[EURO Journal on Computational Optimization 9 (2021) 100005](https://doi.org/10.1016/j.ejco.2021.100005)

Contents lists available at [ScienceDirect](http://www.ScienceDirect.com/)

EURO Journal on Computational Optimization

journal homepage: [www.elsevier.com/locate/ejco](http://www.elsevier.com/locate/ejco)

Sample average approximation for risk-averse problems: A virtual power plant scheduling application

Ricardo M. Lima[a](#_bookmark0),[∗](#_bookmark5), Antonio J. Conejo[b](#_bookmark1), Loïc Giraldi [a](#_bookmark0),[e](#_bookmark4), Olivier Le Maître [c](#_bookmark2), Ibrahim Hoteit[d](#_bookmark3),

Omar M. Knio[a](#_bookmark0)

a *Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi* *Arabia*

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

c *Centre de Mathématiques Appliquées, CNRS, Inria, Ecole Polytechnique, Palaiseau, France*

d *Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Saudi Arabia*

e *CEA, DES, IRESNE, DEC, Cadarache F-13108 Saint-Paul-Lez-Durance, France*

a r t i c l e i n f o a b s t r a c t

*Keywords:*

Sample average approximation

Risk-averse stochastic programming Virtual power plant

In this paper, we address the decision-making problem of a virtual power plant (VPP) involving a self-scheduling and market involvement problem under uncertainty in the wind speed and electricity prices. The problem is modeled using a risk-neutral and two risk-averse two-stage stochastic programming formulations, where the conditional value at risk is used to represent risk. A sample average approximation methodology is integrated with an adapted L-Shaped solution method, which can solve risk-neutral and specific risk-averse problems. This methodology provides a framework to understand and quantify the impact of the sample size on the variability of the results. The numerical results include an analysis of the computational performance of the methodology for two case studies, estimators for the bounds of the true optimal solutions of the problems, and an assessment of the quality of the solutions obtained. In particular, numerical experiences indicate that when an adequate sample size is used, the solution obtained is close to the optimal one.

# Introduction

The optimal operation and electricity market involvement of virtual power plants (VPPs) is currently an active field of research. VPP refers to an aggregation of distributed energy resources (DER) that interacts with [the electricity market as a single entity (Awerbuch and Preston, 1997; Pudjianto et al., 2007). These DERs involve generators with small ca-](#_bookmark52) pacities or that face significant trading risks if operating by themselves. For example, a wind farm trading in an electricity market is subject to uncertain wind speeds, electricity prices, and imbalance costs, which may involve high risk. Aggregating multiple and diverse generators into a VPP creates an entity with a single power generation profile and a larger power capacity than the constituting units. A VPP has flexibility due to the complementarity of its diverse resources, and capacity to de- velop market intelligence to optimize its participation in the electricity market (through offers in the pool and by signing forwards contracts). However, to define a single power output profile, a self-scheduling prob- lem considering all constituting generators, their interactions, and their

generation constraints needs to be solved, which is more diﬃcult than solving individual generator problems. The VPP concept is particularly relevant for the integration of distributed renewable energy resources in power systems to mitigate flexibility limitation and power output un- certainty. Additional details on the technical and commercial functions of VPPs can be found in [Pudjianto et al. (2007)](#_bookmark84), [Morales et al. (2014)](#_bookmark74), and [Jansen et al. (2008)](#_bookmark59). In [Section 1.1](#_bookmark6), a review on VPP optimization is presented. Within the different aspects of a VPP operation, we are concerned with the optimal scheduling of the generation units and the interaction with the electricity market.

In the present work, we propose a sample average approximation [(SAA) methodology (Kleywegt et al., 2001; Shapiro and Homem-De- Mello, 2000) to solve the risk-averse stochastic programming problem](#_bookmark66) that describes the operation of a VPP. Three important aspects in the im- plementation of the SAA are studied: 1) the development of an eﬃcient solution methodology; 2) the impact of the sample size on the perfor- mance of the methodology; and 3) the determination of point estimates and confidence intervals for the solutions.

∗ Corresponding author.

[*E-mail addresses:* ricardo.lima@kaust.edu.sa (R.M. Lima), conejonavarro.1@osu.edu (A.J. Conejo), loic.giraldi@cea.fr (L. Giraldi), olivier.le- maitre@polytechnique.edu (O. Le Maître),](mailto:olivier.le-maitre@polytechnique.edu) [ibrahim.hoteit@kaust.edu.sa](mailto:ibrahim.hoteit@kaust.edu.sa) [(I. Hoteit),](mailto:olivier.le-maitre@polytechnique.edu) [omar.knio@kaust.edu.sa](mailto:omar.knio@kaust.edu.sa) [(O.M. Knio).](mailto:olivier.le-maitre@polytechnique.edu)

<https://doi.org/10.1016/j.ejco.2021.100005>

Received 22 October 2020; Received in revised form 7 February 2021; Accepted 11 March 2021

2192-4406/© 2021 The Authors. Published by Elsevier Ltd on behalf of Association of European Operational Research Societies (EURO). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* 1. *Literature review: optimization of VPPs*

In the literature, VPP models with distinct portfolios of generation units, different interactions with the electricity market, and, conse- quently, various sources of uncertainty can be found. In terms of inter- actions with the electricity market, the main features captured by VPP models include:

1. [forward contracts (Lima et al., 2015; Pandžić et al., 2013a; Shaban- zadeh et al., 2016);](#_bookmark67)
2. [offers to the day-ahead market (Baringo and Baringo, 2017; Dab- bagh and Sheikh-El-Eslami, 2016; Kardakos et al., 2016; Moghad- dam et al., 2013; Pandžić et al., 2013b; Rahmani-Dabbagh and Sheikh-El-Eslami, 2016; Shayegan-Rad et al., 2017; Tajeddini et al., 2014; Zamani et al., 2016);](#_bookmark53)
3. [offers/bids to the balancing market (Dabbagh and Sheikh-El-Eslami, 2016; Kardakos et al., 2016; Tajeddini et al., 2014);](#_bookmark60)
4. offers to spinning reserves ([Dabbagh and Sheikh-El-Eslami, 2016](#_bookmark60)); and
5. strategic offering considering the market clearing and other market players’ strategies ([Kardakos et al., 2016](#_bookmark64)).

The operational decision-making problem of VPPs involves decisions under uncertainty due to the nature of the renewable energy sources and electricity prices. Therefore, stochastic optimization approaches constitute a natural framework to address this type of problems. The most used approaches are based on stochastic programming and ro- bust optimization. Some examples include the application of risk-neutral [stochastic programming models (Lima et al., 2018; Pandžić et al., 2013a;](#_bookmark68)

in [Lima et al. (2018)](#_bookmark68) does not use an SAA approach, but rather focuses on risk-averse problems and their solutions using the L-Shaped method ([Van Slyke and Wets, 1969](#_bookmark85)). In that work, 14 variants of the L-Shaped method using combinations of single and multiple optimality cuts and eﬃcient parallel implementations are discussed, without considering SAA.

In this work, we propose a more elaborate approach that eﬃciently relies on multiple samples to assess solutions and determine confidence intervals on the relevant performance indices of the VPP operation. No such work is available for VPP applications.

* 1. *Literature review: sample average approximation*

The term sample average approximation is used in the literature to refer to a problem that approximates a stochastic optimization one ([Shapiro and Homem-De-Mello, 2000](#_bookmark82)), but it is also used to indicate a methodology where the solution of that approximate problem is per- formed multiple times and increasingly accurate solutions are obtained ([Kleywegt et al., 2001](#_bookmark66)). The repetition enables the calculation of con- fidence intervals of the optimal objective function value, and also the assessment of the solution.

The detailed characterization and statistical properties of the SAA methodology, namely consistency and rates of convergence to optimal objective function values and solution, have been stud- ied by [Mak et al. (1999)](#_bookmark69), [Shapiro and Homem-De-Mello (2000)](#_bookmark82), [Kleywegt et al. (2001)](#_bookmark66), and reviewed in detail in [Shapiro (2009)](#_bookmark76) and [Homem-de Mello and Bayraksan (2014)](#_bookmark70). In these works, expressions for the convergence to the optimal value and solution are developed and ex-

[2013b](#_bookmark60)[), risk-averse stochastic programming (Dabbagh and Sheikh-El-](#_bookmark68)

pressions relating the probability of an optimal solution *𝑥*∗

to be equal

[Eslami, 2016; Kardakos et al., 2016; Lima et al., 2018; Moazeni et al.,](#_bookmark60)

to the true solution *𝑥*∗

*𝑁*

with the sample size are established. However,

[2015; Moghaddam et al., 2013; Tajeddini et al., 2014), robust optimiza-](#_bookmark60) [tion (Lima et al., 2015; Rahimiyan and Baringo, 2016; Shabanzadeh et al., 2015), and recently, a hybrid approach based on stochastic pro-](#_bookmark67) gramming and robust optimization ([Baringo and Baringo, 2017](#_bookmark53)).

When risk-neutral or risk-averse stochastic programming models are used, the problem is solved for a random uncertainty sample. The main assumption is that the sample used is representative of the full distribu- tion of the uncertainties in the problem. In the works above, the sample size ranges from less than 100 elements to a maximum of 400, except in the work [Lima et al. (2018)](#_bookmark68) where samples of up to 25,000 elements were used. In some works, a reduced sample size was obtained using a scenario-reduction method ([Dupačová et al., 2003](#_bookmark61)). Typically, the size of the sample is limited by the capability to solve the problems using the corresponding extensive form and a branch and cut solver for mixed- integer linear programming (MILP) problems. Compared with the works cited above, our VPP model considers a time horizon of one week di- vided into periods of one hour, whereas most works above consider the 24 h of the next day. The time horizon of one week avoids the myopic view of a single-day horizon, which neglects the temporal coupling of some VPP constraints that condition next day operations. Therefore, this extended horizon enables considering constraints of thermal units, e.g., minimum up-time and minimum down-time, and water mass balances in the reservoir of hydro units across consecutive days, at the cost of a larger model.

We emphasize that risk-averse stochastic programming models are

more complex to solve than risk-neutral stochastic programming; there- fore, the impact of the sample size and the length of the time hori- zon on the computing time of risk-averse problems are comparatively more critical. Some works propose an out-of-sample analysis that pro- vides an additional characterization of the solution resulting from the stochastic programming model ([Baringo and Baringo, 2017](#_bookmark53)). Compared to an out-of-sample analysis that performs a study for a single solu- tion, the SAA methodology includes a first stage with multiple opti- mization replications using different samples to generate alternative solutions and to provide a point estimate of the upper bound on the true objective function value (for a maximization problem). The work

as stated in [Kleywegt et al. (2001)](#_bookmark66) and [Linderoth et al. (2006)](#_bookmark71), it is im- practical to determine some of the constants in those expressions. Fur- thermore, the expressions do not provide an estimate of the computing time required to obtain the approximated solutions and point estimates of the bounds. Therefore, computational experiments are essential to determine the trade-off between the computational resources and the accuracy obtained from the sample size. This type of analysis was per- formed, for example, in [Linderoth et al. (2006)](#_bookmark71) and [Verweij et al. (2003)](#_bookmark86). [Wang (2007)](#_bookmark88) and [Wang and Ahmed (2008)](#_bookmark89) were the first to propose SAA methods for risk-averse stochastic programming problems. They focused on min-max problems with expected value objectives and prob- lems with constraints involving expected values. For these problems, they proved that the results of their SAA problems converge exponen- tially fast to the true results of the stochastic problems as the sample

size increases.

A related method to this work, which has been extended to inte- grate an SAA methodology with risk measures, is stochastic dual dy- [namic programming (SDDP) (](#_bookmark83)[Pereira and Pinto, 1991](#_bookmark77)[). In Philpott and de Matos (2012) and](#_bookmark83) [Shapiro](#_bookmark83) [et al. (2013) are developed risk-](#_bookmark83) averse SDDP methods to address multistage linear stochastic program- ming models describing power systems planning problems. In risk- averse SDDP, the upper bound estimators are known to be weak ([Shapiro, 2011](#_bookmark78)), and in this regard, importance sampling was proposed in [Kozmík and Morton (2015)](#_bookmark72) to improve the quality of those estimators (for a minimization problem).

Regardless of the approach used, it is generally essential to develop an eﬃcient SAA methodology that can handle a large number of replica- tions as well as large sample sizes to calculate tight confidence intervals. Furthermore, an eﬃcient methodology is paramount to study the impact of the sample size on the computational performance of the methodol- ogy. This is a central theme of the present work.

* 1. *SAA applied to a risk-averse VPP optimization model*

The application of SAA methodologies to the scheduling problems of VPPs has not been reported in the literature. To fill the gap, we propose

an SAA methodology to solve these problems that is based on methods from [Linderoth et al. (2006)](#_bookmark71). Specifically, a distinct class of risk-averse problems from those considered in [Linderoth et al. (2006)](#_bookmark71) is addressed, and thus, insights that are not found in [Linderoth et al. (2006)](#_bookmark71) and [Verweij et al. (2003)](#_bookmark86) are analyzed and discussed.

In this work, new avenues to solve the VPP problem are explored, focusing on the application of an SAA methodology that leads to com- prehensive results for the VPP, including point estimates, confidence in- tervals, and the assessment of the solution obtained. The main contribu- tions of this work are twofold: 1) providing new and eﬃcient strategies to improve the performance of the SAA methodology; and 2) carrying out a comprehensive computational study applied to a VPP risk-averse two-stage stochastic programming problem. Specifically, an eﬃcient ini- tialization strategy to handle multiple optimization replications that re- duces the computing time is proposed. In addition, a new method to reduce the number of times the bound estimation stage is ran is dis- cussed. As a result, a detailed analysis of the effect of the sample size shows that, unexpectedly, increasing the sample size for the optimiza- tion leads to a reduction of the overall computing time in one of the procedures implemented. This reduction occurs because savings in the bound estimation stage compensate the additional time spent on the optimization with larger samples.

The paper is organized as follows. In [Section 2](#_bookmark8), the problem state- ment and a description of the VPP studied are presented. The stochastic models with a focus on the objective functions are described in [Section 3](#_bookmark7). The SAA approach and the implementation of two solution procedures are presented in [Section 4](#_bookmark13), while the sampling techniques are shown in [Section 5](#_bookmark27). The computational results are discussed in [Section 6](#_bookmark29) and conclusions are summarized in [Section 7](#_bookmark50). In a supplementary document, appendices with the deterministic model of the VPP, flow diagrams of the solution procedures used, and additional extensive computational results are made available.

# Problem statement

In this work, we revisit the problem considered in [Lima et al. (2018)](#_bookmark68) but using an SAA methodology. The problem addresses the optimal operation and electricity market involvement of a VPP that consists of a thermal unit, a wind power farm, and a pumped-storage hydroelectric plant. This configuration provides the flexibility of two dispatchable units, two renewable sources, and storage to manage the energy available and to respond to the market. The VPP participates in the electricity market as a single entity by selling and buying electricity in the market pool and through weekly contracts. The time horizon is one week, divided into 168 hourly periods.

The VPP generates electricity through the three units and may con- sume electricity to pump back water to the upper reservoir of the hydro plant. Uncertainty in the wind speed and the electricity prices is consid- ered, with a constant value for each hourly period.

The decision sequence of the VPP involves a two-stage framework. In the first-stage, the decisions are made before the beginning of the time horizon, where the VPP has to decide on the self-scheduling of the thermal unit and the electricity to buy or sell through forward contracts. The self-scheduling of the thermal unit involves only the commitment, defined as the periods where it is up or down. The VPP can choose to buy or sell electricity from two different weekly contracts, each one in- volving fixed prices and quantities. In the second-stage, the decisions for the full week are aggregated into one stage. These decisions are related to the dispatching of the thermal and hydro plants and to the additional electricity to buy or sell in the electricity pool (short-term electricity market). The dispatch involves the hourly determination of the power to generate by the thermal unit when it is up and the power generation or consumption by the pumped-storage hydro plant.

The objective of the VPP is to maximize the operational profit by de- termining an optimal solution concerning: a) the commitment, dispatch, and coordination between units in each period; b) the electricity bought

and sold in each period; and c) the electricity bought or sold through contracts. From the decision-maker perspective, the optimal first-stage variables are the most relevant. These decisions are implemented and fixed during the time horizon, whereas the optimal second-stage vari- ables for the 168 periods represent recourse actions based on the realiza- tions of the uncertain parameters. In practice, during the week, the VPP needs to solve additional decision problems to define the bidding strat- egy for the specific electricity market pool (day-ahead market, intra-day market, balancing market). These problems have a shorter horizon and use updated information for wind speed and electricity prices.

The thermal unit has lower and upper bounds on the power gen- eration, minimum up-time and down-time, start-up and shutdown and power-up and power-down ramp rate limits. Regarding costs, this unit has fixed and variable generation costs, hot and cold start costs. The state of the thermal unit before the beginning of the time horizon is known. The hydro plant has lower and upper limits on the volume of water of the upper reservoir, and an upper bound on the pumped flow of water.

# Risk-averse stochastic models

In this section, the main characteristics of the risk-averse stochas- tic programming formulations for the VPP problem are described. The deterministic version of the VPP optimization problem is described in Appendix A, where the objective function, constraints, variables, and input data are presented. Below, we focus on building a general risk- averse stochastic programming formulation that will be used in the SAA methodology outlined in [Section 4](#_bookmark13).

* 1. *VPP general stochastic programming formulation*

The stochastic programming formulation of the VPP problem is an extension of the deterministic VPP problem, which results from con- sidering the electricity prices and the wind power output as random variables. In the deterministic VPP problem, the aim is to maximize the profit defined as the difference between the revenues from selling elec- tricity and the costs of operation and buying electricity. In the stochastic version, the profit is a random function, and thus the aim is to maximize

a functional of the profit. We start by denoting the profit function *𝑓* as

*𝑓* (*𝑥*+*, 𝑥*−*, 𝑢, 𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*)*, 𝑠*(*𝜉*)*, 𝜉*) ∶= (*𝑐*+)*⊤𝑥*+ + (*𝑐*−)*⊤𝑥*− + *𝑐*0*⊤𝑢*

+ *𝑐̃*0*⊤*(*𝜉*)*𝑦*+(*𝜉*) − *𝑐̃*0*⊤*(*𝜉*)*𝑦*−(*𝜉*) + *𝑐⊤𝑠*(*𝜉*) (1)

where *𝜉* ∶ Ω → ℝ*𝑟* is a random vector defined on the probability space (Ω*, , 𝑃* )*,* with Ω being the set of all possible outcomes, a *𝜎*-algebra and *𝑃* a probability measure. The vectors *𝑐*+*, 𝑐*−*, 𝑐*0*,* and *𝑐* ∈ ℝ*𝑛*1 denote deterministic vectors with known parameters, whereas, *𝑐̃*0(*𝜉*) is a ran- dom vector. The vectors *𝑥*+*, 𝑥*−*, 𝑢, 𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*)*, 𝑠*(*𝜉*) ∈ ℝ*𝑛*1 represent the

T T

variables of the VPP model.

