

Multi-Kalman filter to wind power forecasting

Paulo Salgado
CITAB/ECT Departamento de
Engenharias
Universidade de Trás-os-Montes e Alto
Douro
Vila Real, Portugal
psal@utad.pt

Getúlio Igrejas
Escola Superior de Tecnologia e Gestão
Instituto Politécnico de Bragança,
Bragança
Bragança, Portugal
igrejas@ipb.pt

Paulo Afonso
Instituto de Telecomunicações
ESTGA, Universidade de Aveiro
Agueda, Portugal
pafnaa@ipb.pt

Abstract— Wind power forecasting methods are important for the safety of wind renewable energy utilization. However, because wind power is weather dependent and, thus, can be variable and intermittent over different time-scales, it's difficult to build accurate and robust predictive models. In this context, an accurate model is a relevant contribution for a reliable large-scale wind power integration. This paper explores a new approach using a multi-model Kalman filter to provide an estimate of the average hourly wind speed, in a 24 hours horizon. The K-means algorithm is used to obtain the characteristics curves of wind speed each time the sub-model is used for forecasting. The accuracy of the proposed forecasting model is compared with other statistical methods, namely some that are usually considered suitable, robust and accurate.

Keywords—wind speed forecasting, Kalman filter.

I. INTRODUCTION

The use of wind power has been growing quickly around the globe and, particularly, in the last decade in Portugal. The strong investment in infrastructures was reflected in the fact that in the last year renewable energies accounted for 57% of the electricity supply in Portugal (REN - Rede Energética Nacional), ranging from hydroelectric power (28%) to wind (22%), biomass (5%) and photovoltaic (1.4%). Non-renewable production, however, accounted for 43% of consumption, divided between coal and natural gas [1].

The problem with wind power is that it is climate dependent, therefore, irregular and with strong variations over different time-scales. Consequently, it is difficult to estimate with reasonable certainty the electric energy that can be produced [2]. Moreover, with the increased prevalence of wind power in the energy network it is augmented the unpredictability factor of energy production. This point can only be overcome if there is the ability to predict the power that will be injected into the electrical network [3].

Additional to the technological upgrading of wind power generation systems and reduction of production costs, the improvement of forecasting methods of wind power are fundamental for a more reliable integration into a large-scale energy market, but also for small-scale systems production of self-consumption. This search for prediction accuracy stimulates researchers to develop accurate short-term models for wind power estimation. In last decade research allowed obtaining good wind speed forecasting systems. These can be classified as physical models [4] and statistical methods [5]. The first group is suitable for long-term prediction while the second presents a good performance in short-term prediction. Physical models consider geographic and local terrains characteristics and weather variables to estimate the future wind speed and generated power. The

statistical methods use statistical tools to predict the wind speed and wind power. Recently, hybrid methods combining mathematical models for weather forecasting and statistical techniques have been proposed [6]. In this last group are included the autoregressive model (AR), the moving average model (MA), the autoregressive moving average model (ARMA), the autoregressive integrated moving average model (ARIMA) [7], the Box-Jenkins methodology and the Kalman filter [8].

Alternatively to the classical time series analysis, that models the statistical relationship among the data, mainly soft computing (or machine learning) approaches using fuzzy systems [9], artificial neural networks (ANN) [10] or its recurrent version (RNN) [11], but also other models, like gray predictors or support vector machines (SVM) [12] and Bayesian methods [13] have been used with success.

The aim of this work is to analyse and compare different approaches in what concerns to performance and model estimation time to predict wind speed, along a time series difficult to predict. The forecasting models that will be treated henceforward, are well described in the literature. We focused on the AR, ARMA and the Linear Regressor. Results obtained using these techniques were compared with the proposed Multi-Kalman-Filters (MKF), which structure is based on the K-means classification algorithm. In previous work [14] the hybrid model with NN-Fuzzy Clustering has already been tested in the same way. All these approaches are supervised and, therefore, capable to model the process from input/output data samples. Features used in the train are the mean hourly wind speed values, in $m.s^{-1}$, measured on the location of Alvadia, district of Vila Real, Portugal. It covers a period from 1st of January 2007 to 17th of March 2009, representing 19368 h in total. This and other data is available from SNIRH (National System of Water Resources) [15].

