Report

R-24-09

October 2024

Semi-distributed hydrological modelling of Lake Eckarfjärden drainage basin

Water flux estimations based on HYPE-modelling

Alexander Koutsouris

SVENSK KÄRNBRÄNSLEHANTERING AB

SWEDISH NUCLEAR FUEL
AND WASTE MANAGEMENT CO

Box 3091, SE-169 03 Solna
Phone +46 8 459 84 00
skb.se

SVENSK KÄRNBRÄNSLEHANTERING

ISSN 1402-3091
SKB R-24-09
ID 2048247
October 2024

Semi-distributed hydrological modelling of Lake Eckarfjärden drainage basin

Water flux estimations based on HYPE-modelling

Alexander Koutsouris
Swedish Environmental Research Institute, IVL

Keywords: Hydrological modelling, HYPE, Eckarfjärden.

This report concerns a study which was conducted for Svensk Kärnbränslehantering AB (SKB). The conclusions and viewpoints presented in the report are those of the author. SKB may draw modified conclusions, based on additional literature sources and/or expert opinions.

This report is published on www.skb.se

Abstract

Svensk Kärnbränslehantering AB (SKB) is in the permitting process for the final repository for spent nuclear fuel (SFK) and is therefore examining the current and future hydrology at Forsmark. The regional hydrology in Forsmark has previously been conceptualized and modelled using physical and fully distributed numerical models. SKB now seeks to use a complementary hydrological model to generate input data for the biosphere radionuclide transport model which is then used to estimate potential exposure doses of radionuclides released from the repository.

This report evaluates the capability of the HYPE model to simulate hydrological fluxes, using the Lake Eckarfjärden drainage basin in Östhammar, Sweden, as a study area. The results demonstrated that HYPE manage to capture the main hydrological dynamics within Lake Eckarfjärden drainage basin. Average KGE value for the two separate evaluation periods was 0.74 for streamflow, 0.62 for lake stage and an NSE value of 0.55 for groundwater. Model performance for lake stage was likely penalized for the short calibration and evaluation period while the performance score for groundwater is likely affected by comparing simulated drainage basin average to a single observation well. Furthermore, HYPE tracks hydrological fluxes and storages within the model, which means that the contributions of groundwater, overland flow etc. can be assessed.

Overall, the model effectively represented processes such as infiltration, snow accumulation and melt, and water routing. The agreement between model simulations and observed data indicated its capability to capture the main dynamics of the hydrological system. This suggests that HYPE can provide estimates of hydrological fluxes in similar drainage basins.

Sammanfattning

Svensk Kärnbränslehantering AB (SKB) befinner sig i tillståndsprocessen för slutförvaret för använt kärnbränsle (SFK) och undersöker därför den nuvarande och framtida hydrologin i Forsmark. Den regionala hydrologin i Forsmark har tidigare konceptualiserats och modellerats med hjälp av fysiska och fullständigt distribuerade numeriska modeller. SKB söker nu använda en kompletterande hydrologisk modell för att generera indata till biosfärens radionuklidanstransportmodell, som sedan används för att uppskatta potentiella exponeringsdoser av radionuklider som släpps från deponeringsplatsen.

Denna rapport utvärderar HYPE-modellens förmåga att simulera hydrologiska fluxer, med Eckarfjärdens avrinningsområde i Östhammar, Sverige, som studieområde. Resultaten visade att HYPE lyckas fånga den huvudsakliga hydrologiska dynamiken inom Eckarfjärdens avrinningsområde. Genomsnittlig KGE-värde för de två separata utvärderingsperioderna var 0,74 för vattenflöde, 0,62 för sjövattenstånd och ett NSE-värde på 0,55 för grundvatten. Modellens prestanda för sjövattenstånd påverkades troligen negativt av den korta kalibrerings- och utvärderingsperioden, medan prestandan för simulering av grundvatten troligen påverkas av att simulerad grundvattennivå för hela avrinningsområdet jämfördes med en enskild observationsbrunn. Vidare spårar HYPE flöden och magasinering av vatten, vilket kan tillämpas för att uppskatta fördelningen av Eckarfjärdens tillrinning mellan olika källor såsom grundvatten eller ytliga flöden.

Sammanfattningsvis representerade modellen effektivt processer som infiltration, snöackumulering och smältning samt vattentransport. Överensstämmelsen mellan modellsimulationer och observerade data indikerade dess förmåga att fånga de huvudsakliga dynamikerna i det hydrologiska systemet. Detta tyder på att HYPE kan tillämpas för uppskattning av hydrologiska flöden i liknande avrinningsområden.

Contents

1 Introduction

1.1 Background

SKB is updating its understanding of the current and future hydrology at Forsmark. The regional hydrology in Forsmark has previously been conceptualized and modelled using physical and fully distributed three-dimensional numerical models. With help from those modelling efforts, a number of biosphere objects have been identified, which are defined as areas in the future landscape where radionuclides released from the repository are expected to reach the surface through ground water discharge. Each biosphere object also has an associated drainage basin.

In order to estimate the potential exposure to radionuclides, there is a need to simulate the rainfall-runoff response of these drainage basins to be used as input for parameterizing the biosphere radionuclide transport and accumulation model (henceforth referred to as the Biosphere model). A hydrological model can here be used for forward modelling of the hydrological fluxes in yet-unseen landscapes, the formation of which is primarily driven by isostatic uplift.

The required output from the hydrological model is seasonal to annual water balances for surface and groundwater flows. The output fluxes of the hydrological model should also correspond to the input fluxes of the Biosphere model in order to avoid a need for interpretation between these models. The main fluxes are 1) catchment area evapotranspiration, 2) surface water, 3) shallow groundwater and 4) deep ground water.

The HYPE (HYdrological Predictions for the Environment, Lindström et al. 2010) model has been selected as a potential candidate for semi-distributed model for simulating water fluxes in the Forsmark area and assess uncertainty bounds for the hydrological fluxes. Particularly as an alternative to more complex, physical, hydrological models. To investigate its suitability to simulate the hydrological fluxes at Forsmark, it was applied to Lake Eckarfjärden drainage basin, located in the Forsmark area within the municipality of Östhammar.

This report aims to:

- Investigate HYPE's capability to simulate the hydrological fluxes in the Lake Eckarfjärden drainage basin.
- Investigate the possibility to use a simplified model structure to simulate hydrological fluxes to biosphere objects at Forsmark, as an alternative to more complex physical models.

1.2 HYPE

HYPE is a semi-distributed hydrological model for simulating the water balance and nutrient fluxes within drainage basins. The model is actively developed by the Swedish Meteorological and Hydrological Institute (SMHI) and the model code is open source. HYPE has been applied at various scales from small drainage basins to national, continental, and global scales (e.g., S-HYPE and WWH; Arheimer et al., 2020; Strömqvist et al., 2012). A drainage basin may be subdivided into multiple sub-basins in HYPE, which can be connected by rivers or regional groundwater flow. Sub-basins are further discretized into classes based on land-use and soil type[1] properties[2] called "Soil type Land use Combinations" or "SLCs" (Figure 1-1).[3] The SLCs are combinations of similar land-uses and soil types across the model domain.

[1] HYPE use the term "soil type", which here is used synonymous to regolith. Regolith is defined as unconsolidated, loose, heterogeneous superficial deposits covering solid rock, including soil.

[2] These properties can be (generally) described as land use, topography, vegetation/crop type, soil type, soil depth, number of soil layers, distance between the soil surface and the nearby stream bottom.

[3] An SLC can viewed as a "hydrological response unit" (HRU) which is the smallest spatial unit used in the parameterization of a lumped hydrological model.

HYPE uses a lumped parameter modelling approach to simulate the hydrological response to meteorological forcing (precipitation and evaporation/evapotranspiration) for each sub-basin. This hydrological response can be characterized as changes in surface-, soil- and groundwater storages which in turn generate “runoff” which is then routed to the sub-basin outlet as “surface water”. This discretization of a drainage basin into sub-basins, combined with the lumped parameter approach used for modelling hydrological response within the sub-basins, is why HYPE is considered a semi-distributed hydrological model.

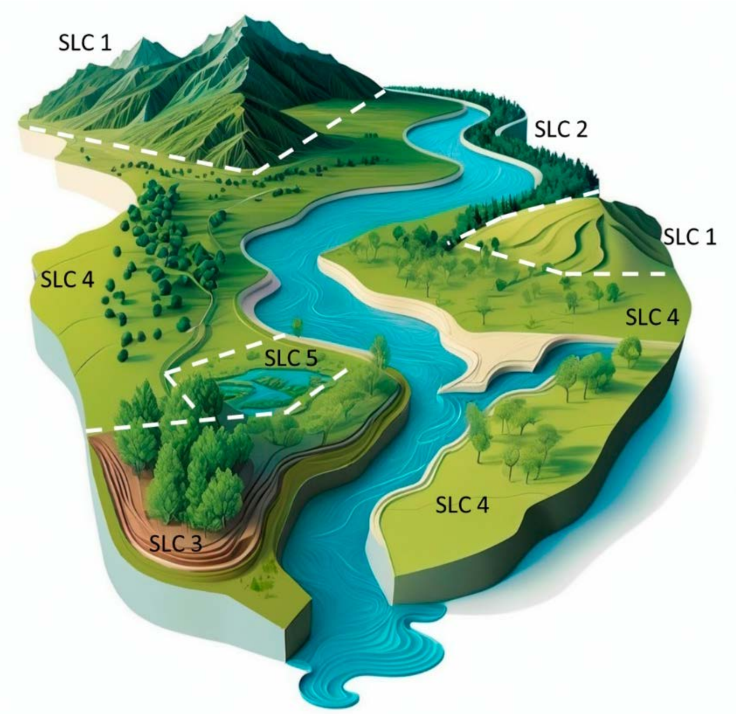

Figure 1-1. *Example of a hydrological drainage basin in HYPE with Soil type Land use Combinations (SLCs) represented. Modified from image generated with OpenArt.ai.*

1.2.1 Runoff generation

A HYPE model may be divided into sub-basins; the hydrological response in each sub-basin is calculated using a lumped parameter approach, i.e. the parameters of the model represent the spatially averaged characteristics of the hydrological system. Spatial averaging of parameters is done via the classification of the SLCs. Each sub-basin contains one or more SLCs which are defined by the modeler. The SLCs are then represented in the sub-basin as a fraction of the total area and do not have a topological relation to each other, which means that water is not routed between SLCs within a sub-basin. Each SLC can be divided vertically into one to three soil layers where each layer has a different thickness, set of soil hydraulic properties and runoff response (Figure 1-2). Storage changes due to evaporation from the soil layers decrease with depth and is limited to the upper two layers. Soil water is generated through infiltration of precipitation (P), snow melt or ponded surface water. Groundwater, in turn, is generated through percolation of soil water when the water content of the soil is above the specified field capacity. Within each SLC, each soil layer may be defined using a unique parameter set which determines each soil layer's infiltration, percolation and runoff generation. Lakes, rivers and glaciers are special types of SLCs, treated differently from land SLC in the model.

Within HYPE, "runoff" is characterized as surface water generated from either groundwater or excess surface water (Figure 1-2). Surface runoff may be generated when the ground water table rise above the surface level or due to excess infiltration (i.e. precipitation rate exceeds the infiltration rate). It is also generated as a function of the groundwater level and the parameters related to groundwater recession. Runoff from groundwater is generated when soil water is above field capacity and the groundwater level within an SLC is above a set "stream depth" level (i.e. drainage level) for that SLC. Water stored below the stream depth level does not contribute to runoff generation. HYPE does not explicitly separate runoff into base flow and quick flow. However, runoff generated during periods between precipitation and snowmelt events, and during periods of low flow can be assumed to be base flow. Therefor parameters governing low flows (e.g. soil recession coefficients and stream depth) can be assumed to relate to base flow.

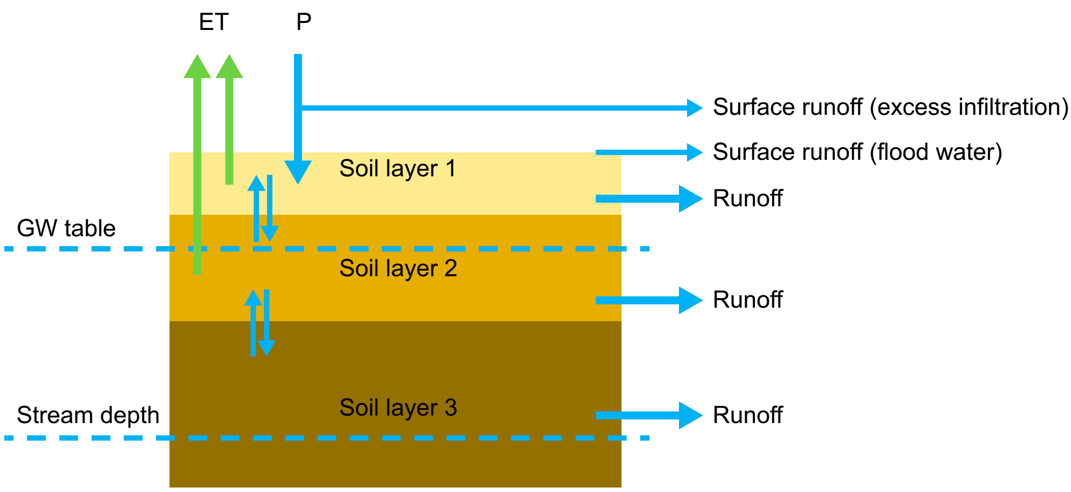

Figure 1-2. *Water fluxes and runoff generation in an SLC for the model setup. GW is groundwater, P is precipitation and ET is evapotranspiration (modified from www.smhi.net/hype/wiki).*

In HYPE, all runoff is routed downstream as surface water via "local-" and "main-rivers" (see Section 1.2.2 and Figure 1-3. Routing of water flow through HYPE). This means that, effectively, all modelled discharge between sub-basins in HYPE is represented as surface water and that the main HYPE model does not explicitly model groundwater discharge between sub-basins. There is a HYPE module available which accounts for drainage of groundwater from the lowest soil box of SLCs (Figure 1-3) to a regional groundwater flow to a downstream sub-basin and/or an outlet lake (more information given in Appendix A). However, the amount of water drained from this layer is defined by the model user. This module has not been used in the model setup for this study (see model setup, Section 2.1).

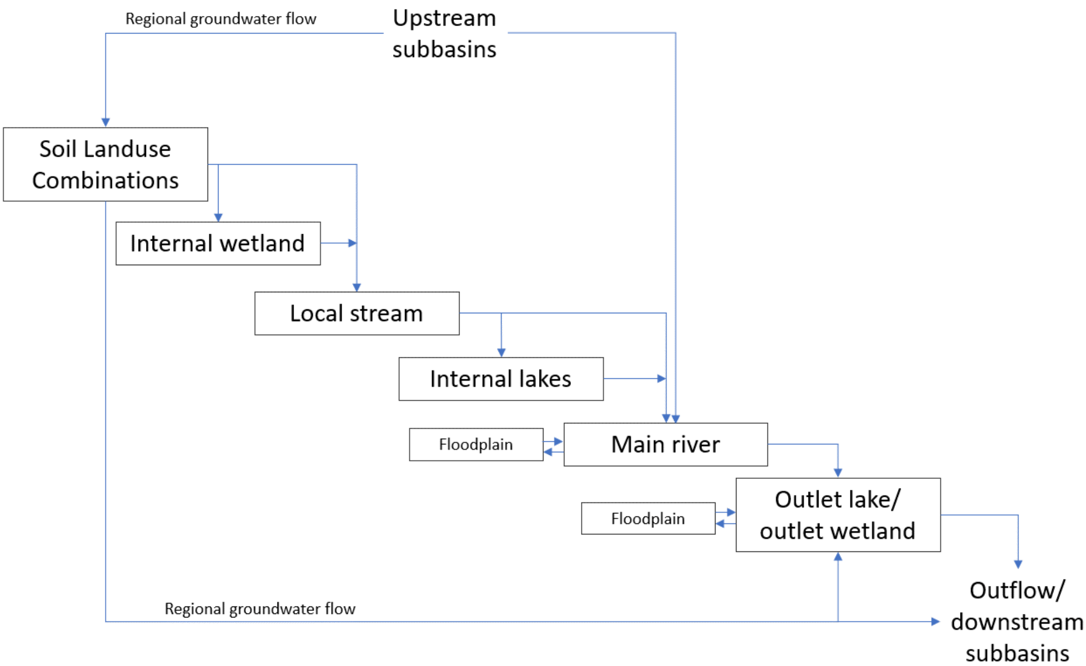

Figure 1-3. *Routing of water flow through HYPE. Lakes, wetlands, floodplains, and regional groundwater flow are optional. All flows between storages are surface water flows unless otherwise specified. All storages are subjugated to precipitation and evapotranspiration. Figure based on figure(s) presented on the HYPE wiki pages (hype.smhi.net/wiki).*

1.2.2 Rivers and lakes

As HYPE is a semi-distributed, lumped parameter model, there is no topological relationship between lakes, rivers and SLCs within a drainage basin/sub-basin, but rather as storage entities which receive and discharge runoff. HYPE accounts for six different types of surface water reservoirs within a sub-basin: "internal lakes", "outlet lakes", "main rivers", "local rivers", "internal wetlands" and "outlet wetlands". These are special SLCs in HYPE and hydrological processes for these SLCs are governed by separate calculations. Furthermore, there are a few different ways to represent wetlands in HYPE, these are further described in Section 1.2.3.

Local rivers represent ditches and streams which can only receive runoff from SLCs within a sub-basin. In HYPE, all of the local rivers within a sub-basin are modeled as a single, "conceptual" surface waterbody with a characteristic length which is a function of the sub-basin area. All runoff generated within the sub-basin is routed to a local river (Figure 1-3). Similar to a local river, a local lake is meant to represent all of the lakes contained within the sub-basin boundaries and can receive and delay runoff from the local rivers. All runoff which is routed to a local lake is done so via the local river (this is accounted for as a constant percentage of the runoff within the local river). The main river in HYPE is the waterway within the sub-basin that can receive runoff from upstream sub-basins and convey water to downstream sub-basins and outlet lakes. The main river is represented as a lumped model within a sub-basin. The main river of a sub-basin receives water either internally from local rivers, internal lakes, and flood plains or externally from upstream sub-basins' outlet lake, outlet wetland and main river. The outlet lake or outlet wetland only contribute to the main river of downstream sub-basins.

Water may also flow from surface waters to floodplains and back if the floodplain module is used. The floodplain module is not used in the model setup for this study (see model setup, Section 2.1). All runoff from the local rivers and local lakes is routed to the main river. An outlet lake in HYPE is a lake that is situated at the sub-basin outlet. An outlet lake receives all its runoff via the main river and, possibly, from regional groundwater flow. The discharge from the outlet lake, conveyed to downstream sub-basins, is calculated based on a rating curve as a function of water level above the lake outlet threshold level.

1.2.3 Wetland representation

There are multiple ways of representing wetlands in HYPE. They can be assigned an SLC and be represented as a percentage of the drainage basin surface area and with specific soil characteristics. Wetlands also have a special land class used for nutrient modelling. Finally, wetlands adjacent to the outlet lake could also be represented as a floodplain using the floodplain module (Appendix A).

Wetlands can be represented as an SLC in HYPE by assigning parameter values reasonable for a wetland to a specific SLC (SLC 9 in the model setup used in this study). Using this method, water retention and release is determined as a function of groundwater table as well as parameters related to groundwater recession. This method was used to represent wetlands in the model setup used in this study (see model setup, Section 2.1.3).

Wetlands represented as a special land class use three vertical soil boxes and allow storage of water above ground level. Water is released from the wetland as a function of water level, the wetland threshold, and a rating curve for the wetland. There are two types of wetlands: internal wetlands and outlet wetlands. Internal wetlands represent all inland wetlands and receive their water as a fraction of local runoff. Outlet wetlands receive their water from the sub-basin main river and release its water to the main river of the downstream sub-basin. Precipitation and evaporation also impact the water balance of both internal and outlet wetlands. The capability to account for wetlands in HYPE as a special land class was not used in the model setup for this study. Outflow from Lake Eckarfjärden is be assumed to be dominated by lake levels and lake dynamics rather than the dynamics of the minor wetlands adjacent to Lake Eckarfjärden, which excludes the use of an outlet wetland. Internal wetlands were not used due to the minor presence of wetlands in Lake Eckarfjärden drainage basin (1 %) and the added model complexity to represent internal wetlands.

The floodplain module (Appendix A) can be used to represent wetlands adjacent to the outlet lake. While there are wetlands adjacent to Lake Eckarfjärden, the flood plain module was not implemented in this model setup for this study (see model setup, Section 2.1) as the added complexity of this module was deemed unnecessary to fulfill the objectives of the study.

2 Model implementation

2.1 Conceptual model for Eckarfjärden

Lake Eckarfjärden drainage basin was conceptualized as a single drainage basin with Lake Eckarfjärden as an outlet lake. The streamflow gauge used to evaluate simulated drainage basin streamflow is located 180 m downstream from the Lake Eckarfjärden outlet (Section 2.1.2). The disparity between lake outlet and stream gauge is assumed to have negligible impact on the capability to simulate streamflow considering that that the difference in catchment area for the lake outlet and for the stream gauge is < 1 % (Section 2.1.2). The land area of Lake Eckarfjärden drainage basin was conceptualized as 10 SLCs (Section 2.1.1), where the dominating SLC was forested till (SLC 6 in Table 2-2). Wetlands were conceptualized as an SLC (see Section 2.1.3), contributing to runoff in the same way as other SLCs (see Section 1.2.1).

In the model setup used for Lake Eckarfjärden water is routed through 1) infiltration of precipitation into the soil, 2) runoff generated from SLC in the local river, 3) runoff in main river, represented as 1D based on river length, 4) inflow to outlet lake and 5) outflow from outlet lake and drainage basin (Figure 2-1; Section 2.1.3). As the water is routed through the drainage basin, parameters related to the local and main river acts as a delay and dampening of the water flux. It should also be noted that while HYPE tracks water contribution from each soil box to the local river, there is no direct water flux between a specific SLC's soil box and the outlet lake.

