Modelling and assessment of the combined technical impact of electric vehicles and photovoltaic generation in radial distribution systems

J. C. Hernández[a], F. J. Ruiz-Rodriguez[b], F. Jurado [c]*

[a] *Department of Electrical Engineering, University of Jaén, 23071 EPS Jaén, Spain*

[b] *Department of Electrical and Thermal Engineering, University of Huelva, 21004 Spain*

[c] *Department of Electrical Engineering, University of Jaén, 23700 EPS Linares, Jaén, Spain*

Abstract:

Photovoltaics (PVs) provide new opportunities for radial distribution systems (RDSs) that feed electric vehicle charging stations (EVCSs). However, the accurate assessment of the combined technical impact is problematic because of the uncertainties of sources/loads. In previous research, we developed a technique to assess the impact of PV generation. This new study presents a general analytical technique (GAT) that evaluates the combined impact for an extended time frame. Specifically, the GAT effectively assesses the fulfilment of technical requirements for weekly RDS operating variables as specified in regulations. As our main objective is to improve the assessment accuracy of the EV and PV interaction in RDSs, the weekly assessment was extended to a one-year

* Corresponding author. Tel.: +34 953 648518; fax: : +34 953 648586.

E-mail addresses: jcasa@ujaen.es (J.C. Hernandez), francisco.ruiz@die.uhu.es (F.J. Ruiz-Rodriguez), fjurado@ujaen.es (F. Jurado).

time period, during which it is possible to capture the total uncertainty. Also, correlation of input variables is handled.

The computational cost of the GAT is lower than that associated with Monte-Carlo simulation, which is used to confirm the GAT accuracy. Although the results focus on an RDS located in Spain, GAT is applicable to any RDS and is scalable to different penetration levels. The numerical results show the impact of different correlated and non-correlated case studies on the voltage profile, apparent power flow in lines, and real loss.

Keywords: Distribution system; electric vehicle charging station; photovoltaic power system; probabilistic load flow; probability density distribution; uncertainty.

Nomenclature

List of symbols

a_{i_i} real constants

$\boldsymbol{C}$ [1] multivariable input random variable of the RDS

COP Copula

d any given day of the week ($d = 1,...,7$)

$\boldsymbol{d}$ daily distance covered by an EV

d_r maximum range of an EV

$\boldsymbol{E}(\boldsymbol{E}')$ initial SOC of the battery at the beginning of a recharge cycle, being $\boldsymbol{d}(\boldsymbol{L_1})$ the the distance covered by the EV

E discrete SOC of the battery

$f_{C_i}(\boldsymbol{C}_i)$ PDF of the continuous univariate random variable $\boldsymbol{C_i}$

$f^*_{C_i}(\boldsymbol{C}_i)$ PMF of the discrete univariate random variable $\boldsymbol{C_i}$

$F_{C_i}(\boldsymbol{C}_i)$ CDF of the univariate random variable $\boldsymbol{C_i}$ $(F_{C_i}(\boldsymbol{C}_i) = \int_{-\infty}^{\infty} f_{C_i}(\boldsymbol{C}_i)\, d\boldsymbol{C}_i))$

[1] Bold letters are used to random variables

$F_{C_i}^{-1}$ inverse distribution function of the univariate random variable $\boldsymbol{C_i}$

$g_{t_k'}(\boldsymbol{t_k'})$ PDF of the random variable EV parking duration

G_{ij} (B_{ij}) real (imaginary) part of the element in the bus admittance matrix

$\boldsymbol{G_{g\_\beta}}$ global irradiance on a β tilted surface

$G_{t_k'}(\boldsymbol{t_k'})$ CDF of the random variable EV parking duration

$\boldsymbol{H_g}$ ($\boldsymbol{H_d}$) global (diffuse) irradiation

$h_{t_k}(\boldsymbol{t_k})$ PDF of the random variable charging start time of the EV battery

i, j, k any given number

$K_C(\xi)$ cumulant-generating function of random variable $\boldsymbol{C}$

$\boldsymbol{k_d}$ hourly diffuse fraction

$\boldsymbol{K_t}$ daily clearness index

$\boldsymbol{L_i}$ ($\boldsymbol{l_i}$) distance travelled of the *i*th trip for an EV in km (p.u.)

$M_C(\xi)$ moment-generating function of random variable $\boldsymbol{C}$

m any given month ($m = 1,...,12$)

n any given RDS node

n_n number of RDS nodes

n_{cl} any given commercial load node of the RDS

n_{ev} number of RDS nodes with EV

n_f number of lines of the RDS

n_{il} any given industrial load node of the RDS

n_l number of RDS nodes with load

n_s sample size

n_{p_k} any given RDS node with a representative parking lot

n_{pv} number of RDS nodes with a PV unit

n_{rl} any given residential load node of the RDS

n_{rv} random variable number

N_c number of 10 - min intervals required for full charging process of the EV battery

P_{ev} discrete charging power level of the EV battery

$\boldsymbol{P}_{\boldsymbol{ev}_i}$ charging power of an *i*th single EV

$P_{ev}^{\Diamond}$ continuous charging power of the EV battery

$P_g\,(Q_g)$ real (reactive) generation power of a traditional generator

$\boldsymbol{P}_{\boldsymbol{loss}}$ total real loss in the RDS

$\boldsymbol{P}_{\boldsymbol{l}}\,(\boldsymbol{Q}_{\boldsymbol{l}})$ real (reactive) load power

$\boldsymbol{P}_{\boldsymbol{pv}}$ PV power

$\boldsymbol{P}_{\boldsymbol{tev}}$ total charging power for a given set of EVs

q any given10 - min interval in a day $(q = 1,...,144)$

r any given number

$\boldsymbol{S}_{\boldsymbol{f,i\text{-}j}}$ apparent power flow in line connecting node *i* and *j*

$\boldsymbol{t}_{\boldsymbol{k}}$ charging start time of the EV battery

$\boldsymbol{t}_{\boldsymbol{k}}^{'}$ EV parking duration in which the EV battery is charged

t_x *x*th 10 - min interval

t_1 internal time in a full charging process of the EV battery

t_2 time required for a full charging process of the EV battery

$\boldsymbol{U}_{\boldsymbol{i}}\,(\boldsymbol{U}_{\boldsymbol{C}_i})$ univariate uniform distribution (associated with random variable $\boldsymbol{C}_{\boldsymbol{i}}$)

V_i voltage magnitude at node *i*

$\boldsymbol{W}_{\boldsymbol{C}_i}$ univariate standard normal distribution associated with random variable $\boldsymbol{C}_{\boldsymbol{i}}$

x any given number

$\boldsymbol{X}$ RDS state variables (nodal voltage angles and magnitudes) $\boldsymbol{X} = [\boldsymbol{V}_1, \delta_1, \boldsymbol{V}_2, \delta_2, ,..., \boldsymbol{V}_j, \delta_j]^{\mathrm{T}};\ j = 1,..., n_n$

$\boldsymbol{X}^*$ RDS output variables (apparent power flows in lines and total real loss in the RDS) $\boldsymbol{X}^* = [\boldsymbol{S}_{\boldsymbol{f,i_1\text{-}j_1}}, \boldsymbol{S}_{\boldsymbol{f,i_2\text{-}j_2}}, ..., \boldsymbol{S}_{\boldsymbol{f,i_k\text{-}j_k}}, \boldsymbol{P}_{\boldsymbol{loss}}]^{\mathrm{T}};\ k = 1,..., n_f$

$\boldsymbol{Z}$ RDS operating variables $\boldsymbol{Z} = [\boldsymbol{X}\ \ \boldsymbol{X}^*]^{\mathrm{T}} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, ..., \boldsymbol{Z}_j]^{\mathrm{T}};\ j = 1,..., 2 \cdot n_n + n_f + 1$

$\alpha_{\boldsymbol{C}}^{\overbrace{i_1 \ldots i_r}^{r}}$ rth - order moment of the random variable $\boldsymbol{C}$

$\alpha_{\boldsymbol{V_n^m}}^{\overbrace{1 \ldots 1}^{r},\ \mathbf{GAT}}$ ($\alpha_{\boldsymbol{V_n^m}}^{\overbrace{1 \ldots 1}^{r},\ \mathbf{MCS}}$) rth-order moment of the RDS operating variable $\boldsymbol{V_n^m}$ obtained by the (GAT) MCS

$\varepsilon_{\alpha_{\boldsymbol{V_n^m}}^{\overbrace{1 \ldots 1}^{r}}}$ individual relative error of the rth-order moment of the RDS operating variable $\boldsymbol{V_n^m}$

$\kappa_{\boldsymbol{C}}^{\overbrace{i_1 \ldots i_r}^{r}}$ rth-order cumulant of the random variable $\boldsymbol{C}$

λ number of EVs in a given set

$\mu_{\boldsymbol{d}}$ expected value of the random variable $\boldsymbol{d}$

$\xi_{\boldsymbol{pv}}$ PV cell electrical efficiency

ρ_{C_i,C_j} correlation coefficient in matrix $\sum_{\boldsymbol{C}}$

$\sigma_{\boldsymbol{d}}$ standard deviation of random variable $\boldsymbol{d}$

$\sum_{\boldsymbol{C}}$ ijth correlation matrix of multivariate random variable $\boldsymbol{C}$

φ_i singleton probability for an ith single EV

$\boldsymbol{\phi}$ CDF of the univariate standard normal distribution

ϕ^{-1} inverse of the univariate standard normal distribution

ω weight

δ_{ij} phase angle of voltage from node i to j

Subscripts

ev electric vehicle

l load

n nth RDS node

n_{cl} n_{cl}th commercial load node of the RDS

n_f number of RDS lines

n_{il} n_{il}th industrial load node of the RDS

n_n number of RDS nodes

n_{p_k} n_{p_k}th RDS node with a parking lot

n_{rl} n_{rl}th residential load node of the RDS

pv photovoltaic

q qth 10-min interval in a day

t_x xth 10-min interval (t_x)

Superscripts

d dth day of the week

m mth month

Abbreviations

CDF cumulative distribution function

EVCSs electric vehicles charging stations

GAT general analytical technique

MCS Monte Carlo simulation

PDF probability density function

PMF probability mass function

PLF probabilistic load flow

PRLF probabilistic radial load flow

PV photovoltaic

RDS radial distribution system

SOC state-of-charge

1. Introduction

The development of electric vehicles (EVs) is currently driven by the need to decrease reliance on foreign oil supplies and CO_2 emissions [1,2]. However, their large-scale use involves the massive integration of EVCSs in traditional RDSs. This integration involves numerous technical challenges [3,4]. One of the most important challenges concerns the heterogeneous sources that generate the power for EVCSs. CO_2 emissions can be reduced by the inclusion of renewable power in these RDSs [1,2,4] (e.g. PV power). Although renewable power in RDSs with EVCSs makes the generating system more sustainable [3,4], the assessment of this interaction is problematic because of the inherent

uncertainties associated with the sources/loads involved. Precisely for this reason, this issue has not as yet been successfully addressed. Moreover, the uncertainty parameters may have a considerable level of correlation.

Many probabilistic studies have analysed the potentially negative technical impacts of EVCSs on RDSs such as the following: (i) transformer and cable thermal loading [5-10]; (ii) stability and node voltage profile [5,8,11-17]; (iii) power line losses [5,10,15]; (iv) system power demand [9,14-21]; (v) system reliability and costs [22-24]; (vi) harmonics and unbalance [10]. Similar probabilistic work analysed how PV units produced various adverse technical impacts on RDSs [20,25-28].

Until now, the negative technical impacts of EVCSs or PV units have been minimized by demanding interconnection requirements (EVCS [4] and PV [29]), which are based on probabilistic assessments for worst-case scenarios, i.e. peak load or any other representative snapshot (e.g. PV units [30-32]). These probabilistic assessments limited to one or only a few snapshots provide the statistical characterization (PDF and CDF) for each snapshot or time interval; however, they do not reflect the aggregated statistical behaviour over a longer time period [33].

An accurate assessment of the technical impact of EVCSs and PV units on RDSs should include the investigation of all possible inputs. This is only possible if an aggregated approach of multiple snapshots is used. In fact, the current statistical characterization of time-variant variables in regulation [34,35] for RDSs is performed from the aggregation of actual stochastic values in each time subinterval of a given time frame (usually 10-min intervals[2] over a one-week time period). This provides a single

[2] The thermal effects of overload are related to the 10-min load level.

statistical index [36] that is compared with a regulation limit (e.g. node voltage [34], cable thermal limit[3]).5

A review of the literature [5-28] indicates that the accurate assessment of the EVCS and PV impact has still not been performed from a probabilistic perspective. Thus, major shortcomings found include: (i) a time interval that is not adjusted to the 10-mins required in regulations [34,35] (e.g. 1 min [11], 0.5 h [17], 1 h [5-9,12-14,16,17-26,28]); (ii) a time frame restricted to a low number of hours/days, which leads to an inaccurate assessment (e.g. 1 h [25,26], 3 h [11], 1 day [5,8,9,11,13- 19,21- 24,28], 2 days [6,7], 7 days[20], and 25 days [12]); (iii) assessment of technical impacts without considering their probability of occurrence as stated regulations [34,35]; in other words, the probability distribution of RDS operating variables was not calculated (usually averaged values [6-19,21,22,24]); (iv) the correlation of the input variables is rarely considered [17,20,25,27]; (v) the statistical behaviour aggregated in longer time frames has not been investigated, especially when this is the most important shortcoming to be surmounted.

