Truncated Copula-based Scenario Generation to Model Spatiotemporal Correlations of Wind and Solar Power Plants in Large-Scale Systems

Yitong Liu, *Student Member, IEEE,* Mingguo Hong, *Member, IEEE,* Junbo Zhao, *Senior Member, IEEE,* Xiaochuan Luo, *Senior Member, IEEE,*

***Abstract*—The increasing penetration of wind and solar generation introduces significant uncertainties for the planning and operation of power systems. Thus, it is critical to generate reliable, realistic wind and solar energy scenarios to support risk assessment and identify potential extreme events. This paper develops an R-vine copula framework to capture the complex spatiotemporal dependence among wind and solar plants in large scale systems. Since the large number of random variables poses significant challenges for vine construction, a truncation strategy is proposed to reduce computational complexity while preserving the dominant dependence structure. The proposed approach is validated using real data from ISO New England. Comparison results with the multivariate Gaussian copula show that the truncated R-vine copula substantially reduces computational burden and accurately captures tail dependence, making it more suitable for risk-aware power system applications.**

***Index Terms*—Solar and wind, correlation modeling, scenario generation, truncation copula, tail dependence, vine copula.**

I. Introduction

INCREASING integration of renewable energy sources into modern power systems yields significant operational uncertainties. As a result, conventional deterministic analytical approaches are no longer adequate in capturing the stochastic nature introduced by renewable generations [1]. To effectively assess and mitigate potential operational risks, it is necessary to generate multiple stochastic scenarios that represent plausible realizations of renewable power outputs, which constitute a fundamental prerequisite for the subsequent risk analysis.

When generating such scenarios, it is crucial to recognize that different generation sites exhibit inherent spatiotemporal dependencies. In other words, their outputs are not entirely independent due to shared meteorological conditions and geographic proximity. Accounting for these dependencies is essential for ensuring that the generated scenarios remain realistic and statistically consistent, while avoiding redundant or physically implausible cases. Copula theory provides a powerful framework for modeling multivariate dependencies among random variables [2], [3]. Generally, copulas can be categorized into two main families. One is elliptical copulas, such as the Gaussian copula and Student-t copula, and another is Archimedean copulas, e.g., the vine copula family. Elliptical copulas are computationally efficient and easy to implement, such as the multivariate Gaussian copula, but they are limited to representing symmetric and linear dependencies, and thus fail to capture tail-dependent or asymmetric relationships often observed in data of wind power output. In contrast, vine copulas offer much greater flexibility, as they can capture nonlinear, asymmetric, and tail-dependent correlations by decomposing high-dimensional dependency structures into a series of pairwise copulas [4].

Vine copulas have been used to model the relationship between multiple wind/PV power outputs [5]–[7]. Several studies have used vine copulas to generate renewable energy scenarios. Ref. [8] utilizes a D-vine copula to generate the correlated samples that reflect the dependent structures between the inputs, such as loads and renewables, for probabilistic load-margin assessment. Ref. [4] also models the dependence among wind farms using vine copula and feeds the generated scenarios into a downstream OPF-based operational programming. A time-varying regular vine mixed copula model is established in [9] to capture the dynamic correlations. The work in [10] generates time-coupled wind power scenarios for wind farms. The work of [10] also compares the performance of D-vine, C-vine, and Gaussian copulas, and concludes that a D-vine has turned out to be ideal for modeling the temporal dependence structure of wind power forecasting errors. Ref. [11] utilizes a sequential method to model R-vine. Although it is not necessarily equivalent to a global optimum, it facilitates reducing the computational burden. To enable efficient fitting and sampling of large datasets, clustering and principal component transformation are introduced, followed by a truncation to pairwise dependencies in [12].