2.1.1 Soil type Land use Combinations

SLCs in HYPE are user defined. Six types of soil and three types of vegetation were used for the hydrological model of Lake Eckarfjärden (Table 2-1; Figure 2-2; Figure 2-3). Land use classification was based on the vegetation map and soil map. Open was interpreted as "mire" where it overlapped with organic soils and as grass in other occurrences. "Forrest" and "water" in the land use classes corresponds with the vegetation classes. Together land use, soil and vegetation classes were aggregated into 10 SLCs to represent Lake Eckarfjärden drainage basin (Table 2-2). The SLCs were formed based on soil and land use data from SKB (Table 2-4) and were reclassified using GIS into the classes detailed in Table 2-2 (see also Appendix B for reclassification scheme). Combined SLCs were then created using the topological relation between soil types, land use and vegetation type. Some combinations of SLCs were integrated into more dominant SLCs due to their small representation within Lake Eckarfjärden drainage basin and to decrease the granularity of the model setup used in this study.

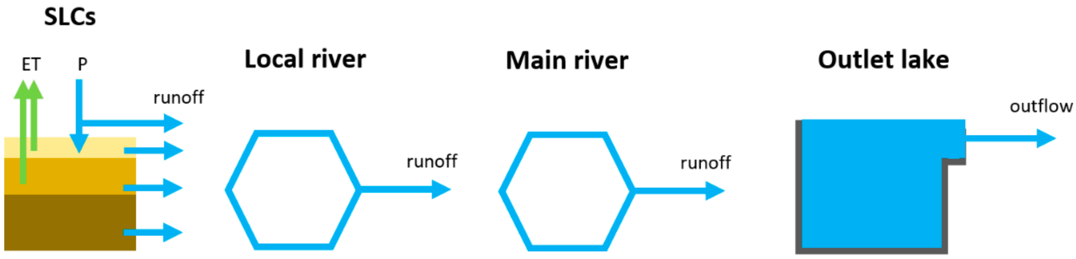

Figure 2-1. *Schematic representation of runoff generation in HYPE model implemented for this study. Runoff generated from all SLCs are first routed through a single conceptual representation of local rivers and then through a main river. The main river feeds an outlet lake where the outflow from the drainage basin is generated.*

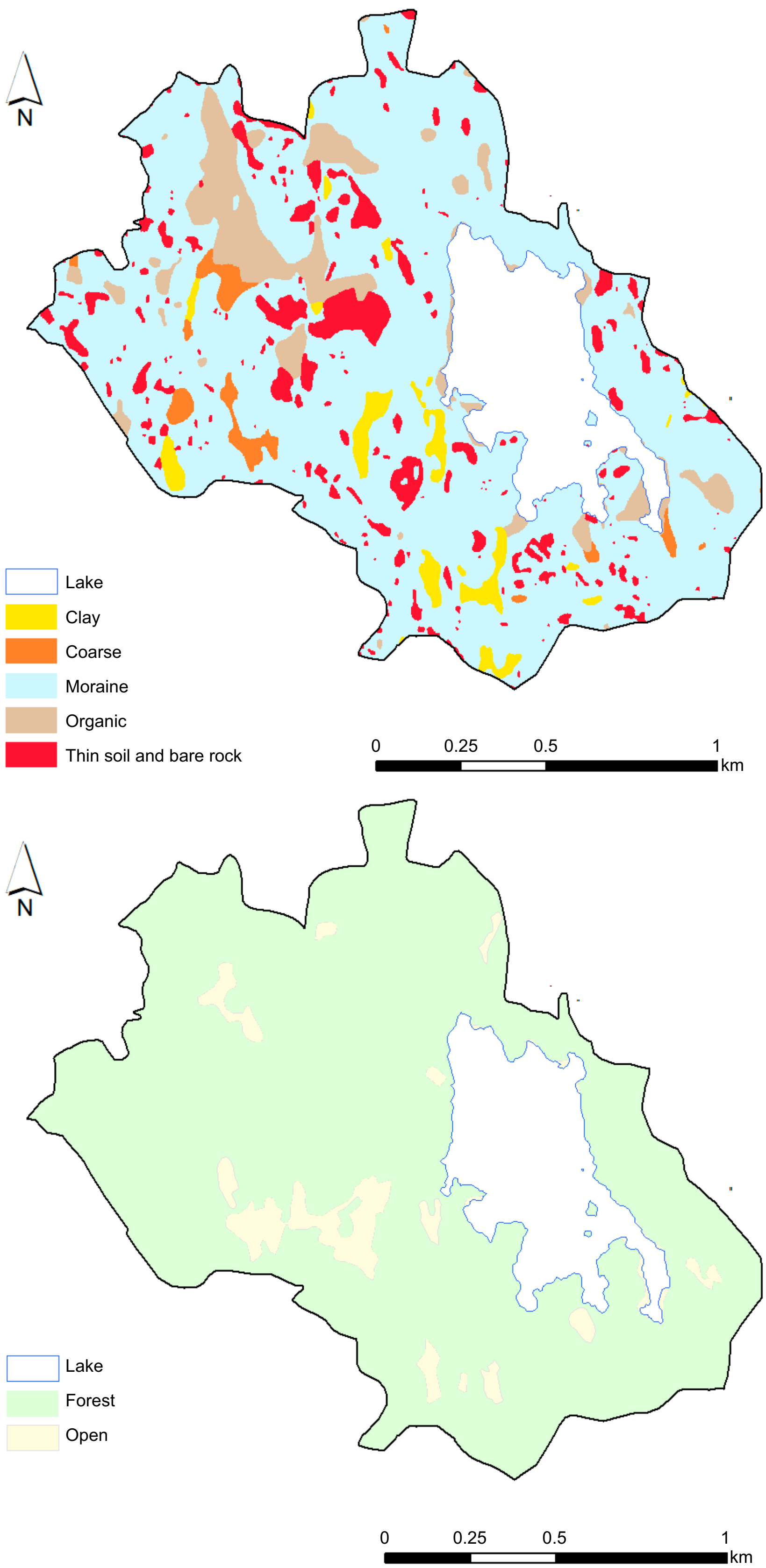

Figure 2-2. *Top: Soil map for soil types within Lake Eckarfjärden drainage basin. Soil types have been aggregated into 5 classes. Soil cover, scale 1:25 000 from Jordartskartan © SGU. Bottom: Vegetation map within Lake Eckarfjärden drainage basin. Land use, scale 1:15 000 from Fastighetskartan © Lantmäteriet. Extent of lake Eckarfjärden corresponds to a a lake-level of +5.37 m a.s.l.*

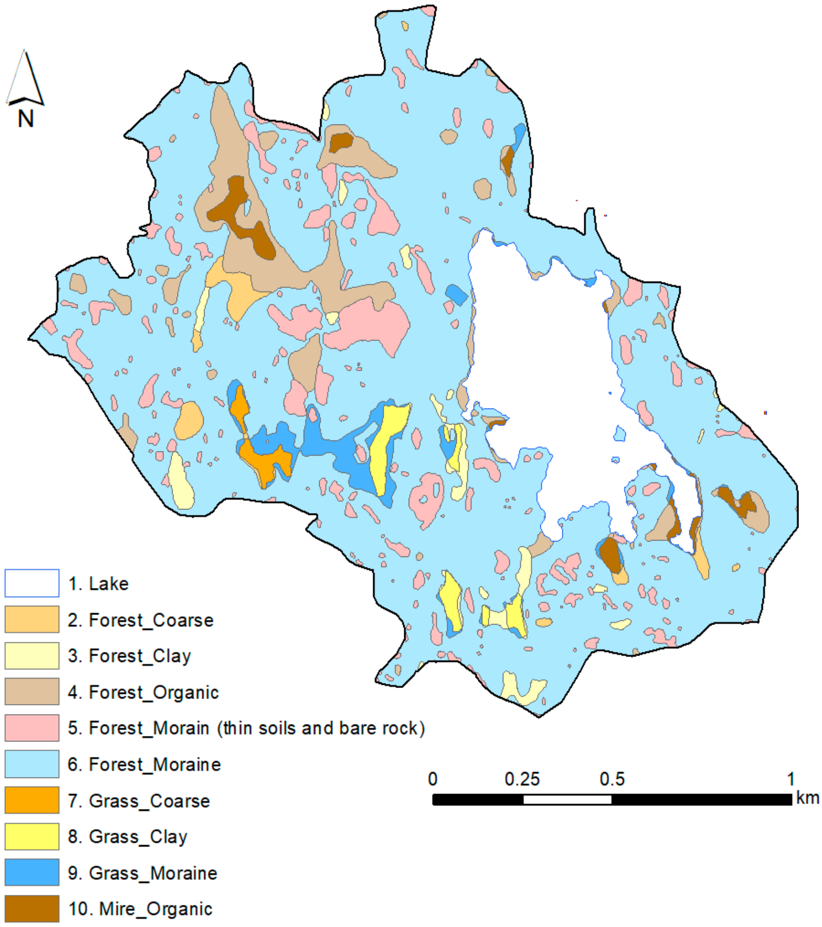

Figure 2-3. *Soil type Land use Combinations (SLCs) used for representing Lake Eckarfjärden drainage basin in the used in the model setup for this study.*

Each SLC was divided into a maximum of three vertical layers (Table 2-2). First, soil stratification data was used to determine the average thickness of soil types in the soil profile within each SLC. Then, the profile was lumped into up to three layers based on soil types and thicknesses of the soil strata (layer 1–3 in Table 2-2). In general, the first layer corresponds to an upper layer of a few decimeters and acts as a quick-responding upper layer. The thickness of soil layer 2 in an SLC was defined on a case-by-case basis depending on the characteristics of the soil (Appendix B). Layer 3 represent the remaining soil down to the bedrock. Each SLC may also have a unique stream depth, here a stream depth of 1 m was used for all SLCs except for the SLC representing thin soils which have a stream depth of 0.2 m.

Table 2-1. Soil, land use and vegetation types that were used to identify the main SLCs within Lake Eckarfjärden drainage basin.

Land use	Soil type	Vegetation type
Lake	Organic	Open
Mire	Clay	Forest
Forest	Coarse	Water
Grass	Moraine	
Other land use	Thin soil and bare rock	
	Water	

Table 2-2. Soil type Land use Combinations (SLCs) used in model setup. "Layer 1–3" corresponds to the distance from the soil surface to the bottom of that soil layer.

SLC	Land use	Soil type	%	layer 1 [m]	layer 2 [m]	layer 3 [m]
1	Lake	Water	12.3	-	-	-
2	Forest	Coarse	1.2	3	0.5	1.9
3	Forest	Clay	1.6	3	0.8	1.8
4	Forest	Organic	6.5	3	0.6	1.4
5	Forest	Moraine*	7.5	0.5	-	-
6	Forest	Moraine	65.7	3	0.3	1.3
7	Grass	Coarse	0.5	3	0.2	1.6
8	Grass	Clay	1.1	3	1	1.7
9	Grass	Moraine	2.2	3	0.3	1.3
10	Mire	Organic	1.3	3	0.7	1.4

* Thin soils and bare rock.

2.1.2 Lake representation

Lake Eckarfjärden is here defined as an outlet lake (i.e., the outlet of the drainage basin) even though the outlet of the lake is approximately 180 m upstream from the discharge station (PFM002668) used for stream flow data in the hydrological model. As mentioned previously, the difference in catchment area for Lake Eckarfjärden and for the discharge station is < 1 % and thus assumed to be negligible. The area of the lake was set to 0.284 km^2, which is representative for a lake level of +5.37 m a.s.l, and an average lake depth of 0.91 m (Table 2-4).

A comparison of streamflow data at the drainage basin outlet and Lake Eckarfjärden water levels show that the relationship between the two are non-stationary (Figure 2-4), particularly during the period of 2011–2013. There is a drift in lake stage measurements relative to calculated streamflow, which may be a result of changes in the stream channel, the lake outlet threshold or a drift in measured data. Lake data for this period was therefore omitted from the model calibration. Even so, the non-stationarity is still present in the data set post 2013 and likely has an adverse impact on model performance as the lake rating curve in HYPE assumes stationarity. The lake rating curve in HYPE is given as:

$$q_t = k \times (w_t - w_0)^p$$

where q_t is the streamflow at timestep t, k is a constant, is an exponent p, w_t is water level at timestep t and w_0 is the lake outlet threshold. The parameters for k and p were automatically calibrated in the model setup (see Section 2.3 and Appendix C) and the lake outlet threshold was set to 5.28 m. Literature values for lake threshold is 5.34 (Bosson et al., 2008), though simulated lake levels indicated a clear need to lower the model value threshold. The need to lower the threshold may be due to the non-stationarity in the relationship between lake level and streamflow.

2.1.3 Wetland representation

Wetlands in the Lake Eckarfjärden HYPE model are represented as an SLC with "Mire" as land use and "Organic" as soil type (SLC 10 in Table 2-2). Thus, water flow from the wetland is routed from the soil boxes, through the "local stream" network, continuing to "main river" and then to the lake in the model setup. The model setup for Lake Eckarfjärden thereby does not distinguish between wetlands directly connected, i.e. adjacent, to the lake and wetlands further upstream. The soil parameter ranges used in the calibration for the Wetland SLC can be found in Appendix C.

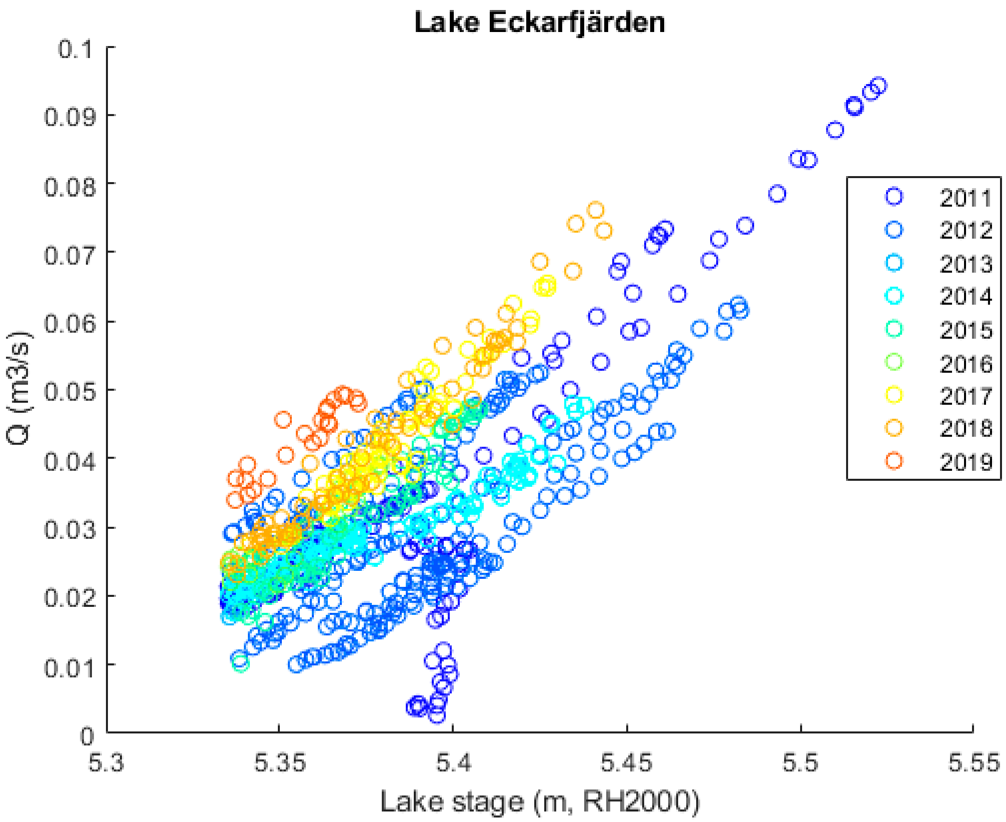

Figure 2-4. *Lake stage and lake outlet discharge relationship at the outlet of Lake Eckarfjärden. The colors correspond to the year of the measurement point.*

2.2 Data

2.2.1 Time series data

The model setup for Lake Eckarfjärden is driven by time series data of precipitation (P), potential evaporation (ET_p) and temperature (T) data (Table 2-3). Precipitation and ET_p constitutes influx and outflux of water respectively. Temperature is used in HYPE for determining whether precipitation falls as snow or rain and for the snow-melt routine. Time series of streamflow at the drainage basin outlet, lake level of Eckarfjärden and groundwater level at a single well (Table 2-3) were used for calibrating and evaluating the model. Time series data on snow water equivalent was used as auxiliary data to assess the performance of the snow-melt routine.

The hydro-meteorological data used in the model setup of this study is shown in Table 2-3, Figure 2-5 and Figure 2-6. Meteorological data from two stations (PFM006281 and PFM010700) was combined into unified time series by using data from PFM006281 where available (2013 to 2021) and using data from PFM010700 for filling the gaps in the time series (Figure 2-6). Data was available for the period 2003 to 2015 for the latter station. ET_p data was provided by SKB, the specific reference/ methodologies describing calculation of ET_p is done using the Penman (1948) equation following the recommendations from Eriksson (1981). This method of estimating ET_p has been used in previous hydrological modelling efforts (Bosson et al., 2008). Negative potential evaporation values, reflecting condensation, were adjusted to zero as negative values are ignored by HYPE. Streamflow data was available from the gauging station PFM002668, located circa 180 m downstream from the Lake Eckarfjärden outlet and was available for the time period 2004–2019. Time series for snow water equivalent were available from the measurement station AMF000072 for the period 2004–2019. Measurements are ongoing at several of stations included in the dataset used here, thus the time series could be prolonged in later applications if so required.

Finally, timeseries were extended backwards in time copying observed data from the period of 2004-01-01 to 2007-05-12 and was used to extend the spin-up period of the model for the calibration (see Section 2.3).

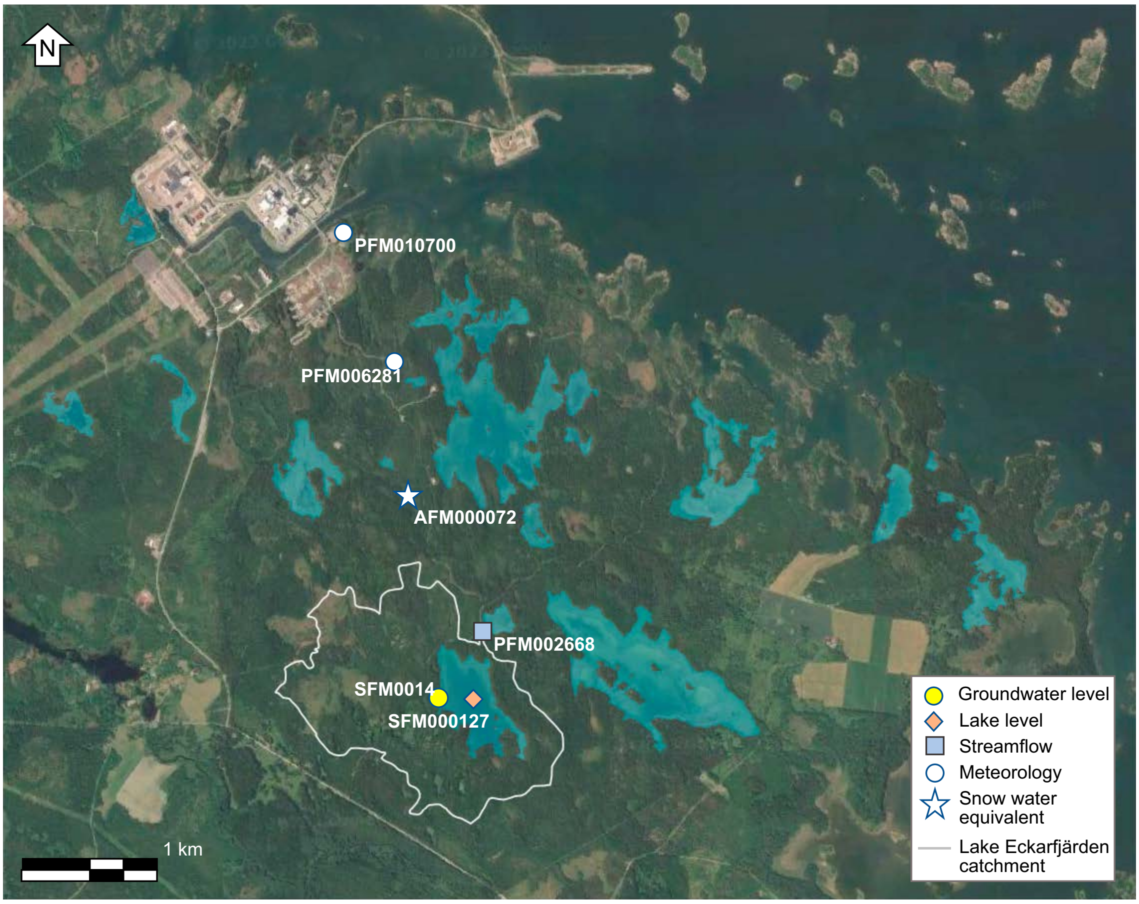

Figure 2-5. *Location of monitoring positions for the hydrological modeling of Lake Eckarfjärden. Lake Eckarfjärden watershed is shown with a light grey line. The background are satellite images from Google Maps.*

Table 2-3. Time series data used for the hydrological modelling of Lake Eckarfjärden.

