A Novel Probabilistic Technique for Optimal Allocation of Photovoltaic Based Distributed Generators to Decrease System Losses

Engy A. Mohamed[1*], Yasser G. Hegazy[1], Mahmoud M. Othman[2**]

Abstract

This paper presents a novel algorithm for modeling photovoltaic based distributed generators for the purpose of optimal planning of unbalanced distribution networks. The proposed algorithm utilizes sequential Monte Carlo method in order to accurately consider the stochastic nature of photovoltaic based distributed generators. An efficient algorithm based on Firefly optimization method is proposed for optimal placement of photovoltaic based distributed generators in unbalanced distribution network. The proposed optimization algorithm aims to minimize the annual energy loss by determining the optimal locations of photovoltaic distributed generators. The proposed algorithms are implemented in MATLAB environment and tested on the IEEE 37-node feeder. Several case studies are conducted to prove the effectiveness of the proposed algorithms. The results obtained are presented and discussed.

Keywords

distributed generation, firefly method, optimization, planning

[1] Information Engineering and Technology Faculty,
German University in Cairo

[2] Electrical Power and Machine Department, Faculty of Engineering,
Ain shams University, Cairo, Egypt

[*] First corresponding author, e-mail: engy.adel-mohamed@guc.edu.eg

[**] Second corresponding author, e-mail: mahmoud.othman@eng.asu.edu.eg

1 Introduction

Photovoltaic (PV) based power stations are good choice for replacement of the traditional electrical energy generation as it is infinite and less pollutant source of energy. However, due to its stochastic nature, PV increases the network uncertainties. The PV power is difficult to be accurately simulated because it is strongly correlated to the climate, ambient temperature, season, time and geography [1]. Thus, a probabilistic model of the PV power is needed in order to simulate the actual behavior of these stations.

Models that considers the stochastic nature of the PV power can be classified into two categories; analytical methods [2-6] and Monte Carlo based techniques [7-10]. Authors in [2] Presented a modeling method that based on dividing the solar irradiance into states; finding the average solar irradiance and consequently the most likely power of each hour of the day after consecutive mathematical equations based on the photovoltaic module. In [3] the stochastic nature of the photovoltaic was handled by offering unsymmetrical two point estimation method and it was compared by symmetrical two point estimation method, Gram-Charlier and Latin Hypercube method. The authors in [4] presents a methodology to model PV based power stations for reliability studies by combining Markov Chain and Monte Carlo method for the generation of a multistate PV model based on the transition probability matrix. Reference [6] presents a chronological probability model of photovoltaic (PV) generation on the basis of conditional probability and nonparametric kernel density estimation. Reference [7] described an approach based on Monte Carlo Method to evaluate the uncertainty of the passive parameters of double diode photovoltaic cell using manufacturer's data for the panels, measured environmental parameters and semi empirical equations. The authors in [8] presents a Monte Carlo based strategy for modeling PV power generators considering their dependency with other renewable sources. In [9] a method based on the pseudo-sequential Monte Carlo simulation technique had been proposed to evaluate the reserve deployment and customers' nodal reliability with high PV power penetration. In [10] a Monte Carlo based model which presents

market- based optimal power flow (OPF) with different combination of generation and load demand.

Optimal planning of PV based distributed generators (DG) in distribution a network is an important task in order to gain maximum benefits. Several studies were performed for optimal planning of renewable sources [11-14]. An optimal power flow was solved using non-linear programming in [11] for optimal allocation of renewable distributed generation. Authors in [12] proposed a mixed integer nonlinear programming method for optimal allocation of wind DG unit. A continuous stochastic optimal allocation of wind power considering load uncertainty was presented in [14].