The vectors *𝑥*+*, 𝑥*−*,* and *𝑢* correspond to the first-stage decisions. *𝑥*+*,*

*𝑥*− capture the power to sell and buy through the contracts, respectively. The revenues related to the contracts are represented by (*𝑐*+)*⊤𝑥*+*,* while the costs by (*𝑐*−)*⊤𝑥*−. The vector *𝑢* encapsulates the binary variables that define the operation mode of the thermal unit. Therefore, the term *𝑐*0*⊤𝑢*

represents the startup costs, shutdown costs, and fixed costs of operation of the thermal unit.

The second-stage decisions are represented by the vectors *𝑦*+(*𝜉*)*,*

*𝑦*−(*𝜉*)*, 𝑠*(*𝜉*)*,* and *𝑟*(*𝜉*). Note that *𝑟*(*𝜉*) is not part of the objective function. The vectors *𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*) denote the energy to sell or buy to the market, and the vector *𝑐̃*0(*𝜉*) denotes the uncertain hourly electricity prices. The vector *𝑠*(*𝜉*) is the electricity generated by the thermal unit. Finally, the vector *𝑟*(*𝜉*) captures the energy generated or consumed by the pumped-

storage hydro unit and the volumes of water associated with this unit.

The stochastic model of the VPP aims at maximizing a functional of the profit subject to the constraints that define the region of operation of the thermal unit, pumped-storage hydro plant, and the two contracts

that are offered to the VPP:

and *𝐹𝑓*

is the cumulative distribution of *𝑓, 𝐹* (*𝑤*) = *𝑃* {*𝑓* ≤ *𝑤*}.

*𝑥*+ *,𝑥*− *,𝑣,𝑢,𝑦*+max− (*𝜉*)*,𝑠*(*𝜉*)*,𝑟*(*𝜉*)

*𝑓*

(*𝜉*)*,𝑦*

*𝜓* [*𝑓* (*𝑥*+*, 𝑥*−*, 𝑢, 𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*)*, 𝑠*(*𝜉*)*, 𝜉*)]

From an optimization perspective, there are two additional defini- tions that are relevant. The first is due to [Pflug (2000)](#_bookmark79), which defines

s.t. *𝐴*+*𝑥*+ + *𝐴*−*𝑥*− + *𝐵𝑣* ≤ *𝑏*

the CVaR through the optimization problem

*𝐸𝑢* ≤ *𝑔*

*𝐶 𝑢* + *𝐷𝑠*(*𝜉*) ≤ *𝑑,* a.s.

1−*𝛼 𝑓*

*𝜂*

*𝜂*

1 − *𝛼*𝔼

*𝜂*

*𝑓*

*,* (6)

CVaR

[ ] = max { − 1

[( −

)+]}

*𝑠*(*𝜉*) − *𝑦*+(*𝜉*) + *𝑦*−(*𝜉*) + *𝑟*(*𝜉*) − *𝑥*+ + *𝑥*−

+*ℎ*(*𝜉*) = 0*,* a.s.

*𝐽𝑟*(*𝜉*) ≤ *𝑗,* a.s.

*𝑥*+*, 𝑥*− ∈ ℝ*𝑛*1 *, 𝑣, 𝑢* ∈ 𝔹*𝑛*1 *, 𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*)*,*

(2)

where (⋅)+ = max{⋅*,* 0}. In [Pflug (2000)](#_bookmark79), it is shown that the optimal value of *𝜂* in [(6)](#_bookmark9) is the VaR1−*𝛼* [*𝑓* ]. The second definition shows that the func-

[tion that defines the CVaR of losses is a convex function (Rockafellar and Uryasev, 2000), which in our specific case leads to the equality](#_bookmark88)

+ max {CVaR [*𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*)]}

*𝑞𝑢𝑎𝑑𝑠*(*𝜉*) ∈ ℝ*𝑛*1 *, 𝑟*(*𝜉*) ∈ ℝ*𝑛*1 *,*

+

where 𝔹 ∶= {0*,* 1}*, 𝑣* are the binary variables associated with the choice

*𝑥,𝑧,𝑦*(*𝜉*)

= max

1−*𝛼*

{*𝜂* − 1 {[*𝜂* − *𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*)]+}}*.* (7)

𝔼

of blocks in the contracts, *ℎ*(*𝜉*) ∈ ℝ*𝑛*1 defines the uncertain wind power,

*𝐴*+*, 𝐴*− ∈ ℝ*𝑚*1 ×*𝑛*1 *, 𝐵* ∈ ℝ*𝑚*1 ×*𝑛*1 *, 𝐶* ∈ ℝ*𝑚*2 ×*𝑛*1 *, 𝐷* ∈ ℝ*𝑚*2 ×*𝑛*1 *, 𝐸* ∈ ℝ*𝑚*3 ×*𝑛*1 *, 𝐽* ∈

*𝑥,𝑧,𝑦*(*𝜉*)*,𝜂*

1 − *𝛼*

ℝ*𝑚*4 ×*𝑛*1 are matrices with known parameters, and *𝑏* ∈ ℝ*𝑚*1 *, 𝑑* ∈ ℝ*𝑚*2 *, 𝑔* ∈

ℝ*𝑚*3 *, 𝑗* ∈ ℝ*𝑚*4 are vectors with known parameters.

In Problem [(2)](#_bookmark10), the first constraint covers the contracts. The second

constraint represents the relationships among the binary variables as- sociated with the commitment of the thermal unit. The third constraint captures the region of operation of the thermal unit, including minimum

up and down times, and power generation limits. The fourth constraint represents the energy balance to the VPP. Finally, the last constraint

Based on the properties of the CVaR, we define an objective function representing the maximization of a combination of the expected profit and CVaR:

max *𝜓* [*𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*)] ∶= max 𝔼 (1 − *𝛽*)*𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*)

{ [

*𝑥,𝑧,𝑦*(*𝜉*) *𝑥,𝑧,𝑦*(*𝜉*)*,𝜂*

( 1 )]}

+*𝛽*

*𝜂* − 1 −

[*𝜂* − *𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*)]+

*,* (8)

*𝛼*

captures the region of operation of the pumped-storage hydro plant. See Appendix A for details.

For the sake of clarity on the exposition of the SAA methodology, a compact formulation of the model with a simplified notation is in-

troduced. We set *𝑥* ∶= (*𝑥*+*, 𝑥*−)*, 𝑦* ∶= (*𝑦*+(*𝜉*)*, 𝑦*−(*𝜉*)*, 𝑠*(*𝜉*)*, 𝑟*(*𝜉*))*, 𝑧* ∶= (*𝑢, 𝑣*)*,*

*𝑐* ∶= (*𝑐*+*, 𝑐*−)*, 𝑐* ∶= *𝑐*0*, 𝑐̃* ∶= *𝑐̃*0 with proper dimensions:

max *𝜓* [*𝑓* (*𝑥, 𝑧, 𝑦*(*𝜉*)*, 𝜉*) ∶= *𝑐⊤𝑥* + *𝑐⊤𝑧* + *𝑐̃⊤*(*𝜉*)*𝑦*(*𝜉*)]

≤

where *𝛽* ∈ [0*,* 1] is a parameter that defines the weights of the expec- tation and CVaR of *𝑓* . We are particularly interested in the solution of Problem [(3)](#_bookmark12) with [(8)](#_bookmark11) defined with *𝛽* = (0*,* 1) and *𝛽* = 1 (maximization of

the CVaR of the profit) using an SAA approach.

Detailed analyses on the properties of the CVaR and VaR metrics and their integration into stochastic programming can be found in [Pflug (2000)](#_bookmark79), [Rockafellar and Uryasev (2000)](#_bookmark88), [Rockafellar (2007)](#_bookmark87), and

*𝑥,𝑧,𝑦*(*𝜉*)

s.t. *𝐴𝑥* + *𝐵𝑧*

*𝑏*

≤

*𝐶𝑧* + *𝐷𝑦*(*𝜉*) *𝑑,* a.s.

*𝑦*(*𝜉*) + *𝑥* = *ℎ*(*𝜉*)*,* a.s.

*𝑥* ∈ ℝ*𝑛*1 *, 𝑧* ∈ 𝔹*𝑛*1 *, 𝑦*(*𝜉*) ∈ ℝ*𝑛*1 *,*

(3)

[Sarykalin et al. (2008)](#_bookmark90). Risk metrics for multi-stage stochastic program-

[ming, such as the expected CVaR can be found in Pflug and Ruszczyn- ski (2005),](#_bookmark81) [Shapiro](#_bookmark80) [et al. (2009), and](#_bookmark81) [Mello](#_bookmark75) [and Pagnoncelli (2016).](#_bookmark81) Also, the work in [Alonso-Ayuso et al. (2018)](#_bookmark51) provides extensive com- parisons on multi-period risk metrics.

+ +

where *𝑥, 𝑧* denote first-stage variables and *𝑦*(*𝜉*) second-stage variables, *ℎ*(*𝜉*) ∈ ℝ*𝑛*1 and *𝑐̃*(*𝜉*) ∈ ℝ*𝑛*1 are random vectors, *𝑐, 𝑐,* and *𝑐̃* are determinis- tic vectors with known parameters, *𝐴* ∈ ℝ*𝑚*1 ×*𝑛*1 *, 𝐵* ∈ ℝ*𝑚*1 ×*𝑛*1 *, 𝐶* ∈ ℝ*𝑚*2 ×*𝑛*1 *,* and *𝐷* ∈ ℝ*𝑚*2 ×*𝑛*1 are matrices with known parameters, and *𝑏* ∈ ℝ*𝑚*1 and

# Sample average approximation

For the generic problem formulation

*𝑑* ∈ ℝ*𝑚*2 are vectors with known parameters.

*𝑤*∗ = max

*𝑥,𝑦*(*𝜉*)*,𝑧*∈



{𝔼*𝜉* [*𝜙*(*𝑥, 𝑦*(*𝜉*)*, 𝑧, 𝜉*)]}*,* (9)

*3.2. Risk-averse VPP stochastic models*

where  ∶= {*𝑥, 𝑦*(*𝜉*)*, 𝑧 𝐴𝑥* + *𝐵𝑧* ≤ *𝑏*; *𝐶𝑧* + *𝐷𝑦*(*𝜉*) ≤ *𝑑, 𝑎.𝑠.*; *𝐹 𝑦*(*𝜉*) + *𝐺𝑥* =

*ℎ*(*𝜉*)*, 𝑎.𝑠.*; *𝑥* ∈ ℝ*𝑛*1 *, 𝑧* ∈ 𝔹*𝑛*1 *, 𝑦*(*𝜉*) ∈ ℝ*𝑛*1 }*,* we define the sample average ap-

|

{ ∑

+

+

the CVaR of the profit in the objective function. The CVaR is a coher-

Based on [(3)](#_bookmark12), we consider a risk-averse formulation that involves

[proximation problem as (Kleywegt et al., 2001; Shapiro and Homem-De- Mello, 2000):](#_bookmark66)

[ent risk measure with relevant properties in terms of convexity (Artzner](#_bookmark54)

[et al., 1999; Pflug, 2000; Rockafellar and Uryasev, 2000). Therefore, it](#_bookmark54) has clear advantages over non-convex risk measures. Note that CVaR

∗ = max

*𝑁*

*𝑤*

*𝑥,𝑦,𝑧*∈

*𝑁*

1 *𝑁*

*𝑁*

*𝜙*(*𝑥, 𝑦, 𝑧, 𝜉𝑛*)}

*,* (10)

is not the only possibility to measure the risk; other measures have been suggested including the variance, value at risk, drawdown, or

*𝑛*=1

for a sample with independent and identically distributed elements

 ≤

*𝜉𝑛* from the distribution of *𝜉,* and where ∶= {*𝑥, 𝑦*(*𝜉𝑛*)*, 𝑧 𝐴𝑥* + *𝐵𝑧*

*𝑏*; *𝐶𝑧* + *𝐷𝑦*(*𝜉* ) ≤ *𝑑,* ∀*𝜉*

buffered probability of exceedance; see [Pflug and Römisch (2007)](#_bookmark82) and

[Shang et al. (2018)](#_bookmark91). Also note that many works in the literature apply

*𝑛 𝑛 𝑛*

*𝑁*

*𝑛 𝑛*

∈ Ξ; *𝐹 𝑦*(*𝜉* ) + *𝐺𝑥* = *ℎ*(*𝜉* )*,* ∀*𝜉*

∈ Ξ; *𝑥* ∈ ℝ+ *, 𝑧* ∈

| *𝑛*1

the CVaR to a loss function, whereas in this work the CVaR is applied +

𝔹*𝑛*2 *, 𝑦*(*𝜉𝑛*) ∈ ℝ*𝑛*3 *, 𝑛* = 1*,* … *, 𝑁* }*,* and *𝜙* stands for *𝜙* .

*𝛽*

∗ ∗ ∗ ∗

to a profit one, which leads to different definitions below as compared

The optimal value and solution of Problem [(9)](#_bookmark14) are *𝑤* and *𝑥 , 𝑦 , 𝑧 ,*

respectively, which are approximated by *𝑤*∗ and *𝑥*∗ *, 𝑦*∗ *, 𝑧*∗

obtained

with the works cited. The terms “average value at risk” and “expected

from [(10)](#_bookmark15). Note that *𝑤*∗

*𝑁 𝑁 𝑁 𝑁*

shortfall” are also used to refer to the CVaR ([Rockafellar, 2007](#_bookmark87)).

The CVaR for the (1 − *𝛼*) quantile of random variable *𝑓* is defined [as the conditional expectation of *𝑓* for *𝑓* VaR1−α[f] (Rockafellar and](#_bookmark88)

≤

[Uryasev, 2000):](#_bookmark88)

CVaR1−*𝛼* [*𝑓* ] = 𝔼 *𝑓 𝑓* ≤ VaR1−*𝛼* [*𝑓* ] *,* (4)

[ | ]

where VaR1−*𝛼* [*𝑓* ] is the value at risk defined as

VaR1−*𝛼* [*𝑓* ] = max{*𝑤 𝐹* (*𝑤*) ≤ 1 − *𝛼*}*,* (5)

| *𝑓*

*𝑁* is a random variable that depends on the sample

of *𝜉𝑛*.

The objective function [(8)](#_bookmark11) involve the CVaR measure, which may

not resemble the generic formulation in [(9)](#_bookmark14). However, note that [(8)](#_bookmark11) is reformulated as a function of the expectation operator.

Using the SAA methodology, two types of inference statistics for the solutions are calculated: 1) point estimates and confidence intervals for

the upper and lower bounds on the optimal objective function value *𝑤*∗;

2) a point estimates of the upper bound on the gap between *𝑤*∗ and *𝑤*∗ .

*𝑁*

**Table 1**

Case 1. Optimization results for the formulation with *𝛽* = 0. The results are ordered by the optimal objective function value (*𝑤̂ 𝑁,𝑚* ). *𝑀* = 30*, 𝑁* = 10. **Bold** – set of distinct first-stage solutions.

First-stage variables (aggregated)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝑚* | *𝑤̂ 𝑁,𝑚* (€) | ITER | GAP (%) | *T* (s) | UT (%) | SUP | SD | SELLC (MW) | BUYC (MW) |  |
| 23 | 2,437,900 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **155** | **0** |  |
| 22 | 2,754,781 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **50** | **55** |  |
| 27 | 2,791,618 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 50 | 55 |  |
| 18 | 2,807,663 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 50 | 55 |  |
| 17 | 2,808,236 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 50 | 55 |  |
| 2 | 2,828,660 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 50 | 55 |  |
| 19 | 2,881,600 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **50** | **110** |  |
| 20 | 2,950,385 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 50 | 110 |  |
| 3 | 2,987,171 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **0** | **160** |  |
| 9 | 3,051,161 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 21 | 3,075,317 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 28 | 3,086,196 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 26 | 3,098,970 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 13 | 3,112,520 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 7 | 3,141,942 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 4 | 3,182,338 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 14 | 3,225,259 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 12 | 3,241,059 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 15 | 3,245,519 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 25 | 3,251,400 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 10 | 3,273,009 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 8 | 3,311,875 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 16 | 3,326,421 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 29 | 3,331,775 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 160 |  |
| 11 | 3,371,868 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **0** | **215** |  |
| 6 | 3,525,814 | 2 | 0.0000 | 1 | **100** | **0** | **0** | **0** | **265** |  |
| 30 | 3,543,351 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 265 |  |
| 5 | 3,682,847 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 265 |  |
| 1 | 3,738,467 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 265 |  |
| 24 | 3,743,889 | 2 | 0.0000 | 1 | 100 | 0 | 0 | 0 | 265 |  |

*𝑚* – Optimization replication index, *𝑤̂ 𝑁,𝑚* - optimal objective function value for optimization repli- cation *𝑚,* ITER – number of iterations of the L-Shaped method, GAP – gap between the upper and

lower bound within the L-Shaped method, *T* – elapsed wall-clock time, UT – percentage of up-time of the thermal unit, SUP/SD – number of startups/shutdowns of the thermal unit, SELLC/BUYC – power sold/bought through contracts.

Besides, for the specific application, the solutions are complemented with the corresponding point estimates and confidence intervals for the expected profit and the CVaR of the profit.

jective function value *𝑤*∗*,* the following notation is used: a) for the upper To evaluate the inference statistics on the bounds of the optimal ob- bound, *𝑁* denotes the sample size, *𝑛* is the index of an element of the sample, *𝑀* is the number of replications, and *𝑚* is the replication index;

b) for the lower bound, *𝑁* ′ denotes the sample size, *𝑛*′ is the index of an element of the sample, *𝑇* and *𝑇* ′ are the number of replications, and *𝑡* and *𝑡*′ are replication indices. For the upper bound on the gap between

*𝑤*∗ and *𝑤*∗ *, 𝑁, 𝑛, 𝑀,* and *𝑚* are used with the same purpose as in a).

rences. These results do not support a clear-cut decision on the best solution to choose. Obviously, the variance of these results is due to the size of the sample used for the optimization. However, with a small sample and the methodologies presented in this work, it is possible to provide a set of inference statistics and to assess the solutions. This in- formation can help the selection of first-stage solutions.