Data type	Station name	Start	End	Reference
Streamflow (Q)	PFM002668	2004-12	2019-10	SKBdata-21-032 (1:1)
Precipitation (P)	PFM006281	2013-04	2021-03	SKBdata-21-032 (0:6)
	PFM010700	2003-05	2015-06	SKBdata-21-032 (0:9)
Potential evaporation (ET_p)	PFM006281	2013-04	2021-03	SKBdata-21-032 (0:8)
	PFM010700	2003-05	2015-06	SKBdata-21-032 (0:11)
Temperature (T)	PFM006281	2013-04	2021-03	SKBdata-21-032 (0:7)
	PFM010700	2003-05	2015-06	SKBdata-21-032 (0:10)
Lake stage	SFM000127	2011-03	2019-10	SKBdata-21-032 (1:2)
Snow depth	AFM000072	2004-01	2019-04	SKBdata-22-070 (0:1)
Groundwater, soil	SFM0014	2003-05	2020-09	SKBdata-21-034 (1:1)

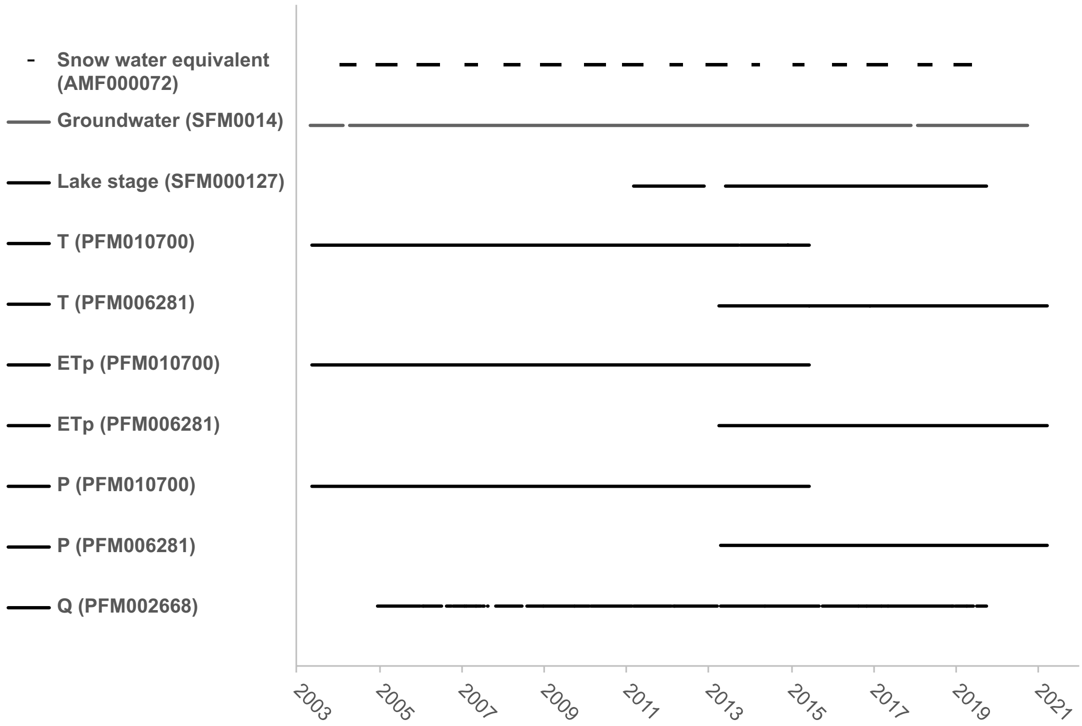

Figure 2-6. *Data availability for time series datasets used in the HYPE model of Eckarfjärden.*

2.2.2 Spatial data

Spatial data includes topography for delineating the catchment area as well as soil type, land use and vegetation data for defining SLC spatial coverage and parameter values (Table 2-4). Data on soil type, depth and extent was obtained from the SKB soil model (Table 2-4).

Table 2-4. Spatial data used for the hydrological modelling of Lake Eckarfjärden.

Data type	Resolution/scale	Reference
Soil	1:25000	Petrone and Sohlenius (2020)
Vegetation	1:15000	Löfgren (2010)
DEM	20 m	Petrone and Strömgren (2020)
Lakes	*	Brydsten and Strömgren (2005)

* Not known, data delivered as polygon.

2.3 Model calibration

The model uses a spin-up period between January 2000 and December 2004. A portion of this period is composed of synthetic data (Section 2.2.1). Since the HYPE model was set to initiate the model run with filled water storages, the spin-up period serves to drain the water storages to realistic levels.

Two model parameter calibrations were made and evaluated due to the short period of overlapping data for lake levels and streamflow (January 2014 through October 2019). The calibration and evaluation periods are shown in (Figure 2-7). Calibration period A covers 2005-01-01 through 2016-06-30 and calibration period B covers 2005-01-01 through 2014-01-01 and 2016-07-01 through 2018-12-31 (Figure 2-7). By having two separate calibrations and evaluations the amount of data on which model performance is evaluated on is extended, and therefore gives a better indication of the models robustness. The results from both calibrations and evaluations were assessed. However, since the results mostly showed the same strengths and weaknesses in the model implementation, only the results from calibration period A are shown in the results section (Section 3). The results for calibration period B is shown in Appendix D in order to make the model performance and conclusions drawn clearer to the reader. Areas where the two evaluations differ are reported in the results section (Section 3).

The calibration was automated using the Differential Evolution Markov Chain method (DE-MC) for optimizing 15 of the parameter types used in HYPE (see Appendix C parameter lists and values). The DE-MC combine the use of Differential evolution for genetic algorithms for global optimization and Markov Chain Monte Carlo to generate samples from the posterior distribution function.

The main idea of the DE-MC approach is to combine the Differential Evolution algorithm (DE), which is a global optimization algorithm over the parameter space, with a Markov Chain Monte Carlo (MCMC) technique to generate samples of parameter sets from the previous generation of parameter sets (Sherri et al., 2019). The optimization scheme was set to generate 400 populations using the DE and then use MCMC to perform 1 500 realizations of each population (Appendix C). This results in 600 000 realizations out of which 400 parameter sets are evaluated using the defined objective function (the best performing within each population). Here the objective function was weighted using the variables streamflow, lake water level and groundwater level (Table 2-5). Kling-Gupta efficiency (KGE) was used for evaluation of streamflow and lake water level. Nash-Sutcliffe efficiency (NSE) was used for evaluation of groundwater level. NSE was here used as an alternative to KGE since it rendered better groundwater representation in initial calibration attempts. Finally, the parameter set with the highest weighted performance score was then accepted as the resulting parameter setup for the model setup within this study (see Appendix C for parameter values).

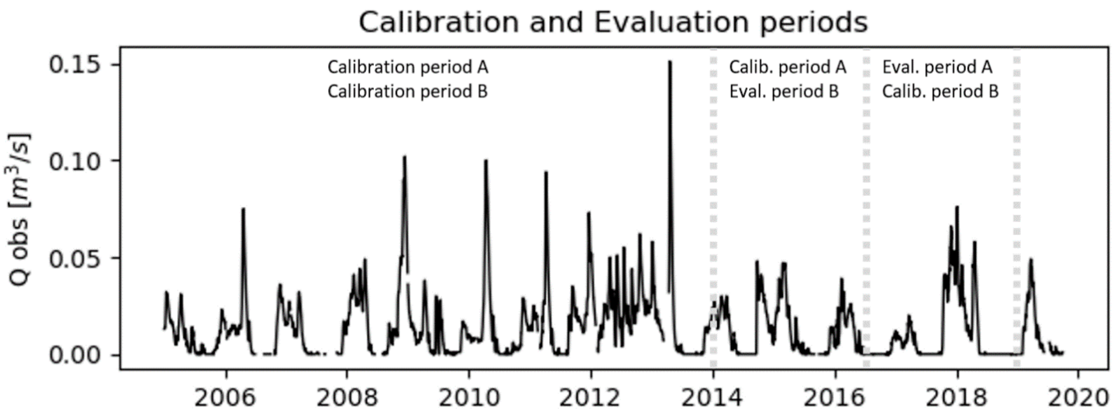

Figure 2-7. *Visualization of calibration and evaluation periods. Background shows observed streamflow downstream of Lake Eckarfjärden (PFM002668).*

In total, 92 parameters in HYPE were defined, out of which 66 parameters were automatically calibrated (Appendix C). Default values were used for the remaining 26 HYPE parameters. Parameter ranges for the automatic calibration were based on a combination of sensitivity analyses, analyzing parameter values against the objective function (Table 2-5) as well as recommended parameter ranges from SMHI (unpublished[4]).

Table 2-5. Objective function for automatic calibration of selected parameters for the Lake Eckarfjärden HYPE model.

Variable	Criterion	Weight
Streamflow	KGE*	0.5
Lake water level	KGE*	0.4
Groundwater level	NSE**	0.1

* Kling-Gupta efficiency.

** Nash-Sutcliffe efficiency.

2.4 Model evaluation

The evaluation period was constrained by the overlapping data period for the streamflow and lake stage, which covers January 2014 through October 2019. This means that the available time series for calibration and evaluation of model performance with regards to lake level is rather short. To increase the information on model performance, two calibrations and evaluations were made for different time periods. Together, evaluation period A and B cover the full years of lake level data (2014–2018). The selected time periods for the evaluations also allows for a break between evaluation period A and B during a time of zero flow (Figure 2-7, Section 2.3). Evaluation period A covers the time period of 2016-07-01 through 2018-12-31 and evaluation period B covers the time period of 2014-01-01 through 2016-06-30 (Figure 2-7). These evaluation periods were preceded by a spin-up period starting from 2005-01-01.

The evaluation period was evaluated using the same model set up and evaluation criteria as the calibration period (see Section 2.3). The results for evaluation period A are shown in the results section (Section 3) while the results for evaluation period B is shown in Appendix D.

[4] A list of unpublished recommended parameter ranges for some global HYPE parameters and some soil parameters were received at the SMHI HYPE in 2022. Available soil types were “coarse, medium, fine”.

3 Results

Daily simulations of streamflow showed a generally good agreement with observations for the evaluation period with regards to streamflow and lake stage. Groundwater level did not have a high performance score, though this is to be expected since the HYPE drainage basin was calibrated against the water level time series from a single groundwater well.

Simulated streamflow for both calibration periods had an average KGE of 0.85, simulated lake stage had an average KGE of 0.60 for the calibration periods and simulated groundwater had an average NSE of 0.27 (Table 3-1). The weighted performance (see Section 2.3) had an average of 0.69. Simulated streamflow for both evaluation periods had an average KGE of 0.74, simulated lake stage had an average KGE of 0.62 for the evaluation periods and simulated groundwater had an average NSE of 0.55 (Table 3-1). The weighted performance (see Section 2.3) had an average of 0.67.

Insights from calibration and evaluation period A mostly match that of calibration and evaluation period B. Therefore, detailed results from calibration and evaluation period A are shown below while the detailed results from calibration and evaluation period B are shown in Appendix D. Aspects of the results where the two calibration and evaluation periods differ are described in the following results section.

Table 3-1. Performance of the Lake Eckarfjärden HYPE model for the two calibration and evaluation periods.

Variable	Criteria	Calibration A	Calibration B	Evaluation A	Evaluation B	Weight
Streamflow	KGE	0.84	0.86	0.71	0.76	0.5
Lake stage	KGE	0.52	0.67	0.66	0.57	0.4
Groundwater level	NSE	0.31	0.23	0.65	0.45	0.1
Objective function	Weighted	0.65	0.73	0.68	0.65	-

3.1 Model performance, calibration period A

3.1.1 Streamflow

The simulated daily streamflow of the Lake Eckarfjärden drainage basin for calibration period A captures the main hydrological dynamics for calibration period. Simulated streamflow for the calibration period had a KGE value of 0.84. Looking at the time series, most of the observed streamflow peaks are captured in the simulation (Figure 3-1). A notable exception is during the end of 2015 where almost no simulated streamflow is generated despite clear peaks in the observed streamflow. The cumulative error of runoff for calibration period varies between approximately ±50 mm and no clear trend is seen in the cumulative error (Figure 3-1). A more detailed visual examination of the time series (not shown), supported by the cumulative error, indicate that there are timing issues with the onset and offset of simulated streamflow peaks. This can also be seen in the flow duration curve Figure 3-2 where zero flow constitutes a larger part of the flow duration curve compared to observed streamflow. A linear regression (Figure 3-2) of simulated and observed streamflow also shows that there is some hysteresis in the relationship between the simulated and observed streamflow, this is likely due to the models difficulties in timing the streamflow generation and due to streamflow being measured further downstream from the lake outlet.

A comparison of the flow duration curves for simulated and observed streamflow for calibration period (Figure 3-2) show that simulated streamflow performs particularly well for flows in the range of 0.02 – 0.105 m^3/s (mid flows to parts of the high flows). The simulated streamflow is not able to capture flows of > 0.105 m^3/s (the highest peaks) and tends to overestimate flows < 0.02 m^3/s (mid to low flows). As mentioned previously, zero flow is also more prevalent in the simulated time series. for simulated streamflow, only 72 % of the streamflow exceeded 0 m^3/s, while observed streamflow showed that 80 % of the streamflow exceeded 0 m^3/s (Figure 3-2).

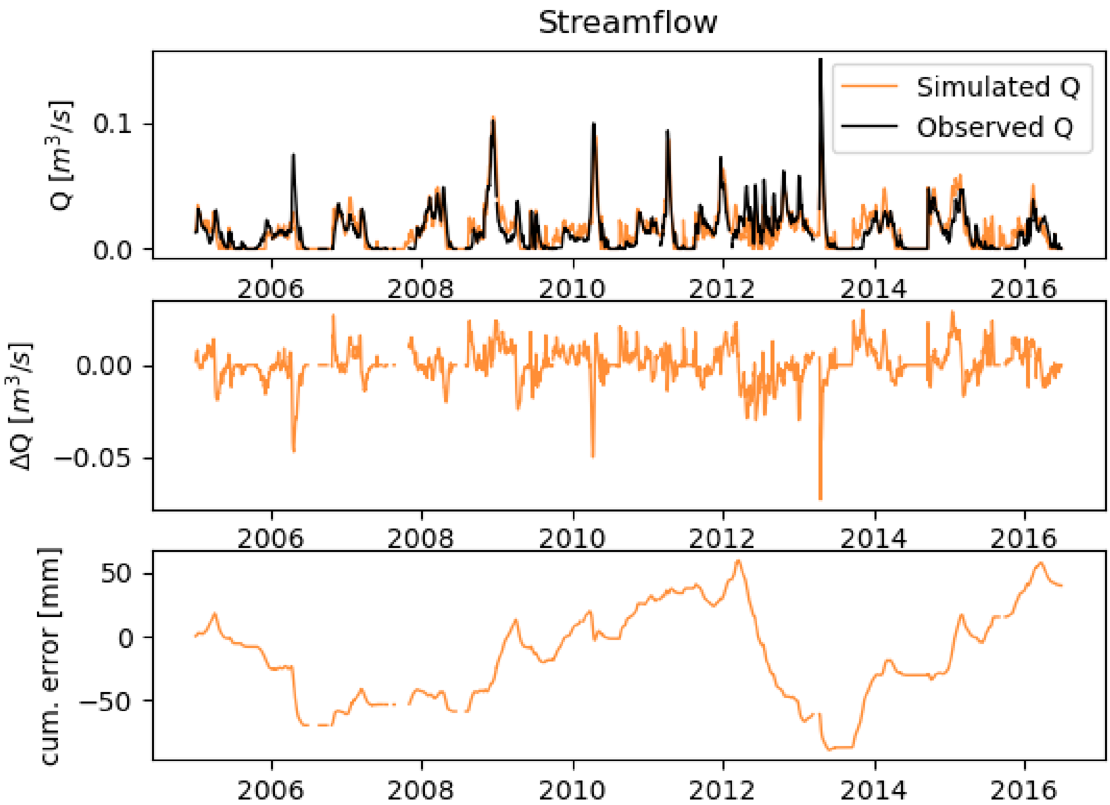

Figure 3-1. Model performance for calibration period A. Top: Simulated streamflow (orange line) and observed streamflow (black line). Mid: error in streamflow estimations (R observed – R simulated). Bottom: cumulative error of runoff [mm].

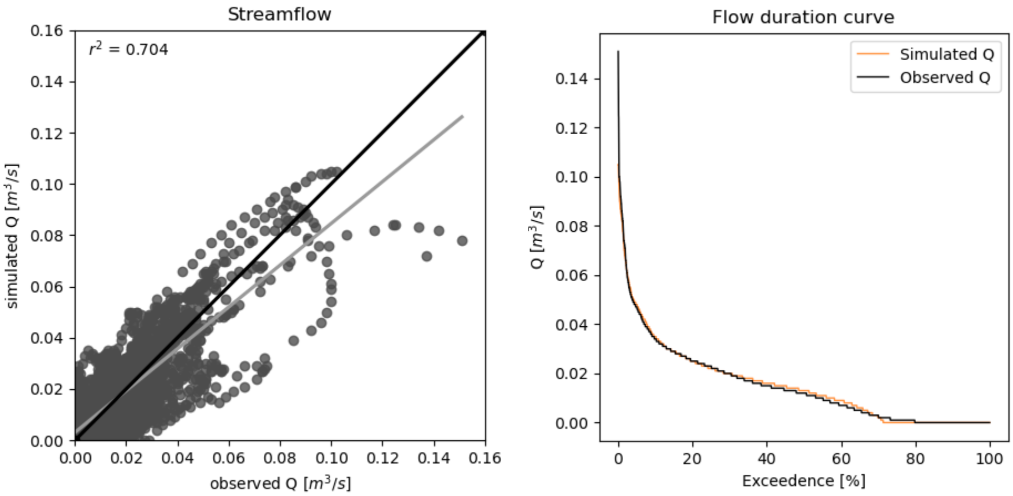

Figure 3-2. Left: linear regression of simulated and observed streamflow for calibration period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Flow duration curve of simulated (orange line) and observed (black line) streamflow for calibration period A.

Looking at the long-term average of monthly streamflow (Figure 3-3), the simulated streamflow tends to both under- and overestimate the streamflow depending on the month. The relative error for each specific month varies between +40 % and −50 % (Figure 3-3). In general, the simulated streamflow tends to overestimate streamflow during the autumn and winter months and underestimate during spring and summer months. The underestimation during the spring is most evident during April

(Figure 3-3) where the spring flood is clearly underestimated. An aggregation of all daily streamflow values for a particular month (Figure 3-4) indicates that the range of possible streamflow for each month is mostly well captured. Even so, the high flow events during April are clearly underestimated and the summer flow peaks during June and July tends to be overestimated as well.

The monthly average streamflow for calibration period B, however, manage to capture the spring flood to a larger extent and perform well at matching the other months of the year (Appendix D, Figure D-3). This is also reflected in the possible streamflow values for each month (Appendix D, Figure D-4).

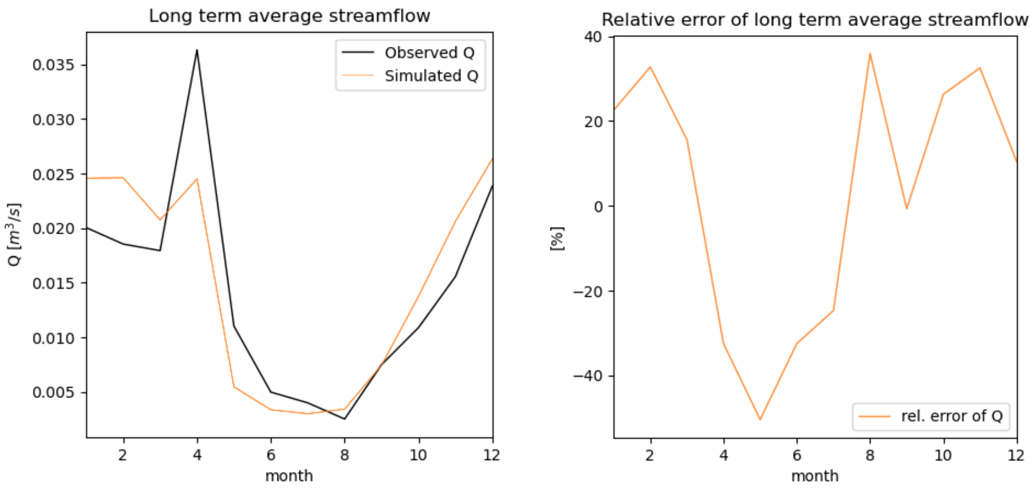

Figure 3-3. *Left: Monthly average streamflow for calibration period A. Both simulated (orange line) and observed (black line) streamflow is shown. Right: Relative error ((Q simulated – Q observed)/Q observed) of monthly average streamflow for calibration period A.*

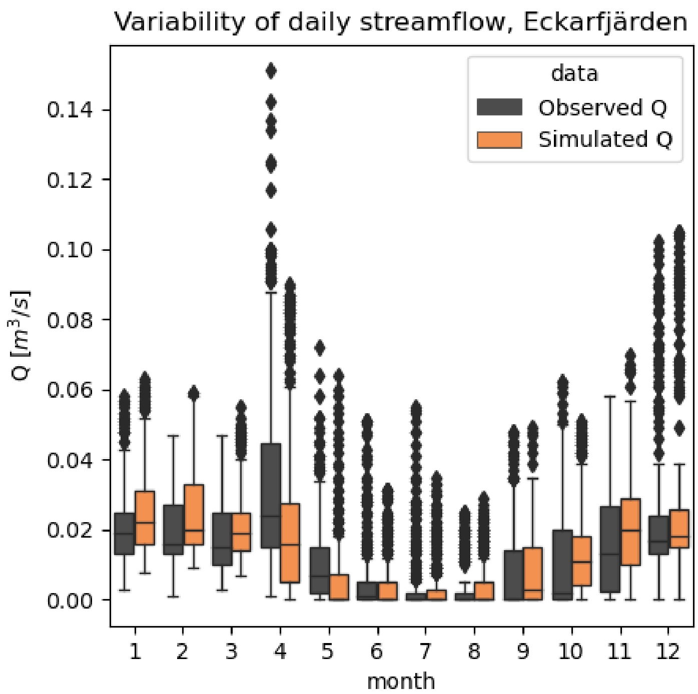

Figure 3-4. *Daily streamflow for the calibration period A, for each month of the year. Simulated streamflow is shown in orange and observed streamflow is shown in grey.*

The simulated annual streamflow has an r^2 value of 0.72 when compared to observed data for calibration period (Figure 3-5). Simulated streamflow tends to be overestimated in spring and early summer months while simulated streamflow tends to be underestimated the rest of the year. The relative error ranges between −50 % to +36 %. A potential explanation for this pattern is the non-stationary relationship between Lake Eckarfjärden lake stage and streamflow at the measured gauge (see Section 2.1.2).

