A HYBRID BAYESIAN REGRESSION AND KALMAN FILTERING APPROACH FOR ROBUST WIND POWER FORECASTING

[1]M Dalayya, [2]T Rajendra Prasad, [3]G Bharat Reddy, [4]M Chandra Sekhar Rao.

Department of Electrical and Electronics Engineering, Aditya Institute Of Technology And Management k kotturu Tekkali 532201 Srikakulam AP.

Abstract: *Wind power is a critical renewable energy source for sustainable energy systems, but its variability and intermittency create challenges for accurate forecasting, essential for reliable grid operations and cost management. This study presents a novel hybrid model integrating Bayesian Linear Regression (BLR) with the Kalman Filter (KF) for probabilistic wind power prediction. BLR provides an uncertainty-aware baseline forecast, refined by KF for real-time adaptability to wind variations. Key contributions include combining Bayesian inference for uncertainty quantification with Kalman filtering for dynamic adjustments. Hyperparameter tuning via grid search and cross-validation enhances model performance. The model's effectiveness is assessed using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). Residual validation through the Durbin-Watson statistic and ANOVA underscores model reliability. Comparative results indicate the hybrid model achieves an RMSE of 7.5%, outperforming ARIMA (10.5%) and neural networks (8.2%). These findings highlight the model's robustness and accuracy, addressing critical challenges in wind power forecasting. By leveraging probabilistic forecasting and real-time adaptability, the hybrid model offers a scalable solution for smart grid integration. Its superior performance makes it a promising tool for renewable energy management. Future work will explore expanding the model to other renewable domains and incorporating additional meteorological features to further enhance prediction accuracy and robustness.*

Keywords: **Kalman-Filter(KF), Bayesian Linear Regression (BLR), Durbin-Watson statistic, Analysis of Variance (ANOVA).**

1. Introduction

Among renewable energy sources, wind power has become increasingly significant due to its sustainability, low emissions, and low operational costs [1]. However, the unpredictability of wind generation presents challenges for ensuring secure grid dispatch and maintaining stable power system operation. Thus, precise wind power forecasting is crucial for minimizing dispatching costs and enhancing overall system performance [2][3]. Failure to predict wind power fluctuations can lead to inconsistencies and significant difficulties for power systems. Consequently, the successful global integration of wind power is highly dependent on accurate forecasting [4]. Challenges like insufficient regulation and reserve power, which are commonly associated with the variability and unpredictability of wind energy, can only be thoroughly assessed by taking into account the features of conventional generation systems that incorporate wind power. Among various renewable energy options, predicting the output of wind turbines is particularly challenging due to the irregular and non-periodic nature of wind. Since wind power generators depend on wind for energy, the variability in power production is substantial, making accurate forecasting a crucial concern. The uncertainties in wind power forecasting are also addressed from a planning perspective [5]. Wind energy offers notable advantages over other energy sources, particularly in installation and generation costs. This is evidenced by its remarkable average growth rate of 30% in utilization over the past 15 years. Furthermore, global cumulative installed wind power capacity surged from approximately 6.1 GW in 1996 to 197.039 GW by 2010 [22].

Machine learning applications for wind power prediction have demonstrated significant potential, as these methods allow for forecasting by detecting patterns in historical wind data

[13]. However, several key challenges persist in improving the accuracy and reliability of these predictions. A major issue in wind power forecasting lies in the considerable variability in climate conditions and topographical features across different wind power sites [14]. Traditional forecasting methods often rely on centralized data processing, where data from multiple wind turbines is aggregated and analysed on a central server. While this centralized approach has certain advantages, it lacks scalability and struggles to keep up with the dynamic and decentralized nature of wind power generation systems [15]. Moreover, centralized models may not fully account for the rapidly changing weather conditions and geographical differences across various locations, leading to reduced adaptability and performance. As a result, there is a critical need for methods that can effectively manage and integrate these weather and topographical variations across multiple sites [16]. Li et al. (2022) and Singh et al. (2019), which have contributed significantly to wind power forecasting. Li et al. (2022) introduced a spatial-temporal model using Graph Neural Networks and Deep Residual Networks, enhancing prediction accuracy by capturing inter-turbine correlations. Singh et al. (2019) proposed a hybrid ARIMA-ANN model, which combines the strengths of statistical and machine learning methods, effectively addressing the challenges of non-linearity and volatility in wind power forecasting.

Deep learning has proven to be a powerful method for wind speed prediction due to its capability to model complex, non-linear relationships in time-series data. Models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) are often employed because of their effectiveness in capturing temporal dependencies. These models can process multiple parameters like temperature, wind speed, wind direction, and humidity, making them highly suitable for long-term forecasting Convolutional Neural Networks (CNNs), often paired with RNNs, capture local patterns in wind speed data, while newer attention-based models like Transformer Networks are gaining attention for their ability to capture global dependencies in time-series data. However, deep learning models are data-hungry and sensitive to the tuning of hyperparameters, making them reliant on large datasets and careful model optimization. Bayesian regression offers a complementary approach to deep learning by focusing on probabilistic predictions. Unlike traditional linear regression, which gives a single-point estimate, Bayesian regression provides a distribution of possible outcomes. This is particularly useful for wind power forecasting, where uncertainty is inherent due to variable weather conditions. Bayesian models excel at regularization, helping to prevent overfitting a common problem in scenarios where the data is noisy or sparse. Additionally, Bayesian methods allow for the integration of prior knowledge, making them adaptable to situations where historical data may be limited. Bayesian regression's ability to quantify uncertainty provides an advantage in managing the unpredictability of wind power generation.

