Risk-based security assessment of transmission line overloading considering spatio-temporal dependence of load and wind power using vine copula

Khuntia, Swasti R.; Rueda, José L.; van der Meijden, Mart

DOI
10.1049/iet-rpg.2018.6091

Publication date
2019

Document Version
Accepted author manuscript

Published in
IET Renewable Power Generation

Citation (APA)
Khuntia, S. R., Rueda, J. L., & van der Meijden, M. (2019). Risk-based security assessment of transmission line overloading considering spatio-temporal dependence of load and wind power using vine copula. *IET Renewable Power Generation*, *13*(10), 1770-1779. https://doi.org/10.1049/iet-rpg.2018.6091

Risk-based security assessment of transmission line overloading considering spatio-temporal dependence of load and wind power using vine copula

Swasti R. Khuntia [1*], Jose L. Rueda [1], Mart A.M.M. van der Meijden [1,2]

[1] Department of Electrical Sustainable Energy, Delft University of Technology, The Netherlands
[2] TenneT TSO B.V., The Netherlands
[*] s.r.khuntia@tudelft.nl

Abstract: Location of wind power plants and demand centres are not always close by; hence, the transmission of energy puts a burden on existing grid infrastructure. This unwanted burden necessitates transmission lines to operate more and more frequently close to their operating limits. To alleviate such situations, this research addresses the advantages of modelling spatio-temporal dependence of load and wind power using vine copula. Probabilistic AC optimal power flow is performed on a modified IEEE 39-bus system with significant wind penetration. Real load and wind power data from a U.S. utility is mapped onto the test-case to achieve realistic results. Load flow calculation can help in performing steady-state voltage and overload evaluations for post-disturbance system conditions. Because the security level of a power system is determined by the likelihood and severity of security violation. In this research, the probability of line overload is calculated from load flow and the severity function describes the risk of line overloading. Two case studies depicting future operating conditions of massive wind power penetration with reduced fossil fuel and nuclear power generation are considered. Simulation results prove the advantage of addressing spatio-temporal dependency to quantify the overload risk index, which is treated as a security indicator.

1. Introduction

The electric power system infrastructure is subjected to increasing stress due to fundamental changes in both generation and demand side. A larger picture of such stress is encountered in the form of transmission loading pattern.

- On the *generation* side, integration of stochastic generation sources in the form of Renewable Energy Sources (RES) is a challenging task for Transmission System Operators (TSOs). Among the existing RES, wind power has gained significant attention among TSOs because of three reasons, namely, (i) *large Wind Power Plants (WPPs) can be connected to bulk power system at the transmission level*, (ii) *large size WPPs are being built/planned in regions with high potential for wind energy and TSOs have to facilitate their integration*, and, (iii) *TSOs are more and more limited by local constraints to build new transmission infrastructure in time*. This comes at a time when the electric power industry is undergoing an energy transition, especially the increased penetration of RES and decentralized generation while discarding fossil fuels to achieve a greener future in the form of a low-carbon power system. Wind power is characterized by variability and uncertain nature, and its generation capacity is dependent on geographic location. Over the last decade, electricity production from WPPs has reached significant levels in several regions of Europe and the U.S. [1]-[3]. The European Wind Energy Association expects 50% of the electrical energy demand to be met by wind energy by 2050 [4]. In addition to wind on land, a substantial proportion of wind capacity from the North Sea is anticipated: 150 GW by 2030 and 350 GW by 2050. In [5], it was learned that expansion of WPPs in terms of farm-size and unit capacity is significant at transmission level and its integration possesses challenge for TSOs in terms of infrastructure management, operation, and security. The participation of WPPs into existing grid is different from conventional generators in terms of location and output generation which is uncertain and variable. It adversely affects day-ahead operational planning decisions which introduce a level of risk for TSOs. Variation in wind power hamper power system operation in real-time when WPPs are unable to deliver the required reserve capacities in real-time. Embedding of WPPs raise concerns in terms of planning and upgrading of existing infrastructure in terms of size, location and distribution of WPPs [6].
- On the *demand* side, the advent of new technologies and a growing number of variable generation sources at the distribution level is challenging itself. The traditional power system network was designed for the passive load without any plans for communication and digitalization. It is supplemented with the spatial distribution of demand centres which are not always located close to generating sites.

As the location of demand centres and WPPs are not always close-by, the non-dispatchable sources cannot be easily managed or curtailed. This spatial dispersion imposes a burden on the TSOs who have to operate the existing infrastructure with uncertainty from both ends, i.e., demand and generation. An example is the country of Germany, where the wind power generation is concentrated in the northern part, solar power generation is concentrated in the southern part and load centres are mostly in the mid-western and southern parts of the country [7]. The stochastic power with spatial diversity is causing destabilization of the electric grids (e.g., potential blackouts, weakening voltage). It can be deduced that though RES reserves are significant, their location is non-uniformly distributed and often far from load centres. As the location of WPPs and load centres are not always close by, transmission of energy puts a burden on

transmission links in the existing grid infrastructure. Complexity in terms of inter-spatial dependence and temporal correlation of load and wind power, therefore, impose a challenging operational threat to TSOs. In such cases, the system is being asked to perform in ways it was not designed for, eventually resulting in performance under increasing stress. As the system becomes heavily loaded and vulnerable to disturbances, security of supply is at risk. It calls for an overload analysis. Literature study reveals some commonly used indices like overload, cascading overload, low voltage and voltage instability [8]. To account for the mentioned developments in reliability management, TSOs need to re-evaluate the system security [9].

Though achievements have been made in terms of an efficient forecast of future load and wind power generation, there are other vital concerns corresponding to wind power such as spatio-temporal dependence, variability, non-normality, non-stationarity, non-dispatchable (unless there is adequate storage) and seasonal patterns to name a few. Wind speed is temporally correlated at one location and for different locations wind speed is both spatially as well as temporally correlated [10]. A statistical space-time model considering terrain, wind speed and direction was proposed in [11]. Spatial dependence of wind for transmission line overloading is found in [12]. It is to be noted that spatial dependency within multiple wind-farms as studied in [13] is out of the scope of this research. Addressing temporal correlation for both load and wind power can be found in [14-16]. To the best of our knowledge, no study has reported the importance of addressing spatio-temporal dependence of load and wind power till date. It is important to address the spatio-temporal dependence from transmission system point of view when TSOs are facing stagnant expansion planning because (i) *weak transmission capacity causes reduced integration capacity of WPPs* and (ii) *redundant transmission capacity results in resource waste*.

In such cases, the risk of transmission line overloading and voltage instability cannot be avoided in future. And to overcome the burden of transmission line overloading, this research will facilitate TSOs to operate the grid within security limits considering spatio-temporal dependency. The key contributions of this research are twofold:

i. A novel attempt to assess transmission line overloading risk while addressing spatio-temporal dependency using vine copula is made in this research. The sampling algorithm uses real data from nineteen spatially distributed load and two wind power zones spanning three years horizon from a U.S. utility to model the joint probability distribution.
ii. The sampled output is mapped onto a modified IEEE 39-bus system to achieve realistic results based on two case studies representing future scenarios of high wind power penetration with reduced fossil fuel and nuclear power generation. RBSA is performed on transmission line overloading by performing probabilistic AC optimal power flow (OPF) and considering spatio-temporal dependence of load and wind power. Risk quantification of overloading is achieved as a product of probability and severity of overload.