In this work we propose a new short-term forecasting multi-model in a parallel structure that is a result of an identification process based on the analysis of historical time to find patterns in wind speed time series. This task was performed using the K-means clustering algorithm that recognizes from wind speed series the half-daily pattern sequences. The main inputs for the learning process are wind data collected on selected meteorological stations. The wind speed values from the previous hours are grouped into clusters according to their similarity. After the recognition by the clustering algorithm a Time Variant State Space Model (TVSSM) is associated to each cluster and its parameters are tuned using a Kalman filter method.

The assessment of this new forecast approach, together with the suggested identification method, has shown excellent results in short-term prediction of wind velocity. The proposed MKF approach to forecast wind speed can be

extended to the wind power generation by using power curves, typically an equation that relates the wind turbine's power output with the cubic power value of wind speed, $P \propto v^3$. By safety condition and limitation of turbines, the range of wind speed allowed is limited. Wind speed has also long periods with low intensity values that do not contribute to the production of electricity. In this situation, it is a sign purely random and of resilience. This problem has less impact on the wind energy signal. So, due to this fact, wind power forecast models are almost times more accurate. In addition, their learning processes are also more robust, convergent and faster.

The organization of this paper is the following: Section 2 presents a modified version of the K -means algorithm to be used in recognition of wind speed curves. This version is a fetch to adapt it to the time series data. Section 3 presents various methods used in forecasting wind speed. In Section 4 we propose the MKF algorithm. Its performance is described in the next section with the presentation of the computational results. Finally, the last section presents the main conclusions of this study and about MKF algorithm.

II. RECOGNITION OF WIND SPEEDS CURVES

The K -means treats each observation in data as an object having a location in space. It performs a data partition by grouping them by a similarity criterion. In built groups, their objects are close to each other and, as far as possible, from objects in other clusters.

The K -means clustering algorithm [16] is one of the most widely used clustering methods, being based on the concept of hard-partition. It partitions a data set $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathbf{R}^n$, representing a finite set of feature vectors, with $\mathbf{x}_i \in \mathbf{R}^m$ where m is the dimension of the features space and n is the number of objects (measurements), into K clusters. Each cluster is represented by a prototype $\mathbf{c}_i \in \mathbf{R}^n$, $1 \leq i \leq K$. The membership degree $u_{i,k} \in \{0,1\}$ indicates the degree of belongingness of data object \mathbf{x}_k to prototype i , with $u_{i,k} = 1$ if it belongs to cluster k or $u_{i,k} = 0$ if it does not. All membership degrees form a membership matrix $U \in \mathbf{R}^{K \times n}$. Each sample point \mathbf{x}_k satisfies the constraint $\sum_{i=1}^K u_{i,k} = 1$. Clustering is formulated as a constrained minimization problem, imposed in the membership functions:

$$J_m = \sum_{k=1}^n \sum_{i=1}^K u_{i,k}^m \left(d_{ik}^2 + \eta(1 - u_{i,k-1}) \right) \quad (1)$$

where $d_{ik} = d(\mathbf{x}_k, \mathbf{c}_i)$ is the distance from a feature point \mathbf{x}_k to the center of the i^{th} cluster, \mathbf{c}_i . η is a parameter that penalizes the change of clusters between elements of consecutive instants, but without effect when the new sample remains in the same cluster.

In this paper, the feature vector $\mathbf{x}_k = [v_{k-1} \dots v_{k-12}]^T - v_k$ represents the past 12 hours wind speed values, taken from

the k^{th} instant. The half-daily wind speed \mathbf{x}_k is related by the membership vector $\mathbf{u}_k = [u_{1k} \dots u_{ik} \dots u_{ck}]^T$.

The main aim of the proposed forecasting model with a sub-model for each cluster is to find the relations between the pair wise vector \mathbf{u}_k and the matrix of speed wind prototypes \mathbf{c}_k . In the forecasting process the cluster is used to classify the wind speed profile to select the sub-model used for forecasting. Tuning is done using a Kalman filter.

The time series data is partitioned by K -means algorithms, for $K=20$. The centers of clusters i are shown in Figure 1. Figure 2 evidences the probability of exchange between clusters, from one instant to the next.