2 Modelling strategy

The flowchart of the proposed algorithm used for modeling PV power is summarized in Fig. 1. The modeling strategy is divided into historical data processing followed by solar irradiance simulation using proper cumulative distribution function, and then the calculation of the simulated PV powers is performed. Finally, Monte Carlo convergence is applied to obtain the most likelihood values of PV powers at each hour. The proposed strategy is discussed in the following subsections.
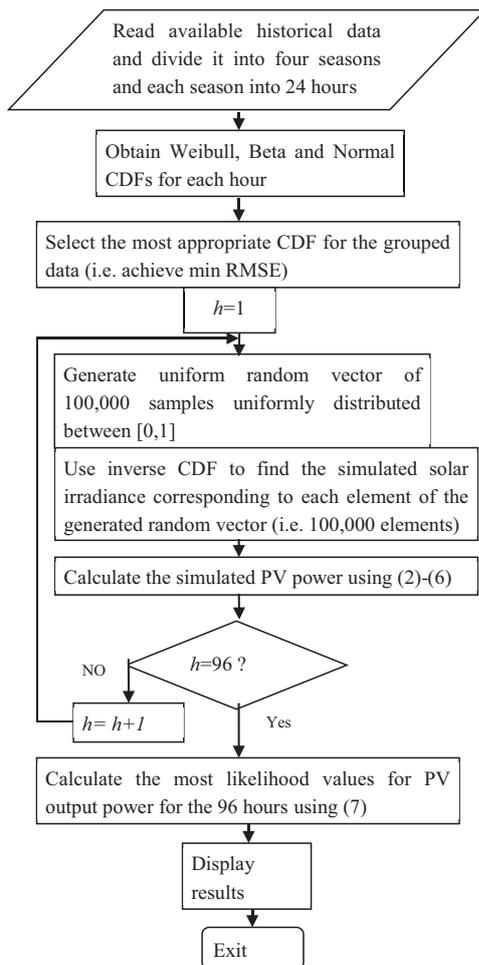

Fig. 1 Modelling strategy

2.1 Historical Data Processing

Three years of historical data between the years 2001-2003 at different locations of the same site are used in this study. Six readings of the solar irradiance and ambient temperature were taken at each hour during the three years. The available data is seasonally divided (i.e. each season data is separated). The data representing each season is further subdivided into 24-h segments (time segments), each referring to a particular hourly interval for the entire season. Thus, there are 96 time segments for the year (24 for each season). Considering a month to be 30 days, each time segment then has 1620 irradiance (3 years × 30 days per month × 3 month per season × 6 readings per hour).

2.2 Solar irradiance modelling

Depending on the 1620 solar irradiance collected to represent solar irradiance in each hour, different cumulative distribution are tested to evaluate the most appropriate cumulative distribution function (CDF) to fit the random phenomenon of the irradiance data. Three types of the most famous probability functions (Weibull, Beta, and Normal) are constructed based on the given data (each hour of the four seasons have its own CDF). A comparison based on the percentage error of solar irradiance of each cumulative distribution function to the actual data is calculated, The root-mean-square error (RMSE) between the actual CDF and simulated CDF is used for estimation of the most appropriate simulated CDF to the actual data. RMSE is calculated for each type of CDF using (1)

$$\text{RMSE} = \sqrt{\frac{1}{Nh} \sum_{h=1}^{Nh} \sum_{j=1}^{Nd} \left(CDF_{act}^{h,j} - CDF_{sim}^{h,j} \right)^2} \qquad (1)$$

Where $CDF_{act}^{h,j}$, $CDF_{sim}^{h,j}$ are the value of actual CDF and simulated CDF respectively at hour h and data j, Nd is the total number of data in each hour (i.e. 1620 data) and Nh is the total number of hours (i.e. 96 hours).

The RMSE is calculated for all random variables (i.e. solar irradiance) for each type of CDF (i.e. Beta or Weibull or Gaussian). The type of CDF that achieves the minimum RMSE is selected to simulate the solar irradiance.

2.3 Simulated random variables of solar irradiance

A uniformly distributed random number vector of 100,000 values bounded between 0 and 1 is generated. At each random number the corresponding solar irradiance random variable is obtained from the CDF (i.e. the inverse CDF is obtained at each random number). For each of the obtained 96 CDF, a vector of 100,000 random variables is calculated representing the simulated solar irradiance at each hour at each season. (i.e. 96 simulated random variables vectors are calculated each contain 100,000 random variables).

