# A New Evaluation Metric Reflecting the Lead-Lag Scenarios in Wind Power Forecasting


Jinlu Yuan, Fei Wang
*Department of Electrical Engineering*
*North China Electric Power University*
Baoding 071003, China
jinluyuan@ncepu.edu.cn
feiwang@ncepu.edu.cn

Miadreza Shafie-khah
*School of Technology and Innovations*
*Department of Electrical Engineering*
*University of Vaasa*
Vaasa 65200, Finland
miadreza@gmail.com

Zhao Zhen
*State Key Lab of Power System*
*Department of Electrical Engineering*
*Tsinghua University*
Beijing, China
georgiazhz@foxmail.com

João P. S. Catalão
*Faculty of Engineering*
*University of Porto and INESC TEC*
Porto 4200-465, Portugal
catalao@fe.up.pt



2021 IEEE Industry Applications Society Annual Meeting (IAS) | 978-1-7281-6401-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/IAS48185.2021.9677172



***Abstract*—Wind power forecast evaluation matters greatly as wind power has an ever-increasing proportion in the power system. Generally speaking the forecasting result can be divided into lead-lag scenarios and common scenarios which depends on whether the wind process is predicted on time. During the lead-lag scenarios the errors usually change from large positive numbers to negative ones (or the opposite), especially in the both ends of the period. Compared with the common scenarios in the same value of root mean square error (RMSE), large changes in errors from positive to negative in a short time can cost nearly two times of spinning reserve but get the same assessment score. For power system the two scenarios should be evaluated differently, however, few metrics in the evaluation can indicate the lead-lag scenarios in that they dispose the errors ignoring the signs or time continuity of the errors, or analysis the errors in a macro-scale sight like 24 hours horizon scale. This paper proposes a new metric based on RMSE which detects the changes of signs of errors in a process of moving average. Except for normal advantages like objectivity, adaptability, unity, symmetry and stability, the new metric has the ability to reflect both the lead-lag scenarios and common scenarios. The new metric can be used in the evaluation of wind and solar power, load, price, demand response forecasting and the process of neural network parameter training.**



***Keywords*—*Forecasting evaluation metric, lead-lag scenario, wind power forecasting, root mean square error, sign of error, moving average***


## I. Introduction

The development of human society needs kinds of energy sources like wood and coal, and the energy sources can be briefly divided into two categories, non-renewable and renewable energy sources [1]. With the scientific progress human are able to make use of the renewable energy to a greater extent [2], and human will achieve power systems with extremely high levels of variable renewable energy [3].

Over the last decade the penetration of the renewable energy generation in the power system has been increasing prominently [4]. The renewable energy generation is mainly divided into two parts, photovoltaic (PV) and wind power generation [5]. They both have intrinsic volatility and intermittency, which leads to adverse effect on power grid dispatching operation inescapably [6]. So various kinds of forecasting methods have recently emerged in order to reduce the negative effect caused by the uncertainty [7]. Forecasting methods can be divided into physical and statistical methods [8], and in an other angel methods can be divided into point and probability forecasting [9]. Lots of forecasting methods are carried out [10]–[15].

In the field of designing the WPF strategies, in order to proof availability and accuracy of the proposed forecasting model the author of it usually make comparison with some benchmarks where the metrics are used [16]. In [17], Wang used three metrics, which are RMSE, MAE and COR, to show the advance of proposed forecast model, and in [18] MAPE and MBE were used. Ozkan used NMAE and NRMSE for evaluating his novel WPF model [20] as well as Yang used NMAE [19].

To evaluate different WPF strategies, except for common metrics above-mentioned, there are lots of novel metrics and evaluation systems in the perspective of error distribution. In [21] Chen used Diebold-Mariano (DM) test methods for the evaluation of wind power forecasting models. Chen considered the difference impact of positive and negative errors and proposed asymmetric DM test. This can result in the deliberate adjustment in WPF by the plants. In [22] Bessa reported using criteria based on entropy of the prediction error distribution are more suitable than using the traditional minimum square error criterion to evaluate the accuracy in training WPF models. In [23] and [24] Yao used skewness and kurtosis and some classic metrics with the maximizing deviations method and grey correlation analysis to evaluate the performance of the prediction result.

These metrics describe the errors from a macro perspective, however, in [25] Ma divided the forecasting errors into two categories that are horizontal and longitudinal errors (these concepts needed reconsidered). It was an advance in forecasting evaluation but it had some defects. On the one hand they did not discover the essence of the problem. On the other hand they used r to reflect this problem which was unbefitting.

Generally speaking the forecasting result can be divided into lead-lag scenarios and common scenarios which depends on whether the wind process is predicted on time. The word lead-lag is widely used in different research fields such as economics [26], automatic control [27] and load management [28]. The concept of lead-lag is that two curves (stock price, power output, voltage, etc.) are very closely related after a shift in time. It is noted that the value of r of the two curves is usually small though they look very relevant.

