

# Probabilistic Clustering of Wind Energy Conversion Systems Using Classification Models

Paula Odete Fernandes<sup>1,2</sup> (✉) and Ângela Paula Ferreira<sup>1,3</sup>

<sup>1</sup> Polytechnic Institute of Bragança, Campus de Santa Apolónia,  
Apartado 1134 5301-857, Bragança, Portugal  
{pof, apf}@ipb.pt

<sup>2</sup> UNIAG and NECE, Covilhã, Portugal

<sup>3</sup> CISE - Electromechatronic Systems Research Centre,  
University of Beira Interior, Covilhã, Portugal

**Abstract.** This research intends to give insights on the pattern aggregation of wind energy conversion systems technologies through identification of homogeneous groups within a set of wind farms installed in Portugal. Pattern aggregation is performed using Hierarchical Cluster Analysis followed by Discriminant Analysis, in order to validate the results produced by the first one. The clustering support matrix uses three independent variables: installed capacity, net production and capacity factor, in a per year basis. Cluster labelling allows the identification of two homogenous groups of wind farms, whose main attributes are based on the technological conversion system trend: (1) asynchronous generator based technology and (2) direct driven synchronous generator based technology, with higher capacity factors.

**Keywords:** Wind farms · Wind turbine generators · Cluster analysis · Discriminant analysis

## 1 Introduction

The integrated use of renewable energy sources and mature technologies in power systems contributes to strengthen domestic economies, by reducing dependency on imported fossil fuels and dramatically reduce greenhouse-gas emissions, over time.

In this scenario, the exploitation of wind energy resources play an important key role, capable of decarbonising the power sector. Wind energy is a clean source and enfolds an environmental friendly technology. Its renewable character and the fact it does not pollute during the operational phase makes it one of the most promising energy source in reducing environmental problems at both global and local levels.

The technological maturity of wind energy conversion systems contributes to a small economic differential cost with regard to conventional technologies. In fact, wind energy systems are more competitive than other renewable energies, apart from for hydro energy systems. This scenario suggests that, in the following decades, wind energy will remain the main commitment to new electrical generation capacity.

According to the Global Wind Energy Council (GWEC), installed wind power capacity has grown to cumulative worldwide installation level of 369,5 GW, with 51,5 GW alone installed in 2014. Europe's total installed capacity has been surpassed for the first time by Asian markets (India, China). Nevertheless, by the end of 2014, the wind power capacity installed would produce energy enough to cover approximately 10% of the Europe's electricity consumption, in a normal wind year [1]. Portugal accounts for about 5% of the wind energy capacity installed in European Union, with approximately 4,9 GW of accumulated installed capacity in 2014 which is capable to generate about 15% of the electrical energy consumption [2].

Despite the advantages inherent to a renewable energy source, wind energy has also some unfavourable conditions. The small power density of the wind leads to wide and material extensive turbines, thereby hindering the on-site assembly and the electrical infrastructure. Technological developments are being made in order to increase unit power wind generators to optimize the impact on the ground for onshore wind farms or for offshore applications [3].

Concerning the prime source, wind is stochastic in nature and essentially ruled by random meteorological changes. Due to its intermittent and unpredictable behaviour, wind energy conversion systems are not dispatchable, i.e., they do not have the ability to produce electrical energy following load requirements. The inherent variability of wind power is also raising concerns regarding the reliability and cost-effectiveness of the transmission and distribution power systems while supporting large wind farms [4]. High penetration levels of wind energy implies structural changes in power systems as, for instance, the usage of storage systems and/or coupling hydro and wind systems to smooth the output pattern [5]. Progress and improvements are being made in renewable energy integration into power systems. Development of energy storage provided, for instance, by bi-directional Vehicle-to-Grid technologies and intelligent networking will allow a greater penetration potential of wind energy.

The life span of the first wind farms is coming to an end, which implies repowering processes aiming an augmented efficiency and reliability of the wind turbines. From the available technologies of the conversion systems, it is far from clear which of them is the optimal.

The motivation for this work is supported by the absence of a deterministic certainty in allocating outputs of wind farms, regarding the technological conversion system trends [6,7]. Therefore, this study aims at giving some insights into technological approaches for wind turbines, using probabilistic clustering to identify homogeneous groups.

