VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2025, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 13, Number 3, July-September 2025 *pp:198-210*

A Novel Approach toward Windspeed Forecasting using an Advanced Deep Learning Framework with Explainable AI

Syed Azeem Inam[1*], **Hassan Hashim**[1], **Asif Mehmood Awan**[1], **Haider Rajput**[1], **Saddam Umer**[1]

[1]Department of Artificial Intelligence and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

Keywords:
Physics-Informed Neural Networks; Wind Speed Prediction; Explainable AI (XAI), SHAP and LIME Interpretation, Renewable Energy Optimization; Climate Modelling.

Journal Info:
Submitted:
August 20, 2025
Accepted:
September 23, 2025
Published:
September 27, 2025

Abstract

Accurate wind speed forecasting is vital for optimizing renewable energy deployment and for advancing our understanding of climate dynamics. Traditional machine-learning approaches often neglect the fundamental physical principles driving atmospheric processes, which limit their robustness and ability to extrapolate beyond the training domain. This investigation introduces an innovative PINN architecture that integrates deep learning techniques with established meteorological theory to improve both predictive fidelity and interpretative clarity. The framework embeds a temperature-sensitive physical constraint directly within the optimization objective. This formulation guarantees that the predictions remain consistent with thermodynamic equilibrium. Structured as a four-layer sequential network with 13 inputs, two hidden (64 neurons each), and a single output unit, the PINN outperformed eight competitive baseline architectures ranging from Bayesian ridge regression to gradient boosting and various hybrid architectures trained on a suite of handcrafted covariates, including wind-shear terms, mean humidity, and temperature-interaction derivatives derived from multi-year records spanning the climatically distinct locales of Badin, Dadu, and Rohri. Observed reductions in mean-squared error and mean absolute error were dramatic, and the coefficient of determination rose to an impressive 0.99. Furthermore, the application of XAI techniques, specifically SHAP and LIME, identified temperature and humidity as the dominant predictors, corroborating the physical consistency of the model while ensuring operational transparency for users. This study establishes an integrative linkage between data-driven learning methodologies and established domain expertise, resulting in a robust and interpretable decision-support tool for both energy system planning and climate impact assessment.

***Correspondence author email address:** syed.azeem@smiu.edu.pk
DOI: 10.21015/vtse.v13i3.2230

1 Introduction

Accurate prognoses of wind speed are not only crucial to the faster implementation of wind power grids but also to the enhanced knowledge of climate dynamics and its extensive environmental impacts on healthcare [1, 2], as wind power plays a pivotal role in the decarbonization initiatives globally and serves as a strategic center when it comes to reducing greenhouse gas emissions and minimizing the severe climate outcomes [3–6]. For wind speed forecasting, the typical ensemble models use either traditional statistical methods or purely data-driven approaches. Although both are highly successful techniques, they have been critically critiqued when subjected to the high standards of predictions due to their highly negligible attentiveness to dynamical consistency. Statistical approaches often overlook the principles of fluid mechanics and thermodynamics describing the winds [7, 8], whereas, given the chaotic and nonlinear nature of the atmosphere, the simpler ML models are also prone to overfitting [9], as they can be well-adjusted to training data, but can cause distortions to the variability that is present in real-world observations, ultimately reducing the accuracy of the model when utilized in a forecast [10].

Owing to the limitations of conventional forecasting models involving boundary sensitivity and computational scalability, there has been a strong emphasis on the utilization of Physics-Informed Neural Networks (PINNs). PINNs enforce physical principles by incorporating weak-form equations by harnessing the power of deep neural networks, resulting in a predictive model that adheres to conservation laws and dissipative phenomena. Recent studies have successfully employed PINNs to wind speed forecasting by embedding Navier-Stokes and thermodynamic equations directly within the architecture, yielding improved robustness to the observational data [11, 12]. This integration of experimental data with the knowledge of thermodynamic principles not only enhances predictive accuracy but also offers richer interpretability, consequently redefining standard procedures within meteorological forecasting.

The present study fuses physical constraints and meteorological insights within the PINN architecture, engineered explicitly for wind speed forecasting across multiple climatic zones. Alongside its focus on prediction fidelity and accuracy, the study incorporates SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) to explain and interpret the outputs of the architecture under consideration. This incorporation enhances the model transparency, allowing the decision-making to be more intelligible and credible to the stakeholders [10]. The dual characterization of the present study lies in the development of PINN architecture while simultaneously embedding XAI capabilities, delivering a holistic wind speed forecasting solution that integrates heightened predictive accuracy with nuanced explanatory strength, synergizing the idea of prioritizing the deployment of trustworthy AI technologies.

2 Related Work

The evolution of wind forecasting methodologies has moved from traditional Numerical Weather Prediction (NWP) systems toward modern, increasingly ML-driven approaches. Initially, these forecasts operated on comprehensive, physics-based models that numerically integrated the governing equations of synoptic-scale motion, boundary-layer flow, and turbulent diffusion, thereby generating wind fields that were temporally and spatially resolved and grounded in canonical meteorological theory. Although such models produced meteorologically plausible wind fields, their deployment in real-time forecasting revealed shortcomings in both accuracy and computational speed, primarily due to their intensive resource consumption and their pronounced responsiveness to perturbations in initial conditions [14, 15]. In response, the forecasting community shifted to hybrid approaches that fused empirical data with physical insight, employing strategies such as Autoregressive Integrated Moving Average (ARIMA) models and Support Vector Machine (SVM) classifiers. These semi-empirical models achieved modest improvements; however, they remain vulnerable to overfitting in data-sparse environments and exhibit

diminished behavior during abrupt regime transitions, thereby underscoring the need for forecasting frameworks that are both generalized across differing climatic conditions and able to adapt dynamically over extended temporal horizons [16, 17].