Various methods of probabilistic load flow (PLF) have been presented in literature to evaluate the impact of the input uncertainties in power systems. Of these methods, the most frequently used are: (i) MCS [5-9,11-18,22]; (ii) Latin hypercube sampling [25]; (iii) analytical techniques (e.g. fast Fourier transform and cumulants method combined with series expansion [26,27]); and (iv) point estimate methods [16]. The first two are relatively computationally expensive and time-consuming, when compared with the other two. In previous research [27], we developed an analytical technique for RDSs that was computationally more effective than the mainstream MCS [37], and which was just as accurate. Firstly, linear models were defined to handle the nonlinearities of the power

[3] Specific standard for each cable.

system with a lower computational burden. Then, convolution techniques [38-40] were implemented to obtain the statistical information of RDS operating variables. Finally, Cornish-Fisher expansion [41,42] estimated their statistical characterization. However, this research was specifically conceived for RDSs with correlated PV units and loads. Furthermore, it was only able to evaluate the technical impact for one 10-min interval and did not take into account longer time simulations.

This new study further develops the research in [27] and presents an innovative GAT that provides the accurate assessment of the combined technical impact of EVCSs and PV units on RDSs. As main goal, the method in GAT pursues to overcome each of the shortcomings previously highlighted. Thus, GAT carries out this impact assessment in a time frame that covers a one-year period, with a weekly assessment on a 10-min basis each month of the year. In this way, GAT is able to accurately capture the effects of the daily/weekly/monthly variability of input random variables. Besides, GAT takes account of correlation of input variables and provides probability distributions for RDS operating variables. In short, new contributions with respect to previous research in [27] and technical literature are the following: (i) an improved probabilistic model of the EV charging load at home and the development of this probabilistic model for parking lots; (ii) an innovative GAT that manages stochastic time series of the node and EV charging load as well as the PV generation for a one-year period with an affordable computational burden; and (iii) the comparison of different case studies where the interaction of renewable power (specifically PV) with EVCSs in RDSs affords a better understanding of its single or combined technical impact.

The rest of the paper is organized as follows: Section 2 presents the statistical background; Section 3 describes the probabilistic models; Section 4 presents the correlation of input variables; Section 5 outlines the GAT used to evaluate the weekly

operating variables in RDSs throughout the entire year; Section 6 describes the proposed test system; Section 7 presents the simulation results; and finally, Section 8 gives the conclusions that can be derived from this study. Additionally, the appendices clarify technical details.

2. Statistical background: moments and cumulants

For a better understanding of the proposed GAT, this section provides a short introduction to probability theory, which also justifies the use of moments and cumulants. For more details, references [38,43] can be consulted.

The moments of a random variable are a set of descriptive constants of its probability distribution which are useful for measuring its properties and, in certain circumstances, for specifying it [38,42,43]. However, they are not the only set of constants for the purpose, or even the best set. Another series of constants, the so-called cumulants, have properties which are more useful from the theoretical standpoint.

Let $\boldsymbol{C}$ be a random variable, whose components are $\boldsymbol{C}_1,\boldsymbol{C}_2,...,\boldsymbol{C}_{n_{rv}}$, with PDF $f_C(\boldsymbol{C})$, or PMF $f_C^*(\boldsymbol{C})$. Then, the moment of order one, two, and so on about the origin are defined as [38,42,43]:

$$\alpha_{\boldsymbol{C}}^{\overbrace{i_i}^{1}}=\mathrm{E}\left[C_{i_i}\right];\qquad \alpha_{\boldsymbol{C}}^{\overbrace{i_i i_j}^{2}}=\mathrm{E}\left[C_{i_i}C_{i_j}\right];\qquad \alpha_{\boldsymbol{C}}^{\overbrace{i_i\ldots i_r}^{r}}=\mathrm{E}\left[C_{i_i}\ldots C_{i_r}\right]$$

-continuous case- -discrete case- (1)

$$\alpha_{\boldsymbol{C}}^{\overbrace{i_i\ldots i_r}^{r}}=\int_{-\infty}^{\infty}\overset{r}{\leftrightarrow}\int_{-\infty}^{\infty}C_{i_i}\ldots C_{i_r}\cdot f_C(\boldsymbol{C})\,d\boldsymbol{C};\qquad \alpha_{\boldsymbol{C}}^{\overbrace{i_i\ldots i_r}^{r}}=\sum_{i_i=1}^{\infty}\overset{r}{\leftrightarrow}\sum_{i_r=1}^{\infty}C_{i_1}\ldots C_{i_r}\cdot f_C^*(\boldsymbol{C})$$

Superscripts denote components and not powers. These indices need not take on distinct values and there may, of course, be more than n_{rv} indices, implying repetitions. Thus, e.g., $\alpha_{\boldsymbol{C}}^{i_i i_i}$ is the mean square of $\boldsymbol{C}_{i_i}$ and $\alpha_{\boldsymbol{C}}^{i_j i_j i_j}$ is the mean cube of $\boldsymbol{C}_{i_j}$, the i_j-th component of $\boldsymbol{C}$.

In probability, it is convenient the use of the moment-generating function because it encodes the moments of a distribution in its MacLaurin expansion. This characteristic function, $M_{\boldsymbol{C}}(\xi)$, associated to a random variable $\boldsymbol{C}$ is defined as:

$$M_{\boldsymbol{C}}(\xi) = \mathrm{E}\left[e^{\xi \boldsymbol{C}} \right] \tag{2}$$

The MacLaurin expansion is as follows:

$$M_{\boldsymbol{C}}(\xi) = 1 + \xi_{i_i} \alpha_{\boldsymbol{C}}^{i_i} + \xi_{i_i} \xi_{i_j} \alpha_{\boldsymbol{C}}^{i_i i_j} / 2! + \xi_{i_i} \xi_{i_j} \xi_{i_k} \alpha_{\boldsymbol{C}}^{i_i i_j i_k} / 3! + \xi_{i_i} \xi_{i_j} \xi_{i_k} \xi_{i_l} \alpha_{\boldsymbol{C}}^{i_i i_j i_k i_l} / 4! + \ldots, \tag{3}$$

which can be assumed convergent for all $|\xi|$ sufficiently small. The moments are just the partial derivatives of $M_{\boldsymbol{C}}(\xi)$ evaluated at $\xi = 0$.

The moments of a given component of $\boldsymbol{C}$ gives information regarding the shape of the marginal distribution of that component. The cross moments measure anisotropy.

The cumulants of a distribution are a set of constants that provide an alternative to the moments to its characterization [43]. Unlike the moments, these cumulants are not directly ascertainable by summatory or integrative processes -as in (1)-. To find them it is thus necessary either to find the moments and employ relationship formulas [38,42,43] or to derive them from the cumulant-generating function that is defined as:

$$K_{\boldsymbol{C}}(\xi) = \log M_{\boldsymbol{C}}(\xi) \tag{4}$$

which has an expansion:

$$K_{\boldsymbol{C}}(\xi) = 1 + \xi_{i_i} \kappa_{\boldsymbol{C}}^{i_i} + \xi_{i_i} \xi_{i_j} \kappa_{\boldsymbol{C}}^{i_i i_j} / 2! + \xi_{i_i} \xi_{i_j} \xi_{i_k} \kappa_{\boldsymbol{C}}^{i_i i_j i_k} / 3! + \xi_{i_i} \xi_{i_j} \xi_{i_k} \xi_{i_l} \kappa_{\boldsymbol{C}}^{i_i i_j i_k i_l} / 4! + \ldots, \tag{5}$$

The cumulants in the expansion, denoted by $\kappa_{\boldsymbol{C}}^{i_i}$, $\kappa_{\boldsymbol{C}}^{i_i i_j}$, $\kappa_{\boldsymbol{C}}^{i_i i_j i_k}$ and so on, are just the partial derivatives of $K_{\boldsymbol{C}}(\xi)$ evaluated at $\xi = 0$.

Roughly speaking, mixed cumulants have an interpretation in terms of dependence or independence: univariate cumulants have a simple interpretation in terms of the shape of the marginal distribution. The first cumulant is the mean or expected value, the second

cumulant is the variance, and the third cumulant is a measure of asymmetry. Mixed cumulants could be interpreted in terms of the shape of the joint distribution. Besides, when random variables can be partitioned into different independent blocks, then all mixed cumulants involving indices from different blocks are zero.

The proposed GAT (see Section 5), in a first stage is used to concisely model and characterize input random variables as well perform load-flows in an RDS to determine its output random variables. This is accomplished by applying linearized load-flow equations, and finally reconstructing distributions through the use of approximations. For calculation purposes, the GAT must know the moments and cumulants of the random variables. Nonetheless, the GAT works better with cumulants than moments for the following reasons [38]:

- Most statistical calculations with cumulants are simpler than calculations with moments. This is the case when determining the cumulants of output variables from the cumulants of input variables subject to linear transformation [38] (Appendix C).
- The Cornish-Fisher expansion used for approximations to the distributions is best expressed with cumulants.
- When approximate normality is involved, higher-order cumulants can usually be neglected but not higher-order moments.

3. Probabilistic models

3.1. PV power model

The probabilistic PV power model in [27] is specified for 10-min intervals. Thus, the meteorological random variables involved for each nth RDS node $(n = 1,...,n_{pv})$, mth month, and xth 10-min interval (t_x) are the global and diffuse irradiation $(\boldsymbol{H}^{m}_{g,n,t_x}, \boldsymbol{H}^{m}_{d,n,t_x})$.

Global irradiation $\boldsymbol{H}^{\boldsymbol{m}}_{\boldsymbol{g,n,t_x}}$ is obtained from the variable, daily clearness index $\boldsymbol{K}^{m}_{\boldsymbol{t,n}}$ [26,27], whereas diffuse irradiation $\boldsymbol{H}^{\boldsymbol{m}}_{\boldsymbol{d,n,t_x}}$ is obtained from the variable, hourly diffuse fraction $\boldsymbol{k}^{m}_{\boldsymbol{d,n}}$ [26,27]. Both irradiations are used to build the new random variable, global irradiance on a β-tilted surface for each nth RDS node, mth month, and xth 10-min interval (t_x) $\boldsymbol{G}^{\boldsymbol{m}}_{\boldsymbol{g\_\beta,n,t_x}}$. Subsequently, the PV random power $\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,n,t_x}}$ for each xth 10-min interval (t_x), mth month, and nth RDS node $(n = 1,...,n_{pv})$ is obtained from the global irradiance on a β-tilted surface and the PV cell electrical efficiency $\boldsymbol{\xi}^{\boldsymbol{m}}_{\boldsymbol{pv,n,t_x}}$, which depends on the global irradiance. The variables $\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,n,t_x}}$ form the vector $\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,t_x}}$ $(=[\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,1,t_x}}, \boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,2,t_x}}, ..., \boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,n_{pv},t_x}}]^{\mathrm{T}})$.

The cumulant method [38] (Appendix C) is applied to obtain the cumulants of the random variable $\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,n,t_x}}$ from those of $\boldsymbol{K}^{m}_{\boldsymbol{t,n}}$ and $\boldsymbol{k}^{m}_{\boldsymbol{d,n}}$. Then, series expansion [26,27] provides the statistical characterization $(\mathrm{PDF}, \mathrm{CDF})$ of the PV random power $\boldsymbol{P}^{\boldsymbol{m}}_{\boldsymbol{pv,t_x}}$.

3.2. Load model

Currently, certain distribution network operators are involved in the massive deployment of smart meters in RDSs to measure the electrical load. Accordingly, the stochastic load model in this research is directly based on smart meter measurements over a period of several years, i.e. historical data [26,27]. This makes it possible to statistically characterize $(\mathrm{PDF}, \mathrm{CDF})$ a typical load profile, real and reactive load power, $(\boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{l,t_x}}$ $[=(\boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{l,1,t_x}}, ..., \boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{l,n_l,t_x}})^{\mathrm{T}}]$, $\boldsymbol{Q}^{\boldsymbol{d,m}}_{\boldsymbol{l,t_x}}$ $[=(\boldsymbol{Q}^{\boldsymbol{d,m}}_{\boldsymbol{l,1,t_x}}, ..., \boldsymbol{Q}^{\boldsymbol{d,m}}_{\boldsymbol{l,n_l,t_x}})^{\mathrm{T}}])$ for each xth 10-min interval (t_x), dth day of the week, mth month, and nth RDS node $(n = 1,...,n_l)$. This approach makes it possible to build an actual load profile (see [26,27]) by extracting all of the exogenous

information, namely atypical data (e.g. official holiday, important events, etc.) and weather data.

3.3. Model of EV charging load

3.3.1. Modelling of the stochastic nature of EV charging load

Three random variables are required to determine the charging demand of a single EV at any instant: (i) the charging start time (t_k) [18,19]; (ii) the initial state-of-charge (SOC) of the battery (E) [18,19]; and (iii) parking duration in which the EV battery is charged (t_k'). Moreover, this charging demand depends on the EV battery charging characteristics, which may vary, depending on battery type and charging mode [44]. Nowadays, most EVs on the market use lithium-ion batteries [45]. Their generic charging profile and related battery SOC are shown in Fig. 1 [18,46].