Clustering wind farms allocate different units into a group which contains some common characteristics, which may be used to reduce the size and the order of mathematical models and also to perform pattern classification into extensive multi-dimensional data set [8-11].

Previous studies had been performed on pattern aggregation of wind farms using probabilistic clustering [12,13]. This work addresses the problem through a methodology based on two multivariate analyses: Hierarchical Cluster Analysis and Discriminant Analysis. In order to identify the clusters characteristics, it is also performed an exploratory descriptive analysis. In addition to previous publications, this

work considers an increased reference data set of wind farms and extended time span in order to establish a comparative analysis with previous results.

The paper is organized as follows: next section presents technological trends on wind energy conversion systems, Section 3 overviews the clustering and validation methodologies, Section 4 applies the proposed approach to a case study, presents the main results and discussion and, finally, Section 5 rounds up the paper with the main conclusions.

## 2 Wind Energy Conversion System Technologies

Despite the fact first development of commercial wind energy technology began in the late 1930s, only after the oil crises of the 1970s, there have begun economical incentives to develop the technology further [14]. Since the 1980s, there has been a significant consolidation of the design of wind turbines. This section describes main design styles in wind energy conversions systems and points out the technological trends of the conversion system drive train.

### 2.1 Design Styles

The mainstream commercial market uses horizontal axis wind turbines, meaning the rotating axis is parallel to the ground. This option is inherently more efficient than vertical axis. Concerning the number of blades, the aerodynamic efficiency and reduced acoustic noise emission establish three-bladed rotor design.

Other important issue related with the design of a wind energy conversion systems is the mean of limiting rotor power in high operational wind speeds. There are two main approaches: stall and pitch control.

In stall regulated machines, speed regulation is intrinsic to the aerodynamic design, without any change of the rotor geometry. Under this control approach, wind turbine runs at approximately constant speed even when the wind speed is high, without producing excessive power. The constant speed is achieved through the connection of the electric generator to the grid. Regarding this aspect, the grid behaves like a large fly-wheel, holding the speed of the turbine nearly constant irrespective of changes in wind speed.

Pitch control involves pitching the blades (i.e., turning the wind blades about their main axis) in order to regulate the power the rotor extracts from wind. This control involves an active control system, which should sense the blade position and defines appropriate changes of blade pitch, according to the measured output power.

Another important and decisive design issue of the wind turbines is the use of variable rotational speed versus fixed speed, with consequences on the overall performance of the system [15,16].

The constant speed turbine designs consist on generators operating at fixed speed when producing power, directly connected to the utility grid which, through the generator, holds the speed constant. This concept makes use of Squirrel Cage Induction Generators (SCIG) with a geared drive train to adapt the rotational speed to the fre-

quency of the grid. With this design, the wind energy capture and also the power quality in the utility grid are reduced.

Variable speed wind energy systems allows operation below rated power, enabling increased energy capture, and also above rated power, even over a small speed range, which can substantially ease pitch system duty and reduce output power variability. This exploitation mode of wind energy systems improves the power quality when compared with constant speed systems. Variable speed wind energy systems may be implemented using synchronous or asynchronous generators, allowing wider or narrower wind speed ranges, respectively.

Solutions based on asynchronous generators, the so called Doubly Fed Induction Generators (DFIG), with the stator windings directly connected to the grid and a partial scaled electronic converter between the rotor and the grid, allow a low to moderate variation of the rotor speed. Since the power converter is partially scaled, typically one third of the rated power of the system [17], this solution is somewhat cost effective but, on the other hand, there are limitations to control effectively the grid variables, which translates in a deficient quality power system [18]. It should be pointed out that this concept uses a geared drive train to match the low rotational speed promoted by wind velocities to the higher efficient rotational speed of this generator type.

Solutions based on Synchronous Generators (SG) use full scaled electronic converters. The electrical energy is generated at variable frequency (strictly related to the rotational speed of the rotor) and then converted to the frequency of the grid. This concept takes advantage of the wide speed range operation allowed by the full scale converter between the generator and the grid, which also allows boosting the grid stability and performance. Additionally and when compared with DFIG, this type of generators requires lower ratio gearboxes, or even its omission, which translates in higher reliability and lower maintenance costs [18]. Gearboxes are one of the most expensive components of the wind turbine system and require significant repair or overhaul before the intended life span of the entire system is reached. Thus, the simplification introduced in the drive train by the absence of these components improves significantly the reliability of the wind turbine system and helps bring the cost of wind energy back to a decreasing trajectory [19].

