

This is the accepted manuscript, which has been accepted by IEEE for publication © 2014. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The full reference is:

“Assessing the Applicability of Uncertainty Importance Measures for Power System Studies”

R. Preece, J.V. Milanović

IEEE Transactions on Power Systems, 2015

DOI: [10.1109/TPWRS.2015.2449082](https://doi.org/10.1109/TPWRS.2015.2449082)

Assessing the Applicability of Uncertainty Importance Measures for Power System Studies

Robin Preece, *Member, IEEE*, and Jovica V. Milanović, *Fellow, IEEE*

Abstract—This paper critically evaluates a number of uncertainty importance measures for use in power system stability studies. Sensitivity analysis of uncertain system parameters is vital as new technologies proliferate and the total level of system uncertainty grows. Accurate assessment of the importance of different uncertainties can guide power system operators towards parameters which will require the greatest levels of mitigation or increased monitoring in order to reduce the uncertainty and its subsequent impact. Local and global sensitivity analysis techniques are described and evaluated within this paper, including non-parametric methods, variance-based approaches, and distribution-based techniques. The techniques are illustrated using a large 295-bus realistic network model of a generic distribution system. Numerical experiments on dynamic models are used in order to assess the impact of uncertainties on the mitigation of system frequency excursions using single-site and distributed energy storage devices.

Index Terms—Energy storage, frequency stability, importance, probabilistic analysis, sensitivity analysis, uncertainty.

I. INTRODUCTION

POWER systems are becoming increasingly uncertain and complex as new technologies proliferate. These uncertainties may affect many different aspects of power system modelling, analysis, and operation and their impacts and importance must be carefully analyzed and quantified. This rise in uncertainty means that traditional deterministic approaches towards stability and security assessment no longer adequately represent the true performance of the system. Also, the complex and evolutionary nature of networks means it is not necessarily possible to predetermine which conditions will lead to worst-case scenarios and ensure system adequacy for these points. Probabilistic studies provide a means to account for this uncertain variation, and accurately quantify the effects that emerging technologies will have on system stability.

The importance of probabilistic approaches towards power system stability analysis has been highlighted in previous research [1]–[3]. However, it may not be feasible (or even possible) to accurately model all of the uncertainties that exist within a power system. There are many potential sources of uncertainty within power systems including (but not limited to) natural temporal variations, forecast errors, monitoring er-

rors, and physical system parameter estimation. There are two main forms of uncertainty: aleatory and epistemic. The aleatory uncertainty, also known as irreducible uncertainty and variability, represents the inherited random behavior of power systems [4]. The epistemic uncertainty, also called reducible uncertainty and state of knowledge uncertainty, models the uncertainty in parameters estimation due to data shortages or model simplifications [4]. Each form of uncertainty has its representation and quantification methods. For some of these uncertainties, it may be possible to produce an accurate model for the uncertainty based on historical values or data tolerance values from manufacturers. However, for some parameters, the level of uncertainty may be unquantifiable without additional monitoring, measurement, or analytical effort. More importantly, it may not even be necessary to accurately model all uncertainties as many may have little impact on the system phenomena of interest, despite adding considerable computational burden.

The assessment of uncertainty importance using sensitivity analysis (SA) provides information about which inputs will have the greatest effect on an observed system output. With respect to power system uncertainties, this may reveal which uncontrolled variables require mitigation devices (for example, the installation of energy storage to provide some control over intermittent renewable energy sources). Alternatively, the importance information may highlight the need for improved monitoring of parameters, or modelling of system variations (for example, through increased estimation accuracy for forecasts). It is also possible to assess how the importance of different parameters varies as the level of uncertainty grows or reduces which may highlight particular parameters which critically affect system performance under certain conditions.

Sensitivity analysis within power systems has often focused on linear system approaches applied to oscillation damping controller development such as [5]. These linear SA approaches (occasionally utilizing eigenvalue analysis) have also been applied to voltage stability problems in [6] and also in [7] and [8] where extension to quadratic sensitivities is also used to rank contingency importance. A thorough analysis of the use of trajectory sensitivities to assess the influence of system uncertainties on transient behavior is presented in [9]. In [9], the focus is placed on efficiently estimating the bounds of possible system performance from a known point with given levels of uncertainty using first order approximations. This work, [5]–[9], demonstrates the possible benefits of sensitivity analysis, although it is primarily limited to linear approximations around local conditions. As power systems become increas-

This work was supported by the Research Councils UK, through the Hub-Net consortium (grant number: EP/I013636/1).

R. Preece and J.V. Milanović are with the School of Electrical and Electronic Engineering, The University of Manchester, PO Box 88, Manchester, M60 1QD, UK. (email: robin.preece@manchester.ac.uk, milanovic@manchester.ac.uk).

ingly complex, more *global* SA techniques are required in order to fully assess the impact of uncertainties across the full range of possible uncertain values. Furthermore, the increasing penetration of converter-interfaced generation and loads and complex control architectures requires an SA approach which does not assume linear system behavior and is able to include the nonlinearities inherent within modern power systems.

This paper describes a variety of different uncertainty importance indices which can be used to assess the sensitivity of a power system output to uncertain inputs. This work presents the first comparative analysis of this wide variety of SA techniques to power system applications. This research is of critical importance as power systems are increasingly being analyzed using probabilistic methods. This novel application of these SA approaches to determine the importance of uncertainties within power networks enables the development of further probabilistic system analysis. SA has multiple potential power systems applications including highlighting sources of uncertainty which can dominate system performance and require physical mitigation solutions, or identifying where knowledge gaps or errors are resulting in overly conservative operating practices.

The techniques applied within this work are illustrated using examples of differing complexity to demonstrate the applicability of the SA methods under varying analysis problem formulations and constraints. A complex 295 bus realistic network model of a generic distribution system is utilized including large-scale distributed energy storage devices. The importance of uncertainties is assessed with respect to frequency excursions within this work. It must be noted, however, that this represents just one of numerous potential applications including (but not limited to) further stability assessment, state estimation, probabilistic security constrained optimal power flow (PSCOPF), or reliability assessment. A further area of application includes probabilistic methods for studying the cascading mechanism of blackouts. The proposed global sensitivity analysis methods could potentially be incorporated with techniques such as the CASCADE model [10], ORNL-PERSC-Alaska (OPA) model [11], and the Manchester model [12] in order to establish critical uncertainties affecting cascading failures, though this will require further research and investigation.

