POLITECNICO
MILANO 1863

SCHOOL OF CIVIL, ENVIRONMENTAL AND LAND MANAGEMENT ENGINEERING
MASTER OF SCIENCE IN ENVIRONMENTAL AND LAND PLANNING ENGINEERING

ASSESSING THE VALUE OF HYDROCLIMATIC SERVICES FOR HYDROPOWER MEGADAMS

Master Thesis by:
Saoudi Yousra
Matr. 898360

Advisor:
Prof. Matteo Giuliani

Co-Advisor:
Dr. Louise Crochemore

Academic Year 2018 – 2019

Acknowledgments

I would first like to thank my advisor Professor Matteo Giuliani who guided me during this thesis. His enthusiasm in sharing his knowledge and ideas have stimulated my interest in water resource management. This thesis would not have been possible without his great support, dedication, and patience.
A special gratitude goes to Dr. Louise Crochemore from the Swedish Meteorological and Hydrological Institute who kindly provided us the forecast and climate scenarios data and helped us with precious advices in developing this work.
Many thanks to the all the research group of Environmental Intelligence for global change at Politecnico di Milano for sharing their tips.

I also want to thank my dearest friends Martina, Elisabetta, Laura, Sonia, Beatrice, and Roberta who have taken part in this journey. We shared joy and happiness during this five years and your unconditional love helped me to overcome difficult moments. I am very grateful to have met you. You have made me a better person.

Deep and special thanks go to my family who has always supported me to achieve this goal. To my parents, who have made a lot of sacrifices to allow me to pursue my dreams and have taught me to be patient in life. To my sister Leila and my brother Mohamed for their constant affection and encouragements. To my little sister Nora, who always reminds me the beauty of being a child.

Abstract

Population expansion, economic development, and increased frequency and intensity of extreme climate events are challenging water system management in many regions worldwide. These trends emphasize the need of accurate medium- to long-term predictions to timely prompt anticipatory water operations. State-of-the-art forecasts proved to be skillful over seasonal and longer time scales especially in regions where climate teleconnections, such as El Niño Southern Oscillation, or particular hydrological characteristics, such as snow- and/or baseflow-dominance, enable predictability over such long lead times. Recent studies have investigated the value of seasonal streamflow forecasts in informing the operations of water systems in order to improve reservoir management strategies. However, how to best inform the operations of hydropower systems is still an open question because hydropower reservoir operations can benefit from hydroclimatic services over a broad range of time scales, from short-term to seasonal and decadal time horizons, for combining daily and sub-daily operational decisions with strategic planning on the medium- to long- term. In this work, we propose a machine-learning based framework to quantify the value of hydroclimatic services as their contribution to increasing the hydropower production of the Grand Ethiopian Renaissance Dam (GERD) in Ethiopia. The GERD, with an installed capacity of more than 6,000 MW is considered the largest hydroelectric power plant in Africa and the seventh largest in the world. Its construction is part of the Ethiopian strategic hydropower development plan that aims to serve the growing domestic and foreign electricity demands. The quantification of the forecast value relies on the Information Selection Assessment framework, which is applied to a service based on bias adjusted ECMWF SEAS5 seasonal forecasts used as input to the World-Wide HYPE hydrological model. First, we evaluate the expected value of perfect information as the potential maximum improvement of a baseline operating policy relying on a basic information with respect to an ideal operating policy designed under the assumption of perfect knowledge of future conditions. Second, we select the most informative lead times of inflow

forecast by employing input variable selection techniques, namely the Iterative Input Selection algorithm. Finally, we assess the expected value of sample information as the performance improvement that could be achieved when the inflow forecast for the selected lead time is used to inform operational decisions. In addition, we analyze the potential value of forecast information under different future climate scenarios. Results show that the GERD operation informed with seasonal forecasts leads to a small improvement in the annual hydropower production in both historical and future periods. This potential gain becomes larger when we focus on the performance during the heavy rainy season from June to September (Kiremt season), and it further amplifies in the future scenario. The added production obtained with the forecast-informed operations of the GERD may represent an additional option in the current negotiations about the dam impacts on the downstream countries.

Riassunto

La crescita della popolazione, lo sviluppo economico e le condizioni climatiche sempre piú avverse stanno mettono a dura prova la gestione delle risorse idriche in molti bacini idrografici nel mondo. In questo contesto, la necessitá di avere previsioni accurate di medio-lungo periodo che consentano di anticipare le decisioni nella gestione dei sistemi idrici é in costante crescita. I recenti sistemi previsionali sono ormai capaci di fornire previsioni accurate su orizzonti stagionali, specialmente in regioni dove le teleconnessioni climatiche, come El Niño Southern Oscillation, o particolari caratteristiche idrologiche, consentono la prevedibilitá su periodi di tempo lunghi. Recenti studi hanno investigato il valore delle previsioni di afflusso stagionale per la gestione dei sistemi idrici. Tuttavia, si sa ancora ben poco di come meglio informare i sistemi idroelettrici in quanto essi traggono beneficio dai servizi idroclimatici su diverse scale temporali, dal breve periodo fino a orizzonti di tipo stagionale e decadale. In questa tesi, viene proposto una procedura basata su tecniche di Machine Learning con l'obiettivo di quantificare il valore dei servizi idroclimatici nel contribuire ad aumentare la produzione energetica della Grand Ethiopian Renaissance Dam (GERD) situata in Etiopia. La diga ha una capacitá installata di 6,000 MW ed é considerata il piú grande impianto idroelettrico in Africa e settimo al mondo. La sua costruzione fa parte del piano di sviluppo strategico etiope che ha lo scopo di fronteggiare la crescente richiesta domestica ed estera di elettricitá. La quantificazione del valore delle previsioni si basa sull' Information Selection Assessment framework, applicato ad un servizio che dipende dalle previsioni stagionali fornite dal sistema SEAS5 dell'European Centre for Medium-Range Weather Forecasts usate come input al modello idrologico World-Wide HYPE. Inizialmente, abbiamo valutato il massimo potenziale miglioramento di una gestione del sistema basata su poche informazioni nei confronti di una politica ideale che opera sotto la perfetta conoscenza delle condizioni future. Successivamente, abbiamo selezionato l'orizzonte previsionale piú rilevante utilizzando un algoritmo di input variable selection. Infine, abbiamo valutato l'effettivo aumento di produzione energetica che puó essere ottenuta quando

le previsioni selezionate vengono utilizzate per informare la gestione del sistema. Abbiamo inoltre analizzato il potenziale valore che le previsioni stagionali avranno in futuro considerando diversi scenari di cambiamento climatico. I risultati mostrano piccoli spazi di miglioramento nella produzione energetica annua sia nelle condizioni storiche che in quelle future. Tuttavia, questo potenziale guadagno incrementa valutando la performance durante la stagione delle piogge da giugno a settembre (Kiremt season), e viene ulteriormente ampliato nel futuro. Questa produzione energetica aggiuntiva puó presentare un'opzione nella attuale negoziazione sugli impatti futuri che la diga avrá principalmente su Egitto e Sudan.

Contents

List of Figures

List of Tables

Acronyms

ANNs Artificial Neural Networks

BOP baseline operating policy

CGMs General Circulation Models

CMIP5 Coupled Model Intercomparison Project 5

DDP Deterministic Dynamic Programming

DPS Direct Policy Search

ECMWF European Centre for Medium-Range Weather Forecasts

EMODPS Evolutionary MultiObjective Direct Policy Search

ENSO El Niño-Southern Oscillation

ESP Ensamble Streamflow Prediction

EVPI Expected Value of Perfect Information

EVSI Expected Value of Sample Information

GDP Gross Domestic Product

GERD Grand Ethiopian Renaissance Dam

HydroGFD Hydro Global Forcing Dataset

IIS Iterative Input Selection

IOP informed operating policy

ISA Information Selection Assessment

IVS Input Variable Selection

KNN K-Nearest Neighbour

MB Model Building

MISO multi-input-single output

MOEAs MultiObjective Evolutionary algorithms

NAO North Atlantic Oscillation

NWP Numerical Weather Prediction

POP perfect operating policy

RBFs Radial Basin Functions

RCP Representative Concentration Pathway

SDP Stochastic Dynamic Programming

SISO single-input-single output

SMHI Swedish Meteorological and Hydrological Institute

SSTs Sea Surface Temperatures

UBNB Upper Blue Nile Basin

1 Introduction

1.1 Global context

The rapid transformations occurring worldwide are putting high pressure on the availability of freshwater resources (*Rodell et al.*, 2018). One of the great challenges of the 21^{st} century is represented by climate change: constant rise in average temperatures and changes in the precipitation pattern are altering the water cycle, thus leading to an increase in the intensity and frequency of extreme climate conditions. It is expected that water scarcity and stress cases will grow in many regions around the world. The tendency is to observe increased occurrences of more severe and less predictable flood and drought events resulting in negative impacts in the society such as poor health, low productivity, food insecurity, and constrained economic development (*IPCC*, 2014). In particular, developing countries and third world countries will face the gravest risks from the changing climate since they are poorly equipped to adapt and prevent environmental threats.
On top of that, world population is expected to increase from current 7.6 billion to 8.6 billion in 2030, 9.8 billion in 2050 and 11.2 billion in 2100, with most of the population expansion concentrated in urban areas. While developed countries are growing at slower rates, especially Europe, more than half of global population increase, between now and 2050, is expected to occur in Africa (*United Nations*, 2017). The pressure given by global population growth, urbanization expansion, and rising economic prosperity will consequently increase food, energy, and water demands, thus further accentuating the stress on fresh-

water resources.
In this context, the importance of water-control structures, such as water reservoirs is growing. Hydropower systems can play a significant role in decreasing vulnerabilities of water and minimizing conflicts among multiple water uses. Moreover, hydropower represents the major source of renewable energy, corresponding to 58% of the total global renewable energy generation (*IRENA*, 2015b). However, managing existing water infrastructures more efficiently by adopting different technologies and methods, rather than planning new dams, represents a low-cost and flexible solution that can improve the resilience of the system against the increased variability of future extreme conditions (*Castelletti et al.*, 2008; *Giuliani et al.*, 2016b).
The recent advances in the skill of hydroclimatic services for the water sector have created significant opportunity to adopt these tools in supporting decision-making. Climate services generate a wide range of information on past, present, and future climate and they have the potential of becoming a supportive and flourishing market providing products that can better inform decision makers at all levels, from public administrations to business operators, when taking decisions for which the implications of a changing climate are an issue (*EC*, 2015). Notably climate services can provide accurate medium- to long-term climate predictions that can be used in water systems to timely prompt anticipatory operations, thus reducing the negative impacts given by extreme climate events.
Recent studies have investigated the value of seasonal streamflow forecasts in informing the operations of water systems to improve reservoir management strategies (e.g., *Anghileri et al.*, 2016; *Giuliani et al.*, 2019; *Nayak et al.*, 2018). State-of-the-art forecasts proved to be skillful over seasonal and longer time scales especially in regions where climate teleconnections, such as El Niño-Southern Oscillation (ENSO), or particular hydrological characteristics, such as snow- and/or baseflow-dominance, enable predictability over such long lead times. However, how to best inform the operations of hydropower systems is still an open question. Hydropower reservoirs operate often for multiple purposes (e.g. domestic/agricultural water supply, environment protection, tourism, flood protection, etc.) and, for this reason, they can benefit from hydroclimatic services over a broad range of time scales. Short-term forecasts (from few hours up to 2-3 days ahead) can improve flood control, which may induce spill of water with losses of production and, consequently, of economic revenue; medium-range forecasts (up to 7-15 days ahead) can support the optimal management of the production; long-term predictions (months ahead) can help in anticipating the effects of seasonal changes in water availability and implement drought

management plans (*Ramos et al.*, 2019).

1.2 Objective of the Thesis

The purpose of this thesis is to evaluate the potential value of seasonal forecasts in informing the operation of hydropower megadams under historical conditions and future scenarios affected by climate change. The objective of the analysis is to understand in which circumstances hydropower system can effectly benefit from long-term forecasts in terms of increasing hydropower production. The assessment of the forecast value relies on the Information Selection Assessment (ISA) framework (*Giuliani et al.*, 2015), which consists of three steps:

1. Quantification of expected value of perfect information as the potential maximum improvement of a baseline operating policy relying on a basic information with respect to an ideal operating policy designed under the assumption of perfect knowledge of future conditions.

2. Selection of the most informative lead times of inflow forecast by employing input variable selection techniques, namely the Iterative Input Selection algorithm (*Galelli and Castelletti*, 2013).

3. Assessment of the expected value of sample information as the performance improvement that could be achieved when the inflow forecast for the selected lead time is used to inform operational decisions.

The quantification of the forecast value under historical conditions relies on seasonal streamflow forecasts generated by a hydroclimatic service based on bias adjusted ECMWF SEAS5 seasonal forecasts used as input to the World-Wide HYPE hydrological model (*Arheimer et al.*, 2020); conversely, a synthetic forecast generation model was developed to simulates the existing forecast system under different climate change projections and estimate the future forecast value.
The proposed methodology is tested on a real case study, i.e. the Grand Ethiopian Renaissance Dam (GERD) in Ethiopia. The GERD has an installed capacity of more than 6,000 MW and is considered the largest hydroelectric power plant in Africa and the seventh largest in the world. The dam construction, started in 2011, is part of the Ethiopian strategic hydropower development plan that aims to serve the growing domestic and foreign electricity demands. Moreover, the GERD reservoir is expeted to partially buffer the effects of extreme events (floods and droughts) currently occuring in the country given by the

high variability of hydro-meteorological process largely correlated to ENSO phenomenon.

1.3 Thesis Outline

The thesis is structured into the following chapters.

Chapter 2 provides an overview of the state-of-the-art about metereological and hydrological forecasting in water systems management. It explores the types of existing forecasts, discussing skill, lead time, uncertainty, and the relation between forecast skill and forecast value. A focus on seasonal forecasting systems in the Upper Blue Nile Basin (Ethiopia) is also provided.

Chapter 3 reports a comprehensive description of the Grand Ethiopian Renaissance Dam case study, analysing the socio-economic development of Ethiopia, the hydrological and climatological characteristics of the Upper Blue Nile Basin, and the benefits and concerns raised by the dam construction. Then, we provide a description of the mathematical model of the system dynamics and the formulation of the objective function adopted for representing the interests of Ethiopia. Finally, the data used in the study are described.

Chapter 4 illustates the proposed methodology, which is composed of the following three main blocks of the ISA framework: 1) Quantification of the expected value of perfect information; 2) Lead time selection; 3) Assessment of the expected value of sample information. The procedure is repeated for both historical and future conditions.

Chapter 5 reports the results following the proposed 3-blocks framework for both historical and future conditions. A final comparison of the forecast value in the two periods is provided.

Chapter 6 summarizes the conclusions of this study and suggests possible directions for further research.

2

State-of-the-art

Most existing water reservoirs operating rules are conditioned upon a simple information set, including the day of the year and the storage, and, in few cases, also the previous day inflow (*Hejazi et al.*, 2008). However, since these operating rules are designed under a narrow range of forcing conditions, this type of approach could lose reliability and performance when these conditions are not met.

A practicable and zero-cost alternative to improve the management of most water systems is the direct use of observational data that describes the current conditions of the water system as indicated by *Denaro et al.* (2017). Their study, applied to a water reservoir that operates for flood control and water supply, showed that considering solely the Snow Water Equivalent (SWE) as additional information to the existing operating policy, allows for almost 10% improvement in the system performance. In fact, SWE informs the operating policies about the amount of water stored at high altitudes, which will be available during spring and allows to accumulate water for providing reliable irrigation over the summer.

However, relying on observational data to predict future conditions is not always possible. It might happen that the equipment fails or that the basin is located in an area where the conditions are hard to be measured (*Toth et al.*, 2000). In this context, forecast models or hydroclimatic services can provide advance information to potentially improve decision-making in water resource management and increase societal benefits, especially over the long term (*Block*, 2011). Along the past fifty years, there was a considerable improvement in weather

forecasting, with great advances in lead times, forecast skill and value. Forecasts are expected to be even more helpful over the next years, when extreme weather conditions will be more frequent and intense due to climate change (*Dai*, 2011). Several studies have been decently published analysing the multiples contributions of forecasting to the water resources operations, exploring different time horizons, different catchment types, and different operational objectives including flood control, water supply, hydropower generation, and combination among them.

2.1 Forecast lead time

Forecasts are usually divided in two categories: short-term and long-term or seasonal forecasts. The first range from one day to five days, rarely exceeding one week, while the second range from months to multiple years. Forecasts are usually skillful over short lead time and predictability tends to decrease over longer lead times (*Denaro et al.*, 2017). However, the value of a forecast does not only depend on the time horizon but also on other factors such as reservoir capacity, the operating objectives, and the catchment type.
Short-term forecasts are more valuable when the reservoir has a capacity smaller than its annual inflow volume, and it is operated for short-term operation purposes such as flood protection (*Saavedra Valeriano et al.*, 2010): a short lead time is generally sufficiently accurate to characterize the fast dynamics of the peak flow and to predict its magnitude and timing. Short-term forecasts can be based on direct observational data through the use of radars and gauges to capture the precipitation states or on more advanced models, such as numerical weather prediction and ensemble prediction systems that ensures accuracy, quickness, reliability, and robustness of the forecast system (*Adamowski*, 2008; *Cloke and Pappenberger*, 2009).
Nayak et al. (2018) demonstrated the value of short-term forecasts for the operations of Folsom Reservoir (California) for both flood control and water supply objectives. Results showed that water supply can be markedly improved without increasing the risk of floods due to the prediction of major flood events by a forecast model. However, they also found that longer-lead forecasts (30-day) do not improve performance much beyond the short-term (3-day) forecasts, even under perfect information.
On the other hand, long-term forecasts are more valuable for large reservoirs with medium to long-term operating objectives such as hydropower generation and water supply (*Anghileri et al.*, 2016). Long-term forecasts mostly rely on slow climate dynamics and, consequentially, require different weather infor-

mation to be generated. Seasonal forecasts proved to be skillfull in some specific areas where climate teleconnection, such as El Niño-Southern Oscillation (ENSO) may enable predictability over longer lead times. During the last few decades, there has been wide recognition that natural oscillations in areas of the Pacific Ocean have a significant impact on the patterns of weather and climate around the world (e.g., *Eltahir*, 1996; *Zaroug et al.*, 2014). The most dominant among these oscillations is known as the El Niño Southern Oscillation, a coupled ocean-atmosphere phenomenon observed over the tropical Pacific Ocean with a 2-7 years return period. ENSO involves Sea Surface Temperatures (SSTs) oscillation along the years, with a warm-phase called El Niño and a cold-phase called La Niña (figure 2.1).
Several studies attempted to use oceanic and atmospheric variables as predictors in seasonal hydrologic forecasting. Forecasts based on ENSO were shown to keep the skill from 6 to 9 months of lead time for some regions along the coastline of the Pacific Ocean (*Hamlet and Lettenmaier*, 1999). However, while in some locations such as the United States (e.g., *Hamlet and Lettenmaier*, 1999), western South America (e.g., *Poveda et al.*, 2011), or East Africa (e.g., *Block and Rajagopalan*, 2007), ENSO teleconnection is well defined, some regions are not substantially affected by ENSO. For example, in Europe local conditions depend on the concurrent state of other climate signals, particularly the North Atlantic Oscillation (NAO) (e.g., *Kingston et al.*, 2006).

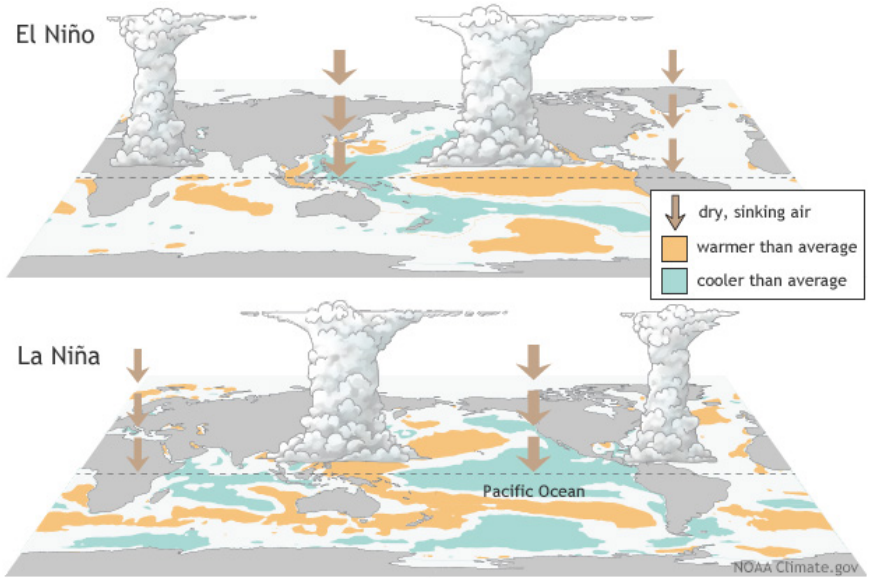

Figure 2.1: *Representation of both El Niño and La Niña, figure by National Oceanic and Atmospheric Administration.*

A recent study by *Giuliani et al.* (2019) detected the presence of potential teleconnections between multiple climate signals (e.i., ENSO and NAO), with the local precipitation in the Lake Como basin (Italy) and exploited these teleconnections to generate skillful forecast of local meteorological variables on a seasonal time scale. Results showed that the use of streamflow forecasts improves the performance over the baseline solutions by 44%; this gain increases 59% when observed preseason SST is directly used in informing water system operation.
Long term forecasts prove to be skillfull also in snow-dominated river basins as indicated by *Anghileri et al.* (2016). They developed a forecast-based adaptive control framework for a water system characterized by large inter-annual variability in precipitation and temperature. Results showed that season-long Ensamble Streamflow Prediction (ESP) forecasts improved operations with respect to the baseline, remaining 35% below the perfect forecast value and that the inter-annual component of the ESP forecast contributes 20-60% of the total forecast value. Moreover, they investigated which type of forecast could generated most value with respect to the baseline under the assumption of different stress situations: for high demands with respect to the storage capacity or inter-annual carryover, only a perfect forecast would be able to properly anticipate the streamflow; seasonal ESP forecast value is highest when the demand is medium-high and the reservoir is relatively small; inter-annual forecasts worked best for big storages while their contribution is negligible for small reservoirs.

2.2 Types of forecast models

Hydrologic forecasts are typically generated via either dynamic, process-based climate models with outputs (e.g., precipitation and temperature) fed into hydrologic models, or via empirical, data-driven models which produce either meteorological forecasts to feed hydrologic models or directly predict future streamflows. The first approach models the physical dynamics between the principal interacting components of the hydroclimatic system (*Yuan et al.*, 2015). A classical example is the numerical weather prediction (NWP). NWP focuses on taking current observations of weather and processing these data with computer models to forecast the future state of weather. However, there are limits to how far into the future NWP could perform, since they are sensitive to the initial conditions. Small errors in the initial conditions can lead to large errors in the forecast. Furthermore, predictability is limited by model errors due to the approximate simulation of atmospheric processes of the state- of-

the-art numerical models (*Buizza*, 2002). A way to overcome this problem is to conceive an ensemble of forecasts or an ensemble prediction system (EPS), in which multiple predictions of possible states are put together, providing different values of forecasting and thus giving an indication of the range of possible future climate state. The concept of ensembles can be applied also for streamflow forecasting, called ensemble streamflow prediction (ESP), where historical sequences of climate data (precipitation, potential evapotranspiration, and/or temperature) at the time of forecast are used to force hydrological models, providing a plausible range of representations of the future streamflow states (*Harrigan et al.*, 2018).
Over the last decade, a new approach that combine the ESP with General Circulation Models (CGMs) into a climate model-based seasonal hydrologic forecasting is receiving more attention. CGMs describe the global behaviour of atmosphere, ocean, and land in an integrated way and they are skillful in predicting large-scale precursors, such as ENSO, over long lead times (*Yuan et al.*, 2015). In general, physical models are accurate, since parameters of these models have direct physical interpretation, and their values might be established through field or laboratory investigation and they can simulate a wide range of flow situations. However, they require more data to be constructed, and cannot integrate observed data directly to improve model results (*Elsafi*, 2014).
The second class of forecast models are the data-driven models, also known as black box models (*Sharma*, 2000). They are based on modeling the statistical relationship between input and output data, without explicitly considering the physical processes that are involved. An example of this approch is the linear regression model based on large-scale climate predictors correlated in a linear way. These forecasting models suffer from two main drawbacks. Regression coefficients are largely influenced by a small number of outliers due to limited data length and correlations between predictors and, moreover, nonlinear relationships cannot be captured. An alternative proposed by *Block and Rajagopalan* (2007) was to use local polynomial regression model, a nonparametric methods, to overcome the drawbacks of traditional linear regression. The method first identifies the K nearest neighbours (observed data points) in proximity of a specific point of interest, then it locally parametrizes the data with a polynomial function, and finally adds random residuals with the same standard deviation of the polynomial to generate an ensemble. This approach allows to minimize outliers' disturbance, detection of local correlations, and elimination of multicollinearity.
Other non-linear, data-driven methods commonly used are the K-Nearest Neighbour Method (KNN) and Artificial Neural Networks (ANNs), where no a priori

relationship between parameters and observed values has to be hypothesised and no knowledge of the underlying process is needed. The KNN method exploits the closeness between the most recent observations and K 'similar' sets of observations chosen in an adequately large training sample (*Toth et al.*, 2000). ANNs are computing systems inspired by the biological neural networks that are trained for recognizing relationships between a set of inputs and outputs. Many studies showed ANNs overcoming the more basic methods such as linear regression (*Toth et al.*, 2000; *Elsafi*, 2014). It is worth mentioning that despite data-driven hydrological methods are becoming increasingly popular in forecasting applications due to their rapid development times, minimum information requirements, and ease of real-time implementation, they are restricted to stationary data, which makes them potentially unreliable when projected into future conditions affected by climate change (*Adamowski*, 2008).

2.3 Forecast uncertainty

The last decade has seen growing research in producing probabilistic hydrometeorological forecasts and increasing their reliability (*Ramos et al.*, 2013). Forecasts can better anticipate hydrologic extremes and can contribute to inform the operation of water systems in order to increase the performance of water management strategies. However, they cannot provide an exact prediction of future conditions due to many sources of uncertainty, which especially increases with the increase of forecast lead time. Generally, the major source of uncertainty of a long-term forecast is the one related to the meteorological input (*Cloke and Pappenberger*, 2009). Observational uncertainty, related to data collection, is also relevant because of the temporal and spatial uncertainties that can alter the system. Moreover, the geometry of the system and approximations in the construction of forecast models can further reduce the forecast accurancy.
However, knowing the uncertainty information can play an important role in the decision-making process as indicate by *Ramos et al.* (2013). Their study showed how decisions may differ when providing forecasted value with the associated uncertainty. The results suggested that the higher the uncertainty of a prediction, the more conservative and risk averse the decisions taken were. Moreover, with uncertain information, more optimal decisions were also taken, improving the overall performance of the system.
Another study by *Giuliani et al.* (2020) showed that the added value of a forecast might be undermined if end-users are not able to adequately interpret the uncertainty associated to the forecast ensemble. They explored increasing levels of risk aversion in the use of forecasts by informing the operating policy with

the mean, the 25^{th} and 10^{th} percentiles, and the minimum of the forecast ensemble. Their results showed that a solution that uses the ensemble minimum produces a 0.14% higher performance than using the ensemble mean, whereas the 25^{th} percentile of the ensemble generates a 0.31% reduction of system performance. These numerical results pointed out the importance of adopting integrated frameworks that include decision models and account for end-user behavioral factors capturing different perception of risk and uncertainty.

