SSR Journal of Engineering and Technology (SSRJET)

Homepage: https://ssrpublisher.com/ssrjet/

Volume 3, Issue 4, 2026 ISSN: 3049-0383

Adaptive Hybrid Markov Chain–LSTM Framework for Multi-Horizon Wind Speed Forecasting with Bias–Variance Optimization

Peter Ayodimeji Ogunsola, Ibrahim Abba & Opeyemi Amusan

Department of Engineering, Federal University of Science and Technology, Kabo, Kano State, Nigeria

Received: *21.03.2026* | ***Accepted:*** *19.04.2026* | ***Published:*** *20.04.2026*

***Corresponding author:** Peter Ayodimeji Ogunsola

DOI: *10.5281/zenodo.19665276*

Abstract **Original Research Article**

The coexistence of stochastic regime transitions and nonlinear temporal dependencies has made accurate wind speed forecasting in complex terrain remains challenging. This study proposes an adaptive hybrid Markov Chain–Long Short-Term Memory (MC–LSTM) framework for multi-horizon wind speed forecasting, designed to address the limitations of standalone statistical and deep learning models. While Markov models capture stochastic state transitions effectively, they struggle with long-term dependencies, whereas LSTM networks model temporal patterns but may exhibit smoothing effects and error accumulation over extended horizons. To overcome these limitations, a bias–variance optimization perspective is introduced to guide the integration of both models within a unified adaptive ensemble framework. Using high-resolution NASA POWER hourly wind data for a complex terrain site, the proposed model dynamically adjusts the contribution of each component across forecasting horizons (t+1, t+3, t+6, and t+24). The results show that the hybrid framework consistently outperforms individual models in terms of RMSE, MAE, and coefficient of determination (R^2), while exhibiting slower error propagation across increasing forecast horizons. Statistical validation using the Diebold–Mariano test confirms the significance of performance improvements. The findings highlight that the proposed framework provides a more stable and generalizable forecasting approach by balancing model bias and variance, making it suitable for power system planning and renewable energy integration under uncertain wind conditions.

Keywords: Hybrid forecasting, Wind speed prediction, LSTM, Markov chain, Bias–variance.

I. INTRODUCTION

The growing penetration of renewable energy into modern power systems is primarily driven by the global transition toward sustainable energy generation. Among the available renewable resources, wind energy has gained significant attention due to its environmental benefits, technological maturity, and wide geographical availability. However, the inherent variability and stochastic nature of wind present major challenges for power system operation, energy planning, and optimal turbine control. These challenges necessitate the development of accurate and reliable wind speed forecasting models to ensure efficient energy management and grid stability [1], [2]. Early approaches to wind speed forecasting were predominantly based on statistical time-series models such as

persistence methods, autoregressive (AR), moving average (MA), and autoregressive integrated moving average (ARIMA). While these models are computationally efficient and offer interpretability, their reliance on linearity and stationarity assumptions limits their effectiveness in real-world wind systems, which are inherently nonlinear and non-stationary. Consequently, their performance degrades significantly in complex and highly variable environments [3]. To address these limitations, machine learning and deep learning techniques such as artificial neural networks (ANN), support vector regression (SVR), and long short-term memory (LSTM) networks have been widely adopted. Among these, LSTM models have demonstrated strong capability in capturing temporal dependencies and nonlinear relationships in wind speed data. Their ability to model both short-term fluctuations and long-term dependencies makes them particularly suitable for time-series forecasting tasks [4], [5], [6]. More advanced architectures such as gated recurrent units (GRU), bidirectional LSTM, and transformer-based models further enhance predictive performance by modeling complex temporal interactions [7], [8], [9]. However, despite their advantages, these models are computationally demanding, data-intensive, and often prone to overfitting, especially under highly variable wind conditions.

In contrast, stochastic models such as Markov chains offer an alternative perspective by representing wind speed dynamics as transitions between discrete states. These models are particularly effective in capturing short-term persistence and regime-switching behavior inherent in wind processes [10], [11]. Extensions such as higher-order Markov models and semi-Markov frameworks have been proposed to improve predictive capability [12], [13], [14]. Nevertheless, the discretization process introduces subjectivity and potential information loss, while the limited memory structure restricts their ability to capture long-term dependencies and nonlinear interactions. As a result, Markov models typically exhibit high bias but low variance, limiting their standalone performance in complex environments. Given the complementary strengths and weaknesses of these approaches, hybrid models that integrate stochastic and machine learning techniques have attracted increasing research attention. By combining probabilistic structure with nonlinear learning capability, hybrid models aim to achieve improved forecasting accuracy. However, many existing hybrid approaches rely on fixed or heuristic weighting strategies, which fail to adapt to changing wind regimes. Furthermore, the absence of a rigorous theoretical foundation in most hybrid frameworks makes their performance improvements largely empirical rather than systematically justified [12], [15], [16], [17], [18]. Another critical limitation in existing studies is the focus on relatively homogeneous terrains, where wind patterns are more predictable. In contrast, complex terrains which are characterized by elevation variations, surface roughness, and orographic effects introduce significant uncertainty and variability in wind dynamics where the Mambilla Plateau in Taraba State, Nigeria represents a typical example of such environments. This highlights the need for models capable of capturing both nonlinear temporal behavior and stochastic regime transitions under complex environmental conditions. From a theoretical perspective, the bias–variance trade-off provides a useful framework for understanding the limitations of existing models and the potential benefits of hybridization. Markov chain models generally exhibit higher bias but lower variance, while deep learning models such as LSTM exhibit lower bias but higher variance due to their flexibility. Therefore, an appropriately designed hybrid model can improve generalization performance by balancing these two components effectively [19]. Motivated by these gaps, this study proposes an adaptive hybrid Markov Chain–Long Short-Term Memory (MC–LSTM) framework for multi-horizon wind speed forecasting, in complex terrain, whereby integrates stochastic state-transition modeling with nonlinear temporal learning and employs an adaptive ensemble weighting mechanism to dynamically balance model contributions. The framework is validated using real wind speed data from the Mambilla Plateau, providing insights into its effectiveness in capturing terrain-induced variability and improving multi-horizon forecasting performance. Unlike conventional hybrid forecasting approaches that rely on fixed or heuristic weighting schemes, this work

introduces a bias–variance optimization perspective to systematically guide the integration of Markov and LSTM models and dynamically adjusts model contributions across different forecasting horizons, improving robustness and reducing error propagation in longer-term predictions and the performance improvements are rigorously validated using the Diebold–Mariano test, ensuring that gains are statistically significant rather than incidental. Although demonstrated on a complex terrain dataset, the proposed methodology is designed to be transferable to other geographical regions and renewable energy forecasting applications. Although the model is evaluated using site-specific data, the proposed framework is inherently generalizable and can be extended to other regions and renewable energy forecasting problems with similar stochastic and temporal characteristics.

II. MATERIALS AND METHODS

2.1 Study area and data description

This study focuses on site-specific wind speed forecasting using high-resolution meteorological data obtained from the NASA POWER database for the Mambilla Plateau in Taraba State, Nigeria. The region is characterized by pronounced diurnal cycles, seasonal variability, and terrain-induced wind fluctuations, which introduce significant nonlinear and stochastic behavior into the wind field. These characteristics make the dataset particularly suitable for evaluating hybrid forecasting models. Hourly observations spanning January 2025 to February 2026 were utilized to ensure adequate temporal coverage of both short-term dynamics and longer-term variability. The dataset includes three key atmospheric variables: wind speed at 50 m height (WS50M), surface pressure (PS), and air temperature at 2 m (T2M). These variables collectively capture both local wind characteristics and broader atmospheric influences relevant to wind evolution.

2.2. Data processing

Raw meteorological data often contain inconsistencies, scale heterogeneity, and missing entries, which can adversely affect both probabilistic and learning-based models if not properly addressed. To mitigate these issues, a structured preprocessing pipeline was implemented, including missing value treatment, normalization, and state discretization. These steps ensure compatibility between the Markov-chain and LSTM components while improving model stability and generalization performance.

2.2.1. Missing value treatment

Data quality control was performed to ensure reliability of the input series. Missing values were initially handled using linear interpolation (1), which provides a simple and effective approach for short gaps.

$$x_t = x_{t-1} + \frac{x_{t+1}-x_{t-1}}{2} \quad (1)$$

However, for longer gaps, linear interpolation alone may not adequately preserve temporal structure, particularly in wind regimes with strong diurnal and seasonal patterns. To address this limitation, a seasonally informed imputation strategy was employed (2), where values were estimated using observations from similar temporal contexts (e.g., the same hour across adjacent days). This approach helps maintain consistency with the underlying atmospheric behavior.

$$v_t = \frac{1}{N_s}\sum_{k\in S_t} v_k \quad (2)$$

where x_{t_0} and x_{t_1} represent the nearest available observations surrounding the missing time step, and x_t denotes the interpolated value. Historical observations corresponding to the same temporal context (for example, the same hour of day across adjacent days) were also considered when longer gaps required seasonally informed imputation.S_t

2.2.2. Normalization

To stabilize training and ensure numerical efficiency, all input variables were normalized using Min–Max scaling:

$$x_t^{norm} = \frac{x_t - x_{min}}{x_{max} - x_{min}} \quad (3)$$

This transformation constrains values to the interval [0, 1], improving convergence of the LSTM model.

