Wind missing data arrangement using wavelet based techniques for getting maximum likelihood

Antonio J. Zapata-Sierra[a], Alejandro Cama-Pinto[b], Francisco G. Montoya[a], Alfredo Alcayde[a], Francisco Manzano-Agugliaro[a,1,*]

[a] *Department of Engineering, CEIA3, Universidad de Almeria, Almeria (Spain)*
[b] *Engineering School, Universidad de la Costa, Barranquilla (Colombia)*

Abstract

Long time series of wind data can have data gaps that may lead to errors in the subsequent analyses of the time series. This study proposes using the wavelet transform as a system to verify that a data completion technique is correct and that the data series behaves correctly, enabling the user to infer the expected results. Wind speed data from three weather stations located in southern Europe were used to test the proposed method. The series consist of data measured every 10 minutes for 11 years. Various techniques are used to complete the data of one of the series; the wavelet transform is used as the control method, and its scalogram is used to visualize it. If the representation in the scalogram has zero magnitude, it shows the absence of data, so that if the data are properly filled in, then they have similar magnitudes to the rest of the series. The proposed method has shown that in case of data series inconsistencies, the wavelet transform can identify the lack of accuracy of the natural periodicity of these data. This result can be visually checked using the WT's scalogram. Additionally, the scallograms provide valuable information on the variables studied, e.g. periods of higher wind speed. In summary, the wavelet transform has proven to be an excellent analysis tool that reveals the seasonal pattern of wind speed in periodograms at various scales.

Keywords:
Wind data, Wavelet Transform, FFT, missing data, renewable energy, data filling

*Corresponding author
Email addresses: ajzapata@ual.es (Antonio J. Zapata-Sierra), acama1@cuc.edu.co (Alejandro Cama-Pinto), pagilm@ual.es (Francisco G. Montoya), aalcayde@ual.es (Alfredo Alcayde), fmanzano@ual.es (Francisco Manzano-Agugliaro)
[1] Tel: +34950015693

1. Introduction

The scarcity of fossil fuels and the environmental concerns of climate change related to their use have contributed to the development and use of the two most important renewable energies in the world, i.e., solar and wind energy, being wind energy one of the most widespread in the world, depending on availability [1, 2]. Each MWh of wind energy prevents the emission of at least 500 kg of greenhouse gases [3, 4, 5], and since it was first developed in the 1980s, wind farms have experienced a worldwide increase of more than 1500%, reaching a total installed capacity of 432 GW by the end of 2015 [6, 7].

Wind energy production is related to the quality and quantity of the wind speed data [8]. The quality depends on whether the data set is reliable and uniform, while the quantity is related to the data recording time which is usually shorter than the lifetime of the structure of the wind generator and is useful for modelling the worst wind load case expected on the structure during its service life [9]. For these reasons, the development of wind energy requires better understanding of the collection of wind speed time series [10].

Energy models based on environmental data often require a full time series of meteorological data, e.g. wind speed, so reconstruction of missing data is a key issue in the functionality of these models. It is not unusual for weather stations to fail, and therefore techniques are needed to fill the gaps in the data series to use them as input data in the models. Several approaches have been pursued by researchers, first of them is using the same weather station: interpolation using classic statistical models as Linear Regression [11], Auto Regression [12], or Auto Regression Integrate Moving Average [13]. Sometimes for extreme events different types of statistical distributions are used such as: Gumbel, Exponential, Gamma, Normal or Lognormal [14]. Another approach is to use data from nearby stations. So, where data from several neighboring weather stations are interrelated, as example deterministic methods: trend surface analysis (TSA), the inverse distance weighting (IDW) [15, 16], the spatial regression test (SRT) [17], local polynomial (LP) [18], thin plate spline (TPS) [19]. Another group of methods are the geostatistics ones [20]: kriging (provides a solution by taking account of the spatial correlation), ordinary kriging (assumes the mean is unknown, focuses on the spatial component and uses only the samples in the local neighborhood for the estimate), universal kriging (assumes the presence of a trend in average values across the study area) or cokriging (involves more complicated calculations than kriging and the detailed principles are well explained by [21]).

Wind speed is a random meteorological phenomenon that changes with geographic location and time of day, month, year, etc., and whose trend in time and space is difficult to predict [22]. The fast Fourier transform (FFT) is widely used to assess the frequency content of the time series of wind speed data and provides the power spectral density (PSD) [23]. This technique is used to investigate the complicated properties associated with the distribution between the frequency (Hz) and magnitude of the power spectrum (dB). However, most signals, including geophysical time series such as average wind speed, are complex and are considered non-stationary processes because of their many non-stationary (transient) characteristics, such as drift, sudden changes, event starts and ends. These characteristics are usually the most important part of the signal and need to be analysed in order to understand the physical phenomena hidden behind the signal [24]. FFTs are useful for extracting frequencies from a stationary or transitory signal, as well as their predominance in the entire time series (for example, wind speed), in order to investigate the properties associated with the distribution between the frequency (Hz) and magnitude of the power spectrum. It produces averaged spectral coefficients that are time independent and useful for identifying dominant frequencies in a signal; however, it cannot capture wind speeds that vary over time. Therefore these signals, which in nature have irregular or time-limited characteristics, are considered non-stationary. FFT may not be practical or efficient for wind data. In addition, the FFT is limited by the fact that a single window analysis cannot detect signal characteristics that are much longer or much shorter than the window size [25]. Short-Time Fourier Transform (STFT) was developed in attempt to analysis the non-stationary signal. However, using STFT with narrow window to capture high temporal resolution leads to poor frequency resolution [26]. The WT becomes a powerful analyzing tool for stationary, non-stationary, intermittent time series, especially, to find out hidden short events inside the time series [27, 28]. Because of its advantages, the WT have been applied in the various fields such as wind data analysis [29, 30]. Therefore, a representation is required that can follow the spectrum of the signal as it varies with respect to time, this is the case of the WT [31].

Wavelet transform (WT) was first introduced and formulated by Morlet et al. in 1982 [32], and Grossmand and Morlet [33]. Wavelets are functions that satisfy certain mathematical requirements and are used for the representation of data. Wavelets are very suitable for data approximation with variations or sudden discontinuities [33]. The basic idea of wavelets is to analyze functions according to scales [34]. In wavelet analysis, the scale used to analyze the data plays a special role. Algorithms using wavelets process data at different scales or resolutions [35]. If a signal or function is observed using a wide window, small details are not observed; on the other hand, if the window used is narrow, then they can be observed. In parsing by wavelets, those windows automatically adjust when the resolution changes [36]. As with the Fourier transform, the WT uses internal products to measure the similarity between the original signal and an analysing function; specifically, a correlation is made between the original signal and the chosen wavelet [37, 38]. The wavelets have been successfully applied to different studies of meteorological and climatological series to analyse their time scales of variability, underlining the advantages of this technology compared to Fourier transform analysis. The most interesting difference between the two transforms is that the size of the window in the WT varies according to the scale, while in the FFT the window is fixed. This allows a better location in time-frequency. Another difference is the base of functions in FFT are always the same while in the WT it is possible to choose different wavelet mothers that will originate different bases of functions. This allows to choose a suitable base according to the problem under study. In this way, signal analysis using wavelets bases provides immediate access to signal information that remains hidden from other analysis tools. In summary, these functions can slide through the time variable and change their length such that large windows are used for capturing phenomena at low frequencies and short windows for high frequencies. It has been seen that wavelets are a mathematical tool to decompose functions hierarchically. The most widely known applications are: data compression [39, 40]; flexible representation of multi-resolution curves [41], this is of special application in diagnosis in medicine for radiological issues [42]; or retrieve three-dimensional surfaces from unknown objects in applications such as visual inspection [43], stand-alone navigation or robot control [44]. Therefore, the application areas of the transformed wavelet are so varied and, although most of the wavelet's theory has already been developed, new applications can still be searched for. This makes wavelets a useful and interesting tool.

