Renewable Energy 116 (2018) 367–383

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

An over-limit risk assessment of PV integrated power system using probabilistic load flow based on multi-time instant uncertainty modeling

B. Rajanarayan Prusty, Debashisha Jena*

Department of Electrical and Electronics Engineering, National Institute of Technology Karnataka, Surathkal, 575025, India

ARTICLE INFO

Article history:
Received 23 May 2017
Received in revised form
23 August 2017
Accepted 27 September 2017
Available online 28 September 2017

Keywords:
Ambient temperature
PV generation
Risk assessment
Temperature-augmented probabilistic load flow (TPLF)
Uncertainty modeling

ABSTRACT

In this paper, the risk assessment of a PV integrated power system is accomplished by computing the over-limit probabilities and the severities of events such as under-voltage, over-voltage, over-load, and thermal over-load. These aspects are computed by performing temperature-augmented probabilistic load flow (TPLF) using Monte Carlo simulation. For TPLF, the historical data for PV generation, ambient temperature, and load power, each collected at twelve specific time instants of a day for the past five years are pre-processed by using three linear regression models for accurate uncertainty modeling. For PV generation data, the developed model is capable of filtering out the annual predictable periodic variation (owing to positioning of the Sun) and decreasing production trend due to ageing effect whereas, for ambient temperature and load power, the corresponding models accurately remove the annual cyclic variations in the data and their growth. The simulations pertaining to the aforesaid risk assessment are performed on a PV integrated New England 39-bus test system. The system over-limit risk indices are calculated for different PV penetrations and input correlations. In addition, the changes in the values of TPLF model parameters on the statistics of the result variables are analyzed. The risk indices so obtained help in executing necessary steps to reduce system risks for reliable operation.

1. Introduction

In recent years, power systems are more often operating under highly unpredictable conditions due to the integration of various renewable energy sources (RESs). Among the RESs, PV generation is greatly favored because of its ability to generate power at varying capacities. This results in uncertainty that sets a higher requirement on system security during planning and operation. Further, geographically nearby PV generations are correlated feed-in. The increase in uncertainty effect due to high PV penetrations and their associated correlations cause system variables to violate the limit and make the system vulnerable. Hence, risk assessment by computing risk indices based on over-limit probability and severity to recognize system weakness more realistically is entailed [1]- [2]. The calculation of risk indices are accomplished with the help of probabilistic load flow (PLF) with respect to input uncertainties and correlations. The accuracy of the computed risk indices depends on the accuracy of the PLF results. The following are the three major requirements to achieve accurate PLF results.

i) Application of an accurate uncertainty handling method,
ii) Establishment of an accurate power system model, and
iii) Accurate modeling of input uncertainties.

The various methods used for PLF are categorized as, numerical methods, analytical methods, approximate methods, and hybrid methods [3]. Monte Carlo simulation (MCS), a typical numerical method is considered as a reference for accuracy comparison of other PLF methods [3–22]. MCS provides numerical estimation of result variables based on random statistical sampling and solves the PLF problem by a series of deterministic routines.

The establishment of an accurate power system model is highly essential in PLF. A majority of the PLF studies except for [6] assume transmission branch resistance as constant. But, branch resistance depends on branch temperature which in turn is a function of a set of factors that are probabilistic in nature; among which ambient temperature is dominant. In order not to overlook the temperature related errors, temperature-augmented load flow (TLF) captures

* Corresponding author.
E-mail address: bapu4002@gmail.com (D. Jena).

https://doi.org/10.1016/j.renene.2017.09.077
0960-1481/

electro-thermal coupling effect of transmission branches [23]. The first proposal on sensitivity matrix-based temperature-augmented PLF (TPLF) model is cited in Ref. [22] and the usefulness of probability distributions of TPLF result variables for various power system studies is detailed in Table 1.

In case of TPLF, ambient temperatures of the temperature dependent branches (TDBs) are included in the input vector in addition to the bus power injections. This increases the total number of input random variables (RVs), all of which may not be modeled by any specific parametric distributions. Hence, it is an uphill task to accurately model the probability distributions and to include the associated correlations. Assumption of some parametric probability distributions to quantify the input uncertainties may not always be suitable in all cases. On the other hand, a more realistic probabilistic modeling, incorporating past experiences can be achieved from the historical data. The authors in Refs. [4] [13], [22] [24], have performed uncertainty modeling of input RVs at a particular instant of time. The uncertainty modeling of peak load power (at 7 p.m.) [4], maximum PV generation (at noon) [13] [22], [24], and ambient temperature (at noon) [22] is performed for PLF. In order to remove the trend from load samples, a fitting curve using a set of standard functions is used [4]. In Ref. [13], the periodic effect due to annual positioning of the Sun is removed from PV generation samples by filtering out the daily, seasonal, and annual periodic components whereas; in Ref. [24] the removal of predicable lowest frequency annual periodic component is accomplished with the help of a linear clear sky model. In Ref. [22], ambient temperature data is probabilistically modeled by filtering out the lowest frequency periodic component of one cycle/year. The undertone of removing the trend and the periodic effect from the historical data essentially is not to attribute their variations to a movement in uncertainty.

Although the authors in Ref. [22] successfully have augmented temperature effect in PLF analysis, the influence of variation of TPLF model parameters on the statistics of result variables is overlooked. At a specific time of the day PV generation depends on the geographical and environmental conditions of that location. At different time instants, the production patterns are different and the clear sky model for eliminating the periodic effect as proposed in Ref. [24] may not be suitable as it accounts for only the Sun's height which alone is not adequate. Hence, an accurate clear sky model taking into account the other important factors such as the Sun's direction and the angle of incidence of solar radiation deserves research attention. Similarly, multi-time instant uncertainty modeling of ambient temperature and load power needs to be equally regarded for TPLF. Further, the analysis of the impact of various PV penetrations and different input correlations on TPLF results is imperative in making the over-limit risk assessment more realistic. With this motivation, investigations are performed on the following objectives.

i) An accurate uncertainty modeling of PV generation, ambient temperature and load power at multiple time instants.
ii) An analysis of the effect of various PV penetrations and the variations of TPLF model parameters on the statistics of the result variables.
iii) Over-limit risk assessment considering various PV penetrations and input correlations.

In Section 2, the application of MCS for TPLF is systematically detailed. The input uncertainties are probabilistically modeled and correlation effects are discussed in Section 3. In Section 4, various types of over-limit risk indices are elaborated. In Section 5, modified New England 39-bus power system is used to analyze the effect of PV penetration and the value of model parameters on statistics of result variables. In addition, the system over-limit risk indices are computed for various PV penetrations and input correlations. Finally, the concluding remarks are given in Section 6.

2. PLF in temperature-augmented power system model

The power system model as developed for TLF is the basis for TPLF using MCS. TLF assumes that the power system is both in electrical and thermal steady state. It is a general conception that electrical dynamics is neglected in load flow. Again, the thermal dynamics of the branch conductors is assumed short as compared to the changes in conductor loading over time. TLF model can be developed either by considering branch resistance [25] or branch temperature [23] as state variable. The consideration of the latter simplifies the mathematics required for modeling and is computationally more efficient. The transmission branches having non-zero series resistance are referred to as TDBs. The variation in branch reactance due to temperature variation is assumed negligible as in Ref. [23]. The modeling steps of TLF are explained as under.

The resistance of a transmission branch i − j (branch connecting i^{th} bus and j^{th} bus) is expressed as,

$$R_{i-j} = R_{Ref,\ i-j}\left(\frac{T_{i-j} + T_{F,\ i-j}}{T_{Ref,\ i-j} + T_{F,\ i-j}}\right) \tag{1}$$

where T_{i-j} is the conductor temperature of the branch i − j, T_F is the temperature constant, $R_{Ref,\ i-j}$ and $T_{Ref,\ i-j}$ are the reference values of R_{i-j} and T_{i-j} respectively. According to thermal resistance model, T_{i-j} is expressed as,

$$T_{i-j} = T_{Amb,\ i-j} + T_{Rise,\ i-j} = T_{Amb,\ i-j} + R_{\theta,\ i-j}P_{Loss,\ i-j} \tag{2}$$

where, T_{Amb} and T_{Rise} are the ambient temperature and branch temperature rise above T_{Amb} respectively and R_θ is the thermal resistance. By using (2) let us define,

$$T'_{i-j} = T_{i-j} - R_{\theta,\ i-j}P_{Loss,\ i-j} = T_{Amb,\ i-j} \tag{3}$$

Since the real and reactive bus power injections (P and Q respectively) are specified, the mismatch equations ΔP and ΔQ

Table 1
Usefulness of probability distributions of the TPLF result variables.

Result variable	Adequacy indices
Bus voltage magnitude	Steady state under and over voltage probabilities can be obtained.
Branch temperature	Probability of branch temperature above the allowable maximum limit i.e., thermal over-load probability can be ensured.
Generator bus reactive power	Capability of the system to maintain bus voltage magnitudes at desired level can be evaluated.
Branch power flow	Steady state overload probabilities of the transmission branches can be identified to take decisions regarding reinforcement plans and operations.
Slack bus power	Probability of slack bus power exceeding the limit can be known.

respectively can be explicitly expressed. But, it is difficult to define the branch temperature mismatch equations since their values are unknown *a priori*. However, an equation for T' can be defined for the TDBs. Here, T' is the measure of the difference between the present value of T and the calculated temperature by using the state variables i.e., bus voltage angle δ and bus voltage magnitude $|V|$. From (3) it is clear that, the value of T' is equal to T_{Amb}. The calculation of R_θ is detailed in Ref. [22] as the ratio of the rated temperature rise $T_{\text{Rated rise}}$ to the corresponding rated loss $P_{\text{Rated loss}}$. The value of $P_{\text{Rated loss}}$ is either specified or is calculated using the fully loaded resistance R_{Hot} measured at worst-case T_{Amb} i.e., $T_{\text{Amb-wc}}$. The expression of R_{Hot} for a TDB using (1) is given as,

$$R_{\text{Hot, i-j}} = R_{\text{Ref, i-j}}\left(\frac{T_{\text{Amb-wc, i-j}} + T_{\text{Rated rise, i-j}} + T_{\text{F, i-j}}}{T_{\text{Ref, i-j}} + T_{\text{F, i-j}}}\right) \tag{4}$$

The expression for $P_{\text{Loss, i-j}}$ [23] is given as,

$$P_{\text{Loss, i-j}} = g_{\text{i-j}}\left(|V_{\text{i}}|^2 + |V_{\text{j}}|^2 - 2|V_{\text{i}}||V_{\text{j}}|\cos\delta_{\text{ij}}\right) \tag{5}$$

where $g_{\text{i-j}}$ is the conductance of branch i – j; $\delta_{\text{ij}} = \delta_{\text{i}} - \delta_{\text{j}}$. Substituting (5) in (3) yields,

$$T'_{\text{i-j}} = T_{\text{i-j}} - R_{\theta,\text{ i-j}}\, g_{\text{i-j}}\left(|V_{\text{i}}|^2 + |V_{\text{j}}|^2 - 2|V_{\text{i}}||V_{\text{j}}|\cos\delta_{\text{ij}}\right) \tag{6}$$

The mismatch equations as the difference between the specified and calculated values are given as,

$$\begin{aligned} \Delta P_{\text{i}} &= P_{\text{Sp, i}} - P_{\text{i}} \\ \Delta Q_{\text{i}} &= Q_{\text{Sp, i}} - Q_{\text{i}} \\ \Delta T'_{\text{i-j}} &= T_{\text{Amb, i-j}} - T'_{\text{i-j}} \end{aligned} \tag{7}$$

where, $P_{\text{Sp, i}}$ and $Q_{\text{Sp, i}}$ respectively are the specified real and reactive power injections at an i$^{\text{th}}$ bus, P_{i} and Q_{i} at an i$^{\text{th}}$ bus are calculated by using the basic load flow equations. In terms of temperature-augmented Jacobian matrix, mismatch equations in (7) are expressed as,

$$\begin{pmatrix} \Delta P \\ \Delta Q \\ \Delta T' \end{pmatrix} = \begin{pmatrix} \frac{\partial P}{\partial \delta} & \frac{\partial P}{\partial |V|} & \frac{\partial P}{\partial T} \\ \frac{\partial Q}{\partial \delta} & \frac{\partial Q}{\partial |V|} & \frac{\partial Q}{\partial T} \\ \frac{\partial T'}{\partial \delta} & \frac{\partial T'}{\partial |V|} & \frac{\partial T'}{\partial T} \end{pmatrix} \begin{pmatrix} \Delta\delta \\ \Delta|V| \\ \Delta T \end{pmatrix} = (J) \begin{pmatrix} \Delta\delta \\ \Delta|V| \\ \Delta T \end{pmatrix} \tag{8}$$

The error component of state vectors using (8) is obtained as,

$$\begin{pmatrix} \Delta\delta \\ \Delta|V| \\ \Delta T \end{pmatrix} = \left(J^{-1}\right) \begin{pmatrix} \Delta P \\ \Delta Q \\ \Delta T' \end{pmatrix} \tag{9}$$

where, J is the state vector Jacobian matrix.

