ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20231019001038
Original scientific paper

A Wind Power Forecasting Model Incorporating Recursive Bayesian Filtering State Estimation and Time-Series Data Mining

Peng LIU*, Tieyan ZHANG, Furui TIAN, Yun TENG, Chuang GU

Abstract: To enhance the precision of wind power forecasting and the integration of renewable energy, a wind power prediction model, synthesising recursive Bayesian filtering state estimation with time-series data mining, was developed. Initially, the Autoregressive Integrated Moving Average Model (ARIMA)-Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH) model was utilised for mining historical wind power data and establishing a model. Subsequently, the double-parameter t-distribution was employed to fit the prior estimation error and observation error, which integrated observational information with prior estimates through a sophisticated recursive Bayesian filtering approach, culminating in the formulation of a robust predictive model. Validation of this model was conducted using a diverse dataset, encompassing wind farms with varying capacities and distinct time intervals. Simulation outcomes reveal that this model's forecasting accuracy markedly surpasses that of conventional methodologies. Notably, an enhanced predictive precision is observed in wind farms with larger capacities, particularly when shorter intervals of observational data are employed. This model demonstrates significant potential for advancing the accuracy and efficiency of wind power forecasting, a critical element in the optimization of renewable energy utilization.

Keywords: data mining; information fusion; probability distribution fitting; recursive bayesian filtering; wind power forecasting

1 INTRODUCTION

Enduring attention has been afforded globally to energy and environmental issues. The national "Fourteenth Five-Year" plan emphasizes the intensification of non-fossil energy development, the establishment of novel power systems, enhancement of grid intelligence, and promotion of grid adaptability to extensive, centralized new energy [1]. Within the new energy system, wind power encompasses the largest proportion of installed capacity, with forecasts predicting wind power generation to reach 234.5 billion kWh at the end of 2023, of which the incremental wind power generation will achieve 109.3 billion kWh, accounting for 46.8% of the total wind power generation in 2023. With the escalating proportion of wind power in energy systems, its inherent variability and unpredictability increasingly challenge the stability of these systems. Given the strategic emphasis on national energy policies and the escalating integration of wind power, the precision of wind power generation forecasts emerges as a critical factor [2-4]. In the realm of wind power forecasting methodologies, two primary categories have been identified: statistical and physical approaches. The physical approach, predicated on the wind turbines' inherent characteristics, transforms meteorological data into power generation estimates based on turbine parameters. Conversely, the statistical approach, dependent on historical data analysis, extracts correlations to predict power generation. The accuracy of physical methods, not reliant on extensive historical data, hinges significantly on the model's precision and parameter selection. This reliance introduces limitations to the predictive accuracy, particularly in wind farms situated in regions characterized by complex physical phenomena, which pose challenges for precise modeling. In contrast, statistical methods analyze patterns in historical data to formulate non-linear relationships, thus bypassing the limitations of inaccurate physical descriptions. These methods are predominantly applied in short-term and ultra-short-term forecasting, incorporating techniques such as time-series extrapolation and artificial intelligence-based prediction methods [5, 6]. Reference [7] used time series analysis method; an ARIMA-GARCH model was developed. However, this model exhibits restricted adaptability to dynamic changes and inadequately captures the clustering of wind power data fluctuations over time, thereby lacking the necessary flexibility and timeliness in adjusting to variations in wind power output. Studies [8-12] on the distribution characteristics of wind power and wind speed, by deriving the distribution characteristics satisfied by wind speed, subsequently determine the prediction error of wind power output. Such methods, which obtain a statistical model of wind power output, can describe wind speed characteristics but cannot directly and effectively inform judgements in research on wind power uncertainty. Literature [13, 14] proposed an algorithm combining time series analysis and Kalman filtering (FKE) to predict measured wind speed signals. However, the FKE corresponding system should satisfy homogeneity and superposition, and wind farm output signals belong to irregular nonlinear systems. In the domain of wind power forecasting, the integration of Kalman filtering with time series analysis, while innovative, encounters inherent limitations due to linearization. This aspect was observed when direct forecasting methods were applied to wind power, highlighting a crucial area for improvement. The utilization of Long Short-Term Memory (LSTM) networks for multivariate wind power forecasting [15] marks a significant advancement. However, the efficacy of LSTM networks is somewhat diminished due to their struggles with capturing seasonality and periodicity. These networks exhibit heightened sensitivity to anomalies and face substantial challenges in hyperparameter tuning, which compromises their suitability for practical applications and diminishes their timeliness. Further research led to the application of Transformer networks in wind power forecasting [16]. Notably, the computational complexity of these networks escalates with increased sequence lengths, raising concerns over computational resource constraints in processing extended datasets. Additionally, the complexity associated with the excessive number of model

parameters in Transformer networks impedes the achievement of the desired timeliness in wind power forecasting. To address these challenges, this study leverages the ARIMA-FIGARCH model for data mining. This model adeptly captures the long-term correlations present in time series data without burdening computational resources, particularly in the estimation of prior results. The deployment of a recursive Bayesian state estimation framework significantly enhances the timeliness of forecasts. This framework proficiently facilitates the integration of observational data with prior estimates, which is contingent upon representing both prior estimation errors and observational errors within a unified dimension. Adopting a double-parameter t-distribution for probability fitting, this research effectively aligns prior estimation errors and observational errors in a singular dimensional framework. Such an approach enables continuous and dynamic fusion and updating of observational data with prior estimates, culminating in a substantial improvement in the timeliness and accuracy of predictive modeling.

2 ESTABLISHING THE ARIMA-FIGARCH MODEL

Wind power output showcases temporality, volatility, uncertainty, and heteroscedasticity. To capture the persistent volatility inherent in the data, modelling through ARIMA-FIGARCH is undertaken. The Autoregressive Integrated Moving Average Model (ARIMA) is utilized to seize the trend and periodicity of the data. Meanwhile, the Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH) model aims to capture persistent volatility within the time series. A mathematical model, constituted through the conjoint application of ARIMA and FIGARCH, is apt to describe the characteristics of wind power time series with heightened precision [17-22].

