World Academy of Science, Engineering and Technology
International Journal of Energy and Power Engineering
Vol:9, No:8, 2015

A Novel Probablistic Strategy for Modeling Photovoltaic Based Distributed Generators

Engy A. Mohamed, Yasser G. Hegazy

***Abstract*—**This paper presents a novel algorithm for modeling photovoltaic based distributed generators for the purpose of optimal planning of distribution networks. The proposed algorithm utilizes sequential Monte Carlo method in order to accurately consider the stochastic nature of photovoltaic based distributed generators. The proposed algorithm is implemented in MATLAB environment and the results obtained are presented and discussed.

***Keywords*—**Comulative distribution function, distributed generation, Monte Carlo.

I. 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 consider 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.

Engy A. Mohamed is with the Information Engineering and Technology Faculty, German University in Cairo, Cairo, Egypt (e-mail: Engy.adel-mohamed@guc.edu.eg).

Yasser G. Hegazy is Dean of the Faculty of Information Engineering and Technology, German University in Cairo, Cairo, Egypt (e-mail: Yasser.higazi@guc.edu.eg).

This paper presents a Monte Carlo based modeling algorithm for modeling PV power for the sake of optimal planning of PV based DG. The proposed algorithm is implemented in MATLAB environment and the results and the consequent discussions prove the effectiveness of the proposed algorithm.

II. Modeling 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.

A. 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 $\times$ 30 days per month $\times$ 3 month per season $\times$ 6 readings per hour).

B. Solar Irradiance Modeling

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)

$$\mathrm{RMSE} = \sqrt{\frac{1}{Nh}\sum_{h=1}^{Nh}\sum_{j=1}^{Nd}(CDF_{act}^{h,j} - CDF_{sim}^{h,j})^2} \quad (1)$$

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.

C. 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).

D. 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 (2)-(6):

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

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

$$V = V_{oc} - K_v \times T_c \quad (4)$$

$$FF = \frac{V_{MPP} \times I_{MPP}}{V_{oc} \times I_{sc}} \quad (5)$$

$$P_s = N \times FF \times V \times I \quad (6)$$

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.

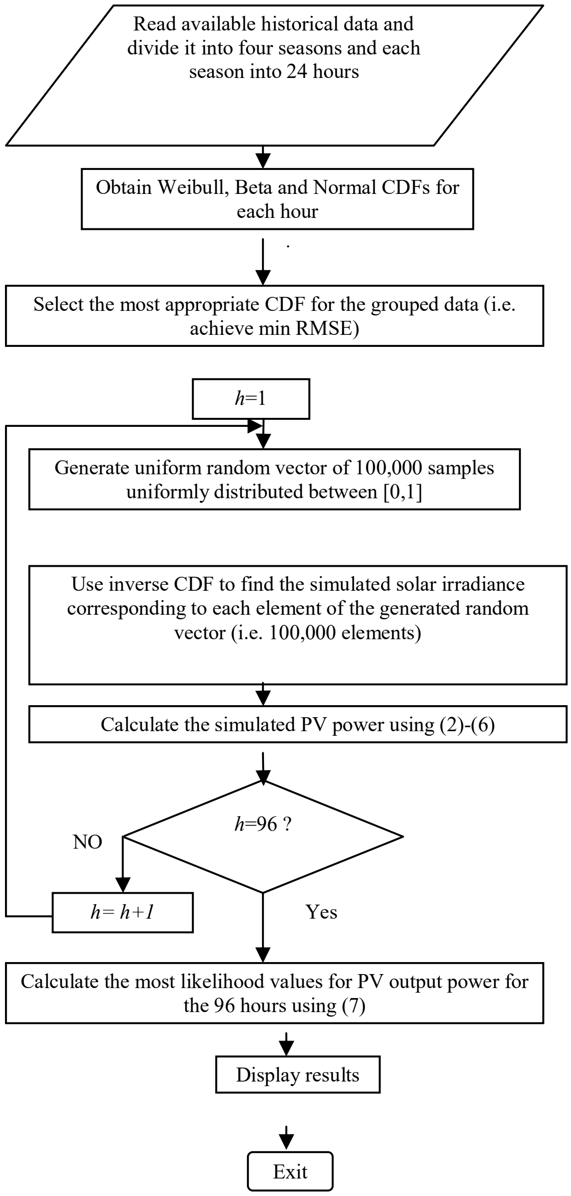

Fig.1 Modeling strategy

E. 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 equation (7). For the very high number of simulations (i.e. 100,000) the Monte Carlo convergence is guaranteed.

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

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).

