Model performance of a conceptual hydrological model over the Nile Basin

effect of parameter uncertainty on the performance of PCR-GLOBWB 2

Department of Physical Geography

Faculty of Earth Sciences

Utrecht University

Master Thesis

Erwin van der Haar

5949629

29/05/2023

Contents

List of Figures

List of Tables

Abstract

The Nile River is the major river basin in the northeast of Africa with an ever-growing population dependent on it for its water needs. These needs are only expected to grow in future. Predicting Nile discharge is hard as discharge of the river itself can vary a lot from year to year. Furthermore, Global Hydrological Models, such as PCR-GLOBWB2, also struggle to accurately predict Nile flows and tend to overestimate discharge, most likely due to poor data availability. The Nile River itself can be divided into several basins, with the Blue Nile being the most important in terms of discharge, with 60-70 % of discharge originating here, but this comes mostly in the summer months. The rest of the year the White Nile, a much larger basin with its origins in the region around Lake Victoria, is the main source of water.

To improve model performance of the PCR-GLOBWB 2 it was chosen to improve input data and parameterization. By using updated data soil data (SoilGrids), topographic data (MERIT-DEM) and meteorological data (WFDE5) and using an updated river routing method. The results were validated by GRDC data, but as these are limited, both spatially and temporally it was chosen to seek an additional way of validating data. In this case GLEAM was used to validate evaporation of the model. Combining of both validation methods was used to better understand the discrepancies in model performance for the Nile Basin.

Results still showed poor results for the run without any improvement as expected. It showed that the White Nile performance was very poor, owing to an underestimation of evaporation, mainly in the months of no precipitation. This was observed in some other basin as well. The Blue Nile performance was better but still has overestimations of discharge, especially in the drier months.

Validation of the results with GLEAM showed good options to close the water balance and get better temporal and spatial data despite some of its current limitations.

Introducing new datasets and new parameterization did not improve model results as expected and the poor datasets were not the main cause of the poor performance. The same regions which performed poorly for the standard run also performed poorly for the other runs, indicating an underlying problem. This has to do with the evaporation in the model, especially in the Sudd swamps, where flooding does not occur as expected leading to larger discharge and decreased water storage. Here it was identified that missing model descriptions is a large problem and that this process needs to be better described. For future, many improvements can be made to model and by this, combined with better validation methods, the description of PCR-GLOBWB 2 of the hydrology of the Nile River could be improved.

1 Introduction

The Nile River basin is the main river basin in the northeast of Africa spanning eleven countries from Lake Victoria in the south to the Mediterranean Sea where it ends. It is the longest river in the world with 6670 km and covers 3 255 000 km^2 across several different climate zones: from tropical in the south to arid in the north. The river Nile itself is usually subdivided into three parts with several smaller subbasins (see Figure 1); The White Nile, The Blue Nile which come together at Khartoum and form the Main Nile.

1.1 The Discharge of the Nile

The Blue Nile is most important source when looking discharge, the river originates in the Ethiopian Highlands and contributes about 48.3 km^3 yearly, which is about 60 - 70 % of total Nile discharge. The Atbarah is another major river flowing from the Ethiopian Highlands, contributing about 11.1 km^3 (~13 %) to total flow, showing the importance of the Ethiopian Highlands. However, the flow regime of the rivers originating in the Ethiopian Highlands is strongly seasonal, causing dry seasons and wet seasons. The wet season lasts from June to October and nearly all rain falls in this season. This is illustrated by the fact that average discharge of the Blue Nile per month for the dry season (November-May) is less than 1 km^3 for most months. (Sutcliffe & Parks, 1999).

In the dry season the inflow from the White Nile to the Main Nile at Khartoum becomes important, contributing on average about 2 km^3 a month. This means that from November to May, 70-90 % of total discharge in the Main Nile is from the White Nile. Overall yearly contributions to the Main Nile from the White Nile are 26 km^3 (Conway & Hulme, 1993; Di Baldassarre et al., 2011).

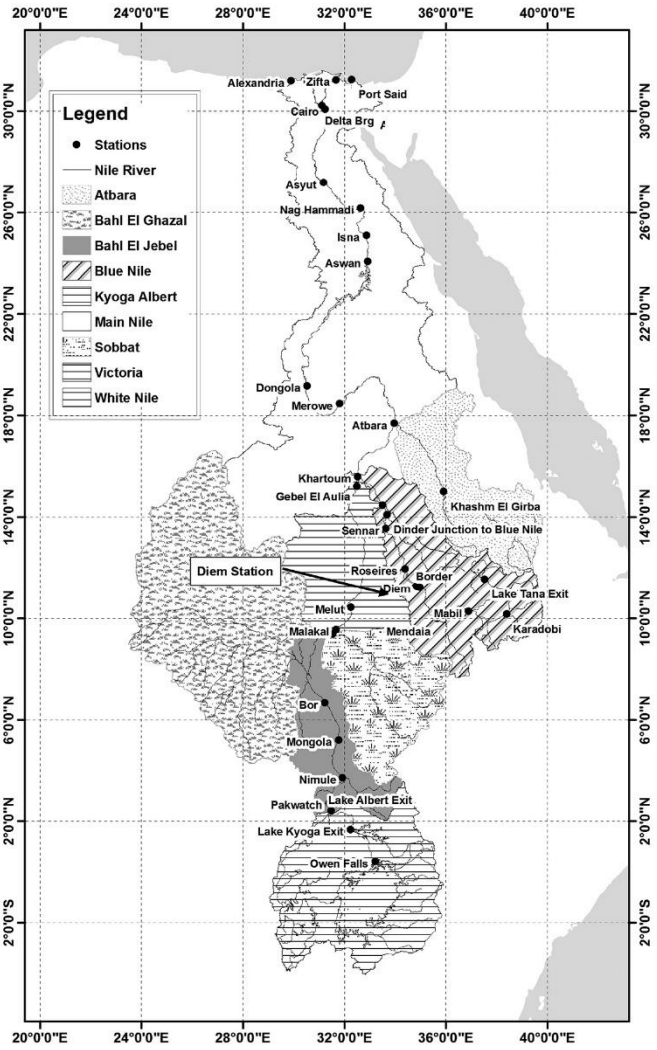

Figure 1: The Nile and its tributaries (di Baldassarre et al., 2011)

The Ethiopian Highlands and other mountainous regions deserve special attention. As these regions are the major source of water for the catchment the Nile is clearly dependent on water towers. With only 45 % of the Nile catchment being mountainous (>1000 m), while more than 90 % of the water originates in these areas (Viviroli and Weingartner, 2008).

Nile discharge is variable throughout the years. Long historical records of annual minimum water levels from 622 to 1284 CE have been studied. Hurst (1951) has shown that the variability in the discharge is not a random process, but rather has some inherent long-term memory (see Figure 2). This has been named the Hurst Phenomenon and is found not only in the Nile River, but in many

natural systems. For the Nile this means that periods of low Nile flow and periods of high Nile flow tend to cluster together to form longer dry and wet periods (Koutsoyiannis, 2005). The El Nino-Southern Oscillation (ENSO) has been hypothesised to be the cause of the Hurst Phenomenon for the Nile. (Eltahir, 1996).

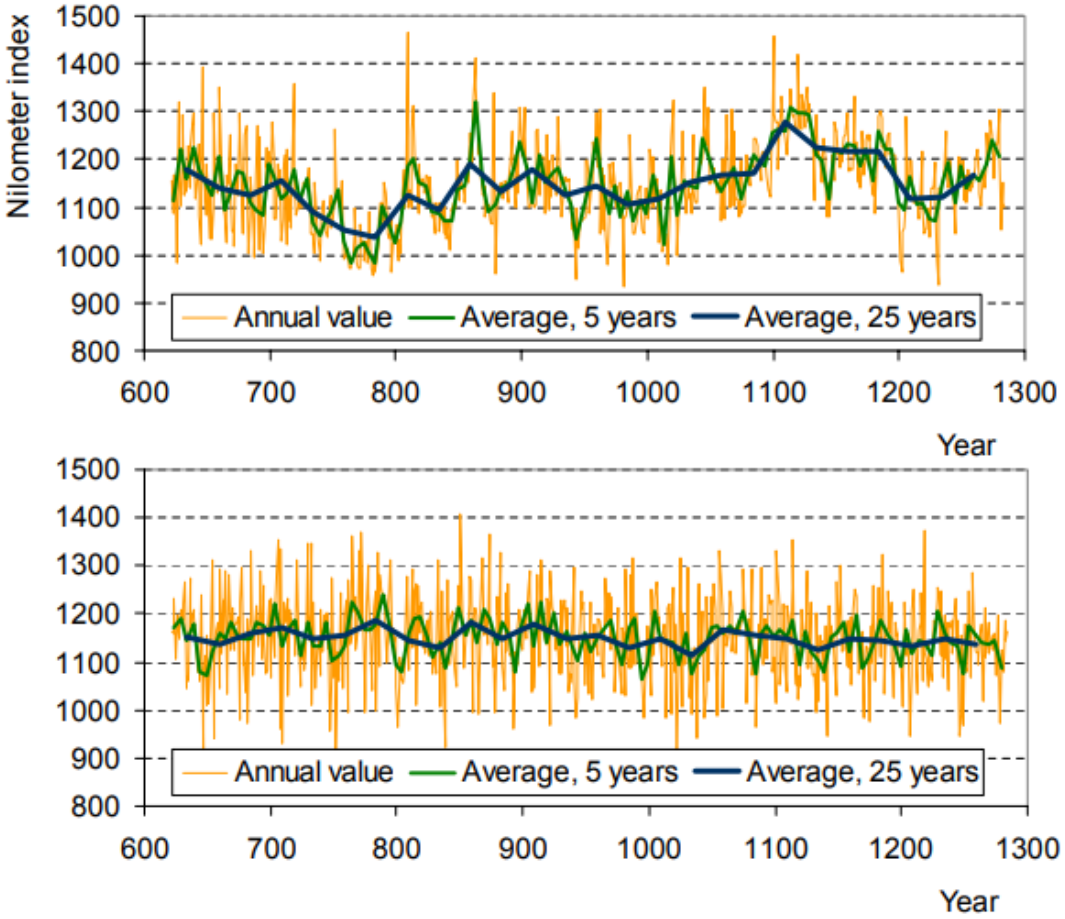

Figure 2: Observed minimum Nile levels and trends (above) compared to random white noise (below) (Koutsoyiannis, 2005)

The projected effect of climate change on Nile River discharges is not certain, but most models predict an increase in total precipitation for the Blue Nile Basin, one of the most important sources. Although the amount of rain days will decrease the intensity will increase (Misiani et al., 2020). The rain season is also thought to last longer and shift from June-September to earlier in the year, which would result in increased discharge and lower peak discharge by 2100. (Roth et al., 2018). However, the exact changes of climate change on the rest of the Nile Basin are uncertain (Di Baldassarre et al., 2011).

1.2 Water Demands on the Nile

The Nile River is the major source of water for the riparian countries. Irrigation demands account for 70 % of water use globally, but this increases up to 95 % for developing countries (Oestigaard, 2012). As the countries in the Nile basin are some of the poorest countries in the world (Di Baldassarre et al., 2011) a large part of their water demand will be irrigation. Table 1 shows the amount of irrigated land of the Nile's riparian countries and from this most of the irrigated land can be found in Egypt, Sudan, and South Sudan. Especially since Egypt and Sudan are located in arid regions with only limited precipitation, so the Nile is the most important source of water for these countries. Egypt relies on the Nile River for 95 % of its primary freshwater usage (AbuZeid, 2020). In Sudan in 2005, which included South Sudan, 85 % of the population depends on water from the Nile (Hamad & El-Battahani, 2005).

Table 1:. Irrigated land area per riparian country in the Nile basin (Abtew & Melesse, 2014)

Country	Area in Nile Basin (km2)	Percent of country	Irrigated land in the basin (ha)	Irrigable land (ha)
Burundi	13,000	46	50	80,000
D.R. Congo	22,300	1	80	10,000
Eritrea	25,700	21	5,800	150,000
Ethiopia	366,000	32	32,100	2,220,000
Egypt	307,900	33	2,923,200	4,420,000
Kenya	52,100	9	9,800	180,000
Rwanda	29,400	83	3,300	150,000
Sudan and South Sudan	1,943,100	78	1,930,300	2,750,000
Tanzania	118,400	13	14,100	30,000
Uganda	238,700	98	4,100	202,000
Total	3,107,600		4,927,830	10,192,000

Population in the Nile region is expected to increase from 400 million to near to 1 billion people by 2050 (United Nations, 2017). Egypt's population increasing from about 100 million to about 150 million and Sudan's population increasing from 40 million to 80 million. Coupled with an only small increase in projected Nile discharge as a result of climate change, will cause increased water scarcity in the region as the same amount of water has to provide this increasing amount of people (Coffel et al., 2019). This is especially troubling as water stress is already high for these regions with dry years already capable of causing damage (Wada et al., 2011).

The unpredictability and seasonality of Nile discharge has been known for some time and has sometimes proved to be disastrous for agriculture in Egypt, which was dependent on the water arriving in summer. Especially as conditions of low flow tend to occur several years in a row (Hurst Phenomenon) To counteract the vulnerability of agriculture to these unpredictable discharges several schemes were developed to improve year-round water availability to Egypt's agriculture (Sutcliffe et al., 2016).

One of these was the possibility of constructing a large reservoir behind the Aswan (low) dam to act as over year storage. The concept of over year storage means that in low flow years the flow can be

supplemented with stored water from the reservoir and that the seasonality of the Nile can be reduced to a steady outflow (Sutcliffe & Parks, 1999). Figure 3 shows that water arriving at Aswan before the construction of the dam and flowing out after Aswan was not stable, with the construction of the Aswan High dam, finished by the 1970s, this changed and outflow became more regular, even though inflow remained irregular. The dam itself causes 10 km^3 of evaporative losses each year, but overall increases water availability to Egypt.

The construction of the Grand Ethiopian Renaissance Dam (GERD) has sparked some controversy in the region. The GERD in the Blue Nile is built on the border of Sudan and Ethiopia. Ethiopia itself hopes the project will help its economy by providing hydroelectric power (Carlson, 2013). Egypt and Sudan are concerned that Ethiopia can dictate access to water resources for these countries as it is constructed on the most important tributary of the Nile. Although the dam can also be used to regulate water in dry periods which are frequent in the Ethiopian Highlands. This will increase the total storage of Nile water for dry periods and increase total water availability for downstream countries if managed well (Wheeler et al., 2020).

1.3 Issues of predicting future water demand for the Nile

Predicting future water availability in the Nile River catchment with regard to changing water demand is necessary to quantify water stress and prepare for effects of population growth and anthropogenic climate change. Information of water availability will be important to ensure access to water resources in future. One such ways of predicting future water needs is by using global hydrological models. Global Hydrological models can look at water availability as a result of climate change. When this information of water availability is linked to water resource models it is possible to look at effects of socio-economic developments and feedbacks. Various models such as WGHM (Döll et al., 2003a), Mac-PDM.09 (Gosling & Arnell, 2011) and PCR-GLOBWB 2 (Sutanudjaja et al., 2018) and more (e.g., Bierkens, 2015; Sood & Smakhtin, 2015) already exist which do just that. These models all simulate the global terrestrial water cycle at a catchment scale, using a gridded model that calculates the water balance. Models like this require input data to solve the water balance. PCR-GLOBWB 2 (Figure 3), for example, is composed of several modules. These modules require data like soil types, time series of climate data, crop data from the FAO and more from which they can produce water balance calculations and results.

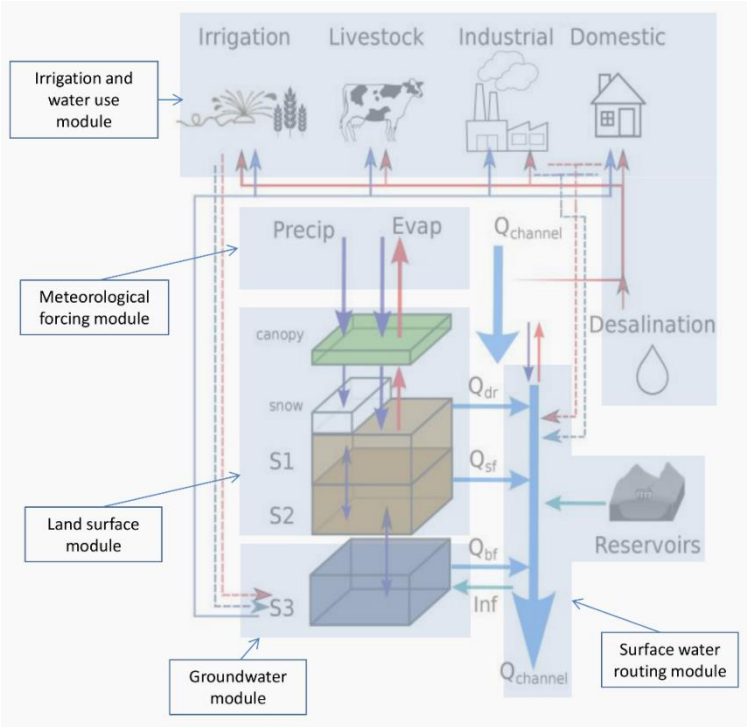

Figure 3: PCR-GLOBWB 2 model structure (Sutanudjaja et al., 2018)

However, for the Nile and similar semi-arid basins (e.g., Niger), GHMs tend to overestimate discharge significantly. Such an overestimation is reported for PCR-GLOBWB that is applied without any calibration and amounts to an overestimation of almost one order of magnitude van Beek et al., 2011). Similarly, the WGHM model, which is tuned by adjusting runoff coefficient to get discharges within 1 % of observed discharge values, still reported overestimations of discharge in the Nile Basin at the maximum accepted runoff coefficient. This hints at a possible model incompleteness for processes in the Nile (Döll et al., 2003).

Several hypotheses for the poor performance of GHMs for the Nile River Basin have been proposed. Van Beek et al., (2011) argues that runoff generation might be overestimated as evaporation from ponds and other sources might be underestimated, which is also reported in Döll et al., (2003) for the WGHM model. Underestimation of evaporation in the Sudd region might also contribute to this lack of evaporation and overestimation of discharge for the entire length of the river. Furthermore, runoff depth overestimation already occurs in the headwaters of the Nile and persists along the entire course van Beek et al., 2011; van Wirdum, 2017).

Another issue for modelling the Nile River basin is the lack of data. Firstly, validating model performance, usually done by comparing river gauges to model discharge, can only be carried out very limited as discharge data is only available up till 1984 in most cases (di Vittorio & Georgakakos, 2021). This is reflected in the work of (van Wirdum (2017) where only 22 suitable gauging stations were deemed suitable for determining Nile performance for PCR-GLOBWB 2. This gives a very limited view of the spatial performance of a model.

Secondly, the input data is often of poor quality, owing to limited research conducted in many of remote parts of the basin. Added to this the large scale of the model as well as the variability between the different subbasins makes PCR-GLOBWB 2 struggle to accurately simulate Nile flows (van Wirdum, 2017). Important inputs for PCR-GLOBWB 2 such as precipitation are based on in situ gathered data. However, these are limited providing poor input for the model as well as the quality of this data not being the same for all the subbasins.

Improving model performance for the Nile requires: evaluating available data that can be used for validation, investigating new potential ways of validating the performance and seeking improved and more recent datasets for forcing.

As no new public discharge data is likely to be released due to the political situation in the region (di Vittorio & Georgakakos, 2021), other ways of validating model performance will have to be investigated. Earth Observation (EO) can be used to cover the Nile Basin more fully. Sadly, discharge cannot directly be inferred from this type of data. However, other parts of hydraulic cycle can be measured this way. Evapotranspiration is one of these parameters that is very suitable for validating model results with by using e.g. the GLEAM dataset (Martens et al., 2017). The GRACE satellites are already used to validate water storage in model studies, such as in Sutanudjaja et al., (2018). However, resolution of GRACE is coarse and the roughly ten years' operational history is relatively short (Tangdamrongsub et al., 2017) and this makes it impossible to validate both with GRACE and with discharge data for the Nile Basin.

Optimizing parameters is a way of increasing model results. By using more recent and accurate datasets, which are closer to the reality will model performance may increase. Especially in regions with poor in situ data availability improved datasets which use newer computational and AI techniques might yield better results.

1.4 Objective

The main objective of this thesis will be to determine the cause of the poor performance of PCR-GLOBWB as an exponent of other GHMs over the Nile Basin. The performance will be evaluated in terms of differences between model results and validation data, using observed discharge at gauging stations for the primary validation and using earth observation data as a secondary source in order to cover the Nile Basin in full over a more recent period. This may reveal some processes which are not incorporated in the model but do contribute to the simulated overestimation of discharge. The secondary objective of this research is to explore the possibility of improving model performance by incorporating alternative datasets on forcing (WFDE5), soil data (SoilGrids) and terrain characteristics (MERIT-DEM) as well as using different ways of routing. The premise of this analysis is that the process description in PCR-GLOBWB 2 is adequate and can be adapted by improved parameterization that still falls within the bounds of likely parameter values. In case this does not hold, alternative process descriptions can be proposed.

Eventually the results of this thesis may hopefully be a guideline for increasing model performance not only for the Nile Basin on PCR-GLOBWB 2, but also for other GHMs over data scarce regions elsewhere in the world.

1.5 Research Questions

- What is the cause of poor performance of discharge generation in PCR-GLOBWB 2 when compared to validation data?

- What is the potential of using EO data for closing the spatial and temporal gaps in measurements for the Nile River Catchment as well as improving process understanding by providing more information?

- What is the effect of introducing different datasets on model performance for PCR-GLOBWB 2 for the Nile River Catchment?

-What is the effect of using the KWE for the routing module for predicting flooding behaviour in the Nile Basin?

- Which processes, if any, are additionally required to be modelled in the PCR-GLOBWB 2 framework to overcome the difference between model results and validation data?

1.6 Approach

To fully answer the research questions the following chapters were made. The following chapter will consist of three parts. First, an explanation of the hydrology of the Nile Basin will be explained. This will focus on the several subbasins to determine the unique behaviour of each basin separately as well as its contribution to the main Nile flow. Second, an explanation of global hydrological models and their historical development will be given. PCR-GLOBWB 2 will be examined here as well as an exponent of hydrological models. Furthermore, this chapter will also focus on the new datasets that will be introduced to try and improve model performance.

Chapter 3 details the methodology, starting with the way of validating the model by both river discharge data as well as earth observation data, in this case evaporation. Followed by an explanation of the five different model runs; the standard run to which the rest of the runs will be compared, the three runs with new datasets and the one run which has a different routing method.

Chapter 4 shows the results of the analysis performed for the five different runs. Both in terms of the comparison to the discharge data per subregion of the Nile and in a spatial comparison of evaporation data for the whole Nile basin.

