FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Optimisation approaches for operational decision-making in a watershed system with interconnected dams

Tiago Gonçalves Vaz

Mestrado em Engenharia Eletrotécnica e de Computadores

Supervisor: Maria Beatriz Brito Oliveira

July 24, 2023

Abstract

It is known that a big contributor to climate change is the burning of fossil fuels, due to its emission of greenhouse gases. In the energy production sector, this activity can be reduced by increasing the quantity and efficiency of renewable energies, such as solar panels, windmills, and hydropower plants.

This paper aims to optimize operational decision-making in watershed systems. To do so, a systematic representation of watersheds by a network of different types of connection points was constructed and a mathematical model in the form of MILP (Mixed-Integer Linear Programming) was developed. This model allows for the optimization of water allocation, reservoir operations, and decision-making processes within the watershed system. By considering various constraints such as water demand, reservoir capacities, environmental factors, and energy costs, the model seeks to find an optimal solution that maximizes profit while minimizing water consumption. The model was implemented in Python using the PuLP library, and CPLEX was used to solve it.

Several instances were created to do methodological tests on the model, which include real and simulated watersheds. Although the model did not face problems when optimizing for small watersheds, finding a solution for larger watersheds was taking substantially much more time. In light of this, two solution methods based on the R&F (Relax-and-Fix) heuristic were proposed: Partial R&F and Full R&F. They were tested against the model and each other in the various simulated watersheds to access their performance, and to determine the optimal values of their two parameters (Step and Overlap). In these tests, the OC (opportunity cost) is another parameter that also varied. It is this parameter that allows the model to minimize its water consumption, as the larger its value is, the less profitable it becomes to generate power or discharge water. The tests showed that, for low OC, both R&F approaches were vastly superior to the model in terms of execution time, while finding near-optimal solutions (Full R&F still had faster execution times than Partial R&F). For high OC values, even the model was giving fast results.

To access the viability of the model being used in a decision support system, some tests were made with real instances taken from the Cávado watershed, situated in Portugal. Each instance corresponded either to a drought, flood, or normal date. Regarding the methodological tests, the Full R&F only reached a solution a few times while Partial R&F and the model did not have much difficulty reaching the optimal/near-optimal solution. But the main intent of these tests was to compare the behavior of the model with that of the real Cávado watershed. In general, the model followed what the Cávado watershed did in the various climate scenarios, which suggests that the model could reasonably be used in a decision support system.

Resumo

É sabido que um dos principais factores que contribuem para a alteração climática é a queima de combustíveis fósseis, devido à sua emissão de gases com efeito de estufa. No sector da produção de energia, esta atividade pode ser reduzida através do aumento da quantidade e da eficiência das energias renováveis, como os painéis solares, os turbinas eólicas e as centrais hidroelétricas.

Este trabalho tem como objetivo otimizar a tomada de decisões operacionais em sistemas de bacias hidrográficas. Para tal, foi construída uma representação sistemática das bacias hidrográficas por uma rede de diferentes tipos de pontos de ligação e foi desenvolvido um modelo matemático na forma de MILP (Mixed-Integer Linear Programming). Este modelo permite otimizar a atribuição de água, a exploração das albufeiras e os processos de decisão no sistema de bacias hidrográficas. Ao considerar várias restrições, como a procura de água, as capacidades dos reservatórios, os factores ambientais e os custos de energia, o modelo visa encontrar uma solução ótima que maximize o lucro e minimize o consumo de água. O modelo foi implementado em Python usando a biblioteca PuLP, e o CPLEX foi usado para resolvê-lo.

Foram criadas várias instâncias para efetuar testes metodológicos ao modelo, que incluem bacias hidrográficas reais e simuladas. Embora o modelo não apresentasse problemas na otimização para pequenas bacias hidrográficas, a obtenção de uma solução para bacias hidrográficas maiores demorava substancialmente mais tempo. Perante este facto, foram propostos dois métodos de solução baseados na heurística R&F (Relax-and-Fix): R&F Parcial e R&F Completo. Estes métodos foram testados em comparação com o modelo e entre si nas diferentes bacias hidrográficas simuladas para avaliar o seu desempenho e determinar os valores óptimos dos seus dois parâmetros (Step e Overlap). Nestes testes, o CO (custo de oportunidade) é outro parâmetro que também variou. É este parâmetro que permite ao modelo minimizar o seu consumo de água, uma vez que quanto maior for o seu valor, menos rentável se torna turbinar ou descarregar água. Os testes mostraram que, para valores baixos de OC, ambas as abordagens R&F foram muito superiores ao modelo em termos de tempo de execução, encontrando soluções quase ótimas (o R&F Completo ainda teve tempos de execução mais rápidos do que o R&F Parcial). Para valores elevados de OC, até o modelo estava a dar resultados rápidos.

Para avaliar a viabilidade da utilização do modelo num sistema de apoio à decisão, foram efectuados alguns testes com instâncias reais retiradas da bacia hidrográfica do Cávado, situada em Portugal. Cada instância correspondia a uma data de seca, cheia ou de normalidade. No que diz respeito aos testes metodológicos, o R&F Completo apenas alcançou uma solução algumas vezes, enquanto o R&F Parcial e o modelo não tiveram muita dificuldade em alcançar a solução ótima/ quase ótima. Mas o principal objetivo destes testes foi comparar o comportamento do modelo com o da bacia hidrográfica real do Cávado. Em geral, o modelo seguiu o comportamento da bacia hidrográfica do Cávado nos vários cenários climáticos, o que sugere que o modelo pode ser razoavelmente utilizado num sistema de apoio à decisão.

Acknowlegments

First of all, I would like to thank my supervisor Prof. Beatriz Brito Oliveira, for the continuous support, guidance, and patience that was needed to keep me on track. I am also grateful for her availability and promptitude to answer my questions and doubts.

To Luís Brandão, for helping me and for sharing his insights into dam management, and in the Cávado watershed in particular.

To my Friends, for helping me overcome the beginning hurdles, as I was still stumbling around.

Finally, to my Family: to my mom, Estela, for her unwavering support, to my dad, for the availability to find specialized people for me to talk to, and to my brother, for the graphic design, and for the reminders to work when procrastination was too strong.

Tiago Vaz

Contents

List of Figures

List of Tables

Abbreviations

APA	*Agência Portuguesa do Ambiente*
BRKGA	Biased Random-Key Genetic Algorithm
CNPGB	*Comissão National Portuguesa das Grandes Barragens*
EU	European Union
GCM	General Circulation Model
LP	Linear Programming
MIP	Mixed Integer Programming
MILP	Mixed Integer Linear Programming
NEMO	Nominated Electricity Market Operator
OC	Opportunity Cost
OF	Objective Function
OMIE	*Operador del Mercado Ibérico de Energía - Polo Español*
R&F	Relax-and-Fix
SNIRH	*Sistema Nacional de Informação de Recursos Hídricos*
SSVM	Smooth Support Vector Machine
USA	The United States of America

Chapter 1

Introduction

This chapter will present the problem at hand, give some context and motivation to solve said problem, define the objectives of this work and inform of the structure of the document.

1.1 General Context and Motivation

Renewable energy is energy that is generated from natural resources such as sunlight, wind, water, and geothermal heat, which are replenished naturally. These sources of energy are considered renewable because they are constantly restored and are not depleted by use. Renewable energy has the potential to significantly reduce our reliance on fossil fuels, which are finite resources and are a major contributor to climate change. However, it is important to consider the potential impacts of renewable energy projects, such as their potential impact on the environment and local communities. It is also important to ensure that renewable energy is developed and used sustainably, considering the long-term availability of the natural resources used to generate it.

Hydropower has long been a reliable energy source since the windmill's invention, and the most well-known technique for harvesting power through the natural cycle of water is the hydropower dam. Humans have constructed thousands of dams throughout the years, especially in the last century, as knowledge and engineering were improving. These giant structures form a barrier in the natural flow of rivers, which causes the formation of large bodies of water called reservoirs [1].

Dams have several benefits such as regulating river flows, helping flood control, and generating hydropower, among others [1]. The reservoirs in them contained supply water for biological consumption (human, animal, and agricultural), industrial use, recreation, fishing, water treatment, and sanitary use [2]. Despite these apparent advantages, dams can be severely damaging to the watershed in which they are built, for example: besides being a massive barrier to fish migration, they alter the water supply to downstream flora and fauna, and drown nearby flora when they are first built [1]. With this in mind, there is a growing understanding that the maintenance of the ecological equilibrium has to be considered when managing power generation [3].

Optimizing the operation of existing dam systems has become key in dealing with growing water demands because of the rising costs to build dams [4] and the several environmental issues that slow the starting process.

1.2 Problem Definition and Objectives

This problem handles the setup of a watershed with a river, its tributaries and multiple interconnected dams. In this context, the problem is to optimize the operational management of each dam according to some predefined objective like maximizing profit or energy production or even maintaining secure water levels in case of extreme weather. The decisions are bound by constraints, decision factors and uncertainties such as the weather and the energy markets.

Dams are generally used to generate hydropower or for environmental purposes, and in a system of interconnected dams, this management can be optimized further due to their global view of the system. At any time, the actions that can be taken in each dam are to drive water through turbines to generate power, to discharge water to the river (i.e., to let water through without generating power), to pump water back from a downstream dam or to do nothing. Each of these actions has its advantages and disadvantages depending on the price of energy, which is not constant throughout time, environmental needs such as minimum and maximum water flow, the expected weather conditions (which besides varying throughout time can also be uncertain, so there is a need to work with margins of error), the structural integrity of the physical system, like minimum and maximum levels of water on each dam, flood prevention, among others.

The objective of this work is to create a model of a general watershed system and develop an optimization algorithm to support the decision-making of the operational management of the dams in said system. An application for this work is to aid in optimizing real systems of interconnected dams and checking what possible outcomes a certain decision would have on the system, therefore helping to ponder whether that is the right action to take. This could help generate more profit and better manage a watershed, environmentally speaking.

1.3 Document Structure

Besides the introduction, this report contains five more chapters. Chapter 2 will present the state-of-the-art, discuss the approaches found and contextualize this work among the current solutions. Chapter 3 will detail and explain the mathematical model. Chapter 4 will propose an alternative solution method to solve the model. Chapter 5 will present the results of the methodological tests as well as compare the behavior of the model with that of a real watershed. Finally, chapter 6 will access the results obtained, summarize the most important conclusions and present some ideas for future work.

Chapter 2

Literature Review

This chapter will present the literature reviewed in preparation for this work. It explores the several factors that affect operational decision-making when operating a dam, like energy costs and weather conditions. It also discusses and analyzes the approaches found, in order to innovate and guide when designing the model and accessing its capabilities.

2.1 Dam Operation Management

Several factors can influence the decision to release water from a dam or reservoir. One of the main factors is the demand for water downstream of the dam. If there is a high demand for water for irrigation or other purposes, more water may be released from the dam. Another factor that can influence when water is released from a dam is the level of the reservoir. If the water level in the reservoir becomes too high, water may need to be released to prevent the dam from overflowing. Similarly, if the water level in the reservoir becomes too low, water should not be released. Environmental considerations also play a role in when water is released from a dam. For example, if the water level in a downstream river is low, more water may need to be released to maintain a healthy ecosystem. Finally, energy costs can also influence when water is released, as it may be more profitable to generate electricity when the price is the highest, or even to pump water from a downstream reservoir and only later generate electricity when the price is higher (considering the cost it took to pump the water).

The main natural factor that alters the water level in a reservoir is the water inflow, that is, the water that comes from upstream and stops at the reservoir. Although there are water inflow forecast techniques, there is always an error [5] which can come in the format of data uncertainty (rounding and/or measuring errors) and intrinsic uncertainty. Forecasting techniques rely on computational algorithms, along with present and past data, so the further into the future these algorithms try to predict, the greater the prediction error [6].

2.1.1 Downstream Water Needs

As dams regulate the water flow, they can be directly responsible for the abundance or lack of water downstream [7]. All works reviewed assume or imply that there is always a constant, minimum, flow through the dam (either by discharge or by electricity-generating turbines), that is given and that serves to supply industries and agriculture, and contains the minimum environmental flow, i.e., the minimal flow needed to satisfy the environmental needs downstream.

2.1.2 Energy Needs, Markets and Prices

Energy is a very important resource in the current society. It can be found in many different forms, such as coal, natural gas, and electricity, through the fusion and fission of atoms, among others. In a country, it is most commonly transmitted as electricity through the power grid. Despite every country having a power grid, in general, the lower the percentage of the population that has access to it, the worse their living conditions and the poorer that country will be [8]. Power grids are very important because they provide a reliable and consistent source of electricity to homes, businesses, industries, and other organizations, so they are essential to any country. Power grids also allow for the integration of renewable energy sources such as solar and hydropower, which can help reduce dependence on fossil fuels and decrease the production of greenhouse gases.

In the power grid, the amount of generated electricity and the amount of consumed electricity must be the same at all times [9]. Hydroelectric power plants such as dams generate electricity by using the kinetic energy of falling water to drive a turbine. This electricity is then immediately transferred to the power grid, making these power plants very efficient at balancing the power grid. Because energy consumption is not constant and renewable energy supply is intermittent, energy storage is critical in a 100% renewable future.

Pumped-storage hydropower is a common method of energy storage and it helps further optimize profit according to the variability of energy prices [10]. This seems reasonable as it is still unfeasible to store and transport large quantities of electricity [11] in batteries or even the power grid due to losses.

In the European Union (EU), the energy market is based on auctions where Nominated Electricity Market Operators (NEMOs) can bid for energy in a certain slot of time. It can be divided into three separate time periods [11]:

- Long-term markets: where NEMOs can buy energy at least one month in advance.
- Wholesale or spot markets: also called day-ahead or intraday markets.
- Balancing markets: balance with precision the supply and demand of energy in real-time.

In the EU, day-ahead markets, as the name implies, are auctions that happen every day, year-round, and that determine the hourly price of energy for the next day. In intraday markets, the bought energy is being used on the same day, and it can be sold up until one hour before being delivered. Depending on the member state, there are batches of 15, 30 and 60 minutes of energy, nonetheless, every member state uses the 60-minute batches [11][12].

In Brazil, there are day-ahead markets [13], as well as in the United States of America (USA), where there are also intraday markets. The latter has batches of 5 and 60 minutes to balance supply and demand [14]. China is in the process of implementing day-ahead and intraday (real-time) markets, with some provinces in the pilot stage from 2017 [15].

These countries/economies were discussed here because of several factors like population, economic health (investing in renewable energy does not show immediate payback), current use of hydropower and its share in the total energy production. China can benefit greatly from the use of hydroelectric power plants as its climate and area are very favourable to this power generation method. Also, China is an emerging superpower. Brazil already passed the USA in terms of hydropower generation, being only behind China, and the Brazilian hydropower sector alone meets more than 75% of the electricity demand in the country [16]. The USA and the EU are already demolishing dams for environmental reasons [17], so optimizing the operation of their existing dams is even more relevant.

With this information, it is possible to develop an optimization model that relies on day-ahead markets to know the energy prices a few hours in advance.

2.1.3 Weather Forecast

Within the effects of the weather, the ones that affect mostly dam operation are the precipitation and the temperature. The latter in conjunction with pressure and irradiation make the reservoir (and river) water evaporate, while precipitation adds water to the reservoir. Evaporation is a very important factor to consider especially in arid and semi-arid climates, where losses of water by evaporation can reach 25% of the total amount of consumption water allotted per year [18][19].

