Elena SÁIZ-MARÍN[1], Enrique LOBATO[1]

IIT, ICAI School of Engineering, Universidad Pontificia Comillas (1)

Optimal voltage control by wind farms in distribution networks using regression techniques

***Abstract**. Due to the growth of the wind power penetration, the voltage control by this technology must be evaluated. The aim of this control is the minimization of the distribution losses while keeping away the system from voltage collapse. This paper investigates if it is possible to assess accurately the optimal voltage control of wind farms using simple regression functions instead of running an Optimal Power Flow (OPF) each time. The methodology proposed is illustrated considering an actual distribution network comprising 13 wind farms of the Spanish power system.*

Streszczenie** W artykule zbadano możliwość kontroli napięcia w farmie wiatrowej na podstawie funkcji regresji. Metodę zbadano na przykładzie sieci dystrybucyjnej 13 farm wiatrowych. (**Sterowanie optymalną wartością napięcia w sieci dystrybucyjnej farm wiatrowych przy wykorzystaniu funkcji regresji)

Keywords: Artificial Intelligence, Voltage Control, Wind Power.
Słowa kluczowe: sztuczna inteligencja, funkcja regresji, farmy wiatrowe.

Introduction

Wind power has experienced a wide development throughout the world due to technological advances in wind turbines and favourable policy incentives. Spain is the fourth largest country in wind power installed capacity, with 19813 MW at the end of year 2010 [1]. The Spanish Ministry of Industry, Tourism and Trade considers as a probable scenario 29000 MW of installed wind power capacity for the year 2016 [2]. The wind production has reached 16000 MW and the penetration level has been more than 50% [3].

Within this worldwide and national framework, wind power poses increasing challenges to the planning and operation of power systems. For maintaining the stability of a power system the control of frequency and voltage are required. Currently, since the wind power technology does not provide these services the Transmission System Operators have limited the penetration of wind energy into the grid. However, current technology developments enable the design of operation and control strategies of wind turbines to provide such grid services. In fact, even though currently the wind generation does not provide these services, REE, the operator of the Spanish system, is developing new operational procedures that consider the provision of frequency and voltage by these technologies. Consequently, the provision of them by wind power must be evaluated.

Nowadays, technology developments enable the design of operation and control strategies of wind turbines to provide frequency and voltage control. The most common technology is the doubly fed induction generator (DFIG). In DFIG the stator winding of the generator is coupled to the grid whereas the rotor winding is coupled to a power electronic converter. In this way the electrical and mechanical frequency are decoupled, because the power electronic convertor compensates the difference between mechanical and electrical frequency by injecting a rotor current with variable frequency. Consequently variable-speed operation becomes possible [4]. One important advantage of this technology is the ability to provide reactive power control without installing additional capacity support [5].

Assuming that wind generators are able to provide voltage control, this paper will focus on the optimal voltage control of the distribution network which connects the different wind farms of an area to the transmission network bus. The aim of this control is the minimization of the distribution losses and keeping away the system from voltage collapse. For controlling the voltage profile by wind power there are two possible strategies:

- Set point of voltage magnitude, in this case the reactive power is controlled to achieve a specific voltage value
- Set point of constant power factor, in this case the reactive power is controlled to achieve a specific power factor

For computing the optimal adjustment of the reactive power and voltage set point of each wind farm there are various possibilities. The most simple and usual one is using a commercial OPF, as for example the PSS/E program [6, 7, 8, 9]. This software determines the settings of a selected group of control variables to achieve an optimal steady-state operation of a power system. However, several drawbacks can be identified. Firstly, it needs a license. Secondly, it might take excessive computational time if the OPF is to be run on real-time.

This paper investigates if it is possible to assess accurately enough the optimal voltage control of wind farms using simple regression functions instead of running an Optimal Power Flow (OPF) each time. It is proposed to evaluate the relation between selected explanatory variables and the objective function. In order to perform this relation three steps are needed: (a) a data base of optimal operational scenarios built using a genetic algorithm, (b) selection of explanatory variables, and (c) developing of regression rules of the optimal voltage control.

