Data-Informed Scenario Generation for Statistically Stable Energy Storage Sizing in Isolated Power Systems ⋆ ,⋆⋆

Spyridon Chapaloglou[a,*], Damiano Varagnolo[b], Francesco Marra[c], Elisabetta Tedeschi[a,d]

[a] *Department of Electric Power Engineering, Norwegian University of Science and Technology, O.S. Bragstads Plass 2 E, 7034 Trondheim, Norway*
[b] *Department of Engineering Cybernetics, Norwegian University of Science and Technology, O.S. Bragstads Plass 2 E, 7034 Trondheim, Norway*
[c] *Equinor R&T Electrical Technology department, Arkitekt Ebbels 10, 7005, Trondheim, Norway*
[d] *Department of Industrial Engineering, University of Trento, Via Sommarive, 9, 38123 Povo, Italy*

Abstract

We consider the problem of how to generate uncertainty scenarios to be used in energy storage sizing problems in isolated power systems. More precisely, we consider storage sizing formulations where both loads and generation are stochastic, where no closed form analytical expression is available, and where the presence of multiple discrete random variables makes the sizing problem mixed integer and with combinatorial search spaces. We thus propose and characterize a data-driven scenarios selection strategy that mitigates the computational issues associated with these types of storage sizing problem formulations while guaranteeing statically stable optimal solutions. Specifically, the approach works by first learning the distribution of the uncertainties of the loads and generation starting from field data, and then generating, through the learned distri-

⋆Project / research funded by VISTA - a basic research program in collaboration between The Norwegian Academy of Science and Letters, and Equinor.
⋆⋆This scientific paper was supported by the Onassis Foundation - Scholarship ID: F ZP 056-1/2019-2020.
⋆Corresponding author
URL: spyridon.chapaloglou@ntnu.no (Spyridon Chapaloglou), damiano.varagnolo@ntnu.no (Damiano Varagnolo), fmarr@equinor.com (Francesco Marra), elisabetta.tedeschi@ntnu.no (Elisabetta Tedeschi)

bution, an optimal set of uncertainty scenarios that are subsequently used in a two-stage stochastic programming reformulation of the original sizing problem. The workflow does not impose arbitrary structures to the correlation among the uncertainties, nor does it lump these in a single parameter; thus, it is suitable for systems with any load characteristics. Moreover, the approach ensures to reach a solution that is statistically close to the one that would be computed if the original problem was solvable and not computationally intractable. As a case study, we analyze the problem of designing an energy storage system for a wind powered oil and gas platform to minimize the expected daily system operational costs. Numerical simulations showed that the proposed methodology leads to higher quality solutions compared with other scenario selection strategies. This reveals realistic estimations of the expected benefits, while also highlighting the risk-management limitations, when solving the risk-constrained version of the problem.

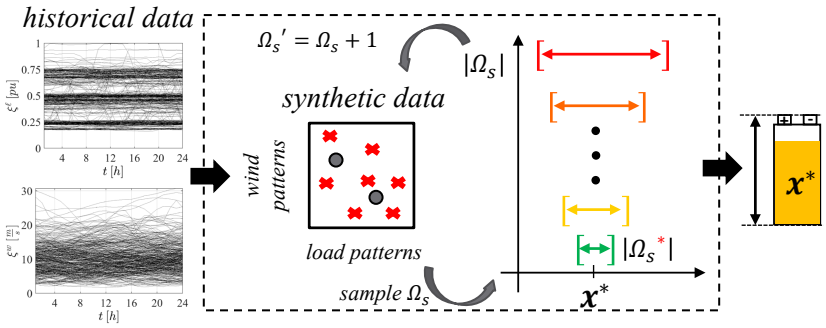


Figure 1: Graphical abstract

Keywords: Scenario reduction, stochastic programming, mixed-integer optimization, energy storage, offshore installations, renewable energy, multiple uncertainties, wind uncertainty, demand volatility

1. Introduction

Modern power systems are characterized by high levels of uncertainty and operational complexity. This stems from the interaction of intermittent en-

ergy production from Renewable Energy Sources (RESs) with electricity consumption patterns. Under this perspective, and to achieve a more reliable, cost-effective and environmentally friendly operation, Energy Storage Systems (ESSs) are recognized as an essential component of future grids [1, 2] and are especially valuable for non-interconnected systems. However, determining the optimal size of an ESS when considering the expected future system operational patterns is not a trivial task [3, 4, 5]. This is because the problem should account not only for costs and other important limiting factors, but also for the fact that the uncertainties (especially load requirements and generation) are statistically distributed, and their distribution cannot be learned perfectly.

A potential strategy to deal with the presence of such uncertainties is to employ stochastic optimization formulations [6, 7]. When integrating uncertainty into the storage sizing problem by assuming parametric distributions [8, 6], there exist inherent limitations on how well the available data can be described and generalized. On the contrary, data-based methods have shown the ability to better generalize and replicate more complex patterns found in historical datasets [9]. Energy storage sizing problems that consider the expected daily system operation are often formulated as two-stage Mixed-Integer Linear Programming (MILP) problems where day-ahead unit commitment and/or economic dispatch are determined over a scenario set that includes uncertainty realizations [10, 9, 11]. These approaches are indeed inherently less conservative than worst-case alternatives [12]. However, they often lead to problem formulations that are computationally intractable [13] because the cost to be optimized should consider all the scenarios that have been observed in the past. In other words, assuming for example we have a dataset comprising years of measured loads and disturbances. Here, such formulations would naturally embed all the scenarios, leading to a computationally intractable combinatorial problem. However, in cases where the resulting problems are linear, (LP) [14] such issues do not exist and one is not typically constrained by computational tractability.

To mitigate this issue, one potential approach is reducing the number of scenarios to be considered in the stochastic optimization formulation [11, 15].

However, the algorithm for selecting such alternative subsets of scenarios should lead to a final sizing solution that has low sensitivity to the considered subset. Otherwise, the quality of the final solution is intuitively of less value since it is arbitrated by the scenarios selection algorithm and not by the dataset itself (see also [16] for a more formal discussion of this point). For that reason, to guarantee the reliability of the final sizing, the scenarios selector should not only properly generate/select such scenarios, but also keep a sufficient number of them to minimize the inevitable information loss due to the scenarios sampling procedure [15, 17].

In the literature, one may thus find various ways of selecting scenarios subsets, e.g., [18, 19, 20, 21, 22, 23, 24, 25, 9, 11]. Highlighting the main categories that most of these methods fall, we have scenario generation by: random sampling of historical data, optimal scenario reduction, moments matching, and clustering to representative patterns [25, 9, 11]. As another example, [17] proposed the scenario map method as an alternative scenario reduction technique that considers RES and load uncertainty through the system net load. However, net load data do not capture the whole uncertainty space since they represent a conditional slice of it, especially when the underlying individual uncertain profiles present high variance (such as in isolated industrial power systems with wind power). In addition, specific correlation values between RES and load were imposed in [26, 27, 28] and Cholesky decomposition was then used to model dependent profiles, as commonly done in several studies. To the best of our knowledge, all available literature reports that when dealing simultaneously with both RES and load uncertainties, these factors are either lumped together in a single parameter before using some scenarios reduction method (namely the net-load as in [29, 30]), or the proposed approach imposes some specific structure to the correlation among the uncertainties. However, the validity of such assumptions is debatable, as they depend on the specific power system application. Those assumptions typically rely on conclusions from specific analyses [31] which *i)* targeted large interconnected power systems (implying that such scenarios selection strategies are not suitable for the specific case of isolated

power systems with non-standard load patterns) and *ii)* are not guaranteed to be time-invariant. In fact [32] showed that, for wind energy, the observed correlation levels are much lower than the ones observed between load and solar energy, due to its less consistent and predictable behavior.

As mentioned although it is possible to generate high quality profiles for individual uncertain parameters (i.e., RES or load individually), most of the state-of-the-art stochastic optimization formulations for the sizing problem do not effectively explore the whole space of combined uncertainties and often do not investigate how big this exploration should be to ensure that the final computed size is statistically close to what one would get by solving the original problem, including all possible scenarios from the available datasets. As such, to the best of our knowledge, there is currently no literature that addresses, in a general way, how to select and combine RES and load profiles (and their associated uncertainties) over each other from a dataset that is too large to be used in its entirety, which is typical for two-stage MILP sizing formulations.

Contributions

This paper proposes a methodology for performing this selection that considers different rearrangements of profiles from the reduced subsets (samples) as conditional draws (instances) of the total uncertainty space (population). Therefore, the proposed approach intuitively aims to perform more accurate modelling/exploration of the combined uncertainty space than the approaches existing in the literature. This will make the sizing solution statistically closer to what would be computed (if we had the computational capability) with the whole dataset.

In summary, we propose an integrated methodology that can be applied to energy storage sizing problems under stochastic optimization frameworks, and that generates subsets of scenarios that intrinsically consider the relationship between the optimization problem and the input data. The novelty of the approach is that it generates minimum-cardinality subsets that implicitly account for the effect of coupling load and RES uncertainties through opportune sta-

tistical evaluations of the solutions to the original optimization problem. In addition, the entire methodology is purely data-driven, and it avoids any strong assumptions on the individual uncertainties and their coupling. The proposed methodology is compared against other scenario subset generation methods, demonstrating its superiority to achieve preferable statistical properties for the obtained optimization solutions. As a test case study, we address the sizing of a techno-economically feasible ESS for an isolated offshore oil and gas (O&G) platform, that includes onsite power generation from Gas Turbines (GTs) and integrates wind power.

The remainder of the paper is organized as follows: section 2 formulates the optimization problem, section 3 describes the proposed scenario generation methodology, and section 4 presents the numerical results and an investigation of which effects the worst-case scenarios included in the reduced subset have on the solutions. Finally, section 5 presents an overall summary and concluding remarks.

2. Formulating the BESS sizing problem using risk-aversion considerations

This section formulates the BESS sizing problem as a stochastic optimization problem (SP) that integrates considerations on risks by introducing a risk-aversion user-defined parameter. The following subsections introduce the quantities needed to arrive at the formulation that is summarized at the end of the section.

