



**Politecnico
di Torino**

Master of Science in Land and Environmental Engineering

Master Degree Thesis

**Global hydrological models and
the Great Alpine Region(GAR):
review and performance
evaluation of the ISIMIP model
ensemble**

Supervisors

prof. Alberto Viglione

prof. Jost-Diedrich Graf Von Hardenberg

Dott. Luigi Cafiero

Candidate

Lorenzo PIPAN

ACADEMIC YEAR 2022-2023

Abstract

One of the direct consequences of climate change lies in its impacts on the hydrological cycle: the recurrence of extreme positive and negative anomalies of precipitation is set to become more and more frequent, while the rising of global temperatures is already impacting on freshwater availability, or excess, in many regions of the world. The Great Alpine Region (GAR), i.e. the geographical area comprising the European Alps and their close proximity, is in this context a particularly impacted spot as it is showing high levels of both exposure and vulnerability to climate change. This is especially relevant because of the importance of the Alps for the freshwater resources across the whole continent. In this thesis we perform a thorough data analysis on hydrological data at $0.5^\circ \times 0.5^\circ$ resolution for the ISIMIP-2a protocol's model ensemble over the period 1971-2000 in the GAR. The purpose of this work is to provide insights into the current capability of Global Hydrological Models (GHMs) and their reliability on the selected scale and area. To achieve that, we perform an exploratory analysis of the data providing graphs and maps of different hydrological signatures for the region. Secondly, we compare the models' output data with a large observational dataset provided by the Global Runoff Data Center (GRDC). Lastly, the models' performance is ranked over two different experiments and we set a link between the performance of the models and the morphological properties of the catchments. To this end, daily runoff generated at each pixel of the model's grid is compared to daily runoff data from 141 catchments of the GAR. Three hydrological signatures are selected to perform the comparison: a quantile for the high flows (Q_{95}), one for the low flows (Q_5) and lastly the Mean Monthly Runoff curve. For each signature the modified Kling Gupta Efficiency (KGE') is computed to account for the misfit between the observed and modeled data. The performance on the three signatures allows for a relative ranking of the ensemble participating the project. Results highlight that the ensemble mean tends to outperform each single model in general but has specific downsides related especially to the representation of the monthly regime curve. All models struggle to represent correctly the low flows quantiles while the high flows quantiles are more consistently reproduced by the models. Mean monthly regime curves rank in the middle of the three indicators and show two main important results: one is the consistent negative bias in mean values across all models, which overall considerably underestimate the total volume of runoff produced, secondly the difficulty of the models to correctly reproduce the seasonal timing of both peak and low flows. This indicates the limits of using current state of the art GHMs in the context of a complex region such as the GAR, pointing towards the difficulty

of representing complex mechanisms of snow melt in complex topography.

Contents

Abstract	I
Introduction	2
0.1 General background	2
0.2 Aim of the work and research questions	3
1 Theoretical Background	4
Theoretical Background	4
1.1 Overview of ISIMIP	4
1.1.1 ISIMIP water sector	5
1.1.2 Climate input: Reanalysis datasets	6
1.1.3 Socio Economical Scenarios	9
1.1.4 The GHMs	10
1.2 Background on the GAR	12
1.2.1 Overview of the Alps and climate change	12
1.2.2 GAR, climate modelling and future projections	13
1.2.3 Major Rivers of the GAR	15
2 Data and Methods	19
Data and Methods	19
2.1 ISIMIP data over the GAR	19
2.2 Hydrological signatures over the GAR	22
2.3 GRDC data for the alpine region	26
2.3.1 From discharge to generated runoff	27
2.4 Methods	29
2.4.1 Performance evaluation metrics	29
2.4.2 Software used	32
3 Results	34
Results	34
3.1 Results for selected representative catchments	34

3.1.1	Border of Slovenia and Italy: the River Soca	34
3.1.2	Example of catchment in the high Alps: The spawn of river Rhone at Gletsch	36
3.1.3	Catchment affected by human impacts: Durance at Serre Poncon	37
3.2	Performance Evaluation of "NOSOC" model ensemble	40
3.3	Performance Evaluation of "VARSOC" model ensemble	46
3.3.1	Summary table of results	53
3.3.2	Link between performance and morphology of the catchments	54
3.3.3	Link between performance of specific models on different indicator	56
4	Discussion	59
	Discussion	59
4.1	Discussion of the results	59
4.2	Discussion on the methodology adopted	62
4.3	GAR modeling and climate change	63
4.4	Conclusions	66

List of Figures

1.1	General scheme of ISIMIP, source ISIMIP	4
1.2	The water cycle, credits source USGS	5
1.3	Graph of input data to GHMs	10
1.4	Scheme of modelled processes, source [9]	11
1.5	European alps with orography, credits [8]	14
1.6	GAR regionalization and principal river basins	15
1.7	Danube statistics	16
1.8	Rhine statistics	17
1.9	Po statistics	17
1.10	Rhone statistics	18
2.1	Flood october 2000, generated runoff	20
2.2	Flood october 2000, simulated discharge	21
2.3	Mean annual specific runoff	22
2.4	Pardè Range	23
2.5	Slope of FDC	24
2.6	Normalized High Flows	25
2.7	Normalized Low Flows	26
2.8	Stations' location	27
2.9	Area and Elevation of catchments	27
2.10	Catchment of river Var at Malaussene	28
2.11	Example of plotting the annual hydrograph for graphical inspection	30
2.12	Boxplot,source: Towards Data Science	30
2.13	Scatterplot of Area vs Elevation of Catchments	31
3.1	Soca River at Kobarid	35
3.2	Soca River at Kobarid	35
3.3	Rhone at Gletsch	36
3.4	Rhone at Gletsch	37
3.5	Durance at Serre Poncon	38
3.6	Durance at Serre Poncon	38
3.7	Map of the site with pixel comparison	38
3.8	Monthly regime curve with the correct pixel	39

3.9	Results for the three signatures, nosoc	40
3.10	Results for Monthly regime, nosoc	41
3.11	Subcomponents:monthly regime	41
3.12	Results for Q_5 , nosoc	43
3.13	Subcomponents: Q_5	43
3.14	Results for Q_{95} , nosoc	45
3.15	Subcomponents: Q_{95}	45
3.16	Results for the three signatures, varsoc	46
3.17	Results for Monthly regime, varsoc	47
3.18	Subcomponents:Monthly Regime varsoc	47
3.19	Difference between socio economical scenarios,Monthly regime	48
3.20	Results for Q_5 , varsoc	49
3.21	Subcomponents: Q_5 varsoc	49
3.22	Difference between socio economical scenarios, Q_5	50
3.23	Results for Q_{95} , varsoc	51
3.24	Subcomponents: Q_{95} varsoc	51
3.25	Difference between socio economical scenarios, Q_{95}	52
3.26	Monthly regime vs morphological properties	54
3.27	Q_5 vs morphological properties	54
3.28	Q_{95} vs morphological properties	55
3.29	Ensemble Mean different KGE	56
3.30	DBH different KGE	57
3.31	Matsiro different KGE	58
4.1	Results of Gudmundsson et al[9] for quantiles evaluation	60
4.2	Relative change in Mean annual specific runoff GAR	63
4.3	Monthly regime for the historical and future (RCP 8.5) Soca at Kobarid . .	64
4.4	Change in Pardè range for the GAR	64
4.5	Monthly regime for the historical and future (RCP 8.5) Rhone at Gletsch .	65

List of Tables

- 1.1 Summary of reanalysis datasets 7
- 1.2 Summary of socioe economical scenarios 9
- 1.3 List of participating models 13

- 3.1 KGE Score for Soca catchment 36
- 3.2 KGE Score for Rhone at Gletsch 37
- 3.3 KGE Score for Durance at Espinasses 39

Introduction

The Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) is an international research initiative that provides a framework for assessing and comparing the impacts of climate change across various sectors of the environment and society. The project aims at filling the gap between the future climate projections, which have long been available, e.g. through the Climate Model Intercomparison Project (CMIP) and the consequences of such projections on both anthropic and natural systems. This thesis presents a thorough data analysis of the ISIMIP output dataset for the water sector for the region of the European Alps, GAR (4-19° E, 43-49° N)[2] for the historical period 1971-2000. Beyond the data analysis, showcasing the general feature of the dataset, the purpose of the thesis is to evaluate the performance of the Global Hydrological Models (GHMs) that participate in the project in terms of representing the intricacies of the alpine hydrology. This is done through a rigorous statistical analysis of the model's output data versus an observational dataset. The objective is to provide new insights into the current capabilities of GHMs while focusing on the GAR. This work finds its relevance in the context of the efforts to adapt and mitigate impacts of climate change on human society. The choice of the region is strategic: The Great Alpine Region was referred to as the "Water Towers of Europe" (European Environmental Agency [33]), being the spawning point of some of the most relevant sources of freshwater of the continent.

0.1 General background

Global Hydrological Models are computer-based numerical models that focus on resolving the movement of water through the Earth's surface. They do so by conceptualizing the relevant processes and solving mathematical equations on a discretized representation of the land masses. Over the years they have been recognized as valuable tools for many research topics and practical purposes relevant to our interactions with the hydrosphere (e.g Doll et al. [6]). Their areas of application range from pure scientific research, for instance climate impact projections, to measures related to water resource planning, such as agriculture and flood management. The general idea behind this work is to understand the functioning of the GHMs while applied to the GAR. The main hypothesis is that the Alps are an area where GHMs will have shortcomings related to:

1. Complex morphologies with very steep elevation gradients

2. Presence of processes related to snowmelt, glaciers and permafrost environment

Indeed, in a similar study focusing on the evaluation of GHMs in 6 Pan-arctic watersheds, Gädeke et al [12] found out that, in general, all models reproduced poorly the hydrological processes of the arctic environment, especially the correct seasonality of discharges, which is related to permafrost and snowmelt mechanisms. The interest for this analysis is therefore two-folded: on one hand, the alpine hydrology is and will be highly impacted by climate changes effects , with severe repercussions on the economies on both local and continental scale[8] [33]; on the other side, the hydrological modelization of an area with such heterogeneous geomorphological features is a particularly challenging task which we speculate current GHMs will have trouble representing and believe requires further research. As other studies pointed out when evaluating GHMs inside specific regions, e.g. [7] and many more, the models have difficulties in correctly reproducing observed values. To understand in a comprehensive way which are the limitation of the models and their advantages a thorough performance evaluation is carried out: three selected hydrological signatures are compared against observed values from river gauging stations located in the GAR. The adopted evaluation metric is the modified Kling Gupta Efficiency (KGE'), [11],selected for its ability to measure the misfit between observed and modeled values while at the same time providing for a finer interpretative tool because of its subdivision of the misfit in terms of three components: timing volume and variability.

0.2 Aim of the work and research questions

The scope of the work can be summarized as: understanding how Global Hydrological Models work, in particular those who adhere to ISIMIP, when applied to the specific region of interest of the GAR. The main topics we cover are therefore the following:

1. Divulge the general structure of both ISI-MIP and of GHMs
2. Explore how the ISIMIP data set represent hydrological signatures for the GAR
3. Evaluate the performance of the GHMs when applied to the GAR on the selected hydrological signatures
4. Discuss the possible reasons for shortcomings and good performance of models and the potential uses in the context of climate change projections

Each of the following chapters will tackle one of the listed objectives. In conclusion, this thesis provides a big picture of the current state of the global hydrological models, while at the same time delving deeper into their limitation and advantages through the performance evaluation on the selected area.

Chapter 1

Theoretical Background

In this chapter we give a top-to-bottom view into the ISIMIP project. Firstly, we describe the generalities of ISIMIP. Secondly, we delve deeper into the specific components of the project by addressing their structure. The elements under study are the reanalysis datasets and socio economical scenarios for the input side and the Global Hydrological Models on the output side. Each of these element presents non-trivial information which we discuss in order to have a complete background and understand the performance evaluation set-up and its results. Lastly, we explain some of the features of alpine climate and hydrology in order to focus once again the attention on the study area.

1.1 Overview of ISIMIP

ISIMIP was designed by the Potsdam Institute for Climate Impact Research (PIK) in collaboration with the International Institute for Applied Systems Analysis (IIASA) to help answer two main questions:

- How is climate change affecting human society and natural systems today?
- How will it do so in the future?

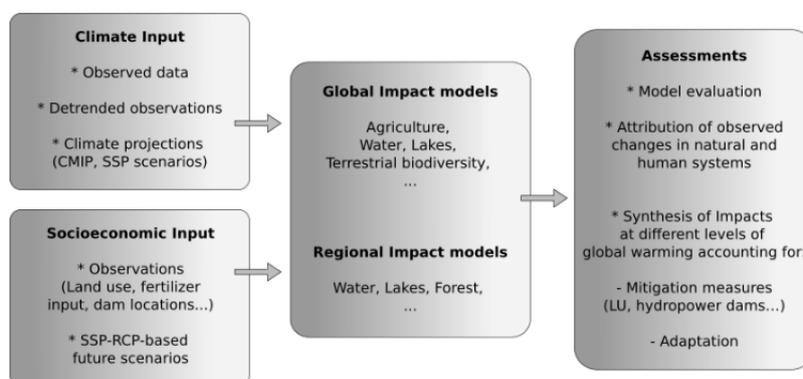


Figure 1.1: General scheme of ISIMIP, source ISIMIP

The project currently gathers more than 100 impact modellers' team from across the globe and has been cited in many publications, most notoriously the IPCC AR5 and 6.[9] The experiments are organized in successive rounds, the experiment ISIMIP-2 is the latest for which results are publicly available. The general scheme, presented in figure 1.1, is that climate and socio economical data are the input to the impact models to then produce outputs in the form of effects on selected sectors. Global impact modelling is the core of the project, with models having a $0.5^\circ \times 0.5^\circ$ resolution (50 x 50 km at the equator), although specific focus regions also feature regional models runs at higher resolution.

1.1.1 ISIMIP water sector

The water sector entails the modelization of the water exchanges from the atmosphere through the land masses, as schematized in fig. 1.2. It is arguably one of the most direct declination of climate impacts.

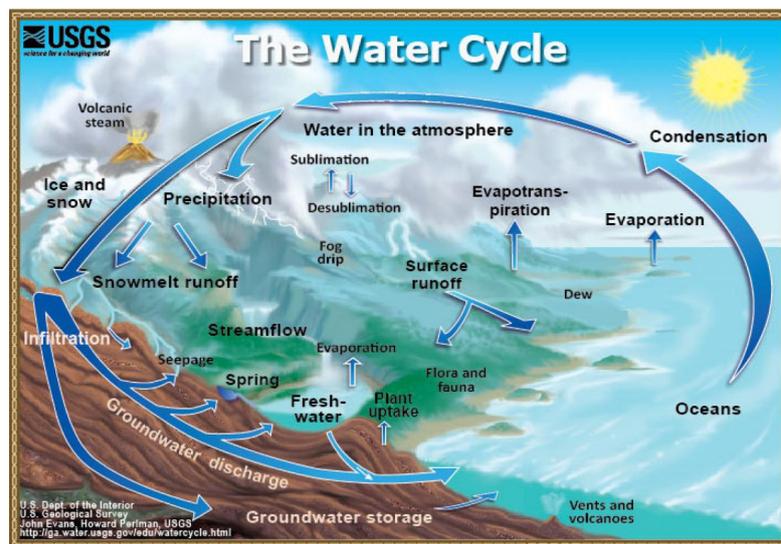


Figure 1.2: The water cycle, credits source USGS

For the Water Sector in ISIMIP2 there are two separate batches of experiments. One is the validation run (protocol 2a) while the other is the run for the future projections (protocol 2b). In the 2a protocol the climatic input is in the form of reanalysis data, i.e. historically observed meteorological data which has been gridded, interpolated and gap filled in order to provide a spatially and temporally dense input for the model to run at the selected space and time resolution. In the protocol 2b instead, the input are the output data of the climate models, therefore simulated climatic variables.

In terms of socio-economical input the main distinction in the scenarios presented for the 2a and 2b protocols is that socio economical scenarios for the future have an impact on the climate input themselves, being that anthropogenic climate change effects will vary depending on the strategies and measures of adaptation that human society will employ. These scenarios are the Representative Concentration Pathways (RCPs) and Shared Socio

Economical Pathways (SSPs) introduced by the IPCC. In the 2a set of experiments instead, the socio economic scenarios relate to the including or not of the most realistic representation of all the anthropogenic interferences to the water cycle. This interferences can be roughly summarized as:

- Human impacts on land use
- Human impacts on water abstraction

The realistic and consistent representation of all the impacts that humans have on the hydrosphere as well as their changes over the historical period should provide for the most loyal and true to observed values results. However, this has to be verified, as including elements of realism also means having to parameterize more sub-pixel-scale processes that cannot be explicitly modelled as they occur at a much finer scale than that of the grid of the models, therefore introducing new sources of error. This thesis focuses on the historical run or ISIMIP 2a. This is because the purpose of the thesis is to delve deeper into hydrological modelling.

1.1.2 Climate input: Reanalysis datasets

In the ISI-MIP 2a experiment's round, the first common input that we present are the climatic datasets that are used to force the impact models. Those are datasets common to all sectors as all impact models take climatic input and use it to derive its effects on specific sectors.

Table 1 presents a list of all the datasets available for ISIMIP along with a summary description of each of them. The general feature, as previously explained, is that all of them are reanalysis products. Reanalysis entails combining past observational data with numerical weather prediction models to generate a comprehensive and plausible description of the Earth's atmosphere over a specific period. Unfortunately the three most recent products based on the more advanced ERA-interim or even more the product based on ERA-5 were not used in the ISIMIP 2a round; as their were only recently published. Therefore the choice was to evaluate the GHMs forced by the WATCH WFD dataset which was among the most consolidated dataset used in other similar studies [9]. Also from the results of existing literature, we considered to use only one forcing dataset. Other studies agreed on the fact that the contribution to the overall uncertainties coming from the choice of the dataset is less relevant than the contribution from the choice of a different GHM[12]. However, that could be also due to the fact that similar products based on similarly dated reanalysis methodologies presented the same limitations. It remains to be seen in the next round of ISIMIP whether that will be the case with the newly developed products which have already given promising results in the context of impact modelling [5].

WATCH WFD dataset The Water and Global Change (WATCH) project is an initiative focused on evaluating and understanding the terrestrial water cycle. Its primary objective

Name	Reanalysis	Years	Resolution, coverage	Bias target	Primary Comments
WATCH-WFDEI	ERA-40, ERA-interim	1901- 2016	0.5°, Land	CRU, GPCC	Combined WFD(1901-1978) and WFDEI (1979-2016) discontinuity present in the data
GSWP3-W5E5	ERA5	1901-2016	0.5°, Land+Ocean	CRU, GPCC, GPCP	Combined GSWP3(1901-1978) and W5E5 (1979-2016) discontinuity present in the data minimized via bias adjustment methods
GSWP3-EWEMBI	ERA-interim	1901-2016	0.5°, Land+Ocean	CRU, GPCC, GPCP, SRB	Combined homogenized GSWP3 (1901-1978) and EWEMBI (1979-2016) discontinuity present in the data minimized via bias adjustment methods
GSWP3	20CR	1901-2010	0.5°, Land+Ocean	CRU, GPCC, GPCP,SRB, CPC-Unified CRU, SRB, TRMM,	Based on dynamical downscaling
PGMFD v2.1 (Princeton)	NCEP/NCAR Reanalysis 1	1901-2012	0.5°, Land+Ocean	GPCP, WMO validated against GSWP2	
WATCH (WFD)	ERA-40	1901-2001	0.5°, Land	CRU, GPCC	

Table 1.1: *Summary of reanalysis datasets*

is to assess various hydrologically important variables, including evaporation, soil moisture, and runoff. This evaluation is conducted using land surface models and general hydrological models. On of the first step of the project focused on creating the required meteorological forcing data as input named WATCH Forcing Data (WFD). This is the dataset that is also used in the ISIMIP project and that we adopted for the performance evaluation.

We now briefly describe the general properties of the WFD. The WFD covers the time period from 1958 to 2001, with data based on the 40-year European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). ERA (ECMWF Re-Analysis) is a series of global atmospheric reanalysis datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). In particular ERA-40 is the second product of the ERA series after ERA-15, it has a coarser resolution, around 125 km globally, then the required 50 km (or 0.5°) of the impact models and for this reason it was processed in order to comply with the spatial resolution of the GHMs.

The WFD consists of subdaily, regularly gridded data with a half-degree resolution in latitude and longitude. It covers the period from 1958 to 2001. The WFD includes various meteorological variables, such as wind speed at 10 m, air temperature at 2 m, surface pressure, specific humidity at 2 m, downward longwave radiation flux, downward shortwave radiation flux, rainfall rate, and snowfall rate. The data is stored at 67,420 points over land (excluding the Antarctic) and follows the netCDF format using the Assistance for Land-Surface Modelling Activities (ALMA) convention.

The creation of the WFD for the late twentieth century involved bilinear interpolation of each variable from the ERA-40 grid to the CRU land-sea mask at a 0.5° resolution. Elevation corrections were applied to the interpolated temperature, surface pressure, specific humidity, and downward longwave radiation in sequential order, taking into account the elevation differences.

The 2-meter temperatures in ERA-40 were known to lack some climatic trends and exhibit biases. Therefore, bias correction was performed on the monthly average interpolated and elevation-corrected temperatures using CRU TS2.1 gridded observations.

CRU TS (Climatic Research Unit Time-Series) is a dataset produced by the Climatic Research Unit (CRU) at the University of East Anglia in the United Kingdom. It provides global gridded climate data at various temporal resolutions. The CRU data used for bias correction includes adjustments for inhomogeneities between stations and incorporates the correlation length of the variables. However, it may still have limitations and rare inhomogeneities. Offsets and outliers in the CRU temperature data were identified and removed prior to their use in bias correction. Monthly diurnal temperature ranges were also corrected using the CRU data.

Advantages and limitations The WFD dataset has been used priorly in impact modelling studies [12] and comes from a well-established climate research project such as the ERA; those are the main reasons why it was selected. Another advantage is the fact that it presented no discontinuities for the selected period 1971-2000 giving the opportunity to utilize a full 30 years of daily data for the analysis, which we deemed as another important factor especially in the context of working with extreme high and low flows.

Apart from the advantages and opportunities it is also crucial to be aware of the main limitations of the dataset in order to correctly interpret the results of the following analysis: as previously pointed out there is a well known issue in the variable of precipitation in the ERA-40 analysis. That is due to the coarse resolution and it is particularly evident in the case of alpine area.[15] Indeed, studies analyzing the precipitation data for the alpine region have shown that a significant negative bias exist in the representation of the precipitation over mountainous catchments e.g Adam et al [26]. The authors of the studies have identified the coarse resolution as being the source of this error, because the data at that spatial scale failed to reproduce highly localized phenomena of both high and low intensity, which are also a typical feature of small alpine catchment.

1.1.3 Socio Economical Scenarios

Impact modelling entails reproducing a wide array of realistic properties of the physical world. Among these properties are the impact of anthropic presence on the environment. Those elements are included in the socio economical scenarios.

In Table 2 the scenarios are listed along with their main characteristics. In figure 1.3 instead, a graph show a representation of both climate and socio economical input providing a better understanding of the input side.