The rest of the paper is organized as follows: Section 2 presents a background on risk-based security assessment (RBSA) studies. Section 3 introduces vine copula for spatio-temporal modelling and describes the sampling algorithm. Section 4 focuses on database generation and preparation for RBSA studies. Section 5 discusses the two case studies along with result analysis. Finally, section 6 concludes the research work.

2. Background on risk-based security assessment (RBSA) studies

Bulk interconnected power systems with distributed and geographically isolated generators and demand centers constitute a majority of the power network. With increasing RES and other DERs, the present day power systems are dynamic in nature with network topology changing more and more frequently with the change in demand. As nuclear power plants and fossil fuel plants are phased out to include more RES in form of massive wind power penetration, uncertainty in load demand actuates the power network to operate at loading limits; thus, making it susceptible to blackout under minor/major disturbances. In order to operate the power system economically, the state of the system has to be identified as secure/insecure. Security assessment studies aim to balance the system security as well as the economy for power system operation. Power system security can be divided into two, namely, static and dynamic security. Static security analysis targets steady-state post-disturbance conditions, namely it is assumed that the system reaches operating equilibrium after a disturbance and it is checked whether system limits are violated. Dynamic security analysis targets system stability after a disturbance, and therefore it is investigated whether the system can reach a new state of equilibrium after a disturbance. Sometimes static security reliability assessment can be referred by literature as adequacy assessment, and dynamic security reliability assessment can be met simply as security reliability assessment. A classification hierarchy is shown in Fig. 1. Another way of classification is based on assessment techniques. The three schemes of security assessment are, namely, (i) deterministic security assessment that is more traditional and considers a set of most credible contingencies resulting in high operating costs, (ii) probabilistic security assessment that considers probabilistic indices LOLP (Loss of load probability for likelihood of events) and EENS (Expected energy not supplied for both likelihood as well as severity of events), and (iii) risk-based security assessment (RBSA) that considers both likelihood and severity of events allowing the power system to operate closer to or beyond its limits. Further, RBSA is categorized as; (i) static RBSA that considers the risk of overload and voltage violations and (ii) dynamic RBSA that considers the risk of instability in terms of voltage and swing transient. The deterministic security assessment methods get usually many conservative results in overload analysis. With recent advancement and more adoption of risk theory in power system, the risk assessment method is gradually evolving and acknowledged.

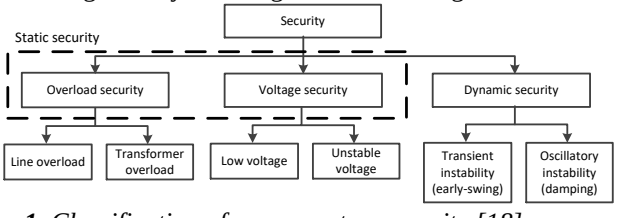


Fig. 1. *Classification of power system security [18]*

Risk assessment by computing risk indices based on over-limit probability and severity to recognize system weakness more realistically is entailed in this research. In regard to transmission line overload, literature study reveals the study of static RBSA with 'N-1' contingency [17], risk visualization using Poisson distribution [8, 18], online static RBSA with forecasted operating condition [19], possibility and severity of risk occurrence [20]. Risk assessment is performed in four steps [8]: (i) *describe an index that represents system risk*, (ii) *select a system state and calculate its probability*, (iii) *evaluate the outcome of the system state*, and (iv) *calculate the risk index*. For any operating condition, the risk associated with ith state S_i at time t is calculated for all possible values of probability and severity associated with it and can be written as [18],

$$Risk_t = \sum_{i=1}^{n} Prob_t(S_i) Sev_t(S_i) \quad (1)$$

where $Prob_t(S_i)$ is the state probability, $Sev_t(S_i)$ is the associated severity of state i and n is the total number of system states. The line overload possibility can be measured by the probability mass function of line flows. Ref. [18] categorized severity function into three types, namely, discrete, continuous and percentage of rating violation severity function. The concept of severity functions has been used in recent studies [11] to investigate transmission line overloading in power systems with wind and load-power generation correlation.

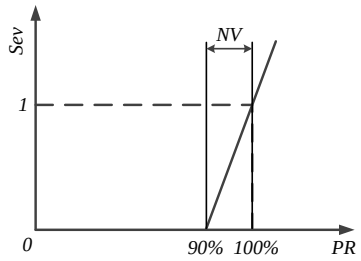


Fig. 2. *Linear severity function of line overload (PR: Percentage of Rating, NV: Near Violation)*

Fig. 2 shows the severity function as a continuous function of line power flow as the percentage of rating (PR) and severity (Sev), written as a function of j-th branch apparent power flow F_j and rated apparent power flow F_j^{max},

$$PR = \frac{|F_j|}{F_j^{max}} \times 100\%$$
$$Sev_j = \begin{cases} 0, & (PR/100) < 0.9 \\ 10(PR/100) - 9, & (PR/100) \geq 0.9 \end{cases} \quad (2)$$

When the line power flow exceeds 90% of its rating, near violation for overload takes place which increases linearly as power flow exceeds the limit. Such a linear relationship is related to the fact that saturation effect on machines and transformers are neglected, line impedances are predominantly inductive (for high voltage transmission network, which is the focus of this paper), voltage magnitudes are around nominal value, and there is minor angle difference in bus voltages, which lead to a linear dependency between active power transfer through a transmission line and the difference between bus voltages. Severity function signifies the extent of security violation and helps in quantifying the severity. In case of the transmission line, severity function is defined for each line and the power flow in the line determines the associated risk. For each line, the severity function defined in eq. (2) evaluates to 1 at the deterministic limits, i.e., 100% of line apparent power flow rating. It is to be noted that higher risk values do not necessarily indicate a larger interruption of security of supply and vice-versa. For example, a 130% overload of a transmission line might have higher risk value than a 130% overload of another transmission line but it does not necessarily lead to increasing the penalty incurred due to unavailability of security of supply.

In this research, calculation of risk indices is accomplished with probabilistic AC load flow. As such, the accuracy of computed risk indices depends on the accuracy of load flow results. Load flow analysis provides an effective tool to analyse the relationship between the bus injection fluctuation in terms of near violation and overload in system operation state. Presence of WPPs introduces uncertainty in the normal operating state. In addition to wind power, uncertainty in load contributes equally towards the problem. Hence, the line flows are eventually influenced by the bus injections. We adopt a multivariate analysis of addressing the uncertainty as explained in the next section.

3. Vine copula for spatio-temporal modelling and joint normal transformation

It was learned that both load and wind power vary across the year and depend on the geographical area. To capture the inter-spatial dependence and temporal correlation between the variables, a joint probability distribution is defined using vine copula. Copula functions are the most general method for dealing with dependency between variables. Learning dependencies over correlation is important to study spatio-temporal dependence. Use of vine copula is not new in the field of electric power systems, but to our knowledge, it has not yet been applied for spatio-temporal modelling of load and wind power to assess transmission line overloading risk. Literature study reveals the use of vine copulas to tackle power system uncertainty [21-22]. Use of copula theory in RBSA is reported in [23].