The model so developed is referred to as a single slack bus TLF model. The main advantage of this model is that, the transmission branch temperature is directly calculated from the power flow solution using Newton-Raphson method. In TLF, the elements in the state vector are updated by using the errors estimated from (9) in each iteration. This continues until all mismatches fall within a presumed tolerance. MCS solves the TPLF by a series of N_S TLFs each time considering a set of values for input RVs. Simultaneously, N_S number of samples for each result variable is obtained. Finally, the probability distributions of the result variables are approximated.

3. Modeling of input uncertainty and description of input correlation

The circumstance in which uncertainty modeling is carried out mainly depends on the type of its application i.e., time instant model or time period model. Therefore, it is essential to decide the type of uncertainty model for implementation. Here, time-instant model is adopted. An input RV at time t is modeled as a probability distribution. The historical data for the past five years (2012–2016) is used for uncertainty modeling. The input RVs in this study includes PV generation, T_{Amb}, and load power.

3.1. PV generation uncertainty modeling

The historical data of PV generations are collected from three different locations of USA (refer Table 2) situated in the Northern hemisphere (21924 hourly values for each location) [22], [26]. The intent is to accomplish an accurate uncertainty modeling at each time instant from 7 a.m. to 6 p.m. on hourly basis. Since, the data considered for modeling is at a specific time of the day; it conflates variability (predictable PV generation variation across the year) and uncertainty (unpredictable variation due to meteorological conditions). Hence, the historical data observed at a particular time instant for several years cannot be used alone to inform the uncertainty of PV generation at that time instant. This calls for the preprocessing of the observed data by filtering out the predictable periodic effect. The stochastic variability left after removal of the effect is termed as the uncertainty component. Consequently, the variance of PV generation post processing is less than the preprocessed value [24].

The existing approach [24] to eliminate the periodic effect from data, considers the actual PV generation as a linear function of $\sin\theta_S$, where θ_S is the solar elevation angle. But, such an approach is less accurate in tracing the periodic variations (skewed and/or multimodal) of data collected at different time instants and at various locations. The function $\sin\theta_S$ is the measure of the Sun's height. But, the Sun's direction (as defined by solar azimuth angle γ_S) also has an effect on PV generation. Moreover, the research carried out in Ref. [27] indicate that array tilt θ_T significantly dominates the PV generation pattern since it affects the angle of incidence of solar radiation θ_I. The aforementioned factors cannot be ignored in the design of a linear clear sky model. The actual data in the model considered as a linear function of a set of terms, $T_1 = \sin\theta_S$, $T_2 = \sin\gamma_S$ and $T_3 = \cos\theta_I$ is proposed. In order to account for the decreasing trend in power production due to aging of

Table 2
Technical and geographical details of the PV arrays.

	PV Array 1	PV Array 2	PV Array 3
Capacity	10 kW	7.65 kW	8.88 kW
Orientation/Tilt	South/30°	South/25°	South-West/30°
Location/Post code	Monkton/21111	Parkesburg/19960	Lincoln/19365
Coordinates	39.578° (N), 76.614° (W)	39.959° (N), 75.917° (W)	38.87° (N), 75.423° (W)

Note: All the three locations are in Eastern Time zone.

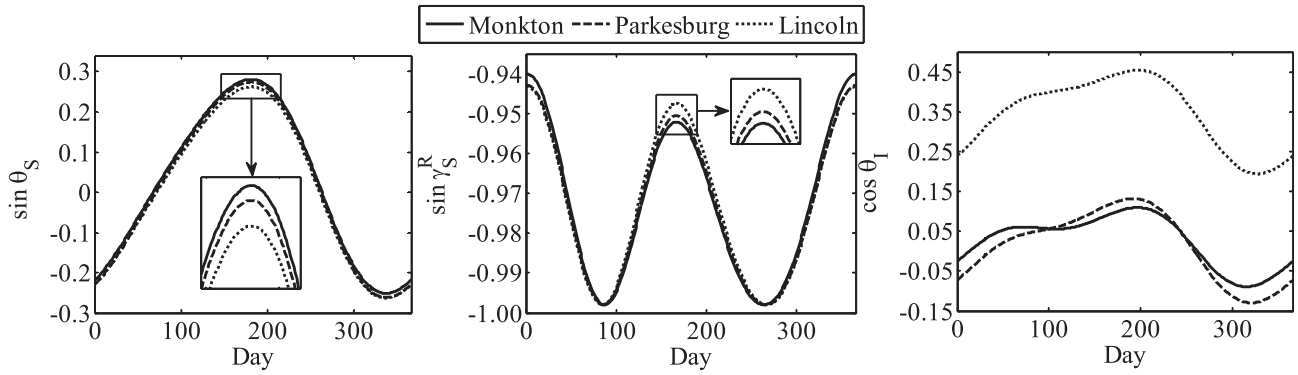


Fig. 1. Variations of $\sin\theta_S$, $\sin\gamma_S^R$ and $\cos\theta_I$ at 6 p.m. for one year.

PV array, two extra terms $T_4 = \mathrm{DN}$ and $T_5 = \mathrm{DN}^2$ are included where, "DN" is the day number of the year (DN=1 for 1st January 2012).

The linear model considering the above five terms is developed as,

$$P_{M1} = a_P T_1 + b_P T_2 + c_P T_3 + d_P T_4 + e_P T_5 + f_P \tag{10}$$

where a_P, b_P, c_P, d_P, e_P and f_P are the unknown parameters of the model which are estimated using the least squares method. The sum of squares of the errors, E is defined as,

$$E\,(a_P,\ b_P,\ c_P,\ d_P, e_P, f_P) = \sum_{k=1}^{\mathrm{nd}} \{P_{PV}(k) - P_M(k)\}^2 \tag{11}$$

where, P_{PV} is the observed PV generation, 'nd' is the total number of data samples collected at a particular time instant for five years. The values of a_P, b_P, c_P, d_P, e_P, and f_P are obtained by solving a set of following equations:

$$\frac{\partial E}{\partial a_P} = 0,\ \frac{\partial E}{\partial b_P} = 0,\ \frac{\partial E}{\partial c_P} = 0,\ \frac{\partial E}{\partial d_P} = 0,\ \frac{\partial E}{\partial e_P} = 0 \text{ and } \frac{\partial E}{\partial f_P} = 0 \tag{12}$$

The expressions for θ_S and γ_S are respectively given as,

$$\theta_S = \sin^{-1}(\sin\theta_{LA}\sin\theta_D + \cos\theta_H\cos\theta_{LA}\cos\theta_D) \tag{13}$$

$$\gamma_S = \sin^{-1}\left(\frac{-\sin\theta_H\cos\theta_D}{\cos\theta_S}\right) \tag{14}$$

where, θ_{LA} is the latitude, θ_D is the declination angle, θ_H is the hour angle. For the Northern hemisphere, θ_D is calculated as,

$$\theta_D = 23.45^\circ \sin\left(\frac{\mathrm{DN}+284}{365}\times 360^\circ\right) \tag{15}$$

The value of θ_H can be calculated by converting clock time to solar time [28]. A step by step formulation is provided underneath.

Step-1: Equation of time (EOT) is calculated as,

$$\mathrm{EOT} = 9.87\sin(2\mathrm{D}) - 7.53\cos(\mathrm{D}) - 1.5\sin(\mathrm{D}) \text{ where D} = \left(\frac{360}{365}\right)(\mathrm{DN}-81), \tag{16}$$

Step-2: For the Western longitudes, local solar time (LST) is calculated as,

$$\mathrm{LST} = \mathrm{LT} + (4\mathrm{min./deg.})\left(\theta_{LO}^{STM} - \theta_{LO}\right) + \mathrm{EOT} \tag{17}$$

where, LT is the local clock time, θ_{LO} is the local longitude of the location, θ_{LO}^{STM} is the local longitude of standard time meridian which is calculated as,

$$\theta_{LO}^{STM} = 15^\circ\left[\frac{\theta_{LO}}{15^\circ}\right] \tag{18}$$

where $\left[\frac{\theta_{LO}}{15^\circ}\right]$ returns the nearest integer to $\frac{\theta_{LO}}{15^\circ}$.

Step-3: Finally, θ_H is calculated as,

$$\theta_H = 15^\circ(\mathrm{LST} - 12) \tag{19}$$

The criteria used to obtain the required value of γ_S with reference to the true North [27] is given as,

$$\gamma_S^R = \begin{cases} 180^\circ - \gamma_S, & \cos\theta_H \geq \tan\theta_D/\tan\theta_{LA} \\ 360^\circ + \gamma_S, & \cos\theta_H < \tan\theta_D/\tan\theta_{LA} \end{cases} \tag{20}$$

By using (20), the expression for θ_I is given as,

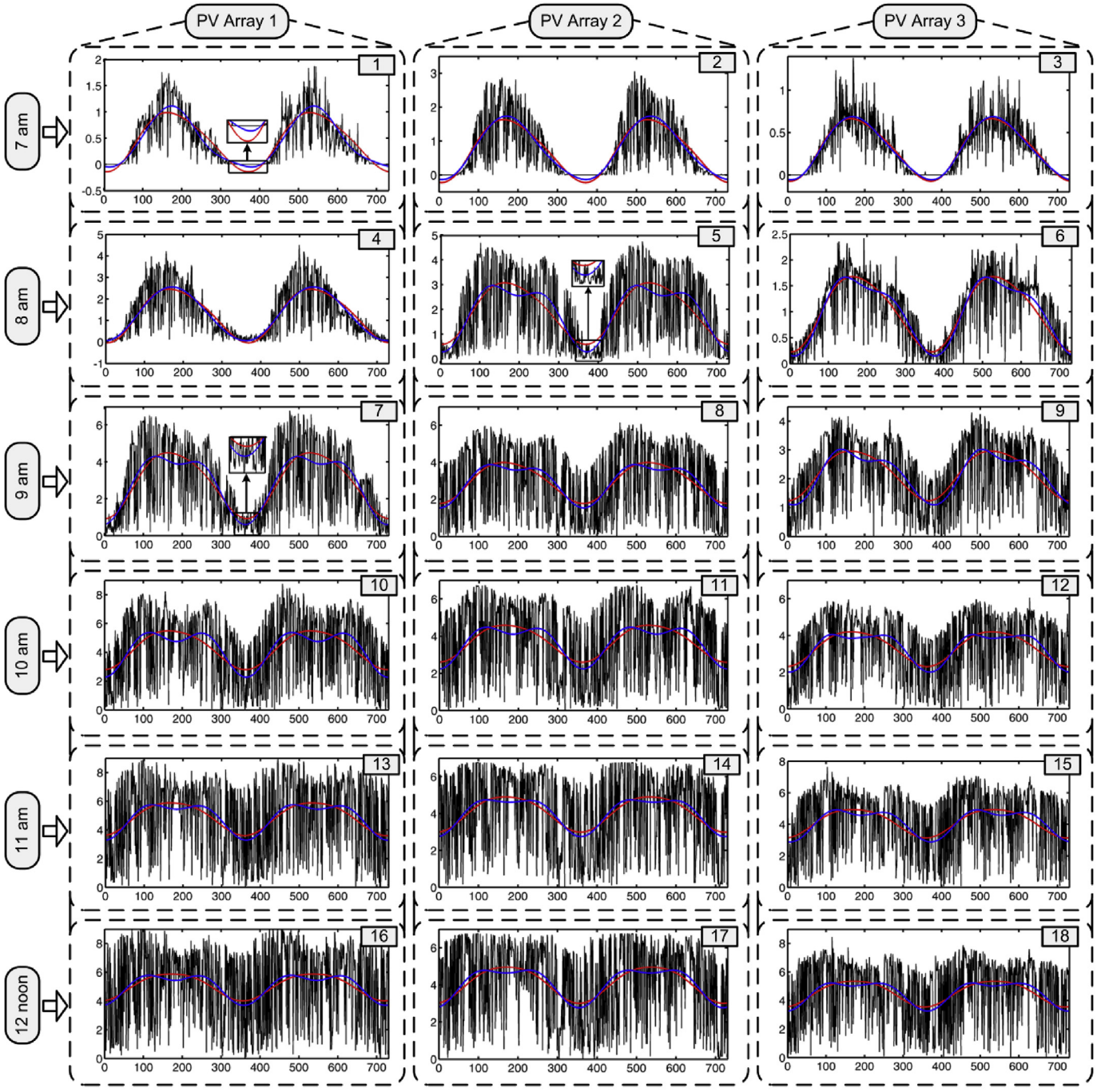
$$\theta_I = \cos^{-1}\left\{\sin\theta_S\cos\theta_T + \cos\theta_S\sin\theta_T\cos\left(\gamma - \gamma_S^R\right)\right\} \tag{21}$$