Scenario Name	Scenario description and comments
nosoc	No human influences except for year-2000 constant land-use patterns. No anthropogenic water abstraction (e.g., irrigation), no reservoirs/dams. No population and GDP data prescribed.
pressoc	Present-day human impact runs: only climate varies; keep all other settings (population, GDP, land-use, technological progress, etc.) constant at year 2000 values. This run will be used to quantify adaptation pressure under current socioeconomic conditions. For water models, pressoc includes present-day irrigation and other water uses / reservoirs.
varsoc	Not only climate but also population, GDP, land-use, technological progress, etc. varies over the historical period.It is the most complete and realistic representation possible

Table 1.2: *Summary of socio economical scenarios*

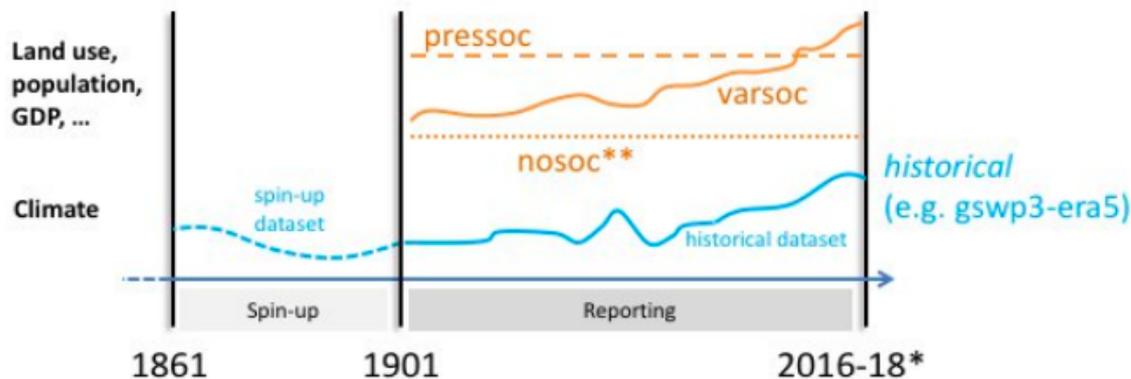


Figure 1.3: *Graph of input data to GHMs*

Discussion on Socio Economical Scenarios: VARSOC and NOSOC While, as stated, providing a complete and realistic representation of the world, including the its anthropic elements, is indeed one of the core prerogative of impact modelling, in contrast for example with pure climate modelling. Their correct representation should not be given as a a priori verified assumption. In fact in some of the reference studies [9] the model run are taken as only "nosoc" because of the small size of the catchments and the alluded similarity between the modelling of nearly undisturbed catchments and the model of singular pixels of the grid. In others, such as [7] or [12] the distinction is not even highlighted. A single study [30] was found that systematically compared the performance under the different scenarios. While this study did found out that the inclusion of human impacts lead to a sensible increase in performance of the models versus observed data. It was noted how this improvement was more pronounced for highly managed catchments then nearly natural ones. Considering the above notions we opted to verify also these assumption and validate the two scenarios separately in order to understand their limitation and advantages in the context of the specificity of the alpine catchments.

1.1.4 The GHMs

We have already defined GHMs in terms of their aim of resolving the terrestrial part of the water cycle. In this section we delve deeper into their structure and list all the models participating in the ISIMIP project.

First and foremost a distinction is due: the rough definition given above for the term Global Hydrological Models, actually applies to at least two different typologies of tools that were historically developed for different purposes. While Hydrological Models, both global or regional, were developed for the purpose of water resource management, Land Surface Models (LSMs), conceptualize the same processes but for the purpose of providing a lower boundary to atmospheric models in the wider context of climate modelling. The main distinction between the two is the fact that LSMs can be coupled to atmospheric models and put

emphasis on describing the vertical heat, water and sometimes carbon fluxes; HMs on the other hand, focus more on describing the lateral movement of water. In summary LSMs solve both the energy balance and water balance; GHMs solve only the water balance. Haddeland et al.[13]provides a study in which the performance of the two are thoroughly compared. Besides the different original purpose they all represent the hydrosphere and adopt similar schemes and parametrizations. Furthermore, model development and advancements continue to blur the boundaries between these categories as models become more integrated and comprehensive in representing Earth system processes. Therefore for the purpose of this study the wide definition of GHMs is extended to both LSMs and GHMs.

Structure of the models adhering to ISIMIP In general, o the GHMs resolve on the afore-mentioned $0.5^\circ \times 0.5^\circ$ grid the balance equation 1.1 for all the grid cells

$$\frac{dS}{dt} = P - E - Q \quad (1.1)$$

Where:

- $\frac{dS}{dt}$ = changes in storage
- P = precipitation
- E = evapotranspiration
- $Q = Q_s + Q_g - Q_a =$ runoff sum of surface (s) and groundwater (g) components minus water abstracted (a)

Figure 1.4 further illustrates the general scheme applied by most models. It must be however noted how each of them applies different methods and parametrizations of the same processes. Table 1 presents the list of all the models participating in the ISIMIP 2a experiment along with their reference paper/ technical manual.

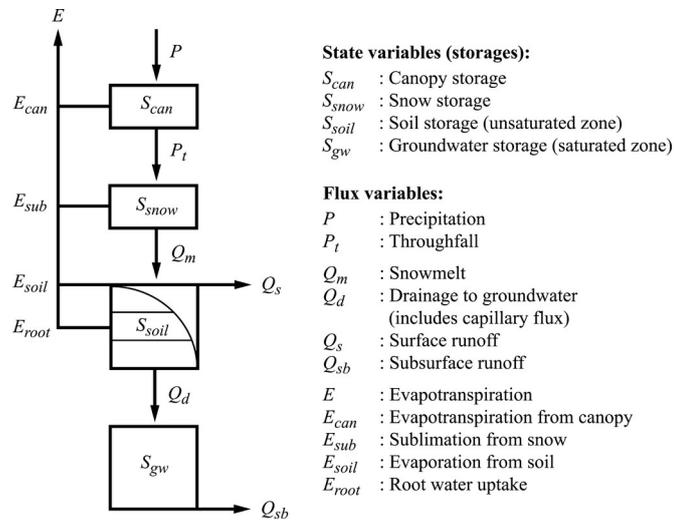


Figure 1.4: Scheme of modelled processes, source [9]

In table 2, in italic, we highlight the model which fall under the classification of LSMs. We also show the different modelization of the snow scheme. This information showcases the main difference between the two, with GHMs having the distinction between rain and snow precipitation based on the daily value of temperature. On the other hand LSMs need radiation input at a finer temporal scale (hourly or 6-hours) in order to evaluate the heat fluxes and radiative balance. This distinction is also relevant in the context of our study area being interested by snow precipitation and should be kept in consideration while interpreting the results. For instance Haddeland et al [13], in their study focusing on the comparison of the two kinds of models, found that the physically based radiative balance method applied by the LSMs was consistently reproducing a smaller snow water equivalent values compared to the degree day approach.

1.2 Background on the GAR

1.2.1 Overview of the Alps and climate change

The Alpine ridge extends for 800 kms and has a width of 200 km on average. It presents unique morphological characteristics such as high peaks reaching up to 4800 meters above sea level interspersed by low valleys (see figure 1.5). It is one of the dominant feature of the European landscape. For centuries it has been an area of interest for scientists, starting from the work of Horace Benedict de Saussure who already in the eighteen century, therein laid the basis for the discipline of mountain meteorology [23]. Over the course of the last century the interest has morphed into a widespread alarm due to the fact that the measured effect of global warming on the area amounted to a regional increase in temperature of 2° Celsius already by 2009, indicating a warming rate twice as much as the average for the Northern emisphere. [33]

This worrisome climatic evidence intertwines with other reasons for concern: the Alps are a densely populated area with high economic relevance for the European continent. Many sectors such as agriculture, hydropower generation and tourism, are threatened by such a fast rate of change that could be destructive for the Alpine natural and anthropic systems. A more specific and, in the context of this study, more notable vulnerability factor for the GAR, lies in the fact that the 4 major European river, the Danube, Rhine, Rhone and the Po have their spawn in the alpine area and are fed majorly by the runoff generated in the Alps. More precisely the Danube receives on average 26% of its discharge from the alps while for the Po river the Alps' contribution reaches values of 53%. This proves the utmost importance of the GAR and the possible catastrophic effects of the disruption of its environmental functions at both local and continental scale.

List of participating models	Socio Economical Scenarios	Snow Scheme	Main technical reference
<i>CLM 4.0</i>	nosoc	Energy Balance.	Oleson et al[20]
DBH	nosoc varsoc	Degree day	Tang et al [29]
<i>H08</i>	nosoc varsoc	Energy balance	Hanasaki et al [14]
<i>JULES-W1</i>	nosoc	Energy balance	Best et al [4]
LPJmL	nosoc varsoc	Degree day	Sitch et al [24]
<i>MATSIRO</i>	nosoc varsoc	Energy balance	Takata et al [28]
MPI-HM	nosoc	Degree day	Stacke et al [25]
<i>ORCHIDEE</i>	nosoc	Energy balance	Guimberteau et al [10]
PCR-GLOBWB	nosoc	Degree day	Sutanudjaja et al [27]
<i>VIC</i>	nosoc	Energy balance	Liang et al[1]
WAYS	nosoc	Degree day	Mao et al[17]
<i>WEB-DHM-SG</i>	nosoc	Energy balance	Qi et al[22]
WaterGAP2	nosoc varsoc	Degree day	Muller et al [19]

Table 1.3: *List of participating models*

1.2.2 GAR, climate modelling and future projections

For the purpose of this work, what needs to be noted is the complexity and heterogeneity at different scales that characterizes the climatology of the study area. Indeed the GAR features notable variations in climate variables across all three space dimensions (latitude

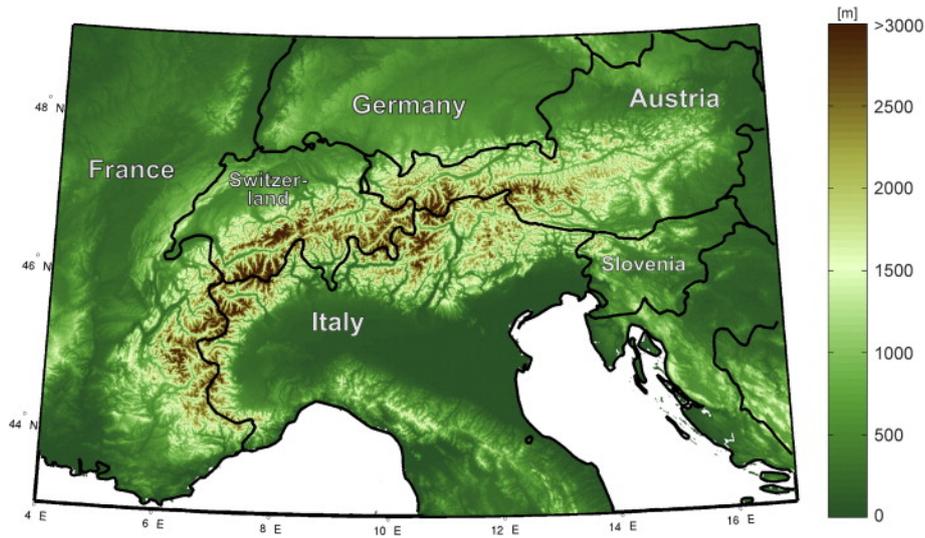


Figure 1.5: *European alps with orography, credits [8]*

and longitude and elevation). It experiences frequent occurrences of intense precipitation events, which can result in various hazards. The presence of permanent and transient snow and ice in the higher altitudes is of great significance. Additionally, the complex topography often leads to distinct flow phenomena influenced by the surrounding mountains and hills. One of the latest and more successful efforts at characterizing the climatology of the Alps is the Histalp project [2]. Its aim was to provide for a spatially and temporally dense meteorological database for the GAR. Among its results there is also the regionalization of the GAR into 4 main sectors, i.e. North and South West sectors and North and South East sectors. The different macro regions were retrieved via Principal Component Analysis (see [32]) the meteorological variables available. What was achieved was the subdivision of the GAR region in 4 different regions each with its peculiar climatology. This result is visible in figure 1.6 which was retrieved from [33].

The results of these studies highlight the complexity intrinsic to modelling the GAR. In the case of the hydrosphere in particular, even on the historical trend recent analysis have shown differentiated results. Most notably, in the NW sector the overall yearly precipitation has increased (+9%) while in the SE a significant decrease has been detected (-9%) over the 20th century [2]. The differences are relevant also on the temporal interannual scale, while precipitation in winter are projected to increase, especially the rain fraction over the snow, precipitation in summer is instead projected to decrease with very dry summers, especially in the South corners becoming more and more frequent[8]. Logically these changes will continue to impact the runoff of rivers. The decrease in the snow cover and its duration through the winter, will have important effect on the decrease of summer runoff for all the rivers which are heavily fed by the Alps.



Figure 1.6: GAR regionalization and principal river basins

1.2.3 Major Rivers of the GAR

To further expand on the GAR's hydrology, we focus on this section on the characteristics of the 4 major rivers spawning in the area. These are, as previously mentioned, the rivers Danube, Rhone, Po and Rhine which flow into all of the 4 climatic sub-regions highlighted by figure 1.6.

Danube (North East)

The Danube has the second largest river basin in Europe (after river Volga) and is the main tributary to the Black Sea. The river crosses three climatic zones due to its extensive length from west to east: the Mediterranean climate, the continental climate with lower precipitation and colder winters, and the Atlantic climate with high precipitation. Germany, Austria, Slovak Republic, Hungary, Croatia, Serbia and Montenegro, Bulgaria, Romania, Moldova, and Ukraine are among the ten nations that the Danube passes through or borders.

The Danube begins in the Black Forest (Schwarzwald) of Germany, at an elevation of roughly 1000 meters. It is interesting to note that in its higher course around Immendingen, the Danube loses roughly half of its discharge to the Rhine basin through underground passageways (reduction from 12 to 6 m³/s). After traveling 2,780 kilometers, the Danube arrives in the Black Sea, mostly flowing to the south-east. The Danube discharges water at its mouth on average at a rate of roughly 6,500 m³/s in Tulcea [31]. Austria's portion of the Danube catchment is primarily enclosed by the Alps, and it produces 1 448 m³/s, or 22%, of the basin's overall runoff (ICPDR, 2005).

Table 3.3 Characterisation of the Danube river basin

Length of river	2 780 km
Total catchment area	801 463 km ²
Inhabitants within catchment area	~ 81 million
Concerned states	Albania (126 km ²), Austria (80 423 km ²), Bosnia Herzegovina (36 636 km ²), Bulgaria (47 413 km ²), Croatia (3 4965 km ²), Czech Republic (21 688 km ²), Germany (56 184 km ²), Hungary (93 030 km ²), Italy (565 km ²), the former Yugoslav Republic of Macedonia (109 km ²), Moldova (12 834 km ²), Poland (430 km ²), Romania (232 193 km ²), Serbia and Montenegro (88 635 km ²), Slovak Republic (47 084 km ²), Slovenia (16 422 km ²), Switzerland (1 809 km ²), Ukraine (30 520 km ²)
Important tributaries	Lech*, Naab, Isar, Inn*, Traun*, Enns* Morava/March, Raab/Rába, Vah, Hron, Ipeľ/Ipoly, Sió, Drau/Drava*, Tysa/Tisza/Tisa, Sava, Tamis/Timis Morava, Timok, Jiu, Iskar, Olt, Yantra, Arges, Ialomita, Siret, Prut
Major uses	Navigation, hydropower, industry (chemical, food, textile, metal, paper, car, service etc.), agriculture, drinking water, local recreation
Note:	* Alpine tributaries.

Figure 1.7: Danube statistics, source ICPDR via EEA 2009

Rhine (North West)

The river Rhine connects the Alps with the North Sea. It forms the most important cultural axis of central Europe [3], with cities such as Basilea, Dusseldorf, Rotterdam and Strasbourg being located on its course. The two headstreams of the Rhine, which are the "Vorderrhein" and the "Hinterrhein" both spawn in the canton of Grisons in Switzerland in an area with high peaks with elevations reaching 3000 m.s.l.. The region where it flows it's highly populated and after the confluence of the two spawning rivers the anthropic effects become highly significant, such as water for industry and hydropower (see figure for details). The river morphs its regime from nival-glacial to pluvial down its course due to addition of significant amounts of water by tributaries not spawning in the Alps; nevertheless, the alpine section, which constitutes only 15% of the total area of the catchment, impacts for 34% of the contribution to the Rhine, with peaks in summer of percentage up to 50% due to the decreased contribution of the pluvial signal in respect to the reliable contribution of snow and glacier melt. This last contribution is also vital in the sense of providing good quality water for drinking purposes.

Po(South East)

The Po river spawns in Piedemont Region in the Monviso Mountain at 2022 m.s.l. It flows in a large delta outlet into the Adriatic Sea. Most of its basin its in Italy (70 000 km²) while a little portion of it lies in Switzerland (4 000 km²). It is surrounded by the Alps in the North West and the Apennines on the South. The approximate distribution of environment typology is 2/3 mountainous and 1/3 for the Po river basin. Its regime derive of a complex interaction between the influences of the two aforementioned mountainous areas (Alps and Apennines) and the Mediterranean climate. The contribution of the Alps accounts

Length of river	1 320 km
Total catchment area	185 000 km ²
Inhabitants within catchment area	50–58 million
Nine riparian* states	Germany (approximately 100 000 km ²); Switzerland, France and the Netherlands (20 000–30 000 km ²); Austria and Luxemburg (approximately 2 500 km ²); Italy, Liechtenstein and Belgium (a small percentage)
Headstreams	'Vorderrhein' and 'Hinterrhein'
Important tributaries	Aare, Ill, Main, Nahe, Mosel, Ruhr, Lippe
Six major sections	Alpine Rhine (confluence — Lake Constance); High Rhine (Lake Constance–Basel); Upper Rhine (Basel–Bingen); Middle Rhine (Bingen–Cologne); Lower Rhine (Cologne–Dutch border); Rhine Delta
Major uses	Navigation (possible from Rotterdam to Basel; ships up to 3 000 t) hydropower, cooling water, industry (chemical, food, textile, metal, paper, car, service, etc.), agriculture, drinking water, local recreation

Note: * States located on the banks of the Rhine river.

Figure 1.8: *Rhine statistics, source IKSR via EEA 2009*

on average for 53% of the river's discharge as already stated; among this contribution of note is the contribution of around 600 km² of glacier areas. Its mean discharge at the outlet is around 1600 m³/s. The water of the Po river is heavily under pressure by industry, agriculture and hydropower production. Across the Alps 174 reservoirs manage on average 2 766 million m³ per year.

Length of river	657 km
Total catchment area	74 000 km ²
Inhabitants within catchment area	~ 16 million (whole basin)
Concerned states	Italy (70 000 km ²) and Switzerland (4 000 km ²)
Important tributaries	Adda*, Agogna, Banna, Chisola, Dora Baltea*, Dora Riparia, Grana del Monferrato, Lambro, Maira, Malone, Meletta, Mincio*, Oglio*, Olona, Orco, Pellice, Rotaldo, Sangone, Sesia, Stura del Monferrato, Stura di Lanzo, Tanaro, Terdoppio, Ticino*, Varaita
Major uses	Hydropower, cooling plants, industry (chemical, food, textile, metal, car, service, etc.), irrigation for agriculture, drinking water, local recreation. Navigation limited due to the presence of reduced sections

Note: * Principal alpine tributaries.

Figure 1.9: *Po statistics, source Po River basin Authority via EEA 2009*

Rhone(South West)

The Rhone is the second largest input of freshwater to the Mediterranean after the Nile. Its spawn it's in the homonymous glacier located in the Swiss canton of Valais, in an area comprising peaks of up to 4000 m of elevation. Its basin its influenced by climatic input of both Mediterranean, Continental and even Polar origin. It's tributaries are mostly alpine streams and therefore its regime is mostly nival/glacially regulated. It flows through the largest water body in western union, Lake Lemman, which regulates its regime and serves as a mitigating factor for flood downstream of it. The first section of the river before lake Lemman, is the one which is most influenced by the alpine tributaries with a high contribution in the

spring discharge due to snowmelt. The last section, in France and especially after Avignon the Mediterranean influence becomes stronger with extreme precipitation event often causing flooding events in the tributaries mostly rivers Isere and Durance. Its average discharge is close to $1700 \text{ m}^3/\text{s}$ near the outlet.

Length	812 km (Swiss part: 261 km)
Total catchment area	95 500 km ² (Swiss part: 10 100 km ²)
Inhabitants within catchment area	16 million (Swiss part: 1.2 million)
Riparian states	Switzerland, France
Important tributaries	Saone, Isère, Durance, Ain (in Switzerland: Vispa, Dranse, Arve)
Five major sections	Valais (Wallis) transect; Lake Léman; Jura transect; Provence transect; Rhone delta (Camargue)
Major uses	Hydropower, water for cooling power stations, irrigation for agriculture, industry, recreation, navigation (up to Lyon; beyond Lyon on the Saone and on the Rhone-to-Rhone canal system)

Figure 1.10: *Rhone statistics, source Encyclopaedia Britannica via EEA 2009*

Impacts of Climate Change on the Major GAR rivers, past observations and future perspective We have stated how each of those 4 major rivers is majorly impacted by the Alpine climatology. In different proportions and sections, all of the mentioned rivers are under a significant effect by the mechanisms of snow and glacier melting. Moreover these major streams support the lively-hood of millions of people in the European continent and also sustain a considerable portion of its biodiversity. Over the 20th century climate change has already affected the hydrology of the GAR and measure of adaptation have become more and more compelling. As stated there are however difference in the modality and degree of which these impact are distributed. The north and south quadrants of the Alps have shown differentiated results with the north being affected by increasing precipitation in autumn and winter. This for instance is shown in increasing trends for winter in the northern rivers, Danube and Rhine, especially close to their spawn. The southern rivers have been instead showing more significant impacts especially in the summer decrease which has been consistent and sizeable, affected by the lesser degree of snowmelt mitigation. The heatwave of 2003 has created numerous problem for the hydro power systems in all the major rivers, as water was too scarce and too hot to sustain a significant production over the summer months. With the future projections foretelling a situation of really hot and dry summers such as that of 2003 being as common as one every two summers in some scenarios [8], it is clear how the GAR and its water resource needs to be protected from further damages to both human and natural systems.

Chapter 2

Data and Methods

In this chapter we present the datasets that were used. The first section is dedicated to showcasing the features of the input/ output data of the ISIMIP GHMs ensemble. The second section instead presents the observational dataset that was used for comparison, this second sets of data comprises of gauge data from 179 stations. Finally we give an overview of the methods used to 1) setup the performance evaluation 2) carry out the performance evaluation.

2.1 ISIMIP data over the GAR

We downloaded the data from the ISIMIP repository (<https://data.isimip.org/search/>) which gives the possibility to directly download the data and even perform preliminary operations on it by cutting out geographical boxes of interest. I downloaded data for all 13 models participating in the nosoc scenario and the 5 model participating in the varsoc scenario. The variables are the following:

- Total runoff “qtot”, $qtot = \text{rainfall} + \text{snowmelt} - \text{evapotranspiration}$;

- Simulated discharge “dis”; simulated river discharge in specific points inside the grid.

In order to explore the dataset in a qualitative manner and get a first feel for the downloaded variables I performed some simple preliminary analysis. Following, is a sequence of images (timelapse) following the famous flood event which affected the Piedmont Region in October 2000. Both variables are observed in their evolution in a time window encompassing the days between the 14th and the 17th of October.

