Master Thesis
TVVR 19/5015

Hydropower forecast in Ethiopia using the HBV model

A case study in Omo-Gibe river basin

Youssef Faouzi

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University

Hydropower forecast in Ethiopia using the HBV model

A case study in Omo-Gibe river basin

By:

Youssef Faouzi

Master Thesis

Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University
Box 118
221 00 Lund, Sweden

Water Resources Engineering
TVVR-19/5015
ISSN 1101-9824

Lund 2019
www.tvrl.lth.se

Master Thesis
Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University

Swedish Title:	Vattenkraftsprognos i Etiopien med HBV-modellen
English Title:	Hydropower forecast in Ethiopia using the HBV model
Author:	Youssef Faouzi
Supervisor:	Magnus Persson, Water Resources Engineering at LTH; Stefan Söderberg, Hydrological Research and Forecast at Refinitiv
Examiner:	Linus Zhang, Water Resources Engineering at LTH
Language:	English
Year:	2019
Keywords:	Hydropower; Omo-Gibe; HBV model; rainfall-runoff; Ethiopia

Acknowledgments

I would like first to express my gratitude to my parents for having supported me throughout this thesis project, support without which this work would have been daunting. I would like to extend this gratitude to my supervisor Magnus Persson for his valuable advice and suggestions during this project, to Stefan Söderberg at Refinitiv for offering me a chance to be part of this project and overseeing the direction this project was taking, to the whole team at Refinitiv for lending a helping hand whenever they could especially Marion Guegan for her considerable help and feedback on this work.

Abstract

Ethiopia produces 86% of its electricity from hydropower source, still they experience frequent power shortages. As they are part of the East African Power Pool, this power shortage is usually made up for by buying energy from other members of the pool. Forecasting the energy available for usage in the future allows Ethiopia to buy energy ahead of time when prices are low, but also to sell surplus electricity to neighboring countries. This is the degree project I was offered to investigate into at the company Refinitiv. To do so, a modified HBV model was used to predict the hydroelectricity, using as input one of the two weather datasets (CPCP and CFSR), both provided by the company, and a Q-target series against which the computed runoff is evaluated. The task of constructing a Q-target series meant that an extensive search for runoff data was necessary. After finding such data and entering the Q-target series in the model, the model's performance was assessed during the calibration and validation parts. The model using CFSR data was found to perform better than CPCP's by a substantial margin. Nonetheless, the results are promising for both models. Suggestions for further work in this subject relate to the quality of the runoff data used, as the data used in this project originated from a secondary source, and missing runoff data was filled in using unknown interpolation equations. For this reason, it is recommended to get the raw data from the primary source, namely the Ministry of Water, Irrigation and Energy. This would allow for greater flexibility during the construction of the Q-target series.

Contents

1. Introduction

1.1 Background

Ethiopia generates 86% of its electricity from hydropower, the economically feasible potential was estimated by the Ethiopian government to be around 45 GW, of which only 8.5% is currently utilized. The remaining untapped potential is huge and can help achieve Ethiopia's goal of becoming a lower middle-income country by 2025. For instance, it could make Ethiopia a reliable regional power hub, through selling energy to neighboring countries when in surplus of energy. An important feature of hydropower is that it usually, but not always, relies on water stored in reservoirs, which can be used for irrigation purposes and thus moderate the effects of the frequent drought periods Ethiopia is subject to (Power Africa, 2016).

As Ethiopia is part of the East African Power Pool, a 10-country power pool, the infrastructure for power trade between the pool power members is already present. A 2030 projection has Ethiopia the country with the most power surplus of all countries of the pool, this is contingent on Ethiopia's efforts in building more hydropower plants by 2030. In order to do so, a model-based estimation, rather than a crude estimate, of the hydropower potential that the many rivers of Ethiopia offer is clearly warranted (Deloitte, 2015).

This is the task Refinitiv (formerly known as Thomson Reuters) has set for me to investigate, since they supply hydropower generation forecasts for a number of countries in the world using a modified version of the HBV model.

The HBV model is a hydrological model that can be used to compute river discharge using only precipitation and temperature data for a given catchment area (Bergström, 1992). After calibration, this model uses the precipitation and temperature as inputs and gives as output the estimated river runoff. This inflow to reservoirs can be then linked to the energy produced by hydropower plants.

The HBV model needs observed runoff at first to calibrate its parameters: observed river and stream flows are needed to calibrate the model, the longer

the data series, the more accurate the model gets. Once the model is calibrated for a certain river basin or area, weather forecast data can be used to predict river flows, which in turn is used to predict the energy that can be generated from hydropower in the near future.

Predicting the available energy in the near future is useful since it allows to meet potential power shortages by buying electricity from neighboring countries, or on the contrary sell surplus electricity to neighboring countries, in essence it helps minimize the costs of electricity production.

1.2 Aims and objectives

The aim of this degree project is to build and analyze the performance of a hydrological model that can be used to forecast the energy production from hydropower source in the Omo-Gibe river basin in Ethiopia.

In order to achieve this goal, the work proper was divided into two main objectives: those related to building the Q-target series, and those to calibrate the HBV model and run it.

- *Construct the Q-target series:*
  1. Find hydrological data
  2. Select streamflow stations
  3. Weigh streamflow stations
  4. Construct the Q-target series

- *Calibrate and validate the model:*
  1. Choose the appropriate precipitation and temperature polygons
  2. Calibrate the HBV model
  3. Validate the model
  4. Compare the performance of the model

1.3 Methodology

A literature study has been done to learn about the energy market in Ethiopia, the amount of hydroelectricity produced each year, the biggest energy-producing power plants, and Ethiopia's current energy wholesale to neighboring countries.

Then, the HBV model is introduced along some parameters. This rainfall-runoff model needs precipitation and temperature data as input and a Q-target series in order to calibrate the model. Two precipitation datasets were provided by Refinitiv, this resulted in calibrating two distinct models each one using a different precipitation dataset, this also meant that one model could be compared to the other. The source of weather data is discussed at length in the data collection section. As for building a Q-target series, an extensive search has been followed to find streamflow data for hydrological stations. Some of the criteria applied for choosing the right gauging stations are: nature of the flow – natural or disrupted by human activities, distance to power plants, stations upstream of the hydropower plant in question, quality and quantity of data.

Once the Q-target series was built, the calibration process was carried with changes made to the model parameters until no substantial improvement to the r^2 value could be achieved. Both models were validated then and their performance analyzed and compared against each other, before concluding the project with some suggestions on how the model could be improved upon.

1.4 Limitations

Some limitations are inherently present in this kind of work. The first limitation to consider relates to the building of the Q-target series as it is not only dependent on the choice of the streamflow stations to use but also the weight given to each station. As there is no exact method to follow in order to choose and weigh the stations, this step remains to some extent subjective.

Secondly, both precipitation and temperature data used as input are not measured data but modelled data, this means that the input is synthetic and therefore adjustments are an intrinsic part of the input. Furthermore, the HBV model itself makes assumptions and therefore uses a set of parameters that may not be relevant to the case at hand. Finally, the energy prediction step doesn't consider whether or not the energy produced by the hydropower plants happens when said power plants are used to their fullest capacity.

2. Literature study

2.1 Introduction of Ethiopia

Ethiopia is a country located in East Africa, specifically in the region of the Horn of Africa, and is home to about 110 million inhabitants which makes it the second most populous country in Africa. Even though Ethiopia is a landlocked country, it is commonly referred to as 'Water tower of Africa' because of the great water reserves it possesses. For instance, the Blue Nile river, which accounts for 80% of the streamflow of the Nile river by volume, has its source in Ethiopia (Champion and Manek, 2019).

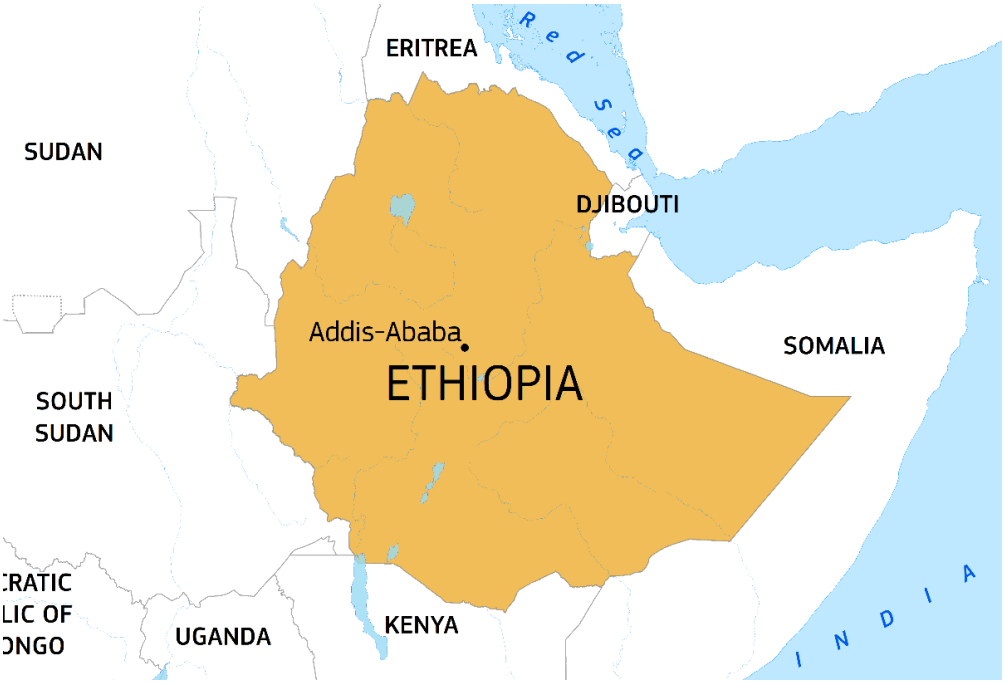

Figure 1. Map of the Horn of Africa (European Civil Protection and Humanitarian Aid Operations, 2019)

This control over the availability of water in downstream countries is a source of frequent discontent, as it constitutes a considerable leverage over neighboring countries most notably Egypt, which relies on the Nile river for the wellbeing of its agricultural sector (Champion and Manek, 2019). Still, Ethiopia itself is susceptible to frequent droughts especially in the eastern part of the country, and it harms its own agricultural sector, which

accounts for about 75% of the labor force and represents 40 % of its GDP, its biggest exports are coffee and oilseeds (Central Intelligence Agency, 2019).