Vine copula follows a nested tree structure where every node in the tree is a bivariate (two dimensional) copula. Conditional on the lower tree, the vine structure is then built up based on conditional, bivariate copulas. Thus, vine copulas allow to flexibly combine bivariate copulas to multivariate copulas leading to distributions of higher dimensions that are aware of separating distances across space and time. By definition [24], a vine copula on n variables is a nested set of trees T_j where the edges of the j^{th} tree become the nodes of the $(j+1)^{st}$ tree for $j = 1, \ldots, n$. A regular vine on n variables is defined as a vine in which two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node. Each edge in the regular vine may be associated with a conditional rank correlation and a copula, and each node with a marginal distribution. A regular vine can be either D – vine where each node in T_j has a degree of at most 2 or C (canonical) – vine in which each tree T_j has a unique node of degree $n - i$. Fig. 3 shows the C-vine of five trees. A C-vine selects a root node in each tree, and all pair-

wise copulas connecting with this node are modeled and conditioned on all of the previous root nodes. Such a modeling scheme is appropriate for spatio-temporal modeling of complex multivariate dependence structures with mixed types of dependencies, such as asymmetries and tail dependencies, since each pair-copula can belong to a different parametric copula function.

Vine copulas approximate multivariate copulas through bivariate building blocks (Fig. 3). The advantage of the spatio-temporal vine copula is its flexibility in the selection of copula families through bivariate spatio-temporal copulas. Bivariate spatio-temporal copulas are a convex combination of different copula families that are parameterized by spatial and temporal distance (Fig. 4 and equation 3). This changing dependence structure allows for instance to preserve spatial symmetry within each time step while allowing for a directional effect across time. Such an aspect is deemed important for improvement of risk calculation (e.g. less possibility of overestimation).

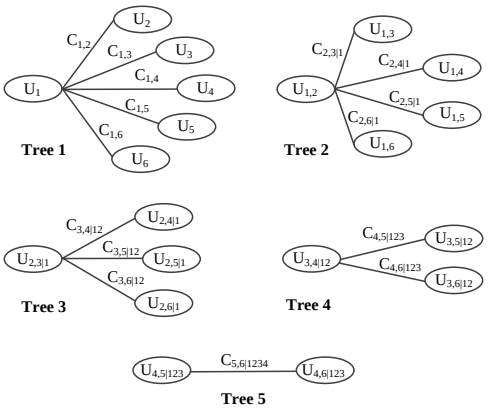


***Fig. 3**: C-vine of CDF ($u \in [0,1]$) with five trees*

The building blocks of the vine copula are composed out of convex combinations of bivariate copulas. The weight of the convex combination, as well as the copulas' parameters, are defined by the distance over space and time, thus modelling spatial and temporal correlation. The key benefit of copula functions resides in the separation between the dependence structure and the univariate distributions of the variables. This property allows copula functions to model dependencies for any type of distribution functions. Copula functions separate the dependency structure from the standalone distribution functions followed by each variable. To understand the spatio-temporal modeling using vine copula model, a spatio-temporal random field H is considered such that

$$H: S \text{ x } T \text{ x } \Phi \longrightarrow \mathbb{R} \qquad (3)$$

where S corresponds to the spatial domain, T corresponds to temporal domain and both with an underlying probability space Φ. For a section of the spatio-temporal random field defined as $H = (h(s_0, t_0), h(s_1, t_1) \dots h(s_n, t_n))$ of size $n + 1$, the section consists of one pivotal location and its n-neighbors in distinct spatio-temporal locations $(s_0, t_0), (s_1, t_1) \dots (s_n, t_n) \in S \text{ x } T$. Normally some spatial locations would be sampled at multiple time instances. And as the dependence structure changes over space and time, the first tree of the vine is realized by spatio-temporal bivariate copulas. The rest of the vine, i.e. the vine of the variables conditioned under the value of the central location, is modeled as some n-dimensional C-vine. To understand the functional capability of C-vine, Fig. 4 shows the spatio-temporal n-dimensional C-vine. The temporal extension of the spatial copula at different time lags for 3 spatial locations with Euclidean distance defined as $h_E \coloneqq \|s_i - s_j\|$, $s_i \; \forall \, i, j \in \{0,1,2,3\}$ & $t_C = 1 \dots n$. Every curved connection is modeled by the same spatio-temporal copula C_{h_E, t_C} but with different spatial and temporal distances, h_E and t_C deduced from the indicated spatio-temporal locations. It is already assumed that marginals are stationary and combining them with multivariate copula results in a multivariate distribution of the spatio-temporal random field. And this multivariate distribution is later used for application studies like simulation or prediction.

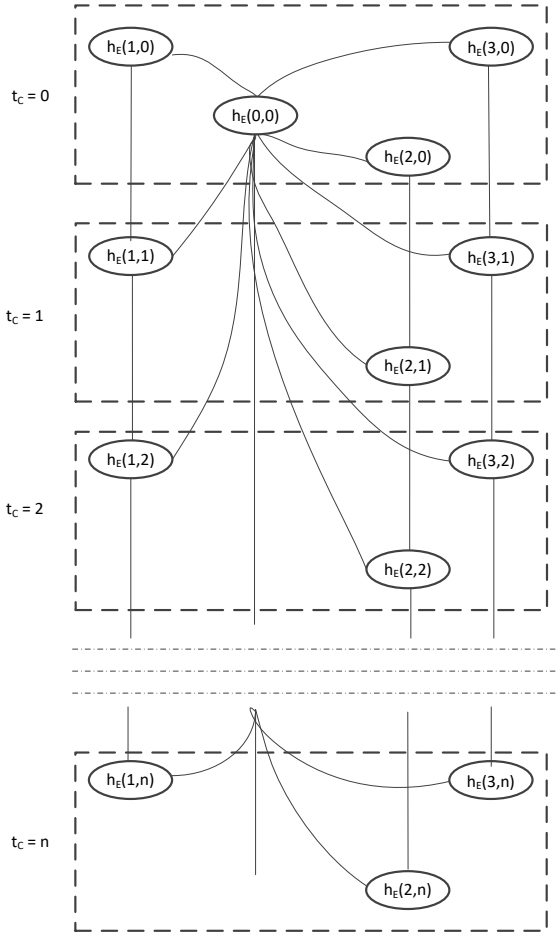


***Fig. 4:** Spatio-temporal n-dimensional C-vine*

The sampling algorithm employed in this research is inspired by ref. [25] and is described in *Algorithm 1*. Considering a high-dimensional dataset of size (M x N) representing M temporal data points and N spatially distributed variables (load and wind power sites). The first step is transforming the high-dimensional data to low-dimension to ease the computational bottleneck. This is achieved by clustering and followed by feature extraction. Selection of clustering technique is based on the goodness of clustering test and feature extraction is performed using singular value decomposition (SVD) [26]. SVD is preferred over widely used principal component analysis because of its

efficient and robust numerical methodology. Feature extraction is followed by dependence modelling using vine copula. In order to evaluate the performance of vine copula on the basis of two-sample tests, a resampling method is employed to randomly generate comparison samples from the historical and simulated datasets. Sets of historical and simulation datasets based on cluster size and weight were drawn randomly, and the process was repeated for each of the variables. Performing inverse ECDF transformations on the sampled output to retrieve high dimensional data is achieved at the end of this algorithm. For a chosen sample size S, the output sample is a joint normal distribution of dimension S x N. Implementation of the algorithm with real data for RBSA is explained in the next section.