3.1.2 Lake level

Observed and simulated time series data for calibration period A indicate that the simulated lake stage represents the observed lake stage adequately (Figure 3-6) with an KGE value of 0.52. The main period of difference between simulated and observed lake stage is during the end of 2015 where lake stage is underestimated. The difference between observed and simulated lake stage ranges between −0.10 m and +0.11 m (Figure 3-6). The linear regression also indicates adequate representation of the simulated lake stage with an r^2 value of 0.67 (Figure 3-7).

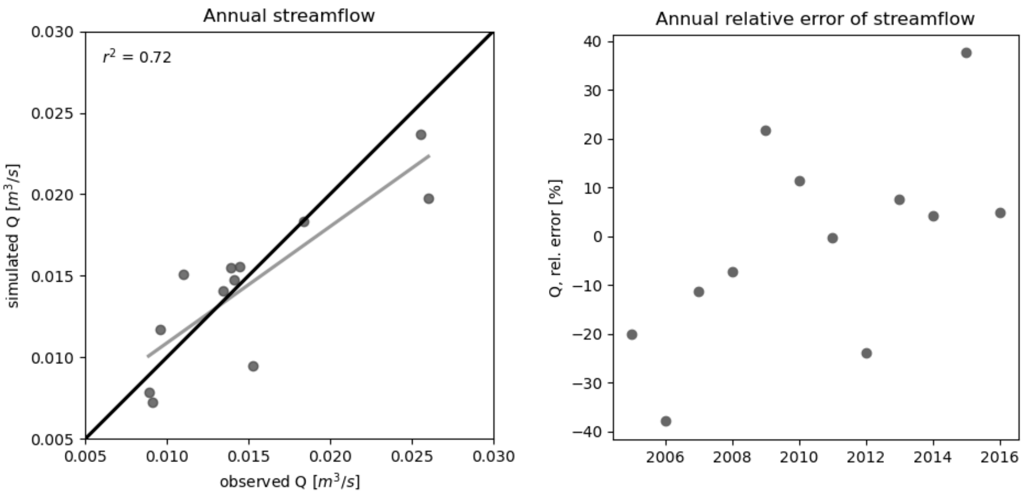

Figure 3-5. *Left: linear regression of simulated and observed annual streamflow for calibration period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Relative error [%] of simulated and observed annual streamflow for the calibration period A.*

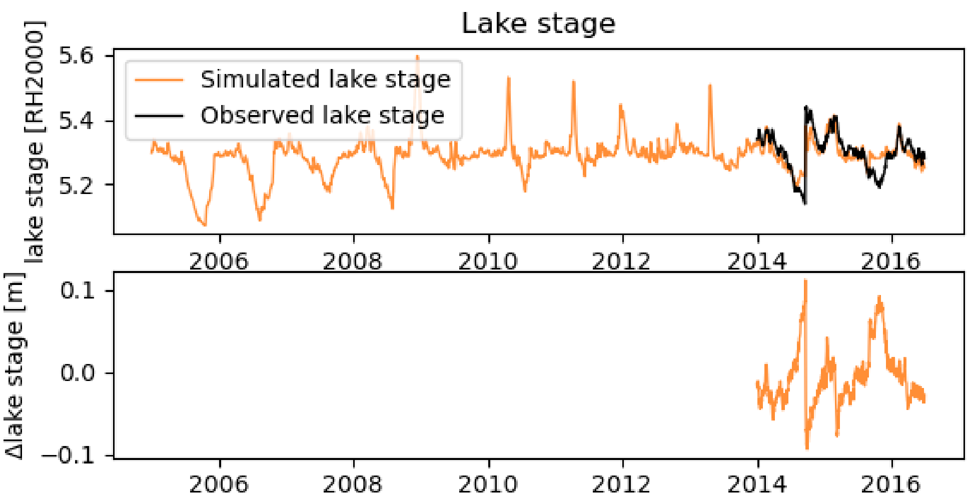

Figure 3-6. *Model performance for calibration period A. Top: Simulated lake stage of Eckarfjärden (orange line) and observed lake stage of Eckarfjärden (black line). Bottom: error in estimated lake stage (observed – simulated).*

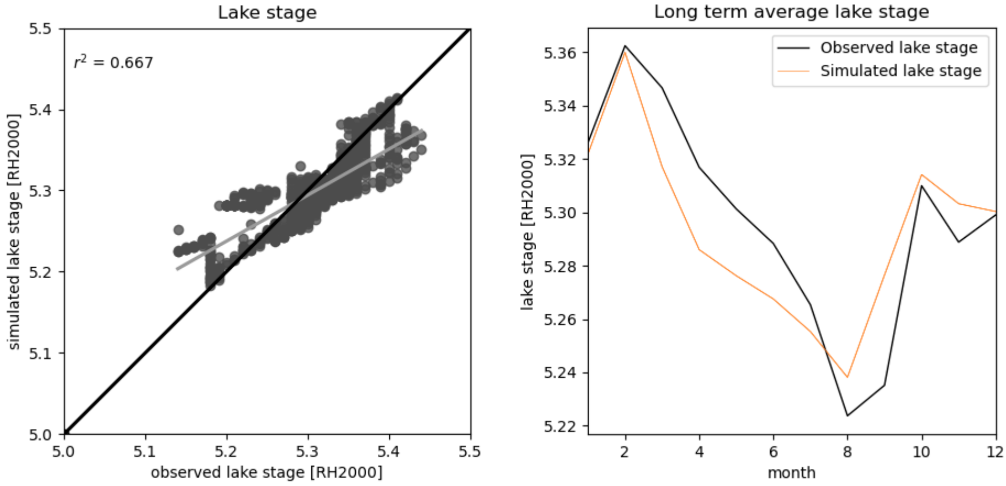

Figure 3-7. *Left: linear regression of simulated and observed lake stage for Eckarfjärden for calibration period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Lake Eckarfjärden monthly average lake stage for calibration period A. Both simulated (orange line) and observed (black line) lake stage are shown.*

The long-term monthly averages of lake stage show the simulated lake stage for calibration period A captures the main seasonal dynamic of Lake Eckarfjärden. There is a tendency to underestimate the lake stage during the spring and summer months and an overestimation of the lake stage during the autumn and winter months. An aggregation of all lake stage values for a particular month (Figure 3-8) indicate that the range of possible lake stages for each month tends to underestimate both the lake level and that possible ranges during the spring months while underestimating the possible ranges during the winter months (Figure 3-8).

The results from calibration period B show that the simulated lake levels instead are overestimated for all months except March through June (Appendix D, Figure D-7). This may be a result of overfitting to the short period of lake level data that was available for calibration (2.5 years). For better stability in model performance longer overlapping time periods for lake level and streamflow may be required.

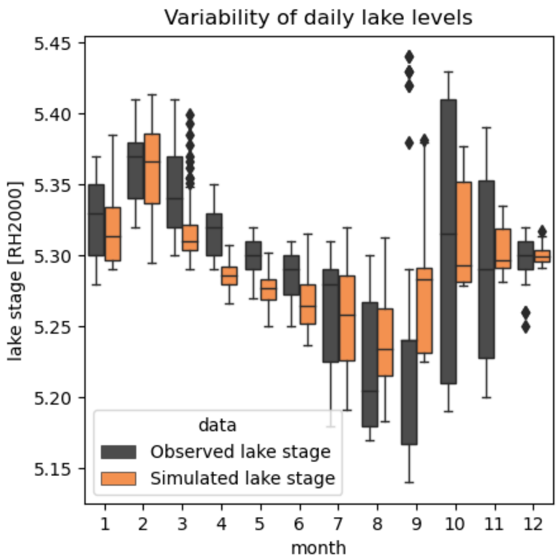

Figure 3-8. *Eckarfjärden daily lake stage for calibration period A, for each month of the year. Simulated lake stage is shown in orange and observed lake stage is shown in grey.*

The influx of water to Lake Eckarfjärden can be divided into its constituents (Figure 3-9). According to the HYPE model realization used in this study, 30–50 % is generated from the upper two soil layers (via "local stream") depending on the month, and the third soil layer is not activated at all in the surface runoff generation. The lack of runoff generation from the third soil layer is related to the model parameter "stream depth", which determines lowest possible drainage level and here is set to a level above the third soil layer for most SLCs.

The saturated overland flow, which occurs when the groundwater level reach above ground varies between 15–40 % during the months of October through April, and 0–2 % during the months of May through September. Overland flow due to excess infiltration constitutes up to 6 % of the water influx to lake Eckarfjärden according to the HYPE model. Finally, the direct precipitation to Lake Eckarfjärden constitutes 10 – 60 % of the water influx depending on the month. The direct precipitation is also likely overestimated since HYPE does not allow for a dynamic lake surface area.

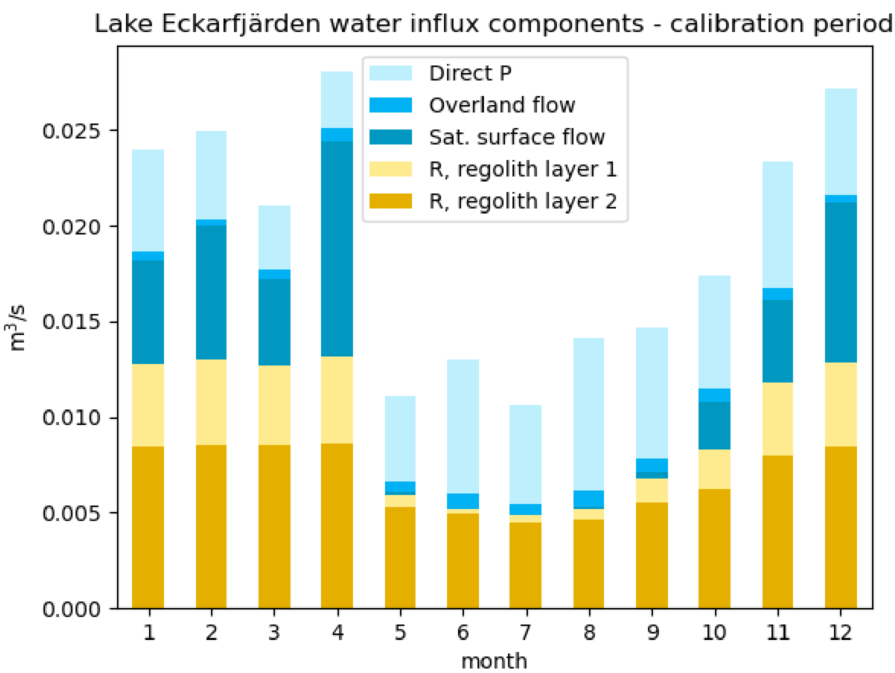

Figure 3-9. *Water influxes to Lake Eckarfjärden for calibration period A. "Overland flow" is infiltration excess* of *rainfall and snowmelt, "Sat. surface flow" occurs when the groundwater level is higher than the ground level, "R, soil layer 1" and "2" are groundwater contributions from the upper two soil layers. "Soil layer 3" does not contribute to runoff in this model realization.*

3.1.3 Groundwater

Time series of groundwater level below ground surface (hence forth "groundwater levels") shows that the general seasonality of observed groundwater is captured by the simulated groundwater for calibration period (Figure 3-10). The observed groundwater is generally closer to the ground surface than the simulated groundwater stage. This difference is, however, not surprising considering that the simulated groundwater represents an average for the whole drainage basin while the observed groundwater is measured at a point. The true groundwater level for the whole drainage basin is unknown. Even so, the general dynamics of point measurement can be used as a rudimentary assessment of the performance of the simulated groundwater. The NSE value, which was used in the calibration scheme, for the groundwater levels was 0.31. The performance score of the simulated groundwater is low and likely penalized by the difference in mean groundwater level as well as timing issues. The linear regression of groundwater levels shows a r^2 value of 0.63 (Figure 3-11).

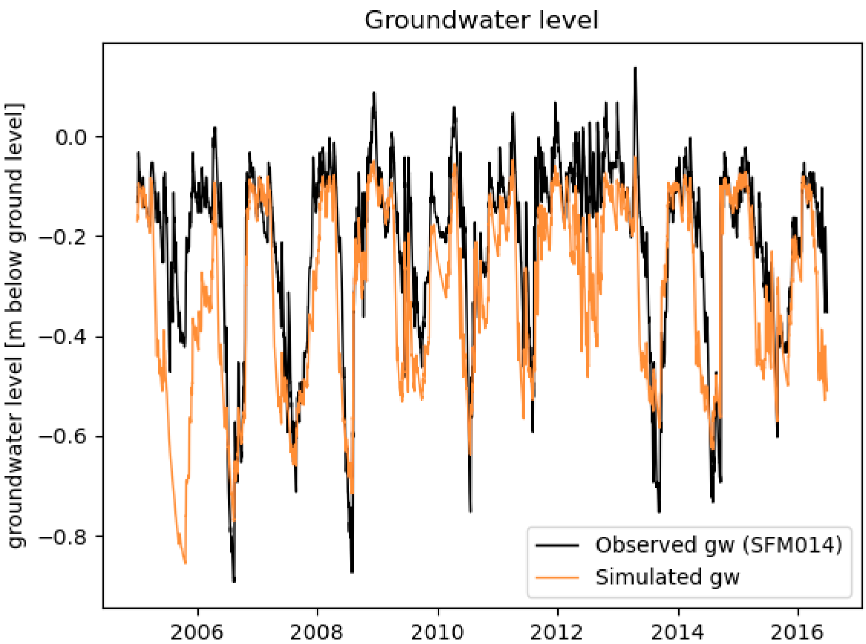

Figure 3-10. *Simulated groundwater level (orange line) and observed groundwater level (black line) for calibration period A.*

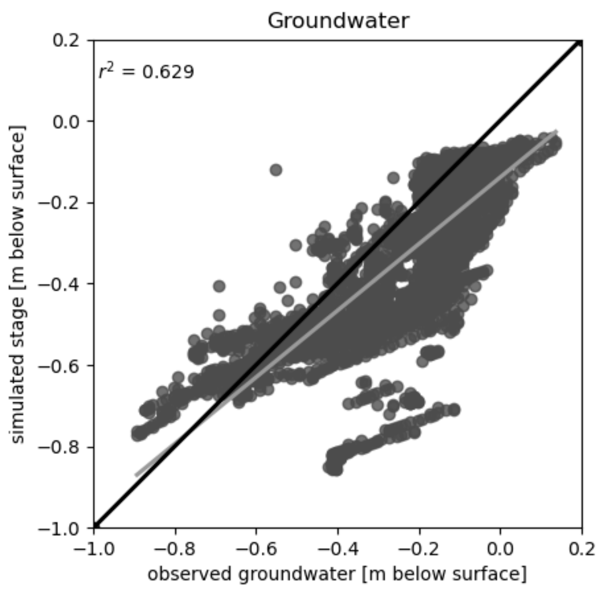

Figure 3-11. *Linear regression of simulated and observed groundwater level for calibration period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

3.1.4 Snow

Snow depth and snow water equivalents were used as auxiliary data as quality control for the snow melt parameters (Figure 3-12). A general comparison of both snow depth and snow water equivalents time series showed that there is a good agreement between simulated and observed snow. This is supported by the linear regression of simulated and observed snow depth, with an r^2 value of 0.91 (Figure 3-13).

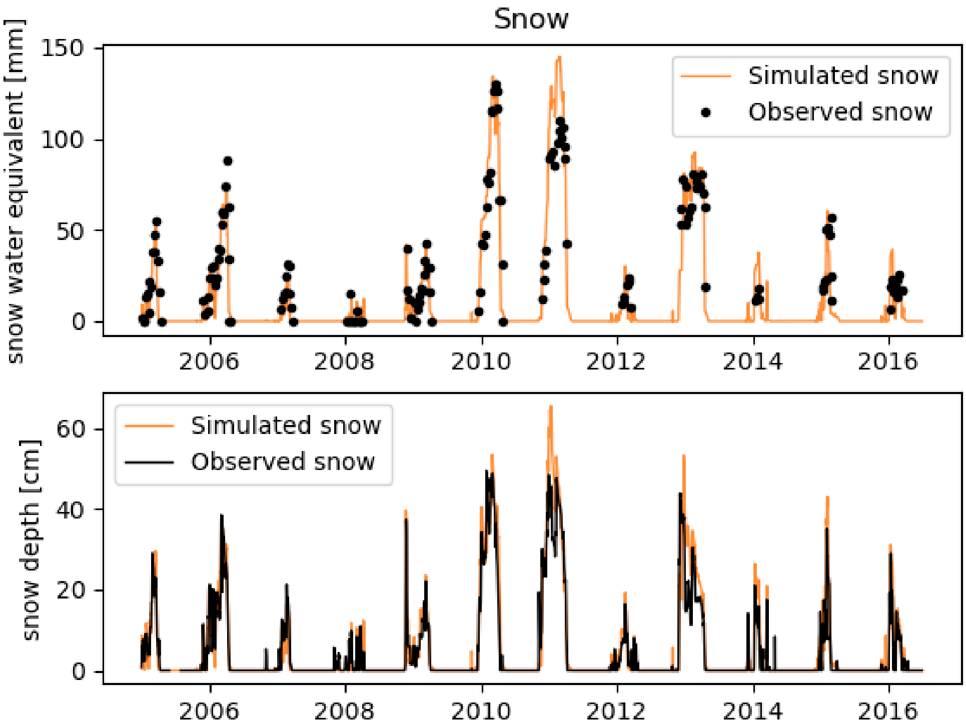

Figure 3-12. Snow time series data. Top: Simulated snow water equivalent (orange line) and observed snow water equivalent (black line) for calibration period A. Bottom: Simulated snow depth (orange line) and observed snow depth (black line) for calibration period A.

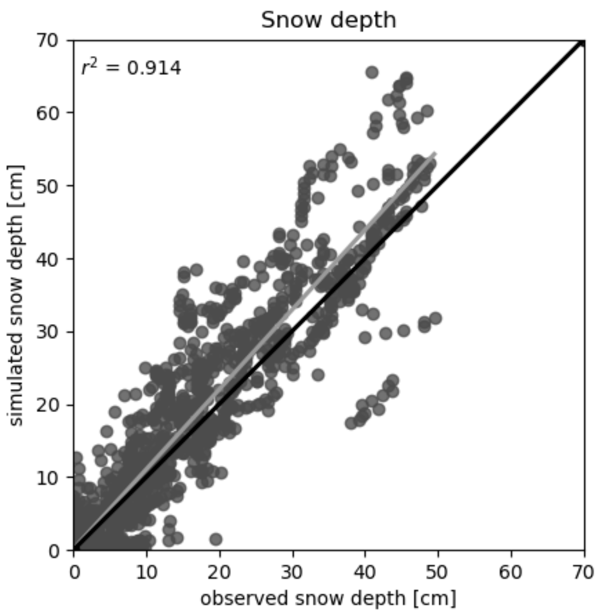

Figure 3-13. Linear regression of simulated and observed snow depth for calibration period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.

3.2 Model performance, evaluation period A

Daily simulations of streamflow showed a generally good agreement with observations for the evaluation period with regards to streamflow and lake stage where the main dynamics of the streamflow is captured (Table 3-1). Groundwater levels also had a generally good agreement with observed ground water levels (Table 3-1). Simulated streamflow for both evaluation periods had an average KGE of 0.74, simulated lake stage had an average KGE of 0.62 for the evaluation periods and simulated groundwater had an average NSE of 0.55 (Table 3-1). The weighted performance (see Section 2.3) had an average of 0.67.

3.2.1 Streamflow

The simulated daily streamflow of the Lake Eckarfjärden drainage basin for evaluation period A captures the main hydrological dynamics for evaluation period. Simulated streamflow for the evaluation period had a KGE value of 0.71. Looking at the time series, observed streamflow peaks are captured in the simulation albeit delayed (Figure 3-14). The cumulative error of runoff for evaluation period had a positive trend with a cumulative error of approximately 100 mm at the end of the evaluation period (Figure 3-14). The flow duration curve also shows a general overestimation of simulated streamflow for almost all flow conditions.

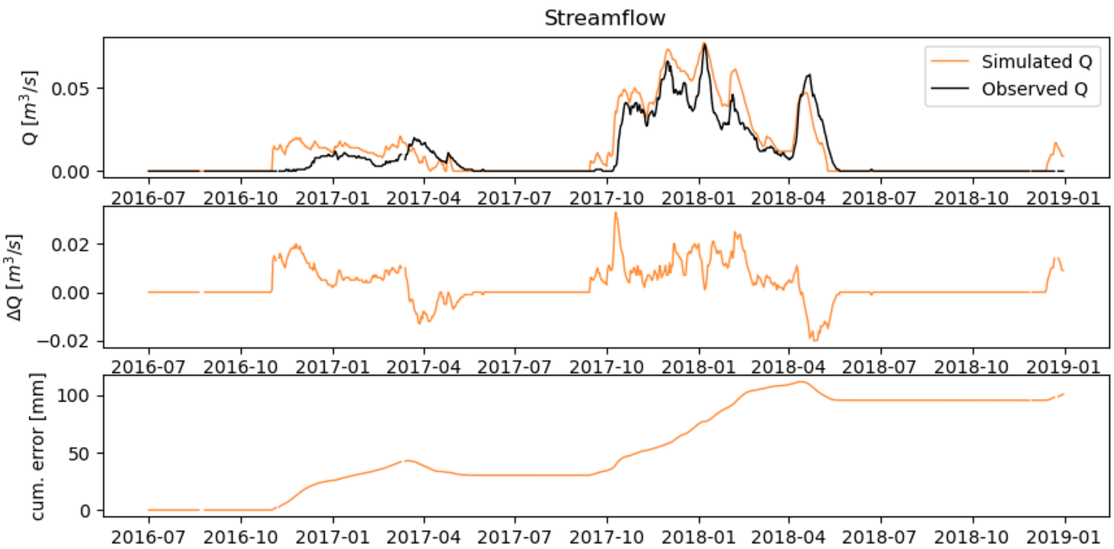

Figure 3-14. *Model performance for evaluation period A. Top: Simulated streamflow (orange line) and observed streamflow (black line). Mid: error in streamflow estimations (R observed – R simulated). Bottom: cumulative error of runoff [mm].*

A linear regression (Figure 3-15) of simulated and observed streamflow also show that there is some hysteresis in the relationship between the two. This is likely due to the model's difficulties in timing the streamflow generation and due to streamflow being measured further downstream from the lake outlet. The r^2 value for the linear regression was 0.87.

