

# Applying Wavelet Filters in Wind Forecasting Methods

José A. Domínguez-Navarro <sup>1,\*</sup>, Tania B. Lopez-Garcia <sup>1</sup> and Sandra Minerva Valdivia-Bautista <sup>2</sup>

<sup>1</sup> Department of Electrical Engineering, EINA, University of Zaragoza, 50018 Zaragoza, Spain; 782726@unizar.es

<sup>2</sup> Centro Universitario de Ciencias e Ingenierías (CUCEI), Universidad de Guadalajara (UDG), Guadalajara 44160, Mexico; minerva.valdivia@academicos.udg.mx

\* Correspondence: jadona@unizar.es

**Abstract:** Wind is a physical phenomenon with uncertainties in several temporal scales, in addition, measured wind time series have noise superimposed on them. These time series are the basis for forecasting methods. This paper studied the application of the wavelet transform to three forecasting methods, namely, stochastic, neural network, and fuzzy, and six wavelet families. Wind speed time series were first filtered to eliminate the high-frequency component using wavelet filters and then the different forecasting methods were applied to the filtered time series. All methods showed important improvements when the wavelet filter was applied. It is important to note that the application of the wavelet technique requires a deep study of the time series in order to select the appropriate family and filter level. The best results were obtained with an optimal filtering level and improper selection may significantly affect the accuracy of the results.

**Keywords:** wavelet transforms; forecasting methods; wind energy



**Citation:** Domínguez-Navarro, J.A.; Lopez-Garcia, T.B.; Valdivia-Bautista, S.M. Applying Wavelet Filters in Wind Forecasting Methods. *Energies* **2021**, *14*, 3181. <https://doi.org/10.3390/en14113181>

Academic Editor: Simon J. Watson

Received: 18 April 2021

Accepted: 26 May 2021

Published: 29 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

When the penetration of wind power into the network reaches a certain level, system operators have difficulties in balancing generation with demand. To help address this issue, it is necessary to apply forecasting methods to estimate the wind power generated in the next few hours and days.

Several methods have been used to forecast the wind speed: stochastic methods such as the AR [1] and ARIMA [2–4] process or heuristic methods such as the Kalman filter [5,6], neural networks [7–9], and neuro-fuzzy systems [10,11]. Among all methods, neural networks are the most widely used by researchers.

It is difficult to compare different methods if they do not use the same dataset and the same performance indexes. A typical approach for comparison is to use the persistence model as the reference [12]. Table 1 illustrates the improvement achieved by different forecasting methods compared to the persistence method (although different forecast horizons are used, the list gives an idea of the range of improvement). These improvements are rarely over 25%.

Wind speed series have considerable uncertainty because of weather fluctuations and the added instrument uncertainty. This uncertainty and noise make it difficult to improve the forecasts. There are several strategies to process the data [13–15]. Some authors have used wavelet transforms [16] to process the time series. Most authors [17–31] have used wavelets to decompose the time series into sub-series, called approximation and details; applied the forecasting method to each sub-series, and finally, summarized the forecasting results to obtain the final solution. The advantage of this method comes from the sub-series having an improved performance with respect to the original series. A few authors [32–37] have used other wavelet filtering techniques to eliminate the high frequency variations and smooth the time series. In all papers, the authors selected the wavelet family and decomposition level without too much justification. For example, the cited authors only

used one wavelet family. These works used the wavelet transform as an auxiliary technique and did not study them at sufficient depth.

**Table 1.** Review of forecasting methods.

Authors	Forecast Method	Forecast Horizon	Improvement vs. Persistence Model
[1]	AR	–	–
[2]	ARMA	1–10 h	12–20%
[3]	ARMA	20''–10''	5–12%
[4]	ARMA	6 h	19%
[5]	Kalman filter	1'	4–10%
[6]	Kalman filter	–	–
[7]	ANN	1'–10'	11–8%
[8]	Spatial/ANN	2–2 h	15–25%
[9]	ANN	1'–1 h	–
[10]	Neuro-fuzzy	6–36 h	46%
[11]	Neuro-fuzzy	10'	5%

In this paper, the wavelet transform was analyzed thoroughly. This work demonstrated that the selection of the wavelet family and decomposition level were far more important than they have been given credit for thus far. The improvement obtained was greater than that achieved with most new forecasting methods. The result was applied to the three main forecasting methods currently used, namely statistical, neural network, and fuzzy methods. These were applied to several forecast horizons and sample times. In all cases, the results obtained were improved for each method when the optimal wavelet filter was applied. Finally, the main contribution of the paper is to highlight the importance of data processing and to propose it as an additional phase in the forecasting method, so that both steps are optimized together.

The rest of this paper is structured as follows. Section 2 explains the basic concepts of the wavelet transform. Section 3 presents the different forecasting approaches. Section 4 describes the forecasting approach proposed. In Section 5, the comparison criteria to evaluate the improvement of each method are explained. Section 6 presents the results for the different methods considered. Finally, Section 7 draws the main conclusions of this research.

## 2. Basic Concepts of the Wavelet Transform

Fourier analysis is commonly used to help analyze different types of signals. With this method, a signal  $f(t)$  is expressed as a linear decomposition of real-valued functions of  $t$ , as shown in Equation (1),

$$f(t) = \sum_k a_k \phi_k(t) \quad (1)$$

where  $a_k$  are the real-valued expansion coefficients and  $\phi_k(t)$  are a set of real-valued functions of  $t$  called the expansion set. In Fourier series, these are  $\sin(k\omega_0 t)$  and  $\cos(k\omega_0 t)$  with frequencies of  $k\omega_0$ .

### 2.1. Wavelet Transform

An introduction to the wavelet transform can be found in [16]. In Equation (2), the signal is already decomposed into coefficients  $a_{j,k}$  and functions  $\Psi_{j,k}(t)$ , which depend on parameters  $j$  and  $k$ ,

$$f(t) = \sum_k \sum_j a_{j,k} \Psi_{j,k}(t) \quad (2)$$

where  $\Psi_{j,k}$  are the wavelet expansion functions and  $a_{j,k}$  is the discrete wavelet transform of  $f(t)$ , or the set of expansion coefficients.

The wavelet expansion functions, or family of wavelets are generated from a mother wavelet  $\Psi(t)$  by scaling and translation:

$$\Psi_{j,k}(t) = 2^{j/2}\Psi(2^j t - k) \tag{3}$$

where the parameter  $k$  translates the function and the parameter  $j$  scales it. Figure 1 shows the translating and scaling operations of the function  $\Psi(2^j t - k)$ .

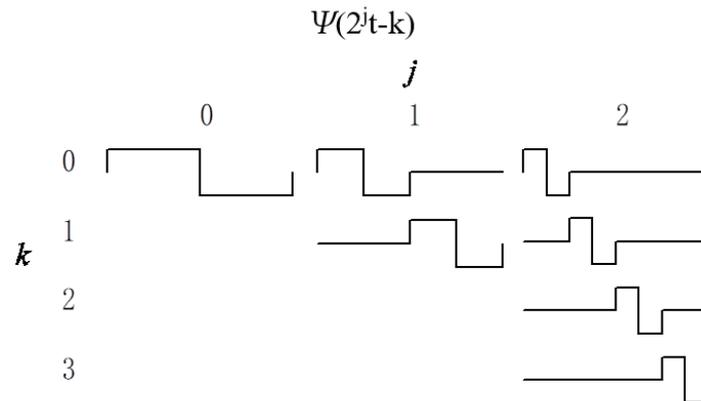


Figure 1. Translating and scaling operations of  $\Psi(2^j t - k)$  function.

### 2.2. Multi-Resolution Formulation of Wavelet Systems

In multi-resolution analysis, the resolution of the approximation of  $f(t)$  depends on the choice of  $j$  in Equation (2). For a value of  $j = j_0$ , the equation is:

$$f(t) = \sum_k c_{j_0,k} \varphi_{j_0,k}(t) \tag{4}$$

For low values of  $j$ , the approximation of  $f(t)$  can represent only coarse information. In multi-resolution formulation,  $\varphi_{j_0,k}(t)$  are called scaling functions. If we want to represent detailed information, then high values of  $j$  are required.

However, there is another way to describe a signal with better resolution without increasing  $j$ . This new approach consists of describing the differences between the approximation and the original signal with a combination of other functions called wavelets  $\Psi_{j,k}(t)$  and the coefficients  $d_{j,k}$ , as shown in Equation (5). The parameters  $k$  and  $j$  indicate the translation and scaling of the function.

$$f(t) = \sum_k c_{j_0,k} \varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{J-1} d_{j,k} \psi_{j,k}(t) \tag{5}$$

$$= \sum_k c_{j_0}(k) 2^{j_0/2} \varphi(2^{j_0} t - k) + \sum_k \sum_{j=j_0}^{J-1} d_j(k) 2^{j/2} \psi(2^j t - k)$$

There are several packets of scaling functions  $\varphi(t)$  and wavelets  $\Psi(t)$ , as shown in Equation (6) (see Figure 2), which were chosen depending on the signal that has to be approximated.

$$\varphi(t) = \sum_n h_0(n) \sqrt{2} \varphi(2t - n) \tag{6}$$

$$\psi(t) = \sum_n h_1(n) \sqrt{2} \psi(2t - n)$$

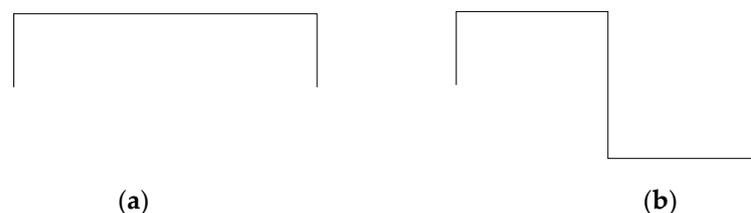


Figure 2. Haar packet. (a) Haar scaling function,  $\varphi(t)$ ; (b) Haar wavelet function,  $\Psi(t)$ .

In Equation (6), the coefficients  $h_0(n)$  and  $h_1(n)$  with  $n \in \mathbb{Z}$ , are a sequence of real numbers called filter coefficients. The process is similar to digital filters, where  $h_0(m - 2k)$  acts as a low-pass filter and  $h_1(m - 2k)$  acts as a high-pass filter. Figure 3 shows the decomposition process of  $c_j$  in:  $c_{j+1}$  (low frequency) and  $d_{j+1}$  (high frequency).

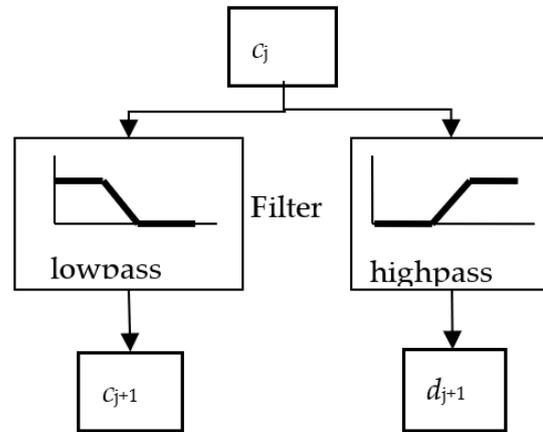


Figure 3. Decomposition process.