Although leveraging the flexibility of the vine copula to model multivariate dependencies has become a common choice, the approach faces significant computational challenges when the number of random variables is large. This means that vine copulas suffer from scalability issues when dealing with a large number of units or extended forecast horizons. In such cases, the vine copula requires a substantial number of tree layers, and determining the pairwise connections (edges) within each layer becomes computationally intensive. The computational burden and memory requirements grow exponentially, making full vine modeling impractical for large-scale systems. In the existing studies, either the dimensionality of the random variables was not sufficiently large to encounter

This work was supported by ISO New England.

Yitong Liu and Junbo Zhao are with the Department of Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA. (Email: yitong.liu.th@dartmouth.edu, junbo.zhao@dartmouth.edu).

Mingguo Hong and Xiaochuan Luo are with ISO New England, Holyoke, MA 01040 USA.

979-8-3315-8138-1

this challenge, or lack of an explicit criterion for determining the truncation level.

To address these challenges, this paper proposes an R-vine copula–based scenario generation framework for wind and solar power. The marginal distributions of individual units are fitted using forecast percentiles (e.g., P10, P50, and P90), ensuring that the generated scenarios are consistent with forecast uncertainty levels. Furthermore, to mitigate the computational burden associated with high-dimensional dependency modeling, a truncation strategy is developed to balance between correlation-capturing accuracy and computational efficiency. The proposed approach is validated using real historical data from the ISO New England with many wind and solar generation plants.

II. The Proposed Truncated Vine Copula for Scenarios Generation

Copulas capture the dependency structure among multiple random variables, providing mathematical functions to model the joint distribution of multivariate variables by separating the marginal behavior of each variable from its dependence structure. Suppose there are n random variables x_1, x_2, ..., x_n, and their marginal distributions are F_1, F_2, ..., F_n, respectively. Then, based on the copula theory, the cumulative joint distribution of all random variables can be expressed as a function of the univariate marginal distributions with copula C, i.e.,

$$F(y_1, y_2, \ldots, y_n) = C\big(F_1(y_1), F_2(y_2), \ldots, F_n(y_n)\big) \qquad (1)$$

Vine copulas model the dependence structure among variables layer by layer through a series of bi-variate copulas and their conditional relationships. This allows for flexibility in applying diverse bi-variate copulas, such as the Gumbel copula, to capture specific features like tail dependence. The modeling of the R-vine copula and its truncation method are introduced below.

A. Vine Copula

A vine copula decomposes a multivariate joint distribution into the product of its marginal distributions and a series of bi-variate conditional distributions. In this way, the n-dimensional variable is composed of $n-1$ trees, where each tree consists of an edge set E_i that represents bi-variate copulas, and a node set N_i. For the first tree $N_i = \{1, ..., n\}$, and for trees $i = 2, .., n$ and $N_i = E_{i-1}$. Then based on Theorem 4.2 of [13], the multivariate joint density is:

$$f(x_1, \cdots, x_n) = \left[\prod_{k=1}^{n} f_k(x_k)\right]$$
$$\left[\prod_{i=1}^{n-1} \prod_{e\in E_i} c_{e_a,e_b|D_e}\left(F_{e_a|D_e}(x_{e_a} \mid D_e), F_{e_b|D_e}(x_{e_b} \mid D_e)\right)\right] \qquad (2)$$

where $f(x_1, \ldots, x_d)$ denotes the joint probability density function (PDF) of the n-dimensional random vector $(x_1, \ldots, x_n)$; $f_i(x_i)$ denotes the marginal PDF of the i-th random variable x_i; $e \in E_i$ denotes an edge in the set E_i, which is associated with a bivariate conditional dependence; e_a and e_b are two nodes of the edge e, which are called conditioned nodes, and these nodes correspond to variables x_{e_a} and x_{e_b}; D_e denotes the conditioning set associated with edge e, on which the pair (x_{e_a}, x_{e_b}) is conditioned; $c_{e_a,e_b|D_e}$ is the conditional pair-copula density function describing the dependence between x_{e_a} and x_{e_b} given the conditioning variables D_e; $F_{e_a|D_e}(x_{e_a} \mid D_e)$ is the conditional cumulative distribution function (CDF) of x_{e_a} given D_e, and its value lies in the interval $[0, 1]$; $F_{e_b|D_e}(x_{e_b} \mid D_e)$ is the conditional CDF of x_{e_b} given D_e. In summary, in an R-vine, each edge e defines a pair-copula that characterizes the conditional dependence between the variables x_{e_a} and x_{e_b} given the conditioning set D_e, and the corresponding copula density is $c_{e_a,e_b|D_e}$. The entire R-vine distribution is constructed by combining all these conditional pair-copulas to form the joint distribution of the n-dimensional random vector.