where, γ is the azimuthal angle of the PV array. For the South and the South-West facing PV arrays the values of γ are 180° and 225° respectively. Finally, the residual i.e., the unpredictable component of the PV generation is obtained as,

$$\varepsilon_{PV} = P_{PV} - P_{M1} \tag{22}$$

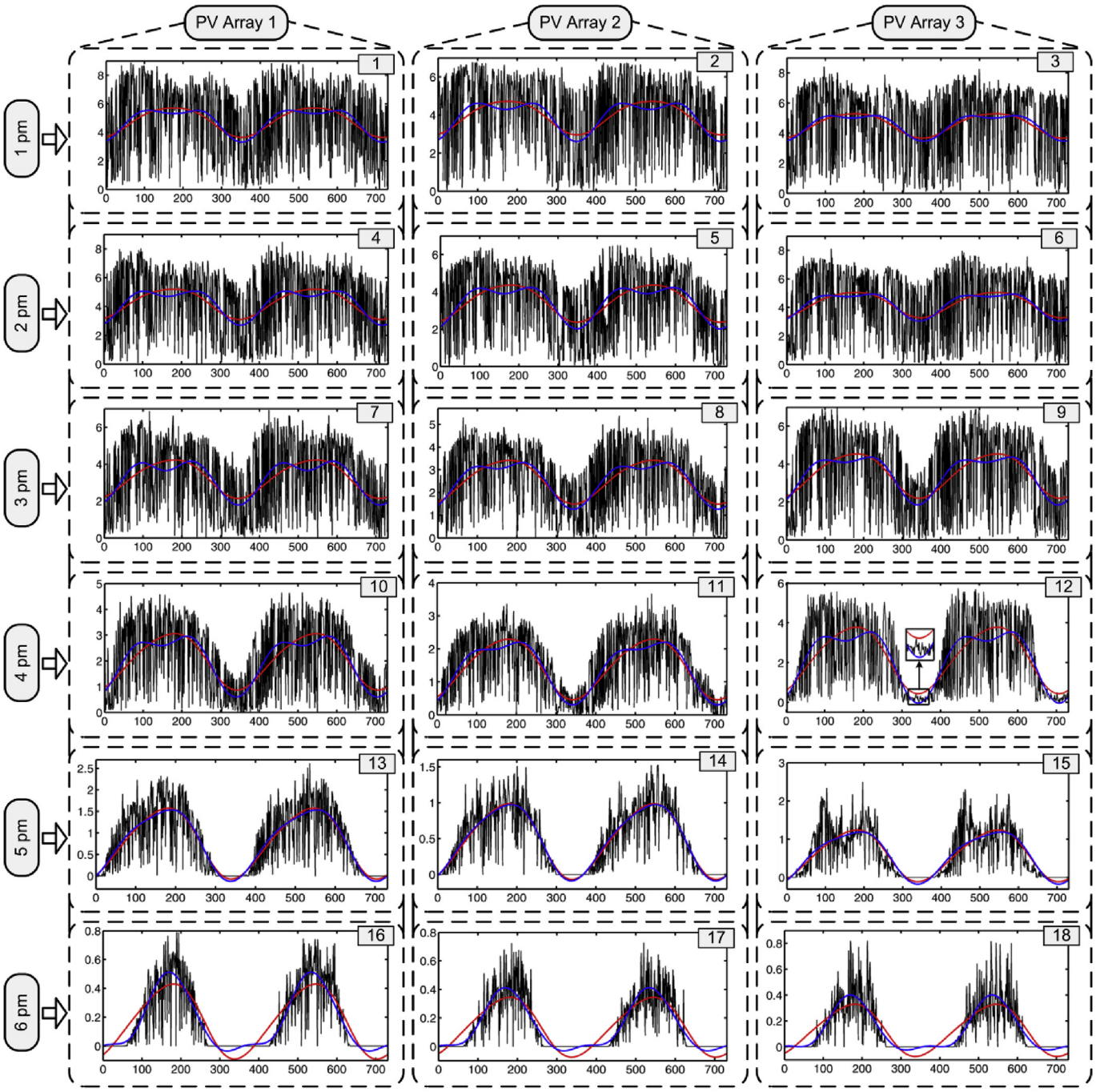
Henceforth, the model proposed in Ref. [24] is referred to as model-1 and the proposed model in (10) as model-2. The plots for $\sin\theta_S$, $\sin\gamma_S^R$ and $\cos\theta_I$ using (13), (20) and (21) respectively at 6 p.m. for one year is shown in Fig. 1. Since all the three places are situated geographically not far from each other, $\sin\theta_S$ and $\sin\gamma_S^R$ plots are nearly identical. However, a slight variation is due to the varying geographical coordinates. On the other hand, $\cos\theta_I$ for Lincoln is significantly deviating from other two owing to its differing orientation (refer Table 2). The slight difference between the plots of Monkton and Parkesburg is because of the dissimilar tilt angles. The variations depicted in Fig. 1 are different at different instants of time. A fitting curve to the observed data using these three functions at any particular time instant captures accurately the periodic pattern.

3.1.1 Preprocessing of PV generation data: The first step is to remove the effect of daylight time shifting (popularly known as daylight saving time) from the data. The observed PV generations on a daily basis (at multiple time instants) for all the three places are shown in Fig. 2. Due to space constraint, plots of only two years data (2012 and 2013) are shown. However, the uncertainty modeling considers five years of data. The PV generation at a particular instant of time is found to be periodic and the patterns are extremely differing at various time instants and locations. The estimated periodic variations using model-1 and model-2 are also shown for all the cases. Three different colors are used in the plots. The black color represents the actual PV generation; the red color and the blue color represent the periodic patterns as captured by model-1 and model-2 respectively. The mean values, standard deviations and percentage coefficient of variations (% CVs) of PV generations pre and post processing at the three arrays for twelve



(a) PV generation productions during morning hours (7 am to noon at an interval of one hour).

Fig. 2. Observed PV generation plots at twelve time instants for two years (2012 and 2013).



(b) PV generation productions during afternoon hours (1 pm to 6 pm at an interval of one hour).
Note: For all the plots, x-axis: Day, y-axis: Actual PV Generation (kW).

Fig. 2. (*continued*).

time instants are plotted in Fig. 3 considering five years data. The % CV is calculated as,

$$\%CV = \frac{\text{Standard deviation value}}{\text{Mean value}} \times 100 \tag{23}$$

The accuracy of model-2 in tracking the periodic pattern as compared to model-1 is quite evident from Fig. 2. The plots of mean value, standard deviation and % CV of PV generations of the three arrays (refer Fig. 3) have nearly the same trend. The maximum mean value and standard deviation of the PV generation occurs at noon for PV arrays 1 and 2 whereas; for PV array 3 it occurs at 1 p.m. The % CV plot is opposite to that of the mean value plot. The highest and the lowest values of % CV occur at 6 p.m. and noon respectively for all the three PV arrays. The periodic patterns as depicted by model-1 and model-2 are removed from the data using (22) to obtain the unpredictable component of PV generation. The % CV pre and post processing at all the time instants using both the models is provided in Table 3. During 10 a.m. to 3 p.m., the reduction in % CV

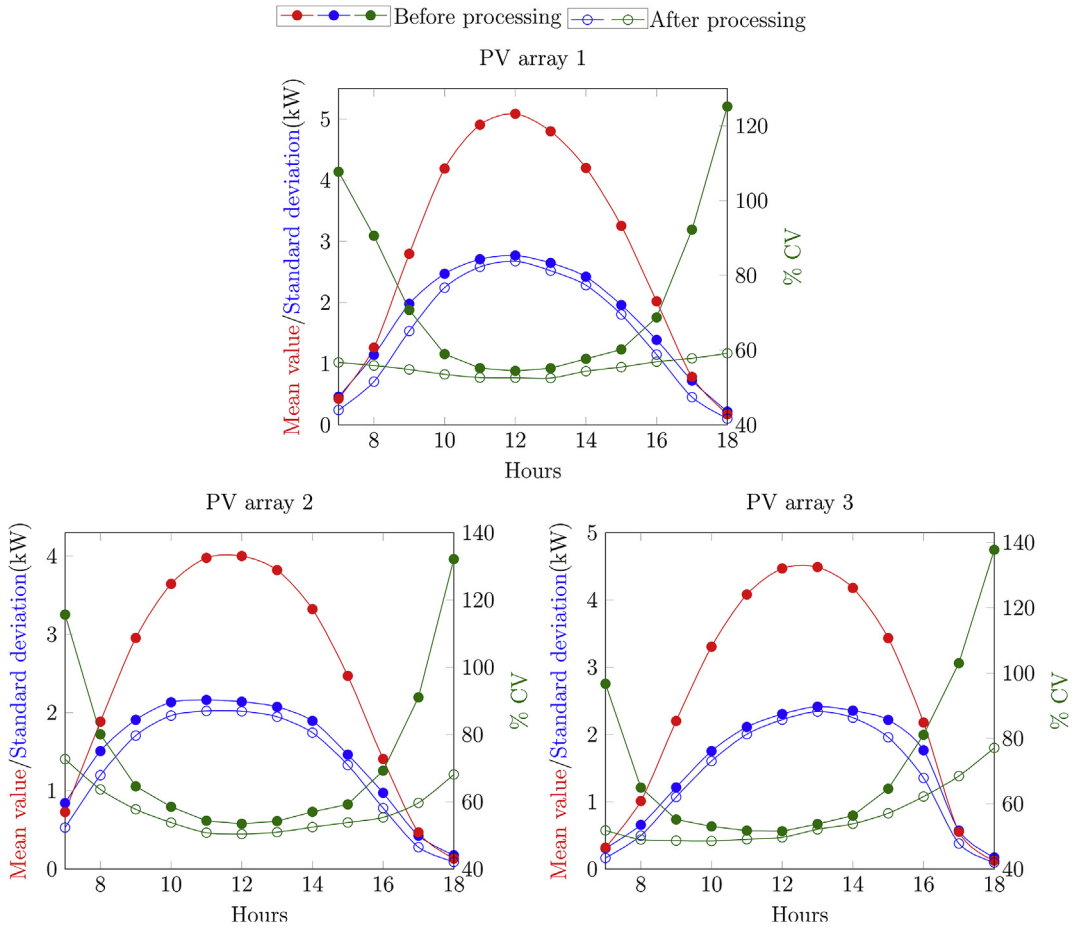


Fig. 3. Mean values, standard deviations and % CVs of PV generations pre and post processing.

post processing is less compared to that at remaining time instants.