The  $j+1$  level scaling coefficients are:

$$\begin{aligned} c_{j+1}(k) &= \sum_m h_0(m - 2k)c_j(m) \\ d_{j+1}(k) &= \sum_m h_1(m - 2k)c_j(m) \end{aligned} \tag{7}$$

These expressions represent the approximation and details of the signal for a  $j + 1$  level scaling, and  $m = 2k + n$  is a sequence of real numbers.

This process can be repeated iteratively to reduce the high-frequency component as shown in Figure 4.

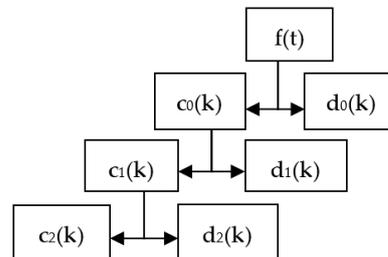
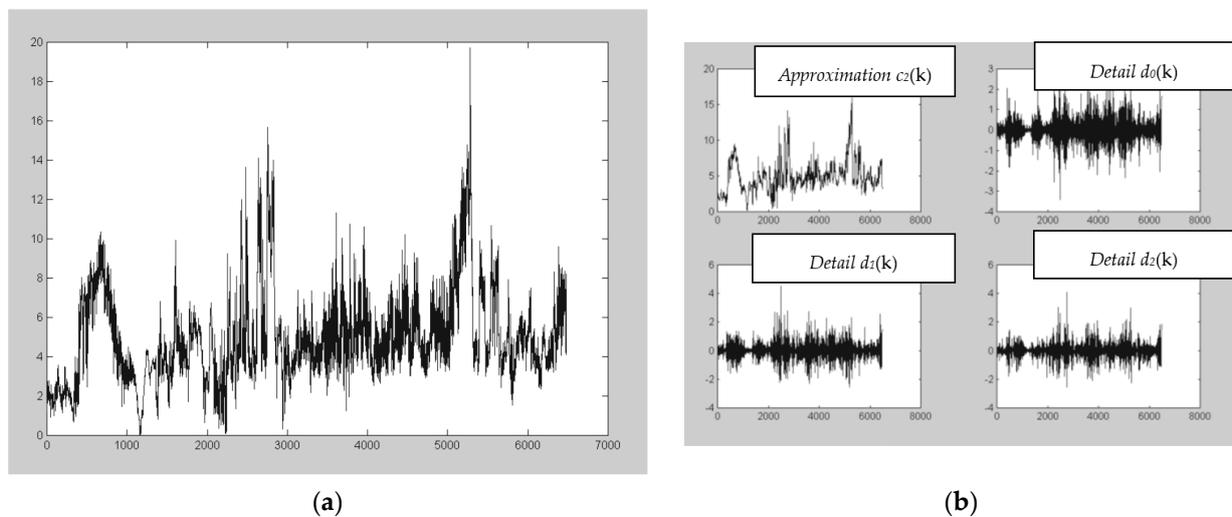


Figure 4. Iterative decomposition.

In Figure 5b, we can see the approximation  $c_2(k)$  and the details  $d_2(k)$ ,  $d_1(k)$ , and  $d_0(k)$  of the original signal  $f(t)$  are shown in Figure 5a.

Wind speed time series have a high-frequency component due to wind gusts, measurement errors, and random events as well as a low-frequency component with slower variation. The high-frequency component of the signal introduces a lot of noise into forecasting methods, causing them to perform poorly. If this component is eliminated and the forecasting methods are applied to an approximation with only the low frequency component, improved results can be obtained.



**Figure 5.** Approximation and details of a signal. (a) Original signal; (b) Approximation  $c_2(k)$  and details  $d_2(k)$ ,  $d_1(k)$  and  $d_0(k)$ .

### 2.3. Wavelet Families

There is a large number of wavelets. The selection of the wavelet function depends on the problem and the properties of the wavelet function [16]. The main properties are its region of support and the number of vanishing moments. The region of support affects its localization capabilities, whereas the vanishing moments limit the ability of the wavelet to represent the information of a signal. In this paper, the wavelet families used were: Haar, Daubechies, Symlet, Coiflet, BiorSplines, and Meyer.

There are some methods to select the optimal wavelet family, but they have been developed for specific applications and it is not certain that they can be applied to forecasting problems:

- In the cross-correlation method [38], the optimum wavelet maximizes the cross-correlation between the signal of interest and the wavelet;
- In the energy method [39], the aim is to maximize the energy of the signal of interest; and
- In the entropy method [40], the best wavelet minimizes the entropy of the signal of interest.

## 3. Forecasting Models

The wavelet filter was applied to several forecasting methods, namely regression, neural network and fuzzy models.

### 3.1. Persistence Model

In the persistence model, the variable value in  $t + \Delta t$  is equal to the variable value in  $t$ . Due to its simplicity, this model was used as a reference.

$$y_t = y_{t-1}. \quad (8)$$

### 3.2. Regressive Model

This model [41] is based on the multiple regressions that study the relations between a dependent variable and a set of independent variables. Among the independent variables, there are exogenous variables such as temperature, and intrinsic variables like the historical values of these variables. When the model only uses the historical values of these variables, it is called the auto-regressive temporal series model. In this work, the model used the historical values:

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i}. \quad (9)$$

where  $\alpha_i$  represents the auto-regressive parameters and  $p$  is the number of past values.

### 3.3. Neural Network Model

Neural networks [42] are auto-adaptive dynamic systems that are able to find nonlinear relations between several variables. The model used is a multilayer perceptron that gives good results in forecasting problems:

$$y_t = \sum_j w'_{kj} g \left( \sum_i (w_{ji} y_{t-i} - \theta_j) - \theta'_k \right) \quad (10)$$

where  $\theta'_j$  and  $\theta'_k$  are the layer thresholds;  $w_{ji}$  and  $w_{kj}$  are the layer weights;  $i$  and  $j$  are the number of neurons in each layer; and  $g$  is the activation function.

### 3.4. Fuzzy Model

The fuzzy model [43] is based on concepts of fuzzy sets theory, fuzzy rules of type if-then, and approximation reasoning:

$$\begin{cases} R_j: \text{if } \cap_i (y_{t-i} \text{ is } \tilde{A}_i) \text{ then } u_j = \sum_i (p_i y_{t-i}) \\ y_t = \sum_j \left( \frac{\omega_j u_j}{\bar{\omega}_j} \right) \end{cases} \quad (11)$$

where  $\omega_i$  is the normalized firing strengths;  $u_i$  is the functions that depend on the inputs  $y_{t-i}$ ;  $\tilde{A}_i$  is the fuzzy set that represents the input variables; and  $p_i$  is the membership grade of each input  $y_t$  in  $\tilde{A}_i$ .

## 4. Forecasting Approach

Wind time series have high variability due to the intrinsic uncertainty of the wind; this variability negatively influences the forecasting result. In this paper, the adopted approach, illustrated in Figure 6, consists of using an optimized filter based on wavelets to de-noise the data that are used to train the chosen forecasting method. The filter is optimized with a genetic algorithm that selects the best wavelet family and the optimal decomposition level.

The algorithm receives as inputs the wind speed data and the prediction method to be used.

In step 1, a random population of individuals is created. Each individual contains the information of the parameters of the prediction method, the wavelet family, and the level of decomposition.

In step 2, each individual in the population is evaluated. The evaluation has three phases.

The first phase consists of applying the wavelet filter to the input data with the wavelet family and the level indicated by each individual. The data are divided into training and test sets. The original time series is filtered with the wavelet transform and is decomposed into an approximation component and several details of the signal. The approximation component has improved behavior in comparison to the original series in the forecasting process. Therefore, only the approximation component is used in the next phase and the details are discarded. In this phase, there are two important decisions to analyze: the best wavelet family to use and the optimal filter level. These questions have not been answered in the technical literature.

The second phase consists of training the prediction method with the parameters indicated by each individual and the training dataset. A forecasting method is applied only to the approximation component. In this paper, three methods were used to forecast the time series: autoregressive, neural networks, and fuzzy models.

The third phase consists of evaluating the prediction method, already trained, with the test dataset.

In step 3, the best individuals are selected based on the error obtained in the evaluation with the test data.

In step 4, the crossover and mutation operators of the genetic algorithm are applied that give rise to a new population.

Steps 2 through 4 are repeated until the termination criterion is reached, which is the number of generations or iterations of the genetic algorithm.

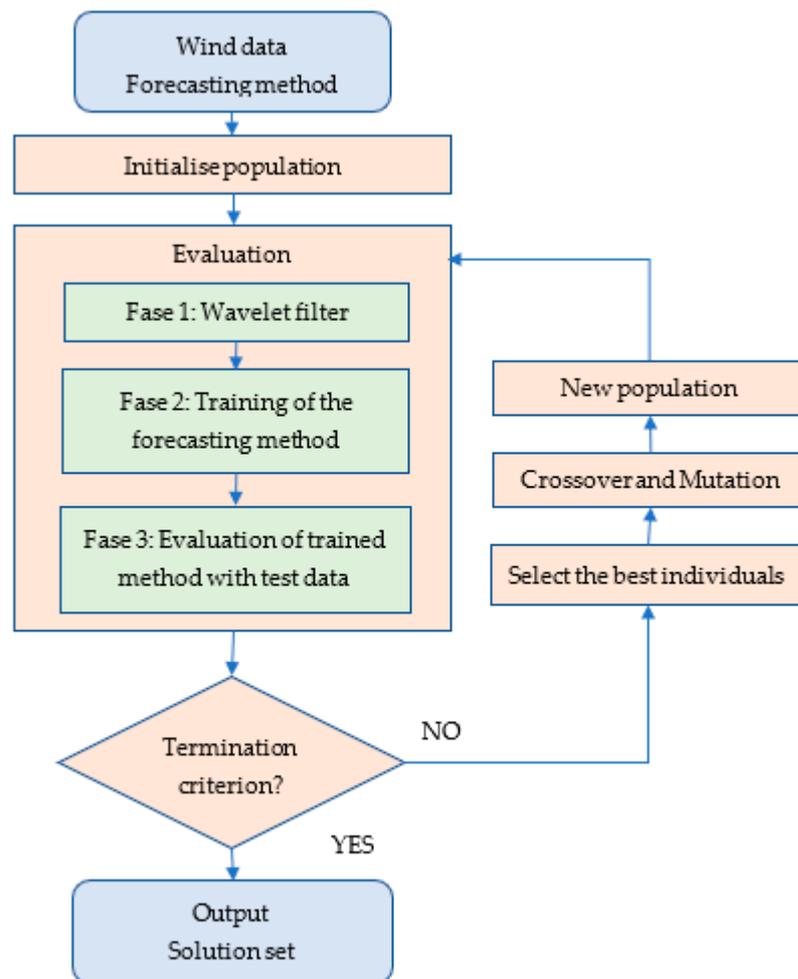


Figure 6. Flow chart of the forecasting approach.