Chapter 5 discusses the results of the analysis. This will be used to explore the model performance and to identify regions and processes where the model performance is not optimal. This chapter will also offer a discussion on the evaluation using GLEAM and offers some suggestions for improving model performance for the Nile Basin.

This thesis concludes with the overall findings in Chapter 6.

2 Literature Chapter

2.1 The Hydrology of the Nile River Basin

The Nile River is the longest river in the world at about 6700 km. It extends over eleven countries: Egypt, Sudan, South Sudan, Ethiopia, Eritrea, Uganda, Kenya, Tanzania, Burundi, Rwanda, and the Democratic Republic of Congo. This covers a wide variety of climatic zones, from Tropical climates in the South to Arid climates for the Northern parts of the Nile.

These different climatic conditions have an impact on the distribution of the precipitation in the catchment, which in turn has an impact on the runoff generated. And is highly variable within the catchment. With high precipitation mainly occurring in the southern regions and the Ethiopian highlands and almost no precipitation in the northern regions.

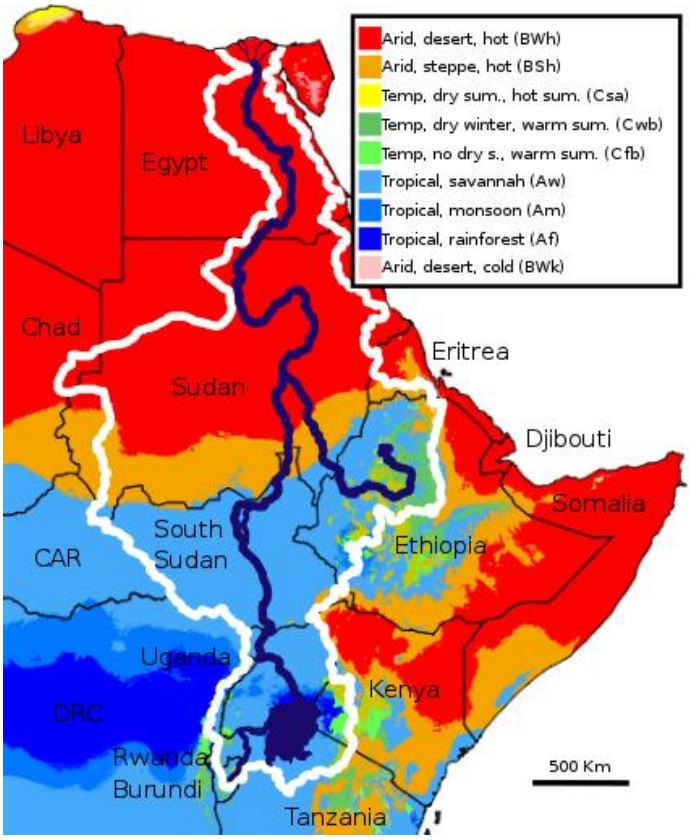

Figure 4: Climate zones in the Nile Catchment (Beck et al., 2020)

Eight subbasins contribute to Nile flow. The Blue Nile and the White Nile basin are most important for Nile flow, but other basins also have an impact on Nile flow. Therefore, all the basins will be described individually and linked to the flow they contribute to the next section or Main Nile. The flow data is based on gauging stations in the whole Nile catchment, with some gauging stations being active for longer than others. Sadly, since 1984 no public river discharge data was made available (di Vittorio & Georgakakos, 2021). Nile flow is highly variable between years, and the discharge presented in this chapter is based on long term averages.

2.1.1 The Subbasins

2.1.1.1 Lake Victoria

The southernmost sources of the Nile are in the Lake Victoria region. The Mufumbiro mountains in Rwanda and Burundi drain into the Kagera river, which is one of the rivers feeding Lake Victoria. These mountains act as water towers, but precipitation is high in the lowlands as well in this region. There are various tributaries to the lake, the Kagera River being the largest, and providing about 6km^3 yearly to Lake Victoria However, precipitation over the lake is the most important source of water, accounting for about 85% of all gain to Lake Victoria (Sutcliffe & Parks, 1999). Peaks in precipitation occur in March-May and November-December and precipitation is highest over the lake itself (Conway & Hulme, 1993). The effect of Lake Victoria to total Nile flow is to provide a baseflow. Measured outflow at the Owen Fall Dam (since 1951) shows a steady outflow of 28.6 km^3 per year (Figure 11). Precipitation amounts increased since 1960 which coincides with an increase in lake levels and in turn an increase in discharge from Lake Victoria Figure 5).

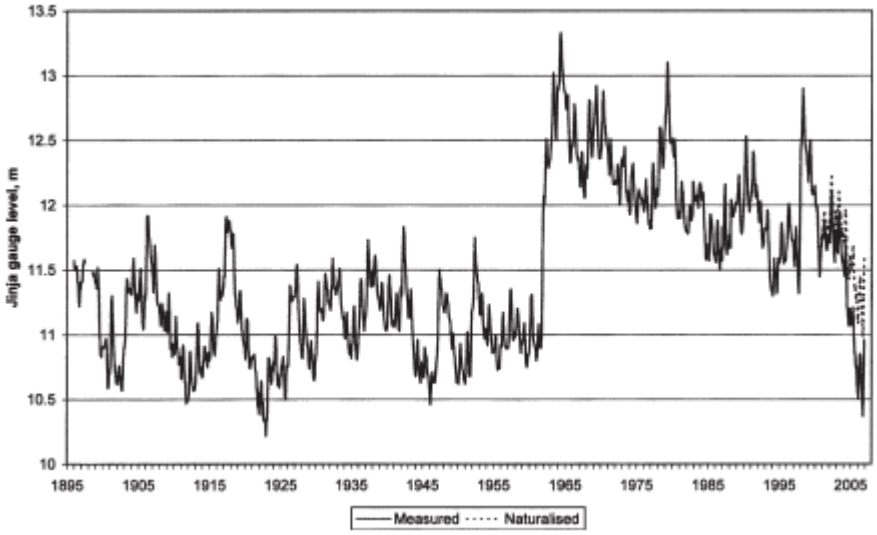

Figure 5: gauge level at outflow of Lake Victoria as an analogy of discharge (Sutcliffe, 2009)

2.1.1.2 Northern Great lakes: Lake Albert & Lake Kyoga

The water from Lake Victoria flows north into Lake Kyoga, where generally lake evaporation is balanced by inflow and precipitation (Conway & Hulme, 1993). From here it flows into Lake Albert before leaving the area into the Sudd. Lake Albert is also fed by waters coming from Lakes Edward and George and the Ruwenzori mountains through the Semliki River as well as from other tributaries. Even though high evaporation rates are expected in Lake Albert, outflow of Lake Albert into the Bahr el Jebel is higher than outflow of Lake Victoria. The increase in discharge is thought to be due to the Semliki River feeding Lake Edward (Sutcliffe & Parks, 1999). In total about 32.8 km^3 a year leaves Lake Albert into the Albertine Nile (Figure 11).

2.1.1.3 The Sudd/Bahr el Jebel

The Albertine Nile comes together with the Aswa River, adding 3 km^3 yearly, and becomes the Bahr el Jebel. At Mongalla discharge is measured at about 36km^3 yearly, upstream from Mongalla the river enters the Sudd Swamps. This region is hydrologically important as it is thought that about 50% of flow entering the swamps evaporates. The swamps themselves are in a very flat area with a slope of only 0.1 m/km (Conway & Hulme, 1993). There is one rainy season from April to November and yearly precipitation ranges from 900 mm/year in the south to 800 mm/year in the north. Temperatures are high; 30-33°C in the dry season and 26-28° C in the rainy season These high temperatures account for the high evaporation rates in this area (Mohamed et al., 2006). This evaporation is high year-round, as shown by water balance calculations of Mohamed et al. (2004) in Figure 6. Here evaporation from SEBAL algorithm based on satellite observations were compared to long-term averages of precipitation from Sutcliffe & Parks (1999). There is a distinction to be made between the permanent swamps and the seasonal swamps. The permanent swamps are close to the

river, whereas the seasonal flooded swamps are further away. Since the 1960s the extend of the permanent swamps has tripled (by 1999) due to an increase of Lake Victoria levels which has increased the area of the swamps to about 30 000 - 40 000 km^2 (Sutcliffe & Parks, 1999). This shows that the permanent swamps are mainly fed by upstream water entering the low area. The seasonal swamps are fed by rainwater ponding in the rainy season and seasonal torrents adding to river volumes between the Lake Albert outflow and Mongalla, where the discharge measurements are more seasonal than at the exit of Lake Albert (Sutcliffe, 2009). Yet, backwater effects of the Blue Nile floods and other barriers further downstream on the White Nile also add to the increased flooding in these regions (Sutcliffe & Brown, 2018).

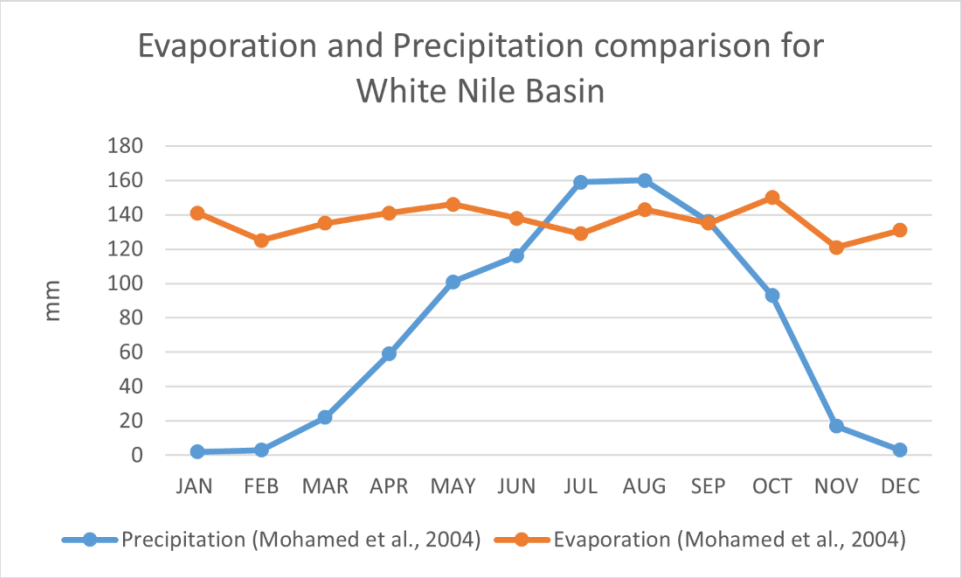

Figure 6: Monthly precipitation (Sutcliffe & Parks, 1999) and evaporation from SEBAL algorithm combined in Mohamed et al. (2004) *for the Sudd Swamps.*

Although some water flows back into the main channels after the flooding most of the water is lost due to evaporation in the Sudd. Recorded losses due to evaporation are on average 47.2 % between 1905 and 1960. As a result of the higher inflows and increased extend of flooding, more water was lost in the period 1961-1983 where average losses were 57.7 %. The construction of the Jonglei canal to bypass the Sudd and reduce losses was started in the 1950s, but never finished. However, if it was implemented, the decrease in flow to the Sudd would decrease swamp extend (Sutcliffe & Parks, 1999). Outflow at the end of the Sudd Swamps is measured at 16.1 km^3 a year, less than half the inflow at Mongalla.

2.1.1.4 The Bahr el Ghazal

The Bahr el Ghazal joins the Bahr el Jebel from the west in the northernmost reaches of the Sudd swamp. The basin is the largest of all the subbasins and has an average rainfall of about 1200 mm a year, which is relatively high. However, the contribution of the Bahr el Ghazal to the flow of the Nile is minimal. As with the Sudd, water spills into flat floodplains and swamps where much of the water evaporates (Conway & Hulme, 1993). There is a possibility of the waters from the Bahr el Ghazal swamps spilling into the Sudd Swamps, connecting the two areas (Di Baldassarre et al., 2011). Overall, contributions of water from the Bahr el Ghazel to the flow of the Nile are minimal (0.3 km^3 a year) due to the high precipitation rates due to the evaporation in the swamps and floodplains as can be seen in Figure 7.

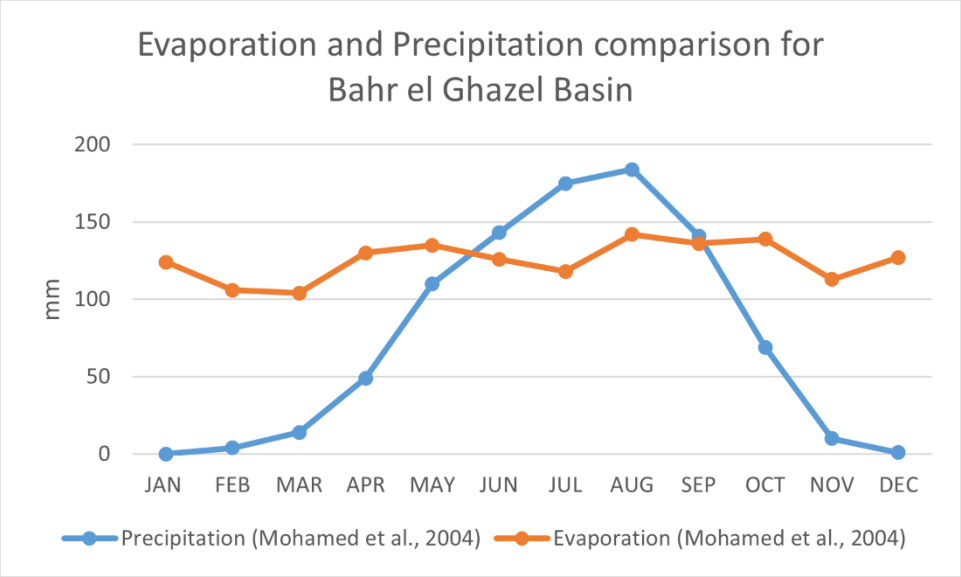

Figure 7: Monthly precipitation (Sutcliffe & Parks, 1999) and evaporation from SEBAL algorithm combined in Mohamed et al. (2004) *in the Bahr el Ghazel basin.*

2.1.1.5 The Sobat

The Sobat River emerges from the confluence of two major tributaries: The Baro and the Pibor. These two rivers have very different origins. The Baro drains a part of the western Ethiopian Highlands and adds more water in total. It is also seasonal with higher discharges between April and October, coinciding with the rainy season of the Ethiopian Highlands. Yearly precipitation ranges from 1300 mm in the plains to 2400 mm in the Ethiopian Highlands. However, the Baro spills into a plain at the foot of the Highlands and feeds the Machar Marshes (Conway & Hulme, 1993). This has the effect that a large part of the water which enters this plain and the marshes there evaporates.

The Pibor drains the flat plains east of the Bahr el Jebel and down to the Ugandan and Kenyan borders. Here flow is less seasonal as precipitation falls more spread out over the year. Furthermore, precipitation is lower than for the Baro river catchment, with only about 800 – 900 mm a year falling in this flat and warm area.

The Baro contributes more water to the White Nile through the Sobat than the Pibor. The Baro contributing 9.5 km^3 a year and the Pibor only contributing about 3.25 km^3 (Sutcliffe & Parks, 1999). However, due to the floodplains and marshes present in these tributaries, some losses occur before the confluence with the main Sobat. Exact data is unknown, but estimates range from 30 % loss for the Baro and 14 % for the Pibor (Conway & Hulme, 1993).

From the confluence of the Baro and Pibor the river is known as Sobat and merges with the White Nile. It adds 13.7 km^3 a year to total volume of flow, which is about half of White Nile flow at this point. The total water balance of the Sobat shows high precipitation, especially in the summer months and high year-round evaporation (Figure 8).

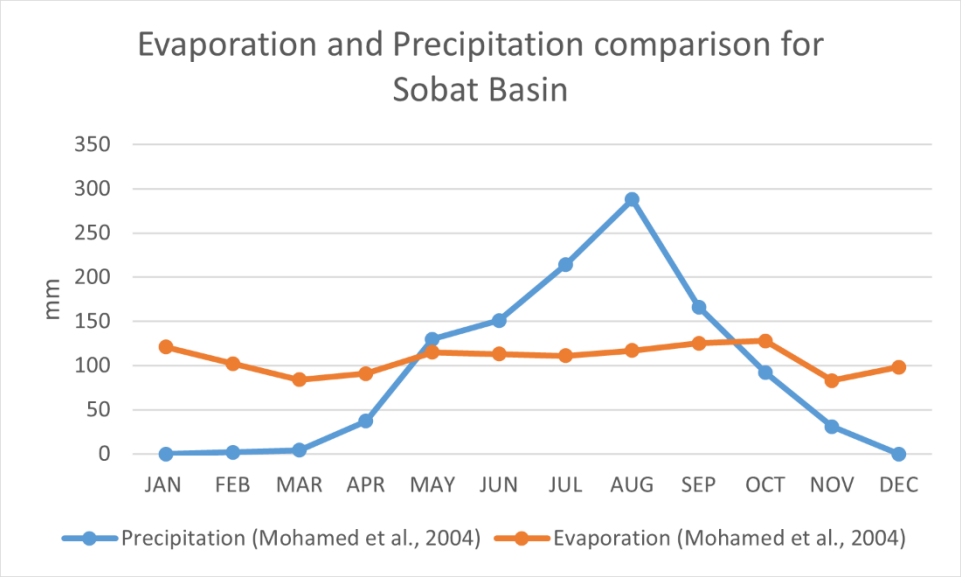

Figure 8: Monthly precipitation (Sutcliffe & Parks, 1999) and evaporation from SEBAL algorithm combined in Mohamed et al. (2004) *in the Bahr el Ghazel basin.*

2.1.1.6 The White Nile

Even though, the White Nile is called so after the confluence of the the Bahr el Ghazel and the Bahr el Jebel, it is only past the confluence with the Sobat that the river leaves the swamps of the Sudd and becomes different. To illustrate this, the White Nile between Malakal and Khartoum will be discussed separately here.

Between the mouth of the Sobat and the confluence with the Blue Nile the river falls 13 m over a reach of 840km and thus is in flat land (Sutcliffe & Parks, 1999). As the river flows north the climate becomes more arid and open water evaporation increases

Inflow into this stretch of the White Nile is from the Sudd swamps, showing a reduced signal of Lake Victoria levels, and the Sobat River, which adds seasonality to the White Nile flow. The gauging station at Malakal shows a different pattern than the combination of the Sobat and Sudd outflow would give in figure 11. This is most likely due to the different dates of measurement, 1905-1983 vs 1907-1995 (Sutcliffe & Parks, 1999). Only in exception years do other tributaries add to this stretch of the White Nile. Observed differences between inflow and outflow between Malakal and Mogren are measured to be about 2 km^3 a year, which is a reduction due to evaporation. The construction of the Jebel Aulia dam adds another 2.5km^3 a year to measured evaporation (Conway & Hulme, 1993). In total White Nile contributions contributing to main Nile flow at Khartoum are 26 km^3 (Sutcliffe & Parks, 1999).

2.1.1.7 The Blue Nile

The Blue Nile is the most important tributary of the Nile when considering volume of flow. The water originates from the Ethiopian Highlands with elevations of 2000 – 3000 m, with some peaks of over 4000 m (Sutcliffe, 2009a). The source of the Blue Nile is the Little Abbay which drains from part of the Ethiopian Highlands into Lake Tana. From Lake Tana the Blue Nile flows into the Blue Nile Gorge for 900 km until it reaches Sudan. Several other tributaries, such as the Dabus and Didessa, join the Blue Nile in the gorge. Downstream of the Roseires dam in Sudan the river is joined by two more rivers; the Rahad and the Dinder. Both these rivers also have their headwater in the Ethiopian Highlands (Conway & Hulme, 1993).

Average yearly precipitation of the Blue Nile Catchment up to Roseires is 1600 mm. Local precipitation increases southwards, with 1000mm a year near the Ethiopian-Sudanese border to 1800 mm in the 'loop' of the Blue Nile from Lake Tana, to about 2400 mm in the upper reaches of the Didessa river. The rainfall has a strong seasonality with the migration of the ITCZ, the rainy season is therefore between July and October and 90 % of the runoff occurs between these months (Sutcliffe & Parks, 1999).

As most runoff is generated in the rainy months, the Blue Nile adds a strongly seasonal signal to Main Nile flow. In the lowlands of Sudan between the Ethiopian border and Khartoum some losses occur, and discharge of the Blue Nile is 0.5 km^3 less than at the entrance of the Roseires reservoir across the Ethiopian border. There is a loss of water here even with the contributions of the Dinder and Rahad rivers. One of the main reasons is the Roseires reservoir, causing high evaporation as well as it being used for irrigation. At the convergence with the White Nile at Khartoum roughly 26 km^3 yearly is added by the White Nile, spread more averagely over the year. The Blue Nile adds 48.3 km^3 a year, mainly in the rainy season. The combined discharge after Khartoum shows this seasonality strongly, with a baseflow of White Nile flows (Figure 9) (Sutcliffe & Parks, 1999).

2.1.1.8 The Atbarah

The Atbarah, or sometimes called the Black Nile, is the northernmost tributary to feed into the Nile. The river is ephemeral and only flows during the rainy season of the Ethiopian Highlands. From here it flows north and meets the Main Nile at the town of Atbarah. As it leaves the Highlands of Ethiopia no further water is gained to the Atbarah, rather, only water is lost through evaporation as the river runs through semi-arid and arid climates. Yearly contribution of the Atbarah to Nile flow is 11 km^3 a year making it a significant tributary.

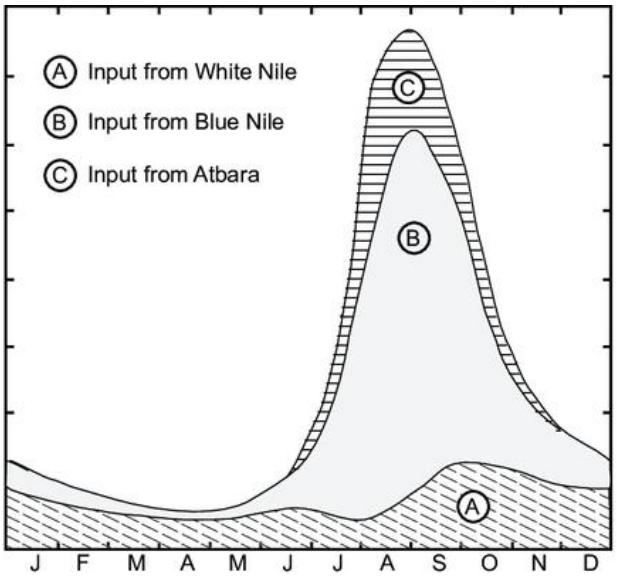

Figure 9: Mean monthly discharge of the Blue Nile, the White Nile and the Atbarah, and total discharge for the main Nile, based on 1912-1936 averages(Williams et al., 2022)

2.1.1.9 The Main Nile

Khartoum to Wadi Halfa

The Main Nile from Khartoum to Wadi Halfa, just south of the reservoir of the Aswan High Dam, is a stretch of 1500 km through increasingly arid environments. Average rainfall for Khartoum is 142 mm a year. For the area around the town of Atbarah, where the Atbarah tributary joins the Nile, it is only

64 mm a year and at Wadi Halfa average annual precipitation is negligible (Sutcliffe & Parks, 1999). In this stretch the Nile no longer gains water and only loses it through high open water evaporation and water used for irrigation. However, losses in practice are limited due to limited irrigation occurring in this region as well as the relatively small size of the river itself. Significant inflow through precipitation only occurs in exceptional storms.