The integration of deterministic physical modeling with contemporary machine learning techniques has recently emerged as a powerful paradigm for wind speed forecasting, marking a significant milestone in operational meteorology. Physics-Informed Neural Networks (PINNs) epitomize this hybrid strategy, utilizing multi-level deep learning frameworks that seamlessly weave conservation equations and empirical turbulence laws into the loss-driven optimization process. The curricular structure of PINNs requires the governing partial differential equations to be embedded directly within the aggregate loss term, thus rigorously tying the neural network's predictions to conservation principles while retaining the nonlinear expressiveness characteristic of deep neural architectures [18, 19]. Such a design yields a discernible enhancement in the extrapolation capability over diverse aerodynamic regimes and a corresponding gain in the skill of extended multi-hour predictions, achieved by methodically harmonizing measurement sequences with a mechanically consistent dynamical foundation [20]. Despite their long-term potential, PINNs currently face noteworthy barriers regarding interpretability and formal validation. While refinements in predictive fidelity have been documented, the pathways through which the models generate results, along with their performance stability across diverse operational domains, remain insufficiently quantified. Integrating contemporary explainable artificial intelligence (XAI) techniques such as SHAP and LIME into the PINN scaffold may provide valuable insights into internal model behavior and illuminate the degree to which encoded physical constraints shape the outcomes [21, 22]. Concurrently, the existing literature often neglects a rigorous and cohesive fusion of quantitative physical insights into the development of PINN loss functions, commonly relegating the intricate couplings among meteorological quantities to secondary consideration, which is a deficit that compromises the representative fidelity of modeled atmospheric dynamics [16].

This study addresses two key shortcomings in meteorological wind forecasting. Firstly, it integrates temperature-dependent physical constraints directly into the PINN framework, a modification that demonstrably improves both spatial and temporal predictive accuracy, as the reformulated architecture exhibits superior generalization performance, consistently outperforming previous designs. Secondly, it conducts an extensive evaluation of XAI by utilizing SHAP and LIME.

3 Dataset Description and Feature Engineering

This study utilizes meteorological data acquired from the Pakistan Meteorological Department, with observational coverage from three climatically distinct locations, Badin, Dadu, and Rohri, chosen to reflect regional variability in weather behavior. The original records include daily maximum and minimum temperatures, relative humidity at two distinct observation times (midnight and noon), wind speed at the same intervals, cumulative daily rainfall, and the year of measurement. For the specific scope of this study, the objective was to develop a PINN capable of accurately predicting mean daily wind speed. As such, the mean wind speed was selected as the dependent variable and was excluded from the predictor set to avoid direct leakage of target information.

A systematic feature engineering process was employed to expand the dataset beyond its raw meteorological measurements and embed domain-relevant physical insights. One of the earliest transformations involved computing the wind speed difference, defined as the variation in wind speed between noon and midnight, capturing intraday variability in momentum exchange and boundary layer turbulence. This variable enhances the sensitivity of the model to diurnal aerodynamic shifts, which are instrumental in defining convective fluxes and surface-level ventilation. Derived temperature-based features, such as temperature difference and humidity difference, were also included to provide contextual information. These

indirectly influence wind speed patterns through their effect on thermal gradients and pressure differentials, which drive wind generation. Additionally, interaction terms such as humidity-windspeed interaction and windspeed-temperature interaction were constructed to reflect the combined influence of thermodynamic and dynamic processes on atmospheric flow. To model seasonality explicitly, the year variable was transformed into cyclical sine and cosine encodings, thereby enabling the model to capture annual weather cycles without artificial discontinuities between December and January. This allowed PINN to learn long-term temporal dependencies associated with monsoonal transitions and seasonal shifts in pressure belts, both of which affect wind dynamics.

Other relevant augmentations included cumulative rainfall, capturing hydrological memory effects that influence roughness and local pressure systems, rainfall ratio, highlighting the proportional change in precipitation over time, and windspeed-humidity ratio, which approximates the drying power of the air and modifies atmospheric stability. These derived indicators collectively served to model the latent relationships between energy balance, moisture transport, and aerodynamic behavior. By enriching the original dataset with physically meaningful and temporally informed features, the model was equipped to recognize nonlinear, multidimensional relationships underlying the evolution of windspeed. This approach promoted generalizability across distinct climatic zones and preserved the interpretability of model behavior within the context of fundamental atmospheric processes. A detailed list of derived predictor features is provided in Table 1.

4 Methodology

In the present study, a PINN is deployed to forecast wind speed by embedding governing physical laws directly within the structure of a deep learning framework. The architecture is specifically configured to resolve the intricate, nonlinear behavior of the atmospheric dynamics, while simultaneously guaranteeing that the resultant wind patterns conform to conservation laws for momentum and pressure-gradient forces.

The model is organized as a Fully Connected Feed-forward Neural Network comprising three sequential layers, i.e., an input layer, two hidden layers, and a final output layer. The input layer accommodates an input vector of 13 environmental variables, which includes temperature differentials, wind speed, humidity ratios, and rainfall fractions, each of which is hypothesized to exert an influence upon wind speed variability. After ingestion, the input features are propagated through the hidden layers, where they are transformed by the Rectified Linear Unit (ReLU) activation function. The first hidden layer expands the input space to 64 dimensions, whereas the subsequent hidden layer maintains this dimensionality, producing a second 64-dimensional vector. The final output layer condenses this representation to a single scalar that estimates daily wind speed. The overall topology comprises 5,121 tunable parameters. The architecture of the proposed PINN architecture is illustrated in Figure 1, and its model specification is shown in Table 2.

To embed prior scientific knowledge into the learning process, physics-based constraints are integrated directly into the loss function of the proposed model. One such constraint is motivated by the observation that lower temperatures are typically associated with calmer atmospheric states, notably reduced wind activity. To reflect this relationship, the study imposes the following conditional constraint:

$$\hat{W}^{(i)}_{mean} \leq 1 \quad \text{if} \quad T^{(i)}_{mean} < 10$$

This is encoded in the loss function using an indicator mask that activates only for samples where the mean temperature is below 10°C. The corresponding wind constraint loss term is defined as:

$$L_{wind} = \frac{1}{n} \sum_{i=1}^{n} 1_{T^{(i)}_{mean}<10} \left(\max(\hat{W}^{(i)}_{mean} - 1, 0) \right)^2$$

Here, $1_{T^{(i)}_{mean}<10}$ is the indicator function, selecting only those instances where the physical constraint should apply.

