Probabilistic Load Flow for Distribution Networks with Photovoltaic Generators Part 1: Theoretical Concepts and Models

S. Conti, and S. Raiti
D.I.E.E.S. - Università degli Studi di Catania, Viale A. Doria, 6, 95125 Catania, (*Italy*)

***Abstract*--Connections of distributed generation (DG) systems to distribution systems are increasing in number, though they may often be associated with the need of costly grid reinforcements or new control issues to maintain optimal operating conditions. Appropriate analysis tools are required to check distribution networks operating conditions in the evolving scenario. Load flow (LF) calculations are typically needed to assess the allowed DG penetration level for a given network in order to ensure that voltage and current limits are not exceeded. The present paper deals with the solution of the LF problem in distribution networks with photovoltaic (PV) DG. Suitable models for the active power produced by PV DG units and the power absorbed by the loads are to be used for representing the uncertainty for solar energy availability and loads variation. The proposed models have been incorporated in a radial distribution Probabilistic Load Flow (PLF) program that has been developed by using Monte Carlo techniques. In this paper the theoretical concepts and the software implementation are described. The developed program allows probabilistic predictions of power flows at the various sections of distribution feeders and voltage profiles at all nodes of a network. In the second part of the paper the application to a practical case study is presented to show the results of the proposed method.**

Index Terms--**Power Distribution, Distributed Generation, Photovoltaic Generators, Probabilistic Load Flow Analysis.**

I. INTRODUCTION

The European Renewable Energy Study (TERES), commissioned by the European Union (EU) to examine the feasibility of EU CO_2-reduction goals and the EU renewable energy targets, found that around 60% of the renewable energy potential that can be utilized until 2010 can be categorized as decentralized power sources in terms of distributed generation (DG) [1].

DG can be defined as electric power generation within distribution networks or on the customer side of the network. Network integration of DG is a very complex issue that is significantly different from traditional network integration of power generation into transmission networks. Actual distribution networks are designed as radially operated, passive systems. For this reason connection of DG to the distribution system may be associated with costly grid reinforcements or new control issues to maintain optimal operating conditions (e.g. acceptable voltage quality, correct protections operation, etc.).

With regard to LV distribution networks, an example of grid-connected renewable energy sources is provided in many European Countries by the implementation of the "Photovoltaic (PV) Roof-Tops" programs, supported by national governments.

At the LV level, one of the most relevant issues is the quality of the voltage supplied to customers, according to the European Standard EN 50160 [2], especially in terms of possible overvoltage caused by connection of generators [3], [4].

Appropriate Load-Flow (LF) calculations are needed to assess the allowed DG penetration level for a given network in order to ensure that the maximum voltage at the point of common coupling (PCC) and lines current carrying capacity are not exceeded [5].

However, when connection of generators based on renewable energy (such as solar energy) is considered, it is not possible to achieve a realistic evaluation of where and when overvoltages may occur by simply using traditional deterministic LF (DLF) analysis.

In fact, this analysis (recalled in Section 2) is normally based on some selected combinations of consumer loads and PV power productions. Consequently, it does not take into account the statistical variation of loads and solar radiation.

Then, to solve this problem, the paper presents a Probabilistic Load Flow (PLF), based on Monte Carlo techniques, for radial distribution networks with PV DG (Section 3).

The procedure incorporates suitable models for the active power produced by PV DG units and the power absorbed by the loads in order to represent the uncertainty for solar energy availability and load variation.

In Section 4 the development of a software tool to implement the method proposed will be presented, in order to assess voltages and currents in a distribution network calculated hour by hour throughout a year.

In the second part of this paper [6] the application to a practical case study will be presented to show the results of the proposed method.

II. DETERMINISTIC LOAD-FLOW (DLF)

Let us consider a three-phase symmetrical, radial distribution network with n nodes and n branches, where we define as "nodes" the points of load connections, the points of change in the line characteristics and the junctions, and as "branches" the conductor segments between two nodes.

