Probabilistic Power Flow Considering Wind Speed Correlation of Wind Farms

Han Yu
University of Calgary
AB, Canada, T2N1N4
hayu@ucalgary.ca

Bill Rosehart
University of Calgary
AB, Canada, T2N1N4
rosehart@ucalgary.ca

Abstract - The development of renewable energy tends to introduce uncertainty into power system operation. An obvious example is wind power, which in many systems now plays an important role in the overall generation mix. Since wind speed correlation tends to have a significant impact on the operation of power systems, a method of simulating correlated wind speed in different wind farms for modeling is of importance. This paper proposes using an extended Latin hypercube sampling algorithm to simulate correlated wind speeds for different wind farms when solving probabilistic power flow problems. The proposed method employs rank numbers of the sampling points to generate correlated wind speed samples of different wind farms, which avoids generating negative wind speed values in transforming uncorrelated samples to correlated samples, thus improving sampling accuracy. Simulation results using test systems show that the proposed method is effective, simple and easy to be applied to other correlated renewable energy state sampling problems. The illustrations also show the necessity of taking wind speed correlations into consideration in probabilistic power flow.

Keywords - Probabilistic power flow, Latin hypercube sampling, Wind generation, Wind speed correlation

1 INTRODUCTION

PROBABILISTIC power flow (P-PF) is a useful tool to evaluate the impact of uncertainties in power system operation and planning. Generally in P-PF, load variations and forced outage of generations are taken into account. With the development of renewable energy, the intermittent and inconstant renewable power output tends to introduce more uncertainties into power system. An obvious example of renewable energy is wind power, which in many systems now plays an important role in the overall generation mix. Since wind speeds are usually correlated in adjacent wind farms, a P-PF method with consideration of the wind speed uncertainty and correlation is of importance [1].

Early calculating methods of P-PF include convolution [2] and Monte Carlo simulation (MCS) with random sampling (RS) [3]. In order to alleviate the computational burden of P-PF, researchers developed many methods which include the cumulant method [4], the fast Fourier transform method [3], the Von Mises method [5], the point estimate method [1] and the first-order second-moment method [6] etc. Although each P-PF solving method above has its own special advantages, MCS is still generally agreed to be the most robust, flexible and accurate approach. Therefore, it is of great value to keep the advantage of MCS and improve its calculation efficiency.

Recently, MCS with Latin hypercube sampling (LHS) is employed to deal with power system reliability analysis [7] and P-PF problem [8] and found to be more efficient than RS. However, uncertain wind power and wind speed correlations are not considered in [8].

There are three main methods employed to simulate correlated wind speeds in the literature. One is Markov chain Monte Carlo [9], which is employed in chronological simulation. In reference [10], Feijóo et al. proposed the other two methods to simulate correlated wind speeds in steady-state power system analysis, in which the first method is more simple, time-saving and widely used than the second one since it avoids large numbers of repeated number counting in the latter. The steps of Feijóo's first method can be summarized as follows [10].

(1) Uncorrelated wind speed values are generated according to the distributions by random sampling;
(2) The simulated values are normalized by subtracting the mean and dividing by the standard deviation;
(3) The correlated wind speed values are achieved by multiplying the decomposed correlation matrix with the normalized simulated values and adding the subtracted mean.

The advantage of this method is that the correlated wind speed samples can be obtained directly from uncorrelated samples. The disadvantage is that some negative wind speed values appear in the results, which should be discarded; meanwhile, the distributions of simulated wind speeds have a small difference from the assumed distribution. Since Feijóo's first method is essentially a RS method, and for simplification, it is called "RS method" in this paper.

This paper adapts a method to simulate correlated wind speeds more efficient and accurate than the RS method. It extends the LHS method from [8] to consider the wind speed correlation between different wind farms. Rank numbers of the sampling points are employed to generate correlated wind speed samples of different wind farms, which avoids generating negative wind speed values in transforming uncorrelated samples to correlated samples. The proposed method is tested using the IEEE 14-bus system and the IEEE 118-bus system. Simulation results show that the proposed method is effective, simple and easy to be applied to other correlated renewable energy state sampling problems. Illustrations also show the necessity of taking wind speed correlations into consideration in probabilistic power flow.

The remaining parts of this paper are organized as follows. Latin hypercube sampling is introduced in Section 2. The proposed method of inducing wind speed correlations is described in Section 3. In Section 4, the algorithm of P-PF which takes into consideration both wind farm output uncertainty and correlated wind speed is proposed. In Section 5, the proposed method is examined using two test systems. Finally, conclusions are given in Section 6.

2 OVERVIEW OF LATIN HYPERCUBE SAMPLING

The following is brief introduction to LHS. Detailed information on this approach can be found in [11].

Let $X_1, \cdots, X_K$ represent the total K random input variables and X_k belongs to $X_1, \cdots, X_K$. The method of calculating samples of X_k with LHS is as follows.

