Accepted Manuscript

Risk-averse formulations and methods for a virtual power plant

Ricardo M. Lima, Antonio J. Conejo, Sabique Langodan,
Ibrahim Hoteit, Omar M. Knio

PII: S0305-0548(17)30307-6
DOI: [10.1016/j.cor.2017.12.007](https://doi.org/10.1016/j.cor.2017.12.007)
Reference: CAOR 4371

To appear in: *Computers and Operations Research*

Received date: 14 December 2016
Revised date: 24 October 2017
Accepted date: 8 December 2017

Please cite this article as: Ricardo M. Lima, Antonio J. Conejo, Sabique Langodan, Ibrahim Hoteit, Omar M. Knio, Risk-averse formulations and methods for a virtual power plant, *Computers and Operations Research* (2017), doi: [10.1016/j.cor.2017.12.007](https://doi.org/10.1016/j.cor.2017.12.007)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Highlights

- Risk-neutral and risk-averse stochastic programming formulations are considered.
- Implementation of decomposition methods to handle the CVaR.
- Wind ensembles used to characterize the wind speed uncertainty.
- Extensive computational results for performance and risk management analysis.
- The parallel solution of the sub-problems is paramount to obtain efficient methods

Risk-averse formulations and methods for a virtual power plant

Ricardo M. Lima^{a,*}, Antonio J. Conejo^c, Sabique Langodan^b, Ibrahim Hoteit^b, Omar M. Knio^a

^a*Computer, Electrical and Mathematical Sciences & Engineering Division*

^b*Physical Science and Engineering Division,*

King Abdullah University of Science and Technology, (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia

^c*Integrated Systems Engineering-Electrical and Computer Engineering, The Ohio State University, OH, USA*

Abstract

In this paper we address the optimal operation of a virtual power plant using stochastic programming. We consider one risk-neutral and two risk-averse formulations that rely on the conditional value at risk. To handle large-scale problems, we implement two decomposition methods with variants using single- and multiple-cuts. We propose the utilization of wind ensembles obtained from the European Centre for Medium Range Weather Forecasts (ECMWF) to quantify the uncertainty of the wind forecast. We present detailed results relative to the computational performance of the risk-averse formulations, the decomposition methods, and risk management and sensitivities analysis as a function of the number of scenarios and risk parameters. The implementation of the two decomposition methods relies on the parallel solution of subproblems, which turns out to be paramount for computational efficiency. The results show that one of the two decomposition methods is the most efficient.

Keywords: Optimization under uncertainty; stochastic programming; conditional value at risk; energy; virtual power plant.

1. Introduction

Risk-averse methodologies enhance the capabilities of general stochastic programming by introducing a risk measure in the formulation of the optimal operation of virtual power plants. The risk measure provides additional means to quantify the uncertainty, and to drive the solution methods to solutions with a specified level of risk. The risk measure can be incorporated in the objective function, or as a constraint to limit the risk associated with the solutions (Kall and Mayer, 2011).

Single-stage risk-averse stochastic programming formulations are described in detail in Kall and Mayer (2011). Two-stage risk-averse stochastic programming problems are more complex than the corresponding risk-neutral formulations, from the point of view of the formulation

*Corresponding author.

Email addresses: `ricardo.lima@kaust.edu.sa` (Ricardo M. Lima), `conejonavarro.1@osu.edu` (Antonio J. Conejo), `sabique.langodan@kaust.edu.sa` (Sabique Langodan), `ibrahim.hoteit@kaust.edu.sa` (Ibrahim Hoteit), `omar.knio@kaust.edu.sa` (Omar M. Knio)

and size. Therefore, to solve large-scale two-stage risk-averse stochastic programming problems, decomposition methods are crucial. However, although two-stage risk-averse stochastic programming formulations are well established, the corresponding decomposition methods have only been investigated recently (Kall and Mayer, 2011; Noyan, 2012)

In general, risk-neutral stochastic programming formulations use the expectation of a random variable as the measure of quality, which for example, translates into the minimization of the expected cost, or maximization of the expected profit. These formulations and the corresponding properties and solution methods are described in Birge and Louveaux (2011) and Kall and Mayer (2011).

Among the several risk measures studied in the literature – variance, shortfall probability, expected shortage, value at risk, stochastic dominance – the conditional value at risk (CVaR) (Rockafellar and Uryasev, 2000) has relevant properties, which make CVaR a convenient risk measure for optimization models. The properties of the CVaR and its integration with optimization problems have been studied and compared with the Value at Risk (VaR) by several authors (Pflug, 2000; Rockafellar, 2007; Rockafellar and Uryasev, 2000; Sarykalin et al., 2008; Shapiro et al., 2009).

The introduction of the risk measure in the formulation, requires an adaptation of the decomposition methods developed for risk-neutral formulations, such as the L-Shaped method (Birge and Louveaux, 1988; Van Slyke and Wets, 1969). The pioneering work of Schultz and Tiedemann (2006) implemented a Lagrangian relaxation to handle the integer variables in the recourse problems of a risk-averse stochastic programming problem involving the CVaR. Ahmed (2006) proposed decomposition methods based on extensions of the L-Shaped method to handle two convex risk measures. The author focuses on methods to address formulations with the mean-absolute semi-deviation or the mean quantile deviation. In fact, one of the decomposition methods proposed in Ahmed (2006) is implemented in this work. Fabian (2008) developed a decomposition method based on level-type methods (Lemaréchal et al., 1995) and the polyhedral representation proposed by Künzi-Bay and Mayer (2006) to handle the CVaR in the objective function and constraints. Miller and Ruszczyński (2011) focused on a specific type of problem with unresolved uncertainty after the second-stage decisions are made. Noyan (2012) implemented and studied variants of two L-Shaped decomposition methods for a risk-averse stochastic programming problem with the CVaR measure. In addition, Noyan (2012) proposed two expressions to evaluate the Value of Stochastic Solution (VSS) and Expected Value of Perfect Information (EVPI) for risk-averse stochastic programming problems. Fabian et al. (2015) proposed a decomposition method that employs aggregate models, which do not use all the information from the solution of the recourse problems, and is based on bundle level methods (de Oliveira and Sagastizbal, 2014). Our literature review shows that there have been two parallel directions on the development of decomposition methods to handle risk measures, one based on the work of Ahmed (2006), and the other based on the work of Fabian (2008).

Recently, Homem-de Mello and Pagnoncelli (2016) and Zhang et al. (2016) have addressed risk-averse multiple stage stochastic programming using CVaR as risk measures. Zhang et al. (2016) have extended the decomposition methods proposed by Ahmed (2006) and Noyan (2012)

to deal with multiple stages.

In this work, we address the risk-averse optimal operation of a virtual power plant (VPP) and the efficient implementation of decomposition methods to solve large-scale problems. A VPP consists of a group of units that are dispersed, but they are coordinated within the group in order to interact with the exterior as a single unit. Additional characteristics of VPPs can be found in Lombardi et al. (2009) and Morales et al. (2014). In this work, the VPP consists of a flexible but comparatively expensive thermal unit, a stochastic weather-dependent but comparatively cheap wind power plant and a pumped-storage hydroelectric plant. This is a versatile and flexible VPP because it provides storage, two dispatchable units, a renewable source, and interactions between the pumped-storage hydroelectric plant and the other units, and between the pumped-storage hydroelectric plant and the market. Such VPP allows us to efficiently illustrate the features of the proposed solution algorithms. Needless to say, alternative VPP configurations can be readily analyzed using the proposed algorithms. The coordination of the VPP involves the integrated scheduling of the units, and the interaction with the electricity market as a single entity.

Recently, the optimal operation of VPPs has been the subject of several works. The optimization framework for the optimal operation of VPPs, naturally, involves optimization under uncertainty using stochastic programming or robust optimization (Ben-Tal and Nemirovski, 1998; Bertsimas et al., 2011). Some authors have developed two stage risk-neutral stochastic programming formulations (Pandzic et al., 2013a,b), while others proposed two stage risk-averse stochastic programming using the CVaR as a risk measure (Dabbagh and Sheikh-El-Eslami, 2015, 2016; Kardakos et al., 2016; Moazeni et al., 2015; Moghaddam et al., 2013; Tajeddini et al., 2014). Lima et al. (2015) proposed a two-stage adaptive robust optimization approach, and Shabanzadeh et al. (2015) a robust optimization based approach.

The models of the VPPs reported in the literature exhibit different features, at the level of the composition of the VPP, and also at the level of the interaction with the electricity market. Regarding the composition of the VPPs, most authors consider a renewable generation source, a dispatchable unit, and a storage unit that can be a cluster of batteries or a pumped storage hydro unit. In terms of interactions with the electricity market, several features have been considered, namely: a) utilization of forward contracts to hedge against the volatility of the electricity market (Lima et al., 2015; Pandzic et al., 2013a; Shabanzadeh et al., 2015); b) offers to the day-ahead market (Dabbagh and Sheikh-El-Eslami, 2015, 2016; Kardakos et al., 2016; Moghaddam et al., 2013; Pandzic et al., 2013b; Tajeddini et al., 2014; Zamani et al., 2016); c) offers to the balancing market (Dabbagh and Sheikh-El-Eslami, 2015, 2016; Kardakos et al., 2016; Tajeddini et al., 2014); d) offers of spinning reserves (Dabbagh and Sheikh-El-Eslami, 2016); and e) strategic offering considering the market clearing and other market players' strategies (Kardakos et al., 2016). The time horizon considered in most works covers 24 hours, with the exception of the models that include forward contracts, where the time horizon is 168 hours.

In the works cited above related with the VPPs and stochastic programming, the number of scenarios postulated is in the range of 25 to 1000 scenarios. In addition, in those works that have used risk-averse stochastic programming formulations, the utilization of decomposition

algorithms is not reported. In general, the utilization of decomposition methods to solve risk-averse stochastic programming problems in energy systems has been limited, and to the best of our knowledge they have not been applied to the optimal operation of VPPs.

This paper has five main contributions: 1) a detailed comparison of two formulations for risk-averse stochastic programming problems; 2) efficient implementation of decomposition methods to solve large scale problems; whereby a significant computational improvement is obtained; 3) utilization of a wind ensemble to forecast and quantify uncertainty in the wind speed, establishing the link between ensemble member and a scenario used in stochastic programming; 4) detailed analysis of the computational performance of the extensive form and decomposition methods; relevant insights are provided regarding the use of single- and multiple-cuts in the decomposition methods; 5) sensitivity and risk management analysis as a function of the parameters of the two studied risk-averse formulations.

The structure of the paper is as follows. In Section 2, we start by defining the problem statement and the main assumptions of the modeling of the VPP operations. In Section 3, we present the generic formulation of the problem. In Section 4, we describe the risk-neutral and risk-averse stochastic programming formulations. In Section 5, we derive and present the decomposition methods. In Section 6, we report the detailed results of the computational performance of the formulations and decomposition methods, and sensitivity and risk analysis. We focus on the main results and conclusions in Section 7. In the appendices, we provide the model and data that describes the VPP, and the Auto Regressive Integrated Moving Average (ARIMA) models developed to forecast the electricity prices.

2. Problem statement

In this work, we address the optimization of the scheduling operations and electricity market involvement of a VPP that comprises of a thermal unit, a pumped storage hydro unit, and a wind farm. The goal is to determine the energy output of the thermal unit, the energy output or input of the hydro unit subject to the feasible region of operation of these units, and the energy sold or bought through contracts or the market, so that the operational profit is maximized. The operational profit results from the sale of electricity to the pool of the market, and from energy sold through forward contracts minus the costs of operation of the thermal unit, and cost of electricity bought by contract or in the pool.

The operational region of the thermal unit is basically defined by the minimum up and down times and bounds on the power output, while the region of the hydro unit is constrained by the water mass balances, bounds of the reservoir water levels, and the energy generation and consumption functions. We assume that the system can operate as a virtual power plant, which means that the hydro unit can use the energy from the wind farm and from the thermal unit to pump back water to the upper reservoir; and that there are no operational costs associated with the wind farm and hydro plant.

We consider that the wind power generated by the wind farm and electricity prices in the pool are given as random variables. We rely on a 51-member wind ensemble generated by an

atmospheric circulation model to forecast the wind speed, and on forecasts of the electricity prices generated from specific ARIMA models.

The forward contracts are given by a fixed setup, which means that they are defined by a curve of blocks of constant energy and price.

The time horizon considered is one week with a time resolution of one hour. The time resolution is applied to the unit operations but also to the electricity market involvement in the pool. The decision framework involves two stages of decision, the first before the beginning of the time horizon, and the second-stage during the time horizon. Defined in the first-stage are the time periods that the thermal unit will be generating energy, and consequently the periods where the startups and shutdown occur, and the decisions related to the contracts. In the second-stage, the decisions are associated with the dispatch of the thermal and hydro units, and the involvement with the electricity market pool.

3. Deterministic model

A general formulation for the deterministic optimization problem described in the problem statement can be expressed as:

$$\begin{aligned}
 & \max_{x,z,y} \quad c^T x + \bar{c}^T z + \tilde{c}^T y \\
 & s.t. \quad Ax + Bz \leq b \\
 & \quad \quad Cz + Dy \leq d \\
 & \quad \quad Fy + Gx = h \\
 & \quad \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, y \in \mathbb{R}_+^{n_3},
 \end{aligned} \tag{1}$$

where the goal of problem (1) is to maximize the profit defined as $c^T x + \bar{c}^T z + \tilde{c}^T y$, $\mathbb{B} := \{0, 1\}$, x , z and y are vectors with dimensions n_1 , n_2 , and n_3 , respectively, $c \in \mathbb{R}^{n_1}$, $\bar{c} \in \mathbb{R}^{n_2}$, and $\tilde{c} \in \mathbb{R}^{n_3}$ denote deterministic vectors with known parameters, $A \in \mathbb{R}^{m_1 \times n_1}$, $B \in \mathbb{R}^{m_1 \times n_2}$, $C \in \mathbb{R}^{m_2 \times n_2}$, $D \in \mathbb{R}^{m_2 \times n_3}$, $F \in \mathbb{R}^{m_3 \times n_3}$, $G \in \mathbb{R}^{m_3 \times n_1}$ are deterministic matrices with known coefficients, and $b \in \mathbb{R}^{m_1}$, $d \in \mathbb{R}^{m_2}$, $h \in \mathbb{R}^{m_3}$ are deterministic vectors with known parameters. The relation between the general formulation in (1) and the specific model of the VPP is made in the remaining of this section. In addition, we further refer the reader to Appendix A, where we present and describe the constraints and data of the VPP.

In (1), vector x represents the power sold and bought through contracts, and z includes the binary variables associated with the discrete selection of the blocks of constant power and price of the contracts offered to the VPP, plus the binary variables that define the startups, shutdowns, and on/off status of the thermal unit. Vector y denotes the continuous variables that encompass the power output of the thermal unit, the power output of the hydro unit, the power consumption of the hydro unit, the turbined flow of the hydro unit, the pumped flow of the hydro unit, the volume of water in the reservoir of the hydro unit, and the power sold and bought in the pool of the electricity market.

The vectors c and $\bar{c}$ denote the fixed setup of blocks of energy and price of the contracts and fixed costs of the startup, shutdown, and generation of electricity of the thermal unit. The vector

$\tilde{c}$ denotes the revenues and costs of selling and buying electricity in the pool, and the variable costs of generation of electricity of the thermal unit. We represent the wind power by vector h , and the remaining parameters in (1) define the operation of the generation units, namely efficiencies of generation and consumption of energy by the hydro unit, maximum capacity of the reservoir of the hydro unit, maximum turbined and pumped flow of water, bounds on ramp rates of the thermal unit, minimum up and down times of the thermal unit, and minimum and maximum generation capacity of the thermal unit. See Appendix A for a full list of the parameters.

The first constraint in (1) models the choice of contracts and the region of operation of the thermal unit in terms of the minimum up and down time, hot and cold startup costs, shutdown costs, and logical relations between the binary variables that define the startup, shutdown, and on/off status. These constraints do not involve the continuous variables of the thermal unit, neither the variables of the hydro unit. The second and third constraints in (1) define the regions of operation of the thermal and hydro unit. These constraints encompass the bounds on the power output and the modeling of the generation cost of the thermal unit, the water mass balances for the reservoir and the power generation and consumption functions of the hydro unit.