Looking at the long-term average of monthly streamflow during the evaluation period (Figure 3-16), the simulated streamflow tends to both under- and overestimate the streamflow depending on the month. The relative error for each specific month with flow varies between + 60 % and −70 % (Figure 3-16). In general, the simulated streamflow tends to overestimate streamflow during the autumn and winter months and underestimate during spring months. An aggregation of all daily streamflow values for a particular month (Figure 3-17) indicate that the range of possible streamflow for each month tends to have a larger span and higher flows compared to observed streamflow. Even so, the main seasonal dynamics are captured in the simulated streamflow.

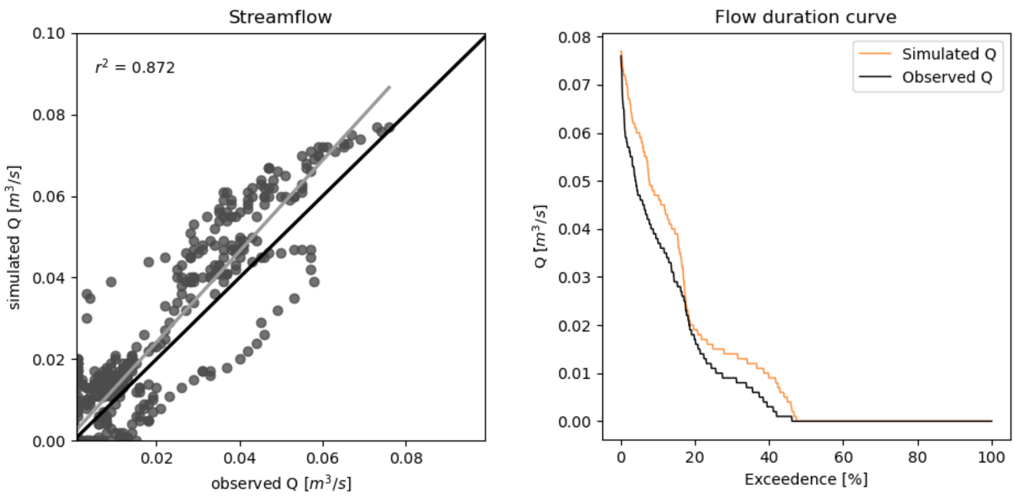

Figure 3-15. *Left: linear regression of simulated and observed streamflow for evaluation period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Flow duration curve of simulated (orange line) and observed (black line) streamflow for evaluation period A.*

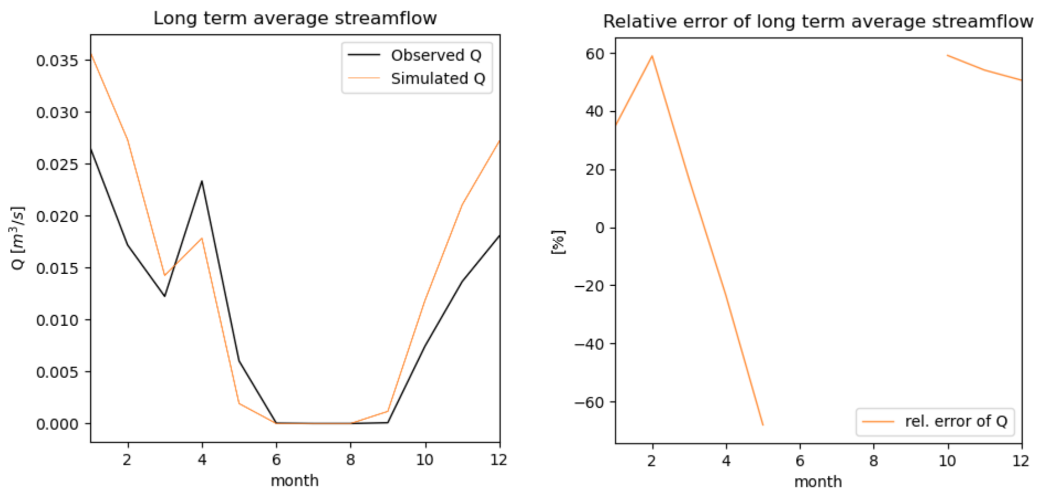

Figure 3-16. *Left: Monthly average streamflow for evaluation period A. Both simulated (orange line) and observed (black line) streamflow is shown. Right: Relative error ((Q simulated – Q observed)/Q observed) of monthly average streamflow for months with streamflow during evaluation period A.*

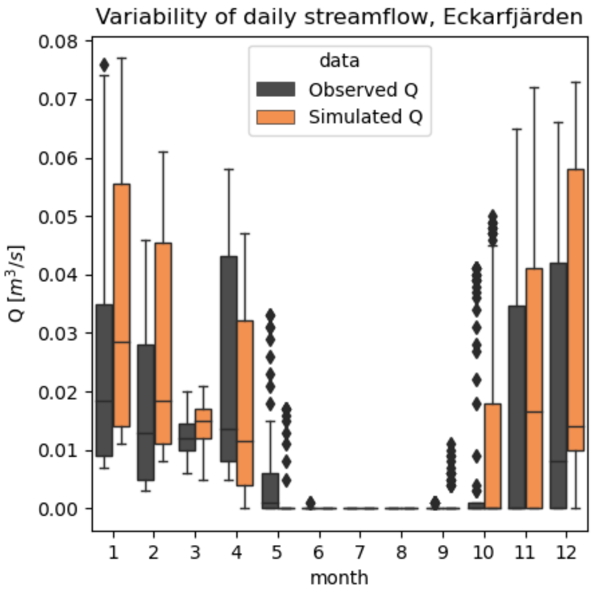

Figure 3-17. *Daily streamflow for the evaluation period A, for each month of the year. Simulated streamflow is shown in orange and observed streamflow is shown in grey.*

3.2.2 Lake level

Simulated lake levels for the evaluation period (Figure 3-18) has an KGE value of 0.66 when compared to observed lake levels. The main periods of difference between simulated and observed lake stage are during the autumn of 2016 and 2018 where lake stage is overestimated. The difference between observed and simulated lake stage ranges between −0.08 m and +0.20 m (Figure 3-18). The linear regression also indicates adequate representation of the simulated lake stage with an r^2 value of 0.75 (Figure 3-19).

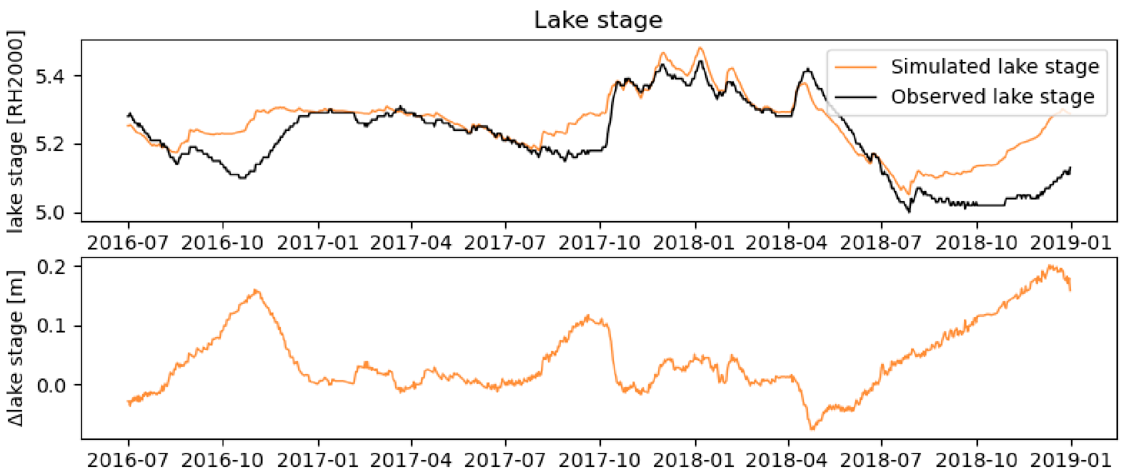

Figure 3-18. *Model performance for evaluation period A. Top: Simulated lake stage of Eckarfjärden (orange line) and observed lake stage of Eckarfjärden (black line). Bottom: error in estimated lake stage (observed – simulated).*

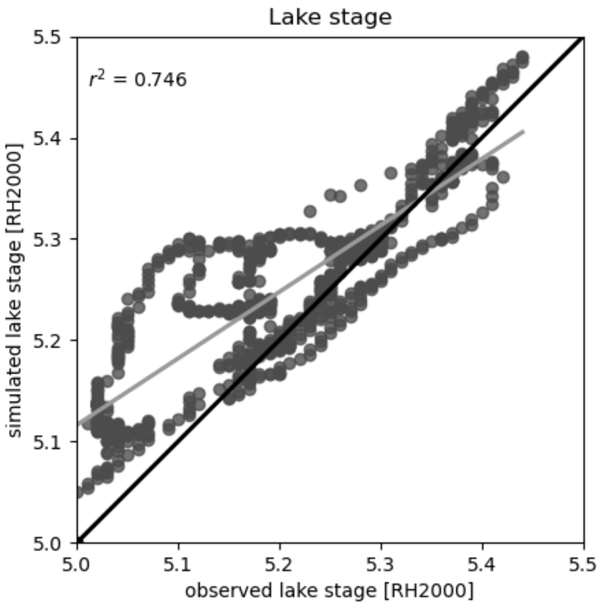

***Figure 3-19.** Linear regression of simulated and observed lake stage for Eckarfjärden for evaluation period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

The long-term monthly averages of lake stage show the simulated lake stage for evaluation period A captures the main seasonal dynamic of Lake Eckarfjärden though overestimates lake levels during the months of July through December. An aggregation of all lake stage values for a particular month (Figure 3-20) indicates that the distribution magnitude of simulated lake levels for each month is similar to observed lake levels, albeit elevated during July through December (Figure 3-20).

In contrast to the calibration period, monthly averages for lake levels are better represented in evaluation period B compared to evaluation period A (Appendix D, Figure D-20). In fact, the monthly averages for lake stage, resulting from evaluation period A, are similar to that of calibration period B. This indicates that the difference in performance in simulation long term monthly averages of lake stage between the two calibration and evaluation periods are mostly due to interannual variability in forcing and calibration variables rather than model structure.

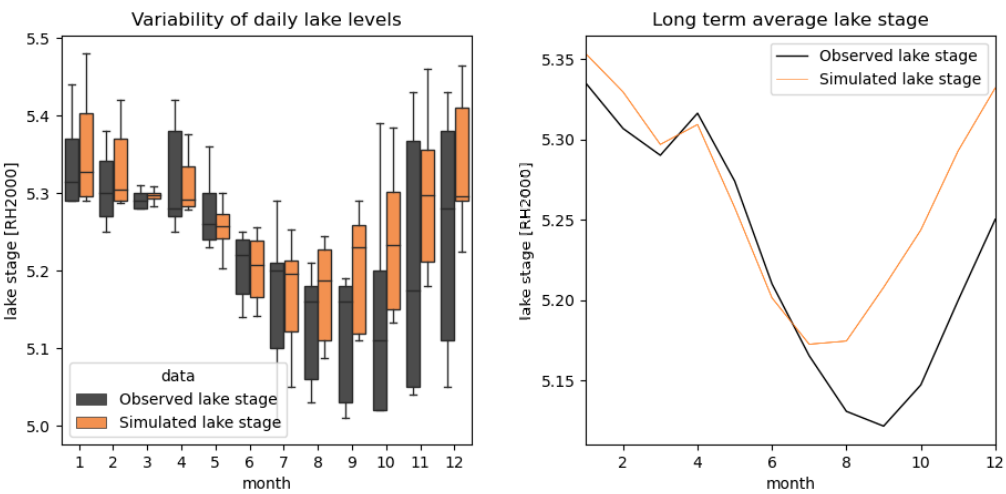

***Figure 3-20.** Left: Eckarfjärden daily lake stage for evaluation period A, for each month of the year. Simulated lake stage is shown in orange and observed lake stage is shown in grey. Right: Lake Eckarfjärden monthly average lake stage for evaluation period A. Both simulated (orange line) and observed (black line) lake stage are shown.*

The influx of water to Lake Eckarfjärden can be divided into its constituents (Figure 3-21). The lake influxes for the evaluation period are nearly identical to the fluxes of the calibration period. According to the HYPE model realization used in this study, 45–60 % of the water entering the lake is generated from the upper two soil layers (via "local stream") depending on the month and the third soil layer is not activated at all in the surface runoff generation. The lack of runoff generation from the third soil layer is related to the model parameter "stream depth", which determines lowest possible drainage level and here is set to a level above the third soil layer for most SLCs.

The saturated overland flow, which occurs when the groundwater level reach above ground varies between 15–40 % during the months of October through April, and up to 2 % during the months of May through September. The saturated overland flow during October through April is likely highly overestimated. Overland flow due to excess infiltration constitutes up to 6 % of the water influx to lake Eckarfjärden according to the HYPE model. Finally, the direct precipitation to Lake Eckarfjärden constitutes 10 – 60 % of the water influx depending on the month. The direct precipitation is also likely overestimated since HYPE does not allow for a dynamic lake surface area.

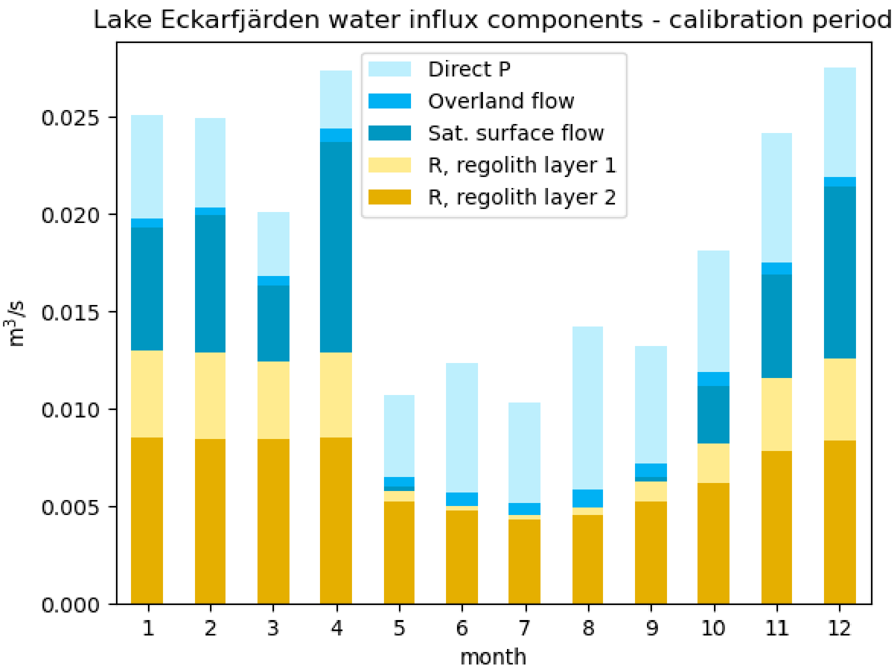

Figure 3-21. *Water influxes to Lake Eckarfjärden for evaluation period A. "Overland flow" is excess precipitation and snowmelt, "Sat. surface flow" occurs when the groundwater level is higher than the ground level, "R, soil layer 1" and "2" are groundwater contributions from the upper two soil layers. "Soil layer 3" does not contribute to runoff in this model realization.*

3.2.3 Groundwater

Time series of groundwater level below ground surface (hence forth "groundwater levels") shows that the general seasonality of observed groundwater is captured by the simulated groundwater for evaluation period (Figure 3-22). The main difference is the recession period during spring and early summer months. The recession curve for the observed time series shows a convex shape while the simulated groundwater levels show a concave shape. One possible explanation for the difference is that observed groundwater levels are measured at a point near Lake Eckarfjärden, which means that it is fed groundwater from upstream areas. Another possible explanation is that the groundwater level at the measured point is dependent on the lake level, considering the proximity of the groundwater well to Lake Eckarfjärden. The NSE value, which was used in the evaluation scheme, for the groundwater levels was 0.65. The linear regression of groundwater levels shows a r^2 value of 0.63 (Figure 3-23).

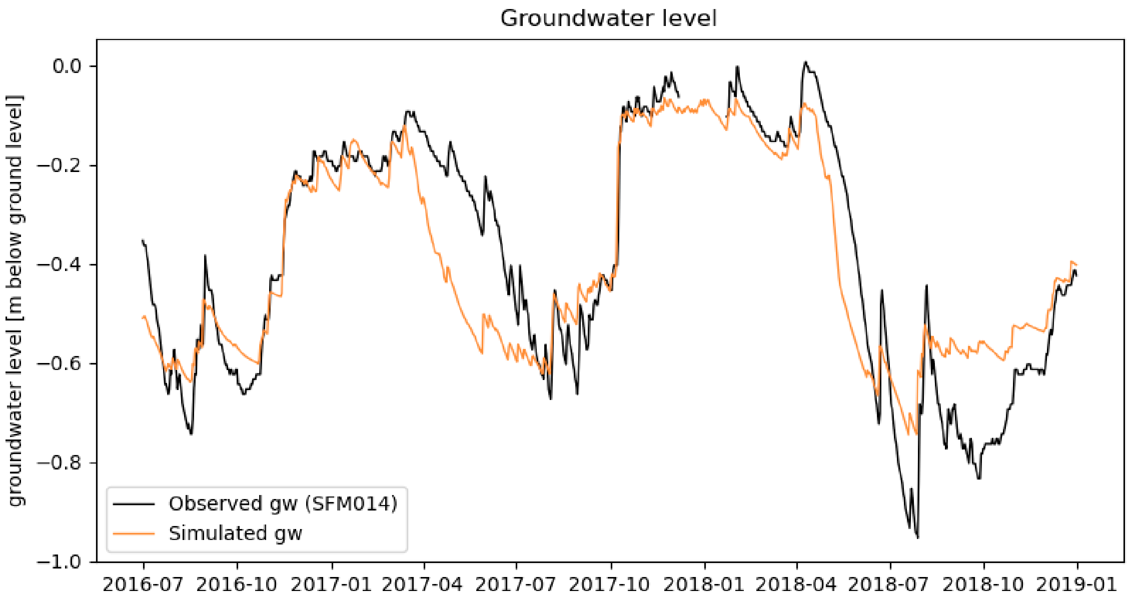

Figure 3-22. *Simulated groundwater level (orange line) and observed groundwater level (black line) for evaluation period A.*

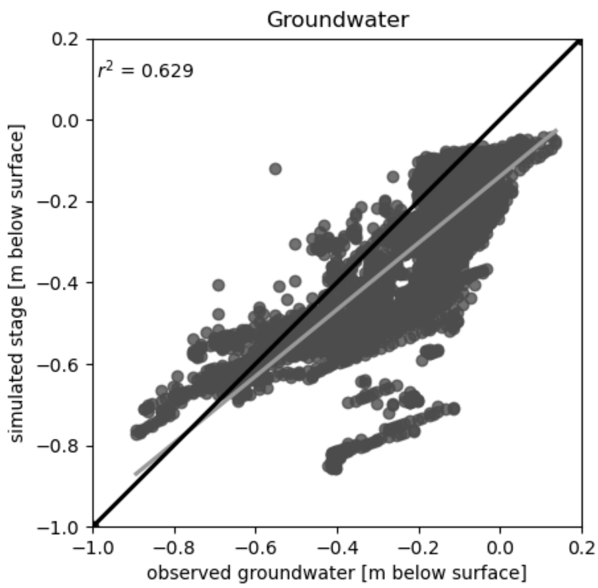

Figure 3-23. *Linear regression of simulated and observed groundwater level for evaluation period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

3.2.4 Snow

Snow depth and snow water equivalents were used as auxiliary data as quality control for the snow melt parameters (Figure 3-24). A general comparison of snow water equivalent time series showed that there is a good agreement between simulated and observed snow for evaluation period A. This is supported by the linear regression of simulated and observed snow depth, with an r^2 value of 0.90 (Figure 3-25).

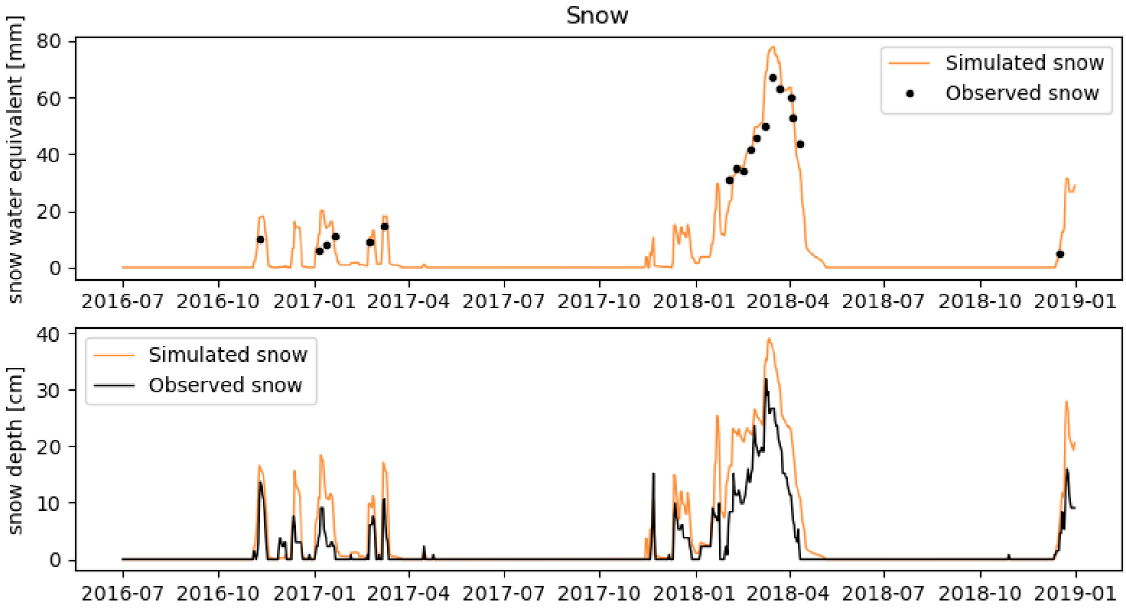

Figure 3-24. *Snow time series data. Top: Simulated snow water equivalent (orange line) and observed snow water equivalent (black line) for evaluation period A. Bottom: Simulated snow depth (orange line) and observed snow depth (black line) for evaluation period A.*

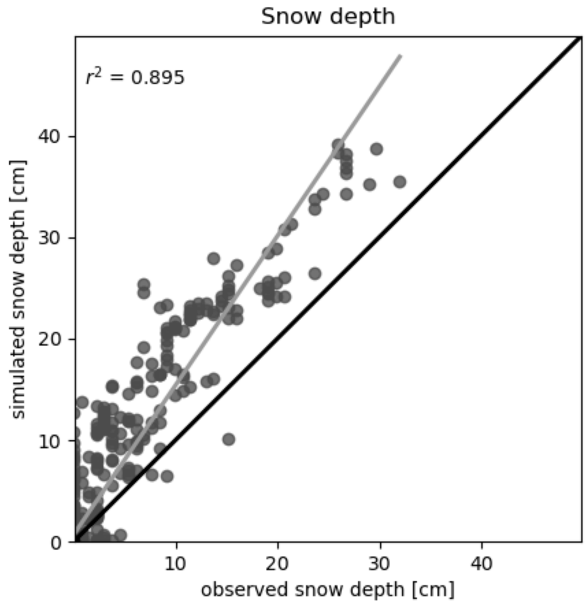

Figure 3-25. *Linear regression of simulated and observed snow depth for evaluation period A. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

4 Suggestions for future improvements

One key factor for improving hydrological model performance is the availability of extensive datasets. In the case at hand, the model could benefit from an extension of the streamflow data beyond its current limit of 2019-10-01. By the inclusion of more recent streamflow data, a better overlap can be achieved between the streamflow and lake data time series. This extended dataset would enable more comprehensive model calibration and validation, leading to enhanced accuracy in capturing hydrological dynamics of Lake Eckarfjärden drainage basin.