Algorithm 1:	Modelling high-dimensional spatio-temporal dataset using vine copula
	Inputs: High dimensional dataset of size (M x N), representing M data points and N features.
	Outputs: S x N dimensional sampled dataset
1	Perform clustering to partition the high dimensional data. k number of clusters are selected after performing goodness of clustering (GoC) test on sample size S
2	Feature extraction of k clusters using singular value decomposition (SVD)
3	Calculate copula function and construct vine copula models for k clusters. Choice is different copula functions can be tested and goodness of fit (GoF) test is performed to select the best copula function
4	Simulate the copula function for k clusters using cluster weight obtained in Step 1
5	Reconstruct dataset from low to high dimension with all features using eigenvectors from Step 2
6	Output as S x N dimensional sampled dataset
7	**End**

4. Database generation and preparation

For this research, publicly available load and wind power data are taken from U.S. regional transmission operator [27]. Aggregated zonal load data (nineteen numbers) and wind power data (two numbers) spanning three years with an hourly resolution from the three market zones (*MIDATL*, *WEST* and *SOUTH*) are used. To study the spatial correlation, geographical coordinates of zones are needed. Since the exact coordinates are treated confidential, an approximated weighted centroid is defined to locate the 'centre' of load and wind power generated zones. More details on data collection can be found in the supplementary material [28].

To visualize the complexity, scatter plot with marginal histograms of four load (*AP*, *CE*, *DAY*, *DUQ*) and one wind power (*WEST*) zone under the *WEST* zone is shown in Fig. 5. The marginal histograms (in the diagonal) reveal non-Gaussian nature while the scattered plots reveal non-linear dependencies and suggest a weak correlation as well. Use of copula functions allow modelling dependency between variables that do not follow the same distributions, inclusive of non-normal distributions, which was evident from scatter plot of load and wind power in Fig. 5.

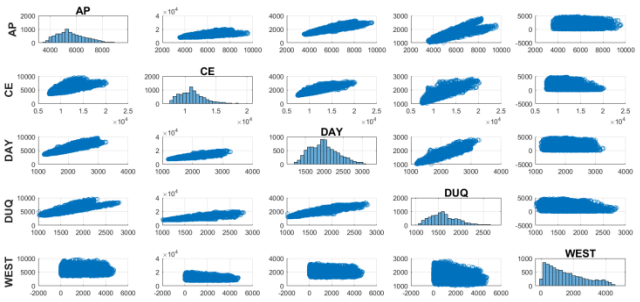


Fig. 5: *Scatter plot with marginal histograms of original data of four load zones (AP, CE, DAY, DUQ) and one wind power zone (WEST)*

The proposed risk assessment is validated using the modified IEEE 39-bus system with a base value of 100 MVA. A modified version of the IEEE 39-bus system with WPPs and updated conventional generation capacities is considered in this research. The original test case consists of 39 buses, 10 conventional generators and two WPPs at bus 34 and 37, and 46 transmission lines with net demand in the network of $6254.23MW$ and net generation capacity of $7367MW$. The modified topology with wind farms and divided zones according to real data is shown in Fig. 6. Next step is to map the real-life data onto the IEEE 39-bus system by scaling the real-life data to match test-case parameters. A scaling ratio is defined as,

$$Scale\ Ratio = \frac{\text{maximum coincident peak demand of real data}}{\text{sum of active power demand across all buses in test} - \text{case}} \quad (4)$$

Considering large-scale wind power penetration in future as part of energy transition, generation capacity of conventional generation sources (such as nuclear and fossil fuels) is lowered [29-30] and is met by wind power generation. Generation cost data is obtained from [31]. In addition, there is no topological change in terms of addition of transmission lines which gives us the option to assess the overloading risk on existing grid network. The two cases studied in this research are:

Case A: The first case considers a 7.5% increase in system load to $6725.05MW$. The total generation is $8167.87MW$ including $2640MW$ of wind at bus 34 and 37 respectively. For *MIDATL*, the net load is $2355.13MW$ and net generation is $4877.871MW$ including $1760MW$ of wind power. For *WEST*, the net load is $3752.1MW$ and net generation is $3500MW$ including $880MW$ of wind power. For *DOM*, there is only an increase in demand to $617.82MW$.

Case B: The second case considers a 9% increase in system load to $6817.11MW$. The total generation is $7712.86MW$ including $2760MW$ of wind. For *MIDATL*, the net load is $2447.19MW$ and net generation is $3792.861MW$ including $1760MW$ of wind power. For *WEST*, the net load is same as *Case A* and net generation is $3360MW$ including $1000MW$ of wind power. For *DOM*, the net load and generation is same as in *Case A*.

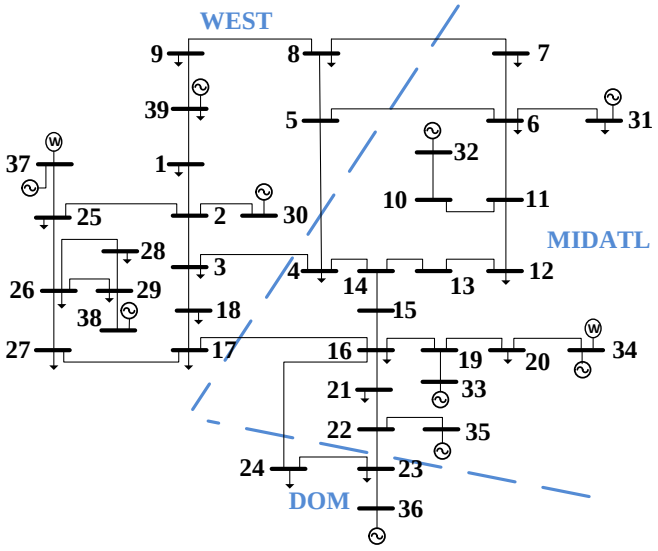


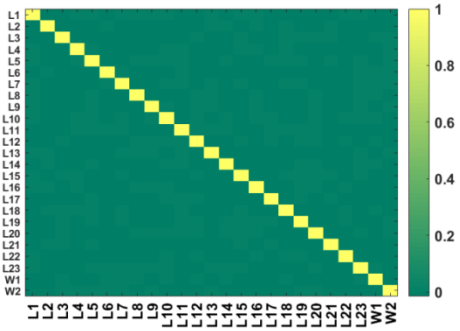
Fig. 6: *Modified IEEE-39 bus system divided into three market zones (MIDATL, WEST, DOM)*

For the two cases, a comparative analysis of non-correlated and correlated load and wind power samples is performed. In the case of non-correlated samples, synthetic wind power data using Weibull distribution and random samples of load taking the load at each bus as mean values is considered. This also serves as benchmark data to assess the correlated samples. The formula for the probability distribution function (PDF) of wind power is expressed mathematically as:

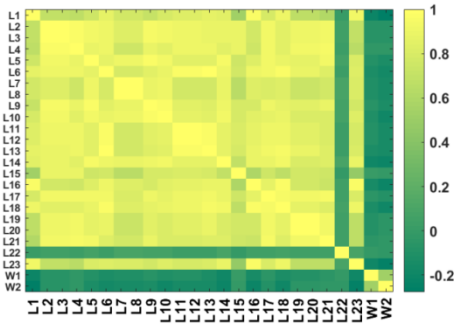
$$f(v;\gamma,\alpha) = \frac{\gamma}{\alpha}\left(\frac{v}{\alpha}\right)^{\gamma-1} exp\left\{-\left(\frac{v}{\alpha}\right)^{\gamma}\right\} \qquad (5)$$

$$v > 0, \gamma > 0, \alpha > 1$$

where, v is the wind speed, γ is the shape parameter and α is the scale parameter. Though α significantly depends on wind farm location, we consider a single value as the sole intention of this research work is to generate aggregated zonal wind power and thereby not considering different wind farm sizes. For the non-correlated cases, $\gamma = 2$ and $\alpha = 11$ is used to generate non-correlated wind power for the probabilistic AC OPF [32]. Similarly, random load samples are generated from the mean values of load at each bus [31]. Technically, generated samples of load and wind power are non-correlated and a correlation plot for *Case A* is shown in Fig. 7(a).