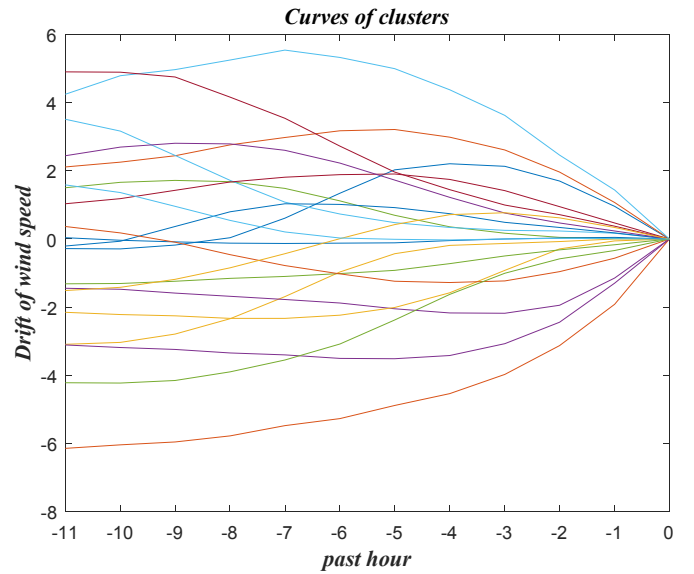


Figure 1. Prototypes (centers) of clusters.

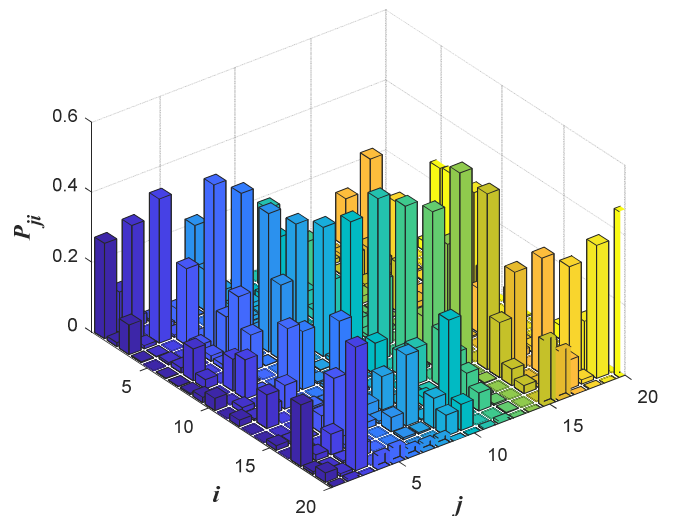


Figure 2. Probability for exchange clusters in each time instant

III. THE FORECASTING MODELS

In this section, we introduce the concept of wind speed forecasting model. Moreover, the mathematical models related with the different approaches are also presented.

All forecasting models were implemented according to the general expression presented in equation 2, meaning that their output value is the current value of the time series (v_k) and depends on the past measurements.

$$v_{k+h} = F(v_{k-1}, \dots, v_{k-p}) \quad (2)$$

In equation 2, F represents any one of the three models used, h is the headed time and p the order.

In this work, the aim of the forecasting model is to predict a variable v_k (wind speed, at the instant k) from independent past values v_{k-i} , $i = 1, \dots, p$ which are assumed as known. Results obtained from the application of the linear and non-linear models were confronted with the those from persistent method. The accuracy results of the forecasting models will be measured by the sum of errors expressed in absolute L_1 norm or Root Mean Square (RMS or L_2). Both error measures are easier to interpret, and they are expressed in the same units as the forecasted variable for a time horizon of n samples:

$$L_1 = \frac{1}{n} \sum_{k=1}^n |\hat{v}_k - v_k| \quad (3)$$

$$L_2 = \sqrt{\frac{1}{n} \sum_{k=1}^n (\hat{v}_k - v_k)^2} \quad (4)$$

Skill score can use L_1 or L_2 , including a normalized version, by presenting them as a percentage improvement (equation 5).

$$\text{Imp}_{\xi}^{\text{ref}} = \frac{\xi^{\text{ref}} - \xi}{\xi^{\text{ref}}} \times 100 \quad (\%) \quad (5)$$

where Imp is equal to the improvement with respect to ξ^{ref} (value for the reference approach, here in this work the persistent model response) and ξ the value for the advanced approach. Values below 100 correspond to accurate forecasts.

A. Persistent model

In the Persistent model, it is assumed that the forecast value of the time series is the last measured one.

$$v_t = v_{t-1} \quad (6)$$

Due to its simplicity, it is used as benchmarking for the various approaches.

B. AR model

The AR model structure is given by the following equation:

$$v_k + \sum_{i=1}^p a_i v_{k-i} = e_k \quad (7)$$

AR model parameters (a_i , $i = 1, \dots, p$) are estimated using variants of the least-squares method.