2.1 Establishment of the ARIMA Model

The ARIMA ($p\ d\ q$) consists of an AR (Autoregressive Process) of order p, an MA (Moving Average Process) of order q, and d as the order of differencing [23-31]. The expression can be written as:

$$\left(1-\sum_{i=1}^{p}\alpha_i L^i\right)(1-L)^d\ y_t = \alpha_0 + \left(1+\sum_{j=1}^{q}\beta_i L^i\right)\varepsilon_t \tag{1}$$

where, y_t signifies wind power time series, α represents the coefficients of autoregressive terms, β signifies the coefficients of moving average terms, and $\mathcal{E}_t$ stands for the residual terms.

Prior to modelling, the time series undergoes the Augmented Dickey-Fuller Test (ADF). The examination of time series stationarity forms the core of this analysis, utilizing a statistical method designed to ascertain the presence of a unit root within the time series data. The ADF test, serving this purpose, is structured around two hypotheses. The null hypothesis (H0) posits the existence of a unit root, indicative of non-stationarity within the series. In contrast, the alternative hypothesis (H1) asserts the absence of a unit root, suggesting stationarity. The calculation of the ADF statistic, integral to this test, is derived from the ADF model. Concurrently, the critical value is ascertained, contingent upon the chosen level of significance. Should this critical value be surpassed by the absolute value of the ADF statistic, H0 is consequently rejected, leading to the classification of the series as stationary. Conversely, failure to surpass this threshold confirms the series' non-stationary nature. Crucially, the determination of the need for differencing, aimed at smoothing the series, hinges on the identification of the number of unit roots. This aspect is pivotal in informing the selection of the model parameter d, thus shaping the analytical framework of the study. Values for parameters p and q are determined by solving the AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF), respectively, with these lag orders representing temporal distance between the current and past time points. The ACF of the AR(p) model demonstrates a tailing characteristic, while its PACF truncates after order p. For the MA(q) model, the ACF truncates post order q, and its PACF exhibits a tailing characteristic. The autocorrelation coefficient's expression is delineated as Eq. (2):

$$\gamma_s = \overline{\rho}_s = \sum_{t=s+1}^{T}\left(y_t-\overline{y}\right)\left(y_{t-s}-\overline{y}\right)\Big/\sum_{t=1}^{T}\left(y_t-\overline{y}\right)^2 \tag{2}$$

where, γ_s signifies the autocorrelation coefficient, and $\overline{\rho}_s$ represents its estimated value, derived from the sample data, with s denoting the interval between sequences.

Eq. (3) articulates the expression for the partial autocorrelation coefficient:

$$y_t = \phi_{s1}y_{t-1} + \phi_{s2}y_{t-2} + \cdots + \phi_{ss}y_{t-s} + \cdots \tag{3}$$

where, ϕ_{ss} is indicative of the partial autocorrelation coefficient between y_{t-s} and y_t, encapsulating the partial autocorrelation for a span of s periods.

The determination of the model's p and q orders is conducted through an examination of ACF and PACF. This process is followed by the construction of the model's likelihood function, which quantifies the probability of data occurrence under specific model parameters, typically predicated on the assumption of independent distribution of observational data. The maximization of the log-likelihood function facilitates the estimation of model parameters. Utilizing the Maximum Likelihood Estimation (MLE) method, the model parameters that optimally elucidate the observational data are ascertained, thereby maximizing the data's probability of occurrence within the stipulated model framework. Optimal parameters are subsequently identified through the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The model's completeness is evaluated by the Ljung-Box Q test, which ascertains whether residuals constitute a white noise sequence. The H1 hypothesis of this test postulates that all $\overline{\rho}_s$ are null, implying that the residuals form a white noise sequence. In contrast, the H0 hypothesis contends that at least one $\overline{\rho}_s$ is non-zero, suggesting the residuals do not form a white noise sequence. The Q statistic is computed based on the sample autocorrelation coefficients, adhering to a chi-squared distribution with

($s - n$) degrees of freedom. The specific formula for the Q statistic is outlined in Eq. (4):

$$Q = N(N+2)\sum_{k=1}^{s}\frac{\gamma_k^2}{N-k} \sim \chi_{s-n}^2 \tag{4}$$

where, N denotes the number of samples, n represents the count of unknown parameters within the model, and s is typically selected as 8, 14, or 24, contingent upon the actual size of the sample.

2.2 Establishment of the FIGARCH Model

The FIGARCH model is expressed as:

$$\begin{cases} y_t = E\left(y_t \mid \psi_{t-1}\right) + \varepsilon_t \\ \varepsilon_t = \upsilon_t h_t^2 \\ h_t^2 = \alpha_0 + \beta(L)h_t + \left[1 - \beta(L) - \alpha(L)(1-L)^d\right]\varepsilon_t^2 \\ \alpha(L) = 1 - \sum_{i=1}^{q}\alpha_i L^i \\ \beta(L) = \sum_{j=1}^{p}\beta_j L^j \\ (1-L)^d = 1 - dL - \frac{1}{2!}d(1-d)L^2 - \\ \qquad -\frac{1}{3!}d(1-d)(2-d)L^3 \ldots \end{cases} \tag{5}$$

where, L denotes the lag operator, and d represents the order of integration $(0 \le d \le 1)$. Given that $(0 < d < 1)$, $(1-L)^d \varepsilon_t^2$ is modelled as an ARMA process, where $\alpha(L)$ and $\beta(L)$ coefficients can be adapted to capture the memory of data volatility [24]. The magnitude of d can reflect the duration of volatility memory.

Modelling in the FIGARCH model is focused on residuals, initially subjecting the squared residuals of the ARIMA model to the Ljung-Box Q test. Should it not represent white noise, the Lagrange Multiplier Test (LM) is required, applying a moving average to the squared residual terms, determining the fit quality (R-squared, R^2) of the approximating function, and constructing a statistic $L = TR^2 \sim \chi^2$ to test for the presence of ARCH effects. Ultimately, the order of the FIGARCH is ascertained using ACF, PACF, AIC, and BIC, with coefficients of the FIGARCH model estimated through MLE.