Suppose there are N_G generation units and each has D data points from historical dataset. The objective is to generate scenarios for the upcoming T-hour forecast horizon. Therefore, the random vector $\mathbf{x}$ is composed of $x_i(t)$ for $i = 1, ..., N_G$ and $t = 1, ..., T$, and $\mathbf{x} \in \mathbb{R}^{N_G \times T}$. The historical data point corresponding to each $x_i(t)$ is $\tilde{x}_{i,m}(t)$ where $m = 1, ..., D$. Then, the dataset for modeling the R-vine structure is:

$$\mathbf{X} = \begin{bmatrix} \tilde{x}_{1,1}(1) & \cdots & \tilde{x}_{1,1}(T) & \cdots & \tilde{x}_{N_G,1}(1) & \cdots & \tilde{x}_{N_G,1}(T) \\ \tilde{x}_{1,2}(1) & \cdots & \tilde{x}_{1,2}(T) & \cdots & \tilde{x}_{N_G,2}(1) & \cdots & \tilde{x}_{N_G,2}(T) \\ \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \tilde{x}_{1,D}(1) & \cdots & \tilde{x}_{1,D}(T) & \cdots & \tilde{x}_{N_G,D}(1) & \cdots & \tilde{x}_{N_G,D}(T) \end{bmatrix}$$

where $\tilde{x}_{i,m}(t)$ denotes the m-th data point of i-th unit's power output at time t; $\mathbf{X} \in \mathbb{R}^{D\times(N_G\times T)}$, and in $\mathbf{X}$, each column corresponds to all observations for a specific unit–timestep pair, while each row contains all unit–timestep values associated with a single observation. To be used for R-vine modeling, this data matrix must first be transformed into the $[0, 1]$ domain using the empirical probability integral transform. Then, the R-vine structure is determined by selecting pairs with the strongest dependence based on the Akaike information criterion (AIC) [14] (eq. (3)), which ranks candidates by their log-likelihood while incorporating a penalty that favors lower-parameter families.

$$AIC(K) := -2loglik(K) + 2n_{\text{parameters}}(K) \qquad (3)$$

where K denotes the K-th layer tree; $n_{\text{parameters}}$ is the number of parameters; and the log-likelihood is computed as the sum of the logarithmic copula densities evaluated at all observations:

$$loglik(\lambda) = \sum_{i=1}^{n}\sum_{l=1}^{\lambda}\sum_{e\in E_l} \ln\left[c_{e_a,e_b|D_e}\Big(F_{e_a|D_e}(x_{e_a}), F_{e_b|D_e}(x_{e_b})\Big)\right] \qquad (4)$$

B. Truncation of Vine Copula for Improved Computational Efficiency

An n-dimensional variable corresponds to a vine copula with $n-1$ layers of trees and in the k-th tree, there are $n-k$ edges (pair-copulas). The computational cost increases substantially in high dimensions to construct each tree and solve parameters for each pair-copula [15]. There, truncation is needed for high-dimensional vine copulas to release the computational burden.