Omo-Gibe river basin

Ethiopia is divided into 12 basins out of which 9 are river basins. In this work the focus is on the Omo-Gibe river basin (shown in figure 2), which is home to about 6.5 million people and has a size of just under 80000 km^2. Omo river represents the largest river of this basin, while Gibe river is the largest tributary to the Omo river, contributing about 90% of the total volume of the Omo river. (Legass, 2016)

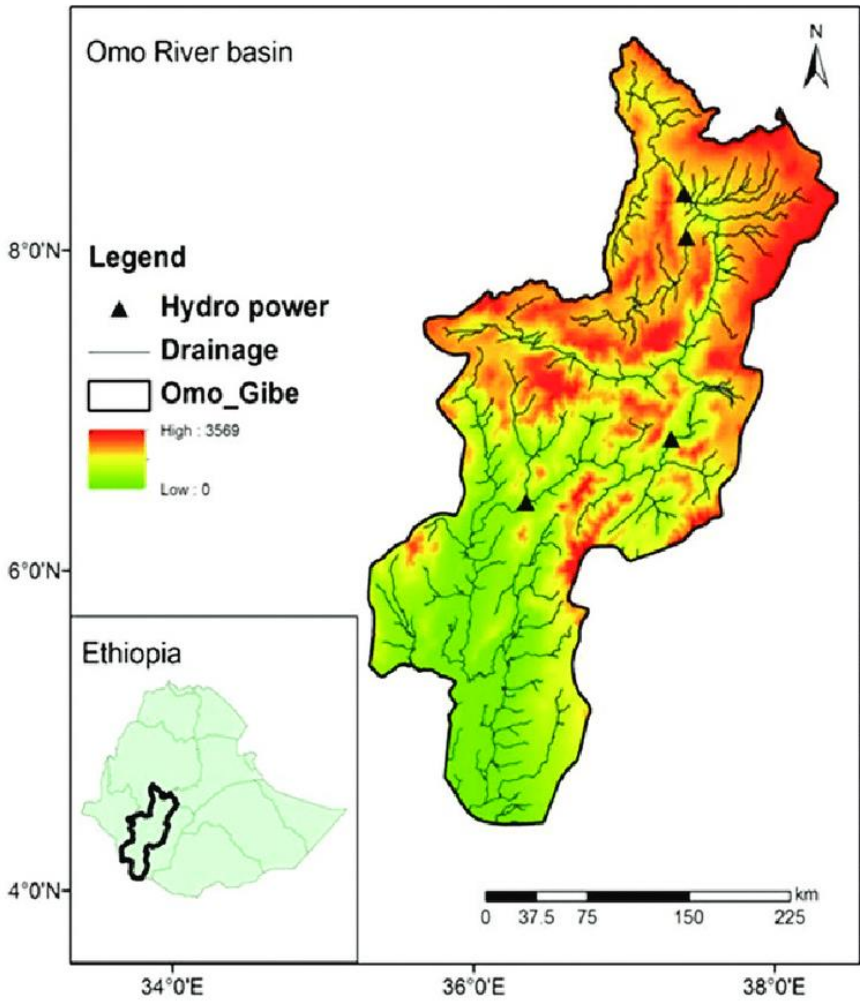

Figure 2. Topographical map of Omo-Gibe river basin *(Legass, 2016)*

2.2 Energy market of Ethiopia

2.2.1 Introduction

The energy market of Ethiopia is dominated by the Ethiopian Electric Power (EEP) which is a state-owned company that produces the electricity through the power plants it owns. While there is a private sector, the number of privately-owned power plants is very low compared to the public-owned ones in terms of generating capacity most likely because of the fact that the country has the third-highest public investment in the world, but the sixth-lowest private investment (World Bank, 2016). The Ethiopian Electric Utility (EEU) is given the task of distributing the electricity. The electrification rate stands at 25% with 85% of the urban population has access to electricity while only 5.3% of the rural population shares the same privilege. The low electricity access and substantial distribution losses coupled with the fact that the electricity demand sees about a 20% growth each year put a lot of additional stress on the national grid (African Development Bank, 2018).

2.2.2 Power generation capacity

Only 8.5% of Ethiopia's hydropower potential (45 GW) is utilized. This potential is a crude estimate given by the government, not a research-based one (Power Africa, 2016). This represents a huge untapped potential. The government is aware of it and, as the power generation stood at only 4180 MW in 2015/2016, follows an aggressive stance with regards to building more hydropower dams as exemplified in their latest *Growth and Transformation Plan* released in 2016 where the government set the power generation capacity target to 17,347 MW of which 13,817 MW comes from hydropower source (Federal Democratic Republic of Ethiopia, 2016).

Hydropower plants are scattered all around the country as can be seen from figure 2:

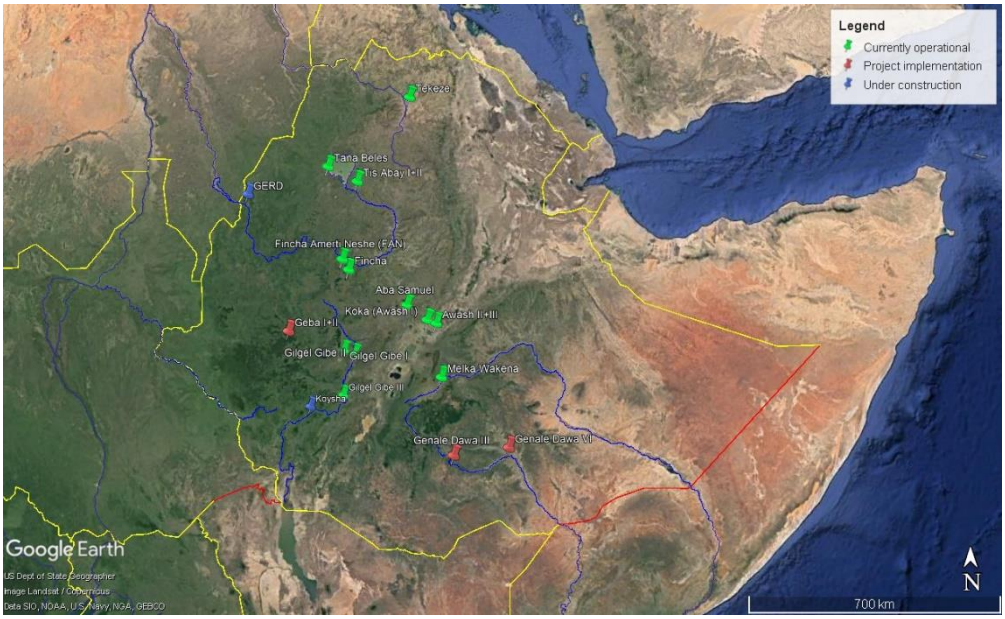

Figure 2. The location of current and future hydropower plants in Ethiopia

The chart below sums up the different hydropower plants in Ethiopia with the installed capacity and its current status. Varying plants in terms of installed capacity can be found from the small-scale Aba Samuel's 6.6 MWe to the ambitious and giant GERD's 6450 MWe (Ethiopian Electric Power, 2018).

Table 1. Current and future hydropower plants, their location and their nominal capacity

Latitude	*Longitude*	*Name*	*Installed capacity (MWe)*	*Status*
8.788°N	*38.706°E*	*Aba Samuel*	*6.6*	*operational*
8.468°N	*39.156°E*	*Koka (Awash I)*	*43*	*operational*
8.393°N	*39.352°E*	*Awash II+III*	*64*	*operational*
9.561°N	*37.413°E*	*Fincha*	*134*	*operational*
9.789°N	*37.269°E*	*Fincha Amerti Neshe (FAN)*	*95*	*operational*
7.831°N	*37.322°E*	*Gilgel Gibe I*	*184*	*operational*
7.757°N	*37.562°E*	*Gilgel Gibe II*	*420*	*operational*
6.844°N	*37.301°E*	*Gilgel Gibe III*	*1,870*	*operational*
7.225°N	*39.462°E*	*Melka Wakena*	*153*	*operational*
11.82°N	*36.92°E*	*Tana Beles*	*460*	*operational*
13.348°N	*38.742°E*	*Tekeze*	*300*	*operational*
11.486°N	*37.587°E*	*Tis Abay I+II*	*84.4*	*operational*
5.51°N	*39.718°E*	*Genale Dawa III*	*254*	*operational, but out of use for social reasons*
5.68°N	*40.93°E*	*Genale Dawa VI*	*257*	*project implementation*
8.211°N	*36.073°E*	*Geba I+II*	*385*	*project implementation*
11.214°N	*35.089°E*	*GERD*	*6,450*	*under construction, 65% complete (as of 4/2018)*
6.584°N	*36.565°E*	*Koysha*	*2,160*	*under construction*

As predicted, the lion's share of Ethiopia's energy production comes from hydropower, e.g. 86% in 2015. Figure 3 shows the total energy production and

the share of each energy source. Hydropower is and remains the predominant source of energy in Ethiopia (Ethiopian Electric Power, 2018).

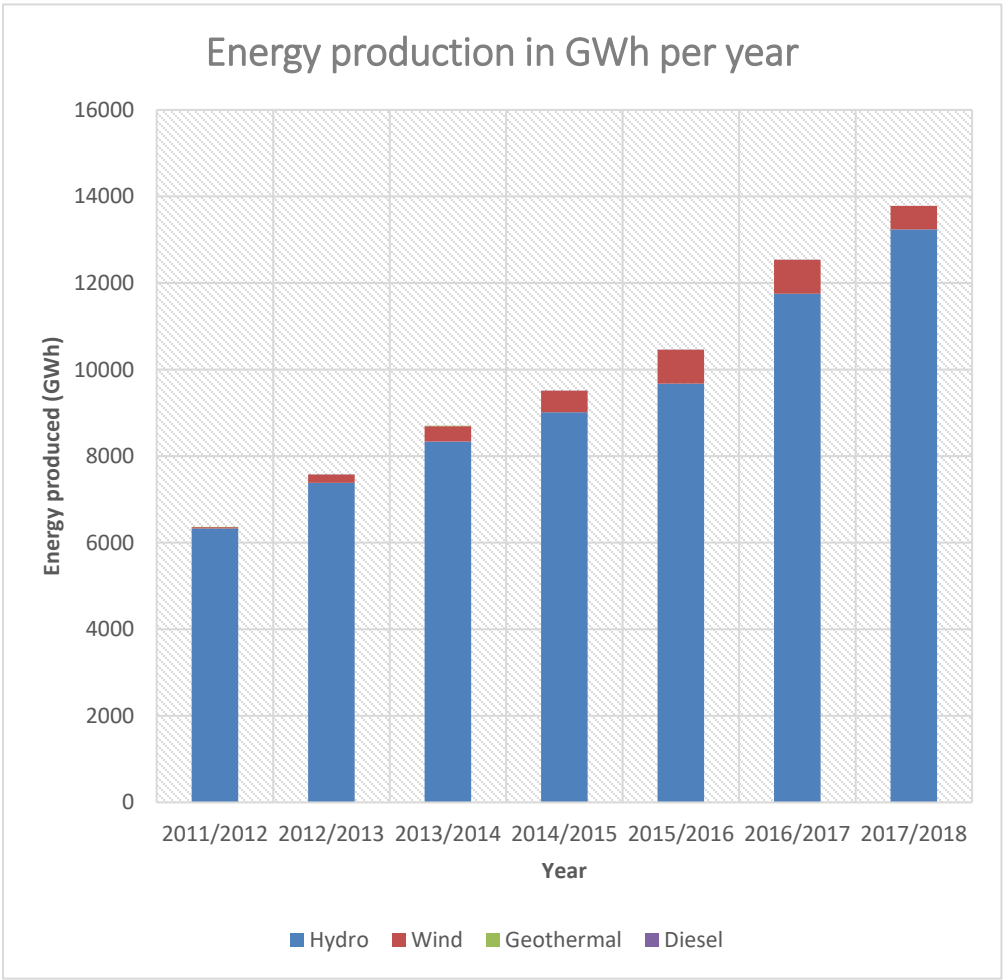

Figure 3. Energy produced in GWh per year in Ethiopia

Not all power plants are equal in terms of energy generated though as can be seen from figure 4 which shows the share of energy produced by hydro power plant in the period extending from September 2017 to August 2018 (Ethiopian Electric Power, 2018).

The share of the 3 Gilgel Gibe power plants, all of which are located in the Omo-Gibe river basin, stands at 58% of the total energy produced, adding Tana Beles increases this value to 80%.