2.4 Forecast value and skill

In the literature, the term forecast value indicates the operational value of using a forecasting system to support water management (*Anghileri et al.*, 2016). The forecast value is measured in terms of system performance improvement as defined by the operating objectives (*Murphy*, 1993). The forecast skill is instead defined as the ability of a model to accurately predict the reality within given upper and lower bound (*Hamlet and Lettenmaier*, 1999).
Over the past 50 years, there has been a continuous improvement with the developing of new technologies and models that contributed in greatly advancing forecast skills (*Lynch*, 2008). However, an increase in forecast skill does not necessarily translate into an improvement in the forecast value.
In a study on farmers decisions based on forecasts, *Li et al.* (2017) identified a clear non linear relationship between skill and value. Their study showed that some institutional forecast products attain both high forecast quality and high decision performance, but in many cases the decisions were still optimal even though farmers were informed by products with low forecast quality. Moreover, farmers' different perceptions of risk were shown to be a relevant variable influencing the system performance: while a risk-neutral profile presented an increase of 3% with respect to the baseline case, a risk-prone profile experienced 10% of increment.
In another study developed in the Lake Como system (Italy), a regulated lake operated for flood protection and irrigation supply, *Giuliani et al.* (2020) identified an exponential relation between the increase in forecast skill and the resulting gain in system performance. Results showed a 10 to 1 relation between skill and forecast value, meaning that large gains in forecast skills are necessary to generate moderate gains in end-user profit.

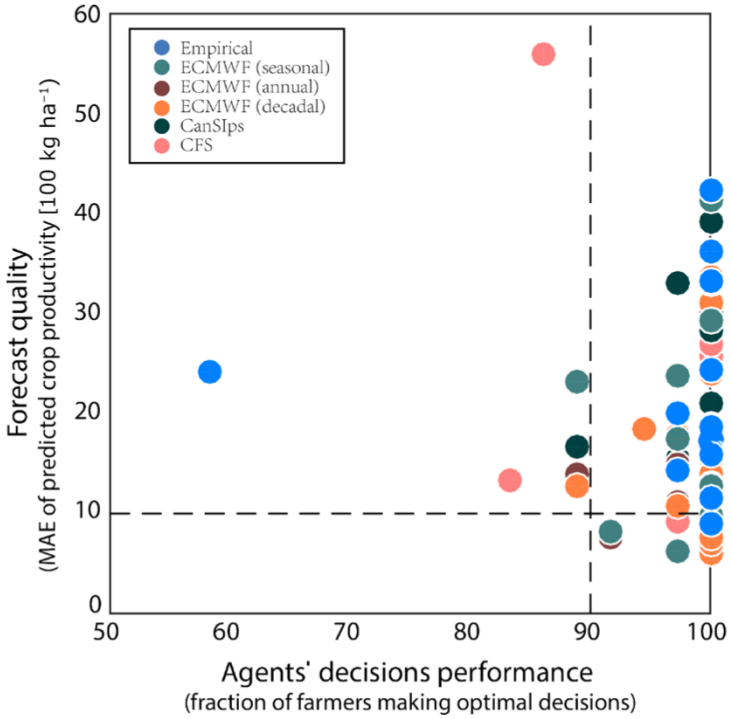

Figure 2.2: *Scatterplot of forecast quality of predicted crop productivity (y axis) and farmers' crop decisions performance (x axis) under different forecast products. Figure by Li et al. (2017).*

2.5 Seasonal forecast in the Upper Blue Nile Basin

In the previous section, we saw how global climate oscillations and particularly ENSO may contribute to extending forecast lead times. Since the aim of this thesis is to assess the value of streamflow seasonal forecasts in informing the operation of the Grand Ethiopian Renaissance dam which is located on the Blue Nile river in Ethiopia, we briefly discuss in this section state-of-art studies applied in the Upper Blue Nile Basin (UBNB) in order to provide evidences of the potential of seasonal forecasting in this region.

ENSO is significantly correlated with rainfall variations over the eastern side of the African continent (*Camberlin et al.*, 2001). Especially, Ethiopian seasonal precipitations are characterized by a strong intra- and interannual variability that are mainly driven by ENSO. This hydrologic variability resulting from low-frequency oscillations can produce extended periods of above- or below-normal precipitation making the country particularly vulnerable to the effects of flood and drought events. In this context, a successful seasonal forecasting system would have great economic and social value.

Interannual variability of precipitation within the UBNB has been investigated by previous researchers. *Eltahir* (1996) found that 25% of the natural variability

in the annual flow of the Nile is associated with ENSO and proposed using this observed correlation to improve the predictability of the Nile floods. *Amarasekera et al.* (1997) showed that ENSO episodes are negatively correlated with the floods of the Blue Nile. *Abtew et al.* (2009) found that high rainfall is likely to occur during La Niña years and low rainfall conditions during El Niño years. He also discovered that extremely dry years are highly likely to occur during El Niño years and extremely wet years are highly likely to occur during La Niña years. *Seleshi and Zanke* (2004) reported that the June-September rainfall in the Ethiopian Highlands is positively correlated to the Southern Oscillation Index (SOI) and negatively correlated to the equatorial eastern Pacific SST. Finally, *Zaroug et al.* (2014) highlighted the impact of timing and sequence of El Niño and La Niña on the drought and flood conditions over the UBNB. Results showed that ENSO exerts a significant influence on the June-September rainfall season and that 83% of El Niño events starting in April-June results in droughts while there is a 67% chance for occurrence of an extreme flood when an El Niño event is followed by a La Niña. They also recommended to use the main rainy season from June to September, called also the Kiremt season, in the seasonal forecasting of the Blue Nile due to its highly correllation with ENSO phenomenon.

Few studies attempted to use oceanic and atmospheric variables as predictors in seasonal hydrologic forecasting focusing on the June-September rainfall in Ethiopia. *Elsanabary and Gan* (2014) developed a rainfall predictive tool for forecasting June−September precipitations at weekly time steps in order to prevent flood events. The framework combines a wavelet principal component analysis (WPCA) to identify regions of Sea Surface Temperatures (SSTs) that are strongly teleconnected to the seasonal rainfall, an artificial neural networks-genetic algorithm to forecast the rainfall season over the UBNB from the above-selected SSTs at one-season lead time, and a statistical algorithm to disaggregate the seasonal rainfall forecasts to weekly rainfall. Their results showed that the forecasted seasonal rainfall agreed well with the observed data for the UBNB with a correlation value between 0.68 and 0.77. As an extension to their work, they used the disaggregated, weekly rainfall data to drive a basin-scale hydrologic (rainfall-runoff) to forecast the streamflow of the UBNB at up to one season lead time, which is useful for an optimal allocation of water among various competing users in the river basin.

Block and Rajagopalan (2007) developed a robust framework for generating ensemble forecasts of the Kiremt season precipitation in the UBNB. A suite of SST-related predictors that capture various aspects of the summer rainfall, including the large-scale ENSO teleconnections were first identified. Using the

best set of predictors, cross-validated ensemble predictions of the Kiremt season rainfall were generated from a local polynomial regression model. The forecast ensembles demonstrated significant skills during extreme wet and dry years compared to climatological forecasts utilized by the Ethiopian National Meteorological Services Agency.
In another study, *Block* (2011) also demonstrated the improved economic value and reliability resulting from a seasonal climate forecast coupled with hydropower system applied to the Upper Blue Nile Basin in Ethiopia. He explored forecast value under different managerial risk preferences and he found that even risk-averse actions, if coupled with forecasts, exhibits superior benefits and reliability compared with risk-taking tendencies relying on climatology. Moreover, a hydropower sensitivity test revealed a propensity toward poor-decision making when forecasts over-predict wet conditions: if the forecast predicts a wetter than normal year and is wrong, the cost is higher than when the forecast predicts a dryer than normal year and is wrong. If the reservoir operators release too much water downstream, expecting rainfall that doesn't come, then the following year there might not be enough water to meet production needs, while having held back too much water in expectation of a dry year that turns out to be wet has a smaller impact.

In summary, several studies explored the hydrologic variability of the Blue Nile river connected to ENSO, highlighting the potential for long-term forecast system in anticipating extreme events. However, just a few studies (e.g., *Block*, 2011) tried to assess the economic and social value of these forecasts in informing operational decisions and especially hydropower system operation. This quantification of the forecast value is however extremely important in a country like Ethiopia where hydropower represents the main source of energy and that can also contribute to the mitigation of adverse climate extremes.

3

Study Site

3.1 Ethiopia

Ethiopia, officially named the Federal Democratic Republic of Ethiopia, is a country located in the northeastern part of the Horn of Africa. As shown in figure 3.5, it is bordered by Eritrea to the north and northeast, Djibouti to the east, Somalia to the east and southeast, Kenya to the south, Sudan and South Sudan to the west. With over 112 million inhabitants, Ethiopia is the most populous landlocked country in the world and the second-most populated nation on the entire African continent after Nigeria. It occupies a total area of 1,100,000 km^2, and its capital, and also largest city, is Addis Ababa.

3.1.1 A growing country

Over the last decade, Ethiopia has made important development gains in education, health, food security, and economic growth. In 2015, the World Bank highlighted that the country had witnessed a rapid economic growth with Gross Domestic Product (GDP) increasing about 10.9% between 2004 and 2014, making the country one of the fastest growing economies in the world (figure 3.1). This improvement was mostly driven by government investment in infrastructure, as well as sustained progress in the agricultural and service sectors. Despite all this progress, Ethiopia still remains one of the poorest countries in the world, with an estimated annual per capita income of $772 (*WorldBank*, 2019). According to the United Nations' Committee for Development Policy (*CDP*, 2018), Ethiopia is listed among the least developed countries. Besides,

Ethiopia is also facing an increase in population which has grown from 18 millions in 1950 to almost 100 millions in 2015, with a population growth rate of 2.85%, a birth rate of 36.5 births per 1,000 total population and, a death rate of 7.7 deaths per 1,000 total population (*CIA*, 2016). The most recent United Nations estimates predict an increase of the population from the actual 112 millions to 205 millions in 2050 and about 295 millions in 2100 (figure 3.2). Moreover, the country is home to the second largest refugee population on the continent, hosting 928,600 registered refugees from South Sudan, Somalia, Sudan, Eritrea, and Kenya. Ethiopia is also experiencing a rapid urbanization with an increase in urban population of 8.1% from 1975 to 2000 and, from the most recent estimates, the phenomenon is destined to persist, with an urbanization rate estimated at 4.64% between 2015 and 2020 (*CIA*, 2016). In addition, extreme climatic and weather conditions, such as droughts and floods, have become increasingly common and costly in the country, leading to severe impacts on agricultural production, livestock, water resources, and human health. In this context, Ethiopia has to provide services and guarantee a better quality of life for a constantly increasing population that continues to move from rural to urbanized areas and that will be vulnerable to climate change impacts. This will require several economic investments, to ensure sufficient food, water availability, and easy access to electricity.

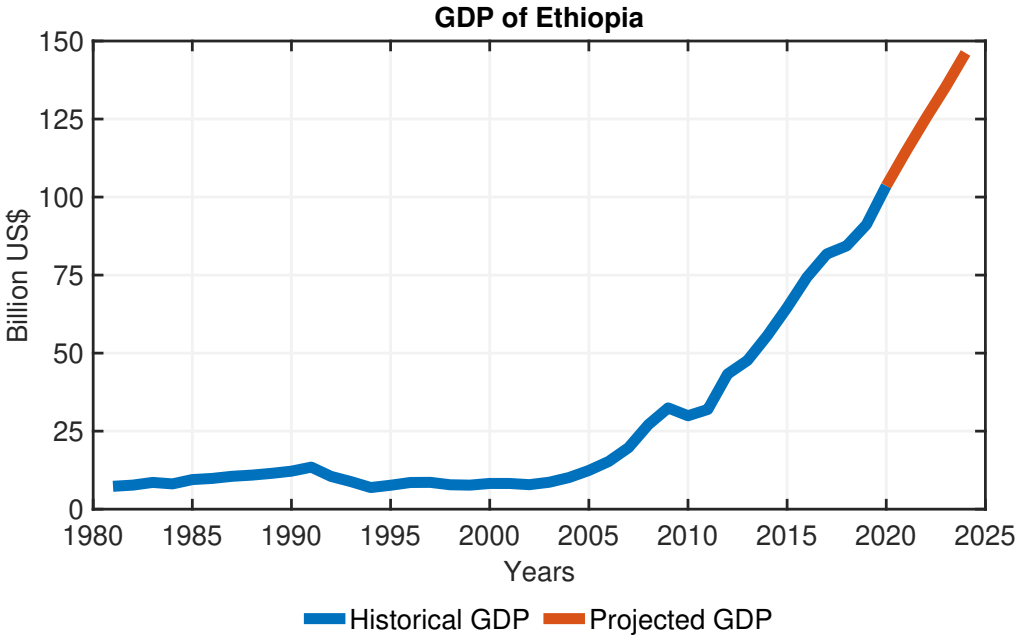

Figure 3.1: *Ethiopian Gross Domestic Product. Data from WorldBank (2019).*

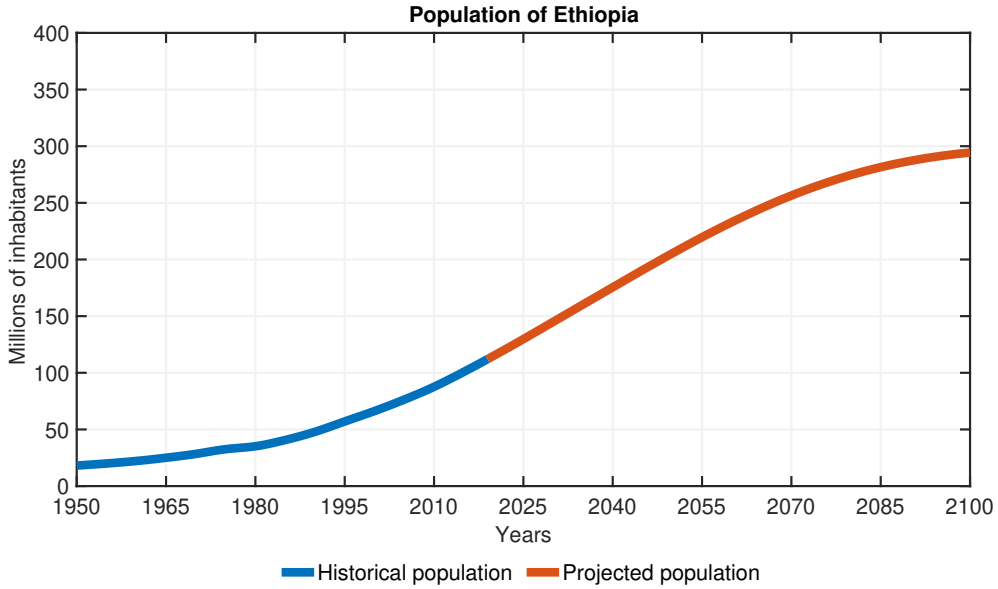

Figure 3.2: *Ethiopian Population. Data from UnitedNations (2019).*

3.1.2 Energy and hydropower

Ethiopia has 14 major rivers, including the Nile, which flow from its highlands and represent the largest water reserve in Africa. Ethiopia's hydropower capacity has been described as a *blue gold* that could contribute to economic growth of the country (*Verhoeven*, 2011). Ethiopia ranks second only to the Democratic Republic of Congo in economically feasible hydropower production among all the African nations (*King and Block*, 2014). According to *CIA* (2016), hydroelectric plants in the country represent around 86% of the total installed electricity capacity, while the remaining electrical power is generated from fossil fuels (3%) and other renewable sources (11%). Several reasons are behind Ethiopia's motivation to develop reservoir capacity and hydropower. First, the climate and hydrology of the country is extremely variable, on both interannual and intra-annual time scales, and floods and droughts are a recurring problem. A strong correlation between annual rainfall and the GDP growth rate has been demonstrated (*Bank*, 2006), and because the majority of Ethiopians rely on rainfed agriculture, this variability leads to severe vulnerabilities. Therefore the development of storage may partially buffer the effects of such climate extremes (*Zhang et al.*, 2015). Second, energy demand is expected to increase from 605 MW in 2012 to an estimated 2,540 MW in 2030 due to population, quality of life, and urbanization growth (figure 3.3). In this context, Ethiopia aims at becoming a regional power hub by exploiting its exceptional renewable energy resource and is undertaking a larger program of dam construction on several rivers (*Block and Strzepek*, 2012). Currently, Ethiopia acquires the majority of its energy from 15 different hydropower plants ranging in capacity from 11.4 to 1,870 MW, totaling 4,336.6 MW. The Ethiopian Electric Power Corporation *EEPCo* (2013), state-owned electric power producer, is implementing additional hydropower projects, in order to increase the internal hydropower production of nearly 15 GW. Three dams are under construction (table 3.1) and the most ambitious of these projects is the Grand Ethiopian Renaissance Dam (GERD) on the Blue Nile River, that, with its 6 GW of nominal hydropower capacity, will increase the electricity generation capacity in the country of 138%.

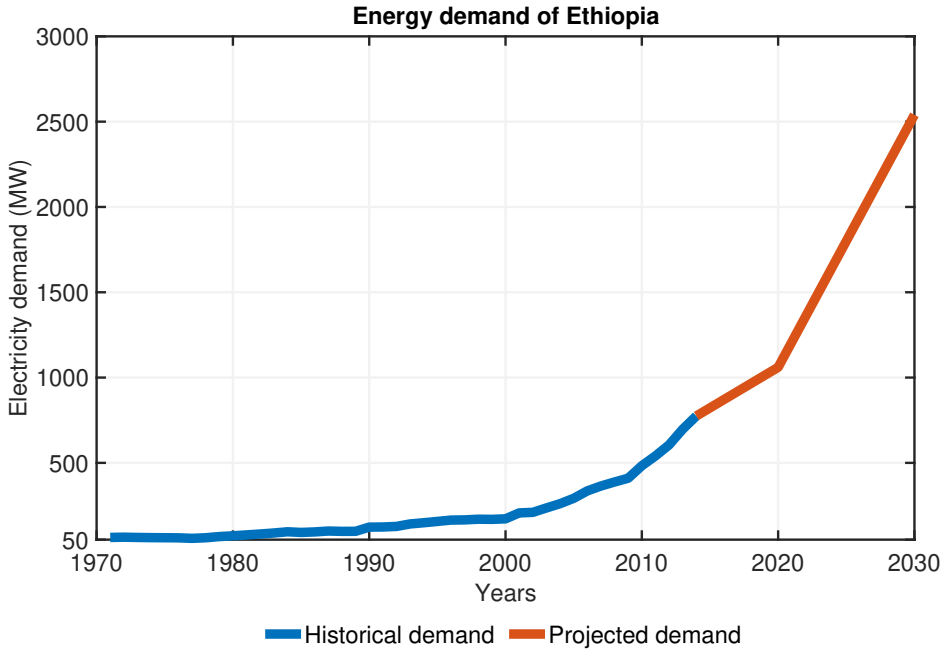

Figure 3.3: *Ethiopian electricity demand. Historical data from (WorldBank, 2019), estimate demand data from Ethiopia Growth and Transformation Plan.*

Table 3.1: *Operating and under construction (U.C.) dam in Ethiopia.*

Dam	1^{st} use	2^{nd} use	Commissioned	River	Reservoir Capacity (Km^3)	Installed capacity (MW)
Aba Samuel	Hydropower	Flood Control	1932	Akaki	0.035	-
Koka	Hydropower	Flood Control	1969	Awash	1.9	42
Fincha	Hydropower	Drinking water	1973	Fincha	0.65	100
Melka Wakena	Hydropower	Drinking water	1989	Shebelle	0.75	153
Gibe I	Hydropower	Flood Control	2004	Gilgel Gibe	0.92	180
Beles	Hydropower	Irrigation	2010	Beles	9.12	460
Tekeze	Hydropower	Flood Control	2010	Tekeze	9.3	300
Gibe II	Hydropower	-	2010	Omo	0.92	420
Gibe III	Hydropower	Flood Control	2015	Omo	14.7	1870
Genale Dawa III	Hydropower	Flood Control	2017	Ganale	2.6	256
Genale Dawa VI	Hydropower	Irrigation	U.C.	Ganale	0.18	257
GERD	Hydropower	Flood Control	U.C.	Blue Nile river	74	6,000
Koysha	Hydropower	Irrigation	U.C.	Omo	6	2,200

3.2 Grand Ethiopian Renaissance Dam (GERD)

The GERD's construction project started in 2011 and today is almost completed. The dam builder is the Italian company Salini Costruttori, which also served as primary contractor for the Gilgel Gibe II, Gibe III, and Tana Beles dams. The GERD, with an installed capacity of more than 6,000 MW and a catchment area of nearly 200,000 km^2 (*King and Block*, 2014), is considered to be the largest hydroelectric power plant in Africa and the seventh largest in the entire world (*IRENA*, 2015a).

The GERD is located 700 km northeast of the capital city Addis Ababa, in the Benishangul-Gumaz region of Ethiopia, along the Blue Nile River and not far from the border with Sudan. The Ethiopian Upper Blue Nile Basin is the largest Ethiopian river basin in terms of discharge, the second largest in terms of area and it hosts the largest tributary of the Main Nile (*Conway*, 2000). The Blue Nile River begins at Lake Tana in the northwestern Ethiopia highlands and is joined by many tributaries before reaching the Sudanese border. Then, it continues 650 km north-west until it converges with the White Nile in the city of Karthoum (Sudan), forming the Nile River. The Nile then enters Lake Nasser on the Sudanese-Egyptian border, created by the High Aswan Dam, and subsequently flows through Egypt to the Mediterranean Sea (figure 3.4). Although the White Nile Basin size is more than five times that of the Blue Nile Basin, the latter contributes significantly more streamflow to the main Nile River. According to *El-Fadel et al.* (2003) and *Conway* (1997), the majority (about 80%) of the annual Nile flow to Lake Nasser originates from the Ethiopian highlands, and the Blue Nile alone, with its 1,450 km (800 of which are inside Ethiopia) provides about 53% (50 km^3) of the annual flow.
The flow of the Blue Nile varies considerably over its yearly cycle due to the large inter-annual variation in precipitation driven mainly by El Niño- Southern Oscillation (ENSO) phenomenon (*Block and Rajagopalan*, 2007). Ethiopia is characterized by three main seasons (*Avery and Eng*, 2012):

- dry season from October to February (Bega season);
- a light rainy season from March to May (Belg season);
- a heavy rainy season from June to September (Kiremt season).

During the dry season the natural discharge of the Blue Nile can be as low as 113 m^3/s, while during the wet season the peak flow of the Blue Nile often exceeds 5,663 m^3/s in late August.

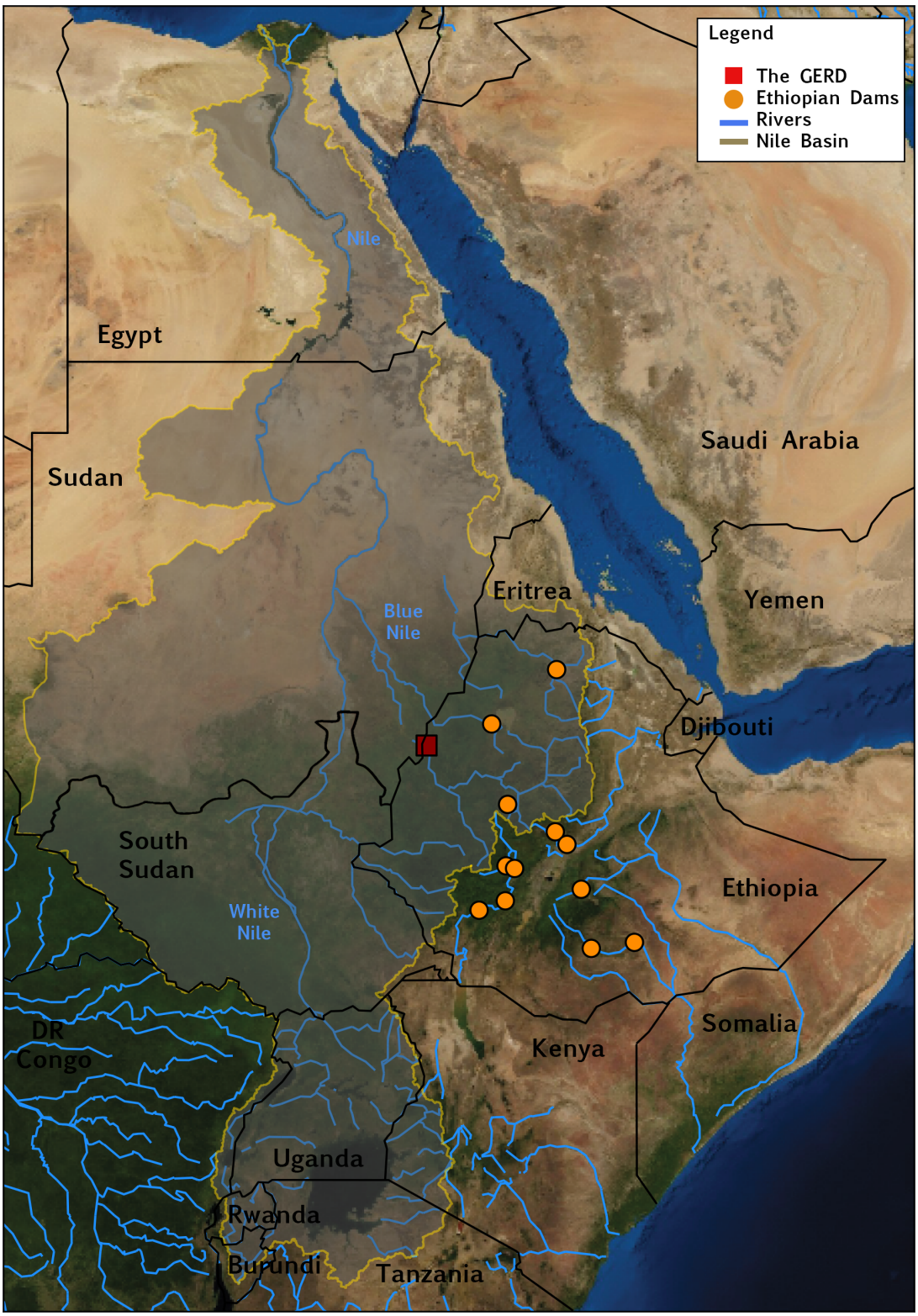

Figure 3.4: *Map of the Nile River Basin and ethiopian dams.*

Figure 3.5 shows some technical data of the GERD. The GERD reservoir is to be place in a deep gorge, producing a relatively small surface area considering its volume, and in a region with relatively cooler temperatures, allowing for a lower evaporation when compared to reservoirs further downstream, particularly those in desert locations (*Taye et al.*, 2016). The main dam ground level will be at an altitude of about 500 m asl, corresponding roughly to the level of the river bed of the Blue Nile. Counting from the ground level, the main gravity dam will be 155 m high (from 500 to 655 m asl), 1,780 m long and composed of roller-compacted concrete. A curved 5.2 km long and 50 m high rock-fill saddle dam will support the main dam and reservoir. The reservoir will have a storage capacity of 74 km^3 and a surface area of 1,874 km^2, when at the full supply level of 640 m asl. Hydropower generation can run between reservoir levels of 90 m, the so-called minimum operating level, and 140 m above ground level, the full supply level. The first 90 m of the height of the dam will be a dead height for the reservoir, leading to a dead storage volume of the reservoir of 14.8 km^3. A system of three spillways safeguards the project against a 30,200 m^3/s peak discharge, an event not considered to happen at all, as this discharge volume is the so-called Probable Maximum Flood. All waters from the three spillways are designed to discharge into the Blue Nile before the river enters Sudanese territory. The first 2 spillways activates when the level reaches 140 m above ground level, while the 3^{rd} one, located on the right abutment of the saddle dam, activates at 142 m, when the combined discharge of the other two spillways equals the 1,000-year flood. Two powerhouses located at the toe

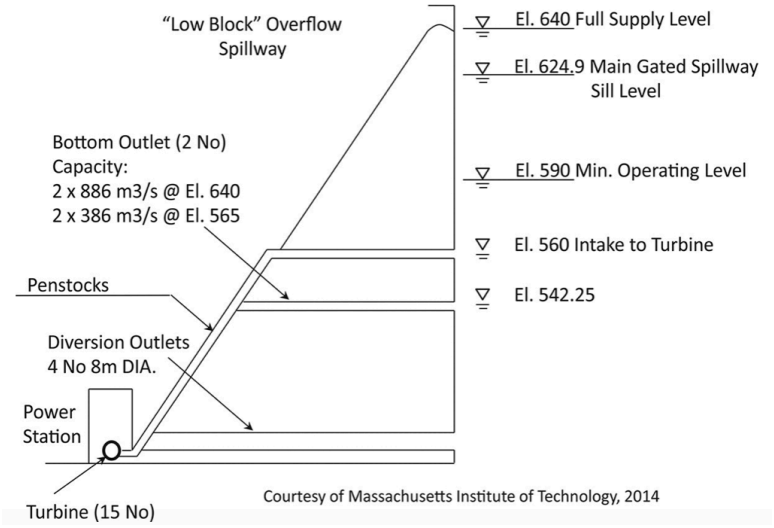

Figure 3.5: *Cross-section of the Grand Ethiopian Renaissance Dam with assumed hydraulic capacities. Source: MIT (2014).*

of the dam will house 16 Francis turbines at 375 MW each, totaling 6,000 MW in capacity, for an expected annual generation of 15 TWh. Once completed, the reservoir will take from 5 to 15 years to fill with water and different filling and managing reservoir policies will directly impact the millions of people in downstream countries, who rely on the Blue Nile water (*King and Block*, 2014).

