Influence of Stochastic Dependence on Small-Disturbance Stability and Ranking Uncertainties

Kazi N. Hasan, *Member, IEEE* and Robin Preece, *Member, IEEE*

***Abstract*—A high level of stochastic dependence (or correlation) exists between different uncertainties (i.e., loads and renewable generation), which is nonlinear and non-Gaussian and it affects power system stability. Accurate modeling of stochastic dependence becomes more important and influential as the penetration of correlated uncertainties (such as renewable generation) increases in the network. The stochastic dependence between uncertainties can be modeled using 1) copula theory and 2) joint probability distributions. These methods have been implemented in this paper and their performances have been compared in assessing the small-disturbance stability of a power system. The value of modeling stochastic dependence with increased renewables has been assessed. Subsequently, the critical uncertainties that most affect the damping of the most critical oscillatory mode have been identified and ranked in terms of their influence using advanced global sensitivity analysis techniques. This has enabled the quantification and identification of the impact of modeling stochastic dependence on the raking of critical uncertainties. The results suggest that multivariate Gaussian copula is the most suitable approach for modeling correlation as it shows consistently low error even at higher levels of renewable energy penetration into the power system.**

***Index Terms*—Copula, correlation, probabilistic assessment, small-disturbance stability, sensitivity analysis, stochastic dependence, uncertainty.**

NOMENCLATURE

BVC	Bivariate Clayton copula.
BVF	Bivariate Frank copula.
BVG	Bivariate Gumbel copula.
$C(\bullet)$	Copula function.
cdf	Cumulative distribution function.
Ind	Independent probability distributions.
JN	Joint normal distribution.
MVG	Multivariate Gaussian copula.
MVT	Multivariate Student's t copula.
N	Number of Monte Carlo simulations.
p	Number of uncertain (input) parameters.
pdf	Probability density function.
R_w	Weighted correlation coefficients.
S_i	Index of Sobol 1st order effect.
ST_i	Index of Sobol total effect.
$T(\bullet)$	Multivariate student's t distribution.
X	All input parameter set.
$\underline{\mathrm{X}}_{\mathrm{i}}$	ith input parameter.
$\overline{\mathrm{X}}$	Mean value of input parameter samples.
Υ	All output parameter set.
Υ_i	ith output parameter.
ε_{ARMS}	Average root mean square error.
ρ	Index of Pearson correlation coefficient.
$\Phi(\bullet)$	Multivariate normal distribution.

Manuscript received July 11, 2017; revised September 28, 2017; accepted November 30, 2017. Date of publication December 4, 2017; date of current version April 17, 2018. This work was supported by the EPSRC, U.K., through the Identifying Critical Uncertainties in Power Systems project (EP/N004310/1). Paper no. TPWRS-01004-2017. *(Corresponding author: Kazi N. Hasan.)*

The authors are with the School of Electrical and Electronic Engineering, University of Manchester, Manchester M60 1QD, U.K. (e-mail: kazi.hasan@manchester.ac.uk; robin.preece@manchester.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2017.2779887

I. INTRODUCTION

INCREASED proportions of renewable energy sources (RES) lessen the flexibility of power system operation due to the intermittent nature and spatiotemporal dependence among the RES sources. The inherent variations of RES in combination with the variation in system load bring more challenges in system operation in terms of balancing supply-demand, maintaining stability and minimizing RES curtailment. A high level of stochastic dependence (or correlation) exists between different loads and RES and these dependencies are nonlinear and non-Gaussian [1]. With the increased penetration of RES in the network, this stochastic dependence within loads and RES is becoming more prominent [1], [2].

Independent probability distributions or random sample generation techniques do not take account of correlation and may lead to notable error, and hence consideration of dependence structure is necessary [2]–[7]. Linear correlation among parameters does not reflect the true stochastic dependence among parameters and subsequently may lead to non-optimum solutions for real systems. Copula theory provides one effective way of modelling stochastic dependence between random variables. The stochastic dependence can be modelled using (a) copula theory (i.e., bivariate and multivariate copula), as well as through (b) joint probability distribution (such as multivariate joint normal distribution). All of these methods have been implemented and thoroughly analyzed in this study.

Different copula families reveal different dependence structures among correlated uncertainties. The most widely used copula families are Archimedean and elliptical copula [8]. Archimedean (i.e., Clayton, Frank, Gumbel) is suitable for representing complicated dependence structure, though restricted to two dimensions only. Elliptical copula (i.e., Gaussian, Student's t) can be extended to higher dimensions, representing multiple marginal distributions. A detailed discussion and study of the relevance of these methodologies and underlying theory with respect to power system applications are presented in Section II-D.

The suitability of copula modelling techniques varies depending on the structure of the dependence among parameters (such as symmetry, asymmetry, tail dependence etc.). The stochastic dependence within load, wind, and PV has been previously modelled through bivariate methods, using Clayton [7], [9], Frank [7], [9], and Gumbel [7], [9], [10] copulas. Multivariate methods such as Gaussian [2], [7], [9], [11] and Student's t [7] copulas have also been applied to power systems studies. Additional techniques which have also been used to capture correlation include joint (Gaussian) distribution [12], vine copula [6], pair copula [4], and dependent discrete convolution (DDC) [3]. The complexity and computational burden of the implementation of these later methods (i.e., vine and pair copula, and DDC) limit their widespread application [13]. Hence this paper implements and compares the most commonly used bivariate and multivariate copulas and multivariate joint normal distribution.

These modelling techniques have been previously applied to model the dependence between two sets of variables, such as wind-wind [1], [9], [11], [14], load-wind [2], load-load [7], load-PV [10], wind-PV [15] and PV-PV [4]. The level and structure of the stochastic dependence within the abovementioned interdependence and intra-dependence can vary significantly as will be shown in Section II. Though valuable, these previous works are limited to modelling two-factor dependence only. There is no work which assesses the impact of load-wind-PV stochastic dependence on a multivariate platform.

The stochastic dependence modeling techniques have previously been applied to load flow [9], generation scheduling [11], transmission planning [2], demand response [7], and load LVRT (low-voltage ride-through) applications [10]. Copula theory and joint probability distribution have had limited applications in power systems as mentioned above due to their complexity. In particular, the application of stochastic dependence is rare in power system stability assessment, despite a growing focus on probabilistic analysis. There is no work in the existing literature which models stochastic dependence of load-wind-PV (through copula theory and joint distribution) to assess their impact on power system stability.