In order to build a data base representing all the possible scenarios, random power generation for the wind farms is considered. Although a scenario in which all the wind farms have the same percentage of utilization (Actual power generation / Maximum power generation) is more realistic, due to the fact that they are located in the same area, the random power generation is considered for giving robustness to the study. For each random power generation the case is converged for different voltage set points of the bus which connects the distribution network to the transmission system and wind farms are optimally adjusted using genetic algorithms [10, 11, 12, 13].

The explanatory variables that are selected in this paper are the losses from each wind farm to the transmission network bus. This variables will be used with the intention of generalized the study for other areas. In this study transformer taps, shunts reactors and capacitors are not taking into account. Nevertheless, it will be analysed in future studies.

Finally, regression rules are derived for the optimal voltage set points of each wind farm. The methodology proposed in this paper is illustrated considering an actual distribution network comprising 13 wind farms of the Spanish power system.

The paper is organized as follows. In section II the methodology is explained. The results obtained are illustrated in section III. Finally, the conclusions are drawn in section IV

Methodology

The methodology is composed of three steps (a) a data base of optimal operational scenarios built using a genetic algorithm, (b) selection of explanatory variables, and (c) developing of regression rules of the optimal voltage control. This methodology will be applied to an actual distribution network comprising 13 wind farms of the Spanish power system.

A. Optimal operational scenarios building

The scenarios are created giving random values to each wind farm active power production. These values are obtained multiplying a random number (between zero and one) generated for each wind farm by its active power capacity. Moreover, nine voltage set points of the transmission network bus are considered (the lowest voltage considered is 0.97, incremented in steps of 0.01 until achieving the maximum value of 1.05). Each scenario is converged and the distribution losses computed. A total number of 56 cases of wind power production, each one with 9 different voltage set-points of the transmission network bus, sum up to a total of 504 scenarios to be optimized.

For each converged scenario, a genetic algorithm is run to adjust the control variables of the wind farm in order to minimize distribution losses. The paper considers two possible control variables, the reactive power or the voltage set point of each wind farm. The application example that will be presented will outline the advantages and disadvantages of each possibility. Transformer taps, shunts reactors and capacitors are not taking into account.

The population is formed by M individuals. The number of individuals is heuristically determined. Each individual has N chromosome, one for each wind farm. Each chromosome represents the control variable of the corresponding wind farm, either the voltage set point or the reactive production of the wind farm. According to this fact, each individual represents a possible solution of the voltage control. The population is randomly initialized between the limits of the variable of control.

After creating a random initial population of M individuals the structure of each iteration (generation) of the genetic algorithm is the following: (a) selection of M individuals that go to the matting pool, (b) crossover of couples of individuals of the matting pool, (c) mutation of individuals with the intention of exploring new areas trying to do not become stuck at a local optimum.

The idea of the selection in a genetic algorithm is that the best individuals have more possibilities to be selected. In this paper the best individual is always selected into the matting pool and the rest (M-1), according to the roulette wheel. For constructing the roulette wheel firstly the probability of each individual is calculated according to the following formula:

(1) $$P_i = (1/F_i^2)/\sum_j^n(1/\ F_j^2\ )$$

, where Pi is the probability of each individual and F_i is the objective function of the individual which is also named Fitness. The exponent is used for prioritizing the selection of the best individuals, which are those with the lowest fitness. Once that the probabilities are calculated they are sorted and the cumulative probability is obtained. Then (M-1) random numbers are generated and the (M-1) individuals are selected.

The idea of the crossover is generating individuals that are similar to their parents. First two parents are determined randomly in between the selected population. Another number is generated randomly and if it is less or equal than the crossover probability, the parents are crossed obtaining two sons. Otherwise the parents remain unaltered within the iteration. The two sons are obtained according to the following formulas:

(2) $$X_{iSon1}=0.5X_{iParent1}+0.5X_{iParent2}$$

(3) $$X_{iSon2}=1.5X_{iParent1} - 0.5X_{iParent2}$$

where X_i corresponds to chromosome i of the individual.