2.2.3. State discretization for Markov modelling

The continuous wind speed series was discretized into a finite number of states using quantile-based partitioning (4). This approach ensures that each state contains approximately equal numbers of observations, improving the statistical robustness of transition probability estimation.

$$S_i = [q_{i-1}, q_i), i = 1,2, \dots, K \quad (4)$$

where q_k represents the kth quantile used to define the state boundaries.$q_i i^{th}$

Unlike uniform discretization, which may impose arbitrary boundaries, the quantile-based method adapts to the empirical distribution of the data. This is particularly important in complex terrain environments, where wind variability is non-uniform. By aligning state boundaries with the data distribution, the discretization better reflects actual wind regimes and enhances the sensitivity of the Markov model to realistic transitions.

Each state corresponds to a wind speed interval:

$$s_i = [v_i^{low}, v_i^{high}) \quad (5)$$

The mapping from continuous wind speed to discrete states is defined as:

$$X_t = s_i \text{if } v_t \in [v_i^{low}, v_i^{high}) \quad (6)$$

The choice of state width is crucial. A uniform discretization is given by:

$$\Delta v = \frac{v_{\max} - v_{\min}}{K} \quad (7)$$

However, in complex terrains where wind variability is non-uniform, adaptive discretization based on statistical dispersion provides better representation:

$$\Delta v = \sigma_v \quad (8)$$

This ensures that the discretization reflects the inherent variability of the dataset rather than imposing arbitrary intervals. As a result, the Markov model becomes more sensitive to realistic wind transitions.

2.3. Markov-chain model

The Markov Chain model provides a stochastic framework for modeling wind speed evolution based on transition probabilities between discrete states, and has been widely applied in wind speed modeling due to its ability to capture regime-switching behavior [10], [11], [12]. It assumes that the future state depends only on the current state, which is a reasonable approximation for short-term wind dynamics.

2.3.1. State definition

Each wind speed observation is mapped to a discrete state:

$$s_t \in \{S_1, S_2, \dots, S_K\} \quad (9)$$

This mapping transforms the continuous time series into a symbolic representation, facilitating probabilistic analysis. The state-based representation allows the model to capture regime-switching behavior, such as transitions between low, moderate, and high wind conditions.

2.3.2. Transition probability matrix

The core of the Markov model lies in the transition probability matrix, which quantifies the likelihood of transitioning from one wind state to another.

The transition probabilities are formally defined as:

$$P_{ij} = P(X_{t+1} = s_j \mid X_t = s_i) \quad 10)$$

In practical implementation, these probabilities are estimated empirically from observed state transitions within the training dataset:

$$P_{ij} = \frac{N_{ij}}{\sum_{j=1}^{K} N_{ij}} \quad (11)$$

where N_{ij} denotes the number of observed transitions from state i to state j.$N_{ij} S_i S_j$

. This empirical formulation ensures that the transition matrix reflects the underlying stochastic dynamics of wind behavior, consistent with established Markov-based wind modeling approaches [12], [20].

2.3.3. State prediction

Once the transition matrix is established, future states can be predicted using probabilistic propagation:

$$\pi_{t+1} = \pi_t P \quad (12)$$

where π_t denotes the state probability distribution at time t.$\pi_t t$

For practical forecasting, a deterministic approximation is often used:

$$\hat{X}_{t+1} = \arg\max_j P_{ij} \quad (13)$$

The predicted wind speed is then obtained by mapping the state to its representative value:

$$\hat{v}_{t+1}^{MC} = \frac{v_j^{low} + v_j^{high}}{2} \quad (14)$$

This approach provides a computationally efficient means of capturing regime-based wind dynamics and provides robust baseline predictions with low variance, making it suitable for integration into the hybrid framework.

2.4. LSTM-based machine learning model

Unlike traditional models, LSTM networks can learn long-term relationships and complex patterns, making them well-suited for time series forecasting. LSTM networks have been extensively used for wind speed forecasting due to their ability to capture long-term temporal dependencies and nonlinear relationships in time-series data [4], [6], [7]. The LSTM model is employed to capture nonlinear temporal dependencies in wind speed data.

Input representation

A sliding window approach is adopted to construct input sequences:

$$X_t = [x_{t-n}, x_{t-n+1}, \dots, x_{t-1}] \quad (15)$$

where L denotes the input sequence length.n

This representation allows the model to learn temporal dependencies over multiple time steps. By incorporating historical information, the LSTM can capture both short-term fluctuations and longer-term trends.

2.4.1 LSTM formulation

Wind speed exhibits strong temporal dependencies, including short-term persistence and long-range correlations. Traditional models fail to capture these dependencies effectively, which motivates the use of Long Short-Term Memory (LSTM) networks. The LSTM architecture introduces memory cells that regulate information flow through gating mechanisms:

The forget gate determines which past information to discard:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (16)$$

The input gate controls how much new information is incorporated:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (17)$$

The candidate cell state represents new information:

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (18)$$

The cell state is updated as:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (19)$$

Finally, the output gate determines the hidden state:

$$h_t = o_t \odot \tanh(C_t) \quad (20)$$

These gating mechanisms enable the model to capture both short-term fluctuations and long-term temporal patterns, which are critical for accurate wind forecasting.

2.4.2. Prediction

The LSTM output is mapped to wind speed forecast:

$$\hat{y}_{LSTM} = W_o h_t + b_o \quad (21)$$

This formulation allows the model to produce continuous-valued predictions, capturing nonlinear relationships inherent in wind dynamics.

2.5. Proposed hybrid MC-LSTM model

The proposed hybrid MC–LSTM model as shown in figure 1, combines the complementary strengths of stochastic and machine learning approaches. The Markov chain component provides a probabilistic representation of regime transitions, offering stability and low variance, while the LSTM component captures nonlinear temporal dependencies and complex sequential patterns.

The hybrid prediction is formulated as a weighted combination of the two model outputs (22), subject to the constraint (23), ensuring that the final prediction remains within a physically meaningful range.

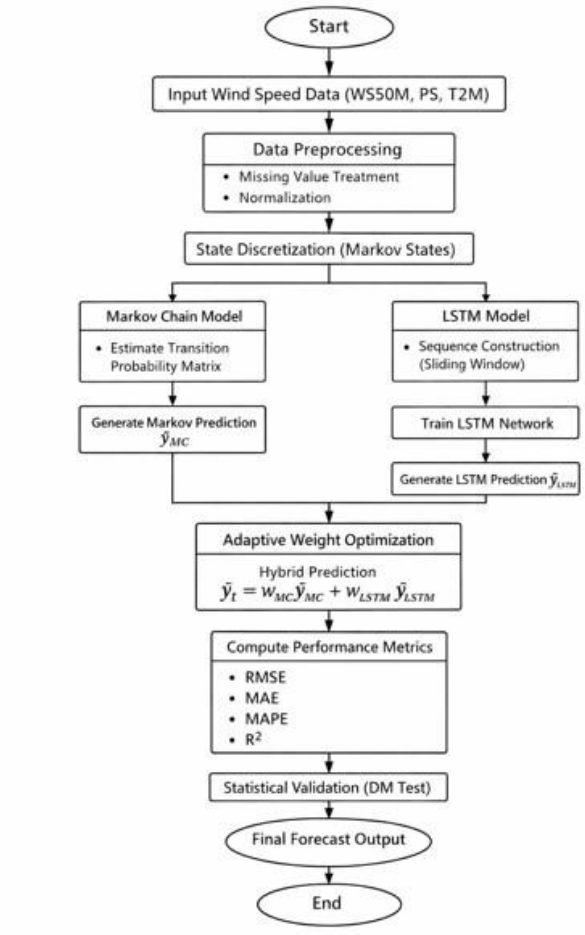

Fig.1 The proposed hybrid MC–LSTM forecasting framework

2.5.1. Weighted ensemble formulation

The hybrid prediction is formulated as a convex combination of the two model outputs:

$$\hat{v}_t^{Hybrid} = w_{MC}\hat{v}_t^{MC} + w_{LSTM}\hat{v}_t^{LSTM} \quad (22)$$

subject to:

$$w_{MC} + w_{LSTM} = 1, 0 \leq w_{MC}, w_{LSTM} \leq 1 \quad (23)$$

This formulation ensures that the final prediction remains within the feasible range of the individual model outputs.