Some earlier works by the author contributed to the identification of efficient computational techniques for modelling system uncertainties. An efficient computational technique based on Latin hypercube sampling to estimate the probability of small-disturbance instability has been presented in [16]. Implementation of a wide range of local and global sensitivity analysis techniques to identify the most important uncertainties in a power system has been presented in [17]. A comparison of three efficient estimation methods namely, point-estimate, Cumulant-based and probabilistic collocation methods to assess their performance in power system stability has been documented in [18]. However, none of these works have considered stochastic dependence among system uncertainties.

This is the first study to model stochastic dependence using a variety of methods in order to assess their impact on the accuracy of system stability measures (specifically the pdfs of critical mode damping). Additionally, this research performs a rigorous check of the stochastic dependence modelling by using parameter ranking with sensitivity analysis as an additional measure of the accuracy of the methods.

The impact of higher penetrations of renewable energy sources (RES) has been analyzed with respect to the impact on critical mode damping. The accuracy of stochastic dependence modeling at diverse levels of RES is presented, which highlights the importance of modelling stochastic dependence and using the right method for low/high RES penetration into the network. Following the modelling of stochastic dependence during the sampling procedure, the priority ranking of critical uncertainties to identify those having the greatest impact on small-disturbance stability has been completed using the global Sobol sensitivity analysis technique [19]. This has revealed how the stochastic dependence among uncertainties and their modelling approaches affects parameter ranking and can lead to (in)accurate results. An accuracy measure based on the correlation coefficients among ranking matrices has also been presented to determine the robustness of the ranking.

This paper makes the following novel contributions:

1) A clear demonstration of modelling stochastic dependence among system uncertainties using six different techniques to explicitly capture correlation.
2) A comparison of the accuracy of stochastic dependence modelling through bi-variate and multi-variate copula and joint normal distribution.
3) Assessment of the significance of modelling stochastic dependence for variable levels of RES penetration and identification of appropriate modelling tools.
4) The evaluation of critical uncertainties through power system stability indicators considering the impact of stochastic dependence.
5) Identification of the most appropriate method for modelling stochastic dependence for small-disturbance stability studies based on thorough error analysis against realistic data.

II. Modelling Stochastic Dependence

In order to assess the accuracy of different methods for modelling stochastic dependence, a benchmark data set is required. A data set has been collected from various sources which includes system loads at different substations, as well as wind speeds and solar irradiation at different weather stations. This benchmark raw data has been re-generated using:

1) (i) independent probability distributions (Ind) where the correlation is ignored (current common practice);
2) Some *multivariate copulas*: (ii) Gaussian (MVG) (iii) Student t (MVT);
3) A number of *bivariate copulas*: (iv) Clayton (BVC), (v) Frank (BVF), (vi) Gumbel (BVG); and
4) (vii) multivariate joint normal distribution (JN).

A. System Uncertainties

Hourly power system load, wind and solar profile have been obtained across the entire annual cycle to capture the weather pattern and spatio-temporal dependence [20], [21]. Fig. 1 presents the correaltion coeficient matrix representing

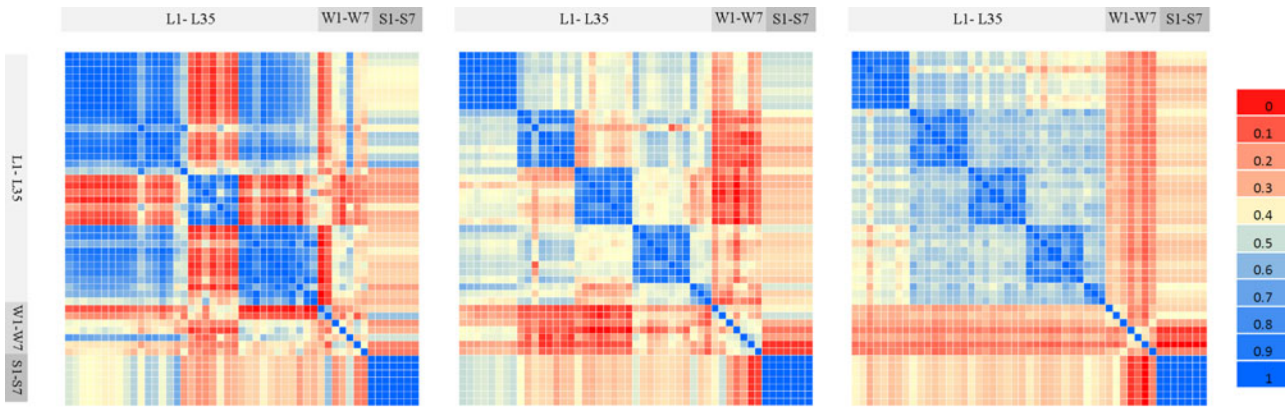

Fig. 1. Stochastic dependence pattern of power system load, wind speed and solar irradiance, in different timeframes of a (a) day, (b) month and (c) year.

the intra-dependence and interdependence among load demand, wind speed and solar irradiance, detailed further below.

B. Correlation Among System Uncertainties

Fig. 1 is a 49×49 matrix of Pearson correlation coefficients between the 49 considered system uncertainties within this work. A similar figure could be produced for any data set and any test system. In Fig. 1, row/column 1–35 are loads, 36–42 are wind speeds and 43–49 are solar irradiance. It can be seen that a high level of correlation exists between some sets of system uncertainties, such as load-load and PV-PV within row/column 1–35 and 43–49 respectively. The yearly correlation has been used in this study for stochastic dependence modelling, however, the method would not change if different durations of (daily, weekly, monthly etc.) correlation parameters were selected.

The correlation coefficients are changing with the duration of data samples as shown in Fig. 1, which represents data for a (a) day, (b) month and (c) year. The levels of intra-dependence and interdependence within some sets of system uncertainties vary as the duration of the system uncertainties changes. Similarly, the different granularity of the data samples may change the correlation coefficients (which cannot be produced due to the unavailability of the data). Hence, by changing the granularity and duration of the data samples, there could be either lower or higher levels of dependence among system uncertainties. These lower or higher levels of dependence do not change the focus and scope of the paper. Rather, the paper emphasizes the value of correlation modelling by appropriate approaches.