II. MEASURING UNCERTAINTY IMPORTANCE

The importance of uncertain variables in terms of their impact on system performance can be assessed in different ways. This sensitivity analysis aims to determine the level of influence that an uncertain input variable X_i (or set of uncertain input variables) will have on an observed system output Y . In the following descriptions, p refers to the number of uncertainties, and n refers to the number of simulations used to assess each uncertainty.

A. Non-parametric techniques

The simplest global methods than can be used to assess uncertainty importance are based on input-output correlation and are referred to as non-parametric (as they do not compare dis-

tribution parameters). The approaches described within this section are applicable for systems with either independent or dependent random variables. It has been noted [13], [14] that the ability of these non-parametric techniques to assess uncertainty importance (and subsequently rank variables based on their impact on the system output) is described by the coefficient of determination R_y^2 associated with a linear regression of Y with X . This statistic describes the proportion of variance in Y that is explained by the regression model in X , with the remainder explained by unrecorded parameters or phenomena. Broadly, if R_y^2 is high, we can expect accurate uncertainty ranking based on importance.

1) Pearson Correlation Coefficient – ρ_{XY}

The Pearson correlation coefficient describes the *linear* dependence between X and Y . It is the most commonly used measure of correlation and is often referred to simply as the *correlation coefficient*. The Pearson correlation coefficient ρ_{XY} is calculated using (1), in which $cov(\bullet)$ is the covariance between variables, and σ is the parameter standard deviation.

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \quad (1)$$

2) Spearman Rank Correlation Coefficient – ρ_{X_r, Y_r}

The Spearman rank correlation coefficient is the Pearson correlation coefficient between the ranked versions of the variables X and Y . First, X and Y are translated into ranks (X_r and Y_r) and then ρ_{X_r, Y_r} is determined as above. This produces a value of one if the variables are monotonically related (regardless of the linear nature of their relationship) and is less affected by outliers than the Pearson correlation coefficient.

3) Partial Correlation Coefficient – $\rho_{XY.Z}$

The partial correlation coefficient removes the effects of a set of *control* variables (Z) to describe the degree of association between X and Y . It is determined as the correlation between the residuals R_X and R_Y resulting from the linear regression of X with Z and Y with Z respectively. Partial correlation can be calculated using Pearson or Spearman coefficients as above. Values are determined for individual variables X_i by forming Z as a set consisting of all other uncertainties variables ($Z = X_{-i}$).

B. Variance-based Techniques

The prerequisite of a highly linear model imposed by the non-parametric approaches may not always be achieved in practice. It is therefore necessary to define importance indicators which are independent of model linearity. The two methods described below use the variance of the observations Y as the basis for assessing uncertainty importance.

It should be noted that these measures can be determined for a set of variables as well as single parameters. This is not possible with the previously discussed non-parametric techniques and represents a methodological advantage.

1) First Order Effects – SI

The first method calculates the first order, or main, effects (SI) according to [14] and is given by (2). This measure is also

described in analogous forms in [15] and [16].

This first order effect describes the percentage reduction in variance that is achieved if an uncertain parameter X_i is precisely specified (i.e. no longer a random variable).

$$SI_i = 1 - \frac{E[V(Y|X_i)]}{V[Y]} \quad (2)$$

In (2), $V[Y]$ denotes the variance of Y with all uncertain parameters, and $E[V(Y|X_i)]$ denotes the expectation of the conditional variance of Y given X_i has a fixed value. This conditional expected variance is taken over all X_j , $j \neq i$, weighted by the density of X_i .

2) Total Effects – ST

The first order effects describe the partial variance associated directly with one uncertain variable. There also exist *interaction terms* which describe the combined effects of multiple parameters. For a system with independent inputs, the total output variance can be fully decomposed into a sum of terms as in (3) where the first order effects described above represent the first term S_i . The higher order terms are also presented which represent the interactions between multiple uncertainties. In (3), S_{ij} is the 2nd order interaction term relating to any two uncertainties i and j , where $j \neq i$. Similarly, S_{ijk} is the 3rd order interaction term relating to any three uncertainties i , j , and k , where $k \neq j \neq i$. These interactions will exist up to the p^{th} order with p uncertainties.

$$1 = \sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{k>j} S_{ijk} + \dots \quad (3)$$

For a purely linear model, the interactions – all terms in (3) except the first – would all equal zero and so the total observed variance would equal the sum of the first order effects. Another measure, the total effects ST_i [14] is defined as the sum of all terms in (3) which contain the subscript i . This measure describes the percentage of variance that remains if all parameters except X_i are specified and only X_i is a random variable. It is also described by (4), in which X_{-i} represents the vector of all X_j where $j \neq i$, (i.e. all parameters except X_i). The similarity with SI_i defined in (2) is very apparent. This is exploited in order to calculate both SI_i and ST_i without the need for determining all the interaction terms in (3).

$$ST_i = \frac{E[V(Y|X_{-i})]}{V[Y]} \quad (4)$$

3) Determining Variance-based Sensitivity Indices

Practically, the first order and total effects can be determined through quasi Monte Carlo (MC) based sampling of the model. An important result from [17] is that the total effects can be determined without the need to determine all the interaction terms (which total $2^p - 1$ and quickly become unmanageable as the number of uncertainties p increases).

We can define a new complimentary set C_{-i} as the sum of all terms in (3) that do not contain the subscript i . From this

definition, it can be seen that (5) is true, as ST_i and C_{-i} will, between them, include every sensitivity term.

$$1 = ST_i + C_{-i} \quad (5)$$

The set C_i can be similarly defined as equal to S_i , and therefore equal to the first order effects SI_i .

When running n iterations of the model during the MC process, the n -dimensional vector of outputs Y is produced. More specifically, Y can be described as a function of the uncertainties, $Y(X_1, X_2, \dots, X_i, \dots, X_p)$, in which $X_1, X_2, \dots, X_i, \dots, X_p$ are n -dimensional vectors containing the random values of each uncertainty. A new output vector Y'_i can be defined as the following function: $Y'_i(X'_1, X'_2, \dots, X_i, \dots, X'_p)$, in which the n -dimensional vector for X_i remains the same, and all other uncertainties are resampled from their distributions. Similarly, we can define Y_{-i}' as $Y_{-i}'(X_1, X_2, \dots, X_i, \dots, X_p)$ in which only X_i is resampled and all other parameters (X_{-i}) are remain the same. In these definitions, resampling is denoted by the prime symbol (').