The cumulative distribution function is employed to calculate the representative sample values of X_k:

$$y_k = F_k(x_k) \tag{1}$$

where y_k indicates the probability of $X_k \leq x_k$. Since y_k increases from 0 to 1, for assigned sample size N, the range of y_k can be stratified into N equal and non-overlapping intervals, the length of which is $1/N$. In each interval, the midpoint is selected as a representative value. Then the sample values of X_k are calculated by substituting these representative y_k values into the inverse function of equation (1), i.e. the n^{th} sample of X_k, which is represented as x_{kn}, is calculated by

$$x_{kn} = F_k^{-1}\left(\frac{n-0.5}{N}\right) \tag{2}$$

The samples of X_k are assembled in a row vector $[x_{k1}, x_{k2}, \ldots, x_{kN}]$. After all the random variables are sampled, a preliminary $K \times N$ sampling matrix $\boldsymbol{X}$ is achieved.

Essentially, the sampling process of LHS is to divide the distribution area of a random variable into N section with equal probability, and then select a representative value as a sample from each section. The advantage of LHS is that it ensures the whole distribution area of X_k can be entirely covered by the samples. Compared with RS, LHS achieves much higher sampling efficiency [11].

After the LHS samples of $X_1, \cdots, X_K$ are achieved, the correlations between random variables $X_1, \cdots, X_K$ are induced to LHS samples of $X_1, \cdots, X_K$. The details are explained in the next section.

3 INDUCE DESIRED CORRELATIONS BETWEEN SAMPLES OF DIFFERENT RANDOM VARIABLES

The section describes the proposed extension of the LHS method to correlations between different random variables. The correlation between any of the two random variables X_k and X_g in $X_1, \cdots, X_K$ is defined as follows:

$$\begin{aligned} r_{kg} &= \frac{cov(X_k, X_g)}{\sigma_k \sigma_g} \\ &= \frac{E[(X_k - \mu_k)(X_g - \mu_g)]}{\sigma_k \sigma_g} \end{aligned} \tag{3}$$

where $E()$ is an operator to get mean value, $cov()$ is an operator to get covariance, μ_k and μ_g are mean values of X_k and X_g respectively, and σ_k and σ_g are standard deviation of X_k and X_g respectively. A correlation coefficient matrix $\boldsymbol{C}$ is formed by the correlations between any the random variables in $X_1, \cdots, X_K$ as follows:

$$\boldsymbol{C} = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1K} \\ r_{21} & r_{22} & \ldots & r_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ r_{K1} & r_{K2} & \ldots & r_{KK} \end{bmatrix} \tag{4}$$

3.1 Method of inducing desired correlations between samples of different random variables

Some approximations are made for the purpose of inducing correlations to Latin hypercube samples simulated with the method introduced in Section 2. Since rank correlations between random variables' samples are consistent with and close to the real correlations, the problem of inducing correlation $\boldsymbol{C}$ to sampling matrix $\boldsymbol{X}$ is simplified as making the rank correlation of $\boldsymbol{X}$ to be $\boldsymbol{C}$ [12]. In other words, the above problem becomes permuting the elements in the initial sampling matrix $\boldsymbol{X}$ according to a rank matrix $\boldsymbol{S}$ whose associated row correlation matrix is $\boldsymbol{C}$. Here, the rank matrix is defined to be a matrix whose elements in each row are permutation of number $1, \cdots, N$. The steps of permuting the sample matrix to get the desired correlation $\boldsymbol{C}$ are as follows [12]:

(1) Generate a $K \times N$ matrix $\boldsymbol{S}_0$ whose correlations between the rows are minimized close to zero.
An initial $\boldsymbol{S}_0$ is obtained by randomly placing $1, \cdots, N$ in each row. If the correlations between the rows of $\boldsymbol{S}_0$ are evaluated, it will be found that small undesired correlations exist. Therefore, those correlations are further minimized. Assuming $\boldsymbol{C}_0$ is the associated correlation coefficient matrix of $\boldsymbol{S}_0$, as a positive-definite symmetric matrix, $\boldsymbol{C}_0$ can be decomposed by Cholesky decomposition as follows:

$$\boldsymbol{C}_0 = \boldsymbol{L}_0 \boldsymbol{L}_0^T \tag{5}$$

where $\boldsymbol{L}_0$ is a lower triangular matrix. Then $\boldsymbol{S}_0$ is updated to make the correlations between the rows are closer to zero than the initial $\boldsymbol{S}_0$ by the following method:

$$\boldsymbol{S}_0 \leftarrow \boldsymbol{L}_0^{-1} \boldsymbol{S}_0 \tag{6}$$

After the above transformation, unlike the initial $\boldsymbol{S}_0$, the elements in the updated $\boldsymbol{S}_0$ are not assured to be integer and positive.

(2) Judge the condition of coefficient matrix $\boldsymbol{C}$.
If $\boldsymbol{C}$ is identity, i.e. the input random variables are uncorrelated, assign $\boldsymbol{S}_0$ to $\boldsymbol{S}'$, then go to Step 4 to

calculate the final rank matrix, where $\boldsymbol{S}'$ represents an intermedium matrix whose associated correlation matrix has been improved to close to $\boldsymbol{C}$; otherwise, the input random variables are correlated, go to Step 3 to induce the correlation $\boldsymbol{C}$.

