
**Assessing the impact of climate change on
streamflow in catchments across the United
States, using an easily reproducible modeling
approach**

By

Vincent Hoogelander

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

December 7th, 2022

Student number: 4709330
Thesis committee: Dr. Rolf Hut (Chair, TU Delft)
Dr. Markus Hrachowitz (TU Delft)
Dr. Stef Lhermitte (TU Delft)

Abstract

In this thesis, an easily reproducible modeling approach was developed for assessing the climate change impact on streamflow. This approach was tested by using it to assess the impact of climate change on streamflow in 5 different contrasting catchments across the United States. Many studies show that climate change is expected to influence streamflow regimes all over the world. However, these studies are often difficult to reproduce because the modeling approaches used are usually only locally applicable. In the approach used in this study, hydrological model calibration and validation were done using open-accessible ERA5 forcing together with observed streamflow data provided by the GRDC. The model performed best in a mountainous catchment, while the worst performance was found in a dry catchment and a catchment containing several lakes. The low performances here are mainly caused by imperfect forcing data used for calibration and the neglect of lake processes. The climate change impact analysis used forcing from two CMIP6 models with the SSP245 and SSP585 scenarios. The projections showed significant changes in streamflow in colder regions, which are most likely related to changing snow melt processes. The main finding in warmer regions is that streamflow is generally expected to decrease in the drier periods. Changes of streamflow in these regions are most likely related to changes in precipitation and evaporation processes. However, results remain very uncertain due to disagreements between climate models and sometimes doubtful performance of the hydrological, caused by oversimplification of the model and imperfect ERA5 calibration data. The designed modeling approach facilitates reproducibility of climate change impact analyses in a wide range of catchments using different climate models and scenarios. Its use makes it easier to expand similar analyses to a large ensemble of these aspects.

Acknowledgements

This MSc thesis is the final part for fulfilling the graduation requirements of my MSc Civil Engineering at TU Delft. Within this Master, I followed the track Water Management with a specialization in Hydrology.

In this way, I would like to express my gratitude to a number of people. Firstly, many thanks goes to my daily supervisor, Rolf Hut. I have learned a lot from your expertise, and your ideas and feedback helped me a lot throughout the project. Also, your guidance beyond the technical aspects really helped me at times during the project. Additionally, I would like to thank the rest of my committee, Markus Hrachowitz and Stef Lhermitte, for providing me with feedback during our meetings. Also, thanks goes to Jerom Aerts (PhD student at TU Delft) and Peter Kalverla (Netherlands eScience Center) for their technical support during the project.

I also want to thank my parents and brother for their support throughout my whole study career. Lastly, special thanks goes to Mariska, who gives me unconditional support and love at all times.

*Vincent Hoogelander
Delft, December 2022*

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Study goals	2
1.3	Thesis outline	2
2	Methods	3
2.1	Study area	3
2.2	Hydrological model	5
2.3	Input data and climate simulations	5
2.4	Model calibration and validation	6
2.5	Climate change impact analysis	7
2.5.1	Annual, seasonal and monthly changes	7
2.5.2	Changes in minimum and maximum annual streamflow	7
2.5.3	Uncertainty analysis	8
2.6	Reproducible modeling approach	8
3	Results	10
3.1	Model calibration and validation	10
3.2	Historical streamflow simulations	11
3.3	Climate change impact analysis	13
3.3.1	Annual, seasonal and monthly changes	13
3.3.2	Changes in minimum and maximum annual streamflow	15
3.3.3	Conclusions and linkages of hydrological changes per catchment	18
3.3.4	Uncertainty analysis	19
4	Discussion	21
4.1	Historical streamflow simulations	21
4.1.1	Calibration and validation	21
4.1.2	Historical simulations using CMIP6 climate models	23
4.2	Climate change impact projections and uncertainties	23
4.3	Reproducible modeling approach	24
5	Conclusion	27
A	Implementation of the HBV-mountain model to the eWaterCycle platform	33
A.1	Add a BMI to the hydrological model	33
A.2	ESMValTool	33
A.3	Docker container	35
B	Climate change impact analysis tool	36
B.1	Features of the HBV-mountain BMI model	36
B.2	Code structure	37
B.3	Increasing computation speed for model calibration	38
B.4	Available CMIP6 models	39
C	Calibration and validation of the hydrological model	40
C.1	Objective functions	40
C.2	Parameter constraints	41
C.3	CMA-ES calibration iterations	42
D	Additional data	43
D.1	Calibration and validation results	43
D.2	Climate change impact analysis	44

Introduction

1.1. Introduction

Climate change is expected to influence climate, atmospheric and hydrological processes all over the world (IPCC, 2021). For example: precipitation patterns and snowmelt processes will be altered as a result of increasing air temperatures, leading to changes in streamflow regimes (Hanus et al., 2021, Roudier et al., 2014, Huang et al., 2020, Hirpa et al., 2019, Neves et al., 2020, Ndhlovu and Woyessa, 2021, Brunner et al., 2020, Lane et al., 2022). Impacts on runoff signatures due to hydrological changes can be assessed using a combination of hydrological and climate models. In this thesis, "climate models" will be used as a catch all for global recirculation models, including re-analysis datasets. The most common approach used in studies focusing on climate change impact on streamflow is to use one hydrological model in a limited amount of catchments. Another approach that can be found in the literature is the use of large multi-model ensembles, disregarding individual model performances. The study of Krysanova et al. (2018) argues, however, that this second approach is less trustworthy regarding the analysis climate change impact assessment, as these studies often use an ensemble of both good and poor performing models. Studies that focus both on using a wider variety of hydrological models and individual model performances in different contrasting catchments are relatively limited, however. Besides the possibility of using different hydrological models, studies related to climate change impact can differ in the number of used climate models, scenarios and catchments. In the work of Clark et al. (2016) it is found that the results of climate change impact studies bring along a lot of uncertainties related to these aspects. In climate change impact studies, a part of these uncertainties are frequently addressed, but in the field of hydrological models it is often difficult for researchers to tackle these uncertainties by using multiple hydrological models. The reason for this is that hydrologists are often hindered by limited model availability, as many existing hydrological models are only locally applicable and are programmed in different languages (Hut et al., 2017; C. Hutton et al., 2016). This makes it difficult to reproduce the hydrological model in another catchment, and to make an assessment of the performance and uncertainties of predictions between different hydrological models and catchments, together with different climate model combinations. Still, in the literature it is mostly found that the climate model uncertainty is more dominant than the hydrological model uncertainty (De Niel et al., 2019, Her et al., 2019). Nonetheless, working with both multiple hydrological, climate models and catchments, while taking individual model performances into account, could enhance the overall credibility of future projections.

The eWaterCycle platform aims to solve the problem of limited hydrological model availability by making hydrological models more FAIR (Findable, Accessible, Interoperable and Reproducible). In this platform, hydrologists are able to share their work and add new models, so that these models can be accessed by other hydrologists through a common interface in Python, where users are able to easily adapt some model acts as well (Hut et al., 2022). To make the hydrological models FAIR, all hydrological models on the eWaterCycle platform can be accessed through a Basic Model Interface (BMI) (E. Hutton et al., 2020) within a Jupyter Notebook environment. In this way, different hydrological models (possibly written in different programming languages) can be easily reproduced in the same manner. Forcing input for the hydrological models on the platform can be generated using the Earth System Model Evaluation Tool (ESMValTool) (Righi et al., 2020). Using this tool, forcing input from large climatic datasets can be easily and quickly generated for use in specific hydrological models without the need of different complex functions. This FAIR approach enhances the flexibility for modellers, as they are being able to use a wide variety of hydrological models for their research without the need of focusing on the more complex model technology.

The aim of this thesis was to make a climate change impact assessment on streamflow by building on the FAIR modeling approach of eWaterCycle. This is done by developing a flexible and reproducible approach for conducting reproducible climate change impact analysis on streamflow. This thesis report demonstrates the functionalities of this approach by performing climate change impact assessments on streamflow in five contrasting catchments across the United States, using two climate models from the CMIP6 dataset with two future climate scenarios (Eyring et al., 2016). The approach is built on the HBV-mountain hydrological model (Hanus et al., 2021), for which a BMI is added such that it can be used in the same way as the hydrological models on the eWaterCycle platform. Although this study still uses only one hydrological model, the approach can be redesigned such that more different BMI hydrological models can be used, so that multiple models can be used to reproduce similar analysis.

1.2. Study goals

The first goal of this thesis work was to develop a flexible modeling approach for conducting climate change impact analysis on streamflow in a reproducible way. This is done by firstly adding a BMI to the HBV-mountain hydrological model, as was developed in the work of Hanus et al. (2021), such that this model and the approach can be implemented on the eWaterCycle platform. The purpose of this first goal is to increase the flexibility for conducting climate change analyses on the eWaterCycle platform.

The second goal of this thesis was to use this approach to assess the impact of climate change on streamflow in five different contrasting climates located accross the United States. This is done by using two climate models from the CMIP6 dataset with two different future SSP scenarios. The purpose of this second goal was to demonstrate the functionalities and flexibility of the approach by showing the calibration process in the catchments and different results of the climate change impact assessment.

1.3. Thesis outline

This report is structured as followed: Chapter 2 discusses the methods used in this study. More specifically, Sections 2.1 - 2.5 discuss the methods used for the model calibration and climate change impact assessment, while Section 2.6 highlights the usage and some of the functionalities of the reproducible modeling approach. The results of calibration and climate change impact assessment are presented in Chapter 3, and these results are discussed in more detail in chapter 4. In the discussion section, Sections 4.1 and 4.2 focus on the found results in the climate change impact assessment, while Section 4.3 focuses on how the designed modeling approach and how this approach relates to the found results. The final conclusions are outlined in Chapter 5.

2

Methods

This chapter discusses the methods used for the climate change impact analysis, and how these methods are made reproducible. Section 2.1 focuses on the study area, Section 2.2 describes the hydrological model, and the input data for model calibration and making the streamflow simulations are described in Section 2.3. The calibration and validation process is explained in Section 2.4. The methods for the climate change impact analysis are explained in Section 2.5. Lastly, Section 2.6 presents the usage and some of the functionalities of the reproducible modeling approach.

2.1. Study area

This study focused on the climate change impact on streamflow in five different catchments located across the United States. Most of these selected catchments are located in different climate zones, according to the Köppen climate classification (Beck et al., 2020, Köppen, 1936). Furthermore, the selected catchments differs in the elevation levels, land use proportions, drainage area and long term mean discharge. Land cover types are taken from the 2001 USGS National Land Cover (NLCD) Database (Homer et al., 2007), catchment elevations are taken from the Digital Elevation Model of NASA SRTM (NASA, 2013). The NLCD land cover data is classified in the four different land cover types used in the HBV-mountain hydrological model (section 2.2). The riparian zone is estimated by assuming that the wet pixel area the NLCD data is equal to the riparian zone area, which is, given the spatial resolution of the NLCD data (30m) and the relatively small size of most of the catchments, a simplification of the method of using small fixed-width buffers for estimating riparian zones (e.g. Salo et al. (2016)). However, using this method likely underestimates the true riparian area in larger streams, as the riparian buffers are often larger here compared to smaller streams (Song et al., 2021). All catchment data are summarized in table 2.1 and the locations can be found in figure 2.1. Figure 2.2 shows examples of the elevation and land cover map of the Big Rock Creek catchment. This catchment is distinguished by its relatively dry climate. According to NLCD land cover data, the Thunder Creek catchment area consists of 11% glacier, and is the only one of the five catchments with glacial cover. This percentage of glacier cover is slightly less than the percentage glacier cover of the Thunder Creek catchment used in the studies of Fountain and Tangborn (1985) (14%). Streamflow in this catchment is predominantly influenced by glacier melt (Tangborn, 1980). Although the Kawishiwi catchment is located in the same climate zone as the Youghiogheny catchment, this catchment is selected because of its distinct elevation levels, low temperatures, and land use proportions: the Kawishiwi river passes through several lakes, which is the reason of its relatively large percentage of riparian zone. Also, the small difference in the minimum and maximum elevation level in this catchment differs from the other catchments.



Figure 2.1: Locations of all selected catchments in this study

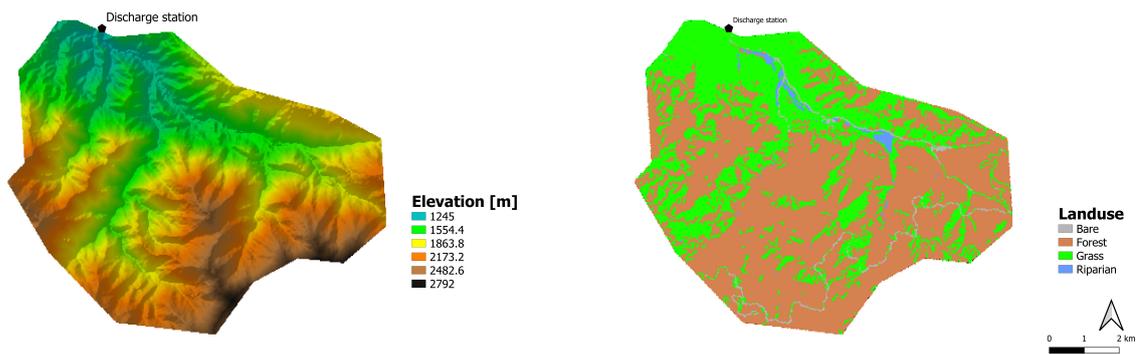


Figure 2.2: Elevation (left) and landuse (right) map of the Big Rock Creek catchment

Table 2.1: Catchment data

	Big Rock Creek	Thunder Creek	Conasauga river	Kawishiwi river	Youghiogheny river
GRDC station number	4118220	4146370	4149430	4113900	4123345
Drainage area (km²)	59.3	271.9	652.68	657.9	347.1
Elevation (m)	1245 - 2792	396 - 2746	187 - 1284	436 - 527	716 - 1024
Min-max temperature (°C)	5.5 - 24.5	-6.0 - 12.8	3.4 - 25.3	-13.5 - 18.4	-2.4 - 20.2
Annual precipitation (mm y⁻¹)	361	1755	1369	786	1418
Long term mean Discharge (m³ s⁻¹)	0.78	17.9	13.8	5.0	9.2
Bare (Glacier) (%)	2 (0)	32 (11)	6 (0)	0 (0)	4 (0)
Forest (%)	62	46	71	65	60
Grass (%)	35	22	21	0	30
Riparian (%)	1	1	1	35	6
Climate classification	Csa Temperate, dry and hot summer	Dfc Cold, no dry season, cold summer	Cfa Temperate, no dry season, hot summer	Dfb Cold, no dry season, warm summer	Dfb Cold, no dry season, warm summer

2.2. Hydrological model

The HBV-mountain hydrological model is used for the climate change impact analysis tool, which is developed and used in the study of Hanus et al. (2021). In order to ensure the model can be used in a reproducible way, a Basic Modeling Interface (BMI) is added to the hydrological model (E. Hutton et al., 2020), of which the workflow can be found in appendix A. This process based, semi-distributed hydrological model consists of four different hydrological response units (HRUs): a bare or sparsely vegetated zone, forested hillslope, grassland hillslope and a riparian zone. The hydrological fluxes are quite similar in all units, except for the bare rock zone, which is distinguished by a glacier storage component and where interception storage is negligible (see figure 2.3). Besides, water can re-enter the soil from the slow reservoir in the riparian zone. The model is designed so that the land use and elevation proportions can be adjusted accordingly to the actual proportions in the catchment, as found in table 2.1.

The model uses precipitation and temperature as forcing input data, the parameter sets for each response unit, and some initial settings which are mainly catchment depended. These settings mostly concern catchment elevation, land use and amount of daily sunlight hours. In case of missing evapotranspiration data, the model estimates the daily potential evapotranspiration using the Thornthwaite method, which only uses the temperature and sunlight hours per day for the estimation of evapotranspiration (Thornthwaite, 1948). The equifinality of the model is reduced by keeping some of the model parameters the same for each HRU. The parameter set consists of a total of 20 different parameters, of which 8 are kept the same for all HRUs, giving the model a total of 20 degrees of freedom. The catchment elevations are incorporated in the model in the form of four elevation bands, which differs from snow melt processes due to different temperatures. A more detailed description of the hydrological model can be found in the paper of Hanus et al. (2021).

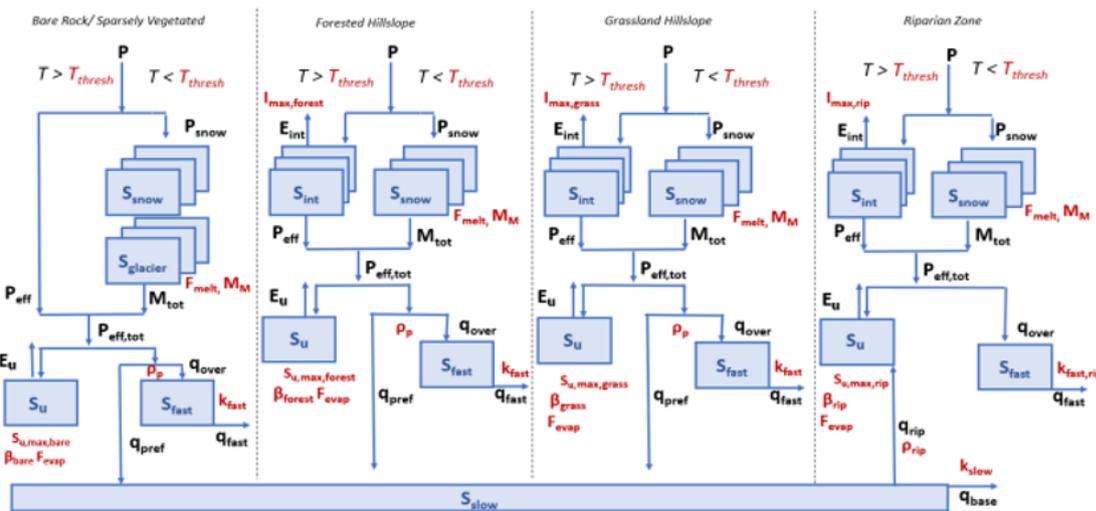


Figure 2.3: HBV-mountain model hydrological model used in this study (Hanus et al., 2021)

2.3. Input data and climate simulations

Using eWaterCycle's included datasets, streamflow observation data are taken from the Global Runoff Data Centre (GRDC, <https://www.bafg.de/GRDC>, last access: 11 October 2022). In case of the selected catchments, all discharge data is originally provided by the US Geological Survey (USGS). Precipitation and temperature data for model calibration is taken from the ERA5 global reanalysis dataset, which is also available within eWaterCycle (Hut et al., 2022). The ERA5 dataset is an ECMWF reanalysis product, whose data are global estimates developed based on a combination of historical (satellite and in-situ) observations together with advanced modelling and data assimilation systems (Hersbach et al., 2020). Due to the large complexity of integrating glacier retreat in the modeling process (Li et al., 2015), it is chosen to keep the glacier extent in the Thunder Creek constant over time. The selected time range for the calibration and validation data is from 1986 - 2005.

Two different climate models from the CMIP6 dataset are used for the climate change impact analysis:

the GFDL-CM4 model from GFDL-NOAA (USA) (Held et al., 2019) and the MPI-ESM1-2-HR model from MPI (Germany) (Gutjahr et al., 2018), with the years 1975 – 2005 as reference period and 2070 – 2100 as future period. Note that the start of the reference period is 10 years earlier than the start of the calibration period, this is done in order to expand the amount of data so that it can be compared with a longer future period. Before the CMIP6 data is used in the ERA5-calibrated model, the historical and future data is bias corrected with a Linear Scaling method using monthly precipitation and temperature averages (Lenderink et al., 2007, Azman et al., 2022). Precipitation data is bias-corrected using equation 2.1, and temperature data is corrected using equation 2.2. In these equations, $P_{\text{corrected}, d, m}$ and $T_{\text{corrected}, d, m}$ are the bias-corrected precipitation and temperature on the d^{th} day in the m^{th} month, $P_{d, m}$ and $T_{d, m}$ are the original precipitation and temperature data on the d^{th} day in the m^{th} month, $P_{\text{ERA5}, m}$ and $T_{\text{ERA5}, m}$ are the mean monthly precipitation and temperature of the ERA5 calibration data in the m^{th} month, and P_m and T_m are the original precipitation and temperature data in the m^{th} month.

$$P_{\text{corrected}, d, m} = P_{d, m} \left(\frac{P_{\text{ERA5}, m}}{P_m} \right) \quad (2.1)$$

$$T_{\text{corrected}, d, m} = T_{d, m} \left(\frac{T_{\text{ERA5}, m}}{T_m} \right) \quad (2.2)$$

Future climate simulations are made for two different Shared Socioeconomic Pathways (SSPs): SSP245 and SSP585 (Eyring et al., 2016). The SSP245 scenario assumes a medium level of radiative forcing with an increase of approximately 3°C by the end of the 21st century, while the SSP585 assumes a high level of radiative forcing with an increase of approximately 5°C by the end of the century (Tebaldi et al., 2020).