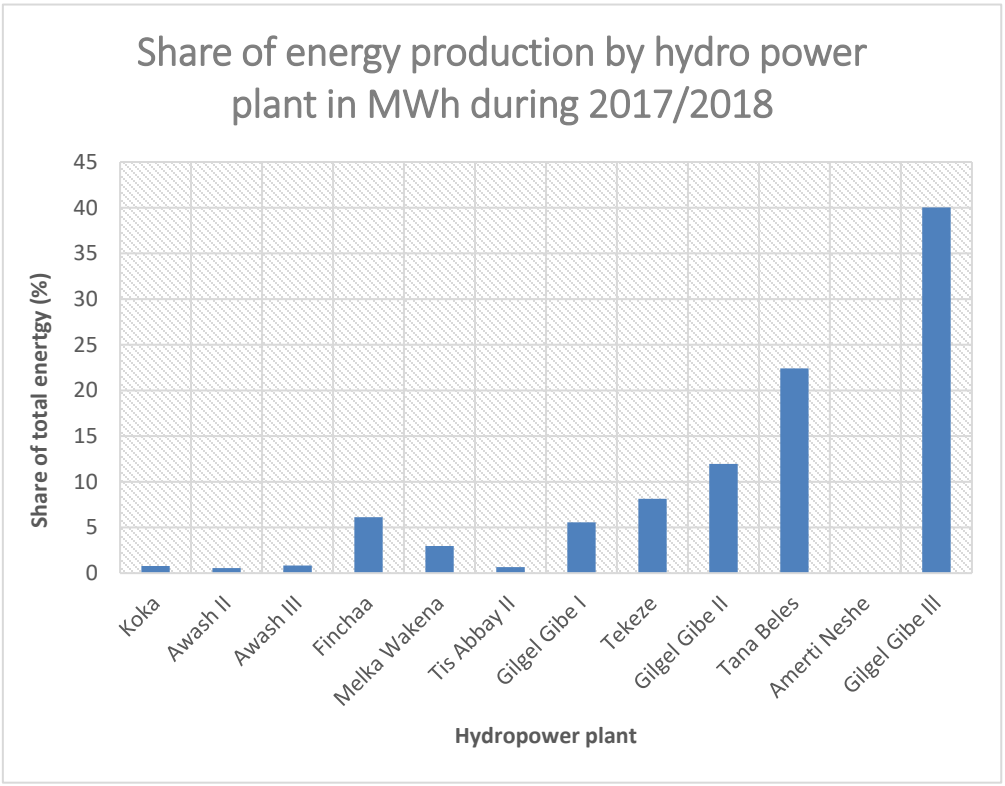

Figure 4. Share of energy production by hydropower plant in MWh during 2017/18

The East African Power Pool is a 10-country power pool whereby members of the EAPP pool their energy so any surplus production isn't wasted but is transferred to nearby countries who can make up for any energy shortage. Since energy is a vital resource, reliance on it keeps growing, and being a member of an energy pool remains the surest way to provide energy for domestic, commercial and industrial purposes. Ethiopia, and neighboring countries Sudan, Djibouti and Kenya are all members of the pool. A 2030 projection has Ethiopia as the country with the most power surplus of all countries of the pool, and as Ethiopia builds more power plants, the energy trade with the other members can only increase as can be seen in table 2 (Ethiopian Electric Power, 2018):

Table 2. Energy trade in GWh between Ethiopia and neighboring countries.

INTERCONNECTIONS	11/12	12/13	13/14	14/15	15/16	16/17	17/18
ETHIO – DIJBOUTI	331	386	267	379	271	492	911
ETHIO - KENYA	1	1	1	1	2	-	-
ETHIO - SUDAN	-	175	340	381	427	759	520
TOTAL (GWH)	**332**	**562**	**608**	**761**	**700**	**1,251**	**1,431**

Figure 5 shows the total energy transferred from Ethiopia to neighboring countries and EAPP members Sudan and Djibouti from September 2011 to August 2018 (Ethiopian Electric Power, 2018).

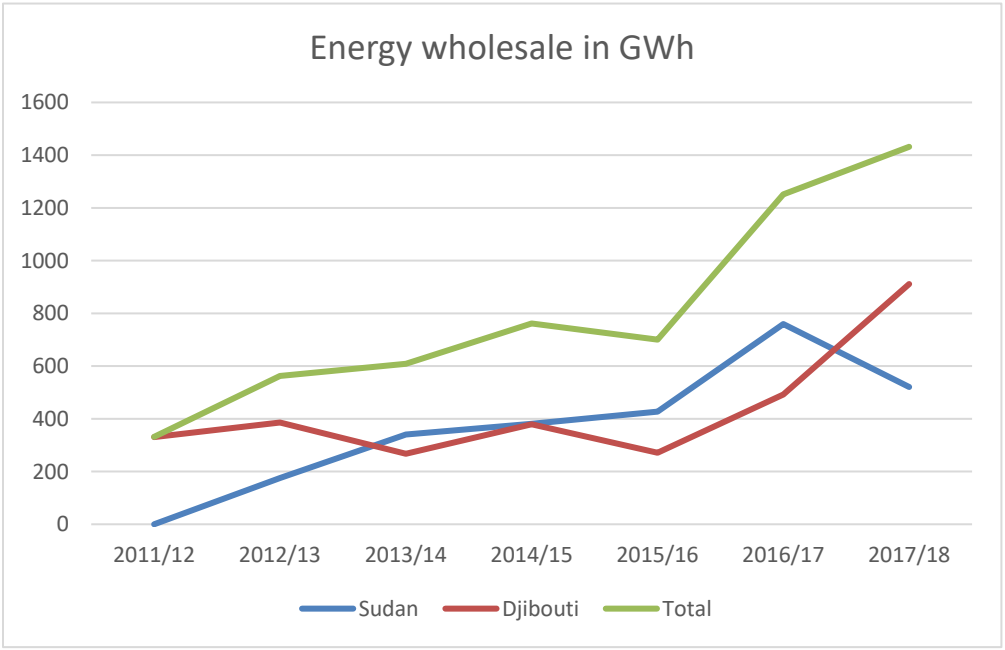

Figure 5. Energy wholesale from Ethiopia to Sudan and Djibouti in GWh

2.3 HBV model

The HBV model is a hydrological runoff model developed in 1972 by SMHI (Swedish Meteorological and Hydrological Institute) used in Scandinavia but also around the world. This model has seen applications in different fields like real time forecasting, groundwater simulations, design floods and water balance studies amongst others. (Bergström, S., 1992)

The model used to estimate the runoff in the Omo-Gibe river basin is based on the HBV model and resembles in a simplified manner the one pictured in figure 6. The model is composed of two distinct but interrelated boxes, the soilwater box and the runoff box. This model needs the precipitation and the temperature as inputs, while it outputs the computed runoff.

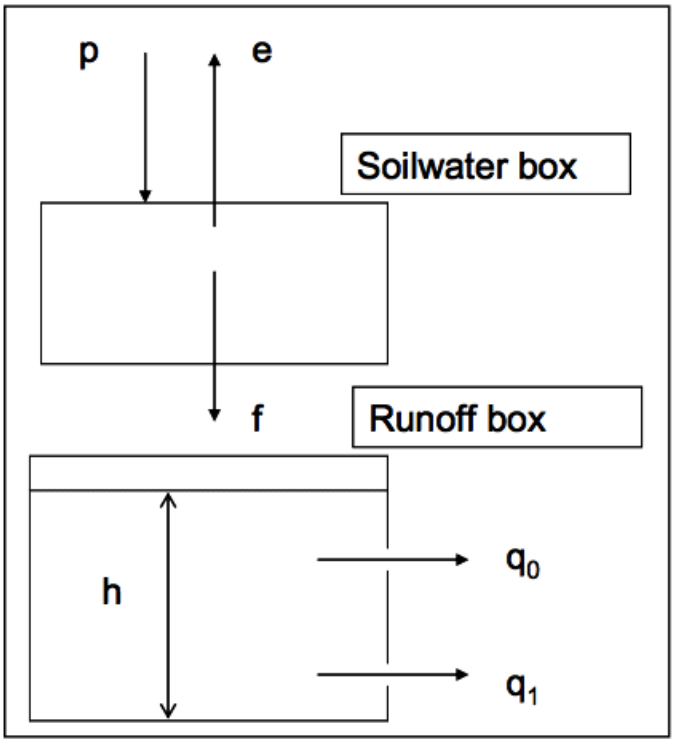

Figure 6. Example of an HBV model (Persson, 2018a)

Soilwater box

The soilwater box simulates the reaction of soils to precipitation. The water in the soilwater box is usually estimated using mathematical equations that are interrelated. The number of equations used increases with the complexity of the model but so does the number of input data needed, variables used, and time to compute the program. Below are simple equations which combined together are used to describe the effect of precipitation on soils.

$$\frac{dh_s}{dt} = p - e - f$$

The variation in water level h_s in the soilwater box is equal to the difference between the precipitation p, evapotranspiration e, and percolation to runoff box f. It should be noted that the precipitation p can be either in the form of rainfall or snowmelt.

While rainfall is typically a given data, snowmelt is usually computed, the equation below can be used to do just so:

$$s_m = \max(0, C_d * (T - Temp_{ref}))$$

In essence, this equation describes the fact that if the temperature T is above a reference temperature $Temp_{ref}$ (usually set to 0°C) then the snow is melting, and its value is equal to this difference in temperature times a model parameter C_d. So, the precipitation p mentioned above is really equal to the sum of the rainfall r and snowmelt s_m.

Usually, the evapotranspiration and percolation too are modelled with parameters that control the influence of temperature, the soil moisture content, the potential evapotranspiration and the field capacity on them.

The evapotranspiration is calculated using the following equation:

$$e = p_e * \min(1, \frac{h_s}{h_p})$$

From the equation, the maximum evapotranspiration e that can be reached is p_e which is the potential evapotranspiration and is a model parameter. Unless the water level in the soil h_s is above a threshold h_p (also a model parameter) which means that there is enough water in the soil and therefore the soil is saturated, the evapotranspiration will be less than the maximum value (pe). When the soil is unsaturated, the soil eats away from that potential evapotranspiration.

Percolation is calculated using the following equation:

$$f = p * (\frac{h_s}{FC})^b$$

From the equation, percolation happens when there is precipitation and there is water in the soilbox. Percolation increases as either precipitation or the

water level h_s increases, but also if FC is set to a lower value. This is, of course, a simplification of what happens in reality, but a good approximation, nonetheless. Both FC and b are model parameters.

Runoff box

The runoff box simulates the reaction of a reservoir with two outlets to an inflow of water (the percolation f), the bottom outlet has a slow and steady kind of flow q_1 (called slow runoff) and the top one has a fast response q_0 (called fast runoff).

The variation in the water level h is calculated using the continuity equation below:

$$\frac{dh_r}{dt} = f - q$$

q being the runoff, and f is the percolation described in the section above.

While the equation for the runoff is:

$$q = \frac{h_r}{T} + \frac{h_r - H_t}{T_{fast}}$$

T and T_{fast} are time constants, and H_t is the threshold level at which point the fast response is triggered.

The model parameters are summed up in the chart below (table 3):

Table 3. Model parameters and their respective boxes

BOX	PARAMETER
SOILWATER (SNOWMELT)	C_d, $Temp_{ref}$
SOILWATER (EVAPOTRANSPIRATION)	p_e, h_p
SOILWATER (PERCOLATION)	FC, b
RUNOFF	T, T_{fast}, H_t

Some of these parameters are limited in respect to the value they can carry, for example T is always going to be bigger than T_{fast}, while T_{ref} is usually going to hover around 0 °C. Others like C_d can be estimated if the type of soil is made known, for example C_d is equal to 2 mm/°C/day for forested areas and C_d is between 3-6 mm/°C/day for open fields. Of course, these values can be

exceeded or not even reached. When this happens, the model becomes more analytical than physically based (Persson, 2018b).

All these model parameters are determined during the calibration phase, the idea is that the model can simulate the runoff of a particular basin by using a specific set of model parameters that distinguish the reaction of this basin to weather conditions. After the calibration process, this set of parameters is saved for future uses to predict the runoff using the temperature and precipitation data.

3. Data collection and preparation

In this section, I describe both data collection and preparation tasks. The former refers to how the required data for the model was collected: precipitation and temperature data have been provided by Refinitiv, and an extensive search on the internet has been made to find data pertaining to streamflow rates and energy produced. The latter deals with the structuring of the data to match the format of the model's input.

