Master thesis

Modelling reservoir dynamics in a data scarce semi-arid region

Erik van Welsenes

October 25th, 2023

Delft University of Technology
Faculty of Civil Engineering
Water Management

Masterthesis

Modelling reservoir dynamics in a data scarce semi-arid region

Erik van Welsenes

in partial fulfillment of the requirements for the degree of

Master of Science
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on November 1st, 2023 at 13:15

Thesis Committee

Dr.	Markus Hrachowitz	Delft University of Technology
MSc.	Michel Zuijderwijk	Witteveen+Bos
Ir.	Daan te Witt	Witteveen+Bos
Prof. dr.ir.	Nick van de Giesen	Delft University of Technology

Institution

Delft University of Technology
Faculty of Civil Engineering
Water Management
The Netherlands

Preface

This thesis represents the description of my MSc thesis project and is written to obtain the degree of Master of Science in Civil Engineering - Water Management at Delft University of Technology. I carried out this research project with the help of my supervisors from Delft University of Technology, but also with the help of the group Water Nuisance at Witteveen+Bos, where I did an internship during the research period. Witteveen+Bos is working on investigating different rehabilitation measurements to increase the water availability around 4 reservoirs in Burkina Faso. With my thesis, I had the chance to contribute to this project, by investigating how to create a reliable model to simulate the reservoir storage fluctuations.

My thesis can be read by anyone interested in hydrological modelling in any data scarce region, as the methods applied and tested are not only valid for the Nakambé Catchment but any other catchment. In my thesis, I compared two models, even though I put a lot of effort and time into trying to improve the models, due to the limited data the improvements were only very limited and nothing can be concluded with 100% certainty. That this is the case and how to cope with this is one of the main things I have learned during my thesis. The assignments during my study always worked out and the data needed was nearly always available, however, in reality this is not the case.

I would like to thank Witteveen+Bos for giving me the opportunity to work together and get to know the company. I had my own project during my time at Witteveen+Bos, however, I felt included within Witteveen+Bos. Already, on my third day, I was welcome to the department trip and during my whole time at Witteveen+Bos, I was welcomed to team meetings, coffee breaks, and the usual lunch walks. It was always interesting to hear about other projects, and people were also interested in my project, which made me feel like a member of the team instead of an outsider.

I would like to express my gratitude to all of my supervisors: Markus Hrachowitz, Michel Zuijderwijk, Daan te Witt and Nick van de Giesen. Special thanks to my daily university supervisor Markus, who was always quick to understand the graphs I showed and really helped me to better think about and understand the obtained results, and helped a lot with what steps to take to improve the model and results. Next, I want to thank Daan with whom I had a lot of interesting discussions about the results, and for all his help with the Wflow model. Michel was always very busy but if needed, I could always ask for help and want to thank him for the practical solutions he offered during all the discussions.

Lastly, I want to thank my family and friends for the help and motivation they gave me during my whole thesis.

Erik van Welsenes
Delft, Oktober 25, 2023

Summary

In order to evaluate measures to increase water availability in and around reservoirs it is necessary to have reliable reservoir water storage models. For that reason, it is necessary to assess the validity of these models. The aim of this research is to do this assessment of two models in data scarce semi-arid regions. The Nakambé catchment is used as a case study. The study covers important aspects of hydrological modelling, such as model input (data) selection, hydrological model choice, calibration, model performance testing, parameter sensitivity, and reservoir water storage simulation.

For the selection of the optimal model forcing data diverse precipitation products are reviewed: CHIRPS, ERA5, and local measurements. Among the evaluated datasets, CHIRPS emerges as the superior choice validated against the local measurements. With respect to the potential evaporation, the combination of ERA5 and local measurements results in the most suitable potential evaporation data, leveraging the temporal and spatial aspects of ERA5 and the absolute values of the local measurements.

Comparing a lumped hydrological model (HBV) with a distributed model (SBM Wflow) in simulating river discharge reveals that the HBV model outperforms its counterpart in simulating discharge. This contrast in performance is attributed to potential overparameterization in the Wflow model, coupled with the complexities of parameter estimation in data-scarce areas. The HBV model, while bearing simplifications, benefits from a more comprehensive calibration process. The model performance is strongly influenced by the calibration efficiency, where the significantly shorter simulation time of the HBV model facilitates an extensive Monte Carlo sampling-based calibration, in contrast to Wflow's time consuming manual parameter adjustment.

Additionally, the sensitivity analysis showed that in the HBV model, the parameters affecting actual evaporation are the most sensitive one. This emphasizes the importance of accurately simulating this component for the proper model performance. The Wflow model exhibits strong equifinality due to the many parameters within the model. The complexity of this model made it impossible to test all parameters and therefore only some parameters are tested.

Both reservoir water storage models studied, the HBV Reservoir Water Storage Model (HBV_RWSM) and the Wflow reservoir module, can effectively simulate reservoir water storage fluctuations, although they differ in how the components are calculated. Due to data limitations, it is impossible to determine which, if any, of the models is correct. However, based on the downstream discharge the HBV_RWSM displays a more promising performance.

In conclusion, the HBV model outperformed the SBM_Wflow model in simulating discharge due to its simplicity and ease of calibration. Sensitivity analyses highlighted the significance of accurately representing actual evaporation. Both water balance models, the HBV_RWSM and the Wflow reservoir module, performed similarly concerning the NSE values. The fluxes contributing to the water balance in the two reservoir water storage models differ significantly. The lack of data on these fluxes makes it impossible to determine which models performs best. Data limitations remain a significant hurdle in model evaluation, emphasizing the need for additional data collection, particularly upstream and downstream of the reservoir, to enhance reliability and reduce uncertainties.

Contents

1 Introduction **1**
- 1.1 Background . . . 1
- 1.2 Problem statement . . . 1
- 1.3 Report structure . . . 2

2 Study area & Data **4**
- 2.1 Study Area . . . 4
  - 2.1.1 Sahelian hydrological paradox . . . 6
- 2.2 Static input data . . . 6
  - 2.2.1 Topography . . . 6
  - 2.2.2 Land use . . . 7
- 2.3 Dynamic input data . . . 8
  - 2.3.1 Precipitation . . . 8
  - 2.3.2 Potential evaporation . . . 9
- 2.4 Validation and Calibration data . . . 13
  - 2.4.1 Discharge . . . 13
  - 2.4.2 Reservoir water storage . . . 14
- 2.5 Focus location & period . . . 14

3 Methodology **16**
- 3.1 Data selection . . . 16
  - 3.1.1 Precipitation . . . 16
  - 3.1.2 Potential Evaporation . . . 17
  - 3.1.3 Discharge . . . 17
  - 3.1.4 Reservoir water storage . . . 18
- 3.2 Discharge models . . . 18
  - 3.2.1 HBV model . . . 18
  - 3.2.2 Wflow . . . 20
  - 3.2.3 Calibration . . . 21
  - 3.2.4 Model performance testing . . . 22
  - 3.2.5 Validation . . . 23
- 3.3 Sensitivity analysis . . . 23
  - 3.3.1 HBV model . . . 23
  - 3.3.2 Wflow . . . 24
- 3.4 Reservoir water storage models . . . 24
  - 3.4.1 HBV_RWSM . . . 24
  - 3.4.2 Wflow reservoir module . . . 26

4 Results **28**
- 4.1 Data selection . . . 28
  - 4.1.1 Precipitation . . . 28
  - 4.1.2 Potential Evaporation . . . 32
  - 4.1.3 Discharge . . . 33
  - 4.1.4 Reservoir water storage . . . 35
- 4.2 Hydrological models . . . 35
  - 4.2.1 Validation . . . 38
  - 4.2.2 Budyko Framework . . . 39
- 4.3 Sensitivity Analysis . . . 40
  - 4.3.1 HBV model . . . 40
  - 4.3.2 Wflow . . . 42
- 4.4 Reservoir water storage . . . 45
  - 4.4.1 HBV reservoir water storage model . . . 45
  - 4.4.2 Wflow reservoir module . . . 48
  - 4.4.3 Comparison the HBV_RWSM & the Wflow reservoir module . . . 49

5 Discussion **51**
5.1 Data selection . . . 51
5.1.1 Precipitation . . . 51
5.1.2 Potential evaporation . . . 52
5.1.3 Discharge . . . 53
5.1.4 Reservoir water storage . . . 54
5.2 Hydrological models . . . 54
5.2.1 HBV . . . 54
5.2.2 Wflow . . . 55
5.2.3 Hydrological models compared . . . 56
5.2.4 Validation . . . 57
5.3 Reservoir water storage . . . 58
5.3.1 Reservoir water storage model performance . . . 59
5.4 Lessons learned and improvement ideas . . . 60

6 Conclusion & Recommendations **61**
6.1 Conclusion . . . 61
6.2 Recommendations . . . 62

7 Appendix **68**
7.1 Water levels & Height-Volume relations . . . 68
7.1.1 Water levels . . . 68
7.1.2 Height Volume relation . . . 68
7.2 Parameter ranges HBV model . . . 68
7.3 Equation of HBV model . . . 70
7.4 Wflow equations . . . 70
7.5 Length and crop coefficient of each crop type . . . 71
7.5.1 Length of growth stages in days . . . 71
7.5.2 Kc value . . . 71
7.6 Precipitation correlation graphs . . . 71

List of Figures

1 Overview map of Upper White Volta basin with measurement locations of precipitation, potential evaporation, and discharge. . . . 4
2 Climate zones in Burkina Faso, where 600 and 900 are annual rainfall in mm 5
3 Example landscape of both climate zones 5
4 Average monthly precipitation and potential evaporation at Ouagadougou 6
5 (a) DEM and river stream pattern of the study area & (b) Land use map of study area . . 7
6 Precipitation temporal coverage overview with the names of all measurement locations and the corresponding locations. White indicates that there is no data available 9
7 Daily and monthly average potential evaporation (2018-2020) for all potential evaporation products. The legend provides the locations of the local measurement, for the gridded products the grid cell located at Ouagadougou is plotted 10
8 Potential evaporation temporal coverage overview (b) and locations (a). The names represent the location where the measurement is taken. 11
9 Annual average potential evaporation for ERA5 (a), GLEAM_a (b), GLEAM_b (c), GAIPE (d) and CRU (e) on the same scale. 11
10 Annual average potential evaporation for ERA5 (a), GLEAM_a (b), GLEAM_b (c), GAIPE (d) and CRU (e). Note that in all graphs yellow is high and blue is low but the scale is not the same. 12
11 Discharge temporal coverage overview (b) and locations (a). The names represent the location where the measurement is taken. 13
12 Discharge time series of the three discharge measurement locations within the catchment . 13
13 Water storage over time in the reservoirs Toece and Loumbila 14
14 Water surface area over time in the reservoir Toece 14
15 Overview map of the Toece reservoir and its catchment 15
16 Overview of the complete model structure 16
17 Schematisation of the HBV model structure 19
18 Overview of the different processes and fluxes in the Wflow model. 21
19 Relation between precipitation and effective precipitation (Allen et al., 1998) 25
20 Correlation between CHIRPS, ERA5 and the local measurements at the north (a), middle (b) and south (c) of the study area. 29
21 Results of the IDW interpolation of the local measurements with a) one day example and b) one year example 30
22 Spatial variation in absolute and relative differences of ERA5, CHIRPS and the local measurements (Local measurement - ERA5 or CHIPRS) 31
23 Temporal variation in absolute and relative differences of gridded data and the local measurements (Local measurement - gridded data) over time. The red box in figure (c) and (d) indicates the relevant period. 32
24 Effect of the combination of ERA5 and local measurements at (a) Ouagadougou, (b) Ouahigouya and (c) annual average in study area 33
25 Discharge in m^3/s and mm/d at Ramsa (a & c) and Yilou (b & d) for the modelled time period 34
26 Discharge out of reservoir Toece based on Yilou and spillway equation, where (a) show the results for peak calibration and (b) for timing calibration. 34
27 Effect of sedimentation on the reservoir volume in Toece 35
28 Discharge HBV simulation 36
29 Discharge Wflow simulation 36
30 Hydrograph of the discharge simulations per year 38
31 Discharge HBV validation 39
32 Discharge Wflow validation 39
33 Budyko Framework 40
34 Sensitivity analysis, performance ranges of the HBV model 41
35 10% variation parameter effect 42
36 Wflow results with increase Rooting depth 42
37 Wflow results with increase Rooting depth and increase porosity 43

38 Wflow results with decrease Rooting depth and increase porosity 43
39 Wflow results with decrease infiltration capacity 44
40 Influence of different cyclic infiltration capacities 45
41 Average crop water demand per month 46
42 Legend of the fluxes graphs 47
43 Water storage (a) and fluxes (b) of the HBV_RWSM without constraint 47
44 Water storage (a) and fluxes (b) of the HBV_RWSM with constraint 47
45 Water storage (a) and fluxes (b) of the HBV_RWSM with constraint and fraction of the Evaporation and Crop water demand 48
46 Water storage (a) and fluxes (b) of the Wflow reservoir module 48
47 Reservoir water storage simulations of both HBV_RWSM and Wflow reservior module . . 49
48 Comparison of reservoir outflow of both reservoir storage models and observed data 50
49 Annual values of the precipitation and potential evaporation which are used as input for the model 53
50 Observed water storage in reservoir Toece during the period of 2011-2016 59
51 Water level over time in Toece and Loumbila 68
52 Height volume relations of the reservoirs Toece and Loumbila 68
53 Heigth/Area relation of the Toece reservoir 69
54 Equations behind the HBV model for each reservoir 70

List of Tables

1 Overview of the different potential evaporation datasets which are available 12
2 Parameter and ranges used for calibration of the HBV model 19
3 Wflow calibration parameters 21
4 Differences between HBV_RWSM and the Wflow reservoir module 27
5 Cumulative discharge 35
6 NSE of reservoir water storage with different warm-up periods 49

1 Introduction

1.1 Background

Burkina Faso is a landlocked country in West Africa, in the Sahelian region. The country has a semi-arid climate, meaning that it is typically dry, sunny, and hot during most time of the year which is alternated with an intense rainfall period around August. This causes a water deficit, even though the annual precipitation is not particularly low (Ibrahim et al., 2012)(Dai et al., 2004). The water deficit limits the availability of food and water and constrains economic growth (UNDP, n.d.) (Fowe et al., 2015). The effects of the water deficit contributes to low the economic figures of Burkina Faso which is one of the poorest countries in the world, with 41 % of the population living below the poverty line (in 2018). The agriculture sector contributes to 35% of the country's GDP and accounts for 80% of employment (INSD, 2013).

Hundreds numbers of small local reservoirs have been constructed to increase the water availability during the dry period (Gerald et al., 2019). Water stored in these reservoirs is used for irrigation to supplement rain-fed and dry-season agriculture and for domestic purposes (Annor et al., 2009 & Cecchi et al., 2008 & Cecchi et al., 2009 Cecchi et al., 2008). These small reservoirs were not built at the same time and were constructed by different agencies (Liebe et al., 2005), which leads to a lack of overall management efficiency. Together with poor monitoring and institutional support, this results in a lack of information on the capacity, operational management, and limited maintenance (Liebe et al., 2005).

An important current disadvantage of small and shallow reservoirs is that they have a relatively high surface/volume ratio, which makes them more susceptible to evaporation. A case study in southern Burkina Faso showed that 60% of the water in the reservoir is lost to evaporation (Fowe et al., 2015). These high evaporation rates are caused by the high temperatures, which are projected to slightly increase according to IPCC scenarios (Masson-Delmotte et al., 2021).

In addition, siltation, aggravated by land erosion, has caused the reservoirs to become shallower and has decreased their storage capacity (Schmengler, 2011).

The combination of shallow reservoirs, climate change, and increasing water demand will increase water scarcity for millions of people (Schmengler, 2011). The problem of water scarcity can already be seen in the fluctuations in crop prices. For example, onion prices are 6 times higher during the dry season than during the wet season (Viard-Cretat et al., 2019).

1.2 Problem statement

To increase the water availability in and around these reservoirs the RVO's Department of International Development started the Develop to Build (D2B) program. As part of this program, Witteveen+Bos has been granted the project to rehabilitate nine reservoirs in Burkina Faso ((DWS), 2023) (Neterlands-Agency-Enterprice, 2023). However, due to the unsafe situation in Burkina Faso, only four of the nine reservoirs are investigated.

Creating a reliable hydrological model that can simulate discharge and reservoir fluctuations remains notoriously challenging, particularly in data scarce regions (Spence, 2010 & Davies and Beven, 2015 & Grande et al., 2022). Not only is each catchment different, but also varying discharge responses to precipitation have been observed to depend on season or antecedent wetness conditions (Meerveld and McDonnell, 2006). This study aims to gain insight into the modelling of reservoir dynamics in data scarce

semi-arid regions. To model reservoir dynamics several aspects are important such as which data and model to use, and how to validate the model outcome. This leads to the following research question and associated sub-questions:

How to accurately simulate reservoir dynamics using a hydrological model in data scarce semi-arid regions? – A case study for the Nakambé catchment

1. **What methods can be used to select most reliable model forcing data?**
   The first step to setting up a good hydrological model is to collect reliable input data. In this thesis, the forcing is defined as daily precipitation and potential evaporation, using discharge as the calibration data. Local data is available, however, it contains many gaps and the quality of the measured values is unknown. Remote sensing and reanalysis data are additional options for data that can be used in data scarce regions, like the Nakambé catchment. This sub-question explores various data comparison methods to select the best forcing data as input for the models.

2. **Comparing a lumped and a distributed hydrological model; which simulates the discharge best?**
   In this study, two different conceptual models are tested, namely a lumped model (HBV) and a distributed model (Wflow_SBM). Addressing this question provides insight into the underlying mechanisms of each hydrological model and why a certain outcome is simulated. Additionally, it also provides a better understanding of the assumptions and uncertainties within the models. Furthermore, the performance of both hydrological models is compared.

3. **What is the sensitivity of both hydrological models to parameter choices?**
   In both models, assumptions are made and a different way of assessing the parameters is performed. The effect of these choices and parameterization is crucial for their output. The effect of these different choices and assumptions is tested by performing a sensitivity analysis. This gives insight into the impact and thus importance of the choices made.

4. **What is the sensitivity of the different water balance models and which reservoir water storage model simulates reservoir dynamics best?**
   The water balance equation is used to model reservoir dynamics. The Wflow model includes a built-in reservoir module based on the water balance, whereas, for the HBV model, a dedicated water balance model needs to be developed. Although both models use the water balance equation, there are differences in the definition of some of the components. This sub-question investigates which components of the water balance have the most impact on modelling the reservoir dynamics.

Combining all four sub-questions gives insight into the overall process of the selection of data, the calibration, the validation, and the sensitivity of a water balance model.

This conclusion is obviously based on the Nakambé catchment only. However, this study gives insight into the process of how this is achieved, which leads to recommendations for other catchments.

1.3 Report structure

The remaining sections of this report are structured as follows. Chapter 2 provides an overview of the study area, followed by a description of the available data. Chapter 3 outlines the methodology, which is structured according to the sub-questions. First, it starts with an explanation of the different comparison

methods for data selection, followed by a description of both models and how discharge and reservoir fluctuations are modelled. It ends with an explanation of the sensitivity analysis and how the performance of the models is assessed. Chapter 4 presents the results according to the order of the previous chapter. These results are then interpreted and put into perspective in Chapter 5. The final chapter, Chapter 6, contains the key findings of the case study and provides recommendations for improvements within the Nakambé case study and for other catchments.

2 Study area & Data

This chapter first provides background information, such as the climate and morphological characteristics of the study area, and then gives an overview of the data used. Three types of data are used; local measurements, satellite products, and reanalysis data. The spatial and temporal performance of each data source is discussed. Figure 1 gives an overview of the study area, the locations of the reservoirs and their catchments, and the locations of the precipitation, potential evaporation, and discharge measurement stations.

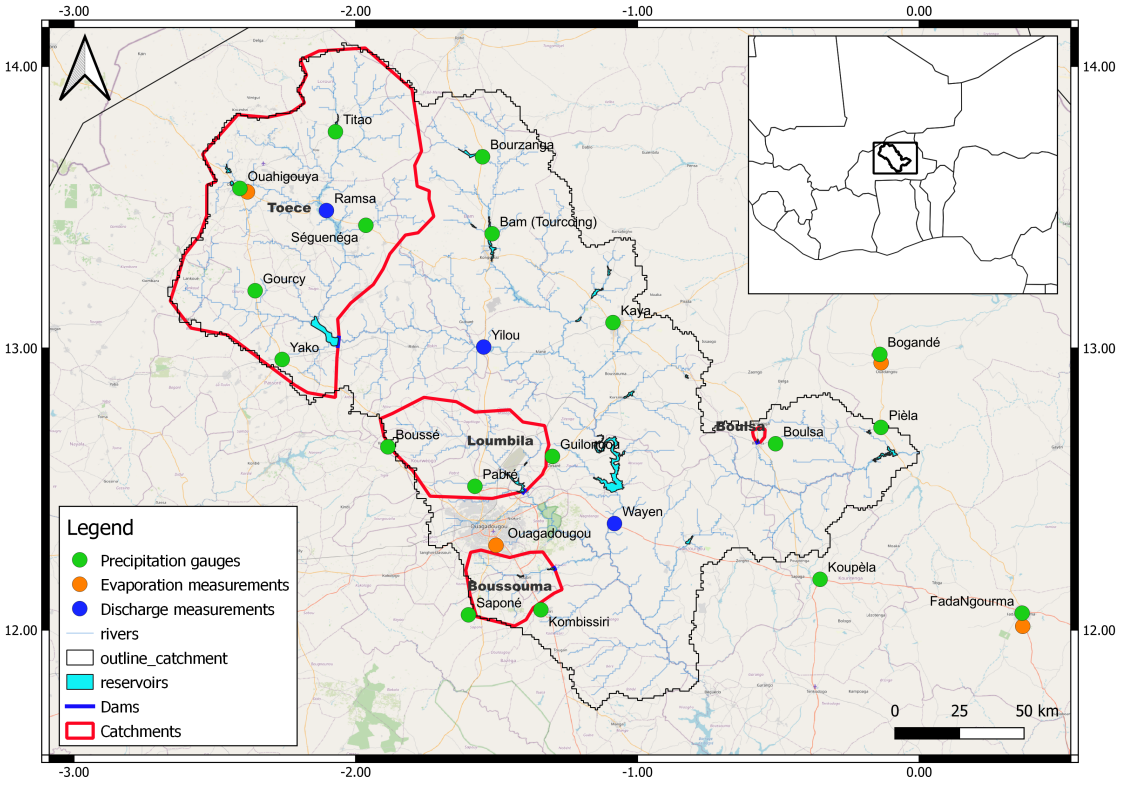


Figure 1: Overview map of Upper White Volta basin with measurement locations of precipitation, potential evaporation, and discharge.

The data can be divided into four categories: static and dynamic input data, and validation/calibration data. The static input data, which remains constant throughout the whole modelling period, will be discussed first. Next, the dynamic input data, data which varies over time and is used as input for the models, will be addressed. Last, the validation and calibration data, discharge, and reservoir water storage data are delved into. This section only gives an overview of all the different data that is used, the choice of which data to use as input for the models will be discussed later.

2.1 Study Area

The study area is the Nakambé River, also known as the White Volta, which is part of the Volta basin. The Nakambé River is an ephemeral river that originates in the North of Burkina Faso and flows Southwards to Ghana. The Nakambé river flows through two climate zones, in the north in the Sahelian zone and in

the middle and south of the catchment in the Sudano-Sahelian zone, as shown in Figure 2 (Ibrahim et al., 2012).

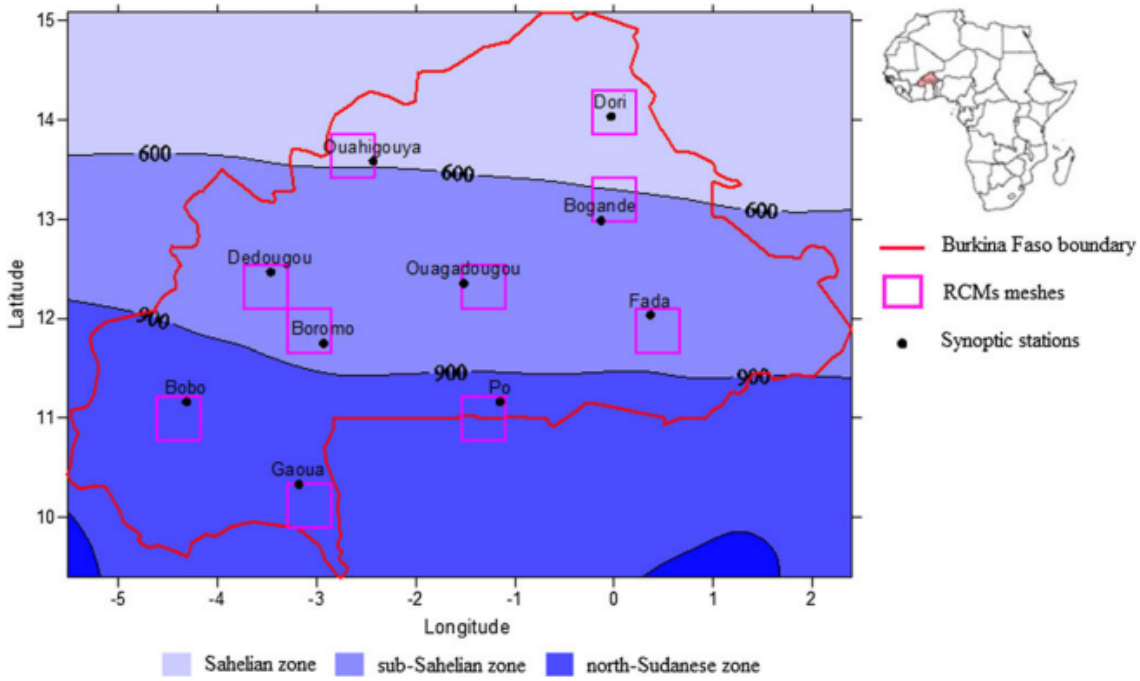


Figure 2: Climate zones in Burkina Faso, where 600 and 900 are annual rainfall in mm

The Sahelian zone is the transition zone between the Sahara and the Savanna. There is a semi-arid climate, which is characterized by a drought period of almost 8 months and an annual rainfall that typically varies between 400 and 600 mm. The climate is typically hot, sunny, dry, and relatively windy all year long (Ibrahim et al., 2012). The vegetation is adapted to these conditions and consists of sparse grasses, shrubs, and acacia trees.

The Sudano-Sahelian zone (called the sub-Sahelian zone in figure 2) is the transition zone from the Sahelian to the Sudanian climate zone. The Sudano-Sahelian zone is typically wetter than the Sahelian zone, with annual rainfall ranging from 600 to 900 mm. The conditions are slightly more favorable for agriculture (Ibrahim et al., 2012). Figure 3 shows an example of how these two climate zones typically look.

(a) Sahelian zone **(b)** Sudano-Sahelian zone

Figure 3: Example landscape of both climate zones

As mentioned above, both climate zones have a distinct dry and wet season. Figure 4, which shows the average monthly precipitation in Ouagadougou, clearly shows that the wet season is around August. Next

to precipitation, figure 4 also shows the potential evaporation. The potential evaporation is almost always (significantly) larger than the precipitation, indicating a water-limited system (Grande et al., 2022).

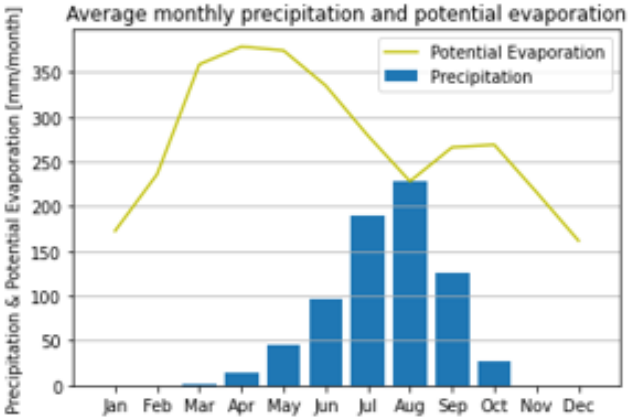


Figure 4: Average monthly precipitation and potential evaporation at Ouagadougou

2.1.1 Sahelian hydrological paradox

Understanding the hydrological principles within the study area is of great importance to set up hydrological models. Within the Sahelian region, two paradoxical hydrological processes have been noticed over the last five decades that may affect model performance. The first paradox occurred during the 'Great Drought' of the 1968-the 1990s when runoff increased significantly while precipitation decreased (Wendling et al., 2019). The second paradox, occurred with the recovery of the precipitation, during which the runoff coefficient continued to increase, even though a re-greening trend took place in the Sahel (Descroix et al., 2013). Descroix et al., 2018 explains the first paradox as mainly driven by land use/land cover changes and the second paradox as caused by both land use/land cover changes and climate evolution.

2.2 Static input data

Static input data are the topography and land use. Figure 5 shows both the topography and land use maps. The topography is assumed to be constant, as morphological changes are typically very slow and therefore no significant changes are expected during the modelling period. Land use is assumed to be constant as well. However, in reality, land use is not constant throughout the 20 year modelling period. As mentioned in section 2.1.1, land use changes over time, as this is one of the explanations for the Sahelian paradoxes. However, this assumption is made to keep the modelling of the hydrology of the Nakambé catchment feasible.

2.2.1 Topography

An accurate representation of the topography is essential for hydrological modelling as it directly influences the flow patterns and water movement within a catchment. In this research, the Shuttle Radar Topography Mission (SRTM) Version 4 Digital Elevation Model (DEM) dataset developed by Jarvis, A., 2008 has been used as the source for topographic information. The SRTM v4 DEM provides comprehensive coverage of global terrain data with a spatial resolution of approximately 90 meters. The dataset has been widely used and validated in various hydrological studies, ensuring its reliability and suitability for hydrological

modelling applications (Yamazaki et al., 2017). The STRM v4 DEM dataset was obtained in 2000, the beginning of the modelling period.

2.2.2 Land use

The selection of an appropriate land use dataset is crucial for the accurate representation of the spatial distribution of land cover and land use. In this thesis, the GlobCover 2009 product is selected as the land use input for the hydrological model (European Space Agency (ESA), 2009). GlobCover 2009 is a global land cover dataset derived from satellite imagery acquired by the Medium Resolution Imaging Spectrometer (MERIS) sensor aboard the European Space Agency's Envisat satellite. It provides detailed information on land cover classes at a resolution of 300 meters, making it suitable for capturing the variability of land use patterns at a regional scale.

The use of the GlobCover 2009 dataset offers several advantages for incorporating land use information into the hydrological model. Firstly, the dataset covers a wide range of land cover classes, including forests, croplands, grasslands, urban areas, water bodies, and bare soil. This comprehensive coverage enables the representation of different land use types and their respective hydrological characteristics, allowing for a more realistic simulation of the hydrological processes at the catchment scale. In addition, the dataset has undergone extensive validation and quality assessment to ensure its reliability and accuracy for hydrological applications. The use of a validated and widely accepted land use product such as GlobCover 2009 enhances the credibility of the model results (Congalton et al., 2014 & Tsendbazar et al., 2015).