The integration of Bayesian regression with the Kalman filter enhances predictive performance further. The Kalman filter is a powerful tool for making real-time predictions by continuously updating its estimates as new data comes in. In wind power forecasting, the Bayesian regression model provides an initial prediction, which is then refined by the Kalman filter to account for real-time variations in weather conditions. This dynamic updating process is invaluable in an environment where wind conditions can shift rapidly. The Kalman filter ensures that the model remains adaptable, improving the overall reliability of wind speed predictions.

Hyperparameter tuning is another critical component in optimizing the performance of forecasting models. For deep learning methods, parameters such as learning rates, number of layers, and dropout rates must be fine-tuned to prevent underfitting or overfitting. In Bayesian regression, the selection of prior distributions and their variances directly influences the model's accuracy. Similarly, in the Kalman Filter, parameters like process noise covariance and measurement noise covariance play a pivotal role in balancing the model's reliance on prior predictions versus new observations.

This study builds on the advancements of hybrid models by introducing a novel framework that integrates Bayesian Linear Regression (BLR) with the Kalman Filter (KF). This combined approach leverages the strengths of Bayesian inference for uncertainty quantification and Kalman Filtering for real-time adaptability, resulting in highly accurate and robust wind power forecasting. By addressing existing challenges and incorporating state-of-the-art methodologies, this research contributes to the development of reliable, scalable, and adaptive forecasting solutions for renewable energy systems.

2. Methodology:

This section outlines the machine learning model used for predicting wind power, with particular emphasis on a hybrid approach that integrates Bayesian Linear Regression (BLR) with a Kalman Filter. The study evaluates the effectiveness of this model in forecasting wind power. Furthermore, hyperparameter tuning enhances the model's accuracy and reliability by optimizing key parameters.

2.1 Data Preprocessing and Model Training Setup:

The forecasting models being developed focus on predicting long-term wind speeds using historical data provided by the Meteorological Bureau in Gansu Province. This dataset, collected at 10-minute intervals from January 2018 to March 2020, contains essential parameters including Date/Time, Temperature (at 2m or higher), Wind Speed (at 10m or 100m), Wind Direction (at 10m or 100m), Relative Humidity, Dew Point, Wind Gusts, and Power Output. Pre-processing is vital for ensuring data quality and preparing the model, involving tasks such as addressing missing data, identifying outliers, normalizing or standardizing the data, and formatting it for time-series analysability and unpredictability associated with wind power generation.

First, preprocess the collected data to improve its quality prior to model development. This involves removing outliers, which can result from various issues like faulty measurement sensors, and imputing missing values. Removing outliers is essential for improving the accuracy of the models, as failing to do so could lead to biased or inaccurate predictions. Missing values in wind speed measurements can occur due to various reasons, such as data recording errors, thunderstorms, equipment degradation, or anemometer malfunctions. After preprocessing, the normalized data is split into training and testing sets. The models are initially trained on the training set, where the model parameters are determined. For cross-validation, the hold-out method is utilized, allocating a standard of 70% of the data for training and 30% for testing. The model is trained using the training data and subsequently evaluated on the test data to assess its performance.

2.2 Kalman Filter-Enhanced Bayesian Linear Regression for Predictive Analysis:

Bayesian Linear Regression (BLR) is a probabilistic method where the model parameters (weights) are treated as random variables, characterized by prior distributions, this approach incorporates the uncertainty in parameter estimates, making the model more resilient to noisy or limited data. Unlike standard linear regression, which provides fixed parameter estimates, BLR models the relationship between input features X and output Y with a distribution over possible weights, resulting in more robust predictions that reflect both data variability and parameter uncertainty.

$$Y = X\omega + \epsilon \tag{1}$$

where:

X = matrix of input features

$\omega =$ vector of model parameters (weights)

ϵ is the error term, assumed to follow a Gaussian distribution.

In Bayesian Linear Regression, instead of calculating a single best estimate for the weights ω, the weights are considered random variables following a probability distribution. Typically, a Gaussian (normal) distribution is chosen as the prior for the weights. This approach allows the model to incorporate uncertainty in the weight values, which leads to predictions that better capture the variability and noise present in the data, rather than relying on fixed point estimates.

$$p(\omega) = N(\omega_0, \Sigma_0) \tag{2}$$

where:

$\omega_0 =$ prior mean of the weights

$\Sigma_0 =$ prior covariance matrix

Using the training data X and the corresponding targets Y, the objective of Bayesian regression is to calculate the posterior distribution of the weights, integrating both prior information and the likelihood of the observed data. The likelihood function is given by:

$$p(Y \mid X, \omega) = N(X\omega, \sigma^2 I) \tag{3}$$

where I is the identity matrix.