In this work, we are concerned with risk-averse stochastic versions of the problem (1), where h and $\tilde{c}$ are the random variables corresponding to the wind power and the electricity prices in the pool, respectively.

4. Stochastic models

In this section, we extend the deterministic problem (1) to consider uncertainty in the parameters h and $\tilde{c}$. This optimization problem under uncertainty is solved using stochastic programming approaches. We consider three formulations: 1) a risk-neutral formulation; 2) a risk-averse formulation that maximizes a combination of expectation and risk measure; 3) a risk-averse formulation that minimizes a risk measure. In terms of solution methods, we consider the direct solution of the full models, and three decomposition methods based on Benders Decomposition (Benders, 1962).

The considered optimization problems are typical two-stage problems, with first- and second-stage variables. The first-stage variables correspond to here and now decisions (Birge and Louveaux, 2011), which are invariant with the realization of the random variables. In the studied problem, the first-stage variables are associated with the decisions made before the start of the week, namely those pertaining to contracts and to the startups and shutdowns of the thermal unit, which are denoted by x and z in (1). This means that the first-stage variables are fixed during the week. The second-stage variables, called wait and see decisions (Birge and Louveaux, 2011) depend on the realizations of the wind power and electricity prices. These variables allow the decision maker to adopt actions after the realization of the uncertain parameters in order to maximize the objective function. The wait and see decisions of the VPP are the power output of the thermal unit, the power output of the hydro unit, the power

consumption of the hydro unit, the turbined and pumped flow of the hydro unit, the volume of water in the reservoir of the hydro unit, and the power sold and bought in the pool of the electricity market; these decisions correspond to the variable vector y in (1).

4.1. Stochastic programming formulation

We start by defining a random vector $\xi : \Omega \rightarrow \mathbb{R}^r$ associated with a probability space $(\Omega, \mathcal{F}, P)$. In this space, Ω is the set of all possible outcomes, $\mathcal{F}$ is a set of events, and $P : \mathcal{F} \rightarrow [0, 1]$ is a probability measure. In the context of problem (1), we consider the two random vectors h and $\tilde{c}$ as $h(\xi)$ and $\tilde{c}(\xi)$ defined as $h : \mathbb{R}^r \rightarrow \mathbb{R}^{m_3}$ and $\tilde{c} : \mathbb{R}^r \rightarrow \mathbb{R}^{n_3}$, respectively. In stochastic programming, it is common to assume that the random vector ξ is described by a discrete probability measure with a cumulative distributed function $F_\xi(v) = P(\xi \leq v)$. Therefore, the vector ξ has a finite support set $\mathcal{S} \subset \mathbb{R}$, which is the smallest closed set such that $P(\xi \in \mathcal{S}) = 1$. The elements of $\mathcal{S}$ are typically denoted by scenarios with $\xi_s, \forall s \in \mathcal{S}$ with probabilities p_s such that $\sum_{s \in \mathcal{S}} p_s = 1$. To simplify the nomenclature, we define the realizations of the random vectors as $h_s := h(\xi_s)$ and $\tilde{c}_s := \tilde{c}(\xi_s)$.

We define the function $f : \mathbb{R}^{n_1} \times \mathbb{B}^{n_2} \times \mathbb{R}^{n_3} \times \mathcal{S} \rightarrow \mathbb{R}^{|\mathcal{S}|}$, which is associated with the objective function of problem (1):

$$f(x, z, y, \xi) := c^T x + \bar{c}^T z + \tilde{c}^T(\xi) y(\xi), \quad (2)$$

where x and z are the first-stage variables, and $y(\xi)$ denote the second-stage variables that depend on the random vector ξ . The variables x and z are independent of the realizations of the random vector.

We define two functions to evaluate the objective function of our problem as $\rho_{\mathbb{E}} : \Upsilon \rightarrow \mathbb{R}$ and $\rho_{\text{CVaR}} : \Upsilon \rightarrow \mathbb{R}$, where Υ is a linear space defined in the probability space $(\Omega, \mathcal{F}, P)$ used to evaluate random vectors. The function $\rho_{\mathbb{E}}$ evaluates the expectation, while ρ_{CVaR} evaluates the CVaR measure. The CVaR of the random variable $f(x, z, y, \xi)$ for the $(1-\alpha)$ -quantile of the cumulative distribution function of $f(x, z, y, \xi)$ is the conditional expectation of the random variable to be less than the $\text{VaR}_{1-\alpha}(f(x, z, y, \xi))$,

$$\rho_{\text{CVaR}}(f(x, z, y, \xi)) := -\text{CVaR}_{1-\alpha}(f(x, z, y, \xi)) := -\mathbb{E}_\xi[f(x, z, y, \xi) | f(x, z, y, \xi) \leq \text{VaR}_{1-\alpha}], \quad (3)$$

where the VaR of $f(x, z, y, \xi)$ for the $(1-\alpha)$ -quantile is defined as

$$\rho_{\text{VaR}}(f(x, z, y, \xi)) := -\text{VaR}_{1-\alpha}(f(x, z, y, \xi)) := -\max\{v | F_v(v) \leq 1 - \alpha\}. \quad (4)$$

We illustrate the concept of CVaR and VaR in Figure 1, and we introduce an alternative definition in Section 4.1.2.

Based on $\rho_{\mathbb{E}}$ and ρ_{CVaR} , we define the function $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ that is used to evaluate the decisions (x, z, y) :

$$\psi(x, z, y) := (1 - \beta)\rho_{\mathbb{E}}(f(x, z, y, \xi)) - \beta\rho_{\text{CVaR}}(f(x, z, y, \xi)), \quad (5)$$

where $\beta \in [0, 1]$ is a parameter that sets the weights of $\rho_{\mathbb{E}}$ and ρ_{CVaR} . The optimization of ψ maximizes the expected profit and minimizes the risk. Given the parameterization of ψ , we

Figure 1: Example of a profit distribution with $\text{VaR}_{1-\alpha}$ and $\text{CVaR}_{1-\alpha}$ marked as a dashed line and a continuous line, respectively, for $\alpha = 0.95$.

will consider three specific cases:

1. risk-neutral approach: ψ defined with $\beta = 0.0$, which leads to:

$$\psi_0(x, z, y) := \rho_{\mathbb{E}}(f(x, z, y, \xi)), \quad (6)$$

2. risk-averse approach: ψ for $\beta \in]0, 1[$, with ψ as in (5);

3. risk-averse approach: ψ with $\beta = 1.0$, which results in:

$$\psi_1(x, z, y) := -\rho_{\text{CVaR}}(f(x, z, y, \xi)). \quad (7)$$

Note that the function f in (2) represents profit, not losses. This is relevant because the properties of risk measures of gains differ slightly from the properties of the risk measures of losses; see a discussion about this issue in Rockafellar (2007).

4.1.1. Risk-neutral formulation

We first introduce a risk-neutral two-stage stochastic programming formulation, and later we focus on the risk-averse formulations. Using the notation described in section 3 and defining $\beta = 0.0$, we aim to maximize $\{\psi_0(x, z, y) := \rho_{\mathbb{E}}(f(x, z, y, \xi)) := \mathbb{E}_{\xi} [c^T x + \bar{c}^T z + \tilde{c}^T(\xi)y(\xi)]\}$

subject to the constraints of problem (1):

$$\begin{aligned}
 & \max_{x,z,y} \mathbb{E}_\xi [c^T x + \bar{c}^T z + \tilde{c}^T(\xi)y(\xi)] \\
 & s.t. \quad Ax + Bz \leq b \\
 & \quad Cz + Dy(\xi) \leq d, \quad a.s. \\
 & \quad Fy(\xi) + Gx = h(\xi), \quad a.s. \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, y(\xi) \in \mathbb{R}_+^{n_3}.
 \end{aligned} \tag{8}$$

The variables x and z are independent of the realizations of the random vector, and therefore, due to the properties of the expectation operator (Ross, 1987), the objective function can be simplified to $c^T x + \bar{c}^T z + \mathbb{E}_\xi [\tilde{c}^T(\xi)y(\xi)]$.

Assuming that the random vector ξ is approximated using a finite discrete distribution, the expectation determined in respect to ξ , $\mathbb{E}_\xi [\tilde{c}^T(\xi)y(\xi)]$, is evaluated as:

$$\mathbb{E}_\xi [\tilde{c}^T(\xi)y(\xi)] := \sum_{s \in \mathcal{S}} (p_s \tilde{c}_s^T y_s), \tag{9}$$

where $\tilde{c}_s := \tilde{c}(\xi_s)$ and $y_s := y(\xi_s)$, and thus, problem (8) can thus be written as:

$$\begin{aligned}
 & \max_{x,z,y_s} c^T x + \bar{c}^T z + \sum_{s \in \mathcal{S}} (p_s \tilde{c}_s^T y_s) \\
 & s.t. \quad Ax + Bz \leq b \\
 & \quad Cz + Dy_s \leq d, \quad \forall s \in \mathcal{S} \\
 & \quad Fy_s + Gx = h_s, \quad \forall s \in \mathcal{S} \\
 & \quad y_s \in \mathbb{R}_+^{n_3}, \quad \forall s \in \mathcal{S} \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2},
 \end{aligned} \tag{10}$$

which is denoted as extensive form of the stochastic program in Birge and Louveaux (2011). The total number of constraints of problem (10) is equal to $m_1 + m_2 \times |\mathcal{S}| + m_3 \times |\mathcal{S}|$, which means that this can easily become a large problem if a considerable number of scenarios is used.

Problem (8) can be rewritten in a typical two-stage formulation decomposed into two problems:

$$\begin{aligned}
 & \max_{x,z} c^T x + \bar{c}^T z + Q(x, z) \\
 & s.t. \quad Ax + Bz \leq b \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2},
 \end{aligned} \tag{11}$$

where $Q(x, z) := \mathbb{E}_\xi [Q(x, z, \xi)] = \sum_{s \in \mathcal{S}} [p_s Q(x, z, \xi_s)]$ is the recourse function value, and the recourse problem is defined as

$$Q(x, z, \xi_s) = \left. \begin{aligned} & \max_{y_s} \tilde{c}_s^T y_s \\ & s.t. \quad Dy_s \leq d - Cz \\ & \quad Fy_s = h_s - Gx \\ & \quad y_s \in \mathbb{R}_+^{n_3} \end{aligned} \right\} \quad \forall s \in \mathcal{S}. \tag{12}$$

4.1.2. Risk-averse formulation

In the context of the VPP that we consider, the risk measures the expected profit of the worst $(1 - \alpha)\%$ profits obtained for a given distribution of the random variables and a set of first-stage decisions. Therefore, for a given distribution of the random variables, the role of the risk measure in the objective function is to drive the optimization to make first-stage decisions that minimize the risk of obtaining low profits for some outcomes of the random variables. A set of first-stage decision variables with high risk means that some outcomes of the random variables may lead to lower profits than a first-stage solution with low risk. For the VPP studied, the first-stage decisions associated with the contracts are an instrument to hedge against the volatility of the electricity prices in the pool. The results presented in Section 6.4 further quantify and help to discuss risk management for the VPP.

In this work, we consider two risk-averse stochastic programming formulations that differ in the objective function. The first formulation maximizes the function (5), while the second maximizes the function (6). In both functions, the risk measure used is the CVaR. In the first formulation, the goal is to maximize a combination of the expected profit and of the conditional value at risk of the profit. While in the second, the goal is to maximize the CVaR of the profit.

Rockafellar and Uryasev (2000) proposed a more amenable definition of the CVaR for optimization than the one in (3):

$$\text{CVaR}_{1-\alpha}(Q(x, z, \xi)) := \max_{\eta \in \mathbb{R}} \left\{ \eta - \frac{1}{1-\alpha} \mathbb{E}_{\xi} [(\eta - Q(x, z, \xi))^+] \right\}, \quad (13)$$

where η is the value at risk, and $(\cdot)^+$ represents $\max\{\cdot, 0\}$. The problem defined in (13) together with the propositions (4.1) and (4.2) provide important features of the CVaR, making it a suitable risk measure to be integrated in optimization formulations.

Proposition 4.1. (Rockafellar and Uryasev, 2000) Problem (13) is a concave optimization problem.

The proof of Proposition (4.1) is based on the following: a) $Q(x, z, \xi)$ is piecewise concave in (x, z) b) $(\cdot)^+$ is a convex operator, thus $(\eta - Q(x, z, \xi))^+$ is convex because $\eta - Q(x, z, \xi)$ is convex; c) taking the expected value of a convex function preserves convexity (Kall and Mayer, 2011). The detailed proof is given in Rockafellar and Uryasev (2000) and Kall and Mayer (2011, p. 155).

Proposition 4.2. (Kall and Mayer, 2011) The $\text{CVaR}_{1-\alpha}$ is the optimal value of the following problem:

$$\text{CVaR}_{1-\alpha}(\tilde{c}^T(\xi)y(\xi)) := \max_{\eta} \left\{ \eta - \frac{1}{1-\alpha} \mathbb{E}_{\xi} [V(x, z, \xi, \eta)] \right\}, \quad (14)$$

where $V(x, z, \xi, \eta)$ is the random variable defined as

$$V(x, z, \xi, \eta) := \max_{v(\xi)} \{-v(\xi) | v(\xi) \geq \eta - \tilde{c}^T(\xi)y(\xi), v(\xi) \in \mathbb{R}_+\}. \quad (15)$$

Proposition 4.2 shows that problem (13) can be solved as a two-stage stochastic programming problem, where η is a first-stage variable, and v is a second-stage variable.

Assuming: a) that ψ is defined as in (5) with $\beta \in]0, 1[$; b) that ξ has a finite discrete distribution, which implies have a finite number of realizations of $v_s := v(\xi_s)$ with probability p_s ; c) that the formulation of the CVaR presented in proposition 4.2 is adapted to a finite number of realizations; and d) the translation invariance and positive homogeneity properties of CVaR (Pflug, 2000; Rockafellar and Uryasev, 2002); we define the following risk-averse two-stage stochastic programming formulation:

$$\begin{aligned}
 \max_{x, z, y_s, v_s, \eta} \quad & c^T x + \bar{c}^T z + (1 - \beta) \sum_{s \in \mathcal{S}} (p_s \tilde{c}_s^T y_s) + \beta \left[\eta - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} (p_s v_s) \right] \\
 \text{s.t.} \quad & Ax + Bz \leq b \\
 & Cz + Dy_s \leq d, \quad \forall s \in \mathcal{S} \\
 & Fy_s + Gx = h_s, \quad \forall s \in \mathcal{S} \\
 & v_s \geq \eta - \tilde{c}_s^T y_s, \quad \forall s \in \mathcal{S} \\
 & v_s \in \mathbb{R}_+, y_s \in \mathbb{R}_+^{n_3}, \quad \forall s \in \mathcal{S} \\
 & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R}.
 \end{aligned} \tag{16}$$

An equivalent formulation can be written as

$$\begin{aligned}
 \max_{x, z, \eta} \quad & c^T x + \bar{c}^T z + (1 - \beta) \sum_{s \in \mathcal{S}} [p_s Q(x, z, \xi_s)] + \beta \left\{ \eta - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} [p_s (\eta - Q(x, z, \xi_s))^+] \right\} \\
 \text{s.t.} \quad & Ax + Bz \leq b \\
 & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R},
 \end{aligned} \tag{17}$$

with $Q(x, z, \xi_s), \forall s \in \mathcal{S}$ given by (12).

The second risk-averse stochastic programming formulation uses $\beta = 1.0$ and $\psi_1(x, z, y) := -\rho_{\text{CVaR}}(f(x, z, y, \xi))$. This formulation maximizes the CVaR of the profit and it is expressed as

$$\begin{aligned}
 \max_{x, z, y_s, v_s, \eta} \quad & c^T x + \bar{c}^T z + \eta - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} (p_s v_s) \\
 \text{s.t.} \quad & Ax + Bz \leq b \\
 & Cz + Dy_s \leq d, \quad \forall s \in \mathcal{S} \\
 & Fy_s + Gx = h_s, \quad \forall s \in \mathcal{S} \\
 & v_s \geq \eta - \tilde{c}_s^T y_s, \quad \forall s \in \mathcal{S} \\
 & v_s \in \mathbb{R}_+, y_s \in \mathbb{R}_+^{n_3}, \quad \forall s \in \mathcal{S} \\
 & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R},
 \end{aligned} \tag{18}$$

or alternatively as

$$\begin{aligned} \max_{x,z,\eta} \quad & c^T x + \bar{c}^T z + \eta - \frac{1}{1-\alpha} \sum_{s \in \mathcal{S}} [p_s (\eta - Q(x, z, \xi_s))^+] \\ \text{s.t.} \quad & Ax + Bz \leq b \\ & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R} \end{aligned} \tag{19}$$

with $Q(x, z, \xi_s)$, $\forall s \in \mathcal{S}$ defined as in (12).