2.1. Objective function and cost modelling

Let in general $\boldsymbol{x}$ be the first-stage decision variables expressing the BESS size in terms of power and capacity rating with cost $\boldsymbol{c}$, $\boldsymbol{y}$ the second-stage optimal decision about how to operate a plant with cost $\boldsymbol{q}$, and $\tilde{\boldsymbol{\xi}}$ the stochastic process variables introducing uncertainty into the parameters of the problem. Note that we explicitly consider two types of uncertainty: $\tilde{\boldsymbol{\xi}}^{\ell}$ to model randomness in the

platform's load, and $\tilde{\boldsymbol{\xi}}^w$ to model randomness in the wind speed at the specified offshore location. Notation-wide, we let $\tilde{\boldsymbol{\xi}} = \left[\tilde{\boldsymbol{\xi}}^\ell \;\; \tilde{\boldsymbol{\xi}}^w\right]^T$. Given this notation, the BESS SP sizing can be initially phrased as

$$\begin{aligned} \text{SP:} \;\; & \min_{\boldsymbol{x}}\{\mathbb{E}[f(\boldsymbol{x};\tilde{\boldsymbol{\xi}})]\} \\ \text{where,} \;\; & f(\boldsymbol{x};\tilde{\boldsymbol{\xi}}) = \boldsymbol{c}^T\boldsymbol{x} + \min_{\boldsymbol{y}}\{\boldsymbol{q}^T\boldsymbol{y}\} \\ \text{s.t.} \;\; & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} + \boldsymbol{C}\tilde{\boldsymbol{\xi}} \leq \boldsymbol{0}. \end{aligned} \tag{1}$$

Note this formulation includes all the randomness related to the stochastic process variables, leading to coupled constraints through data matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$.

This problem can be formulated as a Mixed-Integer Linear Program by deriving the deterministic equivalent of (1) and discretizing the continuous stochastic process $\tilde{\boldsymbol{\xi}}$ over a set Ω containing all the possible realizations of the uncertain parameters (scenarios). The discretized process is denoted as $\widehat{\boldsymbol{\xi}}$. Through this, thus, we can derive the Mixed-Integer Linear Program

$$\begin{aligned} \text{SP:} \;\; & \min_{\boldsymbol{x},\boldsymbol{y}(\omega)} \{\boldsymbol{c}^T\boldsymbol{x} + \sum_{\omega \in \Omega} \pi_\omega \; \boldsymbol{q}^T(\omega)\boldsymbol{y}(\omega)\} \text{ where,} \\ \boldsymbol{q}^T\boldsymbol{y}(\omega) = & \sum_{g \in N_G} \sum_{t \in T} C_{GT} P^g_{GT,t,\omega} + C^{on}_{GT} u^g_{t,\omega} + C^{start}_{GT} z^g_{t,\omega} \\ \text{s.t.} \;\; & \boldsymbol{A}(\omega)\boldsymbol{x} + \boldsymbol{B}(\omega)\boldsymbol{y}(\omega) + \boldsymbol{H}(\omega)\widehat{\boldsymbol{\xi}}(\omega) \leq \boldsymbol{0}, \; \forall\, \omega \in \Omega\,. \end{aligned} \tag{2}$$

Note then that the structure of (2) is such that the first stage investment cost for the BESS decision (i.e., $\boldsymbol{c}^T\boldsymbol{x}$) can be divided into distinct capacity-related ($C_{B,E}$) and power-related ($C_{B,P}$) components. Moreover, the investment is amortized into a daily basis through Capital Recovery Factor (CRF) as

$$\boldsymbol{c}^T\boldsymbol{x} = \frac{r \cdot (1+r)^p}{(1+r)^p - 1}[C_{B,P} \;\; C_{B,E}] \begin{bmatrix} P_B \\ E_B \end{bmatrix} \tag{3}$$

where r is the daily interest rate (derivable from an annual interest rate), p is the recouping period ($p = 365L$), L is the investment lifetime, and P_B and E_B are the BESS power rating and capacity, respectively. The second stage decision $\boldsymbol{q}^T(\omega)\boldsymbol{y}(\omega)$ relates to the operation of the platform's GTs for each scenario ω (whose occurrence probability is assumed to be π_ω). The second stage control

variables $\boldsymbol{y}(\omega)$ and corresponding costs $\boldsymbol{q}^T(\omega)$ are given as

$$\begin{aligned}
\boldsymbol{y}(\omega) &= \left[\boldsymbol{P}_{B,\omega}\ \boldsymbol{P}^g_{GT,\omega}\ \boldsymbol{u}^g_\omega\ \boldsymbol{z}^g_\omega\ \boldsymbol{P}_{D,\omega}\right]^T \quad \forall\, \omega \in \Omega \\
\boldsymbol{q}^T(\omega) &= \left[\mathbf{0}\ \mathbf{C_{GT}}\ \mathbf{C^{on}_{GT}}\ \mathbf{C^{start}_{GT}}\ \mathbf{0}\right] \quad \forall\, \omega \in \Omega \ \text{ where,} \\
\mathbf{C_{GT}} &= \left\{\left(\frac{C_{NG}}{\rho} + \mu C_{CO_2}\right)\alpha_g + C_{O\&M}\right\}_{t\in T} \\
\mathbf{C^{on}_{GT}} &= \left\{\left(\frac{C_{NG}}{\rho} + \mu C_{CO_2}\right)\beta_g\right\}_{t\in T}, \\
\mathbf{C^{start}_{GT}} &= \left\{C^{start}_{GT}\right\}_{t\in T}
\end{aligned} \tag{4}$$

where $\boldsymbol{P}_{B,\omega}$ and $\boldsymbol{P}_{D,\omega}$ are the BESS and (controllable) dump power vectors respectively; $\boldsymbol{P}^g_{GT,\omega}$, $\boldsymbol{u}^g_\omega$, $\boldsymbol{z}^g_\omega$, are the GT power, ON/OFF status and startup indicator variables for each generator g and all time periods T respectively; μ is the ideal combustion coefficient of natural gas (NG) to CO_2; C_{NG} is the fuel (NG) sale value per normal cubic meter; ρ is the NG density at standard temperature and pressure conditions; C_{CO_2} is the CO_2 tax per kg of released CO_2; $C_{O\&M}$ is the estimated operation and maintenance (O&M) costs per generated kWh; C^{start}_{GT} is the estimated startup cost of a generator per event; and α_g and β_g are the estimated regression parameters of the fitted model to data explaining fuel consumption ($\dot{m}_f$) as a function of the GT loading (P_{GT}), approximated with the commonly used affine map $\dot{m}_f = \alpha_g \cdot P_{GT} + \beta_g$.

To complete the formulation of the optimization problem, we then proceed with describing its constraints.

2.2. Energy system operation constraints

The daily system operation can be explicitly considered in the optimization problem (2) by constraining the feasible solution space. From a methodological perspective, these constraints can be derived from modelling the operation of each sub-component, for a given scenario tree Ω, as shown in the following.

The Gas Turbine operation is represented through the following set of equations, where

$$\begin{gathered}
u^g_{t,\omega}\gamma\overline{P}^g_{GT} \le P^g_{GT,t,\omega} \le u^g_{t,\omega}\overline{P}^g_{GT}, \\
\forall\, t \in T,\ g \in N_G,\ \omega \in \Omega
\end{gathered} \tag{5}$$

is the box constraint for the GT power, $\overline{P}^{g}_{GT}$ is the maximum operation capacity of each GT, and γ is the allowed minimum technical operational ratio. To model the realistic case of an offshore O&G platform, we consider the presence of four identical GTs with the same maximum rating and the same technical minimum γ set to be 20%. The GT power ramping constraint was described as

$$\left| P^{g}_{GT,t,\omega} - P^{g}_{GT,t-1,\omega} \right| \leq R,\ \forall\, t \in T,\ g \in N_G,\ \omega \in \Omega \tag{6}$$

where R is the allowed ramping rate of the GT. Recalling that this study considers an hourly discretization of the time series, we set R as $\overline{P}_{GT}$ for all four generators. The minimum required time to start-up a GT after a shut down is modeled as

$$\begin{aligned} u^{g}_{GT,t-1,\omega} - u^{g}_{GT,t,\omega} &\leq 1 - u^{g}_{GT,k,\omega}, \\ \forall\, t \in T,\ g \in N_G,\ \omega \in \Omega, & \\ k = \{t+1, \ldots, \min(t + T_{off} - 1, T)\} & \end{aligned} \tag{7}$$

where T_{off} is the minimum GT off-time after a shutdown (here set to 4 hours, to the best of our knowledge a value representing typical field setups). The startup events of the GTs are modelled as

$$\begin{aligned} u^{g}_{GT,t,\omega} - u^{g}_{GT,t-1,\omega} &\leq z^{g}_{GT,t,\omega}, \\ \forall\, t \in T,\ g \in N_G,\ \omega \in \Omega & \end{aligned} \tag{8}$$