For accurate modeling, it is also essential to have well-constrained model parameters that accurately represent the specific characteristics of different land cover types. The Lake Eckarfjärden drainage basin is dominated by forested moraine, which means that current HYPE implementation has parameters that are well-constrained for this land cover. However, for less dominant land cover types such as wetlands, the current parameter ranges are less constrained through the calibration. To address this, it is recommended to identify other drainage basins dominated by wetlands and set up HYPE models for those drainage basins. This approach would provide valuable data to refine the model parameters specific to wetland-dominated drainage basins and improve the model's representation of wetland-related processes.

Another area of improvement is to ensure comparability with previous hydrological modeling efforts in the Forsmark area through the inclusion of all meteorological stations. Previous modelling efforts have used an average of all meteorological station as model forcing, while the present study used nearest available meteorological station.

Finally, the HYPE model has the capability to track the fluxes of the main constituents in lake water influx. These constituent fluxes can subsequently be used as inputs for other models. However, the contribution of each constituent is sensitive to the model setup. Factors such as the thickness of soil boxes as well as stream depth (drainage level) impacts constituent contributions to the outlet lake. While incorporating groundwater levels as a calibration variable helps to constrain the model in this aspect, conducting an end-member analysis to determine lake fluxes would further constrain the HYPE model with regards to lake water fluxes.

5 Conclusions

Based on the analysis conducted in this study, several key conclusions can be drawn regarding the investigation of HYPE’s capability to simulate hydrological fluxes and the possibility of utilizing its simplified model structure for simulating hydrological fluxes to biosphere objects at Forsmark.

Firstly, the results demonstrate that HYPE is an effective tool for modeling the hydrological fluxes of Lake Eckarfjärden. The model was able to, to a satisfactory extent, simulate runoff as well as lake and groundwater levels. The agreement between the model simulations and observed data indicates that HYPE is capable of capturing the main dynamics of the hydrological system within the studied drainage basin.

Furthermore, the study revealed that HYPE’s representation of key hydrological processes, such as infiltration, snow accumulation and melt, and routing of water through the drainage basin, is consistent with the observed patterns. This suggests that HYPE incorporates appropriate parameterizations and algorithms to adequately simulate these processes, leading to reliable estimates of hydrological fluxes.

Secondly, in terms of investigating the possibility of using the simplified model structure of HYPE to simulate hydrological fluxes to biosphere objects at Forsmark, the findings indicate promising potential. The simplified model structure of HYPE allows for efficient and computationally inexpensive simulations while still capturing the essential hydrological dynamics.

Overall, this study demonstrates the effectiveness of HYPE in simulating hydrological fluxes and highlights the potential of utilizing its simplified model structure for simulating hydrological processes to biosphere objects at Forsmark. The model provides reliable estimates of various hydrological fluxes and can serve as a valuable tool for hydrological studies.

For further improved results, time series of observed variables should be extended in order to allow for a longer overlap of streamflow and lake levels. It should also be noted that forested moraine is the dominant Soil type Land use Combination (SLC) by a large margin, which means that the calibrated parameters are less robust for other SLCs. Parameterization of a drainage basin dominated by other SLCs, such as wetlands, could improve the parameter robustness for these SLCs at Forsmark.

References

SKB's (Svensk Kärnbränslehantering AB) publications can be found at www.skb.com/publications.

Arheimer B, Pimentel R, Isberg K, Crochemore L, Andersson J C M, Hasan A, Pineda L, 2020. Global catchment modelling using World-Wide HYPE (WWH), open data, and stepwise parameter estimation. Hydrology and Earth System Sciences, 24(2), 535–559. DOI: 10.5194/hess-24-535-2020

Bosson E, Gustafsson L-G, Sassner M, 2008. Numerical modelling of surface hydrology and near-surface hydrogeology at Forsmark. Site descriptive modelling, SDM-Site Forsmark. SKB R-08-09, Svensk Kärnbränslehantering AB.

Brydsten L, Strömgren M, 2005. Measurements of brook gradients and lake thresholds. Forsmark site investigation. SKB P-04-141. Svensk Kärnbränslehantering AB.

Eriksson B, 1981. Den potentiella evapotranspirationen i Sverige. SMHI Rapport RMK 28. (In Swedish.). DiVA, id: diva2:1462578

Lindström G, Pers C, Rosberg J, Strömqvist J, Arheimer B, 2010. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. Hydrology Research 41.3–4, 295-319. DOI: 10.2166/nh.2010.007

Löfgren A (ed), 2010. The terrestrial ecosystems at Forsmark and Laxemar-Simpevarp. SR-Site Biosphere. SKB TR-10-01. Svensk Kärnbränslehantering AB

Penman H L, 1948. Natural evaporation from open water, bare soil and grass. PRSL, Ser. A 193-120-146. DOI: 10.1098/rspa.1948.0037

Petrone J, Sohlenius G, 2020. Baseline Forsmark – Depth and stratigraphy of regolith. SKB R-17-07. Svensk Kärnbränslehantering AB.

Petrone J, Strömgren M, 2020. Baseline Forsmark – Digital elevation model. SKB R-17-06. Svensk Kärnbränslehantering AB.

Sherri M, Boulkaibet I, Marwala T, Friswell M I, 2019. A Differential Evolution Markov Chain Monte Carlo Algorithm for Bayesian Model Updating. In: Dervilis, N. (eds) Special Topics in Structural Dynamics, Volume 5. Conference Proceedings of the Society for Experimental Mechanics Series. Springer, Cham. DOI: 10.1007/978-3-319-75390-4_9

Strömqvist J, Arheimer B, Dahné J, Donnelly C, Lindström G, 2012. Water and nutrient predictions in ungauged basins: set-up and evaluation of a model at the national scale. Hydrological Sciences Journal, 57(2), 229–247. DOI: 10.1080/02626667.2011.637497

Appendix A

Additional HYPE modules

A1 Regional groundwater module

The regional groundwater module allows for a water flux from the third soil layer to downstream sub-basins and the outlet lake but does not feature any storage of water. The regional groundwater flow is determined by a recession coefficient and is then routed based on an assigned fraction to downstream sub-basins and/or the outlet lake. Regional groundwater to downstream sub-basins is added to the bottom soil box and may further upwell into soil boxes above it.

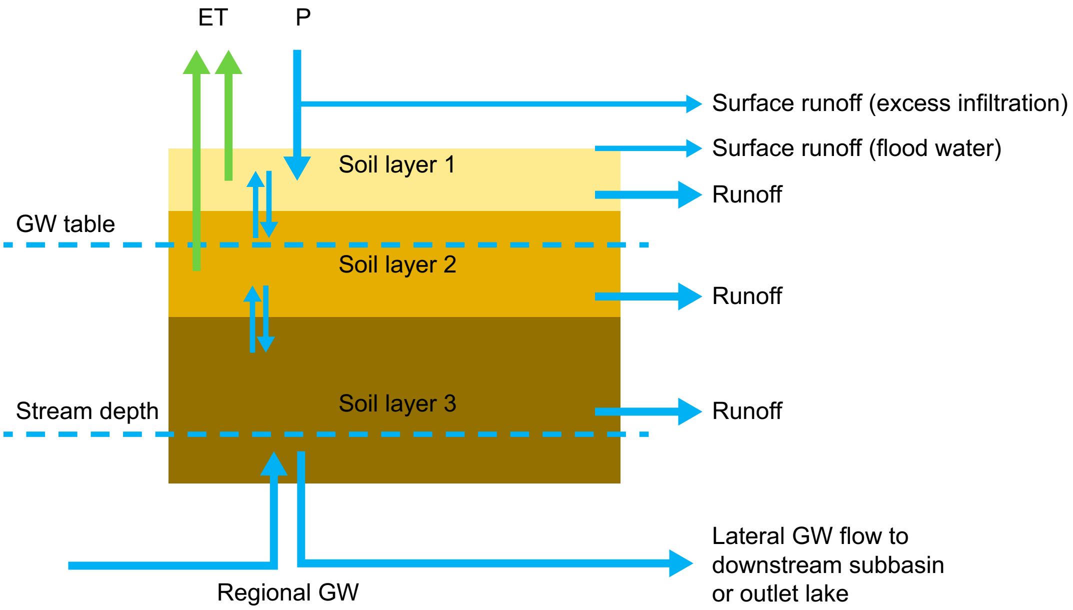

Figure A-1. *Water fluxes in an SLC when the module for regional groundwater flow is enabled. GW is groundwater, P is precipitation and ET is evapotranspiration (modified from smhi wiki).*

A2 Floodplain module

A floodplain module allows a floodplain to be represented as a part of the main river or lake area, Floodplain dynamics are simulated using either a simple flood plain representation which floods when the river or lake stage is higher than a set threshold, or a more complex floodplain representation with soil routines. The floodplain with soil routines calculates flow from soils to the non-flooded area of the floodplain (Figure A-2). Re-infiltration from the flooded areas of the floodplain to the soil may occur. At a certain threshold, the floodplain spills either back into the main river or into the outlet lake, depending on the model setup.

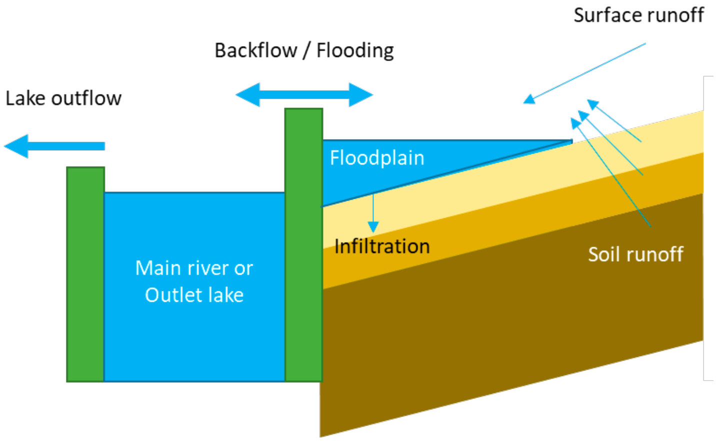

Figure A-2. *Visualization of the floodplain model in HYPE when using the floodplain soil routine.*

Appendix B

Reclassification schemes of soil type, land use and vegetation types

B1 Reclassification of soil types

Spatial extent of soil types was based on SGU's soil map (1:25 000, Table 2-4). The soil types were reclassed into five soil typees for the HYPE model (Table B-1). GIS and the SKB soil model (Table 2-4) were then used to calculate average soil thickness of each HYPE soil type (Table B-2). The average thickness of the soil types was then further aggregated to a maximum of 3 layers to help guide parameterization of HYPE's soil boxes (Table B-3).

Table B-1. Reclassification scheme, SGU soil types to aggregated classes used the HYPE model setup for this study.

Soil type	HYPE soil type	HYPE Soil no.
Peat	Organic	1
Clay	Clay	2
clay gyttja-gyttja clay	Clay	2
Clay, unspecified	Clay	2
Glacial clay, unspecified	Clay	2
Glacial silt, with banded clay	Clay	2
Gyttja	Clay	2
Gyttja, clay-gyttja	Clay	2
Postglacial fine sand, unspecified	Clay	2
Washed material, clay, silt	Clay	2
Filling material	Coarse	3
Filling material, unclassified	Coarse	3
Glacial coarse silt	Coarse	3
Glacial sediment	Coarse	3
Glacial sediment, coarse silt block	Coarse	3
Glacial sediment, gravel, cobble, sand	Coarse	3
Postglacial fine sand	Coarse	3
Postglacial gravel	Coarse	3
Postglacial midsized sand coarse sand	Coarse	3
Postglacial midsized sand gravel	Coarse	3
Cobble, boulders (glacial and post glacial)	Coarse	3
Washed deposit	Coarse	3
Washed deposit, fine sand	Coarse	3
Washed deposit, gravel	Coarse	3
Washed deposit, medium sand, coarse sand	Coarse	3
Washed deposit, sand	Coarse	3
Washed deposit, cobble boulder	Coarse	3
Washed deposit, unspecified	Coarse	3
Till	Till	4
Till, clayey sandy	Till	4
Till, low clayey content	Till	4
Till, sandy	Till	4
Till, sandy-silty	Till	4
Till clay and/or till, clayey	Till	4
Bedrock	Thin soil and bare rock	5
Bedrock, surface	Thin soil and bare rock	5
Crystalline bedrock	Thin soil and bare rock	5
Thin soil and bare rock	Thin soil and bare rock	5

Table B-2. Spatial average of soil type thickness [m] within each SLC. The Nomenclature "Zx" corresponds to the SKB soil model (Petrone and Sohleneus, 2020).

SLC	Peat Z2	Gyttja Z4c	Clay gyttja Z4a	Postglacial sand and gravel Z3a	Glacial clay Z4b	Fine Sediment Z4+Z3a	Till Z5
1	0.1	0.7	0.3	0.3	0.6	1.2	2.7
2	0.1	-	-	0.4	0.6	0.9	3.9
3	-	-	-	0.1	0.3	0.4	3.1
4	0.6	-	-	0.1	0.3	0.4	3.4
5	-	-	-	-	-	-	0.4
6	-	-	-	-	-	-	2.7
7	0.2	-	-	0.3	0.5	0.7	4.1
8	0.3	-	-	0.2	0.5	0.7	3.1
9	0.7	-	-	0.1	0.2	0.4	3.6
10	0.1	-	-	-	0.1	0.1	2.8

Table B-3. Spatial average of soil type thickness [m] within each SLC, after aggregation of soil types.

Aggregated soil type thickness [m]				
SLC	Organic	Coarse	Fine	Till
1		2.0		5.9
2		0.5	1.9	5.9
3			0.8	3.9
4	0.6		1.4	4.8
5				0.5
6				2.7
7	0.2		1.6	5.7
8		1.0	1.7	4.8
9	0.7		1.4	5.0
10			0.3	3.1

Model and calibration parameters

Calibration parameter ranges are shown in Table C-1 and calibrated parameters for Lake Eckarfjärden are shown in Table C-2. Drainage basin parameters and SLC spatial coverage are shown in Table C-3 and Lake description and lake parameters are shown in Table C-4.

Table C-1. Calibration parameters (optpar-file) used for the of HYPE for this study, using the Differential Evolution Markov Chain method. The first value of each parameter (global, land use and soil) is the upper boundary of the parameter range, the second value is lower boundary, and the third value is the minimum step for parameter change.

Parameter	Parameter values			
Differential Evolution Markov Chain				
ngen	1 500			
npop	400			
gammascale	1			
sigma	0.1			
crossover	0.8			
accprob	0			
Global				
preccorr	−0.05			
preccorr	0.05			
preccorr	0.01			
cevpcorr	−0.05			
cevpcorr	0.05			
cevpcorr	0.01			
rivvel	0.1			
rivvel	1			
rivvel	0.01			
damp	0.1			
damp	0.7			
damp	0.01			
gratp	0.001			
gratp	2			
gratp	0.001			
gratk	0			
gratk	2			
gratk	0.001			
Landuse	**lake**	**mire**	**forest**	**grass**
cmlt	1	1	1	1
cmlt	4	4	2.5	4
cmlt	0.01	0.01	0.01	0.01
ttmp	−1	−1	−1	−1
ttmp	2	2	2	2
ttmp	0.02	0.02	0.02	0.02
kc	0.8	0.8	0.8	0.8
kc	1.2	1.2	1.2	1.2
kc	0.01	0.01	0.01	0.01

Table C-1. Continued.

Parameter	Parameter values					
Soil type	**organic**	**clay**	**coarse**	**till**	**thin soil and bare rock**	**water**
wcfc	0.025	0.175	0.025	0.08	0.08	0.05
wcfc	0.1	0.25	0.1	0.175	0.175	0.5
wcfc	0.001	0.001	0.001	0.001	0.001	0.001
wcwp	0.05	0.05	0.05	0.05	0.05	0.05
wcwp	0.5	0.5	0.5	0.5	0.5	0.5
wcwp	0.001	0.001	0.001	0.001	0.001	0.001
wcep	0.2	0.1	0.2	0.03	0.03	0.03
wcep	0.4	0.35	0.4	0.4	0.4	0.4
wcep	0.01	0.01	0.01	0.01	0.01	0.01
mperc1	45	15	45	20	20	15
mperc1	75	30	75	75	75	30
mperc1	0.1	0.1	0.1	0.1	0.1	0.1
mperc2	20	20	20	20	20	20
mperc2	75	75	75	75	75	75
mperc2	0.1	0.1	0.1	0.1	0.1	0.1
rrcs1	0.05	0.05	0.05	0.05	0.05	0.05
rrcs1	0.6	0.6	0.6	0.6	0.6	0.6
rrcs1	0.001	0.001	0.001	0.001	0.001	0.001
rrcs2	0.05	0.05	0.05	0.005	0.05	0.05
rrcs2	0.6	0.6	0.6	0.1	0.6	0.6
rrcs2	0.001	0.001	0.001	0.001	0.001	0.001
srrate	0	0	0	0	0	0
srrate	0.02	0.02	0.02	0.02	0.02	0.02
srrate	0.01	0.01	0.01	0.01	0.01	0.01

Table C-2. Parameters (par-file) used for HYPE in this study.

Parameter	Parameter values			
Global				
rivvel	0.362			
tcalt	0.6			
epotdist	3			
gldepi	3			
sdnsnew	0.1			
snowdensdt	0.004			
gratp	1.771			
gratk	1.447			
cevpcorr	0.034			
preccorr	0.011			
ratcorr	0.065			
damp	0.183			
Landuse	**lake**	**mire**	**forest**	**grass**
cmlt	2.96	1.23	1.79	2.69
ttmp	1.81	0.97	0.05	0.77
srrcs	0.10	0.10	0.10	0.10
kc	1.04	0.34	0.15	0.06
pcluse	0	0	0.10	0
fepotsnow	1	1	1	1
frost	2	2	2	2

Table C-2. Continued.

Parameter	Parameter values					
Soil type	**organic**	**clay**	**coarse**	**till**	**thin soil and bare rock**	**water**
Tot por	0.7	0.48	0.48	0.45	0.3	0.43
wcfc	0.03	0.24	0.01	0.01	0.05	0.30
wcwp	0.18	0.23	0.25	0.22	0.09	0.38
wcep	0.37	0.14	0.30	0.17	0.07	0.27
mperc1	70.61	18.81	34.95	22.44	12.57	14.58
mperc2	64.45	23.44	69.07	35.69	51.81	43.03
rrcs1	0.047	0.035	0.218	0.001	0.200	0.557
rrcs2	0.553	0.459	0.437	0.003	0.073	0.551
srrate	0.019	0.015	0.013	0.012	0.015	0.017
sfrost	1	1	1	1	1	1

Table C-3. Drainage basin parameters and SLC spatial coverage (geodata-file) used for HYPE in this study. The fraction number for the SLCs are the areal extent within Lake Eckarfjärden drainage basin.

Parameter	Parameter value
Area	2 277 477
Rivlen	1 350
Slope_mean	3
Elev_mean	9
SLC_1	0.12
SLC_2	0.01
SLC_3	0.02
SLC_4	0.06
SLC_5	0.08
SLC_6	0.66
SLC_7	0.01
SLC_8	0.01
SLC_9	0.01
SLC_10	0.02
Lake depth	0.75
Icatch	0.98
parreg	1

Table C-4. Lake parameters (lakedata-file) used for HYPE in this study.