2.4 Calculation of the simulated PV power

The output power of the PV array is dependent on the solar irradiance and ambient temperature of the site as well as the characteristics of the module itself. At each value of the calculated random variables the corresponding PV power is calculated using Eqs. (2)-(6)

$$T_c = T_a + S_a \left(\frac{N_{OT} - 20}{0.8} \right) \tag{2}$$

$$I = S_a \left[I_{sc} + K_i \left(T_c - 25 \right) \right] \tag{3}$$

$$V = V_{oc} - K_v \times T_c \tag{4}$$

$$FF = \frac{V_{MPP} \times I_{MPP}}{V_{oc} \times I_{sc}} \tag{5}$$

$$P_s = N \times FF \times V \times I \tag{6}$$

Where: T_c is cell temperature °C, T_a is average hourly ambient temperature °C, S_a is simulated solar irradiance kW/m², N_{OT} is nominal operating temperature of cell °C, I is module current (A), I_{sc} is short circuit current (A), K_i is current temperature coefficient A/°C, V is module voltage (V), V_{oc} is open-circuit voltage (V), K_v is voltage temperature coefficient (V/°C), FF is fill factor, V_{MPP} is voltage at maximum power point (v), I_{MPP} is current at maximum power point (A), P_s is simulated output power of the PV module, and N is the number of modules per array.

2.5 Monte Carlo simulations convergence

The most likelihood value of the obtained 100,000 random powers at each hour is achieved by running Monte Carlo simulation. The dynamic average is calculated using Monte Carlo convergence Eq. (7). For the very high number of simulations (i.e. 100,000) the Monte Carlo convergence is guaranteed.

$$P_{ave} = \frac{1}{N_s} \sum_{j=1}^{NS} P_s(j) \tag{7}$$

Where P_{ave} is the most likelihood value of the PV power calculated at each hour at each season, P_s is the PV power random variable and NS is the total number of simulations (100,000).

3 Problem Statement

The optimization problem under study can be stated as follows:

3.1 Objective Function

The optimization problem under study aims to determine exactly the optimal PV based distributed generators locations for the sake of minimizing the distribution feeder annual energy loss. Since each time segment represents 90 h (30 days per month 3 months per season), the objective function can be described as follows using (8)

$$\text{Minimize the annual energy losses} = \sum_{h=1}^{96} P_{loss,h} \times 90 \tag{8}$$

Where $P_{loss,h}$ is the power loss at each hour h for the 96 hours under study (4 seasons × 24 hours per season)

3.1.1 Technical Constraints

• Voltage limits: voltage at each bus should be within a permissible range usually:

$$0.95 \, p.u. \leq V \leq 1.05 \, p.u. \tag{9}$$

• Lines thermal limit (line Ampacity): it represents the maximum current that the line can withstand at certain DG penetration, exceeding this value leads to melting of the line.

$$I_{flow} \leq I_{Thermal} \tag{10}$$

• Substation limit: this constraint represents the maximum apparent power that the substation can provide.

$$S_{substation, flow} \leq S_{substation, max} \tag{11}$$

• Power balance: the sum of input power should be equal to the sum of output active power in addition to the active power loss. The input power may include the DG active power and the active power supplied by the utility. The active output power is the sum of loads active power.

$$P_{substaion} + \sum P_{DG} = \sum P_{loads} + P_{loss} \tag{12}$$

3.1.2 Assumptions

To formulate an accurate planning strategy that determines the optimal location of PV based DG units, the following assumptions are made.

• All the renewable DG units are working at a unity power factor.
• All buses in the system under study are subjected to the same meteorological conditions.
• The energy can be supplied to the substation if the sum of the renewable DGs power exceeds the system demand.

