Les Cahiers du GERAD ISSN: 0711–2440

Optimizing strategies for short-term hydropower scheduling using a blackbox optimization framework

M. Daadaa, S. Séguin, M.F. Anjos, K. Demeester

G–2023–38

August 2023

Citation suggérée : M. Daadaa, S. Séguin, M.F. Anjos, K. Demeester (Août 2023). Optimizing strategies for short-term hydropower scheduling using a blackbox optimization framework, Rapport technique, Les Cahiers du GERAD G– 2023–38, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2023-38) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

Suggested citation: M. Daadaa, S. Séguin, M.F. Anjos, K. Demeester (August 2023). Optimizing strategies for short-term hydropower scheduling using a blackbox optimization framework, Technical report, Les Cahiers du GERAD G–2023–38, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2023-38) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2023
– Bibliothèque et Archives Canada, 2023

Legal deposit – Bibliothèque et Archives nationales du Québec, 2023
– Library and Archives Canada, 2023

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Optimizing strategies for short-term hydropower scheduling using a blackbox optimization framework

Maissa Daadaa [a, b]

Sara Séguin [a, b]

Miguel F. Anjos [c, b]

Kenjy Demeester [d]

[a] *Université de Québec à Chicoutimi, Chicoutimi (Qc), Canada, G7H 2B1*

[b] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[c] *University of Edinburgh, Edinburgh, United Kingdom, EH8 9YL*

[d] *Rio Tinto, Saguenay (Qc), Canada, G7S 4R5*

maissa.daadaa1@uqac.ca
sara.seguin@uqac.ca
anjos@stanfordalumni.org
kenjy.demmester@riotinto.com

August 2023
Les Cahiers du GERAD
G–2023–38

Abstract : This paper presents a study on the best possible use of optimization models for the short-term hydropower scheduling problem. Different deterministic and stochastic models have been proposed to schedule turbines for an entire planning horizon. In this paper, we take a different perspective on the problem by investigating how to determine the best optimization model during the planning horizon in order to maximize the total energy production. Specifically, we use a mixed-integer linear programming formulation as a deterministic model, where the inflows are represented by a median scenario for each day, and a stochastic multistage mixed-integer linear programming formulation as a stochastic model, where the uncertainty of the inflows is represented using a scenario tree. For the stochastic model, the stages are aggregated to reduce the computation time. We formulate the problem of choosing the best optimization model as a blackbox optimization problem. Computational experiments with two powerhouses in series, each with five turbines, for a 10-day rolling-horizon are presented. The solutions provided by the blackbox solver are analyzed, including an evaluation of the expected inflow volumes. These analyses shed light on the selection of the most appropriate optimization model. The results show that the choice of optimization model can indeed be influenced by the observed variability of inflows.

Keywords : Hydropower scheduling, short-term hydropower optimization, mixed-integer linear programming, stochastic multi-stage programming, blackbox optimization

Acknowledgements: This work was supported by a Mitacs-Rio Tinto Accelerate Internship Award.

Nomenclature

The following notations are used in the paper:

Sets	
$i \in \{1, 2, .., S\}$	indexes of the set of scenarios.
$n \in \{1, 2, .., N_i\}$	indexes of the set of nodes for each scenario i.
$c \in \{1, 2, .., C\}$	indexes of the set of powerhouses.
$l \in \{1, 2, .., U^c\}$	indexes of the set of powerhouses upstream of each powerhouse c.
$j \in \{1, 2, .., J_n^c\}$	indexes of the set of turbines associated to the node n and powerhouse c.
$b \in \{1, 2, .., B^c\}$	indexes of the set of combinations of each powerhouse c.
$k \in \{1, 2, .., K_c^b\}$	indexes of the set of efficiency points associated to powerhouse c and combination b.
$d \in \{1, 2, .., T\}$:	index of the day in the rolling-horizon.
$m \in [1, 2, .., Nc^{max}]$	index of the time change.
$p \in [1, 2, .., Nc^{max} + 1]$	index of period between times changes
$t_m \in [0, .., T]$	the value of the time change
Parameters	
$P_{k,n}^c$	power output of powerhouse c at node n and point k (MW).
$q_{k,n}^c$	water discharge of powerhouse c at node n and point k (m^3/s).
π_i^c	probability of scenario i for powerhouse c.
ξ_n^c	inflow of powerhouse c at node n (m^3/s).
β	conversion factor from (m^3/s) to (hm^3/h).
θ^c	estimated energy losses from maximum storage (MW) at powerhouse c.
ϵ^c	start-up penalty of turbine (MW) at powerhouse c.
N_{max}^c	maximum number of start-ups for powerhouse c.
V_{max}^c	maximum volume of reservoir c (hm^3).
v_{ini}^c	initial volume of reservoir c (hm^3).
v_{final}^c	final volume of reservoir c (hm^3).
Δt	the duration of the stage (h).
$A_{n,k,j}^c = \begin{cases} 1 & \text{if the turbine } j \text{ of powerhouse } c \text{ at the point } k \text{ is activated at node } n. \\ 0 & \text{otherwise.} \end{cases}$	
Nc^{max}	maximum number of changes
EBB_d:	the energy produced at day d
$a_p = \begin{cases} 1 & \text{if the deterministic model is used at period p} \\ 0 & \text{otherwise.} \end{cases}$	
Decision variables	
$y_{k,n}^c = \begin{cases} 1 & \text{if the point } k \text{ is chosen at node } n \text{ for powerhouse } c. \\ 0 & \text{otherwise.} \end{cases}$	
$z_{j,n}^c = \begin{cases} 1 & \text{if the turbine } j \text{ of powerhouse } c \text{ is started at node } n. \\ 0 & \text{otherwise.} \end{cases}$	
v_n^c volume of the reservoir of powerhouse c at node n (hm^3).	
d_n^c water spillage at powerhouse c and node n (m^3/s).	

1 Introduction

Short-term hydropower scheduling is concerned with defining the optimal strategy for daily energy production by defining the optimal values of the water discharges, the reservoir volumes, and the operational status of the turbines in order to achieve an objective such as maximizing total energy production, minimizing total cost, or maximizing total profit, subject to operational constraints, including limits on turbine startups. The primary purpose of limiting the number of turbine starts in a powerhouse is to ensure efficient and reliable operation of the turbines and to minimize potential damage or wear to the turbines. A variety of deterministic and stochastic models have been proposed for this problem. A recent overview of the proposed models can be found in [31].

In deterministic models, all the necessary parameters are known in advance. Different formulations and methods have been proposed in the literature [34]. Mixed-integer linear programming (MILP) is widely used to solve hydropower scheduling problem because it involves making decisions on discrete variables, such as when to start and stop the turbines for each powerhouses. MILP can handle these discrete decisions by incorporating integer variables into the optimization model, ensuring that the generated solutions are realistic. In [16], a MILP formulation is used to solve hydroscheduling problem in the case of a deregulated market, taking into account head variations, hydraulic losses, and mechanical and electrical losses in the turbine generators. The formulation was tested on real cases derived from the Brazilian system with constant inflows over the planning horizon. In [10], the authors developed a novel MILP formulation based on the efficiency curves of the turbine generators to solve the hydroscheduling problem in order to maximize the energy produced and penalize the start-ups. The formulation is based on the definition of a pair of points representing the maximum efficiency for water discharge and the power produced for each combination of active turbines since the maximum power generation is reached at these points. The inflow forecasts are considered as known and available from a historical database. The method was tested on a real case from the Saguenay-Lac-St-Jean hydroelectric system in Québec, Canada. In [23], another deterministic MILP approach was developed under the assumption that inflow uncertainties can be ignored in a horizon of days to weeks [11]. Other studies propose MILP formulations considering irregular forbidden zones of operation [15, 38]. In [8], binary variables are used to deal multiple vibration zones and performance curves of the unit.

Another optimization method commonly used for hydropower scheduling is dynamic programming. Dynamic programming solves the problem by considering multiple sub-problems so that the optimal solution can be obtained using recursion. In [33], dynamic programming was used to determine the maximum power output generated by a given combination of active turbines. This information is then used as input for a two-phase optimization process that determines the optimal water discharge, reservoir volume and set of turbines in operation at each period of the planning horizon. The time steps are the hours, the states are the number of turbines in operation at each step, and the decision variables are the number of turbine starts and stops at each step. In [36], dynamic programming is used to determine the number of turbines in operation on an hourly basis in order to maximize the power produced. The inconvenience of this method is that the problem quickly becomes difficult to solve as the number of variables increases [4].

Other methods used to solve deterministic hydroscheduling problem include Lagrangian relaxation [17, 34], genetic algorithms [21, 30], and the combination of outer approximation and Benders decomposition [27].