The nodes can be numbered according to the following rule: the "origin" of the network (typically a MV/LV secondary substation) takes the number *0* and it is not considered in the subsequent calculation as we assume that the voltage at this node, $\overline{V_0}$, is known exactly. The other nodes are numbered sequentially imposing that a "receiving" node takes a number higher than the "sending" node nearer to it. The terms "receiving" and "sending" are used under the assumption that in a traditional radial network, i.e. without distributed generators, the power flow is directed from a lower to a higher number. The branches are identified by the same number as their receiving node, as shown in Fig. 1.

This numbering method allows a simple storage of the network structure in a single square matrix (called incidence matrix, *[A]*) whose dimension is *(nxn)*. In particular, the rows corresponds to the *n* branches and the columns to the nodes. The elements of *[A]* describe the network topology and are equal to *1* if the node corresponding to column *j* is fed through the branch corresponding to row *i*; *0* otherwise.

The calculation of the branch flows is easily obtained applying the "mesh method" for network analysis. It can be easily shown that:

$$\left[\overline{J}\right] = [A]\left[\overline{I}\right] \tag{1}$$

where

$\left[\overline{I}\right]$ is the vector of the load currents, dimension (*nx1*);

$\left[\overline{J}\right]$ is the vector of the branch currents, dimension (*nx1*).

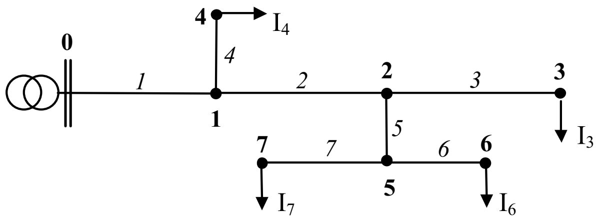

Fig. 1. One-line diagram of a three-phase symmetrical, radial distribution network.

The network complex impedance is equal to:

$$\left[\overline{Z}\right] = [A]^t \cdot \left[\overline{Z_b}\right] \cdot [A] \tag{2}$$

where $\left[\overline{Z_b}\right]$ is the diagonal matrix (*nxn*) whose elements are the complex impedances of the corresponding branches.

For the *i-th* node the complex power $\overline{S}_i$ is defined as:

$$\overline{S}_i = \overline{V}_i \overline{I}_i^* = P_i + jQ_i \tag{3}$$

where

$\overline{V}_i$ is the voltage phasor at node *i*;

$\overline{I}_i$ is the current phasor at node *i*;

$\overline{I}_i^*$ is the complex conjugate of $\overline{I}_i$;

P_i is the net real power in node *i*;

Q_i is the net reactive power in node *i*.

The real and reactive powers P_i and Q_i can be regarded as the difference between the powers absorbed by the load (P_{Li}, Q_{Li}) and the powers delivered by the generators (P_{Gi}, Q_{Gi}) connected at node *i*:

$$\begin{aligned} P_i &= P_{Li} - P_{Gi} \\ Q_i &= Q_{Li} - Q_{Gi} \end{aligned} \tag{4}$$

Then, the voltage at the i-th node can be evaluated by means of the following expression:

$$\overline{V_i} = \overline{V_o} - \sum_{j=1}^{n} \overline{Z_{ij}} \cdot \frac{(P_{Lj} - P_{Gj}) - j(Q_{Lj} - Q_{Gj})}{\overline{V_j}^*} \tag{5}$$

Assuming phasor $\overline{V_o}$ on the real axis, neglecting the phase difference between the node voltages (which is commonly done in MV and LV distribution networks analysis) and considering the real components of expression (5) we obtain:

$$\overline{V_i} = \overline{V_o} - \sum_{j=1}^{n} \left\{ \frac{1}{V_j} \left[R_{ij}(P_{Lj} - P_{Gj}) + X_{ij}(Q_{Lj} - Q_{Gj}) \right] \right\} \tag{6}$$

for *i=1...n*, where R_{ij} and X_{ij} are respectively the real and imaginary part of $\overline{Z}_{ij}$, generic element of matrix $\left[\overline{Z}\right]$.