4 Methodology

The firefly optimization algorithm is inspired from the natural behavior of the fireflies; a firefly of the maximum brightness has the largest ability to attract other fireflies. The brightness of a firefly is affected or determined by the landscape of the objective function. For a maximization problem, the brightness can simply be proportional to the value of the objective function [15], whilst for a minimization problem; the brightness is inversely proportional to the value of the objective function.

The optimization algorithm is summarized in Fig. 2 and described in the following steps:

Step (1): Generate a set of random fireflies bounded inside certain preset region. Each firefly has one dimension represents the PV DG location.

Step (2): Calculate the annual energy losses corresponding to all initial DG locations and the seasonal hourly power schedule of the PV DG by performing unbalanced load flow for the 96 hours.

Step (3): Evaluate the brightness of each firefly based on the type of the objective function (i.e. the brightest firefly is that achieve minimum energy loss).

Step (4): Calculate the distances between all fireflies and the brightest one, the distance between two fireflies i, j (r_{ij}) for the current optimization problem can be calculated as follow:

$$r_{ij} = \sqrt{\left(Loc_i - Loc_j\right)^2} \quad (13)$$

Where loc_i and loc_j are the DG location of fireflies i, j respectively.

Step (5): Move the less attraction (brightness) firefly i towards the more attractive firefly j by using the following equation:

$$Loc_i = Loc_i + \beta_0 \exp\left(-\gamma r_{ij}^2\right) \times \left(Loc_j - Loc_i\right) + \alpha\left(rand - \frac{1}{2}\right) \quad (14)$$

Where the first term is the current firefly position, the second term is used to update the firefly position based on the brightness of the fireflies and the third term is used to randomize the movement of firefly. β0 is the initial attractiveness, γ is the absorption coefficient, the values of these parameters are accurately tuned to solve the optimization problem efficiently. α is a randomization parameter that decrease at each iteration by Eq. (15) and rand is a random number generator uniformly distributed between [0,1].

$$\alpha^{k+1} = \alpha^k\left(\frac{1}{2k_{max}}\right)^{1/k_{max}} \quad (15)$$

Step (6): Evaluate the brightness of the fireflies at the updated positions by calculating the objective function.

Step (7): Repeat steps (3)-(5) until reach the maximum number of iterations (k_{max}).

5 Test cases and results
5.1 Model of PV based DGs

The modeling procedure presented in section II is applied to the 3 years solar irradiance data. The continuous Beta, Weibull and Gaussian CDFs functions are compared to the discrete actual CDF obtained using numerical integration of the real data PDF; different CDFs at different seasons and different selected hours are compared to the actual CDFs and the RMSE is calculated for each season and presented in Table 1. It is obvious from the table that the Beta CDF is the most appropriate CDF to simulate the random behavior of the solar irradiance. Figure 2 shows a sample of graphical comparison between the actual data, Beta, Weibull, and Normal distributions, it can be clearly concluded from the shown figure that Beta distribution is the most fitting distribution to the actual data which emphasizes the results obtained from the RMSE calculations.

The advantage of the beta distribution over the actual data CDF is that beta distribution is invertible, so that, the simulated solar irradiance could be obtained correspondingly to any random number uniformly distributed between [0, 1].

Table 1 RMSE for different seasons

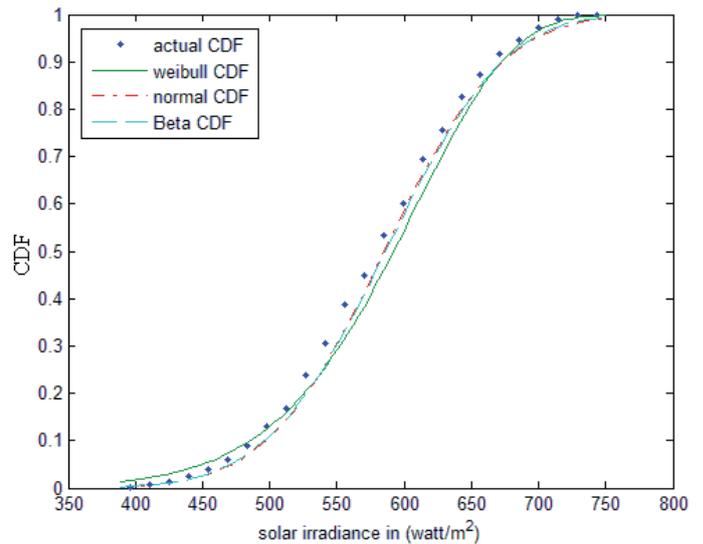| Season	Summer	Winter	Fall	Spring
Beta	0.033364	0.092075	0.075747	0.056225
Weibull	0.055025	0.194151	0.196547	0.146322
Gaussian	0.058844	0.193642	0.194813	0.148391

