

Quasi-Monte Carlo technique in global sensitivity analysis of wind resource assessment with a study on UAE

Cite as: J. Renewable Sustainable Energy **11**, 053303 (2019); <https://doi.org/10.1063/1.5120035>

Submitted: 15 July 2019 . Accepted: 19 September 2019 . Published Online: 07 October 2019

Olga Tsvetkova, and Taha B. M. J. Ouarda



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Wind turbine wake intermittency dependence on turbulence intensity and pitch motion](#)

Journal of Renewable and Sustainable Energy **11**, 053302 (2019); <https://doi.org/10.1063/1.5097829>

[Wind energy at remote islands in arctic region—A case study of Solovetsky islands](#)

Journal of Renewable and Sustainable Energy **11**, 053304 (2019); <https://doi.org/10.1063/1.5110756>

[Oxidation states in perovskite layers formed using various deposition techniques](#)

Journal of Renewable and Sustainable Energy **11**, 053504 (2019); <https://doi.org/10.1063/1.5108656>

CAPTURE WHAT'S POSSIBLE
WITH OUR NEW PUBLISHING ACADEMY RESOURCES

Learn more →



Quasi-Monte Carlo technique in global sensitivity analysis of wind resource assessment with a study on UAE

Cite as: J. Renewable Sustainable Energy **11**, 053303 (2019); doi: 10.1063/1.5120035

Submitted: 15 July 2019 · Accepted: 19 September 2019 ·

Published Online: 7 October 2019



View Online



Export Citation



CrossMark

Olga Tsvetkova^{a)} and Taha B. M. J. Ouarda

AFFILIATIONS

Canada Research Chair in Statistical Hydro-Climatology, INRS-ETE, 490 rue de la Couronne, Quebec, Quebec G1K 9A9, Canada

^{a)} Author to whom correspondence should be addressed: olga.tsvetkova@ete.inrs.ca. Telephone: +1 418 654 2691.

ABSTRACT

The present paper bridges mathematical modeling and wind resource assessment (WRA). Sensitivity analysis (SA) links portions of output variance to the variance in each input variable. Global SA (GSA) explores inputs globally. One-at-a-time SA is dominating in WRA, while GSA is often overlooked. Compared to traditional methods, GSA offers potential improvement by the means of the quasi-Monte Carlo (QMC) technique with its elaborate sampling designs enabling faster convergence. The main novelty of this work is twofold: the use of QMC in WRA and Sobol method as a variance-based GSA method in WRA. This paper is among a few using GSA in WRA. Two case studies were conducted. One shows that QMC with sampling based on Sobol sampling outperforms Latin hypercube sampling and pseudorandom sampling. It also shows that in terms of accuracy of results, the brute-force calculation of Sobol sensitivity indices (SI) should be used whenever the model allows it; otherwise, SI can be estimated. Another case study demonstrates a valid GSA study for WRA at Masdar City, United Arab Emirates. The results suggest that the influence of the variance in Weibull parameters on annual energy production (AEP) might be overestimated, while found to be responsible for 2% of AEP uncertainty, and the influence of the variance in air density might be overlooked, while found to account for 94%. WRA studies would benefit greatly from using the QMC and Sobol method. The Sobol method is a universal GSA method, providing valid results for nonlinear models typical for WRA, and QMC provides global scalability, model independence, and flexibility in uncertainty quantification.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5120035>

NOMENCLATURE

a	Availability of a wind turbine for energy production	GOF	Goodness-of-fit
A	$N \times k$ sample matrix of input variables used	GSA	Global sensitivity analysis
$a_{j,k}$	Binary coefficients in Sobol sequence generation	IRR	Internal rate of return
$A_B^{(i)}$	$N \times k$ matrix where all columns except for the i -th column come from A and the i -th column comes from B	k	Number of input variables
AEP	Annual energy production at Masdar City	LCOE	Levelized cost of energy
B	$N \times k$ sample matrix of input variables used	LHS	Latin hypercube sampling
c	Shape parameter of the two-parameter Weibull distribution	$m_{j,k}$	Sobol sequence generation
CFD	Computational fluid dynamics	MAE	Mean absolute error as defined in Ref. 2
ECDF	Empirical cumulative distribution function	$MAE(\hat{S}_i)$	Mean absolute error of first-order sensitivity indices ²
EL	Electric losses	$MAE(\hat{S}_T)$	Mean absolute error of total effect sensitivity indices ²
FiT	Feed-in tariff	MC	Monte Carlo
FOWF	Floating offshore windfarm	MCP	Measure-correlate-predict
g_i	Terms of analytical G-function used as the test model	MLE	Maximum likelihood estimation
		N	Sample size used in MC/QMC experiments (number of model runs)
		N_r	Number of replicas of a Monte Carlo/quasi-Monte Carlo experiment
		NPV	Net present value

OAT	One-at-a-time
OWF	Offshore wind farm
O&M	Operation and maintenance
p	Atmospheric pressure
PDF	Probability density function
PRNG	Pseudorandom number generator
PRS	Pseudorandom sampling
QMC	Quasi-Monte Carlo
R_{spec}	Specific gas constant of dry air, $R_{spec} = 287.058 \text{ J/kg K}$
s	Number of strata in the case of Latin hypercube sampling
S_i	Sobol first-order sensitivity indices ³
s_j	Order of a polynomial used in Sobol sequence generation
\hat{S}_i	Estimates of Sobol first-order sensitivity indices with the use of a Saltelli estimator ²
\underline{S}_{T_i}	Sobol total effect sensitivity indices ³
\bar{S}_{T_i}	Estimates of Sobol total effect sensitivity indices with the use of a Jansen estimator ^{1,2}
SA	Sensitivity analysis
SBSS	Sampling based on Sobol sequences
SI	Sensitivity indices
SS	Stratified sampling
t	Absolute air temperature
T	$2N \times 2k$ sample matrix of input variables used to implement radial design described by Saltelli in Ref. 2
UA	Uncertainty analysis
UAE	United Arab Emirates
V	Total variance of a model
V_i	A variance based on the first-order effect for a factor X_i
$v_{j,k}$	Direction numbers in Sobol sequence generation
\bar{V}_{T_i}	A variance based on the total effect for a factor X_i
w	Wind speed time series
WRA	Wind resource assessment
WRF	Weather research and forecasting
W2	The two-parameter Weibull distribution
X	$N \times k$ matrix of all input variables
X_i	Generic input variable, vector of $N \times 1$
$x_{i,j}$	j th component of the i th point in a Sobol sequence
$X_{\sim i}$	$N \times (k - 1)$ matrix of all factors but X_i
Y	Vector model output of $N \times 1$
λ	Scale parameter of the two-parameter Weibull distribution
ρ	Air density

I. INTRODUCTION

Sensitivity analysis (SA) is a study of apportioning the uncertainty in the output variable of interest to the uncertainty in each input variable.³ Although SA is an important practice related to uncertainty analysis (UA)³ and is meant to complement UA,⁴ SA may be performed perfunctorily.^{4,5} The problems with SA practices at large are unveiled across engineering fields by recent systematic reviews of Feretti *et al.*⁵ and Saltelli *et al.*⁶ This paper looks at how SA is performed in wind resource assessment (WRA) and offers possible improvement to current practice. Global SA (GSA) explores each input variable over its entire range of interest and provides valid

results for any model.³ The main drawback of GSA as Wagener and Pianosi have pointed out “is the possibly significant impact of the chosen input distributions, which should be carefully scrutinized.”⁷ Although “a sign of improvements in the take up of GSA” is marked, GSA is “still applied by a minority of researchers.”⁵ Instead, easy to implement one-at-a-time (OAT) methods⁸ that vary one parameter while keeping all others fixed are widely used^{4–6} although one should be cautious of such methods as they are “predicated on assumptions of model linearity.”⁶ Saltelli and Annoni voiced reasonable concern about OAT SA results by publishing geometric proof of its inefficiency as “a convincing ... argument against OAT.”⁴ Nonetheless, researchers “seem reluctant to abandon this [OAT] practice.”⁴ Although an attempt to popularize GSA and increase SA quality in the environmental modeling community has already been made by Pianosi *et al.* by introducing a Matlab toolbox for GSA,⁹ it has come to the authors’ attention that OAT is commonly used in WRA,^{10–15} and hence, the authors intend to continue this effort.

It is expensive to explore variable input space as the number of dimensions grows; yet, it is not a reason to practice OAT for nonlinear models. In fact, GSA (variance-based or other) offers ways to globally explore high-dimension models with the use of the quasi-Monte Carlo (QMC) technique.¹⁶ To conduct a variance-based GSA study, it is necessary to run a Monte Carlo (MC) or a QMC experiment. MC is a “universal uncertainty quantification tool with applications across all engineering fields with only computational cost being a limiting factor.”¹⁷ QMC helps mitigate this barrier by using elaborate sampling designs as the results of a QMC experiment converge faster than those of a MC experiment.^{17–19} The difference between a MC and a QMC experiment lies in how random numbers are generated. A MC experiment uses common pseudorandom sampling (PRS), while QMC uses other sampling strategies, for example, Latin hypercube sampling (LHS) or sampling based on low discrepancy Sobol sequences (SBSS).

WRA mostly deals with estimates (often point^{20–22}) of wind power or annual energy production (AEP) cubically proportional to the wind speed. MC in WRA is usually used for windfarm layout optimization^{23,24} or profitability UA^{10,25–29} often with no SA accompanying. In wind turbine specific studies, MC was also used for stochastic blade flutter analysis,³⁰ turbine reliability,³¹ and fatigue load^{32,33} assessments. To the best of the authors’ knowledge, SA in WRA commonly uses OAT methods^{10–15} that disregard model nonlinearity.^{4–6} Only a few WRA studies have used GSA.^{34–36} Di *et al.* used multivariate adaptive regression for optimizing numerous parameters of a weather research and forecasting (WRF) model for more accurate wind power prediction.³⁴ McKay *et al.* used the Fourier amplitude sensitivity test for wind power output.³⁵ Ulker *et al.* studied sensitivity-guided decision-making for wind farm micro-siting, and Sobol method was part of the study.³⁶ It is worth mentioning a broader study on the levelized cost of energy (LCOE) of different energy sources (including onshore and offshore wind) that used the Sobol method.³⁷

The main reason for using QMC in WRA is that it enables the application of GSA methods, such as the Sobol method, which account for model nonlinearity. There are also several benefits to using QMC in general, such as global scalability, model independence, and flexibility. The first benefit is that QMC is what makes GSA practically feasible in the case of numerous input variables. Second, QMC provides a way to conduct a GSA regardless of how computationally demanding the model of interest is. In application to WRA, one can model wind

with a distribution (Weibull^{38,39} or other^{40,41}) or using more time-consuming computational fluid dynamics (CFD)^{42–44} or WRF^{45,46} models. The third benefit is flexibility. One might have limited data about the variables of the output model or assumptions about the underlying distributions of input variables. One might use parametric or nonparametric models to describe input variables. Whichever the case, one can incorporate all the knowledge about a particular site available into a GSA study with the use of QMC. As more data or knowledge becomes available, the GSA study can be elaborated and more accurate results can be obtained. There is also a side benefit to using QMC for GSA—the same QMC experiments yield both UA and GSA results. UA shows how uncertain the estimate of the output variable is and GSA apportions output uncertainty to uncertainty in the input variables.

