Joined Probabilistic Load Flow and Sensitivity Analysis of Distribution Networks based on Polynomial Chaos Method

Giambattista Gruosso, Roberto S. Netto, Zheng Zhang, Luca Daniel, and Paolo Maffezzoni

***Abstract*— Due to the statistical uncertainty of loads and power sources found in smart grids, effective computational tools for probabilistic load flow analysis and planning are now becoming indispensable. In this research, we describe a unified simulation framework that allows quantifying the probability distributions of a set of observation variables as well as evaluating their sensitivity to potential variations in the power demands. The proposed probabilistic technique relies on the generalized Polynomial Chaos algorithm and on a region-wise aggregation/description of the time-varying load profiles. It is shown how detailed statistical distributions of some important figures of merit, which includes voltage unbalance factor in distribution networks, can be calculating with a $\approx 100\times$ acceleration compared to standard Monte Carlo analysis. In addition, it is highlighted how the associated sensitivity analysis is of guidance for the optimal allocation and planning of new loads.**

***Index Terms*— Polynomial Chaos method, Probabilistic load flow, Sensitivity analysis, Unbalanced networks, Uncertainty quantification.**

I. Introduction

Smart distribution grids are expected to provide new types of services, e.g. charging of electrical vehicles, while exploiting new forms of distributed power generation, e.g. higher penetration of renewable energy sources. An increasing active role of consumers is also envisaged: they will be able to alter their usage patterns in order to follow the trends of electricity prices [1]. Such novel features imply larger variations and statistical uncertainty in power demands and generations thus increasing the relevance of efficient/robust simulation tools for probabilistic analysis and planning [2], [3]. Probabilistic Load Flow (PLF) analysis methods consist in using probabilistic models for the power loads as well as in replacing deterministic load flow simulation with proper stochastic analysis methods. The basic and reference stochastic method remains Monte Carlo (MC) simulation even though it can be computationally demanding due to the great numbers of samples it requires to represent uncertain inputs. MC simulation combined with advanced sampling methods [4] can alleviate the computational burden for statistical problems of small size. Efficient analytical/approximate stochastic techniques have been recently investigated in the field of power systems [5]–[10]. **The point estimate method [5], [6] approximates the statistical moments of some observation variables but it does not provide their detailed probability distribution. The cumulant method [7], [8] works only for linear (or almost linear) problems which is not the general case. FOr these reasons, in this paper, we focus on the** Polynomial Chaos (PC) method [9], [10]. The basic idea behind such techniques is that of approximating the relationships that exist between the uncertain inputs and the node voltages by means of surrogate models. The surrogate model can then be exploited within a MC iterative procedure in place of the time consuming deterministic power flow. The polynomial chaos method has been recently employed to efficiently derive the Probability Density Function (PDF) and cumulative probability of line voltages by considering constant power loads at a fixed time instant [9], [10]. In this paper, we build on these recently presented techniques as well as on generalized Polynomial Chaos (gPC) Theory [11] to extend its application/usage. In our analysis, we consider realistic load power profiles, described by quasi static time series over given observation time windows, and we aggregate the profiles in areas, or geographic regions, of the network infrastructure [12].

The original contribution of the proposed method could be sumarized in the following issues:

1. **A region-wise uncertainty quantification analysis is proposed where load uncertainty in each region and for each one of the three phase lines is represented by a single independent random variable. This approach allows dealing with the relevant case of 1-phase loads.**

2. **We exploit the gPC paradigm to evaluate the PDF of a set of Quantities of Interests (QoI) that affect the quality of the network. Such QoI can include the peak and minimum voltage at some observation nodes, and over some observation time windows, as well as the peak value of other figure of merits such as the Voltage Unbalance Factor (VUF).**
3. **The last issue considered in this paper is connected with the analysis of the impact of new load on the grid, identifying the regions in which the new load insertion could be usefull for example for voltage balancing. The increase or decrease in power demand at a given line phase in one of the regions, for instance due to the allocation of new 1-phase loads, reverberate on the nodal voltages in other regions (and phase lines) in a way that is not easy to be predicted. This is fundamental in the case of charging of electrical**

G. Gruosso, R. S. Netto and P. Maffezzoni are with Politecnico di Milano, DEIB, I20133 Milan, Italy, Z. Zhang and L. Daniel are with Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA.

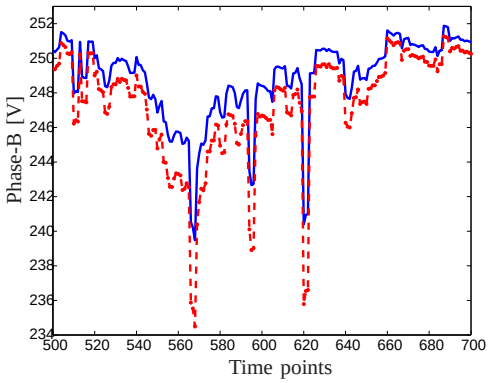

Fig. 1. Effect on the phase B of a positive variation of the loads connected to Phase B. The nominal behavior (solid line) is compared with the effect of load variation (dashed line). Time points indicate minutes.

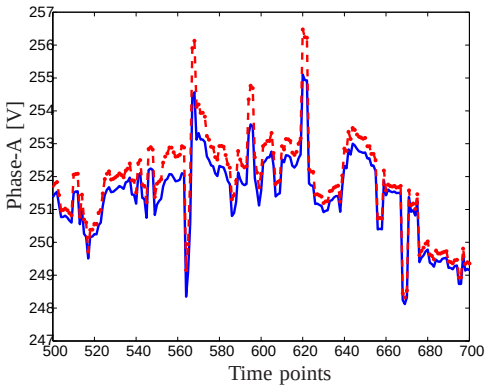

Fig. 2. Effect on the phase A of a positive variation of the loads connected to Phase B. The nominal behavior (solid line) is compared with the effect of load variation (dashed line).

vehicles, or when a storage system is integrated in the grid: therefore perational planning, and proper allocation of new loads, thus requires region-wise and phase-wise sensitivity analysis tools able to efficiently foresee the effects that variations in the aggregated loads of a region can have on the whole network. Several numerical methods for sensitivity analyses in power distribution lines have been provided in the literature that exploit the Jacobian matrix used in power flow calculation [13] or adopt a perturb and observe approach [14]. The first category of techniques are intrusive methods that need the access to the simulation code. In this paper, we concentrate on the second category of methods. We show how an efficient region-wise sensitivity analysis can be naturally derived by exploiting the same gPC simulation framework used for PLF. As a result, the PLF analysis and region-wise sensitivity can be implemented in a homogeneous framework.