Typical forecasting scenarios examples are shown in Fig. 1. The orange curve is a lead-lag scenario. The blue curve is a common scenario. The black curve is a curve of actual output.

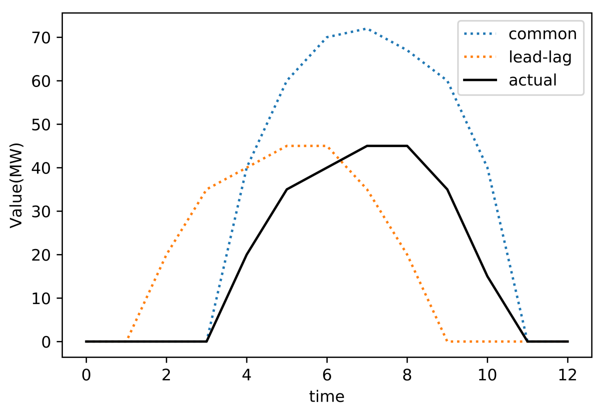


Fig. 1. Two categories of forecasting scenarios

The lead-lag curve has an intersect with the actual curve and the errors turn from positive to negative. If this lead-lag curve goes through 2 units shift in time it will meet the actual curve. On the contrary, the common curve can not conduct a similar shift and it can often be summarized as either larger or smaller in a given period of time. It is noted that the errors caused by lead-lag scenarios and common scenarios are named as lead-lag errors and common errors respectively, and we mainly focus on the lead-lag scenarios in several hours which can obviously occupy the spinning reserve of the power grid.

The lead-lag errors and common errors both have negative effects on the flexibility and reliability of the electrical power system. Large lead-lag errors in several hours will have a more serious impact in the same situation of common ones because the output of the scenarios will go through a process from surplus to vacancy in a short time, in another word lead-lag scenarios can occupy more spinning reserve of the grid relatively. However, little attention has been paid to this problem, the power dispatching system angle. The current evaluation methods can hardly reflect the lead-lag errors and this is why a new metric is needed.

Aiming at the problem, this paper proposes a new metric for the evaluation of short-term wind power forecasting based on RMSE which detects the changes of signs of errors in a process of moving average. The metric is named as sign root mean square error (SRMSE), and it can indicate both the lead-lag and the common errors.

## II. Analysis of Power Forecasting Metrics

To evaluate the adequacy of the prediction result and the performance of the forecasting model, different metrics are needed [29]. Eligible metrics should generally possess the following characteristics [30].

- Objectivity. The metrics should objectively reflect the true usability and capability of prediction in different renewable energy stations.
- Adaptability. The metrics should have a wide adaptability to different circumstances like situations in solar and wind energy stations, normal and transitional processes.
- Unity. The metrics should be practicable for different statistical period, in another word, daily and monthly evaluation can both use this metric.
- Symmetry. The metrics should reflect the quantity of the deviation rather than the sign of the deviation. The evaluation cannot induce the stations to make a prediction with preference.
- Stability. The promotion of forecasting qualify should be enhanced by imposing a penalty on sudden change of bias which possibly indicates that the spinning reserve of the power system is used out, and the stability is on the edge.

There are some classic metrics that have been used very widely in regions of China like mean absolute error (MAE) (2), root mean square error (RMSE) (3), weighted root mean square error (WRMSE) (4), mean absolute percentage error (MAPE) (5), and correlation coefficient (r) (6). e is defined as the difference between the forecasting value and actual value.

$$e_i = \hat{y}_i - y_i \tag{1}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{n}|e_i| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(e_i)^2} \tag{3}$$

$$WRMSE = \sqrt{\sum_{i=1}^{n} e_i^2 * \frac{|e_i|}{\sum_{i=1}^{n}|e_i|}} \tag{4}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}|\frac{e_i}{y_i}| \tag{5}$$

$$r = \frac{COV(y, \hat{y})}{\sqrt{V(y)}\sqrt{V(\hat{y})}} \tag{6}$$

$\hat{y}$ and $y$ are the predictive value and the actual value of wind power output respectively in the equations above, and $e_i$ is the difference between $\hat{y}_i$ and $y_u$. $\bar{e}$ is the average value of $e_i$, and $N$ is the number of output points everyday. $COV(y, \hat{y})$ and $V(y)$ are the covariance between $(y, \hat{y})$ and the variance of $y$ respectively.

Fig. 2 is the differences between forecasting value and actual value in the situation of Fig. 1.