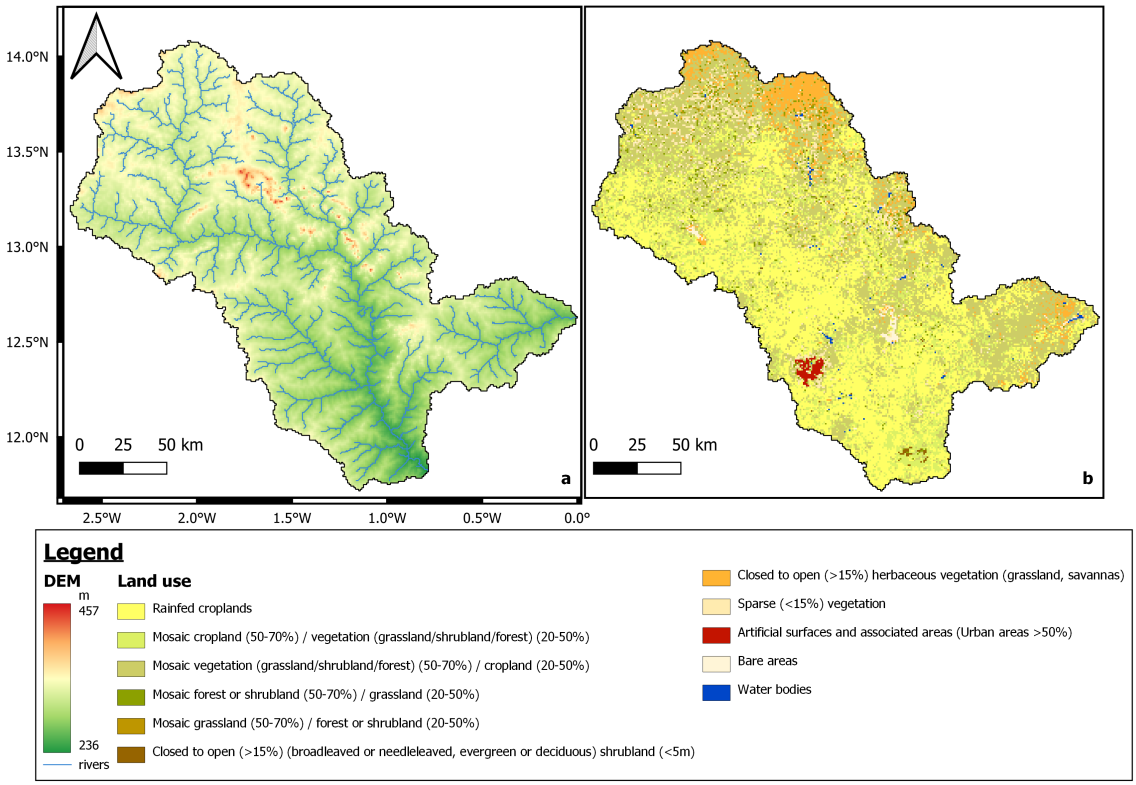


Figure 5: (a) DEM and river stream pattern of the study area & (b) Land use map of study area

2.3 Dynamic input data

2.3.1 Precipitation

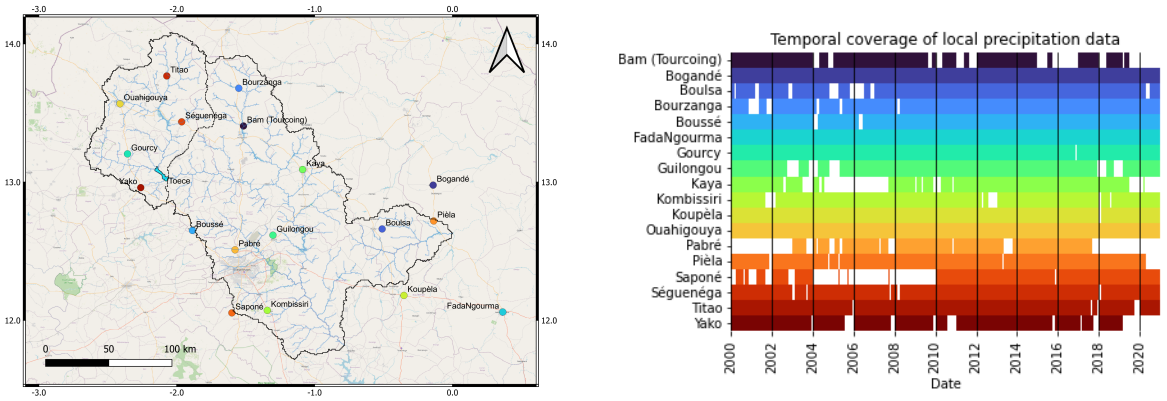
As shown in Figure 1, the Upper White Volta basin has a relatively dense network of precipitation observation stations, which is unique in developing countries like Burkina Faso. In addition to local measurements, satellite-ground (CHIRPS) and climate reanalysis (ERA5) data are available. ERA5 and CHIRPS have a grid size of 0.25° (27.75 km) and 0.05°(5.55 km) respectively. Both ERA5 and CHIRPS contain daily values ranging from 2000 to 2021.

CHIRPS (Climate Hazared Groep InfraRed Precipitation with Station data) is obtained by combining satellite, local observations, and climate model data. CHIRPS data predicts more events and less precipitation per event. This phenomenon is caused by the averaging effect, where the intensity of individual events is smoothed (Funk et al., 2015).

ERA5 is a reanalysis product that combines a large number of historical observations into global estimates using advanced modelling and data assimilation systems (ECMWF, n.d.). ERA5 calculates the precipitation as a by-product of the model behind ERA5. According to Lavers et al., 2022, ERA5 cannot model the highest observed precipitation totals, but it can generally capture their locations and patterns. Furthermore, they found that ERA5 performs less in the intertropical convergence zone during the summer period. Notably, the study area is located within this intertropical convergence zone. This geographical context adds significance to our assessment of ERA5's performance.

Local measurements can be considered more reliable, although they also do have uncertainties. Local measurements may be more reliable, but they do have time gaps, as can be seen in figure 6b, which shows the time coverage of all the measurement locations. Another disadvantage of the local measurements is the fact that they are point measurements rather than a grid covering the whole catchment, which is the case for the other datasets. To obtain a spatial coverage of the precipitation, an interpolation is needed, which creates errors in the spatial representation. Especially in areas like Burkina Faso, where local showers are a common phenomenon according to Dembélé, 2023. These local showers are not captured correctly with an interpolation distance of approximately 50 km.

To summarize, the advantage of the satellite and reanalysis data over local measurements is the spatial and temporal coverage of the data, but local measurements can be considered as more reliable data. Different ways of dealing with the gaps can used to get the most reliable input data in the report, the methods for selecting the best data are explained in section 3.1.1.



(a) **(b)**

Figure 6: Precipitation temporal coverage overview with the names of all measurement locations and the corresponding locations. White indicates that there is no data available

2.3.2 Potential evaporation

Just like precipitation, there are local observations and satellite/reanalysis (gridded) datasets available for potential evaporation. Potential evaporation has a less dense network of local measurement data points, as can be seen in figure 1. However, potential evaporation is less spatially variable than precipitation, which makes interpolation more realistic. The local measurements are obtained using the Pan evaporation measurement.

The available satellite and reanalysis products are ERA5, GLEAM_a, GLEAM_b, CRU, and GAIPE. The various available satellite and reanalysis datasets differ substantially in terms of absolute values and their spatial distribution when compared to each other and the local measurements. Three figures are used to show the extent of this variability among the potential evaporation data products.

Figure 7 presents the daily and monthly averaged potential evaporation of the last three years (2018-2020) of the available data products. The local measurements are named after the location where it is measured, namely Ouagadougou, Ouahigouya, Bogandé, and Fada Nguorma. It clearly illustrates that both GLEAM products (respectively the brown and pink line) and ERA5 (purple line) constantly observe lower potential evaporation rates than the local measurement at Ouagadougou (blue line).

Figure 9 and Figure 10 provide further insights by showcasing the average annual potential evaporation across the entire study area. In Figure 9, each plot is uniformly scaled, which clearly shows the differences between the gridded products. This matches with the observation from Figure 7, that both GLEAM products and ERA5 observe lower values than GAIPE and CRU.

In contrast, Figure 10 introduces a varying scale among different products to emphasize the spatial pattern perceived by each gridded product. In those plots, the color scheme is standardized, with yellow displaying high values and blue representing low values. Figure 7 shows that GAIPE and ERA5 show similar spatial variance with high potential evaporation rates in the northern region and lower rates in the southern part of the study area. In contrast, both GLEAM products show an inverse pattern relative to GAIPE and ERA5 with higher potential evaporation rates in the southern region of the study area. CRU, on the other hand, shows a more east-west variance in the potential evaporation rates within the study area.

Overall, these figures collectively illustrate the significant differences between the potential evaporation data products. The following paragraphs will delve deeper into the gridded products, explaining how each product obtains the potential evaporation and where the differences come from.

ERA5, GLEAM_a, and GLEAM_b all have relatively low annual potential evaporation values but do have good temporal and spatial coverage. GLEAM (Martens et al., 2017) versions 3.7a and 3.7b are used. The difference between a and b is that a is based on satellite and reanalysis data (MSWX net radiation and air temperature) and b is based on satellite data only. As mentioned above, ERA5 is a reanalysis product and uses the Penman-Monteith equation to calculate the potential evaporation. GLEAM uses the Priestley-Taylor equation, which is a simplification of the Penman-Montheith equation. The Priestley-Taylor equation introduces α, which represents an exponential decrease in evaporation as soil water content decreases (Flint and Childs, 1991 & Mercer, 2018). Thus, GLEAM is also dependent on the soil moisture content, which partly follows the amount of precipitation. Therefore, the α could explain why the north-south pattern varies between the different products, as described in Table 1.

CRU (Climate Research Unit) (Harris et al., 2014) and GAIPE (Global Aridity Index and Potential Evaporation) (Trabucco and Zomer, 2019) have reasonable annual values and good spatial coverage. Both products are based on the Penman-Monteith equation but in the CRU dataset, the local vegetation conditions are added to the Penman-Monteith equation. The disadvantage of both is that neither has daily values, CRU has monthly averages and GAIPE is a climate dataset that only contains the 30 year average annual potential evaporation.

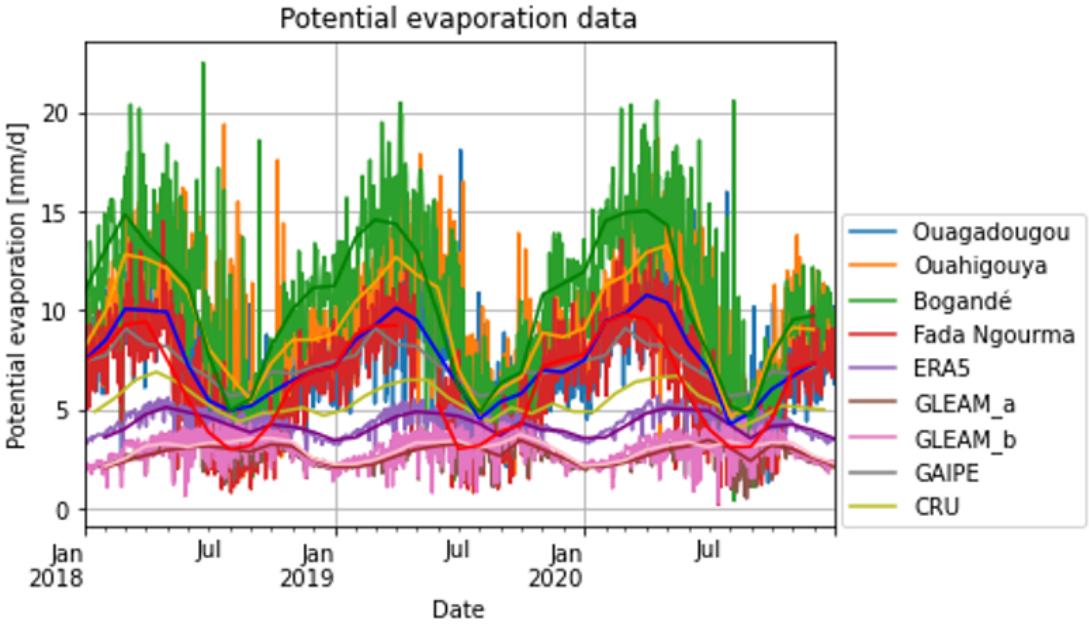
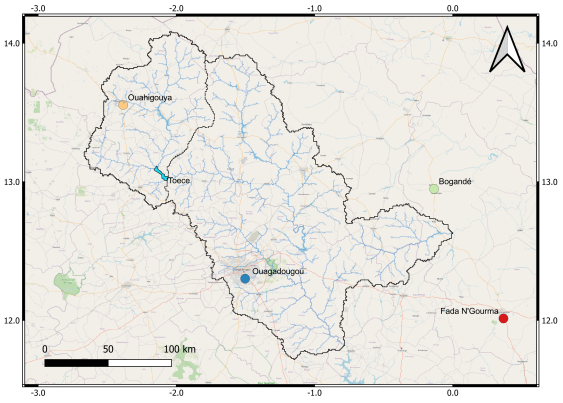
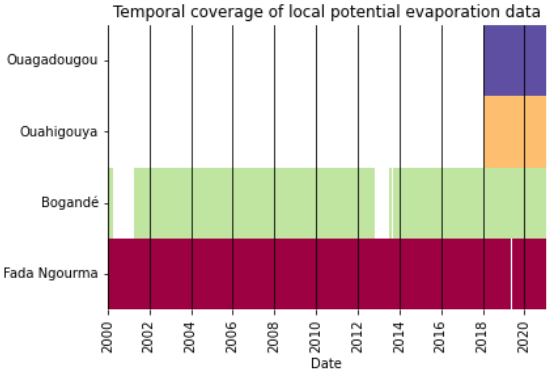


Figure 7: Daily and monthly average potential evaporation (2018-2020) for all potential evaporation products. The legend provides the locations of the local measurement, for the gridded products the grid cell located at Ouagadougou is plotted

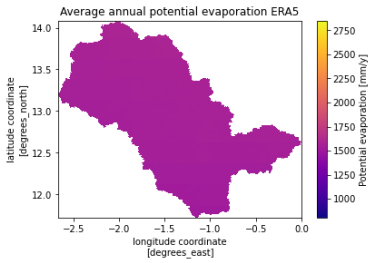


(a)

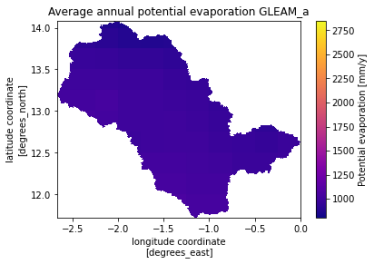


(b)

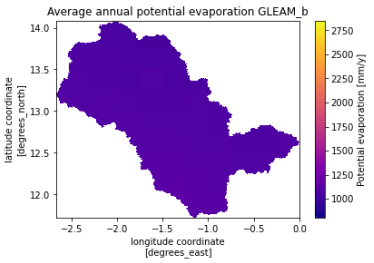
Figure 8: Potential evaporation temporal coverage overview (b) and locations (a). The names represent the location where the measurement is taken.



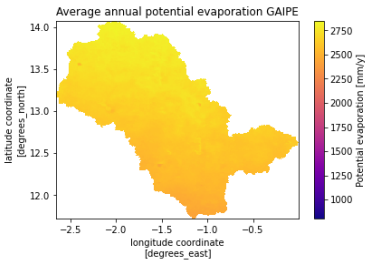
(a) ERA5



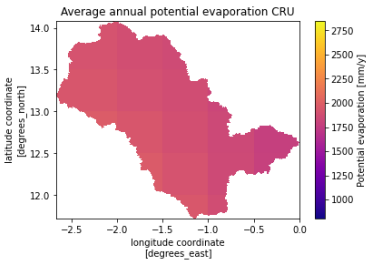
(b) GLEAM_a



(c) GLEAM_b

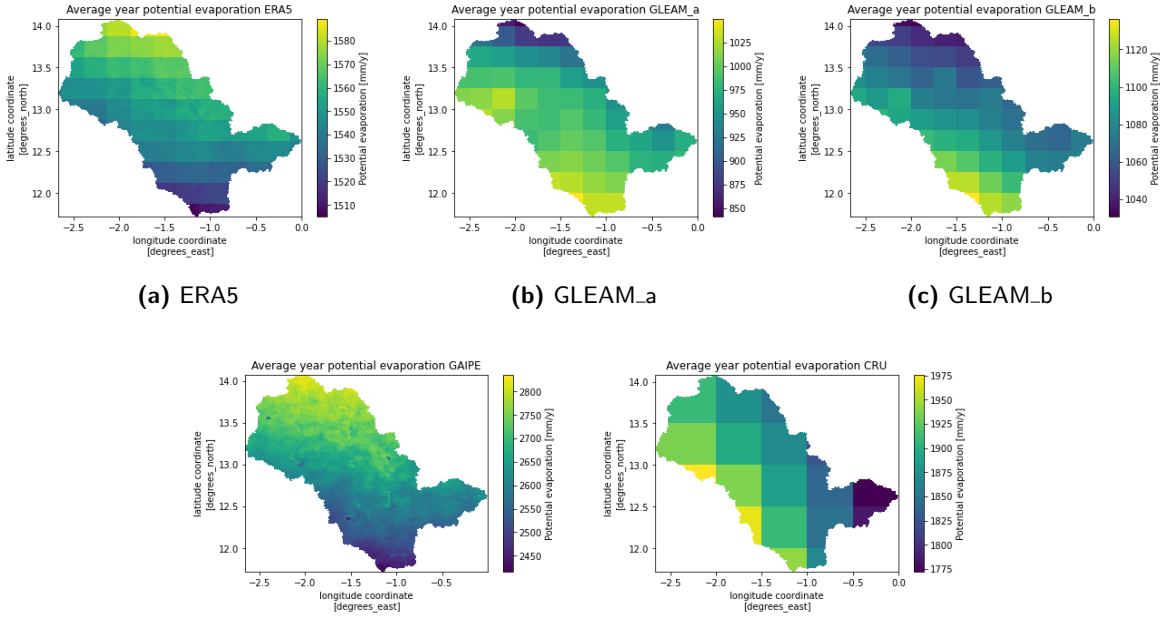


(d) GAIPE



(e) CRU

Figure 9: Annual average potential evaporation for ERA5 (a), GLEAM_a (b), GLEAM_b (c), GAIPE (d) and CRU (e) on the same scale.



(a) ERA5 **(b)** GLEAM_a **(c)** GLEAM_b

(d) GAIPE **(e)** CRU

Figure 10: Annual average potential evaporation for ERA5 (a), GLEAM_a (b), GLEAM_b (c), GAIPE (d) and CRU (e). Note that in all graphs yellow is high and blue is low but the scale is not the same.

To provide an overview of all available data, Table 1 summarises the temporal coverage, spatial resolution, spatial variance, and annual mean for all potential evaporation products. M. Dembélé indicated that at least a potential evaporation of 2000 mm per year can be expected and Amisigo, 2005 even suggest annual values of 2500mm. The local stations do measure values above this minimum annual indication given by M. Dembélé. However, as figure 8b shows, the temporal coverage of the local measurements is relatively low.

As described above, each dataset has its own advantages and disadvantages. How to obtain the best input for the hydrological models is explained in section 3.1.2

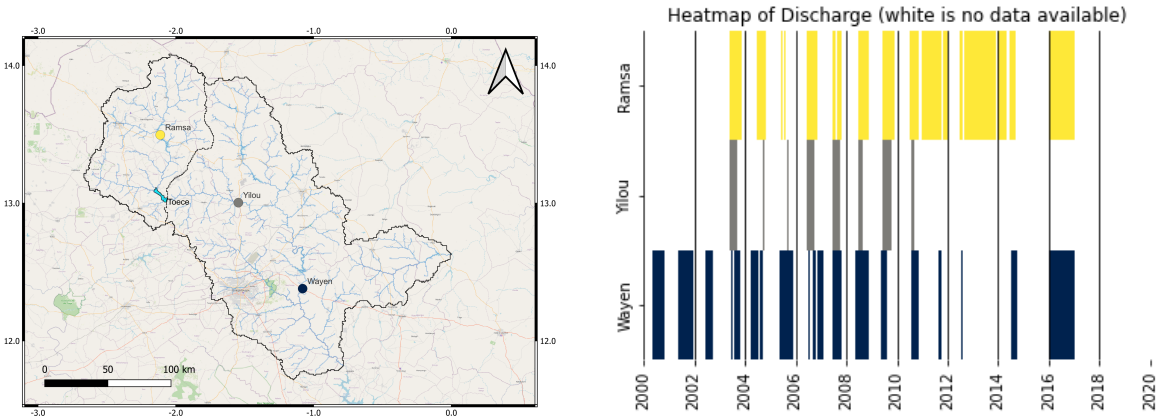
Data type	Product or location	Temporal resolution	Spatial resolution [km]	Spatial variance	Annual average [mm/y]
local stations	Ouagadougou	monthly, last 3 years daily	point	-	2726.4
	Ouahigouya	daily last 3 year	point	-	3378
	Bogandé	daily 20 year	point	-	3766
	Fada Ngourma	daily 20 year	point	-	2258
Reanalysis	ERA5	daily	27.75	high values south	1545
Satellite	GLEAM_a	daily	27.75	high values south	1010
	GLEAM_b	daily	27.75	high values south	1092
	CRU	monthly data	55.5	high values north	1871
	GAIPE	climate data	1	high values north	2443

Table 1: Overview of the different potential evaporation datasets which are available

2.4 Validation and Calibration data

2.4.1 Discharge

There are three discharge monitoring sites: Ramsa, Yilou, and Wayen, blue dots in figure 1. The temporal coverage overview in figure 11b shows that, all measurement stations lack measurements during the dry season and that during the wet season, the data is also very limited. The gaps can be explained by several factors; the data is not collected automatically but needs to be collected by local people, the water level is below the measurement station or there is no discharge at all, which is not reported as zero but as no value. As some of the rivers are intermittent streams or have very low discharges during the dry season, the local people do not see the importance of collecting the flow data. Therefore, it is not known if the measurement gap in the dry season is no available data or no discharge at all (Dembélé, 2023).



(a) **(b)**

Figure 11: Discharge temporal coverage overview (b) and locations (a). The names represent the location where the measurement is taken.

Not only are the time series of the discharges incomplete, but the values also raise some questions with regard to their quality. Figure 12, shows the daily measured discharge in m^3/s, for all three measure sites. It is expected that, by moving downstream, the discharge increases due to a larger catchment area contributing to the flow. However, the discharges at Ramsa and Yilou appear to be of comparable magnitude, despite Yilou being located more than 100 km downstream of Ramsa.

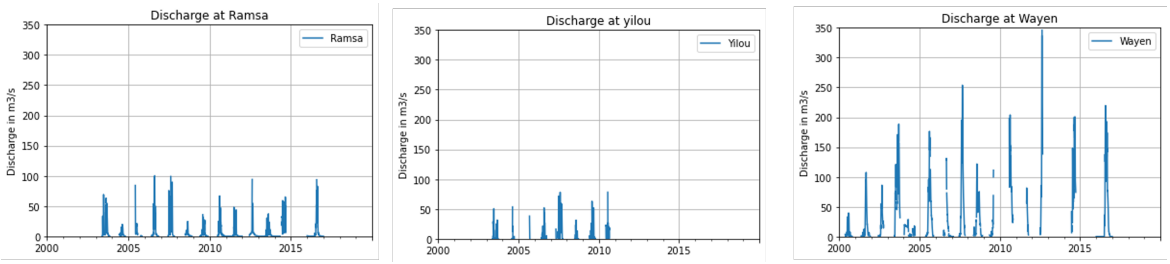


Figure 12: Discharge time series of the three discharge measurement locations within the catchment

2.4.2 Reservoir water storage

Water level measurements are available for two of the four reservoirs. No data is available for the reservoirs Boussoumma and Boulsa. However, according to satellite images from Global Surface Water (Pekel et al., 2016) and local information, the Boulsa reservoir completely dries out during the dry seasons. The two largest reservoirs, Toece and Loumbila, do not completely dry out during the dry season. Figure 13 shows the water storage over time in the reservoirs Toece and Loumbila and Figure 14 the water surface area of reservoir Toece. The Height-Volume & Height-Area relation and water levels, which are used to derive the water storage data, can be found in Appendix 7.1.

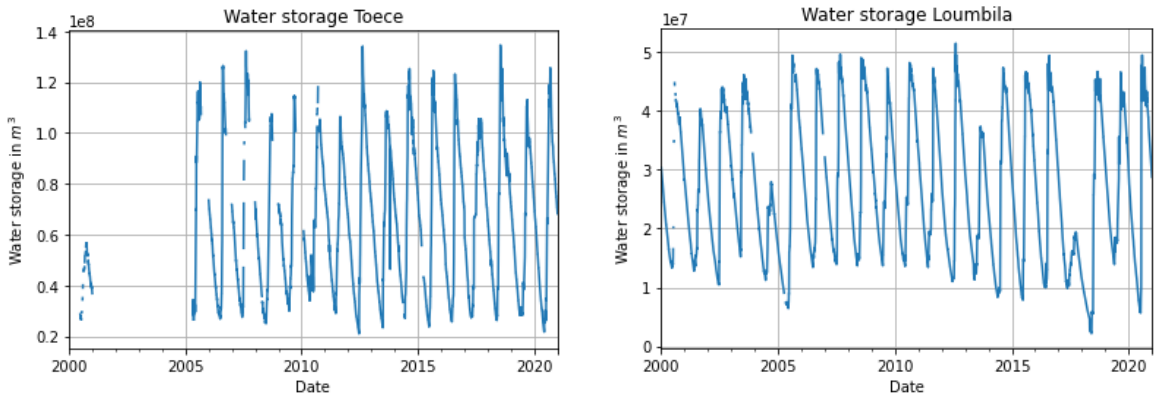


Figure 13: Water storage over time in the reservoirs Toece and Loumbila

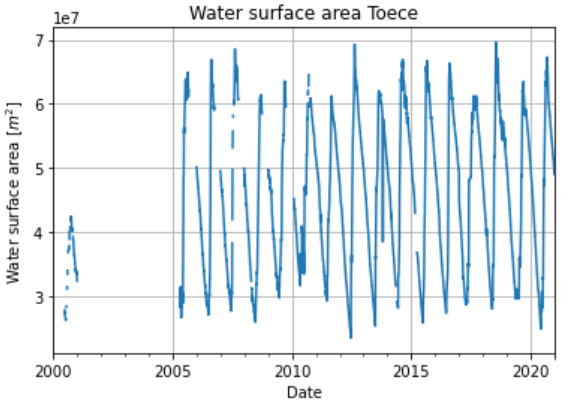


Figure 14: Water surface area over time in the reservoir Toece

2.5 Focus location & period

The primary objective of this study was to investigate the reservoir water storage of four different reservoirs, namely Toece, Loumbila, Boussouma, and Boulsa, within the study area. Figure (1) shows that only discharge measurements from upstream and downstream of the Toece reservoir are available. As the discharge time series are missing, it was impossible to validate the in and outflow of the reservoirs Loumbila, Boussouma, and Boulsa. As a result, the focus of the study shifted only to the Toece reservoir and its catchment. Figure 15 displays the Toece reservoir, its catchment, and the locations for all local

precipitation, potential evaporation, and discharge measurements.

Apart from reducing the study area, the investigation period is reduced as well. The Ramsa and Yilou discharge time series are essential for the Toece reservoir water storage modelling. Figure 10 shows most discharge measurements are available during the period from 2006 and 2011. Therefore, the modelled time frame is the five-year period between 2006 and 2011.

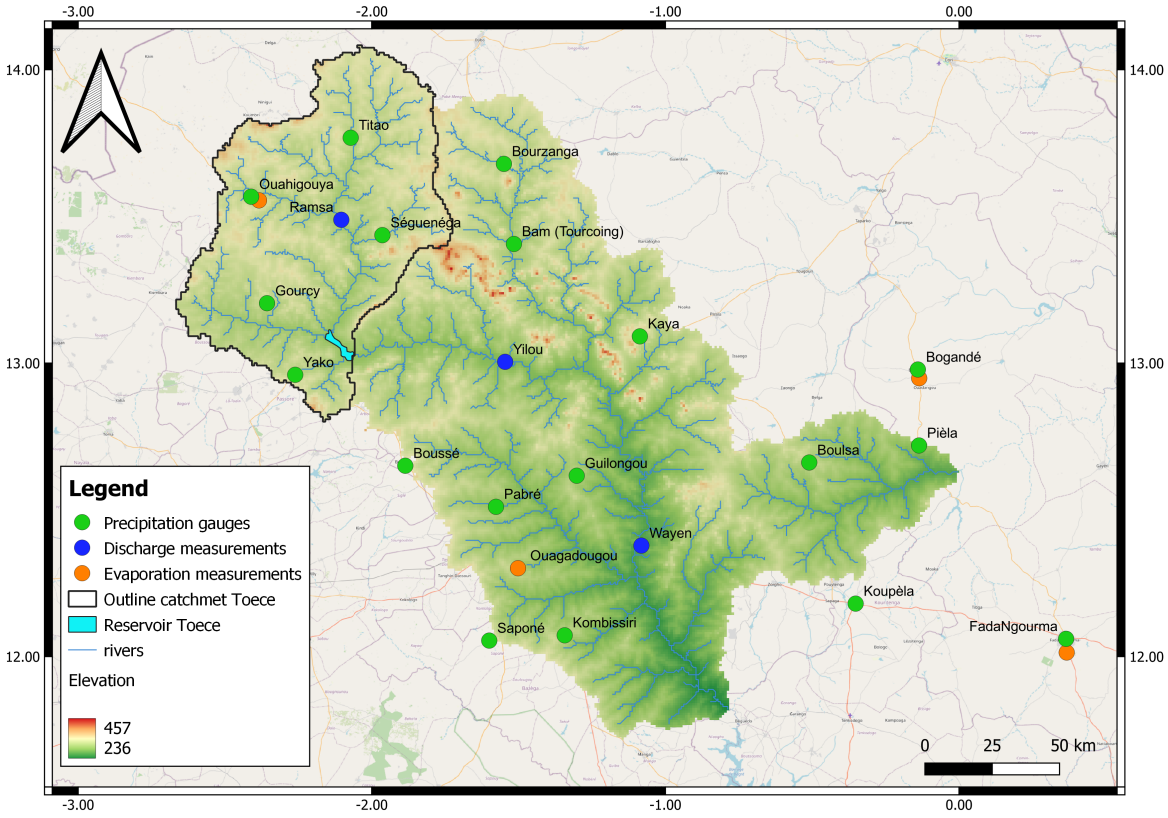


Figure 15: Overview map of the Toece reservoir and its catchment

3 Methodology

This chapter explains the methodologies applied. The four sections match the sub-question's subject description. Consequently, first, the data selection process is explained. This is then followed by elaborating on the two distinct models that are used for modelling the discharges, namely HBV and Wflow. Their mechanisms, principles, parameters as well as the calibration methods employed to calibrate both models are discussed in detail. Thirdly, it describes how the sensitivity analysis is executed. Last, two different water balance models used are described and all methods to obtain the components of the water balance will be discussed.

Figure 16 provides an overview of all the steps taken to achieve the final result, the reservoir water storage. A distinction has been made between the two model paths: the HBV model, which is always displayed in orange, and Wflow, which is always displayed in green. In addition to the overview, the figure also highlights some differences in the input used for the models. These differences will be explained in more detail in the upcoming chapter.

