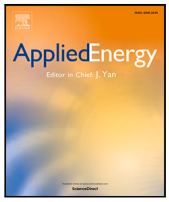
Applied Energy 302 (2021) 117538

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Stable stochastic capacity expansion with variable renewables: Comparing moment matching and stratified scenario generation sampling

Stian Backe *, Mohammadreza Ahang, Asgeir Tomasgard

Department of Industrial Economics and Technology Management, The Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Keywords:
Power system modelling
Stochastic programming
Multi-horizon representation
Scenario generation
Stability testing
Moment matching

ABSTRACT

This paper examines the importance of including operational scenarios representing short-term stochasticity in the long-term capacity expansion models with high shares of variable renewables. As scenario generation routines often are probabilistic, for example based on sampling, it is crucial that they ensure stable results in the capacity expansion model, so that it is the underlying uncertainty that decides the optimal solution, and not the approximation of that uncertainty in the model. However, it is unclear which operational scenario properties that are important to ensure good results and stability in stochastic models. This paper evaluates three sampling-based scenario generation routines in a multi-horizon stochastic capacity expansion problem representing the European electricity system. We compare the use of stochastic versus deterministic modelling with high shares of variable renewables. Further, we perform in-sample and out-of-sample stability tests on 90 scenario trees for each routine, and we compare the routines' ability to produce stable system costs and capacity investments when approximating the optimal value from the real distribution. Results show that stochastic modelling with more than 80% share of variable renewables leads to more investments in both dispatchable and variable renewable capacity compared to deterministic modelling, which means that stochastic modelling should be used with very high shares of variable renewables. The scenario generation routine based on stratified sampling increases stability with the same number of operational scenarios compared to its alternatives, and scenario generation routines using stratified sampling should be further explored.

1. Introduction

The European energy roadmap 2050 specifies the level of greenhouse gas (GHG) emissions to be 80–95% of 1990 levels by 2050. This roadmap is aligned with keeping global warming below 2 °C. Hence, the European power system as an important part of the European energy system [1] will experience high share of variable renewable energy sources (VRES) to achieve net zero emissions in 2050 [2]. VRES such as wind and solar are not deterministic, meaning it is not possible to reliably forecast weather condition decades into the future. The variability and uncertainty of VRES will affect the power system operation [3], and a high VRES share in the energy mix calls for the consideration of high temporal resolution [4] and stochasticity of VRES supply and load when modelling energy systems [5,6]. This study focuses on short-term stochasticity and the comparison of different stochastic scenario generation routines, which influence power system operations and therefore also long-term investments.

To understand the impact of short-term uncertainty related to VRES, energy system models apply stochastic programming [7] using representative scenarios in an approximate scenario tree formulation [8]. The scenario tree formulation results in a dramatic growth in the computational effort with many discrete scenarios representing realizations of stochastic processes, and the trade-off between the granularity of a model and computational effort should be taken into consideration [9]. Planning power system generation and transmission expansion in the long-term under uncertainty can be done endogenously with two levels of decisions: long-term strategic investment decisions and short-term operational decisions [10], where explicit representation of operational decisions influence and improve the strategic decisions [11], including the consideration of different climatic scenarios [12]. Although Ludig et al. [13] conclude that considering deterministic short-term operations with increased temporal resolution seems to stabilize results, Seljom and Tomasgard [14] and Pineda and Morales [15] show that a stochastic representation of short-term VRES reveals limitations with deterministic modelling of short-term operations. Ringkjøb et al. [16] confirm the findings in [14] with a larger case study, and they recommend using stochastic models for long-term modelling of

* Correspondence to: Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Høgskoleringen 1, 7491 Trondheim, Norway.
E-mail address: stian.backe@ntnu.no (S. Backe).

https://doi.org/10.1016/j.apenergy.2021.117538
Received 9 May 2021; Received in revised form 30 July 2021; Accepted 3 August 2021
Available online 18 August 2021

Abbreviations

EEV	Expected result of using the expected value solution
EV	Expected value problem
RAM	Random access memory
SD	Standard deviation
SGR	Scenario generation routine
SP	Stochastic problem
VRES	Renewable energy sources
VSS	Value of stochastic solution

energy system with high shares of VRES. Thus, several capacity expansion models are considering stochastic VRES, e.g., EMPIRE [17–20], TIMES [14,16], EMPS [21], and E2M2 [22].

The collection of representative scenarios in stochastic capacity expansion models, the scenario tree, is generated with a scenario generation routine (SGR). Most of the SGRs are probabilistic, for example based on sampling, meaning that they provide a different tree every time the routine is used, even for the same underlying data. The difference in optimal solutions and expected total costs for the energy system model when using different trees from the same SGR with the same underlying data, reflects the *stability* of the results. Also, as the scenario trees are approximations of the real underlying distributions, the model solution also likely deviates from the true solution of the true stochastic program. In a well-functioning SGR, increasing the tree size should lead to convergence towards stable solutions and expected costs.

SGR stability is evaluated in repeated numerical experiments estimating in-sample and out-of-sample stability [23]. The in-sample stability test evaluates whether the SGR produces similar optimal objective function values when generating different scenario trees. The SGR is in-sample stable if different scenario trees of the same size with the same underlying data result in similar objective function values. The out-of-sample stability test evaluates whether the first-stage decisions produced when solving the problem using different trees from the SGR yield similar optimal objective function values in the 'true' problem. It is commonly hypothesized that if the statistical properties of the scenario tree and the real world align, stability can be achieved with smaller trees. Kaut and Wallace [23] claim that practical performance, instead of theoretical properties, should be considered when evaluating SGRs, because the usefulness of the SGR depends on the structure of the model.

In this paper, we present a case study using the multi-horizon stochastic programming model EMPIRE [17–20] to model capacity expansion for the European power system from 2020 to 2050 subject to carbon emission constraints in line with the decarbonization pathway laid out in [24]. The focus of this paper is not on the resulting technology mix itself, but rather on the difference between using stochastic modelling compared to deterministic modelling. Further, we focus on the difference between results using three sampling-based SGRs with increasingly high shares of VRES. We perform in-sample and out-of-sample tests [23], compare the results across the SGRs, and evaluate which SGR that performs best in our context. We also investigate how capacity expansion decisions are impacted by the different SGRs, and how this links to the stability of the SGRs.

The main contributions of this paper are the following:

1. We present the impact of using single deterministic operational scenarios compared to several stochastic operational scenarios in a long-term multi-horizon capacity expansion model with high shares of VRES.
2. We present which SGR that yield most stability in energy system modelling with VRES.

The structure of the paper is as follows: Section 2 presents existing research related to scenario generation and its application in stochastic capacity expansion models. Section 3 presents EMPIRE and the details of the three SGRs, while Section 4 presents the design of our case study. Section 5 presents and discusses the results from our case study, while Section 6 concludes the paper.

2. Background

Stochastic optimization methods are applied to study the impact of uncertainty on decision making, in particular within finance and energy [25]. Giglio [26] presents an early attempt of designing models that support capacity expansion decisions for a facility subject to uncertain demand and facility lifetime. Hennessy [27] study how risk attitudes and potential rewards impact the optimal capacity choice in a two-stage problem.

Long-term planning of power systems with high share of VRES should take into account the stochasticity of different processes in order to find robust and stable solutions. Caramanis et al. [28] present one of the first attempts to deal with non-dispatchable generation in long-term capacity expansion models. Pineda and Morales [15] develop a stochastic expansion model representing a two-stage electricity market to consider wind forecasting errors, and they find that endogenous forecasting uncertainty leads to less VRES compared to no forecasting uncertainty. As discussed in Section 1, it is recommended to consider a wide range of possible stochastic scenarios to reflect the uncertain nature of VRES. Today, we are not able to use such a large tree in our model due to computational effort, thus SGR's are designed to generate stochastic scenarios representative of true tree. A good SGR produces stable results while also capturing crucial stochastic process characteristics with minimal computational effort. The design of the SGR itself often rely on historical observations, statistical properties, and/or expert's opinion of stochastic processes. We consider two SGR categories: statistical-based SGRs and measure-based SGRs.

Statistical-based SGRs construct scenarios given statistical information about stochastic processes. Quan et al. [29] use Monte Carlo simulation [30] based on prediction intervals and empirical distribution functions to construct VRES scenarios applied in a unit commitment problem. Ottesen and Tomasgard [31] construct scenarios for scheduling energy flexibility by assuming a certain variation from a forecast profile. Simulation-based moment methods are statistical-based SGRs that construct scenarios with known statistical moments of stochastic processes [32]. Høyland et al. [33] present an SGR constructing multivariate scenario trees matching specified moments and correlations for financial optimization. Jin et al. [34] use the method in [33] in a two-stage stochastic mixed-integer capacity expansion model to generate scenarios for demand and gas price. Ponomareva et al. [35] present a less computationally demanding SGR for financial optimization based on constructing scenario trees that exactly matches specified mean, covariance, and third and fourth moments of a random vector. Hochreiter and Pflug [36] discuss the potential failure of moment matching methods to represent the stochastic parameters by demonstrating that two distributions with the same first four moments can have totally different shape.

Measure-based SGRs produce discrete samples directly from historical data rather than assuming statistical properties, and, given sufficient measurement data, these SGRs therefore can preserve more information about stochastic processes than statistical-based SGRs. Gil et al. [37] and de Oliveira et al. [38] use historical hydrological data as the basis for scenario generation in power generation models, and they further perform scenario tree reduction [39] to find a subset of scenarios that best match the statistical properties of the whole tree. Xu et al. [40] use machine learning in their SGR to generate hydro flow scenarios based on historical observations, and they further use a Monte Carlo method for scenario tree reduction. Jin et al. [41] perform scenario reduction based on historical data to represent long-term uncertainty

in wind profiles in a capacity expansion model. Park and Baldick [42] use a Gaussian copula method to generate scenarios based on historical load and wind data in a capacity expansion problem, and they use a fast backward method to reduce the scenario tree. Skar et al. [18] randomly sample historical electricity load and VRES normalized production profiles in the EMPIRE model, and they further perform a quality check on the scenario tree to ensure similar mean and variance as the historical data. Seljom and Tomasgard [14] perform random sampling of historic data related to VRES availability and electricity prices in the TIMES model, and they further sample many candidate scenario trees in [43] and pick the one which best matches the four first moments of all historical data. Seljom et al. [44] use similar historical sampling, and they also sample heat demand scenarios that are simulated using regression models based on temperature data.

