Genetic Fuzzy Learning of local wind conditions

A. Agüera Pérez, J. J. G. de la Rosa, J. G. Ramiro, J. C. Palomares Salas

Research Unit PAIDI-TIC-168. University of Cadiz. Electronic Area. Escuela Politécnica Superior
Avda. Ramón Pujol, S/N. E-11202-Algeciras-Cádiz (Spain)
Phone/Fax number:+0034 956 028020, e-mail: josecarlos.palomares@uca.es,

Abstract. The wind climate measured in a point could be described as the regional wind climate of the zone forced by local effects derived from topography, roughness and obstacles in the surrounding area. This paper presents a method that allows to find fuzzy relations between the local wind climate in a location and the physical geography of the area. The fuzzy sets used in this work will connect relief elements with a probability relationed with the chances of flow over the concerned element. Thus, evaluating the heights and slopes of the terrain around the selected point, each fuzzy system will generate a possible local wind climate, which will be compared with the real one. In order to obtain the best agreement, a population of these fuzzy systems will be improved using Genetic Algorithms.

Keywords

Regional Wind Climate; Local Wind climate; Fuzzy Logic; Genetic Algorithm; Low quality data.

1. Introduction

The interest in the prediction of wind paremeters have been increased in the last decade in the same volume that wind farms have been installed. This fact is due to wind turbines have proved to be an efficient system to produce renewable energy in sites with enough potential. Hence, knowing with accuracy the wind resources available in an area, the production of this energy could be considered or ruled out, minimizing the risk of error and the costs derived from a measurement campaign with negative results.

Once real data of meteorological stations is acquired, Computational Fluid Dynamics (CFD) is the most frequent method used to generate a mesoscalar evaluation of wind potential [1][2]. CFD solves the fluid mechanic equations over a terrain with high computational cost and losing certainity when complex geography and chaotic dynamic arises. This model also needs high quality data to work efficiently because errors in measurements could be amplified. In this way, CFD needs high requirements to work and it is very sensible to errors in initial and boundary conditions. So, the most of the met stations scattered over a selected area (destinied to agriculture, fire detection, pollution,...) will be rejected to be used in a CFD model because of their low quality.

Fuzzy Logic, Artificial Neural Networks and other adaptative tools are statistical structures able to work with low requirements and high tolerance to possible errors [3]. Thus, using these techniques in the wind resource assessment, the data quality could be replaced with data quantity and the deterministic prediction with a probabilistic one, more inline with the atmospheric dynamic [4], [5].

2. Location and Wind Data

To illustrate this paper, a met station belonging to a net of agriculture measurement has been selected. This station is located at Jimena (Andalucia, Spain), inmersed in a complex terrain, and it registers daily mean wind speed and direction thanks to a wind vane and an anemometer placed at 2 meters above ground level. In the same way, three neighbor stations with similar characteristics have been collected in order to build a regional wind climate: Estepona, Vejer de la Frontera and Villamartín (Figure 3).

3. Regional Wind Climate

The regional wind climate used in this paper, which will be forced by the terrain conditions, is generated as a linear combination of real data registered at the stations around Jimena. This combination is based in interpolation functions for geostatistics variables (defined by G. Matheron in [6]), and its application to Atmospheric Sciences suggested by Thiebaux and Pedder [7]. Zekay Sen [8] have used these works to calculate the monthly mean wind speed in different areas in Turkye obtaining good results.

In this paper these techniques are used to generate a wind frequency rose calculating wind vectors in a point (v_{jimena}) as linear combination of real wind vectors recorded at the three stations mentioned before (v_1, v_2, v_3):

$$\vec{v}_{jimena} = a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 \qquad (1)$$

The coefficients a_i are normalized and indicate the contribution of each station to the final result. These coefficients will be calculated using studied geostatic weighting functions and normalizing:

$$a_i = \frac{W(r_i)}{\sum_j W(r_j)} \qquad (2)$$

Where r_i represents the distance between Jimena and the i*th* station, and $W(r_i)$ is the weight function which adopts this form (Thiebaux and Pedder):

$$W(r_i) = \begin{cases} \left(\dfrac{R^2 - r_i^2}{R^2 + r_i^2}\right)^\alpha & r_i < R \\ 0 & r_i > R \end{cases} \qquad (3)$$

The parameter α can be changed in order to modulate the distribution, and R is the distance that we choose beyond that the evaluated point is not affected. The values of this parameters (R=50 Km, $\alpha=2$) have been selected according with A. Agüera *et al*[9]. In this last work, it is possible to find detailed information of the interpolation system used here.

