COMPARISON OF TWO LOAD FLOW APPROXIMATION METHODS FOR DISTRIBUTION PLANNING

Gianni CELLI
University of Cagliari – Italy
celli@diee.unica.it

Fabrizio PILO
University of Cagliari – Italy
pilo@diee.unica.it

Simona RUGGERI
University of Cagliari – Italy
simona.ruggeri@diee.unica.it

Heloise BARAFFE
EDF R&D – France
heloise.baraffe@edf.fr

Josselin FOURNEL
EDF R&D – France
josselin.fournel@edf.fr

Gilles MALARANGE
EDF R&D – France
gilles.malarange@edf.fr

Juliette MORIN
EDF R&D – France
juliette.morin@edf.fr

ABSTRACT

The integration of no-network planning alternatives into distribution system planning tools requires a proper modelling of load and generation behaviours within network calculation. To capture the operational aspects that can affect the planning stage, the time variability of demand and generation has to be explicitly represented in the planning calculations. Moreover, to reduce the computational burden and take account of the uncertain production and consumption of electricity, stochastic models have to be implemented. Two approaches are commonly proposed in the scientific literature: the clustering of demand and generation variability in a year taking account of the occurrence probability of each cluster, and the identification of typical daily profiles that characterise the weekly and seasonal customers' behaviour during a year modelling their uncertainties through suitable probability density functions.

In the paper, the comparison of these two approaches have been shown, underlining qualities and faults (e.g., accuracy, computational time burden, their applicability). The comparison has been performed considering a realistic distribution network, in presence of distributed generation (DG).

INTRODUCTION

The increasing presence of DG together with load growth, and the development of information and communication technologies (ICT) are driving the evolution of electricity distribution from passive to active distribution networks. This revolution is deeply changing the common approach adopted to plan, design and operate the distribution networks. Indeed, the typical utility planning procedures, based on network reinforcement to cope with the worst-case scenario ("Fit & Forget" approach"), can bring today to excessive investments justified only by critical operating conditions that are extremely rare to happen. Therefore, as underlined by the scientific community [1] and remarked by the recent European "Winter Package" [2], it is essential to go beyond this traditional planning perspective by choosing modern methodologies able to include among the planning solutions the potential support coming from the integration of new technologies and the active management of Distributed Energy Resources (DER).

In order to incorporate operation into the distribution network planning, the behaviour of the customers cannot be still represented with a simple snapshot (average or worst-case conditions) but their temporal variability has to be modelled within network calculations. Indeed, the effectiveness of some Smart Grid (SG) solutions (like online network reconfiguration, reactive control, demand response, management of energy storage systems, …), their frequency of application to solve operational issues (e.g. line overcurrent or excessive voltage deviation) and their correct sizing can be estimated only by taking into account the different load and generation operations within the day, along the week and among the seasons. An exact assessment would require the execution of a load flow calculation sequentially for each operating point over one or several years), but this approach is unfeasible for planning studies, due to computation time constraints. For this reason, suitable simplifications are proposed in order to reduce the computational burden and preserve the quality of the solutions.

A second important aspect of these new planning tools is the ability to correctly deal with all the uncertainties that can affect the distribution system planning scenario. Specifically, in this paper it has been considered the uncertainties on customers' behaviour (intermittent power production from renewables, DER availability, and demand deviation). Stochastic models are used to represent these uncertainties, explicitly integrate the risk concept in the system design, and help planners to take objective and transparent decisions.

The aim of this paper is to compare the results of two MV distribution system planning tools, based on different data modelling techniques, on a realistic test case in terms of nodal voltages and line currents:

- The first one uses a clustering of input data (generation and load) to reduce the number of load-flows, and associate an occurrence probability to each cluster to model the stochastic behaviour of customers;
- The second tool is based on the definition of typical daily profiles and implements a probabilistic load flow

(PLF) where the stochastic behaviour of customers is approximated with normal distributions.

DATA MODELLING

One of the most interesting tasks of the planning procedure is related to data modelling. To capture the operational aspects that can affect the planning stage, the time variability of demand and generation has to be explicitly represented in the planning calculations. The challenge is to find to what extent operational aspects need to be modelled in planning.

Although fine granularities (e.g., minute by minute) are capable of capturing detailed operational aspects, when it comes to planning calculations (especially in medium- to long-term horizons), they often prove to be unnecessary and can also lead to time-consuming processes. On the contrary, simplistic representations will not bring the benefits from operational strategies to the planning studies, affecting the quality of the results [3].

