

Forecasting Congestions in Interconnected Power Systems with High Shares of Renewable Energy: A Probabilistic Approach using Copulas

Aiko Schinke-Nendza ¹ and Christoph Weber ²

¹House of Energy Markets and Finance

²Affiliation not available

October 30, 2023

Abstract

This paper proposes a novel approach to forecast congestions in high-voltage grids with high shares of distributed photovoltaic (PV) infeed. The approach is based on a physical PV model using intra-day numerical weather prediction (NWP) input data. Subsequently, probabilistic forecasts are generated based on Kernel density estimators (KDE) and Copula, describing the multivariate spatial dependencies for the marginal distributions of forecasting and approximation errors. Finally, a probabilistic power flow (PPF) using a linearized AC version is proposed, combining the benefits of high accuracy with high computational performance. To assess and quantify the overall advantages of this approach, a case study is carried out for an existing power system.

Forecasting Congestions in Interconnected Power Systems with High Shares of Renewable Energy: A Probabilistic Approach using Copulas

Aiko Schinke-Nendza, *Student-Member, IEEE*, and Christoph Weber, *Member, IEEE*

Abstract—The integration of increasing amounts of volatile renewable energy sources (RES) induces a growing demand for reliable forecasts of congestion to support renewable curtailment and redispatch decisions. Therefore, this paper proposes a novel approach to forecast congestions in high-voltage grids with high shares of distributed photovoltaic (PV) infeed. This approach is based on a physical PV model using intra-day numerical weather prediction (NWP) input data. Subsequently, probabilistic forecasts are generated based on Kernel density estimators (KDE) and Copula, describing the multivariate spatial dependencies for the marginal distributions of forecasting and approximation errors. Finally, a probabilistic power flow (PPF) using a linearized AC version is proposed, combining the benefits of high accuracy with high computational performance. To assess and quantify the overall advantages of this approach, a case study is carried out for an existing power system. To benchmark the proposed approach comparisons to standard multivariate normal (MVN) distributions (with and without correlations), as well as to AC and DC PPFs are compared. Non-parametric probabilistic metrics are used for evaluation, comparing accuracy and precision of individual marginal distributions, and taking the spatial dependency between forecasts at different locations into account.

Index Terms—Congestion, Forecast, Photovoltaic, Copula, Spatial Dependency, Probabilistic Power Flow

I. INTRODUCTION

OVER the past decades the infeed from renewable energy sources into electric power systems all over the world has progressed rapidly. For instance, photovoltaic (PV) became one of the leading technologies in Germany, comprising more than 49 GW of installed capacity by 2019. High shares of RES contribute obviously to reach decarbonization goals. However, most of the RES are decentralized, rather small-scale units, connected to the lower voltage levels and providing highly volatile infeed. Therefore, system operators are presently facing new challenges, especially, in terms of operational tasks such as congestion management. One appropriate support to system operators to handle these challenges are detailed forecasts focusing not only on the infeed patterns but also on the resulting power flows in the grid.

In terms of forecasting PV generation, various methods have been proposed over the past years focusing on different application areas [1]. Generally, those methods can be divided into deterministic and probabilistic forecasts [2]. In contrast to deterministic models, probabilistic forecasts contain additional information on the uncertainty of the prediction. However, when considering probabilistic forecasts of PV generation physical, statistical and hybrid approaches are distinguished, using either parametric or non-parametric techniques [2]. In this context, statistical approaches for PV modelling provide accurate and reliable results, especially when detailed information on individual PV plants is not available but metering data instead [3]. However, in terms of power system operations, e.g., when considering renewable curtailment as part of the system operator's congestion management, detailed information on individual RES units is required [4]. Hence, physical PV models are beneficial for the system operator, predicting the individual probabilistic state of the relevant PV units.

These PV forecasts enable system operators to ensure a stable operation of the corresponding power system, i.e., to anticipate congestions before their occurrence and take countermeasures in advance. In European electricity markets these countermeasures comprise in particular redispatching, countertrading and renewable curtailment of renewable energies [5]. The U.S. wholesale electricity markets instead rely on locational marginal pricing (LMP), i.e., taking congestions into account during market clearing [6]. Regardless of the regulatory framework and market design, it is essential for system operators to predict potential congestions at an early stage, taking the increasing uncertainties into account, in order to derive the necessary countermeasures.

In this context the probabilistic power flow (PPF) offers the possibility to take those uncertainties, e.g., resulting from probabilistic forecasts or network approximations, into account. In recent years, several methods for PPF have been developed and investigated in detail, as discussed by Chen et al. [7]. Generally, present investigations in this context focus on new mathematical approaches [8], network outages [9], power

This work was supported by the Federal Ministry of Education and Research (BMBF) of Germany within the framework of the joint project "Verteilnetze" (Grant No. 05M18PGA).

A. Schinke-Nendza is with the House of Energy Markets and Finance, University of Duisburg-Essen, Universitätsstr. 12, 45117, Essen, Germany (e-mail: aiko.schinke-nendza@uni-due.de).

C. Weber is with the House of Energy Markets and Finance, University of Duisburg-Essen, Universitätsstr. 12, 45117, Essen, Germany (e-mail: christoph.weber@uni-due.de).

system stability [10] and computational efficiency [11]. In terms of modelling stochastic dependencies of forecasting errors and uncertainties [12], either spatial [13] or temporal dependencies are considered [14]. In addition, it is rather common for power system applications to assume multivariate normal distributions when considering such dependencies [15, 16].

Within this paper, a novel approach is proposed for probabilistic congestion forecasting in the presence of large amounts of renewables, in occurrence PV systems. Thereby a hybrid approach is developed to derive multivariate probabilistic RES generation forecasts. In a first step a deterministic PV forecast is carried out using numerical weather prediction (NWP) data, based on a physical PV model with pre-processing of site-specific parameters (cf. Section II). Subsequently, the marginal distributions of the forecasting and approximation errors are simulated (cf. Section III) using kernel density estimators (KDE), after introducing a concept for eliminating the heteroscedasticity resulting from the temporal dependence. Furthermore, a copula approach is used to model the spatial dependence between those errors to cope for the resulting uncertainty. Therefore, different copulas, namely the Gaussian and Student's t, are applied hereinafter and benchmarked to the assumption of multivariate normal distribution (with and without spatial dependencies). Finally, the PPF is specified (cf. Section IV), including empirically validated shift factors for power imbalances to boundary nodes. To meet real-world performance requirements, a fully linearized version is designed that is benchmarked (cf. Section V) to the AC-PPF and the commonly used DC-PPF in terms of computational efficiency and accuracy (cf. Section VI).

