Canale Farini is likely fed by precipitation during all the year, but the discharge is increased during summer by diverting water from the Dora Baltea River. This explains why the simulated discharge values agree much more with the observed ones during all the year except the summer period. The only exception is in 2023 - when the difference values and the Canale Farini values almost coincide – due to the huge gap of the Canale Farini discharge data from 24/05/2023 to 24/08/2023.

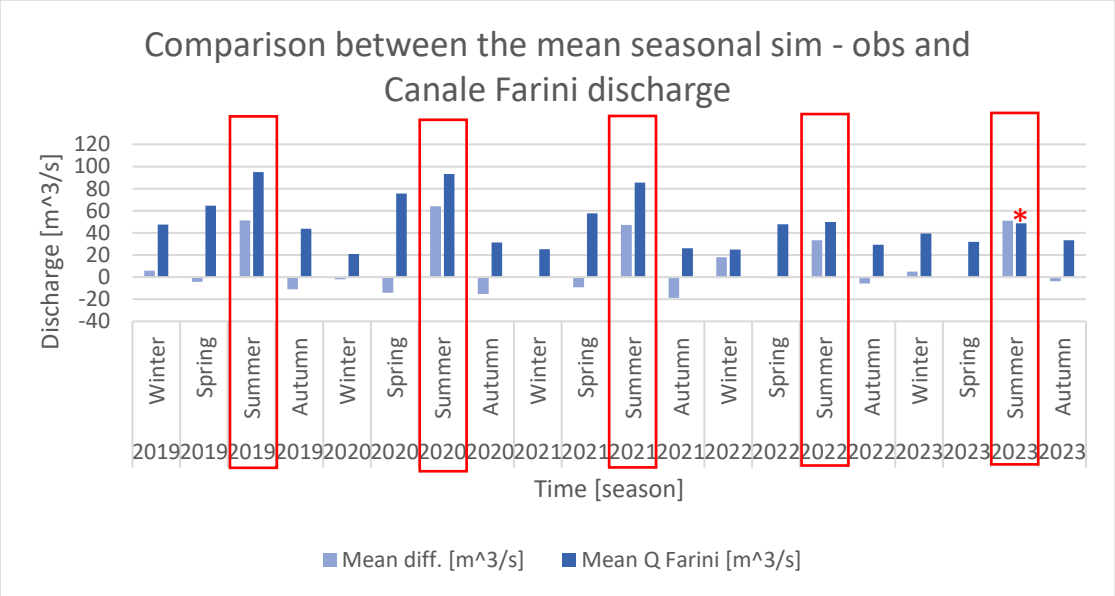



*Figure 5.41, histogram of the comparison between the mean season sim – obs and Canale Farini discharge during 2019-2023 period, the red asterisk over the discharge of the canal in 2023 indicates the gap of missing data during the summer period affecting the comparison*

*Table 5.5, table of the comparison between the mean seasonal sim – obs and Canale Farini discharge, and the cumulative precipitation simulated in the outlet sub-basin during 2019 – 2023 period.*

| Year | Season | Mean diff. [m^3/s] | Mean Q Farini [m^3/s] | Cum. Prec. 1078 [mm] |
|---|---|---|---|---|
| **2019** | Winter | 5.83 | 47.4 | 125 |
| **2019** | Spring | -4.42 | 64.5 | 238 |
| **2019** | Summer | 51.2 | 94.9 | 167 |
| **2019** | Autumn | -10.9 | 43.6 | 426 |
| **2020** | Winter | -1.86 | 21 | 54 |
| **2020** | Spring | -14 | 75.5 | 226 |
| **2020** | Summer | 64.1 | 93.1 | 207 |
| **2020** | Autumn | -15.2 | 31.2 | 150 |
| **2021** | Winter | -0.592 | 25.1 | 110 |
| **2021** | Spring | -9.13 | 57.6 | 142 |
| **2021** | Summer | 47.3 | 85.3 | 200 |
| **2021** | Autumn | -18.7 | 26.2 | 207 |
| **2022** | Winter | 18.1 | 24.9 | 49 |
| **2022** | Spring | -0.857 | 47.7 | 105 |
| **2022** | Summer | 33.2 | 49.8 | 117 |
| **2022** | Autumn | -5.82 | 29.3 | 114 |
| **2023** | Winter | 4.91 | 39.4 | 33 |
| **2023** | Spring | 0.585 | 32 | 279 |
| **2023** | Summer | 51.1 | 48.7 | 265 |
| **2023** | Autumn | -3.75 | 33.3 | 163 |

Thus, the differences between the simulated and observed values during the summer period could be considered as a good proxy of the water diverted from the river to the canal. It is clear that values differ years to years and this approach cannot be taken as sure but as a good estimation.

Therefore, the analysis reported above could be used as a reliable source to correct the simulated data and improve the results of the model including the water diverted by the canal. Two approaches have been carried out to evaluate a correction factor for the simulated discharge: a simpler mean difference and a more rigorous linear regression. The mean difference approach is based on evaluating the mean summer difference between the simulated and observed discharge values and the mean summer simulated discharge over the 5-year period. The correction factor $F_{cor}$ is evaluated using *Equation (5.1)*.

$$F_{cor} = 1 - \frac{Mean\ summer\ difference}{Mean\ summer\ Q\ simulated} \qquad (5.1)$$

The resulting $F_{cor}$ is of 0.47 and it can be applied to the simulated discharge values in order to reduce the bias. It is quick to implement and can capture the overall bias between the observed and simulated discharge.

The linear regression approach models the relationship between the observed and simulated discharge values using a linear regression which provides a correction based on the strength of the relationship. With the linear regression the trend line can be obtained, and the slope *m* and the intercept *b* can be used to adjust the simulated discharge (see *Figure B1.4 in Appendix B1*). The resulting equation gives as slope 0.47 and as intercept -0.40 $m^3/s$. This approach should capture more the variation cross the points but since the slope is the same value obtained with the mean difference approach and the intercept is extremely low, the two approaches give equal results.

Thus, applying the $F_{cor}$ of 0.47 at the simulated discharge values during the summer period, the results were improved by lowering down the overestimation. The timeseries over the entire period of correction is reported in *Figure 5.42*.

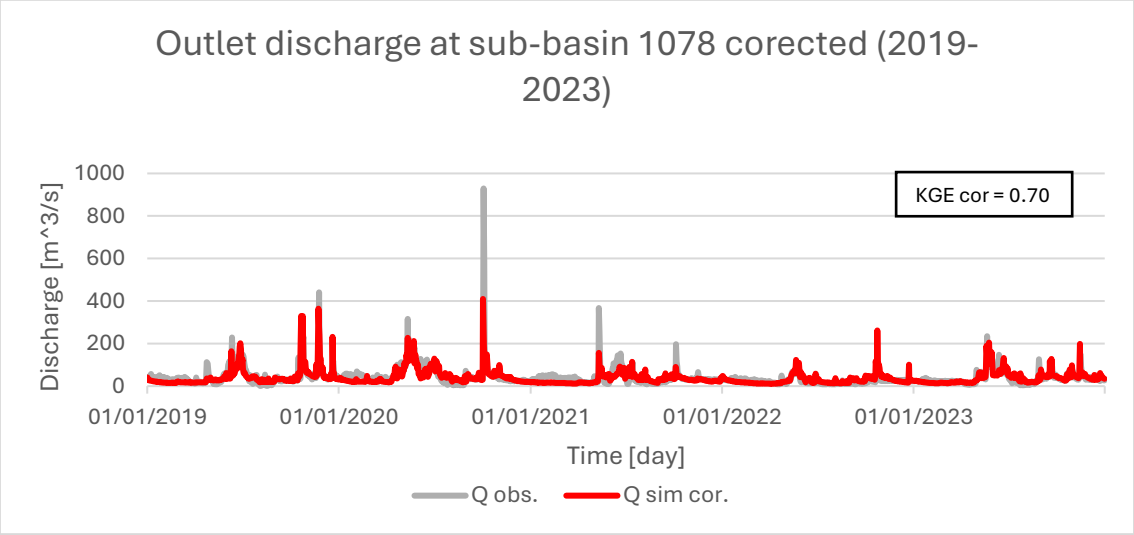



*Figure 5.42, outlet discharge at sub-basin 1078 corrected (2019 2023)*

The resulting KGE – evaluated over the entire period of modelling - has been improved of 0.02 from the uncorrected results. In this way the water drawn by the Canale Farini has been taken into consideration as an overall mean of the differences between the observed and simulated flow which are in accordance with the discharge values of the canal as reported before. *Figure 5.43* reports the zoom on the year 2023 to look better at the results.

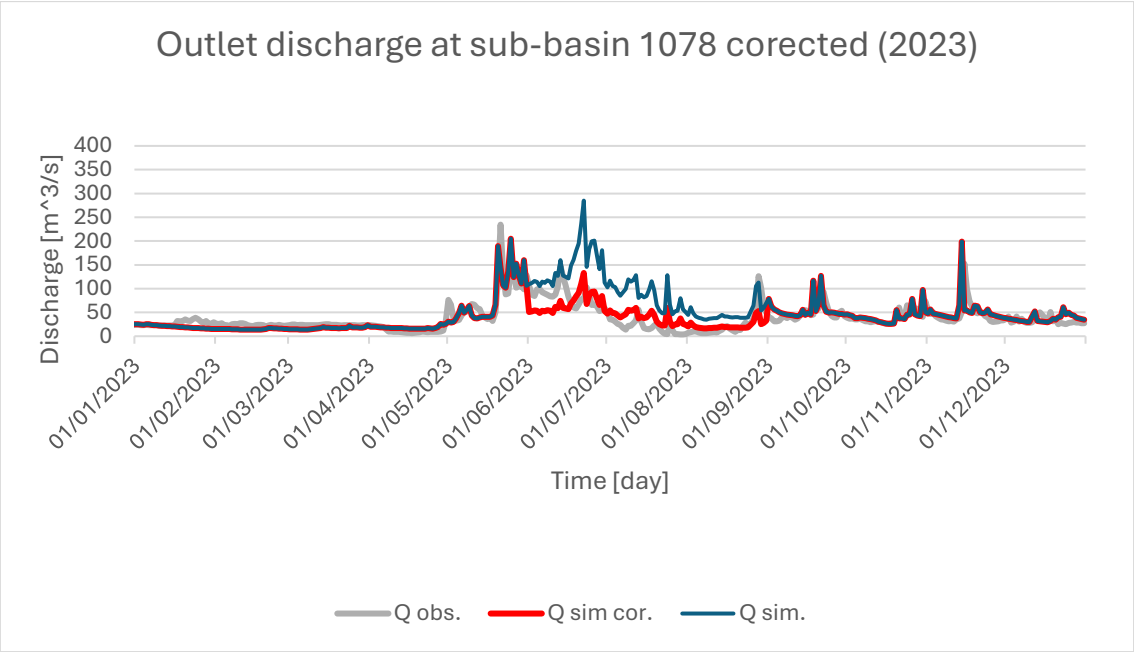



*Figure 5.43, comparison of the simulated, corrected and observed outlet discharge in 2023.*

The red and blue curve identifies, respectively, the corrected and uncorrected flow. The difference is obviously evident only during the summer period since is the modified one. The discharge is lowered down and it better follows the observed values, even though in the first part of the summer the flow seems to be lowered down too much. This could be related to the fact that the overall mean difference is been used and it certainly cannot capture the daily fluctuations. It remains a fact that the proxy evaluated water used by the canal seem to be a good compromise and gives better modelling results.

This is much more evident looking at the difference between the cumulative volumes over the year 2023. *Figure 5.44* shows the comparison of the cumulative curves before and after the correction. The blue curve identifies the uncorrected discharge, and it evidences the higher values modelled during the summer period. Whereas the red curve – showing the corrected values – agrees much more with the observed values reaching remarkably similar quantity of cumulative volume at the end of the year. A slightly underestimation during the first part of the year is evident and in accordance with the negative percentage of bias given as result from the entire modelling.

The comparison of the scatterplots of the corrected and uncorrected timeseries over the period 2019-2023 are reported in *Figure 5.45* and it evidences also the improvement. The corrected cloud is less dispersed around the line indicating a better adherence of the observed and simulated data.

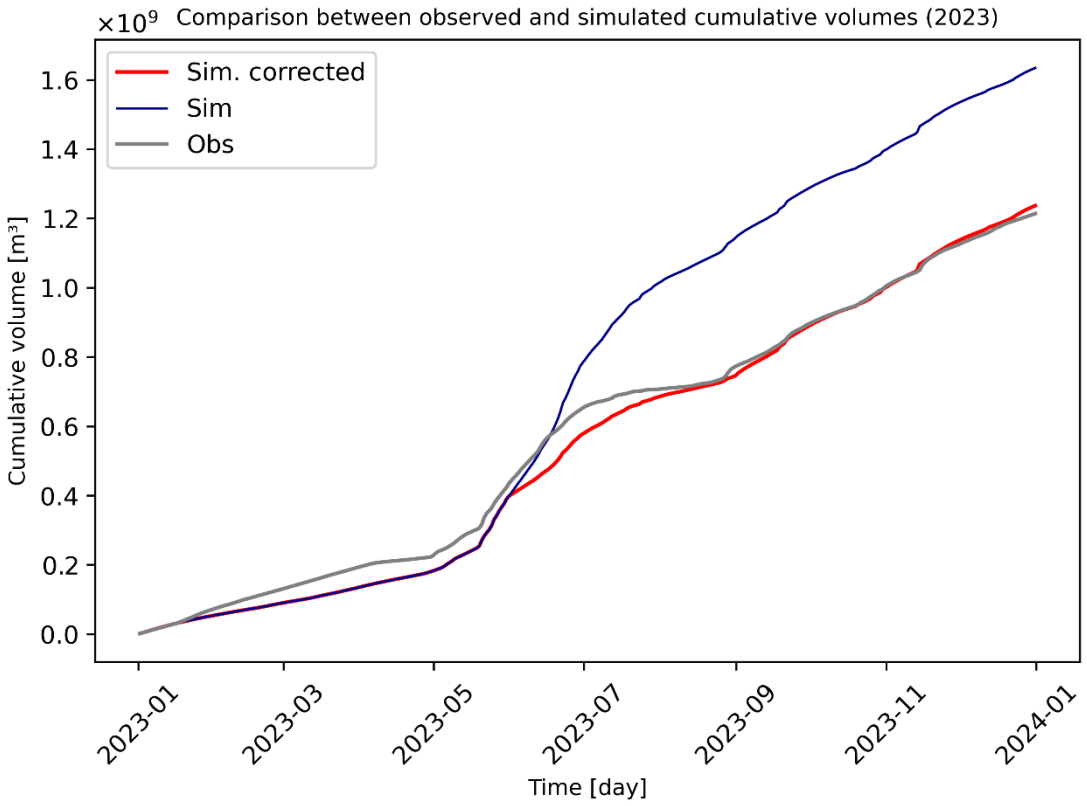



*Figure 5.44, comparison between observed, simulated, and corrected cumulative volumes at the outlet basin.*

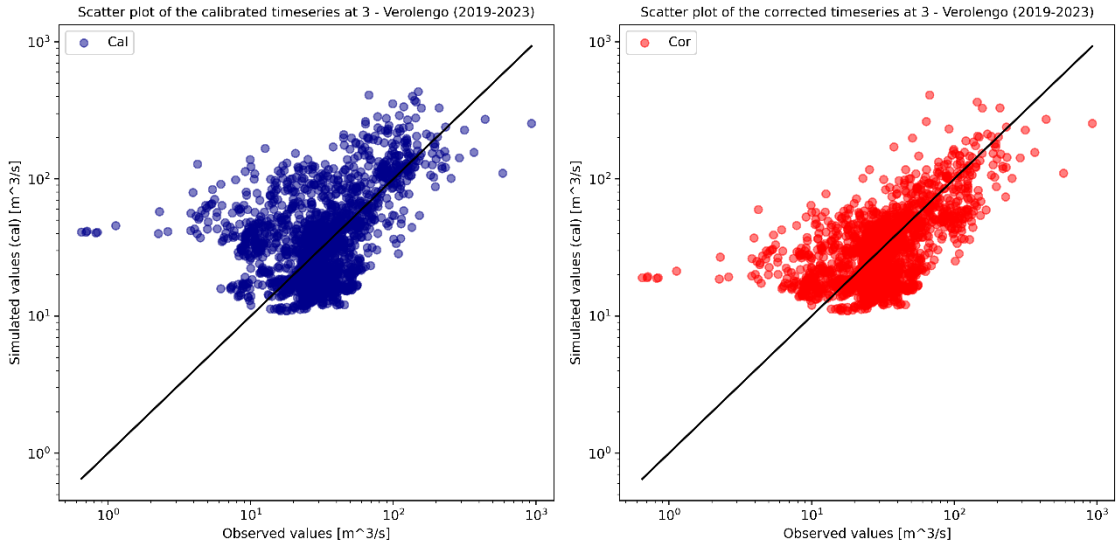



*Figure 5.45, comparison of the simulated and corrected scatter plot over the 2019-2023 period*

Assessing the Dora Baltea water used by the Canale Farini for irrigation purposes is fundamental to have clearer information about the water availability and uses of the territory. The post processing analysis reported above improved the overall result, but it has been done for only a 5-year period since they are the only data available of the Canale Farini discharge – necessary to make the comparison of the proxy of water diverted.