A possible use of these equations is calculation of the allowed PV DG penetration level for a given network in order to ensure that the voltage limits at the point of common coupling (PCC) and the lines current carrying capacity are not exceeded [5].

To do this the values of loads and generations at each node, e.g., for each hour of the day, must be known and provided as inputs to the solving program. Typically, the available inputs are the loads and generations *mean* values (or *expected* values) for each hour, i.e. the possible stochastic variations that can be introduced in the absorbed and produced powers by customers' behaviour and meteorological conditions are not taken into account.

As a consequence, especially in the presence of DG based on renewable energy sources, such as solar energy, it is not possible to achieve a realistic evaluation of where and when overvoltage and/or overcurrent occurs in a distribution network during a year by simply using a DLF analysis, which is based on some selected combinations of consumer loads and PV power productions.

III. PROBABILISTIC LOAD-FLOW (PLF)

As outlined in the previous Section, a realistic investigation of node voltages is possible if the statistical variations of loads and productions are taken into account. This is done by means of appropriate statistical models for loads and PV generators productions, which will be presented in the following subsections. The consumer loads and PV generator productions are assumed to be mutually independent.

Once the statistical models are defined in terms of probability density functions (pdf), Monte Carlo Simulations (MCS) involves repeating the simulation process using in each simulation a particular set of values of the random variables (loads and PV productions at each node of the considered distribution network, hour by hour) generated in accordance with the corresponding pdf, as explained in Section 4.

A sample from a MCS is similar to a sample of experimental observations. Therefore, the results of a PLF based on MCS in terms of power flows at the various sections of the feeders and voltage profiles at all nodes of the network, may be treated statistically, and the methods of statistical estimation and inference are applicable.

A. Probabilistic load model

The load is assumed to be a random variable ($\boldsymbol{P_L}$) normally distributed within each hour of a given month [7]. Then, the probability density function of $\boldsymbol{P_L}$ is given by the following expression:

$$f_{P_L}(\boldsymbol{P_L}) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(\boldsymbol{P_L}-\bar{\boldsymbol{P}}_L)^2}{2\cdot\sigma^2}} \tag{7}$$

where
$\bar{\boldsymbol{P}}_L$ is the mean value;
σ is the standard deviation.

This means that an expected value and a standard deviation specify the load at each load-point and for each hour.

Note that, in order to improve the paper legibility, the variables that depend on the hour of the day and on the month of the year are written in boldface.

All loads are fully correlated and follow the same pdf. Hence a single pdf can be used for the total load of the network. The feeder load is then distributed among the nodes in proportion to their contract rated capacity.

B. Probabilistic PV Generator Model

The amount of solar radiation that reaches the ground, besides on the daily and yearly apparent motion of the sun, depends on the geographical location (latitude and altitude) and on the climatic conditions (e.g. cloud cover). Many studies have proved that cloudiness is the main factor affecting the difference between the values of solar radiation measured outside the atmosphere and on earthly surface. To account for the difference between these two values, a *daily* or a *hourly clearness index* is used. The *hourly clearness index*, $\boldsymbol{k_t}$, is defined as

$$\boldsymbol{k_t} = \frac{\boldsymbol{I_t}}{\boldsymbol{I_o}} \tag{8}$$

where
$\boldsymbol{I_t}$ is the irradiance on an horizontal plane [kW/m^2],
$\boldsymbol{I_o}$ is the extraterrestrial total solar irradiance [kW/m^2].

Known $\boldsymbol{k_t}$ it is possible to determine the *irradiance* on a surface with inclination β to the horizontal plane, $\boldsymbol{I}_\beta$ [kW/m^2] [8], [9].