Symmetry and orthogonality are among the characteristics of the families of wavelets that stand out because these increase the computing speed, which is beneficial. While the Fourier transform is used to find the frequency spectrum of a signal assumed to be stationary, the WT is suitable for non-stationary signals because it breaks down the signal on a time-frequency grid at different resolutions which results in a technique known as multiresolution analysis [45]. The WT breaks down a time domain signal $f(t)$ into a function with scale variables a and with shifting or translation variables b [46].

The wavelet analysis method allows the use of long time intervals where more precise low frequency information is needed and of shorter intervals where high frequency information is needed [47, 48]. In wavelet analysis, a specific wavelet function that is most similar to the desired function is selected, and the changes from one time period to the next time period of the function can be defined by matching a wavelet function and changing the scales and positions of that function [49]. Thus, with this method it is possible to capture the characteristic of the wind speed variation that changes over time based on auto-calculation and changes in seasonal patterns [50]; it can also be used to study the efficiency of a wind park [51]. The WT provides a revealing snapshot of the time-frequency localization that enables understanding of the inherent characteristics of wind [52, 53, 54].

Wavelet techniques have shown their applicability in the study of a signal generated from a time series of the average wind speed, for the extraction of information related to its frequency components that vary over time [55, 56]. By applying the WT on the wind signal, the temporary interactions of its frequency components are detected, in contrast to the FFT, which is not entirely appropriate for non-stationary signals such as wind speed [37]. In addition, with the time-frequency plane of the wavelet scalogram, the information in the time and frequency domain is more detailed in comparison to the FFT [57], which provides almost no signal information [58].

Furthermore, weather stations may suffer measurement interruption due to sensor malfunction [59, 60, 61], an occurrence that is worsened for stations located in isolated, rural areas, such as those studied in this research. This fact needs to be properly resolved because otherwise it complicates the analysis and the quality of the obtained results. Some authors use signal processing applications such as WT and FFT for spectral estimation or reconstruction of the signal [62] using these mathematical techniques, it is even possible to predict faults based on previous sets of stored data [63, 64, 65]. So, techniques such as Weibul or Raleigh could also be used as an alternative method for it [66, 67]. Scripts implemented in MATLAB and that use the wavelet toolbox, and FFT among other functions, are frequently used to view the results [68, 56].

Therefore, the lack of data from a time series is a key factor in the subsequent analysis of time series data [69]. Given the usefulness of the wavelet transform for the analysis of wind time series, the objective of this work is that once the lost wind data are completed the WT ensures that these do not adversely affect seasonal patterns.

2. Material and methods

2.1. Available data

This study was carried out using the average wind speed data recorded every 600 seconds (10 minutes), over a period of 11 years, from January 1, 2002 until December 31, 2012. The data were obtained from weather stations located in the province of Almeria, called "Collado de Yuste" (37.22633N, 2.430768W, 1866 masl), "Solana del zapatero" (37.31286N, 2.430768W, 1116.1 masl) and "Calar Alto" (37.22099N, -2.548748W, 2151 masl), hereinafter referred to as CY, SZ and CA (the latter with data only until August 2010), respectively. These areas are characterized by a dry Mediterranean climate, and the region is used as reference because the wind characteristics that reach lower altitudes are known. In addition, based on the data obtained in Collado de Yuste, the average wind speed over 11 years is 4.51 m/s, which makes Collado de Yuste a favourable candidate location for the installation of wind farms (Figure 1). This study area was selected because it has three nearby stations with a large amount of data, which may make it possible to check the proposed methodology.

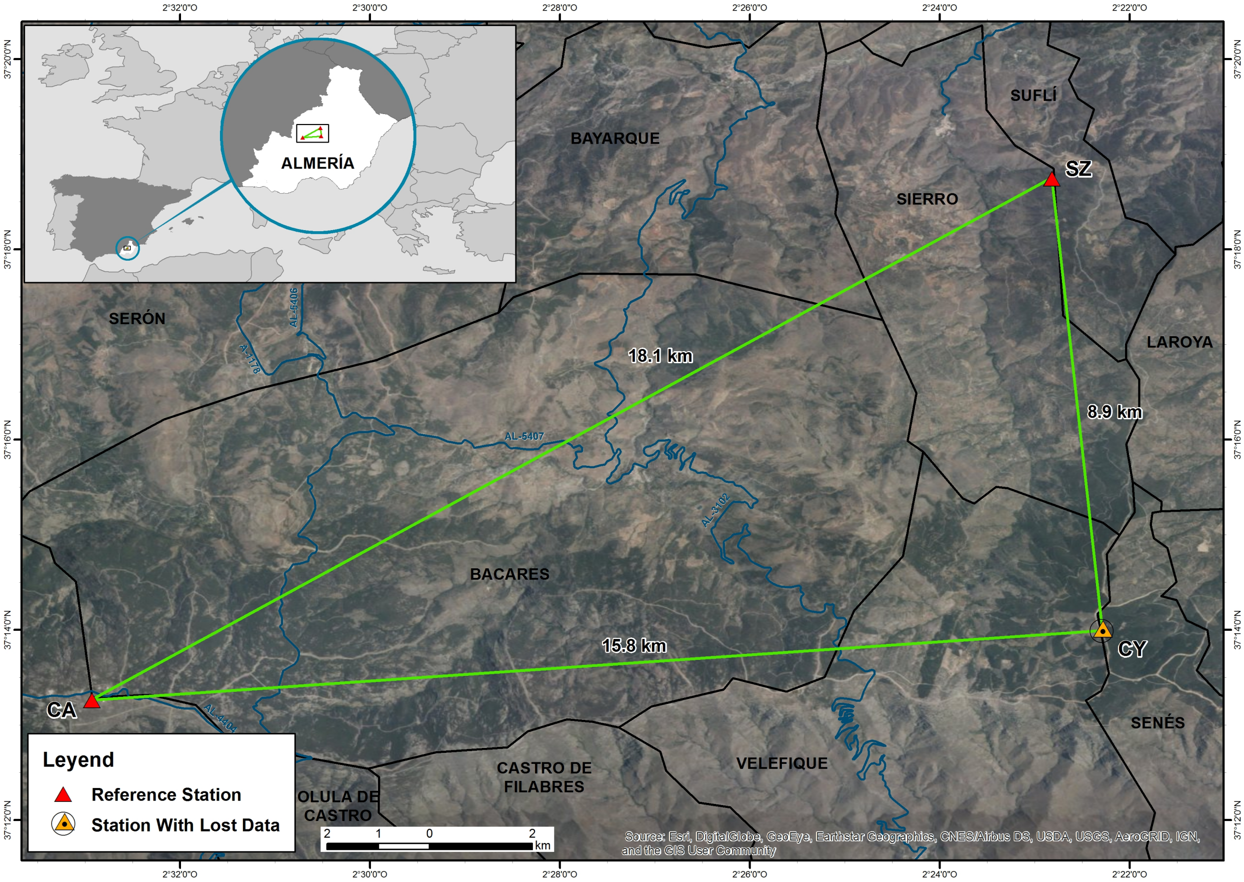

Figure 1: Locations and average wind speed at the Collado de Yuste, Solana del zapatero and Calar Alto weather stations in Almería.

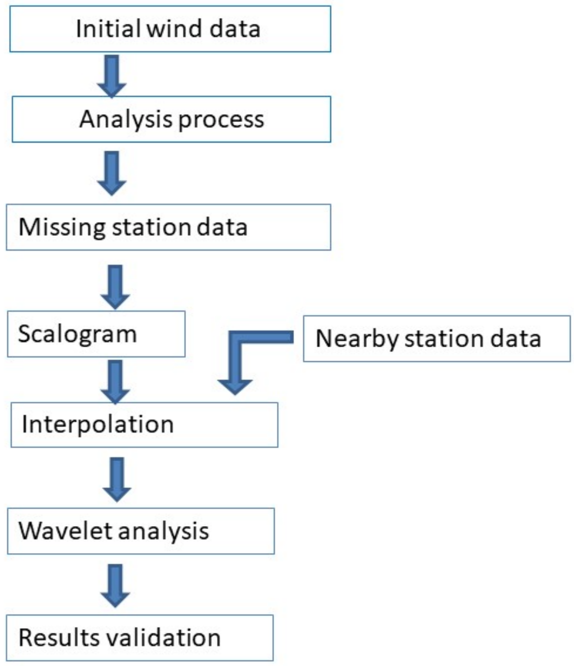

Figure 2: Proposed methodology

2.2. Methods

Figure 2 summarizes the proposed method. First, a scalogram based on the WT to detect the relevance of the potential lost data is developed. Subsequently, an interpolation is performed using a method that uses data from nearby stations, first verifying the correlation coefficient between nearby stations and the station to interpolate. Lastly, the WT is applied to the new data set to test the data validity.