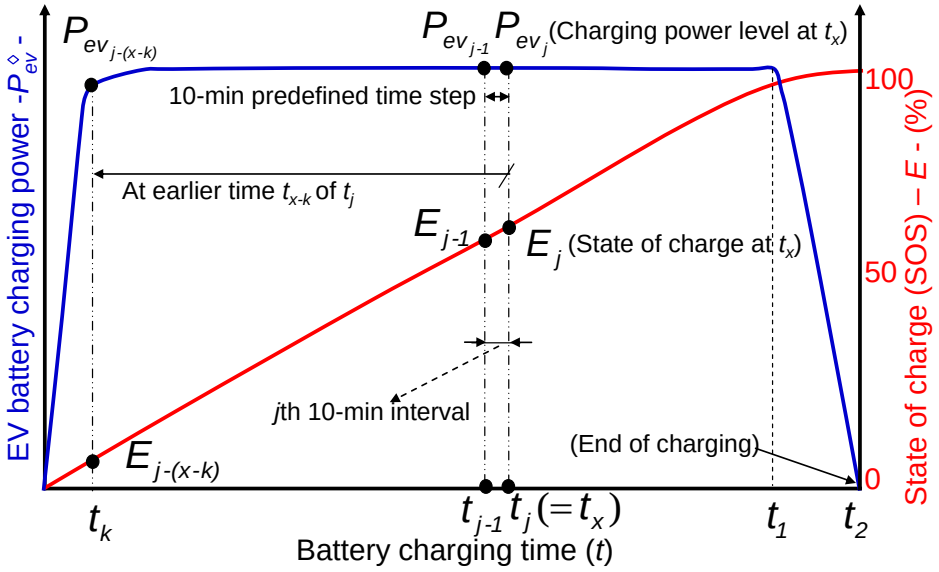

Fig. 1. Generic charging profile and SOC of a lithium-ion EV battery.

Table 1 shows the typical duration values of the charging process for an EV with a typical battery pack of 25-kWh for the four EV battery charging modes in [47].

Table 1
Parameters of EV battery charging profile.

Mode	#1	#2	#3 and #4
Continuous charging power ($P_{ev}^{\diamond}$)	3.5 kW	6.6 kW	40 kW
t_1 [a]	6.3 h	3.6 h	0.50 h
t_2 [a]	8 h	4 h	0.75 h
N_c [b]	48	24	5 [b]

[a] Times according to Fig. 1; [b] Rounded upwards because charging time is a decimal

The initial SOC of an EV battery depends on the use of the EV and can be regarded to be a random variable, depending on the distance travelled. Usually, the random variable distance ***d*** is represented by a lognormal type distribution [6,17-19] with zero probability of occurrence for all negative distances. The PDF is given by:

$$f_d(\boldsymbol{d};\mu_d,\sigma_d)=\frac{1}{\boldsymbol{d}\cdot\sigma_d\sqrt{2\pi}}e^{-\frac{(\ln\boldsymbol{d}-\mu_d)^2}{2\sigma_d^2}};\quad \boldsymbol{d}>0 \tag{6}$$

The random variable, initial SOC of EV battery ***E***, at the beginning of a recharge cycle can be expressed by:

$$\boldsymbol{E}=1-\frac{\boldsymbol{d}}{d_r};\quad 0<\boldsymbol{d}<d_r \tag{7}$$

Considering (6) and (7), the PDF for the random variable ***E*** is:

$$f_{\boldsymbol{E}}(\boldsymbol{E};\mu_d,\sigma_d)=\frac{1}{(1-\boldsymbol{E})\sigma_d\sqrt{2\pi}}e^{-\frac{[\ln(1-\boldsymbol{E})+\ln d_r-\mu_d]^2}{2\sigma_d^2}};\quad 0<\boldsymbol{E}<1 \tag{8}$$

To assist calculations, the continuous charging power $P_{ev}^{\diamond}$ shown in Fig. 1 is discretized into predefined 10-min intervals. The corresponding discrete charging power level P_{ev_j} for the *j*th 10-min interval (associated with time t_j) can thus be expressed by:

$$P_{ev_j}=\frac{1}{t_j-t_{j-1}}\int_{t_{j-1}}^{t_j}P_{ev}^{\diamond}(t)dt;\quad 1\le j\le N_c \tag{9}$$

Based on the discrete charging power level P_{ev_j}, it is possible to obtain the corresponding discrete SOC of the battery E_j before charging starts at each 10-min interval (Fig. 1). Thus, when an EV battery with a SOC E_j is charged at time t_j $(=t_x)$, its charging power is P_{ev_j}. If the charging process starts at an earlier time t_{x-k} of t_j, $(k \le x)$, then at time t_k, the SOC is $E_{j-(x-k)}$ and the charging power level is $P_{ev_{j-(x-k)}}$. The charging power P_{ev_j} at any *x*th 10-min interval (associated with time t_x) could be caused not only by the charging process starting at time t_x with a SOC E_j but also by those starting at any earlier time t_{x-k} $(k < x)$ with a lower SOC $E_{j-(x-k)}$.

Generally speaking, the random variables, charging start time $\boldsymbol{t_k}$, and initial battery SOC $\boldsymbol{E}$ are independent [18,19]. Consequently, the singleton probability φ_i that random variable, $\boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{ev_i,n,t_x}}$ the charging power of an *i*th single EV at any *d*th day of the week, *m*th month, *x*th 10-min interval (t_x), and *n*th RDS node $(n = 1,...,n_{ev})$ will operate at charging power level P_{ev_j} can be expressed as:

$$\varphi^{d,m}_{i,n,t_x}(P_{ev_j}) = \sum_{k=1}^{x} f_{\boldsymbol{E}}(\boldsymbol{E_{j-(x-k)}}) \cdot h^{d,m}_{\boldsymbol{t_k},n}(\boldsymbol{t_k}) \times \left[1 - G^{m}_{t'_k,n}(t'_k)\right]; \tag{10}$$

$$N_c \le x \le 144; \quad x\text{-}k \le j; \quad \boldsymbol{t'_k} = \boldsymbol{t_x} - \boldsymbol{t_k}; \quad x \le j$$

The term $1\text{-}G^{m}_{\boldsymbol{t'_k},n}(\boldsymbol{t'_k})$ represents the probability that the vehicle is still in the parking lot and is being charged though the charging event began at an earlier time.

From (1) and (10), the moment of order *r* of the random variable $\boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{ev_i,n,t_x}}$ can be obtained by:

$$\alpha^{\overbrace{1\ldots1}^{r}}_{\boldsymbol{P}^{d,m}_{ev_i,n,t_x}} = \sum_{j=1}^{\infty} \overset{r}{\overleftrightarrow{\ldots}} \sum_{j=1}^{\infty} P_{ev_j} \overset{r}{\overleftrightarrow{\ldots}} \cdot P_{ev_j} \cdot \varphi^{d,m}_{i,t_x,n}(P_{ev_j}) = \sum_{j=1}^{N_c} P^{r}_{ev_j} \cdot \varphi^{d,m}_{i,t_x,n}(P_{ev_j}) \tag{11}$$

However, when the number of EVs increases (random variables), according to the Central Limit Theorem, the distribution of the sum of these random variables tends to a normal distribution with *r*th-order moments, regardless of the original probability distribution. Thus, for a set λ of EVs with the same battery-charging profile, the non-crossed *r*-order moment of a normal random variable representing the total charging power $\boldsymbol{P}^{\boldsymbol{d,m}}_{\boldsymbol{tev,n,t_x}}$ is:

$$\alpha^{\overbrace{1\ldots1}^{r}}_{\boldsymbol{P}^{d,m}_{tev,n,t_x}} = \sum_{i=1}^{\lambda} \sum_{j=1}^{N_c} P^{r}_{ev_j} \cdot \varphi^{d,m}_{i,t_x,n}(P_{ev_j}) = \lambda \cdot \alpha^{\overbrace{1\ldots1}^{r}}_{\boldsymbol{P}^{d,m}_{ev_i,n,t_x}} \tag{12}$$

3.3.2. Charging demand of a single EV in households

Much research has focused on the coordinated charging of EVs [18]. Typically, the conclusion is that the electricity tariff structure determines the most economical start time for EV battery charging.

3.3.3. Parking patterns

Three random variables are used to stochastically model a realistic parking pattern that defines vehicle arrival at the parking lot (start of the charging process), parking duration in which the EV battery is charged, and finally, the distance travelled at its arrival. A representative case approach is followed since even though there are detailed vehicle mobility models [48], they usually do not include any mobility/parking pattern information.

To assess the stochastic behaviour of a large number of vehicles that can stop at a parking lot, it is assumed that all vehicles travel from a starting point (i.e. home or parking garage) with a given initial SOC and that the route finishes at the same point. This route

usually has three sequences (trips) and two stops, one of which is a long stop at a parking lot to recharge [11,14,65]. This stop may be at the end of the first or second trip. The first trip always begins at home (distance travelled L_1, km); the second trip is between the first and second stop (L_2, km); and the third trip finishes at the end point (L_3, km). Vehicles could arrive at the parking lot at the end of the first or second trip. The route assumption continues to be the same but with the added consideration that the first two trip lengths are random variables ($\boldsymbol{l_1}$ and $\boldsymbol{l_2}$) with normal distributions, and the PDFs are $f_{l_1}(\boldsymbol{l_1})$, $f_{l_2}(\boldsymbol{l_2})$ (Fig. 2). The first trip length ($\boldsymbol{l_1}$) ranges from 0 to 1 p.u. of the entire daily driving distance $\boldsymbol{d}$. Therefore, this length $\boldsymbol{L_1}$ (km) is a new random variable which is given by:

$$\boldsymbol{L_1} = \boldsymbol{d} \cdot \boldsymbol{l_1} \tag{13}$$

Consequently, the random variable, initial SOC of EV battery arriving at parking lot $\boldsymbol{E'}$, is:

$$\boldsymbol{E'} = 1 - \boldsymbol{d} \cdot \boldsymbol{l_1} / d_r; \quad 0 < \boldsymbol{d} \cdot \boldsymbol{l_1} < d_r \tag{14}$$

The second trip length ($\boldsymbol{l_2}$) ranges from 0 to 1 p.u. of the rest of the travel ($1-\boldsymbol{l_1}$). Consequently, the PDF of the second trip length is depends on the first trip. The length of the third trip is no longer a variable when the first two parts are determined ($l_3 = 1-\boldsymbol{l_1}-\boldsymbol{l_2}$).

Any correlation can be inferred between the random variables, total and partial length travelled. Therefore (14) can be linearized [49] in order to apply the cumulant method [38] (Appendix C), which can be used to obtain the cumulants [43] (statistical information) of the new random variable $\boldsymbol{E'}$.

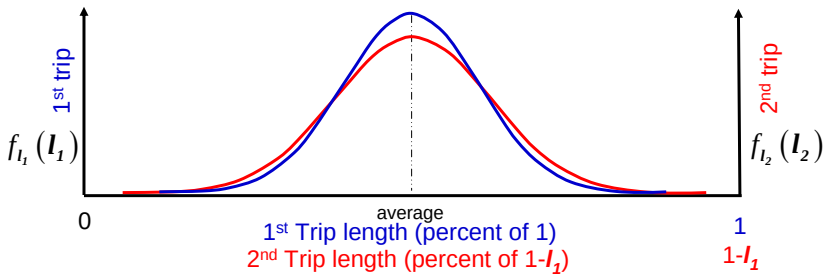

Fig. 2. PDF for the first and second trip length of an EV.

A binary variable is used to locate the long parking stop at the end of the first or second trip. When two sequences are involved, a new random variable, aggregated length (i.e. linear combination of two independent random variables) must be accounted for.

Since certain municipal/private parking lots have smart meters that record whether vehicles are entering or exiting, it was possible to calculate the number of incoming and outgoing vehicles for the entire day. This gave a statistical characterization of vehicle arrival (incoming vehicles) at the parking lot, based on measurements obtained over various years with smart meters. Accordingly, the PDF of the vehicle arrival model $h^{d,m}_{\boldsymbol{t_k},n_{pk}}(\boldsymbol{t_k})$ was modelled as a Gaussian random variable at each n_{p_k}th representative parking lot, which changes each *d*th day of the week, *m*th month, and *k*th 10-min interval (t_k). Similarly, the PDF(CDF) of the random variable $\boldsymbol{t_k'}$, parking duration at each n_{p_k}th representative parking and the *m*th month $g^m_{\boldsymbol{t_k'},n_{p_k}}(\boldsymbol{t_k'})$ ($G^m_{\boldsymbol{t_k'},n_{p_k}}(\boldsymbol{t_k'})$) was determined (obtained from the license plate documentation).

4. Characterization of the correlation among input variables

RDS input variables may or may not be correlated with each other. Generally speaking, correlation is stated through the correlation coefficient matrix $\sum_{\boldsymbol{C}}$. The *ij*th element of $\sum_{\boldsymbol{C}}$ is the correlation coefficient ρ_{C_i,C_j} [4]:[50]

$$\rho_{C_i,C_j} = \frac{\mathrm{Cov}\ (\boldsymbol{C}_i,\boldsymbol{C}_j)}{\sigma_{C_i}\sigma_{C_j}} = \frac{\mathrm{E}\left[\left(C_i-\mu_{C_i}\right)\left(C_j-\mu_{C_j}\right)\right]}{\sigma_{C_i}\sigma_{C_j}} \tag{15}$$

For the RDS nodes, the load correlation due to cyclic-human activities is reasonable [51]. PV powers are also correlated because of common factors such as weather conditions [52].

