How to determine the Portfolio Effect based on wind regime dependency: European examples

J. M. Marco[1], C. Triviño[1], G. Gil[1], A. Garrad[2], L. Landberg[2]
[1]Garrad Hassan Ibérica S.L.U., SPAIN
[2]Garrad Hassan & Partners Ltd., UK

jose.marco@garradhassan.com
circe.trivino@garradhassan.com

Summary: Important uncertainties associated with wind variability may be mitigated when considering several wind farms within a portfolio rather than in isolated. If two wind regimes are dependant, the deviations from the average of each wind regime are expected to be added when considering them together; however, if the wind regimes are independent these deviations may compensate each other (the so called "portfolio effect"). This is an important feature for investors and operators when quantifying the associated uncertainties of the predicted energy yield of the overall portfolio. A methodology to determine the degree of dependency between wind regimes to analytically evaluate the "portfolio effect" is discussed in this paper using different wind regimes based on wind speed and production data.

Correlations of representative series are undertaken between every pair of wind farms so a Pearson coefficient matrix is built defining the degree of dependence between every element of the portfolio. The coefficients are explained in meteorological and geographical terms, so higher correlation corresponds to sites with similar wind regimes. This resulting correlation matrix combined with the individual wind variability uncertainty of each wind farm provides the (reduced) uncertainty of the portfolio associated with the future variability of the wind.

The benefit which is observed in real portfolios of wind farms may be as large as 25% in the overall uncertainty for a 1-year period.

Keywords: Wind regime, Uncertainty, Dependency, Correlation, Portfolio effect.

1 INTRODUCTION

Due to the natural variability of wind, the annual energy yield of a given wind farm may change drastically from year to year depending on the wind regime of the area. A study undertaken [1] has estimated that annual mean wind speeds over the long term may be assumed to be normally distributed with a standard deviation of 6 %. The following standard statistical method is used:

Historic data period uncertainty = (Annual average variability)/$\sqrt{(N)}$

Where:
Annual average variability is 6 % or a region or site specific number as defined above.
N is the number of years of valid data in the period assumed to represent the long term.

The same logic used to derive the uncertainty in the historic wind speed period due to annual wind speed variability is used to define the variability in future period of interest. In this analysis a future period of 1 year duration is considered although other periods may be considered.

When considering the uncertainty of the production estimation of a wind farm, the uncertainty associated with these two sources, historic wind speed period and variability in the future annual mean wind speed, is significant and consequently, the variation in energy production can be considerable.

The different variability of the wind regimes in a portfolio may produce a compensation of the wind farms individual uncertainties if the wind regimes are independent or conversely can compound the uncertainty if they are inter-dependant on each other. It is therefore necessary to analyse the dependency/independency between wind regimes in order to determine how this "portfolio effect" acts in the variability uncertainty associated with a portfolio of wind farms. In such a way, the "portfolio effect" can result in a reduction in the uncertainty of the overall portfolio compared to the uncertainty associated with each individual wind farm within a group.

A methodology to determine the dependence or independence between wind regimes is discussed in this paper and is used to analytically evaluate the “portfolio effect” associated with a portfolio of wind farms globally distributed.

In the first place, basic uncertainty and statistical concepts will be reviewed. A breakdown of uncertainties associated with wind farm production estimation to determine which are related to wind variability and therefore may be affected by the “portfolio effect” and those which are not. An assessment of the expected portfolio effect therefore requires establishing the statistical level of independence between projects and groups of projects for each source of uncertainty identified. Assuming that the correlations between projects is linear, the level of independence is typically assessed through a correlation analysis where the resulting Pearson coefficient, r, defines the level of dependency.