Investments in large dam projects are often controversial. Supporters often denote benefits such as flood control, water supply, and hydroelectric generation, while critics frequently highlight permanent ecosystem changes, modification of river flows, reduction of fish passage and sediment accumulation, and community displacement (*Taye et al.*, 2016). The GERD's construction has been one of the most controversial issues in northern Africa, because of the potential negative effects for Sudan and Egypt, whose populations also live in the Nile Basin and depend on its water. Due to its massive dimensions, the GERD will surely change the Nile water flow, decreasing the downstream water availability at least during its filling period, constituting a breach of the 1959 Nile Treaty. In 1959, in preparation for the construction of the High Aswan Dam, Egypt and an independent Sudan signed one of the most important agreements for the geopolitical situation of the region, the 1959 Nile Waters Agreement for full control utilization of the Nile waters. This agreement (*Treaty*, 1959) allocated 55.5 billion and 18.5 billion m^3 of the Nile River water to Egypt and Sudan, respectively. Moreover, annual water loss due to evaporation and other factors were agreed to be about 10 billion m^3 and no consumptive allocations to other upstream riparian countries (Ethiopia, Uganda, Kenya, Tanzania, Rwanda, Burundi, Democratic Republic Congo (DRC), and Eritrea) were prescribed. The upstream Nile basin nations, including Ethiopia, contested these treaties (*Waterbury*, 2008) and, in 1999, all Nile Basin countries (Burundi, DRC, Egypt, Ethiopia, Kenya, Rwanda, Sudan, Tanzania, and Uganda) came together to form the Nile Basin Initiative (*NBI*, 2012) with the objective to establish a basin-wide water management institution (*Conway*, 2000). Since May 2010, Ethiopia and the other upper riparian states have launched the Cooperative Framework Agreement (CFA) to ensure an equitable and reasonable utilization of the Nile waters and the obligation to avoid significant harm to its riparian neighbors. However, Egypt and Sudan have opted not to sign the CFA, highlighting the tensions that existed even before the announcement of the GERD.

Here, we briefly summerize the potential benefits and concerns raised on the GERD's construction. The GERD will be the hub for clean and renewable energy supply for Ethiopia and other African countries. According to the Ethiopian government (*MoFED*, 2010), hydroelectricity from the GERD is expected to primarily satisfy national demand in order to reduce energy poverty among the almost 60% of Ethiopians who lack access to electricity, while surplus energy will be exported to East Africa countries to improve the coverage of the electricity. GERD's benefits are not limited with power supply but the dam can bring advantages also to other sectors: the volume of the reservoir will be two to three times that of Lake Tana (Ethiopia), which allows to expect up to 7,000 tonnes of fish annually, representing a significant food and commercial source for the region. Moreover, the reservoir could become an hotspot for tourism (*ENA*, 2017) and navigation. The dam could also serve as a bridge across the Blue Nile, complementing another bridge that was under construction in 2009 further upstream (*Ethiopia*, 2011). From a social point of view, the GERD will develop local industry, and it has already created up to 12,000 jobs during the construction phase (*Yihdego et al.*, 2017). In addition, the dam's ability to regulate hydrologic variability will lead to a reduction in property losses due to flooding and will allow for increased agricultural production in downstream countries, especially in Sudan, ensuring reliable all season supply to irrigation schemes, and reducing harvest losses caused by water shortages during dry periods. Water conservation is another major benefit given by the GERD which will minimize the evaporation losses from dams located in less favourable downstream desert locations. Since much of the Nile water is presently lost to evaporation and infiltration as it flows north toward the Mediterranean Sea (*Guariso and Whittington*, 1987), such losses can be significantly reduced storing water in Ethiopia. The evaporation rate in Ethiopian highlands is equal to 1,150 mm/year while in Sudan it reaches 2,500 mm/year (*Block*, 2011), which is similar to the one of High Aswan Dam (i.e. 2,400 mm/year) (*Mulat and Moges*, 2014). These differences may save 25% of the current evaporation losses conserving water in the GERD rather than the High Aswan dam (Egypt). Moreover, the GERD is expected to reduce sediment loads thus extending reservoir life in Sudan and Egypt and decreasing silt build-up in irrigation canals (*Taye et al.*, 2016). It is estimated that Sudan could save $50 million per year in dredging costs (*Tesfa*, 2013) and that silt and sedimentation reduction could improve water quality and reduce treatment costs for drinking water supply (*Taye et al.*, 2016).
However, one of the main issues in debate about the GERD construction is the time of filling the reservoir. Egypt and Sudan will be especially vulnerable to the time frame of filling the dam, which may take anywhere from 5 to 15 years.

Egypt fears a temporary reduction of water availability thus impacting its hydropower, water supply, and irrigation systems. On the other hand, despite Sudan has recognized the benefits that the dam is likely to provide in the country, there are concerns that the sediment yield reduction will decrease downstream soil fertility, particularly affecting flood-recession agriculture with estimated losses in recession agricultural land on the order of tens of thousands of hectares (*Taye et al.*, 2016). Moreover, the GERD will also bring some drawbacks even in Ethiopia itself. Exporting the electricity from the dam towards other countries would require the construction of massive transmission lines to major consumption centers such as the Sudanese capital Khartoum, located more than 400 km away from the dam. A cost of 4.8 billion US$ has already been spent for the GERD construction which corresponds to about 5% of Ethiopia's GDP and adding power transmission lines will require more funding. Moreover, the dam is displacing thousands of people in the Benishangul-Gumuz Region. An estimated 5,110 people were resettled from the downstream area in 2013 and close to 20,000 people are being relocated. The majority of the affected population are indigenous people of Gumuz and Berta, who live in some of the worst economic conditions with little access to development infrastructure (*Rivers*, 2012).

Since the aim of this thesis is to asses the value of seasonal hydroclimatic services in informing the GERD operation to increase hydropower production, we neglect in the following analysis the presence of Sudan and Egypt and we assume that the GERD reservoir is already filled and operated in regime conditions.

3.3 GERD Model

The conceptual model of the GERD system is composed of the regulated GERD reservoir and its hydropower plant. The system dynamics is based on a mass balance equation of the water storage using a monthly timestep:

$$s_{t+1} = s_t + n_{t+1} - r_{t+1} \tag{3.1}$$

where s_t is the storage of the reservoir at time t, while n_{t+1} is reservoir's inflow and r_{t+1} is the released water volume from t to t+1. The time notation indicates the instant when the variable is deterministically known. The reservoir storage is measured at time t, while the water volumes entering and exiting the reservoir are evaluated in the time interval [t; t+1) and so they are known at time t+1. The released volume is a nonlinear function of storage, inflow, and the release decisions u_t at each time step, which accounts for any possible deviation of the actual release r_{t+1} from the decision u_t due to unintentional spills or any other physical legal constraint (*Soncini-Sessa et al.*, 2007). The effective release of the GERD is calculated as:

$$r_{t+1} = min(R_{max}(s_t), max(R_{min}(s_t), u_t)) \tag{3.2}$$

where R_{min} and R_{max} are the minimum and maximum monthly GERD release, respectively, which integrate over the time step the instantaneous functions r_{min} and r_{max} that depend on the reservoir's level (h_t), that is directly linked to the storage (s_t), defined as:

$$r_{min} = \begin{cases} 0[m^3/s] & h_t \leq 140m \\ 22785.7[m^3/s] & h_t \leq 142m \\ 24285.7[m^3/s] & h_t > 142m \end{cases} \tag{3.3}$$

$$r_{max} = \begin{cases} 0[m^3/s] & h_t \leq 90m \\ qTMax & h_t \leq 140m \\ 22785.7[m^3/s] & h_t \leq 142m \\ 24285.7[m^3/s] & h_t > 142m \end{cases} \tag{3.4}$$

where $qTMax$ is the maximum turbine discharge, equal to 4,985.7 m^3/s. According to equation 3.1, evaporation is neglected due to the lack of reliable data. The release decision variable u_t in equation 3.2 is determined by a closed

loop policy, defined as a periodic sequence of control laws with period T which depend on the reservoir storage and previous month inflows:

$$u_t = \mu_t(s_t, n_t) \qquad p = [\mu_0(\cdot), ...\mu_{T-1}(\cdot)] \tag{3.5}$$

In this work, we consider one stakeholder which is Ethiopia and its corrispondent objective is formulated as maximization of the GERD hydropower production, defined as:

$$J_{0,H}^{HP} = \frac{1}{H}(\sum_{t=0}^{H-1} g_{t+1}^{HP}) \tag{3.6}$$

where

$$g_{t+1}^{HP} = \frac{\eta \cdot g \cdot \gamma \cdot r_{t+1} \cdot (h_t - h_{down})}{1000} \tag{3.7}$$

where η is the turbines' efficiency equal to 0.95, g is the gravity acceleration, approximated at 9.8 m/s^2, γ is water specific weight equal to 1000 Kg/m^3, r_{t+1} is the GERD actual release, h_t is the reservoir water level, and h_{down} is water level downstream the main dam equal to 10 m.

3.4 Available data

The model of the GERD reservoir and power plant is based on the data reported in tables 3.2 and 3.3. These data have been retrieved mostly from the Nile Basin Initiative (*NBI*, 2012), except for the spillway (*Ferraro et al.*, 2015), and the hydropower production (*EEPCo*, 2013).

Table 3.2: *Grand Ethiopian Renaissance Dam's data.*

Number of turbines		16
Tot. max. month discharge (turbines)	[Mm^3/month]	12,923
Tot. max. month discharge (turbines)	[m^3/s]	4,985.7
Nominal hydropower production	[MW]	6,000
Minimum operational level	[m]	90
Maximum operational level	[m]	140
Turbine efficiency		0.95
Downstream water body level	[m]	10
Number of spillways		3
1^{st} spillway activation level	[m]	140
2^{nd} spillway activation level	[m]	140
3^{rd} spillway activation level	[m]	142
1^{st} spillway discharge	[m^3/s]	15,000
2^{nd} spillway discharge	[m^3/s]	2,800
3^{rd} spillway discharge	[m^3/s]	1,500

Table 3.3: *Grand Ethiopian Renaissance Dam's elevation, storage and surface.*

Elev. [m]	Storage [Mm3]	Surface [Mm^2]
500	0	3
510	10	11
520	20	29
530	50	61
540	750	111
550	2,000	180
560	3,000	272
570	6,000	387
580	9,800	531
590	15,000	703
600	21,500	905

The GERD inflow data are derived from monthly inflow observations collected at El Diem during the time period 1965-2017 (*Wheeler et al.*, 2018). Figure 3.6 illustrates the historical inflow data over the period 1965-2017, along with the ciclostationary mean (red line) and shows that inflows exhibit a clear periodic pattern with large hydrologic variability during August, September, and October where inflow values range from less than 1,000 to more than 8,000 m^3/s.

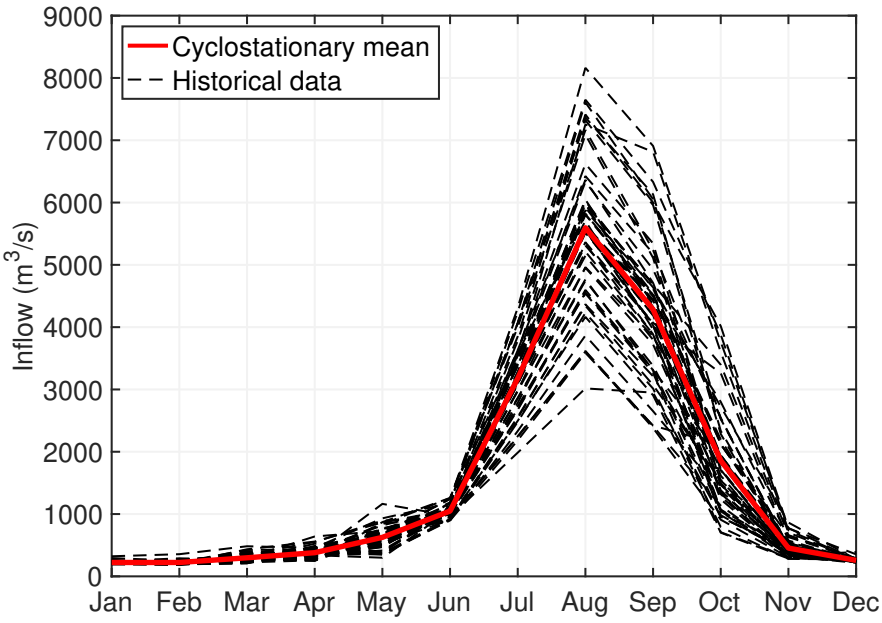

Figure 3.6: *Monthly inflow: cyclostationary mean and historical data (1965-2017).*

The data related to forecasted inflows in the historical period are provided by the Swedish Meteorological and Hydrological Institute (SMHI), which generates seasonal streamflow forecasts from the World-Wide HYPE hydrological model using as input data bias adjusted ECMWF SEAS5 seasonal forecasts. Bias-adjustment is conducted on all members of SEAS5 using the Hydro Global Forcing Dataset (HydroGFD) as reference, an internal SMHI operational system for generating corrected re-analysis fields of precipitation and temperature. The World-Wide HYPE model is a semi-distributed catchment model that aims to reproduce streamflow and water balance at global scale (*Arheimer et al.*, 2020). It covers an area of 135 million km^2, divided into some 131,300 catchments following the river networks, with an average size of catchments 1020 km^2. The model uses a large number of open databases and is calibrated against time-series of various sources of observations (both from in situ monitoring and Earth observations). The hydrological forecasts are delivered once

a month in the form of a 25-member ensemble with a 7-month lead time from 1993 to 2014.

The data related to climate change scenarios are also provided by SMHI. They correspond to a monthly mean value of the expected percent change of water discharge defined as

$$\delta = 100 * \frac{\text{Future period} - \text{Reference period}}{\text{Reference period}} \tag{3.8}$$

for three future time periods (i.e., 2011-2040, 2041-2070, 2071-2100) using as reference period the 1971-2000 and considering two Representative Concentration Pathway (RCPs), RCP 4.5 (moderate emission scenario) and RCP 8.5 (high emission scenario). These scenarios are originated from different CMIP5 General Circulating Models (table 3.4) where precipitation and temperature were bias adjusted against HydroGFD prior to running World-Wide HYPE model. These data were used to compute future projected inflows, as follows:

$$\text{Future inflows} = \text{Historical inflows} + \frac{\delta * \text{Historical inflows}}{100} \tag{3.9}$$

where *Historical inflows* are the observed GERD inflows over the time period 1993-2014. Prior to compute the projected inflows (equation 3.9), a statistical analysis based on historical data was first performed in order to remove climate models that simulates historical inflows that are largely out of the range of the historical hydrological variablity. To combine delta values of the remained climate models, the median is computed in order to filter possible anomalous spikes. Moreover, a correction factor calculated as the ratio between the simulated inflows by climate models and observed inflows over historical period, was applied to the future inflows of the first three months (January, February, March) of the year in order to remove a systematic bias introduced by the model that clearly overestimates the observed inflow values.

Table 3.4: *CMIP5 General Circulating Models considered reporting modeling center and model name.*

Modeling center	**Model name**
Beijing Normal University	BNU-ESM
Centre National de Recherches Mété orologiques and Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	CNRM-CM5
Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Bureau of Metereology (BOM), Australia	ACCESS1.0 ACCESS1.3
Irish Centre for High-End Computing	EC-EARTH
Institute Pierre-Simon Laplace	IPSL-CM5A-LRI IPSL-CM5A-MR IPSL-CM5B-LR
Met Office Hadley Centre	HadGEM2-CC HadGEM2-ES
Max Planck Institute of Meteorology	MPI-ESM-LR MPI-ESM-MR
Norwegian Climate Centre	NorESM1-M
NOAA Geophysical Fluid Dynamics Laboratory	GFDL-CM3 GFDL-ESM2G GFDL-ESM2M
Beijing Climate Center, China Metereological Administration	BCC-CSM1.1 BCC-CSM1.1(m)

4

Methods and tools

In this thesis, we propose a machine-learning based framework to quantify the value of hydroclimatic services as their contribution to increasing the hydropower production of the Grand Ethiopian Renaissance Dam. A flowchart of the methodology used in this work is illustrated in figure 4.1. We quantify the forecast value by relying on the Information Selection Assessment framework (ISA) (*Giuliani et al.*, 2015). First, we evaluate the expected value of perfect information as the potential maximum improvement of a baseline operating policy relying on a basic information with respect to an ideal operating policy designed under the assumption of perfect knowledge of future conditions. Second, we select the most informative lead times of inflow forecast by employing input variable selection techniques, namely the Iterative Input Selection algorithm (*Galelli and Castelletti*, 2013). Finally, we assess the expected value of sample information as the performance improvement that could be achieved when the inflow forecast for the selected lead time is used to inform operational decisions. The potential value of forecast information is assessed under historical hydroclimatic conditions and under future climate chage scenarios in order to understand if forecasts are able to improve the system performance and to mitigate the negative impacts of climate change. For this purpose, we developed a synthetic forecast model that allows the generation of inflow forecasts over future time periods.

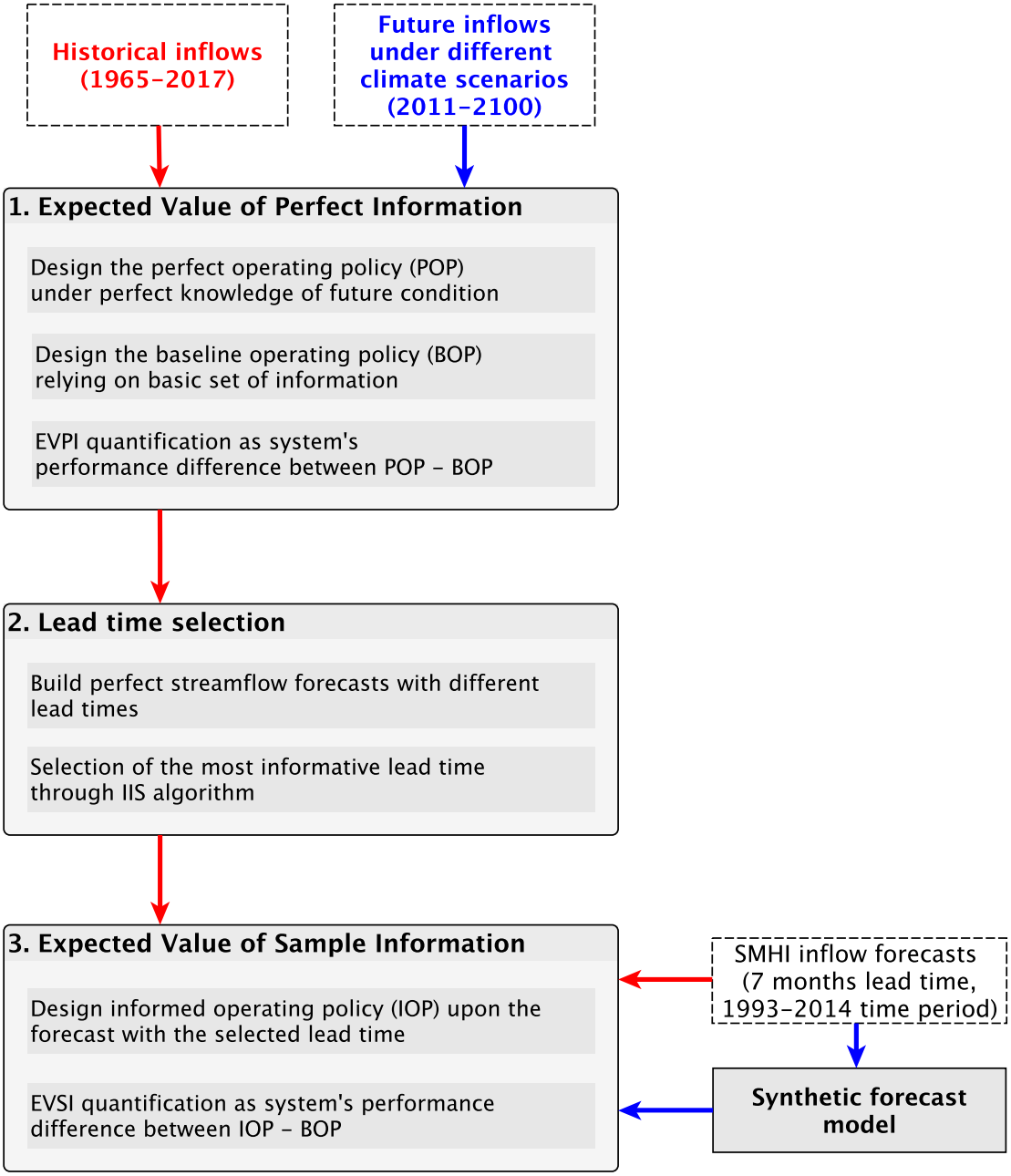

Figure 4.1: *Methodological flowchart of the thesis.*

4.1 Information Selection and Assessment (ISA) framework

The ISA framework (*Giuliani et al.*, 2015) was first designed to identify and select the most valuable information from a set of observational data for informing water systems' operations (figure 4.2). The aim was to remove data with observational error and estimation biases that could possibly negatively affect the system performance. The process can also be used to support the identification and selection of the most valuable forecasts. The analysis is performed considering two type of conditions: historical hydroclimatic conditions and future conditions under climate change and it consists of three main steps as follows:

1. **Expected Value of Perfect Information**
   We evaluate the Expected Value of Perfect Information (EVPI) as the potential maximum improvement of a baseline operating policy (BOP) relying on a basic set of information with respect to an ideal operating policy (POP) designed under the assumption of perfect knowledge of future conditions.
   The policy design problem is formulated as follows:

$$p* = arg \min_{p} | - J| \tag{4.1}$$

   where p is a closed loop operating policy (see equation 3.5) and J is the GERD hydropower production defined in equation 3.6.
   Since we are dealing with just one objective, the EVPI can be computed as a scalar difference:

$$EVPI = J^{\text{POP}} - J^{\text{BOP}} \tag{4.2}$$

   where J^{POP} is the perfect operating policy performance obtained by solving problem 4.1 with the full trajectory of historical inflow deterministically known over the entire evaluation horizon. The problem is solved via Deterministic Dynamic Programming (DDP) (*Bellmann*, 1957), while J^{BOP} is the value of the objective functions that could be obtained by an operating policy relying on a basic set of policy inputs, namely month of the year, GERD's storage, and previous month inflow. This latter was designed by means of Evolutionary MultiObjective Direct Policy Search (EMODPS) (*Giuliani et al.*, 2016a), thus defining the operating policy within a class of function p_θ and searching for the best parametrization $\theta \in \Theta$ for the given class (see Section 4.2.2).

2 **Lead time selection**

The input candidates can be different forecast models, different forecast variables, or different forecast lead times. Since this thesis is focusing on assesing the value of long-term or seasonal forecasts, we considered, as possible candidates, streamflow forecasts at different lead times from one month up to seven months. We select the most informative lead times of inflow forecast by employing input variable selection techniques, namely the Iterative Input Selection algorithm (IIS) (*Galelli and Castelletti*, 2013) coupled with Extremely Randomized Trees (*Geurts et al.*, 2006);(*Galelli and Castelletti*, 2013). The IIS algorithm is based on an iterative procedure, which allows the ranking of the candidate input variables according to their significance in explaining the output variable (see Section 4.2.3). The algorithm is run multiple time in order to filter the randomness associated to the construction of the extra-trees models that are used as regression model in describing the non linear input-output relationship. We considered as input variables perfect seasonal streamflow forecasts over different monthly lead times up to a maximum of 7 months ahead and as output variable the difference between the sequence of reservoir releases derived under the perfect operating policy and the baseline operating policy.

3 **Expected Value of Sample Information**

The Expected Value of Sample Information (EVSI) is assessed as performance improvement that could be achieved when the inflow forecast for the selected lead time is used to inform operational decisions. To perform the forecast informed operating policy (IOP) we used again the Evolutionary MultiObjective Direct Policy Search method. The Expected Value of Sample Information is computed as:

$$EVSI = J^{\mathrm{IOP}} - J^{\mathrm{BOP}} \tag{4.3}$$

where J^{IOP} is the improved operating policy performance obtained by solving problem 4.1 when the inflow forecast $F_{t+\tau}$ with the selected lead time τ is used in informing the operation of the hydropower system (i.e., $u_t = p(t, s_t, n_t, F_{t+\tau})$). For historical conditions, SMHI forecasts were considered, while for future conditions, syntetic forecasts were generated (see Section 4.2.4). The process consists in comparing forecasted inflows to the observed ones over the historical period in order to model the forecast errors and propagating the forecast uncertainty to any inflow trajectory.