Table 1. Engineered Features for the Dataset with Computational Formulas

Feature	Formula
Wind Difference	Windspeed at 1200 – Windspeed at 0000
Mean Temperature	(Max Temperature + Min Temperature)/2
Mean Humidity	(Humidity at 0000 + Humidity at 1200)/2
Temp Difference	Max Temperature – Min Temperature
Rainfall Ratio	Rainfall / Maximum Rainfall
Year Encoding	Sine & Cosine Components
Wind-Humidity Ratio	Mean Wind / (Mean Humidity + ε)
Interaction Features	Multiplicative Combinations

Table 2. Specification of the Proposed Neural Network

Layer	Type	Input Shape	Output Shape	Parameters	Activation	Purpose
Input	-	(batch, 13)	(batch, 13)	0	-	Feature ingestion
Dense-1	Linear	(batch, 13)	(batch, 64)	896	ReLU	Feature transformation
Dense-2	Linear	(batch, 64)	(batch, 64)	4,160	ReLU	Non-linear mapping
Output	Linear	(batch, 64)	(batch, 1)	65	Linear	Windspeed prediction

The overall composite loss function used to train the PINN model integrates the standard data-driven mean squared error loss with these physics-based terms:

$$L_{total} = L_{data} + \lambda_{wind} L_{wind}$$

where λ_{wind} is a hyperparameter controlling the weight of the physics loss, this composite loss ensures that the model learns patterns that are not only statistically robust but also consistent with established physical relationships.

The architecture was trained using the Adam optimizer configured with a learning rate of 0.01 and a training horizon of 500 epochs. This optimizer was selected due to its efficacy in accelerating convergence across deep learning models and its robustness in computing unbiased gradient estimates. Entire datasets were processed as single batches in every training iteration. Such a full-batch approach guarantees that the computed gradients reflect the global landscape of the loss surface and consequently yield consistent and stable updates. The ensuing exposition aims to provide a detailed account of the training mechanics. To reconcile data fidelity with underlying physical constraints in the loss composition, a physics weighting factor of = 0.1 was applied. This scalar coefficient was judiciously chosen to moderate the influence of empirical error terms, thereby discouraging overfitting to the data, while simultaneously ensuring that the governing physical laws are satisfied within the executed representations. The details of the training hyperparameters, along with the justification for their implementation, are presented in Table 3.

Computational complexity (see Table 4) remains a decisive metric in evaluating the performance of the predictive model. Both the forward and backward passes exhibit linear dependence on the number of parameters, yielding a complexity of O(n × p). In contrast, constraint evaluation maintains a complexity of O(n), being independent of the parameter count. Consequently, the overall complexity of the training process is rendered O(E × n × p), where E denotes the epoch count (500), n the batch size, and p the parameter count (5,121). This architecture is calibrated to guarantee that the training remains computationally manageable while still achieving a high predictive accuracy.

Furthermore, to evaluate and validate the performance of the model on three different datasets, 5-fold cross-validation was performed. Figure 2, Figure 3 and Figure 4 shows that all folds produce consistent predictive performance, with low variation in windspeed pre-

Table 3. Training Hyperparameters with Description and Justification

Parameter	Symbol	Value	Description	Justification
Learning Rate	η	0.01	Adam optimizer step size	Balanced convergence speed
Training Epochs	E	500	Maximum iterations	Sufficient for convergence
Physics Weight	λ	0.1	Constraint penalty coefficient	Balance data-physics trade-off
Batch Processing	–	Full batch	Entire dataset per iteration	Stable gradient computation
Optimizer	–	Adam	Adaptive moment estimation	Robust for neural networks

Table 4. Details of Computational Complexity

Operation	Time Complexity	Space Complexity	Dominant Factor
Forward Pass	$O(n \times p)$	$O(p)$	Input Dimension $\times$ Samples
Backward Pass	$O(n \times p)$	$O(p)$	Gradient Computation
Physics Constraints	$O(n)$	$O(1)$	Constraint Evaluation
Total Training	$O(E \times n \times p)$	$O(p)$	Epochs $\times$ Samples $\times$ Parameters

dictions, confirming that the model is generalizable to different data splits.

5 Results

5.1 For the Dataset of Badin

5.1.1 Analysis of the Model Performance

The quantitative assessment in Table 5 shows the superior predictive ability of the Physics-Informed Neural Network (PINN) for modeling mean windspeed in Badin. The PINN implementation performed excellently, yielding an R^2 of 0.999964, a mean squared error (MSE) of 0.000207, a root mean square error(RMSE) of 0.014372, and a mean absolute error (MAE) of 0.010815 for a near-perfect agreement with measured wind speed values throughout the study. These results suggest that the PINN can reconcile the physical continuity of wind movement with empirical patterns, thereby achieving high generalization capability. Even more interesting is that both Bayesian Ridge Regression and Linear Regression produce statistically similar estimations with R^2 values of 0.999963 and 0.997218, respectively, with slightly elevated error-based measurements. This constitutes additional evidence for the argument that windspeed patterns in Badin are still well approximated by linear estimates. However, the ensemble methods of Gradient Boosting, Random Forest, and XGBoost all had minor declines in fitting results with RMSE estimates of 0.1259 – 0.1408 and R^2 of ~0.9965. The Hybrid Complex Neural Network and Standard Neural Network show the most severe decline with RMSEs greater than 0.35 and R^2 values of just under 0.9748, which indicate that there is either not enough regularization or that the models may be overly flexible. These findings confirm the capacity of PINNs and conventional regressors to outperform deeper architectures in domains dominated by stable, low-variance signals like windspeed.

