



Norwegian University of
Science and Technology

Department of Civil and Environmental Engineering

Evaluation of the performance of the hydrological model WEAP in Norway

Anna Szöllösi

BSc student in Civil Engineering

Supervisors:

Tor Haakon Bakken

Professor, NTNU Department of Civil and Environmental Engineering

Dr. Baranya Sándor

Associate professor, BME Department of Hydraulic and Water Resources Engineering

Budapest, 2025

PROJECT REPORT

Name:	Anna Szöllősi	Semester:	2024/25/2
-------	---------------	-----------	-----------

Title:	Evaluation of the performance of the hydrological model WEAP in Norway
--------	--

Task description:

Hydrological models transform climatic input into hydrological information such as streamflow, snow accumulation and snow melt, groundwater levels and more, and can be used as a basis for flood warning, drought calculations, instream habitat assessments and climate impact studies. The majority of the hydrological studies in Norway, including daily inflow calculations by the hydropower companies as well as the daily forecasting carried out by the national authorities, are based on the HBV-model, or some variant based on the concept of the HBV-model. HBV has the advantage that it is tailor-made for cold climate conditions, as we experience in Norway, but has limitations in representing the human interventions (e.g. hydropower operations, water withdrawals, etc.). WEAP is a model developed by the Stockholm Environment Institute (SEI) and has been applied to a large number of cases internationally, but to a limited extent in Norway. WEAP consists of a hydrological model describing the “supply” of water, and a set of routines describing the “demands”, i.e. the use of water. WEAP also has the advantage that climatic input data, a digital terrain model, land cover data and future climatic data can be directly downloaded for the selected case study region, as such can potentially be very quick to configure. The purpose of this study is to test the model for climatic and hydrological conditions in Norway.

The overall objective of the thesis is to evaluate the performance of the hydrological model WEAP in Norway. WEAP has mostly been used under conditions different than in Norway, i.e. with different land use and catchment characteristics and limited snow/glacier accumulation and runoff. In order to evaluate the performance of the model, the following tasks need to be carried out:

1. Identify a suitable catchment for the performance assessment
2. Compile all relevant data for the purpose of testing the performance
3. Analyze/quality check the data to be used, i.e. compare the built-in datasets in WEAP with the data available from national data sources in Norway
4. Calibrate/validate WEAP for the selected test catchment by use of suitable statistical criteria
5. Identify and describe quality of the calibration/validation
6. Examine the impact of climate change on the selected test catchment

Supervisors:

Professor Tor Haakon Bakken	NTNU	100 %
Baranya Sándor	BME Department of Hydraulic and Water Resources Engineering	

Table of Contents

1.	Introduction	5
1.1.	Overview	5
1.2.	Research Questions	7
1.3.	Research Methodology	7
2.	Data and method	8
2.1.	Generalities and presentation of the WEAP model	8
2.2.	Søgne basin	10
2.3.	Available data, division of time period	11
2.4.	Examined statistical variables	13
2.5.	Workflow	14
2.6.	Future climate scenarios	16
3.	Results and evaluation	19
3.1.	Data sources and initial analyses	19
3.1.1.	Precipitation data comparison	19
3.1.2.	Temperature data comparison	22
3.1.3.	Snow depth comparison	24
3.1.4.	Conclusion of the evaluation of climate data	25
3.2.	Calibration of the model for the whole period	25
3.2.1.	Calibration of Kc	25
3.2.2.	Calibration of other parameters	27
3.2.3.	Optimal parameter values	31
3.2.4.	Calibration of melting and freezing point	32
3.3.	Seasonal calibration of the model for summer and winter periods	39
3.3.1.	Calibration of Kc	40
3.3.2.	Calibration of other parameters	41
3.3.3.	Calibration of Melting and Freezing Points	45
3.3.4.	Optimal parameter values	46
3.3.5.	Equifinality – GLUE	48
3.4.	Model validation	49
3.5.	Assessment of future climate scenarios	50
4.	Conclusion	55
4.1.	Data sources and initial analyses	55
4.2.	Calibration of the model for the whole period	55
4.3.	Seasonal calibration of the model for summer and winter periods	56
4.4.	Model validation	56
4.5.	Assessment of future climate scenarios	57
5.	References	59

Abstract

Name: Anna Szöllösi

Title of the thesis: Evaluation of the performance of the hydrological model WEAP in Norway

Supervisors: Professor Tor Haakon Bakken
NTNU Department of Civil and Environmental Engineering
Baranya Sándor
BME Department of Hydraulic and Water Resources Engineering

Semester: 2024/25/2

Water resource management plays a crucial role in ensuring sustainable development, especially in regions affected by climate change, such as Norway. The aim of this study was to assess the applicability of the WEAP (Water Evaluation and Planning) model under Norwegian climatic and hydrological conditions. As part of the research, the Princeton dataset built in WEAP was compared with meteorological data from the Norwegian SeNorge and NKSS sources. The results showed that the local SeNorge data more accurately reflect precipitation and temperature characteristics, and were therefore used as input for the model.

The model was calibrated for the period 1987–2014, achieving moderate accuracy ($NSE = 0.553$; $PBIAS = -6.209\%$). A separate calibration for the summer period yielded improved performance ($NSE = 0.719$), whereas the model calibrated for the winter months performed less reliably ($NSE = 0.496$). The model was validated for the 1974–1986 period ($NSE = 0.483$), indicating a limited but usable predictive capability outside the calibration range.

For future streamflow simulations three global climate models and multiple SSP (Shared Socioeconomic Pathways) scenarios were used to project changes for the 2060–2090 period. The simulations suggest a 29–39% decrease in average annual streamflow compared to the reference period. A particularly notable decline is projected for spring flood peaks, due to reduced snow accumulation and earlier snowmelt. The significant summer streamflow reduction implies increased drought risk and water supply uncertainty.

The findings highlight the impacts of climate change on Norway's water resources and emphasize the need for adaptive development of water management systems. The study underscores the importance of locally calibrated hydrological models and high-resolution regional meteorological data for reliable forecasting of future water management challenges.

1. Introduction

1.1. Overview

Water has always been a significant part of human life. It is well understood that without water, there is no life. Water is present in every subtle cycle of our lives; there is no area where it does not appear. Over time, it has increasingly become a part of the economy, as it is used not only for drinking water supply and irrigation but also for electricity generation through hydropower plants.

In recent decades, due to climate change and population growth, water has become an increasingly central issue. Demand has significantly increased, leading to new challenges in water management. It is crucial to address these challenges within the framework of sustainable economic development and to define future goals accordingly. In this context, the goal of sustainable water management for today's society should be the utilization of resources in a way that ensures their availability for future generations. We bear a significant responsibility in this regard, and several important objectives can be articulated: sustainable management and preservation of water resources, optimization of water consumption, maintenance of water quality, and development of water management infrastructure. In water management, it is essential to ensure that all sectors – including industry, agriculture, households, and natural ecosystems – have access to water while reducing unnecessary water use and preventing the depletion of water resources. To address future challenges, flexible water management systems are necessary to ensure the efficient and sustainable use of water resources.

In Norway, where more than 70% of the country's major rivers are used for electricity generation through hydropower plants, water plays a prominent role. It is important to understand the timing and extent of extreme water flows. Dry periods in southern Norway mainly occur during the summer, due to high temperatures and low precipitation. In northern regions, dry periods are more common in winter, when precipitation is stored as snow until the snow melts in spring. Flood periods mainly occur during the rainy autumn months and during the spring snowmelt.

Today, when addressing challenges, we cannot overlook climate change, which is making extreme events - floods, droughts - more and more frequent. These issues are of great significance for the protection of life and property, as well as for the preservation of ecosystems. In Norway, the average annual temperature is projected to increase by 1.6 to 4.6 °C by the end of the century, while average annual precipitation is expected to increase by 8–18% (Hanssen-

Bauer et al., 2017) As a result of climate change, floods caused by rainfall are expected to become more frequent, while floods resulting from snowmelt may become less common. Snow-covered months are likely to be significantly shorter, and due to increased temperatures, snowmelt is expected to occur earlier (Hanssen-Bauer et al., 2017). In areas where autumn and winter floods are already dominant, significant increases in flood peaks are anticipated. In some regions where snowmelt-induced floods have been typical, floods caused by autumn and winter rainfall may play an increasingly significant role in the future (Lawrence & Hisdal, 2011). Estimating and forecasting these high-water flows are crucial to aid in flood management.

The complexity of water management necessitates the use of various technologies, as different models can help to manage water efficiently and sustainably. For example, hydrological models that simulate the movement and distribution of water under various environmental conditions are fundamental tools in water management. These models can be used to analyze and identify the key processes of water balance in a river basin and thus determine the most appropriate water management strategy to ensure sustainability.

Hydrological models can also aid in the integrated management of water management systems. In Norway, the most widely used hydrological model is the HBV model (Hydrologiska Byråns Vattenbalansavdelning), developed by the Swedish Meteorological and Hydrological Institute (SMHI) in 1972. A significant advantage of this model is its applicability in cold climates, where snow and ice melt play a crucial role in the water balance of catchment areas. However, a drawback is that the model does not account for human interventions, such as hydropower plants or water withdrawals, and therefore may not always provide an accurate representation of reality (Bergström, 1976, 1995). In contrast, the WEAP model (Water Evaluation and Planning) offers the possibility for more complex management of water supply and consumption. The WEAP software was developed by the Stockholm Environment Institute (SEI) in 1988. The software allows the modelling of water management depending on the available data quality and quantity. Additionally, it can model various future water management scenarios to examine how certain climate scenarios affect water distribution and quantity (Sieber, 2006). The use of WEAP is common in Africa, China, America, Europe, and around the world. Applications of WEAP model can be seen in research of Bañares et al. (2024), Li et al. (2015), Mahamadou Mounir et al. (2011), Sardar Shahraki et al. (2016), Winter et al. (2017). It is applied for estimating water management strategies, calculating future water demands, assessing the impacts of climate change, determining irrigation water requirements, and much more.

Although the WEAP model has been applied in Norwegian contexts (Olabiwonnu, 2020; Olabiwonnu et al., 2022; Pramudith et al., 2022), no one has yet examined its accuracy. In Scandinavian countries, the HBV model is much more prevalent and handles climate data more accurately; however, due to its shortcomings (not accounting for human interventions), I considered it worthwhile to examine the more complex WEAP model. In this thesis, the following was examined: how the WEAP model handles the Norwegian climate and snowmelt, whether the model can be calibrated to suit cold-weather conditions.

The aim of my thesis is to test the WEAP model under Norway's specific climatic and hydrological conditions, with particular attention to snow and ice accumulation, variations in water flow, and the applicability of modeling in managing floods and low water flows. In my research, I will examine how the WEAP model can be applied in Norwegian water management practices and how various hydrological and climatic factors influence the model's performance.

1.2. Research Questions

1. Is the WEAP model applicable in Norway?
 - a. How accurately does the model represent snow accumulation and snowmelt processes?
 - b. How reliable is the built-in climate data provided by the model?
 - c. Can the model be calibrated appropriately to be transferable to other catchments with similar characteristics?
2. What is the impact of climate change on characteristic streamflow in Norway, and is it possible to model this using the WEAP software?

1.3. Research Methodology

1. Identify a suitable catchment for the performance assessment
2. Compile all relevant data for the purpose of testing the performance
3. Analyze/quality check the data to be used, i.e. compare the built-in datasets in WEAP with the data available from national data sources in Norway
4. Calibrate/validate WEAP for the selected test catchment by use of suitable statistical criteria
5. Identify and describe quality of the calibration/validation
6. Examine the impact of climate change on the selected test catchment

2. Data and method

2.1. Generalities and presentation of the WEAP model

The WEAP (Water Evaluation and Planning System) model was developed by the Stockholm Environment Institute (SEI) in 1988. WEAP is an integrated hydrological and planning model that combines watershed-level hydrological processes with water management systems. It is capable of directly handling climate data and simultaneously modeling both water demand and supply. With GIS support, built-in ecological and economic analysis modules, and compatibility with other models (e.g., MODFLOW, QUAL2K), it enables complex simulations (Sieber, 2006).

WEAP operates on the principle of water balance calculations and is widely applicable to a range of systems – whether for urban or agricultural water management, smaller catchments, or complex river systems. It can address a wide range of problems, such as sectoral water demand analysis, water rights and allocation priorities, groundwater and river simulations, reservoir operations, hydropower management, pollution monitoring, water needs of ecosystems, etc.

The application of WEAP begins with defining the scope of the study, which includes specifying the time period, spatial boundaries, and system components. The second step is to establish the current situation ("Current Accounts"), entering current water demand, pollution, water resources and infrastructure data. Afterwards, future scenarios are developed based on different policies, costs and impact factors, and then these scenarios are compiled by comparing different assumptions and regulatory options. The final step involves evaluating the adequacy of water supply, cost-benefit analysis, compliance with environmental goals, and uncertainty assessment (Sieber, 2006).

A key foundation of the model is the development of user-defined scenarios. After recording the current situation ("Current Accounts"), a "Reference Scenario" or "business-as-usual" can be developed based on different economic, demographic, hydrological and technological trends. From this point, additional scenarios can be created based on different policies or development directions, which allow exploration of various future assumptions. These scenarios are based on different "What if?" questions. What happens if the rate of population growth or economic development changes? What happens if new sources of pollution emerge? What is the impact of more efficient irrigation techniques? How does climate change affect the hydrology of the system? The different scenarios can be displayed side by side and compared to see how each change affects the water management system (Sieber & Purkey, 2015).

In WEAP, there are five different methods available for modeling processes - such as evaporation, runoff, infiltration, and irrigation demand - in watershed areas. These methods are (1) the Rainfall Runoff and (2) Irrigation Demands Only versions of the Simplified Coefficient Approach, (3) the Soil Moisture Method, (4) the MABIA Method, and (5) the Plant Growth Model or PGM. The choice of method is influenced by the complexity of the research and the data available.

In this thesis, the Soil Moisture Method was chosen, which offers an appropriate level of complexity, is widely used, and has the important feature of allowing snow accumulation modeling. The method interprets the catchment area as two soil layers. The different variables and the principle of the method are presented in Figure 1. Precipitation and melted snow reaches the ground in the form of water, part of which runs off as surface runoff and part of which infiltrates into the topsoil. In this layer, some of the water continues to flow to the lower soil layer, while some of it accumulates as groundwater. Finally, the remaining water from the lower soil layer also collects as groundwater.

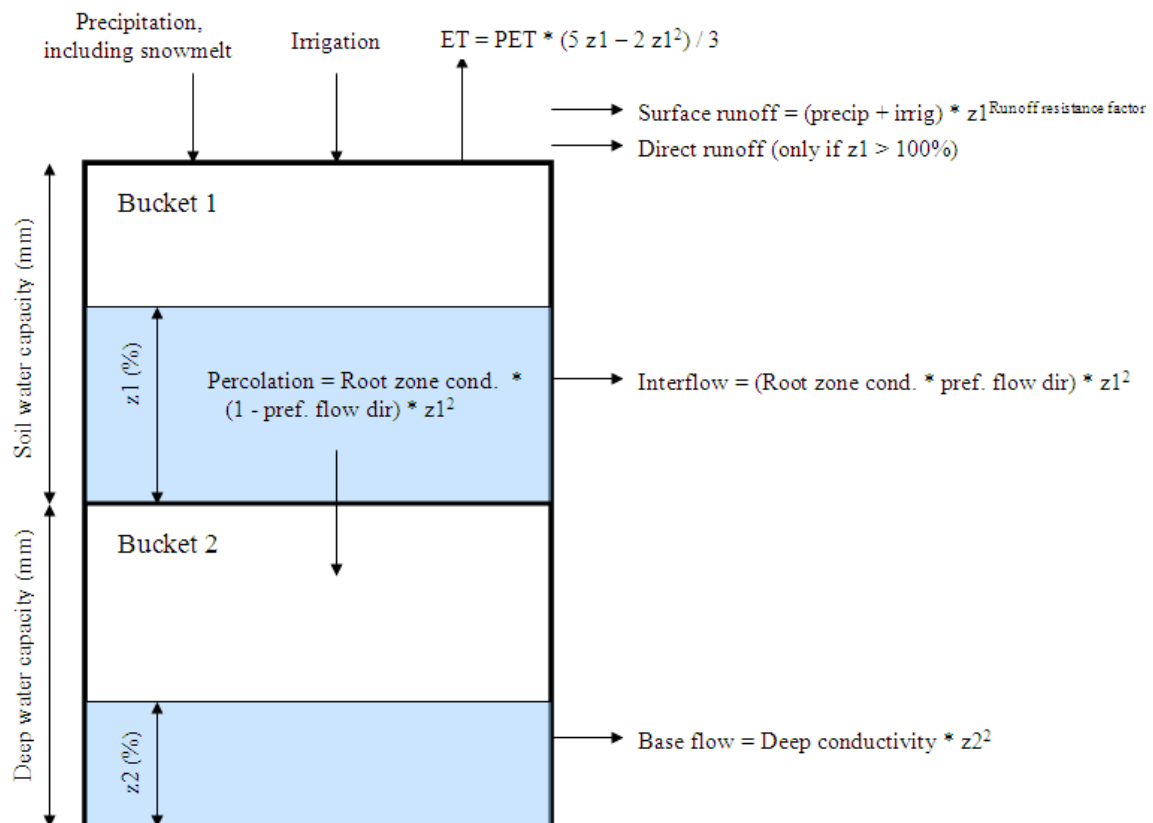


Figure 1 The two-bucket model (Sieber & Purkey, 2015)

2.2. Søgne basin

For building the model, the Søgne River (also known as *Songdalselva*, *Søgneelva*, or *Songa*) was selected. It is approximately 55 km long and located in southern Norway. A key characteristic of this river is that it is unregulated, meaning it maintains its natural flow regime and conditions, which makes it particularly suitable for hydrological modeling. The availability of climate data for rivers in southern Norway within the WEAP software also played an important role in selecting this particular river for the study. In addition, the Norwegian Water Resources and Energy Directorate (NVE) defined a Hydrological Reference Dataset (HRD), which lists rivers that are suitable for climate change modeling based on both regulation status and availability of measured data (Bækø J, Pedersen I, 2024). The Søgne River is included in this list and is therefore considered appropriate for climate change studies.