Since the periods of available data are relatively short, the investigation has focused on the correlation of monthly mean wind speeds as a proxy for determining the level of correlation. The use of monthly data, however, is recognised to introduce the potential for autocorrelation due to the seasonality inherent in most wind regimes. For example, many wind regimes exhibit winter peaking seasonal distributions of wind speed and a correlation of absolute monthly data will in that case exhibit a level of correlation whether one physically exists or not. In order to avoid such artifices, the seasonal pattern of wind speeds was estimated from the available data and used to remove the seasonality (normalise) the individual monthly mean wind speeds. The use of annual mean wind speeds for determining the correlation would be ideal since this effect would be removed, however, as mentioned before, the amount of annual data available is small and therefore it was not practical to follow this route for the analysis reported here.

Thus, as explained in section 2, the reason for using monthly averages is that they are “independent” enough and sufficiently “normally distributed” for the use of Parametric techniques, and they allow having a higher number of data pairs for the correlations.

Finally, the results will be discussed together with the applications and limitations of the methodology.

2 BACKGROUND

The overall uncertainty in an individual energy assessment is typically due to the following sources, each of which is considered statistically independent of the other factors:

- Anemometry accuracy (Type 1)
- Site wind shear estimation and vertical extrapolation (Type 1)
- Correlation analyses to adjust to the long-term (Type 1)
- Wind rose uncertainty (Type 1)
- Wind flow and wake modelling over the site (Type 1)
- Historical data period uncertainty (Type 2)
- Future wind speed annual variability (Type 2)

The uncertainties can therefore be considered to be broadly: (1) those due to the available data and analysis methodology (“model uncertainties”), and (2) those due to the variability of the wind. Any “portfolio effect” in terms of the uncertainty in the expected energy production arises due to the statistical independence of these contributing sources between projects and therefore requires establishing the statistical level of independence between projects and groups of projects for each source of uncertainty identified.

The level of independence can only be estimated analytically for (2) type uncertainties and therefore assumptions have been made for the type (1) uncertainties considering them as dependant. The overall portfolio uncertainty is the result of adding both types of uncertainty (1 and 2) independently [2].

In order to determine the level of independence of (2) type uncertainties, monthly mean wind speed data sets and observed production data at 75 sites across Europe are available. Each data set is assumed to be a sample belonging to a population and therefore, a Parametric Inferential Statistics analysis needs to be undertaken meaning that conclusions about a whole population can be drawn from the sample. When using Inferential Statistics, the correlation coefficient is given together with a confidence interval, which contains the value of the population parameter (with a concrete significance level) and, at the same time, this interval expresses how representative the sample is.

Some requirements [3, 4] need to be fulfilled by the data series in order that the use of Parametric techniques is appropriate; these are listed below:

(1) Both the population and the sample must fit to a normal distribution. It is difficult to know what the population fits to, however according to [5], it is common for monthly and annual mean wind speed distributions to be normally distributed when the sample size is sufficiently large. In order to verify the normality of every sample, the Kolmogorov-Smirnov normality test has been applied. In those series where a normal distribution is not fully achieved, symmetry is required [3].

(2) Independency of the samples, which means that there is no relation between individual elements in the sample or, in other words, no autocorrelation exists. The key point is to determine if the inherent "dependency" of the monthly samples is sufficiently strong to invalidate this method and therefore the Ljung-Box test has been applied to determine the degree of autocorrelation. Additionally, the bivariate distribution must be normal in order to apply this kind of test to the Pearson coefficient. It is known that in the case of two independent and normally distributed samples the normality of the bivariate distribution can be assumed.

Once the level of independence has been established, the combination of uncertainties to estimate the overall uncertainty level is defined as:

$$\text{Variance – Covariance matrix} = \begin{pmatrix} \sigma_1^2 & r_{12}\sigma_1\sigma_2 & \dots & r_{1j}\sigma_1\sigma_j & r_{1m}\sigma_1\sigma_m \\ r_{21}\sigma_2\sigma_1 & \dots & & & \\ \dots & & \sigma_i^2 & & \\ r_{i1}\sigma_i\sigma_1 & & & \dots & \\ r_{m1}\sigma_m\sigma_1 & & & & \sigma_m^2 \end{pmatrix} \quad \text{[Eqn 1]}$$