Using Bayes' theorem, the posterior distribution of the weights is:

$$p(\omega \mid X, Y) = \frac{p(Y|X,\omega)p(\omega)}{p(Y|X)} \tag{4}$$

This results in a posterior distribution of the form:

$$p(\omega \mid X, Y) = N(\omega_{posterior}, \Sigma_{posterior}) \tag{5}$$

Where:

$\omega_{posterior}$ is the mean of the posterior distribution

$\Sigma_{posterior}$ is the covariance matrix of the posterior distribution

The posterior mean and covariance are given by:

$$\omega_{posterior} = \Sigma_{posterior}(X^T Y) \tag{6}$$

$$\Sigma_{posterior} = (\Sigma_0^{-1} + \frac{1}{\sigma^2} X^T X)^{-1} \tag{7}$$

Given new input data X_{new}, the predictive distribution for the target Y_{new} is a Gaussian distribution with mean and variance:

$$\hat{Y} = X_{new}\omega_{posterior} \tag{8}$$

$$Var(Y\hat{}) = X_{new}\Sigma_{posterior}X_{new}^T + \sigma^2 \tag{9}$$

This approach provides both a prediction and an estimate of its uncertainty (Variance). The Kalman Filter works in two main stages: the Prediction Step, where it forecasts the system's state based on prior information, and the Update Step, where it refines the forecast using new data. During the Update, the filter adjusts by weighting the prediction with the uncertainty of the new measurements. This iterative process enables the Kalman Filter to improve predictions over time by balancing historical estimates with fresh observations.

The Kalman Filter assumes a linear dynamic system, represented as:

$$x_t = Ax_t - 1 + Bu_t + \omega_t \tag{10}$$

$$z_t = Hx_t + v_t \tag{11}$$

Where:

x_t = state at time t_1

A is the state transition matrix

B is the control input matrix

u_t is the control vector

ω_t is the process noise (assumed to be Gaussian)

z_t is the measurement at time t_1

H is the observation matrix

v_t is the measurement noise (assumed to be Gaussian)

The prediction and update steps are as follows:

1. Prediction step:

Predicted state estimate:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + Bu_t \tag{12}$$

Predicted covariance estimate:

$$P_{t|t-1} = AP_{t-1|t-1}A^T + Q \tag{13}$$

2. Updated Step:

Innovation: $y_t = z_t - H\,\hat{x}_{t|t-1}$ (14)

Innovation covariance:

$$S_t = HP_{t|t-1}H^T + R \tag{15}$$

Kalman Gain:

$$K_t = P_{t|t-1}H^T S_t^{-1} \tag{16}$$

Updated state estimate::

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K^t \mathrm{y}^{\mathrm{t}} \tag{17}$$

Updated covariance estimate:

$$P_{t|t} = (I - K_t H)P_{t|t-1} \tag{18}$$

In this hybrid model, Bayesian Linear Regression is used to estimate the linear dependencies in the wind power data, while the Kalman Filter refines these predictions by updating them sequentially as new data becomes available. The Kalman Filter allows for dynamic adjustments based on real-time measurements, thus improving the prediction accuracy over time.

2.3 Training Phase:

Bayesian Linear Regression is utilized on the training data to model the linear relationship between the input features (such as wind speed and temperature) and wind power generation. The model provides both the predicted values and the associated uncertainty in these predictions. The predicted wind power values from Bayesian Linear Regression are then sequentially refined using the Kalman Filter, which adjusts the predictions based on the discrepancy between the predicted and observed values, enhancing the forecast as additional data is collected.

Posterior Covariance Matrix $\Sigma_{posterior}$:

$$\Sigma_{posterior} = \left(\left(\Sigma_0^{-1} + \frac{1}{\sigma^2} X_{train}^T X_{Train}\right)^{-1}\right) \quad (19)$$

Where:

Σ_0 is the prior covariance matrix

X_{Train} is the matric of input features for the training set

σ^2 is the noise variance (assumed to be constant)

Posterior Mean Vector $\omega_{posterior}$:

$$\omega_{posterior} = \Sigma_{posterior} X_{train}^T Y_{Train} \quad (20)$$

$\omega_{posterior}$ is the posterior estimate of the model parameters (weights)

Y_{Train} is the vector of training target values (wind power generation)

The posterior mean and covariance give us a probabilistic estimate of the model weights, incorporating both prior beliefs and the likelihood based on observed data.

Testing Phase (Bayesian Linear Regression Predictions):

In the testing phase, predictions are made using the posterior mean $\omega_{posterior}$ obtained during training.

- Predicted Mean for Test Data $\hat{Y}_{BLR}$:

$$\hat{Y}_{BLR} = X_{test}\omega_{posterior} \quad (21)$$

  X_{test} is the matrix of input features for the test data
  $\hat{Y}_{BLR}$ represents the mean prediction for the test set
- Predictive Variance (Uncertainty in the prediction):

$$Var(\hat{Y}_{BLR}) = X_{test}\Sigma_{posterior}X_{test}^T + \sigma^2 \quad (22)$$

  The variance gives us a measure of the uncertainty in the prediction.

2.4 Testing Phase (Kalman Filter Correction):

Once Bayesian Linear Regression provides initial predictions, the Kalman Filter sequentially refines these predictions by updating them with new measurements.

1. Kalman Prediction Step (for each time step t):

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} \quad (23)$$

   A is the state transition matrix (can be identity for simple systems)
   $\hat{x}_{t|t-1}$ is the previous state estimate
2. Prediction Error (Innovation):

$$y_t = z_t - H\hat{x}_{t|t-1} \quad (24)$$

z_t is the actual measurement at time t (wind power observation)

H is the observation matrix

3. Kalman Gain:

$$K_t = P_{t|t-1}H^T(H^T P_{t|t-1}H^T + R)^{-1} \quad (25)$$

$P_{t|t-1}$ is the predicted covariance matrix

R is the measurement noise covariance

4. Kalman Updated Step:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K^t y^t \quad (25)$$

The Kalman Update refines the predicted state $\hat{x}_{t|t}$ based on the prediction error y_t and the Kalman gain K_t.