3.1 Data collection

3.1.1 Precipitation

Both precipitation datasets have been provided by Refinitiv. The first dataset (referred to by CPCP) comes from the US federal agency Climate Prediction Center (CPC). The Climate Prediction Center both monitors and forecasts climate variables like temperature and precipitation (National Weather Service - Climate Prediction Center, 2019).

The second dataset comes from the Climate Forecast System Reanalysis (CFSR) model developed by the US National Centers for Environmental Prediction (NCEP), CFSR is a meteorological model that considers "**the interaction between the Earth's atmosphere, oceans, land and sea ice"** (National Centers for Environmental Prediction - Climate Forecast System, 2019).

None of the datasets provides observed precipitation values. While it is true that CPCP data is based on observed weather data, this data is subject to some statistical fiddling and in the case of CFSR data, it is based on a meteorological model altogether.

Both datasets had the KML file format, therefore the software Google Earth was used to open these files and display their respective content. Once opened, Google Earth shows a gridded Earth (figure 7) where cells or polygons are about 3000 km^2 in size, and are either blue, yellow or red colored. Blue polygons correspond to cells with a ratio of observed data to calculated data

above 0.5, yellow to cells with a ratio lower than 0.5, and red to cells where no observed data is available.

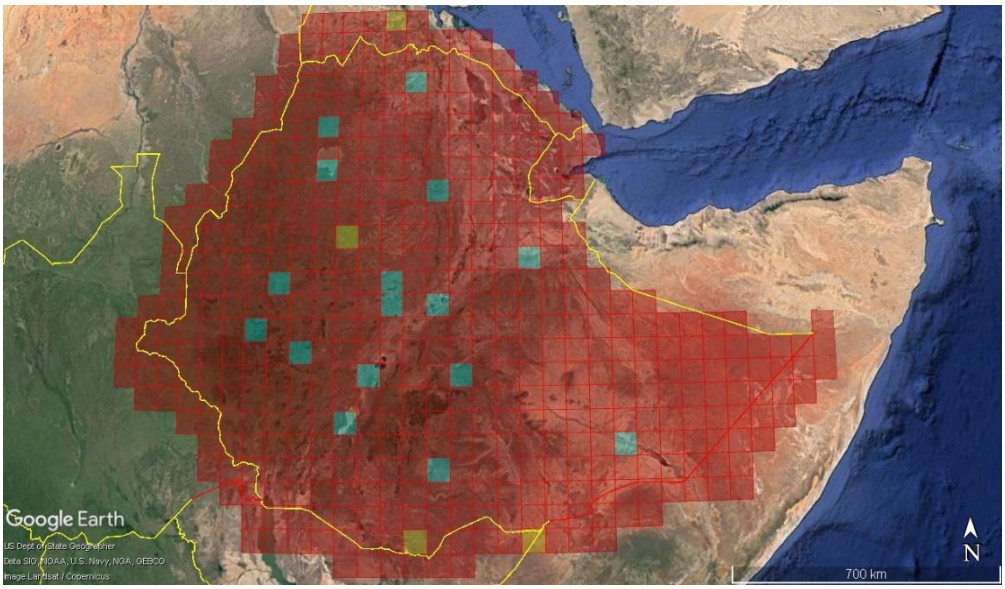

Figure 7. Different kinds of precipitation polygons available in Ethiopia (Blue polygons correspond to cells with a ratio of observed data to calculated data above 0.5, yellow to cells with a ratio lower than 0.5, and red to cells where no observed data is available)

3.1.2 Temperature

The temperature data, too, was provided by Refinitiv, in the case of temperature only one dataset was made available, and it is CFSR-based. The reason behind this is that temperature doesn't vary a lot from one day to another, the greatest daily variation and unpredictability is mostly seen in precipitation patterns. CFSR, being a fully-fledged meteorological model, is less susceptible to missing observed data - say if a particular weather station stops working - than CPCP is, the latter relying heavily on observed data.

Once opened in Google Earth, the resulting grid looks the same as the precipitation's, again with the blue colored cells referring to polygons with a ratio of observed data to calculated data above 0.5, yellow ones to those with a ratio below 0.5, and finally, the red color for polygons with no observation data (figure 8).

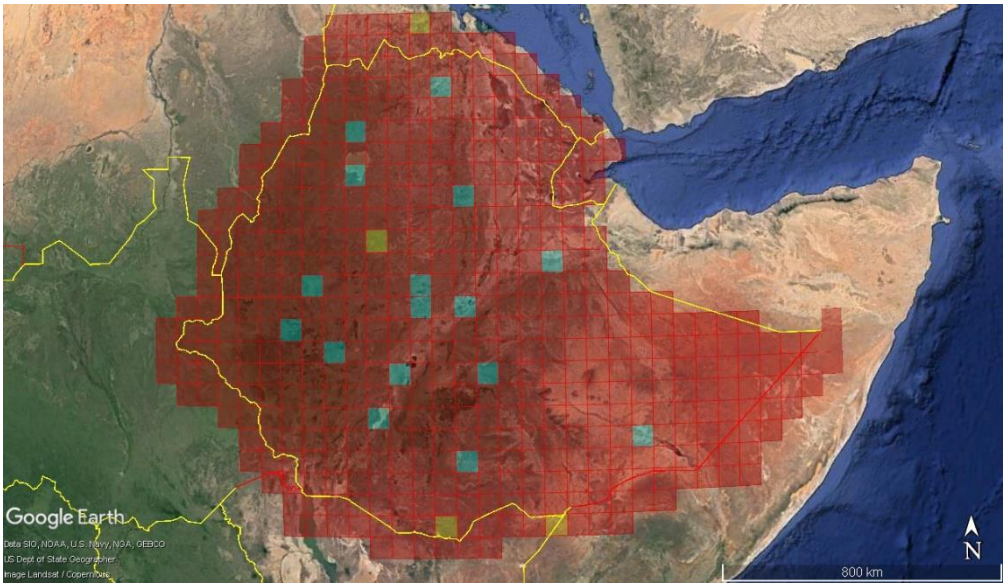

Figure 8. Different kinds of temperature polygons available in Ethiopia (Blue polygons correspond to cells with a ratio of observed data to calculated data above 0.5, yellow to cells with a ratio lower than 0.5, and red to cells where no observed data is available)

3.1.3 Streamflow

The source of the streamflow data (no pun intended!) was a tedious one, as this kind of data in the hydrological world is prized, it's only too natural to find difficulties in the process of searching for it.

Streamflow data can be found on the Global Runoff Database, a river discharge database ran by the Global Runoff Data Center (GRDC). The GRDC compiles streamflow data from different data providers around the globe. A catalogue of the stations can be found on their website in a KMZ file format. The stations can be seen in the figure below represented by a white pin, some major streams in Ethiopia are represented by a blue line (GRDC, 2018).

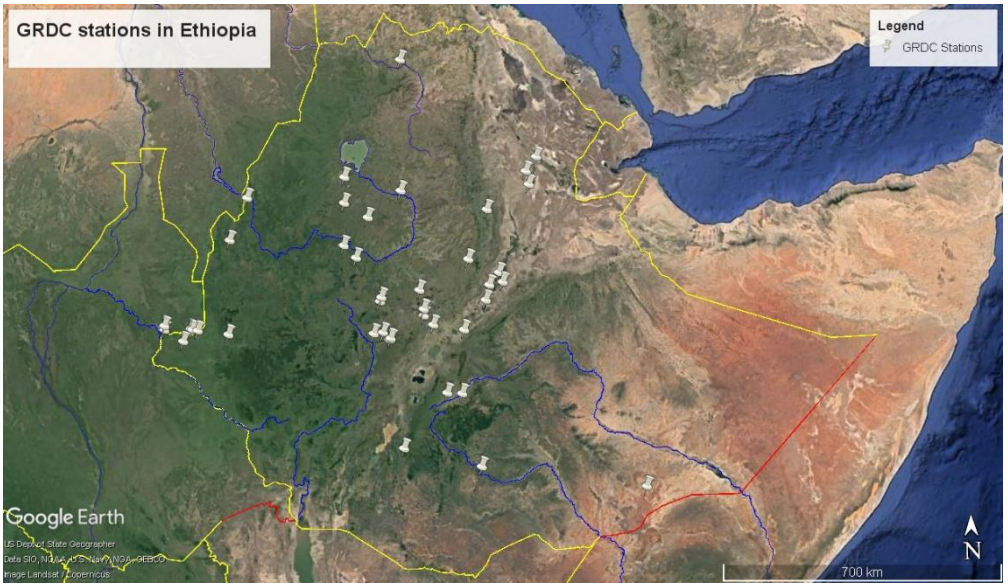

Figure 9. Location of GRDC stations in Ethiopia

A data request was made in order to get flow data for many rivers in Ethiopia. An order form was filled in, after specifying the GRDC No. of the station, the river name, the station name, the period, and the resolution (Monthly/Daily). All available daily data of Ethiopia for the 1978-2018 period, and all available monthly data of Sudan and Somalia for the 1978-2018 period was requested. The reason behind adding neighboring countries Sudan and Somalia is that there is no data for some streams in Ethiopia and this can be remediated by gathering discharge data downstream the rivers in neighboring countries. The requested data was received shortly after the request was made. As expected, data for most rivers was limited to monthly data, and large parts of stations with daily data missing.

Streamflow data across Ethiopia is recorded by Ethiopia's Ministry of Water, Irrigation and Energy (MOWIE). During the work period, multiple crashes of their website were observed. At first when the site was inaccessible, the "Wayback Machine", a digital archive of the Web, was used to access past versions of this website in order to find information related to hydrological data. Various contact emails were found on the website and were promptly contacted, to no avail probably because of the fact that it was an older version

and the emails were no longer valid. When the website was up again, new emails were fetched and contacted, but no response could be obtained.

Another probable source of hydrological information was the Ethiopian Institute of Water Resources and so it was contacted and, in this instance too, no answer was received.

One final solution was to scour the Internet for this kind of data from secondary sources. This extensive search yielded some findings at long last: a 2011 Master thesis report done by Daniel Asefa in Addis Ababa University under the tile “Water Use and Operation Analysis of Water Resource Systems in Omo Gibe River Basin” contained flowrates for the gauging stations shown in the figure below. The values are monthly flowrates from 1985 to 2006 but were limited to those stations present in the Omo Gibe river basin. Missing data in some stations has already been filled using an interpolation of flowrates from nearby stations (Asefa, 2011).

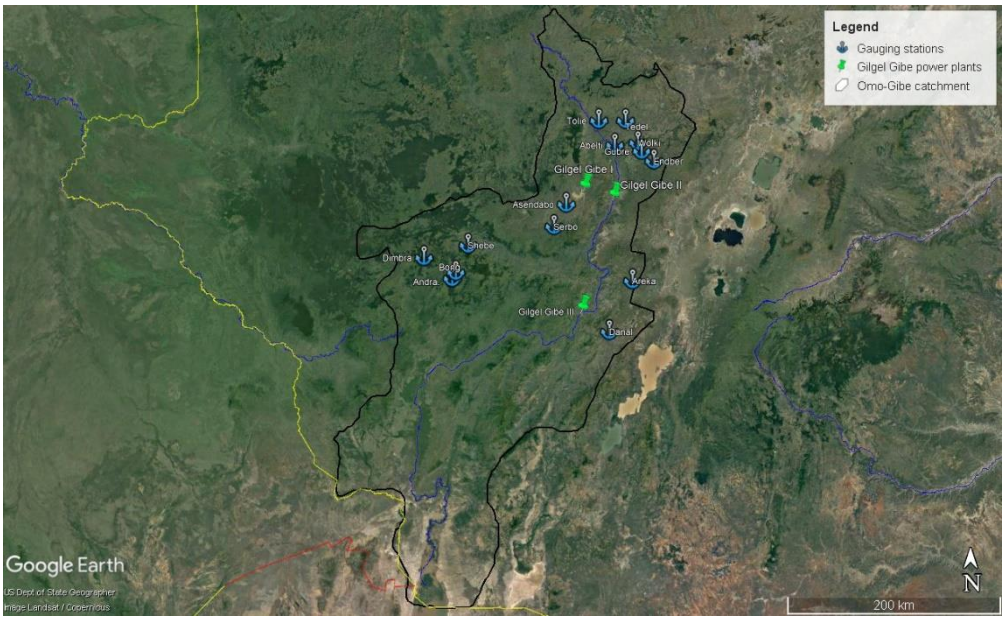

Figure 10. Location of streamflow stations in Omo-Gibe river basin

3.1.4 Energy

The currently installed hydropower park in Ethiopia was found on the Ethiopian Electric Power (EEP) website. Planned plants and the ones under

construction were found on other websites. These three types of plants are all shown in figure 11 (Ethiopian Electric Power, 2018) (IEA, 2019).

