

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/382182212>

# Integrating Physics-Informed Vectors for Improved Wind Speed Forecasting with Neural Networks

Conference Paper · July 2024

CITATIONS

9

READS

194

4 authors, including:



**La Sandhu**

Central South University

22 PUBLICATIONS 42 CITATIONS

SEE PROFILE



**Ebrahim Shahzad Awan**

Western Sydney University

14 PUBLICATIONS 40 CITATIONS

SEE PROFILE



**Engr. Sharjeel Abid Butt**

International Islamic University, Islamabad

10 PUBLICATIONS 254 CITATIONS

SEE PROFILE

# Integrating Physics-Informed Vectors for Improved Wind Speed Forecasting with Neural Networks

Laeq Aslam  
*School of Automation  
Central South University  
Changsha, Hunan, PRC  
204608004@csu.edu.cn*

Runmin Zou  
*School of Automation  
Central South University  
Changsha, Hunan, PRC  
rmzou@csu.edu.cn*

Ebrahim Awan  
*International Islamic University  
Islamabad, Pakistan  
ebrahimawan@gmail.com*

Sharjeel Abid Butt  
*School of Automation  
Central South University  
Changsha, Hunan, PRC  
214608010@csu.edu.cn*

**Abstract**—This paper introduces an approach to enhance wind speed prediction by integrating Physics-Informed Vectors with neural network architectures, specifically Long Short-Term Memory and Temporal Convolution Networks. It also proposes a hybrid decaying loss function aimed at improving the efficiency of model training and its predictive performance. The methodology utilizes physical knowledge and atmospheric data from various global locations to develop predictive models. Evaluations conducted using datasets from Canada, Chile, Kazakhstan, and Mongolia illustrate the advantages of including Physics-Informed Vectors. The incorporation of these vectors leads to improvements in Mean Squared Error, Mean Absolute Error, and R2 Score across different volumes of data. The analysis reveals improvements of up to 8.43% in Mean Absolute Error, 16.39% in Mean Squared Error, and 0.82% in R2 Score for models based on Long Short-Term Memory. For models based on Temporal Convolution Networks, improvements of up to 17.27% in Mean Absolute Error, 29.24% in Mean Squared Error, and 1.55% in R2 Score were observed. The introduction of a custom loss function, which merges mean squared error with Physics-Informed Vector estimates, aids in hastening the convergence of models by modulating the influence of these vectors during the training phase. The study underscores the effectiveness of incorporating physics-informed techniques into machine learning for predicting renewable energy sources, thereby opening paths for further research and application in this domain.

**Index Terms**—Physics-Informed Vectors, Wind Speed Prediction, Neural Network Architectures, LSTM, TCN, Renewable Energy Forecasting

## I. INTRODUCTION

Wind energy is a key agent in renewable energy sources but its unpredictable, impulsive, and uncertain nature poses the major challenge to use it as a sustainable and reliable source of power. Accurate wind speed forecasting is essential for stable power generation, weather prediction, and integrating wind energy into the grid, enhancing energy security and reducing reliance on non-renewable sources. It also improves operational and financial planning in the energy sector, lowering costs. Despite the limitations of physical models and the need for condition-specific customization in probabilistic and machine learning algorithms, Physics-Informed Machine Learning and Neural Networks (PIML/PINN) present effective solutions for addressing complex constraints in this research area.

A recent study explores wind forecasting through supervised learning, which faces challenges at ground level. By analyzing wind data via anemometers located in Italy, the research found that optimal algorithms differ by location due to factors like variable types and model linearity, enhancing the understanding of atmospheric physics for tailored algorithm design [1]. Another innovative method, the graph-based PIGNN-CFD, utilizes CFD simulations to predict urban wind fields rapidly. This approach overcomes the traditional CFD models' limitations of time-intensive computations and non-scalability, allowing predictions across large urban areas [2]. In contrast, traditional forecasts often use Numerical Weather Prediction (NWP) models [3], [4], which can be error-prone due to numerical simplifications. Machine Learning (ML) techniques, particularly those integrating physics, significantly improve wind energy forecasting accuracy [5]. Similarly, Physics-Informed Neural Networks (PINNs) offer a quicker alternative for modeling wind fields in wind farms by reconstructing inflow velocity, validated against sparse simulated data [6]. Additionally, using Physics-Informed Artificial Intelligence (PIAI) surrogates enhances power forecasting by augmenting incomplete data, increasing accuracy and reliability with GANs and other models [7]. A comprehensive review also outlines how Physics-Informed Machine Learning (PIML) could supplement or replace expensive simulations in turbulent flows, particularly through neural networks for complex high Reynolds number flows in fluid mechanics [8].