The impact of the sample size *𝑁* and the number of replications *𝑀*

on the probability of obtaining an optimal solution to the true problem

work, it is claimed that the relation between *𝑁* and this probability from the SAA problem is discussed in [Kleywegt et al. (2001)](#_bookmark66). In that

is problem specific, and that beyond a given *𝑀,* it would be better to

*𝑁* ′

Note that the symbols *𝑡, 𝑡 ,* and *𝑇* are not related to the symbols *𝑡* and *𝑇*

used in Appendix A to refer to time periods and time horizon.

* 1. *Motivating example*

In this section, we illustrate the variance of the optimal objective function values and first-stage solutions obtained using different sam- ples having the same size. The samples and results used in this exam-

increase *𝑁* .

* 1. *Upper bound on the optimal objective function value 𝑤*∗

The approximation of *𝑤*∗ is done by calculating point estimates of lower and upper bounds on *𝑤*∗. An upper bound on *𝑤*∗ is defined by the

≥

well-known relation 𝔼[*𝑤*∗ ] *𝑤*∗ ([Mak et al., 1999](#_bookmark69)). The value of the

ple are part of the computational experiments described in [Section 6](#_bookmark29).

𝔼[*𝑤*∗ ]

*𝑁*

is estimated by the statistical estimator

*𝑁*

*𝑈𝐵𝑁,𝑀*

that is determined

[Table 1](#_bookmark16) outlines the results of 30 optimization replications, each with a

sample of size *𝑁* = 10*,* for the maximization of the expected profit.

by solving *𝑀* optimization problems [(10)](#_bookmark15)

The variability of these results can be highlighted with the results

of two samples. The first sample (*𝑚* = 1) has optimal objective function

*𝑤̂ 𝑁,𝑚* = max

{ 1 ∑*𝑁*

*𝜙*(*𝑥, 𝑦, 𝑧, 𝜉𝑛,𝑚*)}

*,* ∀*𝑚* ∈ *𝑀,* (11)

value of 3,738,467 and first-stage solution {100*,* 0*,* 0*,* 0*,* 265}*,* whereas

*𝑥,𝑦,𝑧*∈

*𝑁,𝑚*

*𝑁 𝑛*=1

for *𝑚* = 2*,* the optimal objective function value is 2,828,660 with first- stage solution {100*,* 0*,* 0*,* 50*,* 55}. This table also shows that the first-

stage solution {100*,* 0*,* 0*,* 0*,* 265} corresponds to the higher profits, and

the solution with {100*,* 0*,* 0*,* 0*,* 160} has the higher number of occur-

and by using *𝑤̂ 𝑁,𝑚* to estimate the *𝑈𝐵𝑁,𝑀* through

1 ∑*𝑀*

*𝑈𝐵𝑁,𝑀* = *𝑀*

*𝑤̂ 𝑁,𝑚.* (12)

*𝑚*=1

The sample variance estimator is

1 2 1 ∑*𝑀* ( )2

(*𝑠*

)

=

*𝑤̂*

– *𝑈𝐵𝑁,𝑀*

*.* (13)

*𝑚*=1

*𝑁,𝑀*

*𝑀* (*𝑀* − 1)

*𝑁,𝑚*

≥ *𝑔𝑎𝑝*(*𝑥̂ , 𝑧̂* )*.* (22)

 is calculated using 𝔼[*𝑤*∗ ] as a valid statistical upper bound on

*𝑁 𝑁*

[

*𝑤*∗*,* 𝔼[*𝑤*∗ ] ≥ *𝑤*∗*,* and 𝔼[*𝑤̂*

*𝑁*

] as an estimator of 𝔼 *𝜙*(*𝑥̂*

*, 𝜉*)]. In

*𝑁*

*𝑁*

*𝜉*

*𝑁*

*, 𝑦, 𝑧̂*

*𝑁*

Thus, *𝑈𝐵* is a statistical estimator of 𝔼[*𝑤̂*

] with the 100(1 − *𝛼*′)%

[Mak et al. (1999)](#_bookmark69), it is proposed to use the upper limit of the confidence

*𝑁,𝑀 𝑁*

confidence interval

interval of the estimator of  as an upper bound on the *𝑔𝑎𝑝*(*𝑥̂ , 𝑧* ).

*𝑁*

*𝑁*

*𝐿𝑈 𝐵𝑁,𝑀 , 𝑈 𝑈 𝐵𝑁,𝑀* ∶= *𝑈 𝐵𝑁,𝑀* − *𝑡𝑀* −1*,𝛼*′ *𝑠𝑁,𝑀 , 𝑈 𝐵𝑁,𝑀* + *𝑡𝑀* −1*,𝛼 𝑠𝑁,𝑀 ,*

[ ] [ ]

(14)

where *𝑡𝑀* −1*,𝛼*′ is the critical value from the *t*-distribution.

* 1. *Lower bound on the optimal objective function value 𝑤*∗

We define *𝐿𝐵𝑁* ′ *,𝑚,𝑇* as an estimator of a lower bound on *𝑤*∗*,* which is associated with a first-stage solution from the optimization replication

*𝑚* denoted by (*𝑥̂𝑁,𝑚 , 𝑧̂𝑁,𝑚* ). The calculation of *𝐿𝐵𝑁* ′*,𝑚,𝑇* is done using *𝑇*

independent samples of size *𝑁* ′*,* as follows

This upper limit calculation is presented in [Procedure 2](#_bookmark24). Note that this procedure uses the same sample to calculate the two terms on the left hand side of the inequality in [(22)](#_bookmark18), and the upper bound on the gap is always positive. Additional details regarding variants of evaluation pro- cedures and convergence analysis of the gap estimator can be found in [Bayraksan and Morton (2006)](#_bookmark55) and [Mak et al. (1999)](#_bookmark69).

*4.5. Implementation*

The practical computation of the statistical point estimates and confidence intervals described in [Sections 4.2](#_bookmark17)–[4.4](#_bookmark23) is presented in [Procedures 1](#_bookmark25) and [2](#_bookmark24), and the corresponding diagrams in Appendix B. These two procedures involve two stages denoted as optimization stage

*𝑤̂*

*𝑁* ′ *,𝑚,𝑡*

= max

*𝑁 ,𝑚,𝑡*

*𝑦*∈ ′

*𝑇*

∑

1

*𝑁* ′

{

*𝑁* ′

*𝑛*′=1

∑

*𝜙*(*𝑥̂*

*𝑁,𝑚*

*, 𝑦, 𝑧̂*

*𝑁,𝑚*

*, 𝜉𝑛*′ *,𝑡* )}

*,* ∀*𝑚* ∈ *𝑀,* ∀*𝑡* ∈ *𝑇 ,*

(15)

and bound estimation stage. In line 3 of [Procedures 1](#_bookmark25) and [2](#_bookmark24), a risk- neutral or a risk-averse stochastic programming problem is solved. Each of these problems can be solved directly using an MILP solver or by decomposition. We adopt the latter and use the L-Shaped method pro- posed in the seminal work of [Van Slyke and Wets (1969)](#_bookmark85). This choice

*𝐿𝐵*

*𝑁* ′ *,𝑚,𝑇*

= 1 *𝑤̂*

*𝑇 𝑡*=1

*𝑁* ′ *,𝑚,𝑡*

*,* ∀*𝑚* ∈ *𝑀.* (16)

is supported by a previous successful application of this method in solv- ing large scale problems ([Lima et al., 2018](#_bookmark68)). The L-Shaped method is

Note that the estimator *𝐿𝐵*

*𝑁* ′ *,𝑚,𝑇*

is associated with the solution of the

an extension of Benders decomposition ([Benders, 1962](#_bookmark56)) for two-stage

stochastic programming problems, which is straightforward to describe

optimization replication with index *𝑚*. To generate a single estimator of

the lower bound (from the *𝑀* bounds *𝐿𝐵𝑁* ′ *,𝑚,𝑇* ), the first-stage solution

using problem [(23)](#_bookmark20):

that corresponds to the maximum value of *𝐿𝐵𝑁* ′ *,𝑚,𝑇* is selected and for

this solution a new estimator is calculated using *𝑇* ′ independent samples of size *𝑁* ′:

*𝑐*

*𝑥,𝑧,𝜂*

*𝑥*

max

*⊤* + *⊤*

+ (1 −

) ∑*𝑁* [ (

*𝑛*)]

{ 1 ∑*𝑁* ′ }

+*𝛽*

*𝜂* − 1 −

*𝑝* (*𝜂* − *𝑄*(*𝑥, 𝑧, 𝜉𝑛*))+

*𝛼 𝑛*=1

∑ [ *𝑛*

1

*𝛽*

*𝑁*

*𝑛*=1

*𝑝𝑛𝑄 𝑥, 𝑧, 𝜉*

(23)

*𝑤̂ 𝑁* ′ *,𝑡*′ = max

*𝑦*∈ ′

*𝑁 ,𝑡*

*𝑁* ′

*𝑛*′ =1

*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉𝑛 ,𝑡*)

*,* ∀*𝑡*′ ∈ *𝑇* ′*,* (17)

*𝑠.𝑡. 𝐴𝑥* + *𝐵𝑧* ≤ *𝑏*

*𝑥* ∈ ℝ*𝑛*1 *, 𝑧* ∈ 𝔹*𝑛*2 *, 𝜂* ∈ ℝ*,*

*𝑐 𝑧*

{

]}

′

where (*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁* ) is the first-stage solution selected. For this first-stage

solution, the estimator of the lower bound is given by

1 ∑

*𝑇* ′

+

where *𝑄*(*𝑥, 𝑧, 𝜉𝑛*) is defined as

*𝑄*(*𝑥, 𝑧, 𝜉𝑛*) ∶= max *𝑐̃⊤𝑦𝑛* ⎫

*𝑠.𝑡. 𝐷𝑦𝑛*

≤ *𝑑* − *𝐶𝑧*⎪

∀*𝑛* = 1*,* … *, 𝑁,* (24)

*𝐿𝐵* ′ ′ =

*𝑁 ,𝑇*

*𝑤̂*

′ ′ *,* (18)

*𝑦𝑛 𝑛* ⎪

and the sample variance of *𝐿𝐵* ′

*𝑇* ′ *𝑡*′ =1

*𝑁 ,𝑡*

′ through

*𝑦𝑛* = *ℎ𝑛* − *𝑥* ⎬⎪

2 2 1

(*𝑠𝑁* ′*,𝑇* ′ )

=

*𝑇* (*𝑇* − 1)

where the first problem is an MILP problem followed by a collection of LP subproblems that constitutes a sample element-wise decomposi-

∑*𝑀* (

*𝑚*=1

*𝑁 ,𝑇*

)2

*𝑤̂ 𝑁* ′ *,𝑡*′ − *𝐿𝐵𝑁* ′ *,𝑇* ′

*,* (19)

*𝑦𝑛*

∈ ℝ*𝑛*3

with the confidence interval defined as

+

⎭

[*𝐿𝐿𝐵𝑁* ′ *,𝑇* ′ *, 𝑈𝐿𝐵𝑁* ′ *,𝑇* ′ ] ∶= [*𝐿𝐵𝑁* ′ *,𝑇* ′ − *𝑡𝑇* −1*,𝛼*′ *𝑠𝑁* ′*,𝑇* ′ *, 𝐿𝐵𝑁* ′*,𝑇* ′ + *𝑡𝑇* −1*,𝛼*′ *𝑠𝑁* ′ *,𝑇* ′ ]*.*

(20)

This approach to calculate the point estimate of the lower bound is based on [Linderoth et al. (2006)](#_bookmark71).

* 1. *Upper bound on the gap between 𝑤*∗ *and* 𝔼*𝜉* [*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*)]

of a solution (*𝑥̂𝑁 , 𝑧̂𝑁* )*,* which defines an estimator of an upper bound on In [Mak et al. (1999)](#_bookmark69), it is derived a method for assessing the quality

the following gap:

*𝑔𝑎𝑝*(*𝑥̂𝑁 , 𝑧̂𝑁* ) ∶= *𝑤*∗ − 𝔼*𝜉 𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*) *.* (21) This gap represents the difference between the true optimal solution *𝑤*∗ and the true optimal solution 𝔼*𝜉 𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*) for (*𝑥̂𝑁 , 𝑧̂𝑁* ). The upper

[ []](#_bookmark69)

[ ]



bound, *,* on the *𝑔𝑎𝑝*(*𝑥̂ , 𝑧̂* ) is defined as ([Mak et al., 1999](#_bookmark69)):

*𝑁*

*𝑁*

tion, meaning one subproblem for each element *𝑛* = 1*,* … *, 𝑁* of the sam- ple. This formulation moves the variable *𝑦𝑛* to the subproblems. The

L-Shaped method is an iterative method, where in each iteration the following steps are performed:

1. A relaxation of problem [(23)](#_bookmark20), called master problem, is solved. This

volving *𝑄*(*𝑥, 𝑧, 𝜉𝑛*)*,* and thus, its objective function value provides an master problem is constructed by outer-approximating the terms in-

solution of this master problem is the first-stage variables: (*𝑥, 𝑧*). upper bound on the objective function value of problem [(23)](#_bookmark20). The

1. The subproblems [(24)](#_bookmark21) are solved independently for fixed values of the first-stage variables obtained from the master problem.
2. The VaR and CVaR of the second stage profit are evaluated based on the distribution of the second stage profits (the objective function values of the subproblems), and they are used in the next step.
3. The upper and lower bounds are calculated and the method termi- nates if the gap between these bounds reaches a specified threshold.
4. The dual variables solution of the subproblems are used to build the

outer-approximations to be added to the master problem in the next

 ∶= 𝔼[

max

*𝑥,𝑦,𝑧*∈ *𝑁*



1 *𝑁*

*𝑁 𝑛*=1

{ ∑

*𝜙*(*𝑥, 𝑦, 𝑧, 𝜉𝑛*)} −

1 *𝑁*

*𝑁 𝑛*=1

∑

*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉𝑛*)]

iteration. In the next iteration, these outer-approximations “cut” the solution of the master problem obtained in step 1.

**Procedure 1** Proposed implementation of an SAA method. Input: *𝑀*

samples of size *𝑁* : *𝜉𝑛,𝑚, 𝑛* = 1*,* … *, 𝑁, 𝑚* = 1*,* … *, 𝑀* ; *𝑇* samples of size

*𝑁* ′: *𝜉𝑛*′ *,𝑡 , 𝑛*′ = 1*,* … *, 𝑁* ′*, 𝑡* = 1*,* … *, 𝑇* ; *𝑇* ′ samples of size *𝑁* ′: *𝜉𝑛*′ *,𝑡*′ *, 𝑛*′ =

[Mak et al., 1999). Input: (*𝑥̂*](#_bookmark55)

[gap between *𝑤*∗ and 𝔼*𝜉 𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*) (Bayraksan and Morton, 2006;](#_bookmark55) **Procedure 2** Proposed implementation of the upper bound on the

1*,* … *, 𝑁* ′*, 𝑡*′ = 1*,* … *, 𝑇* ′; and confidence level. Output: first-stage solu-

tion, point estimates of bounds on the optimal objective function value

*𝑤*∗*,* and confidence intervals.

1: **for** *𝑚* = 1 **to** *𝑀* **do**

2: Initialize the L-Shaped method with (*𝑥̂𝑏𝑒𝑠𝑡 , 𝑧̂𝑏𝑒𝑠𝑡* ), if this solution

[[ ]](#_bookmark55)

[) from](#_bookmark55) [Procedure](#_bookmark25) [1; *𝑀* samples of size](#_bookmark55)

[*𝑁*](#_bookmark55)

[*, 𝑧̂*](#_bookmark55)

[*𝑁*](#_bookmark55)

*𝑁* : *𝜉𝑛,𝑚, 𝑛* = 1*,* … *, 𝑁, 𝑚* = 1*,* … *, 𝑀* ; and confidence level. Output: upper bound on the *𝑔𝑎𝑝*(*𝑥̂𝑁 , 𝑧̂𝑁* ) = *𝑤*∗ − 𝔼*𝜉* [*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*]

1: **for** *𝑚* = 1 **to** *𝑀* **do**

2: Initialize the L-Shaped method with (*𝑥̂𝑁 , 𝑧̂𝑁* ) **or** if available, with

(*𝑥̂𝑏𝑒𝑠𝑡 , 𝑧̂𝑏𝑒𝑠𝑡* )

is available

*𝑁,𝑚*

*𝑁,𝑚*

*𝑁,𝑚*

*𝑁,𝑚*

3: Solve using the L-Shaped method each MILP problem with the

3: Solve using the L-Shaped method each MILP problem with the

sample with size *𝑁* of i.i.d. *𝜉𝑛,𝑚*:

*𝑁,𝑚*

sample with size *𝑁* of i.i.d. *𝜉𝑛,𝑚*:

{

}

*𝑁*

*𝑤̂*

= max

{ 1 ∑*𝑁*

*𝜙*(*𝑥, 𝑦, 𝑧, 𝜉𝑛,𝑚*)}

(28)

*𝑤̂*

∑

*𝑁,𝑚*

= max

*𝑥,𝑦,𝑧*∈



1 *𝜙*(*𝑥, 𝑦, 𝑧, 𝜉𝑛,𝑚*)

*𝑁*

(25)

*𝑁,𝑚*

*𝑥,𝑦,𝑧*∈

*𝑁 𝑛*=1

4: Let (*𝑥̂*

*, 𝑦̂𝑛*

*𝑁,𝑚*

*, 𝑧̂*

*𝑛*=1

) be the optimal solution from (25)

4: Let (*𝑥̂*

*𝑁,𝑚*

*𝑛*

*𝑁,𝑚*

*, 𝑦̂*

*, 𝑧̂*

*𝑁,𝑚*

) be the optimal solution of (28)

*𝑁,𝑚*

*𝑁,𝑚*

*𝑁,𝑚*

5: **if** *𝑤̂ 𝑁,𝑚 > 𝑤̂ 𝑁,𝑚*−1 **then**

5: **if** *𝑤̂ 𝑁,𝑚 > 𝑤̂ 𝑁,𝑚*−1 **then**

6: (*𝑥̂𝑏𝑒𝑠𝑡 , 𝑧̂𝑏𝑒𝑠𝑡* ) ← (*𝑥̂𝑁,𝑚 , 𝑧̂𝑁,𝑚* )

6: (*𝑥̂𝑏𝑒𝑠𝑡 , 𝑧̂𝑏𝑒𝑠𝑡* ) ← (*𝑥̂*

*, 𝑧̂* )

*𝑁,𝑚 𝑁,𝑚*

*𝑁,𝑚*

7: **end if**

*𝑁,𝑚*

*𝑁,𝑚*

*𝑁,𝑚*

7: **end if**

8: Solve each LP subproblem with the sample with size *𝑁* of i.i.d.