The result of this process is a temporal serie of wind vectors created with real information of the area. This serie can be analyzed in order to test the accuracy of the prediction. In figure 1 the real wind frequency rose at Jimena is compared with the rose of the regional wind climate generated by linear combinations of data from stations of the area.

As it is possible to see, the regional wind climate builded with this method gives an important aproximation of the real wind measured in the meteorological station of Jimena. But, in spite of this general similarity, there is an important difference when the expected incoming direction is E and Jimena measures SE. This is due to the linear interpolation predicts a theoretical wind in which no information of geographical conditions have been inserted. However, the wind measured in a point is affected by local topography, and its effect may be evaluated in order to get a serious prediction. The objective of this work is correct the regional wind climate introducing local geographic conditions using fuzzy systems.
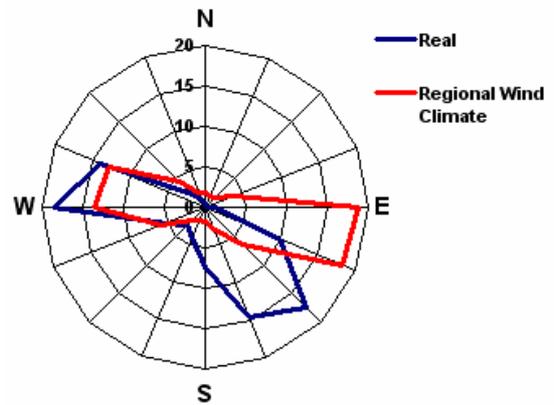
Fig.1. Real and predicted wind frequency roses at Jimena.
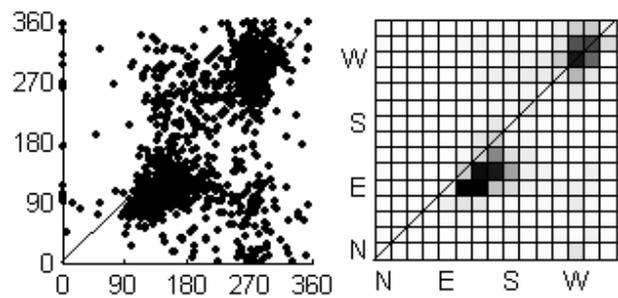
Fig.2.The left graphic represents the correlation between real data measured at Jimena (x-axis) and the regional wind generated by linear interpolation (y-axis). The right one is the same graphic, but grouped in 16 sectors.

The correlation between real and generated values must be analyzed in order to extract preliminary conclusions of the deviations produced (Fig. 2).

In general, there is high correlation, and wind directions from W are very well predicted. But, it is clearly visible how directions around E are displaced to the S. This effect have been observed and commented in the discussion of figure 1, but the information in the correlation graphic is better. From this graphic it is possible to extract diagrams like figure 3, where the real measurements in Jimena are represented when the regional wind climate predicts E. There is a probability that some directions expected as E were measured in Jimena as W. This effect could not be infered from the roses graphic.
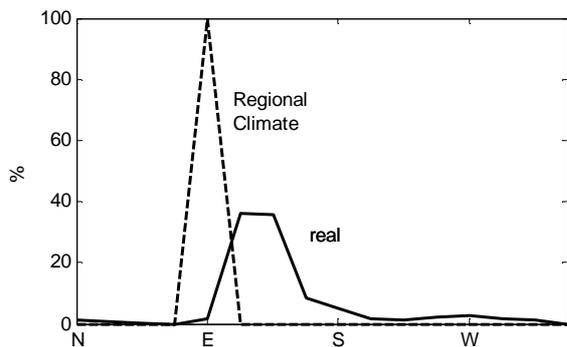
Fig.3. Percentage of real measurements in each direction when the regional wind climate indicate E.