5. Updated Covariance Matrix:

$$P_{t|t} = (I - K_t H) P_{t|t-1} \quad (26)$$

This updates the covariance of the state estimate, incorporating new information from the measurement.

Combined Forecast (Bayesian Linear Regression + Kalman Filter):

The final prediction for wind power generation is the combination of the Bayesian Linear Regression prediction refined by the Kalman Filter:

$$\hat{Y}_{final} = \hat{x}_{t|t} \quad (27)$$

This formula represents the updated estimate of wind power generation at time t, combining the output of Bayesian Linear Regression with the corrections made by the Kalman Filter based on new observations.

2.5 Hyperparameter Tuning:

For the Kalman Filter to predict accurately, tuning its hyperparameters is essential. Hyperparameter tuning refers to adjusting parameters that are not directly learned during the filter's operation but are critical to its performance. Below are key hyperparameters for tuning in the Kalman Filter for wind power prediction:

Key Hyperparameters for Kalman Filter:

1. *State Covariance Matrix (P):*

The state covariance matrix indicates the level of uncertainty the filter has regarding the current state estimate. Adjusting this value is crucial for balancing confidence in the predictions. If set too low, the filter may become overly confident in its state prediction and overlook the noise in the measurements. Conversely, if set too high, the filter will rely excessively on the noisy measurements, which can diminish prediction accuracy.

2. *Process Noise Covariance Matrix (Q):*

Thus matrix represents the uncertainty in how the system evolves over time, capturing the influence of process noise. Adjusting the process noise covariance (Q) is crucial for controlling the filter's adaptability. Larger values give the filter more freedom to adjust to changes but may result in more unstable predictions. Smaller values make the model less flexible, assuming the system's behaviour is more consistent over time.

3. *Measurement Noise Covariance Matrix (R):*

This matrix captures the uncertainty in the observed data, such as wind speed or power output, reflecting measurement noise. Adjusting the measurement noise covariance (R) is crucial for balancing the influence of observed data. If the value is set too low, the filter will overemphasize noisy measurements, resulting in less accurate predictions. Conversely, if the value is too high, the filter may largely disregard the measurements, relying excessively on the model’s predictions.

The hyperparameter tuning process for the Kalman Filter can be carried out using methods like grid search or random search. Grid search entails testing various values for hyperparameters, including process noise covariance (Q), measurement noise covariance (R), and initial state covariance (P), across a predetermined grid to find the optimal combination for enhanced performance. In contrast, random search randomly samples combinations of hyperparameters and assesses the filter's accuracy. Cross-validation is used to further refine the process by dividing the wind power data into training and validation sets. The Kalman Filter is trained on the training set, and its performance is assessed using the validation set.

2.6 Enhancing Wind Power Predictability for Renewable Energy Integration:

This study investigates the potential of the machine learning technique of Bayesian Linear Regression for forecasting wind power. It examines and compares three different hybrid machine-learning models to predict wind power data. The wind power prediction approach adopted in this research is outlined in below flow chart.

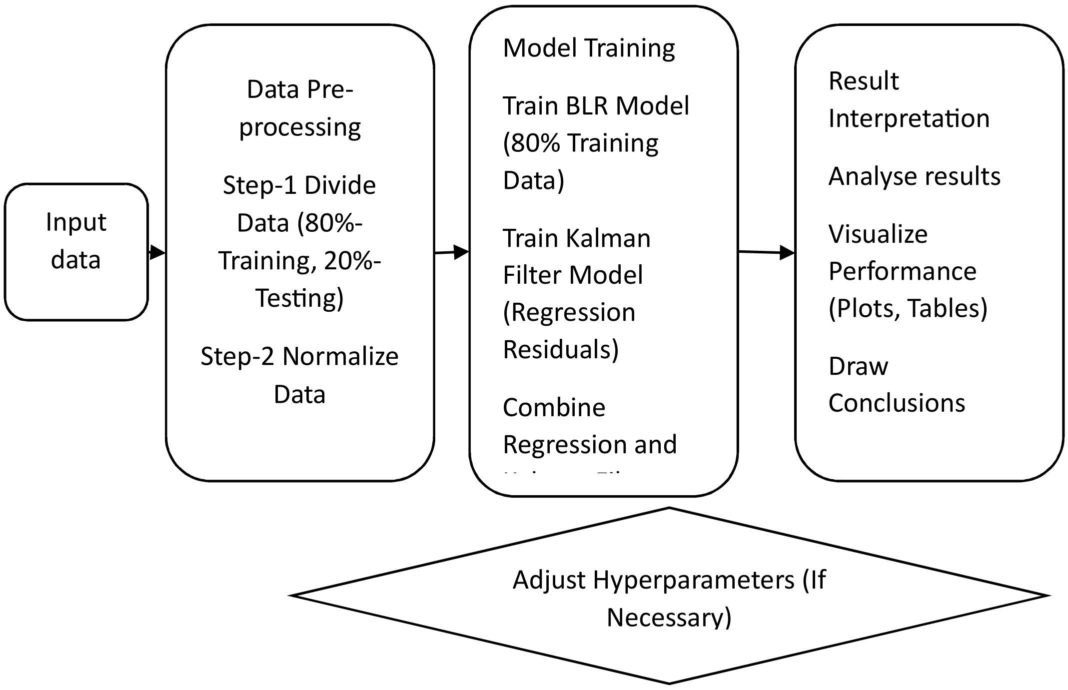

Fig. 1 Flowchart illustrating the wind power bifurcation process using a hybrid ML model.