4.2. Remarks

1. The optimization of the function in (5) aims to maximize the expected value of f and minimize the risk of f . Therefore, in this function, we assume that $\min\{\rho_{\text{CVaR}}\} \equiv \max\{-\rho_{\text{CVaR}}\}$.
2. The definition $\rho_{\text{RCVaR}}(f(x, z, y, \xi)) := -\text{CVaR}_{1-\alpha}(f(x, z, y, \xi))$ assumes that the risk measure is evaluated over gains and not losses, and thus, the risk of $f(x, z, y, \xi)$ increases when the conditional value at risk of $f(x, z, y, \xi)$ decreases. Therefore, for a given distribution of the random variables, the risk associated with one set of first-stage variables is greater than the risk of another set of first-stage variables, if the expectation of the $(1 - \alpha)\%$ worse profits of the former is lower than the expectation of the $(1 - \alpha)\%$ worse profits of the latter.
3. The risk-neutral formulation (11) and the risk-averse formulation (19) are specific cases of the formulation (17); we can obtain the former with $\beta = 0.0$ and the latter with $\beta = 1.0$. However, the individual treatment of each formulation is relevant in order to adapt efficiently the solution methods.
4. The formulation in (17) uses two parameters, β and α . The first parameter sets the weights of the expectation and risk in the objective function, and the second defines the $(1 - \alpha)$ quantile over which the CVaR is evaluated.
5. The second risk-averse formulation given in (19) is similar to one of the formulations described in Pineda et al. (2008). These authors suggest that the maximization of the CVaR of the profit can be an alternative to the maximization of a combination of expected profit and CVaR of the profit. Pineda et al. (2008) discuss and compare the formulations, but they do not compare the computational performance of both formulations, neither using the extensive form or a decomposition algorithm.
6. The formulation in (19) does not require the parameter β , it is only parameterized over $\alpha \in [0, (|\mathcal{S}| - 1)/|\mathcal{S}|]$. This parameter allows to consider a risk-neutral formulation with $\alpha = 0$, which transforms (19) into an equivalent problem to (10). On the other extreme, with $\alpha = (|\mathcal{S}| - 1)/|\mathcal{S}|$, the formulation only maximizes the profit of the single scenario s , $f(x, z, y, \xi_s)$, such that $f(x, z, y, \xi_s) < \text{VaR}_{1-\alpha}(f(x, z, y, \xi))$, $\forall s \in \mathcal{S}$. The extreme case of maximizing the profit of the worst scenario can be viewed as a robust optimization problem, as suggested by Fabian (2008).

5. Solution methods

The two risk-averse formulations in (16) and (18) can be solved directly by an MILP solver, without requiring any decomposition method. However, for a large number of scenarios, the risk-averse stochastic programming formulations can become difficult to solve directly, and therefore, a decomposition method may be more efficient.

The decomposition methods used in this work are based on Benders Decomposition (Benders, 1962), more specifically the single-cut and multiple-cut L-Shaped methods (Birge and Louveaux, 1988; Van Slyke and Wets, 1969) for stochastic programming.

There are two main ideas behind these methods: 1) decomposition of the original problem in a master problem level and a subproblem level; and 2) in the master problem, the recourse problem is outer-approximated by a set of constraints that are generated and added to the master problem in each iteration of the decomposition method. The master problem is a relaxation of the original problem, and thus its optimal objective function value provides an upper bound on the objective function value of the original problem. Meanwhile, a lower bound is calculated based on the first-stage variables and optimal objective function values of the subproblems. The detailed calculation of the bounds is provided latter.

The methods described in this section use the following sets:

$\mathcal{S}$ - set of discrete scenarios;

$\mathcal{K}$ - set of iterations in the decomposition methods;

$\mathcal{M}$ - set of optimality cuts added to the master problem.

5.1. L-Shaped method adapted to handle the CVaR

The L-Shaped method proposed by Van Slyke and Wets (1969) for two-stage stochastic programming problems is described in detail in Birge and Louveaux (2011). This method was initially developed for two-stage linear stochastic programming problems, but it can also be adapted for problems with mixed-integer variables in the first-stage (Birge and Louveaux, 2011; Wollmer, 1980). In the context of a multi-cut L-Shaped method the master problem of (10) is given by the following formulation:

$$\begin{aligned}
 & \max_{x, z, \theta_s} \quad c^T x + \bar{c}^T z + \sum_{s \in \mathcal{S}} \theta_s \\
 & \text{s.t.} \quad Ax + Bz \leq b \\
 & \quad \theta_s \leq p_s \left[(d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right], \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad \theta_s \in \mathbb{R}, \quad \forall s \in \mathcal{S} \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2},
 \end{aligned} \tag{20}$$

while the subproblem is given by

$$\left. \begin{array}{ll} \min_{\mu_s, \lambda_s} & \left(d - Cz^k \right)^T \mu_s + \left(h_s - Gx^k \right)^T \lambda_s \\ \text{s.t.} & D^T \mu_s + F^T \lambda_s \geq \tilde{c}_s \\ & \mu_s \in \mathbb{R}_+^{m_1}, \lambda_s \in \mathbb{R}^{m_3} \end{array} \right\} \quad \forall s \in \mathcal{S}. \quad (21)$$

In problem (20), the second constraint is known as an optimality cut, and represents an outer-approximation of the second-stage expected recourse profit. We will designate this type of cut as expected value (EV) optimality cut. In this constraint, μ_s^m and λ_s^m are the dual variables obtained from problem (21) for fixed valued of (x, z) : (x^k, z^k) . A new set of optimality cuts are added to the master problem on each iteration of the L-Shaped method. Problem (21) is the dual of problem (12), which has an objective function that depends on the first-stage variables, and a feasible region that is independent of the first-stage variables. For a maximization problem, the L-Shaped method assumes that $Q(x, z)$ is piecewise linear concave, and therefore, it can be outer-approximated by the addition of optimality cuts. In our problem we assume that there is complete recourse, since for any (x^k, z^k) there exists y_s , $\forall s \in \mathcal{S}$ such that the problem (21) is bounded. This means that for any commitment and contract decided in the beginning of the week, the producer can always meet the contract obligation. This is possible because there are no restrictions on buying electricity from the pool to satisfy the electricity contract. Therefore, in problem (20), we do not include feasibility cuts as described in Birge and Louveaux (2011) or Kall and Mayer (2011).

The construction of the relaxation of the risk-averse stochastic programming problem (17) has to be adapted to account for the risk measure in the objective function. Following Ahmed (2006), Noyan (2012) proposed two methods based on the L-Shaped method for the decomposition of problem (17). We will highlight the multi-cut versions of these two methods.

In the first method, the piecewise concave term of the objective function in problems (17) or (19):

$$\left\{ \eta - \frac{1}{1-\alpha} \sum_{s \in \mathcal{S}} [p_s (\eta - Q(x, z, \xi_s))^+] \right\} \quad (22)$$

is outer approximated for (x, z) by the third and fourth constraints in (23), which we denominate by CVaR1 optimality cuts. These optimality cuts are added at each iteration of the decomposition method to the master problem. Note that the variable η is considered a second-stage variable. Therefore, for the first risk-averse formulation, given by (17), the master problem

using the CVaR1 optimality cuts is defined as

$$\begin{aligned}
 & \max_{x, z, \theta_{1s}, \theta_{2s}} \quad c^T x + \bar{c}^T z + (1 - \beta) \sum_{s \in \mathcal{S}} \theta_{1s} + \beta \sum_{s \in \mathcal{S}} \theta_{2s} \\
 & \text{s.t.} \quad Ax + Bz \leq b \\
 & \quad \theta_{1s} \leq p_s \left[(d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right], \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad \theta_{2s} \leq p_s \left(\eta^m - \frac{1}{1 - \alpha} v_s^m \right), \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad v_s^m \geq \eta^m - \left[(d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right], \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad v_s^m \geq 0, \quad \forall s \in \mathcal{S} \\
 & \quad \theta_{1s} \in \mathbb{R}, \theta_{2s} \in \mathbb{R}, \quad \forall s \in \mathcal{S} \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2},
 \end{aligned} \tag{23}$$

while, for the second risk-averse formulation, defined in (19), the master problem is as follows:

$$\begin{aligned}
 & \max_{x, z, \theta_{2s}} \quad c^T x + \bar{c}^T z + \sum_{s \in \mathcal{S}} \theta_{2s} \\
 & \text{s.t.} \quad Ax + Bz \leq b \\
 & \quad \theta_{2s} \leq p_s \left(\eta^m - \frac{1}{1 - \alpha} v_s^m \right), \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad v_s^m \geq \eta^m - \left[(d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right], \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \quad v_s^m \geq 0, \quad \forall s \in \mathcal{S} \\
 & \quad \theta_{2s} \in \mathbb{R}, \quad \forall s \in \mathcal{S} \\
 & \quad x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}.
 \end{aligned} \tag{24}$$

The decomposition algorithm using the master problems (23) or (24) is presented in Algorithm (1).

The second method assumes that η is a first-stage variable, which follows the idea of Proposition 4.2, and therefore, we have to outer-approximate the following expression of the CVaR term in the objective function of problems (17) and (19):

$$\sum_{s \in \mathcal{S}} [p_s (\eta - Q(x, z, \xi_s))^+]. \tag{25}$$

To handle the operator $(\cdot)^+$ and to outer-approximate (25), Ahmed (2006) proposed the third constraint in (27), which we designate as CVaR2 optimality cut, where σ_s^m indicates the recourse function values that are less than η^k . Therefore, σ_s^m is evaluated from the following expressions:

$$\sigma_s^m = \begin{cases} 1 & \text{if } \eta^k - Q(x^k, z^k, \xi_s) \geq 0 \\ 0 & \text{if } \eta^k - Q(x^k, z^k, \xi_s) < 0 \end{cases}, \quad \forall s \in \mathcal{S}. \tag{26}$$

Thus, the master problem for the risk-averse formulation (17) considering the CVaR2 optimality

cuts is given by:

$$\begin{aligned}
 \max_{x,z,\theta_{1s},\theta_{2s},\eta} \quad & c^T x + \bar{c}^T z + (1 - \beta) \sum_{s \in \mathcal{S}} \theta_{1s} + \beta \left(\eta - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} \theta_{2s} \right) \\
 \text{s.t.} \quad & Ax + Bz \leq b \\
 & \theta_{1s} \leq p_s \left[(d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right], \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \theta_{2s} \geq p_s \left[\eta - \left((d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right) \right] \sigma_s^m, \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \theta_{1s} \in \mathbb{R}, \theta_{2s} \in \mathbb{R}, \quad \forall s \in \mathcal{S} \\
 & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R},
 \end{aligned} \tag{27}$$

and the master problem for the second risk-averse formulation, defined in (19), is as follows:

$$\begin{aligned}
 \max_{x,z,\theta_{2s},\eta} \quad & c^T x + \bar{c}^T z + \eta - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} \theta_{2s} \\
 \text{s.t.} \quad & Ax + Bz \leq b \\
 & \theta_{2s} \geq p_s \left[\eta - \left((d - Cz)^T \mu_s^m + (h_s - Gx)^T \lambda_s^m \right) \right] \sigma_s^m, \quad \forall s \in \mathcal{S}, \forall m \in \mathcal{M} \\
 & \theta_{2s} \in \mathbb{R}, \quad \forall s \in \mathcal{S} \\
 & x \in \mathbb{R}_+^{n_1}, z \in \mathbb{B}^{n_2}, \eta \in \mathbb{R}.
 \end{aligned} \tag{28}$$

One interesting detail of the formulations (27) and (28) is that the number of CVaR2 optimality cuts added to the master problem in each iteration of the decomposition method is not constant. The value of η determined in the master problem and σ_s calculated after the evaluation of all subproblems determine the number of CVaR2 cuts that are added to the master problem in a given iteration. The details of the implementation of the decomposition method using the master problems (27) or (28) are presented in Algorithm (2). For brevity, we just show the multiple cut versions of the decomposition methods; however, we implement also single-cut versions of these algorithms.

In Algorithms (1) and (2) the evaluation of the CVaR is performed using equations (32) and (36), respectively. The CVaR is then used in the evaluation of a lower bound on the objective function of the original problem.

Algorithm 1 Multi-cut L-Shaped method adapted to handle the CVaR using the master problem (23) or (24) (Noyan, 2012).

Sets: $\mathcal{S}$ - scenarios, $\mathcal{K}$ - iterations, $\mathcal{M}$ - optimality cuts.

```

{Initialization}
 $LB = -\inf, UB = +\inf$ 
 $k := 0, m := 0, \mathcal{M} := \emptyset$ 
while  $|UB - LB| / |LB| > \epsilon$  do
    {Solve the master problem, for  $k = 0$ , set  $\theta_{1s} = 0, \theta_{2s} = 0, \forall s \in \mathcal{S}$ }
        Problem (23) or (24)

    {Let  $(x^k, z^k, \theta_{1s}^k, \theta_{2s}^k)$  be the optimal solution of the master problem}
    {Calculate the upper bound}

        
$$UB = \min \left\{ UB, c^T x^k + \bar{c}^T z^k + (1 - \beta) \sum_{s \in \mathcal{S}} \theta_{1s}^k + \beta \sum_{s \in \mathcal{S}} \theta_{2s}^k \right\} \quad (29)$$


    {Solve subproblem  $s, \forall s \in \mathcal{S}$ }
        Problem (21)

    {Let  $(\lambda_s^k, \mu_s^k)$  be the optimal solution of the subproblem  $s$ }
    {If problem (23) is used then add  $|\mathcal{S}|$  EV optimality cuts to the master problem}
     $m := m + 1, M := M \cup m$ 
        
$$\lambda_s^m = \lambda_s^k, \quad \forall s \in \mathcal{S} \quad (30)$$

        
$$\mu_s^m = \mu_s^k, \quad \forall s \in \mathcal{S} \quad (31)$$


    {Calculate the  $\eta^m$  for the  $\alpha$ -quantile of the distribution of  $Q(x^k, z^k, \xi_s), \forall s \in \mathcal{S}$ }
    {Add  $2 \times |\mathcal{S}|$  CVaR1 optimality cuts to the master problem}
    {Calculate the  $\text{CVaR}_{1-\alpha}^m$  of the distribution of  $Q(x^k, z^k, \xi_s), \forall s \in \mathcal{S}$ }

        
$$\text{CVaR}_{1-\alpha}^m = \eta^m - \frac{1}{1-\alpha} \sum_{s \in \mathcal{S}} p_s \left( \eta^m - \left[ (d - Cz^k)^T \mu_s^m + (h_s - Gx^k)^T \lambda_s^m \right] \right)^+ \quad (32)$$


    {Calculate the lower bound}

        
$$LB = \max \left\{ LB, c^T x^k + \bar{c}^T z^k + (1 - \beta) \sum_{s \in \mathcal{S}} \left\{ p_s \left[ (d - Cz^k)^T \mu_s^m + (h_s - Gx^k)^T \lambda_s^m \right] \right\} + \beta \text{CVaR}_{1-\alpha}^m \right\} \quad (33)$$


end while

```

Algorithm 2 Multi-cut L-Shaped method adapted to handle the CVaR using master problem (27) or (28) (Noyan, 2012).

Sets:= $\mathcal{S}$ - scenarios, $\mathcal{K}$ - iterations, $\mathcal{M}$ - optimality cuts.