When the individual random variables of interest have different marginal distributions, a multivariate joint distribution function can be created using a copula function [53,54], which represents the correlation structure between variables. A copula is a function that joins univariate distribution functions to form multivariate distribution functions. Thus, the random variables $\boldsymbol{C}_i, (i=1,...,n_{rv})$, with CDFs $F_{C_i}(\boldsymbol{C}_i)$, are joined by copula *COP* if their joint distribution function $F_{COP} = F_{C_1,C_1,...,C_{n_{rv}}}$ can be expressed as:

$$F_{COP}(\boldsymbol{C_1},\boldsymbol{C_2},...,\boldsymbol{C}_{n_{rv}}) = F_{C_1,C_2,...,C_{n_{rv}}}[F_{C_1}^{-1}(\boldsymbol{U_1}),F_{C_2}^{-1}(\boldsymbol{U_2}),...,F_{C_{n_{rv}}}^{-1}(\boldsymbol{U}_{n_{rv}})] \tag{16}$$

4.1.1. Generation of correlated input samples using a Copula function

The method used to generate correlated inputs is based on the generation of multivariate correlated random numbers [53,54]. These numbers are used to numerically obtain both the non-crossed and crossed moments/cumulants required in the GAT, and

[4] The rank correlation measures the monotonic relationship between the variables for non-normal marginal; it should be used as a more accurate measurement of correlation [50].

also to run the MCS process that checks the results of the GAT. Summarising, the methodology in [53,54] is composed of the following steps:

Step 1) Transform random variables $\boldsymbol{C}_i$ to uniform variables using their CDFs:

$$\boldsymbol{U}_{\boldsymbol{C}_i} = F_{\boldsymbol{C}_i}(\boldsymbol{C}_i) \tag{17}$$

Step 2) Transform uniform variables $\boldsymbol{U}_{\boldsymbol{C}_i}$ to normal variables using an inverse standard normal distribution (Nataf transformation):

$$\boldsymbol{W}_{\boldsymbol{C}_i} = \phi^{-1}(\boldsymbol{U}_{\boldsymbol{C}_i}) \tag{18}$$

Step 3) Estimate the correlation matrix $\sum_{\boldsymbol{W_C}}$ of $\boldsymbol{W_C}$ from the known correlation matrix $\sum_{\boldsymbol{C}}$ of input vector $\boldsymbol{C}$ [55- 58]. The Gauss-Hermite quadrature in [59] can be used for this purpose.

Step 4) Generation of n_s correlated random input samples from a multivariate standard normal distribution $\boldsymbol{W_C}$ with a given correlation matrix $\sum_{\boldsymbol{W_C}}$, forming the arrays $W_{C_1} \in \Re^{n_{pv}+n_l, n_s}$. Each element can be written as $W_{C_{1,ij}}$.

Step 5) Transform the generated values $W_{\boldsymbol{C_{1,ij}}}$ back to the uniform domain $\boldsymbol{U}_{\boldsymbol{C_1}}$ by applying the standard normal CDF as follows [58]:

$$\boldsymbol{U}_{\boldsymbol{C_{1,i}}} = \phi(W_{\boldsymbol{C_{1,i}}}) \tag{19}$$

Step 6) Transform the generated uniform values back to the original domain $\boldsymbol{C_{1,i}}$ by applying the inverse of the respective CDF:

$$\boldsymbol{C_{1,i}} = F_{\boldsymbol{C}_i}^{-1}(\boldsymbol{U}_{\boldsymbol{C_{1,i}}}) \tag{20}$$

5. General analytical technique (GAT) to evaluate the weekly operating variables in RDSs during one year

Fig. 3 shows a flowchart of the GAT, which is explained in the following sections. The GAT calculates the worst weekly statistical characterization (PDF, CDF) of the RDS operating variables on a 10-min basis throughout the whole year. The GAT is designed as a two-stage process.

In the first stage, the GAT determines the statistical characterization (PDF, CDF) of RDS operating variables throughout the whole year by means of a loop. This loop is repeated 12,096 times, which corresponds to an assessment based on 10-min intervals for seven days of each month. In each iteration, moments and cumulants of RSD input variables for a specific 10-min interval are determined. Subsequently, a probabilistic radial load flow (PRLF) together with the Cornish-Fisher expansion is performed for this interval.

In the second stage, the GAT uses an aggregating method to determine weekly statistical characterizations for each month, based on the finite mixture distribution [60]. A subsequent inspection during all of the months reveals the worst weekly behaviour for each RDS operating variable.

5.1. Stage 1: annual assessment

The main input data are the RDS data, annual data of random variable $\boldsymbol{d}$, allocation data of EVs to nodes, and for each node, if relevant, the following: (i) monthly data of global and diffuse irradiation as well as the monthly PDF of random variable $\boldsymbol{t}_k^{'}$; (ii) daily PDF of random variable $\boldsymbol{t}_k$; (iii) historical load data for each *x*th 10-min interval.

5.1.1. Determining of moments/cumulants of RSD input variables

A. Statistical distributions (PDF or PMF and CDF) of input variables.

(i) *PDF/CDF of PV power at nodes with PV:*

The monthly data of global and diffuse irradiation make it possible to build the PV power PDF/CDF throughout the year, according to Section 3.1.

(ii) *PDF/CDF of load power at nodes with load:*

The load power PDF/CDF throughout the year is built according to Section 3.2.

(iii) *PMF/CDF of charging power at nodes with EVs:*

From the allocation of EVs to different node types, the input data of random variables $\boldsymbol{d}$, $\boldsymbol{t_k}$, $\boldsymbol{t_k'}$, make it possible to build the charging power PMF/CDF at nodes with EVs throughout the year, according to Section 3.3.

B. Treatment of the correlation of the random inputs

Correlated input samples of PV powers and node loads are generated using a copula function (see Section 4.1.1.).

C. Moments and cumulants of RSD input variables.

For correlated case studies, the above correlated samples are used to numerically obtain the non-crossed and crossed moments/cumulants up to the seventh order (see Section 2). However, distributions directly determine the moments/cumulants of the EV load power. For uncorrelated case studies, the distributions determine the moments/cumulants for all the variables.

5.1.2. Probabilistic radial load flow (PRLF)

The main idea of the PRLF in [27] is to use approximate formulas for calculating the cumulants of the RDS operating variables for each *x*th 10-min interval (t_x), *d*th

representative day of a week, and *m*th month of the year ($\kappa^{\overbrace{1\ldots1}^{r}}_{Z^{d,m}_{1,t_x}},\ldots,\kappa^{\overbrace{1\ldots1}^{r}}_{Z^{d,m}_{2\cdot n_n+n_f+1,t_x}}$) based on the cumulants of the input variables. Input cumulants include those of PV powers ($\kappa^{\overbrace{i_1\ldots i_r}^{r}}_{P^{m}_{pv,t_x}}$), node loads ($\kappa^{\overbrace{i_1\ldots i_r}^{r}}_{P/Q^{d,m}_{l,t_x}}$) and EV charging loads ($\kappa^{\overbrace{1\ldots1}^{r}}_{P^{d,m}_{ev_i/tev,t_x}}$). The RDS operating variables $\boldsymbol{Z}^{d,m}_{t_x}$ ($=[\boldsymbol{Z}^{d,m}_{1,t_x},\boldsymbol{Z}^{d,m}_{2,t_x},\ldots,\boldsymbol{Z}^{d,m}_{2\cdot n_n+n_f+1}]^{\mathrm{T}}$) include the system states $\boldsymbol{X}^{d,m}_{t_x}$ (voltage angle and magnitude at the n_n nodes) and the system outputs $\boldsymbol{X}^{*d,m}_{t_x}$ (apparent power flows in the n_f lines together with RDS total real loss).

The PRLF firstly involves the linear approximation of load-flow equations in order to get a computational efficiency as basic advantage of the proposed GAT. The detailed load-flow formulation is given in Appendix A. The method used for the linear approximation of load-flow equations is provided in Appendix B.

The cumulant method [38] (Appendix C) is then used to determine the cumulants of the RDS operating variables. Firstly, it is applied to the linearized equations (B.7) to determine the cumulants of system states $\boldsymbol{X}$ from the cumulants of input variables $\boldsymbol{C}$. This is done by solving this linear system of equations for each order of the cumulants of the input variables. Likewise, the cumulants of system outputs $\boldsymbol{X}^{*}$ are obtained from cumulants of $\boldsymbol{X}$ by applying the cumulant method to (B.8).

5.1.3. Cornish-Fisher expansion

Once the cumulants of the distributions of RDS operating variables are known, it is possible to reconstruct their PDFs and CDFs by using the Cornish-Fisher expansion [41,42] i.e., ($f_{Z^{d,m}_{1,t_x}},\ldots,f_{Z^{d,m}_{2\cdot n_n+n_f+1,t_x}}$), ($F_{Z^{d,m}_{1,t_x}},\ldots,F_{Z^{d,m}_{2\cdot n_n+n_f+1,t_x}}$). PDFs and CDFs are stored for future handling by the finite mixture distribution.

5.2. Stage 2: weekly statistical characterization by means of finite mixture distribution

When the time period that is studied (one week of each month) coincides with multiple 10-min intervals, an aggregation procedure can be applied to find the resulting distribution for this studied time period. The resulting distribution is usually obtained by the mixture [60] of the 10-min interval distributions included in the studied period. The finite mixture distribution [60] is a useful way of describing heterogeneity in the distribution of a random variable $\boldsymbol{Z}$, drawn from more than one parent population $\boldsymbol{Z_p}$ of different features. Thus, the value of the final outcome is randomly selected from the underlying values, each of which has a certain probability of selection. Given a finite set $d*q$ of distributions for the *j*th random variable at the *m*th month, $(\boldsymbol{Z}^{1,m}_{j,t_{x+1}},...,\boldsymbol{Z}^{1,m}_{j,t_{x+q}} \vdots \boldsymbol{Z}^{d,m}_{j,t_{x+1}},...,\boldsymbol{Z}^{d,m}_{j,t_{x+q}})$, coming from the *d* (seven) days and *q* (144) 10-min intervals in a day, with PDFs $(f_{\boldsymbol{Z}^{1,m}_{j,t_{x+1}}},...,f_{\boldsymbol{Z}^{1,m}_{j,t_{x+q}}} \vdots f_{\boldsymbol{Z}^{d,m}_{j,t_{x+1}}},...,f_{\boldsymbol{Z}^{d,m}_{j,t_{x+q}}})$, or the corresponding CDFs $(F_{\boldsymbol{Z}^{1,m}_{j,t_{x+1}}},...,F_{\boldsymbol{Z}^{1,m}_{j,t_{x+q}}} \vdots F_{\boldsymbol{Z}^{d,m}_{j,t_{x+1}}},...,F_{\boldsymbol{Z}^{d,m}_{j,t_{x+q}}})$, and weights $(\omega^1_1,...,\omega^1_q \vdots \omega^d_1,...,\omega^d_q)$, such that $\omega^d_q \geq 0,\ (d=1,...,7;\ q=1,...,144)$ and $\sum_{d=1}^{7}\sum_{q=1}^{144}\omega^d_q = 1$, the mixture distribution of the *j*th resulting random variable for the seven days (weekly distribution) of each *m*th month $\boldsymbol{Z}^m_j$ can be represented by either the density function , $f_{\boldsymbol{Z}^m_j}$, or the distribution function, $F_{\boldsymbol{Z}^m_j}$, as a sum:

$$
\begin{aligned}
f_{\boldsymbol{Z}^m_j} &= \sum_{d=1}^{7}\sum_{q=1}^{144}\omega^d_q f_{\boldsymbol{Z}^{d,m}_{j,t_{x+q}}}; \quad j=1,...,2\cdot n_n+n_f+1 \\
F_{\boldsymbol{Z}^m_j} &= \sum_{d=1}^{7}\sum_{q=1}^{144}\omega^d_q F_{\boldsymbol{Z}^{d,m}_{j,t_{x+q}}}; \quad j=1,...,2\cdot n_n+n_f+1
\end{aligned}
\tag{21}
$$

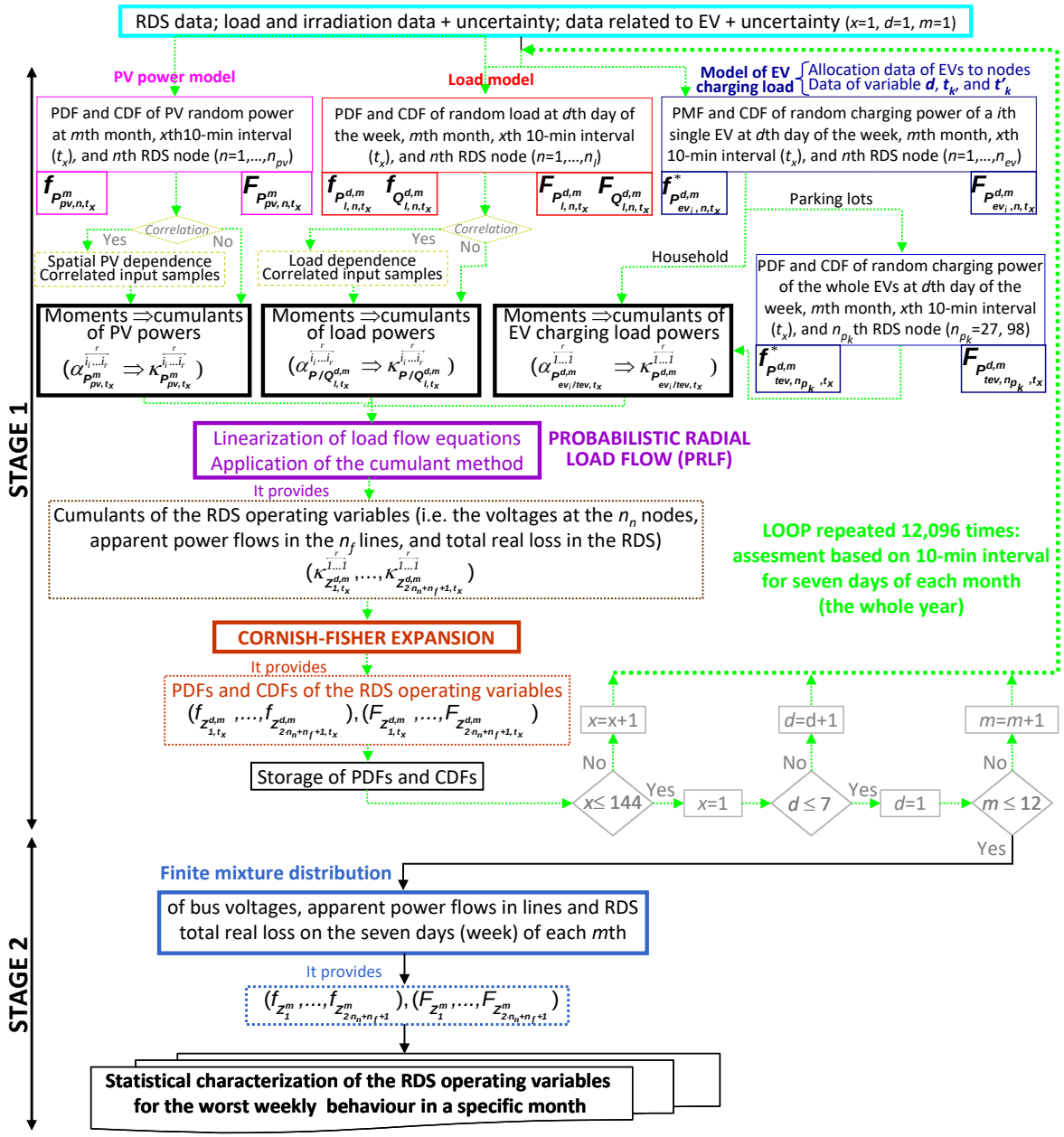

Fig. 3. Flowchart of the GAT to evaluate the worst behaviour of weekly operating variables in RDSs throughout the year.