Surely, a more rigorous approach could be used by adding during the geomorphology analysis a sub-basin in correspondence of the canal in order to assess the water discharge arriving at the confluence. In this way, the discharge could be exactly registered where the canal takes water from the river and thus, it could give much more reliable data to make the correction. In addition, a new calibration process using the data corrected can be done and a better result of the parameters estimation could be accessed. Using this approach the model could surely be improved.

# 6 CONCLUSIONS

The aim of this work was to simulate the hydrological cycle of the Dora Baltea catchment in order to quantify the water availability and gives reliable information for a better management of the source. The study was conducted using the physically based semi-distributed model GEOframe which gives the possibility to evaluate the quantity of water flowing in the basin at a daily scale resolution in a period of 27 years (1996-2023).
The entire basin, according to the functioning of the model used, was sub-divided into hydrological response units (HRUs), each of them schematized as a system of reservoirs then connected together to assess the water balance of the entire catchment. The software is divided into a series of six steps: first the geomorphology is assessed, and the catchment is divided into sub-basins, then the temperature and precipitation data are spatially interpolated using Kriging in order to have timeseries for each sub-basin. After that, the radiation and the evapotranspiration budgets are estimated. At the end, all the different components are connected together as a series of layers to have the final simulation of discharge values and finally, a multi-site calibration step concludes the entire analysis in order to adapt the parameters of the model to the specific studied territory.

The results obtained have been analysed using six error metrics including the Pearson correlation and the percentage of bias. The key one – used also as objective function during the calibration step – is the Kling-Gupta efficiency (KGE) which overall, it has been considered robust, having a value between 0.60 and 0.80. The accuracy of the entire model is good since the percentage of bias at the basin outlet is of -2.24% indicating an extremely low underestimation of the discharge. As well as the final comparison between the sum of the observed and simulated discharge over the entire period of modelling, resulting in $\sum Q_{obs} \backslash \sum Q_{sim} = 1.02$. The model results with a good adherence between the observed and the simulated discharge values and the objective of the entire work was accomplished: reducing the computational demand by using daily time resolution and increasing the minimum area of each sub-basin. Temperature and precipitation data interpolation has been considered particularly good looking at the literature comparison – especially with the ISPRA data [23]. Kriging performance and results are in accordance with major of the historical meteorological data. The estimation of radiation and the related evapotranspiration was not easy comparable with reality since an hourly time resolution is required. However, looking at the overall results and considering a daily mean of such quantities which are challenging to estimate, the results seem to be comparable with the literature ones. Indeed, evapotranspiration seems to be good represented according to ISPRA source [38], having a low underestimation in the sub-basins with high elevation.
The calibration steps have played a crucial role in the entire building of the model guarantying the optimal result of parameters estimation of the Embedded Reservoir Model – implemented as last step in the software to obtain the discharge timeseries. Six calibrations have been performed – four upstream and two downstream- and the resulting KGE at the gauging station calibrated have been increased meanly of 0.37 in comparison with the uncalibrated performances. All the calibrations have been performed firstly with

just 1 round and 100 steps and then with 3 rounds and 200 steps. The increase of accuracy between the first and second attempt is evident looking at the discharge timeseries which gives a better agreement between the observed and simulated values and also by looking at the increase of 0.01-0.10 in the KGE value.
The highest performance was obtained in the first downstream calibration resulting with KGE of 0.79 and 0.81 for the calibration and validation period, it gives proof of a good modelling of the hydrological processes once the flow arrives in the plain regions where the main course of the Dora Baltea is located. The upstream calibrations result with a KGE of 0.57-0.62 which is good considering all the snow processes involved and the difficulties in modelling mountainous regions. The result in terms of KGE value at the outlet basin is of 0.68 for the calibration period which is considered good comparing the value obtained for the Valle d'Aosta part inside the model developed by "*Autorità di Bacino del Fiume Po"* in the Po district [52]. The reference model proposes a finer discretization of the basin and more calibrations which leads to a KGE of 0.76 at the outlet basin – which is 40 km above the basin outlet of the current model. The difference – even though not directly comparable since the outlet is different – is of only 0.08 and, considering the higher accuracy of the reference model compared with the one developed, it has been considered particularly good.
The validation period was chosen before the calibration, having a back casting performance. Overall, the validation KGE values are a bit lower than the calibration ones, but downstream they gave better performance.
Finally, the water balance analysis was done resulting with a reliable performance. The balance was not totally closed in all years, but the highest unbalance is of 10% during the calibration period. However, the groundwater contribution was highly reduced from the uncalibrated and calibrated result, having a better adherence with the reality of the territory which has not so many aquifers. The water balance gives worst results during the validation period due to the presence of some extreme events – like the flood in 2000 and the drought in 2003 – and the absence of measured data. Whereas the modelling of the timing of solid or liquid precipitation as well as the snow melting has been modelled exceptionally good looking at the values of some mountainous sub-basins that have the highest accumulation of snow during winter and the lowest during summer, as expected from an alpine basin.

The model quantifies the water flowing inside the Dora Baltea catchment with a good adherence with the reality which gives the possibility to use it to access the water availability in the territory. The work contributes to the LIFE CLIMAX PO project which "*aims at a climate-smart management of water resources in the Po River basin*" [48] and this model could be used as the start point of one of the possible solutions to address the future challenges in the water management of the territory.
The concluding post-processing analysis of the "Canale Farini" aims to evaluate the quantity of water diverted for irrigation by the canal since no direct data are available. The study assesses the quantity of water that could be drawn by the canal during summer from 2019 to 2023. It was carried out considering the comparison between the simulated and observed data at the outlet basin where a visible excess of water has been simulated

by the model during the summer periods. That difference quantity was considered a good proxy of the water diverted by the canal 5 km above the outlet. Indeed, it is comparable with other simulations of the discharge of the canal which have the highest quantity during summer. In addition, the summer mean difference has been used to evaluate a correction factor in order to improve the simulated discharge. Applying the factor in the summer period, the timeseries result in a better agreement with the observed ones and, above all, the yearly cumulative volume adheres much better to the measured one. This post-processing approach lays the foundation for the inclusion of the anthropogenic structures built inside the catchment. Thus, it could be the starting point for a better implementation of the quantification inside the model and give access to a quantitative analysis of the Dora Baltea water available to agriculture and balance it with all the other competitive uses.

On the other hand, the model is far to be definitive, and indeed some detailed could be added. The model simulates the natural flow of the Dora Baltea River without considering all the anthropogenic structures present in the territory, such as the several dams and some canals which divert water for irrigation purposes. By adding dams - which plays a crucial role in the hydrological lamination of the basin - the reservoirs of the river could be assessed with a much more reliability. Indeed, all the upstream calibrations performed in correspondence of such zones where dams are present, evidenced the lower obtained accuracy in discharge timeseries since the modelled flow does not take into consideration the managing of the discharge done by the dam. As well as the assessment of the water diverted by Canale Farini which could be modelled better by introducing a point at the confluence during the first step of the geomorphology analysis in order to have access to the discharge arriving exactly at the canal. Thus, performing a better post-processing estimation of the water diverted.
Furthermore, the exclusion of the *Lago di Viverone* at the end of the catchment could improve the final performance of the model by giving better results in the water balance. Indeed, the lake has been excluded since it seems to not give a such important contribute to the river unless with groundwater recharge. However, information is missing, and a further investigation is needed to include it inside the model.

In addition, precipitation data registered by the meteorological station have a vast number of gaps due to detection instrument issues related to the low temperature leading to ice formation. The results could be improved by introducing heated instruments and so having much more reliable data on which to base the spatial interpolation performed by the software.
Finally, the radiation estimation was carried out maintaining the default setting of the software without investigate the possibility to introduce the cloud cover contribution or a diversification of the territory. For instance, the albedo effect could be increased at highest elevation due to the presence of snow and ice, whereas lowered down in the plain part of the territory. In this way, the radiation budget could be improved in accuracy without an overall overestimation due to the cloudless sky effect and, as consequence, a better modelling performance of the evapotranspiration processes could be obtained. In

addition, the number of the calibrated sections could be increased to obtain sets of specific parameters for as many sub-basins as possible.

In conclusion, the project accomplished the objectives set by achieving good accuracy in the results while decreasing the computational demand and, simultaneously, the time required to construct the model. The performances are robust, and the study provides meaningful quantitative information of the water budget of the Dora Baltea Basin usable for a better understanding of the water availability of the territory. The model should be a key step for the building of a future sustainable management of the water resources. With the hope of achieving even better results by making improvements, the project lays the foundation for future developments.

# REFERENCES

[1] Tarquini, Simone, Ilaria Isola, Massimiliano Favalli, Alessandro Battistini, and Giulia Dotta. 2023. 'TINITALY, a Digital Elevation Model of Italy with a 10 Meters Cell Size (Version 1.1)'. Application/zip,image/tiff. Istituto Nazionale di Geofisica e Vulcanologia (INGV). Accessed May 20, 2024 https://doi.org/10.13127/TINITALY/1.1

[2] 'Arpa Piemonte - Bacini Idrografici'. 2024. Geoportale Piemonte. Accessed 20 May 2024. https://www.geoportale.piemonte.it/geonetwork/srv/api/records/arlpa_to:12.01.01-D_2011-06-21-11:30.

[3] 'QGIS Development Team'. 2024. 'QGIS Geographic Information System'. Windows. http://qgis.org
QuickOSM plugin. 2024. 'QuickOSM Plugin for QGIS'. https://plugins.qgis.org/plugins/QuickOSM/.

[4] Centro Funzionale Regione Autonoma Valle d'Aosta. 2024. 'Dataview - Centro Funzionale Regione Autonoma Valle d'Aosta' [Online]. Dataview VDA. Accessed 21 May 2024 https://presidi2.regione.vda.it/str_dataview.

[5] ARPA Piemonte. 2024. 'Arpa Piemonte – Accesso Richiesta Dati' [Online]. ARPA Piemonte - Agenzia Regionale per La Protezione Ambientale. Accessed 21 May 2024. https://www.arpa.piemonte.it/rischi_naturali/snippets_arpa_graphs/map_meteoweb/?rete=stazione_meteorologica.

[6] Bancheri Marialaura, Rigon Riccardo and Manfreda Salvatore. 2020. 'The GEOframe-NewAge Modelling System Applied in a Data Scarce Environment'. *Water* 12 (1): 86. https://doi.org/10.3390/w12010086.

[7] Formetta, Rigon GWS-2024. 'GEOframe: GEOframe Winter School 2024: Day 08-01-2024'. Hydro-geomorphology: Part 1, Part 2, Part 3. [Online]. Accessed 17 August 2024. https://geoframe.blogspot.com/2024/01/geoframe-winter-school-2024-day-08-01.html.

[9] Orlandini, Stefano, Giovanni Moretti, Marco Franchini, Barbara Aldighieri, and Bruno Testa. 2003. 'Path-based Methods for the Determination of Nondispersive Drainage Directions in Grid-based Digital Elevation Models'. *Water Resources Research* 39 (6): 2002WR001639. https://doi.org/10.1029/2002WR001639.

[10] 'Tangential Curvature'. n.d. Accessed 17 August 2024. https://surferhelp.goldensoftware.com/gridops/tangential_curvature.htm.

[11] Autorità di Bacino del Fiume Po. 'Linee Generali di assetto idrogeologico e quadro degli interventi Bacino della Dora Baltea'. [Online]. Accessed 21 August 2024. https://www.adbpo.it/PAI/3%20-%20Linee%20generali%20di%20assetto%20idraulico%20e%20idrogeologico/3.3%20-%20Elaborato%20Piemonte/Dora_Baltea.pdf.

[12] ‘Dora Baltea’. 2024. In *Wikipedia*. Accessed 24 August 2024. https://en.wikipedia.org/w/index.php?title=Dora_Baltea&oldid=1241477648.

[13] ‘Dora Baltea’. 2024. From openstreetmap.org, OpenStreetMap contributors. In *Wikipedia*. Accessed 24 August 2024. https://it.wikipedia.org/wiki/Dora_Baltea#/media/File:Doire_Balt%C3%A9e.png.

[14] Crosio, Roberto. ‘Ripercorrendo l’area Del Canale Cavour’. Le Vocazioni Ambientali Del Vercellese, Terra d’acque. Dagli Interventi Di Camillo Cavour Ai Giorni Nostri. Accessed 26 August 2024. https://www.roberto-crosio.net/1_VERCELLESE/turismo_carta_canale.htm.

[15] ‘Presa del Canale Farini a Saluggia | Aree protette Po piemontese’. Accessed 26 August 2024. http://www.parcopopiemontese.it/pun-dettaglio.php?id=970.

[16] Virdee, T. S., and N. T. Kottegoda. 1984. ‘A Brief Review of Kriging and Its Application to Optimal Interpolation and Observation Well Selection’. *Hydrological Sciences Journal* 29 (4): 367–87. https://doi.org/10.1080/02626668409490957.

[17] Formetta, Giuseppe, and Riccardo Rigon. n.d. ‘GEOframe: GEOframe Winter School 2024: Day 09-01-2024. Kriging: Part 1’. Accessed 27 August 2024. https://osf.io/78h4u.

[18] Formetta, Giuseppe, and Riccardo Rigon. n.d. ‘GEOframe: GEOframe Winter School 2024: Day 09-01-2024. Kriging: Part 3’. Accessed 27 August 2024. https://osf.io/8fe9j.

[19] Formetta, Giuseppe, and Riccardo Rigon. 2024. ‘GEOframe: GEOframe Winter School 2024: Day 09-01-2024. Kriging: Part 2’, January. https://osf.io/https://osf.io/mz32s.

[20] Bancheri, Marialaura, Francesco Serafin, Michele Bottazzi, Wuletawu Abera, Giuseppe Formetta, and Riccardo Rigon. 2018. ‘The Design, Deployment, and Testing of Kriging Models in GEOframe with SIK-0.9.8’. *Geoscientific Model Development* 11 (6): 2189–2207. https://doi.org/10.5194/gmd-11-2189-2018.