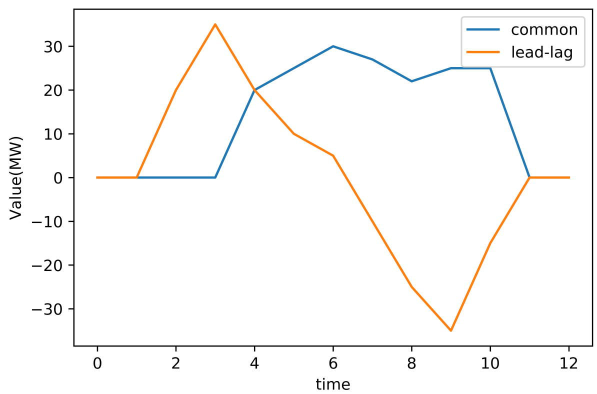


Fig. 2. The errors of two scenarios

The values of metrics above are calculated as shown in Tab. I.

TABLE I
COMPARISON OF METRICS IN TWO CONDITIONS

| metric | common | lead-lag | distinction |
|---|---|---|---|
| MAE | 13.384 | 13.461 | no |
| RMSE | 18.372 | 18.239 | no |
| WRMSE | 25.394 | 26.497 | no |
| MAPE | 0.847 | inf | no |
| r | 0.986 | 0.516 | yes |
| SRMSE(7) | 18.372 | 24.501 | yes |

It is observed that except for r the others cannot indicate the lead-lag errors due to their values are close. The reasons are that on the one hand these metrics ignore the signs of the errors, on the other hand they ignore the time continuity of the errors, in another word, errors from the shuffled series in the same order will get the same result.

r shows their difference, however, r has its shortcomings. r does not directly reflect errors but it indicates the linear relationship between the forecasting and the actual value divided by their respective mean value, so the value of r is between -1 and 1, which results in difficulty in employ. It is not objective due to the result cannot be used to evaluate qualities of the prediction at each plant. In Central China Region r is used in a particular way that a day will be denoted as unqualified if the intraday r is smaller than 0.68. This method has little adaptability nor quantification because it cannot distinguish the situation when r is 0.2 and r is 0.5 and it can rigidly judge the day qualified with r being 0.69. From another point, r cannot focus on some period of time during the day but the lead-lag errors which are in the successive hours would cause harm to the power system.

## III. METHODOLOGY

### A. The form of data

The short-term wind power forecasting predicts the day-ahead output of the wind turbine generators of plant. The time resolution is 15 minutes so there should be 96 values every day as well as 96 values every prediction.The data to be evaluated are two series, one is the prediction series and the other is the actual series. It is noted that the proposed metric can be used in any data with lengths that is greater than the moving step size like ultra short-term (16 points) and monthly evaluation (more than 1000 points).

### B. The design of the metric

As in (1), $\hat{y}_i$ and $y_i$ are elements of the prediction series and the actual series respectively. $e_i$ is the difference of $\hat{y}_i$ and $y_i$, which is an element of the error series. In the following parts $e_i$ is named as the original error series.

In order to detect the changes of error sign we use the central moving average method where the weights are different. As in (7) and (8), $a_i$ and $b_i$ are the value of different central moving average. Parameter k is moving step size and it is usually an odd number due to the reason that $e_i$ in the center of the step with the same length of $(k-1)/2$ points distributed in both sides.

$$a_i = \frac{1}{k}\sum_{j=1}^{k} |e_{i+j-(k+1)/2}| \tag{7}$$

$$b_i = |\frac{1}{k}\sum_{j=1}^{k} e_{i+j-(k+1)/2}| \tag{8}$$

For example, given $e_1, e_2, e_3, e_4$ as an error series where n=4. if k=3, then we get $a_2 = (|e_1| + |e_2| + |e_3|)/3$ and $a_3 = (|e_2| + |e_3| + |e_4|)/3$. There are not $a_1$ or $a_4$, and the lengths of the two new series are both $n - k + 1$ with indexes being from $(k+1)/2$ to $n - (k+1)/2$ +1.

At present we have three series, the primary error series e and the two moving average series a and b, which is the fatal step. The lead-lag errors can be disclosed by the two additional series. The difference between series a and b can indicate the deviation of forecast trend, which can improve the performance of the original RMSE.

Three different ways of using a and b are described as follows, which are numbered as I-SRMSE, II-SRMSE and III-SRMSE.