The mutation is done for introducing new “genetic material”. First one individual is chosen randomly in between the resulting crossed population within the iteration. A random number is generated and if this number is lower or equal than the mutation probability the individual is mutated. Otherwise the selected individual remains unaltered within the iteration. The chromosome of the individual to be mutated is chosen with another random number between zero and the number of chromosomes N. For the problem under study of this paper, it has been tested that in order to avoid being stuck at a local optimum at first, it is interesting to use a high mutation probability in order to explore all the areas. However at the end the mutation probability must be small in order to do not make the algorithm instable. In this paper a lineally decreasing mutation probability within each iteration is employed to enhance the performance of the genetic algorithm.

This process must be repeated the necessary generations for achieving convergence. The optimal generation in which the algorithm should be interrupted is the one that corresponds to the bend point of the evolution of the mean Fitness. The data base is built applying the genetic algorithm in order to obtain the optimal power flow of each scenario.

B. Selection of explanatory variables

The contribution of each wind farm to the distribution losses depend on both the electrical distance of the wind farm to the transmission grid and the load factor of each line. Thus, the explanatory variables that are selected in this paper correspond to the losses from each wind farm to the transmission network bus P_{lossi}, taking into account both the impedance of the distribution network and the current that go through the network.

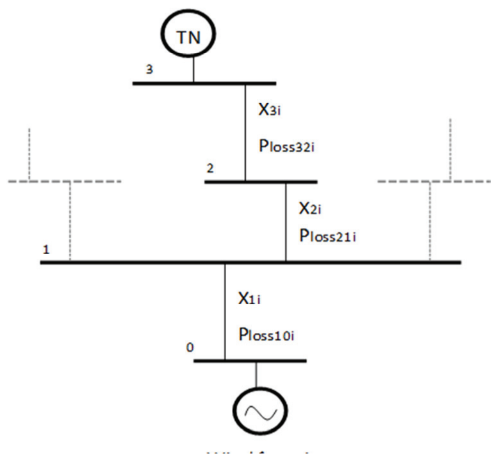

Fig. 1. Diagram of the distribution network

Consequently, the losses from wind farm i to the transmission network bus (TN), will be computed as follows:

(4) $P_{lossi}= P_{loss10i} + P_{loss21i} + P_{loss32i}$

The explanatory variable considered a radial configuration of the distribution network. Future research will consider which other variables must be taken into account if the network is not radial.

C. Developing of regression rules

Finally, regression rules are derived for the optimal reactive production of each wind farm. Reactive production Q_i is approximated as a second grade polynomial depending on the active power loss of the wind farm Plossi:

(5) $Q_i=c_0+c_1 P_{lossi}+c_2 P^2_{lossi}$

where c_0, c_1, and c_2 are the polynomial coefficients.
This paper will assess whether these regression coefficients are independent of both the voltage set point of the transmission network bus and of the topology of the distribution network containing the wind farms under study, or, on the contrary, they have to be computed and adjusted for each new case to reach admissible precision.

Results

The methodology proposed in this paper is illustrated considering an actual distribution network comprising 13 wind farms of the Spanish power system. This section will be structured as follows: (a) generation of scenarios and optimization of them with the application of the genetic algorithm, (b) computation of the explanatory variables and (c) analysis of the relation between the objective function and the control variables. In addition a possible control strategy without running each time an OPF will be proposed.

A. Optimal operational scenarios building using genetic algorithms

For each of the operational scenarios, each individual has thirteen chromosomes (one for each wind farm). Each chromosome represents the control variable. The paper will analyze the advantages and disadvantages of using as control variable the reactive power or the voltage set point of each wind farm.

a.1 Selection of the wind farm control variable

In case of using reactive power as control variable, wind farm nodes are represented as PQ type. Thus, final bus voltages V_{bus} may fall outside their limits.

(6) $0.95 \leq V_{bus} \leq 1.05$

In order to remove individuals containing wind farms with bus voltages outside their limits, there exist two possibilities:

- Eliminate the infeasible individuals. Nevertheless, since this situation occurs quite frequently in the case study, in order to achieve a population representative with enough feasible individuals, much time is being required.
- Introduce a penalization factor in the objective function. This is the method that has been followed in the paper. However, introducing a penalization factor reduces the voltage violation magnitude, but does not guarantee that all voltages fall within their limits.