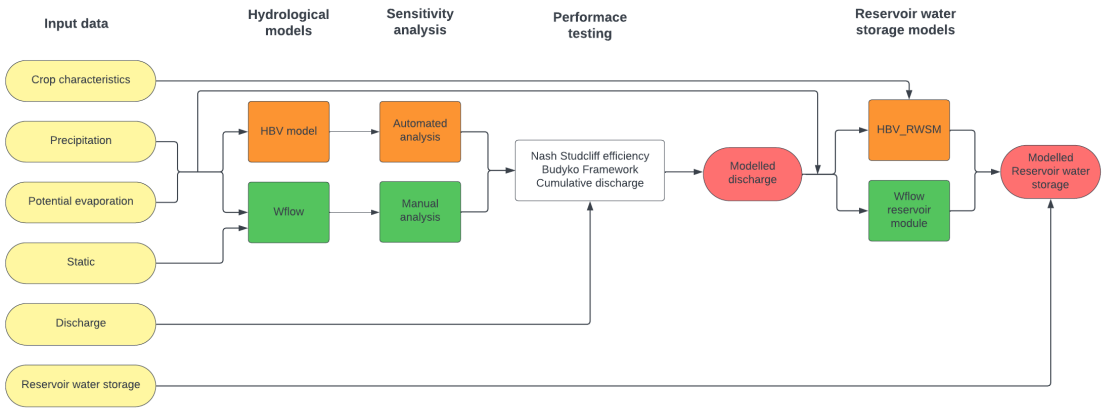


Figure 16: Overview of the complete model structure

3.1 Data selection

3.1.1 Precipitation

Section 2.3.1 presented three precipitation products: local measures, ERA5 and CHIRPS. Local measurements are point measurements and therefore need to be interpolated to obtain a spatial coverage of the entire study area. There are several interpolation methods, such as Thiessen polygons, Kriging, and Inverse Distance Weighing (IDW). In the Nakambé catchment, the precipitation pattern is characterized by a short, intense, and local pattern which makes interpolation difficult. In this study, IDW is chosen as the interpolation method, because the exponent in IDW allows the length of the spatial impact to be varied (Kurtzman et al., 2009). Thiessen polygons assign the same value to the whole polygon and therefore miss the effects of these local precipitations more than IDW. Kriging and IDW showed comparable results in terms of hydrological modelling (Ly et al., 2013), which allows us to favor the simple approach of IDW, see Equation 1. Another advantage of this method is that the spatial interpolation also fills the time gaps, which Thiessen polygons do not do, as they assign missing values to the whole polygon.

$$z = \frac{\sum \frac{z_i}{d(x, x_i)^p}}{\sum \frac{1}{d(x, x_i)^p}} \tag{1}$$

In which z is the predicted value , d the distance, x is the unknown point, x_i is the known point, z_i is the value of the known point, and p the power.

Point-to-pixel correlation

To compare the performance of ERA5 and CHIRPS, daily, monthly, and annual correlation plots are made for each location. Daily data are available and monthly and annual values derived from the daily data are also examined. The days in which the local measurements miss data are also excluded from ERA5 and CHIRPS.

Spatial difference

The spatial relative difference indicates where differences occur. This is based on the interpolated local measurement data. Using $p_r = (p_{local} - p)/p_{local}$ the relative difference between the local measurements and ERA5 and CHIRPS is calculated. The relative difference is used instead of the absolute difference because it also gives an indication of the performance of the values rather than only showing the differences.

Timing differences

Next to the spatial differences, it is also important to understand the temporal differences, as they determine the timing of flows in the river. The monthly timing is examined as the daily timing varies too much for a useful comparison for both ERA5 and CRHIPS. Precipitation is averaged over the whole catchment and the absolute and relative difference per month is calculated.

3.1.2 Potential Evaporation

As figure 7 and table 1, there are significant differences between the potential evaporation products. According to Dembélé, 2023, an annual potential evaporation of at least 2000 mm can be expected in the study area. As each product alone is not sufficient due to temporal gaps, low values, or poor spatial performance, a combination of ERA5 and the local measurements is made. The spatial and temporal pattern of ERA5 is scaled to the values of the local measurements, combining the strengths of these products to obtain the best input data. This is done by first performing an IDW interpolation on the local measurements to create a spatial representation of the local measurements. The following equation is then used to scale ERA5 to the local measurements:

$$PET = (PET_{ERA5)} - \overline{PET_{ERA5,t}}) \cdot \frac{\sigma(PET_{ground,t})}{\sigma(PET_{ERA,t})} + \overline{PET_{ground,t}} \tag{2}$$

Where PET is the potential evaporation, the mean and standard deviation is taken over the time of the modelling period.

3.1.3 Discharge

Discharge data is very scarce, as can be seen in figure 11b. Consequently, the main focus centers on the period spanning from 2006 to 2012, during which most data are available. Unfortunately, throughout the whole model period, data are absent during the dry periods. It is therefore assumed that discharge is zero or negligible during this period. It is acknowledged that especially, further downstream in the catchment

this assumption may not be valid, however, the constraint imposed by the restricted data availability makes this assumption necessary.

In contrast, although more data are available during the wet season, they also still contain gaps. In addressing these, two different approaches are employed for the calibration and validation of simulations. The first approach involves not calibrating and validating the gaps, thereby accounting for the absent data points. Alternatively, a straightforward linear integration strategy is employed as an alternative methodology.

Spillway discharge calculation

The reservoir water storage is controlled by a spillway in the dam of the reservoir. The spillway equation is used to calculate the discharge of water over the spillway. There are different spillway structures, the Toece reservoir is regulated with a horse-shoe spillway, where water flows over the crest. The spillway equation uses the water level to determine the discharge over the spill and is given by:

$$Q = m \cdot B \cdot \sqrt{2g} \cdot \Delta H^{\frac{3}{2}} \tag{3}$$

Where m is a constant for the structure shape which is 0.4 for the horse-shoe structure, B represents the width of the spillway 104.5 m, g the gravitational acceleration in m/s^2, and ΔH represents the difference between the water level and crest height in meters.

The equation requires measurements of the water level and crest height from the same reference height. As this is not the case here, the spillway equation becomes inapplicable. However, there is a workaround as discharge data is available at Yilou. Assuming a direct link between the start of flow at Yilou and flow over the spillway allows us to link a water level to the start of flow over the spillway. This method generates a validation discharge out of the reservoir.

3.1.4 Reservoir water storage

The bathymetry of the reservoirs is measured only once at the end of the original 20-year modelling period. Siltation is recognized as a serious environmental threat in Burkina Faso, therefore, it must be taken into account (Schmengler, 2011). A prior study, conducted by Witteveen+Bos, modelled the erosion and sediment accumulation in the reservoir. Annual sediment rates obtained from this study are used to account for the siltation over the years. The Toece reservoir has a modelled annual sediment inflow rate of 152000 $m^3/year$.

3.2 Discharge models

Two discharge models are used. The two models are a version of the Hydrologiska Byråns Vattenbalansavdelning (HBV) model and the SBM_Wflow model (hereafter referred to as Wflow). Firstly, the structure and parameters of the HBV model will be presented. Secondly, the same is done for the Wflow model. Lastly, the calibration method employed will be described.

3.2.1 HBV model

The HBV model is a lumped conceptual model concept; in this study, a variation of the HBV model based on Gharari et al., 2014 is used. The used type of HBV model consists of four buckets, each representing different concepts in the hydrological system (Fig 17). Each bucket represents different storage, namely

interception, unsaturated, fast (i.e. preferential flow and saturation overland flow), and slow reservoirs (i.e. groundwater). The HBV model uses precipitation and potential evaporation as forcing data and discharge as calibration and validation data. Figure 17 shows the model structure, which is further explained below. Appendix 7.3 goes into more detail about the equations behind the processes.

The precipitation enters through the interception reservoir (S_i), which has a maximum storage capacity (I_{max}) and represents interception by vegetation. Excess precipitation, known as effective precipitation (P_e), is distributed between the unsaturated (S_u) and fast reservoirs (S_s). The interception reservoir losses water through interception evaporation (E_i) which is dependent on the potential evaporation (E_p) and water stored in the interception reservoir.

The unsaturated reservoir (S_u) resembles the soil moisture content in the root zone and is parameterized by the maximum soil moisture capacity ($S_{u,max}$). The unsaturated reservoir receives a fraction of the effective precipitation, depending on the coefficient rho (ρ), which is determined by a power function with exponent β. The unsaturated reservoir supplies the slow reservoir (S_s) through percolation (Q_{us}). Percolation is a linear function of the available moisture in the unsaturated zone and the maximum percolation capacity (P_{max}). The other outflow from the unsaturated zone occurs through plant transpiration (E_a). Transpiration is assumed to be moisture-constrained, when the soil moisture content is lower than zero, no transpiration can occur. If the storage in the unsaturated zone is larger than zero the transpiration is determined by a function dependent on the potential evaporation and the available water in the unsaturated zone.

The effective precipitation that cannot be stored in the unsaturated reservoir flows into the fast reservoir. Q_f, which is determined by the coefficient k_f, and it accounts for the river discharge from the fast reservoir. The slow reservoir's contribution to the discharge is denoted as Q_s, determined by the coefficient k_s. Finally, a time-lag function is included based on the parameter $Tlag$.

In table 2 all eight model parameters are shown, in appendix 7.2 the chosen ranges are explained. The discharge data is used to calibrate the 8 parameters; this will be discussed further in section 3.2.3.

The advantage of the HBV model is that only precipitating and potential evaporation data are needed. The simplicity of the model also has its drawbacks, in particular concerning the areal representation, a fact which limits the use of distributed data. There are also physical shortcomings, such as the lack of an interception routine and the lack of elevation correction (Lindström et al., 1997).

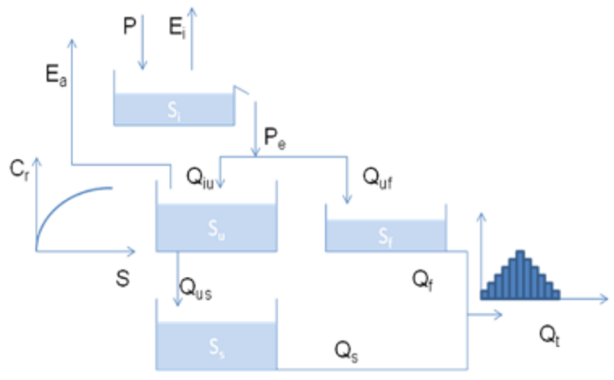


Figure 17: Schematisation of the HBV model structure

Parameter	Range
I_{max}	0 - 4 mm
C_e	0.001 - 1
S_{umax}	1 - 500 mm
$Beta$	0.2 - 2
P_{max}	0.001 - 4 mm/d
Tl_{lag}	0.001 - 2 d
k_f	0.0001 - 2 d
k_s	0.0001 - 1 d

Table 2: Parameter and ranges used for calibration of the HBV model

3.2.2 Wflow

SBM_Wflow (hereafter referred to as Wflow) is a distributed model which is developed by Deltares. Wflow is categorized as a fully distributed hydrological model as the data input, calculations, and output are grid-based (Seizarwati and Syahidah, 2021). The input data required by Wflow are the static data: Land use, topography (DEM), and Leaf Area Index (LAI) and the dynamic data precipitation, potential evaporation, and discharge. The SBM hydrological concepts consist of sub-models describing the water balance equation for each grid cell. The Wflow model simulates the runoff of each grid cell, and the discharge in the river is simulated with a kinematic wave function. Its structure is shown in figure 18. There are three main modules used to simulate the discharge: the rainfall interception module, soil module, and kinetic wave module (Deltares, 2023).

1. The rainfall interception module is based on an analytical approach described by Gash, 1979. The amount of water needed to saturate the canopy is defined as:

$$P = \frac{-RS}{E_w} ln[1 - \frac{E_w}{R}(1 - p - p_t)^{-1}] \tag{4}$$

   in which R is the average precipitation intensity on the saturated canopy, E_w the average evaporation from the wet canopy, p the canopy gap fraction which depends on the Leaf Area Index (LAI), p_t is 0.1 * p and S the canopy capacity.

2. In the soil module, the computation of water balances within the unsaturated and saturated zones is grounded on the TOPOG_SBM mode. Actual infiltration happens when there is storage in free surface water and remaining capacity in an unsaturated zone. The infiltration capacity depends on the input parameter describing the infiltration rate, which will be discussed later. The soil is considered as a bucket with a certain depth (z_t), divided into a saturated (S) and an unsaturated (U) store, expressed in units of depth. The top of the S store forms a pseudo-water table at depth z_i. The unsaturated store (U) is divided into storage (U_s) and deficit (s_d). All infiltrating precipitation first enters the U store. The transfer of water from the unsaturated to the saturated store (st) is controlled by the saturated hydraulic conductivity and is calculated according to:

$$st = K_{sat} \frac{U_s}{s_d} \tag{5}$$

   where K_{sat} is the saturated hydraulic conductivity. When the value of st is negative, the water will flow from the saturated zone to the unsaturated zone. However, horizontal water exchange also takes place between the hydrological units

3. A kinematic wave module is used to route water over the surface and through streams and rivers. For the calculation, this module implements the Manning equation.

A more elaborate explanation of the equations can be found in Appendix 7.4. Wflow allows the addition of different modules to suit better to the modelled location. There are modules such as a snow and glacier melt module, a simple reservoir, a lake module, and a crop growth modules available. Here, only the reservoir module is used and will be described in section 3.4.2.

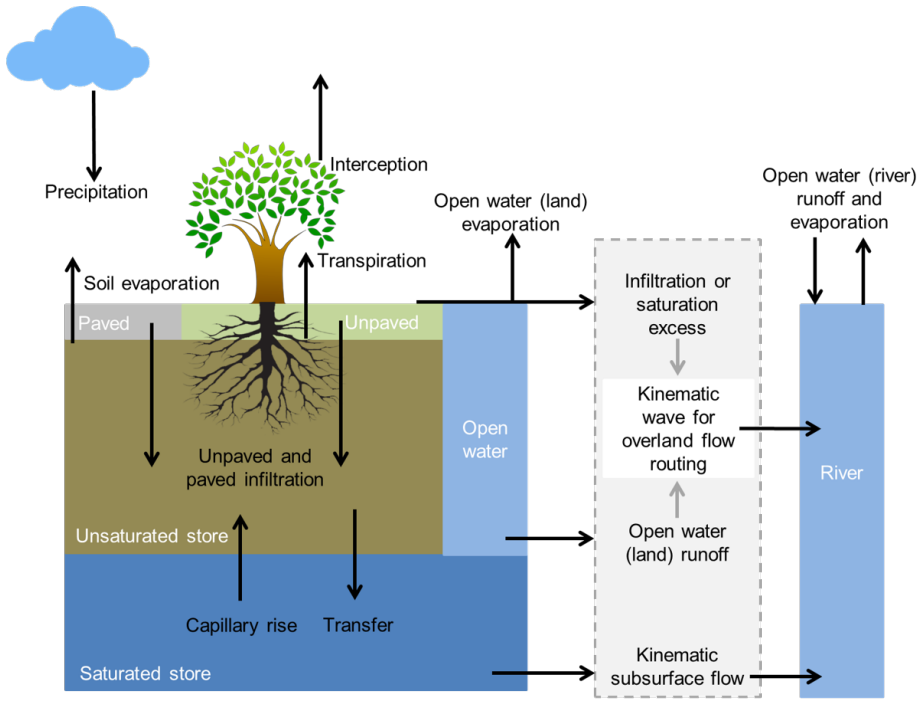


Figure 18: Overview of the different processes and fluxes in the Wflow model.

3.2.3 Calibration

For the HBV model, a Monte Carlo sampling is used to calibrate the model parameters, and 200000 iterations are used to find the best model parameters. Monte Carlo sampling of the Wflow model is not feasible due to the larger time required for a simulation, which is at least a factor of one hundred thousand larger, and moreover, the large number of parameters indicates many more runs are needed. Based on expert knowledge and the documentation of Deltares, 2023 several parameters are selected to be used for the calibration. Table 3 lists these parameters, what they represent, and their influence on the result.

Parameter	Represents	Effect
KsatHorFrac	Horizontal hydraulic conductivity at the surface of the soil	Height of peak flows and base flow
KastVer	Vertical hydraulic conductivity at the surface of the soil	Height of peak flows and base flow
f	Scale parameter that influence the subsurface flow	Tegression of peak flows
rootingDepth	Depth of which vegetation can extract water	Cumulative flow and peaks flow height
InfiltCapSoil	Infiltration capacity soil	Timing & height of peaks and cumulative flow
thetaS	Porosity of the soil	height of peak flows
N	Manning roughness of land cells	Timing of peak flows
N_river	Manning roughness of river cells	Timing of peak flows

Table 3: Wflow calibration parameters

The infiltration capacity parameter is assumed to vary during the year due to crust formation (Hoogmoed and Stroosnijder, 1984). Therefore, the infiltration capacity is a cyclic parameter with a different infiltration capacity for each month.

3.2.4 Model performance testing

The model performance is determined according to three standards, i.e. the Nash-Sutcliffe efficiency (NSE) of the discharge (Nash and Sutcliffe, 1970), Budyko framework (Budyko, 1974), and cumulative discharge. The NSE is a statistical metric used to assess the accuracy and performance of the discharge. Plotting the model simulations in the Budyko framework will give insight into the long-term actual evaporation values and the cumulative flow will provide insight into the amount of discharge that is predicted.

Nash-Sutcliffe Efficiency

The Nash–Sutcliffe efficiency (NSE) is a statistical measure that is used to assess the performance of hydrological models. The NSE is used because it is recommended by several well-founded articles (Legates and McCabe Jr, 1999 & ASCE, 1993) and as it is very commonly used it provides extensive information on reported values (Moriasi et al., 2007). However, a disadvantage, which has to be taken into account is that, as Willems et al., 2014 points out, the NSE is strongly influenced by time shifts between the simulated and observed discharge values. The NSE is calculated as follows:

$$NSE = 1 - \frac{\sum(Q_m - Q_{obs})^2}{\sum(Q_{obs} - \bar{Q_{obs}})^2} \tag{6}$$

Where Q_m is the modelled discharge and Q_{obs} is the observed discharge. An NSE value of 1 means a perfect fit and an NSE of zero means no more accuracy than predicting the mean value of the observed data (Jain and Sudheer, 2008). In hydrology, the acceptance threshold for the NSE is a value above 0.6. In this study, a warm-up period of one year is used, thus the year 2006, is not included in the calculation of the NSE value. The Wflow model runs for the complete 20 years originally suggested. However to make a fair comparison, in the Wflow model, the year 2006 is also not take into account.

In addition to the NSE, the NSElog is another model performance indicator that can be used. The NSElog is more forgiving of peak-flow errors while being more sensitive to low-flow perturbation. Combining both the NSE and NSElog, can improve model understanding and give better insight in both peak and low flows. However, the NSElog is not used in this study due to the lack of data during low flows.

Budyko Framework

The Budyko framework can be used to check the long-term water balance and long-term actual evaporation. The Budyko framework is based on the idea that the long-term average actual evaporation is mainly determined by the balance between the water supply (precipitation) and evaporative demand (potential evaporation) (Wang et al., 2016). Budyko, 1974 proposed the following expression for estimating the average actual evaporation:

$$\frac{E_a}{P} = [\frac{E_p}{P} tanh(\frac{E_p}{P}^{-1})(1 - exp(-\frac{E_p}{P})]^{0.5} \tag{7}$$

Where E_a, E_p, and P are the long-term actual evaporation, potential evaporation, and precipitation respectively.

According to Budyko, 1974, all catchments whether dry or wet, should be located nearby his proposed relation.

Cumulative discharge

The annual cumulative discharge of the observations and simulations are compared. This is done to prevent putting too much weight on only the NSE values. The cumulative discharge comparison mitigates the disadvantage of the NSE, which is sensitive to the timing of peak flows. The annual cumulative discharge is not sensitive to timing differences and is therefore used as an additional performance testing parameter.

3.2.5 Validation

To validate both hydrological models, the model performance outside of the interest period was investigated. Both models are run with the precipitation and potential evaporation data from the years 2011 to 2016. Assessing the accuracy of the discharge predictions of both models during this period allows us to evaluate the model's reliability and its ability to make accurate predictions.

3.3 Sensitivity analysis

The annual cumulative discharge of the observations and simulations are compared. This is done to prevent putting too much weight on only the NSE values. The cumulative discharge comparison mitigates the disadvantage of the NSE, which is sensitive to the timing of peak flows. The annual cumulative discharge is not sensitive to timing differences and is therefore used as an additional performance testing parameter.

Given that the two hydrological models, as described in section 3.2, differ significantly in terms of parameter count and computational run time, the sensitivity analysis of each model differs. This section begins by outlining the methodology used for the HBV model sensitivity analysis, followed by an explanation of the approach utilized for the sensitivity analysis of the Wflow model.

3.3.1 HBV model

To gain a better understanding of the HBV model's behavior and to ensure its robustness and reliability in producing accurate predictions, two sensitivity analyses will be carried out. These analyses serve as valuable tools for understanding the model's behavior and assessing its responsiveness to variations in input parameters.

The first sensitivity analysis employs Monte Carlo sampling. The method allows for testing the established parameter ranges and how wide of a range of a parameter performs well. The Monte Carlo simulation is run and the resulting NSE value for all parameters is plotted. The outcome of this procedure is a set of Nash-Sutcliffe Efficiency (NSE) values of the discharge, each corresponding to a specific parameter combination. Plotting these NSE values visualizes the model's performance across the entire parameter range.

To delve even deeper into the model's behavior and identify the effects of each parameter on the model output, a secondary sensitivity analysis is conducted. This analysis focuses on the eight model parameters, with a particular emphasis on their individual impact on two critical model outputs: the NSE and cumulative flow. Each parameter is methodically subjected to incremental increases and decreases, allowing one to observe the resultant variations in model performance. The range of variation for each parameter is set at 10% of its best-performing value.

The primary objective of this sensitivity analysis is to assess the relative importance and sensitivity of each model parameter. Identifying parameters with higher sensitivity gives valuable insights into the model's equifinality tendencies. Parameters with higher sensitivity will indicate a reduced chance of equifinality, implying that a smaller set of parameter combinations leads to the same accuracy. This understanding has significant implications for model calibration, as it informs us which parameters require closer attention and more precise calibration efforts.

In essence, this sensitivity analysis not only contributes to the understanding of the HBV model's behavior but also serves as a step toward refining the model's calibration process. Differentiating the individual

contributions of each parameter allows for to enhance of the model's reliability and accuracy in discharge simulations.

3.3.2 Wflow

The sensitivity analysis of the Wflow model is less extensive than the HBV model due to the longer simulation time. The sensitivity of the Wflow model is based on varying a selection of parameters that seem to be important based on expert knowledge and documentation. By doing this, the influence on the NSE, cumulative discharge, and visual performance of these parameters are tested. Due to the number of parameters in the Wflow model, only several parameters are tested and no insight into all parameter sensitivities is gained.

The parameters on which the sensitive analysis carried out are:

- Rooting depth
- Porosity
- Infiltration capacity
- Hydraulic conductivity
- Manning roughness

3.4 Reservoir water storage models

To simulate the reservoir dynamics, two reservoir water storage models have been employed, the HBV_RWSM and the Wflow reservoir module. Both models are based on the water balance equation. This section begins by discussing the water balance equation, followed by an elaboration of the differences between the two models used.

The water balance is described by:

$$\Delta S_{\Delta t} = P_{\Delta t} - E_{\Delta t} + Q_{in,\Delta t} - Q_{out,\Delta t}(-I_{\Delta t} - Inf_{\Delta t}) \tag{8}$$

Where $\Delta S_{\Delta t}$ is the change of water storage in the reservoir, $P_{\Delta t}$ and $E_{\Delta t}$ are the precipitation and actual evaporation, $Q_{\Delta t}$ the in and out flow of the reservoir, $I_{\Delta t}$ is the water abstraction for irrigation and $Inf_{\Delta t}$ is infiltration into the soil beneath the reservoir. All units are m^3, the subscript Δt means that all components are over period Δt; here fixed at 1 day.

3.4.1 HBV_RWSM

The HBV Reservoir Water Storage Model (HBV_RWSM) uses the simulated discharge from the HBV model for Q_{in}, t. The HBV_RWSM uses the same precipitation and potential evaporation as in the hydrological models. However, the water balance model requires the volume of water, which is why the precipitation is multiplied by the maximum water surface area. The maximum water surface area is used since it is assumed that even if the actual water surface area is smaller than its maximum, the precipitation that falls within the reservoir basin will still contribute to the water volume in the reservoir instead of infiltrating the reservoir bed. This assumption does not hold for the evaporation. The Volume/Area ratio of the reservoir is used to calculate the water surface area for each time step, which is then used to calculate the

volume of evaporation. It is assumed that the actual evaporation of a water body is equal to the potential evaporation.

The discharge out of the reservoir is determined based on the discharge at Yilou. To estimate the discharge out of the reservoir areal interpolating is used, the discharge at Yilou is multiplied with $A_{catchmentToece}/A_{catchmentYilou}$. The discharge into the reservoir is simulated with the HBV model. It is assumed that the catchment above Ramsa responds the same as the catchment above the reservoir Toece to precipitation and evaporation. This assumption allows to use of the HBV model parameters calibrated at Ramsa to simulate the discharge at the reservoir Toece. The alternative method involves utilizing the calibrated parameters of the HBV model, which are based on Ramsa, to simulate the discharge generated between Toece and Yilou. The modelled discharge can then be subtracted from the discharge at Yilou to obtain the outflow of the reservoir Toece.

Irrigation demand

The main purpose of the reservoir is to provide water for irrigation to allow agriculture in such dry climates. Diagnostic reports of the reservoirs provide information regarding the area utilized for each crop type, or the area is deduced from the crop yield. When the crop type and area specifications are available, it is possible to determine the water demand by:

$$I = k_c \cdot (E_p - P_{eff}) \cdot A \tag{9}$$

Where I is the irrigation demand in m^3, k_c is the crop coefficient, E_p the potential evaporation m, A is the area in m^2 used for the crop, and P_{eff} is the effective precipitation in m, which is the amount of precipitation that can be used by the vegetation.

The relation between the effective precipitation and precipitation varies depending on the precipitation amount. Additionally, the relationship depends on the location and, ideally, should be measured within the study area. However, since there is no available data in the study area, a general relationship, as provided by the Food and Agriculture Organization (FAO) (Allen et al., 1998) is used instead. This relationship is shown in figure 19.

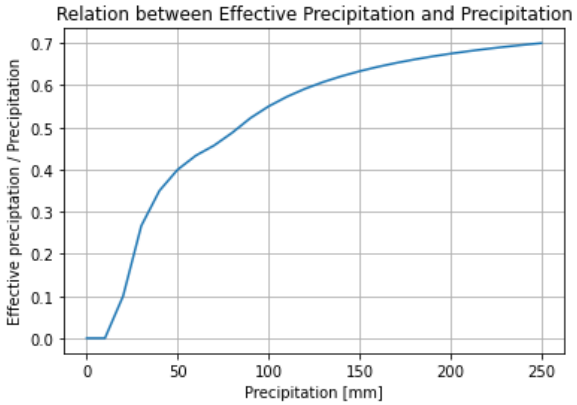


Figure 19: Relation between precipitation and effective precipitation (Allen et al., 1998)

The k_c values represent the integrated effects of changes in leaf area, plant height, crop characteristics, and degree of canopy. The k_c value is determined through experimentation (Pereira and Hrachowitz, 2011 & Allen et al., 1998). The k_c varies over time as the crop's characteristics change during its growth. Each crop has four distinctive growth phases: initial, development, mid-season, and late-season. A unique k_c

value is used for each growth phase. Appendix 7.5 lists the length and k_c value for each phase of each crop type, which is obtained from the FAO (Allen et al., 1998).

Three different variations of the HBV_RWSM are employed. Initially, the method described previously is applied to simulate the water storage in the Toece reservoir. However, to ensure a fair and consistent comparison, two constraints are integrated in the HBV_RWSM, aligning with the Wflow reservoir module. These constraints are:

1. The reservoir water storage cannot be below zero
2. The reservoir water storage cannot be higher than a maximum of $1.2e^8\ m^3$. If the water storage exceeds this threshold, the excess water is added to the discharge out of the reservoir (Q_{out}).

The final method introduces a scaling factor for both evaporation and crop water demand. These factors are calibrated based on the Nash-Sutcliffe Efficiency (NSE) value derived from the reservoir water storage time series, a warm-up period of one year is used to cancel out the effect of the starting conditions.

The potential evaporation is measured above the land surface and it is assumed that this is equal to the actual evaporation above a water body. According to Van de Giesen et al., 2010, the potential evaporation measure above land is not equal to the actual evaporation of larger water bodies. They suggest that the stability of an internal boundary layer of colder air above the reservoir may be the cause behind this phenomenon. Measurements taken in the Volta Basin indeed suggest that the ratio lies somewhere between 0.5 and 1 (Compaoré et al., 2008). This rectifies the use of a scaling factor for the actual evaporation out of the reservoir.

Furthermore, there is significant uncertainty regarding the crop harvest areas. To address this uncertainty, the scaling factor is applied to the irrigation outflow from the reservoir.

3.4.2 Wflow reservoir module

Wflow has a built-in reservoir module that simulates the reservoir water storage over time. The Wflow reservoir module simplifies the reservoir by representing it as a rectangular box, whose surface area is provided as input data (parameter: ResSimpleArea), $2.00e^7 m^2$. This simplification assumes a constant area that is used to compute the volumes of precipitation and potential evaporation from the input data. Wflow assumes that the actual evaporation out of a water body is equal to the potential evaporation, and thus fails to take into account the phenomenon of the reduced actual evaporation due to the boundary layer above the reservoir.

The reservoir receives inflow from the simulated discharge of the upstream catchment. The outflow from the reservoir is controlled by several factors including:

- The maximum storage capacity (ResMaxVolume)
- The minimum flow requirement downstream of the reservoir (ResDemand)
- The maximum discharge released if the water level is below spillway (ResMaxRelease)
- Target fraction full (of maximum storage) for the reservoir (ResTargetFull)
- Target minimum full fraction (of maximum storage) for the reservoir (ResTargetMinFrac)

The module assumes that in the case of a full reservoir, any excess water will immediately flow out of it. If the water volume is between the maximum and minimum storage limits, the outflow is determined by the downstream demand and the maximum discharge release when the water level is below the spillway. If the water storage is below the minimum required water storage level, there is no outflow.

The Wflow reservoir module lacks an irrigation outtake component. However, by defining an infiltration capacity ($Inf_{\Delta}t$), unaccounted water losses can be incorporated. The infiltration capacity is defined in mm and is subsequently multiplied by the reservoir area to calculate the volume of water loss.

Table 4 gives an overview of differences between HBV_RWSM and WFlow reservoir module.

Flux	**HBV_RWSM**	**Wflow reservoir module**
Precipitation	maximum reservoir area ($6.95e^7m^2$)	fixed constant area ($4.00e^7m^2$)
Evaporation	water surface area	fixed constant area ($4.00e^7m^2$)
Qin	HBV parameters used	Wflow discharge calculation
Qout	Discharge from Yilou & Maximum volume	Downstream demand & Maximum volume
Irrigation	Crop water demand	[-]
Infiltration	[-]	Infiltration constant (0.002 m/d)

Table 4: Differences between HBV_RWSM and the Wflow reservoir module

4 Results

In this chapter, the research results are presented in the same structure as used in the methodology. The first section (4.1) describes the results of the data selection. The second section (4.2) will provide the results obtained from the hydrological models. Section 4.3 describes the sensitivity analysis of the parameters in the hydrological models. Finally, section 4.4 will compare the results of the water balance simulations.

4.1 Data selection

The data selection first covers the precipitation data. This is then followed by the results obtained from the potential evaporation combination of local measurements and ERA5. Then, the discharges derived form the spillway equation are shown. Lastly, the effects of sedimentation on the reservoir water storage are presented.

4.1.1 Precipitation

Three different methods were utilized to analyze and select the most reliable precipitation data as input for the hydrological models. ERA5 and CHIRPS are evaluated by their performances at the location of the local measurements, the spatial variation, and temporal variation.