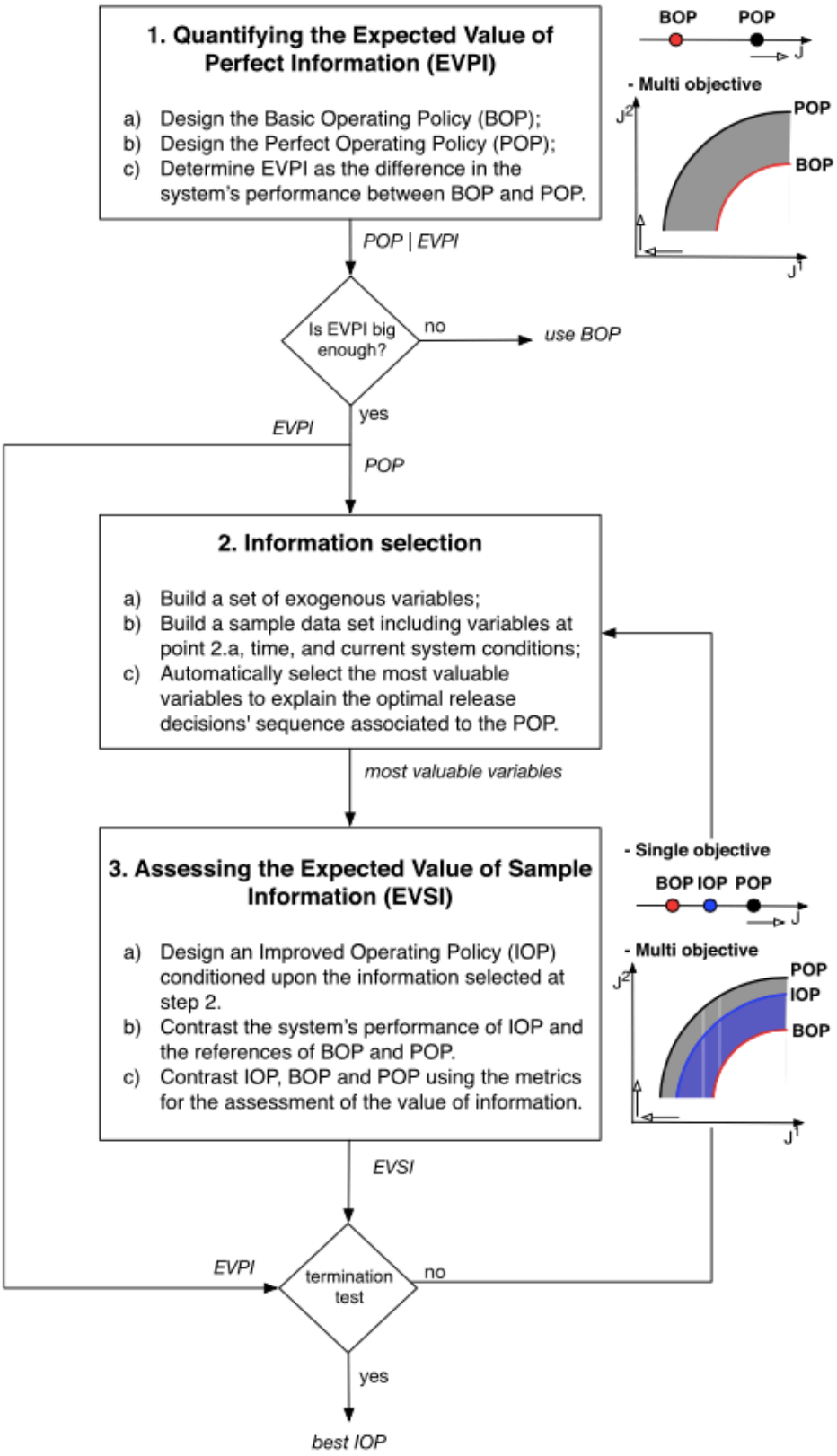

Figure 4.2: *Flowchart of the Information Selection and Assessment (ISA) framework. Figure from (Giuliani et al., 2015)*

4.2 Methods and Tools

In this section, we provide an overview of the algorithms that were used to perform the main steps of the framework illustrated in figure 4.1.

4.2.1 Deterministic Dynamic Programming (DDP)

The perfect operating policy is designed by assuming the sequence ε_{t+1} of future disturbance known over the entire time horizon and solving problem 4.1 via deterministic dynamic programming (DDP)(*Bellmann*, 1957). DDP recursively computes the optimal cost-to-go defined as

$$H_t(x_t) = min_{u_t} \Psi_{\varepsilon_{t+1}} [\Phi[g_t(x_t, u_t, \varepsilon_{t+1}), H_{t+1}(x_{t+1})]] \tag{4.4}$$

where H_t is the optimal cost-to-go at time t with respect to the state x_t (i.e. the reservoir storage), Ψ (i.e. expected value) is a statistic used to filter the disturbance, Φ is the average operator for aggregation over time, g_t is the monthly production, u_t the release decision variable and ε the reservoir's inflow vector. Once the function $H^*(\cdot)$ is known, the optimal policy is completely defined as:

$$m_t^*(x_t) = argmin_{u_t} \Psi_{\varepsilon_{t+1}} [\Phi[g_t(x_t, u_t, \varepsilon_{t+1}), H_{t+1}(x_{t+1})]] \tag{4.5}$$

Due to its recursive nature, this method for computing the optimal cost-to-go (and thus determining the optimal policy) can be applied only if the objectives are expressed by separable cost functionals. This is why the design indicators must be separable as well (*Soncini-Sessa et al.*, 2007).

4.2.2 Evolutionary MultiObjective Direct Policy Search (EMODPS)

Optimal management policies for water reservoir operation are usually designed using Stochastic Dynamic Programming (SDP). However, the adoption of SDP in complex real-world problems is challenged by three major limitations, known as curse of dimensionality, curse of modeling, and curse of multiple objectives. The curse of dimensionality, first introduced by *Bellmann* (1957) means that computational cost grows exponentially with state, control and disturbance dimension. For this reason SDP would be inapplicable when the dimensionality of the system exceeds two or three storages (*Loucks et al.*, 2005). The curse of modelling, introduced by *Tsitsiklis and Van Roy* (1996), means that exogenous information can be considered only as a stochastic disturbance, independent in time, with and associated probability density function, or they must be dynamically modeled (i.e. including additional state variables), thus

adding to the curse of dimensionality. The curse of multiple objectives indicates that computational cost has a factorial grow rate with the number of objectives considered (*Powell*, 2007), since SDP is a single objective approach, and it must be reiteratively used to solve multi-objective problems.
The Evolutionary MultiObjective Direct Policy Search (EMODPS) approach was introduced to alleviate the restrictions of the three main curses of SDP by combining Direct Policy Search (DPS), nonlinear approximating networks, and multiobjective evolutionary algorithms to design Pareto-approximate closed-loop operating policies for multipurpose water reservoirs (*Giuliani et al.*, 2016a). DPS (*Rosenstein and Barto*, 2001) is a simulation based approach, which assumes that the operating rule belongs to a given family of parameterized functions and searches the optimal solution in the policy parameters space. Nonlinear approximating networks, such as Artificial Neural Networks (ANNs) and Gaussian Radial Basin Functions (RBFs), are often selected since they are able to deal with high system's complexity, to provide more flexible structures, and to avoid restricting the search for the optimal policy to a subspace of the decision space that could not include the optimal solution. As demonstrated in *Giuliani et al.* (2016a), RBFs outperform ANNs when dealing with water resources system because they have a smoother shape, a limited domain of parameters, and allow coordination mechanisms in reservoir networks. In the case of using RBFs to parameterize the policy, the release decision u_t is defined as

$$u_t = \alpha + \sum_{i=1}^{N} \omega_i \varphi_i(I_t) \tag{4.6}$$

where N is the number of RBFs $\varphi_i(\cdot)$ and ω_i the weight of the i-th RBF. The single RBF $\varphi_i(\cdot)$ is defined as follow:

$$\varphi_i(I_t) = \exp\left[-\sum_{j=1}^{M} \frac{((I_t)_j - c_{j,i})^2}{b_{j,i}^2}\right] \tag{4.7}$$

where M is the number of input variables I_t and c_i ,b_i are the M-dimensional center and radius vectors of the i-th RBF. According to *Busoniu et al.* (2010) the centers of the RBF must lie within the bounded input space $c \in$[-1; 1] and the radii must strictly be positive $b \in$[0; 1].
The parameter vector can so be defined as

$$\theta = [\alpha, c_{i,j}, b_{i,j}, \omega_i] \tag{4.8}$$

and the numbers of policy parameters is:

$$n_\theta = N(2M + n_u) + 1 \tag{4.9}$$

When DPS problems involve multiple objectives, they can be coupled with multiobjective optimization methods, such as MultiObjective Evolutionary algorithms (MOEAs). MOEAs evolve a Pareto-approximate set of solutions by mimicking the randomized mating, selection, and mutation operations that occur in nature (*Coello et al.*, 2007). The algorithm works as follows: given a dataset *D*, a model structure *M*, an objective (fitness) function *J(θ, D, M)* at first the MOEA randomly generates a population of P individuals ($\theta^1, \theta^2, ..., \theta^P$) and for each individual in the population (i=1,...P), it computes the objective function $J(\theta^i, D, M)$. The algorithm then selects a subset of individuals based on their fitness value (using the Pareto-optimality principle) calculated in the previous step. Once the best individuals are selected, a new population has to be generated, applying to the best individuals genetic operators, such as crossover, mutation and replacement. The cross-over generates new offspring by random combination of the selected individuals, while mutation randomly modifies some offspring components. The replacement operator replaces some (or all) the population individuals by the offspring to obtain a new population. The new generation is then evaluated and the previous steps are repeated until a termination condition is achieved and the algorithm returns the population that minimize the objectives values (figure 4.3).

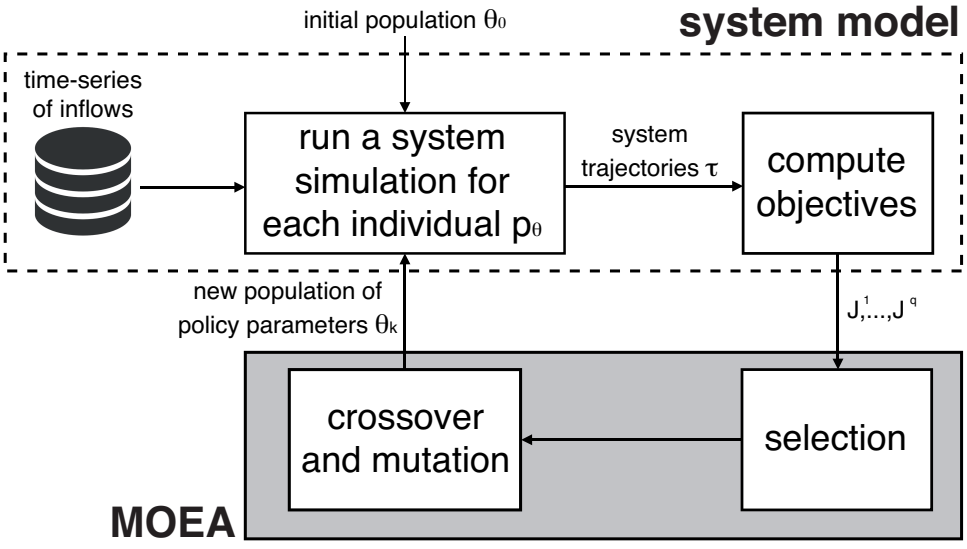

Figure 4.3: *Schematization of the evolutionary multiobjective direct policy search (EMODPS) approach; dashed line represents the model of the system and the gray box represents the MOEA algorithm. Figure from (Giuliani et al., 2016a)*

The EMODPS employs the Borg MOEA which has been demonstrated to meet or exceed the performance of other state-of-the- art MOEAs (*Salazar et al.*, 2016). The Borg MOEA assimilates several design principles from existing MOEAs and introduces several novel components (*Hadka and Reed*, 2013), which include:

- **multiple recombination operators** to enhance search in a wide assortment of problem domains. The Borg MOEA is not a single algorithm but a class of algorithms whose genetic operators are adaptively selected based on the problem and the decision variable (*Hadka and Reed*, 2013), adapting well to real-world problems where information about the most performing operator is unknown a priori.
- ε**-box dominance archive** for maintaining convergence and diversity throughout search. The Borg MOEA divides the objective space into hyperboxes with a side length ε, and if two or more solutions reside in the same hyperbox, the algorithm keeps only the non-dominated one, adding it to the archive.
- ε**-progress** which is a computationally efficient measure of search progression and stagnation. The Borg MOEA verifies that the optimization periodically produces at least one solution whose improvement exceeds the minimum threshold ε to avoid stagnation; if this condition is not met, appropriate actions are taken to either revive search, through the restart mechanism, or terminate the algorithm (*Hadka and Reed*, 2013).

4.2.3 Iterative input selection algorithm (IIS)

Input variable selection (IVS) is an important issue associated with the development of several hydrological applications. Selecting irrelevant, redundant inputs from a large set of candidates to characterize a preselected output can greatly influence model accuracy and add unnecessary model complexity influencing its reliability. IVS methods can be distinguished between model-based (or wrapper) and model-free (or filter) approaches. Each of these methods has different benefits and disadvantages.
The model-based approach relies on the idea of calibrating and validating a number of models with different sets of inputs and to select the set that ensures the best model performance. Generally, they could achieve better performance since they are tuned to the specific interactions between the model class and the data. However, computational costs are very high: a large number of calibration and validation processes must be performed to select the best combination

of inputs and so the method does not scale well to a large data set (*Kwak and Choi*, 2002);(*Chow and Huang*, 2005). Moreover, the optimal set of inputs obtained with a particular model is not guaranteed for other model with different class and architecture (*Maier et al.*, 2010).
On the other hand, in the model-free algorithms, the variable selection is based on the information content of the candidate input data set, as measured by interclass distance, statistical dependence, or information-theoretic measure. This approach has a strong computational efficiency; however, the significance measure is generally monotonic and, thus, without a predefined cutoff criterion, the algorithm tends to select very large subsets of input variables, with high risk of redundancy. Given the high number of candidate input variables in these problems, model-free methods are generally preferred over model-based approaches (*Maier et al.*, 2010).
Galelli and Castelletti (2013) developed a new method named tree-based iterative input variable selection algorithm (IIS) that incorporates some of the features of model-based approaches into a fast model-free method in order to handle very large candidate input sets. The IIS algorithm is capable of selecting the most significant and nonredundant inputs with strong ability in characterizing nonlinear relationships, computational efficiency, and scalability with respect to input dimensionality.
The IIS algorithm is composed of the following steps (figure 4.4)

1. Given the sample data set D and n candidate inputs, the IIS algorithm runs an input ranking (IR) algorithm to sort the n candidate inputs according to a nonlinear statistical measure of significance (e.g., the explained variance). In principle, the first variable in the ranking should be the most significant in explaining the output. In practice, in the presence of several potentially significant, but redundant, inputs, their contribution to the output explanation is equally partitioned and they might not be listed in the very top positions. To reduce the risk for mis-selection, the first p variables $x^j \in X^p \subseteq X^i$ j=1, ...,p in the ranking are individually evaluated in the following step.

2. The relative significance of the first p-ranked variables is assessed against the observed output and p single-input-single-output (SISO) models $f^j(\cdot)$ j=1, ...,p, are identified with an appropriate model building (MB) algorithm and compared in terms of a suitable distance metric (e.g., coefficient of determination, mean-squared error, etc.) between the output y and each SISO model prediction $f^j(x^j)$. The best performing input x^* among the p considered is added to the X_y^i set of the variables selected to explain the y.

3. A MB algorithm is then run to identify a multi-input-single output (MISO) model $m(\cdot)$ mapping the input variables X_y^i so far selected into the output y.

The procedure is repeated using the residuals between (y- $m(X_y^i)$) the output of the model and the observed data as the new output variable in steps 1 and 2, and these operations are iterated until either the best variable returned by the IR algorithm is already present in the input selected set X_y^i or the performance of the underlying model, measured with the metric D, does not significantly improve. The reevaluation of the ranking on the model residuals every time a candidate variable is selected ensures that all the candidates that are highly correlated with the selected variable, and thus may become useless, are discarded (*Galelli and Castelletti*, 2013).

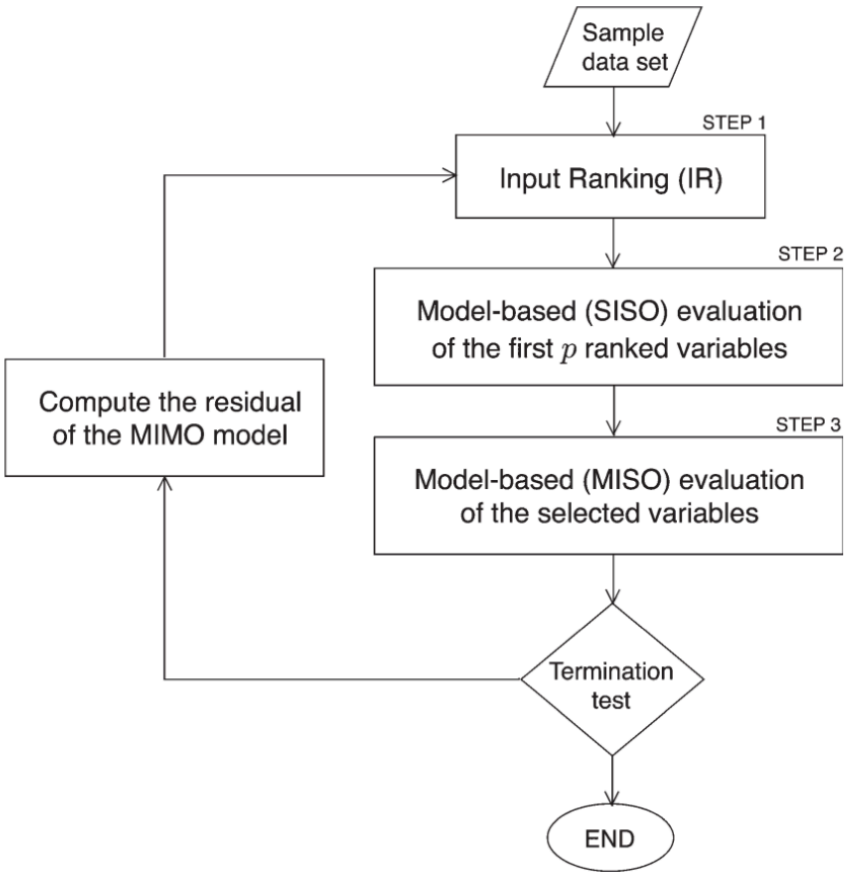

Figure 4.4: *Flowchart of the Iterative Input Variable Selection (IIS) algorithm (Galelli and Castelletti, 2013).*

The IIS algorithm is combined with Extremely Randomized Trees (*Geurts et al.*, 2006) as regression model that describes the relationship between input and output variables. Tree-based regressors are characterized by a tree-like structures, composed of decision nodes, branches, and leaves, which form a cascade of rules leading to numerical values. The tree is obtained by first partitioning at the top decision node, with a proper splitting criterion, the set of the input variables into two subsets, thus creating two branches. The splitting process is then repeated in a recursive way on each derived subset, until some termination criterion is met. When this process is over, the tree branches represent the hierarchical structure of the subset partitions, while the leaves are the smallest subsets associated to the terminal branches. Each leaf is finally labeled with a numerical value.
There are three parameters M, K, and n_{min} that characterize the model building algorithm and diversely affect the ensemble performance and overall method efficiency. M defines the number of trees that compose a forest. This parameter controls the reduction of variance aggregate model. High values of M reduce the variance of the final estimate, but they also considerably increase the calculation time. K identifies the number of regression input chosen randomly at each node and it controls the level of randomness in the tree building process. Its value can be chosen from 1 to n, with n being the number of input variables. The smaller K, the stronger the randomization of the trees and the weaker the dependence of their structure on the values of the output variable in the training dataset. In the extreme case, when K is equal to 1, the splits (cut-directions and cut-points) are chosen in a totally independent way of the output variable and the method builds totally randomized trees. Finally, n_{min} parameter specifies the minimum sample size for splitting a node. The threshold n_{min} is used to balance bias and variance reduction. Large values of n_{min} lead to small trees, with high bias and small variance; conversely, low values of n_{min} lead to fully-grown trees, which may overfit the data. The optimal tuning of n_{min} can depend on the level of noise in the training dataset: the noisier the outputs, the higher the optimal value of n_{min} should be. A value of n_{min} between 5 and 50 represents a robust choice in a broad range of typical conditions (*Geurts et al.*, 2006).

4.2.4 Synthetic forecast generation

Synthetic forecast ensembles have increasingly been used in the context of long-term planning exercises and multidecadal climate projections (*Herman et al.*, 2015; *Kasprzyk et al.*, 2013; *Steinschneider et al.*, 2015). As an example, *Nayak et al.* (2018) developed synthetic, short-term meteorological forecasts for a past period when forecasts were not available. Synthetic precipitation and temperature forecasts were then converted into streamflow value by adopting a run-off model. *Dumont Goulart* (2019) extended *Nayak et al.* (2018) work by directly generating a synthetic inflow forecast into future periods and assessing forecast value under different climate change scenarios. In this thesis, we adopted the same method to produce synthetic traces of an existing forecast model, by using the K-Nearest Neighbor algorithm (KNN).
The model requires two datasets, one is the simulation dataset corresponding to the projected inflows, and the second is the training dataset which includes historical observations that have corresponding forecast data. The algorithm measures the Euclidean distance between each inflow value in the simulation dataset, Q_t, and all the inflows from the training dataset observed in the same month as Q_t. Only the K closest values are then considered when selecting the best match Q_t*, and these values do not have equal weight, as they follow a kernel density function. The closer to the observed data a value is, the higher its weight is and the more probable it is to be chosen. Moreover, for each observation q_t in the training dataset, the forecast error can be computed as :

$$\hat{\varepsilon}_{q,t} = \log \hat{q}_t - \log q_t \tag{4.10}$$

where $\hat{\varepsilon}_{q,t}$ is the error originated by the difference between the observed q_t and the forecasted $\hat{q}_t$ inflows. The formulation in eq. 4.10 assumes an additive residual in logarithmic form based on the results obtained by *Dumont Goulart* (2019). Then, the model generates the synthetic forecast for the simulation period by using the forecast errors of the observations in the training data associated by the KNN algorithm.
It is highly important to remind that this approach implicitly assumes that the forecast error residual propagation of the synthetic forecast is stationary over time.

4.3 Experiment setting

The assessment of the forecast value under historical hydrologic conditions was performed over the period 1993-2014 for which the SMHI forecasts are available, while we focus on the period 2071-2100 for the climate change analysis.

The baseline and informed operating policies for both historical and future conditions are solved by the EMODPS method. The operating policies are defined as Gaussian Radial Basis Functions (RBFs). For baseline operating policies, the RBFs were defined using 4 basis (N), 3 inputs (M) and 1 output (n_u), for a total of 29 parameters. Conversely, for the informed operating policies, the RBFs were defined using 6 basis (N), 5 inputs (M) and 1 output (n_u), for a total of 67 parameters (equation 4.9). The policy parameters are optimized using the self-adaptive Borg MOEA (*Hadka and Reed*, 2013), by setting the Number of Function Evaluation (NFE) equal to 1,000,000 and by considering 10 repetitions (NSEEDS) of the optimization process to filter the uncertainty caused by the random operators of the Borg MOEA.

In the selection of the most informative lead time, perfect streamflow forecasts with different lead times were considered as input variables in the IIS algorithm. Perfect forecasts of cumulative inflows from 1965 to 2016 were computed over seven different lead times, ranging from 1 month to 7 months ahead. In addition to the cumulative future streamflow, both the minimum and maximum from 2 months to 7 months are also included to account for the extreme events. The 19 considered input variables are summarized in table 4.1. We consider as output variable the difference between the sequence of reservoir releases derived under the perfect operating policy and the baseline operating policy obtained under historical hydrological conditions. IIS algorithm parameters were set as: numer of trees in the forest M=500, minimum sample size for splitting a node n_{nmin}= 20, number of regression input K= 19 and number of IIS algorithm runs n_{run}=15. Further analysis justifying the use of the parameters is reported in Appendix A.2.

Table 4.1: *Set of perfect seasonal inflow forecasts over different lead times.*

Name	Description
Lead 1, ..., 7	Cumulative future inflow over 1,. . . ,7 months
Leadmin 2, ..., 7	Minimum future inflow over 2,. . . ,7 months
Leadmax 2, ..., 7	Maximum future inflow over 2,. . . ,7 months

5

Results

In this chapter the numerical results obtained are presented for the historical and future conditions by adopting the Information Selection Assessment framework. A final comparison of the forecast value under these two time periods is also provided.

5.1 Historical conditions

This section reports the results obtained following the three blocks of the ISA framework for assessing the value of seasonal streamflow forecasts in informing the GERD operation during the historical time period 1993-2014.

The first step aims at quantifying the expected value of perfect information (EVPI) as the gap between the Baseline and Perfect Solutions, which provides a measure of the maximum improvement in the system performance. The obtained value of EVPI is reported in table 5.1: the ideal solution designed under the assumption of perfect knowledge of future inflows allows gaining 283.41 GWh/y in the annual hydropower production with respect to the Baseline Solution, which corresponds to a 1.73% improvement. The maximum space for increasing the hydropower production of the GERD baseline operations not informed by any forecast is relatively small, making the use of additional information like streamflow forecast non promising. After this first result, we focused on the system's performance during the wet season from June to September (Kiremt season), when most of hydropower production is allocated be-

cause of high inflow values. This season, as reported in section 2.5, was recommended by state-of-art researches for seasonal forecasting due to the correlation to ENSO phenomenon. Figure 5.1 (a) illustrates the monthly inflow boxplots over the period 1965-2017, along with the corresponding cyclostationary mean (red line), and shows the strong hydrologic variability existing in the wet season (Kiremt season) of the basin. In addition, figure 5.1 (b) reports the cyclostationary mean of the hydropower production obtained under the perfect (black line) and baseline (blue line) operating policies over the period 1965-2017, showing that most of hydropower production occurs during the June-September season. However, the perfect operating policy keeps the hydropower production higher during the wet season, while the baseline operating policy produces more energy during the dry season from November to May. Table 5.1 summarizes the system's performance improvement between Perfect and Baseline Solutions comparing the annual and June-September (wet season) hydropower production and the corresponding absolute gain and relative gain. The performance difference between Perfect and Baseline Solutions evaluated only in the Kiremt season is equal to 18.25%, suggesting that forecasts could have a great value in this particular time of the year.