The majority of the power and energy system modelling literature with stochastic programming do not present stability tests as part of the analyses. To the author's knowledge, the exception is Seljom and Tomasgard [14] who use 90 wind scenarios in a stability test with the TIMES model, and the same authors [45] perform more extensive in-sample and out-of-sample evaluations and compares stability with sample average approximation [46]. Although stability tests have been presented in previous research for similar models, e.g., Seljom and Tomasgard [14,45], we have not identified any research that use stability testing to compare different SGRs for the same model. Therefore, the novelty of this study and our contribution to the literature is to design and compare different SGRs for the same modelling framework. The SGRs that we develop and present in this paper can be replicated and applied in similar models.

In this paper, we evaluate SGR stability in energy system modelling with high shares of VRES, and we compare the performance of a statistical-based SGR and a measure-based SGR. In all three SGRs tested in this paper, representative time series are randomly sampled from historical data. The statistical-based SGR uses stratified sampling, where historical data are clustered into strata that are sampled from. The measure-based method uses moment matching, where scenario trees are produced by approximating moments of all historical data. Note the difference between the two SGR designs: In the former, statistical properties are used to modify the sampling itself, while in the latter, statistical properties are used to select scenario trees after the sampling is done. The EMPIRE model and the three SGRs tested in this paper are presented in the following section.

3. Method

This section presents the overarching structure of the EMPIRE model (Section 3.1), and the details related to the three SGRs that are used and compared in terms of stability (Section 3.2).

3.1. EMPIRE

The EMPIRE model is used to analyse the transition pathway towards a power system with large shares of VRES. An open version of EMPIRE can be downloaded from [47].

EMPIRE represents the power system by a network structure of nodes and arcs. The nodes represent markets, balancing demand and supply of electricity, and the arcs represent transmission exchange between these markets. EMPIRE optimizes capacity expansion of power system assets within nodes and arcs, including generation, storage, and transmission between markets, towards minimizing discounted investment costs and expected operational costs. EMPIRE considers asset lifetime, and initial asset capacity is defined in nodes and arcs. For further mathematical details on the objective function and the constraints in EMPIRE, please refer to Skar et al. [18].

EMPIRE considers scenarios for operation and market clearing in a linear two-stage stochastic program [7]: The first-stage decisions represent long-term investments, and the second-stage decisions represent operations in short-term time steps within multiple seasons and scenarios. To enable both multiple long-term periods for investments and multiple stochastic scenarios for operations while preserving computational tractability, EMPIRE is a multi-horizon stochastic program [48,18]. Because we do not consider long-term uncertainty, the only condition for using the multi-horizon approach in EMPIRE is that any realization of short-term uncertainty within one long-term period does not affect short-term uncertainty in later long-term periods. Correlations between short-term time steps within the same long-term period is included. We consider this reasonable as the outcomes of variables for stochastic wind or solar generation in time periods in 2020 do not influence the same generation in 2030 or 2050.

Fig. 1 illustrates a scenario tree in EMPIRE: a collection of outcomes over short-term time steps. Each scenario in the scenario tree contains a sequence of realizations of stochastic processes for short-term time steps. The operational time series are temporally and spatially correlated for each stochastic process, and cross-correlated across the stochastic processes. Each scenario consists of two categories of seasons: S regular seasons, e.g. winter or summer, and P peak seasons, e.g. high load situations. For each stochastic process, l^{reg} data points represent each regular season and l^{peak} data points represent each peak season. We refer to different outcomes of the same season as scenarios, and there are equally many scenarios for all seasons. The season-scenario structure is the same across all the long-term investment periods. We assume there is no correlation across seasons and long-term periods. The collection of time series representing the stochastic processes for all seasons, scenarios, and long-term periods is the scenario tree that is used as input to solve one instance of EMPIRE.

Scenarios contain unique realizations of stochastic processes for every short-term time step within every long-term period. VRES have scenario dependent availability in each short-term time step for all long-term periods. The stochastic scenarios include realizations of the following six stochastic processes:

- Solar generation,
- Onshore wind generation,
- Offshore wind generation,
- Hydro run-of-river generation,
- Regulated hydro generation,
- Electricity load.

For solar, wind, and hydro run-of-river, the stochastic scenarios consist of hourly capacity factors that express how much of the installed capacity is available for electricity production in a specific hour. For regulated hydro, each stochastic scenario contains seasonal potentials restricting how much electricity that can be produced in one week, which reflects different hydro inflows and reservoir levels.

Although the electricity load is a stochastic process in EMPIRE, the total electricity demand for a region is the same for all scenarios within each long-term period. Hourly electricity load is therefore *shifted* in every scenario such that the electricity demand for a market is the same for all scenarios in each long-term period, see [20] for details.

3.2. Scenario generation in EMPIRE

3.2.1. Basic SGR

The basic SGR in EMPIRE is probabilistic and sampling-based, and it starts with gathering historical data for the stochastic processes considered. Our sampling frame is the historical realizations of the stochastic processes listed above with hourly resolution from 2015–2019. The basic SGR is summarized in Fig. 2 and illustrated in Fig. 3.

To preserve autocorrelation in the basic SGR (Fig. 2), the chronology of all historical time series is respected. However, we partly break this correlation by sorting time series to always start a scenario on the same hour of the week. To preserve cross correlation, the same historic time periods are used for all stochastic processes and all regions in

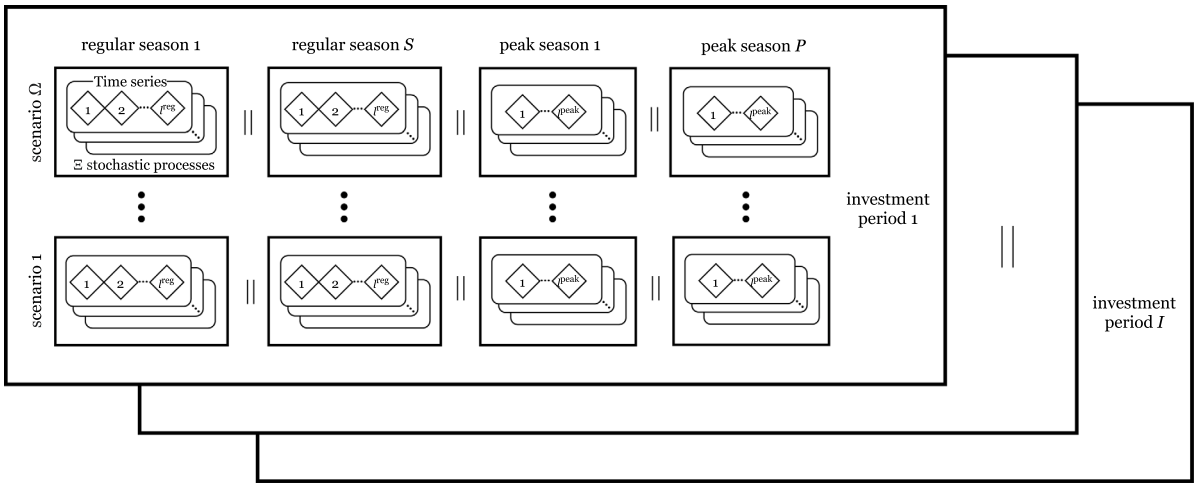


Fig. 1. Illustration of a scenario tree in EMPIRE: a collection of outcomes for short-term time steps uniquely defined as chronological time series by season, scenario, and long-term investment period.

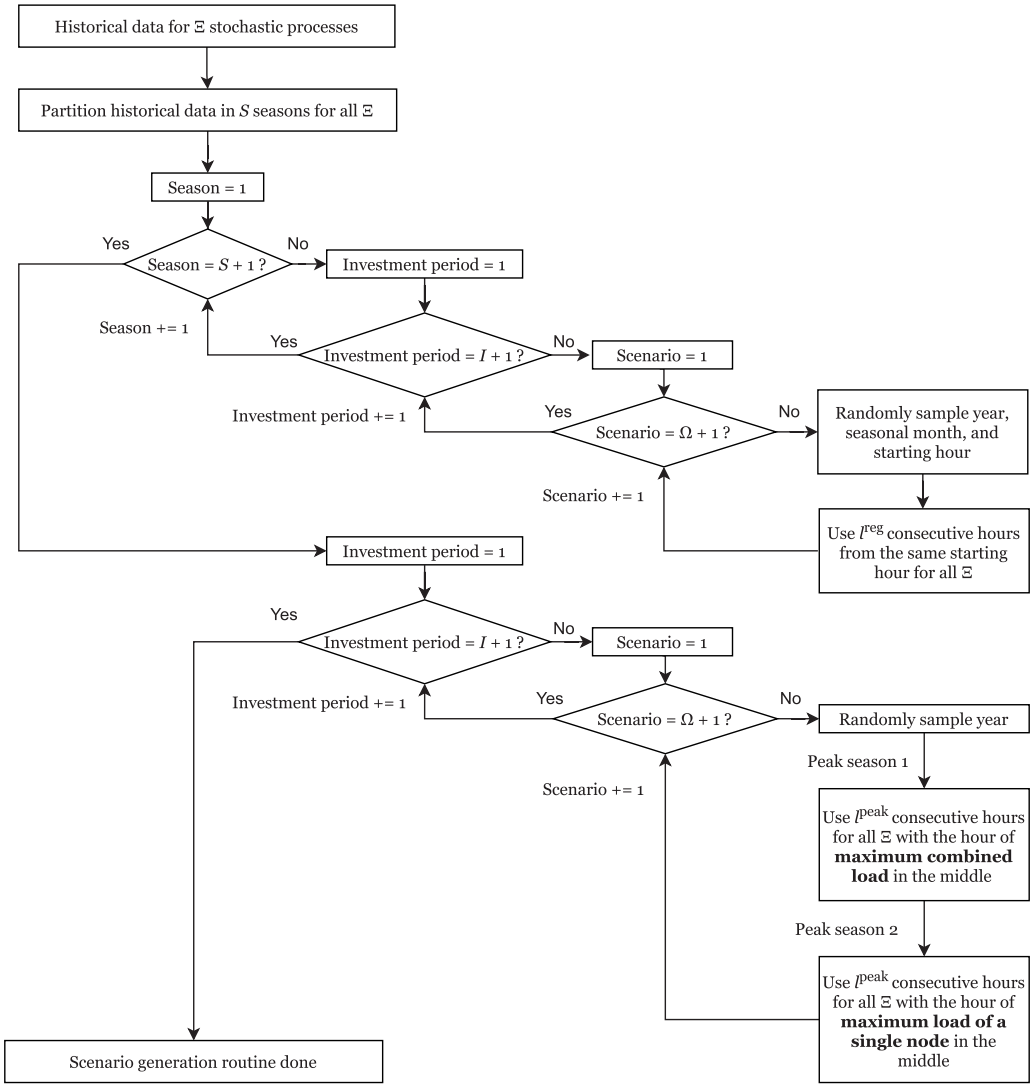


Fig. 2. Schematics of the basic scenario generation routine algorithm based on random sampling within each season in EMPIRE.