*1) I-SRMSE:* As in (9), the value of RMSEe is the root mean square value of the error series, and similar definitions are used in RMSEa, RMSEb and RMSEc subsequently. Thus I-SRMSE comes out. This method disposes the error series independently and amends the RMSE of error series by twice calculation RMSE at last.

$$\begin{aligned}
&RMSEe=\sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i^2}\\
&RMSEa=\sqrt{\frac{1}{n-k+1}\sum_{i=(k+1)/2}^{n-(k+1)/2+1}a_i^2}\\
&RMSEb=\sqrt{\frac{1}{n-k+1}\sum_{i=(k+1)/2}^{n-(k+1)/2+1}b_i^2}\\
&I-SRMSE=RMSEe+RMSEa-RMSEb
\end{aligned}\tag{9}$$

*2) II-SRMSE:* As in (10), II-SRMSE puts the error series a and b together by using their difference to form a new seires c. It amends RMSE of the original error series e by once calculation RMSE at last.

$$\begin{aligned}
&c_i=a_i-b_i\\
&RMSEc=\sqrt{\frac{1}{n-k+1}\sum_{i=\frac{(k+1)}{2}}^{n-\frac{(k+1)}{2}+1}c_i^2}\\
&II-SRMSE=RMSEe-RMSEc
\end{aligned}\tag{10}$$

*3) III-SRMSE:* As in (11), III-SRMSE directly amends the original error series then calculates its RMSE value.

$$\begin{aligned}
&e_i=\begin{cases}|e_i|+a_i-b_i, \frac{(k+1)}{2}\le i\le n-\frac{(k+1)}{2}+1\\ |e_i|, else\end{cases}\\
&III-SRMSE=RMSEe
\end{aligned}\tag{11}$$

All of these three different methods are based on RMSE and when k=1 they are equal to RMSE, thus they can possess the advantages of RMSE. More importantly, they have the ability to reflect both the lead-lag and common errors. But it has shortcomings that it is not suitable when the loss is proportional to the error, and it costs more time in the usage of training the neuronal network, and it can be only used for point forecasting evaluation.

## IV. CASE STUDY

There are 4 subsections in the case study. Firstly in subsection A we choose the data. Secondly in subsection B we compare three types of SRMSE in different values of k and choose the most appropriate one. Then in subsection C .At last in subsection D we make a comprehensive summary of the new metric using all of 5 stations.

### *A. Data sources*

Based on RMSE so the proposed metric can be used in any scene of forecasting evaluation in principle. Photovoltaic power and wind power are the main forms of renewable energy generation. They are quite different that PV power has less variability and intermittency and the prediction of it is smoother and steadier than wind power, which is the reason why there are less conspicuous lead-lag errors in prediction for wind power as shown in Fig. 3and PV power as shown in Fig. 4 which come from Hebei Province, China. Therefore, the WPF data are more suitable to demonstrate the improvement and advantages of SRMSE.

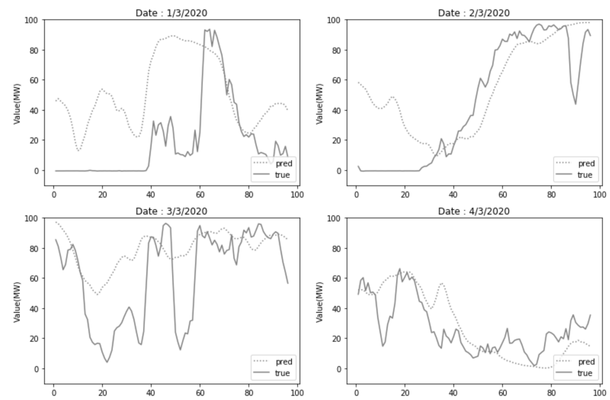


Fig. 3. Examples of wind power prediction

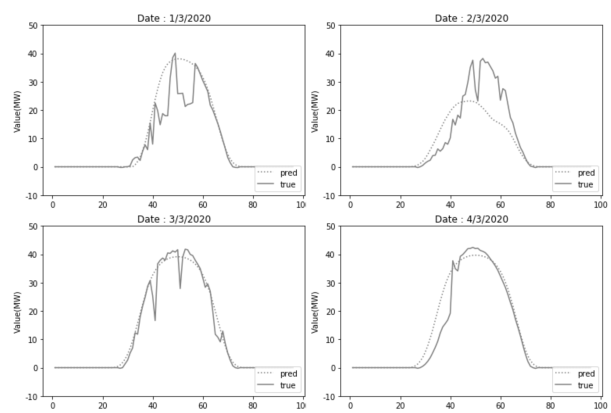


Fig. 4. Examples of PV power prediction

The forecasting and actual output data are from 5 wind power generation stations in Hebei Province, China. The prediction provided is short-term forecasting and the time resolution is 15 minutes. We choose the data in March 2020 of wind station 1 (W1) in order to select k. Station W1 was built up on a flat area in April of 2018 and has 100 MW for total installed capacity, no energy storage capacity, 47 wind generation units.

### *B. Parameter selection*

Parameter k is the moving step size. For short-term forecasting there are 96 values every day so the range of k is from

1 to 95 mathematically. It is noted to say that RMSE can be regarded as a special case of SRMSE especially when k is 1.