## 2.2 Technological Trends

Under the premise of high variability and intermittency of wind speed, the actual demand on power quality issues compels for generators featuring variable speed, which is the dominant trend in the actual market.

Comparing partial speed range systems, promoted by DFIG, and full-range variable speed drives based on SG, the later bring some attractions, specially on operational flexibility and power quality issues, but also have some drawbacks related with the higher power of the electronic converter, with the same rating of the generator [20]. In fact, there was never a clear case for full variable speed range on economic grounds, with small energy gains being offset by extra costs and also additional losses in the power converter.

Another technological trend is related with direct driven generators, i.e., gearless systems. The direct drive systems of Enercon [21] are long established, and gearless systems or with low ratio gearboxes, using Synchronous Permanent Magnet Generator (SPMG) technology have emerged in recent years [22]. In fact, some manufacturers that in past had based their technology on asynchronous generators are now moving to SPMG with full scaled converter [23].

Permanent magnet technology allows a higher power-to-volume-ratio, and fully rated power converter based systems can be applied without design hardware modifications in both 50 Hz or 60 Hz power systems, which increases flexibility for international developers operating in multiple wind markets [14].

Concerning the power control in high operational wind speeds, the design issues of pitch versus stall and degree of rotor speed variation are evidently connected. The stall-regulated design remains viable, but pitch control offers potentially better output power quality, while overall costs of both systems remain similar [14].

### 3 Clustering and Validation Methodology

Clustering, i.e., partitioning objects/cases into similar groups, is a problem with several alternatives in mathematics as well as in applied sciences. Cluster analysis aims at recognizing groups of similar records and, therefore, helps to discover distribution of patterns and interesting correlations in data sets [24, 25].

To attain the main goal of this research study, a hierarchical cluster analysis has been applied, using the methodology proposed by Ward [24], which is the most commonly used for problems similar to the one under analysis, providing a more consistent solution and it is also recommended for quantitative variables measured on a ratio scale. In this methodology, an objective function, defined as the sum of squares of deviations of the individual observations compared with the average of the group, is minimized, aiming at creating groups which have maximum internal cohesion and maximum separate external distance [24]. This method uses the variance to evaluate distances between clusters, which results in an efficient approach when compared with other hierarchical methods (for instance, nearest neighbour, furthest neighbour and median clustering).

The Ward’s distance between clusters  $C_i$  and  $C_j$ ,  $D_w(C_i, C_j)$ , is the difference between the total within cluster sum of squares for the two clusters separately, and within cluster sum of squares, which results from merging the two clusters in cluster  $C_{ij}$  [24, 25], i.e.,

$$D_w(C_i, C_j) = \sum_{x \in C_i} (x - r_i)^2 + \sum_{x \in C_j} (x - r_j)^2 - \sum_{x \in C_{ij}} (x - r_{ij})^2 \tag{1}$$

where  $r_i$  is the centroid of  $C_i$ ,  $r_j$  is the centroid of  $C_j$  and  $r_{ij}$  is the centroid of  $C_{ij}$ .

To implement a dissimilarity measure between subjects, it is selected the Euclidean Distance Squared. The distance is defined as the square root of the sum of the squared differences between the values of  $i$  and  $j$  for all the selected variables  $(1, 2, \dots, p)$ , [26]:

$$D_{ij} = \sqrt{\sum_{k=1}^p |x_{ik} - x_{jk}|^2} \quad (2)$$

where  $x_{ik}$  is the value of the variable  $k$  for cases  $i$  and  $x_{jk}$  is the value of the variable  $k$  for cases  $j$ .

An agglomerative algorithm is used in order to produce a sequence of clustering schemes of decreasing number of clusters at each step. The clustering scheme produced at each step results from the previous one by merging the two closest clusters into one.