5.1.2 SHAP Analysis

The SHAP summary plot (Figure 5) for the model predicting mean windspeed at Badin, it is clear that time-related variables are the overriding influence on the model's predictions. The most important feature is the year, as there is an evident trend in the data: the more recent years (red dots), the higher the predicted wind speeds, and, conversely, blue dots show decreased predicted wind speeds in earlier years. The feature Year Sine is another essential feature, suggesting that wind speed predictions show good sensitivity to time based on the year. Other features, such as the Humidity Windspeed Interaction feature, Windspeed Humidity Ratio, and Humidity Ratio, impact predictions; however, not nearly as much as time-based variables. On the other hand, the features Mean Temperature and Cumulative Rainfall have minimal impact, as shown in the summary plot,

Table 5. Model Performance Table for Mean Windspeed Prediction in Badin.

Algorithm	MSE	RMSE	MAE	R^2
Physics-Informed NN (PINN)	0.000207	0.014372	0.010815	0.999964
Bayesian Ridge	0.000208	0.014431	0.010847	0.999963
Linear Regression	0.015854	0.125912	0.088457	0.997218
Gradient Boosting	0.016732	0.129353	0.085930	0.997064
Random Forest	0.019834	0.140833	0.092242	0.996519
XGBoost	0.040935	0.202323	0.160547	0.992816
Hybrid Complex NN (HCNN)	0.127162	0.356598	0.283641	0.977685
SVR	0.135705	0.368381	0.232175	0.976185
Neural Network (NN)	0.143365	0.378636	0.312540	0.974841

with SHAP values for each of these features concentrated around zero. This analysis supports the point made earlier that for Badin, the forecasting model principally used long-term trends and seasonal cycles to predict windspeed, with all other individual weather measurements (or features) being of less significant value.

5.1.3 LIME Explainability Analysis

The mean wind speed forecast for Badin based on the LIME approach (see Figure 6) has a critical interaction between humidity and temperature as an important factor that reduces wind speed forecasts. With these low interaction influences, the model generates an excessive negative interaction. At the same time, the lack of correlation between wind speed and temperature, and the cumulative effect of rainfall under 134.50 mm, all contribute to this wind speed forecast reduction. This combination of negative influences suggests that the algorithm is reacting to a well-known meteorological coupling. Then, the average downward push is primarily increased as influenced by a high humidity ratio, a humidity difference of over -39.00%, and any data records from prior to the year 2002. On the other hand, a specific range of the Humidity Windspeed interaction term, a low Year Sine value, and wind speed difference under 1.60 m/s all contribute positively and drive the forecast. However, their collective effect on the mean forecast adjustment seems weak and minor compared to the negative factors driving change. All in all, these findings suggest that wind speed forecasting in Badin may strongly depend on evolving meteorological signals, particularly when stagnant humidity-temperature cycles develop and lead to mean forecasts that are lower than other variables suggest.

5.2 For the Dataset of Dadu

5.2.1 Analysis of the Model Performance

The performance metrics in Table 6 reflect a similar trend for Dadu. The PINN model achieved an R^2 of 0.999158, with an MSE of 0.000695, RMSE of 0.02637, and MAE of 0.019474, demonstrating exceptional agreement with empirical wind speed patterns and underscoring the utility of embedding physical priors. Bayesian Ridge followed closely, with R^2 = 0.999114. Linear Regression and Gradient Boosting performed adequately, with R^2 values of 0.988790 and 0.987247, respectively, but with higher error metrics. The XGBoost and Random Forest models showed slightly diminished accuracy, with R^2 scores of 0.980377 and 0.985104. The SVR, Hybrid Complex Neural Network, and Standard Neural Network lagged significantly, with R^2 values as low as 0.947084, highlighting the limitations of purely empirical architectures in capturing the underlying thermodynamic consistency required for accurate wind speed forecasting.

5.2.2 SHAP Analysis

The SHAP summary plot (Figure 9) for the mean windspeed at Dadu shows that the temporal covariates "Year" and "Year Cosine" have the largest SHAP deviation magnitudes, demonstrating their driving relation to the explanatory variables. More recently observed calendar years share positive SHAP values, meaning that these years were associated with increased pro-

Table 6. Model Performance Table for Mean Windspeed Prediction in Dadu

Algorithm	MSE	RMSE	MAE	R^2
Physics-Informed NN (PINN)	0.000695	0.02637	0.019474	0.999158
Bayesian Ridge	0.000732	0.027047	0.019138	0.999114
Linear Regression	0.009255	0.096203	0.075728	0.988790
Gradient Boosting	0.010529	0.102613	0.067578	0.987247
Random Forest	0.012299	0.110899	0.069325	0.985104
XGBoost	0.016201	0.127283	0.085853	0.980377
Neural Network (NN)	0.030872	0.175705	0.146076	0.962607
SVR	0.031325	0.176989	0.113837	0.962059
Hybrid Complex NN (HCNN)	0.043688	0.209018	0.173960	0.947084

jected mean wind speed. In contrast, earlier calendar years shared negative SHAP values, reducing wind speed estimates. The mean temperature and some interaction terms, including Windspeed Temperature Interaction, are also influential, but have smaller magnitudes and do not have SHAP value distributions that varied widely, ranking below temporal terms in terms of importance. Thus, the graphical evidence suggests that long-term trends and seasonal cycles primarily drove this mean wind speed model to project future mean wind speed in Dadu based on observed patterns in the training data.