6. Test system description

6.1. RSD, node load, and PV data

This study uses the ENDE 100 RDS, see [32] (Fig. 4). The choice of this RDS is the result of the analysis of multiple RDSs to capture the effects of their diversity [32]. This permits the specification of feeder parameters (e.g. feeder length, number of customers,

area supplied, etc.), which could explain the occurrence of technical problems (voltage and thermal) due to distributed generation. This representative RDS has 100 nodes, 99 lines, and 24 laterals located in the region of Andalusia (Spain). Its total area is over 375.3 km^2. The farthest nodes (0-97) are at a distance of 34.2 kms from each other. The RDS load is proportional (0.0059%) to that in Spain. It contains a mix of residential (55.4%), commercial (17.4%) and industrial (27.2%) customers. The mix selected is explained by the fact that the EV charging load is more likely to be spatially clustered in residential areas than in commercial or industrial centres [8].

The allocation and sizing of PV units in the ENDE 100 RDS [32] are shown in Fig. 4 (i.e. 13 PV units with a 9.52% PV penetration). These are the result of a multi-objective optimization approach previously formulated by the authors in [31,32].

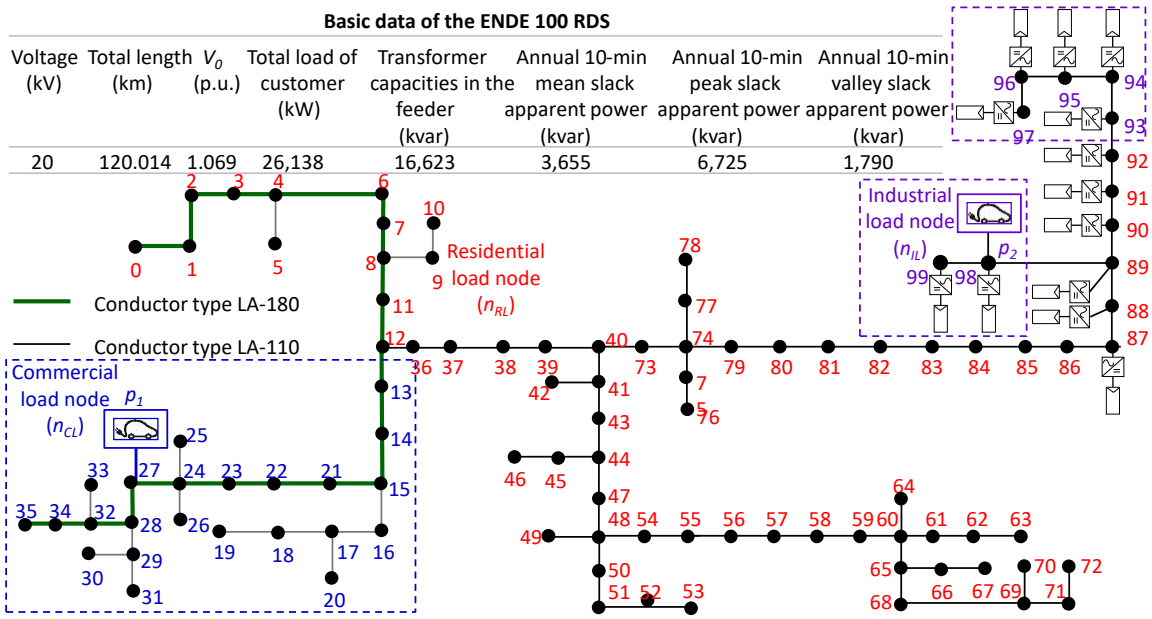

Fig. 4. Single-phase diagram of the ENDE 100 RDS with two parking lots.

Specific data of global and diffuse irradiation [61], according to the specific geolocation of each *n*th RDS node, were used to build the random variables hourly diffuse fraction and daily clearness index. For correlated case studies, the spatial PV

dependence structure was modelled from a spatial PV correlation study in the area surrounding the ENDE 100 RDS (at nodes 18, 19, 64, 86, and 96). This permitted the estimation of the correlation coefficients $\rho_{P_{pv,i,10\text{-}min},P_{pv,j,10\text{-}min}}$ by extrapolating the actual distance between PV units:

Nodes	87	88	89	90	91	92	93	94	95	96	97	98	99
87	1	0.951	0.917	0.894	0.860	0.815	0.769	0.724	0.701	0.633	0.610	0.894	0.871
88	0.951	1	0.962	0.940	0.906	0.860	0.815	0.769	0.746	0.678	0.656	0.940	0.917
89	0.917	0.962	1	0974	0.940	0.894	0.849	0.803	0.781	0.712	0.690	0.974	0.951
90	0.894	0.940	0.974	1	0.962	0.917	0.871	0.826	0.803	0.735	0.712	0.951	0.928
91	0.860	0.906	0.940	0.962	1	0.951	0.906	0.860	0.837	0.769	0.746	0.917	0.894
92	0.815	0.860	0.894	0.917	0.951	1	0.951	0.906	0.883	0.815	0.792	0.871	0.849
93	0.796	0.815	0.849	0.871	0.906	0.951	1	0.951	0.928	0.860	0.837	0.826	0.803
94	0.724	0.769	0.803	0.826	0.860	0.906	0.951	1	0.974	0.906	0.883	0.781	0.758
95	0.701	0.746	0.781	0.803	0.837	0.883	0.928	0.974	1	0.928	0.906	0.758	0.735
96	0.633	0.678	0.712	0.735	0.769	0.815	0.860	0.906	0.928	1	0.974	0.690	0.667
97	0.610	0.656	0.690	0.712	0.746	0.792	0.837	0.883	0.906	0.974	1	0.667	0.644
98	0.894	0.940	0.974	0.951	0.917	0.871	0.826	0.871	0.758	0.690	0.667	1	0.974
99	0.871	0.917	0.951	0.928	0.894	0.849	0.803	0.758	0.735	0.667	0.644	0.974	1

Data for real and reactive load profiles were collected at each node in 10-min intervals by using smart meters over a five-year time period. For correlated case studies, the load correlation matrix [62,63] was obtained from synchronous 10-min measurements.

6.2. EV data

Based on the expected daily driving distance by private/company/fleet vehicles in Spain [64,65], the annual mean and standard deviation of the random variable distance ***d*** considered were equal to 35 and 9.6 km, respectively (70 and 19.2 km / 58.5 and 16.4 km). Furthermore, 7.5% were assumed to be company vehicles and 1.95% fleet vehicles [64,66].

Previous research either considered a single type of EV in the analysis or used several types (e.g. cars, vans, and sports utility vehicles [8,67]) with assumed market-share percentages and battery capacities. For simplicity, this study only considered one type. Thus, all EVs were assumed to have a typical battery pack of 25-kWh and an all-electric range of up to 128.7 km (80 mi), i.e. its typical maximum distance d_r.

The number of EVs distributed at each RDS node depended on the node type: (i) residential load node (n_{rl}); (ii) commercial and industrial load node (n_{cl}, n_{il}). In the first node type, this number was based on the average Spanish household power demand (approximately 0.897 kW/year [64,68]) and the average number of vehicles per household (1.87 in 2014 in Spain [64]). These two statistics made it possible to approximately calculate the number of EVs in residential area nodes from the annual average active power of each node. The number of EVs was added to the residential area nodes. In the commercial and industrial area, two parking lots were located at node #27 (n_{p_1}) and #98 (n_{p_2}) (see Fig. 4). The resulting number of EVs was proportional (0.0057%) to the number of EVs in Spain.

As observed in EV technology roadmap studies [67], there is a high level of uncertainty regarding how fast this technology will penetrate the market. If penetration is moderate, the EPRI [69] states that by 2020, 30% of the total number of vehicles will be EVs. This research was thus based on a 30%-penetration level.

The modelling of the charging start time in households, i.e., PDF for each n_{rl}th residential load node $h^{d,m}_{\boldsymbol{t_k},n_{rl}}(\boldsymbol{t_k})$ was roughly bell-shaped. Accordingly, the resulting normal distributions for the Spanish pattern had a mean in the interval 0:00-1:30 a.m. and a standard deviation of 4.1-6.2 h depending on that of the *d*th day and *m*th month of the year [18,64,66]. On the other hand, the PDF of the random variable, parking duration at home $g^{m}_{\boldsymbol{t_k}',n_{rl}}(\boldsymbol{t_k}')$ was also determined from data for Spain [64,66].

To model vehicle arrival at parking lots, two municipal parking lots located in the ENDE 100 RDS area were selected. They are representative of two parking lot types in Spain [65]: (i) p_1 was located in a commercial area (325 parking spaces, turnover rate of

5.6); (ii) p_2 was located in an industrial area (275 parking spaces, turnover rate of 1.8). Survey data for vehicle arrival were collected at each parking lot in 10-min intervals by using smart meters over a four-year time period. As an example, the mean PDF value of the vehicle arrival model for a representative April Monday $h^{1,4}_{t_k,n_{pk}}(t_k)$ is shown in Figs. 5a and 5b. The PDF of the random variable, parking duration $g^{4}_{t_k,n_{pk}}(t_k')$, is also shown.

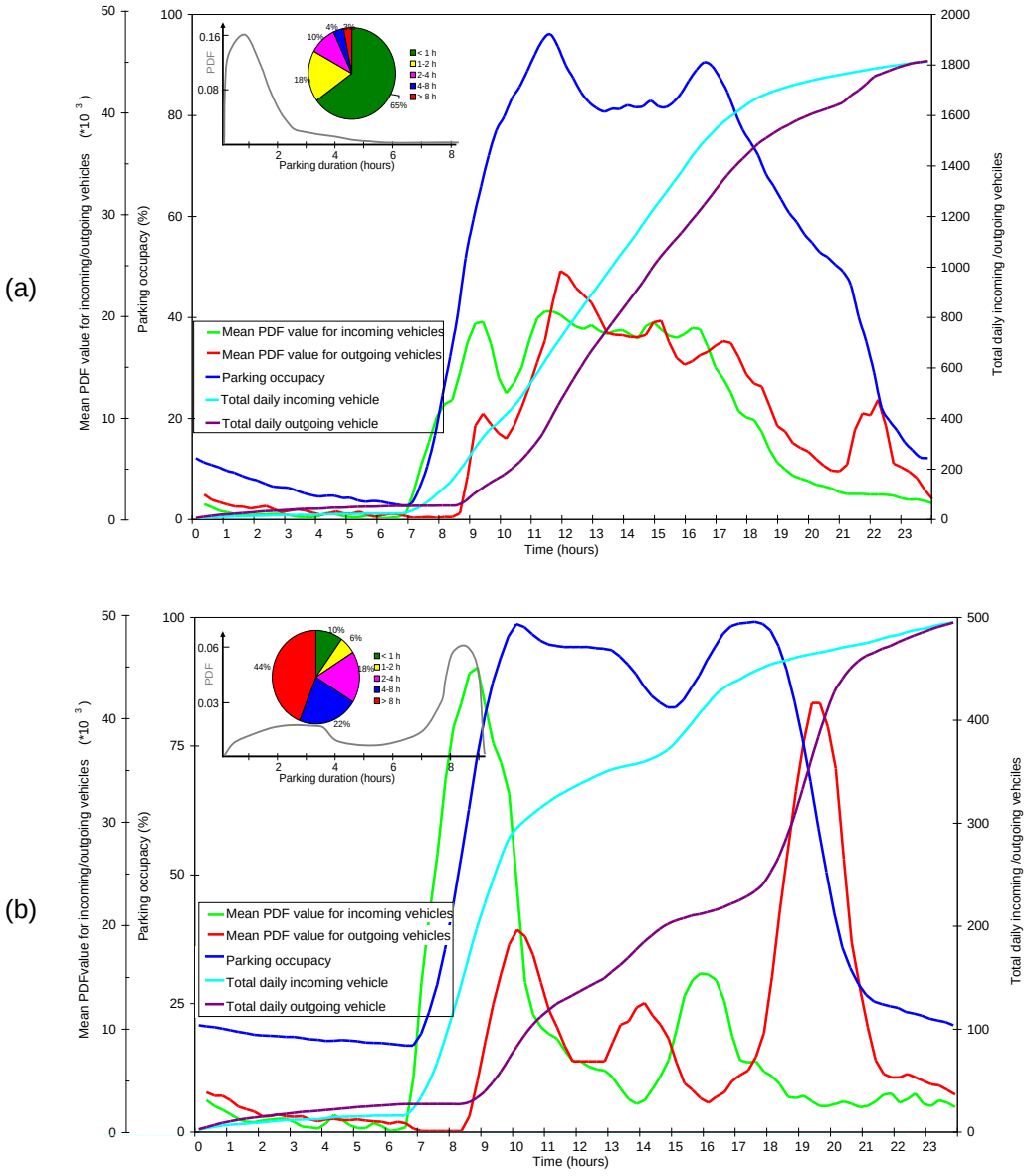

Fig. 5. Parking survey results on a representative April Monday: (a) Parking lot in a commercial area; (b) Parking lot in an industrial area.