The remainder of this paper is organized as follows: in Sec. II we shortly review the deterministic power flow problem and illustrate, with an example, the aims and importance of variability analysis. In Sec. III, the region-wise approach is outlined, while in Sec. IV the gPC-based uncertainty quantification method and sensitivity analysis are described. In Sec. V, we provide more details about the gPC implementation in connection with the Stochastic Testing (ST) selection of sampling points. Finally, in Sec. VI we report simulation results for a IEEE benchmark distribution netwrok case study. In particular, we prove how the proposed methodology can be of guidance for the optimal allocation of new loads in the network.

II. Power Demand Variability and Motivation of the Work

In this paper, we refer to a generic low voltage distribution network made of N_l lines and N buses and designed to provide the prescribed power flows at the network terminals. Deterministic load flow analysis is mathematically formulated as a set of nonlinear equations of the type:

$$\mathbf{F}_n(\vec{\mathbf{V}}) = \mathbf{S}_n - \mathbf{V}_n \sum_{i=1}^{N} \mathbf{Y}_{ni}\mathbf{V}_i^* = \mathbf{0} \qquad (1)$$

for $n = 1, \ldots, N$. In (1), $\mathbf{S}_n = P_n + jQ_n$ denotes the complex power **injection** at node n where P_n and Q_n are the active and reactive powers respectively at network terminations. $\mathbf{V}_n$ denotes the nth node voltage phasor, while $\mathbf{Y}_{ni}$ are the entries of the bus admittance matrix. Node voltage phasors are collected into vector $\vec{\mathbf{V}}$. Power demand at terminations vary in time and thus powers $P_n(t)$, $Q_n(t)$ are functions of time. For a given observation time window (e.g. a day or a week), that is discretized into a sequence of N_t equally-spaced time instants $t_m = m \cdot \Delta t$, power demand is thus specified as given power profiles $P_n^0(t_m)$, $Q_n^0(t_m)$ for $m = 1, \ldots, N_t$. Node voltage waveforms $\mathbf{V}_n(t_m)$ are thus calculated for such nominal load condition by repeatedly solving the nonlinear problem (1) over the sequence of time instants t_m **using the simulation platforms OpenDss [15]. This software performs quasi-static time series (QSTS) simulations, i.e. the chronology of loads at adjacent time points is accounted for by enforcing the dependency of the solution on the hystory of loads.**

A. Aims of the analysis

Due to the uncertainty of power demand, actual power profiles exhibit variations around their nominal values that can be described statistically. Such a statistical variability of power loads induces fluctuations in the node voltages that may bring them out of the safe limits and compromise the quality of the service. The peak and minimum magnitude values assumed by node voltages over some observation time windows are thus crucial QoI in order to check if the network is operating properly.

In order to better explain these concepts, we now present some preliminary results obtained by simulating the benchmark *IEEE European low voltage test feeder* [16] using OpenDSS simulation software [15]. The topology and details about such a network will be given in Sec. VI and in Fig. 6. Here, we anticipate that such a network is a 3-phase network

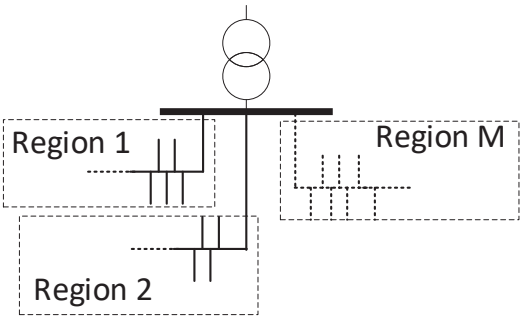

Fig. 3. Example of subdivision of the grid into M Regions, each one of them containing loads connected to the 3 phases.

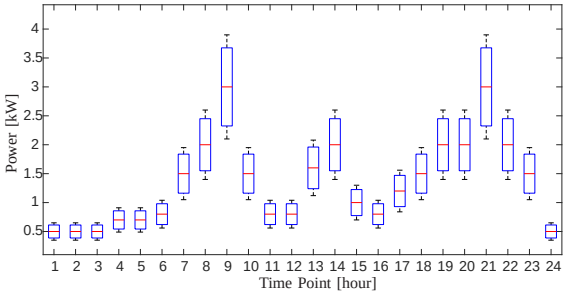

Fig. 4. Load Variation: for each time istant we consider a random gaussian variation of the value

with the possibility of assigning the terminal powers either as 3-phase or 1-phase loads. In this work, we assume that 55 1-phase loads are connected to the network and distributed among the three-phase lines. The network is first simulated with nominal load profiles connected and node voltages are calculated over a certain time window. Second, the total power demand for all of the loads connected to phase line B are increased by a 10% factor and node voltages are recalculated.

Fig. 1 shows the nominal and perturbed waveforms for the phase-B voltage at node 898 over the time window from 8:20 AM to 11:40 AM: the increase in the total power demand on phase-B line results in a reduction in the phase-B voltage with a decrease of the peak value and a more pronounced decrease of the minima. Fig. 2, instead, reports the simulated nominal and perturbed waveforms for the phase-A voltage at node 898 for the same power perturbation: the increase in total power demand on phase-B line results in an increase of phase-A voltage both in peaks and minima.

This example highlights two important issues. A first issue is that, in general, *the variation in power demand of the loads connected to a given phase line can affect the node voltages on all of the three phase lines with fluctuations in the peaks and minima that are difficult to be predicted a priori*. The problem is made more complex since power demands in different areas of the distribution network and at different phase lines can vary independently and in a random way giving rise to a great number of combination loads and scenarios. A second issue is that *variations in the power demand of loads connected to a given phase line induce node voltage variations of different sign on the three phase lines*. This tends to introduce/exacerbate possible voltage unbalance. Such a problem occurs for instance when charging of electrical vehicle in residential dwelling [17]. The figure of merit used to quantify network unbalance is the voltage unbalance factor whose definition is reported in the next section [18].