Figure 20 illustrates the point-to-pixel correlation at measurement stations in Ouahigouya (North), Guilongou (Middle), and Kombissiri (South) together providing a general representation over the study area, while supplementary correlation graphs of all other measurement locations are presented in Appendix 7.6.

On a daily basis, both ERA5 and CHIRPS show slight or, in certain instances, insignificant correlations with the local measurements. However, on a monthly or annual temporal scale, a significant correlation is observed with a correlation coefficient of 0.9 at some locations. The correlation between CHIRPS and the local measurements on monthly and annual temporal scales is always higher than that of EAR5. ERA5 tends to underestimate the precipitation, with ERA5 values mostly lying below the perfect correlation line. In contrast, CHIRPS does not show a constant tendency to overestimate or underestimation the precipitation compared to the local measurements.

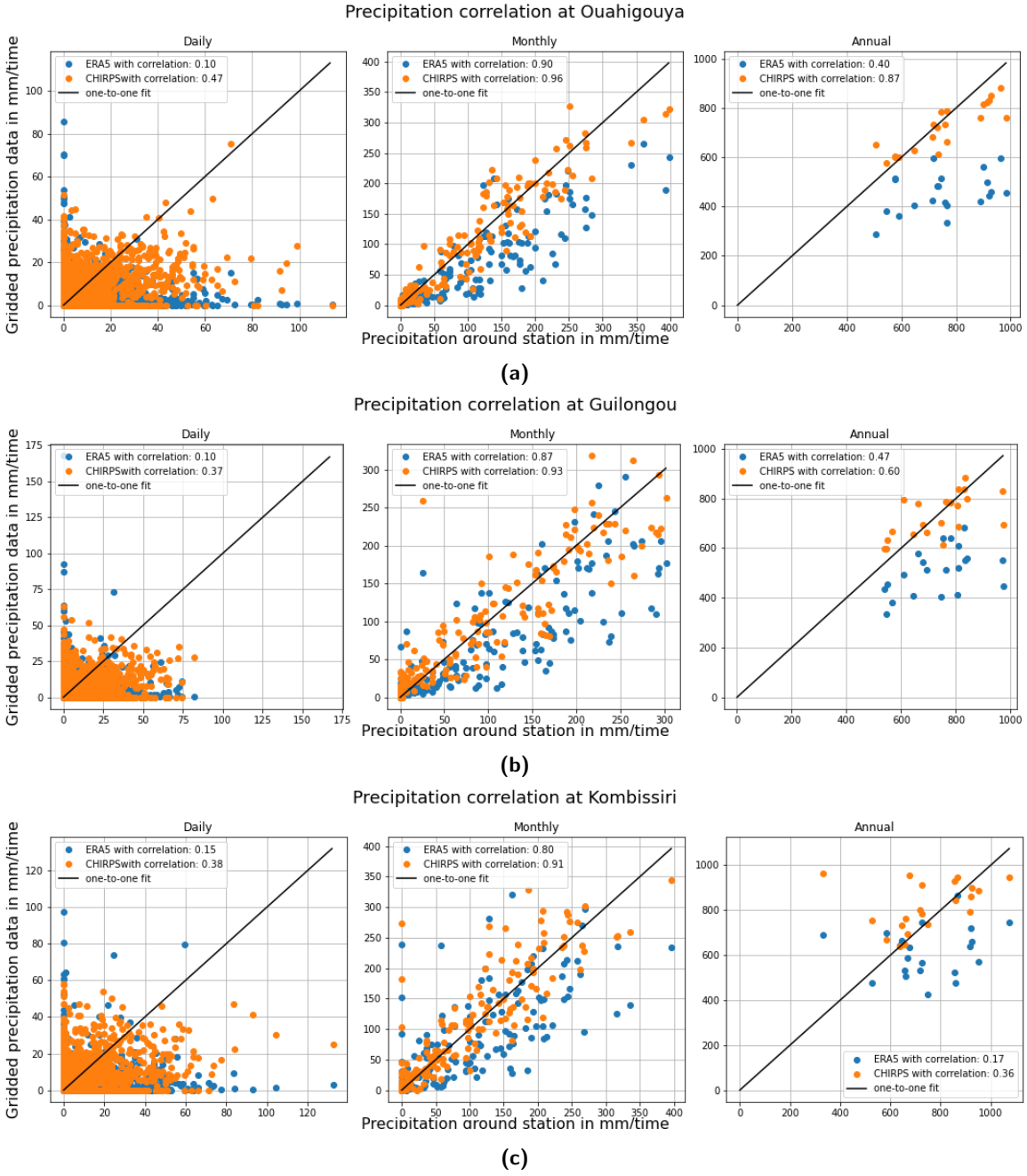
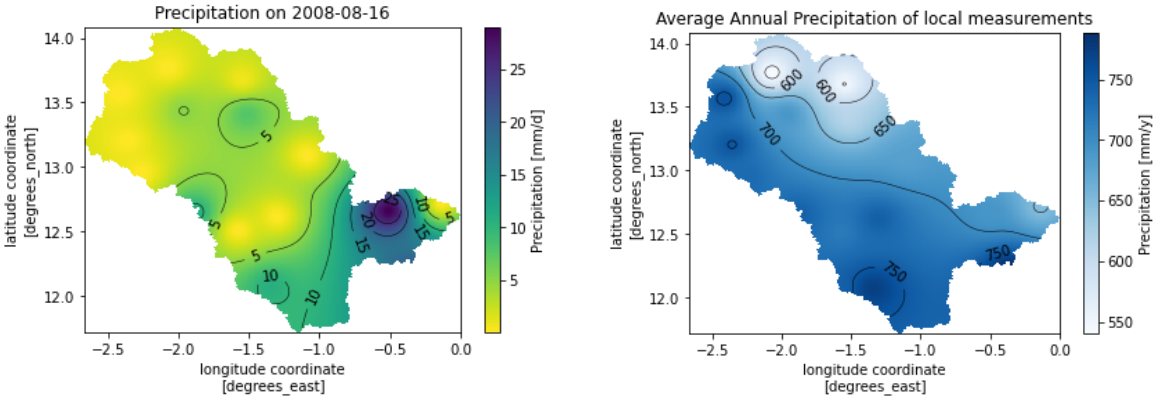


Figure 20: Correlation between CHIRPS, ERA5 and the local measurements at the north (a), middle (b) and south (c) of the study area.

To examine the spatial performance of ERA5 and CHIRPS compared to the local measurements, first, the Inverse Distance Weighting (IDW) interpolation is conducted. The IDW function relies on a power parameter that enables control of how known measurement locations impact the interpolated values, based on their distance from the output point. Typically, a power of 2 is employed. In this study, various powers are compared, a power of 2 shows the most reliable results, as raising the parameter to 3 or 4 increases the spatial weight of the known point, assigning this value to a larger expanse. Figure 21 illustrates the results of the IDW for a single day and year.



(a) **(b)**

Figure 21: Results of the IDW interpolation of the local measurements with a) one day example and b) one year example

The trend of ERA5 observed in the correlation graphs is similarly observed in the spatial pattern of ERA5. Figure 22 illustrates the spatial pattern of the relative differences between CHIRPS or ERA5 and the interpolated local measurements. To derive the difference, ERA5 or CHIRPS is subtracted from the local measurements. The graphs depict depict a blocked pattern in ERA5's spatial pattern, this is the pattern of the grid cells of ERA5. CHIRPS' grid size is significantly smaller which results in that CHIRPS does not show this pattern. In terms of spatial distribution, ERA5 constantly underestimates the precipitation by roughly 4000 mm/year compared to the local measurement. CHIRPS measurements, on the other hand, either overestimate or underestimate the precipitation when compared to the local measurements. Both, the absolute and relative difference plots of CHIRPS, as shown in Figure 22, indicate a trend the in north-south distribution. Under-predictions are present in the northern region, while over-predictions are present in the southern region, with a few small areas being an exception.The average annual differences for CHIRPS and ERA5 are -2 mm and 215 mm, respectively, when averaged over the whole catchment.

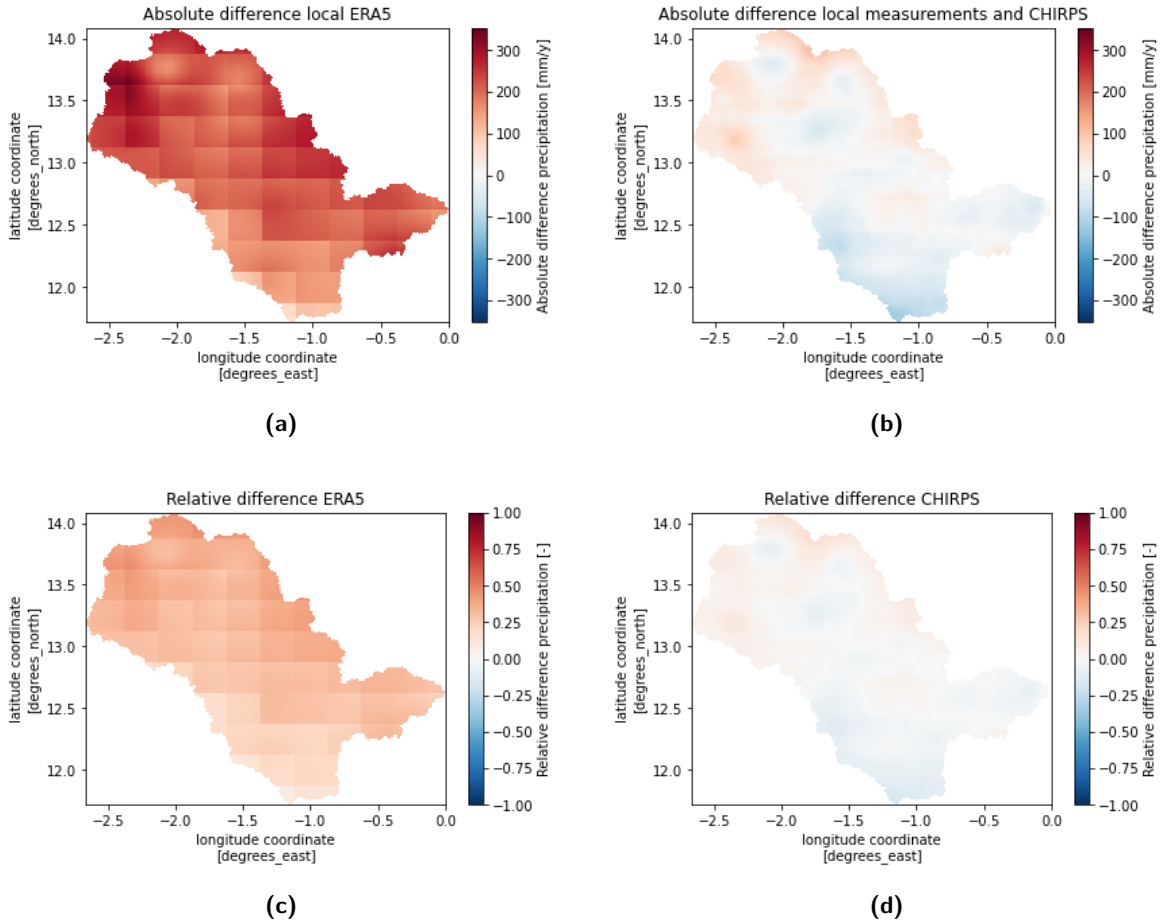


Figure 22: Spatial variation in absolute and relative differences of ERA5, CHIRPS and the local measurements (Local measurement - ERA5 or CHIPRS)

Besides the correlation and the spatial differences the temporal differences are investigated. The timing of the precipitation dictates the timing of the river flows and is thus very important. To capture a comprehensive overview of the temporal performance, the focus is on the monthly intervals, as the correlation graphs show that a daily comparison might be less meaningful. By aggregating the precipitation data across the entire catchment, the monthly averages are derived and shown in figure 23. The temporal variation in the relative difference in precipitation is also investigated. However, as figure 23, because during the dry months, the precipitation is very small, the relative differences become enormous. Therefore, only the values between -1 and 1 are plotted and the focus is placed on the months within the red square. It can be seen that during the wet season, again, CHIRPS scores closer to the local measurements.

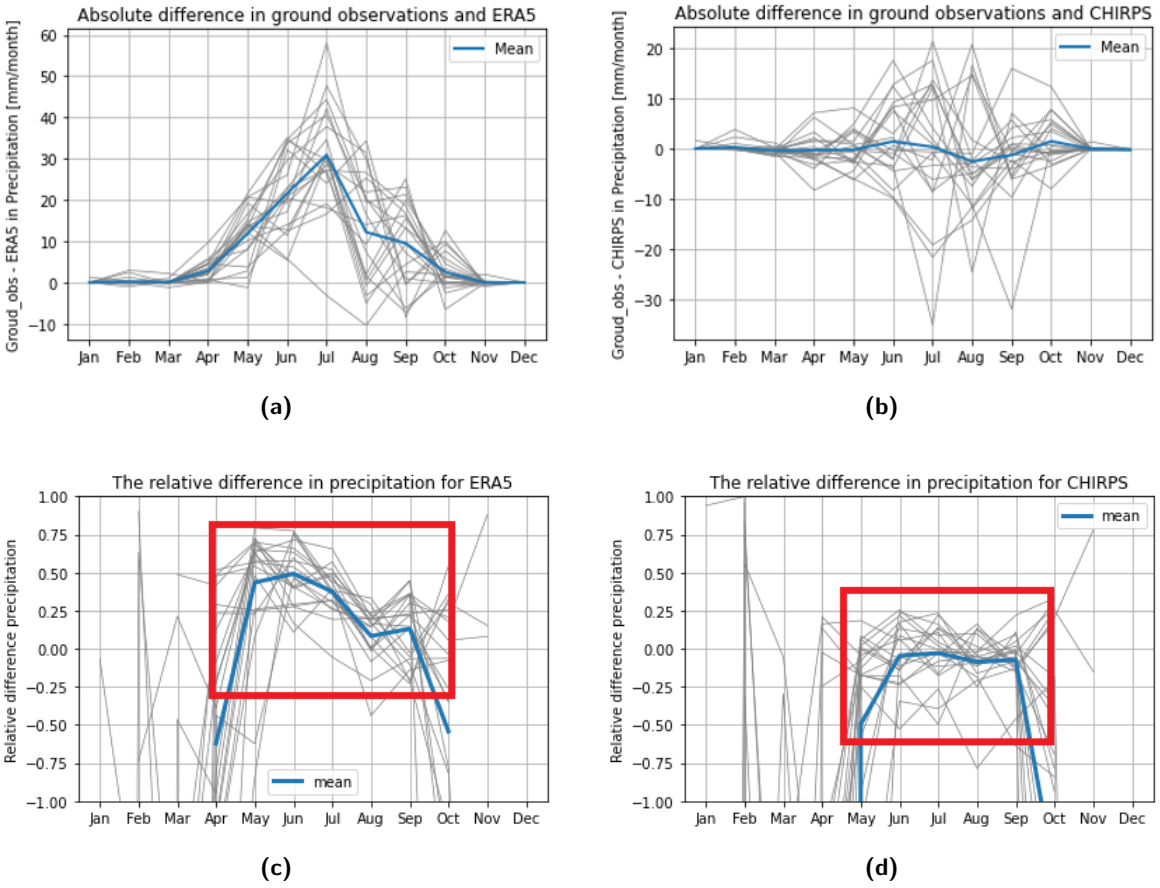


Figure 23: Temporal variation in absolute and relative differences of gridded data and the local measurements (Local measurement - gridded data) over time. The red box in figure (c) and (d) indicates the relevant period.

4.1.2 Potential Evaporation

The results of the potential evaporation analysis, as presented in Figure 7 and Table 1, highlight substantial differences among the various potential evaporation products. As cited by Dembélé, 2023, the study area experiences an anticipated annual potential evaporation of at least 2000 mm. However, a single product does not provide comprehensive temporal coverage, or spatial coverage and meets the requirement of 2000m potential evaporation annually.

To address these challenges and leverage the strengths of available data sources. The methodology for obtaining the most reliable potential evaporation combines the ERA5 reanalysis data with locally measured values. By assimilating spatial and temporal patterns from ERA5 and local measurements, a more robust dataset is received for subsequent analyses. To show the results of the adjusted potential evaporation figure 24 shows the local measured, ERA5 and the adjusted potential evaporation of two locations; Qougadougou and Ouahigouya. As a consequence of the adjustment, the potential evaporation became negative on some days, which is not possible. To cope with this, a minimum value of 2 mm/d is set.

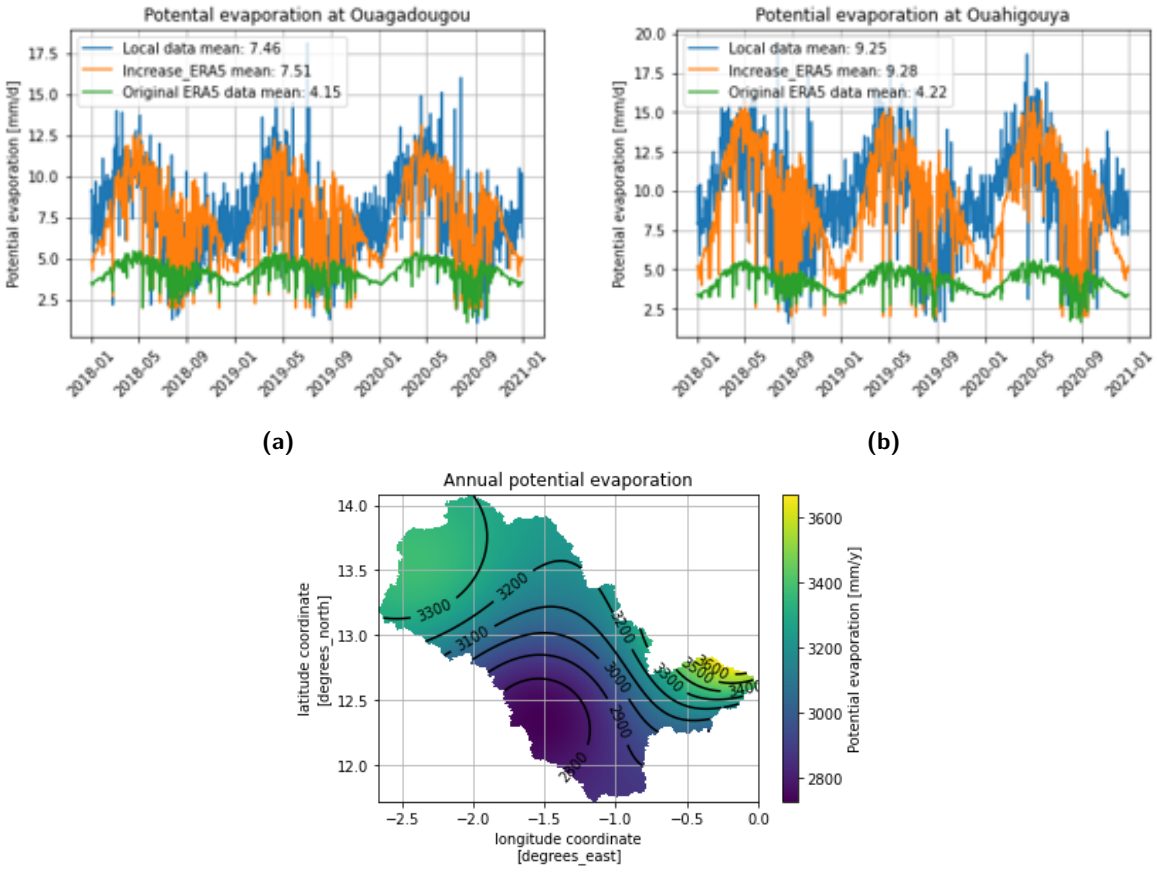


Figure 24: Effect of the combination of ERA5 and local measurements at (a) Ouagadougou, (b) Ouahigouya and (c) annual average in study area

4.1.3 Discharge

The focus period of the study is from 2006 to 2011, as these are the years when most discharge data are available for Ramsa and Yilou. The discharge for these years of Ramsa and Yilou is shown in figure 25. The figure shows the measured discharge in m^3/s and mm/d. Figures a and b clearly show that the discharge at Yilou is in the same order of magnitude as Ramsa even though Yilou is approximately 100km downstream of Ramsa, this effect is also shown in Figures c and d as the area correct discharge is shown. The area corrected discharge at Yilou is significantly lower than at Ramsa. Besides the differences in the discharge, the figure also clearly shows the significant missing data periods in the time series for both locations. To deal with the time gaps at Ramsa, the hydrological models are used to simulate the continuous discharge. At Yilou, due to unknown reservoir operations, this did not provide good enough simulations of the discharge. As described in the method section 3.1.3 the spillway equation in combination with the water levels of reservoir Toece is used to simulate the discharge out of the reservoir Toece. The results are shown below in figure 26. The outcome of the spillway equations is compared to the discharge measured at Yilou which is corrected for the catchment area of reservoir Toece. As the reference level of the water level in reservoir Toece and crest height are unknown, the crest height is varied to obtain the best fit. Figure 26a shows the results of fitting to the magnitude of the peaks. The magnitude of the peak is simulated decently, however, the timing of the start of the river flowing after the dry season is too

late and the peaks are simulated too late. Fitting to the timing of the discharge, figure 26b, shows that the peaks are overestimated. For both results, the retention of the discharge peaks is significantly slower when compared to the measured discharge at Yilou.

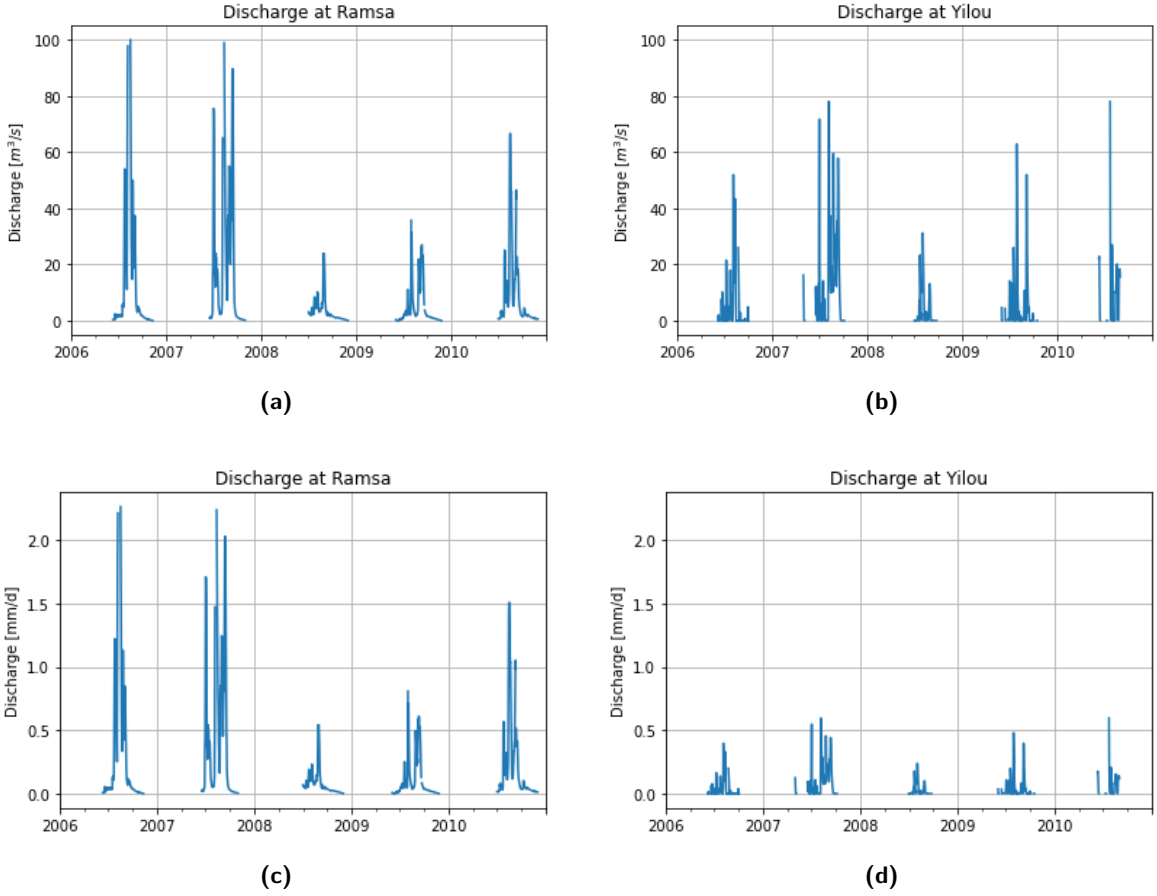


Figure 25: Discharge in m^3/s and mm/d at Ramsa (a & c) and Yilou (b & d) for the modelled time period

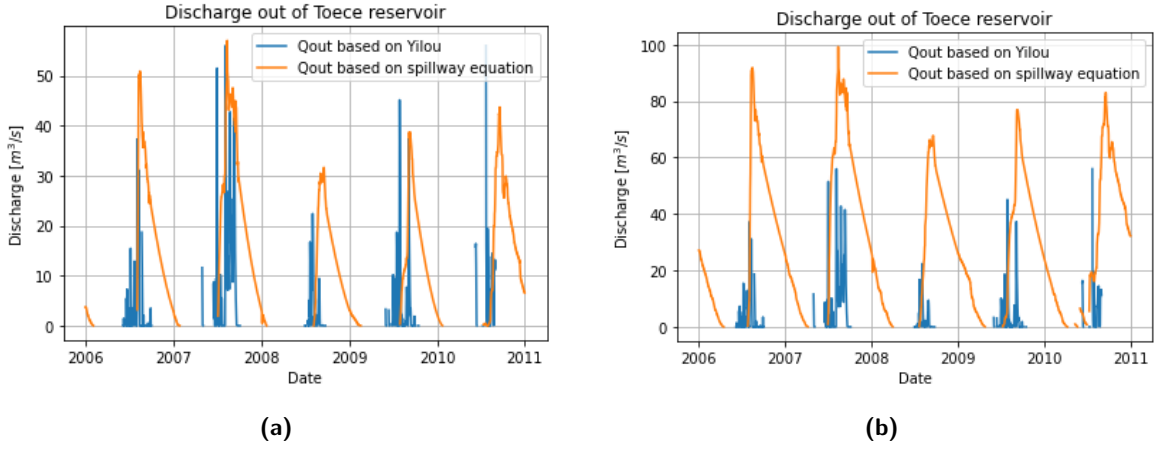


Figure 26: Discharge out of reservoir Toece based on Yilou and spillway equation, where (a) show the results for peak calibration and (b) for timing calibration.

4.1.4 Reservoir water storage

While the exact timing of the bathymetry measurements for the Toece reservoir remains uncertain, it is presumed to have taken place in 2020, at the end of the data time series. Including the annual sedimentation rate of 150000 $m3/y$ into the volumetric analysis of the Toece reservoir indicates a marginal decline in the water storage over time.

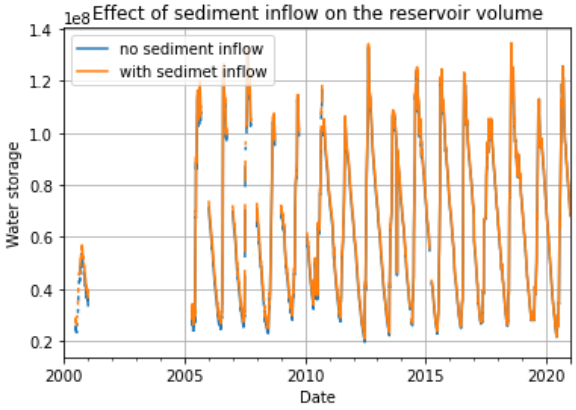


Figure 27: Effect of sedimentation on the reservoir volume in Toece

4.2 Hydrological models

In this section, the results of both hydrological models discharge simulation are delved into. First, the overall results are discussed, mainly focusing on the NSE. Subsequently, the yearly results of both models are presented, in which the focus lies on differences in anomalies with the observed discharge, focusing mainly on the timing of the peaks.

First, the NSE values of the discharge are considered, which offer a quantitative measure of the agreement between the models and the observed discharge data, using the first year, 2006, as a warm-up period. The hydrograph and corresponding NSE values of both models are presented in Figures 28 and 29. It's worth noting that, on the whole, the HBV model outperforms the Wflow model in terms of NSE values, 0.36 and 0.22 respectively.

In addition to the NSE values of the discharge, this study reports the cumulative flow. The cumulative flow is calculated both with and without the warm-up period. Table 1 displays the cumulative discharge over the period in millimeters. It's noteworthy that the inclusion or exclusion of the year 2006 significantly impacts the cumulative discharge. With the inclusion of 2006, both models underestimate the cumulative discharge, whereas excluding it results in an overestimation by both models. The HBV model (113%) aligns more closely with the observed cumulative discharge, whereas Wflow (128%) overestimates cumulative discharge to a greater extent. The reported observed cumulative discharge is based on interpolated discharge, as it provides a legitimate basis for comparison with modelled discharges and avoids the problem of missing data.

