



Bias Correction of Climate Simulations to Assess Climate Change Impacts on Low Flows in the Rhine River

Ahmed Abdelnour

Bias Correction of Climate Simulations to assess Climate Change Impacts on Low Flows in the Rhine River

by

Ahmed Adil Abdelnour

To obtain the degree of Master of Science at Delft University of Technology,

To be publicly defended on the 8th of July 2022 at 11:00.

Student number: 5246563

Project duration: 3.12.2021 – 8.7.2022

Thesis committee:	Prof. Dr. Remko Uijlenhoet	TU Delft (Chair)
	Dr. Astrid Blom	TU Delft
	Dr. Frederiek Sperna Weiland	Deltares
	Prof. Dr. Albrecht Weerts	Deltares - WUR

An electronic version of this thesis can be found at:

The cover photo is obtained from: <https://www.flickr.com/photos/bballandephotographe/34006942002>



Preface

The world is experiencing unprecedented challenges due to climate change, that in turn has jeopardized water uses worldwide. This poses the need for proactive water management strategies that help in alleviating these issues and making the world more sustainable. I have tried my best to utilize my interest in modelling and programming to reflect upon one of these global challenges in this project.

I have learned so much in the past seven months from my supervisors from TU Delft and Deltares. Special thanks go to Remko for chairing my committee and for the critical inputs. I also want to thank Astrid for teaching me how to look at things from a different perspective and summarize the message of this work in a clear manner. Thanks to Albrecht for helping me with the setup of the Rhine model and to his fruitful suggestions during our biweekly meetings. Finally, I would like to thank Frederiek for being my daily supervisor, her extremely positive attitude and helping me with every issue that I have encountered during the project. It has been a true pleasure to be guided by the four of you during this journey!

My time at TU Delft has come to its end, a journey that went quite quickly. I made many friends and enjoy memorable experiences. Although I have done most of my courses online in the first year, I got to do a lot of things during the second year; being part of the Global initiative and the AWC, doing my internship at WEP and my graduation project in collaboration with Deltares.

Aside from technical aspects, I would like to dedicate this work to my wonderful parents (Ghadda and Adil). Thank you for being my role model, always believing and helping me in becoming a better person. I also want to emphasize my gratitude to my siblings (Ayman, Alaa and Nour). Thank you all for the endless support and making this journey bearable while being far from home!

Ahmed Adil,

Delft, July 2022.

Abstract

Regional climate models (RCMs) simulations are used in hydrological (climate-change) impact assessment studies. However, RCMs exhibit noticeable deviation from observation, and can show large variation in ensemble projections (biases). The objective of this study is twofold, first to assess the robustness of two high-skill bias-correction methods; empirical quantile mapping (QM), and scaled distribution mapping (SDM) in improving the hydrological modelling of the Rhine basin. The second is to assess the potential impacts of climate change on low flows at Lobith based on RCP8.5 scenario. The two correction methods are applied to correct the systematic bias from five climate datasets from the Coordinated Downscaling Experiment in Europe (EURO-CORDEX) covering the Rhine domain, using high resolution gridded datasets (1 km²) spanning from 1979 to 2019.

The bias-corrected simulations from the hydrological model provided more accurate discharge estimates than the wet biased simulations, with an average error of less than 100 m³/s at Lobith. The correction methods are also capable of correcting unprecedented temperature and precipitation values, making them useful in climate assessment studies in the Rhine river. However, it appears that the accuracy of the bias correction depends on the parent GCM, performance of the raw RCM and the skill of the hydrological model in estimating discharges at the point of interest. In addition to that, the drizzling effect could not be reduced using these methods.

Noticeable climate change impacts at Lobith are found using the bias-corrected projections. These projections suggest that low flows are going to be more frequent and longer in the coming 38 years. Unprecedented discharges (< 700 m³/s) are projected to occur at least 50 times between 2020 - 2060. This is coupled by a decrease in the long-term mean annual flow by 100 m³/s and a slight shift in the seasonality of low flows (2 weeks shift).

The general hydrograph at Lobith is set to change due to climate change for the period (2020 – 2060), with relatively higher discharges from early June to end of August followed by relatively lower discharges in the last four months of the year. Water levels are projected to decline in average by 20 cm (early June to the end of August) and increase in average by 30 cm (end of August till to end of December). The study recommends the need of combining bias correction, the feedbacks in the climate system (land use changes) and climate adaptation strategies to study these effects further.

Table of Contents

1 Introduction.....	11
1.1 The Motive: inconsistency with reality	11
1.2 Conceptual background	12
1.3 Research scope.....	12
1.4 Research question.....	13
1.5 Reader guide.....	14
2 Area Description and Data.....	15
2.1 Study Area.....	15
2.1.1 The Rhine.....	15
2.1.2 Lobith.....	17
2.2 Data	17
2.2.1 Observations	17
3 Methods.....	25
3.1 Bias-Correction.....	25
3.1.1 Empirical Quantile Mapping	26
3.1.2 Scaled Distribution mapping	29
3.2 Historical experiment.....	32
3.3 Non-Stationarity under scope.....	33
3.4 Performance-based weighting.....	35
3.4.1 Climate Model Weighting by Independence and Performance.....	35
3.4.2 Reliability Ensemble Averaging.....	37
3.5 Hydrological Model.....	38
3.5.1 Additional Correction for temperature	39
3.6 Performance Criteria	39
4 Results	40
4.1 Hydrological model.....	40
4.1.1 Observed discharges.....	40
4.1.2 Climate simulations	45
4.1.3 Performance-based weighting.....	46
4.2 Historical experiment.....	48
4.2.1 Quantile mapping and scaled distribution mapping.....	48
4.2.2 Long-term statistics.....	53
4.3 Non-stationarity.....	55
4.3.1 Precipitation	55

4.3.2 Temperature.....	57
4.4 Projecting climate change impacts by 2060.....	58
4.4.1 Lobith.....	58
4.4.2 General climate trends in the Rhine basin.....	65
5 Discussion.....	69
5.1 Hydrological model.....	69
5.2 Uncorrected climate simulations.....	69
5.3 Performance-based weighting.....	69
5.4 Bias correction in the historical experiment.....	70
5.5 Non-stationarity testing.....	70
5.6 Future climate change impacts on low flows at Lobith.....	71
5.7 Empirical quantile mapping and scaled distribution mapping.....	71
5.8 Research limitations.....	71
5.8.1 Observational datasets.....	71
5.8.2 Climate simulations.....	72
5.8.3 Hydrological modelling.....	72
5.8.4 Sensitivity of the hydrological model.....	72
5.8.5 Land use changes.....	73
5.8.6 Bias correction.....	73
5.8.7 Lobith.....	73
5.8.8 Climate change adaptation strategies.....	73
5.8.9 Water levels.....	73
5.9 Future Research.....	74
6 Conclusion.....	75
7 Appendix I.....	84
7.1 Makkink Evaporation:.....	84
7.2 Rating Curves.....	84
7.3 Assessment metrics.....	86
7.3.1 Long-term statistics.....	86
7.3.2 Meteorology.....	87
7.3.3 Hydrological model assessment.....	88
7.3.4 Other hydrological metrics.....	89
7.4 General hydrological modelling framework.....	90
7.4.1 Bias correction.....	90
7.4.2 ClimWIP.....	92

7.4.3 REA.....	92
8 Appendix II: Additional Results	93
8.1 Dry spells and SPI analysis	97
9 Appendix III: Kaub, Cochem, and Rees.....	100
9.1 Kaub	101
9.2 Cochem	101
9.3 Rees	102
10 Appendix IX: Rockenau and Andelfingen	103
10.1 Rockenau.....	104
10.2 Andelfingen.....	106
11 Appendix X.....	108
11.1 Spatiotemporal REA.....	108

Nomenclature

Abbreviation	Actual name
GCM	General Circulation Model
RCM	Regional Circulation Model
SDM	Scaled distribution mapping method
DWD	Deutscher Wetterdienst
ERA-5	fifth-generation reanalysis data for global climate and weather
HYRAS	High-resolution grid observational datasets
E-OBS	Observational dataset in Europe
REA	Reliability Ensemble Averaging
Wflow_sbm	Wflow simple bucket model
ClimWIP	Climate Model Weighting by Independence and Performance
ST-REA	Spatiotemporal reliability ensemble averaging
QM	Quantile mapping method
KNMI	Royal Netherlands Meteorological Institute
SMHI	Swedish Meteorological and Hydrological Institute
DMI	Danish Meteorological Institute
CNRM	Centre National de Recherches Meteorologique
IPSL	Institut Pierre Simon Laplace
DSST	Differential split sample test
CDO	Climate Data Operators
SPI	Standard precipitation index

List of Figures

Figure 2-1: Elevation map and main tributaries of the Rhine basin.....	16
Figure 2-2: Location of Lobith within the Rhine basin.....	17
Figure 2-3: Number of gauging stations considered in creating the HYRAS dataset by DWD.	18
Figure 2-4: An example of the HYRAS precipitation dataset at 11/04/1979.	19
Figure 2-5: E-OBS temperature dataset at 11/4/1979.....	20
Figure 2-6: Downscaled ERA5 Makkink evaporation at 11/4/1979.....	21
Figure 2-7: daily Observed discharges at Lobith for the period [1979 - 2018].	22
Figure 2-8: Daily discharge data for additional locations [1979 - 2000].....	23
Figure 3-1: Typical model chain in hydrology.....	25
Figure 3-2: general description of climate simulations	26
Figure 3-3: representation of quantile mapping correction on a monthly scale	28
Figure 3-4: General framework for correcting precipitation projections using SDM.	31
Figure 3-5: DSST for temperature.....	34
Figure 3-6: DSST for precipitation.	35
Figure 3-7: Wflow_sbm schematic representation.	38
Figure 4-1: wflow_sbm vs observed during the period [1979 - 2018] at Lobith.	41
Figure 4-2: Scatter plot of wflow_sbm results vs observed measurements at Lobith for (1979-2018).....	42
Figure 4-3: flow and logarithmic flow duration curves for the observed flows and modelled flows for the period (1979-2018).....	43
Figure 4-4: Maximum number of consecutive low flow days in a year.	44
Figure 4-5: Average daily modelled and observed discharges at Lobith for the period 1999 – 2018.	45
Figure 4-6: Average daily flows at Lobith for the period 1979 – 2018 using an ensemble of four EURO-CORDEX simulations.....	46
Figure 4-7: ClimWIP weighted daily discharges using the raw simulations for (1999 – 2018) relative to the raw simulations and the actual discharges at Lobith.....	47
Figure 4-8: REA weighted average daily discharge for the period (1999 – 2018) relative to the raw simulations and the actual discharges at Lobith.	48
Figure 4-9: wflow_sbm forced with bias-corrected EURO-CORDEX simulations for the period (1999 - 2018) at Lobith compared to the actual flows at Lobith for the same period.	50
Figure 4-10: Average precipitation comparison between the observed, raw, and corrected datasets at and near Lobith during the historical experiment	51
Figure 4-11: Average temperature comparison between the observed, raw, and corrected datasets at and near Lobith during the historical experiment	52
Figure 4-12: daily average discharge envelope of the raw forcing's and corrected forcing's for the period (1999 – 2018) in addition to the daily average discharge of the same period.....	53
Figure 4-13: Long-term statistics at Lobith for the period (1999-2018).	54
Figure 4-14: No of days below threshold (1100 m3/s) for the period (1999 - 2018) at Lobith.....	54
Figure 4-15: Precipitation correction experiment through the DSST for winter and summer periods..	56
Figure 4-16: temperature correction experiment through the DSST for winter and summer periods..	57
Figure 4-17: Daily averaged discharges for the period (2020 - 2060) for the raw and bias-corrected simulations compared to the actual discharges in the reference run (1979 - 2019).....	58
Figure 4-18: average daily modelled discharges for each subset during the period (2020 - 2060) vs the actual observed discharge (1999 - 2019).	61
Figure 4-19: Long-term annual mean discharge at Lobith for the period (2020 – 2060)..	62

Figure 4-20: Histogram of the lengths of low flow periods of the bias-corrected simulations and observed discharges.....	63
Figure 4-21: SR values for the bias-corrected subsets (2020 – 2060) and for the observation (1979 – 2018).	63
Figure 4-22: Projected water levels at Lobith for the period (2020 - 2060) relative to the observed values...	64
Figure 4-23: Projected precipitation trend in the bias-corrected datasets.....	66
Figure 4-24: Temperature trend in the bias-corrected datasets for the period (2020 - 2060).....	67
Figure 4-25: Makkink evaporation trend in the bias-corrected simulations for the period (2020 - 2060).....	68
Figure 7-1: Fitted rating curve at Lobith based (Rijkswaterstaat, 2016) measurements	86
Figure 7-2: Hydrological modelling framework using QM and SDM	90
Figure 7-3: Modelling framework using ClimWIP	92
Figure 7-4: Hydrological modelling framework using REA.....	92
Figure 8-1: Discharges at Lobith obtained by forcing wflow_sbm by raw EURO-CORDEX simulations.....	93
Figure 8-2: SDM performance in the DSST for precipitation.....	94
Figure 8-3: QM performance in the DSST for precipitation.....	94
Figure 8-4: SDM performance in the DSST for temperature	94
Figure 8-5: QM performance in the DSST for temperature.	95
Figure 8-6: Lengths below threshold for the period (2020 - 2060) using bias-corrected EURO-CORDEX simulations.....	96
Figure 8-7: SR Q95 values for summer and winter values	96
Figure 8-8: Dry spell analysis of the bias-corrected simulations at Lobith for the period (2020 - 2060).	97
Figure 8-9: SPI Calculation for the corrected KNMI subset using Q.....	98
Figure 8-10: SPI Calculation for the corrected SMHI subset using QM.....	98
Figure 8-11: SPI Calculation for the projected period (2020 - 2060) for the corrected KNMI subset using SDM.....	99
Figure 8-12: SPI Calculation for the projected period (2020 - 2060) for the corrected SMHI subset using SDM.....	99
Figure 9-1: Description of how Kaub, Rees, Cochem, Rockenau and Andelfingen are distributed within the Rhine basin.	100
Figure 9-2: Long-term statistics at Kaub, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 1000 m ³ /s at Kaub.	101
Figure 9-3: Long-term statistics at Cochem, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 300 m ³ /s at Cochem.....	101
Figure 9-4: Long-term statistics at Rees, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 1500 m ³ /s at Rees.	102
Figure 10-1: Rockenau average daily discharge for the period (2020 - 2060) vs the actual discharges for the period (1979 - 2019).....	104
Figure 10-2: Bias-corrected simulations at Rockenau for the period (2020 - 2060) compared to the actual discharges for the period (1979 - 2019).....	105
Figure 10-3: Andelfingen average daily discharge for the period (2020 - 2060) vs the actual discharges for the period (1979 - 2019).....	106
Figure 10-4: Bias-corrected simulations at Andelfingen for the period (2020 - 2060) compared to the actual discharges for the period (1979 - 2019).....	107
Figure 11-1: Spatiotemporal REA weighting method applied at Lobith based on Kaub, Cochem, and Basel for the period (1998 - 2018).....	110

List of Tables

Table 2-1: Land use classes in the Rhine basin.....	15
Table 2-2: General meteorological characteristics of the Rhine basin	16
Table 2-3: List of used climate simulations.	24
Table 3-1: Temperature correction procedure using SDM	31
Table 3-2: Different periods in the historical experiment	33
Table 4-1: ClimWIP GCMs' weights.	47
Table 4-2: Corresponding RCMs' weights using ClimWIP	47
Table 4-3: Derived RCM weights using REA for the period (1979-2018).....	47
Table 7-1: SPI Values description.....	87

1| Introduction

More severe impacts of climate change were witnessed in the recent decades on the regional and global scales, and of course Europe is no exception (IPCC, 2021). The world is prone to more frequent, and more extreme climate events (Sutton and Hawkins, 2020). The hydrological cycle has been severely disrupted by having higher temperature and emission rates, which led to noticeable changes in the timing and magnitude of floods and droughts (Lobanova et al., 2018).

The Rhine river is one of the major rivers in Europe, originating in Switzerland and flowing into in the North Sea after flowing through the Netherlands. The river has been severely impacted by the effects of climate change in the recent decades, most notably by long drought events (Christodoulou et al., 2020), experiencing the last event in 2018.

According to Hirabayashi et al., 2008 the Rhine is one of the rivers that are set to witness more extreme drought conditions despite having an increase in annual precipitation. Droughts abrupt the entire stability of the Rhine catchment and endanger all the associated water users. The Rhine is crucial for many services within the Netherlands, to name a few examples; the Rhine and the Meuse provide more than 40 % of the drinking water (Sjerps et al., 2017), it is the most densely navigated shipping route across Europe and creates an annual navigation turnover of more than 5 billion euros (CCNR, 2020). There is a need for adaptive water management strategies that cater for all possible scenarios that may occur in the future to preserve the Rhine River.

1.1 The Motive: inconsistency with reality

To assess the potential impacts of these drought events on the Rhine catchment, general circulation models (GCMs) serve as a tool to foresee the future and plan accordingly. GCMs are numerical models with different emission scenarios that consider the governing physical processes in the atmosphere, ocean, land, and cryosphere. However, they provide this information at a very coarse scale (100 – 600 *kms*), which makes it difficult to use their simulations in evaluating effects at the regional scale. To overcome this issue, scientists use these GCMs as a boundary conditions at the regional level to force regional climate models (RCMs) (Sørland et al., 2018), i.e, dynamical downscaling.

RCMs simulations are widely used in impact assessment studies (Warnatzsch and Reay, 2019), as they have proven to give quite reasonable results compared to GCMs. Nonetheless, this is not always the case, as in many studies some RCMs have proven to have a lot of systematic biases (errors), giving very high or very low values when compared to the actual observations at the regional scale (Sørland et al., 2018; Pfeifer et al., 2015).

The presence of these biases in RCMs is caused by different reasons, for example because of the parameterization of the physical process of earth's climate in the parent GCM, the boundary conditions during the downscaling procedure and the applied statistical downscaling method (Reiter et al., 2018; Zhao et al., 2017; Kim et al., 2016).

Many users of the climate data (climate simulations) have demanded some form of bias correction (Glahn and Lowry, 1972). Over the recent years many bias correction methods have been developed, resulting in making bias correction a fundamental step in climate impact research (Lafon et al., 2013; Teutschbein and Seibert, 2013). There have been a lot of efforts in demonstrating the applicability of bias correction methods

in many catchments worldwide, but there are not many on the Rhine (Terink et al., 2010; Teutschbein and Seibert, 2013; Enayati et al., 2021; Pierce et al., 2015; Shrestha et al., 2017).

This research tries to assess if there is any added value in applying bias correction methods to climate simulations in the context of modelling low flows in the Rhine. In addition to whether the corrected projections can be used in real world climate adaptation and mitigation strategies.

1.2 Conceptual background

One may wonder how you can correct the bias (systematic error) in the future considering that we are not yet able to foresee. The simple answer is that we cannot fully do that. However, in the context of climate research, RCM simulations are thought of as from a time-dependent multivariate probability distribution varying in space and time (Maraun, 2016), i.e., due to randomness. The randomness is caused by the stochastic generators applied to the GCMs during the dynamical downscaling procedure to create the RCMs (Thompson and Sieber, 2012). This assumption is used to bring the probability distribution functions of future climate simulations closer to that of the historical observations, by mathematically manipulating the mean, variance, or any statistical feature of the climate simulations (Thiemeßl et al., 2012; Hempel et al., 2013). In simple terms, bias correction methods use historic observations to correct future projections based on statistics.

Bias correction can help to overcome some of the embedded errors within climate simulations, when evaluating it from a hydrological perspective (Wood et al., 2004; Hempel et al., 2013). In operational applications, hydrologists improve their hydrological model by calibration and validation. Nonetheless, when forcing the calibrated model with climate simulations to assess future impacts, the model results may sometimes seem unrealistic (i.e., very high flows in a very dry region). This is where implementing bias correction is of a value in making climate simulations more realistic and subjective to the study area.

The downsides of applying statistical downscaling methods in RCMs' simulations cannot be always averted, because most of the current climate models cannot work on resolutions less than 10 km. However, this is not limiting as many methods were proposed to preserve the main features of the change signal like quantile mapping and scaled distribution mapping (Wood et al., 2004; Jang and Kavvas, 2015; Switanek et al., 2017; Lee et al., 2019).

Another issue surrounding bias correction methods is that their working principles which conceive our climate as stationary (not varying with time), and that certainly not the case. Many studies have tested the applicability of widely used bias correction methods in preserving the change signal and extreme, and all of these studies concluded that simple bias correction methods cannot be relied upon and recommended using higher skill methods (Thiemeßl et al., 2013; Cannon et al., 2015; Reiter et al., 2018).

1.3 Research scope

In this study, two high skill bias correction (BC) methods are applied to EURO-CORDEX climate simulations that cover the Rhine basin: empirical quantile mapping (QM) and scaled distribution mapping (SDM). High skill means that they require more complicated mathematical approach, rather than just altering the signal of the mean or the variance of the projections by multiplying or addition. For further info regarding simple bias-correction methods, the reader can refer to Teutschbein, and Seibert, 2012; Acharya et al., 2013; Pierce et al., 2015; Shrestha et al., 2017. Up to the author's knowledge, the chosen bias-correction methods haven't been wrote out in the Rhine catchment, nor being tested in general against non-stationarity conditions within the Rhine.

Empirical quantile mapping or quantile mapping (QM) has gained a lot of popularity over the recent decade. One of the reasons is it can be coupled with other bias-correction techniques (Enayati et al., 2021) or fitted

within an integrated framework with stochastic weather generators and copula approaches (Li and Babovic, 2019).

Scaled distribution mapping (SDM) is a bias correction proposed by (Switanek et al., 2017) as a replacement for QM, as the latter demonstrated lack of justification for introducing inflation/deflation to the climate change signal. SDM provided a general framework to overcome the drizzling effect imbedded within climate models (Chen et al., 2021). This effect cause climate simulations to produce low precipitation rates in the dry days, i.e. overestimating precipitation rates.

There are other methods (besides bias correction) that are used by climate scientists to obtain a more robust change signal when there is a lot of uncertainty involved, known as performance-based weighting methods. These approaches constrain the change signal based on the ability of climate models to produce specific historic observations (Suh et al., 2012; Christensen et al., 2010; Chen et al., 2017; Sperna Weiland et al., 2021).

Two performance-based weighting methods are used in this thesis, reliability ensemble averaging (REA) and climate model weighting by independence and performance (ClimWIP). These methods are selected for comparing the added value of bias correction the methods proposed by the scientific climate community to strengthen the change in climate signal using raw data (uncorrected).

In hydrological modelling, bias correction methods target the climate simulations used to run hydrological models and they differ from the known model development procedures (section 3.1). The term model development procedures refer to the standard procedures used in improving hydrological models, like calibrating, validating, and refining the modelling framework. The hydrological model wflow_sbm (section 3.5) developed by Deltares (van Verseveld et al., 2020) is used to model the Rhine river, examine the added value of bias correction, and investigate potential climate change impacts in the future.

1.4 Research question

The main goal of this project is to gain a better understanding of the reliability of bias correction methods in modelling low flows in the Rhine river and assess future climate impacts at Lobith. The main research question to achieve this objective is:

'Can the impact of climate change on low flows of the Rhine river reliably be assessed at Lobith using bias corrected EURO-CORDEX projections to 2060?'

This question is addressed by answering five sub-questions which are deemed to cover most of the critical issues, within the context of impact assessment studies in hydrology.

- How do bias (un)corrected simulations perform compared to the actual flows measured at Lobith when using wflow_sbm?
- How do the incorporated bias correction methods account for non-stationarity and preserve the climate signal of the RCMs at Lobith?
- How do bias-corrected simulations perform compared to uncorrected simulations when applying performance-based weighting techniques described in Sperna Weiland et al., 2021 using the chosen EURO-CORDEX ensembles?
- What will the impact of climate change be on low flows in the Rhine river when projecting to 2060 (mid-future) after applying bias correction?
- What are the water levels corresponding to the bias-corrected future flows at Lobith?

1.5 Reader guide

The study area and data are discussed in chapter 2. The applied methods are described in chapter 3. Chapter 4 contains the results derived from these methods. These results are then discussed in chapter 5. The conclusions drawn from this research project along with further recommendations are described in chapter 6.

2| Area Description and Data

The study area is discussed first in section 2.1. The datasets used in this research are then mentioned in section 2.2.

2.1 Study Area

2.1.1 The Rhine

The Rhine flows through Switzerland, Germany, Austria, France, Liechtenstein, and the Netherlands. It serves a population of more than 50 million people who are heavily relying upon. The largest water withdrawals of magnitude by Germany, the Netherlands, and France (Wieriks et al., 1997). Many modifications are introduced in this river with the aim to provide arable land, control flooding, and ensure navigable ways. The river is protected by the convention on the protection of the Rhine (CPR) by all countries sharing the Rhine basin.

Agricultural area (arable land and pastures) constitutes about 50 % of the entire land use within the Rhine basin area, followed by forests (31.7 %) and urban areas (8.8 %) (this includes 50 cities of more than 100,000 residents). The rest of the land use is distributed between natural grassland (4.1 %), freshwater bodies (2.6 %), sparsely vegetated areas (2 %), and finally wetlands (0.2 %) (Uehlinger et al., 2009). The average discharge of the Rhine river is about 2300 [m³/s] (Uehlinger et al., 2009), and the basin has a total area of 180000 [km²]. The average annual rainfall within the entire basin is about 800 [mm/year].

Table 2-1: Land use classification in the Rhine basin

Land use class	Share (%)
Agricultural areas	50
Forests	31.7
Urban areas	8.8
Grassland	4.1
Freshwater bodies	2.6
Sparsely vegetated areas	2
Wetlands	0.2

Table 2-2: General meteorological characteristics of the Rhine basin

Characteristic	Value
Annual rainfall	487 - 1098 [mm]
Average Daily evaporation	0.5-3.75 [mm]
Mean summer temperatures	17 °C
Mean winter temperatures	0.725 °C
Total area	~ 180000 [km ²]
Average discharge	~ 2300 [m ³ /s]

Figure 2.1 illustrates how the Rhine river is distributed between the tributary countries (Strahler order ≥ 5) from Switzerland to the Netherlands.

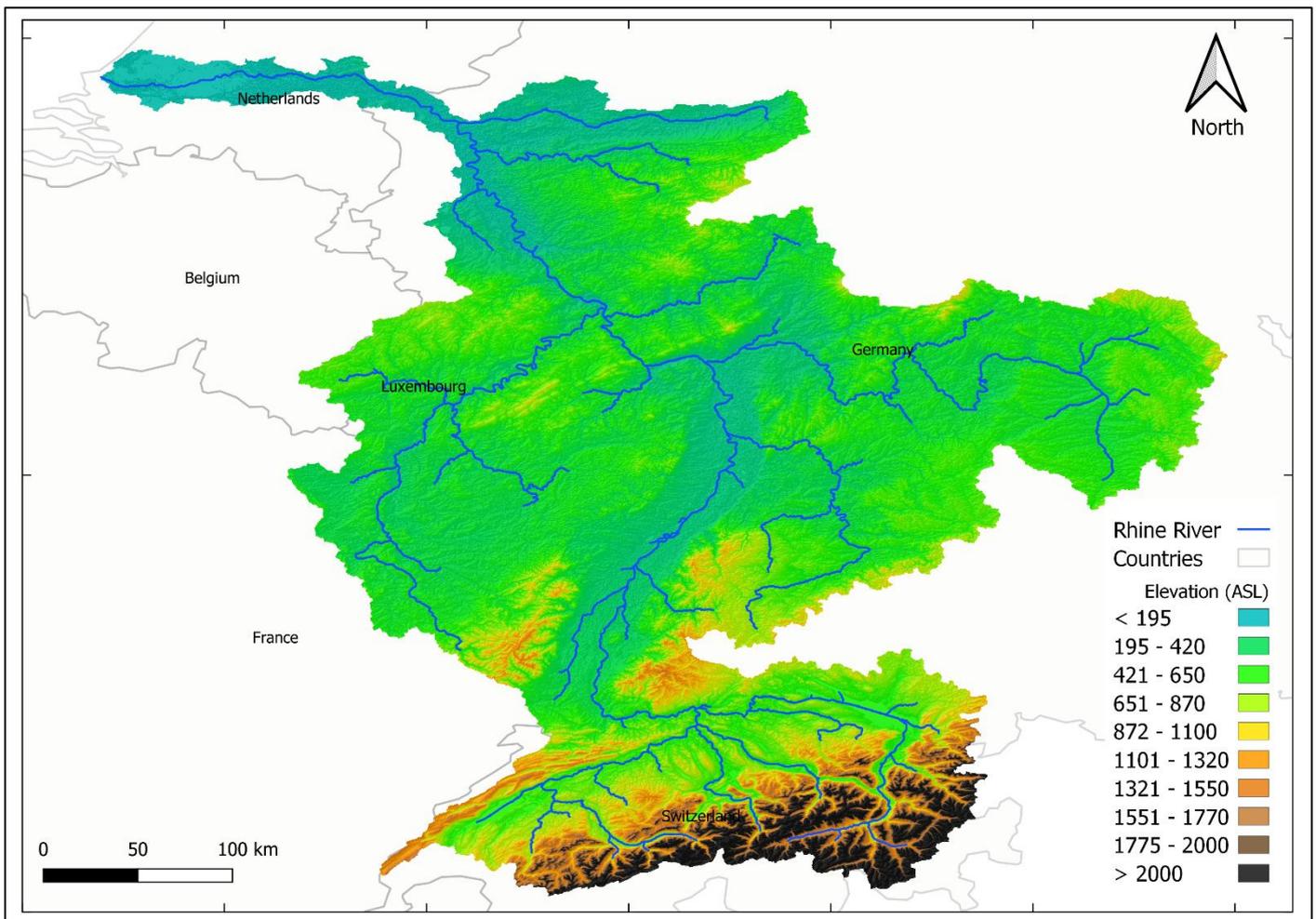


Figure 2-1: Elevation map and main tributaries of the Rhine basin

2.1.2 Lobith

Lobith is a Dutch village situated within the province of Gelderland. This village is approximately 4 km downstream of the German-Dutch border, the Rhine enters the Netherlands at Lobith. Lobith is considered the downstream point of the Rhine catchment before the bifurcations that prolong until the North Sea.

Flow statistics at Lobith have been used in many studies for discharge and extreme value analysis and form the basis for policy making on flood protection and low flows in the Netherlands (Hegnauer et al., 2014; Sperna Weiland et al., 2015). Based on these studies, Lobith is chosen to reflect on the potential future impacts of climate change in the Netherlands.

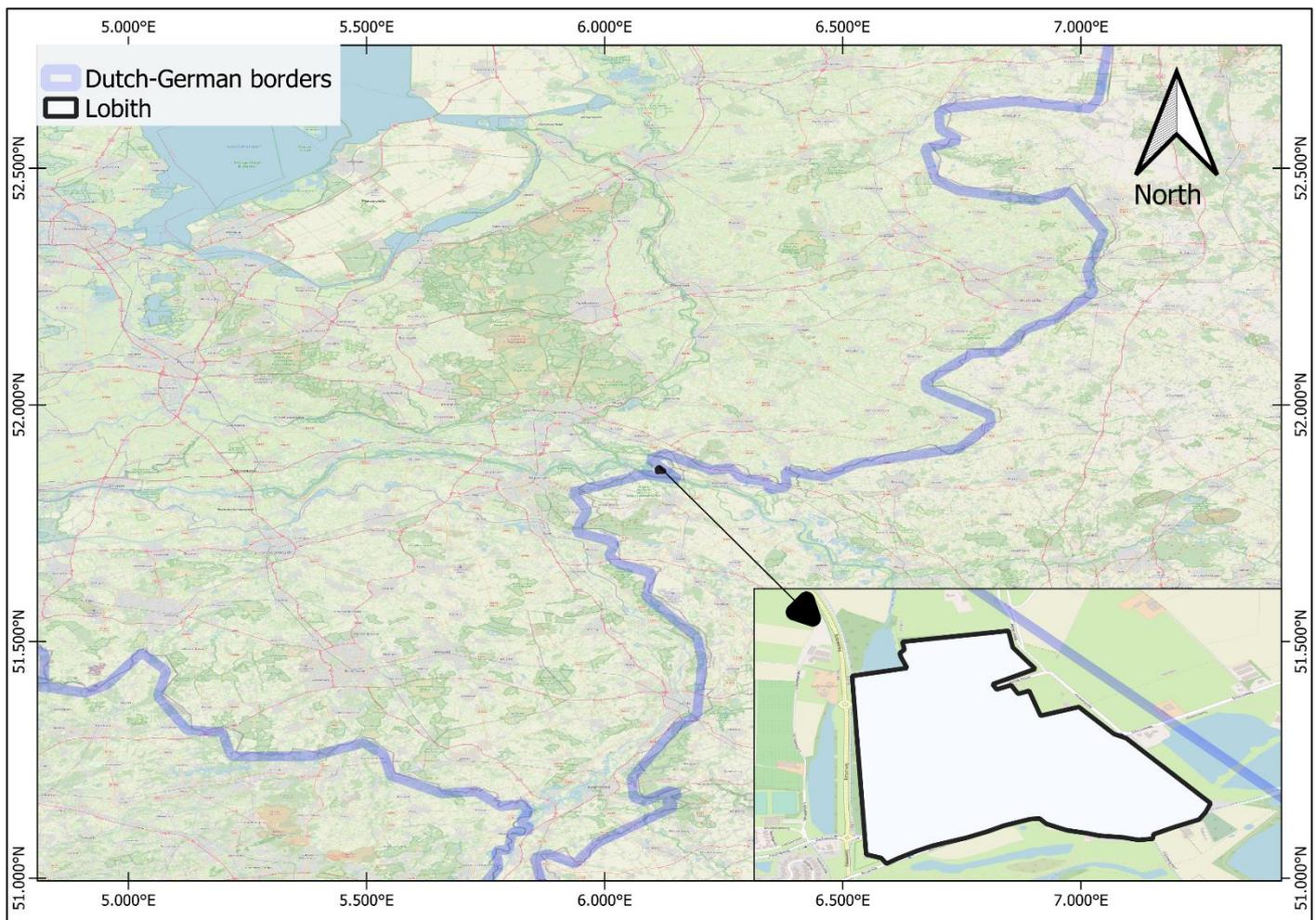


Figure 2-2: Location of Lobith within the Rhine basin

2.2 Data

2.2.1 Observations

The observational dataset is used to apply the bias correction methods. The dataset combines three meteorological variables: precipitation, temperature, and evaporation. Each one of these meteorological observations is selected from a different dataset. The three meteorological variables and their corresponding datasets are discussed in the following section.

Precipitation

The HYRAS dataset covers the entire Rhine catchment at a very fine grid resolution of 1 km², from the 1st of January 1979 till the 31st of December in 2019. HYRAS stands for High-resolution observation grid

observational datasets and it has been developed and maintained by the DWD¹ (German Meteorological Service). HYRAS contains a variety of meteorological variables. For this project only precipitation is used, for running the hydrological model used in this study to model the Rhine basin (section 3.5).

The daily (1 km x 1km) precipitation grids are created based on more than 6100 precipitation stations distributed within the Rhine basin. This is done by applying the REGNIE (translated from German to 'regionalized precipitation amounts') method and validating it using multiple methods to account for seasonal- and spatial-errors. The REGNIE method is a combination of multiple linear regression and inverse distance weighting methods that consider orographical conditions (Rauthe et al., 2013; Van Osnabrugge et al., 2017).

Figure 2.3 shows the distribution of these stations within the Rhine in 1961 and 1991, the highest differences can be seen in the eastern parts of Germany and in most of Switzerland. The figure also identifies the weather stations in the Rhine basin used to generate the HYRAS dataset.

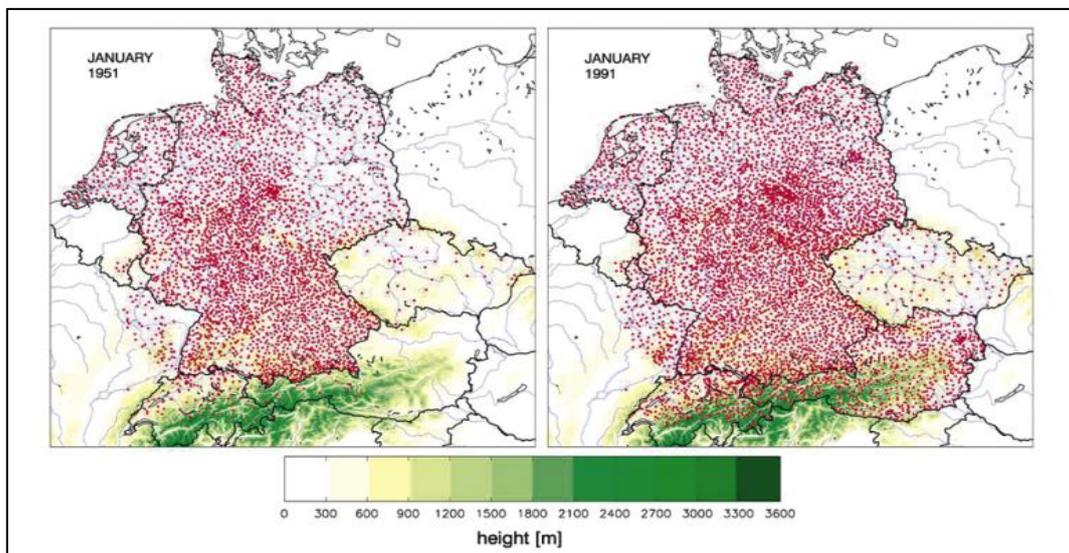


Figure 2-3: Number of gauging stations considered in creating the HYRAS dataset by DWD.

¹ DWD: Deutscher Wetterdienst in German language.

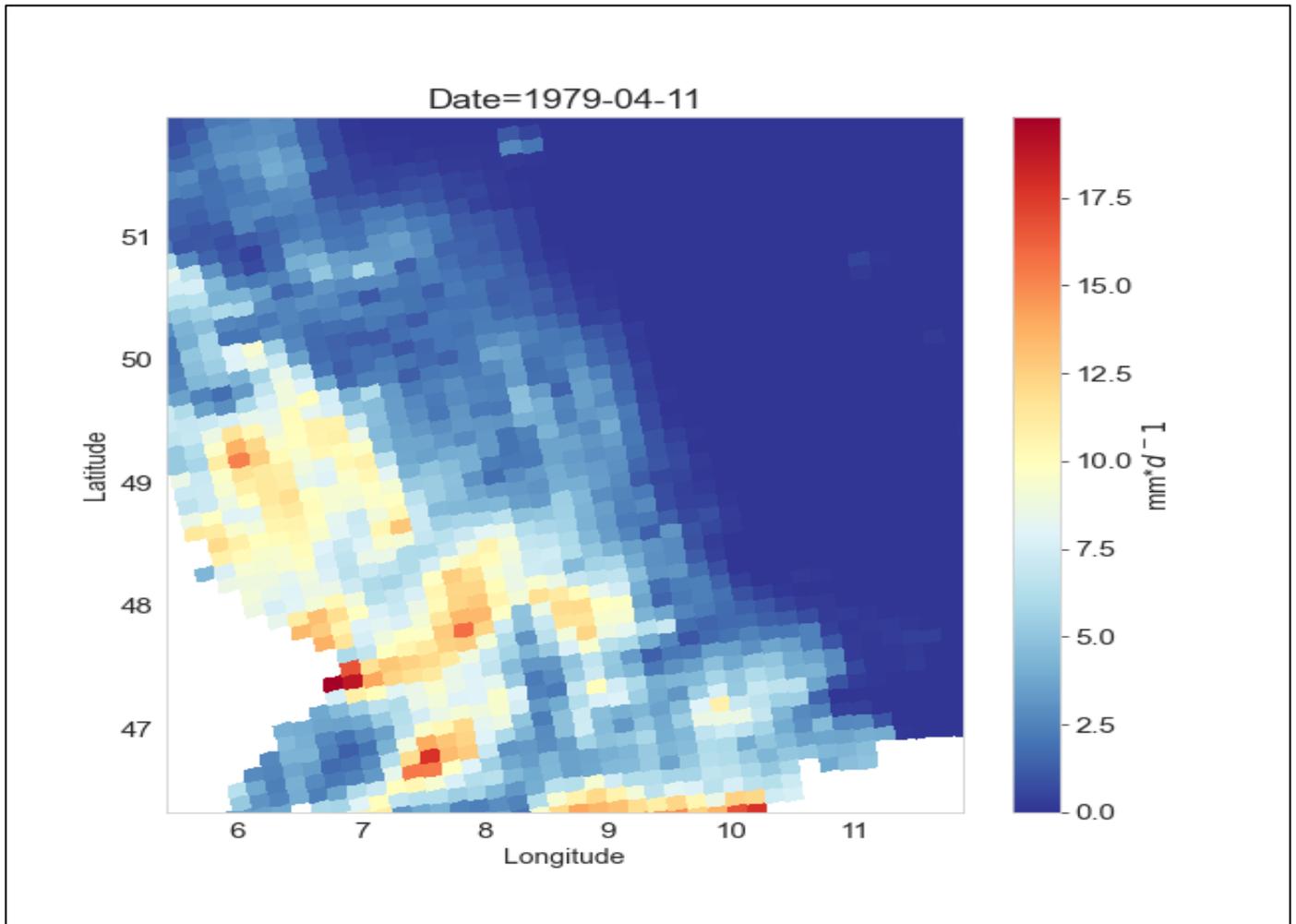


Figure 2-4: An example of the HYRAS precipitation dataset at 11/04/1979.

Temperature

Temperature was obtained from E-OBS. E-OBS is the observational dataset in Europe. It is an interpolated gridded observational dataset created using ECA&D (European Climate Assessment and data). It has a spatial resolution of $[0.1^\circ \times 0.1^\circ \cong 12 \times 12 \text{ km}^2]$ and a daily temporal frequency. The temperature is expressed in $[C^\circ]$. More information on how the dataset is created, and how it compared to other dataset can be found in (Klok and Klein, 2009; Cornes et al., 2018). Central Europe region is known for having the highest station density across all Europe, indicating a higher level of reliability to this dataset. For obtaining the hydrological model simulations the temperature data were downscaled to 1 km grid resolution using a DEM based lapse rate correction (section 3.5).

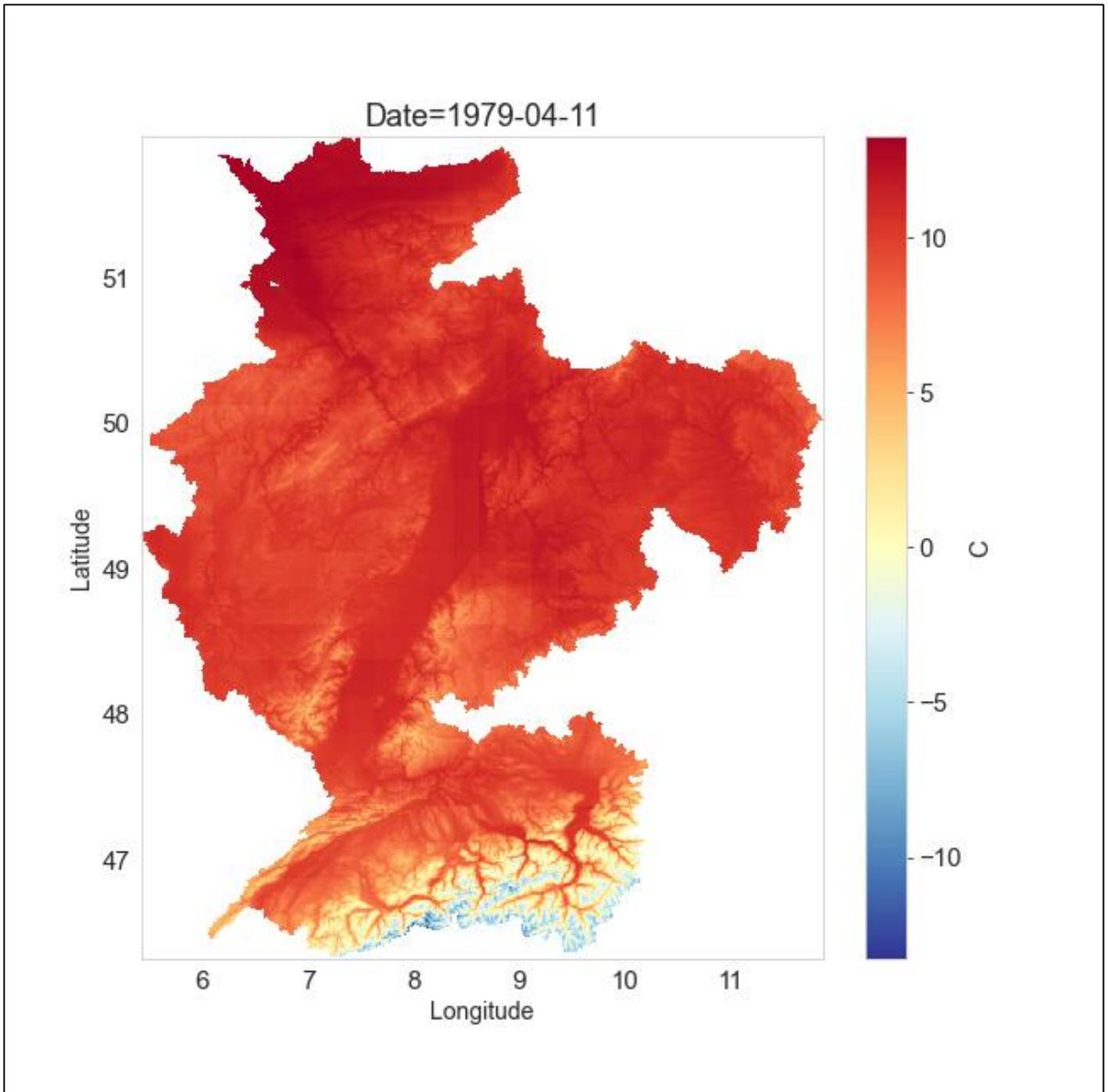


Figure 2-5: E-OBS temperature dataset at 11/4/1979.