## 5. Forecasting Errors

We compared these forecasting methods with the simpler persistence model used as the reference. The measure of the error of each method was calculated by the root mean square error (RMSE),

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (\text{Xpred}_t - \text{Xreal}_t)^2}{n}} \quad (12)$$

where  $\text{Xpred}_t$  is the predicted value in  $t$ ;  $\text{Xreal}_t$  is the real value at  $t$ ; and  $n$  is the number of samples.

The improvement of each method in comparison to the persistence model was calculated with the following equation:

$$\text{Improvement vs. persistence (\%)} = \frac{\text{RMSE}(\text{persistence}) - \text{RMSE}(\text{method})}{\text{RMSE}(\text{persistence})} \quad (13)$$

## 6. Results of Validation Cases

### 6.1. Data Description

The wavelet filter was applied to several forecasting methods: regression, neural network, and fuzzy models. Five wind speed series were used in this work: two series with a high sampling frequency (1'' and 1', respectively) and three with a sampling frequency of 10'. Table 2 shows the main statistic characteristics of these time series and Figures A1–A5 in the Appendix A represent their temporal variation.

Table 2. Time series description.

Time Series	Mean	Standard Deviation	Maximum	Skewness	Kurtosis	Size	Sampling Frequency
I	5.281	1.862	12.597	0.643	3.139	1 month	1''
II	4.951	2.313	19.726	1.212	5.467	1 month	1'
III	8.258	4.709	34.690	1.115	4.815	2 years	10'
IV	8.308	9.883	30.000	1.124	2.873	2 years	10'
V	6.848	3.548	27.020	0.684	3.406	2 years	10'

Several cases were built from these data to observe the performance of the forecasting models when the sampling step ( $\Delta t$ ) and the forecasting horizon (FH) changed. An identifier (ID) was assigned to each case in Table 3.

Table 3. Case description.

T.S.-I			T.S.-II			T.S.-III			T.S.-IV			T.S.-V		
ID	FH.	$\Delta t$	ID	FH.	$\Delta t$	ID	FH.	$\Delta t$	ID	FH.	$\Delta t$	ID	FH.	$\Delta t$
1	1'	5''	11	2 h	5'	21	10'	10'	26	10'	10'	31	10'	10'
2	5'	5''	12	2 h	10'	22	1 h	1 h	27	1 h	1 h	32	1 h	1 h
3	5'	10''	13	6 h	10'	23	6 h	6 h	28	6 h	6 h	33	6 h	6 h
4	5'	20''	14	6 h	30'	24	12 h	12 h	29	12 h	12 h	34	12 h	12 h
5	10'	10''	15	12 h	10'	25	24 h	24 h	30	24 h	24 h	35	24 h	24 h
6	10'	20''	16	12 h	30'									
7	10'	30''	17	12 h	1 h									
8	20'	30''	18	24 h	30'									
9	20'	1'	19	24 h	1 h									
10	1 h	1'	20	24 h	2 h									

### 6.2. Results

Every time series was divided into two sets of the same length: a training set and a test set. The forecasting models were built with the training set and the results presented here were obtained by applying these models to the test set. Moreover, eight inputs were used in all methods.

First, a detailed analysis of wavelet filtering is presented, aiming to answer whether they are helpful with different prediction methods and whether they depend on factors such as the level of filtering, the prediction horizon, and the sampling frequency. Afterward, we analyze whether it is possible to select the wavelet family by any of the methods described in the literature regardless of the prediction method used. Finally, its application is presented in a specific case.

#### 6.2.1. Influence of Wavelet Filters in Several Forecasting Methods

The best obtained results of applying the wavelet filters on time series are presented in Figures 7 and 8. It is shown that regardless of the model, the forecasting horizon, or time step, the performance was much better with the wavelet filter than without it when the optimal wavelet family and level were chosen.

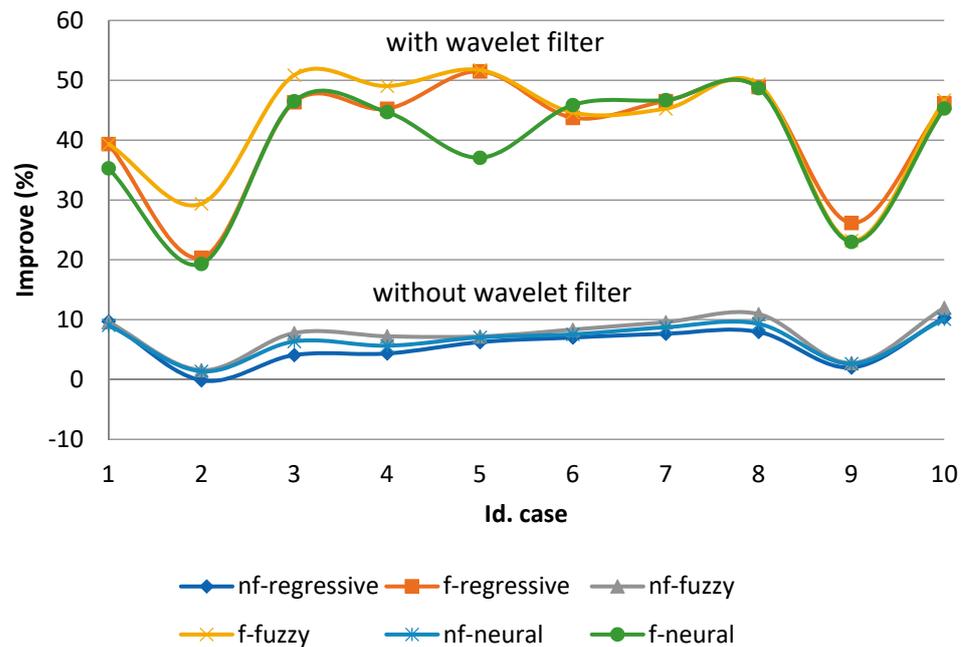


Figure 7. Time series I: Improve versus persistence.

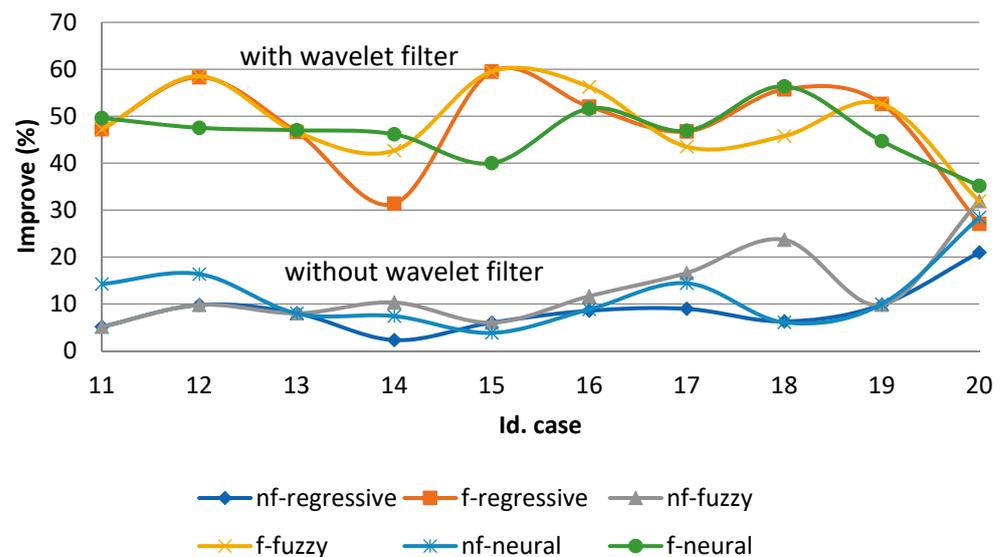


Figure 8. Time series II: Improve versus persistence.

Detailed results are provided in Tables A1–A6 in the Appendix A. It is important to note that the optimal wavelet family and level was different in each case. That is, there was a lot of variability in this point. This fact is contrary to the widespread action among researchers who choose these parameters depending on the application.

### 6.2.2. Influence of Decomposition Level

However, there was a great similarity in the performance of the different wavelet families in each case. Each family reached different levels of improvement, but all families achieved their maximum improvement percentage at a similar level. Figure 9 shows the improvement/level rate for a particular case, with different forecasting methods. Figure 10 shows the improvement/level rate of the same case and method for different wavelet families (see Tables A10–A13 for details).

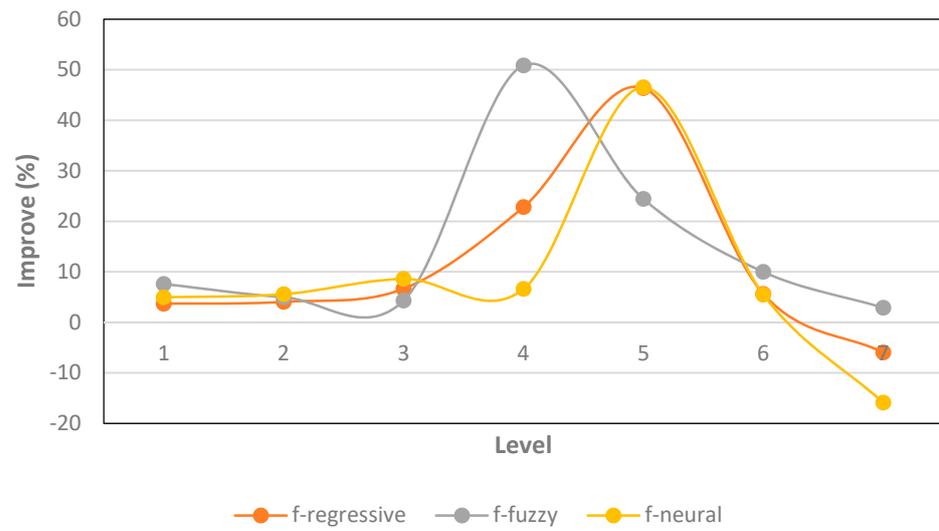


Figure 9. Case 3: Improve versus level for the best family.