(a)



(b)

Fig. 7: *Correlation plot for Case A with* ***(a)*** *non-correlated load (L1…L23) and wind power (W1,W2);* ***(b)*** *correlated load (L1…L23) and wind power (W1,W2)*

For generating correlated samples, real data is modeled using Algorithm I to generate a joint normal distribution with correlated samples. The correlation plot for *Case A* is shown in Fig. 7(b). Inferencing the plot suggests both positive and negative correlation. The correlation coefficients -0.2675, 0 and 1.0 represent slightly negative correlation, perfectly uncorrelated and perfectly correlated. It is important to understand the significance of correlation coefficients. If load and wind power generation were positively correlated, they would tend to increase and decrease at the same time, and adding wind would help the load following task of the power system. On contrary, if the correlation were negative, the wind would tend to decrease when load increases (and vice versa) and this would require more from the load following units in the system. The presence of a negative correlation between load and wind power is of physical significance. It explains the need to balance out the wind power fluctuations in different zones with corresponding load fluctuations to maintain a steady supply.

5. Results and discussions

For result analysis, it is important to note that the samples from vine copula does not replace the need of considering ramping constraints and start-up/shut-down time of generators across time. In addition, the "temporal" in the context of vine copula does not mean including/neglecting inter-temporal implications of ramping constraints. The representation of load and wind power samples as multiple scenarios is applied to perform an offline (or post-mortem) study of transmission line overloading. For the inter-spatial dependencies and temporal correlations between load and wind power, two extremes of *non-correlated* and *completely correlated* samples are considered. Such extreme scenario evaluations will enable us to gain insight into the importance of considering correlation for assessing transmission line overloading risk. Thus, for each *Case A* and *B*, there are two cases of non-correlated and correlated samples. All modeling work is performed in MATLAB (version 2017b) environment using the Matpower package [33] on an Intel Core i7 with 8 cores and 8GB RAM. Matpower's deterministic power flow is used to implement the probabilistic OPF. Probabilistic OPF is built based on a deterministic one because the OPF that Matpower uses is a deterministic OPF. In general, the function of deterministic power flow study can be stated as,

$$z = g(x) \qquad (6)$$

where x is the vector of input variables which includes active power injection P_i at each bus, reactive power injection Q_i at each PQ bus and voltage magnitude V_i at each PV bus and slack bus; z is the vector of output variables which include bus voltage V_i at each PQ bus, bus angle θ_i (except slack bus), branch active power flow P_{ij}, reactive power flow Q_{ij}, and apparent power flow S_{ij}. For the probabilistic power flow problem, the input random variables $x_1, \dots, x_K$ are probabilistic distributions of P_i and Q_i. When wind power is included in probabilistic power flow, an additional random input variable is introduced as the wind power of the WPPs. The output information is probabilistic distributions of V_i , θ_i , P_{ij}, Q_{ij}, and S_{ij}.

The problem of OPF is the allocation of given load amongst the generating units in operation so that the overall cost of generation is minimum. In OPF, the entire set of equality and inequality constraints, all the necessary and sufficient conditions of control parameters etc. must be satisfied thoroughly. The objective function can take various forms such as fuel cost, transmission losses, and reactive sources allocation. The objective function of interest in this research is the minimization of the total production cost of scheduled generating units. Or in other words, it is a simple economic dispatch study where the OPF calls PF calculation function in each iteration. Various techniques have been proposed to solve the OPF problem, for example, non-linear programming, quadratic programming, linear programming, and interior point methods. In this research, the interior point method is used to solve the OPF problem by using Matpower Interior Point Solver (MIPS). For a thorough understanding of MIPS, readers are referred to [34]. A flowchart showing data collection, dataset preparation, generating correlated samples using *Algorithm 1* and running OPF with the calculation of risk indices is shown in Fig. 8.

In this study, temporal dependency is accounted in the samples obtained from the vine copula sampling. Thus, when we use the samples for AC OPF using Matpower, we are running scenarios which are spatio-temporally modeled. All lines are monitored for overload risk though special attention is given to line connecting WPPs and the next nearest bus to rest of the grid, i.e., line 20-34 and line 25-37 for the modified IEEE 39-bus system. Both the lines have a rated capacity of 900MVA. The impacts of load and generation correlation on line overload risk are studied. It can be understood that fluctuation of high wind power can be easily compensated by the grid, provided it is distributed among the strong lines connecting to immediate load sites or the presence of a conventional generator bus that can regulate its generation depending on the needs of high or low wind power generation.

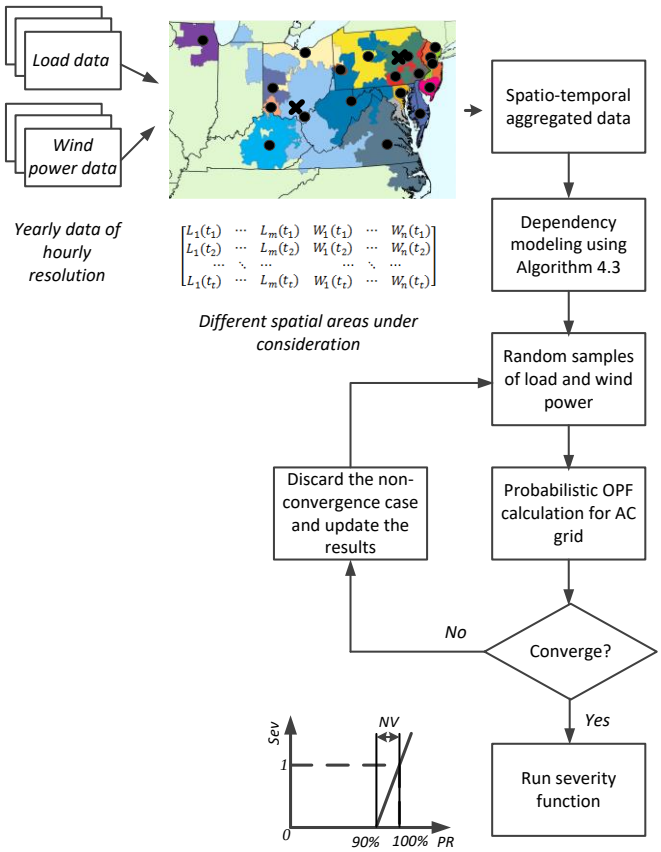


Fig. 8: *Flowchart showing the computation of severity function starting from data collection to running power flow. The map shows the control areas of PJM with load ($L_1 \dots L_m$) and wind power ($W_1 \dots W_n$) zones for t time frames and approximated load centroid (●) and wind power centroid (**X**)*

5.1. Case A

In *Case A*, we consider WPPs at bus 34 (1760MW) and bus 37 (880MW) to compensate the reduced conventional generation as well as load growth. Running probabilistic AC OPF, Fig. 9(a) and Fig. 9(b) show the loading of all the lines for non-correlated and correlated load and wind power samples. Branch indices correspond to Matpower branch indices. The figures give an overview of most loaded lines and we focus on lines 20-34 and 25-37 (indices 34 and 41) as shown in Fig. 10(a) and Fig. 10(b). Because the number of convergence changes for each power flow case, the probability of occurrence vs. apparent power flow of line was mapped. A closer look at the figures reveals the advantage of considering spatio-temporal dependence in terms of line overloading occurrences.