(3) Make a matrix $\boldsymbol{S}'$ whose associated correlation coefficient matrix is $\boldsymbol{C}$.
Similar with $\boldsymbol{C}_0$, matrix $\boldsymbol{C}$ is also positive-definite symmetric, and a lower triangular matrix $\boldsymbol{L}$ is obtained by Cholesky decomposition:

$$\boldsymbol{C} = \boldsymbol{L}\boldsymbol{L}^T \tag{7}$$

Then matrix $\boldsymbol{S}'$ is achieved by

$$\boldsymbol{S}' = \boldsymbol{L}\boldsymbol{S}_0 \tag{8}$$

(4) Calculate the final rank matrix $\boldsymbol{S}$.
In order to employ the information in $\boldsymbol{S}'$ to indicate ranks in the sampling matrix, the rank information of $\boldsymbol{S}'$ is assigned to the final rank matrix $\boldsymbol{S}$:

$$\boldsymbol{S} \leftarrow getRankInRow(\boldsymbol{S}') \tag{9}$$

where $getRankInRow(\boldsymbol{S}')$ denotes getting rank numbers (from 1 to N) in every row of matrix $\boldsymbol{S}'$.

(5) Permute sampling matrix $\boldsymbol{X}$ according to $\boldsymbol{S}$.
The desired correlations between rows of $\boldsymbol{X}$ are induced after arranging the sequence of the sampling values in each row of matrix $\boldsymbol{X}$ according to the sequence numbers in the corresponding row of $\boldsymbol{S}$.

Essentially, the initial $\boldsymbol{S}_0$ indicates the random permutation of without considering the correlations; by equation (6), $\boldsymbol{S}_0$ is updated to minimize the associated correlations; after Step (3) and (4), the final permutation matrix $\boldsymbol{S}$ is achieved whose associated correlation is close to $\boldsymbol{C}$. From above steps, it can be concluded that matrix $\boldsymbol{S}$ is determined by both the original random generated $\boldsymbol{S}_0$ and the correlation coefficient matrix $\boldsymbol{C}$. Since $\boldsymbol{C}$ is a predefined matrix, a certain permutation matrix $\boldsymbol{S}$ comes up from the initial $\boldsymbol{S}_0$.

An example is given as follows. Assume there are two random variables X_1 and X_2 to be simulated with correlation $\boldsymbol{C}$ as

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix},$$

and the sample size is $N = 10$. If the initial randomly generated matrix $\boldsymbol{S}_0$ is

$$\boldsymbol{S}_0 = \begin{bmatrix} 10 & 4 & 1 & 6 & 5 & 2 & 3 & 9 & 8 & 7 \\ 1 & 5 & 7 & 10 & 6 & 4 & 8 & 3 & 2 & 9 \end{bmatrix},$$

after Step (1) to Step (4), the rank matrix $\boldsymbol{S}$ is

$$\boldsymbol{S} = \begin{bmatrix} 10 & 4 & 1 & 6 & 5 & 2 & 3 & 9 & 8 & 7 \\ 5 & 3 & 2 & 10 & 6 & 1 & 7 & 8 & 4 & 9 \end{bmatrix}.$$

If X_1 and X_2 are standard normal distributed, the initial sample matrix $\boldsymbol{X}$ is

$$\boldsymbol{X} = \begin{bmatrix} -1.6 & -1.0 & -0.7 & -0.4 & -0.1 & 0.1 & 0.4 & 0.7 & 1.0 & 1.6 \\ -1.6 & -1.0 & -0.7 & -0.4 & -0.1 & 0.1 & 0.4 & 0.7 & 1.0 & 1.6 \end{bmatrix}.$$

After permuting initial $\boldsymbol{X}$ according to $\boldsymbol{S}$, the final sampling matrix $\boldsymbol{X}$ is achieved as

$$\boldsymbol{X} = \begin{bmatrix} 1.6 & -0.4 & -1.6 & 0.1 & -0.1 & -1.0 & -0.7 & 1.0 & 0.7 & 0.4 \\ -0.1 & -0.7 & -1.0 & 1.6 & 0.1 & -1.6 & 0.4 & 0.7 & -0.4 & 1.0 \end{bmatrix}.$$

The final correlation between the rows of X are

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0.48 \\ 0.48 & 1 \end{bmatrix},$$

which is close to the desired correlation.

3.2 Evaluate the quality of the simulated samples

The quality of the simulated samples can be checked by two kinds of parameters: correlation errors and parameters of the estimated prescribed distribution.