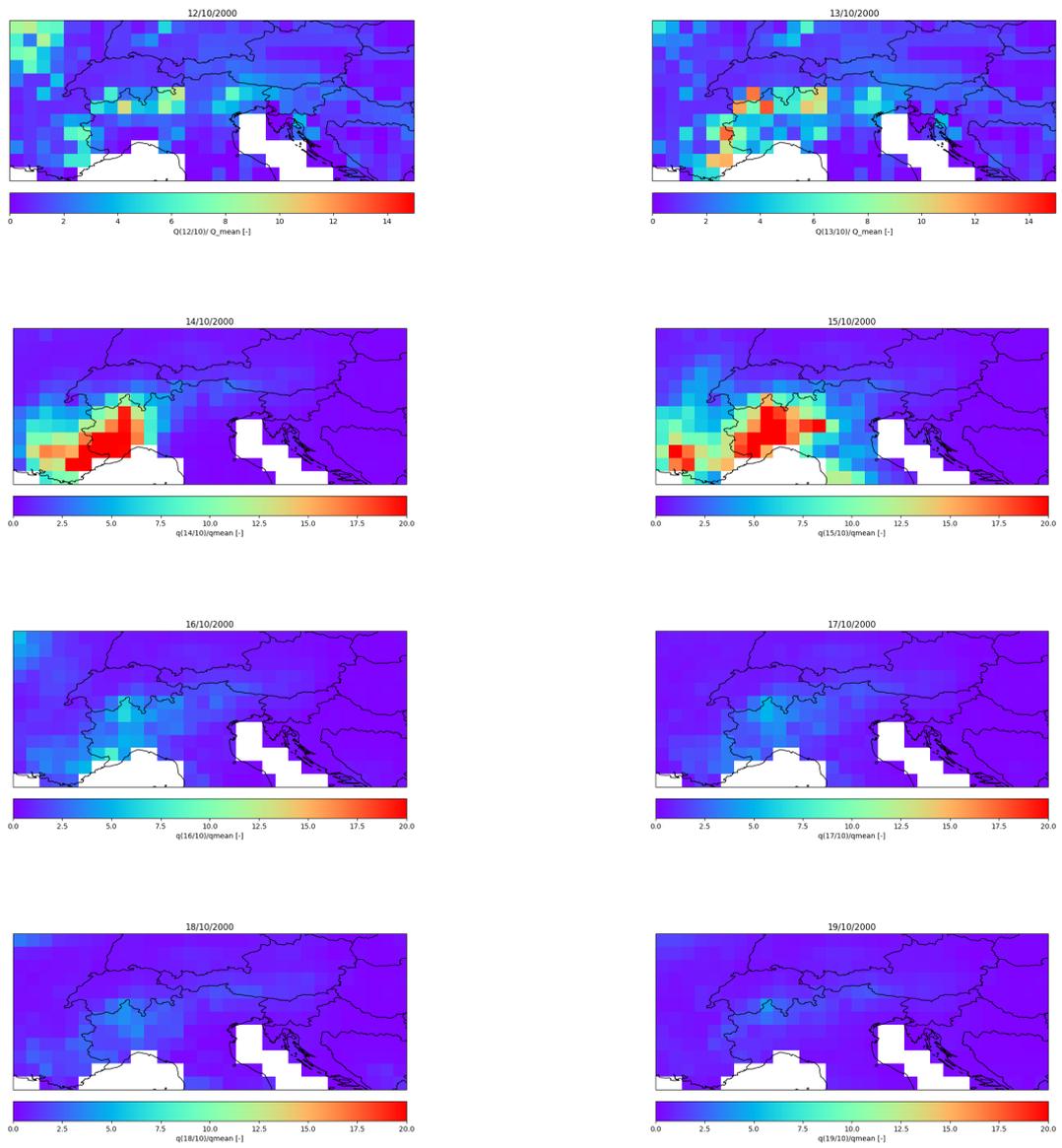


Figure 2.1: *Flood October 2000, generated runoff*

It is possible to visualize how the intense rainfall that fell over the northwest part of Italy resulted in a massive amount of generated runoff over the Piedmont and Valle d'Aosta regions. The Piedmont region in particular, suffered most of the damages. In total there were 23 deaths.

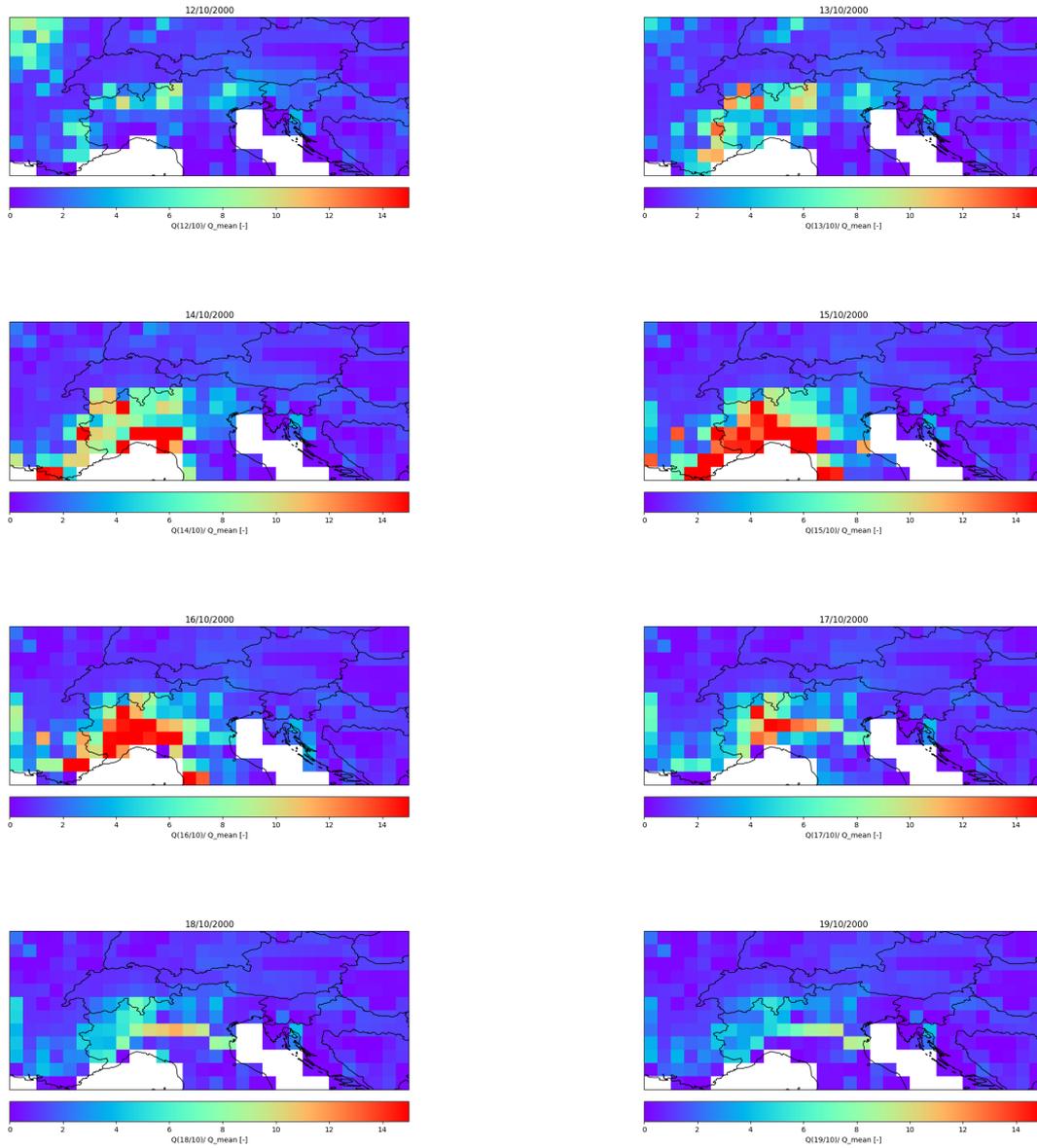


Figure 2.2: *Flood October 2000, simulated discharge*

From the second time lapse we can visualize the evolution of the simulated discharge during the flood. In particular, the signal of the Po river it's clearly visible. Between the 14th and the 17th of October in the metropolitan area of Turin the river reached peak discharges of $2300 \text{ m}^3/\text{s}$. The flood then proceeded downstream with limited damages.

2.2 Hydrological signatures over the GAR

Following we present another possible way of visualizing the data and extrapolate useful information about the region of interest. We will now create spacial maps of 5 selected signatures over the GAR. Such signatures are

1. Mean annual specific runoff
2. Pardè Range
3. Slope of the flow duration curve
4. Normalized High Flows
5. Normalized Low Flows

Mean Annual Specific Runoff (Q_m) is a measure of the generated runoff over any pixel of the region in term of mean annual values. It is calculated via the following equation:

$$Q_m = \frac{365}{T} \sum_{t=1}^N Q(t) \quad (2.1)$$

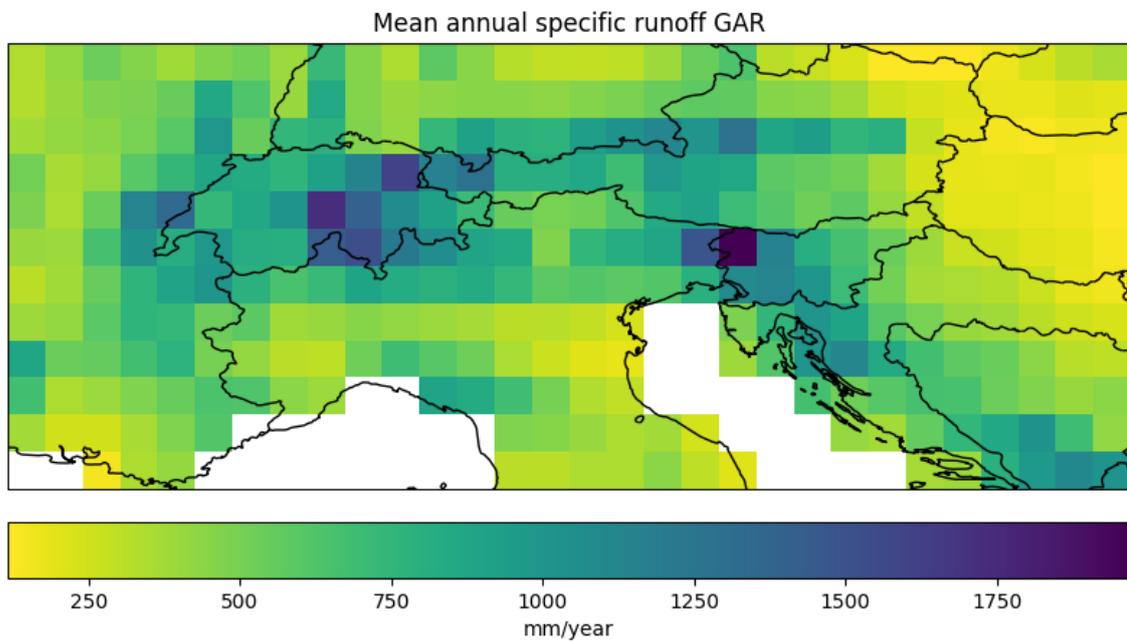


Figure 2.3: *Mean annual specific runoff*

Comments $Q = P - E$ is the long term balance for each pixel that the map in figure 2.1 represent. It is possible to appreciate the differences in annual runoff on the long term for the study area. It is evident how even within the GAR exists very humid catchments with annual runoff close to 2000 mm/ year and dry ones with annual runoff around 400 mm/year.

Pardè Range ΔP is a measure of how much ,throughout the year, the flow generation changes. It is defined as the difference between the maximum and minimum Pardè coefficients:

$$\Delta P = \max(Par_i) - \min(Par_i) \quad (2.2)$$

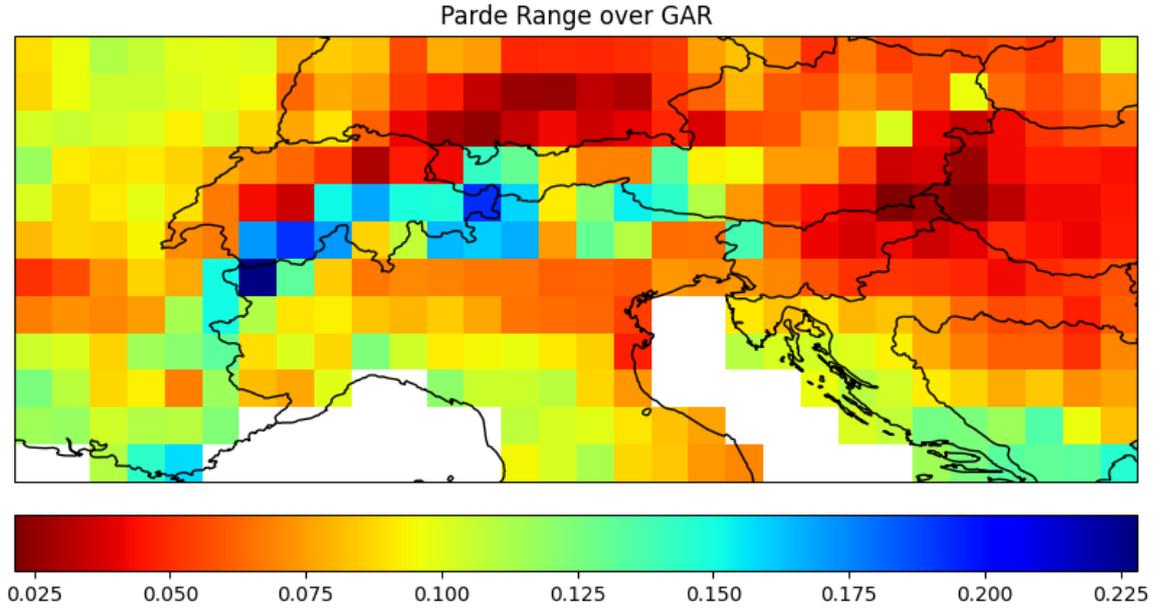


Figure 2.4: *Pardè Range*

With the Pardè coefficient for month i being defined as:

$$\frac{\sum_{t \in M_i} Q(t)}{\sum_{t=1}^N Q(t)} \quad (2.3)$$

Comments The map of the Pardè ranges showcases the different seasonal regimes present in the region. It is possible to appreciate at glance how the highest range exists for regions at high elevation, which indeed present a very high variation between the runoff peak of the spring months and the low values in the summer and winter months. The low Pardè range in the regions to the North of the GAR as well indicates the difference in the continental climate compared to the Mediterranean regions in the south of the Alps.

Slope of the flow duration curve (m_{FDC}) is slope of the flow duration curves in its central part. It indicates how much does a 1% increase in exceedance probability corresponds to in terms of decrease in runoff, in respect to the mean daily value.

$$m_{fdc} = \frac{Q_{0.30} - Q_{0.70}}{40 * Q_m} \tag{2.4}$$

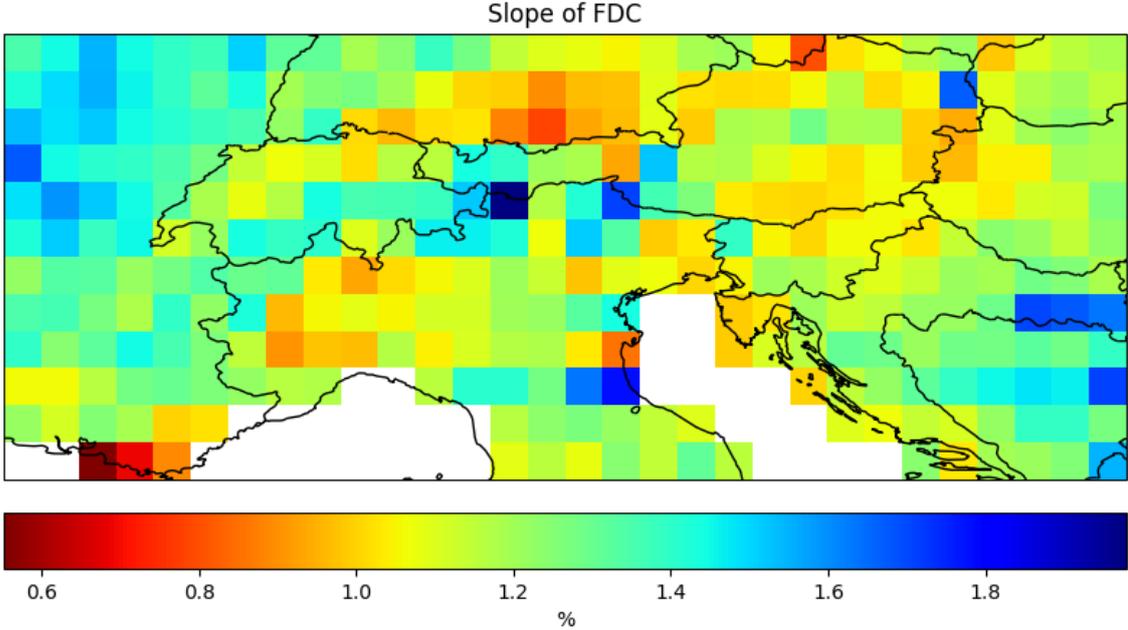


Figure 2.5: *Slope of FDC*

Comments Also the slope of the FDC is an indication of the seasonal variability of runoff, mountainous catchment present FDC curves with a steeper middle section, coherent with the high variability of runoff in snow dominated catchments.

Normalized High flows The 5th quantile (Q_5) divided by the mean daily value.

$$Q_{high} = \frac{Q_{0.5}}{Q_m} \quad (2.5)$$

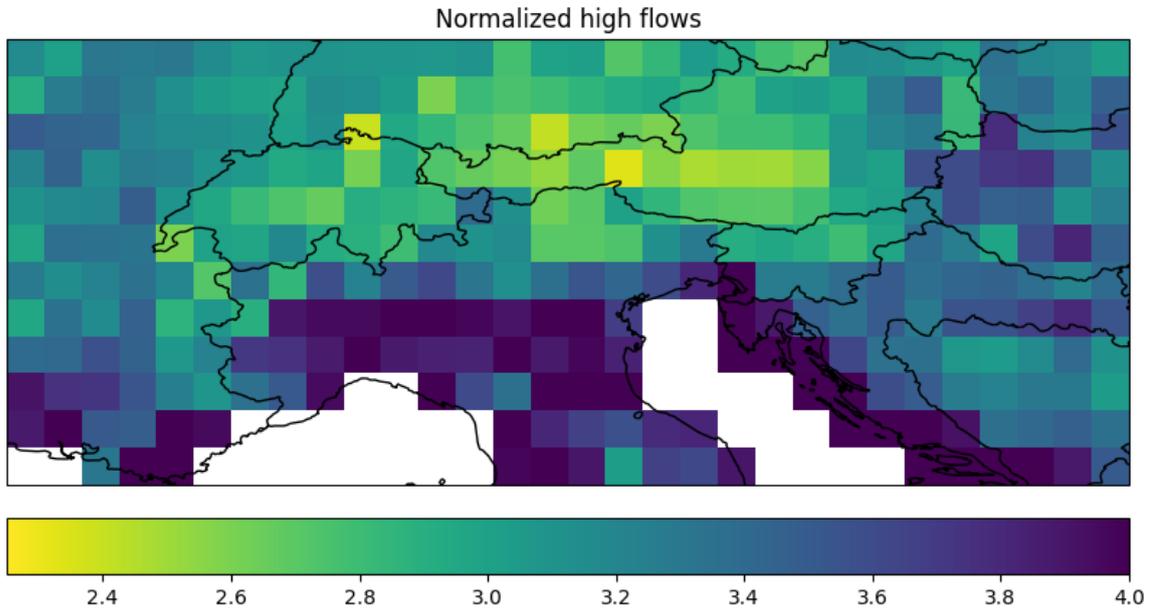


Figure 2.6: *Normalized High Flows*

Normalized Low flows The 95th quantile (Q_{95}) divided by the mean daily value.

$$Q_{low} = \frac{Q_{0.95}}{Q_m} \quad (2.6)$$

Comments Also these signatures showcase a different aspect of the long term FDC therefore an indication of seasonality. Figure 2.4 shows how high are the extreme high values compared to the average conditions. On the contrary figure 2.5 concentrates on the lower end tail of the FDC indicating how low compared to the average are the extreme low flows. Once again we see a pattern in the northern area indicating a more stable flow generation throughout the year, compared to mountainous regions. In this case also the southern regions show case a particularly high deviation of the extremes in respect to average conditions, especially in terms of high flows.

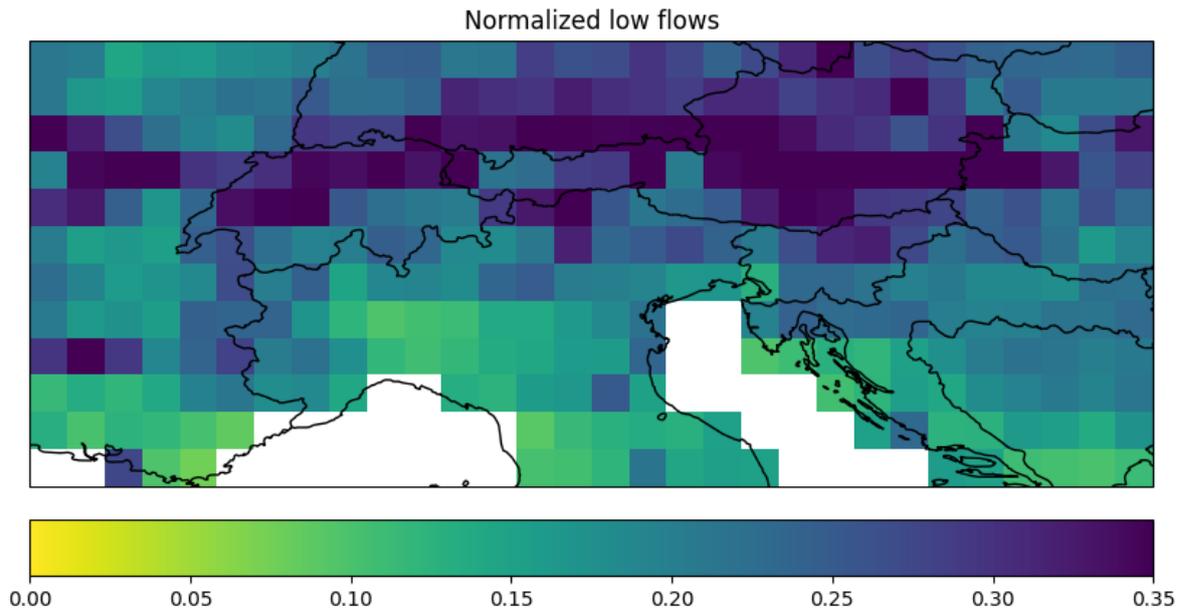


Figure 2.7: *Normalized Low Flows*

2.3 GRDC data for the alpine region

The dataset that we use for comparison was downloaded by the Global Runoff Data Center (GRDC) which also provides a list of stations included in a monitoring project specific to the alpine region named "adaptalp" (website of GRDC: <https://www.bafg.de>). In total the stations from the adaptalp catalogue are 169. We place a cutoff on missing data arbitrarily at 10% and on stations which do not present information about catchment area and elevation. The last cutoff criteria is for stations with area of the catchment above 5000 km^2 , thus catchments that represent the generated runoff for at least two pixels. At the end of this process 141 stations are left.

The location of the stations and other the distributions of area and elevation are represented in the following image and graphs.

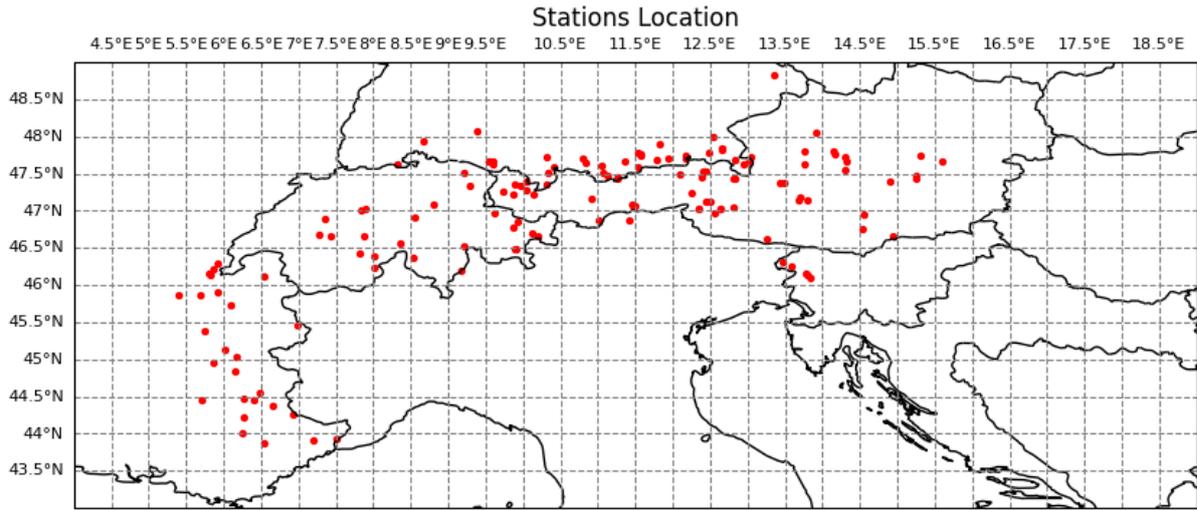


Figure 2.8: Stations' location

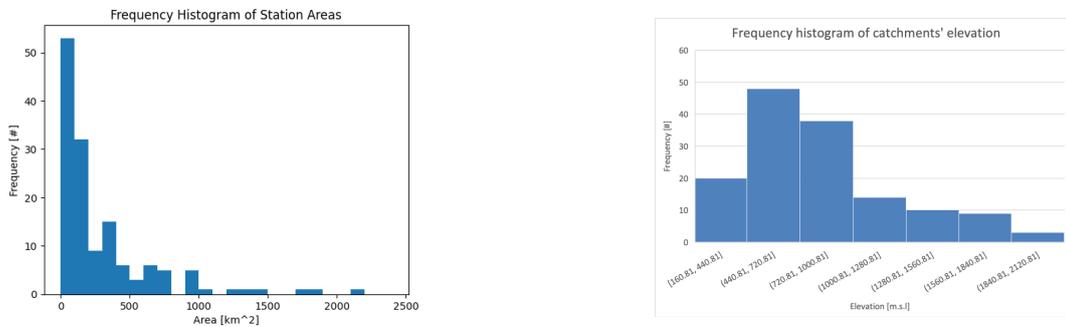


Figure 2.9: Area and Elevation of catchments

2.3.1 From discharge to generated runoff

The data downloaded from the GRDC is discharge data from specific rivers in the closure point of their catchment. Thus all the water that is measured at that section of the river is coming from the relative catchment. In so far as the performance evaluation, we chose to work with the variable generated runoff. Therefore the discharge at each station was divided by the respective area of the catchment. One could argue that this is unnecessary as one of the GHMs' output is simulated discharge. The choice is to compare the generated runoff as the discharge simulated from at each pixel is the sum of all the rivers passing in the pixel itself. For this reason it is very difficult to compare the modeled data to the observed one as the observed data will be but one of the contribution to the total value present in the output of the GHMs. This is particularly troublesome in an area with many small streams and even rivers such as the GAR. We therefore decided in our best judgement to compare the observed generated runoff (calculated as per above) with the generated runoff of the pixel inside which the station is located. This methodology presents other limitations, which we discuss in the next paragraph, but it is independent from the area property.