In stochastic models, some parameters of the problem are uncertain. In hydropower systems, uncertain parameters include the natural inflows, the prices (in the case of a deregulated market), the demand, or other aspects of the problem. Considering the uncertainties allow the producer to obtain a robust and precise solution that reflects reality. Different research papers have looked into short-term hydropower scheduling under uncertainty. A multi-stage mixed-integer linear stochastic program (SMMILP) is one of the more common methods used for this problem. In [18], a SMMILP is developed to solve the hydropower optimization problem subject to the uncertainty of prices and inflows. The objective is to make a balance between current profits and expected future profits in deregulated market. The model is tested with data from a Norwegian hydropower producer and the Nordic power market. In [13], a two-stage mixed integer stochastic program is developed and takes the uncertainty in market prices and both production and physical trading aspects. The first stage determines bids for the day-ahead and the second involves trades in the intraday market and production decisions. In [9], a SMMILP is developed in order to maximize the total energy production subject to the uncertainty of the inflows. The decisions are taken at the beginning of the horizon before knowing the realization of uncertainty, then are adjusted once the uncertainties are known. The uncertainty of inflows is represented by a scenario tree. In hydropower management, the usual way of describing uncertain parameters is by scenario trees. It consists of nodes and paths, where each path represents a scenario with an associated probability and the nodes within the scenarios represent the values of uncertainty. The scenario tree serve as an input to the stochastic model. Therefore, several approaches for generating and reducing scenario trees are proposed including clustering methods[26], moment-matching methods[7], Backward and forward reduction[22] and Monte Carlo method [12].

Another method proposed in the literature to solve hydropower optimization models is stochastic dynamic programming. These models are commonly employed to address hydropower problems within long or medium-term planning horizons. In [29], a stochastic hydroscheduling is solved where the uncertainty of the demand is considered in order to minimize the expected cost. A deterministic equivalent linear program is used to define the water flows, and then a stochastic dynamic programming recursion is applied to give an approximately optimal scheduling for a single station. Other methods proposed in the literature to deal with uncertainty include robust optimization [3] and artificial intelligence [1].

The cited works contribute to advancing the field of hydropower optimization by proposing different techniques for deterministic and stochastic models. The deterministic models offer the advantage of simplicity and computational efficiency compared to the stochastic models. However, they are unable to adequately account for parameter uncertainties, which can lead to suboptimal solutions in the face of unexpected events. To overcome these limitations, a approach that combines deterministic and stochastic methods can be used to find a trade-off between achieving an optimal solution and reducing computational time. To achieve this objective, it is important to determine the appropriate model to use during the planning horizon (either deterministic, stochastic, or both) and the optimal time to transition from one model to the other when a transition is deemed necessary. This decision should be based on some indicators that we aim to determine. To solve this problem, a state-of-the-art blackbox optimization solver is used. Blackbox optimization is a computational tool or algorithm used to find optimal solutions to optimization problems when the underlying objective function or constraints are not explicitly known or easily defined and can only be calculated through a computer code [6]. In [32], blackbox optimization is applied to determine the optimal structure of the scenario tree that maximizes energy production. The scenario tree includes several input parameters, including the number of stages, the number of child nodes for each node, and the aggregation level for each day. Blackbox optimization is used to adjust these parameters in a way that the energy production is maximized.

In this paper, blackbox optimization is used to determine the optimal combination of models and the optimal time to transition from one model to another. For this purpose, a MILP formulation is used as a deterministic model because it provides a flexible framework for modeling and optimizing hydropower scheduling problems. It allows the inclusion of various constraints that enable the representation of real-world operational requirements such as turbine start-up constraints, reservoir limits, etc. Inflows are represented using the median scenario for each day. A SMMILP formulation is used as a stochastic model because it can handle both discrete and continuous decision variables, making it suitable for the complex nature of hydropower scheduling. This modeling approach enables optimization of decisions such as turbine start-up. Moreover, the use of a SMMILP facilitates the application of aggregation techniques to reduce computational complexity, which is the case in this work. The uncertainty of the inflows is represented by a scenario tree that is generated and reduced using the backward reduction method. This method aims to iteratively identify specific scenarios that are eliminated from a full scenario tree in a way that minimizes the distance of the probability distribution between the reduced scenario tree and the original full scenario tree [19]. Moreover, the implementation of the backward reduction method can be done directly using Scenred2/ GAMS [20].

For this purpose, a 10-day rolling horizon benchmark test is developed based on real data from Rio Tinto in the Saguenay region of the province of Québec in Canada. The solution provided by the blackbox solver is analyzed, including an evaluation of the expected inflow volumes, to investigate the choice of optimization models. These analyzes provide valuable insights into the choice of the optimal optimization model. The results show that the choice of optimization model is indeed influenced by the observed variability of inflows.

The paper is organized as follows. Section 2 presents the hydropower optimization models. Blackbox optimization is presented in section 3. The results are discussed in Section 4. Finally, concluding remarks are presented in Section 5.

2 Short term hydropower scheduling

The objective of the hydroscheduling problem is to optimize energy generation by efficiently utilizing the available resources. This entails determining the optimal water discharge, the volume of the reservoirs, and the state of each turbine (on or off) for each powerhouse and at each time stage, subject to some constraints and taking into account different factors such as inflow uncertainty, efficiency, and turbine start-ups. However, the hydropower production function is characterized by nonconvexity and nonlinearity, which poses significant modeling and optimization challenges. To address this complexity, various approximation and linearization techniques have been proposed to effectively represent and analyze the production function in different ways.

In this paper, a novel technique based on efficiency curves of water discharge and produced power is adopted to develop deterministic and stochastic models. The idea is to use the efficiency curves to determine the efficiency points of water discharge for each combination of active turbines, since the maximum generated power is obtained at these points. Therefore, a pair of points of water discharge and produced power is determined. Figure 1 shows an example of the efficiency points for water discharge and power produced for a single combination of 4 and 5 turbines. The model selects one of these points at each decision time to maximise the energy produced. However, to avoid turbine start-ups, a maximum number of turbine changes is imposed to find a viable solution in the practice. This technique was proposed and tested in [10]. The results show that the proposed model outperforms real operational decisions in terms of generated energy. Moreover, by using the efficiency points the number of parameters and variables is reduced and the problem becomes easier to solve. In addition, the producer can directly implement the optimized solution since it is obtained on the efficiency points and thus matches what the engineers want for powerhouses operations. Since our objective is to find a solution that reflects the operational reality, this technique is applied. In this section, the formulations of the proposed deterministic and stochastic models based on this technique are presented.

2.1 Deterministic model

For the deterministic model, a MILP formulation is developed using the efficiency curves to maximize energy produced and penalize turbine stat-ups. The objective is to select a pair of points of maximum efficiency of water discharge and power produced and to find the best combinations of active turbines that maximize the total energy produced. Since no uncertainties are considered, only one scenario is used to represent the inflow. The problem is solved and the solutions for the volume, the water discharge, the produced power, and the combination of active turbines are determined. The

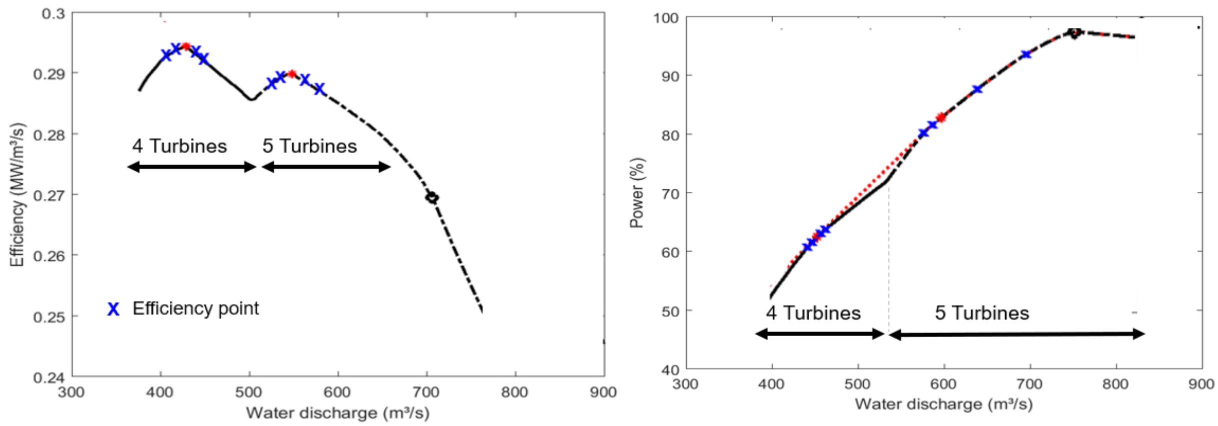


Figure 1: The efficiency curves

volume of the reservoir is updated considering the actual realization of the inflow and the optimization process is repeated during the planning horizon. The decision variables are $y^c_{k,n}$, which allow the model to select the efficiency point k for a given water discharge and power produced, the volume of the reservoir v^c_n, and the state of the turbines $z^c_{j,n}$ for each node n and powerhouse c. The proposed formulation is developed in [10] and is employed in this study. The MILP formulation is given by:

$$\max_{y,v,z}\left[\sum_{c\in C}\sum_{n\in N}\sum_{b\in B^c}\sum_{k\in K^c_b} P^c_{k,n}\times y^c_{k,n} - \sum_{c\in C}\sum_{n\in N}\theta^c\times(V^c_{max}-v^c_n) - \sum_{c\in C}\sum_{n\in N}\sum_{j\in J}\varepsilon^c\times z^c_{j,n}\right]\times\Delta_\tau \tag{1}$$

$$\text{s.t. } v^c_{n+1} = v^c_n + \Delta_\tau\times\left[(\xi^c_n\times\beta) - \sum_{b\in B}\sum_{k\in K^c_b}(q^c_{n,k}\times y^c_{k,n}\times\beta) - (d^c_n\times\beta) + \sum_{l\in U^c}\sum_{b\in B^c}\sum_{k\in k^l_b}(q^l_{n,k}\times y^l_{k,n}\times\beta) + (d^l_n\times\beta)\right],$$