Parameter	Parameter value
Area	283 849
Lake depth	0.91
Rate	0.23
Exp	0.67
w0ref	5.28

Calibration and Evaluation period B

D1 Model performance, calibration period B

D1.1 Streamflow

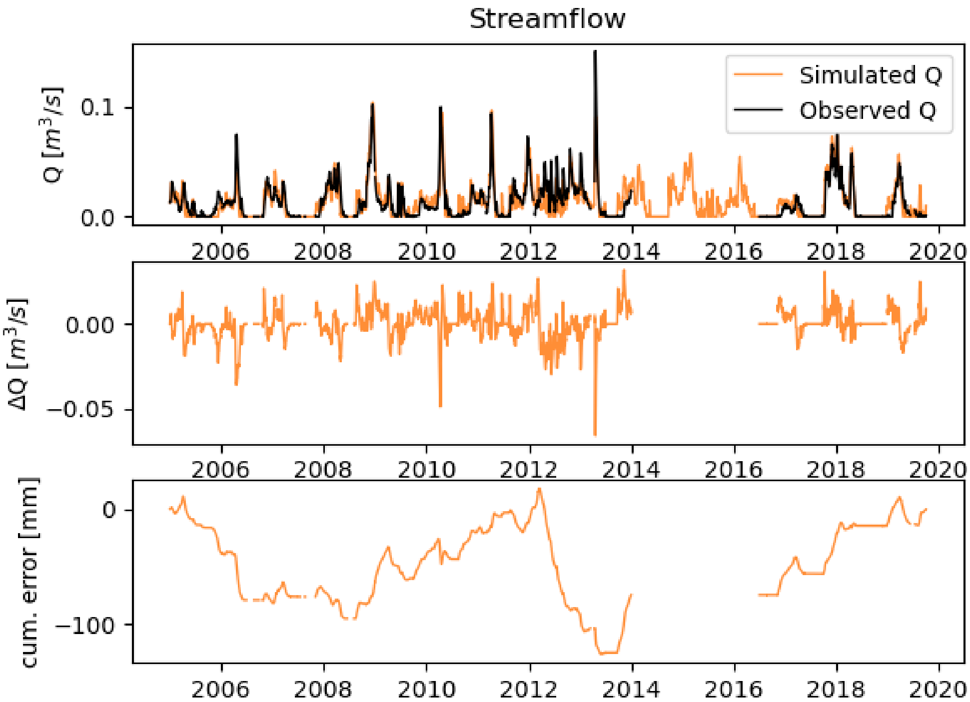

Figure D-1. *Model performance for calibration period B. Top: Simulated streamflow (orange line) and observed streamflow (black line). Mid: error in streamflow estimations (R observed – R simulated). Bottom: cumulative error of runoff [mm].*

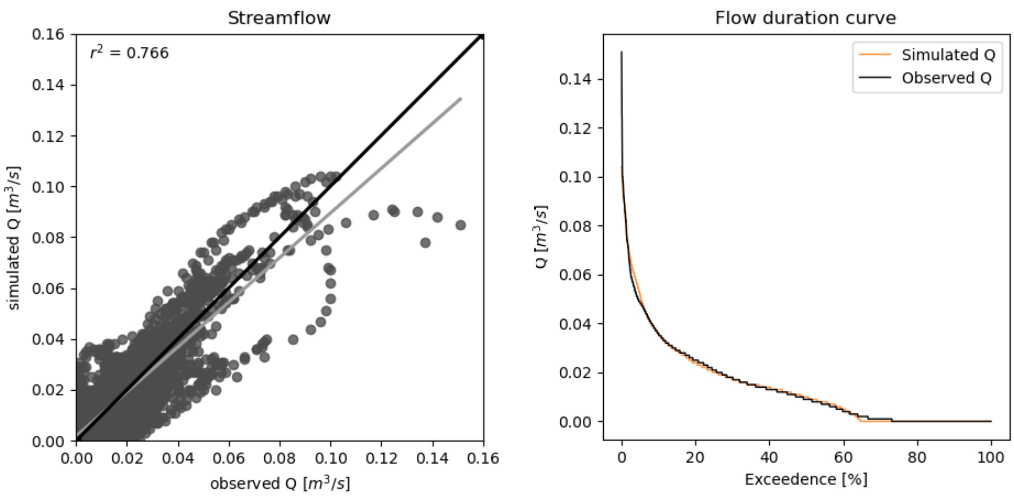

Figure D-2. *Left: linear regression of simulated and observed streamflow for calibration period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Flow duration curve of simulated (orange line) and observed (black line) streamflow for calibration period B.*

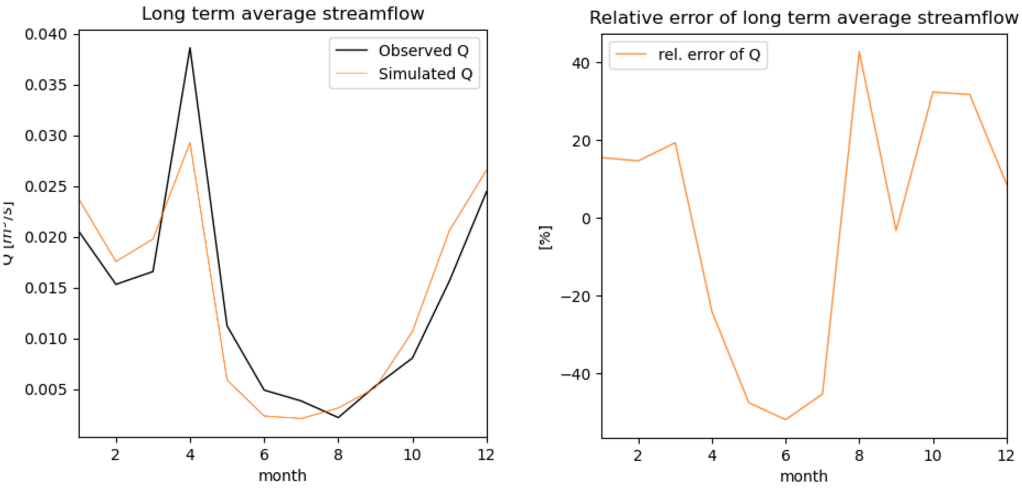

__Figure D-3.__ Left: Monthly average streamflow for calibration period B. Both simulated (orange line) and observed (black line) streamflow is shown. Right: Relative error ((Q simulated – Q observed)/Q observed) of monthly average streamflow for calibration period B.

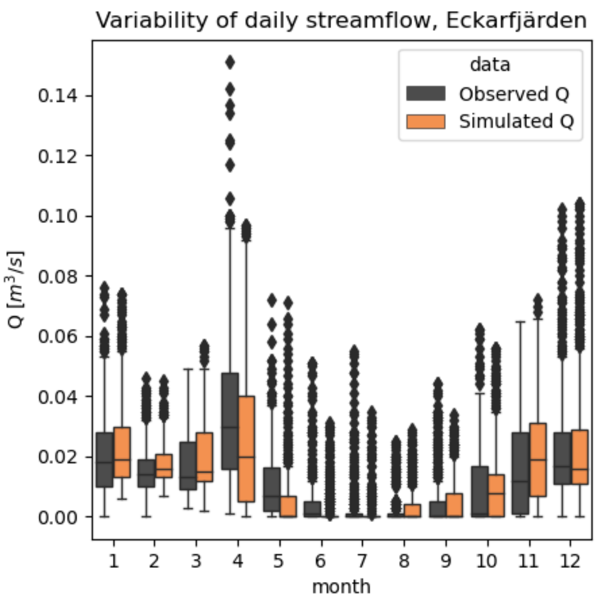

__Figure D-4.__ Daily streamflow for the calibration period B, for each month of the year. Simulated streamflow is shown in orange and observed streamflow is shown in grey.

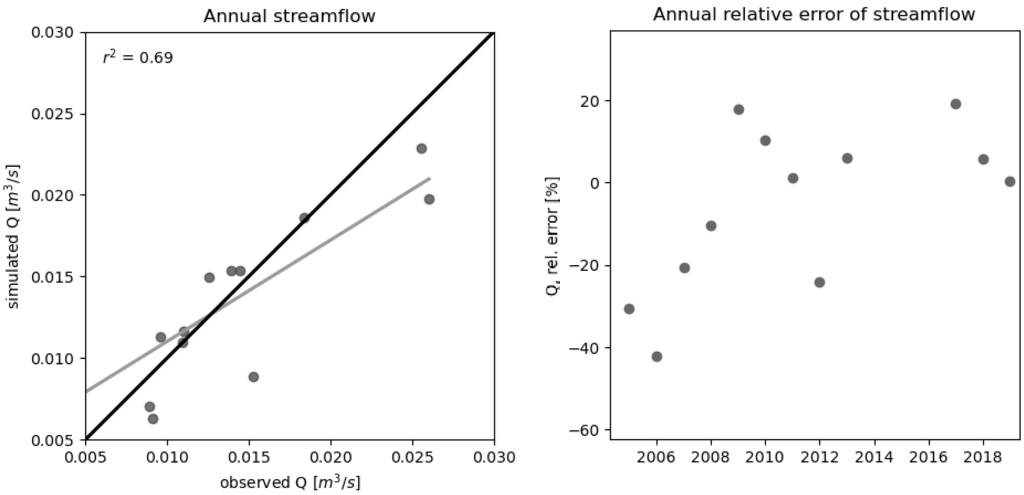

Figure D-5. *Left: linear regression of simulated and observed annual streamflow for calibration period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Relative error [%] of simulated and observed annual streamflow for the calibration period B.*

D1.2 Lake level

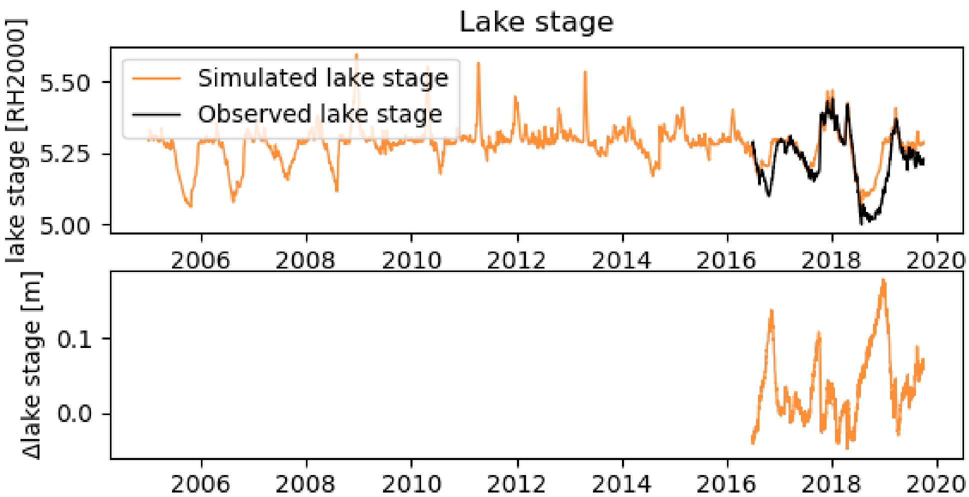

Figure D-6. *Model performance for calibration period B. Top: Simulated lake stage of Eckarfjärden (orange line) and observed lake stage of Eckarfjärden (black line). Bottom: error in estimated lake stage (observed – simulated).*

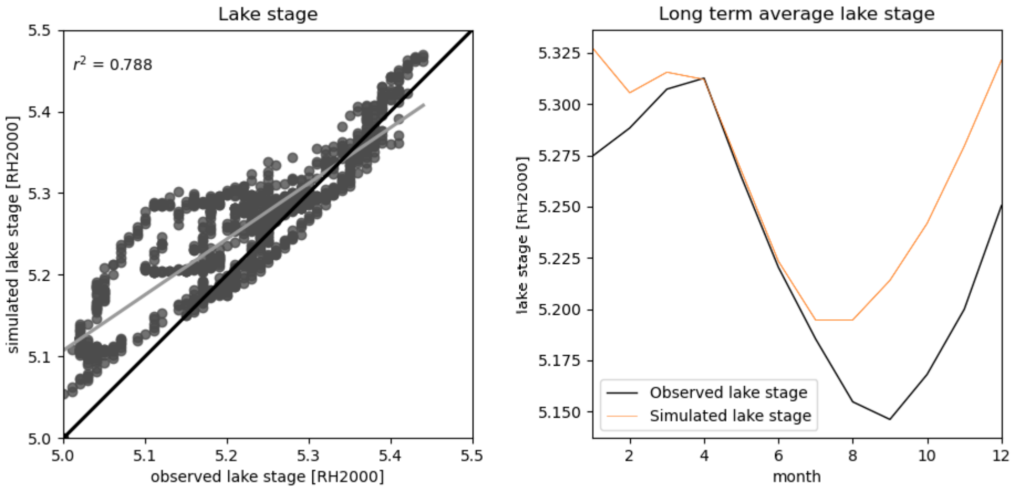

Figure D-7. *Left: linear regression of simulated and observed lake stage for Eckarfjärden for calibration period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Lake Eckarfjärden monthly average lake stage for calibration period B. Both simulated (orange line) and observed (black line) lake stage are shown.*

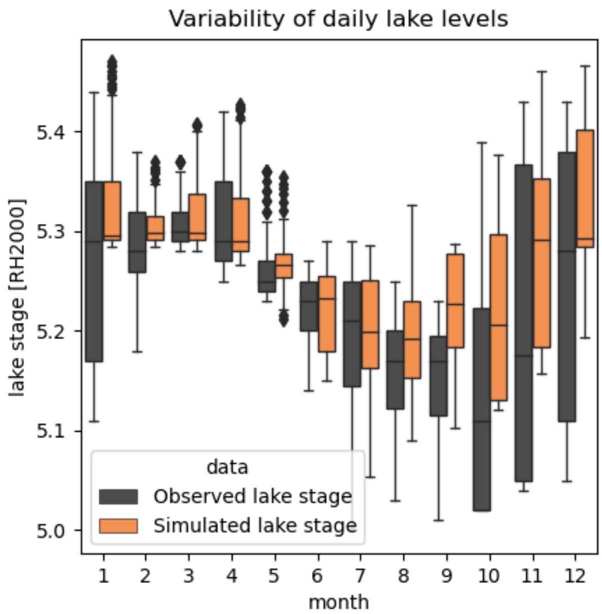

Figure D-8. *Eckarfjärden daily lake stage for calibration period B, for each month of the year. Simulated lake stage is shown in orange and observed lake stage is shown in grey.*

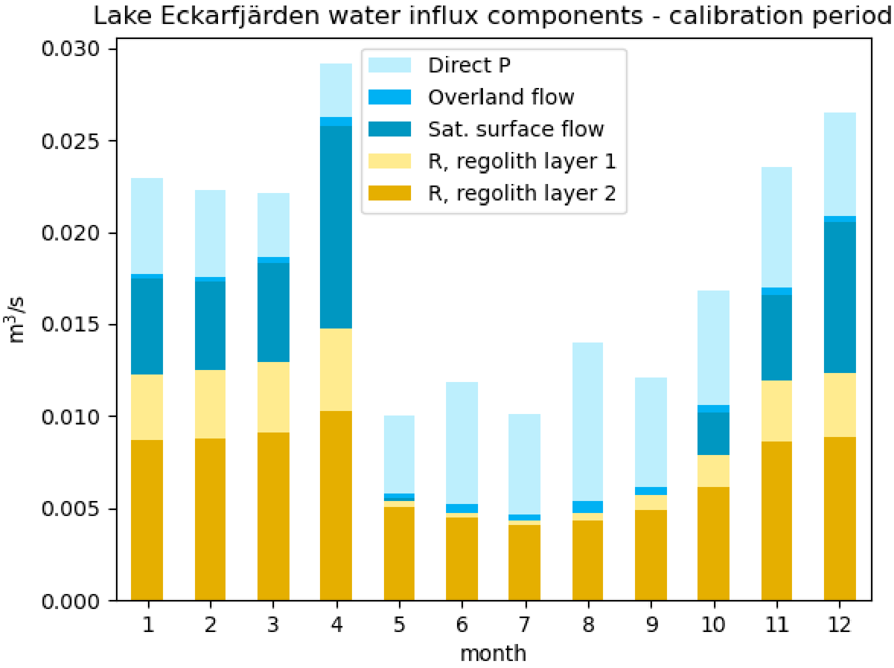

Figure D-9. *Water influxes to Lake Eckarfjärden for calibration period B. "Overland flow" is excess precipitation and snowmelt, "Sat. surface flow" occurs when the groundwater level is higher than the ground level, "R, soil layer 1" and "2" are groundwater contributions from the upper two soil layers. "Soil layer 3" does not contribute to runoff in this model realization.*

D1.3 Groundwater

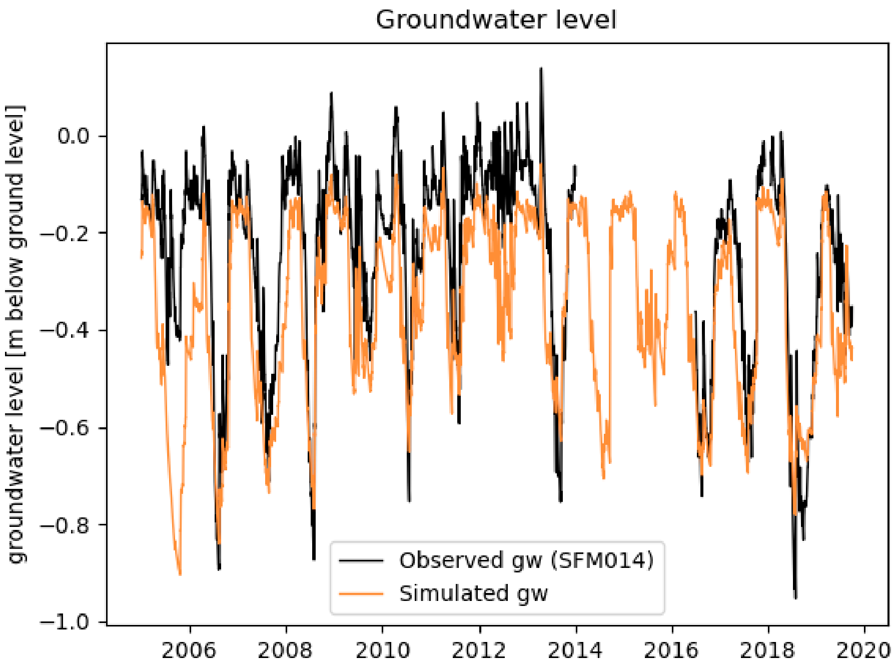

Figure D-10. *Simulated groundwater level (orange line) and observed groundwater level (black line) for calibration period B.*

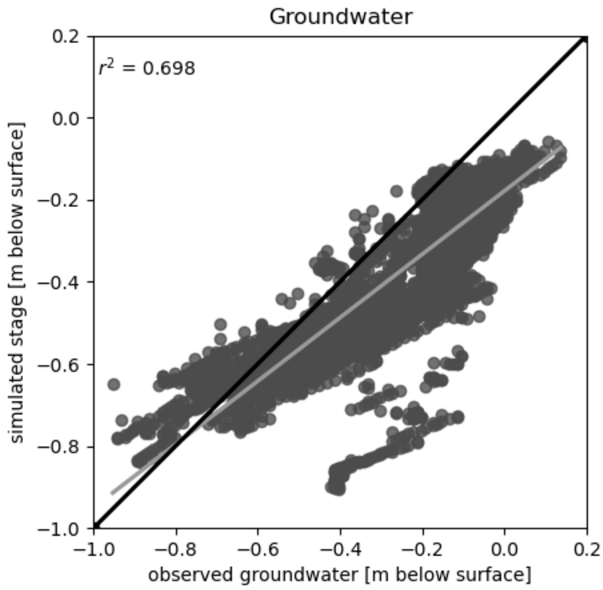

Figure D-11. *Linear regression of simulated and observed groundwater level for calibration period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

D1.4 Snow

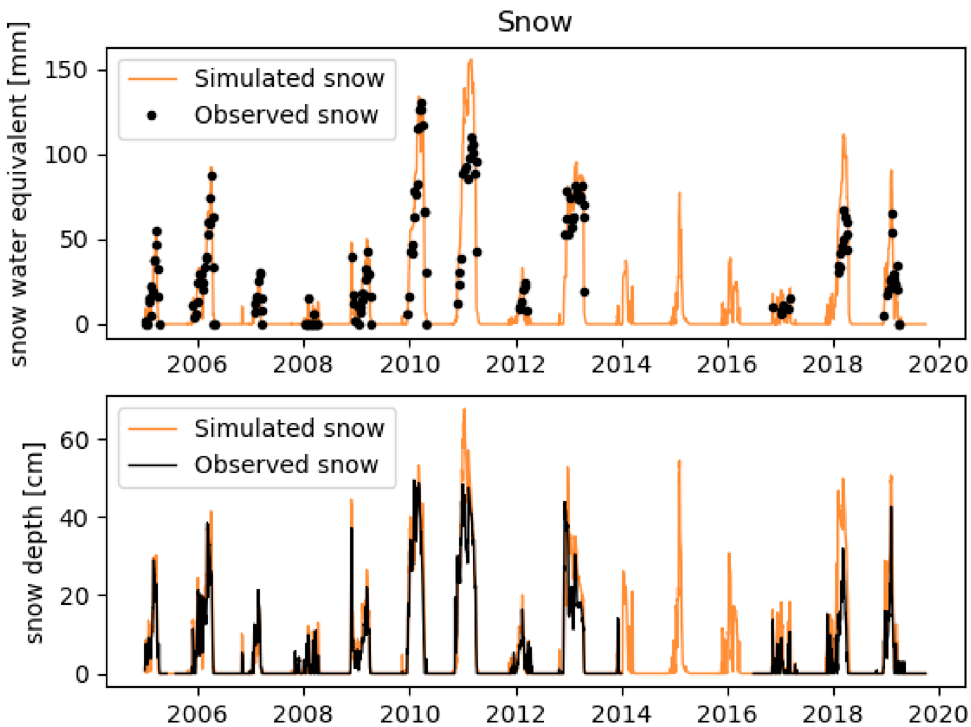

Figure D-12. *Snow time series data. Top: Simulated snow water equivalent (orange line) and observed snow water equivalent (black line) for calibration period B. Bottom: Simulated snow depth (orange line) and observed snow depth (black line) for calibration period B.*

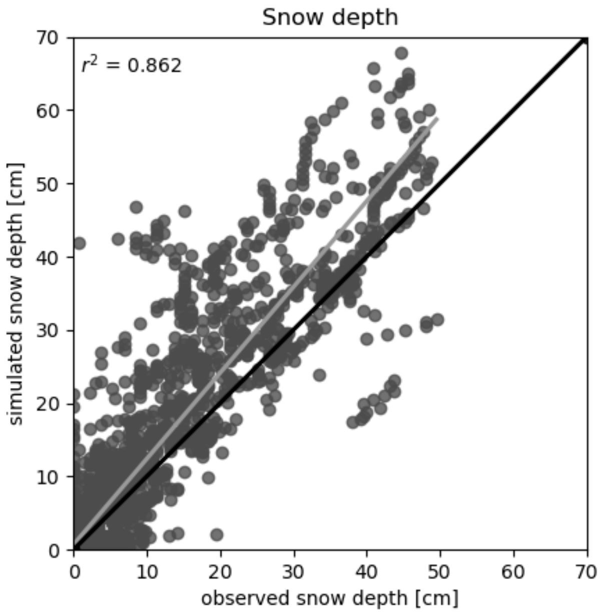

Figure D-13. *Linear regression of simulated and observed snow depth for calibration period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