II. PHOTOVOLTAIC POWER FORECASTING

To forecast the net power injection of photovoltaic (PV) units connected to an electric power system, the physical PV model introduced by Schinke-Nendza et al. [17] is taken as basis. This model uses site-specific hourly NWP data as an input and provides hourly forecasts of the power generation of individual PV units. The structure of the physical PV model used by [17] is depicted in Fig. 1. Using site-specific hourly NWP data, the model provides a rather high accuracy on a single plant level compared to other approaches, e.g., [18]. However, empirical analyses indicate that available data regarding some of the model parameters, such as the nominal power and the panel orientation of PV units, may be erroneous.

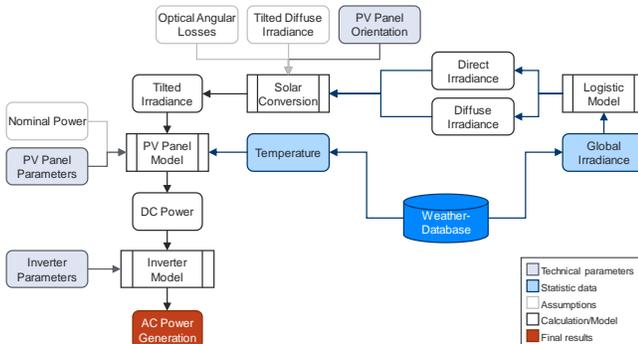


Fig. 1. Structure of the physical PV model used within this paper.

Therefore, a nonlinear least squares (NLS) regression is used here to obtain the best possible fit of the physical model, instead of using a statistical post processing as in [17].

III. MODELING THE SPATIAL INTERDEPENDENCE OF FORECASTING ERRORS

The cumulative net power injection at a certain feed-in location, i.e., a node of the underlying power system, can be obtained by aggregation of multiple PV power forecasts. The relative forecasting error of the PV net power injection at feed-in locations $k \in \mathcal{K} = \{1, \dots, K\}$ during timestep t is defined as:

$$\Delta p_{t,k} = \frac{P_{t,k} - \hat{P}_{t,k}}{P_{0,k}} \quad (1)$$

where $\hat{P}_{t,k}$ is the deterministic forecast of the physical model using NWP data, $P_{0,k}$ is the cumulative rated power, and $P_{t,k}$ is the actual measured power of the PV units at feed-in location k .

A. Upscaling the PV forecasts

Considering feed-in location k , there are $|N_k|$ PV units $n \in N_k$, connected to this node, comprising an actual rated power of $P_{0,k}$. In practice, the available dataset (for measurement and NWP data) yet commonly does not include all $|N_k|$ PV units, but rather a subset N'_k of $|N'_k|$ PV units (with $|N'_k| \leq |N_k|$ and $P'_{0,k} \leq P_{0,k}$). Therefore, the deterministic forecast \hat{P}_k of the physical model as well as the forecasting error $\Delta p_{t,k}$ need to be upscaled to receive reliable forecasts in terms of network congestions, taking all connected PV units into account. In terms of the deterministic forecast, the upscaling is carried out using a linear upscaling coefficient a_{k,P_0} defined as the ratio of the actual cumulative rated power $P_{0,k}$ to the rated power of the subset $P'_{0,k}$:

$$\hat{P}_{t,k} = \hat{P}'_{t,k} \frac{P_{0,k}}{P'_{0,k}} = \hat{P}'_{t,k} a_{k,P_0} \quad (2)$$

Considering the relative forecasting error $\Delta p_{t,k}$ at feed-in location k , the question arises to which degree the individual forecasting errors of the single PV units are stochastically dependent. Our empirical analyses indicate the linear upscaling coefficient to perform the best, i.e., assuming a full correlation of the forecasting errors of the PV units located at a certain node. This assumption of fully correlated forecasting errors outperforms both a correlation-based upscaling and the independency assumption.

B. Elimination of Bias and Heteroscedasticity

To reduce the modelling effort in terms of the copulas and the complexity of the marginal distributions, any bias of the forecasts possibly arising from the physical model is quantified and eliminated for each hour separately. Therefore, the forecasting error is hereinafter assumed to be unbiased (notated as $\Delta \tilde{p}_{t,k}$). Yet, a specific heteroscedasticity, i.e., the time dependence of the variance, may still be observed in the relative forecasting errors $\Delta \tilde{p}_{t,k}$, leading to higher deviations during the midday hours, but lower deviations during the morning and evening hours [14]. Therefore, a novel approach for reducing the heteroscedasticity of the forecasting error is introduced, using the extraterrestrial irradiance as a basis for normalization, following [19]. The unbiased normalized forecasting error $\tilde{\epsilon}_{t,k}$ is then obtained as follows:

$$\tilde{\varepsilon}_{t,k} = \Delta \tilde{P}_{t,k} \frac{\bar{I}_{t,0}}{I_{t,ex}} \frac{P_{t,k} - \tilde{P}_{t,k}}{P_{0,k}} \frac{\bar{I}_{t,0}}{I_{t,ex}} \quad (3)$$

The spatial dependence for all feed-in locations is hence considered hereinafter based on the marginal distributions of $\tilde{\varepsilon}_{t,k}$ using a copula approach.

C. Spatial Dependence Structure

The interdependence of the forecasting errors at multiple feed-in locations $k \in \mathcal{K} = \{1, \dots, K\}$ can be described using a joint distribution function $F(\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_K)$. According to [20] this function may also be written as follows:

$$F(\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_K) = C(F_1(\tilde{\varepsilon}_1), \dots, F_K(\tilde{\varepsilon}_K)) \quad (4)$$

where $C: [0,1]^K \rightarrow [0,1]$ represents a copula and F_k describes the marginal distribution (cumulative distribution function, CDF) of the forecasting error at feed-in location k . In this context, the elements of the input vector can be written as:

$$u_k = F_k(\tilde{\varepsilon}_k) \quad (5)$$

whereby the elements $u_k \sim U(0,1)$ are standard uniformly distributed. By means of the inverse probability function F_k^{-1} , the copula may be expressed as follows:

$$C(u_1, \dots, u_K) = F(F_1^{-1}(u_1), \dots, F_K^{-1}(u_K)) \quad (6)$$

for all feed-in locations $k \in \mathcal{K}$.