In the reviewed state-of-the-art, there are several ways to account for the weather in their models. Votruba et al. [2] in 1989 suggested making a baseline for climate forecast with the data collected in the previous years, but this method quickly fell out of use of the computer algorithms that could effectively predict the weather conditions for the next days. Still, some authors like Allawi et al. [20] and Martinez et al. [19] use that data and specific models, like Class-A pan evaporation data, to predict the monthly and daily evaporation, respectively. Cheng et al. [21] used General circulation models (GCMs) to forecast the weather and even used its output as an input in an hydrological model to predict climate change. So did Zhou et al. [22], but with smooth support vector machine (SSVM). Other authors like Ahmad et al. [23][24] use given weather forecast information in their models. This seems to be the better option because, if there is publicly available information about the weather forecasts, then there is no need to develop a new algorithm to do that.

Based on their literature review, Ahmad et al. [4] suggest that climate change could be taken into account to future-proof the new dam operation models. This challenge appeared in the last decades and it is a destabilizing factor when it comes to dam operation and watershed management in general, due to higher average temperatures and more extreme weather events. Climate change is felt in reservoir operations, but the models developed by Zhou et al. [22] and by Ahmad et al. [24] seem to handle it without trouble.

2.2 Risks

The major risks to the normal functioning of dams are extreme weather events like floods and droughts. In these situations, the operational management of dams gets complex, as any prediction weather related becomes less reliable due to the chaotic dynamics of these natural disasters. One minor risk that is addressable is the erosion and transport of sediments, specially their posterior sedimentation in the reservoir.

2.2.1 Floods

Flood control is important because flooding is the highest risk factor concerning dam breaking and it can cause immense losses in terms of life and property damage [25]. Dams can be used to help prevent flooding in certain areas. When a dam is built across a river, it creates a reservoir of water behind it. During times of heavy rain or snowmelt, the dam can be used to store the excess water, reducing the risk of downstream flooding. However, dams also have their limitations as a flood prevention measure. If the water level in the reservoir rises too high, the dam may not be able to hold all of the water and could fail, leading to even more serious flooding downstream. It is important to note that dams are not a perfect solution to flooding. In some cases, dams may be used in conjunction with other flood prevention measures, such as levees and floodplains.

Some dams already have pre-existing rules for flood control which were based on past floods, but they might not take into account the system as a whole and the time it takes for the water to go from one dam to the next [26].

Floods do not occur very often, but when they happen, it is usually in the same seasons of the year, varying from country to country. For example in China, the wet season (May-October) is when there is more probability of occurring typhoons [26][5], which in turn cause major floods. This likelihood of flooding allows for the development of a model with two standards, one for the flood season and one for the rest of the year. There is always the possibility of using a month-on-month standard, but this may overfit the model like in the work of Zhou et al. [22]. A possible implementation of the double standard model is to, during the flood season, dams would use their storage to prevent major floods, while the rest of the time, dams function normally or with less than desirable amounts of water [27]. Another, more strict, implementation is to have two different ways of operationally managing dams, one in the occurrence of floods and one otherwise [28] instead of seasonally. This is an interesting approach because it gives no time to prepare for floods (in terms of available storage), and yet it might give better results in the case where it is the flood season while it is not flooding, and in the opposite case.

To help dam operators to prepare for possible future floods, there is a real-time forecasting decision support system, which simulates river flows 5 days in advance [27].

2.2.2 Droughts

Droughts are characterized by a lack of water in a region for an extended period of time and, like floods, are natural disasters that humanity can not yet prevent. This can lead to crop failures, which can have serious consequences for farmers and for the food supply. Droughts can also lead to water shortages, which can affect communities and industries that rely on a steady supply of clean water. They bring economic and ecological complications, but their effects can be softened by the correct management of water in reservoirs [7].

There are few works that actively take droughts into account when designing their model. Despite the potential higher price of implementing flood protection, it is still easier and done more often than implementing drought protection. Palmer et al. [29] explain that this could be because there are fewer technical solutions for droughts than there are for floods.

Also like floods and other extreme weather events, drought frequency and intensity are increasing throughout the years due to climate change [18]. In their work, Chen et al. [7] optimized the amount of water supplied when compared to the amount required, and it was the only work in this literature review that took droughts into account in the design of their model. In the simulation constraints, Chen et al. decided to not make the first and last day of the simulation have the same water level in the reservoir, it was deemed unrealistic in drought conditions. This seems reasonable, as the simulation would optimize the operational management of that dam considering that there would be no day after the final day, clearly causing massive issues if its results were to be implemented in a real watershed.

A possible implementation of drought protection could be to treat droughts as floods: with drought seasons and double/multiple standard. Another possible implementation would be to simply have a minimum water level at all times that serves as preventive measure. Many works used this constraint, but not specifically for droughts. This, in addition to the gap in the literature relating to this issue, may suggest that droughts are not relevant enough to be considered in this specific problem and that other constraints and parameters like the rainfall outweigh the impact that droughts could have on the operational decision making process.

2.2.3 Erosion and Sedimentation

Sediment deposition is also a factor that alters the expected dam operation [18]. Sedimentation deposition is when particles of various sizes deposit in the reservoir, reducing its available storage over time. These particles can vary from stone, dirt and sand sediments to biologic bi-products or even human made trash like plastics.

There are already dams that have methods in place to counteract this. One of them is the dead storage one space, which is a specific zone that serves to deposit the sediments, nullifying their effect on the dam operation, but recent attempts at implementing this have mostly been rendered ineffective [18].

In their work, Pinarlik et al. [18] suggest that the sedimentation effect may be negligible (in Türkiye) within a time frame lower than 50 years. Other authors consider sedimentation as a fixed reduction in total capacity, like Yaseen et al. [30]. If sedimentation is to be considered in the development of the model, then it could be represented as a simple linear or quadratic formula to represent the net inflow of sediments, or as an external given parameter varying in time, or just the fixed reduction in total capacity. Of the three possible representations, the second is the most accurate, the third is the most simple to implement and the first is an inbetween solution.

2.3 Multiple Dam Systems

Dam operation management can be considered an open-loop or closed-loop problem [26]. In an open-loop problem, there is an external input that goes through some function/algorithm to calculate the output. In a closed-loop problem, besides the external input, the output of the function/algorithm in the previous iteration also serves as an input (control input), directly influencing the output of the next iteration. In the context of dam operation management, a system with a multiple interconnected dams is considered closed-loop, as downstream dam directly receives the water released by an its upstream dam.

In a watershed with multiple dams, theses can be in series along a river (cascaded dams), or in parallel across its tributaries, or a combination of both. This opens the possibility to pump the water to a reservoir from a downstream dam, allowing for even more optimized water management. Furthermore, cascaded dams are beneficial to flood control because it allows using storage volumes efficiently [31]. In their work, Sampaio et al. [32] define that the actions of generating electricity, pumping water and doing nothing are mutually exclusive.

2.4 Critical Synthesis and Analysis

Table 2.1 maps a selection of works that deal with dam operations according to the different parameters, constraints and decision variables considered. If a work has an "x" or "s" on a certain parameter, that means that this parameter is used in their model. In the case of "s", that parameter is considered to be a random (stochastic) variable. This table serves to have a better understanding of the state-of-the-art and to help visualize where there might exist a literature gap.

The "Nr" column represents the number of dams considered in the model, where n means that the model has the capacity to scale independently of the number of dams, and c means that the dams are in cascade (in series). For example, the first work, from C. C. Wei and N. S. Hsu [26], having a "Nr" of 2 c means that in their work they used two dams in cascade (in the same river). The "Pump" column represents whether that model used Pumped Hydroelectric power-plants or not, that is, if the dams had the possibility of pumping water from a downstream reservoir. The "Natural inflow" column represents the natural inflow of the river upstream of the dam. This can have an "x" or an "s", where "x" just means that this parameter is taken into account and the "s" means that, on top of being used, it is used stochastically (treated like a random variable with an average and variance). The "Installed capacity" column represents which models use the energy capabilities of the dam(s). Maximizing energy production is usually one of the objectives in the models in this column. The "Evaporation" and "Precipitation" columns represent whether the weather effects like evaporation (related with temperature and pressure) and precipitation were taken into account. Finally, the "Energy cost" column represents if the energy cost by time unit was taken into account.

Table 2.1: Situations and operational constraints considered in the most relevant models in the state-of-the-art. "x" and "s" mean the works considered said parameter (stochastically in "s")

Works	Number of dams	Possibility to pump water between dams	Natural inflow constraints	Installed capacity constraints	Evaporation	Precipitation	Energy cost
R. Dittmann et al. [28] (2009)	3		x				
C. C. Wei and N. S. Hsu [26] (2009)	2 c		x				
Y. Zhou and S. Guo [22] (2013)	1			x			
G. S. Sampaio et al. [32] (2013)	n	x	x	x			x
Z. kai Feng et al. [33] (2018)	n c		x	x			
M. Ak et al. [10] (2019)	n c	x		x			
Z. Yaseen et al. [30] (2019)	2 c		x		x		
S. K. Ahmad et al. [24] (2020)	1		s	x			
H. Liu et al. [5] (2020)	1		s	x			
M. Pinarlik et al. [18] (2021)	1		x		x	x	
V. Lai et al. [34] (2022)	1		x	x	x		

Many trivial parameters like minimum and maximum water levels in the reservoir, and minimum and maximum discharge, among others, were omitted from this table because most of the works used them and in a similar way.

From the Table 2.1 it is possible to observe that, despite having multiple dams in their models, only M. Ak et al. [10] and G. S. Sampaio et al [32] used the pumping mechanism of transferring water from a downstream dam to another. The other authors do not reference this mechanism in their articles and not all dams have this mechanism. As seen before when this mechanism is available, it provides space for optimization and it is generally positive to include it in the model.

C. C. Wei and N. S. Hsu [26] and M. Ak et al. [10] did not use information on the natural inflow. Instead, they guided their model solely on the current water level of the dam. This approach has the advantage of not having to specify when or in what way water can be deposited or taken away from the reservoir. The disadvantage is that it is very limiting in terms of the ability to prepare for a future event. H. Liu et al. [5] and S. K. Ahmad et al. [24] on the other hand, not only considered natural inflow but also considered it of stochastic nature, i.e., with an average and deviation per time slot. These works only regard a single dam. In a cascaded dam system, at least half of the dams would have very good information about their inflow because it would be similar to the amount of water its upstream dam let through.

It is possible to see that few authors consider the weather conditions (evaporation and precipitation), and even fewer consider the energy market in their models. No one has, in their work, considered the weather conditions and the energy market in the same work.

Chapter 3

Mathematical Model

The goal of this model is to optimize the integrated operations within a watershed, comprising different connected dams. The interconnection allows to optimize the power production decisions depending on current and future market prices for energy, expected weather conditions, and ecological requirements, among others. The time is represented by discrete periods (of, e.g., 30 minutes). For each time period and each dam, the model can control the power generation (by driving water through the turbines), discharge water, and pump water back from downstream dams (if the infrastructure allows it).

Through a systematic representation of the watershed by a network and different types of connection points, the model considers operational constraints (related to the infrastructure, and water volumes, among others) and ecological requirements and has the goal of maximizing profit within the specified time horizon. The systematic representation of the watershed is built on two concepts: dams, which produce energy, discharge and pump water, and virtual points where the water flow is controlled. Dams can have either continuous or non-continuous downstream flow. In dams with continuous flow, water driven through turbines exits downstream of the dam along with the discharged water. In dams with non-continuous flow, the water is directed to a further downstream point in the river, and only discharged water flows immediately downstream of the dam. Virtual points are points of interest where flows of water might converge to, diverge from, or just pass by. Virtual points can be divided into the following six categories:

- Virtual points that represent the start of the main river and of the tributaries (secondary rivers that debouch into the main river). Their flow throughout time is a given parameter, resulting from the upstream system.

- Virtual points that represent the flow immediately upstream of a dam, thus leading to the volume of water entering the dam at each time period. Some dams are able to pump water back from a downstream dam using the engine that generates power, for a given cost. In these dams, the upstream flow is increased by this amount of pumped water.

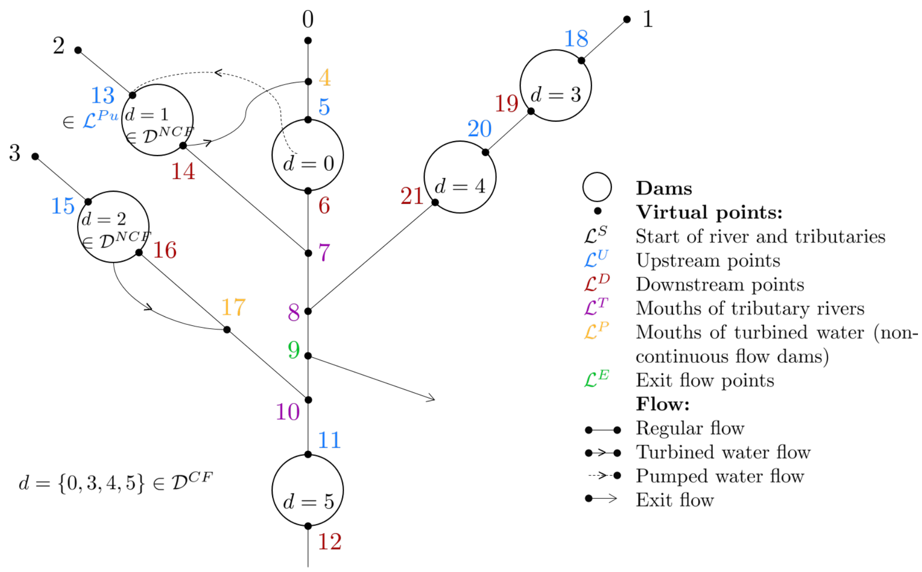


Figure 3.1: Example of a watershed with the systematic representation of dams and flow control virtual points. The notation used in this Figure is introduced later in this Chapter.

- Virtual points that represent the flow immediately downstream of a dam. This flow includes discharged water and, in dams with continuous downstream flow, the water sent through turbines to generate power.
- Virtual points that represent the mouth of the tributary river into the main river. These are points of confluence of flows.
- Virtual points that represent the confluence of the flow of water used to generate power in a non-continuous dam and the river (main or tributary).
- Virtual points that represent points in the river (main or tributary) where there is an outflow of water to serve other uses (e.g., agriculture). This outflow is a given parameter in the model, changing with time.

Figure 3.1 demonstrates an example of a representation of a watershed with three tributary rivers and six dams (two with non-continuous flow).

3.1 Assumptions

The following assumptions were considered when defining the model:

- There is at least one dam in the watershed.
- The effect of tides is not felt in the most downstream virtual point (or any other point).
- Water flows take an integer number of periods to flow between two virtual points.
- There are no residual losses or increments of water between two virtual points.

- The time it takes to start pumping, discharging, and driving water through the turbines is negligible.
- The cost of pumping water is that of the time period when the decision to pump was enacted, independently from how much time the pumped water takes to reach the new upstream dam.

3.2 Indices and Parameters

The following parameters are divided into watershed model parameters, nature parameters, dam specific parameters and operational parameters.