2.2.1. Continous wavelet transform CWT

The WT can be studied using two approaches: the continuous wavelet transform (CWT) [70], in which the variables a and b take continuous values, and the discrete wavelet transform (DWT), in which dyadic and orthogonal scales are used, greatly reducing the redundancy of the CWT. The DWT is particularly useful for noise reduction and data compression, while the CWT is best for feature extraction [71]. For this reason, this paper only discusses CWT [72], which is described by equation 1, where $f(t)$ is the signal to be analysed, $\Psi_{a,b}(t)$ is the mother wavelet scaled by the frequency factor a and localized by the time factor b and Ψ^* is the complex conjugate of the function $\Psi(t)$.

$$CWT_f(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \Psi^* \left(\frac{t-b}{a} \right) f(t) dt \quad \text{where} \quad a > 0 \quad \text{and} \quad -\infty < b < \infty \tag{1}$$

The normalization factor $\frac{1}{\sqrt{|a|}}$ ensures that for all a and b, the energy remains the same. By definition, the CWT is calculated by changing the scale of the window and shifting the window in time; it is then multiplied by the signal that is going to be transformed and integrated over all times. The procedure provides a good frequency resolution for low frequencies, but the time resolution is poor. For higher frequencies, the time resolution is good, but the frequency resolution is poor [73].

More specifically, the Morlet wavelet has proven to be a good choice for the analysis of intermittent oscillations of the average wind speed found in a time series [74] (see Equation 2). It is used specifically in the analysis of meteorological time series, and we use it to analyse the percentage of energy distribution of the scalogram. Using the Morlet wavelet with $w_0 > 5$, a good balance is provided between frequency localization and time localization [75, 76].

$$\Psi_0(n) = \pi^{\frac{-1}{4}} e^{i w_0 n} e^{-\frac{n^2}{2}} \tag{2}$$

where w_0 is the dimensionless frequency and n the dimensionless time.

2.2.2. Wavelet scalogram

The scalogram is a visual method for showing the time and frequency localization of the wavelet coefficients. It is represented by three axes: the x axis represents time, the y axis represents the scale, and the z axis represents the value of the wavelet coefficient [73], which is represented in Figure 3 using distinct colours. The scalogram is used to detect the most representative scales or time instants of a signal, namely, the scales or time instants that contribute the most to total wind energy [77] of the wind speed time series that indicates the spectral sensitivity [25], and in our study, provides a visual representation of the energy distribution of the wavelet coefficients [73]. In contrast to the power spectrum, the scalogram simultaneously reveals the time and frequency information of the wind signal [58] and shows the duration of each frequency component that occurs periodically, quasi-periodically and even randomly.

In signal processing, a scalogram is a method of visually showing a wavelet transform. The scalogram of a time series x in a given scale $a > 0$ is defined in Equation 3 as:

$$S(x) = \|W_a x(b)\| = \sqrt{\left| \int_{-\infty}^{\infty} x(t) \Psi \left(\frac{t-b}{a} \right)^2 \right| da} \tag{3}$$

where the energy of $W_a x(b)$ represents a scale of a. The scalogram allows detection of the most representative scales (or frequencies) or time instants of a signal, namely, the scales or time instants that contribute most to the total energy of the signal [68, 77]. The enveloping curve shown at the bottom of the scalogram

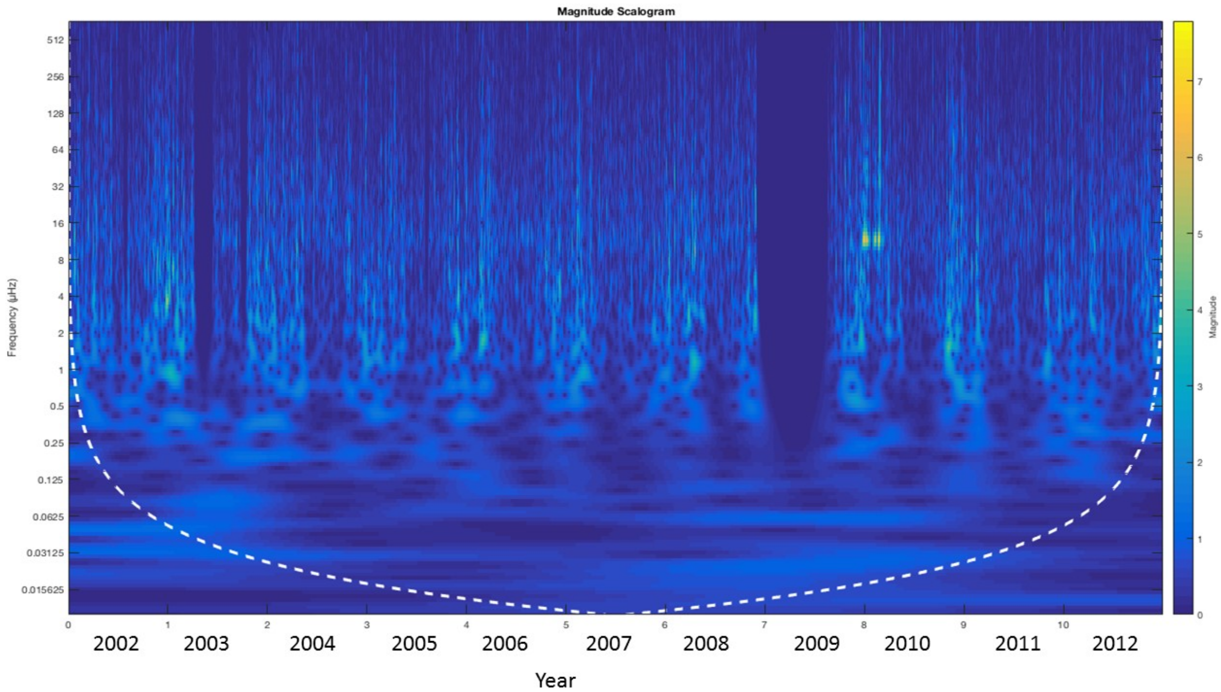

Figure 3: The WT power spectrum of wind speed of CY with original data

is identified with a white dashed line. The estimates within the cone of influence show regions in which the wavelet coefficient is reliable [78].

2.2.3. Data interpolation techniques

2.2.3.1 Interpolation using data from the same station

It can be assumed that the series of daily averages, obtained from the average of all the data available for each selected day, should allow the completion of the series. When a gap is detected, the fragment of the average series is scaled and adapted to the initial and final values. These new interpolated values replace the missing values in their same positions, as shown in Figure 4.

2.2.3.2 Interpolation using data from nearby stations

When reliable data from nearby stations are available, the lost data from the studied station, in our case, station CY (see Figure 1) are completed using Equation 4

$$C_i = \frac{1}{N} \sum_{i=1}^{i=N} \alpha_i A_i \tag{4}$$

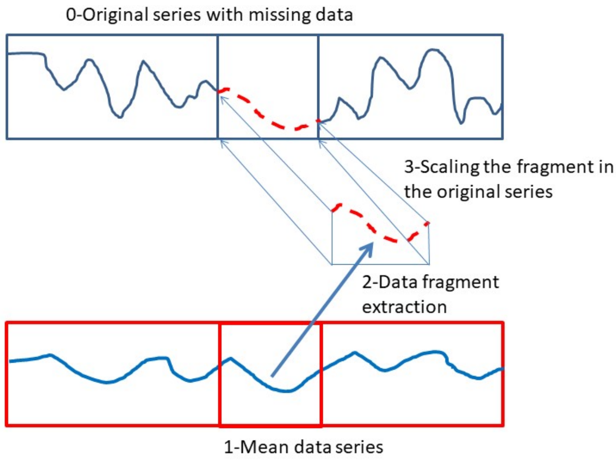

Figure 4: Interpolation of lost data from the same station

where C_i indicates the lost data to be generated, for a specific date and time; A_i indicates the reference data from a nearby station i, for a specific date and time; N indicates the number of locations with existing data for a specific date and time; α_i indicates the weighting coefficient.