Algorithm Fig. 2. Note that unique scenarios are generated for each season and investment period in EMPIRE (Fig. 1), in line with the multi-horizon structure of EMPIRE assuming independent stochastic processes between the long-term periods.

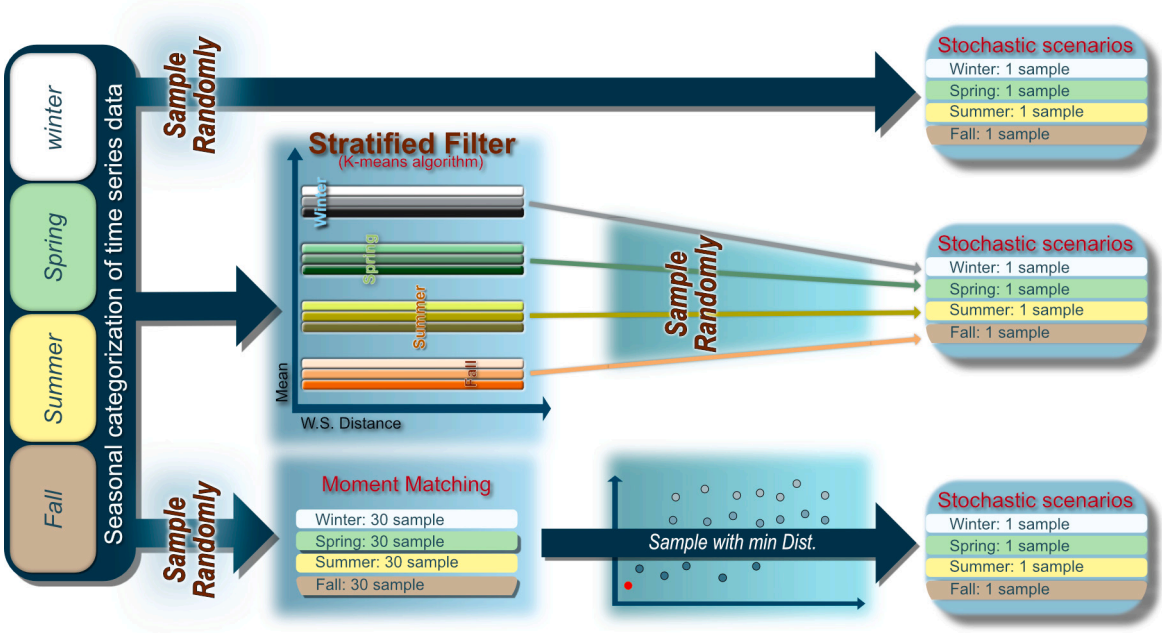


Fig. 3. Illustration of the three scenario generation routines used in this paper: The basic SGR (top), the strata SGR (middle), and the moment SGR (bottom).

We partition our sampling frame into seasonal sub-populations in the basic SGR (Fig. 2). Each seasonal sub-population contains data from the same months for all years, and we sample a random year, month, and starting hour from each season to ensure that all seasons are represented in our scenarios. This feature makes the basic SGR in EMPIRE a stratified sampling method [49]. Note that we assume seasonal independence.

In addition to four seasons, the basic SGR (Fig. 2) includes two peak load situations in every scenario. The first peak season adds the maximum aggregated electricity load situation for all regions for a randomly sampled year. The second peak season adds the maximum electricity load situation for a single region in the same year. Note that the peak seasons contain correlated data for all six stochastic processes, not just electricity load.

For each scenario, all regular seasons are scaled in the objective function of EMPIRE to represent a complete year minus the duration of the two peak seasons. The peak seasons are not scaled and thus represent $2l^{\text{peak}}$ hours of operation in EMPIRE.

In the following, we present two further developments of the basic SGR presented in Fig. 2, namely the strata SGR and the moment SGR (Fig. 3).

3.2.2. Strata SGR

The strata SGR, illustrated in the middle of Fig. 3, improves the representation of different electricity load realizations in the scenario tree by defining strata beyond the seasons. All seasonal sub-populations are further partitioned using k-means clustering [50] based on aggregated electricity load data. We then make sure that each scenario tree contains an equal number of scenarios from every cluster within each season. The strata SGR is an extension of the stratified sampling introduced in the basic SGR: We randomly sample scenarios for each season independently while ensuring that at least one scenario from each cluster is represented in each investment period in the scenario tree. The number of clusters therefore must be divisible by the number of scenarios per investment period. For a more detailed explanation of the strata SGR, see Appendix A.

3.2.3. Moment SGR

The moment SGR, illustrated in the bottom of Fig. 3, focuses on ensuring that electricity load in the scenarios statistically resembles historical electricity load data. Here, the basic SGR (Fig. 2) is run T times producing T scenario trees (Fig. 1). The moment SGR selects the scenario tree among the T candidates that has its first four moments most similar to all historical data. For a more detailed explanation of the moment SGR, see Appendix B.

4. Case study

To enable large scenario trees in EMPIRE when testing the SGRs presented in Section 3.2, we use an instance that is less computationally challenging than in [18–20]. The geographical scope covers Europe as illustrated in Fig. 4, however, we aggregate data from country nodes to represent three European regions, North (blue), West (green), and East (red). The aggregation procedure involves summing country specific input data for each region.

We consider three investment periods each covering ten years representing 2020–2030, 2030–2040, and 2040–2050. The CO2eq. cap is defined for the power sector in total to be fixed per investment period to 1110, 770, and 66 Mton CO2eq. per year, respectively, following [24]. The CO2eq. emissions are assumed to be direct emissions from operations only, and they are estimated for each technology according to [51]. We further assume no CO2eq. emissions from renewable sources including biomass. EMPIRE allows load shedding at the cost of EUR 22,000/MWh following [52]. Short-term operations consist of four regular seasons spanning weeks, each with 168 consecutive hours, and two peak seasons spanning days, each with 24 consecutive hours. We consider 16 generator types and two storage types presented in Fig. 4 with costs according to [53] and initial capacity according to [54]. Fuel costs are according to [1], while transmission expansion costs are according to [18].

As input to the three SGRs, we use hourly historical electricity load data and hydro power production from the ENTSO-E Transparency Platform [55] for the years 2015–2019. Hourly solar availability data for the same years are gathered from simulations based on satellite data [56]. Hourly wind availability come from [57]. The electricity load data in the raw scenario trees are scaled in EMPIRE to match future annual demand projections towards 2050 according to [1]. For solar

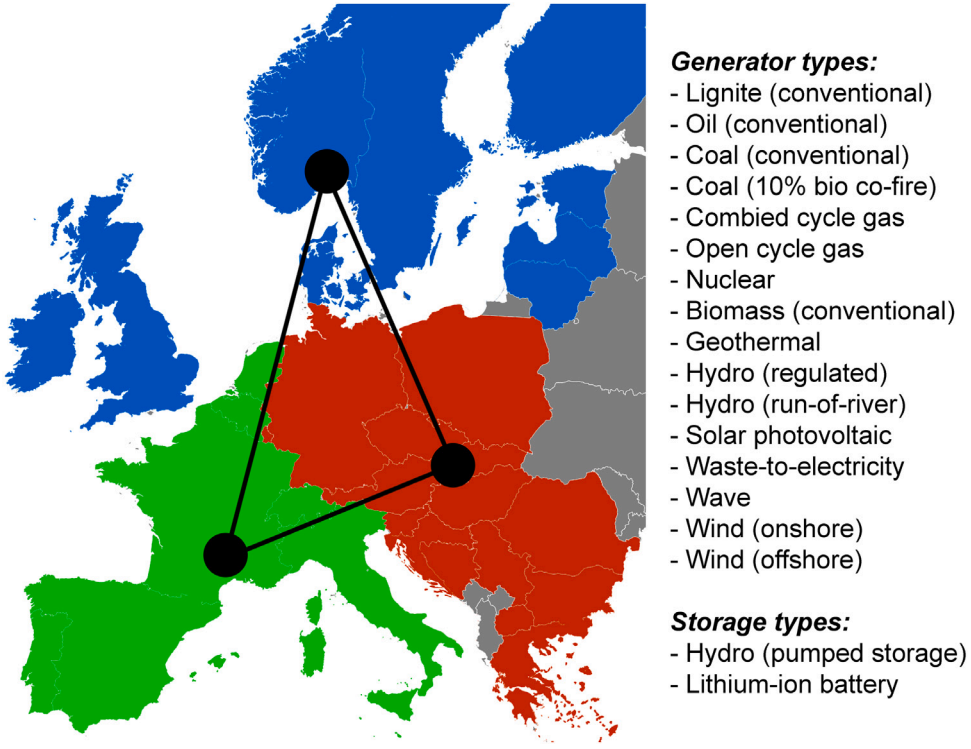


Fig. 4. Technology options in our EMPIRE instances and illustration of the aggregated European regions: North (blue), West (green), and East (Red). Grey countries are not considered.

and wind, we consider representative countries per region such that data from Great Britain, France, and Germany is used for the North, West, and East regions, respectively.

In our case, the 'true' problem is out of reach for two reasons: (1) it relies on data about the long-term future in which the properties of the stochastic processes are uncertain and (2) it is numerically infeasible to represent the true underlying distributions as discrete scenarios. As suggested in [23], we therefore approximate the 'true' problem by solving EMPIRE with a large *reference tree* containing as many realizations of the stochastic processes as possible. Our reference tree is generated using the basic SGR to minimize biases, and it contains 600 scenarios per investment period. With 4+2 seasons, 600 scenarios, and 3 investment periods, the reference tree represents about a quarter of all possible starting hours from the five years of data. The same reference tree is used for all SGRs and instances in the out-of-sample tests.

EMPIRE is implemented in Pyomo [58,59], and all SGRs are implemented in Python. For each SGR, we solve 30×3 instances of EMPIRE with different scenario trees containing 10, 50, and 100 scenarios per investment period. With four regular representative weeks and two peak days, we have $168 \times 4 + 24 \times 2 = 720$ representative hours per scenario, which means the whole planning horizon in EMPIRE contains 21,600 h with 10 scenarios, 108,000 h with 50 scenarios, and 216,000 h with 100 scenarios, and $1,296,000$ hours for the reference tree with 600 scenarios. The actual number of hours in the 30 year planning horizon is 263,000 h.