Where σ_i and σ_j are the standard errors for a given source of uncertainty for Project i and j. Since the Pearson coefficient, r_{ij}, is symmetric $r_{ij} = r_{ji}$, a simplification which can be made for its implementation is to reduce the above matrix to the form below:

$$\text{Variance – Covariance matrix} = \begin{pmatrix} \sigma_1^2 & 2r_{12}\sigma_1\sigma_2 & \dots & 2r_{1j}\sigma_1\sigma_j & 2r_{1m}\sigma_1\sigma_m \\ 0 & \dots & & & \\ 0 & & \sigma_i^2 & & \\ \dots & & & \dots & \\ 0 & 0 & \dots & & \sigma_m^2 \end{pmatrix} \quad \text{[Eqn 2]}$$

The Pearson coefficient has a range of $-1< r <+1$ where +1 represents perfectly linear dependency while a value of 0 represents complete independence.

3 METHODOLOGY AND APPLICATION TO WIND DATA

3.1 Available Data

The analysis of the level of independence between regional wind regimes is ideally undertaken based on an assessment of the correlation of on-site sources of consistent long-term wind speed or production data for each of the projects. However, in many projects, the data available are typically restricted to actual on-site or regional meteorological data over a short period of time and the consistency of data throughout the entire period is sometimes difficult to demonstrate in long-term sources.

A portfolio formed of 75 sites distributed throughout Portugal, Spain, France and Germany has been considered in this analysis, located within the areas shown in Figure 1. A data set of monthly wind speed data as proxy or the real production data of the wind farms in some cases have been considered representative for each wind farm.

In order to avoid potential for autocorrelation due to the seasonality inherent in most wind regimes, the seasonal pattern of wind speeds was estimated from the available data and used to normalise the individual monthly mean wind speeds. The tests described in Chapter 2 were then applied to these series in order verify the use of the parametric techniques on which this methodology is based.

Figure 1 General location of the wind farms of the portfolio in Europe

The overall uncertainty associated to each wind farm has been divided into (1) “model uncertainties” and (2) ”wind variability uncertainties”..

Uncertainty (1): Associated with measurement accuracy of anemometers, mounting arrangements, long-term analysis techniques, modelling and vertical extrapolation among others.

Uncertainty (2): Associated with wind variability of the available period (wind speed data or production data) and the variability in a future period of interest of 1 year.

3.2 Correlation analysis

A correlation of concurrent normalised monthly mean wind speed or production data was undertaken between each combination of project pairs and the resulting Pearson coefficient established. Additionally, a review of the statistical significance of the Pearson coefficient, which is function of both the value of the Pearson coefficient, r, and the number of data pairs, n, was undertaken. The upper and lower limits of the 95% confidence interval were established as follows:

$$r_{upper,lower} = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \pm 1.96\left(\frac{1}{\sqrt{n-3}}\right) \qquad \text{[Eqn 3]}$$

Where the estimated lower limit was negative, it was observed that the values rarely indicated a strong negative correlation and were generally close to zero. Consequently, for such cases the resulting Pearson coefficient values were set to zero, indicating complete independence of the wind regimes, in order to avoid the introduction of artificial statistical influences into the analysis.

As described in Chapter 2, to draw conclusions about the r coefficient, that is, to calculate its confidence interval, it is necessary to assume that the data are independent and normally distributed. In order to demonstrate that these two factors were satisfied, normal distribution and autocorrelation tests (samples in Figures 2 and 3) have been performed for each series.