The following are the major observations from Fig. 2 and Table 3 which illustrates the aptitude of model-2 in tracing the periodic effect due to changing solar position.

i) The insets of 5th and 7th subplots of Fig. 2(a) and the inset of 12th subplot of Fig. 2(b) clearly show that model-2 has the capability to identify the PV generation pattern in the region of low power levels.

ii) During winter season (for the months December, January and February in the Northern hemisphere) PV generation is zero at 7 a.m., 5 p.m. and 6 p.m. It is observed from Fig. 2 that during these instants there is an inaccuracy in identifying the production patterns by both the models. But, at 6 p.m., the performance of model-2 is comparatively better.

iii) At all other time instants, the overall performance of model-2 is found to be superior as compared to model-1. The reduction in % CVs in all the cases (as seen from Table 3)

Table 3
Comparison of % CVs of PV generations pre and post processing.

Time instant	% CV								
	PV Array 1			PV Array 2			PV Array 3		
	BP	AP1	AP2	BP	AP1	AP2	BP	AP1	AP2
7 a.m.	107.74	59.96	56.72	115.63	74.31	72.68	96.72	53.99	51.81
8 a.m.	90.60	57.47	55.89	80.03	64.89	63.65	64.92	51.49	49.01
9 a.m.	70.74	56.07	54.84	64.57	58.32	57.75	55.16	49.58	48.69
10 a.m.	58.96	54.13	53.52	58.50	54.30	53.81	53.05	49.15	48.56
11 a.m.	55.18	52.90	52.65	54.34	51.07	50.79	51.74	49.42	49.10
noon	54.44	52.82	52.62	53.45	50.71	50.39	51.57	49.88	49.68
1 p.m.	55.15	53.10	52.51	54.23	51.44	50.94	53.78	52.27	52.15
2 p.m.	57.66	55.26	54.37	57.00	53.32	52.46	56.36	54.01	53.85
3 p.m.	60.18	56.19	55.46	59.22	54.46	53.83	64.58	57.65	57.06
4 p.m.	68.76	57.99	56.90	69.23	56.40	55.35	81.08	64.48	62.20
5 p.m.	92.24	60.47	57.80	91.03	61.95	59.67	103.02	70.34	68.48
6 p.m.	125.20	64.37	59.16	132.12	73.34	68.11	137.74	81.32	77.08

Note: BP stands for before processing. AP1 and AP2 stands for the values after processing using model-1 and model-2 respectively.

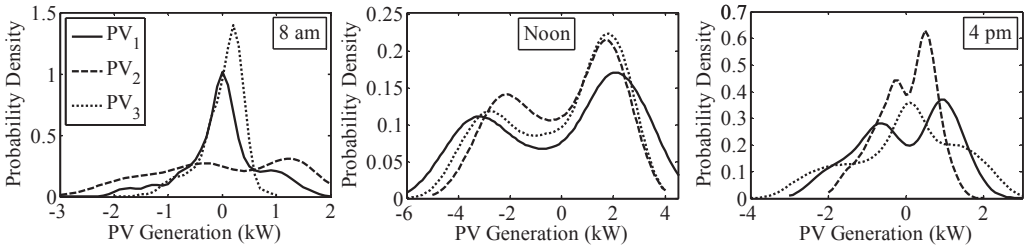


Fig. 4. Comparison of probability density plots of PV generation uncertainty at three different time instants.

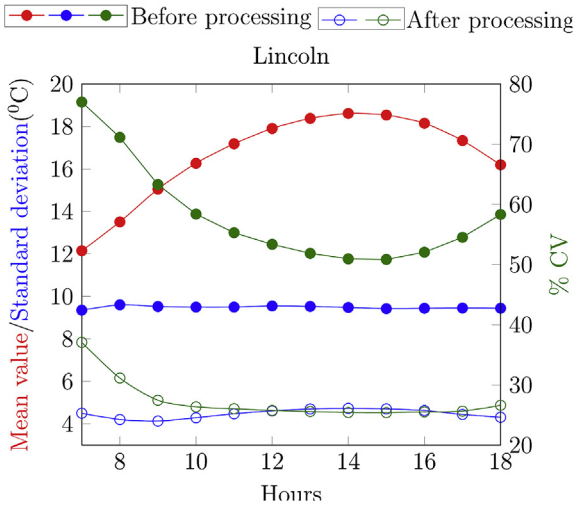


Fig. 5. Mean values, standard deviations and % CVs of T_{Amb} pre and post processing at Lincoln.

further clarifies the accuracy of model-2 in identifying the annual periodic effect at all the time instants.

The probability density plots of residuals of PV generations at 8 a.m., noon and 4 p.m. using model-2 is shown in Fig. 4. At noon, when the % CV value is low, the shapes of probability density plots are almost identical for all the three PV arrays. At 8 a.m. and 4 p.m., % CVs are comparatively higher than that at noon and the shapes of probability density plots are different. While obtaining the unpredictable components of PV generations at the time instants 7 a.m., 5 p.m. and 6 p.m., (22) is applied only during nonzero PV generation periods.

The comparison of plots in Fig. 4 indicates that, the shapes of probability density plots of PV generations are different at different locations and at different time instants. They do not fit to any specific parametric probability distributions. Hence, an assumption of any parametric distribution for PV generation uncertainty may be an obscure approach in any probabilistic analysis.

3.2. T_{Amb} uncertainty modeling

The historical data for T_{Amb} at Lincoln city, USA at twelve time instants are collected from Ref. [29]. Unlike PV generation plots, the T_{Amb} plots at various time instants have almost the same pattern. The mean values, standard deviations and % CVs at various time instants pre and post processing are plotted in Fig. 5. Standard deviation values at all the time instants pre and post processing are nearly the same, because of which the % CV plots pre and post processing are opposite to that of the mean value plot. The observed data is processed by eliminating the periodic effect. To obtain the T_{Amb} uncertainty, a linear regression model is developed which is described as,

$$P_{M2} = a_T + b_T DN + c_T DN^2 + d_T \sin\left\{\left(\frac{2\pi}{365}\right)(DN)\right\} + e_T \cos\left\{\left(\frac{2\pi}{365}\right)(DN)\right\} \quad (24)$$

where, the unknown parameters, a_T, b_T, c_T, d_T, and e_T are estimated using least squares method. Finally, the unpredictable component of T_{Amb} is obtained as, $\varepsilon_T = T_{Amb} - P_{M2}$. The second and third terms in (24) track trend in the data whereas; the last two terms determine the lowest frequency periodic component in the data.

The % CVs pre and post processing are compared in Table 4. It is clearly observed that the periodic effect has a significant impact on the % CVs of T_{Amb}. The various steps in the processing of observed T_{Amb} (at 2 p.m.) are shown in Fig. 6. In Fig. 6(a), the observed T_{Amb} plot with its periodic pattern is shown. The residual after filtering out the periodic pattern with added mean value of observed T_{Amb} is shown in Fig. 6(b). The probability density plot of T_{Amb} considering only the uncertainty component is shown in Fig. 6(c). At all the twelve time instants, Gaussian distribution fits to the data with minimal error as compared to other parametric distributions.

3.3. Load power uncertainty modeling

In case of transmission systems, the uncertainty associated with aggregate load power is generally modeled by considering the historical data [13]. Authors in Ref. [13] have processed load power data taken from Electric Reliability Council of Texas (ERCOT) at noon. So as to obtain the % CV information at other time instants, data are collected from eight weather zones of Texas i.e., Coast (C), East (E), Far West (FW), North (N), North Central (NC), South (S), South Central (SC), and West (W) during 1 a.m.–12 p.m. at an hourly interval [30]. The data represents aggregate load power for all the retail premises in the respective weather zones. The intent is to accomplish an accurate time instant uncertainty model at each hour in a day. The mean values, standard deviations and % CVs at

Table 4
Comparison of % CVs of T_{Amb} pre and post processing.

Time instant		7 a.m.	8 a.m.	9 a.m.	10 a.m.	11 a.m.	Noon	1 p.m.	2 p.m.	3 p.m.	4 p.m.	5 p.m.	6 p.m.
% CV	BP	76.99	71.12	63.29	58.38	55.28	53.35	51.85	50.94	50.85	52.04	54.51	58.32
	AP	37.07	31.14	27.46	26.37	26.05	25.77	25.59	25.42	25.40	25.52	25.64	26.62

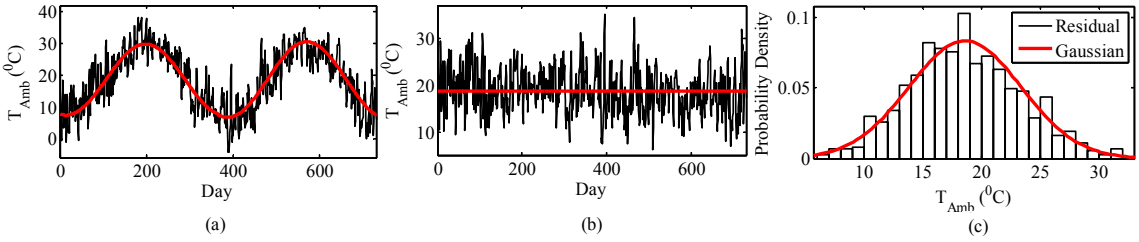


Fig. 6. Processing of T_{Amb} at 2 p.m.

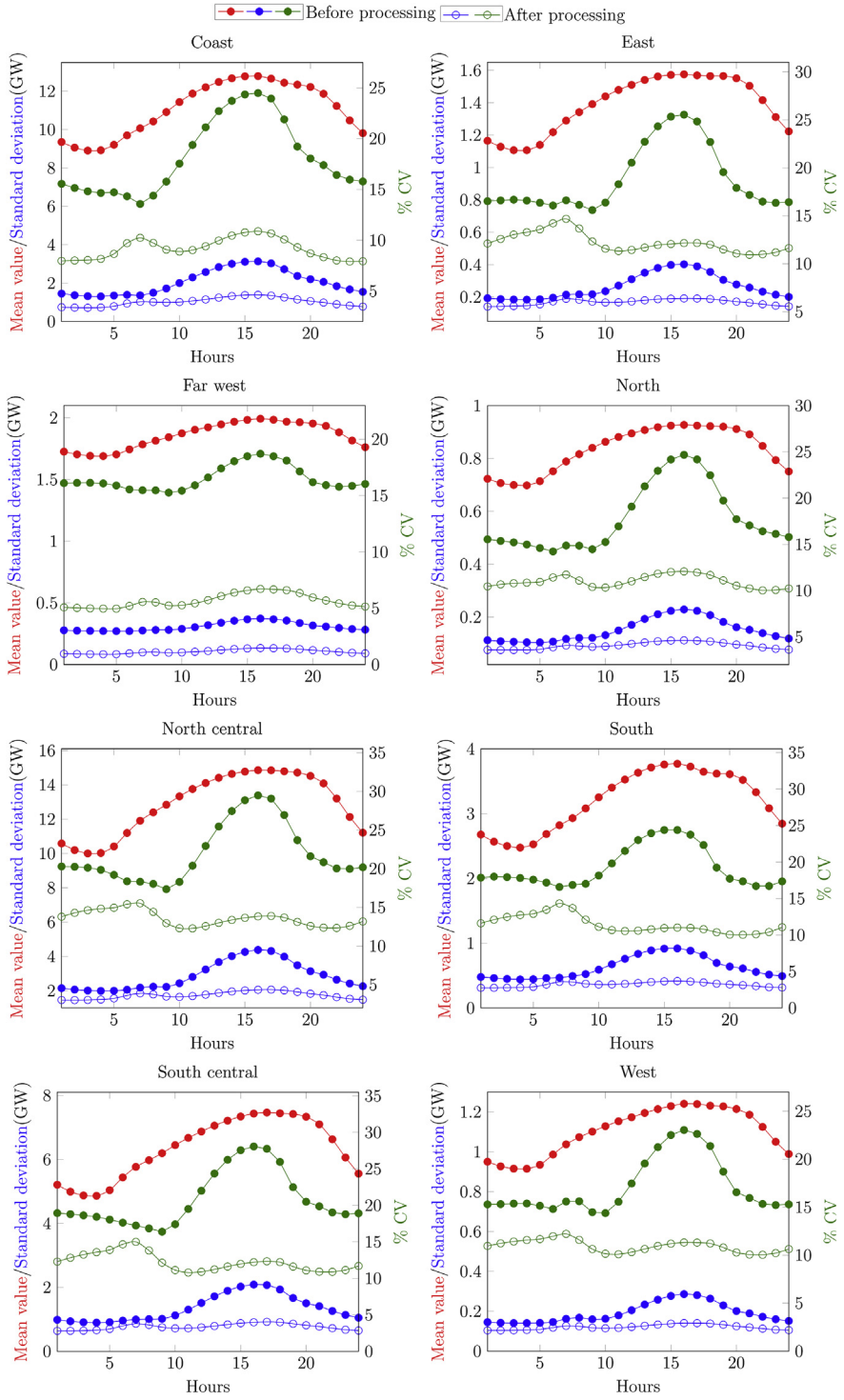


Fig. 7. Mean values, standard deviations and % CVs of aggregate real load powers pre and post processing.

