

Article

A New MCP Method of Wind Speed Temporal Interpolation and Extrapolation Considering Wind Speed Mixed Uncertainty

Xiao Liu ¹, Xu Lai ^{1,*} and Jin Zou ²

¹ School of Water Resource and Hydro Electric Engineering, Wuhan University, Wuhan 430072, Hubei, China; liuxiaowhu@126.com

² Electric Power Research Institute, China Southern Power Grid, Guangzhou 510080, Guangdong, China; zoujin@csg.cn

* Correspondence: laixu@whu.edu.cn; Tel.: +86-134-7613-6201

Academic Editor: Sukanta Basu

Received: 2 July 2017; Accepted: 14 August 2017; Published: 18 August 2017

Abstract: In this paper, a missing wind speed data temporal interpolation and extrapolation method in the wind energy industry was investigated. Given that traditional methods have previously ignored part of mixed uncertainty of wind speed, a concrete granular computing method is constructed and a new Measure–Correlate–Predict (MCP) method of wind speed data temporal interpolation and extrapolation considering all mixed uncertainties is proposed, based on granular computing theory by adopting the cloud model method, support vector regression method, artificial neural network, genetic algorithm, and fuzzy c-means clustering algorithm as tools. The importance of considering mixed wind speed uncertainty and the suitability of using granular computing method are illustrated, and wind speed mixed uncertainty analysis is implemented, then, recommended values and estimation tools for wind speed measurement uncertainty and combined uncertainty are provided. An interpolation case of two practical meteorological sites in central Southern China was used to implement and validate the method proposed in this paper. The following conclusions are reached: (a) by using the method proposed in this paper, mixed uncertainty of wind speed can be considered, comparing to other MCP methods used for purposes of comparison, a better estimation of the wind speed is provided, and most evaluation metrics employed in this analysis were superior to other methods, that is to say, the accuracy of the wind resource assessment improved, and the risks of wind farm construction were reduced; (b) granular computing method is suitable for the issue of wind speed data interpolation and extrapolation considering wind speed mixed uncertainty; (c) mixed uncertainty of wind speed can be divided into three levels, and recommended values of granularity are minimum interval of records, 0.3–0.8 m/s, and 1–3 m/s, respectively.

Keywords: measure-correlate-predict method; mixed uncertainty of wind speed; granular computing theory; cloud model; support vector regression; neural network; genetic algorithm; fuzzy c-means clustering algorithm

1. Introduction

With increasing global warming and environmental problems, wind energy has now become one of the most important sources of green energy. Wind resource assessment is a key procedure in wind farm construction, and wind speed observation is the foundation of wind resource assessment [1,2]. However, in many practical cases, the use of wind speed observation records is restricted by two factors, one is gaps in the wind speed records caused by instrument failure or the destruction of the meteorological mast, the other is lack of long-term wind speed records. To overcome these

drawbacks, one or more reference sites have been chosen and the relationship between the target and reference sites have been constructed using statistical methods (Measure–Correlate–Predict (MCP) methods [3–17]), or physical methods (for example, Wind Atlas Analysis and Application Program (WAsP) [18–20]). As per references [13,20], statistical methods tend to provide higher accuracy and are therefore used widely in practical engineering. The MCP methods model correlations between the target and reference sites using simultaneous records; then, those correlations are used with reference site data to predict missing wind data or long-term wind data of the target site. Previous researchers have provided some effective MCP methods that used different types of functions to model correlation: Linear regression [3–9] has been widely used to model correlation; Derrick [3] proposed MCP method to assess long term wind resources by using linear regression method, and similar works have been done by other researchers [4–7]. MCP methods based on linear regression have been validated in some practical applications [8,9]. Support vector machine (SVM) model was also used for correlation, Mohandes et al. [10] introduced the SVM for wind speed prediction and compared it with the multilayer perceptron (MLP) neural networks, the results proved that the SVM model had less root mean square errors than the MLP model. Ji et al. [11] did further research on SVM, a support vector classifier was utilized to estimate the forecasting error and lower mean square error and mean absolute percentage error than traditional SVM method were obtained. Artificial neural networks (ANN) [12–15] method was another tool to model correlation, Addison et al. [13] investigated the feasibility of using neural networks to make predictions of long term energy yield at a potential wind farm site. Saavedra-Moreno et al. [14] provided very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms. Probabilistic method [16,17] also was used to model correlation, Carta et al. [16] proposed the use of a probabilistic Measure–Correlate–Predict (MCP) method to estimate the long-term wind speed characteristics at a potential wind energy conversion site. García-Rojo [17] indicated an automatic method of calculation of a long-term representative wind climate based on the calculation of the joint probability distribution of the wind at a local station and at a long-term reference meteorological mast.

However, there are some shortages and limitations of these existing methods. Wind speed uncertainty is a mixed uncertainty which mainly consist of randomness and fuzziness, specifically, randomness is an aleatory uncertainty due to stochasticity of nature wind speed, and fuzziness is an epistemic uncertainty mainly caused by measurement. The existing methods have been used to treat wind speed data as accurate quantity, and only randomness has been dealt with probability theory or artificial intelligence method, such as regression or ANN, fuzziness of wind speed would always be ignored. Therefore, inaccurate and unscientific predictions have been obtained by ignoring some components of uncertainties, and then errors of wind resource assessment may occur and risks of wind farm construction are raised.

Granular computing theory (GrC) references and imitates human cognition, and extracts the common theory of problem solving; therefore, it could be used to exploit the tolerance for imprecision, uncertainty, and partial truth to achieve tractability, robustness, low cost solutions, and better rapport with reality [21]. All components of wind speed mixed uncertainty can be considered and the drawbacks of traditional MCP methods can be overcome by using MCP method based on granular computing theory.

For the purpose of considering wind speed mixed uncertainty, this paper constructed a concrete granular computing method, and proposed a new MCP method of wind speed temporal interpolation and extrapolation. As the foundation of the new method, wind speed uncertainty analysis was implemented, recommended values and estimation tools for wind speed measurement uncertainty and combined uncertainty were provided, and the uncertainty hierarchy was determined. The new method used cloud model, support vector regression (SVR), artificial neural network (ANN), genetic algorithm (GA), and fuzzy c-means clustering algorithm (FCM) as tools, and illustrated a detailed procedure of the new method. In order to implement and validate the new method, a wind data temporal interpolation case was studied where four metrics were used: mean relative error, root mean square error, correlation coefficient, and mean relative error of energy

production; and also a comparison of these metrics with those obtained from other methods was implemented.

2. Granular Computing Theory

GrC references and imitates human cognition, and may be regarded as a label for theories, methodologies, techniques, and tools that make use of granules, such as groups, classes, or clusters of a universe in the process of solving a problem, and is a more philosophical way of thinking than a practical methodology to problem solving [21]. Zadeh [22] identified three basic concepts that underlie human cognition: granulation, organization, and causation. Granulation involves the decomposition of a whole into parts; organization involves the integration of parts into a whole; and causation involves the association of cause and effect. There are four steps in the process of using GrC for problem solving, which are partitioning objects into granules, constructing levels, and hierarchy, representing granules and computing with granules [23,24].

The benefits of GrC are evident from its basic guiding principle, which was stated concisely by Zadeh [22] as to “exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness, low cost solutions and better rapport with reality”.

Granules, levels, and hierarchy are the basic elements of granular computing. Specifically, (a) a granule is a clump of points (objects) drawn together by indistinguishability, similarity, proximity or functionality [22,25]; (b) levels consist of granules, and the properties of granules collectively characterize a level of description and understanding; and (c) hierarchy is the sum of the relationships of levels.

3. Uncertainty Analysis of Wind Speed

Uncertainty, an inherent characteristic of the nature of wind, turns wind speed into granules, and then granules constitute levels, and then levels constitute hierarchy. Finally, granules, levels, as well as hierarchy all together constitute the granular world. As the foundation of the granular world is based on uncertainty, an uncertainty analysis of wind speed is the foundation of the proposed method. In this section, the hierarchy is determined, and recommended values and estimation tools for wind speed measurement uncertainty as well as combined uncertainty are provided.