Fig. 2 Comparison of simulated and actual CDF at 09:00 A.M., summer
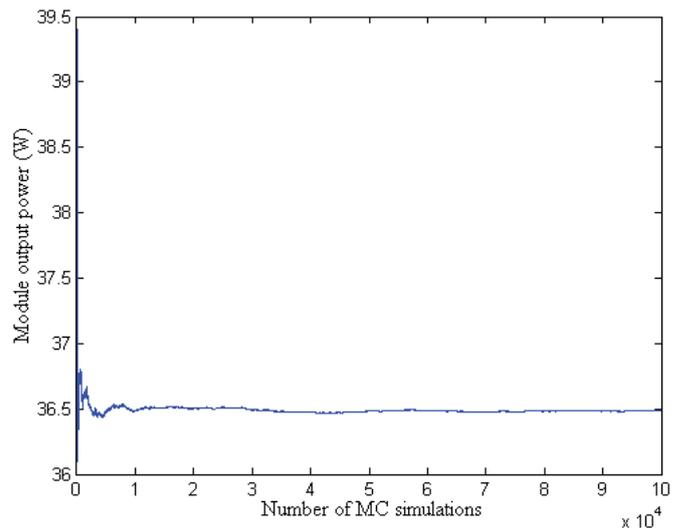

Fig. 3 MC simulations convergence at 12:00 A.M, summer season

The hourly 100,000 Monte Carlo simulations for the random PV power are then obtained using the aforementioned uniformly random number generator and the inverse Beta distributions. Table 2 presents the values of the constants and

parameters required to calculate the PV output power. The converged Monte Carlo results of the PV power of one module for the 96 hours are presented in Table 3. In order to validate the proposed modeling strategy, the converged values of Monte Carlo simulations at all hours are compared to the average values of the PV powers obtained using the actual solar irradiance. Figure 3 shows the converged value of Monte Carlo simulations at a sample hour and Fig. 4 shows the average power obtained using actual solar irradiances at the same hour. The comparison shows that the converged value of Monte Carlo simulations is close to the average obtained using the actual data with the preference of Monte Carlo simulations as the huge number of simulations used guarantees the convergence (i.e. reaching the most likelihood value) unlike the small number of actual data which emphasizes the importance of Monte Carlo method.

Table 2 Characteristics of the PV module

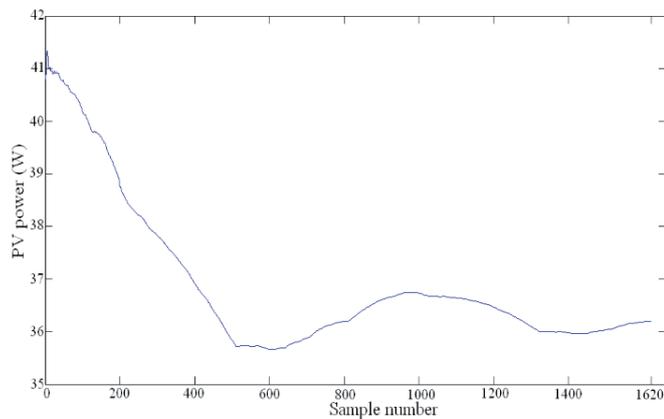| Module characteristics	Features	Module characteristics	Features
Watt peak (W)	53	Current at maximum power (A)	3.05 A
Open circuit voltage (V)	21.7	Voltage temperature Coefficient (mV/°C)	88
Short circuit current (A)	3.4	Current temperature Coefficient (A/°C)	1.5
Voltage at maximum power (V)	17.4	Nominal cell operating temperature (°C)	43

