Decision Making Models for Power Systems Operations Under Uncertainty

by

Carlos Olivos Matus

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 5, 2023

Keywords: unit commitment, power systems operations, stochastic programming, stochastic modeling, optimization

Approved by

Jorge Valenzuela, Chair, Philpott-WestPoint Stevens Professor of Industrial & Systems Engineering, Auburn University
Aleksandr Vinel, Associate Professor of Industrial & Systems Engineering, Auburn University
Daniel Silva, Associate Professor of Industrial & Systems Engineering, Auburn University
Gregory Purdy, Assistant Professor of Industrial & Systems Engineering, Auburn University
Iñigo de la Parra Laita, Associate Professor, Public University of Navarra
Timothy McDonald, Professor of Biosystems Engineering, Auburn University

Abstract

As the power grids worldwide transition from fossil fuel-based power plants to an era dominated by renewable energy sources (RESs), new operational challenges and uncertainties emerge, particularly for Independent System Operators (ISOs) responsible for grid reliability. The unpredictable nature of RESs introduces challenges in energy production, as evidenced by effects like solar power surges and wind ramping effects. Stochastic modeling can effectively address these challenges, aiming to optimize performance metrics such as expected value, worst-case scenario, and risk. Therefore, the development of new models to handle these uncertainties is necessary.

In power system operations, ISOs must ensure the operability and reliability of the system. This is done by planning the energy generation schedule during the day-ahead market in a process known as market clearing. ISOs guarantee a competitive interplay between energy supply and demand in this process, resulting in an energy generator schedule and a competitive market price. This is done by solving the Unit Commitment Problem (UCP), a mathematical optimization problem that guarantees system operability and reliability at a minimal cost.

When planning the energy generation schedule, uncertainties such as demand, energy production from RES, and contingencies must be considered. This is usually done using a scenario-based approach, where realizations of uncertainties are generated from fitted probability distributions to be incorporated into the UCP. Thus, the commitment schedule is generated, such as minimizing the expected cost. However, the quality and complexity of solutions depend on the number of scenarios. As uncertainty rises with increased RESs projects, energy generation becomes unpredictable, necessitating more scenarios for accurate cost approximation. This, in turn, increases computational complexity, highlighting the limitations of the scenario generation method. Therefore, it has become critical to devise new approaches to model this problem in the face of such uncertainty.

This dissertation provides a new methodology to model the Stochastic Unit Commitment Problem (SUCP) that relies directly on the probability distribution of the random variables, the so-called statistical SUCP. In Chapter 2, a dispatch cost function is derived considering a two-stage approach where the commitment decisions are made in the first stage and remain fixed during the planning horizon. In the second stage, the energy is dispatched based on the commitment decisions of the first stage and the random variables. Then, the expected dispatch cost is derived using the probability distribution of the random residual demand. Thus, an analytical function of the expected dispatch cost is formulated. Since this function is nonlinear, a piecewise linear approximation method is used to linearize the model. The breakpoints of the piecewise linear approximation are determined by stability analysis. A sensitivity analysis is performed to assess the behavior of the expected cost under different levels of hourly correlations of the residual demand. Moreover, a reliability analysis assesses the Loss of Load Probability of the resulting commitment schedule. In Chapter 3, another layer of realism is included in the model by incorporating ramping constraints, ensuring a smooth transition in power output levels from fuel-based generators. Since adding these constraints increases the complexity of the model, different solving strategies are proposed to solve realistic instances of the SUCP. Finally, in Chapter 4, a comparison is made between the proposed statistical SUCP and the scenario-based SUCP to assess the computational complexity, optimality, and stability of the solutions. The comparison is evaluated based on various power system sizes, breakpoints, and scenarios, offering insightful knowledge about the benefits of each model.

This dissertation delivers valuable advancements for modeling the SUCP, adding another perspective for ISOs and decision-makers when planning the day-ahead energy generation schedule. The benefits of this research extend to further optimizing power grid operations, reducing costs, and ensuring a reliable transition to renewable energy sources.

Acknowledgments

I would like to express my heartfelt gratitude to my parents, Eduardo Olivos and Marianella Matus. Their support throughout my life and constant encouragement has pushed me forward. They are an integral part of this achievement. My gratitude to my sisters Lorena and Paulina for your continuous care and encouragement. I also extend my warmest thanks to the rest of my family, who consistently uplifted my spirit.

I greatly thank my advisor, Dr. Jorge Valenzuela, who initially invited me to Auburn to pursue my doctoral studies. His thoughtfulness, patience, and support were fundamental to this journey. Furthermore, I extend my gratitude to the professors who served on my committee. I appreciate their dedication, willingness to invest time in my dissertation, and collaborative spirit on research projects. Special thanks to Dr. Daniel Silva, Dr. Alexander Vinel, and Dr. Gregory Purdy for their assistance during these years.

To my friends in Auburn: Pablo Hurtado, Patricia Carcamo, Julia Bitencourt, Ana Wooley, and Duha Ali - your friendship, support, and shared moments of happiness have been pillars of comfort, making this journey more enjoyable and less arduous. I want to thank my friends Evelyn Arrey, Hernan Caceres, and Rocio Palominos, whose persistent encouragement led me on this graduate journey. Thanks for your attentive listening and support during this journey.

I would like to thank Universidad Catolica del Norte and colleagues from the Department of Industrial Engineering in Antofagasta, Chile. Their trust in my abilities as a recent undergraduate fueled my confidence and ambition.

Finally, I am grateful to Auburn University and the many fellow students and professors who crossed paths with me during these four years. The unique spirit and vibrant atmosphere of

this institution have greatly enriched my experience. The memorable moments I was privileged to be a part of were sources of enjoyment and inspiration.

War Eagle!

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In electric power systems, the decision-making process involves several challenges mainly associated with uncertain factors present during the planning horizons. In particular, decisions must be made in advance for different planning horizons to provide households with electric energy. These decisions can be categorized into four specific stages: long-term, medium-term, short-term, and real-time [1]. During the long-term stage, critical investment and facility sizing decisions are made. This process involves planning and determining the size and capacity of power plants, photovoltaic facilities [2], wind farms [3], and battery storage systems [4]. These strategic decisions lay the foundation of a resilient power grid that satisfies future energy demands. Transitioning into a medium-term horizon, decisions are related to system maintenance, ensuring optimal functionality and longevity of the infrastructure [5, 6]. Lastly, the short-term and real-time horizons involve a dynamic decision-making process where various agents interact to secure the right to generate energy, ensuring a steady electricity supply to end consumers [7]. The operation of the power system takes place during these two planning horizons.

1.1 Energy market clearing

In power system operations, the Independent System Operator (ISO) must ensure the operability and reliability of the system. This is done by planning the energy generation schedule that specifies which energy generators owned by competitive Generation Companies (GENCOs) will operate to ensure a reliable and economical power supply. This schedule is generated

during the day-ahead market in a process known as market clearing [8]. Before this process materializes, GENCOs submit daily bids to the wholesale market, competing to acquire the right to dispatch energy according to the mandate of the ISO. Thus, the ISO clears the market, ensuring the correct interplay between energy supply and demand. As a result, by clearing the market, the ISO provides the day-ahead energy market price and the energy generator schedule.

In particular, the market clearing process is done by solving the Unit Commitment Problem (UCP), a mathematical optimization problem designed to establish an energy dispatch schedule that guarantees system operability and reliability at a minimal cost. The UCP considers several technical requirements, including generator unit capacity, mandatory minimum durations for units to remain online and offline, and energy production between periods. This process helps to maintain a steady, cost-effective supply of electricity while adhering to stringent operational and reliability requirements [9].

1.2 Challenges associated to using RES

During the last century, energy production has been generated by fossil-fuel-based generators that require resources such as coal and petroleum. Nevertheless, incentives for using other sources arose since they significantly impact the environment and depend on the worldwide economic context. Thus, producing energy from natural sources such as sun, wind, and biomass started to be attractive for different energy actors due to technological advances during the last two decades [10, 11]. In addition, several policies have been released by governments and world organizations promoting the development of renewable energy projects. For example, in the US, the *Inflation Reduction Act* promotes incentives to reduce renewable energy costs for several organizations. In Europe, the *Green Deal Industrial Plan* accelerates Europe's net-zero transition by supporting the scaling up of manufacturing capacity for climate-friendly technologies. Therefore, it is expected to observe more projects to reach the green transition.

Despite the benefits of energy from natural sources such as sun, wind, and river valleys, these natural sources can generate energy if the right environmental and local conditions allow it. Nevertheless, the natural conditions are not known in advance and can vary during daily operations. Hence, several models have been proposed to deal with uncertainty. In particular,

stochastic modeling is a technique that allows accounting for the inherent uncertainty of certain phenomena in a model. In the context of optimization, by including uncertainty, the decision-maker uses estimations that are accounted as parameters in the optimization model. Thus, different performance metrics, such as expected value, worst-case scenario, and risk, can be optimized using stochastic programming, robust optimization, and chance-constrained methods [12].

In particular, the growth of RESs projects has brought new challenges related to high uncertainty on renewable sources. One example is the duck curve effect observed in power systems with high solar photovoltaic generation. This effect is characterized by a significant drop in electricity demand during midday when solar generation is at its peak, followed by a steep increase in demand during the evening [13, 14]. Similarly, the wind ramping effect in power systems with significant wind energy production leads to sudden and substantial changes in wind power output [15]. Thus, new models and methodologies are needed to overcome uncertainty issues due to the expected increase of variability in energy power generation.

1.3 Uncertainty in the UCP

The energy generation schedule is currently built by solving the UCP, a nonlinear mixed integer programming problem. When uncertainty is included in this problem, stochastic programming, robust optimization, or chance-constrained methods are used to model the uncertainty in the mathematical model, creating the Stochastic Unit Commitment Problem (SUCP).

When modeling the SUCP using the stochastic programming approach, the mathematical formulation aims to minimize the expected commitment and dispatch cost using a two-stage approach. In the first stage, the commitment decisions are made, remaining unchanged during the planning horizon. The cost related to this first stage is derived based on the *hot* and *cold* state of the generator units. In addition, there is a cost for having the generator units on while dispatching minimum capacity. This cost is known in advance based on the commitment decisions.

The second-stage decision corresponds to the energy dispatch process. Since the commitment schedule is already known, the committed units dispatch the necessary energy after

observing the random variables. In particular, the dispatch process is determined by solving an economic dispatch process. Since the realizations of the random variables are not known when the commitment decisions are being made, the stochastic programming approach relies on generating a sample of scenarios usually generated from a probability distribution or data. Even though this method can provide solutions, the quality of them as well as the complexity, relies on the number of scenarios.

In the literature, the uncertainty is represented using scenario generation approaches when using stochastic programming. Several scenarios are sampled from the probability distribution of energy demand and renewable energy production. Then, the expected cost is calculated as the weighted sum between the cost of the realization and its respective probability. As the levels of uncertainty increase due to the inclusion of more RESs projects, the energy generation from these sources becomes more unpredictable. Thus, to obtain a suitable approximation of the expected cost, more scenarios should be considered, increasing the computational complexity of the problem.

1.4 Contributions

This thesis contributes by providing a new mathematical model for the SUCP. This new model does not require scenarios for modeling the expected cost. Instead, the expected cost is derived directly from the probability distribution of the residual demand.

In Chapter 2, a new stochastic approach is proposed for modeling the SUCP. The proposed model of this chapter addresses the uncertainty by deriving an analytical expression of the expected cost when the demand and RESs generation are uncertain. Then, the expected cost is approximated through a piece-wise linear function. Hence, to the best of the author's knowledge, it is the first time the SUCP is addressed through this approach.

In Chapter 3, the statistical model is extended to account for ramping constraints. These constraints ensure that there is a smooth transition to increase/decrease the production of energy from fuel-based generators. In particular, the ramping constraints are known by increasing the complexity of both UCP and SUCP. In addition, generating schedules without ramping capabilities can yield infeasible schedules in practice. Thus, this chapter addresses the shortcomings

of the first model. Since adding ramping constraints increases the complexity of the problem, solving strategies are proposed. These strategies are characterized by ensuring a better bound of the relaxation problem. In addition, a heuristic approach is proposed to generate a feasible solution that can be fed to the solver as a warm-up strategy. Finally, this model is tested on well-known power systems usually used in the literature for benchmarking.

In Chapter 4, a comparison is made between the statistical SUCP and the scenario-based SUCP. This comparison assesses the computational complexity of both models in terms of elapsed time to solve the model and the ability to reach optimality. In addition, we compare the out-of-sample and in-sample stability of both stochastic models. The comparison is made by considering power systems of different sizes, breakpoints, and scenarios for the statistical and scenario-based SUCP. Thus, by determining the benefits of each model, a new tool can be offered to ISOs when planning the day-ahead energy generation schedule.

This dissertation is structured as follows: Chapter 2 presents a new mathematical model for the SUCP. In Chapter 3, we extend the previous model by adding ramping constraints and proposing new solving strategies. In Chapter 4, we compare both the new model with the current scenario-based SUCP. Finally, conclusions and future work are presented in Chapter 5.

Chapter 2

Stochastic Unit Commitment Problem: A Statistical Approach

2.1 Introduction

The Unit Commitment Problem (UCP) is a mathematical optimization problem extensively studied in the power systems and operations research literature. The UCP aims to provide an energy generation schedule at a minimum cost while meeting technical and environmental constraints. The problem has played a relevant role in the liberalization of the electricity market, reducing the cost of generating electric energy [16]. However, solving the UCP is challenging. It is an NP-hard problem that must be solved within a tight time frame, considering the nonlinearity of the cost function and the utilization of binary variables [17]. In addition, real instances of this problem involve hundreds to thousands of generator units [18, 19].

Due to the inherent complexity of the UCP, various mathematical formulations and algorithms have been proposed to solve the UCP within the time constraints of ISOs [20, 21, 22]. Thus, it is possible to extend this problem towards more complicated variations. In particular, the UCP is a stochastic problem in nature due to the random energy demand and renewable energy production. One of the first studies about the stochastic unit commitment problems (SUCP) was proposed in [23] and [24]. In both studies, the authors modeled the problem using stochastic programming and the Monte Carlo simulation method. Under the stochastic programming appraoch, the SUCP is formulated as a two-stage or multistage optimization problem. Under a two-stage model, the commitment decisions are made in the first stage before the random variables are observed. Thus, these decisions remain unchanged over the entire planning horizon. The energy dispatch decisions are made in a second stage in response to the observed values of the random variables. In the case of multistage models, the commitment

decisions are made during each period before the random variables are revealed [25]. The sources of uncertainty in the SUCP are generally the demand, unit failure, and availability of renewable energy. These sources of uncertainty have been represented by generating several scenarios, which may increase the computational complexity of the solution algorithms [26]. The most common approaches for solving the SUCP are Lagrangian Relaxation [27], Column Generation [28, 29], Benders Decomposition [30, 31], and Progressive Hedging [32, 33]. To date, it has been modeled using other stochastic methods, namely chance-constrained [34, 35], and robust optimization [36, 29, 37].

In recent years, integrating Renewable Energy Sources (RESs) in the UCP has become necessary due to the high penetration of these resources in the generation mix. Considering the variability of wind speed, the stochastic modeling of wind energy has been one of the main interests of researchers [38]. In [39], the authors studied the impact of wind power forecasting on the UCP by computing the difference between the estimated cost of the day-ahead Unit Commitment and the actual cost from the Economic Dispatch once the realization of the wind speed was observed. In order to show the benefits of using the SUCP, studies to compare the deterministic UCP and the SUCP have been done [40, 41]. The results have consistently indicated that adopting the SUCP framework can lead to cost savings by reducing the required spinning reserves.

Despite recent improvements in algorithms, formulations, and computer capabilities, the industry has been reluctant to implement stochastic models in their daily operations [32, 42]. An attempt to encourage the industry to implement the SUCP was made in [43]. The authors solved the SUCP considering wind power as the main source of uncertainty. It was found that by implementing the stochastic model the pre-set spinning reserve can be reduced as well as the operating costs. In [33], the authors used a progressive hedging algorithm to speed up the solving process. Thus, the results showed that the proposed algorithm could solve actual instances of the SUCP in less than 25 minutes.

Solar power has also been included in the UCP. In [44], the authors proposed a framework involving a microgrid that produces solar energy generation and a main grid. Under this framework, the main grid imports energy from the microgrid in case of contingency. The authors

used the Benders decomposition approach to solve the problem where the master problem is a traditional UC problem, and the subproblem is a stochastic optimal power flow. In [35], the authors considered wind and PV power generation. The study aimed to reduce the RES forecasting error by controlling the risk of intermittent energy generation. Thus, a chance constraint method was used to ensure that the risk of intermittency was lower than a threshold value.

Another stream of studies has used the advantages of parallel computing in the design of algorithms. In [45], the authors solved a transmission-constrained UCP using a two-stage stochastic optimization formulation considering network failures and wind energy. To solve this problem, a parallel sub-gradient-based algorithm based on the Lagrangian relaxation method was used. In [46], the authors solved a security-constrained SUCP by taking advantage of the structure of the problem and decomposing it into small sub-problems using the Lagrangian relaxation method. The authors highlighted the practical feasibility of implementing the SUCP by generating only a small set of scenarios. Furthermore, in [26], the authors adopted a different approach for solving the security-constrained SUCP by decomposing the problem into three modules: UC, optimal power flow, and a bridge module that enabled communication between the other modules.

Scenario generation has been commonly used to represent uncertainty when solving small and large instances of the SUCP. However, the methodology and the number of scenarios must be carefully assessed to avoid bias in the commitment schedule and prevent increasing the computational burden of the solution algorithms. This Chapter proposes a novel modeling approach to solve the two-stage SUCP that does not require generating scenarios. Using the probability distribution of the residual demand, we develop an analytical function of the expected dispatch cost. To handle the nonlinearity of the function, we use a piecewise linear approximation approach, resulting in a Mixed Integer Linear Programming (MILP) problem. We validate our model by evaluating the optimal unit commitment schedule on the analytical function and conducting simulations. In addition, a stability analysis is performed to determine the number of breakpoints. The Chapter remains as follows: Section 2.2 presents the mathematical model and the analytical expected dispatch cost. Section 2.3 describes methods to solve the problem and

the results of the experiments. Finally, we describe the main findings of this Chapter and future research in section 2.4.

2.2 Formulation

In this section, we provide the mathematical formulation of the SUCP. The model is formulated using the technical constraints provided in [47], adapting the expected cost given the uncertain residual demand.

Sets and indices

I	Set of units, $i \in I$
T	Set of periods, $t \in T$
L	Set of breakpoints, $l \in L$

Parameters

C_i^{hot}	Hot start-up cost of unit $i \in I$
C_i^{cold}	Cold start-up cost of unit $i \in I$
c_i	Energy cost of unit $i \in I$
$F_i(\cdot)$	Dispatch cost function for unit $i \in I$ in period $t \in T$
$H_{it}(\cdot)$	Start-up cost function of unit $i \in I$ during period $t \in T$
P_i^{max}	Maximum energy production of unit $i \in I$
P_i^{min}	Minimum energy production of unit $i \in I$
T_i^{on}	Minimum number of periods of unit $i \in I$ has to be on
T_i^{off}	Minimum number of periods of unit $i \in I$ has to keep off
t_i^{cold}	Number of periods after unit $i \in I$ becomes cold
u_i^{prev}	Binary parameter that indicates the on/off state of unit $i \in I$ before the first period of the planning horizon
Δ_i	Difference between the maximum and minimum energy capacity P_i^{max} and P_i^{min} of unit $i \in I$
κ	Cost of buying energy from other energy markets.
τ_i^{on}	Number of periods unit $i \in I$ has been on prior to the first period of the planning horizon

τ_i^{off} Number of periods unit $i \in I$ has been off prior to the first period of the planning horizon

Random variables

$\tilde{d}_t$ Random demand during period $t \in T$

$\tilde{r}_t$ Random residual demand during period $t \in T$

$\tilde{\omega}_t$ Random production of energy during period $t \in T$

$\tilde{\psi}_t$ Random residual after first-stage decisions during period $t \in T$

Decision variables

u_{it} Binary variable that indicates if the unit $i \in I$ is on/off during period $t \in T$

v_{it} Binary variable that indicates if the unit $i \in I$ is turned on during period $t \in T$

w_{it} Binary variable that indicates if the unit $i \in I$ is turned off during period $t \in T$

y_{it}^{hot} Binary variable that indicates if the unit $i \in I$ starts hot during period $t \in T$

y_{it}^{cold} Binary variable that indicates if the unit $i \in I$ starts cold during period $t \in T$

λ_{itl} Continuous variable that weighs the breakpoint $l \in L$ of unit $i \in I$ during period $t \in T$

η_{itl} Binary variable that indicates if the interval $(l-1, l) \in L$ is selected for the unit $i \in I$ during period $t \in T$

$$\min z = \sum_{t\in T}\sum_{i\in I} H_{it}(\cdot) + \sum_{t\in T}\sum_{i\in I} \mathbf{E}\left[F(\tilde{r}_t)\right] \tag{2.1}$$

$$\text{s.t} \quad \sum_{j=\gamma_{it}^{on}}^{t} v_{ij} \leq u_{it} \quad \forall i \in I, t \in T \tag{2.2}$$

$$\sum_{j=\gamma_{it}^{off}}^{t} w_{ij} \leq 1 - u_{it} \quad \forall i \in I, t \in T \tag{2.3}$$

$$u_{it} = 1 \quad \forall i \in I : u_i^{prev} = 1 \quad \forall t \in \{1, ..., \theta_i^{on}\} \tag{2.4}$$

$$u_{it} = 0 \quad \forall i \in I : u_i^{prev} = 0 \quad \forall t \in \{1, ..., \theta_i^{off}\} \tag{2.5}$$

$$y_{it}^{hot} + y_{it}^{cold} = v_{it} \quad \forall i \in I, t \in T \tag{2.6}$$

$$u_{it} - \sum_{l=t-t_i^{cold}-1}^{t-1} u_{il} \leq y_{it}^{cold} \qquad \forall i \in I, t \in T \tag{2.7}$$

$$u_{it} - u_{it-1} \leq v_{it} \quad \forall i \in I, t \in T \tag{2.8}$$

$$w_{it} = v_{it} + u_{it-1} - u_{it} \quad \forall i \in I, t \in T \tag{2.9}$$

$$u_{it}, v_{it}, w_{it}, y_{it}^{hot}, y_{it}^{cold} \in \{0, 1\} \quad \forall i \in I, t \in T \tag{2.10}$$

The objective function (2.1) aims to minimize the commitment and the expected dispatch costs. The first component, representing the start-up cost, is given by $H_{it}(\cdot) = a_i^{hot} y_{it}^{hot} + a_i^{cold} y_{it}^{cold}$. The second component, denoted as the expected dispatch cost ($\mathbf{E}\left[F_i(\tilde{r}_t)\right]$), depends on the random residual demand during each period. A comprehensive derivation of this cost is presented in Subsection 2.2.1.

Constraints (2.2) and (2.3) ensure the minimum up and down time requirements, where $\gamma_{it}^{\text{on}} = \max\{t - T_i^{\text{on}} + 1, 1\}$ and $\gamma_{it}^{\text{off}} = \max\{t - T_i^{\text{off}} + 1, 1\}$. Constraints (2.4) and (2.5) determine the initial commitment state of the units based on their previous states, where $\theta_i^{\text{on}} = \max\{1, T_i^{\text{on}} - \tau_i^{\text{on}} + 1\}$ and $\theta_i^{\text{off}} = \max\{1, T_i^{\text{off}} - \tau_i^{\text{off}} + 1\}$. Constraint (2.6) ensures that a unit starts in either a hot or cold state. Additionally, constraint (2.7) states that a unit starts in a cold state if it has been off for a duration greater than t_i^{cold}. Constraint (2.8) guarantees that a unit is turned on in the current period if it was turned off in the previous period. Similarly, constraint

(2.9) ensures that a unit is turned off in the current period if it was on in the previous hour. Finally, constraint (2.10) states the integrity of the decision variables.

2.2.1 Modeling the expected dispatch cost

In a competitive electricity market, Generation Companies (GENCOS) submit to the market operator one block with the minimum energy production P_i^{min}, maximum energy production P_i^{max}, and the cost c_i for each unit $i \in I$. Based on the information submitted by the GENCOS, we assume that the market operator dispatches the units economically. Thus, the operator ensures that both the energy demand and the technical constraints are met.

In several markets, RESs operate as non-dispatchable energy. Thus, all the energy they produce is dispatched, acting as a negative demand with no production cost. As a result, the remaining demand, corresponding to the difference between the demand and the production of RESs, is supplied by the thermal units. Hence, the residual demand is given by equation (2.11).

$$\tilde{r}_t = \tilde{d}_t - \tilde{\omega}_t \quad \forall t \in T \tag{2.11}$$

To account for the unmet demand (demand not served by the RESs and thermal units), we create an auxiliary unit with an unlimited capacity. Thus, the ISO incurs a penalty cost of κ per-megawatt-hour (MWh) for the unmet demand, representing the energy purchase from other markets to compensate for the energy shortfall.