Traditionally, techniques of load analysis and synthesis were applied to generate characteristic load profiles used to categorize customers into different classes (approach still used by the majority of the distribution utilities). The results of the load analysis were used to produce suitable approximations of the peak demand, starting from the yearly energy consumption. This traditional representation was used in distribution network planning studies by assuming unique yearly values of demand and generation. Peak values were calculated for classifying worsen operation conditions and used, together with a constant yearly growth rate, in the deterministic 'fit and forget' approach to plan network expansion for a predefined planning period. Average values (if necessary) were considered for Joule losses estimations and reliability analyses. The main drawback of this deterministic approach is that the distribution network is designed assuming the worsen conditions as certain, even if actually they have a very low probability of occurrence.

Two possible approaches are analysed as example of methodologies that could be used to represent the variability of demand and generation [1]: (i) the characterization and clustering of demand and generation variability in a year, and (ii) the identification of typical daily profiles that characterize the seasonal behaviour during a year. The first approach recognizes and clusters different operating conditions from yearly sequences of measurements in order to decrease the computational burden of the planning calculations [5]-[7]. Each state is characterized by a probability of occurrence. The second one extracts some typical representative days from the annual time-series, divides them into elementary intervals (e.g., 1 hour), and repeats network calculations sequentially for each of them [3], [8], [9]. Within these intervals, loads and generators are modelled with suitable probabilistic density functions.

Data clustering

In this approach [7], a clustering method is used by EDF R&D to generate only a limited number of equivalent scenarios from multi-years' time-series inputs for both load and generation. The method is based on a multi-dimensional clustering (one dimension for each type of input). Hence, the time correlation between inputs is taken into account. In short, the method consists of a 5-steps procedure:

1) Create a n-dimensional mesh (typically n=3 when considering, PV, wind generation and consumption). This mesh is characterized by the number of values in each dimensions, giving the total number of classes of the mesh. These classes are ordered and numbered.
2) Swipe and store input values into corresponding classes of the mesh,
3) Evaluate a weighting factor for each classes of the mesh. Once the sweeping procedure is completed, the weighting coefficient is given by the number of occurrences of the values in each class (number of time-steps/year).
4) Select a representative point (herein the barycentre of the points inside a given class) for each class having a non-zero weight. At this step, the representative points with their size are available as depicted in Fig. 1.
5) Launch load-flow calculations only on the representative points having non-zero weight.

The use of weighting coefficient enables 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. Such approach requires much fewer calculations compared to the analysis of the input pattern. It is in fact up to 50–100 times faster without worsening the level of accuracy [7].

The choice of the mesh is empirical, as the results do not linearly follow the number of simulations (dependent upon the size of the mesh). The mesh can also be made finer or coarser in one dimension, depending on the impact of the dimension (load or generation for instance) on output results. A further possible refinement is to build a mesh that is not completely uniform: if the classes on the outer edges are made smaller it is then possible to better simulate these extreme points and this can be useful in some cases.

It could be noted that this approach allows reconstructing the time series of any output variable. To do that, an additional step is added to the pre-processing algorithms: a time series signal is created whilst sweeping the input variables. This signal indicates the class associated with each point of the input time series. This signal allows the reconstruction of the outputs time series.

Identification of typical daily profiles approach

The adoption of typical daily patterns to represent the behaviour of distribution network customers in a year (or in different seasons, between workdays and weekends, …) allows representing the time variability of demand and generation in the planning calculations. Typical days are then divided into elementary intervals (e.g., one minute, five minutes, fifteen minutes or one hour) and the network calculations are repeated sequentially for each of them. The chronological representation of the events allows better modelling actions that need to be assessed in more sequential intervals/states (storage and demand response actions). All these discretization options are useful at the planning stage to highlight the capability of the active management to prevent some constraint violations, deferring possible network investments. A fine discretization can capture some extreme operational conditions that would be smoothed with a rough representation but, generally, the risk of disregarding them in the planning decision can be acceptable compared to the increase of computation burden. For the above reasons, a time step of one hour is commonly used to represent load and generation profiles and to analyses the impact of operation strategies in the planning studies [1], [3].

This approach is proposed in [8], [9]. The tool developed allows the decision maker to perform integrated planning studies (e.g. find the optimal DG arrangement for an existing network) and use this result to plan the medium-term optimal network development and thus identify the optimal network evolution in a given study period with the optimal amount of DG placed in the most convenient points can be found. The tool makes use of graphical and mathematical decision techniques, considers explicitly uncertainties and risks and allows the consideration of multiple and diverse solutions. Such tool can be a useful aid in making decisions in present uncertain and mutable scenarios.