Fig. 10(a) shows a considerable decrease in line overloading risk where the probability of overload during both near violations as well as overload is less than halved. On the other hand, Fig. 10(b) reveals not much success in considering spatio-temporal dependency. Loading of line 25-37 at 100% is marginal. This could be understood as the net load being higher than generation capacity. To understand the risk index of the overall system, Table 1 shows the risk indices for *Case A*. Recalling from equation 1, risk indices correspond to values when the lines are overloaded by 90% or more of their rated capacity. The overload risk index is measured by the probability of line overload and corresponding severity. Line 20-34 which connects 1600MW

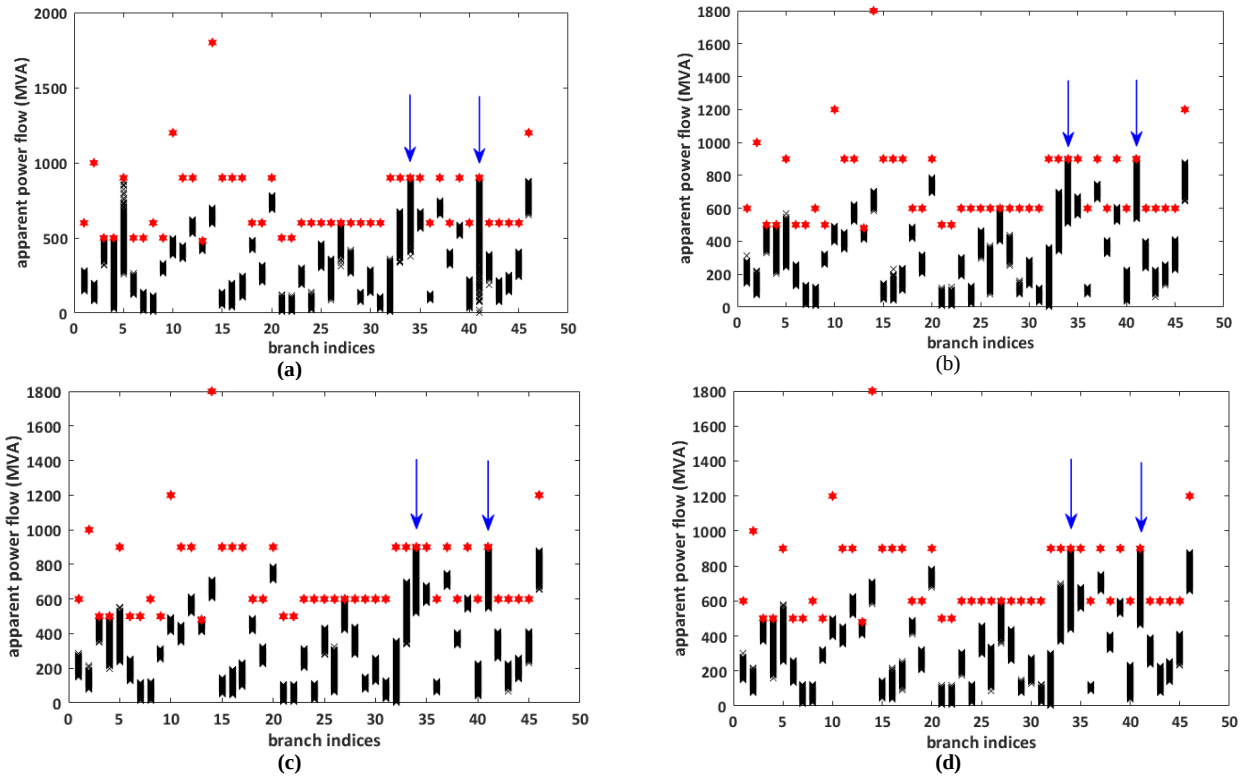
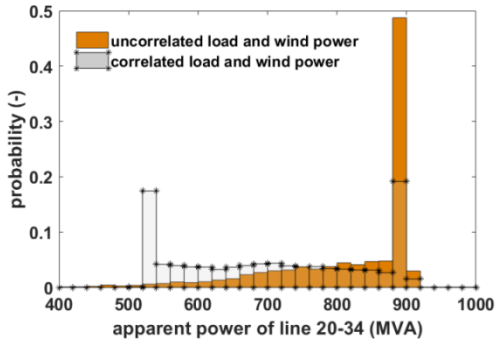
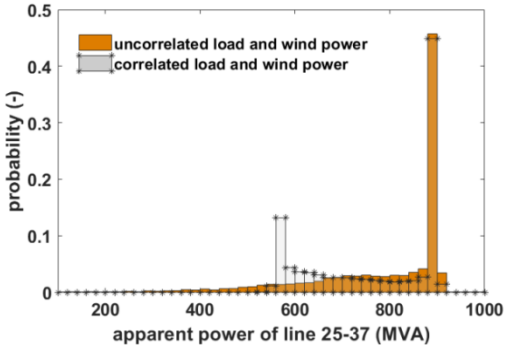


Fig. 9: *Line apparent power flow (MVA) vs. branch indices after running OPF for (a) non-correlated load and wind power in Case A, (b) correlated load and wind power in Case A, (c) non-correlated load and wind power in Case B, (d) correlated load and wind power in Case B. Red marks represent the maximum line capacity. Blue arrows represent the line index under consideration.*

of wind power to rest of the grid shows a remarkable decrease of line overload by more than 50% when the correlation is considered. Similarly, overload risk of line 25-37 decreases by 12.3%. Since the two lines are considered vital when connecting the massive WPPs into rest of the grid, such a decrease in overload risk can be considered beneficial in comparison to the construction of a new line in the same corridor. For the entire system, there is a decrease in the overload risk index from 4.6010 to 3.7840.



(a)



(b)

Fig. 10: *Empirical probability distribution of power flow of (a) line 20-34 for non-correlated and correlated load wind power for Case A (b)line 25-37 for non-correlated and correlated load wind power for Case A. Line rating of both lines is 900MVA*

If we consider the generation facility in *WEST*, influx of 880MW of wind at bus 37 is accompanied with other generation sources at bus 37, 30 and 39 which is responsible for the heavy loading of lines 2-3, 2-25 and 25-37. The risk index of line 25-37 fairly decreases with spatio-temporal modeling. Moreover, variation in wind production is compensated by already available generation at bus 37 that accounts for a slightly low-risk index. Net risk indices of the three lines decrease by 10.38% with consideration of correlation. The other two lines that are heavily loaded are 6-

11 and 16-19. The risk index of 6-11 remains more or less constant with or without considering correlation. It is to be noted that a new load site at bus 6 and an increase of load at buses 7 and 31 without any changes in the generation can be considered as responsible for such overloading. One proposed solution for such case is building an interconnection between bus 34 and 11 or other nearby buses. Overloading of line 16-19 is explained by the addition of WPPs at bus 34 accompanied by generation at bus 33. Consideration of correlation leverages overloading which decreases by 25% and it can be considered advantageous for TSOs.