The formulation of the dispatch cost follows the same idea as the two-stage stochastic optimization, where the commitment decisions are made first and cannot be modified during the planning horizon. Thus, in equation (2.12), we define the residual demand after the first-stage decisions are made. This residual demand accounts for the difference between the total demand and the minimum capacity of the activated units.

$$\tilde{\psi}_t = \tilde{r}_t - \sum_{i \in I} P_i^{min} u_{it} \quad \forall t \in T \tag{2.12}$$

We present the dispatch cost in equation (2.13), which consists of two components: a fixed and a variable component. The former is the cost of dispatching the minimum capacity of the units that are operating. The latter is the cost of dispatching beyond the minimum capacity of the units, which depends on the random residual demand.

The remaining amount to be dispatched depends on the difference between the residual demand and the remaining energy of the previous $i-1$ cheapest units. If for any period $t \in T$ the difference $\tilde{\psi}_t - \sum_{j=1}^{i-1} \Delta_j u_{jt}$ is positive, the amount to be dispatched corresponds to the minimum between the residual demand and the remaining energy of unit $i \in I$. Otherwise, the energy dispatched by the i^{th} unit will be zero.

$$F(\tilde{r_t}) = \sum_{i \in I} c_i P_i^{min} u_{it} + \sum_{i \in I} c_i \min \left[\max(\tilde{\psi}_t - \sum_{j=1}^{i-1} \Delta_j u_{jt}, 0), \Delta_i u_{it} \right] + \kappa \cdot \max(\tilde{\psi}_t - \sum_{j=1}^{|I|} \Delta_j u_{jt}, 0) \quad \forall t \in T \tag{2.13}$$

Then, when applying expectation to the dispatch cost, equation (2.13) remains as follows:

$$\mathbf{E}\left[F(\tilde{r_t})\right] = \sum_{i \in I} c_i P_i^{min} u_{it} + \mathbf{E}\left[\sum_{i \in I} c_i \cdot \min \left[\max(\tilde{\psi}_t - \sum_{j=1}^{i-1} \Delta_j u_{jt}, 0), \Delta_i u_{it} \right] \right] + \mathbf{E}\left[\kappa \cdot \max(\tilde{\psi}_t - \sum_{j=1}^{|I|} \Delta_j u_{jt}, 0) \right] \forall t \in T \tag{2.14}$$

Let $I(x)$ be an indicator function that depends on a nonnegative random variable q, and a nonnegative value x.

$$I(x) = \begin{cases} 1 & x < q \\ 0 & x \geq q \end{cases} \tag{2.15}$$

By applying expectation to the indicator function (2.15), the expected value results in the survival function of the random variable q as shown in equation (2.16).

$$\mathbf{E}\left[I(x)\right] = 1 \cdot \int_x^{\infty} f(q)\, dq + 0 \cdot \int_{-\infty}^{x} f(q)\, dq = P(q \geq x) \tag{2.16}$$

Let the minimum function between two nonnegative numbers q and p be calculated as the following definite integral:

$$\min(q, p) = \int_0^q I(x)\, dx \tag{2.17}$$

Since the resulting values from the minimum and maximum function in equation (2.14) are nonnegative, the above property can be applied. Thus, equation (2.17) holds. Then, the expected dispatch cost is calculated as the integral of the indicator function:

$$\sum_{i \in I} c_i \cdot \int_0^{\Delta_i u_i} \mathbf{E}\left[I(x)\right]\, dx + \kappa \cdot \int_0^{\infty} \mathbf{E}\left[I(x)\right]\, dx \quad \forall t \in T \tag{2.18}$$

Thus, the expected value in (2.18) is calculated as in (2.16), resulting in the following expression:

$$\sum_{i \in I} c_i \int_0^{\Delta_i u_{it}} P(\tilde{\psi}_t - \sum_{j=1}^{i-1} \Delta_j u_{jt} \geq x)\, dx + \kappa \int_0^{\infty} P(\tilde{\psi}_t - \sum_{j=1}^{|I|} \Delta_j u_{jt} \geq x)\, dx \quad \forall t \in T \tag{2.19}$$

Let $S(x)$ be the survival function of the random variable represented by the difference between the demand and the production of RESs. Since $\tilde{\psi}_t = \tilde{r}_t - \sum_{i \in I} P_i^{min} u_{it} \quad \forall t \in T$, expression (2.19) is written as follows:

$$\sum_{i \in I} c_i \int_0^{\Delta_i u_{it}} S_t(x + \sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i-1} \Delta_j u_{jt})\, dx + \kappa \int_0^{\infty} S_t(x + \sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{|I|} \Delta_j u_{jt})\, dx \quad \forall t \in T \tag{2.20}$$

Thus, expression (2.20) is the integral of the survival function shifted for each period. Therefore, it can be centered at the origin as shown in expression (2.21).

$$\sum_{i \in I} c_i \int_{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i-1} \Delta_j u_{jt}}^{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i} \Delta_j u_{jt}} S_t(x)\, dx + \kappa \int_{+\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{|I|} \Delta_j u_{jt}}^{\infty} S_t(x)\, dx \quad \forall t \in T \tag{2.21}$$

Let $\Gamma(x)$ be the indefinite integral of the survival function $S(x)$.

$$\Gamma(x) = \int S(x)\, dx \tag{2.22}$$

Then, expression (2.21) states as follows:

$$\sum_{i \in I} c_i \left[\Gamma_t(\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i} \Delta_j u_{jt}) - \Gamma_t(\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i-1} \Delta_j u_{jt}) \right] + \kappa \left[\mu_t - \Gamma_t(\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{|I|} \Delta_j u_{jt}) \right] \forall t \in T \tag{2.23}$$

Thus, the expected dispatch cost is a nonlinear function that must be transformed into a linear expression.

2.2.2 A linear approximation approach

We employ an incremental piecewise linear approximation model to linearize the expected cost in expression (2.23). This model is based on the method proposed in [48], which introduces a binary variable for each linear segment. Let L be the set of breakpoints with $l - 1$ intervals. The function $\Gamma(x)$ can be approximated using l breakpoints. Let b_l be the breakpoint l^{th} of the function $\Gamma(x)$, where b_l is nonnegative. Then, any value x that belongs to the domain of the function $\Gamma(x)$ can be expressed in linear terms using nonnegative continuous scalars λ_l that sum one. Thus, the function $\Gamma(x)$ is calculated as follows:

$$x = \sum_{l \in L} \lambda_l b_l \tag{2.24}$$

$$\Gamma(x) = \sum_{l \in L} \lambda_l \Gamma(b_l) \tag{2.25}$$

In order to ensure that x has a unique representation, at most two consecutive scalars have to be nonzero. Then, x has one representation as $x = \lambda_l b_l + \lambda_{l+1} b_{l+1}$ if $\lambda_l + \lambda_{l+1} = 1$. Thus, a new variable is created. Let η_l be a binary variable that indicates which segment of the approximated function $\Gamma(x)$ is selected. Hence, the function $\Gamma(x)$ is expressed in linear terms and is included in the objective function. The resulting model is an MILP and is formulated as follows:

$$\min z = \sum_{t\in T}\sum_{i\in I} H_{it}(\cdot) + \sum_{t\in T}\sum_{i\in I} c_i P_i^{min} u_{it} +$$
$$\sum_{t\in T}\sum_{i\in I} c_i \left[\sum_{l\in L} \lambda_{itl}\Gamma_t(b_{tl}) - \sum_{l\in L} \lambda_{i-1tl}\Gamma_t(b_{tl})\right] + \tag{2.26}$$
$$\sum_{t\in T} \kappa \left[\mu_t - \sum_{l\in L} \lambda_{|I|tl}\Gamma_t(b_{tl})\right]$$

s.t $\quad (2.2) - (2.10)$ (2.27)

$$\sum_{l\in L} \lambda_{0tl} b_{tl} = \sum_{k\in I} P_k^{min} u_{kt} \qquad t \in T \tag{2.28}$$

$$\sum_{l\in L} \lambda_{itl} b_{tl} = \sum_{k\in I} P_k^{min} u_{kt} + \sum_{j=1}^{i} \Delta_j u_{jt} \quad \forall i \in I, t \in T \tag{2.29}$$

$$\sum_{l=1}^{L} \lambda_{itl} = 1 \qquad \forall i \in I, t \in T \tag{2.30}$$

$$\lambda_{it1} \leq \eta_{it1} \qquad \forall i \in I, t \in T \tag{2.31}$$

$$\lambda_{itl} \leq \eta_{itl-1} + \eta_{itl} \qquad \forall i \in I, t \in T, l \in \{2, \ldots, L\} \tag{2.32}$$

$$\lambda_{itL} \leq \eta_{itL-1} \qquad i \in I, t \in T \tag{2.33}$$

$$\sum_{l=1}^{L-1} \eta_{itl} = 1 \qquad \forall i \in I, t \in T \tag{2.34}$$

$$\lambda_{itl} \geq 0 \qquad \forall i \in I, t \in T, l \in L \tag{2.35}$$

$$\eta_{itl} \in \{0, 1\} \qquad \forall i \in I, t \in T, l \in L \tag{2.36}$$

The objective function (2.26) minimizes the start-up cost and the expected dispatch cost expressed in linear terms by using the piecewise linear approximation method. Furthermore, since the survival function has different parameters in each period, $|T|$ functions must be approximated in a piecewise manner. The set of constraints (2.27) corresponds to the technical constraints of the UCP described in Section 2.2. Constraint (2.28) and (2.29) set the limits of the integral (2.21) ensuring that the correct value b_{tl} is evaluated in the objective function. The remaining constraints define the rules for constructing a piecewise linear approximation. Equation (2.30) ensures that all λ_l sum up to one. Constraints (2.31), (2.32), and (2.33) indicate

which λ_l are nonzeros according to the selected segment. Constraint (2.34) ensures that only one segment of the piecewise function is selected. Constraints (2.35) and (2.36) correspond to the integrity constraints.

2.3 Results

The results of this section were obtained using Gurobi 9.5 (64-bit) on a computer running Linux with an Intel Xeon E5-2670v2 Ivy Bridge 2.5 GHz using 16 cores and 64 GB of RAM at the Alabama Supercomputer Center. When solving the optimization problems, we set the stopping criteria at 3,600 seconds after starting the optimization or when the optimizer has found an optimal solution. All methods and procedures were coded in Python 3.8.

We test the proposed model using a power system consisting of 20, 60, 100, and 200 units and a planning horizon of 24 hours. The power system data for 20 units is from [49], which is given in Table 2.1. The system consisting of 60, 100, and 200 units corresponds to the 20-units system replicated 3, 5, and 10 times. The cost of the unmet demand is assumed to be 100 $\frac{\$}{MWh}$. The last column of Table 2.1 is the state of the unit at the beginning of the planning horizon. It indicates the number of periods the unit has been on/off, represented by the plus/minus sign, respectively.

Unit	P^{max} **(MWh)**	P^{min} **(MWh)**	c **($/MWh)**	t_{up} **(hour)**	t_{down} **(hour)**	a^{hot} **($)**	a^{cold} **($)**	t^{cold}	**Initial state**
1	455	150	16.19	8	8	4,500	9,000	5	8
2	455	150	16.19	8	8	4,500	9,000	5	8
3	130	20	16.5	5	5	560	1,120	4	-5
4	130	20	16.5	5	5	560	1,120	4	-5
5	130	20	16.6	5	5	550	1,100	4	-5
6	130	20	16.6	5	5	550	1,100	4	5
7	455	150	17.26	8	8	5,000	10,000	5	8
8	455	150	17.26	8	8	5,000	10,000	5	8
9	162	25	19.7	6	6	900	1,800	4	-6
10	162	25	19.7	6	6	900	1,800	4	-6
11	80	20	22.26	3	3	170	340	2	-3
12	80	20	22.26	3	3	170	340	2	-3
13	55	10	25.92	1	1	30	60	0	-1
14	55	10	25.92	1	1	30	60	0	-1
15	55	10	27.27	1	1	30	60	0	-1
16	55	10	27.27	1	1	30	60	0	-1
17	85	25	27.74	3	3	260	520	2	-3
18	85	25	27.74	3	3	260	520	2	-3
19	55	10	27.79	1	1	30	60	0	-1
20	55	10	27.79	1	1	30	60	0	-1

Table 2.1: Power system description

2.3.1 Modeling the demand and RESs production

In power system operation, the hourly stochastic demand is usually modeled using the forecasted value plus an error term which generally fits a Gaussian distribution function with expected value 0 and standard deviation σ [50]. Thus, the actual demand corresponds to the forecasted demand "perturbed" by a random error which probability distribution is the normal distribution function [51, 52], truncated normal distribution function [26, 53], or multivariate normal distribution function [54].

The production of renewable energy depends on the weather and the technology of the renewable source. Thus, probabilistic forecasts are methods developed to estimate the wind and solar power production. When solving the SUCP using scenario generation, the realizations of the random variables are usually sampled (using the Montecarlo simulation method) from a known distribution function such as normal [26], truncated normal [53], multivariate normal [55], uniform [36, 52], beta [56], and Weibull distribution [52].

In our research, we consider a system that includes energy from wind farms. Thus, we model the stochastic residual demand vector $\tilde{\mathbf{r}}$ as shown in equation (2.37). The vector value $\boldsymbol{\mu}$ is the forecasted residual demand, i.e., the difference between the expected demand and the expected wind energy production. The error term ϵ, is assumed to follow a multivariate normal distribution with mean vector 0 and covariance matrix Σ.

$$\tilde{\boldsymbol{r}} = \boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{2.37}$$

Notice that our proposed model does not restrict what forecasting model for the residual demand is used. Therefore, it is assumed a forecast of the residual demand and estimates of the parameters of the error term probability distribution are known. Moreover, the RES production corresponds to solar or wind energy because the forecasted irradiation or wind speed is estimated. The forecasted residual demand used in our problem instances is given in Table 2.2 and is calculated based on the demand and wind power profile from [49] and [54], respectively. We compute the covariance matrix Σ using the standard deviation values in Table 2.2 assuming an Autoregressive model AR(1). Equation (2.38) shows the corresponding covariance matrix

with $|T|$ periods as the planning horizon. For the instances with 60, 100, and 200 units, the mean and standard deviation are amplified by 3, 5, and 10 times, respectively.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \cdots & \sigma_1\sigma_T\rho^{T-1} \\ \sigma_2\sigma_1\rho & \sigma_1^2 & \cdots & \sigma_2\sigma_T\rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_T\sigma_1\rho^{T-1} & \sigma_T\sigma_2\rho^{T-2} & \cdots & \sigma_T^2 \end{pmatrix} \tag{2.38}$$

Hour	Mean (MWh)	Std. Dev. (MWh)	Hour	Mean (MWh)	Std. Dev. (MWh)
1	1,214	45.93	13	2,680	84.85
2	1,286	49.83	14	2,370	81.32
3	1,500	54.78	15	2,264	73.27
4	1,700	60.41	16	1,960	64.54
5	1,766	64.40	17	1,766	64.40
6	1,994	69.14	18	1,930	71.31
7	2,084	72.30	19	2,180	75.29
8	2,240	73.76	20	2,558	87.42
9	2,480	78.92	21	2,354	81.79
10	2,686	84.77	22	1,980	69.57
11	2,744	88.39	23	1,624	56.80
12	2,856	91.14	24	1,506	48.91

Table 2.2: Mean and standard deviation of the residual demand

2.3.2 Computing the piecewise linear function and stability analysis

We compute the breakpoints of the functions $\Gamma(x)$ giving in equation (2.22) using the optimization library PWLF [57] in Python 3.8. The optimization library computes the breakpoints by minimizing the sum-of-square of the residuals using a multi-start gradient optimization method.

In Figure 2.1, we show an example where the function $\Gamma(x)$ corresponding to hour 5 ($\mu_5 = 1,766$ and $\sigma_5 = 64.40$) is approximated by a piecewise linear function using 10 breakpoints. Notice that the multi-start gradient optimization method provides a good fit for the function $\Gamma(x)$.

Since the multi-start gradient optimization method is a local search optimizer, it provides different breakpoints at each run. Therefore, we perform a stability analysis to assess the consistency of this method. Adding breakpoints could increase the computational complexity because of the additional binary variables and constraints. However, more breakpoints could improve

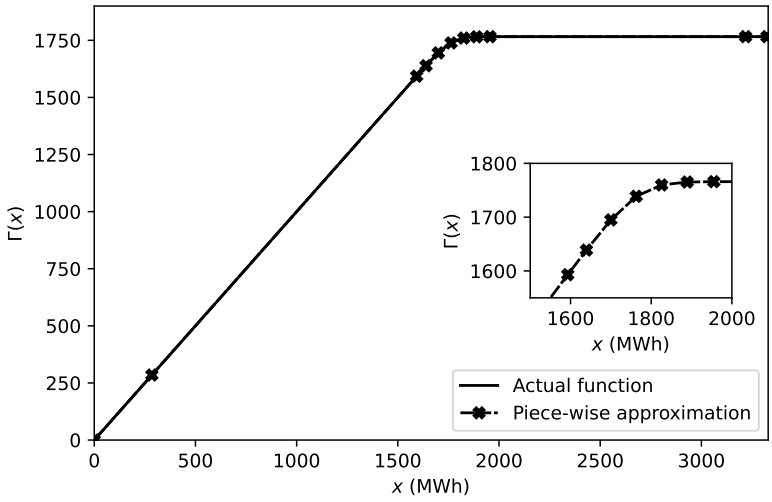


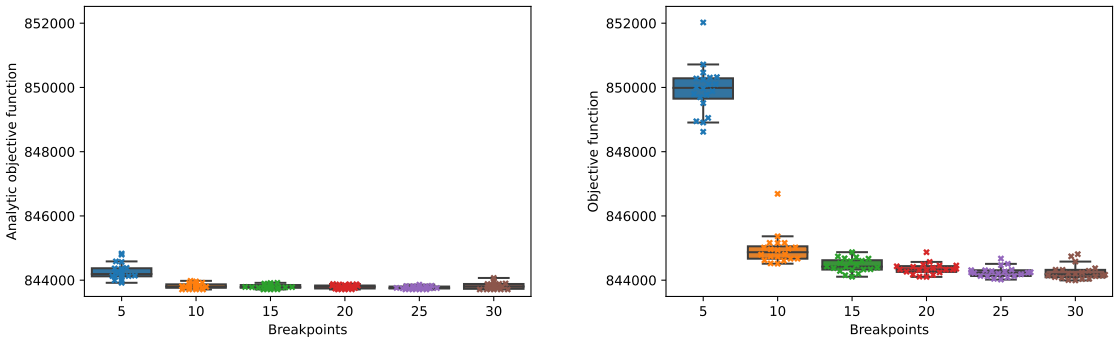
Figure 2.1: Piecewise approximation of function $\Gamma_5(x)$ using 10 breakpoints.

the quality of the objective value. Nevertheless, the question is how many points are needed to obtain a good solution in a reasonable computational time. Thus, we study the in-sample and out-of-sample stability of the breakpoints generation method.

In [58], the authors proposed a method to assess the suitability of a scenario generation method when solving a stochastic programming model. Although our approach does not generate scenarios, it needs to generate a set of breakpoints. Suppose we generate K sets of n breakpoints and solve the optimization problem for each set. If, upon evaluating the optimal commitment schedule using expression (2.23), the results yield nearly identical values across the K sets, the breakpoint generation method is out-of-sample stable. Similarly, the generation method is in-sample stable if the optimal objective value (2.26) is approximately consistent across the K sets.

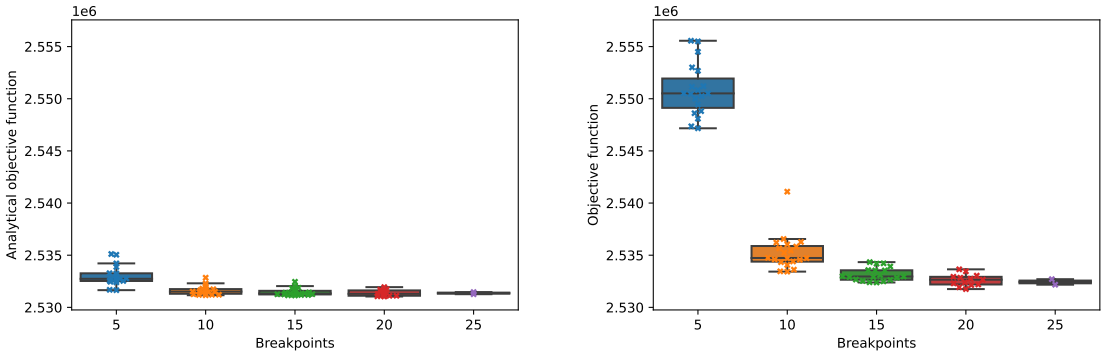
We perform 20 experiments for each combination of units and breakpoints, resulting in 480 experiments. We provide the descriptive statistics of the elapsed computational time and the MIP gap in Tables 2.3 and 2.4, respectively. Figures 2.2, 2.3, 2.4, and 2.5 show the results of the stability analysis for 20, 60, 100, and 200 units, respectively. We observe that when increasing the number of breakpoints, the out-of-sample stability does not considerably change after incrementing from 5 to 10 breakpoints. Thus, solving the problem using more breakpoints does not significantly increase the accuracy. In addition, the in-sample stability increases when

increasing the number of breakpoints, especially when using between 5 to 10 breakpoints. Some instances of 60, 100, and 200 units with 15 to 30 breakpoints did not find an optimal solution. Nevertheless, as shown in Figures 2.2a, 2.3a, 2.4a, and 2.5a, the difference in the results when using between 5 to 15 breakpoints is very small. From the results, we conclude that the proposed solving method has out-of-sample and in-sample stability for five or more breakpoints.



(a) Out-of-sample stability (b) In-sample stability

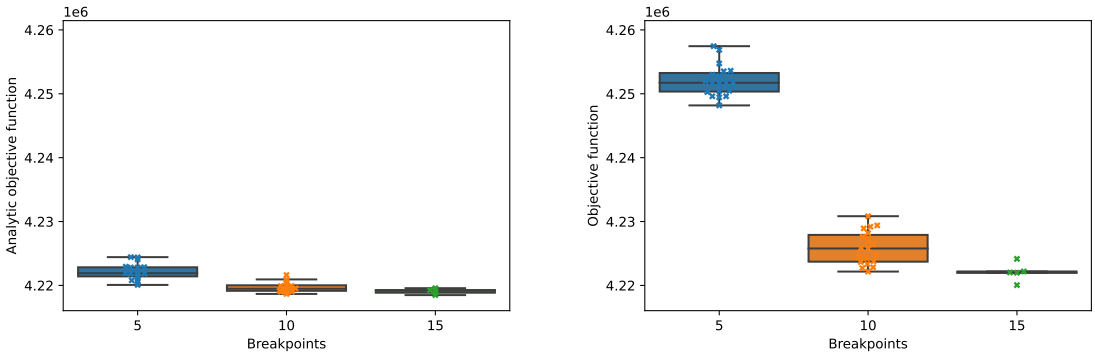
Figure 2.2: Stability analysis for 20 units



(a) Out-of-sample stability (b) In-sample stability

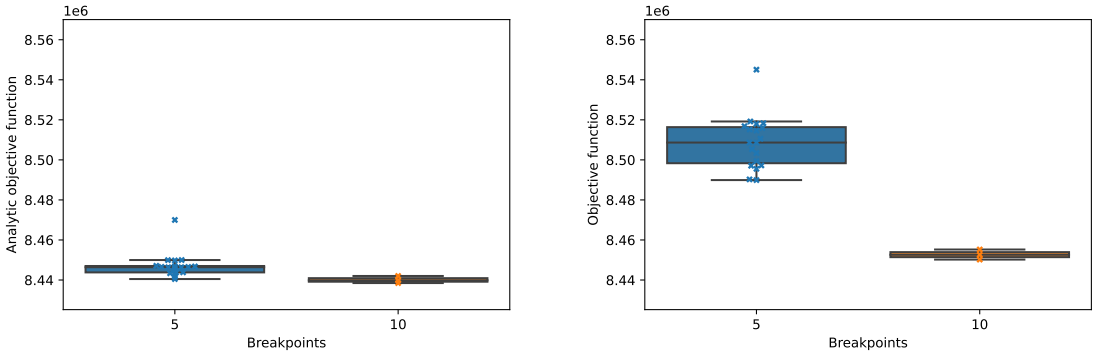
Figure 2.3: Stability analysis for 60 units

In Table 2.3, we show the statistics of the elapsed time, including the time of the PWLF library to approximate the function $\Gamma(x)$, and the time the optimizer reaches either the optimum or the time limit of 3,600 seconds. The SUCP was solved in a few minutes in all test instances with five breakpoints. When increasing the number of breakpoints from 5 to 10, only the instances with 20 and 60 units found an optimal solution in less than 5 minutes. Nevertheless, the



(a) Out-of-sample stability (b) In-sample stability

Figure 2.4: Stability analysis for 100 units



(a) Out-of-sample stability (b) In-sample stability

Figure 2.5: Stability analysis for 200 units

instance with 100 units was solved in less than 45 minutes for more than 75% of the instances with 10 breakpoints.