979-8-3315-8138-1

For vine copula truncation, let λ denote the truncation level. The basic idea is to simplify all higher-order trees above level λ by replacing their pair-copula terms with independence copulas [16], since in practice the conditional dependencies among them tend to be very weak. The AIC gain is computed with different truncation levels to obtain a principled reference for determining an appropriate truncation depth [16]. The formulation of AIC gain for a truncation level λ is:

$$\begin{aligned}\Delta AIC(\lambda) &= AIC(\lambda+1) - AIC(\lambda) \\ &= -2\,\Delta loglik + 2\,\Delta n_{\text{parameters}}\end{aligned} \tag{5}$$

where $\Delta loglik$ denotes the information reward from increasing one layer, and $\Delta n_{\text{parameters}}$ denotes the penalty from increasing new parameters. If $\Delta AIC(\lambda) < 0$, the reward from adding a layer exceeds the penalty, which indicates that the added layer improves the model and should be retained. Otherwise, once the AIC gain becomes positive, adding another layer is not worth it and the vine should be truncated at the current level λ.

In summary, the procedure of determining λ is as follows.

- *Step 1*: Start with $\lambda = 1$ and fit a truncated R-vine copula;
- *Step 2*: Increase λ by one and compute the AIC gain $\Delta AIC(\lambda)$.
- *Step 3*: If $\Delta AIC(\lambda) > 0$, truncate the vine at level λ. Otherwise if $\Delta AIC(\lambda) < 0$, return to *Step 2*.

C. Marginal Distribution Fitting

This section explains how marginal distributions can be estimated from given quantile values. Suppose we have a set $\mathcal{X}$ that is composed of target quantiles x_{p_1}, x_{p_2}, ..., x_{p_m}, together with a corresponding set $\mathcal{P}$ with percentiles p_1, p_2, ..., p_m. For example, $x_{10\%}$ is the quantile of 10%. We aim to find parameter vector $\boldsymbol{\theta}$ of the marginal distribution such that its theoretical quantiles match the target quantiles as closely as possible. Then, the general mathematical formulation of the objective is:

$$\begin{aligned}\min_{\boldsymbol{\theta}}\; J(\boldsymbol{\theta}) &= \sum_{p_i\in\mathcal{P},x_{p_i}\in\mathcal{X}} \left(F^{-1}(p_i;\boldsymbol{\theta}) - x_{p_i}\right)^2 \\ \text{s.t.}\quad & \boldsymbol{\theta}\in\Theta,\end{aligned} \tag{6}$$

where $F^{-1}(p_i;\boldsymbol{\theta})$ is the inverse CDF; Θ is the feasible parameter domain.

Wind power output is generally considered following the Weibull distribution [17], whose inverse CDF of the Weibull distribution is given by:

$$F^{-1}(p;a,l,s) = l + s\left[-\ln(1-p)\right]^{1/a}. \tag{7}$$

where a, l, and s are the parameters of the Weibull distribution, which determine its shape, location, and scale, respectively. Therefore, to fit a Weibull-based marginal distribution, the model is:

$$\begin{aligned}\min_{a,l,s}\quad & \sum_{p_i\in\mathcal{P},x_{p_i}\in\mathcal{X}} \left(l + s[-\ln(1-p_i)]^{1/a} - x_{p_i}\right)^2 \\ \text{s.t.}\quad & a>0,\quad s>0\end{aligned} \tag{8}$$

Since solar power output is generally considered to follow Beta distribution [18], [19], the inverse CDF of the Beta distribution is denoted as $B^{-1}(p;a^B,b^B)$, whose domain is naturally $[0,1]$. Then, the mathematical model for fitting a Beta-based marginal distribution is:

$$\begin{aligned}\min_{a^B,b^B}\quad & \sum_{p_i\in\mathcal{P},x_{p_i}\in\mathcal{X}} \left(B^{-1}(p_i;a^B,b^B) - x_{p_i}\right)^2 \\ \text{s.t.}\quad & a^B>0,\ b^B>0\end{aligned} \tag{9}$$

where the parameters a^B and b^B respectively control the shape of the distribution near 0 and 1, determining its skewness and concentration. A larger a^B causes the distribution to skew toward 1, whereas a larger b^B causes it to skew toward 0.