The QMC technique has been applied in wind related studies,^{18,33,47} but not to WRA specifically. The Sobol method as a variance-based GSA, the focus of this study, has received little attention in WRA.³⁶ Estimation of Sobol indices has not yet been used in WRA. Therefore, the main novelty of this work is twofold: the use of QMC in WRA and Sobol method for GSA in WRA, as both have not received much attention. The present paper aims to fill in this gap by bridging mathematical modeling and WRA, as well as popularize GSA and exemplify a GSA study for WRA. Two case studies were carried out: one using a test model to show that QMC is superior to MC and compare the brute force Sobol method with its estimation and the other using United Arab Emirates (UAE) data to demonstrate how a GSA (and UA) using the QMC and Sobol method could be set up in the context of WRA.

This paper is organized into four main sections: Introduction, Methods, Results, and Conclusions and Future Research. Section I presents the problem associated with SA in WRA. Section II contains three subsections: sampling techniques used in MC/QMC, variance-based GSA, and experimental design. The sampling techniques subsection summarizes sampling techniques used throughout this paper, namely, PRS, LHS, and SBSS. The variance-based GSA subsection reviews the brute force calculation of Sobol first-order and total effect sensitivity indices (SIs) and their estimation. The experimental design subsection contains two parts: the test model and the WRA model in the Masdar City case study. It gives details on the setup of the experimental part of this work and methods used in WRA such as the Weibull distribution to fit wind speed data and to calculate AEP. The assumptions about underlying distributions of input variables to the WRA model are given in this subsection. Section III provides the results of GSA for the test and WRA models. The test model results include the comparison of sampling strategies in the light of GSA in WRA. The WRA model results contain the results of GSA and UA for WRA at Masdar City, UAE. Section IV presents conclusions made along with the suggestions for possible directions for future research.

II. METHODS

In this section, the methods used are explained, the experimental design of the two case studies is presented, and the details about the data and the assumptions used in the UAE case study are provided.

A. Sampling techniques in MC/QMC experiments

Sampling is the foundation of any MC/QMC experiment. It is important to understand how different sampling strategies affect the results of MC/QMC simulations. As mentioned in the introduction, a QMC experiment differs from a MC experiment in the way random

numbers are generated. A MC uses PRS, while QMC uses sampling strategies other than PRS which aim to reduce the computational cost of MC by enabling faster convergence of results. The computational cost of MC is an important barrier to its applicability, but in most cases, it can be overcome with a good sampling strategy.¹⁷

1. Pseudorandom sampling (PRS)

Pseudorandom sampling (PRS) generates a sequence of random numbers using a pseudorandom number generator (PRNG). PRNG is an algorithm for generating a sequence of numbers determined by the initial key, called a seed, for generating a sequence.⁴⁶ The main weakness of random sampling in relation to the problem of GSA is poor dimensional distribution. A random sample typically has gaps and clusters³ which can play a significant role in the results of GSA, meaning that it does not provide a uniform exploration of the variable space. Figure 1 exhibits a two-dimensional pseudorandom sample, and one can see the gaps and clusters of sampled points with the naked eye.

2. Latin hypercube sampling (LHS)

In a GSA problem setting, a homogeneous sample is a critical point as it allows a homogeneous input space exploration, and stratified sampling (SS) aims at avoiding gaps and clusters typical in PRS. Figure 2 exhibits a large two-dimensional LHS sample ($N = 1000$). More homogeneity in Fig. 2 can be observed compared to the pseudorandom sample in Fig. 1.

As a way to overcome the flaws of PRS, stratified sampling (SS) was invented. SS is sampling with a PRNG but at different intervals or strata, providing a somewhat more “homogeneously” distributed sample. LHS is a particular kind of SS. The main goal of LHS is to make sure that each variable is individually stratified over s levels ($s > 2$) and that each level contains an equal number of points.⁴⁸ Figure 3 emphasizes this property. Each of the fifteen horizontal and vertical strata in Fig. 3 contains strictly one point. The rectangles in Fig. 3 outline this very property.

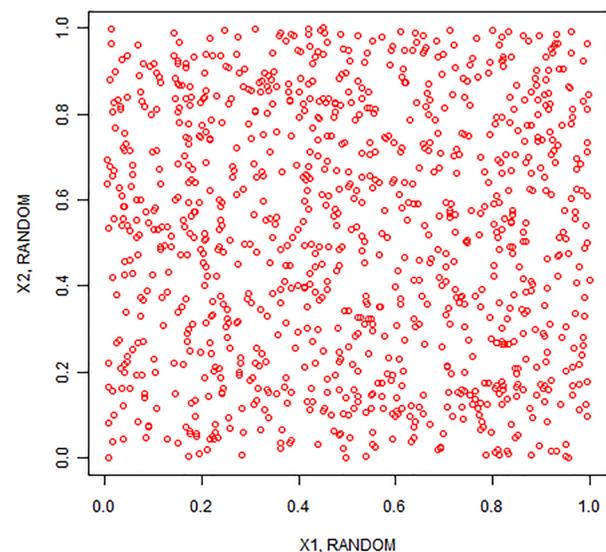


FIG. 1. Two-dimensional pseudorandom sample.

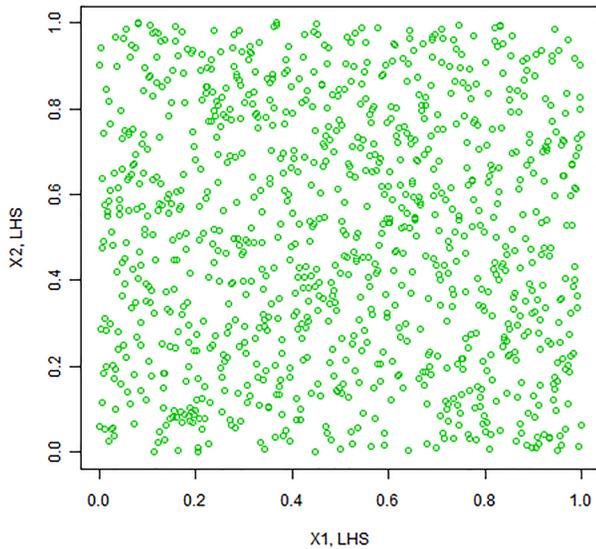


FIG. 2. A two-dimensional LHS sample (N = 1000).

The conditions of LHS require that s divides the sample into size N .³ As a result of such sampling design, LHS has some attractive advantages:³

1. The estimate of the mean converges faster to the true value compared to the one corresponding to a random sample. This means that the smaller sample size is necessary for QMC experiment to converge compared to MC. When the model is computationally expensive, it adds up to a significant time difference between MC and QMC experiments;
2. When the output function is monotonic in each of the inputs X_i , a sample would have less uncertainty with LHS compared to PRS at any sample size N .³

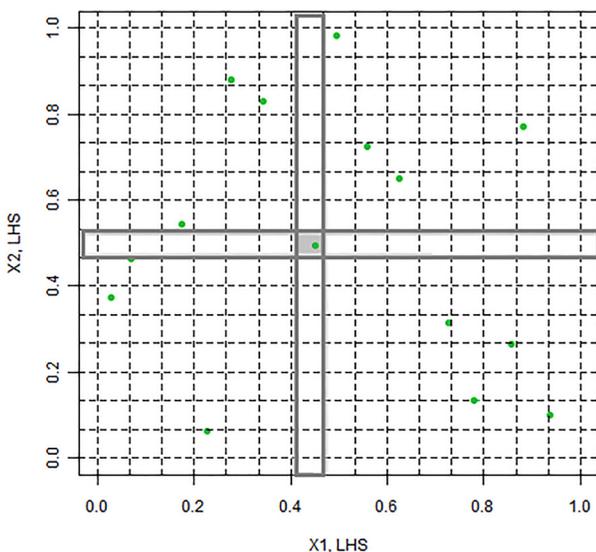


FIG. 3. A two-dimensional LHS sample (N = 15, s = 15).

LHS could be perceived as a compromise between PRS and stratified sampling, and the LHS procedure produces more stable results compared to those of PRS.⁴⁸ The superiority of LHS over PRS has been corroborated often enough, and LHS is generally recommended for use instead of PRS, as it is “an undemanding upgrade, but it substantially boosts convergence and applicability.”¹⁷

3. Sampling based on Sobol sequences (SBSS)

Another sampling used in QMC is sampling based on low-discrepancy sequences, such as Sobol sequences that use two as a base and uniformly partition given interval by a recursive algorithm. The recursive component in the algorithm is the principal difference between SBSS and other sampling strategies. Sobol sequences are examples of low-discrepancy sequences (sequences, with a low distance between sampled points), which were first introduced by Sobol in 1967.⁴⁹ The lower the discrepancy of a sample, the better it is for the purpose of conducting GSA.³ Low-discrepancy sequences have a very important property: as the sequence length N increases, the discrepancy decreases.

Figure 4 exhibits a two-dimensional Sobol sample. One can notice the low distance between sampled points and more order and homogeneity in Fig. 4 compared to LHS (Fig. 2) and to the pseudorandom sample in Fig. 1.

Bratley and Fox provide an implementation algorithm of generating Sobol sequences in Ref. 50. A procedure of defining the j th component of points in a Sobol sequence is summarized below.

A primitive polynomial of degree s_j of the form given in the following equation should be chosen:

$$x^{s_j} + a_{1,j}x^{s_j-1} + \dots + a_{s_j-1,j}x + 1, \quad (1)$$

where coefficients $a_{1,j}, \dots, a_{s_j-1,j}$ are binary.