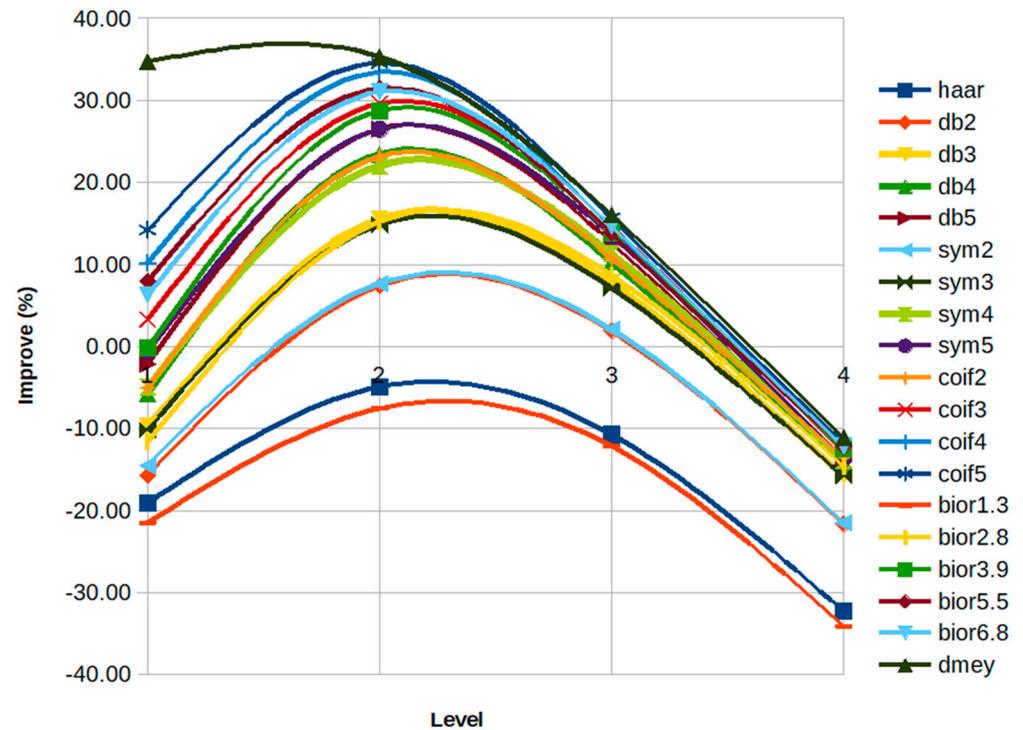


Figure 10. Case 31: Improvement percentage versus level for all families with neural model.

The last results explain why researchers can obtain favorable solutions by applying wavelets, even when they do not select the wavelet family and level accurately.

### 6.2.3. Influence of the Forecasting Horizon

Comparisons made up to now were in percent, because it permitted us to adequately show the difference between whether the wavelet filters were applied or not. However, it is necessary to remember that the error (RMSE) increased with the forecasting horizon, as can be seen in Figures 11–13, although less when the wavelet filters were applied.

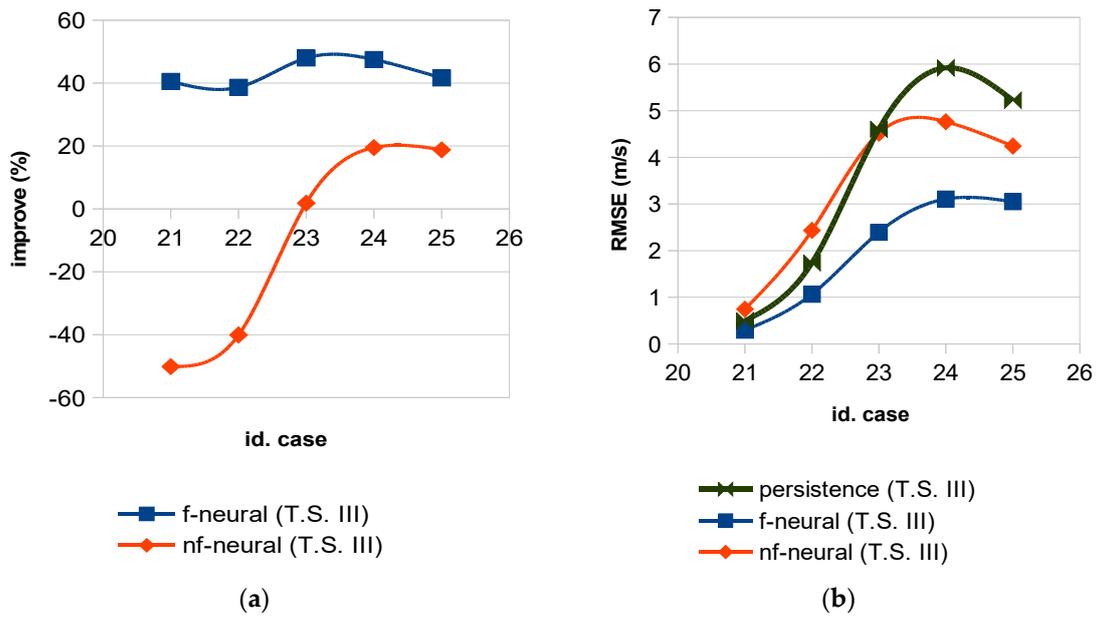


Figure 11. Time series III: (a) Improvement percentage versus persistence; (b) RMSE.

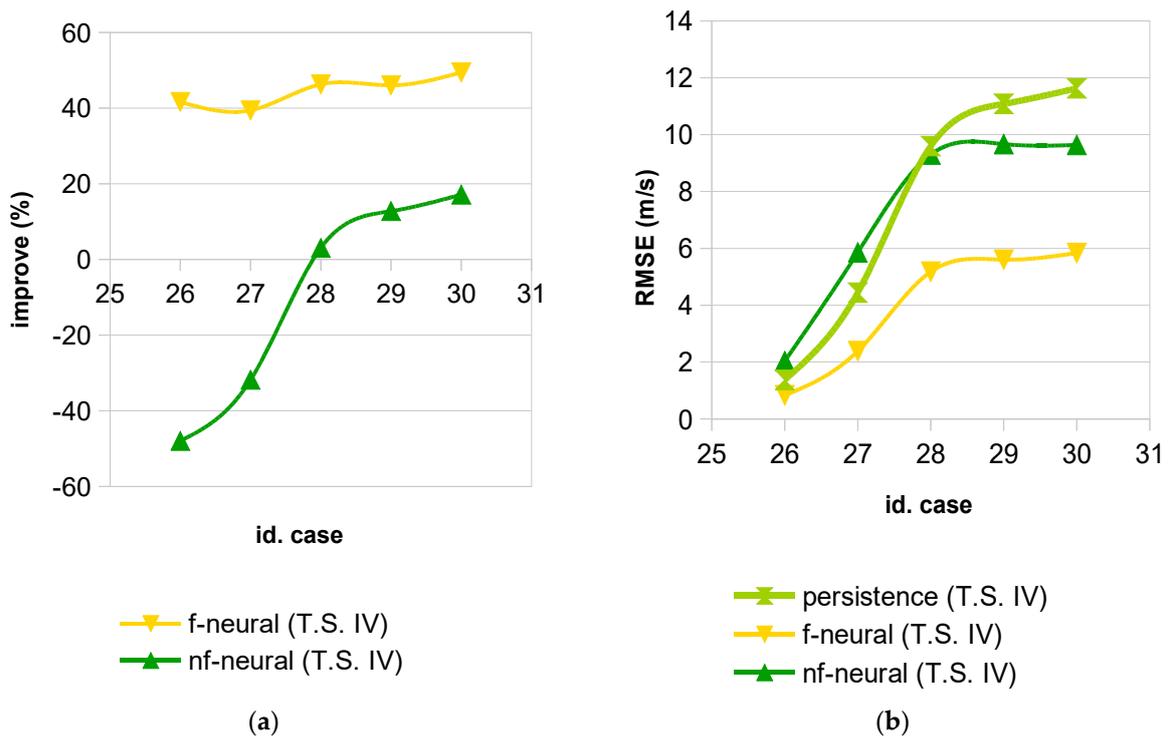


Figure 12. Time series IV: (a) Improvement percentage versus persistence; (b) RMSE.

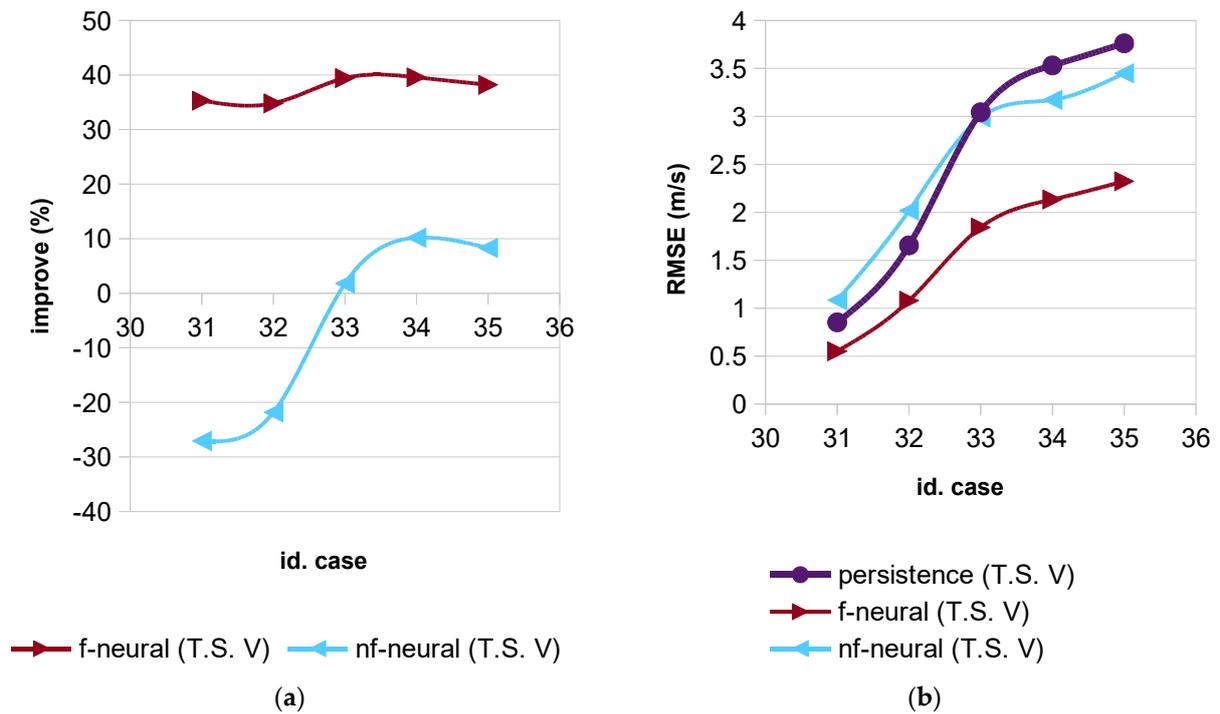


Figure 13. Time series V: (a) Improvement percentage versus persistence; (b) RMSE.

#### 6.2.4. Influence of Different Sampling Frequencies

In Figure 14, it can be seen that with low filtering levels, an important improvement was obtained, but with high filtering levels, information was lost and the improvement decreased or even worsened substantially at high sampling frequencies.