{Initialization}

$LB = -\inf$, $UB = +\inf$

$k := 0$, $m := 0$, $\mathcal{M} := \emptyset$

while $|UB - LB|/|LB| > \epsilon$ **do**

{Solve the master problem, for $k = 0$ set $\theta_{1s} = 0$ and $\theta_{2s} = 0$ }

Problem (27) or (28)

{Let $(x^k, z^k, \eta^k, \theta_{1s}^k, \theta_{2s}^k)$ be the optimal solution of the master problem}

{Calculate the upper bound}

$$UB = \min \left\{ UB, c^T x^k + c^T z^k + (1 - \beta) \sum_{s \in \mathcal{S}} \theta_{1s}^k + \beta \left(\eta^k - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} \theta_{2s}^k \right) \right\} \quad (34)$$

{Solve subproblem $s, \forall s \in \mathcal{S}$ }

Problem (21)

{Let (λ_s^k, μ_s^k) be the optimal solution of the subproblem s }

{If problem (27) is used then add $|\mathcal{S}|$ optimality cuts to the master problem}

$m := m + 1$, $\mathcal{M} := \mathcal{M} \cup m$

Equations (30), (31)

{Calculate σ_s^m from the distribution of $Q(x^k, z^k, \xi_s)$, $\forall s \in \mathcal{S}$ }

$$\begin{aligned} \sigma_s^m &= 1, \text{ if } \eta^k - Q(x^k, z^k, \xi_s) \geq 0 \\ \sigma_s^m &= 0, \text{ if } \eta^k - Q(x^k, z^k, \xi_s) < 0 \end{aligned} \quad (35)$$

{Add $\mu|\mathcal{S}|$ CVaR2 optimality cuts to the master problem}

{Calculate the $\text{CVaR}_{1-\alpha}^m$ of the distribution of $Q(x^k, z^k, \xi_s)$, $\forall s \in \mathcal{S}$ }

$$\text{CVaR}_{1-\alpha}^m = \eta^k - \frac{1}{1 - \alpha} \sum_{s \in \mathcal{S}} \left\{ p_s \left[\eta^k - \left((d - Cz^k)^T \mu_s + (h_s - Gx^k)^T \lambda_s^m \right) \right] \sigma_s^m \right\} \quad (36)$$

{Calculate the lower bound}

$$LB = \max \left\{ LB, c^T x^k + \bar{c}^T z^k + (1 - \beta) \sum_{s \in \mathcal{S}} \left\{ p_s \left[(d - Cz^k)^T \mu_s^m + (h_s - Gx^k)^T \lambda_s^m \right] \right\} + \beta \text{CVaR}_{1-\alpha}^m \right\} \quad (37)$$

end while

6. Computational experiments

6.1. Description of the case studies

In this section, we use two case studies to evaluate the computational performance of the formulations and algorithms presented above. The two case studies involve two generation systems that comprise a pumped-storage hydro unit, a wind farm and a thermal unit. The only difference between the two generation concerns the thermal unit, which has different characteristics. In Case 1, the thermal unit, referred to as G1, has a larger capacity and lower generation costs than the thermal unit used in Case 2, referred to as G2. Furthermore, the minimum up and down time limits are distinct between G1 and G2. The deterministic model that describes the performance of the system in Case 1 and Case 2 and the corresponding data are presented in Appendix A. Each case is solved for two weeks: 1) one week in the Summer of August 2014;

and 2) one week in the Autumn of November, 2014. These two weeks introduce distinct wind and electricity forecasts, and in addition, the first week has a lower initial level of water and a lower natural inflow of water in the reservoir than the second week.

6.1.1. Uncertainty characterization: wind power and electricity scenarios

The wind speed forecast is based on a wind ensemble obtained from the European Centre for Medium Range Weather Forecasts (ECMWF) through The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble project. The ECMWF ensemble comprises 51 members, one unperturbed and 50 perturbed fields. Forecasts are run twice a day with initial times at 00 and 12 UTC, up to 15 days with a variable spatial resolution: about 32 km resolution with 91 vertical levels for first 10 days and about 65 km resolution for thereafter. The unperturbed initial conditions are obtained from ECMWF high-resolution four-dimensional variational (4DVAR) assimilative model run at 16 km resolution, with a 12-hour assimilation window. The perturbations are introduced as a linear combination of Singular Vectors (SVs) and ECMWF Ensemble Data Assimilation (EDA), where SVs account for the fastest-growing perturbations over a 48-hour time interval and EDA includes analysis error estimates provided by the ECMWF 4DVAR model. The model uncertainties in each EDA member are simulated by two stochastic schemes: the Stochastically Perturbed Parameterized Tendency (SPTT, Buizza et al., 1999) to simulate the random model errors due to parameterized physical processes, and the backscatter scheme (SKEB, Berner et al., 2008) to include the upscale energy transfer induced by the unresolved scales on the resolved scales. Since each ensemble member represents an equally likely scenario, all members have equal probability; see Buizza (2014) for more details.

The wind ensemble forecast was then used with a wind-power curve for a specific turbine with a rated power of 2.35 MW to calculate the wind power ensemble. The wind power ensemble for the week of August 25-31, 2014 and for the week of November 24-30, 2014 are presented in Figures 2a and 3a, respectively.

The wind uncertainty quantification using a wind ensemble provides a natural link between the ensemble member and the scenario used in stochastic programming. Specifically, each wind ensemble member is assumed to be an independent wind scenario in stochastic programming. Therefore, in this work, we deal with 51 independent scenarios to quantify the wind uncertainty.

The electricity prices historical data was obtained from the Iberian Peninsula electricity market (Iberian Electricity Market, 2015). The historical data encompasses a time series with 12 weeks of prices, which are used to fit ARIMA models. We consider that the time series are non-stationary, and use differentiation to capture seasonality. The fitted ARIMA models were then used to generate the electricity price scenarios by sampling the error of the ARIMA models for the 168 hours following the twelve weeks of historical data considered. The ARIMA models identified for each week are presented in Appendix B. This approach is very convenient to generate sets of scenarios with different cardinality. Therefore, we have generated the following 8 sets of scenarios for the electricity prices, with sample size $M \in E_M \equiv \{1; 10; 50; 100; 200; 300; 400; 500\}$. In Figures 2b and 3b, we show one set of 100

(a) Wind power ensemble forecast with 51 members.

(b) 100 electricity scenarios forecast. The continuous line is the point forecast, and the dotted lines are the 5% and 95% prediction errors.

Figure 2: Data for the week of August 25-31, 2014.

scenarios for the week of August 25-31, 2014 and another for the week of November 24-30, 2014, respectively. In addition, we also show the 95% prediction interval calculated by the algorithm that we have used to fit the ARIMA model.

Note that the VPP that we consider has no market power as its relative size is small with respect to the considered electric energy system (734.5 MW in Case 1 and 334.5 MW in Case 2 of the VPP vs. 119,600 MW of the system). Thus, the considered VPP has either no or very little influence on the clearing prices. Additionally, note that the price forecasts that we use are based on historical data pertaining to the same market and over the same forecasting horizon of the wind production, and thus correlations among wind productions and prices are

(a) Wind power ensemble forecast with 51 members.

(b) 100 electricity scenarios forecast. The continuous line is the point forecast, and the dotted lines are the 5% and 95% prediction errors.

Figure 3: Data for the week of November 14-30, 2014.

implicitly taken into account. However, the forecasting tools (for wind and price) that we use are independent, and thus the wind forecasting algorithm does not consider the price as an explicative variable, and vice versa, the price forecasting algorithm does not consider the wind production as an explicative variable. Therefore, to generate joint wind-price scenarios, we consider all possible combinations as if these random variables were independent. This is a reasonable approximation provided that the specific wind at the location of the VPP is not highly correlated with the price. We finally note that forecasting techniques are outside the scope of this paper, which focuses on testing and comparing a number of solution techniques for the considered stochastic programming problems, given a scenario set characterizing price

and wind production

For each week, two contracts with three blocks of constant energy and price to sell and three other blocks of constant energy and price to buy are considered. The specific data for the contracts are provided in Appendix A.

6.1.2. Setup of the parameters of the algorithms

The computational experiments performed with the two case studies and two weeks cover a combination of different sets of parameters and scenarios, which allow us to assess the performance of the solution methods, formulations, and to illustrate risk management. In terms of parameters and scenarios, we consider the following:

1. $\beta = 0.0$, which results in a risk-neutral formulation;
2. $\beta \in \{0.1; 0.5; 0.9\} \times \alpha = 0.9$;
3. $\beta = 1.0 \times \alpha \in \{0.1; 0.5; 0.9; 0.95; (|\mathcal{S}| - 1)/|\mathcal{S}|\}$;
4. the electricity prices are independent of the wind speed, and therefore, we consider eight sets of scenarios, each one with 51 wind power scenarios times the number of electricity prices scenarios in E_M .

In terms of solution approach, first we consider the direct solution of the extensive forms with different values of β , α , and number of scenarios. Secondly, we apply the decomposition methods described in Section 5, namely the L-Shaped method for risk-neutral formulations, and Algorithms 1 and 2 for risk-averse formulations. These decomposition methods are applied using the single and multiple versions of the EV, CVaR1, and CVaR2 optimality cuts. Table 1 gives an overview of the combinations used in the computational experiments in terms of values of β , formulations, the structure of the master problem in terms of cuts, and the variants of the algorithms used, and Tables 2, and 3 show the number of constraints and variables of the master problems. Note that the size of the subproblems is the same in each algorithm.

The extension of the risk-neutral formulation to the risk-averse formulation using the CVaR measure keeps the separability of the subproblems, which means that by construction the presented algorithms are amenable to be implemented in parallel. In this work, the models and algorithms are implemented in GAMS and solved with GAMS/CPLEX 12.6.1.0, which allow us to rapidly implement some parallelization techniques at the workstation level. In each algorithm, we implement three strategies to solve the subproblems: 1) no parallelization is used and the subproblems are solved in sequence; 2) the subproblems are solved in parallel using the grid features of GAMS (Bussieck et al., 2009); and 3) the subproblems are solved in parallel using a recent approach provided by GAMS: the grid plus GUSS solver (Bussieck et al., 2012). The grid+GUSS approach relies on the grid features of GAMS to parallelize the solution of the subproblems, and on the GUSS solver to reduce the overhead of the GAMS/CPLEX 12.6.1.0 solver to setup the problems.

Note that in the solution process of the subproblems in parallel, we distribute the subproblems over 40 cores, and therefore, we do not use the concurrent optimization option to enable

Table 1: Match of the formulations with the corresponding master problem and the variants (single- or multiple-cut) of the algorithms implemented.

Formulation		Master problem		Algorithms with single- or multiple-cuts			
β	Eq.	Eq.	Cuts				
				L-Shaped			
0	(11)	(20)	EV	SC	MC		
ALG 1							
$]0, 1[$	(17)	(23)	EV/CVaR1	SC/SC	SC/MC	MC/SC	MC/MC
ALG 2							
$]0, 1[$	(17)	(27)	EV/CVaR2	SC/SC	SC/MC	MC/SC	MC/MC
ALG 1							
1	(19)	(24)	CVaR1	SC	MC	-	-
ALG 2							
1	(19)	(28)	CVaR2	SC	MC	-	-

SC - single-cut; MC - multiple-cut.

Table 2: Number of variables of the master problem to add to $n_1 + n_2$ to obtain the total number of variables within the algorithms presented. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Master problem (11)	Master problem (17)	Master problem (19)
L-Shaped SC	1	-	-
ALG 1 SC/SC	-	$2 + \mathcal{S} /iteration$	-
ALG 1 SC/MC	-	$1 + \mathcal{S} + \mathcal{S} /iteration$	-
ALG 2 SC/SC	-	3	-
ALG 2 SC/MC	-	$2 + \mathcal{S} $	-
ALG 1 SC	-	-	$1 + \mathcal{S} /iteration$
ALG 2 SC	-	-	2
L-Shaped MC	$ \mathcal{S} $	-	-
ALG 1 MC/MC	-	$2 \mathcal{S} + \mathcal{S} /iteration$	-
ALG 1 MC/SC	-	$1 + \mathcal{S} + \mathcal{S} /iteration$	-
ALG 2 MC/MC	-	$1 + 2 \mathcal{S} $	-
ALG 2 MC/SC	-	$2 + \mathcal{S} $	-
ALG 1 MC	-	-	$ \mathcal{S} + \mathcal{S} /iteration$
ALG 2 MC	-	-	$1 + \mathcal{S} $

the parallel utilization of multiple solvers, because we are already distributing the solution of the LP subproblems by the cores. However, we use the parallel options available at the MILP solver level. Therefore, we use eight cores with the deterministic option activated for the solution of the MILP master problems in the decomposition algorithms and for the solution of the MILP extensive form. The utilization of multiple cores is a common practice nowadays, which enables the MILP solver to apply several heuristics in parallel within the branch & bound search. For more details about some of the advantages of using multiple cores, see Lima (2010)

Table 3: Number of constraints of the master problem to add to m_1 to obtain the total number of constraints within the algorithms presented. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Master problem (11)	Master problem (17)	Master problem (19)
L-Shaped SC	$1/iteration$	-	-
ALG 1 SC/SC	-	$(2 + \mathcal{S})/iteration$	-
ALG 1 SC/MC	-	$(1 + 2 \mathcal{S})/iteration$	-
ALG 2 SC/SC	-	$2/iteration$	-
ALG 2 SC/MC	-	$(1 + \mu \mathcal{S})/iteration$	-
ALG 1 SC	-	-	$(1 + \mathcal{S})/iteration$
ALG 2 SC	-	-	$1/iteration$
L-Shaped MC	$ \mathcal{S} /iteration$	-	-
ALG1 MC/MC	-	$3 \mathcal{S} /iteration$	-
ALG1 MC/SC	-	$(1 + 2 \mathcal{S})/iteration$	-
ALG 2 MC /MC	-	$ \mathcal{S} (1 + \mu) /iteration$	-
ALG 2 MC /SC	-	$(1 + \mathcal{S})/iteration$	-
ALG1 MC	-	-	$2 \mathcal{S} /iteration$
ALG 2 MC	-	-	$\mu \mathcal{S} /iteration$

and Lima and Grossmann (2011, 2016).

The optimization runs were performed using a workstation with 40 Intel Xeon CPU E5-2680 v2 @ 2.80 GHz processors, and 125.8 Gb of RAM. The stopping criteria used are 1500 seconds of wall-clock time or 0.5% of the gap between the lower bound and upper bound of the decomposition algorithms. Note that we define by gap the quantity: $|UB - LB| / |LB| \times 100\%$. In the full solution of the stochastic programming formulation, we use 1500 seconds for maximum wall-clock time or 0.5% for the optimality gap. Note that some computational times reported for the direct solution of the extensive form and decomposition algorithms exceed the maximum wall-clock time set. This situation may happen with large problems, when close to the maximum time set the algorithms need to complete some standing computational tasks before terminating.

In a supplementary document, we provide additional detailed results that complement the information in this section regarding the computational performances and solutions obtained.

6.2. Computational performance

6.2.1. Extensive form vs decomposition algorithm

The solution of the extensive form of the risk-neutral and risk-averse problems in (10), (16), and (18) can be addressed directly without using a decomposition algorithm. Therefore, we start by assessing the computational performance of the extensive forms solved with an MILP solver, and afterwards, we compare with the performance of the decomposition algorithms.

In order to further characterize the formulations, we will first analyze the size of the problems associated with the extensive forms and decomposition algorithms. Tables 4 and 5 report the size of extensive forms and the size of the master and subproblems in Algorithms 1 and 2. The size of the extensive forms in terms of the number of equations and variables depends on the

number of considered scenarios. Clearly, the size of the risk-averse formulations is larger than the size of the risk-neutral formulations, due to the inclusion of the CVaR modeling. The risk-averse formulation has $|\mathcal{S}|$ additional constraints and $|\mathcal{S}| + 1$ additional variables with respect to the risk-neutral formulation; see Table 4. Comparing the size of the two risk-averse extensive forms (16) and (18), both formulations have the same size, 10,287,886 constraints, 6,860,359 total variables, and 508 binary variables, but the first formulation has $|\mathcal{S}|$ more terms in the objective function.

In the case of the decomposition algorithms, the size of the subproblems is the same in the L-Shaped, Algorithm 1, and Algorithm 2, independently of considering a risk-neutral or risk-averse formulation. On the other hand, the size of the master problem in the decomposition algorithms is dynamic, which means that the number of variables or constraints increase with each iteration of the algorithm. The size of the master problem depends on the formulation, specific algorithm and the utilization of single- or multiple-cuts, as shown in Tables 2 and 3. In these tables, it is relevant to compare the size of the master problems for the first risk-averse formulation, for example in ALG 1 MC/MC and ALG 2 MC/MC the master problems have 16,486 constraints and 16,159 variables and 6,797 constraints and 11,059 variables, respectively, while in ALG 2 SC/SC it has only 1,188 constraints and 861 variables.