Probabilistic Load Flow Calculation

Load and generation profiles are inherently uncertain. The uncertainty is modelled by representing each point of the profile as a stochastic variable described by a Gaussian distribution. This representation is valid for modelling the load behaviour and it is generally accepted in planning studies for renewables, even if this approximation can become less founded when the DG penetration is high. The main advantage of this assumption is the possibility to implement a simplified PLF to avoid the increasing of the computation time, during the analysis of each network alternative examined by the heuristic algorithm [8], [9]. In the calculation, some proper approximations are introduced, and the radial operating structure of the system is considered. These choices allow significantly simplifying the general techniques used in transmission systems. In this case the following assumptions can be made [9]:

- correlation between loads is linear or null;

- if traditional generators are present (like gas turbine or small industrial combined heat and power plants), no correlation is assumed among them, because no dispatching action for distribution network is considered. If the generators are fed by renewable energy sources (PV plants of wind turbines) correlations may exist due to the relative small distances among them in a MV network and the consequent similar ambient and weather conditions (same solar irradiation or wind speed). In this case, if necessary, the correlation is considered linear.
- local correlations between generators and loads are modelled with linear dependence.

When the power in each node is known, it is possible to calculate nodal currents, by using the nominal voltage. Thus, the voltage in every node can be determined directly, by using nodal currents like random variables, with the equation (1):

$$[V] = [Z] \cdot [I] \tag{1}$$

where Z denotes the impedance matrix.

This approach considers the real and imaginary parts of the nodal current by means of their expected value and standard deviation. By substituting current expected values in system (1), the expected values of the real and imaginary parts of nodal voltages are determined. In order to calculate the nodal voltage variances, it is necessary to substitute in equation (1) the current standard deviations instead of the mean values. Then it is possible to calculate the branch currents simply dividing the corresponding branch voltage drop by the branch impedance. Real and imaginary parts of a branch voltage drop have to be calculated separately by combining the expected values and the standard deviations of the voltage at the edge extremes. Once known the currents flowing in each branch and the voltage of each node through their corresponding PDF (Probability Density Function), it is possible to choose the correct size of each conductor and to verify all the technical constraints (like the voltage profile), taking into account the uncertainties associated to these electrical variables.

CASE STUDY

Description of the case study

For the analysis, two feeders (showed in Fig. 1), presenting different characteristics have been selected. Feeder A (light blue line in Fig. 1) is a typical rural feeder, characterised by long overhead lines with low load demand and 1 generator of 7.2 MVA; while Feeder B is characterised by short buried lines and a high load demand (urban type) and 6.9. MVA wind generator. The feeders' characteristics are summarized in Table 1.

From each feeder, the nodes and the branches supposed to be more critical in the network (i.e., final nodes, nodes where the generators are installed, bifurcation) are

considered.

Table 1. Main characteristics of the feeders analysed

	Feeder A	Feeder B
Num. of Sec. Substations	90	4
Num. of Lines	243	14
OHL lines length	50 km	0
Buried lines length	17 km	4.11 km
Total length	67 km	4.11 km
P load installed	11.63 MW	12.85 MW
P gen installed	6.9 MVA	7.2 MVA

The data considered are three years of 10-min load/generation profiles at the primary substation scale. Following a top down approach, the profiles have been dispatched between all the loads and generators proportionally to their nominal power.

Considering the clustering methodology, an 8*8*8 mesh (wind*PV*load) is used and requires 311 load-flow calculations in PowerFactory. This number is lower than the theoretical number of load flows (8*8*8 = 512), because several classes are not observed in the input profiles, and therefore not calculated.

Regarding the typical days approach, the PLF has been performed varying the numbers of typical days considered:

- 1 typical day (that means 24 PLF);
- 8 typical days: 2 typical days (working/holiday) per season (192 PLF);
- 12 typical days: 3 typical days (working day/Saturday and Sunday) per season (288 PLF);
- 36 typical days 3 typical days (working day/Saturday and Sunday) per month (864 PLF calculation).