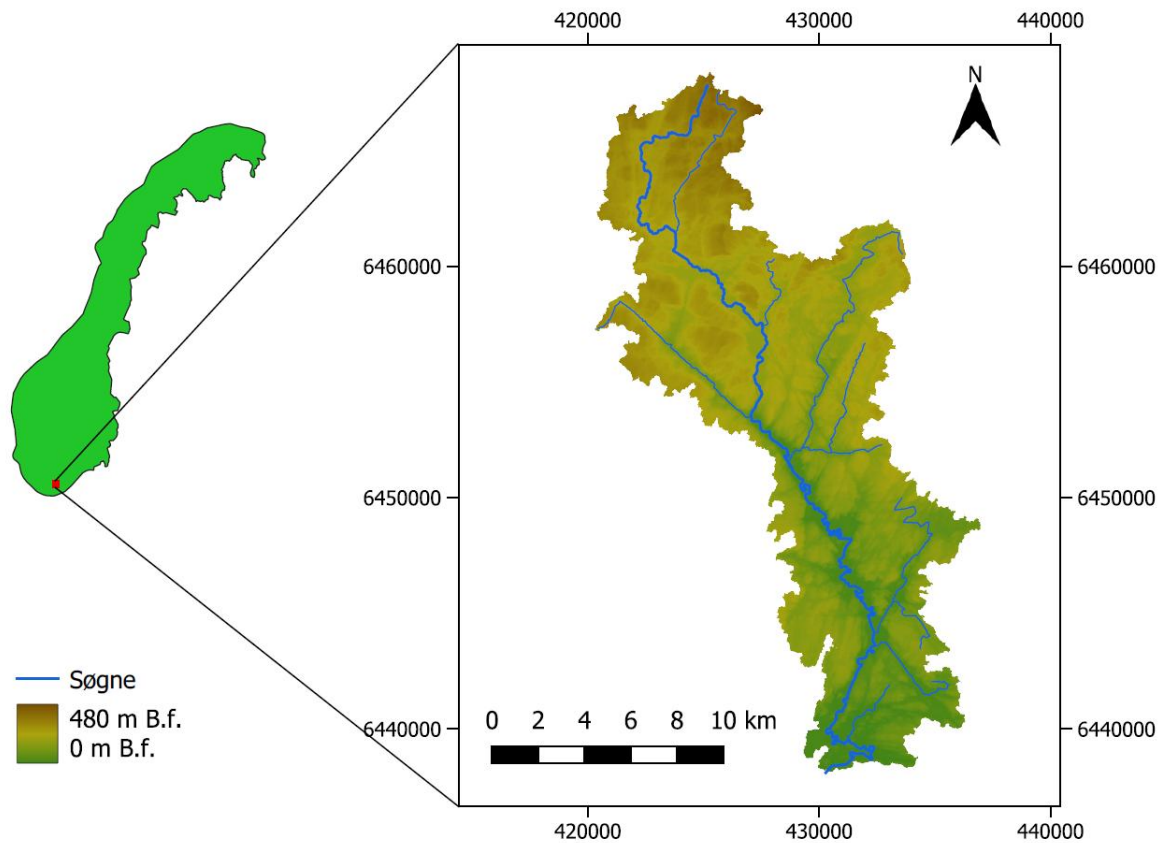


Figure 2 Catchment area of the river Søgne

Figure 2 shows the location of the river within the territory of Norway. The elevation of the catchment area is between 0 and 500 meters. The catchment area is 209 km², and the land use distribution within the catchment is presented in Table 1.

Table 1 Distribution of land use

Land use	Amount (%)
Forest	80,4
Agriculture	5,3
Wetland	5,8
Surface water	4,2
Other	4,3

2.3. Available data, division of time period

During the research, multiple data sources were used, the types and origins of which are summarized in Table 2. The climate datasets integrated into the WEAP modeling software are from a database compiled by Princeton University, which includes variables for precipitation, temperature, relative humidity, wind speed, cloud cover, freezing and melting points, among others.

In order to achieve the highest possible accuracy, I aimed to conduct the analyses primarily at a daily resolution. This required the use of high-resolution meteorological data from Norwegian sources.

The SeNorge platform provided reliable, daily resolved precipitation, temperature and snow depth data, and the Norsk Klimaservicesenter (NKSS) website provided high-quality daily precipitation data. The observed streamflow data required for model calibration were obtained from the NVE Sildre (Norges vassdrags- og energidirektorat) database, which contains measurements from a local gauging station located on the Søgne River.

Table 2 Source and duration of data

Data type	Source	Period (mm/dd/yyyy)		Comment
Precipitation	Princeton	1/1/1948	12/31/2014	
	SeNorge	1/1/1957	Today	
	NKSS	1/1/1948	6/30/2024	
Temperature	Princeton	1/1/1948	12/31/2014	
	SeNorge	1/1/1957	Today	
Snow depth	SeNorge	1/1/1957	Today	
Streamflow	NVE-Sildre	12/14/1973	Today	1992 missing

Table 2 not only presents the types and sources of the data used but also indicates the time periods for which each dataset was available. Figure 3 provides an illustrative graphical representation of these time intervals. The figure shows that the time spans covered by the various data sources differ; however, the period from 1974 to 2014 is the only one during which all the necessary data were simultaneously available. Therefore, this time range provided the most complete and reliable basis for comparative and calibration analyses.

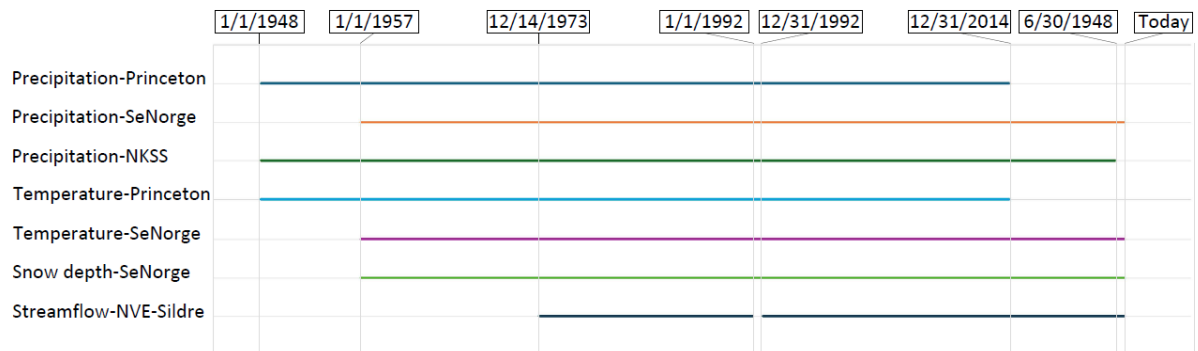


Figure 3 Available time intervals of different datasets

To assess the reliability and predictive capability of the model, I divided the available dataset for the period between 1974 and 2014 into two parts: the period from 1974 to 1987 was used for validation, while the period from 1987 to 2014 was used for calibration. This division follows internationally accepted practice, which recommends allocating approximately one-third of the full dataset for model validation and the remaining two-thirds for calibration (Moriassi et al., 2015).

This approach allows an objective assessment of the model's performance over a period of time unaffected by calibration. It also reduces the risk of "overfitting", i.e. the possibility that the model may become overly tailored to the calibration data and perform poorly during other periods. The methodological principle is supported by several studies and recommendations, emphasizing that the spatial and temporal separation of calibration and validation is a fundamental requirement for the correct application of hydrological models (Refsgaard & Knudsen, 1996).

2.4. Examined statistical variables

Like all hydrological models, the WEAP model needs to be calibrated. The model's adequacy is assessed by comparing simulated values with observed data. Several statistical variables can be used for this comparison; in my research, I applied the NSE (Nash-Sutcliffe Efficiency) and PBIAS (Percent Bias) parameters.

The NSE evaluates how well the simulated values match the observed ones. It indicates how well the graph of observed and simulated data fit the 1:1 line. The NSE is calculated with equation 1:

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2} \right] \quad (1)$$

where Y_i^{obs} is the i-th measured value, Y_i^{sim} is the i-th simulated value, Y^{mean} is the average of the observed data, and n is the total number of observations.

Its value ranges between $-\infty$ and 1, where 1 is the optimal value. A value between 0 and 1 is considered acceptable, while a value below 0 means that estimating by the mean would give a more accurate value (Nash & Sutcliffe, 1970). The categories shown in Table 3 are the guidelines for the evaluation (Moriassi et al., 2015).

The PBIAS (percent bias) expresses the extent to which the model overestimates or underestimates a given variable compared to observations (Gupta et al., 1999). The optimal value of PBIAS is 0; positive values indicate underestimation by the model, while negative values reflect overestimation. The interpretation is also guided by the categories shown in Table 3 (Moriassi et al., 2015). The PBIAS value is calculated with equation 2:

$$PBIAS = \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim}) \cdot 100}{\sum_{i=1}^n (Y_i^{obs})} \right] \quad (2)$$

where Y_i^{obs} is the i-th measured value, Y_i^{sim} is the i-th simulated value, and n is the total number of observations.

Table 3 Categories of model evaluation

Performance Rating	NSE	PBIAS (%)
Very good	$0,80 < \text{NSE}$	$\text{PBIAS} < \pm 5$
Good	$0,70 \leq \text{NSE} \leq 0,80$	$\pm 5 \leq \text{PBIAS} < \pm 10$
Satisfactory	$0,50 < \text{NSE} < 0,70$	$\pm 10 \leq \text{PBIAS} < \pm 15$
Unsatisfactory	$\text{NSE} \leq 0,50$	$\text{PBIAS} \geq \pm 15$

Therefore, the model is considered satisfactory if the conditions $\text{NSE} > 0.50$ and $\text{PBIAS} \geq \pm 25$ are met (Moriassi et al., 2015).

2.5. Workflow

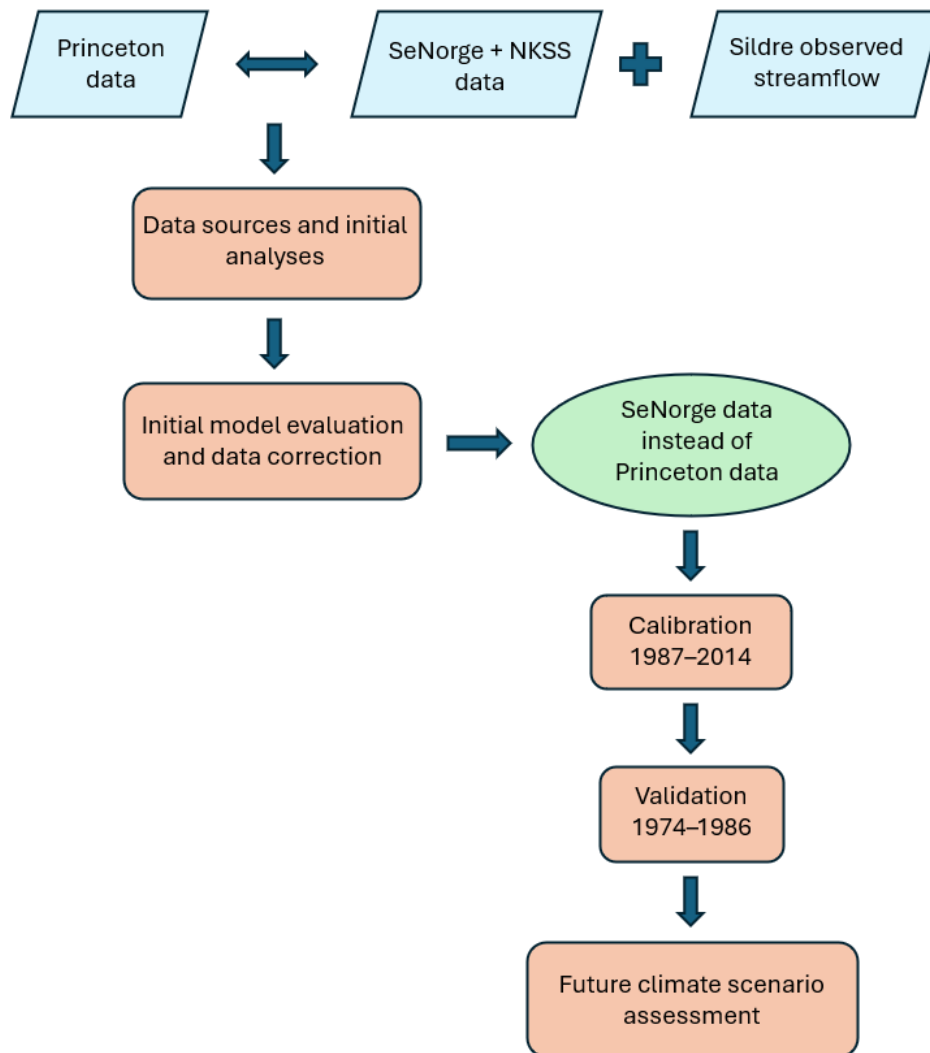


Figure 4 Workflow method of the WEAP model calibration, validation, and future climate scenario assessment

Data sources and initial analyses

As the first step of the modeling process, the climate data integrated into the WEAP software was examined, which originates from Princeton University's database. While these data sets are widely applicable on a global scale, their reliability for use in Norway had not been previously validated. The available dataset covers the period from 1973 to 2014 and includes variables such as precipitation, temperature, humidity, wind speed, and snow depth.

To assess the reliability of these data, I compared the Princeton dataset with:

- Observed streamflow values from the Sildre database,
- Precipitation, temperature, and snow depth measurements from local stations on the SeNorge platform,
- Precipitation data provided by the Norsk Klima Service Senter.

Initial model evaluation and data correction

The initial simulation results, based on Princeton input data, were compared with the observed streamflow. To evaluate the model's performance, two statistical indicators were applied: the Nash–Sutcliffe Efficiency (NSE), and the Percent Bias (PBIAS). These indicators were calculated both for the entire modeling period and seasonally, with separate assessments for the summer and winter periods.

The initial results indicated that the calculated statistical indicators deviated significantly from the generally accepted reference values. As a result, the Princeton dataset was replaced with precipitation and temperature data from the SeNorge database, which was deemed more reliable. Simulations and performance evaluations were then repeated using the updated inputs. Although the recalculated indicators showed notable improvement further model calibration was needed.

Calibration and validation of the model

Model calibration was conducted in multiple stages: first over the full modeling period, and then separately for the summer and winter seasons. The available data covered the period 1974-2014. The calibration was performed for two thirds of the total period (1987-2014), while the remaining one third (1974-1987) was used for model validation.

The calibrated model produced satisfactory results. During the validation phase the performance of the model was examined to ensure the applicability and stability of the model to unknown periods.

Future climate scenario assessment

As the final step in the analysis, future climate scenarios were evaluated based on outputs from various global climate models (GCMs). Their potential impacts on streamflow data were assessed, with particular attention given to low-flow frequency and magnitude, as well as changes in annual average discharge. Based on these results, conclusions were drawn regarding the expected evolution of future hydrological conditions.

2.6. Future climate scenarios

Considering climate change has become a fundamental requirement in both water management and ecosystem protection. Forecasting future streamflow is essential for planning hydropower utilization, modeling ecological processes influenced by low flows (such as fish population survival), and conducting flood risk forecasting and risk analysis.

Shared Socioeconomic Pathways (SSPs)

In assessing future climate scenarios, understanding and applying the Shared Socioeconomic Pathways (SSPs) is essential. The SSPs describe possible future developments in society and economy, which serve as the foundation for simulating future greenhouse gas emissions and climate impacts in climate models. Developed under the CMIP6 project, these scenarios are widely used for climate impact assessments, as well as for evaluating adaptation and mitigation strategies (O'Neill et al., 2016).

The SSPs combine different societal and economic scenarios with various climate policy pathways, quantifying changes in resource use, energy consumption, technological development, population growth, and environmental awareness.

In this study three SSP scenarios were applied, representing different levels of future climate change:

- SSP1-2.6 – A sustainable development pathway characterized by low population growth, environmentally friendly technologies, and strong international cooperation. This is one of the most optimistic scenarios, aiming to limit global warming below 2°C.

- SSP2-4.5 – A middle path scenario, assuming continuation of current trends. It reflects moderate social and economic change and moderate emission reductions, resulting in a mid-range level of climate change.
- SSP5-8.5 – A pessimistic scenario based on rapid economic growth and heavy reliance on fossil fuels. This leads to the highest levels of greenhouse gas emissions and potentially substantial global warming (up to 4–5°C).

Applying these scenarios enables the evaluation of potential future climatic conditions under varying socioeconomic and emission trajectories, with particular focus on their effects on streamflow.

Applied climate models

In this thesis, the objective was to model future streamflow under different climate scenarios using a hydrological model that was calibrated and validated for a Norwegian catchment. For this purpose, three global climate models available within the WEAP system were selected: NorESM2-LM, MPI-ESM1-2-HR, and HadGEM3-GC31-MM.

The NorESM2-LM (Norwegian Earth System Model, version 2, Low Resolution) is a climate model developed by the Bjerknes Centre for Climate Research and the Norwegian Meteorological Institute. It has been specifically optimized for high-latitude, Arctic, and Scandinavian conditions. A comprehensive description of the model is provided by Seland et al. (2020), highlighting its capabilities in simulating temperature, precipitation, and sea ice patterns. Due to its region-specific development and strong performance in Scandinavia, this model was considered a natural choice for studying a Norwegian river.

The MPI-ESM1-2-HR is a high-resolution global climate model developed by the Max Planck Institute for Meteorology. It has been applied and validated within the CMIP6 (Coupled Model Intercomparison Project Phase 6) and HighResMIP (High Resolution Model Intercomparison Project) frameworks. Müller et al. (2018) and Gutjahr et al. (2019) emphasize its excellent performance in simulating temperature and precipitation, particularly across Europe and the Scandinavian region. Its high spatial resolution makes it well-suited for regional impact assessments.