III. Region-wise Probabilistic Analysis

The method that we propose in this paper is independent of the simulation scenario, the loads behaviour and the grid topology. However, for the sake of illustration, we focus on an example based on Low Voltage distribution grids. In our probabilistis analysis, we suppose to partition the network into M disjunct geographic regions R_k, with $k = 1, \ldots, M$ as shown in Fig. 3. **Each region represents an aggregation of loads, e.g., the loads of the same building or block. However, other choices may be considered as well.**

The active power profiles $P_{n_i}(t)$ at all of the nodes n_i internal to a given region R_k, i.e. $n_i \in \mathrm{R}_k$, and connected to a given phase line are modelled in the following way:

$$P_{n_i}(t) = p^0_{n_i}(t)\,[1 + \sigma_{n_i}\,\xi_r] \tag{2}$$

where $p^0_{n_i}(t)$ is the known nominal power profile at node n_i. In (2), ξ_r is a zero-mean Gaussian-distributed **random variable (i.e. the parameter uncertainty)**. having unitary variance that incorporates power demand uncertainty. The quantity σ_{n_i} is a scaling constant that determines the *degree of uncertainty* at node n_i. **In general, two nodes aggregated in the same region may have different degrees of uncertainty.** In this way, the active power $P_{n_i}(t)$ is a stochastic process whose mean value and standard deviation are given by [19]:

$$\begin{aligned} &\langle\, P_{n_i}(t)\,\rangle = p^0_{n_i}(t) \\ &\sqrt{\langle\,(\,P_{n_i}(t) - p^0_{n_i}(t)\,)^2\,\rangle} = \sigma_{n_i}\; p^0_{n_i}(t). \end{aligned} \tag{3}$$

It is worth observing that for a network partitioned in M regions, the number l of independent statistical parameters ξ_r is $l = 3 \times M$. By changing M it is possible to vary the detail of the analysis from the simple case with a single region (and 3 statistical parameters modulating total loads at each phase) to the extreme case where the power demand at each single (1-phase) load is weighted by an independent statistical parameter.

While power demands (2) represent the *inputs* of the probabilistic analysis, the *outputs* are given by a set of q *observation variables*, that are expected to affect the quality of service, generically denoted as V^j, with $j = 1, \ldots, q$.

Such variables can include the peak and minimum values assumed by the phase voltages (or line currents) over some observation time windows. For the case of 1-phase loads, which is considered in this paper, another relevant observation variable is the peak value of the VUF. The percentage VUF is defined as the ratio of the negative voltage sequence component V_n to the positive voltage sequence component

V_p [18], i.e.

$$\text{VUF} = \frac{|V_n|}{|V_p|} \cdot 100. \tag{4}$$

IV. Uncertainty Quantification and Sensitivity Analysis with gPC method

A. Uncertainty Quantification

We formalize the probabilistic problem where the uncertain load power profiles are described by means of l random Gaussian-distributed parameters ξ_r as in (2). Such parameters are collected into the vector $\vec{\xi} = [\xi_1, \xi_2, \ldots, \xi_l]$. **Each realization of the random variables ξ_r corresponds to well determined power profiles and thus to well determined voltage profiles calculated by solving the load flow problem (1). As a result, the jth observation variable $V^j(\vec{\xi})$ (e.g. a node voltage) is a nonlinear function of the random variables $\vec{\xi}$ and thus it is a random variable too.**

The generalized polynomial chaos (gPC) method consists in approximating each observation variable with an order-β truncated series expansion of the type [11]

$$V^j(\vec{\xi}) \approx \sum_{i=1}^{N_b} c_i^j \, H_i(\vec{\xi}), \tag{5}$$

formed by N_b multi-variate basis functions $H_i(\vec{\xi})$ weighted by unknown polynomial chaos coefficients c_i^j.

Each multi-variate basis function is given by the product

$$H_i(\vec{\xi}) = \prod_{r=1}^{l} \phi_{i_r}(\xi_r) \tag{6}$$

where $\phi_{i_r}(\xi_r)$ is a univariate orthogonal polynomial of degree i_r. The form of the univariate polynomials depends on the density function of the rth parameter ξ_r. For Gaussian-distributed variables ξ_r, the associated $\phi_{i_r}(\xi_r)$ are the Hermite polynomials

$$\begin{aligned} \phi_0(\xi_r) &= 1 \\ \phi_1(\xi_r) &= \xi_r \\ \phi_2(\xi_r) &= \xi_r^2 - 1 \\ \phi_3(\xi_r) &= \xi_r^3 - 3\xi_r \\ &\vdots \end{aligned} \tag{7}$$

The PC polynomials satisfy the orthogonality property

$$\langle \phi_i, \phi_j \rangle = \int_{\mathbb{R}} \phi_i(\xi_r)\phi_j(\xi_r)\rho_r(\xi_r)d\xi_r = \delta_{i,j}, \tag{8}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product operator, $\rho_r(\xi_r)$ is the PDF for the variable ξ_r, and $\delta_{i,j}$ is the Kronecker delta function.

For a given number of parameters l and series expansion truncation order β, the degrees i_r of univariate polynomials in (6) forming $H_i(\vec{\xi})$, for $r = 1, \ldots, l$, satisfy the following relation

$$\sum_{r=1}^{l} i_r \leq \beta. \tag{9}$$

As a consequence, for truncation order β and number of parameters l, the number of gPC basis functions in (5) is given by [20]

$$N_b = \frac{(\beta + l)!}{\beta!\, l!}. \tag{10}$$

Once the coefficients c_i^j are computed, the mean value and standard deviation of the jth observation variable $V^j(\vec{\xi})$ can easily be deduced [20]. In addition, the gPC expansion (5) provides a surrogate compact model for the multi-dimensional relationship $V^j(\vec{\xi})$ that links observation variables to random gaussian parameters. As a consequence, the detailed PDF shape of $V^j(\vec{\xi})$ can be determined by running a large number of uncertainty vector realizations $\vec{\xi}^k$ in very short times, i.e. one million of evaluations take a few seconds on a quad-core computer.