One of the most important issue in clustering analysis is the evaluation of the produced results to find the partitioning which best performs the underlying data. To accomplish this requirement, the performed methodology applies a Discriminant Analysis (DA). The basic purpose of DA is to estimate the relationship between a single categorical dependent variable and a set of quantitative independent variables [26]. Discriminant Analysis involves the determination of a linear equation like regression that will predict which group the case belongs to, minimizing the within-group distance and, simultaneously, maximizing the between-group distance, thus achieving maximum discrimination [26]. The basic idea underlying Discriminant Function Analysis is to determine whether groups differ with regard to the mean of a variable, and then to use that variable to predict group membership. Given  $p$  variables and  $g$  groups, it is possible to establish  $m = \min(g - 1, p)$  discriminant functions,  $DF_i$ , in the form given by [27]

$$DF_i = a + \sum_{k=1}^p b_{ik} x_k \quad (i = 1 \dots m) \quad (3)$$

where  $a$  is a constant,  $b_{ik}$  is the discriminant coefficient and  $x_k$  is the independent variable.

The discriminant model has the following assumptions [28]:

1. Multivariate normality, given that data values are from a normal distribution;
2. Equality of variance-covariance within group, *i.e.*, the covariance matrix within each group should be equal;
3. Low multicollinearity of the variables; when high multicollinearity among two or more variables is present, the discriminant function coefficients will not reliably predict group membership.

## 4 Wind Farms Clustering

Wind farms clustering is addressed here by applying the previous described methodology to a case study in order to identify possible clusters and their main attributes.

## 4.1 Case Study

The chosen support variables used to cluster wind farms are the installed capacity, net production and capacity factor, given in a per-year basis.

Installed capacity of a wind farm is the rated power of each wind turbine multiplied by the number of wind turbines within each farm. The rated power of each wind turbine translates the power each drive train is able to convert from mechanical to electrical energy, under given test wind speed conditions. The data set comprises wind farms with installed capacities higher than 10 MW. It should be noted that some wind farms have been submitted to overpowering processes under the considered time span, *i.e.*, installed capacity of each farm in the per-year basis is not constant.

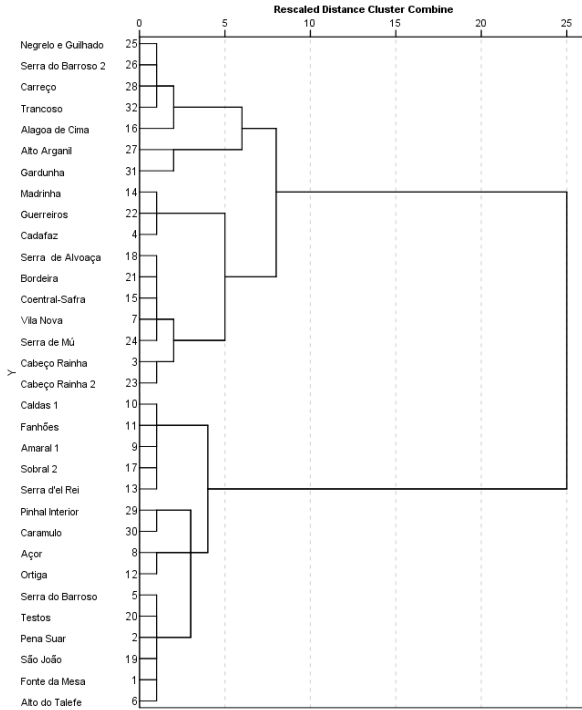
The net production per year measures the output of each farm, in terms of electrical energy delivered to the grid, considering programmed and random unavailabilities (failures) of wind turbines. At present, this output is not constrained by load demand or wholesale markets, as currently regulated.

The capacity factor is the ratio of actual productivity in a year to its theoretical maximum. Higher capacity factors indicate a better utilization of the installed capacity, which helps to reduce investment costs. In fact, capacity factors are particularly important on evaluating the overall economics of wind farms. Typically, capacity factors need to be elevated to values about 50% (or better) to make a modern wind farm commercially viable [18]. The capacity factor relates with the other two variables under analysis, nevertheless its inclusion aims at sensing the wind availability in each farm. Considering a particular farm, with a given number of wind turbines using a specific technological drive train, if the capacity factor increases from one year to another, it means that the meteorological conditions in the later favoured the wind source. On the other hand, comparing wind farms with similar wind availability profile, if they have different capacity factors, it implies that the design technology and layout of the infrastructure are performing differently, assuming that maintenance schemes are similar. From the original data set, two wind farms were excluded from the analysis, given that they performed as outliers regarding this variable, deviating from regular wind energy profiles.