The coefficient α is calculated as the annual median rate of the values of that year for the studied station in relation to reference station i. The median is chosen instead of the average because the median is less affected by outliers and skewed data.

3. Results

Figure 3 shows the scalogram obtained by applying the proposed CWT methodology to all the original data recorded from the studied station CY. A periodic pattern can be seen in areas where there is data, and in areas without data there are homogeneous dark blue bands. The most significant data losses are seen in 2003 (Year 1 in the scalograms) and in 2009 (Year 7), which represent 2.27% and 5.84% of the total data from the 11 years recorded at that station (CY).

The 62,108 values of lost data from station CY were interpolated with data from the same station, and the CWT was then applied. In the scalogram generated from this new set of data (Figure 5), the interpolated areas can be seen easily (see Years 2003 and 2009). Specifically, there is an unnatural periodicity over the 11 years, especially in the periods of time with lost data (Figure 3). Therefore, the procedure to interpolate these values is not entirely appropriate. We believe the reason for this effect is that the average of the data is always more centred than the individual values, and therefore, it does not reproduce the individual behaviour of each data point.

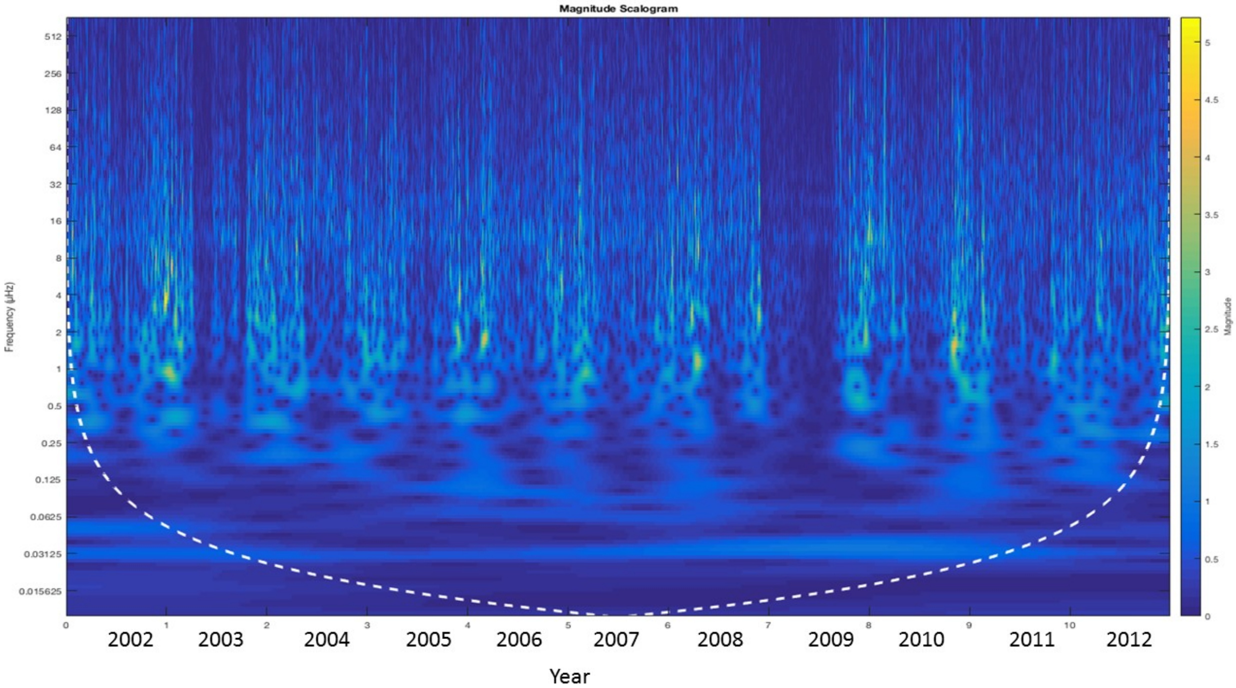

Figure 5: The WT power spectrum of wind speed of CY with interpolated data

Because the interpolation of data with values from the same station is unacceptable, a methodology based on the availability of data from nearby stations is proposed. To do this, it is first necessary to check if there is a good correlation between the stations. For the three stations studied, 2006 (Year 4 in the scalogram) was chosen because the percentage of lost data for all stations is low. The analysis of the normalized cross-correlation of the average wind speed between stations CY-SZ (Figure 6a) and CY-CA (Figure 6b) shows that the Pearson correlation coefficient is close to 1 for both sets of stations; for CY-SZ, the correlation coefficient is 0.8075, and for CY-CA, it is 0.8979. This means that the data sets for stations SZ and CA are closely related to those of CY; in addition, the lag is zero (Figure 6) because the three stations are located near each other (Figure 1) in this case of study. The mean distance is 14 km, with a minimum of 9 km and a maximum of 18 km.

In addition, the monthly correlations between stations CY-SZ and CY-CA was analysed, and it was found that the Pearson coefficient is also close to 1 (Table 1) and that station CA has a stronger relationship with station CY. Next, the lost data from station CY were interpolated and the weighting coefficients for each year were calculated using the 52,560 samples from each year. Table 2 shows the yearly median rates obtained. The median was used to relate the stations because it is less sensitive to outliers or biased values. The analysis was carried out for complete years, but the data series of the CA station only reaches until August

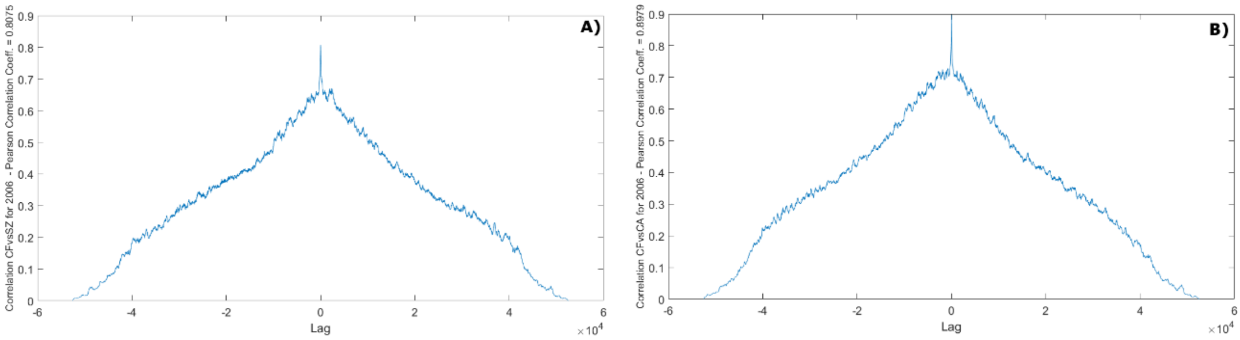

Figure 6: Cross-correlation of the 2006 values from Collado de Yuste station with a) Solana de Zapatero and b) Calar Alto

2010 so, from that moment there is no relationship of CY with CA. The CWT was applied to the completed data for station CY, and the scalogram shown in Figure 7 was obtained. The results obtained show a constant and natural pattern, this time with periodicity and continuity over the 11 years.