The arrival patterns in Figs. 5a and 5b were clearly different as a consequence of driving habits. In Fig. 5a, vehicle arrival was fairly stable during the morning and early afternoon, whereas in Fig. 5b, most of the arrivals occurred in the early morning with a lower peak in the early afternoon. Another important difference was the parking duration. In Fig. 5a, most parking had a duration of less than 1 hour (65%); only 7% of parking was over 4 hours. Parking duration was longer in Fig. 5b (44% of parking was over 8 hours). This indicated that many employees park here all day.

7. Results

The proposed GAT was implemented in MATLAB and tested in the ENDE 100 RDS. The results focused on node voltage 97 (the most critical node for voltage regulation), apparent power flow in the line 0-1 (the most stressed line), and the total real loss. Apparent power flow was normalized by the maximum capacity in the line and the total real loss by the mean annual slack apparent power.

Since our main objective was to improve the assessment accuracy of the EVCSs and PV interaction in RDSs within the context of different customers and correlation of the input variables, various case studies were assessed. More specifically, seven case studies without correlation were studied: (i) #1 base case (without EVs, without PV); (ii) #2 (without EVs, with PV); (iii) #3 (with EVs, without PV); (iv) #4 (with EVs and PV); (v) #5 (with EVs at residential load nodes, without PV); (vi) #6 (with EVs at industrial load nodes, without PV); (vii) #7 (with EVs at commercial load nodes, without PV). Furthermore, for the sake of completeness and because of space limitations, three case studies with the correlation of input variables were evaluated: (i)-#8-(#2-with correlation);-(ii)-#9-(#3-with correlation);-(iii)-#10-(#4-with correlation).

7.1. Proposed GAT versus MCS

The reference method for assessing the accuracy of the analytical techniques is the mainstream MCS [27]. In MCS pseudo-random samples are produced from the stochastic input variable model underlying probability distributions. These samples are then propagated by means of the deterministic system model, which is used in our research to solve a deterministic load flow, with a view to obtaining samples for the output variables of interest. After a certain number of deterministic simulations, the probability distribution of stochastic outputs is directly calculated on the basis of the deterministic samples obtained for each simulation. It should be underlined that the number of simulations needed to obtain an accurate result is independent of system size [70]. Nonetheless, for an accurate representation, many simulations (approximately 10,000 [27]) must be considered. The basic computational part of MCS is deterministic, and there is no need to simplify the non-linear mathematical models to ensure applicability.

The first stage of the proposed GAT includes three error sources compared to the MCS for each 10-min interval. These error sources were the following: (i) the use of a higher or lower number of cumulants to characterize the input variables, i.e., more orders of cumulants or fewer; (ii) the application of the linear approximation of the non-linear load-flow equations; (iii) the use of a shorter or longer Cornish-Fisher expansion with more cumulants or fewer. Although more cumulants produce better results in (i) and (iii), practically speaking, numerical cumulants higher than the fifth-order are rarely required [27]. In our study, the seventh order was the upper limit.

The second stage of the GAT is the same as that of the MCS because it works with the PDFs and CDFs obtained and stored for each 10-min interval. Therefore, it does not add another error source.

The proposed GAT and MCS can only be compared on the basis of their results. The accuracy assessment of our GAT, compared with MCS, took into account the three previously mentioned error sources. This accuracy of the GAT was carried out by means of the individual relative error of the first seven moments for each RDS operating variable. In particular, this error in the voltage magnitude at any *n*th RDS node and *m*th month is given by [17]:

$$\varepsilon_{\alpha_{V_n^m}^{\overrightarrow{1\ldots1}}} = \frac{100 \cdot \left| \alpha_{V_n^m}^{\overrightarrow{1\ldots1},\,\mathbf{GAT}} - \alpha_{V_n^m}^{\overrightarrow{1\ldots1},\,\mathbf{MCS}} \right|}{\alpha_{V_n^m}^{\overrightarrow{1\ldots1},\,\mathbf{MCS}}} \tag{22}$$

The values in Table 2 show the high accuracy level of the GAT for the expected values and standard deviations of the most critical node (node 97) under the assumption of several case studies. The higher moments were nearly accurate. Only the largest error during the 12 months is shown.

Table 2

Maximum individual relative error throughout the months corresponding to the weekly evaluation for the first seven central moments of the 97-node voltage magnitude.

Case studies / Error (%)	$\varepsilon_{\alpha^{1}_{V_{97}^m}}$	$\varepsilon_{\alpha^{11}_{V_{97}^m}}$	$\varepsilon_{\alpha^{111}_{V_{97}^m}}$	$\varepsilon_{\alpha^{1111}_{V_{97}^m}}$	$\varepsilon_{\alpha^{11111}_{V_{97}^m}}$	$\varepsilon_{\alpha^{111111}_{V_{97}^m}}$	$\varepsilon_{\alpha^{1111111}_{V_{97}^m}}$
#8	0.140	0.438	0.627	0.704	0.664	0.520	0.553
#9	0.171	0.545	0.782	0.849	0.730	0.682	0.693
#10	0.193	0.606	0.853	0.973	0.796	0.739	0.771

Table 3 shows the computation time, which highlights the improvement achieved by the more efficient GAT in relation to the MCS. The computer used was an HP Intel® Pentium® dual CPU with 1.60 GHz and 2 GB RAM. The MCS was found to be computationally expensive, and not viable for the required annual assessment. It should be highlighted that the GAT covers a one-year period, with a weekly assessment on a 10-min basis for each month. This means that 12,096 intervals were evaluated in the first stage of the GAT. The MCS evaluated the same number of intervals but each interval comprised 10,000 trials, i.e., 1.209e8 deterministic simulations.

Table 3
Computation times for accurate annual assessment.

Case studies		#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Computation time	GAT (h)	0.03	0.88	0.98	1.29	1.21	1.11	1.10	0.97	1.08	1.43
	MCS (day)	1.29	38.45	41.78	47.52	46.59	44.56	46.04	39.60	43.11	48.01

However, the proposed GAT has certain disadvantages with regard to MCS. It requires extra effort because of programming tasks such as: (i) the computation of the moments and subsequently of the cumulants of input variables; (ii) the computation of the PFDs and CDFs of RDS operating variables by means of the Cornish-Fisher expansion.

7.2. Assessment of the daily technical impact

As a representative example of a particular day and month, Fig. 6 shows the PDFs of the 10-min voltage of node 97 during a July Monday $\boldsymbol{V}_{97,10\text{-}min}^{1,7}$ at case study #10. This represents 144 PDFs for a 24-h duration. It was found that PV support did not help to increase the lowest voltage occurring around midnight. This was because the EV charging demand of residential customers was distributed intensively overnight (from 12:30 to 5:30 a.m.) and also because its node load peak was distributed around 10:00 p.m. In contrast, around 3:00 p.m., the node voltage PDFs reached their maximum value because of PV support and medium load conditions as well as the medium EV charging demand at parking lots.

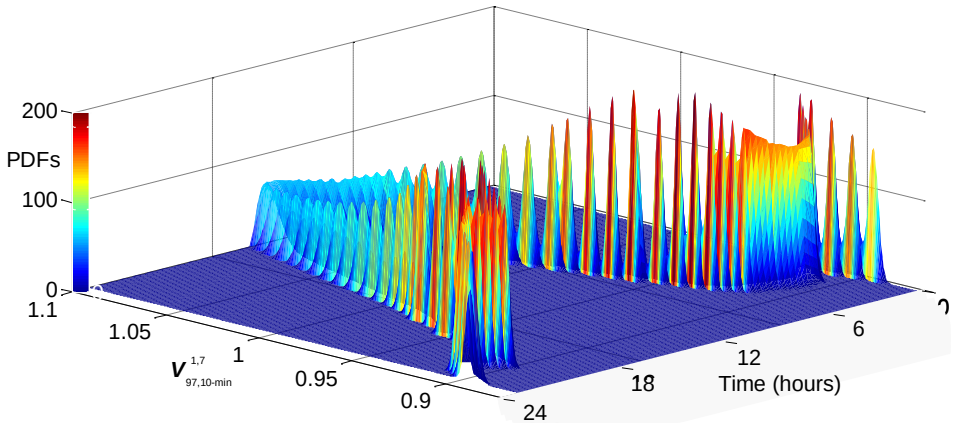

Fig. 6. PDFs of the 10-min voltage magnitude of node 97 during a representative July Monday at case study #10.

7.3. Assessment of the weekly technical impact throughout the year

Figs. 7 and 8 show the comparative assessments of the weekly technical impact of EVCSs and PV interaction on the ENDE 100 RDS in January, with minimum PV power, when voltages reached their lowest values, and also in July with maximum PV power.

7.3.1. Impact on node voltage magnitudes

Fig. 7 compares PDFs in January for the 10-min voltage of node 97 under correlated and uncorrelated case studies. In regards to a single PV impact (#2), the incorrect adjustment of the PV output/load profile was observed in the voltage enhancement with PV units. The voltage was only enhanced during hours with medium voltage levels but not for the worst voltage levels (i.e. approximately 0.95 p.u. in the evening).

Regarding the impact of EVs, the worst case study was #5 (residential customers with EVs), where the impact was highest. A new peak load resulting in lower voltage occurred in the residential area when the smart strategy in [18] was applied. This was the result of the higher percentage of residential customers in the customer mix of the RDS analysed, together with the smart load strategy which concentrated the EV charging demand in a shorter period, as compared to the long period when EVs in parking lots were charging (#6 and #7).

The impact of the EV charging demand on industrial and commercial area is shown at case studies #6 and #7, respectively. More specifically, the parking lot in the industrial area had a higher impact (#6) than the parking lot in the commercial area (#7). This was because the industrial parking lot was located at a feeder end, closer to the monitored node and represented a higher load percentage. As discussed earlier, the addition of the

charging load from parking lots was comfortably absorbed by the RDS without an increase in the peak demand or an important variation in the voltage distribution at node 97.

The PV and EV interaction (#4) increased the dispersion of the PDF as each one introduced a new uncertainty.

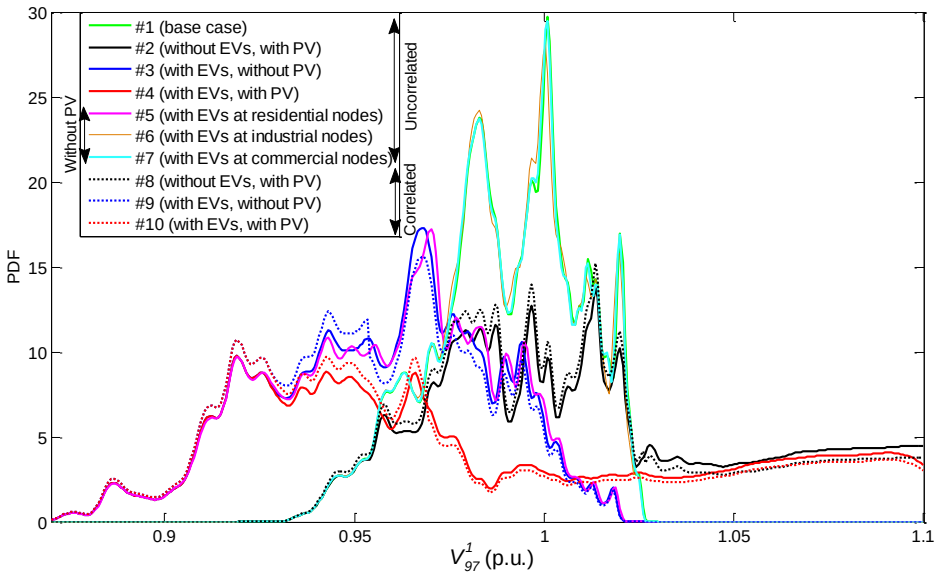

Fig. 7. PDFs for the 10-min voltage magnitude of node 97, corresponding to weekly evaluation in January in different case studies.

The correlation of inputs (#8,9,10) led to lower voltage dispersions compared to those for uncorrelated case studies (#2,3,4), increasing the probability of medium voltage levels. When input variables were uncorrelated, each one introduced a new uncertainty, which increased the dispersion of the resulting PDF.

When January and July were compared (Fig. 8a), the single effect of PV power on the resulting CDF was highest in July. In contrast, the single effect of EVCSs on the CDF was lowest in July.

As can be observed in Fig. 8a, with the 30%-penetration level set, the voltage constraints in [34] were almost met. In July, the lower dispersion of the resulting CDF under the combined effect made the violation of voltage constraints less probable. In fact, for the worst node (node 97) and month (January, 1), the probability of exceeding limits

was less than 7.12%. In July, this probability was less than 0.96%. As a conclusion, the GAT provided the worst case study and index distributions.

7.3.2. Impact on apparent power flow in lines

Fig. 8b reveals that the PV units (#2) led to a shift in the distribution of the 10-min apparent power flow in the line 0-1 towards lower load levels compared to the base case (#1). The opposite outcome was observed because of EVCSs (#3). Thus, the congestion level in the line reached a high level. The combined effect (#4) increased the dispersion of the PDF from each single dispersion. PV and EVCS interaction could mitigate certain overcurrent problems.

Because of correlation between inputs (#10 vs. #4), the synchronisation (the simultaneous increase and decrease) of different power injections was stronger. This caused the probabilities of the low power output section and high power output section of all correlated power injections to become larger. This led to an increase in the power flow variability, which had to be accounted for in the system design. Furthermore, there was also a slight increase in the expected value of the distribution.