The HadGEM3-GC31-MM is part of the third generation of global climate models developed by the UK Met Office Hadley Centre. As described by Andrews et al. (2020), the model was developed under the CMIP6 project and is widely used in European climate scenario analyses,

including Scandinavia. Thanks to its medium spatial resolution and advanced atmospheric parametrization, it is well-suited for supporting regional hydrological studies.

Methodology for Climate Scenario Analysis

To analyze future streamflow changes, daily data were used from the three global climate models (NorESM2-LM, MPI-ESM1-2-HR, HadGEM3-GC31-MM) integrated into the WEAP system. These datasets were available from 2015 to 2100 and were generated according to the different SSP emission scenarios described above.

For evaluation, the 30-year future period of 2060–2090 was selected. This timeframe aligns with the IPCC’s recommendations for climate trend analysis, as longer periods help smooth out natural climate variability, reducing the influence of short-term fluctuations (IPCC, 2021). As a result, the observed changes more accurately reflect long-term climate trends rather than year-to-year variability.

The simulated future streamflow was compared with the past reference period of 1980–2010 simulated by the model. It is important to note that historical reference data were not based on observed measurements, but on the model’s historical simulations. This approach avoids inconsistencies that could arise from combining observed and modeled datasets, which may have structural differences. By comparing both periods within the model’s own framework, the analysis maintains internal consistency, which is crucial for reliable estimation of relative changes.

3. Results and evaluation

3.1. Data sources and initial analyses

The WEAP software provides access to a built-in climate database created by the Terrestrial Hydrology Group at Princeton University. The database contains climate data for the period between 1948 and 2014. Although this data source is widely used in various hydrological and climate modelling studies, it has not been proven that the data can be used with sufficient accuracy in a Norwegian context. Accordingly, validation of the database was considered essential as an initial step in this research.

The data validation was carried out by comparing data from the Princeton University database with data from local Norwegian monitoring stations. As a reference, the SeNorge database was used, which contains precipitation, temperature and snow depth data. The Norsk Klima Service Senter database was also used, which provides local precipitation data. The comparisons have been carried out using Microsoft Excel, applying different visualization techniques such as graphs, pivot tables and histograms.

3.1.1. Precipitation data comparison

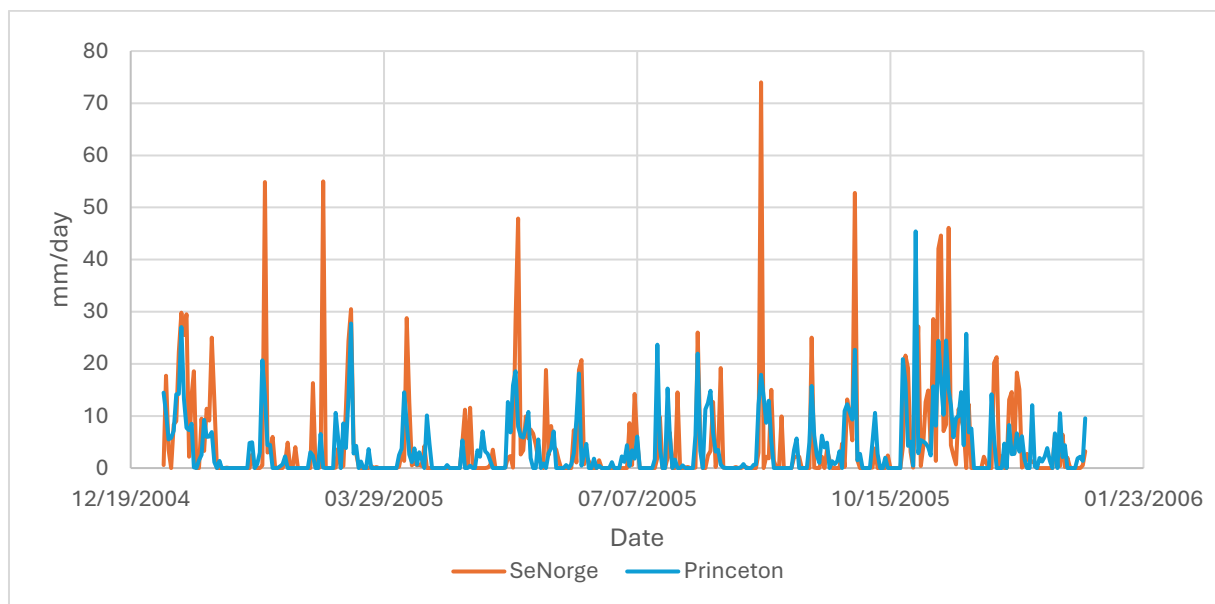


Figure 5 SeNorge and Princeton precipitation data in 2005

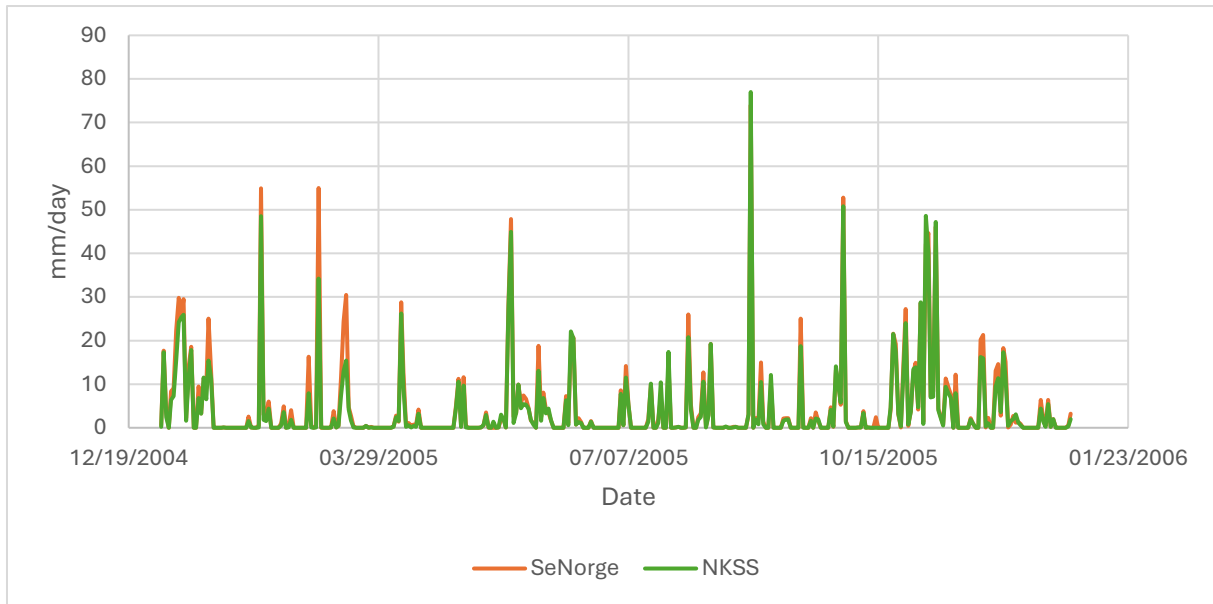


Figure 6 SeNorge and NKSS precipitation data in 2005

Figures 5 and 6 illustrate the precipitation data for a randomly selected year. The SeNorge and Norsk Klima Service Senter measurements show a close correlation: the precipitation maximums occur at the same time, and although there are small differences in intensity, the overall trend of the data is the same.

In contrast, precipitation data from the Princeton database show significant deviations from SeNorge values. The timing and intensity of precipitation peaks do not align clearly, raising concerns about the applicability of the data for hydrological modeling in Norway.

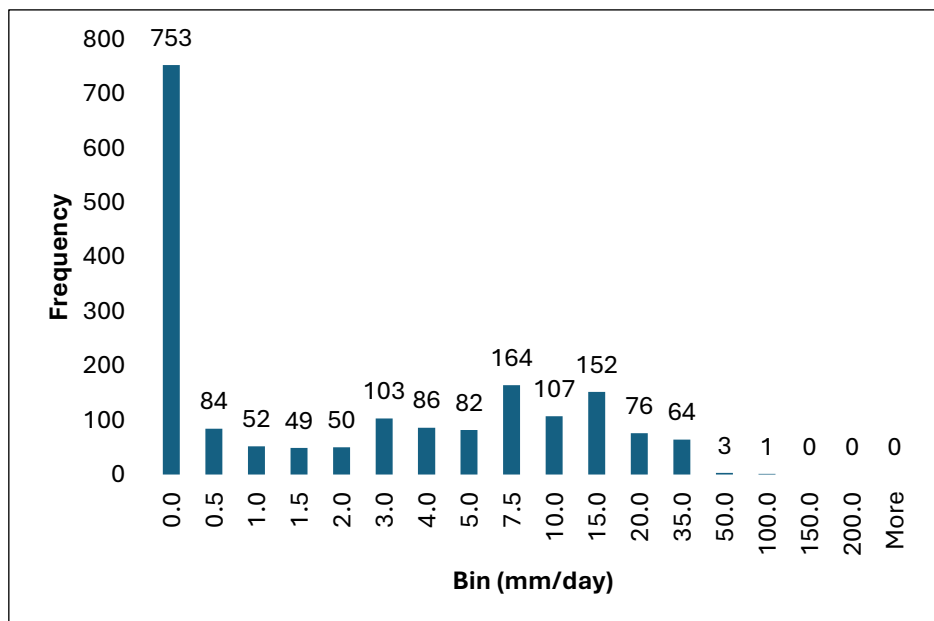


Figure 7 Princeton precipitation data distribution 2005-2009

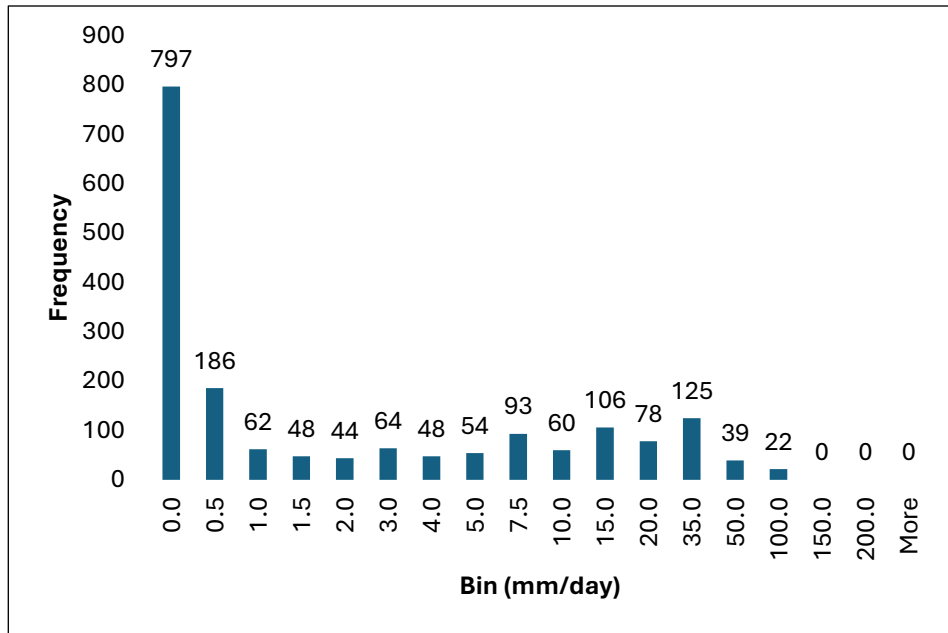


Figure 8 SeNorge precipitation data distribution 2005-2009

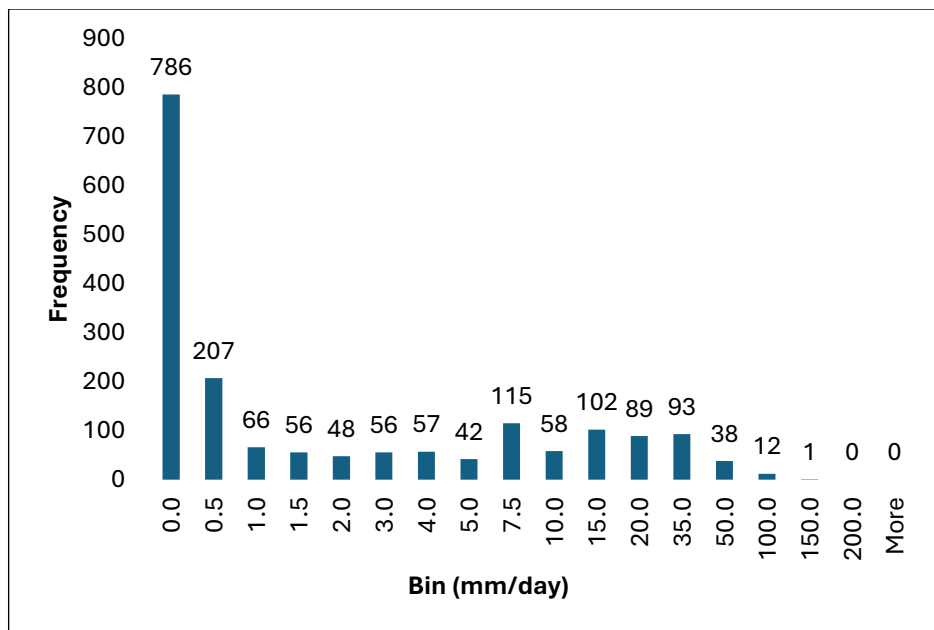


Figure 9 NKSS precipitation data distribution 2005-2009

The histograms (Figure 7-8-9) were created by dividing the data into five-year intervals. The results show that the data from SeNorge and Norsk Klima Service Senter follow a similar distribution, which confirms the consistency of the two data sources. In contrast, the Princeton database shows a significantly different trend.

The precipitation distribution of Princeton data shows a steadier decline, which may indicate that the dataset is smoother or less sensitive to local precipitation intensities. Lower

precipitation values are less frequent in the Princeton database, while the local station data tend to follow the more extreme weather patterns of the northern areas.

Table 4 Annual precipitation totals

Year	Sum / Princeton (mm/year)	Sum / SeNorge (mm/year)	Sum / NKSS (mm/year)
1975	1388.93	1781.60	1565.00
1980	1198.59	1656.50	1284.30
1985	1336.66	1769.10	1474.80
1990	1969.22	2493.00	2250.90
1995	1393.67	1679.00	1456.20
2000	1730.66	2830.10	2589.10
2005	1392.37	1653.10	1439.80
2010	1158.81	1377.10	1235.40
2014	1767.85	2527.50	2495.30

In Table 4 the summarized values represent sums for every year, and in the table every fifth year is presented. A tabular comparison of annual precipitation totals confirms that the SeNorge and Norsk Klima Service Senter data show a strong correlation, while the Princeton database systematically records lower annual precipitation totals. This suggests that the Princeton database may underestimate precipitation levels in Norway, potentially distorting hydrological modeling results.

3.1.2. Temperature data comparison

In addition to precipitation data, temperature data were also compared. In this case, data were only available from the Princeton and SeNorge databases.

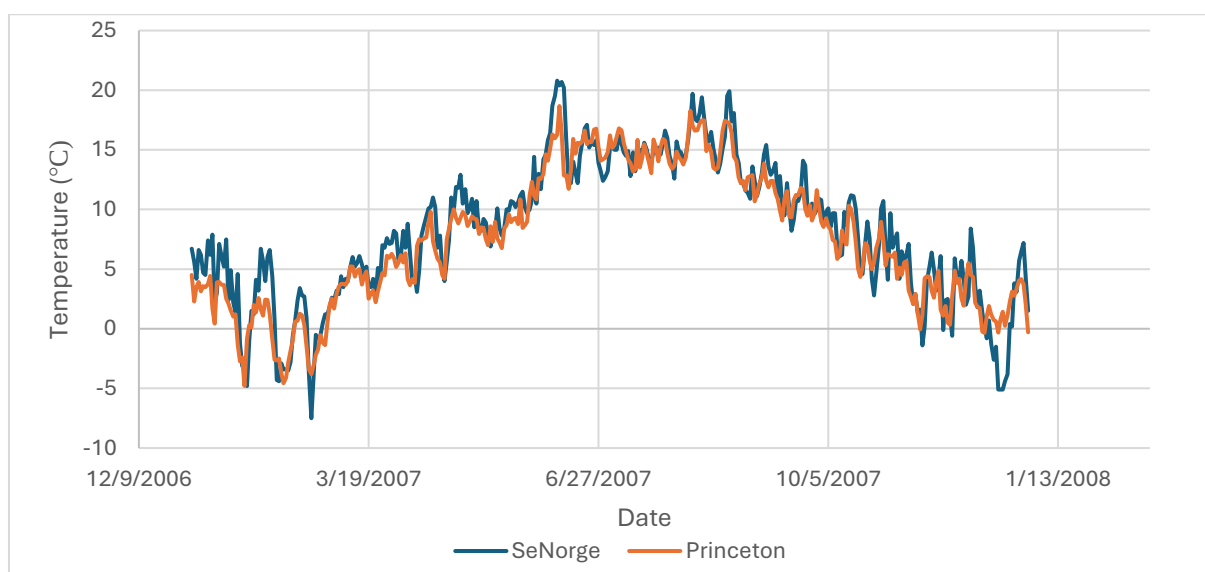


Figure 10 SeNorge and Princeton temperature data in 2007

Figure 10 illustrates the annual temperature variation for a selected year. The trend in the Princeton database follows the seasonal variations well, but there are minor differences, especially in extreme weather conditions. The SeNorge data, on the other hand, seems to be more accurate, as they show larger fluctuations, probably due to the denser network of local meteorological stations and more detailed measurement methods.