The BESS inter-temporal and cycling constraints are instead modeled as

$$E_{B,t,\omega} = E_{B,0} + \sum_{k=1}^{t} P_{B,k,\omega},\ \forall\, t \in T,\ \omega \in \Omega \tag{9}$$

$$0 \leq E_{B,t,\omega} \leq E_B,\ \ |P_{B,t,\omega}| \leq\ P_B\ \forall\, t \in T,\ \omega \in \Omega \tag{10}$$

$$\sum_{t \in T} P_{B,t,\omega} = 0,\ \forall\, \omega \in \Omega \tag{11}$$

where $E_{B,t,\omega}$ is the remaining energy capacity of the BESS at any instant t, for every scenario ω. Moreover, $E_{B,0,\omega} = 0$ shall be considered as the initial capacity of the BESS for every scenario ω. The wind power generation is then

modeled after the basic power curve of the reference case wind turbines, i.e., as

$$W\left(\widehat{\xi}_t^w(\omega)\right) = \begin{cases} 0, & \widehat{\xi}_t^w(\omega) \leq w_{ci} \\ N_w P_w^n \left(\frac{\widehat{\xi}_t^w(\omega)}{w_n}\right)^3, & w_{ci} \leq \widehat{\xi}_t^w(\omega) \leq w_n \\ N_w P_w^n, & w_n \leq \widehat{\xi}_t^w(\omega) < w_{co} \\ 0, & w_{co} \leq \widehat{\xi}_t^w(\omega) \end{cases} \tag{12}$$

$$\forall\, t \in T,\ \omega \in \Omega$$

where w_{ci} is the cut-in wind speed; w_n is the nominal wind speed; w_{co} is the cut-off wind speed; P_w^n is the nominal power of each wind turbine; and N_w is the number of wind turbines in the considered wind farm. Finally, the sub-components interaction is modelled through the power balance equation as

$$\sum_{g\,\in\, N_G} P_{GT,t,\omega}^g + P_{B,t,\omega} - P_{D,t,\omega} = \widehat{\xi}_t^{\ell}(\omega) - W\left(\widehat{\xi}_t^w(\omega)\right) \tag{13}$$

$$\forall\, t \in T,\ \omega \in \Omega$$

2.3. Risk-management formulation

So far, the formulation of problem (1) allows for minimizing the cost under a specific probabilistic disturbance model. However, it does not inherently consider the effects of extreme disturbances. To account for these, we follow the common strategy, e.g., see [33], of penalizing these extreme realizations using the Conditional Value-at-Risk (CVaR) risk measure [34, 10, 26, 35]. In practice, this means that CVaR represents the expected value of the cost for the $1-\alpha$ percentage of worst scenarios. In other words, given an α-quantile, the random cost variable $f\left(\boldsymbol{x};\tilde{\boldsymbol{\xi}}\right)$ from SP (1), and a decision $\boldsymbol{x}$, the $CVaR$ can be calculated after [34] as

$$CVaR_\alpha(\boldsymbol{x}) = \mathbb{E}_{\tilde{\boldsymbol{\xi}}}\left[f(\boldsymbol{x}) \middle| f(\boldsymbol{x}) \geq VaR_\alpha(\boldsymbol{x})\right]. \tag{14}$$

Note that, as soon as the distribution of the random variable is known, eq. (14) can be efficiently computed by solving the linear program [34], i.e.,

$$CVaR_{\alpha}(\boldsymbol{x}) = \min_{\zeta,\ s(\omega)\in\mathbb{R}_{\geq 0}} \left\{ \zeta + \frac{1}{1-\alpha} \cdot \sum_{\omega\in\Omega} \pi_{\omega} s(\omega) \right\} \\ \text{s.t.} \quad f(\boldsymbol{x}; \widehat{\boldsymbol{\xi}}(\omega)) - \zeta \leq s(\omega) \quad \forall\, \omega \in \Omega\ . \tag{15}$$

where $s(\omega)$ are scenario-dependent slack variables and ζ is a helper variable indicating the lowest value of the worst-cases costs. The structure of (15) implies that at the optimal point it holds that $\zeta^* = VaR_{\alpha}(\boldsymbol{x})$. Finally, the optimization problem (2) can integrate risk by introducing a risk-aversion control parameter β, and by reformulating the original problem as

$$\min_{\boldsymbol{x},\boldsymbol{y}(\omega),\zeta,s(\omega)} F\left(\boldsymbol{x}; \widehat{\boldsymbol{\xi}}\right)\ \text{, s.t. } \textit{eq. } (12) - (13) \text{ and } (15) \tag{16}$$

where $F\left(\boldsymbol{x}; \widehat{\boldsymbol{\xi}}\right) = (1-\beta)\left(\boldsymbol{c}^T\boldsymbol{x} + \sum_{\omega\,\in\,\Omega} \pi_{\omega}\ \boldsymbol{q}^T(\omega)\boldsymbol{y}(\omega)\right)$
$+\ \beta\left(\zeta + \frac{1}{1-\alpha} \cdot \sum_{\omega\in\Omega} \pi_{\omega} s(\omega)\right)$.

We note once again that this formulation includes all the scenarios that may have been recorded from historical data. This means that for large datasets, this MILP is computationally unsolvable; consequently, an approach like the one proposed in this paper is needed that selects which scenarios should be included in the formulation in a statistically meaningful way.

3. Methodology

Our goal is to select a number of opportune scenarios that assist in solving a reduced version of eq. (16) such that the solution of the reduced problem is statistically close to the solution of the original (computationally intractable) one. From an intuitive perspective, this requires selecting scenarios that are distributed within the uncertainty space in a way that captures the statistical properties of the potential scenarios that the plant may encounter. Thus, this section proposes an algorithm that serves this selection purpose.

3.1. Accounting for Time Dependencies

To obtain accurate representations of uncertainties, we consider that the generated scenarios must reproduce the cyclostationarity of the historical field data. To infer these, one may consider a dataset $\{\xi^p\}_{i=1:N}$ composed of one-year long hourly *i)* aggregated load ($p = \ell$) time-series from an existing offshore platform (load at 50 MW power range) and *ii)* wind speed ($p = w$) time-series for the same location.

The first step of inferring the probability functions from such historical field data is to infer the structure of the marginal distributions, before inferring the joint ones. Importantly, in our proposed methodology we make no assumptions on the parametric structure of the marginals of the physical variables ($p = \{\ell, w\}$), and rather infer it through a Kernel Density Estimation (KDE) method [33, 36], i.e., independently compute non-parametric smooth representations of the time-dependent marginal empirical distributions of load and wind speed as

$$\widehat{F}^p_{h,t}(\tilde{\xi}^p_t) = \frac{1}{Nh_p} \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\tilde{\xi}_t - \xi^p_i)^2}{2h_p^2}} \quad \forall\, t \in T. \tag{17}$$

To capture the statistical dependency of the load or wind process along its temporal dimension, we then propose to use a copula-based approach, i.e., express the joint distribution of the load or wind vector as an opportune combination of the marginals at the various times t and a copula function for $p = \{\ell, w\}$, so that the joint distribution becomes, consistent with Sklar's theorem,

$$F^p\left(\tilde{\boldsymbol{\xi}}^p\right) = C^p\left(F^p_{h,1}\left(\tilde{\xi}^p_1\right), \cdots, F^p_{h,T}\left(\tilde{\xi}^p_T\right)\right). \tag{18}$$

Assuming then a Gaussian copula for C^p, known to have the potential of adequately modelling temporal dependencies [22], leads to the possibility of using Maximum Likelihood inference to estimate such temporal correlations.

We then propose to leverage copula-based approaches to generate independent profiles from F^p. Specifically, considering that $u^p_t = \widehat{F}^p_{h,t} \sim unif\,[0,1]$, plus using Monte-Carlo sampling and probability inverse transforms (PIT), we propose the generation steps to be:

i. a sampling as $\boldsymbol{z} \sim N\left(0, \widehat{\rho}^{p}_{ML}\right)$,

ii. a unit cube transformation as $\tilde{u}^{p}_{t} = \Phi\left(z_t\right),\ \forall\, t \in T$,

iii. an inverse transformation as $\tilde{\xi}^{p}_{t} = \widehat{F}^{p^{-1}}_{h,t}(\tilde{u}^{p}_{t}),\ \forall\, t \in T$.

Here, Φ is the cumulative distribution function of a standard normal and N is a multivariate normal.

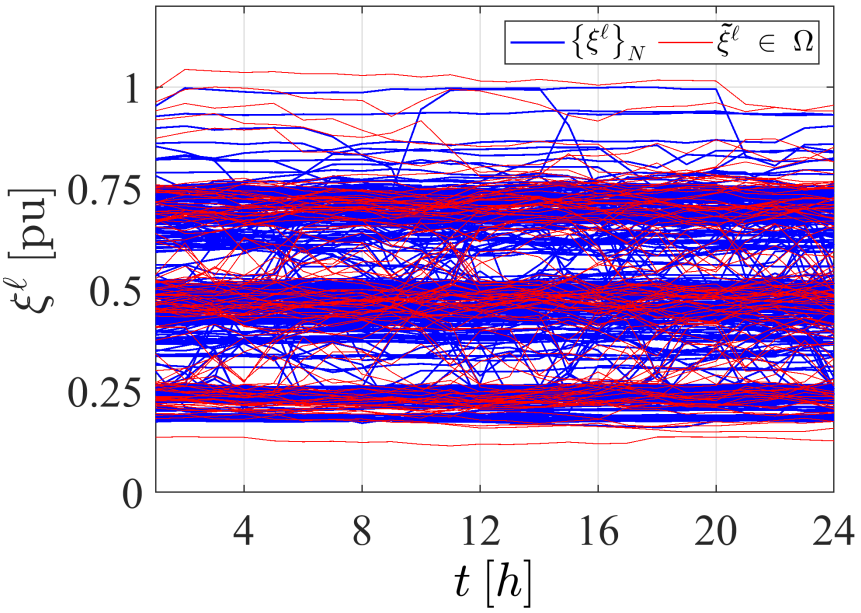


Figure 2: Random load profiles sampling after learning the temporal correlation structure from the dataset and qualitative validation of the capability to reproduce the dataset properties (case for $p = \ell$).

Following the proposed procedure, one can generate sample sets Ω containing representative profiles $\tilde{\boldsymbol{\xi}}^{p}(\omega)$ where $\omega = 1 \ldots |\Omega|$. For $p = w$, the profiles can be directly transformed to wind power through eq. (12). After sampling 100 random profiles from F^{ℓ} and comparing them with the initial dataset (fig. 2), it is possible to observe that the sampled data reproduce the qualitative properties of the original dataset. Thus, the estimated F^{p} can be well approximated through a random sampling approach to generate, arbitrarily, many synthetic profiles.

3.2. Selecting the Scenarios for the considered BESS SP

As indicated above, our goal is to construct scenarios $\tilde{\boldsymbol{\xi}} \in \Omega$ by combining generated profiles $\tilde{\boldsymbol{\xi}}^p(\omega)$ in a way that characterizes the system's uncertainty as well as possible, given a fixed number of scenarios to be included in the sizing Stochastic Problem. To achieve this, the first step is to introduce mathematical tools to assess the usefulness of a given scenario for characterizing the system's uncertainty (section 3.2.1). The second step is to understand how many scenarios are needed to achieve sufficiently good characterization (section 3.2.2).