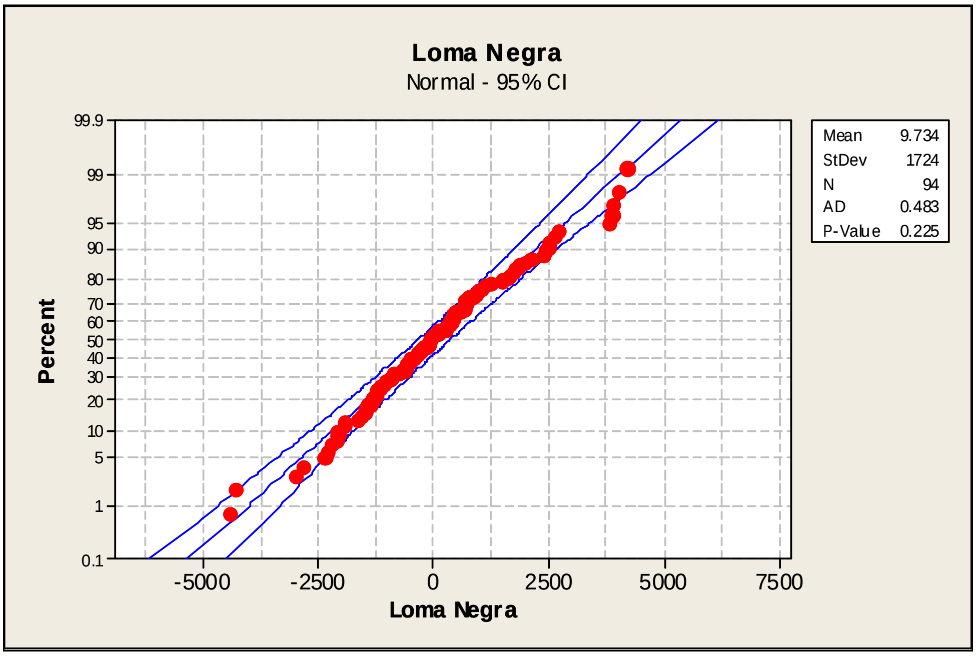

Figure 2 Sample of a Normality Test

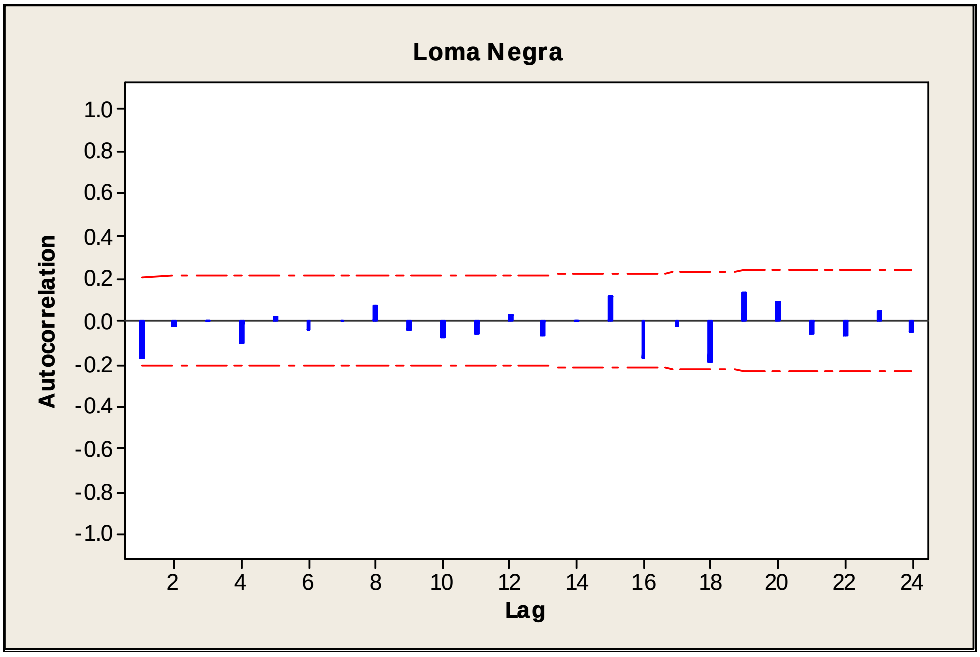

Figure 3 Sample of an Autocorrelation Test

3.3 Covariance matrix

Based on the above correlation analysis and adjustment of the resulting Pearson coefficients as described, a correlation matrix describing the inter-dependency of wind regimes between all projects is defined. From this, the covariance matrices describing each of the terms in [Eqn 2] for the wind variability of the available period and 1-year future wind variability were developed based on the individual project wind variability standard errors calculated.

