Sample average approximation and stability tests applied to energy system design

Pernille Seljom*[a] (pernille.seljom@ife.no) and Asgeir Tomasgard[b] (asgeir.tomasgard@iot.ntnu.no)

a) Department of Renewable Energy Systems
   Institute for Energy Technology (IFE)
   Post box 40
   2027 Kjeller
   Norway
b) Department of Industrial Economics and Technology Management
   Norwegian University of Science and Technology (NTNU)
   7491 Trondheim
   Norway

* Corresponding author + 47 99 02 02 63

Abstract

This paper uses confidence intervals from sample average approximation (SAA) and stability tests to evaluate the quality of the solution of a long-term energy system model with stochastic wind power production. Using poorly designed scenarios can give stochastic model results that depend on the scenario representation rather than the actual underlying uncertainty. Nevertheless, there is little focus on the quality of the solutions of stochastic energy models in the applied literature. Our results demonstrate how too small a sample size can give a poor energy system design and misrepresent the value of the stochastic solution (VSS).

We demonstrate how to evaluate the number of scenarios needed to ensure in-sample and out-of- sample stability. We also show how replication and testing of many candidate solutions using SAA iterations can provide a solution with a satisfactory confidence interval, including when the samples contain fewer scenarios than required for stability. An important observation, though, is that if SAA repeatedly solves the model with a sample size that satisfies in-sample and out-of-sample stability, the confidence interval is narrow, and the solutions are of high quality in terms of providing a tight bound for the optimal solution.

1. Introduction

The Integrated MARKAL EFOM System (TIMES) is a widely used optimisation framework for long-term energy system planning [1-5]. With a higher share of intermittent energy generation from renewables, the focus on how to handle short-term uncertainty in such models is increased [6-8]. Stochastic programming [9] is a mathematical framework that enables us to consider uncertainty and to value flexibility in optimisation models. A stochastic approach to incorporating short-term uncertainty in TIMES was first introduced in [10] and is demonstrated in a user guide in [11]. From an application perspective, this TIMES approach was first used in [12] to take into account the short-term uncertainty related to the intermittent wind power production for Denmark. Since then, the approach has also been used in two other TIMES analyses: a study on the impact of zero energy buildings in the Scandinavian energy system [13] and a study on the impact of policy actions and energy prices on the cost-optimal development of the energy system in Norway and Sweden [14].

The stochastic programming approach provides cost-optimal investment decisions that explicitly consider different outcomes of short-term uncertainty. Nevertheless, most of the TIMES literature, see for example [15-21], does not consider uncertainty but uses a deterministic approach in which investments are based on one operational scenario only. Clearly, a deterministic modelling of short-term uncertainty in long-term energy system models can provide sub-optimal investments, as demonstrated in [12, 22-24], amongst others. However, a major challenge of introducing stochasticity in long-term energy models, like TIMES models, is that the computational effort increases with the number of discrete scenarios used to represent the uncertain parameters.

Most of the stochastic models in the energy systems literature include little or no discussion of the quality of the solution and the number of scenarios used. This is also the case for studies using long-term energy models, such as [13, 14, 22-25], covering the energy sector or sub-systems, such as the electricity sector. This paper uses confidence intervals from sample average approximation (SAA) and stability tests to evaluate the quality of the solution of a long-term energy system model with stochastic wind power production. Stochastic programming provides solutions that value flexibility by using scenarios to represent possible realisations of the uncertainty. The literature on long-term energy models often uses either a deterministic model or a stochastic model with a small number of scenarios. For both stochastic models and deterministic models in applied energy system analysis, there is rarely a discussion about the quality of the solution they provide, or, in fact, what defines a good quality solution.

A premise of our work is that the real world is uncertain and that a true probability distribution representing this real-world uncertainty exists. The quality of a model solution should then ideally be assessed in terms of how it deviates from the optimal solution in the real world. Obviously, a deterministic model, ignoring stochasticity, may lead to arbitrarily poor analyses. It is less reflected on that analyses based on stochastic models may suffer from the same problem. This is because the representation of uncertainty in these models is an approximation (often in terms of a limited number of discrete scenarios) of the true distribution (which can be continuous). If a model uses a poor representation of the true distribution, the results from the stochastic model can depend on the scenario representation rather than the real-world uncertainty.

In this paper, we assume that the true probability distribution can be represented by S discrete scenarios. Our results demonstrate how using too few scenarios to approximate the true distribution can give a poor energy system design and misrepresent the value of the stochastic solution (VSS). One of the reasons that stochastic models often use too small a scenario tree is that the computational efforts increase with the number of scenarios. One method developed to overcome this is the sample average approximation (SAA) method, which uses random sampling from the true distribution, and replication over several iterations, to identify candidate solutions.

The main contribution of this paper is to illustrate the errors that result from an inadequate representation of uncertainty in terms of objective function value, model solutions and estimate of the VSS. In this process, we demonstrate the use of confidence intervals for the optimality gap provided by the SAA [26, 27] algorithm and stability tests [28] to evaluate the quality of the solution provided by a TIMES model of Denmark with stochastic wind power production. The purpose of the paper is to illustrate the typical errors a misrepresentation of uncertainty can lead to, and to discuss how these can be eliminated.

Below, we give a brief methodological overview of methods that are used to generate scenarios and to evaluate the quality of the model solution. Thereafter, we give an overview of energy models that use these methods.

There are several methods to generate a limited number of scenarios from either a specified continuous distribution or a large data set that describes the uncertain parameters. These include random sampling [29], moment matching [30-32] and scenario reduction by distance measures [33-38]. The purpose of using a statistical approach for scenario generation, such as moment matching and distance measures, rather than random sampling is to reduce the number of scenarios required to provide a good quality solution. However, the statistical approaches are more complex, as they use mathematical algorithms to generate scenarios.

Moment matching generates a pre-specified number of scenarios that match for example the first four statistical moments (mean, variance, skewness and kurtosis) of an underlying data set that describe the uncertainty. Consequently, this method presupposes that these moments give an adequate representation of the uncertain parameters. The moments can, however, be insufficient to describe all the characteristics of the uncertain parameters, as completely different distributions can have the same first four moments [39]. The advantage, though, is that the number of scenarios required to represent the first four moments is, by construction, generally lower when using moment matching than when using randomly sampled scenarios.

Methods based on scenario reduction using distance measures take a different approach. They start with a high number of scenarios, created for example by sampling, where the large scenario set represents the true distribution of the uncertainty. Next, the number of scenarios is reduced without distorting these properties. In general, distance-based methods reduce the number of scenarios to a pre-specified size while minimising the distance between the final scenarios and the original distribution (see for example [33, 34, 38]). Distance measures can also be used to find the optimal discretisation of a continuous distribution [40].

The energy literature applies a variety of scenario generation methods. Random sampling is applied to a unit commitment problem in [41], to a power generation expansion in [42] and to an energy system model in [13, 14]. In the latter studies, the sampled scenarios are adjusted to the mean value of the uncertainty. Scenario reduction by distance measures are used in power generation expansion problems in [23, 24, 43, 44], in maintenance scheduling in [45] and in unit commitment problems in [46, 47]. Moment matching is used in a power generation expansion problem in [48], and a combined moment matching and sampling methodology is applied in an energy system model of Denmark in [12] and in an energy system of buildings in [49]. There is also literature that generates scenarios in a different manner and/or does not describe the applied scenario generation method. For example, for power expansion models, [22] pick ten scenarios from the distribution of wind and photovoltaic power (PV) production according to occurrence, and [50] uses the hydro inflow of 75 different climate years to cover outcomes of the uncertainty.

The quality of the scenarios can only be measured by the quality of the resulting solution. In this paper we define this quality based on how close the model optimal solution value is to the optimal solution value for the true distribution defined by S discrete scenarios. Most of the stochastic energy literature lacks a discussion on the quality of the presented model results. This is a methodological weakness since poor scenario representations can give results that depend on the scenario representation rather than the characteristics of the uncertain parameters. There are, however, studies that acknowledge this weakness. The authors in [23] discuss the shortcomings of using few scenarios, including that they do not cover the total range of fluctuations in renewable resources generation, and [13, 14] state that using a higher number of scenarios can improve the quality of the model results.

To evaluate solution quality, stability tests on in-sample and out-of-sample stability are useful [28]. These stability tests indicate whether the optimal solution value is skewed by a misrepresentation of the true underlying stochasticity. Stability tests are used only in a small share of the applied stochastic energy literature. For example, [12] applies both in-sample and out-of-sample stability tests in a Danish TIMES model with uncertain wind power production. Further, there are studies that use in-sample stability only, like [51] with a model for optimal bidding and scheduling in the electricity market and [52] with a stochastic unit commitment model. Although it is not explicitly mentioned, a variant of in-sample stability is also tested in numerous studies that evaluate the model solution value as a function of sample size. This is done in a power expansion problem in [44, 48] and in a bi-level investment problem for a wind power producer in [53]. Stability tests are also applied in other types of models, including supply chain design [54], portfolio selection model [55] and network design [56].

Other solution evaluation methods are directly linked to the scenario generation method. These include the sample average approximation (SAA) method, which uses randomly sampled scenarios. SAA quantifies the quality of a model solution using replication, by repeatedly solving the model using samples with a limited number of scenarios. SAA thereby provides a confidence interval of the optimality gap of a given model solution. SAA is used in a few energy models, including the unit commitment problems in [57, 58], with uncertain wind power production, and a power expansion model in [42], with uncertain natural gas price and electricity demand. The use of SAA is however more prevalent in other types of models, including supply chain design [59, 60], routing problems [61], inventory problems [62] and empty container repositioning [63].