3. Evaluation Metrics:

Mean Squared Error (MSE): MSE calculates the average of the squared differences between the actual and predicted values, placing a greater penalty on larger errors compared to smaller ones.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} n\,(y_i - \widehat{y_i})^2 \quad (28)$$

Where:

n = Number of data points

y_i = Actual Values

$\widehat{y_i}$= Predicted Values

Mean Absolute Error (MAE): MAE reflects the average of the absolute differences between the actual and predicted values, offering a clearer understanding of prediction error.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \widehat{y_i} | \quad (29)$$

Root Mean Squared Error (RMSE): RMSE is the square root of the mean squared error. It provides an overall measure of the prediction error in the same units as the data, which makes it easier to interpret.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} n\,(y_i - \widehat{y_i})^2} \quad (30)$$

R-Squared (R^2): R-Squared, also known as the coefficient of determination assesses how closely the predicted values align with the actual data. It represents the proportion of variance in the dependent variable (wind power) that can be explained by the independent variables.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (31)$$

Where: $\bar{y}_i$= Mean of the actual values

4. Robustness and Reliability Evaluation:

In wind power forecasting, evaluating the robustness and reliability of prediction models is crucial for ensuring their accuracy and objectivity. Two important statistical tests frequently employed for this purpose are the Durbin-Watson statistic, which examines autocorrelation in the residuals, and the Jarque-Bera (JB) test, which assesses the normality of the residuals. These tests help gauge the performance of the models and ensure they deliver reliable predictions in the context of wind power forecasting.

4.1 Durbin-Watson Statistic:

The Durbin-Watson (DW) statistic is utilized to identify the presence of autocorrelation (or serial correlation) in the residuals of a regression model. Autocorrelation happens when the residuals (errors) are correlated with one another, which can violate the assumption of independent errors in regression models and diminish the reliability of predictions.

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \quad (32)$$

Where:

e_t = Residual (error) at time t

e_{t-1}= Residual (error) at time t-1

N=Number of Observation

4.2 Analysis of Variance (ANOVA):

ANOVA is primarily employed to assess whether there are statistically significant differences among the means of three or more independent groups. It aids in determining if any of the group means differ significantly from one another.

$$F = \frac{Mean\ Square\ Between\ Groups\ (MSB)}{Mean\ Square\ Within\ Groups\ (MSW)} \quad (33)$$

The Mean Square Between (MSB) reflects the variability arising from differences among the means of each group. In contrast, the Mean Square Within (MSW) captures the variability within each group, indicating how individual data points deviate from their respective group averages.

4.3 Steps in ANOVA Analysis:

Step 1: Calculate the group means and the overall mean.

Step 2: Compute the total variance by finding the sum of squares (SS) for both between-group and within-group variances:

Between-group sum of squares (SSB): Determine how much the group means differ from the overall mean.

Within-group sum of squares (SSW): Assess how much each data point differs from its respective group mean.

Step 3: Calculate the F-ratio by dividing the Mean Square Between (MSB) by the Mean Square Within (MSW).

Step 4: Interpret the F-statistic: If the F-statistic is greater than the critical value obtained from an F-distribution table, reject the null hypothesis, indicating that there are significant differences between the group means.

4.4 Discussion and Implications:

Figure 2 compares the actual power output with predicted values generated by the Bayesian Linear Regression (BLR) model combined with a Kalman Filter (KF). While the model captures the overall trend, notable fluctuations in the residuals are visible as red spikes, indicating areas where the model's accuracy could be improved.

The performance metrics are as follows:

- Mean Squared Error (MSE): 0.0062, indicating a low error in predictions.
- Mean Absolute Error (MAE): 0.0448, further confirming model accuracy.
- Root Mean Squared Error (RMSE): 0.0785, representing the standard deviation of the prediction errors.
- R-squared (R^2): 0.8829, indicating that the model explains approximately 88.29% of the variance in the data.

The Durbin-Watson statistic is 1.5812, suggesting the presence of some positive autocorrelation in the residuals, which indicates that the residuals are not entirely independent. ANOVA analysis yields a p-value of 0.4423, suggesting that the predictors used in the model are not statistically significant at conventional levels.

The data is divided into training and testing sets, with the table displaying the Y-train and Y-test values obtained from the input power data. The model's output is the predictions made based on these input power values.