3.1. Method for Calculating Uncertainty

Uncertainty is defined as a parameter associated with the results of a measurement that characterizes the dispersion of values that can reasonably be attributed to the measurement, and uncertainties are expressed as standard deviations and are denoted standard uncertainties [26].

The Bins method was used to determine the uncertainties as it is a data reduction procedure where wind speed data are divided into 1 m/s continuous intervals (bins) centered on multiples of 1 m/s, with uncertainties calculated in each bin [27].

Uncertainty consists of some components, and in general, the combined standard uncertainties in bin i can be expressed by

$$u_{c,i}^2 = \sum_{k=1}^M \sum_{l=1}^M c_{k,i} u_{k,i} c_{l,i} u_{l,i} \rho_{k,l} \quad (1)$$

where $c_{k,i}$ is the sensitivity factor of component k in bin i ; $u_{k,i}$ is the standard uncertainty of component k in bin i ; M is the number of uncertainty components in each bin; $\rho_{k,l}$ is the correlation coefficient between component k in bin i , and component l in bin i .

To simplify the above expression, the independence of all uncertainty components was assumed, so $\rho_{k,l} = 0$ when $k \neq l$, and $\rho_{k,l} = 1$ when $k = l$. As the uncertainty components

and combined uncertainty are in the same dimension, $c_{k,i} = 1$. Combined uncertainty can be simply expressed by

$$u_{c,i}^2 = \sum_{k=1}^M u_{k,i}^2 \quad (2)$$

Combined standard uncertainties may additionally be expressed by expanded uncertainties. Referring to the International Organization for Standardization (ISO) guide [26] and assuming normal distributions, combined standard uncertainties were multiplied by a coverage factor of 2 in a 95% confidence level.

3.2. Hierarchy of Uncertainty

Some sources of wind speed uncertainty exist in the issue of MCP method, they are wind speed measurement, the influence of local obstacles, meso-micro scale weather patterns, stochastic variations in wind speed and direction over time and distance, as well as flight delays and atmospheric stability [9]. We should note here that wind speed uncertainty is a mixed uncertainty, which is consisting of randomness and fuzziness.

In practical cases, fuzziness is decided by the campaign of measurement, and they are independent of specific conditions and are most influenced by the instrument of measuring; on the other side, randomness is decided by other sources of wind speed uncertainty, it is strongly dependent on one or multiple specific conditions.

The construction of certain uncertainty levels indicated that the uncertainty components contained in the level are formulated and calculated, so continuity, uniformity, computability, and independence of uncertainty components are demanded. Uncertainty of wind speed measurement, which is represented by fuzziness, meets the requirements and is used to construct a single level; on the other side, randomness is considered in the combined uncertainty rather than a single level.

In this paper, three levels of mixed uncertainty were constructed: bottom level, which ignores all uncertainties; middle level, which includes the measurement uncertainties; and top level, which considers all components of mixed uncertainty.

3.3. Uncertainty of Wind Speed Measurement

To determine and characterize every level of uncertainty, some practical cases were collected. Measurement uncertainty represents fuzziness of wind speed. There are four sources of wind speed measurement uncertainty: anemometer operational characteristics, mounting effects, anemometer calibration, and data acquisition.

3.3.1. Operational Characteristics Uncertainty

Anemometers which were used to wind speed measurement are calibrated in wind tunnel, and these wind tunnel calibrations are performed under controlled environmental conditions with a smooth and low turbulent flow. However, during measurements and operation in the field, these instruments are subject to turbulent flow and environmental conditions that can deviate significantly from wind tunnel calibration conditions. The field conditions may significantly influence instrument characteristics and cause operational characteristics uncertainty.

To calculate operational uncertainties, a classification method of an anemometer was provided [27]. Known general influence parameters on cup anemometer measurements are: turbulence, air temperature, air density, and average flow inclination angle, these influence parameters must be considered in the classification, hence the classification of an anemometer type is divided into two classes, which is dependent on influence parameters. Details are shown in Table 1.

Table 1. Class A and Class B operational ranges.

Classification Category	Class A		Class B	
	Ideal Flat Terrain Sites		Non-Ideal Complex Terrain Sites	
	min	max	min	max
Wind Speed Range (m/s)	4	16	4	16
Turbulence Intensity	0.03	0.12 + 0.48/V	0.03	0.12 + 0.96/V
Turbulence Structure ($\sigma_u / \sigma_v / \sigma_w$)	1/0.8/0.5		1/1/1	
Air Temperature (°C)	0	40	-10	40
Air Density (kg/m ³)	0.9	1.35	0.9	1.35
Average flow inclination (°)	-3	3	-15	15

The classification number k was defined and was used in determination of the operational uncertainty, and a classification, referring to a certain class, is expressed by kA , kB , for example 1.7A or 2.5B. The operational standard uncertainty of a cup anemometer may be derived from the classification number assuming a rectangular uncertainty distribution by using Equation (3).

$$u_{m,i} = (0.05m/s + 0.005 \cdot U_i) \cdot k / \sqrt{3} \tag{3}$$

where $u_{m,i}$ is the operational standard uncertainties in bin i ; U_i is the center of bin i ; and k is the classification number of the anemometer.

In this paper, the classification numbers of five popular cup anemometers, which are widely used in practical cases and represent the average uncertainty of cup anemometers, were collected to determine the approximate level of operational uncertainties, and these cup anemometers were measured by ACCUWIND [28]. The uncertainties are presented in Table 2 and Figure 1.

Table 2. Classification number of five popular anemometers.

Cup Anemometer	Classification Number	
	Class A	Class B
NRG #40	2.4	7.7
RISO P2546	1.9	8.0
Thies FC	1.5	2.9
Vaisala WAA151	1.7	11.1
Vector L100	1.8	4.5

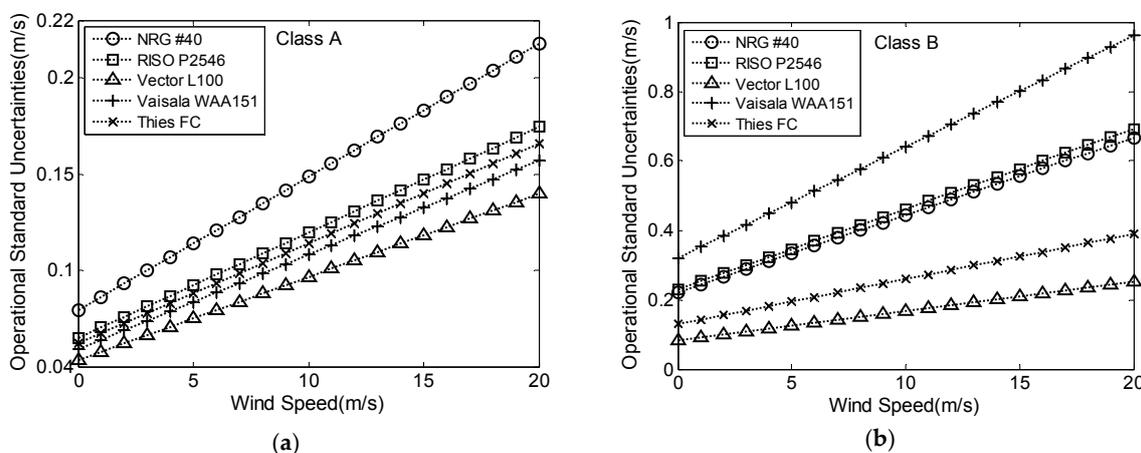


Figure 1. Operational standard uncertainty of five anemometers: (a) in class A operational ranges; (b) in class B operational ranges.

Figure 1 shows that the ranges of operational standard uncertainty are about 0.04–0.21 m/s in Class A and are 0.13–0.96 m/s in Class B, and the average operational standard uncertainties are about 0.1 m/s in Class A and 0.4 m/s in Class B.