C. Conventional ARMA model

The ARMA model, denoted by ARMA(p, q) can be represented as:

$$v_k = c + \sum_{i=1}^p a_i v_{k-i} + \sum_{j=1}^q b_j e_{k-j} + e_k \quad (8)$$

where c is a constant of the model, a_i is the i^{th} autoregressive coefficient, b_j is the j^{th} moving average coefficient, e_k is the error at time period k and represents the observed or predicted value of wind speed at time instant k .

D. Kalman filter

The Kalman filter (KF) consists in an algorithm that uses a model with parameter uncertainty and a series of measurements containing random noise and other inaccuracies, observed over time, to produce estimates.

It works in a two-step process, where the first is for prediction. Here the KF performs estimates of the (*a priori*) current state variables, along with their uncertainties, usually assigned to (*a priori*) estimate covariance matrix. Once a new observation variable value is obtained, the state variable and the covariance matrix are both updated. This process starts with the calculation of the Kalman gain, (*Step 1*), which in conjunction with the error, is used to recalculate (*a posteriori*) an estimate of the output and adjusting the (*a priori*) state vector estimate (*Step 2*). The error covariance matrix is then calculated for the posteriori state vector and its projection obtained (*Step 3* and *4*). For the next iteration an *a priori* estimate of the state vector is made using the model of the system (*Step 5*). The mentioned steps are repeated iteratively. All equations used in this algorithm can be found in [17][18].

The KF model, at time k , assumes the true state, s_k , and is evolved from the previous state ($k-1$) according to

$$s_{k|k-1} = A_k s_{k-1|k-1} + w_k \quad (9)$$

where A_k is the state transition model which is applied to the previous state $s_{k-1|k-1}$; w_k is the process noise, statistically characterized as a multivariate normal distribution with zero means and covariance Q , i.e. $w_k \sim N(0, Q_k)$.

In discrete domain time k , an observation (or measurement) v_k of the true state s_k is made according to:

$$v_k = C_k s_{k|k} + v_k \quad (10)$$

Equation (10) represents the TVSSM model. In this model, v_k is the observation noise (zero mean Gaussian white noise with covariance R) and C_k is the observation model that maps the true state space into the observed space.

Because of its recursive nature, the algorithm can run in real time (here referred as online), but also based on a previous learning process with historical data (offline model). In the first case, it uses only the present input measurements and the previously calculated state. The model parameters are continuously tuned and no additional past information is required. In the last case, parameters of the model are static or restricted to an acceptable range of values, which resulted from an offline learning process. In this work, both processes were used.

In this work, s are the parameters of the forecasting model while $C_k = [v_{k-1} \dots v_{k-12}]^T$ is the vector with the past 12th hours of observed wind speed values.

E. MKF algorithm

The MKF forecasting model (MKFM) is a multi-model structure learned by the K -means algorithm presented in section II, with a number of sub-models equals to the number of clusters. So, MKFM is a non-linear switched model selected by a classifier based on clustering. Each sub-model is a Linear Time-Varying System based on equations (9)-(10). Each parameter of the model is tuned by the KF. For each wind profile, $x_{k-1} = [v_{k-1} \dots v_{k-12}]$, at time $k-1$, a sub-model is selected based on the cluster classifier. This is used to predict the wind speed v_{k+1} , at time $k+1$.

The learning method is divided into three phases:

Step 1 - Structural Learning: first, values of the wind from the last 12 hours are clustered based on the degree of similarity between data. So, the K -means algorithm detects patterns in existing data in order to form the most homogeneous groups. The original time series is partitioned into a number of different clusters, and characterized by the partition matrix U and the cluster centres, c , corresponding to a wind speed typical curve (curves prototypes). Each cluster is associated to a sub-model.

Step 2 - Select sub-model and forecasting: For each wind profile x_{k-1} select a sub-model according to the cluster classifier result. The classification result is expressed by the matrix U . If $u_{i,k} = 1$, the i^{th} cluster is selected. Next, use TVSSM to build the prediction models in the forecasting horizon based on equation (10).

Step 3 - Tuning MKFM parameters: Use Kalman filter for adjusting the weights of parameters of i^{th} model, s , which are state variables of equation (9), where A is asymptotically stable in sense of Lyapunov stability.

