Construction of wind turbine performance model based on SCADA data

Sha Wu¹, Sheng Lu¹, Xin Xie², Shaoping Deng¹, Zixiao Jiang², Jian Wang¹

¹ Powerchina Hubei Electric Engineering Co., Ltd., No.1 Xinqiaosi Road ,jinyinhu Street, Dongxihu District, Wuhan, 430040, China;
² Meteodyn Beijing, 48 Dongzhimenwai Avenue, Beijing 100027, China
E-mail:wushsj@powerchina-hb.com

Abstract. Compared with traditional wind turbine performance evaluation methods such as operating power curve fitting, the wind turbine performance model construction method using SCADA data on machine learning has gradually attracted attention in the industry in recent years. In this technical route, the choice of input SCADA data type will directly affect the accuracy of the wind turbine performance model construction. However, in the past, selecting key SCADA data types in the performance evaluation of wind turbines were mostly subjective judgments or correlation analyses. This paper proposed a key SCADA data type selection method based on mutual information calculation and a wind turbine performance model construction method based on deep neural network training. The selected key SCADA data types based on mutual information were applied to the deep neural network to construct the wind turbine performance model, and the actual SCADA operation data of the wind farm was applied to test the model. The results show that the model construction method's accuracy and generalizability using the above technical route can satisfy the industry demand in the same wind farm with the same wind turbine type.

1. Introduction
With the wind power industry's development, many wind turbines have operated for several years, and wind farms highly recommend efficient O&M and quality improvement. At the same time, wind power grid parity continues to advance. In this context, wind power developers have refined their demand for wind farms post-assessment. Wind turbine performance assessment is an essential part of wind farm post-assessment, but it highly relies on the SCADA system's data collection. The wind turbine SCADA system generally has more than one sensor for each component [1]. It collects different data types, including but not limited to wind speed, wind direction, wind turbine power, blade pitch angle, rotor speed, and temperature. The benefit of the SCADA system and its data collection mode is that researchers can obtain a large amount of data for analysis, so selecting which data types to analyze becomes a critical problem in wind turbine performance assessment [2].

SCADA data types selection is essentially a feature extraction problem, but most researchers have not recognized this in post-assessment. Therefore, recently SCADA data selection still often uses the empirical judgment method. However, it is not easy to obtain the desired result if using this method for wind turbine performance model construction [3]. With further research, the wind power industry gradually uses the method based on feature extraction [4]. For example, the method based on singular value decomposition, which is used widely in image detection and recognition [5]. The method based on the Max margin criterion [6] and the gradient optimization [7]. Furthermore, after SCADA data
selection, the wind turbine performance model construction method is also changed from the fitting of multi-variable function to deep neural networks and machine learning [8].

In this article, a method based on mutual information theory for SCADA data selection is proposed and practiced. In the first part of the article, this method's technical route based on mutual information will be introduced step by step, and we will apply it to a set of real wind farm SCADA data. In the second part of the article, a wind turbine performance model will be constructed using the data types selected. The verification of the model's accuracy and generalizability is also completed in this part.

2. SCADA data types extraction based on mutual information

2.1 Mutual information

Mutual information is proposed in Shannon's information theory, and it can qualify the relevance level between different random variables [9]. Here, the relevance not just means the linear correlation, which is used a lot in statistics. It is a concept much more generalized. In information theory, the mutual information is often related to entropy, which means the amount of information carried by random variables. Taking random variables \( x \) and \( y \) as an example, the relation between their entropy and mutual information is shown in Figure 1.

\[ I(X, Y) = \int \int p(x, y) \log \left( \frac{p(x, y)}{p_x(x)p_y(y)} \right) dx dy \] (1)

In this formula:

- \( p(x, y) \): the joint probability density function of variable \( x \) and variable \( y \)
- \( p_x(x) \): the marginal probability density function of variable \( x \)
\( p_y(y) \): the marginal probability density function of variable \( y \)
\( X \): the domain of variable \( x \)
\( Y \): the domain of variable \( y \)

Furthermore, in the real case, the wind turbine SCADA system cannot output a continuous signal. Each sensor of the system can only record the discrete signal. The definition of mutual information to discrete variables is as in equation (2):

\[
I(x, y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p_x(x)p_y(y)} \right)
\]

According to the mathematic definition of mutual information, if the marginal probability density function of SCADA data types and the joint probability density function between them and the turbine output power can be obtained, the mutual information can then be calculated. However, in the real application, the probability density function is hard to get in the SCADA system. So, if we want to calculate the mutual information for SCADA data types, a method that does not rely on the probability density function should be found.