Main limitation: Representativity of catchments In this study, for the 141 measuring stations only a few catchment's shape file were available. And therefore we opted for comparing each station's data with the pixel within which it falls. This implies assuming a correspondance between modeled and observed data that is only partial. To provide an example, we show the situation for one of the only stations for which the shape file was available in figure 2.8. That is river Var at the station of Malaussane, in France.

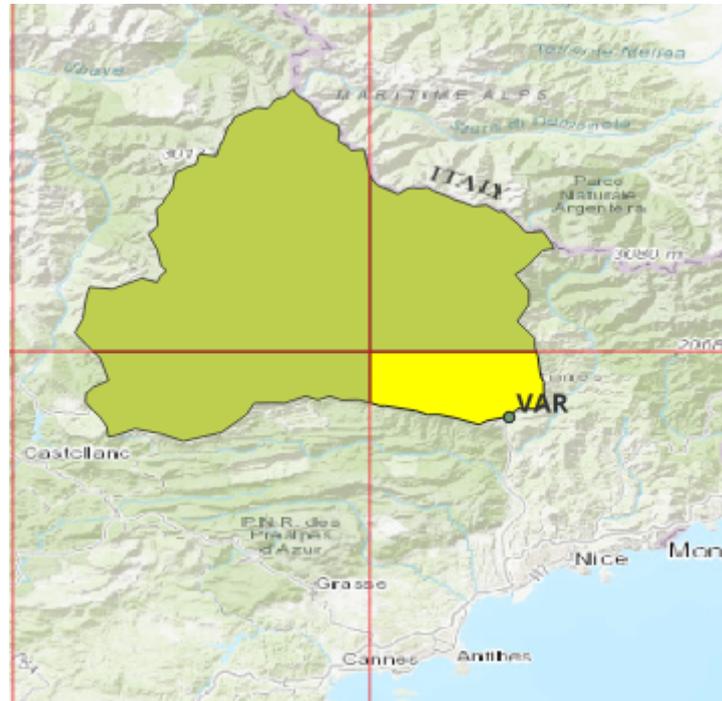


Figure 2.10: *Catchment of river Var at Malaussane*

It is possible to note how, while the measuring station is placed in the pixel in the bottom to the right, the catchment actually falls in 4 different pixels. Correctly and uniquely attributing the generated runoff to one of the 4 pixels given the presented configuration is clearly unfeasible: one would need to multiply the generated runoff by the portion of area that falls in the selected pixel assuming homogeneous runoff generation in the whole catchment, but even then that would represent the contribution to the runoff generation of only that fraction of pixel area (highlighted in yellow), and not of all the other portion of the pixel that falls outside the catchment.

In order to further clarify why it cannot be completely eliminated but must be only accounted for, we present the only two scenarios in which there is perfect match.

1. The catchments have the exact same extension and shape as the pixels and are perfectly centered in their middle point. In this case the generated runoff at the pixel and catchment match 1:1.
2. For each pixel there is 100% observational coverage of every stream of water that passes in it. This is also a completely unrealistic scenario. In fact, measuring river discharge, especially in small and possibly ephemeral streams, is in itself a well known technical issue.

In conclusion, comparing the output of the models with real world data entails accepting a certain degree of mismatch between the two. As explained this is an inherent limiting factor that cannot be completely surpassed and its due to the discretization necessary for the models and their present resolution. Notwithstanding that, it can be accounted for while evaluating the performance of the models as will be explained in the following section.

2.4 Methods

In the previous section, while presenting the dataset, we already showcased the first basic ways of working with data and understand the information that they relay. Given the geolocalized data available it is possible to create different temporal aggregation and provide, as was done, different maps conveying information about long term averages or even inter-annual variability for the whole study area. It is otherwise possible to aggregate the data spatially and inspect the resulting time series, look for trends or step changes. In this section instead we focus on the methods employed in the more quantitative part of the analysis. This is the comparison of the models' output data with the observed data. We also give a brief overview of the statistical tools applied and finally of the softwares used in the analysis.

2.4.1 Performance evaluation metrics

There are many statistical methods to evaluate misfit between observations and modeled data; it is important to remember that no standard procedure exist and each method presents its advantages and disadvantages. One of the founding studies providing a systematic review of the different performance evaluation metrics for hydrological models was provided by Moriasi et al in 2007 [18]. Since then methods have evolved and many authors have provided new methodologies and alternatives for carrying out performance evaluation of models[16] [11]. In this section we therefore give the explanation of the methods chosen along with their functionalities and limitations.

A first coarse distinction is made by Moriasi between graphical and statistical techniques for performance evaluation. In the following paragraph a sample of both will be explained while in the last part the chosen evaluation metrics is discussed in more detail.

Graphical Methods

Graphical methods are the most intuitive and simple evaluation techniques but nevertheless, if coupled with expertise from the model evaluator and in depth knowledge of hydrology, they are a valuable source of information. The most simple kind is plotting the hydrograph of the modeled vs simulated value, which can be already a good indication of how the model is in general reproducing the observations e.g figure 2.9.

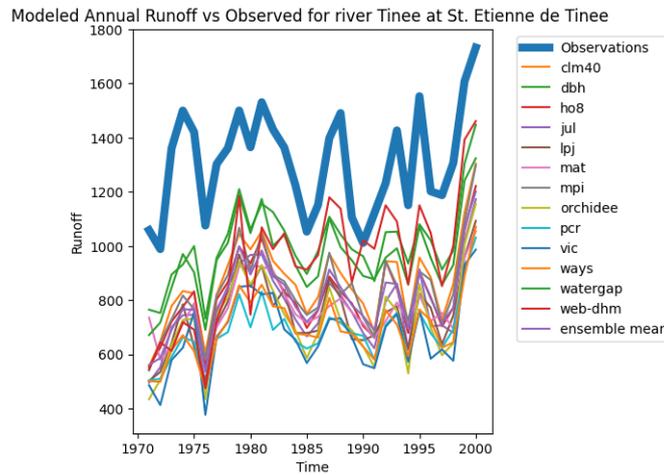


Figure 2.11: Example of plotting the annual hydrograph for graphical inspection

Boxplot or "box and a whisker" plots summarize the information about a distribution by representing in a single line the most significant informations about the distribution. Those informations are:

- Median value: value in the middle of the distribution i.e. 50% of values fall below it and 50% above it
- Q1 or $Q_{0.25}$ the the quantile which is exceeded 25% of the time
- Q3 or $Q_{0.75}$ the quantile which is exceeded 75% of the time
- Interquartile Range (IQR) = $Q1 - Q3$
- Maximum Limit $Q1 + 1.5 * IQR$
- Minimum Limit $Q3 - 1.5 * IQR$
- Outliers, value that fall outside the Minimum and Maximum Limit

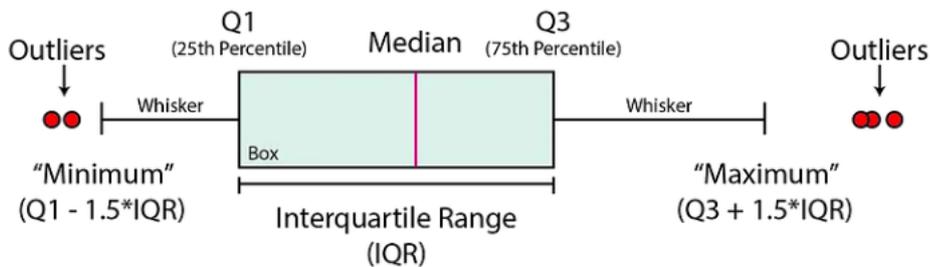


Figure 2.12: Boxplot, source: Towards Data Science

Scatterplots are another way to graphically inspect the data, this time in the form of the relationship between two variables. For instance we can look into the relationship between

elevation and area of the catchment. We might expect from common knowledge to see more extended catchments to be at lower elevation and on the contrary, more elevated catchments to be smaller.

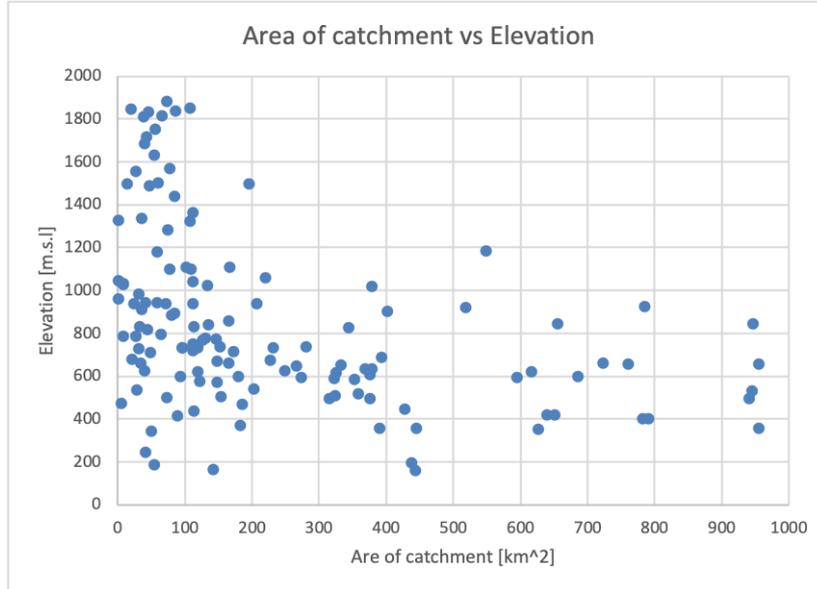


Figure 2.13: *Scatterplot of Area vs Elevation of Catchments*

Statistical evaluation techniques

Statistical evaluation techniques are proposed to provide a numerical, therefore less subjective estimate of the performance of hydrological models. The most common and more widely adopted evaluation metric in hydrology is the Nash Sutcliffe Efficiency [18]. It measures the ratio between the variance of the residuals (noise) versus the variance of the data (information). It ranges from $-\infty$ to 1(included). Value of $NSE = 1$ signifies perfect correspondence between observed and modeled data. Further distinction define value of $NSE > 0$ as acceptable performances and negative values as bad performance. In the context of this study we use a decomposition of the NSE proposed by Kling and Gupta, after whom is named Kling Gupta Efficiency [11].

Kling Gupta Efficiency KGE KGE is a decomposition of the MSE in the form of the Nash Sutcliffe Efficiency. It ranges from $-\infty$ to 1

$$KGE = 1 - \sqrt{s_1 * (1 - r)^2 + s_2 * (1 - \alpha)^2 + s_3 * (1 - \beta)^2} \quad (2.7)$$

Where:

$$r = \frac{\sum_{i=1}^n (q_{obs}(i) - q_{obs,\mu})(q_{sim}(i) - q_{sim,\mu})}{\sqrt{\sum_{i=1}^n (q_{obs}(i) - q_{obs,\mu})^2} \sqrt{\sum_{i=1}^n (q_{sim}(i) - q_{sim,\mu})^2}} \quad (2.8)$$

$$\alpha = \frac{\sigma_{modeled}}{\sigma_{observed}} \quad (2.9)$$

$$\beta = \frac{\mu_{modeled}}{\mu_{observed}} \quad (2.10)$$

The term r , the pearson correlation, indicates the linear correlation between observed and simulated series, the term α about the variability and the term β about the bias. The coefficients s , can be instead used to give different weight to give more emphasis to specific components of the misfit. This can be an interesting perspective, but as of yet it has been scarcely researched, and for simpler interpretation we stick to the most adopted version applying the value of 1 to each of the weight.

Modified Kling Gupta Efficiency KGE' The modified KGE minimizes the effects of cross-correlation between the mean and the standard deviation which is present if there is a bias in the modeled data. We choose this version as we know that there is a bias inherent in precipitation from ERA-40 as mentioned before and as per seen in other studies [26]. The core of the performance evaluation consists in evaluating the modified KGE score for all the selected 141 measuring stations for three different signatures. Those signatures are the one related extreme high Q_5 extreme lows Q_{95} and the mean monthly regime curve. For each of these signatures, each model and each station we have the relative KGE score indicating the misfit between observed and modeled values. The overall results are provided in the Appendices to this thesis.

$$KGE' = 1 - \sqrt{s_1 * (1 - r)^2 + s_2 * (1 - \alpha')^2 + s_3 * (1 - \beta)^2} \quad (2.11)$$

Where:

$$\alpha' = \frac{\sigma_{modeled}}{\sigma_{observed}} * \beta^{-1} \quad (2.12)$$

2.4.2 Software used

The analysis of the datasets was performed using 4 main tools:

1. **Panoply** (<https://www.giss.nasa.gov/tools/panoply/>) is an open source tool provided by nasa that allows for quick visualization of data in the netcdf format. It was used in the beginning, to obtain the first plots and visualize the dataset.
2. **Qgis** was also used in the context of having an interactive map always available.

3. **Python** inside a Jupyter notebook environment was the main tool used for data analysis. The availability of multiple packages such as numpy pandas, matplotlib and more, allowed for ready to use tools for working with large amount of data such as the ones downloaded in the study.
4. **Excel** . The final tables with the output of the performance evaluation where exported in excel. Although python offers an easier and more automated way of working with data; excel was also employed in the final part in order to see the results and begin to interpret them.

The main ones for which the results are shown are Python (especially) and Excel, which are the main tools from which results are obtained. The first two, Panoply and Qgis, where valuable in the context of the fast and readily available visualization of data.

Chapter 3

Results

The section of results is divided into three main sections. First we show the results for specific catchments. We showcase both the performance evaluations results in terms of KGE' scores and plot hydrological signatures in order to look for the specific reasons for bad or good performance at the catchment scale. Secondly we provide the generalized results across all stations and indicators in order to delve deeper into the specifics of model performance on the selected hydrological signatures.

3.1 Results for selected representative catchments

We provide results for selected catchments. Through these case studies we illustrate, even more than the general background given above, the complexities of hydrological modelling of the GAR. With this closer look into some specific and relevant basins of the region we aim to further exemplify the general theme, already highlighted above of the extreme heterogeneity and thus the inherent limits of adopting global models on the selected scale. First we start with the basin of the Soca River, located in the climatic subregion of the South East at the border between Slovenia and Italy. Secondly we present the case study of river Rhone at his spawning point in the homonymous glacier. Then the river Durance, one of Rhone's main affluents, at the artificial lake Serre Poncon. Lastly we present results for river Lech in Lechbruck. In this way we cover all the 4 subregions of the GAR and presents specific results in the context of the various climatologic features that characterize the region

3.1.1 Border of Slovenia and Italy: the River Soca

The Soca River is the first for which we present the results. The river flows from slovenia where it spawns at 1100 meters of elevation in the protected area of mount Triglav. It spans for a total of 136 km out of which 95 are in Slovenia and 42 in Italy where it ends in the Adriatic Sea. The upper basin, is among the wettest area of the whole European continent with yearly precipitation reaching up to 3500 mm/year. The southern part near the outlet features a sub-Mediterranean climate with droughts in the summer months that are not so uncommon. The water of the Soca river is used both by the Slovenian, which have build

reservoirs and hydropower plants in the upper part of the basin, and the Italians which rely heavily on the river both for the energy and agricultural sectors; figure . shows the results of the signatures for the Soca river at the outlet in Kobarid (Caporetto).

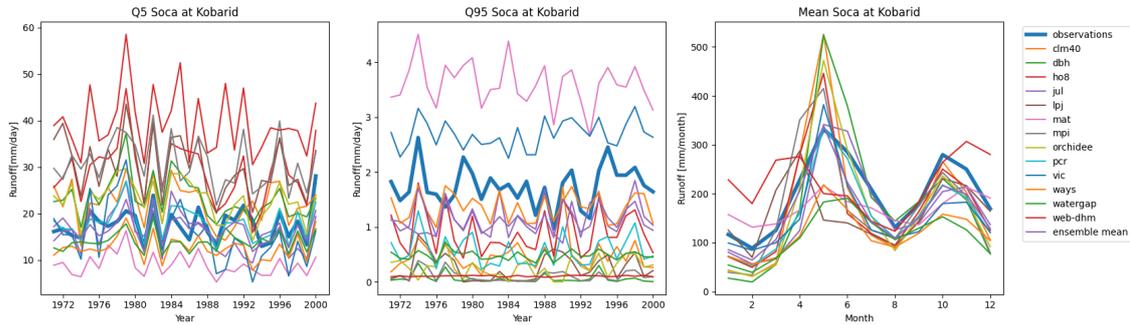


Figure 3.1: *Soca River at Kobarid: results for the high , low flows and monthly regime curve*

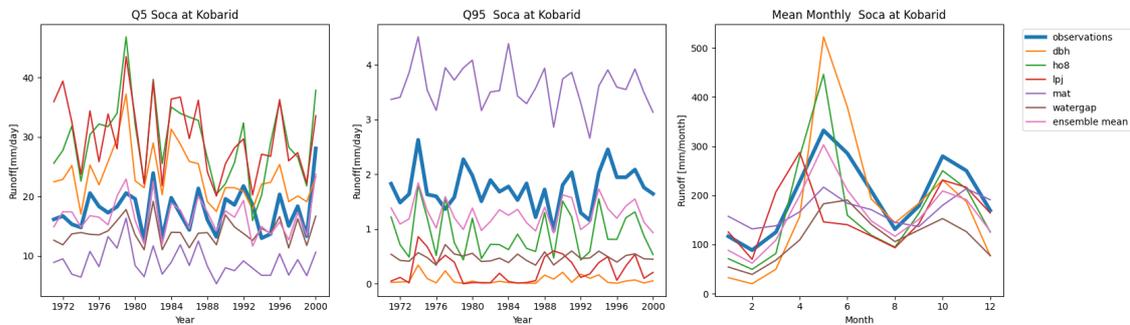


Figure 3.2: *Soca River at Kobarid: results for the high , low flows and monthly regime curve*

It is possible to notice on a first graphical inspection the good results for the model in matching the monthly regime curve for this particular catchment. In particular this catchment show the typical features of the pre-Alpine catchments in presenting two maxima: one in the spring and one in the autumn season; the models mostly follow the timing of the events correctly, only model that anticipates the spring peak and overshoots the autumn peak into almost winter is model web-dhm. Regarding the high flows, the first graphical inspection shows a tendency on the overestimation for this specific catchment, but still a more even spread around the observed value in comparison with the low flows; indeed, the low flows show a clear tendency for most models toward the underestimation with only models VIC and MATSIRO showing a significant overestimation of such values. Table . shows the statistical evaluation results in terms of the signatures and in both the experiments with human impacts parametrized (varsoc) and not parametrized (nosoc).

River Soca at Kobarid		Catchment area (km ²) = 437 Catchment elevation (m) = 194													
Models		clm4.0	dbh	h08	jules	lpjm	matsiro	mpi	orchidee	pcr	vic	ways	watergap2	web-dhm	ensemble
Signatures															
KGE' (nosoc)	Q.5	0.36	0.31	0.22	0.28	0.17	0.30	0.17	0.34	0.71	-0.15	0.63	0.53	-0.13	0.70
	Q.95	-1.20	-4.78	-0.32	-0.05	-2.66	-0.21	-2.73	-1.59	-0.52	-0.06	0.19	0.02	-0.47	0.41
	Monthly	-0.09	-0.09	0.27	0.53	0.39	0.39	0.21	0.04	0.61	0.57	0.59	0.56	0.13	0.73
KGE'(varsoc)	Q.5	NA	0.31	0.22	NA	0.17	0.30	NA	NA	NA	NA	NA	0.53	NA	0.78
	Q.95	NA	-3.67	-0.32	NA	-2.66	-0.21	NA	NA	NA	NA	NA	0.02	NA	0.48
	Monthly	NA	-0.10	0.27	NA	0.39	0.39	NA	NA	NA	NA	NA	0.56	NA	0.78

Table 3.1: KGE Score for Soca catchment

Looking at table 3.1, its possible to see the results o all models for the KGE scores of this specific catchment. The most meaningful result is the appropriate performance of the ensemble mean across all indicators: Monthly Regime KGE' = 0.73, Q₉₅ = 0.41, Q₅ = 0.70. The model DBH expresses the greatest variability going from a decent 0.31 KGE' on the high flows, to a markedly bad fit for the low flows (KGE' = -4.78). The other interesting factor to note is that the models do not show any significant changes while passing from the socio economical scenario no soc(no human impact parametrized) to the scenario varsoc. The only models that shows any change in the results of the KGE' score is once again model DBH which shows a slight decrease in performance on the monthly regime signature (-0.01) but a significant increase in performance for the low flows (+1.11).

3.1.2 Example of catchment in the high Alps: The spawn of river Rhone at Gletsch

We have already presented in chapter 1 the Rhone river as one of the major rivers spawning in the region and in general one of the most important in the Alpine area. We now show the results for the river in its spawning point in the High Alps in Switzerland. The catchment has an area of 38 km² and a mean elevation of 1810 m.a.s.l

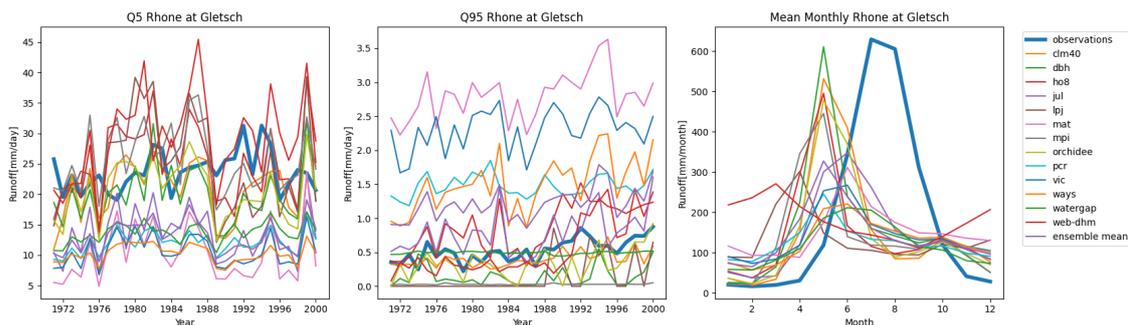


Figure 3.3: Results for the river Rhone at its spawn in Gletsch: results for the high , low flows and monthly regime curve

The most interesting feature that appears from the graphs is the fact that, although models match to a certain degree the shape of the monthly regime curve. They fail to reproduce the correct timing and also amount of water that is actually generated in the catchment. Looking at the last graph on the right in figure 3.4 we can see how the observation curve depicts a significantly larger amount of water compared to all the models. Also all

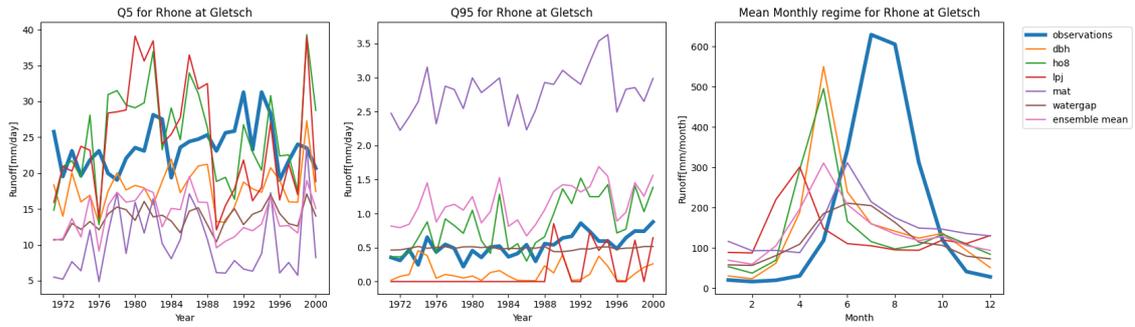


Figure 3.4: Results for the river Rhone at its spawn in Gletsch: results for the high , low flows and monthly regime curve

the models depict the spring peak of runoff given by snow melt way in advance compared to the true values. As far as the high and low flows are concerned we spot a different, and also opposite, result compared to the previous catchment. The high flows, that were previously mostly overestimated, are here mostly underestimated by most models. On the contrary, linking to the previous results, low flows are mostly overestimated the ensemble. Looking at the results in terms of KGE' score on the different indicators, we can see that, while the ensemble mean is not performing worse than all the other models, it is not the best performing model. In this case models JULES, WaterGAP and WAYS, outperform the ensemble mean on most indicators. Once again model DBH is the only one, which shows a marked distinction between the performance in the varsoc and nosoc scenarios, improving its KGE scores on the monthly regime ($KGE_{Monthly,varsoc} = 0.11$, $KGE_{Monthly,nosoc} = 0.09$) on the high flows (+0.10) and on the low flows (+0.16).