3 ESTABLISHMENT OF A PREDICTIVE MODEL BASED ON THE RECURSIVE BAYESIAN FILTERING FRAMEWORK

Recursive Bayesian filtering, anchored in the Bayesian filtering framework [15], incorporates a matrix composed of data points at every timestep in the series model as its state variables. Observations within the model are defined as actual observation value of the leading time interval Δt of the predicted point. Fluctuations between model observation values and the true values of the forecast points adhere to a certain probability distribution, exhibiting variations as the time interval shifts. In this work, a parameterised t distribution is employed to characterise observational noise [10]. The adoption of the double-parameter t-distribution is emphasized, which, in contrast to the conventional normal distribution, demonstrates enhanced robustness and a superior capacity for accommodating outliers. The incorporation of variable double parameters in this distribution introduces a notable increase in flexibility for fitting, while also providing a more accurate representation of uncertainty. It is observed that when the degrees of freedom within the *t*-distribution are constrained, the distribution exhibits a tendency towards a broader scope. This attribute is particularly beneficial in more accurately depicting the uncertainty associated with both prior estimation errors and observational errors. Such precise quantification lays the groundwork for the effective integration of these two error types, a critical component in the realm of predictive analytics. As a result, this approach facilitates the implementation of predictions and updates within the Bayesian framework, enhancing the model's capacity to incorporate and adjust to varying data inputs.

3.1 *t*-distribution with Location and Scale Parameters

To suppress noise, detect trends, smooth data, and enhance data stability, it is essential to implement smoothing processing on wind power time-series data, for which the moving average formula is expressed as:

$$Y_{MA} = \frac{\begin{pmatrix} Y_{t(T/2)+1} + Y_{t-(T/2)+2} \ldots + Y_t + \\ + \ldots + Y_{t-(T/2)-1} + Y_{t-(T/2)} \end{pmatrix}}{T} \tag{6}$$

where, T symbolizes the moving average duration, Y_t represents the average power at time t, and Y_{MA} designates the moving average power with a time window of T.

Subsequent steps involve fitting the probability distribution to the wind power time-series data at varying time intervals, thereby uncovering the volatility pattern of wind power under different time intervals. The probability density function expression is provided in Eq. (7):

$$\begin{cases} f\left(x \mid \mu, \sigma, \upsilon\right) = \dfrac{1}{\sigma\sqrt{2\pi}B\left(\dfrac{\upsilon}{2}, \dfrac{1}{2}\right)}\left(1 + \dfrac{(x-\mu)^2}{\sigma^2\upsilon}\right)^{-\frac{\upsilon+1}{2}} \\ B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t \end{cases} \tag{7}$$

where, υ signifies the degrees of freedom, characterizing the function shape; μ serves as the location parameter, and σ is the scale parameter.

Fig. 1 illustrates the curve graphs when the degrees of freedom (d*f*) are 5, 10, and 30, respectively. Lower degrees of freedom yield a more pronounced fat-tail characteristic of the curve, whereas hgher degrees make the *t*-distribution increasingly approximate the normal distribution.

Parameter μ plays a pivotal role in determining the left-right positioning of the distribution function. Concurrently, the parameter σ is instrumental in dictating the distribution's tendency towards a broader or narrower configuration.

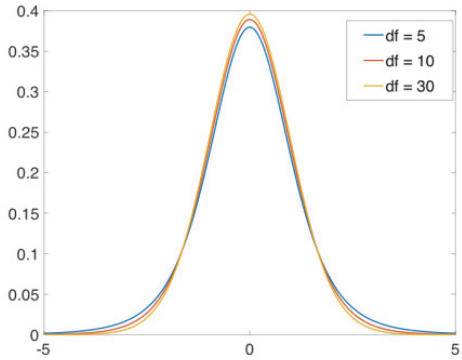


Figure 1 Comparison of t-distributions with different parameters

3.2 Establishment of a Forecasting Model Based on the Recursive Bayesian Filtering Framework

Bayesian filtering utilises prior estimation information and observational data to compute the probability distribution of posterior estimations, thus calculating the optimal estimation of the system's current state [15].

$$P\left(Y_t \mid Z_t\right)=\frac{P\left(Z_t \mid Y_t\right)P\left(Y_t^- \mid Z_{t-1}\right)}{\int P\left(Z_t \mid Y_t\right)P\left(Y_t^- \mid Z_{t-1}\right)\mathrm{d}Y_t} \tag{8}$$

$$P\left(Y_t^- \mid Z_{t-1}\right)=\int P\left(Y_t \mid Y_{t-1}\right)P\left(Y_{t-1} \mid Z_{t-1}\right)\mathrm{d}Y_{t-1} \tag{9}$$

where, Y represents the wind power time series matrix, and Z denotes the observational data matrix.

In practical computations, the time series mathematical model is constituted by nonlinear equations, and both the system state variables and observational variables encompass multidimensional data, rendering Eq. (8) and Eq. (9) unsolvable. Consequently, an assumption that process noise, observational noise, and system state variables adhere to a Gaussian-Laplace distribution is necessary. This allows for an approximate calculation of Eq. (8) and Eq. (9), which, when incorporated into the time series mathematical model, yields Eq. (10) and Eq. (11) as the forecasting model.

$$\begin{cases}\hat{Y}_t^- = \partial\left(y_{t-1}, y_{t-2}, \ldots, y_s, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_s\right)+e_{Y_t} \\ P_{Y_t}^- = \left[\sigma_y \mid f\left(e_{Y_t} \mid \mu_y, \sigma_y, \upsilon_y\right)\right]\end{cases} \tag{10}$$

$$\begin{cases}P_{Z_t} = \left[\sigma_z \mid f\left(Z_t \mid \mu_z, \sigma_z, \upsilon_z\right)\right] \\ K_t = P_{Y_t}^- C_t^{\top}\left(C_t P_{Y_t}^- C_t^{\top}+P_{Z_t}\right)^{-1} \\ \hat{Y}_t = \hat{Y}_t^- + K_t\left(Z_t - \hat{Y}_t^-\right) \\ P_{Y_t} = P_{Y_t}^- - K_t C_t P_{Y_t}^-\end{cases} \tag{11}$$

where, ∂ symbolises the time series mathematical model; $\left[y_{t-1}, y_{t-2}, \ldots, y_s\right]$ represents the wind power data set within window period s, $P_{\hat{Y}_t^-}$ denotes prior estimation error, P_{Z_t} indicates observational error, K_t defines the Bayesian ratio factor, C_t represents the conversion coefficient matrix of Z and Y, $\hat{Y}_t$ illustrates the posterior estimation, and P_{Y_t} signifies posterior estimation error.