Fig. 4 Average of the actual data at 12:00 am, summer season

5.2 Optimal Planning of PV based DGs for Energy Losses Minimization

The proposed algorithm is implemented in MATLAB environment and tested on the IEEE 37-node feeder presented in Fig. 6 to minimize the distribution system annual energy losses through the optimal allocation of the photovoltaic DG.

Six test cases are made; two test cases for the base cases where no DG is connected with different load profiles, four test cases for optimal allocation of PV based DG with different number of PV modules and different load profiles. The test cases are summarized as follow:

- *Test case # 1:* the base case where no DG is connected with peak load profile.
- *Test case # 2:* no DG is connected with variable load profile (i.e. peak load scaled by load scaling factor).
- *Test case # 3:* only one photovoltaic DG of capacity 795kW capacity (15,000 modules) with peak load profile.
- *Test case # 4:* only one photovoltaic DG of capacity 795kW capacity (15,000 modules) with variable load profile.
- *Test case #5:* only one photovoltaic DG of 1590kW capacity (30000 modules) with peak load profile.
- *Test case #6:* only one photovoltaic DG of 1590kW capacity (30000 modules) with variable load profile.

Table 3 PV module converged simulated power

| Hours	Seasons PV output power (W)			
	Winter	Spring	Summer	Fall
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0.0051	2.0309	1.5457	0.0150
7	1.0102	10.3993	9.1009	3.4395
8	7.7385	19.2869	17.7126	10.7027
9	14.8381	26.9237	25.1966	17.2147
10	20.3485	32.4399	30.6294	21.929
11	24.7366	35.6412	34.2144	24.3809
12	26.3658	37.3814	36.509	25.1659
13	25.1997	37.3374	35.9115	24.1829
14	23.1347	33.9647	32.7959	21.012
15	19.2044	28.3968	27.7954	15.2652
16	12.3300	20.5054	20.4046	8.0518
17	4.6891	11.0148	11.2052	1.7354
18	0.3344	2.9472	3.2555	0.0346
19	0	0.0807	0.0901	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	0	0	0	0
24	0	0	0	0
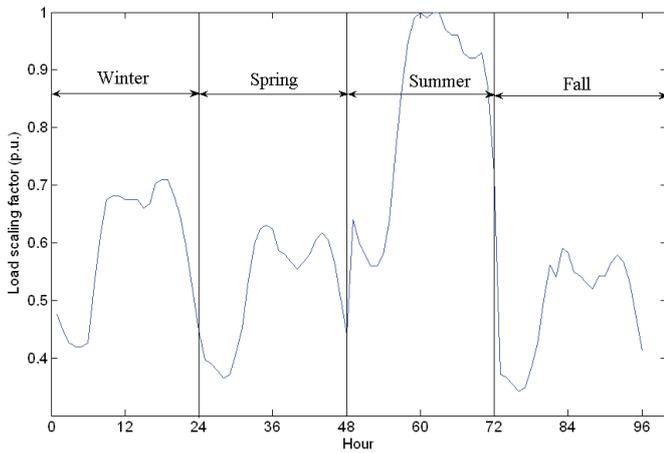
Fig. 5 Load scaling factor [2]
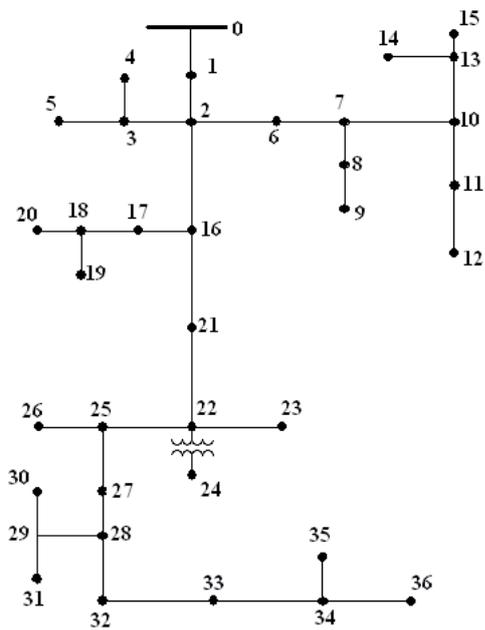
Fig. 6 Renumbered IEEE 37- node feeder