Forecasting models are categorized into statistical [9] and artificial intelligence [10], [11] types, with a focus on Artificial Neural Networks (ANNs) for effective wind speed predictions [12]. Short-term forecasting for smart grids is reviewed, classifying 41 learning-based models into classical, advanced, and probabilistic categories [13]. A novel forecasting framework integrates singular spectrum analysis and VMD with multi-objective optimization for precise predictions [14]. An innovative hybrid deep learning architecture employs an enhanced transformer network for seasonal and stochastic wind variability [15]. Mixed-frequency data is used in a combined model featuring novel evaluation and non-linear combination forecasting modules [16]. A comparative study favors the CNN model for its accuracy and stability over techniques like ARIMA, GM, and LSTM [17]. To address

variable consideration and model optimization, a multivariate system utilizes advanced selection methods and multi-objective optimization [18]. A multi-step approach integrates outlier detection, adaptive data decomposition, and multi-model fusion to enhance forecasting accuracy [19]. Finally, a weighted ensemble model is optimized using an adaptive dynamic grey wolf-dipper throated optimization algorithm for improved accuracy [19]. This work examines the Thermal Wind Equation 1 and the Equation of Motion 2, essential for understanding wind dynamics driven by temperature and pressure gradients.

$$\vec{V}_T = \frac{-R}{f \cdot \Delta p} (\nabla_T \times \Delta T) \quad (1)$$

In this equation,  $\vec{V}_T$  denotes the thermal wind vector, influenced by the specific gas constant  $R$ , the Coriolis parameter  $f$ , pressure difference  $\Delta p$ , and the horizontal temperature gradient  $\nabla_T$ .

$$f \vec{k} \times \vec{V} = -\frac{1}{\rho} \nabla P \quad (2)$$

The Equation of Motion relates the Coriolis effect, air density  $\rho$ , and pressure gradients  $\nabla P$  to wind velocity  $\vec{V}$ . Moreover, humidity modifies these dynamics by affecting air density and temperature. These equations support theoretical wind speed calculations, but real-world applications face challenges from atmospheric non-linearities, humidity changes, and geographical factors. This study introduces two new approaches to improve the accuracy and efficiency of wind speed models.

First, a physics-informed model uses atmospheric data from a 100 km radius in all four cardinal directions, using the strong effect of surrounding conditions on local wind speeds. This method uses external atmospheric variables, including temperature, pressure gradients, humidity, wind speed, and direction, as extra predictive features, improving model accuracy through a detailed spatial analysis.

Second, we develop a custom hybrid loss function that combines future wind speed estimates with mean squared error (MSE), characterized by a decaying property to accelerate the training convergence. This approach not only reduces prediction error more quickly but also integrates atmospheric dynamics into the model training more effectively, leading to improved performance and efficiency in predicting wind speeds. These contributions result in significant progress in atmospheric modeling, offering a more accurate and efficient framework for wind speed prediction

## II. PROPOSED METHODOLOGY

The proposed methodology comprises two main steps: first, computing the Physics-Informed Vector (PIV), followed by its application in predicting wind speed. Figure 1 illustrates the overall block diagram of the scheme. The initial phase involves collecting and concatenating data from temperature, humidity, wind speed, air pressure, and wind direction, sourced from

the four cardinal directions to form vector  $X$ . This physics-informed vector  $X$ , once established, utilizes incoming feature vectors at time  $t$  to generate wind speed estimates for  $t + 1$ . These estimates are then integrated as additional features alongside the original temperature, humidity, wind speed, air pressure, and wind direction of the origin location. Consequently, the enhanced feature set, enriched by the physics-informed estimate, serves as input for the model to refine predictions. Detailed descriptions of each step are provided in the following subsections.