3.3.2. Mounting Effects

A lattice mast with a side mounted anemometer is widely used in practical applications, therefore, anemometers are always influenced by flow distortion from both the mast and the boom. Thrust coefficient, C_T , which depends upon the porosity of the mast and the drag on the individual members, is assumed to evaluate influence. The average C_T of the lattice mast with a square cross-section and shape edge members, square cross-section and round members, triangular cross-section and round members were 0.92, 0.42, and 0.34, respectively. The mounting standard uncertainty can be estimated by Equation (4) below [27].

$$u_{m2,i} = (0.062C_T^2 + 0.076C_T) \cdot \left(\frac{L_m}{R} - 0.082 \right) \cdot U_i / \sqrt{3} \quad (4)$$

where $u_{m2,i}$ is the mounting standard uncertainty in bin i ; C_T is the thrust coefficient of mast; L_M is the mast leg length; and R is the distance between the observation position and the center of the mast.

The results of the center-line relative wind speed of lattice mast with average C_T are shown in Figure 2.

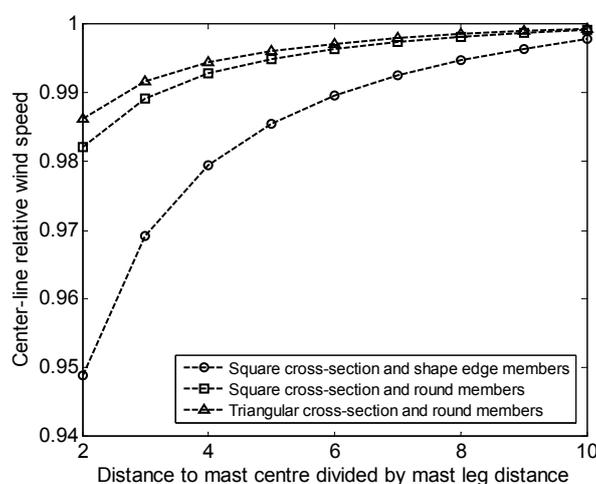


Figure 2. Center-line relative wind speed of lattice mast with average C_T .

In practical cases, the distance between the observation position and mast center was 3–5 times the mast leg length, therefore, we can conclude from Figure 3 that the range of center-line relative wind speed was 0.969–0.996 of free wind speed.

3.3.3. Uncertainty of the Anemometer Calibration

To obtain an overview of the cup anemometer calibration uncertainty, calibration uncertainties of ten commonly used cup anemometers, which have been quantified by previous researchers [29], were collected. The details are shown in Table 3, and the linear regression of average relative uncertainties of ten anemometers is proposed in Figure 3.

Table 3. Relative calibration uncertainty of ten commonly used anemometers.

Type of Anemometer	Relative Calibration Uncertainties of Anemometer Calibration											
	4 m/s	6 m/s	8 m/s	10 m/s	12 m/s	14 m/s	16 m/s	18 m/s	20 m/s	22 m/s	24 m/s	26 m/s
NRG #40	2.65%	2.00%	1.62%	1.49%	1.45%	1.43%	1.38%	1.23%	1.08%	1.43%	1.03%	0.90%
NRG IF3	4.04%	2.74%	1.98%	1.81%	1.57%	1.40%	1.18%	1.12%	1.10%	1.01%	1.07%	0.90%
Risoer cup	2.15%	1.88%	2.00%	1.53%	1.34%	1.29%	1.25%	1.29%	1.19%	1.10%	1.04%	1.11%
R.M Young Wind Monitor	1.47%	1.05%	0.84%	0.76%	0.68%	0.64%	0.61%	0.60%	0.60%	0.61%	0.58%	0.58%
R.M Young Wind Sentry	1.53%	1.06%	1.03%	1.02%	1.08%	1.02%	0.95%	0.94%	0.84%	0.90%	0.84%	0.99%
Second Wind C3	2.74%	2.19%	2.14%	1.66%	1.60%	1.47%	1.51%	1.45%	1.34%	1.31%	1.09%	1.00%
Thies First Class	2.71%	2.24%	1.83%	2.70%	2.29%	1.74%	1.87%	2.21%	1.73%	1.82%	1.76%	1.58%
Vaisala WAA252	2.70%	2.04%	1.90%	1.87%	2.07%	1.89%	1.80%	1.92%	2.05%	1.84%	1.86%	1.68%
Vector V100LK	2.46%	2.13%	2.22%	2.27%	2.04%	2.02%	1.92%	2.01%	1.83%	2.03%	2.15%	1.66%
Vestas Cup	1.50%	1.19%	1.35%	1.10%	1.12%	1.20%	1.14%	0.96%	0.83%	0.85%	1.00%	0.78%
average relative uncertainties	2.40%	1.85%	1.69%	1.62%	1.52%	1.41%	1.36%	1.37%	1.26%	1.29%	1.24%	1.12%

The relative calibration uncertainties in Table 3 were quantified in the wind tunnels at Otech Engineering, Inc. located in Davis, California [29].

The calibration uncertainty of a commonly used anemometer can be concluded from Figure 3 and take the following expression.

$$u_{m3,i} = U_i \cdot (-0.000431 \cdot U_i + 0.021578) / \sqrt{3} \tag{5}$$

where $u_{m3,i}$ is the calibration standard uncertainties in bin i .

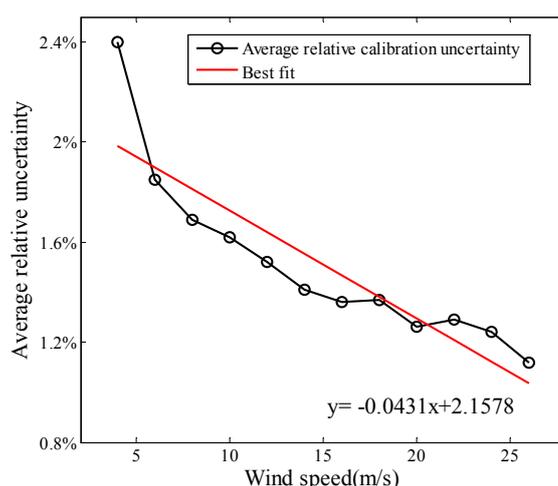


Figure 3. Average relative calibration uncertainty of ten commonly used cup anemometers with best linear regression fit.

3.3.4. Uncertainty of Data Acquisition

Uncertainty of Data Acquisition may come from transmission, signal conditioning, analog to digital conversion, and data processing in the data acquisition system.

Assuming that the uncertainty of data acquisition system was 0.1% of the wind speed measured range, the standard uncertainty from data acquisition is

$$u_{m4,i} = 0.001 \cdot U_r \tag{6}$$

where $u_{m4,i}$ is the data acquisition standard uncertainties in bin i ; and U_r is the wind speed measured range.

3.3.5. Combined Uncertainty of Wind Speed Measurement

The combined uncertainty of wind speed measurement can be derived from Equation (3) to Equation (6), and take the following expression.

$$u_{m,i}^2 = \sum_{k=1}^4 u_{mk,i}^2 = 6.19 \times 10^{-8} \cdot U_i^4 - 6.20 \times 10^{-6} \cdot U_i^3 + \left[8.3 \times 10^{-6} \cdot k^2 + \left((0.062C_T^2 + 0.076C_T) \cdot \left(\frac{L_m}{R} - 0.082 \right) \right)^2 / 3 + 1.55 \times 10^{-4} \right] \cdot U_i^2 + 1.67 \times 10^{-4} \cdot k^2 \cdot U_i + 8.3 \times 10^{-4} \cdot k^2 + U_r \cdot 10^{-6} \tag{7}$$

where $u_{m,i}$ is the wind speed measurement uncertainty in bin i .

In practical conditions, the recommended value of wind speed measurement uncertainty was about 0.3–0.8 m/s.

3.4. Combined Uncertainty

Combined uncertainty of wind speed includes all components of wind speed mixed uncertainty. In this paper, three operational condition classes with two practical meteorological sites were selected to obtain an overview of the combined uncertainty, location, topography, distance, and correlation of the chosen classes are shown in Table 4.

Table 4. Topography, distance, and correlation coefficient of the classes.