Watershed structure model:

$\mathscr{T}$	Set of time periods in the time-horizon.
$\mathscr{D}$	Set of dams in the watershed.
$\mathscr{D}^{CF} \subset \mathscr{D}$	Set of dams with continuous flow of turbined water, i.e. dams whose quantity of water that is driven through the turbines to generate power returns to the river course immediately in the downstream flow.
$\mathscr{D}^{NCF} \subset \mathscr{D}$	Set of dams with non-continuous flow of turbined water, i.e. dams whose quantity of water that is driven through the turbines to generate power only returns to the river course in the following virtual point (with $\mathscr{D}^{NCF} \cup \mathscr{D}^{CF} \equiv \mathscr{D}$).
$\mathscr{D}^{PU} \subset \mathscr{D}$	Set of dams that can pump water back to a previous dam.
$\mathscr{L}$	Set of virtual points of flow control, categorised into six mutually exclusive types, such that $\mathscr{L} \equiv \mathscr{L}^S \cup \mathscr{L}^U \cup \mathscr{L}^D \cup \mathscr{L}^T \cup \mathscr{L}^P \cup \mathscr{L}^E$.
$\mathscr{L}^S \subset \mathscr{L}$	Set of virtual points representing the start of the main river and tributaries.
$\mathscr{L}^U \subset \mathscr{L}$	Set of virtual points immediately upstream of a dam.
$\mathscr{L}^{Pu} \subset \mathscr{L}^U$	Subset of virtual points that are immediately upstream of a dam that receives pumped water from another dam.
$\mathscr{L}^D \subset \mathscr{L}$	Set of virtual points immediately downstream of a dam.
$\mathscr{L}^T \subset \mathscr{L}$	Set of virtual points where tributary rivers pour into the main river.
$\mathscr{L}^P \subset \mathscr{L}$	Set of virtual points where the flow that is driven through the turbines pours back into the river course, for non-continuous flow dams.
$\mathscr{L}^E \subset \mathscr{L}$	Set of virtual points where there is an exit flow for other uses.
$t \in \mathscr{T}$	Index for the time periods (e.g. 30-minute periods) in the time-horizon.
$d \in \mathscr{D}$	Index for the dams in the watershed.
$i, j \in \mathscr{L}$	Indexes for the set of virtual points of flow control throughout the main river and tributaries.
UP_d	Virtual point corresponding to the upstream point of dam d.

DW_d	Virtual point corresponding to the downstream point of dam d.
$Z_{d_1 d_2}$	Binary parameter = 1 if dam $d_1 \in \mathscr{D}^{PU}$ can pump water to dam $d_2 \in \mathscr{D}$ (actually, d_2 is the dam that controls the pumping mechanism and where the pump is located); $= 0$ otherwise.
W_i	Dam to which virtual point $i \in \mathscr{L}^U \cup \mathscr{L}^D \cup \mathscr{L}^P$ refers.
IP_{ij}	Binary parameter $= 1$ if point $j \in \mathscr{L} \setminus (\mathscr{L}^D \cup \{0\})$ is immediately preceded by point $i \in \mathscr{L}$; $= 0$ otherwise.
TB_{ij}	Time (in number of time periods) for a flow to move between point i and j.

Flow and weather parameters:

F_i^0	Steady-state flow (in m^3/s) at virtual point i.
F_d^E	Ecological flow (in m^3/s) to safeguard downstream of dam d.
R_{dt}	Expected rainfall affecting dam d between time periods t and $t+1$, i.e., additional volume of water from rainfall (in m^3).
RP_{it}	Expected rainfall affecting virtual point i between time periods t and $t+1$, i.e., additional flow of water from rainfall (in m^3/s).
OF_{it}	Expected exit flow (in m^3) in point $i \in \mathscr{L}^E$ at time t.

Dam parameters:

V_d^{ini}	Initial volume (in m^3) of dam d.
V_d^{max}	Maximum volume (in m^3) of dam d.
V_d^{min}	Minimum volume (in m^3) of dam d.
TV_d^{max}	Maximum volume (in m^3) to drive water through the turbines of dam d.
TV_d^{min}	Minimum volume (in m^3) to drive water through the turbines of dam d.
FP_d	Maximum flow (in m^3/s) allowed to drive water through the turbines of dam d.
FS_d	Maximum flow (in m^3/s) allowed to discharge water using a spillway of dam d.
FB_d	Maximum flow (in m^3/s) allowed to pump water from dam d.

Operational parameters:

C_t	Cost of energy at t (i.e., between t and $t+1$) in euros per MWh.
GEP_d	Energy generated in dam d when $1m^3$ of water is driven through the turbines to generate power.
GEM_{d_1,d_2}	Energy required when pumping $1m^3$ of water from dam d_1 to dam d_2.
FC_d^P	Fixed cost of starting to drive water through the turbines of dam d to generate power (virtual cost to reduce the number of starts).
FC_d^S	Fixed cost of starting to discharge water from dam d using a spillway (virtual cost to reduce the number of starts).

EP_d	Efficiency of the process of driving water through the turbines of dam d to generate power ($EP_d \in [0,1]$).
EB_d	Efficiency of the process of pumping water back from dam d ($EB_d \in [0,1]$).
AQ_{dk}	Percentage of the maximum flux allowed through the turbines and pumping tubes in level k in dam d ($AQ_{dk} \in [0,1]$).
$OC^{P/S}$	Opportunity cost of driving $1m^3$ of water to generate power (P) OR to discharge through the spillway (S).
Other:	
ST	Number of seconds within a time period.

3.3 Decision Variables

$q_{dt}^{P/S}$	Quantity of water (in m^3) leaving dam d at t (i.e., between t and $t+1$), driven through the turbines to generate power (P) OR discharged through the spillway (S).
$m_{d_1,d_2,t}$	Quantity of water (in m^3) pumped from dam d_1 to dam d_2 at t (i.e., between t and $t+1$).

Auxiliary decision variables:

v_{dt}	Volume of water (in m^3) in dam d at time t.
v_d^E	Volume of water (in m^3) in dam d at the end of the time horizon.
f_{it}	Flow (in m^3/s) in point i at time t.
$\alpha_{dt}^{P/S}$	$= 1$ if water is released in dam d at time t, through the turbines to generate power (P) or discharged through the spillway (S); $= 0$ otherwise.
$\beta_{dt}^{P/S}$	$= 1$ if dam d is starting to release water at time t, through the turbines to generate power (P) or discharged through the spillway (S); $= 0$ otherwise.
ω_{dtk}	$= 1$ if the level k of the turbines is used in dam d at time t; $= 0$ otherwise.
$\lambda_{d_1 d_2 tk}$	$= 1$ if the level k of the pump is used to pump water from dam d_1 to dam d_2 at time t; $= 0$ otherwise.
g_{dt}	$= q_{dt}^S$ (quantity of water leaving dam d trough the spillway) IF water is also released through the turbines in dam d at time t; $= 0$ otherwise.

3.4 Mathematical Model

$$\max \quad \sum_{d\in\mathscr{D}} EP_d \Big(\sum_{t\in\mathscr{T}} q_{dt}^P \cdot C_t \cdot GEP_d \Big) - \sum_{d_1\in\mathscr{D}^{PU}} \sum_{d_2\in\mathscr{D}} \Big(\sum_{t\in\mathscr{T}} C_t \cdot GEM_{d_1 d_2} \cdot m_{d_1 d_2 t} \Big)$$
$$- \sum_{d\in\mathscr{D}} \sum_{t\in\mathscr{T}} \Big(FC_d^P \beta_{dt}^P + FC_d^S \beta_{dt}^S \Big) - \sum_{d\in\mathscr{D}} \sum_{t\in\mathscr{T}} \Big(OP^P \cdot q_{dt}^P + OP^S \cdot q_{dt}^S \Big) \qquad (3.1)$$

$$\textbf{s.t. } q_{dt}^{P/S} \leq \alpha_{dt}^{P/S} V_d^{max} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.2}$$

$$\alpha_{dt}^{P/S} \leq q_{dt}^{P/S} V_d^{max} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.3}$$

$$\beta_{dt}^{P/S} \geq \alpha_{dt}^{P/S} - \alpha_{d,t-1}^{P/S} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \setminus \{0\} \tag{3.4}$$

$$\beta_{d0}^{P/S} \geq \alpha_{d0}^{P/S} \qquad \forall d \in \mathscr{D} \tag{3.5}$$

$$\sum_{d_1 \in \mathscr{D}^{PU}: Z_{d_1 d}=1} m_{d_1 dt} \leq (1-\alpha_{dt}^{P}) V_d^{max} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.6}$$

$$q_{dt}^{P} + q_{dt}^{S} + \sum_{d_2 \in \mathscr{D}: Z_{d,d_2}=1} m_{d,d_2,t} \leq v_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.7}$$

$$v_{d0} = V_d^{ini} \qquad \forall d \in \mathscr{D} \tag{3.8}$$

$$v_{dt} = v_{d,t-1} - q_{d,t-1}^{P} - q_{d,t-1}^{S} + f_{UP_d,t-1} ST + R_{d,t-1} \qquad \forall d \in \mathscr{D} \setminus \mathscr{D}^{PU}, t \in \mathscr{T} \setminus \{0\} \tag{3.9}$$

$$v_{dt} = v_{d,t-1} - q_{d,t-1}^{P} - q_{d,t-1}^{S} + f_{UP_d,t-1} ST + R_{d,t-1} - \sum_{d_2 \in \mathscr{D}} m_{d,d_2,t-1} \qquad \forall d \in \mathscr{D}^{PU}, t \in \mathscr{T} \setminus \{0\} \tag{3.10}$$

$$v_d^E = v_{d|\mathscr{T}|} - q_{d|\mathscr{T}|}^{P} - q_{d|\mathscr{T}|}^{S} + f_{UP_d|\mathscr{T}|} ST + R_{d|\mathscr{T}|} \qquad \forall d \in \mathscr{D} \setminus \mathscr{D}^{PU} \tag{3.11}$$

$$v_d^E = v_{d|\mathscr{T}|} - q_{d|\mathscr{T}|}^{P} - q_{d|\mathscr{T}|}^{S} + f_{UP_d|\mathscr{T}|} ST + R_{d|\mathscr{T}|} - \sum_{d_2 \in \mathscr{D}} m_{d,d_2,|\mathscr{T}|} \qquad \forall d \in \mathscr{D}^{PU} \tag{3.12}$$

$$v_{dt} \geq V_d^{min} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.13}$$

$$v_{dt} \geq TV_d^{min} \alpha_{d,t-1}^{P} - g_{d,t-1} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \setminus \{0\} \tag{3.14}$$

$$v_{dt} \leq V_d^{max} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.15}$$

$$v_{dt} \geq TV_d^{min} \alpha_{dt}^{P} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.16}$$

$$v_{dt} \leq TV_d^{max} + V_d^{max}(1-\alpha_{dt}^{P}) \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.17}$$

$$v_d^E \geq V_d^{min} \qquad \forall d \in \mathscr{D} \tag{3.18}$$

$$v_d^E \geq TV_d^{min} \alpha_{d,|\mathscr{T}|}^{P} - g_{d,|\mathscr{T}|} \qquad \forall d \in \mathscr{D} \tag{3.19}$$

$$v_d^E \leq V_d^{max} \qquad \forall d \in \mathscr{D} \tag{3.20}$$

$$q_{dt}^{P} \leq FP_d \cdot ST \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.21}$$

$$q_{dt}^{S} \leq FS_d \cdot ST \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.22}$$

$$m_{d_1 d_2 t} \leq FB_{d_2} \cdot ST \qquad \forall d_1 \in \mathscr{D}^{PU}, d_2 \in \mathscr{D}, t \in \mathscr{T} \tag{3.23}$$

$$g_{dt} \geq 0 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.24}$$

$$g_{dt} \leq ST \cdot FS_d \cdot \alpha_{dt}^{P} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.25}$$

$$g_{dt} \leq q_{dt}^{S} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.26}$$

$$g_{dt} \geq q_{dt}^{S} - ST \cdot FS_d (1-\alpha_{dt}^{P}) \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.27}$$

$$f_{it} = F_i^0 + RP_{it} \qquad \forall i \in \mathscr{L}^S, t \in \mathscr{T} \tag{3.28}$$

$$f_{it} = \frac{q^S_{W_i t}}{ST} + RP_{it} \qquad \forall i \in \mathscr{L}^D : W_i \in \mathscr{D}^{NCF}, t \in \mathscr{T} \tag{3.29}$$

$$f_{it} = \frac{q^S_{W_i t} + EP_{W_i} \cdot q^P_{W_i t}}{ST} + RP_{it} \qquad \forall i \in \mathscr{L}^D : W_i \in \mathscr{D}^{CF}, t \in \mathscr{T} \tag{3.30}$$

$$f_{it} \geq F^E_{W_i} \qquad \forall i \in \mathscr{L}^D, t \in \mathscr{T} \tag{3.31}$$

$$\begin{aligned} f_{it} = & \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} \geq 0}} f_{j,t-TB_{ji}} + \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} < 0}} F_j^0 + RP_{it} \\ & + \sum_{d_1 \in \mathscr{D}^{PU}: Z_{d_1 W_i}=1} \frac{EB_{W_i} \cdot m_{d_1, W_i, t-TB_{DW_{d_1}, i}}}{ST} \qquad \forall i \in \mathscr{L}^{Pu}, t \in \mathscr{T} \end{aligned} \tag{3.32}$$

$$f_{it} = \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} \geq 0}} f_{j,t-TB_{ji}} + \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} < 0}} F_j^0 + RP_{it} \qquad \forall i \in \Big(\big(\mathscr{L}^U \setminus \mathscr{L}^{Pu} \big) \cup \mathscr{L}^T \Big), t \in \mathscr{T} \tag{3.33}$$

$$\begin{aligned} f_{it} = & \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} \geq 0}} f_{j,t-TB_{ji}} + \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} < 0}} F_j^0 \\ & - OF_{it} + RP_{it} \qquad \forall i \in \mathscr{L}^E, t \in \mathscr{T} \end{aligned} \tag{3.34}$$

$$\begin{aligned} f_{it} = & \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} \geq 0}} f_{j,t-TB_{ji}} + \sum_{\substack{j \in \mathscr{L}: IP_{ji}=1, \\ t-TB_{ji} < 0}} F_j^0 + RP_{it} \\ & + \frac{EP_d \cdot q^P_{d,t-TB_{DW_d, i}}}{ST}, \text{ with } d = W_i \qquad \forall i \in \mathscr{L}^P, t \in \mathscr{T} \end{aligned} \tag{3.35}$$

$$\sum_{k \in \mathscr{K}} \omega_{dtk} = 1 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.36}$$

$$q^P_{dt} = \sum_{k \in \mathscr{K}} \Big(\omega_{dtk} \cdot AQ_{dk} \cdot FP_d \cdot ST \Big) \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.37}$$

$$\sum_{k \in \mathscr{K}} \lambda_{d_1 d_2 tk} = 1 \qquad \forall d_1 \in \mathscr{D}^{PU}, d_2 \in \mathscr{D}, t \in \mathscr{T} \tag{3.38}$$

$$m_{d_1 d_2 t} = \sum_{k \in \mathscr{K}} \Big(\lambda_{d_1 d_2 tk} \cdot AQ_{d_2 k} \cdot FB_{d_2} \cdot ST \Big) \qquad \forall d_1 \in \mathscr{D}^{PU}, d_2 \in \mathscr{D}, t \in \mathscr{T} \tag{3.39}$$

$$q^{P/S}_{dt} \geq 0 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.40}$$

$$m_{d_1 d_2 t} \geq 0 \qquad \forall d_1 \in \mathscr{D}^{PU}, d_2 \in \mathscr{D}, t \in \mathscr{T} \tag{3.41}$$

$$v_{dt} \geq 0 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.42}$$

$$f_{it} \geq 0 \qquad \forall i \in \mathscr{L}, t \in \mathscr{T} \tag{3.43}$$

$$\alpha^{P/S}_{dt} \in \{0,1\} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.44}$$

$$\beta^{P/S}_{dt} \in \{0,1\} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{3.45}$$

The objective function (Equation 3.1) maximizes the profit of the water driven through turbines to generate power, discounting the cost of pumping water back to a different dam. The third component considers fixed costs for starting to discharge water and to drive water through turbines. These are virtual costs to reduce the number of times a dam has to start discharging water or generating power. The last component sets an opportunity cost that is lost with every m^3 of water discharged or driven through the turbines in the time horizon. This lets the model know if it is worth it to store water now to discharge/generate power later or not. For example, if the opportunity cost is 120 euros/MWh and the cost of energy of a given time period is 90 euros/MWh, then the model is discouraged from generating power/discharging at that time period because it would be losing those 30 euros/MWh difference, had it held on to that water for later. Further in the paper, how to define this opportunity cost will be discussed.