As shown in Fig. 5, different values of k will have different effects on I-SRMSE. The blue broken line with legend k=1 is the original RMSE. As k increases from 1 to 33 the values of I-SRMSE start increasing particularly on March 12, 14, 15, 16, 18 and 19 and in the meanwhile stay steady on other days in March. But when k reaches up to 67 the values become distorted for the reason that the values of I-SRMSE surge abnormally on March 2, 4, 8, 10, 17, 21 and 30. Because the value of k is the moving step size. The appearance of abnormal surge in these days can be due to some lead-lag errors which are at a time distance more than 8 hours.

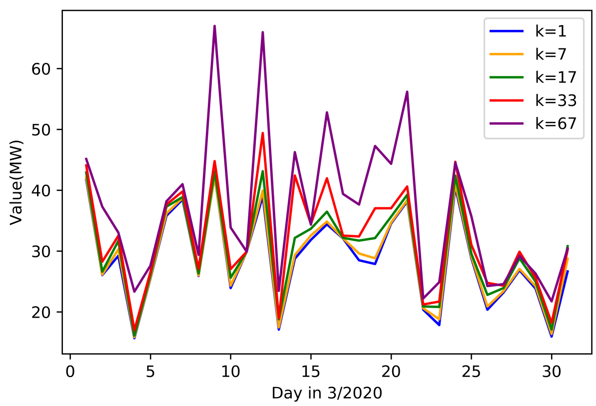


Fig. 5. Comparison of I-SRMSE with different k

As shown in Fig. 6, different values of k will have different effects on II-SRMSE as well. However, compared to I-SRMSE things have changed especially for small k. The differences between SRMSE and original RMSE are greater.

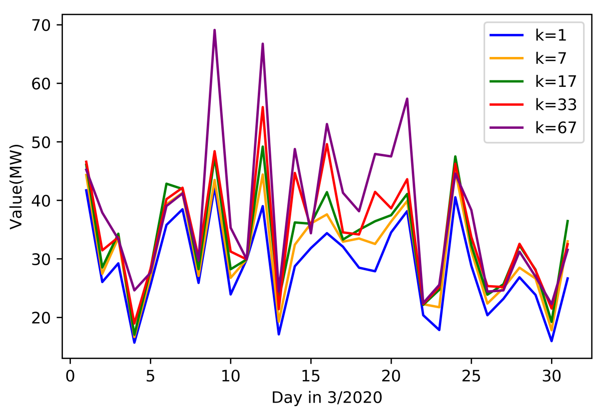


Fig. 6. Comparison of II-SRMSE with different k

As shown in Fig. 7, The differences between III-SRMSE and original RMSE are small even with a large k. This is because the direct amendment in error series is not obvious in the process of average and the larger k is the fewer points are in the moving average. Still, III-SRMSE can reflect lead-lag errors on march 14, 16 and 19.

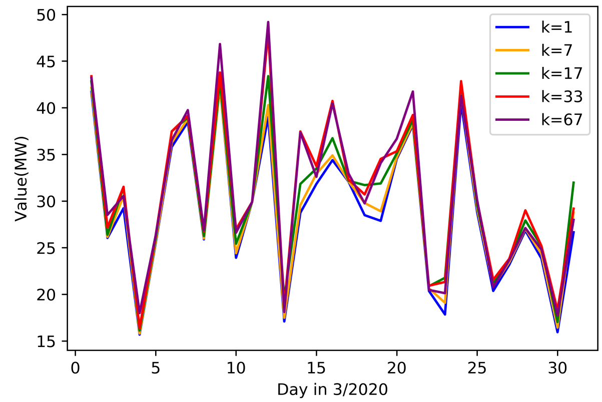


Fig. 7. Comparison of III-SRMSE with different k

We choose the value 7 as the appropriate k from a perspective that the adjustment speed of thermal units is limited and the adjustment capacity is small especially during the extreme peak-valley time which usually lasts for nearly 1.5 hours as shown in Fig. 8.

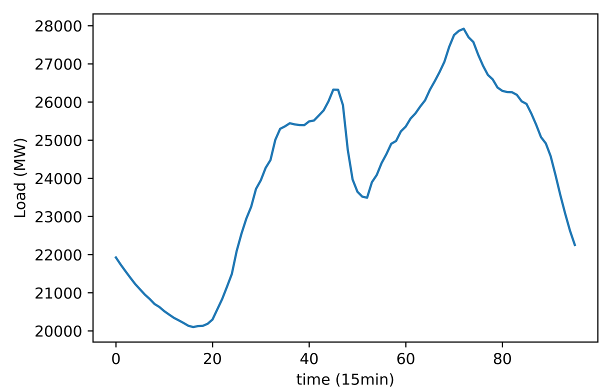


Fig. 8. The average load of Hebei Southern State Grid in 2020

After the selection of k is confirmed we make a comparison of different SRMSE in with the same k=7. II-SRMSE is the most suitable one for the reason that it distinguishes days more obviously.