$$\forall c\in C, \forall n\in N \tag{2}$$

$$\sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n+1}\times A^c_{n+1,k,j} - \sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n}\times A^c_{n,k,j} \le z^c_{j,n}, \qquad \forall c\in C, \forall n\in N, \forall j\in J \tag{3}$$

$$\sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n} = 1, \qquad \forall c\in C, \forall n\in N \tag{4}$$

$$\sum_{i\in S}\sum_{n\in N}\sum_{j\in J} z^c_{j,n} \le N^c_{max}, \qquad \forall c\in C \tag{5}$$

$$v^c_{min} \le v^c_n \le V^c_{max}, \qquad \forall c\in C, \forall n\in N \tag{6}$$

$$v^c_0 = v^c_{ini}, \qquad \forall c\in C \tag{7}$$

$$v^c_N \ge v^c_{final}, \qquad \forall c\in C \tag{8}$$

$$y^c_{k,n}, y^l_{k,n}, z^c_{j,n} \in \mathcal{B}, \qquad \forall c\in C, \forall n\in N, \forall b\in B^c, \forall k\in K^c_b, \forall l\in U^c \tag{9}$$

$$d^c_n, d^l_n, v^c_n \in \mathcal{R}^+, \qquad \forall c\in C, \forall n\in N, \forall l\in U^c. \tag{10}$$

Constraints (2) ensure the water balance of the powerhouses. When the powerhouses are connected, the water discharge $q^l_{k,n}$ and the water spillage d^l_n of the upstream powerhouses are added to the volume v^c_n. Constraints (3) establish the relationship between start-up variables and combination choice. Constraints (4) ensure that only one operating point is chosen for each powerhouse at each node. The number of start-ups are limited using constraints (5). Constraints (6)–(8) specify the bounds on the reservoir volumes. Finally, (9) defines the binary variables and (10) the real variables.

2.2 Stochastic model

A SMMILP formulation is developed for the stochastic model to account for uncertain inflows in order to maximize energy produced and penalize turbine start-ups. The uncertain inflows are represented using scenario trees. For each day of the planning horizon, a set of scenarios is provided to present the forecast of inflows. Based on the distribution of these inflows, a full scenario tree is constructed. This full scenario tree is generated and reduced using the backward reduction method [22, 37]. This method aims to minimize the probability distribution distance between the full scenario tree and the reduced tree by iteratively selecting the scenarios to be removed. The concept is to calculate the distance between the reduced and the full scenario trees for each time period and delete scenarios when the reduced tree is sufficiently close to the full tree within a certain accuracy. This accuracy is defined as the reduction percentage. It defines the desired reduction in terms of the measured distance between the full and reduced scenario trees. For example, if the reduction percentage is 10%, it means that the distance between the reduced tree and the full tree is less than 10%. This implies that the reduced tree retains 90% of the information present in the full tree. Figure 2 shows an example of a full scenario tree with 55 scenarios which is generated and reduced using the backward reduction method, resulting in a final reduced scenario tree with 21 scenarios.

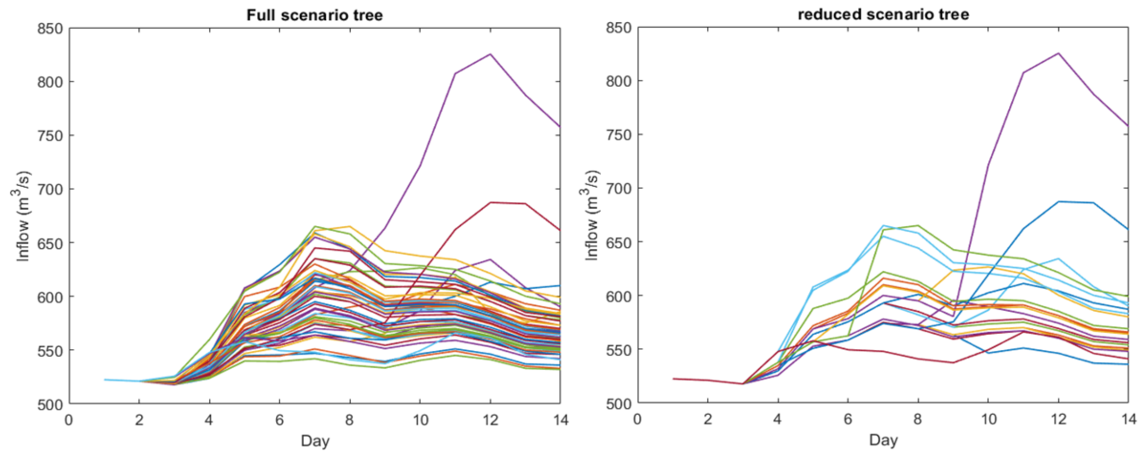


Figure 2: An example of the full and the reduced scenario trees

Backward reduction method is used in several papers [19, 35] and the results prove its effectiveness. In [9] a comparison is made between backward reduction and neural gas methods. A full scenario tree is generated and reduced for each day using the backward reduction and the neural gas methods. Different percentage of reduction (10%, 20% and 30%) are used to build the reduced scenario trees. These reduced scenario trees are used as parameters in the stochastic model. The solutions obtained with both methods are compared, and the results show that the backward method outperforms the neural gas algorithm in terms of preserving statistical information. Moreover, the solutions obtained with 10% of reduction are similar to those obtained with the full tree, but the computation time is higher.

In this paper, the backward reduction method with 10% of reduction is used to generate and reduce the scenario tree. Since the inflow forecasts are updated daily, a scenario tree is generated and then the model is solved for each scenario tree node. Once the actual realizations of the inflows are known, the volume of the reservoir is updated and the scenario tree generation and reduction process and optimization are repeated. The objective is to maximize the total energy production in stage 0 and the expected energy production in future stages. The same technique is used as in the deterministic model, where a pair of points of maximum efficiency of water discharge and power produced, and the best combinations of active turbines are determined. The decision variables are divided into different stages, ensuring that the decisions made in one stage are independent of the information available in subsequent stages. These variables are $y^c_{k,n}$, which allow the model to select the efficiency point k for a given water discharge and power produced, the volume of the reservoir v^c_n, and the state of the turbines $z^c_{j,n}$ for each node n in the scenario tree and powerhouse c. The SMMILP formulation is developed in [10] and is given by:

$$\max_{y,v,z}\Bigg[\sum_{c\in C}\sum_{b\in B^c}\sum_{k\in K^c_b} P^c_{k,0}\times y^c_{k,0} - \sum_{c\in C}\theta^c\times(V^c_{max}-v^c_0) - \sum_{c\in C}\sum_{j\in J}\varepsilon^c\times z^c_{j,0}$$

$$+\sum_{i\in S}\pi^c_i\times\Delta_\tau\times\Big[\sum_{c\in C}\sum_{n\in N_i}\sum_{b\in B^c}\sum_{k\in K^c_b} P^c_{k,n}\times y^c_{k,n} - \sum_{c\in C}\sum_{n\in N_i}\theta^c\times(V^c_{max}-v^c_n)\Big] - \sum_{c\in C}\sum_{n\in N_i}\sum_{j\in J}\varepsilon^c\times z^c_{j,n}\Bigg] \quad (11)$$

$$\text{s.t. } v^c_{n+1} = v^c_n + \Delta_\tau\times\Big[(\xi^c_n\times\beta) - \sum_{b\in B}\sum_{k\in K^c_b}(q^c_{n,k}\times y^c_{k,n}\times\beta) - (d^c_n\times\beta)$$

$$+\sum_{l\in U^c}\sum_{b\in B^c}\sum_{k\in k^l_b}(q^l_{n,k}\times y^l_{k,n}\times\beta) + (d^l_n\times\beta)\Big], \qquad \forall c\in C, \forall n\in N_i, \forall i\in S \quad (12)$$

$$\sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n+1}\times A^c_{n+1,k,j} - \sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n}\times A^c_{n,k,j}\le z^c_{j,n}, \qquad \forall c\in C, \forall n\in N_i, \forall i\in S, \forall j\in J \quad (13)$$

$$\sum_{b\in B^c}\sum_{k\in K^c_b} y^c_{k,n} = 1, \qquad \forall c\in C, \forall n\in N_i, \forall i\in S \quad (14)$$

$$\sum_{i\in S}\sum_{n\in N_i}\sum_{j\in J} z^c_{j,n}\le N^c_{max}, \quad \forall c\in C \quad (15)$$

$$v^c_{min} \leq v^c_n \leq V^c_{max}, \quad \forall c \in C, \forall n \in N_i, \forall i \in S \tag{16}$$

$$v^c_0 = v^c_{ini}, \quad \forall c \in C \tag{17}$$

$$v^c_{N_i} \geq v^c_{final}, \quad \forall c \in C, \forall i \in S \tag{18}$$

$$y^c_{k,n}, y^l_{k,n}, z^c_{j,n} \in \mathcal{B}, \quad \forall c \in C, \forall n \in N_i, \forall i \in S, \forall b \in B^c, \forall k \in K^c_b, \forall l \in U^c \tag{19}$$

$$d^c_n, d^l_n, v^c_n \in \mathcal{R}^+, \quad \forall c \in C, \forall n \in N_i, \forall i \in S, \forall l \in U^c. \tag{20}$$

Constraints (12) ensure water balance at the powerhouses. Constraints (13) are the link between start-up variables and the chosen combination of active turbines considering the set of efficiency points. Constraints (14) force the model to select only one operating efficiency point at each node for each powerhouse. A maximum number of start-ups N_{max} is imposed with constraints (15). Constraints (16)–(18) are the bounds on the reservoir volumes. Finally, constraints (19) define the binary variables and (20) the real variables.