The coefficients $a_{1,j}, \dots, a_{s_j-1,j}$ are used to define a sequence $\{m_{1,j}, m_{2,j}, \dots\}$ consisting of positive integers with the following recursive function:

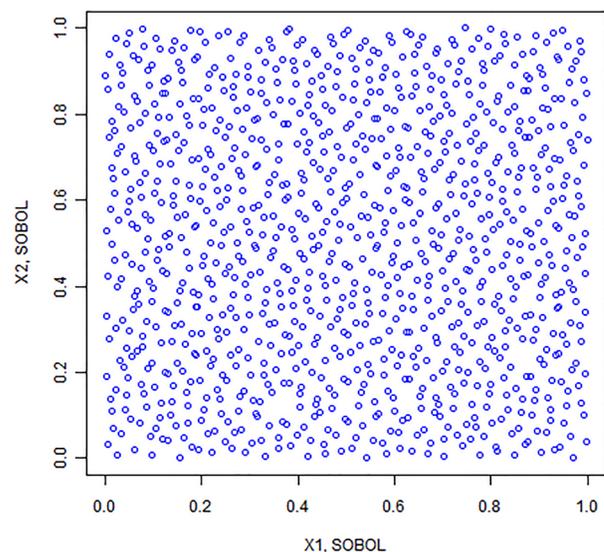


FIG. 4. Two-dimensional Sobol sample (N = 1000).

$$m_{k,j} = 2a_{1,j}m_{k-1,j}XOR2^2a_{2,j}m_{k-2,j}XOR\dots \\ \dots XOR 2^{s_j-1}a_{s_j-1}m_{k-s_j+1,j}XOR 2^{s_j}m_{k-s_j,j}XOR m_{k-s_j,j}, \quad (2)$$

for $k \geq s_j + 1$. XOR is a bit-by-bit exclusive OR operator. Initial values of $m_{k,j}$, $k = 1 \dots s_j$ should satisfy the condition of being odd and less than 2^k .

The so-called direction numbers $v_{k,j}$, $k = 1 \dots s_j$ are defined as

$$v_{k,j} = \frac{m_{k,j}}{2^k}. \quad (3)$$

Finally, the j th component of the i th point in a Sobol sequence is

$$x_{i,j} = b_1v_{1,j}XORb_2v_{2,j}XOR\dots \quad (4)$$

This algorithm is implemented in randtoolbox package in R.⁵¹

4. PRS, LHS, and SBSS comparison

This subsection touches on the quality of space exploration in a MC/QMC experiment and how the properties of samples are responsible for it. To this end, two-dimensional samples produced by PRS (Fig. 1), LHS (Fig. 2), and SBSS (Fig. 4) are compared. A pseudorandom sample (Fig. 1) has many holes and clusters and is the opposite of homogeneity, a property important for thorough input space

exploration necessary for GSA. LHS produces a more homogeneous sample (Fig. 2) compared to PRS; yet, it still has some holes and clusters, although almost not visible with the naked eye. But the homogeneity of a Sobol sample (Fig. 4) is evident compared to both pseudorandom (Fig. 1) and LHS samples (Fig. 2).

Moving on to multidimensional space exploration in a MC/QMC experiment using PRS, LHS, and SBSS, Fig. 5 exhibits projections of the three-dimensional uniformly distributed pseudorandom, LHS, and Sobol samples onto $X1X2$, $X2X3$, and $X1X3$ planes. In Fig. 5, one can see the same shift from nonhomogeneity in a pseudorandom sample to homogeneity in a Sobol sample observed when comparing Figs. 1, 2, and 4, but now in every two-dimensional projection (Fig. 5). Now, imagine that we move to an N -dimensional hypercube of a MC/QMC experiment with N input variables. Obviously, the nonhomogeneity of PRS becomes more of a problem as the sample size grows.

The discrepancy decrease is the reason SBSS produces an increased quality of the multivariate sample in terms of homogeneity and therefore better input space exploration. Figure 5 aims to exhibit precisely this property in the case of a three-dimensional problem. Figure 5 emphasizes the fact that requirements to the quality of a multivariate sample increase as the number of inputs increases, and the famous curse of dimensionality manifests itself (As the dimensionality grows, the volume of the hypercube explodes, making a sample sparse).

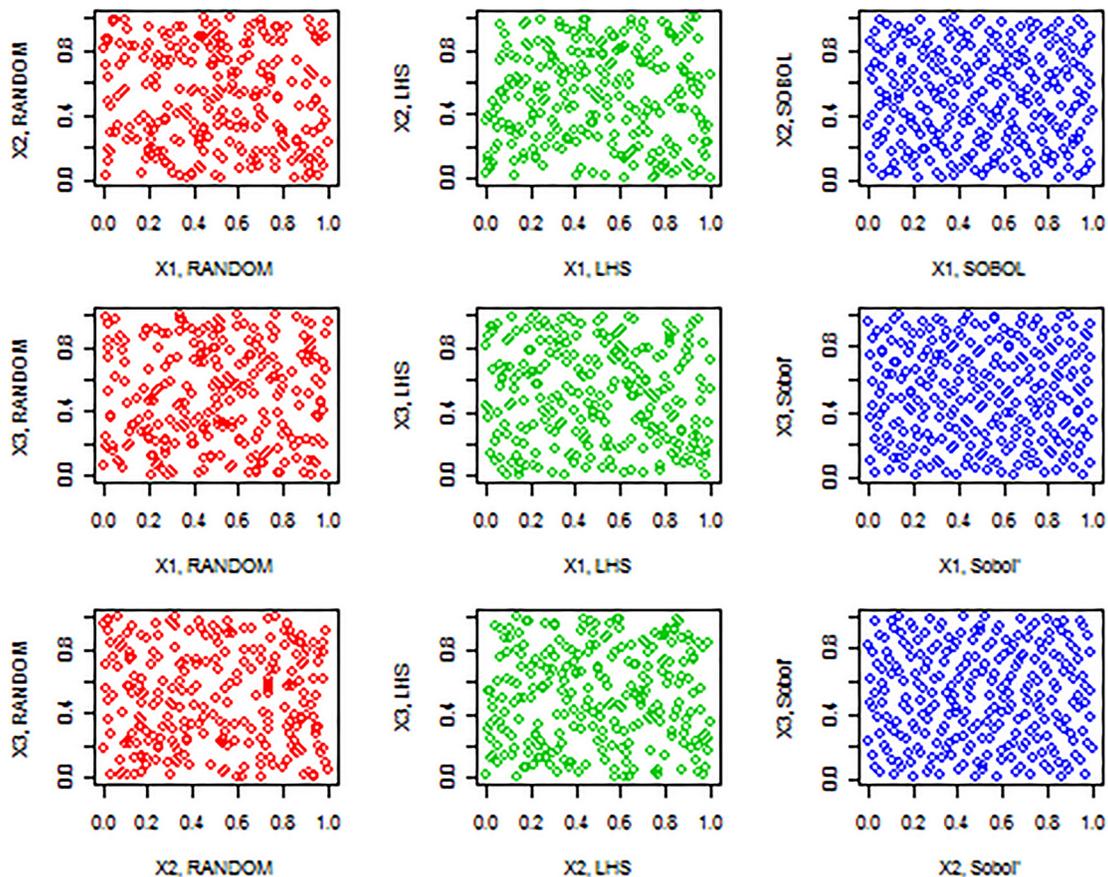


FIG. 5. Projections of three-dimensional random (left), LHS (center), and Sobol (right) samples ($N = 1000$).

This sparsity stands in the way of any method that requires statistical significance. The amount of data necessary to back up a statistically sound result grows with the dimensionality. Tarantola *et al.* investigated this problem in Ref. 52, where they compared convergence properties of SBSS and LHS and recommended using SBSS for most practical cases.

B. Variance-based global sensitivity analysis (GSA)

Variance-based methods constitute a class of GSA based on decomposing the variance of model output. The most interesting features of variance-based GSA methods are that they are independent of models (account for nonlinearity), provide a tool for investigating the influence of the full range of variation of the inputs, and provide the capacity to work with groups of factors.³

The main drawback of GSA as mentioned in the introduction is the impact of assumptions about underlying distributions of the input variables.⁷ Another drawback of variance-based methods is the same as of any MC simulation as they use MC/QMC, i.e., their computational cost in the case of complex models and numerous inputs. Although this is not a concern for the presented UAE case study, it could be if distribution parameters were evaluated for every model run. Although numeric wind models are outside the scope of this paper, if a CFD or a WRF was used to model the wind, SBSS would be useful and would serve as a significant timesaver for the QMC experiment.

Among GSA methods, there are also moment-independent methods⁵³ and density or distribution based methods.^{54,75} For a review of GSA in earth system modeling, refer to the studies by Wagener and Pianosi,⁷ for a review of global and local SA methods—to the study by Borgonovo and Plischke.⁵⁵ Also, Plischke *et al.* published a detailed review of global sensitivity measures.⁵⁶ Guidelines for the application of GSA can be found in Ref. 4. For a comparison of variance-based methods to moment-methods, refer to the study by Khorashadi Zadeh *et al.*⁵³

1. Sobol method: Brute force

Among GSA methods, “the most well-established and widely applied one is probably the variance-based method of Sobol.”⁵³ This method is sometimes referred to as Sobol SI method, Sobol SA or simply Sobol method. The Sobol method is based on calculating global sensitivity indices (SIs)—first-order and total effect—and was first mentioned by Sobol and Kucherenko in 1990.⁵⁷ SI is calculated for each input factor and serve as a measure of importance of this factor in relation to the influence on model output. The term “importance” is used to express the relative contribution of the variable to the uncertainty in the output. Sobol published a review of global SI in Ref. 58.

If a model is given in the form of $Y = f(X) = f(X_1, X_2, \dots, X_k)$, the variance based on the first-order effect for a factor X_i can be written as $V_i = V_{X_i}(E_{X_{\sim i}}(Y|X_i))$, where $X_{\sim i}$ denotes the matrix of all factors but X_i .² The meaning of the inner expectation operator is that the mean of Y is taken over all possible values of $X_{\sim i}$, while keeping the X_i factor fixed. The outer variance is taken over all possible values of X_i . The associated sensitivity measure, first-order sensitivity index, is equal to

$$S_i(Y) = \frac{V_i}{V(Y)} = \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i))}{V(Y)}, \quad (5)$$

where $V(Y)$ is defined as the variance of model output.

The first-order SI defined in Eq. (5) is the main measure used for ranking input factors according to their influence on the model output.

The total effect sensitivity index S_{T_i} of a factor i measures the total effect of this factor on model output, which includes first and higher order effects (or interactions) of factor X_i . One way to visualize the meaning of a total effect sensitivity index is by considering that $V_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i}))$ is the first-order effect of $X_{\sim i}$ so that $V(Y)$ minus $V_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i}))$ must give the contribution of all terms in the variance decomposition which include X_i . This measure gives extra information about the input factors, i.e., how active a factor is in interacting with other factors (dependence of input factors).