Equations 3.2 and 3.3 guarantee the desired relation between $q_{dt}^{P/S}$ and $\alpha_{dt}^{P/S}$, for any given time period t and dam d. Equations 3.2, 3.4 and 3.5 allow to count the number of "starts" for discharging or generating power so that they can be included in the objective function. Equations 3.6 ensures that a dam never pumps water (i.e., receives water from a downstream dam) while driving water through the turbines to generate power at the same time.

Equations 3.7 ensure that the quantity of water discharged, used to generate power, or pumped back to a previous dam is limited by the volume of water available at a dam. For this, Equations 3.8 to 3.10 calculate the volume of water in a dam at each point in time. Equations 3.11 and 3.12 make the same calculation for the time period after the time horizon to ensure the continuity of the results. The water volume must be within bounds at each time period (Equations 3.13 and 3.15).

Equations 3.18 and 3.20 do the same thing, but for the time period after the time horizon. They serve to prevent end-of-horizon effects. This is when decision variables at the time period after the time horizon are not bounded, so the solution goes out of bounds as soon as the time horizon ends. For example, without those equations, a dam could discharge all the water in the last time period, leaving the volume of the next time period (already out of the time horizon) below the minimum allowed.

Also, the volume must be within specific bounds for power to be generated in a dam (Equations 3.16 and 3.17). Equations 3.14 do not allow for power generation with the water beyond the minimum volume to do so. Equations 3.19 do the same, but for the time period after the time horizon to prevent end-of-horizon effects.

Equations 3.24 to 3.27 define the non-linear properties of g_{dt} trough linear equations: g_{dt} is defined as $g_{dt} = q_{dt}^S \cdot \alpha_{dt}^P$, but this equation is non-linear because it multiplies two decision variables. To ensure that the model remains linear (and, consequently, simpler to solve), Equations 3.24 to 3.27 were created. (For more details consult the work of Bisschop J. "AIMMS Optimization Modeling" [35], page 83-85, Chap 7.7)

Equations 3.28 to 3.35 allow to calculate and control the flow of water in each of the virtual points, thus managing the water entering and exiting the dams and the watershed system. The flow in the points that represent the start of the main and tributary rivers in the scope considered is a given value (Equations 3.28). As for the points downstream of a dam, their flow results from the discharged water if the dam is non-continuous (i.e., if the water that is used to generate power does not flow immediately downstream but to a different point), as described in Equations 3.29. In the opposite case, the downstream flow includes also the water used to generate power (Equations 3.30). In those points, it is important to ensure that the flow is at least the value required for the ecological sustainability of the water system (Equations 3.31). In the remaining virtual points, part of the flow results from the flow arriving from previous points. In the case of upstream points whose dams also pump water from downstream dams, the flow coming from previous points is increased by the pumped water (Equations 3.32). In the case of other upstream points, as well as of points that represent the mouth of a tributary river into the main river, the flow is simply the flow coming from previous points (Equations 3.33). Some points represent outflows to fulfill other uses. The flow in those points is decreased by the amount of water that exits the watershed (Equations 3.34). Finally, some points represent the mouth of the main or a tributary river of water that is driven through the turbines of a non-continuous dam to generate power. In those points, the previous flow is increased by this amount of water (Equations 3.35).

Equations 3.36 and 3.37, and 3.38 and 3.39 describe the discrete (non-continuous) nature of generating power and pumping, respectively, in any given period. Equations 3.36 and 3.38 dictate that, at any time t and dam d, only one level of turbines and pumps can be active, respectively. Equations 3.37 and 3.39 define the quantity of water that goes through the turbines and pumps by multiplying the percentage of the maximum throughput of the active level with the maximum throughput allowed in the turbines and pumps, respectively (at any time t and dam d).

Equations 3.40 to 3.45 represent the domain of the decision variables and ensure non-negativity.

Chapter 4

Solution Method

Due to the complexity of the model and the necessity to arrive at good (but not necessarily optimal) solutions, an heuristic solution method is proposed. This chapter will present the Relax-and-Fix algorithm, its implementation, and the reasoning for its choice and use.

4.1 Algorithm Selection

Due to the complexity of the mathematical model, an alternative method to find near-optimal solutions within a reasonable execution time was proposed. After some deliberation, two approaches were seen as the most interesting to be implemented: the Relax-and-Fix algorithm (R&F) and a metaheuristic, such as the Biased Random-Key Genetic Algorithm (BRKGA). Relax-and-Fix is normally used in lot-sizing and scheduling problems, which have a strong temporal component, especially in MIP problems. In the work of K. A. G. Araujo et al. [36], this approach delivered the best results when implemented chronologically or in decreasing order of "criticality". In their work, A. Roshani et al. [37] also achieved good results when iterating through time periods. Metaheuristics like BRKGA and genetic algorithms, on the other hand, are more generalized approaches to optimization problems. To solve MIPs with this approach, one has to encode the inputs into a genome and verify that the output is within the boundaries of the constraints. The former can be as simple as a 1-to-1 correlation, but this is rarely a good option, as one advantage of using this approach is precisely to encode this correlation in a clever way, reducing the number of inputs and with that, the execution time. The latter is the hardest challenge and the most common way to deal with this issue is to make the objective function heavily penalize any solution that does not oblige with the constraints [38]. Ultimately, the approach chosen was the Relax-and-Fix algorithm due to the importance of the time component in this problem and the fact that it uses the structure of the mathematical model, which was also implemented.

4.2 Relax-and-Fix

The Relax-and-Fix approach is a heuristic method used to solve optimization problems. As previously stated, it is particularly effective when dealing with complex problems that have a clear temporal component.

In the Relax-and-Fix approach, the optimization problem is divided into several, smaller, MILP subproblems, that are solved iteratively. A moving window is defined and in each subproblem, every integer and binary variable **after** the window is **relaxed** to real variables (with the same bounds as before), every variable **before** the window is **fixed** to the value optimized for a previous iteration (subproblem solution), and those inside the window are kept unchanged. The process of solving one of these subproblems is called iteration in the remainder of the paper. After an iteration, the variables within the bounds of the window become fixed to the value of the solution just found, the window slides a certain distance (with the maximum distance allowed being the size of the window itself), and a new iteration begins. Variables that are fixed remain constant throughout the subsequent iterations, while other variables can be modified to improve the solution. Fixing variables narrows down the search space and guides the optimization algorithm towards a solution in less execution time, although not necessarily to the optimal solution of the global problem. There is also the possibility to allow overlap, i.e., to make the distance traveled (step) by the window smaller than the window itself. For example, if the window has a size of 4 time periods, and the distance traveled (step) is only of 3 periods, there will be an overlap size of 1 period that will not be relaxed, but will also not be fixed in the current iteration, only in the next one (in this case). A pseudo-code of the Relax-and-Fix implemented is illustrated in Algorithm 1, while Figure 4.1 illustrates the scenario described above.

Algorithm 1 Pseudo-code of the Relax-and-Fix heuristic implemented

```
aux ← InitWindow(fix_window_size)
counter ← 0
while counter < time_horizon do
    model ← model + NewModel(objective_function)
    model ← model + CreateDecisionVariables(counter, window_size)
    model ← model + CreateConstraints(counter, window_size)
    model.integers[counter, counter + window_size] ← IntegerType
    model ← model + AddConstraint(model.integers[0, counter] = aux[0, counter]
    Solve(model)
    aux[0, counter + fix_window_size] ← model.integers[0, counter + fix_window_size]
    if counter + window_size >= time_horizon then
        return : model
    else
        counter ← counter + fix_window_size
    end if
end while
```

In the example in Figure 4.1, the window size is of 4 periods, 3 fixable and 1 non-fixable. This shows the R&F algorithm in the third iteration, where rows 1-6 are already fixed, rows 11-24

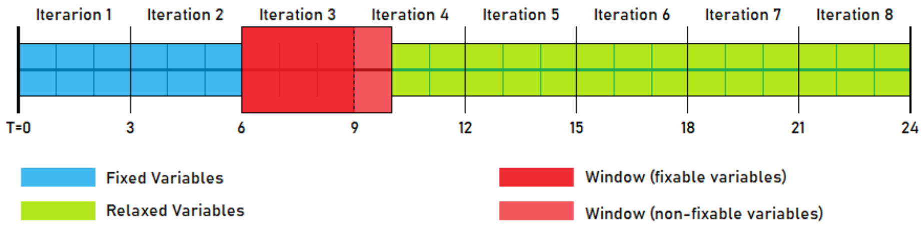


Figure 4.1: Visualization of the Relax-and-Fix heuristic

are relaxed, and rows 7-10 are in the window, where only rows 7-9 will be fixed after the current iteration.

4.2.1 Variants and Implementation

With this heuristic, two variants of the R&F heuristic are proposed: Full R&F and Partial R&F. In Full R&F, every decision variable is, effectively, relaxed, i.e., they were either relaxed or omitted from the constraints (omitting ω causes q^P to be relaxed, without changing its maximum and minimum values). In Partial R&F, only $\alpha^{P/S}$ and $\beta^{P/S}$ are not relaxed, therefore, in the relaxed time periods, the model behaved exactly as a simplified model (not using Equations 3.36 to 3.39) in MILP would if it was possible to use a continuous quantity of water for generating power (q^P) and pumping (m). As instances grew in complexity, the execution times grew exponentially, and this simplified model (without Equations 3.36 to 3.39) also saw its execution times rise, making Partial R&F solutions not desirable for its high execution times. To combat this sharp rise in execution times, the Full R&F variant was developed as well.

In both cases, a few common changes to the mathematical model were made to implement the Relax-and-fix algorithm. Beginning with sets, the following were created:

$\mathscr{T}^F \subset \mathscr{T}$	Set of time periods in the Fixed zone (the decision variables within are already fixed).
$\mathscr{T}^R \subset \mathscr{T}$	Set of time periods in the Relaxed zone (the decision variables within are still relaxed).
$\mathscr{T}^W \subset \mathscr{T}$	Set of time periods in the Window (the decision variables within are not relaxed not fixed). $\mathscr{T}^W \equiv \mathscr{T}^{FW} \cup \mathscr{T}^{NFW}$.
$\mathscr{T}^{FW} \subset \mathscr{T}^W$	Set of time periods in the Fixable part of the Window (the decision variables within will be fixed after the current iteration).
$\mathscr{T}^{NFW} \subset \mathscr{T}^W$	Set of time periods in the Non-Fixable part of the Window (the decision variables within will not be fixed after the current iteration).

And a more appropriate definition for the set $\mathscr{T}$ would be:

$\mathscr{T}$ Set of time periods in the time-horizon, categorised into three mutually exclusive types, such that $\mathscr{T} \equiv \mathscr{T}^F \cup \mathscr{T}^W \cup \mathscr{T}^R$.

Heading to the parameters, the following were added:

Fq^P_{dt} Values of all previously fixed q^P_{dt}.

$Fm_{d_1,d_2,t}$ Values of all previously fixed $m_{d_1,d_2,t}$.

Finally, Equations 3.36 to 3.39 were changed from $t \in \mathscr{T}$ to $t \in \mathscr{T}^W$, effectively relaxing q^P and m at $t \in \mathscr{T}^R$. At $t \in \mathscr{T}^F$, Equations 4.2 and 4.1 were created to fix q^P and m.

$$q^P_{dt} = Fq^P_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T}^F \tag{4.1}$$

$$m_{d_1,d_2,t} = Fm_{d_1,d_2,t} \qquad \forall d_1 \in \mathscr{D}^{PU}, \forall d_2 \in \mathscr{D}, t \in \mathscr{T}^F \tag{4.2}$$

These changes are sufficient to implement the Partial R&F. To implement the Full R&F, more parameters are needed:

$F\alpha^{P/S}_{dt}$ Values of all previously fixed $\alpha^{P/S}_{dt}$.

$F\beta^{P/S}_{dt}$ Values of all previously fixed $\beta^{P/S}_{dt}$.

Some decision variables were also changed and others added; $\alpha^{P/S}$ and $\beta^{P/S}$ were transferred to the real domain (instead of binary), and auxiliary variables were created to help bring $\alpha^{P/S}$ and $\beta^{P/S}$ back to binary once inside the Window:

$aux\alpha^{P/S}_{dt}$ Binary variable to help fix $\alpha^{P/S}_{dt}$.

$aux\beta^{P/S}_{dt}$ Binary variable to help fix $\beta^{P/S}_{dt}$.

Equations 3.44 and 3.45 were discarded, giving way to the following new constraints:

$$\alpha^{P/S}_{dt} \geq 0 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{4.3}$$

$$\beta^{P/S}_{dt} \geq 0 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{4.4}$$

$$\alpha^{P/S}_{dt} \leq 1 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{4.5}$$

$$\beta^{P/S}_{dt} \leq 1 \qquad \forall d \in \mathscr{D}, t \in \mathscr{T} \tag{4.6}$$

$$\alpha^{P/S}_{dt} = aux\alpha^{P/S}_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T}^W \tag{4.7}$$

$$\beta^{P/S}_{dt} = aux\alpha^{P/S}_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T}^W \tag{4.8}$$

$$\alpha^{P/S}_{dt} = F\alpha^{P/S}_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T}^F \tag{4.9}$$

$$\beta^{P/S}_{dt} = F\alpha^{P/S}_{dt} \qquad \forall d \in \mathscr{D}, t \in \mathscr{T}^F \tag{4.10}$$

Equations 4.3 to 4.6 bound the variables to [0,1], Equations 4.7 to 4.8 force the variables to be binary while inside the Window, and Equations 4.9 to 4.10 fix, in the current iteration, the previously optimized values of $\alpha^{P/S}$ and $\beta^{P/S}$.

4.3 Programming and Data Preparation

The model was implemented in PuLP, and the R&F variants were implemented on top of the model (also in PuLP). PuLP is an open-source linear programming (LP) modeling package in Python. PuLP provides a high-level interface to create, solve, and analyze LP problems, and it allows the user to define LP problems by specifying decision variables, objective functions, and constraints concisely and intuitively. It supports a variety of solvers, including both open-source and commercial solvers, such as CBC, CPLEX, GLPK, Gurobi, and SCIP [39]. More detail about computational experiments will be given in the next chapter.