[21] Mohd Aziz, Mohd Khairul Bazli, Fadhilah Yusof, Zalina Mohd Daud, Zulkifli Yusop, and Mohammad Afif Kasno. 2019. ‘Comparison of Semivariogram Models in Rain Gauge Network Design’. *MATEMATIKA* 35 (July):157–70. https://doi.org/10.11113/matematika.v35.n2.1155.

[22] Centro Funzionale Regione Autonoma Valle d’Aosta. 2024. ‘Anagrafica delle stazioni attive’ [Online]. Accessed 21 May 2024 https://cf.regione.vda.it/media/docs/anagrafica-stazioni-attive.pdf.

[23] Braca G., Bussettini M., Lastoria B., Mariani S., Piva F., 2021, "Temperatura media annua media" in Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare – BIGBANG: metodologia e stime. Rapporto sulla disponibilità naturale della risorsa idrica. Page 85. Istituto Superiore per la Protezione e la Ricerca Ambientale, Rapporti 339/21, Roma.

[24] 'Temperature'. Centro Funzionale Regione Autonoma Valle d'Aosta. Accessed 31 August 2024. http://cf.regione.vda.it/it/temperature.

[25] Braca G., Bussettini M., Lastoria B., Mariani S., Piva F., 2021, "Precipitazioni annue medie" in Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare – BIGBANG: metodologia e stime. Rapporto sulla disponibilità naturale della risorsa idrica. Page 80. Istituto Superiore per la Protezione e la Ricerca Ambientale, Rapporti 339/21, Roma.

[26] 'Precipitazioni'. Centro Funzionale Regione Autonoma Valle d'Aosta. Accessed 03 September 2024. https://cf.regione.vda.it/it/precipitazioni.

[27] Formetta, Giuseppe, Marialaura Bancheri, Olaf David, and Riccardo Rigon. 2016. 'Performance of Site-Specific Parameterizations of Longwave Radiation'. *Hydrology and Earth System Sciences* 20 (11): 4641–54. https://doi.org/10.5194/hess-20-4641-2016.

[28] Idso, S. B.: A set of equations for full spectrum and 8-to 14-µm and 10.5-to 12.5-µm thermal radiation from cloudless skies, Water Resour. Res., 17, 295–304, 1981.

[29] Formetta, G., R. Rigon, J. L. Chávez, and O. David. 2013. 'Modeling Shortwave Solar Radiation Using the JGrass-NewAge System'. *Geoscientific Model Development* 6 (4): 915–28. https://doi.org/10.5194/gmd-6-915-2013.

[30] Corripio, J.: Modelling the energy balance of high altitude glacierised basins in the Central Andes., Ph.D. dissertation, University of Edinburgh, 2002.

[31] Corripio, J.: Vectorial algebra algorithms for calculating terrain parameters from DEMs and solar radiation modelling in mountainous terrain, Int. J. Geogr. Inf. Sci., 17, 1–24, 2003.

[32] Iqbal, M.: An Introduction to Solar Radiation, Academic Press, Orlando, FL, USA, 1983

[33] Bancheri, Marialaura. 2016. 'JGrass-NewAge: Net Radiation Component'. SlideShare. 12 April 2016. https://www.slideshare.net/slideshow/jgrassnewage-net-radiation-component/60834675.

[34] Allen, Richard, Luis Pereira, Dirk Raes, and Martin Smith. 'Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements - FAO Irrigation and Drainage Paper 56'. Chapter 2 - FAO Penman-Monteith equation. FAO - Food and Agriculture organization of the United Nations, 1998. https://www.fao.org/4/X0490E/x0490e06.htm.

[35] Allen, Richard, Luis Pereira, Dirk Raes, and Martin Smith. 1998. 'Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements - FAO Irrigation and Drainage Paper 56'. Chapter 6 - ETc - Single crop coefficient (Kc). FAO - Food and Agriculture Organization of the United Nations. 1998. https://www.fao.org/4/X0490E/x0490e0b.htm.

[36] Regione Autonoma Valle d'Aosta. n.d. 'Carta di Copertura del Suolo 2020'. *Geoportale SCT* (blog). Accessed 10 July 2024. https://geoportale.regione.vda.it/download/carta-copertura-suolo/.

[37] Geoportale Regione Piemonte. n.d. 'Dati Scaricabili Da Geocatalogo'. Accessed 10 July 2024. https://geoportale.igr.piemonte.it/cms/.

[38] Braca G., Bussettini M., Lastoria B., Mariani S., Piva F., 2021, "Evaporazione Potenziale Annua Media (LTAA)" in Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare – BIGBANG: metodologia e stime. Rapporto sulla disponibilità naturale della risorsa idrica. Page 95. Istituto Superiore per la Protezione e la Ricerca Ambientale, Rapporti 339/21, Roma.

[39] PSR, Pianeta. n.d. 'L'andamento agrometeorologico di marzo'. www.pianetapsr.it. Accessed 11 July 2024. https://www.pianetapsr.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/2725.

[40] Mastrotheodoros, Theodoros, Christoforos Pappas, Peter Molnar, Paolo Burlando, Gabriele Manoli, Juraj Parajka, Riccardo Rigon, et al. 2020. 'More Green and Less Blue Water in the Alps during Warmer Summers'. *Nature Climate Change* 10 (2): 155–61. https://doi.org/10.1038/s41558-019-0676-5.

[41] Bancheri, M. 2017. 'A Flexible Approach to the Estimation of Water Budgets and Its Connection to the Travel Time Theory.' Ph.D. Thesis, University of Trento, Trento, Italy.

[42] Busti, Riccardo, Daniele Andreis, Giuseppe Formetta, and Riccardo Rigon. 2024. 'GEOframe: GEOframe Winter School 2024: Day 12-01-2024'. Slide 3/27. January 12. https://osf.io/https://osf.io/utzm9.

[43] Formetta, Giuseppe. 2024. 'OMS_Project_Brenta_GWS24_ERM'. GeoFrame [software] Open Science Framework. Retrieved from: https://osf.io/https://osf.io/7yrwg.

[44] Gupta, Hoshin V., Harald Kling, Koray K. Yilmaz, and Guillermo F. Martinez. 2009. 'Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling'. *Journal of Hydrology* 377 (1): 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

[45] Hay, Lauren E., and Makiko Umemoto. 2007. 'Multiple-Objective Stepwise Calibration Using Luca'. Open-File Report. US Geological Survey. https://doi.org/10.3133/ofr20061323.

[46] Martin. n.d. ‘Water and Sanitation.’ United Nations Sustainable Development. Accessed 22 September 2024. https://www.un.org/sustainabledevelopment/water-and-sanitation/.

[47] Chow, Maidment, and Mays. 1988. ‘Applied hydrology’. McGraw-Hill

[48] ‘Il Progetto’. n.d. LifeClimaxPo. Accessed 23 September 2024. https://www.lifeclimaxpo.adbpo.it/il-progetto/.

[49] ‘Homepage’. n.d. Coutenza Canali Cavour. Accessed 1 October 2024. https://coutenzacanalicavour.it/.

[50] ‘Regione Autonoma Valle d’Aosta - GeoNavITG SCT v1.0 - Navigatore Integrabile’. n.d. Accessed 3 October 2024. https://mappe.regione.vda.it/pub/geonavitg/geodighe.asp.

[51] Roati, Gaia, and Giuseppe Formetta. 2024. ‘GEOframe: GEOframe Winter School 2024: Day 11-01-2024. Po River Application’. January 11. https://osf.io/https://osf.io/32mp8.

[52] Roati, Gaia, Giuseppe Formetta, Silvano Pecora, Marco Brian, Riccardo Rigon, and Hervè Stevenin. 2022. ‘Hydrological Modeling and Water Budget Quantification of the Po River Basin through the GEOframe System’. EGU22-12562. Copernicus Meetings. https://doi.org/10.5194/egusphere-egu22-12562.

[53] ‘Project Jupyter’. n.d. Accessed 6 October 2024. https://jupyter.org.

# APPENDIX A

## A1 Figures and tables

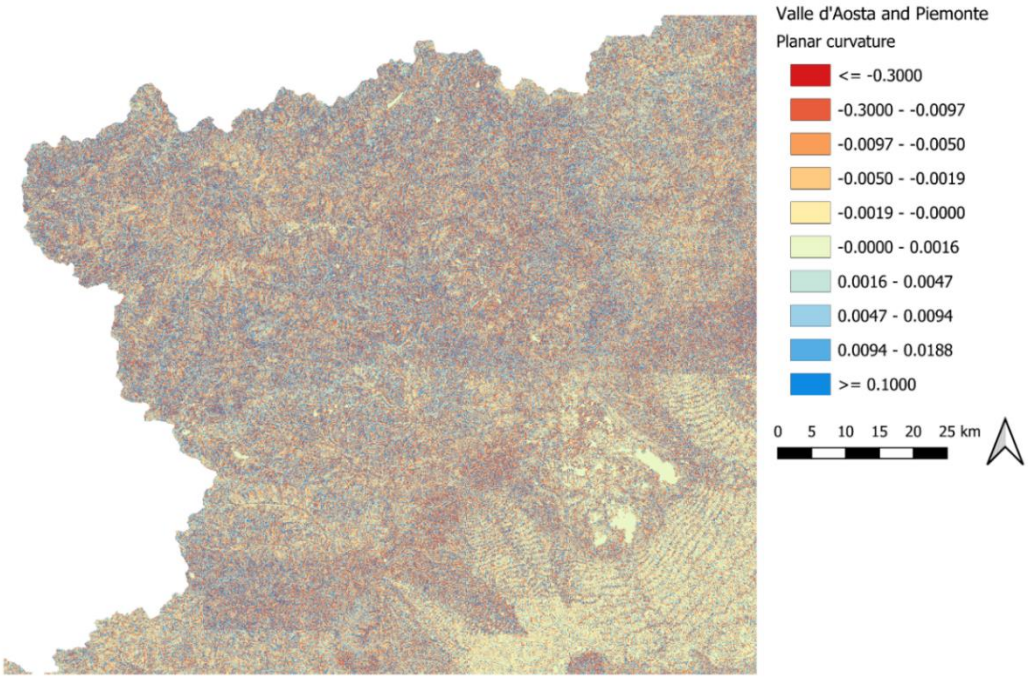



*Figure A1.0.1. planar curvature*

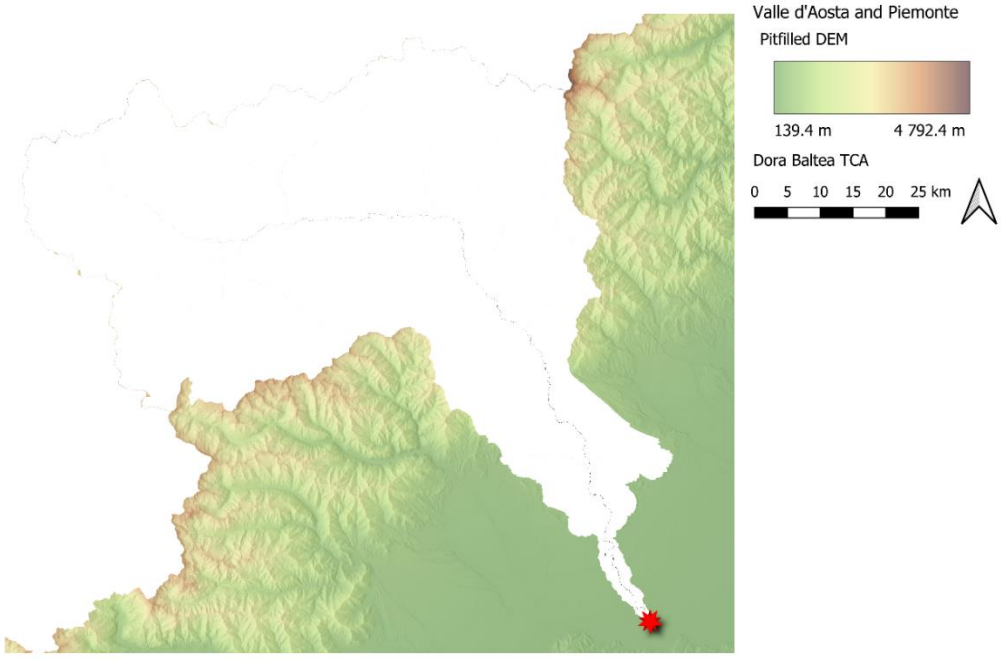



*Figure A1.0.2, Dora Baltea Total Contributing Area*

*Table A1.1, metadata of each sub-basin of the Dora Baltea basin*

| HRU ID | Area [km$^2$] | UTMX [m] | UTMY [m] | Centroid elevation (m.a.s.l.) [m] | Mean elevation. (m.a.s.l.) [m] |
|---|---|---|---|---|---|
| **22** | 63.00 | 408708.11 | 5030566.99 | 230.26 | 284.88 |
| **82** | 61.29 | 394620.77 | 5087596.06 | 2174.36 | 2637.08 |
| **99** | 83.96 | 386052.85 | 5087014.27 | 2433.50 | 2763.61 |
| **148** | 40.80 | 379828.97 | 5082325.47 | 1639.42 | 2481.35 |
| **168** | 44.97 | 393484.10 | 5082028.00 | 1670.39 | 2400.81 |
| **217** | 47.60 | 376014.28 | 5079322.02 | 1483.06 | 2274.00 |
| **220** | 71.39 | 347861.31 | 5080786.20 | 1748.15 | 2505.37 |
| **226** | 77.99 | 402006.18 | 5082218.75 | 2248.44 | 2719.19 |
| **249** | 56.52 | 408633.35 | 5081060.13 | 2489.34 | 2831.02 |
| **258** | 47.20 | 400772.87 | 5074819.28 | 1944.31 | 2198.53 |
| **267** | 69.32 | 356422.84 | 5076547.75 | 1940.80 | 2209.14 |
| **274** | 94.82 | 334734.15 | 5074052.56 | 2634.31 | 2659.34 |
| **276** | 91.73 | 369978.00 | 5079767.63 | 2042.23 | 2202.49 |
| **281** | 40.43 | 362420.99 | 5078519.64 | 1966.76 | 2183.09 |
| **309** | 41.66 | 342638.62 | 5075704.96 | 1786.69 | 2053.77 |
| **347** | 45.24 | 365560.21 | 5073516.60 | 1144.10 | 1669.27 |
| **373** | 85.10 | 390652.76 | 5076048.73 | 1284.07 | 1930.82 |
| **393** | 47.61 | 409122.15 | 5073151.50 | 1708.19 | 2234.62 |
| **404** | 78.60 | 381860.03 | 5074116.14 | 2206.18 | 2105.09 |
| **406** | 42.52 | 344494.48 | 5068984.46 | 1382.74 | 1775.03 |
| **414** | 41.34 | 340231.69 | 5066288.06 | 1665.94 | 1986.74 |
| **436** | 41.50 | 401193.40 | 5069864.51 | 2105.52 | 1900.17 |
| **457** | 40.34 | 386851.24 | 5066254.46 | 518.92 | 1485.24 |
| **461** | 44.90 | 375257.61 | 5066087.56 | 537.93 | 1707.08 |
| **465** | 40.27 | 391185.34 | 5067313.18 | 728.61 | 1429.72 |
| **468** | 56.17 | 383535.46 | 5057910.38 | 2086.69 | 2184.08 |
| **469** | 41.90 | 382188.13 | 5066146.37 | 507.39 | 1045.69 |
| **471** | 41.67 | 378585.59 | 5061311.19 | 2019.72 | 1884.81 |
| **478** | 82.16 | 349185.05 | 5068603.39 | 1040.20 | 2018.40 |
| **481** | 71.61 | 371074.84 | 5066000.82 | 551.99 | 1552.16 |
| **492** | 84.91 | 366326.33 | 5064236.13 | 617.16 | 1504.46 |
| **519** | 57.85 | 334498.09 | 5065458.24 | 2746.46 | 2439.72 |
| **529** | 45.20 | 353494.47 | 5067408.83 | 2056.36 | 2026.91 |
| **550** | 41.29 | 361080.92 | 5064450.58 | 912.66 | 1494.58 |
| **557** | 43.72 | 404721.81 | 5064908.67 | 1854.37 | 1895.07 |
| **562** | 44.77 | 355815.64 | 5062774.05 | 816.61 | 1473.55 |
| **583** | 47.57 | 341871.78 | 5059478.53 | 2208.09 | 2529.44 |
| **599** | 41.40 | 358901.25 | 5056967.49 | 1876.39 | 1691.24 |
| **624** | 75.23 | 395889.23 | 5065300.98 | 408.88 | 1113.39 |
| **628** | 74.84 | 411708.21 | 5064373.99 | 1166.02 | 1935.38 |
| **636** | 40.08 | 349696.80 | 5057986.92 | 1725.76 | 2322.68 |
| **647** | 74.20 | 366133.67 | 5054943.98 | 2023.59 | 2204.03 |

