Weather Forecasting Based on Data-Driven and Physics Informed Reservoir Computing Models

Yslam D. Mammedov (1001955426@ucsiuniversity.edu.my)
UCSI University Faculty of Engineering Technology and Built Environment https://orcid.org/0000-0003-4868-2752

Ezutah Udoncy Olugu
UCSI University Faculty of Engineering Technology and Built Environment

Guleid A. Farah
Georgia Institute of Technology

Research Article

Keywords: Atmospheric disturbance, Echo State Networks, Lorenz system, Physics informed machine learning, Recurrent Neural Networks, Reservoir Computing, Wind speed

Posted Date: September 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-884990/v1

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on November 25th, 2021. See the published version at https://doi.org/10.1007/s11356-021-17668-z.

Weather Forecasting Based on Data-Driven and Physics Informed Reservoir Computing Models

Yslam D. Mammedov[1, *], Ezutah Udoncy Olugu[2, *] and Guleid A. Farah[3]

[1] Department of Industrial & Petroleum Engineering, Faculty of Engineering, Technology & Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia; 1001955426@ucsiuniversity.edu.my
[2] Department of Mechanical Engineering, Faculty of Engineering, Technology & Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia; olugu@ucsiuniversity.edu.my
[3] Department of Computer Science, College of Computing, Georgia Institute of Technology, 30332 Atlanta, GA, United States; gfarah7@gatech.edu

[*] Correspondence: 1001955426@ucsiuniversity.edu.my (Y.D.M.), olugu@ucsiuniversity.edu.my (E.U.O.)

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The data that support the findings of this study are available from the corresponding author upon reasonable request. The code utilized in this study was based on open-source content provided by A. Racca at https://gitlab.com/ar994/robust-validation-esn.

Competing interests: The authors declare no conflict of interest.

Funding: This research was funded by UCSI University through the Pioneer Scientist Incentive Fund (PSIF), grant number Proj-In-FETBE-062.

Authors' contributions: Conceptualization, E.U.O.; Investigation, G.A.F.; Formal analysis, Software and Validation, G.A.F. and Y.D.M.; Funding acquisition and Supervision, E.U.O.

Acknowledgements: We would like to thank A. Racca for helpful discussion.

Abstract

In response to the growing demand for the global energy supply chain, wind power has become an important research subject among studies in the advancement of renewable energy sources. The major concern is the stochastic volatility of weather conditions that hinder the development of wind power forecasting approaches. To address this issue, the current study proposes a weather prediction method divided into two models for wind speed and atmospheric system forecasting. First, the data-based model incorporated with wavelet transform and recurrent neural networks is employed to predict the wind speed. Second, the physics-informed echo state network was used to learn the chaotic behaviour of the atmospheric system. The findings were validated with a case study conducted on wind speed data from Turkmenistan. The results suggest the out-performance of physics-informed model for accurate and reliable forecasting analysis, which indicates the potential for implementation in wind energy analysis.

Keywords: Atmospheric disturbance, Echo State Networks, Lorenz system, Physics informed machine learning, Recurrent Neural Networks, Reservoir Computing, Wind speed

1. Introduction

Due to the continued increase of energy demand, conventional energy sources seem unable to support energy advances in recent years. Global energy consumption is predicted to be grown by around 60% by 2030 (Bahrami et al., 2019). As a result, countries have set a long-term goal for the utilization of renewable energy sources (Zhang et al., 2018). Among the renewable energy options, wind energy is getting prominence as one of the most promising alternatives (Zhang et al., 2020). According to (Bahrami et al., 2019), 121 out of 195 nations use it as a source of electricity with Asia accounted for almost 40%. The authors' specific focus was on promoting the renewable energy exploitation in Turkmenistan by providing the country's first wind speed evaluation. The results highlight the significance of the energy market as Turkmenistan is a main electricity supplier in the Central Asia and its further potential from wind power. Wind speed is a significant element in the wind power production (Hu et al., 2021), and accurate and dependable wind speed and weather forecasting systems are conducive to lowering

operating costs and improving wind power system stability (Zhang et al., 2020). Therefore, there is a need to develop a comprehensive model to characterize the unpredictability and instability of the chaotic nature of weather. Scholars are currently undertaking substantial research and contributing significantly to the area of wind speed forecasting.