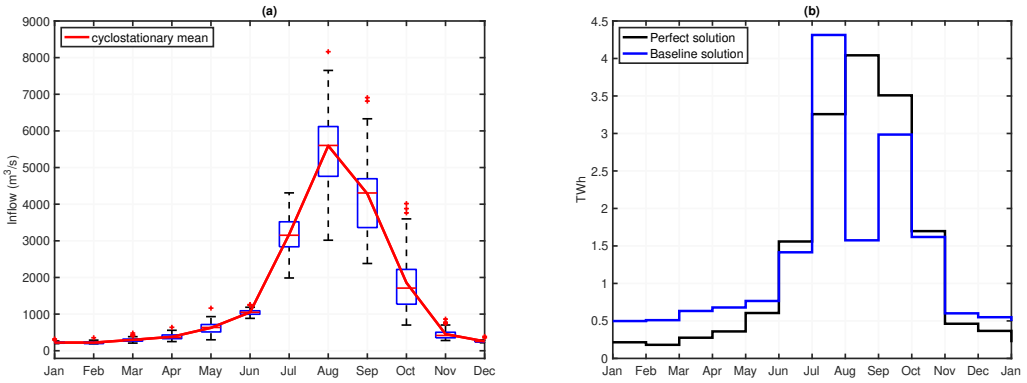

Figure 5.1: *(a) Cyclostationary mean and boxplots of monthly inflow; (b) Cyclostationary mean of the hydropower production obtained under the perfect (black) and baseline (blue) operating policy. Time period: 1965-2017.*

Table 5.1: *Expected value of perfect information: the performance improvement between the perfect and baseline solution considering annual hydropower production and June-September (Kiremt season) hydropower production.*

Policy	Hydropower production	Absolute gain	Relative gain
Baseline solution	16.430 TWh/y	-	-
Perfect solution	16.714 TWh/y	+283.41 GWh/y	+1.73%
Baseline solution Kiremt season	10.761 TWh	-	-
Perfect solution Kiremt season	12.725 TWh	+1963.64 GWh	+18.25%

In the second step of the ISA framework, the selection of the most informative lead time was performed via IIS algorithm. Figure 5.2 reports the results of 15 runs of the algorithm, where the repetition of the experiments aims at filtering the randomness associated to the construction of the extra-trees models. In this parallel axes representation, each variable, corresponding to a candidate forecast with different lead times (see Table 4.1), is represented as a coloured line crossing the three axes at the values of the related performance in terms of frequency of selection, average position, and average relative contribution expressed as coefficient of determination R^2. The reported performance is normalized between their minimum and maximum values and the axes are oriented so that the direction of preference is always upward. Consequently, the most relevant forecast lead time would be represented by a horizontal line running along the top of all the axes, meaning a variable that is selected with high frequency, in the first position, and with the largest relative contribution (*Giuliani et al.*, 2015). Figure 5.2 (a) identifies variables *Lead1* and *Leadmin3* (solid lines), namely inflow value over one month and minimum over three months, as the first and second most relevant forecast, respectively. This result shows that inflow volumes can be more important in the first month, while extremes can prevail at longer time horizons.
In the following analysis, the two selected lead times are considered together to design GERD informed operating policy. Indeed, the cumulated performance of the regression model for 12^{th} run (see figure 5.2 b) shows that the two variables indipendently account for almost an equal R^2 contribution of 38.6% and 39.51%, respectively, leading to a high cumulated coefficient of variation (R^2) of 78.11%.

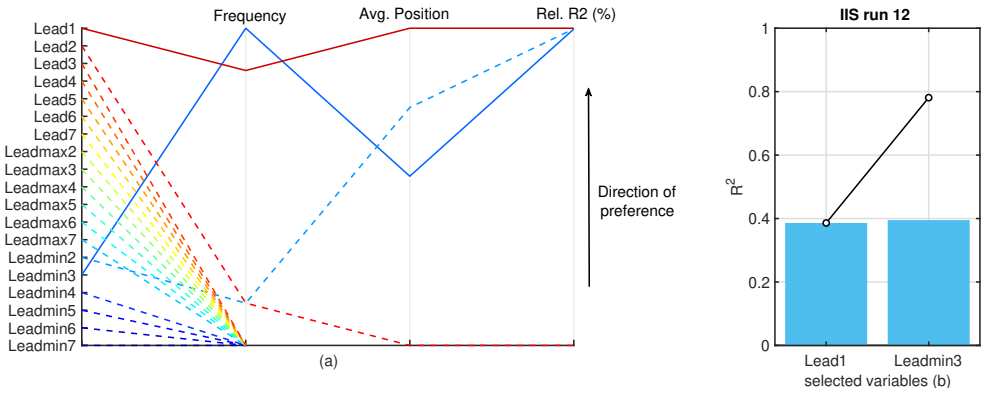

Figure 5.2: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=20: (a) for each lead times, the frequency of selection, the average position, and the average relative contribution in terms of coefficient of determination R^2; (b) Cumulated performance of the regression model for 12^{th} run.*

In the variable selection, we considered as input candidates, perfect forecasts built from historical observational data. However, in order to reproduce a realistic decision making environment, we then test the value of the seasonal forecasts provided by the SMHI. Figure 5.3 illustrates the scatterplot between observed and forecasted inflow over the two selected lead times, showing no particular bias with a Pearson correlation coefficient equal to 0.785 and 0.85 for *Lead1* and *Leadmin3*, respectively. Figures 5.4 (a), (b) compare the trajectory of observed (black line) and forecasted (red line) inflow for *Lead1* and *Leadmin3*. In general, SMHI forecasts tend to underestimate inflow values during the dry season (January-May), while the peak value (August) is overestimated for some years and underestimated for others. Figures 5.4 (c),(d) represent the cyclostationary mean of observed (black line) and forecasted inflow (red line) showing that the two trajectories have the same annual pattern and that forecasts on average underestimate observed inflows.

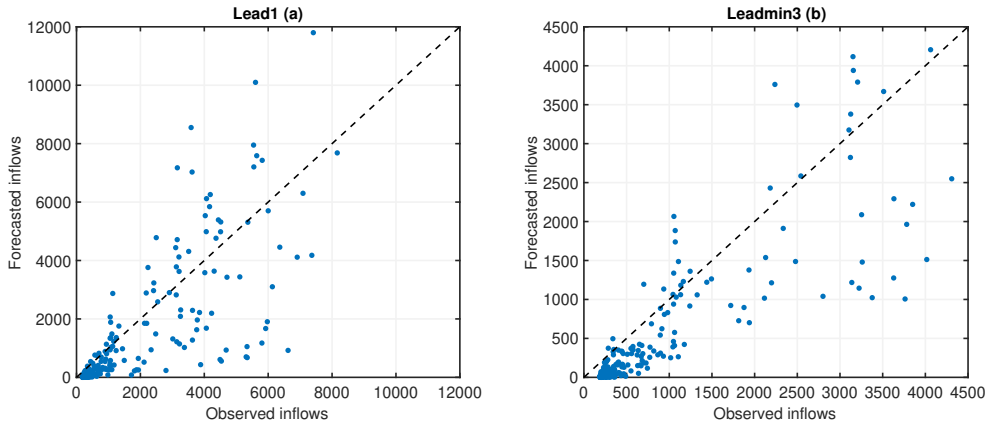

Figure 5.3: *(a) Scatterplot between observed and forecasted inflow over one month; (b) scatterplot between observed and forecasted minimum inflow over three months.*

The last step of the ISA framework quantifies the expected value of sample information as the performance difference between Baseline and Forecast Solutions. Forecasted inflows for *Lead1* and *Leadmin3* are used as inputs to inform the GERD operating policy. Results show that forecasts produce an absolute gain with respect to the baseline operation policy in the annual hydropower production of 80.57 GWh/y, which corresponds to a 0.49% improvement. As expected, this result confirms our first supposition that forecasts slightly improve the system performance if we evaluate the objective over all months of the year. We then restrict the analysis to the system performance during the wet June-September season (Kiremt season) and we obtained a relative gain between the Baseline and Forecast Solutions of 16.66%. This final result shows clearly that forecast information becomes mostly valuable during the wet sea-

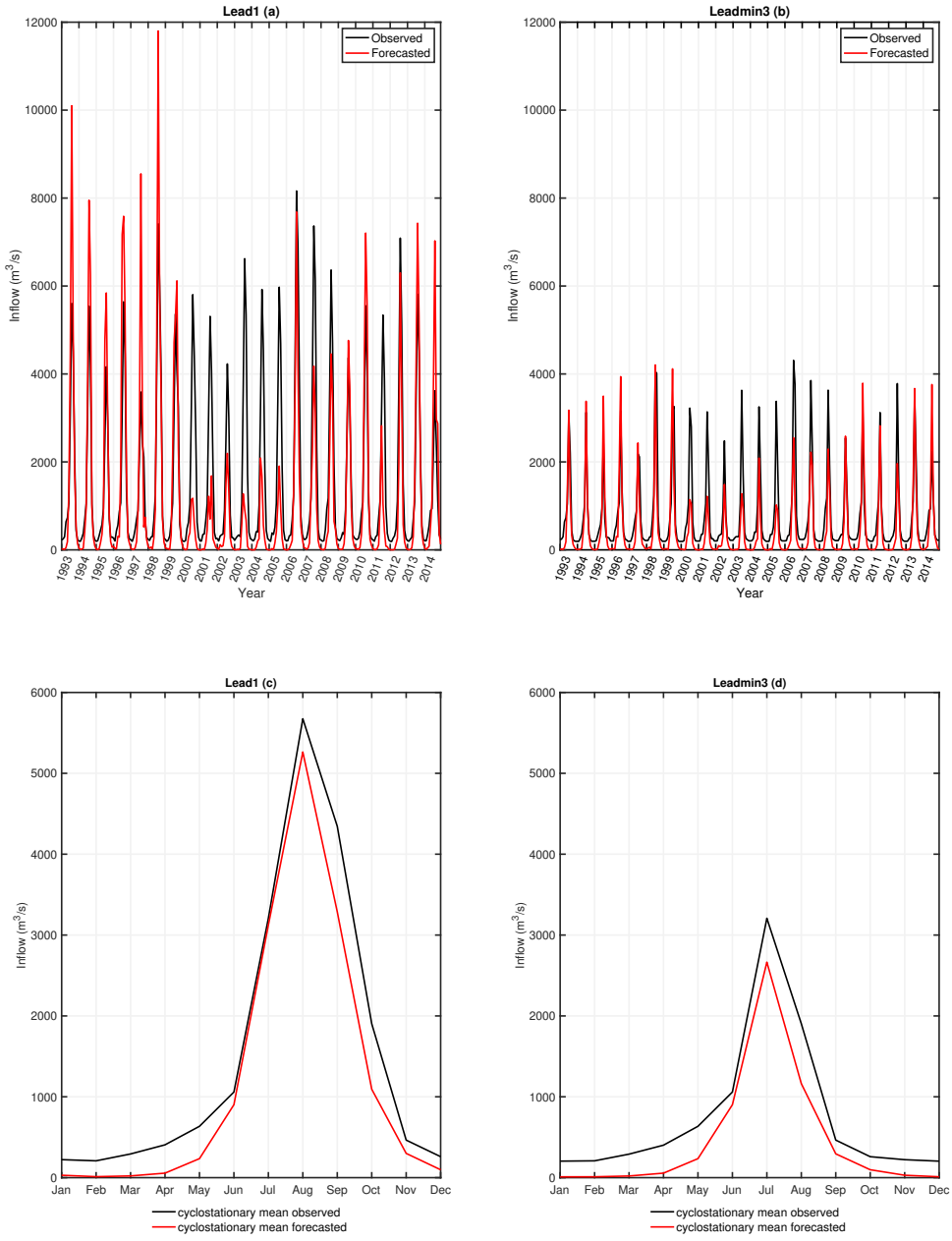

Figure 5.4: *(a) Trajectory of observed and forecasted inflow over one month; (b) Trajectory of observed and forecasted minimum inflow over three months; (c) Cyclostationary mean of observed and forecasted inflow over one month; (d) Cyclostationary mean of observed and forecasted minimum inflow over three months.*

son with gains almost near to Perfect Solution. Table 5.2 summarizes the overall results obtained while assesing the forecast value considering historical conditions (1993-2014).

Table 5.2: *The performance improvement between the Perfect and Baseline Solution and between Forecast and Baseline Solution considering the annual hydropower production and the June-September (Kiremt season) hydropower production.*

Policy	Hydropower production	Absolute gain	Relative gain
Baseline solution	16.430 TWh/y	-	-
Perfect solution	16.714 TWh/y	+283.41 GWh/y	+1.73%
Forecast solution	16.511 TWh/y	+80.57 GWh/y	+0.49%
Baseline solution Kiremt season	10.761 TWh	-	-
Perfect solution Kiremt season	12.725 TWh	+1963.64 GWh	+18.25%
Forecast solution Kiremt season	12.554 TWh	+1793.17 GWh	+16.66%

Additionally, we analyzed the dynamic behavior of the system under different operating policies conditioned over distinct information as reported in figure 5.5. An operating policy with a perfect knowledge of future inflows (black line) is able to maintain the highest level (approximately at 140 m) by storig water during the dry season (January-May) and releasing it in the wet season (June-September), thus producing more energy. Conversely, GERD operating policy informed with basic information (blue line) can not anticipate the high inflow variability occuring in the wet season, thus preferring to release more water during the dry season and leading to values of level that oscillate from 140 m to 135 m. Consequently, the hydropower production in the wet season is lower with respect to the perfect solution (black line). Forecasts (red line) are able to inform the policy about the incoming inflow despite the large hydrological variability, allowing to get higher values of hydropower production than the Baseline Solution, with the trajectory of reservoir level that remains closer to the one obtained with the Perfect Solution. Figure 5.6 illustrates the level, the release and the hydropower production along the entire historical time period 1993-2013. It's important to note in figure 5.6 (a) that for years 1997, 1998, 2009, and 2010 the Forecast Solution (red line) leads to lower levels than under the Baseline (blue) and Perfect (black) Solutions. This is due to the fact that, in these particular years, forecasts largely overestimate the peak of the actual inflow (figure 5.4(a),(b)).

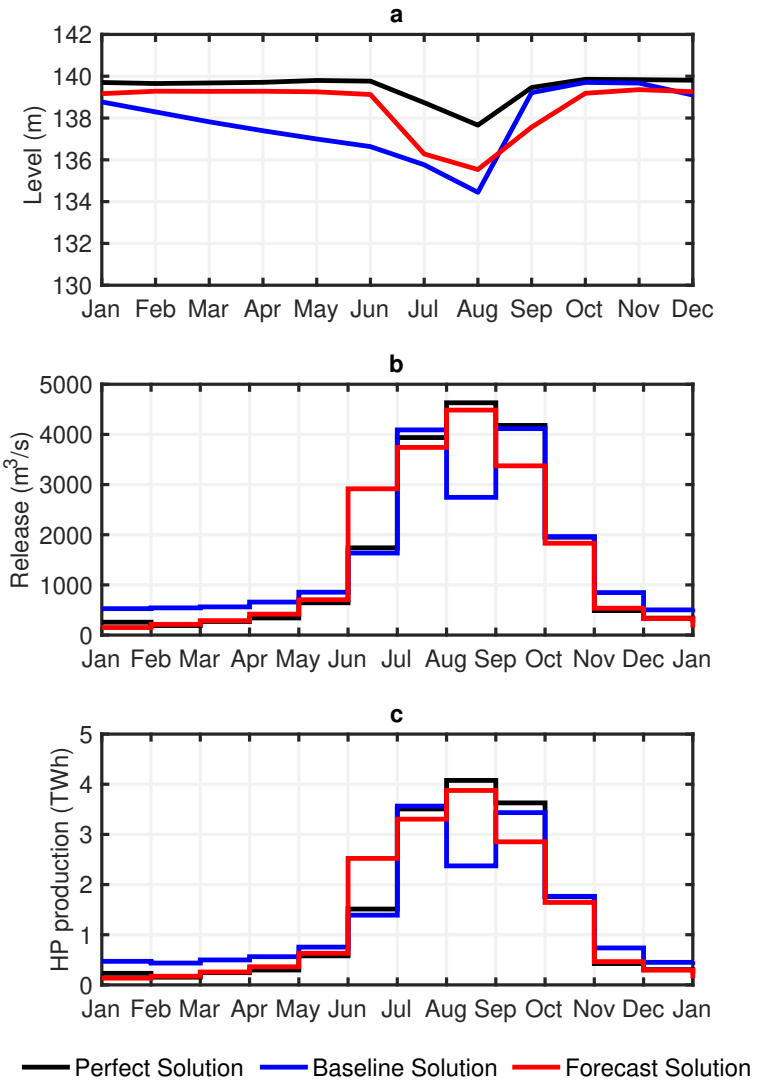

Figure 5.5: *System dynamics under the perfect (black line) baseline (blue) and forecast (red) operating policies: (a) cyclostationary mean of GERD level; (b) cyclostationary mean of GERD release; (b) cyclostationary mean of GERD hydropower production.*

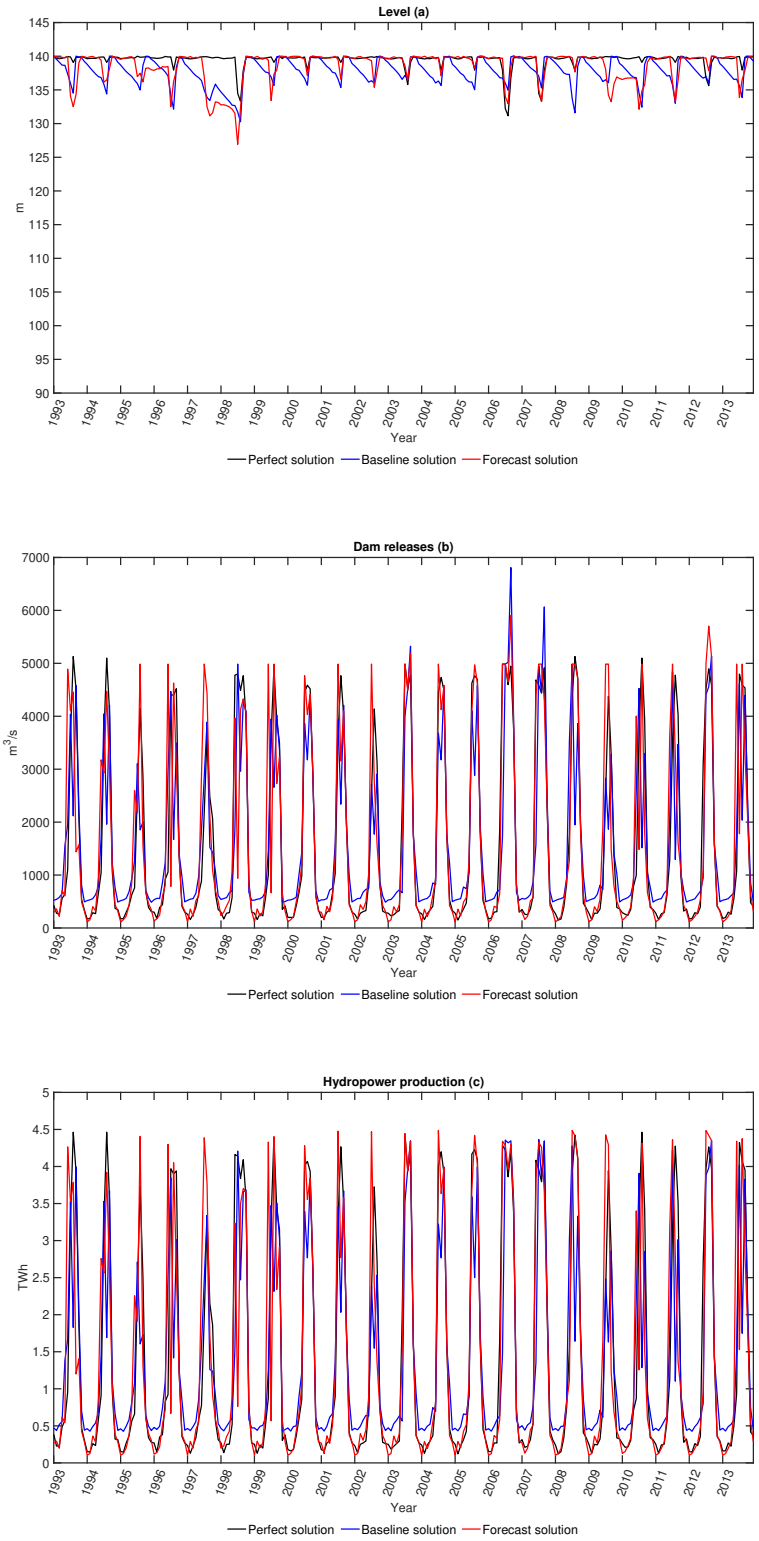

Figure 5.6: *The system dynamics under the baseline (blue line), perfect (black line), and forecast (red line) operating policies for historical time period 1993-2013. (a) Gerd's level; (b) Gerd's release; (c) Gerd's hydropower production.*

5.2 Future conditions

As explained in chapter 4, we adopted the same methodology to assess the value of seasonal forecasts under future climate scenarios. However, to design the perfect, baseline, and forecast operating policy, the trajectories of "observed" and "forecasted" inflows projected into future condition should be generated. We first report the analysis and results that led us to obtain these two trajectories and we then invastigate the forecast value.

5.2.1 Inflow projection

In order to compute the projected inflows, we need to understand which are the climate scenarios to be considered. We anaylized δ values defined as monthly mean value of expected percent change of water discharge (eq. 3.8) given by the different climate models listed in table 3.4 for the RCP 4.5 and the RCP 8.5 scenarios over the time periods 2011-2040, 2041-2070, 2071-2100 as shown in figure 5.7. We can notice that there are some models that return anomalous (high) values of δ. To identify the models producing non realistic scenarios, we compare the historical simulated inflows by each climate model with monthly boxplots of observed data. Results in figure 5.8 show that all models underestimate inflows for the first three months of the year (January-February-March) and there are also models that largely overestimate inflows during the wet season (June-September). Therefore, we removed climate models that produce inflows that are out of the range of historical hydrological variability, in particular, models number 4, 6, and 7, corresponding to models ACCESS1.3, IPSL-CM5A-LRI, IPSL-CM5A-MR 3.4 that simulate value of inflow outside the boxplot of the observations for at least 2 months during the wet season.
However, most of climate models also compute high values of δ in the first three months of the year (dry season). To further adjust these values, we first combined the remained climate models by computing the median value of δ for each month in order to filter possible additional spikes. Then, we calculated the projected inflows following equation 3.9 and, finally, we applied a correction factor to the first three months of the year definied as the ratio between simulated and observed inflows during the historical period. Moreover, we considered as future reference period the 2071-2100 since δ curves for the selected future horizon, as reported in figure 5.7, are quite smooth and also amplify the differences between the future scenarios and the historical period. Figure 5.9 compares the cyclostationary mean of projected inflows for the RCP 4.5 and the RCP 8.5 scenarios with the observed historical data. We can no-

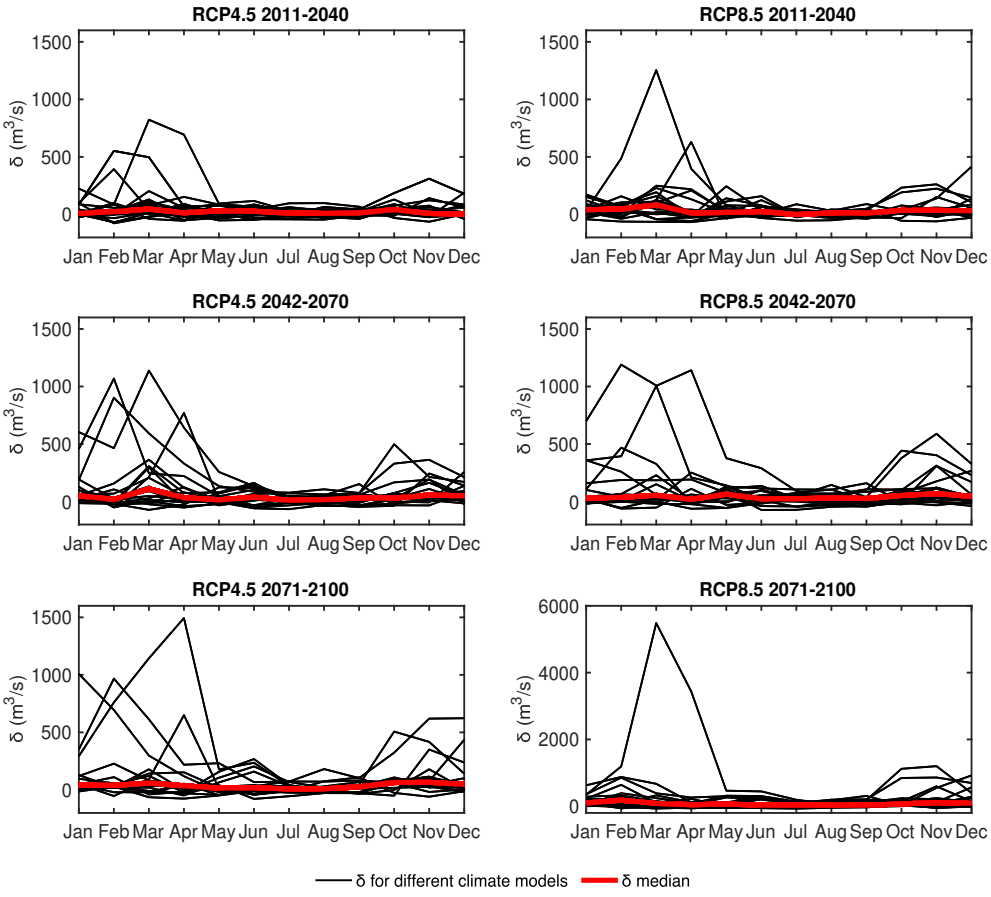

Figure 5.7: *δ monthly mean value of expected percent change of water discharge for each climate scenarios: black lines represent δ values computed by different climate models, red line indicates the median value for each month.*

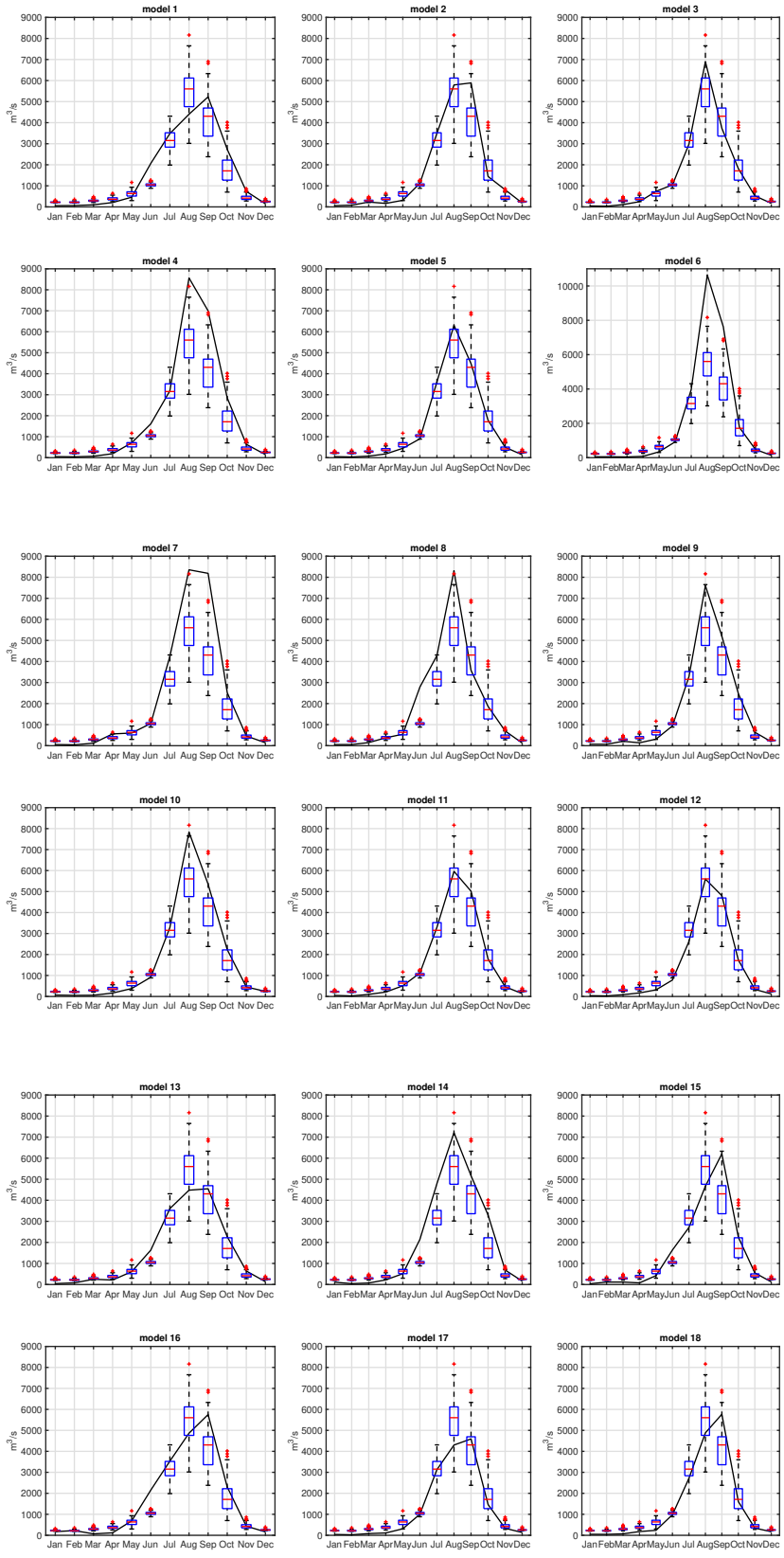

Figure 5.8: *Simulated historical inflows by each climate model compared with the boxplots of historical observed inflows.*

tice that future scenarios will be characterized by a reduction of inflows during the dry season (January-March) and an increase during the wet season (June-September); overall, the annual inflow (figure 5.10) increases with respect the historical period. These results seem reasonable since climate change is expected to accentuate extreme climate conditions (*Coffel et al.*, 2019).