From the results it can be observed that northern of France is quite well correlated with the centre and the west of Germany with Pearson coefficients from 0.58 to 0.83 as identified previously with different techniques in other papers [6].

It is also noticed that there is certain correlation between the wind regimes in the northwest of Spain and the centre and north of France with Pearson coefficients above 0.50, higher than the correlations between several sites within Spain. This is reasonable as the incident fronts from the Atlantic sweep the northwest of Spain first and then the western coast of north Europe while the meteorological mechanisms affecting the northwest and southeast of Spain are rather different.

It is interesting to note the independency between a site in the Ebro Valley and a site in the Iberian System, which are very close together, and the partial dependency of these two sites with Galicia, which is very far away. This statistical result reflects how the same meteorological situation affects different sites in rather different ways resulting in different wind regimes.

The total independency between a site in the Ebro Valley and a site in the Iberian System is explained by the orographic dipole created by the Pyrenees that generates a very canalised flux in the Ebro valley, Cierzo, while low winds are expected downwind the Iberian System. This typical situation produces northeasters (north winds situation) winds in Galicia which contributes to the correlation between Galicia and the Ebro Valley. When the entry of fronts from Galicia occurs (south winds situation), an acceleration of the south wind in the Iberian system is observed which contributes to the correlation between Galicia and this area.

3.4 Overall portfolio uncertainty

The overall future wind speed variability was then established from the covariance matrices as the sum of the terms of the matrix as described by [Eqn 2].

The resulting portfolio wind variability of the period standard error is 139.1 GWh/annum (261.4 GWh/annum not considering portfolio effect).

The resulting portfolio 1-year future wind speed standard error is 219.4 GWh/annum (429.8 GWh/annum not considering portfolio effect).

The relation between the different sources of the "model" uncertainty of the different projects is very difficult to assess as it is not clear how to determine it. For instance, how to determine if the uncertainty of the measurements of two sites are mitigated in some degree when considered together (independent) or the opposite. Same logic applies to the other uncertainties and therefore, assumptions need to be made for the dependency relation between projects for this type of uncertainties which are not related to wind variability.

It is recommended to consider these "model" uncertainties as dependant, as there is no proven way to assess the degree of independency. In such a way, assuming the "model" uncertainties of each project in the portfolio as dependant and the wind variability ones as partially dependant (portfolio effect), the overall 1-year wind speed standard error of the portfolio is calculated to be 488.2 GWh/annum (651.0 GWh/annum not considering portfolio effect).

4 CONCLUSIONS

The "portfolio effect" of the current period and future wind speed variability uncertainty of a portfolio made up of 75 wind farms in different European areas has been assessed in this paper. The analysis is carried out using the real production data or the monthly wind speed data as proxy which have been used to state a level of dependency/independency between sites summarised by a Pearson coefficient matrix. Combining this matrix with the individual uncertainties, it is possible to determine the "portfolio effect" associated with future wind speed variability uncertainties. Results are summarised below:

	Estimated Energy Production [GWh/y]	**Wind variability of the available period (b) [GWh/y]**	**Future 1 yr variability Uncertainty (b) GWh/y]**	**Model uncertainty (a) [GWh/y]**	**Overall uncertainty [GWy]**
No "portfolio effect"	3792.7	261.4	429.8	413.3	651.0
"Portfolio effect"	3792.7	139.1	219.4	413.3	488.2
Benefit due to "portfolio effect" in the wind speed variability					25.0%

Figure 4 shows how lower uncertainties imply higher energy estimations for the same exceedance levels P75 and P90 for the portfolio energy estimation, considering or not considering "Portfolio Effect" for the wind variability uncertainties.