1.1 Related works

Four types of wind power forecasting approaches have been presented so far: physical modelling, statistical methods, artificial intelligence (AI) models, and their hybrids (Wang et al., 2019). First category models are based on the physical processes in the atmosphere. (Srivastava and Bran, 2018) investigated aerosol concentration by using physical parameterizations to describe the dynamics of the atmosphere. (Zhang et al., 2019) employed computational fluid dynamics to simulate the wind's flow pattern impact on pollutant dispersion. Physical models do not need to be trained on prior data; thus, they are practical for use in new wind farms (Hu et al., 2021). On the other hand, they are unsuitable for small regions or short-term forecasting and need a large amount of data. Statistical models, including linear regression and an autoregressive integrated moving average (Hu et al., 2021), prediction consistent with earlier observations. Notable applications include: (Snoun et al., 2019) utilized the Gaussian atmospheric model to accurately retrieve the short-range distribution of wind speed. (Natarajan et al., 2021) compared probability distribution models for performance monitoring in Indian wind farms. Although they outperform physical models, statistical models are unable to predict nonlinear patterns. An alternative solution is to employ AI-based forecasting methods. They have made significant progress in the field of time series forecasting (Liu et al., 2020). The primary advantage of AI models is their high learning capacity and nonlinear mapping capability (Wang et al., 2019). The artificial neural networks (ANN) and their Deep learning structure are frequently employed in wind speed forecasting (Liu et al., 2020). (Zhang et al., 2021) proposed a wind speed prediction scheme based on long and short-term memory neural networks. (Cui et al., 2019) used backpropagation neural network to improve wind speed forecasting accuracy. The advantage of employing ANN over physical and statistical approaches is due to the approximation of nonlinear functions. The adaptive method changes its internal structure; thus, the prediction model may imitate the potential logic of original information.

Although deep learning algorithms perform well in time series analysis, (Liu et al., 2020) highlighted its disadvantage is a single model to learn entire wind speed conditions. Therefore, recent studies employ hybrid models that combine several approaches to reach more accurate results (Wang et al., 2019). (Hu et al., 2021) suggests three models' structure namely, (I) decomposition by employing variational mode decomposition, (II) optimization by using differential evolution, and (III) forecasting the assembly of decomposed variables in the echo state network (ESN). (Liu et al., 2020) uses similar structure with empirical wavelet transform utilized for decomposition, then employs reinforcement learning to ensemble deep learning algorithms. (Tian, 2020) suggests improved ESN with grey wolf optimization to improve accuracy and reduce errors in wind speed time-series. (Wang et al., 2019) uses similar techniques, whereas forecasting was performed before assembling decomposed variables. (Gupta et al., 2021) utilized hybrid methods to predict short-term wind speed in five Indian wind farms. One unique approach was presented by (Zhang and Pan, 2020), author uses a hybrid approach by incorporating Elman-radial Basis function and Lorenz disturbance to acquire more accurate results. Their study employs wavelet transform (WT) and ESN with ensemble techniques.

In addition, AI techniques have a promising advantage over physical and statistical models, however, there is a drawback of robust in reliability for scientific applications and decision-making processes. The main obstacle highlighted in (Kashinath et al., 2021), is that the model's deficiency in observing the physical phenomena of the system. (Tian, 2020) stated that scholars analysing wind speed rarely consider its chaotic nature, where the system has strong nonlinearity and uncertainty. There is an interesting contribution made to resolve this issue in (Zhang et al., 2018), the study utilizes physics model namely, the Lorenz system to describe stochastic volatility of the wind speed model. However, the proposed model considers physical phenomena as a separate segment of prediction analysis. A more recent attempt to resolve this issue emerged in novel physics-informed machine learning (PIML), which incorporates the data and mathematical models, and implements them through AI algorithms (Kashinath et al., 2021). The PIML training is based on additional information from physical principles that enables it to satisfy invariants of the continuous space-time domain for better accuracy and improved generalization (Karniadakis et al., 2021). (Kashinath et al., 2021) highlighted the challenge of learning the nonlinear dynamics of weather and climate phenomena and outperformance of PIML in the resolution and complexity of weather prediction models.