IV. RESULTS

In this section, the results obtained from testing the presented five algorithms are discussed. They are the Persistent, the ARIMA, the AR, the RLS and the proposed MKF in run online and offline versions. All models forecast the wind speed mean one hour ahead.

The main results are now presented and summarized in Table I. They are analysed using the RMS error and compared with the benchmark persistent method according to the percentage of improvement defined on equation 5. The last column of the table contains additional information about architecture, i.e. the number of sub-models and number of parameters for each model.

The results of the ARIMA model are in concordance with the work of Torres and co-workers. They compared the ARIMA based models with the persistence-based models [19]. It was shown that for short-term prediction horizons the persistence model presents best results than the ARIMA based models. Only for long-term horizons the ARIMA models present better performance. The order of the ARIMA model was chosen aiming to improve the accuracy of the forecasting response.

The performance of the AR model was shown to be significantly better than the ARIMA and 6.8% better than the Persistent model. Due to the similarity between the AR and RLS forecasting models their results are practically the same. The proposed MKF model was shown to be the best forecasting model. It was 65% better than the persistent model and 60% better than the AR/RLS models.

In Figure 3 is plotted the response of all forecasting models for an interval period of 100 hours.

Figure 4 shows the number of sub-model of MKF (number of cluster) selected for forecasting and tuned by the Kalman filter.

The analysis suggests that the non-linear MKF models present better performance than the linear AR and ARMA, suggesting greater ability to understand the characteristics of this particular time series. Their lower estimate errors, both for the train and test periods, evidence that the proposed model, known the values of wind pattern curves and the cluster classifier, has a good performance.

TABLE I.
FORECASTING ERRORS FOR THE TEST PERIOD

Model	L2 (ms ⁻¹)	Imp (%)	L1 (ms ⁻¹)	Imp (%)	Architecture (parameters)
Persistent	0.95241	0	0.61451	0	-
ARIMA	1.12123	-17.7	0.73950	-20.3	ARMA(5,0,5)
AR	0.90549	4.9	0.57282	6.8	AR(12)
RLS	0.90548	4.9	0.57256	6.8	RLS(12)
MKFM(1)	0.27398	71.2	0.19321	68.6	MFF(20,12)
MKFM(2)	0.32903	65.5	0.22768	63.0	MKF(20,12)

V. CONCLUSIONS

In this paper a comparison of various time series forecasting approaches, based on the mean hourly wind speed, was presented. Results obtained indicated that the proposed MKF based models outperformed the respective linear and all non-linear approach when compared with the L1 and L2 errors. Moreover, results suggest that the proposed method is a robust and accurate forecast model with RMS errors in the range below of 0.2 ms⁻¹, meaning that it can be used for simulation and to predict wind speed with 1 day anticipation and showing no degradation in the considered lead time.

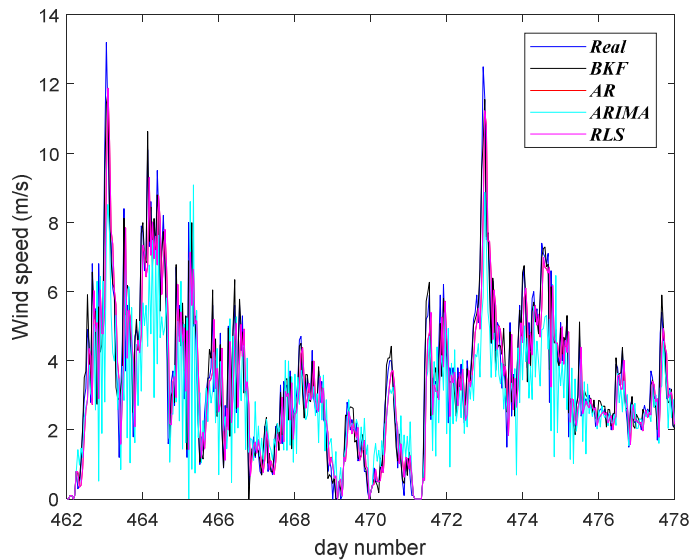


Figure 3. Wind speed forecasting of the 4 models and real wind speed wind (blue line) (test phase).

