A Multi-Timescale Wind Power Forecasting Method Based on Selection of Similar Days

WANG Shen-zhe[1], GAO Shan[2], ZHAO Xin[2], ZHANG Ningyu[1]
(1. Jiangsu Power Supply Company, Nanjing, Jiangsu Province, 210000; 2. School of Electrical Engineering, Southeast University, Nanjing, Jiangsu Province, 210096)

***Abstract*—In the progress of grid planning schemes selection, annual wind power data in different timescale is needed in the calculation of characters. In this paper, a multi-timescale wind power forecasting method based on selection of similar days is proposed. The annual wind power will be clustered by scale of days, and then the LS-SVM forecasting model of each cluster will be trained separately. In the forecasting progress, a similar day of the forecasted day is selected in its cluster as training data. The forecasting model of this cluster will be used to get forecasting results, the time-resolution of which is in level of hour. The forecasting wind power of each day will be combined into annual wind power. A case studied shows that the proposed forecasting method is correct and effective.**

***Index Terms*—multi-timescale; similar days; clustering; least square support vector machine; wind power; forecasting**

I. I NTRODUCTION

Decision-making is an important part of power network planning, which depends on different kinds of assessment index. With large scale of wind power connected, the effect of wind power should be considered and various wind power prediction data will be needed in the definition and calculation of assessment index. Different index need different data. For example, capability index need annual wind power data in the resolution of hour, when power quality index need annual wind power data in minute. Generally wind power data in hour can be predicted by long-term method and wind power data in minute will be get by short-term method. The existing method will be unfitted when two resolutions of wind power data needed at the same time. New wind power data prediction method which can provide annual data in two resolutions should be built up.

In order to break up the limit of time-series prediction method, the prediction in one year time long will be separated and companied. The prediction method based on similar day selection always has the opinion that wind power in different days can be correlated and regularized. By using certain size of similar day wind power data as base data and new

prediction model can be trained, by which the annual wind power prediction can be separated into each days. In day scale, wind power can be predicted in the resolution of minute. The combination of each day prediction data will be the annual wind power prediction data with resolution in minute.

The actual wind power prediction method can be mainly divided into two kinds. One kind of it get wind speed prediction result firstly, which will be transferred to wind power prediction data by speed-power curve. The other kind can get prediction result directly. In this paper, the first kind of method will be used.

At present, there is also much research on wind power prediction based on similar-days. A wind power prediction method based on similar-periods was presented in [4], which has low prediction accuracy. A wind power prediction method based on similar-days and LSLM in [5], in which the most similar-days in year can't be selected. The prediction method based on similar-day NP was presented in [6], by which wind power prediction error was use to get prediction model. A prediction model based on similar-days was presented in [7], in which each kind of cluster uses the same kind of prediction model.

In this paper, the wind speed data in year-long time will be analysis in cluster firstly. Then LS-SVM prediction model will be built separately for each cluster. In the progress of prediction of future days, we firstly choice similar-days wind power data as the historical data from the cluster which the future day belong to. Then the LS-SVM prediction model for the chosen cluster will be used to get predicted wind power data in the future day. After we get prediction wind power of each day, the year-long time prediction result by combined the prediction wind power data in days. In order to avoid the prediction error bring by speed-power curve of signal wind generation. One kind of speed-power curve for whole wind farm was proposed in this paper, by which the wind speed prediction result will be transferred into wind power data.

II. SIMILAR-DAY CLUSTER

This work was supported by the National High-Technology Research and Development Program (863 Program) of China (2011AA05A105)

A. Unsupervised Day-feature Cluster

Wind speed has a certain extent similarity in day-time. Wind speed in same day but different years will be relevant more or less. One kind of similar-day cluster model was built in this paper as the basement of yearly wind speed prediction in multi-time scale. In order to divided the yearly wind speed into several sort, physical quantities which can reflect the character of similar-days should be selected. According to wind speed variation trend, amplitude, and volatility, the similar-day wind speed sample can be constructed as

$$d=\left[d_{s1},d_{s2},\cdots,d_{sn},d_{s\max},d_{s\min},d_{smean},d_{sstd}\right] \quad (2)$$

$d_{s1},d_{s2},\cdots,d_{sn}$ was wind speed in each hour of a day. $d_{s\max}$ was the maximum wind speed in one day and $d_{s\min}$ was the minimum wind speed. $d_{s\,mean}$ was the average wind power . d_{sstd} was the standard deviation of day wind power.