5.2.3 LIME Explainability Analysis

An analysis of the mean windspeed for Dadu using the LIME method (see Figure 8) found that the interaction of humidity and temperature, with a very low interaction effect, was a large driver of the increase in wind speed predictions. Considering this interaction, the model projects a disproportionate increase in the wind speed prediction. Looking at the mean wind speed, there is a high value interaction of wind speed and temperature with a low total rainfall of 119.00 mm, which also increases the forecast prediction. Collectively, these interactions suggest a lot of positive interactions and also demonstrate the model's capacity to handle complex meteorological interactions. The important drivers will be the historic rainfall, low humidity, and low temperature interactions that will affect the trajectory of the mean wind speed. There is a divergent trajectory of the forecast where low rainfall, high differences in wind speed above 1.90 m/s, and low mean rainfall ranging from 0 mm to 3.00 mm have a suppressing effect. However, the suppressing effects from those interactions are additive and statistically unimportant concerning the mean forecast adjustment. Some alternative predictors were looked at, such as mean humidity 50.55%, earlier years before 2009, and high wind speed rainfall ratio, and were shown to have very low positive effects. The findings, therefore, imply that forecasting mean wind speed in Dadu may rely considerably on evolving meteorological signals, especially when stagnant humidity temperature regimes establish themselves, yielding mean forecasts that are systematically higher.

5.3 For the Dataset of Rohri

5.3.1 Analysis of the Model Performance

Scenarios with applied PINN outcomes in Rohri (Table 7) demonstrate a comparable trend. The PINN had the highest R^2 at 0.999511, MSE equal to 0.000389, RMSE equal to 0.019722, and MAE equal to 0.014287, providing the best forecasting strategy on mean windspeed in this area. This degree of accuracy shows the utility of incorporating physical constraints into the modeling process of a variable that shares both seasons and diurnal continuity. The Bayesian Ridge algorithm is very close behind (R^2 = 0.999488), with Linear Regression further back (R^2 = 0.989193), though it still maintains robust predictive ability. Ensemble learners like XGBoost and Gradient Boosting maintain respectable accuracy, with R^2 values of 0.989; however, their corresponding RMSEs are higher in the range of 0.093 and 0.098. At the bottom of the performance rankings, here again are the SVR, Hybrid Complex NN, and Standard Neural Network models. With R^2 scores as low as 0.851,