Groups	Topography	Vegetation	Distance (km)	Correlation Coefficient
simple	plane	grass/crops	8	0.793
complex	hill	forest	40	0.733
extremely complex	mountain	forest	15	0.872

The combined uncertainty of correlation can be characterized by standard deviation, which was calculated from a set of simultaneous reference wind speed data at a specific wind speed of target data, for example, when the wind speed of target data was 5.0 m/s, the set of contemporary reference wind speed had a standard deviation of 1.6 m/s, so the combined uncertainty at the wind speed of 5.0 m/s was 1.6 m/s. Results of the combined uncertainty are shown in Figure 4.

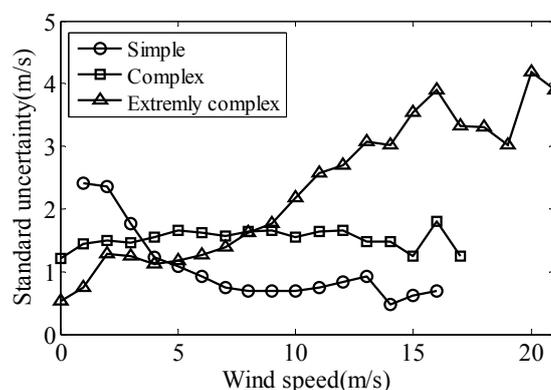


Figure 4. Standard uncertainties of three operational condition classes.

Hence, it can be concluded from Figure 4 that the range of combined uncertainty at each specific wind speed was 1–3 m/s.

4. Proposed Method

In this section, a concrete granular computing method is constructed, and a new Measure–Correlate–Predict (MCP) method of wind speed data temporal interpolation and extrapolation considering uncertainties are proposed. The detailed procedure of proposed method is illustrated as follows.

4.1. Procedure of Proposed Method

The 5 steps of the proposed method are presented below and details are shown in Figure 5: (a) determining the granular hierarchy; (b) granulation; (c) granules representation; (d) granules computing; and (e) interpolation or extrapolation.

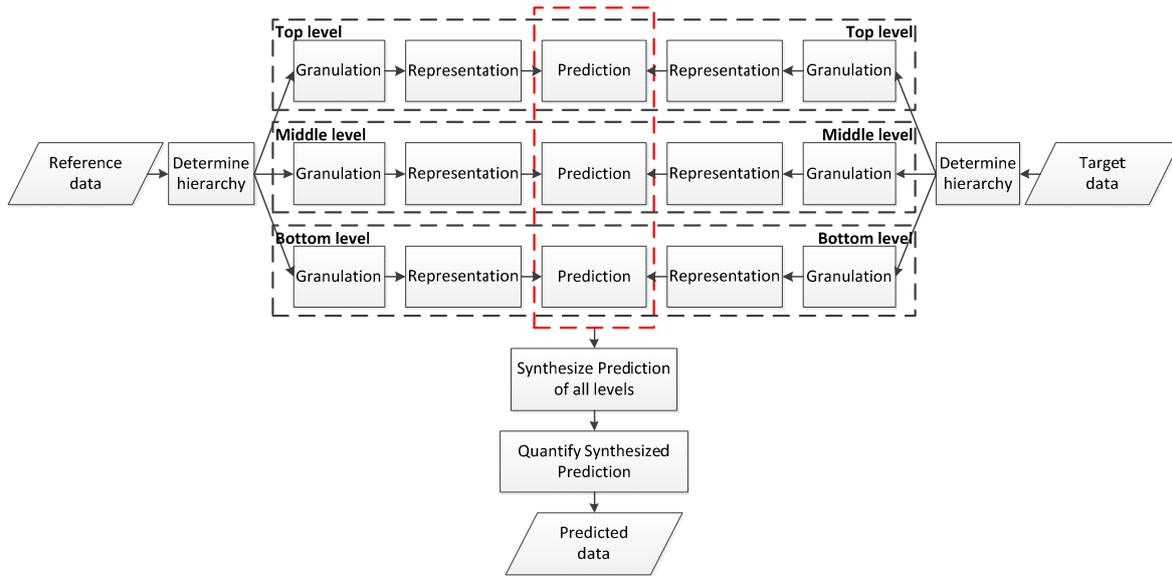


Figure 5. A schematic view of wind speed data interpolation and extrapolation based on the granular computing method.

4.2. Determining the Granular Hierarchy

The foundation of wind speed granules is composed of uncertainties. In this paper, the hierarchy of wind speed granules was decided by the structure of uncertainty, of which three levels of granules were set.

Firstly, in the bottom level, all uncertainties were ignored and wind speed data were supposed to be precise, and then this level was made of basic granules, which were the minimum interval of wind speed recording by anemometer. So, actually, the granule computing in the bottom level was equal to traditional methods. In practical conditions, minimum interval of most anemometer records was 0.1 m/s, so, the recommended value of granularity magnitude in this level was 0.1 m/s.

Secondly, in the middle level, only the continuous, uniform, computable, and independent uncertainties, which were measurement uncertainties, were considered. Therefore, the magnitude of granularity was calculated by Equation (7). In practical conditions, the recommended value was approximately 0.3–0.8 m/s.

Finally, at the top level, all uncertainties in the wind speed interpolation and extrapolation problem were considered, including measurement uncertainty and correlation uncertainty. The magnitude of granularity equal to combined uncertainty, which can be characterized by standard deviation and be calculated from a set of simultaneous reference wind speed data at a specific wind speed of target data (see Section 3.4), was around 1–3 m/s.

4.3. Granulation

This step drew together a clump of objects (points) by indistinguishability, similarity, proximity, or functionality to induce granules. The fuzzy c-means clustering algorithm (FCM), which is an efficient and validated clustering algorithm [30] was adopted in this paper.

The steps of the FCM algorithms are presented below [31]:

- (a) Fix $C, m, A, \|k\|_A$. Choose an initial matrix $U^{(0)} \in M_{fc}$. Then at step $k, k = 0, 1, \dots, LMAX$;
- (b) Compute means $\hat{v}^{(k)}, i = 1, 2, \dots, C$ with equation $\hat{v}_i = \frac{\sum_{k=1}^N (\hat{u}_{ik})^m y_k}{\sum_{k=1}^N (\hat{u}_{ik})^m}; 1 \leq i \leq c$;

(c) Compute an updated membership matrix $\hat{U}^{(k+1)} = [\hat{u}_{ik}^{k+1}]$ with equation

$$\hat{u}_{ik} = \left(\sum_{j=1}^c \left(\frac{\hat{d}_{ik}}{\hat{d}_{jk}} \right)^{2/(m-1)} \right)^{-1}; \quad 1 \leq k \leq N; \quad 1 \leq i \leq c;$$

(d) Compare \hat{U}^{k+1} to \hat{U}^k in any convenient matrix norm. If $\|\hat{U}^{k+1} - \hat{U}^k\| < \partial$, stop. Otherwise set $\hat{U}^k = \hat{U}^{k+1}$ and return to Step (b).

Normally, the granulations of the target and reference data are independent. At each granule level, the cluster count was identified based on the analysis results of the granular hierarchy, and the initial center of the clusters was generated randomly. Next, the FCM algorithm was used to partition the data into clusters. As the results of the FCM algorithm are sensitive to the initial center of clusters [32,33], the genetic algorithm was adopted to optimize the initial center of clusters and was avoided a partial optimum solution (Figure 6).

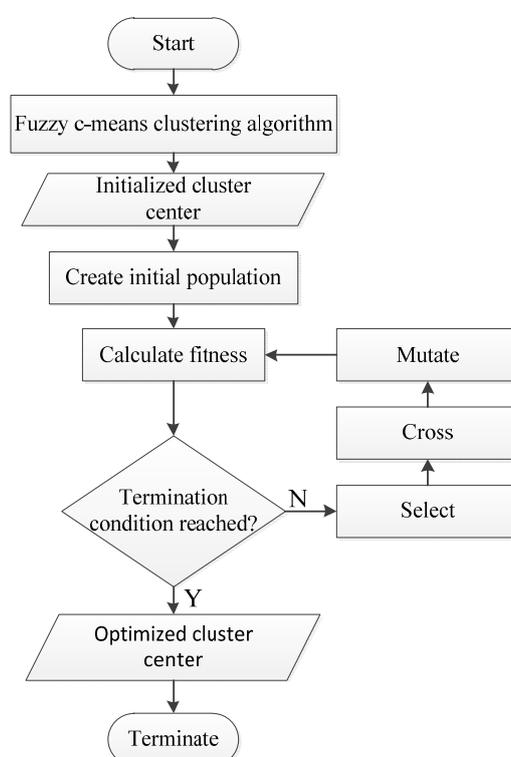


Figure 6. Flow chart of genetic algorithm procedure.