Period	Observed	HBV	Wflow
2006 - 2010	185 mm	169 mm	166 mm
2007 - 2010	128 mm	144 mm	164 mm

Table 5: Cumulative discharge

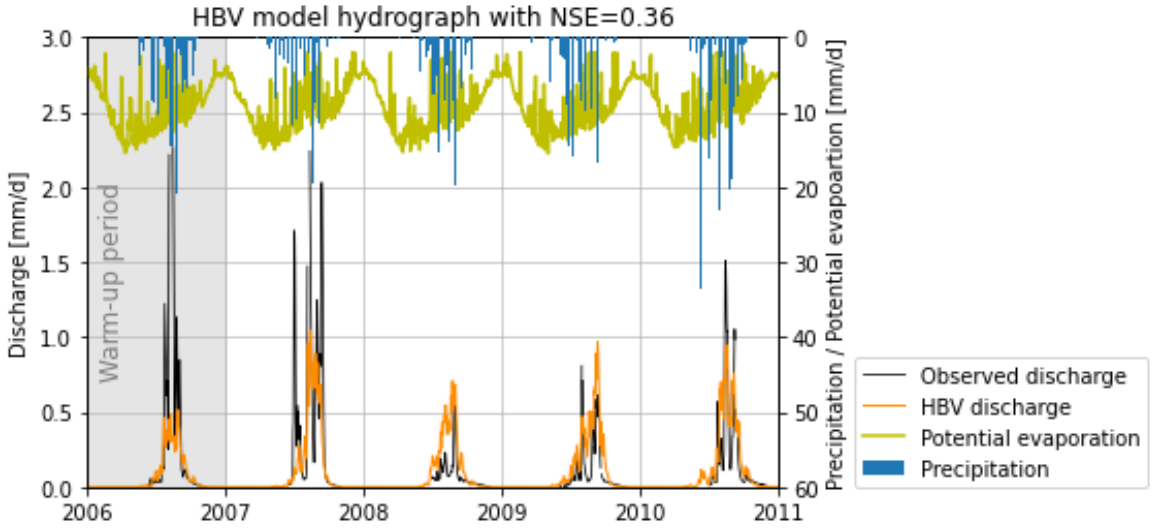


Figure 28: Discharge HBV simulation

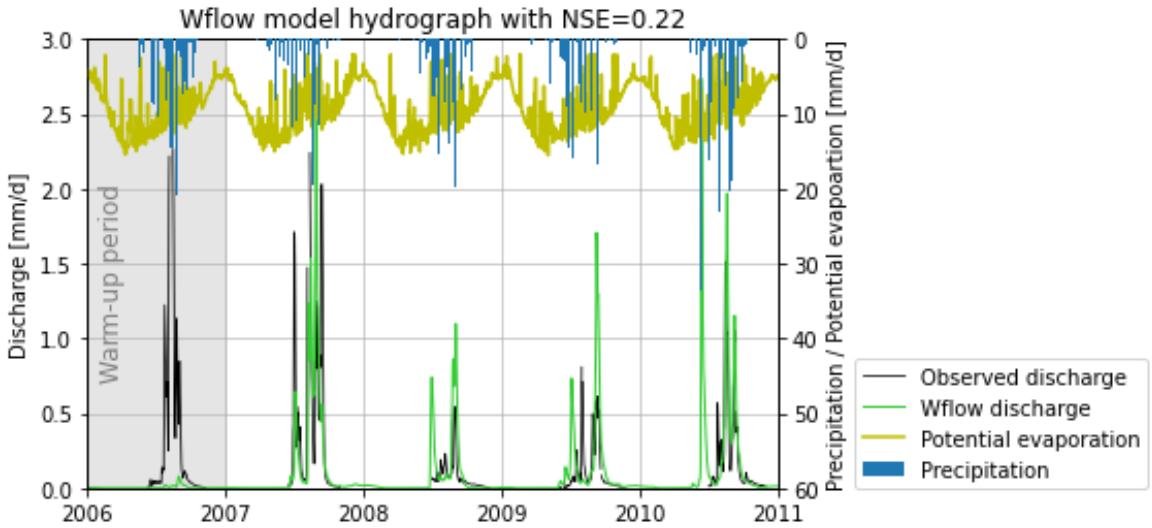


Figure 29: Discharge Wflow simulation

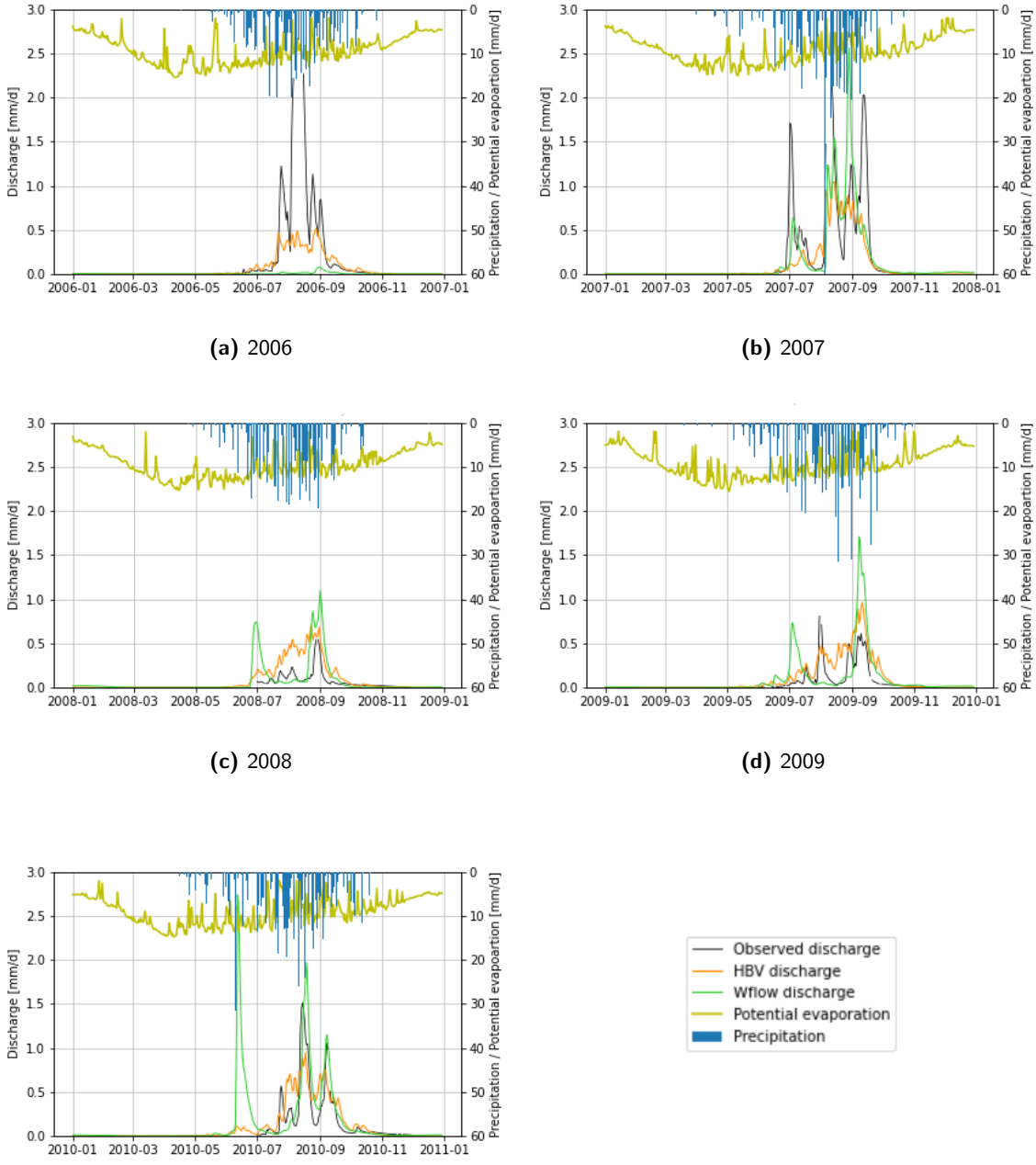
Yearly Anomalies and Model Performance

Figure 30 shows the results of the HBV and Wflow model in one graph per year. Based on the figure, the following anomalies draw attention.

- **Year 2006:** Anomalies begin with the year 2006, which stands out as an exceptional case in the Wflow simulation as nearly no discharge is simulated. While the HBV model effectively captures the timing of peak flow and regression, it significantly underestimates the magnitude of the peak.
- **Year 2007:** Moving on to 2007, we see that the HBV model falls short in capturing the first peak, whereas the Wflow model fares better in this aspect. Nevertheless, both models miss the final peak, despite relatively accurate predictions between those peaks.

- **Year 2008:** 2008 is an intriguing year by it self as there is an significant reduction in measured discharge, a pattern reflected in the simulations of both models. Notably, Wflow predicts an additional peak that was not measured. Throughout 2008, the HBV model consistently overestimates discharges, while Wflow offers predictions that are closer to the observed values.
- **Year 2009:** In 2009, the Wflow model predicts the peak flow occurring too early, while the HBV model accurately captures the timing but lacks precision in modeling the regression following the first peak. Both models, however, perform reasonably well in predicting the second peak timing.
- **Year 2010:** Finally, in 2010, Wflow predicts an unmeasured peak, though no measurements are available for this period. Overall, both models demonstrate respectable performance during this year.

In a broader context, it's noteworthy that both models effectively capture the regression trends in discharge over the entire study period. These results shed light on the strengths and weaknesses of both the HBV and Wflow models in simulating observed discharge patterns. While the HBV model consistently yields higher NSE values, the Wflow model exhibits specific strengths in replicating certain peak events.



(a) 2006 **(b)** 2007

(c) 2008 **(d)** 2009

(e) 2010 **(f)** legend

Figure 30: Hydrograph of the discharge simulations per year

4.2.1 Validation

Figure 31 and 32 display the model simulations for the years 2011 to 2015. The NSE is also presented in the figures. The NSE values are of similar magnitude to those from the calibration period, indicating that the models exhibit consistent performance in a different timeframe than the one used for model calibration.

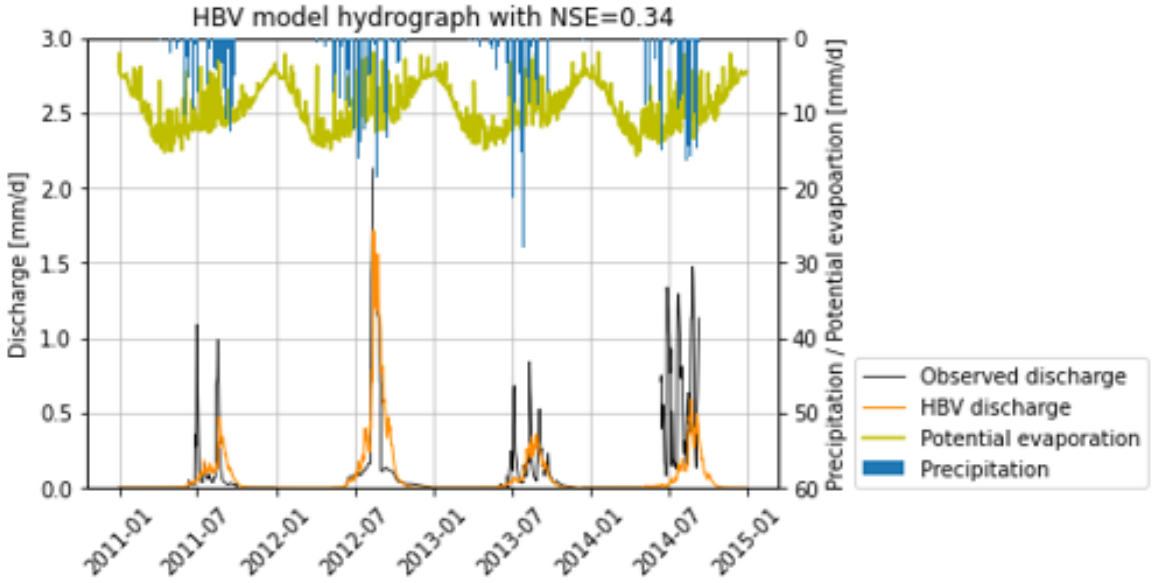


Figure 31: Discharge HBV validation

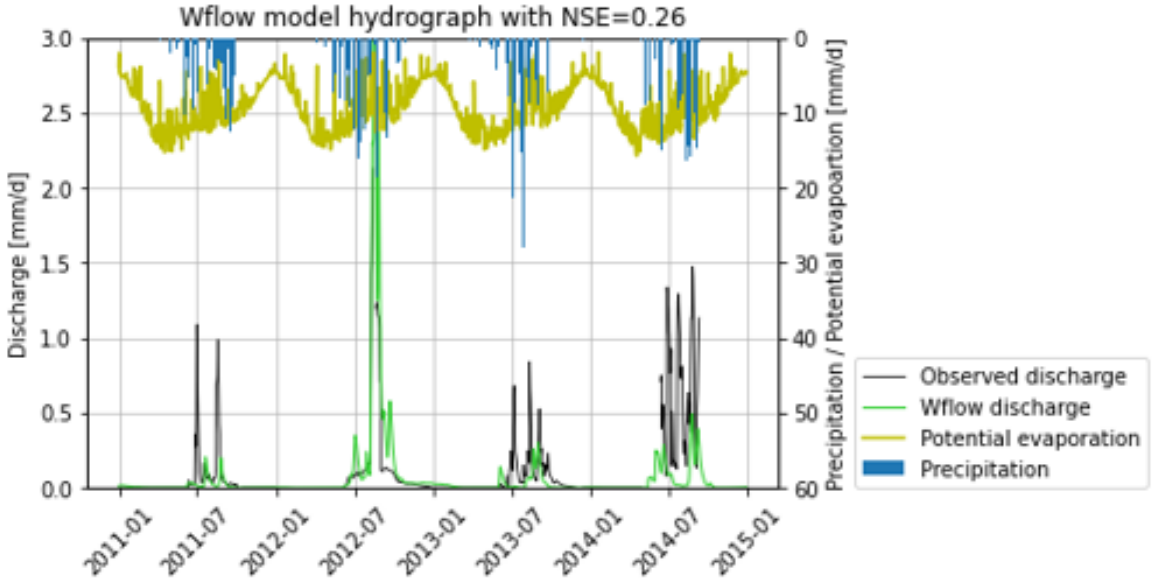


Figure 32: Discharge Wflow validation

4.2.2 Budyko Framework

Figure 33 shows the Budyko framework analysis, plotting the HBV (orange) and Wflow (green) model outputs alongside the observed data. The analysis shows a high degree of agreement between the three; the Wflow model, HBV model, and observed data within the Budyko framework. As both models and the observed data are based on the same data, Ep/P explains that there is no variance in the direction. Ep/P does however, very if one would compare different catchments or different datasets. Only small differences are observed in the Y direction. However, these are consistent with the cumulative discharge over time, which explains the differences. However, it all three are below the proposed curve of Budyko (Budyko, 1974).

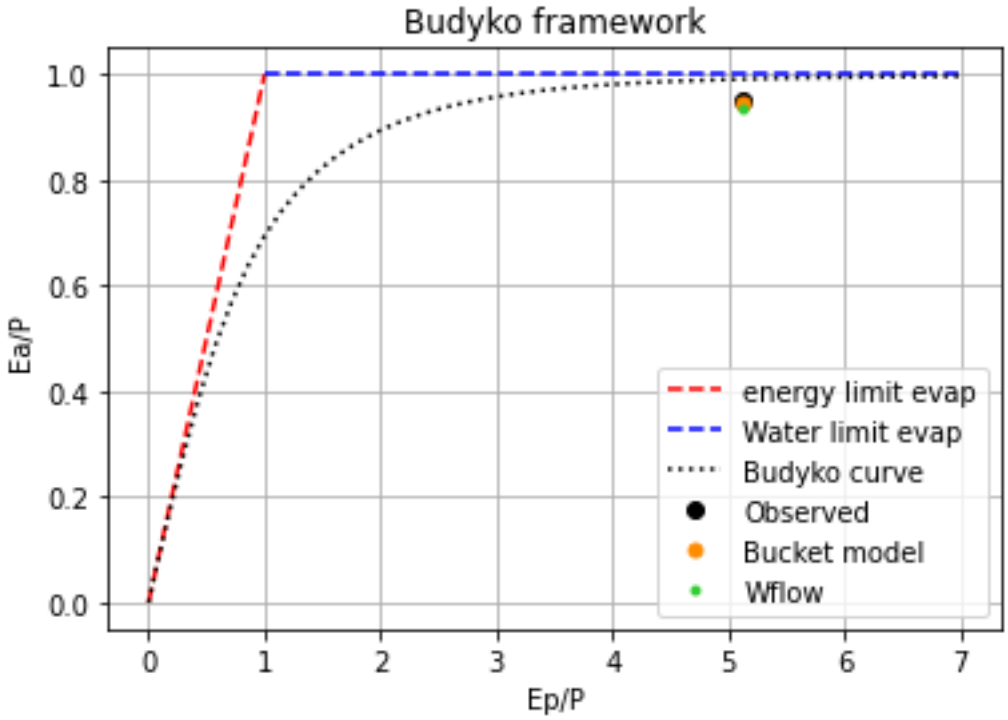


Figure 33: Budyko Framework

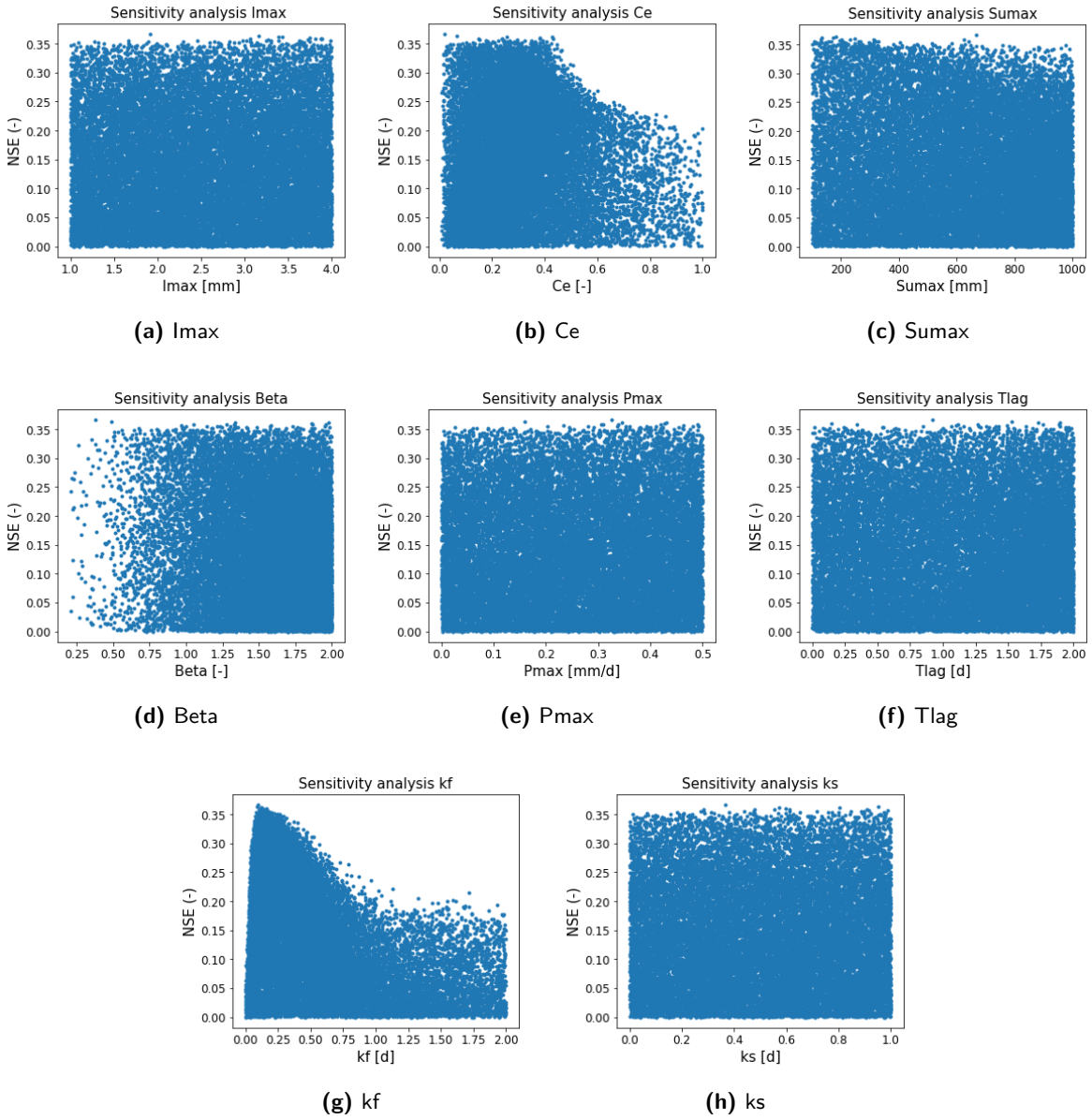
4.3 Sensitivity Analysis

This section presents the results of the sensitivity analysis. Firstly, the sensitivity of the HBV model is discussed, where all eight model parameters are extensively tested. This is followed by the sensitivity analysis of the Wflow, which shows the effects of several parameters.

4.3.1 HBV model

The first sensitivity analysis employs the Monte Carlo sampling. Through the Monte Carlo simulation, parameter values are systematically sampled within their predefined ranges, and assessed the resulting Nash-Sutcliffe Efficiency (NSE) values for discharge. These NSE values, corresponding to specific parameter combinations, were then graphically represented. The results are shown in Figure 34, which gives a visualization of the model's performance across the entire parameter spectrum.

Figure 34 illustrates that within the selected parameter ranges of 5 model parameters, namely Imax, Sumax, Pmax, Tlag, and Ks, the performance of the HBV model remains consistent. However, for the remaining three parameters, this is not the case. Figure 34b demonstrates that a Ce higher than 0.5 significantly leads to poorer NSE model performance. For a Ce lower than 0.5, the model performance remains consistent. A similar pattern is observed for the parameter Beta (Figure 34d), albeit less pronounced. With a low beta, the likelihood of good model performance is not excluded, but it is considerably lower than when the beta falls within the range of 1 to 2. Figure 34g reveals that the kf NSE analysis exhibits a peak between 0.08 and 0.15. Kf values outside this range result in notably poorer NSE model performance.



(a) Imax **(b)** Ce **(c)** Sumax

(d) Beta **(e)** Pmax **(f)** Tlag

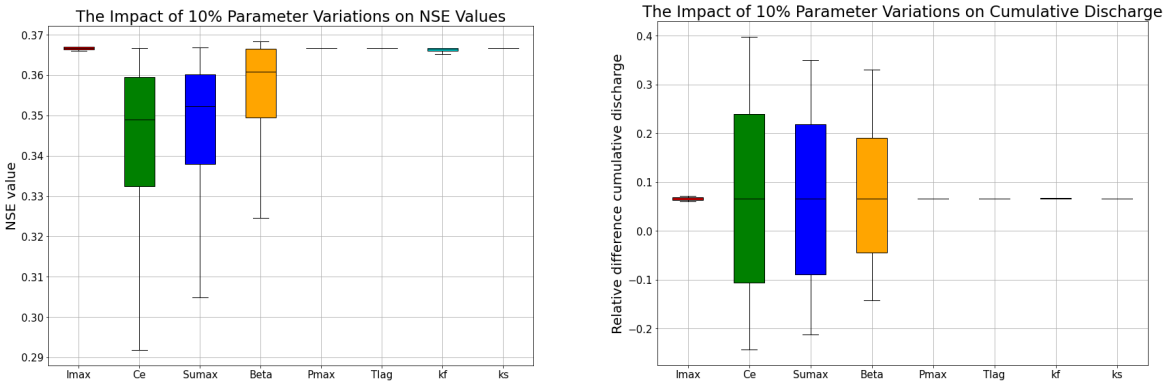
(g) kf **(h)** ks

Figure 34: Sensitivity analysis, performance ranges of the HBV model

Figure 35 presents the secondary sensitivity analysis conducted, which focuses on the effect of individual variation of each parameter on two critical model outputs: the NSE and cumulative flow. To achieve this, each parameter is adjusted with 10% of the best-performing parameter, both upward and downward.

In figure 35 the box represents the interquartile range, which extends from the first quartile to the third quartile of the data. The height of the box represents the spread of the middle 50% of the data. The line in the box is the median. Whiskers are lines extending from the box to the outer points within a certain distance of the box, in this case, the whiskers show the 5th and 95th percentiles.

Conducting a parameter-by-parameter variation while keeping the remaining parameters constant highlights that Ce, Sumax, and Beta have the most substantial influence on the modeled discharge. These three parameters exhibit the highest degree of variation in both the NSE and the relative cumulative discharge, indicating their pronounced impact on model performance.



(a) NSE

(b) cumulative discharge

Figure 35: 10% variation parameter effect

4.3.2 Wflow

The Wflow sensitivity analysis is conducted manually, as opposed to the Monte Carlo sampling used for the HBV model. In this section, several parameters are discussed that exert significant influence. The impact of these parameters will be delved into, how they impact factors such as the timing and magnitude of peaks, cumulative discharge, as well as variations across different years.

Rooting depth

The rooting depth determines how deep vegetation can extract water. With a higher rooting depth, vegetation can still absorb and evaporate water through transpiration from deeper soil layers, reducing the discharge significantly. Figure 36 shows the results of the Wflow with the rooting depth parameter increased by 50%. Even though the NSE value of 0.25 is larger than the Wflow run described in section 4.2, the modelled cumulative discharge is 42 mm/5y, which is almost three times as low as the observed cumulative discharge.

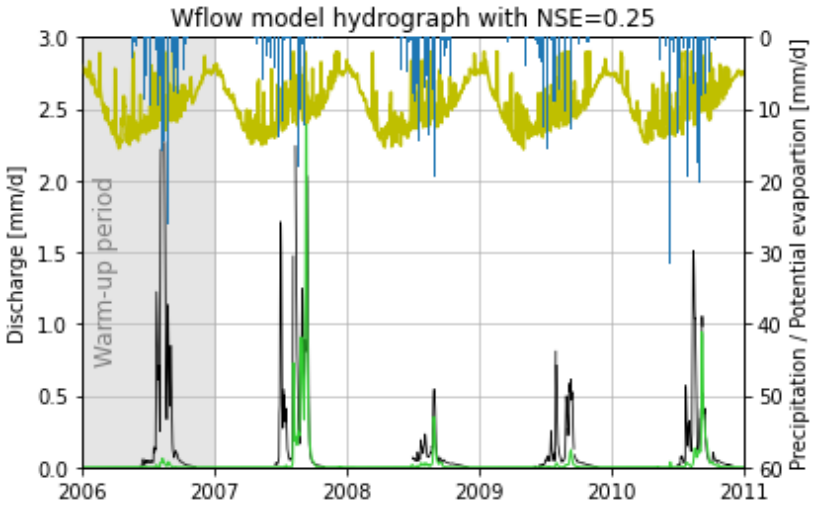


Figure 36: Wflow results with increase Rooting depth

Porosity

The Wflow model's porosity varies between land use types. To adjust the porosity, Wflow allows porosity

scaling. Increasing the porosity leads to higher water retention and less discharge. Figure 37 illustrates the effects of increased porosity and decreased rooting depth. The experiment aims to investigate the extent to which these two parameters counteract each other. Cumulative discharge and NSE are respectively 125 mm and NSE. In addition to the model's performance metrics, certain observations are noteworthy. Particularly striking is the small discharge wave at the beginning of the dry season in 2007, which was not documented due to a lack of data, making it impossible to establish definitive conclusions.

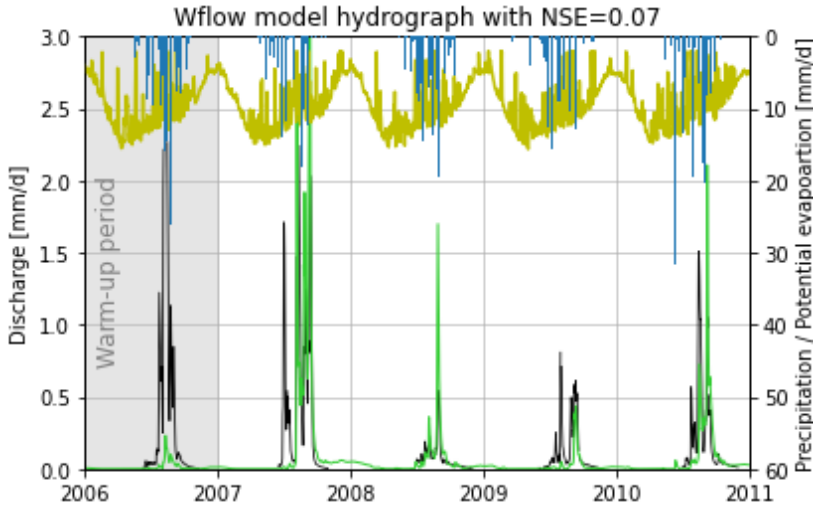


Figure 37: Wflow results with increase Rooting depth and increase porosity

In the Wflow simulations, the year 2006 always underpredicts the discharge. By slightly increasing porosity and reducing rooting depth, it is possible to achieve a better simulation of the discharge in 2006. However, as Figure 38 illustrates this adjustment results in a significant overestimation of peaks in the other years. Resulting in a cumulative discharge approximately twice as large, around 285, compared to the measured cumulative discharge. The NSE value is -4.91, which is indicating a poor fit. Similar to the previous simulation, the discharge wave modelled at the beginning of the dry period is not observed.

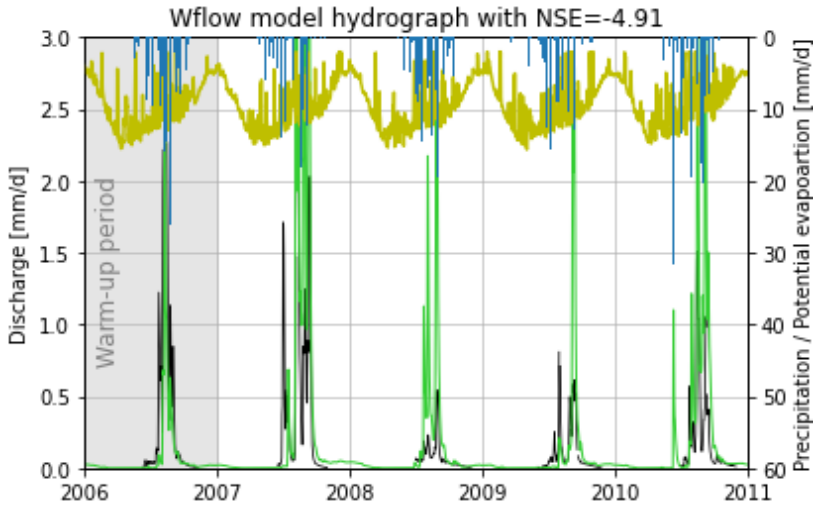


Figure 38: Wflow results with decrease Rooting depth and increase porosity

Infiltration capacity

Another parameter that is investigated is the soil's infiltration capacity. Decreasing the infiltration capacity

results in a reduction in the infiltration, meaning there is more immediate overland flow which results in a faster discharge response to precipitation. In many simulations, the first discharge peak after the dry period is absent. Figure 39 shows the results of decreasing the infiltration capacity with a scale of 0.5. The first peaks are now simulated, however, all other peaks are overestimated and the overall model performance is scoring low on NSE (-4.02) and is double the observed cumulative discharge (278 mm).

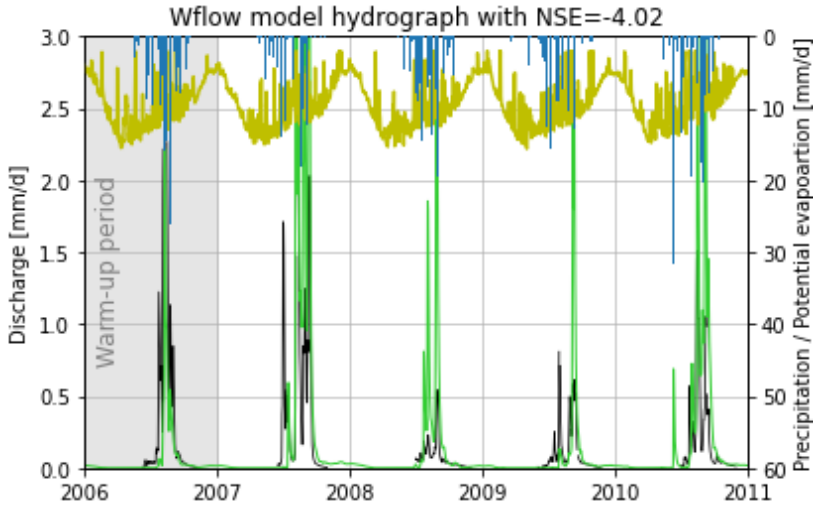


Figure 39: Wflow results with decrease infiltration capacity

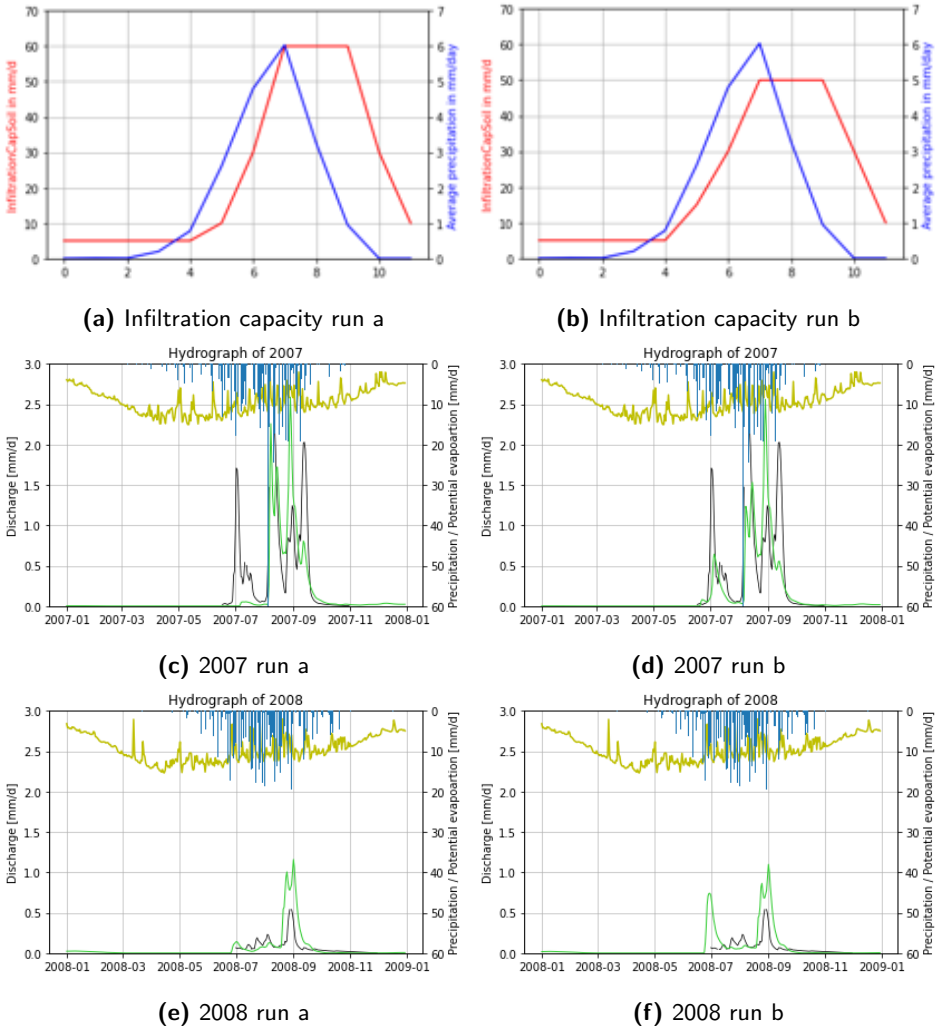
As all other peaks are overestimated, a monthly cyclic infiltration capacity is used, as described in section 3.2.3, due to crust formation. Figure 40 a & b shows the different infiltration capacities used, with all other parameters kept the same in the two runs. The overall differences between the two runs are very small with NSE values of 0.24 and 0.22, respectively a and b. The cumulative flow is also very comparative as both ranges within a 10% difference of the observed cumulative flow. However, there is a difference in predicting the first discharge peak. Run A does not predict a peak at the beginning of the rainy season, while run b does. When comparing with the data, something notable becomes apparent, as shown in Figure 40. In 2007, run a indeed misses the first measured peak, whereas run b predicts it. However, in 2008, the same pattern emerged, but there was no peak in the data, making the run a better match with the measured data. Thus, it varies from year to year which model setup best predicts the discharge.

