European Association for the Development of Renewable Energies, Environment and Power Quality

International Conference on Renewable Energies and Power Quality (ICREPQ'09)

Valencia (Spain), 15th to 17th April, 2009

Prediction of photovoltaic generation for distribution network planning

[1]Miquel Ramon-Marin, [2]Andreas Sumper, [2]Roberto Villafáfila-Robles, [2]Carlos González-de-Miguel

[1]Endesa Distribución Eléctrica S.L., Planificación y Calidad de la Red, Av. Vilanova 12, 4ª Planta, 08018 Barcelona, Spain

[2]E.U. d'Enginyeria Tècnica Industrial de Barcelona – Dept. Ingeniería Eléctrica

C/Compte Urgell, 187, 08036 Barcelona, Spain

miquel.ramon@endesa.es, andreas.sumper@citcea.upc.edu, roberto.villafafila@citcea.upc.es, carlos.gonzalez@citcea.upc.edu,

***Abstract*— Nowadays utilities must optimize the utilization of their distribution networks, analyzing possibilities from technical, economic and legal point of view, and have to face several challenges. One of them is the strong presence of renewable generators, which are promoted by administrations and represent and answer to several social needs, as security of energy supply, the appearance of a new local industrial sector operating globally, and growing price of conventional primary sources. But this means a revolution in the way transport, and specially distribution networks are planned, developed and exploited. The present paper begins an analysis of the impact of this revolution from the distributor's point of view, situating the actual boundary conditions in Spain and focusing on medium voltage network.**

Index Terms- Distribution network planning, renewable generators, electricity demand, medium voltage, prediction of photovoltaic generation.

I. INTRODUCTION

Motivated by electrical market liberalization, and several legislation modifications promoting renewable generators with feed-in tariffs, in a few years Spain has evolved from a vertical integrated electrical network and business model, to a new horizontal distribution network and business model.

From an only sense of power flow from generation to consumption, with manageable generation from conventional sources and a predictable consumption, now there is a liberalized market poll where several agents can offer their capacity of generating electricity, and therefore is implemented a real horizontal network model [1], with diverse senses of power flow according to diverse variables.

Before the same utility often owned generation centrals, transport and distribution networks, separating different operators only by geographical boundaries, but now, generation is increasingly distributed, there is a Transport System Operator (TSO) who guaranties access to the network to any agent interested in, and which, in addition, has exponential increasingly presence of renewable generators of several technologies, connecting, according to legislation classification by power, at all voltage levels of the network, affecting it in several different manners.

It is therefore necessary to look for solutions that face these challenges [2-5] and allow finding real dimension of influence of these generators in distribution network, being possible to approach this situation from multiple points of view, as for example:

Quality of the product: Voltage, Harmonics, Frequency.

Quality of service: Automatization, TIEPI, NIEPI.

Losses: In network, transformation, generation.

Reactive Power management: production, retribution.

System reliability: risks, criticisms, contingencies analysis.

Network point of connection optimization, which depends on point of view: distributor, generator, consumer.

Investment optimization: network planning, infrastructures utilization, activities retribution as transport, distribution, generation.

This article wants to cope with problems associated with optimization of investment, concerning planning of new and utilization of existing electrical distribution infrastructures, motivated by different distributed generation (DG), based on renewable sources, connected to it:

- Photovoltaic (PV) generators.
- Mini hydraulic generators.
- Wind generators.

Most of renewable generators in Spain are medium or small sized, and don't give to the utility detailed generation measures, so suppositions must be done concerning their generation in a given moment, when planning engineers make their network studies and simulations. Those suppositions about how generators behave are well known for conventional generation, usually big and predictable, but not for new renewable generators. Suppositions are in danger of increasing the error committed as renewable generators increase their presence, and hindering planning tasks.

This article focuses on modelling network elements, namely PV generators, studying generation and irradiation historical hourly data, and evaluating their correlation, in order to reduce error committed when generation predictions are done.