The Aswan High Dam and Lake Nasser Reservoir

The last gauging station before the Egyptian border used to be Wadi Halfa. However, as the Aswan dam reservoir filled it was moved upstream to Dongala. The Nile in Egypt is characterized by the Aswan High Dam. Up to the Lake Nasser the Nile flow is somewhat natural and after the Aswan High Dam the flow is completely controlled due to the regulating of the dam (Sutcliffe & Parks, 1999).

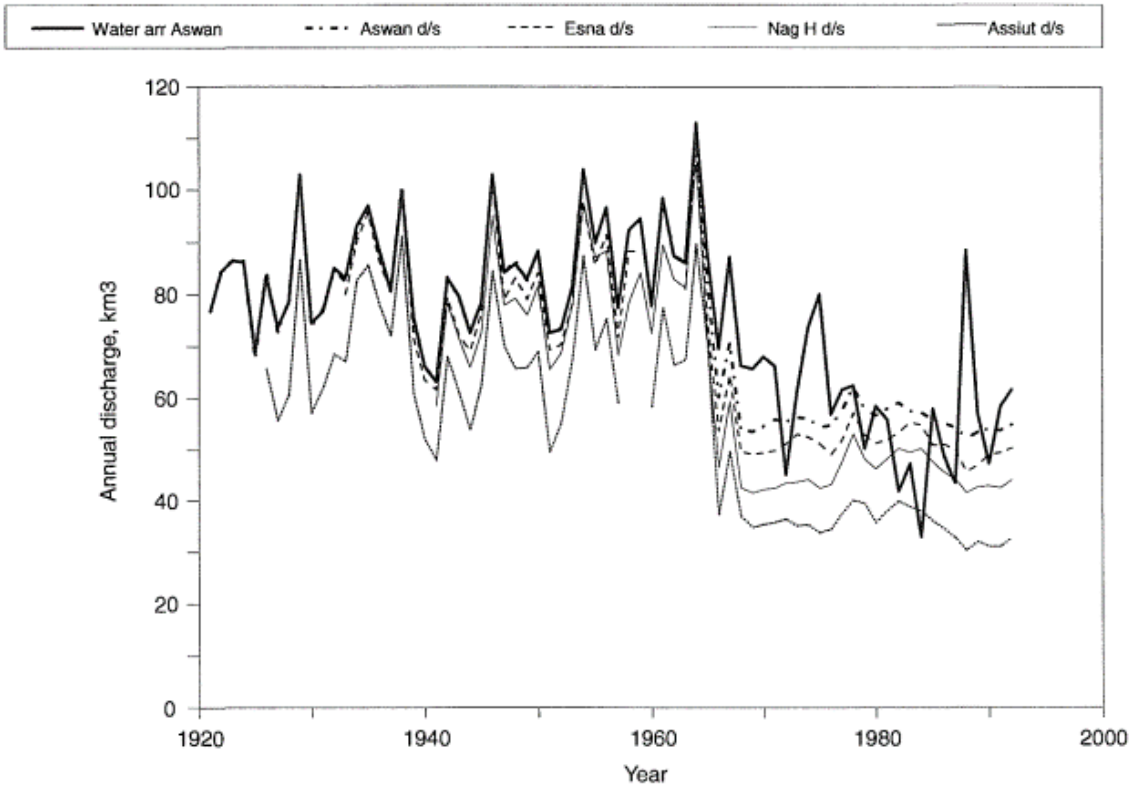

Figure 10: Average annual discharge at the Aswan dam and various stations in Egypt (Sutcliffe & Parks, 1999)

The Aswan high Dam was finished in 1970, but since 1964 the dam already had an influence on river flow. As from this time the reservoir was slowly filling. There were many reasons for building of the dam such as: Control of water flows and regulation for irrigation needs, protection of Nile delta from floods and droughts, hydropower generation, increased agricultural production, improvements of navigability of the Nile River and some more (Abu-Zeid & El-Shibini, 1997).

Annual flow entering Lake Nasser is 84.4 km^3 a year. Outflow before construction of the Aswan HighDam was more seasonal, owing to the seasonality of the Blue Nile. Since 1964 outflow has become more stable year-round (figure 10). Evaporation losses over the reservoir have been determined to be around 10 km^3 a year with bed infiltration losses to be about 1 km^3 a year (Sutcliffe & Parks, 1999).

Past Aswan: Egyptian Nile

The last stretch from the outflow of the Aswan High dam to the Mediterranean Sea is in an arid environment. Inflow through rainfall or tributaries is negligible. Agricultural, industrial and domestic water use reduces the flow in the Nile. Average discharge into the Nile Delta for the years 1959-1964 was 42.9 km^3, so about 40 km^3 was used for these purposes up to the Nile Delta, after which some more is used in the Delta itself. With the building of the Aswan High Dam water use increased to the point that only 18 and 21 km^3 flowed into the Nile Delta in 1982 and 1984 respectively. In the Nile Delta more water is used and outflow into Mediterranean is only 2 - 4 km^3 a year. However, 6 – 7 km^3 a year is estimated to be needed to mitigate salt-water intrusion (Hamza, 2006).

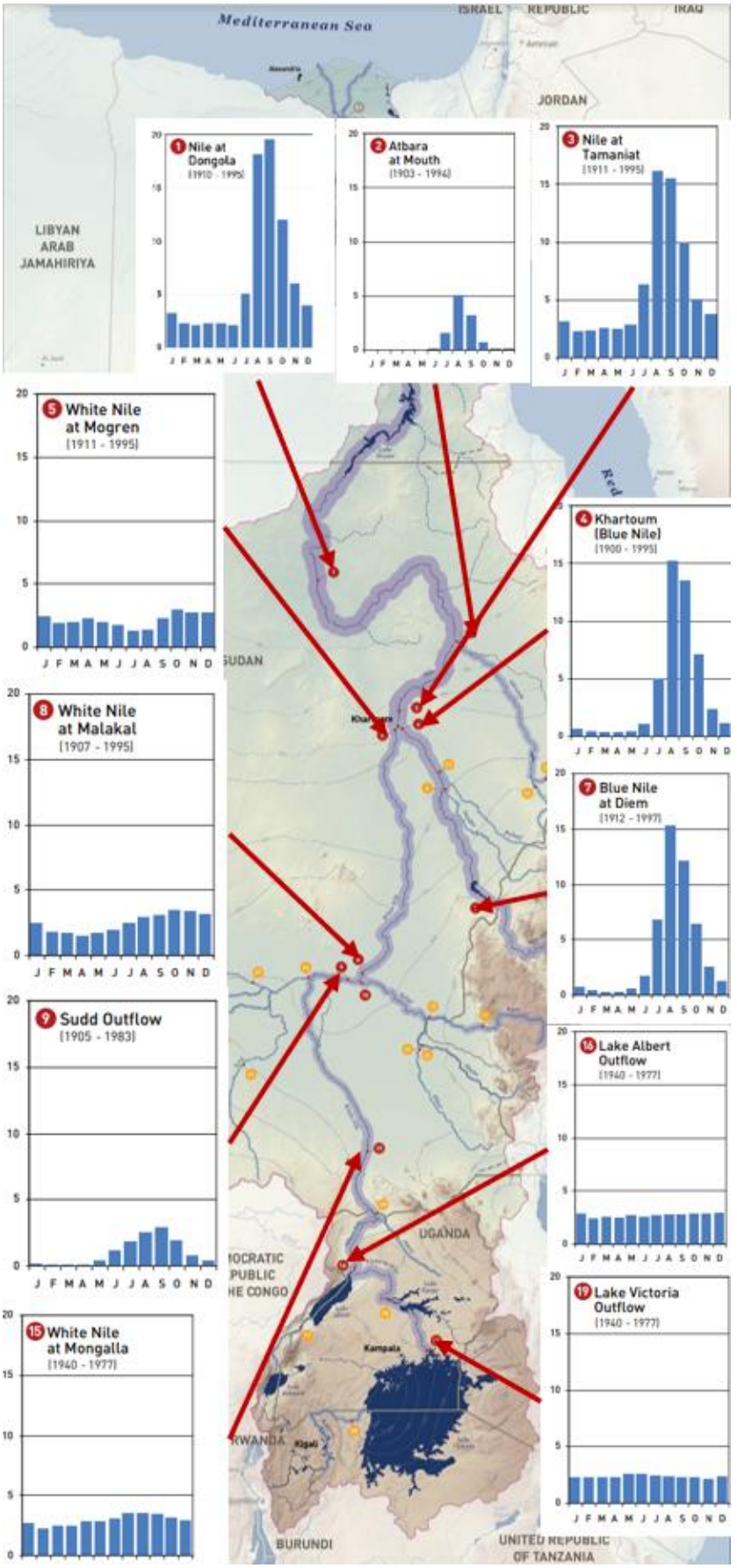

Figure 11: Several gauging stations in the Nile Catchment after (Sutcliffe & Parks, 1999)

2.2 Hydrological Models

The study will be carried out by using a hydrologic model, PCR-GLOBWB 2, to simulate the flow of the Nile. The major advantage of a model is that in situ discharge data is not required to predict the availability of water at a certain point if all variables are known. The development and usage of hydrological models will be discussed in this chapter to identify the history of usage of these models and to compare the results. This in turn to illustrate how PCR-GLOBWB 2 is going to model the Nile flow for this study.

2.2.1 Development of various Hydrological Models

Hydrological models have come a long way since the first leaky bucket model (Manabe, 1969), which was used in global circulation models to link atmospheric processes to surface hydrology. This model was a simple bucket which would yield runoff when soil moisture exceeded field capacity. Furthermore, evaporation from the soil was put equal at atmospheric demand, in general this was a very simple model. Further complexity of this simple model was added through the years as described in detail by Bierkens, (2015) leading to more and more complex land surface models (LSMs). In short, the later versions of these LSMs gained increased complexity when it comes to vegetation and soil processes. Especially the soil processes proved to be very important for the interaction between the land surface and the atmosphere.

More recently, there is a change with most LSMs to become land earth surface models (LESM). These combine the energy and momentum exchange of the older LSMs with full soil hydrology, vegetation phenology and dynamics and carbon cycling. These LESMs can be integrated into fully integrated climate models (Bierkens, 2015).

Still LSMs and LESMs tend to model soil-atmosphere interaction and are not linked to water resources. This was done by the first versions of global hydrological models/macroscale hydrological models, which looked at blue water availability (i.e., surface water and ground water) (Bierkens, 2015). Upon adding water use to these models so called Global Hydrological and Water Resource Models came about. A few examples of these are WGHM (also called Watergap (Döll et al., 2003), Mac-PDM.09 (Gosling & Arnell, 2011) ,and PCR-GLOBWB 2 (Sutanudjaja et al., 2018) and more (e.g., Bierkens, 2015; Sood & Smakhtin, 2015). These models coupled blue water availability to needs and can be used to evaluate future needs.

The lines between the different models, and whether they are strictly MHMs, GHWM, LSMs, LESMs, is not always clear and definite and are only useful for quick sorting of the different models. As most of the models perform roughly the same task, calculating water balances. Some of these models have begun to be much the same, this has been the result of several groups converging to a same process description as a result of limited datasets being available but also due to convenient reincorporation of other model code (Bierkens, 2015).

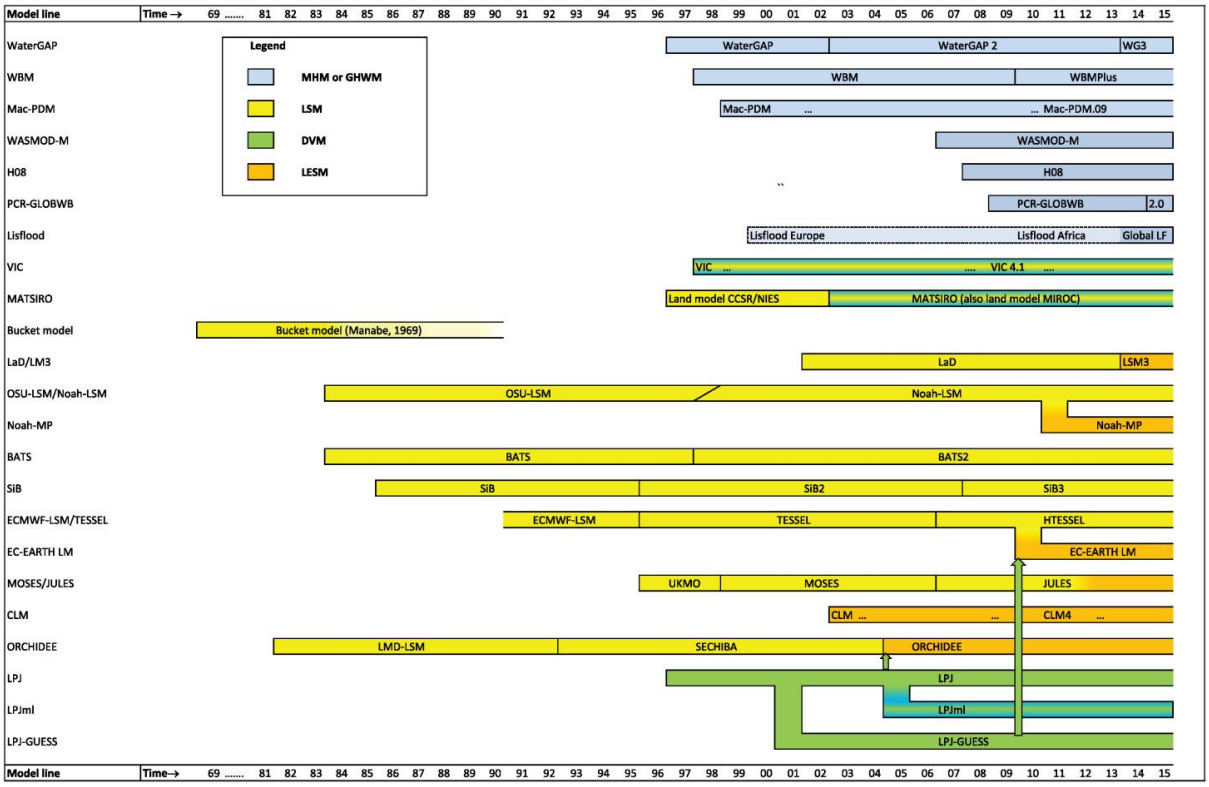

Figure 12: Development history of several hydrological models (Bierkens, 2015)

2.2.2 Current Global Hydrological Models

The newest versions of various GHMs are presented in figure 12. These can look at water availability and connecting this to water resource modelling. This to look at effects of climate change on water availability. But these can also be used for looking at the effects of socio-economic developments and feedbacks on water resource usage.

Several projects exist to compare model results, as all these different model different structures resulting in slightly different outcomes. The WATCH (Water and Global Change) project aimed at bringing better understanding of the current and future water cycle for scientist and in this compared the outcomes of several models all forced with the same WATCH forcing data set (Haddeland et al., 2011; Harding et al., 2011). Several more model comparisons of the different GHMs exist (see Bierkens, 2015; Schellekens et al., 2017). These show a lot of agreement within models, but also a few differences. Most of the current models use a 5 arcminutes resolution on a gridded map with a time scale of a day. Table 2 shows a comparison of different model functions for different models.

Table 2: Comparison of different properties of different hydrological models (Schellekens et al., 2017)

Model	Interception	Evaporation	Snow	Soil layers	Groundwater	Runoff	Reservoirs/lakes	Routing	Water use	Time step
HTESSEL-CaMa	Single reservoir, potential evaporation	Penman–Monteith	Energy balance, 1 layer	4	No	Saturation excess	No	CaMa-Flood	No	1 h
JULES	Single reservoir, potential evaporation	Penman–Monteith	Energy balance, 3 layers	4	No	Saturation and infilt. excess	No	No	No	1 h
LISFLOOD	Single reservoir, potential evaporation	Penman–Monteith	Degree-day, 1 layer	2	Yes	Saturation and infilt. excess	Yes	Double kinematic wave	Yes	1 day
ORCHIDEE	Single reservoir structural resistance to evaporation	Bulk PET (Barella-Ortiz et al., 2013)	1 moisture layer, 1–5 thermodynamic layers	11	Yes	Green-Ampt infiltration	No	linear cascade of reservoirs (sub-grid)	irrigation only	900 s energy balance, 3 h routing
PCR-GLOBWB	Single layer, subject to open water evaporation	Hamon (tier 1) or imposed as forcing	Temperature based melt factor	2	Yes	Saturation excess	Tier 1 only lakes	Travel time approach	Not in tier 1	1 day
SURFEX-TRIP	Single reservoir, potential evaporation	Penman–Monteith	Energy and mass balance, 12 layers	14	Yes	Saturation and infilt. excess	No	TRIP with stream	No	900 s for ISBA, 3600 s for TRIP
SWBM	No	Inferred from net radiation	Degree-day, 1 layer	1	No	Inferred from precipitation and soil moisture	No	No	No	1 day
W3RA	Gash event-based model	Penman–Monteith	Degree-day, 1 layer	3	Yes	Saturation and infiltration excess	No	Cascading linear reservoirs	No	1 day
WaterGAP3	Single reservoir	Priestley–Taylor	Degree-day, 1 layer	1	Yes	Beta function	Yes	Manning–Strickler	Yes	1 day
HBV-SIMREG	No	Penman 1948	Degree-day, 1 layer	1	No	Beta function	No	No	No	1 day

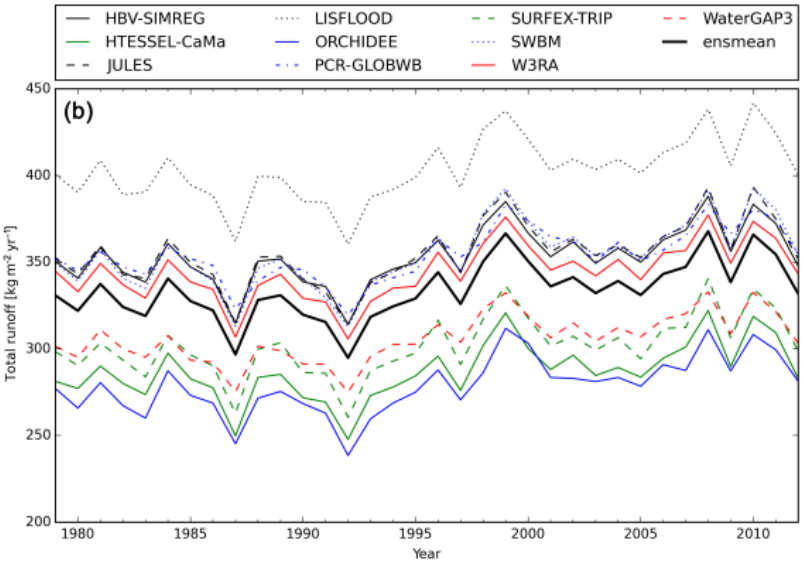

Figure 13: comparison of global runoff for several hydrological models (Schellekens et al., 2017)

Results of model comparisons like those from the WATCH project (Haddeland et al., 2011) or the ensemble used in eartH2Observe project (Schellekens et al., 2017) show similarities between models, but also still major differences. Showing relatively the same behaviour and results, but still different results (Figure 13). Good data is a must for all GHMs and is required for three major processes within GHMs: Model parameterization, meteorological forcing, and validation (Bierkens, 2015).

There is a distinction to be made between calibrated and uncalibrated models. Uncalibrated models change nothing to the data output and in this way show a model that can be used everywhere without a need to be calibrated. An example of uncalibrated models is PCR-GLOBWB 2. Calibrated models, on the other hand, use calibration to make their data fit the results and by doing so make the model more reliable. A downside of this is that calibration has to be done for every basin to

make it fit, and it shows that the model does in general now work. An example of a calibrated model is WaterGap. (Schellekens et al., 2017)

2.3 PCR-GLOBWB 2 an example of a GHM

Here PCR-GLOBWB 2 will be discussed as an example of other GHMs, this to (1) fully uncover the structure of such a model and the data input required and (2) to introduce the model that will be used in this research.

PCR-GLOBWB 2 is a grid based GHM implemented in Python using open source PCRaster Python routines (Karssenberg et al., 2010). The model currently runs at two different resolutions: 5 arc minutes and a coarser 30 arc minutes resolution (Sutanudjaja et al., 2018).

PCR-GLOBWB 2 has a modular structure with the possibility of modules being turned on and off making the model more flexible. The different modules are presented in Figure 3 and will be discussed below.

2.3.1 Meteorological forcing module

The meteorological forcing module of PCR-GLOBWB 2 uses a time series of different meteorological parameters. These include precipitation, temperature, and reference evaporation. This reference evaporation can either be prescribed to the model or calculated and is later used in the land surface module. Calculation of reference evaporation can be done in two ways, either by using Hamon (1963) if only daily mean temperature or using Penman-Monteith following FAO guidelines (Allen, 1998) if net radiation, wind speed, and vapour pressure deficit are additionally available. This is then used to calculate land-cover-specific potential evaporation based on crop factors of the various land cover types according to the FAO guidelines (Allen, 1998) see van Beek (2008) for details on these calculations and forcing. Precipitation is divided into either snow or rain depending on the temperature of the precipitation and temperature drives snowmelt processes (Sutanudjaja et al., 2018).

Meteorological data is required and is based on the CRU TS 3.2 data set (Harris et al., 2014) and is downscaled. The process of downscaling is further described in Sutanudjaja et al. (2018) and van Beek (2008). A list of general required data input of both meteorological parameters and other parameters is included as Appendix A.