The computational performance of the extensive form for the risk-neutral and risk-averse formulations is presented in Tables 6 and 7, respectively. The MILP solver can close the gap of most of the risk-neutral extensive forms for problems with the number of scenarios ranging from 51 to 25,500. However, there are three exceptions, namely for Case 2, Week 1 with 25,500 scenarios and Case 2 Week 2 with 20,400 and 25,500 scenarios.

When risk-averse formulations are considered, the extensive form becomes less efficient. For example, for Case 1, Week 1 with 5,100 scenarios, the extensive form solves the problem with $\beta = 0.0$ in 269 s with gap 0.00%, but it cannot solve the problem with $\beta = 0.1$ in 1500 s, see Tables 6 and 7, respectively. Also, there is a set of problems for which the solver did not return a meaningful MILP gap. This is indicated by the symbol * in Table 7. This situation occurs when the solver cannot compute the solution of the initial relaxation at the root node of the branch & bound tree or the solver cannot find an integer solution within the maximum time set.

In order to evaluate the impact of the maximum wall-clock time set on the solution of the extensive form with $\beta \in \{0.1; 0.5; 0.9\}$ and 5,100 scenarios, we increase the maximum wall-clock

Table 4: Size of the extensive form.

Formulation	Scenarios	β	NCNST	NVAR	0-1 NVAR
Extensive	5,100	0.0	10,282,786	6,855,258	508
Extensive	5,100	0.5	10,287,886	6,860,359	508
Extensive	5,100	1.0	10,287,886	6,860,359	508
Extensive	25,500	0.0	51,409,186	34,272,858	508

NCNST - number of constraints plus objective function; NVAR - number of total variables; 0-1 NVAR - number of binary variables.

Table 5: Size of the problems within the decomposition algorithms in the second iteration. 5,100 scenarios are considered. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Problem	β	NCNST	NVAR	0-1 NVAR
L-Shaped SC	Master	0.0	1,187	859	508
L-Shaped MC	Master	0.0	6,286	5,958	508
ALG 1 SC/SC	Master	0.9	6,288	5,961	508
ALG 1 SC/MC	Master	0.9	11,387	11,060	508
ALG 1 MC/SC	Master	0.9	11,387	11,060	508
ALG 1 MC/MC	Master	0.9	16,486	16,159	508
ALG 2 SC/SC	Master	0.9	1,188	861	508
ALG 2 SC/MC	Master	0.9	1,698	5,960	508
ALG 2 MC/SC	Master	0.9	6,287	5,960	508
ALG 2 MC/MC	Master	0.9	6,797	11,059	508
ALG 1 SC	Master	1.0	6,287	5,960	508
ALG 1 MC	Master	1.0	11,386	11,059	508
ALG 2 SC	Master	1.0	1,187	860	508
ALG 2 MC	Master	1.0	1,697	5,959	508
-	SP [†]	-	1,345	2,017	0

NCNST - number of constraints plus objective function; NVAR - number of total variables; 0-1 NVAR - number of binary variables. SP[†] - Subproblem, where the size is from problem (21). The size of the master problems corresponds to the second iteration of the decomposition algorithms.

Table 6: Computational results for the extensive form for $\beta = 0.0$ as a function of the number of scenarios. **Bold** - runs that did not meet the gap criterion within 1,500 s.

SCN	Case 1				Case 2			
	Week 1		Week 2		Week 1		Week 2	
	G (%)	T (s)	G (%)	T (s)	G (%)	T (s)	G (%)	T (s)
51	0.00	5	0.00	5	0.00	4	0.00	5
510	0.00	22	0.00	22	0.00	21	0.00	21
2,550	0.00	118	0.00	119	0.00	107	0.00	120
5,100	0.00	269	0.00	270	0.00	247	0.00	321
10,200	0.00	590	0.00	614	0.00	622	0.00	900
15,300	0.00	931	0.00	980	0.00	1,156	0.00	1,897
20,400	0.00	1,333	0.00	762	0.00	2,030	0.59	2,114
25,500	0.00	1,747	0.00	929	1.10	2,317	*	*

SCN - Number of scenarios; G - MILP gap between the best integer objective function value and the objective function value of the best node remaining; T - Elapsed wall-clock time; * - Solver aborted during optimization.

Table 7: Computational results for the extensive form as a function of the number of scenarios and β . Maximum wall-clock time set to 1500 s for all instances, except in the last four rows, where it is set to 7200 s. $\alpha = 0.9$.

SCN	β	Case 1				Case 2			
		Week 1		Week 2		Week 1		Week 2	
		G (%)	T (s)	G (%)	T (s)	G (%)	T (s)	G (%)	T (s)
51	0.1	0.00	7	0.00	8	0.00	14	0.00	25
51	0.5	0.00	7	0.00	8	0.00	16	0.00	24
51	0.9	0.00	7	0.00	9	0.00	17	0.00	24
51	1.0	0.00	7	0.00	12	0.00	16	0.00	35
510	0.1	0.00	221	0.00	291	0.00	1,171	*	1,566
510	0.5	0.00	452	0.00	288	0.00	1,434	*	1,545
510	0.9	0.00	445	0.00	656	*	1,542	*	1,548
510	1.0	0.00	1,190	0.00	1,468	*	1,534	*	1,517
2,550	0.1	*	1,611	*	1,586	*	2,083	*	1,588
2,550	0.5	*	1,643	*	1,590	*	2,119	*	1,596
2,550	0.9	*	2,944	*	3,928	*	3,425	*	1,669
2,550	1.0	*	1,666	*	1,729	*	1,650	*	1,629
5,100	0.1	*	1,742	*	2,952	*	3,828	*	2,930
5,100	0.5	*	2,929	*	1,776	*	3,642	*	1,757
5,100	0.9	*	2,562	*	1,683	*	3,496	*	2,042
5,100	1.0	*	2,075	*	2,208	*	2,331	*	2,048
$T_{\max} = 7,200$ s									
5,100	0.1	*	7,446	*	7,213	*	9,781	*	7,434
5,100	0.5	*	7,450	*	8,617	*	9,291	*	7,379
5,100	0.9	*	13,265	*	16,147	*	7,639	*	7,404
5,100	1.0	*	7,774	*	8,000	*	7,762	*	7,458

SCN - Number of scenarios; G - MILP gap between the best integer objective function value and the objective function value of the best node remaining; T - Elapsed wall-clock time; * - Solver did not return a meaningful gap value.

time to 7200 s, but the MILP solver still cannot find an integer solution, see Table 7.

The decomposition algorithms described in Section 5 were implemented using sequential and parallel solution strategies. We start by comparing the extensive form with the sequential and parallel solution strategies for $\beta = 0.0$. Figure 4 compares the wall-clock time required by the extensive form with the L-Shaped algorithm using three approaches to solve the subproblems: 1) sequential; 2) in parallel with the GAMS grid; and 3) in parallel with the GAMS grid+GUSS. It is relevant to observe that 1) the L-Shaped method with the sequential approach becomes rapidly inefficient when the number of scenarios increases, it requires 136 s for 51 scenarios, 1,571 s for 510 scenarios and 17,732 s to perform a single iteration with 5,100 scenarios; and 2) the L-Shaped method with the grid approach performs worse than the extensive form for the three sets of scenarios, for example, we would expect a better performance of the grid approach with 5,100 scenarios than the extensive form, however, the former requires 592 s and the latter 269 s. The L-Shaped method with the grid+GUSS approach becomes more efficient than the extensive form when the number of scenarios increases. From Figure 5, it is also clear that

for problems with a larger number of scenarios, the L-Shaped method with the grid+GUSS approach outperforms the extensive form. For example, for 25,500 scenarios the grid+GUSS approach requires 702 s while the extensive form 1,747 s.

Overall, the grid+GUSS approach is the most efficient to handle the solution of the subproblems, and therefore, the computational results presented in the next sections rely only on this approach.

Figure 4: Wall-clock time comparison between using extensive form (Extensive) and the L-Shaped method using the sequential (SEQ), grid (GRID), and grid+GUSS (GUSS) approaches for the solution of the subproblems. Results obtained for Case 1, week 1, and $\beta = 0.0$, with the L-Shaped SC method.

Figure 5: Wall clock time comparison between the extensive form (Extensive) and the L-Shaped method using the grid+GUSS (GUSS) approach for the solution of the subproblems. Results obtained for Case 1, week 1, and $\beta = 0.0$, with the L-Shaped SC method.

Table 8: Performance of the L-Shaped method with single-cuts (SC) and multiple-cuts (MC) for $\beta = 0.0$. **Bold** - runs that did not meet the gap criterion. Stop criteria: 1500 s, 0.5% gap. The explanation of the acronyms for the algorithms is shown in Table 1.

Cases	Cuts	Week 1						Week 2					
		5,100 scn.			25,500 scn.			5,100 scn.			25,500 scn.		
		Iter	G (%)	T (s)	Iter	G (%)	T (s)	Iter	G (%)	T (s)	Iter	G (%)	T (s)
Case 1	SC	2	0.00	42	2	0.00	702	2	0.04	43	2	0.03	700
Case 1	MC	2	0.00	43	2	0.00	708	2	0.04	44	2	0.03	747
Case 2	SC	3	0.35	71	3	0.28	1,346	21	5.36	1,570	4	6.31	2,002
Case 2	MC	3	0.35	91	3	0.28	1,430	7	5.77	2,431	3	6.20	1,524

Iter - Number of iterations; G - Gap between the lower bound and upper bound of the algorithm; T - wall-clock time.

6.2.2. Performance of the decomposition algorithms

In this section, we present the computational results of the L-Shaped method and Algorithms 1 and 2, implemented with single- and multiple-cuts. The following results are organized by formulation, defined by the value of β . First we focus on $\beta = 0.0$, then on $\beta \in \{0.1; 0.5; 0.9\}$, and finally on $\beta = 1.0$.

The performance of the L-Shaped method applied to Case 1 and Case 2 with the formulation with $\beta = 0.0$ is presented in Table 8. The results show that the L-Shaped method can meet the gap criterion in Case 1 for both weeks, and that the method with the grid+GUSS approach can meet the gap criterion for problems with 5,100 and 25,500 scenarios. Regarding Case 2, the L-Shaped method can meet the gap criterion for the problems with 5,100 and 25,500 scenarios of Week 1, but, it terminates with gaps above 5% and 6% for Week 2 for problems with 5,100 or 25,500 scenarios, respectively. Figure 6 shows the profile of the bounds of the L-Shaped method applied to Case2, Week 2 with $\beta = 0.9$ and 5,100 scenarios, and the solution obtained from the extensive form for the same problem. In this figure, it is clear that both the upper and lower bounds are converging slowly to the solution obtained from the extensive form.

Tables 9 and 10 show the performance of Algorithms 1 and 2 with single-cuts (SC) and multiple-cuts (MC) applied to Cases 1 and 2 with the formulations with $\beta \in \{0.1; 0.5; 0.9\}$ and $\alpha = 0.9$. These tables focus on the number of runs that meet the stopping criteria set and the average elapsed wall-clock time of these three runs, out of the three runs for the values of $\beta \in \{0.1; 0.5; 0.9\}$.

The results in Table 9 indicate that the four variants of Algorithm 1 meet the gap criteria in all problems of Case 1, Week 1 and 2, independently of the number of scenarios used. The same table shows that the performance of Algorithm 1 is better than the performance of Algorithm 2 for both weeks and independently of the type of cuts used. It is clear that Algorithm 2 requires larger computational times than Algorithm 1 to meet the same stopping criteria. For example, considering the cases where the three runs meet the gap criterion, the maximum average times for Algorithm 1 are 56 s, 784 s, 55 s, and 814 s, and for Algorithm 2 are 260 s, 1,202 s, 252 s, and 1,297 s for Week 1 with 5,100 and 25,500 scenarios and Week 2 with 5,100 and 25,500

Figure 6: Bounds of the L-Shaped SC method and the solution from the extensive form for the Case 2, Week 2. 5,100 scenarios.

Table 9: Performance of Algorithms 1 and 2 with single-cuts (SC) and multiple-cuts (MC) applied to Case 1 with $\beta \in \{0.1; 0.5; 0.9\}$, $\alpha = 0.9$. **Bold** - at least one run did not meet the gap criterion. Stop criteria: 1500 s, 0.5% gap. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Week 1				Week 2			
	5,100 scn.		25,500 scn.		5,100 scn.		25,500 scn.	
	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)
ALG 1 SC/SC	3	54	3	767	3	54	3	786
ALG 1 SC/MC	3	55	3	766	3	54	3	814
ALG 1 MC/SC	3	56	3	773	3	55	3	802
ALG 1 MC/MC	3	55	3	784	3	54	3	811
ALG 2 SC/SC	3	237	0	1,685	3	203	0	1,801
ALG 2 SC/MC	3	83	3	1,168	3	84	3	1,246
ALG 2 MC/SC	3	260	0	1,724	3	252	0	1,879
ALG 2 MC/MC	3	87	3	1,202	3	95	3	1,297

SR - number of runs of the algorithm that meet the stop criteria of the algorithms, out of three for $\beta \in \{0.1; 0.5; 0.9\}$; $\bar{T}$ - average elapsed wall-clock time.

scenarios, respectively.

In addition, with 25,500 scenarios, the two variants of Algorithm 2 with the single-cut version of the CVaR2 optimality cuts (ALG 2 SC/SC and ALG 2 MC/SC) are not able to close the gap of the three problems considered, see Table 9. Figures 7 and 8 provide additional insights about the progress of the bounds of Algorithms 1 and 2 for Case 1, Week 1 with $\beta = 0.9$, for 5,100 and 25,500 scenarios, respectively. In both cases, 5,100 and 25,500 scenarios, Algorithm 1 SC/SC meets the gap criterion in two iterations and Algorithm 2 SC/MC in three iterations. With 5,100 scenarios, Algorithm 2 SC/SC requires seven iterations, and with 25,500 scenarios it does not close the gap within the time limit. Table 9 also shows that the performance of Algorithm 1 with multiple cuts is not superior to the performance of its single-cut version. However, the

Figure 7: Bounds of the Algorithms 1 and 2 for Case 1, Week 1 with $\beta = 0.9$ and 5,100 scenarios.

Figure 8: Bounds of the Algorithms 1 and 2 for Case 1, Week 1 with $\beta = 0.9$ and 25,500 scenarios.

multiple-cut version of the CVaR2 optimality cuts are necessary to speed-up Algorithm 2, as it can be seen if comparing in Table 9 the algorithm ALG2 SC/SC with ALG2 SC/MC and ALG2 MC/SC with ALG2 MC/MC. The supplementary document complements the results presented with additional information regarding the trade-off between the number of iterations and the elapsed wall-clock time when the single- or multiple-cut versions of the algorithms are used. Figures 7 and 8 and the results in the supplementary document show that the most efficient version of Algorithm 2 requires one more iteration than Algorithm 1 to meet the gap criterion. The smaller number of iterations required by Algorithm 1 results in a computational advantage over Algorithm 2, in all the instances studied. This advantage suggests that the cuts added to the master problem in Algorithm 1 are stronger than the cuts added in Algorithm 2. By construction, the $\text{VaR}(\eta)$ is a second-stage variable in Algorithm 1 that is evaluated from

Table 10: Performance of Algorithms 1 and 2 with single-cuts (SC) and multiple-cuts (MC) applied to Case 2 with $\beta \in \{0.1; 0.5; 0.9\}$, $\alpha = 0.9$, and 5,100 scenarios. **Bold** - at least one run did not meet the gap criterion. Stop criteria: 1500 s, 0.5% gap. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Week 1				Week 2	
	5,100 scn.		25,500 scn.		5,100 scn.	
	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)
ALG 1 SC/SC	3	150	1	1,579	0	2,387
ALG 1 SC/MC	3	169	1	1,816	0	1,956
ALG 1 MC/SC	3	167	2	1,624	0	2,237
ALG 1 MC/MC	3	225	1	1,835	0	2,140
ALG 2 SC/SC	1	1,117	0	2,007	0	1,626
ALG 2 SC/MC	3	326	0	2,113	0	1,690
ALG 2 MC/SC	1	1,230	0	1,769	0	1,843
ALG 2 MC/MC	3	603	0	1,597	0	1,984

SR - number of runs of the algorithm that meet the stop criteria of the algorithms, out of three for $\beta \in \{0.1; 0.5; 0.9\}$; $\bar{T}$ - average elapsed wall-clock time.

the distribution of $Q(x^k, z^k, \xi_s) \forall s \in \mathcal{S}$ in the first iteration. On the other hand, in Algorithm 2, η is a first-stage variable that is assumed to be zero in the first iteration, and therefore, in the second iteration the CVAR2 optimality cuts added to the master problem are not as strong as the cuts added in Algorithm 1. In the second iteration of Algorithm 2, the master problem provides a new value of η , and thus, the CVAR2 optimality cuts added to the master problem in the third iteration use the information of η from the second iteration. Note that in Algorithm 2, η and $Q(x^k, z^k, \xi_s) \forall s \in \mathcal{S}$ define the indicator σ_s^m that determines the CVAR2 optimality cuts added to the master problem. For large-scale problems, e.g., with 25,000 scenarios, the additional iterations required by Algorithm 2 with CVaR2 optimality cuts, by comparison with Algorithm 1, may prevent Algorithm 2 to meet the gap criterion within the maximum time set.