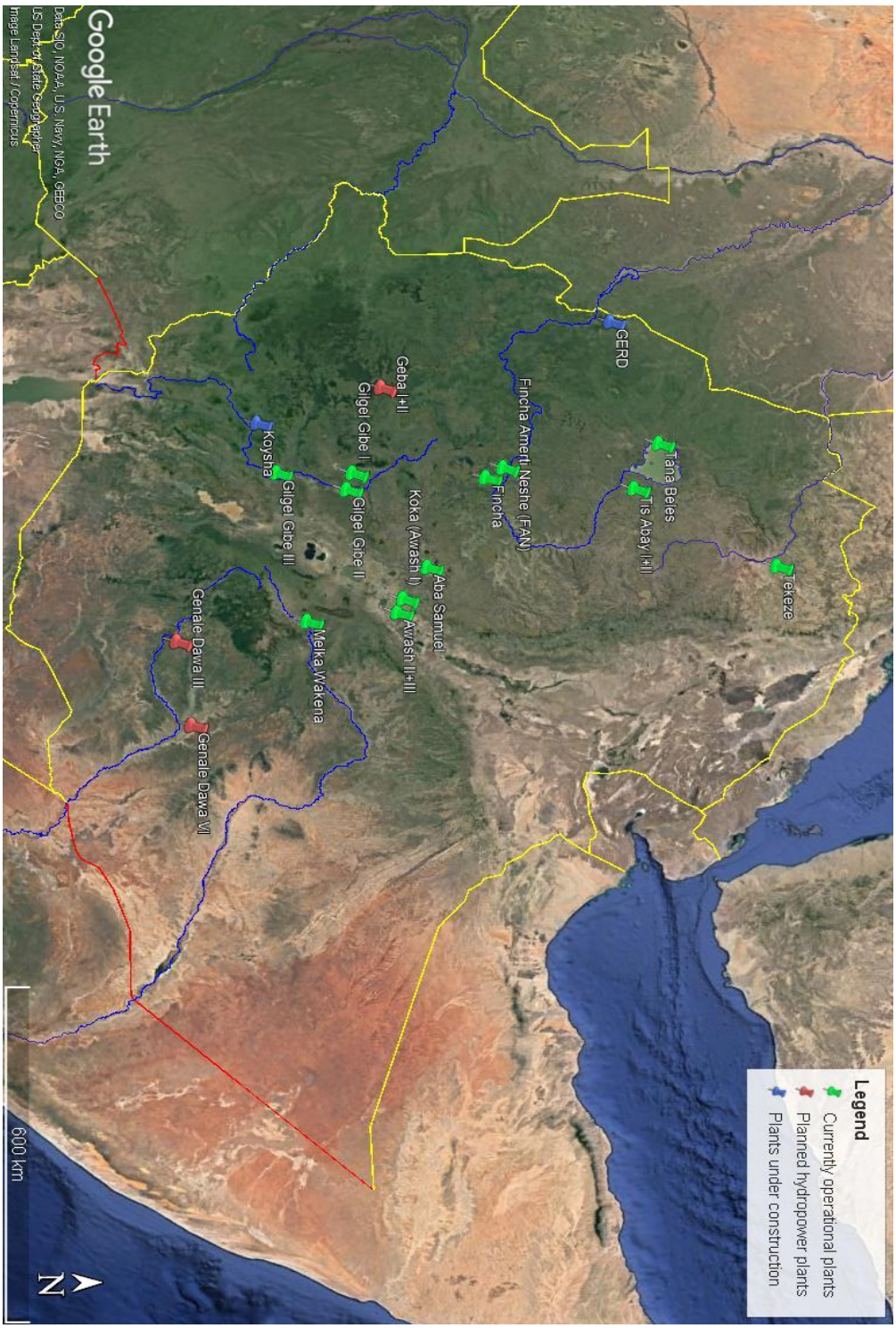

Figure 11. Current and prospective hydropower plants in Ethiopia

Data for the energy produced comes from two different sources: the Ethiopian Electric Power (EEP) covers the energy production for the period from 2011/12 to 2017/18, and the International Energy Agency (IEA) for the period stretching from 1990 to 2016. Values from both sources for the same year were compared and were found to correspond roughly to each other. Unfortunately, this data corresponds to the total energy produced by the hydropower park, except for the year 2017/18 where data for individual plants can be found.

3.2 Data preparation

3.2.1 Precipitation and Temperature

Both meteorological datasets were already correctly formatted, and thus needed no conversion.

3.2.2 Streamflow

Since streamflow rates were given on a monthly m^3/s they were converted to mm daily. In order to achieve this, the original flowrates were converted by accounting for the size of each catchment and using the monthly flow as a constant one.

The different catchments that are part of the Omo-Gibe river basin were found in the study below, then one screenshot of the different catchments was overlayered on top of the Omo-Gibe river basin layer in Google Earth so that as a result the size of the catchments could be determined (See Appendix 2).

For the Asendabo catchment, the source of the image was found in "*The Application of Predictive Modelling for Determining Bio-Environmental Factors Affecting the Distribution of Blackflies (Diptera: Simuliidae) in the Gilgel Gibe Watershed in Southwest Ethiopia*" (Ambelu et al., 2014)

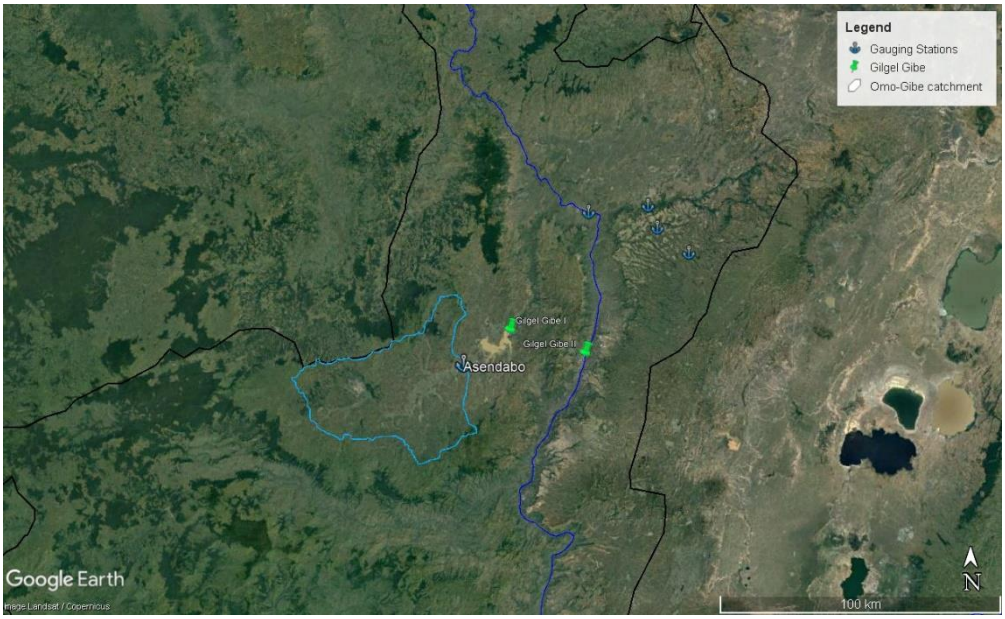

Figure 12. Resulting Asendabo river basin area

For the Abelti catchment, the following source was used: "*Proceeding of The Second National Consultative Workshop on Integrated Watershed Management on Omo-Gibe Basin" (Legass, 2016)*

Figure 13. Resulting Abelti river basin area

Since the data is too big to be formatted manually, a simple MATLAB routine was used to format the data the right way (See Appendix 3 for the actual values).

3.2.3 Construction of the Q-target series

Once runoff data is available, the construction of the Q-target series (also referred to by observed runoff) can start. First, streamflow stations are selected based on a number of criteria: the quality of their respective data (the percentage of missing data), their location in the catchment and their size. Once the number of stations is reduced, a Q-target series can be produced usually by giving a weight to each station depending on their size and quality of their data.

Selection of streamflow stations

Not all the stations shown in figure 10 are useful, some of the stations aren't usable at all since too much data is missing, others are downstream the hydropower plants or too small to effect anything meaningful on the energy produced.

At the end, two gauging stations can be used. The first one "Asendabo" is located just upstream Gilgel Gibe I, and since it's the only river that feeds the Gilgel Gibe I reservoir, it seems appropriate to select this station.

The second station chosen is "Abelti". Even though its stream is affected by the operations done upstream at the level of Gilgel Gibe I, it remains the most important in the whole catchment since the river on which this station is located accounts for the majority of the river flow of the entire basin.

Construction of the Q-target series

The river flows coming from Asendabo and Abelti stations respectively are combined just upstream of the hydropower plant Gilgel Gibe II, this means that both stations have roughly the same impact on downstream hydropower plants. Therefore, the Q-target series is built by simply averaging the observed runoff of the two streamflow stations Asendabo and Abelti.

During the coming calibration and validation processes of the model, the resulting computed runoff is compared with the constructed Q-target series (observed runoff). The smaller the difference between the computed runoff

and the observed runoff the higher the performance of the model. In essence, the Q-target serves the purpose of being the benchmark against which the results of the model are compared.

4. Results

4.1 Temperature and precipitation polygons

First, the temperature and precipitation polygons were chosen. The first criteria for the first batch of the weather polygons to be used was proximity to the streamflow stations and only the polygons located within the catchment could be chosen. The second criteria is the quality of weather data, ranked in order of higher quality, the green-colored polygons come first then the yellow-colored ones, the red ones being of the lowest quality were chosen as a last resort.

The next step is to test each weather polygon alone, this was done while trying a reasonable value for the model parameters. Then, the 3 best performing weather polygons were selected. It was decided to choose 3 polygons in case there is any problem with one polygon there is backup from the other two polygons so that the model would still work correctly albeit in a less-than-optimal fashion.

Finally, the weight to give to each polygon was decided while carrying the calibration phase, which means that changing the values of the model parameters and the weight of the different polygons were carried out simultaneously.

4.2 Calibration and validation of the model

After choosing the temperature and precipitation polygons, the model is ran first for the calibration phase which stretches from 1985 to 1999, and then the validation phase from 1985 to 2006. During the calibration phase, the model parameters are set to different values and then modified many times until different sets of model parameters are tested. The calibration process is stopped when there is no substantial change in the way observed runoff (constructed Q-target series) and computed runoff match even while tweaking the model parameters further. After the calibration is finished, the model is validated using a longer period of time.