2.3 Optimization problem

In the realm of optimization modeling, deterministic models are simpler and more computationally efficient than stochastic models, but they cannot adequately account for parameter uncertainty. On the other hand, stochastic models can effectively capture uncertainty, but are generally more computationally expensive. To overcome these limitations, a combination of deterministic and stochastic methods can be used during the planning horizon to find a trade-off between achieving an optimal solution and reducing computation time. To achieve this objective, it is important to determine the appropriate model to use during the planning horizon (either deterministic, stochastic, or both) and the optimal time to transition from one model to the other when a transition is required.

For example, Figure 3 illustrates different possibilities for the models used in the optimization process. A deterministic model can be used for the whole planning horizon, a stochastic model can be used for the whole planning horizon, or an alternation between the two models can be used. In this example, a single time change t_1 is proposed. While it is theoretically possible to alternate between models every day, but implementing such a solution may not be practical in practice. Therefore, it is crucial to identify the optimal time for transitioning between the different models. To solve this problem, blackbox optimization can be used to determine the appropriate model choice and the optimal transition times between different models.

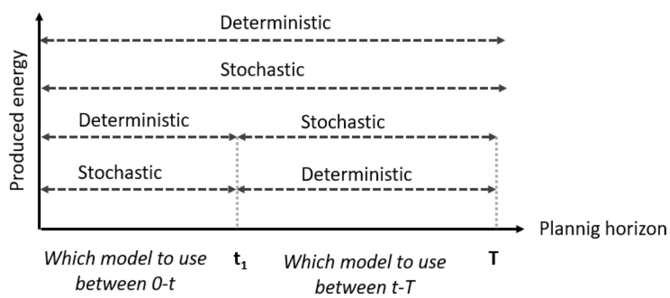


Figure 3: Optimization models

3 Blackbox optimization

In this section, the concept of the blackbox optimization is described, then the formulation of the problem is exposed.

3.1 Blackbox description

In blackbox optimization (BBO), the internal operations or details of the blackbox functions are not explicitly used or modeled. Rather, the optimization algorithm interacts with the blackbox functions by providing input values and returning corresponding output values. The goal is to iteratively search for the optimal solution by exploring the feasible solution space through a series of function evaluations. The general form of an optimization problem is:

$$min_{x \in \Omega} \ f(x) \tag{21}$$

where f is the blackbox.

The mesh adaptive direct search algorithm (MADS) is an algorithm that solves BBO problems [24]. It is an iterative algorithm that evolved from the generalized pattern search (GPS) [5] which in turn evolved from the coordinate search (CS) [2]. The algorithm evaluates the blackbox at some trial points on a spatial discretization called the mesh. The aims is to generate a finite number of trial points on the mesh in order to find a feasible trial point that improves the current best solution. In MADS algorithm, the selection of points is guided by descent directions. The algorithm aims to find points in the search space that move in a direction of descent, leading to improvement in the objective function value. For those purposes, two steps *search* and *poll* are repeated at each iteration until a predefined stopping criterion is reached. In *search* step, trial points are generated anywhere on the mesh with the aim of finding a better solution than the current best one. In *the poll* step, a list of trial points in the vicinity of the current best solution is generated. If a better solution than the current one is found, it is set as the best solution and the mesh size is increased. If the iteration is fails, the next iteration initiated on a finer mesh to improve the solution. MADS is implemented using Nomad software [25] to solve the BBO problem.

3.2 Blackbox formulation

In this paper, BBO is used to determine both the choice of model and the timing of transition from one model to another (when a change is deemed necessary) in order to maximize the total energy produced and penalize turbine start-ups. BBO explores various possibilities and evaluates the performance of the different models to determine the most effective combination. To avoid frequent model changes and to obtain a solution that matches the operational reality, a maximum number of changes Nc^{max} is imposed. According to Nc^{max}, the possible number of the time changes t_m $\forall$ m $\in [1,..,Nc^{max}]$ is defined. The choice of the used model between the time changes is determined using binary variables a_p,$\forall$ $p \in [1,..,Nc^{max}+1]$: 1 means that a deterministic model is chosen during such a period, otherwise, a stochastic model is chosen and therefore the blackbox optimization variables are the times of changes t_m,$\forall$ $m \in [1,..,Nc^{max}]$ and the binary variables a_p,$\forall$ $p \in [1,..,Nc^{max}+1]$.

For example, in the case where only one change is allowed during the planning horizon T (Nc^{max} =1):

- t_1 : is the time change from one model to another (if a change is deemed necessary) as shown Figure 3.
- a_p $\forall$ $p \in [1,2]$: a_1 represents the proposed model between $[0,\ t_1]$ and a_2 represents the proposed model between $[t_1{+}1,\ T]$. For example, if $a_1 = 1$ and $a_2 = 0$, this means that during $[0,t_1]$ the deterministic model is used and during $[t_1{+}1,\ T]$ the stochastic model is used, or, if $a_1 = 1$ and $a_2 = 1$, this means that the deterministic model is used throughout the planning horizon as shown in Figure 3.

Therefore, the blackbox optimization process is described as follows:

1. ***Initialization***: The optimization process starts by initializing the inputs for the blackbox function. These inputs are the time change $t_m \in [0,T]$,$\forall$ $m \in [1,..,Nc^{max}]$ and binary variables a_p,$\forall$ $p \in [1,..,Nc^{max}+1]$.
2. ***Evaluation***: The blackbox function EBB(t_m,a_p), $\forall$ $p \in [1,..,Nc^{max}+1]$, $\forall$ $m \in [1,..,Nc^{max}]$ that maximizes the total energy produced and penalizes the start-ups whole the planning horizon is evaluated according to the initial input set. If the binary variable a_p,$\forall$ $p \in [1,..,Nc^{max}+1]$ is equal to 1, the MILP formulation is used for a specific period based on the given value of the time change $t_m \in [0,T]$. Otherwise, the SMMILP formulation is used. This involves running the function and obtaining the corresponding output value as shown in Figure 4.
3. ***Solution Update***: The obtained output value is compared with the current best solution. If the new output value is better, it becomes the new best solution. The blackbox solver then updates its internal state to reflect this improvement.
4. ***Search for Better Solutions***: The blackbox solver produces new inputs in order to find a better solution than the best one.
5. ***Check for Convergence***: The blackbox solver analyzes the newly obtained output values and compares them to the best solution. If a better solution is found, the process continues to search for further improvements. However, if no better solution is found or a convergence criterion is met (10000 evaluations in this work), the optimization process is terminated.
6. ***Output***: Once the optimization process is completed, the best optimization model to use during the planning horizon and the optimal time changes are returned.

The objective function of the BBO problem is denoted by:

$$\max_{a,t} \quad \sum_{d\in T} EBB_d(t_m,a_p), \qquad \forall p \in [1,..,Nc^{max}+1], \forall m \in [1,..,N_{max}] \tag{22}$$

$$\text{s.t.} \quad \sum_{m\in[1,..,Nc^{max}]} t_m \quad \leq \quad Nc^{max} \tag{23}$$

$$t_{m-1} - t_m \quad \leq \quad 1, \quad \forall m \in [1,..,Nc^{max}] \tag{24}$$

$$t_m \quad < \quad T, \quad \forall m \in [1,..,Nc^{max}] \tag{25}$$

$$a_p \quad \in \quad \mathcal{B}, \quad \forall p \in [1,..,Nc^{max}+1] \tag{26}$$

$$t_m \quad \in \quad \mathbb{N}, \quad \forall m \in [1,..,Nc^{max}] \tag{27}$$

where $d \in T$ is the period (day) in the rolling horizon.

Constraints (23) ensure that the maximum number of time changes is imposed. Constraints (24) and (25) are the bounds on time changes. Constraints (26) define the binary variables and (27) the natural variables.

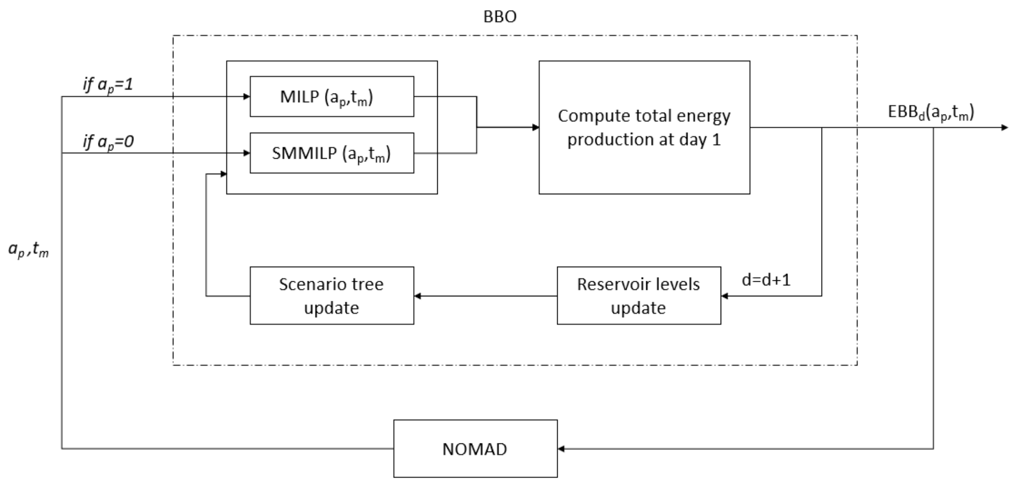


Figure 4: Blackbox

4 Computational results

This section details the case study and the results on which the BBO problem is solved with the objective of defining the optimization model used during the planning horizon to maximize the energy produced and penalize turbine start-ups.