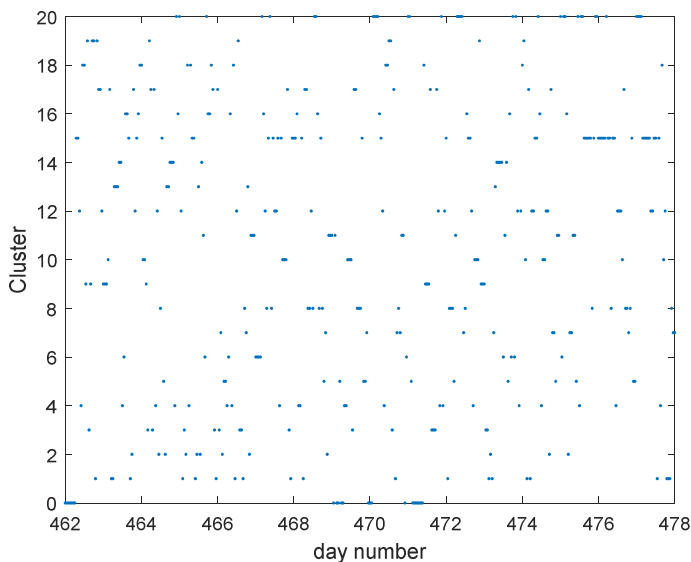


Figure 4. Number of Cluster/KF sub-model used for forecasting of wind speed.

REFERENCES

- [1] https://www.ren.pt/pt-PT/media/comunicados/detalhe/ren_regista_novos_maximos_na_producao_de_energia_eolica/
- [2] M. Negnevitsky, and C. W. Potter, Very Short-Term Wind Forecasting for Tasmanian Power Generation, *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 965 – 972, May 2006.
- [3] M.C. Alexiadis, P.S. Dokopoulos, H. Sahsamanoglou and I.M. Manousaridis, Short-term forecasting of wind speed and related electrical power, *Solar Energy*, vol. 63(1), pp. 61–68, 1998
- [4] Landberg L, Myllerup L, Rathmann O, Lundtang Petersen E, Hoffmann Jørgensen B, Badger J, Gylling Mortensen N. Wind resource estimation – an overview. *Wind Energy*; vol. 6 (3), pp. 261–271, 2003.
- [5] Xiaochen Wang, Peng Guo and Xiaobin Huang, A Review of Wind Power Forecasting Models, *Energy Procedia*, vol. 12, pp. 770–778, 2011.
- [6] Liye Xiao, , Feng Qian and Wei Shao, Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm, *Energy Conversion and Management*, vol. 143, pp. 410–430, 1 July 2017.
- [7] Ergin Erdem, Jing Shi, ARMA based approaches for forecasting the tuple of wind speed and direction, *Applied Energy*, vol. 88, pp. 1405–1414, 2011.
- [8] Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Kallos G, et al. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*; vol. 96(12), pp. 2348–2362, 2008.
- [9] Damousis IG, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion*, vol. 19(2), pp. 352–361, 2004.
- [10] Mabel MC, Fernandez E. Analysis of wind power generation and prediction using ANN: a case study. *Renew Energy*, vol. 33(5), pp. 986–92, 2008.
- [11] Barbounis TG, Theocharis JB. Locally recurrent neural networks for wind speed prediction using spatial correlation. *Information Sciences*, vol. 177(24), pp. 5775–5797, 2007.
- [12] Mohandes MA, Halawani TO, Rehman S, Hussain AA. Support vector machines for wind speed prediction. *Renewable Energy*; vol. 29(6), pp. 939–947, 2004.
- [13] Miranda MS, Dunn RW. One-hour-ahead wind speed prediction using a Bayesian methodology. *Proceedings of the IEEE power engineering society general meeting*, Montreal, Quebec, Canada; 2006.
- [14] Paulo Salgado and Paulo Afonso, Hybrid fuzzy clustering neural networks to wind power generation forecasting, *14th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2013)*, 19–21 November, pp. 359–363, 2013.
- [15] National System of Water Resources (SNIRH) ; <http://snirh.pt/>.
- [16] Anil K Jain, "Data clustering: 50 years beyond k-means", *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [17] Paul Zarchan; Howard Musoff (2000). *Fundamentals of Kalman Filtering: A Practical Approach*. American Institute of Aeronautics and Astronautics, Incorporated. ISBN 978-1-56347-455-2.
- [18] Hamilton, J., *Time Series Analysis*, Princeton University Press. Chapter 13, 'The Kalman Filter, 1994'.
- [19] Torres JL, Garcia A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre. *Sol Energy*; vol.79(1), pp. 65–77, 2005