The wind speed variation trend and amplitude was considered compositely in this sample. The sample vector will be normalized to avoid the error of dimension difference.

$$d_{ji}=\frac{d_{ji}-d_{j\min}}{d_{j\max}-d_{j\min}} \quad (1)$$

$i=1,2,\cdots,m$ Was totality of the sample. $d_{i\min}$ Was the minimum value of vector J, $d_{i\max}$ was the maximum value of vector J.

K-means is the one of the most popular clustering methods at present time, which has high calculation speed and high precision. If the cluster sample is $X=\{x_k\,|x_k\in R^p,k=1,2,\cdots,N,p\in N^*\}$, the classify number is K , part of the sample can be indicated as $Z=\{z_c\,|z_c\in R^p,c=1,2,\cdots,K,p\in N^*\}$. We can get

$$z_c=\frac{1}{K_c}\sum x_k \quad (2)$$

K_c is the capacity of cluster c . The objective function can be defined as

$$Q=\sum_{c=1}^{K}\sum_{k=1}^{K_c}d_{ck}(x_k,z_c) \quad (3)$$

$d_{ck}(x_k,z_c)$ is the distance between the kth element of cluster c and the center of the cluster. The normalized day wind speed vector can be measure by Euclidean distance as

$$d_{ck}(x_k,z_c)=\sqrt{(x_k-z_c)^T(x_k-z_c)} \quad (4)$$

The objective function Q is the quadratic sum of the distance between each cluster element and its center. The target of cluster is to get minimum Q .

An important step of K-means is determining the initial cluster center. In order to avoid the error bring by determining the initial cluster center at random, an initial cluster center selected method based on DKC was proposed in this paper.

The DKC method use $d-$ distance neighbor number and $2d-$ distance neighbor number to determine the DKC value, and also the initial cluster center can be selected.

Definition 1: x_i is one element of one cluster. $d_{\max}$ is the maximum in all d distance close to x_i , which is called $d-$distance.

Definition 2: All the data objectors in the $d-$ radius cluster of x_i was its $d-$ radius neighbor, which can be expressed as:

$$d_{x_i}=\left|x_j\in X\right| \quad dist(x_i,x_j)\le d_{\max},i\ne j \quad (5)$$

Also, the data objectors in 2 $d-$ radius cluster of x_i was its 2 $d-$ radius neighbor, which can be expressed as

$$d_{x_i}=\left|x_j\in X\right| \quad dist(x_i,x_j)\le d_{\max},i\ne j \quad (6)$$

Definition 3: The DKC value of the data objector is the ratio of $2d-$neighbor and $d-$neighbor.

$$DKC_{x_i}=\frac{2d_{x_i}}{d_{x_i}} \quad (7)$$

If the DKC value of one data objector is above the average, which means the area around the objector is sparse. In this situation, this objector can be treated as isolated point. In opposite, the DKC value below the average means the area around the objector is crowd, which also means this objector can be treated as one cluster center.

B. Similar-days selection based on cluster result

Days with similar wind speed was divided by cluster result preliminarily. When wind speed of future days was forecasted, the days in the cluster that the forecasting day belongs which has the similar day-feature will be choose as train sample.

Based on normalized diurnal characteristics vector, the vector of forecasting day is x_0 . The vector of j day in the cluster of forecasting day can be indicated as:

$$\begin{aligned}x_0&=\left[x_0(1),x_0(2),\cdots,x_0(n)\right]^T\\ x_j&=\left[x_j(1),x_j(2),\cdots,x_j(n)\right]^T\end{aligned} \quad (8)$$

The correlation coefficient of factor k about x_0 and x_i can be expressed as"

$$\varepsilon_j(k)=\frac{\min\limits_j\min\limits_k\left|x_0(k)-x_j(k)\right|+\rho\max\limits_j\max\limits_k\left|x_0(k)-x_j(k)\right|}{\left|x_0(k)-x_j(k)\right|+\rho\max\limits_j\max\limits_k\left|x_0(k)-x_j(k)\right|} \quad (9)$$

ρ is resolution ratio, which always can be 0.5. In the integration of correlation coefficient between each factors, similarity of x_0 and x_i can be defined as:

$$F_j=\prod_{k=1}^{n}\varepsilon_j(k) \quad (10)$$

Dominated factor can be identified automatically by the similarity defined by multiplication, which also can avoid the weighting between each factors. Certain amount of sample days can be choose as train data from the cluster the forecasting day belong by similarity.