Taking into account the overall penalization factor the objective function (fitness) is computed as:

(7) $F_i=P_{losses}+K(G/G_t)(V_{bus}-V_{lim})2$

Where P_{losses} correspond to the active losses of the distribution network and K is the penalization factor that has been heuristically determined. A penalization factor that changes in each generation, (G actual generation, G_t total number of generations) proposed by [13] has been employed in (7). As mentioned before, this penalization factor does not imply that the algorithm provides a local optimum that respects bus voltage restrictions. Consequently, this paper will use as control variables the voltage set points of each wind farm.

In case of using voltage set points as control variable, wind farm nodes are represented as PV type. In this case final reactive outputs of wind farms or bus voltages within PQ nodes of the distribution network may fall outside their limits. Within the distribution network under study, this situation is not usual. In this case, the affected individual is not considered.

a. 2. Genetic algorithm

The algorithm has been developed using python [14]. Each individual is converged using the PSS/E program [15]. The parameters have been heuristically determined and are the following ones:

- Individuals M =100
- Crossover probability=0.75

As explained in subsection II.A in this paper a lineally decreasing mutation probability within each iteration is employed to enhance the performance of the genetic algorithm. Fig 2 shows the mutation probability as a function of the number of generations.

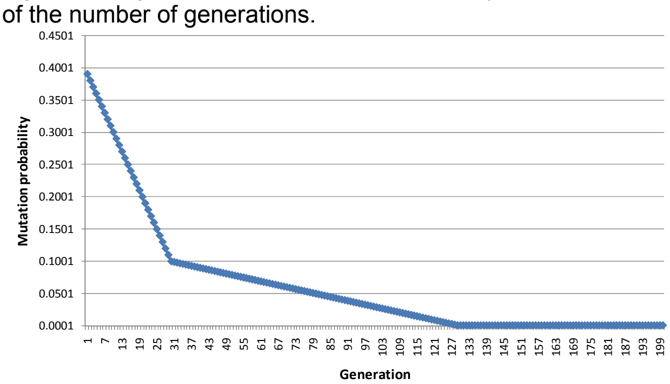

Fig. 2. Mutation probability as a function of the number of generations.

Fig 3 presents the evolution of the mean, maximum and minimum of the fitness within the population of M individuals of one of the randomly generated scenarios. Fig 3 shows that the mean fitness decrease rapidly in the first generations and does not change significantly near convergence.

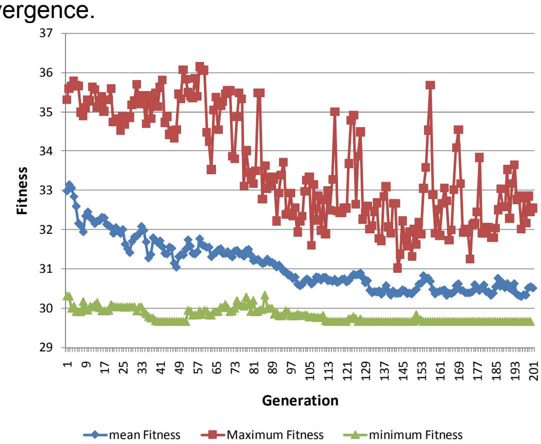

Fig. 3. Evolution of the fitness (mean, maximum and minimum) of one of the randomly generated scenarios.

In order to determine the adequate number of generations, the optimum by the OPF of the PSS/E program has been compared with the mean fitness of the genetic algorithm within each generation. Fig 4 depicts the quadratic mean error between the optimum given by PSS/E and the mean fitness of the genetic algorithm in each generation. A good compromise between computation time and accuracy of the algorithm corresponds to a number of generations located at the bend point of the curve. Thus, in order to optimize each scenario, the number of generations within the genetic algorithm has been set to G_l=100. However, due to the optimum is unknown, the criteria of convergence is established when the mean fitness represented in Fig 3 stabilizes along the number of generations.