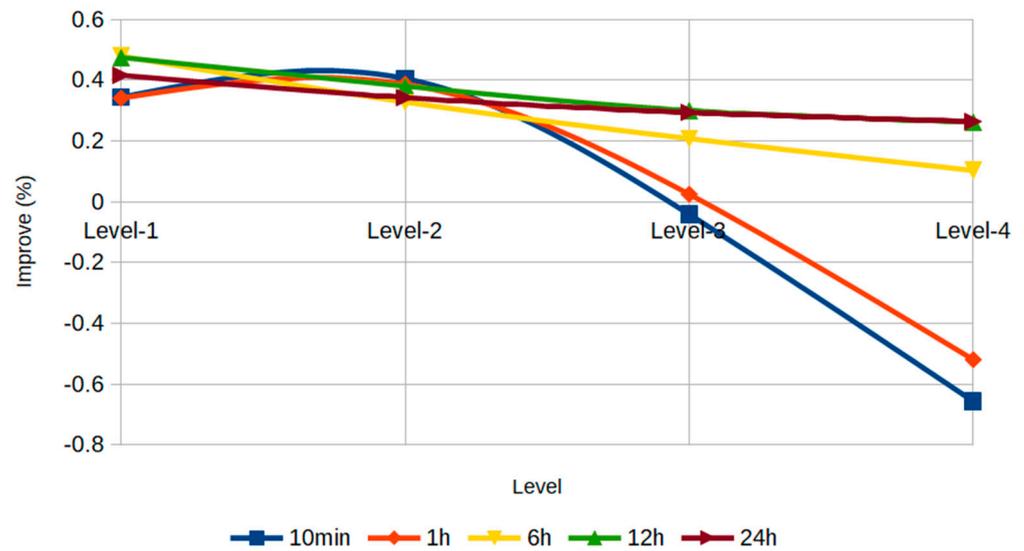


Figure 14. Different sampling frequencies in series III using the dmeiy wavelet filter.

#### 6.2.5. Selection of Optimal Wavelet Family

In Table 4, the wavelet families found by the cross-correlation, energy, and entropy methods are shown in the columns “cross-corr”, “energy”, and “entropy”, respectively. The column “optimum” shows the wavelet family that gave the best results in our tests.

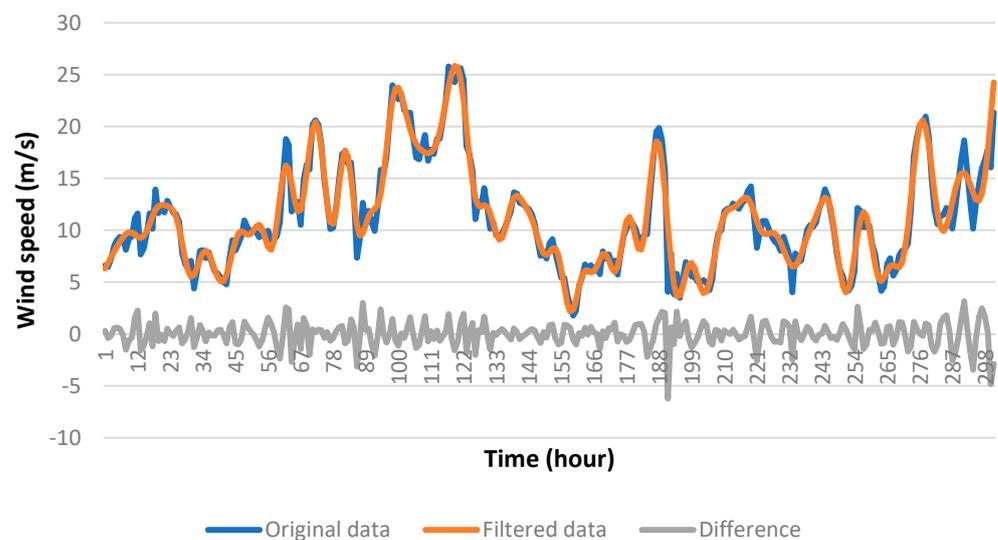
**Table 4.** Selection of wavelet family.

Case	Optimum	Cross-Corr	Energy	Entropy	Case	Optimum	Cross-Corr	Energy	Entropy
1	dmey	Bior5.5	dmey	sym5	21	dmey	Haar-bior1.3	Bior6.8	Bior1.3
2	db2	Bior5.5	dmey	sym5	22	dmey	Haar-bior1.3	Bior6.8	Bior1.3
3	Bior2.8	Bior5.5	dmey	dmey	23	Bior6.8	Bior5.5	dmey	Bior1.3
4	db3	Bior5.5	dmey	db2	24	Bior6.8	Bior5.5	dmey	Bior1.3
5	dmey	Bior5.5	dmey	dmey	25	dmey	Bior5.5	dmey	Bior1.3
6	db5	Bior5.5	dmey	db2	26	dmey	Haar-bior1.3	Bior6.8	Bior1.3
7	Bior5.5	Bior5.5	dmey	Bior1.3	27	dmey	Haar-bior1.3	dmey	Bior1.3
8	db2	Bior5.5	dmey	Bior1.3	28	Bior6.8	Bior5.5	dmey	Bior1.3
9	db2	Bior5.5	dmey	db5	29	coif3	Bior5.5	dmey	Bior1.3
10	Bior3.9	Bior5.5	dmey	db5	30	Bior5.5	Bior5.5	dmey	db2-sym2
11	sym4	Bior5.5	dmey	Bior1.3	31	Bior5.5	Bior5.5	Bior5.5	Bior1.3
12	coif3	Bior5.5	dmey	sym4	32	dmey	Bior5.5	dmey	Bior1.3
13	dmey	Bior5.5	dmey	sym4	33	dmey	Bior5.5	dmey	dmey
14	db4	Bior5.5	dmey	Bior1.3	34	dmey	Bior5.5	dmey	Bior2.8
15	dmey	Bior5.5	dmey	sym4	35	dmey	Bior5.5	dmey	db2-sym2
16	sym5	Bior5.5	dmey	Bior1.3					
17	coif4	Bior5.5	Bior6.8	sym3					
18	dmey	Bior5.5	dmey	Bior1.3					
19	Bior3.9	Bior5.5	Bior6.8	sym3					
20	db4	Bior5.5	Bior6.8	coif5					

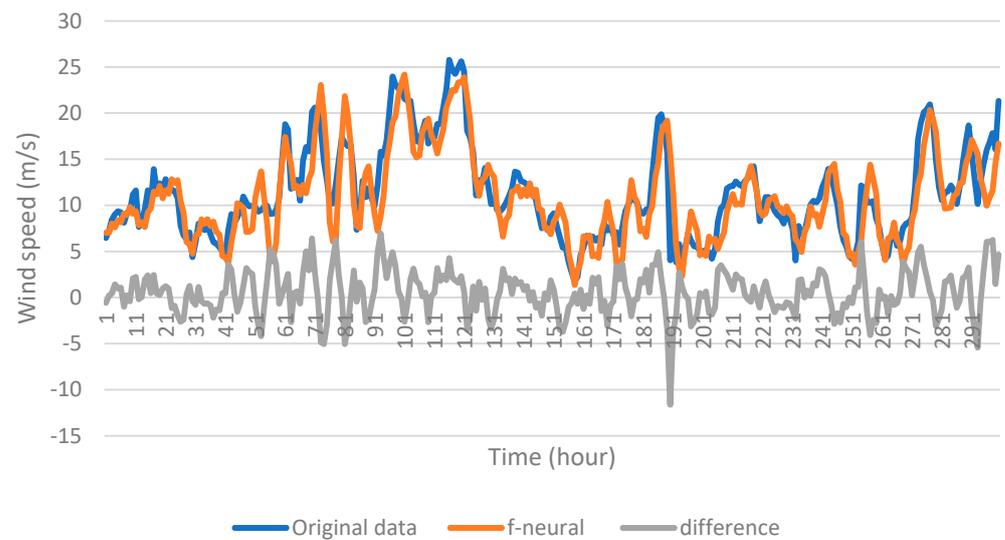
These methods were applied to the time series with poor results. The cross-correlation method obtained the correct result in cases 7, 30, and 31; the energy method in cases 1, 5, 13, 15, 18, 25, 27, 32, 33, 34 and 3; and the entropy method in cases 5 and 33.

### 6.2.6. Applying the Forecasting Approach

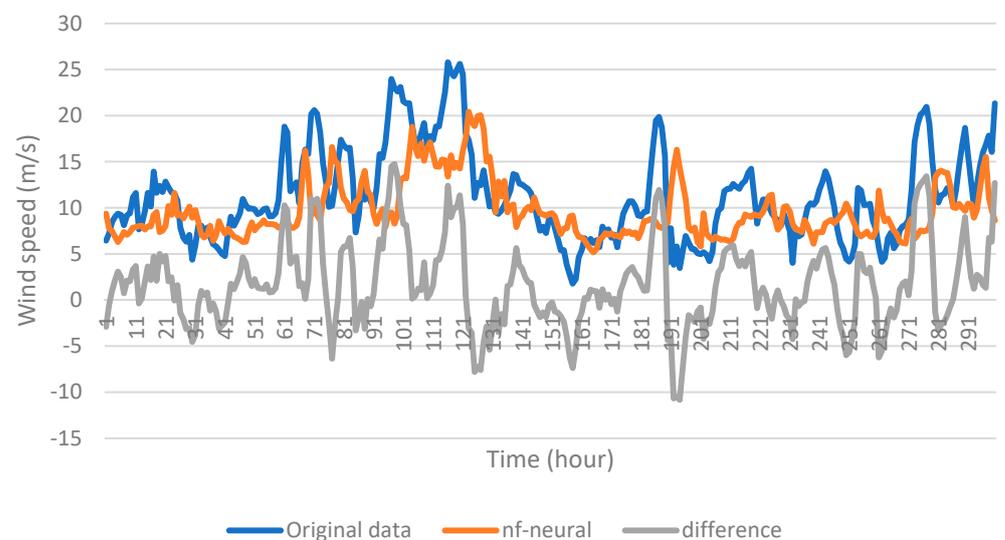
The importance of using filtered data is illustrated in the following example. Figure 15 shows the first 300 data points (to appreciate it in detail) of the original data series of case 22, the data series filtered with the wavelet family “dmey” and a filter level 2, and the difference between the two series. Figure 16 shows the results of the forecast made with the neural network trained with the filtered data training set, and Figure 17 shows the results of the forecast made with the neural network trained with the unfiltered data training set. The results using correctly filtered data were considerably better than those with the unfiltered data.



**Figure 15.** Wavelet filtered process in case 22.



**Figure 16.** Forecasting result for case 22 with the neural network method trained with filtered data.



**Figure 17.** Forecasting result for case 22 with the neural network method trained with unfiltered data.

## 7. Conclusions

In this paper, the forecasting models were applied to the approximation component of the wavelet decomposition and the details components were discarded, as opposed to most authors who use both components in their forecasting models.

A deep analysis of the wavelet filter in results was made, and the conclusions will enable improvements in all forecasting models.

The wavelet filter method was applied to different forecasting models: regression, neural network, and fuzzy models. In all models, this technique (wavelet filter + forecasting model) improved the obtained results compared to the case when only the forecasting model was used. The improvement of these methods versus the persistence method was between 2% and 30%, but with the wavelet filter method, it was between 20% and 50%.