Rhone at Gletsch		Catchment area (km ²) = 38 Catchment elevation (m)= 1810													
Models		clm4.0	dbh	h08	jules	lpjm	matsiro	mpi	orchidee	per	vic	ways	watergap2	web-dhm	ensemble
Signatures															
KGE' (nosoc)	Q_5	-0.16	-0.12	-0.33	-0.59	-0.65	-1.56	-0.13	-0.15	0.00	-0.30	-0.17	0.04	-0.19	-0.14
	Q_95	0.51	-1.47	0.22	0.32	-4.03	-3.36	-0.16	-0.26	-1.08	-2.48	-0.91	-0.36	-0.35	-0.41
	Monthly	0.12	0.09	-0.12	0.36	-0.52	0.23	-0.09	0.09	-0.11	0.12	0.17	0.23	-0.88	0.00
KGE'(varsoc)	Q_5	NA	-0.02	-0.33	NA	-0.65	-1.56	NA	NA	NA	NA	NA	0.04	NA	-0.21
	Q_95	NA	-1.31	0.22	NA	-4.03	-3.36	NA	NA	NA	NA	NA	-0.35	NA	-0.24
	Monthly	NA	0.11	-0.12	NA	-0.52	0.23	NA	NA	NA	NA	NA	0.25	NA	-0.01

Table 3.2: KGE Score for Rhone at Gletsch

3.1.3 Catchment affected by human impacts: Durance at Serre Poncon

The Durance is one of the main affluent of the Rhone. We select to showcase the results as it is in the interesting closing station at the outlet of lake Serre Poncon, which is one of the most prominent artificial reservoir used both for hydropower and agriculture in France. We want to assess the difference in results of the signature in the two scenarios. The catchment area is 3500 km² while its mean elevation is 704 m.a.s.l.

The first thing to note is that the catchment area is now around 50% bigger than the pixel area (3400 catchment area vs 2400 pixel area). We can have a look at the map and

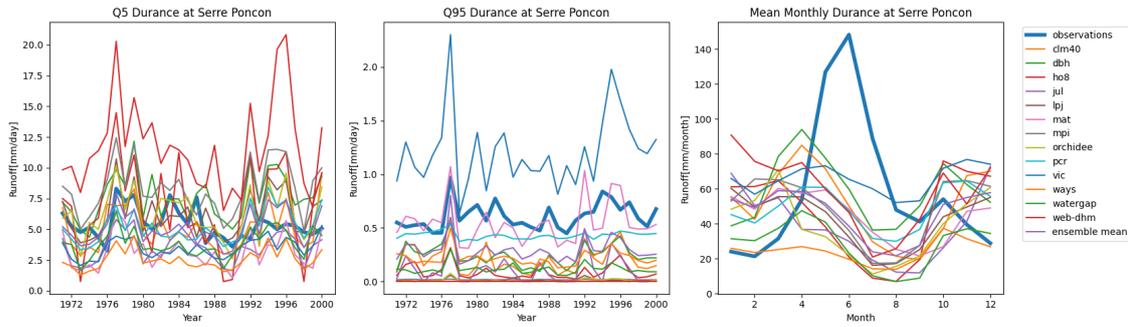


Figure 3.5: Results for the river Durance at the reservoir of lake Serre Poncon

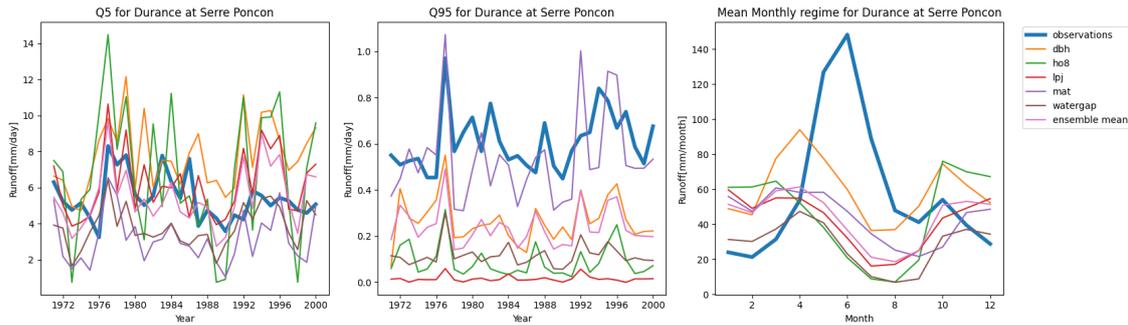


Figure 3.6: Results for the river Durance at the reservoir of lake Serre Poncon

see whether at the lake falls into the correct pixel (figure 3.7).

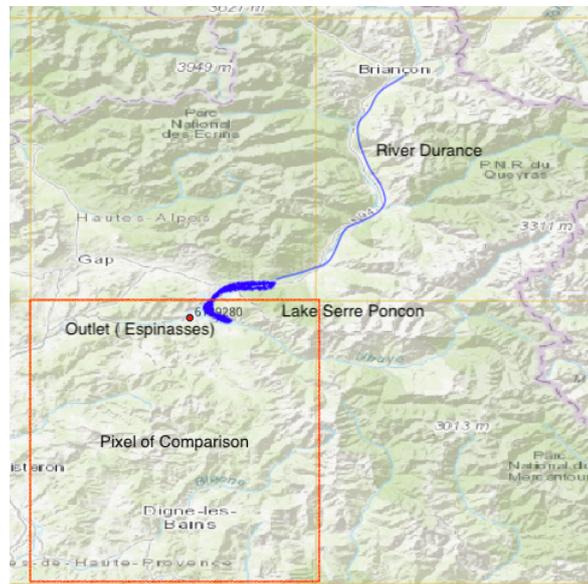


Figure 3.7: We can see that the outlet station falls into a pixel that is more likely not representative of the catchment

We notice how the correspondence between pixel and river and lake is less than ideal, giving already a very plausible explanation for the poor results of this specific catchment. Indeed if we choose to compare with the pixel above and to the right, which apparently entails a larger portion of the catchment and river, as well as the lake itself, we can see a difference in results and a better match in the monthly regime curves (figure 3.8).

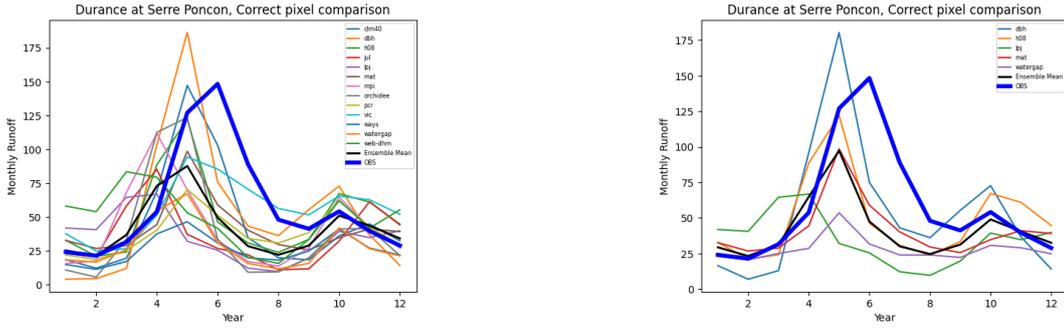


Figure 3.8: *Monthly regime curve with the correct pixel*

The results for the KGE' score given the change in pixel for the ensemble mean are the following:

- $KGE'_{Monthly,nosoc}=0.28$; $KGE'_{Monthly,varsoc}=0.31$
- $KGE'_{high,nosoc}=0.55$; $KGE'_{high,varsoc}=0.43$
- $KGE'_{low,nosoc}=0.034$; $KGE'_{low,varsoc}=0.39$

We can compare the results with the table below with the results for the standard methodology and see that the KGE' score for the ensemble mean in both experiments is drastically lower. Indicating that in this specific case, the pixel within which the outlet station fell, was not the appropriate comparison data for the evaluation. Once again we can comment on the graphs obtained and verify the same pattern of intra-annual runoff regime, with the peak of spring snowmelt that is anticipated by most models.

Durance at Espinasses		Catchment area (km ²) = 3580 Catchment elevation (m) = 704													
		Models	clm4.0	dbh	h08	jules	lpm	matsiro	mpi	orchidee	pcr	vic	ways	watgap2	web-dhm
		Signatures													
KGE' (nosoc)	Q.5	0.23	0.20	-0.43	0.24	0.31	-0.11	0.14	0.33	0.13	0.25	-0.12	0.04	-0.39	0.31
	Q.95	-1.43	-0.30	-2.02	-5.46	-2.57	0.06	-0.26	-0.26	-5.28	-0.10	0.10	-0.34	nan	0.05
	Monthly	-0.06	0.03	-0.54	-0.47	-0.48	-0.10	-0.47	-0.63	-0.15	-0.19	-0.54	-0.24	-0.49	-0.41
KGE'(varsoc)	Q.5	NA	0.14	-0.43	NA	0.31	-0.09	NA	NA	NA	NA	NA	0.19	NA	0.27
	Q.95	NA	-0.13	-1.96	NA	-2.45	0.05	NA	NA	NA	NA	NA	-0.32	NA	-0.10
	Monthly	NA	-0.02	-0.54	NA	-0.48	-0.11	NA	NA	NA	NA	NA	-0.24	NA	-0.35

Table 3.3: *KGE Score for Durance at Espinasses*

3.2 Performance Evaluation of "NOSOC" model ensemble

We now generalize the results obtained with boxplots and scatterplots showing the performance of the ensemble over all the stations. First we present results for naturalized conditions, no human impacts on the hydrology of the catchments.

First the median KGE' for all models over all stations, the following plot shows the results of all the three selected signatures:

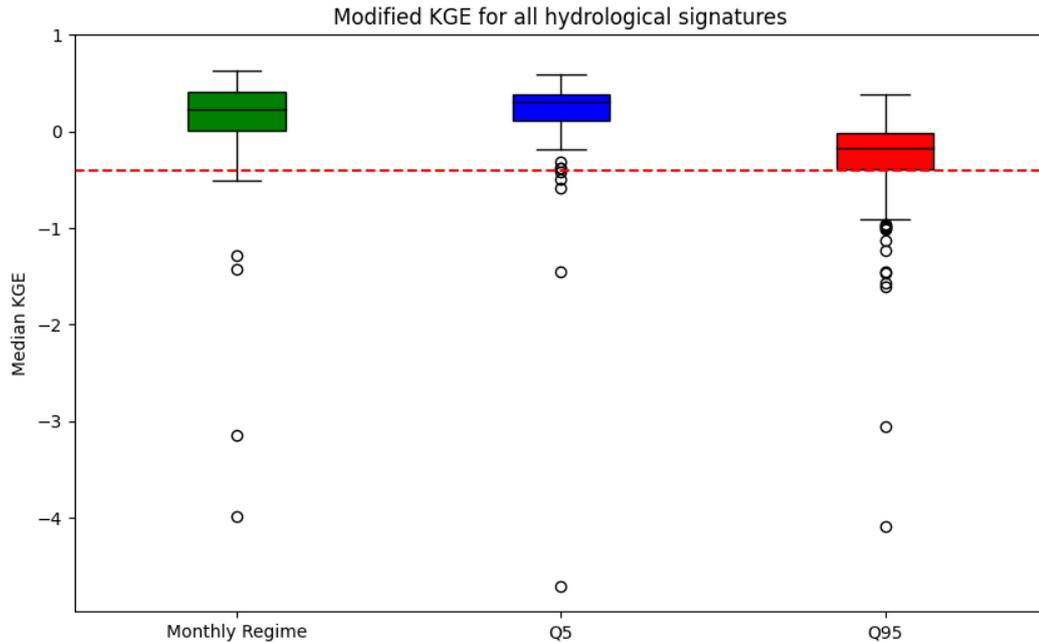


Figure 3.9: Results for the three signatures, nosoc

We can see how the indicator of the low flows reaches over the median performance across the stations the highest performance score on KGE'. It is closely followed by the monthly regime indicator and, quite well below follows the performance over the low flows. We can still appreciate how the box for all the three signatures fall above the line of $1 - \sqrt{2}$ which indicates a better estimating power than the mean of the observation over that signature (for at least 50% of stations). The outliers in the negative sense indicate stations for which the representativeness of the pixel is limited at best; As we have seen in the case of river Durance in Serre Poncon.

Results for indicators over the different models Secondly we expand on the performance over each of the three specific signatures across all stations and models, we start by showcasing the KGE' scoring over all stations for all the 14 models participating in the nosoc experiment.

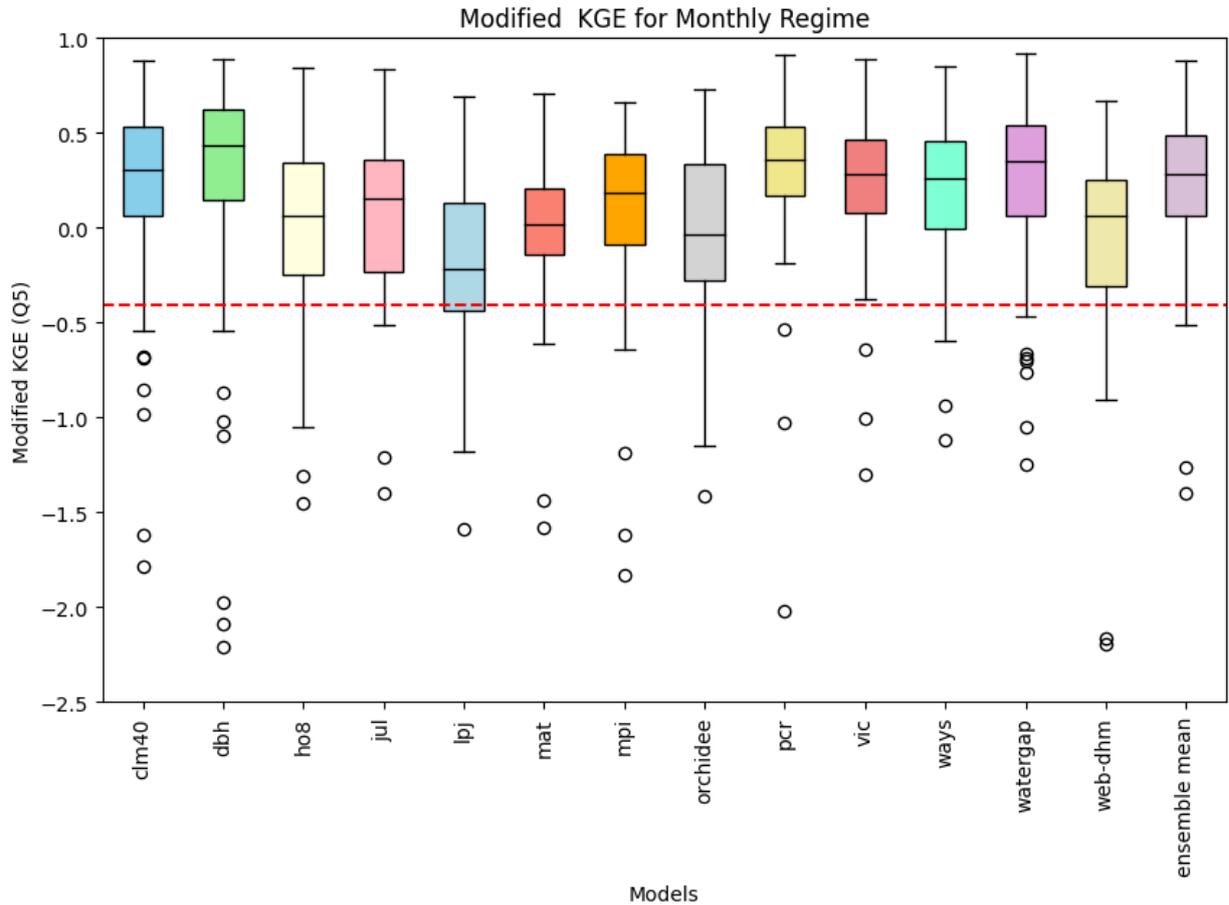
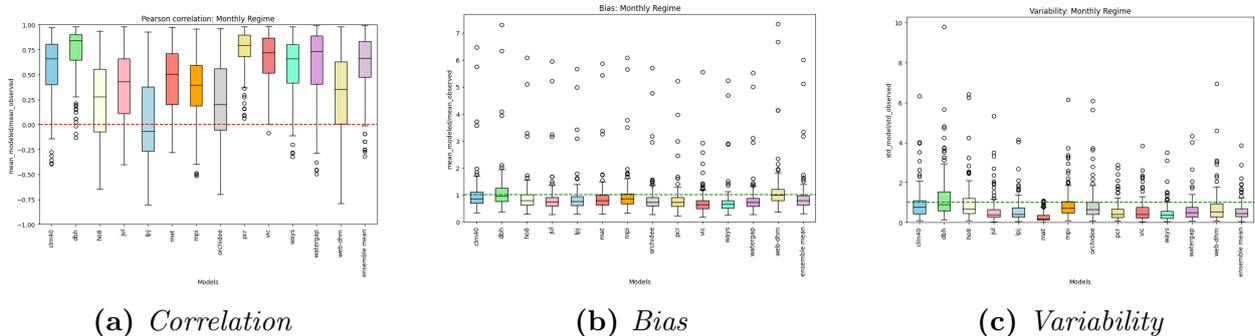


Figure 3.10: Results for Monthly regime, nosoc



(a) Correlation

(b) Bias

(c) Variability

Figure 3.11: Results for the separate subcomponents of the KGE for the monthly regime

Overall the performance of the model ensemble is quite satisfactory on the monthly regime signature. The singular models which performs worse than all the others is model LPJ(median of KGE' over all stations = -0.22), followed by Orchidee(-0.04) and Mat-siro(0.02). The three best performing models are instead models WaterGAP(0.35) , PCR(0.36)

and best overall DBH (median of KGE' over all stations = 0.43). The boxplot of the sub-components show a marked difference in the result for the correlation with the observed data across models, indeed the best performing model (DBH) is also the one which shows a high degree of correlation with the observed data. Thus implying a better representation of the seasonality compared to the other models. The results for bias and variability show instead a tendency on the underestimation for both aspects on this signature across all models. It is interesting to note how the model web-dhm which has the error bias term smaller than most models still ranks as one of the worse models due to its poor representation of intra-annual processes. The ensemble mean ranks 5th on the monthly regime KGE' scores showing still a positive median score of 0.28.

Q₅

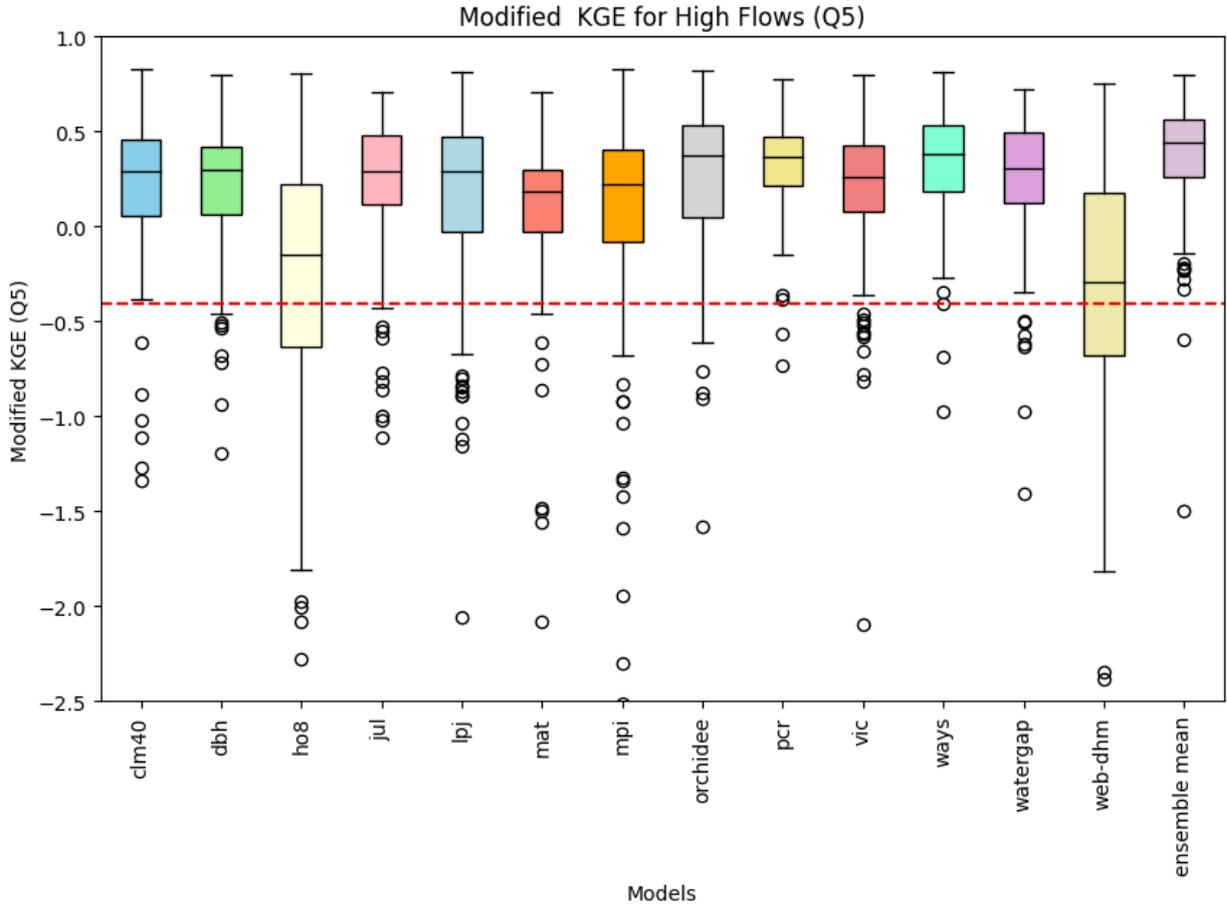
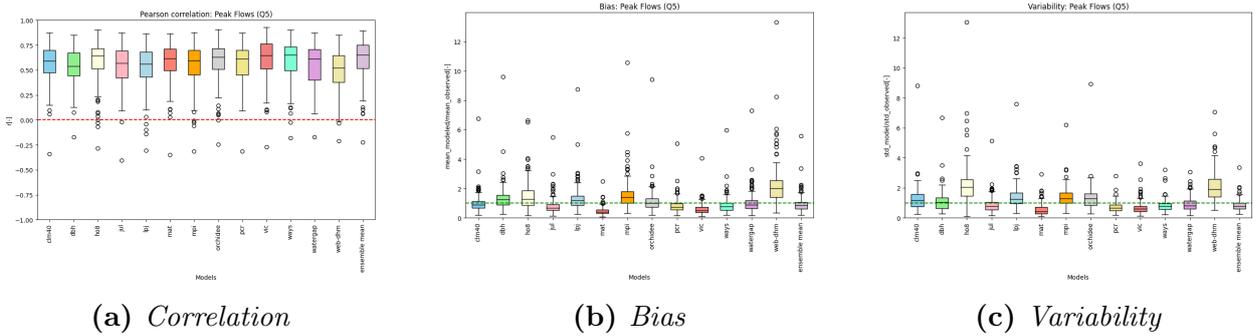


Figure 3.12: Results for Q₅, nosoc



(a) Correlation

(b) Bias

(c) Variability

Figure 3.13: Results for the separate subcomponents of the KGE for Q₅

The high flows signature is, among the selected three, the one over which the models ensemble obtains the highest KGE' scores on average. We can see that in the results of figure 3.12 how most models fall well above the threshold of "poor performance". The three worse performing models over this signature are models web-dhm (KGE' = -0.29), h08 (-0.15) and Matsiro (0.18). The best performing ones are instead Orchidee (0.37) ways (0.38) and the ensemble mean with a median performance across stations of (0.44). In terms of the components we notice how the correlation terms is in general good and also doesnt vary much across models, which is a potential warning sign which we will consider

in the discussion section. The other two terms instead spread between over estimation and underestimation of the runoff. Although the best performing models are characterized by an underestimation while the one performing worse are overestimating the high flows (e.g. h08 and web-dhm, see figure 3.13 b).