A summary of all forcing and observation datasets used in the analysis can be found in table 2.2. All forcing data is generated using the Earth System Model Evaluation Tool (ESMValTool) (Righi et al., 2020), for which an ESMValTool recipe and diagnostics script are made specifically for creating forcing input for the HBV-mountain hydrological model (see appendix A.2). The ERA5 data is accessed via the eWaterCycle platform, while the CMIP6 data is accessed via the High-Performance Computing cluster "Levante" (DKRZ, <https://docs.dkrz.de/doc/>, last access: 26 November 2022).

Table 2.2: Datasets for calibration and climate change impact assessment

Dataset	Calibration		Climate change impact analysis	
	ERA5	GRDC	GFDL-CM4	MPI-ESM1-2-HR
Spatial resolution (lon x lat)	0.1°x 0.1°	- (Station)	1°x 1°	0.9°x 0.9°
Temporal resolution	Hourly	Daily	Daily	Daily
Data	Precipitation Temperature	Discharge	Precipitation Temperature	Precipitation Temperature
Selected time range	1986 - 2005	1986 - 2005	1975 - 2005 2070 - 2100	1975 - 2005 2070 - 2100

2.4. Model calibration and validation

All 20 parameters are calibrated for the selected catchments, which is done by using the GRDC streamflow data together with the ERA5 precipitation and temperature data. All model parameters are process constraint following the suggestions of the study of Gharari et al. (2014), and similar parameter boundaries are used from the study of Hanus et al. (2021) (see appendix C.2). The overall model performance per parameter set is assessed by determining the Euclidean Distance (Hrachowitz et al., 2014, Schoups et al., 2005), which is calculated based on the Nash Sutcliffe efficiency (Nash and Sutcliffe, 1970) of four different individual objectives. The individual signatures are summarized in table 2.3, and the objective functions can be found in appendix C.1. Due to the sensitivity of the NSE method to large outliers in the data, the Flow Duration Curve is determined using the logarithm of the streamflows. The four objectives are used for both the calibration as the validation of the hydrological model.

The Euclidean Distance of the four different objectives is optimized using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) (Hansen, 2007, Ikram et al., 2022). The CMA-ES is an evolutionary algorithm for numerical optimization, which generates a new candidate parameter set solution from a current

solution. This new parameter sets solution is tested based on the objective function, and will be selected for the next iteration in case of a sufficiently good score. The calibration process is done by performing 50 different optimization processes with 50 iterations, resulting in 50 different parameter sets per catchment. Because of possible equifinality of the hydrological model, a set of the best parameter combinations should be used for further analysis, instead of only using the best model fit (Beven and Binley, 1992). Therefore, the 20 best performing parameter sets from the CMA-ES calibration are selected for further climate analysis. During the total calibration and validation time, there was a warming up period of 3 years (1986-1989), and 10 (1989-1999) and 6 (1999-2005) years for calibration and validation respectively. This warming up period was necessary for the model to properly estimate the initial storage conditions in the catchment. To increase the calculation speed, all calculations of the calibration process are done using the High-Performance Computing cluster "Levante" from DKRZ.

$$\sum_{i=1}^n P_{\text{ERA5}} - \sum_{i=1}^n E_{\text{modelled}} = \sum_{i=1}^n Q_{\text{observed}} + \epsilon \quad (2.3)$$

After the model is calibrated for all catchments with the ERA5 data, the results are compared with the observation data using the long term water balance to check for long term biases (ϵ) (equation 2.3). This is done by comparing the precipitation (P_{ERA5}) and the modelled evaporation (E_{modelled}) with the streamflow observations (Q_{observed}).

Table 2.3: Objectives used for calibration and validation of the hydrological model

Signature	Objective function	Note
Timeseries of streamflow	$NSE_{\text{timeseries}}$ $\text{Log } NSE_{\text{timeseries}}$	High flows Low flows
Flow duration curve	$NSE_{\text{flow duration curve}}$	Frequency of streamflow occurrence
Annual runoff coefficient	$NSE_{\text{annual runoff}}$	Precipitation fractionated into streamflow

2.5. Climate change impact analysis

2.5.1. Annual, seasonal and monthly changes

After calibration and running all historical and future streamflow simulations, the results are used for the climate change impact analysis in the catchments. The changes in annual, seasonal and monthly streamflow patterns are investigated by looking at differences in the monthly spread of the modelled total monthly streamflow of all years between the reference and future period. For this, the mean result of all simulations with the different parameter sets and climate models are used. These results are compared to the relative change in mean daily precipitation and the absolute change in mean daily temperature. In order to investigate the differences in the projections between the CMIP6 models, seasonal projections with both separate climate models are projected in the Conasauga catchment. Next to this, the changes in precipitation and temperature under the SSP585 scenario are compared to the relative changes in streamflow in the Big Rock Creek and Thunder Creek catchment. These are very contrasting catchments in terms of land use, total precipitation and temperature. By focusing on these two catchments, the linkage of how precipitation and temperature changes affects streamflow is clearly highlighted. Next to this, changes in the Flow Duration Curve (FDC) between the future and reference period are compared for all catchments. In literature, the FDC is a commonly used indicator for assessing the effect of climate change on streamflow (Leong and Yokoo, 2021, Leta et al., 2018, Nel et al., 2014, Sellami et al., 2016). By looking at the shapes of the FDCs of the reference and future streamflow simulations, changes in the frequency of high and low streamflow flows can be assessed. The analysis is done by using the 7-day moving average of the streamflow simulations. In this way, any potential false signals or occasional fluctuations are smoothed out of the data.

2.5.2. Changes in minimum and maximum annual streamflow

Changes in annual streamflow extremes in all catchments are examined by looking at the magnitude and timing of the annual 7-day minimum and maximum streamflow. This is done by making boxplots of the mean streamflow magnitude during these 7 days of maximum and minimum flow. Next to this, changes in extreme streamflow timing are investigated by using circular statistics (e.g. Hanus et al. (2021), Blöschl

et al. (2017)). In this way, shifts in the timing of the minimum and maximum streamflow can be correctly identified, despite of the turnings of years.

2.5.3. Uncertainty analysis

To check for the uncertainties originating from the different parameter sets, the spread of the future streamflow projections from all parameter sets in the Conasauga catchment is analysed, which is done by using the GFDL-CM4 model for the SSP585 scenario. Next to this, to further analyse the uncertainty coming from the climate models, the results of the two selected CMIP6 models are compared to the total spread of the results of an ensemble of 28 different CMIP6 models, using one parameter set only. Lastly, the difference in spread originating from the different parameter sets and the climate models are compared using boxplots showing the spread of the mean monthly streamflow of all individual years in the future period.

2.6. Reproducible modeling approach

The modeling process is designed such that it is easily and quickly reproducible for all selected catchments. As mentioned earlier in sections 2.2 and 2.3, this is particularly done by adding a BMI to the HBV-mountain hydrological model and using the ESMValTool for generating forcing input. The BMI is an interface which provides a standardized set of functions, allowing users to reuse the model in a straightforward way (E. Hutton et al., 2020). Next to this, the general modeling process is designed in a way that the climate change impact assessment can be done using three different Jupyter Notebooks: one notebook for generating forcing input using ESMValTool, one notebook for model calibration and validation, and one notebook for the analysis. A detailed description of functionalities of each notebook can be found below. The total modeling approach is schematised in figure 2.4. All model code is available from <https://doi.org/10.5281/zenodo.7405678> (Hoogelander, 2022).

Notebook 1: Generating forcing using ESMValTool

In this notebook, forcing input for the HBV-mountain model can be generated using the ESMValTool. Important components of the ESMValTool are the ESMValTool recipe and the diagnostics script: the recipe contains instructions which are needed to generate the desired forcing data. The main instructions mainly include the directory and name of the basin shapefile, the climate dataset from which data is to be retrieved, the desired variables, the start and end year and which diagnostics script should be used. The diagnostics script is a Python script containing the several functions which are used to obtain the desired data in the manner specified in the recipe. Currently, this notebook is designed to generate forcing for this specific model. Before forcing for a catchment can be generated, the catchment shapefile must be stored in the auxiliary data folder (see appendix B.2 for more details). In this notebook, CMIP6 forcing generation is generally done in three steps: firstly, the ESMValTool recipe for the HBV-mountain model is loaded. Currently, the ESMValTool recipe and diagnostics script are not yet available through the ESMValTool's repository, but they can be found on this study's Github repository (see Appendix A.2). Then, a dictionary of all specified CMIP6 models is made, which is done by selecting the CMIP6 models listed in a YAML file. Currently, there are 28 different CMIP6 models listed in this file which can be selected from. Although more CMIP6 models are available on Levante, not all models have precipitation data available on a daily resolution. Besides, not all models have multiple future SSP scenarios included. From all listed CMIP6 models, all models have at least the SSP245 and SSP585 scenario included in their data (see Appendix B.4). The last step is to run the ESMValTool recipe to generate forcing input for the HBV-mountain model. This is done using a function which requires the ESMValTool recipe object, the catchment name or ID, the scenario, the dictionary containing the selected CMIP6 datasets and the start and end year. After running this function, the precipitation and temperature data are stored in a NetCDF file within an ESMValTool output folder, from which the exact location is given after a successful run. In case of generating ERA5 forcing input, the ESMValTool recipe is slightly different, and running this recipe can be done by only specifying the catchment name and start and end year.

Notebook 2: Model calibration and validation

Calibration and validation of the hydrological model are done in the second notebook. Currently, in case all necessary packages are installed, this notebook can be used on a local machine. This notebook is built on the (Julia) HBV-mountain model functions, a wrapped Python BMI, and additional Python preprocessing and calibration functions (See Appendix B.2 for more details). This notebook is used in four steps: first, the ERA5 forcing NetCDF file (generated in the first notebook) is loaded. Then, the second step is to provide all

necessary settings for the calibration. Regarding the required additional data, this includes specifications of where the NLCD landuse map, DEM map, streamflow observations and the catchment's shapefile are stored. For this, it is important that all raster files and the shapefile are projected in the WGS84 coordinate system. Next to this, the amount of CMA-ES calibrations, the iterations per CMA-ES calibration, the start and end year and the number of years for the warming up period are specified in this second step. Then, after specifying all necessary settings, the model is calibrated using a wrapped calibration function, which returns a dataframe containing the calibration scores of the individual objective functions together with the corresponding parameter set. Then, this dataframe is used for the validation, which is again done using a wrapped function. The output of the validation function are the validation results and the simulated streamflow and evaporation over the total calibration and validation period.

Notebook 3: CMIP6 simulations and analysis

In the third notebook, all historical and future streamflow simulations using CMIP6 forcing are produced and the results are analysed. This is done in three main steps: first, the calibration and validation results from the second notebook are loaded, which are in the form of a dataframe containing the calibration and validation scores together with the corresponding parameter set. Additionally, the streamflow and evaporation simulations are loaded so that they can be included in the analysis. Then, in the second step, all historical and future streamflow simulations are made using the CMIP6 data. For this, the locations of the NetCDF files of the CMIP6 forcing and the ERA5 calibration forcing must be given, which are generated using the first notebook. Another option is to loop through all generated CMIP6 models, for which the folders in which the models are located must be specified. In the streamflow simulation runs, all CMIP6 data are bias corrected with the ERA5 data. After running the simulations, the last step is to perform the analysis using the results. In this step, all results discussed in section 2.5 from this report can be reproduced. Although, users are free in making their own analysis after making the streamflow simulations in this notebook.

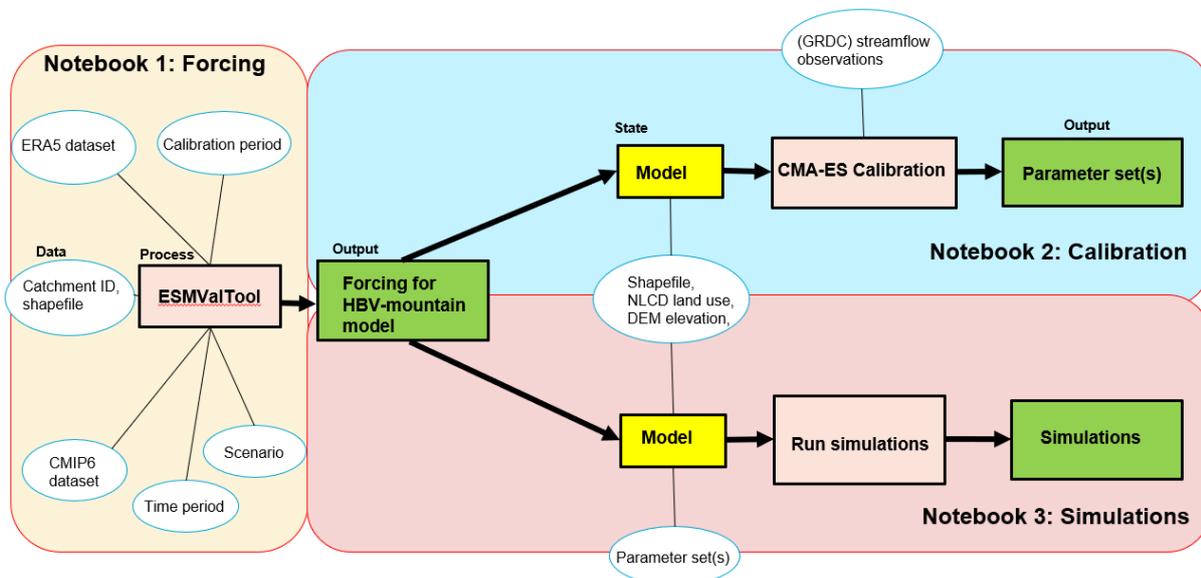


Figure 2.4: Modeling approach

Results

This chapter presents the results found using the methods discussed in Chapter 2. The calibration and validation results are presented in section 3.1. Section 3.2 shows the results of the historical streamflow simulations using the CMIP6 data, and section 3.3 shows the results of the climate change impact analysis. All results are made using the reproducible modeling approach described in section 2.6.

3.1. Model calibration and validation

After running 50 different CMA-ES calibration processes with 50 iterations per catchment, the hydrological model has a Obj_{tot} score between 0.42 - 0.69 and 0.40 - 0.67 for the calibration and validation period respectively, using the 20 best scoring parameter sets in all catchments. The best validation scores are found in the Thunder Creek catchment (0.67), while the Kawishiwi catchment has the lowest validation score (0.40). Mean scores of all individual objectives can be found in table 3.1. The low Obj_{tot} score in the Kawishiwi catchment is mostly due to the low $NSE_{log\ timeseries}$ and $NSE_{annual\ runoff}$ scores. The Big Rock Creek, Conasauga and Youghiogheny catchments show low $NSE_{annual\ runoff}$ as well. The low value in the Big Rock Creek catchment is related to the relatively low score in the $NSE_{timeseries}$, indicating that high flows are poorly captured here. The low $NSE_{annual\ runoff}$ value in the Kawishiwi catchment is mainly due to the relatively low score in both $NSE_{timeseries}$ and $NSE_{log\ timeseries}$. The $NSE_{timeseries}$ score varies among the catchments, with the Thunder Creek and Big Rock Creek catchment showing the highest and lowest score respectively. On the other hand, the Big Rock Creek catchment shows the highest $NSE_{log\ timeseries}$, indicating that the model is good at capturing low flows here. The $NSE_{flow\ duration\ curve}$ scores are relatively high in all catchments, indicating that the model is good in reproducing the streamflow frequency distribution in all catchments.

The simulation results of all catchments for one observational year (2001) together with the streamflow observations can be found in figure 3.1. This figure highlights the model spread of the hydrological model originating from the individual parameter sets in one individual year. The largest relative spread can be found in the Big Rock Creek and Kawishiwi catchment. Although the different climate classifications, both the Conasauga and Youghiogheny catchments have comparable hydrographs, which is likely to be related to their comparable land use proportions and relatively close geographical locations. Next to this, the ERA5-simulation results of all catchments are compared with the observation data by looking at the long term water balance. Figure 3.2 shows cumulative sum of the biases per catchment over the years together with the cumulative biases relative to the cumulative streamflow, and the sum of the absolute biases over the years. The largest relative bias in the long term water balance is found in the Big Rock Creek catchment, while the Kawishiwi catchment has the lowest relative bias in the long term water balance. However, relative to the cumulative streamflow, the absolute error in the Kawishiwi is still high. Total biases in the Big Rock Creek, Thunder Creek and Youghiogheny catchment are all negative, which indicates, according to equation 2.3, that the streamflow is generally underestimated here. Although the Thunder Creek shows good calibration results, the total bias is relatively large compared to other catchments. As the bias is negative here, there is likely a factor underestimating the flow input in equation 2.3. This underestimating factor is related to the neglect of the change in storage in this particular catchment, as glacier melt has a large contributing factor here, indicating that the results of figure 3.2 are not reliable for the Thunder Creek catchment. The Kawishiwi and Conasauga catchments are having a positive bias in the long term water balance, which suggests that streamflow is generally overestimated. Although, as the cumulative change is calculated between two winters, the bias is also affected by neglecting the change in storage, as the

difference in snow storage is likely to be different between the two winter periods. Where the biases in the long term total water balance in most catchments grow irregularly, the cumulative bias in the Thunder Creek catchment increases almost linearly. This indicates that the bias in this catchment is dependent on a constant factor over the years, which is most likely due to the glacier melt processes.

Table 3.1: Calibration and Validation results

	Big Rock Creek		Thunder Creek		Conasauga		Kawishiwi		Youghiogheny	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
$NSE_{timeseries}$	0.31	0.25	0.61	0.58	0.41	0.48	0.54	0.47	0.50	0.44
$NSE_{log\ timeseries}$	0.80	0.80	0.76	0.78	0.70	0.66	0.48	0.32	0.77	0.73
$NSE_{flow\ duration\ curve}$	0.90	0.88	0.92	0.94	0.96	0.94	0.93	0.95	0.97	0.97
$NSE_{annual\ runoff}$	0.09	0.12	0.60	0.67	0.15	0.39	0.43	0.17	0.58	0.26
Obj_{tot}	0.42	0.41	0.69	0.67	0.46	0.56	0.55	0.40	0.65	0.52

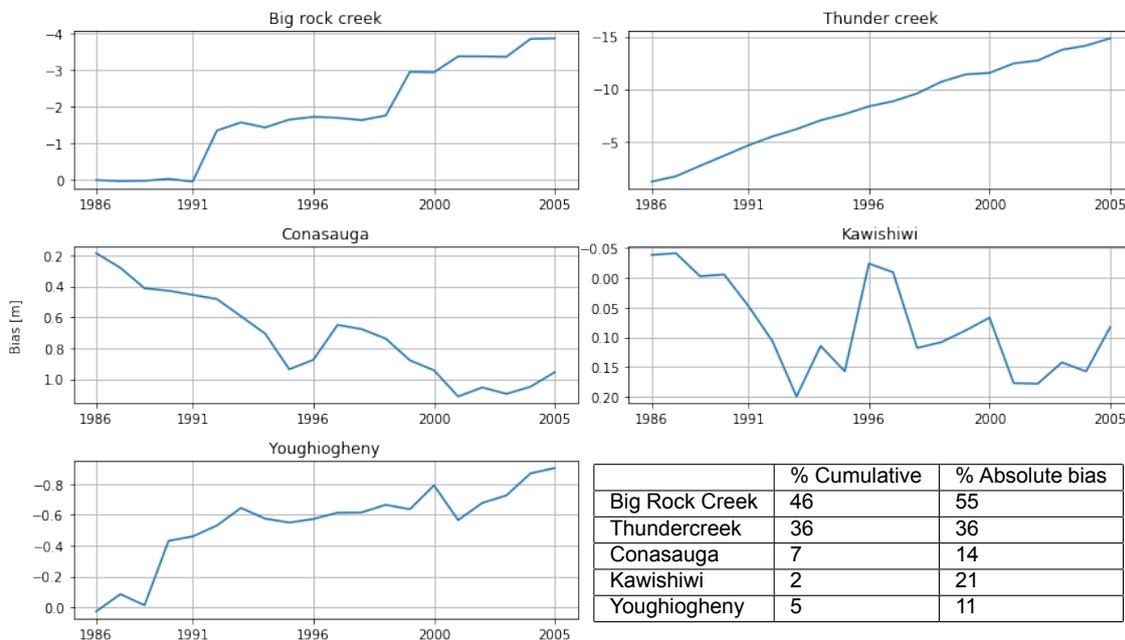


Figure 3.2: Cumulative bias in the long term water balance over the calibration period (equation 2.3)

3.2. Historical streamflow simulations

The 20 best performing parameter sets per catchment are used for making the historical and future climate simulations using the CMIP6 climate models. Results of the annual daily mean streamflow simulations using the raw and bias-corrected CMIP6 models, together with ERA5 simulations in the reference period for the Big Rock Creek and Thunder Creek catchment, can be found in figure 3.3. This figure highlights the effect of the linear scaling of the CMIP6 forcing data, and how this differs between the two CMIP6 models. Both climate models show, without the bias correction, significant differences in the historical climate simulations. In the Big Rock Creek catchment, the MPI-ESM1-2-HR shows the most significant difference between the simulations using raw or bias-corrected data, while in the Thunder Creek catchment this is the case for the GFDL-CM4 model. Whereas in the Big Rock Creek catchment the bias correction mainly affects mean flow magnitude, it mainly influences the timing in the Thunder Creek catchment.