1.2 Contributions

Recognizing the challenges of single models, this paper proposes a hybrid wind speed forecasting approach by simultaneous assessment of two modules. Initially, the wind speed forecasting module was evaluated in three steps: decomposition, optimization, forecasting. First, in the decomposition module, the WT is used to eliminate the noise of original wind speed data and decompose it into several sub-signals with better counters and behaviour. Second, reservoir computing is utilized for all decomposed sub-series. Third, RNN is optimized with RMSprop (Keras) to obtain better results. Then, the dynamic behaviour of atmospheric conditions is simulated by the Lorenz system. First, the chaotic behaviour of atmospheric changes in pressure and temperature is defined by Lorenz equations and incorporated into PIML. Second, physics informed ESN algorithm is employed due to its advantage in describing the chaotic dynamics over traditional reservoir computing methods. Third, the study utilizes chaotic Recycle validation for robust and performance of validation strategy; and Bayesian optimization to compute optimal hyperparameters that are suggested in (Racca and Magri, 2021). Finally, the feasibility and reliability of the proposed forecasting approach are verified by actual wind speed data from Turkmenistan. As a result of the first comprehensive wind energy resource assessment in the region, (Bahrami et al., 2019) suggests that Gazanjyk among four prior cities to have an advantage over other locations for wind energy development and deployment. Following their finding, this study contributes to the renewable energy development of Gazanjyk by proposing a wind forecasting method to secure the stable and reliable performance of wind energy and promote the safety of the power system.

The rest of this article is organized as follows: Section 2 introduces reservoir computing and the structure of PIML. Section 3 shows the implementation steps of the data-driven method. Section 4 describes the effectiveness of physics informed ESN. Section 5 summarizes the conclusion and prospects of study.

2. Methodology

This section proposes a hybrid weather forecasting model that incorporates simultaneous assessment of dynamical behaviour of wind speed and atmospheric system. Section 2.1 presents the framework to capture the uncertainty and volatility of the system. Section 2.2 explains the rationality of employing wavelet decomposition. Section 2.3 introduces RNN and reservoir computing methods for wind speed prediction. Section 2.4 proposes a physics-informed ESN to forecast atmospheric systems.

2.1 Proposed framework

The weather prediction process of the proposed framework is shown in **Figure 1**. The process contains two stages: the wind speed and atmospheric system models. The details of the framework are explained as follows:

Stage 1: the original data on wind speed performance is collected. Then, the decomposition of time-series data is performed by WT. In this step, the wind speed data is separated into approximate sub-signal and the respective detail sub-signals. After, the time-series data is split into training sets for the training model and test sets for validation purposes. Consequently, the reservoir computing approach or RNN is used to construct a wind speed forecasting model.

Stage 2: the Lorenz system is employed to describe the chaotic state of the atmospheric system. The physics-informed reservoir computing is utilized to time-accurately forecast the system. In particular, the ESN is considered to construct chaotic dynamical behaviour represented through the Lorenz system. First, the ensemble of network realizations is employed to address the random initialization of the reservoir for the robustness of the ESN. The validation of the constructed model is evaluated by the chaotic recycle validation (CRV) technique. Second, Bayesian optimization is utilized to correspond to the high hyperparameter sensitivity of ESN. This optimization technique improves reservoir architecture and outperforms conventional methods due to a gradient-free search engine.