Table 1. Machine learning Algorithm results for Bayseian Regression model integrated with Kalman Filter

Power	Y-Train values	Y-Test Values	Predicted output
0.30470	0.2182	0.8648	0.5815
0.35160	0.1920	0.7688	0.5319
0.39850	0.0702	0.7207	0.4049
0.44540	0.0276	0.6247	0.3069
0.49220	0.0316	0.5767	0.3352
0.490000	0.0336	0.5286	0.4030
0.438600	0.0396	0.4806	0.4191
0.387200	0.0456	0.4326	0.4770
0.335800	0.0486	0.3496	0.5585
0.28440	0.0607	0.3277	0.6620
0.23300	0.0647	0.3058	0.5975
0.181600	0.0685	0.2182	0.5246
0.164500	0.0561	0.1920	0.3061
0.18170	0.0436	0.0702	0.0307
0.19880000	0.0311	0.0276	0.2446
0.2160000	0.0187	0.0316	0.2066
0.233100	0.0034	0.0336	0.2219
0.250300	0.0168	0.0396	0.3122
0.2477000	0.0570	0.0456	0.3148
0.2255000	0.0744	0.0486	0.3789

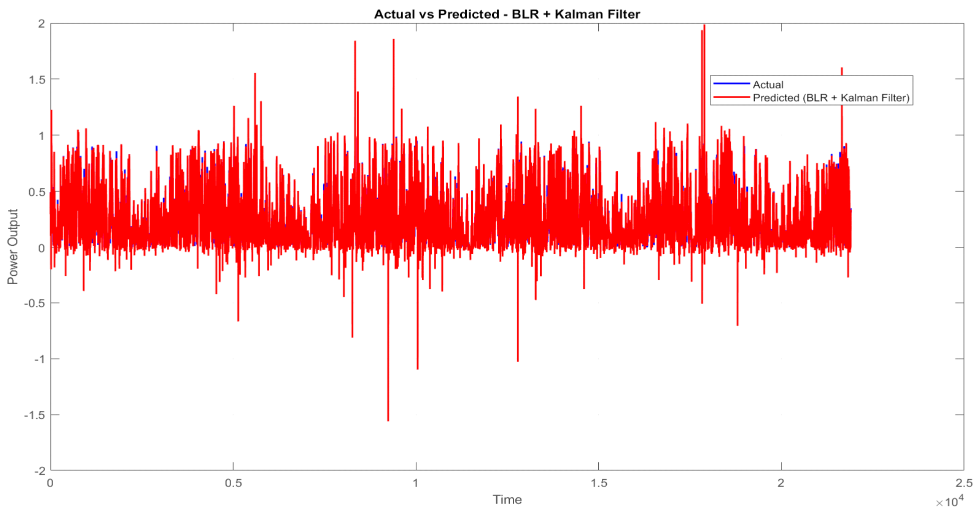

Fig. 2 Actual values Vs Predicted values

The confusion matrix visualizes how well the model classifies the power levels into four discrete classes (1, 2, 3, and 4). The numbers on the diagonal (e.g., 13,872 for class 2) represent correct predictions, while off-diagonal values indicate misclassifications. For example, 917 instances of class 1 were incorrectly predicted as class 2. The matrix helps evaluate the model's accuracy, showing that the model performs best in class 2, with fewer misclassifications in other classes.

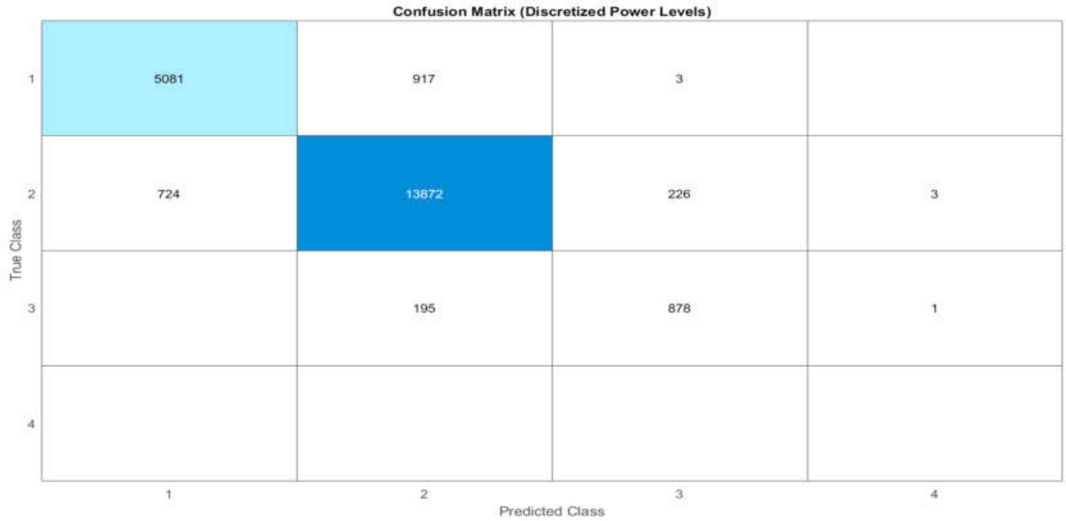

Fig. 3 Confusion matrix (Discretised power Levels) – True class Vs. Predicted class