The optimality gap of clustered scenarios [43] is another solution evaluation method that is linked to distance measures. With this method, the value of a stochastic model is the lower bound of the optimal value, whereas the upper bound of the optimality gap is derived by fixing the first-stage solution to a large scenario tree. However, the lower bound is only feasible for a model with a fixed recourse, implying, for example, that it cannot be used in stochastic models where the uncertainty is related to energy supply efficiency. For energy models, the use of optimality gap of clustered scenarios is limited to [43], a power expansion model covering the west of the Unites States with uncertain hydropower production, PV production, wind power production and electricity demand.

The contribution of this paper is the use of both SAA optimality gap estimates and stability tests to evaluate the quality of the model's solution. We apply SAA to a TIMES model with three different sample sizes to examine how the number of scenarios affects the quality of the investment decisions. To our knowledge, this is the first study that compares and relates stability tests and confidence intervals from SAA to evaluate the quality of the model solution in an energy model. It is also the first time the SAA method has been used in a long-term energy system model with multiple end-use sectors, such as TIMES energy system models.

The remainder of this paper is structured as follows: Section 2 gives a brief introduction of the stochastic methods used in this paper. Section 3 is devoted to a description of the model and the scenario generation methodology. Section 4 shows the computational results of SAA and stability tests and Section 5 details the conclusions.

Abbreviations

bEUR, billion euro; CHP, combined heat and power; DH, district heat; EV, expected value; PV, photovoltaic power; SAA, sample average approximation; TIMES, the Integrated MARKAL-EFOM System; VSS, the value of stochastic solution

2. Stochastic methods

This section describes the stochastic methods that are used in this paper.

2.1 Two-stage model

SAA is primarily used as a solution approach for two-stage stochastic programmes with recourse [9], although it is possible to extend the methodology to a multi-stage setting, as demonstrated by [64].

Here, the true uncertainty distribution is represented by S discrete scenarios that occur with equal probability. The objective function of a two-stage model is shown in Eq.(1). The first-stage variables, x, are independent of the scenarios and the second stage variables, y_s, are scenario dependent. Similarly, the first stage cost parameters, c, are deterministic, while the second-stage costs, d_s, are scenario dependent. In addition, the optimisation considers both scenario independent constraints, $f(x)$, and scenario dependent constraints, $g_s(x, y_s)$. Further, the true optimal value, using the S discrete scenarios, is denoted, v^*, with a corresponding optimal first-stage solution, x^*, and we refer to Eq.(1) as the true distribution problem.

$$
\begin{aligned}
&\min v = c^{\mathrm{T}} x + \frac{1}{S}\sum_{s=1}^{S} d_s^{\mathrm{T}} y_s \\
&s.t \\
&\qquad f(x) \le 0 \\
&\qquad g_s(\mathrm{x}, y_s) \le 0 \qquad\qquad s = 1, 2, ..., S
\end{aligned}
\tag{1}
$$

In TIMES models, both the constraints and the objective are typically linear, although this is not always the case for energy system models. Hence, in what follows, when discussing bounds, we assume that (i) the objective has finite mean and variance, (ii) i.i.d. samples of scenarios can be generated and (iii) the objective can be evaluated exactly for specific values of x, y and realizations of the stochastic parameters.

2.2 The value of stochastic solution

A deterministic model is equivalent to a stochastic model with one scenario, and a model using the expected value of the uncertain parameters as a model input, is denoted the expected value (EV) problem. The value of the stochastic solution (VSS) [65, 66] measures the benefit of using a stochastic model compared to a deterministic approach.

The first step in calculating VSS is to find the first-stage decisions from the EV model, denoted $\overline{x}$. Second, the stochastic model is solved with first-stage decisions fixed to $\overline{x}$, taken from the EV model as shown in Eq.(2). Finally, VSS is the difference in optimal objective function value of this stochastic model with fixed first-stage decisions, $\overline{v}$, and the optimal value of the stochastic model with endogenous first-stage decisions, as defined in Eq.(3).

$$\begin{aligned} &\overline{v}_S = c^{\mathrm{T}}\overline{x} + \frac{1}{S}\min\sum_{s=1}^{S} d_s^{\mathrm{T}} y_s \\ &s.t \\ &\qquad f(\overline{x}) \le 0 \\ &\qquad g_s(\overline{\mathrm{x}}, y_s) \le 0 \qquad\qquad s = 1,2,..,S \end{aligned} \tag{2}$$

$$VSS = \overline{v}_S - v^* \tag{3}$$

2.3 SAA algorithm

The SAA algorithm works by repeatedly solving the two-stage model with a limited number of scenarios, N, sampled from the S scenarios of the true distribution. Below we give a stepwise procedure for the SAA algorithm that is based on [26, 67, 68] and the applied literature in [59, 60, 69]. In this paper and in the following notation, the N scenarios are generated by a crude Monte Carlo sampling method. It is, however, also possible to apply variance reduction techniques within the SAA algorithm, as demonstrated in [62, 67].

Lower bound estimate

Below we describe the SAA used iteratively with M replications for a stochastic two-stage model with recourse. The bounds and algorithm are taken from [26, 67, 68]. First, we discuss how to estimate the lower bound for the optimal value of the true problem with S scenarios. The lower bound estimate is found by solving M two-stage problems each with N samples scenarios, where $N< S$. This lower bound estimate, with its variance, is derived as follows:

a. Generate M independent candidate samples each with N scenarios from the S scenarios.

b. Solve the corresponding optimisation problem for all M candidate samples according to Eq.(4). Here, the optimal objective function value of each sample, m, is denoted $v_{m,N}$:

$$
\begin{aligned}
& v_{m,N} = \min \left(c^{\mathrm{T}} x + \frac{1}{N} \sum_{n=1}^{N} d_{m,n}^{\ \ \mathrm{T}} y_{m,n} \right) \\
& s.t \\
& \qquad f(x) \le 0 \\
& \qquad g_n(\mathrm{x}, y_{m,n}) \le 0 \qquad\qquad n = 1, 2 \ldots, N
\end{aligned}
\tag{4}
$$

and the corresponding optimal first-stage candidate solution is represented by $x_{m,N}$.

Calculate the estimated lower bound, $LB_{M,N}$, that is the average of the optimal objective function value of all M samples, see Eq. (5). The estimated lower bound converges to the true lower bound as M goes to infinity:

$$
LB_{M,N} = \frac{1}{M} \sum_{m=1}^{M} v_{m,N} \tag{5}
$$

c. Estimate the variance of the lower bound $\sigma_{LB}^2(M)$ by the variance estimator $s_{LB}^2(M)$, in Eq.(6):

$$
s_{LB}^2(M) = \frac{1}{\mathrm{M}(M-1)} \sum_{m=1}^{M} \left(v_{m,N} - LB_{\mathrm{M,N}} \right)^2 \tag{6}
$$

d. Derive the confidence interval of the lower bound estimate according to Eq.(7), where $z_{\alpha/2}$ is the critical value of the normal distribution function for a given percentile:

$$
\left[LB_{M,N} - z_{\alpha/2} s_{LB}(M), LB_{M,N} + z_{\alpha/2} s_{LB}(M) \right] \tag{7}
$$

Upper bound estimate

An upper bound for a given first-stage solution, $\hat{x}$, is provided by using this first-stage solution in a two-stage model with N' scenarios, called the reference sample. The N' scenarios are derived by random sampling from the S scenarios, and $S \geq N' > N$. Ideally, we wish this reference sample to be the true distribution, but normally this is not possible. Hence, N' is chosen as large as computationally feasible, and significantly larger than N. This is possible because a two-stage model with a fixed first-stage variable, $\hat{x}$, can be solved as N' separate models, such that the computational efforts related to calculating the upper bound are in most cases moderate compared to solving the full stochastic model.

In the following, $\hat{x}$ is one of the candidate solutions obtained, $x_{m,N}$. An estimate of the upper bound, $UB_{N'}(\hat{\mathrm{x}})$, with its variance, is derived as follows:

a. Find the optimal value of the problem with N' scenarios and a fixed first-stage solution, $\hat{x}$, as described in Eq.(8). This is the estimated upper bound, $UB_{N'}(\hat{\mathrm{x}})$, of this first-stage solution, and the estimated upper bound converges to the true upper bound as N' goes to S.

$$
\begin{aligned}
& UB_{N'}(\hat{\mathrm{x}}) = \frac{1}{N'}\sum_{n'=1}^{N'} c^{\mathrm{T}}\,\hat{\mathrm{x}} + \min(d_{n'}{}^{\mathrm{T}} y_{n'}) = \frac{1}{N'}\sum_{n'=1}^{N'} v_{n'}(\hat{\mathrm{x}}) \\
& s.t \\
& \qquad f(\hat{\mathrm{x}}) \leq 0 \\
& \qquad g_{n'}(\hat{\mathrm{x}}, y_{n'}) \leq 0 \qquad\qquad n' = 1,2,..,N'
\end{aligned}
\tag{8}
$$

b. Apply Eq.(9) to calculate the estimator, $\hat{s}^2_{UB_{N'}}(\hat{\mathrm{x}})$ of the variance of the upper bound $\hat{\sigma}^2_{UB}(\hat{\mathrm{x}})$:

$$
\hat{s}^2_{UB_{N'}}(\hat{\mathrm{x}}) = \frac{1}{\mathrm{N'(N'}-1)}\sum_{n'=1}^{N'}\left(v_{n'}(\hat{\mathrm{x}}) - UB_{N'}(\hat{\mathrm{x}})\right)^2 \tag{9}
$$