III. SVM FORECASTING METHOD

A. SVM introduction

SVM(support vector machine) method is one kind of machine learning method which presented in 1990's based on statistical theory and principle of minimum structure risk. This method is developed quickly in the research of wind speed and wind power forecasting. In comparison to Neural network method, SVM method which has better global optimization and generalized ability is more fitted to solve small sample learning question. The standard SVM method has the defects that the training is too complicated. In recent years, lots of advanced SVM method raised quickly. In the LS-SVM method, model training was replaced by linear equations solver, which enhances the calculation speed greatly.

B. LS-SVM

Assumed training sample sets is $\{x_1, x_2, \cdots, x_n | x \in R^n\}$, the output of training sets is $\{y_1, y_2, \cdots, y_n\}$. If training sample x_i was mapped from characteristic area R^n to R^n , the estimation function is $y(x_i) = w\varphi(x_i) + b$. w is weighting vector, b is constant value.

Resolving estimation value of w and b is equal to solving the optimization question as

$$\min g(w,\xi) = \frac{1}{2}ww + \gamma \sum_{i=1}^{n} \xi_i^2 \qquad (11)$$

$$s.t \quad y_i = w\varphi(x_i) + b + \xi_i$$

γ is punishment coefficient, ξ_i is estimation error. When LaGrange multiplier was introduced, the optimization question will change to:

$$L(w,b,\xi,\lambda) = g(w,\xi) - \sum_{i=2}^{n} \lambda_i (\varphi(x_i)w + b + \xi_i - y_i) \quad (12)$$

$\lambda_i (i = 1, 2, \cdots, n)$ is LaGrange multiplier.

On the basis of Karush-Kuhn-Tucker condition, the first order derivative of LaGrange function is 0. In this situation, we can get

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \\ \dfrac{\partial L}{\partial b} = 0 \\ \dfrac{\partial L}{\partial \xi} = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} w = \sum_{i=1}^{n} \lambda_i \varphi(x_i) \\ \sum_{i=1}^{n} \lambda_i = 0 \\ \lambda_i = \gamma \xi_i \\ w\varphi(x_i) + b + \xi_i = y_i \end{cases} \qquad (13)$$

After ξ and w , the optimization question will be transferred into linear system

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & k(x_1,x_1) + \frac{1}{r} & \cdots & k(x_1,x_n) \\ & & \vdots & \\ 1 & k(x_n,x_1) & \cdots & k(x_n,x_1) + \frac{1}{r} \end{bmatrix} \begin{bmatrix} b \\ \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad (14)$$

Solving this linear system can get

$$y_i = \varphi(x_i) \sum_{i=1}^{n} \lambda_i \varphi(x_i) + \frac{\lambda_i}{\gamma} + b \qquad (15)$$

In order to solving high-dimensional computing problem of the linear system, the inner product will be instead by kernel function $K(x, x_i)$. There are several kinds of kernel function used widely, such as linear kernel function, polynomial kernel function, RBF kernel function and multilayer perception. In this paper, the LS-SVM model will be built on the integration of LS-SVM tool box and matlab program.

IV.EXAMPLE

In order to assurance the accuracy of the forecasting model, the training data for cluster analysis should be adequate. The wind speed and wind power data of one wind farm in china in two years was choose to be training data. The duration of the training data is from Juan 1 2004 to Dec 31 2005. The data in 2004 was used to train forecasting model and the model will be used to get wind speed and wind power data in 2005. The actual data of 2005 can be used to check the accuracy of the model.

A. Cluster result analysis based on similar days

Wind speed data of 2004 was the cluster object. The number of cluster center is 10 and the cluster result was shown as:

Tab1: Similar Day Clustering Result of Wind Farm Ⅰ

cluster	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ	Ⅵ	Ⅶ	Ⅷ	Ⅸ	Ⅹ
Capability	32	14	45	41	37	30	48	42	54	23

In order to determine the applicability of the clustering algorithm, wind speed data of another wind farm in 2004 was used to get cluster result as

Tab2: Similar Day Clustering Result of Wind Farm Ⅱ

cluster	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ	Ⅵ	Ⅶ	Ⅷ	Ⅸ	Ⅹ
Capability	37	20	35	37	30	23	52	45	50	37

The two wind farm is not far, so the cluster result is similar more or less.