4.1 Case study

In this paper, the proposed approach is tested with real data from the Saguenay-Lac-St-Jean hydroelectric system owned by Rio Tinto. It is a company that produces aluminium in the Saguenay region. They own a hydroelectric system with five powerhouses that provides 90% of its energy needs. The remaining energy needs are purchased at a known price from Hydro-Québec, a public company responsible for the generation, transmission and distribution of electricity. This means that all producers must buy and sell to them and negotiate fixed price contracts every year. Therefore, the energy price is not considered and only the uncertainties of the inflows are considered. For the purpose of this paper, the analysis focuses on two in-series powerhouses, chute-savane (CS) and chute-du-diable (CD). Each of these powerhouses is equipped with 5 turbines and has two reservoirs with different capacities. A rolling-horizon methodology is used to validate the optimization approach. The planning horizon of the rolling-horizon is of 10 days. The forecast is for 14 days. For day 1 of the rolling-horizon, the forecasts are for days 1 through 14, for day 2 of the rolling-horizon, the forecasts are for days 2 through 15, and so on. For each day of the rolling-horizon, a set of scenarios is provided to represent the forecast of inflows and the aim is to optimally dispatch the amount of water available in the reservoirs in order to maximize the energy produced and penalize turbine start-ups. Decisions are made hourly for the first day and daily for the next 13 days, and only the solutions for the volume, the water discharge, the produced power and the combination of active turbines for the first hour of the first-stage are retained. Then, the volume of the reservoir is updated considering the actual realization of the inflow and the optimization process is repeated during the 10 days. The objective of this work is to determine the appropriate optimization model, either deterministic, stochastic, or a combination of both, that can be used during the rolling-horizon in order to maximize the total energy produced. The BBO problem is developed to solve this problem. For the deterministic model, the median scenario of all available scenarios is used for every day to represent the inflows, and a scenario tree of 1 node per stage is solved using a MILP formulation. For the stochastic model, the inflow scenarios are generated and reduced using backward reduction method for every day of the rolling-horizon. The problem is solved using a SMMILP formulation. Since stochastic models suffer from long computation times, a small study is conducted to determine the feasibility of using an aggregate stochastic model in order to reduce the computational time. The optimal solutions provided by the BBO solver are analysed to investigate the choice of optimization models during the planning horizon. The formulations are solved using the servers of compute Canada [28] and Xpress solver accessed via Python [14], and the BBO problem is solved using the solver NOMAD [25]. The proposed model is tested using five data sets (July, September, October and December) for the year 2021 and September for the year 2020.

4.2 Aggregated stochastic model

To determine the appropriate optimization model, either deterministic, stochastic, or a combination of both, a BBO problem is used. The main difficulty of the BBO problem resides in the computational time. This is due to the fact that, the SMMILP or/and the MILP formulations are called at each evaluation. The problem is solved at each node of the scenario tree during the 10 days of the rolling-horizon and for 14 forecast days. For the MILP formulation, only one scenario (the median scenario) is used, while for the SMMILP the scenario tree contains several scenarios. Moreover, for every day of the rolling-horizon, the decisions are made hourly for the first day and daily for the next 13 days, which increases the number of optimization variables and the problem become hard to solve. To reduce the computational time, the stages in the SMMILP formulation can be aggregated. Instead of optimizing the decisions for each day, the optimization algorithm focuses on optimizing the decisions for the aggregated intervals. Aggregating days reduces the number of decision variables and constraints in the optimization problem. This leads to faster computations. For that, a small study is conducted in this paper to determine the feasibility of using an aggregated model instead of the stochastic model. However, it's important that the aggregation entails a trade-off between computational efficiency and solution accuracy. Since aggregation can be applied in different ways, two different options (Op1 and Op2) are tested as shows Table 1.

Table 1: The tested possibilities of aggregation

Option	Decisions
Op1	hourly for the first stage, every 2 days for the next 5 stages and 3 days for the last stage
Op2	hourly for the first stage, every 3 days for the next 3 stages and 4 days for the last stage

A comparison between the energy production and computational time for each available option is made to choose the best option that achieves a trade-off between computational time and solution accuracy. Five data sets (July, September, October and December) for the year 2021 and September for the year 2020 are used. Table 2 illustrates the energy produced for the whole planning horizon obtained from the aggregated model with both options (Op1 and Op2) and the non-aggregated model. Negative values indicate that the first option (Op1) provides slightly higher energy compared to the second option (Op2). The results show that the difference in generated energy between the two options is very small. For example, in September 2021, the difference is only 0.02%.

Table 2: The difference of the total energy production (GWh) between non-aggregated and aggregated models

Data sets	Non-aggregated (10^3)	Aggregated Op1 (10^3)	Aggregated Op2 (10^3)	Difference Op1/Op2 (%)
September 2020	101.989	101.901	101.831	-0.06
July 2021	82.6986	82.2410	82.0337	-0.25
September 2021	78.4234	78.3810	78.3653	-0.02
October 2021	72.6545	72.5838	72.5116	-0,09
December 2021	92.045	91.8205	91.751	-0.07

The average time required to optimize a single day in the rolling-horizon process for the different models is calculated and presented in Table 3. The results indicate that optimize with option (Op2) is significantly faster when compared to option (Op1). For example, in September 2020, using the aggregated model with the option (Op2) is 12 times faster than the non-aggregated model and 10 times faster than the option (Op1), moreover the difference between the energy produced with both options is only 0.06%. Based on these results, the decision is made to aggregate the stages using option (Op2), where decisions are made hourly for the first day, every three days for the next three stages, and four days for the last stage. However, only the decisions for the volume of the reservoir, the water discharge, the power generated and the turbines in operation of the first stage are kept.

Table 3: The difference of the computational time between non-aggregated and aggregated models

Data sets	Non-aggregated (min)	Aggregated Op1 (min)	Aggregated Op2 (min)
September 2020	12	10	1
July 2021	12	11	1.2
September 2021	11	9	1.1
October 2021	12	10	1.2
December 2021	12	10	1.2

Let us now compare the solutions of the aggregated model using option (Op2) and the non-aggregated model to evaluate the accuracy of the solutions. Figure 5 shows an example of the proposed decisions for the month of December: the power produced in (MW), the amount of water discharge in (m^3/s), the volume of the reservoir in (hm^3), and the number of turbines in operation for the aggregated model with option (Op2) (dashed line) and for the non-aggregated model (solid line) for two powerhouses CD and CS.

As shown in Figure 5, the solutions proposed with the aggregated and the non-aggregated models are broadly similar. Moreover, the results are obtained 10 times faster.

The results of this small study show that the aggregation of the stage with the option (Op2) offers advantages in terms of computational efficiency and solution quality. This aggregated model is used when solving the BBO problem.

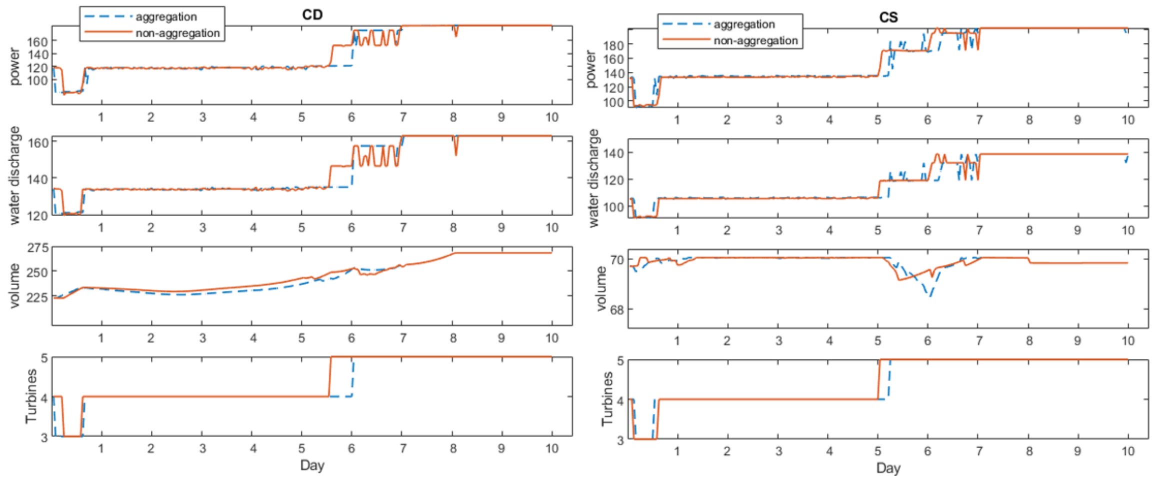


Figure 5: Comparison of the solutions obtained from the aggregated and non-aggregated models

4.3 Blackbox solution

In this section we report the results obtained with the proposed BBO approach. The objective is to define the used optimization model during the planning horizon in a way that maximizes the total energy produced. For that, at each iteration, the solver provides inputs to the blackbox where the MILP and/or SMMILP formulations are called based on the blackbox variables. Then, the solver collects the output (total energy produced) and uses it to generate new inputs to the blackbox in order to find a better solution than the best one. The process is repeated until no improved solution is found or a convergence criterion is met (10000 evaluations in this work). As explained in Section 3.2, the maximum number of time change is defining at the beginning. In this work, two tests are made: first, only one change is allowed during the whole planning horizon. Second, two changes are allowed. The computational experiments are presented in next sections.