Other parameters

KsatHorFrac represents the horizontal hydraulic conductivity at the surface of the soil. Increasing KsatHorFrac will raise the base flow and reduce peak flows. Decreasing KsatHorFrac will lower the base flow and increase peak flows. KsatHorFrac is a parameter that is not estimated based on soil parameters (unlike f and KsatVer), but it assumes a default value of 100. KsatHorFrac can have a large influence on the model performance but it can be difficult to have a well-educated guess on its magnitude.

KsatVer stands for the vertical hydraulic conductivity at the soil surface and is varied based on soil parameters in Wflow. Increasing KsatVer will raise the base flow of the runoff and flatten peak flows.

The Manning roughness of land and river cells can also be varied. Manning roughness is the roughness experienced by the flow and will therefore affect the timing of discharge. In the simulations carried out, Manning roughness did not have a very strong influence on timing. By reducing the roughness, the timing becomes only slightly earlier, the previously mentioned lack of the first discharge peak cannot be simulated by reducing the Manning roughness.



(a) Infiltration capacity run a

(b) Infiltration capacity run b

(c) 2007 run a

(d) 2007 run b

(e) 2008 run a

(f) 2008 run b

Figure 40: Influence of different cyclic infiltration capacities

4.4 Reservoir water storage

This section presents a comprehensive overview of the reservoir water storage fluctuations. The findings are presented through two distinct graphical representations. The first graph illustrates water storage fluctuations within the reservoir, while the second graph delineates the various inflows and outflows. The combination of these visualizations provides a profound understanding of the model's in simulating water storage fluctuations and the underlying mechanisms driving these dynamics.

4.4.1 HBV reservoir water storage model

The HBV Reservoir Water Storage Model (HBV_RWSM) simulates the reservoir water storage based on the precipitation, evaporation, inflow, outflow, and irrigation demand. This section presents the outcome of these simulations. Firstly, the irrigation demand results are shown. Subsequently, the water storage fluctuation simulations will be delved into, which are presented for several scenarios as different methodologies are used.

Figure 41 offers insight into the monthly variations in average crop water demand. The graph presents the

average crop water demand per month per crop type. Notably, a pronounced increase in water demand becomes evident at the end of the dry season. During this critical period, agriculture necessitates the most substantial allocation of water resources to sustain crops. The total crop water demand which is used in the HBV_RWSM is obtained by adding all water demands of each crop type together.

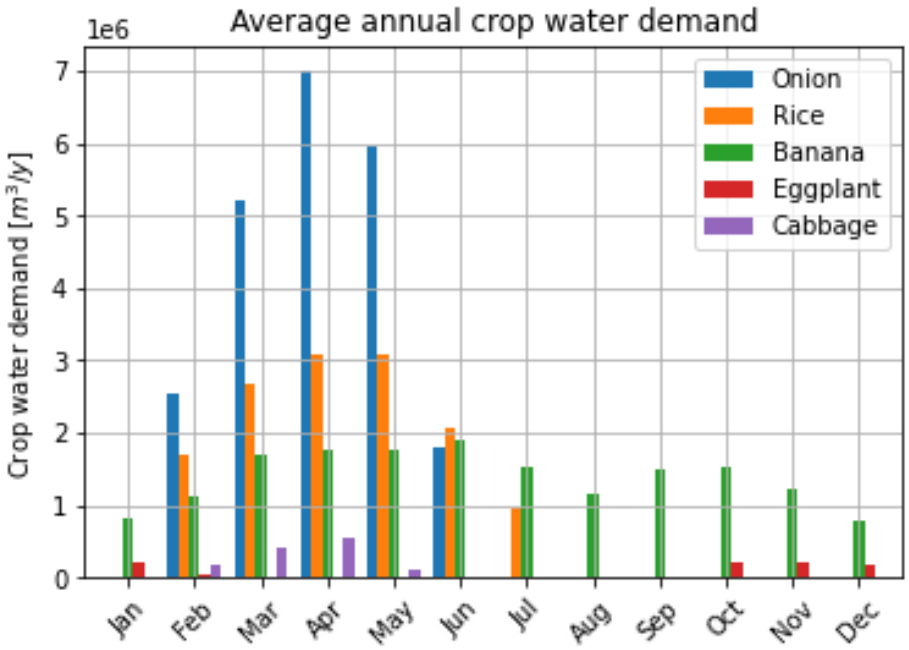


Figure 41: Average crop water demand per month

The HBV reservoir water storage model (HBV_RWSM) operates based on the inflow calculations of the HBV model and the outflow based on the discharge at Yilou. Figure 43 presents the results of the first HBV_RWSM run, which unfortunately does not simulate satisfactory performance. It is noteworthy that the water storage becomes negative in 2006 and 2007, which is physically not possible. Furthermore, during the next three years (2008-2010) the simulated water storage increases over the years while the observed water storage remains constant. Figure 43b shows the corresponding annual fluxes into and out of the Toece reservoir and clearly shows that the inflows are significantly larger than the outflows of the Toece reservoir.

To ensure a fair and meaningful comparison between the HBV_RWSM and the Wflow reservoir module, as outlined in section 3.4.1, two constraints are implemented. These two constraints are:

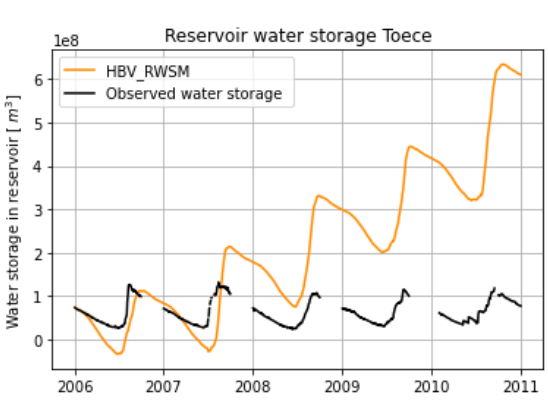
1. The reservoir water storage cannot be below zero
2. The reservoir water storage cannot exceed a maximum of $1.2e^8$ m^3. If the water storage exceeds this threshold, the excess water is added to the discharge out of the reservoir (Q_{out}).

Figure 44 shows the results of the HBV_RWSM with these constraints incorporated. As expected, the peaks of the water storage are simulated more closely in alignment with the observed water storage. However, the out-fluxes are too large as too much water is lost during the dry periods.

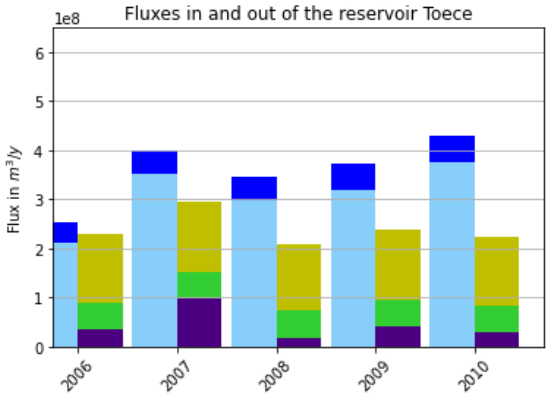
In the final method, a reduction factor is introduced to adjust both the evaporation and water outtake for irrigation. This modification acknowledges the discrepancy between the potential evaporation and the actual evaporation, explained in section 3.4.1, and accounts for the substantial uncertainties in the crop water demand calculations. Using a Monte Carlo sampling and the NSE of the reservoir water storage

resulted in factors of 0.8 and 0.6, respectively for the evaporation and irrigation demand. The results of these factors are shown in figure 45a. The figure shows that except for the year 2006, the peaks and troughs of the water storage fluctuations are simulated decently.

Figure 42: Legend of the fluxes graphs

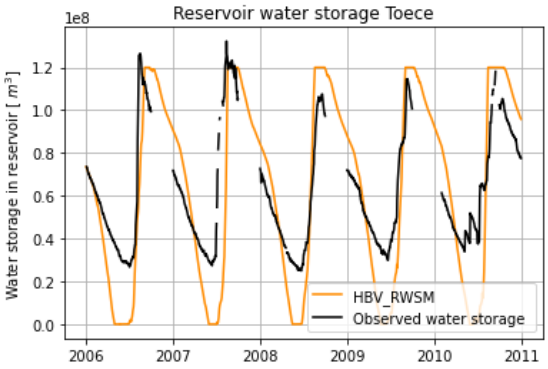


(a) Water storage fluctuation

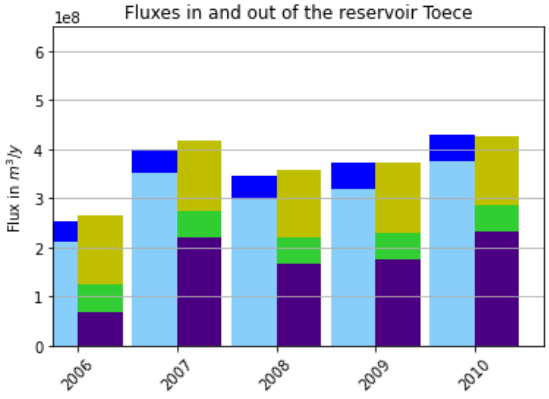


(b) Fluxes

Figure 43: Water storage (a) and fluxes (b) of the HBV_RWSM without constraint

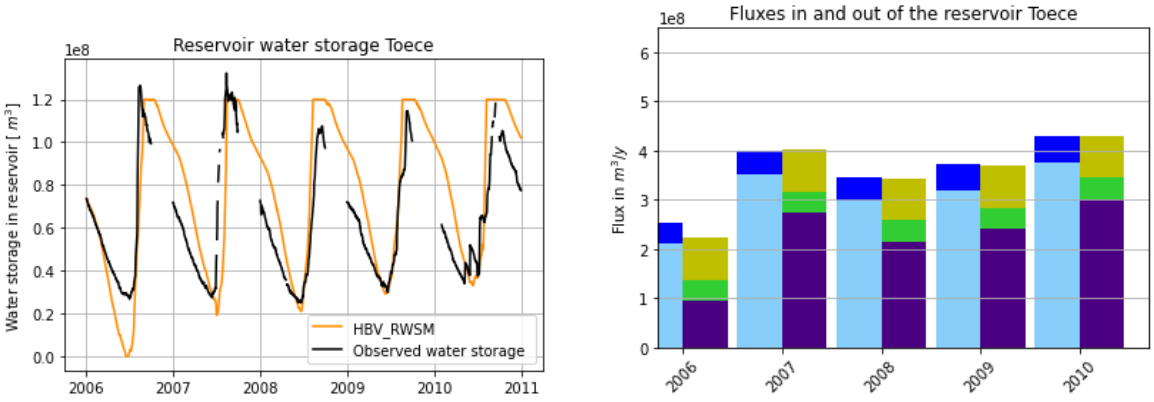


(a) Water storage fluctuation



(b) Fluxes

Figure 44: Water storage (a) and fluxes (b) of the HBV_RWSM with constraint



(a) Water storage fluctuation **(b)** Fluxes

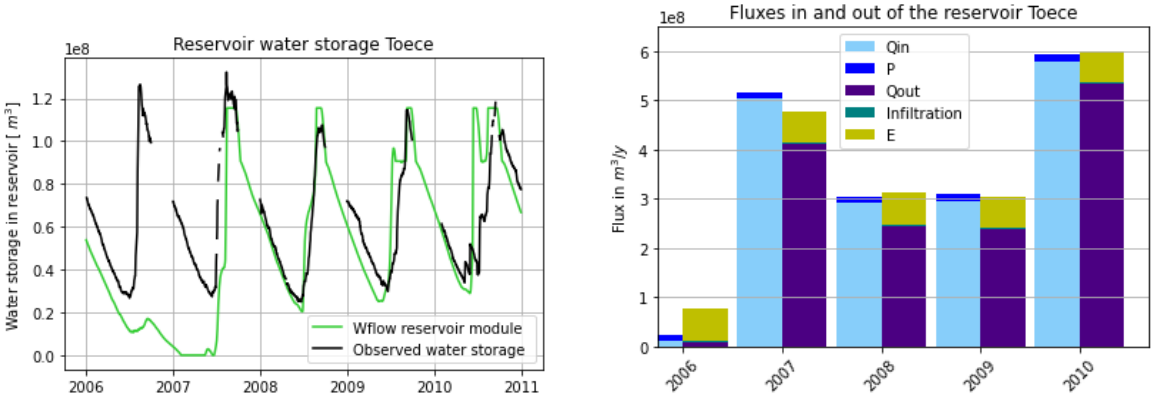
Figure 45: Water storage (a) and fluxes (b) of the HBV_RWSM with constraint and fraction of the Evaporation and Crop water demand

4.4.2 Wflow reservoir module

The Wflow reservoir module simulates the reservoir water storage assuming a constant area and the outflow is determined by water demand and maximum volume of the reservoir. Figure 46 presents the results of the Wflow reservoir module simulation.

Figure 46b provides an overview of the annual fluxes associated with Wflow reservoir module simulation. It shows that the discharge in and out of the reservoir is about 1.5 times larger than in the HBV_RWSM, in the years 2007 and 2010. Precipitation flux is significantly smaller and the evaporation flux is slightly smaller than the HBV_RWSM with scaling factor and significantly smaller than the other two HBV_RWSM simulations.

The Wflow discharge simulations at Ramsa already show almost no discharge in the year 2006. This is transferred to the reservoir water storage simulations as there is almost no inflow in the reservoir, leading to a mismatch with the observed reservoir water storage in 2006.



(a) Water storage fluctuation **(b)** Fluxes

Figure 46: Water storage (a) and fluxes (b) of the Wflow reservoir module

4.4.3 Comparison the HBV_RWSM & the Wflow reservoir module

Figure 47 provides a comparative view of the reservoir water storage simulations of both the HBV_RWSM and the Wflow reservoir module. The filling of the reservoir between 2007 and 2008 exhibits similarities in both simulations. However, a discrepancy emerges in the period after 2008, with the Wflow reservoir module filling the reservoir too rapidly and exhibiting an early peak, which is followed by a small trough to fill up again. This behavior is not observed in the HBV_RWSM simulation.

Furthermore, the depletion of the reservoir in the Wflow reservoir module showcases a concave form, while the HBV_RWSM exhibits a convex form. Additionally, it appears that the HBV_RWSM predicts low periods more closely in alignment with the observations compared to the Wflow reservoir module, which occasionally predicts early and deeper declines in reservoir water storage.

In addition to visual observations, the Nash-Sutcliffe Efficiency (NSE) of the reservoir water storage is evaluated. Table 6 provides an overview of NSE values for various warm-up periods, offering insights into the models performances.

Warmup period	HBV_RWSM	Wflow reservoir module
1 year	0.39	-0.04
1.5 year	0.42	0.42

Table 6: NSE of reservoir water storage with different warm-up periods

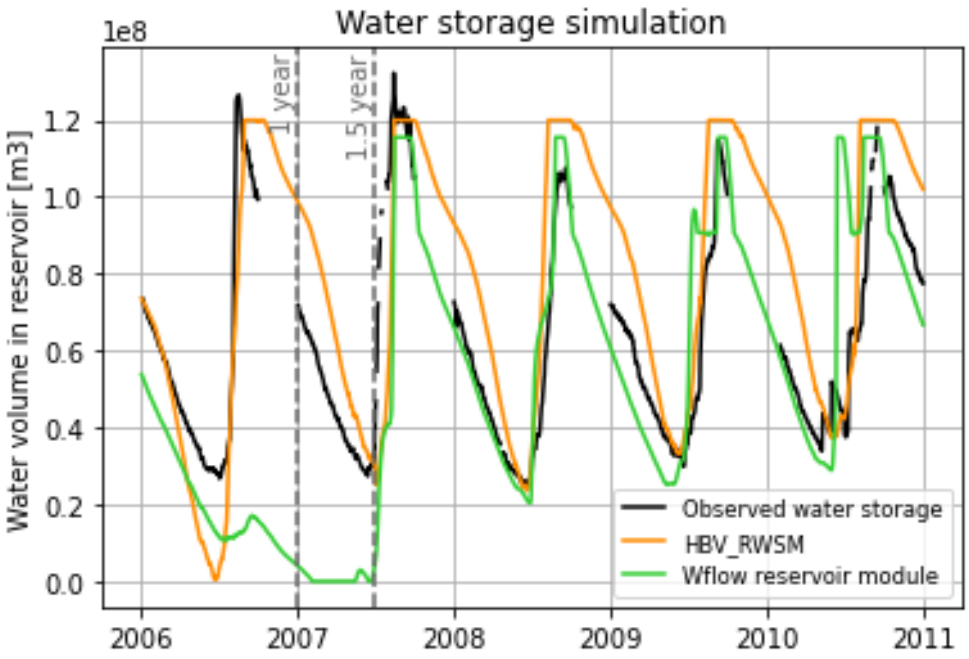


Figure 47: Reservoir water storage simulations of both HBV_RWSM and Wflow reservior module

To gain a better understanding of the results, rather than solely relying on the results of both reservoir water storage simulations, the modelled reservoir outflow is compared to the outflow from the reservoir based on Yilou discharge data. Figure 48 displays these different reservoir outflows. Here, it is once again evident that Wflow underestimates the outflow in the year 2006. In the other years, both models significantly overestimate the reservoir outflow, with Wflow showcasing a greater overestimation compared to HBV_RWSM. The NSE value for this outflow is calculated and is -8.2 and -26.2 for HBV_RWSM and

Wflow, respectively. This indicates that both models are in poor agreement with the observed data from Yilou.

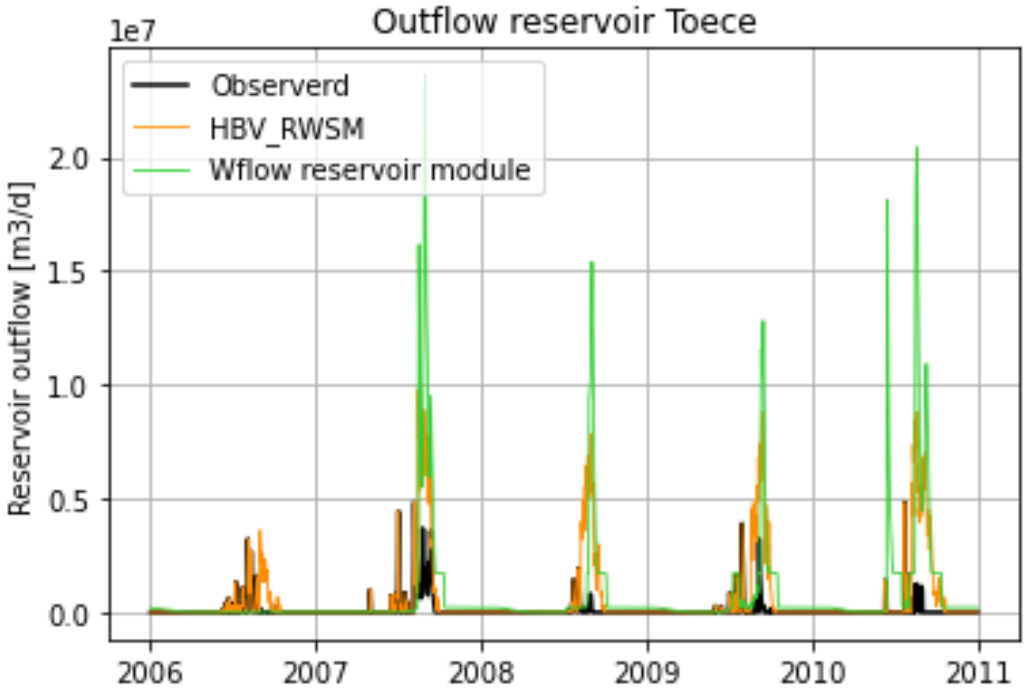


Figure 48: Comparison of reservoir outflow of both reservoir storage models and observed data

5 Discussion

This chapter interprets the results and puts them into perspective. It first follows the same order as in the methodology and results chapters.
In 5.1, the results of the data selection are discussed. Not only are the results delved into, but also improvements to the methodology are discussed. The second section (5.2) will delve into both the hydrological models and the sensitivity analysis. These are combined in one section as the parameter sensitivity is of importance to put the results of the hydrological models into perspective. Next, section 5.3 discusses the reservoir water storage models and their results.

Finally, by taking a step back and viewing the research from a broader perspective, section 5.4 delves into the lessons that can be drawn from this case study. Furthermore, it is explored how the research on reservoir water storage modeling can contribute to the ultimate goal of investigating measures to reduce water scarcity in Burkina Faso.

5.1 Data selection

The data selection focuses mainly on the precipitation (5.1.1), potential evaporation (5.1.2), and discharge data (5.1.3). The differences between local measurements, satellite, and reanalysis data are discussed for both the precipitation and potential evaporation. Next, the effect of lack of discharge data and different methods to cope it this data deficit are considered. Not only the dynamic data are discussed, but also the effects of static data selection and assumptions.

5.1.1 Precipitation

Three precipitation products are utilized in this study, specifically local measurements, CHIRPS, and ERA5. Several different comparison methods, including point-to-pixel correlations, and spatial and temporal differences are employed to evaluate the products and ascertain the most appropriate precipitation product.

Local measurements are generally more reliable than satellite or reanalysis data due to subjective errors in their processing and calibration. However, there are disadvantages in using the local measurements such as time gaps in the data and the fact that local measurements are point measurements and not a grid covering the whole study area. Additionally, the local measurements are obtained using an Excel sheet that lists only the daily value. The measurement method and exact location are unclear, with only the nearby city provided. As a result, the true quality of the data and any associated uncertainties cannot be ascertained.

To spatially compare the local data with CHIRPS and ERA5, the Inverse Distance Weighting (IDW) interpolation method was employed. The advantage of this method is that it fills in time gaps. Different power functions were tested, with a higher power giving more influence to a point over a greater distance. However, due to local showers, this was not desirable. This is why the default power of 2 was chosen. Nevertheless, even with this method, the measurement locations have a significant impact over a large area as the distance between measurement locations exceeds 50 km. Figure 21a illustrates that the entire southern region receives high precipitation because one point in the south recorded a high value. While this may be the reality, it is unlikely. Additionally, in the north, it can be observed that zero values are observed at the measurement stations, while values in between are slightly larger than zero due to the influence of a point further away. It is more likely that the precipitation is zero at this point. Decreasing

the power function further to reduce this effect led to the reverse phenomenon in which low precipitation values are interpolated between two high precipitation measurements. On the daily scale, the locations of the measurement locations can clearly be seen in the precipitation. In the annual averages, this effect is mitigated, and the measurement points are less distinct in the spatial pattern.

Based on the combination of all comparison methods, CHIRPS is selected as the superior precipitation product. The correlation graphs clearly show that CHIRPS consistently exhibits higher correlations than ERA5 at both monthly and annual scales. Daily, both products perform poorly, often showing no correlation. The correlation graphs already reveal that ERA5 frequently underestimates precipitation compared to local measurements. This is also evident in the spatial and temporal differences, which show that ERA5 consistently predicts lower precipitation values than the local measurements. CHIRPS fluctuates around the local measurements. This is evident in the correlation graphs where CHIRPS falls both above and below the one-to-one fit plot. However, in the spatial graphs, it is noticeable that CHIRPS more often underestimates in the north and overestimates in the south. A reason for this can be the previously described phenomenon of assigning high precipitation values to a large area in the south of the study area as fewer measurement locations are present.

Considering the timing, the differences always occur during the wet period, which is logical since all products register zero during the dry period. CHIRPS underestimates and overestimates precipitation differently each year compared to the local measurements. This leads to an overall good average performance for CHIRPS.

CHIRPS has a smaller grid size than ERA5, which likely allows it to better observe local showers than ERA5. Additionally, CHIRPS is calibrated with local measurements and is more focused on precipitation. ERA5 is an atmospheric reanalysis climate dataset, where precipitation is not the main focus but rather a byproduct. Therefore, ERA5 is likely to perform less well than CHIRPS. This is also supported by literature Logah et al., 2023, which found that CHIRPS performs well in this region.

Thus, CHIRPS, when compared to the local measurements, outperforms ERA5 and is thus selected as input data for the models. However, it is good to realize that CHIRPS is selected as the most suited, but this is probably not the exact reality and might be a reason why the results of the hydrological models, which use a daily time scale, differ from the observed discharge. This finding is in line with Brocca et al., 2020 which found that in West Africa the integrated products, such as CHIRPS, outperform both local- and reanalysis-based (ERA5) precipitation estimates.

5.1.2 Potential evaporation

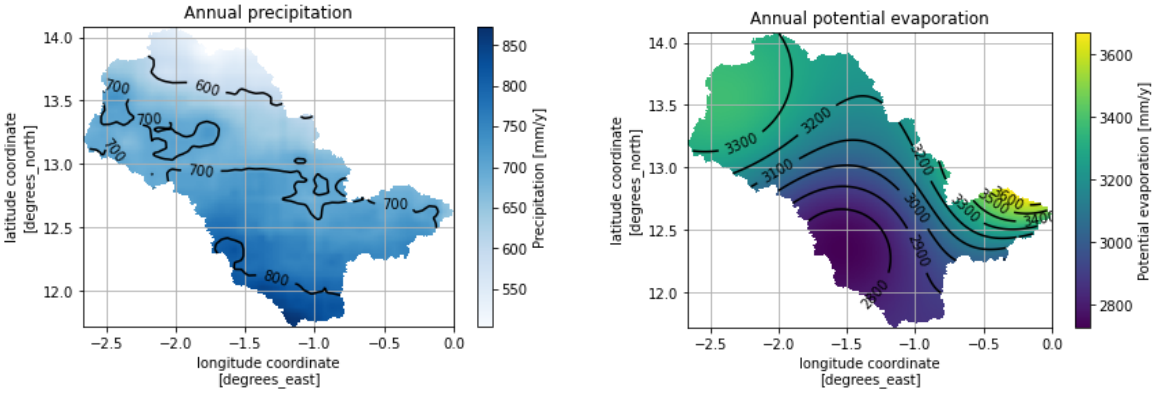
A broarder range of potential evaporation products are considered compared to precipitation. However, as described in sections 2.3.2 and 3.1.2, the differences among these products are substantial. None of these products fully meets all the desired criteria, complete daily coverage, and potential evaporation of at least 2000 mm annually. Consequently, a combination is created of ERA5, which offered good spatial and temporal coverage, and local measurements. This approach yielded values that align with the expected patterns in this region. However, the use of local measurements did introduce an influence on the pattern.

For instance, the measurement taken in Ouagadougou (located in the middle left of the study area) leads to lower values in the vicinity of this location. Resulting in a wave pattern of the contour lines of potential evaporation. This is in contrast with the patterns expected based on products like GAIPE (as shown in Figure 9d). Nevertheless, the combination of ERA5 and local measurements is utilized as input for the model, as it meets the criteria of daily coverage and grid representation. It's worth noting that the existing literature predominantly focuses on actual evaporation, making it challenging to directly compare our results with existing studies. An old study that compares potential evaporation products shows average

potential evaporation rates of 8 mm/d in the climate type of Burkina Faso, which is in the same order of magnitude as the combined potential evaporation product (Anyadike, 1987).

Similar to the precipitation, there is little known about the measurement techniques of the local precipitation measurements. The measurement method used is Pan evaporation but the exact location is not mentioned as only the location of the nearby city is given. Therefore, just like for the local precipitation measurements, the real quality and uncertainties within the data are not fully understood.

Figure 49 shows the annual average of the selected input data for precipitation and potential evaporation. The precipitation clearly shows the north-south pattern, with less precipitation in the north, closer to the Sahara.



(a) **(b)**

Figure 49: Annual values of the precipitation and potential evaporation which are used as input for the model

5.1.3 Discharge

Discharge data is very scarce, and there are no available remote discharge products within the study area. Therefore, the study relies solely on the discharge time series measured at Ramsa and Yilou. The discharge data is provided as a time series with one m^3/s value per day. Both locations, Ramsa and Yilou, are situated near or in reservoirs, which introduces a considerable level of uncertainty regarding the measured values. The exact method of discharge measurement is not known as only the time series is provided, but measuring discharge within a reservoir may not be the most reliable location.

Additionally, the discharge peaks exhibit rapid changes, transitioning from almost no discharge to high discharge within a few days, and then returning to low levels within a similarly short period. This pattern might suggest that discharge is primarily generated through overland flow, which seems unrealistic at this catchment scale. The numerous gaps in the data further contribute to the uncertainties. It is acknowledged that the river is intermittent, however, it remains uncertain whether assuming no discharge during the missing data periods in the dry season is accurate.

In rivers, it is generally expected that discharge increases downstream. However, at Yilou, the discharge is of a similar magnitude to Ramsa, despite a distance of over 100 km between them and the catchment area contributing to Yilou being three times larger than that of Ramsa. This suggests that either human interventions have significantly reduced the discharge or that one or both measurements are inaccurate. Human influences undoubtedly play a role in an area like Burkina Faso, which has numerous reservoirs. However, it is not feasible to assess these influences within the scope of this study, and thus, it is not possible to conclusively determine the cause of this anomaly in discharge measurements.

Given the absence of alternative data sources, the models must be calibrated using the available data. Nevertheless, as mentioned above the reliability of the measured discharge is questionable, resulting in significant uncertainties in the hydrological models.