3.2.1. Ranking the scenarios

Assume Ω to be populated with a number of sampled profiles, and assume our goal to be selecting a subset Ω_s of them. Before delving into details, the overarching structure of our proposed approach is: *i)* rank the samples within Ω based on how representative they are of their kind ($p = \{\ell, w\}$), *ii)* use these ranking scores to map the sampled profiles into a low-dimensional space, *iii)* discretize this space in some statistically optimal way, *iv)* use the scenarios $\widehat{\boldsymbol{\xi}} \in \Omega_s$ from this discretization and solve a computationally tractable version of eq. (16). This whole procedure is illustrated graphically in fig. 3.

We propose starting by ranking the profiles $\tilde{\boldsymbol{\xi}}^p \in \Omega$ using a Kantorovich distance, i.e., using

$$\begin{aligned} D_K^p &= \sum_{\omega \in \Omega \setminus \Omega_k} \pi_\omega \min_{\omega, \omega_k} v^p(\omega, \omega_k), \quad \text{where,} \\ v^p(\omega, \omega_k) &= ||\tilde{\boldsymbol{\xi}}^p(\omega) - \tilde{\boldsymbol{\xi}}^p(\omega_k)|| = \sum_{t=1}^{T} |\tilde{\xi}_t^p(\omega) - \tilde{\xi}_t^p(\omega_k)| \end{aligned} \tag{19}$$

where Ω_k is the dynamically updated set including the k most representative profiles (note that when $k = \left|\Omega\right|$, then Ω_k will be an ordered version of Ω). The usage of a Kantorovich distance is due to its simplicity in providing good and quick heuristic solutions to the optimal transportation problem for scenario reduction algorithms (i.e. Fast Forward Selection (FSS)) [18, 9].

Note now that the rank value of a profile ranked k^{th} out of $\left|\Omega\right|$ can be expressed as the scalar $t\left(\tilde{\boldsymbol{\xi}}^p(\omega_k)\right) = \frac{\left|\Omega\right|-k}{\left|\Omega\right|-1}$, $p = \{\ell, w\}$ where $t(\omega_k) \to 1$ for

profiles that can be thought as "more representative/typical" and $t(\omega_k) \to 0$ for profiles profiles that capture non-casual/non-typical events. In this way, a mapping is enforced, which not only preserves information spanning the whole support of the distribution of each individual random variable (load, wind), but also indicates different possible profile combinations that could actually happen, despite not being observed before. In other words, this mapping expresses a generalized combinatorial space, showing possible ways that "typical" profiles from one uncertain parameter (i.e., load) can happen together with "typical" profiles from the other (i.e., wind) and similarly for "non-casual/non-typical" ones. We note that having the capability of coupling different uncertainties without imposing a particular structure on their covariances is one of the motivations behind this work, since this capability is, to the best of our knowledge, an important component that has not been addressed in the literature until now.

3.2.2. A Clustering Approach for Selecting the Scenarios

Finally, to optimally select the scenarios that *i)* will extend the whole uncertainty space and *ii)* are the minimum (optimal) amount needed, we propose the following procedure.

From the mechanism described in section 3.1, the multivariate data points $(\tilde{\boldsymbol{\xi}}^{\ell}, \tilde{\boldsymbol{\xi}}^{w})$ representing a sampled scenario, are mapped into a unique 2-dimensional space (fig. 4), which contains coupled information for $p = \{\ell, w\}$. We can then sample this space by clustering the data in K groups and considering the centroids of these groups. Assuming that field data is homogeneous (something that empirically is known to hold for wind and load profiles, e.g., fig. 4, given their practical statistical independence), the $\widehat{\boldsymbol{\xi}} \in \Omega_s$ can be determined from any clustering algorithm. In the following considerations we use *k-means*, even

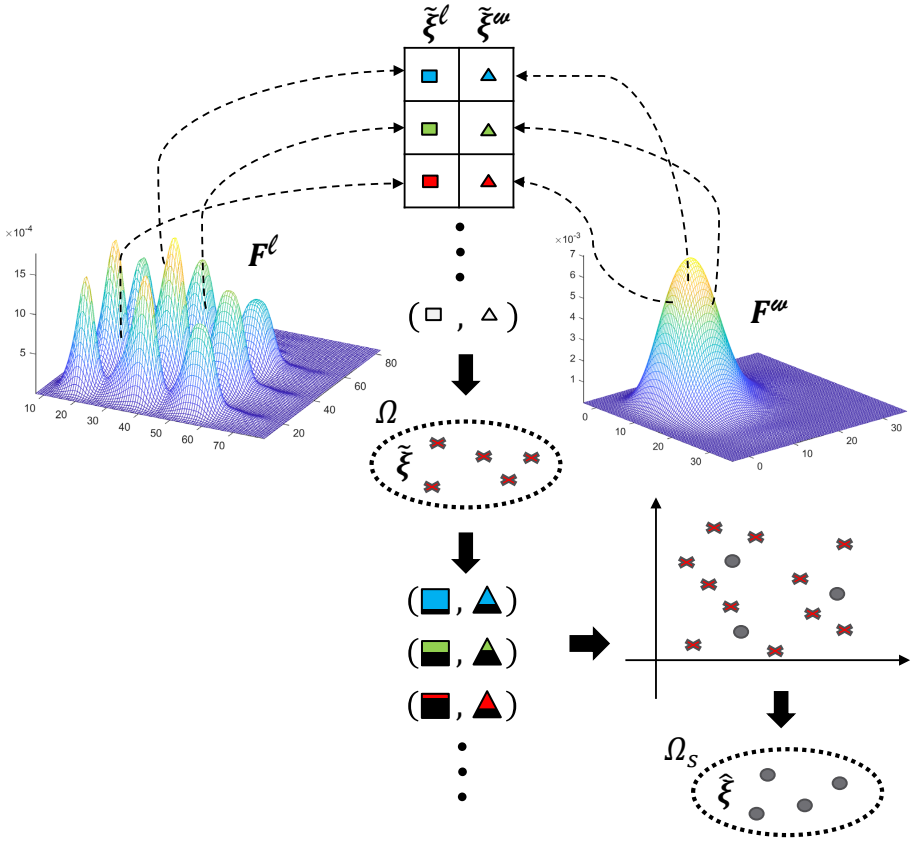


Figure 3: A graphical summary of the proposed procedure for generating the reduced scenario subset Ω_s from the sample space Ω. The starting point is sampling the joint distributions F^p, then joining the load (square) and wind (triangle) realizations together to form scenarios $\tilde{\boldsymbol{\xi}}$. These multi-dimensional vectors are then reduced to 2-dimensional points that are furthermore grouped in $|\Omega_s|$ clusters. The centroids of such clusters are then used as representatives of each such cluster, and thus used to populate the reduced subset of scenarios $\widehat{\boldsymbol{\xi}} \in \Omega_s$. Note that here the distributions F^p have been visualized just for the first two dimensions out of 24

if some other algorithms (e.g., DB-SCAN) may be used. Thus, our setup is

$$\begin{aligned} &\min_{r,\mu} \sum_{i=1}^{|\Omega|} \sum_{k=1}^{K} r_{ik} \, ||\boldsymbol{x}_i - \boldsymbol{\mu}_k|| \, |_2^2, \\ &\text{s.t.,} \;\; \boldsymbol{x}_i = \left(t_i \left(\tilde{\boldsymbol{\xi}}^{\ell}(\omega) \right), t_i \left(\tilde{\boldsymbol{\xi}}^{w}(\omega) \right) \right) \end{aligned} \tag{20}$$

where $r_{ik} \in \{0, 1\}$. Solving eq. (20) returns a set of clusters and the associated

centroids $\boldsymbol{\mu}_k$ (marked as black crosses in fig. 4). This enables selecting the data points that are closest to these centroids as scenarios $\widehat{\boldsymbol{\xi}} \in \Omega_s$ that may be used to formulate the optimization problem eq. (16). This approach enables the coupling of the various uncertainties present in the setup, as well as the ability to find scenarios that statistically cover the whole combinatory space (note that the probability of the various $\widehat{\boldsymbol{\xi}}$ can be computed based on the mass of its corresponding cluster).

The approach then needs to be completed by defining a data-driven way for selecting the number of clusters K such that the solution of problem eq. (15) statistically depends on Ω as little as possible (given a tolerance level of choice). For this, we propose applying bootstrapping techniques to assess the sample statistics of the estimated parameter, namely the problem's objective value for varying K [16]. To summarize, we propose performing tests on the stability of the sizing results based on the following considerations: 1) the stochastic process $\tilde{\boldsymbol{\xi}} \in \Omega$ is approximated by M randomly drawn sample spaces $\tilde{\boldsymbol{\xi}}_m \in \Omega_m$, from which the reduced subspaces $\widehat{\boldsymbol{\xi}}_m \in \Omega_{m,s}$ can be derived as described earlier and illustrated in fig. 3 and $|\Omega_{m,s}| = |\Omega_s| = K$. 2) the optimal number $K^* = \left|\Omega_s^*\right|$ is such that $F(\boldsymbol{x}_m^*; \Omega_{m,s}) \approx F(\boldsymbol{x}_n^*; \Omega_{n,s})\ \forall\ m, n \in 1 \dots M$, where $F\left(\boldsymbol{x}; \widehat{\boldsymbol{\xi}}\right)$ from eq. (16). This means that by choosing $\left|\Omega_s\right| = \left|\Omega_s^*\right|$ we diminish the sensitivity of the solution of the optimization problem on the subsets Ω_s.

In order to estimate Ω_s^*, numerical simulations can then be performed, for example, to obtain the results in fig. 5a and 5b. To assess the statistical dependence of the final sizing solutions on the scenarios selection algorithm, we propose to monitor the statistical properties (range Rg, interquartile range IQR and sample standard deviation s) of the M scenario trees that were created for each $\left|\Omega_s\right|$. In our considered field case, results indicate that $\left|\Omega_s^*\right| = 50$. At this point, one has all the components to code the resulting MILP problem, that in our field case was modeled in *Matlab R2020a* and solved with *Gurobi 9.0.3* in a 28 physical core multi-node cluster with Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60 Hz, 25 GB RAM.