In total, we solve 270×2 instances for the in-sample and out-of-sample test. We assume equal probability for all scenarios within each scenario tree. For the strata SGR, we consider $k = 10$ clusters for all scenario trees, such that each strata is represented 3, 15, and 30 times per season with 10, 50, and 100 scenarios per investment period, respectively. For the moment SGR, we generate $T = 20$ potential trees for each of the 90 instances and use the scenario tree with the smallest relative four-moment distance to the whole sampling frame as defined in (3).

Table 1
Objective function values for stochastic and deterministic instances of EMPIRE, as well as the value of stochastic solution (VSS), for a scenario tree with 200 scenarios.

Problem	Objective [bn EUR]
EV	1142
EEV	16,650
SP	1366
VSS	15,264

5. Results

5.1. Value of stochastic solution (VSS)

We evaluate the value of the stochastic solution (VSS) for a scenario tree with 200 scenarios produced with the basic SGR. The expected value problem (EV) is solved first, which is done by replacing the 200 scenarios with a single scenario consisting of the mean of all realizations of the random variables across the 200 scenarios. After solving the EV, we solve the second stage given the EV's first stage decisions in all 200 scenarios, and the mean of the resulting objective function values is the expected result of using the expected value solution (EEV). The difference between the EEV and the objective function value of the stochastic problem (SP) is the VSS.

Results are presented in Table 1. The VSS is an order of magnitude larger than the SP value. This is because the first stage decisions of the EV are not able to fully supply electricity demand in all hours of the 200 scenarios, which leads to very costly load shedding.

Fig. 5 illustrates the difference between the EEV and the SP in terms of installed generation capacity and expected annual generation in the three investment periods, and Table 2 presents the installed capacity of each technology in the EEV and the relative difference in the SP. Technologies without difference in capacity between the EEV and the SP are not presented in Table 2. The share of VRES grows towards 2050, driven by assumed cost reductions in line with De Vita et al. [53] and emission constraints in line with the European Commission [24], and

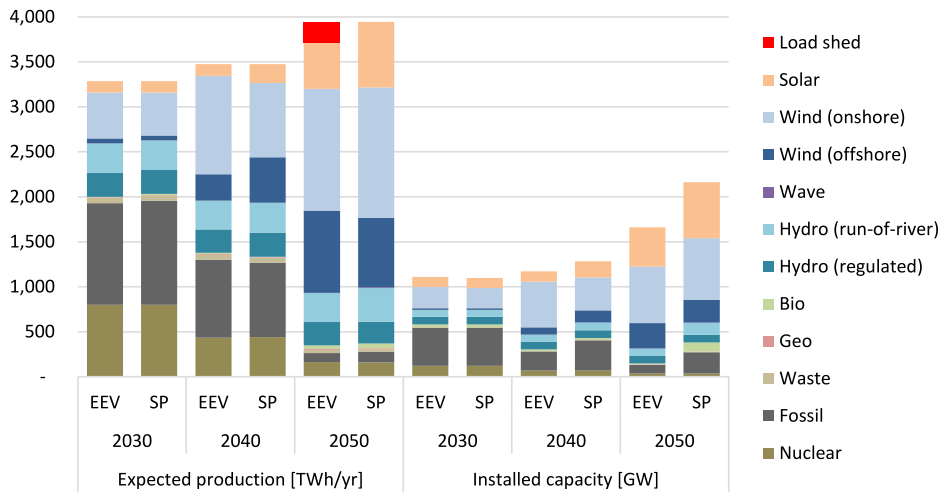


Fig. 5. Expected annual production (TWh/yr) and installed capacity (GW) for Europe as a whole for the expected result of using the expected value solution (EEV) and the recourse problem (SP) for a scenario tree with 200 scenarios.

Table 2
Installed generation capacity for Europe in GW for selected technologies in all investment periods for the EEV and the relative difference in the SP.

Investment period	2020–2030		2030–2040		2040–2050	
Technology	EEV	SP diff.	EEV	SP diff.	EEV	SP diff.
Fossil	420 GW	0%	210 GW	+60%	95 GW	+149%
Bio	31 GW	0%	17 GW	0%	9 GW	+1049%
Hydro (run-of-river)	78 GW	0%	78 GW	+13%	85 GW	+58%
Wind (offshore)	17 GW	0%	81 GW	+65%	280 GW	−10%
Wind (onshore)	237 GW	−5%	506 GW	−28%	629 GW	+9%
Solar	112 GW	0%	114 GW	+56%	436 GW	+43%

the share of expected production by VRES seems to be the main reason for the differences between the EEV and the SP.

Between 2020 and 2030, there are similar results for the EEV and the SP in terms of installed capacity and expected production, and the share of wind and solar is around 20% of the European mix. The similarity between the EEV and the SP in 2030 is because the flexibility from dispatchable generators in the EEV is able to handle load and VRES variations for the 200 stochastic scenarios for all regular and peak seasons. The only difference is 5% less onshore wind capacity in the SP compared to the EEV, which means that the EEV slightly overestimates the cost-optimal amount of onshore wind capacity compared to the SP. In 2030, there is 3% less VRES capacity and 3% less expected VRES production in the SP compared to the EEV.

Between 2030 and 2040, the expected share of wind and solar production is around 40% of the European mix, and the installed capacity of dispatchable generators in the EEV is *underestimated* compared to the SP, especially capacity by fossil, wind offshore, and solar. As opposed to all other technologies, the installed capacity of onshore wind is *overestimated* by 28% in the EEV compared to the SP, which is a larger overestimation than the previous period. In total, there is less onshore and offshore wind capacity in the SP compared to the EEV until 2040, which is aligned with the findings of Seljom and Tomasgard [14] and Pineda and Morales [15], who present a similar comparison between deterministic and stochastic modelling in Danish case studies. Ringkjøb et al. [16] also find that the investments in VRES are higher in the deterministic than the stochastic versions of the TIMES model in a European case study. Although there is 2% less VRES capacity in 2040 in the SP, our case study shows that there is 2% more expected VRES production in the SP compared to the EEV, which is in contrast with [14] and [16]. We find that between 2030 and 2040, there is higher capacity utilization in the SP compared to the EEV, especially for solar and offshore wind (Fig. 5), which also leads to 44% less expected curtailment of surplus VRES.

Between 2040 and 2050, the expected share of wind and solar production is around 70% of the European mix, and the EEV leads to 3% (more than 200 TWh/yr) of expected load shedding (blackouts), which would be unacceptable in practice. The most significant difference is over 10 times more dispatchable bio capacity and 149% more dispatchable fossil capacity in the SP compared to the EEV, which reflects a strong underestimation of the need for dispatchable capacity in the EEV. More installed capacity of dispatchable generators in the SP compared to the EEV does not imply significantly more expected production (Fig. 5), which means that the capacity factor for dispatchable generators are significantly reduced in the SP compared to the EEV. Similar results are shown by Brouwer et al. [60], who suggest that future power markets need revised market designs to ensure system adequacy. Compared to the previous investment period, underestimation of capacity from fossil, bio, run-of-river hydro, and onshore wind is increased. As opposed to the previous periods, offshore wind, rather than onshore wind, is overestimated in the EEV compared to the SP, but the total onshore and offshore wind capacity is higher in the SP compared to the EEV. These findings are in contrast to the findings of Seljom and Tomasgard [14], Pineda and Morales [15], and Ringkjøb et al. [16] who find that there is less VRES in the stochastic model compared to the deterministic model. The conflicting results from [14] could be because of the strict emission constraints after 2040, which results in more than 80% expected production by VRES, i.e., wind, solar, and hydro run-of-river, in our case study. The share of VRES in [16] is 62% in 2050. As wind power is the most cost-efficient solution to respect emission constraints in our study (Fig. 5), and because the emission constraint excludes fossil alternatives, the SP results in 18% more installed VRES capacity than the EEV (Table 2). In 2050, there is 7% more expected production from VRES in the SP compared to the EEV, which also results in 31% higher expected curtailment of surplus VRES in the SP compared to the EEV. Note that our case study does not consider utilization of surplus VRES, e.g., green hydrogen production.

Table 3
Results from in-sample stability test solving 30 instances for each routine and scenario tree size.

Tree size	Routine	Mean (μ)	SD (σ)	Relative SD (σ/μ)
10	Basic	1.35E+12	2.11E+10	1.56%
	Strata	1.37E+12	2.24E+10	1.63%
	Moment	1.35E+12	2.59E+10	1.92%
50	Basic	1.36E+12	1.04E+10	0.76%
	Strata	1.39E+12	1.07E+10	0.78%
	Moment	1.36E+12	1.10E+10	0.81%
100	Basic	1.36E+12	8.19E+09	0.60%
	Strata	1.39E+12	8.68E+09	0.63%
	Moment	1.37E+12	7.26E+09	0.53%

Table 4
Results from out-of-sample stability test solving 30 instances for each routine and scenario tree size. The reference tree contains 600 scenarios.

Tree size	Routine	Mean (μ)	SD (σ)	Relative SD (σ/μ)
10	Basic	1.91E+12	4.99E+11	26.16%
	Strata	1.58E+12	3.79E+11	23.93%
	Moment	1.87E+12	5.56E+11	29.77%
50	Basic	1.41E+12	2.56E+10	1.82%
	Strata	1.39E+12	1.88E+10	1.35%
	Moment	1.42E+12	5.64E+10	3.98%
100	Basic	1.40E+12	2.59E+10	1.86%
	Strata	1.38E+12	2.52E+09	0.18%
	Moment	1.39E+12	1.24E+10	0.89%

There is no capacity expansion of electricity storage in the EV, as opposed to the SP expanding 25 GW charging/discharging capacity and 19 GWh energy storage capacity. Electricity storage is used for load shifting to balance supply and demand in the different scenarios in the SP, while the EV only plans for one realization of load and VRES availability not resulting in any expansion of load shifting capacity.

There is no transmission expansion in neither the SP nor the EV. Expected transmission exchange is 9% higher in the EEV compared to the SP. More utilization of transmission in the EEV is due to a general lack of generation and storage capacity because capacity expansion is done with a narrower scope of different load and VRES situations.

Although the SP favours more VRES surplus and less utilization of cross-border transmission than the EEV, more generation and storage capacity in the SP compared to the EEV means that the very high cost of 3% expected load shedding is avoided. Therefore, the SP does not overestimate the need for generation and storage capacity because the alternative is very costly.