First, since neither LHS nor RS can simulate correlations exactly equal to the desired correlation (although they can closely approximate the correlation), for the purpose of checking the simulating accuracy on the desired correlation, the correlation errors of the samples achieved by the two methods are calculated. Assuming the associated correlation matrix of $\boldsymbol{X}$ is $\tilde{\boldsymbol{C}}$, the correlation error can be measured by ρ, which denotes the mean difference between the element values in $\tilde{\boldsymbol{C}}$ and $\boldsymbol{C}$ as follows:

$$\rho = \frac{\sum_{j=2}^{K}\sum_{k=1}^{j-1}\left|r_{kj} - \tilde{r}_{kj}\right|}{(K-1)K/2} \tag{10}$$

where $\tilde{r}_{kj}$ is the off-diagonal element in $\tilde{\boldsymbol{C}}$.

Second, for the purpose of checking whether or not the simulated samples conform to the prescribed distribution, the samples achieved by RS and LHS are employed to estimate the prescribed distribution parameters, separately. If the estimated parameters are close to the original defined parameters, it means that the distribution of the simulated samples comply with the initial prescribed distribution; otherwise, the distribution of the simulated samples is distorted from the initial prescribed distribution.

4 PROBABILISTIC POWER FLOW CALCULATION CONSIDERING WIND SPEED CORRELATION

Power flow is an important tool to study the steady-state power system operating solution. The function of deterministic power flow study can be stated as

$$\boldsymbol{z} = g(\boldsymbol{x}) \tag{11}$$

where $\boldsymbol{x}$ 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; $\boldsymbol{z}$ 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} and reactive power flow Q_{ij}.

For the P-PF problem, the inputs random variables $X_1, \cdots, X_K$ are probabilistic distributions of P_i and Q_i. When including wind power in P-PF, an additional kind of random input variables are the wind speeds of the wind farms. The output information are probabilistic distributions of V_i, θ_i, P_{ij} and Q_{ij}.

In P-PF studies considered in this paper, the following common assumptions are adopted as follows:

(1) Wind speed distribution is Weibull [13], whose cumulative distribution function is represented as:

$$Y = 1 - e^{-(x/\beta)^{\alpha}} \tag{12}$$

where β and α are scale parameter and shape parameter respectively.

(2) The nonlinear relation between wind speed v and wind farm output P_{wind} is represented as [14]

$$P_{wind} = \begin{cases} P_{rate}\frac{v-v_{ci}}{v_{rate}-v_{ci}} & v_{ci} \le v < v_{rate} \\ P_{rate} & v_{rate} \le v \le v_{co} \\ 0 & otherwise \end{cases} \tag{13}$$

where P_{rate} is the rated wind power, v_{ci}, v_{rate} and v_{co} are cut-in, rated and cut-out wind speed, respectively.

(3) Wind speeds of different wind farms are correlated according to $\boldsymbol{C}$, and the other random variables in power system are uncorrelated [1].

The steps of solving P-PF are summarized as follows:

(1) Read the sample size N, data of the power network structure, probabilistic parameters of load and traditional generations, as well as wind power parameters such as wind speed distribution parameters, wind speed correlation coefficient matrix, and wind generator parameters.

(2) Calculate the total number of input random variables in the system.

(3) Form the initial sampling matrix $\boldsymbol{X}$ by calculating N samples based on LHS method introduced in Section 2 for each input random variable. Separate the initial sampling matrix $\boldsymbol{X}$ into two sub-sampling matrices, in which sub-sampling matrix $\boldsymbol{X}_I$ is formed by the rows of wind speed samples, and sub-sampling matrix $\boldsymbol{X}_{II}$ is formed by the rows of other input random variable samples, respectively. In accordance with the permuting methods, presented in Section 3, update the sub-sampling matrix $\boldsymbol{X}_I$ by inducing the wind speed correlations, and update $\boldsymbol{X}_{II}$ by minimizing the correlations. Finally, combine $\boldsymbol{X}_I$ and $\boldsymbol{X}_{II}$ to form the updated sampling matrix $\boldsymbol{X}$.

(4) Set $n = 1$.

(5) Study the operating state which is determined by the n^{th} column of $\boldsymbol{X}$, i.e. first calculate wind power output with wind speed according to (13), then perform the deterministic power flow evaluation.

(6) Set $n = n + 1$.

(7) If $n \le N$, go to Step 5 to perform the deterministic power flow evaluation based on the operating scenario determined by the next column of $\boldsymbol{X}$; otherwise, move forward to Step 8 since all the sampled scenarios have been dealt with.

(8) Calculated the statistical results of output random variables.

5 EXAMPLE STUDIES

The sampling methods of Monte Carlo simulation and proposed Latin hypercube sampling are tested with the IEEE 14-bus and 118-bus systems. In these studies, the cut-in, rated and cut-out wind speed of wind turbine generators in wind farms are assumed to be 4, 13.61 and 25m/s, respectively [13]. The shape parameter and scale parameter of the wind speed distribution are assumed to be 1.9622 and 11.0086m/s [13]. Since the objective of this paper is to simulate the wind speeds according to a given correlation matrix, the correlation matrix C to cover representative wind speed correlation conditions, i.e. high correlation, midlevel correlation and low correlation. Therefore, the correlation indices are assigned to be 0.8 for high correlation, 0.4 and 0.5 for midlevel correlation and 0.2 for low correlation. The programs are developed with Matlab 2007a on an iMac with Intel Core i7 processor and 8 GB RAM. MATPOWER [15] is adopted to solve deterministic power flow. Since there are random factors in the simulations, except some special remarks, totally 50 trials are performed for each case and statistical results are provided. In Sections 5.1 and 5.2 numerical results are presented for the IEEE 14- and 118-bus test systems. A discussion on these numerical results and general trends are presented in Section 5.3.