B. Sensitivity Analysis

For **notational** simplicity, we denote $V(\vec{\xi}) = V^j(\vec{\xi})$ the jth observation variable and restrict its gPC order expansion to $\beta = 2$. For l independent Gaussian-distributed parameters, the gPC expansion (5) of $V(\vec{\xi})$ contains $N_b = (l+1)(l+2)/2$ terms that can be ordered into linear and nonlinear ones as follows

$$\begin{aligned} V(\vec{\xi}) \quad &\approx c_0 + \sum_{i=1}^{l} c_i \xi_i \\ &+ \sum_{i=1}^{l} \sum_{\substack{k=i+1 \\ k \neq i}}^{l} c_{(l+i)}\, \xi_i\, \xi_k + \sum_{i=1}^{l} c_{(N_b - l - 1 + i)} (\xi_i^2 - 1). \end{aligned} \tag{11}$$

As an instance, for the case of $l = 3$ random gaussian parameters, it results

$$\begin{aligned} V(\xi_1, \xi_2, \xi_3) \quad &\approx c_0 + c_1\, \xi_1 + c_2\, \xi_2 + c_3\, \xi_3 \\ &+ c_4\, \xi_1\, \xi_2 + c_5\, \xi_1\, \xi_3 + c_6\, \xi_2\, \xi_3 \\ &+ c_7\, (\xi_1^2 - 1) + c_8\, (\xi_2^2 - 1) + c_9\, (\xi_3^2 - 1). \end{aligned} \tag{12}$$

The sensitivity of $V(\vec{\xi})$ versus the rth parameter ξ_r results:

$$\left. \frac{\partial V(\vec{\xi})}{\partial \xi_r} \right|_{\vec{\xi} = \vec{0}} = c_r \tag{13}$$

and thus it simply corresponds to the rth linear term coefficient in the gPC expansion.

V. Computing the gPC coefficients

There are two mainstream approaches for computing the gPC expansion coefficients in (5): Galerkin Projection (GP) and Collocation Method (CM) [21]. Galerkin projection is an *intrusive* numerical technique that requires modifying the LF code (1). GP is numerically robust, however the formation and solution of the projection equations require a significant computational effort which limits the practical applicability to problems of small size and with a few statistical parameters (2-3). For such reasons, in this paper we will focus on Collocation methods.

A. Stochastic Collocation

Stochastic collocation, is an approximate technique that allows the application of gPC method to problems with a greater number of statistical parameters. Furthermore, Stochastic collocation is a *nonintrusive* and thus it can be combined with any LF formulation (1) without modifying the implementation codes. According to collocation method, a gPC expansion of the type (5) is adopted for each observation variables V^j. Then, the expansion coefficients c_i^j in the series (5) are calculated by properly selecting a set of testing points where the series expansions (5) are enforced to fit the values of observation variables V_k^j. Stochastic collocation is very efficient however its accuracy depends on the way testing points are selected. In this paper, we will focus on a recently-proposed robust method referred to as Stochastic Testing (ST) method which allows implementation as a non intrusive collocation method [20]. According to ST method the N_b unknown coefficients c_j in the series (5) are calculated by selecting $N_s = N_b$ testing points $\vec{\xi}^k$, for $k = 1, \ldots, N_s$ with the method outlined in the next Sec. V. In each one of the testing points, the observation variable $V_k(t) = V(\vec{\xi}^k)$ is evaluated by running a deterministic LF analysis. Hence, the series expansions (5) are enforced to fit *exactly* (i.e., the polynomials interpolate the samples) the values V_k^j at the testing points. For the jth observation variable, this results in the following linear system

$$\mathbf{M}\,\vec{c}^{\,j} = \vec{V}^j, \tag{14}$$

where $\vec{c}^{\,j} = [c_1^j, \ldots, c_{N_b}^j]^T$ and $\vec{V}^j = [V_1^j, \ldots, V_{N_s}^j]^T$ are the column vectors collecting the unknown coefficients and observation variable values respectively.

The $N_b \times N_b$ square matrix $\mathbf{M} = \{a_{k,i}\} = \{H_i(\vec{\xi}^k)\}$ collects the gPC basis functions evaluated at the testing points, i.e.

$$\mathbf{M} = \begin{bmatrix} H_1(\vec{\xi}^1) & \ldots & H_{N_b}(\vec{\xi}^1) \\ \vdots & \ddots & \vdots \\ H_1(\vec{\xi}^{N_s}) & \ldots & H_{N_b}(\vec{\xi}^{N_s}) \end{bmatrix}. \tag{15}$$

It is worth observing that matrix $\mathbf{M}$, sometimes referred to as the experiment matrix, remains the same for each observation variable, so it is precalculated, inverted and used for any j as follows:

$$\vec{c}^{\,j} = \mathbf{M}^{-1}\,\vec{V}^j. \tag{16}$$

Generalizations of (16) have also been presented in the literature where a number of testing points greater than the number of basis is selected, i.e. $N_s > N_b$ [9], [22].

In this case, the overdetermined system (16) can be solved with a linear regression technique such as the least-squares method:

$$\vec{c}^{\,j} = (\mathbf{M}^T\mathbf{M})^{-1}\,\mathbf{M}^T\,\vec{V}^j. \tag{17}$$

B. Testing points selection

According to gPC+ST method, the selection of the testing points $\vec{\xi}^k$ in the stochastic space is done as to preserve its robustness compared to Galerkin projection method. To this aim, the scalar product (that implements projection) between two polynomials of the series expansion (i.e. the product of such polynomials has degree β at most) is best approximated by a Gauss quadrature formula with $\beta + 1$ nodes

$$\langle \phi_i, \phi_j \rangle = \int_{\mathbb{R}} \phi_i(\xi_r)\phi_j(\xi_r)\rho_r(\xi_r)d\xi_r \approx \sum_{k=1}^{\beta+1} \phi_i(\xi_r^k)\phi_j(\xi_r^k)w_r^k, \tag{18}$$

where ξ_r^k denotes the kth quadrature node and w_r^k the corresponding weight. The $\beta + 1$ quadrature nodes ξ_r^k are thus good testing points for the single uncertainty parameter ξ_r. When the multivariate case with l parameters is concerned, the testing points vectors $\vec{\xi}^k = [\xi_1^k, \xi_2^k, \ldots, \xi_l^k]$ are determined by considering the multi-dimensional grid of all of the possible combinations (i.e. the tensor product) of the univariate quadrature nodes.