4.4. Granules Representation

In this study, the cloud model theory was adopted to represent granules.

The cloud model transfers the procedure of uncertainty between the qualitative concept and quantitative data representation by using natural language [34], and mainly reflects the fuzziness and randomness of the concept within the affairs and human knowledge in the objective world [35].

Suppose that T is the language value of domain u , and mapping $C_T(x): u \rightarrow [0,1], \forall x \in u, x \rightarrow C_T(x)$, then the distribution of $C_T(x)$ in u is called the membership cloud of T , or cloud for short. If the distribution of $C_T(x)$ is normal, it is named the normal cloud model, which is a random number set that obeys the normal distributive rule and has a stable tendency. It is determined by expectation Ex , entropy En , and super entropy He , and reflect the quantitative characteristics of concept $C_T(x)$ [36], Ex determines the center of

the cloud, En determines the range of the cloud and, according to “3 En ” rule[34], about 99.74% of the total cloud drops distribute between $[Ex - 3En, Ex + 3En]$; He determines the cloud drops’ dispersive degree, which means the larger the He is, the more dispersively the cloud drops locate. The sketch map of three digital characteristics of a normal cloud model is shown in Figure 7.

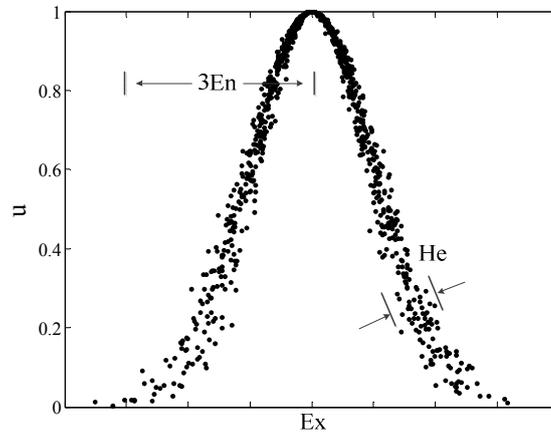


Figure 7. Three digital characteristics of a normal cloud model. u is membership factor.

- backward cloud generator

If data set and certain u_0 are given, then three digital characteristics (Ex, En, He) can be generated by a backward generator [36], of which pseudo-codes is presented below and flow chart is shown in Figure 8.

$$INPUT : \{(x_1, u_0), \dots, (x_n, u_0)\}, u_0$$

$$OUTPUT : \{Ex, En, He\}$$

$$Ex = mean(x)$$

$$En = std(x)$$

$$He = std(sqrt(-(x - Ex).^2 / (2 * \log(u_0))));$$

where $randn(En, He)$ produces a random number with normal distribution whose expectation is En and standard deviation is He .

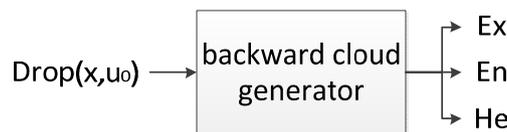


Figure 8. Flow chart of backward cloud generator.

- forward cloud generator

If the three digital characteristics (Ex, En, He) and a certain u_0 are given, then a drop of cloud $drop(x_i, u_0)$ can be generated by a forward generator [36], of which pseudo-codes is presented below and flow chart is shown in Figure 9.

```

INPUT : {Ex, En, He}, n, u0
OUTPUT : {(x1, u0), ..., (xn, u0)}
FOR i = 1 : n
    En' = randn(En, He)
    xi = Ex ± En' · √(-2 · ln(u0))
    drop(xi, u0)

```

where $\text{randn}(En, He)$ produces a random number with normal distribution whose expectation is En and standard deviation is He .

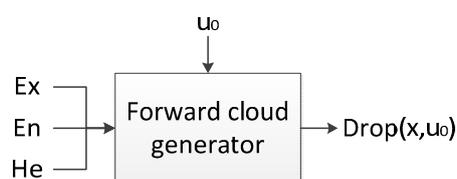


Figure 9. Flow chart of forward cloud generator.

For each granule, every wind speed data record was a drop, and the cloud model digital characteristics (Ex, En, He) can be obtained from all the drops by using the back cloud generator, thereby the three digital characteristics were representative of each granule.

4.5. Granular Computing

Randomness of wind speed was considered in this step. Granular computing is a method of finding the correlation between two worlds of granules, which is complex and non-linear data fitting. There are two alternative methods, artificial neural networks method and support vector regression method, are very suitable for this kind of issue [10–15], both of them were adopted as alternative method in this paper and compared in the Section 5. A brief introduction of the two methods was provided below.

- Artificial neural networks

The topography of the ANN is shown in Figure 10. Very fast training neural-computation techniques [14] were chosen for ANN training. The ANN was trained by the data of two worlds of granules for each level, and three trained ANNs, corresponding to three levels, were the results of this step.

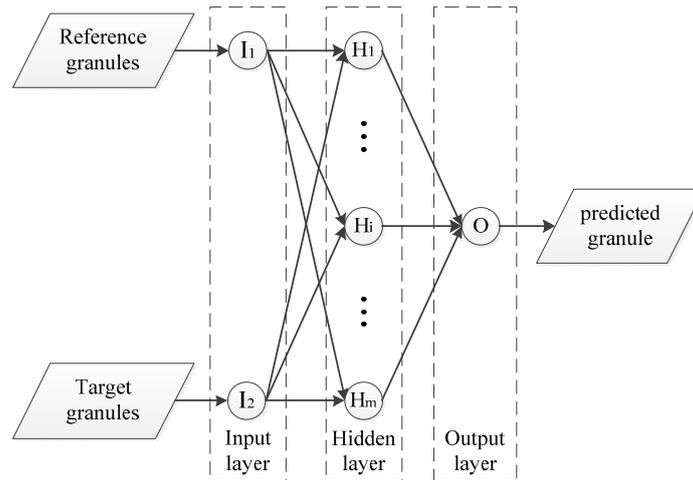


Figure 10. Topological structure of predicted Artificial Neural Network (ANN) in single granular level.

- Support vector regression

The basic idea of support vector machines (SVM) for regression is to map the data x into a high dimensional feature space via nonlinear mapping and to perform a linear regression in this feature space; more detailed description of SVM for regression can be found in Ref. [37]. Radial basis function(RBF) kernels were adopted and five-fold cross validation method was used to determine the C and g .

It is important to note that the correlation constructed was the relationship between two worlds of granules, and the wind speed was in granular format, not in quantitative format.

4.6. Interpolation and Extrapolation

For each level, the contemporary reference data were input into trained ANNs to calculate the target data, which meant that three interpolating or extrapolating data were provided by ANNs.

Another ANN was trained using the contemporary data of reference and target granules for synthesizing the three results provided by the ANNs, and the topography of the synthesized ANN is shown in Figure 11. The Levenberg-Marquardt (LM) algorithm was chosen for this training.

The final results were provided by the synthesized ANN, and the wind data acquired in this step were in granular format. To transfer the granular format wind data to quantitative data, the Y condition cloud generator were used, and finally, the quantitative results were obtained.