2.5.2. Bias-variance theoretical framework

A fundamental challenge in wind speed forecasting lies in balancing model flexibility with generalization capability. Highly simplified models fail to capture nonlinear atmospheric dynamics (high bias), while highly flexible models tend to overfit noise (high variance). This trade-off is formally captured by the bias–variance decomposition.

Let y_t denote the true wind speed and ŷ_t the model prediction. The expected prediction error is decomposed as:$y\hat{y}$

$$\mathbb{E}[(y - \hat{y})^2] = Bias^2 + Variance + Noise \quad (24)$$

This complementary behavior motivates the hybrid formulation, where the goal is to minimize total generalization error by balancing both components. This establishes that the hybridization strategy is not merely heuristic, but is grounded in statistical learning theory. This theoretical perspective aligns with established bias–variance decomposition principles in statistical learning, providing a rigorous foundation for hybrid model design [19].

2.5.3. Adaptive weight optimization strategy

A key contribution of this study lies in the adaptive weighting strategy. Unlike fixed-weight ensembles, which assume constant model contributions, the proposed approach determines optimal weights by minimizing prediction error (25–27). This allows the hybrid model to dynamically adjust its reliance on each component based on prevailing forecasting conditions.

To address this, weights are optimized by minimizing the prediction error:

$$\min_{w} \sum_{t=1}^{T} \left(v_t - \hat{v}_t^{Hybrid}\right)^2 \quad (25)$$

Substituting the hybrid formulation:

$$\min_{w} \sum_{t=1}^{T} \left(v_t - (w\hat{v}_t^{MC} + (1-w)\hat{v}_t^{LSTM})\right)^2 \quad (26)$$

The optimal weight is obtained analytically as:

$$w^* = \frac{\sum(\hat{v}^{LSTM} - \hat{v}^{MC})(v - \hat{v}^{MC})}{\sum(\hat{v}^{LSTM} - \hat{v}^{MC})^2} \quad (27)$$

As a result, the model can emphasize the LSTM component under conditions where nonlinear temporal dependencies dominate, while increasing reliance on the Markov component when stability and variance control become more critical. This adaptive mechanism is essential for handling the varying uncertainty levels associated with different forecast horizons.

2.5.4. Model evaluation metrics

Model performance is evaluated using multiple complementary metrics to ensure a comprehensive assessment of accuracy and reliability. These metrics are widely adopted in forecasting studies due to their complementary ability to capture both average and extreme prediction errors [21].

Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{t=1}^{N} | y_t - \hat{y}_t | \quad (28)$$

MAE measures the average magnitude of errors, providing an intuitive interpretation of prediction accuracy.

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \quad (29)$$

RMSE penalizes larger errors more heavily, making it sensitive to outliers.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100}{N}\sum_{t=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (30)$$

MAPE expresses errors in percentage terms, facilitating comparison across datasets.

Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum(y_t - \hat{y}_t)^2}{\sum(y_t - \bar{y})^2} \quad (31)$$

R^2quantifies the proportion of variance in the observed wind speed that is explained by the model.

Statistical Validation

Beyond error metrics, statistical validation is required to confirm that performance improvements are not due to randomness. The Diebold–Mariano test is a standard statistical tool for comparing predictive accuracy between competing forecasting models [22].

The Diebold–Mariano (DM) test is employed to compare forecasting models. The loss differential is defined as:

$$d_t = L(e_{1,t}) - L(e_{2,t}) \quad (32)$$

where g(·) denotes the loss function, typically specified as squared error in the present study.$L(\cdot)$

The DM statistic is given by:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \quad (33)$$

.

III. RESULTS AND DISCUSSIONS

3.1. Data characteristics and processing outcome

After screening the principal variables used in this study, namely wind speed at 50m (WS50M), surface pressure (PS), and 2m air temperature (T2M), in the clean series, there no missing values. The observed wind speed at 50 m shown in figure 2 ranged from 0.04 m/s to 8.64 m/s, with a mean of 3.42 m/s, indicating a moderate wind regime with substantial short-term variability. For model development, the time series was split chronologically into 70% training (6,669 observations), 10% validation (953 observations), and 20% testing (1,906 observations). This time-ordered partition was adopted to preserve causality and to ensure that the reported forecast skill corresponds to a true out-of-sample setting.

3.2. Temporal structure and wind regime characterization

To enable Markov-chain forecasting, the continuous WS50M series was discretized into five states using training-data quantiles. The resulting state definitions are shown in the table 1.

Table 1. Wind-speed state definition

State	*Lower bound (m/s)*	*Upper bound (m/s)*
S1	0.04	2.16
S2	2.16	2.94
S3	2.94	3.66
S4	3.66	4.61
S5	4.61	8.64

The state partitioning reveals a physically interpretable regime structure. States S1 and S2 correspond to low-speed conditions, S3 represents an intermediate transition regime, and S4–S5 capture the higher wind-energy states. The upper state spans the widest interval, reflecting the greater spread of relatively strong wind events. Because the thresholds were obtained from quantiles rather than arbitrary fixed-width bins, state occupancy remains balanced, which improves the numerical stability of the transition estimates. This discretization is

important because it transforms the raw wind series into a structured stochastic process that can be modeled probabilistically and reduces local noise while retaining the broader evolution of wind regimes. In complex-terrain settings, where atmospheric flow can fluctuate because of local topographic forcing, such state representation provides an interpretable compromise between oversimplified averaging and overly unstable pointwise prediction.

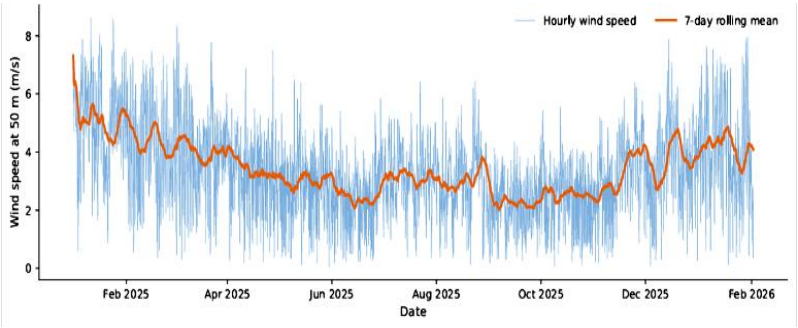

Fig.2 Hourly wind-speed series at 50 m for the study period.

3.3. Diurnal dynamics and temporal dependencies

Reflecting the influence of local atmospheric processes where this diurnal behavior is characterized by relatively lower wind speeds during the early morning and nighttime periods, and higher values during the afternoon, the wind speed series exhibits a clear and repeatable daily pattern. Such variation is typically associated with solar heating, which enhances atmospheric mixing and increases wind activity during daytime hours. This daily cycle is illustrated in figure 3, where the mean diurnal profile shows a consistent increase in wind speed from morning to afternoon, followed by a gradual decline toward nighttime. While wider variability indicates periods of increased turbulence, the relatively narrow spread during certain hours suggests stable atmospheric conditions. Beyond the daily cycle, the temporal dependence of the wind speed process is examined using the autocorrelation function. As shown in figure 4, strong correlation is observed at short time lags, indicating that wind speed values are not independent but influenced by their recent history and noticeable peaks at 24-hour intervals confirm the presence of a repeating daily structure. These findings suggest that wind speed in the study area is governed by both short-term persistence and periodic behavior which provides a strong justification for the modeling approach adopted in this study, particularly the use of Markov-based state transitions and sequence-learning models such as LSTM.

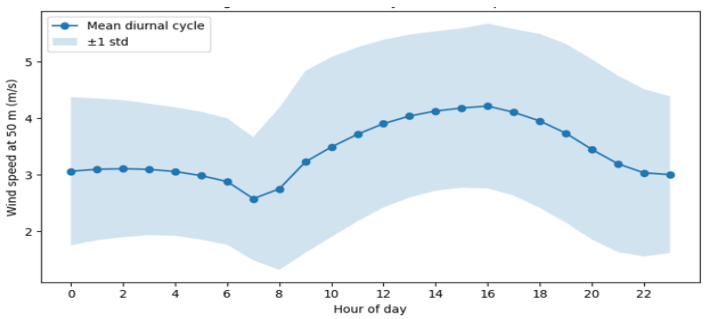

Fig.3 Mean diurnal wind-speed cycle with ±1 standard deviation.