In general, the number $(\beta + 1)^l$ of nodes in the multi-dimensional grid is greater than the number N_b of basis functions defined in (10). A subset of quadrature nodes can thus be selected as testing points to form systems (16) or (17). Stochasting testing selection strategy consists in preferring those quadrature nodes associated to largest Gauss weights and leading to the best (smallest) condition number for the experiment matrix $\mathbf{M}$ [20]. **The flow Chart of the whole procedure is shown in Fig. 5.**

VI. Numerical results

In this section, we focus on low voltage networks having radial topology where the impact of load unbalance is expected to be particularly significant. However, the proposed analysis is general and can easily be extended to other types of electrical systems and other topologies, such as looped or weakly looped ones. The considered network is the *IEEE European low voltage test feeder* [16] described by the circuit reported in Fig. 6 that represents a benchmark case study. This test feeder is radial with a base frequency of 50 Hz, at 230 V (phase voltage)/416 V (line to line voltage). The medium voltage system supplying the substation is modeled as a voltage source with an impedance (Thevenin equivalent) accordling with [16].

The three-phase network has 906 low voltage nodes that are connected by 905 branches. The value of the line impedance and shunt admittance used are defined in [16], but due to the short length of lines (the branches are shorter than one hundred meters) the shunt admittance is neglected and just the series impedance is considered. The original benchmark is provided with 55 1-phase loads that are applied to the nodes represented by marks in Fig. 6 and subdivided as follows: 21 for the phase A (red), 19 for the phase B (Black), 15 for the phase C (Green).

A. Probabilistic analysis

The proposed region-wise uncertainty quantification analysis is applied by partitioning the 55 given loads into three geographic regions as shown in Fig. 6: Region I contains 19 loads (square mark in Fig. 6) distributed among the three phase lines as reported in Table I, Region II has 19 loads (circle mark) while Region III includes 17 loads (cross mark).

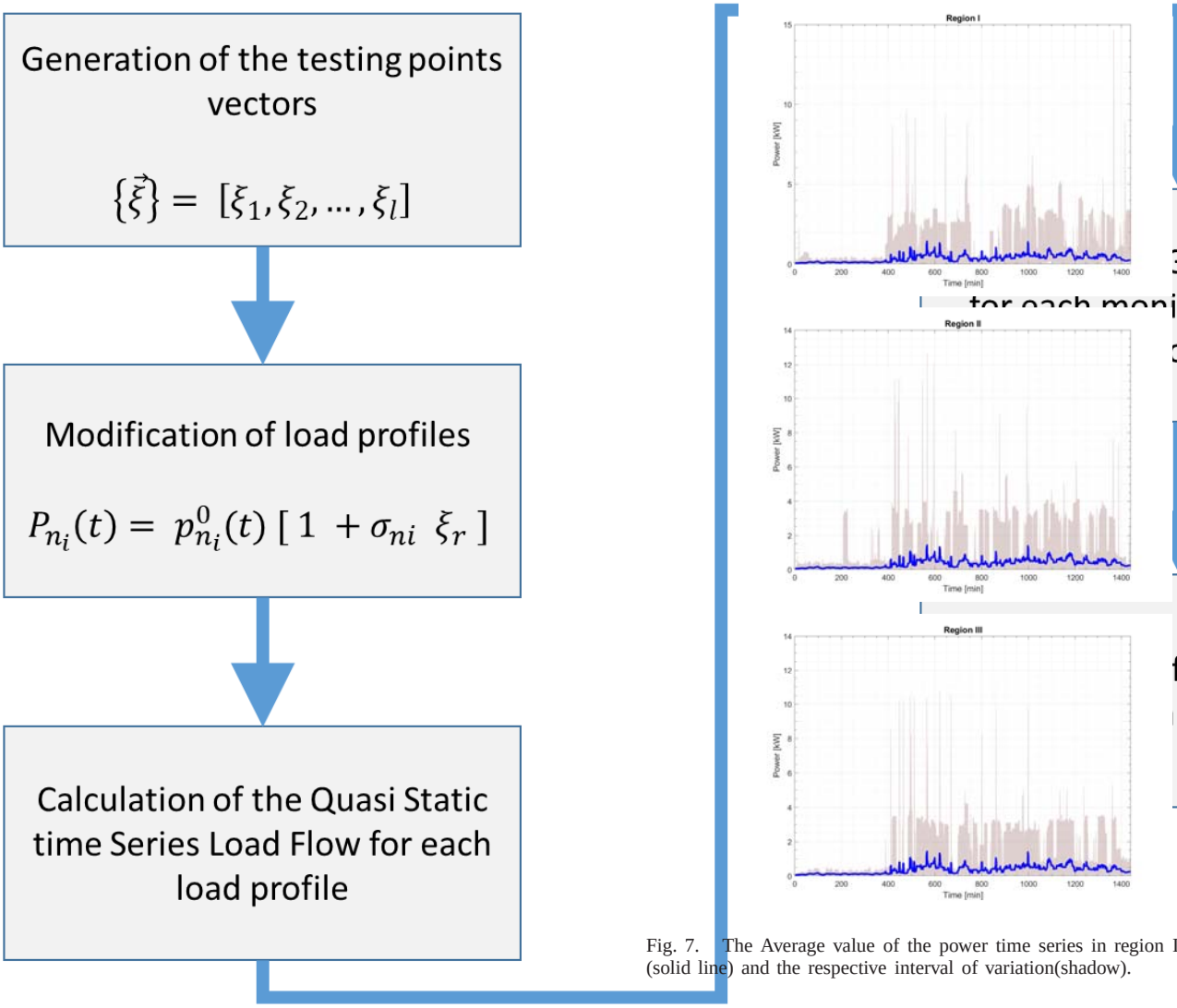

Fig. 5. **Flow chart describing the main step of the proposed methodology**

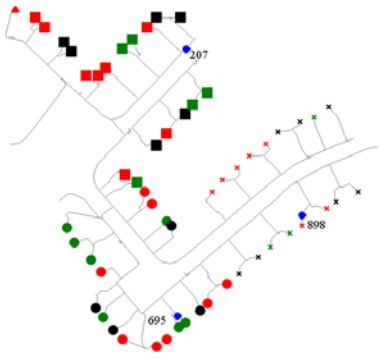

Fig. 6. Topology of the IEEE LV European test feeder. Nodes (207,695,898) in three different regions are monitored in order to determine the effect of the load variation. Mark legend: square= Region I, Circle= Region II, Triangle= Region III. Color Legend: **Red=Phase A**, **Black=Phase B**, **Green = Phase C**.

In what follows, load flow simulations are performed using quasi-static time series power profiles made of 1440 samples (i.e., 24 hours $\times$ 60 minutes) that change for the different

Fig. 7. The Average value of the power time series in region I, II and III (solid line) and the respective interval of variation(shadow).