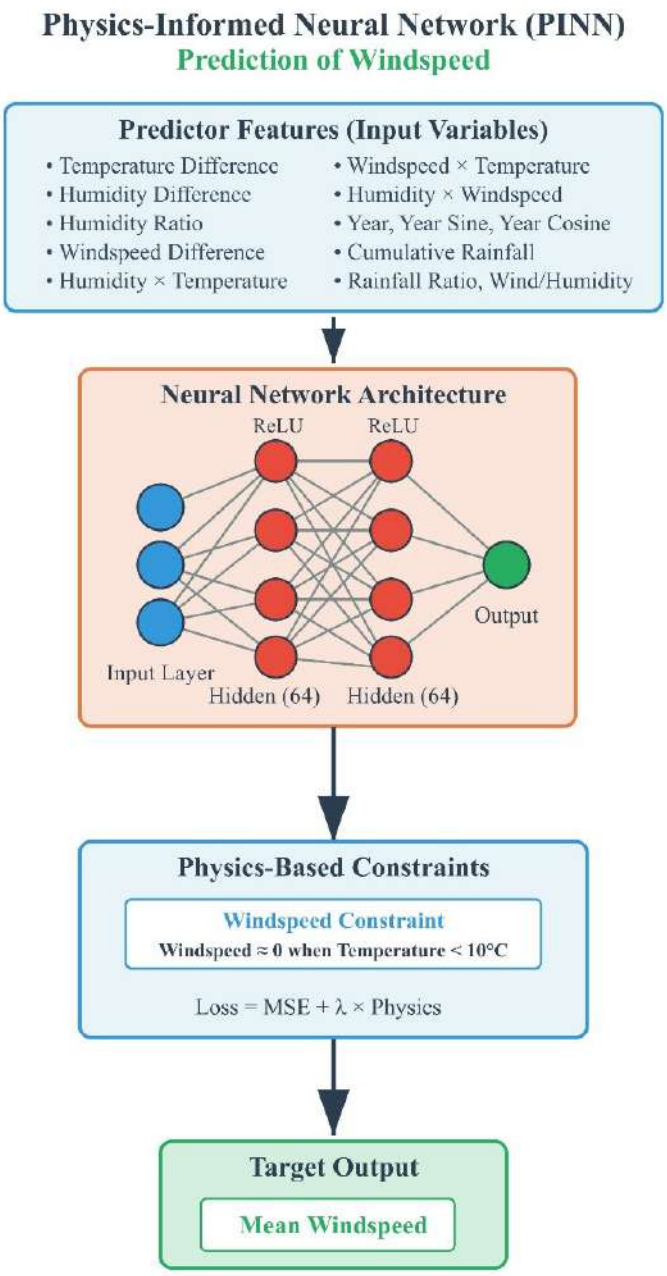

Figure 1. Architecture of the Proposed PINN

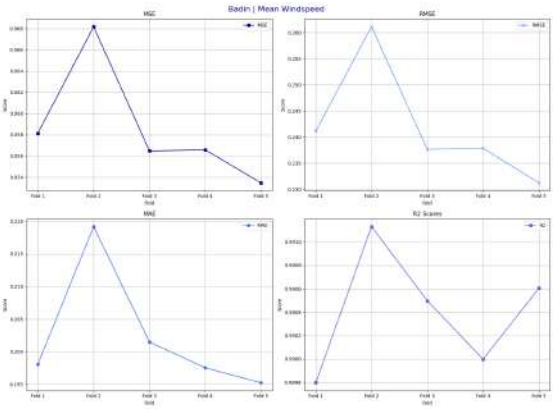

Figure 2. Cross-Validation of PINN Performance for Badin

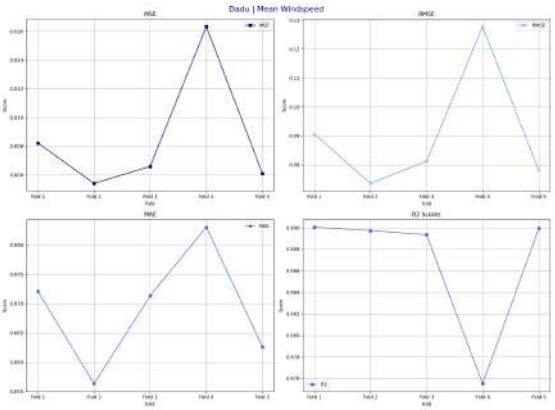

Figure 3. Cross-Validation of PINN Performance for Dadu

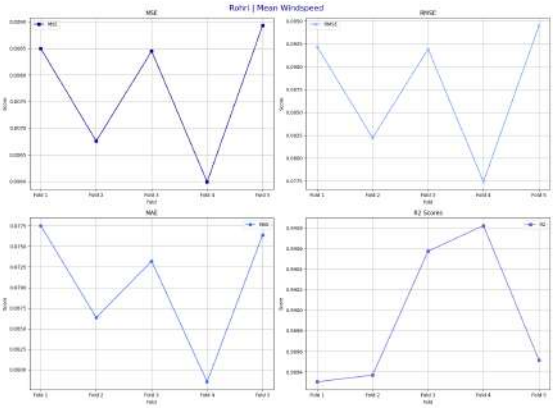

Figure 4. Cross-Validation of PINN Performance for Rohri

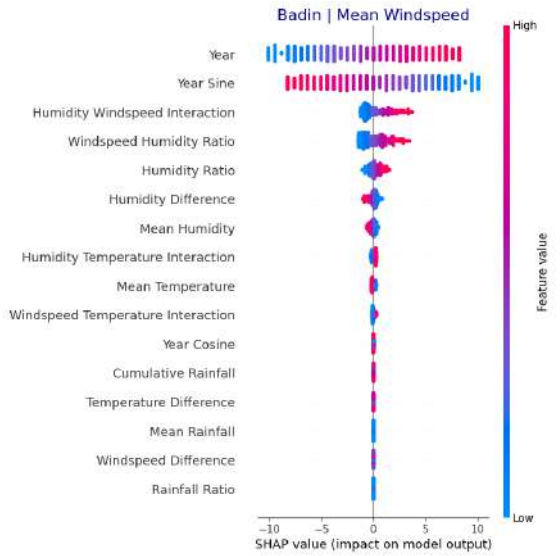

Figure 5. SHAP Values for Interpreting Feature Contributions in the PINN Model for Badin Wind Speed Prediction

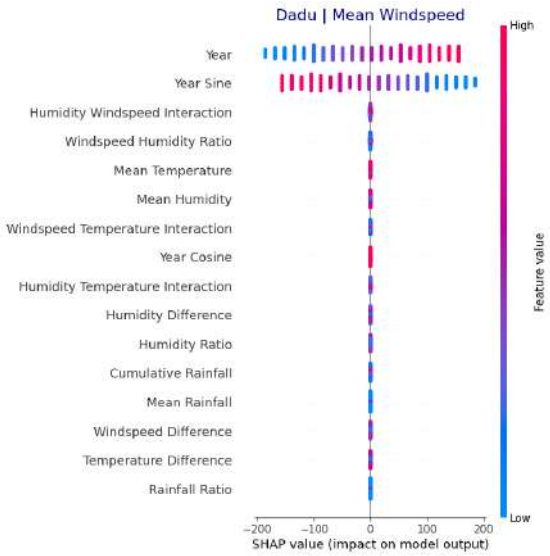

Figure 7. SHAP Values for Interpreting Feature Contributions in the PINN Model for Dadu Wind Speed Prediction and cross validation

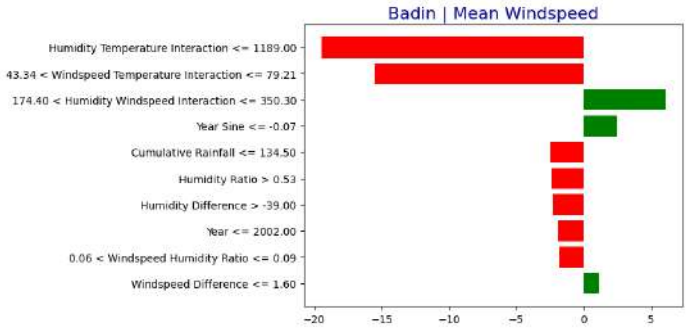

Figure 6. LIME Interpretation of the PINN Model for Badin

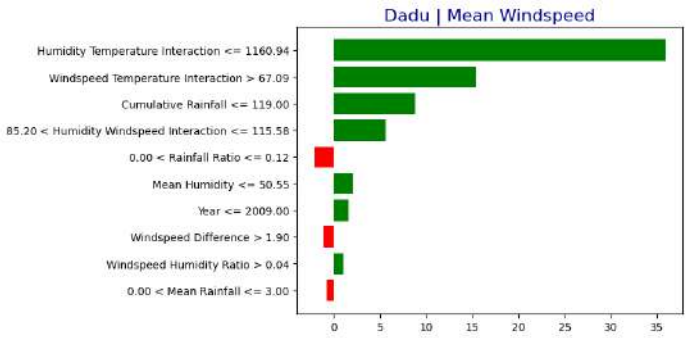

Figure 8. LIME Interpretation of the PINN Model for Dadu

with RMSE values getting up to 0.3436, and MAE values of over 0.264. Their significant performance drop shows that more complex black-box models may not work well for structured and low-noise variables like windspeed, unless preceded with some physical priors or disciplined feature engineering.