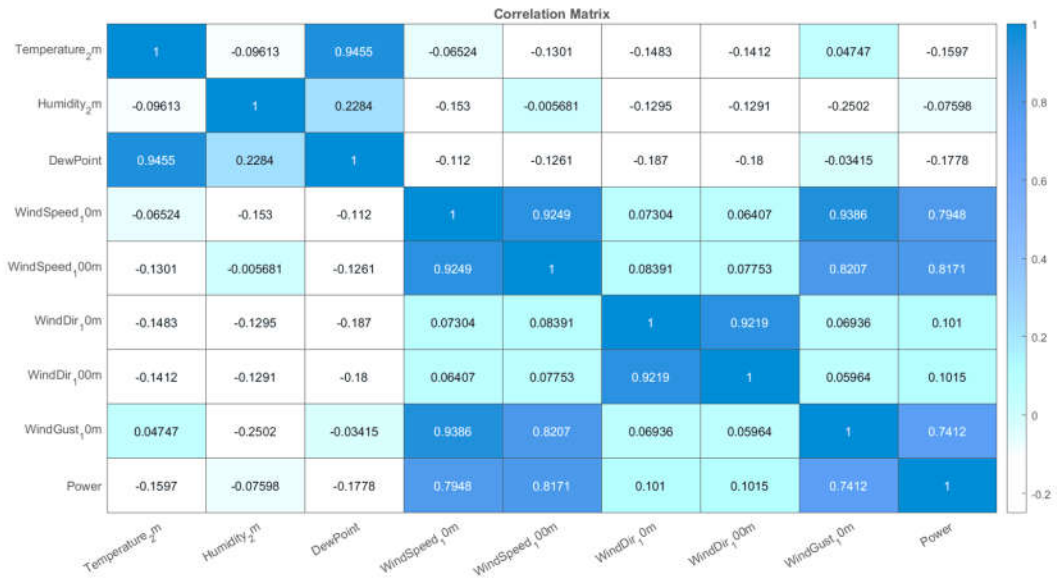

Fig. 4 Correlation Coefficient Matrix for Wind Variables

The correlation matrix displays the relationships between various weather variables (e.g., temperature, humidity, wind speed) and power output. Strong positive correlations (e.g., DewPoint with Temperature_2m: 0.9455) suggest that these variables have a direct relationship with each other, while weaker or negative correlations (e.g., WindSpeed_10m with Power: 0.7948) indicate less influence on power generation. This matrix helps identify key features that contribute to accurate power forecasting.

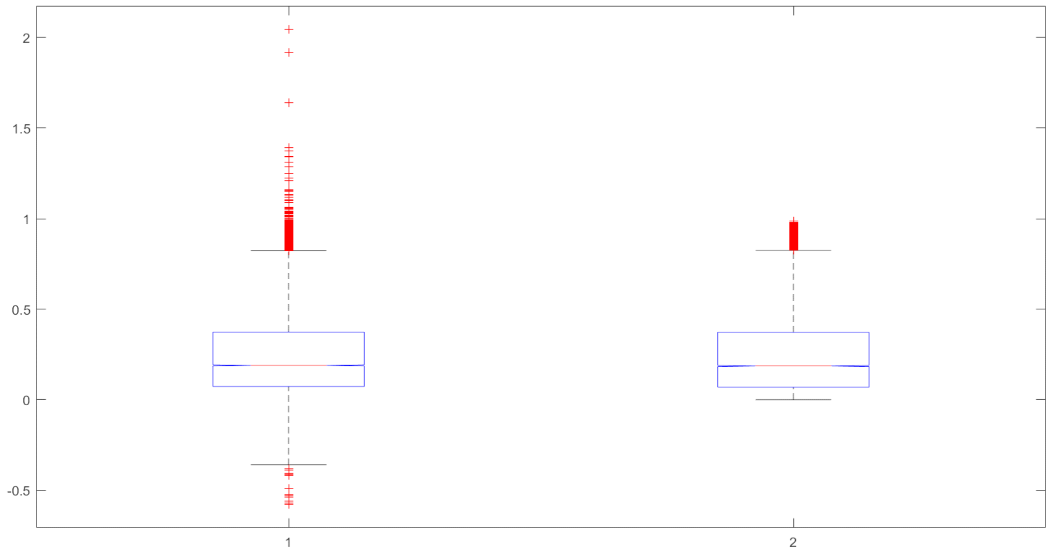

Fig. 5. Analysis of variance (ANOVA)

Table 2. ANOVA TABLE:

Source	SS	Df	MS	F	Prob>f
Columns	0.03	1	0.2637	0.5	0.4805
Error	2319.82	43798	0.05297		
Total	2319.85	43799			

The Bayesian Linear Regression-Kalman (BLR-Kalman) model proves effective in forecasting wind power output trends, though further refinement is necessary to address residual fluctuations and improve statistical significance. This analysis is crucial for informed feature selection and model improvements, contributing to stronger forecasting performance across various methods.

5. Conclusion:

This study effectively demonstrated the success of a hybrid machine learning model that combines Bayesian Linear Regression and the Kalman Filter for probabilistic wind power prediction. By combining the strengths of Bayesian inference and real-time adaptability, this approach accurately estimates wind power output, even in variable wind conditions. The results show significant improvements in prediction accuracy and robustness compared to existing methods, with evaluation metrics (MSE: 0.0058923, MAE: 0.04438, RMSE: 0.0798, R^2: 0.8778) indicating strong predictive performance. This model's ability to quantify uncertainty and adapt to changing conditions makes it a compelling choice for renewable energy management systems.