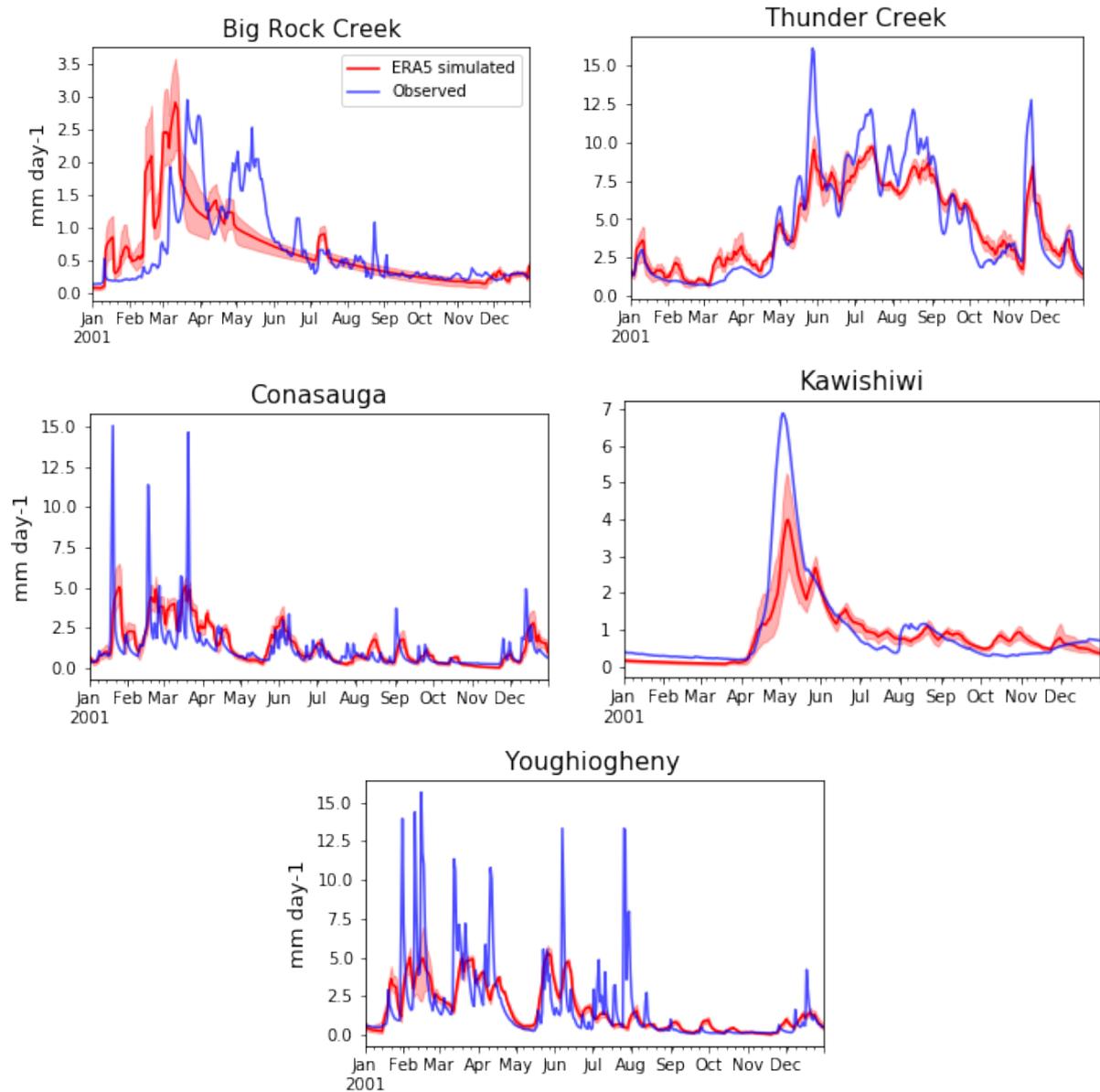


Figure 3.1: Streamflow simulations (with ERA5 forcing data) and observations in the selected catchments in 2001

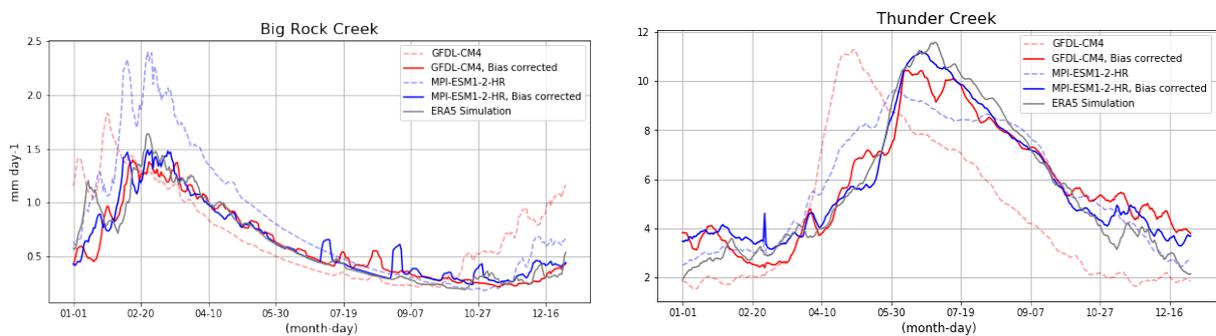


Figure 3.3: Mean daily simulated streamflow in the reference period (1989-2005), using the ERA5 calibration data and two CMIP6 climate models

3.3. Climate change impact analysis

3.3.1. Annual, seasonal and monthly changes

Figure 3.4 shows boxplots of the monthly spread of the modelled total monthly streamflow of all years in the reference (1975-2005) and future (2070-2100) period per catchment, using the mean of the GFDL-CM4 and MPI-ESM1-2-HR simulations. Significant changes in streamflow patterns are projected in the Thunder Creek catchment, where low winter streamflow is expected to significantly increase under both climate scenarios. In the Big Rock Creek catchment, the projections show a large decrease in streamflow during the dry season (May-Nov) under both climate scenarios. The Conasauga and Youghiogheny catchments show similar changes in streamflow, where a increase in streamflow is projected during winter (Dec-Feb) and a decrease in streamflow during the rest of the year in both scenarios. In the Kawishiwi catchment, streamflow is expected to decrease in almost all months in both scenarios, except in February and March.

While future temperature increase is significantly higher in SSP585 than in the SSP245 scenario in all catchments, the severity of precipitation changes differs between the two climate scenarios: in some cases, SSP245 scenario predicts a more extreme increase or decrease in precipitation compared to SSP585. Also, both climate scenarios show inverse predictions in some cases. As stated, the SSP585 simulations show similar but more extreme changes in streamflow compared to the SSP245 simulations in most cases. However, similar to future precipitation predictions, some catchments show inconsistencies between the scenarios in some particular months. For example: in the Conasauga catchment, the SSP245 projection shows a decrease in streamflow in January - March and September - November, while the SSP585 projection shows an increase. In January - March, this difference is also found in the projected precipitation change for both scenarios.

Inconsistencies between change in precipitation and streamflow are commonly found in the results. In the Big Rock Creek catchment for example, precipitation is expected to increase by about 50% in June and July in the SSP245 scenario, while streamflow is expected to decrease. Figure 3.5 makes more visible how the climatic changes under the SSP245 scenarios are related to the streamflow projections in the Big Rock Creek and Thunder Creek catchment. Both catchments have contrasting projections: in the Big Rock Creek Catchment, the increase in temperature correlates to a decrease in streamflow. While there is a significant increase in precipitation in the summer months, this does not result in an increase in streamflow during this time. On the other hand, increasing temperature in the Thunder Creek catchment is accompanied by an increase in streamflow, and the increase in streamflow is larger in proportion to the change in precipitation.

Changes in the FDC per catchment are illustrated in figure 3.6, where the exceedance probability is displayed at a logarithmic scale. A large downward shift in the FDC is projected in the Kawishiwi catchment, while the other catchments show an upward shift in general. The FDC in the Kawishiwi catchment suggests a significant decrease in high flows here, which is most visible under the SSP585 scenario. In both the Big Rock Creek and Youghiogheny catchment, the SSP245 simulations show a larger increase in high flow exceedance. In the Thunder Creek catchment, the difference of change in high flow exceedance between both climate scenarios is less clear. However, the steepening of the FDC slope in the Thunder Creek catchment in the future climate simulations at the high exceedance probability area indicates a decrease in frequency of low flows here, which is more visible in the SSP585 scenario. In the Big Rock Creek catchment, the SSP585 simulations show a decrease in the slope at the high probability area between 0.2 - 0.8, while a change is less clear under the SSP245 scenario.

Figure 3.7 demonstrates the relative seasonal streamflow changes from the GFDL-CM4 and MPI-ESM1-2-HR simulations in the Conasauga catchment separately, for both the SSP245 and SSP585 scenario. The figure highlights the large differences in the projections between the simulations with the individual climate models. Inconsistencies in the streamflow projections in Sep - Nov of the Conasauga catchment are also found in figure 3.7: in the SSP245 scenario, the streamflow projections with both climate models show an increase in streamflow in these months. However, in the SSP585 scenario, the projection using the GFDL-CM4 model shows a very large increase in streamflow, while the projection with the MPI-ESM1-2-HR model shows a small decrease.

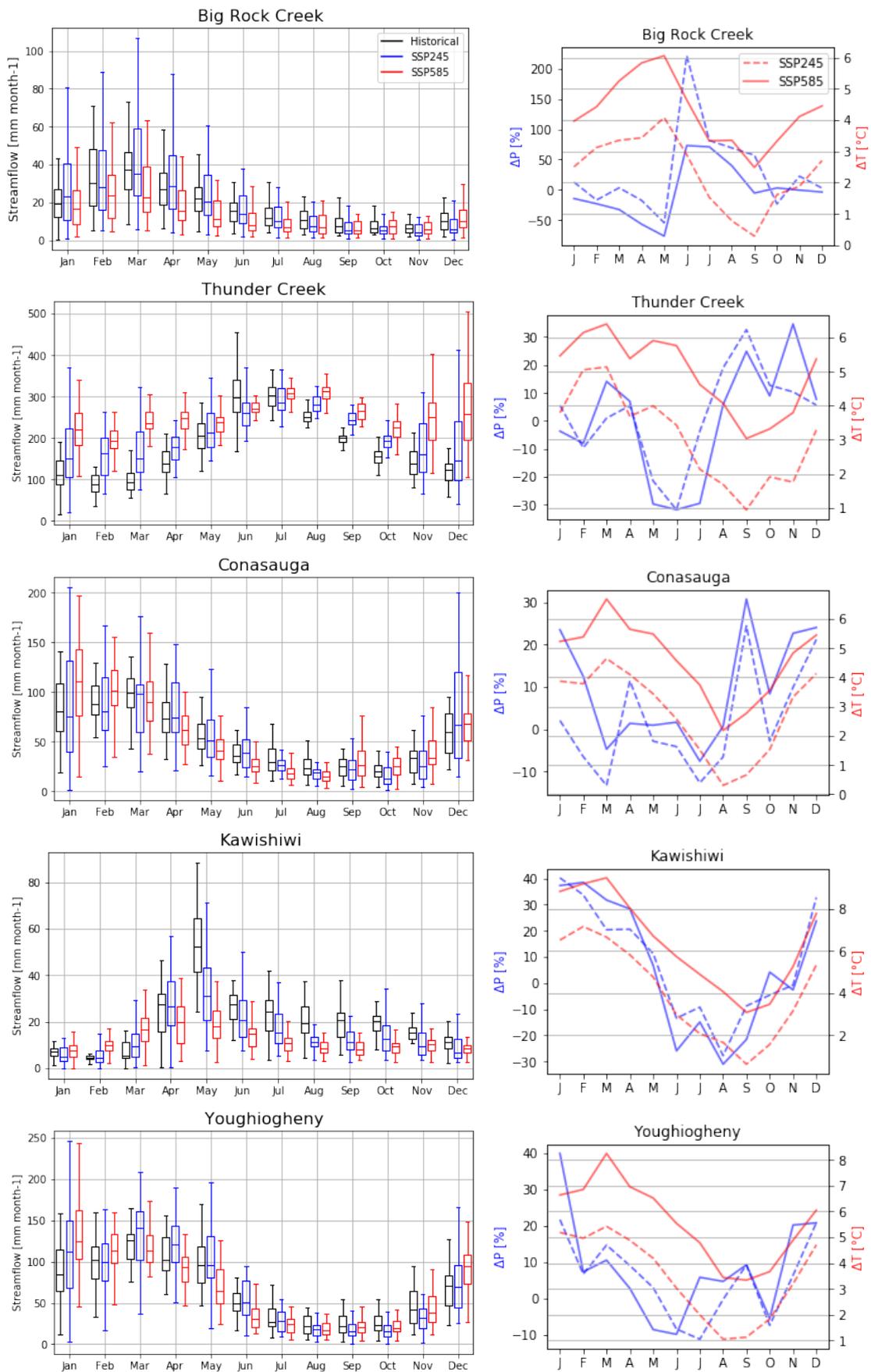


Figure 3.4: Streamflow projections in the reference and future period for two the SSP245 (blue) and SSP585 (red) climate scenarios, using the mean simulation of the two CMIP6 climate models (left), together with the mean relative changes in precipitation (blue) and temperature (red) per catchment (right)

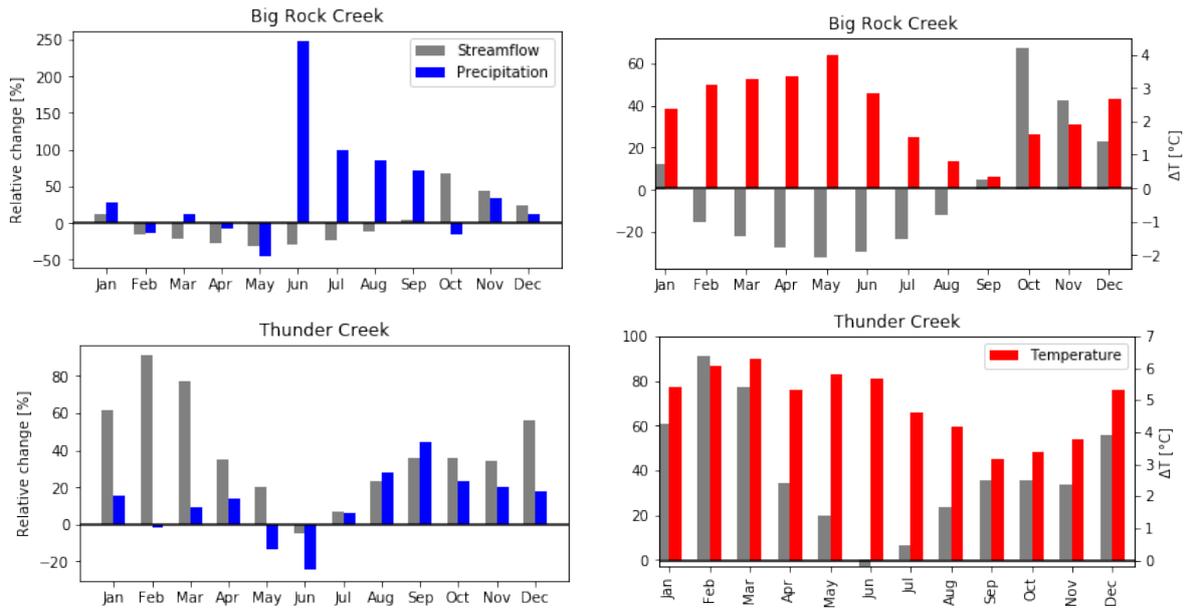


Figure 3.5: Relative changes in streamflow compared to relative changes of precipitation and absolute change in temperature in the Big Rock Creek and Thunder Creek catchment under the SSP245 scenario

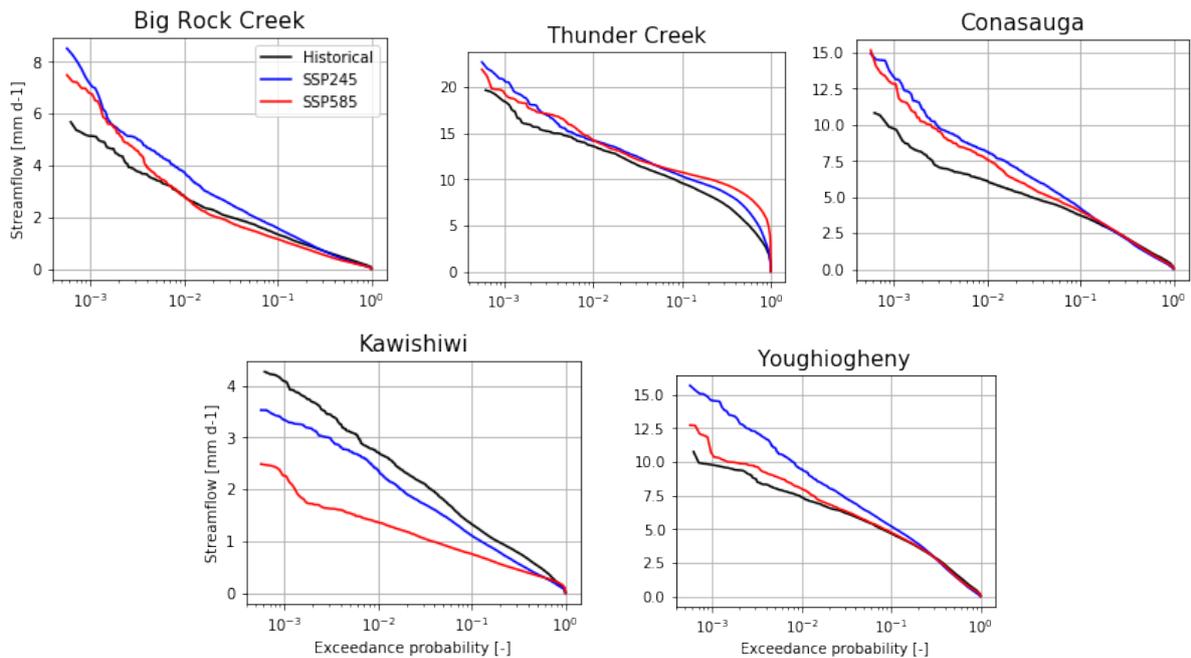


Figure 3.6: Changes in the flow duration curves under both climate scenarios, exceedance probability displayed at logarithmic scale

3.3.2. Changes in minimum and maximum annual streamflow

Changes in 7-days minimum streamflow magnitude under both climate scenarios are presented in figure 3.8. Where in the Big Rock Creek, Conasauga and Youghiogheny catchment both climate scenarios project a decrease in annual low flow, the SSP245 simulations show a larger decrease (36% - 63%) in the median of annual minimum flow than the SSP585 scenario (16% - 25%) in these catchments. In the relatively dry

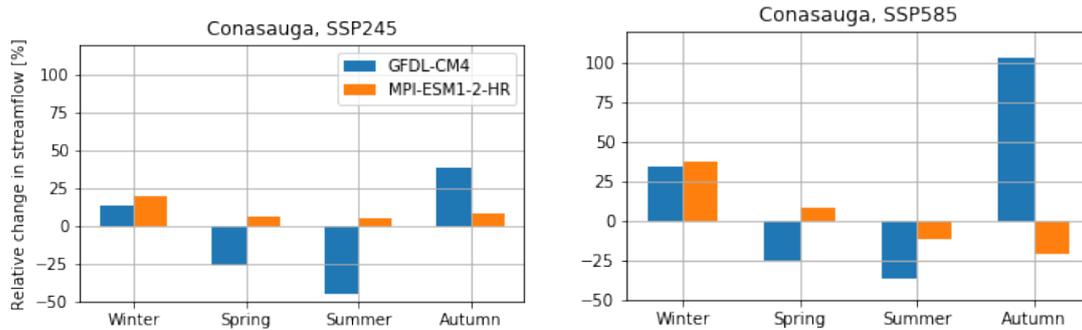


Figure 3.7: Seasonal streamflow changes in the Conasauga catchment between the GFDL-CM4 and MPI-ESM1-2-HR simulations under both climate scenarios

Big Rock Creek catchment, the median of annual low flows is projected to decrease by 36% and 22% in the SSP245 and SSP585 scenario respectively. Both climate scenarios have opposite predictions in the Thunder Creek and Kawishiwi catchment, as the SSP245 simulations show a small decrease in the median of annual low flow magnitude, while the SSP585 simulations predict an increase. Looking at the changes in precipitation and temperature in these catchments, this is most likely caused by the different projected temperature changes: the lower rise in temperature in the SSP245 scenario results in less snow melt in winter compared to the SSP585 scenario, which is the period with low flow in both catchments. However, the spread of annual minimum flow is significantly higher under the SSP245 scenario compared to the reference, indicating that annual minimum flow will vary significantly. The median of the low flows in the Thunder Creek in the SSP585 simulations catchment will increase significantly compared to the reference period (+176%), which follows the high streamflow increases found in figure 3.4.