8: **if** (*𝑥̂𝑁,𝑚 , 𝑧̂𝑁,𝑚* ) is a new solution **then**

9: **for** *𝑡* = 1 **to** *𝑇* **do** {Lower bound estimation}

10: Solve each LP subproblem with the sample with size *𝑁* ′ of

i.i.d. *𝜉𝑛*′ *,𝑡* :

{

*𝜉𝑛,𝑚* :

max

*𝑦*∈

{ ∑

*𝑁,𝑚*

1 *𝑁*

*𝑁 𝑛*=1

*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉𝑛,𝑚*)}

(29)

*𝑤̂*

*𝑁,𝑚*

*𝑁*

*𝑁* ′ *,𝑚,𝑡*

= max

*𝑁 ,𝑚,𝑡*

*𝑦*∈ ′

1

*𝑁* ′

*𝑁* ′

*𝑛*′ =1

∑

*𝜙*(*𝑥̂*

*𝑁,𝑚*

*, 𝑦, 𝑧̂*

*𝑁,𝑚*

*, 𝜉𝑛*′ *,𝑡* )}

(26)

9: Evaluate

1 *𝑁*

∑ [



*𝑁,𝑚* = *𝑁*

*𝜙*(*𝑥̂𝑁,𝑚 , 𝑦̂𝑛*

*, 𝑧̂𝑁,𝑚 , 𝜉𝑛,𝑚*) − *𝜙*(*𝑥̂𝑁 , 𝑦̂𝑛 , 𝑧̂𝑁 , 𝜉𝑛,𝑚*)]

(30)

11: **end for**

10: **end for**

*𝑛*=1

12: Evaluate *𝐿𝐵𝑁* ′ *,𝑚,𝑇* using (16)

13: **if** *𝑚* = 1 **then**

14: (*𝑥̂𝑁 , 𝑧̂𝑁* ) ← (*𝑥̂𝑁,𝑚 , 𝑧̂𝑁,𝑚* )

11: Evaluate the statistical point estimate for the expected gap and sam- ple variance

*𝑀*

∑



15: **else if** *𝑚 >* 1 **and** *𝐿𝐵*

*𝑁* ′ *,𝑚,𝑇*

*> 𝐿𝐵*

*𝑁* ′ *,𝑚*−1*,𝑇*

# then

*𝑁,𝑀*

= 1 

*𝑀*

*𝑁,𝑚*

(31)

16: (*𝑥̂𝑁 , 𝑧̂𝑁* ) ← (*𝑥̂𝑁,𝑚 , 𝑧̂𝑁,𝑚* )

17: **end if**

18: **end if**

(*𝑠𝑀* )2 =

*𝑚*=1

1 *𝑀*

∑ (

*𝑀* (*𝑀* − 1) *𝑚*=1

*𝑁,𝑚*

 2

*𝑁,𝑀*

– 

)

(32)

19: **end for**



12: **return** upper bound *𝑈𝐺𝑁,𝑀*

*𝑁,𝑀*

+ *𝑡*

*𝑀* −1*,𝛼 𝑠𝑀*

20: Evaluate *𝑈𝐵𝑁,𝑀* using (12) and the confidence interval using (14)

= 

21: **for** *𝑡*′ = 1 **to** *𝑇* ′ **do** {Lower bound estimation}

22: Solve each LP subproblem with the sample with size *𝑁* ′ of i.i.d.

*𝜉𝑛*′ *,𝑡*′ :

*𝑤̂ 𝑁* ′*,𝑡*′ = max

*𝑦*∈ ′ ′

*𝑁 ,𝑡*

1

*𝑁* ′

{

*𝑁* ′

*𝑛*′ =1

∑

*𝑁 𝑁*

*𝜙*(*𝑥̂ , 𝑦, 𝑧̂ , 𝜉𝑛*′ *,𝑡*′ )}

(27)

pabilities available. In this work, the parallelization is exploited in the solution of the independent LP subproblems in two places: 1) within the L-Shaped method; and 2) at the lower bound estimation, where for each

first-stage solution, *𝑁* ′ LP subproblems are solved.

23: **end for**

24: Evaluate *𝐿𝐵𝑁* ′ *,𝑇* ′ using (18) and the confidence interval using (20)

In Procedures 1 and 2, we implement an initialization step that uses

25: **return** (*𝑥̂𝑁*

*, 𝑧̂𝑁*

), *𝑈𝐵𝑁,𝑀*

, *𝐿𝐵*

*𝑁* ′ *,𝑇*

′ , and confidence intervals

the best first-stage solution obtained in a replication *𝑚*′ *< 𝑚* to initialize the L-Shaped method in replication *𝑚*. This initialization is indicated in

line 2 of [Procedure 1](#_bookmark25), where (*𝑥̂𝑏𝑒𝑠𝑡 , 𝑧̂𝑏𝑒𝑠𝑡* ) is the best first-stage solution.

*𝑁,𝑚 𝑁,𝑚*

The outer-approximations are known as optimality cuts. These cuts require that all subproblems in step 2 are feasible, otherwise, a feasi- bility subproblem needs to be solved for each infeasible subproblem. The optimal values of the dual variables of these feasibility subprob- lems are then used to build cuts, known as feasibility cuts, to be added to the master problem in the next iteration. In the problems studied in this work, the subproblems are always feasible for all values of the first- stage variables. This is due to the possibility to buy electricity from the pool, which is implemented in the subproblems to satisfy contracts de- cided in the master problem. An algorithmic description, specific prob- lem formulations, expressions for bounds, and details on handling the

variable *𝜂* and the CVaR calculation can be found in [Lima et al. (2018)](#_bookmark68),

and the outer-approximations derivations and convergence analysis in

[Van](#_bookmark65) [Slyke and Wets (1969)](#_bookmark85)[,](#_bookmark65) [Birge and Louveaux (2011)](#_bookmark57)[, and Kall and Mayer (2011).](#_bookmark65)

Clearly, the solution of multiple replications of risk-averse prob- lems in [Procedures 1](#_bookmark25) and [2](#_bookmark24) is a computationally demanding pro- cess, especially for MILP problems. Therefore, implementations of [Procedures 1](#_bookmark25) and [2](#_bookmark24) should include strategies to reduce their overall com- puting time. An attractive approach is to parallelize the implementation of these procedures taking advantage of the modeling and hardware ca-

tion replication (*𝑚* = 1). In contrast, in [Procedure 2](#_bookmark24), by construction, the We note that this first-stage solution is not available in the first optimiza- solution (*𝑥̂𝑁 , 𝑧̂𝑁* ) is available at the beginning of the procedure, thus,

the initialization mentioned in line 2 is active from the first replication. The initialization step entails that in the first iteration of the L-Shaped method, the best first-stage solution is assigned to the first-stage solu- tion, instead of finding a solution using the master problem. With this initialization step, the initial lower bound of the L-Shaped method, cal- culated after solving the LP subproblems, is generally better than the one calculated using the solution of the master problem in the first iter- ation. Note that the corresponding computational gain is multiplied by the number of optimization replications.

In the bound estimation stage, the bounds described in [Sections 4.2](#_bookmark17)–

[4.4](#_bookmark23) are evaluated using the first-stage solutions and the objective func- tion values obtained from the optimization stage. In [Procedure 1](#_bookmark25), the lower bound estimation is computationally more demanding than the upper bound estimation. Therefore, we implement another important time saving strategy: the lower bound estimation is only carried out for

perform the lower bound estimation only for *𝜖*-different first-stage so- new distinct first-stage solutions. Note that an alternative would be to

lutions. This alternative would avoid the computing time for the lower bound estimation for slightly different first-stage solutions. In the dis-

cussion of the results, the impact of avoiding the lower bound estimation for repeated solutions on the computing time is analyzed and situations where it is relevant are identified.

The risk-averse formulations involving the CVaR measure add an ex- tra burden to the L-Shaped method and lower bound estimation. The

tional average of the *𝑄*(*𝑥, 𝑧, 𝜉𝑛*) (defined in [(24)](#_bookmark21)). To reduce this evalu- evaluation of the CVaR measure requires the calculation of the condi-

plemented, where the evaluation of the CVaR for *𝛼 > 𝛼*′ re-uses the in- ation time, an eﬃcient approach in the lower bound estimation is im-

*5.2.2. Continuous formulation*

T P

T Let (Ω*, ,* ) be a probability space, with Ω being the sample space,

a *𝜎*-algebra and P a probability measure. Let *𝐼* = [*𝑎, 𝑏*] be an interval of ℝ. *𝑋* is a square-integrable stochastic process that is assumed to be defined on *𝐼* . The Karhunen–Loève decomposition of the process can be

expressed as

∞

∑ √

*𝑋*(*𝑡*) = *𝜇*(*𝑡*) + *𝜆𝑖 𝑉𝑖* (*𝑡*)*𝜉𝑖 ,* (36)

*𝑖*=1

*𝑖*=1

formation of the distribution of the previous CVaR evaluation for *𝛼*′.

With this overall approach, we provide a full characterization of the

where (*𝑉𝑖* )∞

is a Hilbert basis of *𝐿*2(*𝐼* )*,* the random variables (*𝜉𝑖* )∞ are

*𝑖*=1 ∞

≥ ≥ ≥

solution of our problem, which encompasses the following information:

1. approximation of the optimal solution, in terms of the first-stage variables;

mutually independent, with zero mean and unit variance, and (*𝜆𝑖* )*𝑖*=1

are positive constants in a decreasing order (*𝜆*1 *𝜆*2 ⋯ 0). Under suﬃcient regularity of the covariance *𝐶,* the pairs (*𝜆𝑖 , 𝑉𝑖* )∞ are solutions

*𝑖*=1

of the following Fredholm integral equation of the second kind

∫ ′

1. point estimate and confidence interval for the upper bound on the true optimal objective function value;

*𝑡* ∈*𝐼*

*𝐶*(*𝑡, 𝑡*′)*𝑉𝑖* (*𝑡*′)d*𝑡*′ = *𝜆𝑖 𝑉𝑖* (*𝑡*)*.* (37)

1. point estimate and confidence interval for the lower bound on the

true optimal objective function value;

The random variable *𝜉𝑖* is defined by

1. point estimate on the upper bound on the gap between the true op-

*𝜉* = 1 ∫

*𝑉* (*𝑡*)(*𝑋*(*𝑡*) − *𝜇*(*𝑡*))d*𝑡.* (38)

timal objective function value and the expected optimal objective

*𝑖*

*𝜆𝑖*

√

*𝑖*

*𝑡*∈*𝐼*

*𝑖*=1

function value associated with a first-stage solution;

Note that if *𝑋* is a Gaussian process, then the random variables (*𝜉𝑖* )∞

Therefore, this information provides inference statistics on the solu- tions that are more informative than the single optimal objective func- tion value and corresponding first-stage variables of a deterministic problem.

In the next section, we describe the sampling procedure to generate

follow the standard Gaussian distribution.

1. *2.3. Discretization*

Let (*𝑠𝑖* )*𝑛* be a regular discretization of the interval *𝐼* such that

*𝑖*=0

*𝑎* = *𝑠*0 ≤ *𝑠*1 ≤ ⋯ ≤ *𝑠𝑛* = *𝑏,* and *𝑠𝑖*+1 − *𝑠𝑖* = Δ*𝑡,* (39)

*𝑖*=1

the samples of the wind and electricity prices.

and let (*𝑡𝑖* )*𝑛*

be the midpoints (*𝑠𝑖* − *𝑠*

*𝑖*−1

)∕2. Evaluating the integral in

# Sampling

* 1. *Sampling the electricity prices*

We generate the samples for the electricity prices from an ARIMA model that is fitted to an electricity price time series. Specifically, the samples are drawn by sampling the error term of the ARIMA model; see [Conejo et al. (2010)](#_bookmark58) for a discussion on this type of approach. Specific details for the data used are given in [Section 6.1.1](#_bookmark32).

* 1. *Sampling the wind speed with a truncated Karhunen–Loève expansion*

In this section, we briefly introduce the KLE, and in [Section 6](#_bookmark29),

[Eq. (37)](#_bookmark26) with a midpoint-quadrature formula at every *𝑡𝑗* yields

*𝑛*

∑

*𝐶*(*𝑡𝑖 , 𝑡𝑗* )*𝑉𝑖* (*𝑡𝑗* )Δ*𝑡* = *𝜆𝑖 𝑉𝑖* (*𝑡𝑖* )*.* (40)

*𝑗*=1

Let the matrix *𝐾* ∈ ℝ*𝑛*×*𝑛* and the vector *𝑊𝑖* ∈ ℝ*𝑛* be defined by

*𝐾𝑖𝑗* = Δ*𝑡𝐶*(*𝑡𝑖 , 𝑡𝑗* )*,* and *𝑊𝑖𝑗* = *𝑉𝑖* (*𝑡𝑗* )*.* (41)

The discretized Fredholm integral equation leads to the following alge- braic eigenvalue problem

*𝐾𝑊𝑖* = *𝜆𝑖𝑊𝑖 .* (42)

The time discretization followed by a truncation of the sum yield there- fore the approximations

we provide additional information regarding the wind data used. See [Maître and Knio (2010](#_bookmark73), Section 2.1) for a presentation on KLE.

*𝑋*(*𝑡𝑖* ) ≈ *𝜇*(*𝑡𝑖* ) +

*𝑛 *

*𝜆𝑗 𝑊𝑗𝑖𝜉𝑗* ≈ *𝜇*(*𝑡𝑖* ) +

∑ √

*𝑗*=1

*𝑟*

*𝑗*=1

∑

√*𝜆𝑗 𝑊𝑗𝑖𝜉𝑗 .* (43)

* + 1. *Wind speed model*

Assuming that the wind speed *ℎ*(*𝑡*) is always strictly positive, we

adopt here a log-normal model where *𝑋*(*𝑡*) = log(*ℎ*(*𝑡*)) is a Gaussian pro-

Given a tolerance *𝜖,* the constant *𝑟* is chosen such that

1

( ∑ *𝜆𝑖*

)

*𝑖*=*𝑟*+1 2 ≤

∑

*𝑛*

cess with mean function *𝜇*(*𝑡*) and covariance function *𝐶*(*𝑡, 𝑡*′)*,* which are

*𝑛*

*𝑖*=1

*𝜆𝑖*

*𝜖,* (44)

approximated in our application using a sample (*𝑥𝑖*(*𝑡*))*𝑁* and the empir- 2

ical estimation

*𝑖*=1

ensuring that the relative *𝐿*

error between the *𝑛*-terms and the *𝑟*-terms

*𝜇*(*𝑡*) ≈

1 *𝑁*

*𝑁 𝑖*=1

∑

*𝑥𝑖*(*𝑡*) (33)

*𝑁*

approximations is less than *𝜖*.

# Numerical experiments

The results presented in this section demonstrate the new features

and *𝐶*(*𝑡, 𝑡*′) ≈ 1 ∑(*𝑥𝑖*(*𝑡*) − *𝜇*(*𝑡*))(*𝑥𝑖*(*𝑡*′) − *𝜇*(*𝑡*′))*.* (34)

*𝑁* − 1 *𝑖*=1

Given a time interval uniformly discretized (*𝑡𝑖* )*𝑛* with a resolution

*𝑖*=1

Δ*𝑡* = 1 h, the rest of [Section 5.2](#_bookmark28) is dedicated to the KLE-based approxi-

mation of the wind speed model under the form

implemented in the proposed procedures and provide a complete set of statistical results to better support the decision-maker. The objectives of the experiments are fourfold:

1. to demonstrate the positive impact of the proposed strategies on the computing time;

*ℎ*(*𝑡𝑖* ) ≈ exp

(*𝜇*(*𝑡𝑖* ) +

*𝑟*

*𝑖*=*𝑗*

∑

√*𝜆𝑗 𝑊𝑗𝑖𝜉𝑗*

)*,* (35)

1. to analyze the performance of the optimization stage for risk-averse approaches;
2. to compute high-quality solutions for the VPP problem;

where (*𝜉𝑖* )*𝑟*

*𝑖*=1

are independent random variables drawn according to the

1. to assess those solutions using statistic estimators and confidence

standard normal distribution. intervals.

**Table 2**

Summary of the specifications of the thermal unit in Cases 1 and 2.

|  |  |  |
| --- | --- | --- |
| Thermal unit specifications | Case 1 | Case 2 |
| Initial state | On | Off |
| Maximum power (MW) | 455 | 55 |
| Fixed generation cost (€ /h) | 1000 | 660 |
| Variable generation cost (€ /MWh) | 16.19 | 25.92 |
| Minimum up-time (h) | 8 | 1 |
| Minimum down-time (h) | 8 | 1 |

The first and second objectives above are related to the first contribu- tion of this work and the third and fourth objectives to the second con- tribution. In terms of the organization of the results, we discuss the com- putational performance of [Procedure 1](#_bookmark25) in Section in [6.2](#_bookmark31), the objective function values in [Section 6.3](#_bookmark40), the first-stage solutions in [Section 6.4](#_bookmark46), and the maximum gaps obtained in [Section 6.5](#_bookmark45).

sample size *𝑁* on the computing time of the L-Shaped method and the More specifically, in [Procedure 1](#_bookmark25), we analyze the impact of a) the variance of the solutions; b) the sample size *𝑁* ′ on the computing time and variance of the solutions; c) the number of replications *𝑇* and *𝑇* ′ on

the sample size *𝑁* on the upper bound estimation. the lower bound estimation. In [Procedure 2](#_bookmark24), we study the influence of

In a supplementary document, detailed experimental results that support the ones presented in this section are provided.