1) Gaussian Copula

The spatial dependence structure of the forecasting errors at multiple locations can be modelled, for instance using a Gaussian copula, see [21]. Following [22], the copula $C_{Gauss}: [0,1]^K \rightarrow [0,1]$ is described as the multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{R})$ in \mathbb{R}^K with the density following to:

$$c_{Gauss}(\mathbf{u}|\mathbf{R}) = \frac{1}{\sqrt{\det \mathbf{R}}} \exp\left(-\frac{1}{2} \mathbf{x}^T (\mathbf{R}^{-1} - \mathbf{I}) \mathbf{x}\right) \quad (7)$$

where $\mathbf{R} \in \mathbb{R}^{K \times K}$ is a correlation matrix with diagonal elements equal to one, $\mathbf{I} \in \mathbb{R}^{K \times K}$ the identity matrix, $\mathbf{x} = (\phi^{-1}(u_1), \dots, \phi^{-1}(u_K))^T$ the vector of inverse cumulative standard normal distribution ϕ^{-1} , and u_k the input vector of the Gaussian copula according to (5). The correlation matrix can be estimated approximately, following [22], giving an estimator $\hat{\mathbf{R}}$ close to the maximum likelihood estimation (MLE).

2) Student's t Copula

The spatial interrelation of the forecasting errors at multiple locations can also be modelled using a t copula. Following [23], this type of copula surpasses symmetrical copulas, such as Gaussian copulas, when modelling tail-dependencies of the underlying marginal distributions. Based on the multivariate Student's t distribution for a set of random variables $\mathbf{x} = (x_1, \dots, x_K)^T \in \mathbb{R}^K$, written as $\mathbf{x} \sim t_{K,\nu}(\mathbf{0}, \mathbf{R})$, with ν degrees of freedom, a mean vector equal to $\mathbf{0}$ and a correlation matrix \mathbf{R} as introduced above, the density of the t copula is given by:

$$c_t(\mathbf{u}|\mathbf{R}, \nu) = \frac{\Gamma\left(\frac{\nu}{2}\right)^{K-1} \Gamma\left(\frac{\nu+K}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)^K \sqrt{\det \mathbf{R}}} \cdot \left(1 + \frac{\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}}{\nu}\right)^{-\frac{\nu+K}{2}} \prod_{k=1}^K \left(1 + \frac{x_k^2}{\nu}\right)^{\frac{\nu+1}{2}} \quad (8)$$

where $\Gamma(a)$ is the gamma function, \mathbf{x} the input vector of the t copula with elements $x_k = F_{t_\nu}^{-1}(u_k)$ for $k \in \mathcal{K}$ and u_k

represents the CDF value of the forecasting error at location $k \in \mathcal{K}$ according to equation (5), see [23]. Following [22], the correlation matrix estimator $\hat{\mathbf{R}}$ can be obtained using Kendall's tau estimates in a first step. In a second step, the remaining parameter ν , i.e., the degree of freedoms, can be estimated by fixing $\hat{\mathbf{R}}$ and applying an MLE to maximize the log-likelihood function for the pseudo-sample, which is defined as $\mathbf{x}_n = (F_{t_\nu}^{-1}(u_{n,1}), \dots, F_{t_\nu}^{-1}(u_{n,K}))^T$ with $n \in \{1, \dots, N\}$ and N observations:

$$\ln L(\nu|\hat{\mathbf{R}}, \mathbf{x}) = \sum_{n=1}^N \left(\ln f_{t_{\nu}}(\mathbf{x}_n|\hat{\mathbf{R}}) - \sum_{k=1}^K \ln f_{t_{\nu}}(x_{k,n}) \right) \quad (9)$$

D. Marginal Distributions

The marginal distributions $F_k(\tilde{\varepsilon}_k)$ of the forecast errors at the feed-in locations $k \in \mathcal{K}$, can be estimated using a non-parametric approach. Subsequently a kernel density estimation (KDE) is applied to the marginal distributions, as proposed by Chai et al. for the forecasting errors of PV generation [14]. Using the sample of the forecasting error $\tilde{\varepsilon}_k = (\tilde{\varepsilon}_{k,1}, \dots, \tilde{\varepsilon}_{k,N})$ at location $k \in \mathcal{K}$ with $n \in \{1, \dots, N\}$ and N observations, the kernel density estimator using a gauss kernel is defined as:

$$\hat{f}_k(x) = \frac{1}{Nh_k} \sum_{n=1}^N \phi\left(\frac{x - \tilde{\varepsilon}_{k,n}}{h_k}\right) \quad (10)$$

where ϕ is the density of the standard normal distribution, $\tilde{\varepsilon}_{k,n}$ corresponds to the mean of the n -th kernel function and h_k is the bandwidth of the KDE [24]. The bandwidth h_k , representing a free parameter of the kernel density estimation, can be determined analytically, for instance using the normal distribution approximation following [25]. In this case the optimal estimator of the bandwidth can be obtained as:

$$\hat{h}_k = \left(\frac{4}{3N}\right)^{\frac{1}{5}} \sqrt{\frac{\tilde{\varepsilon}_k^T \tilde{\varepsilon}_k}{N-1}} \quad (11)$$

in case that the forecasting errors are unbiased.

IV. PROBABILISTIC POWER FLOW ANALYSIS

A network consisting of N_B nodes (not necessarily equal to the number of feed-in locations K) and M lines is investigated, assuming a perfectly balanced power system, hence focussing on a single-phase representation. To determine the power flow along the transmission lines of the network, the equations below can be used:

$$p_i = v_i \sum_{j=1}^{N_B} v_j (g_{i,j} \cos \delta_{i,j} + b_{i,j} \sin \delta_{i,j}), \quad i \in N_B \quad (12)$$

$$q_i = v_i \sum_{j=1}^{N_B} v_j (g_{j,i} \sin \delta_{i,j} - b_{i,j} \cos \delta_{i,j}), \quad i \in N_B \quad (13)$$

where p_i is the net active power injection and q_i is the net reactive power injection, both at node $i \in N_B$, v_j is the voltage magnitude, $\delta_{i,j} = \delta_i - \delta_j$ is the difference of voltage angles between node i and node j , and $y_{i,j} = g_{i,j} + b_{i,j}$ is the admittance of the line from node i to node j , consisting of the conductance $g_{i,j}$ and the susceptance $b_{i,j}$, following [7]. In addition, the line parameters are assumed to be time-independent and to not vary with the weather conditions (e.g., temperature, wind speed etc.).