To deal with these uncertainties, the spillway equation is used to obtain the reservoir outflow of reservoir Toece. The use of this method results in significantly slower retention of the discharge peaks. However, the spillway equation requires the water level and crest height as inputs. The crest at Toece Dam consists of three different sections with different heights, and since these heights are not measured from the same reference point as the water level there is information lacking for the spillway equation. By varying the combinations used to calibrate the spillway equation, the timing and magnitude of the peaks cannot be matched to the measured discharge and therefore cannot be used as input to the reservoir storage models.

5.1.4 Reservoir water storage

The uncertainty regarding the timing of bathymetry measurements for the Toece reservoir can influence the height-volume relation used to calculate the water storage capacity. These measurements are presumed to have occurred in 2020, at the end of the available data time series. Incorporating an annual sedimentation rate of 150,000 $m^3/year$, based on a previous study of Witteveen+Bos, sheds light on the reservoir's behavior over time.

The inclusion of sedimentation rates reveals a marginal decline in the water storage within the Toece reservoir. This decline can be attributed to the gradual accumulation of sediment over the years. While the reduction in storage is noticeable, it does not appear to be substantial within the given time frame.

5.2 Hydrological models

In this section, the two hydrological models and their performance are discussed. First, the HBV model will be discussed and is followed by the Wflow model. The section will conclude by comparing both models and their performances based on the Nash-Studcliff efficiency, cumulative discharge, and Budyko framework.

5.2.1 HBV

In this section, we delve into the HBV model, its parameters, and the outcomes of a sensitivity analysis, shedding light on equifinality and parameter sensitivity within the model. The HBV model is explained in Section 3.2.1. It describes the four different reservoirs, the various fluxes between these reservoirs, and the parameters that govern them. The sensitivity analysis in Section 4.3.1 provides further insights into the effect of different parameters.

The sensitivity analysis based on Monte Carlo sampling reveals that five out of the eight parameters, namely Imax, Sumax, Pmax, Tlag, and Ks, can consistently result in a high NSE value within the given range. For these parameters, the HBV model is highly sensitive to equifinality, as there are numerous ways to achieve an equally well-performing solution.

Furthermore, among these five parameters, three parameters were shown to be less prone to equifinality. Regarding the beta parameter, it is evident that a lower beta value has a reduced likelihood of achieving good model performance. Beta influences how the water from the interception reservoir is distributed over the fast and unsaturated reservoirs. Where a large beta will increase the flow towards the fast reservoir. As mentioned in section 5.1.3, the discharge peaks show fast responses to the precipitation. Within the

HBV model, fast responses are captured through the fast reservoir, which explains the tendency to larger Beta values.

Ce values above 0.5 noticeably decrease model performance, shown in figure 34. Ce influences evaporation from the upper layer of the soil, the unsaturated reservoir. A higher Ce value means that a smaller portion of potential evaporation actually results in actual evaporation. Figure 35 illustrates that varying this parameter significantly impacts model outcomes in terms of both NSE values and cumulative discharge. This underscores the critical importance of accurately estimating Ce, particularly in semi-arid regions where evaporation rates are high. The cumulative impact is also strongly affected by the Ce value, with a range of $+/-$ 10% resulting in relative discharge differences ranging from $+$0.4 to -0.25.

Cumulative precipitation is influenced by actual evaporation, as it is the only way for precipitation to exit the system without becoming part of the discharge. The total actual evaporation is determined by the evaporation from the interception reservoir (E_i) and the evaporation from the unsaturated reservoir (E_a). Figure 35 clearly demonstrates that the parameters Ce, Sumax, and Beta have a more significant impact on the cumulative discharge than the Imax parameter. As these are the three parameters that influence the E_a and Imax influences the E_i, this suggests that total actual evaporation is primarily determined by E_a, making Ce, Sumax, and Beta more critical parameters than the Imax.

kf is the parameter that most closely approaches one optimal value, as shown in Figure 34g. This means that the likelihood of estimating this parameter correctly is higher than for other parameters in the model. This can be explained by the steep regression of discharge determined by kf within the HBV model. Figure 35 shows that kf has little influence on cumulative discharge, which is logical since kf determines the regression rather than the amount of discharge. However, the impact on the NSE value is also relatively small. This can be explained by the fact that only 10% of the optimal parameter has been varied to achieve a kf of approximately 0.1. This variation is very small, which may make the effect appear less significant.

Improving the HBV model can probably best be done by adding elevation to the HBV model, even though it adds more parameters to the model, and thus an increase in the equifinality would be expected better model performances can be found(Gharari et al., 2011). They suggest that elevation is a strong indicator for different dominant runoff processes as elevation allows for a more direct link to the hydraulic gradient, which is arguably the most dominant factor for any type of runoff generation. However, Gharari et al., 2011 also mentions that the runoff process is mainly influenced by landscape features scale of approximately 50-100 m. This means a higher DEM resolution than currently used in this study is needed.

Overall it can be concluded that the HBV model is very subjective to equifinality and that the parameters Ce, Beta, and Sumax, which influence the actual evaporation, have the most impact on the model performance.

5.2.2 Wflow

This section delves into the results of the sensitivity analysis conducted for the Wflow model. It's worth noting that this analysis was performed manually, in contrast to the Monte Carlo sampling method used for the HBV model. The primary focus here is to identify and discuss several parameters that exert a significant influence on the model's performance.

The Wflow model used is based on the SBM principle. The Wflow model contains a larger number of parameters compared to the HBV model. The numerous parameters in Wflow provide flexibility for adjusting the model but also lead to equifinality, as it becomes challenging to determine whether the simulation performs well for the right reasons due to the abundance of parameters. Many parameters can have partially offsetting effects.

Wflow consists of three main modules: the rainfall interception module, the soil module, and the kinematic wave module. The selected parameters primarily influence the second module, the soil module. The tested parameters include rooting depth, porosity, infiltration capacity, hydraulic conductivity, and Manning roughness.

It is noticeable that Wflow struggles to simulate the discharge in the year 2006, consistently underestimating it. By significantly reducing the rooting depth, it becomes possible to better simulate this year; however, this change results in a significant overestimation of all other years. Showcasing that more than the rooting depth lies behind the mechanism of Wflow poor 2006 simulation. Varying other parameters that influence the cumulative discharge and peak height such as horizontal and vertical hydraulic conductivity result in the same problem, by better fitting 2006 the other years are simulated poorer resulting in a less overall performance.

In addition to its effect on cumulative discharge, porosity within Wflow also influences the regression. Increasing porosity allows the soil to retain water for a longer period, which subsequently flows into the river. As a result, Wflow simulates a small discharge peak at the beginning of the dry period when no more rainfall occurs. This peak has not been measured and falls within a period of missing data.

The Manning roughness parameter does not influence the discharge generation but influences the timing of the Wflow discharges, as the roughness influences the ease with which discharges flow over land or through the river cells. As the Wflow models timing is often too late, underestimating the first peak, and overestimating the second peak the Manning roughness is varied to influence the timing of the peak. Even though slight changes are visible, decreasing the Manning roughness did not resolve the missing of the first peak.

Finally, the adjustment of infiltration capacity to vary by month aims to replicate the occurrence of the first peak in the discharge. This involves reducing the infiltration capacity during the dry season and subsequently increasing it following the initial rainfall events. This adjustment is based on the concept of crust formation, where the formation of a crust on the soil surface reduces its infiltration capacity, thereby explaining the lower infiltration capacity observed at the end of the dry season. However, it's important to note that while this process is described in the literature, it is likely not the actual mechanism behind the initial discharge peaks. This skepticism arises because the Wflow model operates on a grid size of 1km x 1km, making it unrealistic that water flowing across such a distance of land would remain un-infiltrated.

The two simulations shown with different cyclic infiltration capacities perform equally when the overall performance over the five years is compared. However, Figure 40 shows the differences between the years 2007 and 2008. It clearly depicts the different performances between the two years, where run a misses the first peak in 2007 but performs well in 2008, and run b does simulate the first peak in 2007 but also simulates an unobserved peak in 2008. Run b is selected as better performing since it is assumed that it is better to predict an unobserved peak than missing an observed peak. This assumption arises from the fact that the catchment is strongly influenced by human interventions, as these interventions can abstract or slow down the first discharge peak by filling up upstream reservoirs.

5.2.3 Hydrological models compared

In this section, the performances of both HBV and Wflow are compared. First, the performance testing values are discussed, namely the NSE and cumulative discharge. This will be followed by an examination of the differences between the two models in general and in more detail yearly. Next, the Budyko framework is delved into, and finally, calibration choices in both models are discussed.

Firstly, let's consider the NSE values of the discharge, which offer a quantitative measure of the agreement between the models and the observed discharge data, using the initial year, 2006, as a warm-up period. The

NSE is a widely recognized performance indicator for hydrological models. The NSE values are sensitive to an offset in the timing of the peaks, and as section 4.1.1 discussed, the daily precipitation correlation between the different precipitation products compared is very low. Therefore, it is very likely that the timing of the precipitation is not equal to reality, and as this has a large influence on the discharge timing of the models, not only the NSE values are investigated but the cumulative discharge as well. In general, NSE values above 0.5 or 0.6 are considered 'well-performing models, however, none of both models score an NSE above this threshold, which shows the difficulty of hydrological modeling in data scares regions.

The cumulative discharge is not sensitive to the timing of the peaks and therefore is a good model performance testing value. However, due to gaps in the data, the cumulative discharge becomes less reliable and sensitive to the timing as well. The cumulative is still a valuable indication of the model performance in addition to the NSE value but should be considered the only or most important criterion.

Based on the calculated NSE and cumulative discharge, the HBV model performs better than the Wflow model. In addition to performing better in terms of NSE and cumulative discharge, the HBV model simulates discharge better in 2006, although the peak is significantly underestimated, while Wflow completely misses the year 2006. Both models accurately simulate the regression of the discharge peak. A noticeable difference between the HBV model and the Wflow model is that Wflow clearly simulates distinct peaks, while the HBV model maintains somewhat higher discharge throughout the entire wet season. The peaks in the Wflow discharge result in higher cumulative discharge than the more continuous lower discharges simulated in the HBV model.

The Budyko framework shows that the observed and modelled discharge closely match in the long term. The Budyko framework does not reveal a clear difference between the different models. All discharges are close to the Budyko curve, which suggests that the long-term water balance of the Ramsa catchment is consistent with the Budyko framework.

The HBV and Wflow models are based on different modeling principles, with HBV relying on the HBV principles and Wflow on the SBM model principles. Therefore, the comparison between the two models is not only a comparison between a lumped and distributed model but also a comparison of the performance of these two model types. This also suggests a potential avenue for further research. Comparing the same model type in both lumped and gridded configurations is interesting. This could serve as a basis for future research, where an initial investigation with a simple lumped model determines the most suitable modeling principle. If needed, this could then be extended to a distributed model.

The investigation of which model principle performs best in a semi-arid region is especially interesting as both model types are developed for a different region, where the HBV model is developed for regions with snow and the SBM model generally performs better in steeper areas as the kinematic wave approach assumes that topography is the main controller of discharge, which might not be the case in less steep areas.

As mentioned in section 5.1, there is a substantial amount of literature available on actual evaporation in and around the study area. For further research, it would be beneficial to delve deeper into this literature and explore the use of remote actual evaporation products to compare with the actual evaporation calculated by the hydrological models. This could provide more insights into evaporation and, consequently, lead to better model calibration.

5.2.4 Validation

The NSE values being in a similar range to those during calibration suggest that the hydrological models maintain consistent performance. The ability of the models to perform well in a period different from the

calibration period is a positive sign. It indicates that the models are not overly specialized for a specific set of conditions and can be relied upon for forecasting and decision-making in various temporal contexts.

In 2014, both models significantly underestimated the discharge compared to the observations. Although the precipitation that year was not unusually high compared to other years, the discharge observations indicate a significantly larger discharge. Due to the high probability of errors within the discharge or precipitation data, it is not feasible to dismiss the model based on poor performance in that year.

5.3 Reservoir water storage

The results of the reservoir water storage simulations are discussed section. The differences in methodology between the two models are elaborated and the influences of these differences are discussed. Finally, the performance of both models is compared and discussed.

In an ideal scenario, the water balance would be perfectly closed, and there would be no upward or downward trend in the simulated water storage in the reservoir if it is not measured. However, this is not the case in the HBV_RWSM. The water balance of the Toece reservoir does not close because the sum of incoming fluxes, especially the inflow into the reservoir, exceeds the outgoing fluxes. Due to the measurement stations being so far from the reservoir and the pattern described in section 5.1.3, where the discharge at Yilou is relatively much lower than at Ramsa, there is a lot of uncertainty in the inflow and outflow of the reservoir.

The Wflow reservoir module determines the outflow not based on measurements at Yilou but based on specified reservoir parameters that determine downstream demand and maximum volume. To make a fair comparison, the constraints in the Wflow reservoir module have also been applied to the HBV model. The simulations of the HBV model improve significantly with these constraints, but the outflow is now greatly underestimated, according to these simulations, causing the reservoir to dry up every year, which is not the case in reality.

The two fluxes that affect the outflow are evaporation and crop water outtake. There is significant uncertainty in crop water outtake, as it is estimated based on the area of each crop. Some crops have surface area data available, while others only have a volume of produced crop. Based on these volumes, an area is estimated, but this estimation involves a lot of uncertainty. In addition to crop water demand, evaporation is a major flux out of the reservoir, and as Van de Giesen et al., 2010 described, the actual evaporation is not equal to the potential evaporation above a large water body if the potential evaporation is measured above land. The uncertainties are tackled by calibrating two scaling factors for both out fluxes in the HBV_RWSM, resulting in the final simulation of the HBV_RWSM.

Although Wflow and HBV_RWSM are based on the same water balance, there are many differences in the methods of the two models. In the Wflow reservoir module, the reservoir is simulated as a square box, which means that the surface area is always constant. The surface area is equal to 0.33% of the maximum area based on the height/area ratio of the Toece reservoir. This assumption leads to an underestimation of precipitation and evaporation for a significant part of the time, because it calculates a volume that is too small. However, as the surface area becomes smaller, Wflow overestimates these two fluxes. The HBV_RWSM takes into account the actual water surface, resulting in a better estimate of evaporation. Precipitation in HBV_RWSM is also calculated with a constant surface area, but it is based on the maximum area, assuming that all precipitation falling within the reservoir contributes to the reservoir volume. As a result, precipitation is greater in the HBV_RWSM than in the Wflow reservoir module. However, this effect is not significant if all other fluxes are much larger than precipitation.

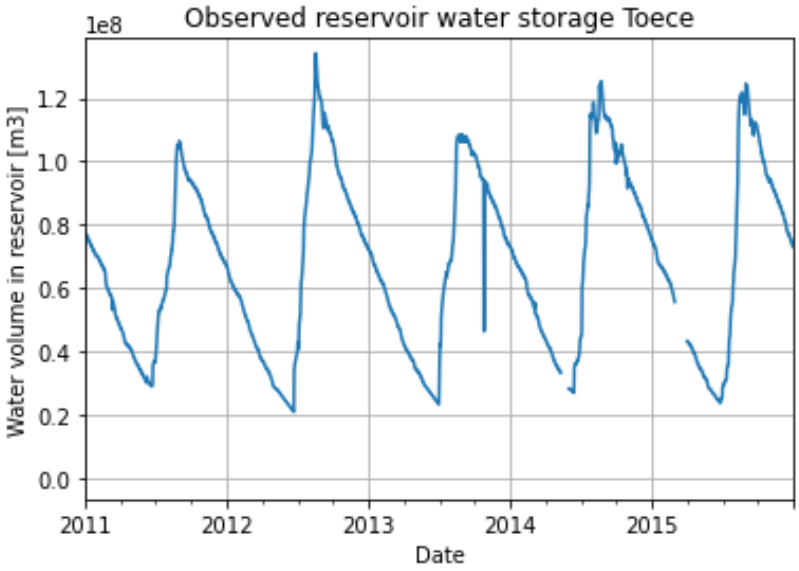


Figure 50: Observed water storage in reservoir Toece during the period of 2011-2016

5.3.1 Reservoir water storage model performance

The reservoir model performance is measured using the NSE value of the storage capacity over time. As depicted in Table 6, the model's warm-up period has a significant influence on the calculated NSE. A warm-up period of one year is justified as it takes one year to mitigate the effect of the initial conditions set within the HBV_RWSM. As discussed in section 4.2, the Wflow model underpredicts the discharge in the year 2006. This underprediction also affects the reservoir water storage simulation, causing the reservoir not to fill up. To reduce the impact of this year, we also looked at the NSE value with a warm-up period of 1.5 years. Both models perform equally well in this scenario.

It's important to note that the observed water storage data is missing at the beginning of the dry season. These periods are not included in the NSE calculations. Figure 50 shows the reservoir water storage in the period 2011-2015, revealing that the regression is nearly linear. This closely aligns with the simulations of Wflow, suggesting that Wflow performs better than the HBV_RWSM.

There is a clear difference in the regression shape of both models. The HBV_RWSM exhibits a convex shape. This can be explained by the fact that initially, the discharge is mainly determined by the outflow and evaporation, and from February/March onwards, crop water demand also contributes significantly to the decline. The Wflow reservoir module simulation shows a more linear/concave shape. The initial steep decline at the beginning of the regression in the Wflow reservoir module simulation is due to the specified reservoir parameters, where the volume is still above the threshold for water to be discharged from the reservoir. Afterward, there is a nearly linear decrease, which can be attributed to the constant surface area of the reservoir in the Wflow reservoir module.

A notable contrast between the two water storage models for reservoirs involves the period during which the reservoir is entirely filled. The HBV_RWSM exhibits a considerably longer period than the other model. Discrepancies in discharge simulations of both models are analyzed in section 5.2.3, explaining this difference. The HBV model produces a steady flow throughout the entire wet season, whereas the Wflow model comprises multiple peaks and low flows. The HBV_RWSM provides a consistent inflow that continually fills the reservoir, a feature that the Wflow model lacks. Consequently, the Wflow model has a shorter maximum fill time.

Next to the regression, the lowest point is of interest as well and this is an important criterion for water shortages. The HBV_RWSM matches the low points more closely than the Wflow reservoir module which is seen as an advantage of the HBV_RWSM over the Wflow reservoir module.

The NSE performance of both reservoir water storage models is similar. However, there are significant differences in the fluxes contributing to the outcome in both models. Investigating the outflow of the reservoir and comparing this with the observed discharges at Yilou shows a very poor relationship between the modeled and observed discharges, with both NSE values of the discharges being negative, with the Wflow reservoir module having the most negative NSE value. As mentioned earlier, there are also many uncertainties in the discharge data at Yilou. Therefore, this poor correlation may not necessarily indicate that the models are not performing well; it could also be attributed to the low quality of the discharge data at Yilou. Due to the lack of reliable data, it cannot be stated with certainty which of the two, if any, accurately simulate the reservoir water storage for the right reasons. However, due to the strong simplifications in the Wflow reservoir module regarding reservoir properties, such as the assumption of a square box for the reservoir, the preference still leans towards the HBV_RWSM because, for these reasons, it appears to be closer to the truth.

5.4 Lessons learned and improvement ideas

This study shows how difficult it is to cope with data scarcity when modelling discharges and reservoir dynamics. However, the study showed, that by extensive data investigation, model adaptations, and simplifications it is possible to simulate the reservoir water storage dynamics decently.

Despite the lack of data, it is possible to simulate the dynamics of reservoir water storage. However, due to the data limitations, it is not possible to determine if this is happening for the right reasons. This makes the next step of modeling the effects of various measures quite challenging. At best, a rough estimate of a relative effect can be made, but even in that case, there will be significant uncertainties. Despite the suggestion that the HBV_RWSM is more reliable, the Wflow model provides the opportunity to implement a wider range of measures in the model.

In this study, however, some limitations and uncertainties need to be acknowledged. One recurring theme in the calibration and performance evaluation of both the hydrological and reservoir water storage models is the lack of discharge data. Collecting additional discharge data is therefore considered the most valuable way to improve the models. For the reservoir water storage models, discharge data directly upstream and downstream of the reservoir is particularly valuable. This eliminates the need for interpolating discharge data from measurement stations to the inflow and outflow of the reservoir, potentially leading to a better understanding of the processes occurring within the reservoir. Further improvements will be discussed in the recommendation section (6.2).

6 Conclusion & Recommendations

6.1 Conclusion

The objective of this research is to gain insight into how reservoir dynamics can be accurately simulated using a hydrological model in data scarce semi-arid regions. To address this objective, four sub-questions were investigated.

What methods can be used to select most reliable model forcing data?
The forcing data considered are precipitation and potential evaporation. the most reliable precipitation product is found by validating two gridded datasets, CHIRPS and ERA5, with local measurements. The gridded datasets are validated based on point-to-pixel correlation, the spatial pattern and the temporal pattern. CHIRPS is found to be the most reliable precipitation product of the considered data products. The local measurements are not used exclusively but rather serve as validation due to the fact that the they are point measurements and time gaps are present. The potential evaporation datasets considered were ERA5, GLEAM_a, GLEAM_b, CRU, GAIPE, and local measurements. As all datasets shows strong differences, the method used in for precipitation does not hold up. The most reliable potential evaporation data were obtained by combining ERA5 and local measurements. This approach leverages the positive aspects of both datasets, i.e. the spatial and temporal pattern of ERA5 is statistically corrected with the values of the local measurements.

Comparing a lumped and distributed model; which simulates the discharge best?
Two distinct hydrological models are used to simulate the discharge: an HBV model (lumped) and the SBM_Wflow model (distributed). Based on the NSE value and the cumulative discharge the HBV model simulates the discharge at Ramsa better. The most likely explanation is the overparameterization of Wflow. Wflow is a distributed model with many model parameters, whereas HBV is a simple lumped model with only eight model parameters. The low amount of parameters in the HBV model is advantageous in data scarce regions, as this ease the calibration of the HBV model, and a Monte Carlo sampling can be used to extensively test all parameters. In the Wflow model, the amount of unknown parameters increase uncertainties as additional estimates and/or assumptions are needed, based on expert knowledge. The Wflow calibration is carried out manually as a Monte Carlo sampling is not possible with this many parameters and the long running time of Wflow.

In conclusion, when the knowledge of the model parameters is insufficient, the incorporation of additional details in the model becomes disadvantageous and increases model equifinality as more solutions can result into similar outcomes.

What is the sensitivity of different hydrological models to parameter choices?
The sensitivity analysis of the HBV model shows that the most sensitive ones are coefficient C_e, exponent$Beta$ and maximum percolation (P_{max}). These three parameters all influence the actual evaporation which shows that correctly simulating the actual evaporation is a key factor in a reliable model. The other parameters all perform well within a wide range, this shows that different parameter sets can result in similar model performance, and thus pointing out there is a degree of equifinality.

For the Wflow sensitivity analysis only several parameters were investigated. As this needs to be done manually instead of with a Monte Carlo sampling (see previous paragraph). The parameters invested are the rooting depth, porosity, infiltration capacity, hydraulic conductivity and manning roughness, which are selected based on the documentation of Deltares (Deltares, 2023). The sensitivity analyse showcases that

different parameter combinations can result in equal model performance, indicating a strong tendency to equifinality as not even all parameters are tested. As not all parameters in the Wflow model are tested, it cannot be concluded with certainty which are the most sensitive parameters. This again shows the advantage of the simpler HBV model.

What is the sensitivity of the different water balance models and which reservoir water storage model simulates reservoir dynamics best?
The sensitivity of the water balance models varies over time as the magnitude of the contribution of each component varies over time. The forced reservoir volume restriction reduces the sensitivity of the discharge, making the water balance less sensitive to the quantity of discharge but mainly to its timing (volume restricting active). Evaporation contributes significantly more to the emptying than precipitation to the filling of the reservoir. Consequently, the sensitivity to evaporation is larger. The same larger impact holds for human abstractions. It seems that, due to the lack of reliable data, it is impossible to determine which, if any, of the models is correct. It looks like they both simulate the water storage well. However, they have significantly different fluxes indicating that both models simulate reservoir dynamics processes differently. As a best guess, based on the more close match with the downstream discharge data at Yilou, the HBV simulates the discharge out of the reservoir better.

In conclusion, based on the performed calibration the HBV model outperforms the SBM_Wflow model in simulating discharge due to its simplicity and ease of calibration. However, this conclusion in influence by the calibration carried out for The sensitivity analyses highlighted the significance of accurately representing actual evaporation. Both water balance models, HBV_RWSM and Wflow, showed equifinality in simulating reservoir dynamics, making it challenging to determine which model is the preferred one. Data limitations remain a significant hurdle in model evaluation, emphasizing the need for additional data collection, particularly discharge data upstream and downstream of the reservoir, to enhance reliability and reduce uncertainties.

Overall, this study shows the difficulties in modelling in data scarce regions.

6.2 Recommendations

The first and most practical recommendation for modelling in data scarce semi-arid regions is to start simple. Beginning with a simple (lumped) hydrological model gives an understanding of the system and the governing hydrological processes, and may provide sufficient information to solve the issue at stake. If this is still insufficient, a high potential improvement may be the inclusion of topography, as suggested by Gharari et al., 2014.

The second recommendation focuses on the data since enhanced reliability of the data likely helps to improve the modelling results. The biggest improvements can be made regarding the discharge data. The availability of a more reliable and continuous discharge time series would enable better calibration and performance testing of the hydrological models. Measuring discharge data both immediately upstream and downstream of the reservoir would be most beneficial, as it also provides a better insight into what happens within the reservoir. Additional precipitation and evaporation data will have less effect as already several products are compared.

Furthermore, to enhance comprehensiveness, a more thorough analysis of the presently existing data can be carried out. Currently, only one dataset per input is considered. A more statistical approach could

generate a new dataset from the available data. Schoups and Nasseri, 2021 introduces a probabilistic data fusion model to achieve water balance closure, which has the potential to improve the current study.

Additionally, an extra model performance test can be introduced. The sensitivity analysis points out that the actual evaporation is an important component to model correctly. Different remote/reanalysis actual evaporation products, such as WAPOR (FAO, 2018), are available which can be used to further assess the correctness of the models as the actual evaporation out of the model can be compared to data.

The study employed a 1-day time step, but adopting a longer time step, such as one month, may prove advantageous. The correlation graph from point-to-pixel depicted in Figure 20 indicates a significantly higher correlation for the monthly time scale. The implementation of a monthly time scale reduces the uncertainty in the daily timing of the input data, which may result in improved accuracy of the hydrological and reservoir water storage model results.

The next recommendation is to conduct a site visit. This will, in particular, offer a better understanding of the uncertainties of the data measurements but will also offer insight into the reality which is modelled, such as understanding the human water abstractions.

There can be reasons one still wants to use Wflow. These may include large heterogeneity within the catchment, requirement of visual presentation of results to nonexperts, testing the effect of (local) measurements to improve water availability. Obviously if less data scarcity exists, Wflow may alse be more suitable. In general, the Wflow model can be improved by improving the model calibration, parameters suggest for further investigation are: soil thickness, M which is the decay of the conductivity with depth, and thataR which is the residual water content. In the case of the Nakambé catchment, adding all the reservoirs in the catchment to the Wflow model could be an effective starting point as human interventions play a severe role in the catchment.

References

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., et al. (1998). Fao irrigation and drainage paper no. 56. *Rome: Food and Agriculture Organization of the United Nations*, *56*(97), e156.

Amisigo, B. A. (2005). Modelling riverflow in the volta basin of west africa: A data-driven framework.

Annor, F., van de Giesen, N., Liebe, J., van de Zaag, P., Tilmant, A., & Odai, S. (2009). Delineation of small reservoirs using radar imagery in a semi-arid environment: A case study in the upper east region of ghana. *Physics and Chemistry of the Earth Parts A/B/C*, *34*(4-5), 309–315. https://doi.org/https://doi.org/10.1016/j.pce.2008.08.005

Anyadike, R. N. (1987). The linacre evaporation formula tested and compared to others in various climates over west africa. *Agricultural and forest meteorology*, *39*(2-3), 111–119.

ASCE. (1993). Criteria for evaluation of watershed models. *Journal of Irrigation and Drainage Engineering*, *119*(3), 429–442.

Brocca, L., Massari, C., Pellarin, T., Filippucci, P., Ciabatta, L., Camici, S., Kerr, Y. H., & Fernández-Prieto, D. (2020). River flow prediction in data scarce regions: Soil moisture integrated satellite rainfall products outperform rain gauge observations in west africa. *Scientific Reports*, *10*(1), 12517.

Budyko, M. I. (1974). Climate and life. *(No Title)*.

Cecchi, P., Gourdin, F., Koné, S., Corbin, D., Etienne, J., & AlainCasenave. (2009). Les petits barrages du nord de la côte d'ivoire: Inventaire et potentialités hydrologiques. *Science et changements planétaires/Sécheresse*, *20*(1), 112–122.

Cecchi, P., Meunier-Nikiema, A., Moiroux, N., & Sanou, B. (2008). Towards an atlas of lakes and reservoirs in burkina faso.

Compaoré, H., Hendrickx, J. M., Hong, S.-h., Friesen, J., van de Giesen, N. C., Rodgers, C., Szarzynski, J., & Vlek, P. L. (2008). Evaporation mapping at two scales using optical imagery in the white volta basin, upper east ghana. *Physics and Chemistry of the Earth, Parts A/B/C*, *33*(1-2), 127–140.

Congalton, R. G., Gu, J., Yadav, K., Thenkabail, P., & Ozdogan, M. (2014). Global land cover mapping: A review and uncertainty analysis. *Remote Sensing*, *6*(12), 12070–12093.

Dai, A., Lamb, P. J., Trenberth, K. E., Humle, M., Jones, P. D., & Xie, P. (2004). The recent sahel drought is real. *International Journal of Climatology*, *24*(11), 1323–1331. https://doi.org/https://doi.org/10.1002/joc.1083

Davies, J. A., & Beven, K. (2015). Hysteresis and scale in catchment storage, flow and transport. *Hydrological Processes*, *29*(16), 3604–3615.

Deltares. (2023). *Wflow*. Retrieved April 6, 2023, from https://www.deltares.nl/en/software-and-data/products/wflow

Dembélé, M. (2023).

Descroix, L., Guichard, F., Grippa, M., Lambert, L. A., Panthou, G., Mahé, G., Gal, L., Dardel, C., Quantin, G., Kergoat, L., et al. (2018). Evolution of surface hydrology in the sahelo-sudanian strip: An updated review. *Water*, *10*(6), 748.

Descroix, L., Moussa, I. B., Genthon, P., Sighomnou, D., Mahé, G., Mamadou, I., Vandervaere, J.-P., Gautier, E., Maiga, O. F., Rajot, J.-L., et al. (2013). Impact of drought and land–use changes on surface–water quality and quantity: The sahelian paradox. *Current perspectives in contaminant hydrology and water resources sustainability*, *2*, 64.

(DWS), D. W. S. (2023). *Rehabilitation large water reservoirs in burkina faso*. Retrieved April 4, 2023, from https://www.dutchwatersector.com/news/rehabilitation-large-water-reservoirs-in-burkina-faso

ECMWF. (n.d.). *Ecmwf reanalysis v5 (era5)*. Retrieved August 1, 2023, from https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5

European Space Agency (ESA). (2009). *GlobCover 2009* (tech. rep.). European Space Agency (ESA). https://due.esrin.esa.int/page_globcover.php

FAO. (2018). *Wapor database methodology: Level 1* (Remote Sensing for Water Productivity Technical Report: Methodology Series). Rome, FAO.

Flint, A. L., & Childs, S. W. (1991). Use of the priestley-taylor evaporation equation for soil water limited conditions in a small forest clearcut. *Agricultural and Forest Meteorology*, *56*(3-4), 247–260.