In particular, for the generation profiles, the generators have been modelled with their typical daily generation profiles that take into account the unpredictability of the primary source by means of normal PDF. In particular, wind generation is represented with a constant mean output power and a high standard deviation, equal in each hour. On the contrary, the photovoltaic plants have a growing production in the morning, a decreasing production in the afternoon and no production during the night. Their standard deviations are variable hourly (low in the sunrise and sunset, and high in the midday hours). In order to obtain the quantiles, the mean value and the standard deviation are combined considering the occurrence of the day during a year.

Qualitative comparison of the approaches

The typical-day approach, compared to the clustering method, has the advantage of taking into account the correlations between consecutive time steps. Therefore, it allows a better modelling of levers with time-step dependency (e.g. storage and demand response). Other levers, such as distributed generation curtailment, can still

be addressed by the clustering approach using results post-processing [10].

Comparison of the methods' accuracy
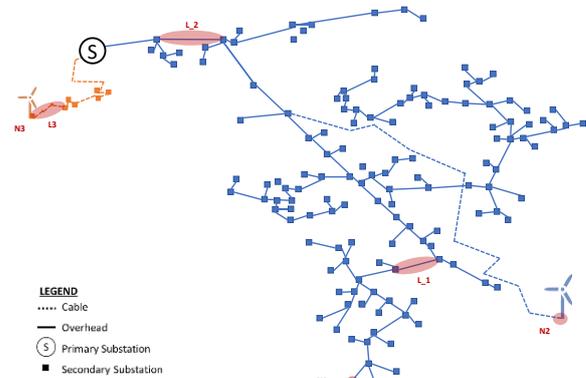
The results obtained using the two methodologies are then

Fig. 1. The feeders analysed.

compared with the reference, which is built performing the sequential exact load flow calculations at a 10-minute time step (157000 load flow calculations) with Power Factory. Fig. 2 shows the box plot of three representative nodes of the network: N1 (Feeder A), N2 (where the wind generator is located in Feeder B) and N3 (where the generator of Feeder B is located), with blue, yellow and green box and whisker respectively, considering the two different approaches (clustering *CLS*, and typical days TD with different numbers of typical days) and the benchmark value (*REF*). Each box represents the extreme values, the 10th quantile, the median and the 90th quantile. From the figure it is possible to see that the clustering allows a good representation of the voltage in the nodes (both the median and the extremes quantiles); while in the second approach, increasing the number of typical days lowers the difference with the exact calculation, both for the median and the last quantiles.

Table 2 summarizes the error in the voltage evaluation of the two approaches compared to the reference values in the three nodes.
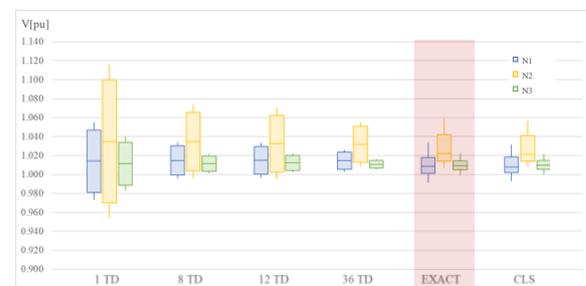

Fig. 2. Box plot of node N1 (Feeder A), blue, N2 (Feeder A) in yellow, N3 in green (Feeder B).

Table 2 – Voltage errors of median, 90th and 100th quantile [%]

		1 TD	8 TD	12 TD	36 TD	CLS
N1	Q50	-0.5%	-0.6%	-0.6%	-0.6%	0.1%
	Q90	-2.9%	-1.2%	-1.1%	-0.6%	0.0%
	Q100	-2.1%	0.0%	0.0%	0.8%	0.2%
N2	Q50	-1.3%	-1.3%	-1.1%	-1.0%	0.0%
	Q90	-5.7%	-2.4%	-2.0%	-0.9%	0.1%
	Q100	-5.6%	-1.4%	-1.0%	0.4%	0.2%
N3	Q50	-0.2%	-0.2%	-0.3%	-0.1%	0.0%
	Q 90	-2.0%	-0.5%	-0.5%	0.0%	-0.1%
	Q100	-1.7%	0.1%	0.1%	0.7%	0.1%

Regarding the currents, the errors are higher for both approaches. Table 3 shows the current error (normalized by the cable rated current). The clustering of different operating conditions from yearly profiles approach has an optimal accuracy in both the evaluations. The typical day approach is less precise, especially when few typical days are considered. Its highest precision is reached for the evaluation of the median, and decreases in the evaluation of the 90th quantile, overestimating the current in the branches. The biggest error occurs in the branches hosting the current generated by the wind generators (lines L2 and L3). Several reasons can explain the higher errors for the PLF method, compared to the clustering approach:

- The simplification of the statistical behaviour of demand and generation with normal distributions,
- The linearization of the load flow,
- No correlations between load and generation.