TABLE I
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

TABLE II
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

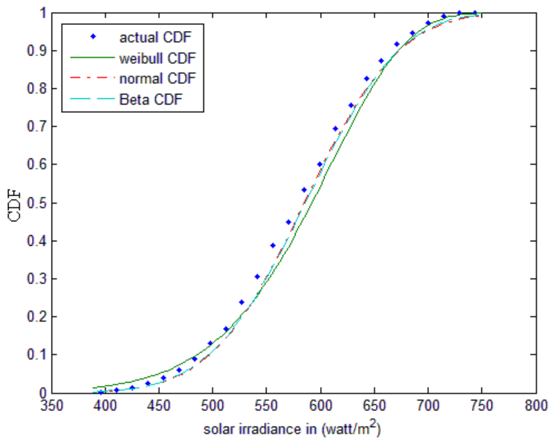

Fig. 2 Comparison of simulated and actual CDF at 09:00 A.M., summer

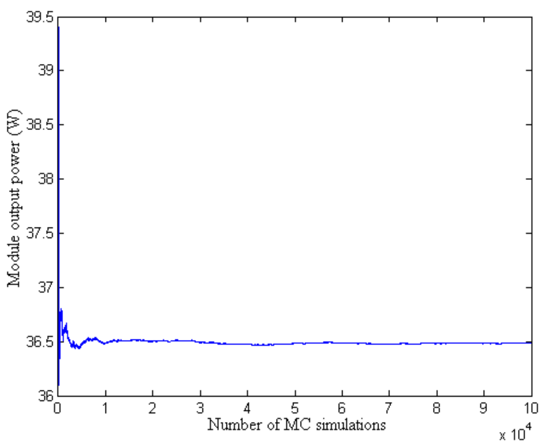

Fig. 3 MC simulations convergence at12:00 A.M, summer season

III. TEST CASES AND RESULTS

A. 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 I. It is obvious from the table that the Beta CDF is the most appropriate CDF to simulate the random behavior of the solar irradiance. Fig. 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].

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 II 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 III.

TABLE III
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

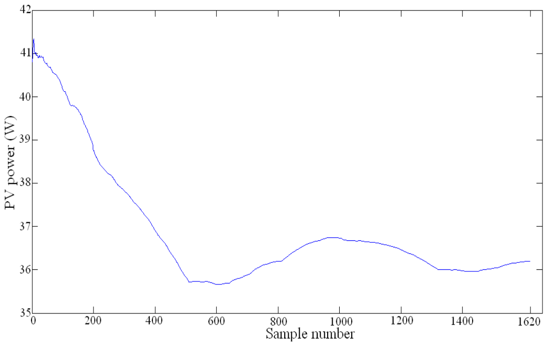

Fig. 4 Average of the actual data at 12:00 am, summer season

B. Validation of the Proposed Model

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. Fig. 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.

IV. CONCLUSION

A novel algorithm for modeling 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 validation of the proposed model is done via comparison of the converged Monte Carlo simulations and the converged actual data. The main contributions of the proposed algorithm are that the modeling algorithm considers the stochastic nature of the PV DGs. Thus, the proposed models accurately simulate their behavior; 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.

REFERENCES

[1] F.J. Ruiz-Rodriguez, J.C. Hernandez and F. Jurado "Probabilistic load flow for radial distribution networks with photovoltaic generators" *IET Renew. Power Gener.*, Vol. 6, no. 2, pp. 110–121, 2012.

[2] Y.M. Atwa, E.F El saadany, M.A. salama and R. seethapathy, "Optimal Renewable Resources Mix for Distribution System Energy Loss Minimization" *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 360- 370, Feb. 2010.

[3] A. Soroudi, M. Aien, and M. Ehsan "A Probabilistic Modeling of Photovoltaic Modules and Wind Power Generation Impact On Distribution Network", *IEEE system journal*, vol. 6, no. 2, pp. 254-259, JUN. 2012.

[4] Radwa Sayed, Y. G. Hegazy, and M. A. Mostafa "Modeling of Photovoltaic Based Power Stations for Reliability Studies Using Markov Chains" International Conference on Renewable Energy Research and Applications, Madrid, Spain, 20-23 October 2013.

[5] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[6] Z. Ren, W. Yan, , X. Zhao, , W. Li, and Juan Yu, "Chronological Probability Model of Photovoltaic Generation" *IEEE Trans. Power Syst.*, vol. 29,no.3, pp. 1077-1088, May. 2014.

[7] F. Attivissimo, A. Di Nisio, M. Savino, and M. Spadavecchia "Uncertainaity Analysis in Photovoltaic Cell Parameter Estimation" *IEEE Trans. On Instr. And measure.*, vol. 61, no. 5, pp.1334-1342, May. 2012.

[8] A. Y. Abdelaziz, Y. G. Hegazy, W. El-Khattam and M. M. Othman, 'Optimal Allocation of Stochastically Dependent Renewable Energy Based Distributed Generators in Unbalanced Distribution Networks', Elec. Power. Syst. Res, Vol. 119, pp. 34-44, Feb. 2015.

[9] Qian Zhao, Peng Wang, Lalit Goel, Yi Ding, Evaluation of nodal reliability risk ina deregulated power system with photovoltaic power penetration, IET Gener.Transm. Distrib. Vol. 3, pp.421–430, 2014.

[10] Pierluigi Siano, and Geev Mokryani, "Probabilistic Assessment of Impact of Wind Energy Integration Into Distribution Network" , *IEEE Trans. Power Syst.* , vol. 28,no. 4, pp. 4209-4217, Nov. 2013.