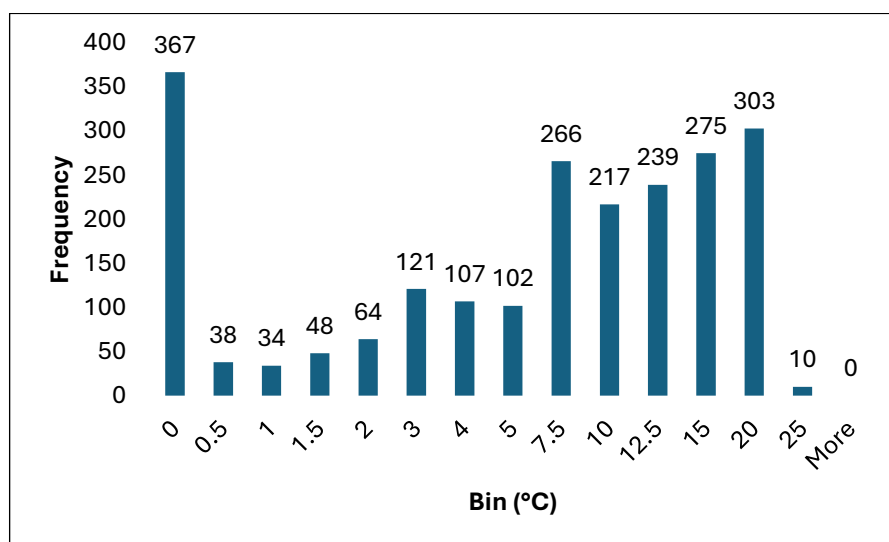


Figure 11 Princeton temperature data distribution 2005-2009

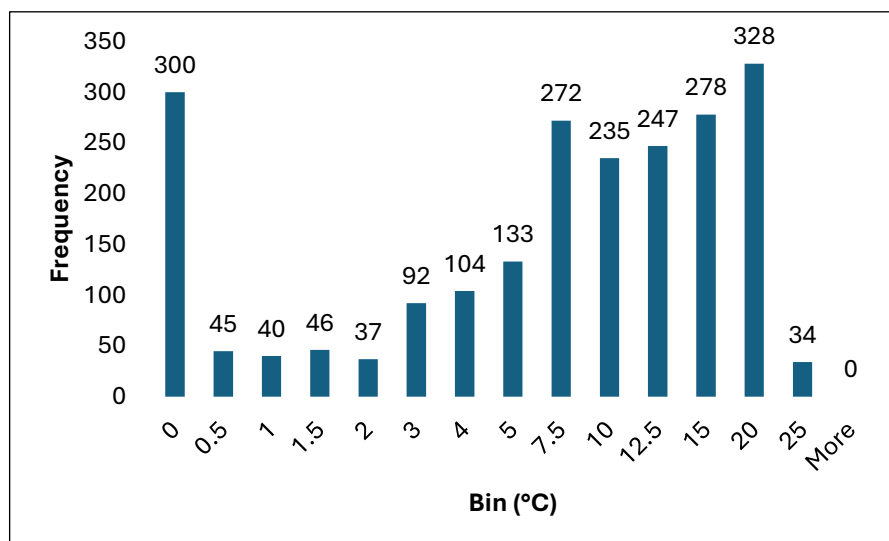


Figure 12 SeNorge temperature data distribution 2005-2009

The data were also compared on histograms (Figure 11-12), breaking the time into periods similar to the precipitation. On these graphs, it can be observed that the Princeton data show a more balanced distribution, with the SeNorge data showing a higher peak concentration. In the Princeton data series, extreme low and high values are less frequent, whereas SeNorge has more data in the higher temperature ranges. So, although the two datasets show similar patterns, the

Princeton data are less extreme. This may suggest that the Princeton model returns smoother values, while SeNorge tracks local variations better.

Table 5 10-year precipitation averages

Year	Average / SeNorge (°C)	Average / Princeton (°C)	Average / Dif- ference (°C)
1973-1983	6.751	6.073	0.677
1984-1993	7.082	6.409	0.672
1994-2003	7.541	6.729	0.812
2004-2014	7.801	7.227	0.574

Based on the 10-year average temperatures (Table 5), the Princeton data showed lower averages than the SeNorge database values in all years. This confirms the assumption that the Princeton model is susceptible to simulating climatic extremes, which can lead to inaccuracies in hydrological modelling at a local scale.

3.1.3. Snow depth comparison

Snow depth data were available from both the Princeton and SeNorge databases.

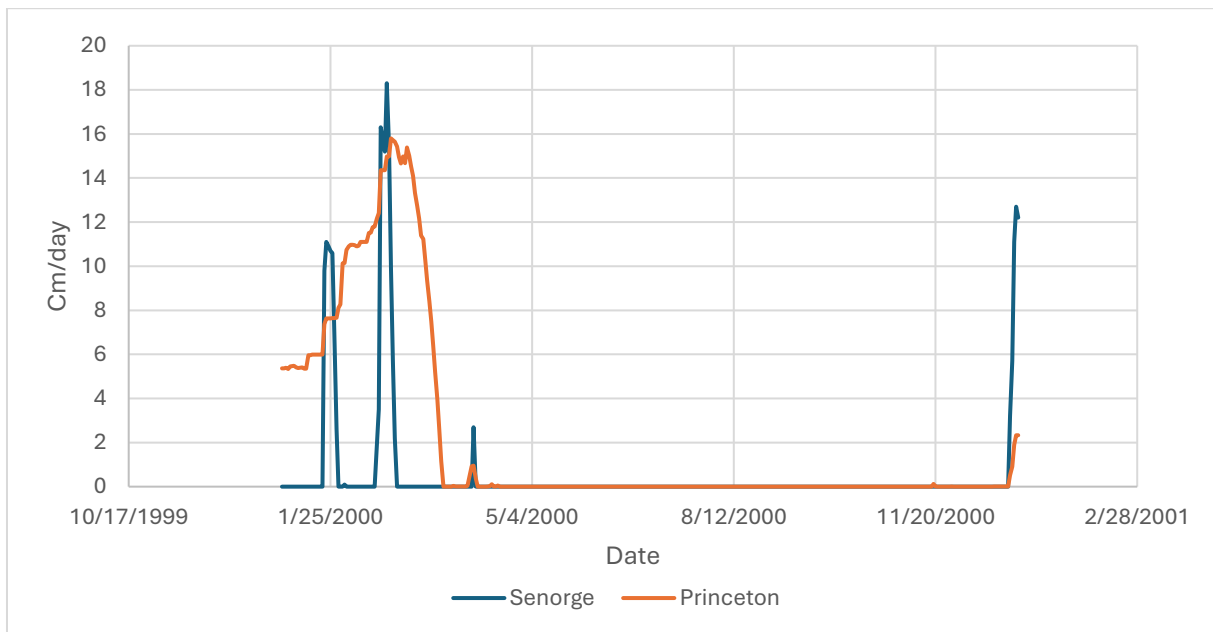


Figure 13 SeNorge and Princeton snow depth data in 2000

Figure 13 shows the daily average snow depth for a randomly selected year. It can be observed that the snow depth data from the SeNorge database decreased faster, meaning that snow melts faster than in the Princeton dataset. This suggests that the temperature model of the Princeton database may differ from the real local conditions.

3.1.4. Conclusion of the evaluation of climate data

Analysis of the data shows that the Princeton database does not provide sufficiently accurate climate data for Norway. Based on comparisons of precipitation and temperature, the SeNorge database was found to be more reliable as it better reflects local measurements.

Therefore, in the modelling analyses, I replaced the Princeton database precipitation and temperature data with the SeNorge data, as they provide higher accuracy and local relevance for the analysis of Norwegian rivers.

3.2. Calibration of the model for the whole period

3.2.1. Calibration of Kc

The first step of the calibration was to adjust the K_c (crop coefficient) parameter. Accurate determination of crop water use is key in hydrological modelling, especially when estimating evapotranspiration. The crop coefficient (K_c) is a dimensionless multiplier that indicates how much water a specific crop or vegetation type transpires relative to the reference evapotranspiration (ET₀) (Pereira, 1998). The actual evapotranspiration (ET_c) is calculated using equation 3:

$$ET_c = K_c \cdot ET_0 \quad (3)$$

The value of K_c changes according to the different growth stages of the vegetation, so the model needs to be capable of capturing these temporal variations. In the WEAP model, the purpose of calibrating K_c is to ensure that the simulated water use of vegetation reflects reality as closely as possible. This is important because it directly affects soil moisture, infiltration processes, surface runoff, and ultimately the accuracy of streamflow simulations (Figure 1). Well-chosen and properly calibrated K_c values are therefore essential for ensuring the reliability of the model's results.

The default value of K_c is 1, which means that the reference evapotranspiration is equal to the actual evapotranspiration. The calibration of the K_c value in this research was performed using the automatic calibration algorithm PEST (Parameter ESTimation) built into the WEAP model. PEST is an iterative optimization tool capable of fine-tuning the model input parameters in a way that minimizes the difference between simulated and measured data (Doherty, 2015). Figure 14 shows the layout of PEST in the WEAP software. Following this initial use, PEST was not applied further, as it mainly supported the visual understanding of parameter sensitivity, but did not define specific optimal parameter values.

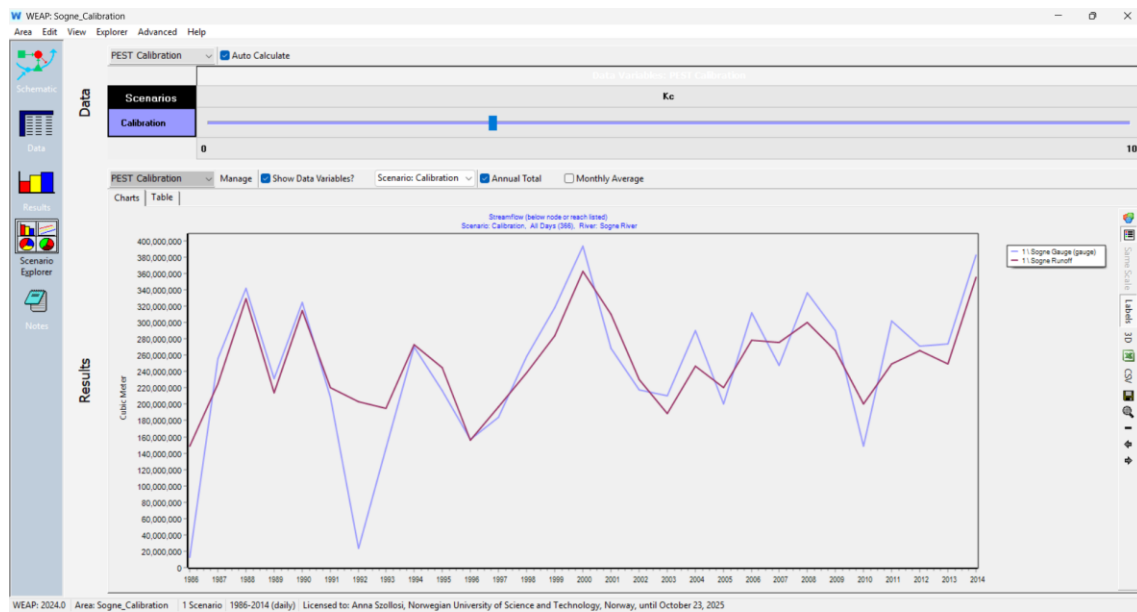


Figure 14 PEST program in the WEAP software

Table 6 Calibration of K_c for the whole period

K_c		
K_c (-)	PBIAS (%)	NSE (-)
2	-15.283	0.398
2,3	-10.535	0.396
2,4	-9.047	0.395
2,5	-7.604	0.393
2,6	-6.203	0.392
2,7	-4.844	0.390
2,8	-3.525	0.388
2,9	-2.243	0.386
3	-0.999	0.384
3.1	0.210	0.381
3.2	1.385	0.379
4	9.711	0.359

Table 6 presents the calibration results of the K_c value, based on the performance indicators PBIAS and NSE. According to the table, the PBIAS value is closest to zero at $K_c = 3.1$, while the NSE also remains within the range generally considered acceptable. The K_c value of 3.1 was kept constant throughout the calibration of all other parameters. This decision was made to maintain a consistent methodological approach across the entire calibration and evaluation process.

3.2.2. Calibration of other parameters

After calibrating the K_c value, further calibration calculations were deemed necessary based on the PBIAS and NSE indicators. In these cases, $K_c = 3.1$ was used, as it provided the most balanced results according to previous analyses. The calibration focused on the following hydrological parameters: Root Zone Conductivity (RZC), Soil Water Capacity (SWC), Deep Conductivity (DC), Deep Water Capacity (DWC), Runoff Resistance Factor (RRF), and Preferred Flow Direction (PFD). These parameters were selected because they have a direct impact on the soil moisture distribution, the dynamics of deep infiltration and the water balance processes simulated by the model. Proper calibration of these variables is essential for improving the model's overall performance.

Table 7 summarizes the default values of the parameters examined during the calibration process. As a starting point, the calibration typically considered half and double the default values, followed by a more detailed analysis using finer intervals. In each case, the most optimal value—based on performance indicators—was highlighted in green.

Table 7 Default value of parameters

Parameter	Default
Root Zone Conductivity (RZC)	20 mm/day
Soil Water Capacity (SWC)	1000 mm
Deep Conductivity (DC)	20 mm/day
Deep Water Capacity (DWC)	1000 mm
Runoff Resistance Factor (RRF)	2
Preferred Flow Direction (PFD)	0,15

Calibration of Root Zone Conductivity

The value of Root Zone Conductivity (RZC) refers to the conductivity of the root zone (top "bucket") under conditions of full saturation, i.e., when relative storage $Z_1 = 1.0$. At this saturation level, water flow divides in two directions: lateral movement within the upper soil layers (interflow), and downward movement toward the deeper soil layers. Different soil types result in varying RZC values. According to Table 8, the statistical performance indicators showed the most optimal results at $RZC = 15$ mm/day.

Table 8 Calibration of Root Zone Conductivity (RZC) for the whole period

RZC+Kc=3,1		
RZC (mm/day)	PBIAS (%)	NSE (-)
10	10.752	0.398
15	4.719	0.393
19	1.022	0.384
20 (default)	0.210	0.381
21	-0.565	0.379
25	-3.345	0.369
40	-10.791	0.336

Calibration of Soil Water Capacity

Soil Water Capacity (SWC) represents the water-holding capacity of the upper soil layer (top "bucket") and is expressed in millimeters (mm). The value of SWC typically ranges between 0 and 1000 mm. According to Table 9, the statistical indicators showed the most optimal performance at SWC = 500 mm.

Table 9 Calibration of Soil Water Capacity (SWC) for the whole period

SWC+Kc=3,1		
SWC (mm)	PBIAS (%)	NSE (-)
200	-10.144	0.555
500	-5.098	0.472
900	-0.739	0.396
950	-0.258	0.388
975	-0.022	0.385
980	0.024	0.384
1000 (default)	0.210	0.381
2000	6.922	0.299

Calibration of Deep Conductivity

Deep Conductivity (DC) represents the conductivity of the deeper soil layer (bottom "bucket") at a relative storage of $Z2 = 1.0$, meaning full saturation. This parameter controls the amount of baseflow generated in the model. In WEAP, the DC value is defined as a single, consistent parameter across the entire catchment area, regardless of soil classification or spatial heterogeneity. Increasing the DC value results in higher baseflow. It is important to note that if backflow occurs towards a groundwater node in the study area, the DC parameter is not taken into account in the model.

Table 10 presents the PBIAS and NSE statistical indicators for various DC values, with the fixed value $K_c = 3.1$. According to the analysis, a DC value of 1000 mm/day was found to be optimal and still hydrologically realistic. Higher values, such as DC = 5000 mm/day, do not reflect realistic soil physical properties; although they may lead to slight improvements in the NSE indicator, they do not yield a significant enhancement in overall model performance, as illustrated in Figure 15.

Table 10 Calibration of Deep Conductivity (DC) for the whole period

DC+Kc=3,1		
DC (mm/day)	PBIAS (%)	NSE (-)
10	0.268	0.386
20 (default)	0.210	0.381
40	0.189	0.380
100	0.178	0.386
200	0.179	0.394
500	0.195	0.407
1000	0.202	0.416
2000	0.206	0.425
5000	0.210	0.436

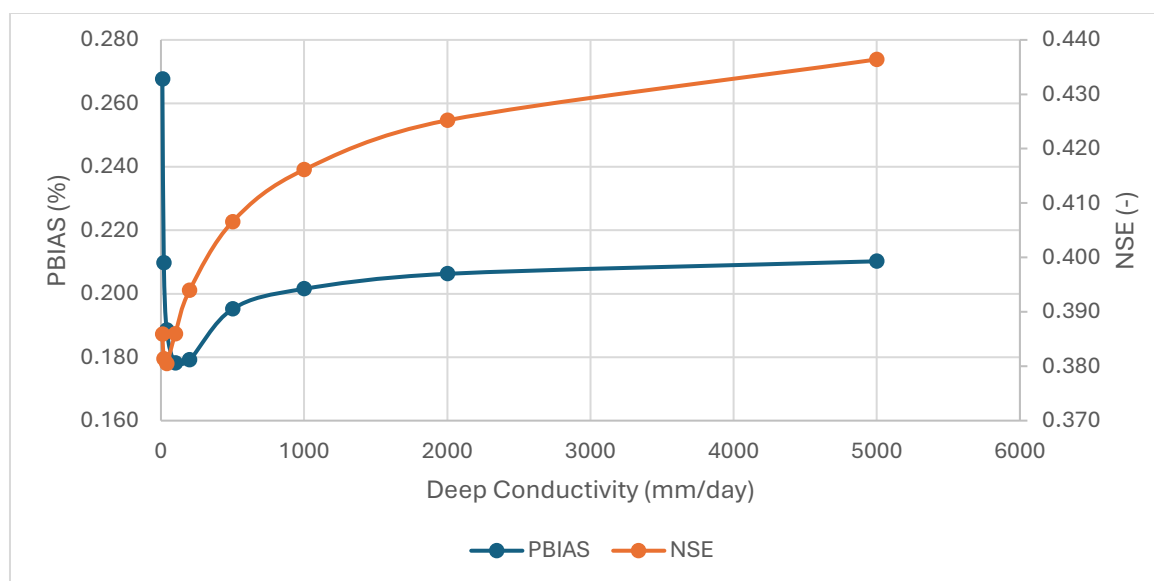


Figure 15 Variation of PBIAS and NSE according to DC

Calibration of Deep Water Capacity

Deep Water Capacity (DWC) represents the actual water holding capacity of the deeper soil layers (bottom “bucket”), expressed in millimeters (mm). This value is given as a single parameter for the catchment area and does not vary according to soil type. In cases where backflow occurs toward a groundwater node within the study area, the DWC parameter can be disregarded. Based on Table 11, $DWC = 10$ mm produced the most optimal results according to the statistical indicators.