Fowe, T., Karambiri, H., Paturel, J., Poussin, J., & Cecchi, P. (2015). Water balance of small reservoirs in the volta basin: A case study of boura reservoir in burkina faso. *Agricultural Water Management*, *152*, 99–109. https://doi.org/https://doi.org/10.1016/j.agwat.2015.01.006

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., et al. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data*, *2*(1), 1–21.

Gash, J. (1979). An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society*, *105*(443), 43–55.

Gerald, F., Cofie, O. O., Cecchi, P., & Barron, J. (2019). Characterization of small reservoirs in burkina faso.

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., & Savenije, H. (2014). Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrology and Earth System Sciences*, *18*(12), 4839–4859.

Gharari, S., Hrachowitz, M., Fenicia, F., & Savenije, H. (2011). Hydrological landscape classification: Investigating the performance of hand based landscape classifications in a central european meso-scale catchment. *Hydrology and Earth System Sciences*, *15*(11), 3275–3291.

Grande, E., Zimmer, M. A., & Mallard, J. M. (2022). Storage variability controls seasonal runoff generation in catchments at the threshold between energy and water limitation. *Hydrological Processes*, *36*(10), e14697.

Harris, I., Jones, P., Osborn, T., & Lister, D. (2014). The climate research unit (cru) time-series (ts) datasets: The 50-year, high-resolution, gridded datasets of the university of east anglia. *Norwich, Climatic Research Unit (CRU), University of East Anglia*.

Hoogmoed, W., & Stroosnijder, L. (1984). Crust formation on sandy soils in the sahel i. rainfall and infiltration. *Soil and tillage research*, *4*(1), 5–23.

Ibrahim, B., Polcher, J., Karambiri, H., & Rockel, B. (2012). Characterization of the rainy season in burkina faso and it's representation by regional climate models. *climate Dynamics*, *39*, 1287–1302.

INSD. (2013). *Institut national de la statistique et de la démographie (insd)*. Retrieved April 4, 2023, from http://www.insd.bf/

Jain, S. K., & Sudheer, K. (2008). Fitting of hydrologic models: A close look at the nash–sutcliffe index. *Journal of hydrologic engineering*, *13*(10), 981–986.

Jarvis, A. (2008). SRTM v4 Digital Elevation Model. *CGIAR-CSI SRTM 90m Database*. http://srtm.csi.cgiar.org

Kurtzman, D., Navon, S., & Morin, E. (2009). Improving interpolation of daily precipitation for hydrologic modelling: Spatial patterns of preferred interpolators. *Hydrological Processes: An International Journal*, *23*(23), 3281–3291.

Lavers, D. A., Simmons, A., Vamborg, F., & Rodwell, M. J. (2022). An evaluation of era5 precipitation for climate monitoring. *Quarterly Journal of the Royal Meteorological Society*, *148*(748), 3152–3165.

Legates, D. R., & McCabe Jr, G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water resources research*, *35*(1), 233–241.

Liebe, J., Van De Giesen, N., & Andreini, M. (2005). Estimation of small reservoir storage capacities in a semi-arid environment: A case study in the upper east region of ghana. *Physics and Chemistry of the Earth, Parts A/B/C*, *30*(6-7), 448–454.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed hbv-96 hydrological model. *Journal of hydrology*, *201*(1-4), 272–288.

Logah, F. Y., Obuobie, E., Adjei, K. A., Gyamfi, C., & Odai, S. N. (2023). Capability of satellite rainfall products in simulating streamflows in the black volta basin. *Sustainable Water Resources Management*, *9*(3), 1–14.

Ly, S., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: A review. *Biotechnologie, agronomie, société et environnement*, *17*(2).

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. (2017). Gleam v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al. (2021). Climate change 2021: The physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, *2*.

Meerveld, H., & McDonnell, J. (2006). Threshold relations in subsurface stormflow: 2. the fill and spill hypothesis (doi 10.1029/2004wr003800). *WATER RESOURCES RESEARCH*, *42*(2), W02411.

Mercer, J. (2018). Penman-monteith and priestley-taylor evaporation. https://doi.org/https://wetlandscapes.github.io/blog/blog/penman-monteith-and-priestley-taylor/

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, *10*(3), 282–290.

Neterlands-Agency-Enterprice. (2023). *Develop2build - d2b*. Retrieved April 4, 2023, from https://english.rvo.nl/subsidies-programmes/develop2build-d2b

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

Pereira, L., & Hrachowitz, I. (2011). Crop water requirements. *Earth Systems and Environmental Sciences*, 3275–3291.

Schmengler, A. C. (2011). *Modeling soil erosion and reservoir sedimentation at hillslope and catchment scale in semi-arid burkina faso* (Doctoral dissertation). Universitäts-und Landesbibliothek Bonn.

Schoups, G., & Nasseri, M. (2021). Gracefully closing the water balance: A data-driven probabilistic approach applied to river basins in iran. *Water Resources Research*, *57*(6), e2020WR029071.

Seizarwati, W., & Syahidah, M. (2021). Rainfall-runoff simulation for water availability estimation in small island using distributed hydrological model wflow. *IOP Conference Series: Earth and Environmental Science*, *930*(1), 012050.

Spence, C. (2010). A paradigm shift in hydrology: Storage thresholds across scales influence catchment runoff generation. *Geography Compass*, *4*(7), 819–833.

Trabucco, A., & Zomer, R. (2019). Global aridity index and potential evapotranspiration (et0) climate database v2. figshare. https://doi.org/10.6084/m9.figshare.7504448.v3

Tsendbazar, N.-E., De Bruin, S., Fritz, S., & Herold, M. (2015). Spatial accuracy assessment and integration of global land cover datasets. *Remote Sensing*, *7*(12), 15804–15821.

UNDP. (n.d.). Burkina faso: Strengthening climate information and early warning systems in africa for climate resilient development and adaptation to climate change. *Environment and Energy*.

Van de Giesen, N., Liebe, J., & Jung, G. (2010). Adapting to climate change in the volta basin, west africa. *Current science*, *98*(8), 1033–1037.

Viard-Cretat, A., Galbusera, S., Maneshi, B., Stein, A., Maffa Diarra, V., Aand Coulibaly, Stéphane, T., & Issoufou, O. K. (2019). Report on entrepreneurship in the sahel: Burkina faso catalystas. *Netherlands Enterprise agency*.

Wang, C., Wang, S., Fu, B., & Zhang, L. (2016). Advances in hydrological modelling with the budyko framework: A review. *Progress in Physical Geography*, *40*(3), 409–430.

Wendling, V., Peugeot, C., Mayor, A. G., Hiernaux, P., Mougin, E., Grippa, M., Kergoat, L., Walcker, R., Galle, S., & Lebel, T. (2019). Drought-induced regime shift and resilience of a sahelian ecohydrosystem. *Environmental Research Letters*, *14*(10), 105005.

Willems, P., Mora, D., Vansteenkiste, T., Taye, M. T., & Van Steenbergen, N. (2014). Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes–part 2: Intercomparison of models and calibration approaches. *Journal of Hydrology*, *510*, 591–609.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853.

7 Appendix

7.1 Water levels & Height-Volume relations

7.1.1 Water levels

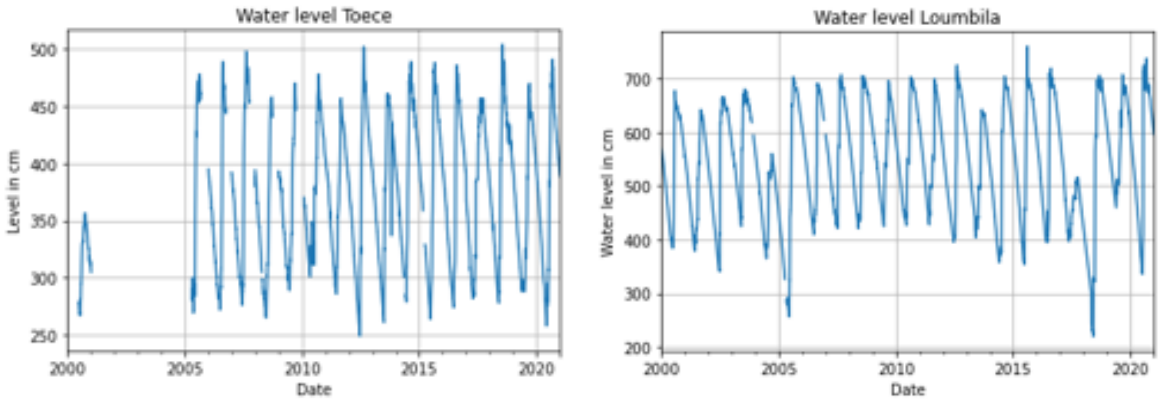


Figure 51: Water level over time in Toece and Loumbila

7.1.2 Height Volume relation

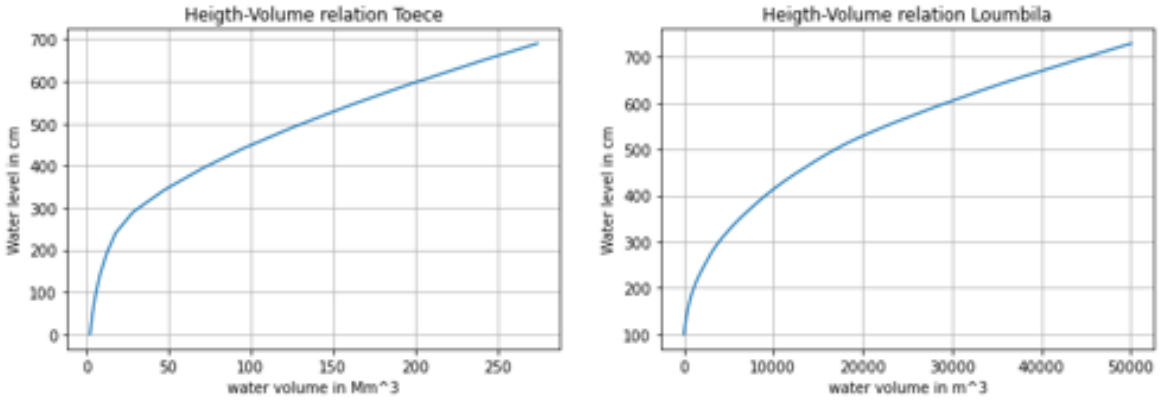


Figure 52: Height volume relations of the reservoirs Toece and Loumbila

7.2 Parameter ranges HBV model

In this appendix the selected ranges that are used in the model will be explained.

- I_{max}: I_{max} is the is the maximum storage capacity of the interception reservoir. Physically, I_{max} is the amount of water that can be stored on the vegetation. Because of the low vegetation in Burkina Faso the range is between 0 and 4 mm.

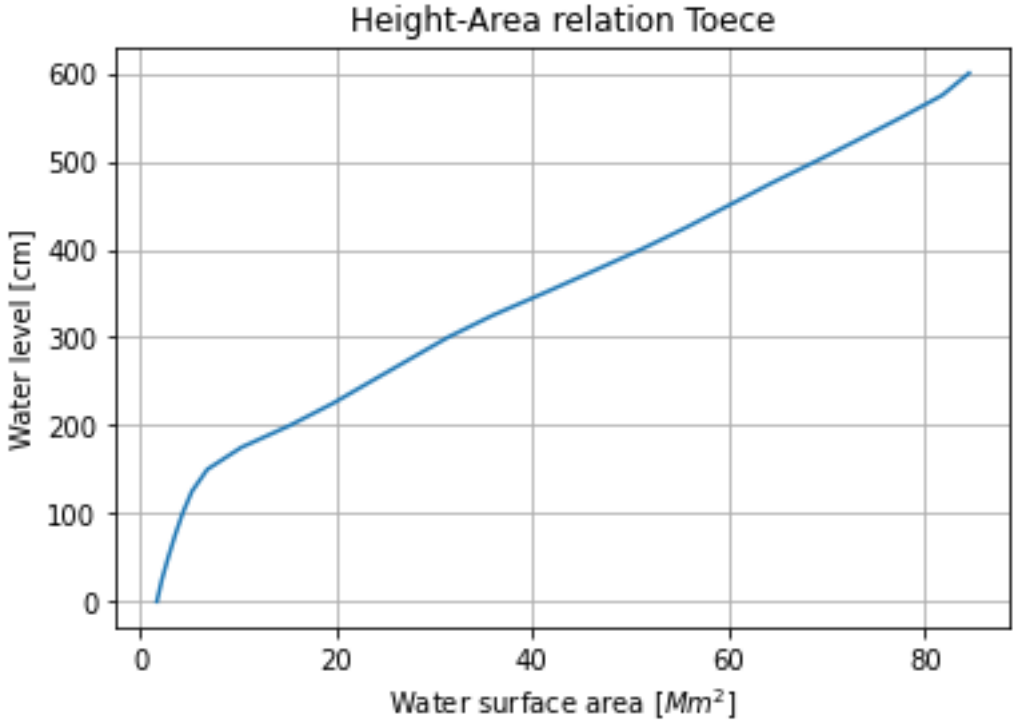


Figure 53: Heigth/Area relation of the Toece reservoir

- **Ce**: Ce describes the relation of evaporation that can occur from the unsaturated reservoir with the following equation:

$$Ea_t = Ep_t * (Su_t/(S_{umax} * C_e)) \tag{10}$$

  The actual evaporation increases with smaller Ce values. Equation 10 shows that Ce cannot be larger than 1 since than the actual evaporation could be larger than the potential evaporation and Ce cannot be zero because this would lead to dividing by zero. There the range varies between 0.001 and 1.

- **Sumax**: Sumax is the maximum storage capacity of the unsaturated soil. The maximum physical possible Sumax can be determined based on the depth * porosity * soil moisture retention.

- β: Beta is a power to the function

$$\rho = (S_u/Sumax)^{\beta} \tag{11}$$

  ρ is then used as fraction of the flow which goes to the unsaturated and to the fast reservoir.

- **Pmax**: determines the amount of percolation from the unsaturated reservoir into the slow reservoir

- **Tlag**: Tlag is used to calculated the lag function of the flow.

- **kf**: kf is the retention coefficient of the fast reservoir. kf is multiplied with the fast storage at a given time to determine the outflow at that given time.

- **ks**: ks is the retention coefficient of the slow reservoir. kf is multiplied with the slow storage at a given time to determine the outflow at that given time.

Reservoir	Water balance equation	Constitutive functions
Interception reservoir	$\frac{\Delta S_i}{\Delta t} = P - Pe - E_i$	$Pe = \begin{cases} 0 & S_i < I_{max} \\ (S_i - I_{\max})/\Delta t & S_i = I_{max} \end{cases}$ $I = \begin{cases} E_{pot} & E_{pot} * \Delta t < S_i \\ S_i/\Delta t & E_{pot} * \Delta t \geq S_i \end{cases}$
Unsaturated reservoir	$\frac{\Delta S_u}{\Delta t} = (1-\rho)Pe - Ea - Perc$	$\rho = \frac{S_u}{S_{u,max}}^{\beta}$ $E_a = min \begin{cases} E_{pot} * \frac{S_u}{S_{u,max} * Ce} \\ S_u \end{cases}$ $Perc = \frac{S_u}{S_{u,max}} * Pmax * \Delta t$
Fast reservoir	$\frac{\Delta S_f}{\Delta t} = \rho * Pe - Q_f$	$Q_f = k_f * S_f$
Slow reservoir	$\frac{\Delta S_s}{\Delta t} = Perc - Q_s$	$Q_s = k_s * S_s$

Figure 54: Equations behind the HBV model for each reservoir

7.3 Equation of HBV model

7.4 Wflow equations

The pseuda-water table at depth z_i such that the value of S at any time is given by:

$$S = (z_t - zi)(\theta_s - \theta_r) \tag{12}$$

Where θ_s and θ_r are the saturated and residual soil water contents, respectively. The division of U into storage U_s and deficit U_d are expressed as follows:

$$U_d = (\theta_s - \theta_r)z_i - U \tag{13}$$

$$U_s = U - U - d \tag{14}$$

The saturation deficit S_d for the soil profile as a whole is defined as:

$$S_d = (\theta_s - \theta_r)z_t - S \tag{15}$$

7.5 Length and crop coefficient of each crop type

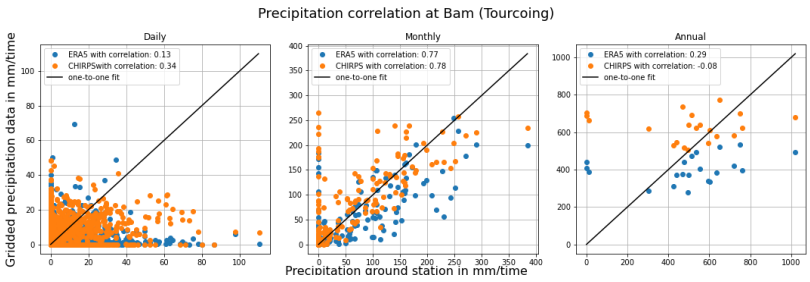
7.5.1 Length of growth stages in days

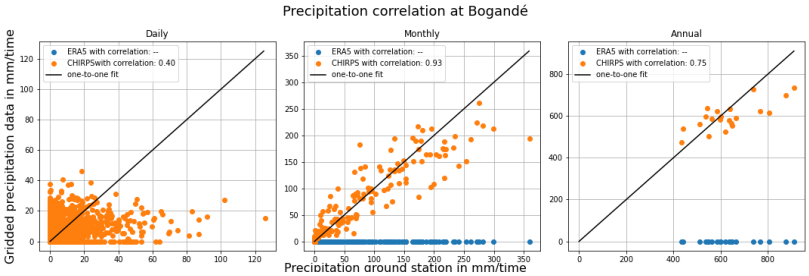
crop type	initial	development	mid season	late stage	Total	sowing date
Onion	20	30	35	40	125	feb
Potato	25	35	45	25	130	feb
Maize	25	40	35	20	120	April
Rice	30	30	70	40	170	feb
Banana	120	60	180	5	365	dec
Eggplant	30	40	40	20	130	oct
Cabbage	25	30	25	15	95	-
Carrot	20	30	40	20	110	Aug
Citrus fruits	60	90	120	95	365	feb
Tomato	15	25	35	35	110	okt

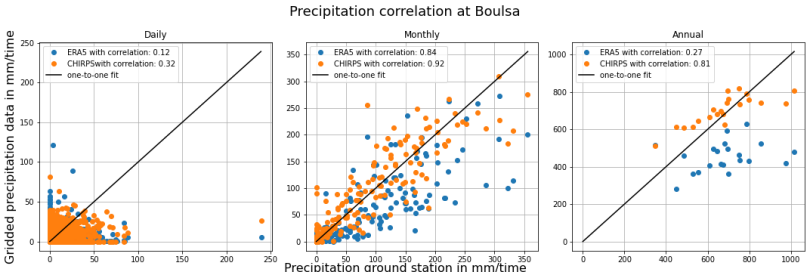
7.5.2 Kc value

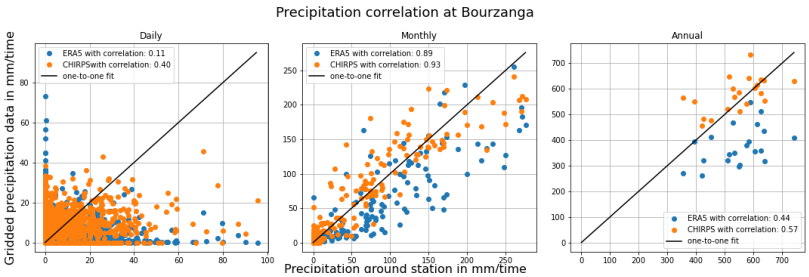
crop type	initial	development	mid season	late stage
Onion	0.5	0.75	1	0.85
Potato	0.35	0.75	1.1	0.9
Maize	0.5	0.8	1.15	0.85
Rice	1.05	1.1	1.2	0.75
Banana	1	1	1.2	1.1
Eggplant	0.6	0.75	1.05	0.9
Cabbage	0.45	0.75	1	0.95
Carrot	0.7	0.8	1.05	0.95
Citrus fruits	0.65	0.6	0.6	0.55
Tomato	0.45	0.75	1.15	0.85

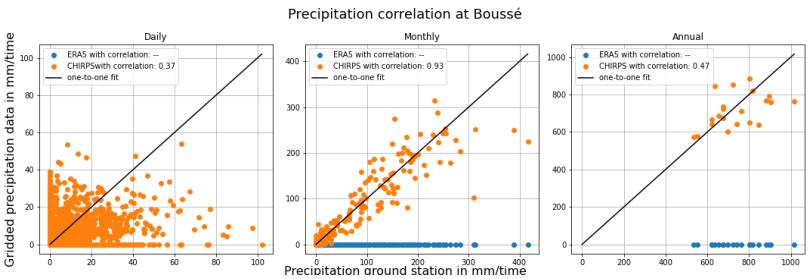
7.6 Precipitation correlation graphs

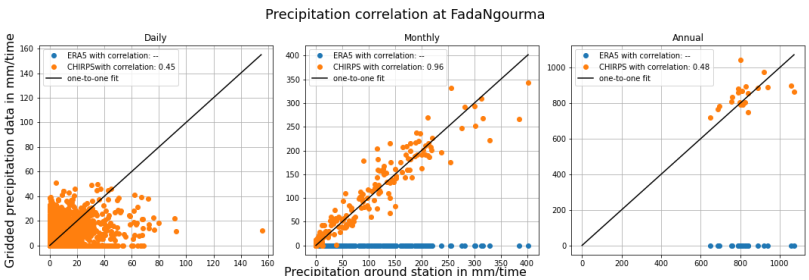


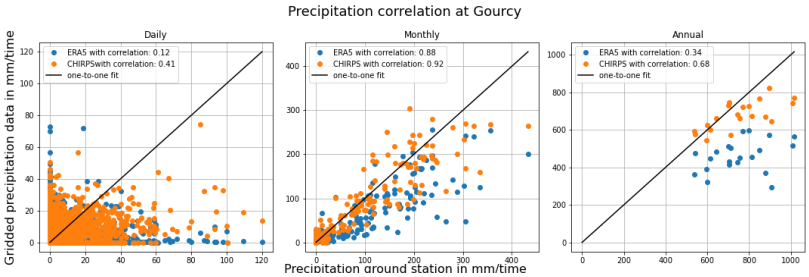

Precipitation correlation at Bogandé
Daily
Monthly
Annual
ERA5 with correlation: --
CHIRPSwith correlation: 0.40
CHIRPS with correlation: 0.93
CHIRPS with correlation: 0.75
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

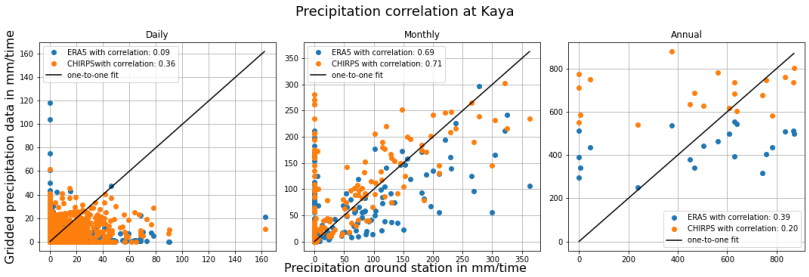

Precipitation correlation at Boulsa
Daily
Monthly
Annual
ERA5 with correlation: 0.12
CHIRPSwith correlation: 0.32
ERA5 with correlation: 0.84
CHIRPS with correlation: 0.92
ERA5 with correlation: 0.27
CHIRPS with correlation: 0.81
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

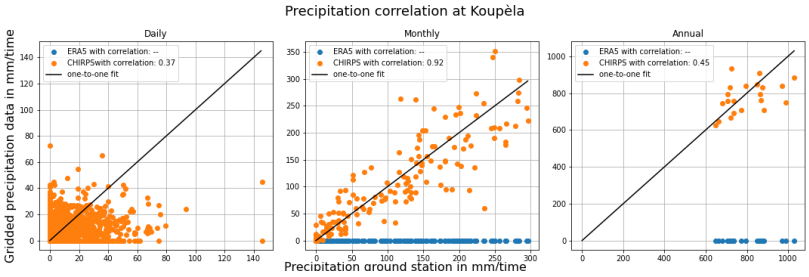

Precipitation correlation at Bourzanga
Daily
Monthly
Annual
ERA5 with correlation: 0.11
CHIRPSwith correlation: 0.40
ERA5 with correlation: 0.89
CHIRPS with correlation: 0.93
ERA5 with correlation: 0.44
CHIRPS with correlation: 0.57
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

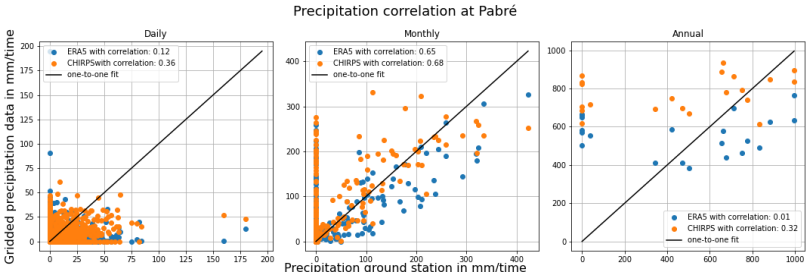

Precipitation correlation at Boussé
Daily
Monthly
Annual
ERA5 with correlation: --
CHIRPSwith correlation: 0.37
CHIRPS with correlation: 0.93
CHIRPS with correlation: 0.47
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

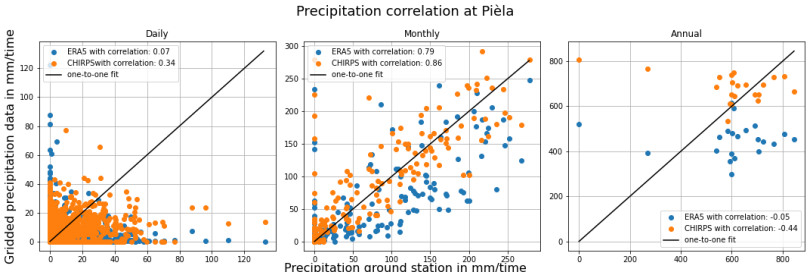

Precipitation correlation at FadaNgourma
Daily
Monthly
Annual
ERA5 with correlation: --
CHIRPSwith correlation: 0.45
CHIRPS with correlation: 0.96
CHIRPS with correlation: 0.48
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

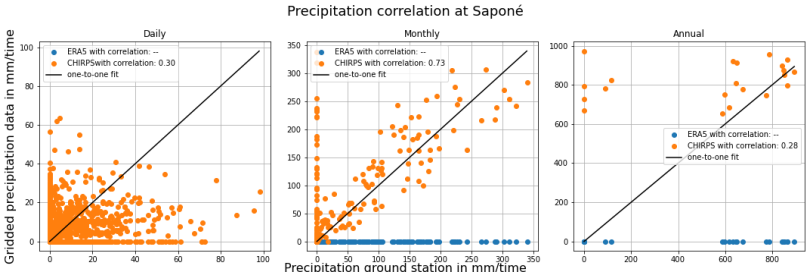

Precipitation correlation at Gourcy
Daily
Monthly
Annual
ERA5 with correlation: 0.12
CHIRPSwith correlation: 0.41
ERA5 with correlation: 0.88
CHIRPS with correlation: 0.92
ERA5 with correlation: 0.34
CHIRPS with correlation: 0.68
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

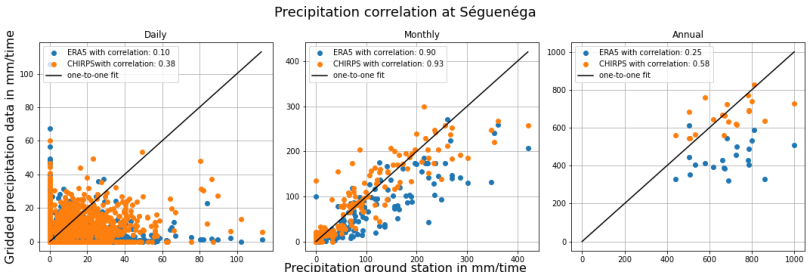

Precipitation correlation at Kaya
Daily
Monthly
Annual
ERA5 with correlation: 0.09
CHIRPSwith correlation: 0.36
ERA5 with correlation: 0.69
CHIRPS with correlation: 0.71
ERA5 with correlation: 0.39
CHIRPS with correlation: 0.20
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

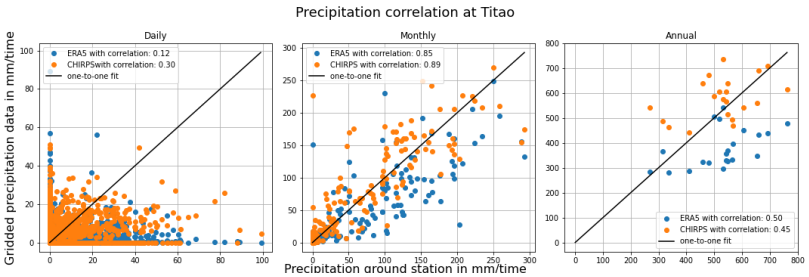

Precipitation correlation at Koupèla
Daily
Monthly
Annual
ERA5 with correlation: --
CHIRPSwith correlation: 0.37
ERA5 with correlation: --
CHIRPS with correlation: 0.92
ERA5 with correlation: --
CHIRPS with correlation: 0.45
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time

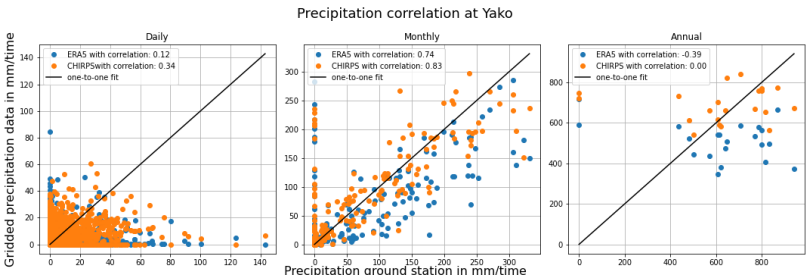

Precipitation correlation at Pabré
Daily
Monthly
Annual
ERA5 with correlation: 0.12
CHIRPSwith correlation: 0.36
ERA5 with correlation: 0.65
CHIRPS with correlation: 0.68
ERA5 with correlation: 0.01
CHIRPS with correlation: 0.32
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time


Precipitation correlation at Pièla
Daily
Monthly
Annual
ERA5 with correlation: 0.07
CHIRPSwith correlation: 0.34
ERA5 with correlation: 0.79
CHIRPS with correlation: 0.86
ERA5 with correlation: -0.05
CHIRPS with correlation: -0.44
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time


Precipitation correlation at Saponé
Daily
Monthly
Annual
ERA5 with correlation: --
CHIRPSwith correlation: 0.30
CHIRPS with correlation: 0.73
CHIRPS with correlation: 0.28
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time


Precipitation correlation at Séguenéga
Daily
Monthly
Annual
ERA5 with correlation: 0.10
CHIRPSwith correlation: 0.38
ERA5 with correlation: 0.90
CHIRPS with correlation: 0.93
ERA5 with correlation: 0.25
CHIRPS with correlation: 0.58
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time


Precipitation correlation at Titao
Daily
Monthly
Annual
ERA5 with correlation: 0.12
CHIRPSwith correlation: 0.30
ERA5 with correlation: 0.85
CHIRPS with correlation: 0.89
ERA5 with correlation: 0.50
CHIRPS with correlation: 0.45
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time


Precipitation correlation at Yako
Daily
Monthly
Annual
ERA5 with correlation: 0.12
CHIRPSwith correlation: 0.34
ERA5 with correlation: 0.74
CHIRPS with correlation: 0.83
ERA5 with correlation: -0.39
CHIRPS with correlation: 0.00
one-to-one fit
Gridded precipitation data in mm/time
Precipitation ground station in mm/time