## *C. Typical days study*

As shown in Fig. 6, in the most of March the differences between RMSE and SRMSE are relatively small. Specially in day 11 the difference is 0, and oppositely the large differences reflect the lead-lag errors. In Tab. II, the values of RMSE, SRMSE, difference and r are calculated in 11 March to 17 March.

The differences on day 11 and day 17 are relatively small (0 and 0.860), and the forecast results are shown in Fig. 9. The appearance of small differences reflects there are less lead-lag errors. 0 difference means the prediction is always greater than or less than the actual values. It is noted that lead-lag errors are few on 17 but r is 0.593 so the relationship between them is not that close.

TABLE II
COMPARISON BETWEEN RMSE AND SRMSE

| DAY | RMSE | SRMSE | difference | r |
|---|---|---|---|---|
| 11 | 29.908 | 29.908 | 0.000 | 0.849 |
| 12 | 39.006 | 44.412 | 5.406 | 0.096 |
| 13 | 17.097 | 19.262 | 2.165 | 0.834 |
| 14 | 28.739 | 32.419 | 3.679 | 0.508 |
| 15 | 31.838 | 36.043 | 4.205 | 0.795 |
| 16 | 34.385 | 37.598 | 3.212 | 0.459 |
| 17 | 32.073 | 32.934 | 0.860 | 0.593 |

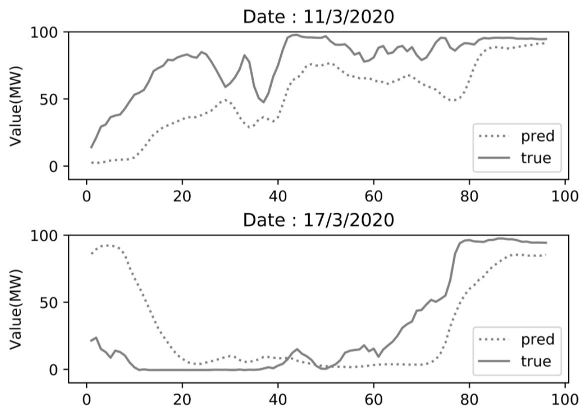


Fig. 9. Forecast result of day 11 and 17

The forecast results of day 14 are shown in Fig. 10. On this day the prediction made a wrong judgement on when did the wind process appear and fade, which will make SRMSE larger as well as exert a negative influence on electricity dispatch system. With this wrong judgement, at the noon of that day there would be a nearly 50MW vacancy of electricity supply and just 4 hours later there would be a nearly 30MW excess produced by this station. The negative influence is suggested to be checked, so SRMSE can solve the practical problem which RMSE cannot.

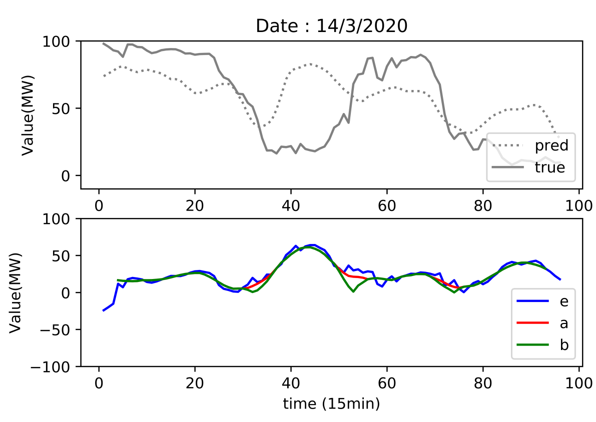


Fig. 10. Forecast result of day 14

*D. A comprehensive summary*

Data in Tab. III show the average value of the normalized RMSE and SRMSE (NRMSE and NSRMSE) and r of 5 power and 6 PV power generation stations from W1 to W5 and P1 to P6 in March 2020. The normalization is by dividing the corresponding capacity at the last step. The column 'Diff. (lift ratio)' displays the difference between two metrics and the percentage of increase in NSRMSE named as lift ratio, which can reflect the extent of lead-lag errors. The column 'unqualified' displays the percentage of that how many days are unqualified with r being less than 0.68 in the evaluation of Central China Region and 64.52% means 20 days have r less than 0.69 in March.