1) Intra-Dependence Within Parameter Groups: Intra-dependence among parameters is mainly linked through the spatiotemporal factors. The load-load intra-dependence is affected by the weather pattern, locality, temperature variation, and daily routine or personal lifestyle of the consumers. Closely located and similar types of (i.e., domestic or commercial or industrial) loads are highly correlated due to the same consumption pattern. The correlation pattern is moderate among different types of (i.e., domestic or commercial or industrial) consumers. The wind-wind intra-dependence could be high if they are closely located. As the distance between the wind farms increases, their intra-dependence tends to decrease. The same is true for PV-PV intra-dependence as in wind. But, when the sampling rate and time scale is longer the PV-PV correlation becomes increasingly high as, in this case, it becomes mainly dependent on day-time duration.

2) Interdependence Among Parameter Groups: The load-wind interdependence is very low for obvious reasons as the consumption pattern of the customers is not related to the spatio-temporal variability of the wind [22]. On the other hand, the load-PV interdependence is at a moderate level as temperature increases (due to high solar irradiance) can result in air conditioning load increase accordingly [23]. It can be seen from Fig. 1 that the wind-PV interdependence is very low which means there is low correlation between the wind speed and solar irradiance across the year [24].

C. Non-Correlated Sampling

Independent modelling refers to the probabilistic modelling of load, wind and PV data ignoring the stochastic dependence among these parameters. This represents the true marginal distribution of the load, wind, and PV, which is obtained by transforming the system uncertainties to the unit square using a kernel estimator of the cumulative distribution function [25].

D. Correlation Modelling by Copula Theory

Copula theory provides an effective way of modeling stochastic dependence (or correlation) between random variables. According to Sklar's theorem: any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula which describes the dependence structure between the variables [26].

In copula modelling, correlated samples are generated by four consecutive stages form the raw data sets: (i) transforming the data to the unit square using a kernel estimator of the cumulative distribution function, (ii) fitting a selected copula to the data and obtaining the copula parameter (i.e., matrix of linear correlation and degrees of freedom), (iii) generating random samples from the selected copula, and (iv) transforming the random sample back to the original scale of the data. The steps (i) and (iv) are the same for all copula and represents a normalization of the data, and subsequent re-scaling, while steps (ii) and (iii) are copula-dependent.

The copula function C can be represented by the multivariate cdf F (cumulative distribution function) and marginal cdf F_i, as in (1) [27]. The different copulas considered in this work will be described one by one.

$$C\left[F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\right] = F(x_1, x_2, \ldots, x_n) \quad (1)$$

1) Multivariate Gaussian Copula: Multivariate Gaussian copula is extensively used in financial modelling, which is given by (2) [8].

$$C(u_1, u_2, ..., u_n; \Sigma) = \Phi_\Sigma\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_n)\right) \quad (2)$$

In (2), Σ denotes a symmetric, positive definite matrix with $diag(\Sigma) = 1$, Φ_Σ is the standard multivariate normal distribution with correlation matrix Σ, and $\Phi^{-1}(\bullet)$ is the inverse of the normal cdf. This copula is known to have radial symmetry.

2) Multivariate Student's t Copula: The multivariate Student's t copula is given by (3) [8].

$$C(u_1, u_2, ..., u_n; \Sigma, v) = T_{\Sigma,v}\left(t_v{}^{-1}(u_1), t_v{}^{-1}(u_2), ..., t_v{}^{-1}(u_n)\right) \quad (3)$$

In (3), Σ denotes a symmetric, positive definite matrix with $diag(\Sigma) = 1$, v is the degree of freedom, $T_{\Sigma,v}(\bullet)$ is the standard multivariate student distribution with correlation matrix Σ, and $t_v^{-1}(\bullet)$ is the inverse of the student cdf. This copula is also known to have radial symmetry.

3) Bivariate Clayton Copula: Clayton copula is defined as in (4) [8].

$$C_\theta(u, v) = \left[\max\{u^{-\theta} + v^{-\theta} - 1; 0\}\right]^{-1/\theta} \quad (4)$$

In (4), $C_\theta(\bullet)$ represents Clayton copula, $\theta \in [-1, \infty)\backslash\{0\}$, and u, v are two corresponding variables. Clayton copula is known to have strong lower tail dependence.

In this work, all bivariate copulas are modelled with respect to load 1. This is an arbitrary choice and other selections for bivariate modelling could be made.

4) Bivariate Frank Copula: Frank copula can be presented as in (5) [8].

$$C_\theta(u, v) = -\frac{1}{\theta}\log\left[1 + \frac{\left(e^{-\theta_u} - 1\right)\left(e^{-\theta_v} - 1\right)}{\left(e^{-\theta} - 1\right)}\right] \quad (5)$$

In (5), $C_\theta(\bullet)$ represents Frank copula, $\theta \in \Re\backslash\{0\}$, and u, v are two corresponding variables.

5) Bivariate Gumbel Copula: Gumbel copula is another bivariate copula and is defined as in (6) [8].

$$C_\theta(u, v) = \exp\left[-\left((-\log(u))^\theta + (-\log(v))^\theta\right)^{1/\theta}\right] \quad (6)$$

In (6), $C_\theta(\bullet)$ represents Gumbel copula, $\theta \in [1, \infty)$, and u, v are two corresponding variables. Gumbel copula is asymmetric and puts more weight in the right tail. It has previously shown superior performance in modelling correlated wind data [11].

E. Multivariate Joint Normal Distribution

The multivariate joint normal distribution (JN) considers the dependence among uncertainties assuming all parameters follow a normal distribution. The probability density function of the p-dimensional JN can be written as (7) [28].

$$\Phi_\Sigma = \left[1\Big/\sqrt{|\Sigma|(2\pi)^p}\right]\cdot e^{-(1/2)(x-\mu)'\Sigma^{-1}(x-\mu)} \quad (7)$$

In (7), x and μ are data series and mean, respectively, both are $1 \times P$ vectors, Σ is the covariance matrix, a $p \times p$ vector, and $(\bullet)'$ denotes transpose matrix. The diagonal elements of Σ contain the variance of each variable, while the off-diagonal elements of Σ contain the covariance between variables.

F. Summary of Alternative Sampling Techniques

The discussion of the abovementioned methods provides an insight of the methodologies, underlying theory, and applicability areas. Fig. 2 shows the stochastic dependence pattern (shown in the unit square) of two arbitrary system uncertainties (load 1 and load 2) when modelled using different sampling techniques. This paper revolves around 3 types of modelling techniques, which consider either 'probability distribution' or 'stochastic dependence', or both, or none of them. These methods are: (1) independent distribution, which considers probability distributions, but neglects the stochastic dependence among uncertainties, (2) joint normal distribution, which considers (normal) probability distributions and stochastic dependence among uncertainties, but neglects non-normal probability distributions, and (3) copula techniques, which considers both probability distributions and stochastic dependence among uncertainties. The yearly correlation has been used in this study for stochastic dependence modelling, however, the method would not change if different time span were selected. The differences in how the stochastic dependence is captured by the different modelling approaches are clearly evident from Fig. 2. This data is used as an input to OPF and subsequent modal analysis to assess system small-disturbance stability.