Since the PV system is usually equipped with a *Maximum Power Point Tracker* (MPPT) and the relationship between the maximum power per unit area of array surface available from the PV system and $\boldsymbol{I}_\beta$ is linear, [10], the power output of the PV system ($\boldsymbol{P_{pv}}$) is given by:

$$\boldsymbol{P_{pv}} = A_C \cdot \boldsymbol{\eta} \cdot \boldsymbol{I}_\beta = A_C \cdot \boldsymbol{\eta} \cdot \left(\boldsymbol{T} \cdot \boldsymbol{k_t} - \boldsymbol{T'} \cdot \boldsymbol{k_t^2}\right) \tag{9}$$

where
A_C is the array surface area [m^2];
$\boldsymbol{I}_\beta$ the *irradiance* on a surface with inclination β to the horizontal plane [kW/m^2];
η is the efficiency of the PV system in Realistic Reporting Conditions (RRC), [11];
$\boldsymbol{T}$ and $\boldsymbol{T'}$ are parameters that depend on inclination β, declination $\boldsymbol{\delta}$, reflectance of the ground ρ, latitude φ, hour angle $\boldsymbol{\omega}$, sunset hour angle $\boldsymbol{\omega_s}$, day of the year n [8], [9].

From (9), if the probability density function $f_{k_t}(\boldsymbol{k_t})$ for the random variable $\boldsymbol{k_t}$ is known, it is possible to obtain the probability density function $f_{P_{pv}}(\boldsymbol{P_{pv}})$ for $\boldsymbol{P_{pv}}$ by applying the fundamental theorem for the function of a random variable [12].

In the present work the expression used for the probability density function of $\boldsymbol{k_t}$ is the one proposed by Hollands and Huget [12].

Depending on the sign of the parameters $\boldsymbol{T}$ and $\boldsymbol{T'}$, the probability density function $f_{P_{pv}}(\boldsymbol{P_{pv}})$ has four different expressions but only two have a physical meaning [8], [9].

In particular, if $\boldsymbol{T}$>0 and $\boldsymbol{T'}$<0 we have:

$$f_{P_{pv}}(\boldsymbol{P_{pv}}) = \begin{cases} \dfrac{\boldsymbol{C} \cdot \left(k_{tu} - \dfrac{1}{2} \cdot (\boldsymbol{\alpha} + \boldsymbol{\alpha'})\right)}{-k_{tu} \cdot A_c \cdot \boldsymbol{\eta} \cdot \boldsymbol{T'} \cdot \boldsymbol{\alpha'}} \cdot e^{\frac{\lambda}{2}\cdot(\alpha+\alpha')} & \text{if } \boldsymbol{P_{pv}} \in \left[0, \boldsymbol{P_{pv}}(k_{tu})\right] \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

while, if T>0 and T'>0 we have:

$$f_{P_{pv}}(P_{pv}) = \begin{cases} \dfrac{C \cdot \left(k_{tu} - \dfrac{1}{2} \cdot (\alpha - \alpha') \right)}{k_{tu} \cdot A_c \cdot \eta \cdot T' \cdot \alpha'} \cdot e^{\frac{\lambda}{2} \cdot (\alpha - \alpha')} & \\ \qquad\qquad if\ P_{pv} \in \left[0, P_{pv}(k_{tu}) \right] & \\ 0 & \\ \qquad\qquad otherwise & \end{cases} \tag{11}$$

where:

$$\alpha = \frac{T}{T'} \text{ and } \alpha' = \sqrt{\alpha^2 - 4 \cdot \frac{P_{pv}}{\eta \cdot T' \cdot A_c}} \tag{12}$$

k_{tu} is the upper bound for k_t;

C, λ are parameters of the probability density function for k_t [12].

In the software implementation, presented in the next Section, the solar radiation at all nodes are fully correlated and follow the same pdf.