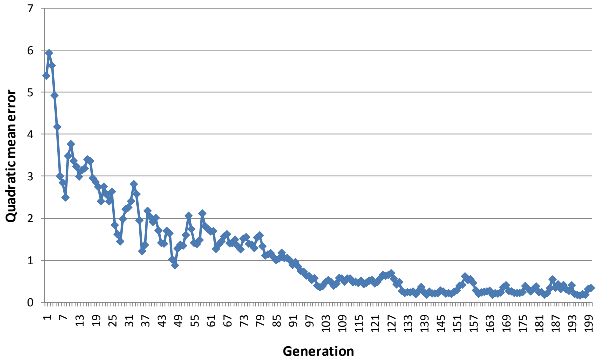

Fig. 4. Evolution of the quadratic mean error between the optimum given by PSS/E and the mean fitness of the genetic algorithm

In addition, for ensuring that the algorithm has converged, it is important to analyze the homogeneity of the total population. Fig 5 contains the histogram representing the fitness of individuals within the population in the initial and final generation (G_l=100). Convergence implies that the fitness variability is small.

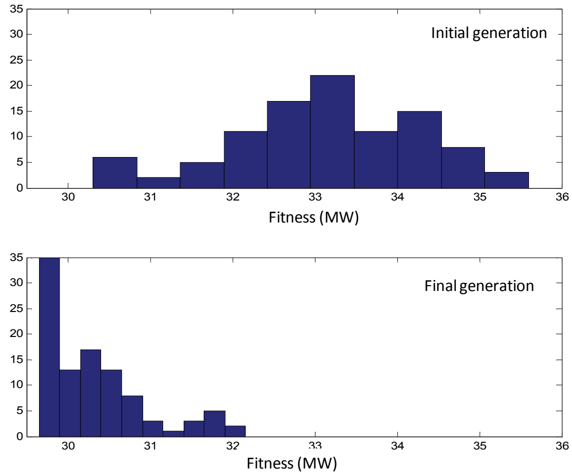

Fig. 5. Histogram representing the fitness of individuals within the population in the initial and final generation (G_l=100).

B. Explanatory variables computation

Once that the genetic algorithm has converged the reactive generation of the wind farm is obtained. With the intention of studying if the control proposed for this area can be generalized for any other distribution network containing wind farms, the losses from each wind farm to the transmission network bus explained in subsection III.B was computed. They are selected because takes into account the impedance of the distribution network and the current that go through the network.

In order to show that the losses from each wind farm to the transmission network bus represents an adequate explanatory variable the results of the 13 wind farms of one of the optimized scenarios is considered in Fig 6 . It shows how the correlation factor R^2 increases significantly when the explanatory variable considered is the active power losses from the wind farm to the transmission network bus instead of using the impedance. Consequently, these variables are representative of the reactive power assignment of the different wind farms.

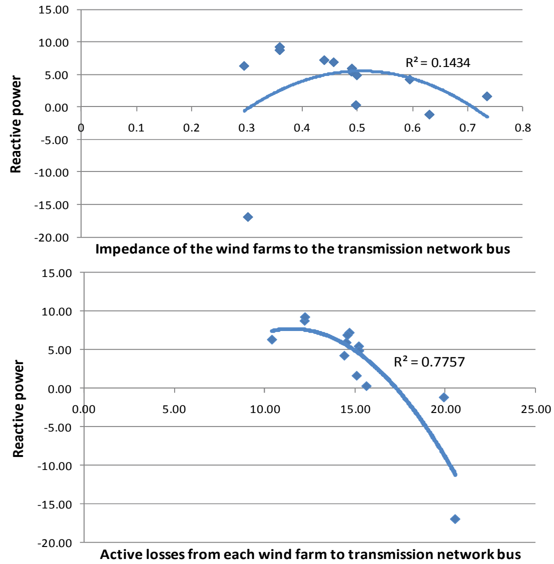

Fig. 6. Relation between the reactive power of the wind farms and the impedance and the active losses of the wind farms to the transmission network bus respectively.

C. Voltage power control using regression rules

This paper investigates if it is possible to assess accurately enough the optimal voltage control of wind farms using simple regression functions instead of running an Optimal Power Flow (OPF) each time. The regression rules have been obtained using the 6552 pairs ($Qi,Plossi$) -56 active scenarios x 9 transmission network voltages x 13 wind farms- using WEKA© [16] and MATLAB©. The overall quadratic regression function of the reactive output of the wind farms as a function of the active losses using the total 6552 pairs of data presents a poor correlation coefficient. Thus, the relationship should be dependent of the voltage of the transmission bus and/or the topology of the distribution network containing the wind farms under study. Fig 7 presents the quadratic regression function of the reactive output of one wind farm as a function of the active losses. It can be appreciated how the correlation coefficient increases significantly respect the previous situation.