2.3.2 Land surface module

The land surface model of PCR-GLOBWB 2 covers the interaction between the vegetation layer, the two soil compartments, the groundwater compartment, and the atmosphere. The standard model run covers four different types of land cover types: tall natural vegetation, short natural vegetation, irrigated crops and paddy-irrigated crops. Vegetation properties vary over the year with monthly climatology and growth of plants, this has an influence on the Leaf area index (LAI) and crop factor, and they have an influence on the water exchange between the plants and the atmosphere (see van Beek, 2008). Note that land cover classes are calculated as a fraction per cell. This is also the case for soil types, so that vegetation properties as well as soil properties vary not only per land cover type but also within cells and between cells. All fluxes to and from the land surface model are calculated separately for every land surface cover type in the cell, and a list of all fluxes in PCR-GLOBWB 2 can be found in Appendix B. Irrigated water needs are based on actual soil water storage (S_1 and S_2) or inundated water storage according to Allen (1998) in the irrigation and water use model (Sutanudjaja et al., 2018).

Soil column fluxes in the soil layers (S_1 and S_2) are driven by degrees of soil saturation and interacts with the groundwater (S_3) layer. Soil parameters are outlined in Appendix and are retrieved from the Soilmap from the FAO (1974). Calculations of vertical groundwater fluxes are shown in van Beek & Bierkens (2009), and this is done for each soil type averaged per cell as per the land cover types.

2.3.3 Groundwater module

The Groundwater module is responsible for the groundwater storage dynamics. It fluxes are outlined in Figure 3 and include groundwater recharge and capillary rise (from the land surface module), groundwater discharge (if storage is positive) or riverbed infiltration (if storage is negative). Groundwater discharge is calculated by a linear-outflow relationship ($Q_{bf}=S_3/J$), with L based on drainage network density and aquifer properties (van Beek & Bierkens, 2009).

It is further possible to link the outcome of PCR-GLOBWB 2 to an integrated groundwater model based on MODFLOW to calculate heads and flow paths. This can be done by one-way coupling, first running the model normally and then linking it to the model (Sutanudjaja et al., 2011). However, fully integrating the two models and letting the MODFLOW groundwater flow model run as the groundwater module is also possible (Sutanudjaja et al., 2014).

2.3.4 Surface water module

The Surface water module is responsible for the routing of surface water throughout the model. First a drainage network is constructed from high resolution topographic maps. Over this drainage network water from the land surface module can be routed. It is also possible to run the surface water module separately and use input from other GHMs for example (Sutanudjaja et al., 2018).

Routing of surface water can be done in three increasingly complex ways in the model. The first method consists of running a simple accumulation of fluxes over the drainage network. This is done in time steps which are longer than the travel time of water of the longest river length. The second method uses an estimation of velocity in the cell based on average discharge of the last five years and Manning's equation, which assumes the energy slopes is the same as the bed slope. Using this velocity an amount of water can be moved through the drainage network daily. This method works best with steep slopes and rivers that are filled year-round. The third method is a more complicated version of the second adding the kinematic wave equation (KWE) of the Saint-Venant equations with the flow described by Manning's equation as before. This third method is more realistic for simulations of flood wave propagation. This is the method that produces flooding of floodplains (Sutanudjaja et al., 2018). When using the kinematic wave equation, it is also possible to get inundation of the floodplain, this is done as in subgrid calculation based on a high-resolution DEM map. Reservoirs, lakes, river channels and inundated floodplains are all subject to open water evaporation and surface water can be subjected to withdrawals for irrigation from the irrigation and water demand module (see Figure 3). The third method also makes an energy routing scheme possible to stimulate surface water temperature (van Beek et al., 2012).

Lakes and reservoir are an important part of the drainage network and are simulated in PCR-GLOBWB 2 as well. Lakes are taken from the Global Lakes and Wetlands database (GLWD) (Lehner & Döll, 2004) and reservoirs are taken from the GRanD database (Lehner et al., 2011). Lakes and reservoirs can extend over several cells are maintained at a stable level throughout all cells. Outflow of lakes is simulated as outflow over a simple broad-crested weir and outflow of reservoirs is based on a release strategy. This strategy is based on aiming average discharge while keeping levels

between a minimum and maximum storage (Wada et al., 2014). However more elaborate strategies are possible too, such as taking downstream water demand into account (Sutanudjaja et al., 2018).

2.3.5 Irrigation and water demand module

The irrigation and water demand module is a fully integrated module in PCR-GLOBWB 2 and is responsible for calculating water demand, withdrawal consumption and return flows within the model structure as seen in Figure 3.

The amount of water needed for irrigation is depended on the crop and on the irrigated area per cell and is calculated according to FAO guidelines (Allen, 1998). The irrigation needs of the crops changes per month. Total irrigated area changes in time according to FAOSTAT data, but the fraction of paddy and non-paddy irrigation stays the same for lack of data. A part of the water irrigated can be lost by transpiration and open-water evaporation or percolation into the soil. Transpiration and evaporation together make up the irrigation water consumption and the percolation acts as a recharge to the groundwater and is a return flow (Sutanudjaja et al., 2018).

Other water demands are divided into three categories: industry, livestock, and household. The water demand is calculated based on population, electricity demand, GDP per capita. Household water demands exhibit a seasonal variable corresponding to temperature. Return flows from industrial and household water is taken into account, livestock fully uses up all the water supplied to it (Wada et al., 2014).

The withdrawal of water is estimated to cover the gross demand calculated, if not enough water resources are available the water withdrawal is scaled down to the available water and is non preferentially scaled back from the three sectors. Water can be withdrawn from three sources: surface water, ground water, and desalinated water. The fractions of surface water and ground water are related to their relative abundance and the desalinated water is a set number see Wada et al. (2011). The availability of surface and ground water is calculated as a 2-year running mean of river discharge and groundwater recharge. If available surface water is first used, up until the discharge is 10 % of long-term average discharge under naturalized flow conditions. This is calculated by running the model first without withdrawal. Groundwater is then abstracted, or if no surface water is available, then it is already abstracted. First renewable groundwater storage is used and if not present, non-renewable groundwater will be used. The amount of groundwater available is limited by pumping capacity. A more detailed scheme of local water distribution is giving in Sutanudjaja et al. (2018).

2.4 Datasets

As mentioned above, GHMs require various good quality input datasets. Part of the input data for PCR-GLOBWB 2 will be further explained as well as a new dataset that might provide more accurate information.

2.4.1 Meteorological Forcing

The current meteorological forcing is the CRU TS 3.2 data set (Harris et al., 2014), which is downscaled to be used in the current version of PCR-GLOBWB 2 by the ERA-Interim reanalysis for the years considered in the current study. The meteorological forcing has three separate inputs into the model: Precipitation, Temperature and Reference Potential Evaporation.

The newest version of the Watch Forcing Data method applied to the ERA-5 reanalysis. This combination resulted in the WFDE5 data set for Precipitation, Temperature and Reference Potential Evaporation. The previous version of this dataset was the WFDE-Interim set, which used the Watch Forcing Method Data method used on the ERA-Interim reanalysis product for meteorological data. This was generated to 0.5 degrees resolution by interpolation of lower resolution data, the ERA-Interim data. The newer WFDE5 uses the higher resolution of ERA5 product to generate the forcing set, which results in higher variability. In general, the errors when compared to observations are smaller for the newer WFDE5 set than the WFDEI set (Cucchi et al., 2020). Since the CRU TS 3.2 is downscaled based on the same method as the WFDEI set, it is expected that using the WFDE5 set will result in this same smaller difference reported for the WFDE5 when used in PCR-GLOBWB 2.

2.4.2 Soil maps

Currently PCR-GLOBWB 2 uses a digitized version of soil map of the world (FAO, 1974). This digital soil map of the world (DSMW) is used to provide some key parameters for the model. These include parameters that are used in infiltration of precipitation and thus are key for correctly modelling rainfall runoff processes as outlined in van Beek & Bierkens (2009).

The digital soil map of the world is a database for soil properties based on samples. However, samples do not spatially cover the entire soil, so many samples are needed. From this a spatial distribution of soil properties can be estimated. This will have some uncertainties in the results which decreases with the amount of soil samples. This means that in 'developed countries' there may be enough point observations to make a reliable soil map. However, in the current 1995 version of the DSMW used in PCR-GLOBWB 2 the number of samples in Sub-Saharan Africa is limited, making it less reliable to use (Sanchez et al., 2009). An improved database was introduced as the Harmonized World Soil Database (HWSD), but this still relied on in situ observations (FAO et al., 2012).

However, new methods now exist to overcome this lack of spatial input data. This is called Digital Soil Mapping (DSM) and consists of building a quantitative numerical model between soil observations and environmental data to model soil formation behaviour. This method of soil mapping has increased in quality over the last few years with more advanced modelling methods including deep learning and artificial intelligence (Poggio et al., 2021).

SoilGrids 2.0 is a 250m scale digital soil map using the newest DSM processes to create a high-quality soil data product that can be used in PCR-GLOBWB 2 to receive the newest soil data for the Nile Basin. A description of the method used to obtain their soil map can be found in Poggio et al., (2021)

2.4.3 Topographic maps

Topographic data is used in the description of several sub grid processes that influence the schemes for runoff–infiltration partitioning, interflow, groundwater recharge, and capillary rise (Sutanudjaja et al., 2018). An improved dataset might decrease the poor rainfall runoff characteristics that were reported to be the cause of the overestimated discharges (van Beek et al., 2011).

Therefore, the implementation of a new DEM to increase the fidelity of the modelled surface and increase model results. Merit-DEM is shown to improve the quality of the topography for flat areas (Modi et al., 2020) and these are also areas where PCR-GLOBWB 2 performs poorly in the Nile catchment. This might improve model behaviour.

3 Methodology

To answer the research questions several model runs will be performed using the PCR-GLOBWB 2 model framework. In total five basic model runs will be performed as outlined in Table 3. The Standard run will be the benchmark run and will be evaluated using the Kling-Gupta Efficiency (KGE) when looking at discharge measurements. It will also be evaluated using GLEAM evaporation data to compare evaporation. The other runs, which will also be validated using the same metrics as the Standard run, are used to evaluate the effect of introducing a new parameter on model performance.

3.1 Validation of model results

Validation of model results is important for using model results in future studies. However, for the Nile Basin this can be difficult as there is limited in situ data available to compare model output with. Two methods were employed to validate the model in this study: one using gauged river discharge data and the other using earth observation data to measure evaporation.

3.1.1 Using discharge data

Comparing observed discharge measurements from gauging sites to modelled discharge is one way of validating model results. Through this the performance of the model can be assessed for locations on the river network. The GRDC (Global River Discharge Centre) has a global database of historical river discharges that can be freely accessed. To assess the accuracy of the modelled discharge the Kling Gupta Efficiency (KGE) was used (Kling et al., 2012). KGE is calculated according to Equation 1.

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2},$$

Equation 1: Calculation of the KGE score for validation of model data (after Knoben et al., 2019)

Where r is the linear correlation between observations and simulations, where σ_{obs} is the standard deviation in observations, σ_{sim} the standard deviation in simulations, μ_{sim} the simulation mean, and μ_{obs} the observation mean. KGE = 1 indicates perfect agreement between simulations and observations. A value of KGE ≥ -0.41 indicates that the model improves upon the mean flow benchmark (Knoben et al., 2019). The closer a value is to KGE = 1, the better the result, the further away, the poorer the performance would be.

For the period of 1979-1989, 37 gauging sites can be found at the GRDC that have (partial) discharge measurements. Most of these measurements only go up to the years 1982 or 1984 although gauging sites in Burundi are often available up to 1990. The locations of the gauging sites are presented in Figure 14 and additional information is provided in Appendix C. Monthly measurements were used to compare the model results of PCR-GLOBWB 2 to GRDC gauging sites.

3.1.1.1 Gauging sites and subregions

The gauging stations are spatially represented in Figure 14. Here they are divided into separate geographic regions based on subregions of the Nile Basin and letter coded accordingly in Table 3. The analysis will be based on these regions for a clearer picture. Additional information about the separate gauging stations is provided in Appendix C.

Table 3: Details of GRDC gauging stations as used in the study. Sorted by the subbasin in which the stations are located. Abbreviations are highlighted in the full name.

Abbreviation	Full Name	No. of gauging stations	Start	End
E	**E**gyptian Nile	6	Lake Nasser	Mediterranean Sea
M	**M**ain Nile	3	Khartoum	Lake Nasser
A	**A**tbarah	2	Ethiopian Highlands	Main Nile (after station **M2**)
B	**B**lue Nile	3	Ethiopian Highlands	Khartoum
W	**W**hite Nile	8	Lake Albert (outflow)	Khartoum
G	(African) **G**reat Lakes	15	Rwanda and Burundi	Lake Victoria (outflow)

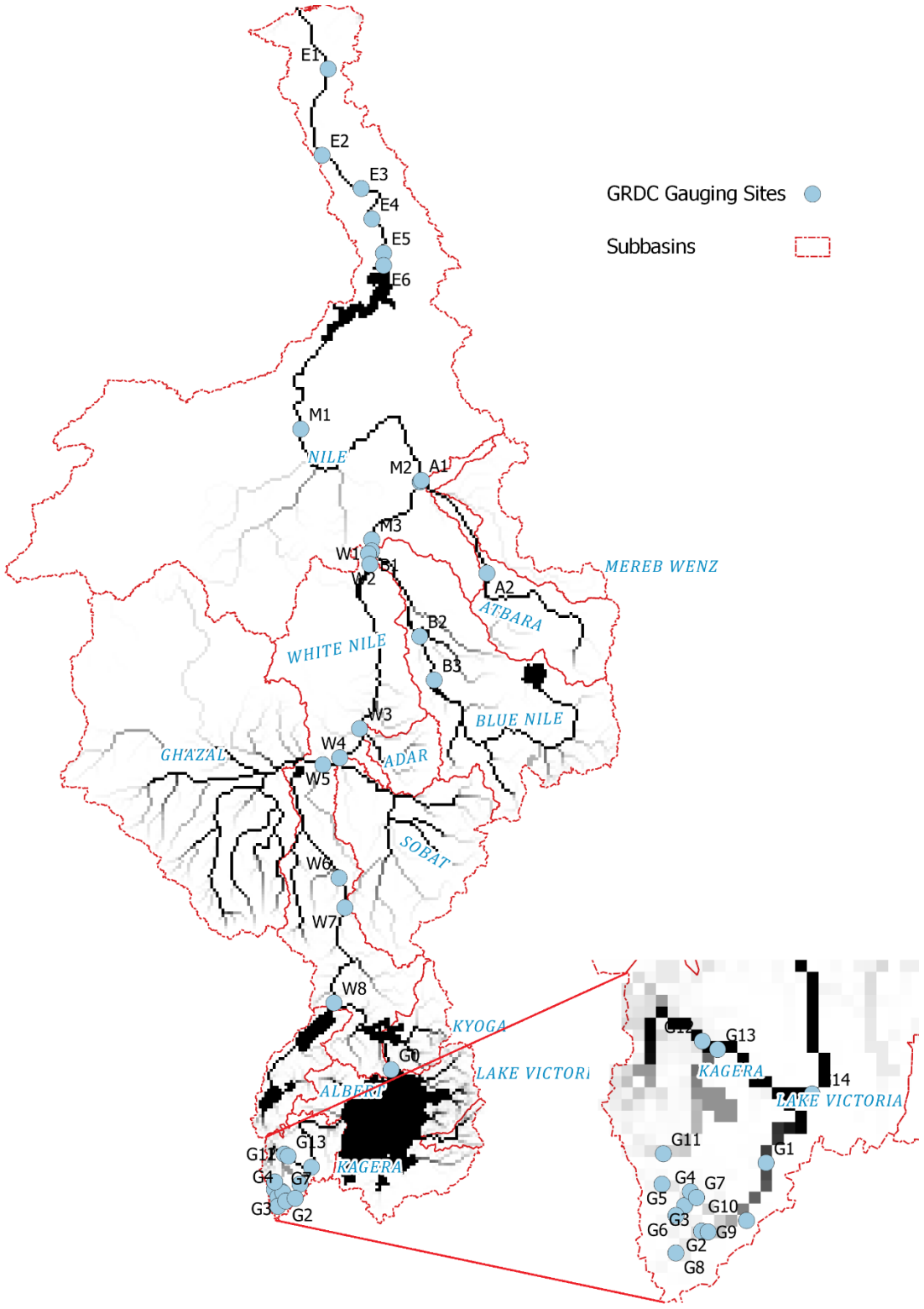

Figure 14: Location of all gauging stations in the Nile Basin used for the validation of discharge data and names of subbasins present in the Nile River as identified by the GRDC. Special zoom on the African Great Lakes region and the spread of stations there. Adapted from GRDC and (Sutcliffe & Parks, 1999)

3.1.2 Validation using Earth Observation data

As validation over the Nile basin using GRDC discharge data is limited on both temporal and spatial scales other methods can be used to validate model results. Several earth observation products can be used. For this study GLEAM was chosen to compare model evaporation to GLEAM evaporation.

3.1.2.1 Description of the GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) is a set of algorithms that can estimate different components of evaporation based on satellite products and observations. The most recent version relevant for this study will be the 3.5a version, which includes a global dataset spanning from 1st of January 1980 until 31st of December 2020. Crucially this means that GLEAM results can be compared to discharge measurements in the period 1980-1984 as well as in the ungauged period 1984-1989. The newest version provides ten different products: Actual evaporation, Soil evaporation, Interception loss, potential evaporation, Snow sublimation, Transpiration, Open-water evaporation, Evaporative Stress, Root-zone soil moisture, and Surface soil Moisture (Martens et al., 2017).

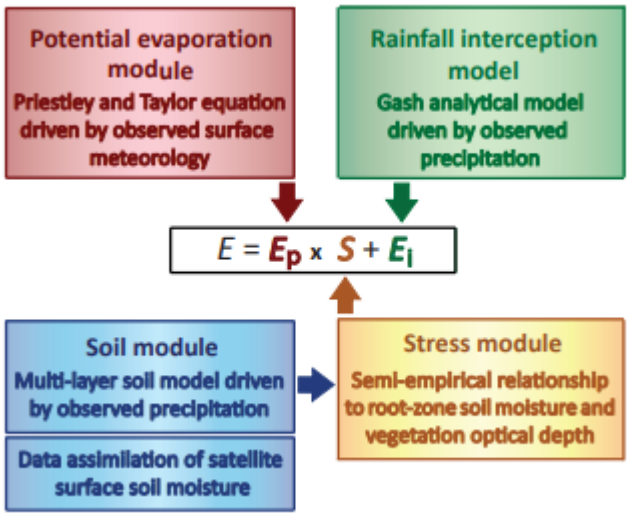

Figure 15: Schematic of model structure of GLEAM (Martens et al., 2017)

What makes GLEAM a possible method for validating model results is its uses of observations for most of its results. The version 3.5a uses a mix of satellite data and reanalysis data. In short, the evaporation is calculated as follows (see Figure 15). First, the Priestley and Taylor equation is uses to calculate Potential evaporation (E_p), this is calculated using meteorologic data for four different landcover types: bare soil, short canopy, and tall canopy. In the 3.5a version this is gathered from the ERA-5 reanalysis set as satellite data for the earliest years is not available. Actual evaporation is calculated using a stress factor (S) and an interception loss (E_i). The interception losses (E_i) are then calculated based on observed precipitation and added. The stress module (S) is calculated based on root-zone soil moisture and observations of microwave Vegetation Optical Depth (VOD), uses for vegetation phenology constraints on evaporation. The VOD is a measure of the plants impact on water usage, lower VOD resulting in lower (S) and thus higher evaporative stress. The root-zone soil moisture availability also impacting the amount of water available for plants and thus being increasing evaporative stress if less is available. Root-zone moisture is calculated in a multi-layered soil model that is driven by observed precipitation amounts. Finally, open water evaporation estimates are based on the Priestly and Taylor equation directly, as no plants or soil properties are involved here (Martens et al., 2017).

3.1.2.2 Validation using GLEAM

Whereas with river discharge gauges comparisons can be made in relatively few points, EO products, such as GLEAM, make it possible to compare results in the entire basin. For the current study this means that the simulated evaporation can be compared everywhere to observed GLEAM evaporation. The output of the comparison can be made in several ways, but in this study, it was chosen to evaluate GLEAM data on a 0.5-degree grid, which is about 5x as large as the 5 arcmin gird used in the model simulation. This was done as the 0.5-degree grid makes spatial differences

noticeable, without cluttering the screen. To exclude any seasonal variability the evaluation was carried out for yearly values.

The difference between the GLEAM dataset and the evaporation produced as output of the model runs of PCR-GLOBWB 2 is measured in two other metrics: the RMSE and the RRMSE. The RMSE measures the average difference between the two values and provides a value of the average difference in mm/year. In Equation 2 the calculation of the RMSE is shown, with H^m indicating the measured value and H^c the calculated value, so the difference between the GLEAM and PCR-GLOBWB 2 output respectively. This in itself gives a good signal of how far off the model is compared to GLEAM, but since the Nile basin encompasses different climatic zones, evaporation can vary considerably. This is why the RRMSE is also used, this gives the relative discrepancy between the observed and modelled results. As shown in Equation 3 the RRMSE is calculated by dividing the RMSE value by the average of the measured value. The RRMSE is used to compare the RMSE to the GLEAM evaporation and is the percentage of the RMSE relative to the GLEAM evaporation.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{H}_d^{i,m} - \overline{H}_d^{i,c}\right)^2}$$

Equation 2: Calculation of RMSE values (Despotovic et al., 2016)

$$\mathrm{RRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{H}_d^{i,m} - \overline{H}_d^{i,c}\right)^2}}{\sum_{i=1}^{n}\overline{H}_d^{i,m}} \times 100$$

Equation 3: Calculation of RRMSE values (Despotovic et al., 2016)

3.2 The model runs

The study consists of comparing five model runs. One benchmark run, three runs which introduce new datasets and one run with the Kinematic Wave equation (KWE) used as the method for routing. For the results the Standard run will be compared to the new dataset runs (Meteo, Soil and Topo run) and separately compared to the KWE run to see the effect of the different routing method.

Table 4: Overview of all model runs performed in the study.

Run number (For reference)	Description of Run/ Reason for run	Reference Name	Changed input	New Dataset used (if applicable)
1	Standard run Benchmark	Standard run		
2	WFDE5 run, effect of different meteorological forcing	Meteo run	-Total precipitation -Air temperature -reference PotentialET	WFDE5 (Cucchi et al., 2020)
3	Soilgrids input data run, effect of updated soil data	Soil run	Upper and lower soil parameters (See Appendix A)	Soilgrids 2.0 (Poggio et al., 2021)
4	New DEM map run , effect of updated DEM data	Topo run	Topographical Parameters (See Appendix A)	Merit-DEM (Yamazaki et al., 2017)
5	Updated river network parameters, for flooding simulation.	KWE run		KWE routing rather than accumulation of fluxes

3.2.1 Model protocol

All model runs have roughly the same basic set-up. The runs are from the 1st of January 1979 to the 31st of December 1989 and the resolution is 5 arcminutes, which translates to a grid size of 100km squares at the equator. For other inputs, Appendix A and B can be assessed. Changes between model runs can be found in Table 4.