The projected changes in 7-days maximum streamflow of figure 3.9 show that maximum flow median will increase in the Conasauga and Youghiogheny catchment under both climate scenarios, with SSP245 simulations showing a more extreme change (+26% - +27%), while in the Big Rock Creek and Thunder Creek catchment this increase is only predicted in the SSP245 scenario (+8% - +27%). In case of the Big Rock Creek catchment, this might be caused by disagreements in projected precipitation between the scenarios: the SSP245 scenario shows a small increase in precipitation in the wettest month (March), while SSP585 projects a decrease. Although the Thunder Creek catchment has a relatively small difference between the scenarios, this might be caused by more snow accumulation in winter under SSP245 compared to SSP585, in combination with faster melting processes in spring under SSP245 compared to the reference, resulting to increasing high flows. In the Kawishiwi catchment, the median of the annual maximum flow is projected to decrease significantly in both climate scenarios (31% - 56%), which is most likely due to a reduction of snow melt.

Next to changes in magnitude of minimum and maximum annual flow, in most cases, both extremes are predicted to shift in time according to the simulations of both climate scenarios. Table 3.2 presents the mean and standard deviation of the minimum and maximum streamflow occurrence over the 30-year period of the reference and future period, and figure 3.10 presents the corresponding mean shifts of the annual flow extremes. The largest shift in annual minimum flow is projected in the Kawishiwi catchment (-30 days in SSP245, -68 days in SSP585). However, this large shift goes along with a significant change in the standard deviation of mean day of minimum flow occurrence in this catchment (+18 days in SSP245, +55 days in SSP585), suggesting that the timing of low flows becomes more unpredictable here. In the Big Rock Creek, Conasauga and Youghiogheny catchment, both SSP scenarios have opposite predictions in shift of minimum flow occurrence: where the SSP245 scenario reveals a later shift, the SSP585 simulations show a earlier shift. The Thunder Creek catchment experiences the largest shift in annual maximum flow (-199 days in SSP245, -175 days in SSP585), accompanying with a significant increase in the standard deviation (+101 days in SSP245, +66 days in SSP585), suggesting an increase in unpredictability and less seasonality in maximum flow occurrence. In most cases, both SSP scenarios have similar predictions in terms of timing of annual maximum flow, although changes under the SSP585 scenario are more extreme.

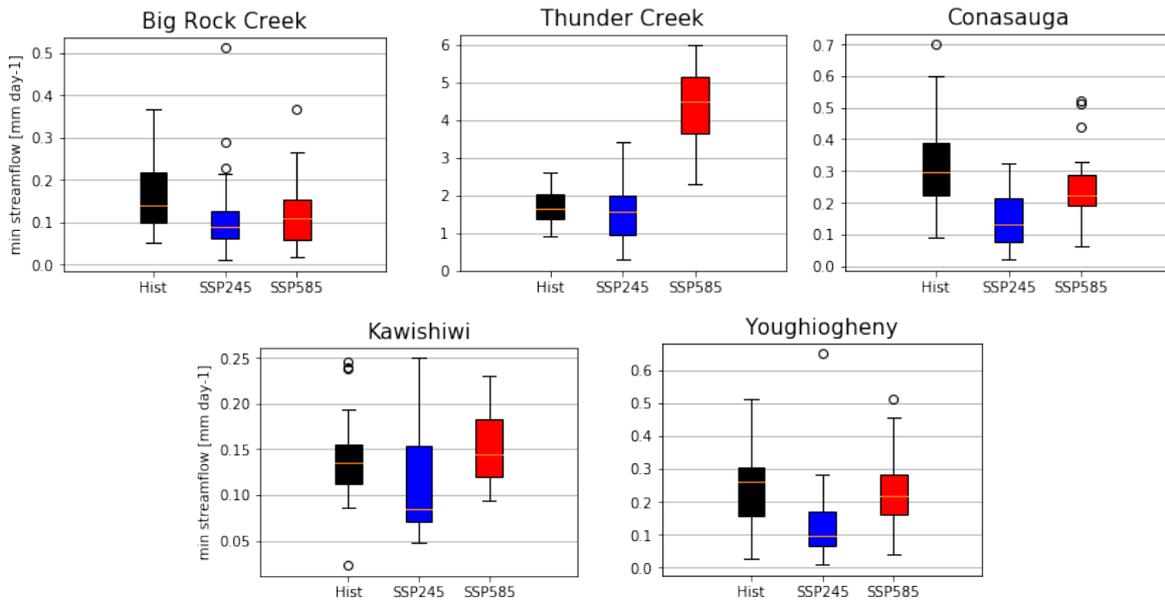


Figure 3.8: Changes in 7-days annual minimum streamflow magnitudes

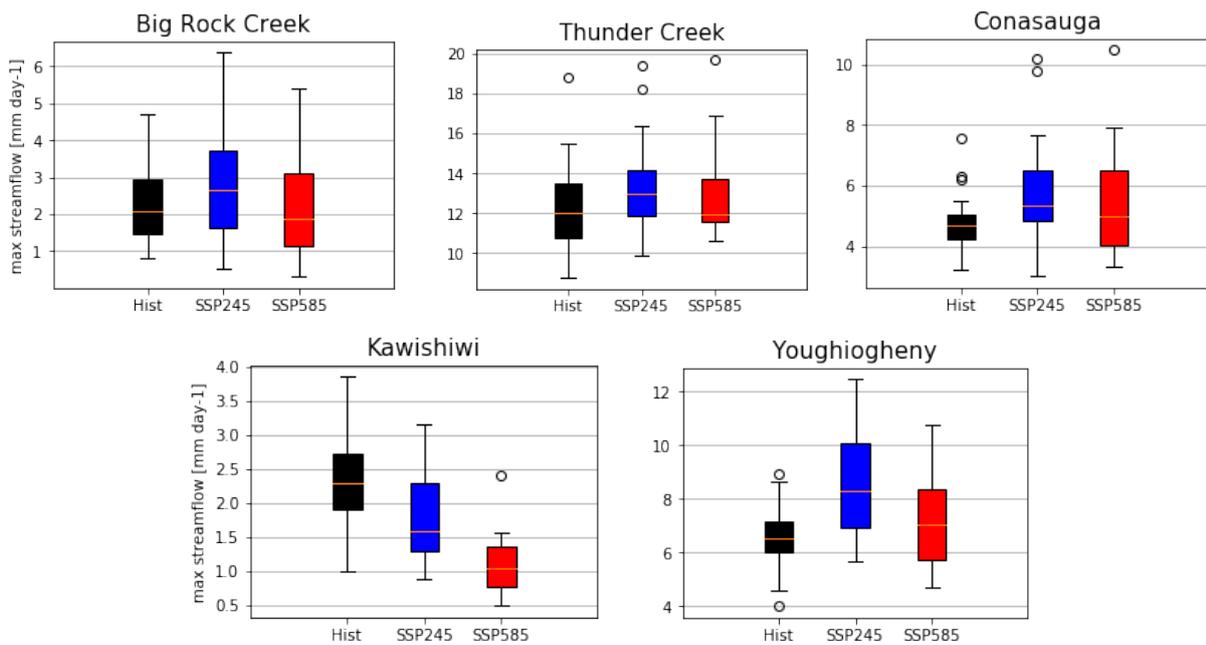


Figure 3.9: Changes in 7-days annual maximum streamflow magnitudes

Table 3.2: Mean and standard deviation of the day of occurrence of annual maximum and minimum flow in the reference (1975-2005) and future (2070-2100) period

	Reference				SSP245		SSP585	
	doy _{minflow}	Date	doy _{maxflow}	Date	doy _{minflow}	doy _{maxflow}	doy _{minflow}	doy _{maxflow}
Big Rock Creek	322 ±44	18 Nov	54 ±45	23 Feb	342 ±30	43 ±40	297 ±59	40 ±66
Thunder Creek	33 ±34	3 Feb	168 ±14	17 Jun	6 ±33	49 ±115	6 ±45	358 ±80
Conasauga	276 ±36	3 Oct	58 ±35	27 Feb	283 ±42	46 ±49	244 ±41	20 ±32
Kawishiwi	64 ±13	5 Mar	126 ±7	6 May	35 ±31	126 ±37	361 ±68	105 ±35
Youghioghenny	257 ±45	14 Sep	64 ±46	5 Mar	271 ±27	67 ±46	243 ±38	26 ±48

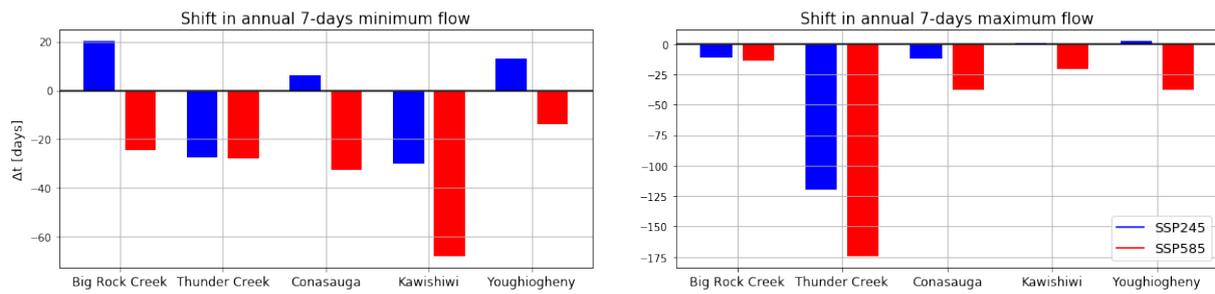


Figure 3.10: Shift in day of annual 7-days annual minimum and maximum flow occurrence under both SSP scenarios

3.3.3. Conclusions and linkages of hydrological changes per catchment

To give an overview, this section summarizes the projected hydrological changes due to climate change per catchment, and draws links and conclusions based on the these.

In the **Big Rock Creek** catchment, the streamflow is projected to generally decrease in most months under both SSP scenarios. Even in the dry months, a significant streamflow decrease is projected, indicating an increasing frequency of streamflow droughts in this area. From January - May, the streamflow decrease can be linked to a decrease in precipitation. Although, while streamflow decreases, precipitation during the dry summer months is projected to increase, which is clearly visible in 3.5. This indicates that the effect of changing evaporation processes or changing precipitation intensities is significantly larger then. However streamflow is generally decreasing in this catchment, more high extremes streamflow levels are projected in the flow duration curve (figure 3.6).

Results suggests that the **Thunder Creek** catchment will experience very significant streamflow changes in the future: streamflow is expected to increase in all months, except for June and July. The increasing streamflow median in most months may be caused by increasing melt processes due to rising temperatures. On the other hand, the decreasing streamflow median in June and July may be due to decreasing snow melt, as less snow will be accumulated during winter months. Next to this, a significant decrease in precipitation is projected in these months, while the streamflow increase from August - October is mainly due to increasing precipitation (figure 3.5). The streamflow increase over the year is also translated to a significant increase in annual maximum flow under SSP585. Surprisingly, under SSP245 this increase is not projected. The timing of maximum flow will shift from June to winter, with a shift of about -123 days in SSP245, to -175 days in SSP585. In addition to extra glacier melt in winter due to rising temperature, more precipitation will directly runoff instead of being stored as snow, resulting in increasing streamflow in winter.

The future simulations of both climate scenarios seem to show many disagreements between the predictions in the **Conasauga** catchment. In some months, these disagreements are caused by disagreements in the precipitation. Figure 3.7 showed that the disagreements were also caused by disagreements between the two climate models, as the signals of streamflow decrease in spring and summer are relatively stronger with GFDL-CM4 than the streamflow increase with MPI-ESM1-2-HR. From March - August, the SSP585 scenario predicts no large change in precipitation, while the streamflow decreases significantly in these months. Possible processes causing this decrease are the decrease of melt processes of snow from winter, and increasing evaporation rates due to rising temperatures. Although SSP245 predicts more change in precipitation, this is not translated in a significant change in streamflow. The FDC suggests that high flows are expected to be more frequent in both climate scenarios. Next to this, annual minimum flow is expected to decrease, while annual maximum flow is expected to increase.

The **Kawishiwi** catchment is expected to see an increase in streamflow from January - February, while streamflow is projected to significantly decrease for the rest of the year. Most importantly, given that the catchment has a relatively cold land climate, the increase of streamflow in winter is likely caused by the significant rising temperatures, resulting in more direct precipitation runoff and less snow storage. Consequently, there will be less melt water in spring and summer, resulting in a significant decrease in streamflow during that time. In addition, precipitation is predicted to significantly decrease from June - September. These processes also translate into a significant projected decrease in the annual maximum streamflow.

Lastly, the **Youghiogheny** catchment show many similarities with the Conasauga catchment in the seasonality. Although they are located in different climate zones, geographically they are close. Nonetheless, the Youghiogheny catchment show less disagreements between the climate scenarios. Streamflow is expected to increase from December - February in SSP585, and until April in the SSP245 scenario. This increase is caused by the projected increase in precipitation and likely be caused by less snow storage during these months. Streamflow is expected to decrease in May and June under both scenarios, which is caused by a decrease in precipitation. Although streamflow will generally slightly decrease from July - October, these changes are not significant. Looking at the expected changes in forcing during that time, this slight decrease is likely caused by rising evaporation rates. The simulations in the SSP245 scenario show a larger spread over the year compared to SSP585. This is also visible in the FDC, where higher flows will occur more common in SSP245 simulations. Next to this, annual maximum and minimum flow is expected to change more extreme in the SSP245 scenario.

3.3.4. Uncertainty analysis

The uncertainty is quantified based on the spread of the results between all parameter sets using one CMIP6 model, and a large ensemble of CMIP6 models using one parameter set. Figure 3.11 presents the spread originating from the different parameter sets (left) and the climate models (right) in the Conasauga catchment under the SSP585 scenario. In these figures can be seen that the spread originating from the climate models is significantly larger than the spread originating from the different parameter sets of the hydrological model. Where the differences between the parameter sets are mainly visible in the magnitude, the simulations with the different climate models show large differences in both timing and magnitude.

The simulations using the different climate models show large differences in the timing of annual maximum flow. To see how these differences relates to the projected changes in timing of annual maximum flow in table 3.2, the spread of annual maximum flow occurrence between the simulations using all climate models is shown in figure 3.12. In this figure, also the day of 7-day maximum flow occurrence using the two selected CMIP6 models are indicated (using one parameter set), together with the historical observation. The simulation using GFDL-CM4 appears to be an extreme projection in terms of annual maximum flow timing compared to the total distribution. The MPI-ESM1-2-HR simulation projects a smaller shift compared to most other CMIP6 models. Nevertheless, the boxplot further indicates that annual maximum flow is very likely to occur earlier in the year in the Conasauga catchment.

Figure 3.12 shows the boxplots of the spread of the simulated total monthly streamflow of all individual years in the future (2070 - 2100) under the SSP585 scenario, using the individual simulations of GFDL-CM4 dataset with the 20 individual parameter sets (blue), and the 28 CMIP6 datasets using one parameter set (black). The large spread of the different results are not directly translated into a much larger spread in streamflow in the individual months: from February - August, the spread of the different climate models is larger than the spread of the individual parameter sets, but in the rest of the year, the spread of the parameter sets is larger.

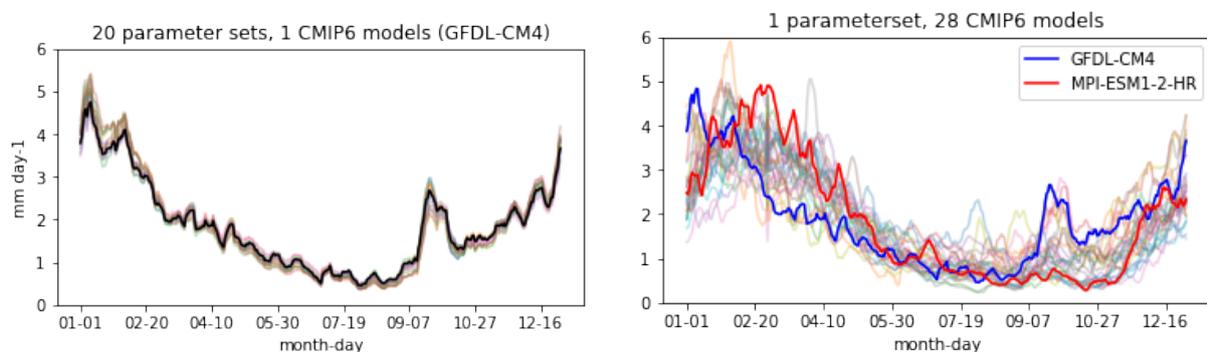


Figure 3.11: Mean daily simulated streamflow in the SSP585 scenario (2070-2099) using 20 parameter sets and the GFDL-CM4 model (left), and using 28 CMIP6 models with one parameter set (right). The black line in the left figure is the result of the selected parameter set used in the comparison between the 28 CMIP6 simulations.

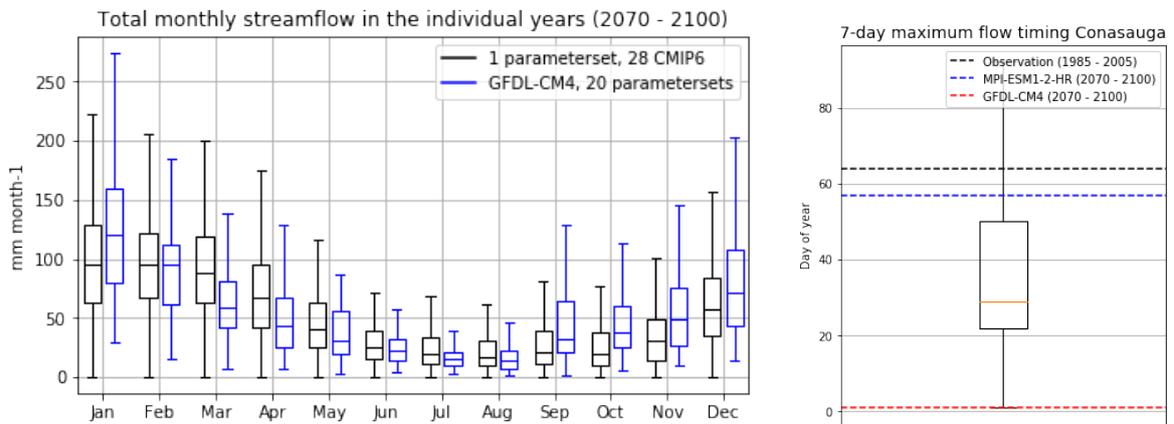


Figure 3.12: Left: boxplots of the total monthly streamflow of all individual years in the future (2070 - 2100) under the SSP585 scenario, using the individual simulations of GFDL-CM4 dataset with the 20 individual parameter sets, and the 28 CMIP6 datasets using one parameter set. Right: spread of annual maximum flow occurrence between the simulations using all climate models in (SSP585), together with the two selected climate models and the observation in the reference period.