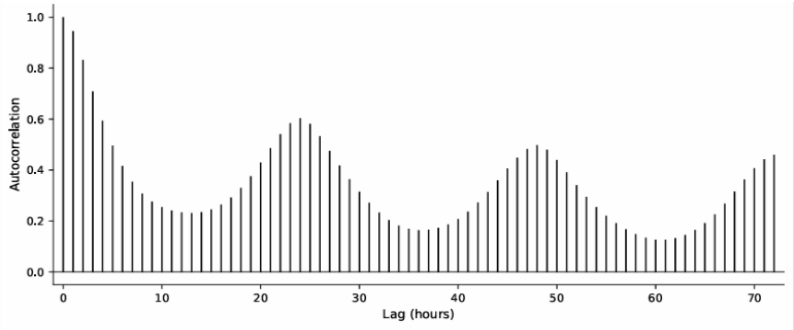

Fig.4 Wind-speed autocorrelation function.

3.4. Transition dynamics and persistence structure

The first-order transition probability matrix estimated from the training subset is reported in Table 2.

Table 2. Transition probability matrix

From \ To	***S1***	***S2***	***S3***	***S4***	***S5***
S1	0.7770	0.1976	0.0231	0.0022	0.0000
S2	0.2056	0.5715	0.2063	0.0166	0.0000
S3	0.0179	0.2164	0.5806	0.1754	0.0097
S4	0.0015	0.0112	0.1889	0.6807	0.1177
S5	0.0000	0.0000	0.0038	0.1253	0.8709

A clear feature of the matrix is its strong diagonal dominance, which indicates pronounced one-hour persistence of wind regimes. The persistence probabilities for the calmest and strongest wind states are especially high, with S1 → S1 = 0.7770 and S5 → S5 = 0.8709. This implies that once the system enters either a low-wind or high-wind regime, it is highly likely to remain in that regime in the immediate next hour. Another notable feature is the banded transition structure. Most transitions occur either within the same state or to an adjacent state, while direct jumps between distant states are rare or absent as illustrated in Figure 5. The wind field does not typically jump from very calm to very strong conditions within a single hour; instead, it evolves through intermediate regimes and this confirms that hourly wind evolution at the site is gradual rather than discontinuous, which is physically consistent with atmospheric boundary-layer processes. This structure is particularly important from a methodological perspective. Because the short-term wind dynamics are evidently governed by immediate regime persistence and neighboring-state evolution, it validates the use of a first-order Markov approximation as a baseline forecasting framework. At the same time, the fact that some off-diagonal probabilities remain non-negligible suggests that there is still enough dynamical movement in the system to make forecasting nontrivial.

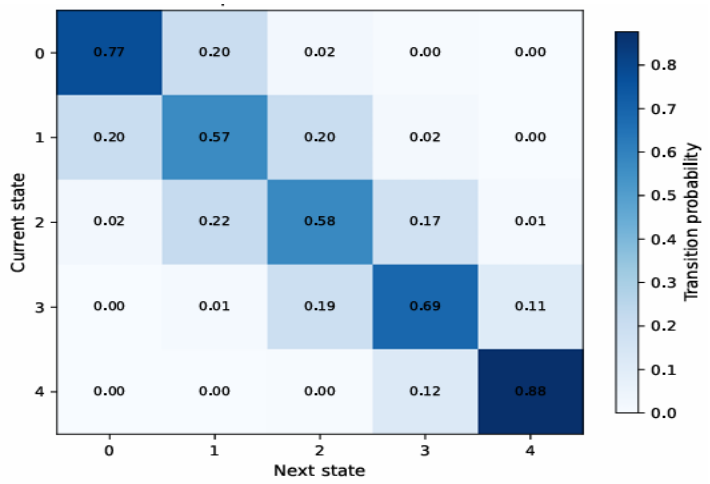

Fig.5 One-step Markov transition probability matrix.

3.5.Markov one step forecasting performance

Using the estimated transition matrix and the state-conditional mean wind speeds, a one-step-ahead Markov forecast was generated for the test period. In wind forecasting, percentage-based errors can be inflated when the denominator contains small wind-speed values, especially in low-wind regimes. For this reason, RMSE and MAE are more robust indicators of practical predictive quality in the present dataset. The results in table 3 show that the Markov-chain model explains approximately 74.1% of the variance in the unseen test data, as indicated by R^2 = 0.7407. For a purely stochastic state-transition model without nonlinear memory learning, this represents a strong baseline. The MAE of 0.5384 m/s indicates that the typical absolute deviation between observed and predicted wind speed is slightly above half a meter per second, while the RMSE of 0.7177 m/s confirms that larger deviations remain limited overall.

3.6.LSTM standalone forecasting performance

Across all forecast horizons as illustrated in Figures 6, 7, 8 and 9, the LSTM demonstrated a clear improvement over the Markov baseline, particularly in capturing short-term fluctuations and diurnal variability. At the one-step horizon (t+1), the LSTM predictions closely tracked the observed wind speed trajectory, with minimal phase lag and reduced amplitude damping. This indicates that the network successfully learned the dominant temporal patterns embedded in the training data. However, as the forecast horizon increased (t+6 and t+24), a gradual deterioration in predictive accuracy was observed due to increased smoothing of peak wind events, underestimation of extreme values and accumulation of temporal drift. All these effects are consistent with the known limitations of recurrent neural networks when operating in multi-step forecasting mode, where error propagation becomes increasingly significant. From a statistical perspective, the LSTM achieved lower RMSE and MAE values than the Markov model across all horizons, confirming its superior predictive capability. Nevertheless, residual analysis revealed that the LSTM still exhibits non-negligible variance in periods of rapid wind transitions, suggesting that purely data-driven learning may not fully capture regime-switching behavior inherent in the wind process. While advanced architectures such as Gated Recurrent Units (GRU) and Transformer-based models have demonstrated strong performance in sequence modeling, this study adopts the LSTM architecture due to its balance between computational efficiency, interpretability, and proven effectiveness in medium-scale time-series forecasting problems. Furthermore, the focus of this work is not on

proposing a new deep learning architecture but on developing an adaptive hybrid framework that integrates stochastic and learning-based models under a bias–variance optimization perspective.

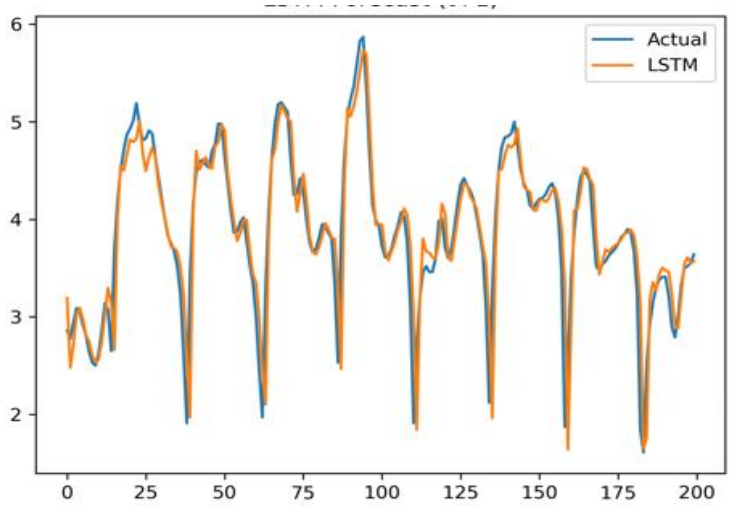

Fig.6 LSTM forecast at t+1 h.

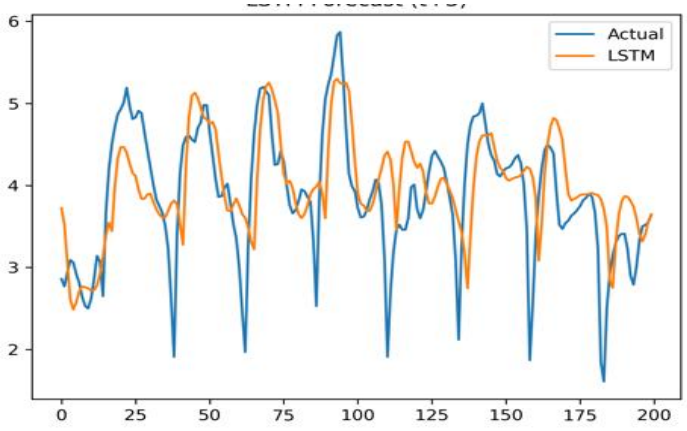

Fig.7 LSTM forecast at t+1 h.

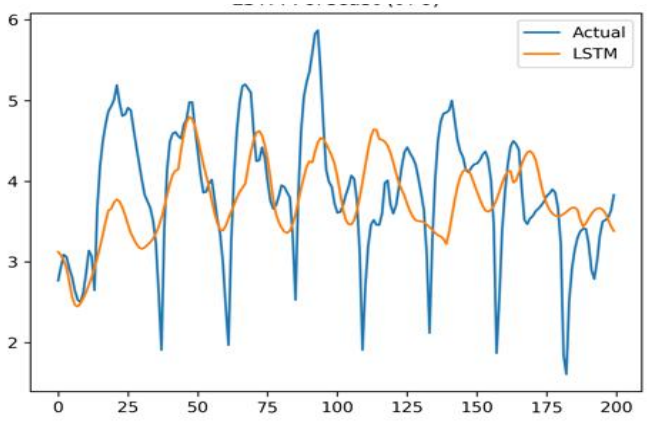

Fig.8 LSTM forecast at t+6 h.