| | | | | | |
|---|---|---|---|---|---|
| **681** | 53.29 | 391435.26 | 5058219.98 | 1758.02 | 1772.20 |
| **704** | 41.84 | 354970.51 | 5052644.82 | 1636.59 | 2228.98 |
| **737** | 46.36 | 348923.15 | 5051808.78 | 2084.67 | 2475.32 |
| **767** | 55.46 | 411617.16 | 5056897.75 | 1037.70 | 1748.67 |
| **788** | 84.19 | 400334.99 | 5057068.84 | 1050.54 | 1254.29 |
| **807** | 39.36 | 387201.36 | 5051828.79 | 2214.27 | 2406.29 |
| **818** | 41.29 | 361315.36 | 5049178.46 | 1978.12 | 2539.62 |
| **821** | 68.26 | 368870.00 | 5046889.03 | 2323.27 | 2645.25 |
| **848** | 40.85 | 375480.46 | 5053960.74 | 2628.56 | 2550.44 |
| **852** | 71.48 | 395220.77 | 5050233.31 | 1479.38 | 1707.52 |
| **853** | 43.10 | 408688.08 | 5053277.23 | 714.96 | 1438.28 |
| **897** | 69.97 | 376785.81 | 5048121.43 | 2510.33 | 2556.25 |
| **905** | 56.71 | 345420.22 | 5046858.71 | 2193.09 | 2675.20 |
| **919** | 40.92 | 402677.25 | 5048986.93 | 795.60 | 1161.52 |
| **936** | 74.82 | 351767.12 | 5043340.04 | 2250.88 | 2646.12 |
| **962** | 80.91 | 359774.25 | 5042399.12 | 1999.03 | 2701.17 |
| **999** | 51.02 | 407667.84 | 5046726.53 | 292.12 | 1205.14 |
| **1060** | 45.79 | 395547.84 | 5043572.81 | 1242.80 | 1905.91 |
| **1064** | 53.28 | 402146.51 | 5039230.98 | 694.86 | 1096.62 |
| **1073** | 210.03 | 415617.57 | 5035048.94 | 237.20 | 381.58 |
| **1077** | 52.43 | 400227.57 | 5034782.66 | 767.41 | 967.78 |
| **1078** | 150.74 | 416843.05 | 5020078.61 | 231.05 | 237.51 |

*Table A1.1, metadata of the meteorological stations in the Dora Baltea Basin*

| ID | Name | UTMX [m] | UTMY [m] | Latitude [°] | Longitude [°] | Elevation (m.a.s.l.) [m] | Temperature Data | Precipitation Data |
|---|---|---|---|---|---|---|---|---|
| **1000** | Arvier - Chamençon | 352606.7853 | 5060047.328 | 45.6783 | 7.10742 | 1238 | ✓ | ✓ |
| **1020** | Aymavilles - Ponte Dora Baltea | 363989.746 | 5063401.088 | 45.7108 | 7.25257 | 618 | ✓ | × |
| **1030** | Aymavilles - Vieyes | 363703.6647 | 5056615.922 | 45.6497 | 7.2508 | 1139 | ✓ | ✓ |
| **1040** | Bionaz - Place Moulin | 382972.7989 | 5084010.108 | 45.8997 | 7.49137 | 1979 | ✓ | ✓ |
| **1080** | Chamois - Lac de Lou | 393058.4801 | 5077703.957 | 45.8446 | 7.62275 | 2020 | ✓ | ✓ |
| **1090** | Champdepraz - Chevrère | 392035.303 | 5059817.853 | 45.6835 | 7.61357 | 1260 | ✓ | ✓ |
| **1100** | Champdepraz - Ponte Dora Baltea | 396715.4042 | 5059549.644 | 45.6818 | 7.67371 | 370 | ✓ | × |
| **1110** | Champorcher - Chardonney | 391538.592 | 5052836.214 | 45.6206 | 7.60875 | 1430 | ✓ | ✓ |
| **1120** | Champorcher - Rifugio Dondena | 387085.265 | 5051992.684 | 45.6123 | 7.55184 | 2181 | ✓ | ✓ |
| **1130** | Cogne - Crétaz | 370610.4245 | 5052556.565 | 45.6145 | 7.34048 | 1470 | ✓ | × |
| **1140** | Cogne - Lillaz | 374571.6067 | 5050319.63 | 45.5951 | 7.39184 | 1613 | ✓ | ✓ |
| **1150** | Cogne - Valnontey | 370278.5061 | 5049340.213 | 45.5855 | 7.33708 | 1682 | ✓ | ✓ |
| **1160** | Cogne - Grand-Crot | 372652.479 | 5049569.294 | 45.588 | 7.36744 | 2279 | ✓ | ✓ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1170** | Courmayeur - Ferrachet | 346902.6546 | 5081139.987 | 45.8668 | 7.02754 | 2290 | ✓ | ✓ |
| **1190** | Courmayeur - Lex Blanche | 331909.4699 | 5070366.013 | 45.7664 | 6.83826 | 2162 | ✓ | ✓ |
| **1200** | Courmayeur - Mont de la Saxe | 343231.2415 | 5075695.417 | 45.817 | 6.98204 | 2110 | ✓ | ✓ |
| **1220** | Courmayeur - Pré-de-Bard | 349875.1715 | 5083012.664 | 45.8843 | 7.06523 | 2040 | ✓ | ✓ |
| **1230** | Fénis - Clavalité | 383241.7406 | 5060254.179 | 45.686 | 7.50058 | 1531 | ✓ | ✓ |
| **1240** | Fénis - Lavodilec | 381952.8675 | 5055699.515 | 45.6448 | 7.48514 | 2250 | ✓ | ✓ |
| **1250** | Gressan - Pila-Leissé | 368261.0979 | 5058140.833 | 45.6643 | 7.30885 | 2280 | ✓ | ✓ |
| **1260** | Gressoney - la - Trinité | 407451.8649 | 5080106.134 | 45.8683 | 7.80761 | 1992 | ✓ | ✓ |
| **1270** | Gressoney - la - Trinité - Eselbode | 408807.5251 | 5075896.53 | 45.8306 | 7.82587 | 1642 | ✓ | ✓ |
| **1300** | Gressoney - Saint - Jean - Weissmatten | 408609.5796 | 5066764.871 | 45.7484 | 7.82505 | 2038 | ✓ | ✓ |
| **1310** | Hone - ponte Dora Baltea | 401759.3699 | 5052289.332 | 45.6172 | 7.73993 | 340 | ✓ | ✓ |
| **1320** | Issime - Capoluogo | 410936.2425 | 5059941.477 | 45.6873 | 7.85621 | 960 | ✓ | ✓ |
| **1340** | La Thuile - La Grande Tete | 337707.8398 | 5061017.704 | 45.6837 | 6.91591 | 2430 | ✓ | ✓ |
| **1360** | La Thuile - Villaret | 340909.9122 | 5063803.385 | 45.7095 | 6.95609 | 1488 | ✓ | ✓ |
| **1370** | Lillianes - Granges | 410900.3279 | 5053385.772 | 45.6283 | 7.85695 | 1256 | ✓ | ✓ |
| **1390** | Morgex - Lavancher | 346131.0439 | 5073199.328 | 45.7952 | 7.02014 | 2842 | ✓ | ✓ |
| **1430** | Nus - Les Iles | 378497.456 | 5066224.284 | 45.7389 | 7.43818 | 534 | ✓ | ✓ |
| **1440** | Ollomont - By | 368135.8862 | 5082274.403 | 45.8814 | 7.30066 | 2017 | ✓ | ✓ |
| **1470** | Pontboset - Fournier | 397055.3501 | 5050964.827 | 45.6046 | 7.67989 | 1087 | ✓ | ✓ |
| **1480** | Pontey - Ponte Dora Baltea | 392441.9757 | 5066223.234 | 45.7412 | 7.61737 | 473 | × | ✓ |
| **1490** | Pré - Saint - Didier - Capoluogo | 343759.0725 | 5069667.736 | 45.7629 | 6.99078 | 996 | ✓ | ✓ |
| **1500** | Pré - Saint - Didier - Plan Praz | 340730.2666 | 5068832.982 | 45.7547 | 6.95213 | 2177 | ✓ | ✓ |
| **1520** | Rhemes-Notre-Dame - Chanavey | 353590.8288 | 5048963.78 | 45.5788 | 7.12338 | 1690 | ✓ | ✓ |
| **1530** | Rhemes - Saint-Georges - Feleumaz | 353742.6595 | 5052183.826 | 45.6078 | 7.12436 | 2362 | ✓ | ✓ |
| **1550** | Rhemes - Saint-Georges - Capoluogo | 356051.2713 | 5056787.727 | 45.6497 | 7.15259 | 1179 | ✓ | ✓ |
| **1570** | Saint-Oyen - Moulin | 360850.0493 | 5075708.786 | 45.8209 | 7.20871 | 1310 | ✓ | × |
| **1590** | Saint-Rhémy-en-Bosses - Crévacol | 356767.9927 | 5076713.206 | 45.8291 | 7.15589 | 2018 | ✓ | ✓ |
| **1600** | Saint-Rhémy-en-Bosses - Gran San Bernardo | 356696.2798 | 5081205.794 | 45.8695 | 7.15363 | 2360 | ✓ | × |
| **1610** | Saint-Rhémy-en-Bosses - Mont-Botsalet | 356268.4144 | 5078092.052 | 45.8414 | 7.14905 | 2500 | ✓ | × |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1620** | Valgrisenche - Menthieu | 348155.1805 | 5052972.941 | 45.6137 | 7.0525 | 1859 | ✓ | ✓ |
| **1630** | Valpelline - Chosoz | 371033.8785 | 5076155.64 | 45.8269 | 7.33963 | 1029 | ✓ | ✓ |
| **1650** | Valsavarenche - Eaux-Rousses | 360190.1515 | 5047478.995 | 45.5668 | 7.20835 | 1651 | ✓ | ✓ |
| **1670** | Valsavarenche - Orvieille | 358943.9889 | 5048963.071 | 45.5799 | 7.19196 | 2170 | ✓ | ✓ |
| **1690** | Valtournenche - Cime Bianche | 398527.7824 | 5085791.457 | 45.9182 | 7.69146 | 3100 | ✓ | × |
| **1700** | Valtournenche - Lago Goillet | 396322.7533 | 5087283.91 | 45.9313 | 7.66271 | 2541 | ✓ | ✓ |
| **1710** | Valtournenche - Grandes Murailles | 391688.5372 | 5087163.328 | 45.9295 | 7.60298 | 2566 | ✓ | ✓ |
| **1720** | Valtournenche - Maen | 392190.2244 | 5079719.455 | 45.8626 | 7.61112 | 1310 | ✓ | ✓ |
| **1730** | Ayas - Alpe Aventine | 400630.6828 | 5080867.473 | 45.8742 | 7.71959 | 2045 | ✓ | ✓ |
| **1760** | Hone - Ayasse | 401191.4866 | 5051687.088 | 45.6117 | 7.73277 | 367 | ✓ | ✓ |
| **1770** | Pré-Saint-Didier - Gare | 343698.5398 | 5069735.961 | 45.7635 | 6.98998 | 1000 | ✓ | × |
| **1780** | Gressoney -la-Trinité - Diga Gabiet | 410534.8258 | 5077538.271 | 45.8456 | 7.8478 | 2373 | ✓ | ✓ |
| **2500** | Arvier - Cooperativa Enfer | 357236.123 | 5062907.523 | 45.705 | 7.16599 | 738 | ✓ | ✓ |
| **2510** | Jovençan - Pompiod | 364975.5136 | 5063146.219 | 45.7087 | 7.2653 | 670 | ✓ | ✓ |
| **2520** | Quart - Ollignan | 373518.2287 | 5067301.597 | 45.7477 | 7.37392 | 650 | ✓ | ✓ |
| **2530** | Saint-Pierre - Lago delle Rane | 359386.1852 | 5069372.42 | 45.7636 | 7.19172 | 2370 | ✓ | × |
| **2560** | Gressoney - Saint - Jean - lago di Seebna | 414814.6735 | 5065854.505 | 45.741 | 7.90497 | 2270 | ✓ | × |
| **3000** | Verrés - Capoluogo | 397649.8272 | 5057678.369 | 45.6651 | 7.6861 | 375 | ✓ | ✓ |
| **3010** | Saint-Denis - Raffort | 389272.8221 | 5067245.709 | 45.7499 | 7.57641 | 840 | ✓ | ✓ |
| **3020** | Villeneuve - S.R. Saint-Nicolas | 360464.7498 | 5063101.137 | 45.7074 | 7.20739 | 839 | ✓ | ✓ |
| **3030** | Morgex - Capoluogo | 347456.0146 | 5069075.661 | 45.7584 | 7.03848 | 938 | ✓ | ✓ |
| **3040** | Gressoney - Saint - Jean - Bieltschocke | 409262.6414 | 5068866.69 | 45.7674 | 7.83305 | 1370 | ✓ | ✓ |
| **3050** | Champorcher - Petit-Mont-Blanc | 391602.6258 | 5053290.745 | 45.6247 | 7.60947 | 1640 | ✓ | ✓ |
| **3060** | Rhemes-Notre-Dame - Chausanne | 352790.6706 | 5046937.234 | 45.5604 | 7.11374 | 1794 | ✓ | ✓ |
| **3070** | Valsavarenche - Pont | 359511.5045 | 5043048.074 | 45.5268 | 7.20093 | 1951 | ✓ | ✓ |
| **3080** | Nus - Saint-Berthélemy - Osservatorio | 381707.9956 | 5071841.752 | 45.79 | 7.47806 | 1675 | ✓ | ✓ |
| **3110** | Courmayeur - Dolonne | 341954.1483 | 5073226.42 | 45.7945 | 6.96642 | 1200 | ✓ | ✓ |
| **3120** | Saint-Vincent - Terme | 395198.1224 | 5067098.602 | 45.7495 | 7.6526 | 626 | ✓ | ✓ |
| **3530** | Bard - Albard | 402526.6235 | 5052032.84 | 45.615 | 7.74982 | 662 | ✓ | ✓ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **3550** | Gressoney -la-Trinité - D'Ejola | 408002.2151 | 5078742.175 | 45.8561 | 7.81496 | 1837 | ✓ | ✓ |
| **3560** | Roisan - Preyl | 369145.1753 | 5071193.394 | 45.7819 | 7.31667 | 935 | ✓ | ✓ |
| **3570** | Brusson - Tchampats | 400309.9199 | 5068425.93 | 45.7622 | 7.71803 | 1288 | ✓ | ✓ |
| **3590** | Valtournenche - Breuil Cervinia | 393494.7446 | 5087343.105 | 45.9314 | 7.62623 | 1998 | ✓ | ✓ |
| **3600** | Ayas - Champoluc | 401233.4588 | 5076701.514 | 45.8368 | 7.72821 | 1566 | ✓ | ✓ |
| **4000** | Aosta - Piazza Plouves | 369588.3878 | 5066182.332 | 45.7369 | 7.32372 | 580 | ✓ | ✓ |
| **4020** | Aosta - Mont-Fleury | 367652.4425 | 5065511.842 | 45.7305 | 7.29903 | 577 | ✓ | ✓ |
| **4050** | La Thuile - Les Granges | 341803.6811 | 5066026.281 | 45.7297 | 6.96684 | 1637 | ✓ | ✓ |
| **4070** | Cogne - Gimillan | 371803.7254 | 5052665.346 | 45.6157 | 7.35575 | 1785 | ✓ | ✓ |
| **4080** | Entroubles - Chavrière | 362709.5 | 5075200.503 | 45.8167 | 7.23278 | 1339 | ✓ | ✓ |
| **4090** | Saint-Christophe - Aeroporto | 372683.9685 | 5066384.979 | 45.7393 | 7.36344 | 545 | ✓ | ✓ |
| **4100** | Saint-Marcel - Surpian | 378992.7089 | 5065959.005 | 45.7366 | 7.44461 | 540 | ✓ | ✓ |
| **4110** | Donnas - Clapey | 403790.2061 | 5049968.562 | 45.5966 | 7.76643 | 318 | ✓ | ✓ |
| **1** | Traversella | 399115.637 | 5042455.607 | 45.52833 | 7.70806 | 1165 | ✓ | ✓ |
| **2** | Cavallaria | 406908.8571 | 5041342.718 | 45.5194 | 7.80805 | 1353 | ✓ | ✓ |
| **3** | Borgofranco d'Ivrea | 410003.4203 | 5040716.385 | 45.51417 | 7.84778 | 337 | ✓ | ✓ |
| **5** | Meugliano | 404765.3087 | 5036295.509 | 45.47369 | 7.78159 | 634 | ✓ | ✓ |
| **6** | Parella Chiusella | 405963.4733 | 5030311.393 | 45.42 | 7.79806 | 260 | ✓ | ✓ |
| **7** | Piverone Lago | 424406.3376 | 5031236.151 | 45.43056 | 8.03361 | 230 | ✓ | ✓ |
| **9** | Candia Lago | 413679.9079 | 5019212.632 | 45.32111 | 7.89861 | 226 | ✓ | ✓ |
| **11** | Verolengo | 422289.5827 | 5003977.775 | 45.185 | 8.01083 | 163 | ✓ | ✓ |