4

Discussion

This chapter discusses the results found in chapter 3. First, Section 4.1 focuses on model calibration and validation and the historical streamflow simulations, Section 4.2 discusses the results of the climate change impact analysis, and section 4.3 focuses on the used modeling approach.

4.1. Historical streamflow simulations

4.1.1. Calibration and validation

The hydrological model was calibrated using forcing data from ERA5 and streamflow data from GRDC. These are both open-source datasets, which makes them well applicable in a reproducible modeling approach. However, as the ERA5 dataset is not based solely on observations, its use for model calibration has some drawbacks. Although ERA5 show good capabilities in reproducing forcing data for hydrological modeling purposes in some areas (Bandhauer et al., 2022, Reder and Rianna, 2021), most biases in the dataset compared to observations are linked to precipitation, and there are some significant biases in the data in some areas of the United States (Tarek et al., 2020). For streamflow in catchments that are predominantly influenced by precipitation, it is expected that this can lead to large errors between simulation and observation. Additionally, ERA5 is a gridded dataset, while precipitation events can have a very high local variability, this gives extra uncertainties in streamflow simulations (McMillan et al., 2012). Particularly in small catchments like the Big Rock Creek catchment, this could lead to complications due its high sensitivity to small space and time scales (Fabry et al., 1994, Mendoza et al., 2016). This sensitivity could explain the relatively low score in the $NSE_{\text{timeseries}}$ in this catchment.

This study used the semi-distributed HBV-mountain hydrological model as developed and utilized in the study of Hanus et al. (2021). This model is developed for hydrological analysis in Alpine catchments, and is based on the widely used HBV hydrological model (Seibert and Bergström, 2022, Bergström and Forsman, 1973). The calibration and validation results have lower scores in the total objective function compared to the results found in the Alpine catchments of Hanus et al. (2021). As discussed, one important reason for this difference in performance is the use of ERA5 forcing data rather than direct observation data for calibration. The NSE scores in the Thunder Creek are also lower, even though this is the catchment having the most similar characteristics compared to the Alpine catchments analysed in Hanus et al. (2021). Although the model is designed in such a way that the land use and elevation proportions can be adjusted accordingly to the actual proportions in the catchment, the calibration results suggest that the HBV-mountain model performs better in more mountainous catchments in case of using ERA5 forcing data for calibration. In contrast to the other catchments, the streamflow in the Thunder Creek catchment is predominantly influenced by glacier melt, and thus relatively insensitive to imperfections in the precipitation forcing. Considering the used modeling approach, this hydrological model was well applicable for reproducible use in different catchments because of its relatively low complexity and little required input data.

The model was optimized using the CMA-ES, which resulted in 50 different optimized parameter sets per catchment, from which the 20 best scoring parameter sets were selected for use in the hydrological analysis. The advantage of CMA-ES is that it is a relatively efficient optimization method, as it optimizes within the search space of attraction, which makes this method more efficient than generating random parameter set samples (Poissant et al., 2017). This makes this method suitable in the reproducible modeling approach, as it is possible to generate a well-performing parameter set within a few iterations. However, a trade-off of using this method is its complexity and premature stagnation (Jin et al., 2020), meaning that

the method is unable to search for parameters outside the bounds of convergence. However, performing multiple calibration processes with this strategy tackles the latter limitation to some extent. After calibration using the CMA-ES optimization, it was found that some catchments had individual objective scores with lower NSE scores than the suggested lower limit of a NSE score ($NSE > 0.50$, for streamflow), as reported in the literature (Moriassi et al., 2007). Although, considering the imperfections of the ERA5 forcing data, the catchments have reasonable scores in the total objective function. Lowest Obj_{tot} scores were found in the Big Rock Creek and Kawishiwi catchment, which are the two most contrasting catchments in terms of climate and land use. As stated, one reason for the low score in the Big Rock Creek catchment is its high sensitivity to small space and time scales, what makes ERA5 forcing less suitable here. Nevertheless, the HBV-mountain model captured low flows and flow frequencies relatively well in this catchment. Additionally, it was found that HBV models produces larger errors in streamflow compared to the precipitation input error (Kobold and Brilly, 2006). In case of the Big Rock Creek catchment, which is likely to have a larger precipitation error compared to the other catchment, this has probably led to extra large streamflow errors. As found in section 3.1, the low Obj_{tot} score in the Kawishiwi catchment is mainly due to low $NSE_{log\ timeseries}$ and $NSE_{annual\ runoff}$ scores. Since the Kawishiwi flows through several lakes, the catchment has a large capacity to sustain water, which places emphasis on low flows in particular (Engeland and Hisdal, 2009). Therefore, given the low performance in capturing low flows, a better representation of the lakes in this catchment would give better calibration results. However, in the currently used modeling approach this has not been implemented yet.

All objectives in the multi-objective function were calculated using the Nash Sutcliffe Efficiency. The drawbacks of using the NSE for model calibration and validation this method are widely discussed in the literature (Gupta et al., 2009, Krause et al., 2005, Jain and Sudheer, 2008). One limitation of this method is its sensitivity to large data outliers, although this limitation is partly tackled by using the logarithm of the streamflow as well. This sensitivity is especially high in catchments having a flashy streamflow regime, like Big Rock Creek, as the differences in magnitude between high and low flows on a small timescale are more significant here. The use of the NSE together with the annual runoff as objective exposed the sensitivity of the NSE to a relatively small set of data points. However, the advantage is that it highlighted the performance over a lower time resolution. Within the multi-objective functions, four different objectives were used for the calibration and validation (timeseries, logarithm of timeseries, flow duration curve and annual runoff). However, using more different objectives for optimization could enhance the model's credibility (Efstratiadis and Koutsoyiannis, 2010), this number of objectives reduced the complexity of the CMA-ES calibration while ensuring recognition of the consistency between some signatures.

In case no other evapotranspiration data is provided, the model uses the Thornthwaite method for estimating daily potential evapotranspiration, which was also done in this study. The advantage of using this method is that it requires less data compared other methods of estimating evapotranspiration (Aschonitis et al., 2022), which makes it well suitable for use in a reproducible modeling approach for climate change impact assessment, as the amount of different available climatic data is often limited in the CMIP6 datasets. However, some studies show that the Thornthwaite method generally overestimates potential evapotranspiration in humid climates and underestimates it in dry climates, and therefore the estimates may be unreliable in some cases (Pereira and Pruitt, 2004, Trajkovic and Kolakovic, 2009, Chen et al., 2005, Lu et al., 2005).

In order to estimate the bias between simulated and observed streamflow, the long term water balance between the ERA5 precipitation, simulated evapotranspiration and observed streamflow was examined (equation 2.3). The bias was estimated by the looking at the difference between observed streamflow and the incoming (ERA5 precipitation) and outgoing (evapotranspiration) fluxes in the catchment, while neglecting change in storage. Results showed that the catchments were all having a growing bias over the years, although the direction of growth differed. There are multiple factors affecting the bias and its direction: firstly, as discussed, a major factor influencing the total bias is the ERA5 precipitation. Secondly, the large bias found in the Thunder Creek catchment in combination with a relatively good model performance suggested that neglecting the change in storage cannot be done in all cases. Although change in glacier extent is assumed to be zero, the large bias suggested that glacier melt is a large contributing factor in the water balance. Due to the constant increase in bias, the results can be linked to a decrease in the size of the glacier in reality. Thirdly, given the uncertainty of the Thornthwaite method for estimating potential

evapotranspiration, the estimation of actual evapotranspiration is likely to have a share in the total cumulative biases of the long term water balance as well. In the relatively dry Big Rock Creek catchment for example, the Thornthwaite method likely underestimated the potential evapotranspiration, which made the total bias in this catchment less negative. In contrast, the potential evapotranspiration in the relatively wet Thunder Creek is likely to be overestimated, making the bias more negative here. Lastly, another factor affecting the bias are streamflow observation errors. Given the order of magnitude of the biases, the use of ERA5 precipitation data is likely to contribute most to the total bias in most catchments.

4.1.2. Historical simulations using CMIP6 climate models

Historical streamflow simulations were made using the GFDL-CM4 and MPI-ESM1-2-HR datasets from CMIP6. These simulations were made with the ERA5 calibrated model, and therefore the CMIP6 datasets were bias-corrected to prevent for systematic misleading of the climate simulations due to anomalies between the ERA5 and CMIP6 data. This study used the Linear Scaling as bias correction method. Figure 3.3 highlighted the importance of bias correction of the CMIP6 data, as large differences between simulations using bias-corrected data and simulations using the raw CMIP6 data were observed. The advantage of using the Linear Scaling method for bias correction is that it is a relatively simple approach to correct for daily mean values and variability. Although the method has proven to be a good corrector of mean daily data, a limitation of this method is that it is not able to correct for other statistical properties, such as the frequency of dry days and the intensity of wet days (Teutschbein and Seibert, 2012).

4.2. Climate change impact projections and uncertainties

The changes in streamflow due to climate change were analysed using the simulations of a reference (1975 - 2005) and future (2070 - 2100) period. All simulations were made using the modeling approach described in section 2.6. This section discusses the results found in the climate change impact assessment by using this approach.

The results showed different projections per catchment. Given the differences in validation scores, the uncertainties of the projections of future hydrology vary between the catchments. First, the calibration results in the Big Rock Creek catchment indicated that the model scores poorly in capturing high flows, which was also highlighted in the simulations of 2001 (figure 3.1). Next to this, the large cumulative bias indicated that the streamflow is generally underestimated here, making it likely that future high flows are underestimated as well. Therefore, the predictions in the wet seasons are more unreliable, including the analysis of change in annual maximum flow. The predictions in the dry season are more reliable, as the model was able to capture low flows relatively well in this catchment. The projected changes in high flows are also more uncertain in the other catchments due to their relatively low $NSE_{\text{timeseries}}$ scores, but to a lesser extent than the Big Rock Creek catchment. Based on figure 3.1, the simulations in some catchments seem to miss many peak flows, which indicates that the exact magnitudes of change are very uncertain. Still, the results gives an indication of the direction of change, and an estimate of the order of magnitude of relative change that can be expected in the catchments.

Results showed that both SSPs often agree on the direction of change, although there are some significant differences in the magnitude of change. Also, some cases showed some disagreements in direction of change. These differences were often caused by the large differences in precipitation data, although in some cases, this was caused by differences between temperature projections. This last behaviour is mostly reflected in the Thunder Creek and Kawishiwi catchment, which are relative cold catchments, as changes in snow accumulation and melt processes are most significant here, which is translated to significant changes in streamflow.

To reduce the complexity of the modeling process, it was chosen to assume no future changes in glacier extent in the Thunder Creek. However, many studies show the importance of considering changes of glacier extent in hydrological impact studies (Li et al., 2015, Luo et al., 2013, Frans et al., 2018). Although some relatively easy approaches on modeling future glacier retreat exist (Seibert et al., 2018, Huss et al., 2010), they require additional input data which is often not readily available. To demonstrate the reproducibil-

ity of the modeling approach, a more simplistic approach is chosen. Nevertheless, this induces a higher uncertainty in the future hydrological projections in this catchment. For example, a study on predicting streamflow changes due to glacier melt in the same catchment predicted less contributing of glacier melt to summer flow due to a decreasing glacier extent at the end of the 21st century (Frans et al., 2018). Given this, the predictions in summer flow are likely to be underestimated, as the currently predicted increases only account for a precipitation and snow melt decrease. On the other hand, neglecting the change in glacier extent does give extra insights on the dynamics on streamflow change caused by these two factors (Van Tiel et al., 2018). Another simplification was the neglect of future land use changes in the modeling approach. Because of the different land use proportions included in the hydrological model, which are all having some different parameters, this additionally increases the uncertainty, as land use changes can have a significant effect on hydrology (Jaramillo and Destouni, 2014). However, future land use changes are a large source of uncertainty (McGlynn et al., 2022). This is, in addition to the same reason for neglecting glacier change, another reason why it was not considered in this study.

Figures 3.3 and 3.7 highlighted the differences in historical and future seasonal projections between the results of both climate models. Although the climate model uncertainty is reduced by using the average of the results with multiple climate models, the results indicate the importance of using more (than two) climate models to increase the reliability of the predictions. This is in line with other studies focusing on the climate change impact on streamflow, where it is found that the climate models are the largest contributing factor on uncertainty on streamflow predictions due to disagreements between the models (Kiesel et al., 2019, Melsen et al., 2018 De Niel et al., 2019). Using a large ensemble of climate models alleviates the impact of uncertainty within individual modes, and would thus give a better representation of the overall climate change prediction. Next to this, the relatively low resolution of the ERA5 compared to some catchment sizes possibly resulted in imperfections of the precipitation input during calibration, as discussed in 4.1.1. This uncertainty also applies to the climate models, which have an even lower resolution compared to ERA5 data (2.2). To see how the uncertainty originating from the different parameter sets relates to the uncertainty of the climate models, section 3.3.4 illustrated the results of 20 different parameter sets using one climate model, and the results of 1 parameter set using an ensemble of 28 CMIP6 models. This is also done to give an indication of how the two selected climate models relate to a large ensemble of CMIP6 datasets in one catchment. The results further indicated that the largest uncertainty originates from the climate model choice. This is specifically visible in the projected mean annual 7-day maximum streamflow occurrence in the SSP585 scenario (figure 3.12). The results here showed that the simulations using GFDL-CM4 showed the most extreme shift in the annual 7-day maximum flow of all CMIP6 models compared to the observation, while the MPI-ESM1-2-HR simulations projected a relatively small shift. Since the results of table 3.2 were made based on the average of the simulations with both climate models, the projected change in annual 7-day maximum flow is in the mean space of all CMIP6 projections in this case, but more different climate models are needed for making a more reliable climate change impact assessment.

4.3. Reproducible modeling approach

The first goal of this study was to design a modeling process in a way that it is easily reproducible in other catchments, and using different climate models and scenarios. This section discusses and summarizes the strengths and weaknesses of the developed modeling approach, and how this relates to the results found in the climate change impact assessment. Additionally, some recommendations are given for further improvements.

First of all, the modeling approach is designed in a way that little input data is required, and that only openly accessible datasets are used. For the model set-up, this includes the NLCD land cover map, (SRTM) elevation map and the catchment shapefile. For model calibration and making the streamflow simulations, this includes the (GRDC) streamflow observations and the ERA5 and CMIP6 datasets, for which the precipitation and temperature data are generated using the ESMValTool recipe. A limitation of using the NLCD dataset is that results are currently only reproducible in the United States, although other areas can be included as well by implementing other land use maps (e.g. CORINE Land Cover for Europe <https://land.copernicus.eu/pan-european/corine-land-cover>, last access: 29 November 2022). Considering the forcing generation, a requirement is that both ERA5 and CMIP6 datasets must be available in the

working environment. In the current used modeling approach, ERA5 data is obtained using the ESMVal-Tool recipe on the eWaterCycle platform, and the CMIP6 data is generated on the HPC cluster "Levante" from DKRZ. Thus, exact reproduction of the forcing generation (following the process of notebook 1 in figure 2.4) is currently only possible with a user account on both platforms.

In the used approach, users can flexibly adjust settings by scenario and catchment area. Because of the efficiency of the CMA-ES calibration process, users are able to quickly generate parameter sets for making future streamflow simulations. As can be seen on figure 2.4, reproducing the results for another catchment is only a matter of downloading the necessary GRDC observation data of this catchment and changing the shapefile (which is automatically given along with the GRDC observation data). Producing results with another climate model is a matter of changing the name of the CMIP6 model. This makes it easy to expand this study to a large ensemble of different catchments, CMIP6 models and future climate scenarios by using this approach.

The modeling approach was used in catchments having a drainage area ranging from 59.3 - 657.9 km², which is relatively low compared to other catchments available in the GRDC dataset. Analysing larger catchments using this hydrological model is less reliable with the used approach, due to increasing precipitation heterogeneity at larger catchment scales. This heterogeneity can be implemented in the modeling approach by using different precipitation zones, as shown in the study of Hanus et al. (2021). In case large precipitation zones are used, a time lag function might be needed as well.

The current modeling approach uses a set of four objective functions for model calibration in all selected catchments. To further increase the flexibility in the calibration process, the modeling approach should be adjusted in a way that users are able to choose for a different set of objective functions per catchment. For example: if modelers are mostly interested in analysing changes in peak flows, the peak distribution (Euser et al., 2013) could be additionally used in the multi-objective function, or in case the streamflow in a catchment is highly influenced by baseflow, the base flow index could be included in the calibration. To implement this flexible calibration approach, more objective functions should be added in the modeling approach.

Despite its flexibility and reproducibility, the designed modeling approach has some limitations, as discussed in sections 4.1 and 4.2. First, although ERA5 data is easily and openly accessible, it reduces the reliability of the results. Second, some simplifications of the hydrological processes in the hydrological model puts major limitations on the modeling approach. In case of the analysed catchments, these limitations mainly include disregarding the glacier retreat processes in the Thunder Creek catchment and the lake processes in the Kawishiwi catchment. In order to make the modeling approach more flexible, the modeling process must be further developed so that these processes can be included. In case of the implementation of glacier retreat, this could be in the form of an additional glacier model that gives the glacier area at a certain time step in a form that can be used in the hydrological model for each new time step. Furthermore, it is not possible to account for land use changes in the modeling process. Change in land use caused for example by transition from forest land to grassland due to deforestation, but also by an increase in developed (urban) area. Urban areas are characterized by their low perviousness, which leads to a very different hydrological response compared to natural conditions (Guan et al., 2015). Although urban land use is not present or very negligible in the selected catchments, the hydrological model is not able to reflect this land use. Nevertheless, in case other catchments that have a more significant percentage of developed area will be analysed or where this is likely to change in future. This can possibly be implemented in the hydrological model in the form of an additional HRU. However, this will require a significant modification of the model codes used in the current the modeling process.

The notebooks for forcing generation, model calibration and making the historical and future streamflow simulations are currently designed specifically for the HBV-mountain hydrological model. Nevertheless, the processes are built on the same methods used for the other hydrological models on eWaterCycle. Forcing generation for other models can be easily done by using the corresponding ESMValTool recipe and diagnostics script. The ESMValTool recipe and diagnostics scripts for the eWaterCycle models are currently available in the ESMValTool environment, but they are designed for generating forcing from the ERA5 and ERA-Interim data sets only. Although changing this for generating CMIP6 data is mainly a matter of chang-

ing the data set name, not all necessary forcing data for the specific hydrological models may be available in the listed available CMIP6 models for the HBV-mountain model (see Appendix B.4). Considering the model calibration and running the streamflow simulations, the model set up is specifically applicable for the HBV-mountain model as well, but because of the BMI structure, the remainder of the processes will be very similar to other BMI models. For CMA-ES calibration of other hydrological models however, other parameter constraints applicable for the specific model must be given as well. Thus, even though the used methods should be applicable for the other eWaterCycle models, using another hydrological model in the modeling process requires some adjustments in the model functions.

5

Conclusion

In this thesis, a reproducible modeling approach for conducting a climate change impact assessment on streamflow was developed and tested in five different contrasting catchments across the United States. The modeling process was designed such that the analysis can be done with three separate notebooks: one notebook for generating forcing using the ESMValTool, one notebook for model calibration and one for making the historical and future climate simulations. In each notebook, the settings can easily and quickly be changed for analysing a different catchment, climate model or scenario. The analysis was done using the HBV-mountain hydrological model, which was calibrated for every catchment using GRDC streamflow data and ERA5 forcing. The model performances were different per catchment: highest calibration and validation scores were found in the Thunder Creek catchment, which is characterized by its glacial cover and large differences in altitude. Lowest scores were found in the relatively dry Big Rock Creek catchment and the Kawishiwi catchment. The low score in Big Rock Creek was caused by inaccurate precipitation data, while in the Kawishiwi catchment this was likely related to the exclusion of the lakes in the hydrological model. The analysis of climate change impact on streamflow was done by using forcing from the GFDL-CM4 and MPI-ESM1-2-HR models from CMIP6 for two future climate scenarios (SSP245 and SSP585). After comparing historical (1975 - 2005) simulations with the future (2070 - 2100) projections, it was found that all catchments will experience changes differently due to climate change. Changes in warmer catchments were mostly related to changes in precipitation and evaporation. Projections showed that these catchments will generally experience a streamflow decrease during the drier periods in both climate scenarios, while changes during the rest of the year vary per catchment and climate scenario. Changes in colder catchments were likely to be related to changing patterns in snow accumulation and melt processes. Projections here showed both an increase and decrease in streamflow, most likely due to an increase in glacier melt and a reduction in snow melt respectively. However, the results remain very uncertain due to imperfections of the ERA5 calibration data, disagreements between climate models and simplifications in the modeling process. The modeling approach presented in this study facilitates reproducibility in assessing the impact of climate change streamflow. This makes it easy to expand a similar analysis to a large ensemble of catchments, climate models and future scenarios.