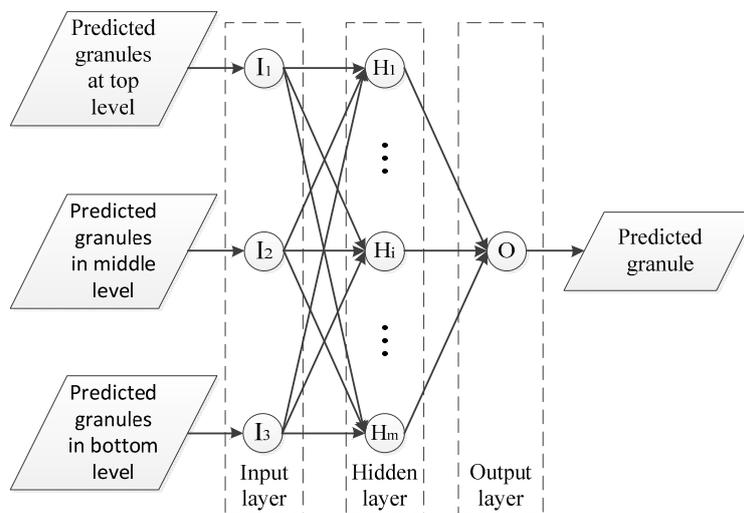


Figure 11. Topological structure of the synthesized ANN.

5. Case Study

5.1. Case Study Description

To implement and validate the granular computing method, a missing wind data interpolation case was utilized. Two meteorological sites located in Hubei province, Central Southern China were chosen; the longitude and latitude, start and end date, equipped instrument, elevation, and local terrain of two sites are shown in Table 5. The two sites are both located in ridgeline with no obstacles around, and the distance between them is 40 km, evaluation map of region of interest is shown in Figure 12.

Data were collected as 1-h values from 1 January 2012 to 31 December 2012, resulting in a total of 8760 records. The correlation coefficient was 0.733, scattered plot of two sites is shown in Figure 13.

The chosen wind farm belongs to a subtropical monsoon climate, located in a complex terrain, and equipped with widely used anemometer, so, gaps in wind speed records are easily emerged due to icing, lightning stroke, equipment failure, and so on. It is a common practical case which is frequently need wind speed interpolation and extrapolation, so, it is a fine representation and the results obtained in this case study can be extended to alternative cases.

In particular, the ten-fold cross validation method was adopted. All records were randomly split into ten mutually exclusive subsets ($D_1, D_2 \dots D_{10}$) of equal size, the proposed method was trained and tested ten times, each time $t \in [1, 10]$, it was trained on D/D_t and tested on D_t , the evaluation metrics of each time was averaged as final evaluation metrics.

Table 5. Description of the two meteorological sites.

Parameters	Reference Site	Target Site
Location	31°32.113' N, 114°09.749' E	31°28.248' N, 114°10.847' E
Elevation (m)	620	430
Local terrain	mountain	hill
Start date	2012/1/1	2012/1/1
End date	2012/12/31	2012/12/31
Anemometer type	NRG 40C	NRG 40C
Mounting height (m)	70/60/50/30/10	70/60/50/30/10

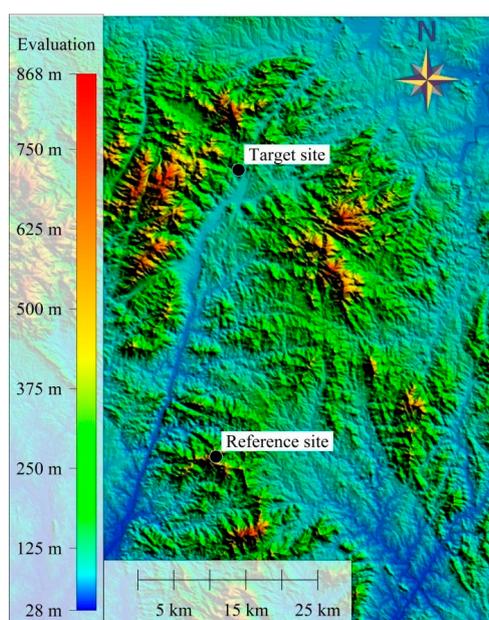


Figure 12. Evaluation map of region of interest. Solid dot show the target and reference site.

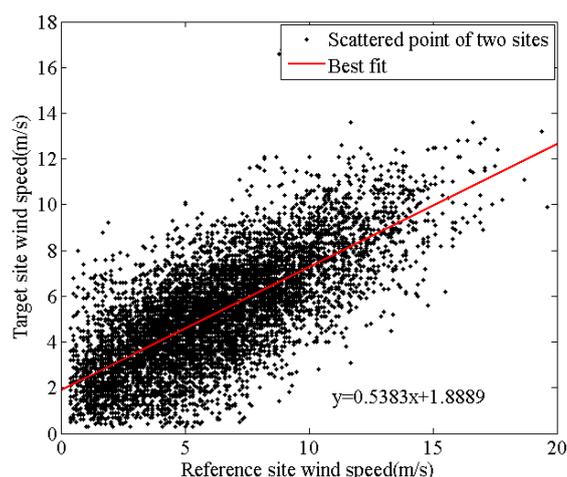


Figure 13. Scattered plot of wind speed of target site versus reference site with best fit line from 1 January 2012 to 31 December 2012.

In addition, a commercial pitch-regulated 2000 kW rated power wind turbine (GW-2000) was used for the estimations in this study (Figure 14); the power curve was published by the manufacturers in catalogues in table form. Since the power curve data are a discrete pair of values of wind speed and power output, the cubic spline interpolation method was used to calculate the output power from wind speed.

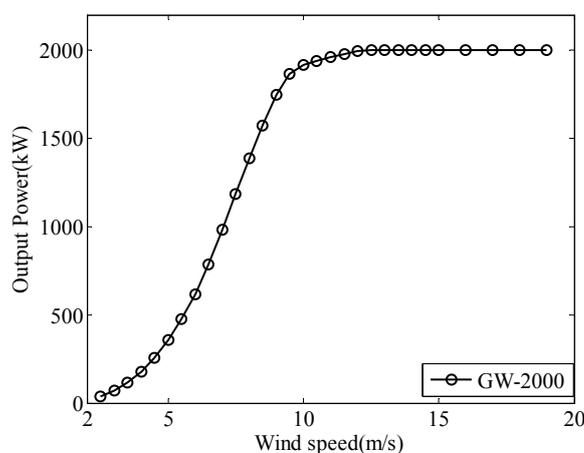


Figure 14. Output power curve of the wind turbine used in the case study.

It is also important to note that in order to increase accuracy, site analysis was carried out in modal terms with respect to wind direction by ideally subdividing the wind data into 16 wind sectors.

5.2. Evaluation Metrics Used

MRE (mean relative error), RMSE (root mean square error), r (correlation coefficient), and MREEP (mean relative error of energy production) were selected as the evaluation metrics using the following expressions.

$$MRE = \frac{\frac{1}{n} \sum_{i=1}^n (V_i - V_{mi})}{\frac{1}{n} \sum_{i=1}^n (V_{mi})} \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (V_i - V_{mi})^2} \tag{9}$$

$$r = \frac{\sum_{i=1}^n [(V_i - \bar{V}) \times (V_{mi} - \bar{V}_m)]}{\sqrt{\sum_{i=1}^n (V_i - \bar{V})^2 \times \sum_{i=1}^n (V_{mi} - \bar{V}_m)^2}} \tag{10}$$

$$MREEP = \frac{\frac{1}{n} \sum_{i=1}^n (P_i - P_{mi})}{\frac{1}{n} \sum_{i=1}^n P_{mi}} \tag{11}$$

where V_i is the counts of data; V is the interpolated wind speed; V_m is the true wind speed; P is the interpolated energy production, $kW \cdot h$; and P_m is the true energy production, $kW \cdot h$.

5.3. Determining the Granular Hierarchy

In this study, the type of anemometer was identified as NRG #40C. As per the NRG #40C white paper [38], the minimum interval was 0.1 m/s; the classification number was 7.7; the range of collected wind speed was 0–17 m/s; and distance to mast center divided by mast leg distance was 0.25. Based on the data, the average measurement uncertainty calculated following the method provided in Section 3.4 was 1.54 m/s, and the measurement and combined uncertainty versus wind speed are shown in Figure 15.