TABLE III
COMPARISON OF 11 POWER GENERATION STATIONS

| No. | NRMSE | NSRMSE | Diff. (lift ratio) | r | unqualified |
|---|---|---|---|---|---|
| W1 | 28.74% | 31.46% | 0.027 (9.47%) | 0.463 | 64.52% |
| W2 | 31.91% | 34.44% | 0.025 (7.93%) | 0.409 | 64.52% |
| W3 | 12.34% | 13.02% | 0.007 (5.52%) | 0.483 | 64.52% |
| W4 | 29.49% | 31.86% | 0.024 (8.06%) | 0.409 | 67.74% |
| W5 | 9.00% | 9.90% | 0.009(10.06%) | 0.388 | 67.74% |
| P1 | 7.11% | 7.93% | 0.008(11.50%) | 0.948 | 3.23% |
| P2 | 7.37% | 8.12% | 0.007(10.08%) | 0.918 | 6.45% |
| P3 | 6.47% | 7.12% | 0.006 (9.96%) | 0.963 | 3.23% |
| P4 | 11.08% | 12.09% | 0.010 (9.10%) | 0.939 | 3.23% |
| P5 | 8.19% | 9.04% | 0.009(10.44%) | 0.880 | 9.68% |
| P6 | 6.70% | 7.42% | 0.007(10.74%) | 0.958 | 0.00% |

PV power is more accurate and steady than wind power because their average values of NSRMSE are 24.14% and 8.62% and ones of r are 0.430 and 0.934. It is noted that values of the unqualified rate of wind power stations are closer than ones of PV power stations. Thus the unqualified rate of r¡0.68 is not suitable for wind power stations. The average values of lift ratio are 8.21% and 10.30% so the lead-lag errors are more obvious in PV power stations.

Among 5 wind power stations the prediction of station W5 can be the most accurate because the values of NRMSE and NWRMSE are smallest, however, the lift ratio of W5 is the greatest and r is smallest. This represents that station W5 has the largest lead-lag errors and it probably only activates few units, which the current evaluation metrics cannot reflect.

Oppositely P3 among 6 PV power stations is of top performance with the smallest values of NRMSE and NWRMSE, a small lift ratio and the biggest r.

## V. CONCLUSION

Different metrics are needed to represent the performance of power prediction and the influences on electrical power system as well as guide the operation of the power system. Lead-lag and common errors both have negative effects on the flexibility and reliability of the system but the current metrics can hardly reflect the lead-lag errors. Proceeding from the actual needs of the power system, this paper proposes a new metric for the evaluation of wind power forecasting named as sign root mean square error (SRMSE). Moving average method is used two times to form two series which detect the signs of errors differently. Then use the statistical properties of new series to amend the original value of RMSE.

The proposed metric can reflect both lead-lag scenarios and common scenarios as well as it inherit good characteristics from RMSE like objectivity, adaptability, unity, symmetry and stability. The new metric can better evaluate the performance of the predictions and guide the operation of the power system but it has shortcomings that it is not suitable when the loss is proportional to the error, and it costs more time in the usage of training the neuronal network, and it can be only used for point forecasting evaluation. The case study demonstrates that the new metric has the ability to reflect both the lead-lag scenarios and the common scenarios. The proposed new evaluation metric named SRMSE has its practical significance.

## Acknowledgment

This work is supported by the National Key R&D Program of China (Technology and application of wind power / photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266).

## References

[1] K. Doubleday, S. Jascourt, W. Kleiber, and B. M. Hodge, “Probabilistic solar power forecasting using bayesian model averaging,” *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 325–337, 2021.

[2] S. R. Sinsel, R. L. Riemke, and V. H. Hoffmann, “Challenges and solution technologies for the integration of variable renewable energy sources—a review,” *Renew. Energy*, vol. 145. Elsevier Ltd, pp. 2271–2285, Jan. 01, 2020.

[3] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, “Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy,” *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, March 2017.

[4] Y. Jia, X. Lyu, P. Xie, Z. Xu, and M. Chen, “A Novel Retrospect-Inspired Regime for Microgrid Real-Time Energy Scheduling with Heterogeneous Sources,” *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4614–4625, 2020.

[5] E. Kabir, P. Kumar, S. Kumar, A. A. Adelodun, and K. H. Kim, “Solar energy: Potential and future prospects,” *Renew. Sustain. Energy Rev.*, vol. 82, pp. 894–900, 2018, doi: 10.1016/j.rser.2017.09.094.

[6] Z. Zhen et al., “Pattern Classification and PSO Optimal Weights Based Sky Images Cloud Motion Speed Calculation Method for Solar PV Power Forecasting,” IEEE Trans. Ind. Appl., vol. 55, no. 4, pp. 3331–3342, 2019, doi: 10.1109/TIA.2019.2904927.

[7] J. M. González-Sopeña, V. Pakrashi, and B. Ghosh, “An overview of performance evaluation metrics for short-term statistical wind power forecasting,” *Renew. Sustain. Energy Rev.*, no. September, p. 110515, 2020.

[8] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, “Current methods and advances in forecasting of wind power generation,” *Renew. Energy*, vol. 37, no. 1, pp. 1–8, 2012.

[9] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006.