This formula assumes that the sample with N' scenarios is independent of the M previous samples. In [26] an alternative bound based on using all the scenarios in the M samples as the reference is given instead.

c. Derive the confidence interval of the upper bound estimate according to Eq.(10), where $z_{\alpha/2}$ is the critical value of the normal distribution function for a given percentile:

$$
\left[UB_{N'}(\hat{\mathrm{x}}) - z_{\alpha/2} s_{UB_{N'}}(\hat{\mathrm{x}}), \mathrm{UB}_{N'}(\hat{\mathrm{x}}) + z_{\alpha/2} s_{UB_{N'}}(\hat{\mathrm{x}})\right] \tag{10}
$$

Optimality gap estimate

Based on the estimated bounds, an estimator of the optimality gap, $GAP_{M,N,N'}(\hat{x})$, and its variance is in accordance to Eq.(11) and Eq.(12) respectively, and the confidence interval of the optimality gap is calculated according to Eq.(13).

$$GAP_{M,N,N'}(\hat{x}) = UB_{N'}(\hat{x}) - LB_{M,N} \tag{11}$$

$$s^2_{GAP_{N,M,N'}}(\hat{x}) = \hat{s}^2_{UB_{N'}}(\hat{x}) + s^2_{LB}(M) \tag{12}$$

$$\left[GAP_{M,N,N'}(\hat{x}) - z_{\alpha/2} s^2_{GAP_{M,N,N'}}(\hat{x}), GAP_{M,N,N'}(\hat{x}) + z_{\alpha/2} s^2_{GAP_{M,N,N'}}(\hat{x})\right] \tag{13}$$

2.4 In-sample and out-of-sample stability

Stability tests are used to evaluate the quality of the scenarios in a stochastic program. If the scenarios result in optimal solutions that satisfy in-sample and out-of-sample stability, the optimal solutions are of good quality [28].

The in-sample stability test [28] investigates whether a scenario generation method creates a set of scenarios that consistently result in similar optimal values. Based on our previous notation, the in-sample stability is defined according to Eq.(14). If M scenario sets are created with N scenarios each, the optimal values, derived by using the different scenario sets, are required to be similar. In addition, it is desirable that the optimal value remain the same when the sample size increases from N to N'. If this is the case, the scenario trees are in-sample stable. Note that the average and standard deviation of in-sample stability values in Eq. 14, for a given sample size N, is identical with the SAA lower bound estimate and standard deviation, and thereby also confidence interval.

$$v_{1,N} \approx v_{2,N} \approx \ldots\ldots\ldots\ldots \approx v_{M,N} \tag{14}$$

The out-of-sample stability test [28] investigates whether a scenario generation method, with the selected sample size, creates scenario trees that provide optimal solutions that give approximately the same optimal value as when using the true probability distribution. One way of checking this is to insert the fixed first-stage solution of each sample *m* with size *N* into an optimization problem that uses the *S* scenarios, representing the true distribution. In some cases, this is too challenging, due to the computational effort or because it is not possible to represent the true distribution. In these situations, an approximation of out-of-sample stability can still be useful. A main purpose of out-of-sample-type tests is to test the solutions on different samples/trees than were used to find them.

If a reference sample, consisting of $N'<S$ scenarios, gives an adequate representation of the underlying uncertainty, out-of-sample stability implies that the estimator for the upper bound, $UB_{N'}(x_{m,N})$, should have approximately the same value for all candidate samples *m*, where the optimal first-stage candidate solution is represented by $x_{m,N}$:

$$UB_{N'}(\mathrm{x}_{1,N}) \approx UB_{N'}(\mathrm{x}_{2,N}) \approx \cdots\cdots\cdots\cdots \approx UB_{N'}(\mathrm{x}_{M,N}) \qquad (15)$$

For a given sample size, N, the average of the values in Eq.(15) and the its standard deviation can be used to derive the confidence interval of the out-of-sample stability. In this paper we use Eq. (15) to derive the out-of-sample stability. An alternative formulation, presented in [26], not using a reference sample based on *N´* scenarios but rather using the *N*M* sampled scenarios, is to compare the upper bound for all k ≠l:

$$UB_k(x_{k,N}) \approx UB_k(x_{l,N}), \quad UB_l(x_{l,N}) \approx UB_l(x_{k,N}) \text{ and finally } UB_k(x_{l,N}) \approx UB_l(x_{k,N}) \quad (16)$$

where $k = 1,\ldots,M \quad l = 1,\ldots,M \quad k \neq l$

In any case, as stated for the upper bound calculations for SAA, a two-stage model with a fixed first-stage variable, $\hat{x}$, can then be solved as separate model for each scenario, making the out-of-sample stability test computationally tractable.

2.5 Link between SAA algorithm and stability tests

Based on the description above, this section further elaborates the link between the SAA algorithm and the stability tests. Assume that the *M* samples with *N* scenarios are used for both the SAA lower bound calculations and in-sample stability tests. The reference sample with *N´* scenarios is used for both the SAA upper bound calculation and the out-of-sample stability test.

Observation 1:

i. The calculation of the SAA lower bound estimate (See Section 2.3) provides *M* different candidate solutions indexed from *1,…, M* and corresponding optimal solution values $v_{m,N}$. These corresponds exactly to the $v_{m,N}$ in the in-sample stability tests in Eq.(14).

ii. The variance estimator of the SAA lower bound, $s_{LB}^2(M)$ from Eq.(6), is identical with the variance for $v_{m,N}$, with *m=1,…,M* in the in-sample stability test.

Observation 2:

i. Each of the first stage solutions $x_{m,N}$ can be used to calculate an upper bound as described in Eq.(8). For sample *m* this corresponds exactly to the $UB_{N'}(x_{m,N})$ used in the out-of-sample stability test in Eq.(15) for a reference scenario set with *N´* scenarios.

ii. There is no immediate link between the out-of-sample stability criteria, requiring that the upper bound estimator $UB_{N'}(x_{m,N})$ for each sample *m=1,…,M* be approximately the same (Eq.(15)), and the SAA upper bound based on the sample variance $\hat{s}_{UB_{N'}}^2(\hat{\mathrm{x}})$ of a single $UB_{N'}(\hat{\mathrm{x}})$.

iii. If the *M* samples with *N* scenarios are out-of-sample stable, the upper bound estimator from SAA, $UB_{N'}(\hat{\mathrm{x}})$, equals the upper bound estimator from the stability test for all *m*. Still, if the sample variance $\hat{s}_{UB_{N'}}^2(\hat{\mathrm{x}})$ is significantly large, this does not guarantee a small enough confidence interval for the optimality gap in SAA according to a user-defined criterion.

iv. Vice versa, if the *M* samples with *N* scenarios are not out-of-sample stable when tested on the *N´* scenario set, the confidence interval for the best candidate solution , $\hat{x}$, can be within a user-defined criterion and only depending on $\hat{s}_{UB_{N'}}^2(\hat{\mathrm{x}})$.

Observation 3:

i. In SAA, the requirement is that the confidence interval for the in-sample stability in the *M* samples corresponds to $s_{LB}^2(M) \le \varepsilon$ with ε as an user-defined criterion, approaching zero. If these *M* sampled scenario sets are used in the SAA, the width of the confidence interval for the lower bound approaches zero.

ii. If the M samples are not in-sample stable, the SAA lower bound may still be acceptable, in terms of a user defined criterion, with the width of the confidence interval depending on $s_{LB}^2(M)$.

3. Model description and scenario generation

First, this section gives an overview of the model structure and assumptions of our TIMES model. Thereafter, we present the stochastic modelling approach, including the generation of the wind scenarios.

3.1 TIMES model structure and assumptions

TIMES is a bottom-up modelling framework, providing a detailed techno-economic description of resources, energy carriers, conversion technologies and energy demand from a social welfare perspective. The model minimises the total discounted cost of the energy system to meet the demand for energy services for the analysed model horizon. The energy system cost includes investments in both supply and demand technologies, expenses related to operation of capacity, fuel costs, income from electricity export and cost of electricity import from external countries.

The model decisions are made with full knowledge of all future events and their probabilities. Each model period is five years within the time horizon from 2010 to 2050. Each model period is divided into 48 time-slices, structured in four seasons, each with a representative day with 12 two-hour periods. Winter is defined as December, January and February, spring is defined as March, April, May, summer is June, July and August and autumn is September, October and November. Investment decisions are made in each model period, and operational decisions are made at the time-slice level. The operational decisions include activity levels for each capacity type, fuel consumption and electricity trade. The model is regionally divided into the two Danish spot price areas: DK1 and DK2. The annual discount rate is set to 4%, and policy instruments like subsidies and tax are excluded.

The heat demand and the non-substitutable electricity demand are exogenous input to the model and is based on energy projections from the Danish Energy Agency [70]. The heat demand is split into three different categories: central heat demand, de-central heat demand and individual heat demand. Central heat is produced in large combined heat and power (CHP) plants, with a flexible heat to electricity production ratio, and is delivered to large district heat (DH) networks. De-central heat is produced in smaller CHP plants, with a fixed heat to power ratio, and in DH plants, and is delivered in smaller and more remote DH grid networks. Individual heat is the remaining heat demand that is met by heating technologies located within buildings, like wood stoves, heat pumps and boilers. The heat demand split does not capture the competition between the different heating options, e.g. whether it is profitable to expand the DH grid to replace the natural gas grid.