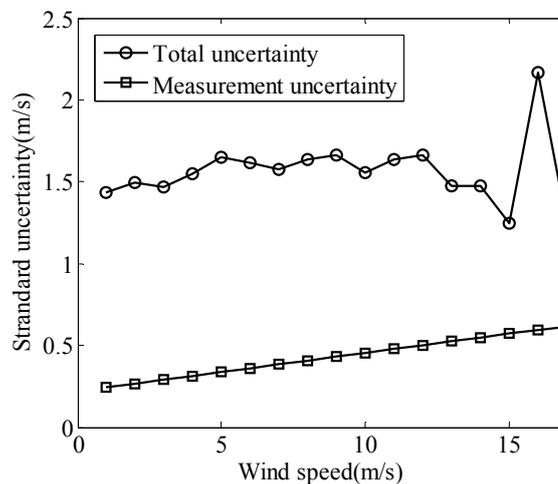


Figure 15. Combined and measurement uncertainty in this case.

The bottom level was constructed by basic granules, which were the minimum interval of wind speed magnitude recording by anemometer. In this case, the range of collected wind speed was 0–17 m/s, and the magnitude of granularity was 0.1 m/s, therefore the number of clusters was equal to 17 m/s divided by 0.1 m/s, it was 170.

In the middle level, only uncertainties of measurement were considered, and the average measurement standard uncertainty, 0.5 m/s, was adopted for the magnitude of granularity, and the range of collected wind speed was 0–17 m/s, so the number of clusters was equal to 17 m/s divided by 0.5 m/s, it was 34.

At the top level, all uncertainties were considered, and the average combined uncertainty, 1.54 m/s, was adopted for the magnitude of granularity, and the range of collected wind speed was 0–17 m/s, so the number of clusters was equal to 17 m/s divided by 1.54 m/s, it was 11.

5.4. Results Analysis

Missing data in this case study were interpolated based on the granular computing method, and a comparison of the results with traditional methods such as linear regression method, variance ratio method, ANN method, and SVR method was proceeded, results are shown in Figure 16 and Table 6.

It can be observed from the results that GrC with SVR method provided a slight better estimation than GrC with ANN method. To be specific, comparing to GrC with ANN method, the correlation coefficient of the GrC with SVR increase 0.51%, and the *MRE*, *MREEP*, and *RMSE* decrease 6.25%, 0.28% and 2.80%, respectively.

It can be likewise observed from the results that all evaluation indexes of GrC with SVR method provided in this paper are superior to the other methods. The linear regression method, widely used in practical engineering, was chosen as the benchmark, and the correlation coefficient of the GrC with SVR method increase 8.48%, and the *MRE*, *MREEP*, and *RMSE* decrease 61.04%, 60.98% and 15.76%, respectively. It can be concluded from Figure 16 that wind speeds predicted by the granular computing method are very close to the true value in tendency, but prediction error increases in circumstances with small or big wind speed at the target site.

The results of the case study show that the accuracy of the GrC with SVR method meet the needs of practical engineering, and has an advantage over other methods. Therefore, the granular computing method is a fine alternative method in the wind speed temporal interpolation and extrapolation problem.

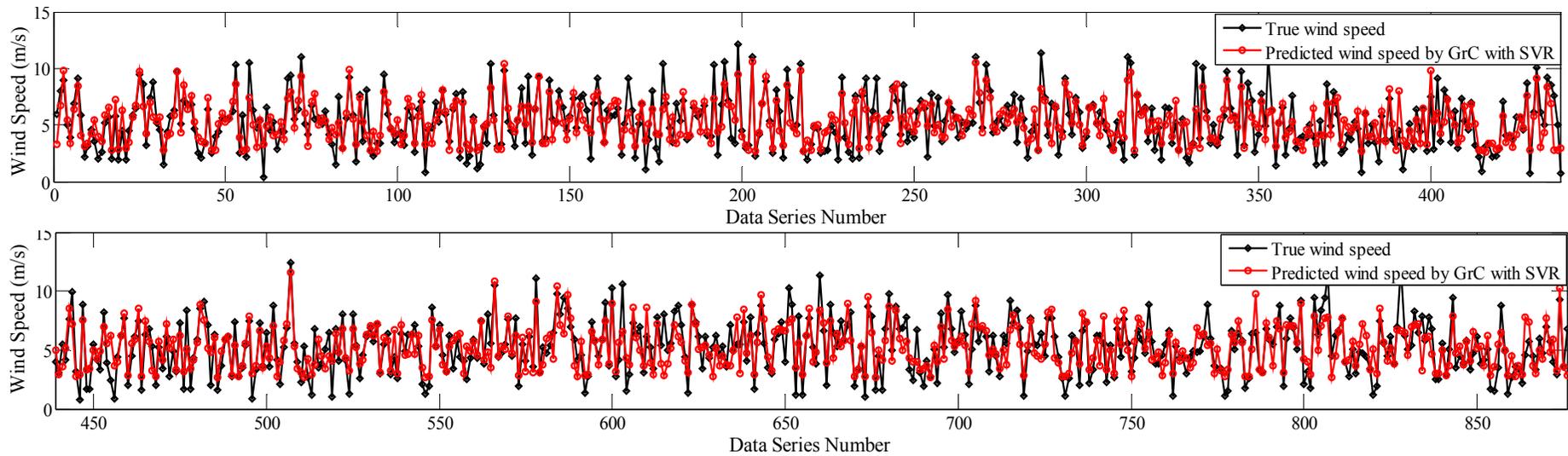


Figure 16. Comparisons between true wind speed and wind speed predicted by the granular computing method. The comparisons are proceeded in one class of ten-fold cross validation processes.

Table 6. Evaluation metrics comparisons between granular method and traditional methods.

Evaluation Index	Linear Regression Method	Variance Ratio Method	ANN	SVR	Granular Computing Method with ANN				Granular Computing Method with SVR			
					Top	Middle	Bottom	Synthesized	Top	Middle	Bottom	Synthesized
					Level	Level	Level		Level	Level	Level	
MRE	-0.77%	-2.01%	-1.42%	-1.37%	-1.75%	-0.85%	-1.42%	-0.32%	-1.19%	-0.79%	-1.37%	-0.30%
r	0.731	0.732	0.779	0.782	0.752	0.745	0.779	0.789	0.768	0.757	0.782	0.793
MREEP	-8.97%	3.46%	-5.95%	-4.93%	-8.93%	-4.61%	-5.95%	-3.51%	-6.87%	-3.18%	-4.93%	-3.50%
RMSE (m/s)	1.65	1.76	1.5	1.47	1.59	1.61	1.5	1.43	1.63	1.53	1.47	1.39

6. Conclusions

In this paper, wind speed mixed uncertainty analysis is implemented, and recommended values and estimation tools for wind speed measurement uncertainty and combined uncertainty are provided, as well as a determination of the uncertainty hierarchy. For the purpose of considering mixed uncertainty of randomness and fuzziness, a concrete granular computing method is constructed and a new MCP method of wind speed temporal interpolation and extrapolation based on granular computing theory is proposed; and a wind speed data temporal interpolation case is provided to implement and validate the new method. Hence, the following conclusions are reached:

1. By using the MCP method proposed in this paper, mixed uncertainty of wind speed had already been considered, thus, a better estimation of the wind speed is provided compared to other methods selected for comparison. In the case study, almost evaluation metrics of interpolation with the proposed method were superior to other methods used in comparison. In comparison to the linear method, the correlation coefficient of the proposed method increased 8.48%, and the *MRE*, *MREEP*, and *RMSE* decreased 61.04%, 60.98% and 15.76%, respectively. The proposed method improved wind resource assessment accuracy and reduced the risks of wind farm construction.
2. Suitability of using granular computing methods for the issue of wind speed data interpolation and extrapolation is proved. By using GrC method, wind speed mixed uncertainty can be taken into account; accurate results and low cost solutions can be derived.
3. Mixed uncertainty of wind speed can be divided into three levels, and recommended values of granularity are minimum interval of records, 0.3–0.8 m/s, and 1–3 m/s, respectively. Also, estimation tools for wind speed measurement uncertainty and combined uncertainty are provided.