In this paper it is assumed that the diferences between regional and real wind climate are a consequence of the physical geography around the station. So, we will try to generate a better prediction developing a system able to correct the regional wind introducing local relief information.

4. Terrain

In figure 4 is represented the process through which an altimetric map image is transformed in the matrix used as terrain input in this model. In the altimetric image each colour is associated to a height above sea level. Then, reading the RGB components of each pixel, a height matrix of the area can be created and represented (figure 4a).

The wind measured in a point could be defined as a vector whith module and direction equal to wind speed and wind direction. Hence, the simmetry of this problem is radial and input arguments should be expressed in polar coordinates. Figure 4b shows the mean heights of the digitized area, fractioned in sectors (M_{ij}) depending on the polar coordinates relatives to the central point (Jimena).

Each M_{ij} is obtained as the mean height of the n pixels inside the ij-sector (1):

$$M_{ij} = \frac{\sum_{n} h(r,\theta)}{n} \qquad (1)$$

The matrix M, composed of Mij, is represented in Figure 4c. As it is shown in the graphic explanation below, each column contains information of one direction. So a turn in the wind direction will be interpreted as a displacement in x-axis. M gives a usefull representation of the physical geography around the calculation point, adapted to the simmetry of the problem, and it is easily applied as input of fuzzy systems.

Fig.4. Area of Jimena and transformations to obtain the model terrain input from a map image.

5. Genetic Fuzzy System

The fuzzy system used in this problem pretends to connect the local wind conditions with terrain characteristics of the area. The vagueness in the terrain description in this scale, the quality of data recorded at the used stations, and the chaotic dynamics inherent to atmospheric events demand a fuzzy treatment of these elements. The proposed system will calculate the probability of possible changes in the direction of the wind analyzing the terrain in those directions. The

matrix M, mentioned in the previous section, has been created attending to these aspects and it helps to evaluate relief characteristics as heights, slopes or complexity in polar coordinates. In order to simplify the problem this study will be focused in the effects of the slopes in the wind deviation. So, two inputs parameters will be needed: a description of the terrain will be done considering the "slope" and the wind behavior will be represented by a "deviation" in the incoming direction. Using these inputs arguments the fuzzy system should infer the value of the output, "probability", which represents the chances that a given wind has to change the direction according to the slopes. References [10], [11] are important bibliography to understand and study in depth the concepts used in this section.

A. Fuzzy Sets

Fuzzy Logic is based on the fuzzyfication of crisp variables obtaining fuzzy sets, so linguistic terms are generally used to build the new description. For example, a fuzzyfication of the parameter "slope", defined in the range [-500 m, 500 m] as the difference of heights between two elements of M, could be done using this fuzzy sets [down, up]. Each fuzzy set is defined in a domain and it is characterized by a membership function with values between 0 and 1 which indicates the degree of membership of an element to the concerned fuzzy set. So the parameter "slope" could be fuzzyficated as it is shown in figure 5. In this figure is also represented a possible election of fuzzy sets for the parameter "deviation", [low, medium, high], defined between 0 and 45 degrees. "Probability", whith range [0, 1] will be divided with two sets [low, high].

B. Fuzzy Rules

Using these fuzzy sets it is possible to to build "IF-THEN" rules connecting the inputs with the output generating a knowledge base. For example:

IF "deviation" is *high* and "slope" is *up* THEN "probability" is *low*.