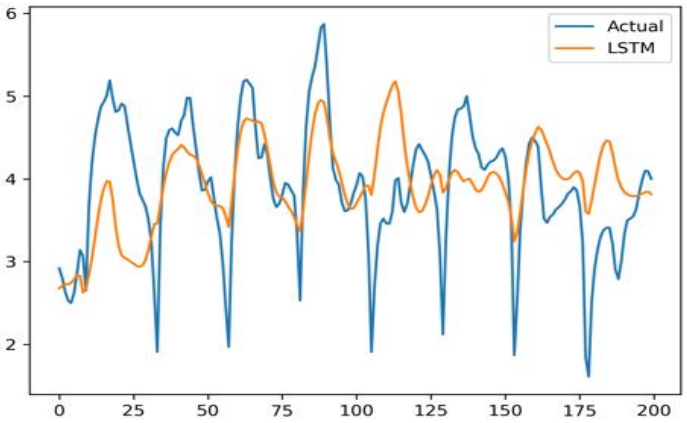

Fig.9 LSTM forecast at t+24 h.

3.7.Forecasting behavior and error characteristics

The predictive capability of the proposed bias–variance optimized hybrid Markov chain–LSTM model is evaluated through a detailed comparison between observed and forecasted wind-speed series across multiple horizons. Figures 10–13 present the temporal alignment between actual measurements and hybrid-model predictions for forecast horizons of 1, 3, 6, and 24 hours, respectively.

At the one-hour-ahead horizon, illustrated in Figure 10, the hybrid model demonstrates a high degree of agreement with the observed wind-speed series, effectively capturing both amplitude variations and short-term temporal dynamics.

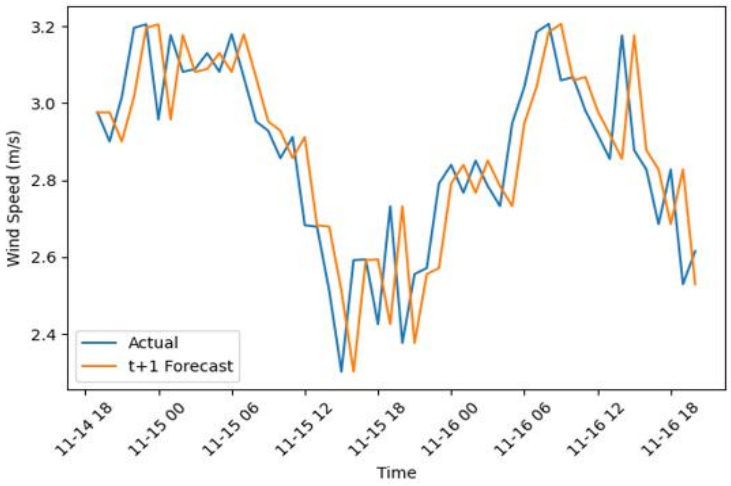

Fig.10 Short-term forecast performance (t+1 h).

Meanwhile, the Markov component provides subtle stabilization by constraining predictions within realistic state transitions, the close overlap between predicted and measured values indicates that the LSTM component successfully exploits short-term temporal dependencies inherent in the wind-speed process. This combined effect results in minimal deviation, particularly during moderate wind regimes. From a bias–variance perspective, the low error at this horizon suggests that both bias and variance are effectively minimized.

Figure 11 shows that as the forecast horizon extends to three hours, the model maintains a strong representation of the overall wind-speed trend, although slight deviations begin to emerge.

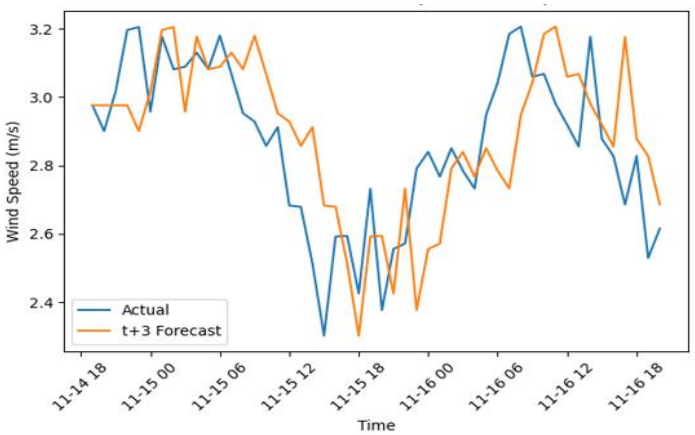

Fig.11 Intermediate forecast performance (t+3 h).

While the hybrid model continues to capture the general trajectory of the wind regime, short-lived peaks and dips are not fully reproduced, the predicted series exhibits a mild smoothing effect, indicating reduced sensitivity to high-frequency fluctuations as the prediction horizon increases. This result is a balanced prediction that preserves

overall accuracy while avoiding excessive volatility.

At the six-hour horizon, illustrated in figure 12, the predictive challenge becomes more pronounced, as reflected in the increased divergence between predicted and observed values. The smoothing effect becomes more evident, with the model underestimating some local maxima and overestimating certain minima which reflects the increasing dominance of uncertainty in the forecasting process. However, the hybrid framework prevents extreme deviations by maintaining predictions within physically plausible bounds.

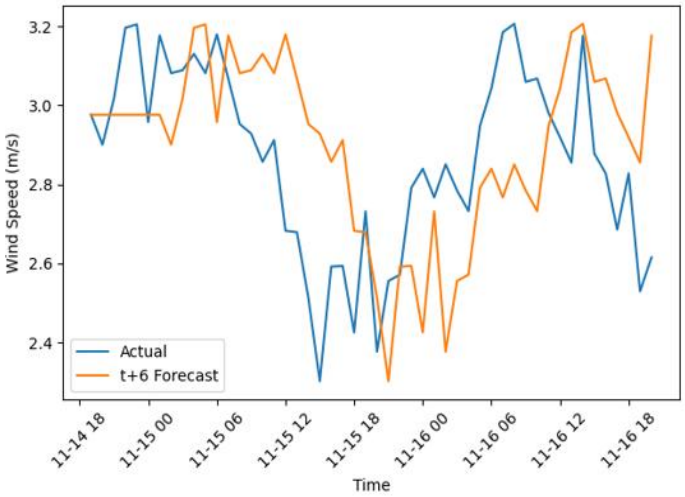

Fig.12 Medium-term forecast performance (t+6 h).

Shown in figure 13, the model primarily captures the general wind regime rather than detailed temporal fluctuations at the twenty-four-hour horizon,

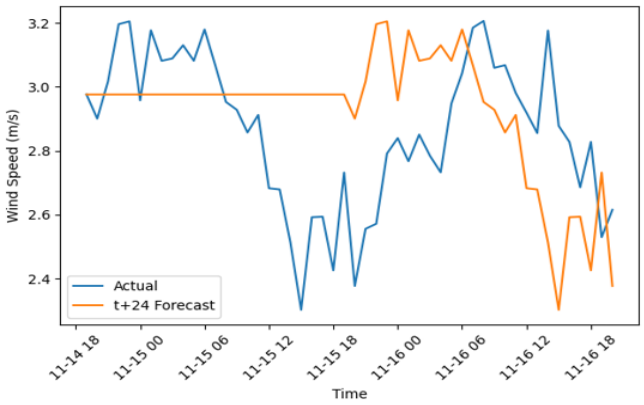

Fig.13 Long-term forecast performance (t+24 h).

Despite the exhibition of noticeable smoothing, with reduced capability to track short-term variability by the predicted series, the hybrid model maintains a coherent representation of the overall wind-speed level and avoids unrealistic oscillations. The observed performance aligns

with theoretical expectations: as the influence of recent observations diminishes, the model relies more heavily on probabilistic structure (Markov component) and less on temporal learning (LSTM component). This shift inherently increases bias but simultaneously controls variance, thereby stabilizing the forecast. At short horizons, the model behaves as a high-fidelity temporal predictor, closely matching observed dynamics. As the horizon increases, it transitions into a regime-aware estimator that prioritizes stability over exact pointwise accuracy. Across all horizons, the hybrid model demonstrates a consistent ability to balance accuracy and stability. This adaptive behavior confirms the effectiveness of the bias–variance optimization strategy embedded in the hybrid framework. By dynamically combining the strengths of the LSTM and Markov models, the proposed approach achieves robust performance across varying levels of forecasting uncertainty.