Initial storage properties for the six land cover classes used is not available for the Nile region. For this reason, a spin up had to be used to generate these initial properties. Rather than using a spin up for each model run. Two spin ups were performed of ten years, using the output of the first as the initial values for the second. The second output was then used as an initial value for all model runs so that all the model runs had the same initial water storage conditions.

4 Results

This chapter shows the different validations of model results. First the evaluation of the GRDC stations will show the scores per region. Second, the results of the comparison of GLEAM to model evaporation results will show be shown for the entire Nile basin.

4.1 Result of validation with discharge data

Monthly modelled discharge data was compared to observed data gathered from the GRDC for all 37 stations used and evaluated using KGE scores. Results are shown here for the subregions, the results per station are in Appendix C. First the effect of introducing new datasets on KGE scores will be shown, followed by the effect of using the KWE routing.

4.1.1 Standard and new input data runs

The Standard run was compared to the runs which introduced new initial datasets. The results are presented below.

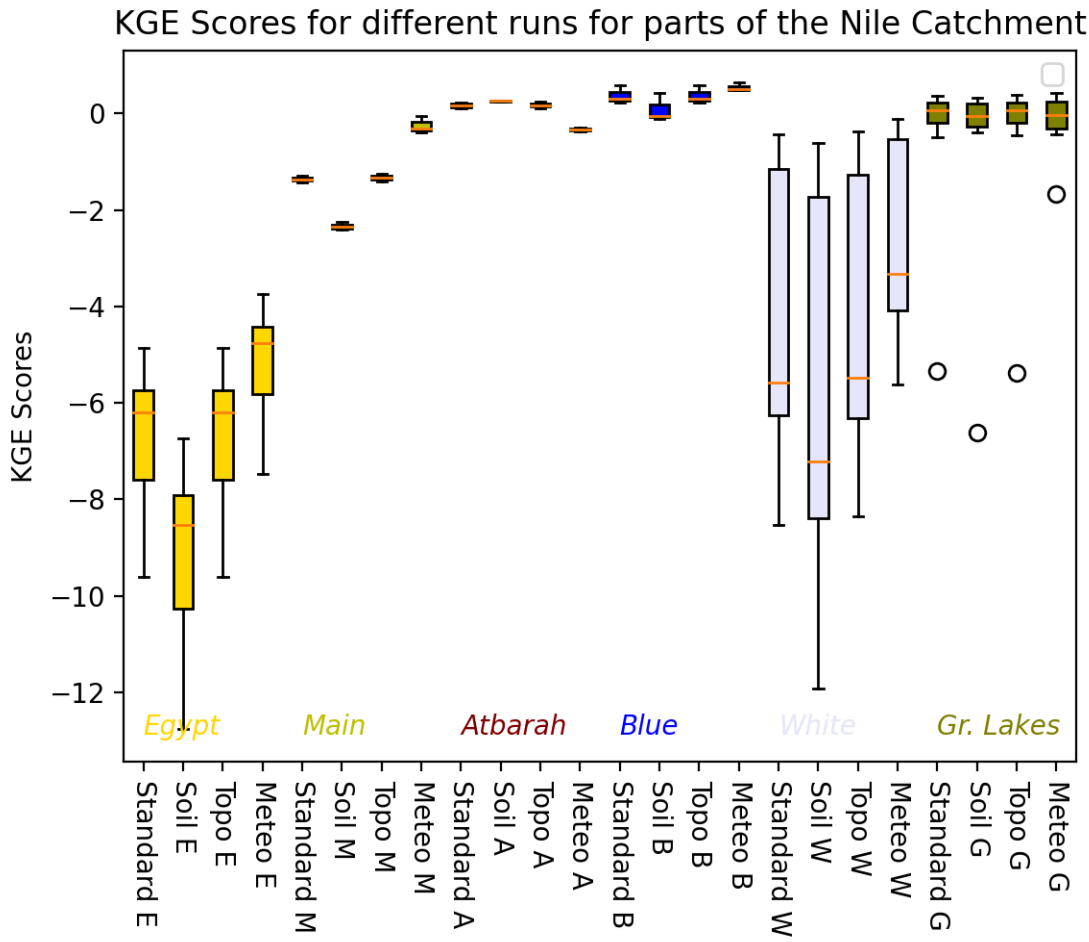

Figure 16: Box and whiskers plot for the KGE scores of every subregion for the standard run and for the three runs with different forcing. From left to right the subbasins are further away from the Nile Delta. The median value is represented by the orange horizontal line in the middle of the box.

When just regarding the Standard run, the performance for many of the subbasins is poor. With many basins having negative reported KGE scores. Although there are some definite differences between the various basins, both in the spread of scores and the scores itself. Furthermore, the Great lakes region has some outliers, displayed as circles in the figure.

Values for the Egyptian part of the Nile show the lowest KGE scores. The White Nile and the Main Nile also show low KGE scores, with the spread being large for the White Nile, but the median is generally on the lower end.

The values for the Main Nile, the Atbarah, Blue Nile and the Great lakes region are shown in Figure 17. This provides a zoom of the top part of Figure 16. This gives a better view of the figures with higher KGE scores, and this highlights the difference for the different runs for these regions. Here the relative high scores for the Atbarah and the Blue Nile are still below 0.5.

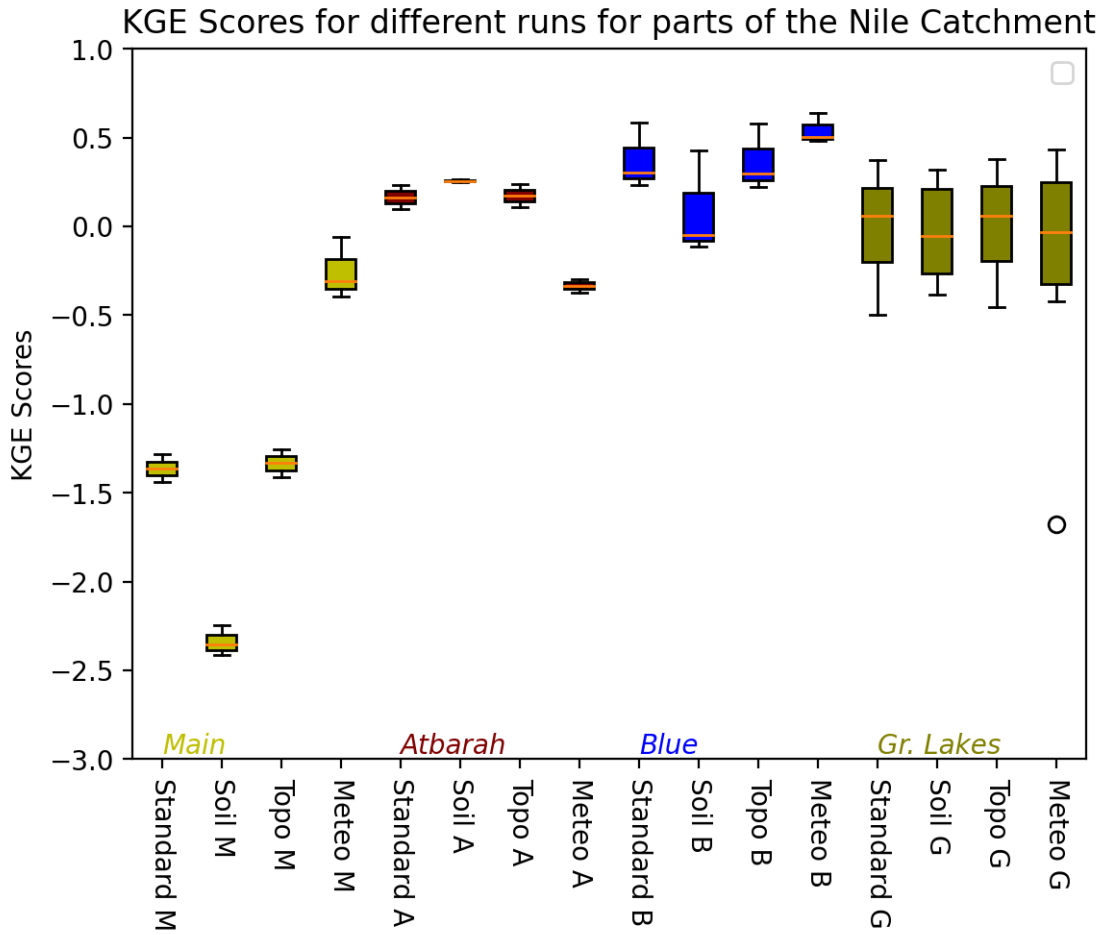

Figure 17: Box and whiskers plot for the KGE scores for the subbasins with higher scores for the standard run and for the three runs with different forcing. From left to right the subbasins are further away from the Nile Delta. The median value is represented by the orange horizontal line in the middle of the box.

Model performance as expressed in KGE scores for the different subregions is variable with the different datasets used. However, blanket statements regarding dataset use and performance cannot be made here as the new input for each subregion has a different effect on the KGE scores. In general, KGE scores for the Soil run are lower for most subregions and KGE scores for the Meteo run are generally higher than the benchmark run. The KGE scores of the Topo run are close to the Standard run for all regions. For the Atbarah region, the results are different than mentioned above. The Meteo run performs worse than the Standard run and the Soil run shows increased values for the KGE scores of the two gauging sites situated there. The Topo run shows the same pattern as the Standard run for most regions, with KGE scores being similar to the Standard run. For The Great

lakes region, differences between the Standard run and the three other runs are less pronounced that for the other subregions.

4.1.2 Results of Kinematic Wave Equation as routing

The comparison between the KGE scores of the Standard run and the run using the KWE as routing are presented in Figure 18. It was thought that using the KWE would increase flooding along the river and make the river more natural. However, KGE scores show the opposite.

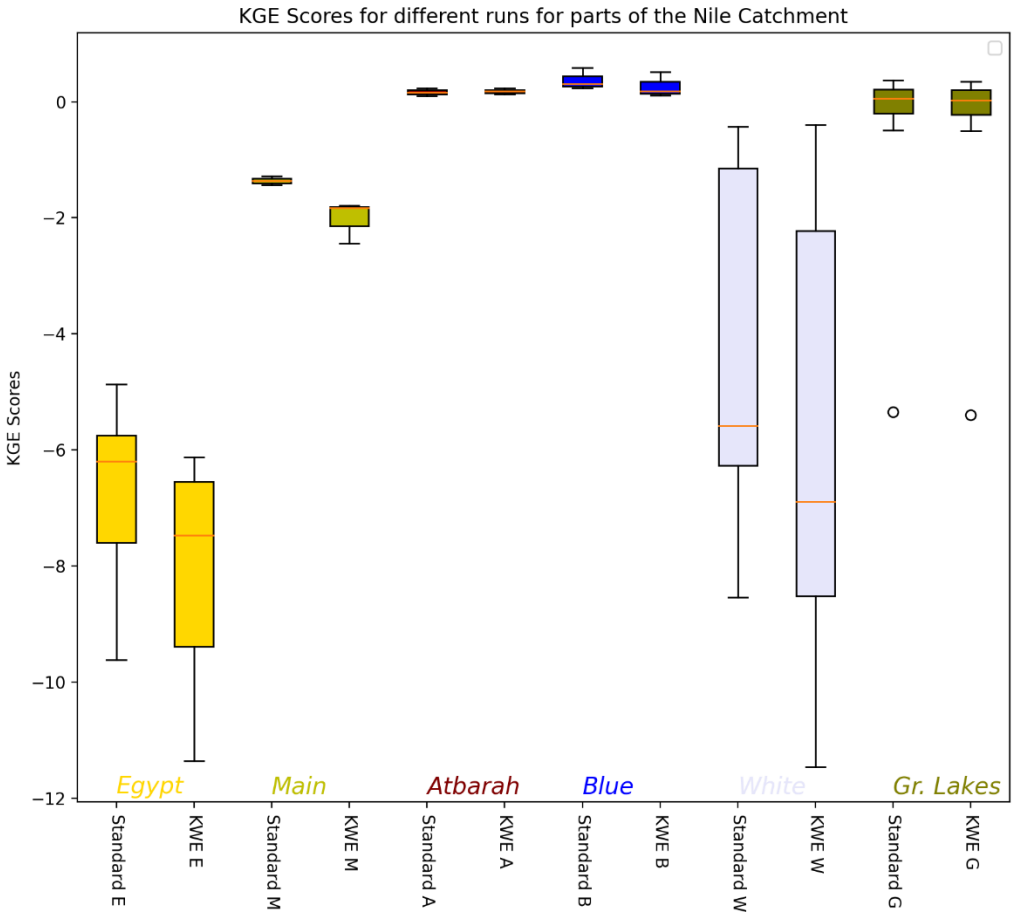

Figure 18: Box and whiskers plot for the KGE scores for all the subbasins for the standard run and for the run using the KWE as the routing. From left to right the subbasins are further away from the Nile Delta. The median value is represented by the orange horizontal line in the middle of the box.

In general, the same pattern of KGE scores is observed as before, with higher scores for the Blue Nile, White Nile and the Atbarah, and lower scores for the other basins. For all the subregions defined, the KGE scores of the Kinematic Wave Equation run were lower than for the Standard run. Of note are the results of for the White Nile, as this is a region which includes large areas of flatlands that are regularly flooded, like the Sudd swamps and the scores here are much worse.

4.2 Results of validation with GLEAM evaporation

The results of the GLEAM validation are presented here. The total evaporation output of PCR-GLOBWB 2 is shown for each run as well as the evaporation as predicted by the GLEAM model. Furthermore, the RMSE in mm/year and RRMSE values are shown for each model run. The results of the four runs will be compared to the Standard run.

4.2.1 The Standard Run

The four figures here give the results for the comparison of GLEAM evaporation to the evaporation output of PCR-GLOBWB 2 in term of RMSE and RRMSE for the standard run. In general evaporation patterns in **a)** and **b)** of Figure 19 are roughly the same. With high evaporations around Lake Victoria, decreasing northwards to almost no evaporation occurring in the deserts next to the Nile in Egypt. The differences between the output of PCR-GLOBWB 2 and GLEAM are expressed in the RMSE and RRMSE values in **c)** and **d)**. The RMSE values for the Standard Run show high values near Lake Victoria indicating a large discrepancy between model output and GLEAM values. The southwestern border of South Sudan also has high values of RMSE as well as in regions of Ethiopia, especially on the border with South Sudan. The general course of the Nile in Sudan and Egypt also shows high RMSE values, especially around Lake Nasser. Then looking at the RRMSE values, they in general seem to be quite high. As these are up to more than 50% in some regions, indicating that the error as described by the RMSE, is larger than half the reported evaporation from GLEAM. However, the regions that have high RMSE values do not generally have high RRMSE values. This is true for the region around Lake Victoria and southwestern South Sudan, although here these can still range up to 20%. Ethiopia does have large regions of high RRMSE values on the same regions as were reported to have high RMSE values. This indicates that here the model output is very different from GLEAM. Lastly, the region above 15°N show values of higher than 50% for the RRMSE.

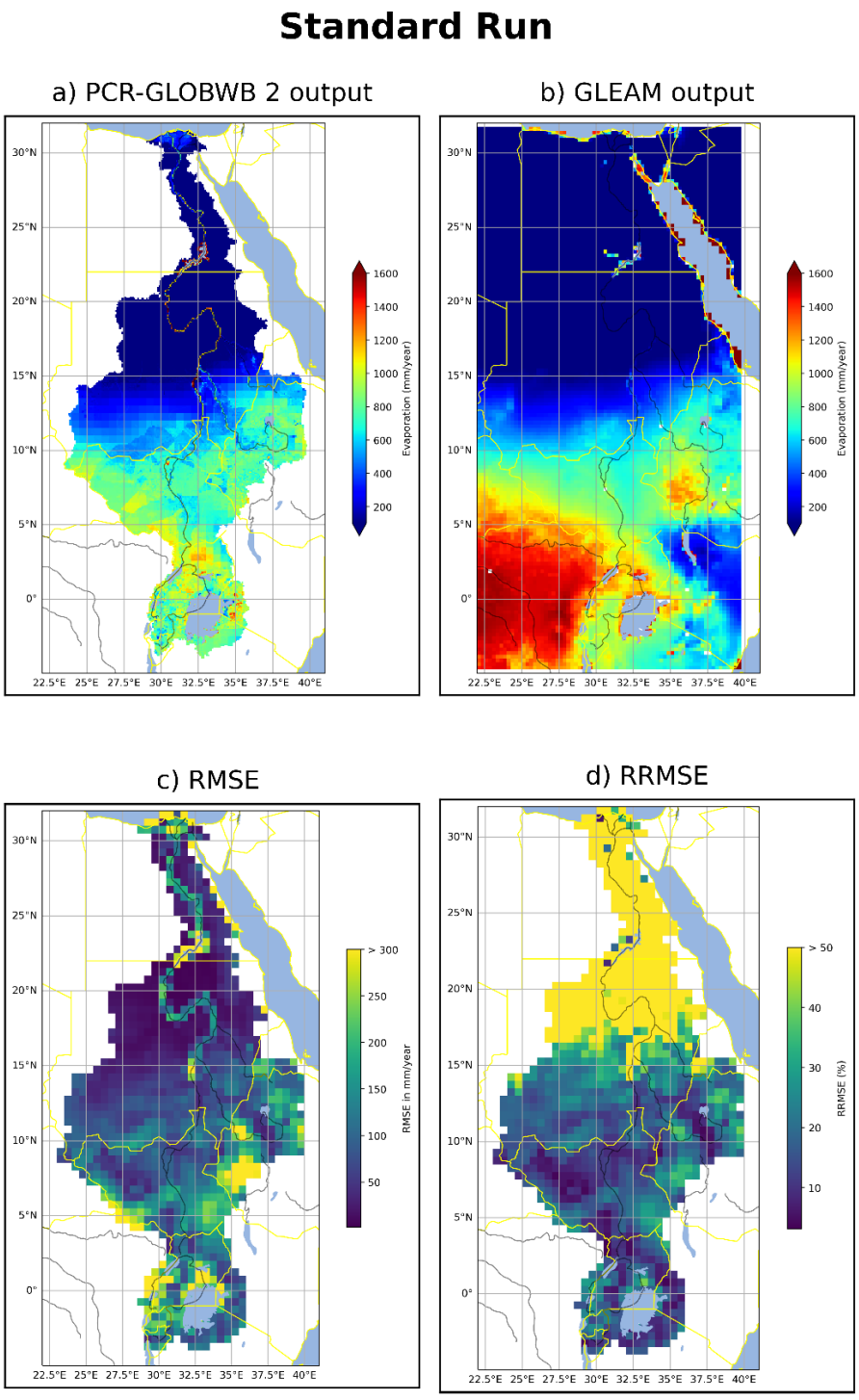

*Figure 19: Maps of the evaporation and statistical values for the Standard run, **(a)** average yearly total evaporation for the Nile Basin as modelled in PCR-GLOBWB 2 for the years 1979-1989, **(b)** average yearly evaporation of GLEAM for the years 1979-1989, **(c)** RMSE values in mm/year for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989 **(d)** RRMSE values in % for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989.*

4.2.2 Topo Run

Compared to the Standard run, the Topographic run shows the same patterns for the RMSE and RRMSE values.

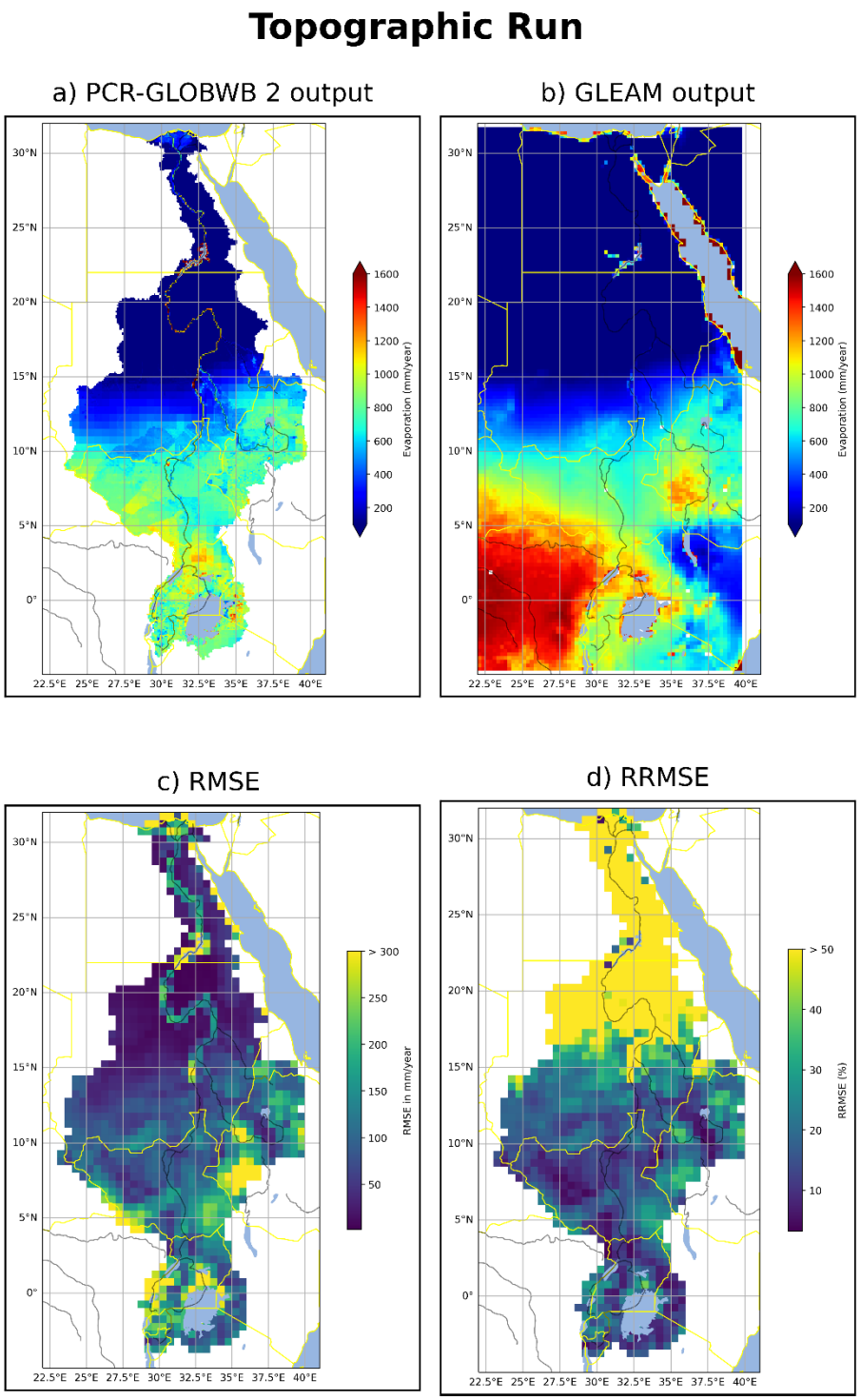

*Figure 20: : Maps of the evaporation and statistical values for the Topo run, **(a)** average yearly total evaporation for the Nile Basin as modelled in PCR-GLOBWB 2 for the years 1979-1989, **(b)** average yearly evaporation of GLEAM for the years 1979-1989, **(c)** RMSE values in mm/year for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989 **(d)** RRMSE values in % for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989.*

4.2.3 Soil Run

Comparing the Standard Run to the Soil Run reveals some differences in the evaporation of the two runs. The evaporation itself is different already in some regions as shown in **a)**. Especially when looking at evaporation in Ethiopia, which is generally lower for the Soil run. Near the lower reaches of the Blue Nile in Ethiopia this has the result that both the RMSE and RRMSE values are higher than for the Standard run, so this shows a larger discrepancy. For the northern part of Ethiopia, it does result in less difference between the model output and GLEAM, resulting in lower RMSE and RRMSE values. Another region where there are large differences between the runs is on the western side of Lake Victoria, all the way to the border with Congo. This has higher values for both RMSE and RRMSE for the Soil run.