1) Test cases # 1 and # 2

The annual energy losses for the base case where no DG is connected is obtained by running the unbalanced load flow for the 96 hours and calculating the power loss for each hour, then the annual energy loss is calculated using (8). The load profile for the second test case is presented in Fig. 6. Table 4 summarizes the results of the two test cases.

2) Test cases # 3, #4, # 5 and # 6

In this sub section, the optimal locations of PV based DG with 15000 modules (i.e. rated power of 795 KW) with two different loading profiles (peak load profile and the aforementioned load scaling factor) are determined. Also, the same cases are repeated with PV based DG with 30000 modules (i.e. rated power of 1590 KW). The optimal location of the PV DG and the corresponding annual energy losses for the four test cases are presented in Table 5 and the annual energy losses when PV DG is connected at the optimal location for the four test cases

are presented in Fig. 7. It can be concluded that a significant reduction of power losses occurs when PV DG is connected to the distribution networks at the optimal location. It is clear from the figures that no reduction occurs at the night time as the PV DG output power is null.

Table 4 Results for test cases #1 and #2

Annual energy loss, test case #1	540.967 MWh
Annual energy loss, test case #2	229.1088 MWh

Table 5 Results for test cases # 3-6

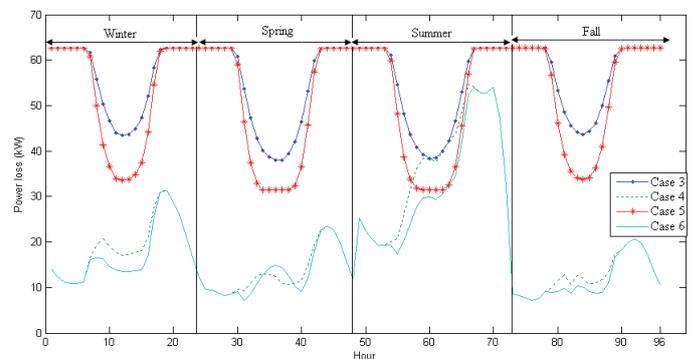|	Test case #3	Test case #4	Test case #5	Test case #6
Optimal location	33	32	32	28
Annual energy losses (MWh)	477.990	175.491	445.131	161.721

Fig. 7 Power losses for test cases 3-6

6 Conclusions

A novel algorithm for modelling the PV based DGs considering their stochastic nature is presented in this paper. The proposed algorithm is based on Monte Carlo method and used to determine a probabilistic hourly/seasonal model for PV DG. Moreover, a Firefly based algorithm is proposed for accurate determination of the optimal locations of the PV based DGs connected to unbalanced distribution network in order to minimize the annual energy losses. The main contributions of the proposed algorithms are that the modelling algorithm considers the stochastic nature of the PV DGs. Thus, the proposed models accurately simulate their behaviour, the rationale behind the proposed model is to include the probabilistic model into a deterministic optimization problem which simplifies the solution of the optimization problem and the firefly algorithm precisely determines the optimal allocation of the PV DG without violating the system constraints.