3.3 Recursive Process of the Tracking Model

The resolution process is executed in three steps: forecasting, updating, and recursion. Eq. (10) signifies the forecasting step of the model, wherein the time series model and residual distribution fitting are employed to facilitate the prediction of the state distribution at the subsequent time step. Eq. (11) delineates the updating step of the model, during which observational data are integrated and corrected with the predicted state distribution, and the probability distribution of the posterior estimation is computed. Subsequently, the posterior estimation distribution is utilized as the prior estimation distribution for the next time step, after which the aforementioned process is repeated, recursively updating the state estimation. Fig. 2 illustrates the updating process, and Fig. 3 provides a flowchart of the algorithmic recursive process.

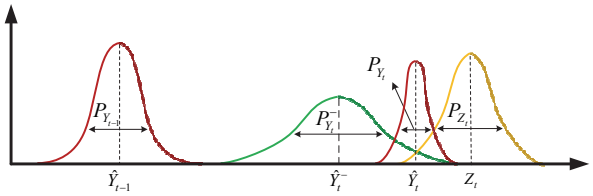


Figure 2 Schematic of state estimation

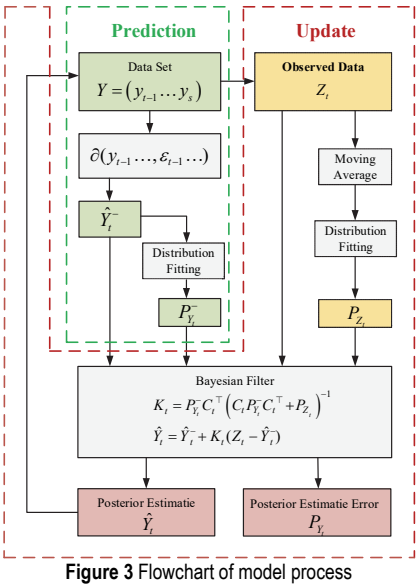


Figure 3 Flowchart of model process

4 CASE STUDY

In a scenario considering two wind farms in Northeast China with installed capacities of 2500 MW and 800 MW

respectively, wind power data from January 1st to 10th, 2023, was utilised as the training set for establishing the prediction model. A comparison between the actual data and predicted data from 00:00 to 02:00 on the 11th was conducted to evaluate the predictive model's accuracy.

4.1 Construction of the ARIMA-FIGARCH Model

An initial application of the ADF test was employed to assess time-series stationarity. Fig. 4 presents the wind power output time series y_t. In the graph, the horizontal axis is designated for time points, with data being systematically collected at 15 minute intervals, encapsulating a full day's dataset. Correspondingly, the vertical axis is utilized to depict the output of wind power. Tab. 1 exhibits the results of the ADF test. With a *P*-value of 0.0146 and exceeding 0.01, the null hypothesis cannot be rejected, implying a non-stationary time series y_t that necessitates smoothing treatment. Empirical evidence, accumulated through extensive experience in this field, indicates that time series data of this nature generally achieve stability post-application of first-order differencing. Fig. 5 illustrates the time series of $y_t^{'}$ after first-order differencing. Subjecting $y_t^{'}$ to the ADF test yielded the results shown in Tab. 2; a P-value of 0.000, which is less than 0.01, thus enabling the rejection of the null hypothesis, deeming the sequence stationary. It is deduced that the parameter *d* should be assigned a value of 1.

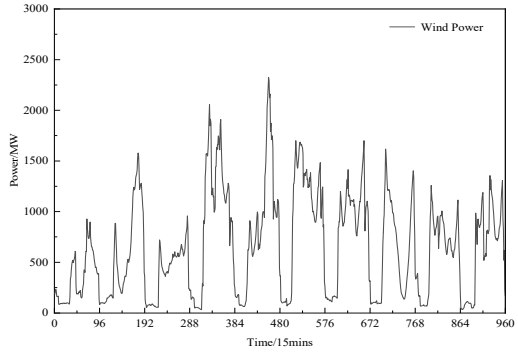


Figure 4 Wind power time series (y_t)

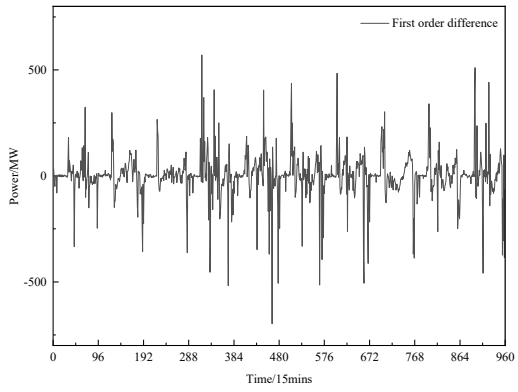


Figure 5 Time series of first-order difference of wind power(y_t')

Table 1 ADF test results of y_t

	Test statistic	Dickey-Fuller critical value		
		1%	5%	10%
$Z(t)$	–3.306	–3.430	–2.860	–2.570
MacKinnon approximate *p*-value for $Z(t)$ = 0.0146				

Table 2 ADF test results of y_t'