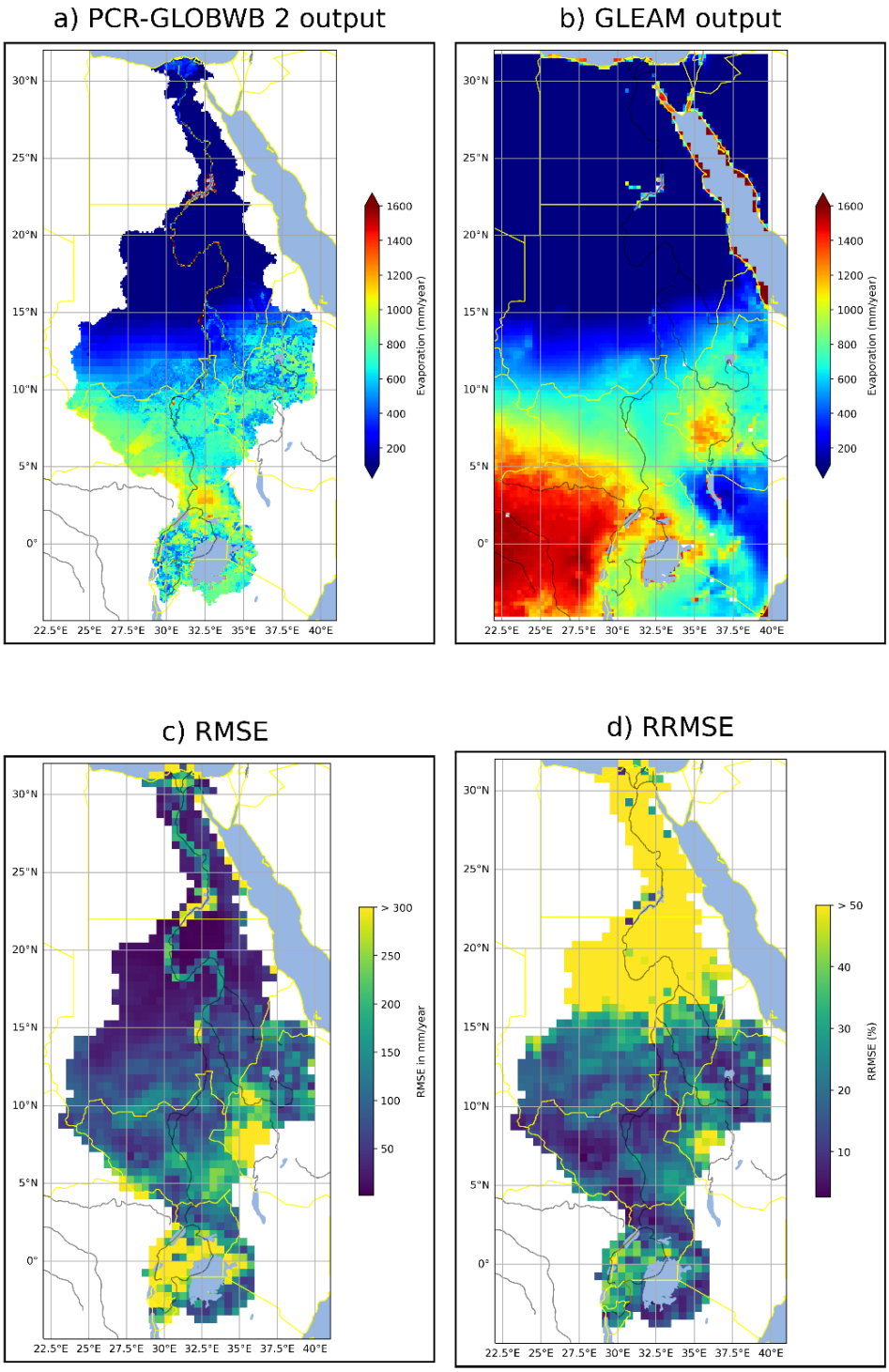

*Figure 21: Maps of the evaporation and statistical values for the Soil run, **(a)** average yearly total evaporation for the Nile Basin as modelled in PCR-GLOBWB 2 for the years 1979-1989, **(b)** average yearly evaporation of GLEAM for the years 1979-1989, **(c)** RMSE values in mm/year for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989 **(d)** RRMSE values in % for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989.*

4.2.4 Meteorological Run

The Meteorological Run shows large differences for the evaporation when compared to the Standard run. In general, evaporation is higher in most regions, and this expresses itself in the comparison to GLEAM. In some regions this results in model output of the Meteorological Run being closer to GLEAM and in others it results in larger discrepancies. In the region around Lake Victoria, the Meteorological Run has lower RMSE scores for the western part. The southwestern border of South Sudan also shows lower discrepancies as outlined by the lower RMSE and RRMSE scores there. This is also observed near the border of South Sudan and Ethiopia. However, there are some regions where the differences are larger too, such as northern Ethiopia as well some spots in the middle of South Sudan, the sudd swamps. Between 10°N and 15°N the RMSE and RRMSE values are generally higher with many spots having RRMSE values getting closer to 50%, indicating very high levels of difference between model output and GLEAM. Indicated by some regions in the west of Sudan as well as large parts of Ethiopia.

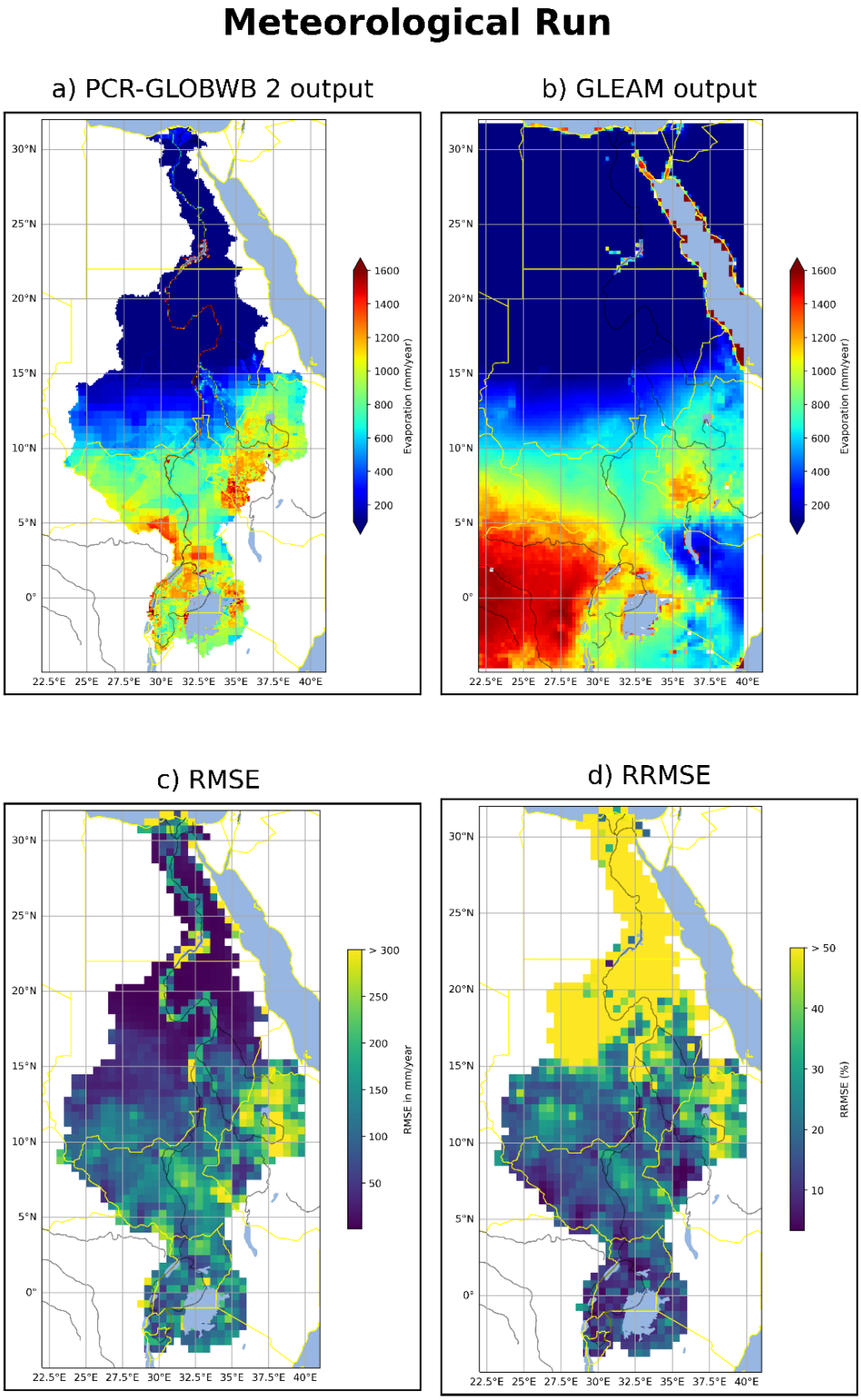

*Figure 22: Maps of the evaporation and statistical values for the Meteo run, **(a)** average yearly total evaporation for the Nile Basin as modelled in PCR-GLOBWB 2 for the years 1979-1989, **(b)** average yearly evaporation of GLEAM for the years 1979-1989, **(c)** RMSE values in mm/year for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989 **(d)** RRMSE values in % for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989.*

4.2.5 Kinematic Wave Equation Run

The run where the Kinematic Wave Equation was used shows the same patterns of evaporation as the Standard Run. The main difference is in the course of the Nile River, where RMSE scores are not as high as in the standard run. Only for around Lake Victoria and for Lake Nasser the RMSE scores are high.

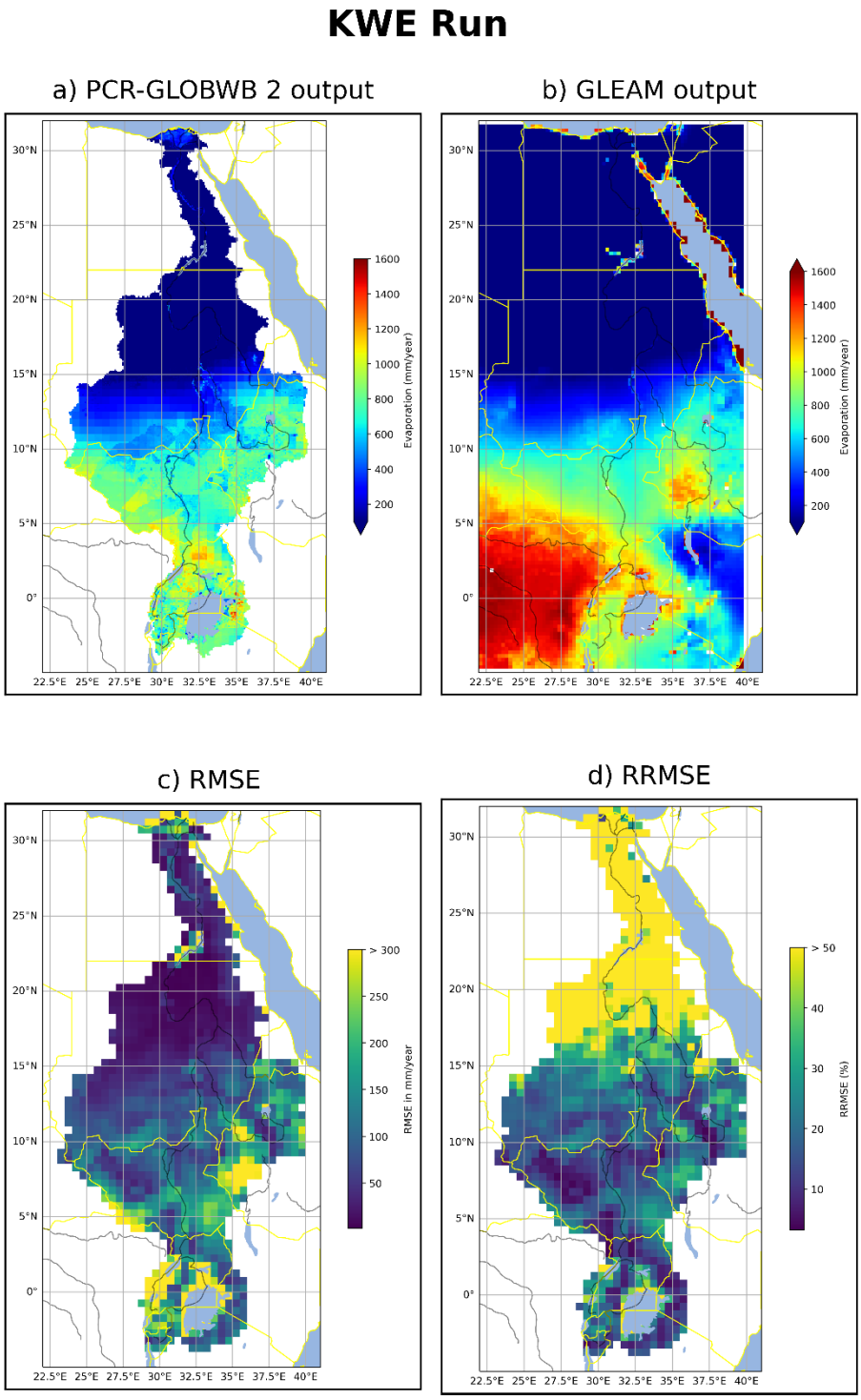

*Figure 23: :Maps of the evaporation and statistical values for the KWE run, **(a)** average yearly total evaporation for the Nile Basin as modelled in PCR-GLOBWB 2 for the years 1979-1989, **(b)** average yearly evaporation of GLEAM for the years 1979-1989, **(c)** RMSE values in mm/year for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989 **(d)** RRMSE values in % for the Nile Basin for 30 arcminute gridcells evaluated for the years 1979-1989.*

5 Discussion

5.1 Evaluation of results

Results of both validation methods show general poor results for the performance of the Nile in PCR-GLOBWB 2, with low KGE scores and high discrepancies between modelled and observed evaporation. The reason for this is different for the various subbasins, but in general is caused by large overestimations of discharge in large parts of the basin. The modelled Hydrology of the Nile will be discussed in terms of the performance of the model from tributaries of Lake Victoria in the south to the Egyptian part of the Nile in the North. This to see if the poor performance of the model can be explained by looking at these two ways of validation and the behaviour of the particular subbasins.

5.1.1 Evaluation of the gauging stations

First an analysis of the methods used to validate river discharge using the KGE also poses some problems which are inherent to the Nile Basin. First of all, the known problem that most of the gauging sites only have data available until 1984, some only until 1982, barring the stations in Burundi. That while most meteorological data products start in 1979, both WFDE5 and the CRU dataset. This provides only a window of a few years in which the analysis can be performed, which is limited, especially for the stations that only have three years of data.

5.1.2 The (small) Great Lakes

The Great Lakes region has the most gauging stations out of all regions, and results for the KGE were not bad, with the median KGE scores of around 0. However, the importance of the main gauged river, the Kagera, is limited. That while the Kagera has 14 out of the 15 gauging sites for this region. The outflow of the Kagera is about 6 km^3 annually is far less than the about 30 km^3 outflow of Lake Victoria. So, the impact of the KGE scores here on the rest of the Nile is limited. The GLEAM comparison here yields some regions which have about 10 – 30 % error in their evaporation when compared to GLEAM. This might explain some of the underestimation of the output here. Furthermore, the main source of water from Lake Victoria is the precipitation falling on the lake itself (Sutcliffe & Parks, 1999), and evaporation is not measured here by GLEAM, so no comparison can be made here. Interestingly, this is one of the regions where there is no overestimation of discharge yet.

Even though for the stations in Burundi data is often available for most of the modelled period, not all stations can be reliably used. As some of these stations have a catchment of only a couple of hundred square kilometres, which is not much when compared to the 100 km^2 square scale of the model. Furthermore, this only encompasses 1 % of the total area of the Nile Basin, so the detailed analysis is interesting here for just this region, not for the whole Nile Basin.

5.1.3 The Water Balance of the White Nile

The largest error in an upstream basin is to be found White Nile. The main cause of this is an overestimation of discharge in this segment of the Nile. A closer look at the average discharge gives a clear picture of what occurs on this stretch, especially when looking at the modelled discharge of the ungauged subbasins of the Sobat and Bahr el Ghazel. Figure 24 shows the overestimations which only increases as the White Nile continues towards Khartoum. Even if the model average yearly inflow at the Sudd is underestimated when compared to the FAO long term averages, the outflow of

the White Nile past its confluence with the Sobat is still much higher than. The large Bahr el Ghazel outflow also indicates that for the whole region of the White Nile discharge is overestimated.

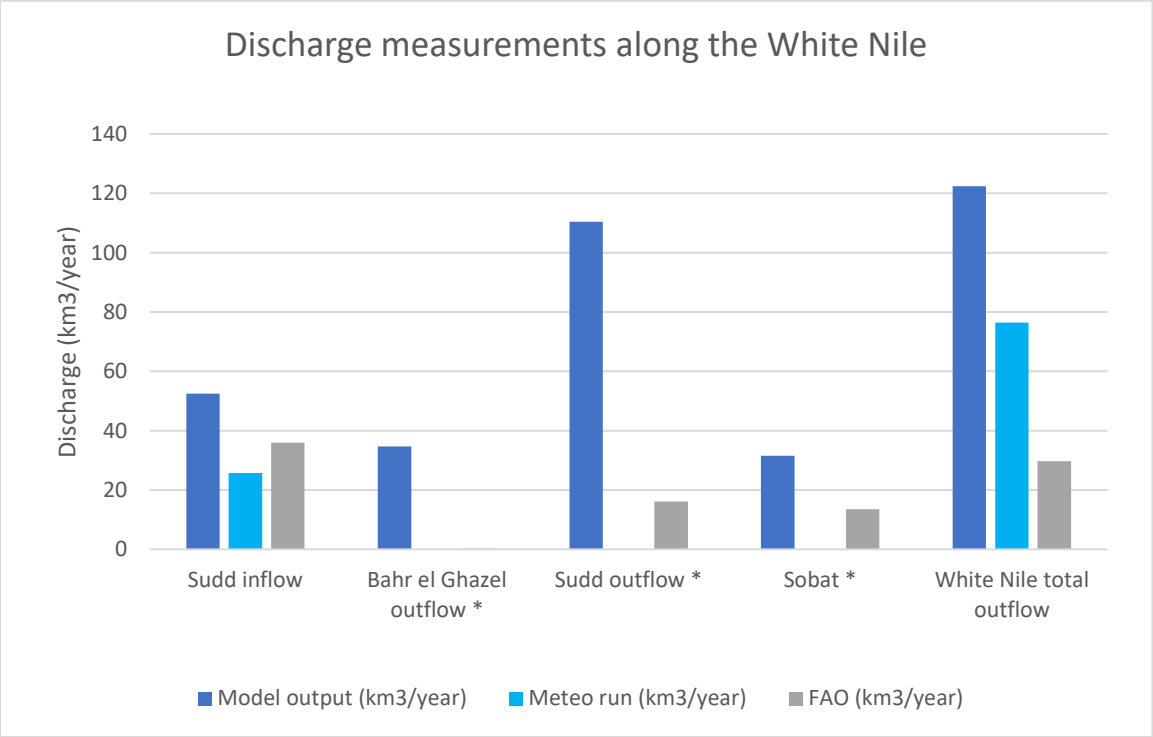

Figure 24: Modelled discharge compared to observed discharge from the FAO. For the Sudd inflow and White Nile total outflow the discharge of the Meteo run was additionally considered. The sites are ordered along from upstream to downstream.

Overestimation of discharge can be due to too much precipitation or too little evaporation, or a combination. Comparing the previous evaporation and precipitation estimates and measurements from Mohamed et al. (2004) (Figures 6 to 8) to the evaporation and precipitation output from the model over the Sudd region yields Figures 25 to 27. This shows that the model underreports evaporation when compared to satellite observations, especially in the drier months. Wetlands such as the Sudd, but also in the Bahr el Ghazel and Sobat subbasins are generally wet throughout the year, and this causes the large evaporation that is generally observed. Precipitation is also much lower for the model years when compared to the reference of Mohamed et al, this is due to the years in the model being years of drought. (Awadallah, 2014) as well as the model considering a larger area for each subbasin. Yet, this still shows much lower evaporation than predicted.

Interestingly, the results of the GLEAM comparison do not show large areas of high discrepancies in this part of the Nile. However, due to the relatively large size of the subbasins here small differences can in the end result in a large overestimation of discharge for the whole of the subbasins. Furthermore, there are still regions with up to 200 mm reported RMSE values, which is not negligible at all. So, although the Sudd does not light up as an area of high RMSE values, there is still likely an error here.