The r^2 value is a measure of how well the computed runoff matches the observed runoff, and it is calculated using the equation below:

$$r^2 = \frac{\sum(\overline{Q_{obs}} - Q_{obs})^2 - \sum(Q_c - Q_{obs})^2}{\sum(\overline{Q_{obs}} - Q_{obs})^2}$$

The maximum achievable r^2 value is 1 and can be achieved if Q_c is equal to Q_{obs}, on the other hand there is no minimum value as it can have a negative value in case the model is poorly calibrated. If the r^2 value is equal to 0, then this means that the first term of the denominator is equal to the second, which in turn means that the model performs just as well as if the mean observed runoff was used instead of calculated runoff.

4.3 Calibration phase

Figure 14 below shows the observed runoff (Q_{obs}) in red and the computed runoff Q in blue using the CPCP dataset. The r^2 achieved in this case is 0.763 which represents a good value, as can be seen from the graph the computer runoff Q and the Q_{obs} match perfectly for 7 years and besides 1987, 1992 and 1997 for which Q and Q_{obs} don't correspond well, the other remaining years see a good match.

The figure below shows also the accumulated difference which starts at 0 in the year 1985 and stands at 266.1 at the end of the calibration phase by the end of the year 1999.

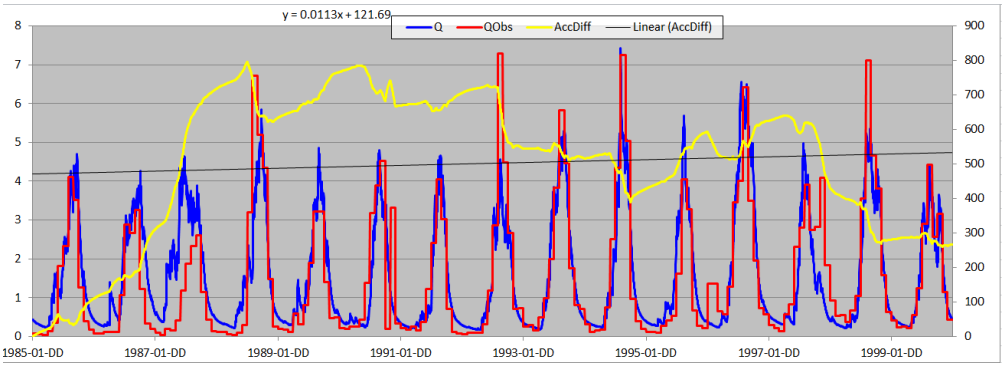

Figure 14. Calculated and observed runoff for the CPCP model during the calibration phase.

On the other hand, the figure below (figure 15) uses the CFSR dataset, here too Q_{obs} and Q_{calc} match quite closely with an r^2 value of 0.8. Even though the r^2 value achieved here is higher than it is the case when using the CPCP dataset, the peaks of Q and Q_{obs} here don't match quite well (9 instances) against 5 instances for the CPCP model. On the other hand, the phase when the observed runoff gets low in the winter every year fits tighter in the CFSR model than in the CPCP model. The CFSR model also achieves a lower accumulated difference in absolute terms than CPCP does, 238.4 mm against 266.1 mm.

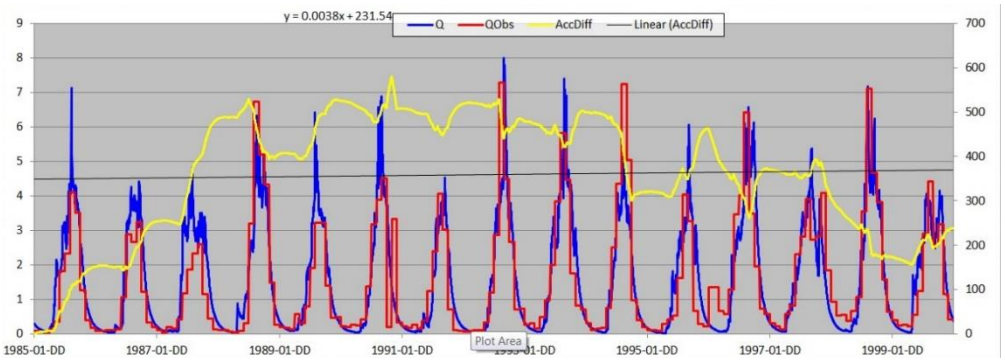

Figure 15. Calculated and observed runoff for the CFSR model during the calibration phase.

4.4 Validation phase

Once the calibration phase is finished, the next step is to validate the model, which means the model is run for a longer period and its strength and accuracy assessed. In this step there is no fiddling of the model parameters, they remain unchanged during the validation phase. The validation period starts in 1985 and extends to 2006.

The figure below (figure 16) shows the results of the CPCP model after validation. The r^2 value drops by 24 % to 0.579 down from 0.763, as in all the new years the model performed only correctly for one year (2006). And this can be seen also in the accumulated difference value, as it increases to 2236.4 mm since 6 out of the last 7 years are poorly simulated.

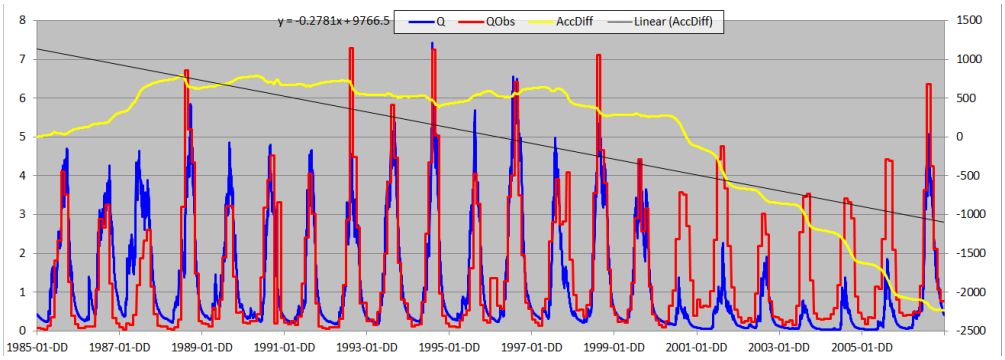

Figure 16. Calculated and observed runoff for the CPCP model during the validation phase.

On the other hand, the CFSR model performs really well, and sees a very minor decrease in the r^2 value down from 0.8 to 0.797. Two years out of the last six are well simulated, which is quite different from the results using the CPCP model.

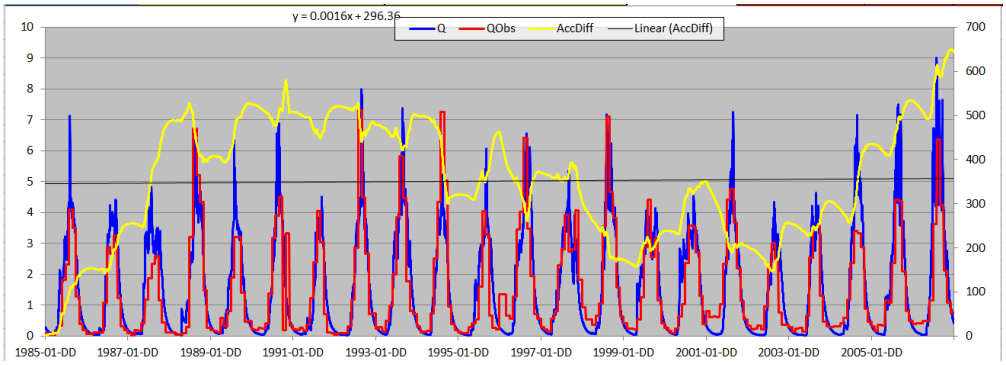

Figure 17. Calculated and observed runoff for the CFSR model during the validation phase.

Both resulting R-squared and accumulated difference values are summarized in the chart below for both datasets CPCP and CFSR.

Table 4. Resulting R^2 and accumulated difference from both CPCP and CFSR datasets

Dataset	CPCP	CFSR
Calibration	$R^2 = 0.763$ AccDiff = 266.1	$R^2 = 0.800$ AccDiff = 238.4
Validation	$R^2 = 0.579$ AccDiff = -2236.4	$R^2 = 0.797$ AccDiff = 644.9

Since both datasets have different precipitation values, it's normal that the resulting r^2, AccDiff and computed runoff are all different as can be seen in table 4 above and figure 18 below which shows the computed runoff computed from 1985 to 2018 using both datasets (the top graph shows the

calculated runoff using CFSR data and the bottom one using CPCP data) (for the actual values, please see Appendix 4).

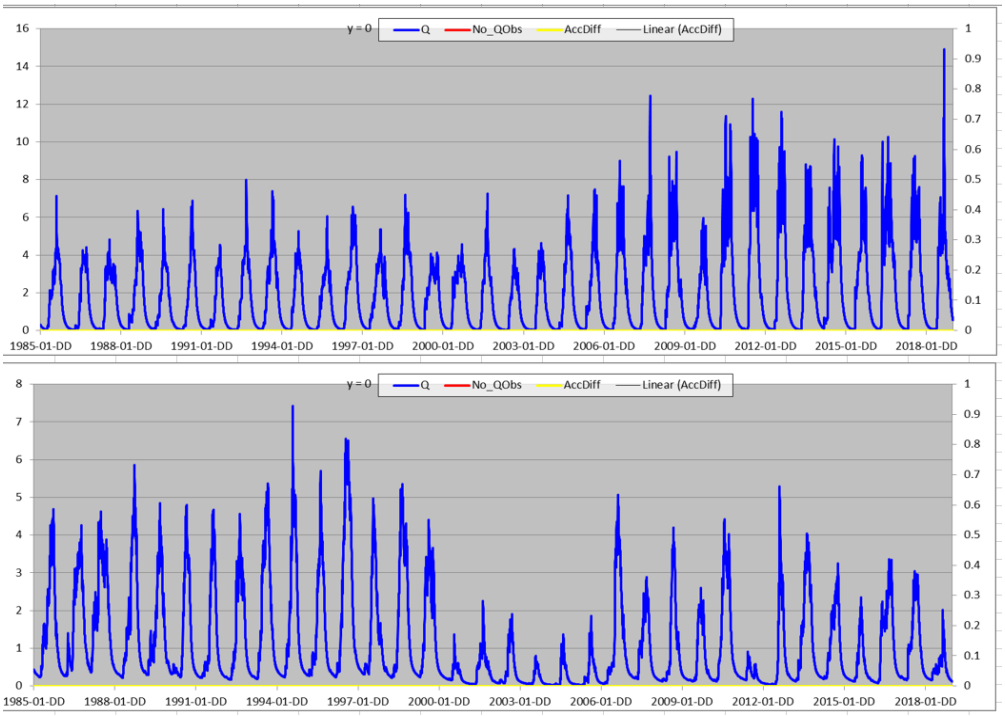

Figure 18. Computed runoff using CFSR (top) and CPCP (bottom) datasets

The energy produced by the 3 hydropower plants in the Omo-Gibe river basin in 2017 amounts to 7450 GWh, while the computed runoff using CFSR for the same year totals 846 mm, which gives a factor of 8.8 GWh/(mm/year). After updating CFSR data and adding it to the model, it's easy to estimate the energy produced in the future using this factor.

4.5 Analysis and discussion

At the end of the calibration process, both CPCP and CFSR resulted in similar R^2 value, 0.763 and 0.8 respectively, however after validation CPCP sees a large decrease in the R^2 value down from 0.763 to 0.579, while CFSR sees a negligible decrease in the R^2 value.

Similar to the change in the R^2 value, the accumulated difference (difference between the computed runoff and observed runoff) observed a large variation (in absolute terms) while using CPCP data, a variation of about 2500 mm as opposed to CFSR with 420 mm.

For the CPCP model, the accumulated difference sees a sharp raise during 1987 which is the worst modelled year of all 15 years. For the remainder of the calibration period besides 1997, the accumulated difference remains fairly stable. In essence, the accumulated difference observes an abrupt change for the badly modelled years and stays stable for well-modelled years. As for the CFSR model, during the third year the accumulated difference sees a sharp increase, so this year (1987) is actually badly modelled for both datasets.