Q₉₅

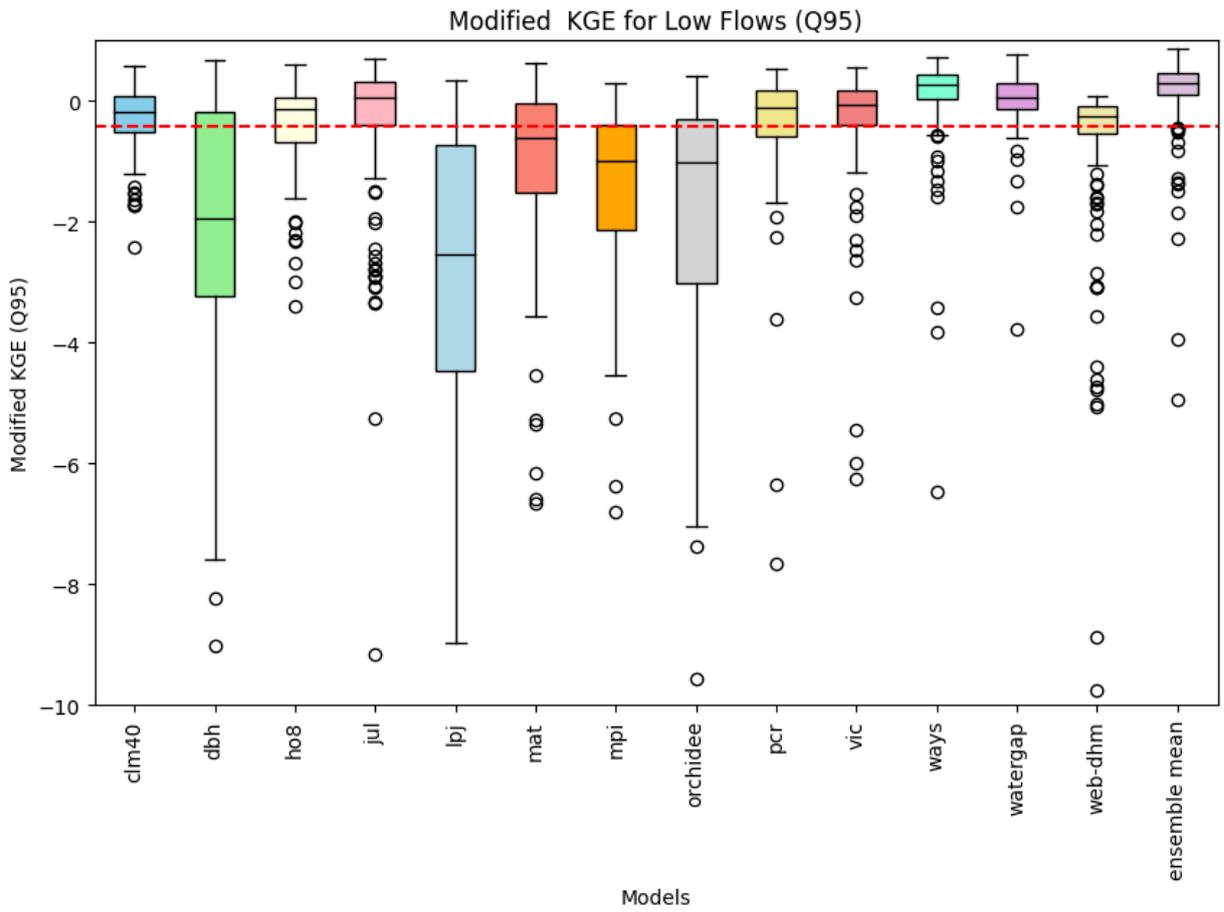
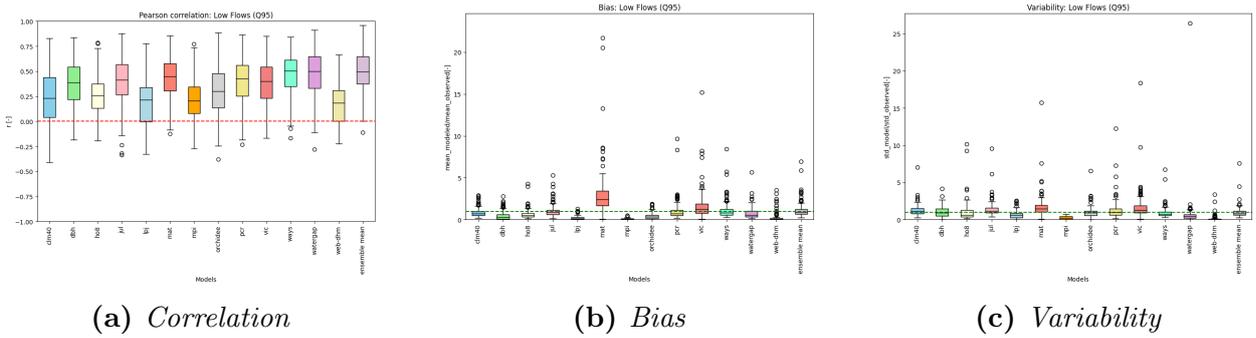


Figure 3.14: Results for Q₉₅, nosoc



(a) Correlation

(b) Bias

(c) Variability

Figure 3.15: Results for the separate subcomponents of the KGE for Q₉₅

3.3 Performance Evaluation of "VARSOC" model ensemble

Results considering human impacts on the hydrology of the catchments:

Likewise, boxplot of all indicators:

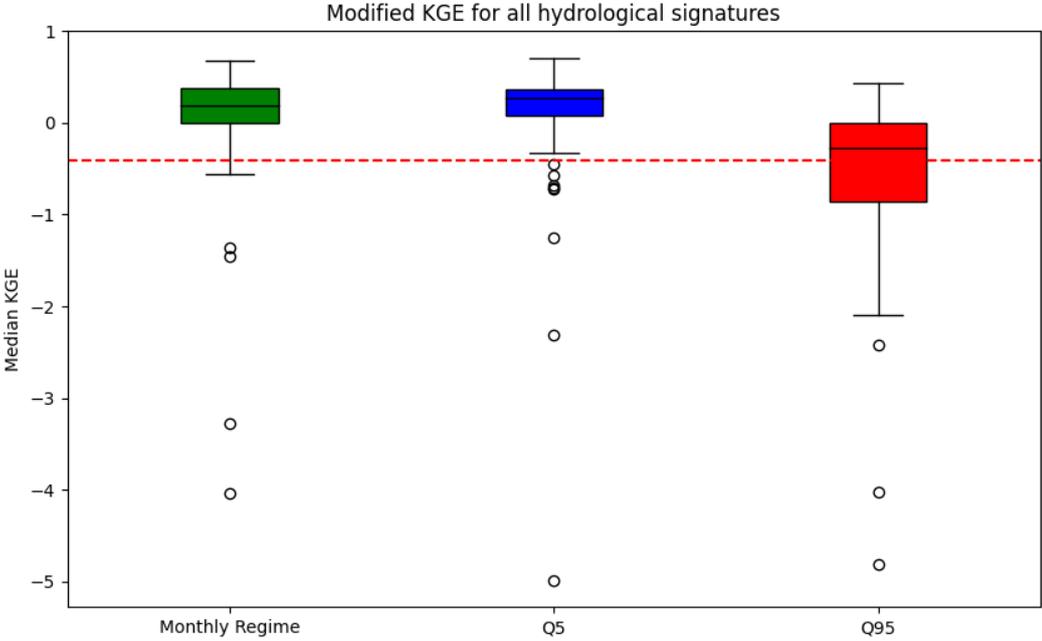


Figure 3.16: Results for the three signatures, varsoc

Monthly regime Secondly the boxplot for each model over all station on every indicator separately . We start with monthly regime:

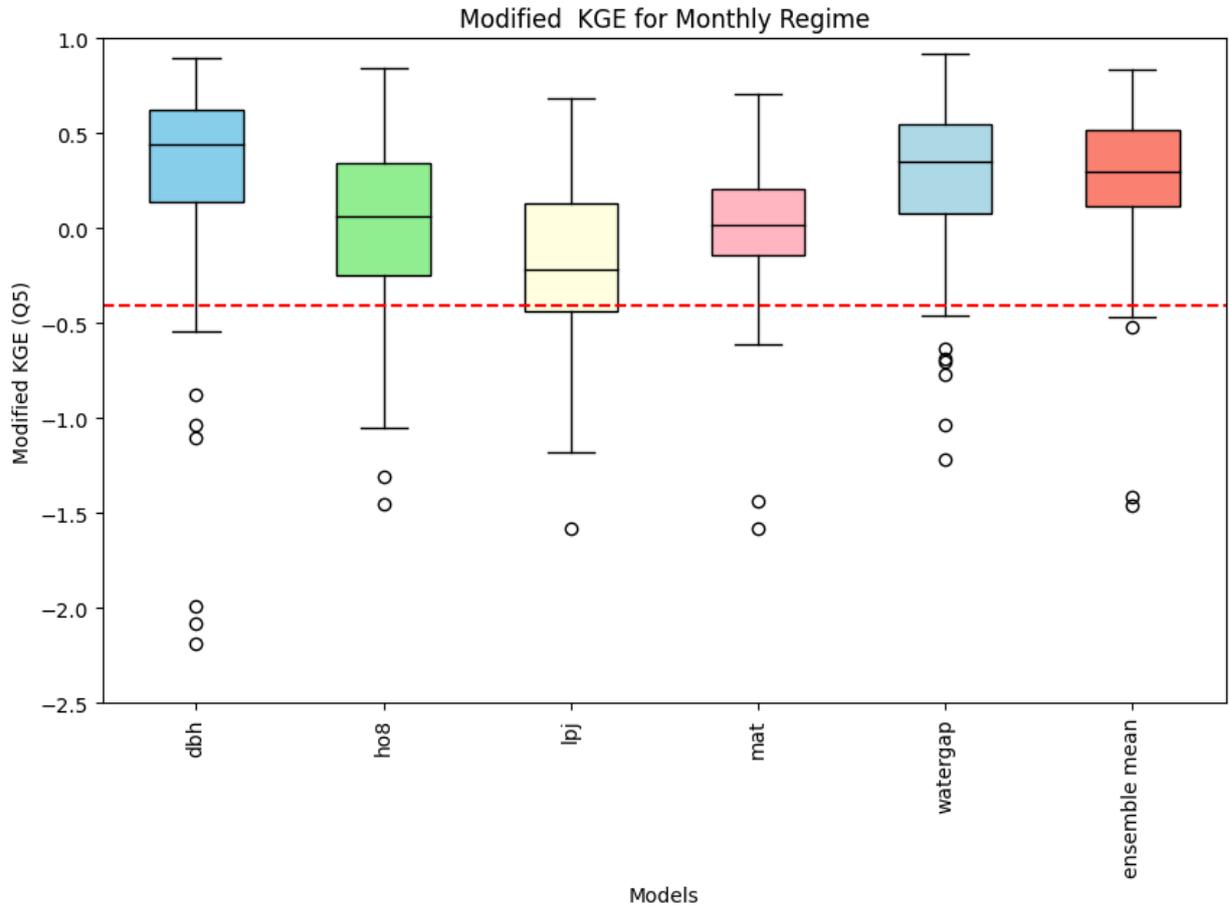


Figure 3.17: Results for Monthly regime, varsoc

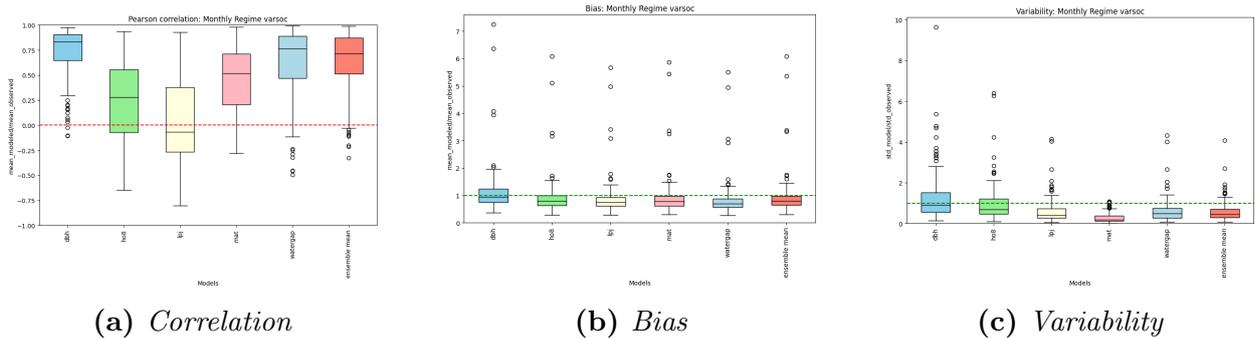


Figure 3.18: Results for the separate subcomponents of the KGE for Monthly regime in the varsoc experiment

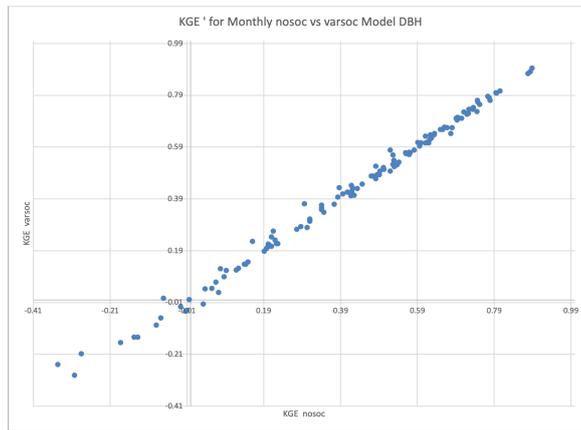


Figure 3.19: *For one of the best performing model we show the difference in the representation of the monthly regime's KGE' in the different scenarios, in this case we can see that while moving from the worse to best performance the good indicator of KGE are more towards the varsoc experiment, although its not solid evidence we can still detect a slight graphical trend*

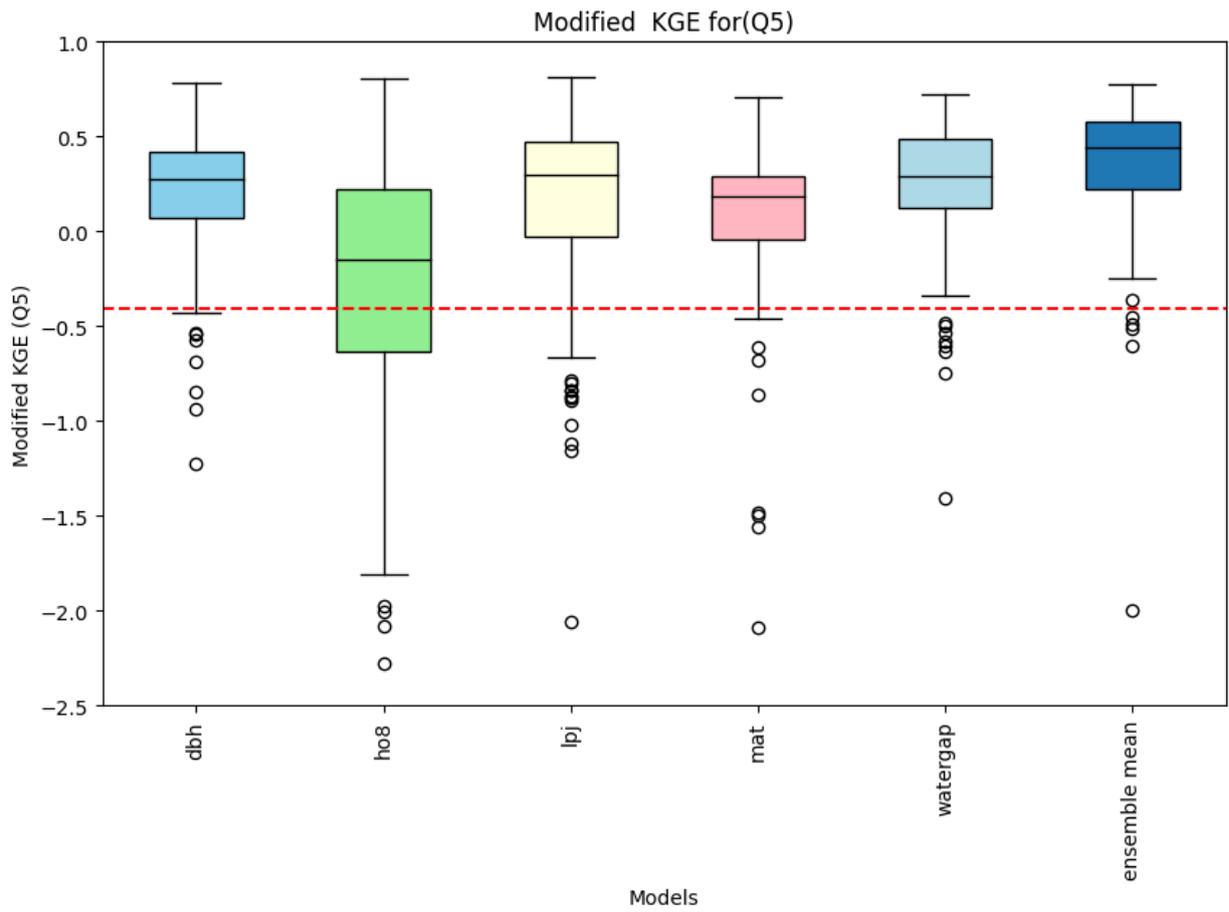


Figure 3.20: Results for Q_5 , varsoc

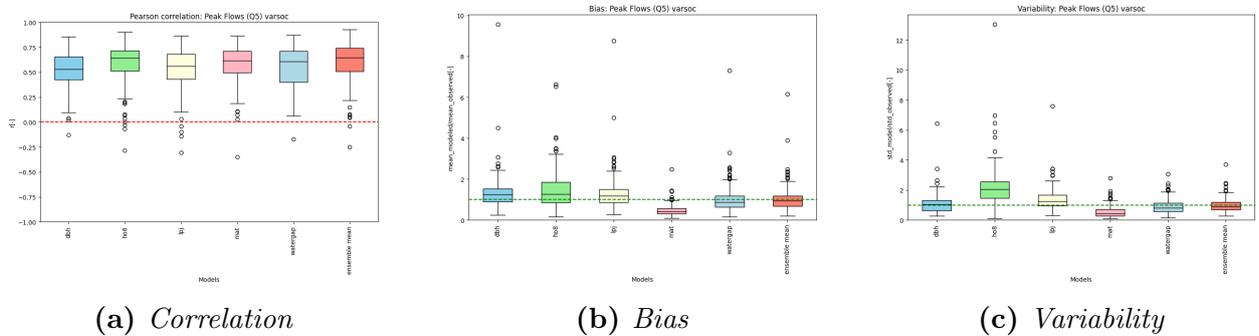


Figure 3.21: Results for the separate subcomponents of the KGE for Q_5 in the varsoc experiment

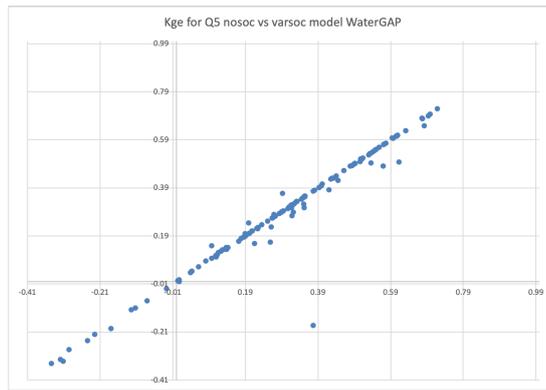


Figure 3.22: For one of the best performing model we show the difference in the representation of the high flows in the different scenarios, slim in terms of the selected KGE' metrics

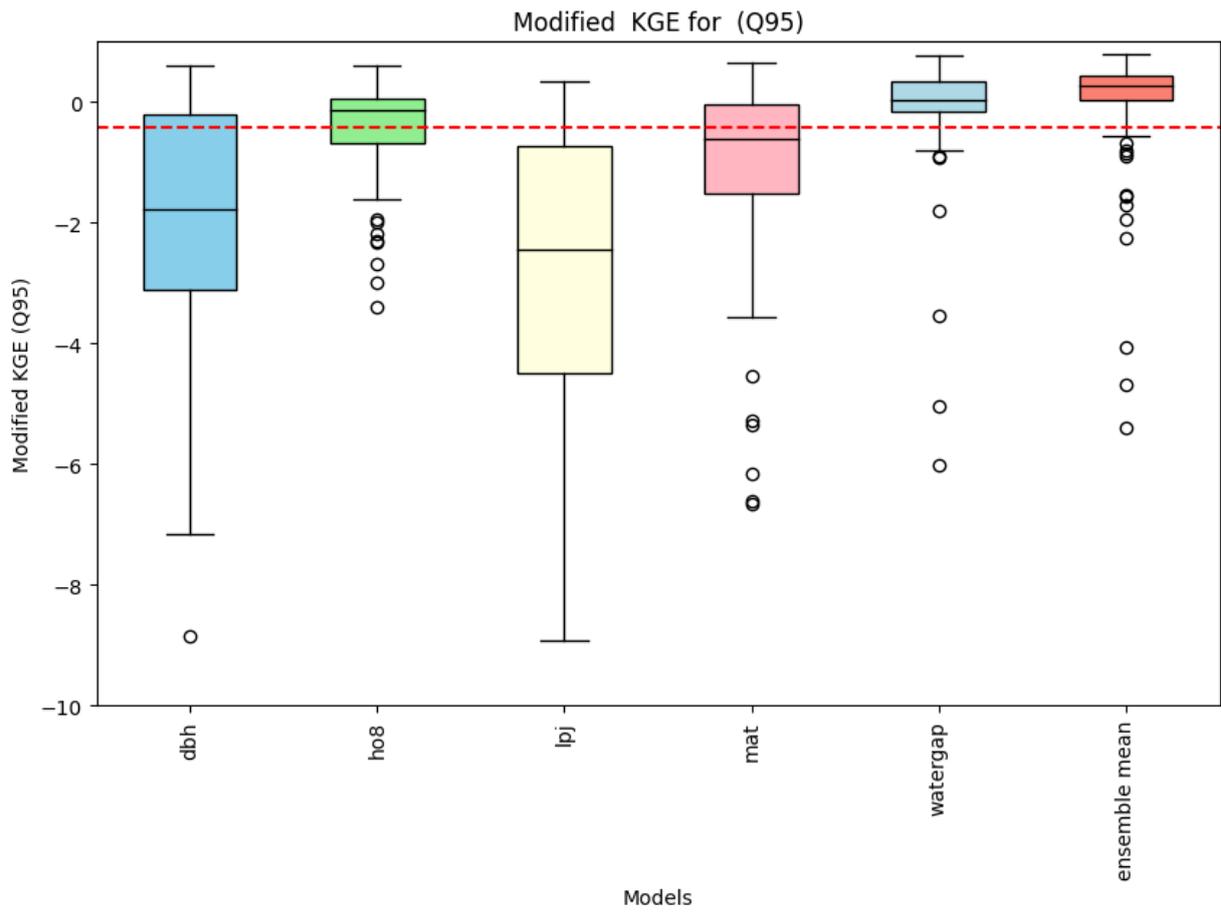


Figure 3.23: Results for Q_{95} , varsoc

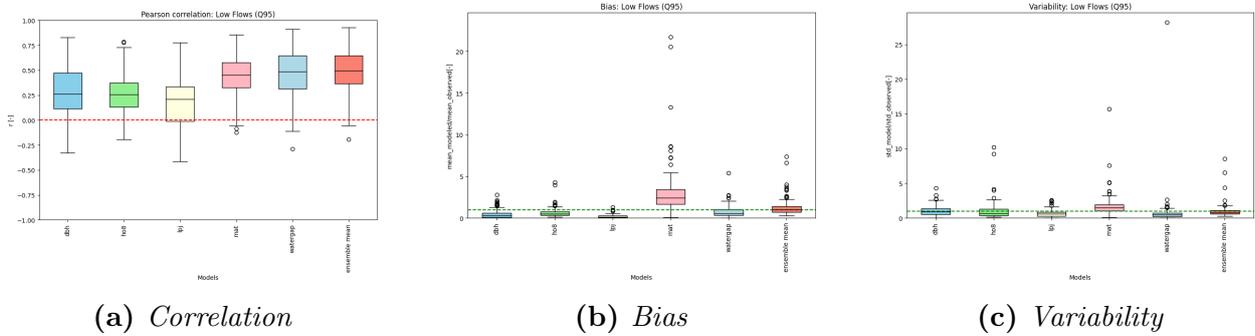


Figure 3.24: Results for the separate subcomponents of the KGE for Q_{95} in the varsoc experiment