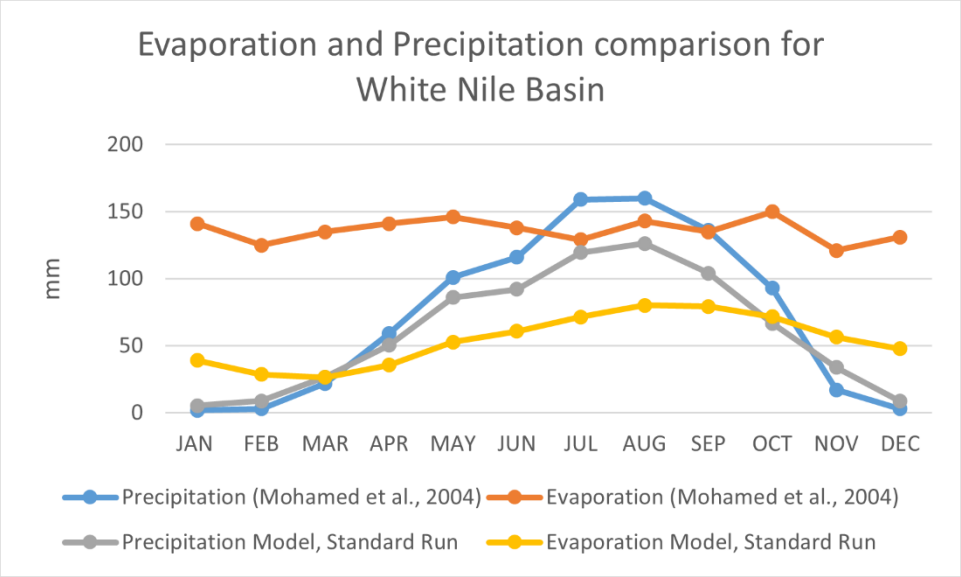

Figure 25: monthly evaporation and precipitation in the Sudd (from Mohamed et al., 2004) compared to output of the standard run of the model over the entire White Nile basin.

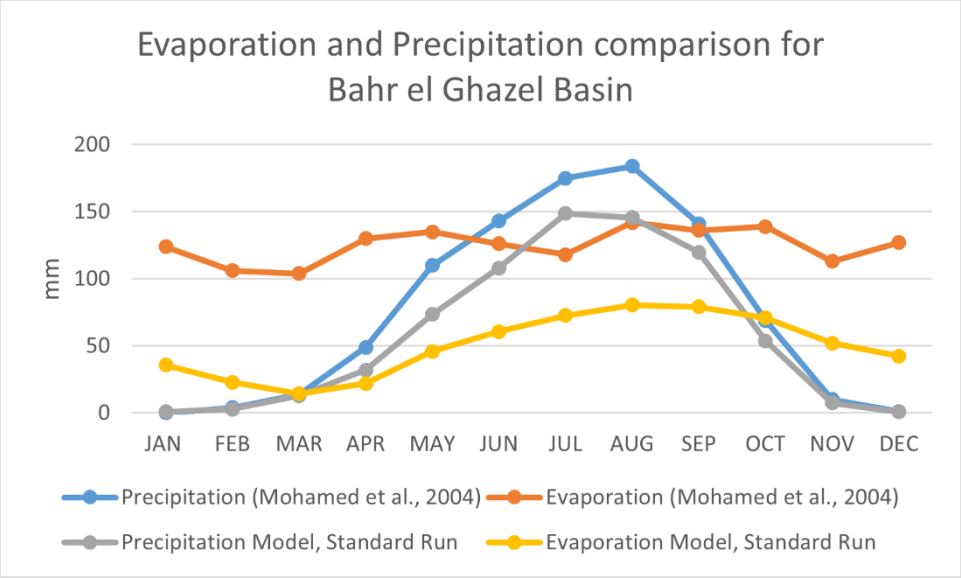

Figure 26: monthly evaporation and precipitation in the southern part of the Bahr el Ghazel (from Mohamed et al., 2004) compared to output of the standard run of the model in the entire Bahr el Ghazel basin.

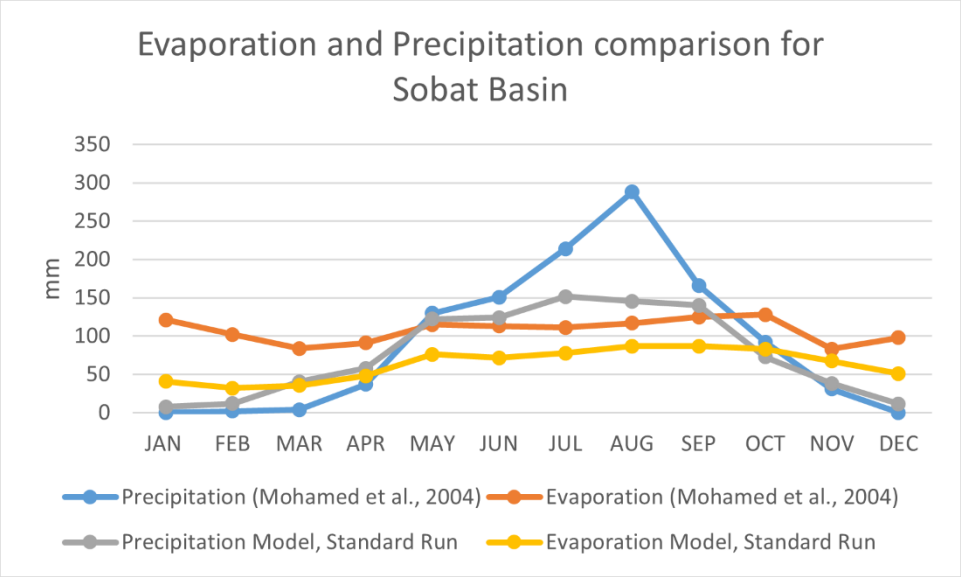

Figure 27: monthly evaporation and precipitation in the Sobat (from Mohamed et al., 2004) compared to output of the standard run of the model in the Sobat basin.

5.1.4 The Blue Nile and the Ethiopian Highlands

Looking at the results for the river originating in the Ethiopian Highlands is important as this is one of the main source regions of the Nile. The Blue Nile and the Atabarah both report the best values of all gauging stations. Showing that the model performs reasonably well in predicting the discharge. The performance of the Blue Nile is slightly better than the performance of the model for the Atbarah. For the evaporation comparison in the Ethiopian Highlands, some regions have quite high values of RMSE and also for the RRMSE. The largest region of high discrepancy is in the upstream part of the Baro river, which flows into the White Nile through the Sobat. However, this is an ungauged basin, so exact performance for this river is hard to gauge. Although it is very likely that due to this the model performance of this river would not be great, as can also be see in Figure 24 and the comparison to long term averages. For the middle part of Ethiopia, following the river from Lake Tana to Khartoum, the difference is not as large. The error in KGE score can thus be explained by the discrepancy reported of evaporation. Yet, the Blue Nile is still too wet in the dry months. The Atbarah performs worse than the Blue Nile, and this is can also be seen in the RMSE and RRMSE scores for North Ethiopia, being generally higher than for the Blue Nile. The same behaviour as in the Blue Nile also exists here where the Atbarah has too high discharge in the summer months.

5.1.5 Past Khartoum: error propagation

The poor results for the Main and Egyptian Nile can be explained in two ways, either by looking upstream, or by looking at the processes occurring there. A large part of the error is due to the White Nile already performing poorly as well as small error in the predicted Blue Nile discharge. These errors compound and for this reason the Main Nile results will be worse. In the Main Nile itself, KGE scores are quite similar for all three gauging sites, which expresses itself in the small plots in Figure 16. This means that between Khartoum and Lake Nasser, no large processes take place to improve or worsen the KGE scores. This is despite the evaporation being poorly modelled when compared to GLEAM. RRMSE scores around the river are generally quite high, but this is where the RMSE shows that actual differences in evaporation are low, so the effect is small.

5.1.6 Egypt and the Aswan dam

When looking at the difference between the Egyptian Nile and the Main Nile there KGE scores become a lot lower. The cause can be found at the Aswan High Dam. Although a dam operation scheme is present in the model, this does not work perfectly as is indicated by the KGE score decrease between before and after the Aswan dam. Secondly, in contrary to the Main Nile, the KGE scores do decrease downstream for the Egyptian Nile. This indicates that the water use is not represented right. However, as the Nile discharge is overestimated, much more water arrives at Aswan than would be realistic, so dam calculations are further hampered by this. Improvement of the behaviour in the Egyptian part of the Nile should start with improving the incoming water at Aswan.

Even though at Aswan part of the issue is due to the overestimation of incoming water, this is likely not the only reason. As stated by van Wirdum, (2017), PCR-GLOBWB poorly captures human influence, and in particular non-natural reservoirs. This can also be seen in other parts of the Nile for the results of the Standard run performed here. In both the Blue and White Nile, KGE scores decrease after large reservoirs such as the Sennar dam in the Blue Nile and the Jebel Aulia dam in the White Nile. The effect might partially be due to the errors of reported inflow but can also be attributed to the representation of reservoir operations by the model. Although compared to the error occurring at Aswan, the effect of these two other dams is smaller and especially for the White Nile it pales in comparison to error owing to the overestimations of discharge.

5.2 The (in)effect of the new datasets

The proper functioning of a hydrological model is dependent on several factors, but proper process description and proper input data are both a requirement. However, as PCR-GLOBWB 2 severely overestimates discharges and underestimates evaporation part of the model description for the Nile catchment does not match reality. For this reason, there are discrepancies between the simulated and observed discharges and evaporation as was shown above.

One way to improve a model's performance is to change parameters. Either updating them to a newer version or with a newer method that is thought to better present reality. The expected outcome of this is that if the input is thought to be better, the outcome should also be better. In this study this was tried to improve the performance of the Nile catchment for PCR-GLOBWB 2. This was done by using the three new datasets, and by using an improved routing method. However, in some cases, such as with the addition to Soilgrids, the outcome was worse than before for large regions. Or with the addition of a newer version of Meteorological forcing, the KGE values increased, but when compared to the GLEAM evaluation values were further off than with the standard parameterisation. Additionally, the introduction of a new dataset often yielded opposite results for various subbasins, such as the Meteo Run improving the performance for the Blue Nile when compared to the Standard run but decreasing performance for the Atbarah. This hints at different processes playing a role. Further exemplified by these regions being too wet in the dry seasons, but this will be touched upon later. Also adding the new routing method did not yield significant improvements.

What is then the cause of this poorer performance. Either the new parameters are worse, or alternatively, the process description itself is not sufficient.

This second reason might be valid in the Nile catchment, as PCR-GLOBWB 2 does not perform well here, reporting a large amount of overestimation of discharge for example. The concept of error propagation applies here in some way again, if the initial assumption is wrong then a calculation using this wrong assumption will result in a faulty answer even if the calculation uses the latest and

newest data to be performed. This might be the case for the worse performing basins in the Nile catchments as well. So, the paramount object should be to improve model description in the Nile catchment and only after that see which parameter might better describe the actual situation in a basin.

This problem also shows why the Meteo run might result in better terms of KGE values, and if that is the only thing that was checked it could be concluded that the Meteo run was objectively better for almost all of the Nile. However, upon comparing to GLEAM observations the Meteo run seems to perform worse in predicting the evaporation of the Nile. The concept of equifinality can be applied here as the outcome of the model in terms of discharge is better, but by decreasing the correctness of the evaporation. If only discharge was considered then the Meteo run would be better, but better for the wrong reason. Some process is not being described as it should, as indicating by the too wet dry seasons of the rivers in the Ethiopian Highlands, which is apparently something that is not fixed by just introducing a new dataset.

Additionally, incorporating the KWE for the routing, and thus allowing for flooding to occur in the model did not improve model behaviour either. Especially since this is an important factor in the hydrology of the Nile, this was expected to increase KGE scores. Hinting at that even this is not as important as the underlying problems within the model.

This is one of the problems for the Nile basin, as generally there is little high-quality information available. Model description is also poor, so changes to the model that would in theory make the model better end up with worse results because the process description is off. Any improvements that might be made by introducing new datasets or routing are overshadowed by larger model errors inherent in the model. And even though an analysis of the effects of the datasets is possible, the results of this would for large reasons be non sensical as the larger errors are the cause of the poor performance, not the data input.

5.3 Earth observation data for model evaluation

Validation of model results using gauging stations is still limited in the Nile River Basin. The temporal and spatial scale on which the analysis can be performed is very limited and to remedy this, it was chosen to try and validate model results with earth observation data, in this case the evaporation product GLEAM. The expected results would be that in basins which perform well, the GLEAM evaporation, expressed in terms of RMSE and RRMSE, would be close to model evaporation and therefore the error would be small. Then, the results of the different runs could be compared, and the influence of the new datasets could also be evaluated in terms of evaporation. However, this was not as clear as was hoped. Following will be an evaluation of this method for different subregions and for the whole Nile in general.

For the White Nile and its tributaries, comparing evaporation of the model to GLEAM is harder. Overestimation of discharge in this region already indicates that there is too much water in the system. However, when looking at regions which perform relatively well it is expected that a good comparison can be made for the evaporation, one of these regions is the Blue Nile and the Atbarah. Yet there are also large regions for those rivers, where there are still poor scores for the comparison metrics. This hints that even though KGE scores are high, the modelling of the accurate water balance in this region might not be accurate. Using the evaporation comparison highlights this and shows the equifinality of discharge measurements. Results might be good, but not for the right reasons.

One limitation of using the RMSE and RRMSE metrics in the current way is that these are directionless. Just indicating the difference, both expressed in mm/year and in % difference from the GLEAM evaporation value, but not whether the model overestimates of underestimates evaporation. For future use, this would be good to add. As for example, there is an underestimation of runoff from Lake Victoria and relatively large values for RMSE, so it is known that the model does not perform well here. The interpretation that this is due to an overestimation must then be made according to the discharge measurement. Showing the limits of using this method of validation independently. Especially since another region, such as the White Nile, might also have the same error in terms of RMSE, but here it is most likely an underestimation of evaporation. This cannot be seen with just the RMSE and RRMSE values, so another metric which does incorporate this direction of difference is better for an analysis of evaporation.

The results for the northern region of the Nile shows the necessity of using both RMSE and RRMSE values. As there is a large difference when looking at percentage difference, but due to the dry and water limited conditions the actual difference, as expressed in RMSE, is limited.

Another limitation is the temporal scale of the data availability and the comparisons that can be made from that. As the discharge gauging validation was done between 1979-1984 and the evaporation comparison was from 1980-1989. In the variable Nile Basin, that difference in years considered could impact performance of the model as those five years (1984-1989) are not gauged. For a better evaluation the current method of evaporation comparison can be repeated in a basin which is gauged during the timespan of the GLEAM observations, to better validate this method.

The version of GLEAM used does not have the open water evaporation of large bodies of water implemented, so this explains why errors are large near Lake Victoria. Furthermore, some gridcells also have missing value, usually near regions of open water, such as Lake Nasser, but not all of the lake is missing. This might explain some of the error around the edges of Lake Victoria, as averages are taking per gridcell, and this can then include a missing value together with a normal version, resulting in an erroneous depiction of the evaporation there. Open water evaporation is a problem for GLEAM itself, as the fraction of open water is determined from the product of Tuanmu & Jetz (2014) .This method has shown to have poor correlation in prediction of open water regions. Linking this back to the Nile, this is problematic, especially regarding the large amount of open water evaporation occurring in regions like the Sudd, where small lakes and inundated areas are dominant. Using a comparison to GLEAM in this region can then result in an incomplete analysis of the evaporation.

Another error in the current method can be traced back to the resolution used, namely 30 arc minute gridcells. A better way of doing this would be using the same gridsize as the model output itself, this would reduce the errors at location such as the borders or near large open water lakes as the results would not be averaged.

All in all, using GLEAM for validation provides interesting results, but has definite room for improvement. Especially in poorly gauged basins such as for the Nile. Furthermore, the direction of difference is also needed for a better insight in missing processes.

However, this method does show the possibility and necessity of looking at the whole of the water balance when validating a model. And in future, it is recommended to look at all the large components of the water balance, especially for the Nile basin, for a better understanding of the hydrology of a system, rather than just looking at one of the separate terms, such as discharge or evaporation.

5.4 Missing descriptions

PCR-GLOBWB 2 model performance for the Nile region is generally poor, with large overestimations of discharge. Especially for the White Nile, but also for the generally better performing Blue Nile. For the White Nile, the cause of this is an underestimation of evaporation, especially in the months with little precipitation. Although the comparison to Mohamed et al. (2004) is not perfect, as these are for different years and slightly different areas, the general pattern is clear and does show this lack of evaporation. The reason for this is the discrepancy between what the river is and how the model portraits the river, as well as some of the structures present in the model.

First, the Sudd region as well as parts of the Bahr el Ghazel and other wetlands regions are more like large inland deltas where the river enters a flat stretch and flow is reduced and backs up to form these huge swamps as well as large backwater effects from upstream are reasons for large flooding. This is not implemented, and in the model, there is just one main river which flows through these regions. This hinders flood behaviour as water drains too easily into these main channels. Furthermore, if flooding does occur, it can only flood the grid cells which immediately border the river itself, so this is limited to 10 km on each side for the current resolution. For the Sudd swamps, as an example, the maximum flooded region can at some points exceed 100 km in width. Furthermore, in the regions that do flood, water cannot infiltrate the soil from the floodplain, and this water either evaporates or drains out, further adding to this underestimation of evaporation as there is no water available to evaporate in the months of no precipitation. Added to this, in the normal set-up of the model no flooding occurs. Only if the KWE is used as the routing method then flooding is properly modelled. Currently the KWE does not work either, so an updated river morphology is also needed, especially as currently the rivers allow for too much water to be transported away.

For a better representation of these wetland regions these points mentioned above need to be implemented in the model structure, only after that the effects of new datasets can be evaluated. Poor river performance for other rivers which have wetlands is also found in other rivers, such as the Niger and the Murray. Although the reported reason for the poor performance of PCR-GLOBWB was water use (van Beek et al., 2011), these rivers also include large inland wetlands.

Several rivers might profit from a scheme like this as in-land deltas are found in the Niger or Murray rivers, where PCR-GLOBWB 2 also performs. Although reported reason for this poor performance is water use, all these rivers include large wetland areas.

Modelling wetlands such as the Sudd has historically proven to be hard. Implementing a more detailed scheme of the actual flow of the myriad of small streams which can be found in these regions. Trying to get this more accurate scheme to work requires more information that currently is not available. The actual extend of the Sudd swamps is one unknown, which most studies have based only on water balance calculations to predict the needed area for the amount of water to evaporate owing to the lack of in situ data available resulting in different reported extends based on different studies (Mohamed et al., 2004; Sutcliffe & Parks, 1999; Wilusz et al., 2017). Added to this that which channels are used depends on upstream inflow as well as on the Blue Nile discharges also adds increased complexity and uncertainty (Sutcliffe & Brown, 2018). Finding a solution that does not increase the complexity of a global model for regional problem remains one of the challenges going forward. Smaller changes to model structures, such as allowing floodwaters to infiltrate and water to pond more easily, are therefore preferable as these require less new data input that is just not available for regions such as the Sudd.

5.5 Future research

Ultimately, the current model predictions of the Nile River are not able to accurately model the behaviour of the Nile River. The implementation of new datasets has not yielded a large improvement in KGE scores for the subbasins and did not result in fixing the errors which occurred in the discharge of the Nile. However, there are some things which can still be tried to improve model outcome. Introduction of higher resolution modelling has resulted in better model performance for Europe (Hoch et al., 2023). This might also work for the Nile if the requisite datasets are available at high enough resolution. This, however, will not improve inundation effects if this is still limited to one grid cell away from the river and will likely decrease performance.

Looking at the water balance for the White Nile revealed some interesting behaviours which explained part of the poor performance. In lieu of gauging data, these water balance exercises can be expanded to the rest of the Nile basin and to other parts of the hydrologic cycle, such as precipitation, can be included to give a better indication of processes in the Nile Basin.

The water balance revealed a large error for parts of the basin for the dry months and a large seasonal difference. Evaluation of evaporation using GLEAM could also be done seasonally to look at those effects.

6 Conclusion

The performance of the global hydrological and water use model PCR-GLOBWB 2 for the Nile Basin is poor. The main cause of this is an overestimation of discharge when compared to discharge measurements from the GRDC. Although the amount of overestimation depends on the basin. The Blue Nile, the most important source of water for the Nile in general, reports a small overestimation of discharge in the dry months. Whereas the White Nile reports overestimations of discharge of more than three times the expected discharge. These errors persisted throughout the basin, with errors occurring from the earliest sources onwards.

Overestimation of discharge is linked to various processes within the model, most importantly for the White Nile, water is not retained in the system year-round, and this decreases evaporation, especially in the months where no runoff occurs. This has to do with lacking process description of flooding, which is an important process here. The model does not represent the wetland region as well and water cannot infiltrate into the ground when flooding occurs, also flooding extend is limited to gridsize next to the river and the area possible to be flooded therefore is too small.

Using the more updated datasets WFDE5 for meteorology, Soilgrids 2.0 for soil parameters and MERIT-DEM for topography has not yielded a large improvement as any improvements were overshadowed by the large inherent model structure errors.

Using the Kinematic Wave Equation (KWE) as the routing method did also not increase model performance and flooding did still not occur as was expected, this is partially to blame on a lack of input data for river morphology and also on the previously mentioned model errors when it comes to flooding behaviour.

Earth observation data was used to cover the spatial and temporal gaps present in the validation using GRDC river discharge gauges. This did show some regions where evaporation from the model showed large discrepancies with the GLEAM evaporation model, which was used as a benchmark to compare against. However, it did give a rough inside of where these errors occurred. In all, it seemed to be a good first step in closing the water balance, even though the results of evaporation comparisons were not as clear.

Improvements to model structures, such as including a better description of the wetland hydrology should be a first step for improving model results in the Nile Basin. Although this might be hard as the input data to properly describe the wetlands in the Nile Basin is limited.

All in all, the Nile Basin remains a hard basin to model properly with large errors occurring in the prediction of discharge. The results vary between different subbasins and the reason for the poor performance is not always the same. Although performance of PCR-GLOBWB 2 for the Nile Basin is poor, the current research has showed where some of the errors lay and where it can be improved for future.

7 References

AbuZeid, K. M. (2020). Existing and Recommended Water Policies in Egypt. In S. Zekri (Ed.), *Water Policies in MENA Countries* (pp. 47–62). Springer International Publishing. https://doi.org/10.1007/978-3-030-29274-4_3

Abu-Zeid, M. A., & El-Shibini, F. Z. (1997). Egypt's High Aswan Dam. *International Journal of Water Resources Development*, *13*(2), 209–218. https://doi.org/10.1080/07900629749836

Allen, R. G. P. L. S. R. D. S. M. (1998). Crop evaporation: Guidelines for computing crop requirements, . *UN-FAO, Rome*.

Awadallah, A. (2014). Evolution of the Nile River drought risk based on the streamflow record at Aswan station, Egypt. *Civil Engineering and Environmental Systems*, *31*. https://doi.org/10.1080/10286608.2013.853747

Bierkens, M. F. P. (2015). Global hydrology 2015: State, trends, and directions. *Water Resources Research*, *51*(7), 4923–4947. https://doi.org/https://doi.org/10.1002/2015WR017173

Carlson, A. (2013). *Who Owns the Nile? Egypt, Sudan, and Ethiopia's History-Changing Dam*.