Table 11 Calibration of Deep Water Capacity (DWC) for the whole period

DWC+Kc=3,1		
DWC (mm)	PBIAS (%)	NSE (-)
10	0.209	0.468
50	0.204	0.442
100	0.199	0.425
250	0.183	0.400
500	0.179	0.384
750	0.192	0.380
1000 (default)	0.210	0.381
2000	0.319	0.392

Calibration of the Runoff Resistance Factor

Runoff Resistance Factor (RRF) is defined as the control surface runoff response, influenced by various environmental factors such as catchment area index and land slope. The RRF typically ranges between 0 and 1000, with a default value of 2.

Since RRF can vary depending on land use categories, its proper calibration is essential for accurate simulation results. Sensitivity analyses were conducted during calibration to determine the statistical performance indicators corresponding to different RRF values, as summarized in Table 12. Although the value of $RRF = 1.5$ appeared theoretically optimal, the simulations indicated that $RRF = 2$ (the default value) yielded the most favorable results in terms of both PBIAS and NSE. It is also important to note that the model showed high sensitivity to changes in the RRF value.

Table 12 Calibration of Runoff Resistance Factor (RRF) for the whole period

RRF+Kc=3,1		
RRF (-)	PBIAS (%)	NSE (-)
1	-7.093	0.439
1,5	-2.432	0.433
2 (default)	0.210	0.381
4	4.040	0.200

Calibration of Preferred Flow Direction

The Preferred Flow Direction (PFD) parameter controls the direction of water movement in the root zone layer (top "bucket"). It can range from 0 to 1, where a value of 1.0 indicates that 100% of the water moves laterally (interflow), while a value of 0 implies that 100% of the water moves vertically downward into the lower soil layer (bottom "bucket") or directly into the groundwater. The PFD value may vary depending on the soil class. The default value in the model is 0.15.

The results of calibration simulations performed for different PFD values are summarized in Table 13. The performance indicators were most favorable at PFD = 0.9, suggesting that lateral water movement is dominant in the study area.

Table 13 Calibration of Preferred Flow Direction (PFD) for the whole period

PFD+Kc=3,1		
PFD (-)	PBIAS (%)	NSE (-)
0,1	0.208	0.373
0,15 (default)	0.210	0.381
0,3	0.215	0.405
0,4	0.218	0.420
0,5	0.219	0.433
0,7	0.214	0.456
0,9	0.168	0.470
1	-0.024	0.472

3.2.3. Optimal parameter values

In the previous chapters, a full model run was carried out using the best-performing parameters highlighted in green, in order to evaluate the success of the calibration and the overall performance of the model. The selected, calibrated parameters are summarized in Table 14.

The value of K_c differs from the previously indicated optimal choice highlighted in green: rather than the originally applied $K_c = 3.1$, a value of 2.7 was used in this final simulation. This adjustment was made because, when the model was run with the complete parameter set, $K_c = 2.7$ yielded the most favorable performance metrics. Compared to earlier simulations, this change was therefore intended to further improve the model's overall accuracy.

Table 14 Optimal values

Parameter	Calibrated value	Unit
K_c	2,7	-
Root Zone Conductivity (RZC)	15	mm/day
Soil Water Capacity (SWC)	500	mm
Deep Conductivity (DC)	1000	mm/day
Deep Water Capacity (DWC)	10	mm
Runoff Resistance Factor (RRF)	2	-
Preferred Flow Direction (PFD)	0,9	-

The simulations resulted in $PBIAS = -5.998\%$ and $NSE = 0.552$, which fall within the "satisfactory" model performance range according to the categorization scheme proposed by (Moriassi et al., 2015). Although the results obtained can be considered satisfactory, the question arises whether further fine-tuning could lead to additional improvements in the model's performance.

3.2.4. Calibration of melting and freezing point

Since the model was not originally developed for Scandinavian climatic conditions, a detailed examination of snow-related hydrological processes—particularly snow accumulation and snowmelt—was deemed necessary. To this end, some critical parameters related to climatic factors, such as melting point (MP) and freezing point (FP), were calibrated. The preliminary expectation before the calibration was that adjusting these factors would improve the model performance.

Melting Point calibration

In the model, the melting point (MP) defines the temperature threshold above which snowmelt begins. The default value is $+5\text{ }^{\circ}\text{C}$; however, the parameter can be adjusted within a range of $-50\text{ }^{\circ}\text{C}$ to $+50\text{ }^{\circ}\text{C}$. During the calibration process, $PBIAS$ and NSE performance indicators were

evaluated across a range of temperature values. Based on the results (see Table 15), a melting point of 4 °C yielded the best model fit, characterized by low PBIAS and high NSE values. Therefore, this value is considered optimal under the given climatic conditions.

Table 15 Calibration of Melting Point (MP) for the whole period

MP+Kc=3,1		
MP (°C)	PBIAS (%)	NSE (-)
10	1.944	0.288
5 (default)	0.210	0.381
4	-0.076	0.385
3	-0.289	0.381
2	-0.459	0.376
0	-0.659	0.364

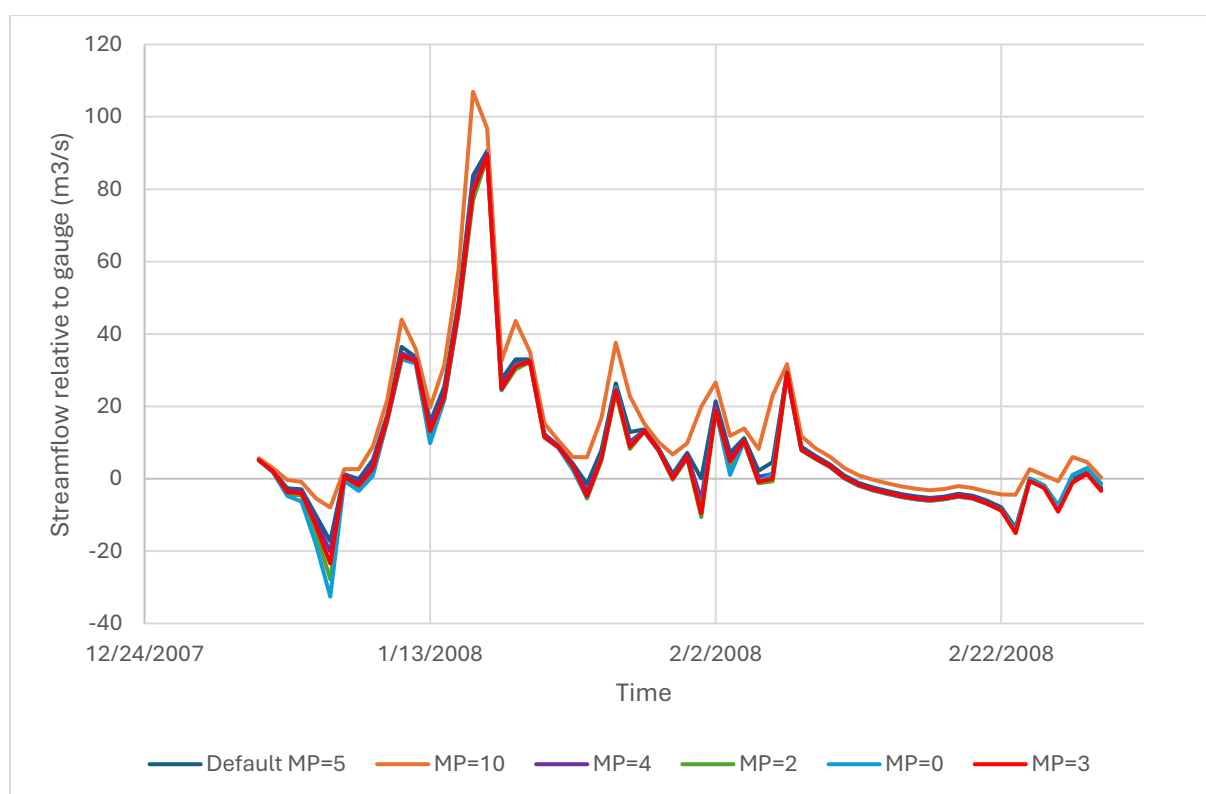


Figure 16 Relative streamflow as a function of Melting Point during the Jan–Feb 2008 Period

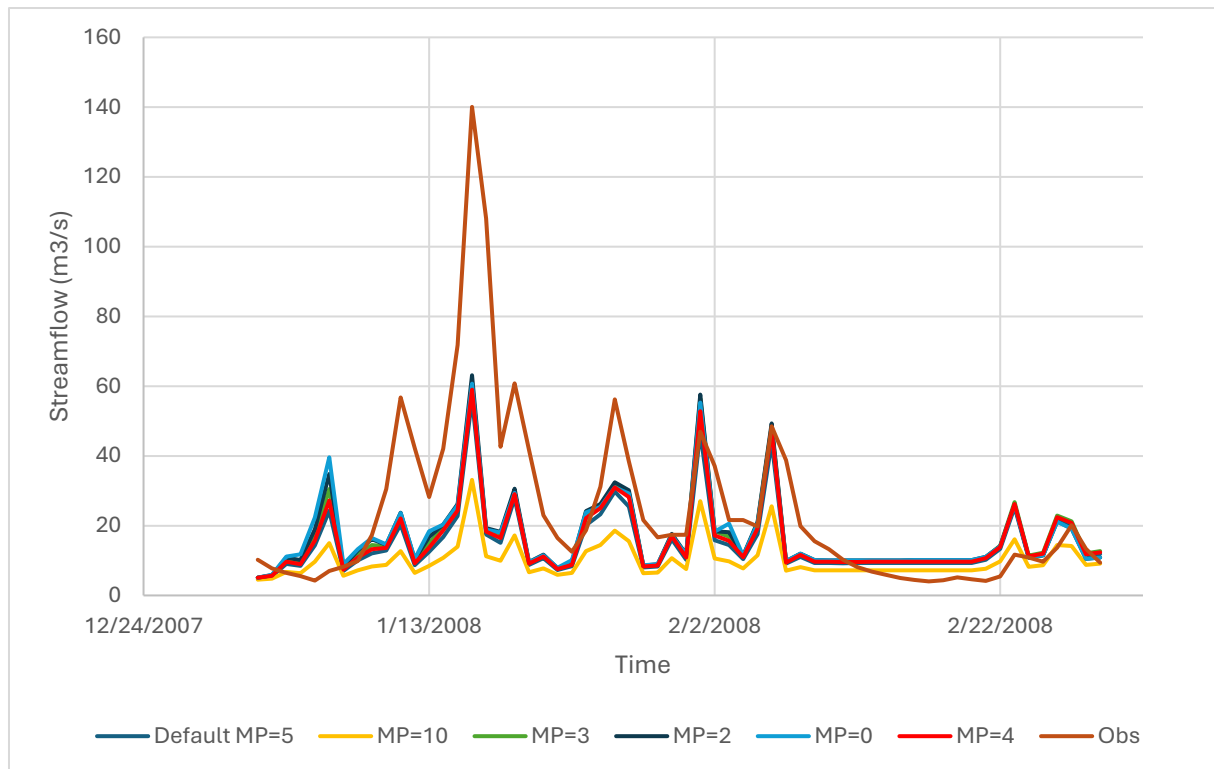


Figure 17 Streamflow as a function of Melting Point during the Jan–Feb 2008 Period

Figures 16 and 17 illustrate snowmelt events during a randomly selected year – specifically, the winter of 2008 – under various melting point (MP) settings. Figure 15 demonstrates the impact of different MP values on relative discharge, while Figure 16 compares the simulated discharge for each setting with the observed discharge (Obs) data.

According to the figures, variations in MP have only a moderate influence on model output. The curves generated under different settings follow very similar trends, with only minor differences observed in the magnitude and timing of peak flows. This suggests that calibrating the melting point does not lead to drastic changes in runoff dynamics; however, based on the NSE and PBIAS indicators, it can still yield modest improvements in model fit.

Freezing Point calibration

In the model, the freezing point (FP) represents the temperature threshold below which precipitation accumulates in solid form (as snow). The default value is -5°C , and the parameter can be adjusted within the same range as the MP parameter. According to the calibration results (see Table 16), a value of -5°C provided the best model performance, indicating that no adjustment of the default setting is necessary.

Table 16 Calibration of Freezing Point (FP) for the whole period

FP+Kc=3,1		
FP (°C)	PBIAS (%)	NSE (-)
-10	-0.306	0.376
-7	-0.058	0.379
-6	0.063	0.380
-5 (default)	0.210	0.381
-4	0.390	0.381
-3	0.608	0.379
0	1.688	0.367

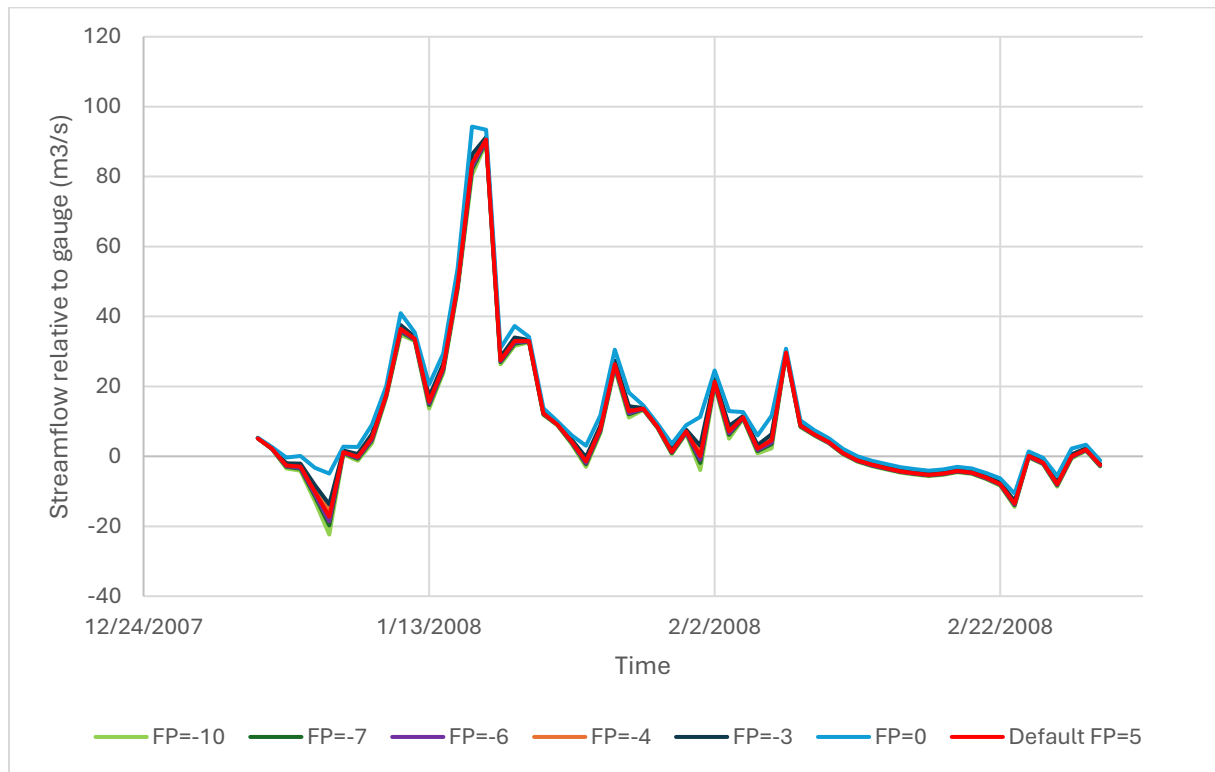


Figure 18 Relative streamflow as a function of Freezing Point during the Jan–Feb 2008 Period

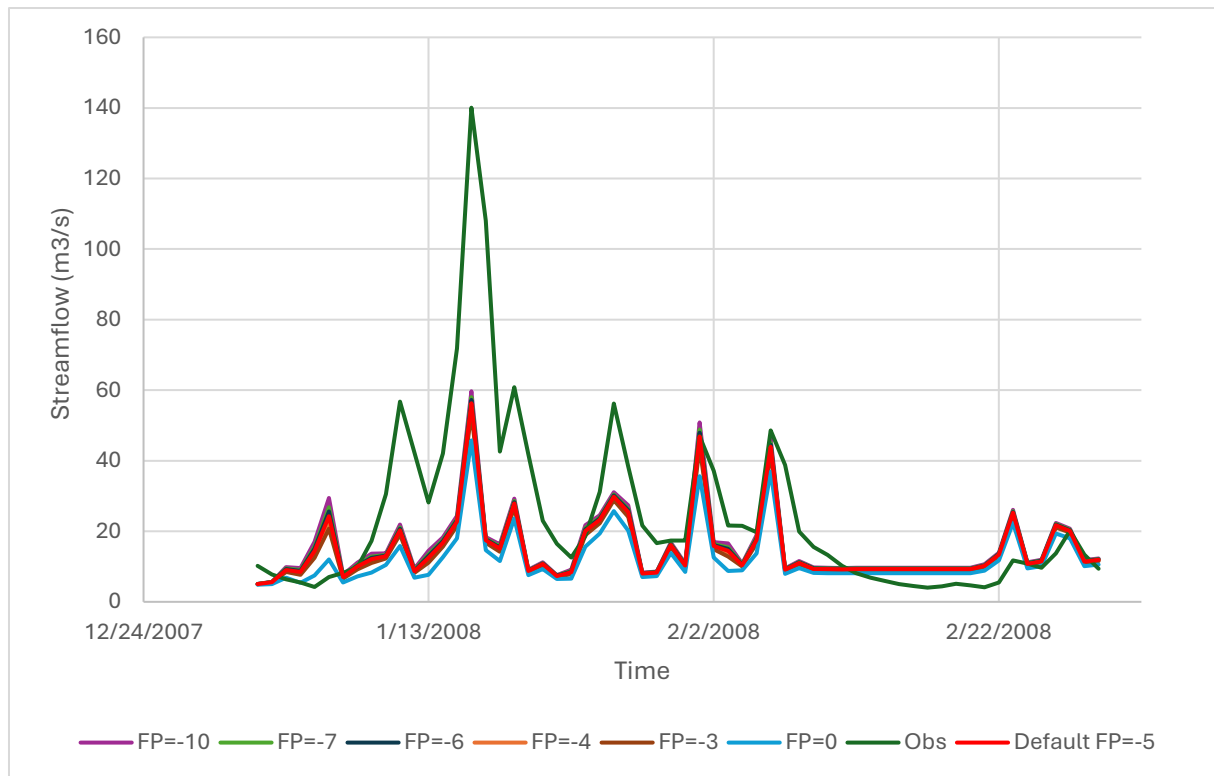


Figure 19 Streamflow as a function of Freezing Point during the Jan–Feb 2008 Period

Figures 18 and 19 present a snowmelt event during a randomly selected year – specifically, the winter of 2008 – demonstrating the effects of varying the freezing point (FP). The figures indicate that changes in the freezing point have a negligible impact on model performance.