References

[1] Ruiz-Rodriguez , F. J., Hernandez, J. C., Jurado, F. "Probabilistic load flow for radial distribution networks with photovoltaic generators." *IET Renewable Power Generation*. 6(2), pp. 110–121. 2012. DOI: 10.1049/iet-rpg.2010.0180

[2] Atwa, Y. M., El-Saadany, E. F., Salama, M. M. A., Seethapathy, R. èOptimal Renewable Resources Mix for Distribution System Energy Loss Minimization." *IEEE Transactions of Power Systems*. 25(1), pp. 360-370. 2010. DOI: 10.1109/TPWRS.2009.2030276

[3] Soroudi, A., Aien, M., Ehsan, M. "A Probabilistic Modeling of Photovoltaic Modules and Wind Power Generation Impact On Distribution Network" *IEEE System Journal*. 6(2), pp. 254-259. 2012. DOI: 10.1109/JSYST.2011.2162994

[4] Sayed, R., Hegazy, J. G., Mostafa, M. A. "Modeling of photovoltaic based power stations for reliability studies using Markov Chains." In: 2013 International Conference on Renewable Energy Research and Applications, Madrid, Spain, Oct. 20-23, 2013, pp. 667-673. DOI: 10.1109/ICRERA.2013.6749838

[5] Eason, G., Noble, B., Sneddon, I. N. "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions." *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*. 247(935), pp. 529-551. 1955. DOI: 10.1098/rsta.1955.0005

[6] Ren, Z., Yan, W., Zhao, X., Li, W., Yu, J. "Chronological probability model of photovoltaic generation." *IEEE Transactions on Power Systems*. 29(3), pp. 1077-1088. 2014. DOI: 10.1109/TPWRS.2013.2293173

[7] Attivissimo, F., Di Nisio, A., Savino, M., Spadavecchia, M. "Uncertainaity Analysis in Photovoltaic Cell Parameter Estimation." *IEEE Transactions on Instrumentation and Measurement*. 61(5), pp.1334-1342. 2012. DOI: 10.1109/TIM.2012.2183429

[8] Abdelaziz, A. Y., Hegazy, Y. G., El-Khattam, W., Othman, M. M. "Optimal allocation of stochastically dependent renewable energy based distributed generators in unbalanced distribution networks." *Electric Power System Research*. 119, pp. 34-44, 2015. DOI: 10.1016/j.epsr.2014.09.005

[9] Zhao, Q., Wang, P., Goel, L., Ding, Y. "Evaluation of nodal reliability risk ina deregulated power system with photovoltaic power penetration." *IET Generation, Transmission & Distribution*. 8(3)., pp.421-430. 2014. DOI: 10.1049/iet-gtd.2013.0340

[10] Siano, P., Mokryani, G. "Probabilistic Assessment of Impact of Wind Energy Integration Into Distribution Networks." *IEEE Transactions on Power Systems*. 28(4), pp. 4209-4217. 2013. DOI: 10.1109/TPWRS.2013.2270378

[11] Ochoa, L. F., Harrison, G. P. "Minimizing energy losses: Optimal accommodation and smart operation of renewable distributed generation." *IEEE Transactions on Power Systems*. 26(1), pp. 198-205. 2011. DOI: 10.1109/TPWRS.2010.2049036

[12] Atwa, Y. M., El Saadany, E. F. "Probabilistic approach for optimal allocation of wind-based distribution generation in distribution systems." *IET Renewable Power Generation*. 5(1), pp. 79-88. 2009. DOI: 10.1049/iet-rpg.2009.0011

[13] Wang, C., Nahrir, M. H. "Analytical approaches for optimal placement of distributed generation sources in power systems." *IEEE Transactions on Power Systems*. 19(4), pp. 2068-2076. 2004. DOI: 10.1109/TPWRS.2004.836189

[14] Novoa, C., Jin, T. "Reliability centered planning for distributed generation considering wind power volatility." *Electric Power Systems*. 81(8), pp. 1654-1661. 2011. DOI: 10.1016/j.epsr.2011.04.004

[15] Yang, X.-S. "Firefly algorithms for multimodal optimization." In: *Stochastic Algorithms: Foundation and Applications*. Vol. 5792, pp. 169–178, 2009. DOI: 10.1007/978-3-642-04944-6_14