* 1. *Description of the case studies*

We consider two cases pertaining to the VPP described in [Section 2](#_bookmark8), which we refer to as Cases 1 and 2. In both cases a similar wind farm and a similar pumped-storage hydro unit are considered. However, the specifications of the thermal unit in each case are different in terms of initial state, operational constraints, and costs. [Table 2](#_bookmark30) specifies the differences in terms of capacity and costs between the thermal unit in Cases 1 and 2.

For the same samples of the wind speed and electricity prices, these differences induce a distinct reaction of the VPP. Furthermore, the higher generation cost of the thermal unit in Case 2 leads to a situa- tion where the gap between the electricity price and generation costs is smaller than the gap in Case 1. The results presented in this section show that the two cases are suﬃciently distinct to give a broad view on the performance of the methods used in this work.

The solutions of interest for the VPP problem are the optimal first- stage solutions and the performance indices expected profit and the CVaR of the profit for different quantiles. Given the number of elements in each sample, and the fact that in practice, the second-stage variables have to be adjusted to the future realization of the random variables, these variables are less meaningful and thus their values are not re- ported.

* + 1. *Wind speed and electricity prices data*

We consider a time horizon of 168 h, which corresponds to the week of August 25–31, 2014. The ARIMA model uses the same structure and fit of the model proposed in [Lima et al. (2018)](#_bookmark68). The electricity price time series has 12 weeks before the week studied, from the Iberian Peninsula electricity market ([Iberian Electricity Market, 2015](#_bookmark62)).

The raw data for the wind speed for a specific wind farm location consists of a wind speed ensemble with 51 members obtained from the European Centre for Medium Range Weather Forecasts (ECMWF). We use the wind speed ensemble and the KLE to generate additional samples of the wind speed for the SAA methodology. The electricity prices are considered independent of the wind speed available at the location of the considered VPP, due to the small capacity of the VPP in each case study, which has no market power to influence the electricity prices.

**Table 3**

Size of the extensive form and problems within the L-Shaped method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Problem | *𝑁* | *𝛽* | NCNST | NVAR | 0-1 NVAR |
| Extensive | 10 | 0.0 | 21,346 | 14,298 | 508 |
| Extensive | 5000 | 0.0 | 10,081,186 | 6,720,858 | 508 |
| Extensive | 5000 | 1.0 | 10,086,186 | 6,725,859 | 508 |
| Master† | 10 | 0.0 | 1187 | 859 | 508 |
| Master† | 5000 | 0.0 | 1187 | 859 | 508 |
| Master† | 5000 | 0.5 | 6188 | 5861 | 508 |
| Master† | 5000 | 1.0 | 6187 | 5860 | 508 |
| Subproblem*⋆* | – | – | 1345 | 2353 | – |

† – Sizes refer to the second iteration of the L-Shaped method. *⋆* – Size refers to the dual of [(24)](#_bookmark21). NCNST – number of constraints plus objec-

tive function; NVAR – number of total variables; 0–1 NVAR – number of binary variables.

* + 1. *Setup of the parameters of the procedures*

The risk-neutral and risk-averse stochastic programming problems are solved with an L-Shaped based method with single optimality cuts.

10,800 s, a maximum gap between the bounds of 1×10−4%*,* and a max- For this method, the stop criteria are a maximum wall-clock time of

imum number of iterations of 5000. Note that this gap is related to the bounds on the objective function value within the L-Shaped method for a given sample. These bounds are not related to the bounds described in [Sections 4.2](#_bookmark17) and [4.3](#_bookmark19).

following parameters *𝑀* = 30 (number of optimization replications), We perform a sensitivity analysis with [Procedure 1](#_bookmark25) combining the

*𝑁* ∈ {10*,* 50*,* 100*,* 500*,* 5000} (size of the samples used in each optimiza- tion replication), *𝑇* ∈ {10*,* 30}*, 𝑇* ′ ∈ {10*,* 30} (number of replications in the lower bound estimation), *𝑁* ′ ∈ {5000*,* 25*,* 000} (size of the samples

used in the lower bound estimation).

We consider four combinations of (*𝛽, 𝛼*) ∈

{(0*,* −)*,* (0*.*5*,* 0*.*9)*,* (1*,* 0*.*9)*,* (1*,* (*𝑁* − 1)∕*𝑁* )} for the optimization replica- tion. The case with *𝛼* = (*𝑁* − 1)∕*𝑁* corresponds to the maximization

of the worst profit consistently with the definition of the CVaR. In the lower bound estimation, for each first-stage solution we evaluate point estimates and confidence intervals of the CVaR of the profit for

*𝛼* ∈ {0*.*9*,* 0*.*95*,* (*𝑁* ′ − 1)∕*𝑁* ′}.

For [Procedure 2](#_bookmark24), we set *𝑀* = 30 (number of replications of the op- timization) and *𝑁* ∈ {500*,* 5000} (size of the samples). All confidence

intervals presented correspond to 95%.

A workstation with 40 Intel Xeon CPU E5-2680 v2 @ 2.80 GHz pro- cessors, and 125.8 Gb of RAM was used and the solution of the LP sub- problems was distributed among the 40 CPUs. The MILP and LP prob- lems were solved with CPLEX 12.7.1.0 using the GAMS/GRID/GUSS ca- pabilities to distribute the solution of the LP subproblems.

* 1. *Computational performance*

In this subsection, we first present the dimensions of the problems considered and discuss the choice of the L-Shaped method over the di- rect solution of the extensive form. Next, the benefits of the proposed initialization of the L-Shaped method are shown and then analysis on the overall performance of [Procedure 1](#_bookmark25) is provided.

lions of constraints and variables for samples with *𝑁* = 5000*,* and that In [Table 3](#_bookmark30), we show that the size of the extensive form reaches mil-

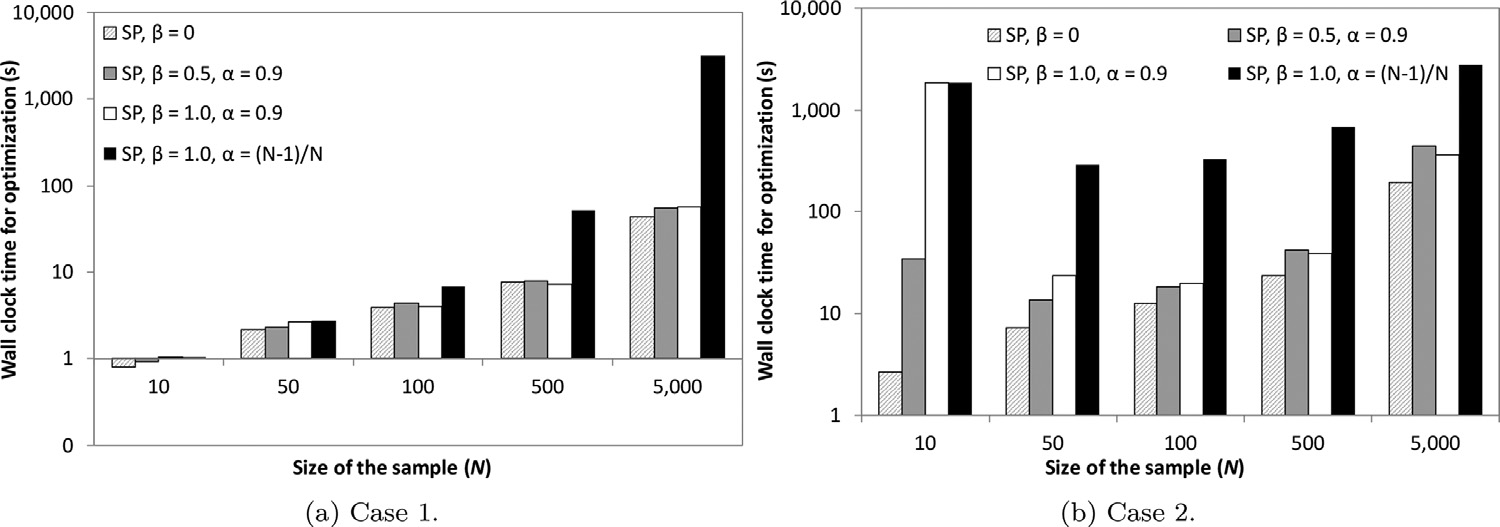
master problem and *𝑁* smaller subproblems. the L-Shaped method decomposes the extensive form into one smaller

We adopt the L-Shaped variant that provided the best performance in [Lima et al. (2018)](#_bookmark68), i.e., Algorithm 1, which corresponds to using a single optimality cut for each of the expectation and CVaR operators and calculating VaR based on the solution of the subproblems, rather than

using VaR as a first-stage variable. In that work it is shown that for *𝛽* = 0*,*

the L-Shaped method with an eﬃcient parallelization solution of the

the extensive form. Whereas for *𝛽* ∈ {0*.*1*,* 0*.*5*,* 0*.*9*,* 1*.*0}*, 𝛼* = 0*.*9*,* sample subproblems is one order of magnitude faster than the direct solution of

**Fig. 1.** Average wall-clock time for 30 opti- mization replications for the combinations of sample sizes and risk parameters studied.

**Table 4**

Ratio between the total wall-clock time for 30 optimization repli-

solution (T0) in the L-Shaped method. *𝑀* = 30. cations using an initial solution (Tinit) and without using an initial

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | *𝑁* = 500 |  |  | *𝑁* = 5000 |  |  |
| Case 1 | Case 2 |  | Case 1 | Case 2 |
| *𝛽* | *𝛼* | Tinit/T0 | Tinit/T0 |  | Tinit/T0 | Tinit/T0 |
| 0 | – | 0.68 | 0.82 |  | 0.63 | 0.82 |  |
| 0.5 | 0.9 | 0.52 | 0.75 |  | 0.56 | 0.80 |  |
| 1 | 0.9 | 0.56 | 0.78 |  | 0.61 | 0.81 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 0.96 | 1.19 |  | 1.06 | 1.09 |  |

sizes of 5100 elements, and maximum wall clock time of 7200 s, the MILP solver applied to the extensive form did not find a feasible solution for four case studies similar to the two considered in this work. We refer the reader to that work for detailed computational results showing the superior performance of the L-Shaped method over solving directly the extensive form.

* + 1. *Initialization of the optimization stage*

[Table 4](#_bookmark34) shows the ratios between the total wall-clock time with and without initialization for 30 optimization replications.

These results cover four combinations of risk parameters using two sample sizes – 500 and 5000 elements – over 30 replications. A ratio smaller than one means that the initialization is effective in reducing the total computing time. Overall, the results suggest that the initializa- tion reduces the required computing time. The exception is the formula-

tion with *𝛽* = 1 and *𝛼* = (*𝑁* − 1)∕*𝑁,* (last row of [Table 4](#_bookmark34)). This peculiar

behavior occurs because convergence is not reached for the first-stage

solution over the replications, and thus, one solution from one repli- cation is not necessarily a good initial solution for another replication. For the remaining risk metrics, there is convergence for the first-stage variables, and therefore, the initialization has a positive impact on com- puting time. This is further discussed in [Section 6.2.4](#_bookmark37).

* + 1. *Performance of the optimization stage*

[Fig. 1](#_bookmark33) a and b shows for Cases 1 and 2, respectively, the average wall-clock time for 30 optimization replications as a function of the risk parameters and size of the sample. These results show that with a sam- ple with 10 elements, the average wall-clock time for Case 1 is approxi- mately 1 s, independently of the risk parameters. However, this comput-

ing time increases with the sample size, with the formulation with *𝛽* = 1*,*

*𝛼* = (*𝑁* − 1)∕*𝑁* exhibiting a steeper increase for an average of 3146 s.

For Case 2 the average wall-clock time for the formulations with the

cuses on the results of the formulations with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* . CVaR is independent of the size of the sample. The next sub-section fo-

* + 1. *Performance of the optimization stage for 𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁*

The formulation with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* exhibits some relevant

results that are worth discussing. First of all, for the formulations with

which requires resources, however, for *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* this evalu- the CVaR, the L-Shaped method evaluates the CVaR in each iteration,

ation is straightforward. In general, comparing the formulations with

*𝛽* = 0 and *𝛽 >* 0*, 𝛼 <* (*𝑁* − 1)∕*𝑁* in terms of computing time, we can

demanding. For example, in [Fig. 1](#_bookmark33)a for *𝑁* = 500*,* the computing times see that the CVaR specific calculation is not computationally the most for *𝛽* = 0 and *𝛽 >* 0*, 𝛼 <* (*𝑁* − 1)∕*𝑁* are similar. On the other hand, for

*𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* the average number of iterations is 5.7, whereas for *𝛽* = 1*, 𝛼* = 0*.*9 it is 2.0. Therefore, the higher computing times ob- tained with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* do not arise from the CVaR evalua-

tion, but rather from the number of iterations executed by the L-Shaped method, as discussed below.

For Case 1, for *𝑁* = 10*,* the L-Shaped method requires on average 2.1

iterations to meet the stopping criteria, while for *𝑁* = 5000 it requires on

average 8.6 iterations. [Fig. 2](#_bookmark35) presents the profiles of the bounds in the

L-Shaped method for the optimization replications with the lower and higher number of iterations within each optimization replication with

*𝑁* = 10 and *𝑁* = 5000. The figures for Case 1 show that for *𝑁* = 10 the

number of iterations ranges from 2 to 4, while for *𝑁* = 5000 it ranges

demands extra iterations, since for *𝛽* = 0 and *𝑁* = 5000*,* the L-Shaped from 2 to 21 iterations. However, it is not the size of the sample that

method requires on average 2.0 iterations to meet the stopping criteria. It is rather the elements within large samples that have an impact on

of the CVaR of the profit with *𝛼* = (*𝑁* − 1)∕*𝑁* . the performance of the L-Shaped method if applied to the maximization

For Case 2, [Fig. 1](#_bookmark33)b shows that the average wall-clock times for the formulations with the CVaR are independent of the size of the sample.

In the specific case of *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁,* for *𝑁* = 10*,* the L-Shaped

ria, whereas for *𝑁* = 5000 it requires on average 11.4 iterations. These method requires on average 101.2 iterations to meet the stopping crite-

results combined with the results in [Fig. 1](#_bookmark33)b indicate that 1) Case 2 forces the L-Shaped method to perform more iterations than Case 1, for the same samples, which is explained by the higher generation costs of the thermal unit in Case 2; and 2) consequently, for the formulations that maximize the worst profit observation, the L-Shaped method requires additional iterations than the other formulations. [Fig. 2](#_bookmark35)c and d shows

number of iterations obtained with *𝑁* = 10 and *𝑁* = 5000. From these the profiles of the optimization replications with the lower and higher

quires few iterations. For example, 10 and 6 for *𝑁* = 10 and *𝑁* = 5000*,* figures, it is also clear that in some replications the L-Shaped method re-

25 for *𝑁* = 10 and *𝑁* = 5000*,* respectively. respectively, while in others it may require a large number – 417 and

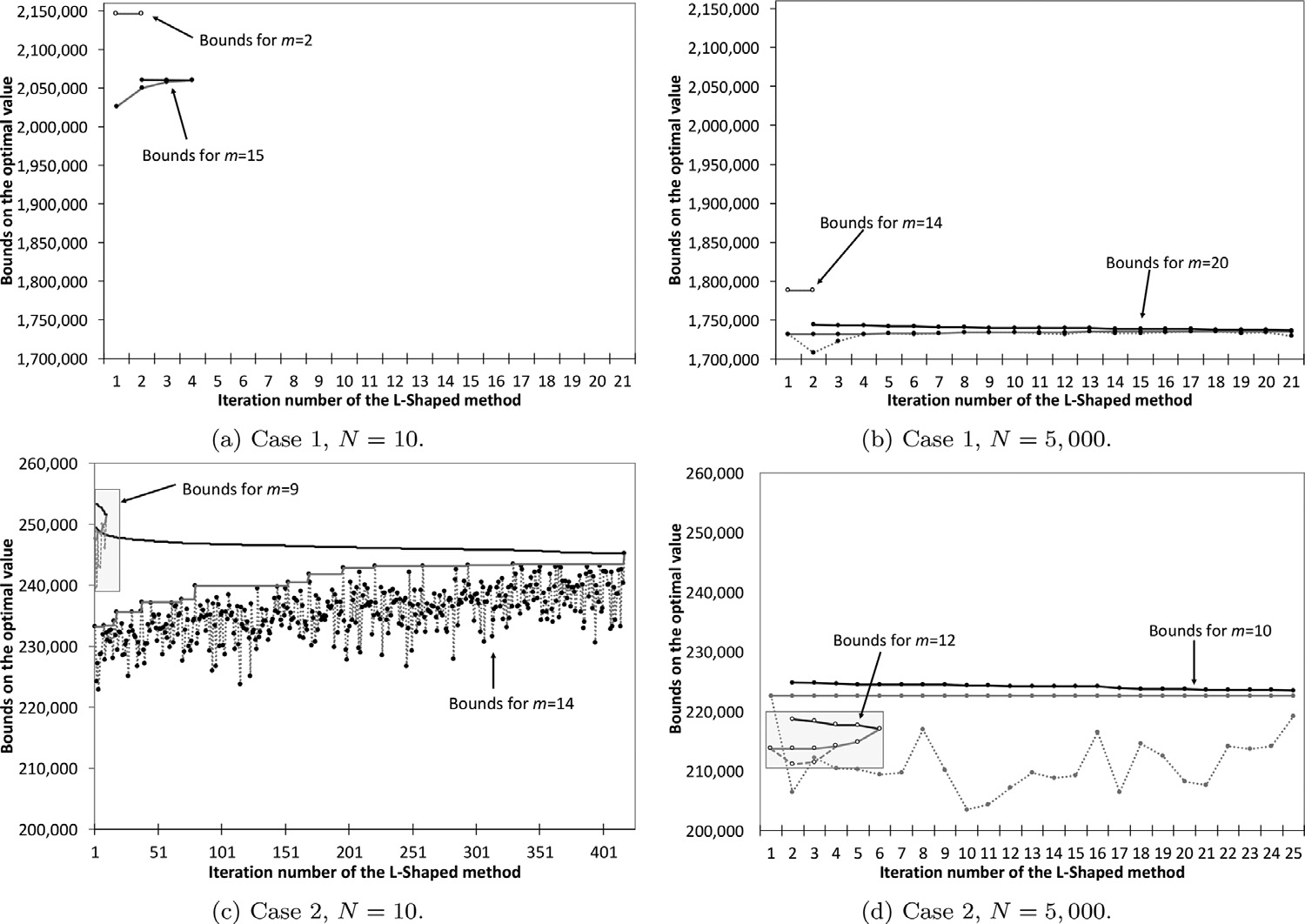
Overall, these results show that:

1. the maximization of the CVaR of the profit with *𝛼* = (*𝑁* − 1)∕*𝑁*

son with *𝛽* = 0. The reason is that in the former problem, the search forces the L-Shaped method to perform more iterations, by compari-

is driven by the worst profit, which forces the L-Shaped method to search for the optimal commitment of the thermal unit or contract se-

case profit. By contrast, the problem with *𝛽* = 0 is driven by the max- lection to respond to the low electricity prices that induce the worst

**Fig. 2.** Upper and lower bounds (continuous lines) from the L-Shaped method. The dashed line is the objective function for the first-stage solution. In each sub-figure, there are the two

higher number of iterations. For *𝑀* = 30*, 𝛽* = 1*,* optimization replications with the lower and and *𝛼* = (*𝑁* − 1)∕*𝑁* .

imization of the expected profit that is less sensitive to low electricity price observations;

1. the size of the samples have an impact on the performance of the L-Shaped method because increasing the sample size increases the likelihood of observations with lower electricity prices;

**Table 5**

Case 1. Number of successful optimization replications and distinct solutions.