References

- Aschonitis, V., Touloumidis, D., ten Veldhuis, M.-C., & Coenders-Gerrits, M. (2022). Correcting thornthwaite potential evapotranspiration using a global grid of local coefficients to support temperature-based estimations of reference evapotranspiration and aridity indices. *Earth System Science Data*, *14*(1), 163–177. <https://doi.org/10.5194/essd-14-163-2022>
- Azman, A. H., Tukimat, N. N. A., & Malek, M. A. (2022). Analysis of linear scaling method in downscaling precipitation and temperature. *Water Resour. Manage.*, *36*(1), 171–179.
- Bandhauer, M., Isotta, F., Lakatos, M., Lussana, C., Båserud, L., Izsák, B., Szentes, O., Tveito, O. E., & Frei, C. (2022). Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in european regions. *Int. J. Climatol.*, *42*(2), 727–747.
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2020). Publisher correction: Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data*, *7*(1), 274.
- Bergström, S., & Forsman, A. (1973). DEVELOPMENT OF A CONCEPTUAL DETERMINISTIC RAINFALL-RUNOFF MODEL. *Hydrology Research*, *4*(3), 147–170. <https://doi.org/10.2166/nh.1973.0012>
- Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Process.*, *6*(3), 279–298.
- Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., ... Živković, N. (2017). Changing climate shifts timing of european floods. *Science*, *357*(6351), 588–590.
- Brunner, M. I., Melsen, L. A., Newman, A. J., Wood, A. W., & Clark, M. P. (2020). Future streamflow regime changes in the united states: Assessment using functional classification. *Hydrology and Earth System Sciences*, *24*(8), 3951–3966. <https://doi.org/10.5194/hess-24-3951-2020>
- Chen, D., Gao, G., Xu, C.-Y., Guo, J., & Ren, G. (2005). Comparison of the thornthwaite method and pan data with the standard penman-monteith estimates of reference et in china. *CLIMATE RESEARCH*, *28*, 123–132. <https://doi.org/10.3354/cr028123>
- Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S. ., Wood, A. W., Fowler, H. J., Prudhomme, C. ., Arnold, J. R., & Brekke, L. D. (2016). Characterizing uncertainty of the hydrologic impacts of climate change. *Current Climate Change Reports*, *2*(2), 55–64. <https://doi.org/10.1007/s40641-016-0034-x>
- De Niel, J., Van Uytven, E., & Willems, P. (2019). Uncertainty analysis of climate change impact on river flow extremes based on a large multi-model ensemble. *Water Resour. Manage.*, *33*(12), 4319–4333.
- Efstratiadis, A., & Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: A review. *Hydrological Sciences Journal*, *55*(1), 58–78. <https://doi.org/10.1080/02626660903526292>
- Engeland, K., & Hisdal, H. (2009). A Comparison of Low Flow Estimates in Ungauged Catchments Using Regional Regression and the HBV-Model. *Water Resources Management: An International Journal, Published for the European Water Resources Association (EWRA)*, *23*(12), 2567–2586. <https://doi.org/10.1007/s11269-008-9397-7>
- Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., & Savenije, H. H. G. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences*, *17*(5), 1893–1912. <https://doi.org/10.5194/hess-17-1893-2013>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, *9*(5), 1937–1958.
- Fabry, F., Bellon, A., Duncan, M. R., & Austin, G. L. (1994). High resolution rainfall measurements by radar for very small basins: The sampling problem reexamined. *J. Hydrol. (Amst.)*, *161*(1-4), 415–428.
- Fountain, A. G., & Tangborn, W. V. (1985). The effect of glaciers on streamflow variations. *Water Resour. Res.*, *21*(4), 579–586.

- Frans, C., Istanbuloglu, E., Lettenmaier, D., Fountain, A., & Riedel, J. (2018). Glacier recession and the response of summer streamflow in the pacific northwest united states, 1960-2099. *Water Resources Research*, 54. <https://doi.org/10.1029/2017WR021764>
- Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., & Savenije, H. H. G. (2014). Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci.*, 18(12), 4839–4859.
- Guan, M., Sillanpää, N., & Koivusalo, H. (2015). Modelling and assessment of hydrological changes in a developing urban catchment. *Hydrological Processes*, 29(13), 2880–2894. <https://doi.org/https://doi.org/10.1002/hyp.10410>
- 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. *J. Hydrol. (Amst.)*, 377(1-2), 80–91.
- Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J., Storch, J.-S., Brüggemann, N., Haak, H., & Stössel, A. (2018). Max planck institute earth system model (mpi-esm1.2) for high-resolution model intercomparison project (highresmip). *Geoscientific Model Development Discussions*, 1–46. <https://doi.org/10.5194/gmd-2018-286>
- Hansen, N. (2007). The cma evolution strategy: A comparing review. https://doi.org/10.1007/3-540-32494-1_4
- Hanus, S. ., Hrachowitz, M. ., Zekollari, H. ., Schoups, G. ., Vizcaino, M. ., & Kaitna, R. . (2021). Future changes in annual, seasonal and monthly runoff signatures in contrasting alpine catchments in austria. *Hydrology and Earth System Sciences*, 25(6), 3429–3453. <https://doi.org/10.5194/hess-25-3429-2021>
- Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., Bushuk, M., Wittenberg, A. T., Wyman, B., Xiang, B., Zhang, R., Anderson, W., Balaji, V., Donner, L., Dunne, K., Durachta, J., ... Zadeh, N. (2019). Structure and performance of GFDL's CM4.0 climate model. *J. Adv. Model. Earth Syst.*, 11(11), 3691–3727.
- Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., & Seong, C. (2019). Uncertainty in hydrological analysis of climate change: Multi-parameter vs. multi-GCM ensemble predictions. *Sci. Rep.*, 9(1), 4974.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Jean-Noël Thépaut. (2020). The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.*, 146(730), 1999–2049.
- Hirpa, F. A., Alfieri, L. ., Lees, T. ., Peng, J. ., Dyer, E. ., & Dadson, S. J. (2019). Streamflow response to climate change in the greater horn of africa. *Climatic Change*, 156(3), 341–363. <https://doi.org/10.1007/s10584-019-02547-x>
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., & Wickham. (2007). Completion of the 2001 national land cover database for the conterminous united states. *Photogrammetric Engineering and Remote Sensing*, 73(4), 337–341.
- Hoogelander, V. (2022). *vhoogelander/BMI_HBVmountain_streamflow_analysis: Thesis (Version v0.1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.7405678>
- Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., & Gascuel-Odoux, C. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resources Research*, 50(9), 7445–7469. <https://doi.org/https://doi.org/10.1002/2014WR015484>
- Huang, Y., Liu, W., Wu, B., Wan, Y., & Zhu, S. (2020). Climate change in ganjiang river basin and its impact on runoff. *IOP Conf. Ser. Mater. Sci. Eng.*, 964(1), 012008.
- Huss, M., Juvet, G., Farinotti, D., & Bauder, A. (2010). Future high-mountain hydrology: A new parameterization of glacier retreat. *Hydrology and Earth System Sciences*, 14(5), 815–829. <https://doi.org/10.5194/hess-14-815-2010>
- Hut, Drost, N. ., van de Giesen, N. ., van Werkhoven, B. ., Abdollahi, B. ., Aerts, J. ., Albers, T. ., Alidoost, F. ., Andela, B. ., Camphuijsen, J. ., Dzigan, Y. ., van Haren, R. ., Hutton, E. ., Kalverla, P. ., van Meersbergen, M. ., van den Oord, G. ., Pelulessy, I. ., Smeets, S. ., Verhoeven, S. ., ... Weel, B. . (2022). The ewatercycle platform for open and fair hydrological collaboration. *GMD*. <https://doi.org/10.5194/gmd-2021-344>
- Hut, van de Giesen, N., & Drost, N. (2017). Comment on “most computational hydrology is not reproducible, so is it really science?” by christopher hutton et al.: Let hydrologists learn the latest computer sci-

- ence by working with research software engineers (rses) and not reinvent the waterwheel ourselves. *Water Resources Research*, 53(5), 4524–4526. <https://doi.org/10.1002/2017WR020665>
- Hutton, C., Wagener, T., Freer, J., Han, D., Duffy, C., & Arheimer, B. (2016). Most computational hydrology is not reproducible, so is it really science? *Water Resources Research*, 52. <https://doi.org/10.1002/2016WR019285>
- Hutton, E., Piper, M., & Tucker, G. (2020). The basic model interface 2.0: A standard interface for coupling numerical models in the geosciences. *Journal of Open Source Software*, 5, 2317. <https://doi.org/10.21105/joss.02317>
- Ikram, R. M. A., Goliatt, L., Kisi, O., Trajkovic, S., & Shahid, S. (2022). Covariance matrix adaptation evolution strategy for improving machine learning approaches in streamflow prediction. *Mathematics*, 10(16). <https://doi.org/10.3390/math10162971>
- IPCC. (2021). *Climate change 2021: The physical science basis. contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, B. Zhou (eds.)]. Cambridge University Press. In Press.
- Jain, S. K., & Sudheer, K. P. (2008). Fitting of hydrologic models: A close look at the nash-sutcliffe index. *Journal of Hydrologic Engineering*, 13(10), 981–986.
- Jaramillo, F., & Destouni, G. (2014). Developing water change spectra and distinguishing change drivers worldwide. *Geophysical Research Letters*, 41(23), 8377–8386.
- Jin, J., Yang, C., & Zhang, Y. (2020). An improved cma-es for solving large scale optimization problem. https://doi.org/10.1007/978-3-030-53956-6_34
- Kiesel, J. ., Gericke, A. ., Rathjens, H. ., Wetzig, A. ., Kakouei, K. ., Jähnig, S. C., & Fohrer, N. . (2019). Climate change impacts on ecologically relevant hydrological indicators in three catchments in three european ecoregions. *Ecological Engineering*, 127, 404–416. <https://doi.org/10.1016/j.ecoleng.2018.12.019>
- Kobold, M., & Brilly, M. (2006). The use of hbv model for flash flood forecasting. *Natural Hazards and Earth System Sciences*, 6(3), 407–417. <https://doi.org/10.5194/nhess-6-407-2006>
- Köppen, W. (1936). *Das geographische system der klimaregionen*. Verlag von Gebrüder Borntraeger. <https://books.google.nl/books?id=hM2uugAACAAJ>
- Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.*, 5, 89–97.
- Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., & Kundzewicz, Z. W. (2018). How the performance of hydrological models relates to credibility of projections under climate change [Cited By :84]. *Hydrological Sciences Journal*, 63(5), 696–720. www.scopus.com
- Lane, R. A., Coxon, G., Freer, J., Seibert, J., & Wagener, T. (2022). A large-sample investigation into uncertain climate change impacts on high flows across great britain. *Hydrology and Earth System Sciences*, 26(21), 5535–5554. <https://doi.org/10.5194/hess-26-5535-2022>
- Lenderink, G., Buishand, A., & van Deursen, W. (2007). Estimates of future discharges of the river rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci.*, 11(3), 1145–1159.
- Leong, C., & Yokoo, Y. (2021). A step toward global-scale applicability and transferability of flow duration curve studies: A flow duration curve review (2000–2020). *J. Hydrol. (Amst.)*, 603(126984), 126984.
- Leta, O. T., El-Kadi, A. I., & Dulai, H. (2018). Impact of climate change on daily streamflow and its extreme values in pacific island watersheds. *Sustainability*, 10(6). <https://doi.org/10.3390/su10062057>
- Li, H., Beldring, S., Xu, C.-Y., Huss, M., Melvold, K., & Jain, S. K. (2015). Integrating a glacier retreat model into a hydrological model – case studies of three glacierised catchments in norway and himalayan region. *Journal of Hydrology*, 527, 656–667. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.05.017>
- Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern united states. *J. Am. Water Resour. Assoc.*, 41(3), 621–633.
- Luo, Y., Arnold, J., Liu, S., Wang, X., & Chen, X. (2013). Inclusion of glacier processes for distributed hydrological modeling at basin scale with application to a watershed in tianshan mountains, northwest china. *Journal of Hydrology*, 477, 72–85. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2012.11.005>

- McGlynn, E., Li, S., F. Berger, M., Amend, M., & L. Harper, K. (2022). Addressing uncertainty and bias in land use, land use change, and forestry greenhouse gas inventories. *Clim. Change*, *170*(1-2).
- McMillan, H., Krueger, T., & Freer, J. (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrol. Process.*, *26*(26), 4078–4111.
- Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., & Teuling, A. J. (2018). Mapping (dis)agreement in hydrologic projections. *Hydrology and Earth System Sciences*, *22*(3), 1775–1791. <https://doi.org/10.5194/hess-22-1775-2018>
- Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., Brekke, L. D., & Rajagopalan, B. (2016). Effects of different regional climate model resolution and forcing scales on projected hydrologic changes. *J. Hydrol. (Amst.)*, *541*, 1003–1019.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.
- NASA. (2013). Shuttle radar topography mission (SRTM) global [Accessed on 12 August 2022].
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — a discussion of principles. *J. Hydrol. (Amst.)*, *10*(3), 282–290.
- Ndhlovu, G. Z., & Woyessa, Y. E. (2021). Evaluation of streamflow under climate change in the zambezi river basin of southern africa. *Water*, *13*(21). <https://doi.org/10.3390/w13213114>
- Nel, J. L., Le Maitre, D. C., Nel, D. C., Reyers, B., Archibald, S., van Wilgen, B. W., Forsyth, G. G., Theron, A. K., O'Farrell, P. J., Kahinda, J.-M. M., Engelbrecht, F. A., Kapangaziwiri, E., van Niekerk, L., & Barwell, L. (2014). Natural hazards in a changing world: A case for ecosystem-based management. *PLoS One*, *9*(5), e95942.
- Neves, G. L., Barbosa, M. A. G. A., Anjinho, P. d. S., Guimarães, T. T., das Virgens Filho, J. S., & Mauad, F. F. (2020). Evaluation of the impacts of climate change on streamflow through hydrological simulation and under downscaling scenarios: Case study in a watershed in southeastern brazil. *Environ. Monit. Assess.*, *192*(11), 707.
- Pereira, A. R., & Pruitt, W. O. (2004). Adaptation of the thornthwaite scheme for estimating daily reference evapotranspiration. *Agricultural Water Management*, *66*(3), 251–257. <https://doi.org/https://doi.org/10.1016/j.agwat.2003.11.003>
- Poissant, D., Arsenault, R., & Brissette, F. (2017). Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. *J. Hydrol. Reg. Stud.*, *12*, 220–237.
- Reder, A., & Rianna, G. (2021). Exploring ERA5 reanalysis potentialities for supporting landslide investigations: A test case from campania region (southern italy). *Landslides*.
- Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., Vegas-Regidor, J., Bock, L., Brötz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Loosveldt Tomas, S., & Zimmermann, K. (2020). Earth system model evaluation tool (ESMValTool) v2.0 – technical overview. *Geosci. Model Dev.*, *13*(3), 1179–1199.
- Roudier, P., Ducharme, A., & Feyen, L. (2014). Climate change impacts on runoff in west africa: A review. *Hydrol. Earth Syst. Sci.*, *18*(7), 2789–2801.
- Salo, J. A., Theobald, D. M., & Brown, T. C. (2016). Evaluation of methods for delineating riparian zones in a semi-arid montane watershed. *J. Am. Water Resour. Assoc.*, *52*(3), 632–647.
- Schoups, G., Hopmans, J., Young, C., Vrugt, J., & Wallender, W. (2005). Multi-criteria optimization of a regional spatially-distributed subsurface water flow model. *Journal of Hydrology*, *311*, 20–48. <https://doi.org/10.1016/j.jhydrol.2005.01.001>
- Seibert, J., & Bergström, S. (2022). A retrospective on hydrological catchment modelling based on half a century with the hbv model. *Hydrology and Earth System Sciences*, *26*(5), 1371–1388. <https://doi.org/10.5194/hess-26-1371-2022>
- Seibert, J., Vis, M. J. P., Kohn, I., Weiler, M., & Stahl, K. (2018). Technical note: Representing glacier geometry changes in a semi-distributed hydrological model. *Hydrology and Earth System Sciences*, *22*(4), 2211–2224. <https://doi.org/10.5194/hess-22-2211-2018>
- Sellami, H., Benabdallah, S., La Jeunesse, I., & Vanclooster, M. (2016). Climate models and hydrological parameter uncertainties in climate change impacts on monthly runoff and daily flow duration curve of a mediterranean catchment. *Hydrol. Sci. J.*, *61*(8), 1415–1429.
- Song, M., Jiang, Y., Liu, Q., Tian, Y., Liu, Y., Xu, X., & Kang, M. (2021). Catchment versus riparian buffers: Which land use spatial scales have the greatest ability to explain water quality changes in a typical temperate watershed? *Water*, *13*(13). <https://doi.org/10.3390/w13131758>

- Tangborn, W. (1980). Two models for estimating climate–glacier relationships in the north cascades, washington, u.s.a. *Journal of Glaciology*, 25(91), 3–22. <https://doi.org/10.3189/S0022143000010248>
- Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over north america. *Hydrol. Earth Syst. Sci.*, 24(5), 2527–2544.
- Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z., Hurtt, G., Kriegler, E., Lamarque, J.-F., Meehl, G., Moss, R., & Ziehn, T. (2020). Climate model projections from the scenario model intercomparison project (scenariomip) of cmip6. <https://doi.org/10.5194/esd-2020-68>
- Teutschbein, C., & 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*, 456–457, 12–29. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2012.05.052>
- Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Soil Science*, 66, 55–94.
- Trajkovic, S., & Kolakovic, S. (2009). Evaluation of reference evapotranspiration equations under humid conditions. *Water Resour. Manage.*, 23(14), 3057–3067.
- Van Tiel, M., Teuling, A. J., Wanders, N., Vis, M. J. P., Stahl, K., & Van Loon, A. F. (2018). The role of glacier changes and threshold definition in the characterisation of future streamflow droughts in glacierised catchments. *Hydrology and Earth System Sciences*, 22(1), 463–485. <https://doi.org/10.5194/hess-22-463-2018>



Implementation of the HBV-mountain model to the eWaterCycle platform

A.1. Add a BMI to the hydrological model

One goal of this thesis work was to rewrite the HBV-mountain hydrological model, developed in the work of Hanus et al., 2021, such that it can be implemented on the eWaterCycle platform. Before a model can be implemented to eWaterCycle, a Basic Model Interface (BMI) must be added to the model. In short, the BMI is an interface, which provides a standardized set of functions, allowing users to reuse the model in a straightforward way (Hutton et al., 2020). Therefore, the first step of adding the model to eWaterCycle is to add a BMI to the model. BMI is currently supported in the following programming languages: C, C++, Fortran, Java, and Python (CSDMS, <https://bmi.readthedocs.io/en/latest/>, last access: 1 September 2022). All model codes from this study were reused in this study (sarah-hanus, 2021), and modified in a way so that the BMI can be added to the model.

Since the original code of the HBV-mountain model used in Hanus et al. (2021) is written in Julia, a BMI implementation for the Julia programming language was written first. This was done by creating a HBV-mountain model Julia mutable object containing all model components (see figure A.1). This object can be used in all BMI functions. Then, the Julia BMI functions were wrapped into a Python BMI, using the Pyjulia Python package. This is done to ensure that the model can be installed together with grpc4bmi inside a Docker container (see appendix A.3). The features of the HBV-mountain (Julia) BMI model and a brief description of the workflow of how this BMI can be used is described in Appendix B.1.

Link to the Julia BMI implementation: https://github.com/vhoogelander/BMI_HBVmountain_streamflow_analysis/blob/master/HBVmountain/General/Refactoring/Julia_BMI_model.jl

Link to the Python BMI wrapping: https://github.com/vhoogelander/BMI_HBVmountain_streamflow_analysis/blob/master/HBVmountain/General/BMI_HBVmountain_Python.py

A.2. ESMValTool

In eWaterCycle, forcing data for the hydrological models is generated using the ESMValTool, which is a tool to generate and analyse climate data from large climatic data sets (Righi et al., 2020). Important components of this are a ESMValTool recipe and a diagnostic script: the recipe contains instructions which are needed to generate the desired forcing data. The main instructions mainly include the directory and name of the basin shapefile, the climate dataset from which data is to be retrieved, the desired variables, the start and end year and which diagnostics script should be used. The diagnostics script is a Python script containing the several functions which are used to obtain the desired data in the manner specified in the recipe. Next to this, an ESMValTool configuration file is needed which contains information needed by ESMValTool to run, which is mostly related to directories of the stored climate datasets, the output of the ESMValTool runs and the auxiliary data.