II. IMPACT OF RENEWABLE GENERATORS ON DISTRIBUTION NETWORKS PLANNING

Distribution network (DN) is complex. DN is usually radial network. Moreover, it has several possible interconnections between lines in case of contingency or demand variability, so its configuration is dynamically exploited.

It is also the backbone that connects low voltage (LV), where massive number of customers and DG are connected, to substations and high voltage grid, having also a lot of medium size costumers and DG directly connected to it, being possible to dispose of historical stored data of load and generation from utilities.

It is also an important decision to have a probabilistic approach to the problem, because generation uncertainty makes deterministic power flow methodology usually used insufficient, as it gives a static "photograph", usually in two generation and demand sceneries, for the maximum and minimum demand and generation sceneries. In horizontally-operated power systems, stochastic generation may vary its production from minimum to maximum in a very short period of time, and demand can vary also depending on costumer behaviour and other parameters [6-9-10]. There are several models proposed for this approach [7-8], as for example Monte Carlo simulation.

Regarding medium voltage (MV) distribution network, utilities usually know measurements in head of MV lines, the Feeder Measure. But this measure includes distributed energy generation, and it is not possible to know the real demand from loads in the line, the Costumers Demand, or even the accumulated demand in the HV/MV substation, because actual legislation in Spain makes not mandatory for medium and small generators to install detailed registered measures that could give vital information to distributor.

So it is a must for utilities to take in consideration several structured suppositions when planning future demand on MV lines or when calculating if it is possible to connect more customers or distributed generators, or any other calculation needed by the utility. Those suppositions about how generators behave are well known for conventional generation, usually big and predictable, but not for new renewable generators, usually medium or small but present in a great number, spread at the entire electrical network in different voltage levels. Suppositions are in danger of increasing the error committed as renewable generators increase their presence, and hindering planning tasks.

This article proposes a methodology to reduce suppositions, making possible to have a simple way to estimate real generation rate of renewable generators present in an area, in one concrete moment of the past, based in historic data available, and even, based in those data, to predict future situations in the DN.

We can express the problem with a simple equation

$$\text{Feeder Measure (F)} = \text{Costumer Demand (D)} - \text{Dist. Generation (DG)} \quad (1)$$

Costumer Demand (D) can be estimated from equation (1). Feeder Measure (F) is known through measure at feeder head at substation. On the other hand, Distributed Generation (DG) can be obtained from power generation measurements or even from other source.

Following this methodology is applied to photovoltaic generation, studying generation and irradiation historical hourly data, and evaluating their correlation, in order to reduce error committed when generation predictions are done.

III. PHOTOVOLTAIC GENERATION PREDICTION

PV distributed generators are having a major impact in MV and LV networks in Spain, because of great number of installations (Spain was the first photovoltaic market in the world in 2008, with more than 3.000 MW photovoltaic plants installed [11]). It is easily scalable and implemented anywhere because of availability of its primer mover, the sun, so most of the installations connected to DN are medium or small and don't have detailed registered measures for the utility.

Following, PV generation behaviour study is done, based on historical data, on the way to get a model to simulate its behaviour in the past or to make predictions in the future.

Figure 1 shows an example of PV power production in Spain. As this figure depicts, real energy production of a plant of this kind is not accurately simulated by supposing two generation sceneries, peak and valley [12]. There are important generation variations, even on the same day, that should been considerate for a detailed study.

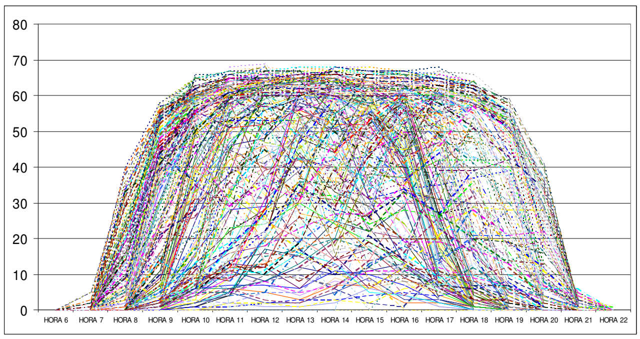

Figure 1. 70 kWp PV power (kW) generation daily cuves along one year.