When implementing the model in PuLP, the first step, and probably the hardest one, was to figure out the best way to encode all the parameters into files, so that the program could read the parameters from files, instead of being hard-coded. This may seem like an easy task, but, because there are at least 23 non-mutually exclusive parameters, it would be very inefficient to make 23 different files for a single instance (an instance, in this case, is an imaginary or real river basin with certain river flows, and energy costs, that the method will solve for the corresponding objective function; more detail on instances on the next section). So, the parameters were divided into these 10 files:

1. The first file is structured as a matrix, where each row corresponds to a dam and each column represents, from left to right, the upstream point, the downstream point, the point to which the water flows if it passes through the turbines, the ecological flow, the initial volume, the maximum volume, the minimum volume, the maximum volume to drive water through the turbines, the minimum volume to drive water through the turbines, the maximum water flow (in m^3/s) allowed to drive through the turbines, to pump, and to discharge, the fixed cost off starting to drive water through the turbines (it is the same for pumping), and the fixed cost of starting to discharge water through the spillway, the efficiency of driving water through the turbines, and the efficiency of pumping water to an upstream dam, and the various fractions of the maximum water flow that are allowed in that dam (in pumping and driving water through the turbines).

2. The second file is also structured as a matrix where each row corresponds to a different set of virtual points according to the model. For example, if the first row contains the numbers 0 and 7, that means These points are parts of the set $\mathscr{L}^S$, meaning they represent the start of the main river or a tributary. Another example is if the fourth row contains the numbers 2, 4, and 6, these points are part of the set $\mathscr{L}^D$, meaning they represent points immediately downstream of a dam.

3. The third file is also structured as a matrix and is like the last one. It serves to identify which dams are part of which subset, The rows also being in the same order as the model.
4. The fourth file is structured as an array, where each row represents the energy cost for that period. Through this file, the program also assesses how many periods there are.
5. The fifth file is structured as a matrix, wherein in each row, the first column informs what point the following information is relevant to and the second column is the flow in said point in meters cubed per second.
6. The sixth file is structured as a matrix, where each row represents a point, the first column identifies the point, and the following columns represent the exit flow that will happen at that point in the time corresponding to the number of the column plus one. This file only has points from the set LE.
7. The seventh file is structured as a matrix, wherein in each row there is a connection between two points, the water flows from the point in column 1 to the point in column 2 and it takes the number in column three periods of time to do that journey. The fourth column is 1 if it is a real connection, or 0 if it is a fake connection, i.e., if water is pumped/driven through the turbines for non-continuous flow (0 for both cases).
8. The eighth file is structured as a matrix, where each row represents a pump connection. The first column represents the dam from which the water is pumped, the second column represents the point to which the pumped water flows end the third column is the dam that receives the pumped water.
9. The ninth file is structured as a matrix, where is row represents A dam. The first column has the dam identifier end the following columns Represent the rain that will occur and that dam in the time corresponding to the number of the column plus one, just like the sixth file.
10. Finally, the tenth file is like the ninth one but refers to each virtual point instead of each dam.

This information makes it possible to build several variables to contain the parameters. Next is the creation of the decision variables with the correct dimensions, type (integer, continuous or binary), and bounds, definition of the objective function, and writing of all the constraints. After solving, the terminal prints interesting variables for debugging, such as the volumes in each dam, the flows at every point, the amounts of water that went through the turbines, the amount that was pumped, and the amount that was discharged. There is an option to save for a file the value of the objective function, the execution time, and some R&F parameters, including Step and Overlap.

When solving with the Relax-and-Fix heuristic, in a given iteration, if ever $\mathscr{T} = \mathscr{T}^F \cup \mathscr{T}^W$, i.e., the Fixed zone and the Window occupy the entire time-horizon, then the current iteration is the last and the program is to stop after the current iteration is complete. This allows the program to have a lower execution time by not trying to find again the solution that was already in the window (at $t \in \mathscr{T}^{NFW}$).

Chapter 5

Computational Tests and Results

In this chapter, methodological tests will be performed to access the performance of the Relax-and-Fix approaches and compare them to the results obtained from MILP. Furthermore, the solution methods will be given several instances that emulate the Cávado watershed in real situations. The objective is to compare the results given by the solution methods with what actually happened in the real Cávado watershed and to extrapolate the overall behavior of the model, gaining insights into its advantages and limitations.

5.1 Test Environment

The mathematical model was implemented in Python (PyCharm Community Edition 2021.2.3) using the PuLP library and solved with CPLEX version 22.1.1.0. The Relax and Fix method was implemented on top of the mathematical model in Python. All computational experiments were conducted on an Asus X507UAR with 8 GB of RAM, and an Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz - 1.99 GHz processor, running on a Windows 10 operating system.

5.2 Test Instances

An instance is defined by its indices and parameters: virtual points, dams, connections between points, between dams, and between dams and points, initial state, exit flows, dam parameters, rain in dams and points, classification of points and dams in their respective sets/subsets, energy costs, time flows, and the number of seconds within a period. Two instances are only considered the same instance if their indices and parameters are equal. This means that two cloned instances are not considered the same if, for example, one of them has a different flow time between two virtual points or if their rains are different, despite everything else being the same.

To validate the model developed and implemented, a first smaller (dummy) instance was created. For example, one test was to put a certain amount of water in a dam in the form of rainfall and check if the volume of water in the dam at that time period increased by the same amount (excluding all the other factors that also make the water level vary). Same for the pumped water:

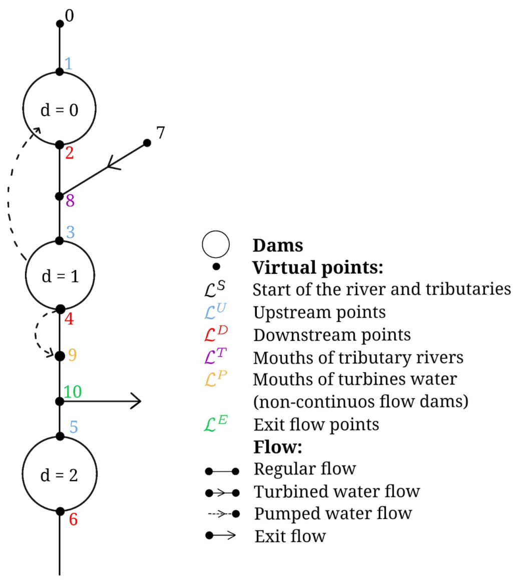


Figure 5.1: Structure of instance 1

the dam from which the water was pumped had to have that amount of water loss in its volume, and vice versa for the dam that received the pumped water, but only after the time that it took for the water to go from the downstream point of the first dam to the upstream point of the second dam.

The structure of an instance is the connection between virtual points, their connection with the dams, the connection between dams, and the classification of points and dams in their respective sets/subsets. The structure of this initial instance is shown in Figure 5.1. This structure was crafted as such to ensure that all the mechanics provided by the model were used simply, making it easy to understand the model in the various tests. To check if the equations for the flows were correctly written, the time of flows between each point was increased to the maximum number of periods while only leaving some connections with reasonable time flows. The test was the following: Forcing every connection to take 5 periods of time except for the connection between points 7 and 8, which only took one period of time. In this setup, the flows in point 8 were the same as the point 7 but with a delay of one period of time, until the 6th period when the flow was the sum of the flow in point 7 at the 5th period and the flow in point 2 at the first period, confirming both that the time flow between points was working and also that points like 8, that have more than one upstream point, were summing their flows as expected.

Several instances were created with this base structure to test the implementation of the model, but afterward, a more realistic instance was created with this structure in order to do some methodological tests like measuring the value of the objective function and the time of execution, called "instance 1". A second instance, called "instance 2", was developed to test the program in a bigger and more complex scenario, with more connections, more tributary rivers (each one with at least

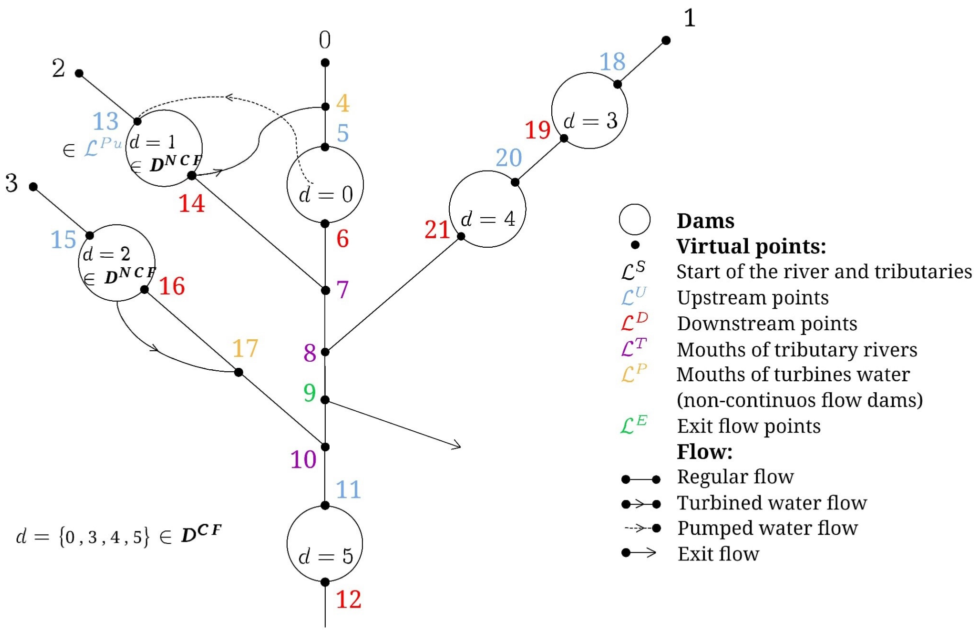


Figure 5.2: Structure of instance 2

one dam), and overall, less linear water flow. The structure of this instance is illustrated in Figure 5.2.

The third and last instance structure created was the Cávado one. As the name suggests, this instance structure represents the Cávado watershed, which will be the case-study later in the paper. This structure will lead to the creation of various instances. The structure of these instances is illustrated in Figure 5.3.

The instances that were created with this base structure are “instance 3”, “Cávado normal”, “Cávado drought”, “Cávado flood 1”, and “Cávado flood 2”. “Instance 3” was the first one created using this structure and it has a mix of real and simulated data as parameters. The other four instances represent real scenarios, with normal, drought, and 2 flood situations, which will be analyzed in detail in Section 5.4.

5.3 Methodological Tests

To assess the performance of the Relax-and-Fix heuristic and compare it to the MILP performance, some methodological tests were conducted. For the MILP, the tests consisted of measuring the value of the objective function and the execution time under different values of the opportunity cost, for instances 1, 2, and 3, while for the Relax-and-Fix heuristic, the tests also changed step and overlap. The variance in OC serves to see the difference in performance between when integer variables for pumping and generating power are more advantageous (low OC) and when continuous variables like for discharging are the advantageous ones (high OC). The variance in step and

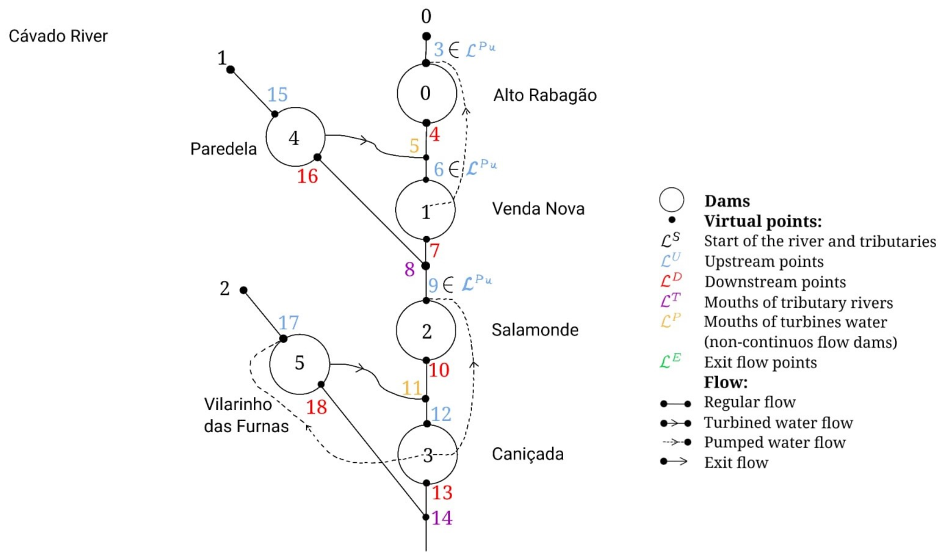


Figure 5.3: Structure of instance 3

overlap is considered to check which combinations lead to the best performance.

In these tests, the opportunity cost for water driven through the turbines ranges from the minimum to the maximum cost of energy in the each time horizon. This minimum-maximum range is 50-110 euros/MWh for instances 1 and 2, and 120-170 for instance 3. The values for the energy cost were all simulated, but they followed the normal variance of energy costs during the day (high costs during lunch and dinner time, and low costs at around 3 a.m.). The opportunity costs for discharged water was always 0.0001% of the OC for power generation. This cost does not penalize discharging as much, but it is enough to discourage the model from discharging unnecessarily.

5.3.1 MILP Performance

Table 5.1 displays the execution times (in seconds) of the MILP model for the different opportunity costs (OC) in euros/MWh and different instances ("I 1" = instance 1). In the second row (OC = 60/-), instances 1 and 2 where tested with OC = 60, but instance 3 was not tested, as the range of energy costs was bigger in instances 1 and 2 (50 to 110) than the range in instance 3 (only 120 to 170). It is clear that the lower the OC, the higher the execution time. This is to be expected because the model will only find driving water through the turbines viable when the OC is lower than the current cost of energy. In that scenario, the discrete/integer component of the turbines will force the algorithm to search for integer solutions, which is much more difficult than searching in an almost continuous solution space.

Regarding the objective function (OF), every combination of instance and OC arrived at the optimal solution, except instance 2 with OC = 50 and OC = 60, as the execution time was capped at 2 hours (7200 seconds). The best-bound solution had a gap of 5.09%, and 2.55%, respectively.

Table 5.1: MILP performance under different OC

OC	Time (sec)			OF Value (E+07)			Gap (%)		
	I 1	I 2	I 3	I 1	I 2	I 3	I 1	I 2	I 3
50/120	144	7200	593	14.78	12.33	11.74	0.00	5.09	0.00
60/-	395	7200	-	11.59	9.17	-	0.00	2.55	-
70/130	79	453	162	8.65	6.42	8.55	0.00	0.00	0.00
80/140	4	24	67	6.05	4.01	5.57	0.00	0.00	0.00
90/150	2	5	21	3.60	2.07	2.94	0.00	0.00	0.00
100/160	2	3	3	1.79	0.73	1.00	0.00	0.00	0.00
110/170	1	3	2	0.65	- 0.00	0.22	0.00	0.00	0.00

5.3.2 Relax-and-Fix Performance

Along with the different OC (opportunity costs) and instances, the Relax-and-Fix tests also varied the steps and overlaps. The steps ranged from 1 to 16 and the overlap from 0 to 11, but the sum of step plus overlap never exceeded 20. This limit was chosen to be 20 because, it is lower enough from 24 (the time horizon) that the execution times are not too large, but it is big enough to see how the execution times and the OF value evolve as the window grows closer to 24 time periods. In some cases, like some tests on instances 1 and 2, the maximum step was lowered to 12 to accelerate the testing phase. Nevertheless, all instances were tested to, at least, a sum of steps plus overlap equating to 16 and a step of 12.