The wind speed similar day can be divided obviously when cluster center number was 10. The cluster result of cluster 4 was shown as:

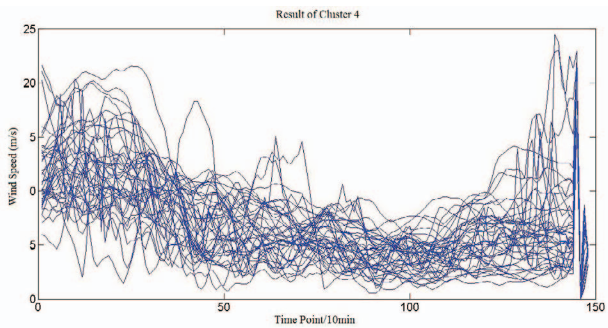

Fig1: Clustering Result Curve of Fourth Category

B.Wind power forecasting based on LS-SVM model

In order to improve the accuracy of forecasting model, about quarter of data nearest to the cluster center was selected as training data. In the progress of forecasting for future days, the cluster model that the forecasted day belongs will be selected firstly, and then training data was selected from the historical data of this cluster kind.

Generally wind farm speed-power relationship of signal wind turbine was also used in wind farm, which has low accuracy in large-scale wind farm. In this paper, the wind speed-power relationship of wind farm was matched by actual wind speed and wind power data of that wind farm. The cftool toolbox of matlab was used in this paper. The fitted result can be shown as:

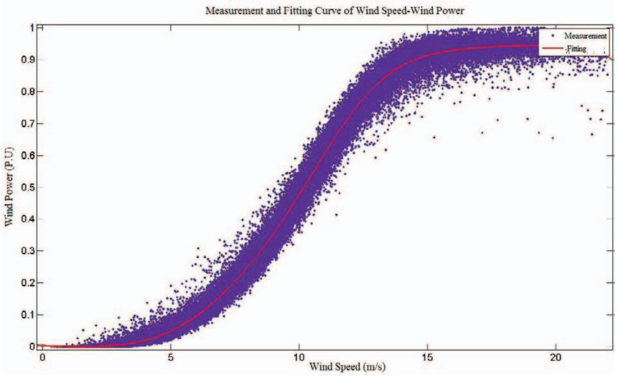

Fig2:Measurement and Fitting Curve of Wind Speed-Power

The comparison of forecasting and actual wind power data was shown as:

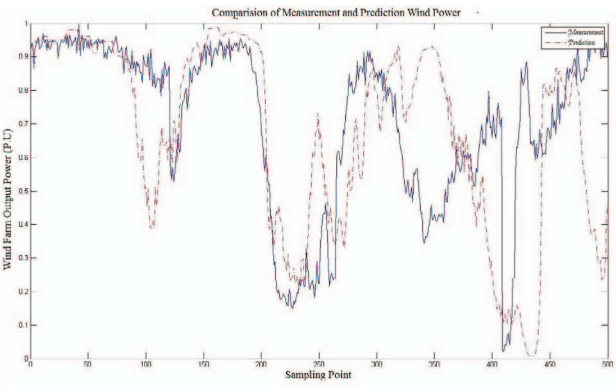

Fig3:Comparision of Measurement and Prediction Wind Power

In this paper, the root-mean square error was used to measure the forecasting error:

$$E = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{Mi} - P_{Li})^2}}{P_{Cap}} \times 100\% \qquad (16)$$

E is root-mean-square error, P_{li} is forecasting wind power of i, P_{mi} is actual wind power, P_{cap} is wind farm capability. Because of the forecasting theory based on similar days, the maximum error, minimum error, average error and standard error was considered separately in this paper. The error of this forecasting method was also compared with persistence method, SVM method and neural network method based on similar day, which shown as

Tab3: Error Comparison of Different Prediction Methods

Method	maximum	minimum	average	standard
Persistence Method	74.47	12.41	35.84	18.13
SVM Method	44.91	13.53	27.84	6.44
NN Method	44.29	11.71	19.25	7.57
Paper Method	43.92	2.84	17.72	8.51

From the data in the tab we can know that the forecasting accuracy of this method is better than the other three methods. In one year time-long, the average forecasting error is 17.83%. In this paper, the annual wind power forecasting data has the time-scale in minute, which cover the shortage of rough time-scale in the traditional long-time wind power forecasting method.

V.CONCLUSION

The multi-timescale wind power forecasting method based on similar days was proposed in this paper. The cluster based on similar-days and LS-SVM method was used in wind speed forecasting with time –scale is 10min. The annual wind speed forecasting data can be combined by forecasting result in day time. The wind speed-power relation can be fitted by actual wind speed and power data, by which the forecasting wind speed can be transferred into wind power data. The advancement of cluster algorithm and SVM model is the point in further research.