4.3.1 Computational experiments

The BBO model is tested using five data sets (July, September, October and December) for the year 2021 and September for the year 2020. Table 4 illustrates the outputs of the blackbox. BBO1 present the obtained energy produced where only one time change is allowed, and BBO2 when 2 time changes are allowed.

Table 4: Total energy production (GWh) from the optimization models

Data sets	Deterministic (10^3)	Stochastic (10^3)	BBO1 (10^3)	BBO2 (10^3)	Proposed model
July 2021	81.3712	82.0337	82.0337	82.0337	Stochastic
October 2021	72.1565	72.5116	72.5116	72.5116	Stochastic
September 2021	78.0841	78.3653	78.3759	78.3759	Deterministic & Stochastic
September 2020	101.451	101.831	101.831	101.999	Deterministic & Stochastic
December 2021	91.727	91.751	91.820	91.827	Deterministic & Stochastic

The results show: for July and October 2021, using the stochastic model during the planning horizon is the best model because maximum energy production is achieved when only the stochastic model is used. For September 2020, December, and September 2021, more energy can be generated by using both models throughout the planning horizon. Table 5 shows the proposed solutions for the cases where both stochastic and deterministic models are used, allowing for a maximum of two changes.

Table 5: Proposed solutions from blackbox optimization

Data sets	Number of changes	Solutions
September 2021	1	6 days stochastic - 4 days deterministic
December 2021	2	6 days deterministic - 2 days stochastic - 2 days deterministic
September 2020	2	1 day stochastic- 1 day deterministic - 8 days stochastic

For example, for September 2021, the blackbox suggests using a stochastic model for the first 6 days and a deterministic model for the next 4 days. Therefore, a time change (at the end of day 6) is proposed, although the maximum number of changes is 2. For December 2021, the blackbox proposes two changes, using the deterministic model for the first six days, the stochastic model for two days, and finally the deterministic model for the last two days.

To evaluate the quality of the solutions provided by the BBO solver using the aggregated model, the results obtained by solving the BBO problem with the aggregated model are compared to the results obtained by the BBO with the non-aggregated model. The results for the month of September 2020 are tested because the computation time in this month is short compared to other months.

Table 6 shows that the total energy produced from the BBO with the aggregated model is quite similar to the total energy produced from the BBO with the non-aggregated model. Moreover, both models propose the same solutions. However, the solution of the BBO with the aggregated model is 18 times faster than the non-aggregated model and allows 6 times less use of memory capacity, confirming the appropriateness of the choice of the aggregation model.

Table 6: Proposed solutions from blackbox model using the aggregated and non-aggregated model

Data sets	Energy produced in GWh (10^3)	Solutions
Aggregated	101.9994	1 day stochastic- 1 day deterministic - 8 days stochastic
non-aggregated	102.0767	1 day stochastic- 1 day deterministic - 8 days stochastic

4.3.2 Analysis of the results

The solutions obtained from the BBO1, BBO2, deterministic and stochastic models for the 5 test cases are analyzed and compared. For example, Figure 6 shows the power produced, the amount of water discharge, the volume of the reservoir and the number of turbines in operation for the two powerhouses CD and CS for the month of December 2021.

According to Table 5, for December 2021, a deterministic model is used for the first six days, followed by a stochastic model for the next two days, and a deterministic model for the last two days. This distinction can be seen in Figure 6, where the curves of the BBO2 and deterministic models match up to day 6. From day 6 to 8, a noticeable change in the curve's shape occurs and after day 8, the curves of the BBO2 and deterministic models overlap again. The same observations hold for the other cases.

To gain insight into the factors influencing the choice of the optimization model, an analysis is conducted on the expected volume of inflows for both the stochastic and deterministic models. For the deterministic model, the expected volume of the median scenario is calculated at each day of the rolling horizon. For the stochastic model, since different scenarios are available, the expected volume is calculated as follows:

$$EV = \sum_{i \in S} \pi_i \times \sum_{n \in N_i} \xi_n \tag{28}$$

where S is the number of scenarios in the scenario tree, π_i is the probability of each scenario $i \in S$ and ξ_n are the values of the inflows at each node $n \in N_i$.

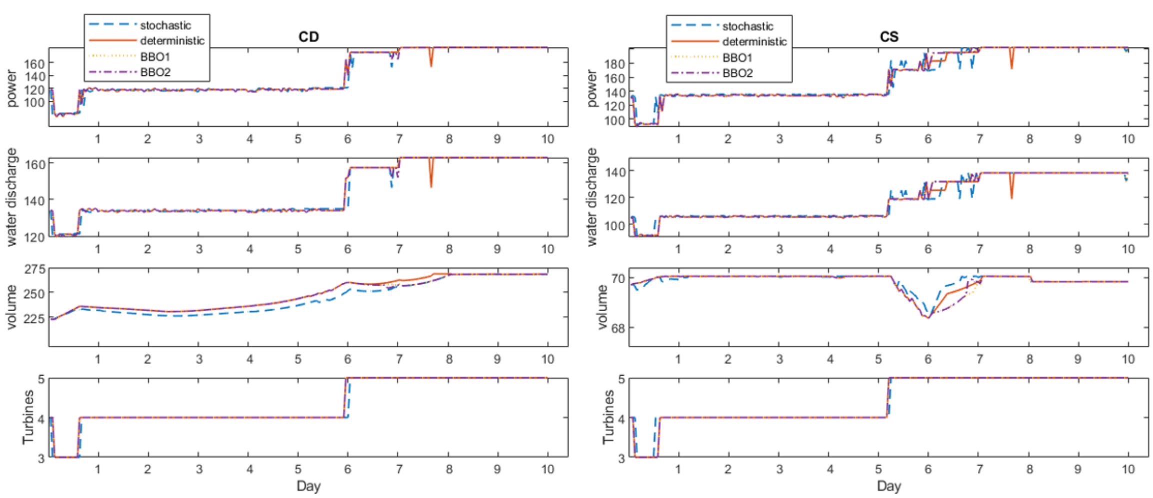


Figure 6: Solutions for stochastic, deterministic, BBO1 and BBO2 models for the December test case

Figure 7 illustrates the expected volume of inflows for both models for the powerhouses CD and CS. The results show that the expected volumes are quite similar for the first six days before a significant discrepancy occurs between the 6th

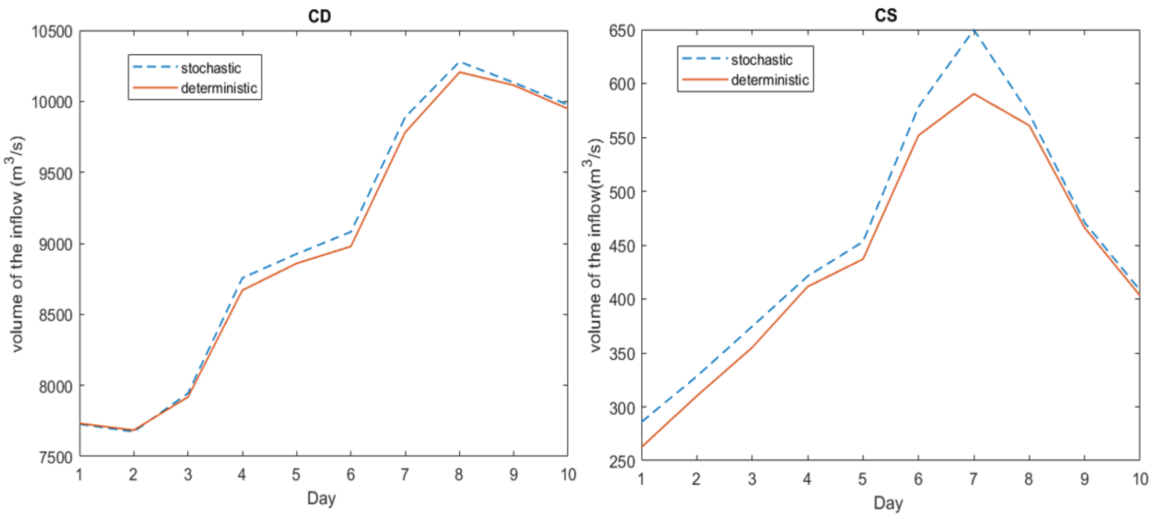


Figure 7: The expected volume of inflows for the stochastic and deterministic models for the December case

and 8th day, especially for the powerhouse CS. After the eighth day, the expected volume of inflows for the stochastic and deterministic models converge again and become quite similar which is consistent with the results obtained in Table 5.