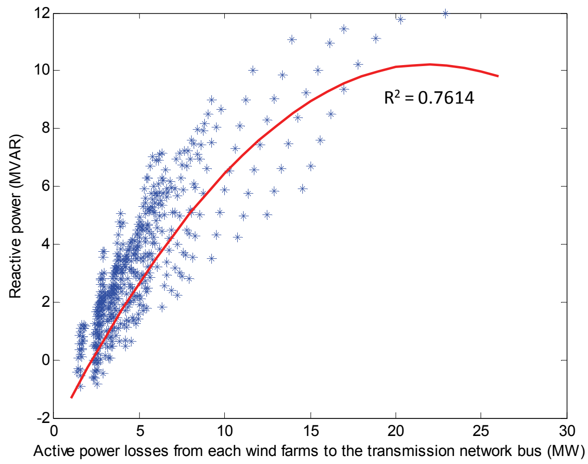

Fig. 7. Relation between the losses from one wind farm to the transmission network node and the reactive power of tone wind farm

In addition Fig 8 presents the quadratic regression functions for three different voltage set points of the transmission network bus of one of the wind farms of the distribution network under study - 56 pairs (Q_i,P_{lossi}) for each voltage set-point. Correlations coefficients increases to values of R^2 = 0.9611, 0.9546, 0.9149 demonstrating that the relationship depends of the voltage and thus, one different regression should be obtained for each set point.

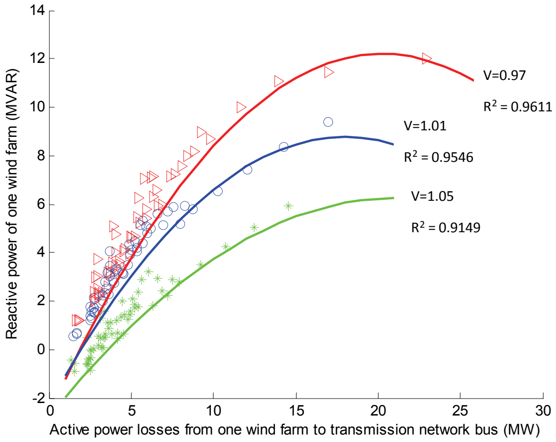

Fig. 8. Relation between the losses and the reactive output of one of the wind farms of the distribution network.

The losses from the wind farm to the slack node were proposed with the intention of generalizing the study to other areas. Nevertheless, although the methodology can be applied to other networks, the regression rules must be built due to the no linearity of the problem and the different reactive capacities of each wind farm

The control proposed using the relations identified is presented in Fig 9.

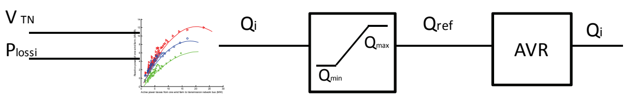

Fig. 9. Control proposed using quadratic regression function for each wind farm and each voltage set-point of the transmission network bus

Finally, in order to validate if the optimal voltage control proposed in this paper using regression functions (one different quadratic function for each farm and voltage set point), the distribution losses obtained by the methodology proposed in this paper is compared with the optimal losses provided by the OPF of the PSS/E package. For this purpose, a different validation set of 10 random scenarios has been formed. Fig. 10 shows how the difference between the active losses obtained by the OPF of the PSS/E program and the estimation method is less than the tolerance of the OPF (0.1 MW)

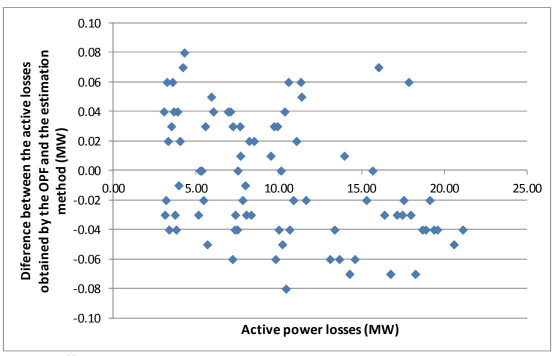

Fig. 10. Difference between the active losses obtained with the OPF and the estimation method.