A. Network Congestions

The constraints in a power system for steady-state operation typically depend on the values of the current magnitudes i_m on the transmission lines $m \in \mathcal{M}$ and the voltage magnitudes v_j at the nodes $j \in N_B$. Assuming the voltages to be uncritical, congestions in a power system exclusively depend on the transmission line currents. Therefore, the following constraints needs to hold:

$$i_m^{max} \geq i_m = \frac{1}{y_{j,l}} \sqrt{g_{j,l}^2 + b_{j,l}^2} \sqrt{v_j^2 + v_l^2 - 2v_j v_l \cos \theta_{j,l}} \quad (14)$$

whereby the maximum permissible current i_m^{max} of line m may contain some additional security margins. For ease of notation, the time index t is dropped here and subsequently.

B. Modelling Boundary Nodes to External Networks

To investigate the impact of the forecasting error on the congestions in an interconnected network, the effects of adjacent networks need to be considered. In case of a forecasting error different from zero, the power imbalance is at least partly shifted to the boundary nodes $b \in N_T$ connecting the considered network with external networks. Therefore, a multivariate linear regression (MVLRL) on available measurement data is proposed in this paper. Building on the adjustment matrix introduced in [26], it models the dependency of power deviations at the boundary nodes from the inner nodes that connect decentralized PV units and loads. The MVLRL is specified as follows:

$$\Delta P_b = \sum_{k \in \mathcal{K}} \beta_{b,k} \Delta P_k + \varepsilon_b \quad (15)$$

where $\beta_{b,k}$ is the adjustment factor at boundary node b given changes at inner node k . The adjustment matrix \mathbf{B} summarizes the shifts in power flows at the boundary nodes as a (linear) function of the changes at the infeed and load nodes:

$$\begin{pmatrix} \Delta P_1 \\ \vdots \\ \Delta P_B \end{pmatrix} = \underbrace{\begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,K} \\ \vdots & \ddots & \vdots \\ \beta_{B,1} & \cdots & \beta_{B,K} \end{pmatrix}}_{\mathbf{B}} \begin{pmatrix} \Delta P_1 \\ \vdots \\ \Delta P_K \end{pmatrix} \quad (16)$$

Furthermore, the error of this approximation is denoted as ε_b and described by a time-invariant marginal distribution $F_b(\varepsilon_b)$. The overall uncertainty ξ_b at boundary node b given a balanced power flow schedule and some PV forecast errors $\tilde{\varepsilon}_k$ can then be described as a superposition of the approximation error ε_b and the forecasting errors of the feed-in locations $\tilde{\varepsilon}_k$ weighted by the adjustment factors, respectively:

$$\xi_b = \varepsilon_b + (\beta_{b,1} \quad \cdots \quad \beta_{b,K}) \begin{pmatrix} \tilde{\varepsilon}_1 P_{0,1} \\ \vdots \\ \tilde{\varepsilon}_K P_{0,K} \end{pmatrix} \frac{I_{ex}}{I_0} \quad (17)$$

To estimate marginal distributions of the approximation error ε_b , the changes in power balances between two consecutive timesteps are considered. In the event of forecasting errors, the same adjustment matrix may be used to describe the corresponding changes in the boundary power flows, assuming the initial power flow to be known perfectly for the expected values of the PV power injection. The additional approximation error $\varepsilon_{t,b}$ then comes on top of the forecasting errors. Since both types of errors have different sources and are derived from different datasets (approximation

error: time series of observed power flow changes; forecasting error: mismatches between prediction and measurement for a certain time step) the errors are appropriately deemed to be stochastically independent.

C. Linearized Power Flow

The AC power flow defined by equations (12) and (13) describes a non-linear problem due to the quadratic dependence on the voltage magnitudes and the sinusoidal dependence on the voltage angles. To substantially improve the computation time of the PPF, a linearized AC power flow is used hereinafter.

1) Network Equations

In this context, a certain state of operation in an upcoming interval (index: *) is estimated based on the expected initial state of operation (index: 0), using a Taylor series:

$$\begin{pmatrix} \theta^* \\ v^* \end{pmatrix} \approx \begin{pmatrix} \theta_0 \\ v_0 \end{pmatrix} + \mathbf{J}^{-1} \left(\begin{pmatrix} p^* \\ q^* \end{pmatrix} - \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \right) \quad (18)$$

where \mathbf{J} is the Jacobian matrix, consisting of the gradients of the active and reactive power differentiated with respect to the voltage angle θ and the voltage magnitude v , see [8]:

$$\begin{pmatrix} \Delta p \\ \Delta q \end{pmatrix} \approx \underbrace{\begin{pmatrix} \nabla_{\theta} p & \nabla_v p \\ \nabla_{\theta} q & \nabla_v q \end{pmatrix}}_{\mathbf{J}} \begin{pmatrix} \Delta \theta \\ \Delta v \end{pmatrix} \quad (19)$$

Hence the inverse Jacobian matrix can be used to determine the change in the voltage magnitudes Δv and angles $\Delta \theta$ following a change in the active and reactive powers Δp and Δq .

2) Line Current

In addition, also equation (18) can be linearized for the initial state of operation, using the expected values of net power injections, as described above. The current i_m on transmission line $m \in \mathcal{M}$ depends on the voltage magnitudes and angles at the beginning of the line $j \in N_B$ and the end of the line $l \in N_B$ (written as $j \leftrightarrow l$, where $j \neq l$ holds). The first-order Taylor approximation for the current on line $m \in \mathcal{M}$ writes

$$i_m \approx i_{m,0} + \nabla f_m(\mathbf{x}_0)(\mathbf{x}_0 - \mathbf{x}^*) \quad (20)$$

with $f_m(\mathbf{x}) = y_{j,l} \sqrt{v_j^2 + v_l^2 - 2v_j v_l \cos(\theta_j - \theta_l)}$ and the vector of input variables $\mathbf{x} = (v_j, v_l, \theta_j, \theta_l)^T$. Using two $M \times N_B$ matrices \mathbf{D}_{θ} and \mathbf{D}_v , the linear approximation of equation (20) may be rewritten:

$$i^* \approx i_0 + (\mathbf{D}_{\theta} \quad \mathbf{D}_v) \begin{pmatrix} \theta^* \\ v^* \end{pmatrix} - \begin{pmatrix} \theta_0 \\ v_0 \end{pmatrix}, \quad (21)$$