Evaporation

The evaporation dataset considered in this project are calculated based on the Makkink evaporation formula (Makkink, 1957) applied to the ERA5 inputs (appendix I). ERA5 is the fifth-generation reanalysis data for global climate and weather, which is developed by the ECMWF (European Centre for Medium Range Weather Forecasts). ERA5 is hourly data with a grid resolution of $[0.25^\circ \times 0.25^\circ \cong 30 \times 30 \text{ km}^2]$, and spans from 1979 until now. The chosen period is between (1979 – 2019) to match the timing available in the used HYRAS and E-OBS datasets.

The reason for choosing Makkink evaporation is because the Royal Netherlands Meteorological Institute (KNMI) is also documenting its evaporation based on the Makkink equation. This is necessary to ensure that this research is following the local context of the Netherlands and the Rhine Basin.

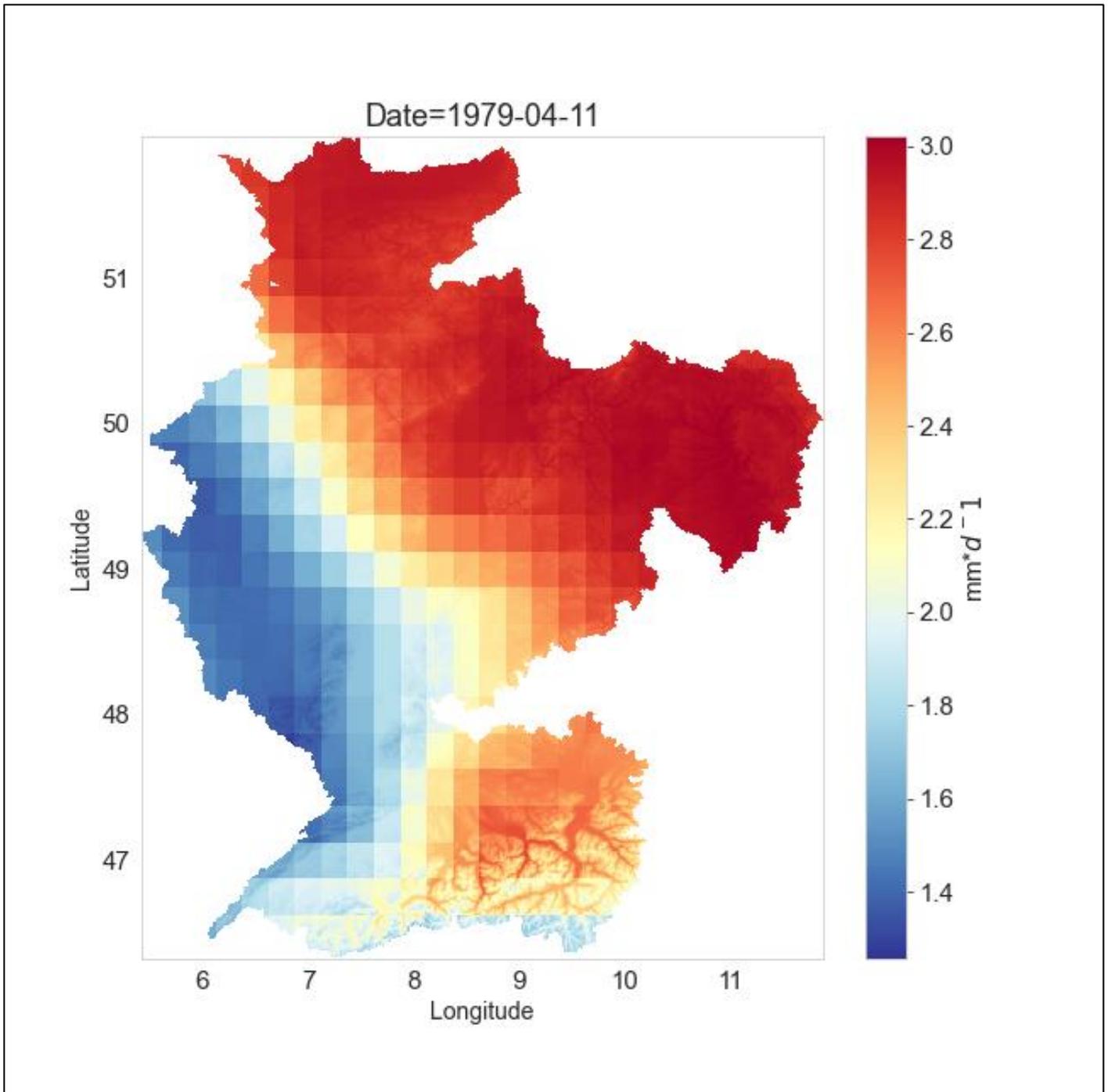


Figure 2-6: Downscaled ERA5 Makkink evaporation at 11/4/1979.

Observed Discharges

An observed discharge timeseries that spans for more than 100 years (1901 – 2019) at Lobith is used (Deltares, 2022). These discharge datasets were created by translating the measured water levels at Lobith into discharges. These daily discharges were obtained by averaging the 23-hour values from (0:00 hour till 23:00 hour). The final product after these procedures is daily discharge values expressed in (m^3/s). Observed flows at Lobith are fundamental for the steps described in the methods section. Part of the dataset is shown in the figure below, within the reference period considered in this project (1979 – 2018).

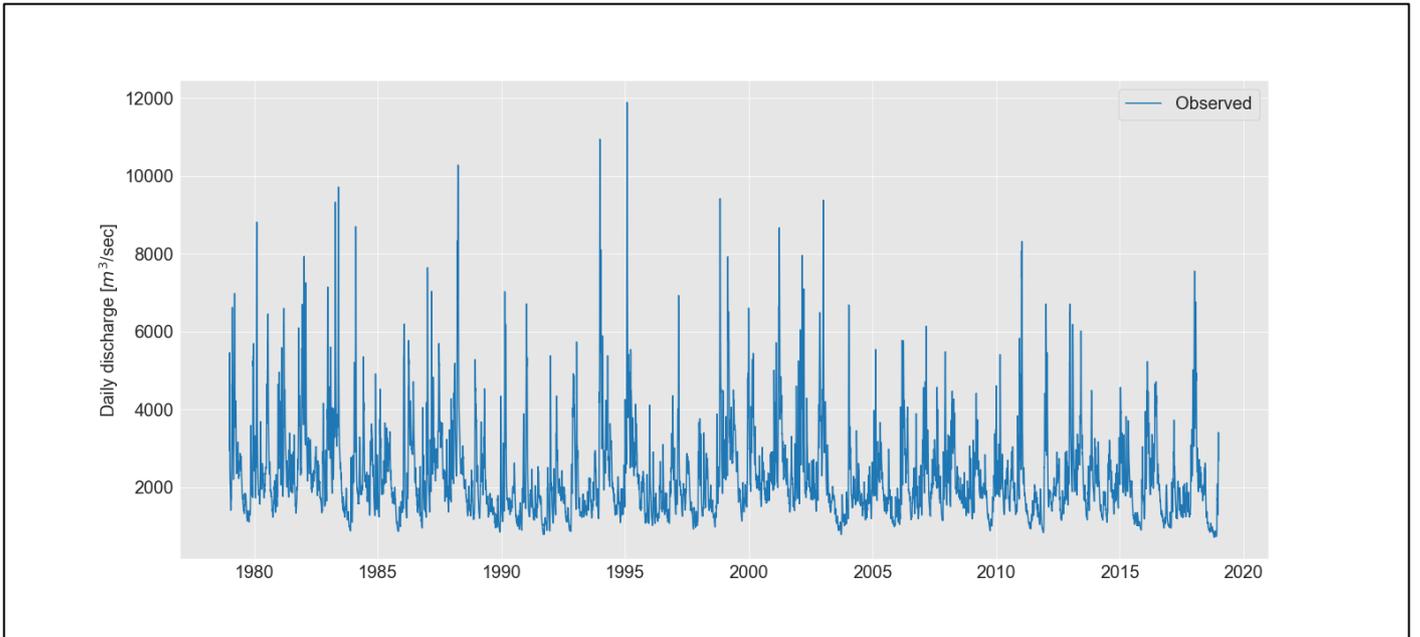


Figure 2-7: daily Observed discharges at Lobith for the period [1979 - 2018].

Additional Discharge data

For the method described in appendix (X), a discharge time series for more locations than Lobith are required to perform the calculations and compute the spatial and temporal weights within the Rhine basin. The chosen locations are Cochem, Kaub, Basel, Maxau, Worms, Koeln, and Andernach. This dataset spans from [1979 – 2000] for seven regions located within the Rhine basin (Deltares, 2022). The reason for choosing these locations is because they are distributed within the Rhine basin and their timeseries is more than 20 years, which is the minimum period for climate related studies.

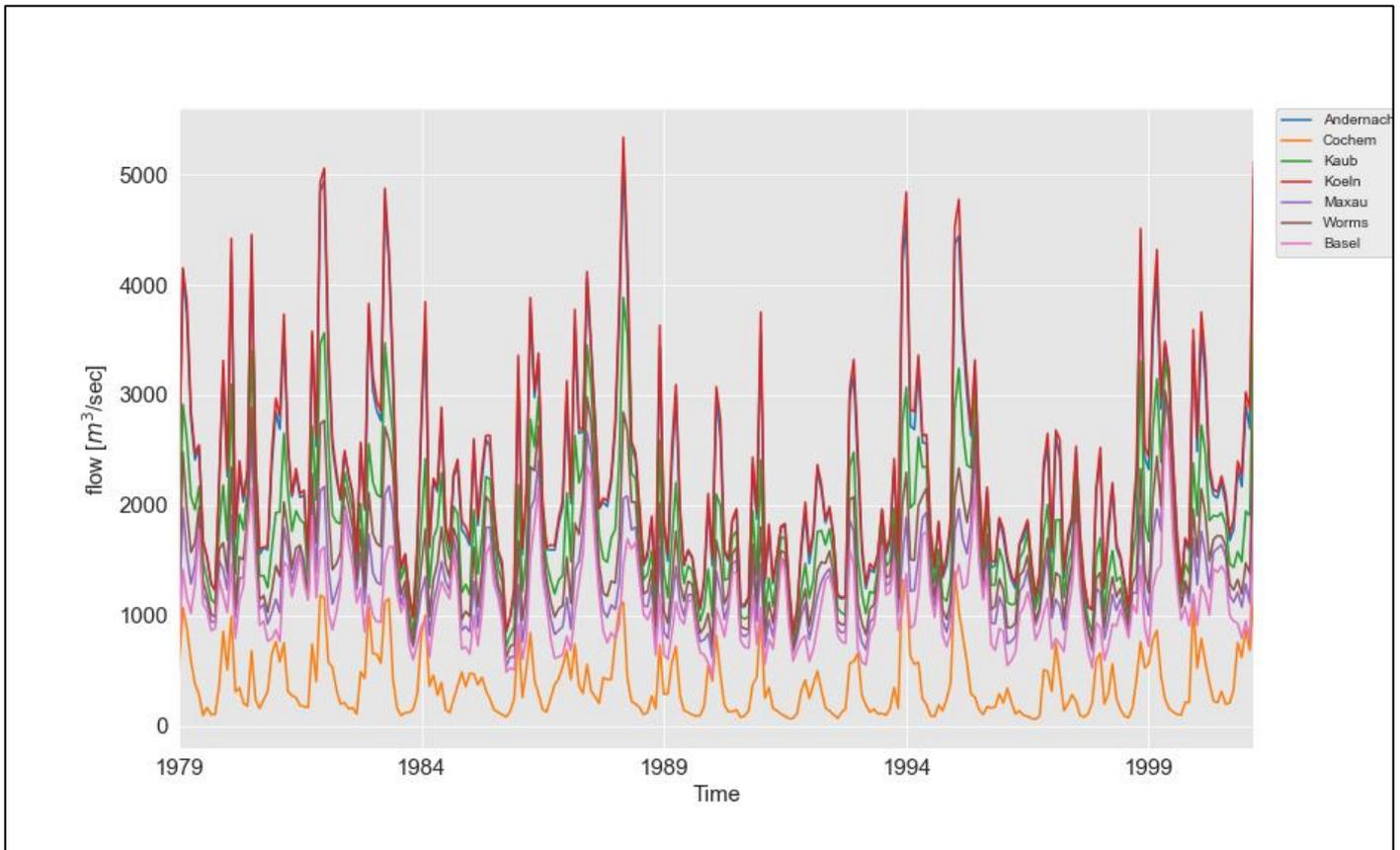


Figure 2-8: Daily discharge data for additional locations [1979 - 2000]

EURO-CORDEX

The Coordinated Regional Climate Downscaling Experiment (CORDEX) has been initiated to transfer climate change information to the local scale (Giorgi et al., 2009), through downscaling GCMs with RCMs. EURO-CORDEX is part of the CORDEX project with focus on the European continent (Jacob et al., 2014).

Within this study five different RCM subsets are chosen; these subsets are selected from twenty different RCMs. These RCMs provided the lowest error values when comparing the modelled discharges to the available observations in nine different catchments distributed within central Europe using wflow_sbm (Visser, R., 2020).

The RCM subsets are reprojected from their raw resolution (0.11 degrees) to a smaller resolution (0.05 degrees) using the hydromt package (0.4.3.dev) developed by Deltares (Beusen, B., 2021). One possible climate change scenario is considered in this project, which is the Coupled Model Intercomparison Project Phase 5 (CMIP5) Representative Common Pathways (RCP8.5) (Riahi et al., 2011). There are two main reasons for choosing RCP 8.5; the first is that remaining error sources have been found to be less when using climate scenarios with a strong change signal (Chen et al., 2017). The second reason is that it starts to become the most realistic climate scenario within the available scenarios in terms of severity and greenhouse gases emissions.

Table 2-3: List of used climate simulations (ID refers to how these subsets are referred to in the results section).

GCM	RCM subset	Period	ID
ICHEC-EC-EARTH	DMI-HIRHAM5		DMI
ICHEC-EC-EARTH	KNMI-RACMO22E	1979	KNMI
ICHEC-EC-EARTH	SMHI-RCA4	–	SMHI
MOHC- HadGEM2-ES	CNRM-ALADIN63	2060	CNRM
MOHC- HadGEM2-ES	IPSL-WRF381P		IPSL

The chosen period for these projections is (1979 – 2060), the start year is chosen as 1979 to make it similar to the used observational dataset, while the year 2060 is chosen based on earlier research that demonstrated that error sources in climate models and their internal climate sensitivity remain logical when projecting to the near future (Chen, et al, 2020; Ehret et al., 2012; Teutschbein and Seibert, 2012). Near future refers to projecting up to 40 years ahead.

These climate simulations contain daily temperature rates in [K], and precipitation in [Kg/ m² /s]. However, they do not directly contain evaporation. To overcome this issue the Makkink formula (Makkink, 1957) is used to calculate evaporation using temperature, pressure [Pa], and incoming downward shortwave radiation [W/m²] as inputs. Temperature had to be expressed in [Celsius] and pressure [hPa] to use the Makkink formula outlined in appendix I.

3| Methods

Two bias-correction methods are applied to the chosen climate simulations. The two methods are discussed in section 3.1 in addition to showing the different mechanisms when correcting each meteorological variable. Section 3.2 discusses how these bias-correction methods are evaluated based on the historical experiment. The applied test for investigating the reliability of bias correction methods in changing climate conditions is described in section 3.3. The applied performance-based weighting techniques are mentioned in section 3.4. Finally, the hydrological model is demonstrated in section 3.5.

3.1 Bias-Correction

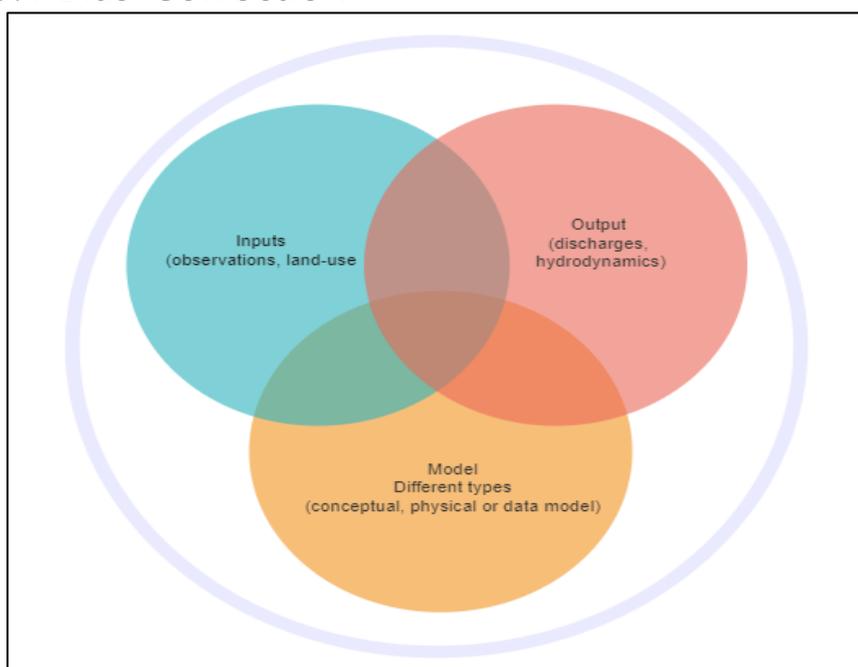


Figure 3-1: Typical model chain in hydrology.

Figure 3.1 describes the general chain in modelling hydrological systems. The general goal of modelling is to predict a certain variable(s) in the future and use these modelled values in different uses (management, assessment, etc.). The system consists of three main attributes: inputs or forcing's that are used by the model, the model that is used to generate predictions and the model outputs. Bias correction deals with inputs that are used by hydrological models. This is different from the standard procedures which is referred to as model development (calibration or validation). The two bias corrections methods adjust the simulations obtained (RCMs in this case).

The type of observations used in this project are gridded datasets (see figure 2.5, 2.6 and 2.7). Each grid cell covers a spatial domain, this domain can be of any size and can be altered based on the user's interest. Climate simulations are in the same gridded format as the observations.

Figure 3.2 outlines the basic structure of climate simulations. The picture denoted as A, outlines the layout of gridded datasets, in each grid cell an average timeseries of a specific meteorological variable is stored (see picture B). This timeseries represents the observations at the size of the grid cell and can be of any time resolution (hourly, daily, monthly, etc.).

Bias correction methods work on a general principle to correct the systematic bias, which is using the historical run to correct the future run (projections in this case). The actual observations and the historical climate simulations are both considered in the historical run, while the future climate simulations are considered in the future run. The reference period is the term used to refer to the historical run and the projection period refers to the future run.

The correction methods utilize the working principles of climate models, as the simulations obtained from climate models are random variables distributed over time at each grid cell. Climate simulations shown in picture C of figure 3.2 can be expressed in the form of cumulative distribution function as outlined in picture D of the same

figure. This provide the possibility to adjust the statistical properties of climate simulations and correcting them, referred to as bias correction of climate simulations. Empirical quantile mapping and scaled distribution mapping are bias correction methods that correct future simulations based on historical observations.

In the following subsections the two bias correction methods used in this research are discussed: empirical quantile mapping and scaled distribution mapping.

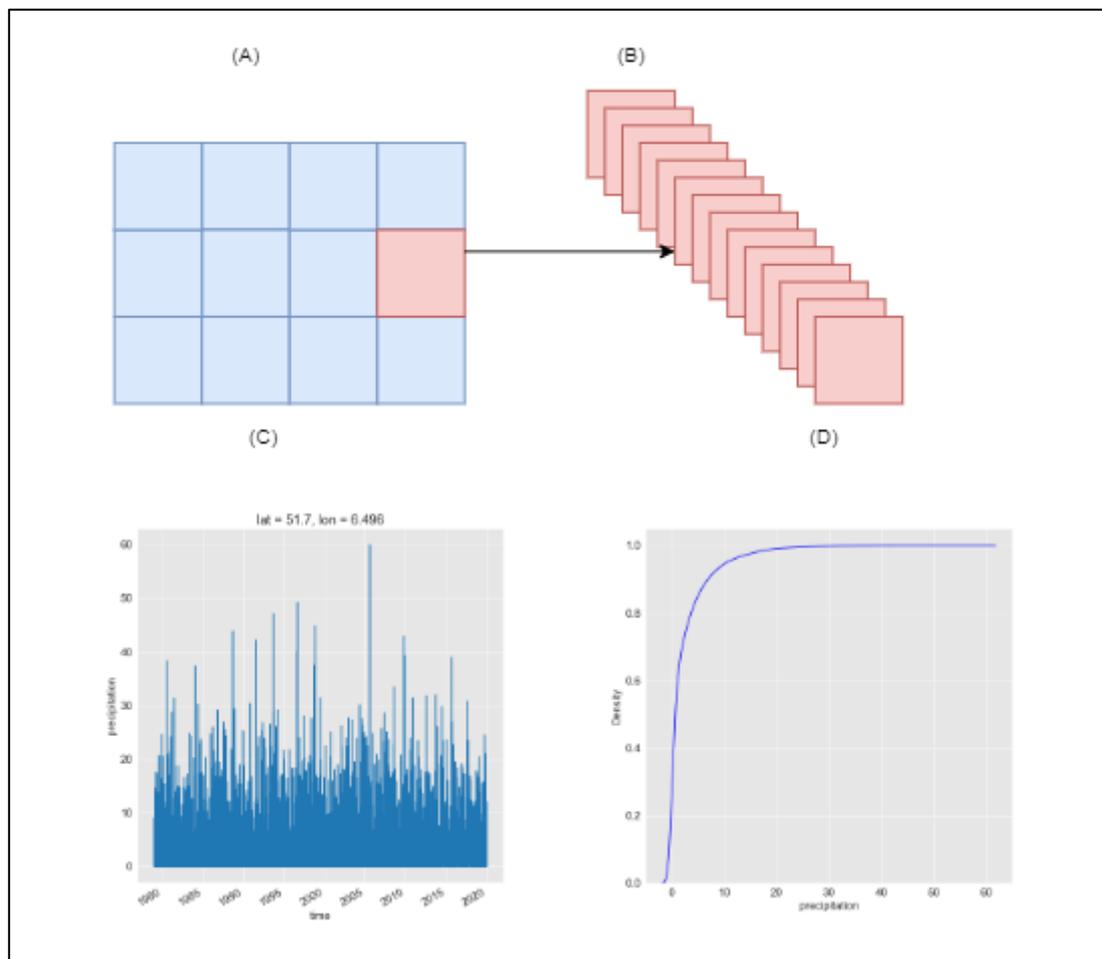


Figure 3-2: general description of climate simulations

3.1.1 Empirical Quantile Mapping

Empirical quantile mapping is essentially mapping of projections probabilities based on the respective probabilities of the observed datasets and historical simulations. The goal of this method is to bring the cumulative probability distribution function of these future projections closer to that of the observations by adjusting it using the cumulative distribution function of the observational datasets. The general formula for applying quantile mapping is listed in equation 3.1.

$$xc(t)_{MODF} = ICDF_{OBS}(CDF_{MODH}(x(t)_{MODF})) \quad (3.1)$$

Where $xc(t)_{MODF}$ is the bias-corrected future climate simulation for the meteorological variable x at time step t . $x(t)_{MODF}$ is the raw projected (future) climate simulation for the meteorological variable x at time step t . $ICDF_{OBS}$ is the inverse cumulative distribution function derived from of the observed time series for the meteorological variable x at the specific grid cell. CDF_{MODH} represents the cumulative distribution function of the raw climate simulation during the reference run (historical run, which is the same reference run in the observation) for x .

It should be noted that, equation (3.1) is applied to each grid cell within the catchment independently, which also considers the spatial heterogeneity within the Rhine, depending on the quality of the observational data. This method has demonstrated a good performance in some regions around the globe, but less so in other regions (Enayati et al., 2021; Reiter et al., 2018; Li and Babovic, 2019; Zhao et al., 2017).

However, most of the available literature suggests that this method performs better when applied to sub-annual scales or seasonal scales (Zhao et al., 2017; Kim et al., 2016). For this reason, the general procedure described below has been refined, to be applied on the monthly scale.

The following equation shows how the general empirical quantile mapping has been altered, instead of considering the entire time series to construct the inverse cumulative distribution, and cumulative distribution functions. Through this method, the projected data that's falls within a certain month will be analysed and corrected based on the corresponding time series in the observation dataset.

$$xc(t)_{\text{MODF}, m} = \text{ICDF}_{\text{OBS},m}(\text{CDF}_{\text{MODH},m}(x(t)_{\text{MODF},m})) \quad (3.2)$$

Where $xc(t)_{\text{MODF},m}$ is the bias-corrected future climate simulation for the meteorological variable x at time step t within month m . $x(t)_{\text{MODF},m}$ is raw projected (future) climate simulation for the meteorological variable x at time step t within month m . $\text{ICDF}_{\text{OBS},m}$ is the inverse cumulative distribution function derived from of the observed time series for x at that grid cell for month m . Finally, $\text{CDF}_{\text{MODH},m}$ is the cumulative distribution function of the raw climate simulation during the reference run (historical run, which is the same as the used reference run in the observation) for x on the month m .

The corrected simulations are then regrouped into one timeseries based on their temporal index (day, month, and the year) and grouped into one dataset, using the mergetime feature in CDO (Climate Data Operators package installed on a Linux machine).

$$xc_{\text{MODF}} = \text{sort}((xc_{\text{MODF}, m1}, \dots, xc_{\text{MODF}, m12}), \text{time}) \quad (3.3)$$

An example of how this method is applied, suppose that the observation dataset and raw simulations for the period (1979 – 1998) are used to correct the systematic bias from the simulations for the period (1999 – 2018) using this method. The reference run consists of the daily observations and raw simulations that occur within the period (1979 - 1998), while the projected run consists of the raw simulations for the period (1999 – 2018).

First, each dataset is divided into 12 datasets (12 month). Each dataset contains the simulations that fall within a specific month, the first dataset is all simulations or observations that occur within the that month (January in this case). In other words, for the case of observations all the observations that have occurred between the first of January to the 31st of January regardless of the year in which the observation took place. Meaning that all the observations that occur in the period (1979 – 1998) in January are listed in separate dataset, the same procedure is repeated in the other 11 months. The same applies for EURO-CORDEX simulations.

Then, these future simulations are corrected from the systematic bias based on equation 3.1. After the correction, the 12 corrected datasets are then merged back into one dataset. This is done by CDO package in Linux.

The CDO package can merge these 12 corrected datasets in one dataset using the mergetime feature. This feature sorts the corrected datasets using the associated time index for each simulation. The bias-corrected dataset for January contains all the corrected simulations that are associated with January, every simulation that occur between the first to the end of January for the period (1999 - 2019). CDO uses the full associated time index to sort the corrected datasets in an ascending order.

An example of the effect of applying this method at a specific grid cell can be seen in figure (3.3). Where the cumulative distribution function of the observed, raw, and future timeseries for each month are shown. In 9 of the 12 months the CDF of the corrected simulations is closer to the CDF of the observed dataset than the CDF of the raw (uncorrected) simulations.

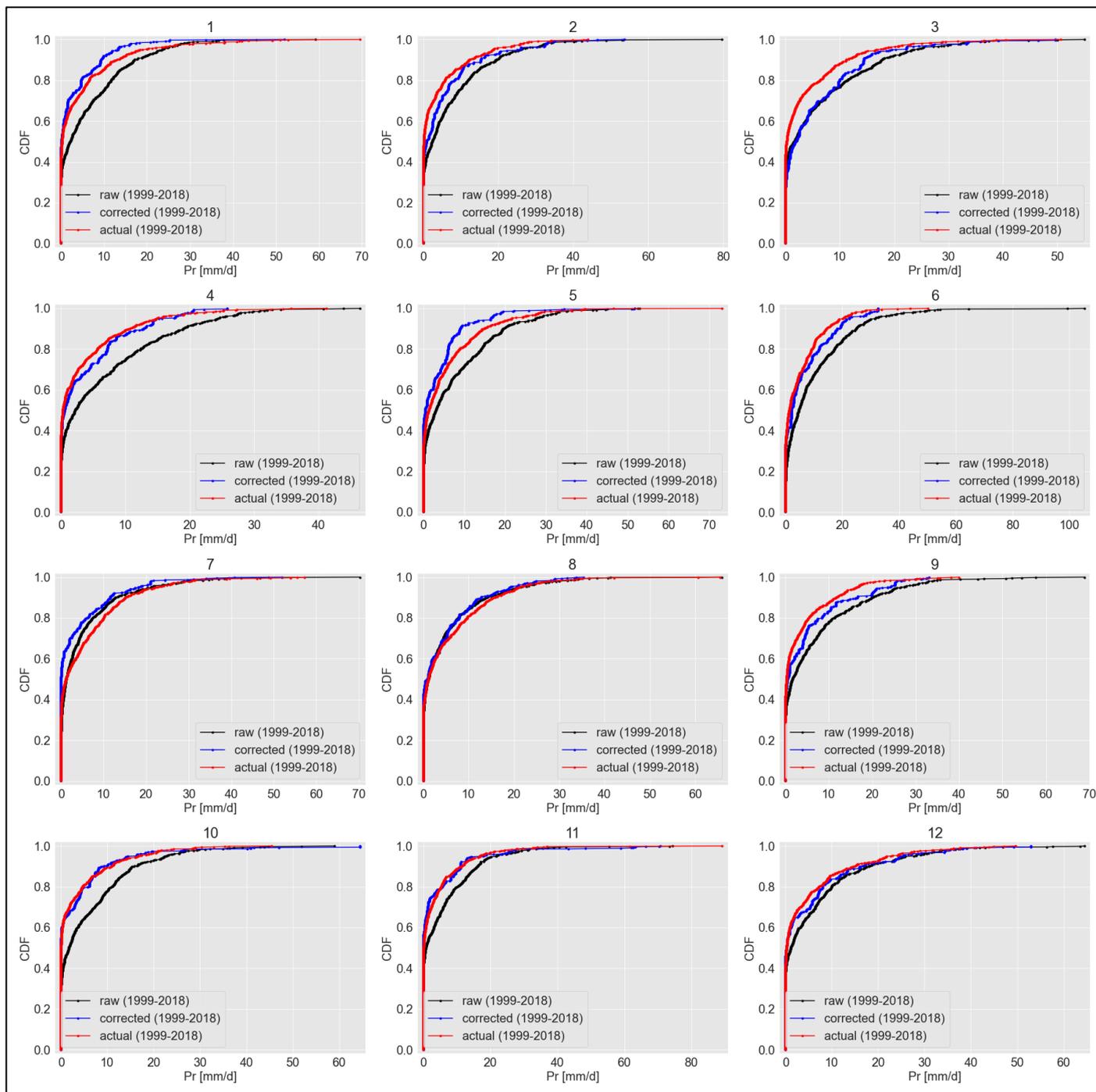


Figure 3-3: representation of quantile mapping correction on a monthly scale

The empirical quantile mapping is used to correct the selected subsets, each subset contains the three meteorological variables: precipitation, temperature, and evaporation. Then, the bias-corrected subsets are used as inputs into the wflow_sbm model (section 3.5). There is no special consideration for each meteorological variable like in the scaled distribution mapping.

3.1.2 Scaled Distribution mapping

Prior to explaining how SDM is applied to the climate simulations, it is important to mention that correcting precipitation and evaporation variables using SDM depends on the fitted statistical distribution. The chosen distribution should be site-specific and suitable for the extremes under consideration. For correcting precipitation and evaporation projections gamma distribution is chosen, while for temperature correction normal distribution (Gaussian) is the selected distribution.

Precipitation

One of SDM main advantages is that it limits the drizzling effect within the climate models, while preserving the change signal as much as possible. To achieve this, a threshold must be set, here a threshold of 0.1 mm is chosen as in (Switanek et al., 2017).

The correction of precipitation is performed using the pyCAT package, which is a climate analysis tool written in python and developed by Wegener Center for Climate and Global Change. However, for SDM methodology to be implemented, seven different steps are required to compute the bias-corrected climate simulations:

First step is applied after setting the minimum threshold (0.1 mm), values that are lower than the threshold are set to zero (non-rainy days). The number of days with rain events can be distinguished from number of days without any rainfall event. The number of rainy days in the bias corrected future runs (RD_{BC}) can be expressed in equation 3.4. The term scaling in this method, assumes that the fraction of rainy days in the future is scaled based on the ratio of the fraction of rainy days during the observation to the fraction of rainy days in the historical run.

$$RD_{BC} = RD_{MODF} * \left(\frac{RD_{OBS}/TD_{OBS}}{RD_{MODH}/TD_{MODH}} \right) \quad (3.4)$$

TD_{OBS} , TD_{MODH} represent the total number of days (non-rainy days are included) in observation and the historical run respectively. RD_{OBS} , RD_{MODH} , and RD_{MODF} are the number of rain days in observation, historical run, and future run respectively.

In the Second step, the gamma distribution is fitted to the precipitation values. The gamma distribution applied probability density function is described in equation 3.5.

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad \text{for } x > 0 \quad (3.5)$$

Where $\alpha > 0$, $\beta > 0$ are the shape and scale parameters, $g(x)$ is the fitted gamma distribution, $x (> 0)$ is the precipitation values, and $\Gamma(\alpha)$ is the gamma function evaluated by the shape parameter.

Then, cumulative distribution function (CDF) is constructed after fitting the probability distribution shown in equation 3.5. Three CDFs are obtained: the CDF of the HYRAS dataset, the raw simulation in the reference run and the future raw simulations in the future run. Then, these CDFs are inverted to obtain the inverse cumulative distribution functions (ICDFs) of the same three datasets.

In the third step the scaling factor (SF_R) is calculated by dividing the fitted raw future model run to the fitted historical run through the raw CDF of the future run, by using the CDFs and ICDFs obtained from the earlier step.

$$SF_R = \frac{ICDF_{MODF}(CDF_{MODF})}{ICDF_{MODH}(CDF_{MODF})} \quad (3.6)$$

Where $ICDF_{MODF}$ and $ICDF_{MODH}$ are the inverse cumulative distribution function for the fitted future run and for the historical run of climate simulations. CDF_{MODF} is the cumulative distribution function of the future raw simulations.

The Fourth step is to calculate the recurrence interval (RI) for the three runs (observation run, historical run, and future run). The recurrence interval (equation 3.7) is similar to how return periods are calculated in the hydrological context.

$$RI = \frac{1}{1-CDF} \quad (3.7)$$

For comparing recurrence intervals across the entire distribution, linear interpolation is performed based on the historical run and the observation (RI_{IMODH} , RI_{IOBS}). This aim of this step is to compare the recurrence interval within the full distribution; therefore, the CDF is used for this purpose. These RI values are then used to scale the number of the rainy days in the future run.

The Fifth step is to calculate the scaled recurrence interval RI_{SCALED} for the future run, through equation 3.8. This step is necessary to ensure that the CDF in the remaining steps is constrained between 0 and 1.

$$RI_{SCALED} = \max \left[1, \frac{RI_{IOBS} * RI_{MODF}}{RI_{IMODH}} \right] \quad (3.8)$$

Then, scaled recurrence interval is used to compute the scaled cumulative distribution function, by inverting equation 3.7. The result is shown in equation 3.9.

$$CDF_{SCALED} = \frac{1}{1-RI_{SCALED}} \quad (3.9)$$

In the sixth step an initial array of bias corrected future values is constructed by using equation 3.10. The array is constructed based on the scaling factor (equation 3.6) and the scaled cumulative distribution function (equation 3.9).

$$BC_{INITIAL} = ICDF_{OBS}(CDF_{SCALED}) * SF_R \quad (3.10)$$

For the final Step the bias corrected values are sorted based on their temporal index and place them as the modelled time series, by doing so the final bias corrected future run is constructed. This is to ensure that the largest bias-corrected values are placed into their original timestep.

All the steps for correcting the precipitation simulations are listed in figure (3.4). The correction procedure is based on statistics. This correction method is designed for correcting climate model simulations that overestimate the number of rainy days and not designed for correcting climate simulations that underestimate the number of rainy days. For the latter case, this method will result in more underestimation of rainy days in the climate simulations.

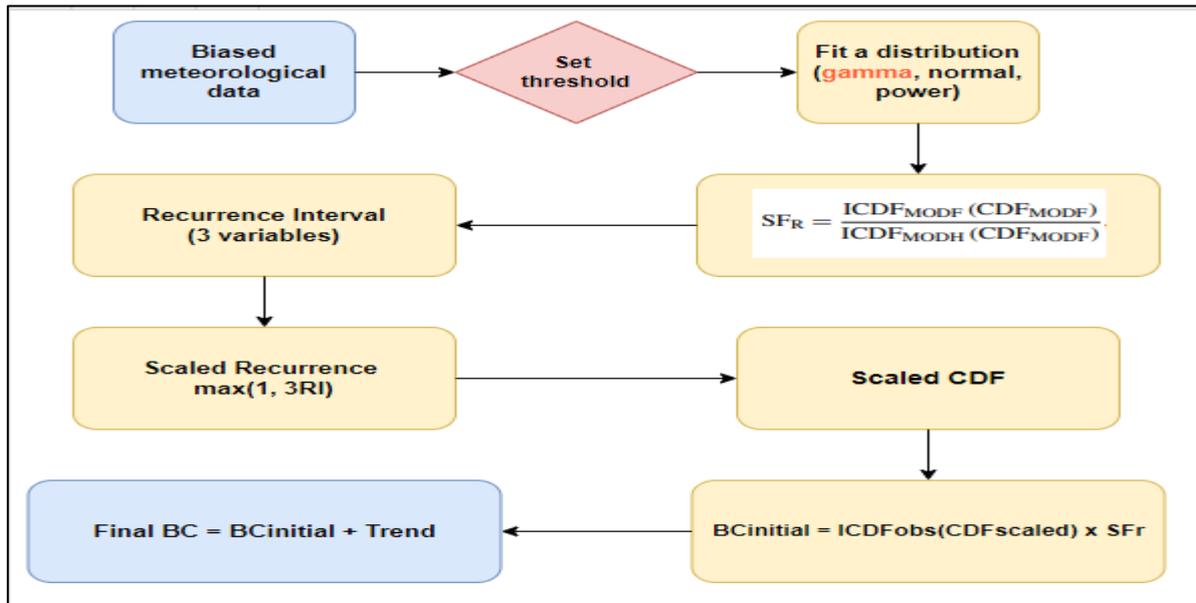


Figure 3-4: General framework for correcting precipitation projections using SDM.

Temperature

Temperature correction procedure differs from the precipitation correction procedure, this is because temperature follows a normal distribution (Harmel et al., 2002). The procedure for correcting the systematic bias from future temperature simulations is described in table 3.1 (Switanek et al., 2017).

Table 3-1: Temperature correction procedure using SDM

Step	Procedure
1	The first step is to detrend the observed, historical run, and future run time series using the linear detrending method. This is to ensure the correction of the future simulations is based on the natural variability in the observation dataset. The suggested modifications from step 2 to step 6 use the detrended datasets.
2	The normal (Gaussian distribution) is fitted to all the detrended time series (observation, historical run, and future run). The results of this fitting are the empirical mean and standard deviation. Three CDFs are then constructed based on the fitted normal distribution (CDF_{OBS} , CDF_{MODH} , CDF_{MODF}). Based on these CDFs, two inverse cumulative distribution functions are derived for the historical raw simulations and the future raw simulations $ICDF_{MODH}$, $ICDF_{MODF}$.
3	The scaling factor (SF_A) between the fitted raw future model distribution and the fitted raw historical distribution for each event within the future run time series is calculated. This is done by applying equation 3.11. Keep in mind it is different than the scaling factor for the case of precipitation (equation 3.6). The scaling factor for temperature is the absolute difference between the future run and the historical run based on the variations in the change signal of temperature. $SF_A = [ICDF_{MODF}(CDF_{MODF}) - ICDF_{MODH}(CDF_{MODF})] * \left(\frac{\sigma_{OBS}}{\sigma_{MODH}} \right) \quad (3.11)$ σ_{OBS} , σ_{MODH} are the standard deviation values calculated from the observed and raw historical time series.
4	Calculate the recurrence intervals (RI) for the detrended time series (observation, historical run, and future run) using equation 3.12. $RI = \frac{1}{0.5 - CDF - 0.5 } \quad (3.12)$

	The reason for why it is different than equation 3.7, is to reflect on the tail's nature in the normal distribution. This is not the case in the gamma distribution used for precipitation correction.
5	<p>The scaled recurrence interval (RI_{SCALED}) for the future run is used to limit the scaled cumulative distribution function between 0 and 1, this is necessary to ensure meaningful values. The recurrence intervals (RI) for the observation, historical run, and future run in step 4 (equation (3.12)), are then used to compute the scaled recurrence interval (RI_{SCALED}) for the future run using equation 3.8.</p> <p>Then final scaled cumulative distribution function (CDF_{SCALED}) is calculated by using the cumulative distribution function of the observations and the scaled recurrence interval (RI_{SCALED}) through equation 3.13.</p> $CDF_{SCALED} = 0.5 + CDF_{OBS} - 0.5 * (0.5 - \frac{1}{RI_{SCALED}}) \quad (3.13)$
6	<p>The initial array that contains bias-corrected future run values is constructed by using equation 3.14. The scaling factor (SF_A) is the correction value after scaling the future run based on the available observation.</p> $BC_{INITIAL} = ICDF_{OBS}(CDF_{SCALED}) + SF_A \quad (3.14)$
7	<p>The initial bias corrected array is then sorted based on the temporal index. The result is modelled values that are placed on their correct timestep. The purpose of this step is to ensure that the largest bias-corrected temperature values are placed into their original timestep. The final bias-corrected dataset is formed by adding the removed trend in step 1 to the sorted initial bias corrected array. Equation 3.15 describes the general process</p> $BC_{Final} = Sort(BC_{INITIAL}) + Trend \quad (3.15)$

Evaporation

Evaporation correction is not considered by (Switanek et al., 2017), however it is an essential part of modelling the Rhine river using wflow_sbm (section 3.5). The EURO-CORDEX simulations

The reason for correcting evaporation instead of deriving it using the bias-corrected temperature, raw projected radiation and raw projected pressure is to prevent combining corrected variables with raw variables, to ensure a fair comparison between raw projections and bias-corrected projections.

Evaporation is corrected using the same steps (gamma distribution) described in the precipitation section with only a minor change. This minor change is introduced because we cannot discard minimum evaporation values from the evaporation datasets because of thresholds, as evaporation in general depends on many other factors (water availability, humidity, incoming solar radiation, etc.).

The applied bias correction procedure for correcting evaporation simulations is the same as the precipitation procedure with a threshold. This step is not applied through the pyCAT package because of the size of the datasets used in this project and setting no threshold to do the same procedure was not possible in pyCAT. To overcome this issue a separate python code that is written based on the same steps outlined by Switanek et al., 2017 in precipitation correction (no minimum threshold) is used. The code combines the use of universal function feature (apply_ufunc) in Xarray library and Dask library to fasten computational time.

3.2 Historical experiment

A historical experiment refers to using the two correction methods to correct the historical climate simulations and assess it based on the recorded measurements.

The period of [1999 – 2018] is assumed to be the future run to be corrected, while the period between [1979 – 1998] is used as the reference run to correct the future run. The choice of choosing these periods is based on two considerations, the first for analysing climate change impacts the minimum period is 20 years. The second factor is these methods are more tested when they have lesser observations (i.e., using 30 years to correct 10 years is thought not to be an accurate representation for this test, as 10 years is lesser than 20 years).

The period between (1979 – 1998) is considered as the reference run and the period between (1999 – 2018) is the future run. For the reference run the raw simulations (no correction) and observational datasets that fall within the same time are used. For the future run the raw simulations for the period (1999 – 2018) are the ones to be corrected. QM and SDM are applied to the same period, to ensure comparability. Table 3.2 shows the periods that are considered in this experiment.

The bias-corrected and raw simulations are used to generate the modelled discharges at Lobith for the period (1999 – 2018). The assessment of the bias-corrected and raw simulations is based on their accuracy relative to the actual discharges at Lobith for the same period (1999 – 2018). This is done by comparing the average daily discharges for the same period (1999 – 2018). The average daily discharges for the period (1999 – 2018) are derived by averaging all the discharges that occur at a certain day within the 20 years period.

Furthermore, the modelled discharges from the raw and corrected simulations are compared relative to the actual discharges at Lobith based using four long-term statistics. These statistics are the long-term annual mean flow, long-term summer mean flow, 7-day minimum flow and lengths below thresholds (appendix I).

The actual meteorological observations at Lobith are used to investigate some of the changes in the shape of the hydrograph before and after the correction relative to the actual observed discharges.

Table 3-2: Different periods in the historical experiment

Reference run		Future run	
Observed data	1979 - 1998	Raw climate simulations	1999 - 2018
Raw climate simulations	1979 - 1998	Bias-corrected simulations	1999 - 2018
		Actual observations	1999 - 2018

3.3 Non-Stationarity under scope

Differential split sample test (DSST) was originally proposed by (Klemeš, 1986) to systematically assess hydrological model transposability. In hydrological modelling, DSST is implemented to test the robustness of a hydrological model in providing reasonable predictions in changing climate conditions. This is done by calibrating the model based on the wet periods and evaluating its performance during the dry periods. Then, the model is calibrated based on the dry periods and evaluated based on wet periods.

DSST was adopted for the first time by Teutschbein and Seibert, 2013 to test the ability of bias corrections methods to account for non-stationary conditions. They have evaluated six bias correction methods based on the ability to correct annual values. To implement the DSST test, they have sorted the years from the driest to the wettest and then applied the correction, and vice versa.

The variations in the annual values due to climate change is low compared to the variations in finer time resolutions like daily, monthly, etc. In this research, DSST is applied differently to evaluate the ability of the

empirical quantile mapping method and scaled distribution mapping method in correcting different seasons. Two different seasons are considered in this project: The period between June to August (known as JJA) and the period between December to February (known as DJF). The selected period on which the correction is performed is the same period discussed in the historical experiment (section 3.2). In bias correction methods, the reference run is used to correct the future.