Now it is possible to map crisp input values into output ones by using the processes of fuzzyfication, fuzzy inference and defuzzyfication.
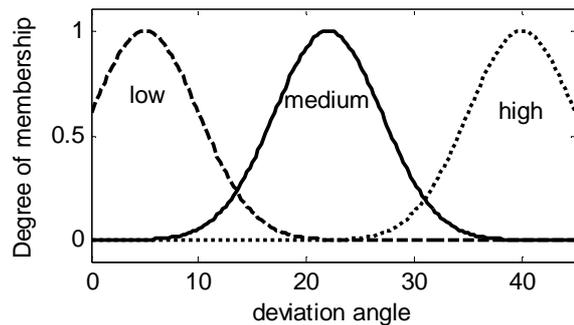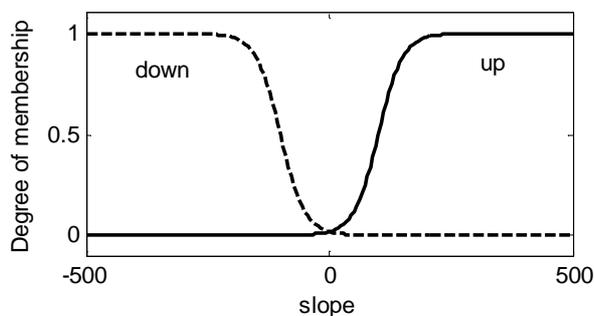
Fig.5. Fuzzy sets over the domain of slopes and deviantion angles.

C. Genetic Algorithm

The membership functions and fuzzy rules determine the response of a fuzzy system to the input arguments. So, the choice of these characteristics is the most important step to get a useful knowledge base. This knoledge base could be defined and adjusted by a human expert or founded using adaptatives tools. In this paper we will try to optimize the fuzzy system characteristics with Genetic Algorithms (GA). So, we must encode the fuzzy system characteristics into a string of values which will be considered a genome of this system. Mutation and Crossover operators can act on these strings generating new fuzzy systems; and a selection operator can select individuals according to their fitness values obtained from a fitness function or objective. The action of these operators brings an iterative process through which an initial population of fuzzy systems could be improved in order to obtain the best fitness value, calculated using a fitnees function that will be presented an analyzed later.

We have chosen 19 parameters (genes) to characterize each fuzzy system (Table 1). Seven of them are relationed with the position and shape of the membership's functions; so, they are continous. The other ones are discrete variables used to build the fuzzy rules of the system.

Table I

Gen	Values	Parameter
1	[0, 1]	Width Slope / down
2	[0, 1]	Width Slope / up
3	[0, 1]	Position Slope / down
4	[0, 1]	Position Slope / up
5	[0, 1]	Width Deviation / low-medium-high
6	[0, 1]	Width Probability / low
7	[0, 1]	Width Probability / high
8	0, 1	Fuzzy rule 1 / input 1
9	0, 1	Fuzzy rule 2 / input 1
10	0, 1	Fuzzy rule 3 / input 1
11	0, 1	Fuzzy rule 4 / input 1
12	0, 1, 2	Fuzzy rule 1 / input 2
13	0, 1, 2	Fuzzy rule 2 / input 2
14	0, 1, 2	Fuzzy rule 3 / input 2
15	0, 1, 2	Fuzzy rule 4 / input 2
16	0, 1	Fuzzy rule 1 / output 1

17	0, 1	Fuzzy rule 2 / output 1
18	0, 1	Fuzzy rule 3 / output 1
19	0, 1	Fuzzy rule 4 / output 1

D. Fitness Function

The prove suggested to test and rank the fuzzy systems is minimize the diferences between the regional wind climate and the real one observed in Jimena, showed before in Figure 3. The fuzzy system which predicts more successfully the real distribution using the E incoming direction will be selected.

Using an iterative process, the regional direction will be corrected by interaction with M, the matrix which contains the processed information of the map (Figure 4 c). In each iteration the wind direction will be defined by its position in M, and the fuzzy system will evaluate the five upper elements of M giving the probability of flow over each element (Figure 6). As result of this, a probability distribution will be calculated, depending on the fuzzy system characteristics. In some way, the obtained probability distribution has been affected by topographic elements.
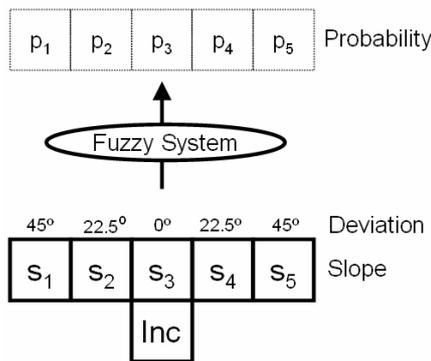
Fig.6. Corrections of directions using the fuzzy system, where "Inc" indicates the incoming direction, s_i and p_i the slope and probability in each direction .