In Table 2.4, we show the descriptive statistics of the MIP gap of the experiments. Notice that instances that did not found an optimal solution report the optimality gap within 3,600 seconds. Notice that the proposed method scales well to 200 units, considered a medium-sized system in the literature. The instance with 100 units found an optimal solution with 15 breakpoints in less than 50% of the cases. A similar result was observed with 200 units and 10 breakpoints. It is important to observe from Figures 2.2 - 2.5 and Table 2.4 that our approach provides a near-optimal solution to the SUCP problem with just five breakpoints.

Units	Breakpoints	Average (s)	Minimum (s)	Median (s)	3rd Quartile (s)	Maximum (s)
20	5	8.32	7.62	8.22	8.65	9.30
	10	35.92	30.49	36.94	38.09	40.79
	15	74.51	56.85	75.79	81.20	86.76
	20	123.69	102.33	122.93	129.62	158.41
	25	305.12	160.49	184.33	300.03	888.65
	30	663.06	254.37	712.80	828.00	1,207.90
60	5	24.17	18.57	20.80	27.55	48.39
	10	130.07	69.62	97.32	115.79	426.15
	15	926.36	153.02	220.21	1339.29	3671.73
	20	2,172.74	213.38	2,598.62	3,712.77	3,752.21
	25	3,634.56	1,350.44	3,764.37	3,775.88	3,816.62
	30	3,831.10	3,790.68	3,828.43	3,847.84	3,888.63
100	5	53.34	38.18	54.29	63.15	74.61
	10	1,430.94	135.86	965.25	2,640.95	3,657.27
	15	2,863.74	319.44	3,689.85	3,696.93	3,701.34
	20	3,728.40	3,704.37	3,725.75	3,740.07	3,774.00
	25	3,770.43	3,734.81	3,771.21	3,778.92	3,803.47
	30	3,796.25	3,769.42	3,792.36	3,800.25	3,840.54
200	5	215.44	132.95	219.58	250.29	283.50
	10	3,518.58	1,818.96	3,683.96	3,685.52	3,691.16
	15	3,716.02	3,694.19	3,707.65	3,714.74	3,839.07
	20	3,770.31	3,731.12	3,765.89	3,776.17	3,851.13
	25	3,799.16	3,766.23	3,798.11	3,810.99	3,829.53
	30	3,851.92	3,815.15	3,851.91	3,867.84	3,889.85

Table 2.3: Elapsed time descriptive statistics

2.3.3 Analyzing the effect of the forecasting error hourly correlation

We study the effect of the forecasting error correlation on the expected cost using the instance of 100 units and 10 breakpoints. We considered three levels of correlation, namely 0.0, 0.5, and 0.8, assumming that the model is solved at hour 0. Thus, the error ϵ_0 at that hour is known. Before solving the SUCP, we compute the conditional forecasting error at each hour given ϵ_0. We assume seven different values for ϵ_0, namely 0, $\pm 3\sigma_{24}$, $\pm 6\sigma_{24}$, and $\pm 9\sigma_{24}$. The three values of the forecasting error correlation and the seven values of ϵ_0 creates fifteen different SUCP cases. In Table 2.5, we show the expected cost obtained using the solution of the optimization method and evaluating the optimal solution using the analytic function (2.23). From Table 2.5, we observe that the difference between expected costs is small for all values of correlation, confirming that our approach is suitable for solving the SUCP.

During the commitment scheduling process, the system operator must consider energy reserves as a backup for generation outages or energy shortages. In deterministic models, the reserves are set as a fixed percentage of the forecasted demand. We calculate the system reserve as a percentage of the demand for different values of the error correlation. In Figure 2.6, we show the system reserve over the planning horizon. Notice that the reserve starts increasing during hours 15 to 19, due to the minimum up-time of the units after the period of high demand

Units	Breakpoints	Average (%)	Minimum (%)	Median (%)	3rd Quartile (%)	Maximum (%)
20	5	0	0	0	0	0
	10	0	0	0	0	0
	15	0	0	0	0	0
	20	0	0	0	0	0
	25	0	0	0	0	0
	30	0	0	0	0	0
60	5	0	0	0	0	0
	10	0	0	0	0	0
	15	3	0	0	0	60
	20	23	0	0	74	78
	25	68	0	76	78	78
	30	74	0	78	79	80
100	5	0	0	0	0	0
	10	3	0	0	0	66
	15	57	0	76	77	79
	20	77	74	78	79	79
	25	79	75	79	80	80
	30	80	76	80	80	81
200	5	0	0	0	0	0
	10	64	0	76	77	78
	15	77	66	78	79	80
	20	80	77	80	80	82
	25	81	78	81	81	82
	30	82	81	82	82	82

Table 2.4: MIP gap descriptive statistics

		Expected cost ($)	
ρ	ϵ_0	**Approximation**	**Analytic**
0.8	-9σ_{24}	4,033,540	4,028,857
0.8	-6σ_{24}	4,097,542	4,093,358
0.8	-3σ_{24}	4,161,113	4,156,461
0.5	-9σ_{24}	4,187,310	4,182,383
0.5	-6σ_{24}	4,201,660	4,194,226
0.5	-3σ_{24}	4,212,360	4,206,177
0.8	0σ_{24}	4,225,727	4,218,857
0	0σ_{24}	4,223,433	4,219,210
0.5	0σ_{24}	4,227,916	4,220,395
0.5	3σ_{24}	4,238,545	4,231,944
0.5	6σ_{24}	4,252,855	4,246,638
0.5	9σ_{24}	4,261,407	4,256,355
0.8	3σ_{24}	4,289,624	4,284,149
0.8	6σ_{24}	4,351,254	4,348,062
0.8	9σ_{24}	4,421,781	4,415,020

Table 2.5: Effect of forecasting error and hourly correlation

(hours 9 to 14). Thus, in periods of high demand, the system uses most of the available capacity, resulting in a lower reserve. After the second period of high demand, the optimal solution turns off units to decrease the dispatching cost.

2.3.4 Analyzing the effect of increasing wind energy production

We assess the impact of both increasing the wind energy production and the unmet demand cost. We increment the wind energy production of the base case in increments of 10% up to 50% and the unmet demand cost from 100 to 200, 500, and 1,000 $\frac{\$}{MWh}$. Table 2.6 shows the expected cost at different levels of the wind multiplier and unmet demand cost. The results show that

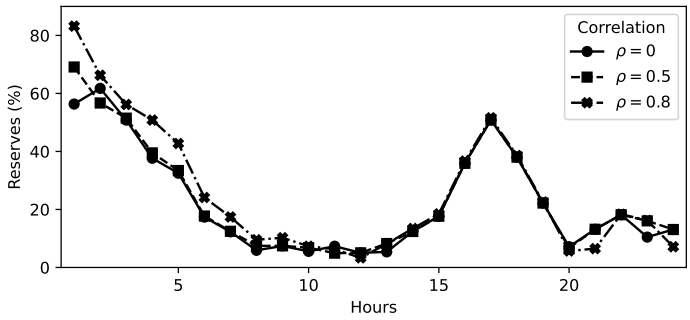


Figure 2.6: Reserves as a percentage of the forecasted residual demand when $\epsilon_0 = -6\sigma_{24}$

increasing the wind energy production by 50% can reduce up to 5% of the expected cost. Moreover, when increasing the unmet demand cost, the effect on the expected cost is negligible, increasing just 1%. Notice that increasing the production of wind energy also increases the uncertainty from this source.

		Increase percentage of wind energy					
		0%	10%	20%	30%	40%	50%
Unmet demand cost ($)	100	4,219,150	4,177,789	4,137,185	4,096,726	4,056,402	4,015,829
	200	4,221,612	4,181,852	4,139,407	4,099,121	4,059,085	4,018,089
	500	4,225,328	4,190,016	4,144,372	4,102,544	4,060,368	4,036,443
	1,000	4,246,800	4,214,777	4,152,279	4,111,149	4,068,103	4,065,353

Table 2.6: Expected cost when increasing wind energy production

2.3.5 Assessing the Loss of Load Probability

The Loss of Load Probability (LOLP) is an important metric to assess the reliability of the system [42]. We study the effect of increasing the cost of unmet demand on the LOLP by considering four values, namely, 100, 200, 500, and 1,000 $\frac{\$}{MWh}$. In addition, we test six values for the standard deviation of the forecast error, namely 3%, 5%, 7%, 10%, 20%, and 30% of the forecasted residual demand. We calculate the system capacity at every hour using the optimal commitment schedule of each instance. Then, using the marginal distribution of the forecast error, we compute the LOLP at each hour. In Table 2.7, we report the average, median, third quartile, and standard deviation of the LOLP over the planning horizon. We evaluate the optimal schedule using the analytic function (2.23) to calculate the actual expected cost.

We observe that when the value of sigma is 0.03, increasing the unmet demand cost results in a slight increment of the expected cost. However, when the value is 0.3, increasing the unmet

demand cost results in a large increment of the expected cost. When both the forecast error variability and unmet demand cost are large, the model turns more units on until using the total capacity of the system. Beyond that limit, the model starts buying energy from other markets. In terms of the LOLP, we observe that the system is reliable, presenting a LOLP lower than 5% on average. Only when the error variability is 30% of the forecasted residual demand the LOLP is higher than 5%. Considering that the high penetration of RESs increases the supply uncertainty [59], a model that provides a less costly and reliable commitment schedule when including variability is advantageous [60].

		LOLP descriptive statistics				
Std. Dev. as a percentage of the mean residual demand	**Unmet demand cost ($)**	**Average (%)**	**Median (%)**	**3rd Quartile (%)**	**Std. Dev. (%)**	**Expected cost ($)**
3%	100	1	0	1	3	4,218,331
	200	1	0	1	1	4,219,215
	500	0	0	0	1	4,223,020
	1,000	0	0	0	1	4,243,806
5%	100	1	0	2	2	4,227,476
	200	1	0	1	1	4,231,235
	500	0	0	0	1	4,247,443
	1,000	0	0	0	0	4,266,372
7%	100	2	1	3	2	4,241,722
	200	1	0	1	1	4,252,557
	500	0	0	0	0	4,264,016
	1,000	0	0	0	0	4,294,136
10%	100	2	1	3	2	4,273,594
	200	1	0	1	1	4,285,507
	500	0	0	0	1	4,320,090
	1,000	0	0	0	1	4,355,814
20%	100	4	2	5	5	4,429,663
	200	4	1	4	6	4,540,944
	500	3	0	4	6	4,857,387
	1,000	3	0	4	6	5,379,034
30%	100	8	3	12	9	4,705,851
	200	7	2	12	9	5,062,468
	500	7	2	12	9	6,116,668
	1,000	7	2	12	9	7,864,118

Table 2.7: LOLP and expected cost when changing unmet demand cost and standard deviation

2.4 Conclusion

This Chapter proposed a new modeling approach to solve the two-stage SUCP model. This new approach does not require scenarios. Instead, the multivariate probability distribution of the error term of the forecasted residual demand was used to model the SUCP. Since the resulting objective function is nonlinear, a piecewise linear approximation was used to transform the

nonlinear SUCP into a MILP. The model that includes the technical constraints of the standard deterministic UCP was solved in a reasonable time using a commercial solver.

We assessed the accuracy of the piecewise linear approximation and determined the number of breakpoints by performing a stability analysis. Experimental results indicated that the linear approximation allows solving real-size problems using a few breakpoints. It was observed that increasing the number of breakpoints did not significantly increase the accuracy of the linear approximation. In addition, in most experiments, the processing time was within the range required by the ISOs.

The reliability of the system was assessed by studying the LOLP at higher levels of the forecast error variability. For most of the experiments, the results showed a LOLP on average of less than 5%. As more renewable energy sources become available, modeling the demand and supply uncertainty becomes highly relevant.

The main contribution of this Chapter is a novel approach to modeling the uncertainty within the SUCP and developing a suitable solution method that provides cost-efficient unit commitment schedules with acceptable levels of reliability. To the best of our knowledge, this is the first time the SUCP is solved without using scenario generation, which adds a new formulation for solving this relevant problem. Future research should compare the advantages and disadvantages of the proposed model to conventional SUCP formulations, which are based on generating scenarios.

Chapter 3

Modeling Ramping Constraints for the Statistical Unit Commitment Problem

3.1 Introduction

In power systems operations, the energy generation schedule is generated by solving the Unit Commitment Problem (UCP), a non-convex mixed-integer linear optimization problem (MILP). This problem has been studied for over four decades by the Operations Research community, and with the increasing integration of Renewable Energy Sources (RESs) into the energy matrix, addressing uncertainty has become more relevant when modeling the UCP. Thus, stochastic modeling is a good approach to incorporate uncertainties and account for disturbances impacting normal operations [61].

By including such uncertainties, decision-making processes achieve a higher level of realism that deterministic models cannot capture. Notably, this enables considering various factors, such as demand management programs, renewable energy production, and outages, which are essential in addressing the complexity of the UCP [62]. In addition, stochastic models can handle the variability of RESs, reducing reliance on manual reserves. Regarding reliability, stochastic schemes outperform deterministic ones due to their ability to accommodate potential scenarios [63]. Thus, adopting uncertainty through stochastic models can result in more accurate, flexible, and resilient solutions for power system operations.

The UCP started to be solved using algorithms such as dynamic programming [64, 65], priority list heuristics [66], and the Lagrangian-relaxation method [67, 68]. However, due to the progress of off-the-shelf solvers, researchers have focused on improving the mathematical formulation to speed up the time of the solution algorithms. In particular, most solvers

use the branch-and-cut algorithm (B&C) to solve Mixed Integer Linear Programming (MILP) optimization models [69]. The B&C algorithm is an enumerate algorithm that begins solving a simplified version of the original problem by relaxing integer variables. Because the algorithm starts with a relaxed problem, having tighter formulations can help reduce the number of potential solutions, tightening the search space. Thus, having a tighter formulation where the solution of the linear relaxation is closer to the solution of the integer model can speed up the solving process [70].

Since mathematical formulations directly impact how the B&C algorithm performs, several improvements for the deterministic formulation of the UCP have been proposed, focusing on enhancing the tightness of the commitment and ramping constraints. In [21], the authors proposed a new formulation for the self-commitment problem that modeled start-up unit trajectory through a tighter and more compact formulation. Later, in [71], the authors extended the previous work by implementing it in a centralized UCP. This implementation resulted in tighter and more compact formulation than in [72, 20, 73], leading to faster solving process. Following the same idea but for the power-based UCP, in [74], the authors derived the convex hull for operating slow and fast-start units, obtaining integer solutions when solving the UCP as a linear problem. At the same time, the self-commitment problem has been used to grasp the improvement of mathematical formulations, which can be extended to centralized UCP [75, 76]. A comprehensive review of the UCP formulations can be found in [18], where the authors consolidated into Julia's programming language library the most relevant formulations of the UCP. Thus, further computational studies can be done and easily implemented.

Many studies have provided better formulations for the Stochastic Unit Commitment Problem (SUCP) through valid inequalities. For example, in [77], cutting planes were proposed for the multi-stage security-constrained UCP. The authors proposed valid lifting and cover inequalities to strengthen the stochastic generation ramping and the load balance polytope. In addition to valid inequalities, the effort in stochastic optimization relies on proposing optimization methods that break down the structure of the problem, creating tractable sub problems that lead to an optimal solution in a reasonable time [78].

Some aspects of the generator units must be considered in the mathematical formulation of the UCP, such as the minimum online and offline time, their power, and ramping capacity. Without considering these elements, the solution could be infeasible in practice, limiting the applicability of the model. To have a more realistic model, in this Chapter, we model the ramping constraints for the statistical SUCP, that constrain the power trajectory of the generator units. We model the nonlinear expected cost using a modeling strategy that yields a logarithmic number of variables. In addition, different solving strategies are proposed to solve realistic instances of the SUCP. The remainder of the Chapter is as follows: Section 3.2 describes the mathematical formulation and Section 3.3 three strategies to solve the problem. We present the description of the power systems and results in Section 3.4. A case study using a power system from the California Independent System Operator (CAISO) is tested in Section 3.5. Finally, conclusions and future research are presented in Section 3.6.

3.2 Mathematical formulation

The following mathematical model minimizes the commitment and expected dispatch costs. We model the technical constraints using a mathematical formulation called "3bin". This formulation describes the transition between commitment states using three binary variables [47]. The model is described as follows:

Sets and indices

I Set of units, $i \in I$

T Set of periods, $t \in T$

L Set of breakpoints, $l \in L$

B Set of digits, $b \in B$

Parameters

C_i^{hot} Hot start-up cost of unit $i \in I$

C_i^{cold} Cold start-up cost of unit $i \in I$

c_i Energy cost of unit $i \in I$

$F_i(\cdot)$ Dispatch cost function of unit $i \in I$ in period $t \in T$

$H_{it}(\cdot)$ Start-up cost of unit $i \in I$ during period $t \in T$

P_i^{max} Maximum energy capacity of unit $i \in I$

P_i^{min} Minimum energy capacity of unit $i \in I$

RU_i Ramping up capacity of unit $i \in I$

RD_i Ramping down capacity of unit $i \in I$

T_i^{on} Minimum number of periods of unit $i \in I$ has to be on

T_i^{off} Minimum number of periods of unit $i \in I$ has to keep off

t_i^{cold} Number of periods after unit $i \in I$ becomes cold

u_i^{prev} Binary parameter that indicates the on/off state of unit $i \in I$ at the beginning of the planning horizon

Δ_i Difference between the maximum and minimum energy capacity P_i^{max} and P_i^{min} of unit $i \in I$

κ Cost of buying energy from other energy markets.

τ_i^{on} Number of periods unit $i \in I$ has been on prior to the first period of the planning horizon

τ_i^{off} Number of periods unit $i \in I$ has been off prior to the first period of the planning horizon

Λ Incidence matrix to model a piecewise linear approximation

Random variables

$\tilde{d}_t$ Random demand during period $t \in T$

$\tilde{r}_t$ Random residual demand during period $t \in T$

Decision variables

u_{it} Binary variable that indicates if the unit $i \in I$ is on/off during period $t \in T$

v_{it} Binary variable that indicates if the unit $i \in I$ is turned on during period $t \in T$

w_{it} Binary variable that indicates if the unit $i \in I$ is turned off during period $t \in T$

y_{it}^{hot} Binary variable that indicates if the unit $i \in I$ starts hot during period $t \in T$

y_{it}^{cold} Binary variable that indicates if the unit $i \in I$ starts cold during period $t \in T$

p_{it} Dispatch of energy from unit $i \in I$ during period $t \in T$

λ_{itl} Continuous variable that weighs the breakpoint $l \in L$ of unit $i \in I$ during period $t \in T$

η_{itb} Binary variable that indicates if the digit $b \in B$ is selected for the unit $i \in I$ during period $t \in T$

$$\min z = \sum_{t\in T}\sum_{i\in I} H_{it}(\cdot) + \sum_{t\in T}\sum_{i\in I} P_i^{min} c_i u_{it} + \sum_{t\in T}\sum_{i\in I} \mathbf{E}\left[F(\tilde{r_t})\right] \tag{3.1}$$

$$\text{s.t} \quad p_{it} \leq (P^{max} - P^{min}) u_{it} \quad \forall i \in I, t \in T \tag{3.2}$$

$$p_{it} - p_{it-1} \leq RU_i \quad \forall i \in I, t \in \{2, |T|\} \tag{3.3}$$

$$p_{it-1} - p_{it} \leq RD_i \quad \forall i \in I, t \in \{2, |T|\} \tag{3.4}$$

$$\sum_{j=\gamma_{it}^{on}}^{t} v_{ij} \leq u_{it} \quad \forall i \in I, t \in T \tag{3.5}$$

$$\sum_{j=\gamma_{it}^{off}}^{t} w_{ij} \leq 1 - u_{it} \quad \forall i \in I, t \in T \tag{3.6}$$

$$u_{it} = 1 \quad \forall i \in I : u_i^{prev} = 1 \quad \forall t \in \{1, ..., \theta_i^{on}\} \tag{3.7}$$

$$u_{it} = 0 \quad \forall i \in I : u_i^{prev} = 0 \quad \forall t \in \{1, ..., \theta_i^{off}\} \tag{3.8}$$

$$y_{it}^{hot} + y_{it}^{cold} = v_{it} \quad \forall i \in I, t \in T \tag{3.9}$$

$$u_{it} - \sum_{l=t-t_i^{cold}-1}^{t-1} u_{il} \leq y_{it}^{cold} \qquad \forall i \in I, t \in T \tag{3.10}$$

$$u_{it} - u_{it-1} \leq v_{it} \quad \forall i \in I, t \in T \tag{3.11}$$

$$w_{it} = v_{it} + u_{it-1} - u_{it} \quad \forall i \in I, t \in T \tag{3.12}$$

$$p_{it} \geq 0 \quad \forall i \in I, t \in T \tag{3.13}$$

$$u_{it}, v_{it}, w_{it}, y_{it}^{hot}, y_{it}^{cold} \in \{0, 1\} \quad \forall i \in I, t \in T \tag{3.14}$$

The objective function (3.1) minimizes the expected commitment and dispatch cost. The first component of the objective function is the start-up cost function, where $H_{it}(\cdot) = C_i^{hot} y_{it}^{hot} + C_i^{cold} y_{it}^{cold}$. The second component is the cost of operating the generator units at minimum capacity. The last component is the expected dispatch cost ($\mathbf{E}\left[F(\tilde{r_t})\right]$) that depends on the random residual demand during each period and is presented in the following subsection. Constraint (3.2) sets the capacity and minimum dispatch of each generator unit. Constraints (3.3) and (3.4) ensure the ramping capacity of each generator unit. Constraints (3.5) and (3.6) set the minimum up and down time respectively where $\gamma_{it}^{on} = \max\{t - T_i^{on} + 1, 1\}$ and

$\gamma_{it}^{off} = \max\{t - T_i^{off} + 1, 1\}$. The set of constraints (3.7) and (3.8) determine the initial commitment state of the units, depending on their previous state where $\theta_i^{on} = \max\{1, T_i^{on} - \tau_i^{on} + 1\}$ and $\theta_i^{off} = \max\{1, T_i^{off} - \tau_i^{off} + 1\}$. Constraint (3.9) ensures that the unit can start either hot or cold. This constraint is related to constraint (3.10) that indicates that a unit starts cold if the number of periods being off is greater than $t_i^{cold}\ \forall i \in I$. Constraint (3.11) ensures a unit is turned on in the current period if it was off during the previous period. Constraint (3.12) does the same but for turning off the unit. Finally, constraint (3.13) and (3.14) ensures the integrity of the variables.

3.2.1 Modeling ramping constraints

This section describes our approach to include ramping constraints on the statistical UCP model. From Chapter 2, the expected dispatch cost states as follows:

$$\mathbf{E}\left[F(\tilde{r}_t)\right] = \sum_{i \in I} c_i \int_{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i-1} \Delta_j u_{jt}}^{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i} \Delta_j u_{jt}} S_t(x)\, dx + \kappa \int_{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{|I|} \Delta_j u_{jt}}^{\infty} S_t(x)\, dx \qquad \forall t \in T \tag{3.15}$$

In order to include ramping constraints, the energy dispatch has to be controlled by a continuous variable. In Chapter 2, the dispatch was represented by the term $\Delta_j u_{jt}$. Thus, the energy dispatch was 0 or Δ_j. However, to provide more flexibility, this parameter can be transformed into a continuous decision variable p_{it} that controls the dispatch to be between 0 and Δ_i. By introducing the variable p_{it}, equation (3.15) can be modified, replacing the term $\Delta_j u_{jt}$. This modification allows the model to effectively manage the energy dispatch between two periods, facilitating the inclusion of ramping constraints. One consequence of this modification is that each generator unit will dispatch its maximum capacity subject to its corresponding ramping constraints. Thus, the expected dispatch cost for each period states as follows:

$$\mathbf{E}\left[F(\tilde{r}_t)\right] = \sum_{i \in I} c_i \int_{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i-1} p_{jt}}^{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{i} p_{jt}} S_t(x)\, dx + \kappa \int_{\sum_{k \in I} P_k^{min} u_{kt} + \sum_{j=1}^{|I|} p_{jt}}^{\infty} S_t(x)\, dx \qquad \forall t \in T \tag{3.16}$$

The expected cost of the new model will be at least the same as the original. Therefore, the model without ramping constraints provides a lower bound for the model with ramping constraints. Since the expected dispatch cost (3.16) is nonlinear, a linear approximation is presented.