4. Discussion

Figure 3 reveals that the scalogram shows zero magnitude between 2008 and 2009 due to the lack of data. Once the data are filled with the right technique (see Figure 7), the magnitudes of the scalogram are different from zero and present values similar to the rest of the previous data series. Figures 8 and 9 are obtained by applying the FFT to the original data and the new dataset of station CY, respectively, with the PSD (power spectrum density) represented on the Y-axis in logarithmic scale and the frequency (Hz) on the X-axis, also in logarithmic scale. The sampling rate is $fs = \frac{1}{T} \approx 0.0017$Hz because each sample was recorded at 10-minute intervals. Figure 8 shows that the resulting cycles do not correspond to a natural cycle which are summarized in Table 3. Figure 9 shows how the natural cycle's patterns of the completed series are the expected, e.g. 24 hours, 7 days, 28 days, or 356 days. Now with the correct data verified by this technique, wind forecasts can be performed and used for instance for the placement of wind farms where wind data are indispensable for their study [79]. It is shown that the peak of year cycle is varied to 401 days, this is caused by the loss of 3% of the data in the total of the 11 years. Therefore a data loss of 2.5% can give an error of 9% in the computation of an annual cycle. Then a threshold of data loss greater than 2.5% gives unacceptable results for using the FFT technique.

Because the FFT is not sensitive to the temporal variation of wind speed, the WT can be used for this time-frequency analysis, as shown in Figure 10. By using the scalogram, it is possible to reveal the areas with most energy, represented by a colour scale according to the magnitude value, which are those that have significant oscillations with annual periods. The upper (11.5 µHz) and lower (0.0317 µHz) horizontal bands demarked

	Pearson Coefficient	
Month	**CY with SZ**	**CY with CA**
January	0.7789	0.7804
February	0.8636	0.8487
March	0.8393	0.9237
April	0.7616	0.9225
May	0.7951	0.9334
June	0.8440	0.9115
July	0.8316	0.9308
August	0.8236	0.9057
September	0.8358	0.9271
October	0.7922	0.9146
November	0.7234	0.8947
December	0.8430	0.9041

Table 1: Pearson coefficient for station CY for 2006

	Median rate	
Year	**Median CY/ SZ**	**Median CY/ CA**
2002	2.6250	0.8424
2003	2.8750	0.8038
2004	2.7222	0.8038
2005	2.5000	0.8193
2006	2.6429	0.8027
2007	2.4348	0.7836
2008	2.3529	0.8038
2009	2.7500	0.8010
2010	2.9286	-
2011	2.8500	-
2012	2.7857	-

Table 2: Median rate (2002-2012)

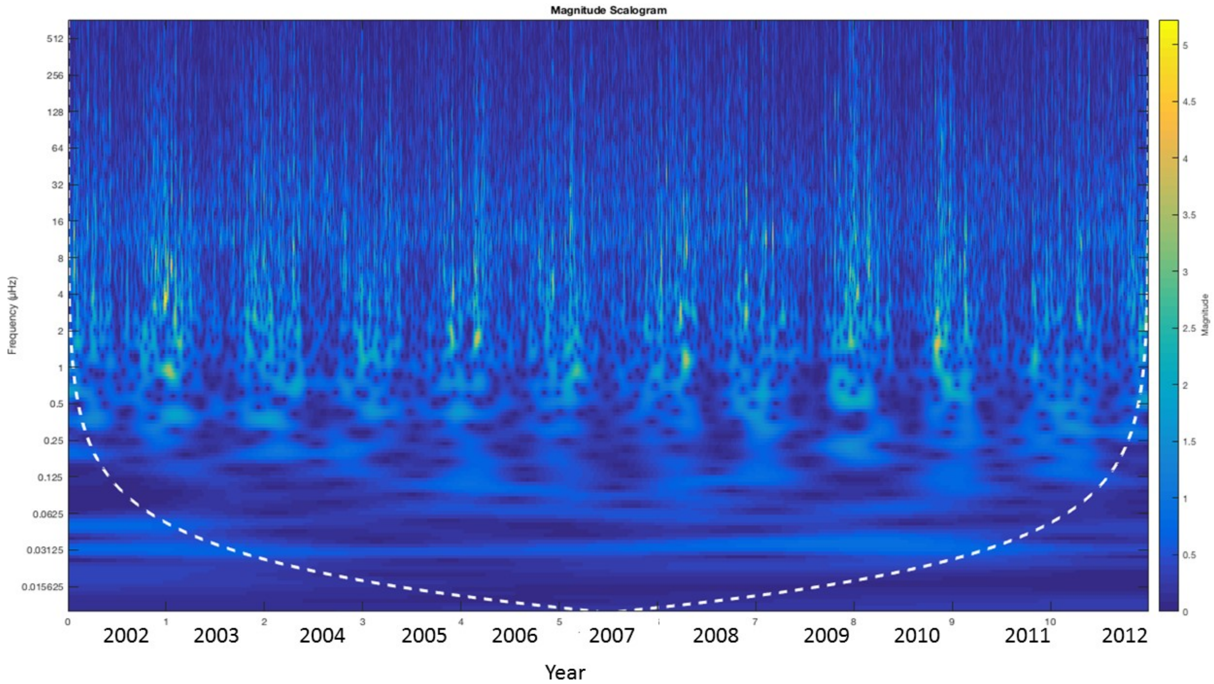

Figure 7: The WT power spectrum of wind speed of the new CY containing data aggregations

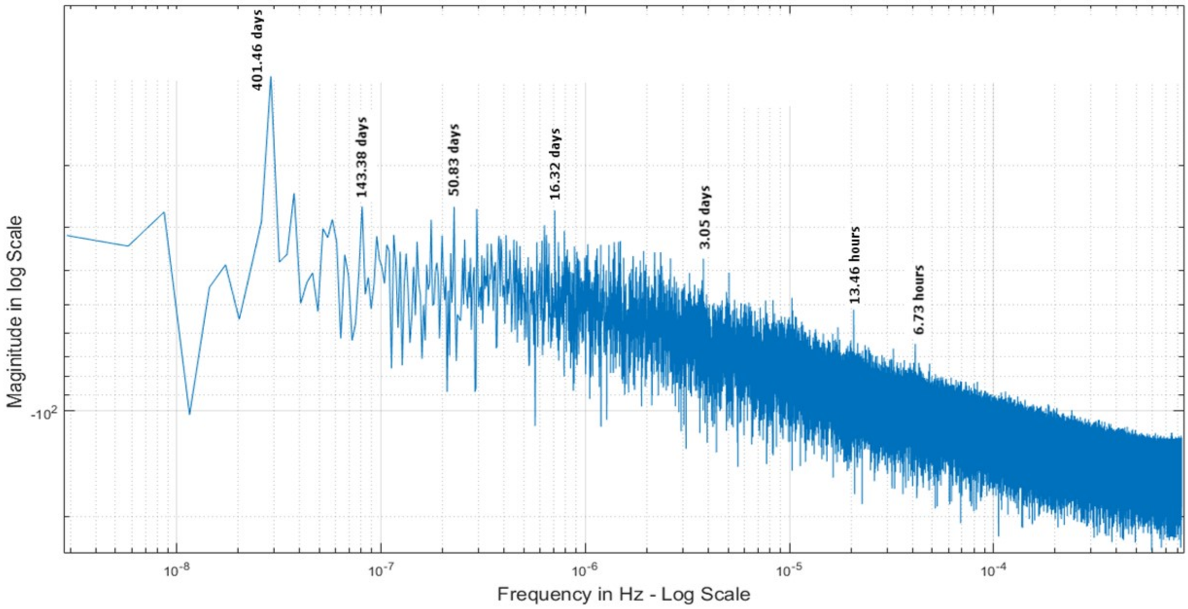

Figure 8: Original values of wind speed PSD for CY (2002-2012)