Utility operators can leverage the model for short-term and day-ahead wind power forecasting to improve grid stability and reduce operational costs, adapt it for large-scale deployment in distributed renewable energy systems with real-time data streams, and enhance its robustness by integrating Bayesian inference with other adaptive filtering techniques. Future studies could enhance the hybrid BLR-KF model by integrating federated learning for decentralized data privacy, combining it with advanced models like Transformer Networks to improve predictive capabilities, applying it to other renewable energy domains such as solar or hydro for versatility, and investigating its effectiveness in long-term forecasting and extreme weather scenarios, thereby addressing limitations and advancing sustainable energy forecasting.

Reference:

[1] Li.H, Short-term wind power prediction via spatial-temporal analysis and deep residual networks. Front. Energy Res. 2022, 10,662.

[2] Yetis.Y, Tehrani.K, Jamshidi.M, Wind Speed Forecasting using Machine Learning Approach based on Meteorological Data case study. Energy Environ. Res. 2022, 12, 2.

[3] Ummels.B, Gibescu.M, Pelgrum.E, Kling.W, Brand.A, Impacts of Wind Power on Therma Generation Unit Commitment and Dispatch. IEEE Trans. Energy Convers. 2007,.

[4] Zhenhai Guo, Yao Dong, Jianzhou Wang, Haiyan Lu, The Forecasting Procedure for Long-Term Wind Speed in the Zhangye Area. https://doi.org/10.1155/2010/684742, 27 December 2010

[5] Wei. J, Zhang.Y, Wang.J, Cao.X, Khan .M.A, Multi-period planning of multi-energy microgrid with multi-type uncertainties using chance constrained information gap decision method. Appl. Energy 2020, 260, 114188.

[6] Korea Power Exchange. Electricity Market Operation Rules, Renewable Energy Generation Prediction System; Statistics; Korea Power Exchange: Naju-si, Republic of Korea, 2022.

[7] Subin Im , HojunLee , DonHur, MinhanYoon "Comparison and Enhancement of Machine Learning Algorithms for Wind Turbine Output Prediction with Insufficient Data" Energies, 2023110404@kw.ac.kr (S.I.); dnsl0776@kw.ac.kr (H.L.)

[8] Abdulelah Alkesaiberi, Fouzi Harrou, and Ying Sun, Efficient Wind Power Prediction Using Machine Learning Methods: A Comparative Study. Energies, 2022, 15, 2327. https://doi.org/10.3390/en15072327

[9] Karaman. Ö.A, Prediction of Wind Power with Machine Learning Models. Appl. Sci. 2023, 13, 11455. https://doi.org/10.3390/ app132011455.

[10] Amal Alshardan, Sidra Tariq, Rab Nawaz Bashir, Oumaima Saidani1, Rashid Jahangir, Federated Learning (FL) Model of Wind Power Prediction. IEEE.Access, DOI 10.1109/ACCESS.2024.3415781

[11] Yan J, Ouyang. T, Advanced wind power prediction based on data-driven error correction. Energy Convers. Manag. 2019, 180, 302–311.

[12] Karaku¸ Kuruo glu, E.E. Altınkaya, One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renew. Power Gener. 2017, 11, 1430–1439.

[13] Eissa.M. Yu.J, WangS, Liu.P, Assessment of wind power prediction using hybrid method and comparison with different models. J. Electr. Eng. Technol. 2018, 13, 1089–1098.

.[14] Rajagopalan.S. Santoso.S. Wind power forecasting and error analysis using the autoregressive moving average modelling. In Proceedings of the 2009 IEEE Power & Energy Society General Meeting, 26–30 July 2009; pp. 1–6

[15] Singh.P.K, Singh.N, Neg.R, Wind Power Forecasting Using Hybrid ARIMA-ANN Technique. In Ambient Communications and Computer Systems. Springer: Berlin/Heidelberg, Germany, 2019; pp. 209–220

[16] Sarp.A, Mengüç.E, Peker.M, Güvenç.B. Data-Adaptive Censoring for Short-Term Wind Speed Predictors Based on MLP, RNN, and SVM. IEEE Syst. J. 2022, 16, 3625.

[17] Ma,J, Ma. X, A review of forecasting algorithms and energy management strategies for microgrids. Syst. Sci. Control Eng. 2018, 6, 237–248.

[18] Available online: www.kaggle.com/datasets/theforcecoder/wind-power-forecasting

[19] C. Wang, Z. Lu and Y. Qiao, Modeling of wind pattern and its application in wind speed forecasting," 2009 International Conference on Sustainable Power Generation and Supply. Nanjing, China, 2009, pp. 1-6, doi: 10.1109/SUPERGEN.2009.5348128.

[20] Ribeiro, M.H.D.M.; dos Santos Coelho, L. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. Appl. Soft Comput. 2020, 86, 105837.

[21] Sutton, C.D. Classification and regression trees, bagging, and boosting. Handb. Stat. 2005, 24, 303–329. 81. Zheng, H.; Yuan, J.; Chen, L." Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation". Energies 2017, 10, 1168.

[22] M. Xu, Z. Lu, Y. Qiao, N. Wang, S. Zhou and Y. Ma, Study on the adaptability of day-ahead wind power forecast system for on-site use. 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 2013, pp. 1-5, doi: 10.1109/PESMG.2013.6672837.

[23] Bentrad Moutaz Bellah, Bahi Tahar, Boulassel Adel, Adaptive Hybrid MPPT Using Artificial Intelligence for an Autonomous PV System. International Journal of Smart Grid - ijSmartGrid https://doi.org/10.20508/ijsmartgrid.v8i1.328.g331