D. Wind and Solar Power Scenario Generation

As an enhanced coverage of sampling space enables the generated scenarios to account for a wider spectrum of potential risks, the Latin Hypercube Sampling (LHS) is adopted as the initial sampling strategy. It guarantees stratified sampling in every dimension, leading to a uniform exploration of the N_G-dimensional unit hypercube. D_s samples are generated by LHS, denoted as $\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{D_s}\}$, where each sample $\mathbf{u}_i$ is an N_G-dimensional vector.

These uniform samples are then transformed into dependent samples that follow the joint distribution encoded by the vine copula through the inverse Rosenblatt transform:

$$v_{M[n-j,j]} = F^{-1}\big(u_{M[n-j,j]}\,\big|\,u_{M[n-j-1,j-1]},\ldots,u_{M[0,0]}\big) \tag{10}$$

where v is the correlated sample, whose dependence structure is determined by the R-vine copula; $F^{-1}(\cdot,\cdot)$ denotes the conditional inverse CDF of the corresponding pair-copula (conditional inverse CDF); M denotes the R-vine structure matrix, which determines the variable generation order and the conditioning sets of each variable, and $M[d-j,j]$ is the position of the variable currently being generated in the structure matrix; the conditioning set $u_{M[n-j-1,j-1]},\ldots,u_{M[0,0]}$ is the set of preceding variables determined by the corresponding row and column of the structure matrix.

Finally, the samples $\mathbf{x}_i$ in the physical domain are obtained by applying the inverse marginal CDFs, which are obtained in the section II.C, to the copula-domain samples $\mathbf{v}_i$, i.e.:

$$\mathbf{x}_i = F_i^{-1}(\mathbf{v}_i) \tag{11}$$

III. Experiment Results

The performances of the proposed method are evaluated using real historical data from ISO New England, which includes 37 wind units and 62 solar units. The percentiles provided by the forecasting department consist of P10, P50, and P90, with a 36-hour forecasting horizon. They are used to fit the marginal distributions of each unit. The horizon for scenario generation is also 36-hour.

For comparison, the performance of the proposed method is evaluated against that of a multivariate Gaussian copula–based scenario generation approach. Here, a brief introduction to the principle of scenario generation based on the multivariate Gaussian copula is given. Gaussian copula models the dependency structure of random variables via multivariate normal distribution. It shows an advantage when modeling high dimensional distributions with less computational burden, but lacks support for capturing tail dependence. Gaussian copula can be fully specified by the correlation matrix $\mathbf{R} \in [-1,1]^{n\times n}$, and the Gaussian copula with parameter matrix $\mathbf{R}$ can be written as $C^{\text{Gauss}}_{\mathbf{R}}(\mathbf{u}) = \Phi_{\mathbf{R}}\left(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right)$ where $\Phi_{\mathbf{R}}$ is the joint cumulative distribution function of a multivariate normal distribution; u_i is the cumulative probability under marginal distribution F_i. With original data matrix X, Z follows a standard normal distribution and can be obtained via $Z_{i,m}(t) = \Phi^{-1}\big(F_i(X_{i,m}(t))\big)$ and the correlation structure among different units is preserved. Then, the correlation matrix can be computed for Z using the Pearson correlation method.

The LHS samples are then transformed into independent standard normal samples by applying the inverse cumulative distribution function of the multivariate standard normal distribution, i.e., $z_i = F_z^{-1}(u_i)$. Next, correlations among the samples are introduced through the Cholesky decomposition $\mathbf{R}$, $\mathbf{R} = \mathbf{C}\mathbf{C}^{\mathrm{T}}$. The correlated Gaussian samples can be obtained by $\mathbf{Z}_c = \mathbf{C}^{\mathrm{T}}\mathbf{Z}$. Finally, with inverse marginal distributions, samples in the physical domain are obtained.