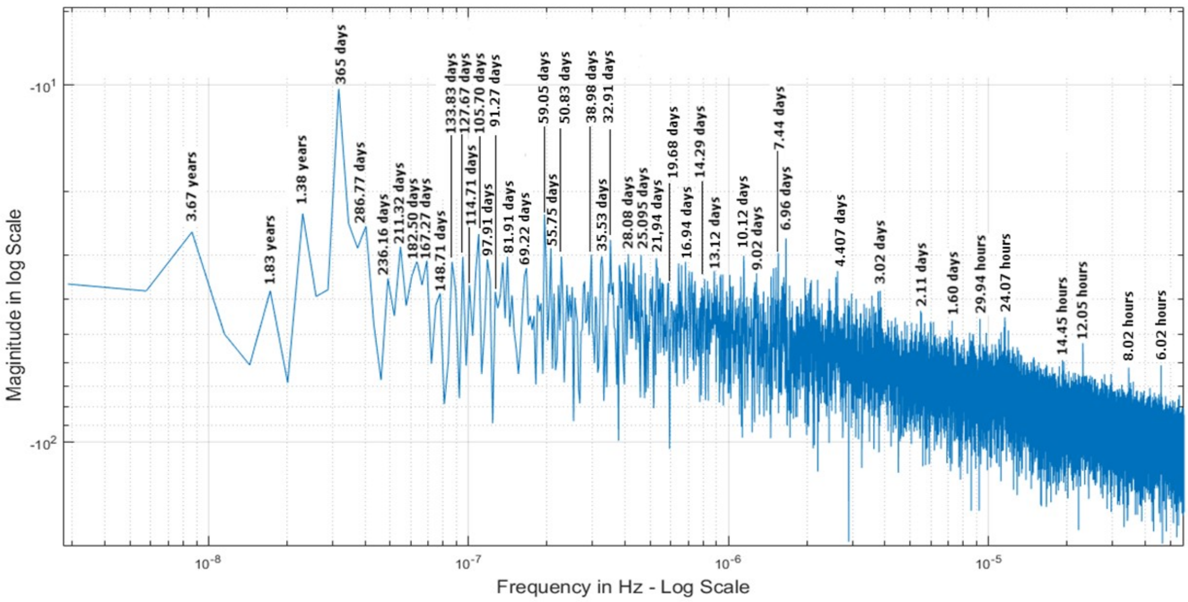

Figure 9: New data set of wind speed PSD for CY (2002-2012)

Frequency	**units**	**Cicle**
365.00	days	1 year
182.50	days	1/2 year
91.27	days	1/4 year
28.08	days	Moon cycle
6.96	days	1 Week
24.07	hours	1 day
12.05	hours	1/2 day
8.02	hours	1/3 day
6.02	hours	1/4 day

Table 3: The main cycles (Figure 9) of the average wind speed obtained with the new dataset for station CY

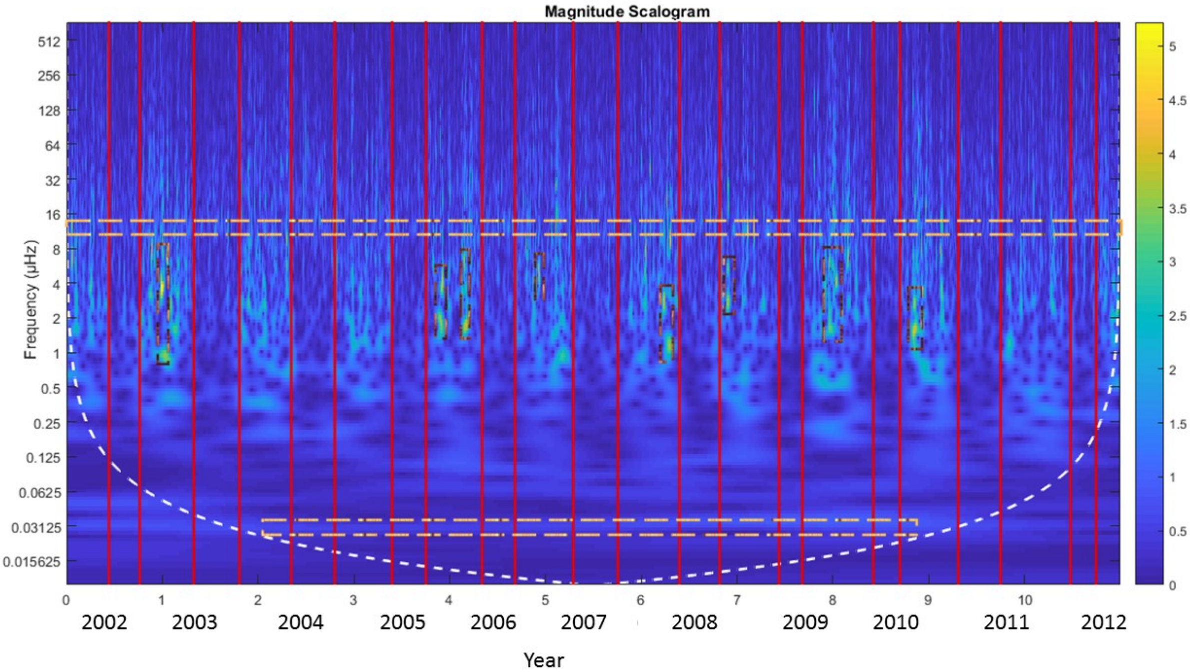

Figure 10: Highlights in the main activities on the windspeed scalogram

with horizontal dashed lines represent the periodicity during a day and a year, respectively. The vertical red lines are the time intervals with most activity, fluctuating from April to June, and from September to November during the eleven years with frequencies ranging between 1 and 8 µHz. Wind speed changes are negligible between the months of December and March, depending on the year. The highest wind speeds are also demarked with dashed squares in the months of December 2002 to January 2003, November to December 2005, February to March 2006, December 2006, March to May 2008, September to December 2008, November 2009 to January 2010, and October 2010 to January 2011.

5. Conclusions

The interpolation of lost wind speed data of a weather station using data obtained from the interpolation of data from other stations can lead to incorrect results that are not easily detected. The behaviour of wind, as well as the cycles and periods of its cycles, are not easy to visualize in graphs. This complexity is partially solved by using the FFT that details the cycles; but this technique does not help if the data series is incomplete. A data loss of 2.5% can give an error of 9% in the computation of an annual cycle. In this case, it is necessary to complete the lost data using an interpolation technique. The missing results can be seen using the scalogram of the WT, and if these results are completed using data from the station itself, the resulting periodicity is not natural. When the data have been interpolated in a satisfactory manner using data from nearby stations, as demonstrated by their correlation coefficient, the scalogram shows the periodicity of the data with an apparent naturalness. Therefore, the use of the WT and its representation through scalograms allows detecting the validity of the interpolation of lost wind data. In addition, the scalograms provide additional information on the variables studied, for example, the periods of highest intensity. As a general conclusion of this study, the WT has proven to be a time and frequency analysis tool that reveals the seasonal pattern of wind speed in periodograms at various scales.

6. Acknowledgements

The authors would like to acknowledge the Andaluz Institute of Agricultural Research and Training (IFAPA in Spanish) for the data and the Ibero-American Postgraduate University Association (AUIP) for its support in this research.

References

[1] Q. Hernández-Escobedo, A.-J. Perea-Moreno, F. Manzano-Agugliaro, Wind energy research in mexico, Renewable Energy 123 (2018) 719–729.

[2] F. Manzano-Agugliaro, A. Alcayde, F. Montoya, A. Zapata-Sierra, C. Gil, Scientific production of renewable energies worldwide: an overview, Renewable and Sustainable Energy Reviews 18 (2013) 134–143.

[3] D. Kiplangat, K. Asokan, K. Kumar, Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition, Renewable Energy 93 (2016) 38–44. doi:10.1016/j.renene.2016.02.054.

[4] J. Clancy, F. Gaffney, J. Deane, J. Curtis, B. Ó Gallachóir, Fossil fuel and co¡inf¿2¡/inf¿ emissions savings on a high renewable electricity system - a single year case study for ireland, Energy Policy 83 (2015) 151–164. doi:10.1016/j.enpol.2015.04.011.

[5] M. Nicholson, T. Biegler, B. Brook, How carbon pricing changes the relative competitiveness of low-carbon baseload generating technologies, Energy 36 (1) (2011) 305–313. doi:10.1016/j.energy.2010.10.039.