III. Test Network and Simulation Details

All techniques for stochastic dependence modelling discussed above are illustrated and analyzed using OPF simulation and small-disturbance stability analysis of the NETS-NYPS test power system [29]. Probabilistic modelling of the input parameters and calculations have been performed in MATLAB using OPF solvers from MATPOWER [30]. Modal analysis has been conducted in DIgSILENT PowerFactory.

A. NETS-NYPS Test System

The simulation network is a substantially modified version of the NETS-NYPS test system (New England Test System – New York Power System) with a high amount of RES penetration. The network has 5 areas, 16 large synchronous machines and 68 buses, as shown in Fig. 3, as well as the addition of variable capacity of RES (as will be discussed in Section IV). Network data, component modelling and more information of the test system are available in [29].

B. Probabilistic Modelling

A sufficient number of Monte Carlo samples is determined through a stopping rule as in (8) [16].

$$\mathrm{E} > \left[\left\{\Phi^{-1}(1 - \delta/2)\cdot\sqrt{\sigma^2(\mathrm{X})/\mathrm{N}}\right\}\Big/\overline{\mathrm{X}}\right] \quad (8)$$

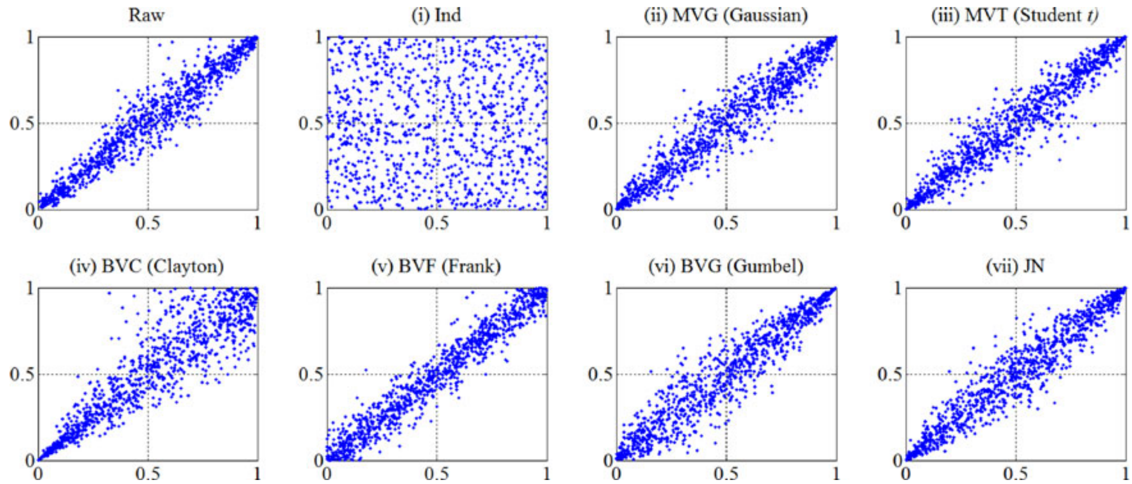

Fig. 2. An illustrative example of modelling stochastic dependence (between load 1 and load 2) through alternative techniques including raw data, (i) independent samples (Ind), (ii) multivariate Gaussian copula, (iii) multivariate Student's t copula, (iv) bivariate Clayton copula, (v) bivariate Frank copula, (vi) bivariate Gumbel copula and (vii) multivariate joint normal distribution.

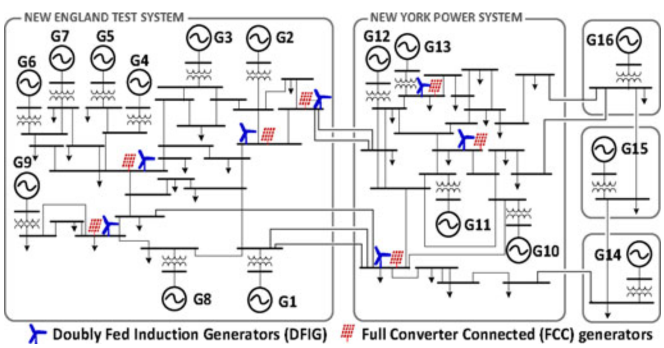

Fig. 3. Modified NETS-NYPS (New England Test System – New York Power System) network with a high amount of wind and solar generation.

TABLE I
NUMBER OF SIMULATIONS REQUIRED USING THE MC STOPPING RULE TO OBTAIN $E = 0.3\%$ WITH 99% CONFIDENCE

Simulation techniques	Raw	Ind	MVG	MVT	BVC	BVF	BVG	JN
Number of simulations	1000	2600	980	1390	910	923	918	958
Computational time (min)	31	57	32	33	32	32	32	37

In (8), $\Phi^{-1}(\bullet)$ represents the inverse normal cdf with a mean of zero and standard deviation of one, $\sigma^2(\bullet)$ is the variance of a sample, and δ is the desired confidence level. As presented in (8), simulations can be stopped if the calculated sample mean error falls below a specific threshold, E.

For the raw data, a 0.3% sample mean error is obtained with N = 1000 simulations with a required simulation time of 31 minutes. Table I presents the number of simulations and computational time for various techniques to obtain the same 0.3% error according to the MC stopping rule presented by (8). These simulation numbers will be used for the remainder of the work to ensure the sampling error is consistent.