Conclusions

A possible voltage control without using an OPF is proposed. This control has been obtained using simple regression rules on a data base that was built with a genetic algorithm. Using these simple rules is not necessary to run each time an OPF. Nevertheless, due to the no linearity of the problem and the different reactive capacities of each wind farm, this study cannot be generalized for other areas.

REFERENCES

[1] Global Wind Energy Council, "Global Wind Energy 2010 Report," Available: www. gwec.net.

[2] Ministerio de Industria, Turismo y Comercio, "Planificación de los sectores de electricidad y gas, 2008-2016. Desarrollo de las redes de transporte." Madrid, 2008.

[3] REE. Estructura de la generación en tiempo real. Available: https://demanda.ree.es/generacion_acumulada.html.

[4] Thomas Ackermann, *Wind Power in Power System.* John Wiley & Sons, Ldt, July 2005.

[5] E. Vittal, M. O'Malley and A. Keane, "A Steady-State Voltage Stability Analysis of Power Systems With High Penetrations of Wind," *Power Systems, IEEE Transactions on,* vol. 25, pp. 433-442, 2010.

[6] R. Keypour, A. A. Rashidi and M. Jazaeri, "Probabilistic voltage stability constrained optimal power flow considering wind farms," in *2010 IEEE Region 8 International Conference on 'Computational Technologies in Electrical and Electronics Engineering' (SIBIRCON 2010),* 2010.

[7] R. A. Jabr and B. C. Pal, "Intermittent wind generation in optimal power flow dispatching," *Generation, Transmission & Distribution, IET,* vol. 3, pp. 66-74, 2009.

[8] C. F. Moyano and J. A. P. Lopes, "Using an OPF like approach to define the operational strategy of a wind park under a system operator control," in *Power Tech, 2007 IEEE Lausanne,* 2007, pp. 651-656.

[9] J. Qiao, Y. Min and Z. Lu, "Optimal reactive power flow in wind generation integrated power system," in *2006 International Conference on Power System Technology,* 2006.

[10] S. R. Paranjothi and K. Anburaja, "Optimal Power Flow Using Refined Genetic Algorithm," *Electric Power Components and Systems,* vol. 30, pp. 1055-1063, 10/01, 2002.

[11] S. C. Liu, J. H. Zhang, Z. Q. Liu and H. Q. Wang, "Reactive power optimization and voltage control using an improved genetic algorithm," in *Power System Technology (POWERCON), 2010 International Conference on,* 2010, pp. 1-5.

[12] Li Ling, Zeng Xiang Jun and Zhang Ping, "Wind farms reactive power optimization using Genetic/Tabu hybrid algorithm," in *Intelligent Computation Technology and Automation (ICICTA), 2008 International Conference on,* 2008, pp. 1272-1276.

[13] Ying-Yi Hong and Yi-Feng Luo, "Reactive power control in distribution system using genetic algorithms," in *Intelligent Systems Applications to Power Systems, 2007. ISAP 2007. International Conference on,* 2007, pp. 1-6.

[14] "Python," Available: www.python.org.

[15] Siemens. PSS/E. *32* Available: www.energy.siemens.com.

[16] The University of Waikako, "WEKA," Available: www.cs.waikato.ac.nz/ml/weka/.

***Authors**: Elena Sáiz-Marín, Instituto de Investigación Tecnológica (IIT), ICAI School of Engineering, Universidad Pontificia Comillas, C/Alberto Aguilera, 23, 28015 Madrid, Spain, E-mail: Elena.Saiz@iit.upcomillas.es; dr Enrique Lobato, , Instituto de Investigación Tecnológica (IIT), ICAI School of Engineering, Universidad Pontificia Comillas, C/Alberto Aguilera, 23, 28015 Madrid, Spain, E-mail: enrique.lobato@iit.upcomillas.es*