7.3.3. Impact on total real loss

Fig. 8c shows the weekly comparison of the PDFs of total 10-min real loss in January and July. As expected, the ENDE 100 RDS was more likely to operate at lower real loss conditions with the interaction of PV and EVCSs (#4) as compared to the single impact of EVCSs (#3). Moreover, the resulting distribution approached a flattened distribution. It should be highlighted that although the single effect of PV power or EVCSs was quite different in January and July, the combined effect originated PDFs of total real loss that were fairly similar in both months. The dispersion of the resulting-PDF increased in correlated case study (#10) vs. uncorrelated case study (#10). Nonetheless, its distribution became more symmetrical (skewness was closer to zero).

(a)

(b)

(c)

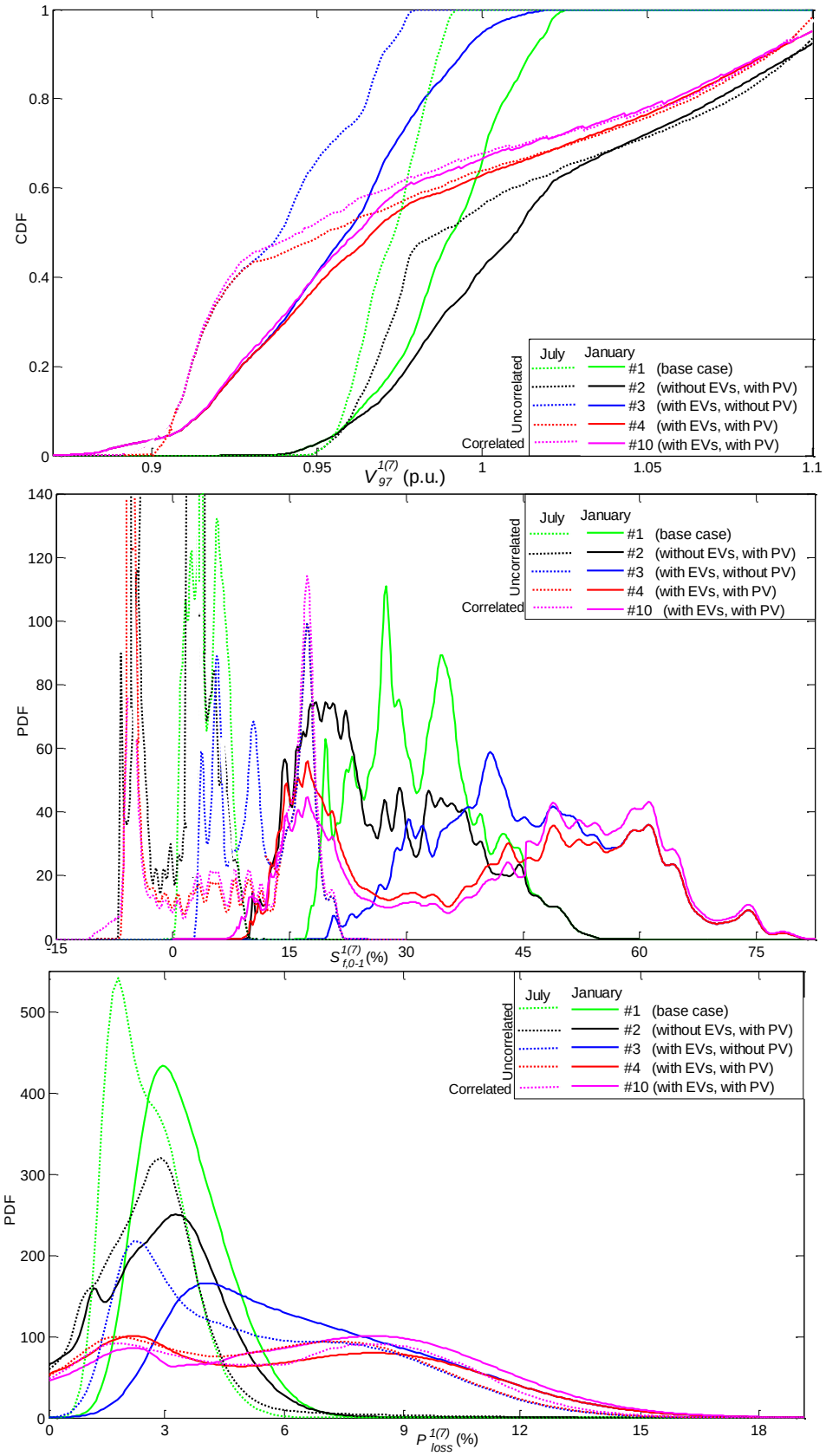

Fig. 8. Weekly evaluation in January/July in different case studies: (a) CDFs for the 10-min voltage of node 97; (b) PDFs of 10-min apparent power flow in line 0-1; (c) PDFs of total 10-min real loss in the ENDE 100 RDS.

8. Conclusion

This paper has presented an innovative GAT, which was tested on an operating ENDE 100 RDS [32]. This new technique effectively assessed the stochastic behaviour of the overall system with loads (node and EV) and PV generation during a one-year period under correlated and uncorrelated case studies. Assessment for the whole year was performed for 10-min intervals to conveniently capture the total uncertainty and determine the improvement of the RDS operating variables. Thanks to this new GAT, the computational burden of the assessment was affordable.

This assessment took into account the typical parameters required for modelling the EV charging load at home/parking lots as well as the node load and PV generation throughout a one-year period. Models of node load, PV power, and EVCS load at home and parking lots were built. Statistical data were obtained from a representative RDS in Spain. This study demonstrated that the combined impact of PV and EVCS on an RDS considerably differed from the individual impact of PV or EVCSs and that generally it improved. Thus, although the variability of the analysed RDS operating variables increased, this burden was compensated by a lower probability that such outputs would be closer to regulatory limits. The results also indicated that the correlation between inputs had a major impact on the RDS operating variables, mainly on power flow and RDS total real loss.

Appendix A. Detailed load-flow formulation

The exact equations of the non-linear load-flow for a power system with PV generation and EV charging loads can be mathematically described by [71]:

$$\begin{aligned} P_{g,n} + P_{pv,n} - P_{l,n} - P_{ev_i/tev,n} &= V_n \sum_{j=1}^{n_n} \left[V_j \left(G_{nj} \cos\delta_{nj} + B_{nj} sin\delta_{nj} \right) \right]; \\ Q_{g,n} - Q_{l,n} &= V_n \sum_{j=1}^{n_n} \left[V_j \left(G_{nj} sin\delta_{nj} - B_{nj} \cos\delta_{nj} \right) \right]; \qquad n = 1,...,n_n \end{aligned} \tag{A.1}$$

Power systems are innately uncertain systems. Accordingly, the main goal of the PLF is to determine the system states $\boldsymbol{X}$ $(=[\boldsymbol{V}_1, \delta_1, \boldsymbol{V}_2, \delta_2, , ..., \boldsymbol{V}_j, \delta_j]^{\mathrm{T}}; {}_{j=1,...,n_n})$ and then, the system outputs $\boldsymbol{X}^*$ $(=[\boldsymbol{S}_{f,i_1\text{-}j_1}, \boldsymbol{S}_{f,i_2\text{-}j_2}, ..., \boldsymbol{S}_{f,i_k\text{-}j_k}, \boldsymbol{P}_{loss}]^{\mathrm{T}}; {}_{k=1,...,n_f})$, as a function of the input random variables $\boldsymbol{C}$ $(=[\boldsymbol{P}_{pv}\ \boldsymbol{P}_l\ \boldsymbol{Q}_l\ \boldsymbol{P}_{ev_i/tev}]^{\mathrm{T}})$. Therefore, (A.1) can be written in a general form:

$$\boldsymbol{C} = g_{PLF}(\boldsymbol{X}) \tag{A.2}$$

The system outputs could be expressed as:

$$\boldsymbol{X}^* = h_{PLF}(\boldsymbol{X}) \tag{A.3}$$

The algorithm in [72] is used here to solve the deterministic load flow in (A.1) since the Newton-Raphson algorithm causes some convergence problems in RDSs. This algorithm provides the expected values of system states $\mu_{\boldsymbol{X}}$ from the expected input values $\mu_{\boldsymbol{C}}$.

Appendix B. Linear approximation of load-flow equations

The linear approximation of the load-flow equations (A.1) or (A.2) is used to obtain the system states $\boldsymbol{X}$ as a linear combination (weighted sum) of system inputs $\boldsymbol{C}$. The linearization process is made around the solution point of the deterministic load flow in [72], namely, the expected values of the system state $\mu_{\boldsymbol{X}}$.

To illustrate this technique, let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two random variables which, at some stage of the computations, are multiplied to give a third random variable $(\boldsymbol{Y} = \boldsymbol{A} \cdot \boldsymbol{B})$. If the deviations of $\boldsymbol{A}$ and $\boldsymbol{B}$ around their expected values $(\mu_{\boldsymbol{A}},\ \mu_{\boldsymbol{B}})$, they are $\Delta\boldsymbol{A}$ and $\Delta\boldsymbol{B}$, respectively. The following can thus assumed:

$$\boldsymbol{A} \approx \mu_{\boldsymbol{A}} + \Delta\boldsymbol{A}; \quad \boldsymbol{B} \approx \mu_{\boldsymbol{B}} + \Delta\boldsymbol{B} \tag{B.1}$$

When second-order terms are neglected $(\Delta\boldsymbol{A} \cdot \Delta\boldsymbol{B})$, the following expression is obtained:

$$\boldsymbol{Y} \approx \mu_{\boldsymbol{B}} \cdot \Delta\boldsymbol{A} + \mu_{\boldsymbol{A}} \cdot \Delta\boldsymbol{B} + \mu_{\boldsymbol{A}}\mu_{\boldsymbol{B}} = \mu_{\boldsymbol{B}} \cdot \boldsymbol{A} + \mu_{\boldsymbol{A}} \cdot \boldsymbol{B} - \mu_{\boldsymbol{A}}\mu_{\boldsymbol{B}} \tag{B.2}$$

This approximation is accurate for cases where the dispersion of the random variables is limited around the mean value. When this is not the case, the variable data will be transformed less accurately.

This technique can be applied to voltage magnitudes and angles in (A.1). For the product of voltage magnitudes is as follows:

$$\begin{aligned} V_n V_j &\approx \mu_{V_n} V_j + \mu_{V_j} V_n - \mu_{V_n}\mu_{V_j} \\ V_n^2 &\approx 2\mu_{V_n} V_j - \mu_{V_n}^2 \end{aligned} \tag{B.3}$$

For angles δ_{nj} , Maclaurin's series are firstly used in *sin* and *cos* functions, and then the linear approximation is applied [70]:

$$\begin{aligned} &sin\delta_{nj} \approx \delta_{nj} - \frac{\delta_{nj}^3}{6}; \quad \rightarrow \quad \delta_{nj} - \frac{\delta_{nj}^3}{6} = a_{nj} + b_{nj}\delta_{nj}; \qquad (e.g.\ a_{nj} = 1 + \mu_{\delta_{nj}}^2/2) \\ &cos\delta_{nj} \approx 1 - \frac{\delta_{nj}^2}{2}; \quad \rightarrow \quad 1 - \frac{\delta_{nj}^2}{2} = c_{nj} + d_{nj}\delta_{nj} \end{aligned} \tag{B.4}$$

Therefore, it can be assumed that:

$$\begin{aligned} V_n V_j sin\delta_{nj} &\approx a'_{nj} + b'_{nj}\delta_{nj} + c'_{nj}V_n + d'_{nj}V_j \qquad (e.g.\ a'_{nj} = 2\mu_{V_n}\mu_{V_j}\mu_{\delta_{nj}}(-1 + \mu_{\delta_{nj}}^2/3)) \\ V_n V_j \cos\delta_{nj} &\approx a''_{nj} + b''_{nj}\delta_{nj} + c''_{nj}V_n + d''_{nj}V_j \end{aligned} \tag{B.5}$$

When these approximations are substituted in (A.1), the following ca be obtained:

$$\begin{aligned} P_{g,n} + P_{pv,n} - P_{l,n} - P_{ev_i/tev,n} &= \sum_{j=1}^{n_n}\left(e'_{nj} + f'_{nj}\delta_n - f'_{nj}\delta_j + g'_{nj}V_n + h'_{nj}V_j\right) \\ Q_{g,n} - Q_{l,n} &= \sum_{j=1}^{n_n}\left(e''_{nj} + f''_{nj}\delta_n - f''_{nj}\delta_j + g''_{nj}V_n + h''_{nj}V_j\right) \end{aligned} \tag{B.6}$$

Coefficients e'_{nj}, e''_{nj}, f'_{nj}, f''_{nj}, g'_{nj}, g''_{nj}, h'_{nj}, and h''_{nj} are computed from-RDS parameters and the expected values of the system states $\mu_{\boldsymbol{X}}$ [70]. Therefore, (B.6), once linearized, is expressed as:

$$\boldsymbol{C} = g_{PLF}^{\Delta}(\boldsymbol{X}) \tag{B.7}$$

Likewise, the linear approximation of (A.3) is expressed as [70]:

$$\boldsymbol{X}^{*} = h_{PLF}^{\Delta}(\boldsymbol{X}) \tag{B.8}$$

Appendix C. Cumulant method

This appendix aims to examine how the cumulant arrays $\kappa_{\boldsymbol{C}}^{i}$, $\kappa_{\boldsymbol{C}}^{i_i i_j}$, $\kappa_{\boldsymbol{C}}^{i_i i_j i_k}$,… of a random variable $\boldsymbol{C}$ change when we make a simple transformation from the original variables $\boldsymbol{C}_1, \boldsymbol{C}_2, ..., \boldsymbol{C}_{n_{rv}}$ to new a variable $\boldsymbol{Z}$. Let $\boldsymbol{Z}$ be a random variable that is a linear combination of $\boldsymbol{C}$ given by:

$$\boldsymbol{Z} = \sum_{i_i=1}^{n_{rv}} a_{i_i} \cdot \boldsymbol{C}_{\boldsymbol{i_i}} \tag{C.1}$$

Applying the cumulant method [38], for example, the third-order cumulant of the variable $\boldsymbol{Z}$ can be expressed as a function of the cumulants of the variable $\boldsymbol{C}$ as:

$$\kappa_{\boldsymbol{Z}}^{111} = \sum_{i_i=1}^{n_{rv}} \sum_{i_j=1}^{n_{rv}} \sum_{i_k=1}^{n_{rv}} a_{i_i} a_{i_j} a_{i_k} \kappa_{\boldsymbol{C}}^{i_i i_j i_k} \tag{C.2}$$

The cumulant method thus replaces the convolution of the random variables by the linear combination of their cumulants. In general, the *r*-order cumulant could be obtained as [38]:

$$\kappa_{\boldsymbol{Z}}^{\overbrace{1...1}^{r}} = \sum_{i_i=1}^{n_{rv}} \overleftrightarrow{...}^{r} \sum_{i_r=1}^{n_{rv}} a_{i_i} \overleftrightarrow{...}^{r} a_{i_r} \cdot \kappa_{\boldsymbol{C}}^{\overbrace{i_i...i_r}^{r}} \tag{C.3}$$

References

[1] Van Vliet O, Brouwer AS, Kuramochi T, Van den Broek M, Faaij A. Energy use, cost and CO_2 emissions of electric cars. J Power Sources 2011;196(4):2298–2310.