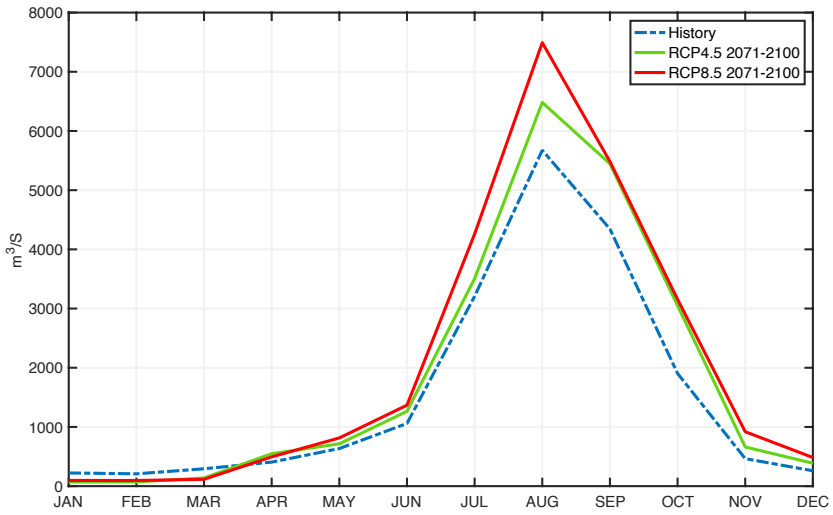

Figure 5.9: *Cyclostationary mean of projected inflows for the RCP 4.5 (green line) and the RCP 8.5 (red line) scenarios in period 2071-2100 compared with the cyclostationary mean of historical inflows (blue dashed line).*

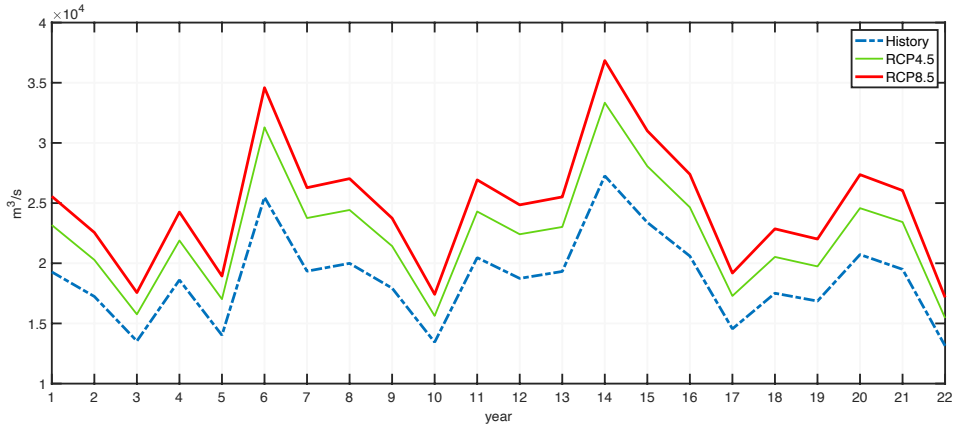

Figure 5.10: *Annual projected inflows for the RCP 4.5 (green line) and the RCP 8.5 (red line) scenarios in period 2071-2100 compared with historical annual inflow (blue dashed line) 1993-2014.*

5.2.2 Forecast projection

To assess the forecast value under future conditions, we also developed a synthetic generator model that mimics the existing SMHI forecasts and produces a forecast ensemble for the future scenarios. The KNN approach described in section 4.2.4 is used to generate a 25-member forecast ensemble working on the additive log-scale residuals. The algorithm compares forecasted inflows to the observed ones over the historical period in order to model the forecast errors and then propagates the forecast uncertainty to the projected inflow. It is worth reminding that this approach assumes stationarity in the forecast skill, assuming that forecasts in the future will be characterized by the same level of accuracy (i.e., same residuals) shown in the historical period. The model is calibrated over the training period 1993-2014, validated over the 1965-2016 and then applied to each scenario of projected inflows for the time period 2071-2092. Two type of forecast lead times are considered, namely one month ahead (*Lead1*) and the minimum over three months (*Leadmin3*) based on the results obtained in section 5.1. Figure 5.11 illustrates a scatterplot between observed and forecasted inflows, including both the real forecasts which are available only over the time period 1993-2014 and the synthetic one generated for the historical period 1965-2016. The figure shows that the synthetic forecasts successfully reproduce the skill of the real forecasts. Figure 5.12 gives further evidence to the quality of the synthetic results by showing that the synthetic forecasts are again similar to the observed ones in terms of residual-observation relationship. Both sets of data are distributed according to a triangular shape with the majority of the errors being close to zero. The histograms presented in figure 5.13 also show some similarities between real and synthetic forecasts. Figures 5.14 and 5.15 report scatterplots between observed and forecasted inflows, including both real forecasts for historical period 1993-2014 (black triangles) and the synthetic forecasts generated for the future climate scenarios in period 2071-2092 (blue circles). In the future scenarios synthetic forecasted inflows deviate more from the historical data and the peak values are amplified due to inflow increase.

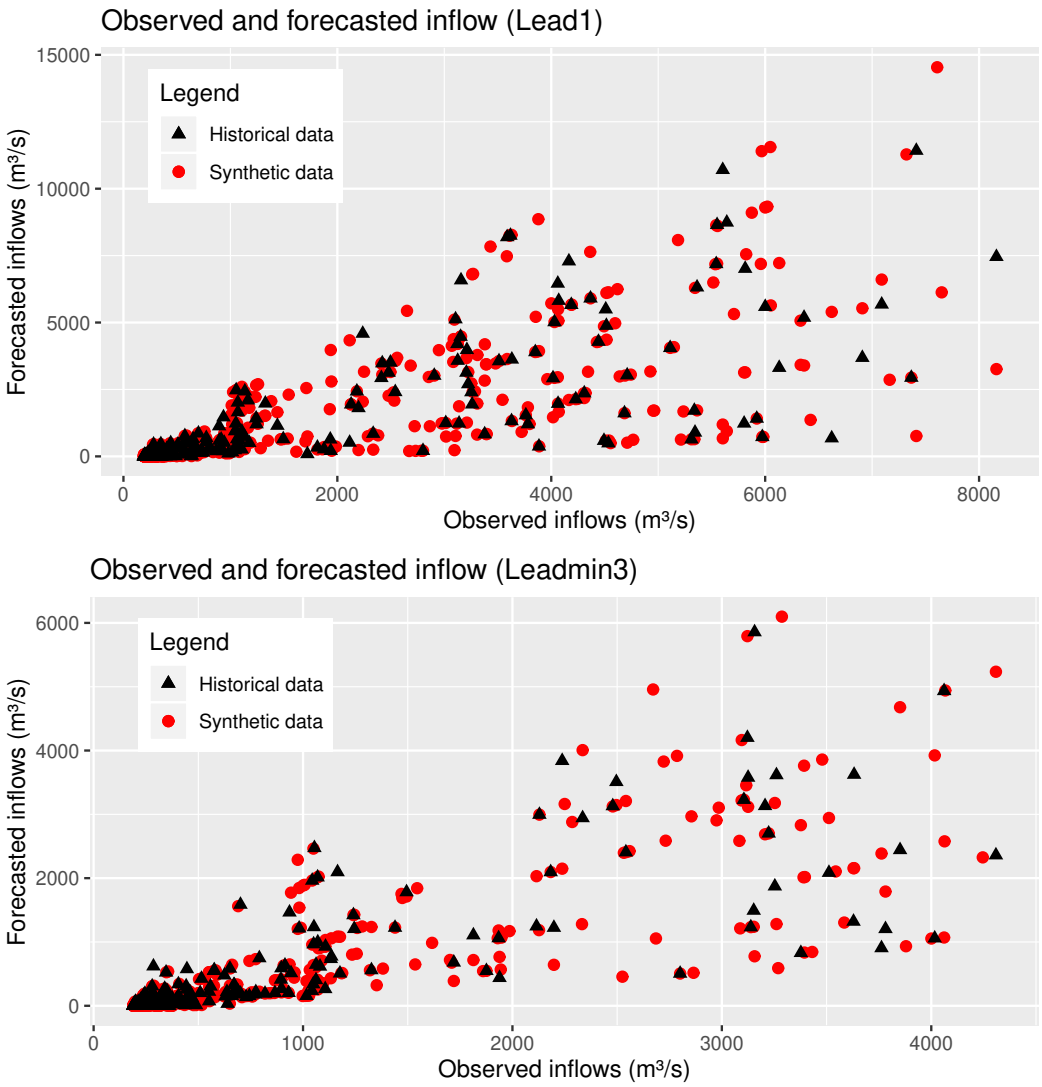

Figure 5.11: *Scatterplot presenting the forecasted inflow based on the observed inflow over one month Lead1 (top panel) and minimum inflow over 3 months Leadmin3 (low panel). Black triangles represent the historical data (1993-2014) and the red circles the synthetic forecast generated for the time period (1965-2016).*

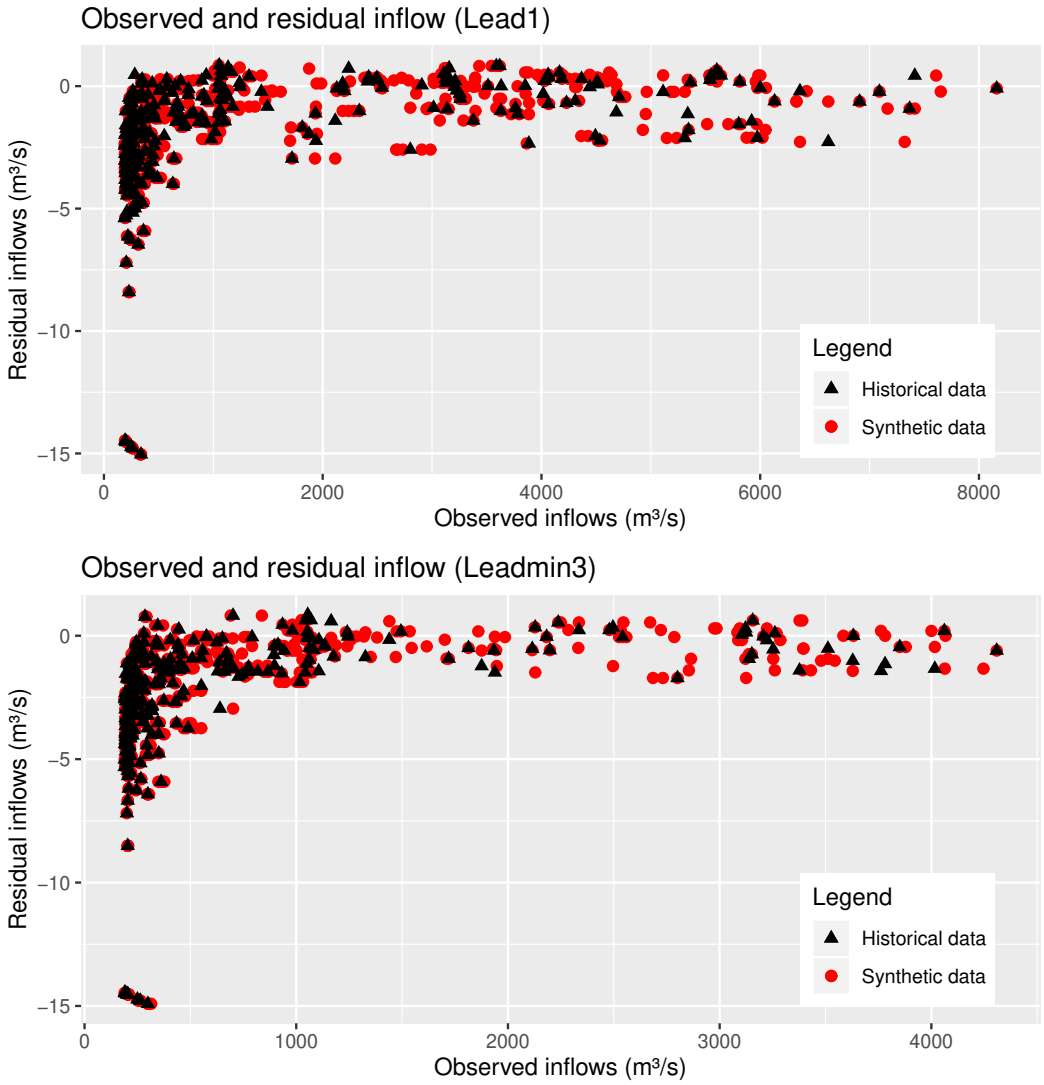

Figure 5.12: *Relationship between the residual inflow and the observed inflow over one month Lead1 (top panel) and minimum inflow over 3 months Leadmin3 (low panel). Black triangles represent the historical data (1993-2014) and the red circles the synthetic forecast generated for the time period (1965-2016).*

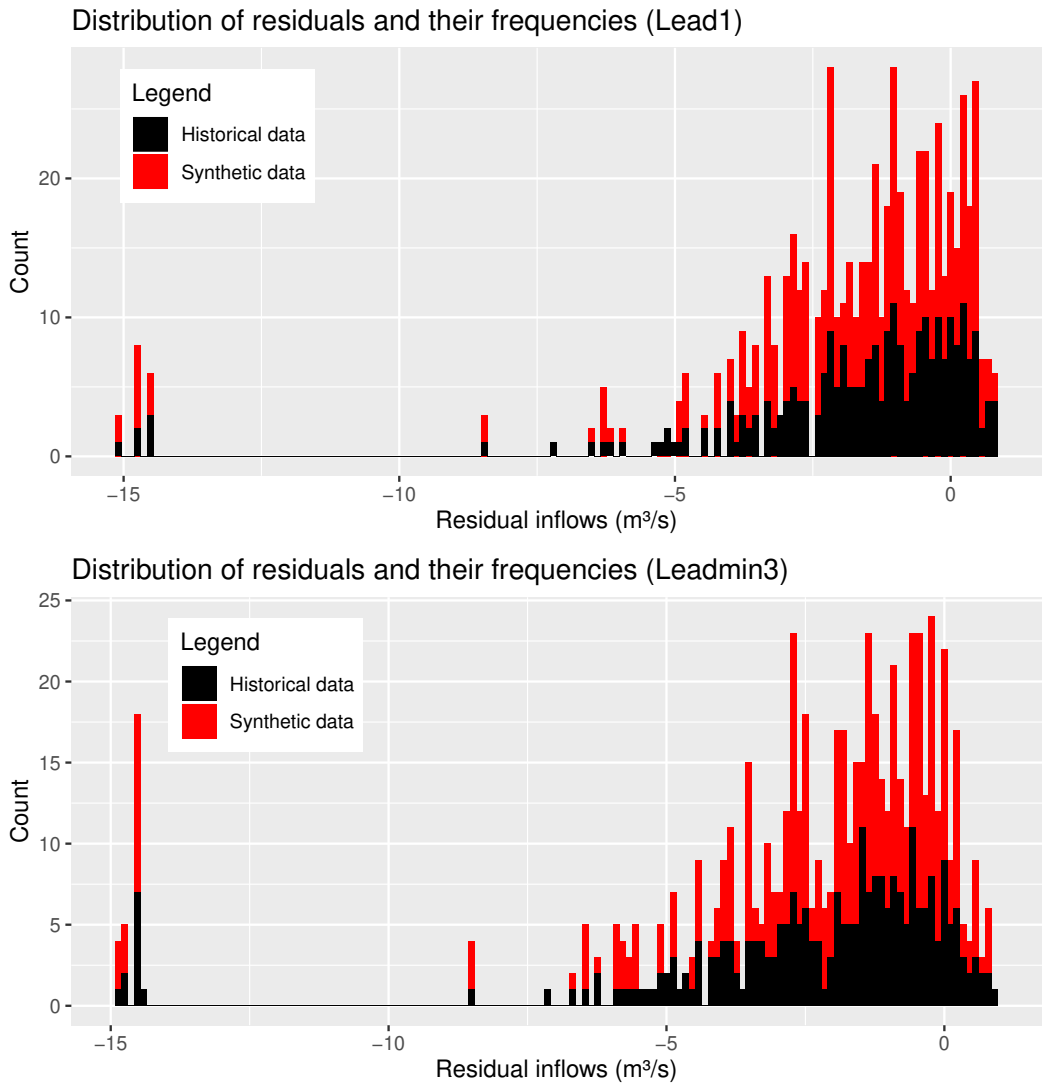

Figure 5.13: *Histogram with magnitude and frequencies of the residuals for both historical (1993-2014) and synthetic cases (1965-2016). Black represents the historical and red the synthetic generated inflow over one month Lead1 (top panel) and minimum inflow over three months Leadmin3 (low panel).*

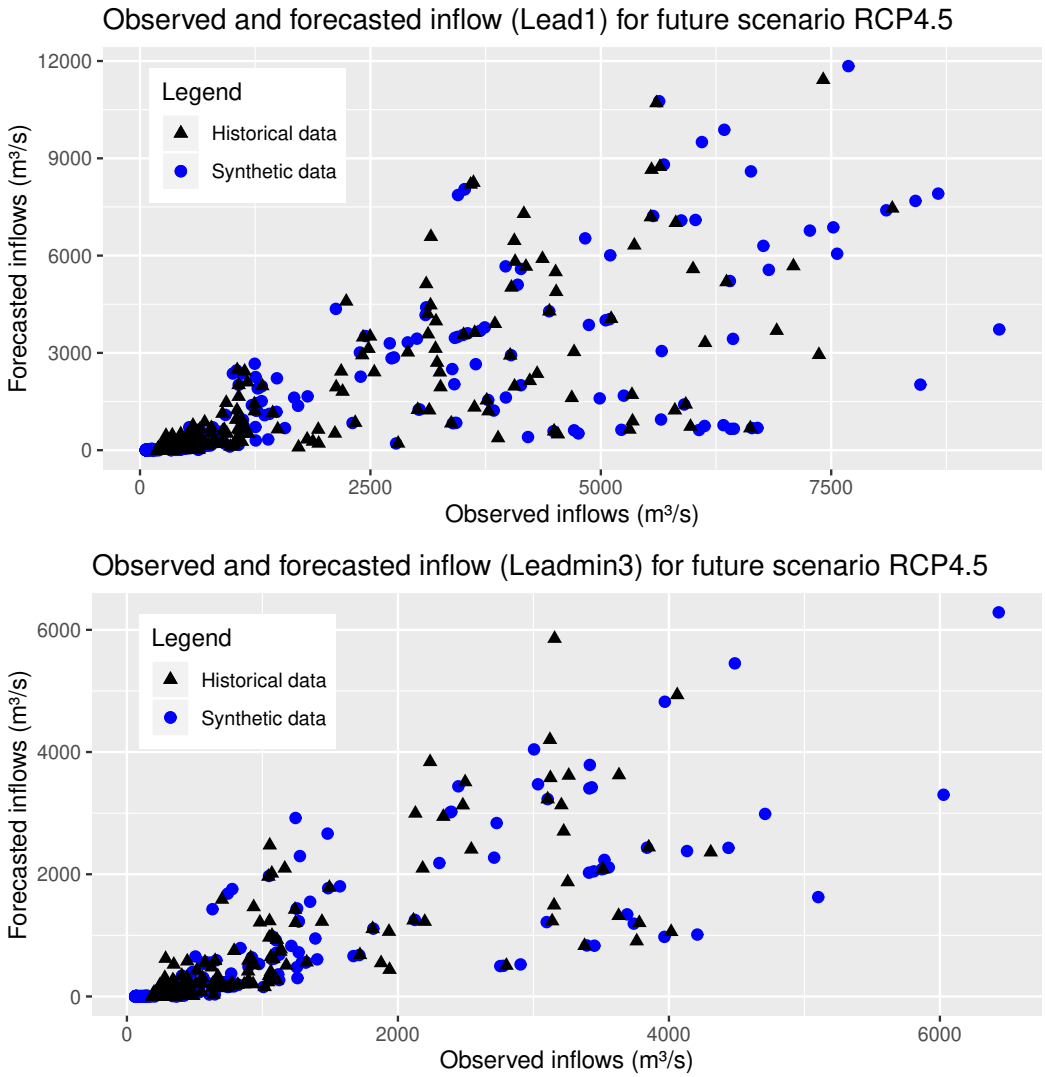

Figure 5.14: *Scatterplot presenting the forecasted inflow based on the observed inflow over one month Lead1 (top panel) and minimum inflow over 3 months Leadmin3 (low panel). Black triangles represent the historical data (1993-2014) and the red circles the synthetic forecast generated for the RCP 4.5 scenario in the time period (2071-2092).*

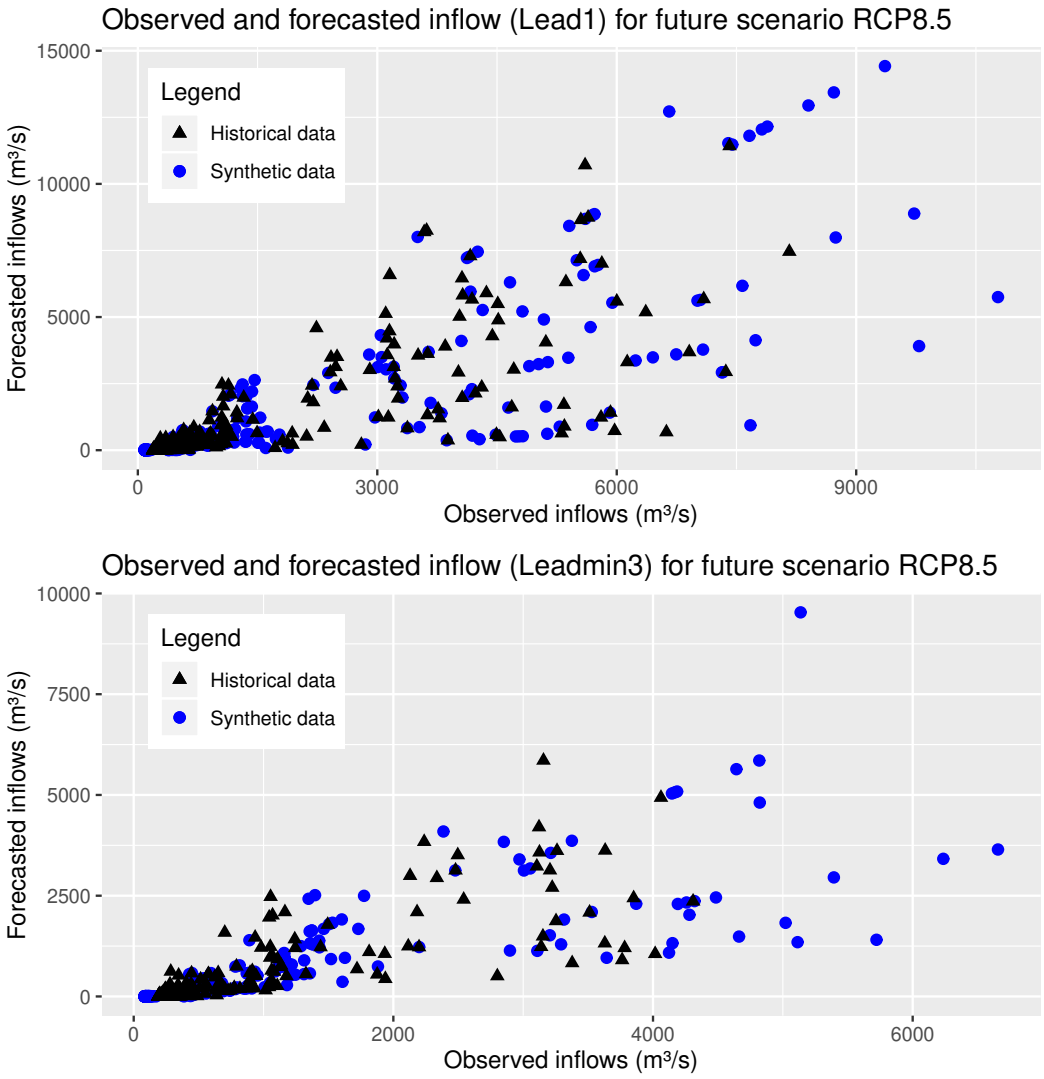

Figure 5.15: *Scatterplot presenting the forecasted inflow based on the observed inflow over one month Lead1 (top panel) and minimum inflow over 3 months Leadmin3 (low panel). Black triangles represent the historical data (1993-2014) and the red circles the synthetic forecast generated the RCP 8.5 scenario in the time period (2071-2092).*

5.2.3 Forecast value

Once we obtained the trajectories of "inflows" and "forecasts" projected into future climate change scenarios, we quantified the EVPI as the performance improvement between the Perfect and Baseline Solutions by obtaining an absolute gain in the annual hydropower production of 690.13 GWh/y for the RCP 4.5 scenario and 625.89 GWh/y for the RCP 8.5 scenario corresponding to a relative gain of 3.62% and 2.96%, respectively. As we did for the historical conditions, we also investigated the system performance during the June-September wet (Kiremt) season. In this case, results show again larger gains in hydropower production corresponding to a 79.31% and 61.58% improvement for RCP 4.5 and RCP 8.5. Finally, we assessed the projected forecast value by computing the difference in system performance between the Forecast and Baseline Solutions. Once again, results demonstrate the high potential value of forecast information during the wet season: we obtained a relative gain in hydropower production of 40.47% and 34.27% for the RCP 4.5 and the RCP 8.5 scenarios, respectively (tables 5.3- 5.4). The system dynamics under the three operating policies for the two climate scenarios is reported in figures 5.16 and 5.17 showing the GERD's level, release, and hydropower production. We can notice that the perfect operating policy (black line) is able to keep a high reservoir level around 140 m by storing water during the dry season and releasing it in the wet season, allowing for high hydropower production. The baseline operating policy (blue line) can not anticipate the large amount of water coming during the wet season and prefers releasing more water in the dry season leading to a trajectory of level that oscillates from 140 m to 130 m. The forecast operating policy (red line) is instead able to predict future inflows and attains a higher hydropower production than the Baseline Solution.

