Electric Power Systems Research 125 (2015) 133–140

Contents lists available at ScienceDirect

Electric Power Systems Research

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

A simple and efficient method for fast analysis of renewable generation connection to active distribution networks

Alberto Pagnetti*, Gauthier Delille

EDF R&D, 1 Av. de Gaulle, 92141 Clamart, France

ARTICLE INFO

Article history:
Received 5 August 2014
Received in revised form 4 February 2015
Accepted 30 March 2015
Available online 24 April 2015

Keywords:
Active distribution networks
Distributed generation
Network planning
Load flow analysis
Time series

ABSTRACT

This paper introduces a new method to analyze the connection of an intermittent generation unit to an active distribution network. This method permits the assessment of the voltage/current constraints that the unit might generate during its operation, and also the study of various smart solutions, alternative to grid reinforcement, like advanced voltage regulation schemes or generation curtailment. To do that, time series are used as inputs (for generation profiles, load profiles on the studied network, etc...) but a clustering technique is employed to generate a limited number of equivalent scenarios in order to reduce the number of points for which a load flow is calculated. This clustering technique allows a significant reduction of the computation time, thus making it possible to simulate very effectively one or multiple years of operation of the distributed generation (DG) unit(s) in a network. These simulations give useful results such as critical yearly quantities (like for instance losses or energy curtailed) that can be used to perform a proper technical and economic analysis.

Two clustering methods are presented: the first neglects any form of correlation between the input time series, whereas the second makes it possible to consider the effect of possible correlations between load and generation. When the second method is used, the time series of the outputs (like energy curtailed, or the voltage at a given point in the network) can also be reconstructed with a good accuracy.

1. Introduction

The connection of any new distributed generation (DG) unit to a distribution network may require grid updates to avoid over-current or over-voltage limit violations: the growing penetration of renewable energy sources (RES) at MV and LV levels results in a significant need for grid development, with subsequent costs [1].

To reduce these burdens, some alternative solutions to grid reinforcement ("no-grid solutions") can be put in place in some cases [2–4,20] if they are technically feasible and more interesting from the economic point of view than "fit and forget" traditional solutions. However, the lack of suitable planning approaches hinders their adoption: the usual connection studies based on the so-called "deterministic methods" or "worst-case approaches", though well suited to design robust grid solutions, are not suitable when it comes to the study of most of the alternatives to grid reinforcement. This is because carrying out proper technical and economic analysis of these options requires, in most cases, probabilistic connection studies in order to assess the technical impact of each solution (yearly amount of curtailed generation for instance) and thus the costs borne by the distribution system operator (DSO) and the generation owner.

Many stochastic load-flow approaches have been already proposed in the literature [5–9] to analyze the effect of intermittent generation on a large network. In these approaches, critical inputs such as loads and generation profiles are characterized using probability density functions. This allows to devise optimal power flow strategies and also to account for reliability of equipment. These approaches have already been used for network planning [10] as well. Other methods use time series as inputs [11–16]. These have the advantages of being simpler to implement, and are adapted to some processes that are time dependent and for which time measurements or predictions exist. A serious drawback is the longer simulation time, if the time series simulation needs to be carried out over a long time period such as one or several years, with a time step of one hour or less.

In this paper, we propose to use time series as inputs to simulate the operation of an active distribution network over one or several years, with significant shares of RES. Similar techniques have already been shown to be useful in order to characterize no-grid alternatives in DG/RES connection studies [2,10,14].

* Corresponding author. Tel.: +33 1 47 65 33 21.
E-mail address: alberto.pagnetti@edf.fr (A. Pagnetti).

http://dx.doi.org/10.1016/j.epsr.2015.03.025
0378-7796/

A new technique is presented in this paper that permits to drastically reduce the simulation time, up to a factor of 100 compared to time series, though with very limited impact on the results of interest for grid planning purposes. This technique is based on a clustering approach [17,18].

Using a case study, the paper illustrates how this method makes it possible to characterize the voltage constraints due to the connection of a new DG unit and gives some output quantities over a year or multiple years of operation to feed detailed technical and economic analysis. We furthermore show that even if a reduced number of representative points are simulated instead of complete year(s), it is possible to reconstruct the whole time series of any of the outputs evaluated in the power flow with a reasonable accuracy.

All the results have been validated comparing our method to a time series analysis based on sequential load-flow evaluation for each point of the time series input profiles.