It has been shown in [17], [18] that as n approaches infinity, (6) and (7) are true. Therefore, C_i and C_{-i} can be approximated as the Pearson correlation coefficients between output vectors Y from complementary model runs (as defined above).

$$\lim_{n \rightarrow \infty} C_i = \rho(Y, Y'_i) \quad (6)$$

$$\lim_{n \rightarrow \infty} C_{-i} = \rho(Y, Y_{-i}') \quad (7)$$

By grouping inputs into vectors of X_i and X_{-i} (where i can be single variable, or a set of inputs), this asymptotic result advocates the following assumptions for large n [17]–[19]:

- For any input group i , the first order effect SI_i is equal to the correlation coefficient (C_i) of the output vectors from two model runs in which all values for variables in X_i are common, but all other inputs use independent samples.
- For any input group i , the total effect ST_i is equal to one minus the correlation coefficient ($1 - C_{-i}$) of the output vectors from two model runs in which all values for X_i are independently sampled, but all other inputs are common.

It is therefore possible, by systematically isolating different input sets one at a time, to determine both the first order and total effects for any input, or set of inputs. This will require a total of $(p+1)n$ simulations in order to establish either the first order, or total, effects for p input sets.

An established alternative is to use Fourier amplitude sensitivity testing (FAST) [20]. It has been shown in [21] that this approach provides sensitivity indices equivalent to the first order effects described by (2). The FAST approach is not described in detail within this paper as initial results were inaccurate and the method was not further explored. However the basic approach is to vary different uncertain parameters at

different frequencies and to therefore encode the parameter identity in its variation frequency [22]. Subsequent Fourier analysis provides the strength of the different parameter variation frequencies in the observed output. Therefore, the propagation of variations at a given frequency provides a measure of the sensitivity of the output to that parameter – and therefore the parameters importance. In this way, the multidimensional problem is transformed to just a single dimension.

It should be noted that FAST application for systems with large numbers of uncertainties can become computationally intensive due to the determination of a set of p integer frequencies (used during the Fourier transformation). These frequencies must be incommensurable to a given order to reduce the error introduced by the FAST approximation. Also, although just one study is required, the number of simulations n required within the study is dependent on the maximum frequency of parameter variation used which increases quickly and non-linearly as p increases [20].

C. Distribution-based Techniques

It has been noted previously, in [14] and [13], that focusing on variance as the only uncertainty importance measure is equivalent to assuming that the variance is sufficient to describe the observed output variability. Examples of different distributions which display the same variance can easily be produced and an ideal measure of uncertainty importance should account for changes to the *entire distribution* – not just one descriptor of the distribution.

1) Borgonovo – δ

The Borgonovo δ index developed in [23] assesses the impact that a reduction in uncertainty has on the shape of the entire distribution. It achieves this by fixing an uncertain parameter X_i at a given value x_i^* and varying the remaining inputs X_{-i} in order to produce a new distribution for the observed output ($Y | X_i = x_i^*$). This distribution can be compared to the original observed distribution of Y (with all parameters uncertain) and the area difference between the two – $s(X_i)$ – can be established. With reference to Fig. 1, $s(X_i)$ is defined as (8), and the importance index δ is defined as (9). In (9) the expectation of $s(X_i)$ is determined by integrating across the full range of X_i , taking into account its distribution, as in (10) [13], [23].

$$s(X_i) = \int |f_Y(y) - f_{Y|X_i=x_i^*}(y)| dy \quad (8)$$

$$\delta = \frac{1}{2} E[s(X_i)] \quad (9)$$

$$E[s(X_i)] = \int f_{X_i}(x_i) \left[\int |f_Y(y) - f_{Y|X_i=x_i^*}(y)| dy \right] dx_i \quad (10)$$

The δ index is limited to the range of $0 \leq \delta \leq 1$, with zero indicating identical density functions and one indicating non-overlapping density functions. Computationally, it requires an MC-based integration across the range of X_i , using N studies for this *outer integration* loop. It can also be computed for sets of inputs, though this will require an additional outer integration loop across all parameters in the set X_i . The total

computational cost is therefore much higher for this index than the others discussed above. This is the price to be paid for a sensitivity measure which accounts for the whole distribution and not just a moment and requires $p n N$ simulations to determine δ indices for all parameters (or for grouped input sets that include all parameters).

For many power systems applications, model evaluation can be computationally intensive, particularly if dynamic studies are required. Therefore, even with small MC simulation numbers for n and N (of less than 1000), this approach will not be suitable due to the computational expense. Some approaches that use efficient sampling, such as Latin hypercube [24] or quasi-random LP τ [25], [26], can be used to reduce n and N . However the dimensionality curse is still a major limitation as p increases due to the need for multiple integration loops.

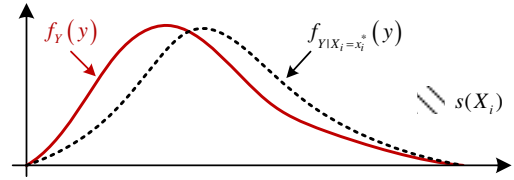


Fig. 1: Example distributions to demonstrate the basis for the derivation of the δ uncertainty importance index.

D. One Factor at a Time – S_i^{loc}

The above methods involve the global random variation of uncertainties across all possible values to produce multiple observations from which correlation, variance measurements, and distributions can be established. It is far more common to find reference to SA which utilizes a local one-at-a-time (OAT) parameter-by-parameter approach. Inferences about the importance of uncertain parameters are made by observing the changes in Y that occur for small changes in X_i (whilst keeping X_{-i} constant at the nominal values \mathbf{x}_{-i}^0).