*𝑀* = 30.

*𝛽* = 0 *𝛽* = 0*.*5*, 𝛼* = 0*.*9 *𝛽* = 1*, 𝛼* = 0*.*9 *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁*

*𝑁* SR DS SR DS SR DS SR DS

1. the thermal plant with higher generation costs induces more itera- tions of the L-Shaped method by comparison with the thermal plant

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10 | 30 | 6 | 30 | 6 | 30 | 3 | 30 | 3 |
| 50 | 30 | 4 | 30 | 4 | 30 | 2 | 30 | 3 |
| 100 | 30 | 3 | 30 | 3 | 30 | 2 | 30 | 3 |
| 500 | 30 | 1 | 30 | 1 | 30 | 2 | 30 | 10 |
| 5,000 | 30 | 1 | 30 | 1 | 30 | 1 | 25 | 18 |

with lower generation costs; and

1. the variability in the number of iterations required by the L-Shaped method in each optimization replication is induced by the differ-

ent characteristics of the observations with lower electricity prices among the samples. By different characteristics, we mean that across the samples, the lower electricity prices occur on different hours of the week and distinct minimum values.

* + 1. *Combined performance of the optimization and bound estimation* *stages*

[Fig. 3](#_bookmark38) presents the total wall-clock time required by [Procedure 1](#_bookmark25) and by its two main stages: optimization and bound estimation. This fig-

ure compares the computational performance obtained with *𝛽* = 0 and

*𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* . It also shows the impact on the computing time of

bound estimation, from {*𝑇 , 𝑇* ′} = 10 and *𝑁* ′ = 5000 to {*𝑇 , 𝑇* ′} = 30 and increasing the number of replications and the sample size in the lower

*𝑁* ′ = 25*,* 000. Using *𝑁* ′ = 25*,* 000 and {*𝑇 , 𝑇* ′} = 30 the overall time in- creases, by comparison with *𝑁* ′ = 5000 and {*𝑇 , 𝑇* ′} = 10*,* but there is no

influence on the optimization time.

These results provide relevant insights regarding the trade-off be- tween the sample size and the overall time required. We can observe that increasing the sample size in the optimization, the overall time

does not monotonically increase. In fact, for *𝛽* = 0*,* there is a minimum

overall time for *𝑁* = 500*,* instead of *𝑁* = 10; see [Fig. 3](#_bookmark38)a and c. This be-

havior is explained by noting that: 1) the lower bound estimation is

replication; and 2) for *𝑁* = 10*,* multiple distinct solutions are obtained, only performed when a new solution is obtained from an optimization whereas for *𝑁* = 500*,* only one distinct solution is obtained. In [Fig. 3](#_bookmark38), as

*𝑁* increases the time required by the optimization stage increases, but not the lower bound estimation time. For *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁,* there is also a minimum time that corresponds to *𝑁* = 10 for *𝑁* ′ = 5000 and

SR – number of successful optimization replications that meet the stop criteria (out of 30); DS – number of distinct solutions.

{*𝑇 , 𝑇* ′} = 10*,* and *𝑁* = 50 for *𝑁* ′ = 25*,* 000 and {*𝑇 , 𝑇* ′} = 30; see [Fig. 3](#_bookmark38)b and d.

of distinct solutions obtained as a function of *𝑁* and the risk parameters We complement these results by presenting in [Table 5](#_bookmark36) the number

for Case 1.

This table shows that increasing the sample size, the problem con-

except for *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* . For *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁,* increasing verges to one distinct solution for all combinations of risk parameters;

cific situation, *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁,* the objective is to maximize the the sample size increases the number of distinct solutions. In this spe-

first-stage solutions for *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* is the hourly commitment single worst observation of the profit. The primary difference among the

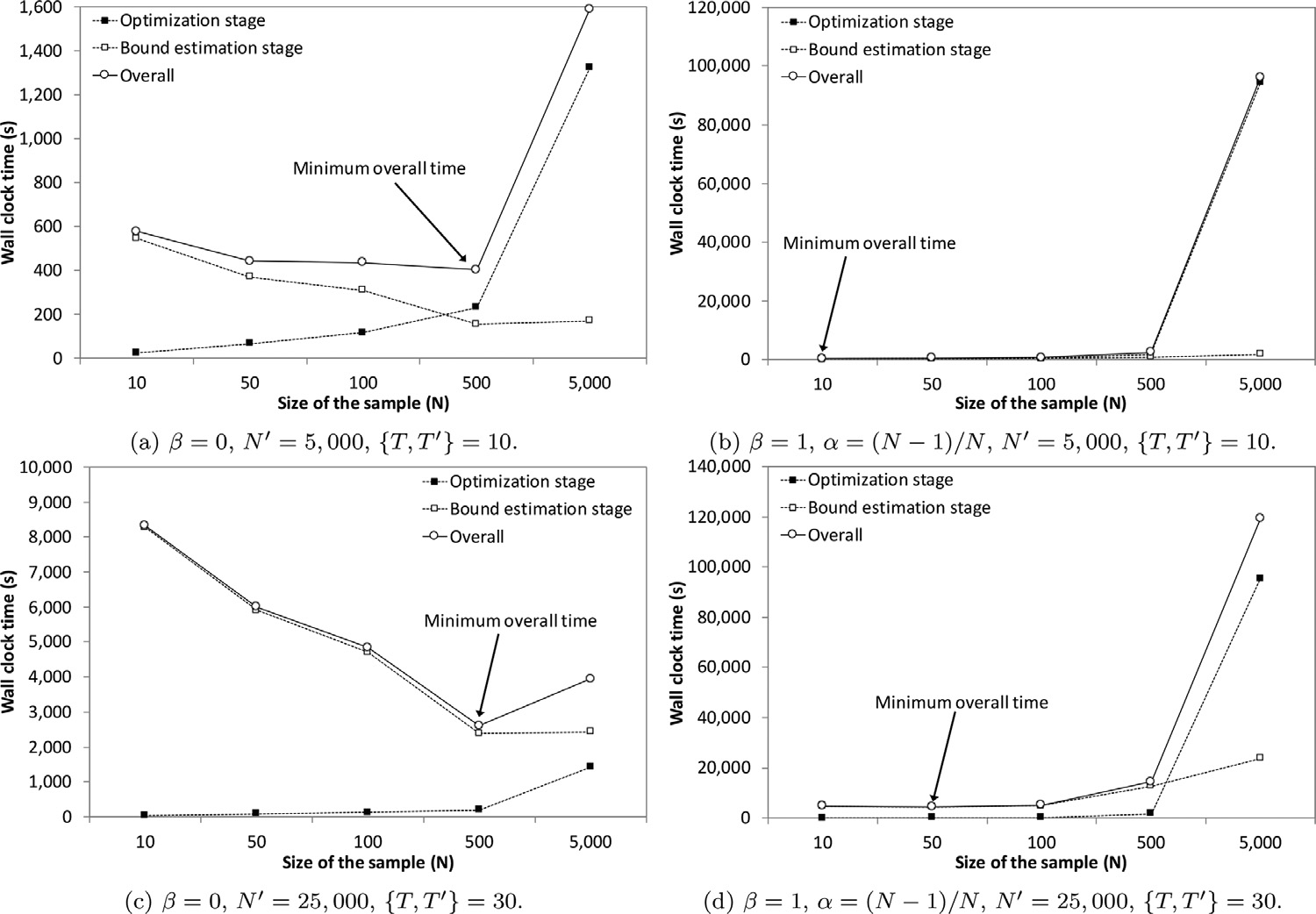
of the thermal plant; see also additional results available in the supple- mentary document. This difference suggests that the thermal plant is prone to shutdown and startup to follow the lowest electricity price ob- servation. The samples with more elements have a higher likelihood of lower electricity prices, which from one sample to the other may occur at different hours of the week. Therefore, distinct first-stage solutions emerge among the optimization replications.

The number of distinct solutions obtained with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁*

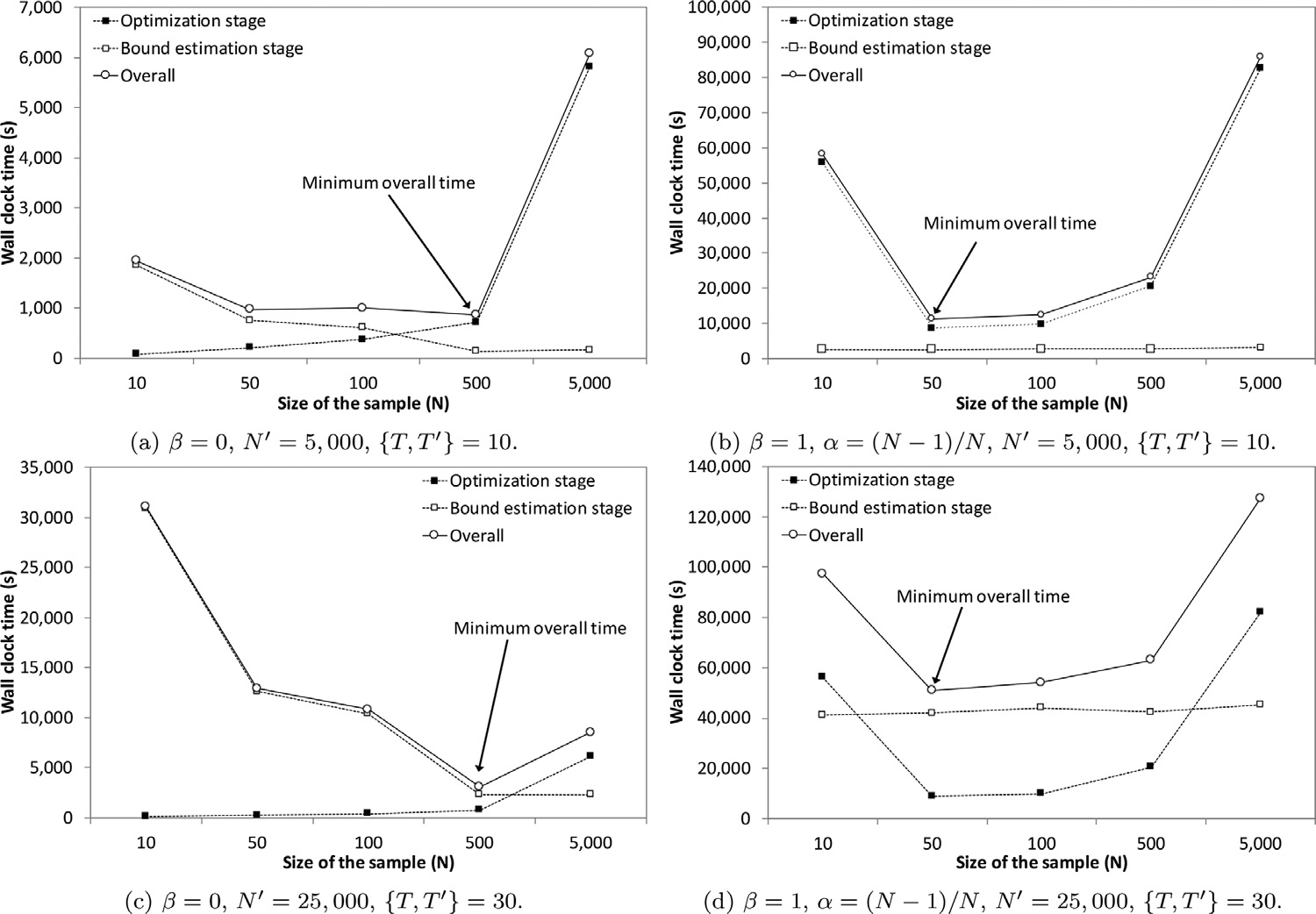
for *𝑁* = 500 also justifies the worst performance of the initialization im-

plemented in the L-Shaped method. Thus, one solution from one opti-

mization replication may not be a good starting point for an optimiza- tion replication with a different sample.

**Fig. 3.** Case 1. Wall-clock time as a function of the size of the sample for the optimization.

*𝑀* = 30.

**Fig. 4.** Case 2. Wall-clock time as a function of the size of the sample for the optimization.

*𝑀* = 30.

For Case 2 with *𝛽* = 0*,* the results show that increasing the sample

size for the optimization, the number of distinct solutions decreases,

[Fig. 4](#_bookmark39)a and c. However, for *𝛽* = 1 with *𝛼* ∈ {0*.*9*,* (*𝑁* − 1)∕*𝑁* }*,* the number which means a decrease in the time for the lower bound estimation; see

of distinct solutions is equal to the number of optimization replications, which means that there is no convergence to a distinct solution; see [Table 6](#_bookmark43). Therefore, the computing time for the lower bound estimation is independent of the sample size for the optimization; see [Fig. 4](#_bookmark39)b and d.

* 1. *Objective function values and its bounds*

value *𝑤*∗ for the different formulations as a function of the samples size This section presents the bounds on the optimal objective function

and the number of replications. [Figs. 5](#_bookmark41) and [6](#_bookmark42) show the limits of the

mal objective function value *𝑤*∗ obtained for Case 1 and Case 2, respec- confidence intervals for the upper and lower bounds on the true opti-

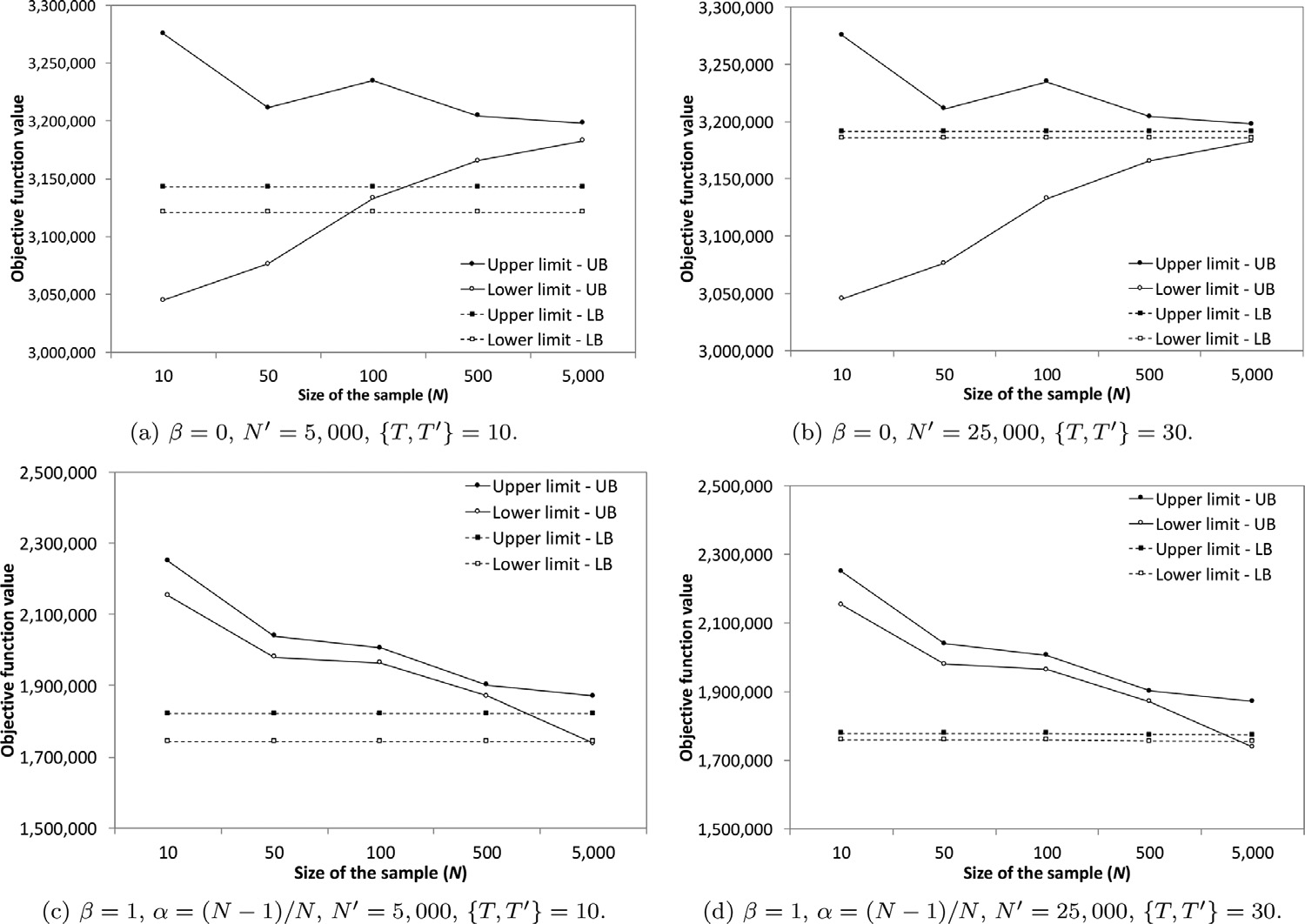
tively. These figures illustrate the impact of the sample size and number

tervals for *𝛽* = 0 and *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* . In the supplementary results of replications used in the lower bound estimation on the confidence in-

document, the bounds for additional risk parameters are presented.