[6] T. Gentils, L. Wang, A. Kolios, Integrated structural optimisation of offshore wind turbine support structures based on finite element analysis and genetic algorithm, Applied Energy 199 (2017) 187–204. doi:10.1016/j.apenergy.2017.05.009.

[7] H. Chitsaz, N. Amjady, H. Zareipour, Wind power forecast using wavelet neural network trained by improved clonal selection algorithm, Energy Conversion and Management 89 (2015) 588–598. doi:10.1016/j.enconman.2014.10.001.

[8] J. Dai, Y. Tan, W. Yang, L. Wen, X. Shen, Investigation of wind resource characteristics in mountain wind farm using multiple-unit scada data in chenzhou: A case study, Energy conversion and management 148 (2017) 378–393.

[9] J. Yan, T. Ouyang, Advanced wind power prediction based on data-driven error correction, Energy Conversion and Management 180 (2019) 302–311.

[10] I. Okumus, A. Dinler, Current status of wind energy forecasting and a hybrid method for hourly predictions, Energy Conversion and Management 123 (2016) 362–371.

[11] D. E. Claridge, H. Chen, Missing data estimation for 1–6 h gaps in energy use and weather data using different statistical methods, International journal of energy research 30 (13) (2006) 1075–1091.

[12] A. T. Ihler, S. Kirshner, M. Ghil, A. W. Robertson, P. Smyth, Graphical models for statistical inference and data assimilation, Physica D: Nonlinear Phenomena 230 (1-2) (2007) 72–87.

[13] W. Ji, C. Chan, J. Loh, F. Choo, L. Chen, Solar radiation prediction using statistical approaches, in: Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on, IEEE, 2009, pp. 1–5.

[14] J. G. L. Segura, Influence of mountains on extreme rainfall in semi-arid areas of europe: The sierra de los filabres (se spain) a case of study, Journal of Food, Agriculture and Environment 11 (3-4) (2013) 2458–2460.

[15] Q. Hernández-Escobedo, F. Manzano-Agugliaro, A. Zapata-Sierra, The wind power of mexico, Renewable and Sustainable Energy Reviews 14 (9) (2010) 2830–2840.

[16] Q. Hernandez-Escobedo, R. Saldaña-Flores, E. Rodríguez-García, F. Manzano-Agugliaro, Wind energy resource in northern mexico, Renewable and Sustainable Energy Reviews 32 (2014) 890–914.

[17] K. G. Hubbard, J. You, Sensitivity analysis of quality assurance using the spatial regression approach—a case study of the maximum/minimum air temperature, Journal of Atmospheric and Oceanic Technology 22 (10) (2005) 1520–1530.

[18] M. Ekhtesasi, Z. Gohari, Determining area affected by dust storms in different wind speeds, using satellite images, Desert 17 (2) (2012) 193–202.

[19] T. McVicar, L. Li, T. Van Niel, M. Hutchinson, X. Mu, Z. Liu, Spatially distributing 21 years of monthly hydrometeorological data in china: Spatio-temporal analysis of fao-56 crop reference evapotranspiration and pan evaporation in the context of climate change.

[20] W. Luo, M. Taylor, S. Parker, A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from england and wales, International journal of climatology 28 (7) (2008) 947–959.

[21] N. Cressie, Fitting variogram models by weighted least squares, Journal of the International Association for Mathematical Geology 17 (5) (1985) 563–586.

[22] T. Arslan, S. Acitas, B. Senoglu, Generalized lindley and power lindley distributions for modeling the wind speed data, Energy Conversion and Management 152 (2017) 300–311.

[23] M. M. Alam, S. Rehman, L. M. Al-Hadhrami, J. P. Meyer, Extraction of the inherent nature of wind speed using wavelets and fft, Energy for Sustainable Development 22 (2014) 34–47.

[24] Y. Falamarzi, N. Palizdan, Y. Huang, T. Lee, Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (wnns), Agricultural Water Management 140 (2014) 26–36. doi:10.1016/j.agwat.2014.03.014.

[25] F. Chellali, A. Khellaf, A. Belouchrani, Wavelet spectral analysis of the temperature and wind speed data at adrar, algeria, Renewable Energy 35 (6) (2010) 1214–1219.

[26] I. Janajreh, L. Su, F. Alan, Wind energy assessment: Masdar city case study, Renewable energy 52 (2013) 8–15.

[27] S. Avdakovic, A. Lukac, A. Nuhanovic, M. Music, Wind speed data analysis using wavelet transform, International Journal of Engineering and Applied Sciences 7 (2011) 116–120.

[28] H. Liu, X.-w. Mi, Y.-f. Li, Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and elman neural network, Energy Conversion and Management 156 (2018) 498–514.

[29] N. D. Anh, et al., Orthogonal-based wavelet analysis of wind turbulence and correlation between turbulence and forces, Vietnam Journal of Mechanics 29 (2) (2007) 73–82.

[30] X.-w. Mi, H. Liu, Y.-f. Li, Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm, Energy Conversion and Management 151 (2017) 709–722.

[31] H. Liu, X. Mi, Y. Li, Smart deep learning based wind speed prediction model using wavelet packet decomposition,

convolutional neural network and convolutional long short term memory network, Energy Conversion and Management 166 (2018) 120–131.

[32] J. Morlet, G. Arens, E. Fourgeau, D. Glard, Wave propagation and sampling theory—part i: Complex signal and scattering in multilayered media, Geophysics 47 (2) (1982) 203–221.

[33] A. Grossmann, J. Morlet, Decomposition of hardy functions into square integrable wavelets of constant shape, SIAM journal on mathematical analysis 15 (4) (1984) 723–736.

[34] A. Graps, An introduction to wavelets, IEEE computational science and engineering 2 (2) (1995) 50–61.

[35] W. Wang, P. McFadden, Application of wavelets to gearbox vibration signals for fault detection, Journal of sound and vibration 192 (5) (1996) 927–939.

[36] K. L. Butler-Purry, M. Bagriyanik, Characterization of transients in transformers using discrete wavelet transforms, IEEE Transactions on Power Systems 18 (2) (2003) 648–656.

[37] G. Strang, Wavelet transforms versus fourier transforms, Bulletin of the American Mathematical Society 28 (2) (1993) 288–305. doi:10.1090/S0273-0979-1993-00390-2.

[38] D. Bayram, S. Seker, Redundancy-based predictive fault detection on electric motors by stationary wavelet transform, IEEE Transactions on Industry Applications 53 (3) (2017) 2997–3004. doi:10.1109/TIA.2016.2622231.

[39] S. Santoso, E. J. Powers, W. Grady, Power quality disturbance data compression using wavelet transform methods, IEEE Transactions on Power Delivery 12 (3) (1997) 1250–1257.

[40] S. G. Chang, B. Yu, M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, IEEE transactions on image processing 9 (9) (2000) 1532–1546.

[41] L.-M. Reissell, Wavelet multiresolution representation of curves and surfaces, Graphical Models and Image Processing 58 (3) (1996) 198–217.

[42] M. Kumar, R. B. Pachori, U. R. Acharya, An efficient automated technique for cad diagnosis using flexible analytic wavelet transform and entropy features extracted from hrv signals, Expert Systems with Applications 63 (2016) 165–172.

[43] H.-D. Lin, Automated visual inspection of ripple defects using wavelet characteristic based multivariate statistical approach, Image and Vision Computing 25 (11) (2007) 1785–1801.

[44] C.-K. Lin, Nonsingular terminal sliding mode control of robot manipulators using fuzzy wavelet networks, IEEE Transactions on Fuzzy Systems 14 (6) (2006) 849–859.

[45] M. Santhosh, C. Venkaiah, D. V. Kumar, Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction, Energy Conversion and Management 168 (2018) 482–493.

[46] H. Singh, R. Mehra, Discrete wavelet transform method for high flux n-γ discrimination with liquid scintillators, IEEE Transactions on Nuclear Science 64 (7) (2017) 1927–1933. doi:10.1109/TNS.2017.2708602.