5.1 IEEE 14-bus system

The data of the IEEE 14-bus system for P-PF can be found in [16]. In the original system, there are totally 23 input random variables. The generator output distributions are binomial, and load distributions are discrete and normal. In analysis in this paper, it is assumed that in total four wind farms are connected to the system, so that the total number of input random variables is 27. The capacities and connected buses of wind farms are shown in Table 1. The wind speeds in these four wind farms are correlated and the correlation matrix $\boldsymbol{C}$ is as follows:

Connected Bus	bus-7	bus-7	bus-8	bus-8
Capacity (MW)	10	15	10	15

Table 1: Wind farms connected to the IEEE 14-bus system

$$\boldsymbol{C} = \begin{bmatrix} 1.0 & 0.8 & 0.5 & 0.5 \\ 0.8 & 1.0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1.0 & 0.8 \\ 0.5 & 0.5 & 0.8 & 1.0 \end{bmatrix}$$

The correlation errors of the samples achieved by both RS and LHS are calculated, which is shown in Table 2.

And the estimated Weibull distribution parameters of correlated wind speeds simulated by RS and LHS are presented in Table 3.

Sample size N	Correlation Error	
	RS	LHS
200	0.0452	0.0220
400	0.0413	0.0182
600	0.0329	0.0169
1000	0.0286	0.0142
2000	0.0266	0.0135

Table 2: Comparison of simulated wind speed correlation error ρ between RS and LHS in the IEEE 14-bus system

Sample size N	RS		LHS	
	shape P.	scale P.	shape P.	scale P.
200	2.0737	11.2892	1.9694	11.0075
400	2.0593	11.2671	1.9659	11.0082
600	2.0666	11.2977	1.9647	11.0084
1000	2.0666	11.2902	1.9637	11.0085
2000	2.0624	11.2520	1.9630	11.0086

Table 3: Estimated Weibull parameters of wind speeds (m/s) simulated by RS and LHS in the IEEE 14-bus system

Power flow on branch 4-7 is significantly affected by wind power since this branch is connected with wind farm directly. For the sake of evaluating the performance of LHS and RS in P-PF with consideration of wind power, the active power flow on branch 4-7, which is denoted as P_{4-7}, is adopted as a representative in this paper. For simplification, statistical parameters $\mu_{P_{4-7}}$ and $\sigma_{P_{4-7}}$ are employed to show the converging condition of the active power distribution, where $\mu_{P_{4-7}}$ and $\sigma_{P_{4-7}}$ are the mean and standard deviation of distribution of P_{4-7}. The convergence curves of $\mu_{P_{4-7}}$ with RS and LHS are shown in Figures 1 and 2 respectively. Meanwhile, the convergence curves of $\sigma_{P_{4-7}}$ are shown in Figures 3 and 4. The general characteristic remained the same for all 50 trials considered.

There are small differences in the final converged P-PF results achieved by RS and LHS because different simulated wind speed lead to different wind power output in power flow calculation. In order to study these differences, P-PFs with large sample size with both methods are performed. The maximum difference of statistical parameters obtained by RS and LHS, which represented as

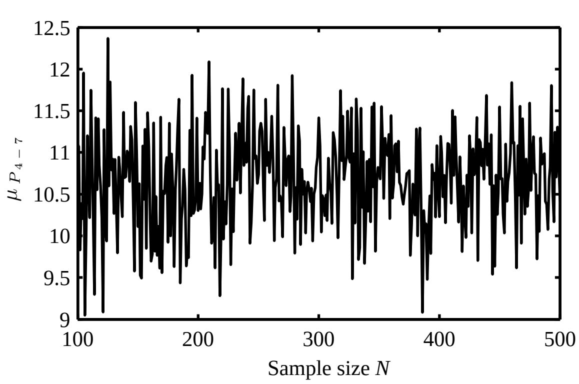

Figure 1: Convergence curves of $\mu_{P_{4-7}}$ by RS.

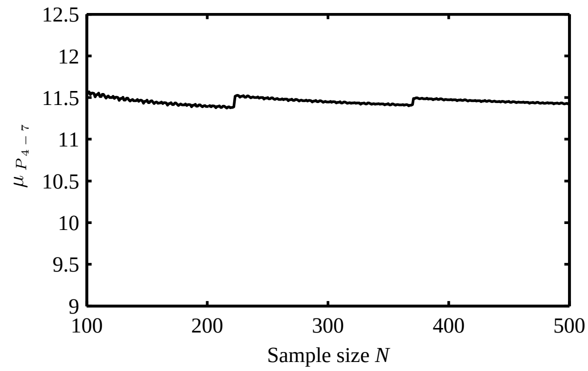

Figure 2: Convergence curves of $\mu_{P_{4-7}}$ by LHS.