	Test statistic	Dickey-Fuller critical value		
		1%	5%	10%
$Z(t)$	–22.402	–3.430	–2.860	–2.570
MacKinnon approximate *p*-value for $Z(t)$ = 0.0000				

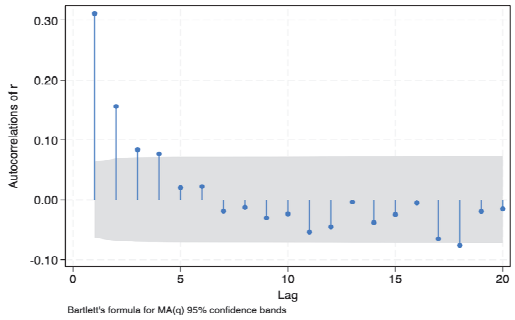


Figure 6 ACF plot of time series

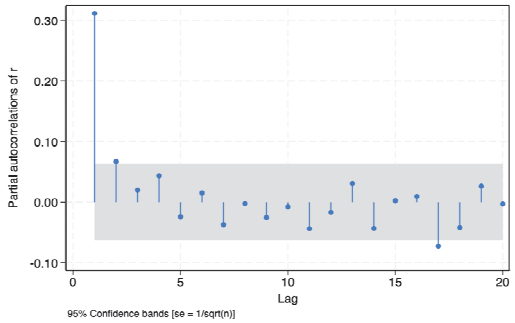


Figure 7 PACF plot of time series

The construction of ACF and PACF for $y_t^{'}$ was undertaken, with Fig. 6 showing the ACF plot and Fig. 7 displaying the PACF plot. The determination of the parameters *p* and *q* is achieved through an analysis of the ACF graph, which is instrumental in reflecting the time series' autocorrelation. It is observed that when the autocorrelation coefficients in the ACF graph truncate at the second order, specifically extending beyond the zero autocorrelation interval at lags 1 and 2, it implies that the model parameter *q* is optimally set to either 1 or 2. Similarly, in the PACF analysis, truncation of the PACF coefficients at the first order, particularly beyond the zero autocorrelation interval at lag 1, indicates an appropriate setting of the model parameter *p* to 1. Hence, ARIMA model coefficients ARIMA(1, 1, 1) and ARIMA(1, 1, 2) may be selected. Utilizing the MLE method for model coefficient determination, and subsequently computing each model's AIC and BIC, presented statistics as per Tab. 3. The outcomes reveal that ARIMA(1, 1, 1) renders the most optimal results. The LjungBox *Q* test was applied to fitted residuals and squared residuals for white noise testing, and *Q*-statistics alongside *P*-value are depicted in Tab. 4.

Table 3 ADF test results of time series of y_t

Model	AIC	BIC
ARIMA(1 1 1)	11584.55	11604.02
ARIMA(1 1 2)	11586.32	11610.65

Table 4 Ljung Box Q test results

	Portmanteau (*Q*) statistic	Prob > chi2(12)
Residual term	6.3474	0.8976
The squared term of the residual	49.4805	0.0000

As indicated in the table, with a *P*-value of 0.8976 for the residual sequence, the null hypothesis cannot be rejected, signifying the residual sequence as a white noise sequence. However, with a *P*-value of 0.0000 for the squared residuals *Q*-test, the null hypothesis is rejected, indicating that the squared residuals are not white noise, necessitating further LM testing. Regression of squared residuals against their lagged values was performed. Given that four points are sampled per hour in this series, regression against the fourth lag of the squared terms was undertaken. The test result of *P*-value = 7.547e–7, which is less than 0.01, allows the rejection of the null hypothesis, thereby concluding that the squared residual sequence exhibits clustered volatility. Similarly, ACF and PACF were employed to estimate the chosen FIGARCH, and through AIC and BIC, an appropriate model was selected, yielding model coefficient FIGARCH(1, 0, 1). Lastly, parameters of the system state equation ARIMA(1, 1, 1), FIGARCH(1, 0, 1) based on historical wind power output data were derived through MLE, as detailed in the Tab. 5.

Table 5 Parameters of ARIMA (1,1,1), FIGARCH (1,0,1) model

	coefficient	OPG std.err.	*z*	$P>\|z\|$	95% conf. interval	
cons	0.2935506	1.397733	0.21	0.834	–2.445955	3.033056
ar L1.	0.6726433	0.0332357	20.24	0.000	0.6075025	0.7377842
ma L1.	–0.2518709	0.0579689	–4.34	0.000	–0.3654879	–0.1382538
arch L1.	1.524278	0.1631342	9.34	0.000	1.204541	1.844015
garch L1.	0.3170058	0.0250157	12.67	0.000	0.267976	0.3660357
cons	73.24259	15.17899	4.83	0.000	43.49231	102.9929

4.2 Probability Distribution Fitting for Prior Estimation and Observational Errors

Fig. 8 displays the results of a moving average of wind power data with a time interval of 15 minutes from a 2500 MW wind farm. Probability distribution fittings were conducted separately for the a priori estimation residuals and the wind power data volatility of the two wind farms at different time intervals. The fitting graphs and parameters are shown in Fig. 9 and Tab. 6, respectively.