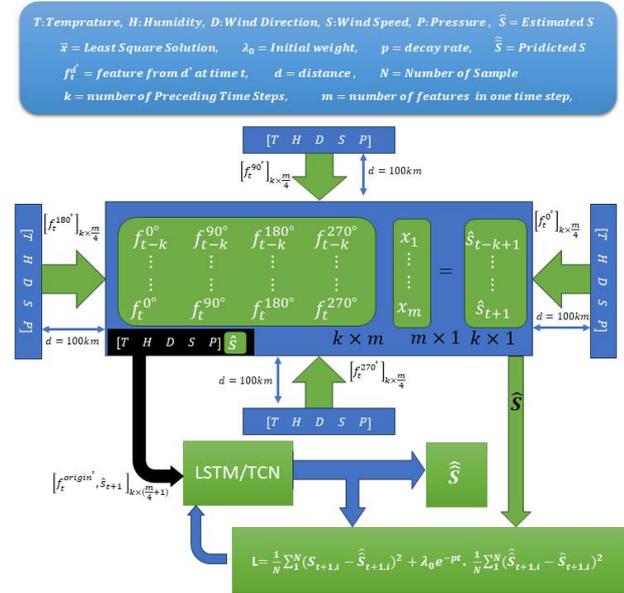


Fig. 1. Block diagram of Proposed Scheme.

### A. Calculation of Physics-Informed Vector (PIV)

The calculation of the Physics-Informed Vector (PIV) begins by assembling feature vectors from the four cardinal directions (North, South, East, West) into matrix  $A$ . Each row in matrix  $A$  represents the combined atmospheric data at a specific time step leading up to the current time  $t$ , structured as in equation 3.

$$A = \begin{pmatrix} f_{t-k}^{0^\circ} & f_{t-k}^{90^\circ} & f_{t-k}^{180^\circ} & f_{t-k}^{270^\circ} \\ \vdots & \vdots & \vdots & \vdots \\ f_t^{0^\circ} & f_t^{90^\circ} & f_t^{180^\circ} & f_t^{270^\circ} \end{pmatrix} \quad (3)$$

where  $k$  represents the total number of historical observations, and each feature vector  $f_t^\theta$  ( $\theta$  being 0, 90, 180, or 270 degrees) contains the parameters Temperature (T), Humidity (H), Wind Direction (D), Wind Speed (S), and Air Pressure (P) from one of the cardinal directions. The target vector  $b$  corresponds to the future wind speed at the central location.

The PIV,  $X$ , is obtained through solving the least mean square problem represented by the equation  $AX = b$ . Given the nature of  $A$  as a tall matrix, where the number of rows

(observations) exceeds the number of columns (features), the direct inverse of  $A$  is not computable. Instead, the solution involves the use of the pseudo-inverse ( $A^+$ ) of  $A$ , provides a means to find the 'best possible' solution to a system of linear equations that does not have a unique solution or might be over determined.

Mathematically, the pseudo-inverse is utilized as given in equation 4 to solve for  $X$ .

$$X = A^+b \quad (4)$$

where  $A^+$  is the pseudo-inverse of  $A$ . The computation of  $A^+$  is typically achieved through the Singular Value Decomposition (SVD) of  $A$ , given by equation 5.

$$A = U\Sigma V^T \quad (5)$$

Here,  $U$  and  $V$  are orthogonal matrices, and  $\Sigma$  is a diagonal matrix containing the singular values of  $A$ . The pseudo-inverse  $A^+$  is then calculated as in equation 6.

$$A^+ = V\Sigma^+U^T \quad (6)$$

where  $\Sigma^+$  is obtained by taking the reciprocal of the non-zero singular values in  $\Sigma$  and then transposing the matrix. In practical terms, this process minimizes the squared differences between the observed future wind speeds ( $b$ ) and the wind speeds predicted by the model ( $AX$ ), thereby finding the least mean square solution.

By employing the pseudo-inverse, we effectively find the best-fit line that maps the atmospheric features from the four cardinal directions to the future wind speed of the central location. This best-fit line, or model, represented by  $X$ , serves as the Physics-Informed Vector (PIV), which encapsulates the relationship between the side features and the future wind speed, denoted as  $\hat{S}_{t+1}$ .

The application of the pseudo-inverse in solving the least mean square problem allows for the incorporation of complex atmospheric data into a coherent model for wind speed prediction, reflecting a sophisticated approach to understanding and forecasting atmospheric dynamics.