Final data set comprises 32 wind farms from two promoters acting in the wind energy sector in Portugal, EDP and GENERG. The information was collected from institutional and technical Annual Reports available from [29,30], including a time span of 4 years (from 2010 till 2013).

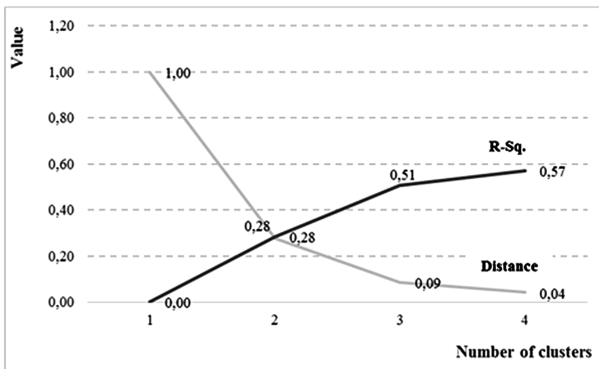
## 4.2 Results and Discussion

The analysis of a dendrogram, using the Ward linkage method, is a common way to anticipate the hypothetical optimal number of wind farms' clusters as well as their composition. This graph also allows observing the distance level at which there is a combination of wind farms and clusters. From the obtained dendrogram, shown in Fig. 1, at the rescaled distance of 15, it is straightforward the definition of two notable groups.



**Fig. 1.** Dendrogram graph for the data set under analysis

In order to strengthen the identification of the optimal number of clusters, it is also used the test R-Squared (R-Sq) based on the analysis of variance (ANOVA) information, namely the ratio between Sum of Squares Between Groups or Clusters and the Total Sum of Squares. The results of the relativized distance between clusters is shown in Fig. 2, from which a solution of two clusters has been chosen, explaining 28% of the total variance. Table 1 outlines wind farms that pose similar features allocated in the 2 identified clusters. The first cluster comprehends 15 wind farms, labelled cluster A, and the second cluster has 17 wind farms, with the label S.



**Fig. 2.** Optimal number of clusters for the data set



**Table 1.** Clustering components based on variables installed capacity, net production and capacity factor

Cluster	Wind Farms		
A	Açor	Fanhões	São João
	Alto do Talefe	Ortiga	Serra d'el Rei
	Amaral 1	Pena Suar	Sobral 2
	Caldas 1	Pinhal Interior	Serra do Barroso
	Caramulo	Fonte da Mesa	Testos
S	Alagoa de Cima	Madrinha	Serra de Alvoaça
	Bordeira	Negrelo e Guilhado	Serra de Mú
	Cabeço Rainha	Guerreiros	Serra do Barroso 2
	Cadafaz	Cabeço Rainha 2	Alto Arganil
	Coentral-Safra	Gardunha	Vila Nova
	Carreço	Trancoso	

The following step consists in application of the Discriminant Analysis to validate the results produced by applying the Cluster Analysis, as previously described. According to this analysis, it is possible to observe significant mean differences for all predictors - independent variables - installed capacity, net production and capacity factor in the clusters A and S - dependent variables. The discriminant function reveals a significant association between groups and all predictors. Regarding the DA results, with only two groups, one Discriminant Function has been produced with an eigenvalue of 3,002. The canonical correlation is given by the multiple correlation between the predictors and the discriminant function. With only one function, the DA provides an index of overall model fit, which is interpreted as being the proportion of variance explained (R-Sq). The relation between the canonical discriminant function and the clusters reveals a greater correlation, approximately 87%. The significance of the discriminant function, evaluated by the Wilks' lambda, is considerably high (p value < 0,001), which means that the model has significant discriminatory power, and provides the proportion of total variability not explained, i.e., the converse of the squared canonical correlation, evaluated in 25%. The classification results reveal that 93,8% of wind farms are classified correctly into 'Cluster A' or 'Cluster S' groups, which can be considered excellent. The wind farms in Cluster S were classified with slightly better accuracy (94,1%) than wind farms in Cluster A (93,3 %).

Table 2 shows the results for the variables into the different clusters, Capacity Factor ( $C_F$ , in %), Net Production ( $W$ , in GWh) and Installed Capacity ( $P$ , in MW) for the time span under analysis.