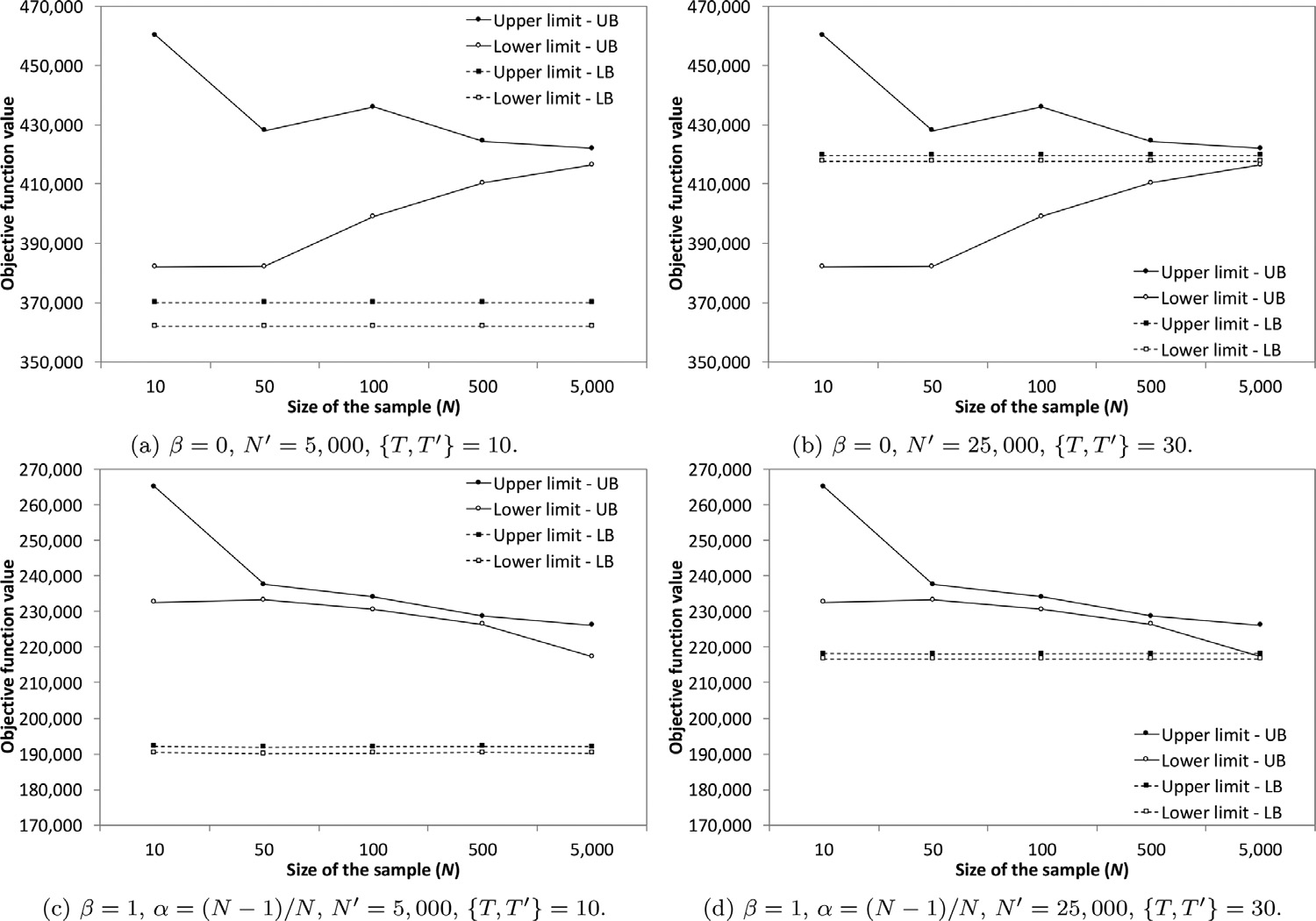
expected profit (*𝛽* = 0), the point estimate of the upper bound on the true A relevant result shows that for the formulation that maximizes the optimal objective function value *𝑤*∗ falls within the confidence interval of the upper bound for all *𝑁* tested. Increasing the sample size reduces

confidence intervals. For the formulation with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁,* the the confidence interval, converging approximately to the center of the

**Fig. 5.** Case 1. Limits of the confidence inter- vals on the point estimate of the lower and up-

per bounds (*𝐿𝐵𝑁*′ *,𝑇* ′ and *𝑈𝐵𝑁,𝑀* ) on the true

optimal objective function value (*𝑤*∗) as a func- tion of risk parameters, *𝑁* ′*, 𝑇 ,* and *𝑇* ′. *𝑀* = 30.

**Fig. 6.** Case 2. Limits of the confidence inter- vals on the point estimate of the lower and up-

per bounds (*𝐿𝐵𝑁*′ *,𝑇* ′ and *𝑈𝐵𝑁,𝑀* ) on the true

optimal objective function value (*𝑤*∗) as a func- tion of risk parameters, *𝑁* ′*, 𝑇 ,* and *𝑇* ′. *𝑀* = 30.

**Table 6**

Case 2. Number of successful optimization replications and distinct solutions.

*𝑀* = 30.

*𝛽* = 0 *𝛽* = 0*.*5*, 𝛼* = 0*.*9 *𝛽* = 1*, 𝛼* = 0*.*9 *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝑁* | SR | DS |  | SR | DS |  | SR | DS |  | SR | DS |  |
| 10 | 30 | 26 |  | 30 | 17 |  | 28 | 30 |  | 28 | 30 |  |
| 50 | 30 | 10 |  | 30 | 30 |  | 30 | 30 |  | 30 | 30 |  |
| 100 | 30 | 8 |  | 30 | 30 |  | 30 | 30 |  | 30 | 30 |  |
| 500 | 30 | 1 |  | 30 | 23 |  | 30 | 30 |  | 30 | 30 |  |
| 5,000 | 30 | 1 |  | 30 | 30 |  | 30 | 30 |  | 27 | 30 |  |

SR – number of successful optimization replications that meet the stop criteria (out of 30); DS – number of distinct solutions.

point estimate of the upper bound overestimates the true optimal objec-

confidence intervals (for different *𝑁* ). Thereby, increasing the sample tive function value, and the point estimate is not included in all the

size for the optimization decreases the value of the point estimate and range of the confidence intervals.

[Figs. 5](#_bookmark41) and [6](#_bookmark42) indicate that the confidence interval for the lower

with the samples with *𝑁* ′ = 5000*,* {*𝑇 , 𝑇* ′} = 10. For example, this can be bound on the true optimal objective function value is underestimated

seen comparing in [Fig. 5](#_bookmark41)a and b the confidence interval for the lower bound represented by the dashed lines. These results indicate that for

*𝑇* ′ = 10 there is a bias in the estimation of the lower bound on the true

optimal objective function value. However, as *𝑁* ′ is increased to 25,000

**Table 7**

Case 1. First-stage solutions and point estimates of the lower bound on the true optimal objective function value (*𝐿𝐵𝑁*′ *,𝑇* ′ ) with confidence intervals. *𝑁* ′ = 25*,* 000*,* {*𝑀, 𝑇 , 𝑇* ′} = 30.

First-stage solutions (aggregated) CI, [LLB,ULB]

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝛽* | *𝛼* | *𝑁* | UT (%) | SUP | SD | SELLC (MW) | BUYC (MW) | *𝐿𝐵𝑁*′ *,𝑇* ′ (€) | LLB (€) | ULB (€) |  |
| 0 | – | 10 | 100 | 0 | 0 | 0 | 160 | 3,188,599 | 3,185,869 | 3,191,329 |  |
| 0 | – | 50 | 100 | 0 | 0 | 0 | 160 | 3,188,599 | 3,185,869 | 3,191,329 |  |
| 0 | – | 100 | 100 | 0 | 0 | 0 | 160 | 3,188,599 | 3,185,869 | 3,191,329 |  |
| 0 | – | 500 | 100 | 0 | 0 | 0 | 160 | 3,188,599 | 3,185,869 | 3,191,329 |  |
| 0 | – | 5000 | 100 | 0 | 0 | 0 | 160 | 3,188,599 | 3,185,869 | 3,191,329 |  |
| 0.5 | 0.9 | 10 | 100 | 0 | 0 | 155 | 0 | 2,454,870 | 2,453,669 | 2,456,071 |  |
| 0.5 | 0.9 | 50 | 100 | 0 | 0 | 155 | 0 | 2,454,870 | 2,453,669 | 2,456,071 |  |
| 0.5 | 0.9 | 100 | 100 | 0 | 0 | 155 | 0 | 2,454,870 | 2,453,669 | 2,456,071 |  |
| 0.5 | 0.9 | 500 | 100 | 0 | 0 | 155 | 0 | 2,454,870 | 2,453,669 | 2,456,071 |  |
| 0.5 | 0.9 | 5000 | 100 | 0 | 0 | 155 | 0 | 2,454,870 | 2,453,669 | 2,456,071 |  |
| 1 | 0.9 | 10 | 100 | 0 | 0 | 315 | 0 | 2,100,859 | 2,100,156 | 2,101,562 |  |
| 1 | 0.9 | 50 | 100 | 0 | 0 | 315 | 0 | 2,100,859 | 2,100,156 | 2,101,562 |  |
| 1 | 0.9 | 100 | 100 | 0 | 0 | 315 | 0 | 2,100,859 | 2,100,156 | 2,101,562 |  |
| 1 | 0.9 | 500 | 100 | 0 | 0 | 315 | 0 | 2,100,859 | 2,100,156 | 2,101,562 |  |
| 1 | 0.9 | 5000 | 100 | 0 | 0 | 315 | 0 | 2,100,859 | 2,100,156 | 2,101,562 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 10 | 100 | 0 | 0 | 315 | 0 | 1,768,962 | 1,759,967 | 1,777,956 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 50 | 100 | 0 | 0 | 315 | 0 | 1,768,962 | 1,759,967 | 1,777,956 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 100 | 0 | 0 | 315 | 0 | 1,768,962 | 1,759,967 | 1,777,956 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 500 | 95.2 | 1 | 1 | 315 | 0 | 1,764,836 | 1,755,568 | 1,774,104 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 5000 | 95.2 | 1 | 1 | 315 | 0 | 1,764,554 | 1,755,197 | 1,773,911 |  |

CI – Confidence interval, LLB and ULB as in [(20)](#_bookmark22), UT – percentage of up-time of the thermal unit, SUP/SD – number of startups/shutdowns of the thermal unit, SELLC/BUYC – power sold/bought through contracts.

and {*𝑇 , 𝑇* ′} to 30, the confidence interval of the lower bound converges with the confidence interval of the upper bound.

Note that for the optimization replications where the L-Shaped method did not meet the stopping criteria, the upper bound from the L-Shaped method is used to evaluate the point estimate on the upper bound on the true optimal solution. This replacement justifies the larger range of the confidence interval for the upper bound for Case 1 with

*𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁, 𝑁* = 5000 in [Fig. 5](#_bookmark41)c and d, and for Case 2 with

*𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁, 𝑁* ∈ {10*,* 5000} in [Fig. 6](#_bookmark42)c and d.

Overall, the convergence of the bounds for *𝑁* ′ = 25*,* 000*,* {*𝑇 , 𝑇* ′} = 30 indicate that high-quality solutions are obtained, which are further

assessed in the next sub-section.

* 1. *First-stage solutions*

[Tables 7](#_bookmark44) and [8](#_bookmark47) provide the best first-stage solution and inference statistics resulting from each set of risk parameters and sample size.

The first stage solutions are given in aggregated form to simplify their presentation. Therefore, to distinguish between equal aggregated solutions that correspond to distinct disaggregated solutions, the point estimates of their lower bounds are provided.

For Case 1, for each set of risk parameters (*𝛽, 𝛼*) ∈

{(0*,* −)*,* (0*.*5*,* 0*.*9)*,* (1*,* 0*.*9)} there is only one distinct solution, inde- pendently of the sample size for the optimization. However, *𝛽* = 1*,*

*𝛼* = (*𝑁* − 1)∕*𝑁* has three distinct solutions, which are identified by the three distinct values of *𝐿𝐵𝑁* ′*,𝑇* ′ *,* see [Table 7](#_bookmark44).

For Case 2 with *𝛽* = 0*,* the same first-stage solution is obtained in-

dependently of the sample size for optimization, while for the other set

of risk parameters, within each set there are slight variations on the power bought through contracts. The slight variations in this variable lead to small variations in the lower bound and confidence interval for the point estimate of the lower bound on the true optimal objective function value.

These results explain why the confidence interval on the point esti- mate for the lower bounds in [Figs. 5](#_bookmark41) and [6](#_bookmark42) are represented by (almost) horizontal dashed lines.

* 1. *Gap between 𝑤*∗ *and* 𝔼*𝜉* [*𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*)]

the upper bound on the gap between *𝑤*∗ and 𝔼*𝜉 𝜙*(*𝑥̂𝑁 , 𝑦, 𝑧̂𝑁 , 𝜉*) *,* as de- We further evaluate each distinct solution using a point estimate on

[ ]

scribed in [Section 4.4](#_bookmark23). [Tables 9](#_bookmark48) and [10](#_bookmark49) show this point estimate for

*𝑁* = 500 and *𝑁* = 5000*,* for Cases 1 and 2, respectively.

For Case 1, a gap of zero is obtained for the solutions with *𝛽* = 0 and

*𝛽* = 0*.*5*,* independently of *𝑁,* and for *𝑁* = 5000 for *𝛽* = 1*, 𝛼* = 0*.*9. For the

other risk parameters, the maximum relative gap is 1.5%.

For Case 2, the solution obtained for *𝛽* = 0 has a gap of zero for both values of *𝑁* . The remaining gaps are relatively small, as the maximum

relative gap is 1.1%.

We also present the wall-clock time required, which consists of the time for the optimization and lower bound estimation. It is clear that the

formulations with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* are more demanding, in partic-

ular for *𝑁* = 5000. The limiting step in this procedure is the optimization stage, which in Case 1 with *𝛽* = 1*, 𝛼* = (*𝑁* − 1)∕*𝑁* and in Case 2 is much

[Figs. 3](#_bookmark38)a, b and [4](#_bookmark39) a, b. For example, in Case 2 with *𝑁* = 500*,* the optimiza- more demanding than the lower bound estimation, as it can be noticed in tion stage requires on average 1216 s and 20,998 s for *𝛼 <* (*𝑁* − 1)∕*𝑁* and *𝛼* = (*𝑁* − 1)∕*𝑁,* respectively, while the lower bound estimation on

average only requires 73 s.

For the assessment of multiple solutions of the same problem, [Procedure 2](#_bookmark24) can be improved by noting that the optimization stage is independent of the solution to assess. Therefore, to assess multiple so- lutions, it is necessary one optimization stage (valid for all solutions) plus one lower bound estimation stage per solution. In this way, and given the computing times for the optimization stage mentioned in the previous paragraph, the computing time required by [Procedure 2](#_bookmark24) could be significantly reduced.

Conceptually, [Procedure 1](#_bookmark25) is oriented to the generation of solutions and specific inference statistics, whereas the objective of [Procedure 2](#_bookmark24) is to assess the solution quality of a given solution. The optimization stage is similar in both procedures, however, the lower bound estimation in [Procedure 1](#_bookmark25) is made for each new solution, whereas in [Procedure 2](#_bookmark24) it is only for the single input solution of the procedure. The common parts between the procedures suggest that they can be merged to take ad- vantage of these similarities to reduce the overall computing time. A merging would require [Procedure 1](#_bookmark25) to be adapted so that during the

**Table 8**

Case 2. First-stage solutions and point estimates of the lower bound on the true optimal objective function value (*𝐿𝐵𝑁*′ *,𝑇* ′ ) with confidence intervals. *𝑁* ′ = 25*,* 000*,* {*𝑀, 𝑇 , 𝑇* ′} = 30.

First-stage solutions (aggregated) CI, [LLB,ULB]

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝛽* | *𝛼* | *𝑁* | UT (%) | SUP | SD | SELLC (MW) | BUYC (MW) | *𝐿𝐵𝑁*′ *,𝑇* ′ (€) | LLB (€) | ULB (€) |  |
| 0 | – | 10 | 100 | 1 | 0 | 0 | 160 | 418,652 | 417,656 | 419,648 |  |
| 0 | – | 50 | 100 | 1 | 0 | 0 | 160 | 418,652 | 417,656 | 419,648 |  |
| 0 | – | 100 | 100 | 1 | 0 | 0 | 160 | 418,652 | 417,656 | 419,648 |  |
| 0 | – | 500 | 100 | 1 | 0 | 0 | 160 | 418,652 | 417,656 | 419,648 |  |
| 0 | – | 5000 | 100 | 1 | 0 | 0 | 160 | 418,652 | 417,656 | 419,648 |  |
| 0.5 | 0.9 | 10 | 100 | 1 | 0 | 100 | 50.89 | 260,050 | 259,967 | 260,132 |  |
| 0.5 | 0.9 | 50 | 100 | 1 | 0 | 100 | 50.81 | 260,048 | 259,966 | 260,130 |  |
| 0.5 | 0.9 | 100 | 100 | 1 | 0 | 100 | 51.15 | 260,055 | 259,971 | 260,138 |  |
| 0.5 | 0.9 | 500 | 100 | 1 | 0 | 100 | 52.00 | 260,061 | 259,975 | 260,146 |  |
| 0.5 | 0.9 | 5000 | 100 | 1 | 0 | 100 | 50.99 | 260,052 | 259,969 | 260,134 |  |
| 1 | 0.9 | 10 | 100 | 1 | 0 | 100 | 38.24 | 239,852 | 239,795 | 239,910 |  |
| 1 | 0.9 | 50 | 100 | 1 | 0 | 100 | 38.68 | 239,869 | 239,811 | 239,927 |  |
| 1 | 0.9 | 100 | 100 | 1 | 0 | 100 | 38.84 | 239,865 | 239,807 | 239,923 |  |
| 1 | 0.9 | 500 | 100 | 1 | 0 | 100 | 38.64 | 239,869 | 239,811 | 239,927 |  |
| 1 | 0.9 | 5000 | 100 | 1 | 0 | 100 | 38.64 | 239,869 | 239,811 | 239,927 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 10 | 100 | 1 | 0 | 100 | 38.22 | 217,371 | 216,616 | 218,127 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 50 | 100 | 1 | 0 | 100 | 38.31 | 217,361 | 216,610 | 218,113 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 100 | 1 | 0 | 100 | 38.00 | 217,389 | 216,617 | 218,161 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 500 | 100 | 1 | 0 | 100 | 38.04 | 217,387 | 216,619 | 218,156 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 5000 | 100 | 1 | 0 | 100 | 38.05 | 217,387 | 216,619 | 218,155 |  |

CI – Confidence interval, LLB and ULB as in [(20)](#_bookmark22), UT – percentage of up-time of the thermal unit, SUP/SD – number of startups/shutdowns of the thermal unit, SELLC/BUYC – power sold/bought through contracts.