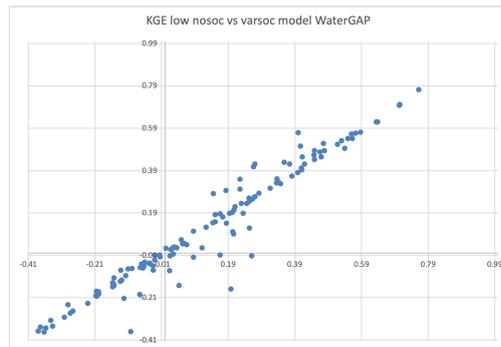


Figure 3.25: *For one of the best performing model we show the difference in the representation of the low flows in the different scenarios, in this case we can see that while moving from the worse to best performance the good indicator of KGE are more towards the varsoc experiment, although its not solid evidence we can still detect a slight graphical trend*

3.3.1 Summary table of results

Model	Monthly regime		Q5		Q95		Rank(based on nosoc)	Rank(based on nosoc)	
	KGE'(median of all stations)		KGE'(median of all stations)		KGE'(median of all stations)				
	Nosoc	Varsoc	Nosoc	Varsoc	Nosoc	Varsoc			
clm40	0.31	NA	0.29	NA	0.19	NA	4	6	8
dbh	0.43	0.44	0.30	0.27	-1.94	-1.78	1	8	13
h08	0.06	0.06	-0.15	-0.15	-0.14	-0.14	11	13	7
jules	0.16	NA	0.29	NA	0.07	NA	9	9	3
lpj	-0.22	-0.22	0.29	0.30	-2.63	-2.63	14	7	14
matsiro	0.02	0.02	0.19	0.19	-0.62	-0.62	12	12	10
mpi	0.19	NA	0.22	NA	-0.99	NA	8	11	11
orchidee	-0.04	NA	0.37	NA	-1.01	NA	13	3	12
pcr	0.36	NA	0.37	NA	-0.10	NA	2	4	6
vic	0.28	NA	0.26	NA	-0.06	NA	6	10	5
ways	0.26	NA	0.38	NA	0.26	NA	7	2	2
watergap	0.35	0.35	0.31	0.29	0.06	0.03	3	5	4
web	0.07	NA	-0.29	NA	-0.34	NA	10	14	9
ens. mean	0.28	0.30	0.44	0.44	0.29	0.28	5	1	1

3.3.2 Link between performance and morphology of the catchments

We now look into the performances of the models and see whether there is a link between morphological features of the catchments, i.e. area and elevation, and the KGE' scoring on the different indicators. We make scatter plots of the KGE' values compared to the area and elevation properties of the catchments.

For each hydrological signature we select the best performing model overall on the selected indicator and use it for the scatterplots.

The scatter plots are limited for the stations that have $KGE' > -0.41$ as values above zero imply that there is some level of representativity of the catchment by the modeled data.

Monthly regime:

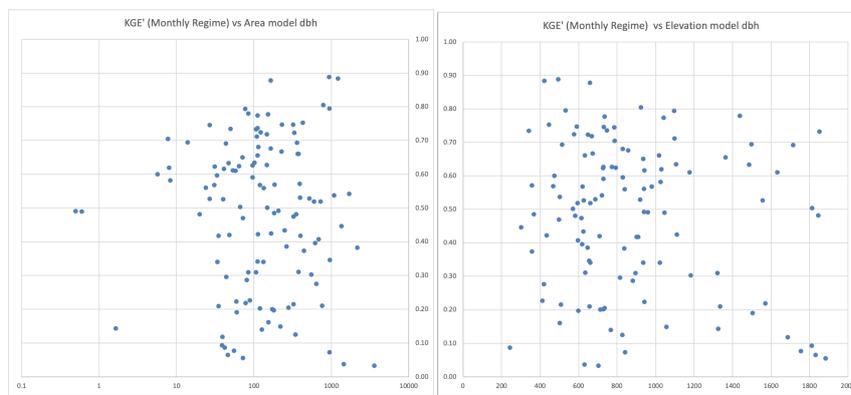


Figure 3.26: *Monthly regime vs morphological properties :no trend is noticeable. Only influence of the skewness of the initial dataset, i.e. many stations with small area*

Q5:

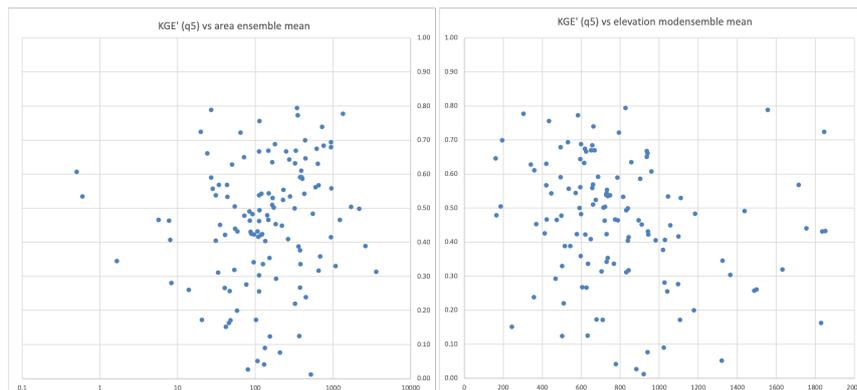


Figure 3.27: *Q₅ vs morphological properties: no trend is noticeable. Only influence of the skewness of the initial dataset, i.e. many stations with small area*

Q95:

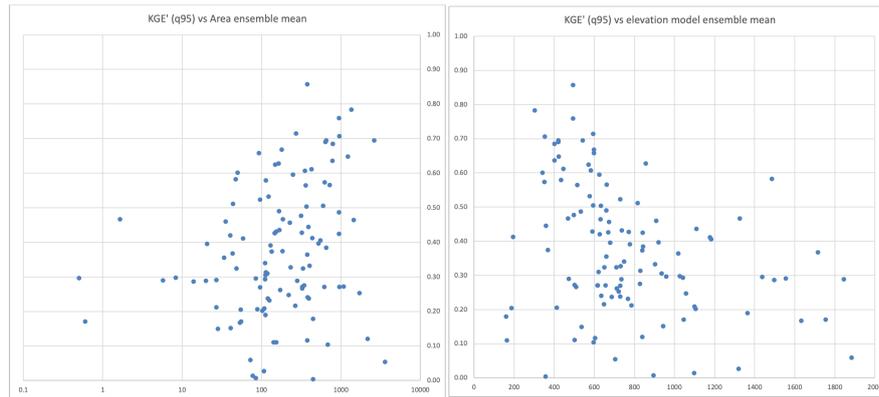


Figure 3.28: Results for Q_{95} vs morphological properties : Only in Q_{95} we can spot a certain trend in the performance of the catchment vs their morphological properties, that is consistent with the fact that low flows are more linked to small scale properties of the catchment such as soil moisture and local geology compared to monthly regime and q_5 .

Comments These attempts at showing a link between morphology and performance of the models show no particular interesting results. Only for Q_{95} is possible to note a decrease in performance with the decrease area, and a decrease of performance with the increase in elevation. The aim is to show that the models struggle more to represent the complex feature of small alpine catchments.

3.3.3 Link between performance of specific models on different indicator

Also in this section we want to show the different results for a selection of models. We use the scatter plot to highlight the different performance on the different indicators. Indeed one would expect to see a good correlation between performing well on any given hydrological signature and performing similarly on the others. But as results have showed this is not the case. In particular we showcase the result for three models.

1. "Ensemble Mean" : the ensemble mean of the whole models in the experiment nosoc is showcased
2. Model DBH
3. Model Matsiro

1) Ensemble Mean

We show the results for the ensemble mean as it outperforms each single model on the quantile signatures achieving a median score on the KGE' between all stations of 0.44 on Q_5 and of 0.29 on Q_{95} . The ensemble mean has a relative rank of 1 in both quantiles but falls a short in the monthly regime where it is ranked 5.

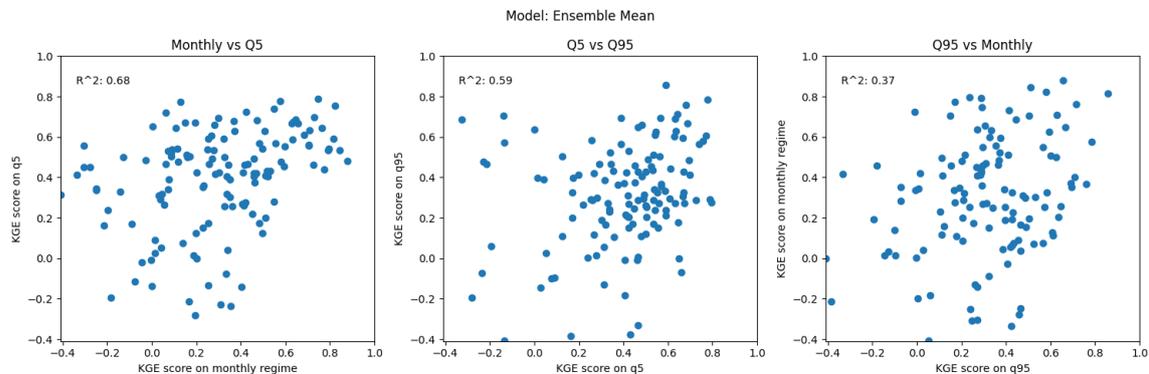


Figure 3.29: Compared results of KGE on the three signatures. Correlation of good performance across indicator is good for Monthly regime and high flows. Drops at

2) DBH

The model DBH outperforms each model in the monthly regime signature, but falls short on the Q_5 (rank 6) and even worse on the Q_{95} where it is ranked 13. It is possible to see how the model presents good performance on the first two signatures and a good correlation (0.74) between the score on those two, first graph in figure 3.16; the correlation drops while comparing the score of the first two signatures with the low flows signature.

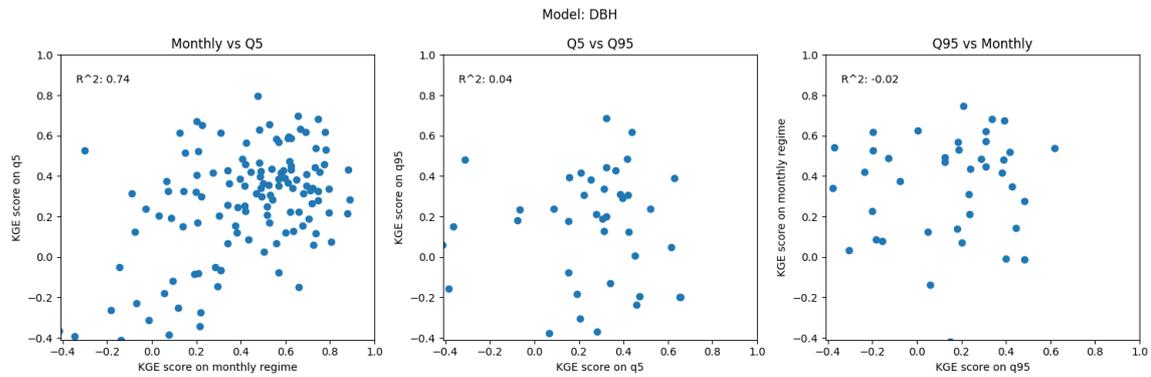


Figure 3.30: Compared results on the different KGE scores for model DBH. It is possible to see a good correlation between good performance on high flows and monthly regime. Correlation drops when comparing with low flows.

2) Matsiro

The model Matsiro is amongst the one who consistently performs badly being ranked 12th on the monthly regime 12th on the Q_5 10th on the Q_{95} . In the following graphs it is possible to appreciate how for this model performance on any signature is nearly uncorrelated with the performance on others because the results are overall really poor.

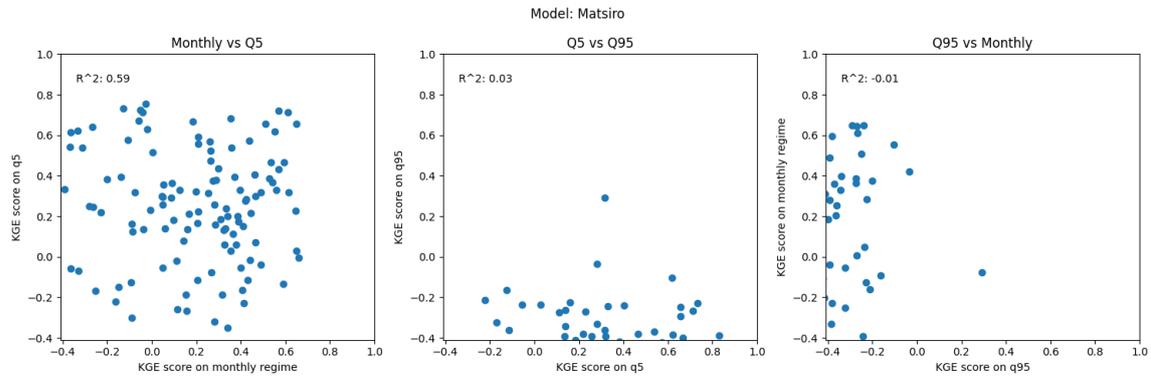


Figure 3.31: *Compared results of KGE score on the different signatures. Possible to appreciate a good correlation between good performance on the high and monthly signature while correlation drops when comparing with the low flows score. Limitin factor is the low score achieved in all stations for Q_{95}*

Chapter 4

Discussion

We now summarize the results of the analysis and draw some conclusions. First and foremost, even before the results of the study we were aware that there are several limiting factor in adopting GHMs to specific regions. As per the results of [7], the GHMs account for a big source of variability and struggle to reproduce correctly the signatures. In this section we firstly discuss the results and place them in the context of current literature then discuss the general use of GHMs for impact projections and resource planning, finally we discuss also the methodology used and the metric adopted (KGE’).

4.1 Discussion of the results

Worst signature: Low Flows(Q_{95}) :one of the main result of this study is the overall bad performance of all the models on the low flows. Indeed if we look across all models and stations it is the indicator for which the scoring is lowest (- 0.16). This discovery is as well coherent with [9] that uses a similar methodology, and although using a different scoring criteria, still finds that the low flows are the most poorly reproduced out of all the quantiles inspected, even finding further proof of consistency in the fact that there is a decreasing trend in performance while moving from the higher flows quantiles to the lower ones (see fig. 4.1).

Our results agree that low flows are difficult for models to reproduce and other studies point out how this could be due to the lower degree of correlation of the low flows with the precipitation events, but higher degree of correlation with small scale properties of the catchment such as the interaction between soil moisture and even vegetation in periods of prolonged droughts.[12] Amongst the model which performs well in this indicator we find the ensemble mean, which clearly exploits the overestimation of low flows by some models (e.g. Matsiro) and the underestimation from most of the others, and places itself in the middle outperforming by a considerable margin all of the other models (median performance across all stations = 0.29). We can see how the models who singularly performs better on these indicators are WaterGAP2 and Ways, both of which fall under the classification of Global Hydrological Model, this could imply the fact that the greater degree of attention to the

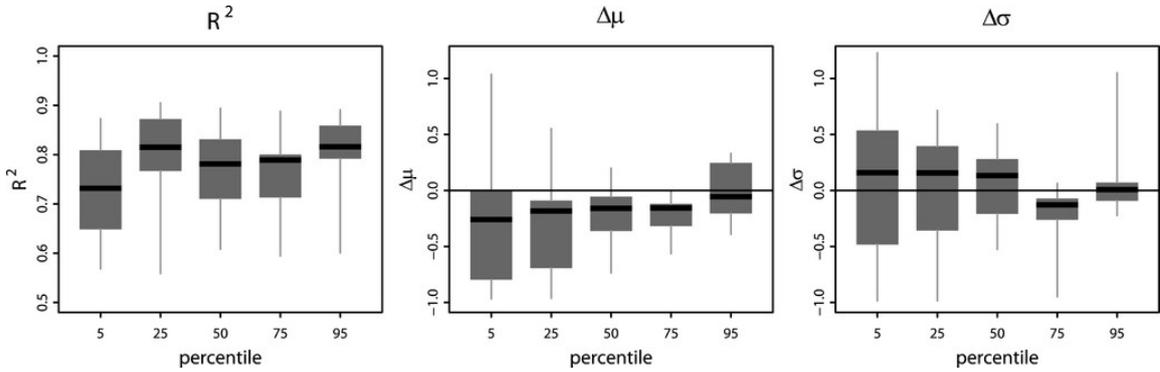


Figure 4.1: Results of Gudmundsson et al[9] for quantiles evaluation: in this picture the American notation of quantiles it's used, therefore the 95th/ 5th quantiles in our study are the 5/95th (swapped) in this graph, deriving from the use of non exceedance probability(American) and exceedance probability (European). Going from the low flows quantiles on the left we can see a marked decrease in the performance on both correlation, relative difference in mean (bias) and relative difference in standard deviation (variability) coherent with the results of our study

lateral fluxes in the models design have payed off in the description of the low flows, that, as said, are heavily dependent on the lateral movement of water and morphological properties than on the energy balance per se, although this is a speculation and need to be further investigated. A more reliable guess as in why those two models outperform the others lies in the fact that specifically models WaterGAP and WAYS are the only models for which a very simple calibration procedure was employed.

When inspecting the bias term, in the present work only the models Matsiro and VIC consistently overestimate the low flows while all the other models present a bias term consistently lower than 1 indicating a systematic underestimation instead. This result is amplified in our study and could be a consequence of the application of GHMs to the very specific and complex Alpine region, thus highlighting the differences while working on areas of complex morphology such as the chosen catchments. Further proving this point is the already highlighted limitation of the employed reanalysis products, which are downscaled to 0.5° but start from a coarser resolution (around 1°) and were proven in other studies [26] to be unable to capture highly localized phenomena of both high and low intensity which are typical of the Alpine area. Indeed also the disproportion between the pixel area(around 2500 km^2 and the catchment areas (median of catchment areas= 134 km^2) could be a decisive reason of difference in this sense. We can imagine that the runoff generation in a very small and elevated alpine catchment could be higher than the surrounding generated runoff in the remaining portions of the pixel. Therefore the leading cause for this underestimation probably lie both in the limitation of the input climatic dataset and in the inherent limitation of the GHMs in describing the hydrological phenomena involved.

Best performing: High Flows(Q_5). The high flows are the signature which overall scores the best over the model ensemble and across all stations (median value of $KGE' = 0.29$). Also this discovery applies to the results of other studies like the ones cited above.

All study tend to point out how models are often design for flood design purposes and that could be one of the reasons why they score higher than other quantiles. Furthermore high flows are well correlated with precipitation events. That is shown also in the high and pretty coherent across models value of the term "r" of the KGE' indicator. This result might also show that using the KGE' for evaluating the high flow might not be the best solutions, as probably the most important aspect of the peak flow signal lies in the volume of water reproduced, more than the variability or correlation of the signal. Indeed a specific indicator for the performance of model over the specific component of peak flow of the hydrograph exist and it is termed Annual Peak Flow Bias (APFB). We simply report its formula and notice the interesting quality of being exactly the KGE score with the weight of coefficients s1 and s2 = 0.

$$APFB = \sqrt{\left(1 - \frac{\mu_{modeled}}{\mu_{observed}}\right)^2} \quad (4.1)$$

Monthly Regime ranks in the middle but tells valuable information about GHMs. Monthly regime results show the negative bias and the difference between models in reproducing the timing of events within the year. In particular its clear to see how snow dominated catchments are poorly represented with the snow melt peak that is consistently anticipated, as it typically occurs between February an March in the modeled timeseries while only between April and May for the observed data. This confirms the difficulty of models of reproducing mechanisms of snowmelt that occurs at high elevation, and its consistent with the study of both Gadeke et al [12] and Giuntoli et al [7] which find this same results for two very different region such as North America and the Artic watersheds. The results of the cited studies as well as the present one point towards the fact that models have a tendency to overestimate the timing of runoff, in particular the mechanisms of snowmelt seem to be represented as way faster than the observational data reports. Another important aspect that is clear is that a poor score on the Monthly regime, more than poor scores on the extremes, shows the representativity of the catchment by the pixel, as was shown in the case study of the Durance river at Serre Poncon.

Differences between scenarios are minimal, quite surprisingly adding the different modelization of the human impacts i.e. considering a scenario with no human impacts and a scenario with historically varying human impacts, did not produce a significant difference in the results of the performance evaluation. We can point to several reasons as of why this could be the case, besides the already stressed limitation of the comparison between pixel and station: the first one is that, as pointed out in a specific study by Veldkamp et al [30], the increase in performance was reported to be significant especially in highly managed catchments. In our study the catchments are mostly small and almost natural catchments, indeed the biggest impact on the stations that are explored would most likely be the ones of reservoir and dams for hydropower production. That is seen for instance while looking at the Rhone at Gletsch , how for model DBH the curve of the monthly regime has a lower

peak in the spring month in the varsoc experiment (580 mm/month) compared to the 600 mm/month of the nosoc experiment. This result is coherent with the presence of reservoirs in the study catchment. The most interesting takeaway was how most model did not present any significant variation of scoring besides the model DBH. This could imply a really limited effort from other models into characterizing human impacts, compared to model DBH which states the evaluation of the effects of human impacts as one of its primary design purpose. Also this result highlight the importance of choosing the correct model for the correct purpose, stressed in many studies (e.g.[13]) and indicating the fact that each model is fine tuned for the representation of specific processes and not all processes combined can be optimized in the context of the same GHM.