PREFERENCE

[1] Wang Xiaolan，Li Hui. Effective wind speed forecasting in annual prediction of output power for wind farm[J]． Proceeding of the CSEE， 2010， 30(8)：117-122．

[2] Liu Chun，Lv Zhenhua，Huang Yuehui，et.al．A new method to simulate wind power time series of large time scale[J]．Power System Protection and Control，2013，41(1)：7-13．

[3] Kusiak A， Zijun Zhang． Short-horizon prediction of wind power: a data-driven approach[J] ． IEEE Transactions on Energy Conversion， 2010， 25(4)：1112-1122．

[4] Wang Xiaolan， Ge Pengjiang． PV array output power forecasting based on similar day and RBFNN[J]． Electric Power Automation Equipment，2013，33(1)：100-103．

[5] Jiang Xiaoliang，Jiang Chuanwen，Peng Minghong，et.al． A short-term combination wind speed forecasting method considering seasonal periodicity and time continuity[J]． Automation of Electric Power Systems，2010，34(15)：75-79．

[6] Meng Yangyang，Lu Jiping，Sun Huali，et.al. Short-term wind power forecasting based on similar days and artificial neural network[J]. Power System Technology，2012，34(12)：163-167．

[7] Yang Xiyun ， Sun Baojun ， Zhang Xinfang ， et.al． Short-term wind speed forecasting based on support vctor machine with similar data[J]． Proceeding of the CSEE，2012，32(4)：35-41．

[8] Ding Zhiyong，Yang Ping，Yang Xi，et.al. Wing power prediction method based on sequential time clustering support vctor machine[J]. Automation of Electric Power Systems，2012，36(14)：131-135．

[9] Wagstaff K，Cardie C，Rogers S，et al．Constrained K-means clustering with backgroud knowledge[C]//Proceedings of the 18th International conference on machine learning，June 28-July 1，2001，Willamstown，MA，USA：577-584．

[10] Chen Fuji ， Jiang Fang ． Research on improved $K-means$ clustering algorithm based on two-distance[J] ． Journal of taiyuan university of technology，2012，43(2)：114-118．

[11] Mohandes M A ， Halawani T O ， Rehman S ， et al ． Support vector machines for wind speed prediction[J]．Renewable Energy，2004，29(6)：939-947．

[12] Li Jin，Liu Jinpeng，Wang Jianjun．Mid-long term load forecasting based on simulated annealing and SVM algorithm[J]．Proceeding of the CSEE，2011，31(16)：61-66．

[13] Chacon J E ， Alberto R C ． A note on the universal consistency of the kernel distribution function estimator[J]．Statistics & Probability Letters，2010，80(17-18)：1414-1419．

[14] Kuh A，Mandic D．Applications of complex augmented kernels to wind profile prediction[C]//IEEE International Conference on Acoustics, Speech and Signal Processing ， 19-24 April ， 2009 ， Taipei ：3581-3584．

[15] Du Ying，Lu Jiping，Li Qing，et.al．Short-term wind speed forecasting of wind farm based on least square-support vector machine[J] ． Power System Technology，2008，32(15)：62-66．

[16] Fu Meiping，Ma Hongwei，Mao Jianrong．Short-term photovoltaic power forecasting based on similar days and least squars support vector machine[J]． Power System Protection and Control，2012，40(16)：65-69．

[17] Li Fangfang，Zhao Yingkai，Yan Xin．Application of matlab LS-SVM toolbox[J]．Computer Applications，2006，26(12)：358-360．

Wang Shenzhe was born in 1987. He received the Bachelor, Master and Ph.D degree from Department of Electrical Engineering, Southeast University, China, in 2007, 2010 and 2014, respectively.

He is now an engineer in JiangSu Power Supply Company. His current research interests include Power system planning and analysis, wind/solar power analysis and control.

Gao Shan was born in 1973. He received the Bachelor degree from Department of Electrical Engineering, ShanDong University, China, in 1997. He received the Ph.D degree from Department of Electrical Engineering, Southeast University, China, in 2002.

He is now an associate professor in the Department of Electrical Engineering, Southeast University, China. His current research interests include Power system planning and analysis, wind/solar power analysis and control.

Zhao Xin was born in 1984. He received the Bachelor, Master and Ph.D degree from Department of Electrical Engineering, Southeast University, China, in 2005, 2008 and 2014, respectively.

He is now an lecturer in the Department of Electrical Engineering, Southeast University, China. His current research interests include Power system planning and analysis, wind/solar power analysis and control.

Zhang NingYu was born in 1985. He received the Bachelor degree from Department of Electrical Engineering, China University of Mining and Technology, China, in 2006. He received the Ph.D degree from Department of Electrical Engineering, Southeast University, China, in 2013.

He is now an engineer in JiangSu Power Supply Company, China. His current research interests include Power system planning and analysis, wind/solar power analysis and control.