The DSST test is applied on the JJA and DJF periods. QM and SDM are used to correct wet periods in the future based on dry periods in the reference run. The assessment is based on the actual observations during the same projected period. Therefore, the bias correction methods are applied based on periods outlined in table 3.2, the only difference here the focus is on seasons and not on the entire dataset.

For the periods listed in table 3.2, the values that fall between the first of June to the end of August (JJA) in the observation dataset, raw and future simulations are listed in separate datasets. The same is applied for the values that fall between the first of December to the end of February (DJF) in the observation dataset, raw and future simulations (see figure 3.5).

It is worth noting that precipitation and temperature are the only variables considered in this test. Evaporation is excluded because the ERA-5 Makkink evaporation dataset and the Makkink evaporation for the climate simulations are calculated differently.

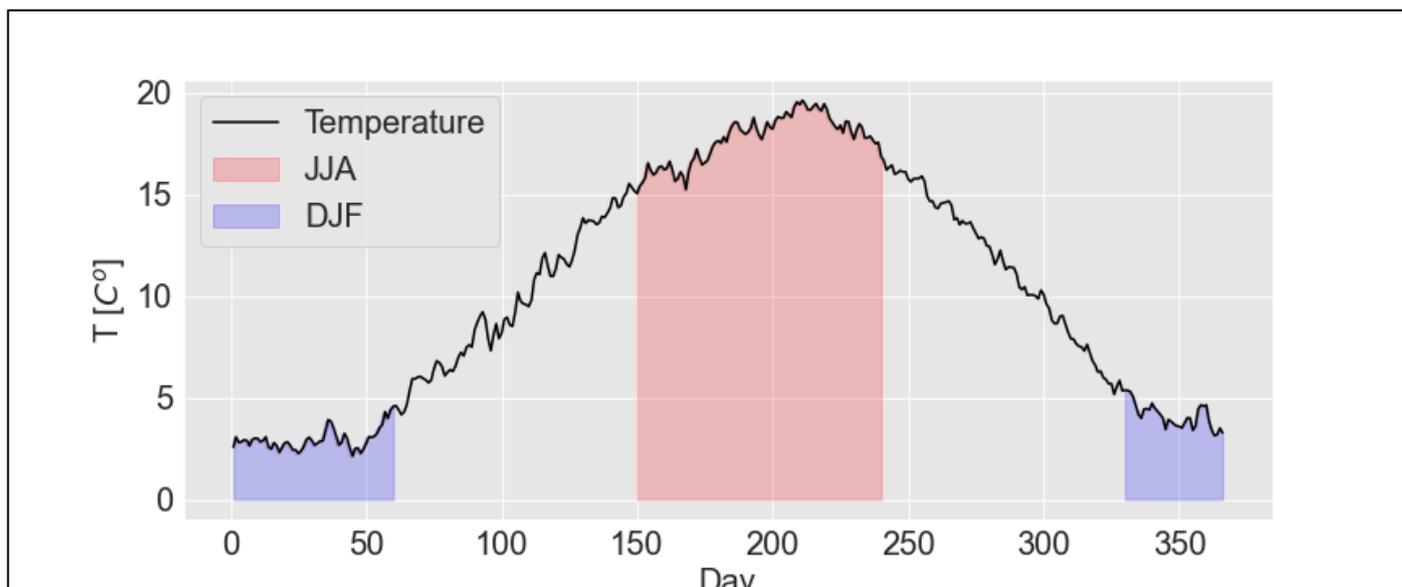


Figure 3-5: DSST for temperature, the red envelope is JJA and the light blue is the DJF period.

For temperature, the wet period at Lobith is the winter period (DJF), while the dry period is the summer period (JJA). QM and SDM are used to correct wet periods (DJF) in the future based on dry periods (JJA) in the reference run, and vice versa. The assessment is based on the computed mean and standard deviation temperature values of the corrected periods relative to the actual observations during the same projected.

For precipitation, it is hard to distinguish between the wet period and the dry period at Lobith. This is because the mean monthly precipitation at Lobith during the summer and winter periods from 1979 – 1998 is 2.1 mm/d (see figure 3.6). Therefore, the period in which the mean monthly precipitation is higher during the validation period (1999 – 2018) will be referred to as the wet period, while the other will be referred to as the dry period in the results section.

The assessment is based on the computed mean and standard deviation precipitation values of the corrected periods relative to the actual observations during the same projected.

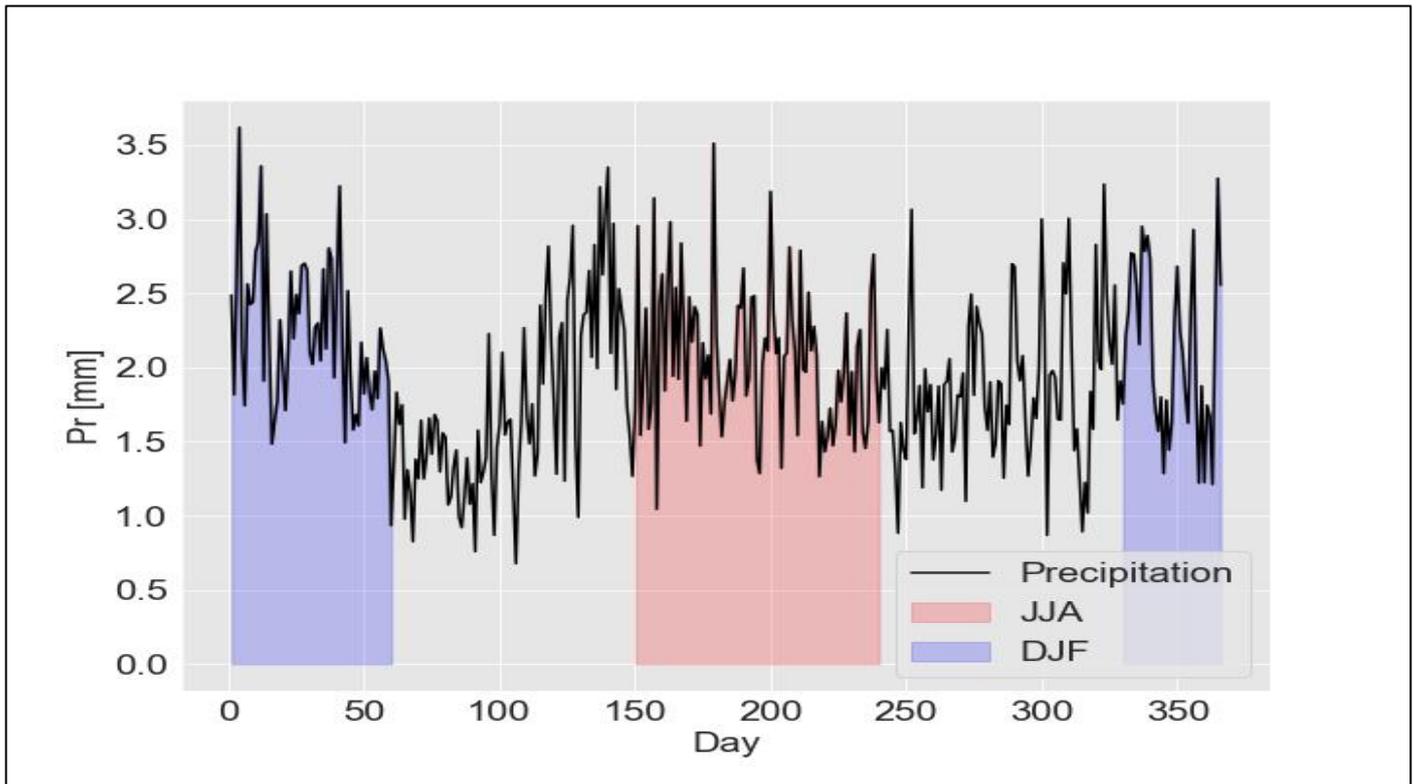


Figure 3-6: DSST for precipitation, the red envelope is JJA and the light blue is the DJF period.

3.4 Performance-based weighting

In the following subsections, two performance-based weightings are discussed: Reliability Ensemble Averaging (REA) and Climate Model Weighting by Independence and Performance (ClimWIP). One that evaluates the raw climate simulations based on the modelled discharge statistics (REA) and another that weight GCMs based on the meteorological variables.

3.4.1 Climate Model Weighting by Independence and Performance

The Climate Model Weighting by Independence and Performance (ClimWIP) is a recent method that has been proposed by (Knutti et al., 2017). ClimWIP assess general circulation models (GCMs) based on independence and performance.

The performance assumption originates from the fact that there are many uncertainties in climate projections, for that it is not valid to assume equal performance weight for each GCM (Knutti, 2010). For this reason, ClimWIP assigns a different weight to each GCM based on their historical performance in projecting the average change in a certain meteorological variable.

The Independence assumption is because within the climate community there are many climate models share a similar building code (Annan and Hargreaves, 2017; Abramowitz et al., 2019). Thus, it deemed not logical to consider models that only perform better without accounting to the fact how they have been built.

The merits of climate models is that they help us in perceiving different possible realities, this assumption is undermined by considering climate models that show high performance and have the same code in any impact assessment studies, while discarding models which have been built different. ClimWIP has been used in assessing many climate impacts assessment studies across the globe (Sperna Weiland et al., 2021; Brunner et al., 2019; Brunner et al., 2020).

The ClimWIP weights the GCMs by evaluating the accuracy of their simulated meteorological variables at a certain region to the actual observations at that region, while considering their independence in simulating these variables.

The general weighting formula for ClimWIP method can be seen in equation 3.16.

$$w_i = \frac{e^{-\frac{D_i^2}{\sigma_D^2}}}{1 + \sum_{j \neq i}^M e^{-\frac{S_{ij}^2}{\sigma_S^2}}} \quad (3.16)$$

Where w_i is the weight for GCM i , D_i is the distance between the specific observational dataset and the climate simulations of GCM i , S_{ij} is the distance between GCM i and GCM j , σ_S is the shape parameter which stands for the performance criteria and σ_D is the shape parameter which stands for the performance criteria.

What is meant by distance in (D_i, S_{ij}) is the error function. In the distance in the ClimWIP method is the mean square error (Brunner et al., 2019). The term D is the error function between observational datasets and each GCM i , while S is the error function between GCM i and GCM j .

The values of σ_S and σ_D are very sensitive and they require special attention. Large σ_D (performance factor) values lead to equal GCM weights and small σ_D values lead to overconfident results in the performance of the GCMs. Small σ_S values indicate that all models are considered as independent and large values consider that the GCM model are dependent. The term dependent refers to GCMs having the same code structure

Precipitation and temperature have been considered as the only meteorological variables to be used in deriving the weights of the GCMs. Evapotranspiration is discarded due to the mismatch in calculating of the Makkink evaporation in the observational datasets and the observations (Makkink, 1959). Additionally, four weights are considered for the RCMs (subset with IPSL ID is excluded from this step due to data issues).

For the case of the Rhine, the value for σ_S and σ_D parameters are chosen based on Brunner et al., (2019) that applied the method over Europe. They have documented the most stable σ_S and σ_D weights that provide reasonable GCM weights per region in Europe. One value for parameter σ_S is used (0.706), while for the performance parameter σ_D a value of 0.607 for temperature, and a value of 0.831 for precipitation in the entire Rhine basin (central Europe) (Brunner et al., 2019).

Based on these parameter values the weighting is performed using the ESMValTool (v2.0) in Linux (Righi et al., 2020). ESMValTool stands for Earth System Model Evaluation Tool. It is an open source tool that provide the possibility of using different recipe, a recipe is a special term used in this tool to refer to the methods which are developed by climate scientists. The applied recipe is the (recipe_climwip_brunner_2019_med.yml) recipe, which is the same method proposed by (Brunner et al., 2019).

The calculated GCM weight is then used to obtain the RCMs' weights based on that parent GCM (the GCM on which the dynamical downscaling of the RCM is applied). In this project there are two parent GCMs: the ICHEC-EC-EARTH model and the MOHC- HadGEM2-ES model. The weighting procedure is described in equation 3.17.

$$R_{g,i} = \frac{w_i}{n_i} \quad (3.17)$$

Where $R_{g,i}$ is the weight of RCM $_g$ within GCM i and n_i is the total number of RCMs within the GCM i .

After obtaining all the RCMs weights, these values are multiplied to their corresponding modelled discharges. Then, they are summed together to come with a one weighted discharge that represents the four RCMS. Following equation 3.18.

$$Q_{TW} = \sum_i^n Q_i * M_{w,i} \quad (3.18)$$

Where Q_{TW} refers to the total weighted discharge, Q_i is the modelled discharges for RCM i , $M_{w,i}$ is the calculated weight for RCM i , and n is the total number of RCMs which are considered.

The general procedure on how ClimWIP is applied to obtain the weighted raw discharges can be found in appendix I.

3.4.2 Reliability Ensemble Averaging

Reliability Ensemble Averaging (REA) has been proposed by (Giorgi and Mearns, 2002), and ever since it has been modified to suit several applications (Giorgi and Mearns, 2003; Tegegne et al., 2019; Sperna Weiland et al., 2021). The reason for REA being popular within the climate community is because it is a multivariate method, meaning it is suitable for weighting RCMs and using them in various applications.

This method has been considered as a weighting method for the chosen RCMs, based on their historical performance in producing accurate modelled discharges, by using the simulations after running `wflow_sbm` (section 3.5). REA is based on achieving two criteria's; divergence, and performance, which are similar to the two criteria listed in ClimWIP.

The REA provides a weight (reliability factor) for each RCM by considering two things: the divergence of the future discharges of the specific RCM to the ensemble mean future discharges (5 RCMs in this case) and the bias in the RCM's raw discharge relative to the current climate (actual discharge).

In simple terms, it assigns a weight to each RCM based on the mean difference between the future modelled discharges of the RCM to the mean of the ensemble (the mean value of all RCMs in the future) and on the mean difference between the modelled discharges in the past to the actual discharges.

The general formula for REA method is described in equation 3.19. The convergence in the future discharges is calculated based on the raw (no bias correction) modelled discharges at Lobith for the period (2020 – 2060). The bias of each RCM is based on the mean difference between the actual discharges for the period (1979 – 2018) to the raw modelled discharges for the same period.

$$w_{\bar{Q},i} = [w_{B,i}^m \cdot w_{D,i}^n]^{1/(m,n)} \\ = \left\{ \left[\frac{\epsilon_Q}{abs(B_{Q,i})} \right]^m \left[\frac{\epsilon_Q}{abs(D_{Q,i})} \right]^n \right\}^{1/(m,n)} \quad (3.19)$$

Where $w_{\bar{Q},i}$ is the weight of RCM i based on discharge statistics and $B_{Q,i}$ is the bias (error) between the projected discharges at a certain location and the observed discharges at that location. This bias is the temporal mean value of the differences between the RCM i simulation and the observed discharges during the reference period (1979 – 2018).

D_{Q_i} is the divergence between RCM i discharges to the average weighted ensemble (all RCMs) discharges, which is the temporal mean of the difference in the RCM i simulation to the average weighted signal of the entire ensemble during the future run (2020 – 2060).

ε_Q is a measure of natural variability, it is calculated by subtracting the maximum and minimum observed values at Lobith based of a moving average window of 20 years. The reason of choosing 20 years window is to avoid the influence of trends (Giorgi and Mearns, 2003; Sperna Weiland at al., 2021).

The parameters m and n are used to obtain the weights based on different criterions. How these parameters are assigned is subjective and based on the user. However, in this research both were set to 1 based on earlier work by (Tegegne et al., 2019; Sperna Weiland at al., 2021).

The calculation procedure is performed in excel after obtaining all the modelled discharges through the wflow_sbm model. Then, these weights are then multiplied by their corresponding modelled discharges from wflow_sbm and summed together to obtain the weighted average discharge that represents all subsets based to the REA method (as in equation 3.18).

3.5 Hydrological Model

Modelled river flows are calculated using the wflow_sbm (sbm stands for simple bucket model) conceptual hydrological model developed by Deltares. Wflow_sbm is free and open source package that can be downloaded easily by following the documentations described here ([How to install · Wflow.jl \(deltares.github.io\)](https://deltares.github.io/Wflow.jl)). The hydrological processes in the model are a refined version from the CQflow model (Köhler et al., 2006). The soil part is largely based on Topog_SBM (Vertessy and Elsenbeer, 1999).

In figure 3.7 below, the different hydrological fluxes and processes are described. The differences between the wflow_sbm and Topog_SBM is the following:

- Accounting for evapotranspiration and interception losses.
- New root water uptake reduction function has been incorporated.
- Accounting for capillary rise, glaciers, and snow build-up.
- The possibility of using multiple soil layers.
- Water routing over an eight-direction network (D8).

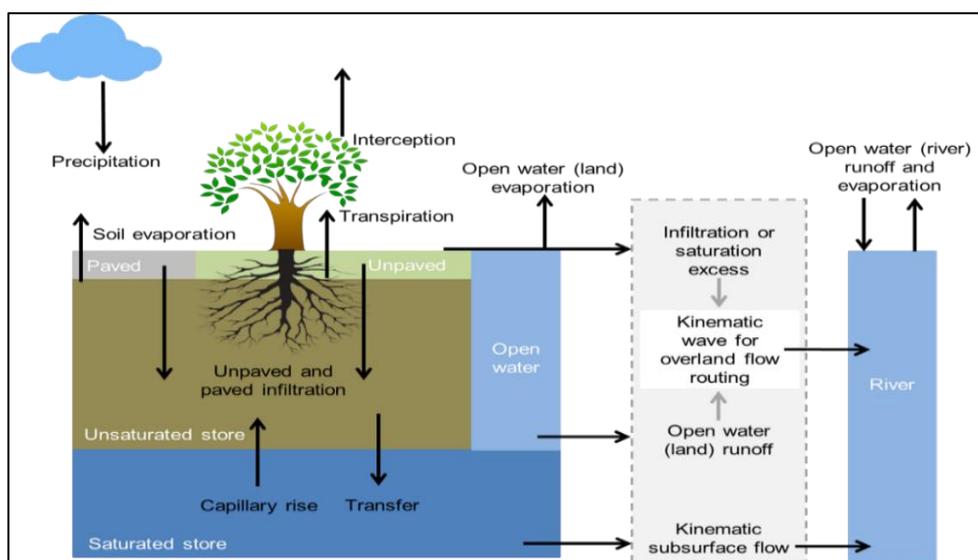


Figure 3-7: Wflow_sbm schematic representation (van Verseveld et al., 2020).

The model which has been used in this study is already calibrated by Deltares (2021). This model is built based on a Julia framework that has proven to be computationally faster than the python framework. It also contains an updated soil parameter map along with other internal modifications. Wflow_sbm is forced by the following forcing's: temperature, precipitation, and evaporation. The outputs of wflow_sbm are daily discharges expressed in [m³/sec]. It's important to mention that, the model is initiated by ERA5 states in the first day of the simulation (January 1st, 1979), this is needed to avoid having the model to start from zero discharges at the first period of the simulation.

The general framework that describe the application of bias correction into the chosen ensembles, along how they were fed into the wflow_sbm can be found in appendix I.

3.5.1 Additional Correction for temperature

Climate models generate their simulations based on a certain orography's that are different from what is observed (Matiu et al., 2019; Van Vooren et al., 2019). Temperature correction is deemed necessary for all of climate simulations (bias-corrected or not), because the target resolution to operate wflow_sbm is [1 km²], while the remapped raw climate simulations share a resolution of about [5 km²]. Moreover, due to having five different orography's for the Rhine basin in these EURO-CORDEX simulations, which are different from the digital elevation model used in the wflow_sbm model. The goal of this step is to ensure that the temperature driven hydrological processes (i.e., unsaturated zone, snow melting) are related to the real elevation, especially in the Alpine part of the Rhine basin.

The applied temperature correction method was the one proposed by (Van Osnabrugge et al., 2019), described through the following equation.

$$T_{x, \text{corrected}} = \gamma^*(H_{\text{wflow_sbm},x} - H_{i,x}) + T_x \quad (3.20)$$

Where $T_{x, \text{corrected}}$ [°C] is the final temperature product at grid cell x , that was corrected based on a constant lapse rate γ of 0.0066 [°C/m]. $H_{\text{wflow_sbm}, x}$ is the orography elevation at grid cell x recorded in the digital elevation file within wflow_sbm, while $H_{i,x}$ is the orography of climate model i at grid cell x . T_x is the raw temperature or bias corrected input at grid cell x .

3.6 Performance Criteria

To assess the robustness of the applied correction methods, many assessment metrics have been used. The reason for choosing these metrics, is to quantify the implications of applying bias-correction methods and how its translated in the low flows' context, in addition to assessing the possible potential impacts by 2060. These metrics are described in detail in appendix I and include KGE, NSE, NSE_{log}, etc.

4| Results

In this chapter the results of this research work are outlined. The performance of the hydrological model `wflow_sbm` at Lobith is first assessed in section 4.1. Section 4.20 describes the implemented historical experiment to compare applying bias correction compared to using raw simulations or weighting them based on two performance-based weighting methods. Subsequently, section 4.30 is about testing these methods in changing climate conditions using DSST. Finally, section 4.4 demonstrates the potential climate change impacts based on the bias-corrected simulations at Lobith.

4.1 Hydrological model

The calibrated hydrological model `wflow_sbm` (van Verseveld et al., 2020) is used to model the hydrology of the Rhine basin (section 3.5). The results obtained from `wflow_sbm` is used for answering the main research question, therefore it is important to evaluate the model performance based on the available observations at Lobith. The available discharge measurements at Lobith for the period (1979-2018) are used to evaluate the model ability in providing accurate discharges at Lobith.

The results of using raw simulations (without correction) as inputs in the `wflow_sbm` for the same period are also presented. This is to evaluate the performance of using the raw subsets in hydrological modelling. As a final step the added value of applying the two performance-based weighting methods (section 3.4) in raw simulations is evaluated based on the discharge measurements at Lobith.

4.1.1 Observed discharges

The calibrated hydrological model (`wflow_sbm`) is forced using the observational dataset between [1979 – 2018] (section 2.2). The model outputs are compared to the observed discharge measurements at Lobith within the same period, the reason for not assessing the model until 2019 is because the observed discharge dataset only spans until mid-March 2019.

As can be seen in figure 4.1, `wflow_sbm` is able to consistently capture the timing of low flows and high flows throughout the 40 years run. The discharge metrics of the modelled flows are reasonable as well, `wflow_sbm` achieved NSE score of 0.78, NSE_{\log} of 0.81 and KGE of 0.82. However, it appears that the model overestimates high flows at Lobith in most of the times and to a lesser extent it underestimates low flows. This can be clearly seen in the scatter plot shown in figure 4.2, as `wflow_sbm` appears to overestimate high flows in average by 20 %, but in lower flows it has lower errors compared to the actual measurements.

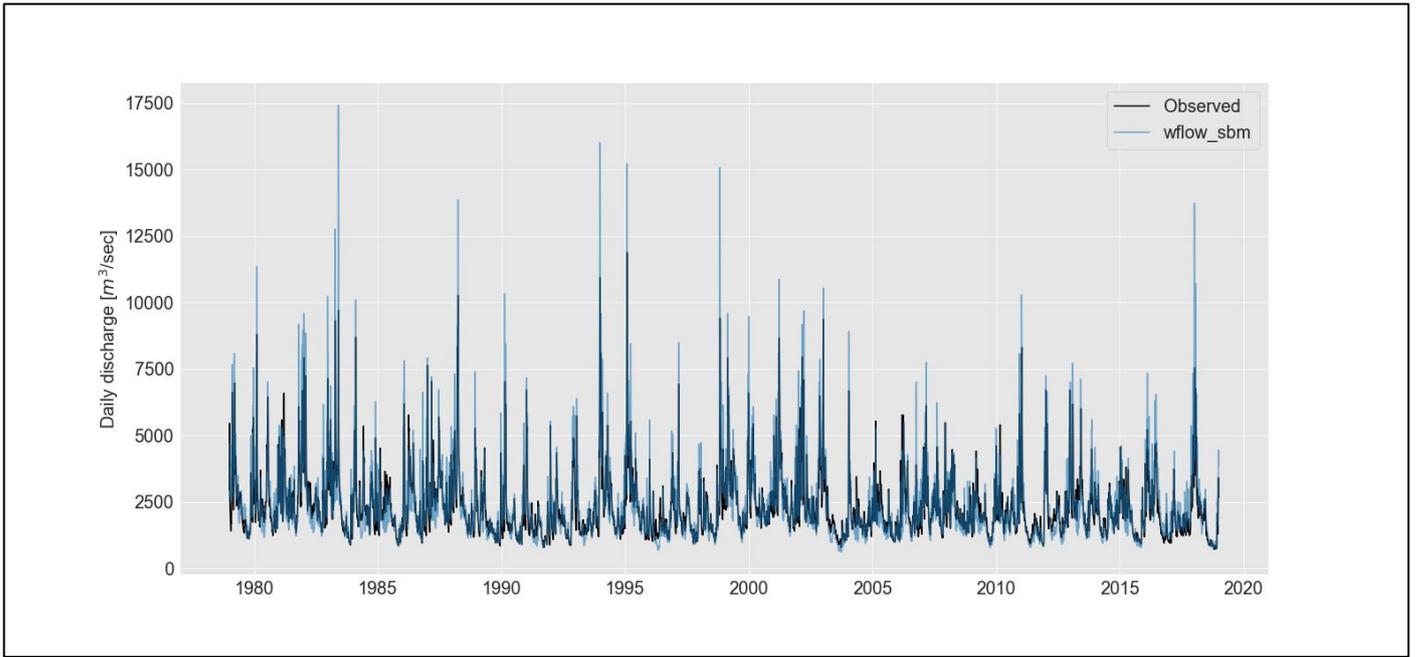


Figure 4-1: wflow_sbm vs observed during the period [1979 - 2018] at Lobith.

In figure 4.2, it appears there is a reasonable correlation between the results of wflow_sbm and actual measurements during the period (1979-2019). The scatter plot indicates the simulated and actual flows correlate well in the range of (900 – 6000 m³/s), however after that there is a disagreement between the two

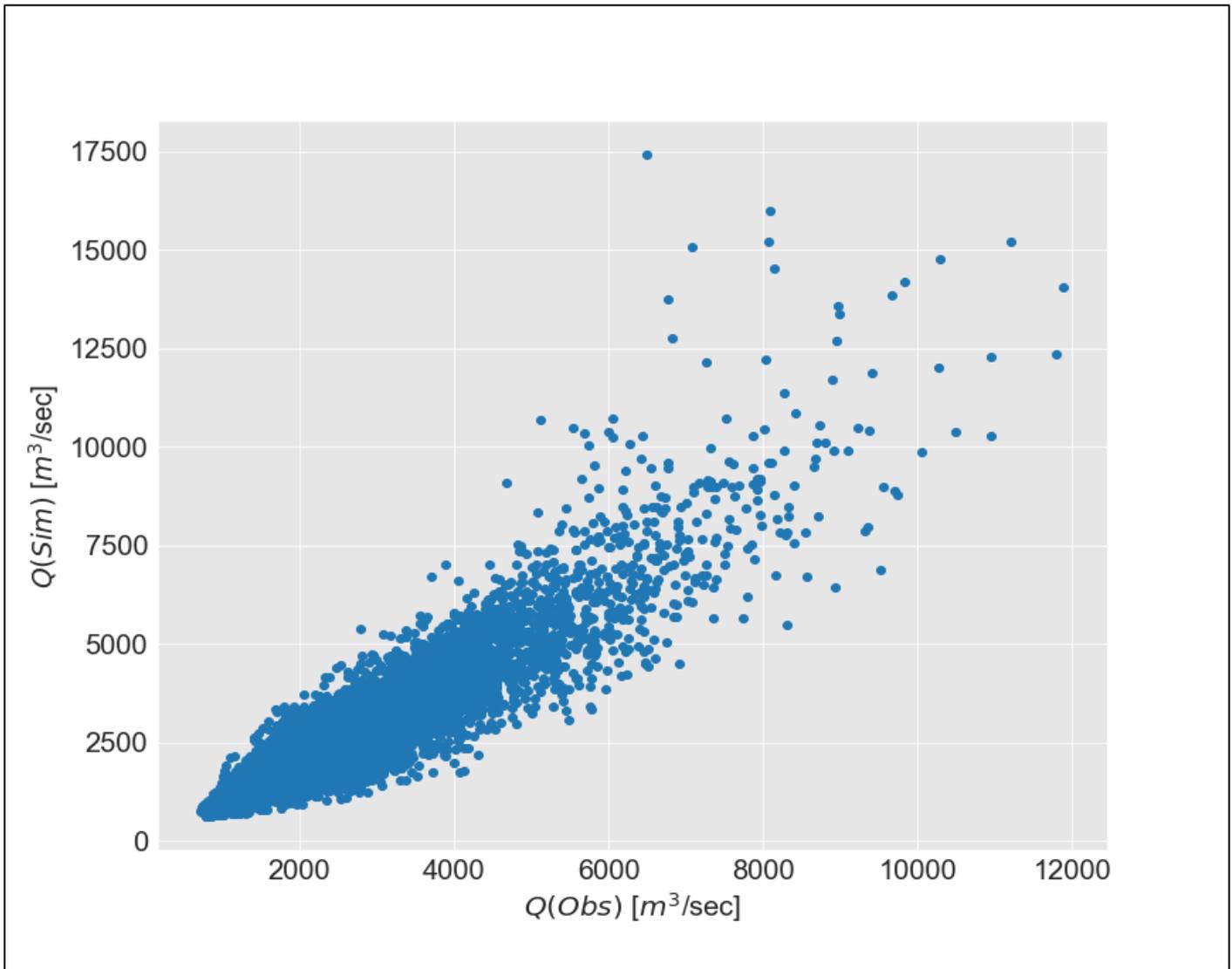


Figure 4-2: Scatter plot of wflow_sbm results vs observed measurements at Lobith for the period (1979-2018).

The flow duration curves, and logarithmic flow duration curves of the observed flows and the modelled flows at Lobith are shown in figure 4.3. The results agree with what is apparent in figure 4.2, that the model overestimates high peaks and underestimates low flows. In the derived flow duration curves of the two timeseries, it appears that the model agrees well with the observation for flows lower than $3500 \text{ m}^3\text{s}^{-1}$.

The logarithmic flow duration curves shown in figure 4.3 describe the ability of wflow_sbm in providing accurate predictions. It can be said that the model slightly underestimates low flows, for flows lesser than $2500 \text{ m}^3\text{s}^{-1}$ and vice versa.

From the flow duration curves in figure 4.3, the 95th percentile flow (Q95) (appendix I) of the measurements is $1100 \text{ m}^3\text{s}^{-1}$ and the one from the model is $1009 \text{ m}^3\text{s}^{-1}$. The difference between the two values is minimal ($90 \text{ m}^3\text{s}^{-1}$). The Q95 of the observations is used as the limiting threshold in low flows analysis, the values below this threshold are described as severe low flows, while values above the threshold are considered within the range of less severe flows to high flows (normal range).

Additionally, the 5th percentile value of the flow (Q5) from the model results is $4858 \text{ m}^3\text{s}^{-1}$ and from the observations is $4552 \text{ m}^3\text{s}^{-1}$. The error in estimating Q5 is three times more than the recorded error in estimating Q95.

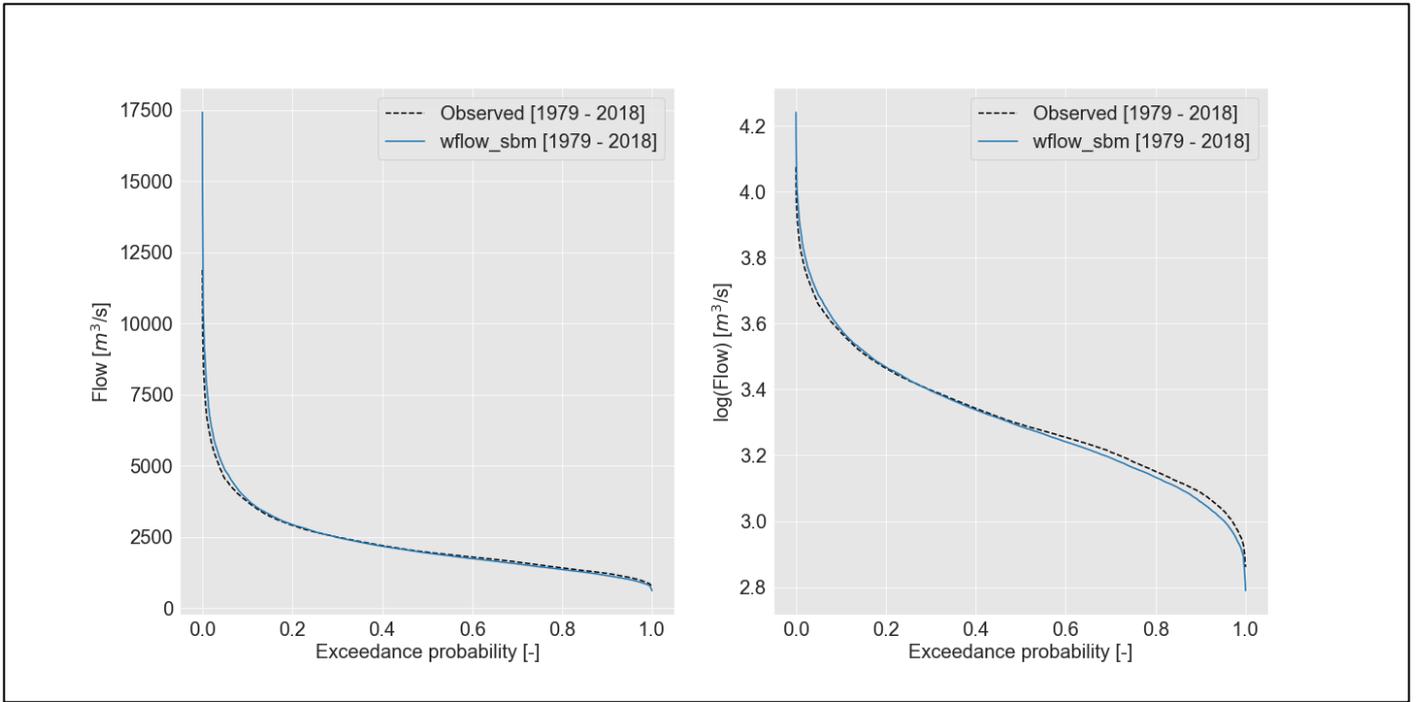


Figure 4-3: flow duration curve and logarithmic flow duration curve for the observed flows and modelled flows for the period (1979-2018).

For applications where low flows are considered, the focus should not only be limited to the model ability to produce marginal values as Q95 or Q5, but also on how it can predict the number of consecutive low flow days in a year or through the entire run. The threshold of $1100 \text{ m}^3\text{s}^{-1}$ is used to count the maximum consecutive low flow events in figure 4.4. The results depicted in figure 4.4 are the maximum number of consecutive low flow events in a year for the period (1979-2018). The results obtained from wflow_sbm and the observed discharge measurements are consistent in most of the run period. There are few times in which the maximum observed consecutive low flow events at Lobith are lower than the modelled, but wflow_sbm has resulted in values that are 20 % higher than the observations in 1991, 1996, 2003, 2009, 2012 and 2014. In general, the consecutive number of low flow events from wflow_sbm are higher than the observations.

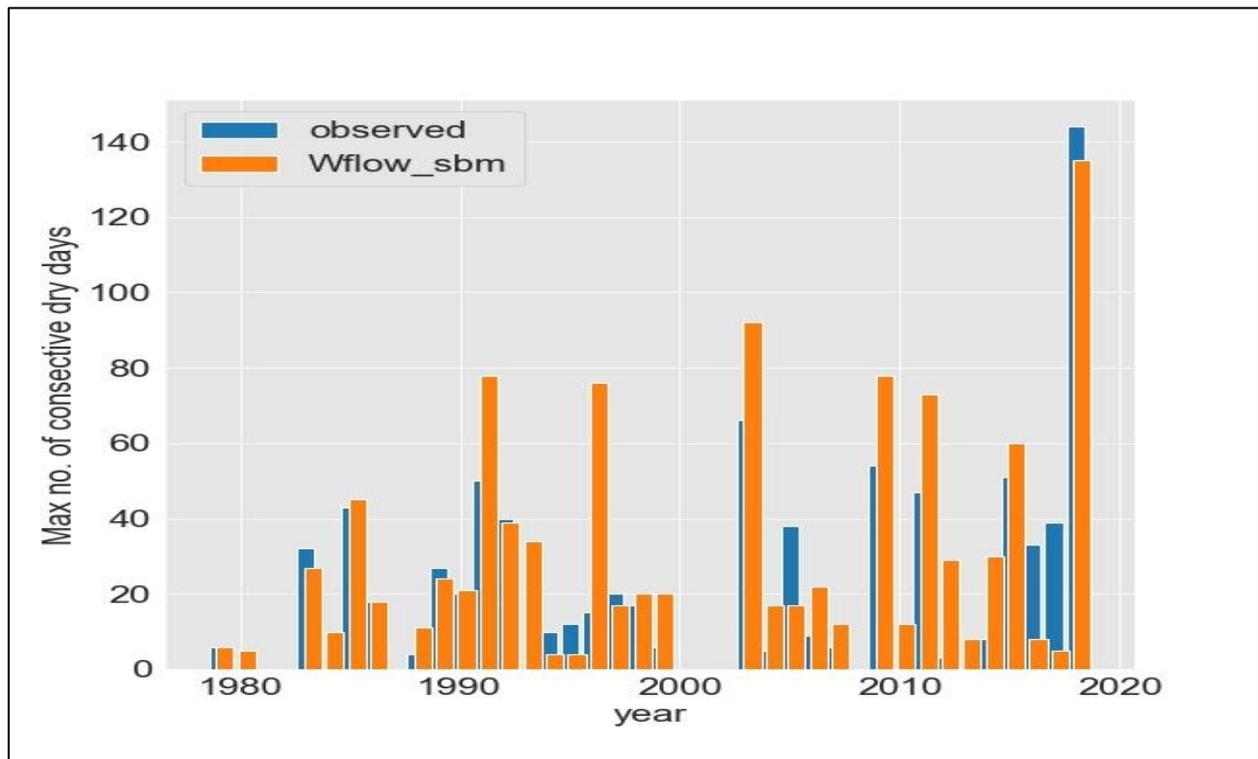


Figure 4-4: Maximum number of consecutive low flow days in a year.

However, the model has provided reasonable results through the entire run. This makes wflow_sbm useful in assessing the effects of low flows at Lobith in the future period. The entire time series shown in figure 4.1 is averaged daily to construct figure 4.5, all the data that falls within a certain day are averaged. The resulted dataset has a length of 366 (leap year is included).

The average daily discharge values will be used in assessing the impacts of climate change on low flows at Lobith instead of a long hydrological timeseries as the one shown in figure 4.1. This is because the level of uncertainty when dealing with the results of climate simulations is high and it cannot be treated like normal daily discharge timeseries, due to the randomness and loose boundary conditions. Therefore, the change in climate is expressed based on an average scale and not daily.

The average daily discharge values of the modelled and observed timeseries for the period (1979-2018) are shown in figure 4.5. In the first two months wflow_sbm provide higher average discharge values than the measured (day 1 to 60). This is inverted in the following four months (day 61 to 180). The two timeseries seems to agree well during the summer period (day 181 to 280). This is due to the hydrological regime being linear during this period, as the snow melt from the Alps is the major source to sustain this period. For the rest of the period (day 281 to 366), the model provided flows that are higher from the average measured flows.

There is a sudden decline in the daily average flow at Lobith at day 60, the model has captured this change as well. The cause of this change is because in average the discharge values between the end of February and start of May are lower compared to the durations before and after that period. This period is found to have lower precipitation values (see the blue line in figure 4.10).

Overall, it appears that the modelled flows during the autumn and winter periods are higher than the observed because of the witness conditions. The measured values are higher than the modelled during the spring period. The model agrees well with observations through the summer period, because of linear response of snow melting.

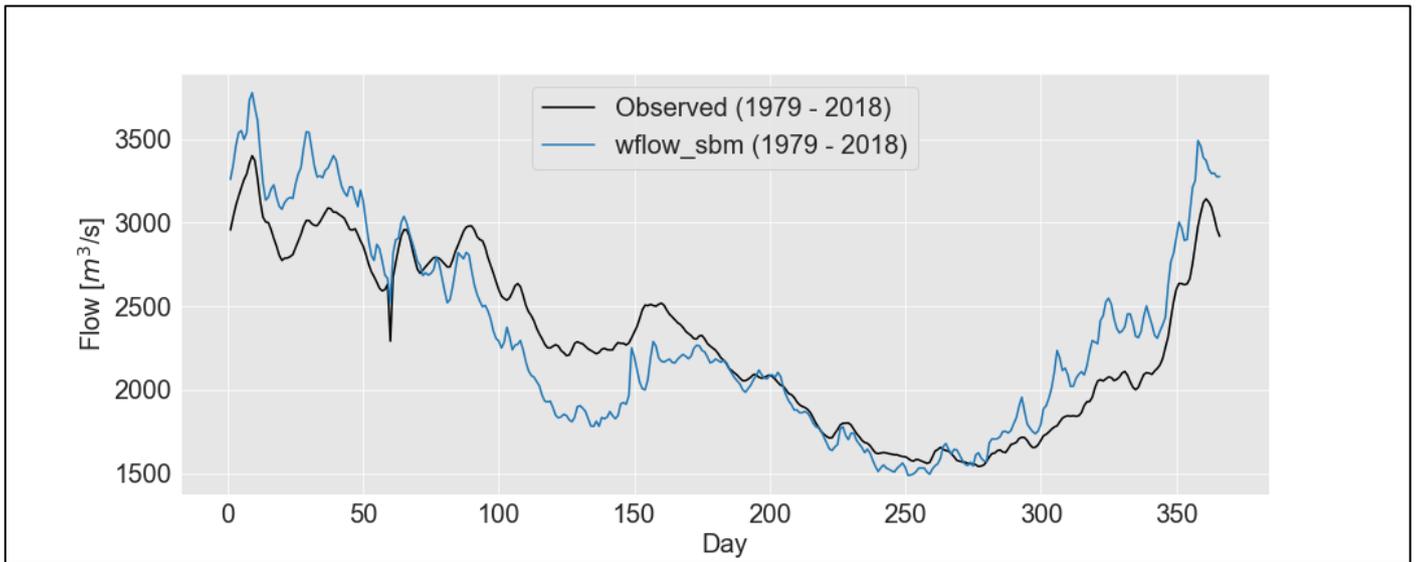


Figure 4-5: Average daily discharges at Lobith for the period 1999 – 2018 for the modelled and observed discharges.

4.1.2 Climate simulations

Raw simulations

The four EURO-CORDEX subsets outlined in section 2.2 (excluding IPSL) are used to obtain the modelled discharges at Lobith for the same period as in section 4.1.1. The results of using these simulations in running wflow_sbm to obtain the modelled discharges at Lobith are shown in figure 4.6. The daily discharges from wflow_sbm are averaged daily to obtain the result shown in figure 4.6.

The blue envelope represents the range where the maximum and minimum average daily discharges of all these simulations lie. The upper band represents the maximum averaged daily discharge of the four simulations for each day. The lower band represents the minimum averaged daily discharge of the four simulations.

The discharges obtained from these raw simulations are higher the observed average daily discharge at Lobith for the same period. Some of the actual average daily discharge lies within the lower boundary of the ensemble's envelope. The period between the day 180 to day 310 could not be captured by the envelope, in addition to few days at the end of January, February and November.

The detailed average daily discharge for each EURO-CORDEX subset is shown in figure 8.1 (appendix II). In figure 8.1, none of the selected simulations succeeded in providing reasonable predictions that can be related to the observations. The average daily discharges from the DMI and CNRM subsets are at least $250 \text{ m}^3\text{s}^{-1}$ greater than the daily average observed discharge. Results from KNMI and SMHI are more accurate, these two subsets gave their best results in the period between the first of March to the end of May.

The raw EURO-CORDEX are not able to provide predictions that can be related to Lobith. One would expect more pessimistic results when applying the chosen scenario (RCP8.50). Although a slight increase in daily averaged flow during some periods due to excessive snow melting or extreme precipitation can be related, but not an increase in the average daily discharge during most of the year by $400 \text{ m}^3\text{s}^{-1}$. The raw simulations are not suitable for the use in modelling the discharges at Lobith based on the fact that they are not able to provide reasonable daily discharges or mimic the general flow regime (hydrograph) at Lobith accurately, in terms of the timing of the flows and their magnitude.

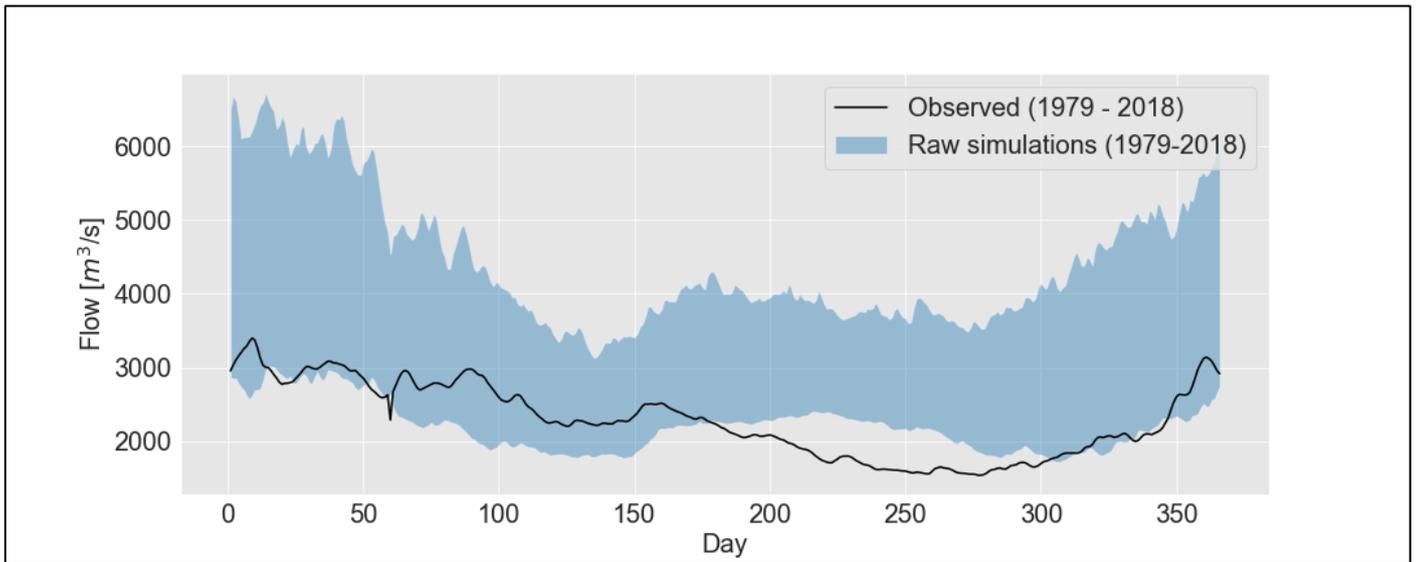


Figure 4-6: Average daily flows at Lobith for the period 1979 – 2018 using an ensemble of four EURO-CORDEX simulations.