Ideally, input factors should be independent of one another, which is hardly the case in real world applications. In the case of independent input factors, the first-order SI is sufficient. When some factors are not independent, the calculation of total effect indices is necessary along with the calculation of first-order SI. The definition of a total effect sensitivity index² is

$$S_{T_i}(Y) = \frac{V_{T_i}}{V(Y)} = \frac{E_{X_{\sim i}}(V_{X_i}(Y|X_i))}{V(Y)} = 1 - \frac{V_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i}))}{V(Y)}. \quad (6)$$

Both first-order and total effect SIs can be calculated using a brute force method, i.e., according to their definitions in Eqs. (5) and (6). Brute force is only applicable in the case of a simple model and a small number of inputs, such as the case study on UAE presented in this paper. As the number of input variables increases or the model becomes too complex, the SI brute force calculation is no longer computationally affordable. In such cases, estimation of SI can be used. For example, if CFD or WRF models were to be used to model wind, estimation of SI can be used instead.

2. Sobol method: Saltelli and Jansen estimators

SI estimation provides ways to estimate SI instead of computing them using brute force which could be computationally not feasible when the curse of dimensionality is an issue for a given experimental design. There is ongoing research on the ways of estimating global Sobol SI.^{59–61} A number of estimation methods are ready for use in Sensitivity package in R.⁵¹ The estimation method of global SI, the authors used for the test model case study, is referred to simultaneous computation of estimates of first-order SI and total effect indices with the use of Sobol sampling, radial design, and Jansen estimator described in Ref. 1.

Model variance $V(Y)$ can be calculated as follows:³

$$V(Y) = \frac{1}{N} \sum_{j=1}^N (f(A_j)^2) - \left(\frac{1}{N} \sum_{j=1}^N (f(A_j)) \right)^2. \quad (7)$$

Saltelli *et al.* compared different schemes for estimating SI in Ref. 2 and found that the first-order SI S_i should be computed according to Eq. (8) and total effect SI S_{T_i} —according to Eq. (9). Equation (8) is also called the Saltelli estimator,² and Eq. (9) is referred to as the Jansen estimator,¹

$$S_i(Y) = \frac{V_{X_i}(E_{X \sim i}(Y|X_i))}{V(Y)} = \frac{\frac{1}{N} \sum_{j=1}^N f(B)_j \left(f(A_B^{(i)})_j - f(A)_j \right)}{\frac{1}{N} \sum_{j=1}^N \left(f(A)_j^2 \right) - \left(\frac{1}{N} \sum_{j=1}^N \left(f(A)_j \right) \right)^2}, \quad (8)$$

$$S_{T_i}(Y) = \frac{E_{X \sim i}(V_{X_i}(Y|X_i))}{V(Y)} = \frac{\frac{1}{2N} \sum_{j=1}^N \left(f(A_B^{(i)})_j - f(A)_j \right)^2}{\frac{1}{N} \sum_{j=1}^N \left(f(A)_j^2 \right) - \left(\frac{1}{N} \sum_{j=1}^N \left(f(A)_j \right) \right)^2}. \quad (9)$$

The matrix T of size $(2k, 2N)$ is generated with the use of SBSS. Then, matrix A would be assigned the value of the upper left matrix of size (k, N) and B —the lower right matrix of the same size as A . This is what Saltelli *et al.* call the radial design,²

$$T = \begin{pmatrix} A & \vdots \\ \dots & B \end{pmatrix}, \quad (10)$$

$$A = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ a_{21} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{Nk} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1k} \\ b_{21} & \dots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{N1} & \dots & b_{Nk} \end{pmatrix}. \quad (11)$$

The matrices A, B are used to form the $A_B^{(i)}$ matrices. The matrices $A_B^{(i)}$ are generated according to Eq. (12); all the columns come from matrix A , except for the i th column which comes from matrix B ,²

$$A_B^{(i)} = \begin{pmatrix} a_{11} & \dots & b_{1i} & \dots & a_{1k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{N1} & \dots & b_{Ni} & \dots & a_{Nk} \end{pmatrix}. \quad (12)$$

The implementation of this estimation procedure is summarized below:

1. Sample the matrices A, B with SBSS according to Eq. (10);
2. Generate $A_B^{(i)}$ matrices according to Eq. (12);
3. Calculate model outputs $Y(A), Y(B)$, and $Y(A_B^{(i)})$;
4. Compute first-order and total effect SI according to Eqs. (8) and (9), respectively.

Sensitivity package in R offers an implementation of this procedure called soboljansen.⁵¹

C. Experimental design

The numerical experiment consisted of two parts:

1. Testing sampling strategies on a test model for which analytical values of SI are available and choosing the best performance variance-based GSA method and sampling strategy according to test model results;
2. Performing GSA using the sampling strategy identified as best performing for the WRA model, i.e., annual energy production “AEP.”

The numerical experiment with the test model is aimed at defining the best GSA method and sampling strategy for the WRA model. To obtain robust results, a high number of replicas N_r of calculating a set of first-order SI S_i are used (number of replicas of experiment $N_r = 1000$). The ranking accuracy is calculated as

$$Accuracy = \frac{N_{identical}}{N_r} \times 100\%, \quad (13)$$

where $N_{identical}$ is the number of replicas when the ranking obtained in a numerical experiment matches the analytical ranking and N_r is the overall number of replicas of MC/QMC experiment.

Saltelli *et al.* used the following definitions of mean absolute error (MAE) for numerical experiments of convergence of first-order and total effect SI for a set of test functions in²

$$MAE(\widehat{S}_i) = \frac{1}{N_r} \sum_{j=1}^{N_r} \sum_{i=1}^k \widehat{S}_i(j) - S_i, \quad (14)$$

$$MAE(\widehat{S}_{T_i}) = \frac{1}{N_r} \sum_{j=1}^{N_r} \sum_{i=1}^k \widehat{S}_{T_i}(j) - S_{T_i}, \quad (15)$$

where $\widehat{S}_i(j)$ and $\widehat{S}_{T_i}(j)$ are estimates of first-order and total effect SI and S_i and S_{T_i} are the analytical values.

In order to compare the results of the test model with the results of Saltelli *et al.* reported in Ref. 2, $MAE(\widehat{S}_i)$ and $MAE(\widehat{S}_{T_i})$ are calculated. All numerical experiments were setup and run using the R Language and Environment for Statistical Computing.⁵¹

1. Test model

A specific G-function described by Saltelli *et al.* in Ref. 2 was chosen to be the test model as the one being closest (multiplicative) to the WRA model of AEP given in Eq. (25),

$$G = G(X_1, X_2, \dots, X_k, a_1, a_2, \dots, a_k) = \prod_{i=1}^k g_i, \quad (16)$$

$$g_i = \frac{4X_i - 2 + a_i}{1 + a_i}, \quad (17)$$

where $a_i = i$, $k = 4$, and all input variables are uniformly distributed in the interval $[0,1]$.

The variance based on first-order effect V_i , the variance based on total effect V_{T_i} , and the total variance V can be calculated according to Eqs. (18), (19), and (20), respectively,¹²

$$V_i = V_{X_i}(E_{X \sim i}(Y|X_i)) = \frac{1}{3(1+a_i)^2} = \frac{1}{3(1+i)^2}, \quad (18)$$

$$V_{T_i} = V_i \prod_{j \neq i} (1 + V_j) = \frac{1}{3(1+i)^2} \prod_{j \neq i} \left(1 + \frac{1}{3(1+j)^2} \right), \quad (19)$$

$$\begin{aligned} V &= \prod_{i=1}^k (1 + V_i) - 1 = \prod_{i=1}^4 (1 + V_i) - 1 \\ &= \prod_{i=1}^4 \left(1 + \frac{1}{3(1+i)^2} \right) - 1. \end{aligned} \quad (20)$$

The analytical form of the first-order SI for the chosen multiplicative test model is

$$S_i = \frac{V_i}{V} = \frac{1}{3(1+i)^2 \left[\prod_{i=1}^4 \left(1 + \frac{1}{3(1+i)^2} \right) - 1 \right]} \quad (21)$$

The analytical form of total effect SI for the chosen test model can be calculated as

$$S_{T_i} = \frac{V_{T_i}}{V} = \frac{\prod_{i \neq j} \left(1 + \frac{1}{3(1+j)^2} \right)}{3(1+i)^2 \left[\prod_{i=1}^4 \left(1 + \frac{1}{3(1+i)^2} \right) - 1 \right]} \quad (22)$$

All QMC experiments for the test function were conducted with the sample size of $N = 1024$.

2. Wind resource assessment (WRA) model in the UAE case study

Wind energy is proportional to the cube of wind speed, intuitively making wind energy most sensitive to the variance in the wind or wind model parameters. Uncertainty is inherent but not limited to the wind. There are also turbine related uncertainties, such as uncertainties associated with the turbine lifetime, necessary turbine maintenance, and losses. The distance to the grid, interest rate, inflation rate, investment cost, and operation and maintenance (O&M) costs also add a great deal of economic uncertainty to wind economics that rely on WRA. What is of interest here is whether it is possible to quantify the uncertainty originating from various sources of uncertainty. To what extent are wind model parameters responsible for the variance in wind energy? These are the research questions this case study is aiming to answer, and luckily variance-based GSA² is tailored to address such research questions.

When designing this case study, the authors looked into how other researchers designed their MC experiments in WRA. For instance, Kwon used variables such as Weibull shape and scale parameters, air density, mean wind speed, surface roughness exponent, and measure-correlate-predict (MCP) residual as inputs to MC simulation for AEP UA.⁶² Montes *et al.* included economic uncertainties to assess wind farm profitability using a MC approach and obtained a probability distribution function (PDF) for the net present value (NPV) of costs and internal rate of return (IRR).²⁵ Amirinia *et al.* conducted a MC based offshore WRA study of the Arabian Gulf.²⁶ They presented UA results as PDF of AEP and used two-scenario economic analysis. Caralis *et al.* used the MC method for assessing the profitability of twelve offshore wind projects. They integrated MC simulation into a classical financial model, thus accounting for various sources of uncertainty related to expected changes in feed-in tariff (FiT) legislation in Greece.²⁷ The results demonstrate how difficult it is to propose a common FiT for all offshore wind farms (OWFs), despite Greece's small size.