### 2.3 Coupula function and Sklar theorem

In Statistic and probability theory, the Coupula function can connect the marginal cumulative distribution function of a set of variables and their joint cumulative distribution function via the Sklar theorem [10].

According to Sklar theorem, if \( P \) is the joint cumulative distribution function of a multidimensional random variable \( x = x_1, \ldots, x_N \). And \( x \) has marginal cumulative distribution function \( \{P_i, i = 1 \ldots N\} \). Such a relation between \( P \) and \( \{P_i\} \) can be established via the Coupula function.

\[
P(x) = C(P_1(x_1), \ldots, P_N(x_N))
\]

In a bi-dimensional variable \( (x, y) \) case, the formula above changes to equation (4).

\[
P(x,y) = C(P_x(x), P_y(y))
\]

According to the probability density function and cumulative distribution function properties, the differential of equation (4) gives the equation (5).

\[
p(x,y) = \frac{\partial^2 p(x,y)}{\partial x \partial y} = \frac{\partial^2 C(P_x(x),P_y(y))}{\partial P_x(x)\partial P_y(y)} \cdot \frac{\partial P_x(x)}{\partial x} \cdot \frac{\partial P_y(y)}{\partial y}
\]

Finally, bring the equation (5) into mutual information definition equation (1).

\[
I(x,y) = \int_x \int_y c(P_x(x), P_y(y)) \log \left( c(P_x(x), P_y(y)) \right) \frac{\partial P_x(x)}{\partial x} \frac{\partial P_y(y)}{\partial y}
\]

\[
c(P_x(x), P_y(y))
\]

is the Coupula density function, and since \( P_x(x) \) and \( P_y(y) \) the marginal cumulative distribution function of \( x \) and \( y \), whose value is distributed in \([0, 1]\). So, by proposing \( P_x(x) = a \) and \( P_y(y) = b \), finally, the mutual information of \( x \) and \( y \) is represented as the equation (7).

\[
I(x,y) = \int_0^1 \int_0^1 c(a, b) \log(c(a, b)) dadb
\]

As mentioned before, the mathematical definition of mutual information is hard to apply because the probability density function and joint probability density function are hard to get from the SCADA system. Coupula function and Sklar theorem make mutual information depending only on Coupula density function and variables' marginal cumulative distribution function. These two functions can be obtained from SCADA data using nonparametric estimation.

### 2.4 Calculation of Coupula based on the empirical distribution

Coupula density function is essentially the differential of Coupula function. In a SCADA sampling system, the Coupula between two data types can be estimated by empirical distribution.

For a set of sampling of variables \( x \) and \( y \) \( \{x_i, y_i\}^N \), the Coupula function of \( x \) and \( y \) using empirical distribution is represented in equation (8) [11].

\[
C(P_x(x) = a, P_y(y) = b) = \frac{\sum_{i=1}^N 1(P_x(x_i) < a, P_y(y_i) < b)}{N}
\]
The marginal cumulative distribution function \( P_x(x_i) \) and \( P_y(y_i) \) can also represent in empirical distribution format.

\[
P_x(x_i) = \frac{1}{N} \sum_{x \in \{x_i\}} 1 \quad (x < x_i) \\
P_y(y_i) = \frac{1}{N} \sum_{y \in \{y_i\}} 1 \quad (y < y_i)
\]

After obtaining the Coupula function, the Coupula density function can get via the kernel density estimation method, which is usually used to estimate an unknown distribution [12].

\[
c(P_x(x) = a, P_y(y) = b) = \frac{1}{nh^2} \sum_{i=1}^{N} K \left( \frac{|P_x(x_i) - a|, |P_y(y_i) - b|}{h} \right)
\]

The \( K \) function is the kernel function used. In this study, a Beta function is chosen.