6. Results

After the minimization process explained before, an optimized fuzzy system has been obtained. The correction at the probability distribution given by this fuzzy system is shown in figure 7. The accumulated probability in each element of M is shown below. This graphic could be interpreted as a distribution of possible trajectories that wind could trace while blowing from the East direction and the final result is the corrected distribution.
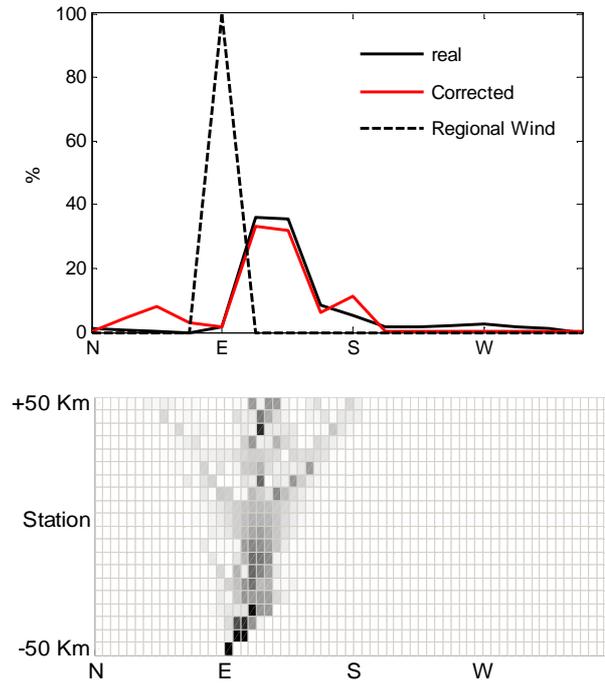
Fig.7. Comparison among regional, corrected and real distributions of directions. Below it is shown the density of trajectories calculated by the optimized fuzzy system.

The optimized fuzzy system selected by the GA presents the distribution of fuzzy sets shown in figure 8, and the four rules obtained are shown in table 2.

Table II

IF	Slope	Deviation		Probability
	Down	Medium	THEN	Low
	Up	High		Low
	Down	High		Low
	Down	Low		High
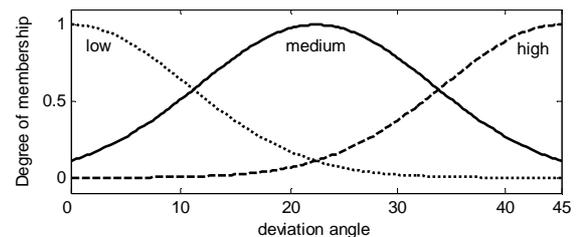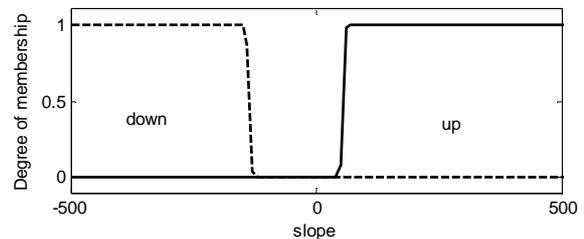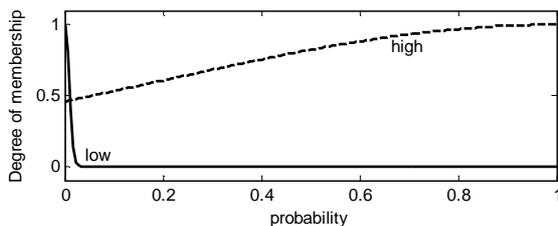
Fig.8. Optimized fuzzy sets.

Note that the "slope" range [-100, 20] is undefined. These elements are not members of "up" or "down", what could be an indicator of the need of a new fuzzy set ("plain").

Due to our system only use two inputs values, a surface of inference could be traced as figure 8 shows, where slope and deviation are connected to the output, probability.