Optimal parameter values

To assess overall model performance, the calibrated melting and freezing points were incorporated into the simulation alongside the previously identified best-performing hydrological parameters. The applied parameter values are summarized in Table 17.

Table 17 Final optimal values for the whole period

Parameter	Calibrated values	Unit
Kc	2,7	-
Root Zone Conductivity (RZC)	15	mm/day
Soil Water Capacity (SWC)	500	mm
Deep Conductivity (DC)	1000	mm/day
Deep Water Capacity (DWC)	10	mm
Runoff Resistance Factor (RRF)	2	-
Preferred Flow Direction (PFD)	0,9	-
Melting Point (MP)	4	°C
Freezing Point (FP)	-5	°C

The results of the updated simulation are PBIAS = -6.209%, NSE = 0.553. These results represent a slight improvement over the previous simulation, particularly in terms of the NSE indicator, which reflects an enhanced predictive capacity of the model. Thus, the calibrated melting and freezing point therefore contribute slightly to fine-tuning the model, especially in cold, snow-covered regions.

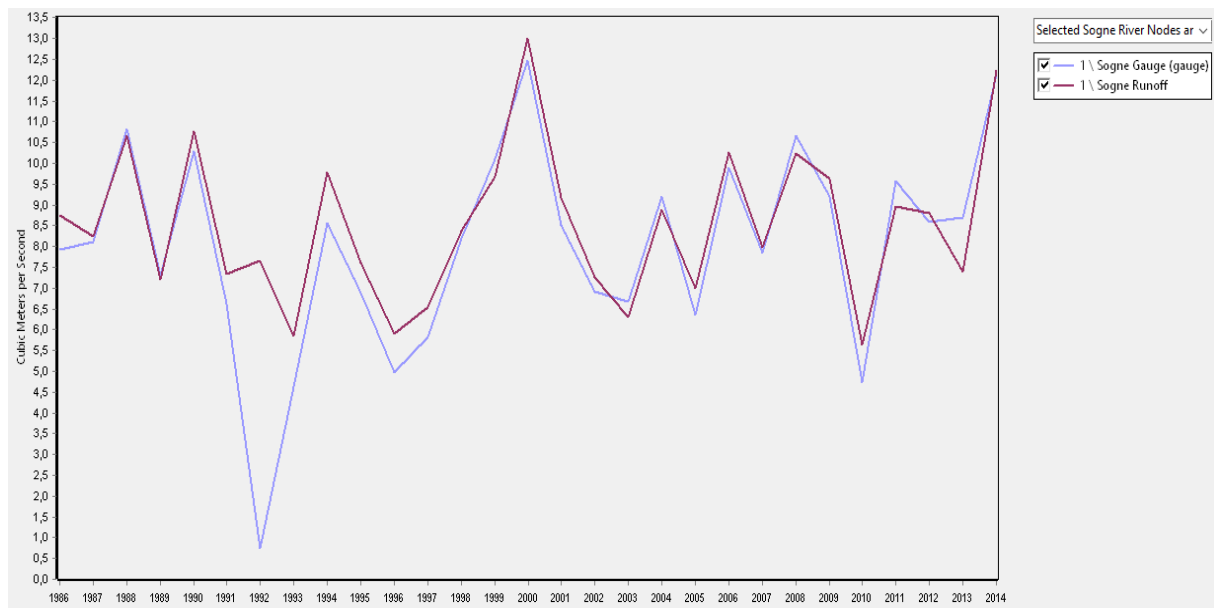


Figure 20 Annual average streamflow using parameters calibrated for the whole period

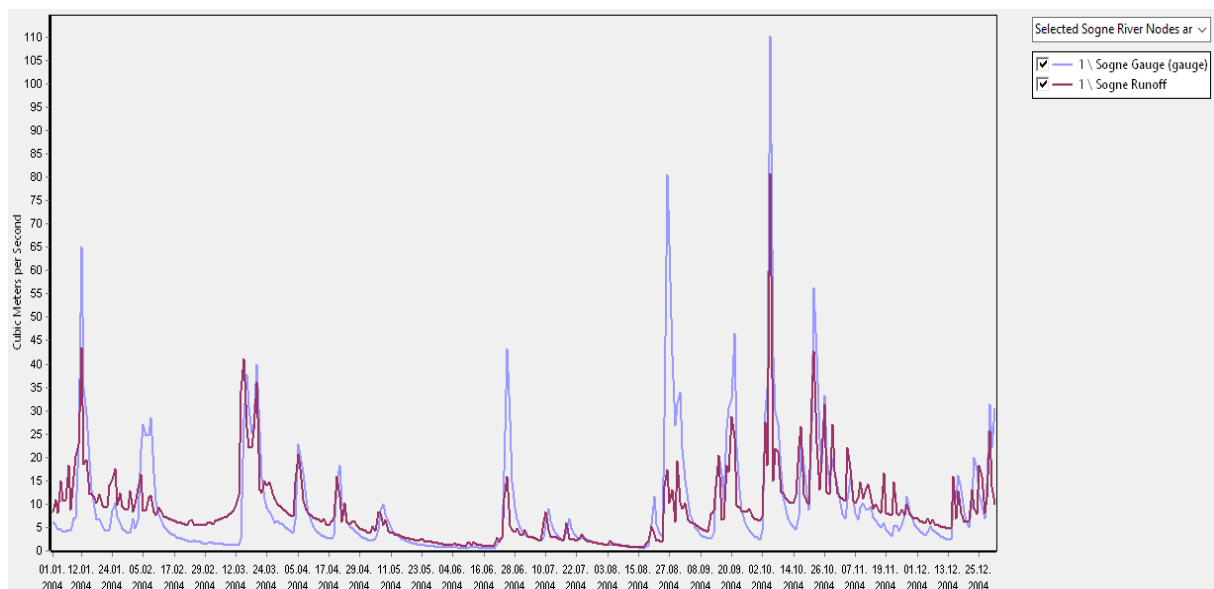


Figure 21 Changes in measured and simulated streamflow in 2004 using parameters calibrated for the whole period

Figures 20 and 21 compare the simulated ("Runoff") and observed ("Gauge") streamflow. The graphs are taken directly from the WEAP software. Figure 20 illustrates the annual average streamflow trends over the period 1986–2014, while Figure 21 presents the daily streamflow time series for a randomly selected year (2004).

Based on Figure 20, it can be concluded that there is generally good agreement between the simulated and observed streamflow. The figure also shows that no observed data was available for the year 1992. Overall, the model is capable of reproducing the interannual variability of

streamflow, effectively capturing fluctuations driven by precipitation and other hydrological factors.

Figure 21 offers a more detailed view of model performance at a daily resolution. For the year 2004, the simulated and observed data show a strong correlation, especially during periods of low and moderate flow. The model adequately represents key features of runoff dynamics, including the system's response to precipitation events and subsequent recession phases. The largest deviations occur during peak flow events. In these cases, observed values are generally higher than simulated ones, suggesting that the model may underestimate runoff associated with extreme precipitation events.

3.3. Seasonal calibration of the model for summer and winter periods

Although the calibration of the model for the whole period has already shown satisfactory results, the question arises whether the performance of the model can be further improved by calibrating the model separately for summer and winter, taking seasonal variability into account. In this approach, snow-related processes such as snow accumulation and snowmelt could be isolated or excluded from individual simulation periods, thereby potentially improving overall model performance. The following section explores this possibility.

For the study, the one-year period was divided into two segments:

- Winter period: November 1 – April 30
- Summer period: May 1 – October 31

Characteristics of the winter period

During winter, temperatures drop significantly, promoting snow accumulation and surface ice formation. Most precipitation falls in solid form (as snow), which is temporarily stored in the catchment area and does not immediately appear in the streamflow. Consequently, discharge levels are generally low, while the spring snowmelt can cause sudden and significant runoff events, potentially resulting in flood peaks.

Characteristics of the summer period

The summer period is characterized by rising temperatures and variable precipitation patterns. In May and June, high discharge may result from the significant water excess generated by snowmelt. In mid-summer, decreasing precipitation and increased evaporation often lead to

lower water levels. In September and October, frequent rainfall typically causes streamflow to rise again.

Given the contrasting hydrological characteristics of these two periods, separate calibration is justified. This approach enables a more accurate representation of seasonal model sensitivity and offers the potential to further improve simulation performance.

Calibration Method

In this case, calibration was performed by applying a consistent set of parameter values across the entire simulation period (a distinct parameter set was used for the summer and winter calibrations, respectively). Using the selected parameter set, I ran the simulation over the full period. From the resulting streamflow time series, only the months corresponding to the target season (i.e., either summer or winter) were extracted. The NSE and PBIAS performance indicators were then calculated based solely on this season-specific subset of the data.

3.3.1. Calibration of Kc

To investigate the seasonal sensitivity of the Kc (crop coefficient) parameter, calibration was performed separately for the summer and winter periods. Model fit was evaluated using the statistical indicators PBIAS and NSE. The results are summarized in Table 18.

Table 18 Calibration of Kc for summer and winter periods

	Kc			
Kc (-)	PBIAS - Nyár (%)	NSE - Nyár (-)	PBIAS - Tél (%)	NSE - Tél (-)
1	-58.264	0.331	-22.259	0.356
2	-26.626	0.364	-9.280	0.368
2,3	-19.317	0.359	-5.886	0.367
2,4	-17.064	0.356	-4.803	0.366
2,5	-14.896	0.353	-3.743	0.365
2,6	-12.809	0.350	-2.706	0.365
2,7	-10.800	0.346	-1.691	0.364
2,8	-8.864	0.343	-0.698	0.363
2,9	-7.000	0.339	0.274	0.361
3	-5.203	0.335	1.226	0.360
3,1	-3.471	0.331	2.158	0.359
3,2	-1.800	0.327	3.070	0.357
4	9.655	0.292	9.740	0.343
5	-	-	1.093	0.491
6	-	-	-5.145	0.489

The values in green represent the best compromise for the model's performance over the period. However, it can be concluded that seasonal calibration did not result in a significant improvement for the summer period; in fact, a decline in the NSE value was observed compared to the calibration performed over the full year (see Table 6 above).

This suggests that seasonal adjustment of the K_c parameter alone can be insufficient to improve model performance, and that further investigation of parameter combinations or seasonal calibration of additional variables may be necessary.

3.3.2. Calibration of other parameters

To further improve model performance, a seasonal calibration of key hydrological parameters was conducted – namely, Root Zone Conductivity (RZC), Soil Water Capacity (SWC), Deep Conductivity (DC), Deep Water Capacity (DWC), Runoff Resistance Factor (RRF), and Preferred Flow Direction (PFD) – separately for the summer period (May 1 – October 31) and the winter period (November 1 – April 30). For the K_c parameter, the original value for the full simulation period ($K_c = 3.1$) was retained for both seasons. Although separately determined K_c values (e.g., 3.2 for summer and 5 for winter) could theoretically provide more seasonally accurate evapotranspiration estimation, the globally applied value was deemed a suitable compromise for the study's objectives and to maintain modeling consistency. As such, the effects of seasonal differences were assessed by adjusting the remaining parameters while keeping K_c constant.

Calibration of Root Zone Conductivity

The RZC parameter was modified in several steps, and model fit for each value was evaluated using the PBIAS and NSE indicators, separately for the summer and winter periods. The results are summarized in Table 19.

Table 19 Calibration of Root Zone Conductivity (RZC) for summer and winter periods

	RZC+K_c=3,1			
RZC (mm/day)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
10	10.561	0.369	10.854	0.364
15	2.294	0.348	6.003	0.367
19	-2.452	0.334	2.861	0.361
20 (default)	-3.471	0.331	2.158	0.359
21	-4.435	0.328	1.483	0.356
25	-7.849	0.316	-0.961	0.347
40	-16.819	0.285	-7.600	0.310

Based on the results, the optimum value of Root Zone Conductivity is 15 mm/day in both periods. This value resulted in the lowest bias (PBIAS) and provided relatively balanced model performance (NSE).

Calibration of Soil Water Capacity

The Soil Water Capacity (SWC) parameter was calibrated to fine tune the hydrological performance of the model, separately for the summer and winter periods. The SWC value was varied over a wide range (200-2000 mm) and the model performance for each value was evaluated using the PBIAS and NSE indicators. The results are shown in Table 20.

Table 20 Calibration of Soil Water Capacity (SWC) for summer and winter periods

	SWC+Kc=3,1			
SWC (mm)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
200	-9.602	0.652	-10.430	0.464
500	-2.955	0.476	-6.233	0.426
900	-2.979	0.351	0.447	0.371
950	-3.212	0.340	1.306	0.365
975	-3.339	0.335	1.733	0.362
980	-3.365	0.334	1.818	0.361
1000 (default)	-3.471	0.331	2.158	0.359
2000	-9.405	0.244	15.565	0.272

Based on the calibration results, SWC = 200 mm for the summer period and SWC = 500 mm for the winter period showed the best balance between PBIAS and NSE.

Calibration of Deep Conductivity

For the Deep Conductivity (DC) parameter, the calibration covered a broad range from 10 to 5000 mm/day. Model performance across this range was evaluated separately for each season. The corresponding PBIAS and NSE values are presented in Table 21.

Table 21 Calibration of Deep Conductivity (DC) for summer and winter periods

	DC+KC=3,1			
DC (mm/day)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
10	-8.979	0.332	5.162	0.365
20 (default)	-3.471	0.331	2.158	0.359
30	0.281	0.332	0.150	0.356
40	3.021	0.334	-1.311	0.355
100	11.081	0.346	-5.593	0.357
200	15.667	0.359	-8.019	0.363
400	18.755	0.373	-9.645	0.371
500	19.467	0.377	-10.018	0.374
1000	20.992	0.390	-10.816	0.382
2000	21.781	0.402	-11.225	0.391
5000	22.217	0.416	-11.450	0.401

The results show that DC = 100 mm/day for the summer period and DC = 200 mm/day for the winter period give the best performance.

Calibration of Deep Water Capacity

Deep Water Capacity values were tested in the range 10-2000 mm. The model fit was evaluated according to Table 22.

Table 22 Calibration of Deep Water Capacity (DWC) for summer and winter periods

	DWC+Kc=3,1			
DWC (mm)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
10	22.322	0.461	-11.496	0.428
50	22.306	0.423	-11.495	0.406
100	21.781	0.402	-11.225	0.391
250	17.913	0.368	-9.202	0.369
500	9.289	0.343	-4.643	0.356
750	1.899	0.333	-0.712	0.355
1000 (default)	-3.471	0.331	2.158	0.359
2000	-12.893	0.337	7.313	0.372

Based on the calibration, $DWC = 500$ mm for the summer period and $DWC = 10$ mm for the winter period give the best performance.

Calibration of Runoff Resistance Factor

The Runoff Resistance Factor (RRF) was varied within the range of 1 to 4. Based on PBIAS and NSE indicators, an RRF value of 1.5 the summer, and RRF value of 2 for the winter provided the most balanced model performance. Table 23 presents the statistical performance metrics associated with different RRF values, using a fixed $K_c = 3.1$. The results indicate that the model is highly sensitive to changes in this parameter.

Table 23 Calibration of Runoff Resistance Factor (RRF) for summer and winter periods

	RRF+K_c=3,1			
RRF (-)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
1	-14.897	0.482	-2.961	0.368
1,5	-7.038	0.430	0.006	0.388
2 (default)	-3.471	0.331	2.158	0.359
4	-1.869	0.088	7.168	0.196

Calibration of Preferred Flow Direction

The Preferred Flow Direction (PFD) parameter was tested over the interval from 0.1 to 1. According to the calibration results shown in Table 24, for the summer period $PFD = 0.5$, for the winter period $PFD = 0,7$ yielded the most favorable outcomes.

Table 24 Calibration of Preferred Flow Direction (PFD) for summer and winter periods

	PFD+K_c=3,1			
PFD (-)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
0,1	-4.404	0.319	2.649	0.351
0,15 (default)	-3.471	0.331	2.158	0.359
0,3	-0.444	0.364	0.564	0.379
0,4	1.792	0.385	-0.615	0.391
0,5	4.232	0.405	-1.905	0.403
0,7	9.867	0.439	-4.896	0.421
0,9	16.828	0.463	-8.651	0.430
1	20.794	0.470	-11.043	0.430

3.3.3. Calibration of Melting and Freezing Points

Snow-related processes, such as snow accumulation and melting, play a critical role in modelling water balance, especially for seasonal distribution. To enhance model performance, the Freezing Point (FP) and Melting Point (MP) parameters were calibrated separately for the summer and winter periods.

Calibration of Freezing Point

The freezing point was examined in the range of -10 °C to 0 °C. Table 25 presents the model performance metrics for different FP values, with the most optimal values highlighted in green.