Acknowledgments: This work was supported in part by the National Natural Science Foundation of China under Grant 51379159 and Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20130141130001.

Author Contributions: Xiao Liu conceived and designed the methodology; Xiao Liu and Jin Zou performed the case studies and analysis; Xiao Liu and Xu Lai wrote the paper.

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

References

1. Burton, T.; Jenkins, N.; Sharpe, D.; Bossanyi, E. *Wind Energy Handbook*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2011.
2. European Wind Energy Association. *Wind Energy—the Facts: A Guide to the Technology, Economics and Future of Wind Power*; Earthscan: London, UK, 2009.
3. Derrick, A. Development of the measure-correlate-predict strategy for site assessment. In *European Community Wind Energy Conference, Proceedings of the 1993 International Conference, Lubeck-Travemunde, Germany, 8–12 March 1993*; Garrad, A.D., Palz, W., Screller, S., Eds.; Stephens, H.S. and Associates: Bedford, UK, 1993; pp. 681–685.
4. Ramsdell, J.V.; Houston, S.; Wegley, H.L. Measurement strategies for estimating long-term average wind speeds. *Sol. Energy* **1980**, *25*, 495–503.
5. Woods, J.C.; Watson, S.J. A new matrix method of predicting long-term wind roses with MCP. *J. Wind Eng. Ind. Aerodyn.* **1997**, *66*, 85–94.
6. Lackner, M.A.; Rogers, A.L.; Manwell, J.F. Uncertainty analysis in MCP-based wind resource assessment and energy production estimation. *J. Sol. Energy Eng.* **2008**, *130*, doi:10.1115/1.2931499.
7. Thøgersen, M.L.; Nielsen, P.; Sørensen, T.; Svenningsen, L.U. An Introduction to the MCP Facilities in WindPRO. EMD International A/S; 2010. Available online: <http://help.emd.dk/knowledgebase/content/ReferenceManual/MCP.pdf> (accessed on 17 June 2017).
8. Khadem, S.K.; Hussain, M. A pre-feasibility study of wind resources in Kutubdia Island, Bangladesh. *Renew. Energy* **2006**, *31*, 2329–2341.
9. Abbes, M.; Belhadj, J. Wind resource estimation and wind park design in El-Kef region, Tunisia. *Energy* **2012**, *40*, 348–357.

10. Mohandes, M.A.; Halawani, T.O.; Rehman, S.; Hussain, A.A. Support vector machines for wind speed prediction. *Renew. Energy* **2004**, *29*, 939–947.
11. Ji, G.-R.; Han, P.; Zhai, Y.-J. Wind speed forecasting based on support vector machine with forecasting error estimation. In Proceedings of the 6th International Conference on Machine Learning and Cybernetics, Hong Kong, China, 19–22 August 2007; pp. 2735–2739.
12. Velázquez, S.; Carta, J.A.; Matías, J.M. Comparison between ANNs and linear MCP algorithms in the long-term estimation of the cost per kWh produced by a wind turbine at a candidate site: A case study in the Canary Islands. *Appl. Energy* **2011**, *88*, 3869–3881.
13. Addison, J.F.; Hunter, A.; Bass, J.; Rebbeck, M. A neural network version of the measure–correlate–predict algorithm for estimating wind energy yield. In Proceedings of the 13th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management, Houston, TX, USA, 3–8 December 2000; pp. 917–922.
14. Saavedra-Moreno, B.; Salcedo-Sanz, S.; Carro-Calvo, L.; Gascon-Moreno, J.; Jimenez-Fernandez, S.; Prieto, L. Very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms. *J. Wind Eng. Ind. Aerodyn.* **2013**, *116*, 49–60.
15. Liu, X.; Lai, X.; Zheng, F. Analysis of the interpolation method for wind speed data based on Reanalysis data. *J. Huazhong Univ. Sci. Technol.* **2017**, *45*, 78–83.
16. Carta, J.A.; Velázquez, S. A new probabilistic method to estimate the long-term wind speed characteristics at a potential wind energy conversion site. *Energy* **2011**, *36*, 2671–2685.
17. García-Rojo, R. Algorithm for the estimation of the long-term wind climate at a meteorological mast using a joint probabilistic approach. *Wind Eng.* **2004**, *28*, 213–223.
18. Angelis-Dimakis, A.; Biberacher, M.; Dominguez, J.; Fiorese, G.; Gadocha, S.; Gnansounou, E.; Guariso, G.; Kartalidis, A.; Panichelli, L.; Pinedo, I.; et al. Methods and tools to evaluate the availability of renewable energy sources. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1182–1200.
19. Bowen, A.J.; Mortensen, N.G. Exploring the limits of WAsP the wind atlas analysis and application program. In Proceedings of the 1996 European Wind Energy Conference and Exhibition, Goteborg, Sweden, 20–24 May 1996; pp. 584–587.
20. Landberg, L.; Mortensen, N.G. A comparison of physical and statistical methods for estimating the wind resource at a site. In Proceedings of the 15th British Wind Energy Association Conference, York, UK, 6–8 October 1993; Pitcher, K.F., Ed.; Mechanical Engineering Publications Ltd.: London, UK, 1993; pp. 119–125.
21. Yao, Y. Perspectives of granular computing. In Proceedings of the 2005 IEEE International Conference on Granular Computing, Beijing, China, 25–27 July 2005.
22. Zadeh, L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst.* **1997**, *90*, 111–127.
23. Bargiela, A.; Pedrycz, W. *Granular Computing: An Introduction*; Springer Science & Business Media: New York, NY, USA, 2003.
24. Lin, T.Y.; Yao, Y.Y.; Zadeh, L.A. *Data Mining, Rough Sets and Granular Computing*; Springer: Berlin, Germany, 2002.
25. Zadeh, L.A. Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. In *Soft Computing—A Fusion of Foundations, Methodologies and Applications*; Springer-Verlag: Berlin, Germany, 1998; Volume 2, pp. 23–25.
26. International Organization for Standardization. *Uncertainty of Measurement—Part 3: Guide to the Expression of Uncertainty in Measurement*; ISO: Geneva, Switzerland, 2008.
27. International Electrotechnical Commission. *Wind Turbines—Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines*; IEC: Geneva, Switzerland, 2012.
28. Risø National Laboratory. ACCUWIND—Classification of Five Cup Anemometers According to IEC61400-12-1. Available online: http://orbit.dtu.dk/fedora/objects/orbit:88359/datastreams/file_7703253/content (accessed on 17 June 2017).
29. Coquilla, R.V.; Obermeier, J. Calibration uncertainty comparisons between various anemometers. In Proceedings of the American Wind Energy Association Wind Power Conference, Houston, TX, USA, 10 July 2008.
30. Pal, N.R.; Bezdek, J.C. On cluster validity for the fuzzy c-means model. *IEEE Trans. Fuzzy Syst.* **1995**, *3*, 370–379.

31. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203.
32. Scheunders, P. A genetic c-means clustering algorithm applied to color image quantization. *Pattern Recognit.* **1997**, *30*, 859–866.
33. Alata, M.; Molhim, M.; Ramini, A. Optimizing of Fuzzy C-Means Clustering Algorithm Using GA. *World Acad. Sci. Eng. Technol.* **2008**, *2*, 670–675.
34. Li, D.Y.; Meng, H.J.; Shi, X.M. Membership clouds and membership cloud generators. *J. Comput. Res. Dev.* **1995**, *32*, 15–20.
35. Di, K.; Li, D.R.; Li, D.Y. Cloud theory and its applications in spatial data mining and knowledge discovery. *J. Image Graph.* **1999**, *4*, 930–935.
36. Li, D.Y.; Du, Y. *Artificial Intelligence with Uncertainty*; CRC Press: Boca Raton, FL, USA, 2017.
37. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Dordrecht, The Netherlands, 2013.
38. Nrgsystems. #40C Anemometer Uncertainty. Available online: <https://www.nrgsystems.com/assets/resources/40C-Anemometer-Uncertainty-AppNote.pdf> (accessed on 17 June 2017).



© 2017 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 (<http://creativecommons.org/licenses/by/4.0/>).