The characterisation of energy technologies, with cost data and efficiencies, are input to the model and are primarily based on [71, 72]. The model includes a set of technologies used to transform energy sources to final demand. It is optional to use electricity in heat and transport technologies and consequently the total electricity consumption is a model decision. Future annual energy prices for biomass, fossil fuels and electricity in the neighbouring countries to Denmark are based on [73]. Further, we use hourly electricity prices from 2014 for Germany, Sweden and Norway, provided by Nord Pool Spot and the European Energy Exchange, to map how the electricity prices in these countries vary within a year. The electricity prices in the two Danish regions are endogenous in the model, as they are the dual values of the electricity balance equation, while the electricity prices in neighbouring countries are exogenous. The Danish transmission capacity to its neighbouring countries, Germany, Norway and Sweden, is provided by [73].

3.2 Stochastic modelling approach

The described stochastic modelling approach is applicable to long-term investment models with short-term uncertainty in general and is not limited to the TIMES modelling framework. In our two-stage model, the first-stage decision variables are investments in new capacity and the second-stage variables are operational decisions. Figure 1 illustrates a scenario tree with this information structure. In the first stage, the realisation of the wind scenarios is unknown, and investments in new capacity for the entire model horizon, from 2010 to 2050, are made. In the second stage, starting after the branching point of the scenario tree, the outcomes of the wind scenarios are known, and operational decisions are made for each of the scenarios for all model periods. Consequently, the investments are identical for all scenarios, whereas the operational decisions are scenario dependent. To consider the different operational situations in the optimisation, the model minimises the investment costs and the expected operational costs over all scenarios. This provides investment decisions that both recognise the expected operational cost and that are feasible for all the scenario realisations of wind power production. The investment and operational decisions are made simultaneously, and we apply a multi-horizon model structure [74], where operational decisions are independent between the 5-year model periods.

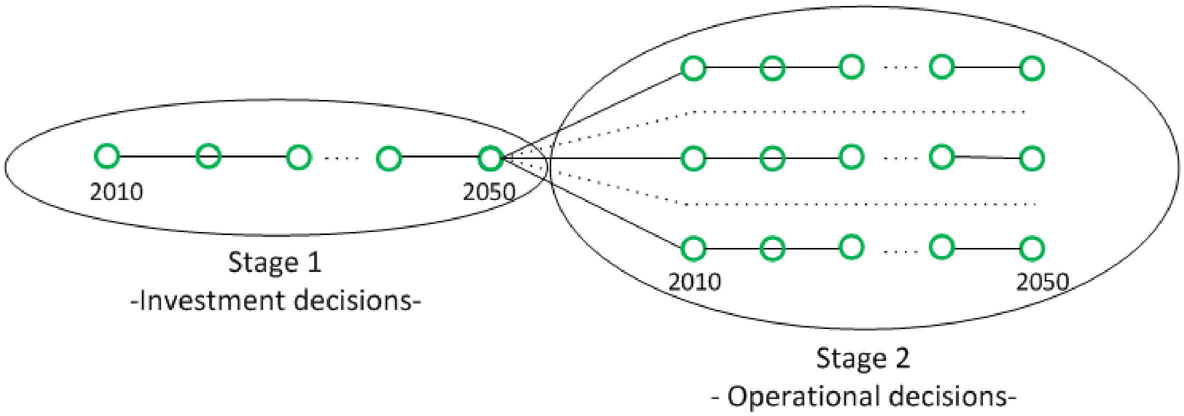

Figure 1: Illustration of the two-stage scenario tree.

3.3 Scenario generation

The only uncertain parameter in our model is wind power availability, which equals the hourly wind power production over the installed capacity. We use historical production and capacity data from 2000 to 2014 for the two Danish Nord Pool regions to estimate the characteristics of this uncertain parameter. As wind turbine design and layout have become more efficient with time, we adjust the wind availability data according to the time, such that the average availability data for the entire historical period equals the annual availability of 2014. We use every second hour from the data set to comply with the time-slice structure of the model with 12 two-hour steps daily. Further, we assume that each historical day of the data set has the same probability to occur (and correspondingly every hour).

The wind scenarios are based on a selection of historical days. First, all the days from 2000 to 2014 are sorted by month and season. Thereafter, a scenario is generated by randomly sampling days within each season, where the model input is the corresponding hourly availability factors of the sampled days. The days are sampled independently for the four seasons, assuming there is no seasonal dependence of wind availability. This procedure is repeated for each five-year period, meaning that each scenario consists of a randomly sampled sequence of days (each with 12 hours). The scenario generation approach ensures a consistent daily correlation since a sampled day consists of 12 chronological two-hour blocks for wind availability that captures the regional correlations, as the same sampled days are used to derive the model input on wind power availability in both model regions. With four seasons and 12 two-hour periods, a scenario consists of 48 stochastic wind parameters for each region and model period. Finally, to cover the monthly variations of wind availability within a season, the scenarios consist of an equal number of days within each month of the four seasons.

The statistical properties of the wind power availability in a sample, using only a limited number of scenarios, can potentially deviate from the true distribution for wind power availability. Figure 2 plots the mean daily wind power availability in fall for the Danish region DK1 for the period for which we have data. We assume the observations represent the outcomes of the true distribution. The figure also shows the maximum, minimum and 75/25 quantile of 1000 different scenario trees generated by random sampling with the following sample sizes, N = 3, 30, 60 and 600. The corresponding variance is plotted in Figure 3. These figures show that the mean value and variance converges to the true distribution when the number of scenarios increases. With a low sample size, the mean value and variance can be identical to the true distribution for one scenario tree but vary highly among the randomly generated scenario trees. We observe a similar trend for all seasons in both model regions. For the sample size of N = 600, the sampled scenarios have approximately the same mean and variance as the true distribution for all scenario trees. Further, we suppose that a reference sample with N' = 600 is adequate to calculate the confidence interval of the model solution with the SAA algorithm.

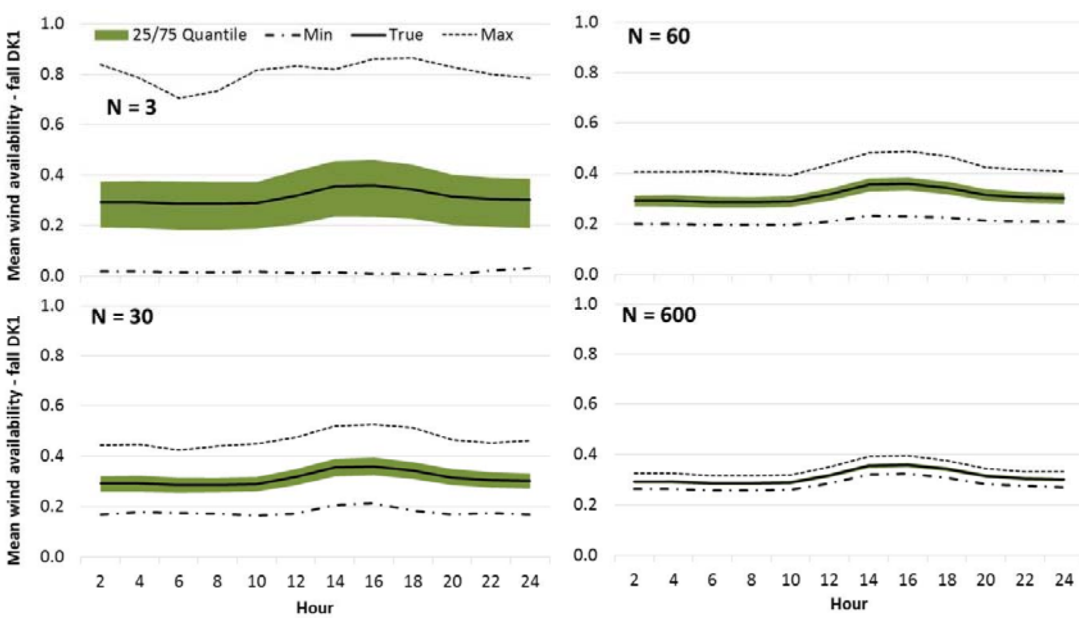

Figure 2: Mean availability factor for fall DK1 for the true distribution as well as the maximum, minimum and 75/25 quantile of 1000 different scenario trees for N = 3, 30, 60 and 600.