Table 25 Calibration of Freezing Point (FP) for summer and winter periods

	FP+Kc=3,1			
FP (°C)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
-10	-1.872	0.332	0.524	0.349
-7	-2.659	0.332	1.319	0.354
-6	-3.027	0.331	1.699	0.357
-5 (default)	-3.471	0.331	2.158	0.359
-4	-4.000	0.330	2.714	0.359
-3	-4.641	0.329	3.386	0.356
0	-7.652	0.325	6.632	0.339

Calibration of Melting Point

The melting point was tested over a range of 0 °C to 10 °C. According to Table 26, the value of 0 °C was optimal for the summer period, and 4 °C for the winter period.

Table 26 Calibration of Melting Point (MP) for summer and winter periods

	MP+Kc=3,1			
MP (°C)	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
10	-9.450	0.313	7.974	0.215
5 (default)	-3.471	0.331	2.158	0.359
4	-2.546	0.332	1.232	0.364
3	-1.792	0.333	0.507	0.357
2	-1.151	0.334	-0.092	0.348
0	-0.367	0.334	-0.813	0.329

3.3.4. Optimal parameter values

The previously described seasonal calibrations resulted in individually optimized parameter values for each season. Simulations were conducted using the green-highlighted parameter combinations to assess hydrological model performance based on PBIAS and NSE indicators. Two simulation configurations were applied: one that included the melting point (MP) and freezing point (FP) parameters, and one that excluded them. The final, best-performing parameter combinations are presented in Table 27 and the final simulation results are shown in Table 28.

Table 27 Final best ideal values for summer and winter periods

Parameter	Calibrated value – Summer	Calibrated value – Winter	Unit
Kc	3,2	5	-
Root Zone Conductivity (RZC)	15	15	mm/day
Soil Water Capacity (SWC)	200	500	mm
Deep Conductivity (DC)	100	200	mm/day
Deep Water Capacity (DWC)	500	10	mm
Runoff Resistance Factor (RRF)	1,5	2	-
Preferred Flow Direction (PFD)	0,5	0,7	-
Melting Point (MP)	0	4	°C
Freezing Point (FP)	-10	-5	°C

Table 28 Final simulation results

	PBIAS – Summer (%)	NSE – Summer (-)	PBIAS – Winter (%)	NSE – Winter (-)
Good	-6.664	0.719	1.093	0.491
Good+MP,FP	-7.621	0.714	1.684	0.496

Based on the results, it can be concluded that applying the calibrated values for freezing and melting points decreased model performance during the summer period, while improving it during the winter. This suggests that temperature-driven snow processes play a more significant role in winter, whereas in summer, these parameters may have a negligible or even disruptive effect on the model.

Although the statistical indicators for the winter period worsened compared to the values obtained for the full period, they still fall within the “satisfactory” model performance category defined by (Moriassi et al., 2015). In contrast, the indicators for the summer period show significant improvement.

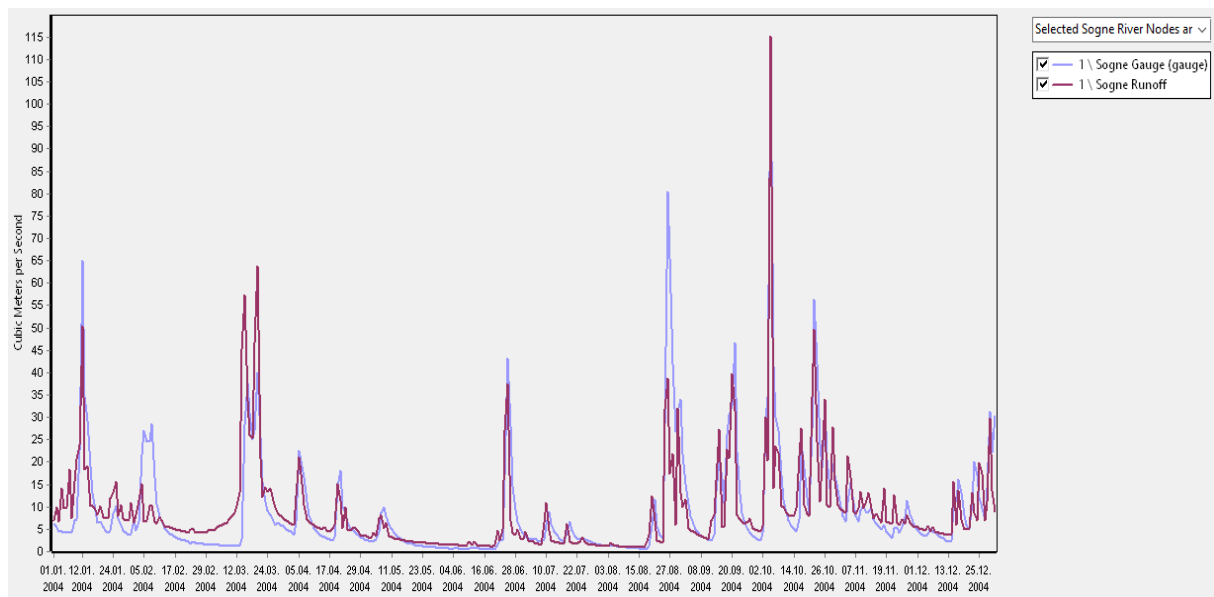


Figure 22 Changes in measured and simulated streamflow in 2004 using parameters calibrated for the summer period

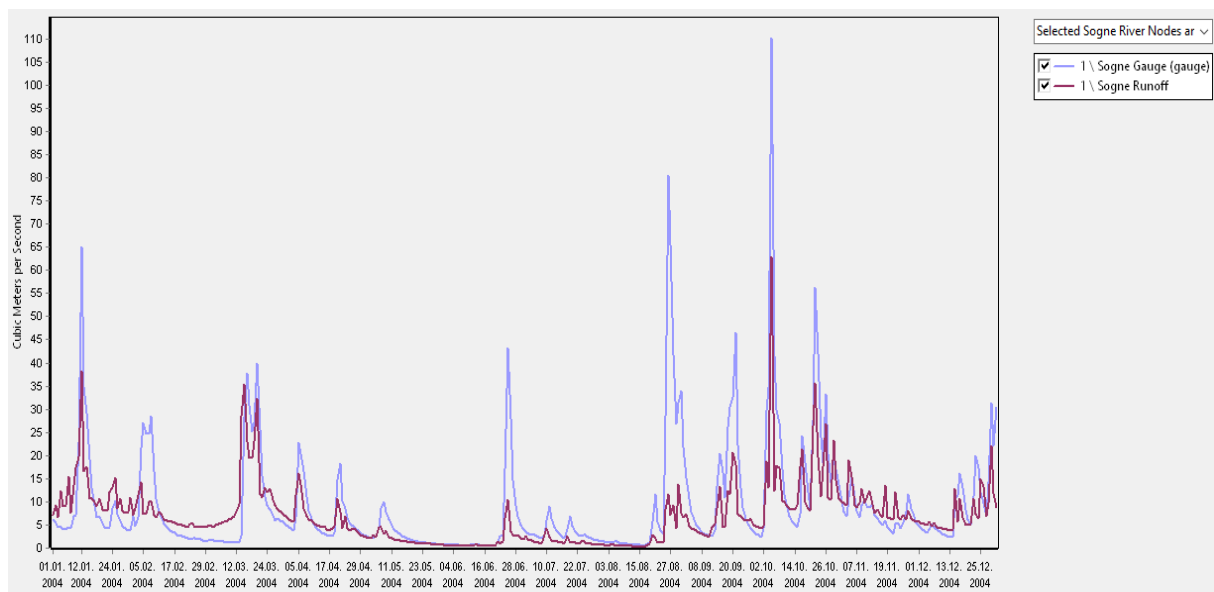


Figure 23 Changes in measured and simulated streamflow in 2004 using parameters calibrated for the winter periods

Figures 22 and 23 present the simulated and observed streamflow for a randomly selected year. The graphs are taken directly from the WEAP software. Figure 22 shows the results of the model calibrated for the summer period, while Figure 23 displays the results from the winter-calibrated model.

In Figure 22, it is evident that the model captures the smaller and medium-scale flow fluctuations relatively well; however, it occasionally underestimates the larger flood peaks. In

some cases, the model's response time is faster than observed, meaning that the simulation begins rising earlier than the actual event. Overall, the model provides reliable results, and the summer calibration generally shows good agreement.

Figure 23 demonstrates that the model is unable to accurately reproduce the larger flood peaks. The simulated values consistently underestimate the observed peak discharges and exhibit a smoother hydrograph shape. While the model still performs well during low-flow periods, it responds less sensitively to rapid changes caused by precipitation events. It can be concluded that the model calibrated for winter conditions fail to reliably simulate the larger flood events, suggesting that winter precipitation (rain, snow, and meltwater) is still not adequately represented in the model's parameterization.

3.3.5. Equifinality – GLUE

Throughout the calibration process, it is essential to consider the principle of equifinality, which represents a major source of uncertainty in hydrological modeling (K. Beven, 2006). Equifinality describes the phenomenon where a given model can produce similarly good fits to observations with different parameter sets – implying that there is not necessarily a single "best" set of parameters.

A widely used method for quantifying model uncertainty and addressing equifinality is the Generalised Likelihood Uncertainty Estimation (GLUE) approach, introduced by K. J. Beven & Binley (1992). Instead of searching for a single optimal parameter set, GLUE accepts multiple parameter combinations that sufficiently match the observations. It assigns likelihood values to these combinations based on model performance allowing statistical interpretation of model output uncertainties and the determination of predictive probability ranges.

The GLUE method thus also helps to reveal the structural uncertainty of the model, as not only the parameters but also different elements of the model setup (e.g. representation of hydrological processes) can affect the equivalence of the results. In this context, it is important to acknowledge that a certain level of uncertainty is inevitable in modeling and should be explicitly addressed as part of the decision-support process.

This emphasizes that, although a reasonably well-fitting parameter combination was identified during the model calibration, it should not necessarily be considered the optimal solution. In line with the principle of equifinality, it is possible that other parameter sets could provide similarly good or even better fits. Therefore, the results of the calibration should be interpreted

not as a definitive solution, but as one of several plausible outcomes, especially when the model is intended to support decision-making processes.

3.4. Model validation

Model validation is an essential step in the model development process. In this study, validation was carried out using the remaining one-third of the available data, covering the period from 1974 to 1986. During the validation, the same calibrated parameter set – originally optimized for the full calibration period (1987–2014) – was applied without any further modification.

No validation was performed using the parameters calibrated separately for summer and winter periods, as the model structure did not support the simultaneous application of season-specific parameters across the entire time series. Since these parameter sets were tailored specifically for summer or winter months, applying them to the full year (and thus for validation purposes) would not have yielded realistic or representative results.

The validation produced a PBIAS of -9.84% and a Nash-Sutcliffe Efficiency (NSE) of 0.483. Although the NSE value falls below the commonly accepted threshold for satisfactory model performance suggested by (Moriassi et al., 2015), the results can still be considered acceptable, particularly given that the validation was conducted for a fully independent, earlier time period. Furthermore, the outcomes align with the general expectation that model performance tends to be lower during validation than during calibration, since the parameter set is optimized specifically for the calibration period and may not perform equally well in a different period with different hydrological characteristics.

3.5. Assessment of future climate scenarios

The three climate models (NorESM2-LM, MPI-ESM1-2-HR, and HadGEM3-GC31-MM) were run for a 30-year future period between 2060 and 2090, and their outputs were compared to simulated historical data for the 1980–2010 reference period.

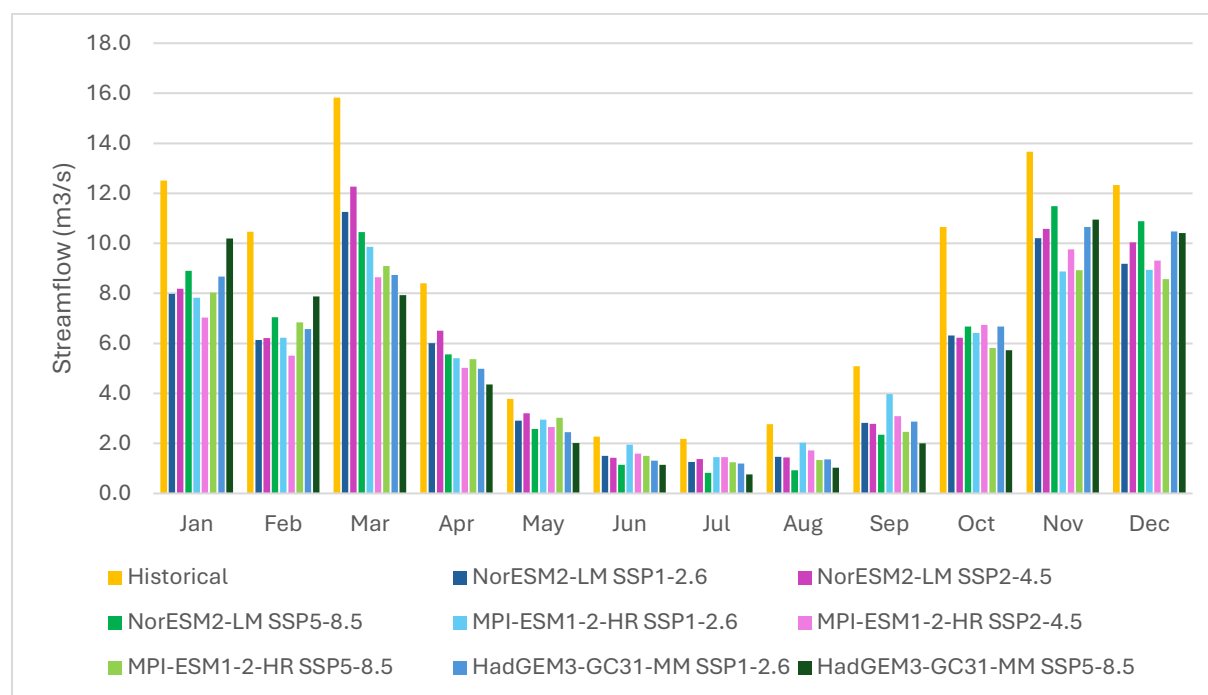


Figure 24 Monthly average streamflow under different scenarios

Table 29 Future streamflow compared to the reference period

		Average values of the 30 years (m³/s)	Change compared to historical where historical data is considered as 100% (%)
	Historical simulation	8.338	100
NorESM2-LM	SSP1-2.6	5.631	67.541
	SSP2-4.5	5.905	70.826
	SSP5-8.5	5.728	68.697
MPI-ESM1-2-HR	SSP1-2.6	5.485	65.789
	SSP2-4.5	5.114	61.335
	SSP5-8.5	5.271	63.215
HadGEM3-GC31-MM	SSP1-2.6	5.503	65.994
	SSP5-8.5	5.408	64.862

Figure 24 illustrates the monthly average streamflow values projected for the future period (2060–2090), based on different SSP scenarios of three global climate models (NorESM2-LM,

MPI-ESM1-2-HR, and HadGEM3-GC31-MM), compared to the simulated values of the reference period (1980–2010).

Compared to the simulated historical data, most future scenarios differ in nearly all months. The most striking changes occur during the winter and early spring months (January–April), where streamflow values are generally lower than in the historical reference. This suggests that snowmelt may occur earlier or that snow accumulation may decrease due to warmer temperatures. Such changes are particularly relevant for spring flooding risks and reservoir management.

During the summer months (June–August), all three models and scenarios project low streamflow values. This supports the widely expected climate trend of more frequent summer droughts, which could have adverse impacts on ecosystems, water use, and agriculture.

In the autumn and early winter months (October–December), streamflow values mostly remain lower or similar to those of the reference period. This may indicate that precipitation in the future will continue to fall mostly in liquid form during these months, yet immediate runoff will not significantly exceed past levels.

Table 29 shows the ratio of future streamflow values compared to the simulated historical reference, which is considered as 100%. According to the future climate scenarios, the calculated streamflow is significantly lower than those of the 1980–2010 reference period. Based on the different emission scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5) applied to the three examined climate models (NorESM2-LM, MPI-ESM1-2-HR, and HadGEM3-GC31-MM), average future streamflow values only reach 61–71% of the reference levels.

This implies that, as a result of climate change, streamflow is projected to decrease by approximately 29–39% during the examined future period (2060–2090), depending on the specific model–scenario combination. The greatest decrease is indicated by the MPI-ESM1-2-HR model under the SSP2-4.5 scenario (only 61.3%), while the smallest reduction is found in the NorESM2-LM model under the same scenario (70.8%).

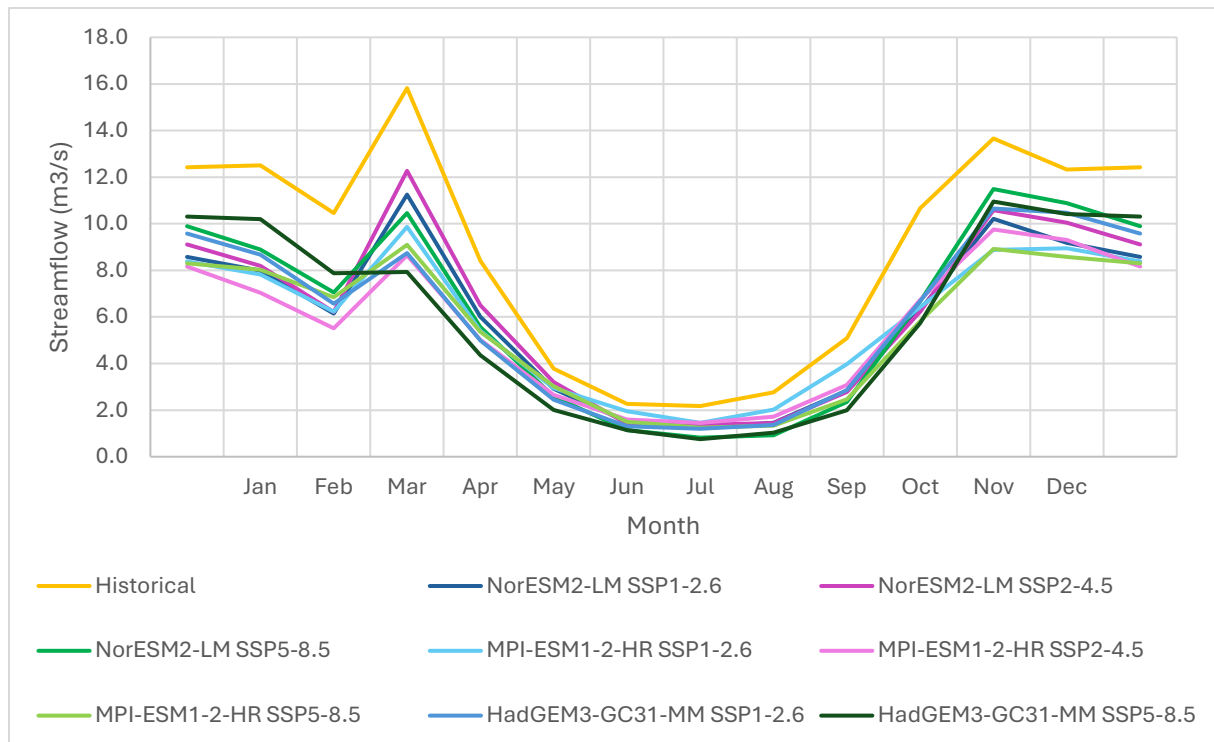


Figure 25 Average monthly streamflow values under different scenarios

Figure 25 clearly illustrates how the annual hydrograph evolves under various climate model–scenario combinations in the future, compared to the results of the historical simulation. The most striking feature shown in the figure is that all future scenarios indicate a flattening of the annual streamflow curve.