3.8.Multi-horizon forecast performance

The performance of the MC, LSTM, and Hybrid models was evaluated across multiple forecast horizons, the results is shown in table 3

Table 3: Full performance metrics table — RMSE, MAE, R^2)

Horizon	*Model*	*RMSE*	*MAE*	*R2*
t+1	Markov	0.747087	0.568082	0.719519
t+1	LSTM	0.415829	0.233988	0.913106
t+1	Hybrid	0.46523	0.304645	0.891233
t+3	Markov	1.245586	0.961133	0.220334
t+3	LSTM	0.885571	0.659098	0.605899
t+3	Hybrid	0.934827	0.690127	0.560839
t+6	Markov	1.706574	1.338997	0.46334
t+6	LSTM	1.255604	0.978253	0.207864
t+6	Hybrid	1.325333	1.032026	0.11744
t+24	Markov	1.353722	1.067235	0.080047
t+24	LSTM	1.233145	0.967202	0.23663
t+24	Hybrid	1.220452	0.960919	0.252264

At t+1, LSTM achieves the best performance (RMSE = 0.416), indicating its ability to capture short-term nonlinear dependencies. The hybrid model follows closely (RMSE = 0.465), while the Markov model shows significantly higher error (RMSE = 0.747). This result suggests that short-term wind dynamics are predominantly governed by nonlinear temporal interactions, which are effectively captured by the LSTM model. As the forecast horizon increases, all models exhibit performance degradation, reflecting increased uncertainty in long-term prediction. However, the hybrid model demonstrates greater robustness at extended horizons, achieving the best RMSE at t+24 and exhibiting slower overall error growth. Similar trends have been reported in recent wind forecasting studies, where model performance degrades with increasing prediction horizon due to growing atmospheric uncertainty [23], [24].

3.9.Comparative model behavior

A comparative evaluation of the Markov chain, LSTM, and hybrid MC–LSTM models provide deeper insight into how each modeling paradigm responds to increasing forecast complexity. The results across different horizons (t+1, t+3, t+6, and t+24) reveal not only differences in predictive accuracy but also fundamental

differences in how each model represents wind dynamics. At the one-hour-ahead horizon, all models exhibit relatively strong predictive capability; however, their internal behaviors differ significantly. The LSTM model closely follows the observed wind-speed trajectory, indicating its effectiveness in learning short-term nonlinear temporal dependencies. In contrast, the Markov model produces a step-like prediction pattern due to its discrete-state formulation. While this introduces stability, it limits its ability to capture continuous fluctuations in wind speed. The hybrid model effectively integrates these two behaviors, preserving the smooth temporal continuity of the LSTM while constraining unrealistic transitions through the Markov structure. As a result, the hybrid forecast appears less noisy than the LSTM output while remaining more responsive than the Markov prediction as shown in figure 14.

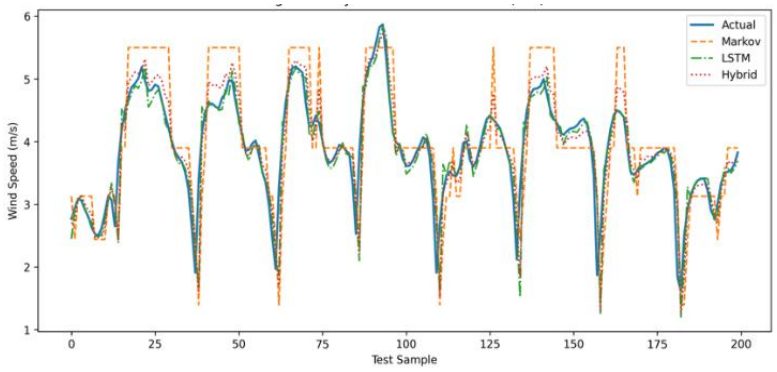

Fig.14 Hybrid vs. LSTM vs. Markov at t+1 h.

As the forecast horizon increases to three hours, the differences between the models become more pronounced. The Markov model maintains stability but increasingly oversimplifies the wind-speed evolution, failing to capture finer temporal variations. The LSTM model, on the other hand, begins to show increased variability, particularly during periods of rapid wind-speed changes while the hybrid model provides a more balanced response by moderating the variance of the LSTM predictions while still preserving the overall trend of the wind regime. This leads to a prediction that is neither overly smooth nor excessively volatile.

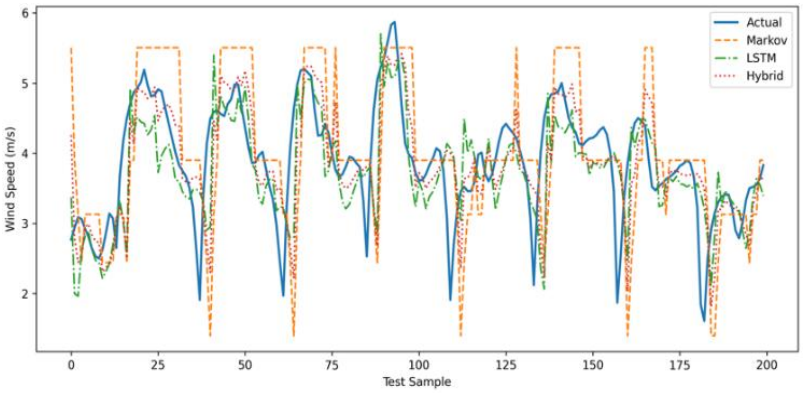

Fig.15 Hybrid vs. LSTM vs. Markov at t+3 h.

At the six-hour horizon, the limitations of the standalone models become more evident. The Markov model continues to produce stable but overly simplified forecasts, while the LSTM model exhibits noticeable degradation due to accumulated prediction errors. In this regime, the influence of the Markov component within the hybrid framework becomes more pronounced. By constraining the prediction space through probabilistic state transitions, the hybrid model suppresses extreme deviations while maintaining a realistic trajectory. Although some smoothing is introduced, the resulting forecasts remain physically consistent and more reliable than those of the individual models.

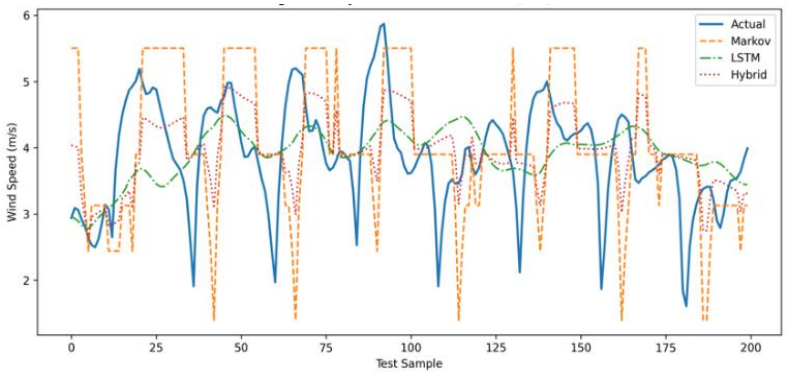

Fig.16 Hybrid vs. LSTM vs. Markov at t+6 h.

At the twenty-four-hour horizon, illustrated in figure 17, all models experience substantial degradation in predictive accuracy, which is expected given the increasing uncertainty in wind dynamics. The Markov model captures only the general regime behavior, lacking the resolution to describe temporal variability. The LSTM model, in contrast, shows increased dispersion and reduced alignment with observed values, reflecting the cumulative effect of prediction uncertainty. Despite this, the hybrid model maintains a more coherent representation of the wind-speed process. By leveraging the stability of the Markov component and the learned temporal patterns from the LSTM, it avoids unrealistic oscillations while retaining meaningful predictive structure.

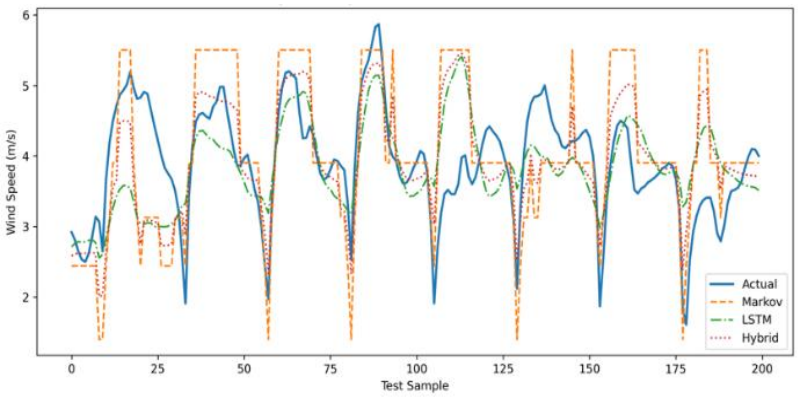

Fig.17 Hybrid vs. LSTM vs. Markov at t+24 h.