Once the Coupula density function between each wind turbine SCADA data types and turbine output power is estimated, their mutual information is estimated approximately.

\[
I(x, y) = \sum_{a,b} c(P_x(x) = a, P_y(y) = b) \log c(P_x(x) = a, P_y(y) = b)
\]

2.5 Application on real SCADA data

The mutual information estimation method introduced above has been applied to real SCADA data recorded over seven months from a wind farm in Hubei, China. This wind farm has 50 wind turbines with 2000kW rated power, and the SCADA system collects 90 different data types, which are recorded in 10 minutes time steps. Due to the SCADA system's availability, some common data types have not been recorded, such as pitch angle and tip speed ratio. A detailed description of 90 data types is in Table 1.

| Data type index | Data type name                        | Data type index | Data type name                        | Data type index | Data type name                        |
|-----------------|---------------------------------------|-----------------|---------------------------------------|-----------------|---------------------------------------|
| 01               | Turbine ID                            | 02              | Time                                  | 03              | Ave wind speed                        |
| 04               | Ave output power                      | 05              | Ave rotor speed                       | 06              | Max rotor speed                       |
| 07               | Min rotor speed                       | 08              | Max total power-on time               | 09              | Min total power-on time               |
| 10               | Max system uptime                     | 11              | Min system uptime                     | 12              | Max failure time                      |
| 13               | Min failure time                      | 14              | Max normal environment time           | 15              | Min normal environment time           |
| 16               | Max production                        | 17              | Min production                        | 18              | Max power consumption                 |
| 19               | Min power consumption                 | 20              | Max maintenance time                  | 21              | Min maintenance time                  |
| 22               | Max wind speed                        | 23              | Min wind speed                        | 24              | Max output power                      |
| 25               | Min output power                      | 26              | Max nacelle temperature               | 27              | Max environment temperature           |
| 28               | Min reactive power                    | 29              | Max reactive power                    | 30              | Ave reactive power                    |
| 31               | Ave environment temperature           | 32              | Min environment temperature           | 33              | Ave yaw angle                         |
| 34               | Ave grid voltage of the A phase       | 35              | Ave grid voltage of the B phase       | 36              | Ave grid voltage of the C phase       |
| 37               | Ave grid current of the A phase       | 38              | Ave grid current of the B phase       | 39              | Ave grid current of the C phase       |
| 40               | Max temperature of the control cabinet| 41              | Max temperature of the pitch cabinet  | 42              | Ave temperature of the control cabinet|
| 43               | Ave temperature of the pitch cabinet  | 44              | Ave wind direction                    | 45              | Max temperature of generator (sensor 1)|
| 46               | Max temperature of generator (sensor 2)| 47              | Max temperature of generator (sensor 3)| 48              | Max temperature of generator (sensor 4)|
| Data type index | Data type name | Data type index | Data type name | Data type index | Data type name |
|-----------------|----------------|-----------------|----------------|-----------------|----------------|
| 49              | Max temperature of generator (sensor 5) | 50              | Max temperature of generator (sensor 6) | 51              | Max temperature of generator front bearing (sensor 1) |
| 52              | Max temperature of generator front bearing (sensor 2) | 53              | Max temperature of generator rear bearing (sensor 1) | 54              | Max temperature of generator rear bearing (sensor 2) |
| 55              | Max temperature of generator magnet (sensor 1) | 56              | Max temperature of generator magnet (sensor 2) | 57              | Max temperature of generator magnet (sensor 3) |
| 58              | Max temperature of generator magnet (sensor 4) | 59              | Max temperature of generator cooling outlet (sensor 1) | 60              | Max temperature of generator cooling outlet (sensor 2) |
| 61              | Max peak acceleration | 62              | Max temperature of generator (sensor 7) | 63              | Max temperature of generator (sensor 8) |
| 64              | Max generator rotation speed on switch side (sensor 1) | 65              | Max temperature of generator (sensor 9) | 66              | Max temperature of nacelle cabinet |
| 67              | Max temperature of pitch capacitor (sensor 1) | 68              | Max temperature of pitch capacitor (sensor 2) | 69              | Max temperature of pitch capacitor (sensor 3) |
| 70              | Min high voltage of pitch capacitor (sensor 1) | 71              | Min high voltage of pitch capacitor (sensor 2) | 72              | Min high voltage of pitch capacitor (sensor 3) |
| 73              | Min low voltage of pitch capacitor (sensor 1) | 74              | Min low voltage of pitch capacitor (sensor 2) | 75              | Min low voltage of pitch capacitor (sensor 3) |
| 76              | Max temperature of generator (sensor 10) | 77              | Max temperature of generator (sensor 11) | 78              | Max temperature of generator (sensor 12) |
| 79              | Max generation time | 80              | Min generation time | 81              | Max grid downtime |
| 82              | Min grid downtime | 83              | Failure status | 84              | Power limitation status |
| 85              | Ave theoretical output power | 86              | Max theoretical output power | 87              | Min theoretical output power |
| 88              | Ave wind speed (redundant) | 89              | Max wind speed (redundant) | 90              | Min wind speed (redundant) |