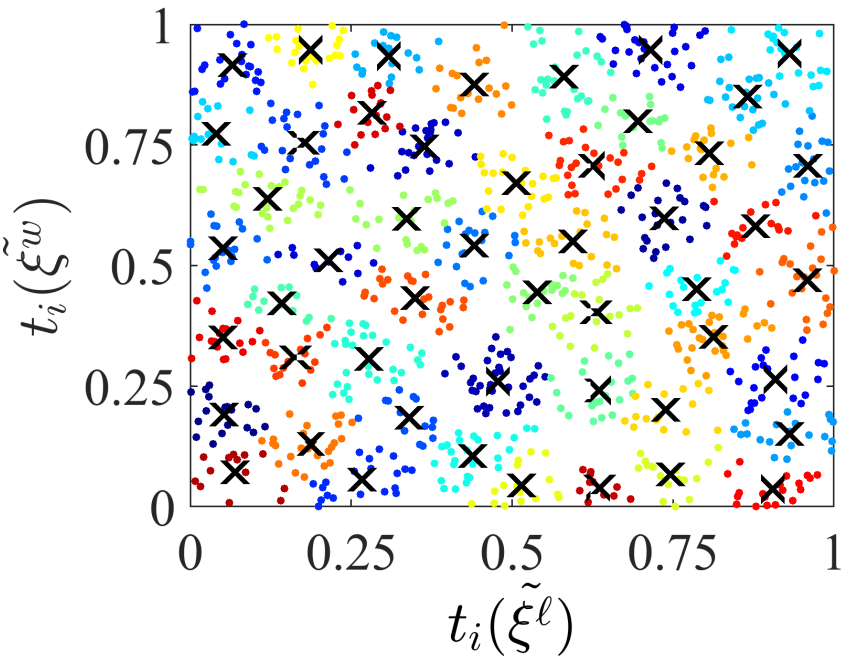
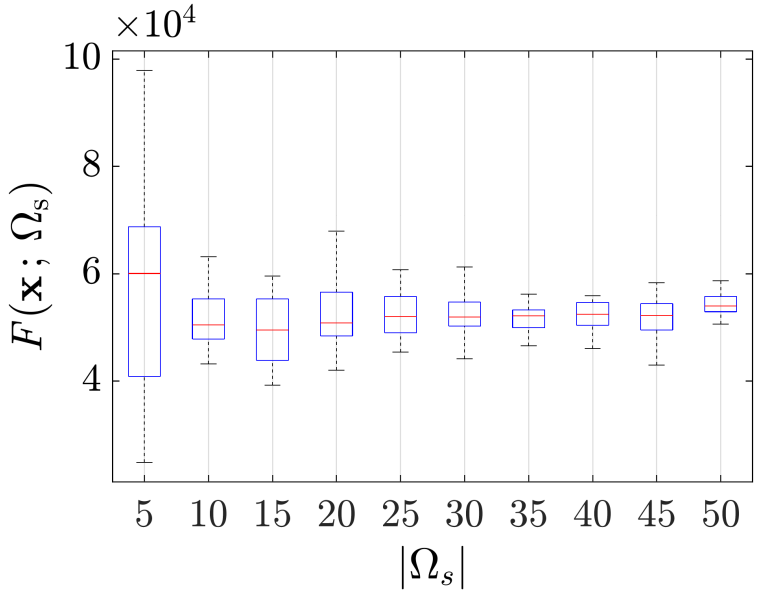
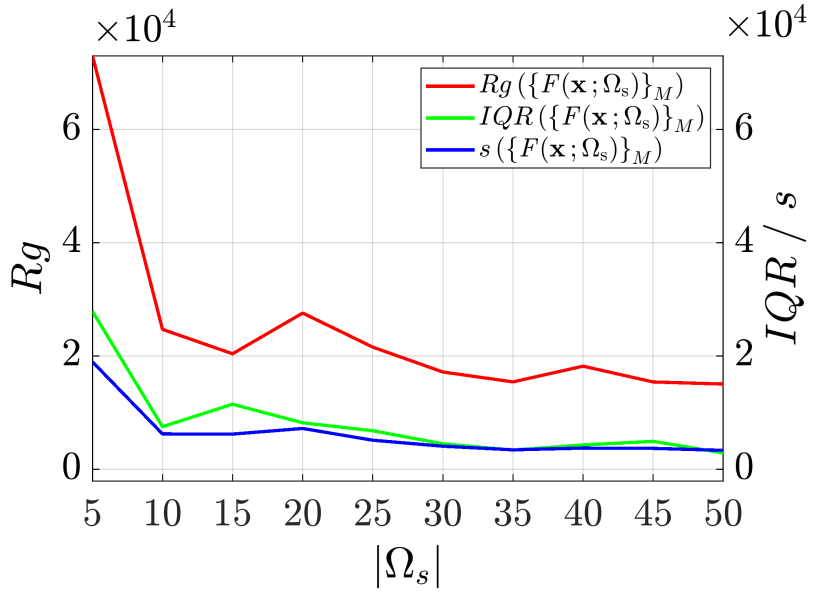


Figure 4: Snapshot of the iterative procedure for deciding $|\Omega_s^*|$. Randomly sampled profiles (load and wind) are represented as points on the 2-dimensional rank space and clustered in $|\Omega_s|$ groups, the centroids of which will populate Ω_s. Different colors correspond to different clusters (like a Voronoi diagram).



(a) Dispersion of the objective values for increasing scenario set cardinality.



(b) Convergence of the statistical indices, indicating the proper scenario set cardinality.

Figure 5: Statistical stability tests for the optimization problem objective and convergence plots for determining $\left|\Omega_s^*\right|$ through the proposed iterative procedure.

4. Results

4.1. Analysing the quality of the solution and associated expected benefits

One of our aims was to demonstrate the effectiveness of the proposed strategy and its advantage over logically simpler alternatives. To do so, we considered the risk-neutral problem ($\beta = 0$ in (16)) and performed numerical tests where several scenario subsets were produced by different typically used strategies [25, 9, 11]. These compared in terms of the stability requirement expressed in section 3.2.2. For a fair comparison $\left|\Omega_s\right| = \left|\Omega_s^*\right|$ was set for all methods. We specify that scenario reduction and selection was a necessary process, since using the whole available dataset to solve eq. (16) is computationally intractable. Thus, one needs to create a criterion for selecting scenarios that is statistically more meaningful than just random selection. That said, the following strategies were compared with the one proposed in this paper:

i. *Data*: under this naïve but straightforward strategy, we performed a random selection of combinations of load and wind profiles as observed in the available dataset. Thus, scenarios were composed of profiles combinations that were uniquely defined by the available historical dataset. As such, we represented uncertainty by conditional draws where a load pattern implies a unique wind pattern.

ii. *Random*: this is a generalization of the *Data* method, where load and wind profiles are allowed to be randomly permuted so that different combinations of load and wind are explored. Again, we used the available datasets (load and wind) but here, different profile combinations can emerge.

iii. *FFS*: this method applies a commonly used optimal scenario reduction algorithm (Fast Forward Selection). The historical profiles (load and wind) found in their corresponding datasets were individually reduced to sets of specified cardinality while redistributing the probability mass of the discarded to the preserved ones, according to their distance. Then, different random profile combinations were explored, considering their relative probabilities.

iv. *H-cl*: with this method, the individual profiles of each random variable (load / wind) were again reduced but this time using an established clustering algorithm: hierarchical clustering. The dominant patterns were distinguished from each variable and stand as representatives of their clusters, the probabilities of which are weighted based on how many profiles are assigned to each group. Then those profiles were again randomly combined to consider their relative probabilities to form scenarios.

v. *SetCorr*: this method was employed to demonstrate the effect of imposing arbitrary correlation between the random variables load and wind, as has been done in several studies [26, 27, 28, 31]. The estimated multivariate probabilities of each variable were used to generate pairs of Monte-Carlo samples (pairs of load and wind profiles) which were then ranked based on the distance of their correlation (Pearson correlation coefficient) from the nominal one. The nominal value was selected in accordance with [26, 27, 28, 31] and was set equal to 0.28. Then, the first $\left|\Omega_s^*\right|$ scenarios (pairs of load and wind profiles) with correlations closest to the nominal value were selected, weighted by the inverse of their distance from it.

vi. *Proposed*: this is the method proposed in this paper and described in section 3.

The results of the comparison are summarized in fig. 6 and table 1. fig. 6 shows that not only is the variation of the solutions gained with the proposed strategy smaller, but the extreme values are also much closer. Those effects are numerically expressed through the statistical indices defined in section 3.2.2 and presented in table 1.

From the direct sampling methods *Data, Random* we observed the high dependence of the optimal objective value to the specific subset used as input to the optimization problem, highlighting the need for a better subset selection methodology. As evident, random permutations may produce different results meaning that the optimal value may be different when using different samples. Similar effects are noticed from methods *FFS, H-cl*. In particular, *FFS* shows

higher bias of the optimal values that may be related to the fact that scenario reduction tends to prefer particular profiles that dominate the problem solution. From *H-cl* we observed a different trend characterized by the higher variance of the optimal values, meaning that the problem is very dependent on the way profiles are combined. Even though profiles can be fairly representative of their own clusters, we observed that for the specified subset cardinality, no representative combinations were identified by this methodology. From *SetCorr* we observed once more the pattern of high variability of the optimal values making clear that considering specific correlation between load and RES does not necessarily imply stable sizing solutions. This is because there could be many different profile combinations with the same correlation that lead to different optimal objective values. Finally, we saw that from our method (*Proposed*) the subset selection was performed in such a way that the dependency of the solution to the input data was minimized. This is because of the optimal exploration of the uncertainty space that was induced by considering all the different possible patterns of load, wind and their combinations. In other words, using the proposed methodology makes the optimization problem solution less sensitive (robust) to the scenario subset selection.