Table 3 – Current errors of median, 90th and 100th quantile [%]

		1 TD	8 TD	12 TD	36 TD	CLS
L1	Q50	0.1%	-0.1%	-0.7%	0.1%	0.3%
	Q90	10.5%	0.1%	0.9%	0.3%	-0.1%
	Q100	12.1%	-0.8%	-4.4%	0.2%	-0.0%
L2	Q50	-5.0%	-4.0%	-4.0%	-5.0%	-1.0%
	Q90	70.0%	25.0%	24.0%	10.0%	-1.0%
	Q100	81.0%	25.0%	23.0%	6.0%	5.0%
L3	Q50	9.6%	9.5%	9.5%	8.5%	-0.3%
	Q90	23.5%	-6.0%	-5.1%	8.4%	2.6%
	Q100	45.2%	26.6%	9.6%	8.5%	2.9%

Both approximation methods offer a substantial reduction of computation time at the expense of results accuracy, especially on current values. The accuracy of both methods can be improved, either by increasing the number of typical days (as shown on Fig. 2, Table 2 and Table 3) or by using a finer mesh for the CLS method. This trade-off is necessary to reach acceptable results for medium or long-term planning studies.

CONCLUSIONS AND FUTURE WORKS

In the paper, the analysis and comparison of two possible data modelling in planning tools are performed. Given a realistic network, characterized by two feeders which present different characteristics (e.g., network topology, cable type, amount of load and generation installed), the methodology based on the identification of typical daily profiles that characterize the seasonal behaviour during a year and the clustering approach are compared with the

time series load flow calculation (157680 load flow calculations) performed with Digsilent PowerFactory. The accuracy of both voltage and current results is higher for the clustering method, which uses less hypotheses on the generation and load profiles. Therefore, this method looks more suitable for planning studies on present-day and near future distribution grids. However, this method is unable to easily account for time dependency. Thus, for prospective studies, the use of typical day profiles could be more suitable. The approach based on the PLF is a promising method, which is limited by the assumption of normal distributions of loads and generation profiles and the load flow linearization.

Another comparison between the tools should be made considering feeders with PV generators; it will allow evaluating if their behaviours are properly modelled.

REFERENCES

[1] CIGRE Technical Brochure, Working Group C6.19, "Planning and Optimization Methods for Active Distribution Systems", 2014

[2] L. Meeus, A. Nouicer, "Technical report on the EU clean energy package", European University Institute July 2018

[3] R. F. Arritt, R. C. Dugan, "Distribution System Analysis and the Future Smart Grid", IEEE Transactions On Industry Applications, Vol. 47, No. 6, November/December 2011

[4] F. Pilo, G. Celli, E. Ghiani, G. G. Soma, "New electricity distribution network planning approaches for integrating renewable", John Wiley & Sons, Ltd, Volume 2, March/April 2013

[5] L. F. Ochoa, C. Dent, G. P. Harrison, "Distribution network capacity assessment: Variable DG and active networks," IEEE Trans. on Power Systems, vol. 25, no. 1, pp. 87-95, Feb. 2010

[6] N. Koutsoukis, P. Georgilakis N. Hatziargyrioy, "Active distribution network planning based on a hybrid genetic algorithm-nonlinear programming method", in proc. CIRED 2017, Glasgow 12-15 June 2017

[7] A. Pagnetti, G. Delille, "A simple and efficient method for fast analysis of renewable generation connection to active distribution networks", EPSR, Volume 125, August 2015, Pages 133-140

[8] G. Celli, S. Mocci, F. Pilo, V. Allegranza, R. Cicoria; "An Integrated Tool for Optimal Active Network Planning", in Proc. of CIRED 2003, Barcelona, Spain, May 12-15, 2003

[9] G. Celli, S. Mocci, F. Pilo, R. Cicoria; "Probabilistic optimization of MV distribution network in presence of distributed generation", in Proc. of 14th PSCC, Sevilla, Spain, June 24-28, 2002

[10] J. Fournel, R. Nedellec, A. Pagnetti, A. Ahram, "On the use of yearly load scenarios to estimate the volume of curtailed generation". in Proc. of IEEE International Energy Conference (ENERGYCON), Leuven April 4-8, 2016 2016 (pp. 1-5).