A. Truncation Selection for Vine Copula

In the ISO New England case, there are 99 units multiplied by 36 time steps, resulting in 3564 random variables. This high dimensionality implies that truncation is required for the vine copula.

First, the truncation-level is selected and illustrated based on the principle proposed in Section II. The variation of AIC gain with respect to the truncation level λ is shown in Fig. 1. It can be found that the AIC gain increases rapidly at the beginning and then starts to level off and changes its sign when λ reaches around 26. As discussed in Section II, an AIC gain below zero indicates that adding one more tree to the vine model provides considerable improvement in goodness of fit than the penalty introduced by the additional parameters. However, once the AIC gain becomes positive, the benefit of adding a new tree is no longer sufficient to compensate for the increased model complexity. Therefore, the truncation level λ should be set to 26 according to Fig. 1.

Then, to further examine whether the chosen truncation level is justified, we computed the Kendall's τ values of all pair-copulas in the full R-vine and analyzed their variations across tree levels. Kendall's τ summarizes the strength of pairwise dependence implied by each pair-copula. As illustrated in Fig. 2, Kendall's τ exhibits a pronounced decay as the tree level increases. The first 26 trees include a substantial number of pairs with $|\tau| > 0.1$, reflecting appreciable dependence. In contrast, nearly all pair-copulas beyond tree 26 exhibit $|\tau|$ values below 0.1 and tend to be close to 0.05. This indicates that the corresponding higher-order conditional dependencies are of negligible magnitude. Since the truncated vine copula coincides exactly with the full vine copula up to selected truncation tree level, truncating the vine at level 26 effectively preserves the essential dependence structure while excluding only very weak and statistically insignificant pair-copulas.

Compared with the theoretical requirement of 3563 trees for a full vine copula, a truncation level of 26 results in a relatively small vine copula. This indicates that the correlations among many of the 99 units are not strong. This observation

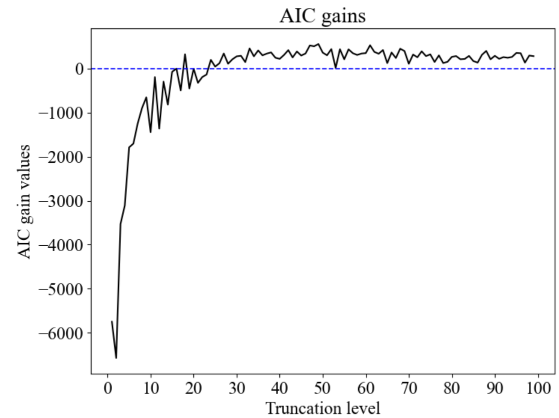


Fig. 1: AIC gains vary with truncation level.

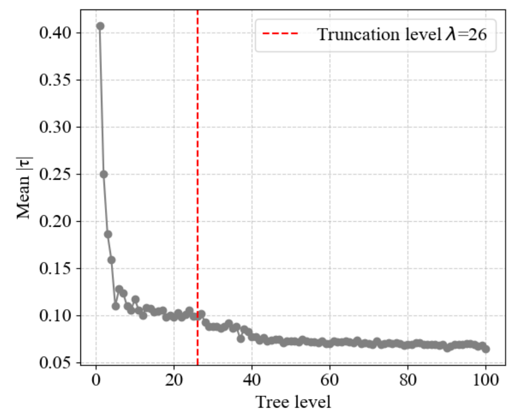


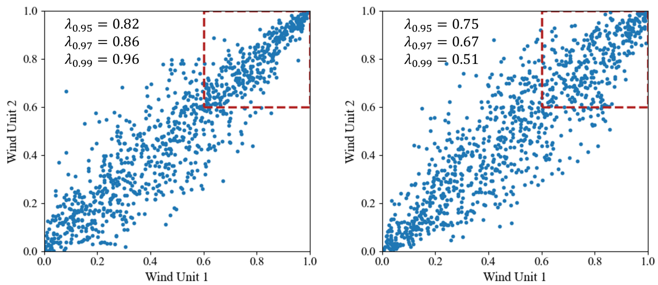
Fig. 2: Decay of dependence strength by tree level.

is consistent with the physical fact that these 99 units are geographically dispersed across different regions of ISO New England, with some located far apart from each other. In the next subsection, we will compare the performance of the truncated vine copula with the Gaussian copula.