7. Conclusions and Future Work

It is clearly visible the accuracy of the distribution obtained with this method, this implies that the optimized fuzzy system contains information about how to correct the measurements from direction E using the slope of terrain. This information could be treated as a real explanation of the phenomenon or as a mathematical coincidence. Anyway, the system predicts the local wind climate with high accuracy evaluating the terrain slopes and the changes in the direction.
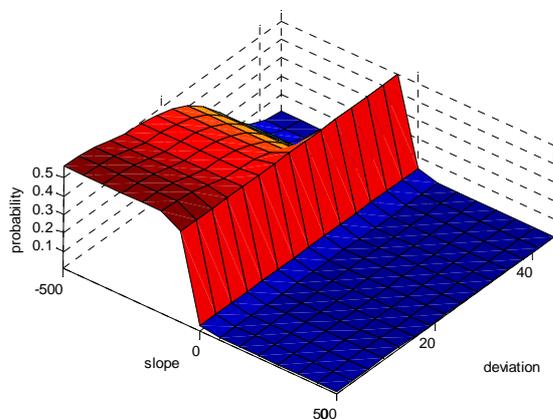
Fig.9. Inference surface.

Observing figure 8, the fuzzy system induces the wind to flow to lower terrains, except if this implies a turn higher than 30 degrees in the direction. Turns of more than 30 degrees are allowed only when the terrain is a plain (slope = [-100, 20]), and when this is the case, there is no favourite direction. The probability is almost 0 when the wind find a positive slope in any direction. These conclusions report a basic description of the wind behavior (real or not) related to the input arguments imposed. So, it is expected that using more variables, a more detailed description could be builded, obtaining a better correlation between real and calculated data.

References

[1] J.M.L.M. Palma, F.A. Castro, L.F. Ribeiro, A.H. Rodrigues, A.P. Pinto; Linear and nonlinear models in wind resource assessment and wind turbine micro-siting in complex terrain; Journal of Wind Engineering and Industrial Aerodynamics 96 (2008) 2308– 2326.

[2] Martín Gastón, Edurne Pascal, Laura Frías, Ignacio Martí, Uxue Irigoyen, Elena Cantero, Sergio Lozano, Yolanda Loureiro; Wind resources map of Spain at mesoscale. Methodology and validation. Wind Energy Department, National Renewable Energy Center (CENER).

[3] J. M. Gutiérrez, Rafael Cano, Antonio S. Cofiño, Carmen M. Sordo; *Redes probabilísticas y neuronales aplicadas a las ciencias atmosféricas*, INM, Ministerio de Medio Ambiente, Madrid.

[4] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, G. Kallos. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering; Journal of Wind Engineering and Industrial Aerodynamics 96 (2008) 2348– 2362.

[5] M. Cellura, G. Cirrincioneb, A. Marvuglia, A. Miraouic; Wind speed spatial estimation for energy planning in Sicily: A neural kriging application; Renewable Energy 33 (2008) 1251-1266.

[6] G. Matheron, *Principles of geostatistics*, Econ. Geol. 58 (1963) 1246-1266

[7] H.J. Thiebaux, M.A. Pedder, Spatial Objective Analysis with Applications in Atmospheric Sciences, Academic Press, New York, 1987, 299 pp.

[8] Zekai Sen, Regional, Ahmet Duran Sahin, Wind energy evaluation in some parts of Turkey, J. Wind Eng. Ind. Aerodyn., 74-76 (1998) 345-353.

[9] A.Agüera, J. G. Ramiro, J. Melgar, J. J. G. de la Rosa, J.C. Palomares, A. Moreno. "Categorization of minimum error forecasting zones using a geostatistic wind model"; *Proceedings Clean Electrical Power 2009 Capri (Italy)*, IEEE.

[10] Lofti A. Zadeh; *Fuzzy sets, fuzzy logic and fuzzy systems*, Advances in fuzzy Systems – Applications and Theory vol..6. World Scientific. 1997.

[11] Elie Sanchez, Takanori Shibata, Lotfi A. Zadeh. *Genetic Algorithms and Fuzzy Logic Systems*. Advances in fuzzy Systems – Applications and Theory vol.7. World Scientific. 1997