What output variables were of interest in SA studies performed in WRA? Several studies on SA of LCOE in WRA have been published in the last few years.^{10–13,29,37} Ioannou *et al.* performed a stochastic financial appraisal of offshore wind farms (OWFs) based on the deterministic lifecycle techno-economic model and reported based on a local SA that the cost of the turbine, the mean time to failure (hence

O&M cost), and foundation cost are most influential on NPV and LCOE.¹⁰ They provide a thorough overview of costs related to OWF and assumptions for modeling them. Castro-Santos and Diaz-Casas used the Oracle Crystal Ball to setup a MC simulation to get the SA results of floating OWF (FOWF).¹¹ Their results indicate that the most important variables are the wind scale parameter and the electric tariff for the LCOE, discounted pay-back period, IRR, NPV, and cost of power. Lerch *et al.* proposed a methodology for calculating LCOE for a specific FOWF and performed an extensive yet local SA with 325 input variables.¹² They found that the turbine, substructure, and mooring system manufacturing cost as well as power cable cost are the most influential variables besides the common discount rate and energy losses. Galvez *et al.* recently reported results of a local SA in WRA for Tabasco, Mexico.¹³ The authors looked at how the change in the discount rate, O&M costs, and investment costs affects LCOE and with the use of SA revealed the need to implement financial mechanisms like FiT to decrease the cost of electricity generation in places with relatively low wind potential.

Although most SA in WRA studies used LCOE^{10–13,29,37} or goodness-of-fit (GOF) measures^{14,15} as the output variable, the authors decided to define AEP as model output because other variables of interest like LCOE, IRR, and NPV all rely on AEP. Conducting GSA with a GOF measure as the output variable is planned to be reported in a following paper.

Masdar City (24.420° N, 54.613° E) is located in the coastal area inside the city of Abu Dhabi, UAE (Fig. 6). The authors chose it for the case study because Naizghi *et al.* have pointed out a high wind power generation potential in the coastal and mountainous regions of the UAE⁶³ and there were temperature and atmospheric pressure measurements available for the location as a nonwind-turbine-specific study was intended. The meteorological station at Masdar City consists of a meteorological tower with 5 calibrated Hersteller Thies Clima cup anemometers,⁶⁴ two wind vanes, one Campbell CR1000 data logger, and three sensors of temperature, barometric pressure, and humidity.⁶⁵ Ten-minute interval data of wind speed, temperature, atmospheric pressure, and humidity were collected at the height of 50 m from August 2008 to July 2007 and from January 2010 to January 2011. These data were used for the QMC experiment for GSA. Extrapolation to the hub height of a hypothetical wind turbine was not necessary as the measurements were taken at the height of 50 m. The length of available data is limited to two years only. A MCP method can be used to extend it, but this step is omitted, as AEP was chosen as the variable of interest. Interannual variability could be taken into account in a later refined GSA study of Masdar City if found fit and necessary.

The importance of accounting for the variance in air density in WRA was pointed out by Jung and Schindler in Ref. 66. Air density was calculated according to Eq. (23) with the collected data of temperature and atmospheric pressure at Masdar City. The Empirical cumulative distribution function (ECDF) of air density (Fig. 7) based on the calculated values was used for distributional modeling of air density. For alternative ways to account for the variability in air density in WRA, refer to Ref. 66,

$$\rho = \frac{p}{R_{spec} T}, \quad (23)$$

where $R_{spec} = 287.058$ J/kg K is the specific gas constant of dry air, t is the air temperature in K, and p is the atmospheric pressure.



FIG. 6. Masdar City location.

The two-parameter Weibull (W2) distribution is widely accepted in the wind energy field.^{38,39} It is also most frequently used and often assumed best-fit distribution for wind data. A lot of researchers used W2 in recent WRA studies.^{21,67–71} The PDF of the W2 distribution is

$$f(w) = \frac{c}{\lambda} \left(\frac{w}{\lambda}\right)^{c-1} e^{-\left(\frac{w}{\lambda}\right)^c}, \quad (24)$$

where w is the sample of wind speed data, c is the shape parameter, and λ is the scale parameter.

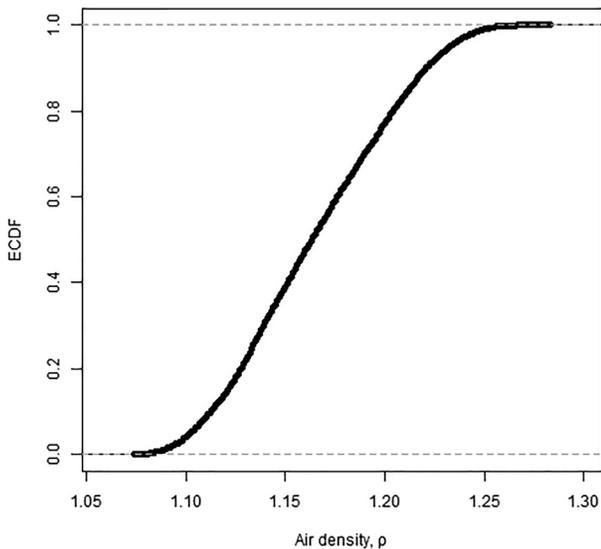


FIG. 7. ECDF of air density ρ at Masdar City.

Ouarda *et al.* showed that the W2 exhibits best GOF for modeling wind speed in the UAE among two-parameter distributions and performs better than most three-parameter distributions, such as the generalized extreme value and the three-parameter lognormal distributions.⁷² Hence, the W2 was chosen to model the wind at Masdar City, UAE.

When W2 is used to model wind and electric losses EL (in percent) are taken into account, the AEP can be calculated according to Eq. (25).³⁸ Generally, electrical losses include cable losses, transformation losses, and losses in the substation,

$$AEP = \frac{(100 - EL)a}{100} \frac{1}{2} \rho \lambda^3 \Gamma\left(\frac{c+3}{c}\right) 8760, \quad (25)$$

where a is the availability of a wind turbine to generate electricity in percent of the time during the year available for electricity production (not down for maintenance),³⁸ ρ is the air density, Γ is the gamma function, c is the shape parameter, and λ is the scale parameter of the W2.

According to Manwell *et al.*, the availability of a wind turbine lies within the 97%–99% interval.³⁸ Madariaga *et al.* report that the electrical losses in wind energy conversion systems are in the range of 3.25%–4.75%,⁷³ and therefore, the expected value of electrical losses was estimated to be 4%. When nothing else apart from a range of variables is known, we make the most general assumption that turbine availability is uniformly distributed in [97, 99] and electrical losses are uniformly distributed in [3.25, 4.75].

Weibull shape and scale parameters were estimated using maximum likelihood estimation (MLE). The results of MLE for wind speed data at Masdar City are given in Table I. For the QMC experiment, we assume that both the Weibull shape and the scale parameter are normally distributed with the mean corresponding to the respective ML estimate and standard deviation—to standard error of the respective

TABLE I. MLE estimates of Weibull distribution parameters for Masdar City wind speed data.

	MLE estimate	Standard error of MLE estimate
Shape parameter of the Weibull distribution, c	2.1073	0.0049
Scale parameter of the Weibull distribution, λ	4.9368	0.0075

MLE estimate. Similar assumptions about Weibull parameters were made by Amirinia *et al.* in a MC UA study of the Arabian Gulf.²⁶ The value of Kolmogorov-Smirnov statistics for the W2 fit to Masdar City wind speed data at 50 m is equal to 0.015.

The input and output variables in QMC experiment with the respective assumptions about the underlying distributions of the variables are summarized in Table II. The authors defined the output variable of interest for WRA in this study as the commonly used AEP in megawatt hour [Eq. (25)]. All QMC experiments for the AEP were conducted using the brute force calculation of Sobol SI with the sample size of $N = 1024$.

III. RESULTS

This section presents the results of the two case studies conducted.

A. Test model results

The results of calculating the accuracy of ranking according to the first-order SI using the brute force method for the test model are given in Fig. 8. Random, LHS, and Sobol samples converge to the analytical ranking with 100% ranking accuracy as the sample size becomes greater than 64. So if the model at hand allows a sample size of 64, then all sampling strategies are equivalent. If a sample size smaller than 64 has to be used, the maximum accuracy can be achieved with SBSS or LHS. One can see that LHS and SBSS always outperform PRS. For small sample sizes of 2 and 4, LHS sampling outperforms SBSS.

TABLE II. Variables used in the WRA model in the UAE case study.

Type	Variable	Definition
Input	Weibull scale, λ	Gaussian (4.9368, 0.0049) according to Ref. 26
Input	Weibull shape, c	Gaussian (2.1072, 0.0075) according to Ref. 26
Input	Availability a , % per year	Uniform [97, 99] according to Ref. 38
Input	Electrical losses EL , %	Uniform [3.25, 4.75] according to Ref. 73
Input	Air density ρ , kg/m ³	The ECDF of air density (Fig. 7)
Output/model	Annual energy production, AEP	$\frac{(100-EL)a}{100} \frac{1}{2} \rho \lambda^3 \Gamma\left(\frac{c+3}{c}\right) 8760$

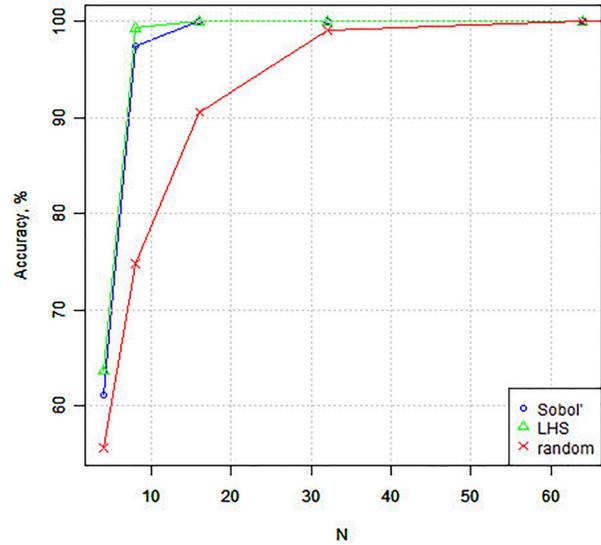


FIG. 8. Accuracy of S_i for the test model using the brute force method with a sample size of $N = 4 \dots 64$ ($N_r = 1000$).

As the sample size increases without reaching the value of 64, LHS performs slightly better than SBSS.

The accuracy of ranking results with an estimation of the first-order SI for the test model is shown in Fig. 9. The accuracy of the estimated indices is significantly lower compared to the brute force method (4.5% as opposed to 100% for $N = 1024$). Therefore, the brute force should be used for the WRA model since the model is not complicated and computational time allows it. One can see that there is a drop in the accuracy with the growing sample size of all sampling strategies. This was not observed in the case of the brute force method.