### B. Incorporating PIV in Wind Speed Prediction

Once  $X$  is calculated, it is used to estimate future wind speed ( $\hat{S}$ ) at any given time  $t$ . This estimation involves multiplying  $X$  by the directional feature vector at time  $t$ — $f_t^{0^\circ}$ ,  $f_t^{90^\circ}$ ,  $f_t^{180^\circ}$ , and  $f_t^{270^\circ}$ —to estimate the wind speed at  $t+1$

$$\hat{S}_{t+1} = (f_t^{0^\circ} \cdot f_t^{90^\circ} \cdot f_t^{180^\circ} \cdot f_t^{270^\circ}) \cdot X \quad (7)$$

This estimate, along with the atmospheric parameters at time  $t$  [ $T, H, D, S, P$ ], is used as input features for predicting wind speed at  $t+1$ . The custom loss function for this model is defined as in equation 8.

$$L = \frac{1}{N} \sum_{i=1}^N (S_{t+1,i} - \hat{S}_{t+1,i})^2 + \lambda_0 e^{-pt} \frac{1}{N} \sum_{i=1}^N (\hat{S}_{t+1,i} - \hat{S}_{t+1,i})^2 \quad (8)$$

Here,  $\hat{S}_{t+1,i}$  is the model's predicted wind speed,  $\hat{S}_{t+1,i}$  is the estimated wind speed from step 1,  $N$  is the number of observations, and  $\lambda_0 e^{-pt}$  controls the contribution of the difference between the model's predictions and the PIV estimates, allowing the model to initially align closely with the PIV estimates. As training progresses and the model parameters are tuned, the loss function's design reduces the search space, enabling finer adjustments towards the actual wind speed.

### C. Optimizing Model Convergence with Hybrid Decaying Loss Function

The strategic modulation by of the PIV-based term in the loss function using equation 8 addresses the core challenge of integrating domain-specific knowledge with predictive modeling. It uses the initial guidance provided by PIV for rapid early adjustments and allows the model to later focus on the empirical data, enhancing both the speed and quality of convergence. This approach balances physical insights with data-driven learning, leading to a predictive model that adjust quickly and performs accurately in wind speed forecasting tasks. These improvements show the usefulness of the hybrid decaying loss function in making the training process more efficient and the predictions more accurate, especially for tasks like wind speed forecasting.

## III. RESULTS AND COMPARISON

In this section, we present a detailed analysis of the performance of Long Short-Term Memory (LSTM) and Temporal Convolutional Network (TCN) models for wind speed prediction. In this study the main objective was to evaluate the impact of integrating Physics-Informed Vectors (PIV) into these models and to examine their efficacy across datasets of varying sizes derived from four different countries that are Canada, Chile, Mongolia, and Kazakhstan. These countries were chosen because they have higher potential for wind energy and their solar energy potential was relatively low, making such areas ideal for deploying wind mills. We primary purposed was to explore the effect of incorporation of PIV influences on model's accuracy, convergence speed, and overall prediction performance.

Hence, we conducted experiments using complete datasets and reduced subsets, representing 75%, 50%, and 20% of the original data. The models' performances were evaluated based on three key metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and R2 Score, both during training and testing phases. This allowed us to not only compare LSTM and TCN models in terms of their predictive accuracy and efficiency but also to understand the benefits and potential limitations of applying PIV in the realm of wind speed forecasting.

### A. Overview of Experimental Setup

Data was collected from each location in the mentioned countries, extending 100 km in all cardinal directions, with the following details and subset sizes:

- **Canada:** Latitude 53.01, longitude -57.02; hourly data from 1998 to 2022, totaling 219,180 hours. Subsets are 219,180 hours (100%), 164,385 hours (75%), 109,590 hours (50%), and 43,836 hours (20%).
- **Chile:** Latitude -53, longitude -70.91; hourly data from 2019 to 2022, totaling 35,070 hours. Subsets are 35,070 hours (100%), 26,302.5 hours (75%), 17,535 hours (50%), and 7,014 hours (20%).
- **Kazakhstan:** Latitude 51.25, longitude 73.42; hourly data from 2017 to 2019, totaling 26,298 hours. Subsets are 26,298 hours (100%), 19,723.5 hours (75%), 13,149 hours (50%), and 5,259.6 hours (20%).
- **Mongolia:** Latitude 47.93, longitude 106.9; hourly data from 2011 to 2015, totaling 43,830 hours. Subsets are 43,830 hours (100%), 32,872.5 hours (75%), 21,915 hours (50%), and 8,766 hours (20%).