Therefore, an analysis of solar irradiation and PV active power generation is done, in order to find in first place correlation between irradiation measures, second between PV plants active power generation, and finally, between solar irradiation and PV active power generation. Pearson correlation is done, and later, in order to eliminate scale differences, Spearman's rank correlation has been applied.

In order to have real data, it has been focused the attention in a specific geographic area, Northeast of Spain, where this kind of generators have had a great penetration in distribution network, and are still growing. The area studied has a predominantly rural demand, with long medium voltage lines, but also includes a big city, where there is a high density urban demand example as well as other smaller semi urban demand concentrations.

On this area, it have been solicited and obtained [13] the hourly solar global irradiance historical data for the period 2006-2008, data corresponding to five automatic meteorological stations, located as squares in the map at Figure 2. Data consists on solar global irradiation measured in Watts/m2 over a flat surface in station location. Detailed description of stations is in Table I, and distance between stations in Table II. They form an irregular square with 33 x 34 km sides, enough area to include most of existing medium voltage lines of the region inside, so analysis made could apply to any MV line over a similar or smaller size area.

TABLE I. METEOROLOGICAL STATIONS (MS)

Station Number	Meteorological stations		
	Height over sea level (m)	*Station working from*	*Sample of irradiation measures*
MS1	245	17/10/1990	Hourly
MS2	441	26/03/2001	Hourly
MS3	261	08/08/2000	Hourly
MS4	1574	17/06/2003	Hourly
MS5	374	02/06/2006	Hourly

TABLE II. DISTANCE BETWEEN STATIONS

Distance between stations (km)				
Stations	***MS1***	***MS2***	***MS3***	***MS4***
MS2	29			
MS3	18	33		
MS4	31	40	49	
MS5	25	54	34	38

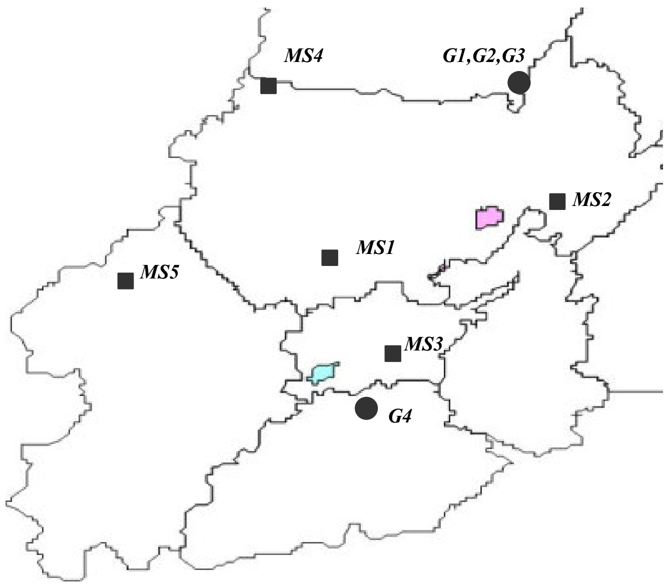

Figure 2. Location map of 5 metereologic stations and PV generators.

It has been also studied telemetric registration of four PV plants located in this area, G1 (100 kW), G2 (50 kW), and G3 (100 kW) located on the same place forming a small solar park in the north of the study area, and a big plant situated in the south, G4 (1.200 kW), separated by 56 km, located as circles in the map at Figure 2. They all where implemented along 2008 during PV market expansion, so we dispose of some moths of measures, depending on plants, on these locations. These measures are taken every 15 minutes, and include separately active and reactive power, generated and consumed.

A. Analisys of irradiation data

The question made here is how different irradiation is, depending on physical location of the plant, to get a better knowledge of variation of the prime mover [12] of the PV generation plant. A sample of it can be seen on Figure 3.

So a Pearson's matrix correlation is done, with all data set 2006-2008, with results seen in Table III. All Pearson p-values have had a result of 0,000. It can be seen that there are very high correlation coefficients between stations, always over 0,93 with the only exception of MS5, witch data set starts on June 2006, but having also a quite high correlation with the rest.