**Table 9**

Case 1. Distinct first-stage solutions and the upper bound on the absolute gap (*𝑈𝐺𝑁,𝑀* ) obtained from [Procedure 2](#_bookmark24).

*𝑀* = 30.

First-stage variables (aggregated) *𝑁* = 500 *𝑁* = 5000

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝛽* | *𝛼* | UT (%) | SUP | SD | SELLC (MW) | BUYC (MW) |  | *𝑈𝐺𝑁,𝑀* | Time (s) |  | *𝑈𝐺𝑁,𝑀* | Time (s) |  |
| 0 | – | 100 | 0 | 0 | 0 | 160 |  | 0 | 254 |  | 0 | 3304 |  |
| 0.5 | 0.9 | 100 | 0 | 0 | 155 | 0 |  | 0 | 277 |  | 0 | 3312 |  |
| 1 | 0.9 | 100 | 0 | 0 | 315 | 0 |  | 563 | 268 |  | 0 | 3312 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 0 | 0 | 315 | 0 |  | 4921 | 1665 |  | 10,618 | 93,279 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 95.2 | 1 | 1 | 315 | 0 |  | 27,188 | 1680 |  | 16,127 | 92,788 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 95.2 | 1 | 1 | 315 | 0 |  | 27,977 | 1667 |  | 16,646 | 92,012 |  |

UT – percentage of up-time of the thermal unit, SUP/SD – number of startups/shutdowns of the thermal unit, SELLC/BUYC – power sold/bought through contracts.

**Table 10**

Case 2. Distinct first-stage solutions and the upper bound on the absolute gap (*𝑈𝐺𝑁,𝑀* ) obtained from [Procedure 2](#_bookmark24).

*𝑀* = 30.

First-stage variables (aggregated) *𝑁* = 500 *𝑁* = 5000

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *𝛽* | *𝛼* | UT (%) | SUP | SD | SELLC (MW) | BUYC (MW) |  | *𝑈𝐺𝑁,𝑀* | Time (s) |  | *𝑈𝐺𝑁,𝑀* | Time (s) |  |
| 0 | – | 100 | 1 | 0 | 0 | 160 |  | 0 | 784 |  | 0 | 7488 |  |
| 0.5 | 0.9 | 100 | 1 | 0 | 100 | 50.89 |  | 160 | 1341 |  | 43 | 15,587 |  |
| 0.5 | 0.9 | 100 | 1 | 0 | 100 | 50.81 |  | 160 | 1381 |  | 45 | 15,033 |  |
| 0.5 | 0.9 | 100 | 1 | 0 | 100 | 51.15 |  | 160 | 1365 |  | 35 | 15,165 |  |
| 0.5 | 0.9 | 100 | 1 | 0 | 100 | 52.00 |  | 182 | 1447 |  | 24 | 15,595 |  |
| 0.5 | 0.9 | 100 | 1 | 0 | 100 | 50.99 |  | 160 | 1444 |  | 40 | 15,332 |  |
| 1 | 0.9 | 100 | 1 | 0 | 100 | 38.24 |  | 79 | 1368 |  | 35 | 13,320 |  |
| 1 | 0.9 | 100 | 1 | 0 | 100 | 38.68 |  | 56 | 1333 |  | 13 | 13,286 |  |
| 1 | 0.9 | 100 | 1 | 0 | 100 | 38.84 |  | 64 | 1306 |  | 18 | 13,543 |  |
| 1 | 0.9 | 100 | 1 | 0 | 100 | 38.64 |  | 55 | 1225 |  | 13 | 13,075 |  |
| 1 | 0.9 | 100 | 1 | 0 | 100 | 38.64 |  | 55 | 1198 |  | 13 | 13,000 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 1 | 0 | 100 | 38.22 |  | 2412 | 20,984 |  | 1365 | 82,972 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 1 | 0 | 100 | 38.31 |  | 2434 | 21,362 |  | 1351 | 84,861 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 1 | 0 | 100 | 38.00 |  | 2371 | 21,144 |  | 1422 | 81,950 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 1 | 0 | 100 | 38.04 |  | 2377 | 20,910 |  | 1407 | 82,132 |  |
| 1 | (*𝑁* − 1)∕*𝑁* | 100 | 1 | 0 | 100 | 38.05 |  | 2379 | 20,945 |  | 1403 | 82,635 |  |

UT – percentage of up-time of the thermal unit, SUP/SD – number of startups/shutdowns of the thermal unit, SELLC/BUYC – power sold/bought through contracts.

lower bound estimation, the required information can be extracted to evaluate the upper bound on the gap defined in [(22)](#_bookmark18).

# Conclusions

We presented and applied a methodology based on SAA to generate first-stage solutions and inference statistics for the optimal operation of a VPP. To cope with the computational complexity of the problems ad- dressed, new strategies were developed to reduce the computing time of the optimization replications and innovations on managing the repe- tition of sampling, optimization, and evaluation of first-stage solutions. We performed a detailed characterization of the overall solution by providing point estimates and confidence intervals for the main quanti- ties of interest, including a) upper and lower bounds for the true optimal objective function value; and b) an upper bound on the gap between the true optimal objective function value and the optimal objective func- tion value for a given first-stage solution. This methodology is applied to formulations involving a parameterized combination of the expected profit and the CVaR of the profit. The results and discussions focus on the extremes of the parameterization – the risk-neutral and risk-averse

solutions.

For specific conditions, we identified a relevant trade-off between the size of the sample used in the optimization and the time spent on the lower bound estimation; specifically, increasing the size of the sample for the optimization reduces the time spent in the lower bound estima- tion and the overall required time. Two reasons justify this behavior:

1) the lower bound estimation is only performed for new first-stage so- lutions; and 2) increasing the sample size for the optimization reduces the number of distinct first-stage solutions. There is one exception to this behavior, which occurs for specific instances of the maximization

of the CVaR with *𝛼* = (*𝑁* − 1)∕*𝑁,* where each optimization replication

provides one distinct first-stage solution.

For the present setup, computational experiments indicated that with the number of replications performed, the five sample sizes tested con- verge to the same solution or a neighboring solution. These results in- dicate that the replications with samples of ten elements can generate first-stage solutions that are the same or close to the best solutions gen- erated from samples of 500 elements. However, the variance of the solu- tions obtained with samples of 10 elements makes it diﬃcult to identify whether the best solution has been obtained.

The inference statistics indicate that some of the optimal first-stage solutions of the SAA problem are relatively close to the true optimal solution.

Although the definition of the sample size and the number of replica- tions that guarantee specific bounds or convergence is problem depen- dent, some practical insights are in order. As general guidelines, given limited time and computational resources, the first issue to address is the sample size for the optimization. On the one hand, if the optimiza- tion replications are relatively eﬃcient, then larger samples should be used to reduce the variance of the solutions, eventually avoiding the lower bound estimation for repeated solutions. On the other hand, more diﬃcult optimization problems may require smaller samples for the op- timization, which increase the variance. However, our results indicate that among the solutions found, there is a high probability of identifying an excellent first-stage solution that can be further evaluated with the lower bound estimation.

Future work can evolve in two directions: 1) a comparison between risk-averse stochastic programming and robust optimization; and 2) the study of alternative risk metrics. Regarding the comparison, the first- stage solutions and point estimates of the expected profit and the CVaR of the profit obtained in this work can be contrasted with the solutions from a robust optimization approach. Additionally, it would be rele- vant to explore alternative risk metrics, such as stochastic dominance ([Escudero and Monge, 2018](#_bookmark63)), and to assess the quality of the solu- tions and the overall SAA methodology performance with alternative risk metrics.

# Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

# Acknowledgments

Research reported in this publication was supported by research [funding from King Abdullah University of Science and Technology (KAUST). Antonio J. Conejo’s contribution is partly supported by NSF](https://doi.org/10.13039/501100004052) project 1808169. We would like to thank two referees for their com- ments and contributions that helped to improve the presentation of this work.

# Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.ejco.2021.100005](https://doi.org/10.1016/j.ejco.2021.100005).

# References

[Alonso-Ayuso, A., Escudero, L.F., Guignard, M., Weintraub, A., 2018. Risk management for forestry planning under uncertainty in demand and prices. Eur. J. Oper. Res. 267, 1051–1074.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0001)

[Artzner,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0002) [P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Math. Fi- nance 9, 203–228.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0002)

[, 1997. In: Awerbuch, S., Preston, A. (Eds.), The Virtual Utility: Accounting, Technology & Competitive Aspects of the Emerging Industry. Springer US.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0003)

[Baringo,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0004) [A., Baringo, L., 2017. A stochastic adaptive robust optimization approach for the offering strategy of a virtual power plant. IEEE Trans. Power Syst. 32, 3492–3504.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0004)

[Bayraksan,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0005) [G., Morton, D.P., 2006. Assessing solution quality in stochastic programs.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0005)

[Math. Program. 108, 495–514.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0005) [Ser. B](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0005)

Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4, 238–252. doi:[10.1007/BF01386316](https://doi.org/10.1007/BF01386316).

[Birge, J.R.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0007), [Louveaux, F., 2011. Introduction to Stochastic Programming. Springer](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0007). [Conejo,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0008) [A.J., Carrión, M., Morales, J.M., 2010. Decision Making Under Uncertainty in](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0008)

[Electricity Markets. Springer US.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0008)

[Dabbagh,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0009) [S.R., Sheikh-El-Eslami, M.K., 2016. Risk assessment of virtual power plants of- fering in energy and reserve markets. IEEE Trans. Power Syst. 31, 3572–3582.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0009)

[Dupačová, J., Gröwe-Kuska, N., Römisch, W., 2003. Scenario reduction in stochastic pro- gramming – an approach using probability metrics. Math. Program. 95, 493–511.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0010)

[Escudero,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0011) [L.F., Monge, J.F., 2018. On capacity expansion planning under strategic and operational uncertainties based on stochastic dominance risk averse management. Comput. Manag. Sci. 15, 479–500.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0011)

Iberian Electricity Market, 2015. [www.omie.es](http://www.omie.es/).

Jansen, J.C., van der Welle, A., Nieuwenhout, F., 2008. Deliverable D3.2.4: The Virtual Power Plant Concept From an Economic perspective: Updated Final Report. Technical Report. Energy Research Centre of the Netherlands (ECN). [http://www.fenix-project.org](http://www.fenix-project.org/)

[Kall, P.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0014), [Mayer, J., 2011. Stochastic Linear Programming. Springer US](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0014).

[Kardakos,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0015) [E.G., Simoglou, C.K., Bakirtzis, A.G., 2016. Optimal offering strategy of a virtual power plant: a stochastic bi-level approach. IEEE Trans. Smart Grid 7, 794–806.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0015)

[Kleywegt,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0016) [A., Shapiro, A., Homem-De-Mello, T., 2001. The sample average approximation method for stochastic discrete optimization. SIAM J. Optim. 12, 479–502.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0016)

[Kozmík,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0017) [V., Morton, D.P., 2015. Evaluating policies in risk-averse multi-stage stochastic programming. Math. Program. 152, 275–300.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0017)

[Lima,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0018) [R.M., Conejo, A.J., Langodan, S., Hoteit, I., Knio, O.M., 2018. Risk-averse formula- tions and methods for a virtual power plant. Comput. Oper. Res. 96, 350–373.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0018)

[Lima,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0019) [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, 457–475.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0019)

[Linderoth,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0020) [J., Shapiro, A., Wright, S., 2006. The empirical behavior of sampling methods for stochastic programming. Ann. Oper. Res. 142, 215–241.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0020)

[Maître,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0021) [O.L., Knio, O.M., 2010. Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics. Springer Science & Business Media.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0021)

[Mak,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0022) [W., Morton, D., Wood, R., 1999. Monte carlo bounding techniques for determining solution quality in stochastic programs. Oper. Res. Lett. 24, 47–56.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0022)

[Homem-de](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0023) [Mello, T., Bayraksan, G., 2014. Monte Carlo sampling-based methods for stochastic optimization. Surv. Oper. Res. Manag. Sci. 19, 56–85.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0023)

[Mello,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0024) [T.H.-d., Pagnoncelli, B.K., 2016. Risk aversion in multistage stochastic program- ming: a modeling and algorithmic perspective. Eur. J. Oper. Res. 249, 188–199.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0024)

[Moazeni,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0025) [S., Powell, W.B., Hajimiragha, A.H., 2015. Mean-conditional value-at-risk opti- mal energy storage operation in the presence of transaction costs. IEEE Trans. Power Syst. 30, 1222–1232.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0025)

[Moghaddam,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0026) [I.G., Nick, M., Fallahi, F., Sanei, M., Mortazavi, S., 2013. Risk-averse prof- it-based optimal operation strategy of a combined wind farm-cascade hydro system in an electricity market. Renew. Energy 55, 252–259.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0026)

[Morales,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0027) [J.M., Conejo, A.J., Madsen, H., Pinson, P., Zugno, M., 2014. Integrating Renew- ables in Electricity Markets. Springer US.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0027)

[Pandžić, H., Kuzle, I., Capuder, T., 2013. Virtual power plant mid-term dispatch optimiza- tion. Appl. Energy 101, 134–141.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0028)

[Pandžić, H., Morales, J.M., Conejo, A.J., Kuzle, I., 2013. Offering model for a virtual power plant based on stochastic programming. Appl. Energy 105, 282–292.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0029)

[Pereira, M., Pinto, L., 1991. Multistage stochastic optimization applied to energy planning.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0030)

[Math. Program. 52, 359–375.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0030)

[Pflug, C., 2000. Probabilistic Constrained Optimization: Methodology and Applications.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0031)

[Springer US, pp. 272–281.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0031)

[Pflug, G., Ruszczynski, A., 2005. Measuring risk for income streams. Comput. Optim. Appl.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0032)

[32, 161–178.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0032)

[Pflug, G.C., Römisch, W., 2007. Modeling, Measuring and Managing Risk. World Scientific Publishing Co. Pte. Ltd..](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0033)

[Philpott, A.B., de Matos, V.L., 2012. Dynamic sampling algorithms for multi-stage stochas- tic programs with risk aversion. Eur. J. Oper. Res. 218, 470–483.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0034)

[Pudjianto, D., Ramsay, C., Strbac, G., 2007. Virtual power plant and system integration of distributed energy resources. IET Renew. Power Gen. 1, 10–16.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0035)

[Rahimiyan, M., Baringo, L., 2016. Strategic bidding for a virtual power plant in the day-a- head and real-time markets: a price-taker robust optimization approach. IEEE Trans. Power Syst. 31, 2676–2687.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0036)

[Rahmani-Dabbagh, S., Sheikh-El-Eslami, M.K., 2016. A profit sharing scheme for dis- tributed energy resources integrated into a virtual power plant. Appl. Energy 184, 313–328.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0037)

[Rockafellar, R.T., 2007. Coherent approaches to risk in optimization under uncertainty.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0038)

[INFORMS 38–61.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0038)

[Rockafellar, R.T., Uryasev, S., 2000. Optimization of conditional value-at-risk. J. Risk 2, 21–41.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0039)

[Sarykalin, S., Serraino, G., Uryasev, S., 2008. Value-at-risk vs. conditional value-at-risk in risk management and optimization. INFORMS 270–294.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0040)

[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. Energy 155, 766–777.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0041)

[Shabanzadeh, M., Sheikh-El-Eslami, M.-K., Haghifam, M.R., 2016. A medium-term coali- tion-forming model of heterogeneous DERs for a commercial virtual power plant. Appl. Energy 169, 663–681.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0042)

[Shang,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0043) [D., Kuzmenko, V., Uryasev, S., 2018. Cash flow matching with risks controlled by buffered probability of exceedance and conditional value-at-risk. Ann. Oper. Res. 260, 501–514.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0043)

[Shapiro,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0044) [A., 2009. On a time consistency concept in risk averse multistage stochastic pro- gramming. Oper. Res. Lett. 37, 143–147.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0044)

[Shapiro,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0045) [A., 2011. Analysis of stochastic dual dynamic programming method. Eur. J. Oper.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0045)

[Res. 209, 63–72.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0045)

[Shapiro,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0046) [A., Dentcheva, D., Ruszczyński, A., 2009. Lectures on Stochastic Programming: Modeling and Theory. In: MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, pp. 253–332.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0046)

[Shapiro,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0047) [A., Homem-De-Mello, T., 2000. On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. SIAM J. Optim. 11, 70–86.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0047)

[Shapiro,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0048) [A., Tekaya, W., Costa, J.P.d., Soares, M.P., 2013. Risk neutral and risk averse stochastic dual dynamic programming method. Eur. J. Oper. Res. 224, 375–391.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0048)

[Shayegan-Rad, A., Badri, A., Zangeneh, A., 2017. Day-ahead scheduling of virtual power plant in joint energy and regulation reserve markets under uncertainties. Energy 121, 114–125.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0049)

[Tajeddini,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0050) [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.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0050)

[Van](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0051) [Slyke, R., Wets, R., 1969. L-shaped linear programs with applications to optimal con- trol and stochastic programming. SIAM J. Appl. Math. 17, 638–663.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0051)

[Verweij,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0052) [B., Ahmed, S., Kleywegt, A.J., Nemhauser, G., Shapiro, A., 2003. The sample av- erage approximation method applied to stochastic routing problems: a computational study. Comput. Optim. Appl. 24, 289–333.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0052)

[Wang,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0053) [W., 2007. Sample Average Approximation of Risk-Averse Stochastic Programs. H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology Ph.D. thesis.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0053)

[Wang,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0054) [W., Ahmed, S., 2008. Sample average approximation of expected value constrained stochastic programs. Oper. Res. Lett. 36, 515–519.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0054)

[Zamani,](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0055) [A.G., Zakariazadeh, A., Jadid, S., 2016. Day-ahead resource scheduling of a re- newable energy based virtual power plant. Appl. Energy 169, 324–340.](http://refhub.elsevier.com/S2192-4406(21)00002-2/sbref0055)