The study was extended to several wavelet families. In all cases, there were improvements, but it was not easy to select the best family. The selection methods did not work for the proposed method.

The filtering level was more important to obtain good results than the wavelet family in most of cases. This optimum level was between 2 and 5 in all wavelet families.

As a final conclusion, it seems necessary to use an optimization algorithm to select the wavelet family and level.

It has become clear that it is not easy to determine the parameters of the data processing methods and that they significantly influence the results obtained. Hence, future research is the joint optimization of the data processing and the forecasting method.

**Supplementary Materials:** Supplementary materials are available online at <https://www.mdpi.com/article/10.3390/en14113181/s1>.

**Author Contributions:** Conceptualization, J.A.D.-N.; Methodology, J.A.D.-N.; Software, S.M.V.-B. and T.B.L.-G.; Validation, S.M.V.-B. and T.B.L.-G.; Writing—original draft preparation, J.A.D.-N.; Writing—review and editing, T.B.L.-G.; Supervision, J.A.D.-N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

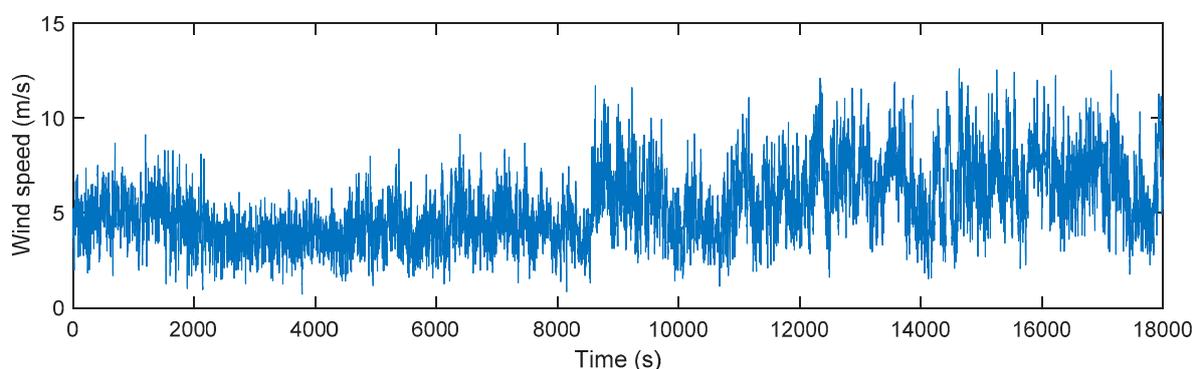
**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in supplementary material.

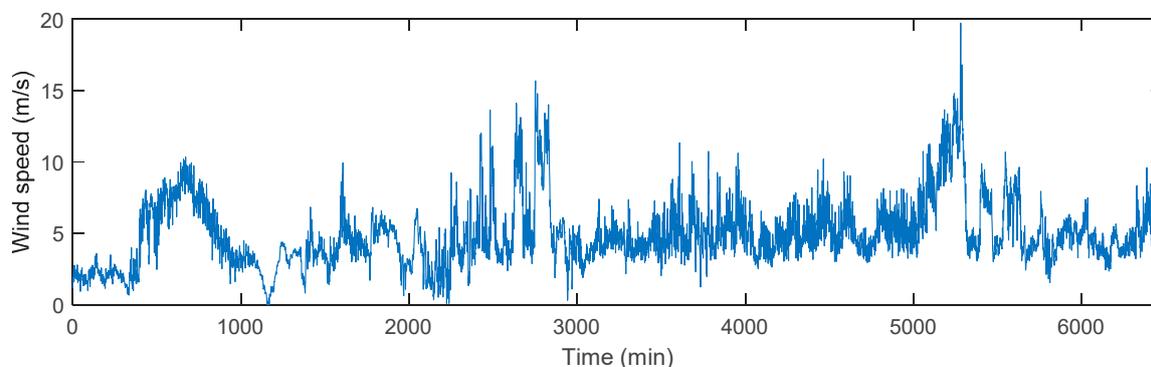
**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

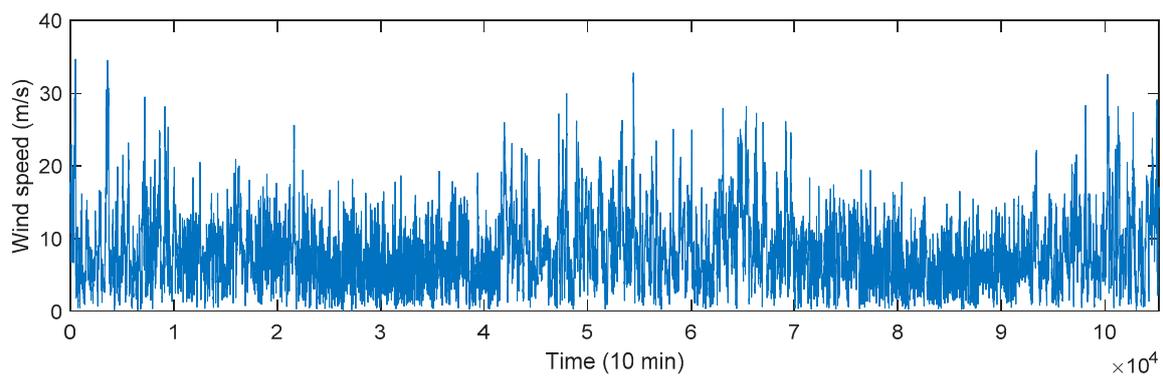
Below are the graphical representations of the measured wind speed data in each time series (Figures A1–A5).



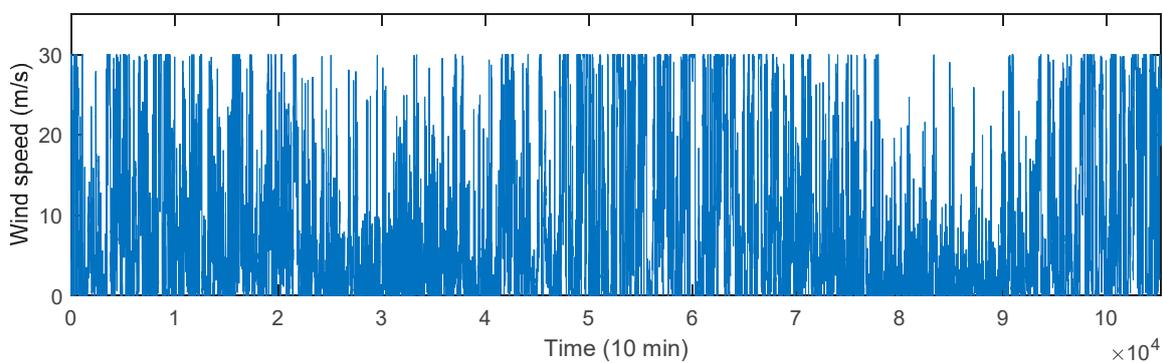
**Figure A1.** Time series 1: Measured wind speed data with 1-s sampling.



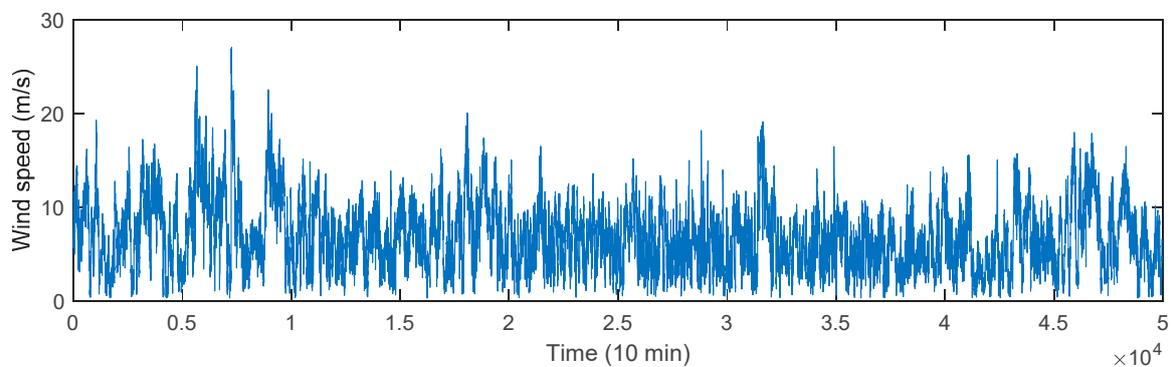
**Figure A2.** Time series 2: Measured wind speed data with 1-min sampling.



**Figure A3.** Time series 3: Measured wind speed data with 10-min sampling.



**Figure A4.** Time series 4: Measured wind speed data with 10-min sampling.



**Figure A5.** Time series 5: Measured wind speed data with 10-min sampling.

In each table, the column FH is the forecasting horizon,  $\Delta t$  is the time step in the look-ahead period, then we have the RMSE of the persistence model, the RMSE of the study model without a wavelet filter, the RMSE of the study model with the wavelet filter, and finally, we have the improvement of the model with and without the wavelet filter versus the persistence model.

**Table A1.** Time series I results of the regressive model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
1	1'	5''	1.255	1.133	0.761	9.71	39.35	mey	3
2	5'	5''	1.017	1.019	0.810	-0.13	20.37	db2	4
3	5'	10''	1.282	1.230	0.688	4.08	46.35	bior2.8	5
4	5'	20''	1.288	1.232	0.705	4.34	45.24	db3	7
5	10'	10''	1.258	1.179	0.610	6.25	51.49	mey	4
6	10'	20''	1.287	1.196	0.724	7.01	43.74	db5	6
7	10'	30''	1.299	1.200	0.694	7.65	46.54	bior5.5	6
8	20'	30''	1.266	1.161	0.647	7.95	48.88	db2	6
9	20'	1'	0.973	0.953	0.718	2.00	26.17	db2	7
10	1 h	1'	1.207	1.083	0.649	10.33	46.18	bior3.9	7

**Table A2.** Time series I results of the fuzzy model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
1	1'	5''	1.255	1.135	0.761	9.58	39.34	mey	3
2	5'	5''	1.017	1.001	0.718	1.57	29.37	db5	4
3	5'	10''	1.282	1.183	0.630	7.76	50.85	mey	4
4	5'	20''	1.288	1.195	0.656	7.22	49.04	bior3.9	6
5	10'	10''	1.258	1.167	0.607	7.20	51.73	mey	4
6	10'	20''	1.287	1.179	0.712	8.33	44.62	db5	6
7	10'	30''	1.299	1.175	0.711	9.58	45.25	db2	6
8	20'	30''	1.266	1.128	0.641	10.93	49.32	db2	6
9	20'	1'	0.973	0.946	0.747	2.76	23.20	db2	7
10	1 h	1'	1.207	1.063	0.643	11.98	46.68	bior3.9	7