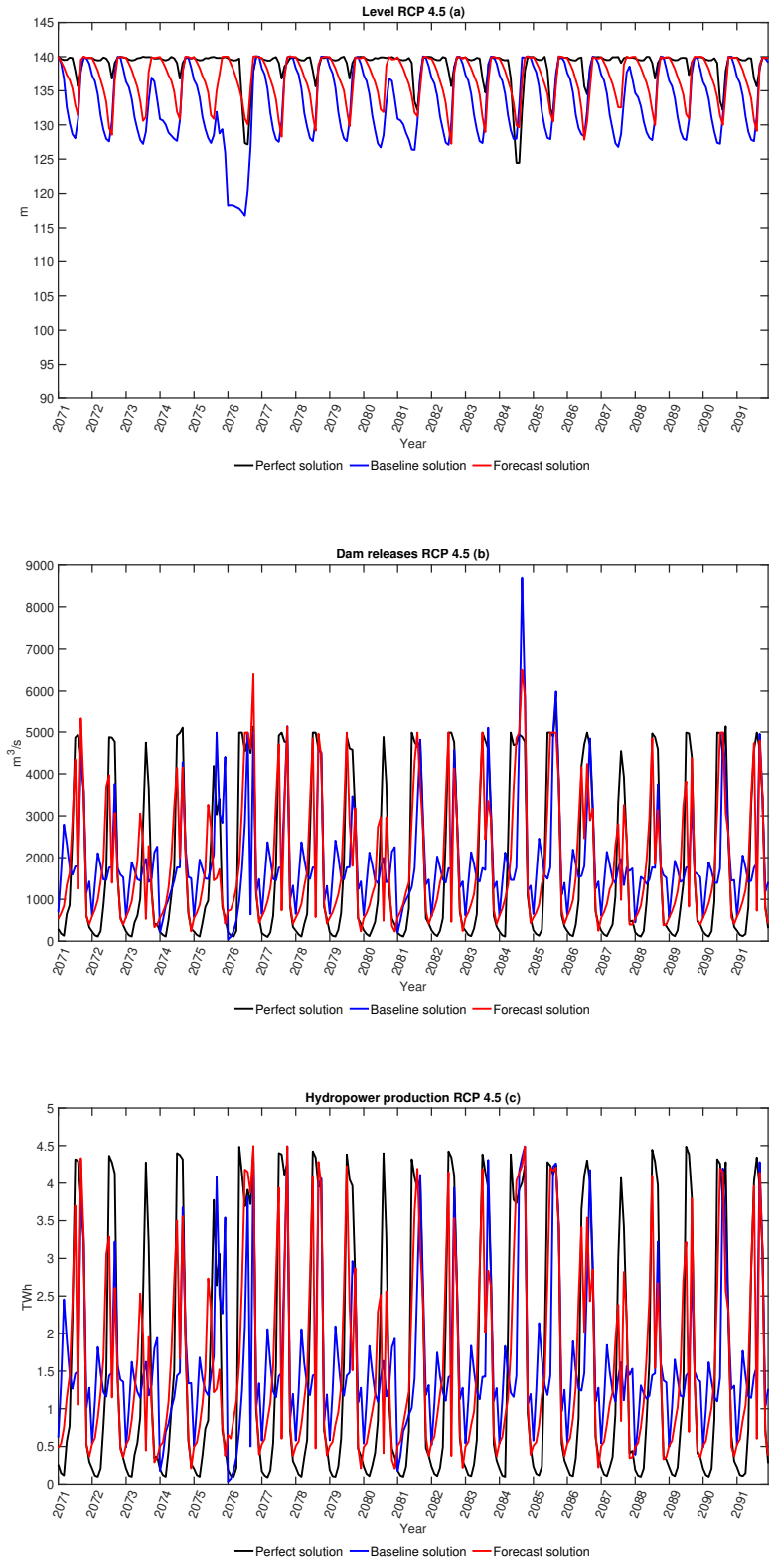

Figure 5.16: *The system dynamics under the baseline (blue line), perfect (black line), and forecast (red line) operating policies for climate scenario RCP 4.5 in time period 2071-2091. (a) Gerd's level; (b) Gerd's release; (c) Gerd's hydropower production.*

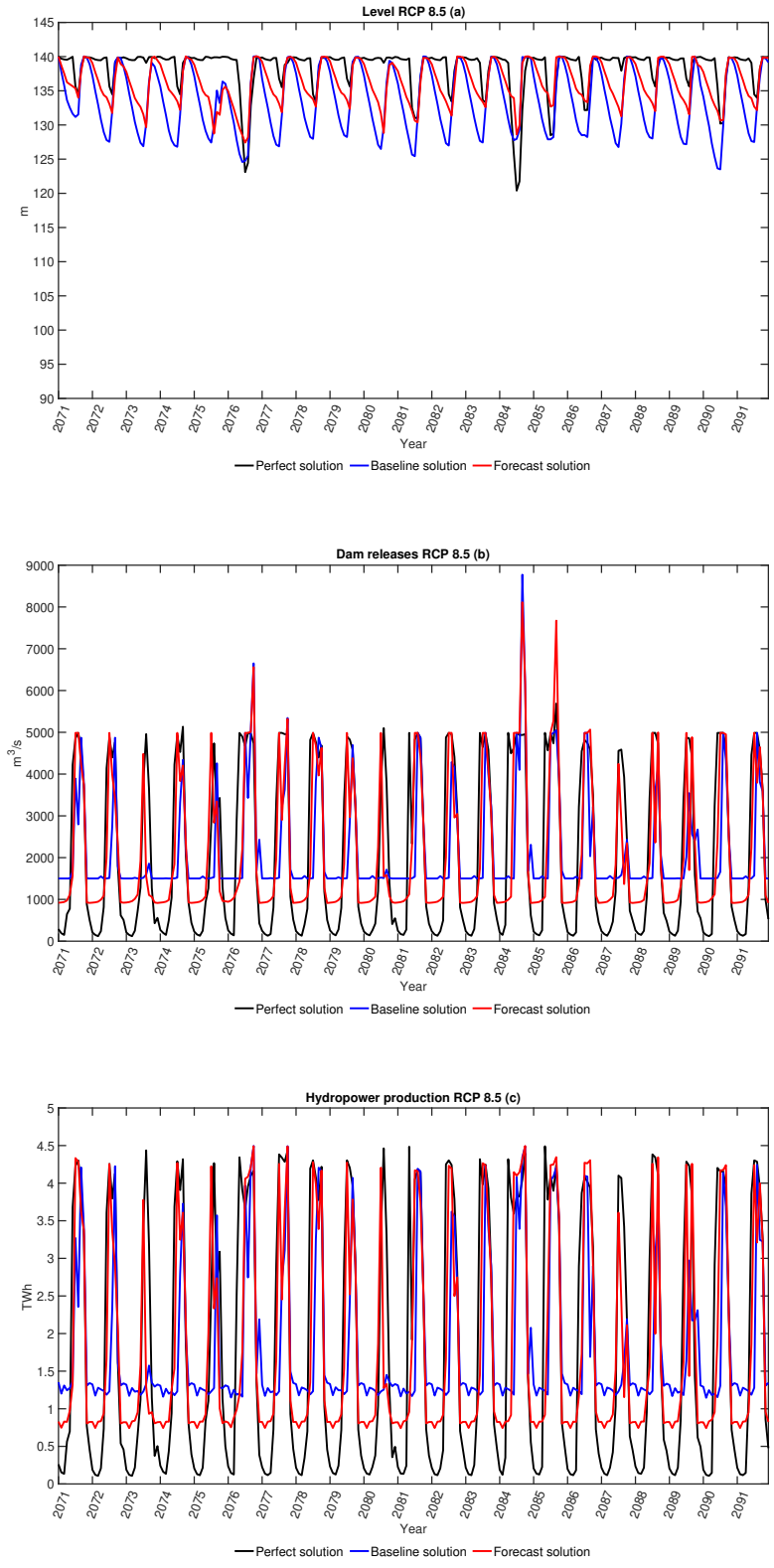

Figure 5.17: *The system dynamics under the baseline (blue line), perfect (black line), and forecast (red line) operating policies for climate scenario RCP 8.5 in time period 2071-2091. (a) Gerd's level; (b) Gerd's release; (c) Gerd's hydropower production.*

5.3 Comparison: historical and future conditions

A final comparison of the system performance under historical and future conditions is reported in figures 5.18- 5.19 and tables 5.3- 5.4, in terms of both annual and wet season hydropower production. We can notice (figure 5.18) that the annual hydropower productions obtained by the different operating policies in the future scenarios are higher than those in the historical period, especially in the RCP 8.5 scenario. In particular, the perfect operating policy attains a production of 19.761 TWh/y and 21.785 TWh/y for the RCP 4.5 and the RCP 8.5 scenarios compared to the historical performance of 16.714 TWh/y. These results can be explained by the large increase of annual inflow volumes in future scenarios. The baseline operating policy attains production slightly lower than the Perfect Solution in both historical and future conditions.
When focusing on the wet (Kiremt) season (figure 5.19), results show that the perfect operating policy in future scenarios leads to values of hydropower production equal to 14.278 TWh and 15.123 TWh for the RCP 4.5 and the RCP 8.5 scenarios with respect to the historical performance of 12.725 TWh. However, the projected hydropower production under the baseline operating policy decreases since the policy is conservative with respect to the increasing inflow caused by climate change. This strategy increases the gap with respect to the perfect operating policy, thus enlarging the space for benefitting from forecast information.
Table 5.3 summarizes the overall results, showing that the maximum space for increasing the annual hydropower production in the future increases mainly for the scenario RCP 4.5, with an improvement between the Perfect and Baseline solutions of 3.62% and a gain between the Forecast and Baseline Solutions of 1.27%. During the Kiremt season (table 5.4), seasonal forecasts lead to very large gains of 40.47% and 34.276% for the RCP 4.5 and the RCP 8.5 scenarios, respectively, confirming their high potential value in projected conditions.

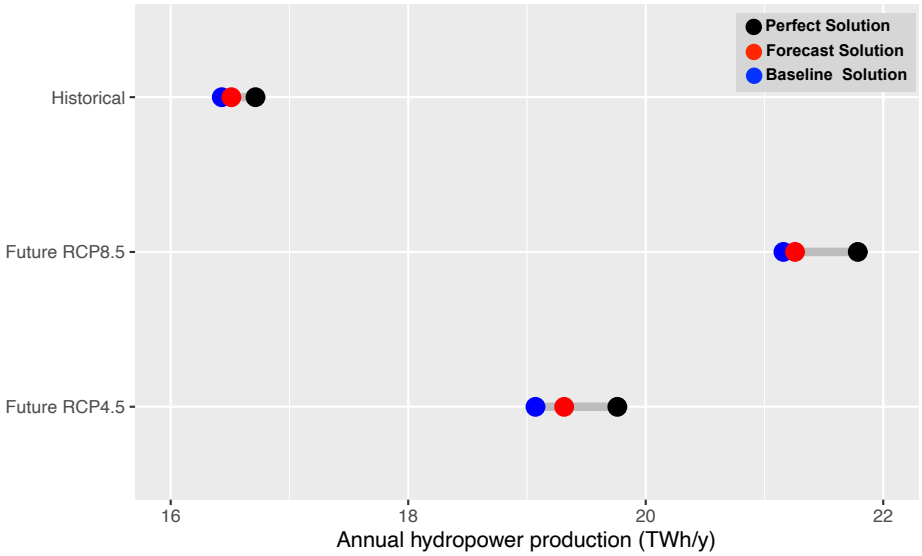

Figure 5.18: *Comparison of the performance improvement between different operating policies under the historical and the future conditions in terms of annual hydropower production.*

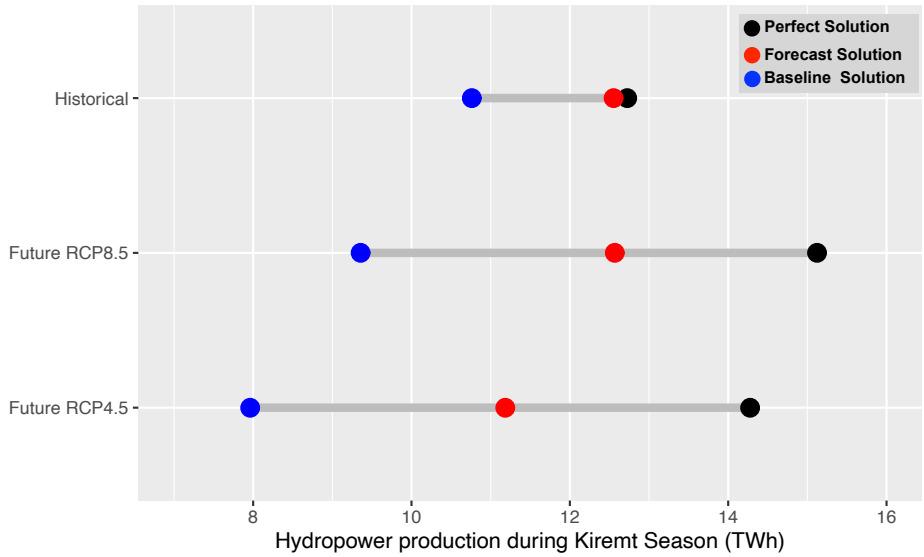

Figure 5.19: *Comparison of the performance improvement between different operating policies under the historical and the future conditions in terms of hydropower production during the wet (Kiremt) season.*

Table 5.3: *Comparison of the performance improvement between different operating policies under the historical and the future conditions in terms of annual hydropower production.*

Period	Policy	Hydropower production	Absolute gain	Relative gain	Economic gain
Historical	Baseline solution	16.430 TWh/y	-	-	-
	Perfect solution	16.714 TWh/y	+283.41 GWh/y	+1.73%	+22.67 M USD/y
	Forecast solution	16.51 TWh/y	+80.57 GWh/y	+0.49%	+6.45 M USD/y
Future RCP 4.5	Baseline solution	19.071 TWh/y	-	-	-
	Perfect solution	19.761 TWh/y	+690.13 GWh/y	+3.62%	+55.21 M USD/y
	Forecast solution	19.313 TWh/y	+241.91 GWh/y	+1.27%	+19.35 M USD/y
Future RCP 8.5	Baseline solution	21.159 TWh/y	-	-	-
	Perfect solution	21.785 TWh/y	+625.89 GWh/y	+2.96%	+50.07 M USD/y
	Forecast solution	21.257 TWh/y	+98.379 GWh/y	+0.465%	+7.87 M USD/y

Table 5.4: *Comparison of the performance improvement between different operating policies under the historical and the future conditions in terms of hydropower production during the wet (Kiremt) season.*

Period	Policy	Hydropower production	Absolute gain	Relative gain	Economic gain
Historical	Baseline solution	10.761 TWh	-	-	-
	Perfect solution	12.725 TWh	+1963.64 GWh	+18.25%	+157.09 M USD
	Forecast solution	12.554 TWh	+1793.17 GWh	+16.66%	+143.45 M USD
Future RCP 4.5	Baseline solution	7.963 TWh/y	-	-	-
	Perfect solution	14.278 TWh	+6315.64 GWh	+79.31%	+505.25 M USD
	Forecast solution	11.185 TWh	+3222.51 GWh	+40.47%	+257.80 M USD
Future RCP 8.5	Baseline solution	9.359 TWh	-	-	-
	Perfect solution	15.123 TWh	+5763.5 GWh	+61.58%	+461.08 M USD
	Forecast solution	12.567 TWh	+3207.99 GWh	+34.276%	+256.64 M USD

The results discussed so far can be integrated to provide an estimate of the economic value of the forecasts used in conditioning the GERD operations. This monetary value can also be considered an estimate of the maximum amount a decision maker might be willing to pay in order to acquire the hydroclimatic service offered by SMHI in order to increase the GERD hydropower production. In this analysis, we transform the production into a gross revenue by using a constant energy price equal to 0.08 USD per kilowatt hour (*Block and Strzepek*, 2010; *Jeuland and Whittington*, 2014). The economic gains obtained under the forecast operating policy correspond to +6.45 Million USD/y for the historical period, +19.35 Million USD/y for the RCP 4.5 and +7.87 Million USD/y for the RCP 8.5 scenarios (table 5.3). These results show that even though relative gains of the annual hydropower production are very low suggesting that forecasts might not be relevant, the energy produced in any case is high leading to large economic values. The economic gains by adopting the forecast operating policy is further expanded during the wet season with values of +143.45 Million USD for the historical period, +257.80 Million USD for the RCP 4.5 and +256.64 Million USD for the RCP 8.5 scenarios (table 5.4).

6

Conclusions

The recent advance in skills of hydroclimatic services are motivating the uptake of these systems in order to improve water resource management strategies. Forecasts can anticipate both short-term events like floods and long-term phenomena such as droughts, and many studies indicate the advantages of using forecast information in water system operations. However, how to best inform hydropower systems operations is still an open question since hydropower can benefit from hydroclimatic services over a broad range of time scales, from short-term to seasonal and decadal time horizons.

This thesis contributes a novel machine-learning procedure to quantify the value of seasonal forecasts as their contribution to increasing the hydropower production of the Grand Ethiopian Renaissance Dam (GERD) in Ethiopia. The quantification of the forecasts value, relies on the Information Selection Assessment framework, which is applied to a service based on bias adjusted ECMWF SEAS5 seasonal forecasts used as input to the World-Wide HYPE hydrological model. First, we evaluated the expected value of perfect information as the potential maximum improvement of a baseline operating policy relying on a basic information with respect to an ideal operating policy designed under the assumption of perfect knowledge of future conditions. Results show that the maximum space for increasing the hydropower production of the GERD baseline operations not informed by any forecast is relatively small, corresponding to 1.73% of improvement with respect the solution given by a perfect operating policy. However, this potential gain becomes larger (18.25%) when we focused

on the performance during the heavy rainy season from June to September (Kiremt season), making room for the uptake of forecast information. This season was recommended by state-of-the-art researches for seasonal forecasting in the region due to the high correlation with ENSO phenomenon which may contribute to extending hydrologic predictability. In the second step, we selected the most informative lead times of inflow forecast by employing the Iterative Input Selection algorithm (IIS) and we identified inflows over one month and the minimum over three months as first and second most relevant lead times, meaning that inflow volumes can be more important in the first month while extremes can prevail at longer time horizons. Finally, we assessed the expected value of sample information as the performance improvement that could be achieved when the inflow forecasts for the selected lead times are used to inform operational decisions. As expected, results show a small improvement of 0.49% when considering an annual hydropower production and a large gain 16.66% when focusing on the system performance during the wet (Kiremt) season. This final result shows clearly that forecast information becomes mostly valuable during the wet season since seasonal forecasts allow to promtly anticipate the hydropower operation by predicting the large hydrological variability that characterizes this particular period of the year.
In addition, we analyzed the potential value of forecast information under different future climate scenarios. We first generate the projected inflows over climate change scenarios and develop a synthetic forecast model to generate streamflow forecasts for the future scenarios. We then applied the same methodology as we did in the historical conditions. The overall results show an increase of annual hydropower production in the future, when evaluating the system perfomance with and without forecasts since annual inflows are expected to increase. However, the gains in the annual production by adopting forecast information are still small, corresponding to 1.27% for the RCP 4.5 scenario and 0.46% for the RCP 8.5 scenario over the period 2071-2100. The forecast value is again larger during the wet season with gains of 40.47% and 34.276% for the RCP 4.5 and the RCP 8.5 scenarios, respectively. These large improvements obtained in the future scenarios with respect to the historical conditions are probably due to the fact that climate change is expected to accentuate extreme events by further decreasing inflows in the dry season and increasing it in the wet season. Therefore, having a forecast system that can anticipate future inflow volumes allows a better allocation of water during the year, thus producing more hydroelectric energy. Our final analysis converted the forecast value into economic terms showing large gains during the Kiremt season of +143.45 Million USD for the historical period, +257.80 Million USD

for the RCP 4.5 scenario, and +256.64 Million USD for the RCP 8.5 scenario.

It is worth reminding that these results were obtained making different assumptions. We first assumed that the Grand Ethiopian Renaissance Dam (GERD) is all completed and the reservoir is fully filled. Today the dam construction is almost finished and it will require from 5 to 15 years to fill the reservoir. Moreover, we neglected the presence of the downstream countries, mainly Sudan and Egypt that will be affected by a reduction of water availability at least during the filling period. In particular, Egypt has voiced concerns that the GERD would negatively impact the country's water supply, especially in light of overpopulation fears, and has thus insisted throughout negotiations that measures has to be put in place to protect downstream countries in case of droughts during the filling process of the dam. An agreement on the GERD will be transformational for the region, resulting in significant transboundary cooperation, regional development and economic integration, and improvement in the lives of the more than 250 million people of Egypt, Ethiopia, and Sudan. In this context, the added production obtained with the forecast-informed operations of the GERD may represent an additional option in the current negotiations about the dam impacts.
Further research effort could be addressed to assess the forecast value considering downstream countries' interest such as the minimization of water deficit for Egypt and minimization of irrigation water deficit for Sudan.
Additionally, since seasonal forecasts lead to an increase of hydropower production especially during the Kiremt season, it could be interesting to translate this gain in production into economic terms by adopting a time-varying price. The energy market in Ethiopia is set politically and most of hydropower dams, electricity network, and the distribution stations are all owned by the governmental monopoly, the Ethiopian Electric Power. The existing electric power generation costs are about 0.09 USD per kilowatt-hour (kWh) and the current price of electricity is at 0.06 USD per kWh, one of the lowest in Africa. So far the difference of 0.03 USD per kWh has been subsidized by the Ethiopian government (*Damte Beyene*, 2019). In this thesis, we assumed a constant energy price equal to 0,08 USD KWh (*Block and Strzepek*, 2010; *Jeuland and Whittington*, 2014); however, this price is expected to rise considerably since the government can no longer afford to subsidize electricity consumption and is changing its economic strategy to guarantee continued interest of foreign investors to invest into electricity generation projects in Ethiopia.
Finally, the overall methodology adopted in this thesis for quantifying forecast value could be applied to other case studies characterized by different hydro-

logical regimes, inflow predictability, and hydropower systems with different installed and storage capacity.

Bibliography

Abtew, W., A. M. Melesse, and T. Dessalegne (2009), El Niño southern oscillation link to the Blue Nile River basin hydrology, *Hydrological Processes: An International Journal*, *23*(26), 3653–3660.

Adamowski, J. F. (2008), River flow forecasting using wavelet and cross-wavelet transform models, *Hydrological Processes: An International Journal*, *22*(25), 4877–4891.

Amarasekera, N., R. F. Lee, E. R. Williams, and B. E. EA (1997), ENSO and the natural variability in the flow of tropical rivers, *Journal of Hydrology*, *200*, 24–39.

Anghileri, D., N. Voisin, A. Castelletti, F. Pianosi, B. Nijssen, and D. P. Lettenmaier (2016), Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments, *Water Resources Research*, *52*(6), 4209–4225.

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

Avery, S., and C. Eng (2012), Lake Turkana & the Lower Omo: hydrological impacts of major dam and irrigation developments, *African Studies Centre, the University of Oxford*.

Bank, W. (2006), Ethiopia: Managing water resources to maximize sustainable growth.

Bellmann, R. (1957), Dynamic programming princeton university press, *Princeton, NJ*.

Block, P. (2011), Tailoring seasonal climate forecasts for hydropower operations, *Hydrology and Earth System Sciences*, *15*(4), 1355–1368.

Block, P., and B. Rajagopalan (2007), Interannual variability and ensemble forecast of Upper Blue Nile Basin Kiremt season precipitation, *Journal of Hydrometeorology*, *8*(3), 327–343.

Block, P., and K. Strzepek (2010), Economic analysis of large-scale upstream river basin development on the Blue Nile in Ethiopia considering transient conditions, climate variability, and climate change, *Journal of Water Resources Planning and Management*, *136*(2), 156–166.

Block, P., and K. Strzepek (2012), Power ahead: meeting Ethiopia's energy needs under a changing climate, *Review of development economics*, *16*(3), 476–488.

Buizza, R. (2002), Chaos and weather prediction January 2000, *European Centre for Medium-Range Weather Meteorological Training Course Lecture Series ECMWF*.

Busoniu, L., D. Ernst, B. De Schutter, and R. Babuska (2010), Cross-entropy optimization of control policies with adaptive basis functions, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *41*(1), 196–209.

Camberlin, P., S. Janicot, and I. Poccard (2001), Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO, *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *21*(8), 973–1005.

Castelletti, A., F. Pianosi, and R. Soncini-Sessa (2008), Water reservoir control under economic, social and environmental constraints, doi: 10.1016/j.automatica.2008.03.003.

CDP (2018), List of least developed countries, `https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/ldc_list.pdf`.

Chow, T. W., and D. Huang (2005), Estimating optimal feature subsets using efficient estimation of high-dimensional mutual information, *IEEE Transactions on Neural networks*, *16*(1), 213–224.

CIA (2016), The world factbook 2016-17,, `https://www.cia.gov/library/publications/the-world-factbook/geos/et.html`.

Cloke, H., and F. Pappenberger (2009), Ensemble flood forecasting: A review, *Journal of hydrology*, *375*(3-4), 613–626.

Coello, C. A. C., G. B. Lamont, D. A. Van Veldhuizen, et al. (2007), *Evolutionary algorithms for solving multi-objective problems*, vol. 5, Springer.

Coffel, E. D., B. Keith, C. Lesk, R. M. Horton, E. Bower, J. Lee, and J. S. Mankin (2019), Future hot and dry years worsen Nile Basin water scarcity despite projected precipitation increases, *Earth's Future*, *7*(8), 967–977.

Conway, D. (1997), A water balance model of the Upper Blue Nile in Ethiopia, *Hydrological sciences journal*, *42*(2), 265–286.

Conway, D. (2000), The climate and hydrology of the Upper Blue Nile River, *Geographical Journal*, *166*(1), 49–62.

Dai, A. (2011), Drought under global warming: a review, *Wiley Interdisciplinary Reviews: Climate Change*, *2*(1), 45–65.

Damte Beyene, A. (2019), Impacts and Drivers of Policies for Electricity Access: Micro-and-Macroeconomic evidence from Ethiopia, `https://energyeconomicgrowth.org/node/236`.

Denaro, S., D. Anghileri, M. Giuliani, and A. Castelletti (2017), Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data, *Advances in water resources*, *103*, 51–63.

Dumont Goulart, H. M. (2019), Assessing the operational value of forecast information under climate change.

EC (2015), A european research and innovation roadmap for climate services, *European Commission*, *702151*.

EEPCo (2013), Ethiopian electric power corporation: Ethiopian energy master plan.

El-Fadel, M., Y. El-Sayegh, K. El-Fadl, and D. Khorbotly (2003), The Nile River Basin: A case study in surface water conflict resolution, *Journal of Natural Resources and Life Sciences Education*, *32*, 107–117.

Elsafi, S. H. (2014), Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan, *Alexandria Engineering Journal*, *53*(3), 655–662.

Elsanabary, M. H., and T. Y. Gan (2014), Wavelet analysis of seasonal rainfall variability of the upper Blue Nile Basin, its teleconnection to global sea surface temperature, and its forecasting by an artificial neural network, *Monthly Weather Review*, *142*(5), 1771–1791.

Eltahir, E. A. (1996), El Niño and the natural variability in the flow of the Nile River, *Water Resources Research*, *32*(1), 131–137.