While the simulated historical data is characterized by a sharp spring peak (in March), this rise is much more moderate and spread out in the future scenarios. This suggests that snowmelt may occur earlier or more rapidly, leading to a less concentrated spring runoff peak compared to the past.

During the summer months (June–August), the models consistently project persistently low streamflow values, indicating that the summer period may become even drier in the future. It is also evident that the differences between the lines are smallest during these months, meaning the models are in strong agreement regarding the future summer water deficit.

In the winter months (November–December), the simulated historical curve is clearly higher, implying that future scenarios do not indicate a significant increase in water availability by the end of the year. This holds true even though milder winters may result in more precipitation falling as rain rather than snow. This suggests that the watershed’s hydrological response at the

end of the year is unlikely to change substantially compared to current conditions, even if the rain–snow ratio shifts.

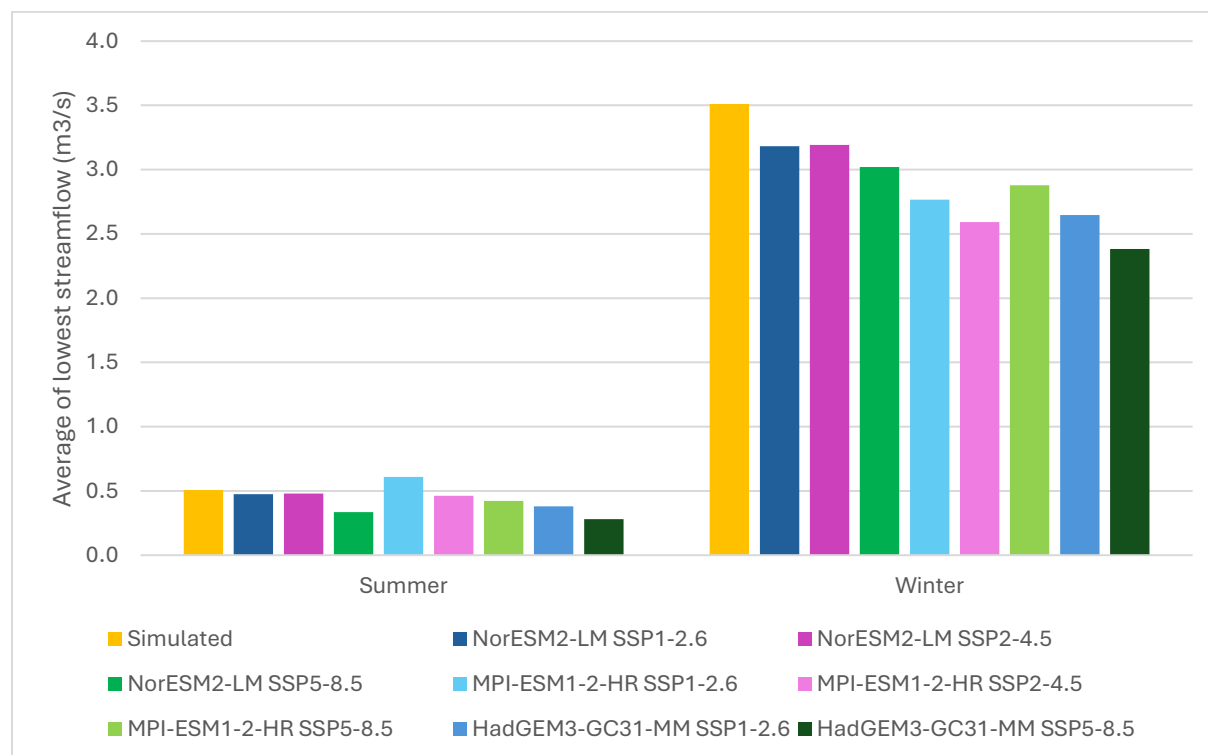


Figure 26 Average summer and winter minimum flows under different scenarios

Table 30 Average summer and winter minimum flows

		Summer (m3/s)	Winter (m3/s)
	Historical simulation	0.508	3.511
NorESM2-LM	SSP1-2.6	0.476	3.183
	SSP2-4.5	0.479	3.192
	SSP5-8.5	0.336	3.021
MPI-ESM1-2-HR	SSP1-2.6	0.608	2.766
	SSP2-4.5	0.463	2.591
	SSP5-8.5	0.423	2.879
HadGEM3-GC31-MM	SSP1-2.6	0.380	2.646
	SSP5-8.5	0.280	2.383

Figure 26 and Table 30 present the lowest summer and winter streamflow values projected under different climate scenarios, compared to the results of the historical simulation. The lowest streamflow values were calculated by selecting the single lowest daily discharge for each year during the summer and winter seasons, respectively. These annual minimum values were then averaged for each climate scenario. As a result, the average minimum summer and winter

streamflow values for each scenario were obtained, offering a clear representation of future trends in low-flow periods.

In the summer season, the historically simulated minimum streamflow was 0.508 m³/s. Compared to this reference value, most future scenarios indicate a decrease. The NorESM2-LM model shows lower values across all examined scenarios, particularly under the SSP5-8.5 pathway, where the projected minimum flow drops to just 0.336 m³/s. The MPI-ESM1-2-HR model shows a slight increase under the SSP1-2.6 scenario (0.608 m³/s), but reductions are observed in the other scenarios. The HadGEM3-GC31-MM model predicts the lowest summer flows in both of its examined scenarios (SSP1-2.6 and SSP5-8.5), especially under SSP5-8.5, where the value declines to 0.280 m³/s. These results suggest that a significant drop in summer flows can be expected, particularly under pessimistic emission scenarios. Such reductions can make summer drought periods more difficult and present serious challenges for water resource management.

In the winter season, the historically simulated minimum streamflow was 3.511 m³/s. Future projections also show generally lower values in winter, although the degree of reduction is less severe than in summer. The NorESM2-LM model projects winter flows around 3 m³/s across all three scenarios, with the lowest value again occurring under SSP5-8.5 (3.021 m³/s). The MPI-ESM1-2-HR model shows a more pronounced decline, especially under SSP2-4.5 (2.591 m³/s). The HadGEM3-GC31-MM model also forecasts reduced winter streamflow, 2.646 m³/s under SSP1-2.6 and as low as 2.383 m³/s under SSP5-8.5.

4. Conclusion

4.1. Data sources and initial analyses

In the initial phase of this study, a comparative analysis was conducted between the global meteorological dataset integrated into the WEAP software – compiled by Princeton University – and local Norwegian data sources. For reference, data provided by SeNorge and the Norsk Klimaservicesenter (NKSS) were used.

The results of this comparison indicated that while the Princeton dataset is easily accessible and globally consistent, it is less reliable at the local level, especially in terms of precipitation intensity and temperature extremes. Due to greater temporal and spatial averaging, it fails to accurately reflect the specific climatic characteristics of Norway.

In contrast, the SeNorge data proved to be more accurate, particularly concerning extreme weather events, which is probably due to the denser network of stations and more detailed and locally calibrated measurement methods. Furthermore, a strong correlation was observed between the SeNorge and NKSS datasets, further supporting the reliability of SeNorge data.

An additional comparison was conducted on snow depth datasets. The results showed a more rapid decrease in snow depth in the SeNorge data, suggesting that snowmelt occurs earlier than indicated by the Princeton dataset. This implies that the Princeton temperature model may not accurately reflect local conditions and fails to capture regional thermal dynamics effectively.

Based on these findings, the SeNorge dataset was selected as the primary source of input data for the hydrological model, as it offers the most reliable representation of Norwegian precipitation, temperature conditions, and snow-related processes.

4.2. Calibration of the model for the whole period

The model was calibrated for the full calibration period (1987–2014). Based on the simulation results, model performance is considered satisfactory: PBIAS = -6.209% and NSE (Nash–Sutcliffe Efficiency) = 0.553. The NSE value exceeds the general acceptability threshold, and the negative PBIAS indicates a slight underestimation of streamflow during the analyzed period. The calibrated parameter values used are summarized in Table 17 (“Final optimal values for the whole period”).

Overall, the model follows the measured data series more accurately at low and medium flows, but during periods of high flows, such as those caused by sudden rainfall or snowmelt, the model tends to underestimate maximum flows.

4.3. Seasonal calibration of the model for summer and winter periods

Based on the seasonal evaluation of the model's performance, it can be concluded that the simulation for the summer period yielded significantly better results ($NSE = 0.719$, $PBIAS = -6.664\%$), while performance during the winter period remained weaker ($NSE = 0.496$, $PBIAS = +1.684\%$) and did not exceed the metrics calculated for the full period.

This can partly be attributed to the fact that the internal hydrological processes in the WEAP model form a non-linear system, meaning that parameters influence each other and may behave differently across seasons. In addition, the difference in model performance between summer and winter months suggests that the model is more sensitive and accurate to the more intense evaporation and water movement in summer, while it is less accurate for snow-related processes in winter.

The lower performance in the winter period is also partly due to the reduction in available data when the time series is divided seasonally, which gives less statistically stable results. Additionally, for statistics calculated over the full calibration period, errors tend to balance each other out. In contrast, this compensatory effect is absent in seasonal analyses, which can lead to a deterioration in statistical indicators.

4.4. Model validation

The model was validated using the remaining part of the time series, independent of the calibration period, covering the years 1974–1986. The validation yielded the following statistical results: $PBIAS = -9.84\%$, $NSE = 0.483$.

The negative PBIAS indicates that the model continued to underestimate streamflow. Although the NSE value falls slightly below the generally accepted "satisfactory" threshold, it still indicates moderate model performance, particularly given that the parameter set was not specifically tuned for this period.

In summary, the results of the validation show that the model can reproduce the streamflow outside the calibration period in a limited but still applicable way. The predictive ability of the model therefore cannot be considered fully generalizable, but it is still capable of identifying basic hydrological patterns and streamflow trends.

4.5. Assessment of future climate scenarios

Future hydrological conditions were analyzed using three global climate models (NorESM2-LM, MPI-ESM1-2-HR, HadGEM3-GC31-MM) under various Shared Socioeconomic Pathways (SSPs). The analyzed period covered 30 years, from 2060 to 2090. For the simulations, parameter values previously calibrated for the full 1987-2014 period were used to ensure comparability with the historical reference period (1980-2010).

Simulation results for future streamflow indicate a systematic decrease in annual average discharge, ranging between 29% and 39% depending on the model and scenario. The most significant change occurs in the magnitude of spring flood peaks, which previously occurred as a result of snowmelt. These peaks are expected to become more moderate in the future due to reduced snow accumulation and earlier melting. Additionally both summer and winter minimum streamflows are also expected to decrease significantly, especially under high-emission scenarios such as SSP5-8.5.

Overall, the projected decline in streamflow, particularly during summer, poses a serious challenge for future water resource management. These findings highlight the need for climate adaptation measures, including expanded reservoir capacity, improved water use efficiency, and strategies to address seasonal shifts in water demand.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, **Professor Tor Haakon Bakken**, for his professional guidance, continuous support, and encouragement throughout the development of this thesis. I also thank **Baranya Sándor**, my co-supervisor, for his assistance and the information he provided.

I am deeply grateful to my **family** and **friends** for their unwavering support, motivation, and belief in me during this journey.

5. References

- Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., Burke, E., Dittus, A. J., Florek, P., Gray, L. J., Haddad, S., Hardiman, S. C., Hermanson, L., Hodson, D., Hogan, E., Jones, G. S., Knight, J. R., Kuhlbrodt, T., Misios, S., ... Sutton, R. T. (2020). Historical Simulations With HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, 12(6). <https://doi.org/10.1029/2019MS001995>
- Bækø Johanne, & Pedersen Anja Iselin. (2024). *Norwegian streamflow reference dataset for climate change studies*.
- Bañares, E. N., Mehboob, M. S., Khan, A. R., & Cacal, J. C. (2024). Projecting hydrological response to climate change and urbanization using WEAP model: A case study for the main watersheds of Bicol River Basin, Philippines. *Journal of Hydrology: Regional Studies*, 54. <https://doi.org/10.1016/j.ejrh.2024.101846>
- Bergström, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments*.
- Bergström, S. (1995). The HBV model. In *Computer Models of Watershed Hydrology* (pp. 443–476). Water Resources Publications.
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1–2), 18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>
- Beven, K. J., & Binley, A. M. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3), 279–298.
- Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models*. Watermark Numerical Computing.
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2), 135–143.
- Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., Von Storch, J. S., Brüggemann, N., Haak, H., & Stössel, A. (2019). Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP). *Geoscientific Model Development*, 12(7), 3241–3281. <https://doi.org/10.5194/gmd-12-3241-2019>

- Hanssen-Bauer, I., Førland, E. J., Haddeland, I., Hisdal, H., Mayer, S., Nesje, A., Nilsen, J. E. Ø., Sandven, S., Sandø, A. B., Sorteberg, A., & Ådlandsvik, B. (2017). *Climate in Norway 2100 Lead authors-a knowledge base for climate adaptation*.
www.miljodirektoratet.no/M741
- IPCC. (2021). *Climate Change 2021: The Physical Science Basis*.
- Lawrence, D., & Hisdal, H. (2011). *Hydrological projections for floods in Norway under a future climate*.
- Li, X., Zhao, Y., Shi, C., Sha, J., Wang, Z. L., & Wang, Y. (2015). Application of Water Evaluation and Planning (WEAP) model for water resources management strategy estimation in coastal Binhai New Area, China. *Ocean and Coastal Management*, 106, 97–109. <https://doi.org/10.1016/j.ocecoaman.2015.01.016>
- Mahamadou Mounir, Z., Ming Ma, C., & Amadou, I. (2011). Application of Water Evaluation and Planning (WEAP): A Model to Assess Future Water Demands in the Niger River (In Niger Republic) HDI: Human Development Index IWRI: International Water Management Institute IWRM: Integrated Water Resources Management MDGs: Millennium Development Goals SEI: Stockholm Environment Institute UNDP: United Nations Development Program WEAP: Water Evaluation And Planning WDI: World Development Indicators. *Modern Applied Science*, 5(1).
www.ccsenet.org/maswww.ccsenet.org/mas
- Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*, 58(6), 1763–1785. <https://doi.org/10.13031/trans.58.10715>
- Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornbluh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., ... Marotzke, J. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in Modeling Earth Systems*, 10(7), 1383–1413.
<https://doi.org/10.1029/2017MS001217>
- Nash, J. E., & Sutcliffe, J. V. (1970). Not to be reproduced by photoprint or microfilm without written permission from the publisher RIVER FLOW FORECASTING THROUGH

CONCEPTUAL MODELS PART I-A DISCUSSION OF PRINCIPLES*. In *Journal of Hydrology* (Vol. 10). © North-Holland Publishing Co.

Olabiwonnu, F. O. (2020). *Assessment of the Role of Hydropower Reservoirs in Sustaining Ecologically Viable Low Flow Conditions under Present and Future Climate*.

Olabiwonnu, F. O., Bakken, T. H., & Anthony, B. (2022). Achieving sustainable low flow using hydropower reservoir for ecological water management in Glomma River Norway. *Sustainable Water Resources Management*, 8(2). <https://doi.org/10.1007/s40899-022-00643-y>

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9), 3461–3482. <https://doi.org/10.5194/gmd-9-3461-2016>

Pereira, L. S. (1998). *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56*. <https://www.researchgate.net/publication/235704197>

Pramudith, S., Supervisor, H., & Bakken, H. (2022). *Assessment of Suitability of WEAP for Studies of the Flood Dampening Effects of Reservoirs in Norway*.

Refsgaard, J. C., & Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32(7), 2189–2202.

Sardar Shahraki, A., Shahraki, J., & Hashemi Monfared, S. A. (2016). An Application of WEAP Model in Water Resources Management Considering the Environmental Scenarios and Economic Assessment Case Study: Hirmand Catchment. *Modern Applied Science*, 10(5), 49. <https://doi.org/10.5539/mas.v10n5p49>

Seland, Ø., Bentsen, M., Olivie, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y. C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schanke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., ... Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, 13(12), 6165–6200. <https://doi.org/10.5194/gmd-13-6165-2020>

Sieber, J. (2006). *WEAP Water Evaluation and Planning System*.

<https://scholarsarchive.byu.edu/iemssconference/2006/all/397>

Sieber, J., & Purkey, D. (2015). *Water Evaluation And Planning System USER GUIDE for*

WEAP 2015. <http://www.weap21.org><http://www.sei-us.org>

Winter, J. M., Young, C. A., Mehta, V. K., Ruane, A. C., Azarderakhsh, M., Davitt, A.,

McDonald, K., Haden, V. R., & Rosenzweig, C. (2017). Integrating water supply constraints into irrigated agricultural simulations of California. *Environmental Modelling and Software*, 96, 335–346. <https://doi.org/10.1016/j.envsoft.2017.06.048>