After the validation period, the accumulated difference actually breaks even around the middle of 2000, but from then and until the end of 2006 the model manages to pile up to 2236 mm worth of water volume, this is due to the fact that the CPCP model constantly underestimates the runoff in the last 7 years of the validation period.

On the other hand, the accumulated difference for the CFSR dataset increases by about 400 mm, in stark contrast to the 2500 mm increase for the CPCP model.

Of course, as is usually the case, the r^2 value and the accumulated difference see a decrease during the validation process and was actually expected, this means that a small variation in the two aforementioned values signifies that the model is strong enough for future uses.

A positive note for the CPCP model is that it does account for the drought period in the early- to mid- 2000s while CFSR does it but only to a small extent. In all actuality, CFSR predicts usually overestimates the runoff while CPCP underestimates it.

Discussion

First, the HBV model is like its name implies a model with inherent limitations, and one of two possibilities present themselves to the user. The first alternative is to work with a simple model and therefore this model won't take into account the intricate interaction between the various physical phenomena at play. The second choice is a complex model that considers this interplay and thus needs more input data. The issues with this kind of model that usually arise are twofold: first, this data is usually too specific to be

available in all parts of the world and second even if such data is found, it might not lead to better results than if the simpler model was used. Even after modelling, the results need a thorough examination since a model is not reality itself, and therefore analyzing data is a great way to enhance the understanding of how the model works and how it could be improved upon.

Only two streamflow stations are used and while it is true that having fewer stations makes the task easier and simpler, this is not optimal in case one station is out of function the user is left with only one station. On the other hand, data for some tributaries might not be available at all and consequently relying on two stations might be the only choice available. Streamflow data for one station, namely Asendabo, was filled in for missing data and so is subject to error, it is much better to have the original intact streamflow data for this station, but it was not possible to get hold of such data.

As climate gets disrupted more frequently to extreme fashion, relying on CFSR might not be the best choice in the long term unless an update or revision to it is implemented and includes the effect of climate change on the weather, otherwise CPCP seems a better choice if the disruption caused by climate change is continuing.

Lastly, the energy generated by the power plants on an individual basis is not available except for the year 2017/18, it is recommended to have such data for other years, and look at the relationship between the energy generated throughout an extended period of time and the computed runoff to have a more accurate factor to predict the energy generated using weather forecast.

5. Conclusion

The objective of this degree project was to build and use the HBV model to forecast hydroelectricity, and as such the results shown by this project are promising. These results depended on two distinct parts: the construction of the Q-target series and the calibration and validation of the model.

The construction of the Q-target series proper is heavily dependent on the quality and quantity of runoff data. An extensive search for this data has been performed, after which runoff data was found for several stations. However, only two stations qualified for the criteria applied to the data as data for most stations was missing and was found to be filled in using unknown interpolation equations. As such, the construction of the Q-target series, and by the same token the results of this project, were inherently limited by the available data or lack thereof.

Nevertheless, the second part of this project, namely the calibration and validation of the model, is quite promising. The CFSR model was found to perform very well since during the calibration phase an r^2 value of 0.8 is considered a good value when building such a model. Furthermore, this value saw only a minor decrease during the validation process, which is a good indicator of the strength of the model.

Some suggestions that could be considered for further work in this project are obtaining data for the gauging stations from the Ministry of Water Irrigation and Energy, which is the operator and owner of streamflow stations in Ethiopia, this will allow access to raw data. The first benefit is the fact that this data comes from a primary source, and therefore reduces the risk of using mishandled data from a secondary source. The second benefit pertains to the fact that when missing data is found for a particular station, the author can decide which interpolation equations can be used in order to fill in the missing data. The construction of the Q-target series is ultimately greatly improved by having access to more better-quality data, since more stations located in the catchment can be used, and different combinations of them can be tested. Finally, getting access to the energy produced on an individual basis for an extended period of time can significantly enhances the accuracy of the energy production forecast.

References

African Development Bank. 2018. “Africa Energy Marketplace – Ethiopian Government Presentation.”

Ambelu, Argaw, Seblework Mekonen, Magaly Koch, Taffere Addis, Pieter Boets, Gert Everaert, and Peter Goethals. 2014. “The Application of Predictive Modelling for Determining Bio-Environmental Factors Affecting the Distribution of Blackflies (Diptera: Simuliidae) in the Gilgel Gibe Watershed in Southwest Ethiopia.” Edited by Arash Rashed. *PLoS ONE* 9 (11): e112221. https://doi.org/10.1371/journal.pone.0112221.

Asefa, Daniel. 2011. “Water Use and Operation Analysis of Water Resource Systems in Omo Gibe River Basin.” MSc Thesis, Addis Ababa University.

Bergström, Sten. 1992. “The HBV Model - Its Structure and Applications.” Edited by SMHI. SMHI Reports Hydrology.

Central Intelligence Agency. 2019. “Africa :: Ethiopia — The World Factbook - Central Intelligence Agency.” Cia.Gov. 2019. https://www.cia.gov/library/publications/the-world-factbook/geos/et.html.

Champion, Marc, and Nizar Manek. 2019. “Death on the Nile Haunts Ethiopia’s Rebirth.” Bloomberg.Com. 2019. https://www.bloomberg.com/graphics/2019-nile-river-ethiopia-dam/.

Ethiopian Electric Power. 2018. “Facts and Figures.” Eep.Gov. 2018. http://www.eep.gov.et/index.php?option=com_content&view=article&id=47&Itemid=264&lang=en.

European Civil Protection and Humanitarian Aid Operations. 2019. Ethiopia. European Civil Protection and Humanitarian Aid Operations. https://ec.europa.eu/echo/where/africa/ethiopia_en.

Federal Democratic Republic of Ethiopia. 2016. “Growth and Transformation Plan II (GTP II) (2015/16-2019/20) Volume I: Main Text.”

GRDC. 2018. “The GRDC - the World-Wide Repository of River Discharge Data and Associated Metadata.” Bafg.De. 2018. https://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html.

IEA. 2019. “Statistics - Global Energy Data at Your Fingertips.” Iea.Org. 2019. https://www.iea.org/statistics/?country=ETHIOPIA&year=2016&category=Renewables&indicator=HydroGen&mode=chart&dataTable=ELECTRICITYANDHEAT.

Legass, Ayalew Talema. 2016. “Proceeding of The Second National Consultative Workshop on Integrated Watershed Management of Omo-Gibe Basin.” In National Conference on Integrated Watershed Management, At Jimma, Ethiopia, Volume: 2. https://www.researchgate.net/publication/313428353_Proceeding_of_The_Second_National_Consultative_Workshop_on_Integrated_Watershed_Management_of_Omo-Gibe_Basin.

National Centers for Environmental Prediction - Climate Forecast System. 2019. “Climate Forecast System.” Noaa.Gov. November 4, 2019. https://cfs.ncep.noaa.gov/.

National Weather Service - Climate Prediction Center. 2019. “Climate Prediction Center.” Noaa.Gov. November 4, 2019. https://www.cpc.ncep.noaa.gov/.

Persson, Magnus. 2018a. “Assignment 2 - Rainfall Runoff Modelling.” Lund University.

Persson, Magnus. 2018b. “Snow Hydrology.” Lund University.

World Bank. 2016. “World Bank Priorities for Ending Extreme Poverty and Promoting Shared Prosperity.” Edited by World Bank.

Appendix 1

Table 5. Energy generated and installed capacity from 1990 to 2017

Year	Energy generated (GWh)	Installed capacity (MW)
1990	1062	412
1991	1082	412
1992	1151	412
1993	1263	412
1994	1354	412
1995	1428	412
1996	1510	412
1997	1566	412
1998	1579	412
1999	1605	412
2000	1646	412
2001	1992	485
2002	2023	485
2003	2280	485
2004	2521	669
2005	2833	669
2006	3259	669
2007	3385	669
2008	3296	669
2009	3524	669
2010	4931	1429
2011	6262	1944
2012	7388	1944
2013	8338	1944
2014	9013	1944
2015	9674	1944
2016	10406	3814
2017	13238	3814

Appendix 2

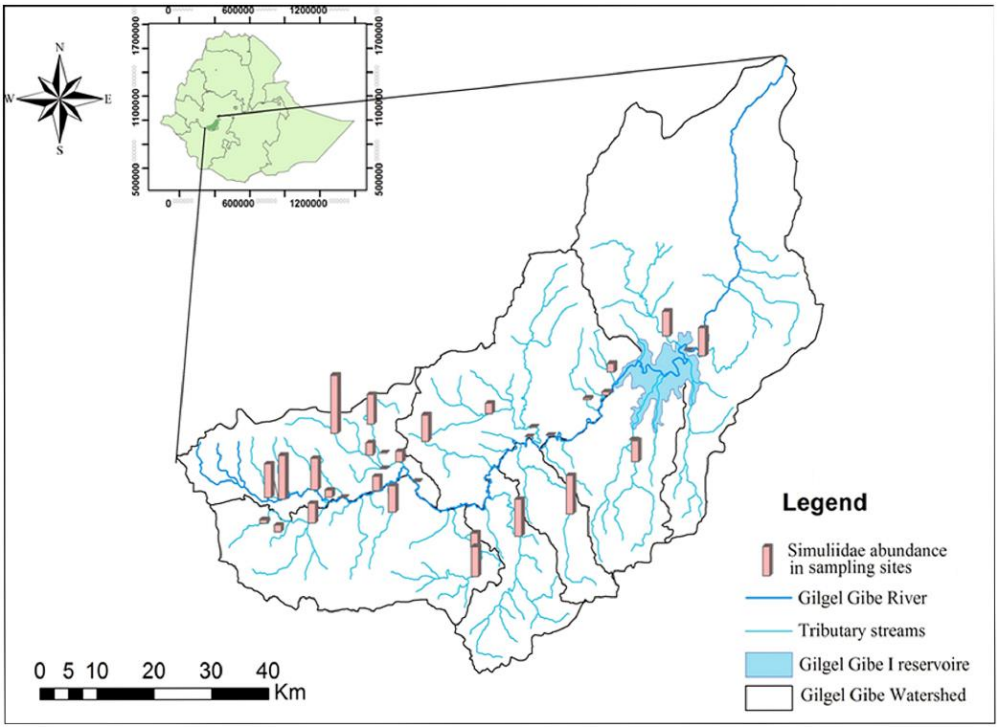

Figure 19. Asendabo river basin (Ambelu et al., 2014)