For the Relax-and-Fix heuristic, instance 1 was tested in Full R&F ("Full (1)") and in Partial R&F ("Partial (1)"), while instances 2 and 3 were only tested in Full R&F ("Full (2)" and "Full (3)"), because, as stated previously, Partial R&F is extremely slow for larger instances (when compared to Full R&F). Naturally, because of the relaxation, there is the possibility that, in the last iterations of a run, there are no solutions in the solution space, which has been constraining as a consequence of the fixation of some variables. For example, if in the last iteration, the fixed variables limit the solution space to a given variable x_t to be in the interval [13,15], but x_t (now inside the window, thus not relaxed) can only be {0,10,20}, then there will be no feasible solution.

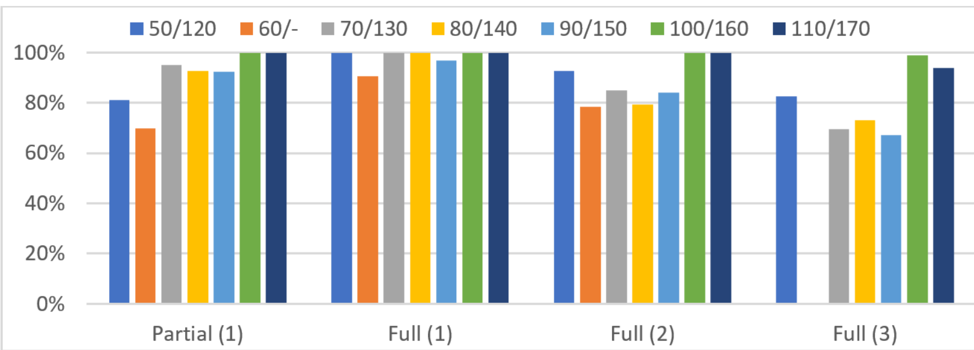
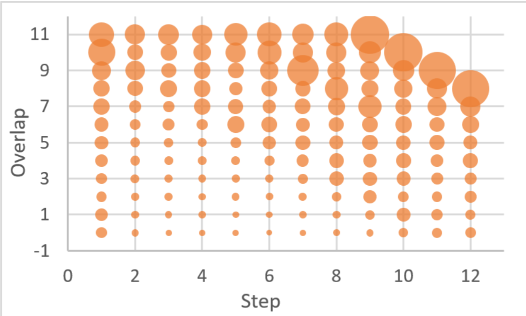
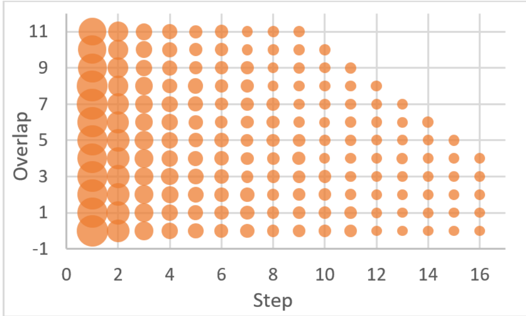


Figure 5.4: Percentage of success in finding a solution among the various combinations of R&F variant and instance ("Variant (instance)"), and OC (colour)



(a) OC = 50



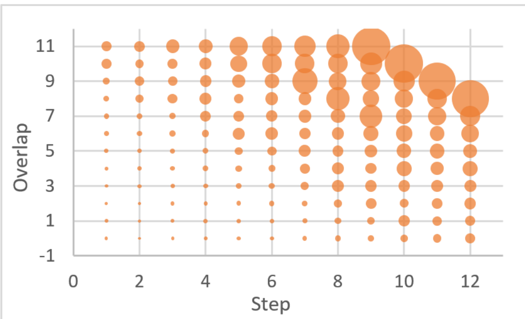
(b) OC = 110

Figure 5.5: Distribution of execution times (area of the bubbles) according to Step and Overlap

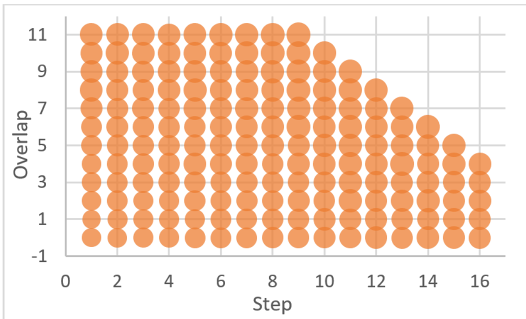
Figure 5.4 shows the percentage of success in finding a solution for several values of OC in the four testing formats. With high OC, seldom is a solution not reached. This makes sense, as the number of integer variables worth using reduces.

Figure 5.5 shows the distribution of execution times, varying Step and Overlap, for two different OC. The area of each bubble corresponds to the execution time of each Step-Overlap pair. Steps are represented in the horizontal axis (1 to 12 in 5.5a and 1 to 16 in 5.5b), and the Overlaps are represented in the vertical axis (0 to 11 in both cases). In particular, Figure 5.5 corresponds to the times obtained when running instance 1 in Full R&F. Despite this, the other instances followed the same pattern, though with more missing points in the graphs, i.e., times when no solution was found.

In case of low OC (5.5a), there is a clear rise in execution time with the rise of the window size (Step + Overlap). When Step = 1, this pattern seems to not apply, but that is due to the higher number of iterations. Figure 5.6a shows that this correlation also applies to Step = 1 when the execution time is divided by the number of iterations. This is to be expected, as the bigger the window, the more non-relaxed variables, the more branches in the tree for the Branch & Bound (B&B) to search, and the higher the execution time.

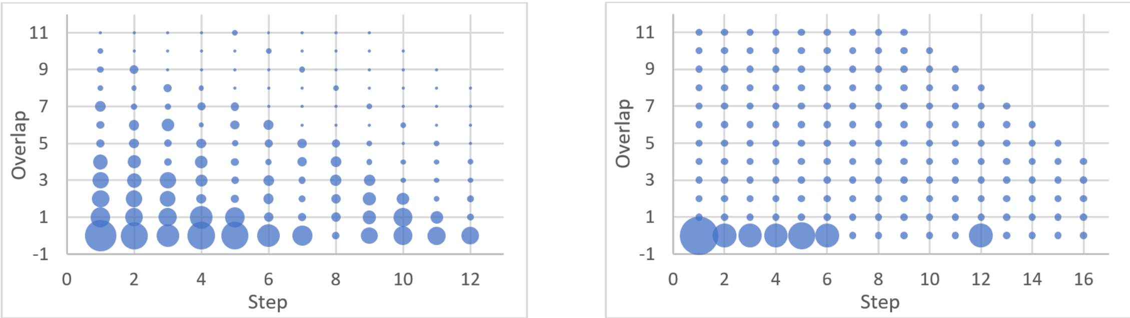


(a) OC = 50



(b) OC = 110

Figure 5.6: Distribution of execution times (area of the bubbles) according to Step and Overlap, per iteration



(a) OC = 50 (b) OC = 110

Figure 5.7: Distribution of gaps (area of the bubbles) of the OF value (in percentage) according to Step and Overlap

In case of high OC (5.5b), there is a decrease in execution time with the increase of the Step, and no notable correlation with the Overlap. The maximum execution time here is 2.663 seconds, in point (1,0). The main cause contributing to the execution time seems to be the number of iterations, as seen in Figure 5.6b, where the execution time is divided by the number of iterations of each run. With a large OC, generating power becomes less profitable and this forces the model to only discharge, which is a non-integer variable. The problem becomes trivial and absent of (many) integer decisions. As such, the window size barely alters the execution time.

Figure 5.7 shows the distribution of gaps, in percentage, of the OF value, varying Step and Overlap, for two different OC. These values also belong to "Full (1)", but the optimal value to which the gaps refer was increased for the sake of visualization in the figures (i.e., in Figure 5.7b the small bubbles all had a gap of 0%, which would otherwise not be visible in the graph). As shown in Figure 5.7a, there is a general trend, as increasing the Step or the Overlap lowers the gap, although less perceptible for the former. To access the relation between the objective function value and the execution time, Figure 5.8b contains the product of the gap with the execution time (according to Step and Overlap), and Figure 5.8a contains the distribution of runs (Step-Overlap pair) with their gap in the horizontal axis and the execution time in the vertical axis. Window sizes smaller than 6 do not seem to yield the best balance of values. With the exceptions of window sizes = 16 and 11/12 (which are dependent on the instance), the bigger the window size, the better the results. Figure 5.8a shows that there are many runs that result in a 1% gap or less, even in less than 10 seconds of execution time.

5.3.3 Performance Comparison

Table 5.1 will be used as a base for comparison between solving the exact optimization problem using a deterministic solver (MILP) and the heuristic procedure developed (R&F). Since the optimal solution was not found in some situations, the OF value to compare will be that of the best solution found within the time limit (in the MILP case).

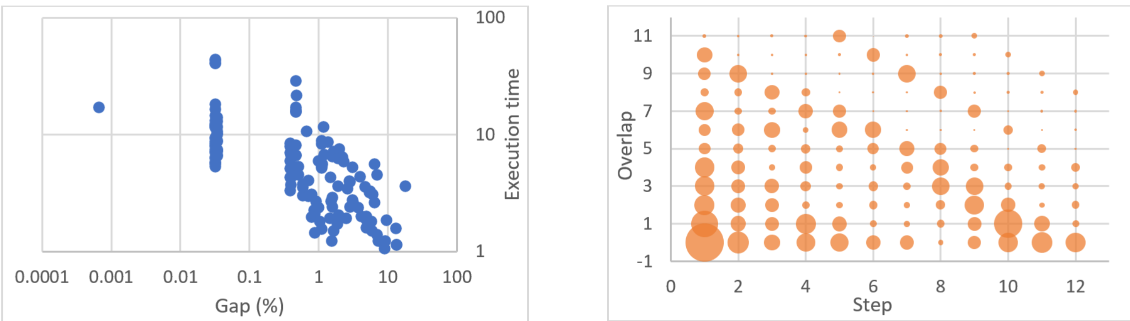
Table 5.2: Time comparison of solutions under 1% gap (in seconds)

Instance	Solution Method	50/120	60/-	70/130	80/140	90/150	100/160	110/170
1	MILP	144	395	79	4	2	2	1
	Parcial R&F	17	10	8	3	2	1	1
	Full R&F	10	10	16	3	1	1	1
2	MILP	7210	7225	453	24	5	3	3
	Full R&F	205	188	und	und	und	2	3
3	MILP	593	-	162	67	21	3	2
	Full R&F	79	-	33	28	9	3	2

Figures 5.9 and 5.10 show the percentage of solutions with an objective function value less than 1% and 2% away from the optimal solution obtained in MILP. These percentages only include feasible solutions.

Judging by Figures 5.9 and 5.10, Partial R&F could be seen as better performing than Full R&F, but with lower rates of solution finding (Figure 5.4), slower execution times, and because Full R&F already delivers very good results in low execution time, it is not deemed as viable of an option to solve the model as Full R&F or MILP. When opportunity costs are high, the execution times are comparable, but when opportunity costs are low, the gap in execution times between Partial and Full R&F widens, the Partial ones being already more than double (215% for OC = 50) than those of Full R&F. This suggests that the execution times of Partial R&F are expected to be much larger than Full R&F with bigger instances like instances 2 and 3. These comparisons were made for the same Step-Overlap pairs and only if the solution was found in both variants of R&F.

Table 5.2 shows the time (in seconds) of the three solution methods over different opportunity costs: MILP, Partial R&F, and Full R&F. The Partial and Full R&F results are an average of all the solution times that had a gap of less than 1%. The value "und" means that there were no solutions with a gap $\leq$ 1% (undefined). In conclusion, for larger instances or low opportunity costs, the R&F approach is preferable to an exact MILP solution approach. Still, the model presents a good performance for smaller instances and high opportunity costs.



(a) Runs' distribution according to their gap and time

(b) Product of time and gap according to Step and Overlap

Figure 5.8: Gap and execution time comparison for "Full (1)", OC = 50

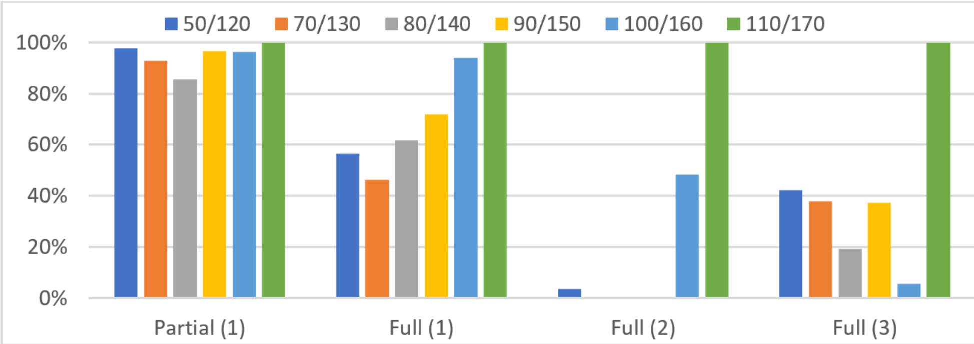


Figure 5.9: Percentage of solutions within 1% of the optimal solution

5.4 Case Study – Cávado Watershed

Here, the Cávado watershed will be the case study, where solutions to instances corresponding to real situations will be analyzed: some methodological aspects will be touched upon, and conclusions about the behavior of the model will be taken by comparing the results with what happened in the real watershed.

5.4.1 Instances and Data

As stated previously, the Cávado watershed is built in the structure of instance 3, and there are four instances with real data: a normal situation, a drought, and two floods. The normal situation is on the days 02/11/2022 to 05/11/2022, and these dates were chosen as "normal" because they sit somewhat in between a drought and a flood, with average rainfall and dam water levels. The drought dates chosen were 15/08/2022 to 18/08/2022, as all of that August and September were considered droughts; the water levels in some dams were almost at the minimum and there was little to no rainfall. For the floods, the dates chosen were 01/02/2021 to 04/02/2021 (Flood 1) and 11/12/2022 to 14/12/2022 (Flood 2).

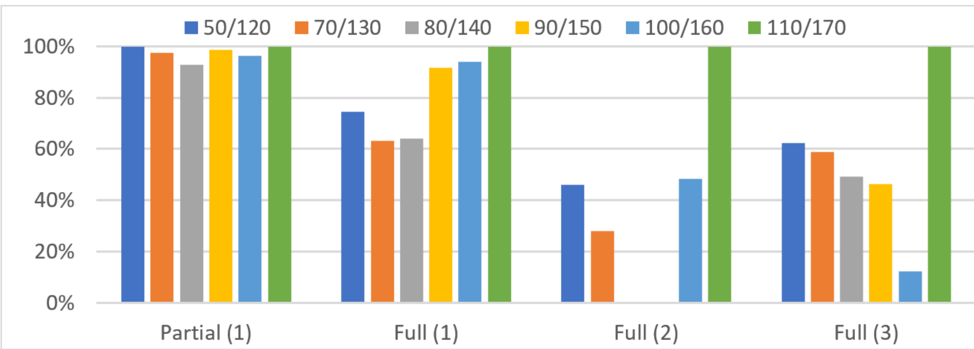


Figure 5.10: Percentage of solutions within 2% of the optimal solution

The data used was obtained from various online sources. The data about the parameters of the dams (minimum volume, maximum discharge, etc), was obtained from CNPGB (Comissão National Portuguesa das Grandes Barragens) [40], SNIRH (Sistema Nacional de Informação de Recursos Hídricos) [41], and APA (Agência Portuguesa do Ambiente) [42]. Data about water volumes, rainfalls, and water flows, among others, was obtained from SNIRH. Finally, data about the hourly and monthly cost of energy was obtained from OMIE (Operador del Mercado Ibérico de Energía - Polo Español) [43]. In these real scenarios, the opportunity costs for power generation and discharging water are the same. For the normal situation, the OC is the average cost of energy for the same month, for the drought situation, the OC is around 125% of the average cost of energy for the same month, and for the flood situations, the OC is around 70% of the average cost of energy for the same month. The opportunity costs were altered from the monthly average energy cost like so to incentivize the model to save water or to generate power now, in case of drought or flood, respectively.