These subsets are used to conduct a comprehensive analysis of the model's performance across different data volumes, providing insights that the PIV performs well even with a limited amount of data. This is particularly important in cases where memory-constrained devices, such as Jetson Nano devices, are used for training models repeatedly. In such conditions, we can calculate a PIV-informed vector using historical data from a certain location, and then, by keeping a limited amount of current data on the device, we can train and perform inference in real-time scenarios.

**Models Tested:** The study utilized two advanced time series prediction models, Long Short-Term Memory (LSTM) and Temporal Convolutional Network (TCN), both with and without Physics-Informed Vector (PIV) integration.

**LSTM:** Featuring 512 units per layer, the LSTM model processes sequences through memory cells ( $C_t$ ) and three gates: input ( $i_t$ ), forget ( $f_t$ ), and output ( $o_t$ ). These components help maintain long-term dependencies within data. Mathematically, LSTM can be given as in equation 9.

$$\begin{aligned}
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i), \\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f), \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o), \\
C_t &= f_t \cdot C_{t-1} + i_t \cdot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c), \\
h_t &= o_t \cdot \tanh(C_t).
\end{aligned} \tag{9}$$

**TCN:** The TCN model includes 512 convolutional filters, a kernel size of 3, and dilation rates of [1, 2, 4, 8, 16] with 'causal' padding. This setup enables the TCN to handle long data sequences by expanding its receptive field exponentially. Mathematically, TCN convolution operation can be given as in equation 10.

$$y_t = \text{ReLU}(W * x_t + b), \tag{10}$$

where  $W$  denotes the convolutional filters, and  $b$  is the bias. The dilation ensures the model's output at any time  $t$  depends only on past inputs, preserving temporal causality.

**Evaluation Metrics:** The model's performance was assessed using:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (S_i - \hat{S}_i)^2 \tag{11}$$

where  $S_i$  are actual wind speeds,  $\hat{S}_i$  are predictions, and  $n$  is the number of observations. MSE measures the average squared prediction errors.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |S_i - \hat{S}_i| \tag{12}$$

MAE quantifies the average absolute errors.

$$R^2 = 1 - \frac{\sum_{i=1}^n (S_i - \hat{S}_i)^2}{\sum_{i=1}^n (S_i - \bar{S})^2} \tag{13}$$

where  $\bar{S}$  is the mean of actual speeds.  $R^2$  indicates the proportion of wind speed variance predictable by the model's variables: temperature, humidity, wind speed, wind direction, and air pressure.

These metrics evaluate the model's accuracy ( $MSE$  and  $MAE$ ) and explanatory power ( $R^2$  Score).

**Hardware Used:** The experiments were conducted on an Ubuntu system equipped with a 3080ti GPU for model training and testing.

$$\text{Improvement \%} = \left( \frac{\text{Metric}_{\text{with PIV}} - \text{Metric}_{\text{without PIV}}}{|\text{Metric}_{\text{without PIV}}|} \right) \times 100 \tag{14}$$

Note that for MSE and MAE, a positive improvement % indicates a decrease in error, while for the R2 Score, a positive improvement % indicates an increase, reflecting better model fit.

TABLE I: Results of PIV Informed Model

Dataset	Subset	MSE (m/s <sup>2</sup> )		MAE (m/s)		R <sup>2</sup> Score	
		LSTM	TCN	LSTM	TCN	LSTM	TCN
Canada	All	0.00160	0.00164	0.020	0.021	0.92	0.91
	75%	0.0016	0.00173	0.020	0.022	0.92	0.91
	50%	0.00164	0.00175	0.020	0.022	0.91	0.91
	20%	0.00173	0.00200	0.021	0.025	0.91	0.89
Chile	All	0.05584	0.07828	0.166	0.207	0.99	0.99
	75%	0.05699	0.07605	0.169	0.202	0.99	0.99
	50%	0.05738	0.08747	0.169	0.216	0.99	0.99
	20%	0.07304	0.14442	0.197	0.284	0.99	0.98
Kazakhstan	All	0.04916	0.05899	0.160	0.179	0.99	0.99
	75%	0.05078	0.06268	0.163	0.183	0.99	0.98
	50%	0.04751	0.08083	0.156	0.208	0.99	0.98
	20%	0.08792	0.14893	0.218	0.288	0.98	0.96
Mongolia	All	0.06321	0.06234	0.174	0.173	0.98	0.98
	75%	0.06571	0.06761	0.178	0.180	0.98	0.98
	50%	0.06693	0.07361	0.178	0.189	0.98	0.98
	20%	0.07024	0.12498	0.183	0.258	0.98	0.96