[10] Y. K. Wu, P. E. Su, T. Y. Wu, J. S. Hong, and M. Y. Hassan, “Probabilistic Wind-Power Forecasting Using Weather Ensemble Models,” *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5609–5620, 2018.

[11] Y. K. Wu, P. E. Su, and J. S. Hong, “Stratification-Based Wind Power Forecasting in a High-Penetration Wind Power System Using a Hybrid Model,” *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 2016–2030, 2016.

[12] K. Li, P. Zhang, G. Li, F. Wang, Z. Mi, and H. Chen, “Day-Ahead Optimal Joint Scheduling Model of Electric and Natural Gas Appliances for Home Integrated Energy Management,” *IEEE Access*, vol. 7, pp. 133628–133640, 2019.

[13] S. Chai, Z. Xu, Y. Jia, and W. K. Wong, “A Robust Spatiotemporal Forecasting Framework for Photovoltaic Generation,” *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5370–5382, 2020.

[14] J. Ma, M. Yang, and Y. Lin, “Ultra-short-term probabilistic wind turbine power forecast based on empirical dynamic modeling,” *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 906–915, 2020, doi: 10.1109/TSTE.2019.2912270.

[15] Y. Yu, X. Han, M. Yang, and J. Yang, “Probabilistic Prediction of Regional Wind Power Based on Spatiotemporal Quantile Regression,” *IEEE Trans. Ind. Appl.*, vol. 56, no. 6, pp. 6117–6127, 2020, doi: 10.1109/TIA.2020.2992945.

[16] X. Zhao, S. Wang, and T. Li, “Review of evaluation criteria and main methods of wind power forecasting,” *Energy Procedia*, vol. 12, pp. 761–769, 2011.

[17] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, “A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework,” *Energy Convers. Manag.*, vol. 212, no. April, p. 112766, 2020.

[18] F. Wang et al., “A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model,”*Energy Convers. Manag.*, vol. 220, no. June, 2020.

[19] M. Yang, S. Zhu, M. Liu, and W. J. Lee, “One parametric approach for short-term JPDF forecast of wind generation,”*IEEE Trans. Ind. Appl.*, vol. 50, no. 4, pp. 2837–2843, July-Aug.2014.

[20] M. B. Ozkan and P. Karagoz, “A novel wind power forecast model: Statistical hybrid wind power forecast technique (SHWIP),” *IEEE Trans. Ind. Informatics*, vol. 11, no. 2, pp. 375–387, Apr. 2015.

[21] H. Chen, Q. Wan, and Y. Wang, “Refined diebold-mariano test methods for the evaluation of wind power forecasting models,” *Energies*, vol. 7, no. 7, pp. 4185–4198, 2014.

[22] R. J. Bessa, V. Miranda, and J. Gama, “Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting,” *IEEE Trans. Power Syst.*, 2009.

[23] Q. Yao, Y. Liu, W. Zhang, and Y. Hu, “Study on the comprehensive evaluation method of regional wind power prediction,” *Chinese Control Conf. CCC*, pp. 10259–10264, 2017.

[24] Y. Qi, L. Yu, B. Kai, S. Rongfu, S. Peng, and W. Yuhui, “A evaluation method of wind farm power prediction based on principal component analysis and entropy methods,” *in Proceedings of the 28th Chinese Control and Decision Conference, CCDC 2016*, 2016, pp. 2132–2136.

[25] R. Ma, L. L. Wang, and S. J. Hu, “Short-term wind power prediction and error analysis,” *Appl. Mech. Mater.*, vol. 448–453, pp. 1851–1857, 2014.

[26] H. C. Xu, W. X. Zhou, and D. Sornette, “Time-dependent lead-lag relationship between the onshore and offshore Renminbi exchange rates,” *J. Int. Financ. Mark. Institutions Money*, vol. 49, pp. 173–183, Jul. 2017.

[27] N. Ghadimi, “Genetically tuning of lead-lag controller in order to control of fuel cell voltage,” Sci. Res. Essays, vol. 7, no. 43, pp. 3695–3701, 2012.

[28] W. Cui et al., “Evaluation and sequential dispatch of operating reserve provided by air conditioners considering lead-lag rebound effect,” *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6935–6950, 2018.

[29] J. W. Messner, P. Pinson, J. Browell, M. B. Bjerregard, and I. Schicker, “Evaluation of wind power forecasts-An up-to-date view,” *Wind Energy*, vol. 23, no. 6, pp. 1461–1481, 2020.

[30] S. Chen, J. Yang, Y. Liu, K. Tang, and P. Song, “Evaluation and Improvement of Power Forecasting Indicators for Renewable Energy Stations,” *in 2019 IEEE 3rd International Electrical and Energy Conference (CIEEC)*, Sep. 2019, pp. 1164–1169.