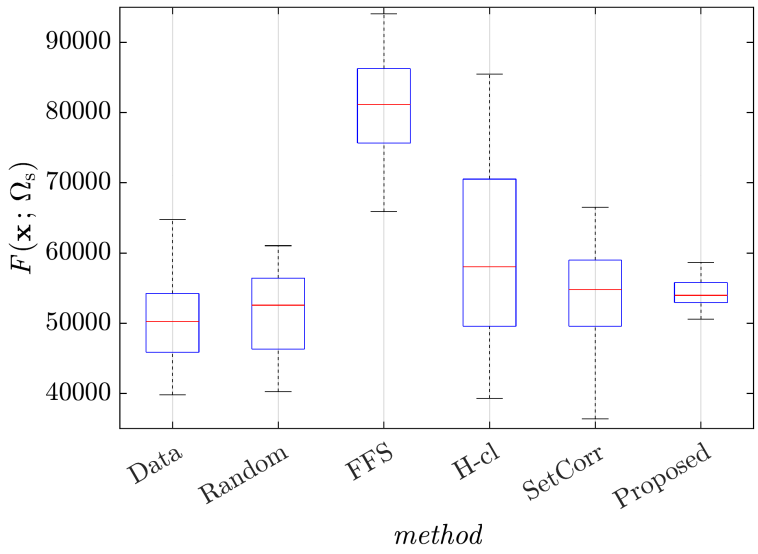


Figure 6: Comparison of optimization solution quality achieved by the proposed methodology against alternative methods of defining reduced scenarios set (load and wind) for the calculated $|\Omega_s^*|$.

Table 1: Statistics of the stability tests comparison

Statistic	Method					
	Data	*Random*	*FFS*	*H-cl*	*SetCorr*	*Proposed*
Rg	24,937	20,758	45,572	72,009	57,653	15,033
IQR	8,376	10,061	10,602	20,958	9,384	2,851
s	6,082	6,210	11,539	16,969	14,893	3,319

We also assessed the value of incorporating uncertainty as proposed in this paper. Here, we calculated the value of stochastic solution (VSS), commonly used in MILP formulations [10, 11]. For that, the expected (average) scenario was first calculated from the generated Ω_s^* set, then the deterministic problem was solved and finally the first stage solution was used to solve the original stochastic problem. It is noteworthy that with the optimal decision from the

deterministic problem there was no BESS size that could diminish the expected operational costs, i.e., $\boldsymbol{x}_d^* = 0$ (d stands for deterministic and s for stochastic). Then the VSS was equal to

$$VSS = F\left(\boldsymbol{x}_d^*; \Omega_s^*\right) - F\left(\boldsymbol{x}_s^*; \Omega_s^*\right) = 1,044 \text{ €} \text{ / day} \tag{21}$$

which means that sizing the storage under the proposed stochastic framework has the potential to reduce the daily operating costs in the reference O&G platform by €$1,044$. In summary, for the whole set Ω_s^* we got the following performance indices results:

$$\begin{aligned}
\frac{F\left(\boldsymbol{x}_d^*; \Omega_s^*\right) - F\left(\boldsymbol{x}_s^*; \Omega_s^*\right)}{F\left(\boldsymbol{x}_d^*; \Omega_s^*\right)} &= 2.09\% \\
\frac{\mathbb{E}\left[V_{CO_2}\left(\boldsymbol{x}_d^*; \Omega_s^*\right)\right] - \mathbb{E}\left[V_{CO_2}\left(\boldsymbol{x}_s^*; \Omega_s^*\right)\right]}{\mathbb{E}\left[V_{CO_2}\left(\boldsymbol{x}_d^*; \Omega_s^*\right)\right]} &= 3.64\% \\
\frac{\mathbb{E}\left[E_D\left(\boldsymbol{x}_d^*; \Omega_s^*\right)\right] - \mathbb{E}\left[E_D\left(\boldsymbol{x}_s^*; \Omega_s^*\right)\right]}{\mathbb{E}\left[E_D\left(\boldsymbol{x}_d^*; \Omega_s^*\right)\right]} &= 6.89\%
\end{aligned} \tag{22}$$

That is, we expect reduction not only in the daily operational costs (2.09%) but also in daily V_{CO_2} emissions (3.64%) and daily dumped energy E_D (6.89%).

4.2. Analysing the effects on the management of the risk

To analyze the impact of extreme (worst-case) events captured by the proposed scenario generation methodology within Ω_s^* on the solution of (16), we performed a comparative analysis by holding the same confidence level $\alpha = 0.8$ and varying $\beta \in [0, 1]$.

The results showed that, as the decision maker becomes more risk-averse, the optimal BESS size is decreased, limiting the capability of the storage to be operated such that it reduces the operational costs. A counter-benefit here is that one gets better (reduced) $CVaR$ values (table 2). However, this reduction is considerably smaller than the variation of the optimal storage decision and its impact on the expected cost. This is illustrated in fig. 7 where the CDF of the cost values $\forall\ \omega = 1 \ldots |\Omega_s^*|$ are plotted for two extreme cases: $\beta =$

0 (risk-neutral) and $\beta = 1$ (risk-averse). We observed that although the $\boldsymbol{x}^*$ decision changes drastically (and this directly impacts the shape of the cost distributions), they are very close (similar) after the given confidence level α is reached and consequently the expectation of the cost for that region ($CVaR$) essentially remains the same.

The reasoning for this can be revealed from fig. 8, where the scenarios are ranked based on their associated cost value and plotted as a contour depending on the value of $\widehat{\xi}^p$. Scenarios associated with low costs ($\omega \to 1$ after sorting) are generally associated with a combination of low loads and high wind power profiles and vice-versa (i.e., for scenarios associated with high costs, $\omega \to 50$). This indicates, in accordance with our expectations, that in cases of high demand and low wind conditions, several GTs need to be operated anyway and whatever decision one makes on the BESS size, this cannot greatly reduce the daily operational cost. Thus, provided that the dataset includes scenarios capable of capturing this phenomenon, a risk-informed decision for the storage size is not very meaningful, as worst cases cannot be avoided (faced) with any storage.

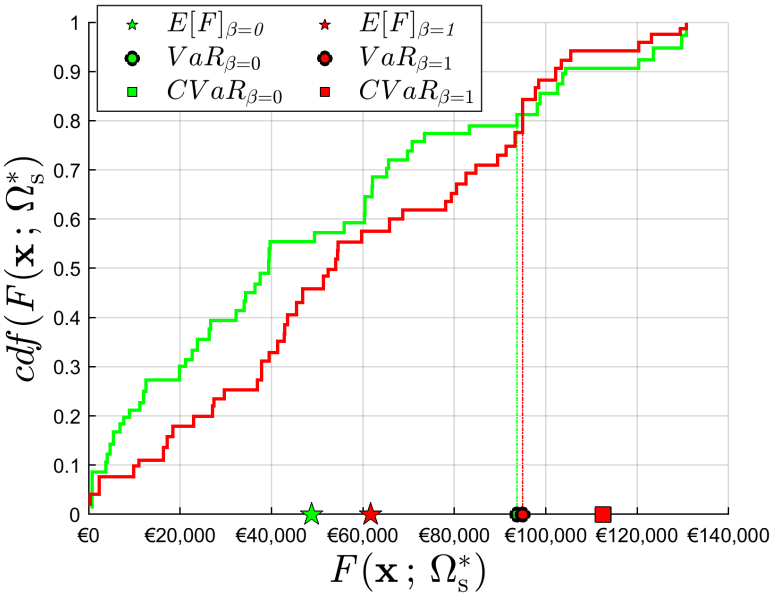
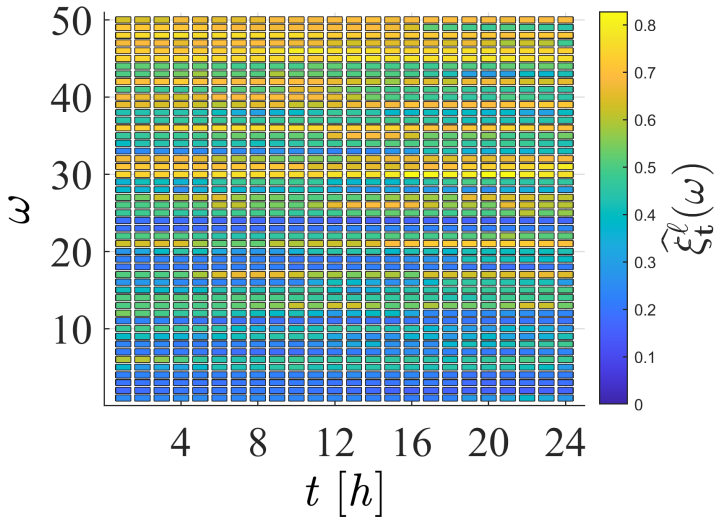
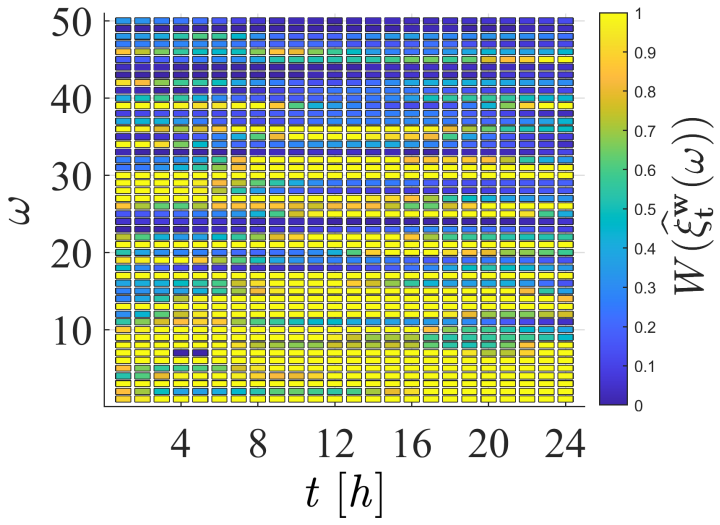


Figure 7: Cumulative distribution of the daily costs for the risk-neutral (green) and risk-averse (red) sizing problem for the optimal scenario subset. Even though risk management affects the shape of the cost distribution changing the mean value, it does not significantly affects the worst-cases tail.