4.1.3 Performance-based weighting

The two methods described in section 3.4 are implemented to the EURO-CORDEX simulations. The ClimWIP method focuses on weighting the GCMs based on meteorological observations, while the REA method weights the climate simulations based on the observed discharges, after running `wflow_sbm` with these simulations. The result of applying each method is discussed on the following subsections. The general framework on which the weighted discharge of the two performance-based weighting methods is obtained can be found in appendix I (section 7.4.2 and section 7.4.3).

The weighting of EURO-CORDEX simulations is based on the period of (1979-2018) as in section 4.1.1. For the case of REA, the future period of (2019-2060) is considered to derive the convergence factor for each climate simulation (section 3.4.2).

Climate model weighting by independence and performance

The computed GCM weights based on the accuracy and independence of the simulated meteorological variables are obtained using the `ESMValTool` (v2.0). The two GCMs' weights are listed in table 4.1 and the corresponding RCMs' weights are listed in table 4.2. These weights are used to derive the weighted average daily discharge for the period (1979-2018) using equation 3.18.

The weights shown in table 4.2 are equal except for the case of CNRM subset. This is because the five EURO-CORDEX simulations considered in this project are based on two GCMs (ICHEC-EC-EARTH and MOHC-HadGEM2-ES). Three RCMs (DMI, KNMI, SMHI) are based on ICHEC-EC-EARTH and two RCMs (IPSL, CNRM) are derived from MOHC-HadGEM2-ES. IPSL subset is discarded from this test because of data issues, leaving the entire weight of MOHC-HadGEM2-ES to CNRM. For this reason, the weight of CNRM is 0.28 and for the rest of subsets is 0.24.

The weighted average daily discharge from ClimWIP weights is shown in figure 4.7. The weighted discharge using ClimWIP has indeed strengthened the change signal of the four climate simulations in a single signal, reducing the uncertainty. The usage of ClimWIP has not helped a lot in overcoming the main issue in section 4.1.2. As the weighted daily discharge of the four climate simulations is close to the observed discharges during the period between the first of March to the end of May. For the rest of the year, the weighted daily average discharge is higher than the daily average observed discharge at Lobith.

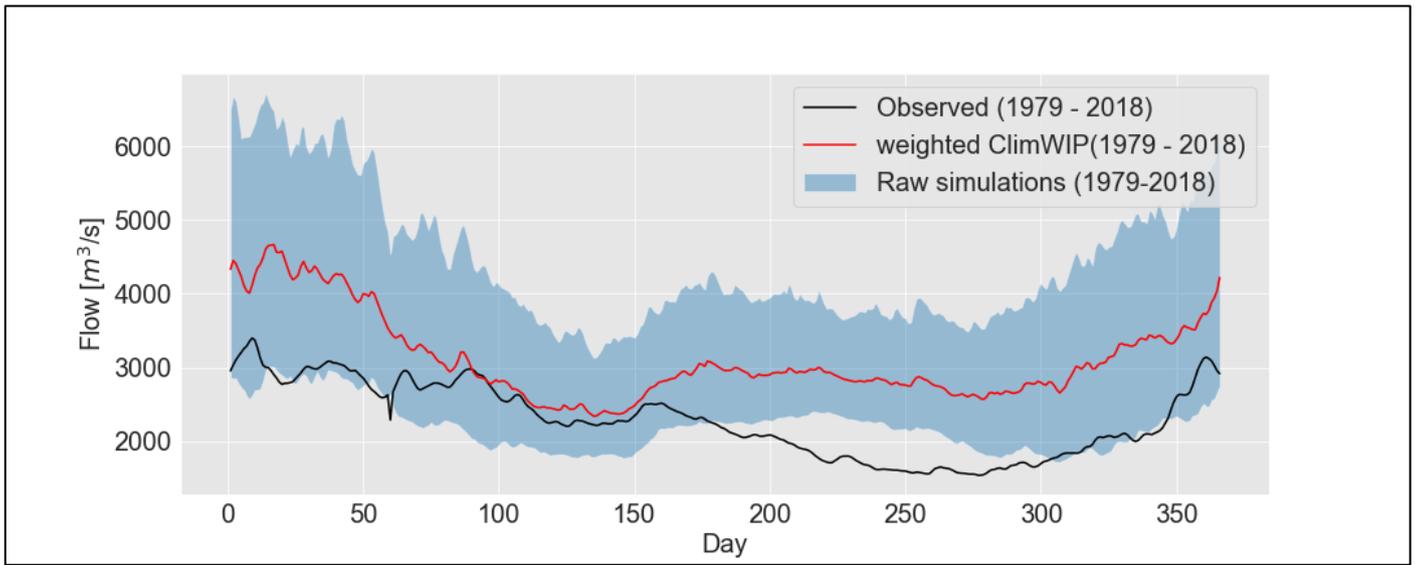


Figure 4-7: ClimWIP weighted daily discharges using the raw simulations for (1999 – 2018) relative to the raw simulations and the actual discharges at Lobith.

Table 4-1: ClimWIP GCMs' weights.

GCM	Weight
ICHEC-EC-EARTH	0.72
MOHC- HadGEM2-ES	0.28

Table 4-2: Corresponding RCMs' weights

RCM subset	Weight
DMI	0.24
KNMI	0.24
SMHI	0.24
CNRM	0.28

Reliability ensemble averaging

The weights derived from the reliability ensemble averaging (REA) is different from that of ClimWIP in two aspects. Firstly, the weighting here is based on the discharges that are obtained after forcing wflow_sbm with the raw simulations (section 4.1.2). The second aspect is in this test a weight is assigned directly to the RCM instead of the parent GCM. The weighting is based on equation 3.19 and the final weighted discharge is obtained through equation 3.18. The weights of the REA method are listed in table 4.3.

The weights of the RCMs using REA are different from the weighting based on GCMs in ClimWIP (table 2.2). The CNRM subsets has the largest value of the RCMs using REA, this is due to its smallest convergence value to the mean of the ensemble in the future period (2019-2060). The RCMs that originate from the ICHEC-EC-EARTH GCM have been assigned different weights. The DMI model has the smallest weight followed by the KNMI. The SMHI model has the highest weight in the selected ICHEC-EC-EARTH subsets.

The weighted average daily discharge using REA is visualized in figure 4.8, the result is like that of ClimWIP. This indicates that these weighting methods cannot overcome the systematic biases in the RCMs.

Table 4-3: Derived RCMs' weights using REA for the period (1979-2018).

ID	Weights
DMI	0.107
KNMI	0.207
SMHI	0.296
CNRM	0.39

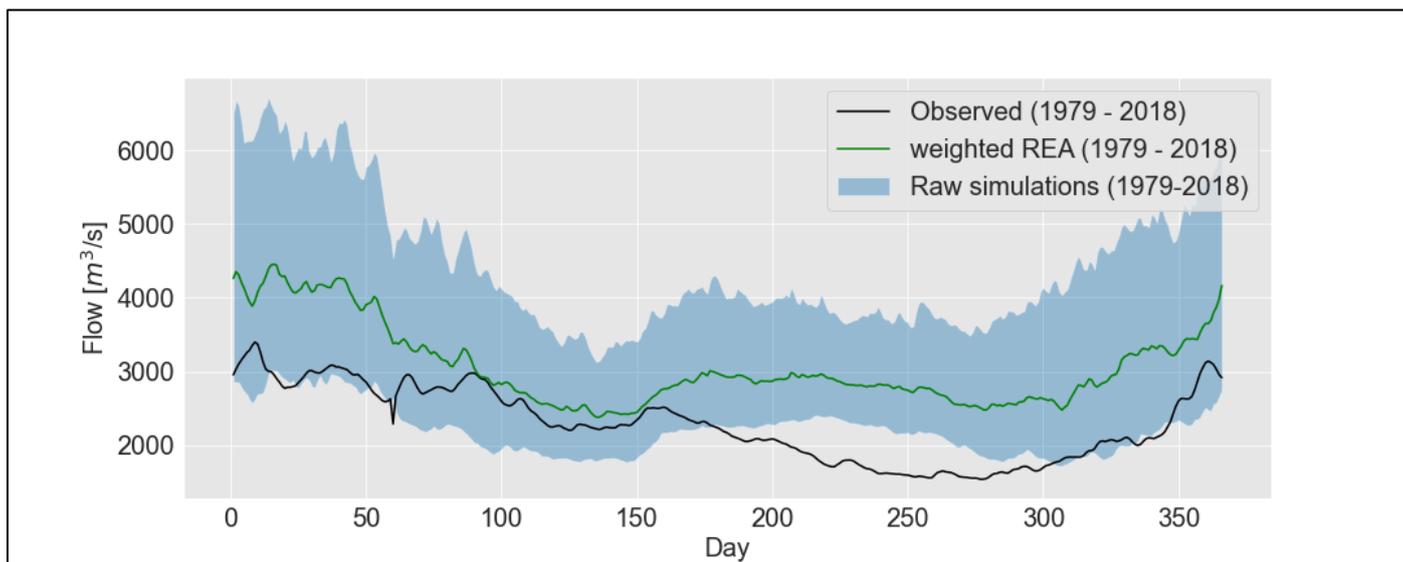


Figure 4-8: REA weighted average daily discharge for the period (1999 – 2018) relative to the raw simulations and the actual discharges at Lobith.

The outputs of the wflow_sbm model when using REA and ClimWIP are very similar because of choosing few RCMs and not having a single RCM that is able to provide average daily discharge estimates lower than the observed. As having lower discharge values and more RCMs could have helped in distributing the weights more, thus altering the weighted result.

4.2 Historical experiment

The historical experiment described in section (3.2) is implemented to evaluate the added value of applying the two bias correction methods compared to using raw climate simulations in hydrological modelling, as inputs. The evaluation is based on the quality of the modelled discharges at Lobith using the bias-corrected simulations. This is coupled by analysing the accuracy of long-term statistics that are relevant in purposes related to hydrology. The results of the empirical quantile mapping method (QM) and the scaled distribution mapping method (SDM) are discussed in the following section 4.2.1. The calculated long-term statistics (appendix I) for all discharges are discussed in section 4.2.2.

4.2.1 Quantile mapping and scaled distribution mapping.

As discussed in section (3.2), the bias correction in the historical experiment is based on the period (1979 - 1998) to derive bias-corrected climate simulations for the period (1999 - 2018). To correct future raw simulations using any bias correction method, there are two datasets that are needed next to raw future simulation. The two are the observational dataset and historical climate simulation of the same period and the same region. The results of this section are corrected based on the reference run (1997 - 2018), see table 3.2.

To validate the bias-corrected simulations the observational datasets of the same period (1999 - 2018) are used.

The results after running the wflow_sbm model with the bias-corrected simulations are shown in figure 4.9. These results are the daily average discharge at Lobith during the period (1999 - 2018) in addition to the actual daily average discharge of the same period, for validation.

The two bias correction methods provided more accurate hydrographs for each RCM compared to using raw simulations (figure 8.1). Most of the models can correctly capture the timing of the actual hydrograph in the second half of the year (day 180 to day 366). Three subsets provided similar performance in the second half

of the year (IPSL, SMHI and KNMI) and two provided lower performance during the same period (DMI and CNRM). There are some differences in the amount of errors in each subset. The three SDM corrected subsets which have the lowest amount of errors are, IPSL, KNMI and SMHI. For the first half of the year, all models failed to provide reasonable flow timing. The QM corrected DMI subset is the closest.

There is a flow trend shared between most of these subsets, which is the sudden jump in flow between day 150 to day 200. To examine the cause of the sudden jump. The precipitation, and temperature datasets at Lobith and in some parts in Germany are analysed. The datasets are observed datasets, raw simulations, and quantile mapping (QM) corrected simulations for the period (1999-2018). The reason for including the raw simulation is to show how the bias corrected simulations in some parts of the Rhine are compared to the observed datasets and to the raw simulations, within this historical experiment.

The result of this analysis is shown in figure 4.10 and figure 4.11.

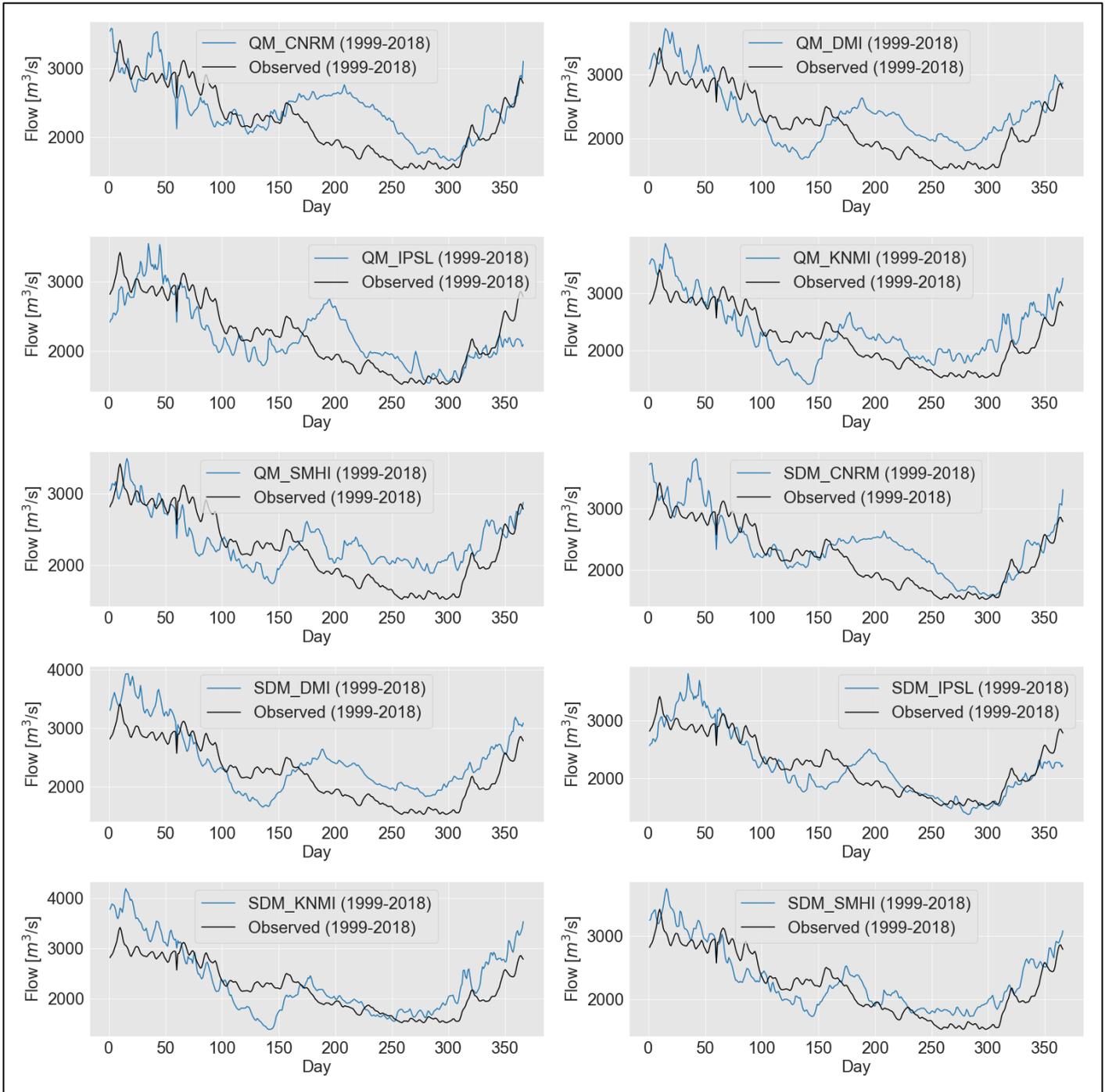


Figure 4-9: wflow_sbm forced with bias-corrected EURO-CORDEX simulations for the period (1999 - 2018) at Lobith compared to the actual flows at Lobith for the same period.

In figure 4.11, It appears that the raw temperature simulations at Lobith are colder than what is observed during the same period, with an average difference of 2 C°. The QM method appears to correct these raw simulations and provide warmer future temperatures. The corrected temperature simulations appear to agree relatively well with the observed temperature for the same period, between day 70 to 300. For the rest of the period, the corrected simulations are warmer than the observed temperature values that are used for validation.

In figure 4.10, the raw precipitation simulation for the future period are higher from the observed precipitation during the same period, with a small exception in August. However, the QM method managed in average to reduce these values and bring them closer to the observation.

The highest average monthly precipitation using bias-corrected simulation within Lobith for the period (1999 – 2018) is 2.52 mm, which is the same average precipitation which is found within the observation. The method also managed to reduce the wet precipitation bias in January and June, but it provided false average values in the rest of the months. It seems that the average error within the entire year has reduced, but the mismatch in the timing of precipitation at Lobith could not be averted. This mismatch is caused by the underlying GCM, which is one of the main source of errors in hydrographs is the GCMs.

The sudden trend depicted in figure 4.9, is due to the fact that within the period of day 150 to 200 (June 1st to July 20th) the highest average precipitation is introduced to the hydrological system and coupled with a constant increase in average temperature (reaching the max at day 200). This translated into higher flows because of higher rainfall and snow melting in the alps. Although the precipitation rates started to fall after day 150, but average temperature continued to increase.

The higher average daily flow of the corrected climate simulations within the period of January to mid-February can be related to having higher average rainfall values than the average observed precipitation at Lobith.

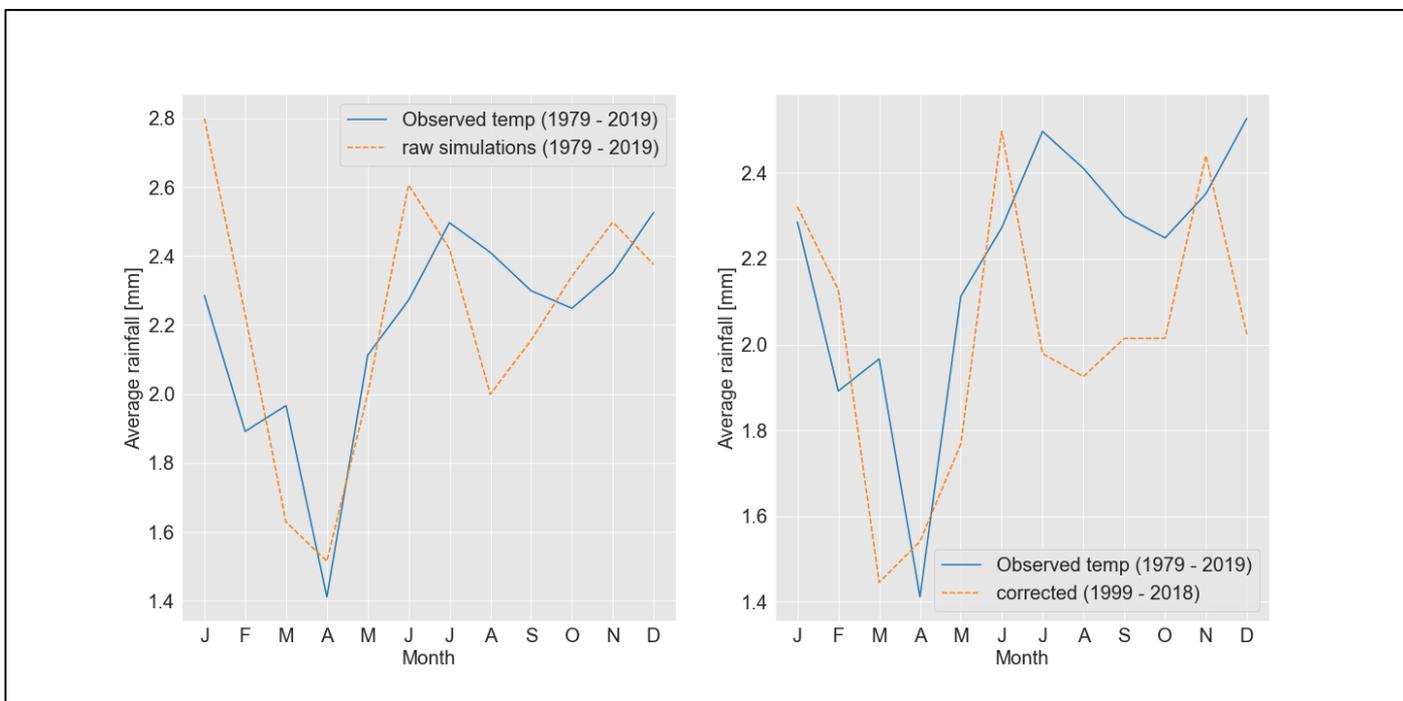


Figure 4-10: Average precipitation comparison between the observed, raw, and corrected datasets at and near Lobith during the historical experiment

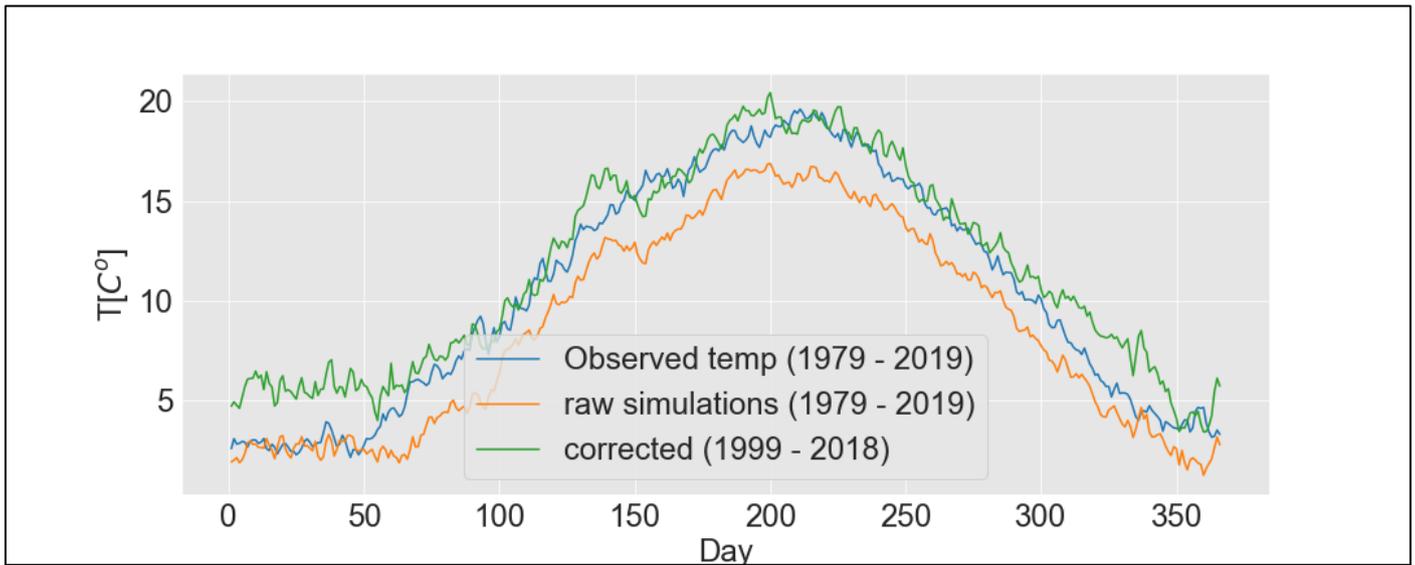


Figure 4-11: Average temperature comparison between the observed, raw, and corrected datasets at and near Lobith during the historical experiment

The ten subsets shown in figure 4.9 are grouped into two ensembles, each ensemble represent the correction method of that ensemble (QM or SDM). The three ensembles (raw, QM corrected, and SDM corrected) are shown in figure 4.12 next to the average daily observed flow for the period (1999 - 2018).

The average daily observed discharge at Lobith is fully contained within the ensemble range in figure 4.12. The two correction methods achieved the same performance in containing the observed discharge and in reducing the higher flows depicted in using raw simulations. QM and SDM have different working principles and yet they managed to produce similar results, indicating the same efficiency in removing the systematic bias.

The bandwidth of the envelope of the two correction methods is the lowest during the summer period (day 130 – day 240), this is due to the fact that during most of the summer and in spring periods the response of the catchment is more linear.

When looking at figure 4.7 and figure 4.8 to figure 4.12, it appears the envelopes of bias-corrected datasets have contained the actual discharges better than the REA and ClimWIP methods. This is again caused by the wet precipitation bias that prevented the raw simulations to be closer to the observed discharge. The weighted signal of the RCMs would have been different if one RCM has provided lower estimates than the observed (i.e., the weighting would be more robust).

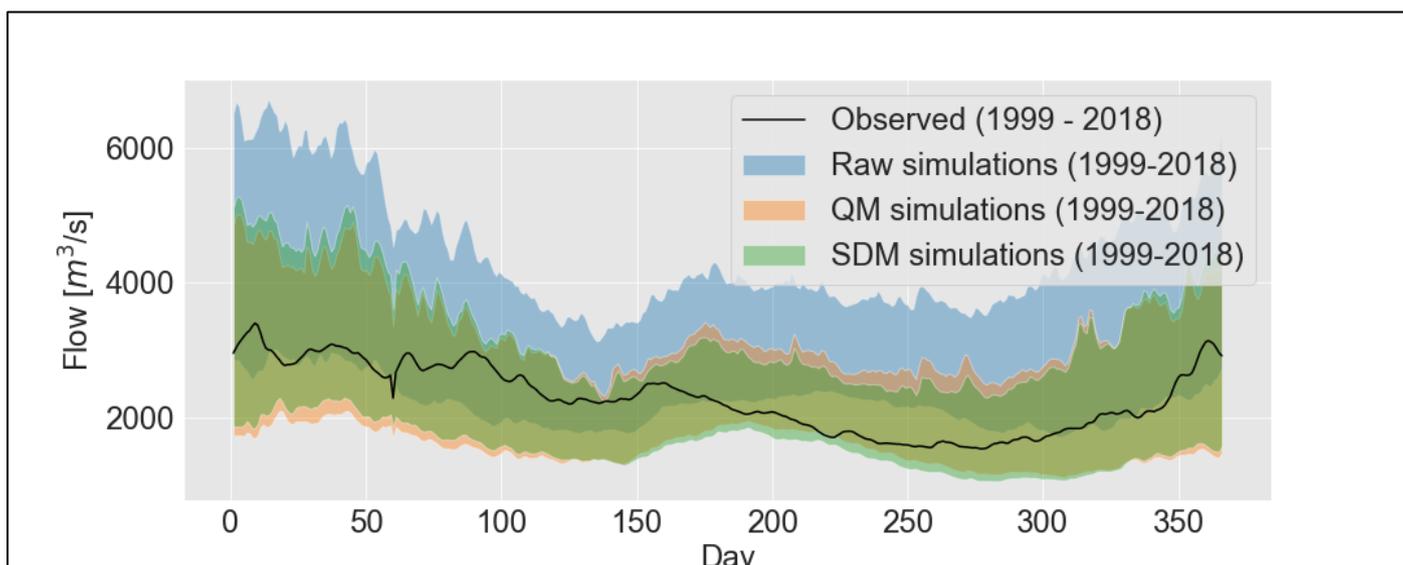


Figure 4-12: daily average discharge envelope of the raw forcing's and corrected forcing's for the period (1999 – 2018) in addition to the daily average discharge of the same period.

4.2.2 Long-term statistics

In this section, the long-term statistics of modelled discharges at Lobith using the bias-corrected simulations for the period (1999 - 2018) are evaluated based on the observed statistics during the same period. The modelled discharges using the raw simulations for the same period are also analysed, to check the robustness of the modelled discharges using bias correction to raw simulations.

Figure 4.13 contains the results of three long-term statistics (minimum 7-day low flow, long-term mean flow during the summer period and long-term annual mean flow). In the figure 4.13(A), the minimum 7-day low flow values of the raw simulations are high compared to the observations. This is because of relatively cold temperature rates and higher precipitation rates (figure 4.10 and figure 4.11). All the bias-corrected simulations (being from QM or SDM) are closer to the observations than the raw simulations. Three SDM subsets provided the most accurate predictions (KNMI, IPSL and SMHI). The bias-corrected simulations using QM exhibited larger errors than SDM subsets.

Figure 4.13(B) includes the long-term mean during the summer period (June to August). The obtained long-term mean during the summer period from the raw simulations is greater than what is obtained from the observed discharges. The bias-corrected simulations provided more reasonable values than the raw ones. The SDM corrected KNMI subset provided the best predictions. SDM corrected subsets resulted in better performance than QM bias-corrected subsets. However, the least performing QM subset performed better than the best performing raw simulations by over 140 m³/s.

The long-term annual mean flow at Lobith is depicted in 4.13(C), where the results of using observed discharge dataset, bias-corrected and raw simulations for the period (1999 - 2018). The results of using bias-corrected simulations are the most consistent here and they are more accurate than using raw simulations. The difference between the average long-term annual mean of the bias-corrected subsets from the actual is the smallest, compared to figure 4.13 (A) and figure 4.13 (B).

The lengths below threshold (1100 m³/s) for the period (1999 – 2018) of the observed discharges, raw and corrected simulations is shown in figure 4.14. Since the raw simulations have produced higher flows in average from what is observed, there are few flow events that are found to occur below this threshold. On the other

hand, the flows from the bias-corrected simulations are more consistent than the raw simulations. Two corrected simulations provided the same range and mean as the observed discharges, the QM-corrected KNMI subset and SDM-corrected CNRM subset.

The results shown in figure 4.13 and figure 4.14 suggest that the two correction methods really depend on the ability of the RCM to provide reasonable results. When looking at the KNMI subset, one can observe that the raw KNMI performed the best compared to the rest of the climate simulations, in predicting these four metrics with the lowest amount of errors. The use of QM and SDM in correcting this subset has made it perform better. Also, it remained to be the best performing subset within the bias-corrected subsets. The same applies for DMI, SMHI and CNRM.

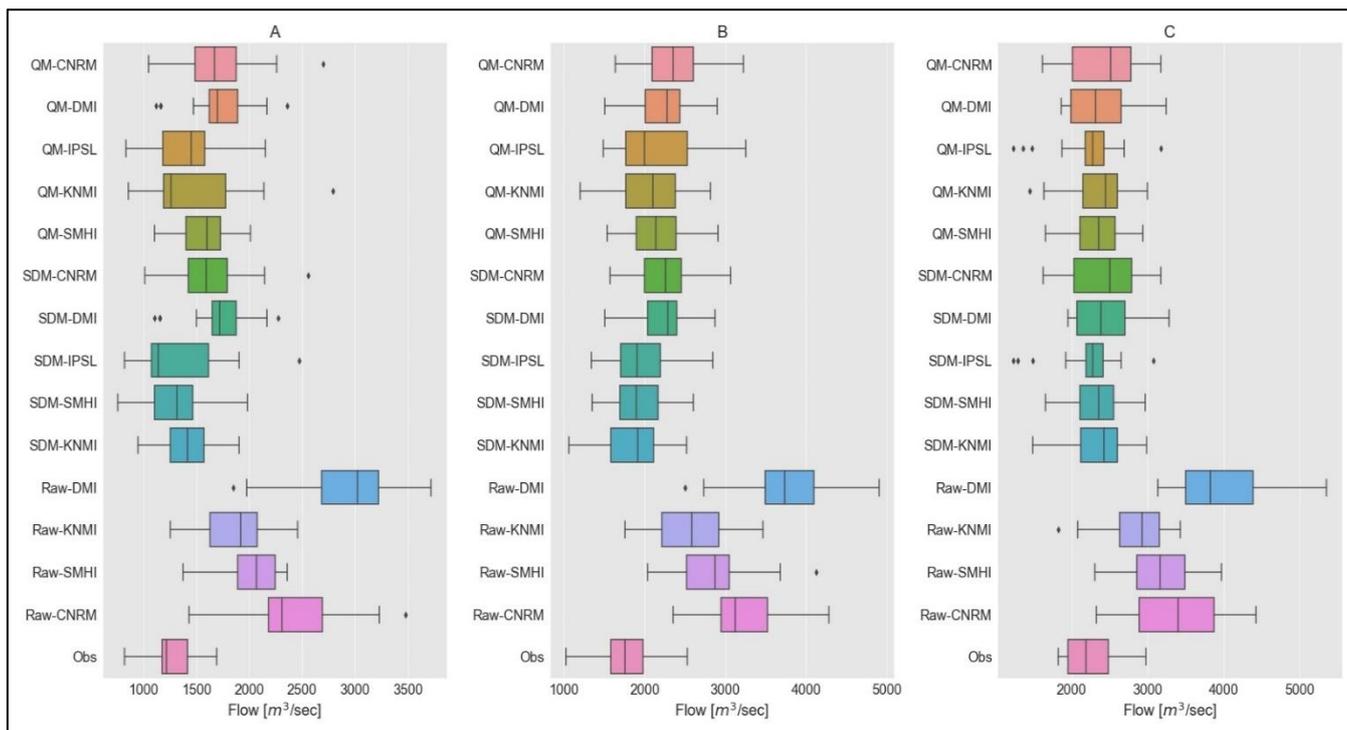


Figure 4-13: Long-term statistics at Lobith for the period (1999-2018). Figure A is the minimum 7-day low flows at Lobith, figure B is the long-term mean flow during the summer period (June to August) at Lobith, and figure C is the long-term annual mean flow at Lobith

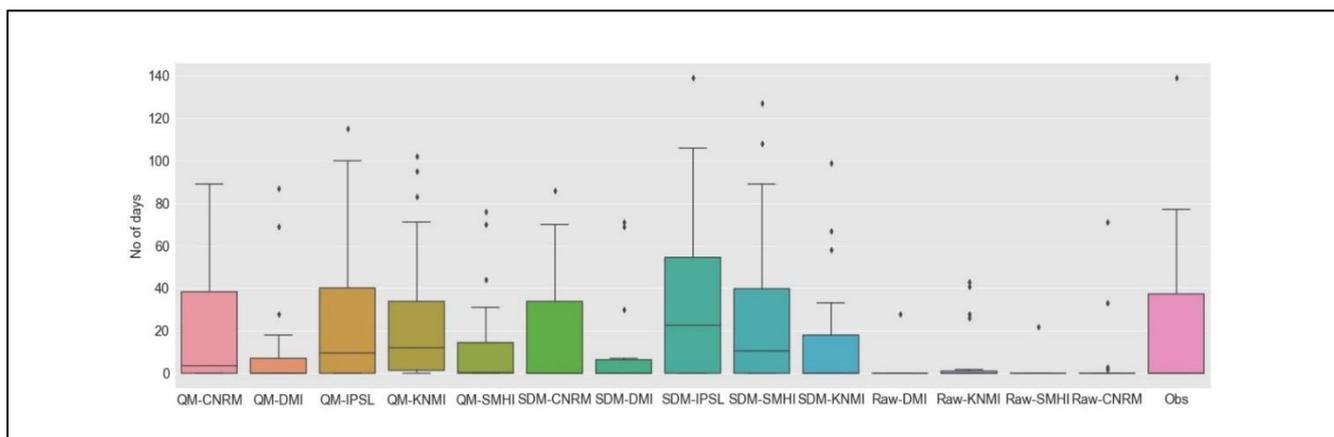


Figure 4-14: No of days below threshold (1100 m³/s) for the period (1999 - 2018) at Lobith.

This indicates that although bias correction methods provided better performance, but that still depends on the RCM of the subset and the underlying GCM.

Although the results depicted in figure 4.13 and figure 4.14 are about Lobith, similar analysis is performed to three other locations distributed within the Rhine basin. The three locations are Kaub, Cochem, and Rees. The reason of selecting these three locations, is because they are considered as bottlenecks for low flows in the Rhine river. In these locations it is shown that applying bias correction has resulted in improving the long-term statistics at two locations (Rees and Cochem), while one location (Kaub) provided better long-term statistics when using raw simulations instead of using bias-corrected simulations (see appendix III for the related graphs).

In general, the applied historical experiment indicated that using bias-corrected result in better performance than using raw simulations. This is based on the low flow analysis performed at Lobith and in three other locations.

4.3 Non-stationarity

The bias correction methods are tested under changing climate conditions as explained in section (3.3). The two bias correction methods (QM and SDM) are applied to two different seasons (JJA and DJF).

The DSST test is only applied to the KNMI dataset because it provided the closest estimates relative to the observations for the historical period (figure 4.9, 4.12 and 4.13). The dataset is dis-aggregated to a common grid of 5 km² to avoid the localized corrections and to allow for consistency and intercomparability.

The DSST applied in the precipitation datasets is first discussed and the test on temperature datasets is discussed later.

4.3.1 Precipitation

Summer period

The summer period (JJA) in the future raw climate simulations between (1999 - 2018) is corrected based on the winter periods (DJF) of the HYRAS dataset and raw simulations during the period (1979 - 1998) at Lobith. The result is shown in figure 4.15.

The QM method has resulted in improving the mean signal of the corrected EURO-CORDEX simulations for the summer period, with a mean precipitation of 2.6 mm and a standard deviation of 4.2 mm/d. This translates into a difference of 0.1 mm for the entire season relative to what is observed in the validation (1999-2018). The difference in estimating the standard deviation is 0.9 mm/day, but still it is better than the error in the standard deviation of the raw simulations (1.3 mm/d).

The SDM has also provided more accurate mean precipitation relative to the future raw simulations. However, the error in the estimated standard deviation using this method is higher than the future raw simulations and quantile mapping simulations.

The two methods have estimated the mean seasonal precipitation value better than the raw future simulations. QM method provided the closest seasonal standard deviation value (4.2 mm/d) relative to the actual observations (5.1 mm/d). However, the estimated seasonal standard deviation value from the SDM method is less accurate than the future raw simulations.

It appears in figure 4.15 that the average mean within the summer period (2.7 mm/d) in the validation period is higher than the mean of the winter period (2.4 mm/d). This makes the summer period the wet period and winter period the dry period at Lobith.

Quantile mapping method resulted in better performance than scaled distribution method in this category. It also appears that the methods can correct precipitation during relatively wet seasons based on relatively dry seasons.

Winter period

The winter period (DJF) in the future raw climate simulations between (1999 - 2018) is corrected based on the summer periods (JJA) of the HYRAS dataset and raw simulations during the period (1979 - 1998) at Lobith. The result is also illustrated in figure 4.15.

QM and SDM mean seasonal estimates are the same value (2.2 mm), both methods underestimated what is observed during the validation period. The raw future simulations estimated mean winter precipitation is identical to what is observed during the same period (2.4 mm).

The estimated seasonal standard deviation (DJF) values by QM, SDM and raw simulations are less than the actual standard deviation in the validation period (4 mm/d).

The two methods did not provide any improvements in the raw future estimates, making them less reliable in correcting precipitation during relatively dry seasons based on relatively wet seasons. This implies that the two methods cannot accurately correct lower precipitation values, thus failing to reduce the drizzling effect of climate simulations.

It is expected that SDM to perform better in correcting overestimated precipitation by climate models and perform less in correcting underestimated precipitation simulations. The results of this test indicate that QM and SDM are less reliable in correcting lower precipitation estimates. This indicate that the drizzling effect of climate models cannot be reduced using these methods, even after introducing the threshold in SDM and using sub-annual scales in QM.

The CDFs for this step can be seen in figure 8.2 and figure 8.3 in appendix II. In general, the CDFs of the corrected simulations are closer to the observed CDF than the future raw simulations.

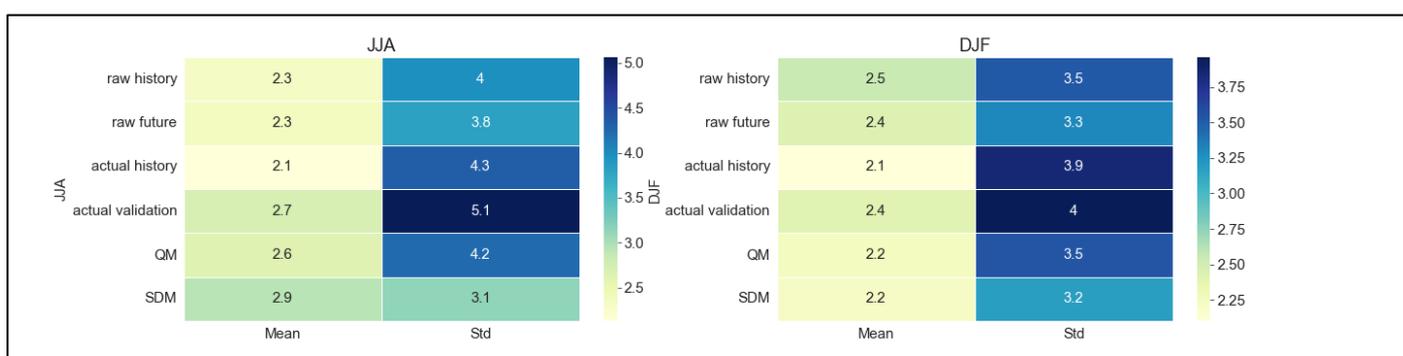


Figure 4-15: Precipitation correction experiment through the DSST for winter and summer periods. Raw history is the raw precipitation simulations in the period (1979 – 1998), raw future is the raw precipitation simulations for the period (1999-2018), actual history is the observed HYRAS measurements for the period (1979-1998), actual validation is the observed HYRAS measurements within (1999-2018), QM is the bias-corrected KNMI simulations using quantile mapping for the period (1999-2018) and SDM is the bias-corrected KNMI simulations using scaled distribution mapping for the period (1999-2018). Mean is the mean value during the specific season and std is the standard deviation during that season.

4.3.2 Temperature

Summer period

The summer period (JJA) in the future raw temperature simulations between (1999 - 2018) is corrected based on the winter periods (DJF) of the E-OBS dataset and raw temperature simulations during the period (1979 - 1998). The result of this correction is shown in figure 4.16.

QM and SDM mean seasonal (JJA) temperature values are 17 C°, while the actual observed mean seasonal temperature at Lobith is found to be 1 C° higher. The raw future simulations provided a mean seasonal (JJA) value of 16 C° and this is 2 C° lesser than the actual value. QM and SDM corrected simulations provided more accurate mean temperature estimates at Lobith.

The standard deviation of the raw and bias-corrected in the summer season is found to be around 3.2 mm/d. This is 0.2 mm/d lesser than the actual standard deviation at Lobith. There is no improvement in estimating the seasonal summer standard deviation when using QM or SDM. But they managed to provide better mean estimates.

The CDFs for this step can be seen in figure 8.4 and figure 8.5 in appendix II. The CDFs of both corrected datasets look better than the CDF of the raw future simulation.

Winter period

The winter period (DJF) in the future raw temperature simulations between (1999 - 2018) is corrected based on the summer periods (JJA) of the E-OBS dataset and raw temperature simulations during the period (1979 - 1998). The result of this correction is shown in figure 4.16.

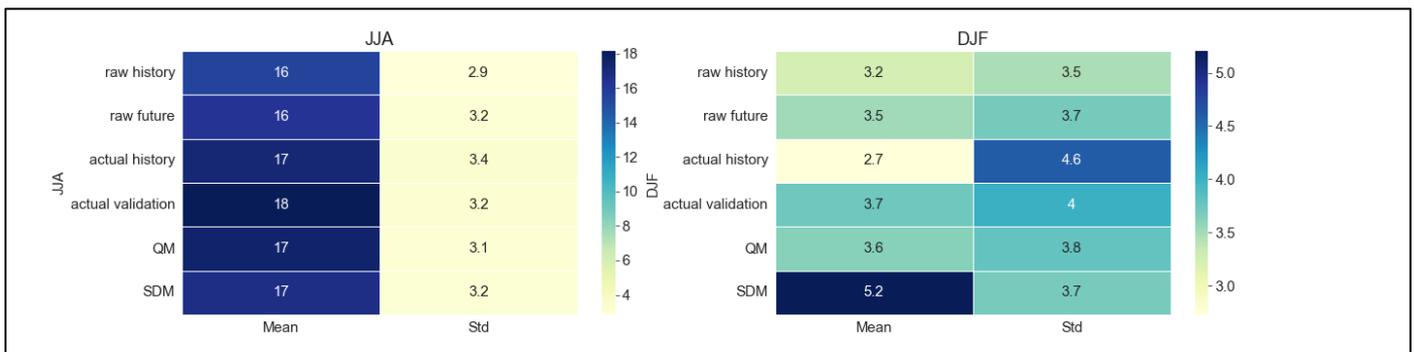


Figure 4-16: temperature correction experiment through the DSST for winter and summer periods. Raw history is the raw KNMI temperature simulations in the period (1979 – 1998), raw future is the raw temperature simulations for the period (1999-2018), actual history is the observed E-OBS measurements for the period (1979-1998), actual validation is the observed E-OBS measurements within (1999-2018), QM is the bias-corrected KNMI simulations using quantile mapping for the period (1999-2018) and SDM is the bias-corrected KNMI simulations using scaled distribution mapping for the period (1999-2018). Mean is the mean value during the specific season and std is the standard deviation during that season.

QM estimated the mean seasonal (DJF) better than SDM, as QM has improved the mean of the raw simulations by increasing it in average by 0.1 mm. SDM has overestimated the mean value and provided a false change signal compared to what actually happened during the validation period. Also, it managed to provide the same standard deviation as the raw future simulations (both 3.7 mm/d), indicating no improvements in the signal.