Such an approach aims to empirically determine the partial derivatives of Y with respect to X_i . This provides the first term in the OAT sensitivity measure S_i^{loc} in (11). The second term in (11) simply weights this local sensitivity by the expected level of variation expected in X_i . This local sensitivity index is often normalized with respect to the largest sensitivity value.

$$S_i^{loc} = \frac{\partial Y}{\partial X_i} \frac{\sigma_{X_i}}{E[X_i]} \quad (11)$$

SA using measures analogous to (11) is typical in power systems analysis and also further afield in engineering, physics and chemistry [14]. However, there is an obvious shortcoming of this approach compared to the previously described global methods – that the sensitivity is only assessed in the local neighborhood of \mathbf{x}^0 and not the full range of uncertainty values. This shortcoming is countered by the significantly reduced computational cost of just $p + 1$ simulations.

E. Summary of Importance Indices

The uncertainty importance indices described are summarized in Table I. This table includes details of whether or not the methods can handle grouped inputs, and also the computational cost (in terms of total simulation numbers required).

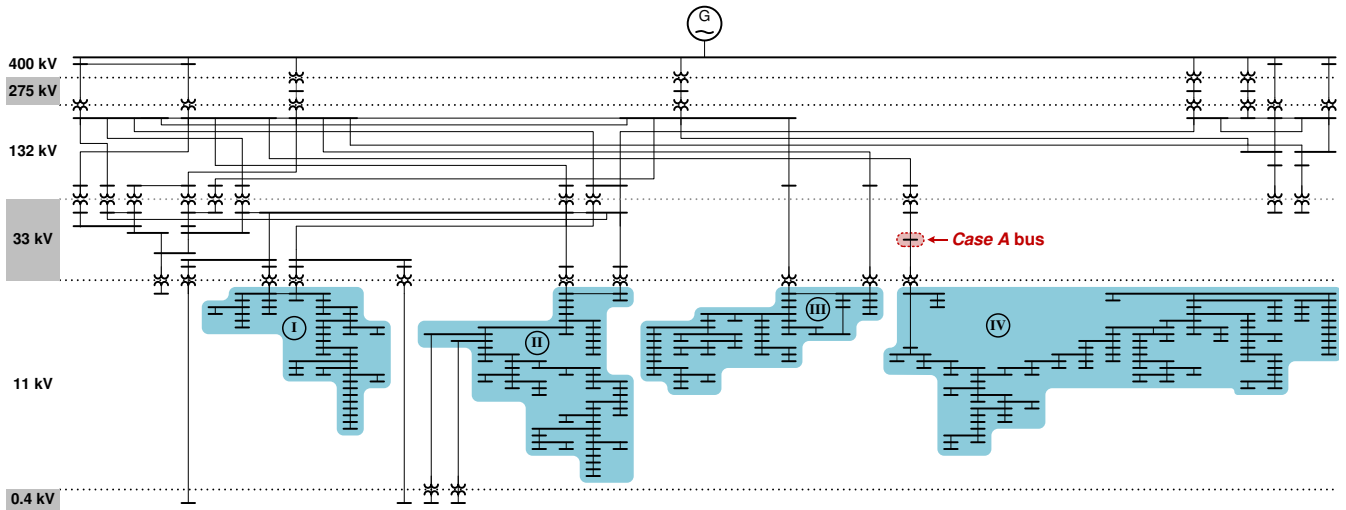


Fig. 2: Network diagram of the 295-bus generic distribution systems (GDS).

TABLE I
SUMMARY OF UNCERTAINTY IMPORTANCE INDICES

Scope	Type	Method	Symbol	Groups	Cost
Local	One At a Time	Weighted Sensitivity	S_i^{loc}	No	$(p+1)$
Global	Non-parametric	Pearson	ρ_{XY}	No	n
		Spearman	ρ_{X,Y_r}	No	n
		Partial	ρ_{XY-Z}	No	n
Variance	First Order Effects	First Order Effects	SI_i	Yes	$(p+1)n$
		Total Effects	ST	Yes	$(p+1)n$
Distribution	Borgonovo	δ		Yes	$p n N^*$

n : number of MC simulations, p : number of variable sets, N : number of outer integration loops (see Section II.C.1)

*For δ , p is equal to the number of variables (not variable sets) as it is necessary to integrate across all variables even if uncertainties are grouped.

It should be noted that this does not represent an exhaustive set of SA techniques and that additional uncertainty screening methods exist, including Plackett–Burman [27], fractional factorial [28], sum-of-trees [29], and multivariate adaptive regression splines [30]. The evaluation of a wide range of approaches has been completed in [31] using an environmental hydrological model, however applications with power systems are not currently available.

III. TEST SYSTEM

The comparative evaluation of the discussed SA techniques is performed on the 295 bus generic distribution system (GDS), shown in Fig. 2. This network comprises four 275 kV transmission infeeds from a large external equivalent network machine. 132 kV and 33 kV meshed sub-transmission networks feed a radial distribution network which predominantly operates at 11 kV, but also at 0.4 kV. The network consists of 295 buses, 276 lines (overhead and cable), and 37 transformers with various winding connections. Detailed description and parameters of the test network can be found in [32]. Pow-

er system modelling is performed using DiGSILENT PowerFactory and all analysis is conducted using Matlab.

A. Frequency Excursions

Dynamic transient simulations are performed in order to assess the frequency excursions (deviations from 50 Hz) that occur following the sudden connection of a load. This load is located at the 400 kV bus within the test system and is sized at 100 MW in order to produce reasonably large frequency excursions (approximately 1% nominal frequency) that may require support from fast acting storage devices. The minimum frequency f_{min} experienced following the load connection is recorded as the observed indicator of system performance (Y).

Frequency excursions are of interest and so it is expected that the generator inertia and governor model parameters will significantly affect f_{min} . The governor model used in this study is shown in Fig. 3(a). It is assumed for this work that the reduced network equivalent generator model has been determined using a method that may introduce some errors. Therefore, the key parameters associated with the generator and governor model are subject to a degree of uncertainty.

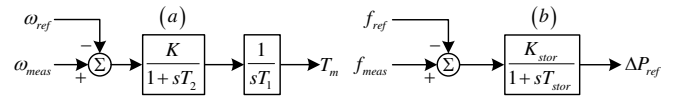


Fig. 3: (a) Network equivalent generator governor model, and (b) simple energy storage frequency support controller.