The SCADA system may have error records or missing records, so it is necessary to filter data before using them in mutual information calculation. In this study, the complete progress of key SCADA data type extraction is shown in Figure 2.
Figure 2. Complete flow chart of Key SCADA data types extraction

In Figure 3, the distribution of Copula density functions between several typical SCADA data types and the turbine output power in the process of key SCADA data types extraction are shown.
For each wind turbine's data type, the mutual information with the Ave output power is calculated. Furthermore, we are taking the averaging value of 50 turbines to sort all data types. Furthermore, some data types have repeating or similar meanings, such as the "Ave wind speed" and "Max wind speed." For these kinds of data types, only one of them is retained. In this study, since a wind turbine performance model will be constructed later using the top-ranked data types, so 7 data types are finally extracted. From the 1st rank to 7th rank, they are: Max temperature of generator rear bearing, Max temperature of generator front bearing, Ave grid voltage of the A phase, Ave reactive power, Ave grid current of the A phase, Ave wind speed, Ave rotor speed.

3. Wind turbine performance model construction by DNN

After selecting key SCADA data types using mutual information, we need to prove this method's validity. In this study, a wind turbine performance model using a deep neural network was constructed. The seven selected key SCADA types are used as the input layer of DNN, and the output layer is the predicted turbine output power. Finally, we compared the output power given by DNN with the real output power, using the RMSE (root mean square error) as an indicator to make the verification.

3.1 DNN model parameters

The deep neural network is a typical technic in machine learning. In this study, we used the selected SCADA data as the network input layer. After spreading through a series of hidden layers, a predicted turbine output power is given by the deep neural network output layer. A typical structure of the deep neural network is shown in Figure 4.
A DNN model has several parameters that can be modified, and different parameter settings can influence the model's performance [14].

In this study, one of the fifty wind turbines in the wind farm is chosen as the reference turbine, which means its SCADA data will be used to train the DNN model (80% of data used as a training set 20% of data used as validation set). An optimization process is allied to three DNN hyperparameters: DNN shape, number of hidden layers, and neuron numbers of the first hidden layer. The optimal value ranges of the above three parameters are shown in Table 2.

Table 2. Deep neural network optimization parameters

| DNN optimal hyperparameters                              | Optimal value range |
|----------------------------------------------------------|---------------------|
| DNN shape                                                | Funnel, Uniform     |
| Number of hidden layers                                  | 2, 3, 4, 5          |
| Neuron numbers of the first hidden layer                 | 64, 128, 256, 512   |

Here, the uniform DNN shape signifies the number of neurons of each hidden layer are equal, and the funnel DNN shape means the number of neurons of each hidden layer decreases by 50%.

The training result of DNN with different hyperparameters on reference turbine SCADA data is shown in Table 3. And two preliminary conclusions can be concluded from the training results:

The optimal hyperparameter configuration is a five hidden layers funnel network with 64 neurons in the first hidden layer. Its final RMSE is 6.95kW, which is 0.35% relative to the turbine's rated power of 2000kW. This RMSE value is small enough to prove that the turbine DNN model constructed by selected SCADA data types is accurate.

Using the data selected by mutual information, the DNN model constructed obtains an available result on the reference turbine's data set. The training results are not overly dependent on hyperparameters' configuration since the largest RMSE of all models is at most 16.77kW, which is 0.84% relative to the turbine's rated power of 2000kW.

Table 3. DNN model training results (on reference turbine)

| Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) | Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) |
|-------------------------|-----------|-----------------------------------------|---------------------------------|-------------------------|-----------|-----------------------------------------|---------------------------------|
| 2                       | Funnel    | 64                                      | 9.23                            | 4                       | Funnel    | 64                                      | 7.75                            |
| 2                       | Funnel    | 128                                     | 8.73                            | 4                       | Funnel    | 128                                     | 7.88                            |
| 2                       | Funnel    | 256                                     | 9.84                            | 4                       | Funnel    | 256                                     | 13.16                           |
3.2 Generalizability verification of trained DNN model

The DNN model of wind turbine trained is already proved has enough accuracy on reference turbine's data set. That means if the wind farm gets new coming SCADA data in the future of this turbine, the trained model can be applied directly. However, whether the trained DNN model can be used for other forty-nine wind turbines requires further verification, which means verifying the model's generalizability.

Suppose the trained DNN model on a reference wind turbine cannot be promoted among others. In that case, the above technical route does not satisfy the condition for use in the industrial field because we need to complete a training process for each turbine separately, which costs much time.