**Table 2.** Summary of descriptive statistics by cluster and variables

Cluster	n.º	Variable	Minimum	Maximum	Mean	Std. Deviation
A	15	$P$ (MW)	10,00	144,00	29,80	37,134
		$C_F$ (%)	23,42	28,36	25,80	1,293
		$W$ (GWh)	22,35	328,80	66,86	83,418
S	17	$P$ (MW)	10,00	114,00	28,98	23,843
		$C_F$ (%)	27,86	32,77	29,690	1,548
		$W$ (GWh)	28,20	280,00	74,14	58,012

From the descriptive statistics it is possible to observe that both clusters have a high dispersion of the installed capacity, with similar mean values. Main difference between clusters is the mean value of the capacity factor, with a low dispersion: cluster A has lower capacity factors than cluster S, meaning that the overall performance of wind farms in the later cluster had a better performance, for similar installed capacities. In consequence, the mean net production of cluster S is higher than the one of cluster A, whereas the high standard deviation observed for this variable in both clusters follows obviously the one observed for the installed capacity.

After the definition of the clusters and their components by the proposed methodology, it is now possible to explore dominant attributes in each cluster, in order to find particular patterns. The attribute looked into each cluster is the technological trend of the energy conversion system utilized in the different farms. It is possible to observe that Cluster A has a predominance of technology based on Asynchronous generators (73,3% of the farms), while in Cluster S the dominant technology is based on direct driven Synchronous generators (76,5% of the wind farms). As previously stated, this technological concept, using full variable speed range, improves substantially the efficiency of the system which corroborates the higher capacity factors observed within this clusters. Moreover, the absence of the gearbox component increases the reliability and allows reduced maintenance schemes which, together with the wide speed range operation, results in an increased capture of the disposable wind energy.

## 5 Conclusions

The main objective of this research was to identify homogeneous groups within a data set of wind farms of two promoters acting in the energy sector in Portugal, based on two multivariate analyses using a support matrix with three independent variables: installed capacity, net production and capacity factor, in a per year basis. In a first stage it is used Hierarchical Cluster Analysis followed by a Discriminant Analysis, in order to validate the results produced by the first one. Based on both methodologies, from the obtained results, it has been possible to identify two clusters, explaining 28% of the total variance.

Regarding the DA outcomes, one discriminant function has been produced with an eigenvalue of 3,002. The relationship between the canonical discriminant function and the clusters reveals a satisfactory positive correlation, about 87%. Also, the significance of the discriminant function, evaluated by the Wilks' lambda test, is significantly high which means that the model has significant discriminatory power, and provides the proportion of total variability not explained, *i.e.*, the converse of the squared canonical correlation is evaluated in 25%. The classification results shows that 93,8% of original grouped cases are correctly classified.

Following the clustering and validation methodology, it has been possible to identify the technological trend based on the wind turbine generator type in each cluster: Cluster A, with asynchronous generator based technology and cluster S, mainly using direct driven synchronous generator based technology.

From the descriptive statistics regarding data set of both clusters, it is possible to infer that cluster S presents higher mean value of capacity factor than cluster A, which is a good indicator that wind farms using the technological trend based on direct driven synchronous generators have a better performance than the ones based on geared asynchronous generators. The latter, may reduce the initial costs of the drive trains, by using partially scaled electronic converters, but as a counterpart, the increase in the converted energy over the lifespan may offset the higher initial cost.