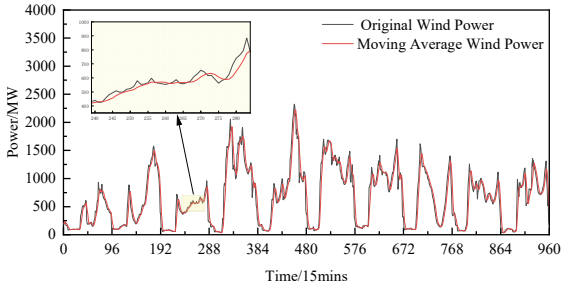
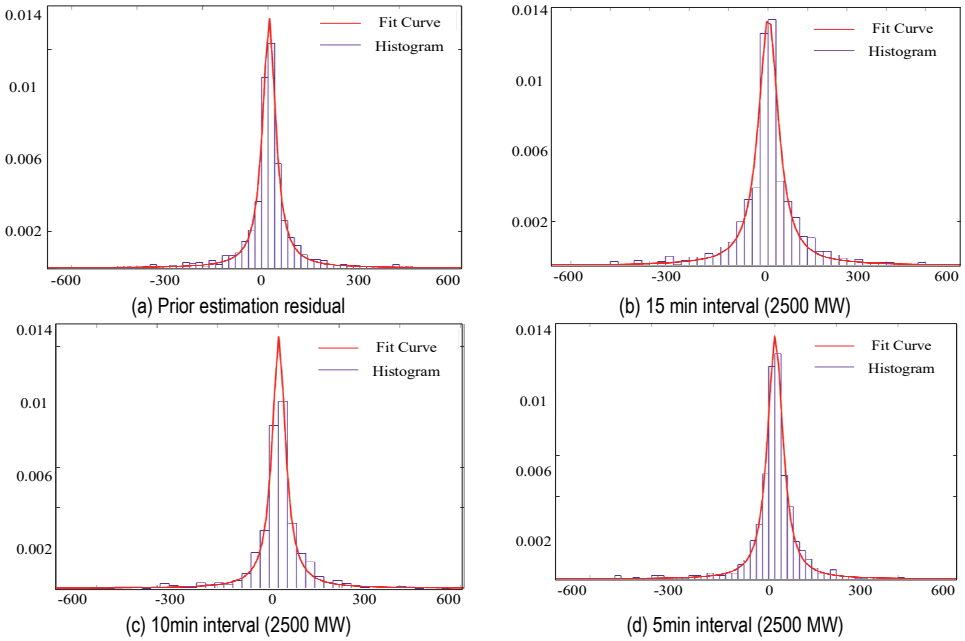


Figure 8 Post-averaging wind power processing data



(a) Prior estimation residual

(b) 15 min interval (2500 MW)

(c) 10min interval (2500 MW)

(d) 5min interval (2500 MW)

(e) 5 min interval (800 MW)

Figure 9 A priori estimation residuals and probability distribution fitting graphs of wind power data at different time intervals and capacities

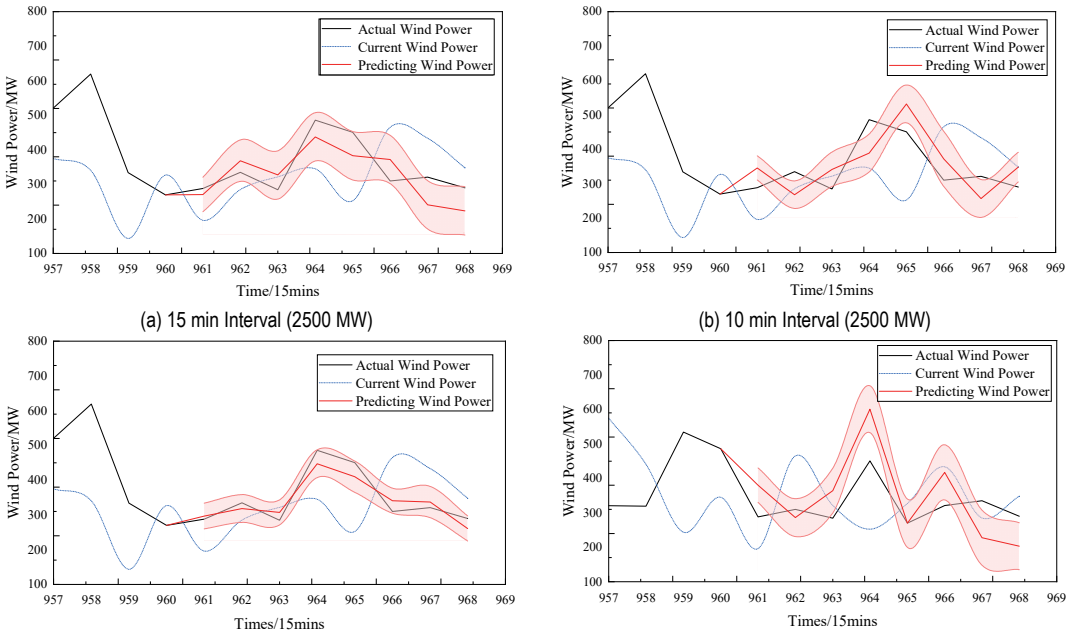
According to Tab. 6, it was discerned that for wind farms with the same installed capacity but different time intervals in their data, smaller intervals correlate with smaller data fluctuations. When time intervals are the same, larger installed capacities exhibit lower data volatility.

Table 6 Fitting parameters for system state equation residuals and wind power data at different time intervals

	μ Estimate	μ Standard Error	σ Estimate	σ Standard Error	v Estimate	v Standard Error
Residual	1.2227	1.0053	25.3765	1.4958	1.0576	0.0709
15 min(2500 MW)	1.2074	1.0489	33.2881	1.4719	1.3382	0.0887
10 min(2500 MW)	0.9456	1.2008	30.0428	2.0439	1.2646	0.1056
5 min(2500 MW)	0.7215	0.9852	16.8545	1.8837	1.0715	0.0807
5 min(800 MW)	1.2075	1.0489	35.2882	1.4719	1.3382	0.0887

4.3 Model Prediction Analysis

Predictions were made for the output of the wind farm from 00:00 to 02:00 on January 11th, based on various installed wind farm capacities and different time interval observations. Fig. 10 showcases the prediction curves and their 95% confidence intervals for each group. Tab. 7 depicts the deviation between the model presented and the day-ahead reported wind power prediction data.