As ESMValTool recipes and diagnostic scripts are often written for very specific purposes (e.g. forcing input for a specific hydrological model), a new recipe is written for the HBV-mountain model. For this, the example recipe + diagnostics script of the MARRMoT hydrological model (Knoben et al., 2019) was used

```

mutable struct HBVmountain_model
  #Storages and discharge
  Discharge::Union{Nothing, Float64}
  Snow_Extend::Union{Nothing, Array{Float64,1}}
  bare_storage::Union{Nothing, Storages}
  forest_storage::Union{Nothing, Storages}
  grass_storage::Union{Nothing, Storages}
  rip_storage::Union{Nothing, Storages}
  Slowstorage::Union{Nothing, Float64}
  Waterbalance::Union{Nothing, Float64}
  Glacier::Union{Nothing, Array{Float64,1}}
  #Model settings
  Area
  bare_parameters::Union{Nothing, Parameters}
  forest_parameters::Union{Nothing, Parameters}
  grass_parameters::Union{Nothing, Parameters}
  rip_parameters::Union{Nothing, Parameters}
  slow_parameters::Union{Nothing, Slow_Parameters}
  Elevation::Union{Nothing, Elevations}
  Total_Elevationbands ::Union{Nothing, Int64}
  Precipitation_gradient ::Union{Nothing, Float64}
  Elevation_Percentage
  #Initial input settings
  bare_input::Union{Nothing, HRU_Input}
  forest_input::Union{Nothing, HRU_Input}
  grass_input::Union{Nothing, HRU_Input}
  rip_input::Union{Nothing, HRU_Input}
  #Forcing
  Precipitation::Union{Nothing, Array{Float64,2}}
  Temperature::Union{Nothing, Array{Float64,2}}
  ETP::Union{Nothing, Array{Float64,1}}
  #Other
  Date::Union{Nothing, Array{Date,1}}
  Current_Date::Union{Nothing, Date}
  Sunhours::Union{Nothing, Array{Float64, 1}}
  Units::Union{Nothing, HBVmountain_model_units}
end

mutable struct HRU_Input
  #inputs (alphabetic order)
  Area_Elevations::Array{Float64,1}
  Area_HRU:: Float64
  Area_Glacier::Array{Float64,1} # smaller than 1
  Elevation_Count::Array{Int64}
  Nr_Elevationbands:: Int64
  Catchment_Elevation::Tuple
  Snow_Redistribution::Tuple
  #Potential_Evaporation::Array{Float64,1}
  Potential_Evaporation_Mean:: Float64
  Precipitation::Array{Float64,1}
  Riparian_Discharge:: Float64 #only necessary for riparian HRU
  Temp_Elevation::Array{Float64,1}
  Total_Effective_Precipitation::Float64
  Total_Interception_Evaporation::Float64
end

mutable struct Parameters
  # parameters (alphabetic order)
  beta:: Float64
  Ce:: Float64
  Drainagecapacity:: Float64 #only necessary for riparian HRU
  Interceptionstoragecapacity:: Float64
  Kf:: Float64
  Meltfactor:: Float64
  Mm:: Float64
  #Percolationcapacity:: Float64 #only necessary for hillslope HRU
  Ratio_Pref:: Float64 #only necessary for hillslope HRU
  Soilstoragecapacity:: Float64
  Temp_Thresh:: Float64
end

mutable struct Slow_Parameters
  Ks:: Float64
  Ratio_Riparian:: Float64
end

```

Figure A.1: HBV-mountain Julia model object (right), individual HRU input object and parameter objects

and adapted for use of the HBV-mountain model. The MARRMoT model requires precipitation and potential evapotranspiration as model forcing input, while the HBV-mountain model requires precipitation and temperature as forcing input. Therefore, the MARRMoT recipe and diagnostic script had to be rewritten in order to obtain precipitation and temperature data using ESMValTool.

The ESMValTool recipe for generating CMIP6 forcing for the HBV-mountain model can be found at: https://github.com/vhoogelander/Thesis/blob/master/ESMValTool/esmvaltool/recipes/hydrology/recipe_HBVmountain.yml.

The diagnostics script corresponding to this recipe can be found at: https://github.com/vhoogelander/BMI_HBVmountain_streamflow_analysis/blob/master/ESMValTool/esmvaltool/recipes/hydrology/recipe_HBVmountain.yml.

Link to the configuration file used on Levante: https://github.com/vhoogelander/BMI_HBVmountain_streamflow_analysis/blob/master/config-user.yml.

A.3. Docker container

The modeling approach is stored in a Docker container with `grpc4bmi` server running as entrypoint. This is done by creating a Docker image using the Dockerfile which can be found in the Github repository. This allows users to use the model without the need of installing all necessary software. Next to this, by using `grpc4bmi`, it is possible to communicate with the model from outside the container. This is necessary for later implementation on the eWaterCycle platform. The used steps in the Dockerfile are briefly summarized below:

1. Create a layer from the `continuumio/anaconda3` base image.
2. Install `grpc4bmi` package (version 0.2.3).
3. Install Julia.
4. Install necessary Python packages, including the BMI packages.
5. Install necessary Julia packages.
6. Install GCC compiler.
7. Install PyJulia using a custom Julia system image. This is necessary because the Python interpreter is statically linked to `libpython`.
8. Run BMI server.
9. Expose `grpc4bmi` port.

Link to Docker hub: <https://hub.docker.com/repository/docker/vhooglander/hbvmountain>



Climate change impact analysis tool

B.1. Features of the HBV-mountain BMI model

As discussed, the Python BMI wraps a Julia BMI which was developed for the HBV-mountain model. This section discusses the functionalities of the most important BMI functions which are needed for the model run, and which processing functions are used in the developed BMI for the HBV-mountain model. All developed code is available from <https://doi.org/10.5281/zenodo.7405678> (Hoogelander, 2022)

HBV-mountain BMI model object

First, an empty HBV-mountain BMI model object is created for use in the Python BMI functions. This object has the structure of the Julia object shown in figure A.1. The user should give the precipitation and temperature data in a NetCDF file while creating the model object. In case the catchment related settings in the model must be preprocessed automatically, the users must give the path to the catchment shapefile, DEM and NLCD land use map as well.

Setup

Creates a configuration file based on the given forcing and paths to the catchment's shapefile, DEM and NLCD land use map. In this function, users can give the parameter set to the model, and the function will store these as a Julia HBV-mountain parameter object which can be used in the Julia functions of the model. The forcing data of the NetCDF file is converted to a CSV file containing the dates, precipitation and temperature and is given to the model object. In addition, the catchment related settings are automatically processed using the shapefile, DEM and land use maps. After all relevant model settings are preprocessed, the function returns the configuration file (Julia object) using a Julia setup function. In case no parameter set or shapefile is given to the empty model object, the setup function returns a configuration file with some default values for the model setup.

Initialize

This function initializes the empty HBV-mountain with the created configuration file of the setup function. In the Python BMI, a initializing function implemented in Julia is invoked, which replaces the empty HBV-mountain model object with the configuration file.

Update

The update function updates the model for one time step, and invokes the BMI update function in Julia, which invokes the Julia model functions. First, in case no potential evapotranspiration data is given, the daily potential evapotranspiration is calculated using the Thornthwaite method. Then, the discharge, total evaporation, snow extend, water balance and storage per HRU at timestep $t+1$ are calculated using the model run function with the model state at timestep t as input. The output of this function is given to the new model state.

Get value

The get value function takes a model variable name, and returns the corresponding value of the model state.

Set value

The set value function takes a model variable and overwrites this specific variable in the model state. The user is for example able to manually give specific forcing data to the model in the form of an array without the need of a NetCDF file. In case all forcing data is manually given to the model, a list of the dates must be given as well.

Finalize

This function clears the model, and should be used when the user is finished with the model runs.

The workflow of how the BMI model is used for the simulations is summarized below:

1. Create a HBV-mountain model object, and give the forcing and the locations of the catchment shapefile, DEM and NLCD map to this object.
2. Create a configuration file, containing some initial model settings.
3. Initialize the HBV-mountain model object using the BMI initialize function with the created configuration file as input.
4. Update the BMI model state using the BMI update function. Time series of streamflow simulations can be created using a loop, by adding the model discharge state to a list within this loop after each model update.

B.2. Code structure

The general folder contains four different notebooks: the three notebooks described in section 2.6, and one example notebook for making a single model run. This section provides some guidance on the code structure on which the functionalities in the notebooks are built.

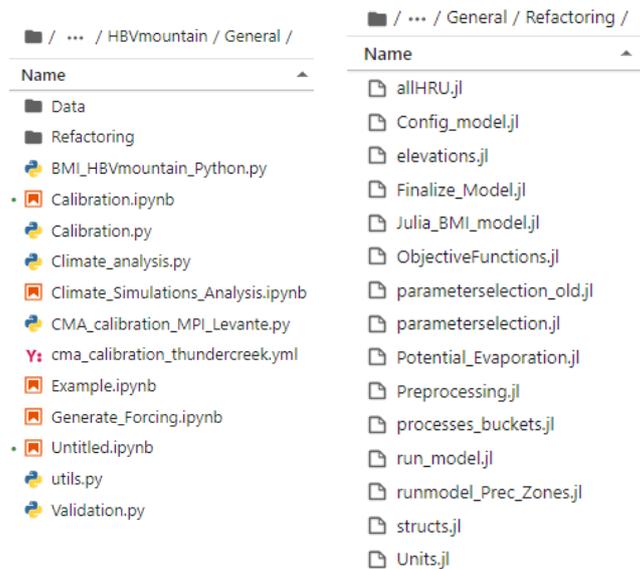


Figure B.1: Code structure in the general folder

Generate forcing

The forcing notebook currently operates only on the HPC cluster "Levante" for generating CMIP6 data and the eWaterCycle environment for generating ERA5 data. The functions in this notebook are mostly built on the ESMValCore Python package. The YAML file containing the list of available CMIP6 models is located in the Data folder. Next to this, the notebook uses a local version of the ESMValTool (not present in the

general folder), containing the necessary recipe and diagnostics script of the HBV-mountain model. This local version of ESMValTool is available on the Github repository of this project. The shapefiles used for generating the forcing data for the specific area must be located in an auxiliary data folder (not present in the general folder). Lastly, the ESMValTool configuration file indicates in which locations the datasets and auxiliary data are stored.

Model Calibration

The model calibration entails notebook 2, and is built on the utils, Calibration and Validation Python files and all Julia files in the Refactoring folder. The Calibration file contains all functions related to the CMA-ES calibration process and the different objective functions. The Validation file contains the validation function, in which the streamflow simulations from the calibration are returned as well. The functions in both files are partly build on several Python packages and preprocessing functions which are imported from the utils file. The preprocessing functions of the utils file are related to the catchment related settings, converting the NetCDF files to forcing for the HBV-mountain model and the bias correction of the CMIP6 data. The BMI model functions used in the calibration and validation processes are built on the Julia_BMI_model file in the Refactoring folder. The Julia BMI function uses the different model functions developed in the work of Hanus et al. (2021). The model run is done in the run_model Julia file, which performs the run of the semi-distributed bucket model. This file is again build on several model functions from the other Julia files.

Climate change impact analysis

The climate change impact analysis of notebook 3 is build on the Climate_Analysis Python file, and again on the utils file and Julia files in the Refactoring folder. The Climate_Analysis file contains the function for running the historical and future streamflow simulations, using the parameter sets obtained from the calibration. Like in the calibration and validation, the model set up is done using the preprocessing functions from the utils file, and the BMI functions for making the simulations are built on the Julia functions from the Refactoring folder.

B.3. Increasing computation speed for model calibration

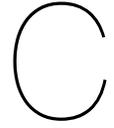
In order to reduce the equifinality of the hydrological model, an ensemble of good performing parameter sets is used for the climate change impact analysis. This done by selecting the 20 best parameter sets of 50 different CMA-ES optimization processes. As each optimization requires many model runs, the total computation speed was increased by performing the different model run functions on the HPC cluster "Levante", using the Message Passing Interface (MPI) for Python (<https://mpi4py.readthedocs.io/en/stable/intro.html>, latest access: 30 November 2022). The Python script used for this can be found in the General folder, and is based on the functions used in the calibration notebook.

Another effective way of reducing the calibration time is to use less iterations in the CMA-ES. The figures of Appendix C.3 show how the score of the multi-objective function evolves over the iterations in all catchments for one CMA-ES calibration process. As can be seen, the objective score did not noticeably increase after only a few iterations in the Big ROck Creek and Kawishiwi catchments for example. Almost all cases show that a sufficiently performing parameter set can be generated after only a few iterations, and that using this a large number of iterations is not necessarily associated with a large performance increase. In the future, the user should firstly determine the extent to which increasing iterations leads to a performance increase, before starting with a large calibration process.

B.4. Available CMIP6 models

Table B.1: CMIP6 datasets listed in the YAML file for generating historical and future forcing for the HBV-mountain model

CMIP6 dataset	SSP119	SSP126	SSP245	SSP370	SSP434	SSP460	SSP585
ACCESS-ESM1-5		X	X	X			X
BCC-CSM2-MR		X	X	X			X
CanESM5	X	X	X	X	X	X	X
CESM2-WACCM		X	X	X			X
CNRM-CM6-1		X	X	X			X
CNRM-CM6-1-HR		X	X	X			X
CNRM-ESM2-1	X	X	X	X	X	X	X
EC-Earth3	X	X	X	X	X	X	X
EC-Earth3-CC			X	X			X
EC-Earth3-Veg	X	X	X	X			X
EC-Earth3-Veg-LR	X	X	X	X			X
FGOALS-g3	X	X	X	X	X	X	X
GFDL-CM4			X				X
GFDL-ESM4		X	X	X			X
IITM-ESM		X	X	X			X
INM-CM4-8		X	X	X			X
INM-CM5-0		X	X	X			X
IPSL-CM6A-LR	X	X	X	X	X	X	X
KIOST-ESM		X	X				X
MIROC6	X	X	X	X	X	X	X
MIROC-ES2L	X	X	X	X			X
MPI-ESM1-2-HR		X	X	X			X
MPI-ESM1-2-LR		X	X	X			X
MRI-ESM2-0	X	X	X	X	X	X	X
NESM3		X	X	X			X
NorESM2-LM		X	X	X			X
NorESM2-MM		X	X	X			X
TaiESM1		X	X	X			X



Calibration and validation of the hydrological model

C.1. Objective functions

The overall model performance per parameter set were assessed by one objective function (equation C.1) based on the Euclidean Distance (Hrachowitz et al., 2014, which was calculated using the Nash Sutcliffe efficiency (Nash and Sutcliffe, 1970) of four different individual signatures. (Nash and Sutcliffe, 1970). All individual objective functions are described below.

$$\text{Obj}_{\text{tot}} = 1 - \sqrt{\frac{\sum_{i=1}^n (1 - NSE_i)^2}{n}} \quad (\text{C.1})$$

Flow timeseries

Optimization of the flow timeseries was done using two different objective functions: firstly, by using the NSE between the modelled and the observed streamflow (equation C.2), and secondly by using the NSE between the logarithms of the modelled and the observed streamflow (equation C.3). The NSE of the timeseries is sensitive to high flows, while the Log NSE is more sensitive to low flows.

$$NSE_{\text{timeseries}} = 1 - \frac{\sum_{i=1}^n (Q_{\text{modelled},i} - Q_{\text{observed},i})^2}{\sum_{i=1}^n (Q_{\text{observed},i} - \bar{Q}_{\text{observed}})^2} \quad (\text{C.2})$$

$$\log NSE_{\text{timeseries}} = 1 - \frac{\sum_{i=1}^n (\log(Q_{\text{modelled},i}) - \log(Q_{\text{observed},i}))^2}{\sum_{i=1}^n (\log(Q_{\text{observed},i}) - \log(\bar{Q}_{\text{observed}}))^2} \quad (\text{C.3})$$

Flow Duration Curve

The flow duration curve shows the percent of time that discharge equals or exceeds a certain level during a specified period (Searcy, 1959). The curve can be determined by ranking all discharges during in a given period. Optimization was done by calculating the NSE between the modelled and observed flow duration curve (equation C.4). To reduce the effect of the high peak flows, the logarithm of the discharges is used for the calculation of the flow duration curve.

$$NSE_{\text{FDC}} = 1 - \frac{\sum_{i=1}^n (FDC_{\text{modelled},i} - FDC_{\text{observed},i})^2}{\sum_{i=1}^n (FDC_{\text{observed},i} - \bar{FDC}_{\text{observed}})^2} \quad (\text{C.4})$$

Annual runoff coefficient

The annual runoff coefficient represents the proportion of total annual precipitation that is converted to streamflow, and can be calculated using equation C.5. The optimization of this signature was done by calculating the NSE between simulated and observed yearly runoff coefficient (equation C.6).

$$RC = \frac{\sum_{i=1}^n Q}{\sum_{i=1}^n P} \quad (\text{C.5})$$

$$NSE_{\text{RC}} = 1 - \frac{\sum_{i=1}^n (RC_{\text{modelled},i} - RC_{\text{observed},i})^2}{\sum_{i=1}^n (RC_{\text{observed},i} - \bar{RC}_{\text{observed}})^2} \quad (\text{C.6})$$

C.2. Parameter constraints

The HBV-mountain model includes 20 parameters which had to be calibrated in the catchment. All model parameters were constraint following the suggestions of the study of Gharari et al. (2014), which were also used in the study of Hanus et al., 2021. The used constraints can be found in table C.1. Furthermore, the parameters followed the additional process constraints as described in equations C.7 - C.11, which were also used in the study of Hanus et al., 2021.