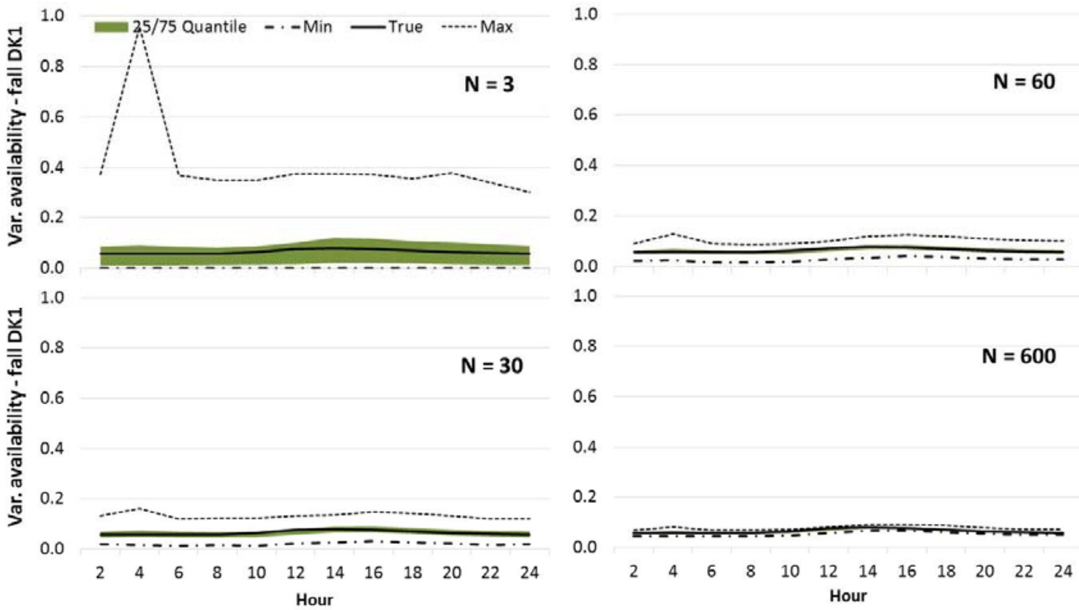

Figure 3: Variance in availability factor for fall DK1 for the true distribution, the maximum, minimum and 75/25 quantile of 1000 different scenario trees for N = 3, 30, 60 and 600.

4. Computational results

We applied the SAA algorithm and the stability tests for ten different candidate samples (M = 10) and for three different sample sizes of N = 3, 30 and 60 scenarios, for a total of 30 different problem instances. Thereafter we evaluated the in-sample and out-of-sample stability for the different scenario tree sizes measured in number of scenarios. The calculations were carried out with a 2.9 GHz CoreI7 Intel processor with 32 GB RAM using the XPRESS solver in GAMS.

4.1 SAA

For a given sample size, N, the scenario set of each candidate sample of the SAA provides a candidate solution to the optimisation problem with a corresponding optimal value. The first-stage decisions from each problem instance and a reference sample at N' = 600, representing a large sample of scenarios are used to estimate the gap between the estimate of the optimal value using N scenarios and the true value. Table 1 shows the optimal value (v), estimated lower bound (*LB*), estimated upper bound (*UB*) and estimated optimality gap (*GAP*) with a 90% confidence interval for all problem instances. For each N=3, 30 and 60, and the number of sample M = 10, the best value of the 10 candidate samples in terms of lowest *UB* and *GAP* is highlighted. The instances are also sorted by ascending *UB*.

Among the instances, *LB* varies only with N as it is an average of all candidate values for that sample size, and *LB* is highest for N = 30. A *UB* estimate is calculated by fixing first-stage decisions of all N*M=30 problem instances in the reference sample, by Eq.(8), and the best *UB* estimate is approximately identical for N = 30 and N = 60. The optimal value improves with an increased sample size as the confidence interval of the optimality gap (*GAP)* narrows as N increases. This is primarily because the variance of the *LB* is reduced with a higher number of scenarios. Note that there is a negative *GAP* for some instances, where *LB* has a larger value than the *UB*. This is because the calculated bounds are estimates of the true bounds. This can be avoided by increasing the value of M and/or N', as the estimated bounds converge to the true bounds when M goes to infinity and N' goes to S.

Among the candidate values, the variance of optimal value v decreases with sample size, and the difference between the maximum and minimum value is *4.0 billion euro (bEUR)*, *1.0 bEUR* and *0.8 bEUR* for N = 3, 30 and 60 respectively. The percentage difference among the candidate values is low, however, as the objective function includes a significant share of costs that are not related to the intermittent characteristics of wind power.

The endogenous wind power capacity in 2020, 2030, 2040 and 2050, as well as estimates of *VSS* for all problem instances, are given in Table 2. For each *N*, the best candidate solution is highlighted, and the instances are sorted by ascending *UB,* as in Table 1. For comparison, we also include the wind capacities provided by a deterministic model, where the expected value of the wind power availability is used as model input.

We estimate VSS for each instance, assuming each scenario sample, with size *N,* is an estimate of the true distribution, providing a solution, $x_{m,N}$, that estimates the true solution, see Eq.(17). This corresponds to a situation where the modeller only solves one model instance of sample size *N,* believing the sample gives a good representation of the true distribution. However, this estimate can deviate significantly from the true VSS as the used sample does not necessarily capture the statistical properties of the uncertainty since $N < S$.

$$VSS\ (x_{m,N}) = \overline{v}_{m,N} - v_{m,N} \qquad (17)$$

For all sample sizes *N*, the best stochastic solution, from the problem instance with the lowest *UB*, has lower wind capacities compared to the capacities provided by the deterministic model. For example, in 2050, the deterministic capacity is *10.9 GW* and the best capacity is *9.5 GW*, *8.9 GW* and *8.8 GW* for *N* = 3, *30* and *60,* respectively. The spread of the wind power capacity among the instances is significant. For example, in 2040, the wind capacity ranging from *– 29%* to *+ 87%* compares to the best stochastic solution among the different problem instances. Furthermore, the spread of the wind capacity for the various candidate samples decreases with the size of *N*. For example, in 2030, the wind capacity ranges from *4.2 GW* to *8.9 GW* for *N* = 3, *4.7 GW* to *7.0 GW* for *N* = 30 and from *6.6 GW* to *8.5 GW* for *N* = 60. The results imply that the stochastic solutions are significantly different than the deterministic solution and that the stochastic solution is highly dependent on the scenario representation. This emphasizes the importance of using a stochastic model and the necessity of evaluating the quality of the stochastic solution.

Table 2 also shows that the estimated *VSS* depends greatly on the scenario representation. The *VSS* ranges from *0.11 bEUR* to *0.84 bEUR* for *N* = 3, from *0.13 bEUR* to *0.43 bEUR* for *N* = 30 and from *0.10 bEUR* to *0.21 bEUR* for *N* = 60. Further, *VSS* of the best candidate solutions decreases with number of scenarios, with *0.18 bEUR*, *0.16 bEUR* and *0.14 bEUR* for *N* = 3, 30 and 60 respectively.

These results demonstrate that if too few scenarios are used, solving a single stochastic instance can lead to substantial and arbitrary errors. Also, the VSS can be considerably misrepresented if we fail to recognise that the scenario number used is too small to give an appropriate representation of the uncertainty. Our computational tests show that this problem is reduced when selecting the best candidate solution However, for our results, VSS is only a small share of the total energy system cost, including the wind power costs (investments and operational costs). For $N = 60$, VSS of the best candidate solution is 1.4% of the wind power costs. Note that this low relative VSS cannot be used to justify using a deterministic modelling approach in general since VSS depends on the model assumptions and methodology. Further, to evaluate the VSS, a stochastic optimisation is required.

Table 1: Optimal objective function (v), estimated lower bound (LB), estimated upper bound (UB) and estimated optimality gap (GAP) for a 90 % confidence interval for all problem instances [billion euro] where the best value in terms of lowest GAP is highlighted.

N = 3	m	V	LB			UB			GAP		
	5	**148.90**	**147.79**	**±**	**0.74**	**148.75**	**±**	**0.14**	**0.97**	**±**	**0.75**
	9	148.44	147.79	±	0.74	148.79	±	0.16	1.00	±	0.76
	4	149.26	147.79	±	0.74	148.81	±	0.15	1.03	±	0.75
	6	147.58	147.79	±	0.74	148.84	±	0.15	1.06	±	0.75
	8	146.97	147.79	±	0.74	148.87	±	0.14	1.08	±	0.75
	3	145.50	147.79	±	0.74	148.87	±	0.16	1.08	±	0.76
	10	147.86	147.79	±	0.74	148.96	±	0.16	1.18	±	0.76
	1	145.48	147.79	±	0.74	148.98	±	0.17	1.19	±	0.76
	7	149.45	147.79	±	0.74	149.20	±	0.11	1.42	±	0.75
	2	148.42	147.79	±	0.74	149.36	±	0.15	1.57	±	0.75
N = 30	**m**	**V**	**LB**			**UB**			**GAP**		
	9	**148.48**	**148.76**	**±**	**0.57**	**148.72**	**±**	**0.15**	**-0.04**	**±**	**0.59**
	10	148.46	148.76	±	0.57	148.73	±	0.15	-0.04	±	0.59
	1	148.34	148.76	±	0.57	148.73	±	0.15	-0.04	±	0.59
	7	148.34	148.76	±	0.57	148.73	±	0.15	-0.03	±	0.59
	4	148.86	148.76	±	0.57	148.77	±	0.13	0.01	±	0.58
	3	149.12	148.76	±	0.57	148.78	±	0.14	0.02	±	0.58
	5	148.92	148.76	±	0.57	148.79	±	0.13	0.02	±	0.58
	8	149.01	148.76	±	0.57	148.82	±	0.13	0.05	±	0.58
	6	148.78	148.76	±	0.57	148.82	±	0.13	0.06	±	0.58
	2	149.33	148.76	±	0.57	148.85	±	0.13	0.09	±	0.58
N = 60	**m**	**V**	**LB**			**UB**			**GAP**		
	5	**148.48**	**148.70**	**±**	**0.37**	**148.72**	**±**	**0.15**	**0.03**	**±**	**0.40**
	1	148.91	148.70	±	0.37	148.72	±	0.14	0.03	±	0.40
	7	148.74	148.70	±	0.37	148.73	±	0.15	0.03	±	0.40
	8	148.33	148.70	±	0.37	148.74	±	0.15	0.04	±	0.40
	9	148.61	148.70	±	0.37	148.74	±	0.14	0.04	±	0.39
	6	148.66	148.70	±	0.37	148.74	±	0.15	0.04	±	0.40

10	149.09	148.70	±	0.37	148.75	±	0.14	0.05	±	0.39
3	148.60	148.70	±	0.37	148.75	±	0.14	0.05	±	0.39
4	148.62	148.70	±	0.37	148.76	±	0.14	0.06	±	0.40
2	148.91	148.70	±	0.37	148.76	±	0.14	0.07	±	0.39

Table 2: Estimated value of stochastic solution, VSS [billion euro] and wind capacity in 2020, 2030, 2040 and 2050 [GW] for all problem instances where the best value in terms of lowest GAP is highlighted.