TABLE I
NUMBER OF LOADS FOR REGIONS AND PHASES

I-A	I-B	I-C	II-A	II-B	II-C	III-A	III-B	III-C
7	7	5	7	5	7	7	7	3

loads. Fig. 7, shows the average value of the active power and the min-max variation interval at each time sample, for the loads in the three Regions I, II, and III, respectively. In such figures, the solid line indicates the average value (i.e. on the ensemble of the 55 loads) while the shadowed area represents the variation interval of the power profiles: all of the loads lie inside that area. Only the active powers are reported in the above figures while the associated reactive powers are assumed to change accordingly with a fixed power factor of 0.9. It is seen how power demand becomes significant from 6:00 A.M. (time equals to 360 minutes in the graphs) to midnight. Within this time range, the power profiles at each time exhibit a wide range of variability thus providing a good benchmark case for testing our method.

Fig. 8 reports the waveforms of the three phase voltages at node 207, used here as the monitoring node for Region I,

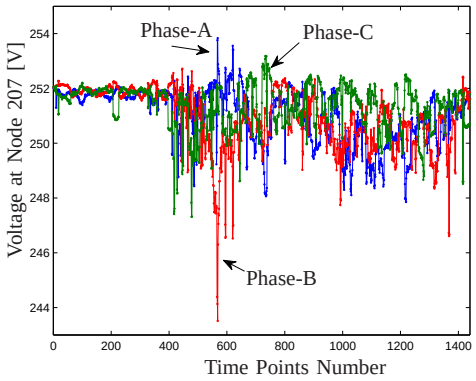

Fig. 8. Time domain evolution of the three phases at node 207.

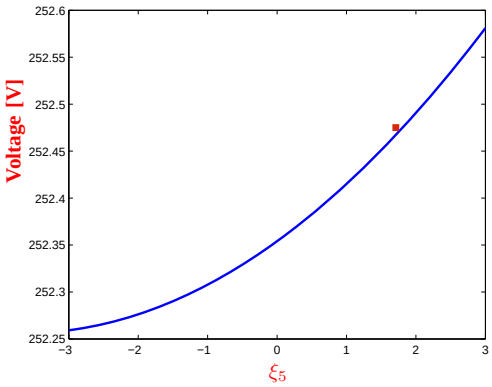

Fig. 9. The line represents the curve of the Phase-C peak value provided by th gPC model vs ξ_5 with the other ξ_r equal to 0. The red square Marker represents the extra sample used for verification.

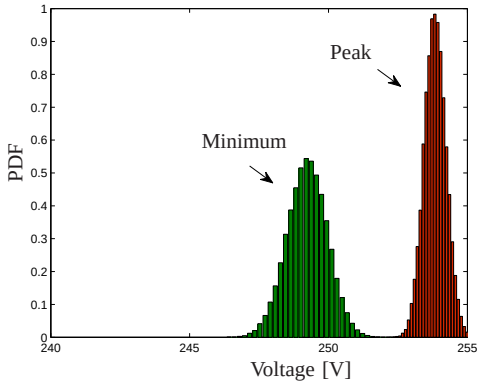

Fig. 10. Statistical distribution of the minimum and peak values for the Phase A at node 207 on the 9:00-10:00 time window.

simulated with OpenDSS with the nominal load profiles provided with the benchmark. Voltage waveforms exhibit sharp fluctuations in time, from 6:00 A.M. on, due to the variations in the load profiles. A first set of QoI is thus represented by the peaks and minima values assumed by the phase voltages over given observation time windows. As an example, we consider the time window from 9:00 A.M to 10:00 A.M. (corresponding to samples from 540 to 600) and with the proposed gPC+ST method, we calculate the statistical distribution of the minima and peak values, assumed over the time window, for the three phase voltages, A, B and C. To this aim, the uncertainty in the load profiles is modelled as in (2) by means of $l = 9$ zero-mean, unitary-variance Gaussian distributed parameters ξ_r that randomly scale the given nominal power profiles. The degree of uncertainty at each node n_i is fixed to $\sigma_{n_i} = 0.2$ for all of the loads and phase lines.

In view of (10), for $l = 9$ stochastic parameters and gPC expansion order $\beta = 2$, the gPC series expansion is made of 55 basis functions. In our implementation we generate 56 testing points in the space of parameters using the stochastic testing method reviewed in Sec. V.A and for each one of them a deterministic load flow analysis is performed. The extra sample point is used in a leave-one-out cross validation error method to check the accuracy of the gPC expansion with $\beta = 2$ [9]. **Fig.9 shows, as an example, the curve provided by the gPC model for one observation variable considered in what follows (i.e. the peak value of the Phase-C at node 207) as a function of one of the parameter uncertainty (i.e. parameter ξ_5) with the other parameters fixed to zero. The red square marker represents the extra sample (i.e. not used to calculate the gPC model). The good fitting of the gPC model shows that truncation order $\beta = 2$ is adequate for the nonlinearities considered in this paper. Figs. 10, 11, and 12 show the calculated PDF for the three phases A, B, C, respectively at node 207.** .

We see how, for the load arrangement provided by the benchmark, Phase-B reaches the lowest voltage levels and exhibits the greatest uncertainty in the minimum value that ranges within the interval $(242, 245)$ V with 90% probability. By contrast, Phase A reaches the highest voltage levels with the peak value that ranges within the interval $(253.5, 255)$ V with 90% probability. Finally, Phase C fluctuates within narrower intervals but its peak value is nonGaussian distributed. This result can be better seen with the aid of Fig. 13 where the statistical distributions of the peak value of Phase-C computed with the proposed gPC and with the reference MC method are reported and compared with the Gaussian distribution of equal mean value and variance.