## B. Results

The impact of dataset size on model performance, specifically using Physics-Informed Vectors (PIV), is detailed in Table I. Models used 12-hour data spans to forecast subsequent time steps, evaluated by Mean Squared Error (MSE), Mean Absolute Error (MAE), and R2 Score across datasets from Canada, Chile, Kazakhstan, and Mongolia, segmented into subsets (All, 75%, 50%, 20%). Results indicate a consistent trend where reduced dataset size leads to increased MSE and MAE, and decreased R2 Score across both LSTM and TCN models, underscoring the challenge of diminished data availability on model accuracy.

For example, in the Canada dataset, LSTM's MSE escalated from 0.00160 (using all data) to 0.00173 (20% data), with a corresponding R2 Score drop from 0.913 to 0.894 in the TCN model. These trends highlight diminished predictive accuracy with lesser data. The lowest errors were noted in the Canadian dataset with the most samples, suggesting a direct correlation between data volume and prediction accuracy. Conversely, despite having fewer samples, the Chile dataset achieved the highest R2 Score, indicating that while larger datasets generally improve accuracy and reduce prediction errors, the R2 Score also reflects other factors such as dataset characteristics or the inherent predictability of the wind patterns, affecting the model's explanatory power.

## C. Models Comparison With and Without PIV

To evaluate the enhancements provided by Physics-Informed Vectors (PIV) on LSTM and TCN models, improvements were measured using MSE, MAE, and R2 Score. The percentage improvement for all metrics is calculated using equation 14.