		Wind capacity, GW				Estimated VSS
N = 3	m	2020	2030	2040	2050	bEUR
	5	**4.7**	**4.7**	**8.6**	**9.5**	**0.18**
	9	4.4	4.4	9.0	8.8	0.11
	4	6.5	6.5	9.0	5.3	0.21
	6	6.3	6.3	10.5	10.4	0.34
	8	4.2	4.2	11.1	10.6	0.21
	3	4.2	4.2	9.7	9.2	0.27
	10	8.9	8.9	8.9	8.6	0.38
	1	7.6	7.6	9.7	6.8	0.30
	7	4.6	4.6	8.5	8.6	0.85
	2	8.3	8.3	4.4	7.5	0.84
N = 30	m	2020	2030	2040	2050	VSS
	9	**5.7**	**5.7**	**8.9**	**8.9**	**0.16**
	10	6.3	6.3	8.7	8.7	0.13
	1	4.5	4.5	8.9	8.9	0.15
	7	6.3	6.3	8.9	9.5	0.13
	4	4.7	4.7	8.7	8.7	0.23
	3	4.7	4.7	8.6	7.5	0.23
	5	4.7	4.7	8.6	8.7	0.25
	8	6.3	6.3	9.0	8.7	0.24
	6	5.3	5.3	8.8	8.8	0.30
	2	7.0	7.0	8.3	8.3	0.43
N = 60	m	2020	2030	2040	2050	VSS

5	**5.9**	**7.9**	**8.8**	**8.8**	**0.14**
1	6.3	6.9	8.6	8.6	0.21
7	4.7	8.4	8.6	8.8	0.15
8	5.0	8.6	9.1	9.5	0.10
9	5.0	8.4	9.1	9.1	0.21
6	6.3	7.8	8.7	9.3	0.13
10	5.0	7.6	8.6	8.6	0.21
3	6.0	6.6	9.1	9.1	0.18
4	6.3	8.2	8.9	7.8	0.15
2	4.8	8.5	8.6	8.5	0.19

	Wind capacity, GW			
	2020	**2030**	**2040**	**2050**
Det.	6.3	8.6	10.9	10.9

4.2 Stability test

This section demonstrates the in-sample and out-of-sample stability of the wind availability scenarios with sample size N =3, 30 and 60. The in-sample stability is illustrated in Figure 4, where the optimal values for all problem instances are plotted relative to the optimal value of the best candidate solution for N = 60. Here, the in-sample stability is improved with a higher sample size, and the relative value ranges from *0.980* to *1.007* for N = 3, *0.999* to *1.006* for N = 30 and from *0.999* to *1.004* for N = 60. Further, the out-of-sample stability is illustrated in Figure 5, which shows the *UB* for all problem instances, relative to the *UB* of the best candidate solution (from the SAA) for N=60. The out-of-sample stability ranges from *1.000* to *1.004* for N = 3, *1.000* to *1.001* for N = 30 and from *1.000* to *1.000* for N = 60. The figures show a higher variance in the in-sample stability than in the out-of-sample stability. The results for this study show that using 60 scenarios to describe short-term uncertainty in wind power production provides both in-sample and out-of-sample stability.

Since the wind power costs are only a small part of the total energy system cost, the influence of the wind availability scenario on the in-sample and out-of-sample stability is marginal. It is therefore more appropriate to show the wind power expenses to isolate the effect of the different model instances. This is illustrated in Figure 6, where the sum of the investment and maintenance costs of the new wind power capacity towards 2050 for all problem instances is plotted. Here, the stability in the wind power costs is improved with a higher sample size, and the costs range from *6.8 bEUR* to *12.8 bEUR* for N = 3, *8.4 bEUR* to *10.2 bEUR* for N = 30 and from *9.3 bEUR* to *10.4 bEUR* for N = 60. Note that the wind power costs are identical for the in-sample and out-of-sample tests since the first-stage decisions, including wind power investments, are set identical in both tests.

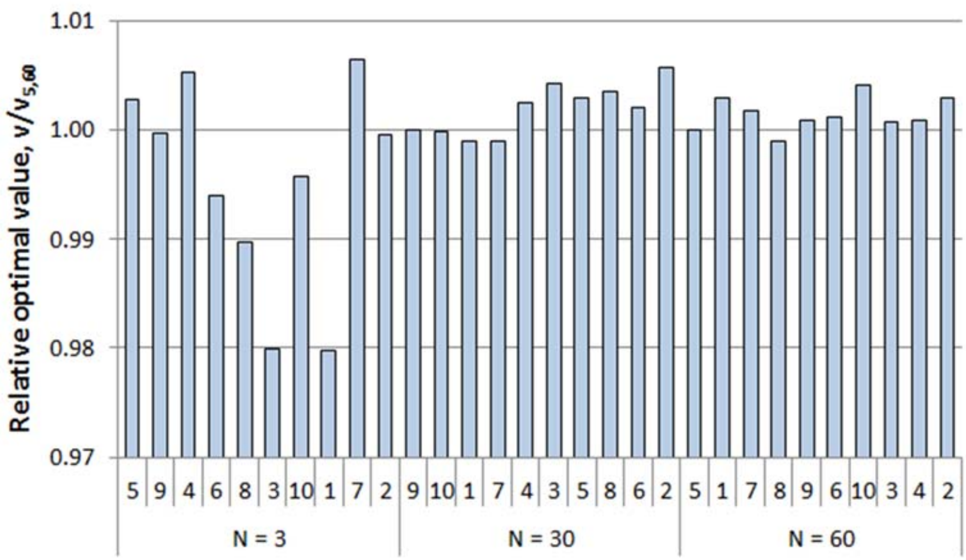

Figure 4: Illustration of in-sample stability.

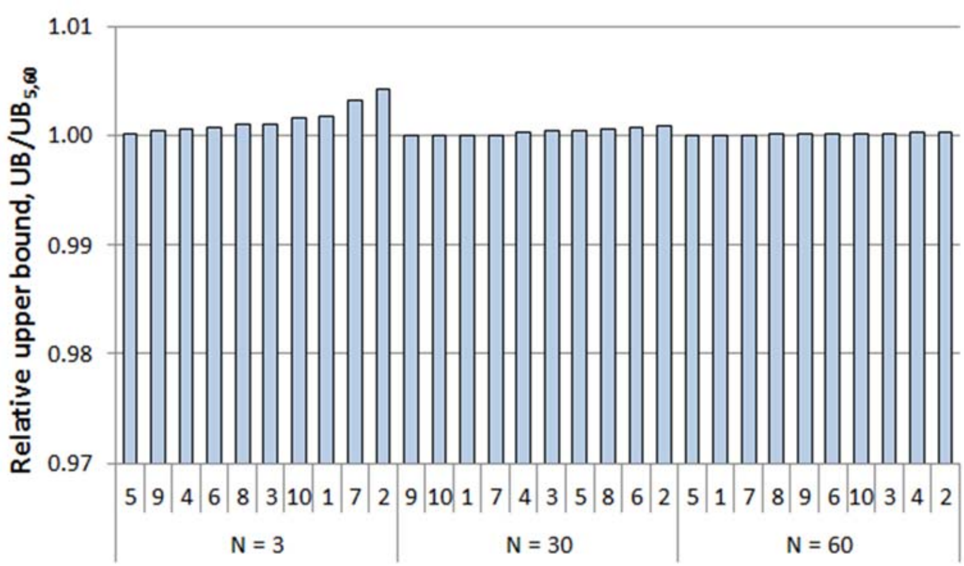

Figure 5: Illustration of out-of-sample stability.

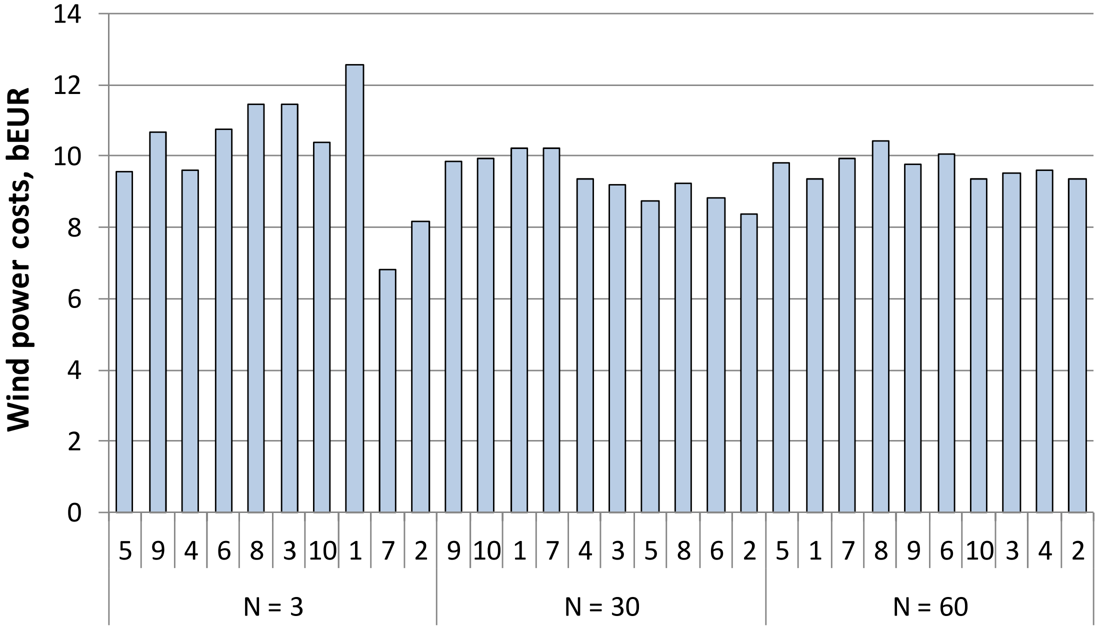

Figure 6: Illustration of wind power costs.