5.3.2 SHAP Analysis

The SHAP summary plot (Figure 9) for the mean wind speed model in Rohri signifies Year and Year Sine as the most significant predictors impacting wind speed accuracy. The extensive range of SHAP values for the year feature indicates that more recent years (i.e., red dots) positively correlate with predicted wind speed, while the more distant years (i.e., blue dots) negatively suppress wind speed forecasts. The Year Sine feature is also highly impactful; however, higher Year Sine values correlate with negative SHAP values, indicating negative impacts on predicted wind speed. The Windspeed Humidity Ratio and the 'Humidity Windspeed Interaction' terms follow with some impact, however much less impactful than Year and Year Sine. The plot shows that other features, including Mean Humidity, Mean Temperature, and various rainfall-related variables, have a much weaker influence, with their SHAP values concentrated near zero. Ultimately, this analysis demonstrates that the wind speed model

Table 7. Model Performance Table for Mean Windspeed Prediction in Rohri

Algorithm	MSE	RMSE	MAE	R^2
Physics-Informed NN (PINN)	0.000389	0.019722	0.014287	0.999511
Bayesian Ridge	0.000407	0.020170	0.014501	0.999488
Linear Regression	0.008589	0.092678	0.078776	0.989193
XGBoost	0.008694	0.093241	0.063456	0.989061
Gradient Boosting	0.009608	0.098021	0.070208	0.987911
Random Forest	0.014081	0.118665	0.074333	0.982282
SVR	0.043886	0.209491	0.133026	0.944781
Hybrid Complex NN (HCNN)	0.110898	0.333013	0.254458	0.860466
Neural Network (NN)	0.118032	0.343558	0.264324	0.851489

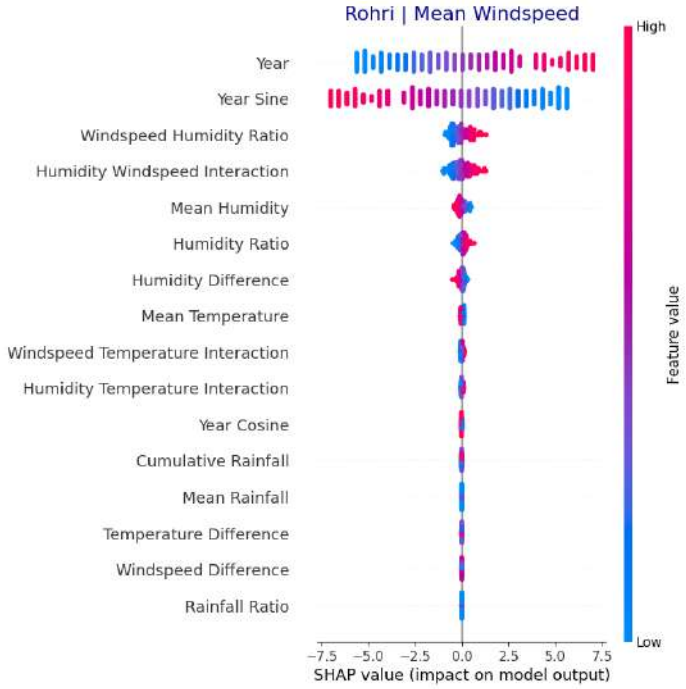

Figure 9. SHAP Values for Interpreting Feature Contributions in the PINN Model for Rohri Wind Speed Prediction

scores results based upon temporal trends and that their seasonal components, while other atmospheric variables are trivial predictors.

5.3.3 LIME Explainability Analysis

The mean windspeed prediction for Rohri, obtained through LIME methodology (see Figure 10), reveals that Humidity Temperature Interaction is the single most influential factor, driving the prediction significantly upward. This effect is counteracted by several

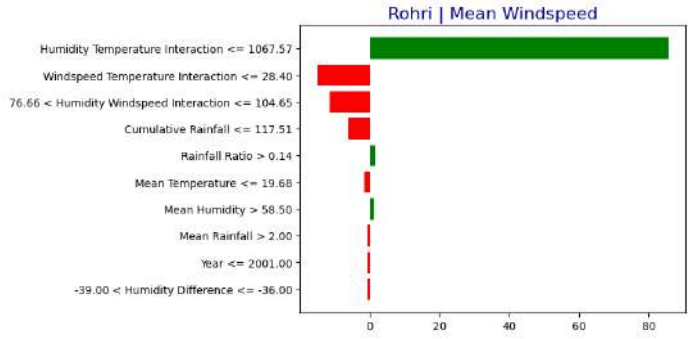

Figure 10. LIME Explanation Plot for Rohri (PINN)

features that push the prediction down, including a low-value Windspeed Temperature Interaction, a specific range for the Humidity Windspeed Interaction, and a low Cumulative Rainfall below 117.51 mm. The model also receives minor positive contributions from a high Rainfall Ratio, a low Mean Temperature below 19.68°C, high Mean Humidity above 58.50%, and Mean rainfall > 2.00 mm. The overall prediction is a result of a strong primary positive driver being tempered by a collection of negative influences.

6 Conclusion

The present study addresses the shortcomings in wind speed forecasting by proposing a novel Physics-Informed Neural Network (PINN), which integrates climate variability with physical laws by incorporating these into the loss function of the model. This incorporation increases the predictive precision and extrapolative robustness.

The model is rigorously validated across three heterogeneous datasets obtained from the Pakistan Me-

teorological Department for the cities of Badin, Dadu, and Rohri in the province of Sindh against state-of-the-art architectures. Outcomes uniformly indicated PINN preeminence, manifesting in systematic error contraction, exemplified by a mean squared error of 0.009 in the Dadu, and high R^2 metrics reaching 0.99, thereby attesting to predictive dependability.

The adoption of XAI tools of SHAP and LIME further the novelty of the approach through interpretative clarity, revealing that temperature and humidity emerge as principal modulators of wind regimes while dissolving the opaque facade often associated with deep learning paradigms. This enhanced interpretability cultivates stakeholder confidence and reaffirms the embedding of physical law within the model.

The flexibility of the proposed architecture to process sparse and noisy observational records, demonstrated by consistent cross-regional accuracies, additionally highlights its value for climate assessments in data-scarce areas. Nonetheless, the framework is not without limitations, as its results remain sensitive to extreme weather anomalies, and reliance on complete batch computation incurs substantial processing costs.

Future investigations can explore hybrid neural architectures to achieve real-time predictive capability and can augment the physical constraint layer to include pressure gradients and seasonal teleconnections. The explicit integration of AI methodology with disciplinary acumen positions this research at the interface of trustworthiness and theoretical rigor.