Decimal	Gray code
0	000
1	001
2	011
3	010
4	110
5	111
6	101
7	100

Table 3.1: Gray code scheme

3.2.2 Modeling a piecewise approximation with a logarithmic number of variables

In this section, we describe a piecewise linear approximation approach that requires a logarithmic number of variables. This approach was proposed in [79], where the authors showed that the incremental formulation that models Special Order Set of type 2 (SOS2) constraints could be modified to have a $\lfloor \log_2 L \rfloor$ binary variables for each function to be approximated, where L is the number of breakpoints of the piecewise linear approximation. In order to implement this modeling approach, we use gray code. Let B be the set of digits in a gray code scheme to represent cardinal numbers. The number of digits required to represent the nonlinear function depends on the selected number of breakpoints. For example, Table 3.1 shows the gray code representation for cardinal numbers from 0 to 7. Thus, 8 breakpoints could be used by the gray code scheme of Table 3.1.

We present an example using Figure 3.1 to explain this modeling approach. First, since the function in Figure 3.1 is approximated using 4 breakpoints, we use the first four decimal of Table 3.1, resulting in only 2 digits of the gray code scheme. We can disregard the first digit because all the decimals numbers between 0 to 3 have a first digit containing 0. Thus, only 2 binary variables are used. Notice that if we use the incremental approximation piecewise method, the model would require 4 binary variables [48].

To represent the approximation mathematically on the formulation, we need to couple the variables λ_l and η. We create this mathematical relationship using an incidence matrix. Let Λ be an incidence matrix that indicates if digit $b \in B$ of the breakpoint $l \in L$ contains the value

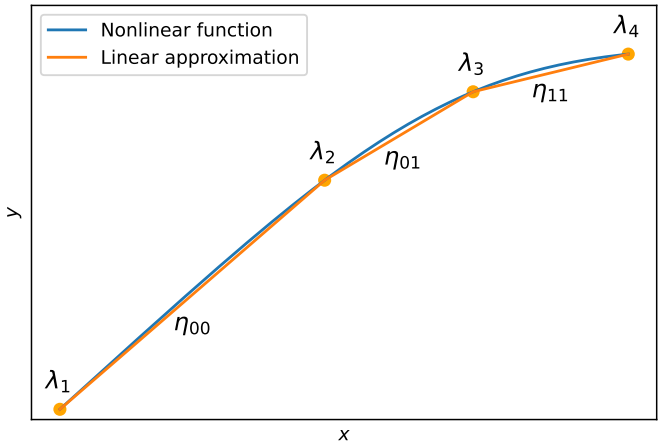


Figure 3.1: Linear approximation represented by gray code scheme

$k \in \{0, 1\}$. The corresponding incidence matrix of Figure 3.1 is expressed in equation 3.17, where the row B_{ij} corresponds to the digit $i \in B$ with the number $j \in \{0, 1\}$.

To represent the nonlinear function in Figure 3.1, we explain how to populate the second column of matrix Λ. By observing Figure 3.1, the second approximation point, λ_2 is contained in both the first and second segments, encoded as 00 and 01 represented by the variables η_{00} and η_{01} respectively. Consequently, for λ_2, the first digit of variables η_{00} and η_{01} only contain the cardinal number 0. As a result, the first and third rows of the matrix in the second column are filled with the number 1. Additionally, the second digit of variables η_{00} and η_{01} contain the numbers 0 and 1. Consequently, the second and fourth rows of the second column in matrix 3.17 are filled with 0 and 1, respectively.

$$\mathbf{\Lambda} = \begin{array}{c} \\ B_{10} \\ B_{11} \\ B_{20} \\ B_{21} \end{array} \begin{array}{c} \begin{array}{cccc} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \tag{3.17}$$

Once the incidence matrix is populated, the following equations describe the piecewise approximation of the function in Figure 3.1.

$$x = \sum_{l \in L} a_l \cdot \lambda_l \tag{3.18}$$

$$y = \sum_{l \in L} f(a_l) \cdot \lambda_l \tag{3.19}$$

$$\sum_{l \in L} \lambda_l = 1 \tag{3.20}$$

$$\sum_{l \in L : \Lambda(b,0,l) \neq 1} \lambda_l \leq \eta_b \quad \forall b \in B \tag{3.21}$$

$$\sum_{l \in L : \Lambda(b,1,l) \neq 1} \lambda_l \leq 1 - \eta_b \quad \forall b \in B \tag{3.22}$$

$$\lambda_l \geq 0 \quad \forall l \in L \tag{3.23}$$

$$\eta_b \in \{0, 1\} \quad \forall b \in B \tag{3.24}$$

By using this approximation scheme, the mathematical model in the previous section is formulated as follows:

$$\min z = \sum_{t\in T}\sum_{i\in I} H_i(\cdot) + \sum_{t\in T}\sum_{i\in I} c_i P_i^{\min} u_{it} + \sum_{t\in T}\sum_{i\in I} c_i \left[\sum_{l\in L} \lambda_{itl}\Gamma_t(b_{tl}) - \sum_{l\in L} \lambda_{i-1tl}\Gamma_t(b_{tl})\right] + \sum_{t\in T} \kappa \left[\mu_t - \sum_{l\in L} \lambda_{|I|tl}\Gamma_t(b_{tl})\right] \tag{3.25}$$

$$\text{s.t} \quad 3.2 - 3.14 \tag{3.26}$$

$$\sum_{l\in L} \lambda_{0tl} b_{tl} = \sum_{k\in I} P_k^{min} u_{kt} \qquad t \in T \tag{3.27}$$

$$\sum_{l\in L} \lambda_{itl} b_{tl} = \sum_{k\in I} P_k^{\min} u_{kt} + \sum_{j=1}^{i} p_{jt} \quad \forall i \in I, t \in T \tag{3.28}$$

$$\sum_{l=1}^{L} \lambda_{itl} = 1 \qquad \forall i \in I, t \in T \tag{3.29}$$

$$\sum_{l\in L:\Lambda(b,0,l)\neq 1} \lambda_{itl} \leq \eta_{itb} \quad \forall i \in I, t \in T, b \in B \tag{3.30}$$

$$\sum_{l\in L:\Lambda(b,1,l)\neq 1} \lambda_{itl} \leq 1 - \eta_{itb} \quad \forall i \in I, t \in T, b \in B \tag{3.31}$$

$$\lambda_{itl} \geq 0 \qquad \forall i \in I, t \in T, l \in L \tag{3.32}$$

$$\eta_{itb} \in \{0, 1\} \qquad \forall i \in I, t \in T, b \in B \tag{3.33}$$

The objective function (3.25) minimizes the start-up cost and the expected dispatch cost linearized by using piecewise linear approximation. The set of constraints (3.26) represents the dispatching and commitment constraints presented in section (3.2). Constraints (3.27) and (3.28) ensure the correct representation of the limits of the integral in expression (3.16). Constraints (3.29), (3.30), and (3.31) ensure that only two consecutive points in each segment of the approximation are selected. Finally, the integrity of the variables is defined in constraints (3.32) and (3.33).

Unit	P^{max}	P^{min}	Δ	c
1	455	150	305	16.19
3	130	20	110	16.50
2	130	20	110	16.60

Table 3.2: 3-units power system

Breakpoints	1	2	3	4
b_l	0	428.94	539.00	715
$\Gamma(b_l)$	0	422.04	485.62	500

Table 3.3: Breakpoints of function $\Gamma(x)$ using $\mu = 500$ and $\sigma = 75$

3.3 Methods to solve the statistical UCP

3.3.1 Indexing the generator units based on economic order

A tighter mathematical formulation leads to a better linear relaxation closing the gap between the solutions of the linear relaxation and integer problem. For this mathematical model, a tighter formulation can be achieved if the units are indexed according to their cost c_i before solving the optimization problem. Consider an example of a 3-unit power system as shown in Table 3.2. This system is analyzed in a single period of demand parameters $\mu = 500$ and $\sigma = 75$. The breakpoints b_l and their respective $\Gamma(b_l)$ function values are described in Table 3.3. For a given schedule $\hat{U}$, we relax the binary variables η_{ib} and solve the economic dispatch problem consisting of the objective function (3.25) and constraints (3.27)-(3.33). We solve the described optimization problem using the predefined order from Table 3.2 and by ordering the units by their cost, c_i. We call the former *predOrder* and the latter *econCostOrder*.

When using *predOrder* and *econCostOrder*, the solutions of the respective linear relaxations differ. These solutions are primarily affected by how the weights λ_l are used to calculate the linear combination of x and its resulting function value $\Gamma(x)$. Therefore, we examine these differences and their implications. The resulting objective function of *predOrder* and *econCostOrder* corresponds to $z_1 = 16,342.60422$ and $z_2 = 16,351.82854$, respectively. Since $z_2 > z_1$, the *econCostOrder* method provides a tighter relaxation. In both cases, all units dispatch their maximum capacity. However, the result of the weight variables λ_l differ.

	Weights				Resulting linear combination	
Units	λ_1	λ_2	λ_3	λ_4	x	$\Gamma(x)$
0	0.55	0.45	0.00	0.00	190.00	186.94
1	0.00	0.40	0.60	0.00	495.00	460.20
2	0.16	0.00	0.00	0.84	605.00	423.03
3	0.00	0.00	0.00	1.00	715.00	500.00

Table 3.4: Linear relaxation solution using *predOrder*

In Table 3.4 and Table 3.5, the weights variables λ_l and the linear combinations of x and $\Gamma(x)$ of each unit are presented for *predOrder* and *econCostOrder* approach, respectively. The columns x and $\Gamma(x)$ are calculated as the sum product between each row of weights and the values in Table 3.3. Thus, a linear combination between the weights λ_l and the breakpoints $(b_l, \Gamma(b_l))$ is calculated.

It is observed in Table 3.4 that when solving the linear relaxation using *predOrder*, the results yield a linear combination for the second unit that does not meet the condition of using only two consecutive λ_l. Even though there are two non-consecutive λs that result in the proper value of x. Its corresponding $\Gamma(x)$ value using λ_1 and λ_4 is not located in the last segment as shown in Figure 3.2. The last green dot uses a projection of a segment that passes through the origin and the last point of the approximation. Thus, when using *predOrder*, the variables η must be binary to enforce that the linear combination is always calculated using two consecutive weights λ.

For the case when using *econCostOrder*, for all the units, the solution of the relaxed problem yield two consecutive λ_l to compute the linear combination of the value of x. Moreover, the linear combinations also yield the proper value of the function $\Gamma(x)$. In Figure 3.2, each orange dot is in each segment of the piecewise linear function. Therefore, using *econCostOrder* yields a tighter linear relaxation which could lead to faster convergence of the B&C algorithm due to higher lower bounds.

3.3.2 Proposing valid cuts for the piecewise approximation

In this section, we introduce a set of valid inequalities to strengthen the formulation that approximates the nonlinear expected cost (3.16). Henceforth, we named this method *ValidInequalities*

	Weights				Resulting linear combination	
Units	λ_1	λ_2	λ_3	λ_4	x	$\Gamma(x)$
0	0.56	0.44	0.00	0.00	190.00	186.94
1	0.00	0.40	0.60	0.00	495.00	460.20
3	0.00	0.00	0.63	0.37	605.00	500.00
2	0.00	0.00	0.00	1.00	715.00	500.00

Table 3.5: Linear relaxation solution using *econCostOrder*

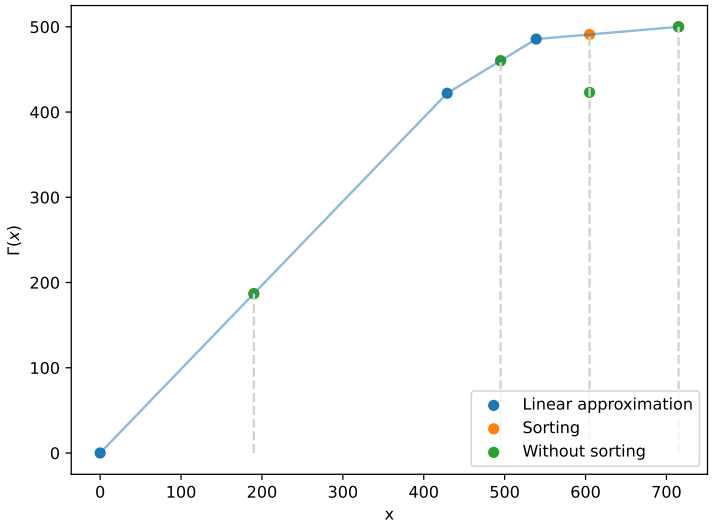


Figure 3.2: Linear combination under two sorting strategies

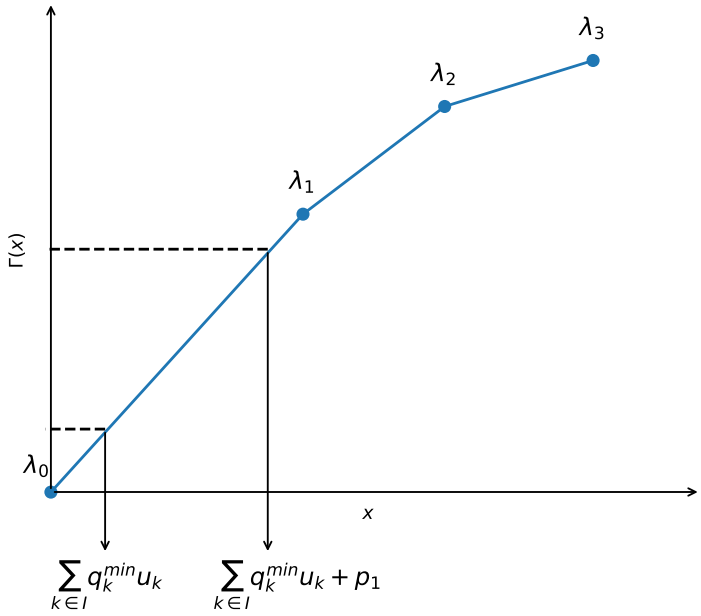


Figure 3.3: Correspondence between the linear approximation and the energy production

In the mathematical formulation presented in section 3.2.2, the variables λ_{itl} and λ_{itl+1} weights the points b_{tl} and b_{tl+1}, respectively. These weights are used to create a linear combination, determining the value on the right-hand side of constraints (3.27) and (3.28). By observing the linear combination in constraints (3.27) and (3.28), an increase in the energy production using the variable p_{it} results in an increase in the cumulative production. This observation implies that the cumulative production up to unit $i + 1$ must be at least equal to the cumulative production up to unit i, as specified in constraints (3.27) and (3.28). The above is shown in Figure 3.3 where the initial energy production increases by the value of p_1. Thus, the energy production along the x-axis consistently moves in the positive direction.

An analysis of the linear combination reveals that as the total production approaches the first breakpoint, the variable λ_{it0} is greater than λ_{it1}. Conversely, as the total production increases and approaches the second breakpoint, λ_{it1} increases in weight while λ_{it0} decreases. Thus, we impose the following constraint:

$$\lambda_{it0} \geq \lambda_{i+1t0} \qquad \forall i \in I, t \in T \tag{3.34}$$

Similarly, as the cumulative production reaches the last breakpoint, the variable λ_{itL} increases and λ_{itL-1} decreases. Thus, the following constraint is imposed:

$$\lambda_{itL} \leq \lambda_{i+1tL} \qquad \forall i \in I, t \in T \tag{3.35}$$

By using both constraints, there is an exact correspondence between the variables λ_{itl} and λ_{i+1tl} for the first and last breakpoints, resulting in a tighter relaxation when solving the original problem.

3.3.3 Generating an initial solution using a priority list heuristic

In this section, we present a hybrid heuristic approach, combining the priority list heuristic with the economic dispatch problem. Thus, a feasible solution is generated and used as a warm-up strategy before initiating the B&C algorithm.

The priority list heuristic method generates a list of generator units arranged in descending order of priority. This ranking determines the sequence in which the generators are started until a specific condition is satisfied. On the other hand, the economic dispatch problem focuses on optimizing the energy dispatch while adhering to technical constraints. In this context, we create an initial commitment schedule using the priority list heuristic. Then, the commitment schedule is fixed in the economic dispatch problem, maximizing the energy dispatch while meeting capacity and ramping constraints. The resulting solution is used as a warm-up strategy, serving as an initial solution.

The pseudo-code 1 describes the priority list method. This method generates an initial schedule by turning on units based on their index and assuming they dispatch maximum capacity. The energy production of each unit is accumulated in the variable g_{cum} for each period and is used to check if the probability of meeting the residual demand exceeds the specified threshold *pbb*. The heuristic stops when the threshold is met, or all units have been turned on. This process is repeated for all periods.

The number of units started in each period may vary depending on the residual demand distribution variance. Higher variance can lead to more units starting up. However, the priority

Algorithm 1 Priority list

for $t \in T$ **do**
  $g_{cum} = 0$
  $i = 0$
  while $P(g_{cum} \geq r_t) \geq pbb$ **and** $i \leq |I|$ **do**
    $g_{cum} = g_{cum} + P_i^{max}$
    $u_{it} = 1$
    $i = i + 1$
  end while
end for

list method does not consider the boundary conditions of each unit expressed in constraints (3.7) and (3.8) and the minimum up/down constraints (3.5) and (3.6). Thus, a heuristic is used to repair the schedule based on the method proposed in [80].

Assume $\theta^{off} = 1$, and $\tau^{off} = 2$, which represent the number of periods a unit should remain off at the beginning of the planning horizon and the number of periods the units must remain off after is shut down, respectively. After executing Algorithm 1, a unit has the commitment status represented by the first schedule in Figure 3.4 for a planning horizon of 8 hours. Based on the parameters of this unit, the schedule in Figure 3.4 is infeasible. First, it violates the time that should remain off at the beginning of the planning horizon (3.8), and the minimum down constraint (3.6). Thus, the repairing heuristic first ensures that the initial commitment status constraint is met by modifying the first period, as shown in Figure 3.4. Once the initial commitment status is repaired, the minimum down constraint is repaired by using the repairing heuristic 2.

First, the repairing heuristic 2 identifies all sequences of up (down) statuses. Then, each sequence is verified to ensure that meets the minimum up (down) constraint. To modify a sequence, equation (3.36) operates by generating a random number between 0 and 1 that is compared to the proportion of the time the unit is on (off) regarding the required time. For example, in Figure 3.5, the first sequence starts at period 2 and finishes at period 3. Since the minimum up constraint is not violated, it continues to the second sequence that starts in period 4. Since the sequence length of the following sequence contains only one period, the minimum down constraint is not met. Thus, there are two alternatives: first, the unit could be turned off

at period 5 or turned on at period 4, which is decided based on equation (3.36). Therefore, a feasible schedule is always generated.

$$\begin{cases} 1 & rand(0,1) \leq \frac{t^{on(off)}}{\tau^{up(down)}} \\ 0 & rand(0,1) > \frac{t^{on(off)}}{\tau^{up(down)}} \end{cases} \tag{3.36}$$

Hours	1	2	3	4	5	6	7	8
Status	1	1	1	0	1	0	1	1

→

1	2	3	4	5	6	7	8
0	1	1	0	1	0	1	1

Figure 3.4: Commitment status of a generator unit

Hours	1	2	3	4	5	6	7	8
Status	0	1	1	0	1	0	1	1

→

1	2	3	4	5	6	7	8
0	1	1	0	**0**	0	1	1

→

1	2	3	4	5	6	7	8
0	1	1	**1**	1	0	1	1

Figure 3.5: Repairing the commitment schedule of a generator unit

Algorithm 2 Sequence repairing heuristic

```
for i in units do
    Generate sequence of on and off periods
    for every sequence do
        if sequence on (off) violates minimum up (down) constraint then
            Repair sequence based on random number
        else
            Constraint is not violated, go to next sequence
        end if
    end for
end for
```

Once the schedule is generated, the dispatch of each unit is generated by maximizing the total dispatch subject to the ramping constraints as shown in model (3.37)-(3.41). Thus, a feasible solution is always generated and can be used as a warm-up solution.

$$\max z = \sum_{t \in T} \sum_{i \in I} p_{it} \tag{3.37}$$

$$\text{s.t} \quad p_{it} \leq (P^{max} - P^{min})\hat{u}_{it} \quad \forall i \in I, t \in T \tag{3.38}$$

$$p_{it} - p_{it-1} \leq RU_i \quad \forall i \in I, t \in \{2, |T|\} \tag{3.39}$$

$$p_{it-1} - p_{it} \leq RD_i \quad \forall i \in I, t \in \{2, |T|\} \tag{3.40}$$

$$p_{it} \geq 0 \quad \forall i \in I, t \in T \tag{3.41}$$

3.4 Testing the proposed solution strategies

3.4.1 Describing the power systems and residual demand model

In this section, we present a description of the systems and the residual demand modeling. We use three power systems to demonstrate the robustness of our methodology. Each system contains information about the generator units and the residual demand.

The first system, referred to as *Kazarlis100*, consists of 100 units and is constructed by replicating five times the 20-unit system detailed in [49]. The second system, named *OrLib100* is a synthetic power system containing 100 units sourced from the OR library *OR-Lib* [81]. Designed to mimic real-world systems, this power system serves as a benchmark for solving the UCP [64]. Lastly, the *Tejada214* system, described in [82], contains 214 units. This system does not include startup and shutdown costs. All three systems are solved within a 24-hour planning horizon, with the cost of unmet demand assumed to be 100 $\frac{\$}{MWh}$.

In this research, the stochastic residual demand, $\tilde{\mathbf{r}}$ is modeled as shown in equation (3.42). The vector value μ corresponds to the forecast of the residual demand, i.e., the difference between the expected demand and the expected renewable energy production. We assume that the error term ϵ follows a multivariate normal distribution with mean vector 0 and covariance matrix Σ.

$$\tilde{\boldsymbol{r}} = \boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{3.42}$$

Notice that the model does not restrict what forecasting model for the residual demand is used. Therefore, it is assumed that a forecast of the residual demand and estimates of the error term distribution parameters are known. Moreover, the RES production corresponds to solar or wind energy because the forecasted irradiation or wind speed is estimated.

For these experiments, the residual demand of the *Kazarlis100* instance is constructed based on the demand and wind power profile from [49] and [54], respectively. The other systems do not consider renewable energy production. Nevertheless, the same assumptions can be applied. We compute the covariance matrix Σ assuming a standard deviation equivalent to 15% of the mean vector. We assume an Autoregressive model AR(1). Equation (3.43) shows the corresponding covariance matrix with $|T|$ periods as the planning horizon.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \cdots & \sigma_1\sigma_T\rho^{T-1} \\ \sigma_2\sigma_1\rho & \sigma_1^2 & \cdots & \sigma_2\sigma_T\rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_T\sigma_1\rho^{T-1} & \sigma_T\sigma_2\rho^{T-2} & \cdots & \sigma_T^2 \end{pmatrix} \tag{3.43}$$

To compute the breakpoints of the functions $\Gamma(x)$ given in equation (3.25), we used the PWLF optimization library [57]. This library determines the breakpoints using a multi-start gradient method that minimizes the sum-of-squares of residuals. Since this method is a local search optimizer, it provides different breakpoints at each run. To ensure accuracy and consistency, we executed 20 approximations using the PWLF library considering 4, 8, and 16 breakpoints, resulting in 60 instances per system. The reasoning behind the number of breakpoints is based on the logarithmic approach to model the piecewise approximation. Using 4, 8, and 16 breakpoints will result in 2, 3, and 4 binary variables per function.

We reformulate the objective function to eliminate the dependence on constants in the objective function when solving the model. This approach disregards the constant value and transforms the objective function into a maximization problem. Thus, instead of minimizing the original expression, we maximize the negative value of the remaining expression. Maximizing the negative achieves the same optimization goal while providing interpretable values in the

objective function. Thus, the objective function to be optimized states as follows:

$$\begin{aligned}\max z = &- \sum_{t\in T}\sum_{i\in I} H_{it}(\cdot) - \sum_{t\in T}\sum_{i\in I} c_i P_i^{\min} u_{it} - \\ &\sum_{t\in T}\sum_{i\in I} c_i \left[\sum_{l\in L} \lambda_{itl}\Gamma_t(b_{tl}) - \sum_{l\in L} \lambda_{i-1tl}\Gamma_t(b_{tl})\right] - \\ &\sum_{t\in T} \kappa \left[-\sum_{l\in L} \lambda_{|I|tl}\Gamma_t(b_{tl})\right]\end{aligned} \tag{3.44}$$

All methods were coded on Python 3.9, and the models were solved using the solver Gurobi 9.5.1. The stopping condition was set at 600 seconds for all experiments in this section.