ENA (2017), The reality of the GERD, `http://ethpress.gov.et/herald/index.php/editorial-view-point/item/7729-the-reality-of-the-gerd`.

Ethiopia, D. (2011), Longest ever bridge in ethiopia under construction, *Retrieved on April*, *25*.

Ferraro, B., A. Bezzi, C. Rossini, and P. Mastrofini (2015), Fast track approach to design and construction at grand ethiopian renaissance dam, `https://www.hydroreview.com/articles/print/volume-23/issue-4/features/fast-track-approach-to-design-and-construction-at-grand-ethiopian-renaissance-dam.html`.

Galelli, S., and A. Castelletti (2013), Tree-based iterative input variable selection for hydrological modeling, *Water Resources Research*, *49*(7), 4295–4310.

Geurts, P., D. Ernst, and L. Wehenkel (2006), Extremely randomized trees, *Machine learning*, *63*(1), 3–42.

Giuliani, M., F. Pianosi, and A. Castelletti (2015), Making the most of data: an information selection and assessment framework to improve water systems operations, *Water Resources Research*, *51*(11), 9073–9093.

Giuliani, M., A. Castelletti, F. Pianosi, E. Mason, and P. M. Reed (2016a), Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations, *Journal of Water Resources Planning and Management*, *142*(2), 04015,050.

Giuliani, M., A. F. Castelletti, R. Fedorov, and P. Fraternali (2016b), Using crowdsourced web content for informing water systems operations in snow-dominated catchments.

Giuliani, M., M. Zaniolo, A. Castelletti, G. Davoli, and P. Block (2019), Detecting the State of the Climate System via Artificial Intelligence to Improve Seasonal Forecasts and Inform Reservoir Operations, *Water Resources Research*, *55*(11), 9133–9147, doi: 10.1029/2019WR025035.

Giuliani, M., L. Crochemore, I. Pechlivanidis, and A. Castelletti (2020), From skill to value: isolating the influence of end-user behaviour on seasonal forecast assessment, *Hydrology and Earth System Sciences Discussions*, pp. 1–20.

Guariso, G., and D. Whittington (1987), Implications of Ethiopian water development for Egypt and Sudan, *International Journal of Water Resources Development*, *3*(2), 105–114.

Hadka, D., and P. Reed (2013), Borg: An auto-adaptive many-objective evolutionary computing framework, *Evolutionary computation*, *21*(2), 231–259.

Hamlet, A. F., and D. P. Lettenmaier (1999), Columbia River streamflow forecasting based on ENSO and PDO climate signals, *Journal of water resources planning and management*, *125*(6), 333–341.

Harrigan, S., C. Prudhomme, S. Parry, K. Smith, and M. Tanguy (2018), Benchmarking ensemble streamflow prediction skill in the UK, *Hydrology and Earth System Sciences*, *22*(3).

Hejazi, M. I., X. Cai, and B. L. Ruddell (2008), The role of hydrologic information in reservoir operation–learning from historical releases, *Advances in water resources*, *31*(12), 1636–1650.

Herman, J. D., P. M. Reed, H. B. Zeff, and G. W. Characklis (2015), How should robustness be defined for water systems planning under change?, *Journal of Water Resources Planning and Management*, *141*(10), 04015,012.

IPCC, C. C. (2014), Mitigation of climate change, *Contribution of working group III to the fifth assessment report of the intergovernmental panel on climate change*.

IRENA (2015a), Africa 2030: Roadmap for a renewable energy future, `https://www.cia.gov/library/publications/the-world-factbook/geos/et.html`.

IRENA, R. E. S. (2015b), International Renewable Energy Agency, *Renewable Energy Target Setting, Abu Dhabi, UAE*.

Jeuland, M., and D. Whittington (2014), Water resources planning under climate change: Assessing the robustness of real options for the Blue Nile, *Water Resources Research*, *50*(3), 2086–2107.

Kasprzyk, J. R., S. Nataraj, P. M. Reed, and R. J. Lempert (2013), Many objective robust decision making for complex environmental systems undergoing change, *Environmental Modelling & Software*, *42*, 55–71.

King, A., and P. Block (2014), An assessment of reservoir filling policies for the Grand Ethiopian Renaissance Dam, *Journal of Water and Climate Change*, *5*(2), 233–243.

Kingston, D. G., G. R. McGregor, D. M. Hannah, and D. M. Lawler (2006), River flow teleconnections across the northern North Atlantic region, *Geophysical Research Letters*, *33*(14).

Kwak, N., and C.-H. Choi (2002), Input feature selection for classification problems, *IEEE transactions on neural networks*, *13*(1), 143–159.

Li, Y., M. Giuliani, and A. Castelletti (2017), A coupled human-natural system to assess the operational value of weather and climate services for agriculture, *Hydrology and Earth System Sciences*, *21*, 4693–4709.

Loucks, D. P., E. Van Beek, J. Stedinger, J. Dijkman, and M. Villars (2005), Water resources systems planning and management–An introduction to methods, models and applications, Studies and Reports in Hydrology series.

Lynch, P. (2008), The origins of computer weather prediction and climate modeling, *Journal of Computational Physics*, *227*(7), 3431–3444.

Maier, H. R., A. Jain, G. C. Dandy, and K. P. Sudheer (2010), Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, *Environmental modelling & software*, *25*(8), 891–909.

MIT (2014), The Grand Ethiopian Renaissance Dam: An opportunity for collaboration and shared benefits of the Eastern Nile Basin.

MoFED (2010), Growth and Transformation Plan (GTP). Addis Ababa: Ministry of Finance and Economic Development (MoFED), Federal Democratic Republic of Ethiopia.

Mulat, A. G., and S. A. Moges (2014), Assessment of the impact of the Grand Ethiopian Renaissance Dam on the performance of the High Aswan Dam, *Journal of Water Resource and Protection*, *2014*.

Murphy, A. H. (1993), What is a good forecast? an essay on the nature of goodness in weather forecasting, *Weather and forecasting*, *8*(2), 281–293.

Nayak, M. A., J. D. Herman, and S. Steinschneider (2018), Balancing Flood Risk and Water Supply in California: Policy Search Integrating Short-Term Forecast Ensembles With Conjunctive Use, *Water Resources Research*, *54*(10), 7557–7576.

NBI (2012), The nile basin initiative (NBI), `https://www.nilebasin.org`.

Poveda, G., D. M. Alvarez, and O. A. Rueda (2011), Hydro-climatic variability over the Andes of Colombia associated with ENSO: a review of climatic processes and their impact on one of the Earth,Äôs most important biodiversity hotspots, *Climate Dynamics*, *36*(11-12), 2233–2249.

Powell, W. B. (2007), *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703, John Wiley & Sons.

Ramos, M., et al. (2019), Hydropower Policy Brief: The benefits of informing hydropower reservoir operations with climate and inflow forecasts, *Tech. rep.*, IMPREX project.

Ramos, M. H., S. J. Van Andel, and F. Pappenberger (2013), Do probabilistic forecasts lead to better decisions?

Reed, P. M., D. Hadka, J. D. Herman, J. R. Kasprzyk, and J. B. Kollat (2013), Evolutionary multiobjective optimization in water resources: The past, present, and future, *Advances in water resources*, *51*, 438–456.

Rivers, I. (2012), Field Visit Report: GERD Project.

Rodell, M., J. Famiglietti, D. Wiese, J. Reager, H. Beaudoing, F. W. Landerer, and M.-H. Lo (2018), Emerging trends in global freshwater availability, *Nature*, *557*(7707), 651–659.

Rosenstein, M. T., and A. G. Barto (2001), Robot weightlifting by direct policy search, in *International Joint Conference on Artificial Intelligence*, vol. 17, pp. 839–846, Citeseer.

Saavedra Valeriano, O. C., T. Koike, K. Yang, T. Graf, X. Li, L. Wang, and X. Han (2010), Decision support for dam release during floods using a distributed biosphere hydrological model driven by quantitative precipitation forecasts, *Water Resources Research*, *46*(10).

Salazar, J. Z., P. M. Reed, J. D. Herman, M. Giuliani, and A. Castelletti (2016), A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control, *Advances in water resources*, *92*, 172–185.

Seleshi, Y., and U. Zanke (2004), Recent changes in rainfall and rainy days in Ethiopia, *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *24*(8), 973–983.

Sharma, A. (2000), Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 3-A nonparametric probabilistic forecast model, *Journal of Hydrology*, *239*(1-4), 249–258.

Soncini-Sessa, R., E. Weber, and A. Castelletti (2007), *Integrated and participatory water resources management-theory*, Elsevier.

Steinschneider, S., R. McCrary, S. Wi, K. Mulligan, L. O. Mearns, and C. Brown (2015), Expanded decision-scaling framework to select robust long-term water-system plans under hydroclimatic uncertainties, *Journal of Water Resources Planning and Management*, *141*(11), 04015,023.

Taye, M. T., T. Tadesse, G. B. Senay, and P. Block (2016), The Grand Ethiopian Renaissance Dam: Source of Cooperation or Contention?, *Journal of Water Resources Planning and Management*, *142*(11), 02516,001.

Tesfa, B. (2013), Benefit of Grand Ethiopian Renaissance Dam Project (GERDP) for Sudan and Egypt.

Toth, E., A. Brath, and A. Montanari (2000), Comparison of short-term rainfall prediction models for real-time flood forecasting, *Journal of hydrology*, *239*(1-4), 132–147.

Treaty, N. (1959), Agreement between the Arab Republic of Egypt and the Republic of the Sudan for the full utilization of the Nile waters, *United Nations Treaty Series*, *64*, 1963.

Tsitsiklis, J. N., and B. Van Roy (1996), Feature-based methods for large scale dynamic programming, *Machine Learning*, *22*(1-3), 59–94.

United Nations, P. D. (2017), World population prospects: the 2017 revision, custom data acquired via website.

UnitedNations (2019), World Population Prospect, `https://population.un.org/wpp/`.

Van Veldhuizen, D. A., and G. B. Lamont (1998), Evolutionary computation and convergence to a pareto front, in *Late breaking papers at the genetic programming 1998 conference*, pp. 221–228.

Verhoeven, H. (2011), *Black Gold for Blue Gold?: Sudan's Oil, Ethiopia's Water and Regional Integration*, Chatham House London.

Waterbury, J. (2008), *The Nile Basin: National determinants of collective action*, Yale University Press.

Wheeler, K. G., J. W. Hall, G. M. Abdo, S. J. Dadson, J. R. Kasprzyk, R. Smith, and E. A. Zagona (2018), Exploring cooperative transboundary river management strategies for the Eastern Nile Basin, *Water resources research*, *54*(11), 9224–9254.

WorldBank (2019), World bank open data, `https://data.worldbank.org`.

Yihdego, Y., A. Khalil, and H. S. Salem (2017), Nile River,Äôs basin dispute: perspectives of the Grand Ethiopian Renaissance Dam (GERD), *Global Journal of Human-Social Science: B, Geography, Geo-Sciences, Environmental Science & Disaster Management*, *17*(2), 1–21.

Yuan, X., E. F. Wood, and Z. Ma (2015), A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development, *Wiley Interdisciplinary Reviews: Water*, *2*(5), 523–536.

Zaroug, M. A., E. A. Eltahir, and F. Giorgi (2014), Droughts and floods over the upper catchment of the Blue Nile and their connections to the timing of El Niño and La Niña events.

Zhang, Y., P. Block, M. Hammond, and A. King (2015), Ethiopia,Äôs Grand Renaissance Dam: Implications for downstream riparian countries, *Journal of Water Resources Planning and Management*, *141*(9), 05015,002.

Zitzler, E., L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Da Fonseca (2003), Performance assessment of multiobjective optimizers: An analysis and review, *IEEE Transactions on evolutionary computation*, *7*(2), 117–132.

Appendix A

Additional material

A.1 EMODPS: runtime analysis

In order to filter the random components of MOEA algortithms, the optimizations were reiteraited for 10 seeds where solutions were then combined into a final reference set and the algorithm is terminated after a fixed number of function evaluations (NFE). With the runtime analysis we want to verify the convergence of the optimizazion problem through three performance metrics, namely the Generetional distance, the Additive ε-indicator, and the Hypervolume indicator accounting respectively for convergence, consistency, and diversity.
Generetional distance (figure A.1 (b)) is a measure of convergence and it consists of the minimum point in the average Euclidean distance vector between solutions in an approximation set and its corresponding nearest solutions in the reference set (*Van Veldhuizen and Lamont*, 1998). A large value for this metric implies that MOEA has failed to find a single solution close to the reference set.
Additive ε-indicator measures the worst case distance (versus the average distance) required to translate an approximation set solution to dominate its nearest neighbor in the reference set (*Reed et al.*, 2013). The indicator is very sensitive

to gaps in tradeoffs and can be viewed as a measure of an approximation set consistency with the reference set (i.e., all portions of the trade-off are present as illustrated in figure A.1 (c)). A low value for this metric is desired as it considers the worst-case distance from the reference set (*Hadka and Reed*, 2013).
The Hypervolume is a measure of diversity and proximity and it quantifies the volume of the objective space dominated by an approximation set; therefore, this metric is to be maximized. The hypervolume calculation is performed across the normalized objective function values, so the values are between 0 and 1, with 1 being the case in which the approximation set dominates the same volume as the reference set (figure A.1 (d)) (*Zitzler et al.*, 2003).
Figure A.2 illustrates the three performance metrics and the contribution to the reference set, for the baseline policy optimization under historical conditions (1993-2014). Each colour line represents one of the 10 seeds of the optimization process with respect to the increase number of function evaluations. Results shows that each metric gets better with the increase of NFE, reaching almost a flat line for all the seeds by the end of the 1 million evaluations, thus confirming that the selected number of function evaluations are enough to reach convergence. For this reason, we designed Baseline and Forecast operating policy for historical and future conditions adopting the same number of optimization seeds and number of function evaluations (NFE).

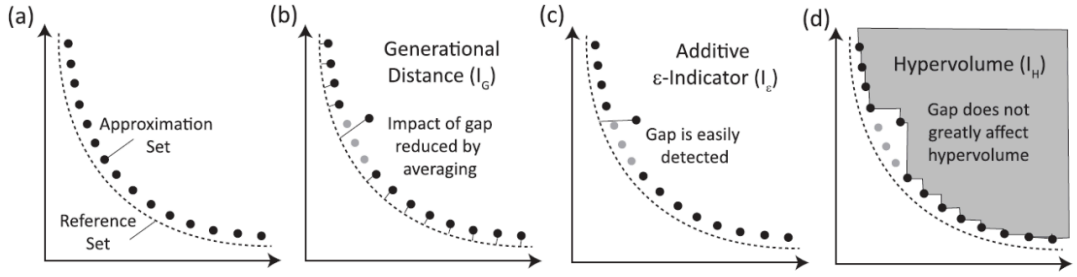

Figure A.1: *llustration of how different metrics capture convergence, diversity, and consistency. (a) A good approximation to the reference set, is indicated by the dashed line. (b) Generational distance averages the distance between the approximation set and reference set, requiring only one point near the reference set to attain ideal performance. (c) The ε-indicator metric focuses on the worst case distance required to translate a point to dominate its nearest neighbor in the reference set. (d) Hypervolume rigorously captures both convergence and diversity making it the most difficult metric to satisfy (Reed et al., 2013)*

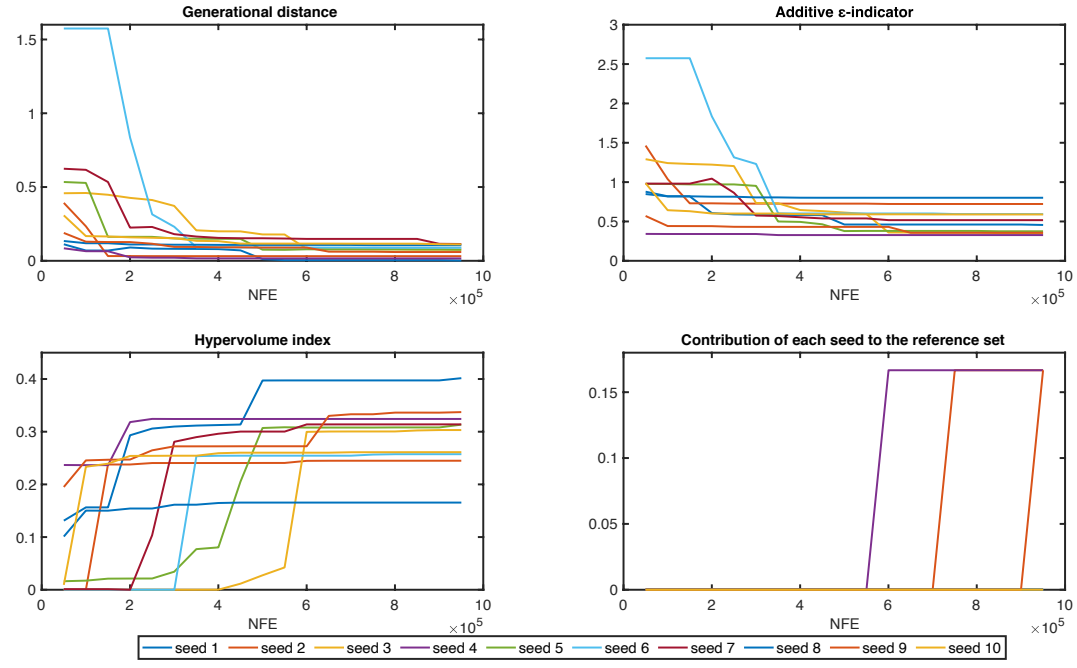

Figure A.2: *Results of the performance metrics for for the baseline policy optimization under historical conditions (1993-2014).*

A.2 IIS algorithm: setting parameters

As reported in section 4.2.3, there are three parameters that characterize the model building algorithm, namely M the number of trees that compose a forest, K the number of regression input chosen randomly at each node, and n_{min} the minimum sample size for splitting a node. However, this latter mainly affects the ensemble performance and overall method efficiency since large values of n_{min} lead to small trees, with high bias and small variance and low values of n_{min} lead to fully-grown trees, which may overfit the data. In this section we present the results obtained by setting M= 500, K=19 (number of input variables) for n_{min} values of 50, 40, 30, 20, 10, 5. The algorithm was run 15^{th} times in order to filter the randomness associated to extra trees models. Results show that with high values of n_{min}, the IIS algorithm ranks the variables in same position through different runs, but the model performance expressed in term of the coefficient of determination R^2 is low. In contrary, low values of n_{min} lead to an increase in R^2 value, but the variable ranking along runs is not same, making difficult to establish which are the most informative variables. A value of n_{min} = 20 is chosen in our analysis since IIS algorithm can reach high model performance preserving almost the same variable ranking among the different runs. The following figures represent information selection results obtained via 15 runs of the IIS algorithms for different value of n_{min}: graph (a) illustrates input variable ranking through different runs, where blocks with different colour represent the selected variable; graph (b) shows model perfermance (R^2) through different runs; graph (c) represents for each input variable (forecast lead times), the frequency of selection, the average position, and the average relative contribution in terms of coefficient of determination R^2.

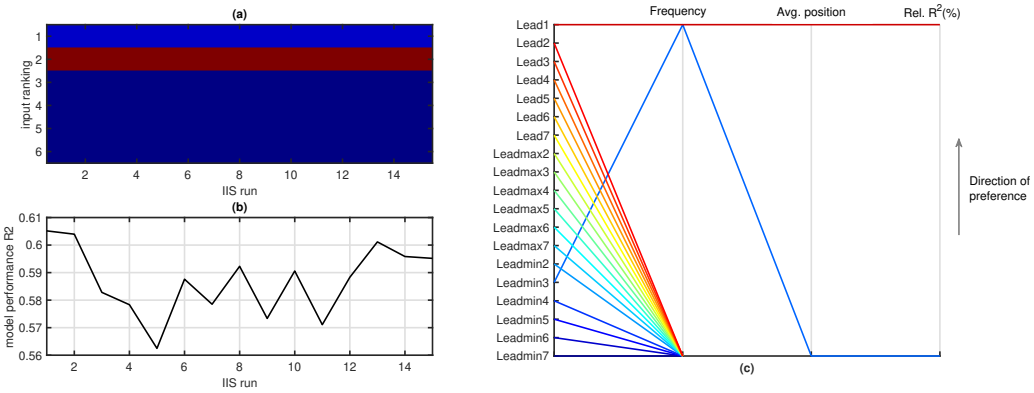

Figure A.3: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=50.*

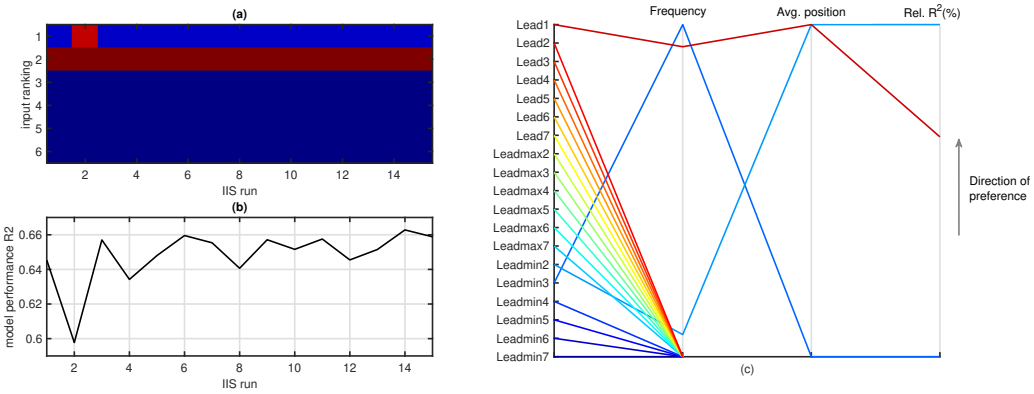

Figure A.4: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=40.*

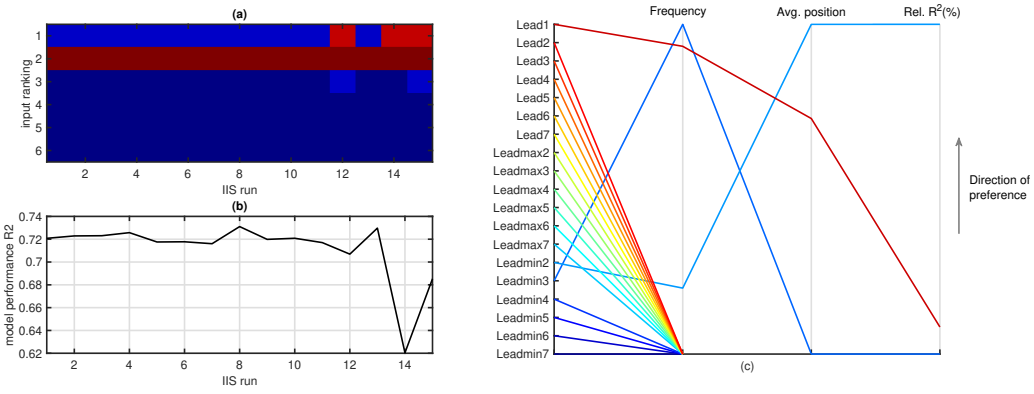

Figure A.5: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=30.*

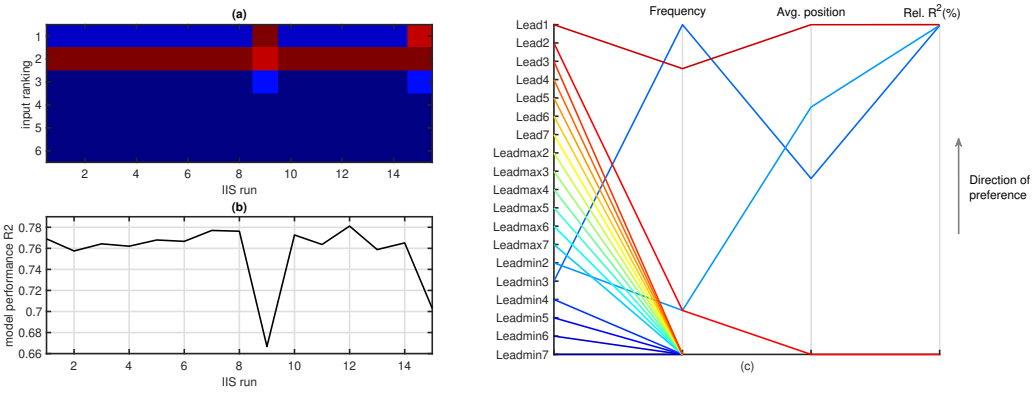

Figure A.6: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=20.*

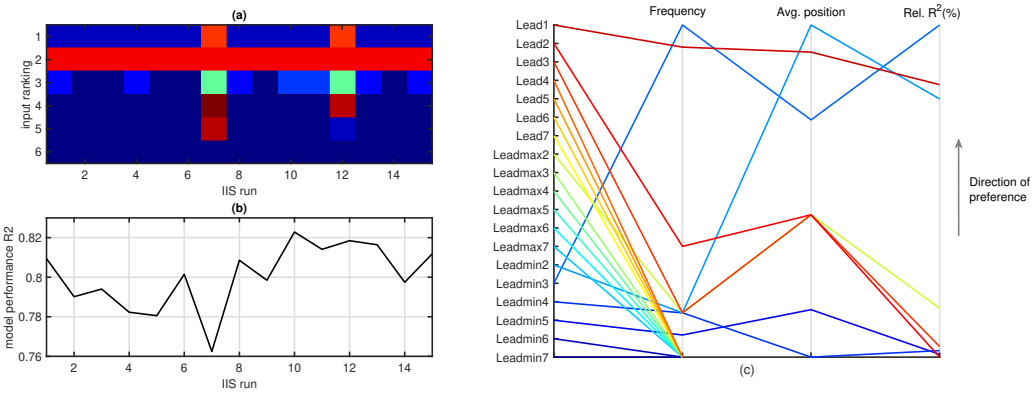

Figure A.7: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=10.*

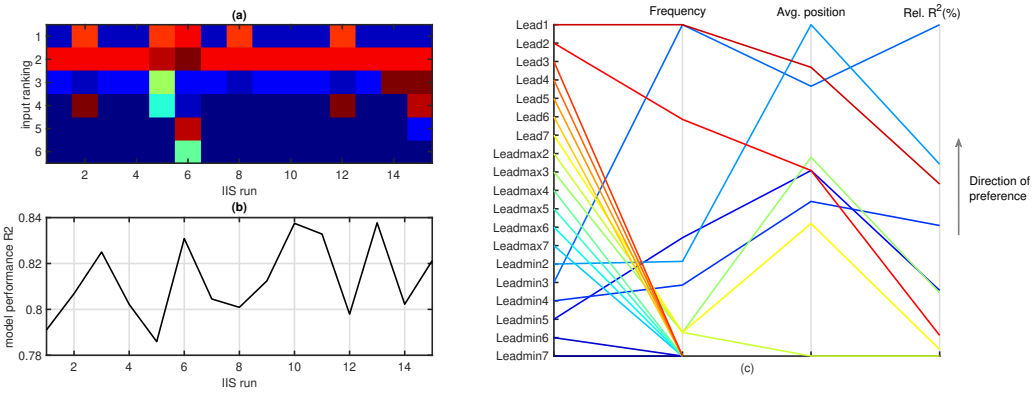

Figure A.8: *Information selection results obtained via 15 runs of the IIS algorithm setting M= 500, K =19, n_{min}=5.*