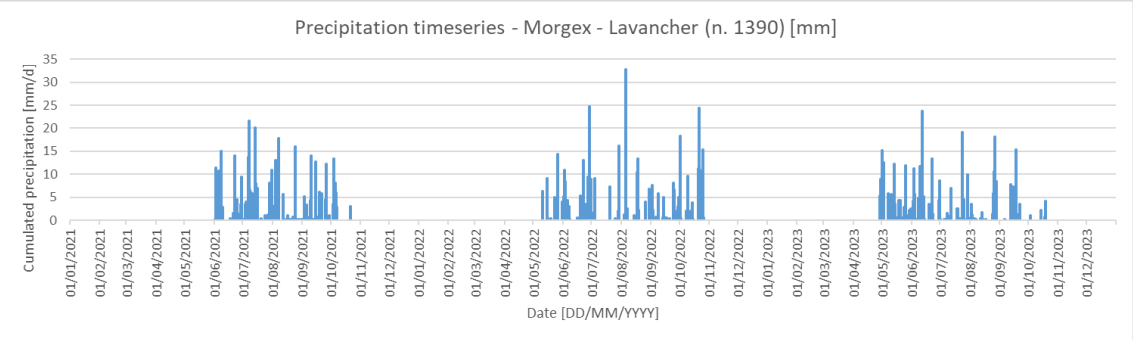



*Figure A1.0.3, precipitation timeseries: station n. 1390 – Morgex – Lavancher*

```
parameter{

 "vreader_station.file"                    "${home}/data/Dora Baltea/temperature_stations.shp"   // .shp meteostations

 "gParamEvaluator.fStationsid"              "ID"                // .shp id field
 "gParamEvaluator.fStationsZ"            "Elevation"             // .shp elevation field
 "gParamEvaluator.tStart"                "1996-01-01 12:00"   // start date
 "gParamEvaluator.tTimeStep"              "1440"                // timeseries timestep

 "gParamEvaluator.doDetrended"              "True"    // detrended krigind with elevation
 "gParamEvaluator.doIncludeZero"            "True"    // include zero value
 "gParamEvaluator.doLogarithmic"            "False"    // use logarithmic of measures (only for precipitation)

 "gParamEvaluator.pSemivariogramType"       "exponential"   // type of empirical semivariogram
 "gParamEvaluator.cutoffDivide"              "20"           // number of bins

 "gParamEvaluator.inHValuesPath"            "${home}/data/Dora Baltea/Temp_formatted.csv"   // .csv temp all stations

}
```

*Figure A1.0.4, GEOframe Java code for temperature variogram parameter estimation*

*Table A1.2, temperature variogram parameters*

| Sill [°C] | Range [m] | Nugget [°C] | Sill with trend [°C] | Range with trend [m] | Nugget with trend [°C] |
|---|---|---|---|---|---|
| 17.15 | 9424 | 0.0 | 4072843 | 294778506408 | 1.27 |

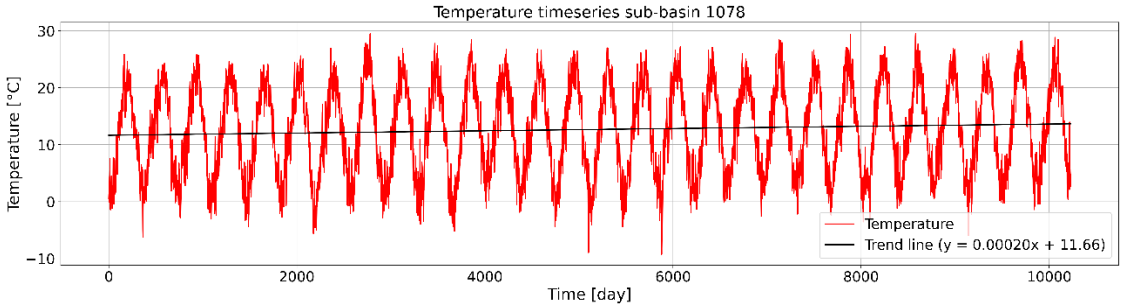



*Figure A1.5, temperature timeseries with trend line of the simulated data at sub-basin 1078*

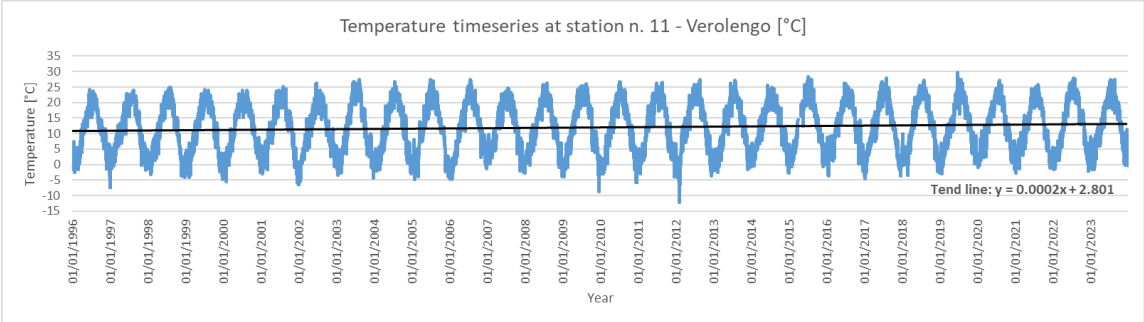



*Figure A1.6, temperature timeseries with trend line of the measured data at station n.11*

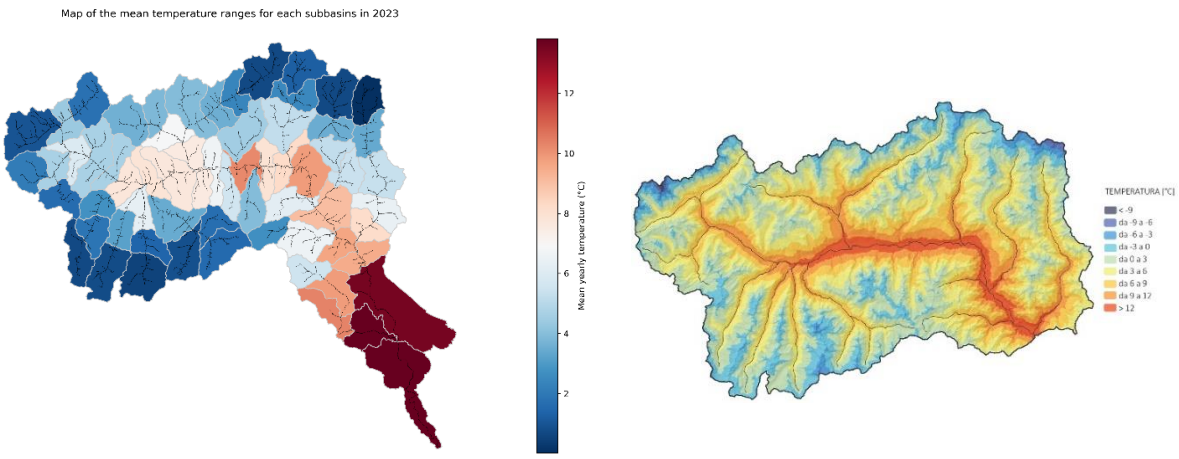



*Figure A1.7, comparison between the interpolated values and the reference data of "Centro Funzionale Regione Autonoma Valle d'Aosta" [24] of the mean temperatures in 2023*

*Table A1.4, precipitation variogram parameters*

| Sill [°C] | Range [m] | Nugget [°C] | Sill with trend [°C] | Range with trend [m] | Nugget with trend [°C] |
|---|---|---|---|---|---|
| 344945764 | 506839557480 | 5.5 | 0.39 | 29262 | 0.1 |

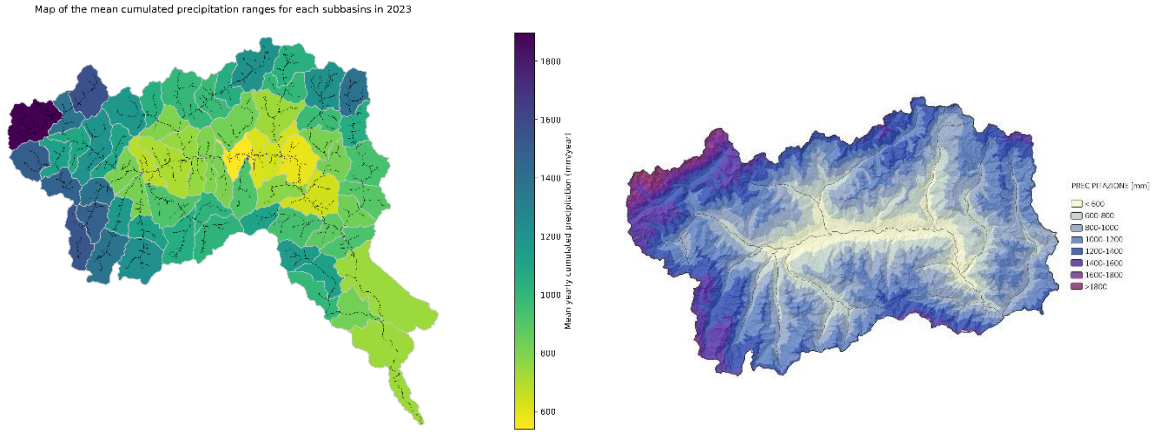



*Figure A1.8, comparison between the interpolated values and the reference data of "Centro Funzionale Regione Autonoma Valle d'Aosta" [26] of the mean cumulated precipitation in 2023*

```
parameter{

    // parameters reader data input air temperature
    "reader_data_airT.file"             "${home}/data/${basin}/${ID}/temp_Krig_${ID}.csv"
    "reader_data_airT.idfield"          "ID"
    "reader_data_airT.tStart"           "${startDate}"
    "reader_data_airT.tEnd"             "${endDate}"
    "reader_data_airT.tTimestep"        "${timestep}"
    "reader_data_airT.fileNovalue"      "-9999"



    "reader_dem.file"                   "${home}/data/${basin}/${ID}/dtm_${ID}.asc"
    "reader_sky.file"                   "${home}/data/${basin}/${ID}/sky_${ID}.asc"
    "reader_centroids.file"            "${home}/data/${basin}/${ID}/centroid_ID_${ID}.shp"

    // parameters of the lwrb  component
    // Idso 1981
    "lwrb.fStationsid"      "${IDsubbasin}"
    "lwrb.X"                0.7
    "lwrb.Y"                5.95
    "lwrb.model"            "6"
    "lwrb.epsilonS"         0.98
    "lwrb.A_Cloud"          0
    "lwrb.B_Cloud"          1

    // parameters of the swrb  component
    "swrb.fStationsid"      "${IDsubbasin}"
    "swrb.tStartDate"       "${startDate}"
    "swrb.doHourly"         "${swrbDoHourly}"   //or false if daily
    "swrb.pCm03"            0.6
    "swrb.pAlphag"          0.9
    "swrb.pVisibility"      80

    // parameters of the netrad  component
    "net.alfa"            0.26
```

*Figure A1.9, parameter section for each component of the radiation*

*Table A1.5, comparison of the simulated and measured total shortwave radiation (sub-basin 414 and station 4050)*

| | **Station** | **Sub-basin** |
|---|---|---|
| **ID** | 4050 – La Thuile – les Granges | 414 |
| **Elevation (m.a.s.l.)** | 1637 | 1665 |
| **Average Total SW radiation ($W/m^2$) [2022-2023]** | 536.83 | 736.06 |
| **Total SW radiation difference sub-basin - station ($W/m^2$)** | 199.23 | |

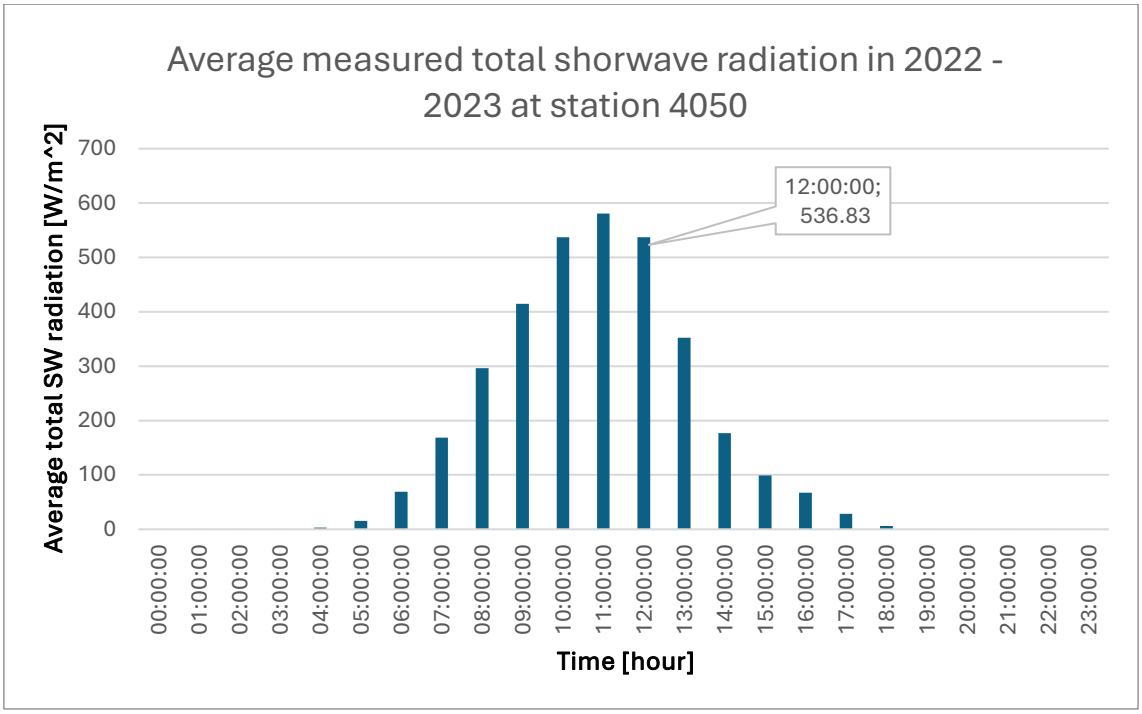



*Figure A1.10, average measured total shortwave radiation in 2022 and 2023 at station 4050*