[2] Tarroja B, Eichman JD, Zhang L, Brown TM, Samuelsen S. The effectiveness of PHEVs and renewable power in support of holistic environmental goals : Part 1-Evaluation of aggregate energy and greenhouse gas performance. J Power Sources 2014;257:461–470.

[3] Haidar AMA, Muttaqi KM, Sutanto D. Technical challenges for electric power industries due to grid-integrated electric vehicles in low voltage distributions: A review. Energy Convers Manage 2014;86:689–700.

[4] Sanchez-Sutil F, Hernandez JC, Tobajas C. Overview of electrical protection requirements for integration of a smart DC node with bidirectional electric vehicle charging stations into existing AC and DC railway grids. Electr Power Syst Res 2015;122:104–118.

[5] Sexauer JM, McBee KD, Bloch KA. Applications of probability model to analyze the effects of electric vehicle chargers on distribution transformers. IEEE Trans Power Syst 2013;28(2):847–854.

[6] Shaaban MF, Atwa YM, El-Saadany EF. PEVs modeling and impacts mitigation in distribution networks. IEEE Trans Power Syst 2013;28(2):1122–1131.

[7] Papadopoulos P, Skarvelis S, et al. Electric vehicles' impact on British distribution networks. IET Electr Syst Transp 2012;2(3)91–102.

[8] ElNozahy MS, Salama MMA. A comprehensive study of the impacts of PHEVS on residential distribution networks. IEEE Trans Sustain Energ 2014;5(1):332–342.

[9] Valsera-Naranjo E, Sumper A, Villafafila-Robles R, Martinez-Vicente D. Probabilistic method to assess the impact of charging of EV on distribution grids. Energies 2012;5:1503–1531.

[10] Aljanad A, Mohamed A. Impact of plug-in hybrid electric vehicle on power distribution system considering vehicle to grid technology: a review. Res J App Sciences Engineering and Technology 2015;10(12):1404–1413.

[11] Rezaee S, Farjah E, Khorramdel B. Probabilistic analysis of plug-in electric vehicles impact on electrical grid through homes and parking lots. IEEE Trans Sustain Energ 2013;4(4):1024–1033.

[12] Leemput N, Geth F, Van Roy J, et al. Impact of electric vehicle on-board single-phase charging strategies on a Flemish residential grid. IEEE Trans Smart Grid 2014;5(4):1815–1822.

[13] Leou R-C, Chun-Lien S, Chan-Nan L. Stochastic analyses of electric vehicle charging impacts on distribution network. IEEE Trans Power Syst 2014;29(3):1055–1063.

[14] Khorramdel B, Khorramdel H, Aghaei J, et al. Voltage security considerations in optimal operation of BEVS/PHEVS integrated microgrids. IEEE Trans Smart Grid 2015;6(4):1575–1587.

[15] Zhou B, Littler T, Meegahapola L, Zhang H. Power system steady-state analysis with large-scale electric vehicle integration Energy 2016;115(1):289–302.

[16] Ahmadian A, Sedghi M, Aliakbar-Golkar M, Elkamel A, Fowler M. Optimal probabilistic based storage planning in tap-changer equipped distribution network including PEVs, capacitor banks and WDGs: A case study for Iran. Energy 2016;112:984–997.

[17] Huang H, Chung CY, Chan KW, Chen H. Quasi-Monte Carlo based probabilistic small signal stability analysis for power systems with plug-in electric vehicle and wind power integration. IEEE Trans Power Syst 2013;28(3):3335–3343.

[18] Qian K, Zhou C, Allan M, Yuan Y. Modeling of load demand due to EV battery charging in distribution systems. IEEE Trans Power Syst 2011;26(2):802–810.

[19] Zhang P, Qian K, Zhou C, Stewart BG, Hepburn DM. A methodology for optimization of power systems demand due to electric vehicle charging load. IEEE Trans Power Syst 2012;27(3):1628–1636.

[20] Nunes P, Farias T, Brito MC. Enabling solar electricity with electric vehicles smart charging. Energy 2015;87(1):10–20.

[21] Carrion M, Zarate-Minano R. Operation of renewable-dominated power systems with a significant penetration of plug-in electric vehicles. Energy 2015;90:827–835

[22] Bremermann LE, Matos M, Peças Lopes JA, Rosa M. Electric vehicle models for evaluating the security of supply. Electr Power Syst Res 2014;111:32–39.

[23] Kavousi-Fard A, Khodaei A. Efficient integration of plug-in electric vehicles via reconfigurable microgrids. Energy 2016;111(15):653–663.

[24] Alipour M, Mohammadi-Ivatloo B, Moradi-Dalvand M, Zare K. Stochastic scheduling of aggregators of plug-in electric vehicles for participation in energy and ancillary service markets. Energy 2017;118(1):1168–1179.

[25] Kabir MN, Mishra Y, Bansal RC. Probabilistic load flow for distribution systems with uncertain PV generation. Appl Energy 2016;163(1):343–351.

[26] Ruiz-Rodriguez FJ, Hernandez JC, Jurado F. Probabilistic load flow for radial distribution networks with photovoltaic generators. IET Renew Power Gener 2012;6(2):110–121.

[27] Ruiz-Rodriguez FJ, Hernandez JC, Jurado F. Probabilistic load flow for photovoltaic distributed generation using the Cornish-Fisher expansion. Electr Power Syst Res 2012;89:129–138.

[28] Kaur A, Nonnenmacher L, Coimbra CFM. Net load forecasting for high renewable energy penetration grids. Energy 2016;114(1)1073–1084.

[29] Hernandez JC, De la Cruz J, Ogayar B. Electrical protection for the grid-interconnection of photovoltaic-distributed generation. Electr Power Syst Res 2012;89:85–99.

[30] Marinopoulos AG, Alexiadis MC, Dokopoulos PS. Energy losses in a distribution line with distributed generation based on stochastic power flow. Electr Power Syst Res 2011;81(10):1986–1994.

[31] Hernandez JC, Medina A, Jurado F. Optimal allocation and sizing for profitability and voltage enhancement of PV systems on feeders. Renew Energy 2007;32(10):1768–1789.

[32] Hernandez JC, Medina A, Jurado F. Impact comparison of PV system integration into rural and urban feeders. Energy Convers Manage 2008;49(6):1747–1765.

[33] Munkhammar J, Widen J, Ryden J. On a probability distribution model combining household power consumption, electric vehicle home-charging and photovoltaic power production. Appl Energy 2015;142(15):135–143.

[34] EN 50160. Voltage characteristics of electricity supplied by public distribution systems; 2000.

[35] IEC 61000-4-30. Testing and measurement techniques - Power quality measurement methods; 2008.

[36] Ortega MJ, Hernandez JC, Garcia OG. Measurement and assessment of power quality characteristics for photovoltaic systems: Harmonics, flicker, unbalance, and slow voltage variations. Electr Power Syst Res 2013;96:23–35.

[37] Kalos MH, Whitlock PA. Monte Carlo methods. 2nd ed. Wiley-VCH; 2008.

[38] P. McCullagh, Tensor methods in statistics. London: Chapman and Hall; 1987.

[39] Papoulis A, Pillai S. Probability, random variables, and stochastic processes. 4th ed. New York: McGraw-Hill; 2002.

[40] Zhang P. Lee ST. Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion. IEEE Trans Power Syst 2004;19(1):676–682.
[41] Cornish EA, Fisher RA. Moments and cumulants in the specification of distributions. Revue de l'Institut International de Statis 1937;5:307–320.
[42] Usaola J. Probabilistic load flow with correlated wind power injections. Electr Power Syst Res 2010;80(5):528–536.
[43] Kendall MG, Stuart A. The advanced theory of statistics. Vol. I. London: Charles Grin and Company Limited; 1963.
[44] Linden D, Reddy TB. Handbook of batteries. McGraw-Hill; 2002.
[45] Godina R, Paterakis NG, Erdinc O, et al. Impact of EV charging-at-work on an industrial client distribution transformer in a Portuguese Island. In Proc 2015 AUPEC; 2010. pp. 1–6.
[46] Madrid C, Argueta J, Smith J. Performance characterization-1999 Nissan Altra-EV with lithium-ion battery. 1999, Southern California.
[47] IEC 61851-1. Electric vehicle conductive charging system - General requirements; 2010.
[48] Härri J. Vehicular mobility modeling for VANET. Hoboken. NJ, USA: Wiley; 2009.
[49] Sanabria IA, Dillon TS. Stochastic power flow using cumulants and Von Mises functions. Int J Elect Power Energy Syst 1986;8(1):47–60.
[50] Ochoa LF, Dent CJ, Harrison GP. Distribution network capacity assessment: Variable DG and active networks. IEEE Trans Power Syst 2010;25(1):87–95.
[51] Verbic G, Cañizares CA. Probabilistic optimal power flow in electricity markets based on a two-point estimate method. IEEE Trans Power Syst 2006;21(4):1883–1893.
[52] Bright JM, Babacan O, Kleiss J, Taylor PG, RCrook. A synthetic, spatially decorrelating solar irradiance generator and application to a LV grid model with high PV penetration. Solar Energy 2017;147(1):83–98.
[53] Papaefthymiou G, Kurowicka D. Using copulas for modeling stochastic dependence in power system uncertainty analysis. IEEE Trans Power Syst 2009; 24(1):40–49.
[54] Kurowicka D, Cooke RM. Uncertainty analysis with high dimensional dependence modelling. New York: Wiley; 2006.
[55] Liu P-L, Kiureghian A. Multivariate distribution models with prescribed marginals and covariances. Probabilistic Engineering Mechanics 1986;1(2):101–112.
[56] Qin Z, Li W, Xiong X. Generation system reliability evaluation incorporating correlations of wind speeds with different distributions. IEEE Trans Power Syst 2013;28(1):551–558.
[57] Cario M, Nelson B. Modeling and generating random vectors with arbitrary marginal distribution and correlation matrix. Northwestern University, Evanston, IL, USA, Apr. 1997, Tech. Rep.
[58] Madsena L, Birkes D. Simulating dependent discrete data. J Statistical Computation Simulation 2011;1 677–691.
[59] Morales JM, Conejo AJ, Perez-Ruiz J. Simulating the impact of wind production on locational marginal prices. IEEE Trans Power Syst 2011;26(2):820–828.
[60] Ray S, Lindsay BG. The topography of multivariate normal mixtures. Annals of Statistics 2005;33(5):2042–2065.

[61] Hontoria L, Aguilera J, An application of the multilayer perceptron: Solar radiation maps in Spain. Sol Energy 2005;79(5):523–530.
[62] Li W. Reliability assessment of electrical power systems using Monte Carlo methods. Springer; 1994.
[63] Li W. Probabilistic transmission system planning. Wiley-IEEE Press; 2011.
[64] National Institute of Statistics of Spain. http://www.ine.es [accessed 01/10/2016].
[65] Pasaoglu G, Fiorello D, Martino A, et al. Driving and parking patterns of European car drivers. A mobility survey. 2012. European Commission. Tech. Rep. EUR 25627 EN.
[66] National Department of Traffic (DGT) of Spain. http://www.dgt.es/es [accessed 01/10/2016].
[67] Gnann T, Plötz P, Kühn A, Wietschel M. Modelling market diffusion of electric vehicles with real world driving data - German market and policy options. Transp Res Part A 2015;77:95–112.
[68] Eurostat. Electricity consumption by households. http://ec.europa.eu/eurostat [accessed 01/10/2016].
[69] Environmental assessment of plug-in hybrid electric vehicles, nationwide greenhouse gas emissions. Tech. Rep. 1015325 EPRI and NRDC 1 (2007) 1–56.
[70] Anders J. Probability concepts in electric power systems. New York: John Wiley and Sons; 1990.
[71] Grainger J. Stevenson WD. Power systems analysis, electrical engineering. McGraw-Hill International Editions; 1994.
[72] Shirmohammamdi D, Hong HW, Semlyen A. A compensation-based power flow method for weakly meshed distribution and transmission network. IEEE Trans Power Syst 1988;3(2):753–762.