The strategies used are as follows:

- ***BaseCase***: The optimization problem is solved without any additional valid inequality or initial solution. Moreover, the system is used in its predefined order without sorting the units by their cost.
- ***validInequalities***: The optimization problem is solved using the set of valid inequalities proposed in Section 3.3.2.
- ***heuristicInit***: The optimization problem is solved by generating an initial solution when starting the B&C algorithm.
- ***econCostOrder***: The optimization problem is solved ordering the generator units by their costs.
- ***hybrid***: The optimization problem is solved combining *heuristicInit* and *econCostOrder*.

For every power system and breakpoint (4, 8, and 16), we generated a total of 20 instances. For each generated instance, we recorded the relaxation value at the beginning of the B&C algorithm and the value of the initial integer solution, which we refer to as the incumbent. Additionally, we stored the MIP gap upon completing the optimization process, reporting either a value of 0 if an optimal solution was achieved or the gap reached within the 600-second time limit. Thus, we calculated the average of these metrics among the corresponding power system and breakpoint combination.

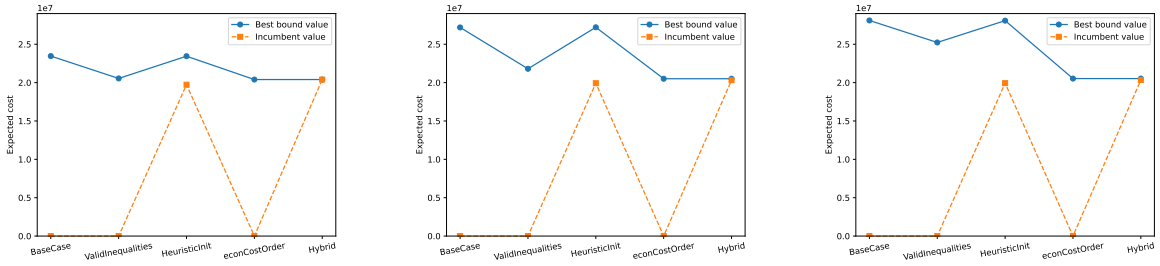
3.4.2 Analyzing the upper and lower bounds

In this section, we analyze the upper bound and incumbent of the objective function using the proposed solving strategies. Figures 3.6, 3.7, and 3.8 show the initial average upper bound and incumbent values obtained at the start of the B&C algorithm (y-axis) for each solving method (x-axis) of the *Kazarlis100*, *OrLib100* and *Tejada214* power systems, respectively.

Figures 3.6a, 3.6b, and 3.6c show the curves of the upper bound and incumbent value of the system *Kazarlis100* using 4, 8, and 16 breakpoints, respectively. The base case presents the highest upper bound and the lowest incumbent value when starting the B&C algorithm. This could lead to a longer time when solving the model due to the time to apply cuts and close subproblem nodes during the B&C algorithm. The upper bound should be lower to avoid exploring subproblem nodes that will not lead to an optimal solution.

When using *validInequalities*, the upper bound is lower than in the base case regardless of the number of breakpoints. However, the effectiveness decreases when adding more breakpoints because the valid inequalities only constrain the transition in the first and last segments of the piecewise functions. Thus, as the number of breakpoints increases, the valid inequalities lose their effectiveness in the upper bound of the objective function. See figures 3.6b and 3.6c. This method produced an unexpected result when applied to systems *OrLib100* and *Tejada214*. Figure 3.7c shows the upper bound and incumbent values of the *OrLib100* system using 16 breakpoints. No feasible solution was found within the specified time limit. The same outcome is observed in figures 3.8b and 3.8c, corresponding to the *Tejada214* system using 8 and 16 breakpoints, respectively. We attribute this outcome to the increased complexity of the relaxation problem when additional constraints are introduced. Since the B&C algorithm solves subproblem nodes iteratively, more constraints lead to prolonged solution times for each subproblem. Although the inclusion of valid inequalities enhances the relaxation of the problem, the base case has fewer constraints. Therefore, it is easier to solve, providing feasible solutions within the time limit.

When using *econCostOrder* instead of using *validInequalities*, the upper bound results in a lower value compared to the *BaseCase* and *validInequalities* methods. This observation is



(a) 4 breakpoints (b) 8 breakpoints (c) 16 breakpoints

Figure 3.6: Average bounds value at root node - *Kazarlis100* power system

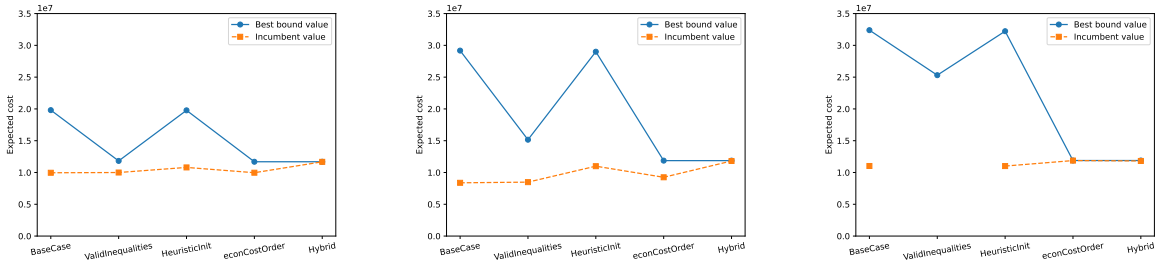
consistent across different numbers of breakpoints, as shown in figures 3.6a, 3.6b, and 3.6c for the *Kazarlis100* power system. The same results are observed in the *OrLib100* and *Tejada214* power systems, as shown in figures 3.7 and 3.8, respectively. Using *econCostOrder* provides a tighter linear relaxation, starting the B&C algorithm from a lower objective value in the linear relaxation.

In our discussion, we have presented methods that improve the upper bound without improving the incumbent value. One way to speed up the incumbent value generation is by feeding an initial solution to the solver. Thus, the B&C algorithm may start from a better solution and provide an estimate of the MIP gap. Figure 3.6a shows that when using *heuristicInit*, the initial incumbent value is higher and closer to the upper bound. This situation also occurs when extending to cases with more breakpoints (see figures 3.6b, and 3.6c). This strategy also leads to a better incumbent value for the other systems (see figures 3.7 and 3.8).

It is important to note that an optimal solution is reached when both the upper bound and incumbent are equal. Therefore, we combine both methods to close both curves and speed up the solving process. It is observed in figures 3.6, 3.7 and 3.8 that by using the *hybrid* method, the upper bound and incumbent value are almost the same independently of the number of breakpoints. Thus, we can guarantee a small MIP gap when using more breakpoints leading to higher accuracy.

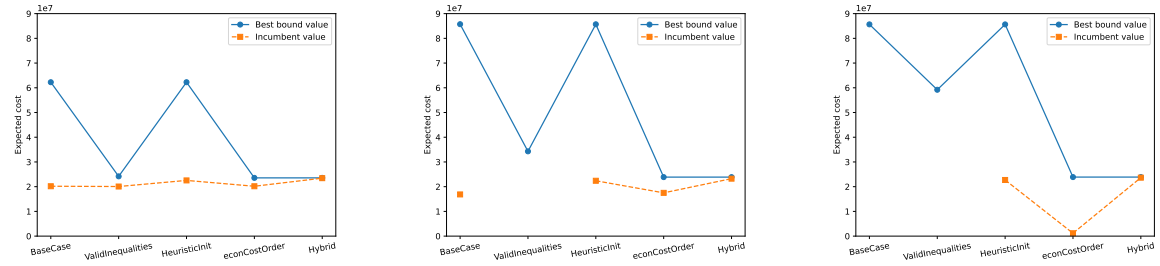
3.4.3 Analyzing the average MIP gap

In this subsection, we analyze the average MIP gap obtained from 20 replicates for each instance, which consisted of the *Kazarlis100*, *OrLib100*, and *Tejada214* power systems using 4,



(a) 4 breakpoints (b) 8 breakpoints (c) 16 breakpoints

Figure 3.7: Average bounds value - *OrLib100* power system



(a) 4 breakpoints (b) 8 breakpoints (c) 16 breakpoints

Figure 3.8: Average bounds value - *Tejada214* power system

8, and 16 breakpoints. The MIP gap of each instance was calculated after a maximum elapsed time of 600 seconds.

Figure 3.9 shows the average MIP gap of the *Kazarlis100* power system for the different methods we used to solve the optimization problem using 4, 8, and 16 breakpoints. For instances with 4 breakpoints, all methods produced an average MIP gap closes to zero. However, as the number of breakpoints increased to 8 and 16, only the *hybrid* method achieved an optimal solution. Specifically, when using 16 breakpoints, the optimization process did not find an optimal solution when using the *econCostOrder* method. In contrast, an optimal solution was achieved when using the *hybrid* method. We attribute this result to the low incumbent value of this instance, as shown in Figure 3.6c.

Figure 3.10 presents the average MIP gap of the *OrLib100* power system. We observed that using 4 breakpoints, the methods *validInequalities*, *econCostOrder*, and *hybrid* resulted in an acceptable average MIP gap. However, when using 8 breakpoints, the average MIP gap more than doubled, indicating a more challenging optimization scenario. Moreover, when utilizing 16 breakpoints, the method *validInequalities* failed to provide any feasible solution. Notice that

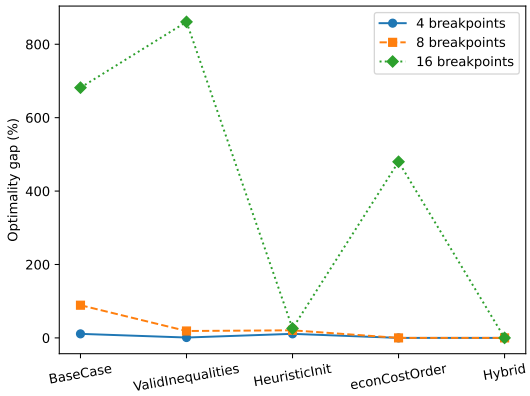


Figure 3.9: Average MIP gap of the *Kazarlis100* power system

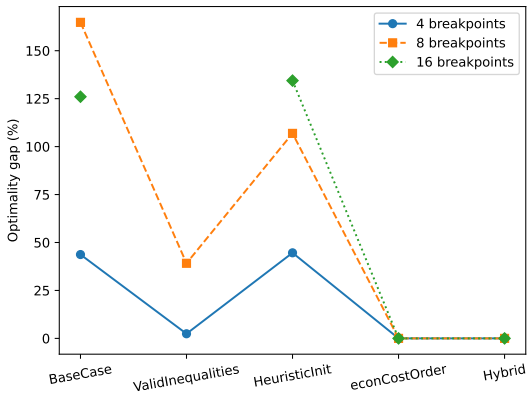


Figure 3.10: Average MIP gap of the *OrLib100* power system

if no feasible solution is found, it is not possible to calculate the MIP gap. The MIP gap requires an incumbent solution as a reference point to compute the distance to the current upper bound. For the *OrLib100* power system, employing *econCostOrder* or *hybrid* methods consistently resulted in optimal solutions for all instances.

Figure 3.11 shows the average MIP gap of the *Tejada214* power system. Like the previous systems, feasible solutions were obtained when using 4 breakpoints. However, when using 8 breakpoints, the *validInequalities* method failed to provide a feasible solution. Moreover, when using 16 breakpoints, both *BaseCase* and *validInequalities* methods did not produce feasible solutions. This outcome can be attributed to the size of the *Tejada214* system, which consists of 214 units, more than double the previous systems. Notably, the *econCostOrder* and *hybrid* methods resulted in smaller average MIP gap, indicating near-optimal solutions.

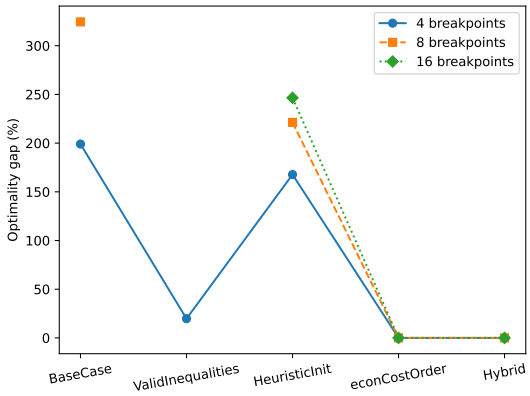


Figure 3.11: Average MIP gap of the *Tejada214* power system

Based on these results, using the *econCostOrder* method, consistently leads to near-optimal solutions due to the favorable upper bound obtained through the initialization of the B&C algorithm. However, as the problem complexity increases with additional breakpoints and a larger number of generator units in the system, relying exclusively on this method may fail to generate feasible or optimal solutions. To address this issue, an initial solution can be provided to the solver to handle larger problem instances. In addition, the solver can measure the quality of this initial solution by computing the MIP gap.

3.4.4 Comparing the logarithmic and incremental piecewise linear approximation methods

In the previous sections, we highlighted the advantages of using the proposed solving strategies to address the statistical UCP, reporting the average upper bound, incumbent value, and MIP gap as performance measures. However, the results were calculated using a logarithmic approach to model the piecewise linear model. In this section, we compare the performance of the SUCP model using two approaches to model the piecewise linear approximation, namely the incremental and the logarithmic approaches. Given the unfavorable outcomes observed in the base case and when using valid inequalities, we provide a comparative analysis only considering the *heuristicInit*, *econCostOrder*, and the *hybrid* strategies.

To perform the comparison, we generated 20 problem instances following the same procedure as in the previous section. These instances consider the power systems *Kazarlis100*, *OrLib100*, and *Tejada214* and 4, 8, and 16 breakpoints. By conducting this comparison, we

Breakpoints	Method	Model	Optimal Solutions	Average MIP Gap (%)	Std. MIP Gap (%)	Average Run Time (s)	Std. Run Time (s)	Average Var. Cost ($)	Average Var. Cost Analytical ($)
4	Heuristic	3bin	0	10.72	0.16	600	0	4,672,776	4,530,589
		3bin + LogApp	0	11.05	0.21	600	0	4,666,729	4,493,503
	sortByCost	3bin	20	0.00	0.00	8	0	4,464,508	4,388,865
		3bin + LogApp	20	0.00	0.00	6	0	4,464,511	4,385,922
	Heuristic + sortByCost	3bin	20	0.00	0.00	7	1	4,464,492	4,389,652
		3bin + LogApp	20	0.00	0.00	5	0	4,464,510	4,385,927
8	Heuristic	3bin	0	12.35	0.18	600	0	4,590,384	4,569,713
		3bin + LogApp	0	20.67	0.40	600	0	4,570,477	4,548,164
	sortByCost	3bin	20	0.00	0.00	110	70	4,360,308	4,339,700
		3bin + LogApp	20	0.00	0.00	83	62	4,360,495	4,339,754
	Heuristic + sortByCost	3bin	20	0.00	0.00	52	24	4,360,431	4,339,793
		3bin + LogApp	20	0.00	0.00	33	8	4,360,392	4,339,798
16	Heuristic	3bin	0	14.99	0.98	600	0	4,853,730	4,845,190
		3bin + LogApp	0	25.76	0.72	600	0	4,706,463	4,700,232
	sortByCost	3bin	20	0.00	0.00	482	67	4,341,543	4,337,528
		3bin + LogApp	0	479.88	137.22	601	0	21,144,847	21,144,031
	Heuristic + sortByCost	3bin	20	0.00	0.00	411	103	4,342,611	4,338,568
		3bin + LogApp	20	0.00	0.00	445	80	4,341,574	4,337,517

Table 3.6: Comparison between models using the Kazarlis power system

aim to obtain insights into the effectiveness of these two modeling approaches to model piecewise linear functions. Henceforth, we call the model using the incremental model as *3binD* and the logarithmic model as *3binLog*.

Tables 3.6, 3.7, and 3.8 provide a comparison between the optimization models *3binD* and *3binLog* using the *Kazarlis100*, *OrLib100*, and *Tejada214* power systems and 4, 8, and 16 breakpoints. For each table, the average and standard deviation of the MIP gap are reported after a maximum of 600 seconds, defined as the stopping condition. If the solver reaches an optimal solution in less than 600 seconds, the reported time corresponds to the elapsed time required to find an optimal solution. In addition, the tables provide the average and standard deviation of the elapsed time. Finally, we assess the accuracy of the statistical UCP approach by reporting the average expected dispatch cost obtained using the solver and its corresponding analytical value determined by evaluating each solution on the analytical expression presented in equation (3.16)

Table 3.6 presents the performance metrics of the optimization model when solving the *Kazarlis100* power system. The results indicate that both models produce similar outcomes regarding the average MIP gap. However, the *3binLog* model did not find an optimal solution when employing the *econCostOrder* method on the model with 16 breakpoints. Regarding the elapsed time, the *3binD* model exhibits a slightly higher average and standard deviation, mainly when using 16 breakpoints and applying the *hybrid* method. Regarding the analytical cost, by increasing the number of breakpoints, the optimization model can provide an accurate result, as the cost by optimization is similar to the analytical one.

Breakpoints	Method	Model	Optimal Solutions	Average MIP Gap (%)	Std. MIP Gap (%)	Average Run Time (s)	Std. Run Time (s)	Average Var. Cost ($)	Average Var. Cost Analytical ($)
4	Heuristic	3bin	0	44.96	0.77	600	0	14,738,374	14,572,867
		3bin + LogApp	0	44.62	0.92	600	0	14,675,589	14,514,742
	sortByCost	3bin	20	0.00	0.00	8	1	14,372,166	14,190,458
		3bin + LogApp	20	0.00	0.00	7	0	14,372,039	14,190,155
	Heuristic + sortByCost	3bin	20	0.00	0.00	10	3	14,372,138	14,190,343
		3bin + LogApp	20	0.00	0.00	8	1	14,372,331	14,190,242
8	Heuristic	3bin	0	57.15	0.37	600	0	14,692,048	14,668,771
		3bin + LogApp	0	106.90	2.32	600	0	14,847,487	14,824,385
	sortByCost	3bin	20	0.00	0.00	17	4	14,194,948	14,170,943
		3bin + LogApp	20	0.00	0.00	14	2	14,194,982	14,170,866
	Heuristic + sortByCost	3bin	20	0.00	0.00	23	8	14,194,795	14,170,793
		3bin + LogApp	20	0.00	0.00	15	1	14,194,915	14,170,855
16	Heuristic	3bin	0	95.97	0.54	600	0	15,499,245	15,492,642
		3bin + LogApp	0	134.41	3.25	600	0	14,937,777	14,932,214
	sortByCost	3bin	20	0.00	0.00	43	16	14,175,041	14,169,469
		3bin + LogApp	20	0.00	0.00	16	2	14,175,117	14,169,516
	Heuristic + sortByCost	3bin	20	0.00	0.00	50	42	14,175,004	14,169,456
		3bin + LogApp	20	0.00	0.00	23	1	14,174,828	14,169,312

Table 3.7: Computational comparison between models using the OrLib power system

Table 3.7 presents the performance metrics of the optimization model when solving the *OrLib100* power system. As in the previous system, the results when using both models are similar. Nevertheless, when applying the *heuristicInit* method with 8 and 16 breakpoints, the *3binD* model presents a much lower average MIP gap. Regarding elapsed time, the *3binLog* model can find an optimal solution in a slightly shorter and less variable time. This is observed when using 16 breakpoints and applying the *hybrid* method. Regarding the accuracy, when using 8 breakpoints, the model presents a difference of $ 24,000, representing around 0.1% of the total cost, and when increasing to 16 breakpoints, this difference is around 0.03%.

Table 3.8 presents the results when solving the model using the *Tejada214* power system. Since this power system has more generator units, the solver neither found an optimal nor near-optimal solution as in the previous power systems. For example, when applying the *heuristicInit* with 4 breakpoints, the result reported a MIP gap on average above 100%. When the model *3binLog* is solved, the elapsed time presents small variability. As in the previous systems, increasing the number of breakpoints increases the accuracy levels. For example, when using the *hybrid* method with 8 breakpoints in the *3binLog* model, the difference between the cost obtained using optimization and its respective analytical cost is around 0.38%, and it decreases to 0.19% when increasing the number of breakpoints to 16.

A surprising finding resulted from the analysis of the studied power systems. Tables 3.6, 3.7, and 3.8 show that when using 16 breakpoints and applying the *heuristicInit* method, the *3binD* model presents a smaller MIP gap compared to the *3binLog* model. However, this difference in the average MIP gap is not accurately observed in the objective value. Surprisingly, despite the smaller gap, the *3binD* model produces a higher objective value than the *3binLog*

Breakpoints	Method	Model	Optimal Solutions	Average MIP Gap (%)	Std. MIP Gap (%)	Average Run Time (s)	Std. Run Time (s)	Average Var. Cost ($)	Average Var. Cost Analytical ($)
4	Heuristic	3bin	0	169.66	0.56	600	0	4,761,526	4,552,037
		3bin + LogApp	0	167.74	1.08	600	0	4,761,526	4,552,037
	sortByCost	3bin	20	0.00	0.00	104	21	3,736,025	3,439,890
		3bin + LogApp	20	0.00	0.00	73	13	3,735,636	3,438,373
	Heuristic + sortByCost	3bin	20	0.00	0.00	119	24	3,735,895	3,439,192
		3bin + LogApp	20	0.00	0.00	84	15	3,735,850	3,438,590
8	Heuristic	3bin	0	171.31	0.63	600	0	4,581,211	4,536,328
		3bin + LogApp	0	221.32	26.97	600	0	4,900,765	4,858,031
	sortByCost	3bin	3	49.51	12.22	591	41	7,042,444	7,028,761
		3bin + LogApp	20	0.00	0.00	201	97	3,427,920	3,414,795
	Heuristic + sortByCost	3bin	19	0.03	0.14	334	127	3,435,111	3,422,397
		3bin + LogApp	20	0.00	0.00	158	67	3,428,222	3,415,148
16	Heuristic	3bin	0	173.87	0.50	600	0	12,450,469	12,434,011
		3bin + LogApp	0	246.54	1.66	600	0	4,584,593	4,561,185
	sortByCost	3bin	0	435.15	385.01	610	14	21,629,829	21,622,724
		3bin + LogApp	20	0.00	0.00	516	119	3,415,435	3,408,917
	Heuristic + sortByCost	3bin	15	0.01	0.01	414	177	3,417,241	3,410,887
		3bin + LogApp	20	0.00	0.00	322	34	3,415,545	3,409,125

Table 3.8: Computational comparison between models using the Tejada power system

Power System	Breakpoints	Model	Columns	Rows	Columns after presolve	Rows after presolve
Kazarlis100	4	3binD	33768	40768	20174	26528
		3binLog	31344	38344	20282	27097
	8	3binD	53160	50464	35749	35288
		3binLog	43464	43192	32446	31900
	16	3binD	91944	69856	65669	51667
		3binLog	65280	48040	57702	40372
Orlib100	4	3binD	33768	40922	17312	23186
		3binLog	31344	38498	17314	23209
	8	3binD	53160	50618	26758	29389
		3binLog	43464	43346	28128	29004
	16	3binD	91944	70010	50769	44254
		3binLog	65280	48194	57275	39554
Tejada214	4	3binD	72072	87295	46737	57854
		3binLog	66912	82135	46748	58634
	8	3binD	113352	107935	85435	80080
		3binLog	92712	92455	73853	69419
	16	3binD	195912	149215	137312	106245
		3binLog	139152	102775	120296	79656

Table 3.9: Problem size comparison before and after presolve

model. We attribute this behavior to the relaxation bound of the *3binLog* model. The solver cannot provide a tight bound for the *3binLog* leading to a higher MIP gap even though the integer solution is better than in the other model.

Table 3.9 presents the average size of each system using 4, 8, and 16 breakpoints for the *3binD* and *3binLog* models, both before and after the presolve. Specifically, the *3binLog* model exhibits a smaller number of columns and rows, indicating a reduced number of variables and constraints. However, after executing the presolve, both models converge to a similar size. The most notable distinction arises when employing 16 breakpoints, where the *3binD* model demonstrates a larger size compared to the *3binLog* model. Therefore, the similarity in the elapsed time reported in Tables 3.6, 3.7, and 3.8 can be attributed to the size of the model after the presolve step.

3.5 Solving the statistical SUCP on the California ISO system

In this section, a case study analyzes the performance and applicability of the model in a realistic power system. The selected system corresponds to the California ISO (CAISO) system, consisting of 610 generator units, 24 hours of planning horizon, and wind energy production equivalent on average between 20% to 40% of the energy demand. This system can be found in [18] and corresponds to the instance named "Scenario400_reserves_0". This system does not consider energy reserve, which is the case analyzed in this Chapter. We considered and scenario for wind energy production corresponding on average to 30% of the energy demand. Regarding the breakpoints, we calculated 20 replicates for 16 and 32 breakpoints using the PWFL library. For the residual demand, the standard deviation is considered as 15% of the mean vector.