The cause of this contradicting behavior in SDM is because the scaling factor (equation 3.11) is very sensitive to the ratio between the standard deviation of the observed temperature values to the standard deviation of

the raw simulations during the reference period ($\frac{\sigma_{OBS}}{\sigma_{MODH}} > 1.3$). This resulted in overestimating the mean signal of the corrected simulations.

The estimated standard deviation in the winter seasonal using QM method is the nearest to the observed value. QM has performed better in correcting relatively cold temperature periods based on relatively hot periods in the reference periods. SDM did not manage to correct lower temperature rates to the level of QM.

4.4 Projecting climate change impacts by 2060

The two bias correction methods (QM and SDM) are utilized to correct the future EURO-CORDEX projections for the period (2020 – 2060) and examine the potential future changes at Lobith by 2060. These future simulations are corrected based on the available observational datasets (in section 2.2.1) and the historical simulations during the historical period (1979 – 2019).

The modelled flows in the future which are obtained by forcing wflow_sbm with these corrected simulations are compared in section 4.4.1, relative to the actual discharges at Lobith between 1979-2019. This is followed by analysing low flow events by some of the metrics that are outlined in appendix I. Then, the water levels translated from the projected discharges are analysed. As a final step the results of the applied regional trend in the bias-corrected simulations are discussed.

4.4.1 Lobith

Discharge

The modelled discharges at Lobith from the corrected simulations for the period (2020-2060) is shown in figure 4.17. The daily discharges for the period (2020 - 2060) are averaged daily to construct this figure. The raw simulations for the same period are shown in this graph to compare with the actual discharges and the discharges from the corrected simulations.

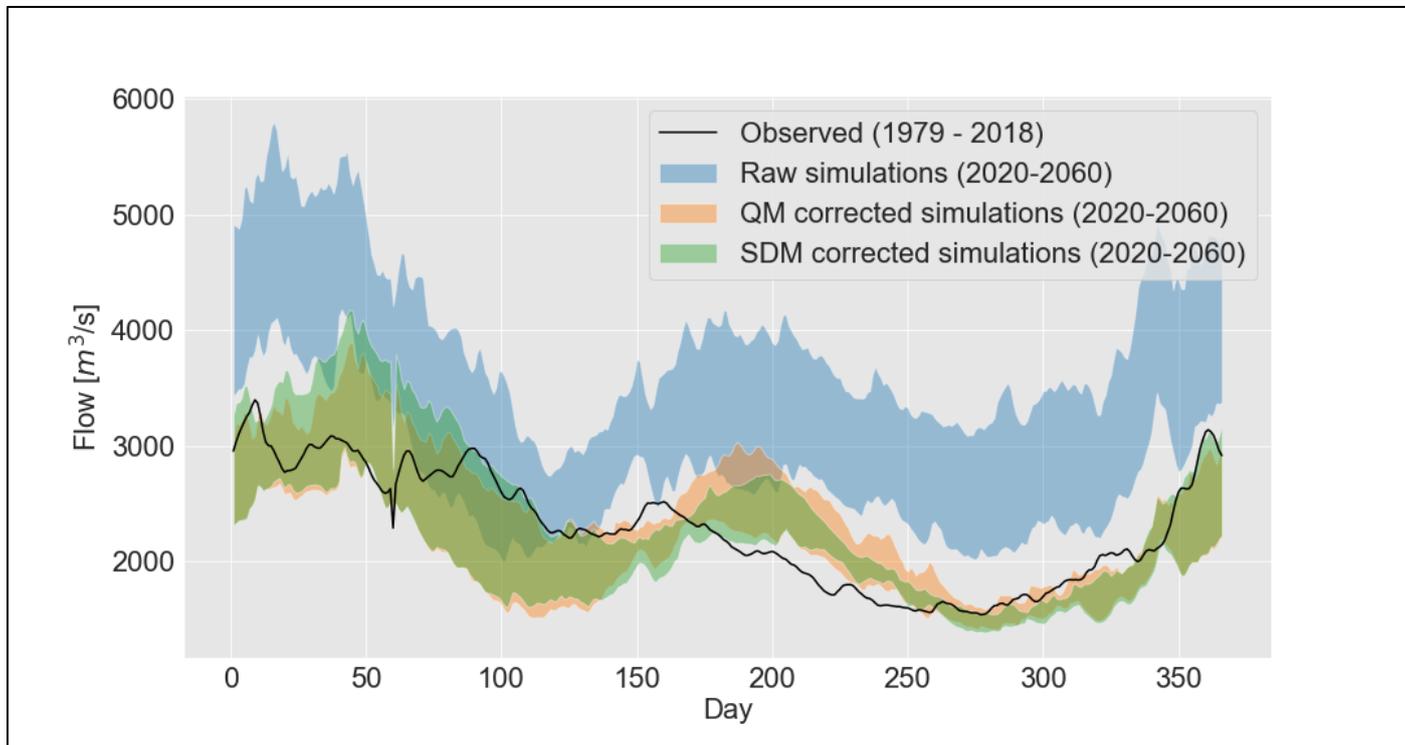


Figure 4-17: Daily averaged discharges for the period (2020 - 2060) for the raw and bias-corrected simulations compared to the actual discharges in the reference run (1979 - 2019).

The raw simulations appear to be higher than the actual discharge at Lobith for the period (1979-2018), like the historical experiment (figure 4.6). An interesting point is that the future raw discharges are higher than

the historical raw discharges. In figure 4.6, one could see that the raw historical ensembles are able to capture some of the observed discharges between (day 60 – 180) and some days in January, November, and December. However, the future raw ensemble discharges are only captured in the period between (day 70 – 130).

The same is exhibited in bias-corrected simulations for the future period, although the envelopes of QM and SDM are matched most of the time. But they are not capturing the observed flows for the period (1979 – 2018), as the corrected projections in the historical experiment (figure 4.12).

In figure 4.17, QM and SDM projections seem to agree well, indicating similar ability in removing the systematic bias. SDM projections are more constrained in the second half of the year compared to QM projections, the situation is inverted for the first half of the year, where QM projections are more constrained.

The bias-corrected projections result in daily discharges that are consistent with the observed discharges from the start of the year until day 140. Then, for the rest of the year the envelopes of QM and SDM ensembles are more constrained. This indicates a potential change in the average daily discharge at Lobith in the future period (2020 – 2060), relative to the historical situation.

In figure 4.12 and figure 4.17, the bias correction methods managed to reduce the bias in high discharges from the raw simulations. but the general hydrograph shape of the corrected ensemble envelope (QM or SDM) is like the raw ensemble envelope. This suggests that QM and SDM can reduce the average bias in the raw simulation, but the uncertainty in the GCM is not reduced. As the subsets that originate from the same GCM provide the same general hydrograph shape (see figure 4.18).

In figure 4.18, all the corrected subsets that originate from the ICHEC-EC-EARTH model (GCM) share the same hydrograph, the subsets are DMI, KNMI and SMHI. Also, IPSL and CNRM share the same hydrograph due to originating from the same GCM, which is MOHC- HadGEM2-ES. It is also noticeable that the QM and SDM discharges from the same subset are similar. This indicate that the two bias correction methods can not alter the GCMs' effects on the hydrology.

Most of the corrected subsets indicate lower flows in the first half of January and higher flows for the second half of January, a decline in the average flows from the end of February to the end of June, followed by higher flows for the period between the first of July till the end of September and a slight decline in the daily average flows until the end of the year.

SMHI corrected subsets (QM and SDM) provided daily average discharges that are lower than the observations by almost 500 m³/s for the first two months. While, IPSL corrected subsets daily average discharges, provided higher discharges between early January to end of March, reaching 1000 m³/s around mid-February.

Although it is hard to overcome the uncertainty imposed by the parent GCM, the ten corrected subsets agree on two things. The first is having relatively higher flows than the observations just before the summer period (day 180) to the end of August. The second aspect is the relatively lower flows than the observations from the end of the summer period till the end of the year.

To investigate the cause of this, two other locations are analysed, Rockenau and Andelfingen. The reason is to evaluate the projected changes in the contribution of snow melting during the summer period at the upper parts of the Rhine. One of these locations is in the Alps (Andelfingen), thus highly influenced by snow melting (see appendix IX). The results at Andelfingen (figure 10.4) suggest that during the period (day 180 – day 200)

most of the corrected simulations indicate an increase in the average daily flow due to snow melting. However, from this day onwards there is a sharp decrease in the contribution of melted snow to the summer flows. The discharges at Andelfingen remains lower than the observed at the end of the year as well.

The projected discharges at Rockenau are lower the daily average discharge of the reference period. All the bias-corrected subsets indicated lower projected discharges at Rockenau for the second half of the year (figure 10.2). Based on the analysis at Rockenau and Andelfingen, it appears the reason of having relatively high discharges from day 180 to day 240 is the larger contribution of melted snow compared to the reference run. This also indicates a change in the timing of the summer season, i.e. shifted two or three weeks earlier than the current start of the summer period (late June or early July).

The reason of having relatively lower discharges at the end of the summer period till the end of the year, is due to lower contribution of melted snow near the end of the summer period. There is also a general decrease in the average projected flows for the first two weeks in the year, but no clear reason is found.

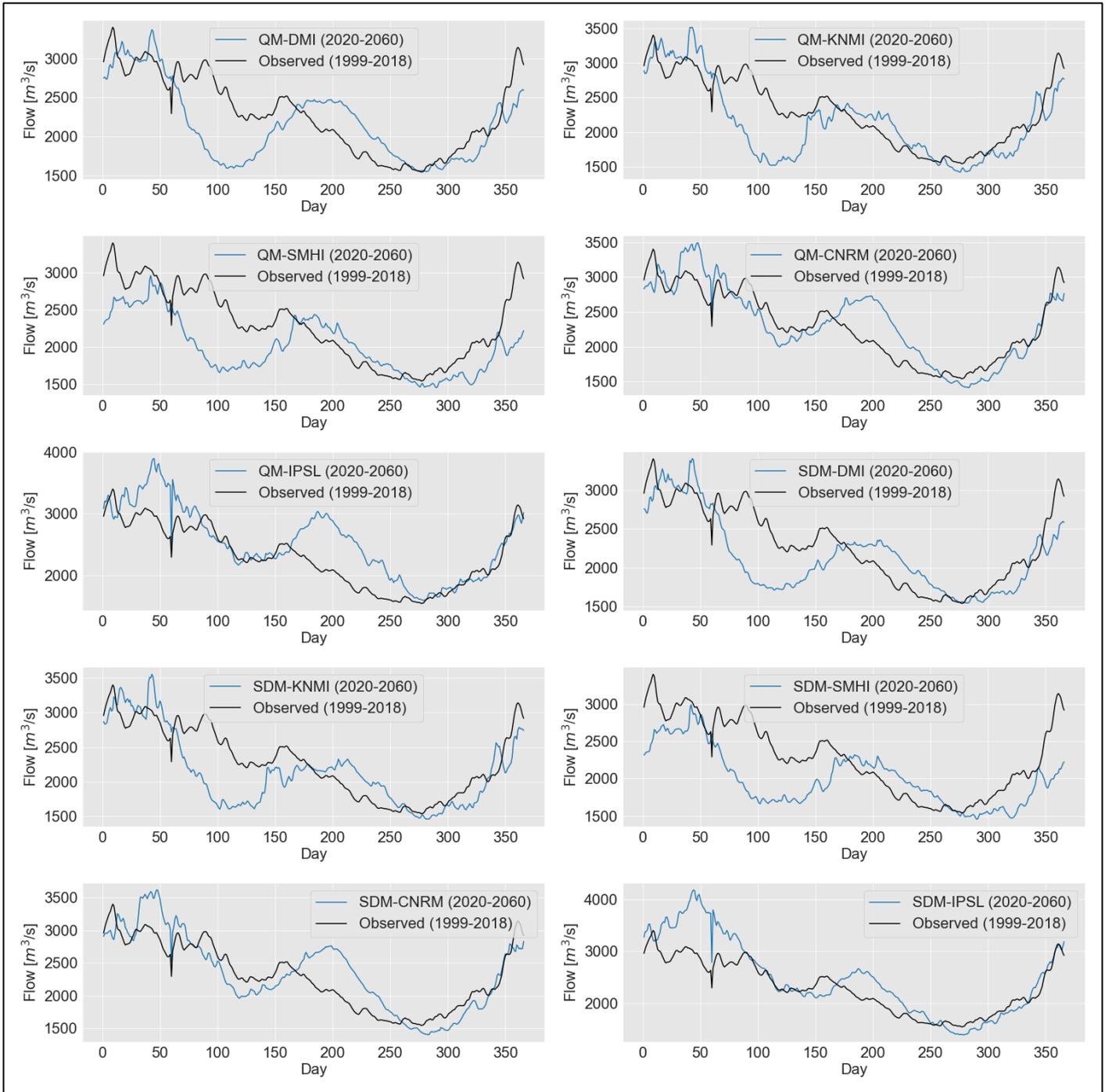


Figure 4-18: average daily modelled discharges for each subset during the period (2020 - 2060) vs the actual observed discharge (1999 - 2019).

Low flows

The daily projected discharges at Lobith for the period (2020 – 2060) are analysed to obtain three metrics that are relevant in assessing low flows at Lobith. The three metrics are the minimum 7-day low flow, annual long-term mean discharge, and the number of lengths below threshold ($1100 \text{ m}^3/\text{s}$). Figure 4.19 contains these metrics.

In figure 4.19(A), minimum 7-day low flow values for all bias-corrected simulations are listed. All the corrected simulations indicated extremely low minimum 7-day low flow values, except for the SDM corrected DMI subset. This is clearer when looking at the illustrated histograms in figure 4.20, which provide a summary of the low flows below the established threshold ($1100 \text{ m}^3/\text{s}$).

Nine of the subsets shown in figure 4.20 contain extremely low flow events at Lobith ($< 700 \text{ m}^3/\text{s}$), that have not been witnessed at Lobith during the period (1979 – 2018). The two corrected KNMI subsets provided extremely low discharges with the highest number of occurrences. The SDM corrected DMI subset is the only subset that has provided lengths that are lower than the observed.

From figure 4.19(B), only Four corrected subsets have provided annual long-term mean discharge values that are close to the observation (SDM-DMI, QM-DMI, QM-KNMI and QM-SMHI). The rest of the corrected subsets provided annual long-term mean discharge values that at least $50 \text{ m}^3/\text{s}$ lesser what is observed during the period (1979 – 2018).

The average of the lengths below threshold in the observed dataset is about 40 days per year. Five subsets have predicted higher average values (QM-KNMI, QM-SMHI, SDM-SMHI, SDM-IPSL and SDM-KNMI), as shown in figure 4.19(C). This indicates that low flow events are going to occur more frequent and last longer. All the corrected subsets have a certain outlier that is longer than 120 days. The comprehensive list of lengths below threshold per climate simulation and year see figure 8.6. In figure 8.6, it can be seen that in the year 2035 and year 2059 are the years with longest number of lengths below threshold in six simulations (QM-DMI, QM-KNMI, QM-SMHI, SDM-DMI, SDM-KNMI and SDM-SMHI).

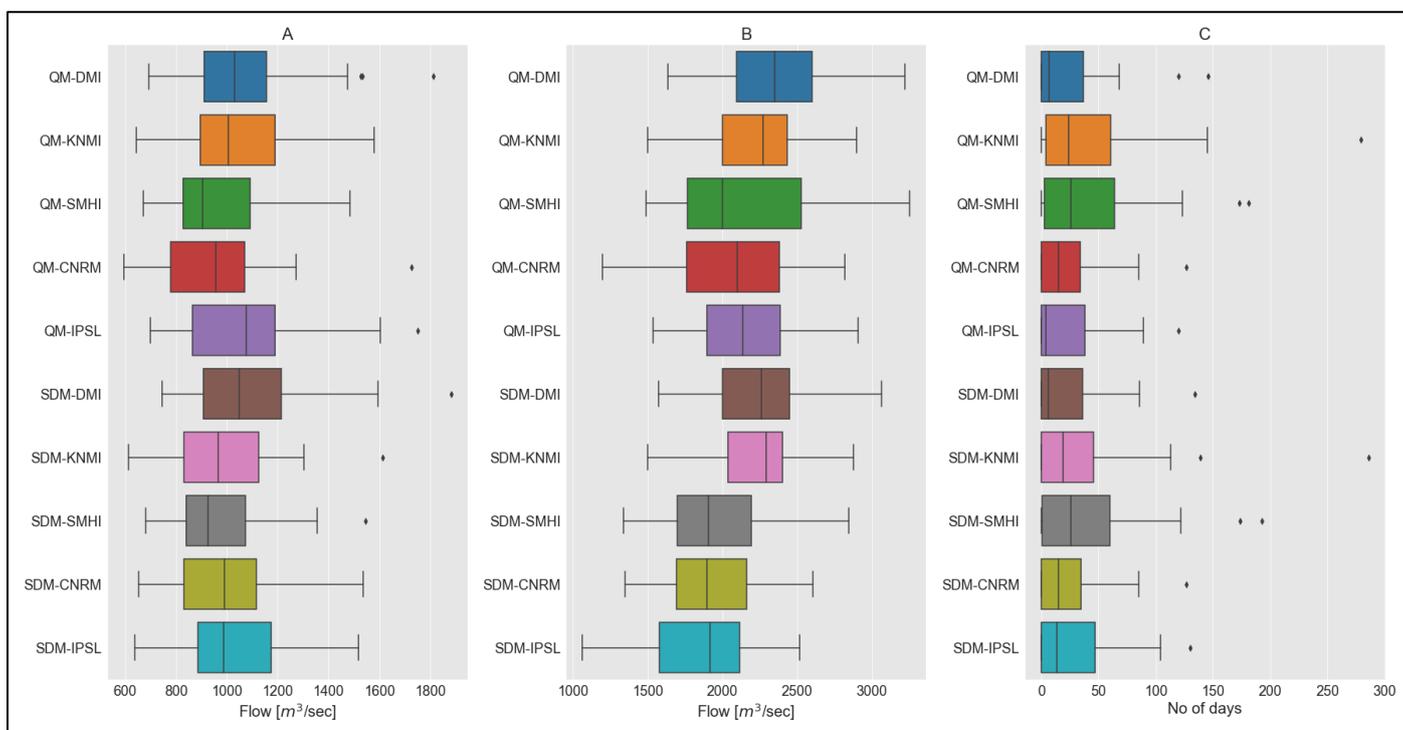


Figure 4-19: Long-term annual mean discharge at Lobith for the period (2020 – 2060). Figure A is the minimum 7-day low flow, figure B is the annual long-term mean discharge at Lobith, and figure C is the lengths below a threshold of $1100 \text{ m}^3/\text{s}$.

Lobith is set to experience unprecedented low flow events in the period (2020 – 2060) relative to the observed period (1979 – 2018) due to climate change (figure (4.20)). To reflect the change in the seasonality of the discharges at Lobith the seasonality ratio is calculated (appendix I), by obtaining the Q95 flow values during the summer period (JJA) and the winter period (DJF). When the SR value is lesser than 1, this indicates low flows during the summer period and high values during the winter periods, and vice versa (Laaha and Blöschl, 2006). The SR value of the observed discharges is about 0.71 as shown in figure 4.21. The detailed values of the Q95 values for the observed and corrected subsets is illustrated in figure 8.7 (see appendix II).

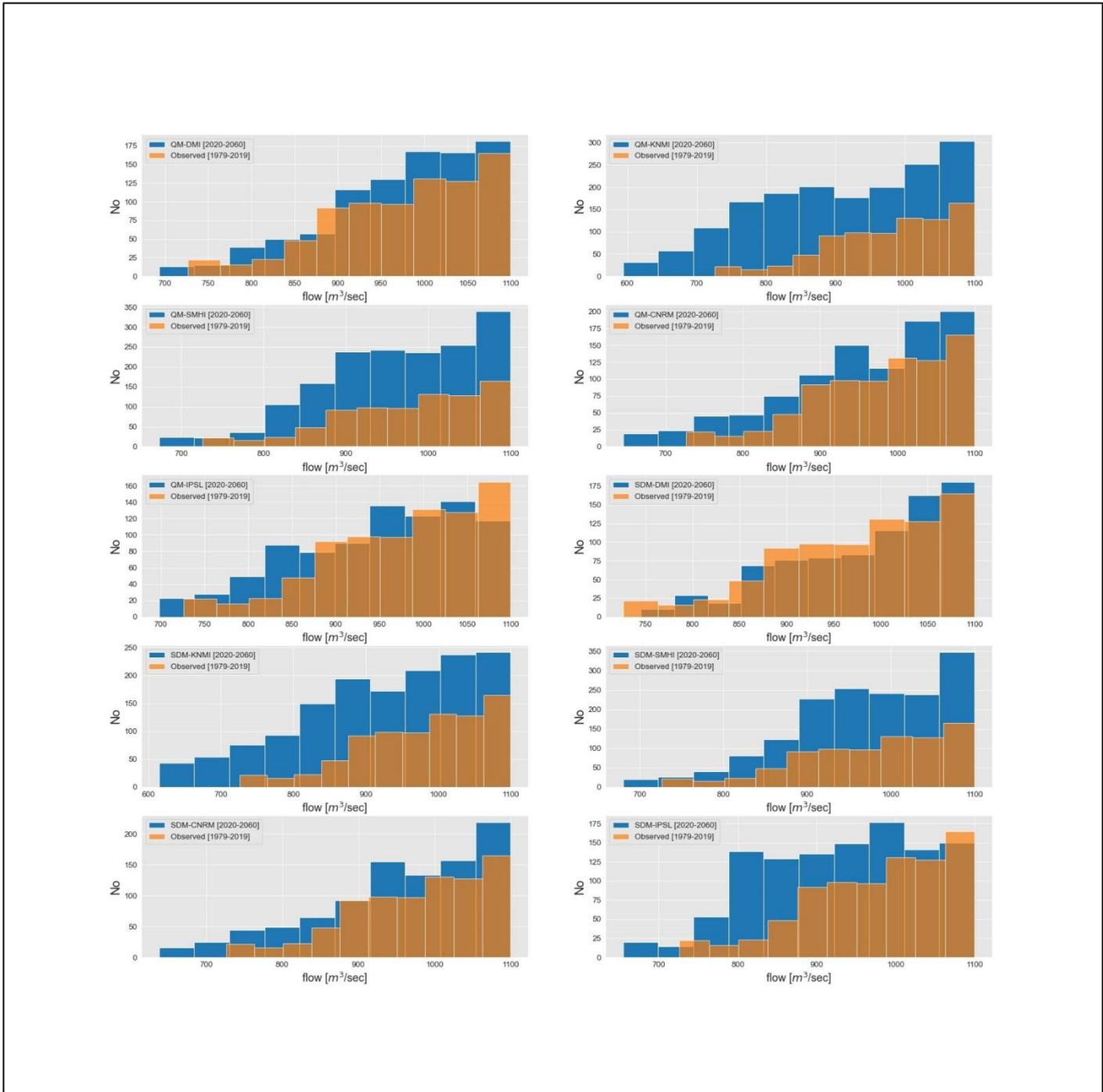


Figure 4-20: Histogram of the lengths of low flow periods of the bias-corrected simulations (2020 - 2060) and observed discharges (1979 - 2018).

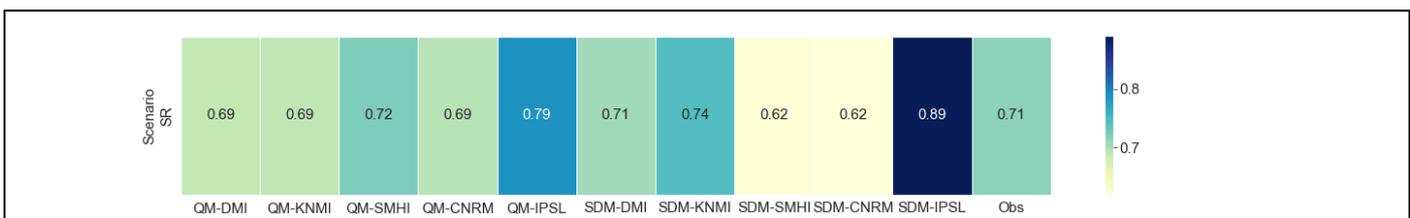


Figure 4-21: SR values for the bias-corrected subsets (2020 – 2060) and for the observation (1979 – 2018).

Six of the corrected subsets indicated stronger seasonality at Lobith during the summer period ($SR \leq 0.71$), that means in the coming period (2020 – 2040) the Rhine river is set to experience lower flows during summer and higher flows in winters. The difference in terms of the SR value is small. However, when looking at figure

8.7 it looks that Q95 values in most of the corrected subsets are lower than the observed value during winter and summer period.

This analysis indicates that besides the slight change in flow seasonality at Lobith, there is an average decline in the flow values for the period (2020 – 2060) relative to the historical period (1979 – 2018).

Water Levels

The fitted equation (A9) in appendix (I) is used to translate the daily projected discharges from wflow_sbm to water levels at Lobith expressed relative to the Normal Amsterdam level (NAP). The result of this translation is shown in figure 4.22.

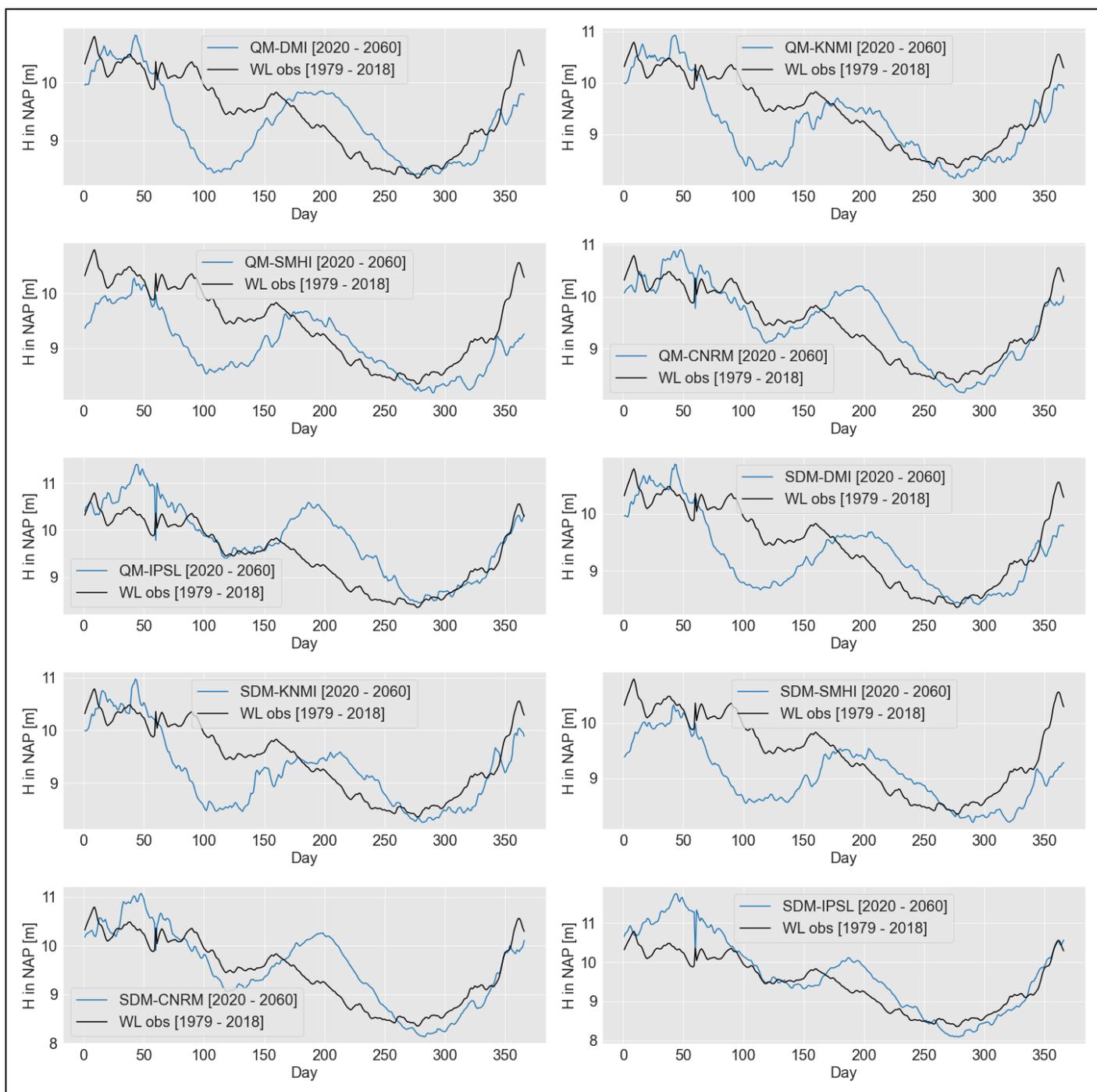


Figure 4-22: Detailed projected water levels at Lobith for the period (2020 - 2060) relative to the observed reference run.

The projected changes in the water levels follow the corresponding change in the projected discharges. For the period from day 150 to day 250, there is a consensus between the corrected simulations in providing relatively higher water levels than the average daily water levels at Lobith for the historical period (1979 – 2018). This is followed by relatively lower water levels until the end of the year. This is due to the lower contribution of melted snow at that period.

Most of the corrected simulations provided higher average daily water levels from mid-January to the end of March than the translated water levels from observed discharges. The corrected SMHI subsets (QM and SDM) is the only corrected subset that showed lower water levels between mid-January to the end of March.

There is an average (all corrected models are included) decline of 20 cm in the average projected water levels relative to the average observed water level from the end of the summer period until mid-January of the next year. There is an average increase in water levels of 30 cm between the first of June to mid-August.

There is a contradicting behavior between the bias-corrected simulations during the period of Mid-January till the first of June, therefore this period is excluded from the water level analysis.

4.4.2 General climate trends in the Rhine basin

In addition to the low flow analysis discussed in the section 4.4.2. The general trend of the three climatic variables (precipitation, evaporation, and temperature) in the Rhine basin is evaluated through a regional Mann-Kendall trend test (Fathian et al., 2016). The test is implemented on the bias-corrected datasets for the period (2020 – 2060). The result of this test in each of the climatic variables are discussed on the following subsections².

The results of this test are thirty sets, most of these results are similar. The distinct results are the only ones communicated in this section.

Precipitation

Most of the bias-corrected datasets indicated no general trend in precipitation, with a small exception in some parts in Germany and Luxemburg (figure 4.23). There is an increase in precipitation in some areas in Germany, however these areas are not within the Rhine basin.

² The shape of the Rhine basin is not shown in the precipitation and evaporation graphs, because of how the bias-corrected netCDF files were written in Python.

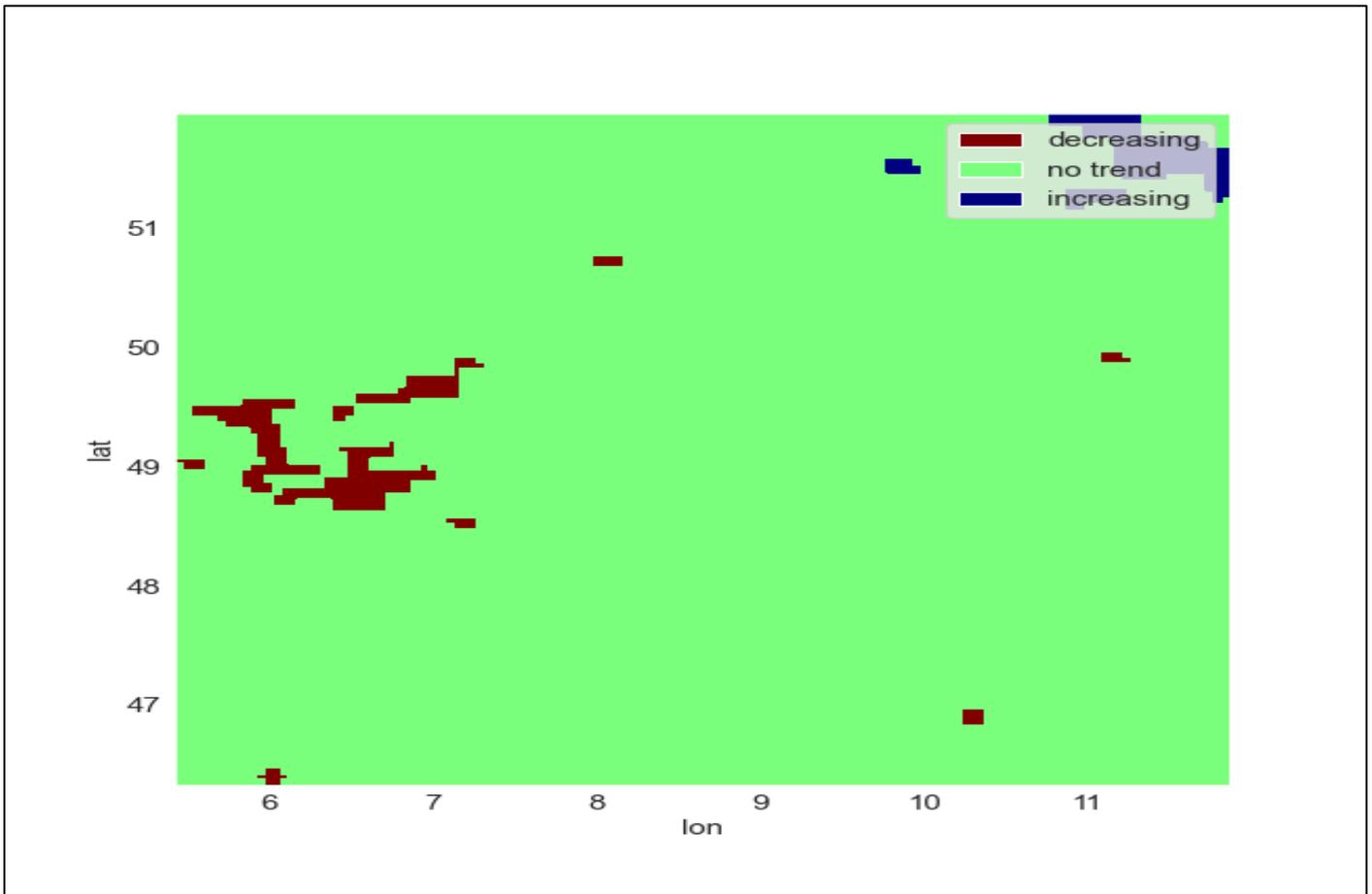


Figure 4-23: Projected precipitation trend in the bias-corrected datasets

Temperature

There is a consensus between the bias-corrected simulations on regional warming within the Rhine basin (figure 4.24). An increasing temperature trend indicates warming. There are few grid cells in Switzerland and Germany in which the test didn't find any general trend. However, based on the RCP8.5 scenario it is expected to experience global warming. This is a prove that the applied bias-correction methods did not distort the change signal of the climate simulations.

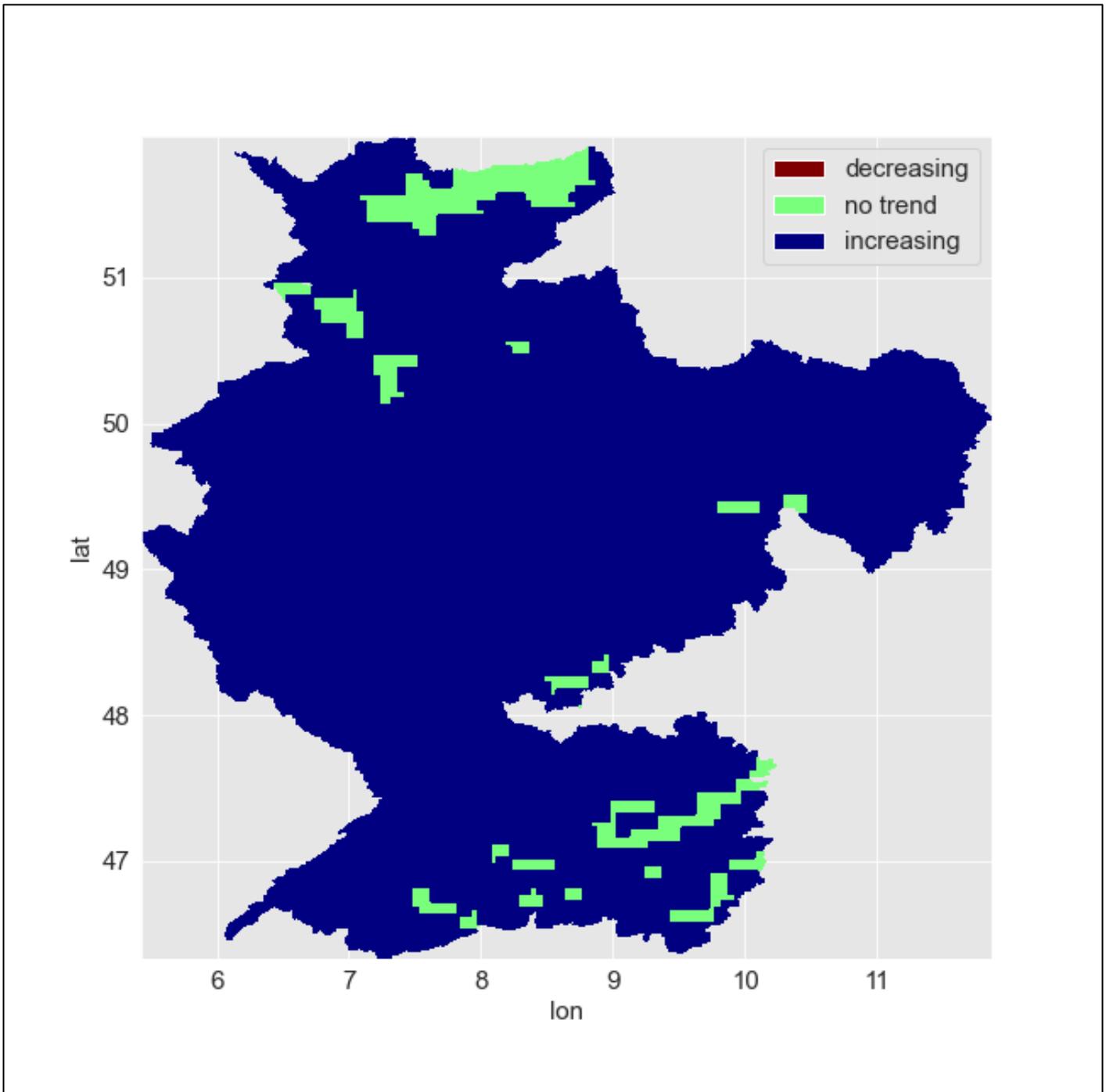


Figure 4-24: Temperature trend in the bias-corrected datasets for the period (2020 - 2060).

Evaporation

In most of the Rhine basin, the regional trend test did not find any noticeable changes in the evaporation trend. However, there are some parts in the Alps which have indicated an increasing trend in evaporation. This is because of the increasing warming in these parts of the Rhine basin that is coupled with the available water in the snowpacks.

Although in figure 4.24, there is a general increase in the temperature rates throughout the Rhine basin. Except for the Alps, none of these grid cells hinted at any increase in potential evaporation based on Makkink formula (Makkink, 1957). There is no clear reason is found for this contradicting results.

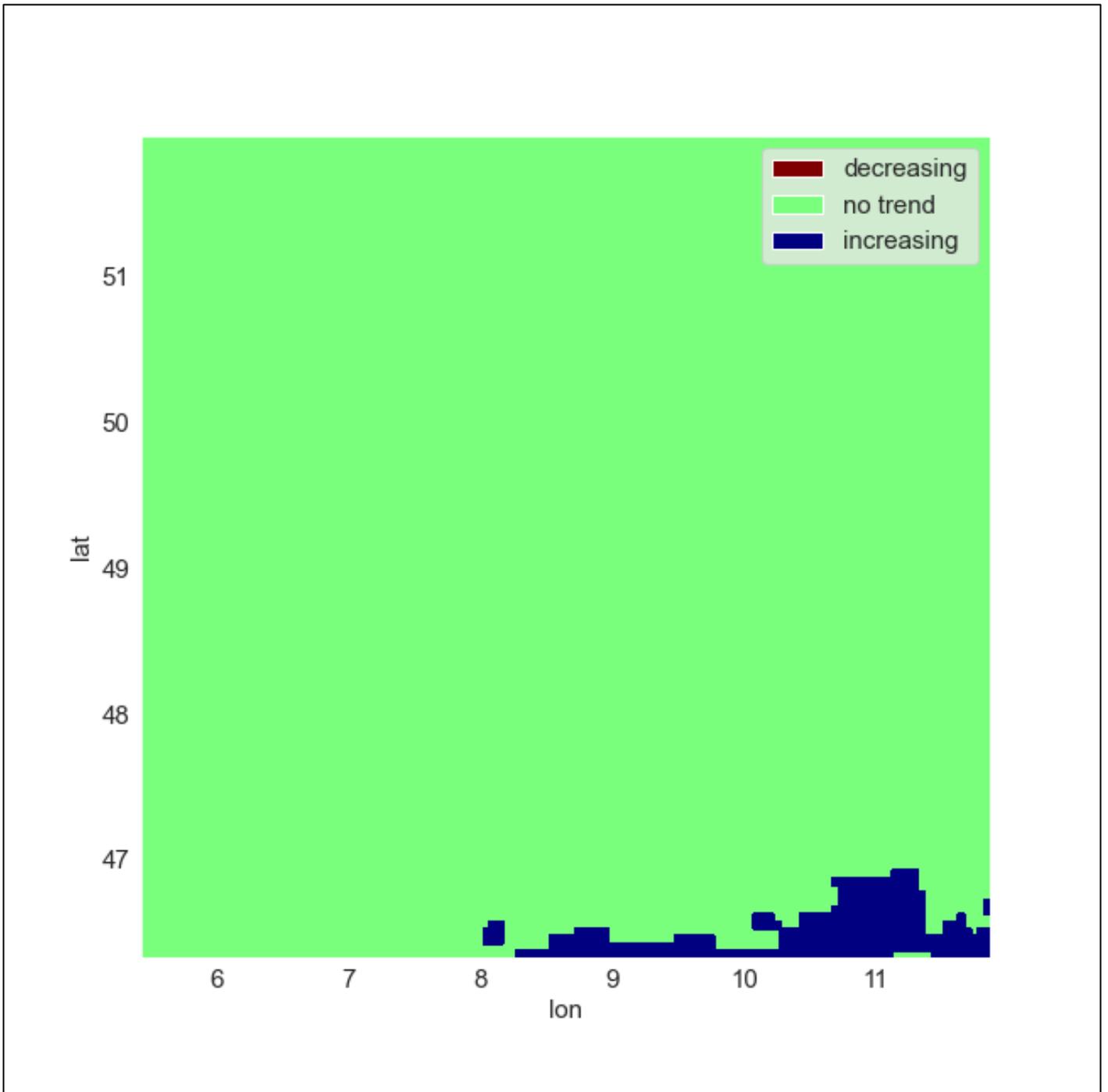


Figure 4-25: Makkink evaporation trend in the bias-corrected simulations for the period (2020 - 2060).

5| Discussion

In section 5.1 some of the results of the calibrated wflow_sbm model are discussed. The results of the modelled discharges using raw simulations for the period (1979 – 2018) are investigated in section 5.2. Section 5.3 elaborates on the limitations of using performance-based weighting methods in this project. Then, section 5.4 assesses the results of the historical experiment. The main findings of the DSST test are investigated in section 5.5. The projected impacts of climate change on the low flows at Lobith are evaluated in section 5.6. Section 5.7 compares the adjusted quantile mapping method to the scaled distribution mapping in the context of modelling the Rhine. The research limitations are mentioned in section 5.8. Finally, section 5.9 indicates further research areas as possible extension to this project.

5.1 Hydrological model

The calibrated wflow_sbm model is used in modelling the hydrology of the Rhine. Reasonable performance metrics ($NSE = 0.7$, $NSE_{\log} = 0.81$ and $KGE = 0.82$) are obtained when evaluating the modelled discharges at Lobith based on the actual discharges for the period (1979 – 2018). It appears that the modelled low flows are slightly underestimated (90 m³/s is the difference in estimating the Q95 value), while the modelled high flows are overestimated (error = 300 m³/s in estimating Q5).

The histogram of the maximum consecutive low flows in a year shows that wflow_sbm can capture most of the dry flow events (figure 4.4). The model resulted in values that are 20 % higher than the observations in six years (1991, 1996, 2003, 2009, 2012 and 2014).

The model is less reliable in studies that focus on high flows (i.e., flooding related studies) and more reliable in studies with a focus on low flows. Since the focus of this study is on analysing low flows, wflow_sbm is found not to be a limiting factor.

5.2 Uncorrected climate simulations

The calculated long-term statistics of the modelled discharges from raw simulations (lengths below threshold, 7-day minimum flow and long-term mean annual discharge) provided estimates that are higher than what the actual discharges have provided. One of the main causes of the high modelled flows, is because the raw temperature simulations for the period (1979 – 2019) are in average 2 C° colder than the observed temperature. This has caused wflow_sbm to exhibit lower losses, thus providing flows that are higher than the observed.

This indicates that these raw simulations cannot be used in analysing the flow changes at Lobith, based on their performance in the reference period (1979-2018).

5.3 Performance-based weighting

The weighting of raw climate simulations based on historical observations, either by meteorological variables (ClimWIP) or observed discharges (REA), appears to restrict the change signal of the raw simulations. The results of these weighting methods shown in figure 4.7 and 4.8, suggest that the two methods are indeed able to restrict the change signal based on the observations. However, these methods are not able to overcome the wet precipitation bias in the underlying RCMs. This has prevented these methods from providing results that are close to the observations.

The two methods are not able to extrapolate from the ensemble range, as the weighted discharge is based on the values of the ensemble. Other methods like the Bayesian inference method (Wu et al., 2014), the

kriging climate change method (Drignei, 2009) and the spatiotemporal REA method (Tegegne et al., 2019) can reduce this problem and lead to more reasonable results.