TABLE III. PEARSON CORRELATIONS BETWEEN STATIONS

Meteorological Stations				
Stations	***MS1***	***MS2***	***MS3***	***MS4***
MS2	0,977			
MS3	0,986	0,974		
MS4	0,935	0,938	0,934	
MS5	0,888	0,874	0,883	0,849

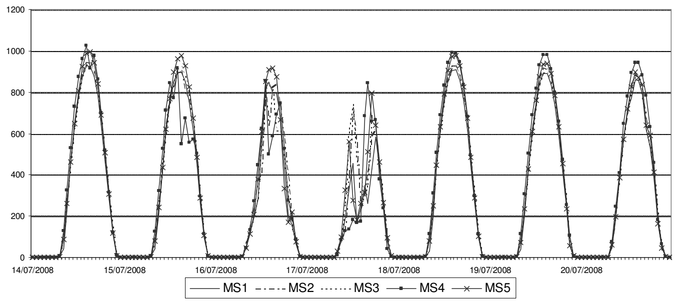

Figure 3. Solar irradiation curves (W/m2) during 2008, week 29.

As said in [12], to get a better correlation, it has also been done a Spearman's ranks correlation matrix, calculating ranks for data. Results can be seen in Table IV, witch don't have great variation with Pearson's, maintaining very high correlation between stations, all of them increasing to be over 0,95 wit the exception of MS5.

TABLE IV. SPEARMAN CORRELATIONS BETWEEN STATIONS

Meteorological Stations				
Stations	***RANKS_MS1***	***RANKS_MS2***	***RANKS_MS3***	***RANKS_MS4***
RANKS_MS2	0,968			
RANKS_MS3	0,972	0,990		
RANKS_MS4	0,956	0,976	0,977	
RANKS_MS5	0,840	0,857	0,860	0,850

To get adjusted correlations between all stations, we have done Spearman's correlation again, eliminating first six months on data set, and results are in Table V.

TABLE V. FINAL SPEARMAN CORRELATIONS BETWEEN STATIONS

Meteorological Stations				
Stations	***RANKS_MS1***	***RANKS_MS2***	***RANKS_MS3***	***RANKS_MS4***
RANKS_MS2	0,969			
RANKS_MS3	0,973	0,989		
RANKS_MS4	0,957	0,976	0,977	
RANKS_MS5	0,970	0,985	0,991	0,979

This very high correlation between irradiation measured by different stations, all over 0,95 (value of 1 would be lineal) is surprising, considering distance and different height from sea level between stations, and is important by itself, as it reveals that during three consecutive years, considering more than one thousand cycles (days), and all weather conditions during the period, measured irradiation has been quite similar in all the stations.

So its possible to deduce that we can know, with a small error, the irradiation in any sunny place in the study area, if we dispose of data from any only station inside the area, in any moment of the year.

B. Analisys of PV power generation data

After that, it has been studied generation data. A sample of it can be seen on Figure 4. The question made in this case is how similar are generation patterns for different plants in the area of study, if they can be considered as independent random variables in a stochastic electrical load flow simulation, or as said in [12], stochastic generators with the same prime movers (sun, wind, etc.) within the same geographical area are not independent.

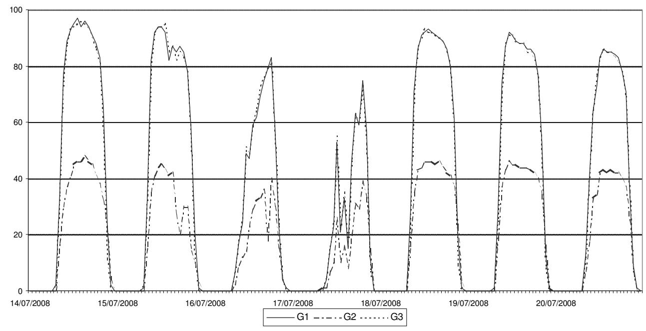

Figure 4. Generation curves (kW) of G1, G2 and G3 during 2008, week 29.