In short, stochastic modelling of the European power system at minimized cost and minimized blackouts, compared to deterministic modelling, comes with the consequences of: more capacity investments in VRES, storage, and dispatchable generators; more VRES surplus; and less utilization of cross-border transmission.

5.2. Stability

Figs. 6 and 7 illustrate the stability test results in box plots. Fig. 6 shows the results from 10 scenarios where the in-sample results are significantly more stable than the out-of-sample results. Fig. 7 shows the results from 50 and 100 scenarios where the out-of-sample results have generally more outliers and more expensive instances than the in-sample results.

Table 3 presents mean and the standard deviation (SD) of the objective function values from the in-sample stability test. The expected objective function value increases and the SD decreases with an increasing size of the scenario tree for all methods. The in-sample stability improvement is larger going from 10 to 50 scenarios than going from 50 to 100 scenarios.

All three methods are relatively in-sample stable with $< 2\%$ relative SD, and the in-sample stability performance is similar for all three methods for each tree size (see Table 3). The strata SGR has a significantly larger expected objective function value compared to the basic and the moment SGR, which indicates that the strata SGR represents more expensive situations. Note that we assume that all scenarios have equal probability in all three SGRs. Especially for the strata SGR, the scenario probability could in addition be tuned for each strata as certain electricity load situations are more or less likely to realize compared to others.

Table 4 presents results from the out-of-sample stability test. Similar to the in-sample test, we observe a decrease in the SD with an increasing size of the scenario tree for all methods. Contrary to the in-sample test, the expected objective function value decreases with larger scenario trees, which is because the first stage decisions are generally performing better in random scenarios when more scenarios are considered in the optimization. There is a large decrease in the SD going from 10 to 50 scenarios, and a smaller decrease going from 50 to 100 scenarios.

The out-of-sample test shows a larger difference between the three methods than the in-sample test. With 10 scenarios, the strata SGR has a significantly lower expected objective function value compared to the basic and the moment SGR. The strata SGR also has the lowest expected objective function value for all tree sizes comparing to the other SGRs, but the difference is less significant for 50 and 100 scenarios.

The largest difference between the three methods is the out-of-sample stability with 10 scenarios, where the strata SGR is most stable (Fig. 6). To study which investment decisions that result in more out-of-sample stability, Fig. 8 shows a box plot for the installed VRES capacity with 10 scenarios in 30 instances for each SGR. For investments in hydro run-of-river, there are small differences between the three SGRs.

Investments in solar and wind capacity are significantly different between the strata SGR and the other two methods (Fig. 8). For onshore wind, the basic SGR results in 2%–6% more investments on average than the strata SGR, while the moment SGR results in 1%–3% more investments on average. For offshore wind, the contrary is the case after 2030: The basic SGR results in 14%–15% less investments on average than the strata SGR, while the moment SGR results in 9%–13% less investments on average. For solar capacity after 2030, the basic and moment SGR results in 6%–27% and 8%–18% more investments on average than the strata SGR, respectively. In short, investments in solar and wind is correlated with stability: less investments in onshore wind and solar, and more investments in offshore wind, results in more out-of-sample stability, which means an improved ability for the power system to adapt to different stochastic scenarios than the ones that are considered in the optimization.

Fig. 9 shows a box plot for the installed dispatchable thermal capacity with fossil fuels and biomass. For both dispatchable generator types, there is less than 1% difference before 2040. After 2040, the basic and moment SGR result in 8% and 6% more fossil capacity on average than the strata SGR, respectively. The contrary is true for bio capacity: the basic and moment SGR result in 17% and 20% less bio capacity on average than the strata SGR, respectively. The total installed dispatchable capacity for both bio and fossil after 2040 is 2% and 1% more on average for the basic and moment SGR than for the strata SGR. Slightly less dispatchable capacity in the strata SGR compared to the other two methods is linked to thermal bio capacity and the emission constraints: when a higher share of the dispatchable capacity does not produce accountable carbon emissions, namely bio capacity, the installed capacity can be better utilized such that less total dispatchable capacity is needed in the strata SGR compared to the other two routines.

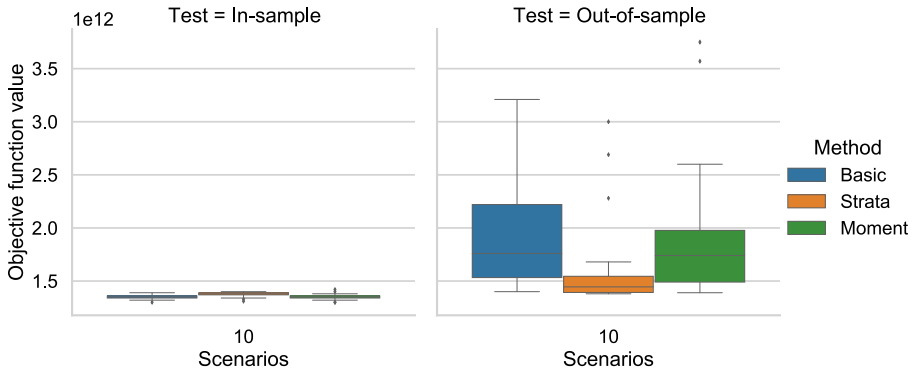


Fig. 6. Box plot of objective function values for 30 instances with 10 scenarios per tree by method and test.

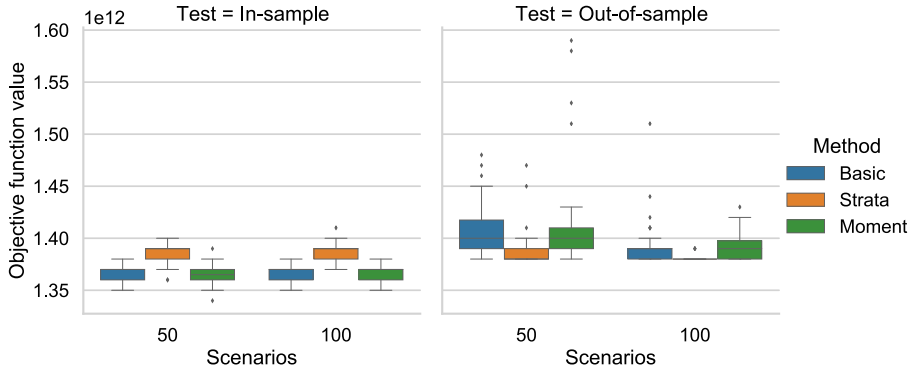


Fig. 7. Box plot of objective function values for 30 instances with 50 and 100 scenarios per tree by method and test.

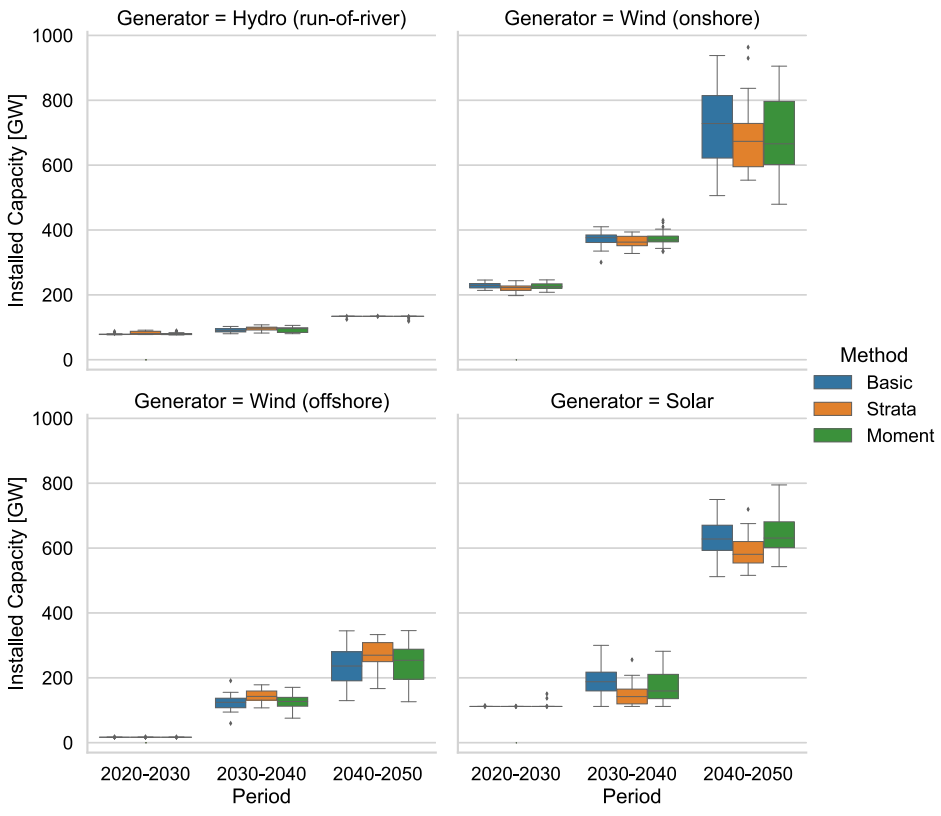


Fig. 8. Box plot of installed VRES capacity for 30 instances with 10 scenarios per tree by method and investment period.

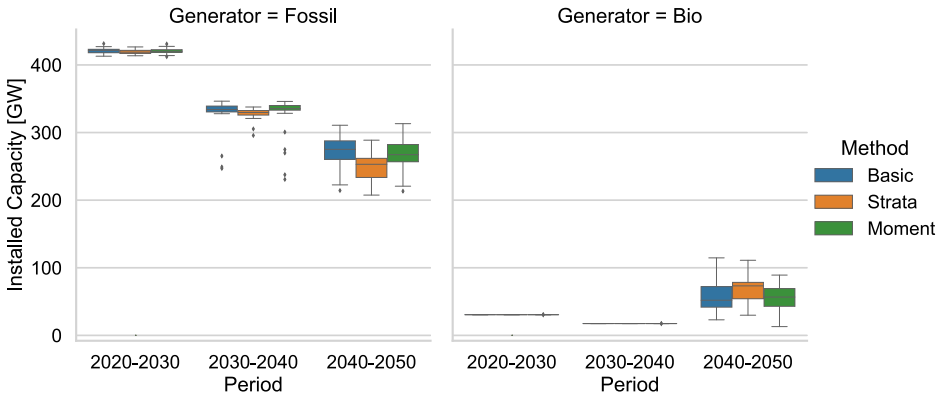


Fig. 9. Box plot of installed thermal capacity (fossil and bio) with 10 scenarios per tree by method and investment period.