Table 5
Comparison of % CVs of aggregate real load powers pre and post processing.

Time instant		7 a.m.	8 a.m.	9 a.m.	10 a.m.	11 a.m.	Noon	1 p.m.	2 p.m.	3 p.m.	4 p.m.	5 p.m.	6 p.m.
Coast	BP	13.57	14.39	15.77	17.55	19.38	21.12	22.71	23.72	24.35	24.49	23.94	21.90
	AP	10.24	9.73	9.08	8.88	9.03	9.41	9.97	10.46	10.77	10.88	10.67	10.08
East	BP	16.61	16.13	15.59	16.37	18.29	20.53	22.72	24.31	25.33	25.54	24.83	22.69
	AP	14.68	13.67	12.32	11.57	11.33	11.44	11.69	11.95	12.05	12.15	12.16	11.99
Far West	BP	15.47	15.47	15.27	15.43	15.92	16.61	17.41	18.05	18.50	18.71	18.49	18.11
	AP	5.56	5.53	5.24	5.27	5.44	5.72	6.07	6.38	6.58	6.71	6.67	6.61
North	BP	14.87	14.86	14.47	15.25	16.95	19.08	21.27	22.96	24.17	24.68	24.18	22.47
	AP	11.73	11.09	10.40	10.33	10.59	11.01	11.47	11.83	12.02	12.08	11.96	11.69
North Central	BP	18.31	18.04	17.35	18.29	20.38	22.94	25.45	27.46	28.86	29.48	29.07	26.94
	AP	15.53	14.42	12.96	12.30	12.28	12.57	13.00	13.37	13.65	13.83	13.89	13.69
South	BP	16.58	16.86	17.01	18.15	19.80	21.51	23.01	23.95	24.39	24.39	23.75	22.32
	AP	14.33	13.69	12.10	11.13	10.68	10.54	10.60	10.78	10.95	11.02	10.97	10.80
South Central	BP	17.23	16.85	16.39	17.41	19.50	22.00	24.36	26.24	27.56	28.07	27.75	25.95
	AP	15.00	13.81	12.14	11.13	10.79	10.90	11.22	11.62	11.96	12.20	12.33	12.21
West	BP	15.59	15.62	14.48	14.38	15.56	17.47	19.55	21.25	22.52	23.04	22.63	21.37
	AP	12.22	11.59	10.64	10.15	10.10	10.32	10.64	10.96	11.19	11.32	11.31	11.22

various time instants pre and post processing are plotted in Fig. 7.

Unlike PV generation patterns which are extremely varying at different time instants, the pattern of aggregate load power time series is approximately same at various time instants comprising of many harmonic components. Further, an increasing growth in the data is observed at all the time instants. A similar set of steps are followed for load uncertainty modeling as applied to T_{Amb}. To obtain load power uncertainty, a linear regression model is developed which is given as,

$$P_{\text{M3}} = a_{\text{D}} + b_{\text{D}}(\text{DN}) + c_{\text{D}}\left(\text{DN}^2\right) + d_{\text{D}}\left(\text{DN}^3\right) + \sum_{h=1}^{5} e_{\text{D},h} \sin\left\{\left(\frac{2\pi}{365}\right)h(\text{DN})\right\} + f_{\text{D},h} \cos\left\{\left(\frac{2\pi}{365}\right)h(\text{DN})\right\} \tag{25}$$

where the model parameters a_{D}, b_{D}, c_{D}, d_{D}, $e_{\text{D},1}\cdots$, $e_{\text{D},5}$, $f_{\text{D},1}\cdots$, $f_{\text{D},5}$ are determined by using least squares method. The load power data at all the 24 h is processed and % CVs at twelve time instants (7 a.m.–6 p.m. at an hourly interval) pre and post processing are presented in Table 5. As in the case with T_{Amb}, the periodic effect has a noticeable impact on % CV values. The unpredictable component is calculated as, $\varepsilon_{\text{D}} = P_{\text{D}} - P_{\text{M3}}$ where, P_{D} is the observed aggregate load power.

In Fig. 8, the observed aggregate load powers with their periodic patterns and growth at arbitrarily selected time instants 10 a.m. and 4 p.m. respectively at Coast for the years 2012–2016 are presented.

It is evident from Fig. 8 that, the periodic patterns as indicated by red color plots accurately has traced the trends and periodic effects in the data. The probability density plots of the uncertainty components of the load data for the same two time instants are shown in Fig. 9. It is found that, Gaussian distribution fits the data with minimal error and the same is applicable to the remaining time instants.

3.4. Input correlation

Spatial correlation between the input RVs are considered for TPLF simulations at various time instants. The degree of correlation is measured by using Pearson product moment correlation coefficient (PMCC). For two RVs X_1 and X_2, the PMCC is represented as, ρ_{X_1, X_2}. The plots of PMCC values between the PV generations as well as between load powers (few cases) at various time instants post processing is depicted in Fig. 10. The shapes of PMCC plots in all the three cases as shown in Fig. 10(a) are nearly the same. The PMCC values are positive due to the common effects of solar radiation, temperature and other environmental factors. The geographical distance between PV arrays 1 and 3 as well as PV arrays 2 and 3 are nearly the same (around 130 km) hence, the PMCC plots resemble closely. The distance 73 km between the PV arrays 1 and 2 is less as compared to the distance in other two cases, because of which the PMCC values in the former case are higher at all the time instants. The aggregate load power data collected from different weather zones comprise of different types of consumers such as, industrial, commercial, residential etc. The positive correlation between the load data series of different weather zones are due to the common environmental factors such as, temperature, sunset, rain fall, etc. and due to social factors such as sporting events, meal time,

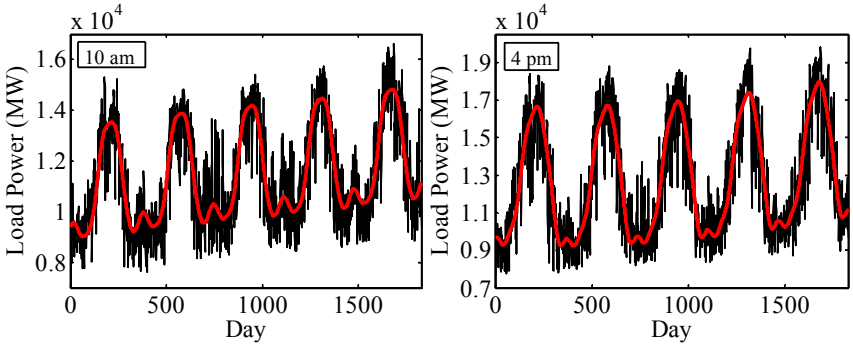


Fig. 8. Observed load power plots at Coast with periodic patterns.

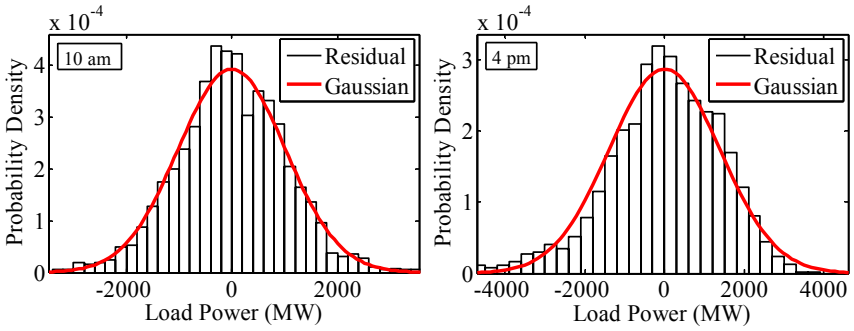


Fig. 9. Probability density plots of residuals of the real load powers as shown in Fig. 8.

working habits, etc. The shapes of PMCC plots for three cases as shown in Fig. 10(b) are nearly the same. The PMCC plots of remaining cases also have the same shapes.

Further, the calculated PMCC values between PV generation and T_{Amb} are extremely low. The values of $\rho_{PV_3,T_{Amb}}$ at various time instants are slightly higher than that of the other two cases $\rho_{PV_1,T_{Amb}}$ and $\rho_{PV_2,T_{Amb}}$. This is because PV array 3 is located in the same area from where the historical data for T_{Amb} is collected.

4. System over-limit risk indices

4.1. Over-limit probability calculations

The over-limit probability of a result variable is defined as the probability of exceeding a predefined limiting value [31]. For a result variable X, the probability of exceeding the higher limit X_{Limit} is denoted as,

$$\text{prob}(X > X_{Limit}) = L(X_{Limit}) = 1 - F(X_{Limit}) \tag{26}$$

where, $L(\bullet)$ and $F(\bullet)$ respectively are the complementary and cumulative distribution functions. Now the probability of falling below the lower limit, $X_{L,\,Limit}$ is denoted as,

$$\text{prob}(X < X_{L,\,Limit}) = F(X_{L,\,Limit}) \tag{27}$$

Hence, from the complementary distribution functions of bus voltage magnitudes, branch apparent power flows and branch temperatures, the values of under-voltage probability (UVP), over-voltage probability (OVP), over-load probability (OLP) and thermal over-load probability (TOLP) are calculated. The system UVP and OVP at the h^{th} hour are calculated as,

$$UVP_h = 1 - \prod_{i=1}^{n}(1 - UVP_{ih}),\ OVP_h = 1 - \prod_{i=1}^{n}(1 - OVP_{ih}) \tag{28}$$

where, "n" is the total number of buses in the system; UVP_{ih} and OVP_{ih} respectively are the UVP and OVP values of i^{th} bus at h^{th} hour. In the similar way, system OLP and TOLP at the h^{th} hour are calculated as,

$$OLP_h = 1 - \prod_{k=1}^{\ell}(1 - OLP_{kh}),\ TOLP_h = 1 - \prod_{k=1}^{\ell}(1 - TOLP_{kh}) \tag{29}$$

where, "ℓ" is the total number of branches in the system, OLP_{kh} and $TOLP_{kh}$ respectively are the OLP and TOLP values of k^{th} branch at the h^{th} hour.

The expected system under voltage and over-voltage times during day time (7 a.m.–6 p.m.) are calculated as,

$$UVT = \sum_{h=1}^{12}(UVP_h \times 1\ \text{hr}),\ OVT = \sum_{h=1}^{12}(OVP_h \times 1\ \text{hr}) \tag{30}$$

Similarly, the expected system over-load time (OLT) and system thermal over-load time (TOLT) during day time are calculated as,

$$OLT = \sum_{h=1}^{12}(OLP_h \times 1\ \text{hr}),\ TOLT = \sum_{h=1}^{12}(TOLP_h \times 1\ \text{hr}) \tag{31}$$

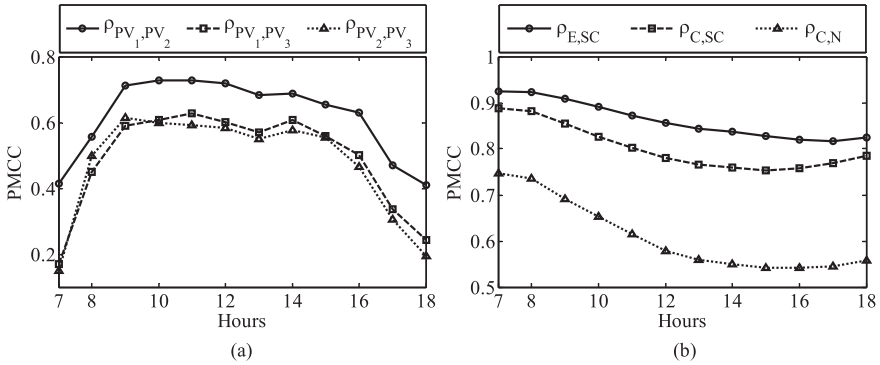


Fig. 10. PMCC between (a) PV generations and (b) a few cases of load powers at various time instants.