**Table A3.** Time series I results of the neuronal model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
1	1'	5''	1.255	1.141	0.812	9.08	35.30	mey	3
2	5'	5''	1.017	1.003	0.821	1.37	19.33	db2	4
3	5'	10''	1.282	1.201	0.686	6.35	46.50	bior2.8	5
4	5'	20''	1.288	1.215	0.712	5.64	44.68	db3	7
5	10'	10''	1.258	1.169	0.791	7.04	37.06	mey	4
6	10'	20''	1.287	1.189	0.696	7.55	45.87	db5	6
7	10'	30''	1.299	1.186	0.692	8.72	46.70	db2	6
8	20'	30''	1.266	1.148	0.649	9.30	48.70	db2	6
9	20'	1'	0.973	0.947	0.749	2.64	22.99	db2	7
10	1 h	1'	1.207	1.085	0.660	10.1	45.31	bior3.9	7

**Table A4.** Time series II results of the regressive model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
11	2 h	5'	0.912	0.865	0.481	5.12	47.19	sym4	3
12	2 h	10'	0.925	0.834	0.385	9.77	58.29	coif3	4
13	6 h	10'	0.730	0.672	0.390	8.00	46.61	mey	4
14	6 h	30'	0.647	0.632	0.444	2.29	31.34	db4	6
15	12 h	10'	6.897	6.482	2.794	6.01	59.49	mey	4
16	12 h	30'	7.346	6.718	3.523	8.55	52.04	sym5	6
17	12 h	1 h	7.968	7.254	4.245	8.95	46.73	coif4	7
18	24 h	30'	15.737	14.750	6.977	6.26	55.69	mey	6
19	24 h	1 h	20.241	18.256	9.592	9.80	52.61	bior3.9	7
20	24 h	2 h	23.967	18.942	17.484	20.96	27.05	db4	7

**Table A5.** Time series II results of the fuzzy model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
11	2 h	5'	0.912	0.865	0.481	5.12	47.19	sym4	3
12	2 h	10'	0.925	0.834	0.383	9.77	58.52	coif3	4
13	6 h	10'	0.730	0.672	0.390	8.00	46.61	mey	4
14	6 h	30'	0.647	0.580	0.370	10.30	42.67	db2	6
15	12 h	10'	6.897	6.482	2.794	6.01	59.49	mey	4
16	12 h	30'	7.346	6.491	3.213	11.63	56.26	sym5	6
17	12 h	1 h	7.968	6.645	4.503	16.59	43.48	coif5	4
18	24 h	30'	15.737	12.013	8.533	23.66	45.77	mey	6
19	24 h	1 h	20.241	18.256	9.592	9.80	52.61	bior3.9	7
20	24 h	2 h	23.967	16.321	16.321	31.90	31.90	haar	1

**Table A6.** Time series II results of the neuronal model.

ID	FH.	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
11	2 h	5'	0.912	0.782	0.459	14.22	49.61	mey	3
12	2 h	10'	0.925	0.774	0.485	16.36	47.51	coif3	4
13	6 h	10'	0.730	0.672	0.387	7.98	46.99	mey	4
14	6 h	30'	0.647	0.599	0.348	7.40	46.15	haar	6
15	12 h	10'	6.897	6.633	4.139	3.82	39.99	mey	4
16	12 h	30'	7.346	6.699	3.555	8.83	51.60	sym5	6
17	12 h	1 h	7.968	6.822	4.232	14.40	46.89	coif5	7
18	24 h	30'	15.737	14.783	6.868	6.06	56.35	mey	6
19	24 h	1 h	20.241	18.224	11.197	9.96	44.68	bior3.9	7
20	24 h	2 h	23.967	17.161	15.545	28.39	35.14	haar	6

**Table A7.** Time series III results of the neuronal model.

ID	FH	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
21	10'	10'	0.497	0.746	0.295	−50.16	40.54	dmey	2
22	1 h	1 h	1.738	2.435	1.066	−40.09	38.70	dmey	2
23	6 h	6 h	4.604	4.518	2.395	1.86	47.97	Bior6.8	1
24	12 h	12 h	5.919	4.764	3.107	19.52	47.50	Bior6.8	1
25	24 h	24 h	5.224	4.241	3.049	18.81	41.63	dmey	1

**Table A8.** Time series IV results of the neuronal model.

ID	FH	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	Level
26	10'	10'	1.372	2.030	0.831	−47.95	39.42	dmey	2
27	1 h	1 h	4.443	5.857	2.386	−31.83	46.29	Bior6.8	1
28	6 h	6 h	9.595	9.301	5.179	3.07	46.03	coif3	1
29	12 h	12 h	11.081	9.665	5.603	12.78	49.44	Bior3.9	1
30	24 h	24 h	11.628	9.634	5.836	17.15	49.81	Bior3.9	1

**Table A9.** Time series V results of the neuronal model.

ID	FH	$\Delta t$	RMSE (m/s)			Improve vs. Persistence (%)		Wavelet	
			Persistence	n-Filter	Filter	n-Filter	Filter	Family	LEVEL
31	10'	10'	0.853	1.084	0.552	−27.07	35.32	dmey	2
32	1 h	1 h	1.656	2.017	1.079	−21.84	34.82	dmey	1
33	6 h	6 h	3.043	2.989	1.843	1.77	39.43	dmey	1
34	12 h	12 h	3.532	3.172	2.134	10.17	39.57	dmey	1
35	24 h	24 h	3.762	3.449	2.325	8.30	38.20	dmey	2

In Tables A10–A13, we can see the effect (in %) of using different wavelet families and different filtered j-level. The best results are marked in bold and underlined. The abscissa axis is the filter level and the ordinate axis is the wavelet family.

**Table A10.** Results of the regressive model (%) in case 3.

Wavelet	Level							
	Family	1	2	3	4	5	6	7
haar		4.08	5.24	9.65	15.45	14.81	6.53	−64.89
daub 2		4.11	4.24	9.01	29.19	17.92	14.64	−75.76
daub 3		3.95	3.97	5.26	18.66	36.62	12.20	−9.48
daub 4		3.73	3.39	6.30	24.22	45.74	8.32	10.82
daub 5		3.63	3.69	5.09	35.18	30.97	24.14	−46.79
symlet 2		4.11	4.24	9.01	29.19	17.92	14.64	−75.76
symlet 3		3.95	3.97	5.26	18.66	36.62	12.20	−9.48
symlet 4		3.77	3.65	5.10	22.10	39.73	5.55	9.77
symlet 5		4.01	3.36	6.67	34.05	26.63	27.39	−64.57
coiflet 2		3.69	3.45	5.39	20.98	42.79	6.81	−6.10
coiflet 3		3.68	3.44	5.36	31.13	43.06	4.89	−0.01
coiflet 4		3.68	3.46	4.11	37.50	38.23	4.41	4.79
coiflet 5		3.68	3.49	4.62	32.45	31.04	4.85	7.47

Table A10. Cont.

Wavelet	Level						
Family	1	2	3	4	5	6	7
bior 1.3	3.60	3.82	6.05	18.12	17.83	12.25	−58.89
bior 2.8	3.67	4.05	6.65	22.81	46.35	5.62	−5.97
bior 3.9	3.71	3.75	3.17	36.75	27.81	22.25	−42.57
bior 5.5	4.04	3.48	3.60	25.45	36.22	11.22	−12.70
bior 6.8	3.66	4.03	5.79	29.49	38.54	8.51	−6.80
meyer	3.64	3.51	3.10	38.44	24.27	10.23	2.87

Table A11. Results of the fuzzy model (%) in case 3.

Wavelet	Level						
Family	1	2	3	4	5	6	7
haar	7.76	8.86	9.02	15.49	14.23	9.12	−63.29
daub 2	7.95	4.77	11.05	30.56	17.60	14.14	−74.53
daub 3	7.73	5.45	6.07	20.51	34.33	10.21	−10.20
daub 4	6.90	4.87	5.54	27.23	43.34	5.67	10.92
daub 5	7.00	4.28	5.81	41.66	30.69	23.73	−46.99
symlet 2	7.96	4.78	11.05	30.56	17.61	14.15	−74.54
symlet 3	7.74	5.45	6.08	20.52	34.33	10.21	−10.21
symlet 4	7.60	5.17	6.16	25.66	38.34	4.52	9.17
symlet 5	7.70	4.81	4.24	39.18	27.83	27.24	−65.07
coiflet 2	7.31	4.94	5.89	22.40	42.09	5.88	−6.99
coiflet 3	7.50	4.91	6.17	34.25	40.49	4.94	0.67
coiflet 4	7.56	4.94	4.97	44.52	35.97	3.96	4.97
coiflet 5	7.59	4.98	5.65	40.20	30.52	4.81	7.52
bior 1.3	7.21	5.41	8.07	20.66	18.05	14.06	−58.78
bior 2.8	7.25	4.79	10.37	25.36	45.00	4.23	−6.61
bior 3.9	7.30	5.27	3.18	44.63	26.84	21.94	−42.45
bior 5.5	8.08	4.56	3.79	29.16	35.24	11.25	−12.30
bior 6.8	7.47	4.90	9.58	31.96	37.50	8.49	−6.41
meyer	7.58	4.97	4.30	50.85	24.42	9.97	2.89

Table A12. Results of the neuronal model (%) in case 3.

Wavelet	Level						
Family	1	2	3	4	5	6	7
haar	6.36	7.57	10.83	−6.02	16.36	2.55	−72.60
daub 2	5.51	−11.23	10.80	5.28	18.06	−15.27	−75.44
daub 3	5.39	4.26	7.07	20.81	36.57	−9.02	−10.14
daub 4	4.93	−17.27	8.84	2.15	45.68	7.39	7.40
daub 5	0.45	5.26	6.60	32.97	30.57	24.66	−72.56
symlet 2	5.51	−11.23	10.80	5.28	18.06	−15.27	−75.44
symlet 3	5.39	4.26	7.07	20.81	36.57	−9.02	−10.14
symlet 4	−0.31	−11.46	7.34	−3.02	40.19	−0.76	9.63
symlet 5	6.06	4.55	8.17	32.26	27.29	27.35	−65.68
coiflet 2	5.25	−3.80	8.25	20.86	3.66	6.88	−9.63
coiflet 3	1.48	−20.58	6.45	1.17	6.55	5.03	1.96
coiflet 4	5.29	−20.46	−0.43	34.05	39.66	4.40	4.62
coiflet 5	−1.08	−22.19	6.00	31.33	33.74	4.78	7.23
bior 1.3	−1.23	−15.93	8.15	−4.01	17.46	−2.32	−58.89
bior 2.8	4.93	5.57	8.59	6.59	46.50	5.48	−15.90
bior 3.9	5.16	4.41	−16.52	4.35	28.43	4.82	−39.62
bior 5.5	5.05	−16.06	5.79	12.66	36.29	11.17	−10.82
bior 6.8	5.06	5.26	8.12	31.26	38.99	0.41	−7.17
meyer	−3.08	−2.65	5.39	34.19	25.64	3.52	2.10