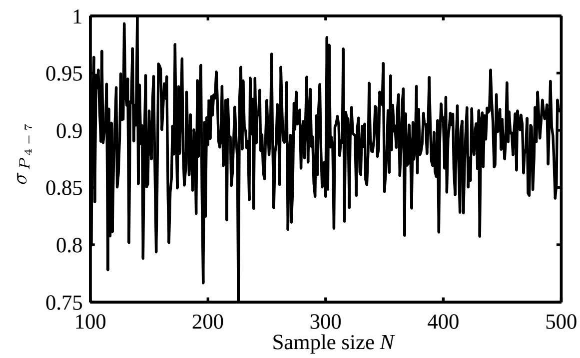

Figure 3: Convergence curves of $\sigma_{P_{4-7}}$ by RS.

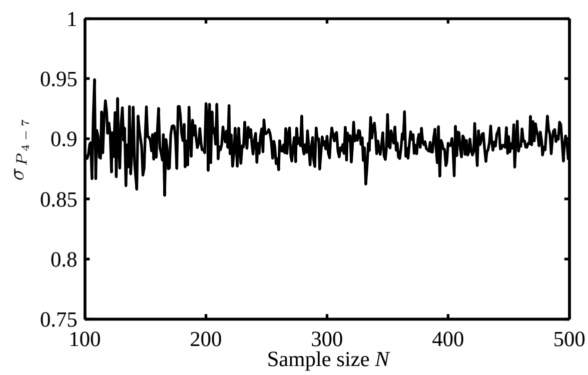

Figure 4: Convergence curves of $\sigma_{P_{4-7}}$ by LHS.

$max(\mu^*_{RS} - \mu^*_{LHS})$ and $max(\sigma^*_{RS} - \sigma^*_{LHS})$, are employed to indicate the differences, where * stands for the type of output random variables which includes V_i, θ_i, P_{ij} and Q_{ij}. The values of $max(\mu^*_{RS} - \mu^*_{LHS})$ and $max(\sigma^*_{RS} - \sigma^*_{LHS})$ in cases of $N = 10,000$ are shown in Table 4. It can be observed that the differences between voltage distribution indices are very minor. However, obvious differences exist in angle and power flow distribution indices. The angle parameters of bus-8 and bus-7 has the first and the second largest distribution differences in both mean and standard deviation. This angle distribution differences leads to power flow distribution differences.

	V_i	θ_i	P_{ij}	Q_{ij}
$max(\mu^*_{RS}-\mu^*_{LHS})$	2.5E-4	0.204	0.96	0.23
$max(\sigma^*_{RS}-\sigma^*_{LHS})$	7E-5	0.049	0.18	0.024

Table 4: Comparison of maximum difference between results of RS and LHS in the IEEE 14-bus system for $N = 10,000$. (* refers to V_i, θ_i, P_{ij} or Q_{ij})

In order to show the importance of considering wind speed correlation in P-PF problem, the power flow distributions on branch 4-7 are evaluated on condition of considering and ignoring wind speed correlations separately, and the distributions of P_{4-7} are shown in Figure 5. The distribution difference of the two curves indicates that considering wind speed correlation is necessary to achieve an accurate P-PF solution.

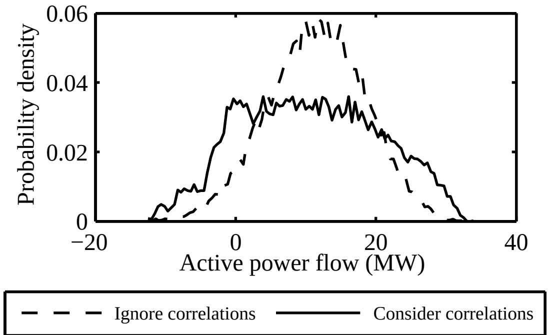

Figure 5: Comparison of distribution of P_{4-7} when wind speed correlations are considered and ignored.

5.2 *IEEE 118-bus system*

The network data of the IEEE 118-bus system are presented in [17]. The deterministic data of load and generator outputs in [17] are assumed to be the mean values in the calculated P-PF. The loads are modelled to be normal distributed and the standard deviations are assumed to be 5% of the mean. Similar with [18], it is assumed that every traditional generation plant is formed by four equal capacity units whose forced outage rates are 0.09. The distribution of generation outputs of a power plant is modelled to be binomial distribution. There are in total 170 input random variables in the original system. Assuming a total of eight wind farms are connected to bus-3, -14, -30, -33, -35, -38, -83 and -115 respectively, each with capacity 100MW, the number of input random variables is increased to 178. For this study, the wind speed correlation efficient matrix $\boldsymbol{C}$ is defined as follows:

$$\boldsymbol{C} = \begin{bmatrix} 1.0 & 0.4 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.0 \\ 0.4 & 1.0 & 0.8 & 0.8 & 0.8 & 0.8 & 0.2 & 0.2 \\ 0.2 & 0.8 & 1.0 & 0.8 & 0.8 & 0.8 & 0.2 & 0.2 \\ 0.2 & 0.8 & 0.8 & 1.0 & 0.8 & 0.8 & 0.2 & 0.2 \\ 0.2 & 0.8 & 0.8 & 0.8 & 1.0 & 0.8 & 0.4 & 0.4 \\ 0.2 & 0.8 & 0.8 & 0.8 & 0.8 & 1.0 & 0.4 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.4 & 0.4 & 1.0 & 0.2 \\ 0.0 & 0.2 & 0.2 & 0.2 & 0.4 & 0.2 & 0.2 & 1.0 \end{bmatrix}$$

The correlation errors of the samples achieved by both RS and LHS are calculated, which is shown in Table 5. And the estimated Weibull distribution parameters of correlated wind speeds simulated by RS and LHS are shown in Table 6.

The P-PF convergence with RS and LHS are evaluated by comparing errors of power flow statistical parameters. The error indices are defined as follows:

$$\varepsilon_{\mu}{}^{*} = \left| \frac{\mu_{simulated}{}^{*} - \mu_{accurate}{}^{*}}{\mu_{accurate}{}^{*}} \right| \times 100\% \tag{14}$$

$$\varepsilon_{\sigma}{}^{*} = \left| \frac{\sigma_{simulated}{}^{*} - \sigma_{accurate}{}^{*}}{\sigma_{accurate}{}^{*}} \right| \times 100\% \tag{15}$$

Furthermore, $\bar{\varepsilon}_{\mu}^{*}$ and $\bar{\varepsilon}_{\sigma}^{*}$ are defined to be the average value of $\varepsilon_{\mu}{}^{*}$ and $\varepsilon_{\sigma}{}^{*}$ respectively, which are adopted to indicate the convergence degree of the P-PF results of the whole system. Assuming the statistical results obtained from sample size $N = 10,000$ are accurate, the maximum and minimum errors of P-PF with sample size $N = 1,000$ are shown in Table 7.

Sample size N	Correlation Error	
	RS	LHS
1000	0.0390	0.0170
2000	0.0362	0.0152
3000	0.0355	0.0133

Table 5: Comparison of correlation error ρ between RS and LHS in the IEEE 118-bus system

Sample size N	RS		LHS	
	shape P.	scale P.	shape P.	scale P.
1000	2.1000	11.3717	1.9637	11.0085
2000	2.0920	11.4047	1.9630	11.0086
3000	2.0946	11.3881	1.9627	11.0086

Table 6: Estimated Weibull parameters of wind speeds (m/s) simulated by RS and LHS in the IEEE 118-bus system

Sample method		Errors	
		RS	LHS
$\bar{\varepsilon}_{\mu}^{V}$	Min	0.0008	0.0003
	Max	0.0035	0.0012
$\bar{\varepsilon}_{\sigma}^{V}$	Min	0.1624	0.1273
	Max	0.5734	0.4741
$\bar{\varepsilon}_{\mu}^{\theta}$	Min	0.1098	0.0106
	Max	1.4464	0.5898
$\bar{\varepsilon}_{\sigma}^{\theta}$	Min	0.5679	0.1711
	Max	3.3609	2.5139
$\bar{\varepsilon}_{\mu}^{P_{ij}}$	Min	1.9377	0.3664
	Max	9.6165	3.0324
$\bar{\varepsilon}_{\sigma}^{P_{ij}}$	Min	1.2690	0.5620
	Max	2.8723	1.4734
$\bar{\varepsilon}_{\mu}^{Q_{ij}}$	Min	0.7987	0.1432
	Max	4.5696	1.0769
$\bar{\varepsilon}_{\sigma}^{Q_{ij}}$	Min	1.5540	0.8044
	Max	3.1407	1.7197

Table 7: Error comparisons (%) of output random variables of the IEEE 118-bus system test (N=1000)

5.3 *Simulation Results Analysis*

It can be observed from Tables 2 and 5 that the correlation errors of LHS samples are about half of RS samples, which means that simulating correlated wind speed with proposed LHS method is more accurate than RS. In Tables 3 and 6, the estimated distribution parameters of simulated wind speeds achieved by LHS are very close to the prescribed Weibull distribution parameters, which means LHS samples retain its original distribution in the process of inducing correlations; while those parameters estimated from RS samples drift away from the prescribed Weibull distribution parameters, which means the distortion of the simulated correlated wind speeds.

The above two advantages of LHS is achieved by using the rank numbers to induce the correlation. First, the sampling values in LHS is unchanged in the process of inducing the correlation, which retains the original simulated distribution. On the contrary, in RS method, the sampling data is used directly to induce the correlation, which lead to the simulated distribution distorted from the original distribution. Second, using the rank numbers in LHS avoids generating the negative wind speeds. Comparatively, in RS method, the generated negative wind speeds is impractical and must be discarded, which causes the correlation errors.