2. Outline of the article

If the connection of a new DG unit is likely to result in the violation of voltage constraints, a number of alternative "no-grid" solutions may be efficient, such as: local voltage control through modulation of the reactive power [19], real-time control of the voltage reference of the on-load tap changer (OLTC) at the main substation, generation curtailment of distributed energy storage. These options are briefly described in reference [2]. Replacing the usual deterministic or "worst-case" planning scenarios like "minimum load, maximum generation" or "maximum load, no generation" by the multiple load-flows of long time series can give interesting results but unfortunately it happens to be very time-consuming as it requires the calculation of a load-flow for each time step of the input time series profiles. For instance, if a network with 500 nodes is considered, the simulation of one year of operation with input data sampled at 10 min-steps requires the computation of 52,560 or 52,704 load-flows, and, depending on the complexity of such network, this may necessitate more than 30 min on a standard Windows PC.

Clearly this may limit thorough analyses particularly if multiple cases are to be tested separately.

The method proposed herein presents the same advantages as the time series methods, but requires much fewer calculations: it is in fact up to 50–100 times faster while maintaining a very similar precision level. Inputs, namely one or multi-year time series for load, generation and primary substation busbar voltages remain the same but, using clustering techniques, these are "compressed" under the form of a shortlist of equivalent simulation scenarios with associated weighting coefficients.

Two simple methods are proposed, that can lead to significant reduction in the simulation time with limited impact on accuracy:

- The first method, based on histograms and discussed in Section 3, disregards any correlation that might exist between the input vectors, it can be used with generic input data when historical profiles are not available or not accurate enough. The time correlation between input vectors is not accounted for, and therefore all possible combinations of input variables are tested and weighted accordingly to the product of their individual occurrence in input vectors.
- The second method, discussed in Section 4, is based on multidimensional clustering and thus takes into account the time correlation between the input vectors. The correlation ensures that only scenarios that really occur in the considered time series are simulated, and thus a further reduction of the number of load-flow calculations is obtained.

Both methods use a limited number of representative points/scenarios instead of huge time series, and a weighting coefficient describes the occurrence of each of these representative scenarios in terms of hours-per-year. The use of this weighting coefficient makes it possible to evaluate precisely the same quantities that would be evaluated using the time series approach, namely the over-current/over-voltage constraint rate/year, the yearly amount of line losses and, when alternatives to grid reinforcement are considered, the amount of curtailed energy/year of reactive power injected/year, etc. It is important that these quantities can be calculated precisely as they are then used to feed economic assessments: cost of the losses, cost of over-sizing the power electronic converters, cost of energy lost due to curtailment, etc.. . .

Since one of the main weaknesses of both methods is that they perform the load-flow calculations without considering the order through which the simulation scenarios follow each other in the time series, and since this piece of information is required to study "no-grid" solutions such as energy storage, an additional step has been put in place to reconstruct, from the calculations made on the set of equivalent scenarios, the time vector of output quantities that would be obtained using a time series approach.

3. Description of method without correlation (H1D)

To outline the method, in this section, only two input variables (generation and load) are used, but the method can be generalized for multiple input variables.

In order to develop this method, the input variables are represented on a histogram (see Fig. 1): the frequencies of the values are normalized to obtain the probability of each class or, in other words, its statistical weight.

Each class is defined as:

$$X_j \le P_{prod}(i) < X_{j+1} \quad j \in [1, N_P] \tag{1}$$

$$Y_l \le P_{load}(i) < Y_{l+1} \quad l \in [1, N_L] \tag{2}$$

where X_j and Y_l are the lower/upper bound of each class, N_P and N_L are the number of classes chosen to describe each input variable.

A representative point (P_j for the generation and C_l for the consumption) for each class is calculated as the mean value of all the values that belong to that class. A weight w_{pj}, w_{cl} is calculated as the probability of occurrence of each class and is associated to each representative point.

All the representative points of each input variables are combined to obtain simulation scenarios through all the possible combinations of generation and load scenarios. As the variables are considered statistically independent, the weight of each equivalent simulation point is obtained by multiplying the weight of each of the input variables.

$$w(i) = w_{pj} \cdot w_{cl} \tag{3}$$

where $w(i)$ is the weighting coefficient of the ith simulation scenario.

This can be summarized by the example in Table 1.

Table 1
Different scenarios obtained by combining the mean values of each of the classes of the input variables. The number of scenarios is $N = N_P \cdot N_C$.

Scenario	Production	Load	Weight (in hour/year)
1	P_1	C_1	$w_{p1} \cdot w_{c1} \times 8760$
2	P_2	C_1	$w_{p2} \cdot w_{c1} \times 8760$
...	...	...	...
N	PN	CN	$w_{pN} \cdot w_{cN} \times 8760$