1) Energy Storage Devices

Energy storage devices can be used in order to support the system frequency during excursions by quickly injecting active power to cover the temporary generation shortfall [33], [34]. A simple proportional controller with a storage system time constant is used in this study shown in Fig. 3(b) which simply injects power when the local monitored frequency deviates from its reference value [34], [35]. This injection of power can only occur if there is sufficient energy stored within the device and the state of charge (SOC) does not fall to zero. The controller parameters of any installed storage devices are

expected to have a significant effect on f_{min} . If a large number of distributed energy storage devices are installed within the network then uncertainty may exist with respect to the controller parameters and SOC of devices.

B. System Operational Scenarios

Two scenarios are considered in order to assess the importance of the considered uncertainties.

1) Case A: Single Storage Unit

A single storage device is installed at the 33 kV bus labelled in Fig. 2. This device is rated at 40 MW, equivalent to a very large battery storage installation [36]. It is assumed that the control settings for this single device will be precisely known and are not subject to any uncertainty. The inertia (H) of the network equivalent generator and also the governor control parameters (K , T_1 , and T_2) are uncertain due to assumed network reduction inaccuracies. Additionally, the level of system loading varies according to the network loading factor α_L . Table II describes the distributions of the *Case A* uncertainties.

Typical transient responses from the equivalent generator and storage device following the load disturbance are shown in Fig. 4 where the minimum frequency f_{min} is labeled.

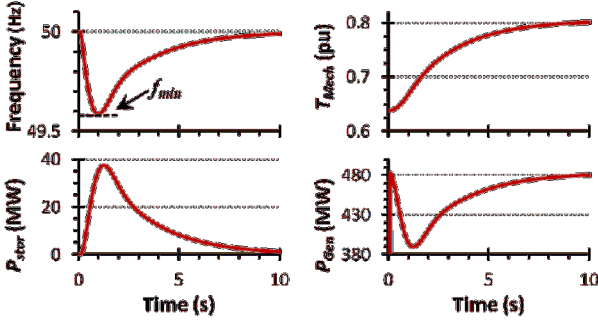


Fig. 4: Typical responses of the system frequency, power output from the storage device (P_{stor}), generator mechanical torque (T_{Mech}), and generator active power (P_{Gen}) following the load disturbance.

2) Case B: Distributed Storage Units

For the second considered scenario, the large single storage device is replaced with 1000 distributed energy storage devices. These identical devices are all rated at 40 kW so that the total installed storage is 40 MW, equivalent to *Case A*. These devices are distributed randomly at the 11 kV system buses with many locations having multiple connected devices.

TABLE II
DISTRIBUTION DETAILS FOR CONSIDERED UNCERTAINTIES

Uncertainty	Distribution	Parameters	Case A	Case B
α_L	Gaussian	$\mu = 0.7$, $\sigma = 0.1$	✓	
H	Gaussian	$\mu = 4$ s, $\sigma = 0.2$ s	✓	✓
K	Uniform	range = [6, 14]	✓	✓
T_1	Uniform	range = [0.8, 1.2] s	✓	✓
T_2	Uniform	range = [0.16, 0.24] s	✓	✓
K_{stor}	Uniform	range = [1, 500]		✓
T_{stor}	Uniform	range = [0.01, 0.50] s		✓
SOC	Uniform	range = [0, 100] %		✓

As with *Case A*, the inertia (H) and the governor control parameters (K , T_1 , and T_2) are modelled as uncertain. Unlike

Case A, variations in loading factor are not considered. However, in addition to the four stated modelling uncertainties, device uncertainties for each of the 1000 distributed energy storage units are also considered. These consist of the storage frequency response parameters based on the controller and device time constant (K_{stor} and T_{stor}). Also considered is the device SOC, which defines whether or not the distributed storage device contains enough energy to support the network frequency during transient excursions. The distributions and associated parameters for these uncertainties are given in Table II. The total number of considered uncertainties for *Case B* is 3004, consisting of the four modelling uncertainties and the three additional uncertainties for 1000 storage units.

IV. RESULTS & DISCUSSION

A. Case A: Single Storage Unit

Considering first *Case A*, with five uncertain parameters ($p = 5$), the number of MC simulations (n) was initially selected as 5000 to ensure system variation was suitably captured. However, when determining the δ importance measure, the need for an inner and outer integration loop meant that this had to be reduced to enable the computation to be completed in reasonable time. For the δ calculation, both n and N (inner and outer integration sample numbers) were set to 100 and so the total number of samples per uncertainty was 10,000 (nN). Note that keeping $n = N = 5000$ would increase the total number of system simulations to 25 million for each uncertainty.

Using 100 samples cannot ensure output variation is suitably captured if pure quasi-random MC sampling is used. To overcome this, more efficient LP τ sampling [25], [26] is used to evenly sample the search space. The obtained weighted results are then used to produce a probability distribution using a kernel smoothing density estimate. It was found that a good estimate can be obtained using very few LP τ samples. Just 100 samples can reproduce an estimate of a Gaussian distribution with an average root mean square error of less than 0.6%. The total time taken to run the simulations for each method is shown in Table III based on computation using an 3.4 GHz Intel Core i7 PC with 8 GB RAM.

TABLE III
TOTAL SIMULATION TIMES FOR CASE A

Methods	S_i^{loc}	$\rho_{XY}, \rho_{X,Y}, \rho_{XY,Z}$	SI_i, ST	δ
Cost	$(p+1)$	n	$(p+1)n$	$p n N$
Simulations	6	5,000	10,000	50,000
Time	7 s	1 h 40 m	8 h 20 m	16 h 40 m

The numerical results produced by the SA methods for *Case A* are shown in Table IV. Also tabulated (in parentheses) are the rankings based on these values. It can be seen that the governor gain K (also labelled X_3) is consistently ranked as the most important uncertainty affecting the maximum frequency deviation. However, the SA methods are not consistent with respect to the second and third ranked uncertainties. It is evident that S_i^{loc} , ρ_{XY} , and $\rho_{XY,Z}$ rank the equivalent generator inertia H (X_2) second most critical, whilst $\rho_{X,Y}$, SI , ST , and δ identify T_1 (X_4) as more important.