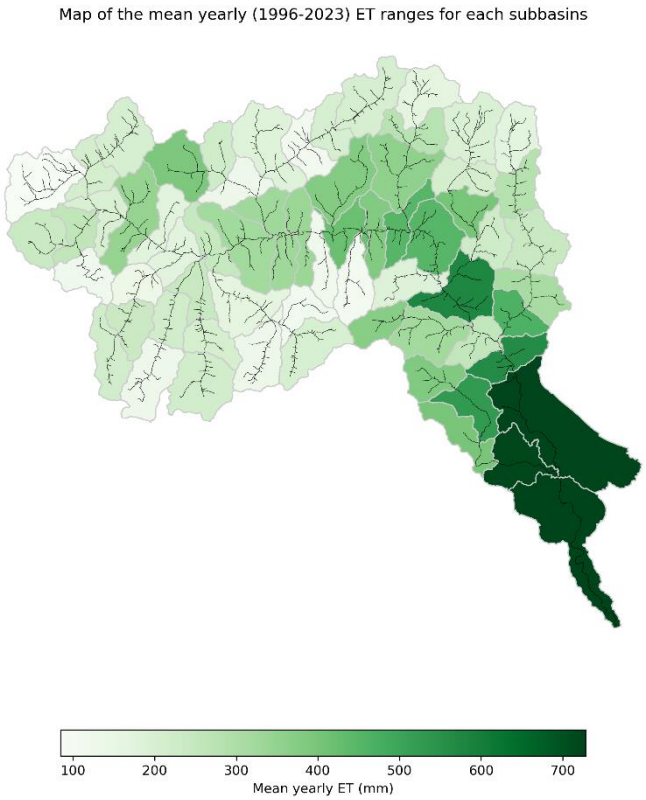



*Figure A1.11, first obtained map of the mean yearly evapotranspiration ranges for each sub-basin.*

# A2 Equations

For the rain snow separator:

$$P_r = \alpha_r \left[ P * arctan(T - T_m) + \frac{P}{2} \right]$$

*(A2. 6)*

$$P_s = \alpha_s (P - P_r)$$

*(A2. 2)*

where:

- $P_r$: rainfall component [$L^1$ $T^{-1}$],
- $P_s$: snow component [$L^1$ $T^{-1}$],
- $\alpha_r$: rainfall measurements errors coefficient [-],
- $T_m$: melting temperature [$\Theta^1$],
- T: temperature,
- P: precipitation,
- $\alpha_s$: snow measurements errors coefficient [-].

For the snow component:

$$\frac{dS_r}{dt} = P_r + m - f - M_d$$

with:

$$m_{degreeday} = \alpha_m (T - T_m)$$

$$f = \alpha_f (T_m - T)$$

$$M_d = (S_r - \alpha_l S_S)$$

*(A2. 3)*

$$\frac{dS_s}{dt} = P_s + f - m$$

*(A2. 4)*

where:

- $\alpha_m$: melting factor [$L^1$ $\Theta^{-1}$ $T^{-1}$],
- $\alpha_f$: freezing factor [$L^1$ $\Theta^{-1}$ $T^{-1}$],
- $\alpha_l$: Water retention capacity coefficient [-],
- $S_r$: rainfall storage [$L^1$],
- $S_s$: snow storage [$L^1$],
- $M_d$: melting discharge [$L^1$ $T^{-1}$].

For the canopy component:

$$\frac{dS_c}{dt} = in_c - aet_c - \ out_c$$

with:

$$in_c = (1-p)M_d$$

$$aet_c = ET_p \min\left(1, \frac{S_c}{k_c LAI}\right)$$

$$T_r \ = pM_d + out_c$$

*(A2. 5)*

where:

- p: free throughfall coefficient [-],
- $k_c$: canopy storage coefficient [-],
- LAI: leaf area index [$L^1$],
- $ET_p$: potential evapotranspiration [$L^1\ T^{-1}$],
- $in_c$: canopy reservoir inflow [$L^1\ T^{-1}$],
- $out_c$: canopy reservoir outflow [$L^1\ T^{-1}$],
- $S_c$: canopy storage [$L^1$],
- $T_r$: througfall [$L^1\ T^{-1}$],
- $aet_c$: actual canopy evapotranspiration [$L^1\ T^{-1}$].

For the rootzone component:

$$\frac{dS_{rz}}{dt} = in_{rz} - \ aet_{rz} - \ R_e$$

with:

$$in_{rz} = \big(1 - \alpha(S_{rz}, p_B)\big)T_r$$

$$in_{ru} = \alpha(S_{rz}, p_B)T_r$$

$$aet_{rz} = \big(ET_p - aet_c\big) \min\left(1, \frac{4}{3}\frac{S_{rz}}{S_{rz,max}}\right)$$

$$Re = g\left(\frac{S_{rz}}{S_{rz,max}}\right)^h$$

*(A2. 6)*

where:

- α: splitting parameter [-],

- g: groundwater recharge coefficient [$L^1T^{-1}$],
- h: groundwater recharge exponent [-],
- $p_B$: spatial variability of the soil moisture capacity [-],
- $ET_P$: potential evapotranspiration [$L^1$ $T^{-1}$],
- $S_{rz,\ max}$: rootzone maximum storage [$L^1$],
- $in_{rz}$: rootzone reservoir inflow [$L^1$ $T^{-1}$],
- $S_{rz}$: rootzone storage [$L^1$],
- $in_{ru}$: runoff reservoir inflow [$L^1$ $T^{-1}$],
- Re: groundwater recharge [$L^1$ $T^{-1}$],
- $aet_{rz}$: actual rootzone evapotranspiration [$L^1$ $T^{-1}$].

For the runoff component:

$$\frac{dS_{ru}}{dt} = in_{ru} - Q_{ru}$$

with:

$$Q_{ru} = c\left(\frac{S_{ru}}{S_{ru,max}}\right)^d \qquad (A2.\ 7)$$

where:

- c: runoff discharge coefficient [$L^1$ $T^{-1}$],
- d: runoff discharge coefficient [-],
- $S_{ru,\ max}$: runoff maximum storage [$L^1$],
- $S_{ru}$: runoff storage [$L^1$],
- $Q_{ru}$: runoff discharge [$L^1$ $T^{-1}$].

For the groundwater component:

$$\frac{dS_{gw}}{dt} = Re - Q_{gw}$$

with:

$$Q_{gw} = e\left(\frac{S_{gw}}{S_{gw,max}}\right)^f \qquad (A2.\ 8)$$

where:

- e: deep groundwater discharge coefficient [$L^1$ $T^{-1}$],
- f: deep groundwater discharge exponent [-],
- $S_{ru,\ max}$: groundwater maximum storage [$L^1$],
- $S_{gw}$: groundwater storage [$L^1$],
- $Q_{gw}$: deep groundwater discharge [$L^1$ $T^{-1}$].

# APPENDIX B

## B1 Figures, tables, and codes

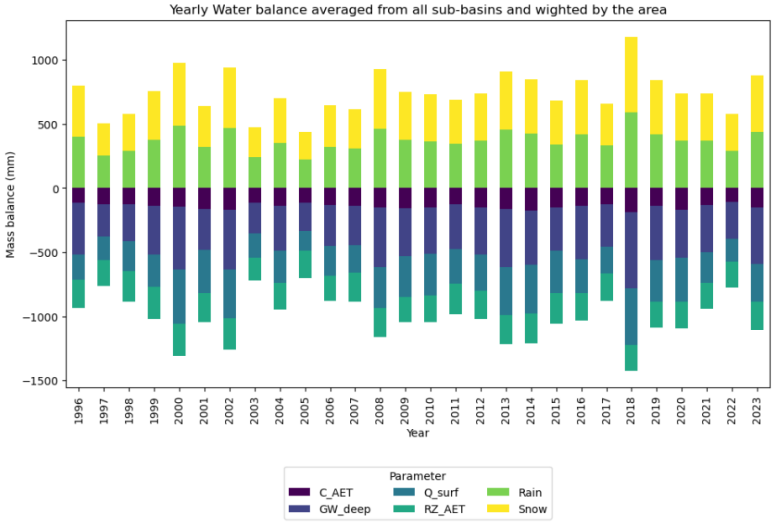



*Figure B1.1, yearly water balance average from all sub-basins and weighted by the area before calibration.*

*Table B1.1, sub-basin IDs and correspondent calibration gauging station*

| Sub-basin ID | Gauging station used for calibration | Sub-basin ID | Gauging station used for calibration | Sub-basin ID | Gauging station used for calibration |
|---|---|---|---|---|---|
| **22** | 3 | **457** | 3 | **737** | 1430 |
| **82** | 3 | **461** | 1430 | **767** | 3 |
| **99** | 1640 | **465** | 3 | **788** | 3 |
| **148** | 1640 | **468** | 3 | **807** | 3 |
| **168** | 3 | **469** | 3 | **818** | 1430 |
| **217** | 1640 | **471** | 1430 | **821** | 1130 |
| **220** | 1430 | **478** | 1430 | **848** | 1130 |
| **226** | 3 | **481** | 1430 | **852** | 3 |
| **249** | 1290 | **492** | 1430 | **853** | 3 |
| **258** | 3 | **519** | 1490 | **897** | 1130 |
| **267** | 1430 | **529** | 1430 | **905** | 1430 |
| **274** | 1430 | **550** | 1430 | **919** | 3 |
| **276** | 1430 | **557** | 3 | **936** | 1430 |
| **281** | 1430 | **562** | 1430 | **962** | 1430 |
| **309** | 1430 | **583** | 1490 | **999** | 3 |
| **347** | 1430 | **599** | 1430 | **1060** | 3 |
| **373** | 3 | **624** | 3 | **1064** | 3 |
| **393** | 1290 | **628** | 3 | **1073** | 3 |
| **404** | 3 | **636** | 1430 | **1077** | 3 |
| **406** | 1430 | **647** | 1430 | **1078** | 3 |
| **414** | 1490 | **681** | 3 | | |
| **436** | 3 | **704** | 1430 | | |

*Table B1.2, upstream calibrated parameter values*

| Parameter Symbol | Parameter Name | Unit | Cal. parameters at station 1490 | Cal. parameters at station 1640 | Cal. Parameters at station 1130 | Cal. parameters at station 1290 |
|---|---|---|---|---|---|---|
| $\alpha_r$ | Rainfall measurements errors coefficient | [-] | 0.727 | 0.275 | 0.446 | 0.571 |
| $\alpha_s$ | Snow measurements errors coefficient | [-] | 1.65 | 0.522 | 1.26 | 1.44 |
| $T_M$ | Melting temperature | [°C] | -0.453 | -0.332 | -0.811 | -0.256 |
| $\alpha_m$ | Combined melting factor | [mm/°/d] | 1.63 | 1.54 | 1.25 | 1.40 |
| $\alpha_f$ | Freezing factor | [mm/°/d] | 1.09 | 0.227 | 0.16 | 0.453 |
| $\alpha_l$ | Water retention capacity coefficient | [-] | 0.0284 | 0.155 | 0.0320 | 0.0608 |
| $k_c$ | Canopy storage coefficient | [-] | 0.294 | 0.311 | 0.267 | 0.202 |
| p | Free throughfall coefficient | [-] | 0.881 | 0.900 | 0.760 | 0.843 |
| $S_{RZ,max}$ | Rootzone maximum storage | [mm] | 198 | 68.4 | 63.3 | 47.9 |
| g | Groundwater recharge coefficient | [mm/d] | 0.00410 | 1.59 | 0.988 | 1.01 |
| h | Groundwater recharge exponent | [-] | 1.57 | 1.45 | 1.47 | 1.12 |
| $p_B$ | Spatial variability of the soil moisture capacity | [-] | 7.20 | 2.36 | 6.98 | 6.08 |
| $S_{RU,max}$ | Runoff maximum storage | [mm] | 42.3 | 76.8 | 11.0 | 44.4 |
| **c** | Runoff discharge coefficient | [mm/d] | 0.855 | 1.93 | 1.21 | 1.74 |
| d | Runoff discharge exponent | [-] | 1.57 | 1.53 | 1.02 | 1.59 |
| $S_{GW,max}$ | Groundwater maximum storage | [mm] | 96.8 | 153 | 56.8 | 50.3 |
| **e** | Deep groundwater discharge coefficient | [mm/d] | 0.0618 | 1.21 | 0.111 | 1.78 |
| f | Deep groundwater discharge exponent | [-] | 1.66 | 1.57 | 1.55 | 1.91 |

*Table B1.3, downstream calibrated parameter values*

| Parameter Symbol | Parameter Name | Unit | Cal. parameters at station 1430 | Cal. parameters at station 3 |
|---|---|---|---|---|
| $\alpha_r$ | Rainfall measurements errors coefficient | [-] | 0.490 | 0.569 |
| $\alpha_s$ | Snow measurements errors coefficient | [-] | 0.446 | 0.983 |
| $T_M$ | Melting temperature | [°C] | 0.584 | 0.671 |
| $\alpha_m$ | Combined melting factor | [mm/°/d] | 1.31 | 2.60 |
| $\alpha_f$ | Freezing factor | [mm/°/d] | 0.554 | 0.410 |
| $\alpha_l$ | Water retention capacity coefficient | [-] | 0.403 | 0.414 |
| $k_c$ | Canopy storage coefficient | [-] | 0.108 | 0.301 |
| p | Free throughfall coefficient | [-] | 0.736 | 0.871 |
| $S_{RZ,max}$ | Rootzone maximum storage | [mm] | 126 | 91.9 |
| g | Groundwater recharge coefficient | [mm/d] | 0.223 | 1.10 |
| h | Groundwater recharge exponent | [-] | 1.81 | 1.47 |
| $p_B$ | Spatial variability of the soil moisture capacity | [-] | 1.40 | 4.52 |
| $S_{RU,max}$ | Runoff maximum storage | [mm] | 29.3 | 74.4 |
| **c** | Runoff discharge coefficient | [mm/d] | 1.63 | 1.98 |
| d | Runoff discharge exponent | [-] | 1.02 | 1.69 |
| $S_{GW,max}$ | Groundwater maximum storage | [mm] | 47.9 | 314 |
| **e** | Deep groundwater discharge coefficient | [mm/d] | 0.375 | 1.17 |
| f | Deep groundwater discharge exponent | [-] | 1.36 | 1.87 |