4.2. Over-limit severity calculations

The severity of over-limit, quantifies the deviation of mean value of the result variable from the reference value. Here, exponential severity functions are implemented. The over-voltage severity function (VSF) at i^{th} bus in h^{th} hour is expressed as,

$$VSF_{ih} = e^{k_V VD_{ih}} - 1 \tag{32}$$

where the factor $k_V = \ln(2)/(|V_{Limit}| - 1)$, VD_{ih} is the voltage deviation of the i^{th} bus at the h^{th} hour and is calculated as, $VD_{ih} = \left|1 - \mu_{|V_{ih}|}\right|$, $\mu_{|V_{ih}|}$ is the mean value of the voltage at i^{th} bus at the h^{th} hour. For $\mu_{|V_{ih}|} = 1$ pu, VSF is zero and its value equals to unity for $\mu_{|V_{ih}|} = |V_{Limit}|$.

Similarly, over-load severity function (OLSF) and thermal over-load severity function (TOLSF) of k^{th} branch at the h^{th} hour are respectively given as,

$$OLSF_{kh} = e^{k_{OL} OLD_{kh}} - 1,\ TOLSF_{kh} = e^{k_{TOL} TOLD_{kh}} - 1 \tag{33}$$

where, the factors $k_{OL} = 2\ln(2)/|S_{Limit,\,k}|$ and $k_{TOL} = 2\ln(2)/|T_{Limit,\,k}|$, $|S_{Limit,\,k}|$ and $T_{Limit,\,k}$ respectively are the apparent power flow limit and branch temperature limit of the k^{th} branch.

In (33), OLD_{kh} and $TOLD_{kh}$ respectively are the over load and thermal over load deviations of the k^{th} branch at the h^{th} hour and are calculated as,

$$OLD_{kh} = \left||S_{Th,\,k}| - \mu_{|S_{kh}|}\right|,\ TOLD_{kh} = \left||T_{Th,\,k}| - \mu_{|T_{kh}|}\right| \tag{34}$$

where, $\mu_{|S_{kh}|}$ and $\mu_{|T_{kh}|}$ respectively are the mean values of apparent power flow and temperature of the k^{th} branch at the h^{th} hour, $|S_{Th,\,i}|$ and $T_{Th,\,i}$ respectively are the threshold values of apparent power flow and temperature of k^{th} branch (taken as 50% of $|S_{Limit,\,k}|$ and $T_{Limit,\,k}$ respectively).

4.3. Calculation of risk indices

The over-limit risk index is calculated as the product of event's over-limit probability and the corresponding severity [1]. The events are under-voltage, over voltage, over-load, and thermal over-load etc. The risk of system over voltage (RSOV) at the h^{th} hour is calculated as,

$$RSOV_h = \sum_{i=1}^{n} (OVP_{ih}) \cdot (VSF_{ih}) \tag{35}$$

Finally, the risk of system over load (RSOL) and risk of system thermal over load (RSTOL) at the h^{th} hour are respectively calculated as,

$$RSOL_h = \sum_{k=1}^{\ell} (OLP_{kh}) \cdot (OLSF_{kh}),\ RSTOL_h = \sum_{k=1}^{\ell} (TOLP_{kh}) \cdot (TOLSF_{kh}) \tag{36}$$

5. Case study and discussion of results

MCS is applied for PLF and TPLF on modified New England 39-bus system as shown in Fig. 11 by considering 30000 samples for each input RV. This number is ascertained by setting variance coefficient to less than 1% for all the result variables [3].

5.1. Power system description and statistical details of input RVs

The power system data is adopted from Ref. [32]. The slack generator is connected at bus 31. The system base power is 100 MVA. The base value of branch temperature is assumed 100 °C; it can however be any convenient value. Its only function is to normalize the temperature scale to aid computational simplicity. It does not carry any relationship with voltage and power base values, implied or otherwise. The three PV arrays as discussed in Section 3.1 are included at buses 26, 27 and 28. The PV arrays are assumed to be not providing voltage support to the system hence, the reactive power generations are zero. The loads connected at buses 9, 18, 21, 23, 24, 25, 26, 27, 28 and 29 are assumed as RVs. The discrete load instants and their corresponding probability values for real and reactive load powers at bus numbers 9 and 18 are specified in Table 6. The real load powers at buses 21, 23, 24, 25, 26, 27, 28 and 29 follow Gaussian distribution with % CVs (post processing) of the load data of eight weather zones as indicated in Table 5 for the respective time instants. The mean values are considered as specified deterministic data. The load power factors at these buses are assumed constant. T_{Amb} is assumed same for all the TDBs. The base case PMCC matrix for TPLF is constituted among 20 input RVs which include PV generations of the three PV arrays, T_{Amb}, real and reactive load powers at buses 21, 23, 24, 25, 26, 27, 28 and 29. The PMCC between PV generations and load power is assumed as 0.3. The other PMCC values are calculated from the historical data post processing.

5.2. Impact of increased penetration of PV generations on the statistics of result variables

The percentage penetration level (% PL) of PV generation (percentage of total system real load power) is expressed as,

$$\%PL = \frac{\text{mean value of total PV generation}}{\text{mean value of total system load}} \times 100 \tag{37}$$

The mean value of total PV generation for a given % PL is obtained by using (37). It is assumed that the obtained mean value corresponds to the time at which maximum PV generation occurs (usually at noon). Then after, it is shared among the three PV units. Let, μ_1^{12}, μ_2^{12}, and μ_3^{12} are the assigned mean values of the PV generations of three arrays based on their ratings as specified in Table 2 where, the superscript "12" corresponds to noon. By adopting this

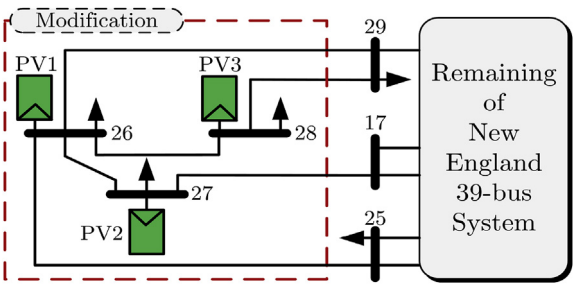


Fig. 11. PV integrated New England 39-bus system.

Table 6
Probabilistic description of discrete load powers.

Bus 9		Bus 18		Probability value
Real (pu)	Reactive (pu)	Real (pu)	Reactive (pu)	
0.04	−0.61	1.32	0.15	0.10
0.05	−0.64	1.46	0.20	0.15
0.06	−0.67	1.58	0.28	0.30
0.07	−0.68	1.66	0.34	0.25
0.09	−0.69	1.70	0.43	0.20

process, the method to compute the mean values of individual PV generations is explained by considering 5% PL as an example. The mean value of the total PV generation by using (37) is calculated as 3.1271 pu (total system real load power is 62.5423 pu). The ratio of ratings of PV_2 to PV_1 and PV_3 to PV_1 is calculated as 0.765 and 0.888 respectively. Hence, the value 3.1271 pu is shared by the PV arrays as, $\mu_1^{12} = 1.1787$ pu, $\mu_2^{12} = 0.9017$ pu, and $\mu_3^{12} = 1.0467$ pu. Now the samples of the PV generation at a k^{th} place i.e., PV_k^{12} corresponding to a given % PL is calculated by multiplying a factor (ratio of μ_k to the mean of actual PV generation at k^{th} place) with all the samples of actual PV generation collected at that location. These set of steps are applied to various other % PLs for obtaining the mean values of PV generations at noon as indicated in Table 7.

In order to analyze the effect of increased PV penetrations on the statistics of result variables, various % PLs such as, base case, 5%, 10%, 15%, and 20% are deliberated. The probability density plots of PV generation pertaining to various % PLs at noon for all the three PV arrays are shown in Fig. 12. It is observed that with an increase in % PL, the variance of PV generation increases. Further, the probability density plots are magnified but the shape remains same. The cumulative probability plots of net real power load at buses 26, 27 and 28 for various PLs are plotted in Fig. 13. The net load power variability in all the three cases is increased with increase in % PL. Further, the lower tails of the distributions are shifted towards the negative axis. In the base case, low PV penetration could not cause multimodality in the cumulative probability of net load power. For the remaining four % PL cases cumulative probability plots of net load are multimodal at buses 26, 27 and 28. It can further be observed that for % PL above 10, bidirectional power injections are evident at buses 27 and 28 whereas the same is noticed above 5% at bus 26.

The buses and branches in the vicinity to the PV array locations are subjected to more uncertainty influences [13]. Hence, the probability distributions of bus voltage magnitudes of buses 26, 27 and 28; branch temperatures and power flows in the branches 26–27 and 26–28 are analyzed. Both PLF and TPLF simulations are carried out at noon by considering various PV penetration cases as described in Table 7 and the results are compared in Table 8. Henceforth, $P_{L,\,i-j}$, $Q_{L,\,i-j}$ and $|S_{L,i-j}|$ are used to represent respectively the real, reactive, and apparent power flows in the branch $i-j$. P_{31} and Q_{31} respectively are the slack bus real and reactive powers. From the comparison of results in Table 8 it is clear that the increase in % PL has a noticeable effect on branch temperatures and branch loadings. The PLF and TPLF results are nearly same for bus voltage magnitudes indicating that the temperature-augmentation has less effect on bus voltage magnitudes. However, temperature-augmentation has a significant effect on the variance of real and reactive power flows. The effect becomes more prominent for increased % PLs. In all the above three cases for bus voltage magnitudes with the increase in % PL beyond 5, mean values are decreased whereas the standard deviation values are increased due to the uncertainty influence of PV generations at these buses. However in the base case, the patterns for mean values and standard deviations are significantly biased as compared to the other cases, this is because of the bidirectional power injections at these buses. TPLF provides the statistical information of branch temperatures of TDBs which PLF fails to do. Due to temperature-augmentation, real power flow variability of the branch 26–27 using TPLF is significantly increased as compared to PLF. There is a fixed pattern to explain the variations in the mean values and standard deviations of real power flow. The slack bus powers are radically changed especially for 15% and 20% penetration cases. The slack bus absorbs the excess real power; this scenario resembles in reality the power export to the nearby systems. The analysis of the other branch power flow distributions indicate that, in most of the cases the increase in % PL leads to bidirectional power flows where the probability distributions are extended to both positive and negative axes. This increase in power flow variability due to the increase in % PL leads to over-limit risks in the system. Hence, system reinforcements are essential.

Table 7
Mean values of PV generations for various % PLs.