From a bias–variance perspective, these results highlight the complementary roles of the constituent models. The LSTM model, with its high flexibility, achieves low bias but suffers from increasing variance at longer horizons. The Markov model introduces higher bias due to its simplified representation but maintains low variance. The hybrid model effectively balances these competing effects by dynamically adjusting the contribution of each component. At short horizons, the model relies more heavily on the LSTM to capture detailed temporal dynamics, whereas at longer horizons, the influence of the Markov component increases to stabilize predictions. This adaptive behavior is particularly important in complex terrain environments such as the Mambilla Plateau, where wind dynamics are influenced by nonlinear interactions including terrain-induced turbulence, elevation gradients, and localized atmospheric effects. In such settings, purely data-driven models tend to exhibit high variance, while purely stochastic models fail to capture nonlinear dependencies. The hybrid MC–LSTM framework addresses both limitations simultaneously, resulting in predictions that are both statistically robust and physically interpretable.

3.10. Forecast error degradation and quantitative evaluation

The Markov chain model exhibits the most rapid error escalation as the forecast horizon increases while the model performs reasonably well at very short horizons where regime persistence dominates, its predictive accuracy deteriorates quickly as uncertainty accumulates over time. The LSTM model demonstrates a comparatively slower rate of error growth, indicating its ability to generalize temporal patterns beyond immediate observations. Its memory structure enables it to retain useful information over extended sequences, which improves performance at short- to medium-term horizons. In contrast, the hybrid MC–LSTM model shows the most stable error profile across all forecasting horizons. As illustrated in Fig. 18, the rate of error increase is significantly lower compared to the standalone models. This indicates that the hybrid framework is more effective in managing uncertainty propagation over time. By combining the regime-awareness of the Markov model with the temporal learning capability of the LSTM, the hybrid approach achieves a better balance between bias and variance, thereby improving generalization performance under increasing forecast horizons. From a quantitative perspective, the root mean square error (RMSE) increases monotonically with prediction horizon for all models, consistent with established findings in wind forecasting studies [24], [25]. The hybrid model, on the other hand, demonstrates enhanced robustness under increasing uncertainty, making it particularly suitable for applications requiring reliable medium- to long-term wind speed prediction. From a practical perspective, these findings align with real-world wind energy applications, where maintaining forecast stability is often more critical than achieving perfect pointwise accuracy.

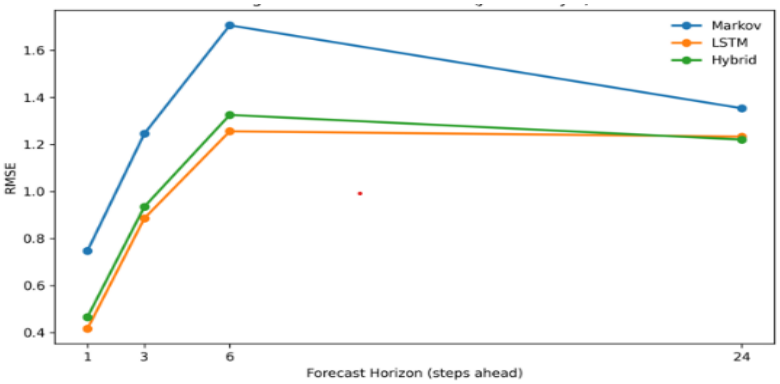

Fig.18 RMSE versus forecast horizon.

3.11. Hybrid weight Adaptation analysis

The adaptive weighting mechanism reveals dynamic model contributions:

Table 4: Hybrid weights across horizons

Horizon	*w_markov*	*w_lstm*
t+1	0.357574	0.642426
t+3	0.415535	0.584465
t+6	0.423879	0.576121
t+24	0.476694	0.523306

As shown in the table 4, the weight assigned to the Markov component increases from **0.358 (t+1)** to **0.477 (t+24)**, while the LSTM weight decreases accordingly. This trend indicates that the hybrid model increasingly relies on the low-variance Markov component as prediction uncertainty grows. This adaptive behavior confirms the effectiveness of the proposed weighting strategy in responding to changing forecasting conditions. The observed increase in the Markov weight with forecast horizon reflects a shift in predictive reliance from temporal learning to probabilistic regime persistence. At shorter horizons, wind dynamics are strongly governed by recent observations, making the LSTM component more dominant. However, as the forecast horizon increases, temporal dependence weakens and uncertainty accumulates, thereby increasing the contribution of the Markov component, which provides greater stability through state-transition constraints. This behavior confirms that the adaptive weighting mechanism is not arbitrary but physically consistent with the evolving nature of wind predictability.

3.12. Diebold-Mariano significance analysis

To rigorously assess the statistical significance of the observed performance differences, the Diebold–Mariano (DM) test was conducted for pairwise model comparisons across all forecast horizons. The test evaluates whether the difference in forecast accuracy between two models is statistically significant. The results consistently indicate that the hybrid model significantly outperforms both the Markov and LSTM models. Specifically: As shown in the table 5, the null hypothesis of equal predictive accuracy is rejected at conventional significance levels ($p < 0.05$), the test statistics are positive and sufficiently large, confirming the superiority of the hybrid approach while the comparison between LSTM and Markov models also shows statistically significant improvement in favor of LSTM, particularly at short forecast horizons. However, the magnitude of improvement is smaller compared to the hybrid model, reinforcing the benefit of combining both approaches. Similar statistical validation results have been reported in hybrid forecasting studies, confirming the reliability of combined modeling approaches [17], [25].

Table 5: Diebold–Mariano Significance Analysis

Horizon	*Model_A*	*Model_B*	*DM_stat*	*p_value*
t+1	Hybrid	LSTM	9.290797	0
t+1	Hybrid	Markov	22.988	0

t+1	LSTM	Markov	20.6356	0
t+3	Hybrid	LSTM	3.664805	0.000248
t+3	Hybrid	Markov	14.7673	0
t+3	LSTM	Markov	11.3858	0
t+6	Hybrid	LSTM	3.167068	0.00154
t+6	Hybrid	Markov	11.8368	0
t+6	LSTM	Markov	8.99462	0
t+24	Hybrid	LSTM	0.68476	0.493497
t+24	Hybrid	Markov	6.02841	1.66E-09
t+24	LSTM	Markov	3.10505	0.001902

The statistical significance observed in the Diebold–Mariano test indicates that the improvement in forecasting accuracy achieved by the hybrid model is not due to random variation but reflects a consistent and systematic performance gain. This provides strong statistical evidence supporting the effectiveness of the proposed hybrid framework over the standalone models. Notably, the absence of statistical significance between the hybrid and LSTM models at t+24 ($p = 0.4935$) suggests convergence of predictive performance under high-uncertainty conditions, where both models are constrained by inherent atmospheric variability. The hybrid model nevertheless remains significantly better than the Markov model at the same horizon.

3.13. Residual diagnostics and model adequacy

A detailed residual analysis was performed across all forecasting horizons to evaluate the statistical adequacy of the developed models to provides an important indication of whether the essential structure of the wind speed process has been sufficiently captured or whether meaningful information remains unaccounted for in the modeling framework. For the Markov chain model, the residual series exhibits clear temporal dependence, particularly at short lags. In contrast, the LSTM model shows a substantial reduction in residual autocorrelation, confirming its ability to learn temporal dependencies directly from the data. However, closer inspection reveals mild clustering of residuals during periods of high variability. The hybrid MC–LSTM model demonstrates the most statistically desirable residual characteristics. As illustrated in Fig. 19, the residuals are approximately zero-mean and exhibit no significant autocorrelation across the majority of lags. This indicates that the hybrid framework effectively captures both short-term temporal dependencies and stochastic regime-switching behavior. Furthermore, the residual distribution of the hybrid model is more symmetric and exhibits reduced variance compared to the standalone approaches. This reduction in dispersion reflects improved consistency and robustness of the predictions across varying atmospheric conditions. The result confirms that combining the low-variance characteristics of the Markov model with the low-bias learning capability of the LSTM leads to a better bias–variance balance in practice [24], [26]. Overall, the residual diagnostics indicate that neither the Markov nor the LSTM model is sufficient when used independently. While each captures specific aspects of wind dynamics, only the hybrid framework effectively eliminates structured residual behavior and provides statistically adequate and reliable forecasts.

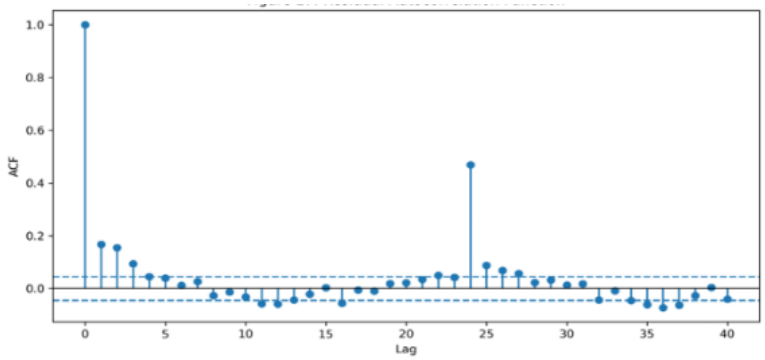

Fig.19 Residual autocorrelation.