TABLE IV
UNCERTAINTY IMPORTANCE VALUES AND RANKING FOR CASE A

Uncertainty	S^{loc}	ρ_{XY}	$\rho_{X,Y}$	ρ_{XY-Z}	SI	ST	δ
$X_1 : \alpha_L$	0 (5)	0.008 (5)	0.011 (5)	0.008 (5)	0.001 (5)	0 (5)	0.047 (5)
$X_2 : H$	0.536 (2)	0.395 (2)	0.367 (3)	0.954 (2)	0.151 (3)	0.191 (3)	0.114 (3)
$X_3 : K$	1.000 (1)	0.800 (1)	0.811 (1)	0.987 (1)	0.651 (1)	0.660 (1)	0.386 (1)
$X_3 : T_1$	0.502 (3)	0.388 (3)	0.375 (2)	0.950 (3)	0.152 (2)	0.232 (2)	0.128 (2)
$X_5 : T_2$	0.208 (4)	0.160 (4)	0.152 (4)	0.776 (4)	0.021 (4)	0.026 (4)	0.071 (4)

These results are displayed graphically in Fig. 5 in which the area of the bubble is proportional to the uncertainty importance value (normalized for each SA method). This visual representation highlights three consistent groups: (i) S^{loc} , ρ_{XY} , and $\rho_{X,Y}$, (ii) SI , ST and δ , and (iii) ρ_{XY-Z} . It is particularly evident that ρ_{XY-Z} produces much higher importance values for many parameters.

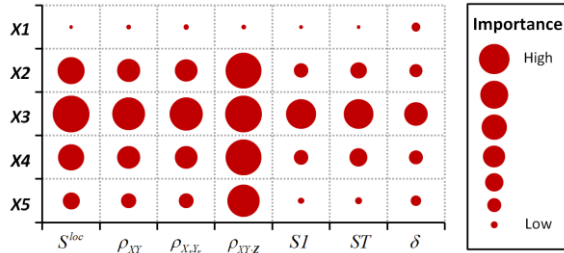


Fig. 5: Case A parameter importance by different SA methods.

With respect to the relative importance of different parameters (i.e. the size of the bubbles in Fig. 5), it can be seen that there is some difference between the different SA methods. Assuming that the δ measure provides the best measure of importance as it is determined based on differences between the entire output distributions, it is evident that the SI and ST indices produce uncertainty importance values which are closest to the δ measure. This is a valuable result, as δ requires large sample numbers as the number of uncertainties (p) increases (see Table I), and these alternative sensitivity measures may provide a viable alternative for practical implementations.

It should be acknowledged that the non-parametric correlation based statistics are still able to provide accurate rankings for the importance of the uncertainties. This can be attributed to the good fit of the linear regression model with an R^2 statistic equal to 0.984. The local SA approach is able to produce very similar results and almost identical ranking despite using just a fraction of the number of sample points. However, whilst this may be true in this instance, it may not be the case for all power systems problems – particularly if the system contains many nonlinearities and discontinuities. It also over estimates the importance of some uncertainties with respect to the δ measure. A key area for further research is the identification of when the use of a local SA approach is, and more importantly is not, appropriate. It is crucial to be able to identify (without the need for full GSA) cases when local SA may

suffice without introducing considerable error.

1) Effect of Sample Size

The effect of the sample size has been investigated to assess whether the high number of simulations ($n = 5000$) is required in order to produce accurate sensitivity assessment. Fig. 6 displays the parameter importance determined using ρ_{XY} and using ST with varying sample sizes. It is evident that the ST measure is slightly more consistent (with bubble sizes varying little). The correlation ρ_{XY} over-estimates the importance of X_1 until the number of samples increases significantly. However, both measures see variations in the ranked order of parameters with X_2 and X_4 swapping second and third place as the number of samples increases. In both cases, sample sizes over 1000 are required to produce the original ranked parameter previously seen.

B. Case B: Distributed Storage Units

For Case B, with 3004 uncertain parameters, some of the approaches used with Case A are unsuitable. Unfortunately the Borgonovo δ measure cannot be determined due to the excessive number of samples (and therefore dynamic power system simulations) required. Even using efficient LP τ sampling with $n = N = 100$ would require more than 30 million samples (which is unrealistic for practical applications). This represents the clearest shortcoming of the δ measure and is acknowledged in [23] as a major limitation.

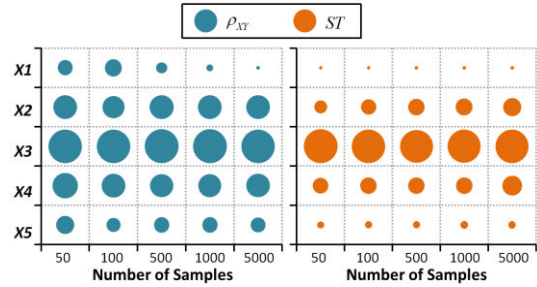


Fig. 6: Parameter importance using ρ_{XY} and ST with varying sample size.

With a large number of uncertainties, it is desirable to perform SA on grouped sets of inputs rather than individual parameters, when assessing importance. Not only does this reduce the number of simulations (for methods excluding Borgonovo), but it also helps to provide more meaningful results. It helps to highlight where effort should be optimally placed. For example, mitigating the issues caused by a group of devices in a particular geographical location, or highlighting the need for more information about a specific group of controller parameters. Of the methods discussed within this paper, SI , ST and δ can be used with grouped uncertainty sets, however δ cannot be implemented due to its aforementioned computational burden. Therefore, SI and ST have been determined for Case B with 1000 distributed storage devices.

1) Grouping by Uncertainty Type

The uncertainties have been grouped into the following sets (note that any desired grouping selection can be made):

- Z_1 : Equivalent generator parameters consisting of the four parameters modelled in Case A (H , K , T_1 , and T_2).

- Z_2 : Storage dynamic response parameters consisting of K_{stor} and T_{stor} for all 1000 distributed storage units.
- Z_3 : Storage SOC parameters for all 1000 storage units.

The results from the completed SA for these groups for *Case B* using the *SI* and *ST* methods is tabulated in Table VI and shown visually as Fig. 7. It is extremely clear that the system frequency response is dominated by the uncertainty of the equivalent generator parameters (Z_1) to such an extent that the dynamic response parameters and SOC are largely irrelevant. It should be noted that the combined installed storage (40 MW) represents approximately 11% of the total system load (350 MW) and that this may be a possible cause for the reduced influence of the storage control parameters on the frequency response. However, as is demonstrated by the following example, the importance of uncertainties is not always immediately obvious.