In Case 2, the relative performance of the algorithms is similar, Algorithm 1 is more efficient than Algorithm 2, but the overall performance of both algorithms degrades, see Table 10. Furthermore, Algorithms 1 and 2 are not able to close the gap of any formulation out of the three for Week 2.

The results obtained with $\beta = 1.0$ and $\alpha \in \{0.1; 0.5; 0.9; 0.95; (|\mathcal{S} - 1)/|\mathcal{S}|\}$ for Cases 1 and 2 are presented in Tables 11 and 12. For Case 1, Week 1, Algorithm 1 takes on average 54 s for 5,100 scenarios, which is not a significant difference compared with the 54 s, 55 s, 56 s, and 55 s for 5,100 scenarios that Algorithm 1 requires to meet the gap criterion for $\beta \in \{0.1; 0.5; 0.9\}$ and $\alpha = 0.9$. However, in Case 1, Week 2 with 25,500 scenarios, the performance of Algorithm 1 degrades when compared with the formulations with $\beta < 1$, with a maximum average wall-clock time of 814 s for $\beta < 1$ and 1,249 s for $\beta = 1$, see Tables 9 and 11.

The results of Case 2 show also some degradation of the performance of Algorithm 1, and confirm the underperformance of Algorithm 2 with single-cuts. For example, for Case 2, Week 1, the average elapsed wall-clock time of ALG 2 SC is 1,378 s with only 1 out 5 problems solved, while the average elapsed wall-clock time of ALG 2 MC is 598 s with 4 out of 5 problems solved,

Table 11: Performance of Algorithms 1 and 2 with single-cuts (SC) and multiple-cuts (MC) applied to Case 1 with $\beta = 1.0$, $\alpha \in \{0.10; 0.50; 0.9; 0.95, (|S - 1)/|S|\}$. **Bold** - at least one run did not meet the gap criterion. Stop criteria: 1500 s, 0.5% gap. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Week 1				Week 2			
	5,100 scn.		25,500 scn.		5,100 scn.		25,500 scn.	
	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)
ALG 1 SC	5	54	5	840	5	54	4	1,249
ALG 1 MC	5	54	5	844	5	54	4	1,199
ALG 2 SC	5	219	0	1,628	5	267	0	1,738
ALG 2 MC	5	83	5	1,264	5	86	4	1,475

SR - number of runs of the algorithm that meet the stop criteria of the algorithms, out of five for $\alpha = \{0.1; 0.5; 0.9; 0.95; (|S - 1)/|S|\}$; $\bar{T}$ - average elapsed wall-clock time.

Table 12: Performance of Algorithms 1 and 2 with single-cuts (SC) and multiple-cuts (MC) applied to Case 2 with $\beta = 1.0$, $\alpha \in \{0.10; 0.50; 0.9; 0.95, (|S - 1)/|S|\}$. **Bold** - at least one run did not meet the gap criterion. Stop criteria: 1500 s, 0.5% gap. The explanation of the acronyms for the algorithms is shown in Table 1.

Algorithm	Week 1				Week 2	
	5,100 scn.		25,500 scn.		5,100 scn.	
	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)	SR	$\bar{T}$ (s)
ALG 1 SC	4	525	2	1,982	0	2,140
ALG 1 MC	5	491	1	1,783	0	2,427
ALG 2 SC	1	1,378	0	1,809	0	1,723
ALG 2 MC	4	598	0	1,954	0	2,184

SR - number of runs of the algorithm that meet the stop criteria of the algorithms, out of five for $\alpha \in \{0.10; 0.50; 0.9; 0.95, (|S - 1)/|S|\}$; $\bar{T}$ - average elapsed wall-clock time.

see Table 10.

It is relevant to observe that the computational experiments with Case 2 provide additional information about the performance of the decomposition algorithms and reveal some of their limitations that are not clear with Case 1. First, Week 1 of Case 2 leads to a degradation of the performance of the methods, in comparison with Case 1. Second, for Case 2, Week 2 with 5,100 scenarios and $\beta = 0.0$, the L-Shaped method stops at the maximum time set with gaps of 5.36% and 5.77% for single-cut and multiple cuts, respectively, while the MILP solver closes the gap of the extensive form in 321 s. Third, for Case 2, Week 2 with 5,100 scenarios and $\beta \in \{0.1; 0.5; 0.9; 1.0\}$, both algorithms are not able to close the gap within the maximum time set of any problem. These results indicate that the combination of the endogenous higher generation cost and the minimum up and down time of the thermal unit in Case 2 combined with the lower electricity prices of Week 2, by comparison with Week 1, force the decomposition algorithms to perform more iterations than in Case 1 to close the gap.

The formulations and algorithms tested with Case 1 and Case 2 are the same and both cases are solved over the same electricity prices and wind speed data. The performance of the tested methods is worse with Case 2 than with Case 1, which suggest that for problems with the same

structure and size, the used data has an influence on the performance of the methods.

However, our results show that Algorithm 1 performs better than Algorithm 2, and that the multiple cut version of the CVAR2 optimality cuts is important for Algorithm 2 in order to reduce the number of iterations and time to meet the gap criterion set. This relative performance of the methods is independent of the two risk-averse formulations, two cases, two weeks, risk management parameters, and number of scenarios. This suggests that for alternative VPP configurations, the algorithms will probably exhibit a similar behavior.

In a completely different application, Noyan (2012) reports that using single-cuts, Algorithm 1 requires less iterations than Algorithm 2, but in terms of computational time there is no clear cut advantage of Algorithm 1. In that study, it is also reported that the multiple cut version of Algorithm 2, by comparison with the single-cut variant, reduces the number of iterations and computational time for a subset of instances, but, the computational time increases for another subset of instances.

In this work, the computational experiments focus on the performance of the algorithms for distinct sets of data for the same problem structure. Other large-scale risk-averse stochastic problems using the CVaR, but with relative differences between the size of the master problem and size and number of subproblems, and master problems with distinct combinatorial complexity may lead to contrasting outcomes of the relative performance of the studied algorithms.

6.2.3. Sensitivity of the results to the number of scenarios

In this section, we analyze the impact of the number of scenarios on the objective function values and in the first-stage decisions of Case 1, Weeks 1 and 2. Note that the objective functions defined in the three formulations considered – risk-neutral, first risk-averse, and second risk-averse – represent, respectively, the expected value of the profit, a combination of expected value of the profit and conditional value at risk of the profit, and the conditional value at risk of the profit. First, we consider the case where the β ranges from zero to one with $\alpha = 0.9$, and afterwards, we consider the case where $\beta = 1.0$ and α ranges from 0.1 to $(|\mathcal{S}| - 1)/|\mathcal{S}|$.

The sensitivity of the objective function to the number of postulated scenarios is presented in Figures 9 and 10, for Case 1, Week 1 and Case 1, Week 2, respectively. In these figures, the objective function of 5 formulations with different values of β are plotted versus the number of scenarios. As expected, for fixed β there is a variation in the objective function values, and for increasing values of β the objective function value decreases. A more detailed analysis shows that in both figures, and in all formulations, there is a clear variation of the objective function between the problems with 51, 510, and 2,550 scenarios, for example for $\beta = 1$ the values are 2,920,958, 2,167,594, and 2,074,453; beyond which, increasing the number of scenarios, the variation in the objective function is within two standard deviations of the average of the objective function value of the set of scenarios with sample size $N \in S_N \equiv \{2, 550; 5, 100; 10, 200; 15, 300; 10, 400; 25, 500\}$. The results obtained for 51 scenarios, where the uncertainty is only due to the wind power, show a relative small sensitivity of the objective function to the value of β , by comparison with the results obtained with larger number of scenarios. This means that for this generation system, the uncertainty of the electricity

Figure 9: Objective function value as a function of the number of scenarios, Case 1, Week 1.

Figure 10: Objective function value as a function of the number of scenarios, Case 1, Week 2.

prices has a significant impact on the objective function values.

In this problem, the first-stage variables indicate the time periods that the thermal unit is generating electricity, the startups and shutdown periods of the thermal unit, and the decisions related with the contracts. The results obtained show that in Case 1, Weeks 1 and 2, the number of scenarios does not have any impact on the thermal unit operations. In all sets of scenarios the thermal unit is up during the whole week, and therefore, there are no startups or shutdowns of the thermal unit. The impact of the number of scenarios is only relevant on the decisions associated with the contracts, namely which contract to sign and the values of the energy sold or bought through each contract. Note that in our formulation, there is a constraint that ensures that each contract can only be used to buy or sell energy, but there is no limitation to buy energy from one contract and use another contract to sell energy.

In Figures 11 and 12, we present the impact of the number of scenarios on the power bought or sold by contracts. In these figures, we assume that the power bought is negative and the power sold is positive. First, it is clear that using risk-neutral formulations, results suggest to buy electricity from contracts, 160 MW in Case 1, Week 1 and Week 2 when the number of scenarios is greater than or equal to 2,550 and 510 in Week 1 and 2, respectively, while more conservative solutions indicate to sell electricity through contracts, between 260 and 315 MW in Week 1 and between 155 MW and 210 MW in Week 2, depending on the number of scenarios considered.

In terms of sensitivity, we can observe in Figures 11a and 12a that for the risk-neutral formulation, after increasing the scenarios to 2,500 and 510, respectively, the impact of the number of scenarios on the energy bought by contract is null. However, when $\beta = 0.1$ is used, in Case 1, Week 1, the set of used scenarios has a considerable impact on the contracts definition. The results show that different number of scenarios may indicate to use one contract to buy and the other to sell, or just use one contract to buy. For three sets of scenarios the solution indicates only to buy electricity through contract, while in the remaining sets electricity is bought and sold through contracts. For Case 1, Week 2, after 51 scenarios the impact is null. In general, for $\beta > 0.5$, the sets of scenarios postulated with more than 510 scenarios do not present significant variations in the contracts definition.

6.3. Value of the stochastic solution and value of perfect information

In this section, we assess the Value of the Stochastic Solution (VSS) and Value of Perfect Information (VPI) for the risk-neutral and the first risk-averse formulation. The VSS and VPI for risk-neutral problems are standard indicators to measure the value of stochastic solutions and methods, but they are not so common in risk-averse formulations. Noyan (2012) is one of the first authors to address the calculation of the VSS and VPI for risk-averse formulations. Note that these indicators are highly dependent on the data of the problem as clearly shown in simple examples in Birge and Louveaux (2011). Table 13 presents the VSS and VPI for the case studies considered for different values of β . These results show that when the objective function

Table 13: VSS and EVPI for different values of β with 5,100 scenarios and $\alpha = 0.9$, for the two cases and two weeks.

C, W	$\beta = 0.0$		$\beta = 0.1$		$\beta = 0.5$		$\beta = 0.9$		$\beta = 1.0$	
	VSS	EVPI	VSS	EVPI	VSS	EVPI	VSS	EVPI	VSS	EVPI
1, 1	0	227,390	27,246	265,668	285,123	269,889	745,185	71,923	882,225	408
1, 2	0	92,351	0	122,206	107,289	134,339	322,321	38,727	384,238	6,666
2, 1	0	232,376	27,072	238,291	231,701	165,611	476,510	52,752	537,984	24,266
2, 2	-14,044	117,290	-16,147	142,120	94,095	122,785	233,014	74,773	271,752	58,762

C - Case; W - Week.

considers only the expected value of the profit, the VSS and VPI are equal to zero, with an exception in Case 2 Week 2, which is negative. However, note that by definition the VSS and VPI must be greater than zero (Birge and Louveaux, 2011). In Case 2, Week 2, the negative

(a) $\beta = 0.0$.

(b) $\beta = 0.1$.

(c) $\beta = 0.5$.

(d) $\beta = 0.9$.

(e) $\beta = 1.0$.

Figure 11: Impact of the number of scenarios in the power bought/sold by contracts (first-stage variable). Case 1, Week 1. Negative scale is used for convenience.

Figure 12: Impact of the number of scenarios in the power bought/sold by contracts (first-stage variable). Case 1, Week 2. Negative scale is used for convenience.

value arises because the stochastic solution is obtained from a solution of a decomposition algorithm that is not able to close the gap. Note that for risk-neutral formulations with uncertainty parameters only in the objective function the VSS is equal to zero (Birge and Louveaux, 2011), which is not exactly the case considered in this work because of the wind power uncertainty, and due to the risk measure in the objective function. Therefore, the results indicate that for the considered data, there is limited value in the risk-neutral formulations. However, for risk averse formulations the nonzero VSS in Table 13 indicate the added value of the risk-averse formulations. For example, Case 1, Week 1, for $\beta = 0.0$ has VSS= 0, while for $\beta = 1.0$ has VSS= 882,225. This means that as the objective function formulation deviates from the maximization of the expected profit to the maximization of the conditional value at risk of the profit, the resulting solutions become more distant from the solution of the deterministic problem.

6.4. Risk management analysis

In this section, we extend the risk analysis to study the impact of the risk parameters β and α on: a) the objective function value; b) the expected value of the profit; and c) the CVaR of the profit. Figures 13 and 14 report the impact of β for two specific cases. Note that for $\beta = 0.0$, the value of the objective function is equal to the expected value of the profit, and for $\beta = 1.0$ the value of the objective function is equal to the CVaR of the profit. In Figure 13, we can observe the trade-off obtained when more weight is put on the maximization of the CVaR, which corresponds to an increase of the CVaR from \$1,196,242 (with $\beta = 0.0$) to \$2,078,467 ($\beta = 1.0$) and a decrease on the expected value of the profit from \$2,998,459 to \$2,510,288. In Figure 14, the maximization of the expected profit without considering the CVaR leads to -\$297,089, while with $\beta = 0.1$ the $\text{CVaR}_{1-0.9}$ increases to \$240,895, while the expected value of the profit decreases from \$348,700 to \$271,655.

Figures 15 and 16 show the results obtained for Case 1, Week 1 and Case 2, Week 1,

Figure 13: Objective function value, expected value of the profit, and conditional value at risk of the profit as a function of β , for Case 1, Week 1, $\alpha = 0.9$ and 5,100 scenarios.

respectively, with the second risk-averse formulation with $\beta = 1.0$. In this formulation, the objective function involves only the maximization of the CVaR of the profit parameterized over α . We can see these charts as an extension of the ones in Figures 13 to 14, since they extend the analysis for $\beta = 1.0$ across different values of α . These charts provide also a relevant information about the expected value of the profit when only the profit of the worst scenario is maximized, which is obtained for $\alpha = (|\mathcal{S}| - 1) / |\mathcal{S}|$. For Case 1, Week 1, if the formulation focuses on maximizing only the worst profit, one obtains the value of \$1,954,807, which leads to an expected value of the profit of \$2,510,288. A similar analysis can be made for Case 2, Week 1, where if we only maximize the expected value of the profit, we obtain an expected value of \$348,700 and a $\text{CVaR}_{1-0.9}$ of -\$297,089, while if the $\text{CVaR}_{1-\alpha \rightarrow 1}$ is maximized, then the expected value of the profit obtained is \$274,471 and the worst profit is \$224,171.