Considering the results of the previous section, an initial solution is constructed using the priority list heuristic. Also, we applied the *sortByCost* method to obtain a tighter linear relaxation. Notice that even though the linear relaxation can be good, given the size of the problem, finding a feasible solution can be complex. Thus, this warming-up strategy can accelerate the finding of a feasible solution.

Table 3.10 shows the average and standard deviation for the MIP gap and the run time. Moreover, it presents the average expected variable cost obtained by using the solver and its corresponding value when evaluating the solution on its nonlinear function. By using 16 and 32 breakpoints, the solver can reach near-optimal solutions independent of the model. However, the solution results in a high average and variable MIP gap when using 32 breakpoints in the *3binLog* model. Although the gap is extremely high, when observing the cost, the *3binD* model is not much higher than the *3binLog* model. Thus, the higher gap is attributed to a poor bound during the optimization. We observe the same result in the previous section when observing the MIP gap and cost on tables 3.6, 3.7, and 3.8. Regarding the accuracy of the solution, when using 16 and 32 breakpoints, there is a deviation of 14% and 10%, respectively, regarding the analytical solution. Thus, higher accuracy levels can be achieved when using more breakpoints.

Breakpoints	Model	Optimal Solutions	Average MIP Gap (%)	Std. MIP Gap (%)	Average Run Time (s)	Std. Run Time (s)	Average Var. Cost ($)	Average Var. Cost Analytical ($)
16	3binD	0	0.36	0.01	602	2	33,575,704	29,452,434
	3binLog	0	0.36	0.01	602	2	33,546,718	29,407,215
32	3binD	0	0.35	0.03	601	1	31,542,551	28,735,913
	3binLog	0	1945.50	839.00	607	4	32,487,493	29,527,021

Table 3.10: Computational comparison between models using the CAISO power system

3.6 Conclusions

This Chapter extended the previous work about the statistical UCP by incorporating ramping constraints. These constraints ensure that the production of energy meets the ramping capability of each unit, incorporating more realism into this model. Furthermore, a new approximation method based on logarithmic piecewise linear approximation is employed to estimate the expected dispatch cost. Different solving strategies are proposed, aiming to solve the problem in a reasonable time. The robustness and applicability of the proposed model are demonstrated by testing on three synthetic power systems and a real power system obtained from the CAISO.

The experimental results showed improvements in the linear relaxation of the optimization problem through the implementation of valid inequalities and the sorting-by-cost strategy. These improvements led to tighter relaxations, resulting in near-optimal solutions. However, it was observed that when more breakpoints were used, only the sorting strategy remained effective in enhancing relaxation and finding an optimal solution. Similarly, when no feasible solution was found, the warm-up strategy generated an initial solution at the beginning of the algorithm, improving the incumbent value. Thus, since sorting the units by cost improved the upper bound and the heuristic approach improved the incumbent, combining both methods solved the problem across all the instances.

A comparative analysis was conducted between the *3binLog* and *3binD* models, with the former using a logarithmic number of variables to represent the piecewise linear approximation. At the same time, the latter corresponds to the conventional approach. It was observed that the *3binLog* model exhibits a significantly reduced number of columns compared to the traditional *3binD* model, resulting in slightly shorter computation times and lower variability when solving the optimization problem through the combined application of the heuristic strategy and sorting the units by cost. However, it was noticed that when employing an increased number

of breakpoints while relying only on the sorting strategy, the *3binLog* model showed poor relaxation, leading to a high MIP gap. This is an interesting finding that must be further studied. In [79], the authors proved that the logarithmic method provides better linear relaxations than the disaggregated method. Nevertheless, we found the opposite. Thus, the conditions of the logarithmic method should be analyzed to determine if the proposed mathematical model has a characteristic that breaks down the conditions of the logarithmic method.

In general, the combined solving approach of applying the heuristic method and sorting the units cost-effectively solved all instances, resulting in optimal or near-optimal solutions within the required solving time frame. Moreover, all the methods also hold for the mathematical model proposed in Chapter 2. Future research should further enhance the realism of the model by incorporating network constraints and transitioning towards a stochastic security-constrained approach.

Chapter 4

A comparison between the Statistical and Scenario-based Stochastic Unit Commitment Problem

4.1 Introduction

In power systems operations, the ISO must ensure the correct functioning and reliability of the system. Thus, it must guarantee the correct dispatching of energy from Generation Companies (GENCOs) to different demand nodes by solving the Unit Commitment Problem (UCP). The UCP is an optimization problem that provides a feasible energy generation schedule meeting technical and environmental constraints at minimum cost [83]. Because of the limited computational power of early computers, the first versions of the UCP did not include the uncertainty of several factors such as demand and availability of generating units. To overcome the risk intrinsic in the dispatching of electricity, the ISO included a reserve requirement in the UCP. Thus, generating an overproduction of energy.

As computers became more powerful, researchers started to propose a stochastic version of the UCP using modeling approaches such as stochastic programming, robust optimization, and chance-constrained.In the case of stochastic programming, the demand behavior is modeled by fitting demand data to a probability distribution and then sampling scenarios from this distribution [28, 84]. Adding renewal energy sources (RES) to the power grid imposes substantially new challenges. Since RESs depend on environmental conditions, the ISO must handle a higher level of variability during the scheduling process. For instance, when using wind power generation, the input of the system can vary from high to low production in a short period, affecting the ramping requirements. Moreover, the correlation between wind power generation

and the demand may be negative. Periods of low demand could result in curtailments or in periods of high demand in load shedding [85].

The modeling of uncertainty in power systems operations has relied on approximating the underlying probability distribution by a set of scenarios which are sampled using techniques such as the Monte Carlo method [86], the Quasi-Monte Carlo method [87] and Sample Average Approximation (SAA). In particular, the SAA method leverages the law of large numbers. The expected value is calculated over many randomly generated scenarios. As the number of scenarios increases, the SAA solution tends to converge to the true solution of the original stochastic problem [88]. In [89], the authors generated wind power scenarios using SAA. In [90], several scenarios were created to represent outages in the system. Then, a Loss of Load Probability (LOLP) chance constraint was included in the model to incorporate load uncertainty.

Although several scenarios are preferred to increase the accuracy of the model, they also increase the computational burden of the solving method. Nevertheless, scenario reduction methods can be used to reduce the computational complexity. In [41], the authors used the WILMAR model to generate a scenario tree. Their method was based on the ARMA time series, which fitted the regional forecast error and used scenario reduction to have a tractable problem. In [91], the authors used a forward selection heuristic to reduce the number of scenarios based on their cost and reliability impacts. The proposed method outperformed classical scenario reduction methods, leading to more reliable schedules. When solving the SUCP using stochastic programming, its inherent structure can be leveraged to address large scale instances effectively. Several decomposition techniques, including Lagrangian relaxation method [92, 86], column generation [28, 29], Benders decomposition [93], and progressive hedging [33], can take advantage of this structure to provide solutions efficiently.

Although a vast literature has addressed the SUCP, most of the methods consider the sampling of scenarios to approximate the expected generation costs. Since this approach is the accepted methodology to solve the SUCP, we use it as a benchmark. In this Chapter, the statistical SUCP presented in Chapter 2 and 3 is compared to the scenario-based model. We consider different size instances and contributions to real settings based on the elapsed time to find the optimal solutions and the stability of the solutions.

The remainder of this Chapter is as follows: Section 4.2 presents the mathematical formulation of both models, *statistical* and *scenario-based*. Section 4.3 compares both models considering a computational and stability analysis. Finally, Section 4.4 presents the final remarks of this Chapter and proposed future research.

4.2 Mathematical formulations

This section presents two mathematical formulations to minimize the expected dispatch cost. The first formulation uses a statistical approach to derive the expected dispatch cost. The second formulation is the SUCP scenario-based approach. Henceforth, we name these methods as *statistical* SUCP and *scenario-based* SUCP.

Sets and indices

I	Set of units, $i \in I$
T	Set of periods, $t \in T$
S	Set of scenarios, $s \in S$
L	Set of breakpoints, $l \in L$

Parameters

C_i^{hot}	Hot start-up cost of unit $i \in I$
C_i^{cold}	Cold start-up cost of unit $i \in I$
c_i	Energy cost of unit $i \in I$
$F_i(\cdot)$	Dispatch cost function of unit $i \in I$ in period $t \in T$
$H_{it}(\cdot)$	Start-up cost of unit $i \in I$ during period $t \in T$
P_i^{max}	Maximum energy capacity of unit $i \in I$
P_i^{min}	Minimum energy capacity of unit $i \in I$
RU_i	Ramping up capacity of unit $i \in I$
RD_i	Ramping down capacity of unit $i \in I$
T_i^{on}	Minimum number of periods of unit $i \in I$ has to be on
T_i^{off}	Minimum number of periods of unit $i \in I$ has to keep off
t_i^{cold}	Number of periods after unit $i \in I$ becomes cold

u_i^{prev} Binary parameter that indicates the on/off state of unit $i \in I$ at the beginning of the planning horizon

Δ_i Difference between the maximum and minimum energy capacity P_i^{max} and P_i^{min} of unit $i \in I$

κ Cost of buying energy from other energy markets.

τ_i^{on} Number of periods unit $i \in I$ has been on prior to the first period of the planning horizon

τ_i^{off} Number of periods unit $i \in I$ has been off prior to the first period of the planning horizon

$S(x)$ Survival function of the residual demand.

$\Gamma(x)$ Indefinite integral of the survival function $S(x)$.

Random variables

$\tilde{d}_t$ Random demand during period $t \in T$

$\tilde{r}_t$ Random residual demand during period $t \in T$

Decision variables

u_{it} Binary variable that indicates if the unit $i \in I$ is on/off during period $t \in T$

v_{it} Binary variable that indicates if the unit $i \in I$ is turned on during period $t \in T$

w_{it} Binary variable that indicates if the unit $i \in I$ is turned off during period $t \in T$

y_{it}^{hot} Binary variable that indicates if the unit $i \in I$ starts hot during period $t \in T$

y_{it}^{cold} Binary variable that indicates if the unit $i \in I$ starts cold during period $t \in T$

p_{it} Dispatch of energy from unit $i \in I$ during period $t \in T$

λ_{itl} Continuous variable that weighs the breakpoint $l \in L$ of unit $i \in I$ during period $t \in T$

η_{itb} Binary variable that indicates if the digit $b \in B$ is selected for the unit $i \in I$ during period $t \in T$

The following mathematical formulation states the constraints for the commitment decisions based on the *3bin* model provided in [47].

$$\min z = \sum_{t\in T}\sum_{i\in I} H_{it}(\cdot) + \sum_{t\in T}\sum_{i\in I} P_{it}^{min} c_i u_{it} + \sum_{t\in T}\sum_{i\in I} \mathbf{E}\left[F(\tilde{r}_t)\right] \tag{4.1}$$

$$\text{s.t} \quad p_{it} \leq (P^{max} - P^{min}) u_{it} \quad \forall i \in I, t \in T \tag{4.2}$$

$$p_{it} - p_{it-1} \leq RU_g \quad \forall i \in I, t \in \{2, |T|\} \tag{4.3}$$

$$p_{it-1} - p_{it} \leq RD_g \quad \forall i \in I, t \in \{2, |T|\} \tag{4.4}$$

$$\sum_{j=\gamma_{it}^{on}}^{t} v_{ij} \leq y_{it} \quad \forall i \in I, t \in T \tag{4.5}$$

$$\sum_{j=\gamma_{it}^{off}}^{t} w_{ij} \leq 1 - u_{it} \quad \forall i \in I, t \in T \tag{4.6}$$

$$u_{it} = 1 \quad \forall i \in I : u_i^{prev} = 1 \quad \forall t \in \{1, ..., \theta_i^{on}\} \tag{4.7}$$

$$u_{it} = 0 \quad \forall i \in I : u_i^{prev} = 0 \quad \forall t \in \{1, ..., \theta_i^{off}\} \tag{4.8}$$

$$y_{it}^{hot} + y_{it}^{cold} = v_{it} \quad \forall i \in I, t \in T \tag{4.9}$$

$$u_{it} - \sum_{l=t-t_i^{cold}-1}^{t-1} u_{il} \leq y_{it}^{cold} \qquad \forall i \in I, t \in T \tag{4.10}$$

$$u_{it} - u_{it-1} \leq v_{it} \quad \forall i \in I, t \in T \tag{4.11}$$

$$w_{it} = v_{it} + u_{it-1} - u_{it} \quad \forall i \in I, t \in T \tag{4.12}$$

$$p_{it} \geq 0 \quad \forall i \in I, t \in T \tag{4.13}$$

$$u_{it}, v_{it}, w_{it}, y_{it}^{hot}, y_{it}^{cold} \in \{0, 1\} \quad \forall i \in I, t \in T \tag{4.14}$$

The objective function (4.1) aims to minimize the overall expected cost, which includes commitment and dispatch costs. The fixed cost is denoted by $H_{it}(\cdot) = a_i^{\text{hot}} y_{it}^{\text{hot}} + a_i^{\text{cold}} y_{it}^{\text{cold}}$. In terms of the expected dispatch cost ($\mathbf{E}\left[F_i(\tilde{r}_t)\right]$), we present two models: the *statistical* SUCP and the *scenario-based* SUCP, presented in Sections 4.2.1 and 4.2.2, respectively. Constraint (4.2) states the capacity and minimum dispatch limits for each generator unit. To ensure ramping capacity, constraints (4.3) and (4.4) are introduced. Furthermore, constraints (4.5) and (4.6) enforce the minimum up and down time, respectively, where $\gamma_{it}^{on} = \max t - T_i^{on} + 1, 1$ and $\gamma_{it}^{off} = \max t - T_i^{off} + 1, 1$. The constraints (4.7) and (4.8) set the initial commitment

state of units based on their previous state, with $\theta_i^{on} = \max 1, T_i^{on} - \tau_i^{on} + 1$ and $\theta_i^{off} = \max 1, T_i^{off} - \tau_i^{off} + 1$. Constraint (4.9) ensures that a unit can start either in the hot or cold state, which relates to constraint (4.10) indicating that a unit starts cold if the number of off periods exceeds t_i^{cold}. Similarly, constraint (4.11) guarantees that a unit is turned on in the current period if it was off in the previous period. In contrast, constraint (4.12) performs the opposite operation for turning off the unit. Finally, constraints (4.13) and (4.14) ensure the integrity of the variables.

4.2.1 Statistical modeling

The expected dispatch cost can be modeled without using scenarios. Instead, we derived the expected dispatch cost function using the probability distribution of the residual demand. We assume that units are dispatched in economic order on their marginal cost c_i. Moreover, if the demand is not met the ISO can buy energy from other markets. The mathematical derivation can be seen in Section 2.2 of Chapter 2. The expected dispatch cost for a particular period $t \in T$ states as follows:

$$\begin{aligned}\mathbf{E}\left[F(\tilde{r_t})\right] = \sum_{i \in I} c_i \left[\Gamma_t(\sum_{k \in I} P_k^{min} y_{kt} + \sum_{j=1}^{i} \Delta_j y_{jt}) - \Gamma_t(\sum_{k \in I} P_k^{min} y_{kt} + \sum_{j=1}^{i-1} \Delta_j y_{jt})\right] + \\ \kappa \left[\mu_t - \Gamma_t(\sum_{k \in I} P_k^{min} y_{kt} + \sum_{j=1}^{|I|-1} \Delta_j y_{jt})\right] \qquad \forall t \in T\end{aligned} \tag{4.15}$$

Notice that $\Gamma(\cdot)$ corresponds to the indefinite integral of the probability survival function of the residual demand $S(x)$. Since this function is nonlinear, we use the piece-wise linear approximation method to generate a mixed-integer linear problem (MILP). Let L be the set of breakpoints for the piece-wise approximation method. Let λ be a continuous variable between 0 and 1 that weights the breakpoints b_l to approximate the $\Gamma(\cdot)$ function as the following

equations:

$$x = \sum_{l \in L} \lambda_l b_l \tag{4.16}$$

$$\Gamma(x) = \sum_{l \in L} \lambda_l \Gamma(b_l) \tag{4.17}$$

The mathematical formulation of the MILP states as follows:

$$\begin{aligned} \min z = & \sum_{t \in T} \sum_{i \in I} H_i(\cdot) + \sum_{t \in T} \sum_{i \in I} c_i P_i^{min} y_{it} + \\ & \sum_{t \in T} \sum_{i \in I} c_i \left[\sum_{l \in L} \lambda_{itl} \Gamma_t(b_{tl}) - \sum_{l \in L} \lambda_{i-1tl} \Gamma_t(b_{tl}) \right] + \\ & \sum_{t \in T} \kappa \left[\mu_t - \sum_{l \in L} \lambda_{|I|tl} \Gamma_t(b_{tl}) \right] \end{aligned} \tag{4.18}$$

$$\text{s.t} \quad (4.2) - (4.14) \tag{4.19}$$

$$\sum_{l=1}^{L} \lambda_{0tl} b_{tl} = \sum_{k \in I} P_k^{min} y_{kt} \qquad t \in T \tag{4.20}$$

$$\sum_{l=1}^{L} \lambda_{itl} b_{tl} = \sum_{k \in I} P_k^{min} y_{kt} + \sum_{j=1}^{i} \Delta_j y_{jt} \quad \forall i \in I, t \in T \tag{4.21}$$

$$\lambda_{itl} \in SOS2 \quad \forall i \in I, t \in T, l \in L \tag{4.22}$$

$$\lambda_{itl} \geq 0 \qquad \forall i \in I, t \in T, l \in L \tag{4.23}$$

The objective function (4.18) minimizes the start-up cost and the expected dispatch cost expressed in linear terms by using the piece-wise linear approximation. The set of constraints (4.19) corresponds to the technical constraints of the UCP presented at the beginning of this section. Constraint (4.2) and (4.3) ensures that the correct value b_{tl} is evaluated in the objective function. Constraint (4.22) states that the weights to formulate the linear formulation must be modeled by using special order set of type 2 (SOS2) to ensure that at most two consecutive λ_l are used. Finally, Constraints (4.23) ensure the integrity of the approximation variables.

4.2.2 Scenario-based modeling

We model the *scenario-based* SUCP as a two-stage problem. In the first stage, the commitment decisions are made. In the second stage, an economic dispatch problem is solved by using the committed units. In particular, the recourse function corresponding to the expected dispatch cost is modeled using a scenario-based approach. Thus, using the SAA method, several scenarios of the random variables are sampled using the Montecarlo sampling method.

Let Ω be the set of sampled scenarios using the Montecarlo method. For each scenario $\xi \in \Omega$, a vector for the residual demand is generated. This vector $\mathbf{R}_\xi = (r_{1\xi}, r_{2\xi}, \ldots, r_{|T-1|\xi}, r_{|T|\xi})$ corresponds to the residual demand for the $|T|$ periods for the scenario $\xi \in \Omega$. In addition, we create a continuous decision variable $p_{it\xi}$ corresponding to the amount of energy to be dispatched for the unit $i \in I$ during period $t \in T$ in the scenario $\xi \in \Omega$. Then, the expected cost is computed as the weighted average between the probability of each scenario and its cost. We have assumed that if the residual demand is not fulfilled, the ISO purchases energy from other markets. Hence, we create another continuous decision variable, $\varepsilon_{it\xi}$, that represents the amount of energy purchased from other markets at a price of κ. The expected cost is represented as follows:

$$\sum_{t\in T}\sum_{i\in I}\mathbf{E}\left[F(\tilde{r}_t)\right] = \sum_{t\in T}\sum_{i\in I} c_i P_i^{min} u_{it} + \sum_{\xi\in\Omega}\sum_{t\in T}\sum_{i\in I}\pi_\xi \cdot (c_i \cdot p_{it\xi} + \kappa \cdot \varepsilon_{t\xi}) \tag{4.24}$$

The following constraints are embedded in the above problem, resulting in the scenario-based SUCP.

$$\min z = \sum_{t\in T}\sum_{i\in I} H_i(\cdot) + \sum_{t\in T}\sum_{i\in I} c_i P_i^{min} u_{it} + \sum_{\xi\in\Omega}\sum_{t\in T}\sum_{i\in I} \pi_\xi \cdot (c_i \cdot p_{it\xi} + \kappa \cdot \varepsilon_{t\xi}) \tag{4.25}$$

$$\text{s.t} \quad (4.5) - (4.14) \tag{4.26}$$

$$p_{it\xi} \leq (P_i^{max} - P_i^{min}) u_{it} \quad \forall i \in I, t \in T, \xi \in \Omega \tag{4.27}$$

$$\sum_{i\in I} P_i^{min} u_{it} + \sum_{i\in I} p_{it\xi} + \varepsilon_{t\xi} \geq r_{t\xi} \quad \forall t \in T, \xi \in \Omega \tag{4.28}$$

$$p_{it\xi} - p_{it-1\xi} \leq RU_g \quad \forall i \in I, t \in \{2, |T|\}, \xi \in \Omega \tag{4.29}$$

$$p_{it-1\xi} - p_{it\xi} \leq RD_g \quad \forall i \in I, t \in \{2, |T|\}, \xi \in \Omega \tag{4.30}$$

$$p_{it\xi} \geq 0 \quad \forall i \in I, t \in T, \xi \in \Omega \tag{4.31}$$

The objective function (4.25) minimizes the fixed and expected costs considering the weighted sum between the probability of occurrence and the cost of each scenario. Constraints (4.26) correspond to the base formulation that guarantees the commitment constraints. Constraint (4.27) sets the upper bound for the dispatch of each unit. Constraint (4.28) ensures that the residual demand is fulfilled by thermal units or using energy purchased from other markets in case of shortage using the variable ε. Constraints (4.29) and (4.30) represent the ramping capability of each unit. Finally, (4.31) states the integrity of the dispatch variable.

4.3 Results

In this section, we present a comparison of both models. To perform the comparison, we use a modified version of the 20-unit system from [49] replicated 5, 10, and 20 times resulting in a 100, 200, and 400-unit system. We assume the residual demand follows a multivariate normal distribution with mean vector μ and covariance matrix Σ. The mean vector is calculated as the difference between the estimated demand and wind energy generation considering profiles from [49] and [54], respectively. Notice that the original profile of the residual demand corresponds to the 20-unit system. Thus, we replicate profiles 5, 10, and 20 times to have an appropriate

residual demand for each system. Finally, the σ vector is calculated as a percentage of the mean vector, and the cost of the unmet demand κ is assumed to be 100 $\frac{\$}{MWh}$

The results of this section were obtained using Gurobi 9.5 (64-bit) on a Linux computer with an Intel Xeon E5-2670v2 Ivy Bridge 2.5 GHz using 8 cores and 64 GB of RAM at the Alabama Supercomputer Center. When solving the optimization problem, we set the optimization criteria considering a wall clock of 600 seconds.

We sampled the scenarios directly from the multivariate normal distribution using the statistical library SciPy [94]. For the *statistical* model, we used the optimization library PWLF [57] to determine the points for the piece-wise linear approximation. Using the multi-start gradient method, the library minimizes the quadratic error between the breakpoints and the function $\Gamma(x)$. We model the SOS2 constraints using the logarithmic piecewise approximation method proposed in [79].

4.3.1 Computational comparison

This section presents a computational analysis comparing the *statistical* and the *scenario-based* models. We generated 20 instances of 4, 8, 16, and 32 breakpoints for the *statistical* model using a power system of 100, 200, and 400 units. For the *scenario-based* model, we generated 20 instances of 50, 100, 200, 500, and 1000 scenarios using a power system of 100, 200, and 400 units.

Table 4.1 presents the descriptive statistics of the MIP gap for both models. The instances of the statistical model are named using the number of points, i.e., "Stat4," and the instances of the scenario generation model are named using the number of scenarios, i.e., "Sce50". As the size of the problem increases due to the number of units, breakpoints, and scenarios, the MIP gap increases, and it is harder to find an optimal solution. In general, the *statistical* model found optimal solutions when using 4 and 8 breakpoints in the systems of 100 and 200 units in 75% of cases. There is an exception when using the model of 400 units, where an optimal solution Was found only when using 4 breakpoints.