Author Contributions

Syed Azeem Inam, Saddam Umer, and Haider Rajput: performed the original writing, software, and methodology. **Syed Azeem Inam** and **Haider Rajput:** handled rewriting, investigation, design methodology, and conceptualization. **Syed Azeem Inam** and **Saddam Umer** contributed related work and results management. **Saddam Umer** and **Haider Rajput** performed proofreading and visualization.

Compliance with Ethical Standards

The authors declare no conflicts of interest. This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

Funding Information

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

References

[1] A. Rahujo, D. Atif, S. A. Inam, A. A. Khan, and S. Ullah, "A survey on the applications of transfer learning to enhance the performance of large language models in healthcare systems," *Discover Artif. Intell.*, vol. 5, no. 1, p. 90, Jun. 2025, doi: 10.1007/s44163-025-00339-0.

[2] S. A. Inam, D. Iqbal, H. Hashim, and M. A. Khuhro, "An empirical approach towards detection of tuberculosis using deep convolutional neural network," *Int. J. Data Mining, Model. Manage.*, vol. 16, no. 1, pp. 101–112, 2024, doi: 10.1504/IJDMMM.2024.136232.

[3] S. A. Inam, S. M. H. Zaidi, A. A. Khan, and S. Ullah, "A neural network approach to carbon emission prediction in industrial and power sectors," *Discover Appl. Sci.*, vol. 7, no. 6, p. 640, Jun. 2025, doi: 10.1007/s42452-025-07257-x.

[4] S. A. Inam *et al.*, "PR-FCNN: A data-driven hybrid approach for predicting PM2.5 concentration," *Discover Artif. Intell.*, vol. 4, no. 1, p. 75, Nov. 2024, doi: 10.1007/s44163-024-00184-7.

[5] A. Adewumi, C. E. Okoli, F. O. Usman, K. A. Olu-lawal, and O. T. Soyombo, "Reviewing the impact of AI on renewable energy efficiency and management," *Int. J. Sci. Res. Arch.*, 2024, doi: 10.30574/ijsra.2024.11.1.0245.

[6] D. Song *et al.*, "Review of AI-based wind prediction within recent three years: 2021–2023," *Energies*, 2024, doi: 10.3390/en17061270.

[7] Y. Liu, Y. Wang, Q. Wang, K. Zhang, W. Qiang, and Q. H. Wen, "Recent advances in data-driven prediction for wind power," *Front. Energy Res.*, 2023, doi: 10.3389/fenrg.2023.1204343.

[8] Y. Xia, "Leveraging AI technology for advancements in wind power," *Te*, 2024, doi: 10.61173/cgtf5f77.

[9] S. A. Inam *et al.*, "A novel deep learning approach for investigating liquid fuel injection in combustion system," *Discover Artif. Intell.*, vol. 5, no. 1, p. 32, Apr. 2025, doi: 10.1007/s44163-025-00248-2.

[10] J. Zhang and X. Zhao, "Spatiotemporal wind field prediction based on physics-informed deep learning and LIDAR measurements," *Appl. Energy*, 2021, doi: 10.1016/j.apenergy.2021.116641.

[11] A. Hamdan, K. I. Ibekwe, V. I. Ilojianya, S. Sonko, and E. A. Etukudoh, "AI in renewable energy: A review of predictive maintenance and energy optimization," *Int. J. Sci. Res. Arch.*, 2024, doi: 10.30574/ijsra.2024.11.1.0112.

[12] C. Cao, R. Debnath, and R. Álvarez, "Physics-based machine learning for predicting urban air pollution using decadal time series data," *Environ. Res. Commun.*, 2025, doi: 10.1088/2515-7620/add795.

[13] C. Fu, J. Xiong, and F. Yu, "Storm surge forecasting based on physics-informed neural networks in the Bohai Sea," *J. Phys. Conf. Ser.*, 2024, doi: 10.1088/1742-6596/2718/1/012057.

[14] A. Hamdan, K. I. Ibekwe, V. I. Ilojianya, S. Sonko, and E. A. Etukudoh, "AI in renewable energy: A review of predictive maintenance and energy optimization," *Int. J. Sci. Res. Arch.*, 2024, doi: 10.30574/ijsra.2024.11.1.0112.

[15] X. Chen, X. Zhang, M. Dong, L. Huang, Y. Guo, and S. He, "Deep learning-based prediction of wind power for multi-turbines in a wind farm," *Front. Energy Res.*, 2021, doi: 10.3389/fenrg.2021.723775.

[16] Y. Wu, Q. Wu, and J. Zhu, "Data-driven wind speed forecasting using deep feature extraction and LSTM," *IET Renew. Power Gener.*, 2019, doi: 10.1049/iet-rpg.2018.5917.

[17] G. Liu *et al.*, "Super-resolution perception for wind power forecasting by enhancing historical data," *Front. Energy Res.*, 2022, doi: 10.3389/fenrg.2022.959333.

[18] C. Fu, J. Xiong, and F. Yu, "Storm surge forecasting based on physics-informed neural networks in the Bohai Sea," *J. Phys. Conf. Ser.*, 2024, doi: 10.1088/1742-6596/2718/1/012057.

[19] S. Zhou *et al.*, "A physical-informed neural network for improving air-sea turbulent heat flux parameterization," *J. Geophys. Res. Atmos.*, 2024, doi: 10.1029/2023jd040603.

[20] R. Liu, Y. Song, Y. Chen, D. Wang, P. Xu, and Y. Li, "GAN-based abrupt weather data augmentation for wind turbine power day-ahead predictions," *Energies*, 2023, doi: 10.3390/en16217250.

[21] B. Qin, X. Huang, X. Wang, and L. Guo, "Ultra-short-term wind power prediction based on double decomposition and LSSVM," *Trans. Inst. Meas. Control*, 2023, doi: 10.1177/01423312231153258.

[22] S. Kumar, J.-C. Tzou, and S.-C. Wu, "Multi-step short-term wind speed prediction using a residual dilated causal convolutional network with nonlinear attention," *Energies*, 2020, doi: 10.3390/en13071772.