4.2 Discussion on the methodology adopted

The KGE' score was selected as an indicator of performance and allows to visualize how the misfit is distributed along its component of bias, timing and variability. That is an interesting property and is the reason why the indicator was proposed in the first place. Kling and Gupta point out [11] how using the KGE decomposition of the MSE, allows to visualize pitfalls where the MSE fails to correctly discriminate between bad and good performance of models. This is not to say that the KGE does not carry the same problem, but in the KGE it is possible to expand on the single value and have a further look into the reasons of a specific value for a given obtained score. This allow for the combination of an "objective" metric such as the KGE' numerical value, with the subjective judgement of the interpreter, who can understand the reasons for a given score and, based on his expertise weight them accordingly. The combination of an objective and a subjective evaluation procedure is suggested as an optimal method by many authors such as [16]. One important disclaimer that has to be pointed out while discussing and drawing conclusion from the results of our analysis is the fact that, while the KGE' scores is indeed an indicator of good performance and it allows to have a deeper view into the causes of both, its results are not directly comparable across signatures. Indeed in this study, we use a one metric for all signature approach, but we have to understand how this limits the possibility of comparing the results as we are talking about different parts of the hydrograph with different underlying hydrological processes. The most obvious example of why they cannot be directly compared lies in the fact that the monthly regime curves are 12 month-long time series, thus evaluating the correlation for a different length (30 years for the high and low flows) has a different mathematical meaning. Also for the low flows and high flows, while the mathematical formulation is similar, and they are more comparable then the monthly regime within themselves. We must remember that the physical processes underlying the condition of peak flows (possibly flood) or low (drought) are significantly different. Indeed the use of specific metrics for the specific signatures gives a better estimates on the actual ability of the models to reproduce the underlying processes in a significant and physically-consistent way. [21]

Another final yet important remark is the fact that, as stated while presenting the data set, there is an inherent limitation which lies in the partial match between pixel and catchment areas. The results can be therefore interpreted also in the sense of providing a list of stations that is well or badly represented by the models.

4.3 GAR modeling and climate change

We can now discuss the roles of GHMs in the context of future climate change projections. We have highlighted some of their limitation and in general showcased their output on selected locations. Using the downloaded data for the experiment 2b in the future scenarios RCP 8.5 (i.e. a scenario where no climate change mitigation is performed). To have a glimpse in the future of the alpine regions hydrology and showcase potential uses of GHMs. We show results for the ensemble mean of the 2b scenarios for the variation of two signature already presented in chapter 2: the mean annual specific runoff, and the Pardè Range. We then plot near them the monthly regime curves for the catchment already selected in the result section of Rhone at Gletsch and Soca at Kobarid.

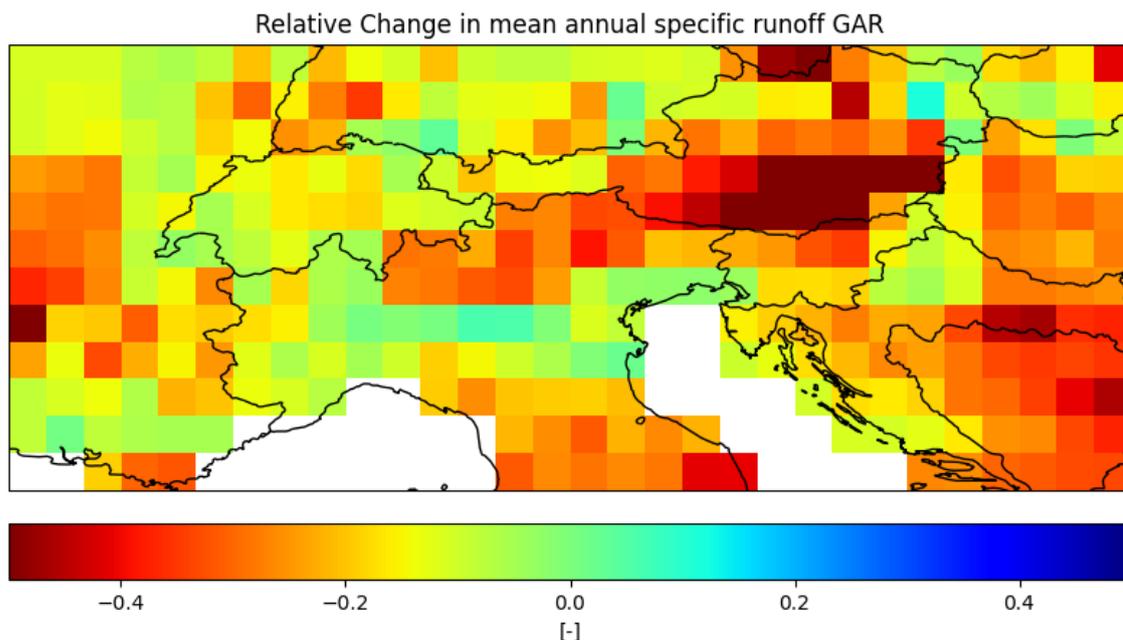


Figure 4.2: Map of the relative change of mean annual specific runoff for the GAR between the mean of the 1971-2000 period and the 2071-2099 period in the scenario RCP 8.5

We can see in figure 4.2 how the change in predicted generated runoff is heterogeneously distributed in the GAR. In particular the south east corner show a marked decrease in runoff between the historical and future period, coherently with what is shown in the predictions for the precipitation of the area in the study by Gobi et al[8]. Furthermore if we investigate into the representative catchment already shown in the previous chapter for the Soca river basin we can see this results in terms of the monthly regime curves.

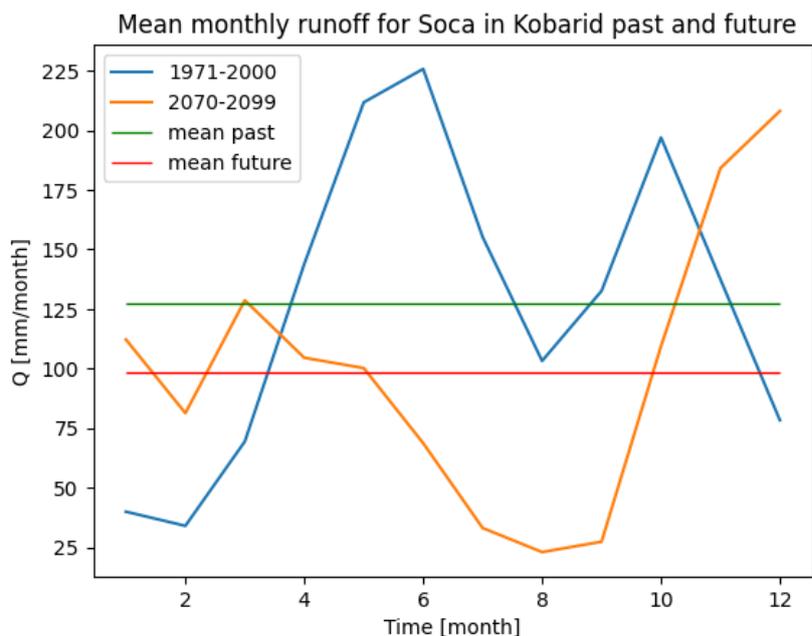


Figure 4.3: Monthly regime curves for the 1971-2000 period and the 2071-2099 period in the scenario RCP 8.5. We can appreciate the decrease in overall runoff as well as the shift of the spring peak towards the winter season due to the increase in temperatures.

We also see the same change but for the Pardè range in order to inspect the spatial distribution of the change in seasonality:

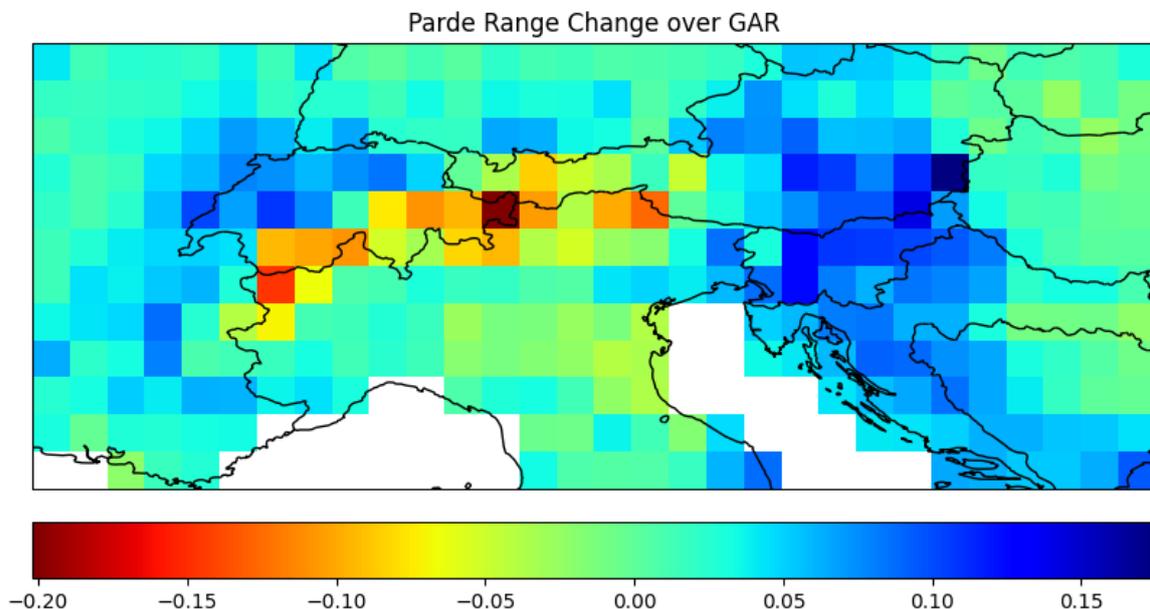


Figure 4.4: Map of the change of Pardè range for the GAR between the mean of the 1971-2000 period and the 2071-2099 period in the scenario RCP 8.5

Where we see how the most marked decrease in the Pardè range is located in the mountainous region, indicating a big shift in seasonality for the snow dominated catchments due to the increase in temperatures indicating the abrupt changes to which the Alps will be potentially affected in the future. To once again see the results in terms of a specific catchment

we look into the already inspected Rhone catchment at Gletsch.

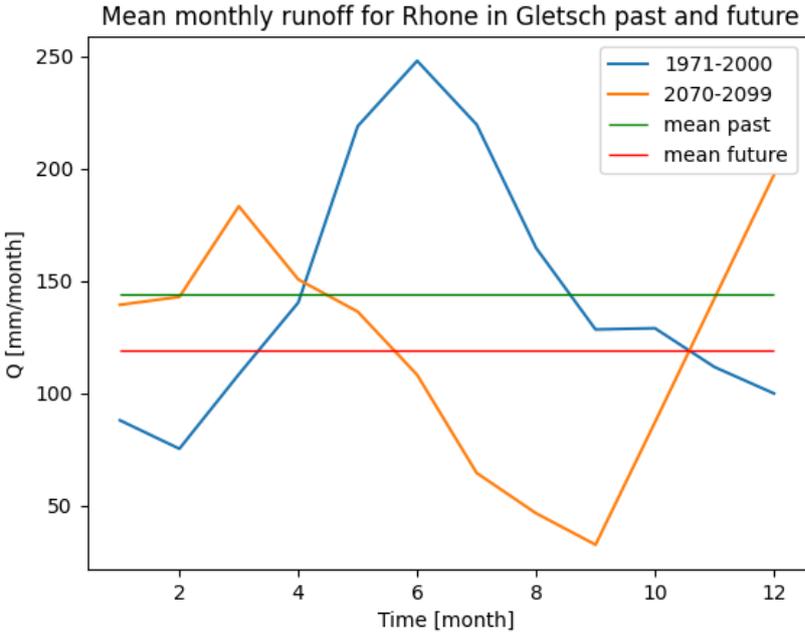


Figure 4.5: Monthly regime curves for the 1971-2000 period and the 2071-2099 period in the scenario RCP 8.5. Significant changes in regime type and volumes of water due to climate change in the unmitigated scenarios for the high alps

4.4 Conclusions

The conclusion of our analysis can be summarized in the following way:

- The KGE' scores on the different indicators highlight a marked decrease in performance of the models passing from the high flow percentile to the low flows'. The monthly regime curves rank in the middle of the two indicators, although is not possible to directly compare the scores of the same metric across indicators as their meaning changes with the underlying processes involved.
- The ensemble mean outperforms each single model in the percentiles and obtains sufficient performance in all signatures, proving as a reliable estimator.
- Significant negative bias exists across all indicators, which is only partly consistent with other studies and could be a symptom of the well known negative bias of precipitation in the reanalysis products across the Alpine areas and in areas at high latitudes. Models such as WaterGAP2, that have undergone a simple calibration procedure show better results across all stations and indicators compared to others.
- Model DBH is the model that consistently reproduces the monthly regime better than all other models. Notwithstanding this its performance drop considerably into the high flows, and even more so on the low flows. This results highlight the goodness of working on a multiple series of signature in order to highlight the different pros and cons of the specific models; this result also points toward the potential of using specific models in specific section of the hydrograph based on their design purpose and calibration methods.
- The results of comparing the models on the different scenarios show a lack of change between the two. This result only partially coincides with other studies and points toward the specificity of the study area and once again highlights the limitation of adopting GHMs in the GAR; but also raising questions about the different representation of human impacts in the models adhering to ISIMIP. One further explanation as of why this difference is small could also lie in the fact that the small alpine catchment used for comparison are de facto considerable as natural catchments.
- Finally, in the context of climate change, GHMs could provide a valuable tool and study such as this one could lay the foundations for understanding impact studies and future hydrological projections, especially with the increase in process representations and detail of GHMs but also of Reanalysis products (e.g. ERA 5) ; the latter in particular have already increased significantly in spatial resolution and the results of the next round of ISIMIP could showcase quite different results compared to the ones showed in this study.

Bibliography

- [1] A simple hydrologically based model of land surface water and energy fluxes for general circulation models - Liang - 1994 - Journal of Geophysical Research: Atmospheres - Wiley Online Library.
- [2] Ingeborg Auer, Reinhard Böhm, Anita Jurkovic, Wolfgang Lipa, Alexander Orlik, Roland Potzmann, Wolfgang Schöner, Markus Ungersböck, Christoph Matulla, Keith Briffa, Phil Jones, Dimitrios Efthymiadis, Michele Brunetti, Teresa Nanni, Maurizio Maugeri, Luca Mercalli, Olivier Mestre, Jean-Marc Moisselin, Michael Begert, Gerhard Müller-Westermeier, Vit Kveton, Oliver Bochnicek, Pavel Stastny, Milan Lapin, Sándor Szalai, Tamás Szentimrey, Tanja Cegnar, Mojca Dolinar, Marjana Gajic-Capka, Ksenija Zaninovic, Zeljko Majstorovic, and Elena Nieplova. Histalp—historical instrumental climatological surface time series of the greater alpine region. *International Journal of Climatology*, 27(1):17–46, 2007.
- [3] Jörg Uwe Belz, Gerhard Brahmmer, Hendrik Buiteveld, H Engel, R Grabher, H Hodel, P Krahe, R Lammersen, M Larina, HG Mendel, et al. *Das Abflussregime des Rheins und seiner Nebenflüsse im 20. Jahrhundert: Analyse, Veränderungen, Trends*. Internationale Kommission für die Hydrologie des Rheingebietes Lelystad, The . . . , 2007.
- [4] M. J. Best, M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding. The Joint UK Land Environment Simulator (JULES), Model description – Part 1: Energy and water fluxes. preprint, Climate and Earth System Modeling, March 2011.
- [5] Marco Cucchi, Graham P. Weedon, Alessandro Amici, Nicolas Bellouin, Stefan Lange, Hannes Müller Schmied, Hans Hersbach, and Carlo Buontempo. WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, 12(3):2097–2120, September 2020. Publisher: Copernicus GmbH.
- [6] Petra Döll, Frank Kaspar, and Bernhard Lehner. A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, 270(1):105–134, January 2003.

- [7] Ignazio Giuntoli, Ilaria Prosdocimi, and David M. Hannah. Going Beyond the Ensemble Mean: Assessment of Future Floods From Global Multi-Models. *Water Resources Research*, 57(3):e2020WR027897, March 2021. Publisher: John Wiley & Sons, Ltd.
- [8] Andreas Gobiet, Sven Kotlarski, Martin Beniston, Georg Heinrich, Jan Rajczak, and Markus Stoffel. 21st century climate change in the European Alps—A review. *Science of The Total Environment*, 493:1138–1151, September 2014.
- [9] Lukas Gudmundsson, Lena M. Tallaksen, Kerstin Stahl, Douglas B. Clark, Egon Dumont, Stefan Hagemann, Nathalie Bertrand, Dieter Gerten, Jens Heinke, Naota Hanasaki, Frank Voss, and Sujan Koirala. Comparing Large-Scale Hydrological Model Simulations to Observed Runoff Percentiles in Europe. *Journal of Hydrometeorology*, 13(2):604–620, April 2012. Publisher: American Meteorological Society Section: Journal of Hydrometeorology.
- [10] Matthieu Guimberteau, Dan Zhu, Fabienne Maignan, Ye Huang, Chao Yue, Sarah Dantec-Nédélec, Catherine Ottlé, Albert Jornet-Puig, Ana Bastos, Pierre Laurent, Daniel Goll, Simon Bowring, Jinfeng Chang, Bertrand Guenet, Marwa Tifafi, Shushi Peng, Gerhard Krinner, Agnès Ducharne, Fuxing Wang, Tao Wang, Xuhui Wang, Yilong Wang, Zun Yin, Ronny Lauerwald, Emilie Joetzjer, Chunjing Qiu, Hyungjun Kim, and Philippe Ciais. ORCHIDEE-MICT (revision 4126), a land surface model for the high-latitudes: model description and validation. preprint, Climate and Earth System Modeling, June 2017.
- [11] Hoshin V. Gupta, Harald Kling, Koray K. Yilmaz, and Guillermo F. Martinez. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1):80–91, October 2009.
- [12] Anne Gädeke, Valentina Krysanova, Aashutosh Aryal, Jinfeng Chang, Manolis Grillakis, Naota Hanasaki, Aristeidis Koutroulis, Yadu Pokhrel, Yusuke Satoh, Sibyll Schaphoff, Hannes Müller Schmied, Tobias Stacke, Qihong Tang, Yoshihide Wada, and Kirsten Thonicke. Performance evaluation of global hydrological models in six large Pan-Arctic watersheds. *Climatic Change*, 163(3):1329–1351, December 2020.
- [13] Ingjerd Haddeland, Douglas B. Clark, Wietse Franssen, Fulco Ludwig, Frank Voß, Nigel W. Arnell, Nathalie Bertrand, Martin Best, Sonja Folwell, Dieter Gerten, Sandra Gomes, Simon N. Gosling, Stefan Hagemann, Naota Hanasaki, Richard Harding, Jens Heinke, Pavel Kabat, Sujan Koirala, Taikan Oki, Jan Polcher, Tobias Stacke, Pedro Viterbo, Graham P. Weedon, and Pat Yeh. Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. *Journal of Hydrometeorology*, 12(5):869–884, October 2011. Publisher: American Meteorological Society Section: Journal of Hydrometeorology.

- [14] Naota Hanasaki, Sayaka Yoshikawa, Yadu Pokhrel, and Shinjiro Kanae. A global hydrological simulation to specify the sources of water used by humans. *Hydrology and Earth System Sciences*, 22(1):789–817, January 2018. Publisher: Copernicus GmbH.
- [15] U. Heikkilä, Anne Sandvik, and Asgeir Sorteberg. Dynamical downscaling of era-40 in complex terrain using the wrf regional climate model. *Climate Dynamics*, 37:1551–1564, 04 2012.
- [16] Valentina Krysanova, Chantal Donnelly, Alexander Gelfan, Dieter Gerten, Berit Arheimer, Fred Hattermann, and Zbigniew W. Kundzewicz. How the performance of hydrological models relates to credibility of projections under climate change. *Hydrological Sciences Journal*, 63(5):696–720, April 2018. Publisher: Taylor & Francis. eprint: <https://doi.org/10.1080/02626667.2018.1446214>.
- [17] Ganquan Mao and Junguo Liu. WAYS v1: a hydrological model for root zone water storage simulation on a global scale. *Geoscientific Model Development*, 12(12):5267–5289, December 2019. Publisher: Copernicus GmbH.
- [18] Daniel Moriasi, Jeff Arnold, Michael Van Liew, Ron Bingner, R.D. Harmel, and Tamie Veith. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50, May 2007.
- [19] Hannes Müller Schmied, Linda Adam, Stephanie Eisner, Gabriel Fink, Martina Flörke, Hyungjun Kim, Taikan Oki, Felix Theodor Portmann, Robert Reinecke, Claudia Riedel, Qi Song, Jing Zhang, and Petra Döll. Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use. *Hydrology and Earth System Sciences*, 20(7):2877–2898, July 2016. Publisher: Copernicus GmbH.
- [20] W. Oleson, M. Lawrence, B. Bonan, G. Flanner, Erik Kluzek, J. Lawrence, Samuel Levis, C. Swenson, E. Thornton, Aiguo Dai, Mark Decker, Robert Dickinson, Johannes Feddes, L. Heald, Forrest Hoffman, Jean-Francois Lamarque, Natalie Mahowald, Guo-Yue Niu, Taotao Qian, James Randerson, Steve Running, Koichi Sakaguchi, Andrew Slater, Reto Stockli, Aihui Wang, Zong-Liang Yang, Xiaodong Zeng, and Xubin Zeng. Technical Description of version 4.0 of the Community Land Model (CLM). 2010.
- [21] Matthias Pfannerstill, Björn Guse, and Nicola Fohrer. Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *Journal of Hydrology*, 510:447–458, 03 2014.
- [22] Wei Qi, Lian Feng, Hong Yang, Junguo Liu, Yi Zheng, Haiyun Shi, Lei Wang, and Deliang Chen. Economic growth dominates rising potential flood risk in the Yangtze River and benefits of raising dikes from 1991 to 2015. *Environmental Research Letters*, 17(3):034046, March 2022. Publisher: IOP Publishing.

- [23] C Schär, TD Davies, C Frei, H Wanner, M Widmann, M Wild, and HC Davies. Current alpine climate. *Views from the Alps: Regional perspectives on climate change*, 21:63, 1998.
- [24] S. Sitch, B. Smith, I. C. Prentice, A. Arneth, A. Bondeau, W. Cramer, J. O. Kaplan, S. Levis, W. Lucht, M. T. Sykes, K. Thonicke, and S. Venevsky. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, 9(2):161–185, 2003. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2486.2003.00569.x>.
- [25] T. Stacke and S. Hagemann. Development and evaluation of a global dynamical wetlands extent scheme. *Hydrology and Earth System Sciences*, 16(8):2915–2933, August 2012. Publisher: Copernicus GmbH.
- [26] Fengge Su, Jennifer C Adam, Kevin E Trenberth, and Dennis P Lettenmaier. Evaluation of surface water fluxes of the pan-arctic land region with a land surface model and era-40 reanalysis. *Journal of Geophysical Research: Atmospheres*, 111(D5), 2006.
- [27] Edwin H. Sutanudjaja, Rens van Beek, Niko Wanders, Yoshihide Wada, Joyce H. C. Bosmans, Niels Drost, Ruud J. van der Ent, Inge E. M. de Graaf, Jannis M. Hoch, Kor de Jong, Derek Karssenberg, Patricia López López, Stefanie Peßenteiner, Oliver Schmitz, Menno W. Straatsma, Ekkamol Vannamete, Dominik Wisser, and Marc F. P. Bierkens. PCR-GLOBWB 2: a 5th arcmin global hydrological and water resources model. *Geoscientific Model Development*, 11(6):2429–2453, June 2018. Publisher: Copernicus GmbH.
- [28] Kumiko Takata, Seita Emori, and Tsutomu Watanabe. Development of the minimal advanced treatments of surface interaction and runoff. *Global and Planetary Change*, 38(1):209–222, July 2003.
- [29] Qihong Tang, Taikan Oki, Shinjiro Kanae, and Heping Hu. The Influence of Precipitation Variability and Partial Irrigation within Grid Cells on a Hydrological Simulation. *Journal of Hydrometeorology*, 8(3):499–512, June 2007. Publisher: American Meteorological Society Section: Journal of Hydrometeorology.
- [30] T I E Veldkamp, F Zhao, P J Ward, H de Moel, J C J H Aerts, H Müller Schmied, F T Portmann, Y Masaki, Y Pokhrel, X Liu, Y Satoh, D Gerten, S N Gosling, J Zaherpour, and Y Wada. Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study. *Environmental Research Letters*, 13(5):055008, may 2018.
- [31] Wikipedia. Danubio — wikipedia, l’enciclopedia libera, 2023. [Online; in data 3-luglio-2023].

- [32] Wikipedia contributors. Principal component analysis — Wikipedia, the free encyclopedia, 2023. [Online; accessed 26-June-2023].
- [33] Zhu Zhongming, Lu Linong, Yao Xiaona, Zhang Wangqiang, Liu Wei, et al. Regional climate change and adaptation—the alps facing the challenge of changing water resources. 2009.