(a) Load profiles



(b) Wind power profiles

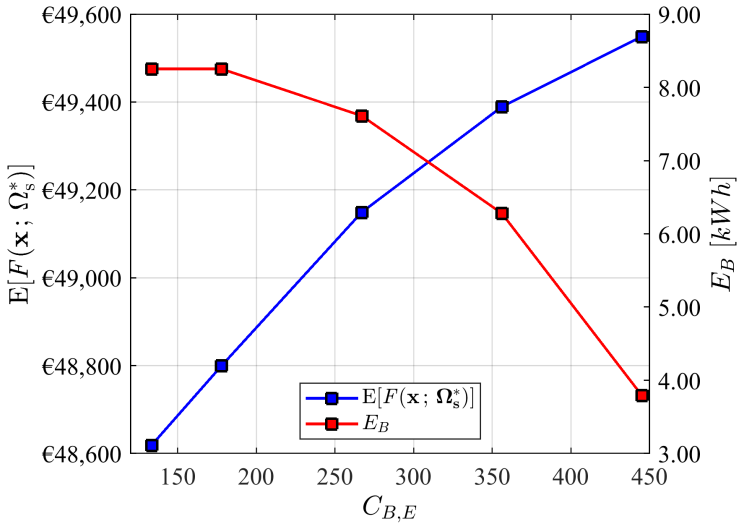
Figure 8: Optimally selected scenarios $\widehat{\boldsymbol{\xi}} \in \Omega_s^*$ from the proposed methodology, ranked based on the corresponding induced cost.

Table 2: Risk management study I: $\alpha = 0.8$

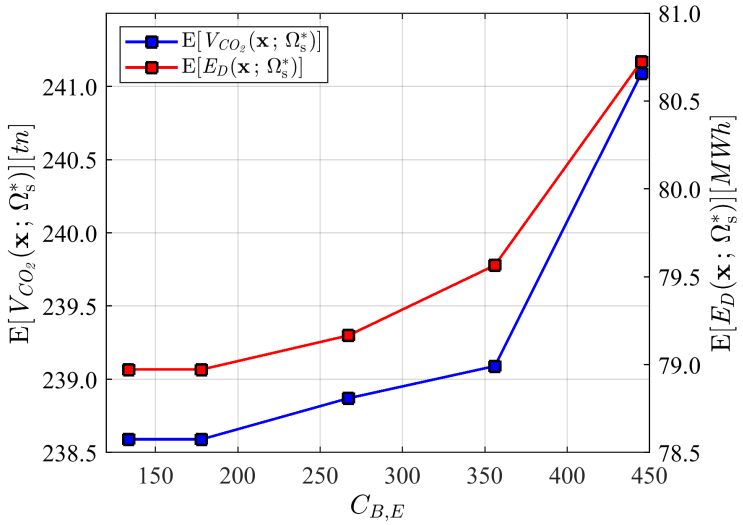
Variable	*Units*	*No BESS*	*With BESS*		
			$\beta = 0$	$\beta = 0.8$	$\beta = 0.95$
E_B	MWh	-	8.252	6.281	3.950
P_B	MW	-	4.873	4.802	3.526
F	€$/day$	49,844	48,718	48,801	49,053
$\mathbb{E}[V_{CO_2}]$	tn/day	247.21	238.59	239.09	241.21
$\mathbb{E}[E_D]$	MWh/day	84.822	78.972	79.565	80.745
VaR	€$/day$	94,892	93,691	93,555	94,925
$CVaR$	€$/day$	112,738	112,545	112,505	112,463

4.3. Analysing the sensitivity of the results on the price of the battery

Finally, in order to address the impact of the sizing parameter values on the decision $\boldsymbol{x}^*$ and on the performance indices referred to section 4.1, we performed a sensitivity analysis on $C_{B,E}$ while preserving the same $C_{B,E}/C_{B,P}$ ratio. Quantifying the effect of the BESS price is important, considering the latest and constantly decreasing trend of the cost [37, 38, 39] and the fact that the system of this case study will be implemented in the near future. fig. 9 shows the results from this analysis.



(a) Cost



(b) Emissions and dumped energy

Figure 9: Sensitivity analysis on battery price $C_{B,E}$ [€/kWh].

As expected, increasing the battery price led to smaller size and in turn higher expected daily operational costs. However, even for the highest battery

price (that refers to the conservative future price estimations), the expected costs were better than the no BESS case. The direct effect of the battery price-size dependency was also evident in the daily expected emissions and curtailed energy. Both the latter indices increased more drastically (almost exponentially) compared to the operational costs, especially following a particular battery price. Finally, below a specific battery price level we saw no significant change in the size decision and therefore in the performance indices.

5. Conclusions and future works

Deciding the techno-economically optimal size of an energy storage for isolated power systems gives rise to combinatorially hard optimization problems when considering unit commitment decisions, load and renewable uncertainties and their interactions without any arbitrary assumption. Considering all available historical data to solve the uncertain optimization problem leads to computational intractability issues, hence there is the need to reduce the number of scenarios that are used in the optimization problem. However, the scenarios selection algorithm should be such that the solution obtained is statistically as close as possible to the one that would be computed if there were no computational issues. This can be of high significance because if the final decision is highly dependent on the selected scenarios, then the confidence of the expected benefits from the energy storage is low. To deal with this problem, this paper proposes a data-informed methodology that designs and selects minimum scenarios subsets such that the solutions gained are of guaranteed statistical quality, while optimally exploring the combinatorial space of different uncertainty outcomes. The proposed methodology was benchmarked against alternative ways of deriving reduced scenario sets and achieved more stable solutions for the same problem complexity. Thus, it provided more realistic estimates of the expected benefits of integrating an energy storage system into a wind powered O&G platform, accounting for a reduction of -2.09% in daily operational costs, -3.64% in daily produced emissions and -6.89% in daily curtailment. In addition, it

was demonstrated that the subsets of scenarios that are generated with the proposed methodology include instances of uncertainty that do not favor the use of energy storage, proving that risk-constrained decisions may promote smaller investments for the storage sizing problem in isolated O&G applications without significant risk reductions.

As for future work, we envision two distinct directions. The first is towards extending the method's applicability for further practical problems by including additional types of renewable energy sources (increasing the dimensionality of the uncertainty space) and by applying the proposed methodology for sizing energy storage for interconnected micro-grids where an additional uncertain variable is the energy pricing profiles. There is then the need to verify if the favorable properties we found for our test case hold true in other power systems with different profile characteristics. The second direction is to find analytical results on the properties of the methodology. The most important one is likely finding results that couple the number of scenarios to be used in the programs with the statistical stability of the sizing results. These theorems may indeed be useful to decide the number of scenarios to be used without performing time consuming stability simulations, but rather using theoretically grounded formula.

6. Appendix

References

[1] Y. Yang, S. Bremner, C. Menictas, M. Kay, Battery energy storage system size determination in renewable energy systems: A review, Renewable and Sustainable Energy Reviews 91 (2018) 109–125. doi:10.1016/j.rser.2018.03.047.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364032118301436

[2] R. Machlev, N. Zargari, N. R. Chowdhury, J. Belikov, Y. Levron, A review of optimal control methods for energy storage systems - energy trading,

Table 3: Parameter Values

Parameter	**Symbol**	**Value**	**Units**
Natural Gas and Gas Turbines			
Natural Gas density	ρ	0.77	$\left[\frac{kg}{m^3}\right]$
CO_2 combustion coefficient	μ	2.682	$\left[\frac{kg_{CO_2}}{kg_f}\right]$
maximum Gas Turbine power	$\overline{P}^g_{GT}$	20.2	$[MW]$
affine fuel consumption coefficient	α_g	172.50	$\left[\frac{kg_f}{MW}\right]$
affine fuel consumption constant	β_g	729.20	$\left[\frac{kg_f}{h}\right]$
operation and maintenance cost	$C_{O\&M}$	5.53	$\left[\frac{€}{MW}\right]$
gas turbine startup cost	C^{start}_{GT}	1217	$\left[\frac{€}{\text{start}}\right]$
CO_2 emissions tax	C_{CO_2}	0.07	$\left[\frac{€}{kg_{CO_2}}\right]$
Natural gas wholesale price	C_{NG}	0.24	$\left[\frac{€}{m^3}\right]$
Wind Farm			
cut in speed	w_{ci}	3	$\left[\frac{m}{s}\right]$
cut off speed	w_{ci}	25	$\left[\frac{m}{s}\right]$
rated wind speed	w_n	12	$\left[\frac{m}{s}\right]$
rated wind turbine power	P^n_w	8	$\left[\frac{m}{s}\right]$
wind farm turbines	N_w	5	$[-]$
Battery Energy Storage System			
interest rate	r	7	$[\%]$
battery lifetime	L	8	$[years]$
battery energy cost	$C_{B,E}$	178.00	$\left[\frac{€}{kWh}\right]$
battery power cost	$C_{B,P}$	89.00	$\left[\frac{€}{kW}\right]$

energy balancing and electric vehicles, Journal of Energy Storage 32 (2020) 101787. doi:10.1016/j.est.2020.101787.
URL http://www.sciencedirect.com/science/article/pii/S2352152X20316248

[3] A. Soroudi, T. Amraee, Decision making under uncertainty in energy systems: State of the art, Renewable and Sustainable Energy Reviews 28 (2013) 376–384. doi:10.1016/j.rser.2013.08.039.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364032113005790

[4] L. A. Wong, V. K. Ramachandaramurthy, P. Taylor, J. Ekanayake, S. L. Walker, S. Padmanaban, Review on the optimal placement, sizing and control of an energy storage system in the distribution network, Journal of Energy Storage 21 (2019) 489–504. doi:10.1016/j.est.2018.12.015.
URL https://linkinghub.elsevier.com/retrieve/pii/S2352152X18303803

[5] M. Hannan, M. Faisal, P. Jern Ker, R. Begum, Z. Dong, C. Zhang, Review of optimal methods and algorithms for sizing energy storage systems to achieve decarbonization in microgrid applications, Renewable and Sustainable Energy Reviews 131 (2020) 110022. doi:10.1016/j.rser.2020.110022.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364032120303130

[6] M. Cao, Q. Xu, J. Cai, B. Yang, Optimal sizing strategy for energy storage system considering correlated forecast uncertainties of dispatchable resources, International Journal of Electrical Power & Energy Systems 108 (2019) 336–346. doi:10.1016/j.ijepes.2019.01.019.
URL https://linkinghub.elsevier.com/retrieve/pii/S0142061518324773

[7] C. Wang, Z. Zhang, O. Abedinia, S. G. Farkoush, Modeling and analysis of a microgrid considering the uncertainty in renewable energy resources, energy storage systems and demand management in electrical retail market, Journal of Energy Storage 33 (2021) 102111. doi:10.1016/j.est.2020.102111.
URL https://linkinghub.elsevier.com/retrieve/pii/S2352152X20319411

[8] V. Jani, H. Abdi, Optimal allocation of energy storage systems considering wind power uncertainty, Journal of Energy Storage 20 (2018) 244–253.

doi:10.1016/j.est.2018.09.017.
URL https://linkinghub.elsevier.com/retrieve/pii/S2352152X18303761

[9] J. Li, J. Zhou, B. Chen, Review of wind power scenario generation methods for optimal operation of renewable energy systems, Applied Energy 280 (2020) 115992. doi:10.1016/j.apenergy.2020.115992.
URL https://linkinghub.elsevier.com/retrieve/pii/S0306261920314380

[10] A. J. Conejo, M. Carrión, J. M. Morales, Decision making under uncertainty in electricity markets, International series in operations research & management science, Springer, New York, 2010, oCLC: ocn662409210.