For September 2021 and September 2020, the same observations hold. For example, Figure 8 illustrates the expected volume of inflows for both models for September 2021. It is worth noting that Table 5 confirms the use of a stochastic model for the first six days, and this distinction is clearly evident in Figure 8 where a noticeable difference in the expected volume of inflows can be observed during this period for the powerhouse CS.

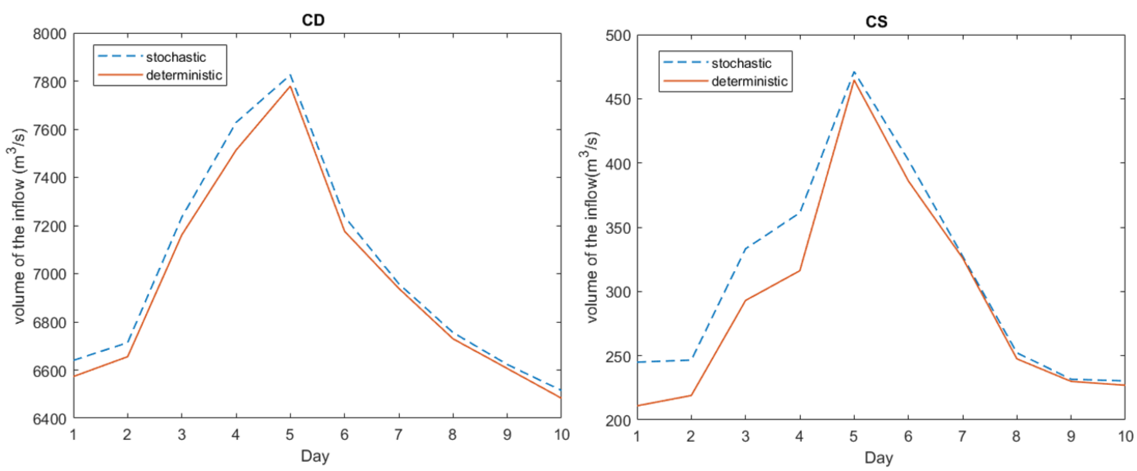


Figure 8: The expected volume of inflows for the stochastic and deterministic models for the September 2021 case

Let us analyze the cases where the stochastic model is throughout the entire planning horizon. For example, Figure 9 illustrates the expected volume of inflows for the month of July. For both powerhouses, it is noticed that a substantial disparity exists between the expected volume of inflows for the stochastic and deterministic models.

According to Figures 7, 8, 9 and Table 5, we notice that the deterministic model is used when the expected volume of inflows from deterministic and stochastic models are quite similar. Otherwise, a stochastic model is used. To gain further clarity, a study of inflow variability was conducted to understand its impact on the choice of optimization model. This analysis can provide valuable insight into how the choice of optimization model may be affected by the observed variability of inflows. For this purpose, a boxplot is used, as shown in Figure 10.

A boxplot is a graph that indicates how the values in the data are spread out based on five sample statistics : the minimum, the maximum, the first quartile(Q1), the median and the third quartile(Q3). The lower box represents the first quartile and the top of the box represent the third quartile. The middle of the box represents the median. The top of the line linking the box to the whisker is the maximum of the data and the bottom is the minimum of the data. Except for the five sample statistics, the boxplot shows the outliers which ares the values that are much higher or much lower compared to the other values in the set (usually 1.5 times the interquartile range IQR =Q3-Q1). Using the boxplot gives an indication of the sample variability by examining the range (length of the entire box whose the length of the whiskers), IQR, or by the outliers.

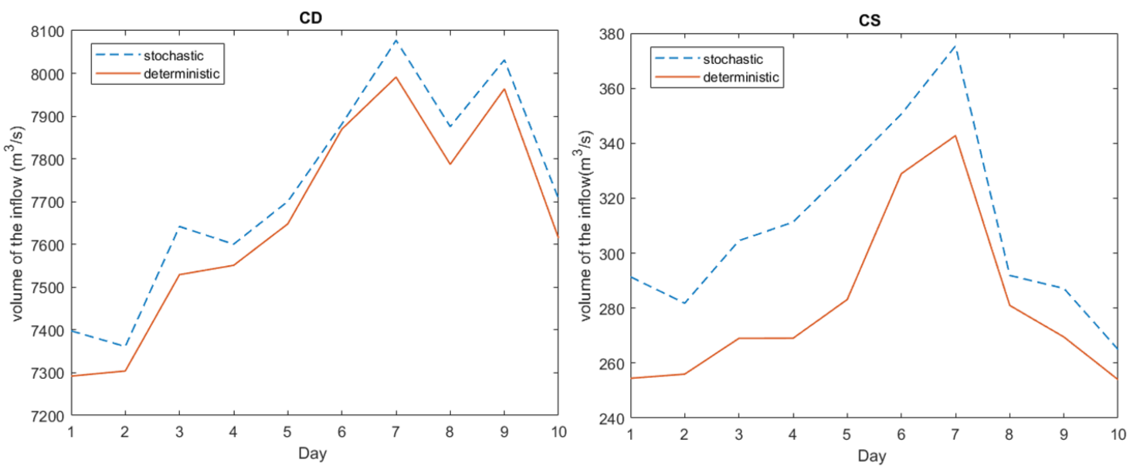


Figure 9: The expected volume of inflows for the stochastic and deterministic models for the July 2021 case

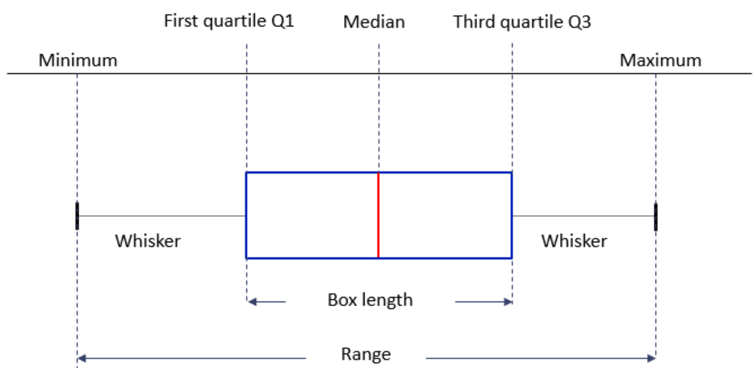
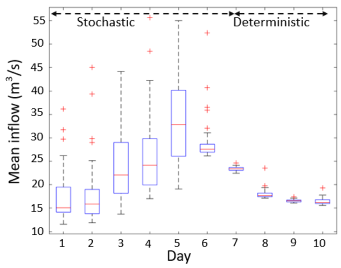


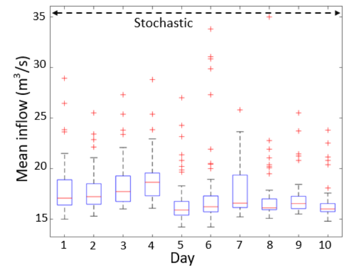
Figure 10: A boxplot

In this paper, the boxplot is used to define the variability of the inflows for the 5 test cases. For that, the mean of each scenario in the scenario tree is calculated and presented with a boxplot for every day in the rolling-horizon. For example, Figure 11 shows the boxplots for September and October 2021. As shown in Table 5, for September 2021 the stochastic model is used for the first 6 days and the deterministic one for the last 4 days. According to Figure 11a, it is observed that the size of the boxplot up to day 6 is very important compared to the sizes of the boxplots from day 7 onward. For example, for day 5, the range of the boxplot =*37 (55-18)* and for day 7, the range = *2*. This means that the variability of the inflows in the first six days is very large. For this reason, the stochastic model is used for these days. The same is noticed for December 2021 and September 2020. In addition, it is noted that the blackbox of the mean inflow suggests the use of the stochastic model when the range of the boxplot exceeds 10.

For October 2021, the stochastic model is used for the entire planning horizon, although the size of the boxplot is not important for these days. This is because there are many outliers during this period (the red dots in the Figure 11b). These outliers represent the extreme scenarios. For this reason, the stochastic model generates more energy as the extreme scenarios are considered unlike the deterministic model. The same is noticed for the month of July. Based on the analyses conducted throughout this study, it appears that the choice between a deterministic model, a stochastic model, or both, depends on certain conditions. When the expected volume of inflows has a high similarity between the deterministic and stochastic models and the variability of inflows is relatively low (i.e., the range of the boxplot of the mean inflows is less than 10 or the number of extreme scenarios is very small), the deterministic model proves to be a viable option, offering the advantage of low computational time. However, when the expected volume of inflows has significant differences and the variability of inflows is high, the stochastic model is strongly recommended. These results demonstrate the importance of carefully considering the degree of variability of uncertainties when selecting appropriate models.