The results presented in this section complement the results in Section 6.2.3, where we analyze the impact of β on the objective function and on the first-stage variables. The results in Section 6.2.3 show that the risk criterion influences the decisions concerning contracts, while risk-neutral solutions direct to buy electricity through contracts and sell the electricity bought and the one generated in the pool. Additionally, more risk-averse solutions point to sell the electricity generated through contracts, see Figures 11 and 12. The values of the CVaR of the profit in risk-neutral solutions show that the VPP is more exposed to lower profits, because the first-stage decision concerning the contracts make the revenues more dependent on the outcomes of the electricity prices in the pool. On the other hand, in risk-averse solutions, a higher CVaR of the profit, by comparison with the risk-neutral, indicates that the VPP is comparatively more protected against low profits because the revenues depend more on the contracts.

Figure 14: Objective function value, expected value of the profit, and conditional value at risk of the profit as a function of β , for Case 2, Week 1, $\alpha = 0.9$ and 5,100 scenarios.

Figure 15: Objective function value, expected value of the profit, and conditional value at risk of the profit as a function of α , for Case 1, Week 1, $\beta = 1.0$ and 5,100 scenarios.

Figure 16: Objective function value, expected value of the profit, and conditional value at risk of the profit as a function of α , for Case 2, Week 1, $\beta = 1.0$ and 5,100 scenarios.

7. Conclusions

We investigated risk-neutral and risk-averse stochastic programming formulations and solution methods for the optimal scheduling and market involvement of a VPP. The random variables are the wind speed (affecting the constraints) and the electricity prices in the pool (affecting the objective function). The wind speed uncertainty is characterized by a wind ensemble with 51 members, and the electricity prices by forecast scenarios generated from an ARIMA model.

We consider two risk-averse formulations, the first maximizes a combination of expected profit and conditional value at risk of the profit, and the second maximizes the conditional

value at risk of the profit. The first needs to be parameterized over the weights of the expected value and conditional value at risk and the quantile considered by the CVaR, while the second formulation just requires the quantile specification. The solution methods implemented encompass the direct solution of the extensive form and decomposition methods. We proposed two algorithms, Algorithm 1 and Algorithm 2, which are extensions of the L-Shaped method that can handle risk-averse formulations involving the conditional value at risk. Comparing with the L-Shaped method, these two algorithms include an additional outer-approximation of the conditional value at risk in the objective function, resulting in additional optimality cuts in the master problem. One of the main differences between these two algorithms is that Algorithm 1 considers the value at risk as a first-stage variable, while Algorithm 2 considers it as a second-stage variable.

The results show that the direct solution of the extensive form can solve the largest risk-neutral formulations tested within the time limits imposed. Furthermore, the direct solution of the risk-neutral extensive forms can compete in terms of computational time with the decomposition methods that solve the subproblems in sequence. However, the direct solution of the risk-averse formulations fails for problems with more than 510 scenarios. The good performance of the direct solution of the extensive forms results from the current state-of-the-art of MILP solvers and their increasing capability to solve large scale problems.

The decomposition algorithms can handle the risk-averse formulations efficiently if the subproblems are solved in parallel with an efficient approach. Algorithm 1 is more efficient than Algorithm 2, and the type of cuts - single or multiple - does not have a significant impact on the performance of Algorithm 1. The higher efficiency of Algorithm 1 results from the smaller number of iterations of its decomposition algorithm, which is due to the quality of the CVAR optimality cuts added to the master problem. These CVAR optimality cuts depend on the value at risk determined at each iteration, which converges to the optimal value faster in Algorithm 1 than in Algorithm 2. Regarding Algorithm 2, the results suggest that the multiple cut version of the CVaR2 optimality cuts are needed to ensure adequate performance. This relative performance between the algorithms is valid over the two cases, each of the two weeks, and the two risk averse formulations with different risk aversion parameters.

The risk-averse formulation that maximizes the conditional value at risk has the advantage of just requiring the definition of one parameter, but, the results demonstrate that it is not computationally superior to the other risk-averse formulation.

The impact of the number of scenarios in the objective function and in the first-stage variables is analyzed for different values of the parameter that defines the weight of the expected profit and conditional value at risk. The results show that for the risk-neutral formulation above 510 scenarios the number of scenarios does not affect the value of the first-stage variables. However, increasing the weight of the conditional value at risk in the objective function, the stability of the solution decreases. A weight of 0.1 of the conditional value at risk and different numbers of scenarios introduce a considerable instability in one of the first-stage variables, but, in general for larger weights of the conditional value at risk, the instability exists but with less variations.

In terms of risk management analysis, a relevant outcome is obtained with one specific instance, Case 2, Week 1, where we compare the expected value of the profit resulting from two different formulations: 1) maximization of the expected profit; and 2) maximization of the worst profit scenario. The difference between the expected profit of 1) and 2) is smaller than the difference between the conditional value at risk of 1) and 2), which means that a significant risk decrease is achieved at the cost of a small decrease in expected profit.

The risk-averse formulations include a risk measure in the objective function and a corresponding parameterization, which distorts the value of the objective function. Therefore, we suggest that whenever risk measures are included in the objective function, the analysis and comparison of different solutions should be based on the values and sensitivities of the first-stage variables, and not only on the objective function value. This suggestion arises from the fact that while different parameterizations of the objective function and different sets of scenarios may lead to solutions with different objective function values, these solutions may have similar first-stage variables.

Future work will involve the comparison of risk-averse stochastic programming and robust optimization from the point of view of uncertainty quantification, risk management, formulations, solution approaches and quantities of interest.

Acknowledgements

The authors acknowledge the support of the Center for Uncertainty Quantification in Computational Science & Engineering at KAUST.

Appendix A. Deterministic model

In this appendix, we describe the specific deterministic model addressed in this work. The goal of the problem is to maximize the operational profit defined as

$$P = \max \sum_f \sum_j \left[\left(\lambda_{f,j}^{sell} f_{f,j}^{sell} - \lambda_{f,j}^{buy} f_{f,j}^{buy} \right) D_f \right] + \sum_t \left[\lambda_t \left(p_t^{sell} - p_t^{buy} \right) \right] - cop. \quad (A.1)$$

$$\sum_{i \in TH} p_{i,t} + \sum_{i \in HY} p_{i,t}^{buy} + p_t^{buy} + \sum_f \sum_j f_{f,j}^{buy} + w_t = \sum_{i \in HY} p_{i,t} + p_t^{sell} + \sum_f \sum_j f_{f,j}^{sell}, \quad \forall t. \quad (A.2)$$

The modeling of the contracts is based on the framework developed in Conejo et al. (2008):

$$f_{f,j}^{buy} \leq F_{f,j}^{buy} \quad \forall f, j, \quad (A.3a)$$

$$f_{f,j}^{sell} \leq F_{f,j}^{sell} \quad \forall f, j, \quad (A.3b)$$

$$\sum_j f_{f,j}^{buy} \leq \sum_j F_{f,j}^{buy} y_f^{buy} \quad \forall f, \quad (A.3c)$$

$$\sum_j f_{f,j}^{sell} \leq \sum_j F_{f,j}^{sell} y_f^{sell} \quad \forall f, \quad (A.3d)$$

$$y_f^{sell} + y_f^{buy} \leq 1 \quad \forall f. \quad (\text{A.3e})$$

The feasible region of operation of the system depends on the features of the thermal and hydro units. The model of the thermal unit is based on the models from Arroyo and Conejo (2000) and Ostrowski et al. (2012):

$$\sum_{tt \geq t - UT_i + 1, tt \leq t} u_{i,tt}^{up} \leq u_{i,t} \quad \forall i, t \geq LM_i + 1, \quad (\text{A.4})$$

$$u_{i,t} + \sum_{tt \geq t - DT_i + 1, tt \leq t} u_{i,tt}^{dn} \leq 1 \quad \forall i, t \geq FM_i + 1, \quad (\text{A.5})$$

where $LM_i = \min\{|T|, U_i\}$, U_i is the number of hours generator i needs to be on at the beginning of the time horizon, $FM_i = \min\{|T|, DM_i\}$, DM_i denotes the number of hours the generator needs to be off at the beginning of the time horizon.

$$P_i^l u_{i,t} \leq p_{i,t} \leq P_i^u u_{i,t} \quad \forall i \in TH, t, \quad (\text{A.6})$$

$$p_{i,t} \leq P0_i + RU_i U0_i + SU_i u_{i,t}^{up} \quad \forall i \in TH, t = 1, \quad (\text{A.7})$$

$$p_{i,t} - p_{i,t-1} \leq RU_i u_{i,t-1} + SU_i u_{i,t}^{up} \quad \forall i \in TH, t > 1, \quad (\text{A.8})$$

$$p_{i,t-1} - p_{i,t} \leq RD_i u_{i,t} + SD_i u_{i,t}^{dn} \quad \forall i \in TH, t > 1, \quad (\text{A.9})$$

$$cop = \sum_{i \in TH} \sum_t cu_{i,t} + A_i u_{i,t} + B_i p_{i,t} + cd_{i,t}, \quad (\text{A.10})$$

$$cd_{i,t} \geq DC_i (1 - u_{i,t}) \quad \forall i, t = 1, T_i^I > 0, \quad (\text{A.11})$$

$$cd_{i,t} \geq u_{i,t}^{dn} DC_i \quad \forall i, t \geq 2, \quad (\text{A.12})$$

$$cu_{i,t} \geq u_{i,t}^{up} HS_i, \quad \forall i, t, \quad (\text{A.13})$$

$$cu_{i,t} \geq \left(u_{i,t} - \sum_{tt \geq t - (DT_i + T_i^c + 1)}^{t-1} u_{i,tt} \right) CS_i \quad \forall i, t > TD_i + T_i^c, \quad (\text{A.14})$$

$$cu_{i,t} \geq \left(u_{i,t} - \sum_{tt < t} u_{i,tt} \right) CS_i, \quad \forall i, T_i^I < 0, (TD_i + T_i^c + T_i^I + 1) < t \leq (TD_i + T_i^c), \quad (\text{A.15})$$

$$1 - u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t = 1, T_i^I > 0, \quad (\text{A.16})$$

$$-u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t = 1, T_i^I < 0, \quad (\text{A.17})$$

$$u_{i,t-1} - u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t > 1, \quad (\text{A.18})$$

$$v_{i,t} - G(-q_{i,t} + qp_{i,t}) = V0_i + GQ_i^{in} \quad \forall i \in HY, t = 1, \quad (\text{A.19})$$

$$v_{i,t} - v_{i,t-1} - G(-q_{i,t} + qp_{i,t}) = GQ_i^{in} \quad \forall i \in HY, t > 1, \quad (\text{A.20})$$

$$ptb_{i,t} - K_i^t q_{i,t} H_i = 0 \quad \forall i \in HY, t, \quad (\text{A.21})$$

$$pp_{i,t} - K_i^p qp_{i,t} H_i = 0 \quad \forall i \in HY, t. \quad (\text{A.22})$$

$$q_{i,t} \leq Q_i^u \quad \forall i \in HY, t, \quad (\text{A.23})$$

$$qp_{i,t} \leq Q_i^u \quad \forall i \in HY, t, \quad (\text{A.24})$$

$$V_{i,t}^l \leq v_{i,t} \leq V_{i,t}^u \quad \forall i \in HY, t, \quad (\text{A.25})$$

$$v_{i,t} \geq V_i^E \quad \forall i \in HY, t = tf, \quad (\text{A.26})$$

$$f_{f,j}^{buy}, f_{f,j}^{sell} \geq 0, \quad \forall f, j, \quad (\text{A.27})$$

$$y_f^{sell}, y_f^{buy} \in \{0, 1\}, \quad \forall f, \quad (\text{A.28})$$

$$u_{i,t}, u_{i,t}^{up}, u_{i,t}^{dn} \in \{0, 1\}, \quad p_{i,t} \geq 0, \quad \forall i \in TH, t, \quad (\text{A.29})$$

$$p_t^{sell}, p_t^{buy} \geq 0, \quad \forall t, \quad (\text{A.30})$$

$$ptb_{i,t}, pp_{i,t}, q_{i,t}, qp_{i,t}, v_{i,t} \geq 0, \quad \forall i \in HY, t. \quad (\text{A.31})$$

The model for the thermal unit includes: a) minimum up and down time limits, Eqs. (A.4), and (A.5), see Rajan and Takriti (2005); b) bounds on the power output, Eqs. (A.6) to (A.9), see Ostrowski et al. (2012); c) modeling of costs related with operation, startups, and shutdowns, Eqs. (A.10) to (A.15), see Arroyo and Conejo (2000); Morales-Espana et al. (2013); Nowak and Romisch (2000); and d) logical relations between the binary variables that define the states of the unit, Eqs. (A.16) to (A.18). The model of the pumped-storage hydro unit encompasses: a) water mass balances for the reservoir at the end of each period, Eqs. (A.19), and (A.20); and b) power generation and consumption functions, Eqs. (A.21), and (A.22). These results in an MILP model, which in some cases is a very tight model as demonstrated in Lima and Novais (2016). The values of the parameters of the thermal unit, pumped-storage hydro unit, and contracts are available in the Tables A.14, A.15, and A.16, respectively.

Table A.14: Base data for the thermal generators.

Case	P^L (MW)	P^U (MW)	UT (h)	DT (h)	T^c (h)	T^I (h)	SU/SD (MW/h)	RU/RD (MW/h)	A_i (\$/h)	B_i (\$/MWh)	HS_i (\$/h)	CS_i (\$/h)
G1	150	455	8	8	5	8	150	91	1000	16.19	4500	9000
G2	10	55	1	1	0	-1	10	11	660	25.92	30	60

Table A.15: Data for the hydro plant.

	H_i (m)	K_i^p (MWs/m ³)	K_i^t (MWs/m ³)	Q_i^{in} (m ³ /s)	Q_i^u (m ³ /s)	$V0_i$ (Hm ³)	V_i^u (Hm ³)	V_i^l (Hm ³)
Week 1	113	0.99	1	0.1	46.5	84	560	10
Week 2	113	0.99	1	5.7	46.5	420	560	10

Table A.16: Price of the energy (\$/MWh) for two contracts for each week.