[11] C. Li, I. E. Grossmann, A Review of Stochastic Programming Methods for Optimization of Process Systems Under Uncertainty, Front. Chem. Eng. 2, publisher: Frontiers. doi:10.3389/fceng.2020.622241.
URL https://www.frontiersin.org/articles/10.3389/fceng.2020.622241/full

[12] X. Kong, J. Xiao, D. Liu, J. Wu, C. Wang, Y. Shen, Robust stochastic optimal dispatching method of multi-energy virtual power plant considering multiple uncertainties, Applied Energy 279 (2020) 115707. doi:10.1016/j.apenergy.2020.115707.
URL https://linkinghub.elsevier.com/retrieve/pii/S0306261920312010

[13] M. Aghamohamadi, A. Mahmoudi, M. H. Haque, Two-Stage Robust Sizing and Operation Co-Optimization for Residential PV–Battery Systems Considering the Uncertainty of PV Generation and Load, IEEE Transactions on Industrial Informatics 17 (2) (2021) 1005–1017, conference Name: IEEE Transactions on Industrial Informatics. doi:10.1109/TII.2020.2990682.

[14] Z. Guo, W. Wei, L. Chen, R. Xie, S. Mei, Sizing energy storage to reduce renewable power curtailment considering network power flows: a distributionally robust optimisation approach, IET Renewable Power Generation 14 (16) (2020) 3273–3280, _eprint: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/iet-rpg.2020.0354. doi:10.1049/iet-rpg.2020.0354.

URL https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-rpg.2020.0354

[15] M. Zhang, X. Ai, J. Fang, W. Yao, W. Zuo, Z. Chen, J. Wen, A systematic approach for the joint dispatch of energy and reserve incorporating demand response, Applied Energy 230 (2018) 1279–1291. doi:10.1016/j.apenergy.2018.09.044.
URL https://www.sciencedirect.com/science/article/pii/S0306261918313539

[16] M. Kaut, S. W. Wallace, Evaluation of scenario-generation methods for stochastic programming, Pacific Journal of Optimization 3 (2) (2007) 257–271.

[17] E. Du, N. Zhang, C. Kang, Q. Xia, Scenario Map Based Stochastic Unit Commitment, IEEE Transactions on Power Systems 33 (5) (2018) 4694–4705, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2018.2799954.

[18] J. Dupačová, N. Gröwe-Kuska, W. Römisch, Scenario Reduction in Stochastic Programming: An Approach Using Probability Metrics.

[19] J. Hu, H. Li, A New Clustering Approach for Scenario Reduction in Multi-Stochastic Variable Programming, IEEE Transactions on Power Systems 34 (5) (2019) 3813–3825, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2019.2901545.

[20] J. Wang, M. Shahidehpour, Z. Li, Security-Constrained Unit Commitment With Volatile Wind Power Generation, IEEE Transactions on Power Systems 23 (3) (2008) 1319–1327, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2008.926719.

[21] A. Jalali, M. Aldeen, Risk-Based Stochastic Allocation of ESS to Ensure Voltage Stability Margin for Distribution Systems, IEEE Transactions on

Power Systems 34 (2) (2019) 1264–1277, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2018.2873774.

[22] P. Pinson, G. Papaefthymiou, B. Klockl, J. Verboomen, Dynamic sizing of energy storage for hedging wind power forecast uncertainty, in: 2009 IEEE Power & Energy Society General Meeting, IEEE, Calgary, Canada, 2009, pp. 1–8. doi:10.1109/PES.2009.5275816.
URL http://ieeexplore.ieee.org/document/5275816/

[23] H. Valizadeh Haghi, S. Lotfifard, Spatiotemporal Modeling of Wind Generation for Optimal Energy Storage Sizing, IEEE Transactions on Sustainable Energy 6 (1) (2015) 113–121, conference Name: IEEE Transactions on Sustainable Energy. doi:10.1109/TSTE.2014.2360702.

[24] Y. Chen, Y. Wang, D. Kirschen, B. Zhang, Model-Free Renewable Scenario Generation Using Generative Adversarial Networks, IEEE Trans. Power Syst. 33 (3) (2018) 3265–3275. doi:10.1109/TPWRS.2018.2794541.
URL http://ieeexplore.ieee.org/document/8260947/

[25] P. Seljom, L. Kvalbein, L. Hellemo, M. Kaut, M. M. Ortiz, Stochastic modelling of variable renewables in long-term energy models: Dataset, scenario generation & quality of results, Energy (2021) 121415doi:10.1016/j.energy.2021.121415.
URL https://linkinghub.elsevier.com/retrieve/pii/S0360544221016637

[26] X. Yang, C. Xu, H. He, W. Yao, J. Wen, Y. Zhang, Flexibility Provisions in Active Distribution Networks With Uncertainties, IEEE Transactions on Sustainable Energy (2020) 1–1Conference Name: IEEE Transactions on Sustainable Energy. doi:10.1109/TSTE.2020.3012416.

[27] M. R. M. Cruz, D. Z. Fitiwi, S. F. Santos, S. J. P. S. Mariano, J. P. S. Catalão, Multi-Flexibility Option Integration to Cope With Large-Scale Integration of Renewables, IEEE Transactions on Sustainable Energy 11 (1) (2020) 48–60, conference Name: IEEE Transactions on Sustainable Energy. doi:10.1109/TSTE.2018.2883515.

[28] S. F. Santos, D. Z. Fitiwi, M. Shafie-Khah, A. W. Bizuayehu, C. M. P. Cabrita, J. P. S. Catalão, New Multistage and Stochastic Mathematical Model for Maximizing RES Hosting Capacity—Part I: Problem Formulation, IEEE Transactions on Sustainable Energy 8 (1) (2017) 304–319, conference Name: IEEE Transactions on Sustainable Energy. doi:10.1109/TSTE.2016.2598400.

[29] F. Pourahmadi, J. Kazempour, C. Ordoudis, P. Pinson, S. H. Hosseini, Distributionally Robust Chance-Constrained Generation Expansion Planning, IEEE Transactions on Power Systems 35 (4) (2020) 2888–2903, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2019.2958850.

[30] J. H. Yi, R. Cherkaoui, M. Paolone, Optimal Allocation of ESSs in Active Distribution Networks to achieve their Dispatchability, IEEE Transactions on Power Systems (2020) 1–1Conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2020.3025991.

[31] G. Sinden, Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand, Energy Policy 35 (1) (2007) 112–127. doi:10.1016/j.enpol.2005.10.003.
URL https://www.sciencedirect.com/science/article/pii/S0301421505002752

[32] Z. Shu, P. Jirutitijaroen, Latin Hypercube Sampling Techniques for Power Systems Reliability Analysis With Renewable Energy Sources, IEEE Transactions on Power Systems 26 (4) (2011) 2066–2073, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2011.2113380.

[33] X. Xu, Z. Yan, M. Shahidehpour, Z. Li, M. Yan, X. Kong, Data-Driven Risk-Averse Two-Stage Optimal Stochastic Scheduling of Energy and Reserve With Correlated Wind Power, IEEE Transactions on Sustainable Energy 11 (1) (2020) 436–447, conference Name: IEEE Transactions on Sustainable Energy. doi:10.1109/TSTE.2019.2894693.

[34] R. T. Rockafellar, S. Uryasev, Optimization of conditional value-at-risk, JOR 2 (3) (2000) 21–41. doi:10.21314/JOR.2000.038.
URL http://www.risk.net/journal-of-risk/technical-paper/2161159/optimization-conditional-value-risk

[35] O. Sadeghian, A. Oshnoei, R. Khezri, S. Muyeen, Risk-constrained stochastic optimal allocation of energy storage system in virtual power plants, Journal of Energy Storage 31 (2020) 101732. doi:10.1016/j.est.2020.101732.
URL https://linkinghub.elsevier.com/retrieve/pii/S2352152X20315693

[36] Z. Zhang, H. Qin, J. Li, Y. Liu, L. Yao, Y. Wang, C. Wang, S. Pei, P. Li, J. Zhou, Operation rule extraction based on deep learning model with attention mechanism for wind-solar-hydro hybrid system under multiple uncertainties, Renewable Energy 170 (2021) 92–106. doi:10.1016/j.renene.2021.01.115.
URL https://linkinghub.elsevier.com/retrieve/pii/S0960148121001221

[37] M. Wilshire, Future trends in energy - Bloomberg New Energy Finance (Aug. 2018).

[38] I. R. E. Agency, Electricity storage and renewables: Costs and markets to 2030, Tech. rep., IRENA, Abu Dhabi (Oct. 2017).

[39] Battery storage to drive the power system transition, Summary, ECOFYS Germany GmbH (2018).
URL https://ec.europa.eu/energy/sites/ener/files/report-_battery_storage_to_drive_the_power_system_transition.pdf