Table C.1: Parameter constraints used for calibration

	Unit	Min	Max
Global parameters			
$T_{\text{threshold}}$	°C	-2.0	2.0
F_{melt}	mm °C ⁻¹	1.0	5.0
M_m	°C	0.001	0.1
ρ_p	-	0.1	0.9
k_{fast}	d ⁻¹	0.1	3.0
k_{slow}	d ⁻¹	0.001	0.1
$F_{\text{evaporation}}$	-	0.4	0.8
Bare parameters			
$S_{u, \text{max, bare}}$	mm	1.0	50
β_{bare}	-	0.1	2.0
Forest parameters			
$I_{\text{max, forest}}$	mm	1.0	5.0
$S_{u, \text{max, forest}}$	mm	50	500
β_{forest}	-	0.1	2.0
Grass parameters			
$I_{\text{max, grass}}$	mm	0.3	2.0
$S_{u, \text{max, grass}}$	mm	5.0	250
β_{grass}	-	0.1	2.0
Bare parameters			
$I_{\text{max, rip}}$	mm	0.3	3.0
$S_{u, \text{max, rip}}$	mm	5.0	250
β_{rip}	-	0.1	2.0
$K_{\text{fast, rip}}$	d ⁻¹	0.2	3.0
ρ_{rip}	d ⁻¹	0.05	0.5

$$I_{\text{max, forest}} > I_{\text{max, grass}} \quad (\text{C.7})$$

$$I_{\text{max, forest}} > I_{\text{max, rip}} \quad (\text{C.8})$$

$$S_{u, \text{max, forest}} > S_{u, \text{max, grass}} > S_{u, \text{max, rip}} \quad (\text{C.9})$$

$$S_{u, \text{max, forest}} > S_{u, \text{max, grass}} > S_{u, \text{max, bare}} \quad (\text{C.10})$$

$$k_{\text{fast, rip}} > k_{\text{fast}} > k_{\text{slow}} \quad (\text{C.11})$$

C.3. CMA-ES calibration iterations

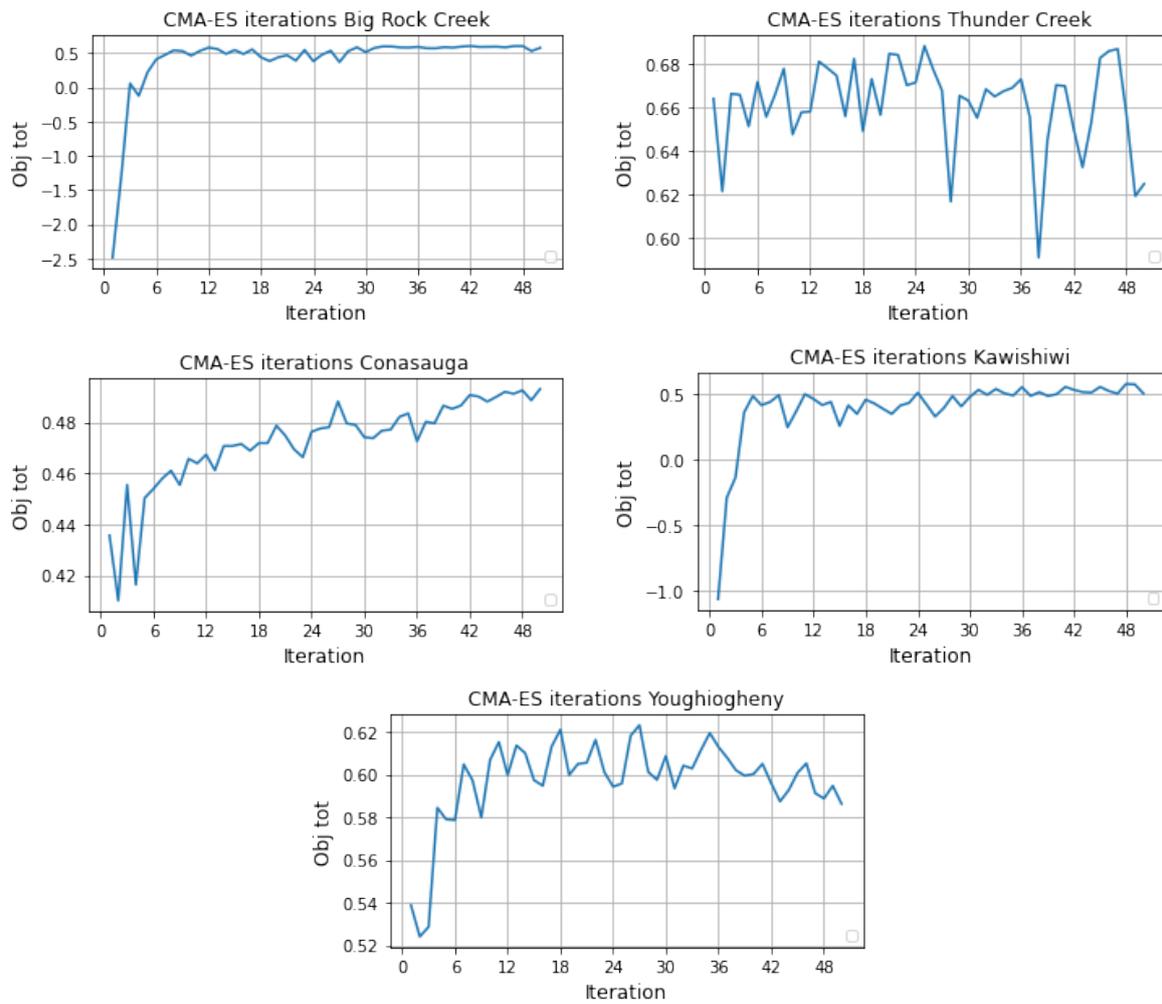


Figure C.1: CMA-ES iterations of one CMA-ES optimization process per catchment

D

Additional data

D.1. Calibration and validation results

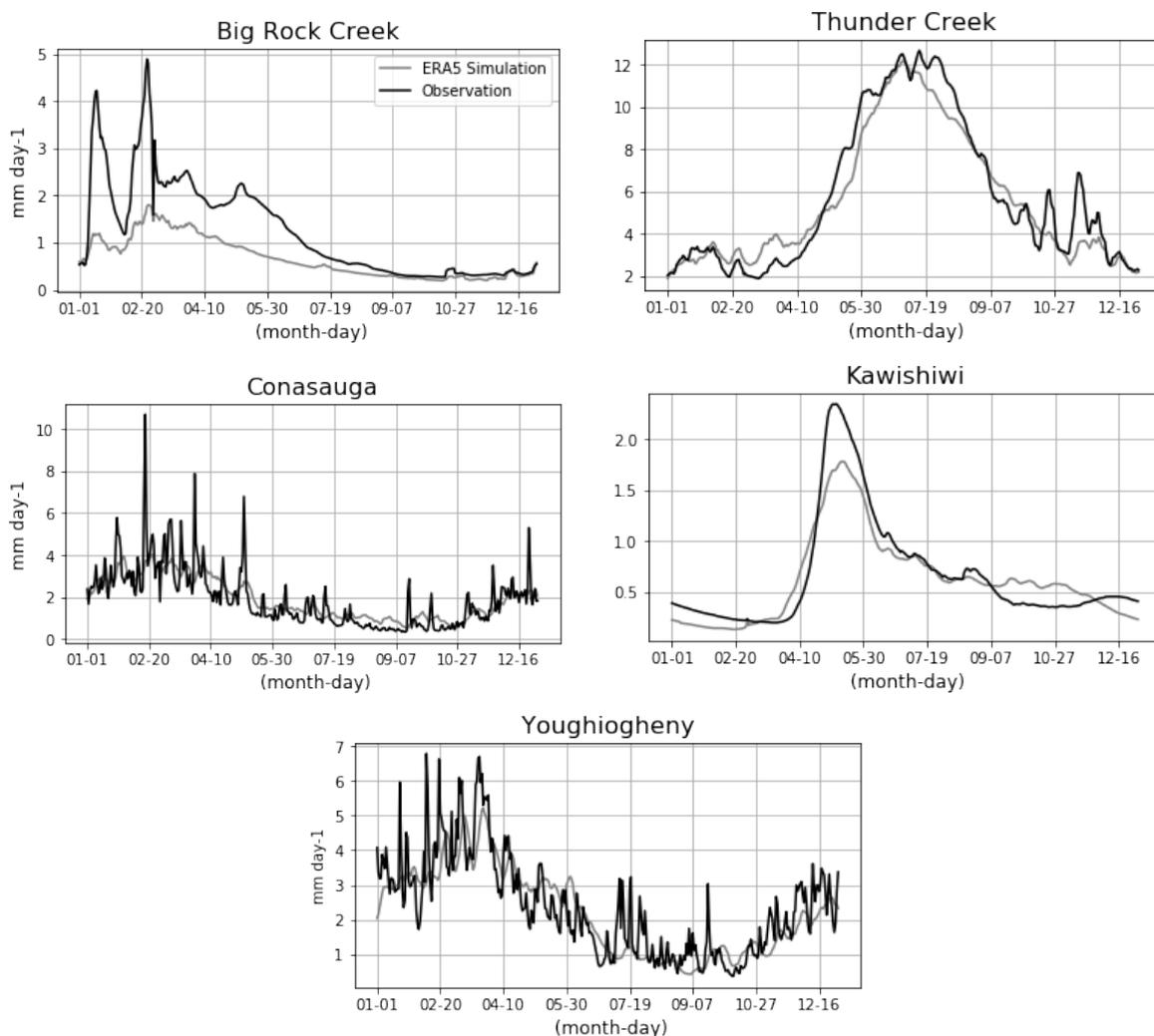


Figure D.1: Daily streamflow using ERA5 simulations and GRDC observations in calibration/validation period (2001)

D.2. Climate change impact analysis

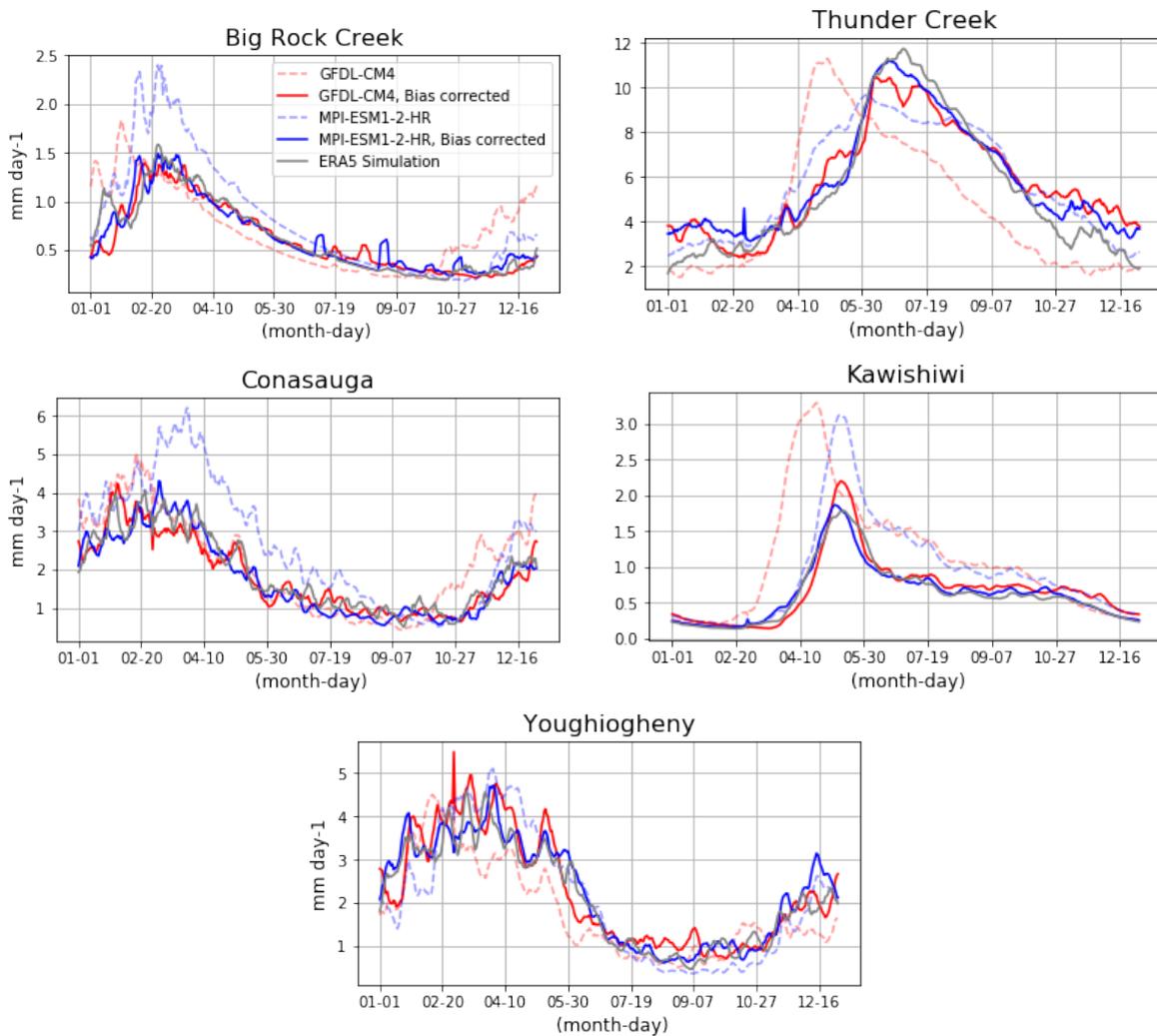


Figure D.2: Mean daily streamflow using ERA5 simulations, raw CMIP6 data and bias-corrected CMIP6 data

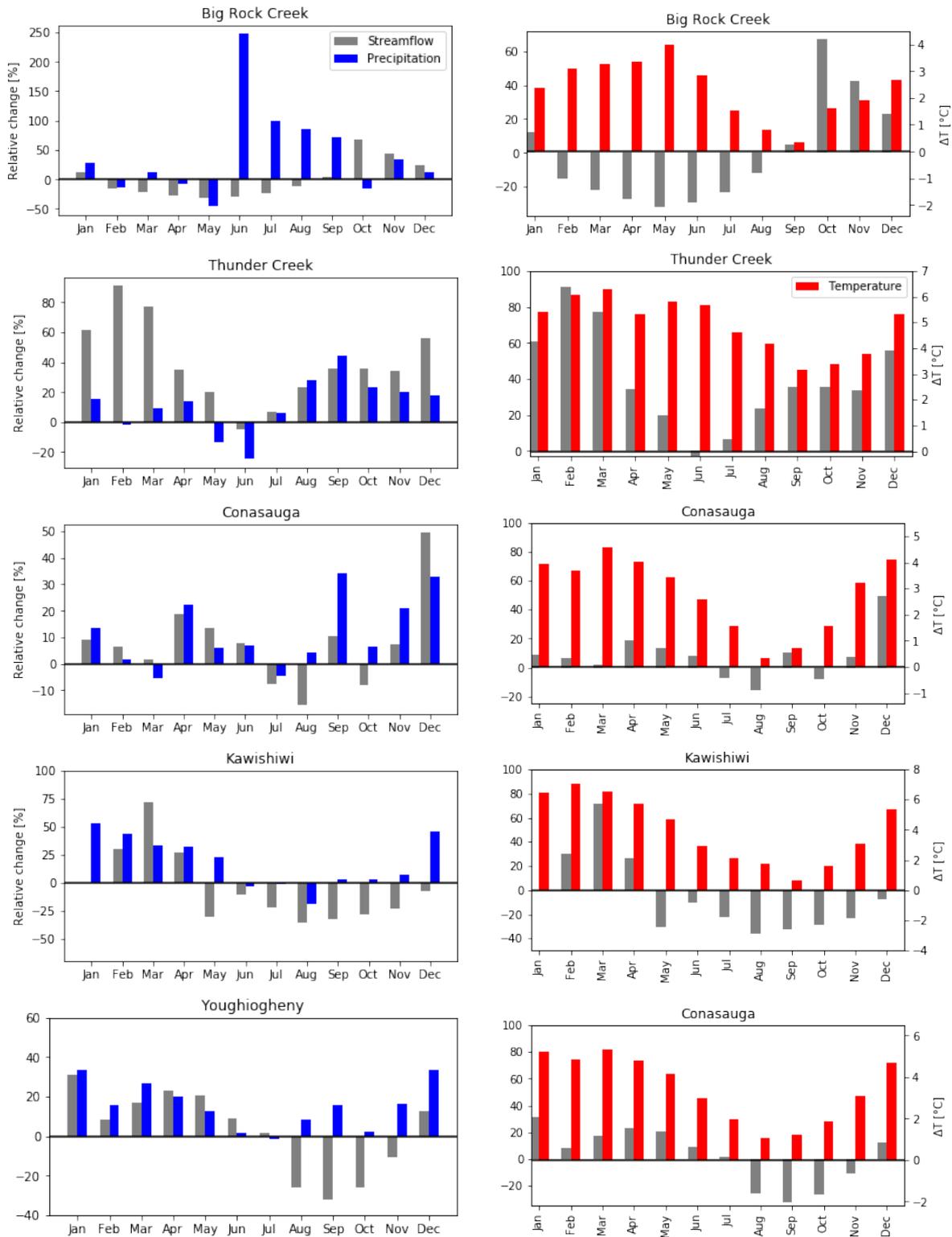


Figure D.3: Relative changes in streamflow compared to relative changes of precipitation and absolute change in temperature under the SSP245 scenario

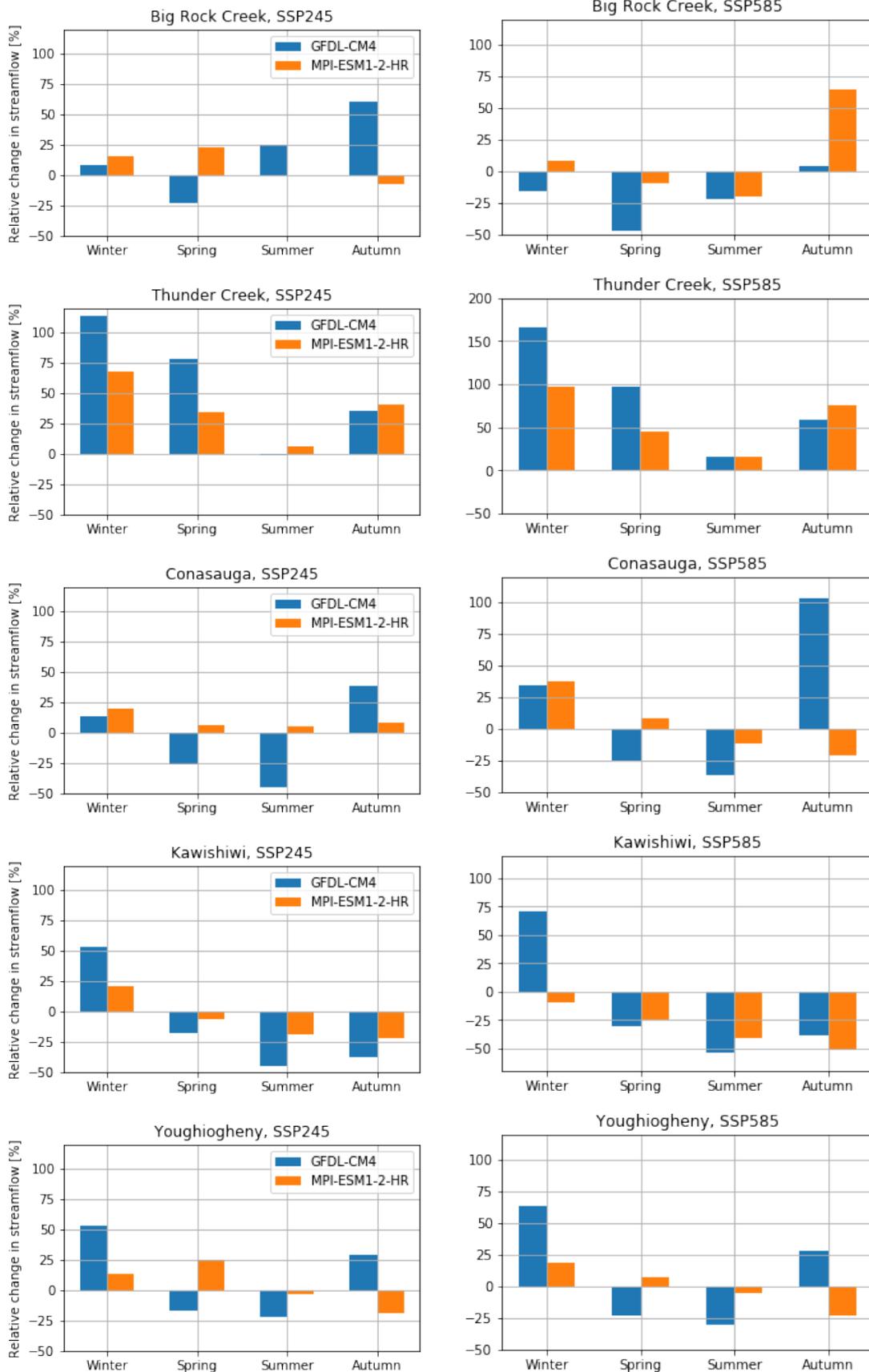


Figure D.4: Seasonal streamflow changes in all catchments between the GFDL-CM4 and MPI-ESM1-2-HR simulations under both climate scenarios

References

- Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., & Savenije, H. H. G. (2014). Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci.*, *18*(12), 4839–4859.
- Hanus, S. ., Hrachowitz, M. ., Zekollari, H. ., Schoups, G. ., Vizcaino, M. ., & Kaitna, R. . (2021). Future changes in annual, seasonal and monthly runoff signatures in contrasting alpine catchments in austria. *Hydrology and Earth System Sciences*, *25*(6), 3429–3453. <https://doi.org/10.5194/hess-25-3429-2021>
- Hoogelander, V. (2022). *vhoogelander/BMI_HBVmountain_streamflow_analysis: Thesis (Version v0.1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.7405678>
- Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., & Gascuel-Odoux, C. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resources Research*, *50*(9), 7445–7469. <https://doi.org/https://doi.org/10.1002/2014WR015484>
- Hutton, E., Piper, M., & Tucker, G. (2020). The basic model interface 2.0: A standard interface for coupling numerical models in the geosciences. *Journal of Open Source Software*, *5*, 2317. <https://doi.org/10.21105/joss.02317>
- Knoben, W. J. M., Freer, J. E., Fowler, K. J. A., Peel, M. C., & Woods, R. A. (2019). Modular assessment of Rainfall–Runoff models toolbox (MARRMoT) v1.2: An open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations. *Geosci. Model Dev.*, *12*(6), 2463–2480.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — a discussion of principles. *J. Hydrol. (Amst.)*, *10*(3), 282–290.
- Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., Vegas-Regidor, J., Bock, L., Brötz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Loosveldt Tomas, S., & Zimmermann, K. (2020). Earth system model evaluation tool (ESMValTool) v2.0 – technical overview. *Geosci. Model Dev.*, *13*(3), 1179–1199.
- sarah-hanus. (2021). *sarah-hanus/hbv-mountain: Semi-Distributed Hydrological Model in Julia Programming Language (Version v1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.4964641>
- Searcy, J. K. (1959). *Flow-duration curves* (-, tech. rep. 1542A). U.S. Govt. Print. Off.,