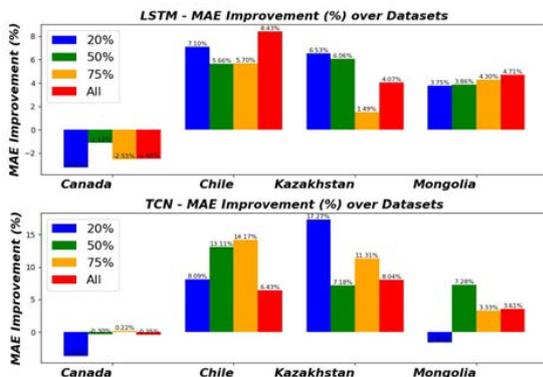


Fig. 2. MAE Improvement Across Datasets and Models

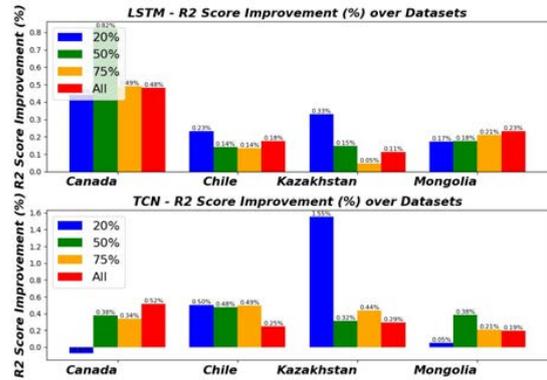


Fig. 3. R2 Improvement Across Datasets and Models

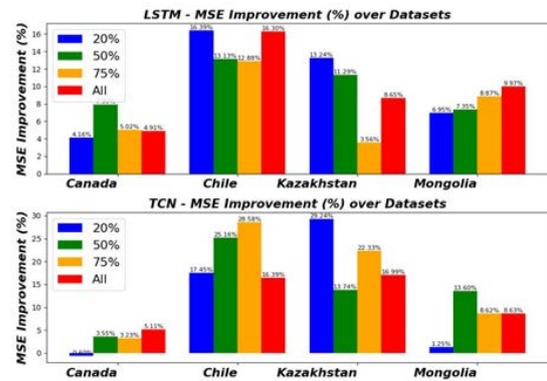


Fig. 4. MSE Improvement Across Datasets and Models

Figure 2 displays enhancements in mean absolute error (MAE) across datasets from Canada, Chile, Kazakhstan, and Mongolia, with the Physics-Informed Vector (PIV) derived from the full dataset in each case. Color coding in the graphs indicates the dataset proportion used for training: red for 100%, yellow for 75%, green for 50%, and blue for 20%. Notably, the PIV-enhanced LSTM model shows improved MAE across all datasets except Canada, with increases from 1.49% to 8.43%. The TCN model sees MAE improvements in Chile, Kazakhstan, and Mongolia, with a performance dip in Canada for the 20% and 50% data subsets.

Figure 4 illustrates mean squared error (MSE) improvements using similar color codes. The LSTM model records MSE enhancements in all datasets, with the smallest gain of 3.56% in Kazakhstan (75% data) and the highest of 16.39% in Chile (20% data). The TCN model also shows MSE improvements, ranging from 1.25% (20% of Mongolia) to 29.24% (75% of Chile), except in some Canada subsets.

Figure 3 demonstrates R2 score improvements, using the same color scheme. The LSTM model sees R2 enhancements across all datasets, with minimal and maximal increases of 0.05% (75% of Kazakhstan) and 0.82% (50% of Canada), respectively. Similarly, the TCN model improves R2 scores

except in the smallest Canada subset, with performance gains from 0.04% (20% of Mongolia) to 1.55% (20% of Kazakhstan).

#### D. Hybrid Decaying Custom Loss Function and Model Convergence

This section details the custom loss function described in Equation 8, which combines mean squared error (MSE) with a Physics-Informed Vector (PIV)-influenced term. The function includes a decaying term  $\lambda_0 e^{-pt}$ , with  $\lambda_0$  set at 0.5 and  $p$  calculated as approximately 0.541, ensuring the PIV influence diminishes after 20 iterations. This decay promotes initial use of PIV insights and focuses on minimizing direct prediction errors over time, enhancing model convergence. Figure 5 demonstrates faster convergence with the PIV-informed LSTM model (blue line) compared to the non-PIV model (orange line).

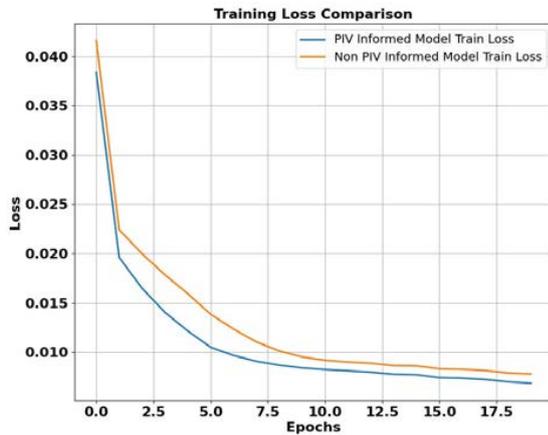


Fig. 5. Convergence Analysis of PIV Informed Model Vs. Non-Informed Models

The hybrid decaying mechanism of the custom loss function leverages physical insights initially, then transitions to pure error minimization, accelerating convergence and ensuring robust predictions. This approach illustrates the effectiveness of integrating domain-specific knowledge into machine learning for complex forecasting tasks like wind speed prediction, highlighting significant convergence improvements in PIV-informed models.

#### IV. CONCLUSION

This work introduces an approach for improving wind speed prediction by integrating Physics-Informed Vectors (PIV) and a hybrid decaying loss function. The proposed method was tested using datasets from various locations with high wind power potential. It demonstrates the advantages of incorporating physical knowledge into machine learning models for managing wind power systems. The custom loss function, which adjusts the influence of PIV over time, effectively accelerates

model training convergence. Although this study used data from the four cardinal directions, future research could explore the effects of including intermediate directions and varying data collection distances to optimize the methodology further.