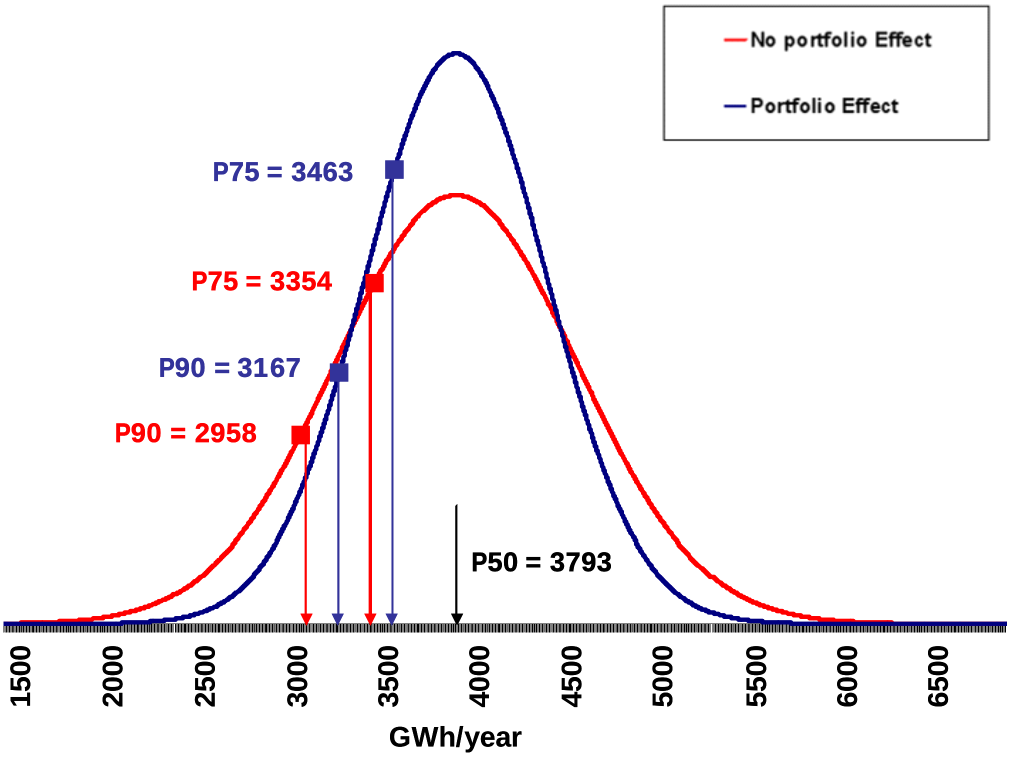

Figure 4 Exceedance levels P75 and P90 for the portfolio energy estimation

It is noted that the portfolio is made up of sites spread throughout Europe in different geographical areas. This geographic dispersion intensifies the independency between wind regimes and therefore increases the observed "portfolio effect" in the uncertainties compared to other portfolios made up of sites with less disperse geographical locations. The scope of this selection is to show the importance of considering this effect when analysing portfolios rather than considering the wind farms as isolated entities. This feature is very important for investors and owners to mitigate wind risk by acquiring a geographically distributed wind farm portfolio.

5 REFERENCES

1. Raftery A., Tindal J. and Garrad A. (1997), Understanding the risks of financing windfarms, Proc. EWEA Wind Energy Conference, Dublin.

2. Kreyszig E. (1983), Advanced Engineering Mathematics, John Wiley and Sons.

3. Montgomery, D.C.; Runger, G.C. (1996). Probabilidad y Estadística Aplicadas a la Ingeniería. McGraw-Hill..

4. Chatfield, C. (1998). Statistics for Technology. Chapman and Hall.

5. Justus, Mani and Mikhail (1979), Internal and Month-to-Month Variations of Wind Speed, Journal of Applied Meteorology.

6. Harman K., Morgan C., (2005), Use of regional wind energy indices to predict long-term wind farm production and to assess portfolio effect, Proc. WREC, Aberdeen.