2) Grouping by Uncertainty Location

An additional grouping is also tested, but uncertainties are categorized based on their geographical location within the network, rather than the uncertainty type. In this study, the four equivalent generator parameters (Z_1 in the previous example) are fixed to more thoroughly investigate the relative importance of the distributed storage devices. The three storage device parameters (K_{stor} , T_{stor} , and *SOC*) for each of the 1000 distributed storage devices are still considered uncertain with the same probability distributions previously described. The geographical groupings are performed according to the areas shown in Fig. 2 with Table V providing the details of the randomly positioned devices within areas I–IV.

The results from the completed SA for these location-based groups for *Case B* using the *ST* method is tabulated in Table VII and shown visually as Fig. 8. In the above studies it has been evident that the *SI* and *ST* indices give almost identical results for these power system examples and so only the *total effects ST* have been calculated for this study to reduce the computational burden.

TABLE V
DETAILS OF GEOGRAPHICAL GROUPING FOR CASE B

Number of	Z_I	Z_{II}	Z_{III}	Z_{IV}
Buses	40	54	43	95
Storage Devices	188	230	174	408

TABLE VI
CASE B UNCERTAINTY
IMPORTANCE (TYPE GROUPS)

Uncertainty Group	<i>SI</i>	<i>ST</i>
Z_1	0.970 (1)	0.971 (1)
Z_2	0.030 (2)	0.030 (2)
Z_3	0.002 (3)	0.002 (3)

TABLE VII
CASE B UNCERTAINTY
IMPORTANCE (LOCATION GROUPS)

Uncertainty Group	<i>ST</i>
Z_I	0.149 (3)
Z_{II}	0.156 (2)
Z_{III}	0.312 (1)
Z_{IV}	0.061 (4)

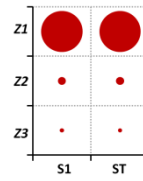


Fig. 7: Case B importance (type groups).

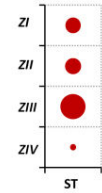


Fig. 8: Case B importance (location groups).

The results from this geographical grouping display some potential unexpected results with respect to the impact of different storage devices on the frequency response of the network. It can be seen that Z_{III} is the most important group despite having the smallest number of storage devices. Additionally Z_{IV} , which has the largest number of connected storage devices, has the smallest impact on system frequency excursions. The remaining groups (Z_I and Z_{II}) have similar importance despite Z_{II} having over 20% more storage devices.

This result emphasizes the necessity of uncertainty sensitivity analysis as it is not at all true that the area with the greatest number of devices is the most important. The nonlinearities in power systems ensure that parameters must be thoroughly investigated to establish their relative importance. In a practical system, such a result could be used to highlight where the greatest effort should be directed in order to mitigate for frequency excursions.

V. CONCLUSIONS

This paper has outlined, compared, and critically evaluated a number of uncertainty importance methods. These methods have been assessed for the suitability to power systems applications and investigated using a large, realistic power system model. This study represents a novel application of global sensitivity analysis (SA) techniques to power systems problems and is of significant value as there is an increasing move towards probabilistic analysis methods.

The significant limitations of the Borgonovo δ measure have been discussed. Despite the use of efficient sampling methods to reduce the computational burden, the index remains too intensive to use in all practical cases except those with very low numbers of uncertainties. The (traditionally used) local sensitivity measure displayed consistent results with global methods, however it is unable to capture complex sensitivities, and it cannot be used with grouped inputs.

The *SI* and *ST* indices produce results which have been shown to be similar to the (impractical) δ measure but can be used with grouped inputs and calculated with reasonable sampling requirements. For many power system applications, these methods will provide the best balance between thorough global SA and practical implementation.

The illustrative examples used in this study, particularly the location-based grouping, have revealed the value of uncertainty importance assessment. This has highlighted that uncertainty analysis could be used to direct increased monitoring and more fully understand the impact of uncertainties. It is also easy to envisage situations where the critical uncertainty associated with a renewable energy source can be identified

and subsequently mitigated for through the installation of energy storage devices. Furthermore, it must be stressed that this represents just one application of uncertainty importance assessment within power system dynamics. The indices described can be applied to any area where probabilistic analysis is used (including steady state analysis such as probabilistic power flow) in order to target mitigation or identify knowledge gaps which presently result in overly conservative system operation.