An extensive suite of 30000 Monte Carlo simulation (MCS) was performed. Such a large sample size is chosen to guarantee a good accuracy of the estimated value. MCS is one way to solve probabilistic OPF which often serves as accuracy reference. MCS firstly samples the random variables and then for each sample a load flow case is solved to obtain all states. Based on the load flow results of all samples, scenarios are generated randomly from PDF. Table 5.2 compares the number of samples converging in the probabilistic OPF problem. The number of correlated cases converging is on a higher side as compared to non-correlated samples. This is certainly helpful as the empirical probability distribution of power flow of lines contains a higher number of feasible states.

Table 1 Comparison of Line Overload Risk Indices for *Case A*

Index	Line	Non-correlated	Correlated
3	**2-3**	0.7688	0.7075
4	**2-25**	0.6024	0.5331
13	**6-11**	0.9841	0.9815
27	**16-19**	0.9571	0.7127
34	**20-34**	0.6746	0.3118
41	**25-37**	0.6128	0.5374
	Total	**4.6010**	**3.7840**

Table 2 Comparison of Non-correlated and Correlated Samples Converging in *Case A*

Total samples	Non-correlated	Correlated
30000	15914	28493

5.2. Case B

Compared to *Case A*, *Case B* considers a slight increase in net wind power penetration while a slight decrease in net system load. WPPs of 1760MW at bus 34 and 1000MW at bus 37 are used to compensate for lowering of conventional (nuclear and fossil fuel) generation units and load growth. Numerically, a 1.5% equivalent increase of wind power penetration and 5.5% decrease in overall generation compared to *Case A* is considered for probabilistic AC OPF. Fig. 9(c) and Fig. 9(d) shows the loading of all the lines after running the power flow. A comparative figure showing the loading of lines 20-34 and 25-37 is shown in Fig. 11(a) and Fig. 11(b). From Fig. 11(a), it is evident that the loading of line 20-34 at 100% significantly decreases when spatio-temporal dependency is considered. A remarkable decrease in loading proves the advantage of addressing correlation. However, Fig. 11(b) shows the heavy loading condition of line 25-37 at slightly higher than 100%. Though there is a decrease in line loading at nearly 100%, an increase in wind power generation in *WEST* still does not compensate for the high zonal net load.

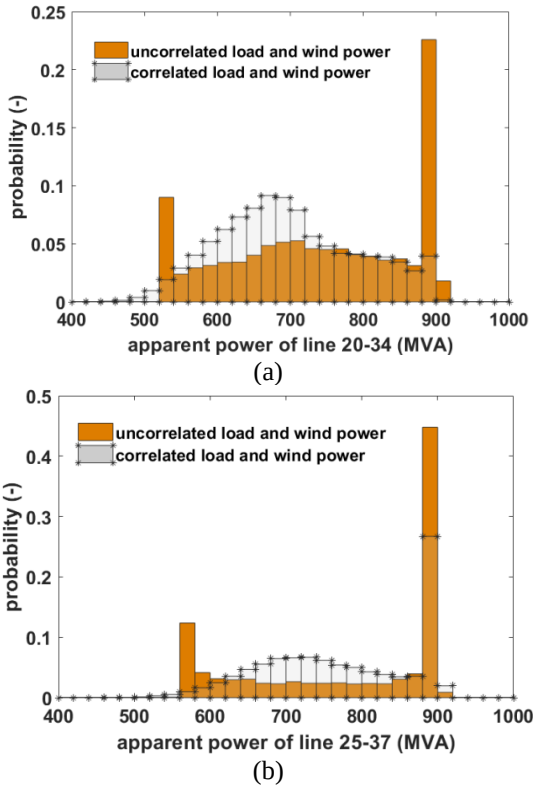


Fig. 11: *Empirical probability distribution of power flow of (a) line 20-34 for non-correlated and correlated load wind power for Case B (b) line 25-37 for non-correlated and correlated load wind power for Case B. Line rating of both lines is 900MVA.*

Table 3 shows the overload risk indices for the total system and some heavily loaded lines. The net system risk index is lowered by 15.23% when the correlation is considered. In both *Case A* as well as *Case B*, it is noticed that the same set of lines are often overloaded. In the *WEST*, risk indices for lines 2-3, 2-25 and 25-37 decreases considerably after accounting for correlation although there is an increase in wind power generation at bus 37. In the *MIDATL*, lines 6-11 and 16-19 show an increase in risk index when the correlation is considered. It can be understood that risk indices are seriously affected with consideration of spatio-temporal dependency. The proposed solution of adding an interconnection between bus 34 and 11 or other nearby buses is still considered to be a proposed solution to combat the increase in risk index. Adding an interconnection can alleviate the overloading on existing lines though finding the exact location is out of the scope of this paper. An increase of load at bus 16 accounts for a marginal increase in risk index for line 16-19. The positive aspect is seen in the form of low overall risk index of the total system with consideration of spatio-temporal dependency.

Table 3 Comparison of Line Overload Risk Indices for *Case B*

Index	Line	Non-correlated	Correlated
3	**2-3**	0.8553	0.7009
4	**2-25**	0.5582	0.4025
13	**6-11**	0.9660	0.9846
27	**16-19**	0.8467	0.8564
34	**20-34**	0.3670	0.1616
41	**25-37**	0.5617	0.4158
	Total	**4.1548**	**3.5217**

Similar to *Case A*, the number of correlated samples converging is higher than the non-correlated samples. The performance of correlated samples is nearly the same with the similar number of converging states ($\sim 28xxx$ states). This indicates the advantage of considering correlated samples as a higher number of feasible states correspond to better insight on system operating condition.

Table 4 Comparison of Non-correlated and Correlated Samples Converging in *Case B*

Total samples	Non-correlated	Correlated
30000	11943	28562

6. Conclusion

To answer the urgent need of relevant tool and risk quantification measures for transmission line overloading, a novel attempt of using vine copula for spatio-temporal modeling to perform a risk-based security assessment of transmission line overloading is presented. Real load and wind power data are mapped onto a modified IEEE 39-bus system representing different market zones of U.S. utility. Use of real data with probabilistic AC OPF analysis gives a more realistic behavior of grid performance in terms of realistic risk indices. The main outcomes of this research can be listed as:

- Use of vine copula to model joint normal distribution addresses spatio-temporal dependency. In order to quantify the risk, this is seen important in future because of the massive integration of wind power into existing grid infrastructure. Joint normal distribution is important as it points out to two key properties: the non-Gaussianity of marginal distributions and the complex dependence structures. The reproducible sampling algorithm generates correlated samples that are then mapped onto the test case for risk assessment. It was also observed that the number of converging states for correlated samples is nearly the same for both the cases. A high number of converging states corresponds to exploring the high number of operating states.
- Probabilistic AC OPF allows measuring the probability of line overload. Overload probabilities contribute significantly to the risk index of both lines and also the whole system. Risk quantification is achieved by combining the probability with the severity of line overload. For studying the overload risk indices for 90% loading or more, the probability of overload is considered and the corresponding probabilistic risk indices are calculated. The proposed risk quantification technique is able to qualitatively interpret the numerical values corresponding to risk indices.
- Ideally, the OPF should consider ramping constraints and generation unit start-up/shut-down constraints, especially in case of high penetration level of wind power generation and short-term operational planning (e.g. day-ahead). However, this research work did not involve a wind penetration level as high as 50% or more of the total generation share, and the goal is to perform OPF from a yearly operational planning perspective. Moreover, if load variability is defined by exclusively considering load data, then it neglects all kind of correlation with wind power generation which is pursued in this research. Such consideration of load and wind power correlation helps in better analyzing the risk metrics.
- The two cases studied in this research can be regarded as future scenarios aiming at low-carbon electricity generation in the form of massive integration of WPPs. Risk indices for the overall system vary significantly for the two cases and a high number is seen for *Case A*. The reason can be understood as an increase in both load and wind power as compared to *Case B*. For both the cases, overloading of lines 6-11 and 16-19 indicate the need for future expansion planning to address the issue.

7. Acknowledgments

The research leading to this result has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 608540 GARPUR project (http://www.garpur-project.eu). The scientific responsibility rests with the authors.

8. References

[1]. 'Technical regulation 3.2.5 for wind power plants with a power output above 11 kw', (Energinet.dk, 2015)
[2]. 'Eirgrid grid code, version 6.0', (EirGrid, 2015)
[3]. Milligan, M., Kirby, B., et al.: 'Review and status of wind integration and transmission in the United States: Key issues and lessons learned', (National Renewable Energy Laboratory, 2015)
[4]. European wind energy association. Wind in power: 2017 European statistics. [Online] https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2017.pdf
[5]. Danish TSO experiences with large scale integration of DERs, *IEA*, Paris February 29, 2016
[6]. Xie, L., Carvalho, P., et al.: 'Wind integration in power systems: Operational challenges and possible solutions', Proc. IEEE, 2011, 99, (1), pp. 214–232
[7]. http://instituteforenergyresearch.org/analysis/germanys-green-energy-destabilizing-electric-grids/
[8]. Ni, M., McCalley, J.D., et al. : 'Software implementation of online risk-based security assessment', IEEE Trans. Power Syst., 2003, 18, (3), pp. 1165–1172
[9]. Khuntia, S.R., Tuinema, B.W., et al.: 'Time-horizons in the planning and operation of transmission networks: an overview', IET Gen. Trans. Dist., 2016, 10, (4), pp. 841-848
[10]. Osborn, D., Henderson, M., et al.: 'Driving forces behind wind', IEEE Power Energy Magazine, 2011, 9, pp. 60-74
[11]. Xie, L., Gu, Y., et al.: 'Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch', IEEE Trans. Smart Grid, 2014, 5, pp. 511-520

[12]. Li, X., Zhang, X., et al.: 'Transmission line overload risk assessment for power systems with wind and load-power generation correlation', IEEE Trans. Smart Grid, 2015, 6, (3), pp. 1233-1242
[13]. Morales, J., Mínguez, R., et al.: 'A methodology to generate statistically dependent wind speed scenarios', Appl. Energy, 2010, 87, pp. 843-855
[14]. Usaola, J.: Probabilistic load flow with correlated wind power injections', Elect. Power Syst. Res., 2010, 80, (5), pp. 528–536
[15]. Morales, J.M., Baringo, L., et al.: 'Probabilistic power flow with correlated wind source', Inst. Eng. Tech., 2010, 5, (4), pp. 641–651
[16]. Abdullah, M.A., Agalgaonkar, A.P., et al.: 'Probabilistic load flow incorporating correlation between time-varying electricity demand and renewable power generation', Ren. Energy, 2013, 55, pp.532-543
[17]. McCalley, J.D., Vittal, V., et al.: 'An overview of risk based security assessment'. Proc. IEEE-PES Summer Meeting, 1999
[18]. Ni, M., McCalley, J.D., et al.: 'Online risk based security assessment', IEEE Trans. Power Del., 2003, 18, (1), pp. 258–265
[19]. Arya, L.D., Titare, L.S., et al.: 'Determination of probabilistic risk of voltage collapse using radial basis function (RBF) network', Elec. Power Syst. Res., 2005, 6, (7), pp. 426-434
[20]. Dai, Y., McCalley, J.D., et al.: 'Annual risk assessment for overload security', IEEE Trans. Power Syst., 2001, 16, (4), pp. 616–623
[21]. Wang, Y., Zhang, N., et al.: 'An efficient approach to power system uncertainty analysis with high-dimensional dependencies', IEEE Trans. Power Syst., 2018, 33, (3), pp. 2984-2994
[22]. Sun, M., Konstantelos, I., et al.: 'C-vine copula mixture model for clustering of residential electrical load pattern data', IEEE Trans. Power Syst., 2017, 32, pp. 2382-2393
[23]. De Jong, M., Papaefthymiou, G., et al.: 'A framework for incorporation of infeed uncertainty in power system risk-based security assessment', IEEE Trans. Power Syst., 2018, 33, (1), pp. 613–621
[24]. Sklar, A.: 'Functions de repartition a n dimensions et leur marges', Publ. Inst. Statis. Univ. Paris, 1959, 8, pp. 229-231
[25]. Khuntia, S.R., Rueda, J.L., et al.: 'Spatio-temporal study for modeling high dimensional future uncertainties: Univariate to multivariate model'. Proc. IEEE PES General Meeting, Portland, 2018.
[26]. Wall, M.E., Rechtsteiner, A., et al.: 'Singular value decomposition and principal component analysis', in 'A practical approach to microarray data analysis' (Springer US, 2003), pp. 91-109
[27]. Weblink: http://www.pjm.com/markets-and-operations/ops-analysis/
[28]. Khuntia SR. Probabilistic security assessment of sustainable power grids: Multivariate analysis and dependence modeling for risk-based security assessment. PhD Thesis. TU Delft; 2018.
[29]. Papaefthymiou, G., Dragoon, K..: 'Towards 100% renewable energy systems: Uncapping power system flexibility', Energy Policy, 2016, 92, pp. 69–82
[30]. Weblink: http://www.pjm.com/renewables/default.html
[31]. Bukhsh, W.A., Zhang, C., et al.: 'An integrated multiperiod OPF model with demand response and renewable generation uncertainty', IEEE Trans. Smart Grid, 2016, 7, (3), pp. 1495-1503.
[32]. Louie, H., Sloughter, J. M.,: 'Probabilistic modeling and statistical characteristics of aggregate wind power', Large scale renewable power Generation, 2014, pp. 19-51. Springer, Singapore.
[33]. Zimmerman, R., Murillo-Sanchez, C., et al.: 'Matpower: Steady- ´ state operations, planning and analysis tools for power systems research and education', IEEE Trans. Power Syst., 2011, 26, (1), pp. 12–19.
[34]. Zimmerman, R.D. Wang, H.: Matpower Interior Point Solver MIPS 1.2.2 User's Manual, 2016 PSERC. Available online: http://www.pserc.cornell.edu/matpower/docs/MIPS-manual-1.2.2.pdf