## References

1. EWEA: Wind in power, 2014 European statistics. <http://www.ewea.org/> (access date: February 2015)
2. ENEOP: Eólicas de Portugal, S.A. <http://www.eneop.pt/> (access date: January, 2014)
3. Jonkman, J., Butterfield, S., Musial, W., Scott, G.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development. National Renewable Energy Laboratory, Technical Report NREL/TP-500-38060 (2009)
4. Karki, R., Billinton, R.: Cost-Effective Wind Energy Utilization for Reliable Power Supply. *IEEE Transactions on Energy Conversion* **19**, 435–440 (2004)
5. Denault, M., Dupuis, D., Couture-Cardinal, S.: Complementarity of Hydro and Wind Power: Improving the Risk Profile of Energy Inflows. *Energy Policy* **37**, 5376–5384 (2009)
6. Arabian-Hoseynabadi, H., Tavner, P.J., Oraee, H.: Reliability Comparison of Direct Drive and Geared Drive Wind Turbine Concepts. *Wind Energy, Wiley Online Library* **13**, 62–73 (2010)
7. Polinder, H., van der Pijl, F.F.A., de Vilder, G.J., Tavner, P.J.: Comparison of Direct-Drive and Geared Generator Concepts for Wind Turbines. *IEEE Transactions on Energy Conversion* **21**, 725–733 (2006)
8. Chicco, G., Ilie, I.S.: Support Vector Clustering of Electrical Load Pattern Data. *IEEE Transactions on Power Systems* **24**, 1619–1628 (2009)
9. Ali, M., Ilie, I.S., Milanovic, J.V., Chicco, G.: Wind Farm Model Aggregation Using Probabilistic Clustering. *IEEE Transactions on Power Systems* **28**, 309–316 (2013)
10. Andrada, M.F., Vega-Hissi, E.G., Estrada, M.R., Garro Martinez, J.C.: Application of k-means clustering, linear discriminant analysis and multivariate linear regression for the development of a predictive QSAR model on 5-lipoxygenase inhibitors. *Chemometrics and Intelligent Laboratory Systems* **143**, 122–129 (2015)
11. Bouguettaya, A., Yu, Q., Liu, X., Zhou, X., Song, A.: Efficient agglomerative hierarchical clustering. *Expert Systems with Applications* **42**, 2785–2797 (2015)
12. Fernandes, P.O., Ferreira, A.P.: Wind farms model aggregation using probabilistic clustering. In: 11th International Conference of Numerical Analysis and Applied Mathematics 2013, ICNAAM 2013, Rhodes, Greece, pp. 618–621. AIP Conference Proceedings (2013)
13. Fernandes, P.O., Ferreira, A.P.: Pattern aggregation of wind energy conversion technologies using clustering analysis. In: 2014 14th International Conference on Computational Science and its Applications (ICCSA), pp. 105–110 (2014)
14. EWEA: Wind Energy – The Facts. Routledge (2009)
15. Hansen, A.D., Iov, F., Blaabjerg, F., Hansen, L.H.: Review of Contemporary Wind Turbine Concepts and their Market Penetration. *Wind Engineering* **28**, 247–263 (2004)
16. Hansen, A.D., Hansen, L.H.: Wind Turbine Concept Market Penetration over 10 Years (1995–2004). *Wind Energy, Wiley Online Library* **10**, 81–97 (2007)

17. Blaabjerg, F., Iov, F., Kerekes, T., Teodorescu, R.: Trends in power electronics and control of renewable energy systems. In: 14th International Power Electronics and Motion Control Conference (EPE-PEMC 2010), Ohrid, Republic of Macedonia, pp. k1–k19 (2010)
18. Mathew, S., Philip, G.S.: *Advances in Wind Energy and Conversion Technology*. Springer (2011)
19. Musial, W., Butterfield, S.: Improving wind turbine gearbox reliability. In: 2007 European Wind Energy Conference, Milan, Italy (2007)
20. Baroudi, J.A., Dinavahi, V., Knight, A.M.: A Review of Power Converter Topologies for Wind Generators. *Renewable Energy*, Elsevier **32**, 2369–2385 (2007)
21. Enercon: Enercon – Technology. <http://www.enercon.de/en-en/21.htm> (access date: June 2011)
22. Conroy, J.F., Watson, R.: Low-Voltage Ride-Through of a Full Converter Wind Turbine with Permanent Magnet Generator. *Renewable Power Generation*, IET **1**, 182–189 (2007)
23. GE\_Energy: Wind Turbines. [http://www.ge-energy.com/products\\_and\\_services/products/wind\\_turbines/](http://www.ge-energy.com/products_and_services/products/wind_turbines/) (access date: June 2011)
24. Ward, J.: Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* **58**, 236–244 (1963)
25. Hardle, W., Simar, L.: *Applied Multivariate Statistical Analysis*, 3rd edn. Springer, Heidelberg (2012)
26. Johnson, R.A., Wichern, D.W.: *Applied Multivariate Statistical Analysis*, 6th edn. Pearson (2007)
27. Klecka, W.R.: *Discriminant Analysis*. Sage Publication, Inc. (1980)
28. Garson, G.D.: *Discriminant Function Analysis*. Statistical Associates Publishers (2012)
29. EDP: Energias de Portugal. <http://www.edp.pt/> (access date: January 2015)
30. GENERG: Energia Natural. <http://www.generg.pt/> (access date: January 2015)