Where the elements $d_{m,j}^{\theta}$ of the matrix \mathbf{D} are defined as

$$d_{m,j}^{\theta} = \begin{cases} \frac{\partial f_m(\mathbf{x}_0)}{\partial \theta_j}, & \forall m \in \{(j,l) | (j,l) \in N_B^2 \setminus \{j \leftrightarrow l\}\} \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

and the elements $d_{m,n}^v$ of the matrix \mathbf{D}_v are defined as

$$d_{m,n}^v = \begin{cases} \frac{\partial f_m(\mathbf{x}_0)}{\partial v_j}, & \forall m \in \{(j,l) | (j,l) \in N_B^2 \setminus \{j \leftrightarrow l\}\} \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

3) Probabilistic Power Flow

Assuming the loads in the power system to be forecasted with a high accuracy, the uncertainties in the net power injection of the PV units $\Delta p_{PV,k}$ at feed-in location $k \in \mathcal{K}$ are larger than the uncertainties in the load $\Delta p_{L,k}$: $\Delta p_{PV,k} \gg \Delta p_{L,k}$. Furthermore, it is assumed that most of the considered PV units

are located in the medium and low voltage levels (without any reactive power management as described in [27]), hence not affecting the reactive power exchange between voltage levels. Therefore, the change in the line currents $\Delta \mathbf{i} = \mathbf{i}^* - \mathbf{i}_0$ depends on the changes of the net power injection of the PV units, i.e., $\Delta \mathbf{p}_{PV,k} = \mathbf{p}_{PV}^* - \mathbf{p}_{PV,0}$ for $k \in \mathcal{K}$, exclusively. Hence, the overall linearized equations for obtaining the transmission line currents based on equations (18) and (21) can be formulated as:

$$\mathbf{i}^* = \mathbf{i}_0 + (\mathbf{A}_\theta \quad \mathbf{A}_v) \begin{pmatrix} \mathbf{J}_{11}^{-1} \\ \mathbf{J}_{21}^{-1} \end{pmatrix} (\mathbf{p}_{PV}^* - \mathbf{p}_{PV,0}) \quad (24)$$

with the inverse Jacobian matrix being segmented into the four $N_B \times N_B$ matrices \mathbf{J}_{11}^{-1} to \mathbf{J}_{22}^{-1} :

$$\mathbf{J}^{-1} = \begin{pmatrix} \mathbf{J}_{11}^{-1} & \mathbf{J}_{12}^{-1} \\ \mathbf{J}_{21}^{-1} & \mathbf{J}_{22}^{-1} \end{pmatrix} \quad (25)$$

The problem of solving multiple non-linear equations thus becomes linear and the solution is obtained directly through matrix operations on the change of the active power net injection of the PV units at the different feed-in locations. The PPF then is carried out using the Monte Carlo simulation introduced below.

D. Monte Carlo Simulation

Within this paper a Monte Carlo simulation (MCS) is proposed, following [28]. Hereinafter the MCS is used to stochastically predict network congestions by applying the probabilistic forecasts of the PV generation as input variables to the power flow analysis. Since finding a closed-form analytical solution for the given mathematical problem may either not be feasible or lead to an over-sophisticated solution, the application of the Monte Carlo method is deemed more appropriate, relying on the central limit theorem and the law of large numbers [28]. In the application, the MCS will be applied to all alternatives: The AC power flow, the linearized AC power flow and the DC power flow. Finally, the different approaches are compared using the performance metrics and benchmark model described in the following section, to evaluate the improvements of the overall approach.

V. PERFORMANCE MEASUREMENT

A. Probabilistic Metrics

To evaluate the goodness of fit of the spatial dependency and to assess the accuracy of multivariate forecasting models predicting the D-dimensional variable \mathbf{z} for a certain timestep t , there are two commonly used probabilistic measures: The energy score and the variogram-based score [21].

1) Energy Score

The energy score (ES) represents a proper multivariate generalization of the univariate continuous ranked probability score and is applicable to evaluate the ensembles of forecasts of the variable $\mathbf{z}_t = (\theta_{t,1}, \dots, \theta_{t,D})^T$:

$$ES_t^{(z)} = \frac{1}{N_S} \sum_{n=1}^{N_S} \|\mathbf{z}_t - \hat{\mathbf{z}}_t^{(n)}\|_2 - \frac{1}{2N_S^2} \sum_{i=1}^{N_S} \sum_{j=1}^{N_S} \|\hat{\mathbf{z}}_t^{(i)} - \hat{\mathbf{z}}_t^{(j)}\|_2 \quad (26)$$

where N_S describes the sample size and $\hat{\mathbf{z}}_t^{(n)}$ is the vector of forecasts for sample n . The ES is a negative scoring rule; hence, lower values indicate a higher forecasting accuracy. However, investigations of Scheurer and Hamill [29] have shown that the

ES is not adequately sensitive to incorrectly specified correlations of multivariate forecasts. It is worth mentioning that the ES is able to detect mean-biased forecasts correctly, but it is not able to indicate incorrect dependency structures and false variances clearly [29].

2) Variogram-Based Score

Especially for evaluating the spatial dependence structure and the spread of multivariate forecasts for the D-dimensional variable θ , the variogram-based score (VS), according to [29], is considered within this paper as a proper scoring rule:

$$VS_t^{(z)} = \sum_{i=1}^D \sum_{j=1}^D w_{ij} \left(\frac{1}{N_S} \sum_{n=1}^{N_S} |z_{t,i}^{(n)} - z_{t,j}^{(n)}|^\gamma - |z_{t,i} - z_{t,j}|^\gamma \right)^2 \quad (27)$$

where γ is the order of the VS (where [29] proposes 0.5 for accurate results) and $w_{j,l}$ is a weighting factor. Just as the ES, this scoring rule is also negative oriented; hence, lower values indicate a higher forecasting accuracy.

3) Diebold-Mariano Test

In addition, the Diebold-Mariano (DM) test can be carried out on VS and ES, according to [30], for testing the null hypothesis that the performance of two considered models is equal, assuming a standard normal distribution to hold for the test statistic, i.e., $t_{DM} \sim \mathcal{N}(0,1)$. For a given the significance level α , the DM test statistic furthermore indicates the preferred model [30].

B. Benchmark Model

Since the probabilistic performance metrics provide a comparative assessment of different models, a simple benchmark model is introduced hereinafter to assess and quantify the improvement resulting from the overall approach introduced in the sections above.