[47] T.-P. Chang, F.-J. Liu, H.-H. Ko, M.-C. Huang, Oscillation characteristic study of wind speed, global solar radiation and air temperature using wavelet analysis, Applied Energy 190 (2017) 650–657. doi:10.1016/j.apenergy.2016.12.149.

[48] H. Liu, H.-q. Tian, Y.-f. Li, Comparison of new hybrid feemd-mlp, feemd-anfis, wavelet packet-mlp and wavelet packet-anfis for wind speed predictions, Energy Conversion and Management 89 (2015) 1–11.

[49] S.-W. Fei, Y. He, Wind speed prediction using the hybrid model of wavelet decomposition and artificial bee colony algorithm-based relevance vector machine, International Journal of Electrical Power and Energy Systems 73 (2015) 625–631. doi:10.1016/j.ijepes.2015.04.019.

[50] H. Shao, H. Wei, X. Deng, S. Xing, Short-term wind speed forecasting using wavelet transformation and adaboosting neural networks in yunnan wind farm, IET Renewable Power Generation 11 (4) (2017) 374–381. doi:10.1049/iet-rpg.2016.0118.

[51] I. Mladenović, D. Marković, M. Milovančević, M. Nikolić, Extreme learning approach with wavelet transform function for

forecasting wind turbine wake effect to improve wind farm efficiency, Advances in Engineering Software 96 (2016) 91–95. doi:10.1016/j.advengsoft.2016.02.011.

[52] M. Alam, S. Rehman, L. Al-Hadhrami, J. Meyer, Extraction of the inherent nature of wind speed using wavelets and fft, Energy for Sustainable Development 22 (1) (2014) 34–47. doi:10.1016/j.esd.2014.02.004.

[53] M. Alam, M. Moriya, H. Sakamoto, Aerodynamic characteristics of two side-by-side circular cylinders and application of wavelet analysis on the switching phenomenon, Journal of Fluids and Structures 18 (3-4) (2003) 325–346. doi:10.1016/j.jfluidstructs.2003.07.005.

[54] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, E. Feitosa, A review on the young history of the wind power short-term prediction, Renewable and Sustainable Energy Reviews 12 (6) (2008) 1725–1744. doi:10.1016/j.rser.2007.01.015.

[55] H. Dai, Z. Zheng, W. Wang, A new fractional wavelet transform, Communications in Nonlinear Science and Numerical Simulation 44 (2017) 19–36. doi:10.1016/j.cnsns.2016.06.034.

[56] P. Sukiennik, J. Białasiewicz, Cross-correlation of bio-signals using continuous wavelet transform and genetic algorithm, Journal of Neuroscience Methods 247 (2015) 13–22. doi:10.1016/j.jneumeth.2015.03.002.

[57] H. Li, F. Xu, H. Liu, X. Zhang, Incipient fault information determination for rolling element bearing based on synchronous averaging reassigned wavelet scalogram, Measurement: Journal of the International Measurement Confederation 65 (2015) 1–10. doi:10.1016/j.measurement.2014.12.032.

[58] Y. He, Y. Liu, Experimental research into time-frequency characteristics of cavitation noise using wavelet scalogram, Applied Acoustics 72 (10) (2011) 721–731. doi:10.1016/j.apacoust.2011.03.008.

[59] M. Tokumitsu, K. Hasegawa, Y. Ishida, Toward resilient sensor networks with spatiotemporal interpolation of missing data: An example of space weather forecasting, Procedia Computer Science 60 (1) (2015) 1585–1594. doi:10.1016/j.procs.2015.08.268.

[60] R. Razavi-Far, E. Zio, V. Palade, Efficient residuals pre-processing for diagnosing multi-class faults in a doubly fed induction generator, under missing data scenarios, Expert Systems with Applications 41 (14) (2014) 6386–6399. doi:10.1016/j.eswa.2014.03.056.

[61] Z. Yang, Y. Liu, C. Li, Interpolation of missing wind data based on anfis, Renewable Energy 36 (3) (2011) 993–998. doi:10.1016/j.renene.2010.08.033.

[62] P. Petrović, N. Damljanović, New procedure for harmonics estimation based on hilbert transformation, Electrical Engineering 99 (1) (2017) 313–323. doi:10.1007/s00202-016-0434-x.

[63] S. Kanarachos, S.-R. Christopoulos, A. Chroneos, M. Fitzpatrick, Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and hilbert transform, Expert Systems with Applications 85 (2017) 292–304. doi:10.1016/j.eswa.2017.04.028.

[64] C. Yu, Y. Li, M. Zhang, An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on elman neural network, Energy Conversion and Management 148 (2017) 895–904. doi:10.1016/j.enconman.2017.05.063.

[65] S. Fornero, N. Kehtarnavaz, M. Swaminadham, D. A. Phillips, Fourier and wavelet transform features for whirl tower diagnostics 4 (1999) 2267–2270 vol.4. doi:10.1109/ICASSP.1999.758389.

[66] F. A. Jowder, Weibull and rayleigh distribution functions of wind speeds in kingdom of bahrain, Wind Engineering 30 (5) (2006) 439–445.

[67] S. Ahmed, H. Mahammed, A statistical analysis of wind power density based on the weibull and ralyeigh models of “penjwen region” sulaimani/iraq, Jordan Journal of Mechanical and Industrial Engineering 6 (2) (2012) 135–140.

[68] Z. German-Sallo, M. German-Sallo, Multiscale analysing methods in electrocardiogram signal processing and interpretation, Procedia Engineering 181 (2017) 583–587. doi:10.1016/j.proeng.2017.02.437.

[69] C. Jung, D. Schindler, Sensitivity analysis of the system of wind speed distributions, Energy Conversion and Management 177 (2018) 376–384.

[70] H. Zheng, Z. Li, X. Chen, Gear fault diagnosis based on continuous wavelet transform, Mechanical systems and signal processing 16 (2-3) (2002) 447–457.

[71] R. R. Singh, R. Mishra, Benefits of dual tree complex wavelet transform over discrete wavelet transform for image fusion, International Journal for Innovative Research in Science and Technology 1 (11) (2015) 259–263.

[72] A. Harkat, R. Benzid, L. Saidi, Features extraction and classification of ecg beats using cwt combined to rbf neural network optimized by cuckoo search via levy flight, in: Electrical Engineering (ICEE), 2015 4th International Conference on, IEEE, 2015, pp. 1–4.

[73] S. Nair, K. Paul Joseph, Wavelet based electroretinographic signal analysis for diagnosis, Biomedical Signal Processing and Control 9 (1) (2014) 37–44. doi:10.1016/j.bspc.2013.09.008.

[74] S. J. Watson, B. J. Xiang, W. Yang, P. J. Tavner, C. J. Crabtree, Condition monitoring of the power output of wind turbine generators using wavelets, IEEE Transactions on Energy Conversion 25 (3) (2010) 715–721.

[75] H. Liu, W. Huang, S. Wang, Z. Zhu, Adaptive spectral kurtosis filtering based on morlet wavelet and its application for signal transients detection, Signal Processing 96 (2014) 118–124.

[76] M. X. Cohen, A better way to define and describe morlet wavelets for time-frequency analysis, bioRxiv (2018) 397182.

[77] S. Yaman, N. Ozturk, U. Çomelekoglu, E. Degirmenci, Determination of dichlorvos effect on uterine contractility using wavelet transform, IRBM 37 (5-6) (2016) 264–270. doi:10.1016/j.irbm.2016.09.002.

[78] A. Grinsted, J. Moore, S. Jevrejeva, Application of the cross wavelet transform and wavelet coherence to geophysical times series, Nonlinear Processes in Geophysics 11 (5-6) (2004) 561–566.

[79] F. G. Montoya, F. Manzano-Agugliaro, S. López-Márquez, Q. Hernández-Escobedo, C. Gil, Wind turbine selection for wind farm layout using multi-objective evolutionary algorithms, Expert Systems with Applications 41 (15) (2014) 6585–6595.