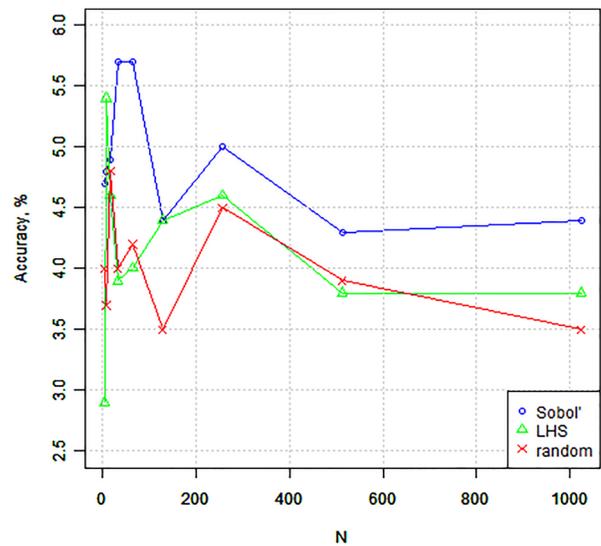


FIG. 9. Accuracy of estimating S_i for the test model with a sample size of $N = 4 \dots 1024$ ($N_r = 1000$).

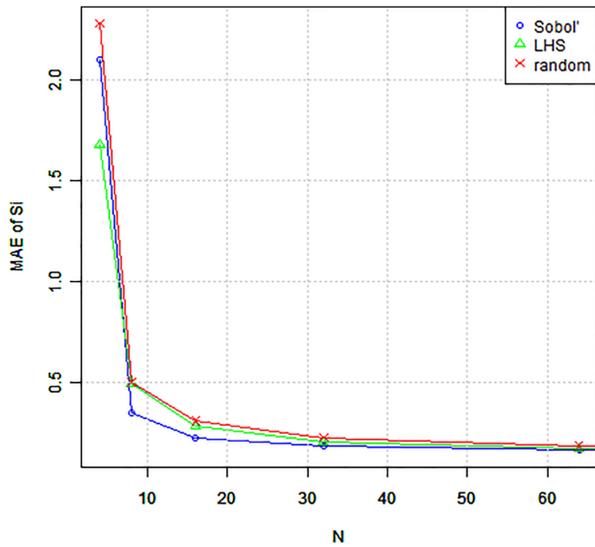


FIG. 10. Mean absolute error of S_i , $MAE(\widehat{S}_i)$, for the brute force method ($N=4\dots64$, $N_r=1000$).

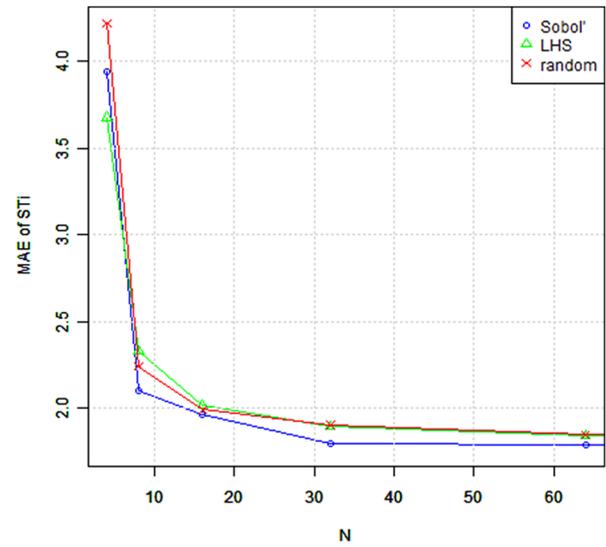


FIG. 12. Mean absolute error of S_{T_i} , $MAE(\widehat{S}_{T_i})$, for the brute force method ($N=4\dots64$, $N_r=1000$).

For LHS, the drop occurs the earliest at a sample size of $N=8$, PRS exhibits the drop at $N=16$, and SBSS shows the drop at $N=64$.

Mean absolute errors (MAEs) were calculated for both sets of indices (first-order and total effect) for the two methods being compared: brute force and best-practice estimation. Figures 10 and 11 show $MAE(\widehat{S}_i)$ as a function of sample size ($N=4\dots64$) for brute force and estimation methods, respectively. Figures 12 and 13 exhibit $MAE(\widehat{S}_{T_i})$ as a function of sample size ($N=4\dots64$) for brute force and estimation methods, respectively. From these graphs, one can see that

- MAE for SBSS is always smaller compared to other sampling strategies for the sample sizes of $N=8, \dots, 1024$;
- LHS outperforms SBSS only with a sample size of $N=8$ which is not commonly used for QMC simulation;
- $MAE(\widehat{S}_i)$ for the estimation method converges to a higher value than in the case of the brute force method;
- $MAE(\widehat{S}_{T_i})$ for the estimation method converges to a higher value than in the case of the brute force method;
- When considering the estimation method, $MAE(\widehat{S}_i)$ is one order smaller than $MAE(\widehat{S}_{T_i})$;

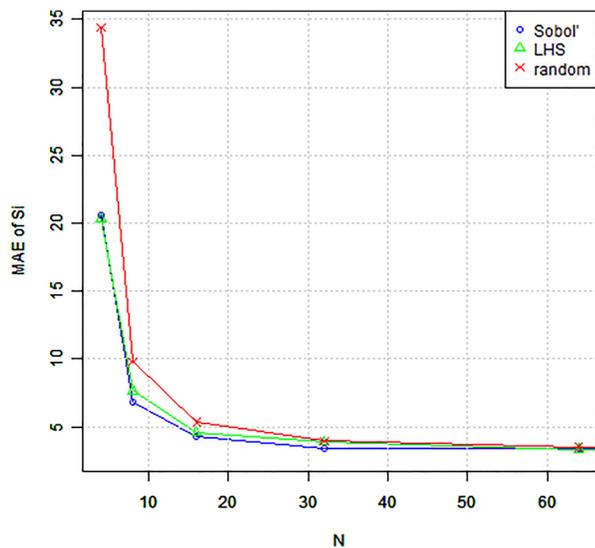


FIG. 11. Mean absolute error of S_i , $MAE(\widehat{S}_i)$, for the estimation method ($N=4\dots64$, $N_r=1000$).

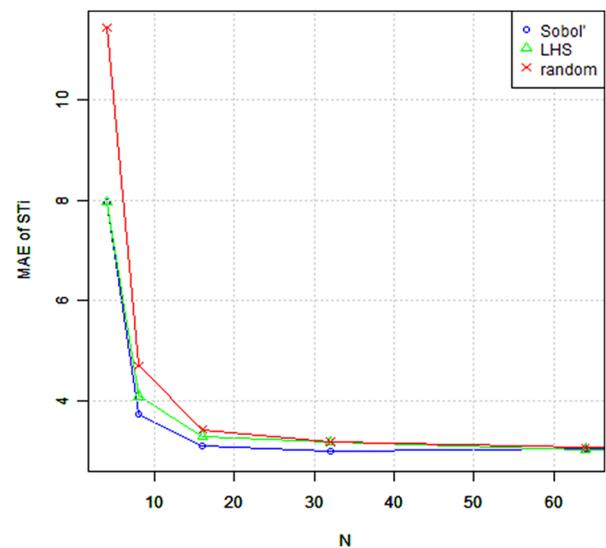


FIG. 13. Mean absolute error of S_{T_i} , $MAE(\widehat{S}_{T_i})$, for the estimation method ($N=4\dots64$, $N_r=1000$).

- When considering the brute force method, $MAE(\hat{S}_i)$ is of the same magnitude as $MAE(\hat{S}_{T_i})$.

Test model results are summarized below:

- (1) Sampling strategies have been tested on the G-function ($a_i = i, k = 4$). LHS and SBSS outperform PRS in all cases tested. In some limited cases, LHS slightly outperforms SBSS, but the difference is not significant. In the majority of cases tested, SBSS exhibits better performance and is therefore recommended for practical use. This finding is similar to the recommendation of Tarantola *et al.* in Ref. 52.
- (2) MAE for SBSS is always smaller compared to other sampling strategies for sample size $N = 8 \dots 1024$. LHS outperforms SBSS only with a sample size of $N = 8$ which is too small for QMC simulation;
- (3) $MAE(\hat{S}_i)$ for the estimation method is always higher than $MAE(\hat{S}_i)$ for the brute force method. $MAE(\hat{S}_{T_i})$ for the estimation method is also always higher than $MAE(\hat{S}_{T_i})$ for the brute force method;
- (4) When considering the estimation method, $MAE(\hat{S}_i)$ is one order smaller than $MAE(\hat{S}_{T_i})$. When considering the brute force method, $MAE(\hat{S}_i)$ is of the same magnitude as $MAE(\hat{S}_{T_i})$;
- (5) GSA brute force and estimation methods have been compared using the G-function ($a_i = i, k = 4$). The brute force method shows 100% accuracy for ranking order at a sample size of $N = 64$ using any sampling strategy. Brute force is recommended for WRA, when wind is modeled with a distribution; else, estimation should be used.

B. The UAE case study results

The GSA results of the AEP at Masdar City are given as a bar chart of total effect SI S_{T_i} in Fig. 14. 94% of variance in the AEP is explained by the variance in air density and 2.7% by the variance in turbine availability, and the Weibull scale parameter and electrical losses contribute the same amount of variance to the AEP, which is equal to 1.6%. The least influential factor to the AEP is the Weibull shape parameter responsible for 0.4% of the AEP variance. Despite the

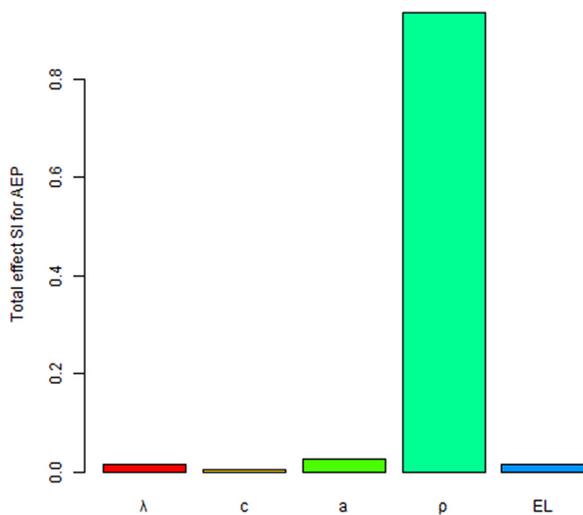


FIG. 14. GSA result or importance of inputs for AEP according to S_{T_i} .