D2 Model performance, evaluation period B

D2.1 Streamflow

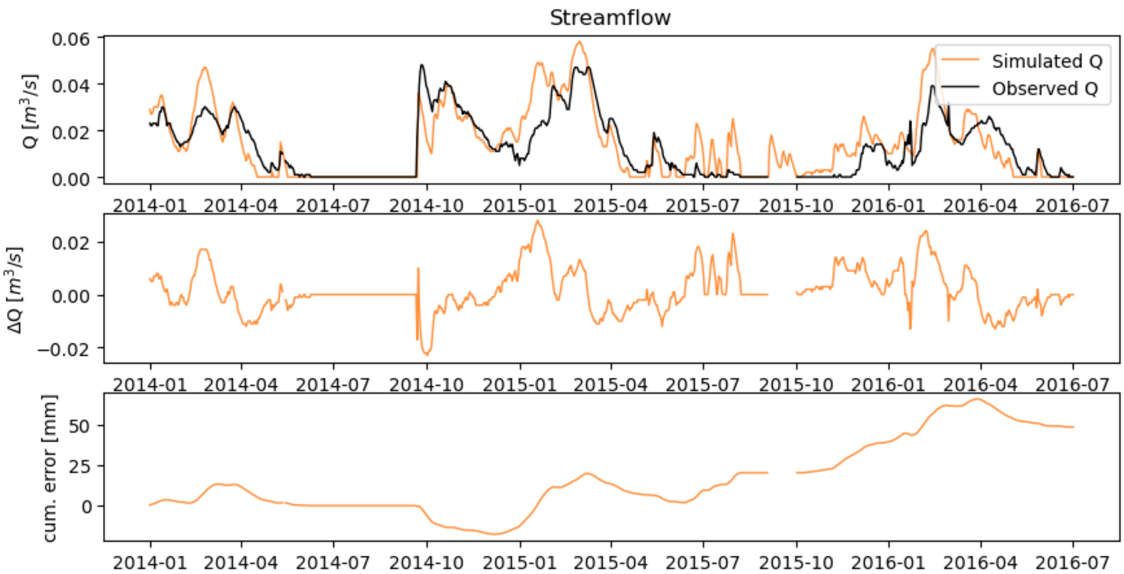

Figure D-14. *Model performance for evaluation period B. Top: Simulated streamflow (orange line) and observed streamflow (black line). Mid: error in streamflow estimations (R observed – R simulated). Bottom: cumulative error of runoff [mm].*

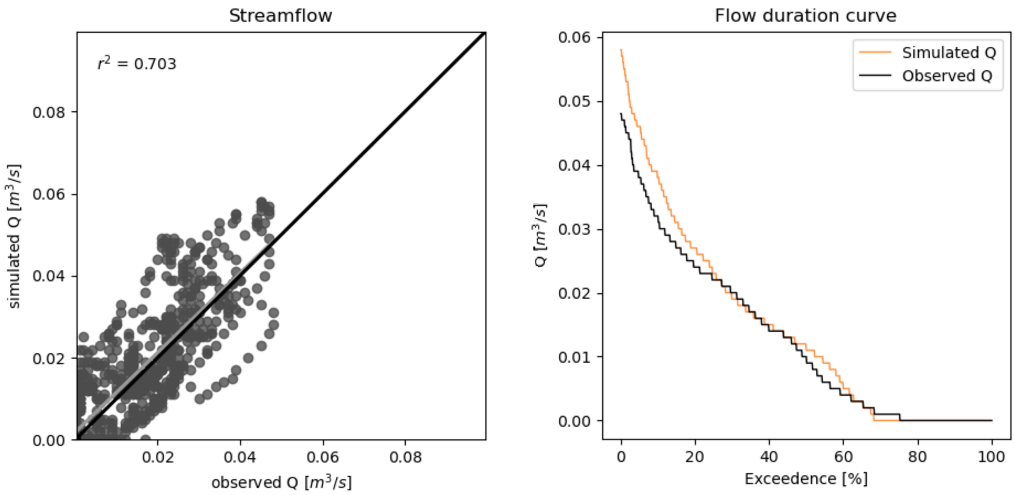

Figure D-15. *Left: linear regression of simulated and observed streamflow for evaluation period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship. Right: Flow duration curve of simulated (orange line) and observed (black line) streamflow for evaluation period B.*

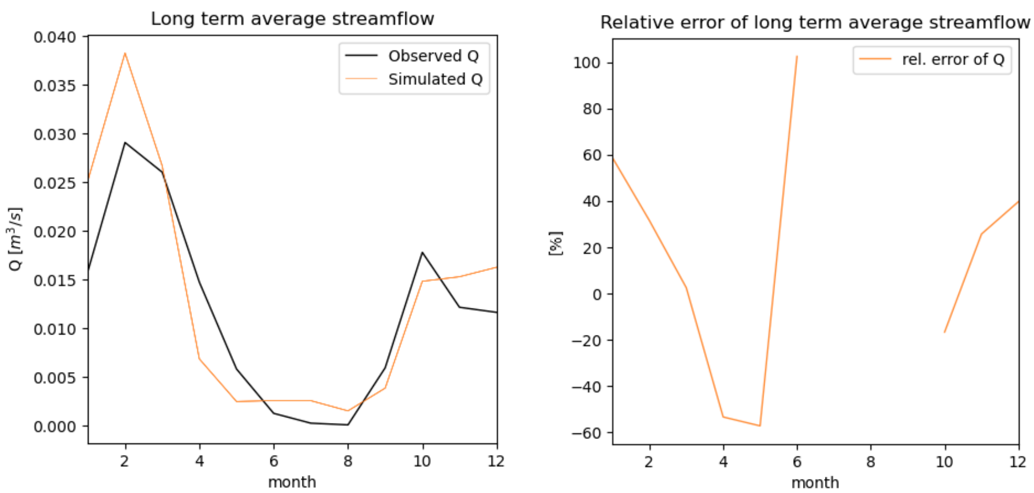

Figure D-16. *Left: Monthly average streamflow for evaluation period B. Both simulated (orange line) and observed (black line) streamflow is shown. Right: Relative error ((Q simulated – Q observed)/Q observed) of monthly average streamflow for months with streamflow during evaluation period B.*

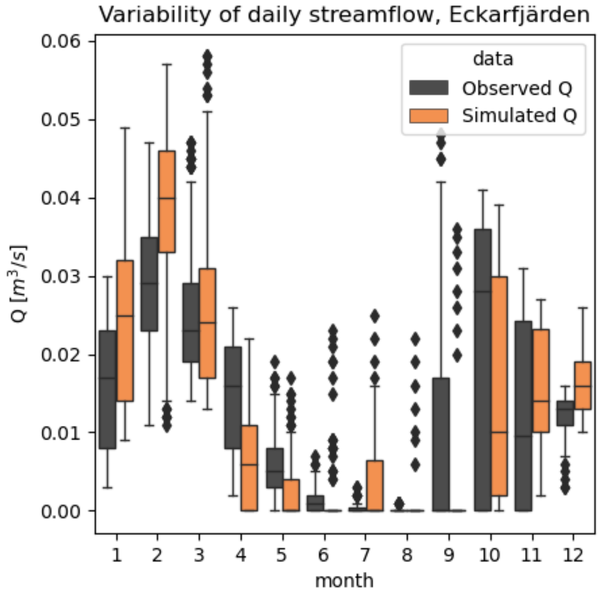

Figure D-17. *Daily streamflow for the evaluation period B, for each month of the year. Simulated streamflow is shown in orange and observed streamflow is shown in grey.*

D2.2 Lake level

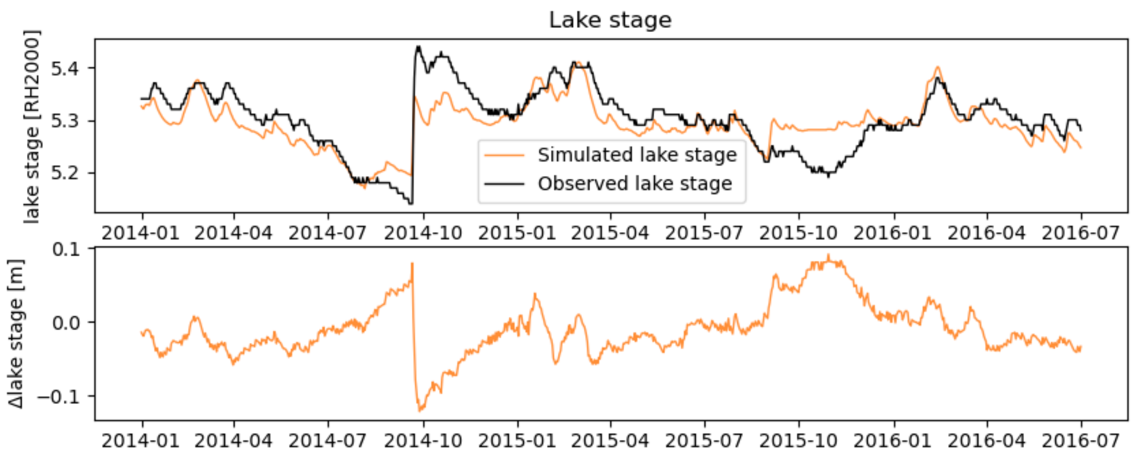

Figure D-18 Model performance for evaluation period B. *Top: Simulated lake stage of Eckarfjärden (orange line) and observed lake stage of Eckarfjärden (black line). Bottom: error in estimated lake stage (observed – simulated).*

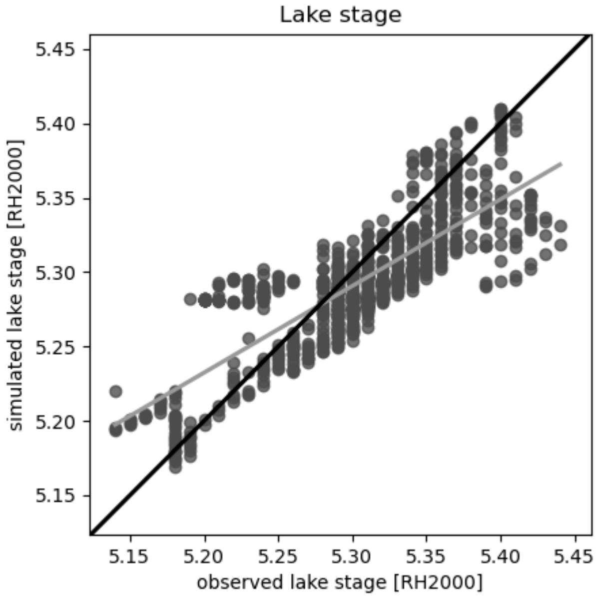

Figure D-19. *Linear regression of simulated and observed lake stage for Eckarfjärden for evaluation period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

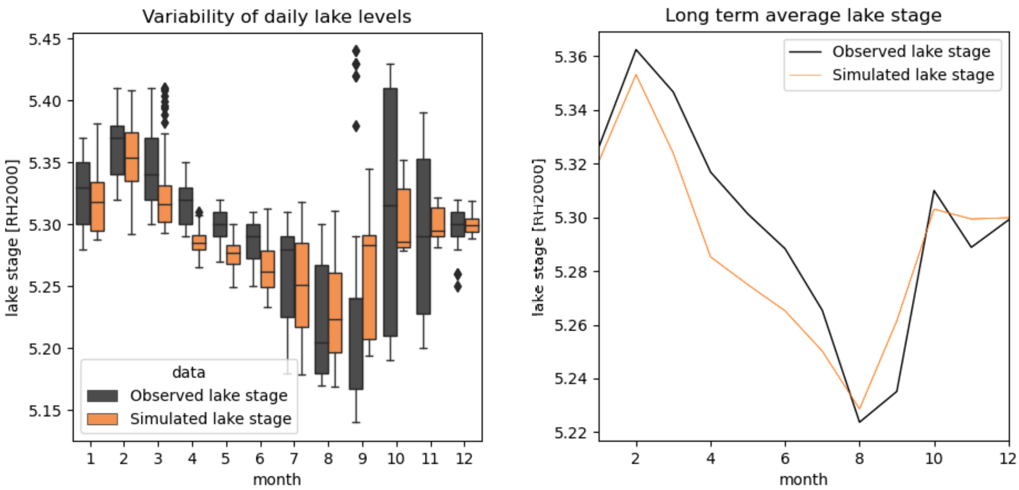

Figure D-20. *Left: Eckarfjärden daily lake stage for evaluation period B, for each month of the year. Simulated lake stage is shown in orange and observed lake stage is shown in grey. Right: Lake Eckarfjärden monthly average lake stage for evaluation period B. Both simulated (orange line) and observed (black line) lake stage are shown.*

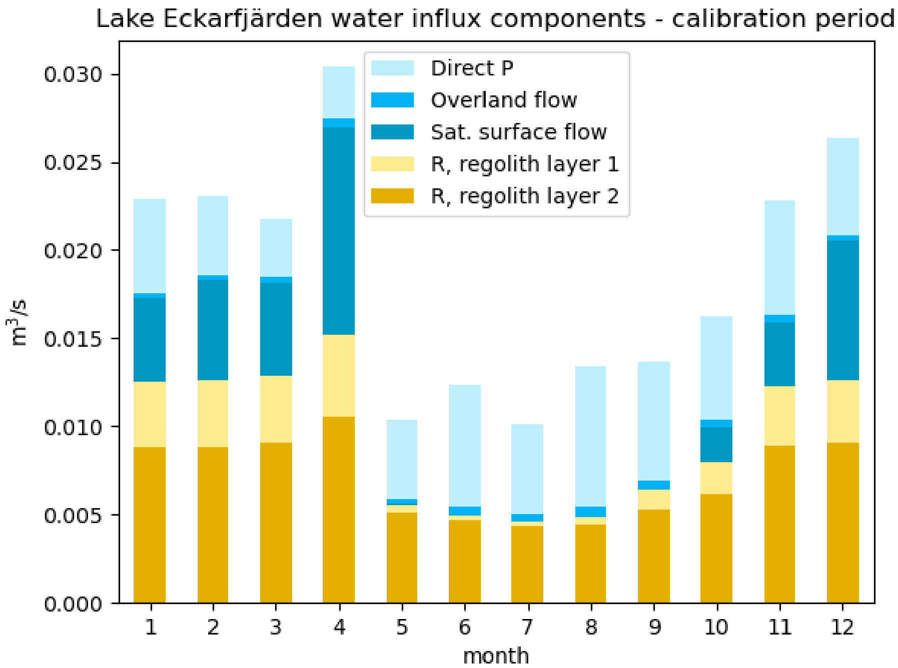

Figure D-21. *Water influxes to Lake Eckarfjärden for evaluation period B. "Overland flow" is excess precipitation and snowmelt, "Sat. surface flow" occurs when the groundwater level is higher than the ground level, "R, soil layer 1" and "2" are groundwater contributions from the upper two soil layers. "Soil layer 3" does not contribute to runoff in this model realization.*

D2.3 Groundwater

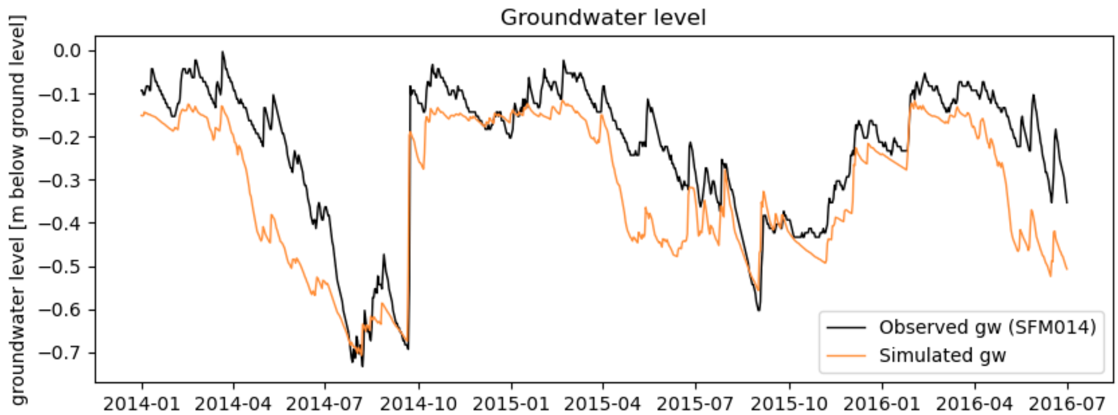

Figure D-22. *Simulated groundwater level (orange line) and observed groundwater level (black line) for evaluation period B.*

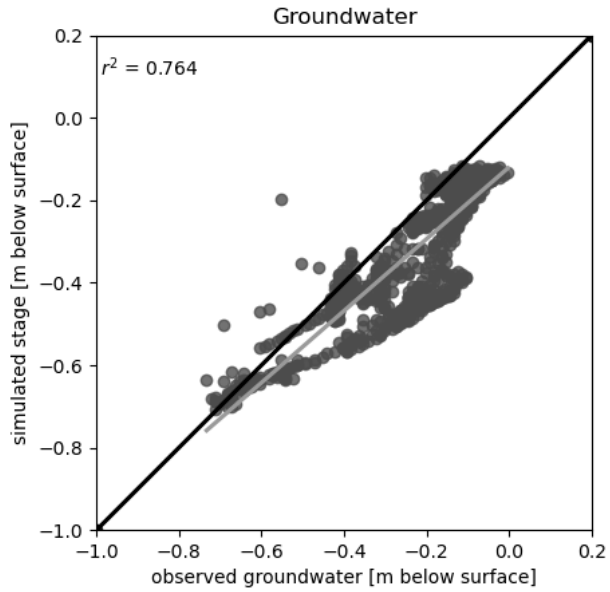

***Figure D-23.** Linear regression of simulated and observed groundwater level for evaluation period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

D2.4 Snow

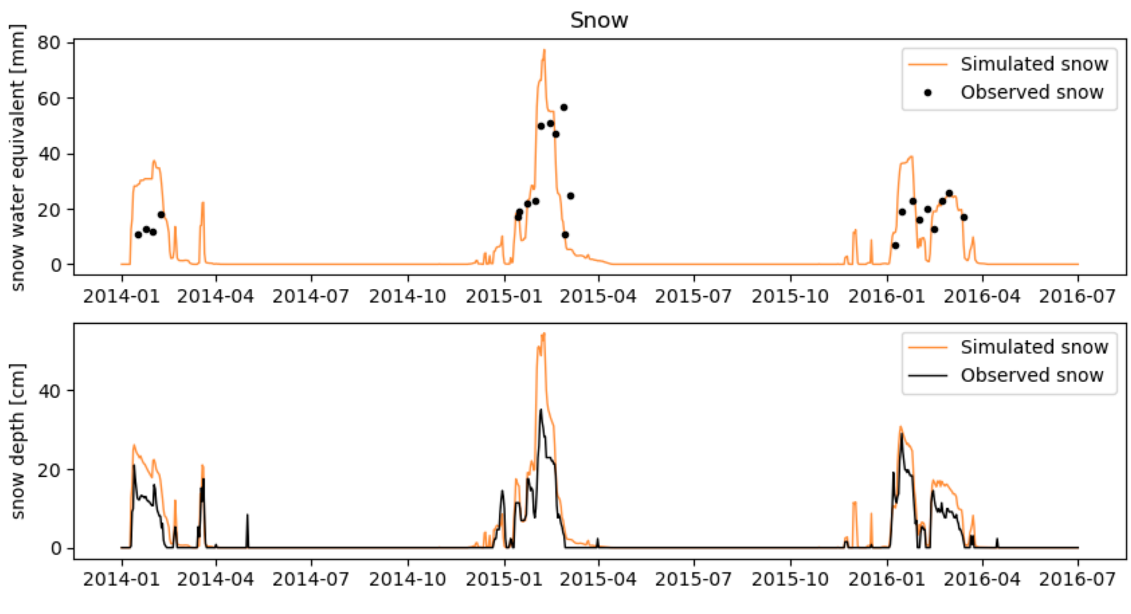

***Figure D-24.** Snow time series data. Top: Simulated snow water equivalent (orange line) and observed snow water equivalent (black line) for evaluation period B. Bottom: Simulated snow depth (orange line) and observed snow depth (black line) for evaluation period B.*

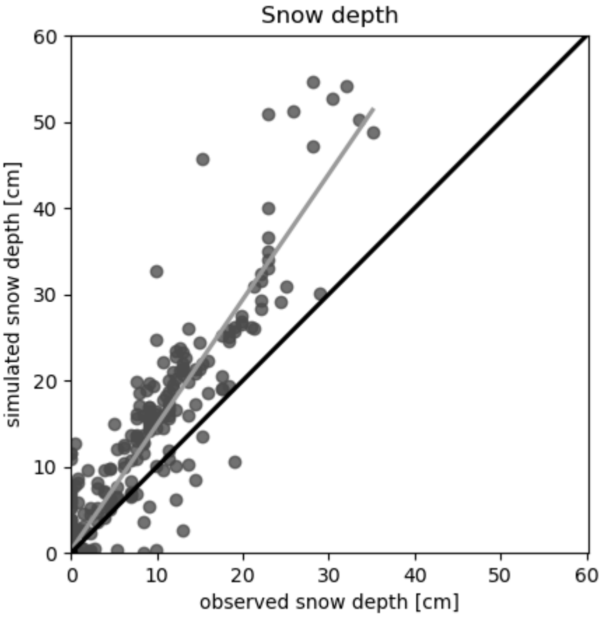

Figure D-25. *Linear regression of simulated and observed snow depth for evaluation period B. The grey line shows the linear regression, the black line shows what would be a 1:1 relationship.*

Arkitektkopia AB, Bromma, 2024

SKB is responsible for managing spent nuclear fuel and radioactive waste produced by the Swedish nuclear power plants such that man and the environment are protected in the near and distant future.

skb.se