An example on how to apply one of these methods is shown in appendix X, where three additional stations (Kaub, Basel and Cochem) are considered in computing the spatiotemporal RCM weights at Lobith for the period (1999 – 2018). The final weighted discharges are better than the ones generated using the original REA method of Giorgi and Mearns, 2003. As the weighted discharge is found to lie in the lower band of the ensemble envelope.

5.4 Bias correction in the historical experiment

The corrected climate simulations for the period (1999 – 2018) resulted in modelled discharges that are closer to what is observed in the same period. QM and SDM methods have provided the same envelope in their daily average discharges for the same period, indicating similar capabilities in reducing the systematic bias.

The comparison between the results of the bias-corrected simulations and raw simulations relative to the observed discharges at Lobith (figure 4.13 and 4.14), shows that the bias correction methods have improved the long-term statistics of the simulations by moving them closer to the observations.

Bias correction methods cannot overcome the embedded errors in GCMs. Although, the corrected subsets indicated a better performance, but the hydrograph shape of each RCM still has a different shape than the actual hydrograph. Furthermore, RCMs that originate from the same GCM share the same hydrograph shape. The corrected DMI, KNMI and SMHI subsets share the same hydrograph because they are derived from the ICHEC-EC-EARTH model. CNRM and IPSL corrected subsets have the same hydrograph, as these RCMs are derived from the same GCM (MOHC-HadGEM2-ES).

Three corrected subsets (KNMI, SMHI and IPSL) have provided better results than the rest. The reason appears to be related to their natural ability to estimate the long-term statistics as raw simulations (without bias correction). The raw KNMI and SMHI subsets have estimated the four metrics (lengths below threshold ($1100 \text{ m}^3/\text{s}$), 7-day minimum flow and long-term mean annual discharge) better than the DMI and CNRM subsets.

This indicates that QM and SDM methods are more of a catalyst to the climate simulations. In other words, it appears that the best performing corrected simulations are the best performing raw simulations. Also, the least performing corrected simulations are the least performing raw simulations. In any case, applying bias correction to the climate simulations has proved to always be better than using raw simulation at Lobith.

This is supported by an applying the same analysis at three other locations (Kaub, Rees and Cochem) within the Rhine basin. The analysis suggests that the bias correction methods have improved the long-term statistics at two locations (Rees and Cochem), while the simulations are better of any correction at Kaub.

The bias-corrected simulations at Kaub have provided estimates with higher errors compared to the raw simulations. The reason appears to be the `wflow_sbm` is not performing well at Kaub compared to Cochem. The NSE value is found to be lower than 0.4 at Kaub, based on the available observations from (1979 – 2000). While at Cochem, the NSE is found to be of 0.56. This suggests that the success of bias correction methods at certain locations depend on the skill of the hydrological model.

5.5 Non-stationarity testing

The DSST indicated that the two methods can work well in certain climatic conditions. Regarding precipitation simulations, QM and SDM can correct extreme precipitation. However, the two method did not provide any improvements when correcting lower precipitation values. Implying that they cannot be relied upon in overcoming the drizzling effect in climate simulations.

For temperature, the two correction methods can correct unprecedented warm temperatures. This is based on the provided estimates relative to the E-OBS. QM is better than SDM in correcting lower future temperature values based on higher temperature. As QM provided more closer estimates to the validation period than SDM. The raw simulation provided better estimates than the SDM in this part of the test.

5.6 Future climate change impacts on low flows at Lobith

The modelled discharges projected a relative increase in the average flows from the start of the summer until end of August. This is followed by lower average discharges until the end of the year. This is due to earlier snow melt because of climate change. Then, at the end of the summer period the snow melt contribution at the Alps is projected to decrease, resulting in lower flows. This is based on the analysis of two other locations (Rockenau and Andelfingen), where the corrected projections indicate lower flows at Andelfingen and Rockenau from the middle of the summer period until the end of the year (see appendix IX).

For the first half of the year, it is difficult to infer the possible impacts by 2060. This is because the modelled discharges do not seem to agree well with each other.

Based on the RCP8.5 scenario, the bias-corrected simulations suggest more extreme low flow events that have not been experienced in the Rhine. The histograms of the modelled discharges shown in figure 4.20 suggest having extremely low discharges that are below 700 m³/s. These discharges are lower than the minimum discharge observed during the period (1979 – 2018). In addition to that, discharges that are in the range of (1100 – 900 m³/s) are going to occur more frequently and last longer.

Considering the extremity of these results, the regional trend test applied to the bias-corrected projections did not indicate a general trend in precipitation and evaporation. However, it has indicated regional warming across the Rhine basin. The reason of this is not fully known. Only some parts in the Alps indicated an increase in potential evaporation, due to having enough water supply for evaporation.

5.7 Empirical quantile mapping and scaled distribution mapping

The two bias correction methods: quantile mapping (QM) and scaled distribution mapping (SDM) have improved the quality of raw EURO-CORDEX simulations in modelling the Rhine river. The similar performance of the two methods throughout this study indicated the presence of the systematic biases in EURO-CORDEX (CMIP5) simulations.

Scaled distribution mapping method is proposed by Switanek et al., 2017 as a better alternative to the empirical quantile mapping method. However, this is not the case for the Rhine basin. Adjusting the QM method to work based on the monthly scale has proven to produce more consistent results than SDM in the Rhine basin. It appears that SDM is more suitable for correcting higher values (precipitation and temperature) and not suitable for correcting lower values (precipitation and temperature). This is mostly related to the scaling factors that are used in this method.

5.8 Research limitations

5.8.1 Observational datasets

The datasets used to correct climate simulations are of high resolution, however there is always uncertainty with data. The first issue regarding the used dataset is the fact that each meteorological variable is selected from a different dataset. However, that cannot be averted as these sets are the ones that cover the entire period for the reference period considered in this study. The other issue is that each of these datasets has a different resolution (31, 12, 1 km²), which adds another level on uncertainty due to the additional interpolation procedure to fit it based on the wflow_sbm resolution (1 km²), except for HYRAS.

E-OBS is often used in bias correction methods, nonetheless it is an interpolated dataset where there are embedded errors from false measurements and interpolation method. E-OBS is found to have lower errors in temperature, and large errors in precipitation (Hofstra et al., 2009), and based on this reason E-OBS temperature dataset is chosen.

The Makkink evaporation calculated from ERA-5 inputs is also subjective to the uncertainty of the input parameters used to calculate Makkink evaporation (pressure, shortwave downwelling radiation, and temperature) (Makkink, 1959). According to (Babar et al., 2019), ERA-5 suffers of large errors in estimating shortwave downwelling radiation and temperature in high altitudes. Additionally, the temperature used to estimate the Makkink is based on ERA-5 and not E-OBS. ERA-5 has the coarsest resolution out of the three, therefore it is also prone to larger errors due to statistical downscaling.

HYRAS precipitation dataset has the finest resolution (1 km²). However, the REGNIE method used in creating this dataset is sensitive to the station density (i.e., works well with high station density). It is also found to underestimate small precipitation rates and estimate rates higher than (1 mm/day) better (Rauthe et al., 2013). This effect might affect the applicability of bias correction methods in correcting lower rainfall rates, and limit SDM method when reducing the drizzle effect of climate simulations.

5.8.2 Climate simulations

The selected EURO-CORDEX simulations are CMIP5, simulations were used to generate the fifth assessment IPCC report. However, there are new CMIP branch, which is CMIP6, that has a better representation of the climate and incorporate socio-economic paths that resulted in more realistic scenarios. CMIP6 is better than CMIP5, and there is no doubt that RCP8.5 scenario of CMIP6 is more realistic and suitable for assessment studies.

The type of analysis presented in this thesis is based on one scenario, which is RCP8.5. The other two scenarios (RCP2.6 and RCP4.5) have not been considered here. It would be worthwhile to consider applying these scenarios in such assessment studies to assess climate change impacts based on all scenarios. In addition to test the applicability of QM and SDM in these scenarios as well.

Two GCMs are selected in this study (ICHEC-EC-EARTH and MOHC- HadGEM2-ES). These subsets were chosen based on the work of (Beusen, B., 2021), as these two models demonstrated the best performance in modelling the Rhine without any correction (out of nine GCMs). There are other GCMs that have not been tested in modelling the Rhine basin, and could have been considered. The expected difference can be seen when applying performance-based weighting, as there will be more distributed weights and its possible that the observed discharge lie within its envelope.

5.8.3 Hydrological modelling

In this study only one model (wflow_sbm) has been considered because of time limitations, however it would be of a great value to model the Rhine basin with more than one model to examine the projected changes in the future using a different model representation. Wflow_sbm is a conceptual model where most of its parameters are based on spatial information, with an aim of little to no calibration. Also, its kinematic wave routing scheme is not the best when modelling flat regions like the Netherlands.

5.8.4 Sensitivity of the hydrological model

When working with different datasets, it would have been wise to examine the hydrological model sensitivity to the errors in the observational dataset. Through this step it would be clearer to identify the pitfalls in the observational datasets, and which one of the meteorological variables is sensitive to the applied bias correction methods.

5.8.5 Land use changes

In any plausible climate assessment study, changes in land use affect the hydrological behavior of river catchments. For example, urbanization and deforestation can lead to an increase in the runoff of the catchment. This due to having lower storage capacity in the canopy or in the unsaturated zone, which can lead to rapid catchment response in the case of extreme rainfall events.

Many researchers have studied the effect of land use changes in the Rhine catchment. Hurkmans et al., 2009 used the VIC model and EURURALIS scenarios to analyse the effects of land use changes. They found out that such changes can lead to an increase in streamflow generation between 2 – 30 (%) in some parts of the Rhine catchment. Bronstert et al., 2007 used a combination of hydrological models and different land use scenarios to examine such changes and reported an increase between 4 – 10 (%) in parts of the Rhine basin. Similar results can be found in Hundecha and Bárdossy, 2004; Pfister et al., 2004.

However, in this research possible changes in land use are not considered when projecting to the future.

5.8.6 Bias correction

The applied bias correction methods in this project are some of the most advanced methods developed worldwide. These methods consider the spatial variability by applying the correction on a grid cell, this approach can be regarded as suitable for correcting temperature and evaporation, but not for correcting precipitation. Precipitation remains the largest source of uncertainty in climate model simulations. Linking temperature and precipitation distributions is a reasonable solution for this issue.

Cannon, 2018 provided a nice improvement to QM by linking the multivariate distributions. Which has resulted in better performance. This linkage is necessary in correcting areas that experience different types of rainfall that can be triggered based on a certain change in temperature, like in the Netherlands.

5.8.7 Lobith

The Rhine is the basin of interest and Lobith as one of the downstream locations at the Dutch border is the main point of interest. Focusing only at Lobith is not truly justifiable as biases may differ throughout the Rhine basin. Although five additional locations are analysed for validating some results (Kaub, Rees, Cochem, Rockenau and Andelfingen). These locations are not analysed in much detail as Lobith. Incorporating more locations in such analysis is more reliable than using Lobith only for investigating climate change impacts in the Rhine basin.

5.8.8 Climate change adaptation strategies

To deal with climate change impacts, many countries drafted their adaptation strategies to cater for this challenge. The Rhine is no exception, considering its importance to the riparian countries. In any climate assessment study, there is a strong need to consider the current adaptation strategies in the study area. Then, examine how these plans can hold by the end of the projection period and check if further refinement is needed. This is also important in raising awareness, exploring adaptation pathways, and assessing the associated implication on the decision-making process. This is not considered in this research. However, the interested reader in knowing the relevant climate adaptation strategies in the Netherlands and the Rhine can refer to the work of Haasnoot et al., 2015; Renner and Meijerink, 2018.

5.8.9 Water levels

Rating curves are used to estimate the projected changes in water levels till 2060. This is not the best option to reflect on the changes on water levels and water depths. To account for the projected changes in discharges

at Lobith, this requires hydrodynamical modelling of the reach near Lobith. This requires accurate bed levels data and additional measurements and was discarded because it was outside the scope of this project.

5.9 Future Research

- Examine the effects of land use changes and climate change in the Rhine using bias corrected projections. This can be done by extending these corrected projections while considering land use changes, using the LUISA modelling platform or EURURALIS (Hurkmans et al., 2009; Sperna Weiland et al., 2021).
- Apply bias correction to precipitation projections using the method of Cannon, 2018, or adjusting the proposed methodologies to account for multivariate statistical distributions.
- Assessing the associated cost of climate change effects in the Netherlands, with a special focus the vital commercial sectors that rely on the stability of the Rhine river (drinking water, navigation, etc.).
- Assessing other (or modifying the proposed) performance-based weighting methods on the Rhine river and make it work on smaller scales (like Lobith). This can be done by incorporating the spatial variability into these methods (similar as ST-REA in appendix X, where three additional locations are considered to compute the weight of each RCM).
- Modelling water availability challenges of the Rhine river due to climate change using CMIP6 simulations, to formulate adaptive water management strategies that increase the preparedness of the vulnerable sectors (drinking water sector, agriculture, and Navigation).
- Assessing the impacts of climate change at Lobith using more than one model. Where a combination of two or three models are used to assess climate change impacts in the Rhine river using bias-corrected projections.
- The projected changes in water levels at Lobith due to climate change using hydraulic and hydrodynamic models.
- Comparing a set of performance-based weighting methods to assess the hydrological response of the Rhine catchment to climate change. In addition to the REA and ClimWIP methods, possible alternatives can be the Bayesian inference method (Wu et al., 2014) and kriging climate change method (Drignei, 2009).

6| Conclusion

The aim of this study is twofold, first to investigate the added value of using two bias correction methods (QM and SDM) in modelling the hydrology of the Rhine river and the second is to assess the impacts of climate change on low flows at Lobith by 2060 using bias-corrected EURO-CORDEX simulations. The investigation is based on the performance of QM and SDM in the historical experiment. The assessment is based on the results generated from the calibrated `wflow_sbm`, after forcing it with the bias-corrected simulations for the period (2020 – 2060). To answer the main research question presented in section 1.4, the following sub-questions are addressed.

- How do bias (un)corrected simulations perform compared to the actual flows measured at Lobith when using `wflow_sbm`?

The raw simulations for the period (1999 – 2018) are corrected using QM and SDM based on the observational dataset during the period (1979 – 1998) and the raw simulations for the same time span. The result of this step is bias-corrected simulations for the period (1999 – 2018). Then, the calibrated `wflow_sbm` model is used to generate all the relevant discharges using raw and corrected simulations for the period (1999 – 2018). Then, these modelled flows from are compared to the actual observations for the same period.

The raw simulations have resulted in daily discharges that are in average 400 m³/s higher than the actual discharges for the same period at Lobith. However, the bias-corrected simulations provided daily modelled discharges that have less errors compared to the actual flows in this period.

Furthermore, the calculated long-term statistics from the corrected discharges contain less errors than the raw discharges relevant to the actual discharges. The improvements of the bias correction are not only constrained at Lobith, but also extended to two other locations, Rees and Cochem. Only one location (Kaub) that indicated lower performance after bias correction. This is linked to `wflow_sbm` being less skillful at that location.

In general, bias-corrected simulations have outperformed uncorrected simulations based on the actual discharges at Lobith. It appears that QM and SDM improved the ability of the climate simulations, but the underlying error in the parent GCM is not reduced using these methods. The improvements in hydrological modelling are also subjective to the reliability of the raw RCM and to the skill (accuracy) of the hydrological model at the specific location.

- How do the incorporated bias correction methods account for non-stationarity and preserve the climate signal of the RCMs at Lobith?

The differential split sample test (DSST) is used to test the reliability of QM and SDM in correcting precipitation and temperature under changing climate conditions. The two methods (QM and SDM) indicated better performance than the raw simulations in correcting relatively high projections based on lower values in the reference run. However, for the reversed case there is a mismatch between the two methods. QM can correct lower temperature simulations based on warmer values better than SDM. However, the two method are not able to correct lower precipitation values compared to higher values, indicating a poor performance in reducing the common drizzle effect in climate simulations.

Overall, QM is better than SDM in performing under different climate conditions. The scaling factors in the SDM prevent it from correcting underestimated simulations, thus limiting its applicability.

- How do bias-corrected simulations perform compared to uncorrected simulations when applying the performance-based weighting techniques described in Sperna Weiland et al., 2021 using the chosen EURO-CORDEX ensembles?

The two weighting methods (REA and ClimWIP) listed in Sperna Weiland et al., 2021 are used to assign the weights to the selected RCMs from the EURO-CORDEX initiative. Although the two weighting have managed to constrain the change signal in the raw ensembles. However, that did not improve the performance of the raw simulations as much as the bias correction. This is because the wet bias is high, which has prevented the two weighting methods from providing reliable predictions at Lobith. The bias-corrected simulations provided more accurate discharges than the weighted raw simulations of the two methods.

- What will the impact of climate change be on low flows in the Rhine river when projecting to 2060 (mid-future) after applying bias correction?

The average discharges for the period (2020 – 2060) are projected to be lower than the average observed discharge in the reference period. The annual long-term mean discharge at Lobith is projected to decrease by at least 100 m³/s. The average lengths below threshold (1100 m³/s) from the corrected simulations are higher than what is observed from the historical period at Lobith (1979 - 2018). Additionally, the projected average minimum 7-day flow is at least 100 m³/s lower than in the historical discharges.

Nine corrected simulations indicated unprecedented low flow events at Lobith (< 700 m³/s), at least 50 times in the next 38 years. Extreme drought events are going to occur more frequently and last longer than what is observed in the historical discharges. There is a slight change in the seasonality of the Rhine river during the summer and winter periods in the future, with lower summer flows and higher winter flows. There is a consensus between the corrected simulations on having a relative decrease in the daily average in the last four months of the year counteracted by higher flows from early June to the end of August. This increase is due to changes in snow melt contribution from the Alps region.

- What are the water levels corresponding to the bias-corrected future flows at Lobith?

The projected water levels are in line with the same changes in future flows. Higher average water levels are projected between the first of June to the end of August. However, from the end of August until mid-January of the next year the average water levels are projected to decrease.

The changes in the projected average water levels is not as extreme as the changes in discharges. An average decline of 20 cm is projected from the end of August until mid-January, while an average increase of 30 cm in water levels is expected between the first of June to mid-August.

The results obtained by combining bias correction with RCP8.5 scenario are alarming. The bias-corrected simulations indicate noticeable climate change impacts at Lobith and over the Rhine basin. This calls for further investigation, by incorporating the feedback in the climate system, analysing the other climate scenarios (RCP2.6 and RCP4.5), current climate adaptation strategies in the Rhine basin with bias-corrected simulations to assess these changes in a comprehensive manner.

References

- Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., ... & Schmidt, G. A. (2019). ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing. *Earth System Dynamics*, 10(1), 91-105.
- Acharya, N., Chattopadhyay, S., Mohanty, U.C., Dash, S.K. and Sahoo, L.N., 2013. On the bias correction of general circulation model output for Indian summer monsoon. *Meteorological Applications*, 20(3), pp.349-356.
- Annan, J. D., & Hargreaves, J. C. (2017). On the meaning of independence in climate science. *Earth System Dynamics*, 8(1), 211-224.
- Babar, B., Graverson, R., & Boström, T. (2019). Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5. *Solar Energy*, 182, 397-411.
- Bellprat, O., Kotlarski, S., Lüthi, D. and Schär, C., 2013. Physical constraints for temperature biases in climate models. *Geophysical Research Letters*, 40(15), pp.4042-4047.
- Beusen, B. L. D. (2021). Plastic transport and the effect of climate change in the Rhine (Master's thesis, Wageningen University and Research).
- Bronstert, A., Bárdossy, A., Bismuth, C., Buiteveld, H., Disse, M., Engel, H., ... & Ritter, N. (2007). Multi-scale modelling of land-use change and river training effects on floods in the Rhine basin. *River Research and Applications*, 23(10), 1102-1125.
- Brunner, L., Lorenz, R., Zumwald, M., & Knutti, R. (2019). Quantifying uncertainty in European climate projections using combined performance-independence weighting. *Environmental Research Letters*, 14(12), 124010.
- Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, 11(4), 995-1012.
- Buishand, T., Velds, C., 1980. *Klimaat van Nederland 1: Neerslag en verdamping*. Royal Netherlands Meteorological Institute (KNMI).
- Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate dynamics*, 50(1), 31-49.
- Cannon, A.J., Sobie, S.R. and Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes?. *Journal of Climate*, 28(17), pp.6938-6959.
- CCNR. (2020). *Waterway in the Rhine. 2020 assessment report*.
- Chen, D., Dai, A., & Hall, A. (2021). The Convective-To-Total Precipitation Ratio and the "Drizzling" Bias in Climate Models. *Journal of Geophysical Research: Atmospheres*, 126(16), e2020JD034198.
- Chen, J., Brissette, F. P., Lucas-Picher, P., & Caya, D. (2017). Impacts of weighting climate models for hydro-meteorological climate change studies. *Journal of Hydrology*, 549, 534-546.

- Chen, J., Brissette, F.P. and Caya, D., 2020. Remaining error sources in bias-corrected climate model outputs. *Climatic Change*, 162(2), pp.563-582.
- Chen, J., Brissette, F.P. and Lucas-Picher, P., 2015. Assessing the limits of bias-correcting climate model outputs for climate change impact studies. *Journal of Geophysical Research: Atmospheres*, 120(3), pp.1123-1136.
- Christensen, J. H., Kjellström, E., Giorgi, F., Lenderink, G., & Rummukainen, M. (2010). Weight assignment in regional climate models. *Climate Research*, 44(2-3), 179-194.
- Christodoulou, A., Christidis, P., & Bisselink, B. (2020). Forecasting the impacts of climate change on inland waterways. *Transportation Research Part D: Transport and Environment*, 82, 102159.
- Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., & Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17), 9391-9409.
- Deltares (2021). Developments 2021 Wflow Julia Meuse and Rhine basins. Technical report.
- Deltares (2022). QL1901-2019. Measurement report.
- Drignei, D. (2009). A kriging approach to the analysis of climate model experiments. *Journal of agricultural, biological, and environmental statistics*, 14(1), 99-114.
- Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K. and Liebert, J., 2012. HESS Opinions" Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, 16(9), pp.3391-3404.
- Enayati, M., Bozorg-Haddad, O., Bazrafshan, J., Hejabi, S. and Chu, X., 2021. Bias correction capabilities of quantile mapping methods for rainfall and temperature variables. *Journal of Water and Climate Change*, 12(2), pp.401-419.
- Fathian, F., Dehghan, Z., Bazrkar, M. H., & Eslamian, S. (2016). Trends in hydrological and climatic variables affected by four variations of the Mann-Kendall approach in Urmia Lake basin, Iran. *Hydrological Sciences Journal*, 61(5), 892-904.
- Giorgi, F., & Mearns, L. O. (2002). Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging"(REA) method. *Journal of climate*, 15(10), 1141-1158.
- Giorgi, F., & Mearns, L. O. (2003). Probability of regional climate change based on the Reliability Ensemble Averaging (REA) method. *Geophysical research letters*, 30(12).
- Giorgi, F., Jones, C. and Asrar, G.R., 2009. Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin*, 58(3), p.175.
- Glahn, H. R., & Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology and Climatology*, 11(8), 1203-1211.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, 377(1-2), 80-91.

- Haasnoot, M., Schellekens, J., Beersma, J. J., Middelkoop, H., & Kwadijk, J. C. J. (2015). Transient scenarios for robust climate change adaptation illustrated for water management in The Netherlands. *Environmental Research Letters*, 10(10), 105008.
- Hamed, K. H. (2008). Trend detection in hydrologic data: the Mann–Kendall trend test under the scaling hypothesis. *Journal of hydrology*, 349(3-4), 350-363.
- Harmel, R. D., Richardson, C. W., Hanson, C. L., & Johnson, G. L. (2002). Evaluating the adequacy of simulating maximum and minimum daily air temperature with the normal distribution. *Journal of applied meteorology*, 41(7), 744-753.
- Hegnauer, M., Beersma, J. J., Van den Boogaard, H. F. P., Buishand, T. A., & Passchier, R. H. (2014). Generator of Rainfall and Discharge Extremes (GRADE) for the Rhine and Meuse basins. Final report of GRADE 2.0. Project Under Commission of the Ministry of Infrastructure and the Environment: Rijkswaterstaat Water, Transport and Environment (RWS-WVL). Deltares, Delft, The Netherlands.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., & Piontek, F. (2013). A trend-preserving bias correction—the ISI-MIP approach. *Earth System Dynamics*, 4(2), 219-236.
- Herschey, R. W., Rumney, G. R., Oliver, J. E., Mosley, M. P., & Holland, P. G. (1998). Chézy formula. *Encyclopedia of Hydrology and Water Resources*, 121-121.
- Hirabayashi, Y., Kanae, S., Emori, S., Oki, T., & Kimoto, M. (2008). Global projections of changing risks of floods and droughts in a changing climate. *Hydrological sciences journal*, 53(4), 754-772.
- Hofstra, N., Haylock, M., New, M., & Jones, P. D. (2009). Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *Journal of Geophysical Research: Atmospheres*, 114(D21).
- Hundecha, Y., & Bárdossy, A. (2004). Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of hydrology*, 292(1-4), 281-295.
- IPCC. (2021). IPCC Sixth Assessment report. (IPCC, 2021) (Tech. Rep.)
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M., Georgievski, G. and Georgopoulou, E., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional environmental change*, 14(2), pp.563-578.
- Jang, S., & Kavvas, M. L. (2015). Downscaling global climate simulations to regional scales: statistical downscaling versus dynamical downscaling. *Journal of Hydrologic Engineering*, 20(1), A4014006.
- Kaspar, F., Schulzweida, U., & Müller, R. (2010, September). Climate data operators” as a user-friendly processing tool for CM SAF’s satellite-derived climate monitoring products. In Conference: EUMETSAT Meteorological Satellite Conference 2010 (pp. 20-24).
- Kim, M.K., Kim, S., Kim, J., Heo, J., Park, J.S., Kwon, W.T. and Suh, M.S., 2016. Statistical downscaling for daily precipitation in Korea using combined PRISM, RCM, and quantile mapping: Part 1, methodology and evaluation in historical simulation. *Asia-Pacific Journal of Atmospheric Sciences*, 52(2), pp.79-89.
- Klemeš, V., 1986. Operational testing of hydrological simulation models. *Hydrological sciences journal*, 31(1), pp.13-24.

- Klok, E. J., & Klein Tank, A. M. G. (2009). Updated and extended European dataset of daily climate observations. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 29(8), 1182-1191.
- Knutti, R. (2010). The end of model democracy?. *Climatic change*, 102(3), 395-404.
- Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*, 44(4), 1909-1918.
- Köhler, L., Mulligan, M., Schellekens, J., Schmid, S., & Tobón, C. (2006). Final Technical Report DFID-FRP Project no. R7991 Hydrological impacts of converting tropical montane cloud forest to pasture, with initial reference to northern Costa Rica.
- Laaha, G., & Blöschl, G. (2006). Seasonality indices for regionalizing low flows. *Hydrological Processes: An International Journal*, 20(18), 3851-3878.
- Lee, M. H., Lu, M., Im, E. S., & Bae, D. H. (2019). Added value of dynamical downscaling for hydrological projections in the Chungju Basin, Korea. *International Journal of Climatology*, 39(1), 516-531.
- Li, X. and Babovic, V., 2019. Multi-site multivariate downscaling of global climate model outputs: an integrated framework combining quantile mapping, stochastic weather generator and Empirical Copula approaches. *Climate Dynamics*, 52(9), pp.5775-5799.
- Livada, I., & Assimakopoulos, V. D. (2007). Spatial and temporal analysis of drought in Greece using the Standardized Precipitation Index (SPI). *Theoretical and applied climatology*, 89(3), 143-153.
- Lobanova, A., Liersch, S., Nunes, J. P., Didovets, I., Stagl, J., Huang, S., ... & Krysanova, V. (2018). Hydrological impacts of moderate and high-end climate change across European river basins. *Journal of Hydrology: Regional Studies*, 18, 15-30.
- Makkink, G. F. (1957). Ekzameno de la formulo de Penman. *Neth. J. Agric. Sci*, 5, 290-305.
- Maraun, D. (2016). Bias correcting climate change simulations-a critical review. *Current Climate Change Reports*, 2(4), 211-220.
- Maraun, D., 2012. Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophysical Research Letters*, 39(6).
- Matiu, M., Petitta, M., Notarnicola, C. and Zebisch, M., 2019. Evaluating snow in EURO-CORDEX regional climate models with observations for the European Alps: biases and their relationship to orography, temperature, and precipitation mismatches. *Atmosphere*, 11(1), p.46.
- McKee, T.B., Doesken, N.J. and Kleist, J. (1993), The relationship of drought frequency and duration to time scales. 8th Conference on Applied Climatology, 17-22 January, Anaheim, CA, pp 179-184.
- McMahon, T. A., & Peel, M. C. (2019). Uncertainty in stage-discharge rating curves: application to Australian Hydrologic Reference Stations data. *Hydrological Sciences Journal*, 64(3), 255-275.
- Pfeifer, S., Bülow, K., Gobiet, A., Hänsler, A., Mudelsee, M., Otto, J., Rechid, D., Teichmann, C. and Jacob, D., 2015. Robustness of ensemble climate projections analyzed with climate signal maps: seasonal and extreme precipitation for Germany. *Atmosphere*, 6(5), pp.677-698.

- Pfister, L., Kwadijk, J., Musy, A., Bronstert, A., & Hoffmann, L. (2004). Climate change, land use change and runoff prediction in the Rhine–Meuse basins. *River research and applications*, 20(3), 229-241.
- Pierce, D.W., Cayan, D.R., Maurer, E.P., Abatzoglou, J.T. and Hegewisch, K.C., 2015. Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, 16(6), pp.2421-2442.
- Prins, J., 2017. The inland navigation analysis system BIVAS: Analysis, validation, and recommendations for improvement. MSc thesis.
- Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. and Gratzki, A., 2013. A Central European precipitation climatology—Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, 22(3), pp.235-256.
- Reiter, P., Gutjahr, O., Schefczyk, L., Heinemann, G. and Casper, M., 2018. Does applying quantile mapping to subsamples improve the bias correction of daily precipitation?. *International Journal of Climatology*, 38(4), pp.1623-1633.
- Renner, T., & Meijerink, S. (2018). Policy entrepreneurs in international river basins—getting climate adaptation on the cross-border water policy agenda. *Regional environmental change*, 18(5), 1287-1298.
- Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N. and Rafaj, P., 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic change*, 109(1), pp.33-57.
- Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., ... & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2. 0—technical overview. *Geoscientific Model Development*, 13(3), 1179-1199.
- Rijkswaterstaat (2016). Lower Rhine weir measurement. Measurement report.
- Schuurmans, H., 2009. Penman-monteith referentieverdamping: Inventarisatie beschikbaarheid en mogelijkheden tot regionalisatie. (in dutch). Tech. rep., FutureWater report 86.
- Shrestha, M., Acharya, S.C. and Shrestha, P.K., 2017. Bias correction of climate models for hydrological modelling—are simple methods still useful?. *Meteorological Applications*, 24(3), pp.531-539.
- Sjerps, R. M., Ter Laak, T. L., & Zwolsman, G. J. (2017). Projected impact of climate change and chemical emissions on the water quality of the European rivers Rhine and Meuse: A drinking water perspective. *Science of the Total Environment*, 601, 1682-1694.
- Smith, P., Beaumont, L., Bernacchi, C.J., Byrne, M., Cheung, W., Conant, R.T., Cotrufo, F., Feng, X., Janssens, I., Jones, H. and Kirschbaum, M.U., 2022. Essential outcomes for COP26. *Global change biology*, 28(1), pp.1-3.
- Snippen, E., Mens, M., Huiink, J., ter Mat, J., 2016: Basisprognoses Zoetwater, Deltares.
- Sørland, S.L., Schär, C., Lüthi, D. and Kjellström, E., 2018. Bias patterns and climate change signals in GCM-RCM model chains. *Environmental Research Letters*, 13(7), p.074017.

- Sperna Weiland, F. C., Visser, R. D., Greve, P., Bisselink, B., Brunner, L., & Weerts, A. H. (2021). Estimating Regionalized Hydrological Impacts of Climate Change Over Europe by Performance-Based Weighting of CORDEX Projections. *Frontiers in Water*, 3, 713537.
- Sperna Weiland, F., Hegnauer, M., Bouaziz, L., & Beersma, J. (2015). Implications of the KNMI'14 climate scenarios for the discharge of the Rhine and Meuse. Comparison with Earlier Scenario Studies.
- Suh, M. S., Oh, S. G., Lee, D. K., Cha, D. H., Choi, S. J., Jin, C. S., & Hong, S. Y. (2012). Development of new ensemble methods based on the performance skills of regional climate models over South Korea. *Journal of Climate*, 25(20), 7067-7082.
- Sutton, R. T., & Hawkins, E. (2020). ESD Ideas: Global climate response scenarios for IPCC AR6. *Earth Syst. Dyn. Discuss*, 2020, 1-4.
- Switaneck, M. B., Troch, P. A., Castro, C. L., Leuprecht, A., Chang, H. I., Mukherjee, R., & Demaria, E. (2017). Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes. *Hydrology and Earth System Sciences*, 21(6), 2649-2666.
- Tegegne, G., Kim, Y. O., & Lee, J. K. (2019). Spatiotemporal reliability ensemble averaging of multimodel simulations. *Geophysical Research Letters*, 46(21), 12321-12330.
- Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., & Uijlenhoet, R. (2010). Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin. *Hydrology and earth system sciences*, 14(4), 687-703.
- Teutschbein, C. and Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of hydrology*, 456, pp.12-29.
- Teutschbein, C. and Seibert, J., 2013. Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?. *Hydrology and Earth System Sciences*, 17(12), pp.5061-5077.
- Themeßl, M.J., Gobiet, A. and Heinrich, G., 2012. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112(2), pp.449-468.
- Thompson, J. M. T., & Sieber, J. (2012). Climate predictions: the influence of nonlinearity and randomness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962), 1007-1011.
- Uehlinger, U. F., Wantzen, K. M., Leuven, R. S., & Arndt, H. (2009). The Rhine river basin.
- Van Osnabrugge, B., Uijlenhoet, R. and Weerts, A., 2019. Contribution of potential evaporation forecasts to 10-day streamflow forecast skill for the Rhine River. *Hydrology and Earth System Sciences*, 23(3), pp.1453-1467.
- Van Osnabrugge, B., Weerts, A.H. and Uijlenhoet, R., 2017. genRE: A method to extend gridded precipitation climatology data sets in near real-time for hydrological forecasting purposes. *Water Resources Research*, 53(11), pp.9284-9303.
- van Verseveld, W., Boisgontier, H., Bouaziz, L., Eilander, D., Haag, A., Hazenberg, P., Hegnauer, M., Imhoff, R., van Osnabrugge, B., Schellekens, J., Sperna Weiland, F., ten Velden, C., Visser, M., Wannasin, C., and Weerts, A.: Wflow_sbm, a spatially distributed hydrologic model: from global data to local applications,

EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-14837, <https://doi.org/10.5194/egusphere-egu2020-14837>, 2020.

- Van Vooren, S., Van Schaeybroeck, B., Nyssen, J., Van Ginderachter, M. and Termonia, P., 2019. Evaluation of CORDEX rainfall in northwest Ethiopia: sensitivity to the model representation of the orography. *International Journal of Climatology*, 39(5), pp.2569-2586.
- Vertessy, R. A., & Elsenbeer, H. (1999). Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization. *Water Resources Research*, 35(7), 2173-2187.
- Visser, R. D. (2020). Impact of climate change on local water resources of European catchments (Master's thesis, Wageningen University and Research).
- Warnatzsch, E. A., & Reay, D. S. (2019). Temperature and precipitation change in Malawi: evaluation of CORDEX-Africa climate simulations for climate change impact assessments and adaptation planning. *Science of the Total Environment*, 654, 378-392.
- Wieriks, K., & Schulte-Wülwer-Leidig, A. (1997, May). Integrated water management for the Rhine river basin, from pollution prevention to ecosystem improvement. In *Natural Resources Forum* (Vol. 21, No. 2, pp. 147-156). Oxford, UK: Blackwell Publishing Ltd.
- Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic change*, 62(1), 189-216.
- Wu, W., Clark, J. S., & Vose, J. M. (2014). Response of hydrology to climate change in the southern Appalachian Mountains using Bayesian inference. *Hydrological processes*, 28(4), 1616-1626.
- Zhao, T., Bennett, J.C., Wang, Q.J., Schepen, A., Wood, A.W., Robertson, D.E. and Ramos, M.H., 2017. How suitable is quantile mapping for postprocessing GCM precipitation forecasts?. *Journal of Climate*, 30(9), pp.3185-3196.

7| Appendix I

In this section, the Makkink equation (Makkink, 1957) used to compute the evaporation is discussed first, followed by the general procedure on which the rating curves at Lobith are calculated. Then, all the metrics used in evaluating the hydrological performance of the wflow_sbm (section 3.5), assessing the added value of bias correction, and estimate the impacts of the meteorology at Lobith and in the Rhine basin are outlined.

7.1 Makkink Evaporation:

EURO-CORDEX simulations do not simulate evaporation directly, raw evaporation simulations are calculated by using the relevant raw simulations. In this thesis, the Makkink formula (Makkink, 1957) is used to calculate evaporation simulations using three simulations: temperature [C°], pressure [hPa] and incoming shortwave radiation [W/m²].

There are many modified Makkink evaporations methods that were proposed, for example (Schuurmans, 2009; Buishand and Velds, 1980). The Makkink evaporation used to calculate evaporation for the used climate simulations for the period between [1979 – 2060] based on three inputs: temperature [C°], pressure [hPa] and incoming shortwave radiation [W/m²]. The Makkink evaporation is calculated based on the formula used by Deltares and described in the equations below (equation A1 to A5).

First calculate the saturated vapor pressure e_{sat} [Pa], based on temperature T [C°].

$$e_{sat} = 6.112 * e^{\frac{17.67 * T}{T + 243.5}} \quad (A1)$$

Second, calculate the slope of the vapor pressure curve S [-].

$$S = e_{sat} * \left(\frac{17.269}{T + 243.5} \right) * \left(1 - \left(\frac{T}{T + 243.5} \right) \right) \quad (A2)$$

Estimate the latent heat of vaporization λ .

$$\lambda = 2.502 * 10^6 - (2250 * T) \quad [J/Kg] \quad (A3)$$

Then, calculate the psychrometric constant by using pressure P [hPa], specific heat capacity of dry air C_p [J/Kg], and latent heat of vaporization λ .

$$\gamma = \frac{C_p * P}{0.622 * \lambda} \quad (A4)$$

Finally, the daily potential evaporation PET expressed in mm /day can be estimated using the following formula:

$$PET = \frac{0.65}{\lambda} * \frac{S}{S + \gamma} * K_{in} * 86400 \quad \left[\frac{mm}{day} \right] \quad (A5)$$

7.2 Rating Curves

Rating curves are used by hydrologists and hydraulic engineers to translate the measured water levels at a certain location to river discharges and vice versa. To be able to project the changes in water levels due to climate change, it's crucial to derive a realistic rating curve formula at Lobith. The water levels and discharge measurements that were commenced during 2016 are used to derive the rating curves at Lobith

(Rijkswaterstaat, 2016). There is a lot of uncertainty involved with using rating curves in the context of flooding or high flows, however since the focus of this research is on low flows. Applying a rating curve formula to compute water levels from the modelled discharges, one can argue that a lesser amount of errors are introduced (McMahon and Peel, 2019), compared to using rating curves in applications that require high flows (i.e., flood modelling).

Usually rating curves are used to translate the measured water levels into discharges. The water levels and discharge measurements are used to obtain the fitted rating curve, based on the following equation:

$$y = a x^p + b \quad (A6)$$

Where y is the water level at the point of interest, x is the discharge value at the point of interest, and the coefficients (a, p, and b) are site specific fitted parameters.

This equation was based on adjusting the normal rating curve derived from the Chezy formula described below (Herschy et al., 1998).

$$Q = c * b * (h - h_0)^{1.50} * i^{0.50} \quad (A7)$$

Where Q is the the discharge, C is the coefficient of discharge, b is the width of the section (weir in the case of available measurement), i is the slope of the river, h₀ is the zero discharge measurement and h is the water level

The water level (h) described in the Chezy formula (equation A7) can be adjusted as in equation A8.

$$h = \left(\frac{1}{c * b * i^{0.5}} \right)^{2/3} * Q^{2/3} + h_0 \quad (A8)$$

The structure of equation A8 is like that of equation A6, the values of the three coefficients (a, b and p) is obtained by applying a fitting procedure on the actual measurements. The fitting was performed using the solver option in excel, as the solution obtained through curve_fit feature in python has yielded a lower R² value.

The fitted rating curve formula for Lobith based on the available measurement (Rijkswaterstaat, 2016) is described in equation A9.

$$y = 0.151414 * x^{0.491511} + 2.85256 \quad (A9)$$

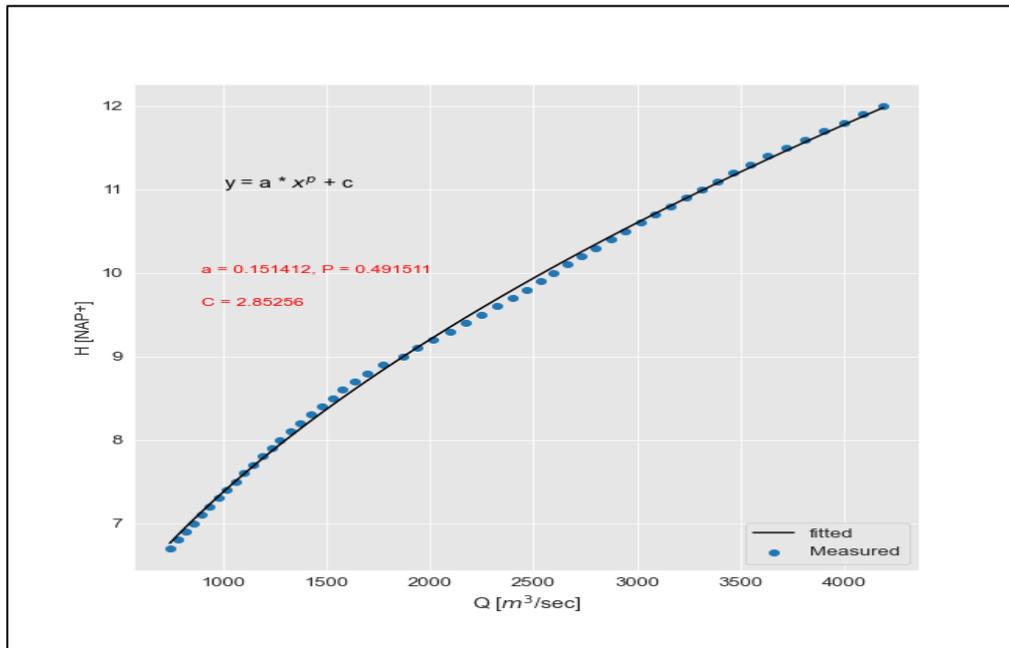


Figure 7-1: Fitted rating curve at Lobith based (Rijkswaterstaat, 2016) measurements

The fitted equation provided a correlation coefficient (R^2) of 0.98. The equation translates flows to water levels in a reasonable manner, until the threshold of 4500 m^3/s . The uncertainty of dealing with the rating curves is high when dealing with high flow. Since in this project is mainly focusing with low flows, using such rating curves is of lesser uncertainty, as the focus is on flows that fall within the range of (900 – 3500) m^3/s .

7.3 Assessment metrics

7.3.1 Long-term statistics

Statistics spanning for more than 20 years is usually used when analyzing climate change. Four metrics have been used throughout this project, three of them consider changes throughout the year, while the remaining one focuses in the changes during the summer season (specifically the period from June to August, known as JJA).

Long-term annual mean

After obtaining the modelled discharges by wflow_sbm, the long-term annual mean is calculated by averaging these daily simulated values throughout the year, then they are averaged again over the long run (≥ 20 years) to obtain the long-term annual mean.

long-term Summer mean

The same procedure used in obtaining the long-term annual mean have been applied, with a minor difference. The difference here is that only the values that fall between the 1st of June to the 31st of August in each year have been considered in calculating the long-term summer mean.

7-day minimum low flow (ND-7)

Moving window of 7-days is applied to the daily discharges, for each year the minimum value is calculated (this value describes the lowest discharges observed during a certain week). This metric is critical in many water-uses like navigation, ecology, and water supply.

Lengths below threshold

What is meant by lengths below threshold is the period (number of days in this case) during which the discharges are below a certain threshold. A threshold of [1100 m^3/sec] is chosen for Lobith, this value is almost 40 % lower than the suggested threshold by Snippen et al., 2016, where the Dutch water systems is

not be functioning properly. The reason for choosing this value is because it is the calculated 95th percentile of the recorded discharges within the period (1901 – 2018)

7.3.2 Meteorology

Three assessment criteria are implemented to assess the possibility of having meteorological droughts at Lobith or other locations. The metrics are standard precipitation index (SPI), the dry spell frequency and the regional Mann-Kendall trend test.

Standard Precipitation Index (SPI)

Standard Precipitation Index (SPI) is employed to track and assess meteorological droughts at a specific location, Lobith in this project. SPI calculates the possibility of having meteorological droughts using only precipitation and was proposed by (McKee et al.,1993). SPI is calculated by fitting a probability density function (gamma in this case) to the frequency distribution of precipitation summed through a certain time window (monthly, bi-monthly, yearly, etc.). The applied probability density function is the same function described in the SDM section.

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad \text{for } x > 0 \quad (\text{A10})$$

Where $\alpha > 0, \beta > 0$ are the shape and scale parameters, $g(x)$ is the fitted gamma distribution, and x is precipitation accumulated over a certain period (≥ 1 month).

The Table below describes the state of extremity based on the derived SPI values. One should note that for each configured time step, there is a corresponding SPI value.

Table 7-1: SPI Values description from Livada and Assimakopoulos, (2007).