Moreover, when using 16 and 32 breakpoints, the *statistical* model failed to find an optimal solution, resulting in a poor MIP gap. In particular, the *scenario-based* model presented

Units	Model	Optimal Solutions	Average (%)	Minimum (%)	Median (%)	3rd Quartile (%)	Maximum (%)
100	Stat4	20	0	0	0	0	0
	Stat8	20	0	0	0	0	0
	Stat16	0	79.08	77.6	79.26	79.51	80.51
	Stat32	0	81.37	80.91	81.45	81.52	81.87
	Sce50	20	0	0	0	0	0
	Sce100	20	0	0	0	0	0
	Sce200	20	0	0	0	0	0
	Sce500	20	0	0	0	0	0.01
	Sce1000	0	65.19	0.14	82.51	82.52	82.53
200	Stat4	20	0	0	0	0	0
	Stat8	7	6.82	0	0	0	75.74
	Stat16	0	80.48	78.65	80.58	80.84	81.45
	Stat32	0	81.73	81.37	81.78	81.81	81.97
	Sce50	20	0	0	0	0	0
	Sce100	20	0	0	0	0	0
	Sce200	20	0	0	0	0	0.01
	Sce500	0	60.85	0.17	82.51	82.52	82.54
	Sce1000	0	99.07	98.95	99.07	99.1	99.17
400	Stat4	20	0	0	0	0	0
	Stat8	0	69.29	55.69	76.69	78.7	81.02
	Stat16	0	81.84	80.7	82.08	82.26	82.54
	Stat32	0	82.34	82.16	82.33	82.42	82.55
	Sce50	20	0	0	0	0	0
	Sce100	0	0.01	0	0	0.01	0.02
	Sce200	0	0.14	0.06	0.14	0.17	0.31
	Sce500	0	99.11	98.96	99.13	99.15	99.21
	Sce1000	0	-	-	-	-	-

Table 4.1: MIP gap descriptive statistics

satisfactory results when using up to 200 scenarios and up to 500 scenarios for the instance of 100 units.

Table 4.2 presents the descriptive statistics of the elapsed time for the different power systems solved by different instances. The instances of the statistical model are named using the number of points, i.e., "Stat4," and the instances of the scenario generation model are named using the number of scenarios, i.e., "Sce50". It can be observed that the *statistical* model using 4 breakpoints was solved in less than 70 seconds for the systems of 100, 200, and 400 units. Then, when using more breakpoints, the elapsed time increased, taking 600 seconds to finish the optimization which was the stopping condition. The above is related to the MIP gap presented in Table 4.1 where the instances using 8, 16, and 32 breakpoints presented higher gaps than when using 4 breakpoints. In Table 4.2, the instances using the *scenario-based* model presented longer times than the *statistical* model. The model presented competitive times compared to the *statistical* model only when 50 scenarios were used.

In general, both models can be accelerated by implementing other solving strategies, such as initial heuristics in the case of the *statistical* model or using the Benders decomposition

Units	Model	Average (s)	Minimum (s)	Median (s)	Third Quartile (s)	Maximum (s)
100	Stat4	5	5	5	5	6
	Stat8	112	23	83	148	374
	Stat16	601	600	601	601	604
	Stat32	603	600	601	601	611
	Sce50	125	76	112	148	229
	Sce100	225	153	208	247	407
	Sce200	454	344	427	509	601
	Sce500	601	600	601	601	602
	Sce1000	601	600	601	601	603
200	Stat4	27	25	26	27	30
	Stat8	543	250	600	601	601
	Stat16	601	600	601	601	602
	Stat32	601	600	601	601	601
	Sce50	302	170	303	338	548
	Sce100	533	288	600	601	603
	Sce200	601	600	601	601	602
	Sce500	602	600	601	602	606
	Sce1000	608	602	606	611	626
400	Stat4	60	56	58	64	65
	Stat8	603	601	602	603	617
	Stat16	601	600	601	602	603
	Stat32	603	600	603	604	606
	Sce50	577	446	600	601	603
	Sce100	601	600	601	601	606
	Sce200	601	600	601	601	602
	Sce500	606	602	604	606	623
	Sce1000	602	602	602	602	606

Table 4.2: Run time descriptive statistics

method to solve the *scenario-based* model. Nevertheless, comparing without these extra strategies allowed equivalently comparing the computational time and their results.

4.3.2 Stability analysis

The stochastic optimization models presented in Section 4.2 are an approximation of the real problem due to the use of breakpoints and scenarios. Thus, to compare both models, we assess the stability of their results using the analytical expected cost described in Equation (4.15). In [58], the authors proposed a method to assess the suitability of a scenario generation approach when solving a stochastic programming model. This method evaluates the stability of solutions obtained from an approximation model compared to the actual stochastic optimization model.

Hence, to evaluate the stability of both the *statistical* and *scenario-based* models, we conduct a out-of-sample and in-sample stability analyses.

Consider a decision vector $\hat{\mathbf{x}}$ representing a commitment schedule. Let $\mathcal{F}(\mathbf{x})$ denote the true value of the expected dispatch cost, while $\hat{\mathcal{F}}_{STAT}(\mathbf{x})$ and $\hat{\mathcal{F}}_{SCE}(\mathbf{x})$ are the approximated values of the expected dispatch cost using the *statistical* and *scenario-based* models, respectively. These functions are used for assessing the out-of-sample and in-sample stability of the models.

Let K denote the number of instances generated, where the *statistical* model produces K sets of breakpoints the *scenario-based* model generates K sets of scenarios. Each instance is solved, and their resulting solutions are evaluated using the true objective function. We conclude that out-of-sample stability has been achieved if the solutions have approximately the same objective value when evaluated with the true objective function. This is described in equation (4.32).

$$\mathcal{F}(\hat{\mathbf{x}}_{\mathbf{k}}) \approx \mathcal{F}(\hat{\mathbf{x}}_{\mathbf{l}}) \qquad \forall k, l \in \{1, \ldots, K\}, k \neq l \tag{4.32}$$

Similarly, in-sample stability is assessed using each of the approximation models. For each formulation, we expect the solutions of the K optimization models to yield approximately the same objective value. Thus, in-sample stability is achieved. This is described in equations (4.33) and (4.34).

$$\hat{\mathcal{F}}_{STAT}(\hat{\mathbf{x}}_{\mathbf{k}}) \approx \hat{\mathcal{F}}_{STAT}(\hat{\mathbf{x}}_{\mathbf{l}}) \qquad \forall k, l \in \{1, \ldots, N\}, k \neq l \tag{4.33}$$

$$\hat{\mathcal{F}}_{SCE}(\hat{\mathbf{x}}_{\mathbf{k}}) \approx \hat{\mathcal{F}}_{SCE}(\hat{\mathbf{x}}_{\mathbf{l}}) \qquad \forall k, l \in \{1, \ldots, N\}, k \neq l \tag{4.34}$$

Notice that in the *statistical* model, the source of uncertainty comes from the method that selects the breakpoints. In contrast, for the *scenario-based* method, the source of uncertainty comes from the distribution itself. Thus, due to the presence of randomness in the generation methods, the outcome is expected to be different every time the piece-wise function is approximated, or a sample of scenarios is generated. Then, when solving the optimization model, the optimal solution of both models may provide a different value when evaluating the commitment schedule $\hat{\mathbf{x}}$ in the function $\mathcal{F}(\mathbf{x})$.

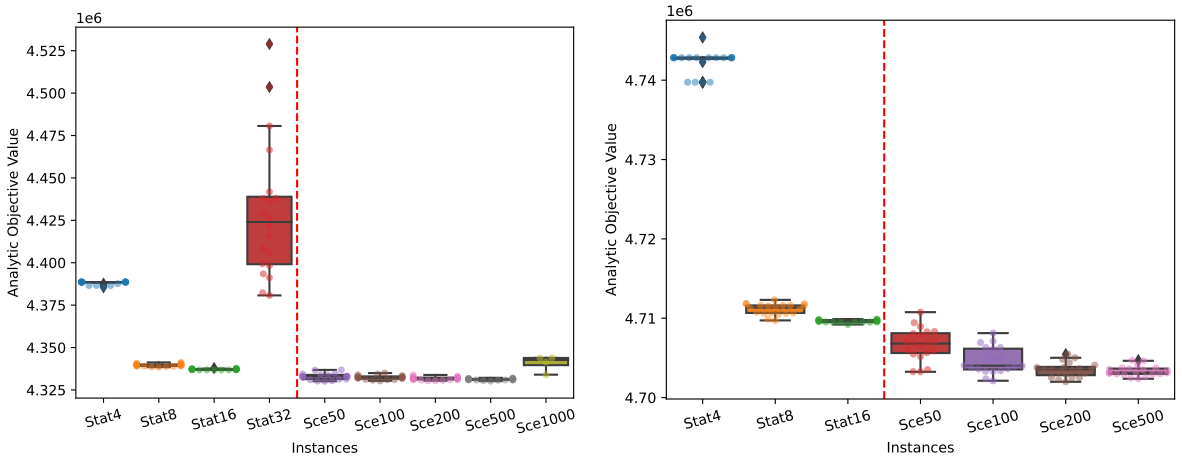
To assess the stability, we use a set K of 20 instances for each number of breakpoints and scenarios. For the *statistical* model, we used 4, 8, 16, and 32 breakpoints, and for the *scenario-based* model, we used sets of 50, 100, 200, 500, and 1000 scenarios. The sets were generated for each system. Thus, a total of 180 instances were generated for each system. In addition, we considered two levels for the parameter σ, namely 0.15μ and 0.30μ.

We solved the *statistical* model using a heuristic described in Chapter 3. This heuristic generates an initial solution when starting the branch-and-cut algorithm. This presents an advantage over the *scenario-based* model. However, this results in more instances within a MIP gap of 1%.

Figures 4.1, 4.2, and 4.3 show a boxplot comparing the out-of-sample stability between instances of the *statistical* and the *scenario* generation models with 100, 200, and 400 units, respectively. For visual purposes, we included only the experiments resulting in a MIP gap below 1%. Some sets were not displayed in the figures because none of their instances were solved within a gap of 1%. In the figures, the x-axis contains the name of the instances, and the y-axis the corresponding analytical objective value. The instances of the statistical model are named using the number of points, i.e., "Stat4," and the instances of the scenario generation model are named using the number of scenarios, i.e., "Sce50".

Figure 4.1a, 4.2a, and 4.3a present a boxplot of the objective values for instances with 100, 200, and 400 units, respectively. The solutions are generated with a standard deviation equal to 15% of the mean residual demand vector. Notice that the solutions became more stable as the number of breakpoints and scenarios increased. There are some exceptions with the solutions of "Stat32" and "Sce1000" in Figure 4.1a, "Stat16", "Stat32", and "Sce500" in Figure 4.2a and, "Stat16", "Stat32" in Figure 4.3a. These instances presented more variability due to a higher MIP gap than others.

When comparing the stability of both models, the *scenario-based* model was less stable than the *statistical* model. Moreover, the marginal improvement of the stability decreased when adding more scenarios. Thus, for the 100-unit system, using 16 breakpoints and 200 scenarios provide a suitable result. For the 200-unit system, 8 breakpoints and 200 scenarios are suitable. Finally, for the 400-unit system, 8 breakpoints and 100 scenarios are suitable. Adding more



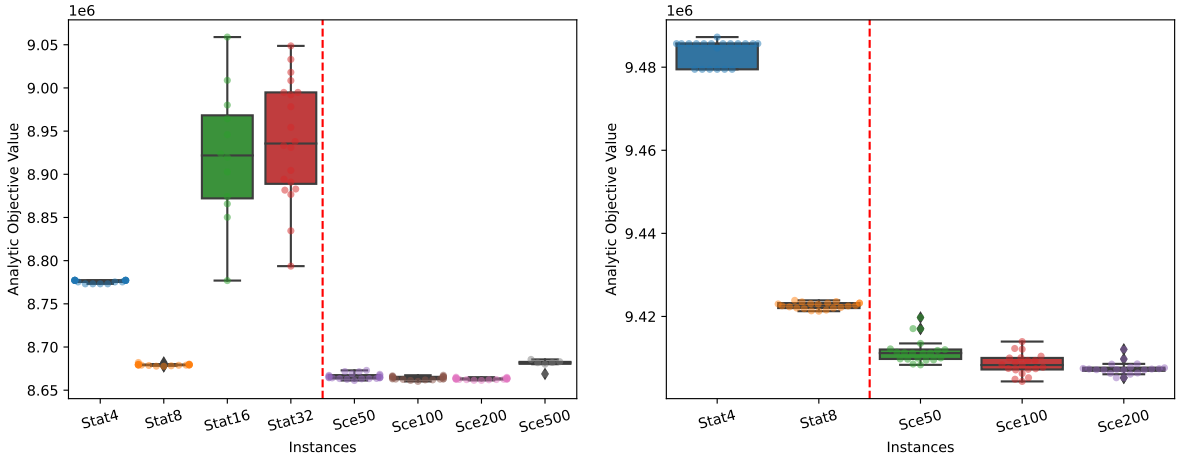
(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.1: Out-of-sample stability 100-Unit system

breakpoints and scenarios would improve the stability and decrease the value of the analytical, objective function. Nevertheless, to do so, other solution methods should be used. In particular, for the *scenario-based* model, Benders decomposition is usually used to solve the problem with many scenarios [95].

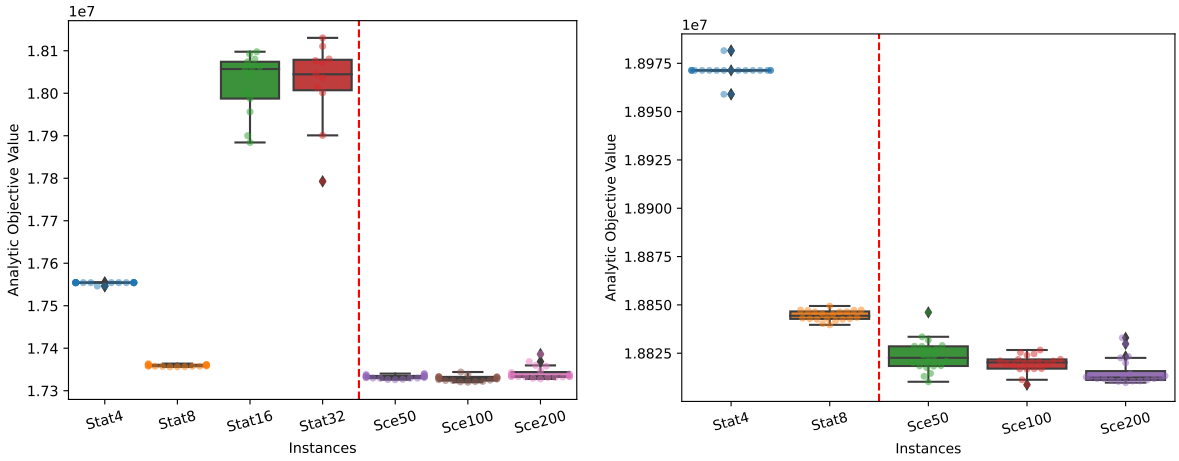
In Figure 4.1b, 4.2b, and 4.3b, the standard deviation is increased to be 30% of the mean vector for the system of 100, 200, and 400 units, respectively. Even though there is high variability, the *statistical* model is slightly more stable than the *scenario-based* model. Notice that the statistical model would need only 8 breakpoints to present better stability than the model with scenarios. We attribute this result to the source of uncertainty when generating scenarios. In particular, the source of uncertainty comes directly from the distribution. Thus, increasing the parameter σ will require more scenarios to reach a stable solution.

Figure 4.4, 4.5 and 4.6 show a boxplot comparing the in-sample stability between instances of the *statistical* and the *scenario-based* models with 100, 200, and 400 units, respectively. The *statistical* model is more stable than the *scenario-based* model for $\sigma = 0.15\mu$ and $\sigma = 0.30\mu$. We attribute this result to how the objective functions for both models are derived. For example, the *statistical* model is derived directly from the probability distribution of the residual demand, and the way is incorporated into the model is by using piecewise linear approximation. Even though fewer breakpoints can result in a poor approximation, they tend to have the same objective value. Thus, when adding more breakpoints, a lower objective value is observed. On



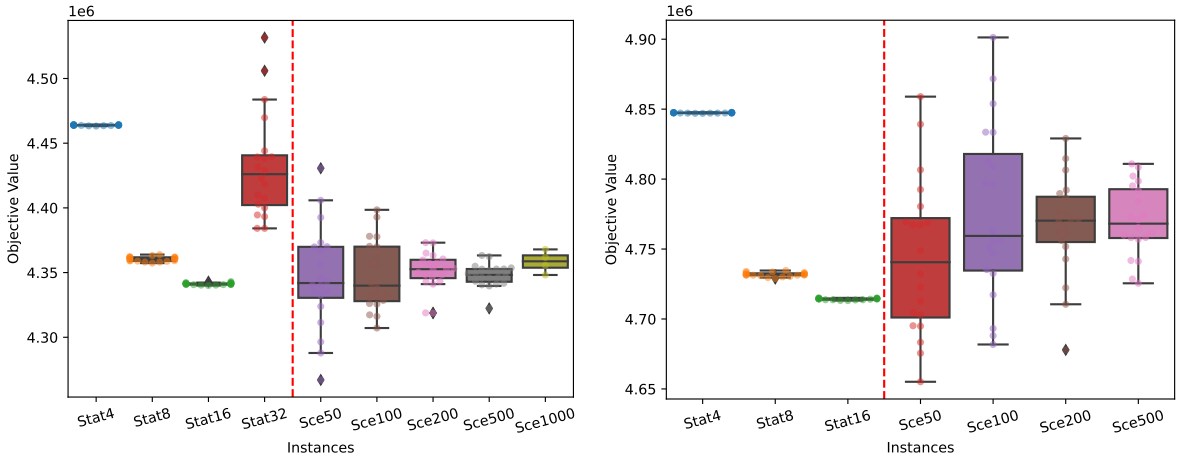
(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.2: Out-of-sample stability 200-Unit system



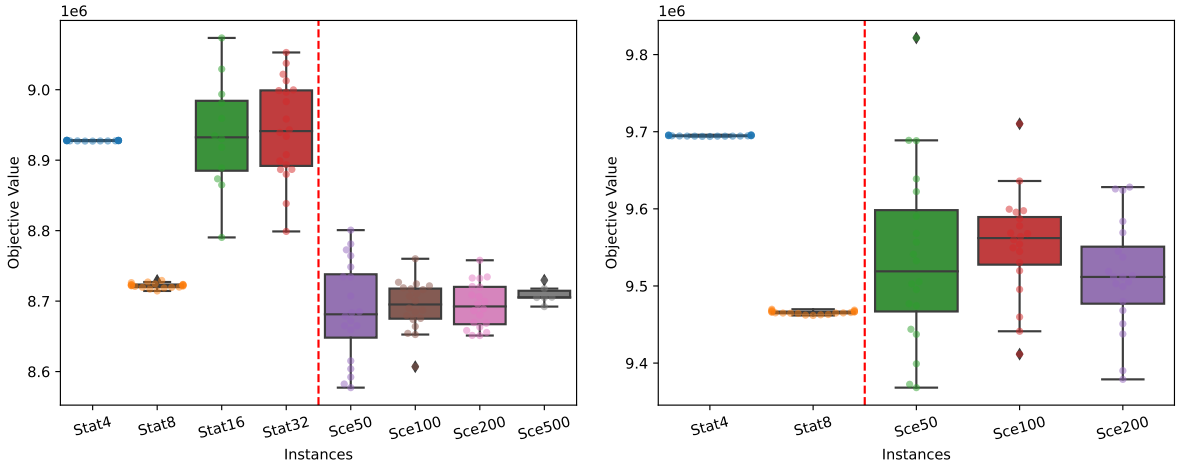
(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.3: Out-of-sample stability 400-Unit system



(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.4: In sample stability 100-Unit system

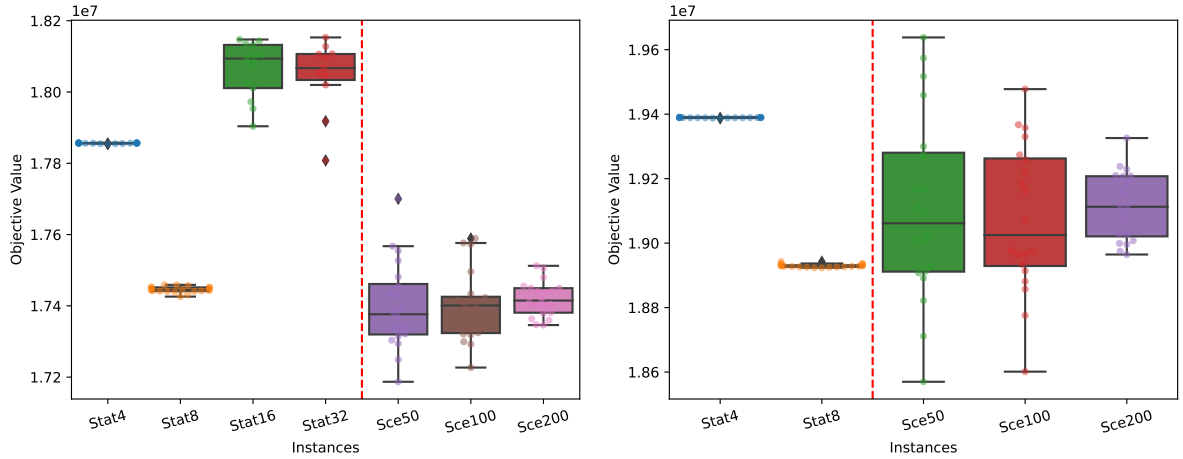


(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.5: In sample stability 200-Unit system

the contrary, the objective function of the *scenario-based* model is derived from samples of the residual demand. Thus, the sample size is directly related to how stable the objective function of the model is. Therefore, to present better in-sample stability, the scenario model should be solved with more than 500 scenarios for the case with $\sigma = 0.15\mu$. This behavior is more severe when increasing the variability to $\sigma = 0.3\mu$.

In general, both models presented similar out-of-sample stability when more breakpoints and scenarios were used, which means both models led to stable solutions to the actual problem. However, when observing the in-sample stability, the *statistical* model is more stable than the



(a) $\sigma = 0.15\mu$ (b) $\sigma = 0.30\mu$

Figure 4.6: In sample stability 400-Unit system

scenario-based model. Thus, the approximated expected value of the *scenario-based* model should not be used to measure the expected cost.

In this section, the objective value of the true value can also be compared using figures 4.1, 4.2, and 4.3. The *scenario-based* presents less expensive objectives with low and high variability in both cases. We attribute this behavior to the piece-wise approximations. Approximating the objective value could prevent finding some optimal solutions on the MILP formulation on the true objective value. This could be addressed by using more breakpoints. Nevertheless, solving instances with more breakpoints is hard and may require other optimization techniques, such as decomposition methods.

4.4 Conclusions

In this chapter, we performed a comparative analysis between the *statistical* and the *scenario-based* models under a two-stage stochastic programming framework. Within this framework, the first stage decisions corresponded to the commitment state of the generator units that remain fixed throughout the planning horizon. Subsequently, dispatch costs were computed upon observing the realizations of the random variables. In particular, for the *statistical* model, the expected cost was derived directly from the probability distribution of the residual demand, resulting in a nonlinear objective function. Thus, a piecewise linear approximation was used to

linearize the model. Conversely, the *scenario-based* model bases its expected cost on a representative sample of scenarios rather than a direct derivation from the probability distribution. For each scenario, an economic dispatch was performed using the committed units. Thus, the expected cost was calculated as the weighted sum between the probability of each scenario and its corresponding cost.

We studied the stability and computational complexity of both the *statistical* and *scenario-based* models. Our findings indicated a dependency on approximation parameters, such as the number of breakpoints and scenarios. Thus, as these parameters are increased, both models presented enhanced out-of-sample and in-sample stability. However, the stability of the *scenario-based* was significantly influenced by the sample size of the scenario set and the level of demand variability. The *scenario-based* model produced volatile in-sample stability results with a small scenario sample at high variability levels. This result suggests that the resulting objective value may not offer a reliable measure of the actual stochastic problem, especially under conditions of high uncertainty. This instability was not observed with the statistical model, which thus establishes itself as an alternative in situations of high variability.

Regarding the solving time, both models presented similar results. In particular, the *statistical* model can be solved in shorter times using a few breakpoints, which undermines the resulting approximation of the actual objective function. However, the proposed *statistical* approach can become a valid alternative for a future where there will be higher uncertainty on energy demand and supply, considering the high penetration of electric vehicles and RESs. Current scenario generation models will require more scenarios to obtain stable solutions and a longer computational time.

Future research should consider real power systems and larger sample sizes for the *scenario-based* model, which could be solved by applying the Benders decomposition method. In addition, including other technologies, such as batteries, could be informative regarding the differences between both models.

Chapter 5

Summary and future research

In this chapter, a summary of this dissertation is presented. We highlight the main contribution as well as the contributions of each chapter. In addition, we mention the limitations of this dissertation and future research.

The accelerating integration of RESs into the energy matrix has revolutionized the field of power system operations. However, this shift has brought a suite of complex challenges. Among these challenges, the uncertainty introduced by RESs into the power grid affects multiple decision-making processes. ISOs prepare the energy generation schedule by solving the UCP, a complex task influenced by random components such as demand and unstable energy production from renewable sources. Including these uncertainty elements is required to ensure the robustness and efficiency of the energy system.