The reference MC method uses $10,000$ runs (i.e. deterministic load flow analyses) selected with a latin-hypercube sampling. With this setting, the peak value distributions provided by the proposed gPC and reference MC method are almost superimposed and the associated standard deviations, i.e. $\sigma_{gPC} = 0.3080$ V and $\sigma_{MC} = 0.3071$ V respectively, match within a relative accuracy of 2%. Since the gPC+ST method only requires 55 deterministic analyses, it introduces a $180\times$ computational speedup factor compared to the reference MC analysis for the same accuracy. Table (II) reports the standard deviations predicted by the MC method for a growing number of samples and the relative error compared to the value provided by the reference MC (i.e. with $10,000$ samples). If a lower order of accuracy for MC method, e.g. a 5% accuracy (which requires about 6000 samples), is accepted the computational cost of MC remains about 100x greater than that of gPC. It is thus reasonable to say that, for

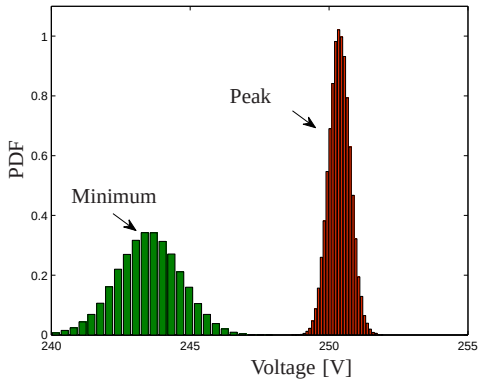

Fig. 11. Statistical distribution of the minimum and peak values for the Phase B at node 207 on the 9:00-10:00 time window.

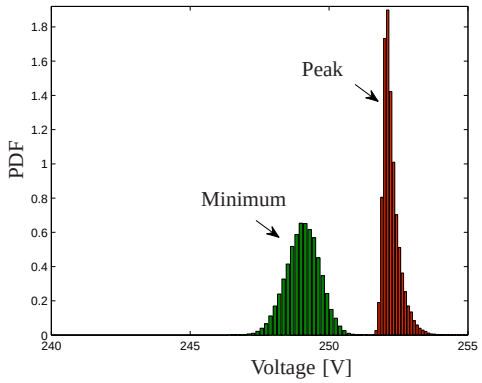

Fig. 12. Statistical distribution of the minimum and peak values for the Phase C at node 207 on the 9:00-10:00 time window.

our example, gPC results about two orders of magnitude faster than MC.

*B. Sensitivity analysis:**a strategy for load guidance***

The probability distributions of the peaks and minima for the three phases at a given observation node provide a measure of their statistical uncertainty over the considered observation time window. Phase voltage uncertainty corresponds to an analog statistical uncertainty in the voltage unbalance factor VUF defined in (4) which represents another relevant QoI. The statistical distribution of the VUF can be derived with the proposed gPC+ST method by using the same 55 deterministic load flow analyses described in the previous subsection VI.A. Furthermore, according to (13), the linear coefficients of the gPC expansion for the VUF observation variable provide de-

TABLE II

CONVERGENCE OF MC VS NUMBER OF SAMPLES

Number of Samples	**2500**	**5000**	**10000**
σ_{MC}	**0.351**	**0.324**	**0.3071**
error	**14%**	**5.5%**	**-**
CPU time for 1440 point timeserie [sec] Intel Core i5 - 3.20 GHz	**12500**	**25000**	**50000**

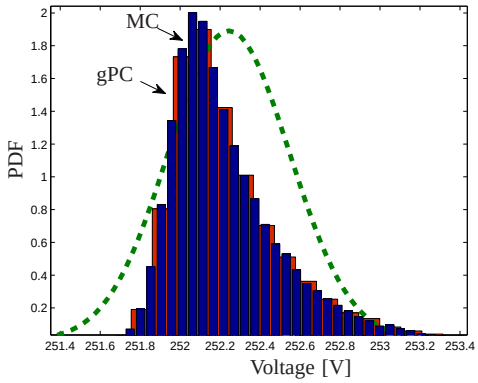

Fig. 13. (Histogram) Detail of the distributions of the Phase-C peak value as computed with gPC and MC method (10,000 samples). (Green Dashed Line) Gaussian distribution of equal mean value and variance.

TABLE III

VUF SENSITIVITY COEFFICIENTS $\times$ 100

	I-A ξ_1	I-B ξ_2	I-C ξ_3	II-A ξ_4	II-B ξ_5	II-C ξ_6	III-A ξ_7	III-B ξ_8	III-C ξ_9
N. 207	-0.25	3.48	-0.82	-0.65	6.98	-0.25	-0.31	5.79	-0.08
N. 695	-0.13	3.68	-1.08	-1.02	12.18	-0.51	-0.41	9.19	-0.14
N. 898	-0.11	3.7	-1.13	-0.98	12.22	-0.54	-0.47	11.63	-0.15

tailed information about its sensitivity with respect to possible variations of the power demands in each region and phase line.

Table III reports the sensitivity coefficients for the three observation nodes 207, 695 and 898 in the three Regions. For a compact notation, such values are reported multiplied by a factor 100. A first important information contained in Table III is the sign of the sensitivity coefficients: a negative sign indicates that an increase in power demand, at that phase and Region, corresponds to a decrease of the VUF and thus to a beneficial effect on load balancing. Viceversa, a positive sign of the sensitivity coefficient indicates that an increase in power demand will move the network towards greater unbalance.

In order to check this result, we repeat the probabilistic analysis with 10 new loads, randomly selected among those provided with the benchmark, connected to the distribution network. Two different allocation strategies, referred to as Case

TABLE IV

NUMBER OF LOADS FOR CASE A)

I-A	I-B	I-C	II-A	II-B	II-C	III-A	III-B	III-C
7	7	8	11	5	7	10	7	3

TABLE V

NUMBER OF LOADS FOR CASE B)

I-A	I-B	I-C	II-A	II-B	II-C	III-A	III-B	III-C
7	10	5	7	9	7	7	10	3

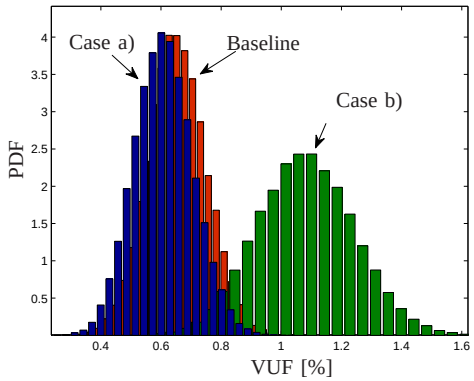

Fig. 14. Statistical distribution of the VUF at node 207.