C. Global Sensitivity Analysis Method - Sobol Total Indices

Sensitivity analysis determines how the input variations propagate through the system to cause variations in the output [31]. A global sensitivity analysis considers the whole range of variation of inputs on the system output. Sobol method is a variance-based method, which is very useful in case of non-linear and non-monotonic models. Sobol total indices are the sum of all the sensitivity indices involving all uncertain factors as presented by (9) [17], [31].

$$S_T = S_i + \sum_{i<j} S_{ij} + \sum_{i \neq j, i \neq k, j<k} S_{ijk} + \ldots \tag{9}$$

Where, S_i is the 1st-order sensitivity index for i, S_{ij} is the 2nd-order sensitivity index describing the interactions between two uncertainties i and $j (\neq i)$. In (9), S_i is the first order sensitivity index and can be expressed by (10), with variance of output with respect to a particular input i by (11).

$$S_i = D_i(\Upsilon)/Var(\Upsilon) \tag{10}$$

$$D_i(\Upsilon) = Var\left[\mathrm{E}(\Upsilon|\mathrm{X_i})\right] \tag{11}$$

Similarly, the 2nd order sensitivity index and variance can be presented by (12) and (13), respectively.

$$S_{ij} = D_{ij}(\Upsilon)/Var(\Upsilon) \tag{12}$$

$$D_{ij}(\Upsilon) = Var\left[\mathrm{E}(\Upsilon|\mathrm{X_i}, \mathrm{X_j})\right] - D_i(\Upsilon) - D_j(\Upsilon) \tag{13}$$

Subsequently, S_{ijk} is the 3rd-order sensitivity index for three uncertainties, $i, j, k\,(i \neq j \neq k)$. These interactions will continue up to pth order for p parameters.

S_i is the Sobol 1st order effect which models the change in variance that is seen when one parameter is no longer considered as uncertain. This is determined by obtaining the correlation coefficient of the output vector from two model runs in which all values for variables in X_i are common, but all other inputs use independent samples. In determining the Sobol total indices, input data set X is partitioned into $X_{\sim i}$ and X_i, where $X_{\sim i}$ is the set of all input variables which include a variation in the ith

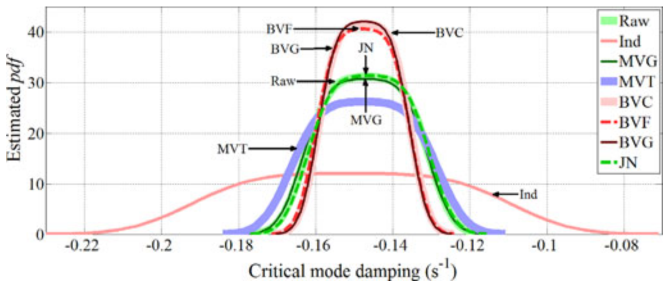

Fig. 4. Kernel smoothing density pdf of critical mode damping (with up to 15% renewable generation).

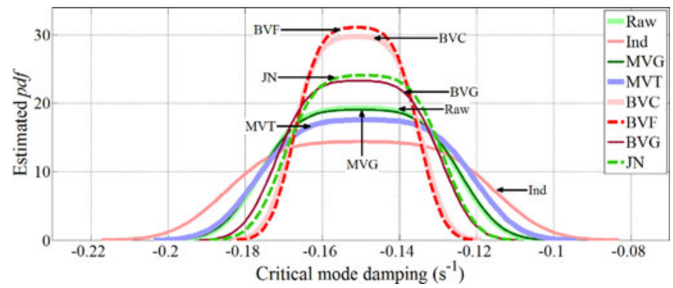

Fig. 5. Kernel smoothing density pdf of critical mode damping (with up to 30% renewable generation).

index of X. The total effect is then calculated by (14).

$$ST_i = 1 - S_{\sim i} \tag{14}$$

where, $S_{\sim i}$ is the sum of the all terms that include the variation in X_i. The Sobol method has been efficiently implemented previously in environmental and hydrological models [32] and recently in power system studies [17], [19].

IV. Results and Analysis

This section discusses the simulation results obtained by a multi-level approach consisting of the correlation modelling, optimal power flow, modal analysis, and subsequent sensitivity analysis. Firstly, the pdfs of oscillation damping will be presented for 15% RES with system input data generated through different techniques discussed in Section II. Secondly, the pdfs of oscillation damping will be presented for 30% RES. These results will compare the accuracy of stochastic dependence modelling techniques by highlighting the impact of increased RES on the accuracy of stochastic dependence modelling techniques.

Thirdly, expanded accuracy measures of stochastic dependence modelling techniques will be presented for 10%, 20%, 30% 40% and 50% RES penetration. These will highlight the impact of increased RES on the accuracy of stochastic dependence modelling techniques for a wider range of considered penetrations. Fourthly, the most influential parameters affecting the oscillation damping will be ranked through the sensitivity analysis technique (as a demonstration for 30% RES). Finally, the accuracy measures of the ranking of uncertainties by alternative modelling techniques will be quantified using a weighted correlation coefficient measure. This will identify the stochastic dependence modelling technique that most accurately captures the true importance ranking of uncertainties.

A. Probabilistic Modal Analysis

Probabilistic modal analysis has been performed with the data sets generated through alternative methods (as discussed in Section II). Fig. 4 presents the obtained pdf of critical oscillation damping with 15% RES. The $pdfs$ from the original raw data set have been compared with the results obtained from copula theory (i.e., bivariate and multivariate copula) and joint probability distribution.

It can be seen from Fig. 4 that the MVG and JN approaches produce the closest results compared to the original raw data. As expected, the independent samples, which ignore the stochastic dependence among uncertainties show the worst performance followed by bivariate copulas. The MVT shows better performance compared to bivariate copulas.

A high amount of normally distributed parameters (i.e., system loads) and relatively low RES penetration (15%) (as presented in Fig. 4) make the JN a suitable option for modeling stochastic dependence among system uncertainties in this case. However, there is a need to further study the impact of higher non-synchronous penetration.

The spread of the critical mode damping represents the stability indication for the simulated system. It can be seen that the system stability is low with the non-correlated samples. If a system is designed and controllers are tuned based on the stability indicators obtained with non-correlated uncertain samples that may lead to poorly designed and non-optimal solutions. The prior discussion in Section II reveals a strong correlation among input parameters considered. Hence, correlated system inputs reflect the true performance indicators. In such cases, appropriate resources can be allocated and design parameters can be selected for secure and optimal operation of the system.

B. Impact of High Amount of Wind and Solar Power

The impact of 30% penetration of RES (non-normally distributed wind and solar power) has been presented in Fig. 5. It can be seen from the estimated pdf of damping that the independent sample has the worst performance followed by bi-variate (Clayton and Frank) copula. The most noticeable difference compared to the 15% RES case is that the performance of JN has deteriorated (contrast with Fig. 4). On the other hand, the performance of bivariate Gumbel copula, which explicitly models tail dependence, is better than the previous case. The multivariate Student's t copula remains unchanged with this increase of RES. Still, the performance of multivariate copulas is better than the bivariate copulas.