Week	Contract	Block Size	Sell blocks			Buy blocks		
			3	2	1	1	2	3
1	1	50	64.77	59.77	54.77	54.77	49.77	44.77
1	2	55	58.29	53.79	49.29	49.29	44.79	40.29
2	1	50	36.15	41.15	46.15	46.15	51.15	56.15
2	2	55	32.53	37.03	41.53	41.53	46.03	50.53

Nomenclature

Sets

F	Forward contracts
HY	Hydro pump-storage generation units
J	Blocks of the forward contracts
I	Generating units
T	Time periods
TH	Thermal generation units

Parameters

A_i, B_i	Production cost function coefficients for unit i (\$/h)
CS_i	Cold start-up cost of unit i (\$/h)
DM_i	Number of periods unit i must be off at the beginning of the time horizon
D_f	Time periods spanned by contract f
DC_i	Shut-down cost (\$)
DT_i	Minimum down time of unit i (h)
FM_i	Minimum number of periods a unit i must be off at the beginning of the time horizon
HS_i	Hot start cost of unit i (\$/h)
LM_i	Minimum number of periods a unit i must be on at the beginning of the time horizon
P_i^l	Minimum power output of unit i (MW)
P_i^u	Maximum power output of unit i (MW)
$P0_i$	Power produced at $t=0$ by unit i (MW)
RD_i	Maximum ramp-down rate of unit i (MW)
RU_i	Maximum ramp-up rate of unit i (MW)
SD_i	Maximum shutdown rate of unit i (MW)
SR_t	Spinning reserve for period t (MW)
SU_i	Maximum start-up rate of unit i (MW)
U_i	Number of periods unit i must be on at the beginning of the time horizon
$U0_i$	Initial state of unit i {on,off}={1,0}
UT_i	Minimum up time of unit i (h)
T_i^c	Cold start hours of unit i (h)
T_i^I	Initial status of unit i (h)
G	Conversion factor between Hm^3 and m^3/s in one hour
H_i	Water head in plant i (m)

K_i^p	Power consumption factor
K_i^t	Power generation factor
Q_i^{in}	Natural inflow of water for plant i (m ³ /s)
Q_i^u	Maximum turbined and pumped flow of water for plant i (m ³ /s)
V_i^u	Maximum volume of water in the reservoir of plant i (Hm ³)
V_i^l	Minimum volume of water in the reservoir of plant i (Hm ³)
V_i^E	Minimum volume of water in the reservoir of plant i at the of the horizon (Hm ³)
$\lambda_{f,j}^{buy}$	Energy price of buying block j of forward contract f (\$/MWh)
$\lambda_{f,j}^{sell}$	Energy price of selling block j of forward contract f (\$/MWh)
Continuous variables	
$cd_{i,t}$	Shut-down cost of unit i in period t (\$)
cop	Total startup, shutdown, production, and online cost of unit i (\$)
cp	Total startup, shutdown and online cost of unit i (\$)
$cu_{i,t}$	Startup cost of unit i in period t (\$)
$f_{f,j}^{buy}$	Power bought through block j of forward contract f (MW)
$f_{f,j}^{sell}$	Power sold through block j of forward contract f (MW)
P	Operational profit of the producer per week (\$)
$p_{i,t}$	Power output of unit i in period t (MW)
$p_{i,t}^{buy}$	Power bought in the pool in period t (MW)
p_t^{sell}	Power sold in the pool in period t (MW)
$ptb_{i,t}$	Power output of the pumped-storage hydro unit i in period t (MW)
$pp_{i,t}$	Power consumption of the pumped-storage hydro unit i in period t (MW)
$q_{i,t}$	Turbined flow of water in plant i in period t (m ³ /s)
$qp_{i,t}$	Pumped flow of water in plant i in period t (m ³ /s)
$v_{i,t}$	Volume of water stored in the reservoir of plant i (Hm ³)
Binary variables	
$u_{i,t}$	On/off status of unit i in period t
$u_{i,t}^{up}$	Startup status of unit i in period t
$u_{i,t}^{dn}$	Shutdown status of unit i in period t
y_f^{buy}	Selection of forward contract f to buy energy
y_f^{sell}	Selection of forward contract f to sell energy
Random variables	
w_t	Wind power output in period t (MW)
λ_t	Pool price in period t (\$/MWh)

Appendix B. ARIMA model for the electricity prices

In this appendix, we describe the ARIMA models developed to fit the historical data of the electricity prices. The application of these models to energy systems is described in Conejo et al. (2010), and a review of ARIMA and alternative methods to forecast electricity prices is given in Weron (2014). We follow the well known approach of Box-Jenkins (Box and Jenkins, 1990) to fit time series, which is based on 1) the identification of the ARIMA model; 2) the estimation of the parameters; 3) the assessment of the quality of the model; and 4) in case the fit is not adequate return to the first step. After an adequate model is found, it is straightforward to forecast the electricity prices by sampling the error of the model. This error is considered to

be white noise with zero mean. We implemented this approach in MatLab using the toolbox ECOTOOL (Pedregal et al., 2012) for the identification, estimation, and assessment of the ARIMA models. The main reason to use ECOTOOL is that it provides the flexibility to test ARIMA models with multiplicative seasonal polynomials using a friendly notation. The time series with the electricity data were obtained from the Iberian Peninsula electricity market (Iberian Electricity Market, 2015), for two different time periods. For both time periods, the time series are not stationary, and therefore, we use differentiation to obtain a constant mean. The selected time series posed some difficulties in the identification of adequate models that would involve uncorrelated errors, which led to some complex models with several parameters. The identified models are presented next for both weeks.

The first time period involves a time series from June 2, 2014 to August 24, 2014, and a forecast for August 25-31, 2014. For this period, the model estimated is the following

$$\tilde{c}_t = C + \frac{\Theta_1(B)\Theta_2(B)}{(1-B)\Phi_1(B)\Phi_2(B)}\epsilon_t, \quad \forall t \in \mathcal{T}, \quad (\text{B.1})$$

where $\tilde{c}_t$ are the observed prices, ϵ_t is the error considered as white noise, and Θ_1 , Θ_2 , Φ_1 , and Φ_2 are given by:

$$\begin{aligned} \Theta_1(B) = 1 - \theta_2 B^2 - \theta_4 B^4 - \theta_7 B^7 - \theta_9 B^9 - \theta_{11} B^{11} - \theta_{14} B^{14} - \theta_{15} B^{15} - \theta_{16} B^{16} \\ - \theta_{17} B^{17} - \theta_{30} B^{30} - \theta_{49} B^{49} - \theta_{73} B^{73} - \theta_{94} B^{94}, \end{aligned} \quad (\text{B.2})$$

$$\Theta_2(B) = 1 - \theta_{24} B^{24} - \theta_{168} B^{168}, \quad (\text{B.3})$$

$$\Phi_1(B) = 1 - \phi_1 B, \quad (\text{B.4})$$

$$\Phi_2(B) = 1 - \phi_{24} B^{24} - \phi_{168} B^{168}, \quad (\text{B.5})$$

where B is the backshift operator defined as $B^n c_t := c_{t-n}$.

The second period consists of a time series from September 1, 2014 to November 23, 2014, and a forecast for November 24-30, 2014, which leads to the following model:

$$\tilde{c}_t = C + \frac{\Theta_1(B)\Theta_2(B)}{(1-B)\Phi_1(B)\Phi_2(B)}\epsilon_t, \quad \forall t \in \mathcal{T}, \quad (\text{B.6})$$

where Θ_1 , Θ_2 , Φ_1 , and Φ_2 are given by:

$$\Theta_1 = 1 - \theta_2 B^2 - \theta_4 B^4 - \theta_6 B^6 - \theta_7 B^7 - \theta_8 B^8 - \theta_{13} B^{13} - \theta_{14} B^{14} - \theta_{15} B^{15} - \theta_{17} B^{17} - \theta_{30} B^{30} - \theta_{44} B^{44}, \quad (\text{B.7})$$

$$\Theta_2 = 1 - \theta_{24} B^{24}, \quad (\text{B.8})$$

$$\Phi_1 = 1 - \phi_1 B - \phi_6 B^6 - \phi_7 B^7 - \phi_8 B^8 - \phi_{13} B^{13} - \phi_{14} B^{14} - \phi_{15} B^{15} - \phi_{17} B^{17} - \phi_{30} B^{30} - \phi_{44} B^{44}, \quad (\text{B.9})$$

$$\Phi_2 = 1 - \phi_{24} B^{24} - \phi_{167} B^{167} - \phi_{168} B^{168} - \phi_{169} B^{169} - \phi_{336} B^{336} - \phi_{504} B^{504} - \phi_{672} B^{672}. \quad (\text{B.10})$$

References

References

- Ahmed, S., 2006. Convexity and decomposition of mean-risk stochastic programs. *Math. Program.* 106 (3, Ser. A), 433–446.
- Arroyo, J. M., Conejo, A. J., 2000. Optimal response of a thermal unit to an electricity spot market. *IEEE T. Power Syst.* 15 (3), 1098–1104.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance* 9 (3), 203–228.
- Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. *Math. Oper. Res.* 23 (4), 769–805.
- Benders, J., 1962. Partitioning procedures for solving mixed-variables programming problems. *Numer. Math.* 4, 238–252.
- Bertsimas, D., Brown, D. B., Caramanis, C., 2011. *Theory and Applications of Robust Optimization*. SIAM Rev. 53 (3), 464–501.
- Birge, J. R., Louveaux, F., 2011. *Introduction to stochastic programming*. Springer.
- Birge, J. R., Louveaux, F. V., 1988. A multicut algorithm for two-stage stochastic linear programs. *European J. Oper. Res.* 34 (3), 384–392.
- Box, G. E. P., Jenkins, G., 1990. *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated.
- Buizza, R., 2014. The TIGGE medium-range, global ensembles. Tech. rep., ECMWF, Shinfield Park, Reading RG2-9AX, UK.
- Bussieck, M. R., Ferris, M. C., Lohmann, T., 2012. *Algebraic Modeling Systems: Modeling and Solving Real World Optimization Problems*. Springer, Berlin Heidelberg, Ch. GUSS: Solving Collections of Data Related Models within GAMS, pp. 35–56.
- Bussieck, M. R., Ferris, M. C., Meeraus, A., SUM 2009. Grid-Enabled Optimization with GAMS. *INFORMS J. Comput.* 21 (3), 349–362, *INFORMS Annual Meeting 2006*, Pittsburgh, PA, 2006.
- Conejo, A., Carrion, M., Morales, J., 2010. *Decision Making Under Uncertainty in Electricity Markets*. Springer US.
- Conejo, A. J., Garcia-Bertrand, R., Carrion, M., Caballero, A., de Andres, A., 2008. Optimal involvement in futures markets of a power producer. *IEEE T. Power Syst.* 23 (2), 703–711.
- Dabbagh, S. R., Sheikh-El-Eslami, M. K., 2015. Risk-based profit allocation to DERs integrated with a virtual power plant using cooperative Game theory. *Electr. Pow. Syst. Res.* 121, 368–378.
- Dabbagh, S. R., Sheikh-El-Eslami, M. K., 2016. Risk Assessment of Virtual Power Plants Offering in Energy and Reserve Markets. *IEEE T. Power Syst.* 31 (5), 3572–3582.
- de Oliveira, W., Sagastizbal, C., 2014. Level bundle methods for oracles with on-demand accuracy. *Optimization Methods and Software* 29 (6), 1180–1209.
- Fabian, C. I., December 2008. Handling cvar objectives and constraints in two-stage stochastic models. *Eur. J. Oper. Res.* 191 (3), 888–911.
- Fabian, C. I., Wolf, C., Koberstein, A., Suhl, L., 2015. Risk-averse optimization in two-stage stochastic models: computational aspects and a study. *SIAM J. Optimiz.* 25 (1), 28–52.
- Homem-de Mello, T., Pagnoncelli, B. K., 2016. Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *Eur. J. Oper. Res.* 249 (1), 188–199.
- Iberian Electricity Market, 2015. www.omie.es.
- Kall, P., Mayer, J., 2011. *Stochastic Linear Programming*, 2nd Edition. Vol. 156 of *International Series in Operations Research & Management Science*. Springer US.

- Kardakos, E. G., Simoglou, C. K., Bakirtzis, A. G., 2016. Optimal Offering Strategy of a Virtual Power Plant: A Stochastic Bi-Level Approach. *IEEE T. Smart Grid* 7 (2, SI), 794–806.
- Künzi-Bay, A., Mayer, J., 2006. Computational aspects of minimizing conditional value-at-risk. *Comput. Manag. Sci.* 3 (1), 3–27.
- Lemaréchal, C., Nemirovskii, A., Nesterov, Y., 1995. New variants of bundle methods. *Math. Program.* 69 (1), 111–147.
- Lima, R. M., 2010. IBM ILOG CPLEX-What is inside of the box? EWO Seminar, Carnegie Mellon University, online access on 21-08-2016: http://egon.cheme.cmu.edu/ewocp/docs/rlima_cplex_ewo_dec2010.pdf.
- Lima, R. M., Grossmann, I. E., 2011. Chemical Engineering Greetings to prof. Sauro Pierucci. AIDIC, Ch. Computational advances in solving mixed integer linear programming problems, pp. 151–160.
- Lima, R. M., Grossmann, I. E., 2016. On the solution of nonconvex cardinality boolean quadratic programming problems: a computational study. *Comput. Optim. Appl.*, 1–37.
- Lima, R. M., Novais, A. Q., 2016. Symmetry breaking in MILP formulations for Unit Commitment problems. *Comput. Chem. Eng.* 85, 162–176.
- Lima, R. M., Novais, A. Q., Conejo, A. J., 2015. Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. An adaptive robust optimization approach. *Eur. J. Oper. Res.* 240 (2), 457–475.
- Lombardi, P., Powalko, M., Rudion, K., 2009. Optimal operation of a virtual power plant. In: 2009 IEEE Power Energy Society General Meeting. pp. 1–6.
- Miller, N., Ruszczyński, A., 2011. Risk-Averse Two-Stage Stochastic Linear Programming: Modeling and Decomposition. *Oper. Res.* 59 (1), 125–132.
- Moazeni, S., Powell, W. B., Hajimiragha, A. H., 2015. Mean-Conditional Value-at-Risk Optimal Energy Storage Operation in the Presence of Transaction Costs. *IEEE T. Power Syst.* 30 (3), 1222–1232.
- Moghaddam, I. G., Nick, M., Fallahi, F., Sanei, M., Mortazavi, S., 2013. Risk-averse profit-based optimal operation strategy of a combined wind farm-cascade hydro system in an electricity market. *Renew. Energ.* 55, 252–259.
- Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., Zugno, M., 2014. Integrating Renewables in Electricity Markets. Springer.
- Morales-Espana, G., Latorre, J., Ramos, A., 2013. Tight and compact MILP formulation for the thermal unit commitment problem. *IEEE T. Power Syst.* 28 (4), 4897–4908.
- Nowak, M. P., Romisch, W., 2000. Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Ann. Oper. Res.* 100, 251–272.
- Noyan, N., 2012. Risk-averse two-stage stochastic programming with an application to disaster management. *Computers & Operations Research* 39 (3), 541 – 559.
- Ostrowski, J., Anjos, M. F., Vannelli, A., 2012. Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE T. Power Syst.* 27 (1), 39–46.
- Pandzic, H., Kuzle, I., Capuder, T., 2013a. Virtual power plant mid-term dispatch optimization. *Appl. Energ.* 101 (SI), 134–141.
- Pandzic, H., Morales, J. M., Conejo, A. J., Kuzle, I., 2013b. Offering model for a virtual power plant based on stochastic programming. *Appl. Energ.* 105, 282–292.
- Pedregal, D. J., Contreras, J., de la Nieta, A. A. S., 2012. ECOTOOL: A general MATLAB Forecasting Toolbox with Applications to Electricity Markets. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 151–171.
- Pflug, C., 2000. Probabilistic Constrained Optimization: Methodology and Applications. Kluwer Academic Publishers, Dordrecht, Ch. Some remarks on the value-at-risk and the conditional value-at-risk,

- pp. 272–281.
- Pineda, S., Conejo, A. J., Carrion, M., 2008. Impact of unit failure on forward contracting. *IEEE T. Power Syst.* 23 (4), 1768–1775.
- Rajan, D., Takriti, S., 2005. Minimum up/down polytopes of the unit commitment problem with start-up costs. Tech. rep., IBM Research.
- Rockafellar, R., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26 (7), 1443 – 1471.
- Rockafellar, R. T., 2007. Coherent Approaches to Risk in Optimization Under Uncertainty. *INFORMS*, Ch. 4, pp. 38–61.
- Rockafellar, R. T., Uryasev, S., 2000. Optimization of conditional value-at-risk. *Journal of Risk* 2, 21–41.
- Ross, S. M., 1987. *Introduction to Probability and Statistics for Engineers and Scientists*. John Wiley & Sons.
- Sarykalin, S., Serraino, G., Uryasev, S., 2008. Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. *INFORMS*, Ch. Chapter 13, pp. 270–294.
- Schultz, R., Tiedemann, S., 2006. Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Math. Program.* 105 (2-3, Ser. B), 365–386.
- Shabanzadeh, M., Sheikh-El-Eslami, M.-K., Haghifam, M.-R., 2015. The design of a risk-hedging tool for virtual power plants via robust optimization approach. *Appl. Energ.* 155, 766–777.
- Shapiro, A., Dentcheva, D., Ruszczyński, A., 2009. *Lectures on stochastic programming : modeling and theory*. MPS-SIAM Series on Optimization.
- Tajeddini, M. A., Rahimi-Kian, A., Soroudi, A., 2014. Risk averse optimal operation of a virtual power plant using two stage stochastic programming. *Energy* 73, 958–967.
- Van Slyke, R., Wets, R., 1969. L-shaped linear programs with applications to control and stochastic programming. *SIAM Journal on Applied Mathematics* 17, 638–663.
- Weron, R., 2014. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecasting* 30 (4), 1030 – 1081.
- Wollmer, R. D., 1980. Two-stage linear programming under uncertainty with 0 – 1 integer first stage variables. *Math. Programming* 19 (3), 279–288.
- Zamani, A. G., Zakariazadeh, A., Jadid, S., 2016. Day-ahead resource scheduling of a renewable energy based virtual power plant. *Appl. Energ.* 169, 324–340.
- Zhang, W., Rahimian, H., Bayraksan, G., 2016. Decomposition Algorithms for Risk-Averse Multistage Stochastic Programs with Application to Water Allocation under Uncertainty. *INFORMS J. Comput.* 28 (3), 385–404.