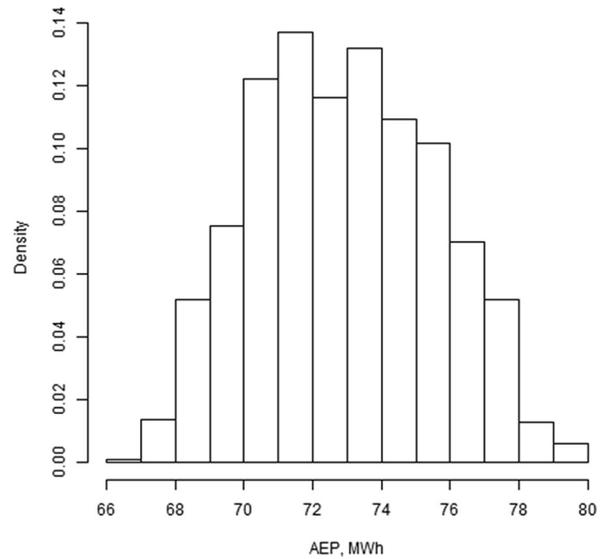


FIG. 15. UA results or histogram of AEP at Masdar City.

AEP calculated in cubical proportion to the Weibull scale parameter, the Weibull scale parameter surprisingly contributes only 1.6% of the variance in the AEP. Thus, research questions posed in Sec. II C 2 have been answered. Note that the variance from all inputs considered adds up to 100%, as the AEP is calculated based on a simple multiplicative model, and proves that the QMC experiment converged. The ranking obtained and the magnitude of difference in variance contribution are surprising. The contribution to the overall variance in AEP by the variance in air density is underestimated and should be further investigated.

The UA results of the AEP at Masdar City are presented in Fig. 15. The PDF of the AEP was obtained using the same QMC experiment ($N = 1024$) setup to obtain the GSA results. Presenting a PDF is always superior to presenting point estimates. From Fig. 15, one can see that the AEP at Masdar City lies in the range of 67–81 MWh. The mean value of the AEP is 72 MWh.

IV. CONCLUSIONS AND FUTURE RESEARCH

The present paper is focused on the Sobol method as a variance-based GSA method for WRA and the aspects of its application. The conclusions from both conducted case studies are:

1. SBSS outperforms LHS and PRS and is recommended for use in QMC in WRA;
2. The brute force SI calculation offers high accuracy of results and is recommended for use whenever the model and number of input variables allow it (similar to the UAE case study); otherwise, their estimation can be used;
3. The UAE case study results suggest that the influence of variance in Weibull parameters on the AEP might be overestimated, while the influence of the variance in air density on the AEP might be overlooked. These results are specific to the case study and should not be generalized without further effort.

The Masdar City case study illustrated that variance-based GSA can be easily, with the means of QMC, and at no extra cost (by using R^{49}) set

up in a WRA study. The wind community would benefit greatly from adapting the use of QMC and the Sobol method, due to the following reasons:

1. First, QMC enables calculating global Sobol SI that provides a valid measure of importance for WRA models nonlinear in their nature;
2. Second, by setting up a GSA study with QMC, a PDF of the output model or the UA results is obtained using the same QMC experiment;
3. Third, it allows for miscellaneous data and information about a site of interest to be incorporated in a GSA study in WRA.

Future research should focus on the influence of air density in a more general context, for instance, through running GSA similar to the one performed for the UAE case study but for a number of sites with longer datasets available. Future efforts should consider wind turbine specific studies with fitting air density to candidate distributions as recently suggested by Jung and Schindler,⁶⁶ modeling air density with the best-fit and calculating AEP with air density corrected power curves.⁶⁶ Including economic uncertainties to the problem setup and conducting GSA for LCOE as defined in Ref. 37 could be carried out with the use of mixture distribution models instead of Weibull for AEP estimation as suggested by Ouarda *et al.* in Ref. 72 and Shin *et al.* in Ref. 74. The idea of combining wind-related and economic uncertainties originated in Refs. 75 and 76. Another interesting direction of research is performing a GSA study of how the quality of wind speed data (missing value percentage, measurement accuracy, and temporal step) affects GOF and/or wind speed distribution parameter estimation accuracy, enhancing the work of Jung and Schindler.¹⁴

ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chair Program. The authors are very grateful to the Editor-in-Chief Carlos F. M. Coimbra and the Editorial Board of Journal of Renewable and Sustainable Energy for their time, valuable input, patience, and understanding. The authors are also very grateful to the two anonymous reviewers for thoughtful and detailed reviews that helped to improve the UAE case study and the overall quality of this manuscript significantly.

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

- ¹M. J. W. Jansen, "Analysis of variance designs for model output," *Comput. Phys. Commun.* **117**, 35–43 (1999).
- ²A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Comput. Phys. Commun.* **181**, 259–270 (2010).
- ³A. Saltelli, M. Ratto, T. Andres, J. Campolongo, F. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis: The Primer* (John Wiley & Sons, Ltd., 2008).
- ⁴A. Saltelli and P. Annoni, "How to avoid a perfunctory sensitivity analysis," *Environ. Model. Software* **25**, 1508–1517 (2010).
- ⁵F. Ferretti, A. Saltelli, and S. Tarantola, "Trends in sensitivity analysis practice in the last decade," *Sci. Total Environ.* **568**, 666–670 (2016).
- ⁶A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, S. Li, and Q. Wu, "Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices," *Environ. Model. Software* **114**, 29–39 (2019).
- ⁷T. Wagener and F. Pianosi, "What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling," *Earth-Sci. Rev.* **194**, 1–18 (2019).
- ⁸See, <https://ec.europa.eu/jrc/en/samo/methods> for Global sensitivity analysis, EU Science HUB; accessed 30 May 2019.
- ⁹F. Pianosi, F. Sarrazin, and T. Wagener, "A Matlab toolbox for global sensitivity analysis," *Environ. Model. Software* **70**, 80–85 (2015).
- ¹⁰A. Ioannou, A. Angus, and F. Brennan, "Stochastic financial appraisal of offshore wind farms," *Renewable Energy* **145**, 1176–1191 (2020).
- ¹¹M. Lerch, M. De-Prada-Gil, C. Molins, and G. Benveniste, "Sensitivity analysis on the levelized cost of energy for floating offshore wind farms," *Sustainable Energy Technol. Assess.* **30**, 77–90 (2018).
- ¹²L. Castro-Santos and V. Diaz-Casas, "Sensitivity analysis of floating offshore wind farms," *Energy Convers. Manage.* **101**, 271–277 (2015).
- ¹³G. Hernández Galvez, R. Saldaña Flores, U. Miranda Miranda, O. Sarracino Martínez, M. Castillo Téllez, D. Almenares López, and A. K. T. Gómez, "Wind resource assessment and sensitivity analysis of the levelised cost of energy. A case study in Tabasco, Mexico," *Renewable Energy Focus* **29**, 94–106 (2019).
- ¹⁴C. Jung and D. Schindler, "Sensitivity analysis of the system of wind speed distributions," *Energy Convers. Manage.* **177**, 376–384 (2018).
- ¹⁵O. Alavi, A. Sedaghat, and A. Mostafaeipour, "Sensitivity analysis of different wind speed distribution models with actual and truncated wind data: A case study for Kerman, Iran," *Energy Convers. Manage.* **120**, 51–61 (2016).
- ¹⁶I. Sobol', "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Math. Comput. Simul.* **55**, 271–280 (2001).
- ¹⁷H. Janssen, "Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence," *Reliab. Eng. Syst. Saf.* **109**, 123–132 (2013).
- ¹⁸X. Xu and Z. Yan, "Probabilistic load flow calculation with quasi-Monte Carlo and multiple linear regression," *Int. J. Electr. Power Energy Syst.* **88**, 1–12 (2017).
- ¹⁹S. Burhenne, D. Jacob, and G. P. Henze, "Sampling based on Sobol sequences for Monte Carlo techniques applied to building simulations," in *Proceedings of the Building Simulation (2011)*, pp. 1816–1823.
- ²⁰M. Shoaib, I. Siddiqui, S. Rehman, S. Khan, and L. M. Alhems, "Assessment of wind energy potential using wind energy conversion system," *J. Cleaner Prod.* **216**, 346–360 (2019).
- ²¹T. Aukitino, M. G. M. Khan, and M. R. Ahmed, "Wind energy resource assessment for Kiribati with a comparison of different methods of determining Weibull parameters," *Energy Convers. Manage.* **151**, 641–660 (2017).
- ²²M. Y. Kazet, R. Mouangue, A. Kuitche, and J. M. Ndjaka, "Wind energy resource assessment in Ngaoundere locality," *Energy Procedia* **93**, 74–81 (2016).
- ²³G. Marmidis, S. Lazarou, and E. Pyrgioti, "Optimal placement of wind turbines in a wind park using Monte Carlo simulation," *Renewable Energy* **33**, 1455–1460 (2008).
- ²⁴L. Wang, J. Yuan, M. E. Cholette, Y. Fu, Y. Zhou, and A. C. Tan, "Comparative study of discretization method and Monte Carlo method for wind farm layout optimization under Weibull distribution," *J. Wind Eng. Ind. Aerodyn.* **180**, 148–155 (2018).
- ²⁵G. M. Montes, E. P. Martin, J. A. Bayo, and J. O. Garcia, "The applicability of computer simulation using Monte Carlo techniques in windfarm profitability analysis," *Renewable Sustainable Energy Rev.* **15**, 4746–4755 (2011).
- ²⁶G. Amirinia, S. Mafi, and S. Mazaheri, "Offshore wind resource assessment of Persian Gulf using uncertainty analysis and GIS," *Renewable Energy* **113**, 915–929 (2017).
- ²⁷G. Caralis, P. Chaviaropoulos, V. R. Albacete, D. Diakoulaki, V. Kotroni, K. Lagouvardos, Z. Gao, A. Zervos, and K. Rados, "Lessons learnt from the evaluation of the feed-in tariff scheme for offshore wind farms in Greece using a Monte Carlo approach," *J. Wind Eng. Ind. Aerodyn.* **157**, 63–75 (2016).
- ²⁸C. bin Li, G. shu Lu, and S. Wu, "The investment risk analysis of wind power project in China," *Renewable Energy* **50**, 481–487 (2013).
- ²⁹V. Gass, F. Strauss, J. Schmidt, and E. Schmid, "Assessing the effect of wind power uncertainty on profitability," *Renewable Sustainable Energy Rev.* **15**, 2677–2683 (2011).
- ³⁰S. Li and L. Caracoglia, "Surrogate model Monte Carlo simulation for stochastic flutter analysis of wind turbine blades," *J. Wind Eng. Ind. Aerodyn.* **188**, 43–60 (2019).