PV Array	Base case	% PL			
		5	10	15	20
PV_1	0.0283	1.1787	2.3574	3.5361	4.7148
PV_2	0.0222	0.9017	1.8034	2.7051	3.6068
PV_3	0.0238	1.0467	2.0934	3.1401	4.1867

5.3. Effect of variation of TPLF model parameters on the statistics of result variables

The base case values of model parameters T_{Ref}, $T_{Rated\ rise}$, and T_{Amb-wc} are 10 °C, 25 °C, and 40 °C respectively. T_{Ref} is the temperature at which branch resistance is initially specified. $T_{Rated\ rise}$ is the rise in temperature that is expected to occur at rated power level. Typically, this is dictated by line sag considerations, allowable operating temperature of insulating materials, etc. T_{Amb-wc} is the worst case value of T_{Amb}. It is used for the estimation of R_{Hot} and therefore estimates $P_{Rated\ loss}$. The above three parameters cannot always be taken as constant and the effect of their variation on the statistics of the result variables would be of interest. The analysis is carried out at noon for 5% PL by considering a few result variables associated with the branches where the effect of temperature dependence is the highest. Due to consideration of temperature dependence, resistances of all the TDBs are increased. In order to quantify the effect of change in first two statistical parameters due to temperature-augmentation, two relative percentage error indices are defined which are given as,

$$e_\mu = \left|\frac{\mu_{WT} - \mu_T}{\mu_{WT}}\right| \times 100,\ e_\sigma = \left|\frac{\sigma_{WT} - \sigma_T}{\sigma_{WT}}\right| \times 100 \tag{38}$$

where, μ_{WT} and μ_T respectively are the mean values of a particular result variable as obtained using PLF and TPLF; σ_{WT} and σ_T respectively are the standard deviation values of a particular result variable as obtained using PLF and TPLF.

There is a significant increase in the value of e_μ for branch temperature. Average e_μ for all the branch temperatures amounts to 123.84%. The higher value is due to the consideration of temperature effect in TPLF which is ignored in case of PLF. On the other hand, the average e_μ of branch power flow and branch power loss respectively amounts to 5.65% and 6.93% whereas, average e_σ of those respectively amounts to 26.29% and 33.39%. The effect of temperature-augmentation on other result variables is less. The values of e_μ and e_σ of branch power flows in case of a few branches

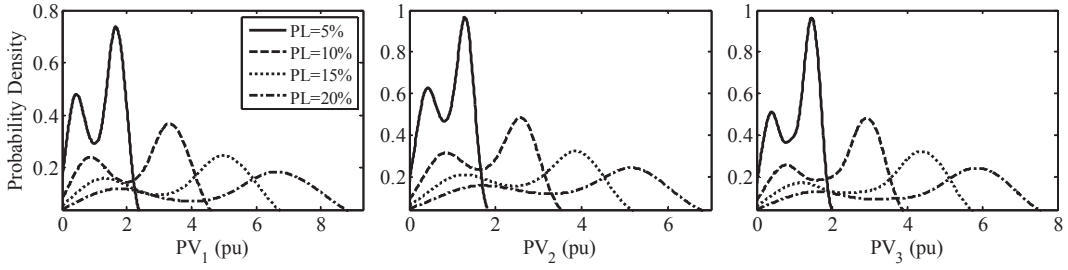


Fig. 12. Comparison of probability density plots of PV generations for various % PLs.

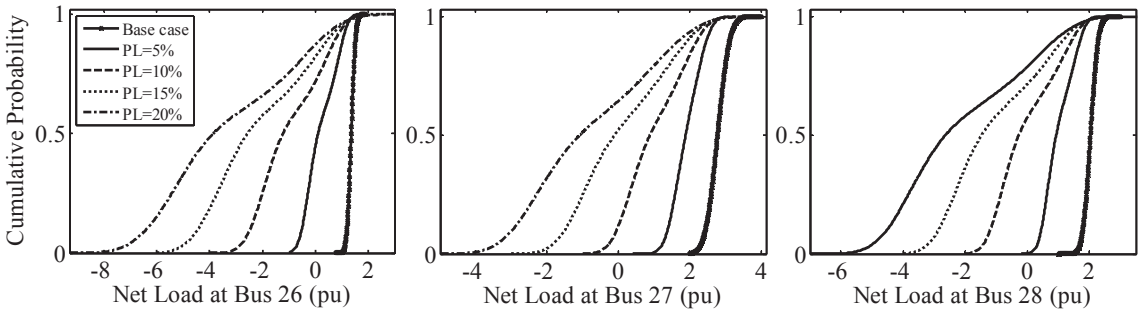


Fig. 13. Cumulative probability plots of net load at buses 26, 27 and 28 for various % PLs.

where the effect of temperature dependence is prominent are indicated in Table 9. It is observed that, the temperature-augmentation has led to increase in power flow variability.

The effect of variations of model parameter values on the average of error indices e_μ and e_σ of power flows (both real and reactive) are provided in Table 10. In majority of the cases, the increase in model parameter values either increases or decreases the values of error indices. But, in a few cases though the effect is prominent, there is no fixed pattern to explain the change in error indices.

5.4. Over-limit risk assessment

The probability distributions of result variables as obtained by TPLF are useful in obtaining the over-limit risk indices under various PV penetrations and input correlations. All the system

Table 8
Comparison of PLF and TPLF results for various % PLs of PV generation.

Result variable		PLF					TPLF				
		Base Case	PL = 5%	PL = 10%	PL = 15%	PL = 20%	Base Case	PL = 5%	PL = 10%	PL = 15%	PL = 20%
$\lvert V_{26}\rvert$	μ	1.0520	1.0524	1.0486	1.0403	1.0267	1.0519	1.0526	1.0493	1.0419	1.0302
	σ	0.0022	0.0016	0.0056	0.0142	0.0285	0.0023	0.0016	0.0050	0.0126	0.0247
$\lvert V_{27}\rvert$	μ	1.0378	1.0375	1.0324	1.0223	1.0063	1.0374	1.0372	1.0324	1.0228	1.0080
	σ	0.0026	0.0022	0.0069	0.0168	0.0331	0.0026	0.0022	0.0066	0.0159	0.0306
$\lvert V_{28}\rvert$	μ	1.0501	1.0512	1.0496	1.0453	1.0376	1.0500	1.0513	1.0502	1.0465	1.0400
	σ	0.0018	0.0010	0.0028	0.0078	0.0165	0.0019	0.0010	0.0024	0.0067	0.0140
T_{26-27}	μ	0.1000	0.1000	0.1000	0.1000	0.1000	0.2185	0.2510	0.2963	0.3559	0.4349
	σ	0	0	0	0	0	0.0498	0.0551	0.0751	0.1169	0.1859
T_{26-28}	μ	0.1000	0.1000	0.1000	0.1000	0.1000	0.1897	0.2020	0.2186	0.2406	0.2698
	σ	0	0	0	0	0	0.0486	0.0500	0.0543	0.0654	0.0846
$P_{L,\,26-27}$	μ	2.5941	3.5371	4.4996	5.4299	6.3454	2.5975	3.5518	4.5194	5.4645	6.3937
	σ	0.3041	0.4963	0.9299	1.3667	1.7890	0.3063	0.4985	0.9353	1.3846	1.8226
$P_{L,\,26-28}$	μ	−1.4228	−2.0524	−2.6951	−3.3272	−3.9718	−1.4200	−2.0514	−2.6925	−3.3267	−3.9537
	σ	0.2702	0.3536	0.6269	0.9270	1.2288	0.2696	0.3525	0.6246	0.9227	1.2251
$Q_{L,\,26-27}$	μ	0.6792	0.6794	0.7288	0.8263	0.9712	0.6840	0.6873	0.7431	0.8502	1.0131
	σ	0.0427	0.0418	0.0886	0.1872	0.3405	0.0421	0.0424	0.0935	0.2007	0.3679
$Q_{L,\,26-28}$	μ	−0.2159	−0.1253	−0.0459	0.0194	0.0744	−0.2117	−0.1171	−0.0315	0.0451	0.1146
	σ	0.0457	0.0475	0.0780	0.1182	0.1698	0.0470	0.0500	0.0839	0.1296	0.1888
$\lvert S_{L,26-27}\rvert$	μ	2.6834	3.6029	4.5595	5.4939	6.4223	2.6879	3.6187	4.5814	5.5318	6.4768
	σ	0.2904	0.4898	0.9276	1.3733	1.8104	0.2926	0.4923	0.9336	1.3929	1.8476
$\lvert S_{L,26-28}\rvert$	μ	1.4418	2.0575	2.6973	3.3298	3.9764	1.4385	2.0561	2.6946	3.3298	3.9597
	σ	0.2589	0.3493	0.6239	0.9254	1.2280	0.2583	0.3484	0.6224	0.9217	1.2257
P_{31}	μ	6.6174	3.6556	0.6932	−2.1857	−4.9783	6.6461	3.6737	0.7318	−2.1042	−4.8050
	σ	1.6653	1.6686	2.5176	3.5503	4.6210	1.6563	1.6504	2.4843	3.4861	4.4277
Q_{31}	μ	2.1702	1.6881	1.6305	1.9856	2.7494	2.1845	1.7048	1.6552	2.0165	2.7717
	σ	0.3973	0.1568	0.1317	0.7033	1.6449	0.3950	0.1538	0.1344	0.7010	1.5991

Table 9
Effect of temperature-augmentation on first two statistical moments of power flows in a few branches.

Branch	% relative change in		Real power		Reactive power		Apparent power	
	Branch resistance	Branch temperature	e_μ	e_σ	e_μ	e_σ	e_μ	e_σ
02–03	08.73	207.91	0.36	2.17	2.42	18.58	0.43	2.64
02–25	08.30	197.72	0.48	1.73	5.54	15.13	0.06	0.56
06–11	10.45	248.70	0.13	1.13	5.62	3.77	0.13	1.13
10–11	07.73	184.11	0.08	0.99	1.56	12.40	0.10	1.09
15–16	07.34	174.80	0.04	0.89	0.67	4.49	0.11	1.01
16–19	08.27	196.83	0.06	5.46	0.97	0.04	0.07	2.16
21–22	07.26	172.78	0.03	0.74	0.23	1.57	0.04	0.85
23–24	06.21	147.98	0.01	0.93	10.49	2.77	0.01	0.94
26–27	06.34	150.91	0.42	0.45	1.16	1.62	0.44	0.50

Table 10
Effect of variations in values of TPLF model parameters on average e_μ and e_σ of power flows.

Result variable	Average of e_μ				Average of e_σ			
	$T_{Ref} = 10\ ^\circ C$	$T_{Ref} = 15\ ^\circ C$	$T_{Ref} = 20\ ^\circ C$	$T_{Ref} = 25\ ^\circ C$	$T_{Ref} = 10\ ^\circ C$	$T_{Ref} = 15\ ^\circ C$	$T_{Ref} = 20\ ^\circ C$	$T_{Ref} = 25\ ^\circ C$
Power flow	1.6932	4.0455	2.9946	4.0364	4.0603	3.6370	3.5356	3.4557
Result variable	$T_{Rated\ rise} = 25\ ^\circ C$	$T_{Rated\ rise} = 30\ ^\circ C$	$T_{Rated\ rise} = 35\ ^\circ C$	$T_{Rated\ rise} = 40\ ^\circ C$	$T_{Rated\ rise} = 25\ ^\circ C$	$T_{Rated\ rise} = 30\ ^\circ C$	$T_{Rated\ rise} = 35\ ^\circ C$	$T_{Rated\ rise} = 40\ ^\circ C$
Power flow	1.6932	1.9870	2.0921	2.2813	4.0603	4.4332	4.7729	5.0793
Result variable	$T_{Amb-wc} = 40\ ^\circ C$	$T_{Amb-wc} = 50\ ^\circ C$	$T_{Amb-wc} = 60\ ^\circ C$	$T_{Amb-wc} = 70\ ^\circ C$	$T_{Amb-wc} = 40\ ^\circ C$	$T_{Amb-wc} = 50\ ^\circ C$	$T_{Amb-wc} = 60\ ^\circ C$	$T_{Amb-wc} = 70\ ^\circ C$
Power flow	1.6932	1.6509	1.6259	1.5471	4.0603	4.2205	3.9791	3.8785

Table 11
UVP and OVP values of load buses at noon for 20% PL.