(a) September 2021



(b) October 2021

Figure 11: The boxplot for the mean inflow for September and October 2021 cases

5 Conclusion

This paper presents a study on the choice of optimization models. The objective is to determine the best model that maximizes the energy produced during the planning horizon. Therefore, deterministic and stochastic models can be used in different periods of the planning horizon. For the deterministic model, the median scenario of all available scenarios is determined for every day and a scenario tree with one node per stage is solved using a MILP formulation. For the stochastic model, the inflow scenarios are generated and reduced using backward reduction method for every day of the planning horizon. The problem is solved using a SMMILP formulation. The stages are aggregated to reduce the computation time. The best optimization model is formulated as a blackbox optimization problem. Therefore, at each iteration, the solver provides inputs to the blackbox where the MILP and/or SMMILP formulations are called. Then, the solver collects the output (total energy produced) and uses it to generate new inputs to the blackbox in order to find a better solution than the best one. The process is repeated until no improved solution is found or a convergence criterion is met (10000 evaluations). The problem is tested for five test cases. First, the aggregated model is compared to the non-aggregated one in order to evaluate the quality of the solution. The results show that the solutions are quite similar. Second, the proposed solutions from the blackbox problem are presented and analyzed. The results show that, for July and October 2021, using the stochastic model for the whole planning horizon allows to produce more energy. This is due to the fact that in these months the expected volume of inflows from the scenario tree method is significantly different from that of the median scenario. Moreover, the variability of inflows is very high (the range of the boxplot of the mean inflows exceeds 10 or numerous extreme scenarios are observed). For the months of September, December 2021, and September 2020, more energy can be generated with both the deterministic and stochastic models. Moreover, the results show that in some periods, where the expected volume of inflows of the scenario tree method is quite similar to that of the median scenario and the variability of inflows is low, the deterministic model is used for these periods. Therefore, it is important to thoroughly consider the degree of variability of uncertainties when selecting the appropriate models for optimal hydropower management. For future research, we intend to conduct experiments with a more extensive data set. Although the current approach was tested with 5 datasets, expanding to a larger and more diverse data set will provide valuable insights and explore a metric that can effectively reveal the best optimization model.

References

[1] Akbari Asanjan, A., Yang, T., Hsu, K., Sorooshian, S., Lin, J., and Peng, Q. Short-term precipitation forecast based on the persiann system and lstm recurrent neural networks. Journal of Geophysical Research: Atmospheres 123, 22 (2018), 12–543.

[2] Amaioua, N. Modèles quadratiques et décomposition parallèle pour l'optimisation sans dérivées. Ecole Polytechnique, Montreal (Canada), 2018.

[3] Anjos, M. F., Conejo, A. J., et al. Unit commitment in electric energy systems. Foundations and Trends® in Electric Energy Systems 1, 4 (2017), 220–310.

[4] Arvidsen, H. I. Hydropower production scheduling using stochastisc dual dynamic programming subject to environmental constraints. Master's thesis, The University of Bergen, 2019.

[5] Audet, C., and Dennis Jr, J. E. Analysis of generalized pattern searches. SIAM Journal on optimization 13, 3 (2002), 889–903.

[6] Audet, C., and Hare, W. Derivative-free and blackbox optimization. Springer (2017).

[7] Backe, S., Ahang, M., and Tomasgard, A. Stable stochastic capacity expansion with variable renewables: Comparing moment matching and stratified scenario generation sampling. Applied Energy 302 (2021), 117538.

[8] Cheng, C., Wang, J., and Wu, X. Hydro unit commitment with a head-sensitive reservoir and multiple vibration zones using milp. IEEE Transactions on Power Systems 31, 6 (2016), 4842–4852.

[9] Daadaa, M., Séguin, S., Anjos, M., and Demeester, K. Quantifying the impact of scenario tree generation meth-ods on the solution of the short-term hydroscheduling problem. Les Cahiers Du GERAD (2022), 2440.

[10] Daadaa, M., Séguin, S., Demeester, K., and Anjos, M. F. An optimization model to maximize energy generation in short-term hydropower unit commitment using efficiency points. International Journal of Electrical Power & Energy Systems 125 (2021), 106419.

[11] Diniz, A. L., and Souza, T. M. Short-term hydrothermal dispatch with river-level and routing constraints. IEEE Transactions on Power Systems 29, 5 (2014), 2427–2435.

[12] Dires, F. G., Amelin, M., and Bekele, G. Inflow scenario generation for the ethiopian hydropower system. Water 15, 3 (2023), 500.

[13] Faria, E., and Fleten, S.-E. Day-ahead market bidding for a nordic hydropower producer: taking the elbas market into account. Computational Management Science 8 (2011), 75–101.

[14] FICO. FICO Xpress Optimization Suite. Available at: https://www.msi-jp.com/xpress/learning/square/optimizer-2015.pdf, 2014.

[15] Finardi, E. C., and Da Silva, E. Unit commitment of single hydroelectric plant. Electric Power Systems Research 75, 2–3 (2005), 116–123.

[16] Finardi, E. C., and Scuzziato, M. R. Hydro unit commitment and loading problem for day-ahead operation planning problem. International Journal of Electrical Power & Energy Systems 44, 1 (2013), 7–16.

[17] Finardi, E. C., Silva, E. L. d., and Sagastizábal, C. Solving the unit commitment problem of hydropower plants via lagrangian relaxation and sequential quadratic programming. Computational & applied mathematics 24 (2005), 317–342.

[18] Fleten, S.-E., and Kristoffersen, T. K. Short-term hydropower production planning by stochastic programming. Computers & Operations Research 35, 8 (2008), 2656–2671.

[19] Follestad, T., Wolfgang, O., and Belsnes, M. M. An approach for assessing the effect of scenario tree approximations in stochastic hydropower scheduling models. In Proc. of the 17th Power System Computation Conference (2011), pp. 271–277.

[20] Gams. GAMS SCENRED2. Available at: https://www.gams.com/latest/docs/T_SCENRED2.html, 2023.

[21] Gil, E., Bustos, J., and Rudnick, H. Short-term hydrothermal generation scheduling model using a genetic algorithm. IEEE Transactions on power systems 18, 4 (2003), 1256–1264.

[22] Growe-Kuska, N., Heitsch, H., and Romisch, W. Scenario reduction and scenario tree construction for power management problems. In 2003 IEEE Bologna Power Tech Conference Proceedings, (2003), vol. 3, IEEE, pp. 7–pp.

[23] Guedes, L. S., de Mendonça Maia, P., Lisboa, A. C., Vieira, D. A. G., and Saldanha, R. R. A unit commitment algorithm and a compact milp model for short-term hydro-power generation scheduling. IEEE Transactions on Power Systems 32, 5 (2016), 3381–3390.

[24] Le Digabel, S. Algorithm 909: Nomad: Nonlinear optimization with the mads algorithm. ACM Transactions on Mathematical Software (TOMS) 37, 4 (2011), 1–15.

[25] Le Digabel, S., Tribes, C., Montplaisir, V. R., and Audet, C. Nomad user guide version 3.9. 1, 2019.

[26] Li, B., Sedzro, K., Fang, X., Hodge, B.-M., and Zhang, J. A clustering-based scenario generation framework for power market simulation with wind integration. Journal of Renewable and Sustainable Energy 12, 3 (2020), 036301.

[27] López-Salgado, C. J., Añó, O., and Ojeda-Esteybar, D. M. Stochastic unit commitment and optimal allocation of reserves: A hybrid decomposition approach. IEEE Transactions on Power Systems 33, 5 (2018), 5542–5552.

[28] of Canada, D. R. A. Digital Research Alliance of Canada. Available at: https://docs.alliancecan.ca/wiki/Technical_documentation, 2022.

[29] Philpott, A. B., Craddock, M., and Waterer, H. Hydro-electric unit commitment subject to uncertain demand. European Journal of Operational Research 125, 2 (2000), 410–424.

[30] Postolov, B., and Iliev, A. Adaptive genetic algorithm for hydro-thermal unit commitment considering the security constraints. International Journal of Electrical Engineering and Computing 4, 2 (2020), 61–69.

[31] Séguin, S. Hydropower optimization. Les Cahiers du GERAD ISSN 711 (2020), 2440.

[32] Séguin, S., Audet, C., and Côté, P. Scenario-tree modeling for stochastic short-term hydropower operations planning. Journal of Water Resources Planning and Management 143, 12 (2017), 04017073.

[33] Séguin, S., Côté, P., and Audet, C. Self-scheduling short-term unit commitment and loading problem. IEEE Transactions on Power Systems 31, 1 (2015), 133–142.

[34] Taktak, R., and D'Ambrosio, C. An overview on mathematical programming approaches for the deterministic unit commitment problem in hydro valleys. Energy Systems 8 (2017), 57–79.

[35] Vardanyan, Y., and Hesamzadeh, M. R. The coordinated bidding of a hydropower producer in three-settlement markets with time-dependent risk measure. Electric Power Systems Research 151 (2017), 40–58.

[36] Wapet, D. E. M., Essiane, S. N., Wamkeue, R., Gnetchejo, P. J., et al. Hydropower production optimization from inflow: case study of songloulou hydroplant. Journal of Power and Energy Engineering 8, 08 (2020), 37.

[37] Xu, B., Zhong, P.-A., Zambon, R. C., Zhao, Y., and Yeh, W. W.-G. Scenario tree reduction in stochastic programming with recourse for hydropower operations. Water Resources Research 51, 8 (2015), 6359–6380.

[38] Zhao, Z., Cheng, C., Jin, X., Liu, L., and Yan, L. A milp model for hydro unit commitment with irregular vibration zones based on the constrained delaunay triangulation method. International Journal of Electrical Power & Energy Systems 123 (2020), 106241.