It has been studied four plants, three forming a small park situated on the same location, G1 (100 kW), G2 (50 kW), G3 (100 kW), and a big plant situated in another place, south from them, G4 (1.200 kW) and separated by 56 km. The study is focused on active power generation. These measures are taken every 15 minutes, but an average every four measures was done, to get hourly data that could be compared with irradiation data set. The data set analyzed is the month of July in 2008, where all data sets match.

As can be seen on Table VI, Pearson's matrix correlation for the considered period gives high values, with all Pearson p-values with a result of 0,000. The smaller correlation results belong to G2, and analyzing data after correlation, it has been find that there were five days at the end of the month in witch G2 was stopped, perhaps because maintenance or breakdown, but even that, G2 has high correlation results with the rest of plants.

TABLE VI. PEARSON CORRELATIONS BETWEEN GENERATORS

Photovoltaic generators			
PV	***G1***	***G2***	***G3***
G2	0,831		
G3	0,975	0,809	
G4	0,857	0,685	0,857

And looking at Spearman's correlations results on Table VII, witch eliminate the difference of scale between nominal power of generation, it can be seen that with the exception of G2, all correlations between stations are over 0,91.

TABLE VII. SPEARMAN CORRELATIONS BETWEEN GENERATORS

Photovoltaic generators			
PV	***RANK_G1***	***RANK_G2***	***RANK_G3***
RANK_G2	0,813		
RANK_G3	0,967	0,793	
RANK_G4	0,925	0,751	0,915

So, eliminating those days that G2 was stopped, it can be seen final Spearman's correlations results on Table VIII, with all correlations between stations over 0,90, including G2.

TABLE VIII. FINAL SPEARMAN CORRELATIONS BETWEEN GENERATORS

Photovoltaic generators			
PV	***RANK_G1***	***RANK_G2***	***RANK_G3***
RANK_G2	0,976		
RANK_G3	0,964	0,944	
RANK_G4	0,919	0,915	0,908

As we could see analyzing irradiation data, there are very high correlation factors between PV plants at the study area, so we can say that generation profiles are quite similar for PV plants at the area, and is quite clear that the supposition that PV generators within the same geographical area can be simulated in a stochastic load flow as independent random values is not correct.

C. Generation versus irradiation

Finally, on the way to formulate a model of how photovoltaic generators behave, is important to analyze relation between final power generation and measured solar global irradiation in the area of study, to know if they could give the Y an X axis of a mathematical function for the area.

TABLE IX. GENERATION VS IRRADIATION

Active power generation (kW) versus irradiation (W/m2)				
	RANK_G1	***RANK_G2***	***RANK_G3***	***RANK_G4***
RANKS_MS1	0,926	0,922	0,905	0,966
RANKS_MS2	0,941	0,936	0,919	0,970
RANKS_MS3	0,929	0,925	0,907	0,973
RANKS_MS4	0,950	0,944	0,927	0,951
RANKS_MS5	0,928	0,923	0,906	0,970

As it can be seen on Table IX, Spearman's correlations between all PV generators and stations are all over 0,90. This clearly shows that there is a relationship between them, near to a linear one (that would be a value of 1), and that it can be used solar irradiation as a very good predictor of photovoltaic generation.

As PV generators population in this area is composed by hundreds of medium and small plants, and the only data that the utility disposes is nominal power of installation and point of connection to distribution network, so it was impossible to know the generating profile of them in a discrete moment of the day.

Now, planning engineers making studies in this area can achieve a good prediction of how photovoltaic where generating, only looking at solar irradiation historical data in the area, even using only data set from an only meteorological station. And therefore, is possible to eliminate photovoltaic Distributed Generation variable from equation (1), having a better prediction of Costumer Demand, being possible to reduce calculation's errors and improving infrastructures utilization and investment optimization.

And also, if it is possible to have a on line measure of solar irradiation, or a meteorological prediction for next hours or days, it would be possible to make a good prediction of how hundreds of medium and small distributed photovoltaic generators are generating at the moment, or will generate on next hours or days, been also useful for other utility's department's as dispatching or maintenance, applying in charges balancing or on operations on the DN.