1) Multivariate Normal Distribution

As a simplified benchmark model, a multivariate normal distribution (MVN) is introduced, following [21]. This allows to identify the benefits of applying the more complex and numerically demanding copula approach for modelling the forecasting errors at multiple feed-in locations and their spatial dependences. A further simplification is to assume the forecasting errors of the multiple feed-in locations to be spatially independent. For both approaches (with and without spatial dependence) the relative forecast errors $\Delta \mathbf{p}_k$ can be described as $\Delta \mathbf{p} \sim \mathcal{MVN}(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p)$. Hence, for each feed-in location and each hour a set of parameters exists: Whereas the mean relative forecasting error $\mu_{p,k}$ should be zero for unbiased forecasts, the variance of the forecasting error $\sigma_{p,k}^2$ and the covariances to the other feed-in locations are summarized in the variance-covariance matrix $\boldsymbol{\Sigma}_p$. When neglecting the spatial dependence between the forecasting errors in the second benchmark, non-diagonal elements become zero.

2) DC Power Flow

The DC power flow is based on the assumption that the voltages at all nodes j and l are close to their nominal value, the phase angle differences $\theta_{j,l}$ are small and losses can be neglected, as described in [6]. Therefore, the relation between nodal net power injections \mathbf{p}_{N_B} and the power flow across the

considered transmission lines can be described using the power transfer distribution factor (PTDF) matrix \mathbf{H} :

$$\mathbf{p}_M = \mathbf{H} \mathbf{p}_{N_b} \quad (28)$$

Based on the previously introduced assumptions $\mathbf{i}_M = \mathbf{p}_M$ holds for per unit values.

VI. APPLICATION

Within this paper an existing power system on the high voltage level (110 kV) is considered as a case study, see Fig. 2. This network consists of 18 buses and 17 nodes, of which 12 nodes represent PV feed-in locations, accounting for an installed capacity of 527.5 MW. The nodes of the considered power system are serially numbered (1 to 17), whereby the boundary nodes connecting to other networks (nodes 4, 6, 14 and 17) are highlighted with arrows. For the power system, measurement data for each node and each transmission line are available in a quarter-hourly temporal solution for one year (here: 2017), i.e., voltages, active and reactive power. Regarding the net power injection of the PV units, measurement data of the active power is available for a sample of PV units comprising 156.2 MW of installed capacity, see Table I.

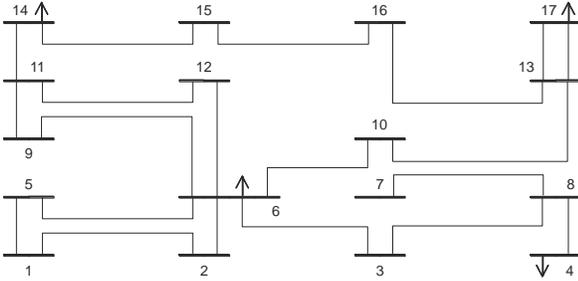


Fig. 2. Investigated segment of the high voltage power system.

TABLE I
DATA OF AVAILABLE PV UNITS IN SAMPLE AND TOTAL

Node No.	Sample Power (in MW)	No. of units in sample	Total Power (in MW)
1	3.30	4	27.16
2	18.44	10	46.91
5	14.23	19	64.02
6	19.49	12	59.03
7	25.44	6	37.13
9	1.58	5	40.87
10	5.34	3	30.42
12	16.25	9	41.82
13	10.68	3	18.47
14	14.09	4	45.29
15	18.01	11	68.13
16	9.35	7	48.25

The measurement data is available in an hourly temporal solution for three months of three years (May to July, 2015 to 2017). For each PV unit, hourly NWP data (global irradiance and temperature) is available in a forecast horizon of 1-3 hours, i.e., the forecasts are updated in intervals of 3 hours and contain information on the next couple of hours.

A. Evaluation Methodology

In a first step, a reference current is calculated for each line based on the available measurement data. Furthermore, the load, the grid voltages and the transfer power to external networks are assumed to be known exactly, i.e., the power

transfer at boundary nodes solely varies due to forecasting and related approximation errors. All models have been trained on the available data of the years 2015 and 2016. In addition, the adjustment matrix and the approximation error for the boundary nodes are determined for January to April 2017 exclusively. Hence, an out-of-sample test of all models is carried out for 2,015 hours of May to July 2017. The MCS hereinafter is based on a sample size of 5,000. Subsequently, the results are evaluated as follows: the KDE-based gaussian copula (GC), the KDE-based student's-t copula (TC), the multivariate normal distribution with correlation (MVN) and without correlation (UMVN) are combined with the different PPF approaches, namely the linearized AC (LAC) PPF as well as AC and DC PPF, as introduced above.

B. Results

In the following, the results of the case study introduced above are presented. Therefore, the sensitivity to the probabilistic PV forecasts, the upscaling method of the forecasting errors, the power flow method and the approximation error are considered. In addition, the computational efficiency of the proposed model is scrutinized.

1) Quality of Probabilistic Forecasts

The multivariate probabilistic PV forecasts based on the UMVN, MVN, GC and TC models are compared in the upper left panel of Fig. 3. Thereby, the average variogram-based score with $\gamma = 0.5$ (VS1: $w_{j,l} = 1$; VS2: $w_{j,l} = \rho_{j,l}$, where $\rho_{j,l}$ is the correlation coefficient of the reference current) and the average energy score (ES) are compared.

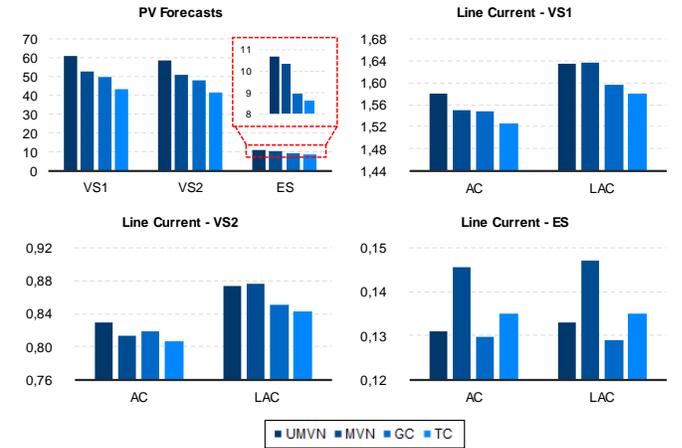


Fig. 3 Variogram-based score (VS1: $w_{j,l} = 1$; VS2: $w_{j,l} = \rho_{j,l}$) and energy score (ES) of the PV forecasts (in MW) as well as the AC and LAC based line current forecasts (in p.u.).