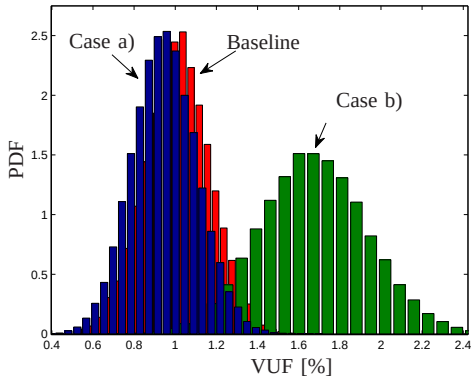

Fig. 15. Statistical distribution of the VUF at node 695.

a) and Case b) and described in Tables IV and Tables V, respectively, are investigated. The allocation of Case a) is done by exploiting the information provided by the gPC-based sensitivity analysis and concentrates the new loads in the Regions and phases having negative sensitivity coefficients with large module. Viceversa, allocation of Case b) is done in defiance of sensitivity analysis, i.e. putting new loads in the Regions and phases with large positive sensitivity coefficients. Figs. 14, 15, 16 show the statistical distributions of the maximum VUF, computed over the time window 9:00-10:00 A.M., for the two different allocation cases and the three observation nodes. It is apparent how Case b) results in a balance deterioration: the mean value and the standard deviation of the VUF increase significantly. At node 898, the allocation of Case b) leads to a 40% probability that the peak VUF will exceed the upper bound of 2%. A similar violation is seen for VUF at node 695. By contrast, the allocation of Case a) results in a reduction of the mean value and standard deviation of the VUF in all of the three observation nodes compared to the baseline case (i.e. 55 loads distributed as in Table I). This shows how sensitivity analysis can help allocating new loads while preserving, or even improving, load balancing in the network.

VII. Conclusion

In this paper, we have described a unified computational framework for probabilistic load flow and sensitivity analyses

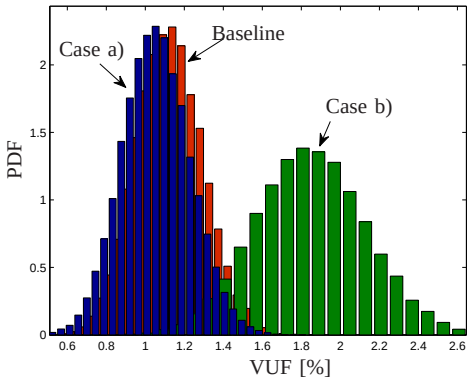

Fig. 16. Statistical distribution of the VUF at node 898.

in distribution networks with uncertain load profiles. Our approach relies on a region-wise uncertainty quantification analysis that aggregates loads within geographic regions of the network and on the usage of the generalized Polynomial Chaos method. We have shown how the detailed PDF of several QoI affecting the quality of service can be calculated with a speed-up factor of $\approx 100\times$ compared to standard Monte Carlo analysis for the same accuracy. In particular, the numerical results have being focused on evaluating the statistical distribution of node voltage peak and minima over a given observation time window and on voltage unbalance factor. Finally, we have proved with an example how the proposed sensitivity analysis can be exploited to properly allocate new (1-phase) loads in the network while preserving, or even improving, load balancing.

References

[1] K. Jhala, B. Natarajan, and A. Pahwa, "Probabilistic voltage sensitivity analysis (pvsa)—a novel approach to quantify impact of active consumers," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2518–2527, May 2018.

[2] Y. Wang, N. Zhang, Q. Chen, J. Yang, C. Kang, and J. Huang, "Dependent discrete convolution based probabilistic load flow for the active distribution system," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1000–1009, July 2017.

[3] J. Tang, F. Ni, F. Ponci, and A. Monti, "Dimension-adaptive sparse grid interpolation for uncertainty quantification in modern power systems: Probabilistic power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 907–919, March 2016.

[4] M. Hajian, W. D. Rosehart, and H. Zareipour, "Probabilistic power flow by monte carlo simulation with latin supercube sampling," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1150–1159, May 2013.

[5] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1594–1601, Nov 2007.

[6] G. Plattner, H. F. Semlali, and N. Kong, "Analysis of probabilistic load flow using point estimation method to evaluate the quantiles of electrical networks state variables," *CIRED - Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2087–2091, 2017.

[7] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2251–2261, Nov 2012.

[8] P. Amid and C. Crawford, "A cumulant-tensor-based probabilistic load flow method," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5648–5656, Sept 2018.

[9] F. Ni, P. H. Nguyen, and J. F. G. Cobben, "Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 694–704, Jan 2017.

[10] Z. Ren, W. Li, R. Billinton, and W. Yan, "Probabilistic power flow analysis based on the stochastic response surface method," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2307–2315, May 2016.

[11] D. Xiu and G. E. Karniadakis, "The wiener–askey polynomial chaos for stochastic differential equations," *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.

[12] D. T. Nguyen, "Modeling load uncertainty in distribution network monitoring," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2321–2328, Sept 2015.

[13] F. Tamp and P. Ciufo, "A sensitivity analysis toolkit for the simplification of mv distribution network voltage management," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 559–568, March 2014.

[14] R. Tonkoski, D. Turcotte, and T. H. M. EL-Fouly, "Impact of high pv penetration on voltage profiles in residential neighborhoods," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 518–527, July 2012.

[15] EPRI, "Opendss." [Online]. Available: http://smartgrid.epri.com/SimulationTool.aspx

[16] IEEE, "Distribution test feeders," may 2015. [Online]. Available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html

[17] G. Gruosso, "Analysis of impact of electrical vehicle charging on low voltage power grid," in *2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles International Transportation Electrification Conference (ESARS-ITEC)*, Nov 2016, pp. 1–6.

[18] P. Pillay and M. Manyage, "Definitions of voltage unbalance," *IEEE Power Engineering Review*, vol. 21, no. 5, pp. 49–51, May 2001.

[19] A. Papoulis and U. Pillai, *Probability, random variables and stochastic processes*, 4th ed. McGraw-Hill, 11 2001.

[20] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel, and L. Daniel, "Stochastic testing method for transistor-level uncertainty quantification based on generalized polynomial chaos," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 10, pp. 1533–1545, Oct 2013.

[21] A. Sandu, C. Sandu, and M. Ahmadian, "Modeling multibody systems with uncertainties. part i: Theoretical and computational aspects," *Multibody System Dynamics*, vol. 15, no. 4, pp. 369–391, 2006.

[22] P. Maffezzoni, Z. Zhang, S. Levantino, and L. Daniel, "Variation-aware modeling of integrated capacitors based on floating random walk extraction," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 10, pp. 2180–2184, Oct 2018.