(a) 15 min Interval (2500 MW)

(b) 10 min Interval (2500 MW)

(c) 5 min Interval (2500 MW)

(d) 5 min Interval (800 MW)

Figure 10 Tracking Results for Wind Farms at Different Time Intervals and Installed Capacities

Table 7 Deviation data between the prediction model and day-ahead reported wind power predictions

	0:15		00:30		00:45		1:00	
	Day-ahead	Forecast	Day-ahead	Forecast	Day-ahead	Forecast	Day-ahead	Forecast
15 min(2500 MW)	35%	15%	21%	18%	26%	13%	40%	20%
10 min(2500 MW)		12%		15%		10%		9%
5 min(2500 MW)		4%		6%		7%		5%
5 min(800 MW)		21%		17%		19%		15%

Simulation results have demonstrated that the model exhibits superior predictive accuracy across various installed capacities and data sampling intervals, outperforming traditional day-ahead forecasting methods. It was observed that, for a 2500 MW wind farm, the predictive accuracy of the model increased with the reduction of the observational data interval. Specifically, with a 15 minute lead time, the forecasted output at 00:15 was 374.56 MW. This forecast accuracy improved as the interval decreased to 10 minutes (345.756 MW) and further to 5 minutes (296.774 MW), juxtaposed against an actual output of 288.134 MW at that time. The error margins for forecasts, using 15, 10, and 5 minute intervals for this wind farm, were approximately 15 - 20%, 10 - 15%, and 5 - 10%, respectively. Furthermore, a comparative analysis at a 5 minute interval between the 2500 MW and 800 MW wind farms at 01:00 (as depicted in Fig. 10 at the horizontal coordinate point 964) revealed that the actual wind power was 486.182 MW. The forecast for the 2500 MW wind farm stood at 447.292 MW, translating to an accuracy rate of 5%, whereas for the 800 MW wind farm, it was 534.798 MW with an accuracy rate of 15%. A horizontal examination of the 800 MW wind farm's forecasts within the 00:15 - 01:00 period indicated an accuracy of about 15 - 20%. This error margin was observed to be larger in comparison to wind farms with higher installed capacities.These findings suggest that the model's accuracy is significantly influenced by both observational and prior estimation errors. A fitting analysis of data variability in wind farms of differing capacities indicated that larger capacities and shorter observational intervals contribute to

reduced observational errors and enhanced accuracy in tracking wind power output.

5 CONCLUSIONS

This research has successfully developed a mathematical model for predicting wind power, utilizing historical wind power generation data in conjunction with the ARIMA-FIGARCH model. The ARIMA-FIGARCH model, adept at capturing long-term correlations in time series, provides precise prior estimation results while optimizing computational resource usage. The employment of a *t*-distribution, characterized by location and scale parameters, for fitting the probability distribution of sample data volatility and mathematical equation residuals, facilitates the simultaneous computation of both prior estimation and observational errors within a unified dimensional framework. This integration forms the cornerstone for merging prior estimates with observational data, a key process within the recursive Bayesian framework. The implementation of recursive Bayesian filtering in this model serves to amalgamate prior estimation information with observational data, thereby enabling the construction of a dynamic predictive model. This model possesses the capability of real-time adjustment based on new observational data, leading to enhanced accuracy in predictive calculations and fulfilling the requirements for ultra-short-term wind power prediction. Simulation results demonstrated that, given a fixed installed capacity of the wind farm, under observations at different time intervals, larger time intervals are associated with larger observational errors, rendering the model predicting less accurate. Conversely, with a constant time interval, a larger installed capacity of the wind farm correlates with more accurate model predicting. In the context of the prevailing energy policies and the continuous upsurge in wind power installed capacity, this predictive model is anticipated to deliver superior forecasting results, particularly in large-capacity wind farms. The model's innovative approach, blending data mining techniques with recursive Bayesian filtering, significantly augments the precision of wind power forecasts. By merging observational data with insights extracted from data mining, the model achieves a high degree of accuracy in predicting future wind power output. This enables power grid dispatch departments to predict wind power supply with increased accuracy, coordinate available power across the entire grid accordingly, mitigate risks introduced by wind power uncertainties, elevate the utilisation of renewable energy sources, and assure the balance of power grid supply and demand.

6 REFERENCES

[1] The 14th Five Year Plan for National Economic and Social Development of the People's Republic of China and the Outline of the Long-Range Goals for 2023. 2021-03-13.

[2] Qian, Z., Pei, Y., Cao, L., Wang, J., & Jing, B. (2016). Review of wind power forecasting method. *High Voltage Engineering, 42*(4), 1047-1060.

[3] Yuan, T. Z., Li, H., & Jia, D. (2022). Modeling and Control Strategy of Wind-Solar Hydrogen Storage Coupled Power Generation System. *Journal of Intelligent Systems and Control*, *1*(1), 18-34. https://doi.org/10.56578/jisc010103

[4] Ge, W., Sun, P., Li, J., Hui, Q., & Kong, X. (2019). Robust estimation model of wind power prediction availability in the period of power system peak load. *High Voltage Engineering, 45*(4), 1281-1288.

[5] Shu, Y., Zhang, Z., Guo, J., & Zhang Z. (2017). Study on key factors and solution of renewable energy accommodation. *Proceedings of the CSEE, 37*(1), 1-8.

[6] Ye, L., Lu, P., Zhao, Y., Dai, B., & Tang, Y. (2021). Review of model predictive control for power system with large-scale wind power grid-connected. *Proceedings of the CSEE, 41*(18), 6181-6197.

[7] Ding, T., Feng, D., Lin, X., Chen, J., & Chen, L. (2017). Ultra-short-term wind speed forecasting based on improved ARIMA-GARCH model. *Power System Technology, 41*(6), 1808-1814.

[8] Villanueva, D., Pazos, J. L., & Feijoo, A. (2011). Probabilistic load flow including wind power generation. *IEEE Transactions on Power Systems, 26*(3), 1659-1667. https://doi.org/10.1109/TPWRS.2010.2096436

[9] Carta, J. A., Ramirez, P., & Velazquez, S. (2009). A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews, 13*(5), 933-955. https://doi.org/10.1016/j.rser.2008.05.005

[10] Ding, M., Wu, Y., & Zhang, L. (2005). Study on the algorithm to the probabilistic distribution parameters of wind speed in wind farms. *Proceedings of the CSEE, 25*(10), 107-110.