A typical approach for handling uncertainty relies on the generation of multiple scenarios. Thus, the expected cost is optimized by generating a feasible commitment schedule that meets technical constraints considering all the scenarios. However, the need for additional scenarios grows accordingly with rising uncertainty levels. This escalation, in turn, amplifies the computational complexity of the problem. Such a surge could notably restrict the efficiency and practicality of the problem-solving process.

Considering these challenges, this dissertation proposed a statistical model that does not require scenarios. Instead, the expected cost was derived using the probability distribution of the residual demand. In addition, we proposed a mathematical formulation to solve the proposed model (Chapters 2 and 3). Finally, in Chapter 4, the proposed methodology was compared to the traditional scenario-based SUCP model.

Chapter 2 introduced this methodology, premised on the assumption that the units dispatch energy in an economic order. As such, a dispatching function was proposed using a two-stage optimization approach where the commitment decisions remain unchanged throughout the planning horizon. Subsequently, the expected cost of the dispatch function was derived, resulting in a nonlinear function that was linearized using an incremental mathematical model. Thus, the mathematical problem could be solved using commercial solvers. The precision of the piecewise linear approximation was assessed by comparing the resulting objective value with its analytical counterpart and employing a Monte Carlo simulation.

Experimental findings suggested that the linear approximation enables the resolution of real-sized problems with only a few breakpoints. Increasing the number of breakpoints did not notably enhance the accuracy of the linear approximation. Furthermore, in most experiments, the processing time fell within the range demanded by the ISOs. We also studied the reliability of the proposed method by analyzing the LOLP with higher levels of forecast error variability. For most of the experiments, the results showed a LOLP on average of less than 5%. One of the limitations of this chapter is the assumption of dispatching in an economic order. Generator units have ramping capabilities, which, in turn, prevent dispatching accordingly to this rule. However, the main methodology can still be used as it enables another perspective to solve stochastic optimization problems. We fixed this limitation in Chapter 3.

Ramping constraints ensure proper energy production between consecutive periods. Failing to account for this characteristic may yield infeasible commitment schedules in practice. Thus, we implemented ramping constraints in Chapter 3, resulting in a more realistic SUCP. Since the resulting model became more complex due to ramping constraints, several strategies were proposed to solve the model, namely, valid inequalities, sorting strategy, and heuristic. The first two strategies improve the linear relaxation of the mathematical model, resulting in a tighter version, which in turn speeds up solving the problem. In particular, the heuristic is used as an initial solution and fed into the solver to improve the incumbent solution. Thus, the proposed strategies helped to accelerate the optimization problem. We tested the strategies using different power systems data, usually used in the literature for benchmarking. In addition, we tested the proposed approach using a real power system from CAISO. Therefore, the model is

shown to be robust in terms of applicability. We noticed that indexing the generator units in an economic order results in a better linear relaxation, which leads to finding optimal solutions. Thus, a hybrid approach using a sorting strategy and an initial heuristic leads to a gap of less than 1%

One of the limitations of this work is that the sorting strategy was proved empirically. Thus, by proving mathematically, this property can be beneficial and implemented more broadly. Another limitation is that using an initial heuristic may lead the solver to get stuck in a particular node of the B&B algorithm, not finding an optimal solution within the required time.

In Chapter 4, we presented a comparative analysis between the statistical and scenario-based models using a two-stage stochastic programming modeling approach. Since the scenario-based model is commonly used in the literature, we compared both models to determine the benefits that the statistical model can present. We compared the computational complexity and stability using both methods.

The main difference between the models is that the statistical model derives the expected cost directly from the probability distribution of residual demand. In contrast, the scenario-based model calculates expected costs by considering a representative sample of scenarios and their corresponding costs, weighted by their probabilities.

We found that increasing the number of breakpoints and scenarios enhances their out-of-sample and in-sample stability. However, the stability of the scenario-based model is significantly influenced by the sample size and demand variability, leading to volatile results with a small scenario sample under high variability conditions. In contrast, the statistical model remains stable under such conditions, making it a reliable alternative in high variability situations. In terms of solving time, both models demonstrate similar performance. However, the statistical model can be solved faster with a few breakpoints, potentially compromising the accuracy of the objective function approximation. Thus, the statistical approach could be a valid alternative, considering the expected increase in uncertainty due to electric vehicles and renewable energy sources. At the same time, the scenario-based model may require larger sample sizes and longer computational times.

This dissertation has some limitations that should be mentioned. The main limitation is that the proposed statistical model does not include security constraints or transmission networks. These constraints ensure proper electricity dispatch. Nevertheless, accounting for those constraints would require modifying and generating a dispatching function for each network bus. This could be a challenging modeling task. Another limitation is that the model was not tested using real data. For example, under the scenario-based approach, one year of data could be considered as the scenarios without needing to sample from a distribution. This could be done on the statistical model by estimating an empirical distribution using available data.

Future research should consider the use of batteries in the statistical model. The transition to green energy production and dispatch is becoming well-established and requires batteries to replace the current fuel-based generator units. This brings more challenges, as modeling the functioning of batteries will require including other constraints in the UCP.

This dissertation presented a novel approach to solving stochastic programming problems, particularly for the SUCP. Thus, the modeling approach could be extended to other areas in the Operations Research field.

References

[1] Swasti R Khuntia, Bart W Tuinema, José L Rueda, and Mart AMM van der Meijden. Time-horizons in the planning and operation of transmission networks: an overview. *IET Generation, Transmission & Distribution*, 10(4):841–848, 2016.

[2] PG Nikhil and D Subhakar. Sizing and parametric analysis of a stand-alone photovoltaic power plant. *IEEE journal of photovoltaics*, 3(2):776–784, 2013.

[3] Hale Cetinay, Fernando A Kuipers, and A Nezih Guven. Optimal siting and sizing of wind farms. *Renewable Energy*, 101:51–58, 2017.

[4] Yuqing Yang, Stephen Bremner, Chris Menictas, and Merlinde Kay. Battery energy storage system size determination in renewable energy systems: A review. *Renewable and Sustainable Energy Reviews*, 91:109–125, 2018.

[5] Hyunchul Kim, Yasuhiro Hayashi, and Koichi Nara. An algorithm for thermal unit maintenance scheduling through combined use of ga, sa and ts. *IEEE Transactions on Power Systems*, 12(1):329–335, 1997.

[6] Amir Abiri-Jahromi, Mahmud Fotuhi-Firuzabad, and Masood Parvania. Optimized midterm preventive maintenance outage scheduling of thermal generating units. *IEEE Transactions on Power Systems*, 27(3):1354–1365, 2012.

[7] Qi Wang, Chunyu Zhang, Yi Ding, George Xydis, Jianhui Wang, and Jacob Østergaard. Review of real-time electricity markets for integrating distributed energy resources and demand response. *Applied Energy*, 138:695–706, 2015.

[8] Mohsen Khorasany, Yateendra Mishra, and Gerard Ledwich. Market framework for local energy trading: A review of potential designs and market clearing approaches. *IET Generation, Transmission & Distribution*, 12(22):5899–5908, 2018.

[9] Yonghong Chen, Aaron Casto, Fengyu Wang, Qianfan Wang, Xing Wang, and Jie Wan. Improving large scale day-ahead security constrained unit commitment performance. *IEEE Transactions on Power Systems*, 31(6):4732–4743, 2016.

[10] Yifeng Chen, Daming Chen, Pietro P Altermatt, Shu Zhang, Le Wang, Xueling Zhang, Jianmei Xu, Zhiqiang Feng, Hui Shen, and Pierre J Verlinden. Technology evolution of the photovoltaic industry: Learning from history and recent progress. *Progress in Photovoltaics: Research and Applications*, 2022.

[11] Mansaku Maeda and David Watts. The unnoticed impact of long-term cost information on wind farms' economic value in the usa.–a real option analysis. *Applied energy*, 241:540–547, 2019.

[12] Warren B Powell. A unified framework for stochastic optimization. *European Journal of Operational Research*, 275(3):795–821, 2019.

[13] Ivan Calero, Claudio A Canizares, Kankar Bhattacharya, and Ross Baldick. Duck-curve mitigation in power grids with high penetration of pv generation. *IEEE Transactions on Smart Grid*, 13(1):314–329, 2021.

[14] Moataz Sheha, Kasra Mohammadi, and Kody Powell. Solving the duck curve in a smart grid environment using a non-cooperative game theory and dynamic pricing profiles. *Energy Conversion and Management*, 220:113102, 2020.

[15] Cristobal Gallego-Castillo, Alvaro Cuerva-Tejero, and Oscar Lopez-Garcia. A review on the recent history of wind power ramp forecasting. *Renewable and Sustainable Energy Reviews*, 52:1148–1157, 2015.

[16] Brian Carlson, Yonghong Chen, Mingguo Hong, Roy Jones, Kevin Larson, Xingwang Ma, Peter Nieuwesteeg, Haili Song, Kimberly Sperry, Matthew Tackett, et al. Miso unlocks billions in savings through the application of operations research for energy and ancillary services markets. *Interfaces*, 42(1):58–73, 2012.

[17] Pascale Bendotti, Pierre Fouilhoux, and Cécile Rottner. On the complexity of the unit commitment problem. *Annals of Operations Research*, 274(1-2):119–130, 2019.

[18] Bernard Knueven, James Ostrowski, and Jean-Paul Watson. On mixed-integer programming formulations for the unit commitment problem. *INFORMS Journal on Computing*, 32(4):857–876, 2020.

[19] Menghan Zhang, Zhifang Yang, Wei Lin, Juan Yu, Wei Dai, and Ershun Du. Enhancing economics of power systems through fast unit commitment with high time resolution. *Applied Energy*, 281:116051, 2021.

[20] James Ostrowski, Miguel F Anjos, and Anthony Vannelli. Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Transactions on Power Systems*, 27(1):39–46, 2011.

[21] Germán Morales-España, Jesus M Latorre, and Andres Ramos. Tight and compact milp formulation of start-up and shut-down ramping in unit commitment. *IEEE Transactions on Power Systems*, 28(2):1288–1296, 2012.

[22] Bernard Knueven, James Ostrowski, and Jean-Paul Watson. A novel matching formulation for startup costs in unit commitment. *Mathematical Programming Computation*, pages 1–24, 2020.

[23] Samer Takriti, John R Birge, and Erik Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497–1508, 1996.

[24] Pierre Carpentier, Guy Gohen, J-C Culioli, and Arnaud Renaud. Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Transactions on Power Systems*, 11(2):1067–1073, 1996.

[25] Jikai Zou, Shabbir Ahmed, and Xu Andy Sun. Multistage stochastic unit commitment using stochastic dual dynamic integer programming. *IEEE Transactions on Power Systems*, 34(3):1814–1823, 2018.

[26] Chunheng Wang and Yong Fu. Fully parallel stochastic security-constrained unit commitment. *IEEE Transactions on power systems*, 31(5):3561–3571, 2015.

[27] Murilo Reolon Scuzziato, Erlon Cristian Finardi, and Antonio Frangioni. Comparing spatial and scenario decomposition for stochastic hydrothermal unit commitment problems. *IEEE Transactions on Sustainable Energy*, 9(3):1307–1317, 2017.

[28] Takayuki Shiina and John R Birge. Stochastic unit commitment problem. *International Transactions in Operational Research*, 11(1):19–32, 2004.

[29] Changhyeok Lee, Cong Liu, Sanjay Mehrotra, and Mohammad Shahidehpour. Modeling transmission line constraints in two-stage robust unit commitment problem. *IEEE Transactions on Power Systems*, 29(3):1221–1231, 2013.

[30] Qipeng P Zheng, Jianhui Wang, Panos M Pardalos, and Yongpei Guan. A decomposition approach to the two-stage stochastic unit commitment problem. *Annals of Operations Research*, 210(1):387–410, 2013.

[31] Ping Che, Lixin Tang, and Jianhui Wang. Two-stage minimax stochastic unit commitment. *IET Generation, Transmission & Distribution*, 12(4):947–956, 2018.

[32] Sarah M Ryan, Roger J-B Wets, David L Woodruff, César Silva-Monroy, and Jean-Paul Watson. Toward scalable, parallel progressive hedging for stochastic unit commitment. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.

[33] Kwok Cheung, Dinakar Gade, César Silva-Monroy, Sarah M Ryan, Jean-Paul Watson, Roger J-B Wets, and David L Woodruff. Toward scalable stochastic unit commitment. *Energy Systems*, 6(3):417–438, 2015.

[34] U Aytun Ozturk, Mainak Mazumdar, and Bryan A Norman. A solution to the stochastic unit commitment problem using chance constrained programming. *IEEE Transactions on Power Systems*, 19(3):1589–1598, 2004.

[35] Sajjad Abedi, Gholam Hossein Riahy, Seyed Hossein Hosseinian, and Arash Alimardani. Risk-constrained unit commitment of power system incorporating pv and wind farms. *International Scholarly Research Notices*, 2011, 2011.

[36] Long Zhao and Bo Zeng. Robust unit commitment problem with demand response and wind energy. In *2012 IEEE power and energy society general meeting*, pages 1–8. IEEE, 2012.

[37] Narges Kazemzadeh, Sarah M Ryan, and Mahdi Hamzeei. Robust optimization vs. stochastic programming incorporating risk measures for unit commitment with uncertain variable renewable generation. *Energy Systems*, 10(3):517–541, 2019.

[38] Saleh Y Abujarad, Mohammad Wazir Mustafa, and Jasrul Jamani Jamian. Recent approaches of unit commitment in the presence of intermittent renewable energy resources: A review. *Renewable and Sustainable Energy Reviews*, 70:215–223, 2017.

[39] Jianhui Wang, Audun Botterud, Vladimiro Miranda, Cláudio Monteiro, and Gerald Sheble. Impact of wind power forecasting on unit commitment and dispatch. In *Proc. 8th Int. Workshop Large-Scale Integration of Wind Power into Power Systems*, pages 1–8, 2009.

[40] Aidan Tuohy, Peter Meibom, Eleanor Denny, and Mark O'Malley. Unit commitment for systems with significant wind penetration. *IEEE Transactions on power systems*, 24(2):592–601, 2009.

[41] Tim Schulze and Ken McKinnon. The value of stochastic programming in day-ahead and intra-day generation unit commitment. *Energy*, 101:592–605, 2016.

[42] Qipeng P Zheng, Jianhui Wang, and Andrew L Liu. Stochastic optimization for unit commitment—a review. *IEEE Transactions on Power Systems*, 30(4):1913–1924, 2014.

[43] J Wang, A Botterud, R Bessa, H Keko, L Carvalho, D Issicaba, J Sumaili, and V Miranda. Wind power forecasting uncertainty and unit commitment. *Applied Energy*, 88(11):4014–4023, 2011.

[44] Robin Broder Hytowitz and Kory W Hedman. Managing solar uncertainty in microgrid systems with stochastic unit commitment. *Electric Power Systems Research*, 119:111–118, 2015.

[45] Anthony Papavasiliou and Shmuel S Oren. Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. *Operations Research*, 61(3):578–592, 2013.

[46] Anthony Papavasiliou, Shmuel S Oren, and Barry Rountree. Applying high performance computing to transmission-constrained stochastic unit commitment for renewable energy integration. *IEEE Transactions on Power Systems*, 30(3):1109–1120, 2014.

[47] Ana Viana and João Pedro Pedroso. A new milp-based approach for unit commitment in power production planning. *International Journal of Electrical Power & Energy Systems*, 44(1):997–1005, 2013.

[48] George B Dantzig. On the significance of solving linear programming problems with some integer variables. *Econometrica, Journal of the Econometric Society*, pages 30–44, 1960.

[49] Spyros A Kazarlis, AG Bakirtzis, and Vassilios Petridis. A genetic algorithm solution to the unit commitment problem. *IEEE transactions on power systems*, 11(1):83–92, 1996.

[50] Jorge Valenzuela, Mainak Mazumdar, and Anoop Kapoor. Influence of temperature and load forecast uncertainty on estimates of power generation production costs. *IEEE Transactions on Power Systems*, 15(2):668–674, 2000.

[51] Miguel Asensio and Javier Contreras. Stochastic unit commitment in isolated systems with renewable penetration under cvar assessment. *IEEE Transactions on Smart Grid*, 7(3):1356–1367, 2015.

[52] Zhi Chen, Lei Wu, and Mohammad Shahidehpour. Effective load carrying capability evaluation of renewable energy via stochastic long-term hourly based scuc. *IEEE Transactions on Sustainable Energy*, 6(1):188–197, 2014.

[53] Zhi Wu, Pingliang Zeng, Xiao-Ping Zhang, and Qinyong Zhou. A solution to the chance-constrained two-stage stochastic program for unit commitment with wind energy integration. *IEEE Transactions on Power Systems*, 31(6):4185–4196, 2016.

[54] GJ Osório, JM Lujano-Rojas, JCO Matias, and JPS Catalão. A new scenario generation-based method to solve the unit commitment problem with high penetration of renewable energies. *International Journal of Electrical Power & Energy Systems*, 64:1063–1072, 2015.

[55] Germán Morales-España, Álvaro Lorca, and Mathijs M de Weerdt. Robust unit commitment with dispatchable wind power. *Electric Power Systems Research*, 155:58–66, 2018.

[56] Yuefang Du, Yuanzheng Li, Chao Duan, Hoay Beng Gooi, and Lin Jiang. Adjustable uncertainty set constrained unit commitment with operation risk reduced through demand response. *IEEE Transactions on Industrial Informatics*, 17(2):1154–1165, 2020.

[57] Charles F. Jekel and Gerhard Venter. *pwlf: A Python Library for Fitting 1D Continuous Piecewise Linear Functions*, 2019.

[58] Michal Kaut and Stein W. Wallace. Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271, 2007.

[59] Ricardo Bessa, Carlos Moreira, Bernardo Silva, and Manuel Matos. Handling renewable energy variability and uncertainty in power systems operation. *Wiley Interdisciplinary Reviews: Energy and Environment*, 3(2):156–178, 2014.

[60] Juan Ma, Vera Silva, Régine Belhomme, Daniel S Kirschen, and Luis F Ochoa. Evaluating and planning flexibility in sustainable power systems. In *2013 IEEE power & energy society general meeting*, pages 1–11. IEEE, 2013.

[61] Wim Van Ackooij, Irene Danti Lopez, Antonio Frangioni, Fabrizio Lacalandra, and Milad Tahanan. Large-scale unit commitment under uncertainty: an updated literature survey. *Annals of Operations Research*, 271(1):11–85, 2018.

[62] Yuping Huang, Panos M Pardalos, Qipeng P Zheng, Yuping Huang, Panos M Pardalos, and Qipeng P Zheng. Two-stage stochastic programming models and algorithms. *Electrical Power Unit Commitment: Deterministic and Two-Stage Stochastic Programming Models and Algorithms*, pages 49–86, 2017.

[63] Semih Atakan, Harsha Gangammanavar, and Suvrajeet Sen. Towards a sustainable power grid: Stochastic hierarchical planning for high renewable integration. *European Journal of Operational Research*, 302(1):381–391, 2022.

[64] Antonio Frangioni and Claudio Gentile. Solving nonlinear single-unit commitment problems with ramping constraints. *Operations Research*, 54(4):767–775, 2006.

[65] Aiying Rong, Henri Hakonen, and Risto Lahdelma. A variant of the dynamic programming algorithm for unit commitment of combined heat and power systems. *European Journal of Operational Research*, 190(3):741–755, 2008.

[66] Tomonobu Senjyu, Tsukasa Miyagi, Ahmed Yousuf Saber, Naomitsu Urasaki, and Toshihisa Funabashi. Emerging solution of large-scale unit commitment problem by stochastic priority list. *Electric Power Systems Research*, 76(5):283–292, 2006.

[67] Chuan-Ping Cheng, Chih-Wen Liu, and Chun-Chang Liu. Unit commitment by lagrangian relaxation and genetic algorithms. *IEEE transactions on power systems*, 15(2):707–714, 2000.

[68] Tao Li and Mohammad Shahidehpour. Price-based unit commitment: A case of lagrangian relaxation versus mixed integer programming. *IEEE transactions on power systems*, 20(4):2015–2025, 2005.

[69] Rimmi Anand, Divya Aggarwal, and Vijay Kumar. A comparative analysis of optimization solvers. *Journal of Statistics and Management Systems*, 20(4):623–635, 2017.

[70] E Robert Bixby, Mary Fenelon, Zonghao Gu, Ed Rothberg, and Roland Wunderling. Mip: Theory and practice—closing the gap. In *IFIP Conference on System Modeling and Optimization*, pages 19–49. Springer, 1999.

[71] Germán Morales-España, Jesus M Latorre, and Andres Ramos. Tight and compact milp formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 28(4):4897–4908, 2013.

[72] Miguel Carrión and José M Arroyo. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Transactions on power systems*, 21(3):1371–1378, 2006.

[73] Deepak Rajan, Samer Takriti, et al. Minimum up/down polytopes of the unit commitment problem with start-up costs. *IBM Res. Rep*, 23628:1–14, 2005.

[74] Germán Morales-España, Claudio Gentile, and Andres Ramos. Tight mip formulations of the power-based unit commitment problem. *OR spectrum*, 37(4):929–950, 2015.

[75] Claudio Gentile, Germán Morales-Espana, and Andres Ramos. A tight mip formulation of the unit commitment problem with start-up and shut-down constraints. *EURO Journal on Computational Optimization*, 5(1-2):177–201, 2017.

[76] Kai Pan, Ming Zhao, Chung-Lun Li, and Feng Qiu. A polyhedral study on fuel-constrained unit commitment. *INFORMS Journal on Computing*, 34(6):3309–3324, 2022.

[77] Ruiwei Jiang, Yongpei Guan, and Jean-Paul Watson. Cutting planes for the multistage stochastic unit commitment problem. *Mathematical Programming*, 157(1):121–151, 2016.

[78] Andrea Fusco, Domenico Gioffrè, Alessandro Francesco Castelli, Cristian Bovo, and Emanuele Martelli. A multi-stage stochastic programming model for the unit commitment of conventional and virtual power plants bidding in the day-ahead and ancillary services markets. *Applied Energy*, 336:120739, 2023.

[79] Juan Pablo Vielma and George L Nemhauser. Modeling disjunctive constraints with a logarithmic number of binary variables and constraints. *Mathematical Programming*, 128:49–72, 2011.

[80] Saeed Moradi, Sohrab Khanmohammadi, Mehrdad Tarafdar Hagh, and Behnam Mohammadi-ivatloo. A semi-analytical non-iterative primary approach based on priority list to solve unit commitment problem. *Energy*, 88:244–259, 2015.

[81] John E Beasley. Or-library: distributing test problems by electronic mail. *Journal of the operational research society*, 41(11):1069–1072, 1990.

[82] DA Tejada-Arango, S Lumbreras, P Sanchez-Martin, and A Ramos. Which unit-commitment formulation is best? a systematic comparison. *IEEE Transactions on Power Systems*, 35:2926–2936, 2019.

[83] Yuping Huang, Panos M Pardalos, and Qipeng P Zheng. Deterministic unit commitment models and algorithms. In *Electrical Power Unit Commitment*, pages 11–47. Springer, 2017.

[84] Pablo A Ruiz, C Russ Philbrick, Eugene Zak, Kwok W Cheung, and Peter W Sauer. Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems*, 24(2):642–651, 2009.

[85] Francois Bouffard and Francisco D Galiana. Stochastic security for operations planning with significant wind power generation. In *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–11. IEEE, 2008.

[86] Lei Wu, Mohammad Shahidehpour, and Tao Li. Stochastic security-constrained unit commitment. *IEEE Transactions on power systems*, 22(2):800–811, 2007.

[87] ZQ Xie, TY Ji, MS Li, and QH Wu. Quasi-monte carlo based probabilistic optimal power flow considering the correlation of wind speeds using copula function. *IEEE Transactions on Power Systems*, 33(2):2239–2247, 2017.

[88] Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on optimization*, 12(2):479–502, 2002.

[89] Qianfan Wang, Yongpei Guan, and Jianhui Wang. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE transactions on power systems*, 27(1):206–215, 2011.

[90] Peng Xiong and Panida Jirutitijaroen. A stochastic optimization formulation of unit commitment with reliability constraints. *IEEE Transactions on Smart Grid*, 4(4):2200–2208, 2013.

[91] Yonghan Feng and Sarah M Ryan. Solution sensitivity-based scenario reduction for stochastic unit commitment. *Computational Management Science*, 13(1):29–62, 2016.

[92] CC Carøe, A Ruszczynski, and R Schultz. Unit commitment under uncertainty via two-stage stochastic programming. In *Proceedings of NOAS*, volume 97, pages 21–30, 1997.

[93] Jiadong Wang, Jianhui Wang, Cong Liu, and Juan P Ruiz. Stochastic unit commitment with sub-hourly dispatch constraints. *Applied energy*, 105:418–422, 2013.

[94] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[95] Martin Håberg. Fundamentals and recent developments in stochastic unit commitment. *International Journal of Electrical Power & Energy Systems*, 109:38–48, 2019.