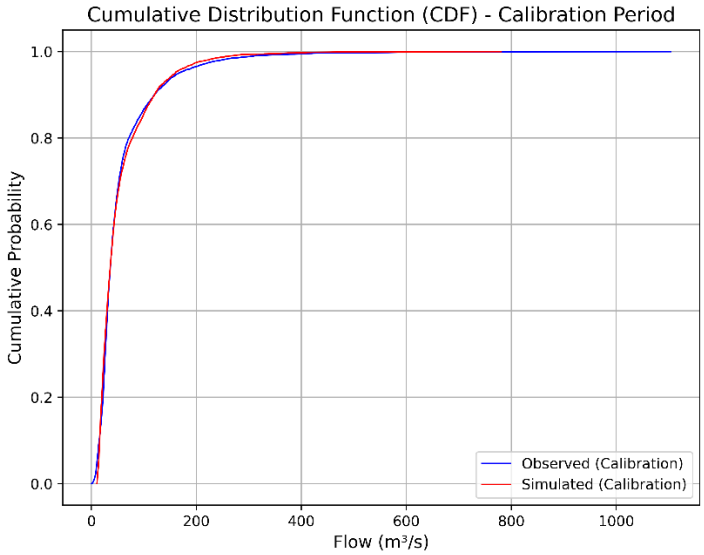



*Figure B1.2, cumulative distribution function (CDF) of observed and simulated discharges at the basin outlet not cut, calibration period 2007-2023*

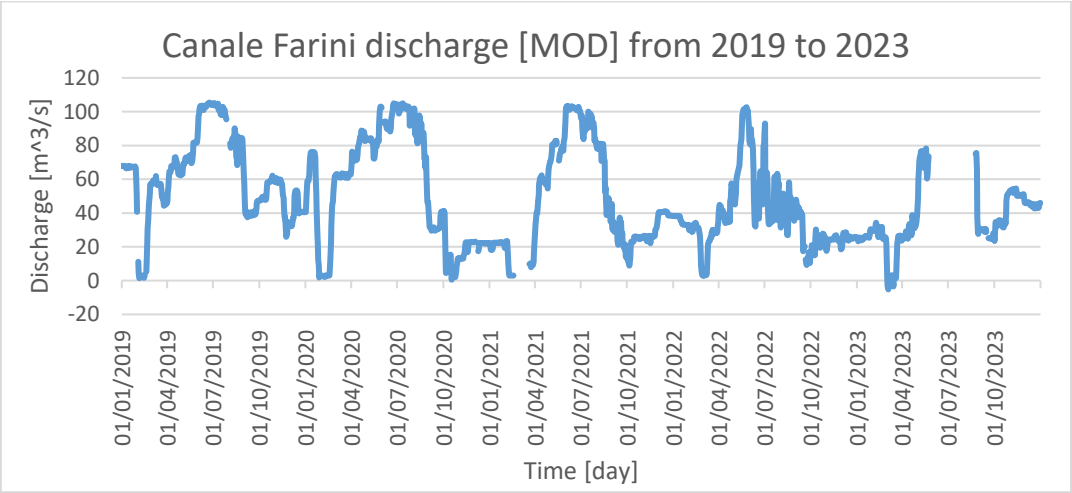



*Figure B1.3, Canale Farini discharge from 2019 to 2023*

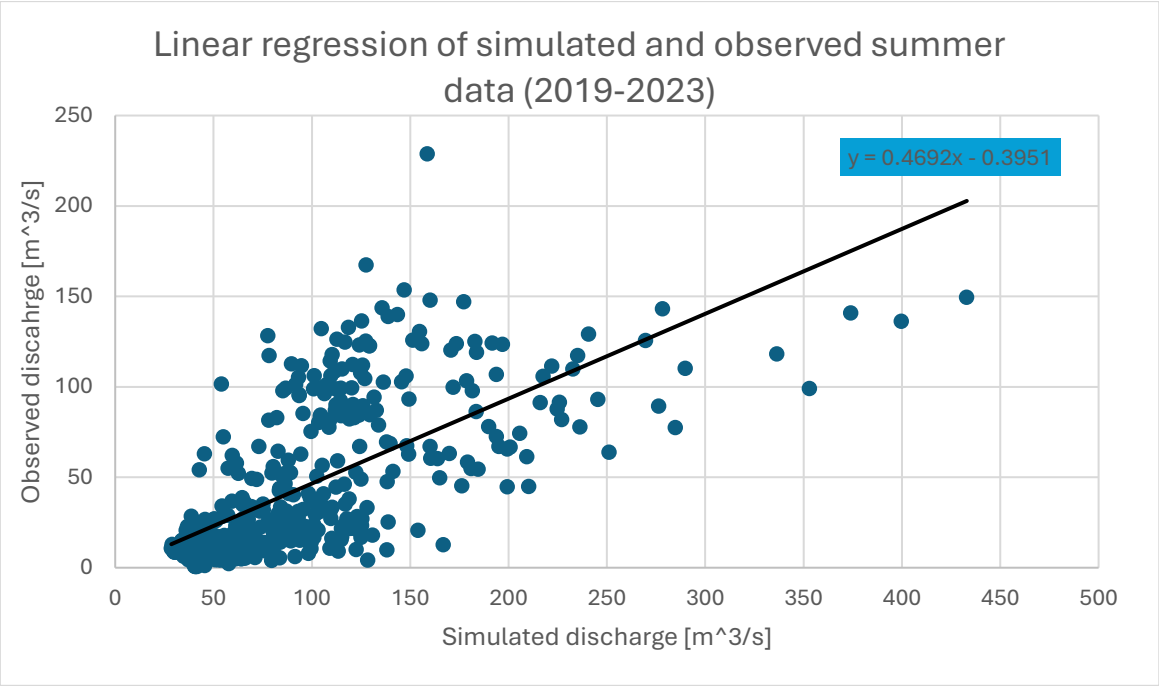
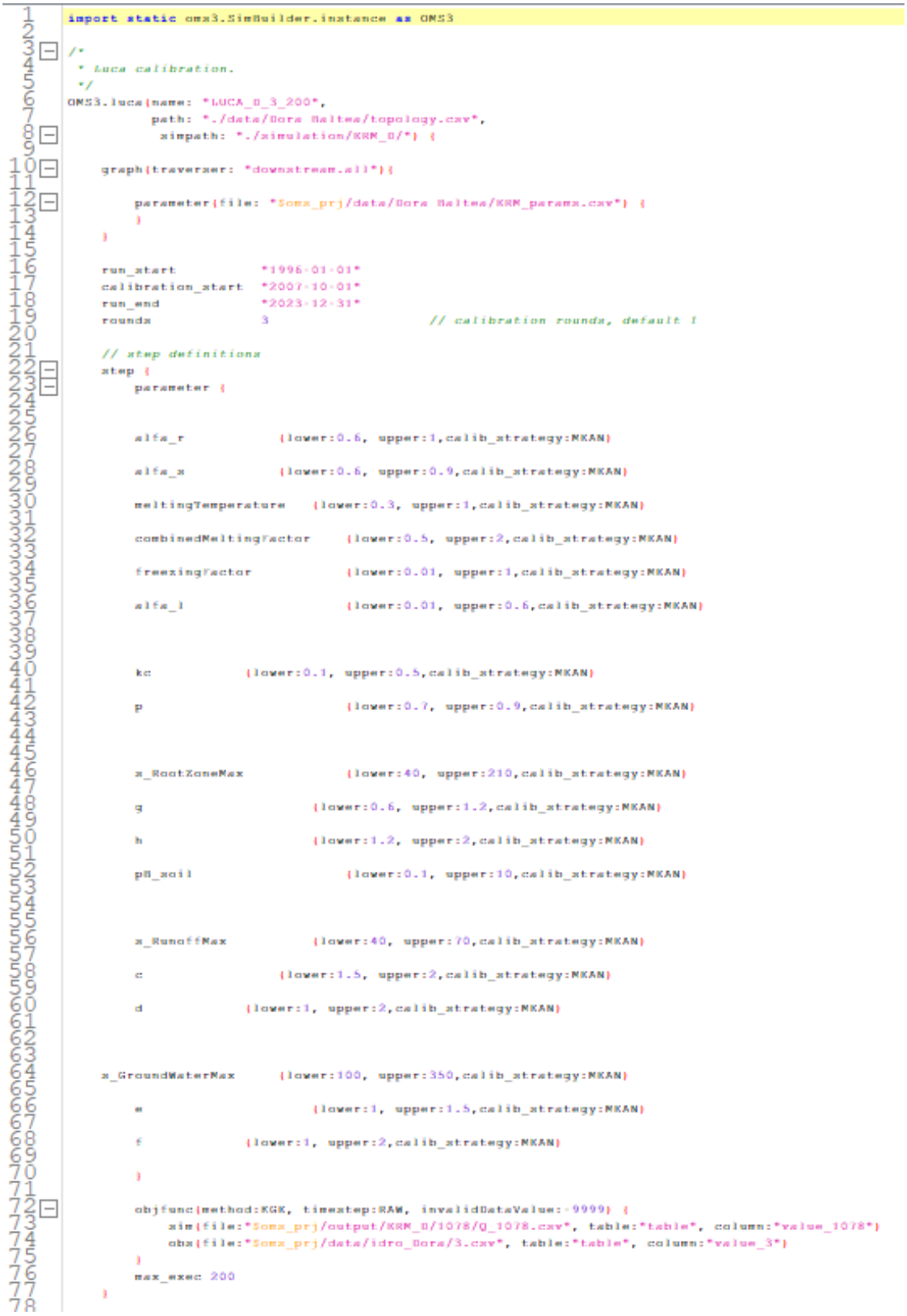



*Figure B1.4, linear regression of the simulated and observed summer data (2019-2023) with the trend line equation*



```
import static oms3.SimBuilder.instance as OMS3

/*
 * Luca calibration.
 */
OMS3.luca(name: "LUCA_0_3_200",
          path: "./data/Dora Baltea/topology.csv",
           ximpath: "./ximulation/KRM_0/") {

    graph(traverser: "downstream.all"){

        parameter(file: "$oms_prj/data/Dora Baltea/KRM_params.csv") {
        }
    }

    run_start           "1996-01-01"
    calibration_start   "2007-10-01"
    run_end             "2023-12-31"
    rounds              3                   // calibration rounds, default 1

    // step definitions
    step {
        parameter {


        alfa_r            [lower:0.6, upper:1,calib_strategy:MKAN]

        alfa_s            [lower:0.6, upper:0.9,calib_strategy:MKAN]

        meltingTemperature    [lower:0.3, upper:1,calib_strategy:MKAN]

        combinedMeltingFactor     [lower:0.5, upper:2,calib_strategy:MKAN]

        freezingFactor            [lower:0.01, upper:1,calib_strategy:MKAN]

        alfa_l                    [lower:0.01, upper:0.6,calib_strategy:MKAN]



        kc            [lower:0.1, upper:0.5,calib_strategy:MKAN]

        p                         [lower:0.7, upper:0.9,calib_strategy:MKAN]



        x_RootZoneMax             [lower:40, upper:210,calib_strategy:MKAN]

        g                     [lower:0.6, upper:1.2,calib_strategy:MKAN]

        h                     [lower:1.2, upper:2,calib_strategy:MKAN]

        p8_soil                   [lower:0.1, upper:10,calib_strategy:MKAN]



        x_RunoffMax           [lower:40, upper:70,calib_strategy:MKAN]

        c                 [lower:1.5, upper:2,calib_strategy:MKAN]

        d             [lower:1, upper:2,calib_strategy:MKAN]



    x_GroundWaterMax      [lower:100, upper:350,calib_strategy:MKAN]

        e                     [lower:1, upper:1.5,calib_strategy:MKAN]

        f             [lower:1, upper:2,calib_strategy:MKAN]

        }

        objfunc(method:KGK, timestep:RAW, invalidDataValue:-9999) {
            sim(file:"$oms_prj/output/KRM_0/1078/Q_1078.csv", table:"table", column:"value_1078")
            obs(file:"$oms_prj/data/idro_Dora/3.csv", table:"table", column:"value_3")
        }
        max_exec 200
    }
```

*Figure B1.5, LUCA calibrator example code structure*

```
import static oms3.SimBuilder.instance as OMS3
def home = oms_prj
def ID=22
def etmodel="FAO"       // ET formula
def qmodel="KRM"    // sim file is either ERM or Dam
def area= 63.00   // Area [km2]


OMS3.sim (name: "22", calibrate: "true") {

resource "$oms_prj/lib"

 model(while:"reader_data_precip.doProcess") {
 //model() {
   components {
         "inNode1077_Q"                "ex0.in1"

        // ------------------------ Classes ------------------------
    "sumDischarge"                  "it.geoframe.blogspot.net3utils.HashMapSum"

    "reader_data_precip"            "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorReader"
    "reader_data_temp"              "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorReader"
    "reader_data_KT"                "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorReader"

    "rainsnowsep"                   "rainsnowSeparator.RainSnowSep"
    "snow"                          "snowReservoir.WaterBudgetSnow"
    "canopy"                        "canopyReservoir.WaterBudgetCanopy"
    "rootzone"                      "rootzoneReservoir.WaterBudgetRootzone"
    "runoff"                        "runoffReservoir.WaterBudgetRunoff"
    "groundwater"                   "groundwaterReservoir.WaterBudgetGroundwater"



    "writer_rainsnowsep_rain"       "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_rainsnowsep_snow"       "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_snow_SWE"               "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_snow_meltingdischarge"  "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_snow_melting"           "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_snow_freezing"          "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_snow_error"             "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_canopy_throughfall"     "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_canopy_aet"             "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_canopy_storage"         "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_canopy_error"           "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_rootzone_aet"           "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_rootzone_storage"       "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_rootzone_recharge"      "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_rootzone_quick"         "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_rootzone_error"         "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_runoff_storage"         "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_runoff_Q"               "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_runoff_Q_mm"            "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_runoff_error"           "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_groundwater_storage"    "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_groundwater_Q"          "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_groundwater_Q_mm"       "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"
    "writer_groundwater_error"      "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "writer_Discharge"              "org.hortonmachine.gears.io.timedependent.OmsTimeSeriesIteratorWriter"

    "outNodeMask"                   "ex0.OutMask"


   }
```