As depicted in Fig. 3 there are significant differences in VS1, VS2 and ES when considering the PV forecasts, indicating the TC to surpass the UMVN, MVN and GC approaches. When using these forecasts for the AC and LAC PPF, eight test cases are obtained. Thereby, the TC still outperforms the other approaches in VS1 and VS2. However, the results for the ES indicate that the MVN and the TC might be mean-biased. Yet the ES is not indicating the improvements of modelling the dependency structure (and the KDE of TC) correctly as emphasized in [29]. Subsequently, the VS is hence considered exclusively. Since the differences of the scores appear to be rather low for both line current forecasts (AC and

LAC PPF), a DM test is carried out. Table II illustrates the resulting test statistics of all possible combinations, whereby a negative sign indicates that the upper model is preferred.

TABLE II
RESULTS OF THE DIEBOLD-MARIANO TEST ON THE VARIOGRAM-BASED SCORE (VS1 & VS2): TEST STATISTIC FOR AC- AND LAC-PPF

Test Case	TC vs.			GC vs.		MVN vs.
	GC	MVN	UMVN	MVN	UMVN	UMVN
AC-VS1	-13.5**	-4.01**	-8.78**	-0.29	-5.69**	-5.8**
AC-VS2	-15.3**	-2.47*	-7.99**	1.67	-3.87**	-5.91**
LAC-VS1	-9.37**	-9.43**	-9.9**	-5.94**	-7.24**	0.38
LAC-VS2	-9.33**	-11.21**	-11.23**	-7.59**	-8.45**	1.61

In this context, the following notation is used indicating the confidence level on which a null hypothesis H_0 of equally performing models is rejected: No asterisk means that the null hypothesis is not rejected, plus sign (+) confidence = 95%, one-asterisk (*) confidence = 99% and two-asterisks (**) confidence = 99.9%. Solely the MVN reveals equal performance to the TC in the AC-VS2 test case

2) Sensitivity to Upscaling of PV Forecasts

To assess the advantages of the proposed upscaling method of the PV forecasts assuming full correlation, a comparative analysis is carried out using the VS. In this context, the results of the upscaling of the PV forecasts assuming *full correlation* of the forecasting errors is benchmarked to a *correlation-based approach*, using the average correlation $\bar{\rho}_k$ of the forecasting errors for each node, and to an upscaling assuming *independent forecasting errors*, i.e., $\bar{\rho}_k = 0$. For this comparison, the results are given in Table III for the full correlation (FC), the correlation-based upscaling (CB) and the independency assumption (IA).

TABLE III
RESULTS FOR DIFFERENT UPSCALING METHOD

Test Case	VS1			VS2		
	FC	CB	IA	FC	CB	IA
TC-AC	1,4961	1,5038	1,5788	0,7893	0,7939	0,8363
TC-LAC	1,5656	1,5689	1,6134	0,8366	0,8381	0,8594

Accordingly, the FC upscaling method (assuming fully correlated forecasting errors at each node) surpasses the correlation-based and independency upscaling methods, for both TC-AC and TC-LAC for VS1 ($w_{ij}=1$) and VS2 ($w_{ij}=\rho_{ij}$). Although the small differences in scores of the FC and CB approaches seem to indicate a comparable forecasting accuracy, the DM test shows that the results are statistically significant at the 99.9% confidence level.

3) Sensitivity to Power Flow Method

Using the DC PPFs as benchmarks for the AC and LAC PPFs, Table IV depicts the results of the TC model in terms of VS1 and VS2.

TABLE IV
RESULTS FOR DIFFERENT POWER FLOW METHODS

Test Case	AC	LAC	DC
TC VS1	1.5250	1.5799	10.8383
TC VS2	0.8062	0.8430	6.0760

As already expected, the results of the PTDF based DC PPF do not come close to the LAC or AC power flow. In opposite,

the VS1 and VS2 results of the TC-LAC are nearby the UMVN-AC results, as illustrated by Fig. 3. When applying a DM test to compare both approaches, i.e., TC-LAC and UMVN-AC, the resulting test statistic t_{DM} are -0.1806 for VS1 and 4.3981 for VS2, respectively. Therefore, both approaches provide equal performance in terms of the VS1, but UMVN-AC surpasses the TC-LAC approach for VS2.

4) Sensitivity to Approximation Errors at Boundary Nodes

The approximation error (AE) can be modelled using the same approaches as used for the forecasting errors. Therefore, each of the approaches TC to UMVN has been carried out for AC- and LAC-PPFs, with and without AE., see Table V.

TABLE V
VARIOGRAM-BASED SCORE (VS1 & VS2) FOR PROBABILISTIC MODELLING WITH (INDICATED AS B) AND WITHOUT APPROXIMATION ERRORS

Test Case	UMVN		MVN		GC		TC	
	VS1	VS2	VS1	VS2	VS1	VS2	VS1	VS2
AC-B	1.581	0.830	1.549	0.814	1.547	0.819	1.525	0.806
AC	1.605	0.845	1.554	0.818	1.542	0.816	1.540	0.816
LAC-B	1.635	0.873	1.637	0.876	1.597	0.851	1.580	0.843
LAC	1.644	0.877	1.639	0.879	1.589	0.848	1.591	0.849

As depicted by Table V, the differences in VS1 and VS2 are rather small, correspondingly the DM test indicates equal performance when comparing the models with AE with those models without AE. In the GC test case, the model without AE even surpasses the model with AE, albeit the difference is not statistically significant.

5) Computational Efficiency

The case study has been carried out using MATLAB / MATPOWER on a Windows Server 2012 R2 Standard with 64-bit operating system, a two kernel 3.50 GHz processor (Intel® Xeon® CPU E5-2637 v2) and 192 GB RAM of memory. Table VI compares the computational efficiency for the different PPF approaches. Thereby, the MCS of the AC-PPF have been carried out sequentially and parallelized.

TABLE VI
COMPARISON OF COMPUTATIONAL TIME FOR LAC-PPF AND AC-PPF APPROACHES (PARALLELIZED AND SEQUENTIAL)

Computational Time	MCS AC-PPF		LAC-PPF
	sequential	parallel	sequential
Average per hour	24.02 s	16.21 s	0.04 s
Total (2015 h)	7.84 h	5.29 h	50.84 s

The results for the computation time of the parallelized MCS AC-PPF opposed the expectations: The additional effort for distributing and allocating large variables (type: 2,015 hours \times 19 lines \times 5,000 samples, i.e., 1.53 GByte per variable) to the virtual kernels, exceeds the gains of parallel computation. Furthermore, the gains of the LAC-PPF are remarkable, taking less than one minute for computing more than 10^6 scenarios.