- ³¹Y. F. Li, S. Valla, and E. Zio, "Reliability assessment of generic geared wind turbines by GTST-MLD model and Monte Carlo simulation," *Renewable Energy* **83**, 222–233 (2015).
- ³²R. Teixeira, A. O'Connor, and M. Nogal, "Probabilistic sensitivity analysis of offshore wind turbines using a transformed Kullback-Leibler divergence," *Struct. Saf.* **81**, 101860 (2019).
- ³³J. Velarde, C. Kramhøft, and J. D. Sørensen, "Global sensitivity analysis of offshore wind turbine foundation fatigue loads," *Renewable Energy* **140**, 177–189 (2019).
- ³⁴Z. Di, J. Ao, Q. Duan, J. Wang, W. Gong, C. Shen, Y. Gan, and Z. Liu, "Improving WRF model turbine-height wind-speed forecasting using a surrogate-based automatic optimization method," *Atmos. Res.* **226**, 1–16 (2019).
- ³⁵P. M. McKay, R. Carriveau, D. S.-K. Ting, and J. L. Johrendt, "Global sensitivity analysis of wind turbine power output," *Wind Energy* **17**, 983–995 (2014).
- ³⁶F. Ulker, D. Allaire, and K. Willcox, "Sensitivity-guided decision-making for wind farm micro-siting," *Int. J. Numer. Methods Fluids* **83**, 52–72 (2017).
- ³⁷T. T. D. Tran and A. D. Smith, "Incorporating performance-based global sensitivity and uncertainty analysis into LCOE calculations for emerging renewable energy technologies," *Appl. Energy* **216**, 157–171 (2018).
- ³⁸J. F. Manwell, J. G. McGowan, and A. L. Rogers, *Wind Energy Explained: Theory, Design and Application* (Wiley, 2009).
- ³⁹P. Wais, "A review of Weibull functions in wind sector," *Renewable Sustainable Energy Rev.* **70**, 1099–1107 (2017).
- ⁴⁰T. B. M. J. Ouarda, C. Charron, and F. Chebana, "Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study," *Energy Convers. Manage.* **124**, 247–265 (2016).
- ⁴¹D. Villanueva and A. Feijóo, "Wind power distributions: A review of their applications," *Renewable Sustainable Energy Rev.* **14**, 1490–1495 (2010).
- ⁴²X.-Y. Tang, S. Zhao, B. Fan, J. Peinke, and B. Stoevesandt, "Micro-scale wind resource assessment in complex terrain based on CFD coupled measurement from multiple masts," *Appl. Energy* **238**, 806–815 (2019).
- ⁴³C. L. Defforge, B. Carissimo, M. Bocquet, R. Bresson, and P. Armand, "Improving CFD atmospheric simulations at local scale for wind resource assessment using the iterative ensemble Kalman smoother," *J. Wind Eng. Ind. Aerodyn.* **189**, 243–257 (2019).
- ⁴⁴E. Arteaga-López, C. Ángeles-Camacho, and F. Bañuelos-Ruedas, "Advanced methodology for feasibility studies on building-mounted wind turbines installation in urban environment: Applying CFD analysis," *Energy* **167**, 181–188 (2019).
- ⁴⁵D. J. Willis, C. Niezrecki, D. Kuchma, E. Hines, S. R. Arwade, R. J. Barthelmie, M. DiPaola, P. J. Drane, C. J. Hansen, M. Inalpolat, J. H. Mack, A. T. Myers, and M. Rotea, "Wind energy research: State-of-the-art and future research directions," *Renewable Energy* **125**, 133–154 (2018).
- ⁴⁶E. B. Barker and J. M. Kelsey, "Recommendation for random number generation using deterministic random bit generators," Report No. 800-90A Rev 1, Gaithersburg, MD, 2012.
- ⁴⁷A. Nikishova, A. Kalyuzhnaya, A. Boukhanovsky, and A. Hoekstra, "Uncertainty quantification and sensitivity analysis applied to the wind wave model SWAN," *Environ. Model. Software* **95**, 344–357 (2017).
- ⁴⁸J. C. Helton and F. J. Davis, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," *Reliab. Eng. Syst. Saf.* **81**, 23–69 (2003).
- ⁴⁹I. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *USSR Comput. Math. Math. Phys.* **7**, 86–112 (1967).
- ⁵⁰P. Bratley and B. L. Fox, "ALGORITHM 659: implementing Sobol's quasirandom sequence generator," *ACM Trans. Math. Software* **14**, 88–100 (1988).
- ⁵¹R.C. Team, <https://www.r-project.org/> for R: A Language and Environment for Statistical Computing (2019).
- ⁵²S. Tarantola, W. Becker, and D. Zeitz, "A comparison of two sampling methods for global sensitivity analysis," *Comput. Phys. Commun.* **183**, 1061–1072 (2012).
- ⁵³F. Khorashadi Zadeh, J. Nossent, F. Sarrazin, F. Pianosi, A. van Griensven, T. Wagener, and W. Bauwens, "Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model," *Environ. Model. Software* **91**, 210–222 (2017).
- ⁵⁴F. Pianosi and T. Wagener, "Distribution-based sensitivity analysis from a generic input-output sample," *Environ. Model. Software* **108**, 197–207 (2018).
- ⁵⁵E. Borgonovo and E. Plischke, "Sensitivity analysis: A review of recent advances," *Eur. J. Oper. Res.* **248**, 869–887 (2016).
- ⁵⁶E. Plischke, E. Borgonovo, and C. L. Smith, "Global sensitivity measures from given data," *Eur. J. Oper. Res.* **226**, 536–550 (2013).
- ⁵⁷I. Sobol' and S. Kucherenko, "On global sensitivity analysis of quasi-Monte Carlo algorithms, Monte Carlo methods and applications," *Monte Carlo Methods Appl.* **11**, 1–9 (2005).
- ⁵⁸I. M. Sobol', "Global sensitivity indices for nonlinear mathematical models," *Rev. Math. Comput. Simul.* **55**, 271–280 (2001).
- ⁵⁹L. Pronzato, "Sensitivity analysis via Karhunen-Loève expansion of a random field model: Estimation of Sobol' indices and experimental design," *Reliab. Eng. Syst. Saf.* **187**, 93–109 (2019).
- ⁶⁰L. Gilquin, E. Arnaud, C. Prieur, and A. Janon, "Making the best use of permutations to compute sensitivity indices with replicated orthogonal arrays," *Reliab. Eng. Syst. Saf.* **187**, 28–39 (2019).
- ⁶¹S. Kucherenko, O. V. Klymenko, and N. Shah, "Sobol' indices for problems defined in non-rectangular domains," *Reliab. Eng. Syst. Saf.* **167**, 218–231 (2017).
- ⁶²S. D. Kwon, "Uncertainty analysis of wind energy potential assessment," *Appl. Energy* **87**, 856–865 (2010).
- ⁶³M. S. Naizghi and T. B. M. J. Ouarda, "Teleconnections and analysis of long-term wind speed variability in the UAE," *Int. J. Climatol.* **37**, 230–248 (2017).
- ⁶⁴Measnet, http://www.measnet.com/wp-content/uploads/2011/06/measnet_anemometer_calibration_v2_oct_2009.pdf for ANEMOMETER CALIBRATION Version 2 October 2009, Measurement (2009).
- ⁶⁵I. Janajreh and I. Talab, "Wind data collection and analyses at Masdar city for wind turbine assessment," *Int. J. Therm. Environ. Eng.* **1**, 43–50 (2010).
- ⁶⁶C. Jung and D. Schindler, "The role of air density in wind energy assessment—A case study from Germany," *Energy* **171**, 385–392 (2019).
- ⁶⁷K. Azad, M. Rasul, P. Halder, and J. Sutariya, "Assessment of wind energy prospect by Weibull distribution for prospective wind sites in Australia," in *Energy Procedia* (Elsevier Ltd., 2019), pp. 348–355.
- ⁶⁸S. Ali, S. M. Lee, and C. M. Jang, "Statistical analysis of wind characteristics using Weibull and Rayleigh distributions in Deokjeok-do Island—Incheon, South Korea," *Renewable Energy* **123**, 652–663 (2018).
- ⁶⁹Z. H. Hulio, W. Jiang, and S. Rehman, "Techno—Economic assessment of wind power potential of Hawke's Bay using Weibull parameter: A review," *Energy Strategy Rev.* **26**, 100375 (2019).
- ⁷⁰J. C. Aririguzo and E. B. Ekwe, "Weibull distribution analysis of wind energy prospect for Umudike, Nigeria for power generation," *Rob. Comput. Integr. Manuf.* **55**, 160–163 (2019).
- ⁷¹M. H. Soulouknga, S. Y. Doka, N. Revanna, N. Djongyang, and T. C. Kofane, "Analysis of wind speed data and wind energy potential in Faya-Largeau, Chad, using Weibull distribution," *Renewable Energy* **121**, 1–8 (2018).
- ⁷²T. B. M. J. Ouarda, C. Charron, J.-Y. Shin, P. R. Marpu, A. H. Al-Mandoos, M. H. Al-Tamimi, H. Ghedira, and T. N. Al Hosary, "Probability distributions of wind speed in the UAE," *Energy Convers. Manage.* **93**, 414–434 (2015).
- ⁷³A. Madariaga, C. J. Martinez de Ilarduya, S. Ceballos, I. Martinez de Alegria, and J. L. Martin, "Electrical losses in multi-MW wind energy conversion systems," *Renewable Energy Power Qual. J.* **1**, 322–327 (2017).
- ⁷⁴J.-Y. Shin, T. B. M. J. Ouarda, and T. Lee, "Heterogeneous mixture distributions for modeling wind speed, application to the UAE," *Renewable Energy* **91**, 40–52 (2016).
- ⁷⁵O. Tsvetkova, *Uncertainty and Sensitivity Analysis for Decision Support in Building Design* (Albert Ludwig University of Freiburg, 2011).
- ⁷⁶S. Burhenne, O. Tsvetkova, D. Jacob, G. P. Henze, and A. Wagner, "Uncertainty quantification for combined building performance and cost-benefit analyses," *Build. Environ.* **62**, 143–154 (2013).