REFERENCES

- [1] J. L. Rueda, D. G. Colome, and I. Erlich, "Assessment and Enhancement of Small Signal Stability Considering Uncertainties," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 198–207, 2009.
- [2] R. Arrieta, M. A. Rios, and A. Torres, "Contingency Analysis and Risk Assessment of Small Signal Instability," *Power Tech, 2007 IEEE Lausanne*. pp. 1741–1746, 2007.
- [3] J. Rueda and I. Erlich, "Probabilistic framework for risk analysis of power system small-signal stability," *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.*, vol. 226, no. 1, pp. 118–133, 2012.
- [4] L. Swiler and A. Giunta, "Aleatory and epistemic uncertainty quantification for engineering applications," in *Proceedings of the Joint Statistical Meetings, 2007*.
- [5] C. Y. Chung, K. W. Wang, C. T. Tse, and R. Niu, "Power-system stabilizer (PSS) design by probabilistic sensitivity indexes (PSIs)," *IEEE Trans Power Syst.*, vol. 17, no. 3, pp. 688–693, 2002.
- [6] N. Amjady and M. Esmaili, "Application of a new sensitivity analysis framework for voltage contingency ranking," *IEEE Trans Power Syst.*, vol. 20, no. 2, pp. 973–983, 2005.
- [7] S. Greene, I. Dobson, and F. L. Alvarado, "Contingency ranking for voltage collapse via sensitivities from a single nose curve," *IEEE Trans Power Syst.*, vol. 14, no. 1, pp. 232–240, 1999.
- [8] S. Greene, I. Dobson, and F. L. Alvarado, "Sensitivity of the loading margin to voltage collapse with respect to arbitrary parameters," *IEEE Trans Power Syst.*, vol. 12, no. 1, pp. 262–272, 1997.
- [9] I. A. Hiskens and J. Alseddiqui, "Sensitivity, Approximation, and Uncertainty in Power System Dynamic Simulation," *IEEE Trans Power Syst.*, vol. 21, no. 4, pp. 1808–1820, 2006.
- [10] I. Dobson, B. A. Carreras, and D. E. Newman, "A probabilistic loading-dependent model of cascading failure and possible implications for blackouts," in *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the, 2003*, p. 10 pp.
- [11] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos*, vol. 12, no. 4, pp. 985–994, Dec. 2002.
- [12] D. S. Kirschen, K. R. W. Bell, D. P. Nedic, D. Jayaweera, and R. N. Allan, "Computing the value of security," *IEE Proc. - Gener. Transm. Distrib.*, vol. 150, no. 6, p. 673, 2003.
- [13] E. Borgonovo, "Measuring uncertainty importance: investigation and comparison of alternative approaches," *Risk Anal.*, vol. 26, no. 5, pp. 1349–61, Oct. 2006.
- [14] A. Saltelli, S. Tarantola, and F. Campolongo, "Sensitivity Analysis as an Ingredient of Modeling," *Stat. Sci.*, vol. 15, no. 4, pp. 377–395, Nov. 2000.
- [15] R. L. Iman, "A Matrix-Based Approach to Uncertainty and Sensitivity Analysis for Fault Trees," *Risk Anal.*, vol. 7, no. 1, pp. 21–33, Mar. 1987.
- [16] T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of nonlinear models," *Reliab. Eng. Syst. Saf.*, vol. 52, no. 1, pp. 1–17, Apr. 1996.
- [17] I. M. Sobol', "Sensitivity estimates for nonlinear mathematical models (English Translation)," *Math. Model. Comput. Exp.*, vol. 1, no. 4, pp. 407–414, 1993.
- [18] I. M. Sobol' and Y. L. Levitan, "On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index," *Comput. Phys. Commun.*, vol. 117, no. 1–2, pp. 52–61, Mar. 1999.
- [19] G. Glen and K. Isaacs, "Estimating Sobol sensitivity indices using correlations," *Environ. Model. Softw.*, vol. 37, pp. 157–166, Nov. 2012.
- [20] G. J. McRae, J. W. Tilden, and J. H. Seinfeld, "Global sensitivity analysis—a computational implementation of the Fourier Amplitude Sensitivity Test (FAST)," *Comput. Chem. Eng.*, vol. 6, no. 1, pp. 15–25, Jan. 1982.
- [21] A. Saltelli, S. Tarantola, and K. P.-S. Chan, "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, vol. 41, no. 1, pp. 39–56, Feb. 1999.
- [22] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, "A methodology for performing global uncertainty and sensitivity analysis in systems biology," *J. Theor. Biol.*, vol. 254, no. 1, pp. 178–96, Sep. 2008.
- [23] E. Borgonovo, "A new uncertainty importance measure," *Reliab. Eng. Syst. Saf.*, vol. 92, no. 6, pp. 771–784, Jun. 2007.
- [24] M. McKay, R. Beckman, and W. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, 2000.
- [25] I. M. Sobol', "Quasi-Monte Carlo methods," *Prog. Nucl. Energy*, vol. 24, no. 1–3, pp. 55–61, Jan. 1990.
- [26] V. A. Maslennikov, J. V. Milanovic, and S. M. Ustinov, "Robust ranking of loads by using sensitivity factors and limited number of points from a hyperspace of uncertain parameters," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 565–570, Aug. 2002.
- [27] R. Plackett and J. Burman, "The design of optimum multifactorial experiments," *Biometrika*, 1946.
- [28] G. E. P. Box and J. S. Hunter, "The 2^k — p Fractional Factorial Designs," *Technometrics*, vol. 3, no. 3, pp. 311–351, Aug. 1961.
- [29] L. Breiman, J. Friedman, C. Stone, and R. Olshen, "Classification and regression trees," 1984.
- [30] H. Pi and C. Peterson, "Finding the embedding dimension and variable dependencies in time series," *Neural Comput.*, 1994.
- [31] Y. Gan, Q. Duan, W. Gong, C. Tong, Y. Sun, W. Chu, A. Ye, C. Miao, and Z. Di, "A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model," *Environ. Model. Softw.*, vol. 51, pp. 269–285, Jan. 2014.
- [32] Y. Zhang, "Techno-economic assessment of voltage sag performance and mitigation," The University of Manchester, UK, 2008.
- [33] G. Delille, B. Francois, and G. Malarange, "Dynamic Frequency Control Support by Energy Storage to Reduce the Impact of Wind and Solar Generation on Isolated Power System's Inertia," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 931–939, Oct. 2012.
- [34] J. Mongkoltanatas, D. Riu, and X. LePivert, "Energy storage design for primary frequency control for islanding micro grid," in *38th Annual Conference on IEEE Industrial Electronics Society (IECON)*, 2012, pp. 5643–5649.
- [35] Chun-Feng Lu, Chun-Chang Liu, and Chi-Jui Wu, "Effect of battery energy storage system on load frequency control considering governor deadband and generation rate constraint," *IEEE Trans. Energy Convers.*, vol. 10, no. 3, pp. 555–561, 1995.
- [36] ABB, "World's Largest Battery Energy Storage System - Fairbanks, Alaska, USA." [online] Available at: www.abb.com, 2011.

Robin Preece (GS'10, M'13) received his BEng degree in Electrical and Electronic Engineering in 2009 and his PhD degree in 2013, both from the University of Manchester, UK. He is currently a Lecturer in Power Systems Engineering at the same institution.

Jovica V. Milanović (M'95, SM'98, F'10) received his Dipl.Ing and his MSc degrees from the University of Belgrade, Yugoslavia, his PhD degree from the University of Newcastle, Australia, and his Higher Doctorate (DSc degree) from the University of Manchester, UK, all in Electrical Engineering. Currently, he is a Professor of electrical power engineering, Deputy Head of the School of Electrical and Electronic Engineering and Head of Electrical Energy and Power Systems Group at The University of Manchester, UK, Visiting Professor at the University of Novi Sad, Novi Sad, Serbia and Conjoint Professor at the University of Newcastle, Newcastle, Australia.