VII. CONCLUSION

This paper proposes a novel approach to provide probabilistic congestion forecasts in high voltage power systems utilizing multivariate copula-based RES-infeed simulations. The improvements presented contribute to increase both, the performance of the congestion forecasts

(upscaling of forecasting errors, reduction of heteroscedasticity, modelling of boundary nodes) and the computational efficiency (application of linearized power flows). The benefits of the proposed model are demonstrated on a real-world case study against commonly used benchmark models. Thereupon, this approach enables system operators to predict potential congestions at an early stage, taking the increasing number of uncertainties into account, in order to derive the necessary countermeasures.

VIII. REFERENCES

- [1] U. K. Das, K. S. Teya, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. van Deventer, B. Horan and A. Stojcevski, "Forecasting of photovoltaic power generation and model optimization: A review," *Renewable and Sustainable Energy Reviews*, Vol. 81, No. 1, pp. 912-928, 2018.
- [2] D. van der Meer, W. J. and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," *Renewable and Sustainable Energy Reviews*, Vol. 81, pp. 1484-1512, 2018.
- [3] R. J. Bessa, A. Trindade and V. Miranda, "Spatial–Temporal Solar Power Forecasting for Smart Grids," *IEEE Transactions on Industrial Informatics*, Vol. 11, No. 1, pp. 232 - 241, 2015.
- [4] Y. Gu and L. Xie, "Fast Sensitivity Analysis Approach to Assessing Congestion Induced Wind Curtailment," *IEEE Transactions on Power Systems*, Vol. 29, No. 1, pp. 101-110, 2014.
- [5] L. Hirth and S. Glismann, "Congestion Management: From Physics to Regulatory Instruments," *Leibniz Information Centre for Economics*, 2018.
- [6] D. Biggar and M. Hesamzadeh, *The economics of electricity markets*, IEEE Press, John Wiley & Sons, 2014.
- [7] P. Chen, Z. Chen and B. Bak-Jensen, "Probabilistic Load Flow: A Review," *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008.
- [8] X. Deng, P. Zhang, K. Jin, J. He, X. Wang and Y. Wang, "Probabilistic load flow method considering large-scale wind power integration," *Journal of Modern Power Systems and Clean Energy*, Vol. 7, pp. 813-825, 2019.
- [9] X. Lu, K.-C. Lam, H. Xu and Z. Jia, "Improved Probabilistic Load Flow Method to Consider Random Generator Outages," *Advanced Materials Research*, Vol. 962-965, pp. 2783-2788, 2014.
- [10] H. Bao, H. Wei and X. Guo, "Solutions to probabilistic analysis of voltage stability in power systems with wind farms using advanced unscented transformation," *International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2016.
- [11] H. Sheng and X. Wang, "Probabilistic Power Flow Calculation using Non-intrusive Low-rank Approximation Method," *IEEE Transactions on Power Systems*, Vol. 34, No. 4, pp. 3014 - 3025, 2019.
- [12] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Power & Energy Society General Meeting*, 2009.
- [13] A. Klettke, I. Petkov and A. Moser, "Modelling of Stochastically Dependent Forecast Errors for Load Flow Simulations in the Transmission Grid Using Convolution," *21st European Conference on Power Electronics and Applications (EPE '19 ECCE Europe)*, 2019.
- [14] S. Chai, M. Niu, Z. Xu, L. L. Lai and K. P. Wong, "Nonparametric conditional interval forecasts for PV power generation considering the temporal dependence," in *IEEE Power and Energy Society General Meeting (PESGM)*, Boston, 2016.
- [15] M. Hojjat and M. H. Javidi D. B., "Probabilistic Congestion Management Considering Power System Uncertainties Using Chance-constrained Programming," *Electric Power Components and Systems*, pp. 972-989, 2014.
- [16] L. Roald und G. Andersson, „Chance-Constrained AC Optimal Power Flow: Reformulations and Efficient Algorithms,“ *IEEE Transactions on Power Systems*, Vol. 33, No. 3, pp. 2906 - 2918, 2018.
- [17] A. Schinke-Nendza, F. von Loeper, P. Osinski, P. Schaumann, V. Schmidt and C. Weber, "Probabilistic forecasting of photovoltaic power supply - A hybrid approach using D-vine copulas to model spatial dependencies," *unpublished working paper*, 2021.
- [18] Y. Saint-Drenan, G. Good and M. Braun, "A probabilistic approach to the estimation of regional photovoltaic power production," *Solar Energy*, Vol. 147, pp. 257-276, May 2017.
- [19] H. M. Shafey, "Prediction of Atmospheric Parameters for Spectral and Total Solar Irradiances," in *Proceedings of ENERGETEX '84, The Global Energy Forum*, Regina, Saskatchewan, Canada, 1984.
- [20] A. Sklar, „Fonctions de épartition à n dimensions et leurs marges,“ *Publications de l'Institut de Statistique de l'Université de Paris*, Vol. 8, pp. 229-231, 1959.
- [21] F. Golestaneh, H. B. Gooi and P. Pinson, "Generation and evaluation of space–time trajectories of photovoltaic," *Applied Energy*, Vol. 176, pp. 80-91, 2016.
- [22] A. McNeil, R. Frey and P. Embrechts, *Quantitative risk management: concepts, techniques and tools*, Princeton: Princeton University Press, 2005.
- [23] S. Demarta and A. J. McNeil, "The t Copula and Related Copulas," *International Statistical Review*, Vol. 73, Issue 1, pp. 111-129, April 2005.
- [24] P. D. Hill, "Kernel estimation of a distribution function," *Communications in Statistics - Theory and Methods*, Vol. 14, Issue 3, pp. 605-620, 1985.
- [25] B. Silverman, *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall/CRC, 1986.
- [26] D. Cremer and C. Zimmer, "Improved modelling of external networks in operation and operational planning of interconnected power systems," *Electrical Engineering* Vol. 83, pp. 291-296, 2001.
- [27] I. Talavera and e. al, "Vertical Reactive Power Flexibility based on Different Reactive Power Characteristics for Distributed Energy Resources," *International ETG Congress 2015*, pp. 42-48, 2015.
- [28] R. Rubinstein und D. Kroese, *Simulation and the Monte Carlo Method*, Hoboken: Wiley, 2017.
- [29] M. Scheuerer and T. Hamill, "Variogram-Based Proper Scoring Rules for Probabilistic Forecasts of Multivariate Quantities," *American Meteorological Society - Monthly Weather Review*, Vol. 143, p. 1321–1334, 2015.
- [30] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, pp. 121-151, 2014.