IV. CONCLUSIONS AND FURTHER WORK

It has been introduced how renewable stochastic generators impact in distributed network planning, and described a simple methodology to face that actual legislation in Spain, makes the utility to have a lack of real data from thousands of generating installations connected to its distribution networks, witch don't give detailed registered measures of is production to the utility. Also has been explained how can be used that methodology to reduce error committed when simulating load flows for planning studies.

Later, that methodology has been applied to photovoltaic distributed generators, analyzing a real geographical area in Northeast of Spain, using historical solar global irradiation and PV generation data set from that area, and concluding that meteorological data has a high correlation in the studied area, that photovoltaic generation plants are clearly not independent random variables and have also big correlation factors, and finally, deducing that solar irradiation, even from an only meteorological station in the area, is a very good predictor for photovoltaic power production, being also very useful to know real Costumer Demand in the past, but also, can be used on line for actual estimations or to make future predictions by dispatching or maintenance during their network operations.

Further work has to be done to validate and generalize the given correlation factors with more extended data, in order to obtain a model of photovoltaic generator for stochastic power flows. It would be also interesting to know the geographical limits of the valid correlations from a station determining its boundaries.

V. ACKNOLEDGEMENT

The authors thank Endesa Distribución Eléctrica S.L and the Servei Metereològic de Catalunya for the data used in this research.

REFERENCES

[1] Jim McDonald. “Adaptive intelligent power systems: Active distribution networks” Energy Policy, 36(12):4346 - 4351, 2008. Foresight Sustainable Energy Management and the Built Environment Project.

[2] Goran Strbac. “Electric power systems research on dispersed generation” Electric Power Systems Research, 77(9):1143 - 1147, 2007. Distributed Generation.

[3] J.A. Peças-Lopes, N. Hatziargyriou, J. Mutale, P. Djapic, and N. Jenkins. “Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities” Electric Power Systems Research, 77(9):1189 - 1203, 2007. Distributed Generation.

[4] G. Pepermans, J. Driesen, D. Haeseldonckx, R. Belmans, and W. D'haeseleer. “Distributed generation: definition, benefits and issues”. Energy Policy, 33(6):787 - 798, 2005.

[5] W. El-Khattam and M. M. A. Salama. “Distributed generation technologies, definitions and benefits”. Electric Power Systems Research, 71(2):119 - 128, 2004.

[6] R. Villafáfila, S. Galceran, B. Bak-Jensen, P. Chen, Z. Chen, S. Sörensen, “Probabilistic assessment of wind power production on voltage profile in distribution networks”. Proc. 9th International Conference on Electrical Power Quality and Utilisation EPQU 2007.

[7] P. Chen, Z. Chen and B. Bak-Jensen, “Probabilistic Load Flow: A Review”, DRPT2008 6-9 April 2008 Nanjing China.

[8] R. Allan and R. Billinton. “Probabilistic assessment of power systems”. Proceedings of the IEEE, 88(2):140 -162, Feb 2000.

[9] P. Chen, Z. Chen, B. Bak-Jensen, R. Villafáfila, S. Sörensen, “Study of Power Fluctuation from Dispersed Generations and loads and its impact on a Distribution Network through a probabilistic approach”. Proc. 9th International Conference on Electrical Power Quality and Utilisation EPQU 2007.

[10] B. Bak-Jensen, John Bech, C.G. Bjerregaard, P. R. Jensen, “Models for probabilistic power transmission system reliability calculation”. Journal IEEE_J_PWRS, volume 14, Aug.1999, pp 1166—1171.

[11] Comisión nacional de la energía (CNE), 2009. Available on line at www.cne.es

[12] G.Papaefthymiou, “Integration of Stochastic Generation in Power Systems”. PhD Tesis, Delft University of Technology, 2007.

[13] Servei Metereològic de Catalunya. Xarxa d’Estacions Meteorològiques Automàtiques. Available on line at www.meteocat.com