The models with different hyperparameter configurations are applied to other 49 wind turbines SCADA data set to verify the generalizability. The Ave value of the RMSE between the predicted output power and the real output power on 49 wind turbines was used as the indicator. The results are shown in Table 4.

The average RMSE of all models on 49 turbines is 27.32kW, the smallest RMSE is 12.59kW, and the greatest RMSE is 115.21kW. Relative to the wind turbine's rated power of 2000kW, the RMSE ratios are respectively 1.37%, 0.63%, and 5.76%. Even compared to DNN models' results of the reference turbine, the RMSE is overall increased, but it is still accurate.

Table 4. Trained DNN Model generalizability verification result

| Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) | Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) |
|-------------------------|-----------|-----------------------------------------|----------------------------------|-------------------------|-----------|-----------------------------------------|----------------------------------|
| 2                       | Funnel    | 512                                     | 8.87                             | 4                       | Funnel    | 512                                     | 16.56                             |
| 2                       | Uniform   | 64                                      | 11.02                            | 4                       | Uniform   | 64                                      | 7.33                              |
| 2                       | Uniform   | 128                                     | 10.82                            | 4                       | Uniform   | 128                                     | 10.66                             |
| 2                       | Uniform   | 256                                     | 7.47                             | 4                       | Uniform   | 256                                     | 16.77                             |
| 3                       | Funnel    | 64                                      | 8.38                             | 5                       | Funnel    | 64                                      | 6.95                              |
| 3                       | Funnel    | 128                                     | 8.05                             | 5                       | Funnel    | 128                                     | 9.43                              |
| 3                       | Funnel    | 256                                     | 13.72                            | 5                       | Funnel    | 256                                     | 10.05                             |
| 3                       | Funnel    | 512                                     | 13.05                            | 5                       | Funnel    | 512                                     | 11.7                              |
| 3                       | Uniform   | 64                                      | 8.56                             | 5                       | Uniform   | 64                                      | 10.98                             |
| 3                       | Uniform   | 128                                     | 7.12                             | 5                       | Uniform   | 128                                     | 12.54                             |
| 3                       | Uniform   | 256                                     | 14.13                            | 5                       | Uniform   | 256                                     | 13.88                             |
| 3                       | Uniform   | 512                                     | 13.95                            | 5                       | Uniform   | 512                                     | 9.89                              |

3.2.5 Generalizability verification of trained DNN model

The DNN model of wind turbine trained is already proved has enough accuracy on reference turbine's data set. That means if the wind farm gets new coming SCADA data in the future of this turbine, the trained model can be applied directly. However, whether the trained DNN model can be used for other forty-nine wind turbines requires further verification, which means verifying the model's generalizability.

Suppose the trained DNN model on a reference wind turbine cannot be promoted among others. In that case, the above technical route does not satisfy the condition for use in the industrial field because we need to complete a training process for each turbine separately, which costs much time.

The models with different hyperparameter configurations are applied to other 49 wind turbines SCADA data set to verify the generalizability. The Ave value of the RMSE between the predicted output power and the real output power on 49 wind turbines was used as the indicator. The results are shown in Table 4.

The average RMSE of all models on 49 turbines is 27.32kW, the smallest RMSE is 12.59kW, and the greatest RMSE is 115.21kW. Relative to the wind turbine's rated power of 2000kW, the RMSE ratios are respectively 1.37%, 0.63%, and 5.76%. Even compared to DNN models' results of the reference turbine, the RMSE is overall increased, but it is still accurate.