**Table A13.** Results of the neuronal model (%) in cases 22 and 31.

Wavelet Family	Levels in Case 22				Levels in Case 31			
	1	2	3	4	1	2	3	4
haar	−30.73	−25.46	−42.13	−30.73	−19.06	−4.92	−10.73	−32.24
daub 2	−26.23	−238.57	−19.91	−26.23	−15.72	7.38	1.90	−21.62
daub 3	−15.85	7.02	−12.60	−15.85	−9.72	15.53	7.77	−15.68
daub 4	−4.14	13.55	−5.95	−4.14	−5.77	23.39	10.34	−14.27
daub 5	−10.55	21.28	−5.52	−10.55	−2.14	26.52	12.67	−13.74
symlet 2	−25.98	−6.49	−19.21	−25.98	−14.49	7.66	2.12	−21.49
symlet 3	−17.80	6.35	−12.78	−17.80	−10.13	14.85	7.11	−15.78
symlet 4	16.53	15.51	−9.61	16.53	−4.92	22.03	11.38	−13.86
symlet 5	−2.92	24.84	−1.95	−2.92	−0.63	26.50	13.43	−13.62
coiflet 2	30.71	33.45	−5.18	30.71	−5.05	23.13	10.78	−13.54
coiflet 3	−2.42	27.34	−0.32	−2.42	3.30	29.61	14.71	−12.76
coiflet 4	6.49	33.27	−1.69	6.49	10.19	33.42	14.54	−12.35
coiflet 5	12.55	35.62	1.93	12.55	14.20	34.52	15.70	−11.70
bior 1.3	−34.18	−27.20	−41.09	−34.18				
bior 2.8	−23.50	4.12	−9.08	−23.50				
bior 3.9	−7.46	26.02	−3.80	−7.46				
bior 5.5	6.31	29.96	−3.36	6.31				
bior 6.8	4.43	28.93	0.26	4.43				
meyer	34.12	38.70	2.48	34.12				

## References

- Brown, B.G.; Katz, R.W.; Murphy, A.H. Time Series Models to Simulate and Forecast Wind Speed and Wind Power. *J. Clim. Appl. Meteorol.* **1984**, *23*, 1184–1195. [\[CrossRef\]](#)
- Torres, J.; García, A.; de Blas, M.; de Francisco, A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Sol. Energy* **2005**, *79*, 65–77. [\[CrossRef\]](#)
- Bossanyi, E.A. Stochastic Wind Prediction for Wind Turbine System Control. In Proceedings of the 7th BWEA Wind Energy Conference, Oxford, UK, 27–29 March 1985.
- Hill, D.C.; McMillan, D.; Bell, K.R.W.; Infield, D. Application of Auto-Regressive models to U.K. wind speed data for power system impact studies. *IEEE Trans. Sustain. Energy* **2012**, *3*, 134–141. [\[CrossRef\]](#)
- Bossanyi, E.A. Short-Term Wind Prediction using Kalman Filters. *Wind Eng.* **1985**, *9*, 1–8.
- Djurovic, M.; Stankovic, L. Prediction of Wind Characteristics in Short-Term Periods. In *Energy and the Environment into the 1990s, Proceedings of the 1st World Renewable Energy Congress, Reading, UK, 23–28 September 1990*; Pergamon Press: Oxford, UK; New York, NY, USA, 1990.
- Beyer, H.G. Short-Term Prediction of Wind-Speed and Power Output of a Wind Turbine with Neural Networks. In Proceedings of the 5th European Community Wind Energy Conference, ECWEC'94, Thessaloniki, Greece, 10–14 October 1994.
- Alexiadis, M.C.; Dokopoulos, P.S.; Sahsamanoglou, H.S. Wind speed and power forecasting based on spatial correlation models. *IEEE Trans. Energy Convers.* **1999**, *14*, 836–842. [\[CrossRef\]](#)
- Flores, P.; Tapia, A.; Tapia, G. Application of a control algorithm for wind speed prediction and active power generation. *Renew. Energy* **2005**, *30*, 523–536. [\[CrossRef\]](#)
- Sideratos, G.; Hatziaargyriou, N.D. An Advanced Statistical Method for Wind Power Forecasting. *IEEE Trans. Power Syst.* **2007**, *22*, 258–265. [\[CrossRef\]](#)
- Potter, C.W.; Negnevitsky, W. Very short-term wind forecasting for Tasmanian power generation. *IEEE Trans. Power Syst.* **2006**, *21*, 965–972. [\[CrossRef\]](#)
- Sfetsos, A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renew. Energy* **2000**, *21*, 23–35. [\[CrossRef\]](#)
- Liu, H.; Chen, C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. *Appl. Energy* **2019**, *249*, 392–408. [\[CrossRef\]](#)
- Qian, Z.; Pei, Y.; Zareipour, H.; Chen, N. A review and discussion of decomposition based hybrid models for wind energy forecasting applications. *Appl. Energy* **2019**, *235*, 939–953. [\[CrossRef\]](#)
- Huang, N.T.; Xing, E.K.; Cai, G.W.; Yu, Z.Y.; Qi, B.; Lin, L. Short-term wind speed forecasting based on low redundancy feature selection. *Energies* **2018**, *11*, 1638. [\[CrossRef\]](#)
- Sidney Burrus, C.; Gopinath, R.A.; Guo, H. *Introduction to Wavelets and Wavelet Transforms*; Prentice Hall: Upper Saddle River, NJ, USA, 1998.

17. Catalão, J.P.S.; Pousinho, H.M.I.; Mendes, V.M.F. Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal. *IEEE Trans. Sustain. Energy* **2010**, *2*, 50–59. [[CrossRef](#)]
18. Liu, H.; Tian, H.-Q.; Pan, D.-F.; Li, Y.-F. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. *Appl. Energy* **2013**, *107*, 191–208. [[CrossRef](#)]
19. Cao, L.; Li, R. Short-term wind speed forecasting model for wind farm based on wavelet decomposition. In Proceedings of the Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, DRPT 2008, Nanjing, China, 6–9 April 2008; pp. 2525–2529.
20. Yao, C.; Yu, Y. A Hybrid Model to Forecast Wind Speed Based on Wavelet and Neural Network. In Proceedings of the International Conference on Control, Automation and Systems Engineering (CASE), Singapore, 30–31 July 2011; 2011; pp. 1–4.
21. Khan, A.A.; Shahidehpour, M. One day ahead wind speed forecasting using wavelets. In Proceedings of the Power Systems Conference and Exposition PSCE '09, Seattle, WA, USA, 15–18 March 2009; pp. 1–5.
22. Mishra, S.; Sharma, A.; Panda, G. Wind power forecasting model using complex wavelet theory. In Proceedings of the International Conference on Energy, Automation, and Signal (ICEAS), Bhubaneswar, India, 28–30 December 2011; pp. 1–4.
23. Tong, J.-L.; Zhao, Z.-B.; Zhang, W.-Y. A New Strategy for Wind Speed Forecasting Based on Autoregression and Wavelet Transform. In Proceedings of the 2nd International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE), Nanjing, China, 1–3 June 2012; pp. 1–4.
24. Li, Y.F.; Wu, H.P.; Liu, H. Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM sub-ordinate computing and IEWT reconstruction. *Energy Convers. Manag.* **2018**, *167*, 203–219. [[CrossRef](#)]
25. Liu, H.; Duan, Z.; Han, F.-Z.; Li, Y.-F. Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm. *Energy Convers. Manag.* **2018**, *156*, 525–541. [[CrossRef](#)]
26. Liu, H.; Mi, X.; Li, Y. Comparison of two new intelligent wind speed forecasting approaches based on Wavelet Packet Decomposition, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Artificial Neural Networks. *Energy Convers. Manag.* **2018**, *155*, 188–200. [[CrossRef](#)]
27. Chitsaz, H.; Amjady, N.; Zareipour, H. Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. *Energy Convers. Manag.* **2015**, *89*, 588–598. [[CrossRef](#)]
28. Esfetang, N.N.; Kazemzadeh, R. A novel hybrid technique for prediction of electric power generation in wind farms based on WIPSO, neural network and wavelet transform. *Energy* **2018**, *149*, 662–674. [[CrossRef](#)]
29. Zhang, Y.; Zhang, C.; Gao, S.; Wang, P.; Xie, F.; Cheng, P.; Lei, S. Wind Speed Prediction Using Wavelet Decomposition Based on Lorenz Disturbance Model. *IETE J. Res.* **2018**, *66*, 635–642. [[CrossRef](#)]
30. Hu, J.; Wang, J.; Ma, K. A hybrid technique for short-term wind speed prediction. *Energy* **2015**, *81*, 563–574. [[CrossRef](#)]
31. Wang, J.-Z.; Wang, Y.; Jiang, P. The study and application of a novel hybrid forecasting model—A case study of wind speed forecasting in China. *Appl. Energy* **2015**, *143*, 472–488. [[CrossRef](#)]
32. Zhang, W.; Wang, J.; Wang, J.; Zhao, Z.; Tian, M. Short-term wind speed forecasting based on a hybrid model. *Appl. Soft Comput.* **2013**, *13*, 3225–3233. [[CrossRef](#)]
33. Liu, D.; Niu, D.; Wang, H.; Fan, L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renew. Energy* **2014**, *62*, 592–597. [[CrossRef](#)]
34. Zhang, Y.; Li, R.; Zhang, J. Optimization scheme of wind energy prediction based on artificial intelligence. *Environ. Sci. Pollut. Res.* **2021**, 1–16. [[CrossRef](#)]
35. Yu, C.; Li, Y.; Zhang, M. An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network. *Energy Convers. Manag.* **2017**, *148*, 895–904. [[CrossRef](#)]
36. Mi, X.-W.; Liu, H.; Li, Y.-F. Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. *Energy Convers. Manag.* **2017**, *151*, 709–722. [[CrossRef](#)]
37. Cheng, L.; Zang, H.; Ding, T.; Sun, R.; Wang, M.; Wei, Z.; Sun, G. Ensemble Recurrent Neural Network Based Probabilistic Wind Speed Forecasting Approach. *Energies* **2018**, *11*, 1958. [[CrossRef](#)]
38. Ma, X.; Zhou, C.; Kemp, I.J. Automated wavelet selection and thresholding for PD detection. *IEEE Electr. Insul. Mag.* **2002**, *18*, 37–45. [[CrossRef](#)]
39. Donoho, D. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* **1995**, *41*, 613–627. [[CrossRef](#)]
40. Coifman, R.; Wickerhauser, M. Entropy-based algorithms for best basis selection. *IEEE Trans. Inf. Theory* **1992**, *38*, 713–718. [[CrossRef](#)]
41. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley: Hoboken, NJ, USA, 1998.
42. Tsoukalas, L.H.; Uhrigh, R.E. *Fuzzy and Neural Approaches in Engineering*; Wiley-Blackwell: Hoboken, NJ, USA, 1997.
43. Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]