Bus	UVP	OVP	Bus	UVP	OVP	Bus	UVP	OVP
1	0	0.1069	11	0.1406	0.0001	21	0	0.1896
2	0.0774	0.0437	12	0.0142	0.0033	22	0	0.0741
3	0.2860	0.0028	13	0.1043	0.0053	23	0	0.0595
4	0.1436	0.0025	14	0.0525	0.0045	24	0	0.7212
5	0.1039	0.0035	15	0	0.0344	25	0	0.2459
6	0.1966	0.0002	16	0.0093	0.0723	26	0.0055	0.1180
7	0.1891	0.0001	17	0.0418	0.0545	27	0	0.2692
8	0	0.0343	18	0	0.1863	28	0	0.2467
9	0.0014	0.0035	19	0	0	29	0	0.0082
10	0.0207	0.0028	20	0	0.0095			

buses and branches are considered for evaluating over-limit risks. Table 11 indicates the UVP (<0.95 pu) and OVP (>1.05 pu) at noon for 20% PL at all the load buses. Table 12 shows OLP ($>|S_{Limit}|$) and TOLP ($>T_{Limit}$) values of TDBs at noon for 20% PL. The value of $|S_{Limit}|$ is the MVA rating of the branch and T_{Limit} is set to 0.5 pu, since a maximum allowable temperature of value 50 °C is typically selected to avoid loss of strength, sag and branch losses, etc. The over-limit probability values are calculated from the respective complementary distribution functions. The system UVP, OVP, OLP and TOLP values for various % PLs are compared in Table 13 with and without considering input correlations. The over-limit probability values differ significantly by accounting for input correlation as compared to the case without considering correlation. The error becomes more noticeable when the % PL increases. It is clearly

Table 12
OLP and TOLP values of TDBs at noon for 20% PL.

Branch	OLP	TOLP	Branch	OLP	TOLP	Branch	OLP	TOLP	Branch	OLP	TOLP
1–2	0	0	5–8	0	0	14–15	0.0930	0.0086	21–22	0	0
1–39	0	0	6–7	0	0	15–16	0.7282	0.5633	22–23	0	0
2–3	0.7734	0.5212	6–11	0.9145	0.7439	16–17	0.0184	0	23–24	0	0
2–25	0.8536	0.6998	7–8	0	0	16–19	0	0	25–26	0.3174	0.0591
3–4	0.6536	0.5421	8–9	0	0	16–21	0	0	26–27	0.6110	0.4132
3–18	0.0580	0.0055	9–39	0	0	16–24	0	0	26–28	0.0118	0.0016
4–5	0.2141	0.0453	10–11	0.7022	0.4937	17–18	0.2140	0.0426	26–29	0	0
4–14	0	0	10–13	0	0	17–27	0.6360	0.5397	28–29	0	0
5–6	0	0	13–14	0	0						

Table 13
System UVP, OVP, OLP and TOLP values at noon with and without considering input correlation.

PL	System UVP		System OVP		System OLP		System TOLP	
	WC	BC	WC	BC	WC	BC	WC	BC
Base case	0	0	1	1	0	0.0027	0	0.0002
5%	0	0	1	1	0.0619	0.1707	0.0098	0.0235
10%	0	0	1	1	0.9203	0.9297	0.4049	0.5029
15%	0	0.0010	0.9998	1	0.9995	0.9998	0.9539	0.9660
20%	0.4765	0.7814	0.9351	0.9609	1	1	0.9991	0.9994

Note: WC: without considering input correlations, BC: Base case correlations.

Table 14
System RSOV, RSOL and RSTOL values at noon for various % PLs.

Risk index	PL	WC	BC
RSOV	Base case	4.1905	4.0715
	5%	4.6464	4.7135
	10%	2.8328	2.9559
	15%	2.1462	2.4382
	20%	1.5556	1.8008
RSOL	Base case	0	0.0008
	5%	0.0422	0.1138
	10%	1.7250	1.8098
	15%	6.0273	5.8387
	20%	12.2257	11.5308
RSTOL	Base case	0	0
	5%	0.0029	0.0069
	10%	0.2850	0.3600
	15%	2.1760	2.2598
	20%	7.5092	7.2365

evident from the comparison of results that with an increase in % PL, system UVP, OLP and TOLP values increase whereas system OVP decreases. Although, the values of OVP at majority of the buses has increased with increase in % PL, there is no fixed pattern to describe the change in OVP with increase in % PL at buses 18, 21, 24, 25, 27 and 28.

Finally, the impact of input correlation on the calculation of system risk indices at noon is compared in Table 14 for various % PLs of PV generations. It is found that, in most of the cases, not accounting for input correlations in analysis leads to underestimation of risk indices. The system RSOV increases when % PL increases from base case to 5% but for PL beyond 5%, system RSOV decreases.

The over-limit risk assessment at the remaining time instants follow the similar set of steps as adopted for noon. Tables 3–5 provide %CV information of input RVs at the remaining time instants.

6. Conclusion

This paper has presented a risk-based power system planning with large scale integration of PV generations. Special focus is given on the accurate uncertainty modeling of input RVs at multiple time instants. The effect of PV generation penetrations and change in value of model parameters on the statistics of result variables are analyzed in detail. Further, the over-limit risk indices are calculated for various PV penetrations and input correlations. Based on the analysis of various results, the following sets of conclusions are drawn.

i) The PV generation has a diverse production pattern. The probability distributions at various time instants are different and do not fit to any specific parametric distribution.
ii) The Probability distributions of ambient temperature and aggregate load power at all the time instants are same and follow Gaussian distribution.
iii) The change in the values of TPLF model parameters has a significant effect on statistics of result variables.
iv) A TPLF study considering higher penetration of PV generation leads to higher variability of branch power flows and branch temperatures as compared to other result variables.
v) Hourly TPLF simulations under various PV penetrations and correlations helps in identifying the critical buses and branches by quantifying risk of over-limit in bus voltages and branch power flows which are essential for the assessment of system reinforcement and reliable operation.

As the extension of this research work, it is planned to accommodate both temporal and spatial correlations in multi-time instant power system studies considering generation dispatch strategy.

References

[1] X. Li, X. Zhang, P. Lu, S. Zhang, Transmission line overload risk assessment for power systems with wind and load-power generation correlation, IEEE Trans. Smart Grid 6 (3) (May 2015) 1233–1242.
[2] M. D. Jong, G. Papaefthymiou, and P. Palensky, A framework for incorporation of infeed uncertainty in power system risk-based security assessment, IEEE Trans. Power Syst., http://doi:10.1109/TPWRS.2017.2687983.
[3] B.R. Prusty, D. Jena, A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach, Renew. Sustain. Energy Rev. 69 (Mar. 2017) 1286–1302.
[4] A.M.L. Da Silva, S.M.P. Ribeiro, V.L. Arienti, R.N. Allan, M.B. Do Coutto Filho, Probabilistic load flow techniques applied to power system expansion planning, IEEE Trans. Power Syst. 5 (4) (Nov. 1990) 1047–1053.
[5] P. Zhang, S.T. Lee, Probabilistic load flow computation using the method of combined cumulants and gram-charlier expansion, IEEE Trans. Power Syst. 19 (1) (Feb. 2004) 676–682.
[6] C.L. Su, Probabilistic load-flow computation using point estimate method, IEEE Trans. Power Syst. 20 (4) (Nov. 2005) 1843–1851.
[7] J.M. Morales, J.P. Ruiz, Point estimate schemes to solve the probabilistic power flow, IEEE Trans. Power Syst. 22 (4) (Nov. 2007) 1594–1601.
[8] J. Usaola, Probabilistic load flow with correlated wind power injections, Elect. Power Syst. Res. 80 (5) (May 2010) 528–536.
[9] Y. Yuan, J. Zhou, P. Ju, J. Feuchtwang, Probabilistic load flow computation of a power system containing wind farms using the method of combined cumulants and Gram-Charlier expansion, IET Renew. Power Gener. 5 (6) (Nov. 2011) 448–454.
[10] M.A. Abdullah, A.P. Agalgaonkar, K.M. Muttaqi, Probabilistic load flow incorporating correlation between time-varying electricity demand and renewable power generation, Renew. Energy 55 (July 2013) 532–543.
[11] G. Valverde, A. Saric, V. Terzija, Probabilistic load flow with nonGaussian correlated random variables using Gaussian mixture models, IET Gen. Transm. Distrib. 6 (7) (Jul. 2012) 701–709.
[12] F.J. Ruiz-Rodriguez, J.C. Hernández, F. Jurado, Probabilistic load flow for photovoltaic distributed generation using the Cornish Fisher expansion, Elect. Power Syst. Res. 89 (Aug. 2012) 129–138.
[13] M. Fan, V. Vittal, G.T. Heydt, R. Ayyanar, Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants, IEEE Trans. Power Syst. 27 (4) (Nov. 2012) 2251–2261.
[14] M. Aien, M. Fotuhi-Firuzabad, F. Aminifar, Probabilistic load flow in correlated uncertain environment using unscented transformation, IEEE Trans. Power Syst. 27 (4) (Nov. 2012) 2233–2241.
[15] G. Carpinelli, P. Caramia, P. Varilone, Multi-linear Monte Carlo simulation method for probabilistic load flow of distribution systems with wind and photovoltaic generation systems, Renew. Energy 76 (Apr. 2015) 283–295.
[16] J.M. Sexauer, S. Mohagheghi, Voltage quality assessment in a distribution system with distributed generation—a probabilistic load flow approach, IEEE Trans. Power Del. 28 (3) (Jul. 2013) 1652–1662.
[17] D. Villanueva, A.E. Feijóo, J.L. Pazos, An analytical method to solve the probabilistic load flow considering load demand correlation using the DC load flow, Electr. Power Syst. Res. 110 (May 2014) 1–8.
[18] N. Gupta, V. Pant, B. Das, Probabilistic load flow incorporating generator reactive power limit violations with spline based reconstruction method, Elect. Power Syst. Res. 106 (Jan. 2014) 203–213.
[19] N. Gupta, Probabilistic load flow with detailed wind generator models considering correlated wind generation and correlated loads, Renew. Energy 94 (Aug. 2016) 96–105.
[20] C. Wu, F. Wen, Y. Lou, F. Xin, Probabilistic load flow analysis of photovoltaic generation system with plug-in electric vehicles, Int. J. Elect. Power Energy Syst. 64 (Jan. 2015) 1221–1228.
[21] B.R. Prusty, D. Jena, Combined cumulant and Gaussian mixture approximation for correlated probabilistic load flow studies: a new approach, CSEE J. Power Energy Syst. 2 (2) (Jun. 2016) 71–78.
[22] B.R. Prusty, D. Jena, A sensitivity matrix-based temperature augmented probabilistic load flow study, IEEE Trans. Ind. Appl. 53 (3) (May/Jun. 2017) 2506–2516.
[23] S. Frank, J. Sexauer, S. Mohagheghi, Temperature-dependent power flow, IEEE Trans. Power Syst. 28 (4) (Nov. 2013) 4007–4018.
[24] M. Fan, V. Vittal, G.T. Heydt, R. Ayyanar, Preprocessing uncertain photovoltaic data, IEEE Trans. Sustain. Energy 5 (1) (Jan. 2014) 351–352.

[25] X. Dong, C. Wang, J. Liang, X. Han, F. Zhang, H. Sun, M. Wang, J. Ren, Calculation of power transfer limit considering electro-thermal coupling of overhead transmission line, IEEE Trans. Power Syst. 29 (4) (July 2014) 1503–1511.
[26] Hourly photovoltaic generation data. [Online]. Available: https://www.pvoutput.org/ladder.
[27] J.V. Dave, P. Halpern, H.J. Myers, Computation of incident solar energy, IBM J. Res. Dev. 19 (6) (Nov. 1975) 539–549.
[28] K.E. Holbert, D. Srinivasan, Solar Energy Calculations. Handbook of Renewable Energy Technology, World Scientific Publishing Co, 2011, pp. 189–204, https://doi.org/10.1142/9789814289078_0008.
[29] Hourly temperature data. [Online]. Available: https://www.wunderground.com/history.
[30] Hourly load data. [Online]. Available: http://www.ercot.com/gridinfo/load/load_hist.
[31] M. Fan, Probabilistic Power Flow Studies to Examine the Influence of Photovoltaic Generation on Transmission System Reliability, Ph.D. dissertation, Arizona State University, Department of Electrical, Computer, and Energy Engineering, Tempe AZ, April, 2012.
[32] S. Frank, J. Sexauer, S. Mohagheghi, Temperature Dependent Power Flow in MATLAB, 2013 [Online]. Available, https://github.com/TDPF/TDPF/blob/master/39bus_data.mat.