Table 5
Computation times to generate scenario trees with the three SGRs with 10, 50, and 100 scenarios.

Routine	Pre-processing	10 scenarios	50 scenarios	100 scenarios
Basic	–	9 s	81 s	253 s
Strata	347 s	9 s	83 s	255 s
Moment	–	161 s	1485 s	4690 s

Table 6
Summary of the problem size, RAM requirement, and solution times for instances with 1, 10, 50, and 100 scenarios.

# of scenarios	1	10	50	100
Constraints (rows)	366,994	3,663,559	18,314,959	36,629,209
Variables (columns)	225,193	2,246,953	11,232,553	22,464,553
Max RAM use	10 Gb	25 Gb	111 Gb	210 Gb
Solution time	9–16 s	228–555 s	1734–3204 s	5007–11, 840 s

5.3. Computational effort

All instances are solved using interior point method (barrier algorithm) [61] without crossover with the Gurobi Solver v9.0.2 [62] running on a computer cluster with CPU 2× 3.5 GHz Intel Xeon Gold 6144 CPU (8 core) and 384 Gb RAM.

Table 5 presents the computation times to generate scenario trees of different sizes for the three SGRs. The basic SGR produces one scenario tree (Fig. 1) in 9 s for 10 scenarios, 81 s for 50 scenarios, and 253 s for 100 scenarios. Recall that 10, 50, or 100 scenarios are generated for three investment periods in one scenario tree. The strata SGR produces scenario trees in the same time as the basic SGR, however, it requires the stratified filter as an input (Appendix A), which takes 347 s to produce. Because the sampling frame does not change throughout this study, the stratified filter only needs to be generated once. The moment SGR is the most expensive routine in terms of time because it produces 20 candidate trees before selecting the one with the smallest relative four-moment distance to the whole sampling frame (Appendix B). Thus, the moment SGR uses 18 times longer to produce one scenario tree compared to the basic SGR.

Table 6 summarizes the problem size, random access memory (RAM) requirement, and solution times for instances with 1, 10, 50, and 100 scenarios. Note that the solution time indicates how long it takes to find the optimal solution after the operational scenario(s) have been generated and the linear program is ready to solve. Because instances differ in terms of their stochastic scenarios, computational time also differs. On average, instances with scenario trees containing 1 scenario solve in 12 s, 10 scenarios solve in 326 s, 50 scenarios in 2394 s, and 100 scenarios in 7470 s. These results indicate that the solution time grows polynomially with the increased number of constraints and variables with more stochastic scenarios in line with the famous findings of Karmarkar [61].

The reason for the high RAM requirement presented in Table 6 is mainly because of large instance sizes with several million constraints and variables. Because available computational power is sufficient for this case study, we have not investigated potential reductions of maximum RAM requirement for this study.

5.4. Discussion

Our results indicate that 10 scenarios are too few to reach out-of-sample stability with any of the three SGRs. The stability improvement when increasing the scenario tree size from 10 to 50 scenarios is significantly larger than from 50 to 100 scenarios, which indicates diminishing stability improvement when growing the tree beyond 50 scenarios. This result is especially important for larger instances as large scenario trees increase the computational challenge (Table 6).

The strata SGR seems to provide solutions of higher quality and stability compared to the basic and moment SGR, which indicates that ensuring the representation of certain electricity load situations is preferable over the moment matching approach. The strata SGR is also computationally cheaper when generating stochastic scenarios than the moment SGR (Section 5.3). Although the strata SGR performs best in our case study, this is not a general conclusion because performance depends both on the mathematical program and the design of the SGR [23]. Future development of the moment SGR could be to generate trees that minimize the relative four-moment distance among all possible trees rather than within a random subset of trees. Performing scenario reduction on a large tree produced by such a moment SGR can be explored.

Ensuring the representation of critical stochastic realizations is worth exploring further in energy system modelling. For example, further work could integrate risk measures from finance in energy system modelling, like conditional value-at-risk [63,64], e.g., Yu and Zheng [65]. As the strata SGR performs best in our case study, it is worth expanding on this method towards other stochastic input data than electricity load, namely ensuring the representation of certain VRES situations. More specifically, stability could be further improved by guaranteeing that samples represent a spectrum of high to low availability of VRES. This could be done for solar, wind, and hydro sources separately, or for all VRES combined. Excess or deficient VRES situations could also be added to the basic SGR as additional peak seasons, since the two current peak seasons only represent peak electricity load. Note that because capacity expansion decisions for VRES are endogenous in EMPIRE, ensuring VRES situations must be done with caution to avoid the model becoming biased towards high or low availability of VRES. Further development of the SGRs in EMPIRE therefore includes improving scenario probability weights in the objective function.

Even though an SGR for EMPIRE is tested to produce good stability, the resulting long-term power system developments are necessarily

biased towards the input data to the SGRs, to EMPIRE, and to how the input data is processed to represent the future. We assume that our historical data input to the SGRs contains all possible realizations of all the stochastic processes, which is probably wrong. Additionally, with the long-term horizon, we are not just dealing with uncertainty related to the stochastic processes themselves, but deeper uncertainty related to the stationarity of the stochastic processes. Future work include exploring stationarity change as wide electrification affect future electricity load profiles and changing weather patterns due to climate change affect both VRES availability and energy demand, especially heating buildings in colder climates. Solaun and Cerdá [66] show that an increasing amount of research is exploring how climate change is impacting renewable energy potentials, especially focusing on Europe. Hamududu and Killingtveit [67] find that European hydro power potential could increase by 1.46% in the north and decrease by 1.82% in the south by 2050. Devis et al. [68] find that wind power output in Europe could change from −12% to +8% by 2050 with the largest decrease in the south, while Jerez et al. [69] find that solar power output in Europe could change from −14% to +2% by 2100 with the largest decrease in the north. It is unclear how such changes will impact EMPIRE results, and we leave it up for future work to explore evolving stationarity for the stochastic processes of electricity loads and short-term VRES availability.

6. Conclusion

This paper presents the impact of different scenario generation routines on results of the multi-horizon power system capacity expansion model EMPIRE. Firstly, we evaluate the value of the multi-horizon structure compared to a deterministic model with high shares of variable renewable energy sources. Secondly, we present three probabilistic scenario generation routines based on randomly sampling historical realizations of cross-correlated stochastic processes, and we evaluate the stability of the scenario generation routines in a case study.

We conclude that the multi-horizon structure outperforms deterministic modelling when modelling power system capacity expansion with high shares of variable renewable energy sources. Our results are in line with results from similar research, namely that deterministic modelling leads to *overestimation* of variable renewable energy sources and *underestimation* of flexible generators and storage compared to stochastic modelling. Further, we contrast findings by similar research when the share of variable renewable energy sources is greater than 80%, where we find that the need for both variable renewable energy sources and flexible generators and storage are *underestimated* with deterministic modelling compared to stochastic modelling. This has important implications for energy planning, as very high shares of variable renewable energy sources will potentially require more capacity than assumed by deterministic models, which will lead to more renewable energy surplus. Although more computational effort is needed, stochastic modelling is therefore recommended for long-term power system capacity expansion planning with high shares of variable renewable energy sources.

Regarding stability, we conclude that more than 10 scenarios are needed for out-of-sample stability for all scenario generation routines in our case study, but that having more than 50 scenarios provides limited improvement on stability. The most stable scenario generation routine in our case study is based on stratified sampling and guarantees the representation of different electricity load situations. Stratified sampling has high potential for further development and requires limited computational effort. Note that conclusions regarding stability are specific to the EMPIRE model and our case study, and similar stability tests ought to be done for other scenario generation routines, models, and/or case studies to ensure stable results.

CRediT authorship contribution statement

Stian Backe: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Mohammadreza Ahang:** Conceptualization, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Asgeir Tomasgard:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

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.

Acknowledgements

We gratefully acknowledge the OpenENTRANCE project which has received funding from the Horizon 2020 research and innovation program under Grant Agreement No. 824260. We also acknowledge the support from the Research Council of Norway through the Research Centre on Zero Emission Neighbourhoods in Smart Cities (FME ZEN) under Grant No. 257660, including support from the centre partners, and through the "Assessment of the value of flexibility services from the Norwegian energy system" under Grant No. 268097.

Appendix A. The strata SGR

In the following, we elaborate on the algorithmic details of the strata SGR that is introduced in Section 3.2.2.

Consider all possible starting hours from the basic SGR presented in Fig. 2 for every season. For every possible scenario j in season s, we take the l^{reg} consecutive hours starting on hour h^{reg}, and let the realizations of the aggregated load over all nodes represent the univariate and uniform distribution $r_{j,s}$. Note that the distribution $r_{j,s}$ spans l^{reg} aggregated load realizations within a single scenario, independent of time of day.

For the distribution $r_{j,s}$, we calculate the mean load, $\mu^{\text{load}}_{j,s}$, over the l^{reg} hours:

$$\mu^{\text{load}}_{j,s} = \frac{1}{l^{\text{reg}}} \sum_{t=1}^{l^{\text{reg}}} \sum_{n \in \mathcal{N}} x^{\text{load}}_{t,n,j,s}, \qquad s \in S,\ j \in J, \tag{1}$$

where $x^{\text{load}}_{t,n,j,s}$ represents the historical electricity load in the tth hour of scenario j, node n, and season s. We use $\mu^{\text{load}}_{j,s}$ to identify whether scenario j generally represents a high or low load situation in season s.

Further, for each season s, we now define the distribution R_s, which is spanning all historical hourly aggregated electricity loads in the whole sampling frame. We then calculate the Wasserstein distance [70], $W_1(r_{j,s}, R_s)$, between $r_{j,s}$ and R_s defined as:

$$W_1(r_{j,s}, R_s) = \inf_{\pi \in \Gamma(r_{j,s}, R_s)} \int_{\mathbb{R}\times\mathbb{R}} |a-b| d\pi(a,b), \qquad s \in S,\ j \in J, \tag{2}$$

where a and b are load measurements within $r_{j,s}$ and R_s, respectively, and $\Gamma(r_{j,s}, R_s)$ is the set of all possible joint probability distributions on $\mathbb{R} \times \mathbb{R}$ whose marginals are $r_{j,s}$ and R_s on the first and second factors, respectively. Pflug [71] also applies the Wasserstein distance in the context of scenario generation. If $W_1(r_{j,s}, R_s)$ is small, it means that there is some similarity between the loads in the scenario and the sub-population of the sampling frame representing season s. If $W_1(r_{j,s}, R_s)$ is large, the scenario represents a load situation that does not occur very often.