B. Comparison of Scenarios Generated by Truncated Vine and Multivariate Gaussian Copula

This section provides a performance comparison between vine-copula–based and Gaussian-copula–based scenario generation. Wind unit 1 and wind unit 2 are two nearby wind units, which means the two adjacent wind units should exhibit a strong upper tail dependence. Figs. 3 (a) and (b) present the vine copula–generated and multivariate Gaussian copula–generated scenario samples, respectively. The upper-tail dependence coefficient $\lambda_U(q) = \mathbb{P}(U_2 > q \mid U_1 > q)$ is also computed to quantitatively assess tail dependence, where U_1 and U_2 denote the samples of the two units in the copula space, and $q \in (0,1)$ is a high quantile threshold. Here, q is set to 0.95, 0.97, and 0.99. The values of λ_U corresponding to the two methods are also annotated in Fig. 3.

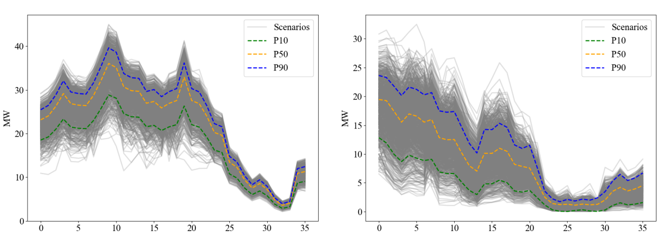
As shown in the red box of Fig.3 (a), the vine copula successfully captures this upper tail dependence, with the generated samples being more concentrated in the high-power region. And λ_U are typically larger than 0.8 and gradually approach 1 as q increases toward 1. This means when one unit produces high power, the other tends to generate high power simultaneously. In contrast, the Gaussian copula, as shown in the red box of Fig. 3 (b), produces samples that are more scattered across the space, and yields relatively small λ_U which further decrease as q increases. This means it does not adequately reflect the tail-dependent characteristics. This is because the Gaussian copula assumes linear correlation and cannot represent joint extreme behavior in the tails. This observation further confirms that the vine copula offers greater flexibility and accuracy in modeling nonlinear dependencies and tail characteristics of wind power outputs, enabling the generation of more physically realistic and statistically consistent scenarios. They serve as more reliable inputs for subsequent system analysis and optimization.



(a) Vine copula (b) Gaussian copula

Fig. 3: Comparison of vine copula and Gaussian copula.

Taking 1000 scenarios as an example, we further illustrate the generated results for the two nearby wind units using Vine and Gaussian copulas in Fig. 4. It can be seen that the generated scenarios generally follow the forecast trend, but exhibit a wider variety of variation patterns. This provides us with insights regarding uncertainty characteristics such as ramping and output spread.



(a) Wind unit 3 (b) Wind unit 4

Fig. 4: Generated time-series power output scenarios for two units.

IV. Conclusion

This paper proposes a scenario generation framework based on truncated R-vine copulas. The R-vine structure is used to model the spatiotemporal dependence of wind and solar power. The proposed method addresses the practical challenge of high dimensionality by truncating higher-order trees, thereby significantly reducing computational burden without sacrificing essential dependence structure. Results obtained using real data from ISO New England demonstrate that the truncated R-vine copula preserves the representation of the dependence compared with the full-vine. It also outperforms the Gaussian copula, particularly in capturing tail dependence. Overall, the approach provides a realistic and computationally efficient tool for uncertainty modeling. Future work will be extending the proposed approach for robust risk analysis and operational decision-making in systems with high renewable penetration.