4.3 Comparison of SAA and stability tests

We observe from the calculations of the SAA confidence intervals that the SAA algorithm also provides the information required to perform the stability tests. The stability tests examine the required sample size needed to provide stability. This implies that the in-sample and out-of-sample tests find the necessary number of scenarios needed to ensure that the solution is not distorted by the scenario representation. The SAA algorithm provides confidence intervals of the optimality gap for a given model solution, and these depend on both the sample sizes used in the iterative process and the number of candidate iterations.

A solution approach where the model is solved once, for a specific sample size, requires a given number of scenarios to achieve stability. In that context, we have two important observations. First, the SAA algorithm does not necessarily require in-sample and out-of-sample stability for the solved sample problems to provide a solution that has an optimality gap that is within a user-defined acceptance level.

This is because a solution an acceptable confidence interval can be produced by testing many candidate solutions identified in the iterations. Second, and equally important, if SAA repeatedly solves the model with a sample size that satisfies in-sample stability, the confidence interval of the best candidate value is narrow, and the solution is thereby of high quality.

5. Conclusions

This paper applies SAA and stability tests to evaluate the quality of the investment decisions provided by a long-term energy system model of Denmark with stochastic wind power production. Whereas the stability tests are applicable to all scenario generation methods, SAA uses randomly sampled scenarios. In a stability perspective, the scenario quality is satisfactory when the solution satisfies in-sample and out-of-sample stability. SAA provides confidence intervals of the optimality gap for given model solutions, and the optimality gap depends both on the sample sizes used in the iterative SAA method and the number of replications. A solution approach whereby the model is solved only once, for a specific sample size, requires a given number of scenarios to achieve stability.

The main contributions of the paper are as follows. First, we address a typical situation in the energy literature wherein a stochastic model is solved with a limited number of scenarios instead of a deterministic model. We show that these stochastic models can provide an arbitrary and wrong estimate of the VSS, compared to a stochastic model with a larger number of scenarios. Second, the scenario representation affects the actual solutions. This is illustrated in our examples with wind power capacity ranging from *–29%* to *+87%* compared to the best stochastic solution in 2040 among the different problem instances. Third, we demonstrate that this problem is reduced when applying SAA since the quality of different candidate solutions are tested based on candidates from independent samples, evaluated using the reference sample. Fourth, we illustrate that when solving single instances instead of using sampling methods like SAA, stability tests offer a good way of ensuring that solutions are meaningful. The challenge, however, is that the number of scenarios needed to achieve stability may be problem specific.

Our results demonstrate how the quality of the model solution and the estimated VSS are improved with a larger sample size. The results of this study show that using 60 scenarios to describe the short-term uncertainty related to wind power production provides both in-sample and out-of-sample stability. We also demonstrate that a solution with a satisfactory user-defined confidence interval, provided by SAA, does not necessarily require stability in the sample size. The quality, in terms of a low confidence interval, is ensured by iteration. On the other hand, a solution with a confidence interval that is lower than a user-defined limit, generated by testing many candidate solutions identified in the SAA iterations, does not necessarily require in-sample and out-of-sample stability for each iteration. An important observation, though, is that if SAA repeatedly solves the model, with a sample size that satisfies in-sample stability, the confidence interval is narrow, and the solutions are of high quality in terms of in-sample and out-of-sample stability as well as providing a tight bound for the optimal solution. Another observation is that there are significant similarities between the calculation of SAA bounds and the stability tests.

Acknowledgements

The authors thank the anonymous reviewer for contributing to improve the quality of the paper considerably with valuable comments and suggestions. The authors are grateful to the Research Council of Norway for funding this work through the following projects: "The future Norwegian energy system in a North-European context" (Grant number: 207067; https://www.forskningsradet.no/prosjektbanken/#/project/NFR/207067/Sprak=en) and "Assessment of the value of flexibility services from the Norwegian energy system" (Grant number: 268097; https://prosjektbanken.forskningsradet.no/#/project/NFR/268097/Sprak=en).

References

1. Loulou, R., et al., *Documentation for the TIMES Model - Part I*. 2005, Energy Technology Systems Analysis Program: http://iea-etsap.org/docs/TIMESDoc-Intro.pdf.
2. Loulou, R., et al., *Documentation for the TIMES Model - Part II*. 2005, Energy Technology Systems Analysis Programme: http://iea-etsap.org/docs/TIMESDoc-Details.pdf.
3. Loulou, R., et al., *Documentation for the TIMES Model - Part III*. 2005, Energy Technology Systems Analysis Programme: http://iea-etsap.org/docs/TIMESDoc-GAMS.pdf.
4. Loulou, R. and M. Labriet, *ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure.* Computational Management Science 2008. **5**: p. 7-40.
5. Loulou, R., *ETSAP-TIAM: the TIMES integrated assessment model. part II: mathematical formulation.* Computational Management Science, 2008. **5**: p. 41-66.
6. Collins, S., et al., *Integrating short term variations of the power system into integrated energy system models: A methodological review.* Renewable and Sustainable Energy Reviews, 2017. **76**: p. 839-856.
7. Connolly, D., et al., *A review of computer tools for analysing the integration of renewable energy into various energy systems.* Applied Energy, 2010. **87**(4): p. 1059-1082.
8. Pfenninger, S., A. Hawkes, and J. Keirstead, *Energy systems modeling for twenty-first century energy challenges.* Renewable and Sustainable Energy Reviews, 2014. **33**: p. 74-86.
9. Kall, P. and S.W. Wallace, *Stochastic Programming*. 1994, Chichester: John Wiley & Sons.
10. Loulou, R. and A. Lehtila *Stochastic Programming and Tradeoff Analysis in TIMES. TIMES Version 3.3 User Note*. 2012.
11. Nijs, W. and K. Poncelet, *Integrating recurring uncertainties in ETSAP energy system models*. 2016, VITO: http://iea-etsap.org/projects/ETSAP%20project%20Recurring%20Risk%20v11.pdf.
12. Seljom, P. and A. Tomasgard, *Short-term uncertainty in long-term energy system models — A case study of wind power in Denmark.* Energy Economics, 2015. **49**(0): p. 157-167.
13. Seljom, P., et al., *The impact of Zero Energy Buildings on the Scandinavian energy system.* Energy, 2017. **118**: p. 284-296.
14. Seljom, P. and A. Tomasgard, *The impact of policy actions and future energy prices on the cost-optimal development of the energy system in Norway and Sweden.* Energy Policy, 2017. **106**: p. 85-102.
15. Zhang, H., W. Chen, and W. Huang, *TIMES modelling of transport sector in China and USA: Comparisons from a decarbonization perspective.* Applied Energy, 2016. **162**: p. 1505-1514.
16. Shi, J., W. Chen, and X. Yin, *Modelling building's decarbonization with application of China TIMES model.* Applied Energy, 2016. **162**: p. 1303-1312.
17. Daly, H.E., et al., *Incorporating travel behaviour and travel time into TIMES energy system models.* Applied Energy, 2014. **135**: p. 429-439.
18. Vaillancourt, K., et al., *A Canadian 2050 energy outlook: Analysis with the multi-regional model TIMES-Canada.* Applied Energy, 2014. **132**: p. 56-65.
19. Lind, A., et al., *Analysis of the EU renewable energy directive by a techno-economic optimisation model.* Energy Policy, 2013. **60**: p. 364-377.
20. Krook Riekkola, A., E.O. Ahlgren, and P. Söderholm, *Ancillary benefits of climate policy in a small open economy: The case of Sweden.* Energy Policy, 2011. **39**(9): p. 4985-4998.
21. IEA, *Nordic Energy Technology Perspectives*. Cities, flexibility and pathways to carbon-neutrality, ed. I.E. Agency. 2016, Paris, France: Nordic Energy Research.
22. Nagl, S., M. Fürsch, and L. Dietmar, *The Costs of Electricity Systems with a High Share of Fluctutating Renewables: A Stochastic Investment and Dispatch Optimization Model for Europe.* The Energy Journal, 2013. **34**: p. 151-179.
23. Spiecker, S., P. Vogel, and C. Weber, *Evaluating interconnector investments in the north European electricity system considering fluctuating wind power penetration.* Energy Economics, 2013. **37**(0): p. 114-127.