With the two measures, $\mu^{\text{load}}_{j,s}$ and $W_1(r_{j,s}, R_s)$, for every possible scenario j in season s, we apply two-dimensional k-means clustering to each seasonal sub-population of the sampling frame. An example of

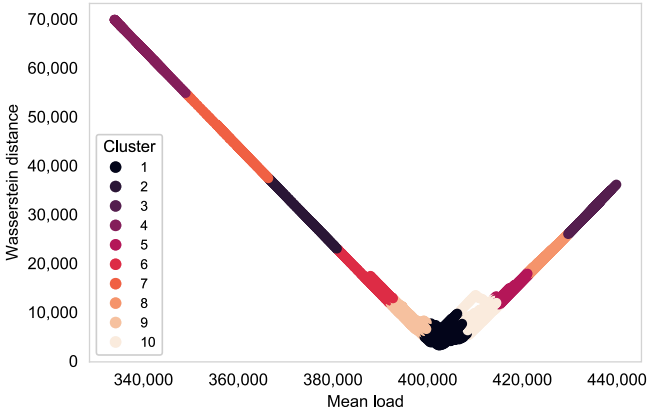


Fig. 10. Illustration of k-means clustering with 10 clusters by aggregated electricity load of all possible scenarios from a season.

the k-means clustering is illustrated in Fig. 10. The k-means clustering produces k clusters for every season s, where the scenarios within one cluster have similar seasonal load mean and similar Wasserstein distance to R_s. Note that data within a cluster are from different years, as we assume no correlations between scenarios.

Appendix B. The moment SGR

In the following, we elaborate on the algorithmic details of the moment SGR that is introduced in Section 3.2.3.

We denote the uniform and univariate distribution $\rho_{n,s,\tau}$ where the support consists of all realizations of electricity loads in node n, season s, and scenario tree τ. For every node n, season s, and candidate scenario tree τ, we calculate the first four moments of $\rho_{n,s,\tau}$: the mean ($m^1_{\tau,n,s}$), the variance ($m^2_{\tau,n,s}$), the skewness ($m^3_{\tau,n,s}$), and the kurtosis ($m^4_{\tau,n,s}$). Further, we denote the uniform and univariate distribution $P_{n,s}$ where the support consists of all realizations of electricity loads in node n and season s in the whole sampling frame, and let the ith moment of $P_{n,s}$ be defined as $M^i_{n,s}$.

We denote the relative four-moment distance $\Delta M(\tau)$ of a candidate scenario tree τ as the following:

$$\Delta M(\tau) = \sum_{n\in\mathcal{N}}\sum_{s\in S} w_{n,s} \sum_{i=1}^{4} \frac{|m^i_{\tau,n,s} - M^i_{n,s}|}{M^i_{n,s}}, \tag{3}$$

where the scenario tree independent weight, $w_{n,s}$, indicate node n's share of aggregated load for all nodes in season s, and it is defined as the following:

$$w_{n,s} = \frac{M^1_{n,s}}{\sum_{n'\in\mathcal{N}} M^1_{n',s}},$$

where $M^1_{n,s}$ is the mean of $P_{n,s}$ for node n and season s.

We calculate $\Delta M(\tau)$ for all T candidate scenario trees. The moment SGR selects the scenario tree, τ^*, that has the smallest $\Delta M(\tau)$ of all the T candidate trees.

References

[1] Capros Pantelis, Kannavou Maria, Evangelopoulou Stavroula, Petropoulos Apostolos, Siskos Pelopidas, Tasios Nikolaos, et al. Outlook of the EU energy system up to 2050: The case of scenarios prepared for European Commission's "Clean energy for all Europeans" package using the PRIMES model. Energy Strategy Rev 2018;22:255–63. http://dx.doi.org/10.1016/j.esr.2018.06.009.

[2] Auer Hans, del Granado Pedro Crespo, Oei Pao-Yu, Hainsch Karlo, Löffler Konstantin, Burandt Thorsten, et al. Development and modelling of different decarbonization scenarios of the European energy system until 2050 as a contribution to achieving the ambitious 1.5 C climate target—establishment of open source/data modelling in the European H2020 project openENTRANCE. Elektrotech Inftech 2020;1–13. http://dx.doi.org/10.1007/s00502-020-00832-7.

[3] Holttinen Hannele. The impact of large scale wind power production on the Nordic electricity system. 2004, URL https://www.vttresearch.com/sites/default/files/pdf/publications/2004/P554.pdf.

[4] Nicolosi Marco, Mills Andrew, Wiser Ryan. The importance of high temporal resolution in modeling renewable energy penetration scenarios. 2010, https://escholarship.org/uc/item/9rh9v9t4. [Accessed 30 July 2021].

[5] Collins Seán, Deane John Paul, Poncelet Kris, Panos Evangelos, Pietzcker Robert C, Delarue Erik, et al. Integrating short term variations of the power system into integrated energy system models: A methodological review. Renew Sustain Energy Rev 2017;76:839–56. http://dx.doi.org/10.1016/j.rser.2017.03.090.

[6] Ringkjøb Hans-Kristian, Haugan Peter M, Solbrekke Ida Marie. A review of modelling tools for energy and electricity systems with large shares of variable renewables. Renew Sustain Energy Rev 2018;96:440–59. http://dx.doi.org/10.1016/j.rser.2018.08.002.

[7] Kall Peter, Wallace Stein W. Stochastic programming. Springer; 1994.

[8] Defourny Boris, Ernst Damien, Wehenkel Louis. Multistage stochastic programming: A scenario tree based approach to planning under uncertainty. In: Decision theory models for applications in artificial intelligence: Concepts and solutions. IGI Global; 2012, p. 97–143. http://dx.doi.org/10.4018/978-1-60960-165-2.ch006.

[9] Poncelet Kris, Delarue Erik, Six Daan, Duerinck Jan, D'haeseleer William. Impact of the level of temporal and operational detail in energy-system planning models. Appl Energy 2016;162:631–43. http://dx.doi.org/10.1016/j.apenergy.2015.10.100.

[10] Das Partha, Mathur Jyotirmay, Bhakar Rohit, Kanudia Amit. Implications of short-term renewable energy resource intermittency in long-term power system planning. Energy Strategy Rev 2018;22:1–15. http://dx.doi.org/10.1016/j.esr.2018.06.005.

[11] Scott Ian J, Carvalho Pedro MS, Botterud Audun, Silva Carlos A. Long-term uncertainties in generation expansion planning: Implications for electricity market modelling and policy. Energy 2021;227:120371. http://dx.doi.org/10.1016/j.energy.2021.120371.

[12] Li Shuya, Coit David W, Felder Frank. Stochastic optimization for electric power generation expansion planning with discrete climate change scenarios. Electr Power Syst Res 2016;140:401–12. http://dx.doi.org/10.1016/j.epsr.2016.05.037.

[13] Ludig Sylvie, Haller Markus, Schmid Eva, Bauer Nico. Fluctuating renewables in a long-term climate change mitigation strategy. Energy 2011;36(11):6674–85. http://dx.doi.org/10.1016/j.energy.2011.08.021.

[14] Seljom Pernille, Tomasgard Asgeir. Short-term uncertainty in long-term energy system models — A case study of wind power in Denmark. Energy Econ 2015;49:157–67. http://dx.doi.org/10.1016/j.eneco.2015.02.004.

[15] Pineda Salvador, Morales Juan M. Capacity expansion of stochastic power generation under two-stage electricity markets. Comput Oper Res 2016;70:101–14. http://dx.doi.org/10.1016/j.cor.2015.12.007.

[16] Ringkjøb Hans-Kristian, Haugan Peter M, Seljom Pernille, Lind Arne, Wagner Fabian, Mesfun Sennai. Short-term solar and wind variability in long-term energy system models-a European case study. Energy 2020;209:118377. http://dx.doi.org/10.1016/j.energy.2020.118377.

[17] Skar C, Doorman G, Tomasgard A. The future European power system under a climate policy regime. In: 2014 IEEE international energy conference. 2014. p. 318–25, http://dx.doi.org/10.1109/ENERGYCON.2014.6850446.

[18] Skar Christian, Doorman Gerard, Pérez-Valdés Gerardo A, Tomasgard Asgeir. A multi-horizon stochastic programming model for the European power system. 2016, https://www.semanticscholar.org/paper/A-multi-horizon-stochastic-programming-model-for-Skar-Doorman/60e07af486807ee3b2426c70cfa5e6d54c8ed15e. [Accessed 30 July 2021].

[19] Marañón Ledesma Héctor, Tomasgard Asgeir. Analyzing demand response in a dynamic capacity expansion model for the European power market. Energies 2019;12(15):2976. http://dx.doi.org/10.3390/en12152976.

[20] Backe Stian, Korpås Magnus, Tomasgard Asgeir. Heat and electric vehicle flexibility in the European power system: A case study of Norwegian energy communities. Int J Electr Power Energy Syst 2021;125:106479. http://dx.doi.org/10.1016/j.ijepes.2020.106479.

[21] Wolfgang Ove, Haugstad Arne, Mo Birger, Gjelsvik Anders, Wangensteen Ivar, Doorman Gerard. Hydro reservoir handling in Norway before and after deregulation. Energy 2009;34(10):1642–51. http://dx.doi.org/10.1016/j.energy.2009.07.025.

[22] Spiecker Stephan, Weber Christoph. The future of the European electricity system and the impact of fluctuating renewable energy – A scenario analysis. Energy Policy 2014;65:185–97. http://dx.doi.org/10.1016/j.enpol.2013.10.032.

[23] Kaut Michal, Wallace Stein W. Evaluation of scenario-generation methods for stochastic programming. Pac J Optim 2007;3(2):257–71.

[24] European Commission. A clean planet for all - A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy. 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0773. [Accessed 30 July 2021].

[25] Bertocchi Marida, Consigli Giorgio, Dempster Michael AH. Stochastic optimization methods in finance and energy: New financial products and energy market strategies, volume 163. Springer; 2011.

[26] Giglio Richard J. Stochastic capacity models. Manage Sci 1970;17(3):174–84. http://dx.doi.org/10.1287/mnsc.17.3.174.

[27] Hennessy David A. Capacity choice in a two-stage problem under uncertainty. Econom Lett 1999;65(2):177–82. http://dx.doi.org/10.1016/S0165-1765(99)00149-4.