References

[1] W. Wu, K. Wang, B. Han, G. Li, X. Jiang, and M. L. Crow, “A versatile probability model of photovoltaic generation using pair copula construction,” *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1337–1345, 2015.

[2] E. M. S. Duque, P. P. Vergara, P. H. Nguyen, A. van der Molen, and J. G. Slootweg, “Conditional multivariate elliptical copulas to model residential load profiles from smart meter data,” *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4280–4294, 2021.

[3] N. Zhou, X. Xu, Z. Yan, and M. Shahidehpour, “Spatio-temporal probabilistic forecasting of photovoltaic power based on monotone broad learning system and copula theory,” *IEEE Transactions on Sustainable Energy*, vol. 13, no. 4, pp. 1874–1885, 2022.

[4] A. B. Krishna and A. R. Abhyankar, “An efficient data-driven conditional joint wind power scenario generation for day-ahead power system operations planning,” *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 3105–3117, 2024.

[5] M. Sun, I. Konstantelos, and G. Strbac, “C-vine copula mixture model for clustering of residential electrical load pattern data,” *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2382–2393, 2017.

[6] Z. Wang, W. Wang, C. Liu, and B. Wang, “Forecasted scenarios of regional wind farms based on regular vine copulas,” *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 77–85, 2020.

[7] A. Arrigo, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, “Embedding dependencies between wind farms in distributionally robust optimal power flow,” *IEEE Transactions on Power Systems*, vol. 38, no. 6, pp. 5156–5169, 2023.

[8] Y. Xu, L. Mili, M. Korkali, K. Karra, Z. Zheng, and X. Chen, “A data-driven nonparametric approach for probabilistic load-margin assessment considering wind power penetration,” *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4756–4768, 2020.

[9] Q. Tu, S. Miao, F. Yao, Y. Li, H. Yin, J. Han, D. Zhang, and W. Yang, “Forecasting scenario generation for multiple wind farms considering time-series characteristics and spatial-temporal correlation,” *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 4, pp. 837–848, 2021.

[10] R. Becker, “Generation of time-coupled wind power infeed scenarios using pair-copula construction,” *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1298–1306, 2018.

[11] Z. Wang, W. Wang, C. Liu, Z. Wang, and Y. Hou, “Probabilistic forecast for multiple wind farms based on regular vine copulas,” *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 578–589, 2018.

[12] I. Konstantelos, M. Sun, S. H. Tindemans, S. Issad, P. Panciatici, and G. Strbac, “Using vine copulas to generate representative system states for machine learning,” *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 225–235, 2019.

[13] D. Kurowicka and R. M. Cooke, *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, 2006.

[14] K. Aas, C. Czado, A. Frigessi, and H. Bakken, “Pair-copula constructions of multiple dependence,” *Insurance: Mathematics and Economics*, vol. 44, no. 2, pp. 182–198, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167668707000194

[15] F. A. Alanazi, “Novel pruning and truncating of the mixture of vine copula clustering models,” *Scientific Reports*, vol. 12, no. 1, p. 19815, 2022. [Online]. Available: https://doi.org/10.1038/s41598-022-24274-7

[16] E. C. Brechmann and C. Czado, “Truncated regular vines in high dimensions with application to financial data,” *Canadian Journal of Statistics*, vol. 40, no. 1, pp. 68–85, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/cjs.10141

[17] T.-H. Yeh and L. Wang, “A study on generator capacity for wind turbines under various tower heights and rated wind speeds using weibull distribution,” *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 592–602, 2008.

[18] D. Q. Hung, N. Mithulananthan, and K. Y. Lee, “Determining pv penetration for distribution systems with time-varying load models,” *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 3048–3057, 2014.

[19] K. Doubleday, S. Jascourt, W. Kleiber, and B.-M. Hodge, “Probabilistic solar power forecasting using bayesian model averaging,” *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 325–337, 2021.

979-8-3315-8138-1