Figures 1 to 4 and Table 7 show that when the sample size is relatively small, P-PF with LHS converges much faster than P-PF with RS, which means that LHS has a much higher sampling efficiency than RS. It can also observed from the results that LHS method works extremely well in estimating the mean values.

While the computational burdens of RS and LHS with the same sampling size are similar, the LHS method provided better results.

6 CONCLUSIONS

The main contribution of the paper is extending the work in [8] to consider the wind speed correlation between different wind farms. A generalized procedure of employing Latin hypercube sampling to deal with probabilistic power flow with consideration of wind speed correlation is provided. The proposed method not only retains the high sampling efficiency of Latin hypercube sampling, but also provides a robust and effective method to sample correlated renewable energy operating scenarios.

REFERENCES

[1] J. M. Morales and J. Perez-Ruiz, “Probabilistic Power Flow with Correlated Wind Sources,” IET Gener. Transm. Distrib., vol. 4, Iss. 5, pp 641-651, 2010

[2] B. Borkowska, “Probabilistic Load Flow,” IEEE Trans. Power App. Syst., vol. PAS-93, pp 752-759, Apr. 1974

[3] R. N. Allan, A. M. L. d. Silva, and R. C. Burchett, “Evaluation Methods and Accuracy in Probabilistic Load Flow Solutions,” IEEE Trans. Power App. Syst., vol. PAS-100, pp 2539-2546, May 1981

[4] A. Schellenberg, W. Rosehart, and J. Aguado, “Cumulant-based Probabilistic Optimal Power Flow (P-OPF) with Gaussian and Gamma Distributions,” IEEE Trans. Power Syst., vol. 20, no. 2, pp 773-781, May 2005

[5] L. A. Sanabria and T. S. Dillon, “Stochastic Power Flow Using Cumulants and Von Mises Functions,” Elect. Power Energy Syst., vol. 8, no. 1, pp 47-60, Jan. 1986

[6] M. Madrigal, K. Ponnambalam, and V. H. Quintana, “Probabilistic Optimal Power Flow,” in Proc. 1998 IEEE Can. Conf. Electrical and Computer Engineering, Waterloo, ON, Canada, pp 385-388, May 1998

[7] P. Jirutitijaroen and C. Singh, “Comparison of Simulation Methods for Power System Reliability Indexes and their Distributions,” IEEE Trans. Power Syst., vol. 23, no. 2, pp. 486-493, May 2008

[8] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, “Probabilistic Load Flow Evaluation with Hybrid Latin Hypercube Sampling and Cholesky Decomposition”, IEEE Trans. Power Syst., vol. 24, no. 2, pp 661-667, May 2009

[9] G. Papaefthymiou and B. Klockl, “MCMC for Wind Power Simulation,” IEEE Trans. Energy Convers., vol. 23, no. 1, pp 234-240, Mar. 2008

[10] A. E. Feijóo, J. Cidrás and J. L. G. Dornelas, “Wind Speed Simulation in Wind Farms for Steady-state Security Assessment of Electrical Power Systems”, IEEE Trans. Energy Convers., vol. 14, no. 4, pp 1582-1588, Dec. 1999

[11] R. L. Iman, “Latin Hypercube Sampling,” in Encyclopedia of Statistical Sciences. Hoboken, NJ: Wiley, 2006, pp 408-411

[12] R. L. Iman and W. J. Conover, “A Distribution-free Approach to Inducing Rank Correlation among Input Variables,” Commun. Statist. -Sim. Comput., pp 311-334, Jan. 1982

[13] 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 Trans. Energy Convers., vol. 23, no. 2, pp 592-602, Jun. 2008

[14] H. Yu, C. Y. Chung, K. P. Wong, and J. H. Zhang, “A Chance Constrained Transmission Network Expansion Planning Method With Consideration of Load and Wind Farm Uncertainties,” IEEE Trans. Power Syst., vol. 24, no. 3, pp 1568-1576, Aug. 2009

[15] R. D. Zimmerman and C. E. Murillo-Snchez, “MATPOWER, a MATLAB Power System Simulation Package”. Ithaca, NY: Version 3.2. Power System Engineering Research Center (PSERC), Sch. Elect. Eng., Cornell Univ., 2007

[16] R. N. Allan and M. R. G. Al-Shakarchi, “Probabilistic Techniques in AC Load-flow Analysis,” Proc. Inst. Elect. Eng., vol. 124, no. 2, pp 154-160, Feb. 1977

[17] Power Systems Test Case Archive. [Online]. Available: http://www.ee. washington.edu/research/pstca

[18] J. M. Morales and J. Perez-Ruiz, “Point Estimate Schemes to Solve the Probabilistic Power Flow”, IEEE Trans. Power Syst., vol. 22, no. 4, pp 1594-1601, Nov. 2007