Table 4. Trained DNN Model generalizability verification result

| Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) | Number of hidden layers | DNN shape | Neuron numbers of the first hidden layer | RMSE on the validation set (kW) |
|-------------------------|-----------|-----------------------------------------|----------------------------------|-------------------------|-----------|-----------------------------------------|----------------------------------|
| 2                       | Funnel    | 64                                      | 26.36                            | 4                       | Funnel    | 64                                      | 42.14                             |
| 2                       | Funnel    | 128                                     | 15.26                            | 4                       | Funnel    | 128                                     | 16.39                             |
| 2                       | Funnel    | 256                                     | 14.76                            | 4                       | Funnel    | 256                                     | 26.06                             |
| 2                       | Funnel    | 512                                     | 115.21                           | 4                       | Funnel    | 512                                     | 28.23                             |
| 2                       | Uniform   | 64                                      | 18.73                            | 4                       | Uniform   | 64                                      | 12.59                             |
| 2                       | Uniform   | 128                                     | 32.81                            | 4                       | Uniform   | 128                                     | 20.3                              |
| 2                       | Uniform   | 256                                     | 32.15                            | 4                       | Uniform   | 256                                     | 19.43                             |
| 2                       | Uniform   | 512                                     | 55.96                            | 4                       | Uniform   | 512                                     | 25.29                             |
| 3                       | Funnel    | 64                                      | 18.82                            | 5                       | Funnel    | 64                                      | 22.41                             |
| 3                       | Funnel    | 128                                     | 20.46                            | 5                       | Funnel    | 128                                     | 24.68                             |
Moreover, we also found that the optimal hyperparameters configuration obtained on the reference turbine is no longer the optimist on the other 49 turbines. So, it can be verified that the wind turbine performance model constructed on the reference turbine is generalized. It can be applied directly to other turbines without any re-training process. Nevertheless, since the optimal hyperparameters configuration obtained on the reference turbine is not the global optimizer. So in practice, the optimal hyperparameters configuration needs to be determined based on all turbines.

4. Conclusion

In this article, a wind turbine performance model construction technical route is introduced. There are two main steps of this route.

The first step is the selection of key SCADA data types. Compared to the empirical judgment method, which is widely used in the wind power industry or the methods based on the theory of feature extraction, this research tries to start from mutual information to select the SCADA data types which are more relevant to the turbine output power. The second step is training the wind turbine DNN model on a reference turbine, including the optimization of hyperparameters configuration. Using RMSE between predicted output power and real output power as an indicator, the trained model's accuracy is proved. And applying the trained model to other turbines in the same wind farm, the trained model's generalizability is also verified.

Due to the limitations of the SCADA system, in this study, some data types that are traditionally considered to have an essential relevance with the output power of wind turbines are missing, such as the pitch angle and tip speed ratio. So, the relevance between them and the output power is still needed to be verified in future work.

Simultaneously, for this wind turbine performance model construction technical route, future work can conduct more in-depth research on the generalizability. This study only verified the trained model's generalizability to the same wind farm for turbines of the same type. We suppose it can be further demonstrated that the model trained by reference turbine can also be promoted in different wind farms or be promoted among different wind turbine types. In that case, this technical route will have a higher value in the wind power industrial field.

References

[1] Sun P, Li J, Wang C and Lei X. A generalized model for wind turbine anomaly identification based on SCADA data 2016. Applied Energy. 168 550–567
[2] Lapira E, Brisset D, Ardakani H D, Siegel D and Lee J. Wind turbine performance assessment using multi-regime modeling approach 2012. Renew Energy. 45 86–95
[3] Hung L P, Alfred R, Hijazi A and Hanafi M. A Review on Feature Selection Methods for Sentiment Analysis 2015. Adv. Sci. Lett. 21 2952–2965
[4] Guyon I and Elisseeff A 2006. An introduction to feature extraction. (Berlin: Springer)
[5] Hong Z Q. Algebraic feature extraction of image for recognition 1991. Pattern Recognit. 24 211–219.
[6] Li H, Jiang T and Zhang K. Efficient and robust feature extraction by maximum margin criterion
2006. *IEEE Trans. Neural Netw.* **17** 157–165.

[7] Ding J, Wen C, Li G and Chua C S. Locality sensitive batch feature extraction for high-dimensional data 2016. *Neurocomputing.* **171** 664–672.

[8] Edzel L, Dustin B, Hossein D A, David S and Jay L. Wind turbine performance assessment using multi-regime modeling approach 2012. *Renewable Energy.* **45** 86-95.

[9] Shannon C. A mathematical theory of communication 1948. *System Technical Journal.* **27** 379-423.

[10] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l’Institut de Statistique de L’Université de Paris, 8, 229–231.

[11] Hominal P and Deheuvels P. Estimation non paramétrique de la densité compte-tenu d’informations sur le support 1979. *Rev. Stat. Appl.* **27** 47–68

[12] Nagler T and Czado C. Evading the curse of dimensionality in multivariate kernel density estimation with simplified vines 2016. *J. Multivar. Anal.* **151** 69–89.

[13] Aurélien G 2017. *Hands-on Machine Learning with Scikit-Learn & TensorFlow.* (America: O'Reilly).

[14] Stuart R and Peter N. *Artificial Intelligence: A Modern Approach, Third Edition.* (New Jersey:Pearson Education)