5.4.2 General Results

Table 5.3 shows the different results under MILP and both R&F variants. In R&F solutions, the format "4,0" means "Step,Overlap". These combinations of Step and Overlap were chosen taking into consideration the performance in the methodological tests (every combination with overlap = 4 or 8), for which were added the runs with an overlap of 12 to increase the number of solution methods (notice that overlap 12 was not used in the methodological tests).

The MILP approach arrived at the optimal solution every time, with fast execution times, except for the 2^{nd} flood, where it took 10 minutes. In the 1^{st} flood situation, the Full R&F solution method arrived at a solution 75% of the runs, while the normal, drought, and 2^{nd} flood situations did not arrive at a solution a single time. Despite most runs not having arrived at a solution, they usually realize that there is no feasible solution only in the last iteration (about 90% of runs). This happens because the solution space gets smaller with every fixed variable, but the relaxed variables can always accommodate for the lost space. When there are no relaxed variables left (i.e., the last iteration), the space lost was too much for any feasible solution to remain. With this in mind, the execution times seem to decrease with the increase of Step, which suggests that the execution times are mainly derived from the number of iterations and that either the solution is obvious or the solution space is very tight. In Partial R&F, this trend also appears (in feasible solutions). This variant very often arrives at a solution (with very low gaps), which is a big contrast to the Full R&F variant. In terms of execution times, Partial R&F is faster in the 2^{nd} flood situation, otherwise, the MILP is the fastest.

In general, the MILP approach works very well, and the other solution methods might not be needed. But in some cases like the 2^{nd} flood, the (Partial) R&F solution method can arrive at a solution extremely close to the optimal solution in just 5% of the time.

Table 5.3: Comparison of solutions from MILP and R&F in real situations

Solution	Normal	OC =	115.38	Drought	OC =	200	Flood 1	OC =	20	Flood 2	OC =	60
Method	time (sec)	FO (E+08)	gap %	time (sec)	FO (E+08)	gap %	time (sec)	FO (E+08)	gap %	time (sec)	FO (E+08)	gap %
MILP	4	4.94	0.00	5	-6.63	0.00	9	5.08	0.00	623	15.66	0.00
Full R&F(4,4)	76	0	100.00	39	0	100.00	39	4.92	3.15	38	0	100.00
Full R&F(4,8)	78	0	100.00	41	0	100.00	42	5.00	1.62	39	0	100.00
Full R&F(4,12)	72	0	100.00	40	0	100.00	45	5.04	0.85	40	0	100.00
Full R&F(8,4)	38	0	100.00	22	0	100.00	23	4.95	2.57	21	0	100.00
Full R&F(8,8)	22	0	100.00	21	0	100.00	21	0.00	100.00	21	0	100.00
Full R&F(8,12)	23	0	100.00	21	0	100.00	27	5.04	0.83	27	0	100.00
Full R&F(12,4)	16	0	100.00	15	0	100.00	18	4.98	1.92	26	0	100.00
Full R&F(12,8)	16	0	100.00	15	0	100.00	16	0.00	100.00	23	0	100.00
Full R&F(12,12)	14	0	100.00	13	0	100.00	18	5.05	0.67	16	0	100.00
Full R&F(16,4)	19	0	100.00	13	0	100.00	14	5.00	1.60	12	0	100.00
Full R&F(16,8)	20	0	100.00	11	0	100.00	15	5.01	1.40	14	0	100.00
Full R&F(16,12)	21	0	100.00	11	0	100.00	16	0.00	100.00	12	0	100.00
Partial R&F(4,4)	77	4.94	0.00	162	-6.63	0.00	4	0.00	100.00	75	15.66	0.01
Partial R&F(4,8)	80	4.94	0.00	143	-6.63	0.00	55	0.00	100.00	87	15.66	0.02
Partial R&F(4,12)	71	4.94	0.00	143	-6.63	0.00	75	5.08	0.03	67	0.00	100.00
Partial R&F(8,4)	40	4.94	0.00	88	-6.63	0.00	43	5.08	0.03	48	15.66	0.01
Partial R&F(8,8)	37	4.94	0.00	76	-6.63	0.00	42	5.08	0.03	45	15.66	0.02
Partial R&F(8,12)	37	4.94	0.00	62	-6.63	0.00	31	0.00	100.00	59	0.00	100.00
Partial R&F(12,4)	28	4.94	0.00	30	-6.63	0.00	32	5.08	0.03	33	15.66	0.01
Partial R&F(12,8)	27	4.94	0.00	40	-6.63	0.00	21	0.00	100.00	30	0.00	100.00
Partial R&F(12,12)	24	4.94	0.00	33	-6.63	0.00	41	0.00	100.00	46	15.66	0.02
Partial R&F(16,4)	20	4.94	0.00	25	-6.63	0.00	23	0.00	100.00	28	15.66	0.01
Partial R&F(16,8)	20	4.94	0.00	24	-6.63	0.00	27	5.08	0.04	29	15.66	0.02
Partial R&F(16,12)	21	4.94	0.00	31	-6.63	0.00	28	5.08	0.04	29	15.66	0.02

5.4.3 Behaviour Analysis

To compare the behavior of the model with that of the real watershed, Tables 5.4 to 5.7 display the average daily flow of water through the turbines (q^P), the spillway (q^S), and the pumps (m), and the volume of water at the 23rd hour of each dam (v), per day in the time horizon. Only the MILP solution was compared to the real watershed because the solution found in (Partial) R&F was either the same or very similar to the MILP one (Table 5.3). The enumeration of the dams is the one in the structure of instance 3 (Figure 5.3). For the real watershed, only the average daily water flow was available, as well as the volume of each dam in the 23rd hour of each day.

One thing to note, that is common throughout all situations, is that there seems to be no minimal ecological flow in the dams of the real watershed (at least one that is imposed hourly or daily). This does not affect much the behavior of the model, as discharging water does not constrain any other flows. In the worst-case scenario, in a drought situation, the model could have been forced to discharge below the minimum permitted volume, therefore leaving the instance without any feasible solutions. This did not happen in the days analyzed.

Table 5.4 refers to the normal situation. In general, the model followed what the watershed did. Although it did not pump water and generated less power, the final volumes were very close, except for dams 3 and 5. Still, their total water volume at the end of the 4th day is within 0.14% of each other. The MILP solution seems to prioritize generating power over discharging water (dams 0 and 4 had less volume of water than the minimum required to start generating power). As dam 1 is well within bounds to start generating power and dam 0 is too far away from the minimum volume to do so, the model determines that pumping water from dam 1 to dam 0 hurts the OF. Perhaps if the OF had an OC with a negative sign for pumping water (because water pumped upstream is never lost, except for the efficiency of pumping), then the model could have tried to pump on these occasions.

Table 5.5 refers to the drought situation. As expected, both the model and the real watershed refrain from spending much water. With high opportunity costs (OC = 200), the model is very conservative about water, especially when compared to the real watershed, which did some questionable decisions, namely, using around four times the ecological flow to generate power on the first day, on dam 4. Considering that the minimum volume allowed in dam 4 is 6220000 m^3, that decision in the middle of August does not seem appropriate (also, smaller bodies of water evaporate with less energy). Of course, there is information that this model does not know, and some more that SNIRH does not hold (Ref [41]); This hidden information could have made this decision obvious/necessary. The model only discharged the ecological minimum, with some power generation when the energy cost was higher than 185 euros/MWh.

Table 5.6 refers to the 1^{st} flood situation. There is a notable increase in power generation when compared to the previous situations. Interestingly, both the model and the real watershed pumped water from dam 1 to dam 0. This does seem like a good solution to "bail-out" dam 1: dam 0 has around 200 million m^3 of free space, while dam 1 can only generate power until 94500000 m^3, and dam 2 until 65000000 m^3. This means that dam 1 cannot simply drive through the turbines/discharge a big flow as it would overload dam 2. Where the optimization of the model comes in is to balance the amount of water pumped to dam 0 with the power generating capabilities of dam 0, which cannot happen at the same time (Equations 3.6). Otherwise, in general, the model prefers to generate power instead of discharging. When compared with the real watershed, the model retains more water in total. In the real watershed, those big flushes of water in dam 3 could be some necessities further downstream that the model is not prepared to handle (for example, maintaining the river tides stable amid a flood).

Finally, Table 5.7 refers to the 2^{nd} flood. Like in the 1^{st} flood situation, dam 1 needs to be bailed out again. This time, dam 2 has more space to work with and it is dam 3 that is close to its limit (167300000 m^3). Dam 0 is well below its minimum volume to generate power, so the model does not see the advantage in pumping water to dam 0; although it happens, it is the last resort (again, it could have been different if the OC also applied to pumped water). If dam 0 is out of the question, then the only way is downstream. The model decides to generate power with the maximum flow (10.5 m^3/s) and discharge much more on the second day. At the end of the 4th day, dams 1 and 3 are almost at full capacity, which could be undesirable if the flood did not stop on the 4th day (which it did not). Dams 2 and 5 also pumped water from dam 3 to prevent it from reaching the maximum volume sooner. Like with the 1^{st} flood, generating power is almost always preferred over discharging water. The real watershed on the other hand pumped about 10 times more from dam 1 to dam 0 and solved the issue with dam 3 by flushing a lot of water to the downstream river.

In conclusion, the model works adequately in normal situations, very well in drought situations due to its reluctance in using water for any means, and also reasonably in flood situations. To be more certain about these conclusions, more real instances would have to be evaluated. Nevertheless, it is positive to know that the model generally follows the management decisions of these infrastructures. This allows, in the future, to rely on its application in a decision support system in this context to support decision-makers, providing them with a more integrated view of the entire watershed and ensuring an adequate solution to the operational problem at hand.

Table 5.4: Comparison of the average daily flow of water through the turbines (q^P), the spillway (q^S), and the pumps (m), and the volume of water at the 23rd hour of each dam, per day, between the MILP solution and the real Cávado watershed during the normal situation

	Normal	MILP				REAL			
Dam		day 1	day 2	day 3	day 4	day 1	day 2	day 3	day 4
0	q^P (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	q^S (m^3/s)	0.3	0.3	0.3	0.3	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	2.9	3.6	2.5	0.0
	v (1000m^3)	125582	125721	125700	125797	125130	126000	126580	126970
1	q^P (m^3/s)	7.7	7.7	9.8	7.7	0.0	0.6	0.0	0.0
	q^S (m^3/s)	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	83485	83032	82261	81909	82680	81350	82360	83330
2	q^P (m^3/s)	16.0	17.0	20.6	16.0	45.6	52.9	47.3	35.3
	q^S (m^3/s)	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	60516	59871	58966	58209	58060	60370	59160	57880
3	q^P (m^3/s)	24.8	26.9	31.9	24.8	68.0	41.4	48.1	30.9
	q^S (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	156059	157561	159427	160780	153010	153540	152300	152650
4	q^P (m^3/s)	0.0	0.0	0.0	0.0	4.0	9.1	4.7	6.3
	q^S (m^3/s)	0.9	0.9	0.9	0.9	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	53591	53685	53706	53831	53830	54290	54810	54960
5	q^P (m^3/s)	18.3	23.3	34.9	18.3	9.4	18.5	8.9	10.5
	q^S (m^3/s)	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	90753	88990	86205	84702	91520	90830	90650	90260

Table 5.5: Comparison of the average daily flow of water through the turbines (q^P), the spillway (q^S), and the pumps (m), and the volume of water at the 23rd hour of each dam, per day, between the MILP solution and the real Cávado watershed during the drought

	Drought	MILP				REAL			
Dam		day 1	day 2	day 3	day 4	day 1	day 2	day 3	day 4
0	q^P (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	q^S (m^3/s)	0.3	0.3	0.3	0.3	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	113444	113744	114005	114239	113150	113150	113060	112970
1	q^P (m^3/s)	0.7	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	q^S (m^3/s)	0.8	0.9	0.9	0.9	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	70323	70240	70213	70148	71100	71090	71100	71100
2	q^P (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	5.0	5.5
	q^S (m^3/s)	2.5	2.5	2.5	2.5	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	42440	42402	42283	42146	42450	42550	42170	42680
3	q^P (m^3/s)	0.0	0.0	0.0	0.0	12.1	12.0	13.3	13.4
	q^S (m^3/s)	4.0	4.0	4.0	4.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	120812	121121	121353	121253	119990	119890	118570	118330
4	q^P (m^3/s)	0.0	0.0	0.0	0.0	3.9	1.8	0.8	1.1
	q^S (m^3/s)	0.9	0.9	0.9	0.9	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	8797	9329	9541	9709	8420	8420	7950	7890
5	q^P (m^3/s)	0.0	0.0	0.0	0.0	2.1	0.0	3.0	3.0
	q^S (m^3/s)	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	52060	52181	52276	52249	51910	51910	51700	51440

Table 5.6: Comparison of the average daily flow of water through the turbines (q^P), the spillway (q^S), and the pumps (m), and the volume of water at the 23rd hour of each dam, per day, between the MILP solution and the real Cávado watershed during the 1st flood

	Flood 1	MILP				REAL			
Dam		day 1	day 2	day 3	day 4	day 1	day 2	day 3	day 4
0	q^P (m^3/s)	2.3	8.1	19.6	23.1	0.0	7.2	0.0	0.0
	q^S (m^3/s)	0.3	0.2	0.1	0.0	0.0	0.0	0.0	0.0
	m (m^3/s)	12.0	6.0	3.0	0.0	8.2	8.5	11.5	5.9
	v (1000m^3)	363834	366519	366785	365349	363340	366640	372960	375980
1	q^P (m^3/s)	2.2	7.0	8.1	5.3	0.5	0.0	0.0	0.0
	q^S (m^3/s)	0.7	0.3	0.2	0.4	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	87865	91454	93581	96008	90610	89150	89190	87300
2	q^P (m^3/s)	18.3	21.1	17.4	22.0	103.1	149.8	168.1	124.1
	q^S (m^3/s)	0.4	0.1	0.5	0.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	61264	64982	64974	64051	61110	64540	62610	60700
3	q^P (m^3/s)	9.9	22.7	26.2	34.0	67.8	67.5	67.4	67.7
	q^S (m^3/s)	2.7	1.3	0.7	0.0	0.0	121.7	249.3	179.3
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	152636	158398	161119	163101	154600	158610	155180	152690
4	q^P (m^3/s)	0.7	4.1	11.6	13.7	2.4	9.6	9.9	13.8
	q^S (m^3/s)	0.9	0.9	0.9	0.9	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	111281	120241	123419	124056	110120	113610	120360	121890
5	q^P (m^3/s)	10.0	26.6	28.3	39.9	1.1	13.7	20.7	30.7
	q^S (m^3/s)	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v (1000m^3)	104058	106090	105923	103869	103900	108220	110600	109550

Table 5.7: Comparison of the average daily flow of water through the turbines (q^P), the spillway (q^S), and the pumps (m), and the volume of water at the 23rd hour of each dam, per day, between the MILP solution and the real Cávado watershed during the 2^{nd} flood

	Flood 2	MILP				REAL			
Dam		day 1	day 2	day 3	day 4	day 1	day 2	day 3	day 4
0	q^P (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	q^S (m^3/s)	0.3	0.3	0.3	0.3	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	4.5	5.3	23.2	31.5	22.7	21.6
	v ($1000m^3$)	197406	202299	204953	207939	197320	202000	210600	214710
1	q^P (m^3/s)	10.5	10.5	8.8	2.6	0.0	0.0	0.0	4.2
	q^S (m^3/s)	0.0	35.6	0.4	0.7	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v ($1000m^3$)	92858	94500	94575	96029	88180	89800	90790	85230
2	q^P (m^3/s)	22.0	22.0	15.6	13.8	0.0	47.2	115.2	154.1
	q^S (m^3/s)	0.0	0.0	0.6	0.8	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	5.5	6.4	0.0	0.0	0.0	0.0
	v ($1000m^3$)	57083	61869	63231	63449	59290	59900	62050	59200
3	q^P (m^3/s)	34.0	34.0	25.5	24.1	68.0	68.0	60.4	68.0
	q^S (m^3/s)	0.0	0.0	1.0	1.2	27.3	110.8	106.3	15.3
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v ($1000m^3$)	149275	159357	162468	166652	135560	129000	132950	143630
4	q^P (m^3/s)	12.3	5.5	8.9	9.6	14.1	13.0	10.5	14.2
	q^S (m^3/s)	0.9	0.9	0.9	0.9	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	v ($1000m^3$)	89744	95664	99921	102844	86650	90400	95250	96530
5	q^P (m^3/s)	39.9	32.4	26.6	15.8	8.4	11.1	24.3	24.5
	q^S (m^3/s)	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0
	m (m^3/s)	0.0	0.0	4.8	5.5	0.0	0.0	0.0	0.0
	v ($1000m^3$)	79189	79336	78989	78743	81020	81400	85970	85740

Chapter 6

Conclusions

In the context of optimizing decision-making in watersheds with multiple dams, this thesis aimed to explore and develop approaches that enhance the efficiency and effectiveness of such processes. The objective was to find suitable methods to address the complexities associated with decision-making in this domain.