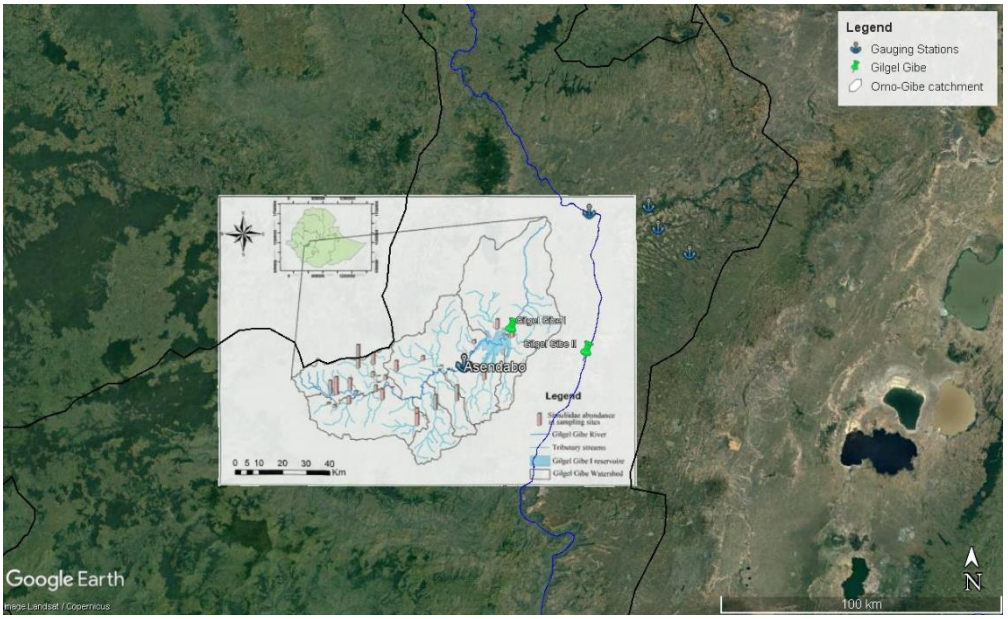

Figure 20. Overlay of Asendabo river basin on Omo-Gibe river basin map

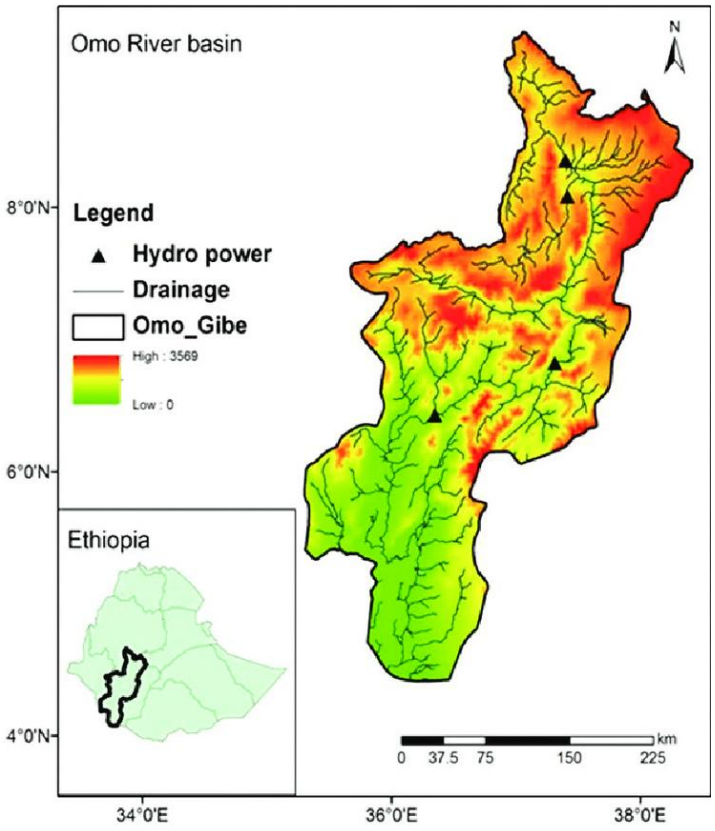

Figure 21. Omo river basin elevation map (Legass, 2016)

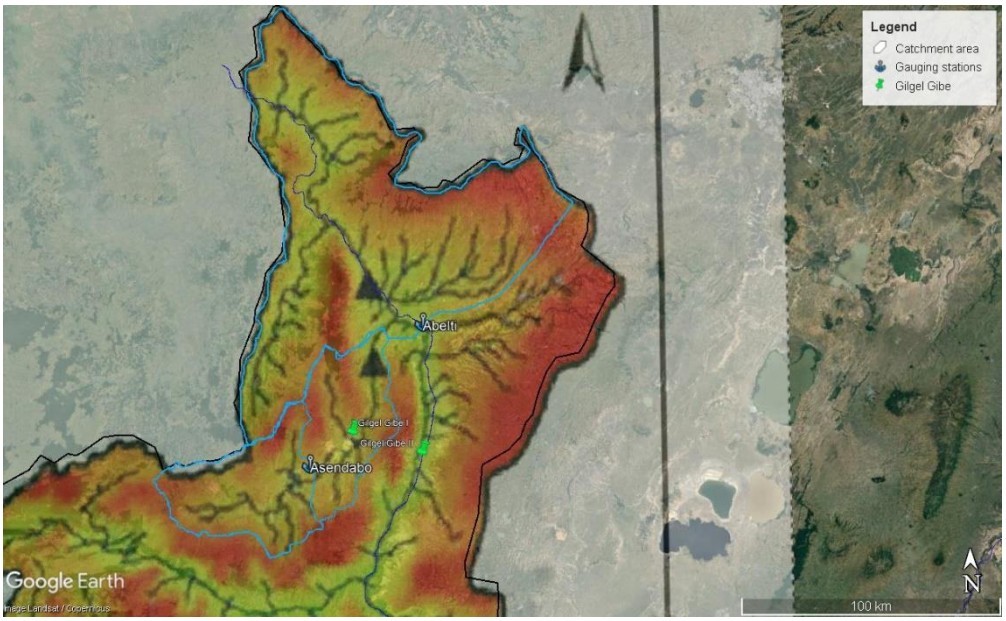

Figure 22. Overlay of the elevation map on Omo-Gibe basin map

Appendix 3

Table 6. Monthly streamflow at Assendabo station from 1985 to 2006

m^3/s	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	2.6	1.7	1.2	5.6	12.8	74.0	69.8	108.1	90.1	36.8	12.7	6.0
1986	2.1	2.6	4.1	3.6	4.4	34.2	74.1	67.8	85.9	32.9	10.2	7.0
1987	3.2	0.5	7.4	5.5	16.0	41.4	57.1	61.3	76.4	35.5	15.2	4.0
1988	3.9	4.3	1.6	1.4	3.4	10.7	56.1	155.5	121.7	94.8	20.5	7.3
1989	4.8	5.0	3.5	21.1	11.4	27.3	52.1	76.3	77.7	41.6	14.9	17.1
1990	8.1	6.4	9.6	8.9	16.3	48.6	91.0	75.9	120.0	5.8	92.5	9.5
1991	6.2	5.2	7.4	4.6	10.5	34.7	67.1	112.7	84.3	20.3	3.4	2.4
1992	0.6	1.5	3.1	3.2	10.7	42.2	85.2	186.8	118.1	74.5	19.9	7.7
1993	7.2	8.5	4.6	14.3	27.1	51.9	94.4	138.2	108.2	49.4	22.9	20.1
1994	5.8	3.2	4.3	5.0	21.4	70.1	126.4	167.5	124.2	27.3	11.3	5.8
1995	3.2	3.0	2.5	8.0	11.6	15.1	45.2	64.8	78.2	18.7	7.9	6.2
1996	36.4	36.4	18.5	16.4	34.6	88.2	87.0	123.0	89.6	45.8	17.1	16.3
1997	6.5	3.2	2.4	17.1	26.2	70.1	69.8	104.2	71.6	71.2	125.3	53.4
1998	28.4	16.1	16.3	11.7	20.6	30.1	84.7	171.2	99.2	84.5	34.2	14.9
1999	10.1	5.9	7.3	6.1	16.6	35.6	79.6	108.7	59.6	72.6	26.0	11.0
2000	6.4	3.8	2.4	7.3	19.9	29.8	62.9	82.8	87.8	75.5	36.5	11.0
2001	22.9	23.6	24.7	24.0	36.6	67.7	117.6	106.5	77.5	49.9	29.8	15.4
2002	9.8	5.9	7.2	9.7	7.1	29.7	57.8	76.4	61.2	21.9	12.2	12.6
2003	10.6	5.3	9.0	10.0	5.7	24.3	81.0	85.2	94.6	30.8	12.7	9.9
2004	6.5	5.0	4.1	6.7	13.4	32.6	65.2	99.1	95.6	77.7	20.5	13.2
2005	9.2	4.8	10.3	8.8	40.7	34.7	68.0	129.5	125.7	51.9	19.7	9.9
2006	7.7	9.2	9.8	12.1	14.4	35.9	121.1	169.1	108.9	52.0	29.5	23.5

Table 7. Monthly streamflow at Abelti station from 1985 to 2006

m^3/s	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	7.9	4.9	2.9	4.8	20.5	62.8	238.7	519.8	461.6	139.2	35.7	14.8
1986	8.2	6.8	8.4	8.8	7.5	97.1	375.2	350.2	409.9	151.1	30.0	14.3
1987	7.3	5.8	12.4	11.8	20.2	82.7	186.2	297.9	278.5	111.1	32.7	11.7
1988	6.7	6.3	4.5	2.2	4.6	19.9	589.3	993.1	758.4	675.9	308.1	29.7
1989	24.4	16.8	12.2	36.6	20.5	62.4	227.5	459.8	453.5	145.3	41.9	35.7
1990	12.9	15.5	15.2	16.8	19.9	95.0	355.7	663.9	567.1	22.3	389.9	37.1
1991	24.0	20.1	28.8	17.5	41.3	141.5	279.8	477.6	354.0	81.3	12.8	9.1
1992	12.1	16.1	8.6	8.9	25.9	80.4	298.4	961.0	568.6	309.2	79.7	167.6
1993	28.0	33.3	13.6	54.3	117.2	257.6	530.1	839.6	624.5	204.2	92.3	81.6
1994	53.8	12.0	16.4	19.0	82.8	292.7	476.1	1071.1	696.3	110.5	44.4	22.3
1995	12.2	11.4	9.3	31.2	45.8	60.2	186.1	782.4	465.9	74.7	30.8	23.8
1996	167.6	167.6	74.2	63.8	157.6	455.7	629.4	1113.2	456.5	284.9	103.5	63.7
1997	47.3	41.8	25.3	59.5	73.5	230.8	383.9	489.2	350.4	376.8	402.8	198.8
1998	92.5	52.9	55.2	34.5	64.7	117.9	505.7	1004.5	745.1	586.3	248.1	92.7
1999	54.7	30.8	27.9	22.4	54.3	139.4	346.5	611.4	377.1	471.2	173.4	55.0
2000	39.1	22.2	14.9	25.4	63.1	114.5	296.2	524.8	477.6	315.9	149.3	43.3
2001	92.4	27.6	31.2	29.3	67.6	258.8	554.8	723.6	588.8	327.0	140.5	61.1
2002	38.4	22.7	28.1	38.2	20.7	94.6	276.6	401.2	343.9	87.7	32.2	24.5
2003	16.5	9.8	13.0	17.5	8.6	37.8	233.0	474.0	440.6	184.9	52.1	29.1
2004	45.4	61.2	67.9	82.4	89.1	141.9	282.3	373.6	368.5	356.9	131.7	52.3
2005	35.9	18.6	40.4	43.3	78.1	99.3	262.0	475.0	488.1	286.1	110.5	78.4
2006	65.5	62.6	56.5	66.0	49.4	110.7	295.1	798.8	555.0	292.7	126.8	76.7

Appendix 4

Table 8. Computed runoff using CFSR and CPCP datasets

	Q (CPCP)	Q (CFSR)
1985	535.482	464.7642
1986	814.593	542.5436
1987	452.897	502.5848
1988	611.2499	623.849
1989	488.5681	484.8227
1990	470.0083	464.642
1991	441.8429	481.259
1992	592.1386	667.8921
1993	614.0423	648.3755
1994	547.7454	415.2983
1995	722.4478	629.3941
1996	611.9962	641.8043
1997	488.4698	623.9397
1998	606.3521	651.7839
1999	303.297	613.8538
2000	126.719	582.5372
2001	141.1265	419.0569
2002	117.155	442.2632
2003	76.3541	554.2938
2004	89.5827	690.2399
2005	319.1816	732.5555
2006	386.7972	700.4985
2007	289.3728	1005.1486
2008	308.2375	546.7717
2009	437.6094	880.1127
2010	274.5921	1188.2591
2011	170.8596	1018.5942
2012	422.8001	895.5815
2013	374.7584	1071.2435
2014	317.7786	934.3338
2015	302.0745	1028.8623
2016	393.9085	955.7579
2017	262.8611	845.5196
2018	58.476	372.1167