[11] Fan, G., Wang, W., Liu, C., & Dai, H. (2008). Wind power prediction based on artificial neural network. *Proceedings of the CSEE, 28*(34), 118-123.

[12] Bludszuweit, H. & Domínguez-Navarro, J. A. (2010). A probabilistic method for energy storage sizing based on wind power forecast uncertainty. *IEEE Transactions on Power Systems, 26*(3), 1651-1658. https://doi.org/10.1109/TPWRS.2010.2089541

[13] Pan, D., Liu, H, & Li, Y. (2008). A wind speed forecasting optimization model for wind farms based on time series analysis and Kalman filter algorithm. *Power System Technology, 32*(7), 82-86.

[14] Qin, L., Dong, H., & Wang, R. (2023). Hybrid energy storage based on Kalman filter and model predictive control to smooth out wind power fluctuation strategy. *Power System Technology*, 2023.

[15] Zhu, Q. M., Li, H. Y., Wang, Z. Q., Chen, J. F., & Wang, B. (2017). Short-term wind power forecasting based on LSTM. *Power System Technology, 12*(41), 3797-3802.

[16] Han, Y. C., Tong, X. Q., & Deng, Y. P. (2023). Probabilistic Distribution Estimation and Temporal Transformer-Based Interval Prediction in Day-ahead Wind Power Prediction. *Proceedings of the CSEE*.

[17] Qu, C., Xu, H., & Tan, Y. (2008). A survey of nonlinear Bayesian filtering algorithms. *Electronics Optics & Control, 15*(8), 64-71.

[18] Sun, Y., Liu, H., & Hu, T. (2023). Short-term wind speed forecasting based onGCN and FEDformer. *Proceedings of the CSEE*, 2023.

[19] Chen, L. & Yang, K. (2013). High-frequency volatility features fotecast model and performance evaluation. *Systems Engineering Theory & Practice, 33*(2), 296-307.

[20] Bi, G., Zhao, X., Li, L., Chen, S., & Chen, J. (2022). A short-term wind speed prediction model based on dual mode decomposition CNN-LSTM integration. *Acta Energiae Solaris Sinica, 44*(3), 191-197.

[21] Ye, F., Wang, Q., Chen, N., Fu, X., & Zhao, Y. (2021). Wind forecast model considering the characteristics of temporal and spatial distribution. *Power System Protection and Control, 43*(2), 1080-1086.

[22] Zhu, Q., Chen, J., Shi, D., Zhu, L., Bai, X., Duan, X., & Liu, Y. (2019). Learning temporal and spatial correlations jointly:

A unified framework for wind speed prediction. *IEEE Transactions on Sustainable Energy, 11*(1), 509-523. https://doi.org/10.1109/TSTE.2019.2897136

[23] Chen, R., Liu, J., Wang, F., Ren, H., & Zhen, Z. (2020, December). Graph neural network-based wind farm cluster speed prediction. *2020 IEEE 3rd student conference on electrical machines and systems (SCEMS)*, Jinan, China, 982-987. https://doi.org/10.1109/SCEMS48876.2020.9352310

[24] Tian, Z., Li, S., Wang, Y., & Gao, X.W. (2016). A short-term wind speed mixed prediction model based on ARIMA and ESN. *Acta Energiae Solaris Sinica, 37*(6), 1603-1610.

[25] Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems, 34*, 22419-22430.

[26] Hu, T., Ma, H., Liu, H., Sun, H., & Liu, K. (2022). Self-attention-based machine theory of mind for electric vehicle charging demand forecast. *IEEE Transactions on Industrial Informatics, 18*(11), 8191-8202. https://doi.org/10.1109/TII.2022.3180399

[27] Zhou, J., Wang, K. X., & Wang, Y. (2023). Fusion of transfer learning and CGAN for Ultra short term power prediction of wind power clusters. *Proceedings of the CSU-EPSA*.

[28] Adedotun, A. F. (2022). Hybrid Neural Network Prediction for Time Series Analysis of COVID-19 Cases in Nigeria. *Journal of Intelligent Management Decision*, *1*(1), 46-55. https://doi.org/10.56578/jimd010106

[29] Arumugam, V. & Natarajan, V. (2023). Time series modeling and forecasting using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average models. *Instrumentation Mesure Métrologie*, *22*(4), 161-168. https://doi.org/10.18280/i2m.220404

[30] Ramli, I., Rusdiana, S., Achmad, A., Azizah, & Yolanda, M.E. (2023). Forecasting of rainfall using Seasonal Autoregreressive Integrated Moving Average (SARIMA) in Aceh, Indonesia. *Mathematical Modelling of Engineering Problems*, *10*(2), 501-508. https://doi.org/10.18280/mmep.100216

[31] Bouhaddour, S., Saadi, C., Bouabdallaoui, I., Sbihi, M., & Guerouate, F. (2023). A novel hybrid approach for daily tourism arrival forecasting: The PROPHET-Bayesian Gaussian Process-Forward Neural Network model. *Ingénierie des Systèmesd'Information*, *28*(4), 833-842. https://doi.org/10.18280/isi.280404

Contact information:

Peng LIU
(Corresponding author)
School of Electrical Engineering, Shenyang University of Technology,
Shenyang 110000, China
E-mail: liupeng@smail.sut.edu.cn

Tieyan ZHANG, Professor, Member IEEE
School of Electrical Engineering, Shenyang University of Technology,
Shenyang 110000, China
E-mail: zty621103@sina.com

Furui TIAN, Engineer
State Grid Zhejiang Electric Power Company, LTD. Zhuji Power Supply Company,
Zhuji311800, China
E-mail: tfr512347083@qq.com

Yun TENG, Professor, Member IEEE
School of Electrical Engineering, Shenyang University of Technology,
Shenyang 110000, China
E-mail: tengyun@sut.edu.cn

Chuang GU, Engineer
Datang Heilongjiang Power Generation Co.,
Ltd. Harbin No.1 Thermal Power Plant,
Harbin 150000, China
E-mail: 381105009@qq.com