IV. SOFTWARE IMPLEMENTATION OF THE PLF ALGORITHM

The authors developed a software tool on MATLAB® platform to implement the PLF calculations.

This program performs a PLF based on MCS, in which from the generated load and the PV power production time series (according to the pdfs previously described) the corresponding node voltages and branch currents time series are obtained, performing subsequent LF calculations.

Eventually, the calculated voltage and current time series are used for deriving the probability distribution of the voltage at each node and the current at each branch.

The application of the European Standard EN 50160 is also possible, since it recognizes the statistical nature of the voltage variations.

One of the most important application of the proposed method is the evaluation of the maximum value of the total PV peak power that can be installed in the nodes of the network without violating voltage and current constraints [5].

To do this it is preliminary necessary to assess the average number of hours in the year with under/overvoltage and overcurrent.

Fig. 2 shows the flow chart explaining the PLF algorithm used to perform this calculation.

The input data are provided to the program in terms of:

- parameters of the pdfs for power required by the loads and produced by the PV generators for each hour;
- electrical and geometrical parameters of the distribution network;
- initial nodal voltage values;
- minimum and maximum voltage values (V_{min} and V_{max}) allowed by EN 50160;
- values of maximum current allowed in each branch (I_{max}) given by the line current carrying capacity.

Based on these detailed input specifications and on the probability density functions previously described in subsections *A* and *B*, the program generates for each hour ($h=1...H$, with $H=8760$, number of hours in a year) of the year ($y=1...Y$) loads and productions statistic values (to ensure reasonable accuracy of the calculation from the Monte Carlo simulations, an opportune number of repetitions Y are to be performed).

For each hour, the following steps are performed:

- a LF calculation is carried out using as inputs the sample values of loads and PV productions. In this way it is possible to obtain for each hour the sample value of the voltage at each node ($V_{h,i}$, $i=1...$n) and the current in each branch ($I_{h,b}$, $b=1...n$);
- a test is done in order to detect undervoltage ($V_{h,i}<V_{min}$), overvoltage ($V_{h,i}>V_{max}$), and overcurrent ($I_{h,b}>I_{max}$);
- the results are saved and made available for subsequent analysis;
- if the counter h is less than H, it is increased by one, otherwise the counter y is increased by one and the new statistic data are generated for the current year.

To reduce the calculation time it is necessary to choose an appropriate initial value for the nodal voltages. The initial value is calculated carrying out a fast DLF (with linear equations) which assumes constant current" model for loads and generators.

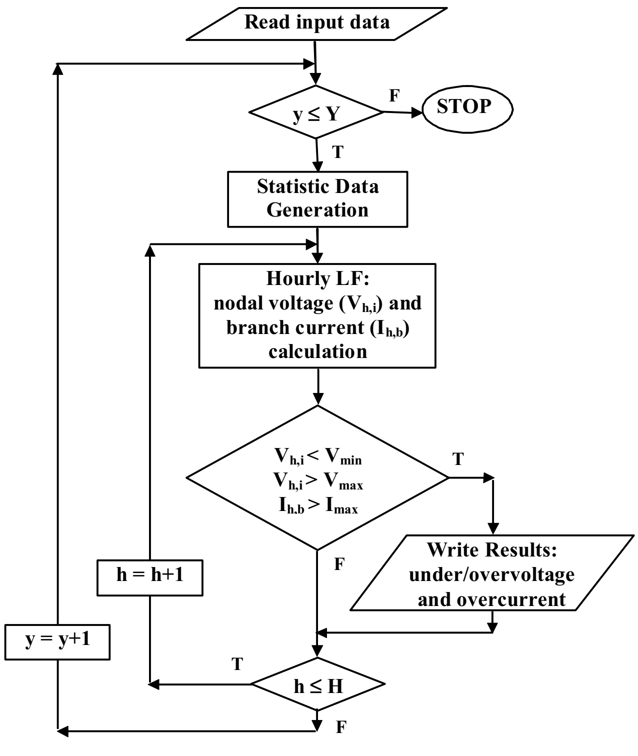

Fig. 2. Flow chart of the PLF algorithm.