Coffel, E. D., Keith, B., Lesk, C., Horton, R. M., Bower, E., Lee, J., & Mankin, J. S. (2019). Future Hot and Dry Years Worsen Nile Basin Water Scarcity Despite Projected Precipitation Increases. *Earth's Future*, *7*(8), 967–977. https://doi.org/https://doi.org/10.1029/2019EF001247

Conway, D., & Hulme, M. (1993). Recent fluctuations in precipitation and runoff over the Nile sub-basins and their impact on main Nile discharge. *Climatic Change*, *25*(2), 127–151. https://doi.org/10.1007/BF01661202

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., & Buontempo, C. (2020). WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, *12*(3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020

Despotovic, M., Nedic, V., Despotovic, D., & Cvetanovic, S. (2016). Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renewable and Sustainable Energy Reviews*, *56*, 246–260. https://doi.org/https://doi.org/10.1016/j.rser.2015.11.058

Di Baldassarre, G., Elshamy, M., van Griensven, A., Soliman, E., Kigobe, M., Ndomba, P., Mutemi, J., Mutua, F., Moges, S., Xuan, Y., Solomatine, D., & Uhlenbrook, S. (2011). Future hydrology and climate in the River Nile basin: A review. *Hydrological Sciences Journal – Journal Des Sciences Hydrologiques*, *56*, 199–211. https://doi.org/10.1080/02626667.2011.557378

di Vittorio, C. A., & Georgakakos, A. P. (2021). Hydrologic Modeling of the Sudd Wetland using Satellite-based Data. *Journal of Hydrology: Regional Studies*, *37*, 100922. https://doi.org/10.1016/J.EJRH.2021.100922

Döll, P., Kaspar, F., & Lehner, B. (2003a). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, *270*(1–2), 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Döll, P., Kaspar, F., & Lehner, B. (2003b). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, *270*(1–2), 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Eltahir, E. A. B. (1996). El Niño and the Natural Variability in the Flow of the Nile River. *Water Resources Research*, *32*(1), 131–137. https://doi.org/https://doi.org/10.1029/95WR02968

FAO. (1974). *FAO-Unesco Soil map of the world Volume I Legend*.

FAO, IIASA, ISRIC, ISSCAS, & JRC. (2012). *Harmonized World Soil Database - HWSD (version 1.2)*.

Gosling, S. N., & Arnell, N. W. (2011). Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. *HYDROLOGICAL PROCESSES Hydrol. Process*, *25*, 1129–1145. https://doi.org/10.1002/hyp.7727

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., … Yeh, P. (2011). Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. *Journal of Hydrometeorology*, *12*(5), 869–884. https://doi.org/10.1175/2011JHM1324.1

Hamad, O. E.-T., & El-Battahani, A. (2005). Sudan and the Nile Basin. *Aquatic Sciences*, *67*(1), 28–41. https://doi.org/10.1007/s00027-004-0767-9

Hamon, W. R. (1963). Computation of direct runoff amounts from storm rainfall. *International Association of Scientific Hydrology Publication*, *63*, 52–62.

Hamza, W. (2006). The Nile Estuary. In P. J. Wangersky (Ed.), *Estuaries* (pp. 149–173). Springer Berlin Heidelberg. https://doi.org/10.1007/698_5_025

Harding, R., Best, M., Blyth, E., Hagemann, S., Kabat, P., Tallaksen, L. M., Warnaars, T., Wiberg, D., Weedon, G. P., Lanen, H. van, Ludwig, F., & Haddeland, I. (2011). WATCH: Current Knowledge of the Terrestrial Global Water Cycle. *Journal of Hydrometeorology*, *12*(6), 1149–1156. https://doi.org/10.1175/JHM-D-11-024.1

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International Journal of Climatology*, *34*(3), 623–642.

Hoch, J. M., Sutanudjaja, E. H., Wanders, N., van Beek, R. L. P. H., & Bierkens, M. F. P. (2023). Hyper-resolution PCR-GLOBWB: opportunities and challenges from refining model spatial resolution to 1km over the European continent. *Hydrology and Earth System Sciences*, *27*(6), 1383–1401. https://doi.org/10.5194/hess-27-1383-2023

Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., & Bierkens, M. F. P. (2010). A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling & Software*, *25*(4), 489–502. https://doi.org/10.1016/J.ENVSOFT.2009.10.004

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424–425*, 264–277. https://doi.org/https://doi.org/10.1016/j.jhydrol.2012.01.011

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Koutsoyiannis, D. (2005). Hydrologic Persistence and The Hurst Phenomenon. In *Water Encyclopedia* (pp. 210–221). John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/047147844X.sw434

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1–4), 1–22.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., & Wisser, D. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, *9*(9), 494–502. https://doi.org/https://doi.org/10.1890/100125

MANABE, S. (1969). CLIMATE AND THE OCEAN CIRCULATION: I. THE ATMOSPHERIC CIRCULATION AND THE HYDROLOGY OF THE EARTH'S SURFACE. *Monthly Weather Review*, *97*(11), 739–774. https://doi.org/10.1175/1520-0493(1969)097<0739:CATOC>2.3.CO;2

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. C. (2017). GLEAM~v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Misiani, H. O., Finney, D. L., Segele, Z. T., Marsham, J. H., Tadege, A., Artan, G., & Atheru, Z. (2020). Circulation Patterns Associated with Current and Future Rainfall over Ethiopia and South Sudan from a Convection-Permitting Model. *Atmosphere*, *11*(12). https://doi.org/10.3390/atmos11121352

Modi, P., Revel, M., Zhou, X., & Yamazaki, D. (2020). MERIT DEM Performs Better for Hydrodynamic Flood Model in Amazon Basin. *AGU Fall Meeting Abstracts*, *2020*, H010-0013.

Mohamed, Y. A., Bastiaanssen, W. G. M., & Savenije, H. H. G. (2004). Spatial variability of evaporation and moisture storage in the swamps of the upper Nile studied by remote sensing techniques. *Journal of Hydrology*, *289*(1), 145–164. https://doi.org/https://doi.org/10.1016/j.jhydrol.2003.11.038

Mohamed, Y. A., Savenije, H. H. G., Bastiaanssen, W. G. M., & van den Hurk, B. J. J. M. (2006). New lessons on the Sudd hydrology learned from remote sensing and climate modeling. *Hydrology and Earth System Sciences*, *10*(4), 507–518. https://doi.org/10.5194/hess-10-507-2006

Oestigaard, T. (2012). *Water Scarcity and Food Security along the Nile : Politics, population increase and climate change*.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

Roth, V., Lemann, T., Zeleke, G., Subhatu, A. T., Nigussie, T. K., & Hurni, H. (2018). Effects of climate change on water resources in the upper Blue Nile Basin of Ethiopia. *Heliyon*, *4*(9), e00771. https://doi.org/10.1016/J.HELIYON.2018.E00771

Sanchez, P. A., Ahamed, S., Carré, F., Hartemink, A. E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A. B., McKenzie, N. J., de Lourdes Mendonça-Santos, M., Minasny, B., Montanarella, L., Okoth, P., Palm, C. A., Sachs, J. D., Shepherd, K. D., Vågen, T.-G., Vanlauwe, B.,

Walsh, M. G., … Zhang, G.-L. (2009). Digital Soil Map of the World. *Science*, *325*(5941), 680–681. https://doi.org/10.1126/science.1175084

Schellekens, J., Dutra, E., la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., … Weedon, G. P. (2017). A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset. *Earth System Science Data*, *9*(2), 389–413. https://doi.org/10.5194/essd-9-389-2017

Sood, A., & Smakhtin, V. (2015). Global hydrological models: a review. *Hydrological Sciences Journal*, *60*(4), 549–565. https://doi.org/10.1080/02626667.2014.950580

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., & Bierkens, M. F. P. (2011). Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin. *Hydrology and Earth System Sciences*, *15*(9), 2913–2935. https://doi.org/10.5194/hess-15-2913-2011

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., & Bierkens, M. F. P. (2014). Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. *Water Resources Research*, *50*(1), 687–705. https://doi.org/https://doi.org/10.1002/2013WR013807

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., & Bierkens, M. F. P. (2018). PCR-GLOBWB~2: a 5arcmin global hydrological and water resources model. *Geoscientific Model Development*, *11*(6), 2429–2453. https://doi.org/10.5194/gmd-11-2429-2018

Sutcliffe, J., & Brown, E. (2018). Water losses from the Sudd. *Hydrological Sciences Journal*, *63*(4), 527–541. https://doi.org/10.1080/02626667.2018.1438612

Sutcliffe, J., Hurst, S., Awadallah, A. G., Brown, E., & Hamed, K. (2016). Harold Edwin Hurst: the Nile and Egypt, past and future. *Hydrological Sciences Journal*, *61*(9), 1557–1570. https://doi.org/10.1080/02626667.2015.1019508

Sutcliffe, J. v. (2009a). The Hydrology of the Nile Basin. In H. J. Dumont (Ed.), *The Nile: Origin, Environments, Limnology and Human Use* (pp. 335–364). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9726-3_17

Sutcliffe, J. v. (2009b). The Hydrology of the Nile Basin. In H. J. Dumont (Ed.), *The Nile: Origin, Environments, Limnology and Human Use* (pp. 335–364). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9726-3_17

Sutcliffe, J. V, & Parks, Y. P. (1999). *The Hydrology of the Nile The International Water Management Institute*.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., Sutanudjaja, E. H., Sun, Y., Xia, T., & Wang, Z. (2017). Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model. *Hydrol. Earth Syst. Sci.*, *21*(4), 2053–2074. https://doi.org/10.5194/hess-21-2053-2017

Tuanmu, M.-N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, *23*(9), 1031–1045. https://doi.org/https://doi.org/10.1111/geb.12182

United Nations, D. of E. and S. A. P. D. (2017). *World Population Prospects: the 2017 Revision, Key Findings and Advance Tables*.

van Beek, L. P. H. (2008). Forcing PCR-GLOBWB with CRU meteorological data. *Utrecht University, Utrecht, Netherlands: Http://Vanbeek. Geo. Uu. Nl/Suppinfo/Vanbeek2008. Pdf*.

van Beek, L. P. H., Eikelboom, T., van Vliet, M. T. H., & Bierkens, M. F. P. (2012). A physically based model of global freshwater surface temperature. *Water Resources Research*, *48*(9). https://doi.org/https://doi.org/10.1029/2012WR011819

van Beek, L. P. H. (Rens), & Bierkens, M. F. P. (2009). *The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification Report*. http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

van Beek, L. P. H., Wada, Y., & Bierkens, M. F. P. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, *47*(7). https://doi.org/https://doi.org/10.1029/2010WR009791

van Wirdum, C. (2017). *Simulating hydrology over complex terrain: improving the performance of PCR-GLOBWB over the Nile Basin*.

Viviroli Daniel and Weingartner, R. (2008). ``Water Towers''—A Global View of the Hydrological Importance of Mountains. In E. Wiegandt (Ed.), *Mountains: Sources of Water, Sources of Knowledge* (pp. 15–20). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6748-8_2

Wada, Y., van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., & Bierkens, M. F. P. (2011). Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research*, *47*(7). https://doi.org/https://doi.org/10.1029/2010WR009792

Wada, Y., Wisser, D., & Bierkens, M. F. P. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40. https://doi.org/10.5194/esd-5-15-2014

Wheeler, K. G., Jeuland, M., Hall, J. W., Zagona, E., & Whittington, D. (2020). Understanding and managing new risks on the Nile with the Grand Ethiopian Renaissance Dam. *Nature Communications*, *11*(1), 5222. https://doi.org/10.1038/s41467-020-19089-x

Williams, M., Snow, M., Self, P., Raven, M., & Cowan, E. (2022). Depositional environments in the White Nile Valley during the last 300,000 years. *Journal of Palaeosciences*, *71*, 19–43. https://doi.org/10.54991/jop.2022.36

Wilusz, D. C., Zaitchik, B. F., Anderson, M. C., Hain, C. R., Yilmaz, M. T., & Mladenova, I. E. (2017). Monthly flooded area classification using low resolution SAR imagery in the Sudd wetland from 2007 to 2011. *Remote Sensing of Environment*, *194*, 205–218. https://doi.org/10.1016/J.RSE.2017.03.005

Yamazaki, D., Ikeshima, D., Neal, J. ~C., O'Loughlin, F., Sampson, C. ~C., Kanae, S., & Bates, P. ~D. (2017). MERIT DEM: A new high-accuracy global digital elevation model and its merit to global hydrodynamic modeling. *AGU Fall Meeting Abstracts*, *2017*, H12C-04.

8 Appendix

Appendix A List of model inputs and parameters 2 (Sutanudjaja et al., 2018)

Description	Symbol	Unit	References/sources
Upper and lower soil store parameters			FAO (2007) soil map; van Beek and Bierkens (2009)
– Soil thickness	Z_1 and Z_2	m	
– Residual soil moisture content	$\theta_{r\text{-}1}$ and $\theta_{r\text{-}2}$	$m^3\ m^{-3}$	
– Soil moisture at saturation	$\theta_{s\text{-}1}$ and $\theta_{s\text{-}2}$	$m^3\ m^{-3}$	
– Soil water storage capacity per soil layer: $SC = Z/(\theta_s - \theta_r)$	SC_1 and SC_2	m	
– Soil matric suctions at saturation	$\psi_{s\text{-}1}$ and $\psi_{s\text{-}2}$	m	
– Exponent in the soil water retention curve	β_1 and β_2	dimensionless	
– Saturated hydraulic conductivities of upper and lower soil stores	K_1 and K_2	$m\ day^{-1}$	
– Total soil water storage capacities $= SC_{upp} + SC_{low}$	W_{max}	m	
Land cover fraction: land cover areas (including extent of irrigated areas) over cell areas	f_{lcov}	$m^2\ m^{-2}$	GLCC v2.0 map (USGS, 1997); Olson (1994a, b); MIRCA2000 data set (Portmann et al., 2010); FAOSTAT (2012)
Topographical parameters	DEM	m	HydroSHEDS (Lehner et al., 2008); Hydro1k (Verdin and Greenlee, 1996); GTOPO30 (Gesch et al., 1999)
– Cell-average DEM	DEM_{avg}	m	
– Floodplain elevation	DEM_{fpl}	m	
Root fractions per soil layer	Rf_{upp} & Rf_{low}	dimensionless	Canadell et al. (1996); van Beek and Bierkens (2009)
Arno scheme (Todini, 1999; Hagemann and Gates, 2003) exponents defining soil water capacity distribution	β_{arno}	dimensionless	Canadell et al. (1996); Hagemann et al. (1999); Hagemann (2002); van Beek (2008); van Beek and Bierkens (2009)
Ratio of cell-minimum soil storage to W_{max}	f_{wmin}	$m\ m^{-1}$	van Beek (2008); van Beek and Bierkens (2009)
Ratio of cell-maximum soil storage to W_{max}	f_{wmax}	$m\ m^{-1}$	van Beek (2008); van Beek and Bierkens (2009)
Parameters related to phenology			Hagemann et al. (1999); Hagemann (2002); van Beek (2008); van Beek and Bierkens (2009)
– Crop coefficient	K_c	dimensionless	
– Interception capacity	$S_{int\text{-}max}$	m	
– Vegetation cover fraction	C_v	$m^2\ m^{-2}$	
Groundwater parameters			GLHYMPS map (Gleeson et al., 2014); van Beek (2008); van Beek and Bierkens (2009)
– Aquifer transmissivity	KD	$m^2\ day^{-1}$	
– Aquifer specific yield	Sy	$m^3\ m^{-3}$	
– Groundwater recession coefficient	J^{-1}	day^{-1}	
Meteorological forcing			van Beek (2008); CRU (Harris et al., 2014); ERA40 (Uppala et al., 2005); ERA-Interim (Dee et al., 2011)
– Total precipitation	P	$m\ day^{-1}$	
– Atmospheric air temperature	T_{air}	°C or K	
– Reference potential evaporation and transpiration	$E_{ref,pot}$	$m\ day^{-1}$	
Others			
– Non-irrigation sectoral water demand (i.e. livestock, domestic, and industrial)		$m\ day^{-1}$	Wada et al. (2014)
– Desalinated water		$m\ day^{-1}$	Wada et al. (2011a); FAO (2016)
– Lakes and reservoirs			GLWD1 (Lehner and Döll, 2004); GRanD (Lehner et al., 2011)

Appendix B List (non-exhaustive) of state and flux variables defined in PCR-GLOBWB 2 (Sutanudjaja et al., 2018)

Description	Symbol	Unit
Interception storage	S_{int}	m
Snow cover/storage in water equivalent thickness (excluding liquid part S_{slq})	S_{swe}	m
Liquid/meltwater storage in the snowpack	S_{slq}	m
Upper and lower soil storages	S_1 and S_2	m
Surface water storage (lakes, reservoirs, rivers, and inundated water)	S_{wat}	m
Groundwater storage (renewable part)	S_3	m
Fossil groundwater storage (non-renewable)	S_{nrw}	m
Total groundwater storage $= S_3 + S_{nrw}$	S_{gwt}	m
Total water storage thickness $= S_{int} + S_{swe} + S_{slq} + S_1 + S_2 + S_{gwt}$	TWS	m
Potential evaporation	E_{pot}	m day^{-1}
Evaporation flux from the intercepted precipitation	E_{int}	m day^{-1}
Evaporation from meltwater stored in the snowpack	E_{slq}	m day^{-1}
Bare soil evaporation	E_{soil}	m day^{-1}
Transpiration from the upper and lower soil stores	T_1 and T_2	m day^{-1}
Total land evaporation $= E_{int} + E_{slq} + E_{soil} + T_1 + T_2$	E_{land}	m day^{-1}
Surface water evaporation	E_{wat}	m day^{-1}
Total evaporation $= E_{land} + E_{wat}$	E_{tot}	m day^{-1}
Direct run-off	Q_{dr}	m day^{-1}
Interflow, shallow sub-surface flow	Q_{sf}	m day^{-1}
Baseflow, groundwater discharge	Q_{bf}	m day^{-1}
Specific run-off from land	Q_{loc}	m day^{-1}
Local change in surface water storage	Q_{wat}	m day^{-1}
Total specific run-off	Q_{tot}	m day^{-1}
Routed channel (surface water) discharge	Q_{chn}	m^3 s^{-1}
Net fluxes from the upper to lower soil stores	Q_{12}	m day^{-1}
Net groundwater recharge, fluxes from the lower soil to groundwater stores	RCH $= Q_{23}$	m day^{-1}
Surface water infiltration to groundwater	Inf	m day^{-1}
Desalinated water withdrawal	W_{sal}	m day^{-1}
Surface water withdrawal	W_{wat}	m day^{-1}
Renewable groundwater withdrawal	W_3	m day^{-1}
Non-renewable groundwater withdrawal (groundwater depletion)	W_{nrw}	m day^{-1}
Total groundwater withdrawal $= W_3 + W_{nrw}$	W_{gwt}	m day^{-1}
Water withdrawal allocated for irrigation purposes	A_{irr}	m day^{-1}
Water withdrawal allocated for livestock demand/sector	A_{liv}	m day^{-1}
Water withdrawal allocated for agricultural sector $= A_{irr} + A_{liv}$	A_{agr}	m day^{-1}
Domestic water withdrawal	A_{dom}	m day^{-1}
Industrial water withdrawal	A_{ind}	m day^{-1}

Appendix C: List of all gauging sites and the KGE scores of each of these sites for all the runs.

Code + Name	Code	Standard	Soil	Topo	Meteo	KGE
E1 EL EKHSASE	E1	-9.61	-12.76	-9.61	-7.47	-11.36
E2 ASSIUT	E2	-7.99	-10.78	-7.99	-6.10	-9.88
E3 NAH HAMMADI	E3	-6.40	-8.70	-6.39	-4.93	-7.89
E4 ESNA	E4	-5.67	-7.76	-5.66	-4.38	-7.05
E5 GAAFRA	E5	-4.87	-6.73	-4.86	-3.74	-6.12
E6 ASWAN DAM	E6	-6.00	-8.37	-6.00	-4.58	-6.38
M1 DONGOLA	M1	-1.44	-2.42	-1.42	-0.06	-2.44
M2 HUBEIDA + HASSANAE	M2	-1.29	-2.25	-1.26	-0.31	-1.79
M3 TAMANIAT	M3	-1.36	-2.35	-1.33	-0.40	-1.83
B1 KHARTOUM (BN)	B1	0.31	-0.05	0.30	0.50	0.18
B2 SENNAR	B2	0.23	-0.11	0.22	0.48	0.11
B3 ROSEIRES DAM	B3	0.58	0.43	0.58	0.64	0.52
A1 KILO 3	A1	0.23	0.25	0.24	-0.37	0.23
A2 DOWNSTREAM KHASM EL GIBRA DAM	A2	0.10	0.26	0.11	-0.30	0.13
W1 MOGRAN	W1	-7.76	-10.41	-7.60	-5.50	-8.67
W2 DOWNSTREAM OF JEBEL AULIA DAM	W2	-8.54	-11.93	-8.36	-5.62	-8.47
W3 MELUT	W3	-5.77	-7.71	-5.61	-3.60	-7.00
W4 MALAKAL	W4	-5.50	-7.24	-5.36	-3.44	-6.79
W5 KILO 3.2	W5	-0.43	-0.61	-0.38	-0.27	-0.40
W6 MALEK	W6	-5.66	-7.21	-5.90	-3.23	-11.46
W7 MONGALLA	W7	-0.76	-1.30	-0.81	-0.12	-2.78
W8 OUTLET	W8	-1.27	-1.87	-1.42	-0.62	-0.56
G0 OWEN RESERVOIR	G0	-0.08	-0.12	-0.04	-0.30	-1.78
G1 BAC (MUYINGA)	G1	0.37	0.24	0.38	0.33	0.35
G2 GITENGA	G2	0.06	0.18	0.06	-0.09	0.05
G3 GITONO	G3	-0.24	-0.19	-0.23	-0.33	-0.24
G4 BURASIRA	G4	-0.31	-0.21	-0.30	-0.34	-0.31
G5 KANBUSORO	G5	0.31	0.14	0.31	0.30	0.30
G6 DISPENSAIRE	G6	-0.16	-0.06	-0.15	-0.32	-0.16
G7 KIBAYA	G7	0.29	0.31	0.30	0.20	0.29
G8 MUYAGE 2	G8	0.19	0.32	0.19	0.16	0.18
G9 MUBUGA	G9	-5.35	-6.63	-5.38	-1.68	-5.39
G10 NYANKANDA	G10	-0.50	-0.32	-0.46	-0.42	-0.51
G11 NGOZI-BUTARE	G11	0.12	0.24	0.13	-0.03	0.11
G12 KIGALI	G12	0.24	0.01	0.26	0.19	0.21
G13 KANZENZE	G13	-0.05	-0.39	-0.05	0.37	-0.15
G14 RUSOMO	G14	0.06	-0.34	0.06	0.43	-0.01