[28] Caramanis Michael C, Tabors Richard D, Nochur Kumar S, Schweppe Fred C. The introduction of nondiispatchable technologies a decision variables in long-term generation expansion models. IEEE Trans Power Appar Syst 1982;(8):2658–67. http://dx.doi.org/10.1109/TPAS.1982.317636.

[29] Quan Hao, Srinivasan Dipti, Khambadkone Ashwin M, Khosravi Abbas. A computational framework for uncertainty integration in stochastic unit commitment with intermittent renewable energy sources. Appl Energy 2015;152:71–82. http://dx.doi.org/10.1016/j.apenergy.2015.04.103.

[30] Glasserman Paul. Monte Carlo methods in financial engineering. Springer; 2003, URL https://www.springer.com/gp/book/9780387004518.

[31] Ottesen Stig, Tomasgard Asgeir. A stochastic model for scheduling energy flexibility in buildings. Energy 2015;88:364–76. http://dx.doi.org/10.1016/j.energy.2015.05.049.

[32] Smith James E. Moment methods for decision analysis. Manag Sci 1993;39(3):340–58. http://dx.doi.org/10.1287/mnsc.39.3.340.

[33] Høyland Kjetil, Kaut Michal, Wallace Stein W. A heuristic for moment-matching scenario generation. Comput Optim Appl 2003;24(2–3):169–85. http://dx.doi.org/10.1023/A:1021853807313.

[34] Jin Shan, Ryan Sarah M, Watson Jean-Paul, Woodruff David L. Modeling and solving a large-scale generation expansion planning problem under uncertainty. Energy Syst 2011;2:209–42. http://dx.doi.org/10.1007/s12667-011-0042-9.

[35] Ponomareva K, Roman D, Date P. An algorithm for moment-matching scenario generation with application to financial portfolio optimisation. European J Oper Res 2015;240(3):678–87. http://dx.doi.org/10.1016/j.ejor.2014.07.049.

[36] Hochreiter Ronald, Pflug Georg Ch. Financial scenario generation for stochastic multi-stage decision processes as facility location problems. Ann Oper Res 2007;152. http://dx.doi.org/10.1007/s10479-006-0140-6, 257–272.

[37] Gil Esteban, Aravena Ignacio, Cárdenas Raúl. Generation capacity expansion planning under hydro uncertainty using stochastic mixed integer programming and scenario reduction. IEEE Trans Power Syst 2015;30:1838–47. http://dx.doi.org/10.1109/TPWRS.2014.2351374.

[38] de Oliveira Welington Luis, Sagastizábal Claudia, Penna Débora Dias Jardim, Maceira Maria Elvira Pineiro, Damázio Jorge Machado. Optimal scenario tree reduction for stochastic streamflows in power generation planning problems. Optim Methods Softw 2010;25(6):917–36. http://dx.doi.org/10.1080/10556780903420135.

[39] Dupačová Jitka, Gröwe-Kuska Nicole, Römisch Werner. Scenario reduction in stochastic programming. Math Program 2003;95(3):493–511. http://dx.doi.org/10.1007/s10107-002-0331-0.

[40] Xu Bin, Zhong Ping-An, Zambon Renato C, Zhao Yunfa, Yeh William W-G. Scenario tree reduction in stochastic programming with recourse for hydropower operations. Water Resour Res 2015;51(8):6359–80. http://dx.doi.org/10.1002/2014WR016828.

[41] Jin Shan, Botterud Audun, Ryan Sarah M. Temporal versus stochastic granularity in thermal generation capacity planning with wind power. IEEE Trans Power Syst 2014;29(5):2033–41. http://dx.doi.org/10.1109/TPWRS.2014.2299760.

[42] Park Heejung, Baldick Ross. Multi-year stochastic generation capacity expansion planning under environmental energy policy. Appl Energy 2016;183:737–45. http://dx.doi.org/10.1016/j.apenergy.2016.08.164.

[43] Seljom Pernille, Tomasgard Asgeir. 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;106:85–102. http://dx.doi.org/10.1016/j.enpol.2017.03.011.

[44] Seljom Pernille, Lindberg Karen Byskov, Tomasgard Asgeir, Doorman Gerard, Sartori Igor. The impact of zero energy buildings on the Scandinavian energy system. Energy 2017;118:284–96. http://dx.doi.org/10.1016/j.energy.2016.12.008.

[45] Seljom Pernille, Tomasgard Asgeir. Sample average approximation and stability tests applied to energy system design. Energy Syst 2019;106:85–102. http://dx.doi.org/10.1007/s12667-019-00363-x.

[46] Kim Sujin, Pasupathy Raghu, Henderson Shane G. A guide to sample average approximation. In: Handbook of simulation optimization. Springer; 2015, p. 207–43. http://dx.doi.org/10.1007/978-1-4939-1384-8_8.

[47] Backe Stian. OpenEMPIRE: Stochastic linear program for investments in the European power system. 2020, https://github.com/ntnuiotenergy/OpenEMPIRE. [Accessed 30 July 2021].

[48] Kaut Michal, Midthun Kjetil T, Werner Adrian S, Tomasgard Asgeir, Hellemo Lars, Fodstad Marte. Multi-horizon stochastic programming. Comput Manag Sci 2014;11(1–2):179–93. http://dx.doi.org/10.1007/s10287-013-0182-6.

[49] Anderson Paul H. Distributions in stratified sampling. Ann Math Stat 1942;13:42–52. http://dx.doi.org/10.1214/aoms/1177731641.

[50] Lloyd Stuart. Least squares quantization in PCM. IEEE Trans Inform Theory 1982;28(2):129–37. http://dx.doi.org/10.1109/TIT.1982.1056489.

[51] Eggleston S, Buendia L, Miwa K, Ngara T, Tanabe K. 2006 IPCC guidelines for national greenhouse gas inventories. In: Volume 2 energy. Stationary combustion, vol. 5, Japan: Institute for Global Environmental Strategies Hayama; 2006 [Chapter 2].

[52] Economics London. The value of lost load (VoLL) for electricity in Great Britain. 2013, http://londoneconomics.co.uk/blog/publication/the-value-of-lost-load-voll-for-electricity-in-great-britain/. [Accessed 30 July 2021].

[53] De Vita Alessia, Kielichowska Izabela, Mandatowa Pavla, Capros P, Dimopoulou E, Evangelopoulou S, et al. Technology pathways in decarbonisation scenarios. 2018, https://ec.europa.eu/energy/sites/ener/files/documents/2018_06_27_technology_pathways_-_finalreportmain2.pdf. [Accessed 30 July 2021].

[54] ENTSO-E. Statistical factsheet 2018. 2018, https://docstore.entsoe.eu/Documents/Publications/Statistics/Factsheet/entsoe_sfs2018_web.pdf. [Accessed 30 July 2021].

[55] ENTSO-E. ENTSO-E transparency platform. 2021, https://transparency.entsoe.eu/. [Accessed 30 July 2021].

[56] Pfenninger Stefan, Staffell Iain. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. Energy 2016;114:1251–65. http://dx.doi.org/10.1016/j.energy.2016.08.060.

[57] Staffell Iain, Pfenninger Stefan. Using bias-corrected reanalysis to simulate current and future wind power output. Energy 2016;114:1224–39. http://dx.doi.org/10.1016/j.energy.2016.08.068.

[58] Hart William E, Watson Jean-Paul, Woodruff David L. Pyomo: Modeling and solving mathematical programs in Python. Math Program Comput 2011;3(3):219–60. http://dx.doi.org/10.1007/s12532-011-0026-8.

[59] Hart William E, Laird Carl D, Watson Jean-Paul, Woodruff David L, Hackebeil Gabriel A, Nicholson Bethany L, et al. Pyomo–optimization modeling in Python, volume 67. 2nd ed.. Springer Science & Business Media; 2017, http://dx.doi.org/10.1007/978-3-319-58821-6.

[60] Brouwer Anne Sjoerd, van den Broek Machteld, Seebregts Ad, Faaij André. Operational flexibility and economics of power plants in future low-carbon power systems. Appl Energy 2015;156:107–28. http://dx.doi.org/10.1016/j.apenergy.2015.06.065.

[61] Karmarkar Narendra. A new polynomial-time algorithm for linear programming. In: Proceedings of the sixteenth annual ACM symposium on theory of computing. 1984. p. 302–11, http://dx.doi.org/10.1007/BF02579150.

[62] Gurobi Optimization LLC. Gurobi optimizer reference manual. 2020, http://www.gurobi.com. [Accessed 30 July 2021].

[63] Rockafellar R Tyrrell, Uryasev Stanislav. Optimization of conditional value-at-risk. J Risk 2000;2:21–42. http://dx.doi.org/10.21314/JOR.2000.038.

[64] Filippi Carlo, Guastaroba Gianfranco, Speranza Maria Grazia. Conditional value-at-risk beyond finance: A survey. Int Trans Oper Res 2020;27(3):1277–319. http://dx.doi.org/10.1111/itor.12726.

[65] Yu Xiaobao, Zheng Dandan. Cross-regional integrated energy system scheduling optimization model considering conditional value at risk. Int J Energy Res 2020;44(7):5564–81. http://dx.doi.org/10.1002/er.5307.

[66] Solaun Kepa, Cerdá Emilio. Climate change impacts on renewable energy generation. A review of quantitative projections. Renew Sustain Energy Rev 2019;116:109415. http://dx.doi.org/10.1016/j.rser.2019.109415.

[67] Hamududu Byman, Killingtveit Aanund. Assessing climate change impacts on global hydropower. In: Climate change and the future of sustainability. Apple Academic Press; 2017, p. 109–32. http://dx.doi.org/10.3390/en5020305.

[68] Devis Annemarie, Van Lipzig Nicole PM, Demuzere Matthias. Should future wind speed changes be taken into account in wind farm development? Environ Res Lett 2018;13(6):064012. http://dx.doi.org/10.1088/1748-9326/aabff7.

[69] Jerez Sonia, Tobin Isabelle, Vautard Robert, Montávez Juan Pedro, López-Romero Jose María, Thais Françoise, et al. The impact of climate change on photovoltaic power generation in Europe. Nature Commun 2015;6(1):1–8. http://dx.doi.org/10.1038/ncomms10014.

[70] Ramdas Aaditya, Trillos Nicolás García, Cuturi Marco. On Wasserstein two-sample testing and related families of nonparametric tests. Entropy 2017;19(2):47. http://dx.doi.org/10.3390/e19020047.

[71] Pflug G Ch. Scenario tree generation for multiperiod financial optimization by optimal discretization. Math Program 2001;89(2):251–71. http://dx.doi.org/10.1007/PL00011398.