V. Conclusions

In this paper theoretical concepts and models have been presented to develop a Probabilistic Load Flow (PLF) using Monte Carlo techniques for radial distribution networks with grid-connected photovoltaic generators.

A software tool has been presented to implement the proposed method that allows to evaluate the effect of connecting PV DG units to the distribution network, with specific regard to slow voltage variations and line current carrying capacity.

The procedure allows to evaluate the penetration limits of PV DG with respect to voltage and thermal constraints. For example, it is possible to calculate the average number of hours with under/overvoltage and overcurrent per year for a LV distribution network with several PV generators.

This makes it also possible to calculate the yearly energy not supplied to the network owing to overvoltage protections tripping and consequent PV generators disconnection to avoid unacceptable overvoltage.

The proposed PLF recognizes the statistical nature of the voltage variations and it is in keeping with the approach of the European Norm EN50160.

In the companion paper [6] a case study is presented to show the application of the proposed method.

References

[1] Grubb M.: *Renewable Energy Strategies for Europe - Volume I, Foundations and Context*, The Royal Institute of International Affairs, London, U.K., 1995.

[2] European Standard EN 50160: *Voltage characteristics of electricity supplied by public distribution systems*.

[3] Conti S., Raiti S., Tina G.: *Small-Scale Embedded Generation Effect on Voltage Profile: an Analytical Method*, IEE Proceedings on Generation, Transmission and Distribution, January 2003, Vol. 150, No. 1, pp. 78-86.

[4] Conti S., Raiti S., Tina G., Vagliasindi U.: *Integration of Multiple PV Units In Urban Power Distribution Systems*, Solar Energy (Elsevier), August 2003, Vol. 75 No. 2, pp. 87-94.

[5] Conti S., Raiti S., Tina G., Vagliasindi U.: *Distributed Generation in LV Distribution Networks: Voltage and Thermal Constraints*, Proc. of the 2003 IEEE PowerTech Conf., 23-26 June 2003, Bologna, Italy.

[6] Conti S., Raiti S.: *Probabilistic Load Flow for Distribution Networks with Photovoltaic Generators. Part 2 - Application to a Case Study*, International Conference on Clean Electrical Power, Capri, Italy, May 21st-23rd, 2007.

[7] Hatziargyriou N.D., Karakatsanis T.S., Papadopoulos M., *Probabilistic Load Flow In Distribution Systems Containing Dispersed Wind Power Generation*, IEEE Transactions on Power Systems, Vol. 8, No. 1, February 1993.

[8] Conti S., Crimi T., Raiti S., Tina G., Vagliasindi U.: *Probabilistic Approach to Assess the Performance of Grid-Connected PV Systems*, Proc. of 7th International Conference on Probabilistic Methods applied to Power Systems, September 22-26, Naples, Italy.

[9] Gagliano S., Raiti S., Tina G.: *Hybrid solar/wind power system probabilistic modeling for long-term performance assessment*, Solar Energy, Vol. 80, Issue 5, May, 2006, pp. 578-588.

[10] Kroposki B., Emery K., Myers D., Mrig L.: *A Comparison of Photovoltaic Module Performance Evaluation Methodologies for Energy Ratings*, Proceedings of 1st WCPEC, Hawaii, USA, December 5-9, 1994.

[11] Gay C. F., Rumberg J. E., Wilson J. H.: *AM/PM-All-Day Module Performance Measurements*, Proc. of 16th IEEE PVSC, San Diego, CA (1982).

[12] Papoulis A.: *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York 1982.

[13] Hollands K.T.G., Huget R.G.: *A probability density function for the clearness index, with applications*, Solar Energy, 30 (1983) pp. 195-209.