Condition	SPI	Probability (%)
Extreme Drought	$SPI \leq -2$	2.30
Severe drought	$-2 \leq SPI \leq -1.50$	4.40
Moderate drought	$-1.50 \leq SPI \leq -1$	9.20
Near normal	$-1 \leq SPI \leq 1$	68.20
Moderately wet	$1 \leq SPI \leq 1.50$	9.20
Severely wet	$1 \leq SPI \leq 1.50$	4.40
Extremely wet	$1 \leq SPI \leq 1.50$	2.30

SPI values are derived using python, and the durations which are considered are three and six months.

Dry Spell analysis

Dry spell corresponds to the number of days with rainfall below 1 mm/day. This metric is essential for assessing the potential occurrence of long drought events like the one experienced in 2018 within the Netherlands. For that a threshold of 1 mm/day is assumed and all the values that are lesser than this threshold within the period of early May to the end of September in a single year are counted. The period is extended than the normal summer period, as the projected discharges indicated a small shift.

The probability of experiencing these spells within a year is calculated. Equation (A11) assumes a Poisson distribution; the exceedance probability of these dry spells is computed using the following equation.

$$P(c) = \exp(-rt) \frac{(rt)^c}{c!} \quad (A11)$$

Where $P(c)$ is the probability of having a number of events exceeding the dry spell duration, r is the rate of exceedance calculated by dividing the number of days exceeding the applied threshold (n_{ex}) in the set years within the number of years considered (y), c is the number of events in interval t and t is the period considered in this case (May to September). This analysis was possible through the Python code written by (Van Der Ent, 2020).

Trend Analysis

As a complementary step to investigate if the Rhine basin is going to be warmer or colder in the coming years under RCP8.5, a regional trend test on the three bias-corrected meteorological variables is applied. The regional trend test is performed by applying the Mann-Kendall (MK) trend test on each grid cell within the climate simulations used in this project. In general, the test is based on testing a certain hypothesis (Hamed, 2008), if the result of this trend is positive it means that there is a positive trend (increase). The contrary can be inferred when having a negative value, implying a decrease in the general trend. Finally, if the result of this test is zero it means that there is no general trend in the grid cell. The trend test was performed using `pymankendall` (package in python). The original MK test is applied by following the outlined steps in (Fathian et al., 2016).

In the case of temperature an increasing trend indicates a warming trend, while a negative temperature indicates a cooling trend. In the case of precipitation, a positive trend at a grid cell hints at an increase in precipitation rates at that grid cell and vice versa when obtaining a negative trend. Positive trend in evaporation indicates an increase in evaporation rates and a negative trend indicates a decrease in the rates in evaporation.

7.3.3 Hydrological model assessment

Three assessment metrics are considered in evaluating the used the performance of the hydrological model `wflow_sbm`. These metrics are chosen to assess the performance of `wflow_sbm` in providing accurate predictions in low and high flows.

Nash-Sutcliffe efficiency

Nash-Sutcliffe efficiency (NSE) is a widely used metric by the hydrologic community. NSE ranges between $-\infty$ to 1, a value of 1 indicates a perfect model. NSE value of 0 indicate that the model describes the observations as good as the mean of the observations. NSE is used to assess `wflow_sbm` performance in producing high flows compared to the observed flows.

$$NSE = 1 - \frac{\sum_{i=1}^N (Q_{m,i} - \bar{Q}_0)^2}{\sum_{i=1}^N (Q_{o,i} - \bar{Q}_0)^2} \quad (A12)$$

Where NSE is the Nash-Sutcliffe efficiency, $Q_{m,i}$ is the modelled flow at time step i , $Q_{o,i}$ is the observed flow at time step i and \bar{Q}_0 is the average of the observed discharge dataset.

Logarithmic Nash-Sutcliffe efficiency

Logarithmic Nash-Sutcliffe efficiency (NSE_{log}) is a modified metric from the NSE, that fluctuates in the same range. NSE_{log} is used to assess `wflow_sbm` performance in producing low flows.

$$NSE_{\log} = 1 - \frac{\sum_{i=1}^N (\log(Q_{m,i}) - \log(\bar{Q}_0))^2}{\sum_{i=1}^N (\log(Q_{o,i}) - \log(\bar{Q}_0))^2} \quad (A13)$$

Where NSE_{\log} is the logarithmic Nash-Sutcliffe efficiency, $\log(Q_{m,i})$ is the logarithmic value of the modelled flow at time step i , $\log(Q_{o,i})$ is the logarithmic value of the observed flow at time step i , and $\log(\bar{Q}_0)$ is the logarithmic value of the average of the observed discharge dataset.

Kling-Gupta efficiency

Kling-Gupta efficiency (KGE) was introduced by (Gupta et al., 2009), KGE incorporates three components (linear correlation, variation, and mean). KGE ranges between $-\infty$ to 1, a value of 1 indicates a perfect model.

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (A14)$$

$$= 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_m}{\sigma_o} - 1\right)^2 + \left(\frac{\mu_m}{\mu_o} - 1\right)^2}$$

Where KGE is the Kling-Gupta efficiency factor, r is the correlation coefficient between modelled discharges and observed flows, α is the ratio of the modelled to the observed standard deviations and β is the ratio of the mean of the modelled flows to the mean of observed discharges.

7.3.4 Other hydrological metrics

Flow duration Curve

Flow duration Curve (FDC) is used by hydrologists to reflect on the proportion of time steps on which a certain flow value is exceeded. FDC is the marginal probability distribution function of the analysed time series. FDC is derived by sorting the discharge data in a descending manner (n values), then a value of m is assigned starting with a value of 1 given to the highest value and increasing this value by 1 when moving downwards. Finally, the exceedance probability (P) is calculated based on the equation (A15).

$$P = \frac{m}{n+1} \quad (A15)$$

P is the probability of exceedance and n being the total count of flow values in the analyzed timeseries.

Logarithmic flow duration curves

The same procedure applied for calculating flow duration curves, the difference here is the logarithmic values of the flow are first calculated and then the similar procedure outlined in equation A15 is followed.

The 95th Percentile flow

The 95th percentile flow (denoted as Q_{95}) is a metric that represents low flows at the analyzed location. Q_{95} is the flow value at which 95 % of the time is exceeded. This value is used to characterize low flows at Lobith.

The 5th percentile flows Q_5

The 5th percentile flow (denoted as Q_5) a metric that represents the high flows at the analyzed location. Q_5 is the flow value at which 5 % of the time being exceeded. This value is used to characterize high flows at Lobith.

Seasonality ratio

To reflect on the changes in seasonality at Lobith considering a changing climate, the seasonality ratio metric is adopted. Seasonality ratio (SR) is calculated based on the flow duration curve (Laaha and Blöschl, 2006). SR is calculated by dividing the 95th percentile flow value observed during the summer period by the 95th percentile flow observed during winter period.

$$SR = \frac{Q_{95th, \text{ summer}}}{Q_{95th, \text{ winter}}} \quad (A16)$$

Having a SR value lesser than 1.0 indicates low flows during summer period (JJA) and high flows during the winter (DJF) which happens to be the case for the Rhine, and vice versa. SR will be used to reflect on the changes in seasonality when projecting to 2060. The Rhine is known for being a snow dominated regime, having high flows during the winter period and low flows during the summer period (Hurkmans et al., 2009). SR is used to reflect on the situation in the future and assess if it is going to remain the same after applying the bias correction methods.

7.4 General hydrological modelling framework

7.4.1 Bias correction

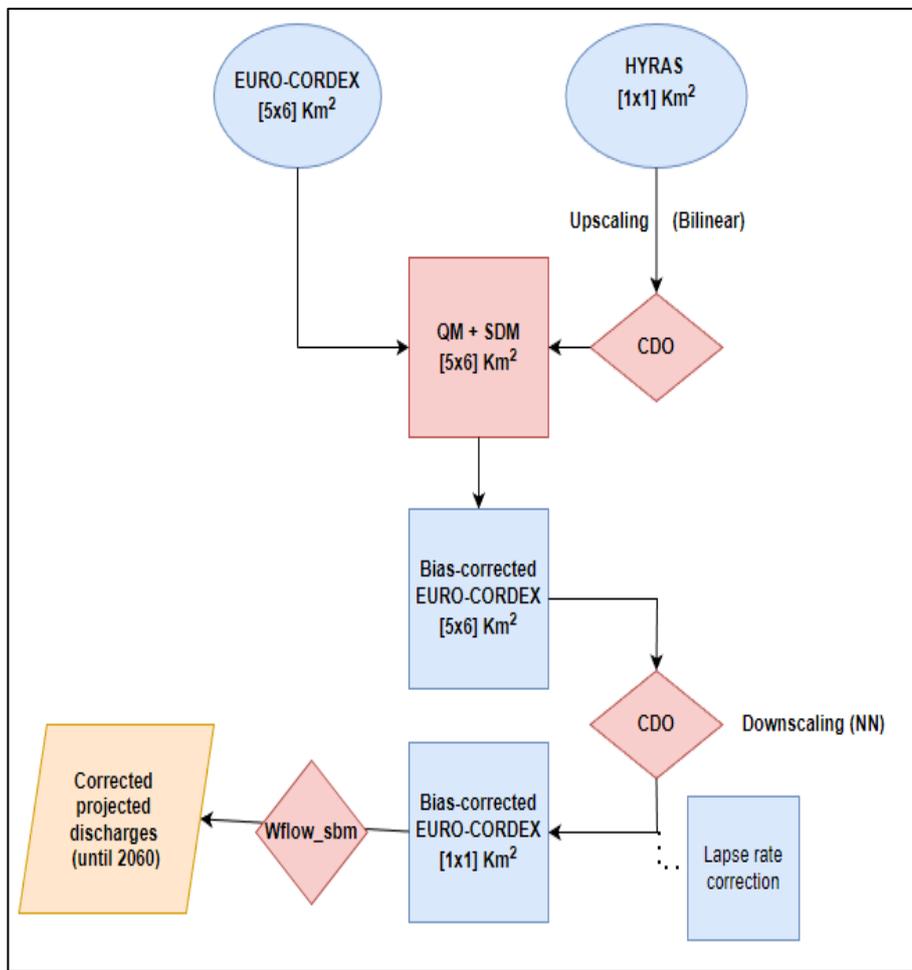


Figure (A1.2) demonstrates the systematic process for producing the bias-corrected discharges in the future. The mismatch in the grid cell resolution of the observational dataset and the reprojected EURO-CORDEX ensemble is solved by re-gridding the observational datasets. The resolution of the observational dataset is coarsened (upsampling, converting from fine resolution to a coarser resolution) using the bilinear upscaling method in CDO (Kaspar et al., 2010). Then, after matching the resolution the bias-correction methods are applied, the product is the bias-corrected meteorological variables with the same resolution as the inputs. After that, CDO is used to downscale the corrected datasets (from coarse resolution to a finer resolution) using the nearest neighbor method (Kaspar et al., 2010). An additional temperature correction is applied using the lapse rate correction equation (equation 3.20). Finally, the three meteorological variables

Figure 7-2: Hydrological modelling framework using QM and SDM

are merged into one file and used to force wflow_sbm to compute the daily (corrected) modelled for the period (2020 – 2060). It should be also noted that this structure is the same structure for the historical experiment. Temperature correction is applied to all simulations (raw or corrected) to ensure the model perform as close as possible to the terrain of the Rhine basin.

7.4.2 ClimWIP

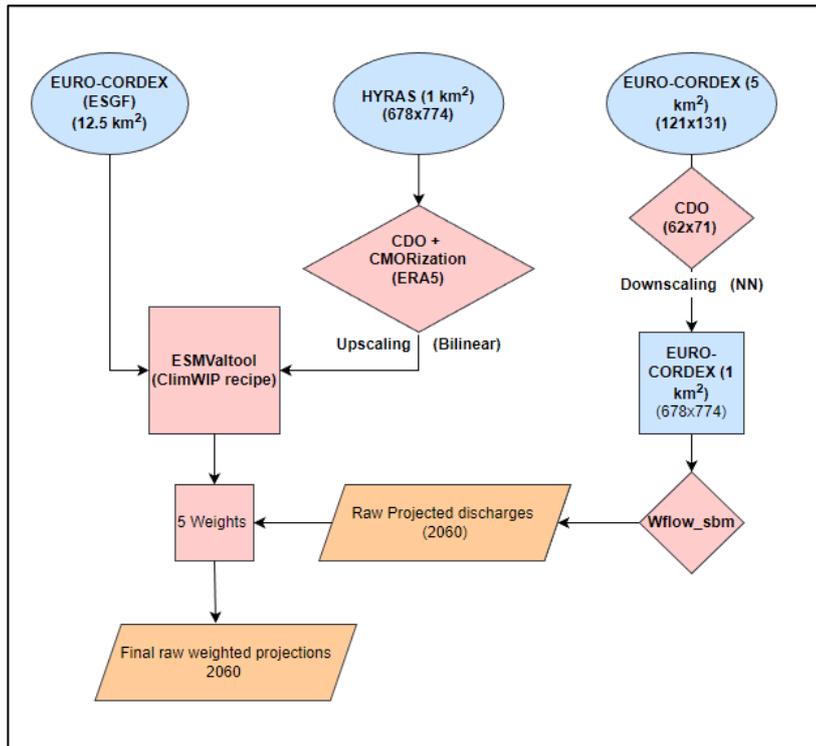


Figure 7-3: Modelling framework using ClimWIP

The ESMValTool is used to compute GCM weights based on ClimWIP (Brunner et al., 2019). However, since the observational datasets does not have the same resolution as the climate simulations, a statistical downscaling is performed using CDO before using these datasets in the softwares package. After obtaining the weights, the final discharge is the weighted average of these raw future simulations as in equation 3.18 (section 3.4.1). The results of this framework is shown in the results section

7.4.3 REA

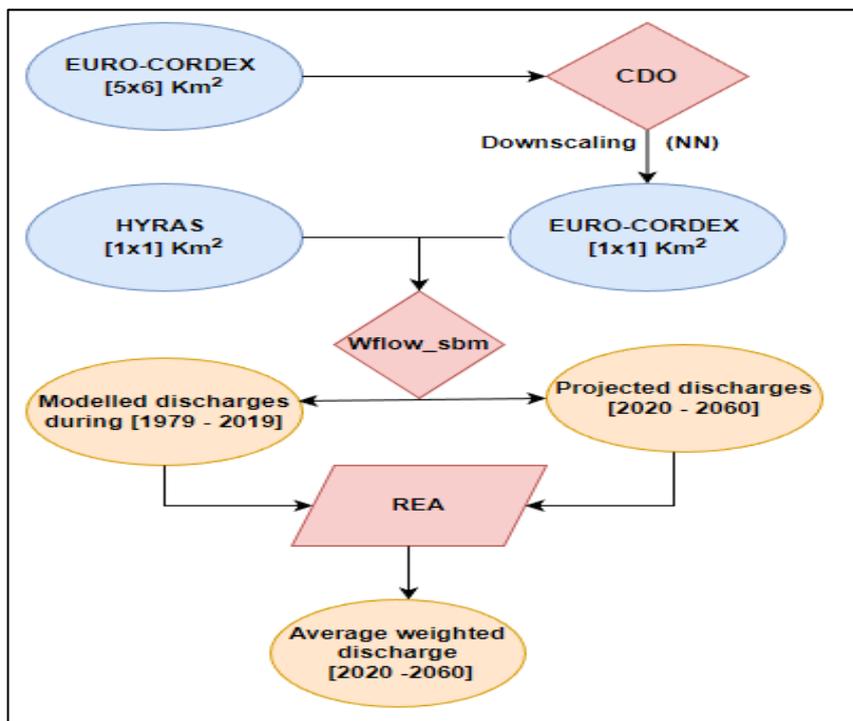


Figure 7-4: Hydrological modelling framework using REA

REA is implemented on the raw modelled discharges (section 3.4.2). Figure (A1.3) shows the applied procedure to calculate the weighted projections to 2060. The raw EURO-CORDEX simulations are remapped (downscaled) to be able to run wflow_sbm based on the 1 km² resolution. Then, the calculated discharges are used to obtain the weights. Then these weights are used to compute the weighted discharge of the ensemble simulations (equation 3.18).

8| Appendix II: Additional Results

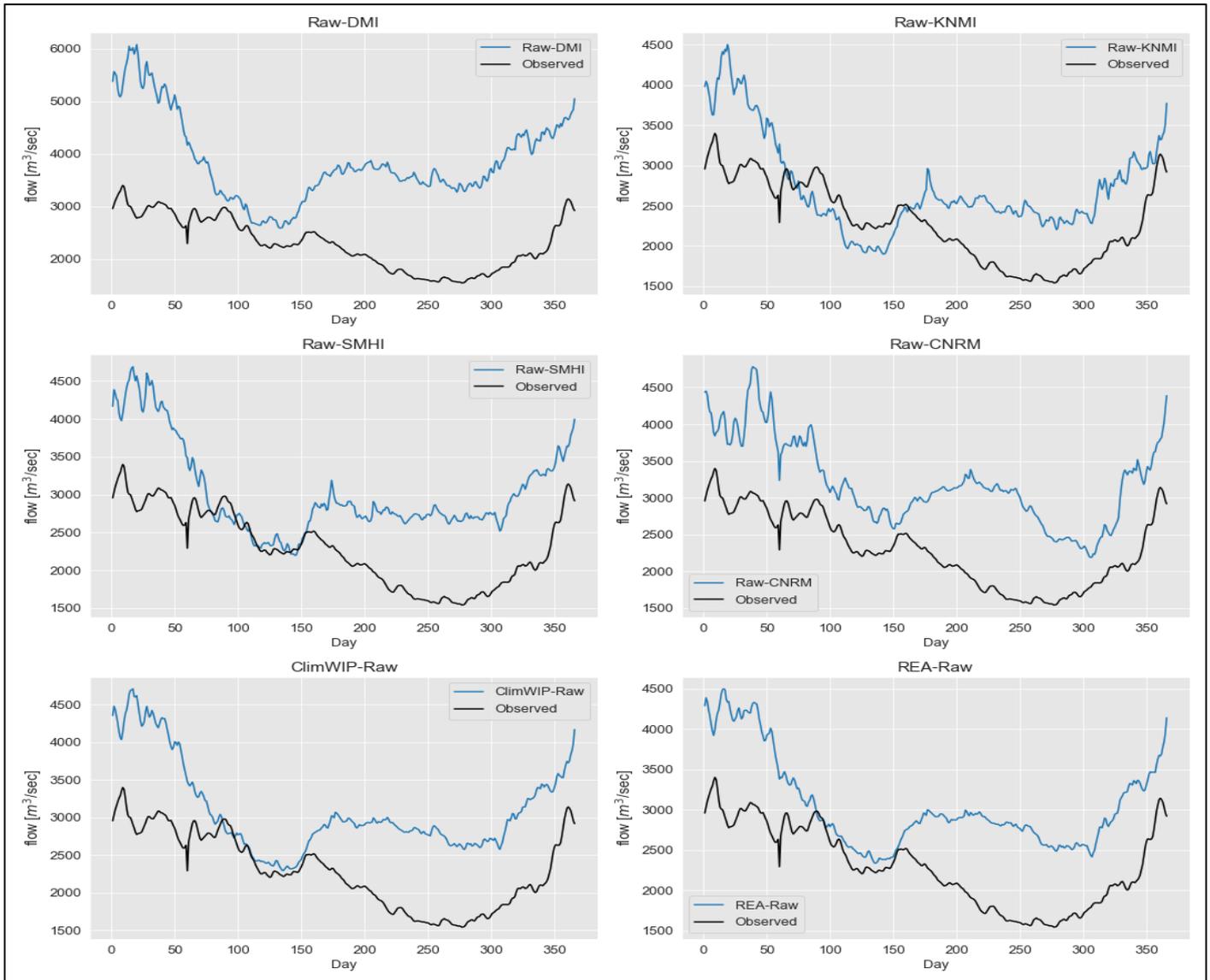


Figure 8-1: Discharges at Lobith obtained by forcing wflow_sbm by raw EURO-CORDEX simulations (1999 - 2018).

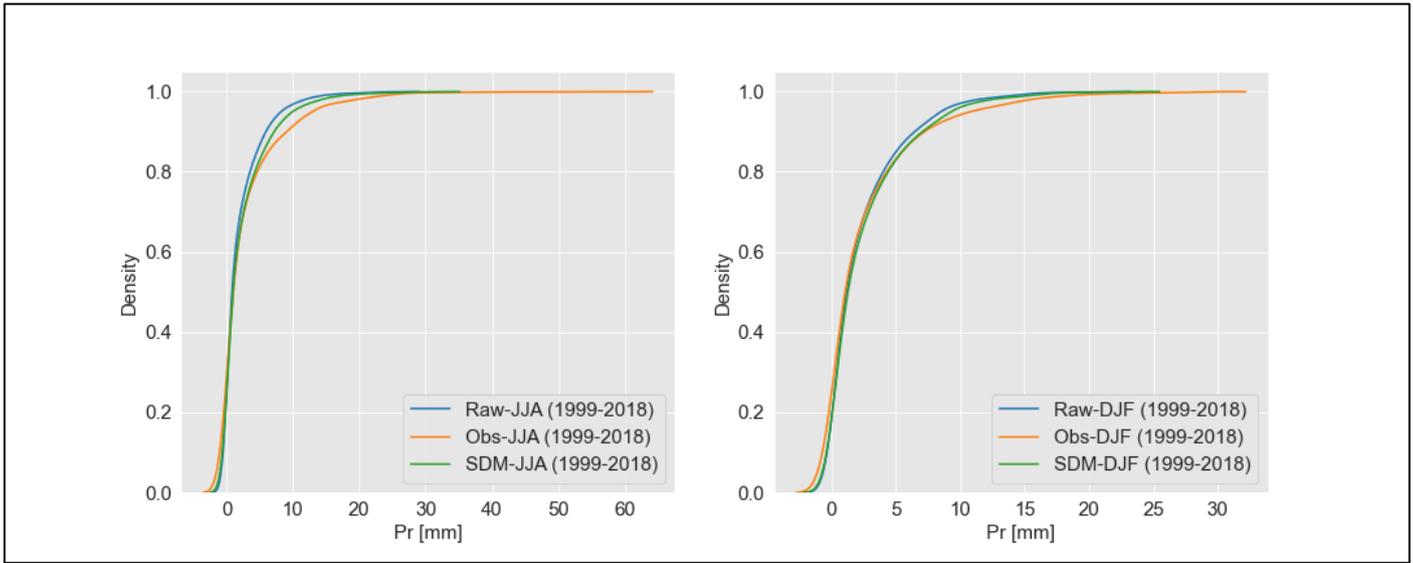


Figure 8-2: SDM performance in the DSST for precipitation.

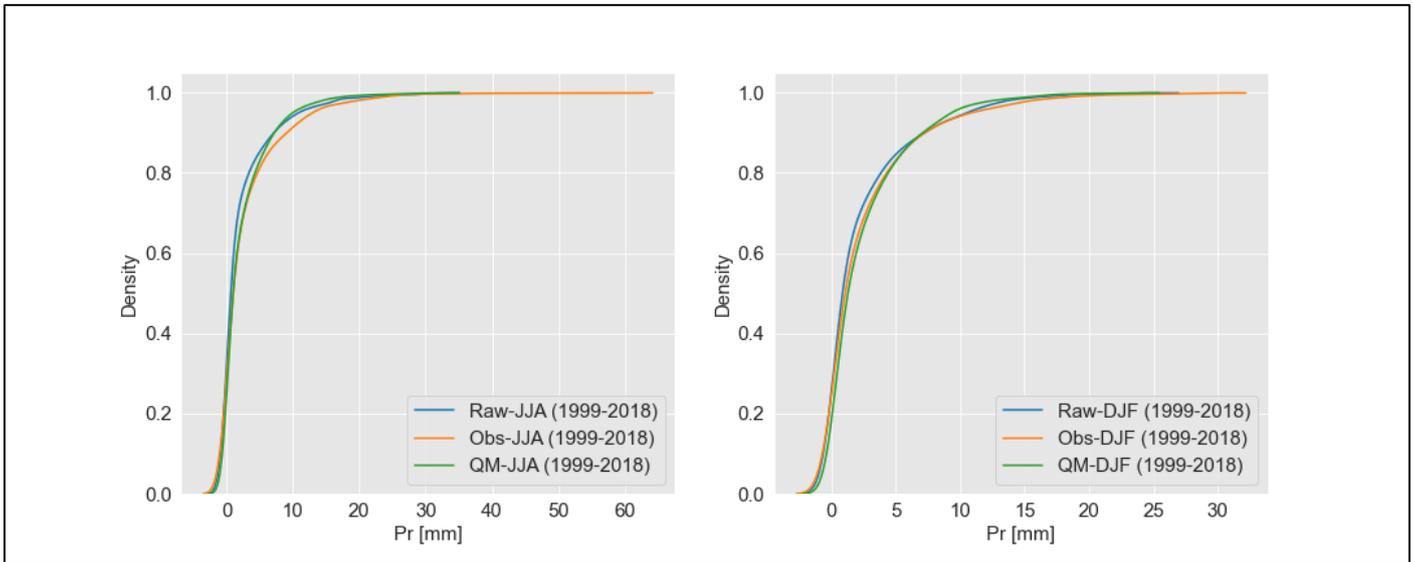


Figure 8-3: QM performance in the DSST for precipitation

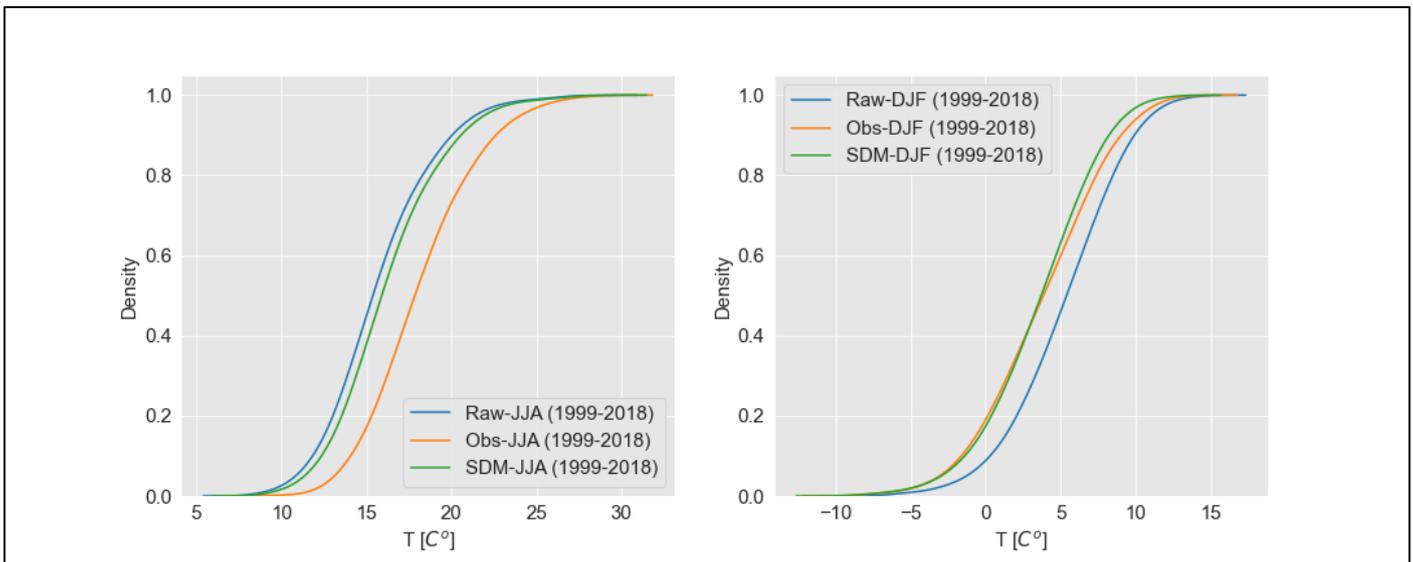


Figure 8-4: SDM performance in the DSST for temperature

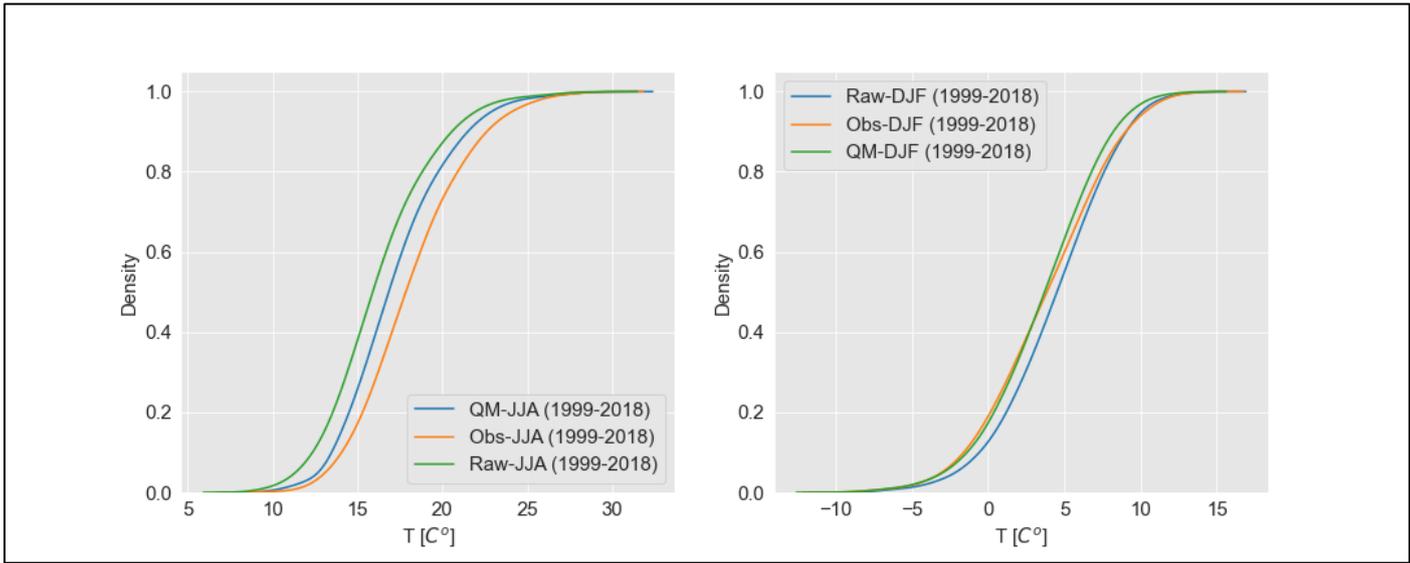


Figure 8-5: QM performance in the DSST for temperature.

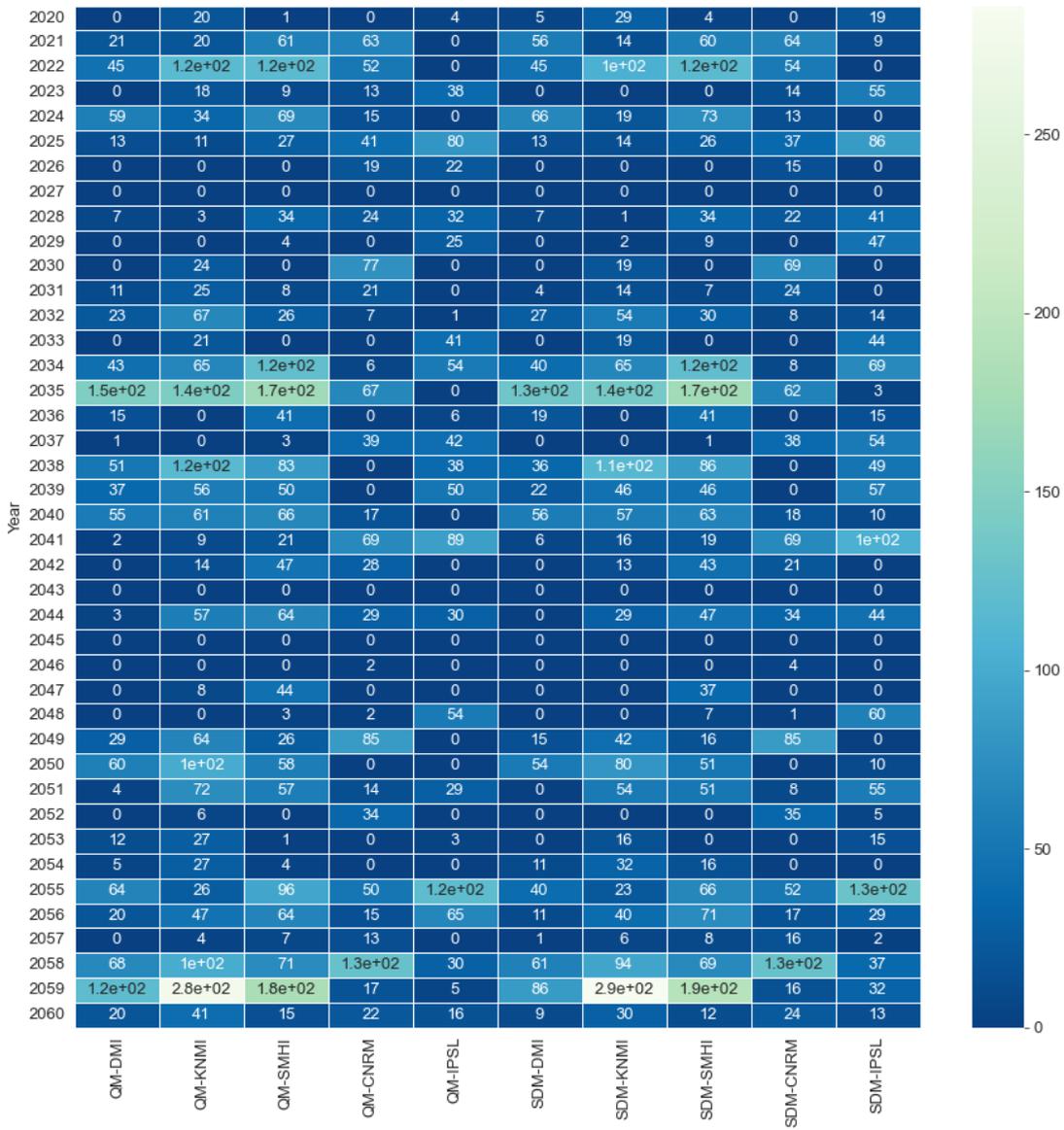


Figure 8-6: Lengths below threshold for the period (2020 - 2060) using bias-corrected EURO-CORDEX simulations

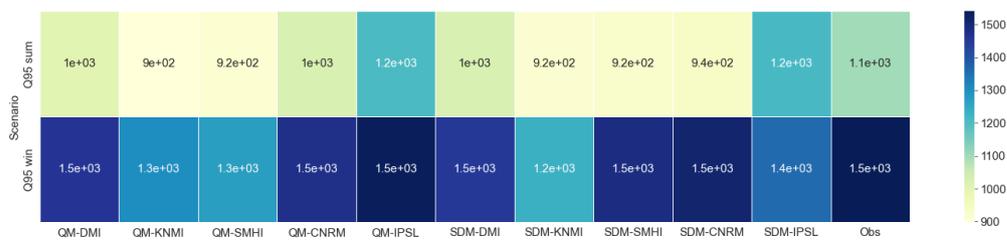


Figure 8-7: SR Q95 values for summer and winter values

8.1 Dry spells and SPI analysis

The bias-corrected simulations at Lobith are analysed to assess the occurrence of dry spells in the future relevant to the actual observations. The results shown in figure 8.8 below, suggest that the occurrence of dry spells between (5 – 10) days is going to be more frequent (six out of the ten corrected simulations). The CNRM and SMHI subsets indicate dry spell durations that are longer than what is observed during the reference period (1979 – 2019).

In the figures 8.9, 8.10, 8.11 and 8.12 the SPI analysis of the KNMI and SMHI subsets is illustrated. The reason for choosing these two subsets in the SPI analysis, is because they showed the highest skill relative to the actual observations during the historical experiment.

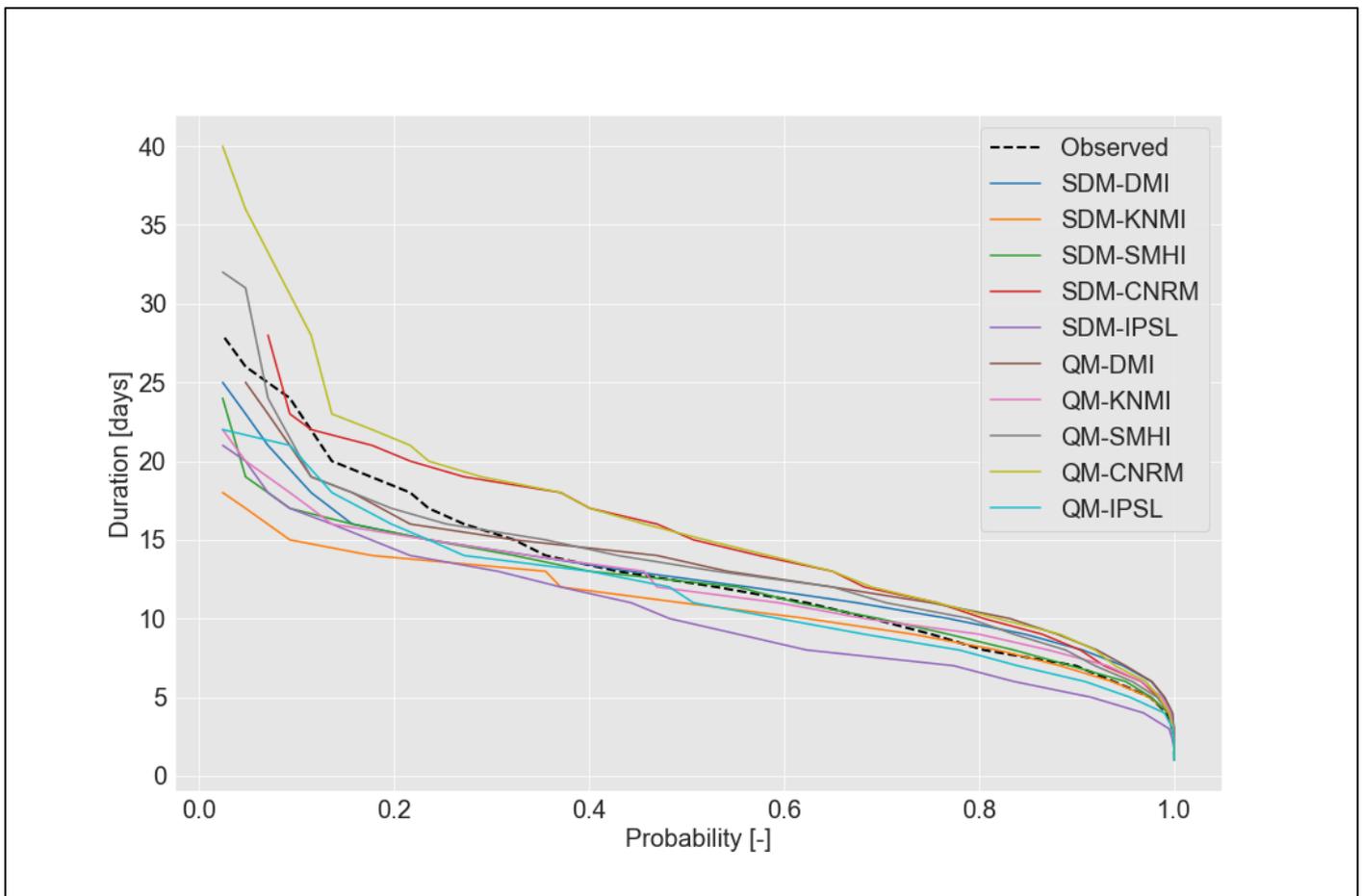


Figure 8-8: Dry spell analysis of the bias-corrected simulations at Lobith for the period (2020 - 2060).

The SPI analysis of the KNMI datasets indicate the highest meteorological droughts are set to on the year 2045 and 2059. The SMHI subsets hint at meteorological droughts during the year 2025, 2035, 2059 and 2060. The same results are found using QM or SDM (with minor differences)

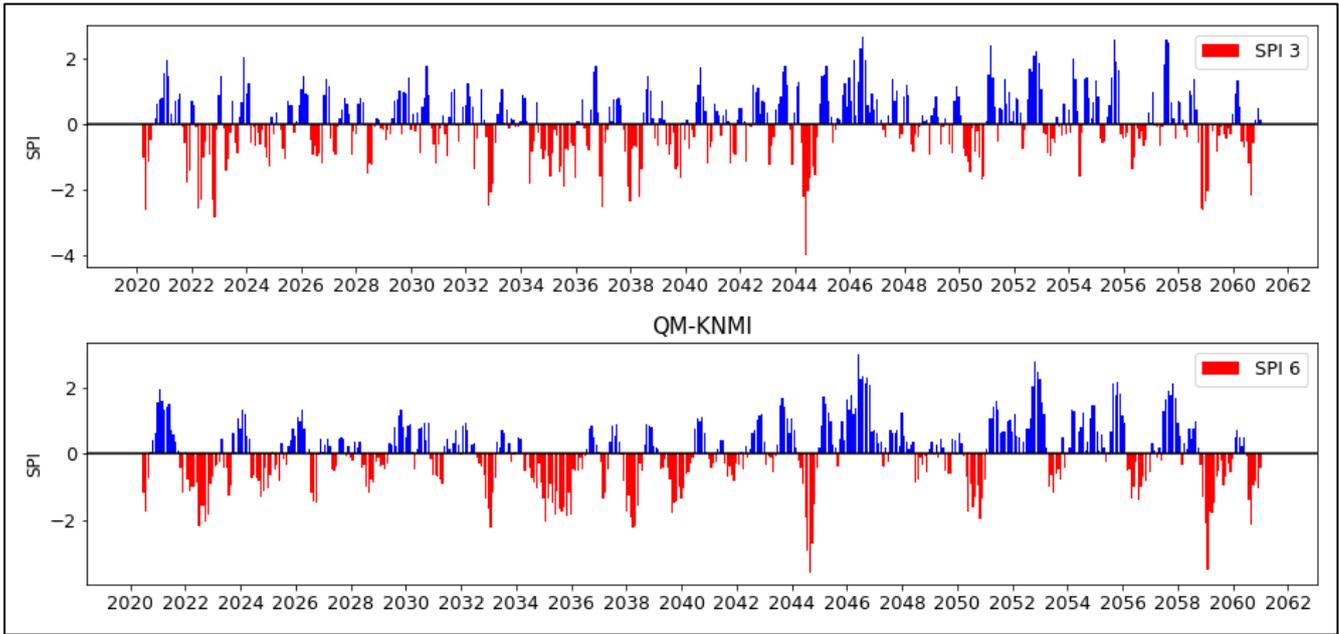


Figure 8-9: SPI Calculation for the projected period (2020 - 2060) for the corrected KNMI subset using QM

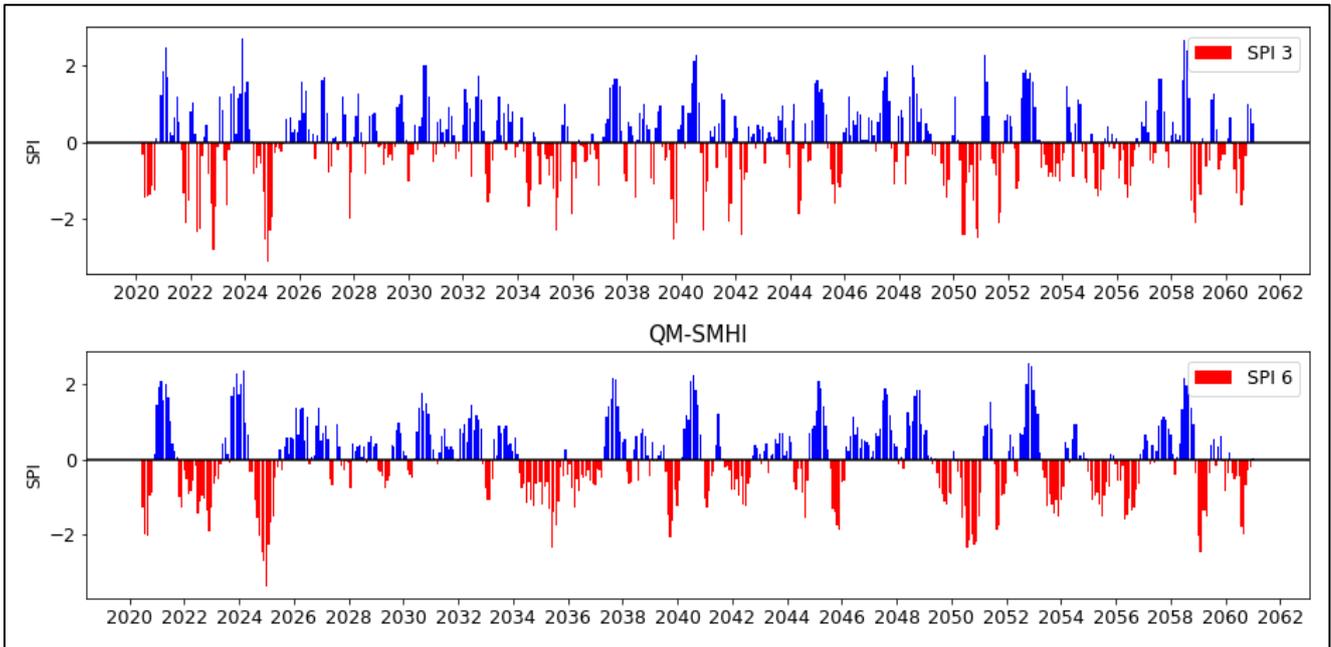


Figure 8-10: SPI Calculation for the projected period (2020 - 2060) for the corrected SMHI subset using QM.