*Figure B1.6, first 68 lines of the ERM code of one sub-basin*

***Timeseries result from the first calibration attempt***

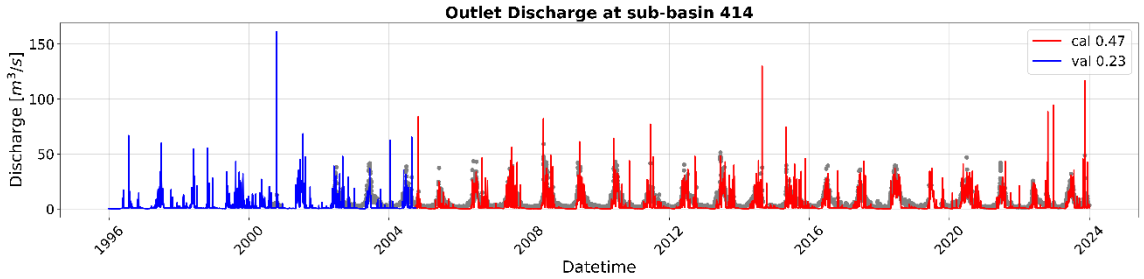



*Figure B1.7, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 1490*

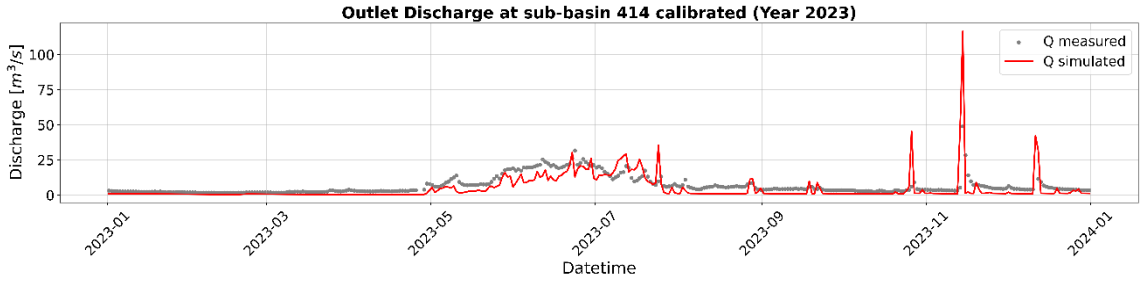



*Figure B1.8, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 1490*

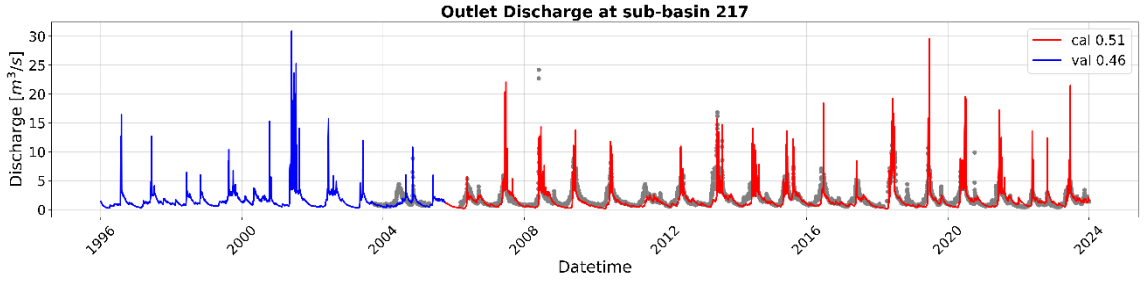



*Figure B1.9, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 1640*

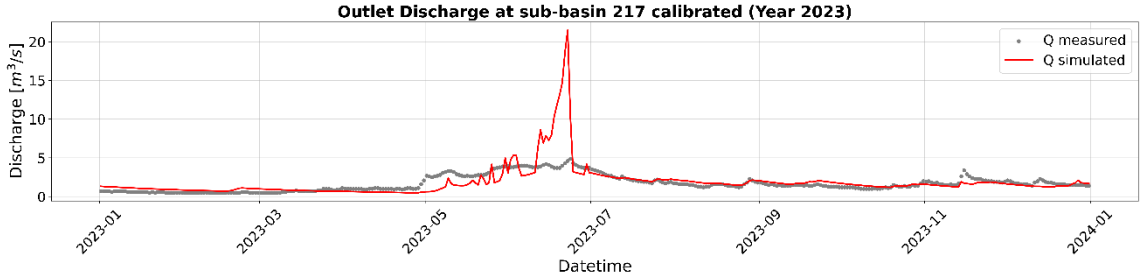



*Figure B1.10, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 1640*

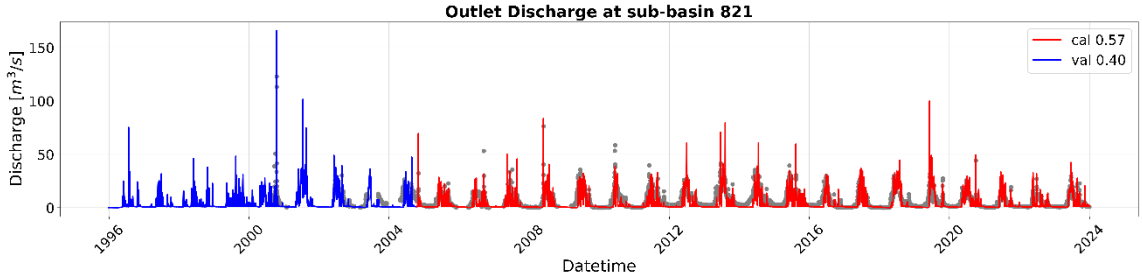



*Figure B1.11, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 1130*

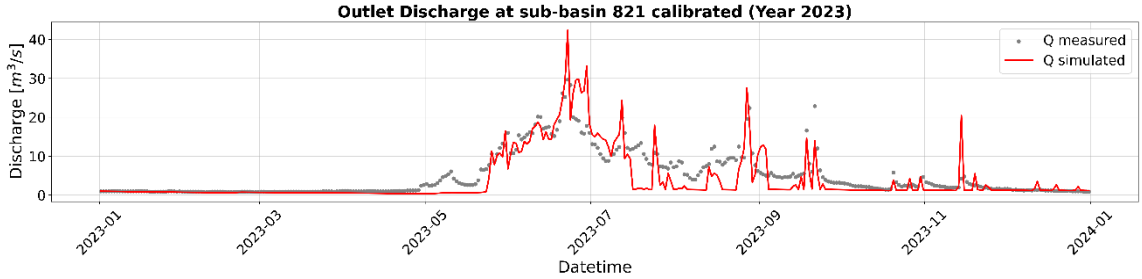



*Figure B1.12, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 1130*

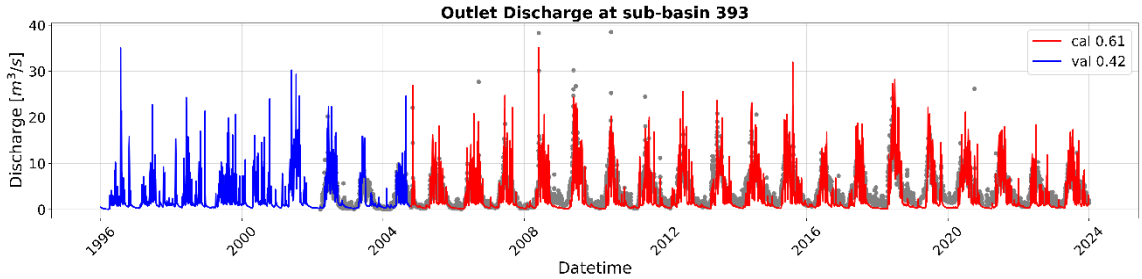



*Figure B1.13, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 1290*

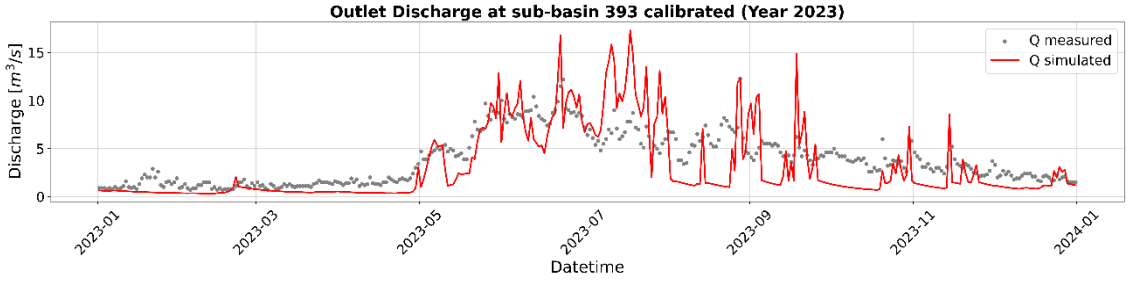



*Figure B1.14, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 1290*

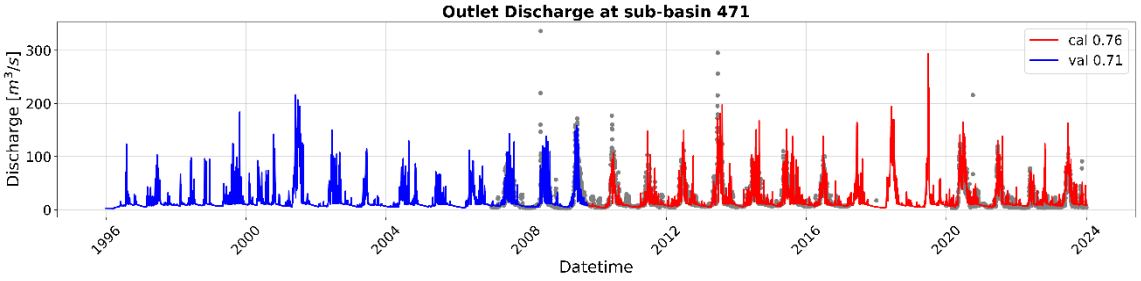



*Figure B1.15, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 1430*

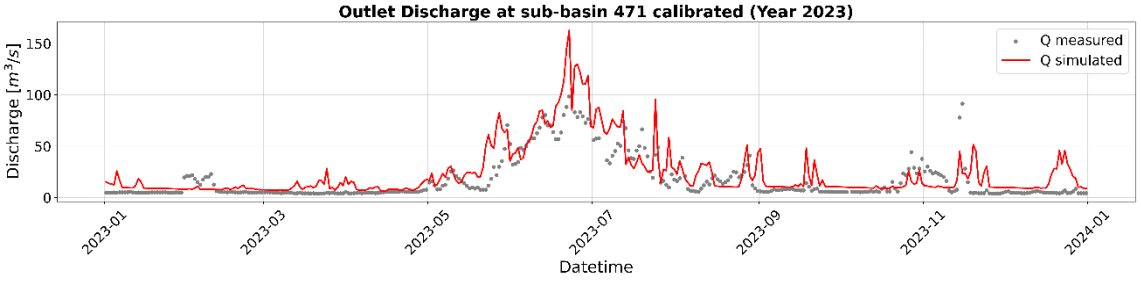



*Figure B1.16, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 1430*

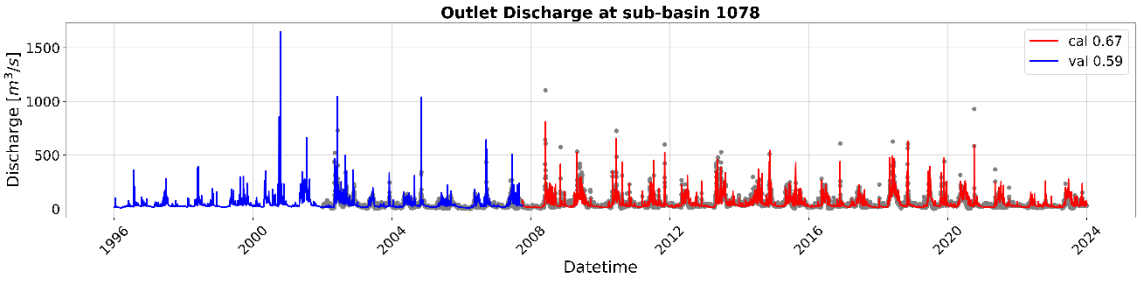



*Figure B1.17, comparison of observed and simulated discharge timeseries between 1996 and 2023, with calibration (first attempt) at station 3*

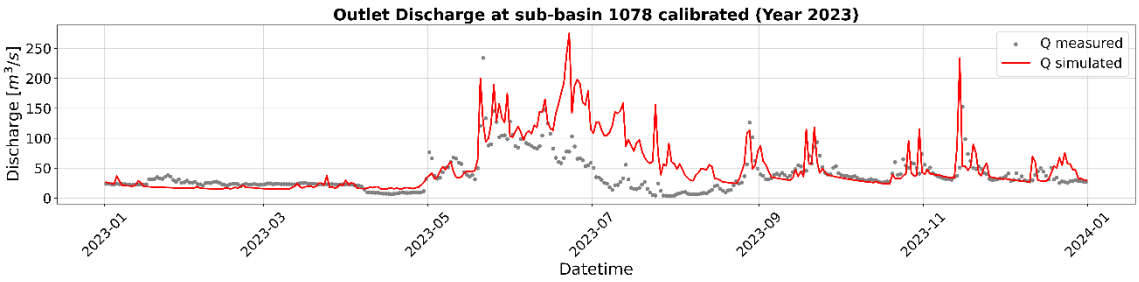



*Figure B1.18, comparison of observed and simulated discharge timeseries in 2023, with calibration (first attempt) at station 3*

## B2 Error Metrics Adopted

- Kling-Gupta efficiency (KGE): evaluates the adherence between the observed and simulated values.

$$KGE = 1 - \sqrt{(r-1)^2 + (a-1)^2 + (b-1)^2}$$

*(C1)*

where:

  - r is the correlation coefficient between the observed and simulated values,
  - a is the variability error, $a=\sigma_S / \sigma_m$ with $\sigma_S$ and $\sigma_m$ as standard deviations,
  - b is the bias error, $b =\mu_S/\mu_M$ with $\mu_S$ and $\mu_M$ as the mean values of measured and simulated data.

KGE=1 indicated the maximum agreement between predicted and observed values.
It is used as the Objective Function during the calibration procedure.

- Pearson Correlation measures the linear correlation between the simulated and observed values. It is defined as:

$$r = \frac{\sum(M_i - \overline{M_\iota})(S_i - \overline{S_\iota})}{\sqrt{\sum(M_i - \overline{M_\iota})^2 \sum(S_i - \overline{S_\iota})^2}}$$

*(C2)*

where: $S_i$ and $M_i$ are the simulated and measured timeseries and $\overline{S_\iota}$ and $\overline{M_\iota}$ are the mean of the simulated and measured timeseries.
r = 1 means linear correlation, r = -1 means inverse linear correlation and r = 0 no linear correlation between data.

- Mean Absolute Error (MAE) is the average of absolute errors telling how much forecasts deviate on average from the observed values. It is defined as:

$$MAE = \frac{1}{N}\sum_{i=1}^{n} |M_i - S_i|$$

*(C3)*

where, $S_i$ and $M_i$ are the simulated and measured timeseries and N is the number of components in the series.
The lower the value is the higher is the goodness of forecast.

- Root-Mean-Square Error (RMSE) provides a measure of the average error between the observed and simulated values.is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - S_i)^2}$$

*(C3)*

where, $S_i$ and $M_i$ are the simulated and measured timeseries and N is the number of components in the series.

- Percent Bias (% bias): measures the average trend of the model to overestimate (positive percentage) or underestimate (negative percentage) the values observed. It is defined as:

$$\%\ bias = 100 * \frac{\sum_{i=1}^{n}(S_i - M_i)}{\sum_{i=1}^{n}(M_i)} \quad (C4)$$

where, $S_i$ and $M_i$ are the simulated and measured timeseries and N is the number of components in the series.

# ACKNOWLEDGEMENTS

Mi è doveroso dedicare questo ultimo spazio dell'elaborato per ringraziare tutti coloro che hanno contribuito alla sua realizzazione e mi hanno supportato durante il mio intero percorso accademico.

Ringrazio la mia relatrice, Prof. Ing. Stefania Tamea, per la possibilità data di realizzare lo studio riportato su una tematica e un luogo a cui tengo. Grazie per il supporto e i consigli ricevuti.

Ringrazio dal profondo del mio cuore mamma, papà e Andrea per essermi stati accanto da sempre e, soprattutto, in questo, ormai volto al termine, percorso universitario. Grazie per tutto il sostegno, l'amore ricevuto e per avermi sempre incoraggiata.

Un ringraziamento speciale va a Matteo che, in questi ultimi due anni, con tanto amore e comprensione, mi ha accompagnato e sostenuto passo dopo passo. Grazie per aver sempre creduto in me anche nei momenti più difficili, per la pazienza e per aver reso tutto quanto più leggero.

A tutti i miei compagni di corso per aver condiviso con me tutte le gioie e dolori di questi ultimi anni. E grazie ai miei amici di sempre che sono stati fonte di sostegno e gioia. Grazie per avermi fatto sorridere anche nei momenti più stressanti.

Infine, un sentito ringraziamento ai i miei nonni, in particolare al mio caro Nonno Gino per essere da sempre stato il mio più grande sostenitore. Spero sempre di rendervi orgogliosi.

Grazie infinite, senza di voi questo percorso non sarebbe stato lo stesso.