To represent a comparison of estimated pdfs, the average root mean square error given by (15) has been calculated [18].

$$\varepsilon_{ARMS} = \sqrt{\left[\frac{1}{N}\sum_{i=1}^{N}\left(F_i^o - F_i^X\right)^2\right]} \tag{15}$$

In (15), F_i^O is the $i-$th value of the pdf obtained using the original samples, F_i^X is the i-th value of the pdf obtained through other approaches, and N is the number of samples considered calculating the $\varepsilon_{\mathrm{ARMS}}$.

The impact of 15% and 30% penetration of RES on the accuracy of the pdfs of damping produced using the different sampling methods is presented in Table II. According to the numerical values of ε_{ARMS}, the methods are ranked in the

TABLE II
AVERAGE ROOT MEAN SQUARE ERROR OF PDF OF CRITICAL MODE DAMPING

Simulation techniques	Ind	MVG	MVT	BVC	BVF	BVG	JN
15% RES	11%	0.5%	3%	7.3%	7%	6.5%	1.2%
30% RES	11%	0.5%	2.8%	7.5%	7.5%	6%	2.8%

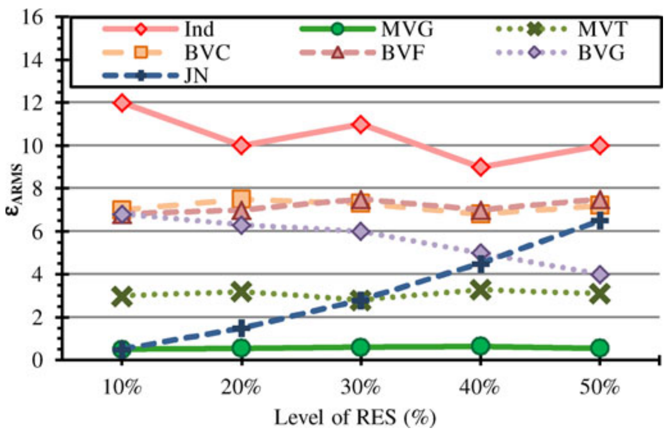

Fig. 6. Variation of $\varepsilon_{\text{ARMS}}$ of pdfs of critical mode damping for various stochastic dependence modelling approaches.

following order for accuracy for a 15% penetration level: MVG, JN, MVT, BVG, BVF, BVC and Ind. MVG maintains its accuracy at both penetration levels having the lowest ε_{ARMS} of 0.5%. As the level of RES penetration increases to 30%, the performance of JN becomes worse with ε_{ARMS} increasing from 1.2% to 2.8%, whereas independent (non-correlated) sampling is always the worst possible choice with ε_{ARMS} at 11%.

To further explore the accuracy of the methods at different levels of RES penetration, a rigorous analysis is performed using a wide range of RES penetration from 10% to 50%.

C. Accuracy of Stochastic Dependence Modelling

Fig. 6 shows the calculated ε_{ARMS} of the pdf of oscillation damping obtained through different stochastic dependence modelling approaches and different levels of RES penetration. According to the numerical values of ε_{ARMS}, the methods are ranked in the following order for accuracy for a 10% penetration level: MVG, JN, MVT, BVG, BVF, BVC, and Ind. The performance of JN deteriorates significantly with the increased levels of RES, whereas Ind always remains the worst possible choice. On the other hand, the performances of bivariate and multivariate copulas are generally maintained across all levels of RES penetration, with the bivariate Gumbel copula which is known to capture wind correlation well becoming better with increased levels of penetration.

The bivariate copulas are less accurate than the multivariate copulas, due to the bivariate nature of modelling dependence with respect to a single system uncertainty (i.e., load 1 in this study). On the other hand, the multivariate Gaussian copula uses the covariance matrix which considers stochastic dependence between all system uncertainties. The multivariate copulas perform particularly well with low error across all levels of RES penetration.

This is a key result, not only highlighting the importance of modelling stochastic dependence, but also the importance of using the right method to model the stochastic dependence and not simply assuming that a JN distribution will be applicable. It has been shown that as the proportion of non-normally distributed uncertainties increases, more advanced techniques (which suitably capture the stochastic dependence as well as accurately representing the marginal distribution) for modelling stochastic dependence must be used.

The changes in power system topology have been implemented in the paper by using 10% to 50% levels of RES (renewable energy sources) penetration into the system. It can be seen from the results that the accuracy of the methods changes slightly, but the trend and the ranking of the accuracy of the methods (as shown in Fig. 6) remain the same.

D. Uncertainty Importance Measures

It has been shown that accurate modelling of correlated uncertainties is vital, this section will now further explore how the priority of system uncertainties is affected by stochastic dependence modelling. This study evaluates the importance of system uncertainties in the assessment of critical mode oscillation damping. Sobol sensitivity analysis technique has been implemented to obtain the parameter ranking.

A heatmap of the ranking of uncertain parameters through the Sobol technique is shown in Fig. 7 (for 30% RES). The columns in Fig. 7 show the importance of individual input parameters on small-disturbance stability through the measure of damping of critical eigenvalues. The rows of Fig. 7 compare the ranking of input parameters by different techniques of stochastic dependence modelling. A general overview of the heat map reveals that the MVG, MVT, and JN have a similar pattern in the darkness of the color shades – this means that they produce similar values of parameter importance. These values are similar to raw data ranking. On the other hand, the three bivariate copula techniques, Clayton, Frank, and Gumbel have similar color spreads which are closer to independent sampling methods.

When loads are modelled with the accurate representation of their stochastic dependence, top ranking loads appear as groups instead of as single parameters (for MVG and MVT). This is an important feature of correlation modeling that can be identified from Fig. 7. The most important parameters are $L_{17} \sim L_{24}$ and $L_{41} \sim L_{49}$ (looking at the results for MVG and MVT). Herein lies the significance of modelling them as correlated parameters. The independent modeling of parameters loses correlation features of the data and may ignore an important parameter of a power system, which might have a high correlation with an influential parameter and subsequent high importance (such as $L_{17} \sim L_{24}$ and $L_{41} \sim L_{49}$ in Fig. 7). The previous heat map of uncertainty correlation in Fig. 1 revealed high correlation among groups of loads which emphasizes the need for correlation modelling.