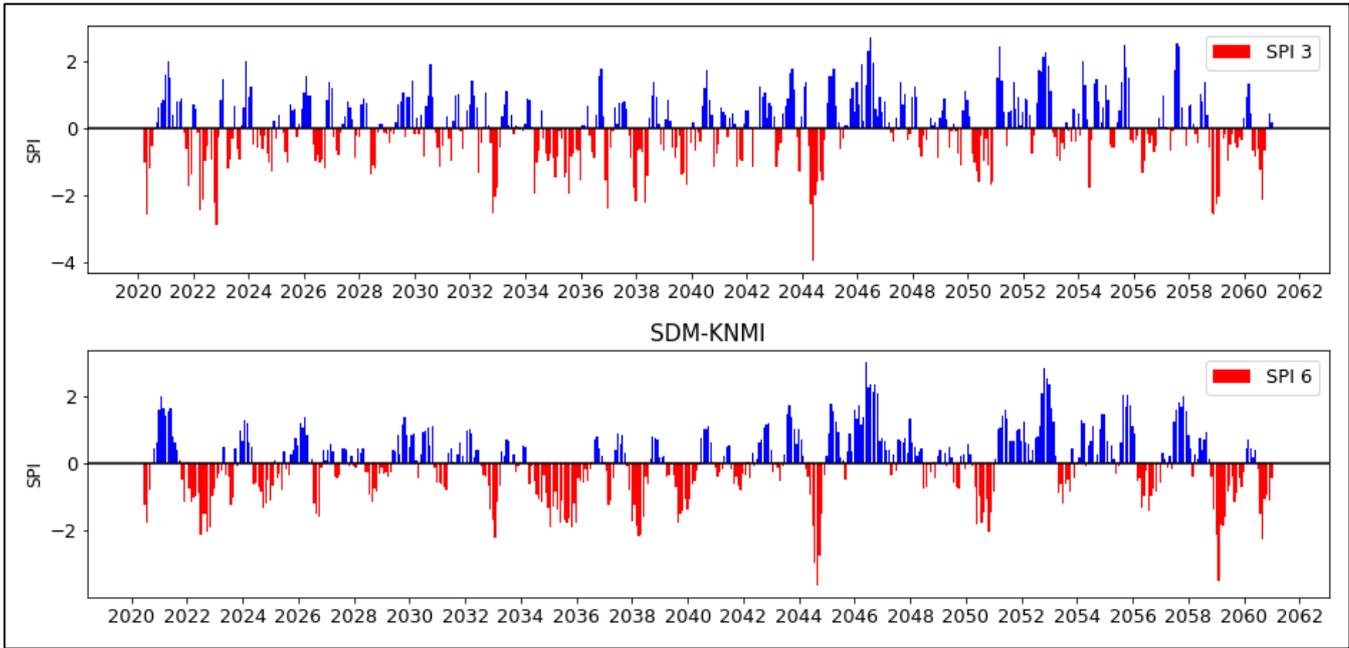


Figure 8-11: SPI Calculation for the projected period (2020 - 2060) for the corrected KNMI subset using SDM

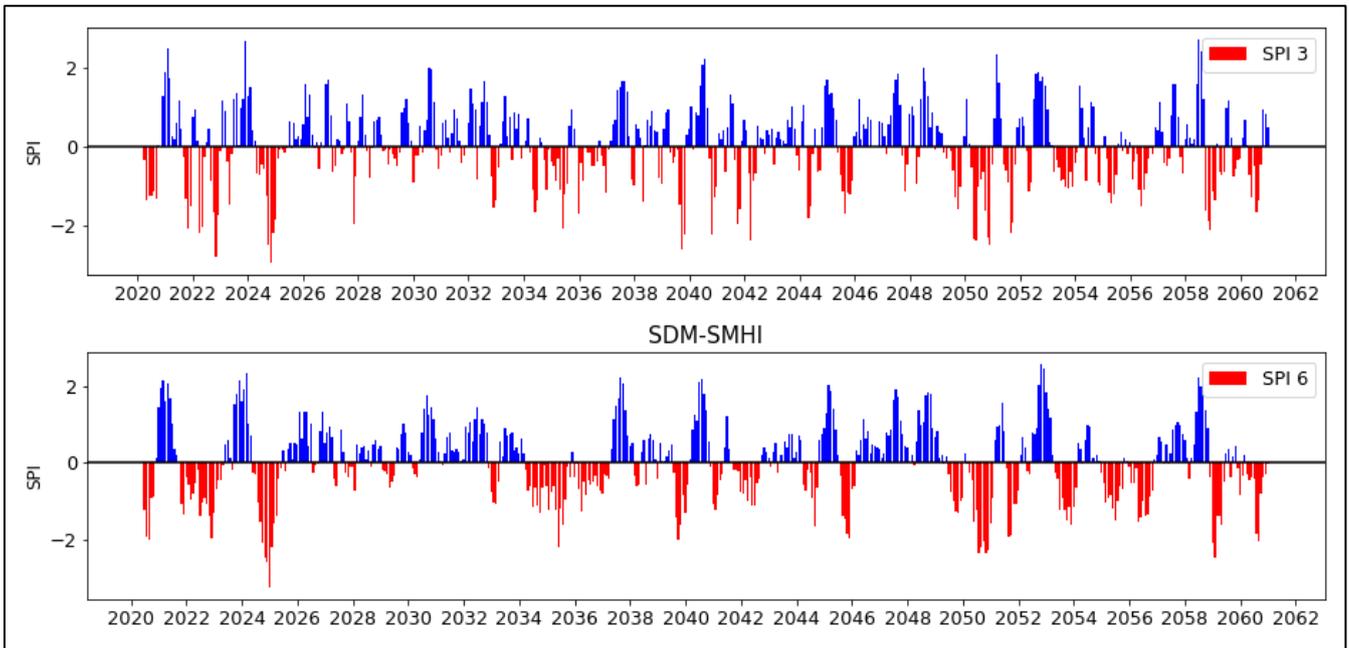


Figure 8-12: SPI Calculation for the projected period (2020 - 2060) for the corrected SMHI subset using SDM

9| Appendix III: Kaub, Cochem, and Rees

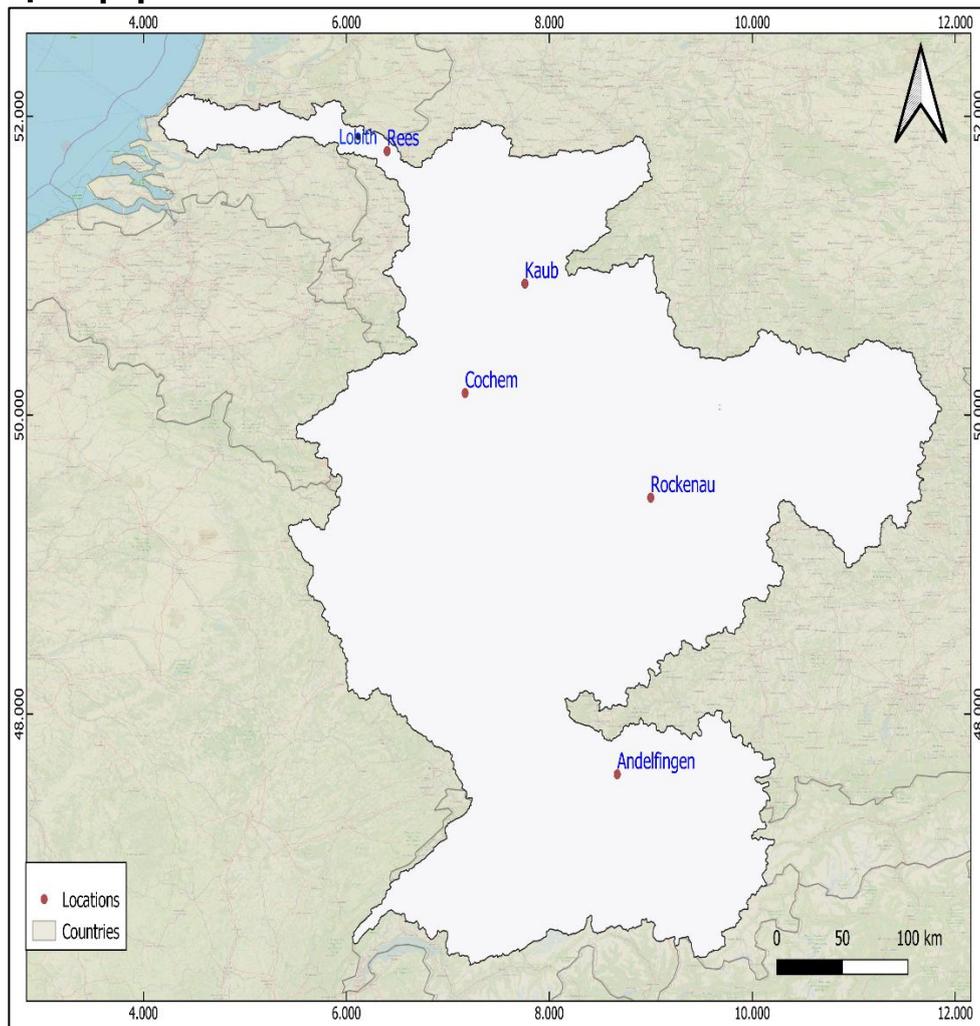


Figure 9-1: Description of how Kaub, Rees, Cochem, Rockenau and Andelfingen are distributed within the Rhine basin.

As an additional check for the historical experiment, a set of three locations are selected for validation (i.e., correcting the projections for the period (1999 – 2018)).

The locations are Kaub, Rees and Cochem. These three towns are in Germany.

The reason for choosing these locations stems from the fact that they are considered as bottleneck areas in the Rhine basin low flow related applications. The analysis is based on three indices (7-day minimum low flow (ND-7), long-term annual mean flow and lengths below threshold³).

As it appears from the figures below for each location, applying bias correction to the simulations resulted in a clear improvement at Rees and Cochem. The calculated stats from the corrected simulations is close to that obtained by forcing the `wflow_sbm` with the actual observations, that can be seen in the 7-day minimum low flow values (ND-7), long-term annual mean flow and lengths below threshold value. However, at Kaub the case is different. As the calculated stats at Kaub are found to less accurate from the actual observations, the raw projections provided more consistent results than the bias-corrected projections. That can be because the `wflow_sbm` is found not to perform well at Kaub compared to the other three locations.

This indicate that the success of the bias correction depends on the hydrological model skill at the considered location. Overall, the bias-corrected projections provide better results than the raw simulations relative to the actual observations.

³ The selected thresholds are $300 \text{ m}^3/\text{s}$ for Cochem, $1000 \text{ m}^3/\text{s}$ for Kaub and $1500 \text{ m}^3/\text{s}$ for Rees.

9.1 Kaub

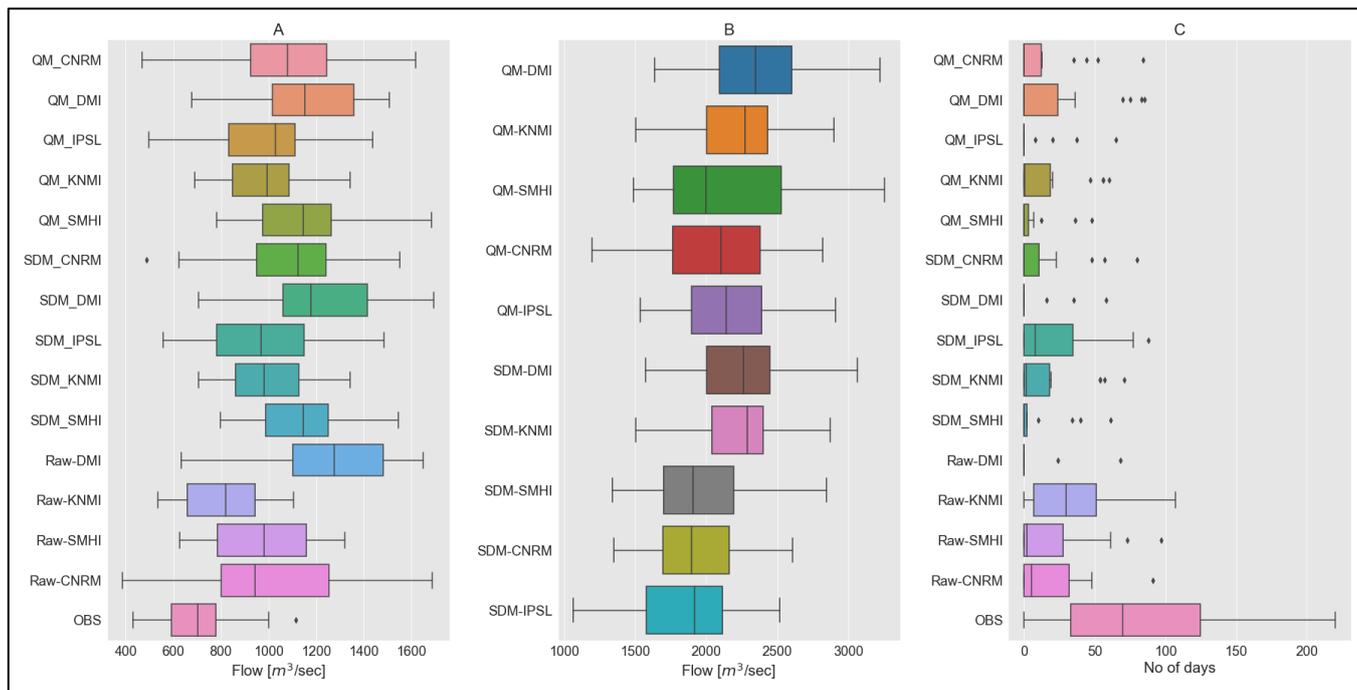


Figure 9-2: Long-term statistics at Kaub, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 1000 m^3/s at Kaub.

9.2 Cochem

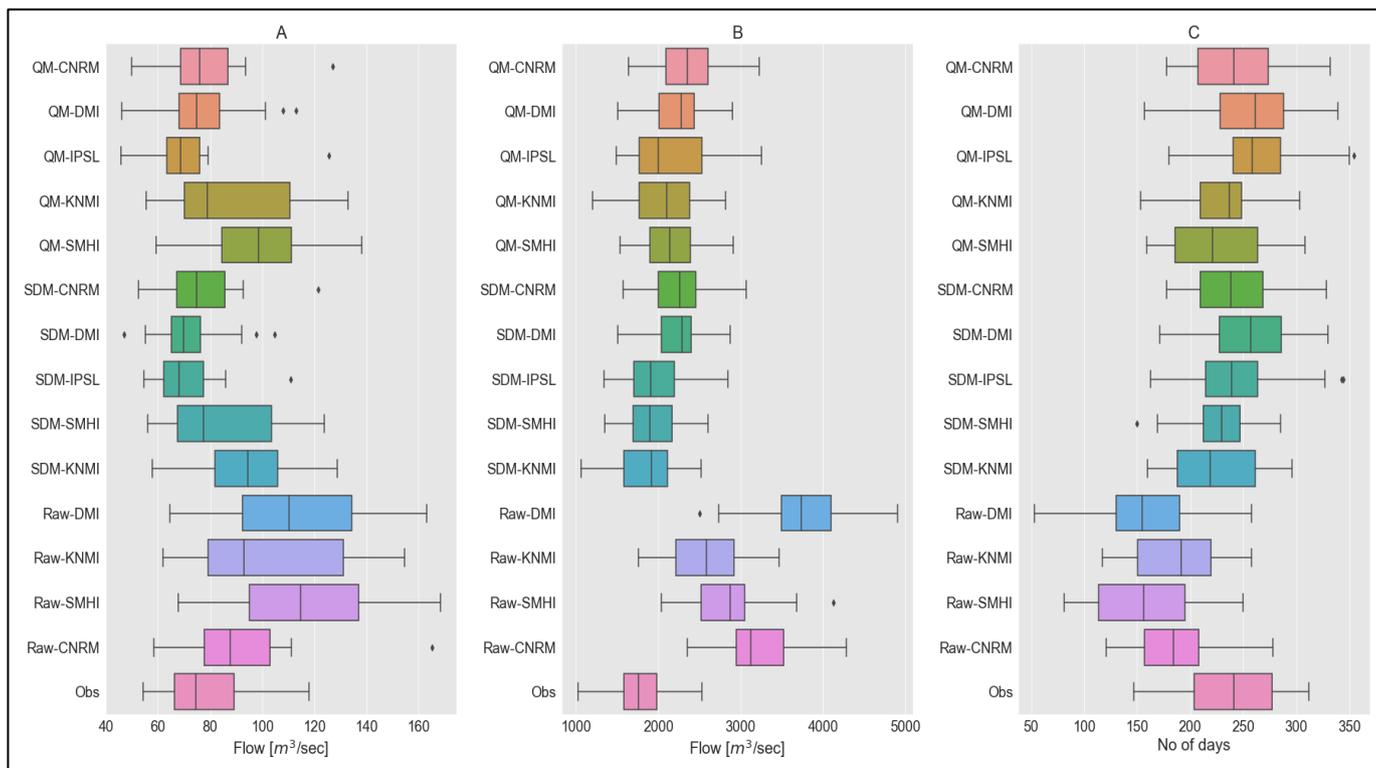


Figure 9-3: Long-term statistics at Cochem, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 300 m^3/s at Cochem

9.3 Rees

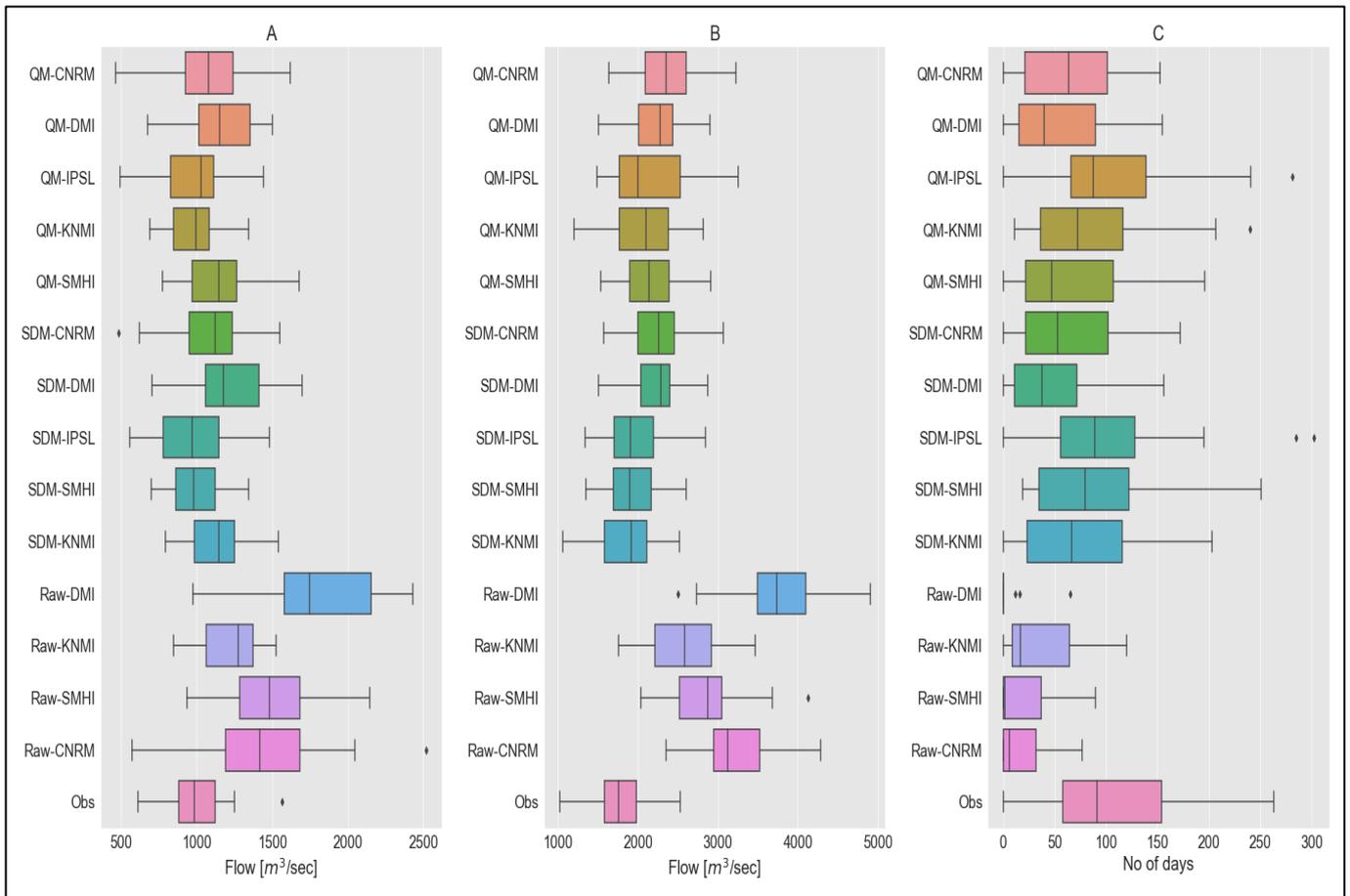


Figure 9-4: Long-term statistics at Rees, A is for 7-day minimum low flow, B is for the long-term mean annual flow and C is the lengths below 1500 m^3/s at Rees.

10| Appendix IX: Rockenau and Andelfingen

Rockenau and Andelfingen are chosen to examine the changes exhibited at Lobith when projecting to 2060. These two locations show a strong dependence on the snow melt from the Alps. As snow is the major contributor to the flow observed at these locations. Rockenau is in Germany, while Andelfingen is in Switzerland.

All the corrected simulations have resulted in lower average discharges at Rockenau and Andelfingen compared to what is observed by forcing `wflow_sbm` with historical observations. These locations are set to experience lower average discharges during summer period, due to the lower contribution from the Alps for the coming period due to climate change. One can expect relatively higher discharges within summer period for in the near future because of excessive snow melting at the start of the summer period.

Once again, the envelope of the two corrected projections is similar, indicating similar performance between QM and SDM in removing the systematic bias from climate simulations. The uncertainty band in QM simulations is lower than in SDM simulations.

10.1 Rockenau

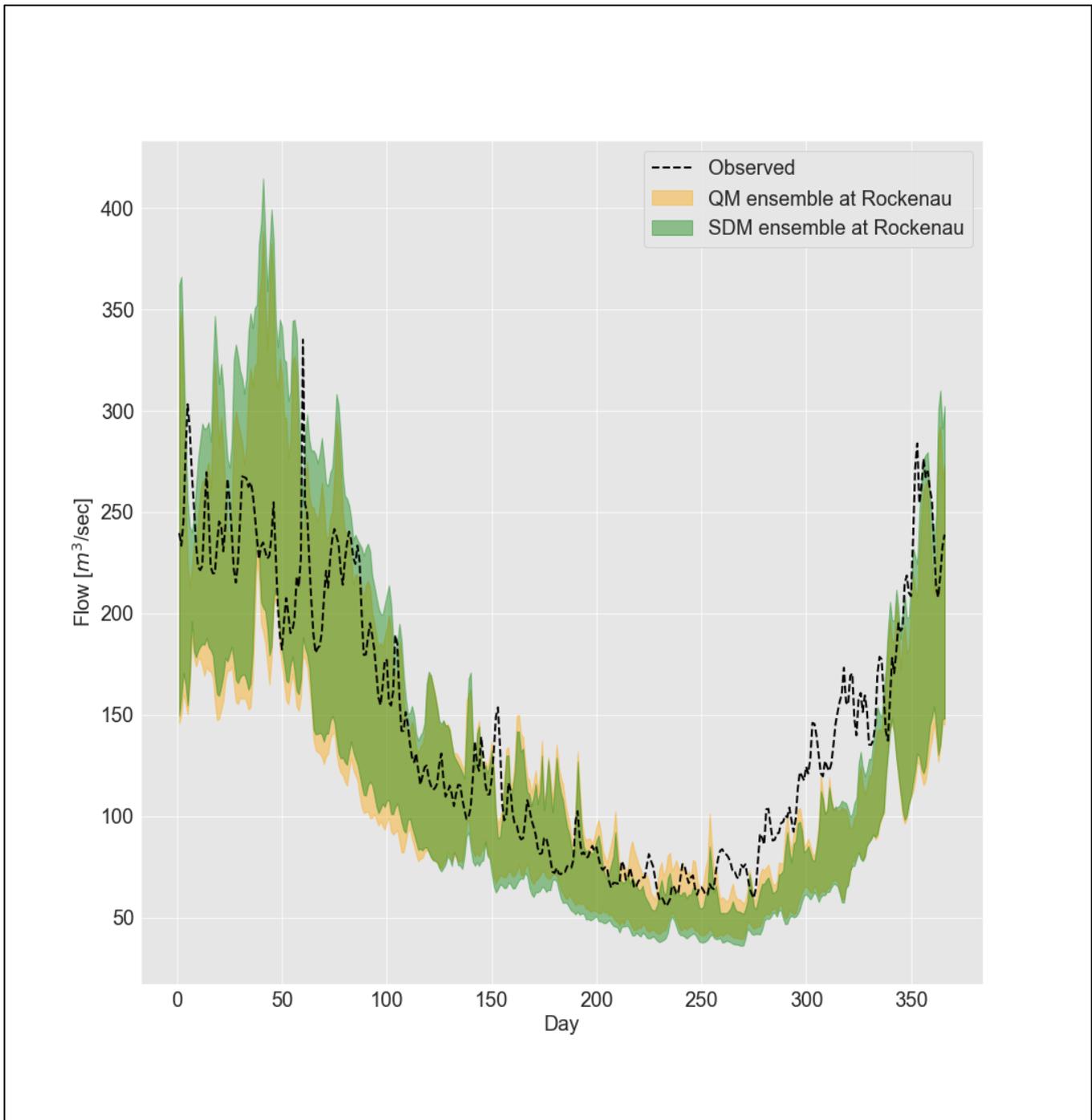


Figure 10-1: Rockenau average daily discharge for the period (2020 - 2060) vs the actual discharges for the period (1979 - 2019).

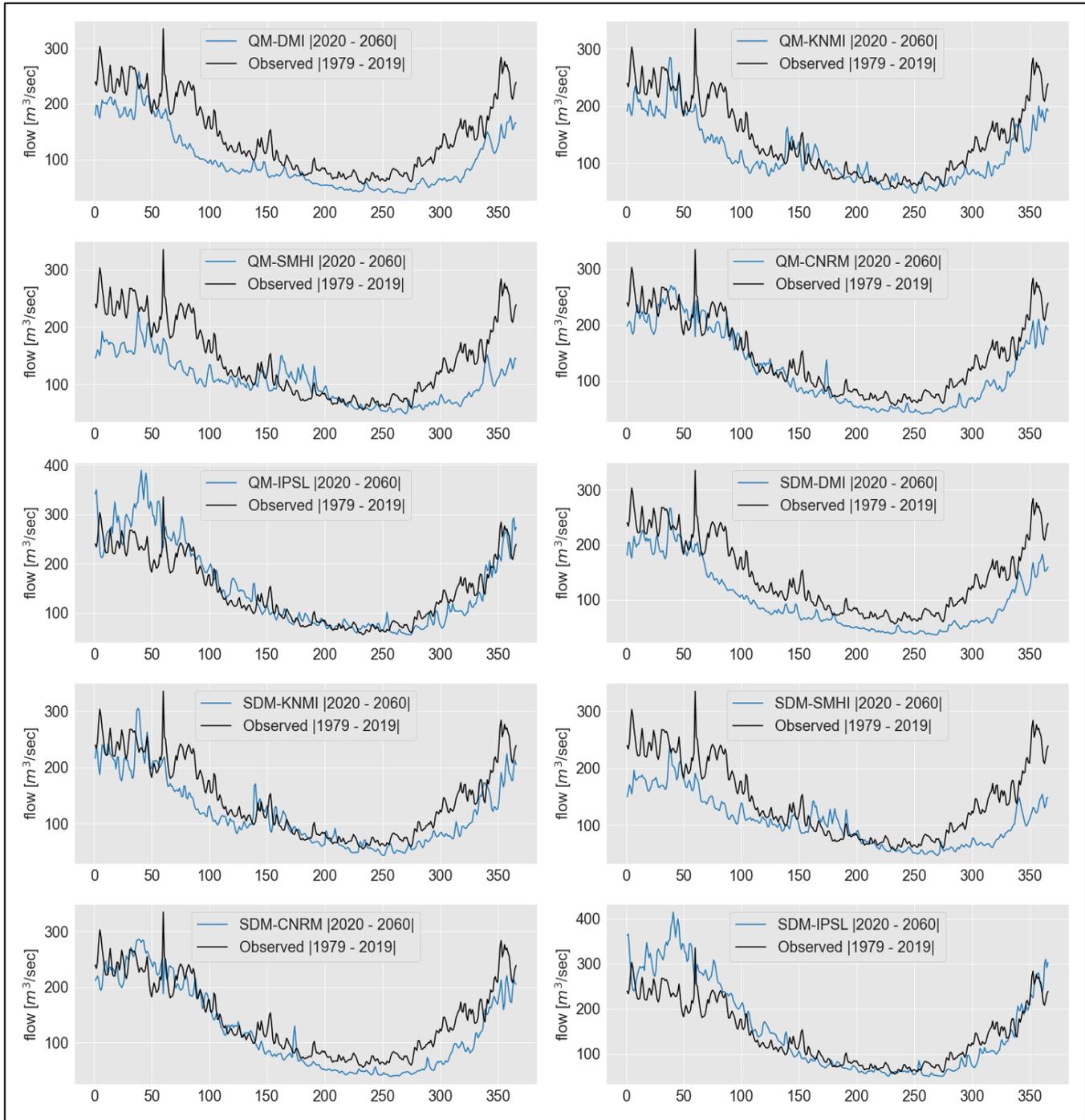


Figure 10-2: Bias-corrected simulations at Rockenau for the period (2020 - 2060) compared to the actual discharges for the period (1979 - 2019)

10.2 Andelfingen

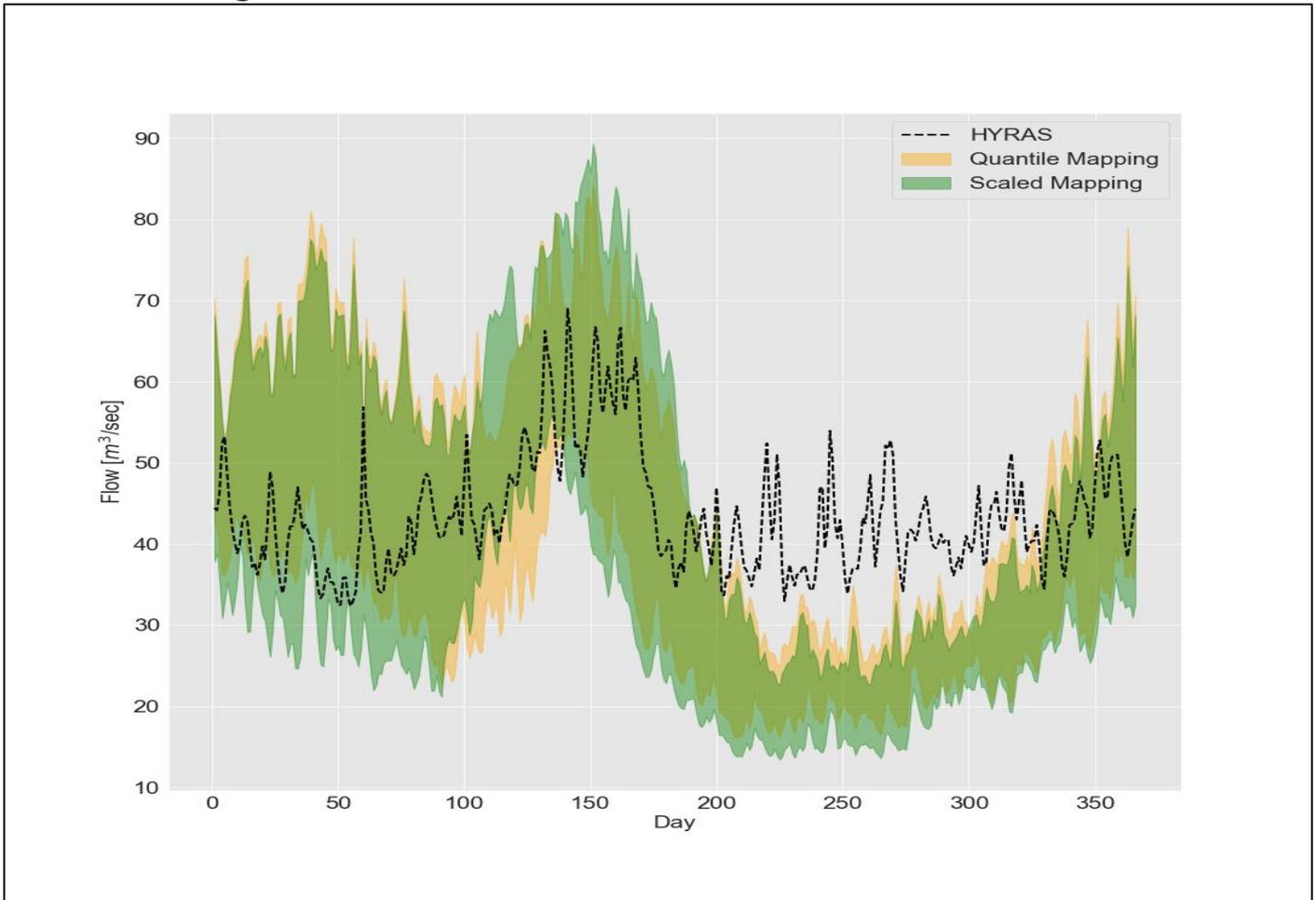


Figure 10-3: Andelfingen average daily discharge for the period (2020 - 2060) vs the actual discharges for the period (1979 - 2019).

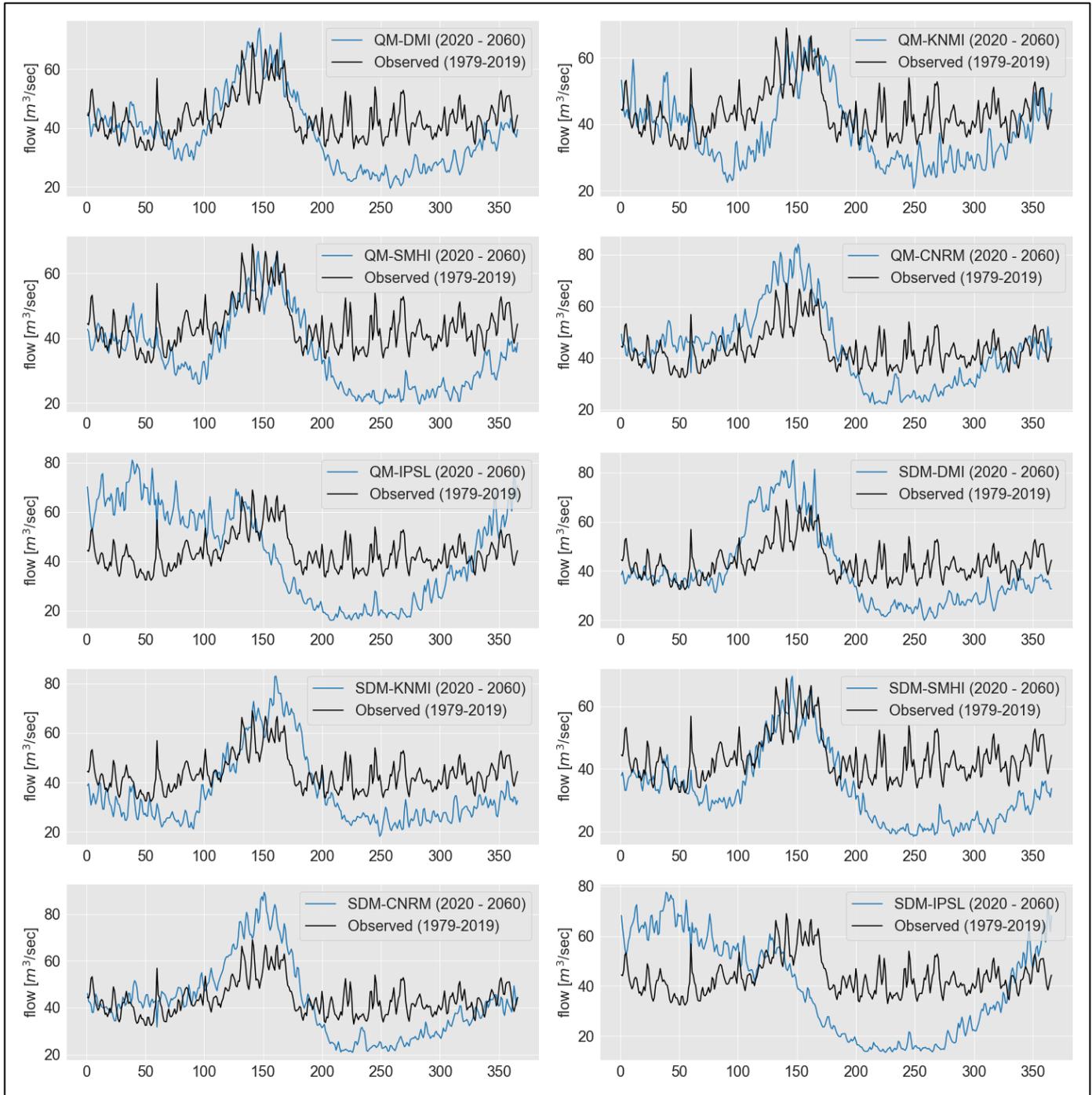


Figure 10-4: Bias-corrected simulations at Andelfingen for the period (2020 - 2060) compared to the actual discharges for the period (1979 - 2019).

11| Appendix X

11.1 Spatiotemporal REA

In extension to the REA method discussed in the methods section (section 3.4.2), two other modifications to this method are introduced. Considering that none of the raw subsets used in this project provided lower values than the actual discharge and the unreliability due to the wet bias, REA and ClimWIP could not compete with QM and SDM.

In this appendix section, the temporal REA, and spatial REA methods are applied following the framework proposed by Tegegne et al., 2019. However, this method is applied to modelled discharges and not to precipitation as in their paper. The goal is to see if there any added value when considering the spatial heterogeneity in the Rhine basin, by incorporating the measurements of three other locations (Basel, Kaub and Cochem) (see section 2.2).

The spatial REA (S-REA), and temporal REA (T-REA) weights are obtained by applying the following expressions, to come with a set of weights that consider the variation in time and space (incorporated into the ST-REA). All while incorporating the information of RCM i at site j , and time t . X and O represent the simulated and modelled discharges respectively.

$$T\text{-REA} = \left\{ \left[\frac{\epsilon^T}{\sqrt{(B_i^T)^2 + \text{var}_i^T}} \right]^m \left[\frac{\epsilon^T}{[\text{abs}(D_i^T)]} \right]^n \right\}^{1/(m*n)} \quad (\text{A17})$$

$$S\text{-REA} = \left\{ \left[\frac{\epsilon^S}{(B_i^S)^2 + \text{var}_i^S} \right]^m \left[\frac{\epsilon^S}{[\text{abs}(D_i^S)]} \right]^n \right\}^{1/(m*n)} \quad (\text{A18})$$

$$ST\text{-REA} = \left\{ \left[\frac{\epsilon^{ST}}{(B_i^{ST})^2 + \text{var}_i^{ST}} \right]^m \left[\frac{\epsilon^{ST}}{[\text{abs}(D_i^{ST})]} \right]^n \right\}^{1/(m*n)} \quad (\text{A19})$$

$$D_i^T = \Delta x_i^T - \frac{\sum_{i=1}^N w_i^T * \Delta x_i^T}{\sum_{i=1}^N w_i^T} \quad (\text{A20})$$

$$D_i^S = \Delta x_i^S - \frac{\sum_{i=1}^N w_i^S * \Delta x_i^S}{\sum_{i=1}^N w_i^S} \quad (\text{A21})$$

$$D_i^{ST} = \Delta x_i^{ST} - \frac{\sum_{i=1}^N w_i^{ST} * \Delta x_i^{ST}}{\sum_{i=1}^N w_i^{ST}} \quad (\text{A22})$$

$$B_i^T = \frac{1}{T} \sum_{i=1}^T \left[\left(\frac{1}{J} \sum_{j=1}^J X_{i,j,t} \right) - \left(\frac{1}{J} \sum_{j=1}^J O_{i,j,t} \right) \right] \quad (A23)$$

$$B_i^S = \frac{1}{J} \sum_{i=1}^J \left[\left(\frac{1}{T} \sum_{t=1}^T X_{i,j,t} \right) - \left(\frac{1}{T} \sum_{t=1}^T O_{i,j,t} \right) \right] \quad (A24)$$

$$B_i^{ST} = \frac{1}{J+T} \left[\sum_{j=1}^J \sum_{t=1}^T (X_{i,j,t} - O_{j,t}) \right] \quad (A25)$$

$$\text{var}_i^T = \frac{1}{T} \sum_{i=1}^T \left[\left(\frac{1}{J} \sum_{j=1}^J X_{i,j,t} \right) - \frac{1}{T} \left(\sum_{t=1}^T \frac{1}{J} \sum_{j=1}^J X_{i,j,t} \right) \right]^2 \quad (A26)$$

$$\text{var}_i^S = \frac{1}{J} \sum_{j=1}^J \left[\left(\frac{1}{T} \sum_{t=1}^T X_{i,j,t} \right) - \frac{1}{J} \left(\sum_{j=1}^J \frac{1}{T} \sum_{t=1}^T X_{i,j,t} \right) \right]^2 \quad (A27)$$

$$\text{var}_i^{ST} = \frac{1}{J+T} \left\{ \sum_{j=1}^J \sum_{t=1}^T (X_{i,j,t} - \frac{1}{J+T} \sum_{j=1}^J \sum_{t=1}^T X_{i,j,t}) \right\}^2 \quad (A28)$$

$$\varepsilon^T = \max(\bar{O}_1^S, \dots, \bar{O}_T^S) - \min(\bar{O}_1^S, \dots, \bar{O}_T^S) \quad (A29)$$

$$\varepsilon^S = \max(\bar{O}_1^T, \dots, \bar{O}_J^T) - \min(\bar{O}_1^T, \dots, \bar{O}_J^T) \quad (A30)$$

$$\varepsilon^{ST} = \max \begin{bmatrix} O_{1,1} & \dots & O_{1,T} \\ \vdots & \ddots & \vdots \\ O_{J,1} & \dots & O_{J,T} \end{bmatrix} - \min \begin{bmatrix} O_{1,1} & \dots & O_{1,T} \\ \vdots & \ddots & \vdots \\ O_{J,1} & \dots & O_{J,T} \end{bmatrix} \quad (A31)$$

Where var_i^S is the spatial variance for RCM i , var_i^T is the temporal variance for RCM i . The measure of natural variability in time ε^T , and in space ε^S . D_i^S is the spatial divergence between the ensemble members, and D_i^T is the temporal divergence, which are expressed as the mean difference between each simulation and the rest of the simulations. B_i^T and B_i^S are the spatial and temporal performance of each RCM i , this is calculated by the mean difference between each RCM's simulation and observations at that point.

These parameters are then used to calculate the spatiotemporal REA weights using the D_i^{ST} which is the spatiotemporal divergence between the ensemble members, ϵ^{ST} natural variability in time and space, var_i^{ST} is the spatiotemporal variance, and finally B_i^{ST} the spatiotemporal performance of RCM i .

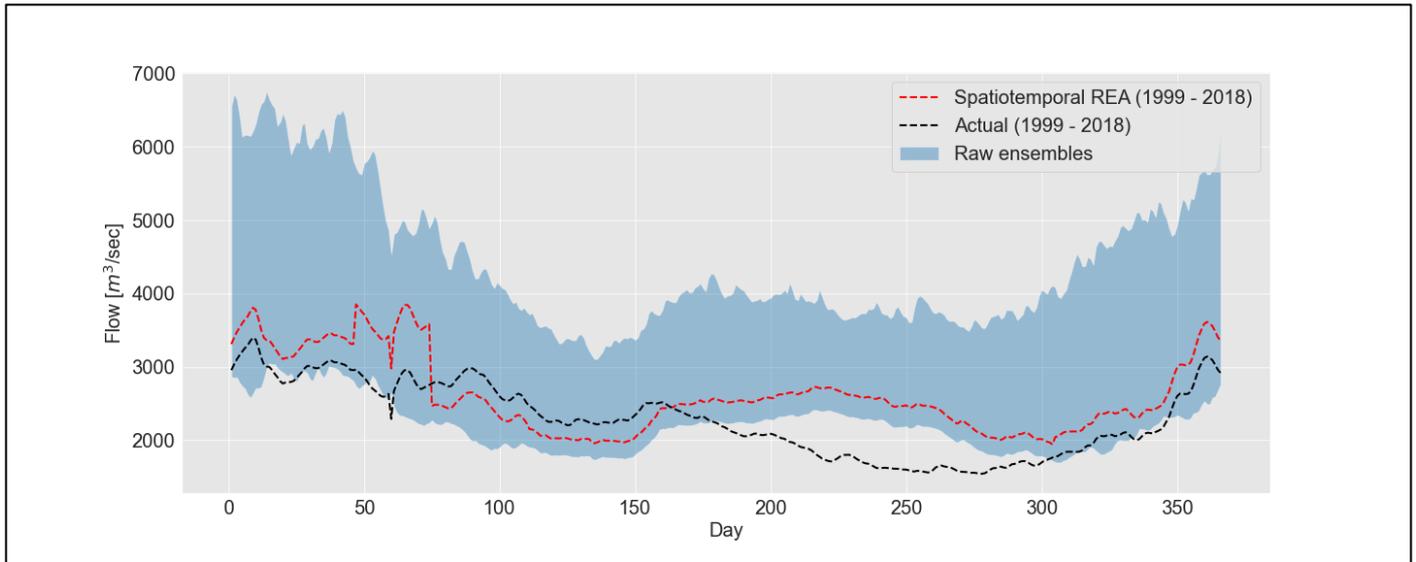


Figure 11-1: Spatiotemporal REA weighting method applied at Lobith based on Kaub, Cochem, and Basel for the period (1998 - 2018).

It appears that incorporating space and time variability when applying performance-based weighting resulted in noticeable improvements in the average change signal of the raw simulations, when compared to using REA and ClimWIP.

The latter methods consider the average change across the entire Rhine basin and generalize it on each grid cell located within. This assumption neglects any consideration of spatial heterogeneity related factors. The error in the weighted average daily discharge at Lobith using the ST-REA compared to the actual daily discharge for the same period (1999 – 2018) appears to be the lower relative to the weighted daily average discharge using ClimWIP and REA for the same period.