To achieve this objective, a mathematical model was created, and then implemented in Python using the PuLP library. Two variants of the Relax-and-Fix heuristic were proposed, Partial R&F and Full R&F. To validate and evaluate the performance of the model and the Relax-and-Fix variants, three different structures were devised, each consisting of at least one instance that was utilized to test the model's capabilities. Notably, the third structure incorporated instances with real data from the Cávado watershed, representing diverse scenarios such as floods, droughts, and normal conditions. This inclusion allowed for a realistic assessment of the model's effectiveness in practical situations. Additionally, a comparison was drawn between the model's outcomes and the actual occurrences in the Cávado watershed.

Throughout the methodology, specific parameters were modified to observe their impact on the execution time and objective function value. The parameter named Opportunity Cost (OC) was altered to investigate the influence of future prices and water availability in the model. Similarly, the parameters Step and Overlap within the Relax-and-Fix approach were systematically adjusted to analyze their effects on the optimization process.

In the methodology section, it was possible to conclude that, to a high enough OC, every instance became easy to solve, and as such, the execution times and the gaps decreased, being similar to every solution method. In the case of R&F, the execution times became proportional to the number of iterations. As the OC decreased, the execution times increased (substantially) and the gaps between the R&F approach and MILP widened. R&F proved to be an excellent alternative to MILP for low OC or big instances.

With real instances, Full R&F rarely reached a solution, while Partial R&F reached almost every time. Furthermore, the gaps of the Partial R&F solutions were always less than 0.05% (very good). As the execution times between Partial and Full were on the same order of magnitude, this suggests that Full R&F is not ideal for instances with a large time horizon (96 time periods in the real instances). The chances of a feasible solution still existing decrease with each iteration of such relaxation.

As models are an abstraction and simplification of reality, there are some points to improve in the proposed approaches. For future work, it could be investigated the impact of adding an opportunity cost to the pumping mechanism. Even though water is not lost when pumping, there are some losses due to efficiency, which could be addressed. It would be expected that the model would behave differently and probably have higher execution times due to an increase in integer decision variables. In this case, the Relax-and-Fix approach would be useful to obtain a good solution in a reasonable time. A more relevant modification to the model could be the acknowledgment of uncertainty in the weather conditions (and in the upstream flows). Future work would also encompass behavior analysis in different watersheds, preferably with more dates for each climate situation. If the parameters are correctly represented, this model could be used in helping the operational decision-making in any watershed and used rather confidently in the case of the Cávado watershed.

Bibliography

[1] S Schmutz and J Sendzimir. *Riverine ecosystem management: Science for governing towards a sustainable future*. Vol. 8. Aquatic Ecology Series. 2018. ISBN: 978-3-319-73250-3. URL: https://library.oapen.org/bitstream/handle/20.500.12657/27726/1002280.pdf?seque.

[2] Ladislav Votruba and Vojtěch Broža. *Water Management in Reservoirs*. Developments in water science. Elsevier Science Publishers B.V., 1989. ISBN: 0-444-98933-1.

[3] Eduardo Ibanez et al. "Enhancing hydropower modeling in variable generation integration studies". In: *Energy* 74 (C 2014), pp. 518–528. ISSN: 03605442. DOI: 10.1016/J.ENERGY.2014.07.017.

[4] Asmadi Ahmad et al. "Reservoir optimization in water resources: A review". In: *Water Resources Management* 28 (11 Sept. 2014), pp. 3391–3405. ISSN: 15731650. DOI: 10.1007/S11269-014-0700-5/FIGURES/4.

[5] Hongrui Liu et al. "A reservoir operation method that accounts for different inflow forecast uncertainties in different hydrological periods". In: *Journal of Cleaner Production* 256 (May 2020). ISSN: 09596526. DOI: 10.1016/J.JCLEPRO.2020.120471.

[6] *How Reliable Are Weather Forecasts? | NOAA SciJinks – All About Weather*. URL: https://scijinks.gov/forecast-reliability/. (accessed: 27/12/2022).

[7] William Chen and Julian D. Olden. "Designing flows to resolve human and environmental water needs in a dam-regulated river". In: *Nature Communications 2017 8:1* 8 (1 Dec. 2017), pp. 1–10. ISSN: 2041-1723. DOI: 10.1038/s41467-017-02226-4.

[8] Evangelos Panos, Martin Densing, and Kathrin Volkart. "Access to electricity in the World Energy Council's global energy scenarios: An outlook for developing regions until 2030". In: *Energy Strategy Reviews* 9 (Mar. 2016), pp. 28–49. ISSN: 2211-467X. DOI: 10.1016/J.ESR.2015.11.003.

[9] Luis M. Abadie and Nestor Goicoechea. "Optimal management of a mega pumped hydro storage system under stochastic hourly electricity prices in the Iberian Peninsula". In: *Energy* 252 (Aug. 2022), p. 123974. ISSN: 0360-5442. DOI: 10.1016/J.ENERGY.2022.123974.

[10] Mümtaz Ak, Elcin Kentel, and Secil Savasaneril. "Quantifying the revenue gain of operating a cascade hydropower plant system as a pumped-storage hydropower system". In: *Renewable Energy* 139 (Aug. 2019), pp. 739–752. ISSN: 0960-1481. DOI: 10.1016/J.RENENE.2019.02.118.

[11] Florence School of Regulation. *Electricity markets in the EU*. URL: https://fsr.eui.eu/electricity-markets-in-the-eu/. (accessed: 17/12/2022).

[12] *NEMOs: Nominated Electricity Market Operators*. URL: https://www.nemo-committee.eu/index. (accessed: 18/12/2022).

[13] *Overview of the Brazilian Energy Market | METRON*. URL: https://www.metron.energy/blog/overview-brazilian-market/. (accessed: 20/12/2022).

[14] *US Electricity Markets 101*. URL: https://www.rff.org/publications/explainers/us-electricity-markets-101/. (accessed: 20/12/2022).

[15] Hongye Guo et al. "Power market reform in China: Motivations, progress, and recommendations". In: *Energy Policy* 145 (Oct. 2020), p. 111717. ISSN: 0301-4215. DOI: 10.1016/J.ENPOL.2020.111717.

[16] Chancery House St et al. "IHA Central Office IHA Regional and National Offices Cover Image: The 944 MW Murum hydroelectric plant Malaysia. Credit: Sarawak Energy". In: (2018). URL: www.hydropower.org. (accessed: 21/12/2022).

[17] Vivek P Kapadia. "Importance of dam operation for downstream river health". In: *Symposium on Sustainable Development of Dams and River Basins*. Vol. 24. 2021, 27th.

[18] Murat Pinarlik, Adebayo J. Adeloye, and Zeliha Selek. "Impacts of Ignored Evaporation and Sedimentation Fluxes at Planning on Reservoir Performance in Operation". In: *Water Resources Management* 35 (11 Sept. 2021), pp. 3539–3570. ISSN: 15731650. DOI: 10.1007/S11269-021-02904-5/FIGURES/16.

[19] David Martínez-Granados et al. "The Economic Impact of Water Evaporation Losses from Water Reservoirs in the Segura Basin, SE Spain". In: *Water Resources Management 2011 25:13* 25 (13 June 2011), pp. 3153–3175. ISSN: 1573-1650. DOI: 10.1007/S11269-011-9850-X.

[20] Mohammed Falah Allawi et al. "Developing reservoir evaporation predictive model for successful dam management". In: *Stochastic Environmental Research and Risk Assessment* 35 (2 Feb. 2021), pp. 499–514. ISSN: 14363259. DOI: 10.1007/S00477-020-01918-6/FIGURES/5.

[21] Chun Tian Cheng and K. W. Chau. "Flood control management system for reservoirs". In: *Environmental Modelling & Software* 19 (12 Dec. 2004), pp. 1141–1150. ISSN: 1364-8152. DOI: 10.1016/J.ENVSOFT.2003.12.004.

[22] Yanlai Zhou and Shenglian Guo. “Incorporating ecological requirement into multipurpose reservoir operating rule curves for adaptation to climate change”. In: *Journal of Hydrology* 498 (Aug. 2013), pp. 153–164. ISSN: 0022-1694. DOI: 10.1016/J.JHYDROL.2013.06.028.

[23] Shahryar Khalique Ahmad and Faisal Hossain. “Forecast-informed hydropower optimization at long and short-time scales for a multiple dam network”. In: *Journal of Renewable and Sustainable Energy* 12 (1 Feb. 2020), p. 014501. ISSN: 19417012. DOI: 10.1063/1.5124097.

[24] Shahryar Khalique Ahmad and Faisal Hossain. “Maximizing energy production from hydropower dams using short-term weather forecasts”. In: *Renewable Energy* 146 (Feb. 2020), pp. 1560–1577. ISSN: 0960-1481. DOI: 10.1016/J.RENENE.2019.07.126.

[25] Wenjun Cai et al. “Flood Risk Analysis for Cascade Dam Systems: A Case Study in the Dadu River Basin in China”. In: *Water 2019, Vol. 11, Page 1365* 11 (7 June 2019), p. 1365. ISSN: 2073-4441. DOI: 10.3390/W11071365.

[26] Chih Chiang Wei and Nien Sheng Hsu. “Optimal tree-based release rules for real-time flood control operations on a multipurpose multireservoir system”. In: *Journal of Hydrology* 365 (3-4 Feb. 2009), pp. 213–224. ISSN: 0022-1694. DOI: 10.1016/J.JHYDROL.2008.11.038.

[27] Sandrine Vidal et al. “Flood Forecast Tool to Help Dam Management from France to Thailand”. In: *Springer Water* (2020), pp. 427–435. ISSN: 23648198. DOI: 10.1007/978-981-15-5436-0_34/FIGURES/5.

[28] R. Dittmann et al. “Optimum multi-objective reservoir operation with emphasis on flood control and ecology”. In: *Natural Hazards and Earth System Sciences* 9 (6 Nov. 2009), pp. 1973–1980. ISSN: 1561-8633. DOI: 10.5194/NHESS-9-1973-2009.

[29] Margaret A. Palmer et al. “Climate change and the world’s river basins: anticipating management options”. In: *Frontiers in Ecology and the Environment* 6 (2 Mar. 2008), pp. 81–89. ISSN: 1540-9309. DOI: 10.1890/060148.

[30] Zaher Mundher Yaseen et al. “A hybrid bat–swarm algorithm for optimizing dam and reservoir operation”. In: *Neural Computing and Applications* 31 (12 Dec. 2019), pp. 8807–8821. ISSN: 14333058. DOI: 10.1007/S00521-018-3952-9/FIGURES/5.

[31] Yong Gun Kim et al. “Effective Optimization-Simulation Model for Flood Control of Cascade Barrage Network”. In: *Water Resources Management* 35 (1 Jan. 2021), pp. 135–157. ISSN: 15731650. DOI: 10.1007/S11269-020-02715-0/FIGURES/12.

[32] G. S. Sampaio et al. “Optimization of the operation of hydro stations in market environment using Genetic Algorithms”. In: *International Conference on the European Energy Market, EEM* (2013). ISSN: 21654077. DOI: 10.1109/EEM.2013.6607278.

[33] Zhong Kai Feng, Wen Jing Niu, and Chun Tian Cheng. “Optimization of hydropower reservoirs operation balancing generation benefit and ecological requirement with parallel multi-objective genetic algorithm”. In: *Energy* 153 (June 2018), pp. 706–718. ISSN: 0360-5442. DOI: 10.1016/J.ENERGY.2018.04.075.

[34] Vivien Lai et al. “Investigating dam reservoir operation optimization using metaheuristic algorithms”. In: *Applied Water Science* 12 (12 Dec. 2022), pp. 1–13. ISSN: 21905495. DOI: 10.1007/S13201-022-01794-1/FIGURES/7.

[35] Johannes Bisschop. *AIMMS optimization modeling*. Lulu. com, 2006.

[36] K. A. G. Araujo et al. “Relax-and-fix heuristics applied to a real-world lot-sizing and scheduling problem in the personal care consumer goods industry”. In: *Operations Research Forum* 4 (2 July 2021). DOI: 10.1007/s43069-023-00230-7.

[37] Abdolreza Roshani, Davide Giglio, and Massimo Paolucci. “A relax-and-fix heuristic approach for the capacitated dynamic lot sizing problem in integrated manufacturing/remanufacturing systems”. In: *IFAC-PapersOnLine* 50 (1 July 2017), pp. 9008–9013. ISSN: 2405-8963. DOI: 10.1016/J.IFACOL.2017.08.1580.

[38] Christian Blum et al. *Hybrid Metaheuristics: An Emerging Approach to Optimization*. Springer Berlin Heidelberg, Jan. 2008. ISBN: 978-3-540-78294-0. DOI: 10.1007/978-3-540-78295-7.

[39] Stuart Mitchell, Michael J. O’Sullivan, and Iain Dunning. “PuLP : A Linear Programming Toolkit for Python”. In: 2011.

[40] CNPGB. *cnpgb21*. URL: https://cnpgb.apambiente.pt/. (accessed: 15/06/2023).

[41] SNIRH. *Sistema Nacional de Informação de Recursos Hídricos*. URL: https://snirh.apambiente.pt/. (accessed: 15/06/2023).

[42] APA. “Região Hidrográfica do Cávado, Ave e Leça (RH2)”. In: (2016). URL: https://apambiente.pt/sites/default/files/_SNIAMB_Agua/DRH/PlaneamentoOrdenamento/PGRH/2016-2021/PTRH2/PGRH_2_RH2_Parte2.pdf.

[43] OMIE. *OMIE*. URL: https://www.omie.es/pt. (accessed: 15/06/2023).