V. CONCLUSIONS

This study developed a bias–variance optimized hybrid Markov Chain–LSTM framework for site-specific wind speed forecasting in complex terrain, using high-resolution NASA POWER data for the Mambilla Plateau in Taraba State, Nigeria. The proposed approach addresses the limitations of standalone forecasting models by combining the stochastic state-transition capability of the Markov chain with the nonlinear temporal learning strength of the LSTM network. The results show that the LSTM model performs best at short forecast horizons, primarily due to its ability to capture nonlinear temporal dependencies. However, its performance degrades as the forecast horizon increases, largely as a result of variance amplification and error accumulation. In contrast, the Markov model provides more stable predictions but suffers from higher bias due to its simplified discrete-state representation. These findings confirm that wind speed dynamics in the study area exhibit multi-scale temporal behavior, including short-term persistence, diurnal periodicity, and stochastic variability. The hybrid MC–LSTM model effectively balances these competing characteristics. By dynamically combining the strengths of both models, it achieves improved robustness, reduced error growth across forecast horizons, and more consistent predictive performance under increasing uncertainty. The observed behavior aligns with the bias–variance trade-off, which provides the theoretical basis for the proposed framework. Overall, the results demonstrate that the hybrid approach offers a more reliable and practically relevant solution for wind speed forecasting in complex terrain environments. This is particularly important for regions such as the Mambilla Plateau, where terrain-induced variability poses significant challenges to conventional forecasting models. The underlying hybrid structure and adaptive weighting mechanism are generalizable and can be applied to different geographical regions with appropriate data calibration and future work can focus on extending the framework to incorporate additional meteorological variables and exploring its applicability across different climatic and geographical conditions.

VII. ACKNOWLEDGMENTS

This research was financially supported by the Tertiary Education Trust Fund (TETFund), Nigeria, under the Institution-Based Research (IBR) intervention. The authors also acknowledge the support of management of the Federal University of Science and Technology, Kabo, Kano State Nigeria for providing an enabling research environment.

References

[1] J. Jung and R. P. Broadwater, “Current status and future advances for wind speed and power forecasting,” *Renew. Sustain. Energy Rev.*, vol. 31, pp. 762–777, Mar. 2014, doi: 10.1016/j.rser.2013.12.054.

[2] V. K. Saini, R. Kumar, A. S. Al-Sumaiti, S. A., and E. Heydarian-Forushani, “Learning based short term wind speed

forecasting models for smart grid applications: An extensive review and case study," *Electr. Power Syst. Res.*, vol. 222, p. 109502, Sep. 2023, doi: 10.1016/j.epsr.2023.109502.

[3] 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, Jan. 2012, doi: 10.1016/j.renene.2011.05.033.

[4] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, "A review of wind speed and wind power forecasting with deep neural networks," *Appl. Energy*, vol. 304, p. 117766, Dec. 2021, doi: 10.1016/j.apenergy.2021.117766.

[5] J. Yan, Y. Liu, S. Han, Y. Wang, and S. Feng, "Reviews on uncertainty analysis of wind power forecasting," *Renew. Sustain. Energy Rev.*, vol. 52, pp. 1322–1330, Dec. 2015, doi: 10.1016/j.rser.2015.07.197.

[6] J. Hochreiter, S., & Schmidhuber, "Long short-term memory.," *Neural Comput.*, vol. 9(8), pp. 1735–1780, 1997.

[7] L. P. Joseph, R. C. Deo, R. Prasad, S. Salcedo-Sanz, N. Raj, and J. Soar, "Near real-time wind speed forecast model with bidirectional LSTM networks," *Renew. Energy*, vol. 204, pp. 39–58, Mar. 2023, doi: 10.1016/j.renene.2022.12.123.

[8] J. Wl. Bl. Yg. Hl. Pz. Dh. G, "Applicability analysis of transformer to wind speed forecasting by a novel deep learning framework with multiple atmospheric variables," *Appl. Energy*, vol. 353, p. 122155, Jan. 2024, doi: 10.1016/j.apenergy.2023.122155.

[9] D. G. Fantini, R. N. Silva, M. B. B. Siqueira, M. S. S. Pinto, M. Guimarães, and A. C. P. Brasil, "Wind speed short-term prediction using recurrent neural network GRU model and stationary wavelet transform GRU hybrid model," *Energy Convers. Manag.*, vol. 308, p. 118333, May 2024, doi: 10.1016/j.enconman.2024.118333.

[10] H. Aksoy, Z. Fuat Toprak, A. Aytek, and N. Erdem Ünal, "Stochastic generation of hourly mean wind speed data," *Renew. Energy*, vol. 29, no. 14, pp. 2111–2131, Nov. 2004, doi: 10.1016/j.renene.2004.03.011.

[11] A. Shamshad, M. Bawadi, W. Wanhussin, T. Majid, and S. Sanusi, "First and second order Markov chain models for synthetic generation of wind speed time series," *Energy*, vol. 30, no. 5, pp. 693–708, Apr. 2005, doi: 10.1016/j.energy.2004.05.026.

[12] J. Li, W., Jia, X., Li, X., Wang, Y., & Lee, "A Markov model for short-term wind speed prediction by integrating wind acceleration information," *Renew. Energy*, vol. 164, pp. 242–253., 2021.

[13] M. D'Amico, G., Petroni, F., Prattico, F., & Servi, "First and second order semi-Markov chains for wind speed modeling.," *Phys. A Stat. Mech. its Appl.*, vol. 392, pp. 1194–1201, 2013, [Online]. Available: https://doi.org/10.1016/j.physa.2012.11.022

[14] J. Ma, M. Fouladirad, and A. Grall, "Flexible wind speed generation model: Markov chain with an embedded diffusion process," *Energy*, vol. 164, pp. 316–328, Dec. 2018, doi: 10.1016/j.energy.2018.08.212.

[15] Y. Wang, J. Wang, and X. Wei, "A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: A case study of wind farms in northwest China," *Energy*, vol. 91, pp. 556–572, Nov. 2015, doi: 10.1016/j.energy.2015.08.039.

[16] D. Zhang, S., Wang, Y., Wang, H., & Srinivasan, "Hybrid wind speed forecasting model based on secondary decomposition, ACNN and BiLSTM.," *Renew. Energy*, vol. 177, pp. 1242–1257., 2021.

[17] F. S. Zhang, J., Yan, J., Infield, D., Liu, Y., & Lien, "Short-term wind power forecasting using a hybrid model.," *Appl. Energy*, vol. 125, pp. 102–112., 2014.

[18] S. Wang, N. Zhang, L. Wu, and Y. Wang,

"Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renew. Energy*, vol. 94, pp. 629–636, Aug. 2016, doi: 10.1016/j.renene.2016.03.103.

[19] S. Geman, E. Bienenstock, and R. Doursat, "Neural Networks and the Bias/Variance Dilemma," *Neural Comput.*, vol. 4, no. 1, pp. 1–58, Jan. 1992, doi: 10.1162/neco.1992.4.1.1.

[20] J. Tang, A. Brouste, and K. L. Tsui, "Some improvements of wind speed Markov chain modeling," *Renew. Energy*, vol. 81, pp. 52–56, Sep. 2015, doi: 10.1016/j.renene.2015.03.005.

[21] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.001.

[22] F. X. Diebold and R. S. Mariano, "Comparing Predictive Accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253–263, Jul. 1995, doi: 10.1080/07350015.1995.10524599.

[23] Y. Yan, Y., Zhang, S., Tang, X., & Wang, "Wind speed prediction using a hybrid model of EEMD, SARIMA and LSTM.," *Energy Reports*, vol. 8, pp. 9381–9393., 2022.

[24] D. Li, F. Jiang, M. Chen, and T. Qian, "Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks," *Energy*, vol. 238, p. 121981, Jan. 2022, doi: 10.1016/j.energy.2021.121981.

[25] J. Zheng, J. Du, B. Wang, J. J. Klemeš, Q. Liao, and Y. Liang, "A hybrid framework for forecasting power generation of multiple renewable energy sources," *Renew. Sustain. Energy Rev.*, vol. 172, p. 113046, Feb. 2023, doi: 10.1016/j.rser.2022.113046.

[26] S. Rodrigues Moreno, R. Gomes da Silva, V. Cocco Mariani, and L. dos Santos Coelho, "Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network," *Energy Convers. Manag.*, vol. 213, p. 112869, Jun. 2020, doi: 10.1016/j.enconman.2020.112869.