#### REFERENCES

- [1] D. Lagomarsino-Oneto, G. Meanti, N. Pagliana, A. Verri, A. Mazzino, L. Rosasco, and A. Seminara, "Physics informed machine learning for wind speed prediction," *Energy*, vol. 268, p. 126628, 2023.
- [2] X. Shao, Z. Liu, S. Zhang, Z. Zhao, and C. Hu, "Piggn-cfd: A physics-informed graph neural network for rapid predicting urban wind field defined on unstructured mesh," *Building and Environment*, vol. 232, p. 110056, 2023.
- [3] T. Howard and P. Clark, "Correction and downscaling of nwp wind speed forecasts," *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, vol. 14, no. 2, pp. 105–116, Jun 2007.
- [4] L. Dong *et al.*, "Studies on wind farms ultra-short term nwp wind speed correction methods," in *2013 25th Chinese Control and Decision Conference (CCDC)*. IEEE, 2013.
- [5] N. M. R. Monteiro, "Improving wind energy forecasting with innovative physics-informed machine learning approaches."
- [6] P. Cobelli, K. Shukla, S. Nesmachnow, and M. Draper, "Physics informed neural networks for wind field modeling in wind farms," in *Journal of Physics: Conference Series*, vol. 2505, no. 1. IOP Publishing, 2023, p. 012051.
- [7] Z. Wu, B. Sun, Q. Feng, Z. Wang, and J. Pan, "Physics-informed ai surrogates for day-ahead wind power probabilistic forecasting with incomplete data for smart grid in smart cities." *CMES-Computer Modeling in Engineering & Sciences*, vol. 137, no. 1, 2023.
- [8] P. Sharma, W. T. Chung, B. Akoush, and M. Ihme, "A review of physics-informed machine learning in fluid mechanics," *Energies*, vol. 16, no. 5, p. 2343, 2023.
- [9] V. Yılmaz and H. E. Çelik, "A statistical approach to estimate the wind speed distribution: the case of gelibolu region," *Doğuş Üniversitesi Dergisi*, vol. 9, no. 1, pp. 122–132, 2008.
- [10] J. Wang and Y. Wang, "A novel wind speed forecasting model for wind farms of northwest china," *International Journal of Green Energy*, vol. 14, no. 5, pp. 463–478, 2017.
- [11] Y. Wang, T. Chen, S. Zhou, F. Zhang, R. Zou, and Q. Hu, "An improved wavenet network for multi-step-ahead wind energy forecasting," *Energy Conversion and Management*, vol. 278, p. 116709, 2020.
- [12] S. M. Valdivia-Bautista, J. A. Domínguez-Navarro, M. Pérez-Cisneros, C. J. Vega-Gómez, and B. Castillo-Téllez, "Artificial intelligence in wind speed forecasting: A review," *Energies*, vol. 16, no. 5, p. 2457, 2023.
- [13] V. K. Saini, R. Kumar, A. S. Al-Sumaiti, A. Sujil, and E. Heydarian-Forushani, "Learning based short term wind speed forecasting models for smart grid applications: An extensive review and case study," *Electric Power Systems Research*, vol. 222, p. 109502, 2023.
- [14] J. Wang, X. Niu, L. Zhang, Z. Liu, and X. Huang, "A wind speed forecasting system for the construction of a smart grid with two-stage data processing based on improved elm and deep learning strategies," *Expert Systems with Applications*, vol. 241, p. 122487, 2024.
- [15] B. S. Bommidi, K. Teeparthi, and V. Kosana, "Hybrid wind speed forecasting using iceemdan and transformer model with novel loss function," *Energy*, vol. 265, p. 126383, 2023.
- [16] Y. Hao, W. Yang, and K. Yin, "Novel wind speed forecasting model based on a deep learning combined strategy in urban energy systems," *Expert Systems with Applications*, vol. 219, p. 119636, 2023.
- [17] X. Li, K. Li, S. Shen, and Y. Tian, "Exploring time series models for wind speed forecasting: A comparative analysis," *Energies*, vol. 16, no. 23, p. 7785, 2023.
- [18] J. Wang, M. Lv, Z. Li, and B. Zeng, "Multivariate selection-combination short-term wind speed forecasting system based on convolution-recurrent network and multi-objective chameleon swarm algorithm," *Expert Systems with Applications*, vol. 214, p. 119129, 2023.
- [19] E.-S. M. El-Kenawy, S. Mirjalili, N. Khodadadi, A. A. Abdelhamid, M. M. Eid, M. El-Said, and A. Ibrahim, "Feature selection in wind speed forecasting systems based on meta-heuristic optimization," *Plos one*, vol. 18, no. 2, p. e0278491, 2023.