24. Sun, N., I. Ellersdorfer, and D.J. Swider. *Model-based long-term electricity generation system planning under uncertainty*. in *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*. 2008. NanJing China.
25. Skar, C., G. Doorman, and A. Tomasgard. *The future European power system under a climate policy regime*. in *Energy Conference (ENERGYCON), 2014 IEEE International*. 2014.
26. Kleywegt, A.J., A. Shapiro, and T. Homem-de-Mello, *The Sample Average Approximation Method for Stochastic Discrete Optimization.* SIAM Journal of Optimization, 2001. **12**(2): p. 479 - 502
27. Shapiro, A., *Monte Carlo Sampling Methods*, in *Handbooks in Operations Research and Management Science*, A. Ruszczynski and A. Shapiro, Editors. 2003, Elsevier. p. 353-425.
28. Kaut, M. and S.W. Wallace, *Evaluation of Scenario-Generation Methods for Stochastic Programming.* Pacific Journal of Optimization, 2007. **3**: p. 257-271.
29. Glasserman, P., *Monte Carlo Methods in Financial Engineering*. 2013: Springer New York.
30. Fleishman, A.I., *A method for simulating non-normal distributions.* Psychometrika, 1978. **43**(4): p. 521-532.
31. Høyland, K., M. Kaut, and S.W. Wallace, *A Heuristic for Moment-Matching Scenario Generation.* Computational Optimization and Applications, 2003. **24**(2): p. 169-185.
32. Ponomareva, K., D. Roman, and P. Date, *An algorithm for moment-matching scenario generation with application to financial portfolio optimisation.* European Journal of Operational Research, 2015. **240**(3): p. 678-687.
33. Dupačová, J., N. Gröwe-Kuska, and W. Römisch, *Scenario reduction in stochastic programming.* Mathematical Programming, 2003. **95**(3): p. 493-511.
34. Heitsch, H. and W. Römisch, *Scenario Reduction Algorithms in Stochastic Programming.* Computational Optimization and Applications, 2003. **24**(2): p. 187-206.
35. Heitsch, H. and W. Römisch, *Scenario tree reduction for multistage stochastic programs.* Computational Management Science, 2009. **6**(2): p. 117-133.
36. Heitsch, H. and W. Römisch, *Stability and Scenario Trees for Multistage Stochastic Programs*, in *Stochastic Programming: The State of the Art In Honor of George B. Dantzig*, G. Infanger, Editor. 2011, Springer New York: New York, NY. p. 139-164.
37. Heitsch, H. and W. Römisch, *Scenario tree modeling for multistage stochastic programs.* Mathematical Programming, 2009. **118**(2): p. 371-406.
38. Pflug, G.C. and A. Pichler, *Approximations for Probability Distributions and Stochastic Optimization Problems*, in *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*, M. Bertocchi, G. Consigli, and M.A.H. Dempster, Editors. 2011, Springer New York: New York, NY. p. 343-387.
39. Hochreiter, R. and G.C. Pflug, *Financial scenario generation for stochastic multi-stage decision processes as facility location problems.* Annals of Operations Research, 2007. **152**(1): p. 257-272.
40. Pflug, G.C., *Scenario tree generation for multiperiod financial optimization by optimal discretization.* Mathematical Programming, 2001. **89**(2): p. 251-271.
41. Quan, H., et al., *A computational framework for uncertainty integration in stochastic unit commitment with intermittent renewable energy sources.* Applied Energy, 2015. **152**: p. 71-82.
42. Jin, S., et al., *Modeling and solving a large-scale generation expansion planning problem under uncertainty.* Energy Systems, 2011. **2**(3): p. 209-242.
43. Munoz, F.D. and J.-P. Watson, *A scalable solution framework for stochastic transmission and generation planning problems.* Computational Management Science, 2015. **12**(4): p. 491-518.
44. Feng, Y. and S.M. Ryan, *Scenario construction and reduction applied to stochastic power generation expansion planning.* Computers & Operations Research, 2013. **40**(1): p. 9-23.
45. Fotouhi Ghazvini, M.A., et al., *Stochastic short-term maintenance scheduling of GENCOs in an oligopolistic electricity market.* Applied Energy, 2013. **101**: p. 667-677.
46. Siahkali, H. and M. Vakilian, *Stochastic unit commitment of wind farms integrated in power system.* Electric Power Systems Research, 2010. **80**(9): p. 1006-1017.

47. Meibom, P., et al., *Stochastic Optimization Model to Study the Operational Impacts of High Wind Penetrations in Ireland.* Power Systems, IEEE Transactions on, 2011. **26**(3): p. 1367-1379.
48. Gil, E., et al., *Generation Capacity Expansion Planning Under Hydro Uncertainty Using Stochastic Mixed Integer Programming and Scenario Reduction.* IEEE Transactions on Power Systems, 2015. **30**(4): p. 1838-1847.
49. Ottesen, S.O. and A. Tomasgard, *A stochastic model for scheduling energy flexibility in buildings.* Energy, 2015. **88**: p. 364-376.
50. Jaehnert, S., et al., *Transmission expansion planning in Northern Europe in 2030—Methodology and analyses.* Energy Policy, 2013. **61**(0): p. 125-139.
51. Ottesen, S.Ø., A. Tomasgard, and S.-E. Fleten, *Prosumer bidding and scheduling in electricity markets.* Energy, 2016. **94**: p. 828-843.
52. Pappala, V.S., et al., *A Stochastic Model for the Optimal Operation of a Wind-Thermal Power System.* IEEE Transactions on Power Systems, 2009. **24**(2): p. 940-950.
53. Baringo, L. and A.J. Conejo, *Risk-Constrained Multi-Stage Wind Power Investment.* Power Systems, IEEE Transactions on, 2013. **28**(1): p. 401-411.
54. Vaagen, H. and S.W. Wallace, *Product variety arising from hedging in the fashion supply chains.* International Journal of Production Economics, 2008. **114**(2): p. 431-455.
55. Fang, Y., L. Chen, and M. Fukushima, *A mixed R&D projects and securities portfolio selection model.* European Journal of Operational Research, 2008. **185**(2): p. 700-715.
56. Hoff, A., et al., *A metaheuristic for stochastic service network design.* Journal of Heuristics, 2010. **16**(5): p. 653-679.
57. Qianfan, W., W. Jianhui, and G. Yongpei, *Price-Based Unit Commitment With Wind Power Utilization Constraints.* Power Systems, IEEE Transactions on, 2013. **28**(3): p. 2718-2726.
58. Zhao, C., et al., *Expected Value and Chance Constrained Stochastic Unit Commitment Ensuring Wind Power Utilization.* IEEE Transactions on Power Systems, 2014. **29**(6): p. 2696-2705.
59. Schütz, P., A. Tomasgard, and S. Ahmed, *Supply chain design under uncertainty using sample average approximation and dual decomposition.* European Journal of Operational Research, 2009. **199**(2): p. 409-419.
60. Santoso, T., et al., *A stochastic programming approach for supply chain network design under uncertainty.* European Journal of Operational Research, 2005. **167**(1): p. 96-115.
61. Verweij, B., et al., *The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study.* Computational Optimization and Applications, 2003. **24**(2-3): p. 289-333.
62. Ünlü, Y. and M.S. Rosetti. *Evaluating variance reduction techniques within a sample average approximation method for a constrained inventory policy optimization problem*. in *Simulation Conference (WSC), Proceedings of the 2011 Winter*. 2011.
63. Long, Y., L.H. Lee, and E.P. Chew, *The sample average approximation method for empty container repositioning with uncertainties.* European Journal of Operational Research, 2012. **222**(1): p. 65-75.
64. Swamy, C. and D.B. Shmoys, *Sampling-Based Approximation Algorithms for Multistage Stochastic Optimization.* SIAM Journal on Computing, 2012. **41**(4): p. 975-1004.
65. Higle, L.J., *Stochastic Programming: Optimization When Uncertainty Matters*, in *Tutorials in Operational Research*. 2005, INFORMS: New Orleans. p. 24.
66. Birge, J.R., *The value of the stochastic solution in stochastic linear programs with fixed recourse.* Mathematical Programming, 1982. **24**(1): p. 314-325.
67. Mak, W.-K., D.P. Morton, and R.K. Wood, *Monte Carlo bounding techniques for determining solution quality in stochastic programs.* Operations Research Letters, 1999. **24**(1–2): p. 47-56.
68. Shapiro, A. and T. Homem-de-Mello, *A simulation-based approach to two-stage stochastic programming with recourse.* Mathematical Programming, 1998. **81**(3): p. 301-325.
69. Linderoth, J., A. Shapiro, and S. Wright, *The empirical behavior of sampling methods for stochastic programming.* Annals of Operations Research, 2006. **142**(1): p. 215-241.
70. Danish Energy Agency, *Danmarks energifremskriving, 2012*. 2012: Copenhagen.

71. NVE, *Kostnader i energisektoren. Kraft, varme og effektivisering,*. 2015, http://publikasjoner.nve.no/rapport/2015/rapport2015_02a.pdf: The Norwegian Water Resources and Energy Directorate.
72. Lind, A. and E. Rosenberg, *TIMES-Norway model documentation*. 2013, Institute for Energy Technology: http://www.ife.no/en/publications/2013/ensys/times-norway-model-documentation.
73. Energinet.dk, *Energinet.dk's analyseforudsætninger 2015-2035*. 2015: http://energinet.dk/SiteCollectionDocuments/Danske%20dokumenter/El/15-00248-7%20Energinet.dk%27s%20analyseforudsætninger%202015-2035%20-%20ekstern%20version.docx%202860511_2_1.pdf.
74. Kaut, M., et al., *Multi-horizon stochastic programming.* Computational Management Science, 2014. **11**: p. 179-193.