Fig. 7 presents a visual representation of the sensitivity measures of the ranking of important parameters. To determine the accuracy of the ranking of system uncertainties by alternative methods to the ranking based on the raw data, a mathematical measure is required. For that purpose, the weighted

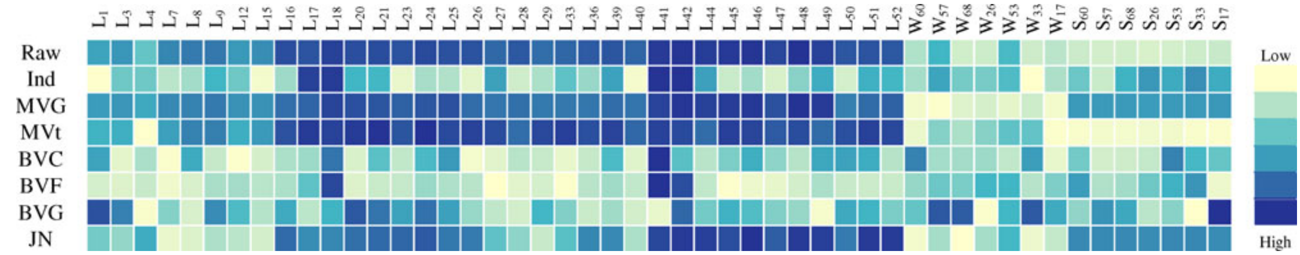

Fig. 7. Heatmap of uncertainty importance measures representing the performance comparison among alternative correlation modelling approaches.

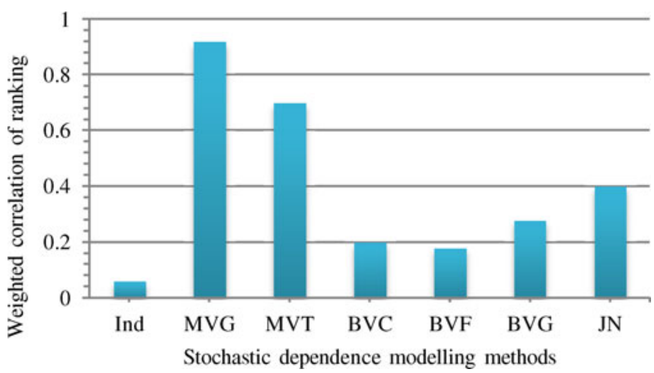

Fig. 8. Weighted correlation coefficients of ranking matrices generated through different stochastic dependence modelling approaches.

correlation coefficients among the ranking matrices of alternative approaches have been obtained by (16) [33].

$$R_w = \frac{\sum_{i=1}^{k} w_i \left(\rho_i^x - \overline{\rho^x}\right)\left(\rho_i^o - \overline{\rho^o}\right)}{\sqrt{\sum_{i=1}^{k} w_i \left(\rho_i^x - \overline{\rho^x}\right)^2 \cdot \sum_{i=1}^{k} w_i \left(\rho_i^o - \overline{\rho^o}\right)^2}} \tag{16}$$

Weighted correlation coefficients have been used in this process to represent different degrees of importance [34]. For example, among 49 parameters (as ranked in Fig. 7 by eight alternative methods), the most important parameter has the maximum weight which is 49 and the least important parameter has the minimum weight which is 1. Hence, the weight vector, w, is taken as uniformly distributed from 1 to 49 (in ascending order) according to ranking. The rationale behind the ascending weight vector is logical and is used to reflect the fact that accurately identifying and capturing the rank order of the most important and influential parameters is more valuable than the least important parameters. Of course, a standard correlation analysis could also be used.

Fig. 8 presents the weighted correlation of ranks compared to the raw data set for the alternative stochastic dependence modelling methods. A value of 1 would represent identical rank orders between the raw data and a given modelling approach. It is clear that the MVG copula method is again confirmed as the best approach as its ranking most closely resembles that obtained from the raw data set. This is followed by the MVT copula and JN distribution. As expected, the independent sample generation technique performs extremely poorly. While the bivariate copula techniques, namely Clayton, Frank, and Gumbel, remain in the middle, where they have moderate performance (measured through R_w). The analysis of uncertainty ranking further confirms the previously obtained results with respect to the performance of the different stochastic dependence modelling methods.

V. Conclusion

Inherent intra-dependence and interdependence among input parameters have been incorporated within the probabilistic modal analysis and priority ranking of uncertainties. Analysis of real data reveals a high level of correlation within the system parameters i.e., within load-load and PV-PV. There is a negligible correlation between system load-wind and wind-PV. These correlation patterns need to be accurately incorporated into the power system analysis.

This study demonstrates the modelling of stochastic dependence among system uncertainties using six different techniques. The impact of the stochastic dependence is assessed by using a system stability indicator represented by the critical mode oscillation damping. The accuracy of all correlation modelling techniques has been presented.

It has been shown that joint normal distribution (JN) is adequate for modelling stochastic dependence of system uncertainties when system RES penetration is low ($\sim$15%). With an increase in RES ($>$15%), the advantages of modelling stochastic dependence through copula theory can be observed. The performance of multivariate Gaussian copula is better than all other methods across all levels of RES penetration.

The results demonstrate that independent modelling, which ignores the existing correlation among uncertainties may lead to very inaccurate solutions. Bivariate copula methods provide moderate accuracy and multivariate copulas are more accurate. With the increased RES, the performance of JN gets worse with increased non-Gaussian uncertainties and the performance of bivariate Gumbel copula gets better.

The yearly correlation has been used in this study for stochastic dependence modelling, however, the method would not change if different time span were selected. The changes in the granularity and duration of the data samples may result in lower or higher levels of dependence among system uncertainties. Varying levels of interactions among system uncertainties and topological changes in the network have been simulated by introducing varying levels of RES penetration into the network and these results have demonstrated the benefits of the copula approaches.

This study, for the first time, has mapped through the stochastic dependence of uncertainties to the uncertainty importance measures. Earlier studies generated the correlated samples and completed stability simulation with the generated samples but did not validate the dependence structure again to the sensitivity analysis. This enables the tracking of propagation of stochastic dependence throughout the power flow, modal analysis and sensitivity simulation processes.

As the proportion of intermittent resources and correlated uncertainties are increasing in power networks, appropriate modelling of stochastic dependence will remain a vital issue in the performance analysis of power systems.

References

[1] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.

[2] H. Park, R. Baldick, and D. P. Morton, "A stochastic transmission planning model with dependent load and wind forecasts," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3003–3011, Nov. 2015.

[3] N. Zhang, C. Kang, C. Singh, and Q. Xia, "Copula based dependent discrete convolution for power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5204-5205, Nov. 2016.

[4] W. Wu, K. Wang, B. Han, G. Li, X. Jiang, and M. L. Crow, "A versatile probability model of photovoltaic generation using pair copula construction," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1337–1345, Oct. 2015.

[5] P. Li, X. Guan, J. Wu, and X. Zhou, "Modeling dynamic spatial correlations of geographically distributed wind farms and constructing ellipsoidal uncertainty sets for optimization-based generation scheduling," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1594–1605, Oct. 2015.

[6] H. V. Haghi and S. Lotfifard, "Spatiotemporal modeling of wind generation for optimal energy storage sizing," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 113–121, Jan. 2015.

[7] M. T. Bina and D. Ahmadi, "Stochastic modeling for the next day domestic demand response applications," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2880–2893, Nov. 2015.

[8] R. Hentati and J.-L. Prigent, *Chapter 4 Copula Theory Applied to Hedge Funds Dependence Structure Determination* (International Symposia in Economic Theory and Econometrics Series). Bingley, U.K.: Emerald, 2010, pp. 83–109.

[9] D. Cai, D. Shi, and J. Chen, "Probabilistic load flow computation using copula and Latin hypercube sampling," *IET Gener., Transm. Distrib.*, vol. 8, pp. 1539–1549, 2014.

[10] N. Saadat, S. S. Choi, and D. M. Vilathgamuwa, "A statistical evaluation of the capability of distributed renewable generator-energy-storage system in providing load low-voltage ride-through," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1128–1136, Jun. 2015.

[11] N. Zhang, C. Kang, Q. Xia, and J. Liang, "Modeling conditional forecast error for wind power in generation scheduling," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1316–1324, May 2014.

[12] H. Park and R. Baldick, "Stochastic generation capacity expansion planning reducing greenhouse gas emissions," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 1026–1034, Mar. 2015.

[13] *Copula Theory and Its Applications*. Berlin, Germany: Springer-Verlag, 2009.

[14] X. Jin, W. Wu, K. Wang, G. Li, and B. Han, "Probabilistic small signal analysis considering wind power correlation," in *Proc. IEEE Power Energy Soc. Gen. Meet.*, 2016, pp. 1–5.

[15] G. Yang, M. Zhou, B. Lin, and W. Du, "Optimal scheduling the wind-solar-storage hybrid generation system considering wind-solar correlation," in *Proc. IEEE PES Asia Pac. Power Energy Eng. Conf.*, 2013, pp. 1–6.

[16] R. Preece and J. V. Milanovic, "Efficient estimation of the probability of small-disturbance instability of large uncertain power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1063–1072, Mar. 2016.

[17] R. Preece and J. V. Milanovic, "Assessing the applicability of uncertainty importance measures for power system studies," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2076–2084, May 2016.

[18] R. Preece, H. Kaijia, and J. V. Milanovic, "Probabilistic small-disturbance stability assessment of uncertain power systems using efficient estimation methods," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2509–2517, Sep. 2014.

[19] K. N. Hasan, R. Preece, and J. V. Milanovic, "Priority ranking of critical uncertainties affecting small-disturbance stability using sensitivity analysis techniques," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2629–2639, Jul. 2017.

[20] ERCOT, Austin, TX, USA. Hourly Load Data Archives. (2017) [Online]. Available: http://www.ercot.com/gridinfo/load/load_hist/

[21] AgriMet. Historical Dayfile Data Access. (2017) [Online]. Available: https://www.usbr.gov/pn/agrimet/webaghrread.html

[22] B. Hu, L. Wu, and M. Marwali, "On the robust solution to SCUC with load and wind uncertainty correlations," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2952–2964, Nov. 2014.

[23] P. Anderson, B. Efaw, and E. McKinney, "A method for determining the relationship between solar irradiance and distribution feeder peak loading," in *Proc. IEEE/PES Transm. Distrib. Conf. Expo.*, 2016, pp. 1–5.

[24] J. Sexauer and S. Mohagheghi, "Hybrid stochastic short-term models for wind and solar energy trajectories," in *Proc. IEEE Green Tech. Conf.*, 2015, pp. 191–198.

[25] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*. New York, NY, USA: Wiley, 2013.

[26] A. Sklar, "Fonctions de répartition à n dimensions etleursmarges," *Publications de l'Institut de Statistique de Université de Paris*, vol. 8, pp. 229–231, 1959.

[27] H. V. Haghi, M. T. Bina, and M. A. Golkar, "Nonlinear modeling of temporal wind power variations," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 838–848, Oct. 2013.

[28] T. T. Soong, *Fundamentals of Probability and Statistics for Engineers*. Chichester, U.K.: Wiley-Interscience, 2004.

[29] G. Rogers, *Power System Oscillations*. Norwell, MA, USA: Kluwer, 2000.

[30] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[31] A. Saltelli, K. Chan, and E. M. Scott, *Sensitivity Analysis*. New York, NY, USA: Wiley, 2008.

[32] B. Iooss and P. Lemaıtre, *A Review on Global Sensitivity Analysis Methods*. New York, NY, USA: Springer, 2015.

[33] E. Borgonovo, "Measuring uncertainty importance: Investigation and comparison of alternative approaches," *Risk Anal.*, vol. 26, pp. 1349–1361, 2006.

[34] Y. Liu, Q. Meng, R. Chen, J. Wang, S. Jiang, and Y. Hu, "A new method to evaluate the similarity of chromatographic fingerprints: Weighted Pearson product-moment correlation coefficient," *J. Chromatographic Sci.*, vol. 42, pp. 545–550, 2004.

Kazi N. Hasan (GS'08–M'14) received the Graduate degree in electrical and electronic engineering from Bangladesh University of Engineering and Technology in 2006, the Master of Engineering degree from the University of Tasmania in 2009, and the Ph.D. degree from the University of Queensland in 2013. He is currently a Postdoctoral Researcher with the University of Manchester, Manchester, U.K. His research focuses on analyzing the impact of renewable energy into the electricity grid.

Robin Preece (GS'10–M'13) received the B.Eng. degree in electrical and electronic engineering and the Ph.D. degree from the University of Manchester, Manchester, U.K., in 2009 and 2013, respectively. He is currently a Lecturer of power systems engineering with the University of Manchester. His research interests include power system dynamic security, mixed AC/DC system performance, and impact of uncertainty on power systems.