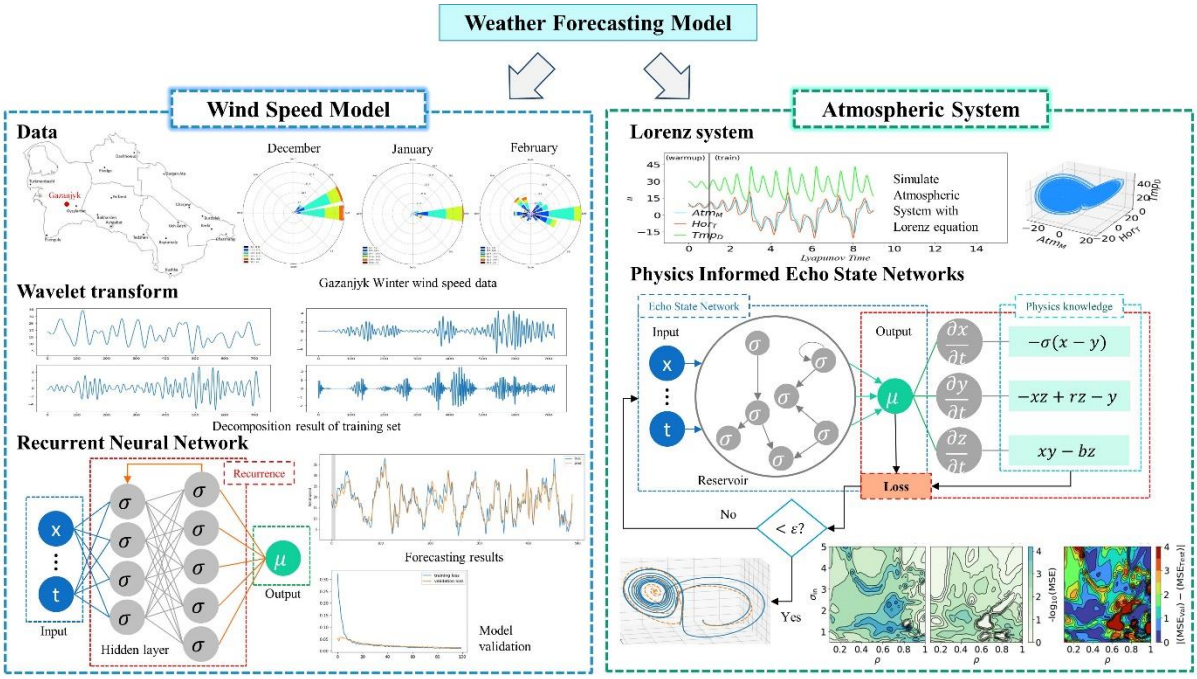

Figure 1. Research Flowchart

2.2 Wavelet transform

The wind speed data consists of randomness, fluctuation, and uncertainty that features the spikes in nonlinear time series (Zhang et al., 2020). Accordingly, these features cause difficulty in wind speed prediction models. To address this problem, wavelet decomposition techniques were employed (Liu et al., 2020; Wang et al., 2019). The WT is a signal decomposition approach that enables accurate time-frequency localization properties (Zhang et al., 2018). More specifically, WT decomposes time series data by generating multiple corresponding filters to reveal the changes or fluctuation of original data (Hu et al., 2021). The general formula of WT can be described as follows:

$$Wavelet\ (i,j) = 2^{-(i/2)} \sum_{t=0}^{T-1} f(t)\varphi[(t - j2^i)/2^i] \quad (1)$$

where i and j represents the scaling and translation parameters of the mother wavelet φ. Besides, t and T are a discrete-time representation of the length of the whole signal $f(t)$.

2.3 Reservoir computing

2.3.1 Recurrent Neural Networks

Wind speed data consist of sequential data, where the prediction model requires to consider information relevant to the previous steps in the sequence (Elsaraiti and Merabet, 2021). Recurrent neural networks (RNN) outperform adaptive neural networks (Bollt, 2021) in learning the long-term dependency for time series forecasting (Kumar et al., 2020). The RNN structure consists of hidden layers distributed across time (Elsaraiti and Merabet, 2021) that enables the achievement of information from the previous state of reading historical data (Duan et al., 2021). In other words, ANN is based on nodes connected between layers and limited to links within a hidden layer. Whereas in RNN this connection is provided, therefore the output has access to the input of the current hidden layer as well as to the output of the previous one. That enables effective learning of time-series data and makes it consistent with wind speed forecasting. Figure 2 illustrates the architecture of RNN.

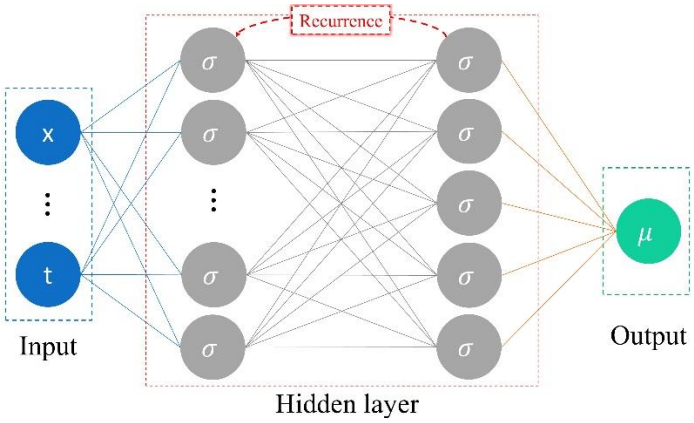

Figure 2. Recurrent Neural Networks

Even though the RNN achieved success in enhanced feature extraction ability in time step prediction models (Duan et al., 2021), it is limited to handle long-term time dependencies (Elsaraiti and Merabet, 2021). That makes RNN's architecture more sensitive toward vanishing or exploding. To address this issue, Jaeger proposed a new approach called reservoir computing (RC) in 2001 and 2002, respectively (Kumar et al., 2020). The main idea of RC is to enable more efficient computation. The simple architecture is consisting of randomly fixed weights that would expand the input vector and using the expansion it would train to fit a linear model. So, the RNN is a structure where the input of hidden layers is carefully designed by random values that just learn the result of several hidden output layers. The training involves linear regression to evaluate weights for an accurate result. The linear combination of input and reservoir state makes RC the best alternative for nonlinear system forecasting of dynamical systems. This approach is also known as echo state networks (ESN).

2.3.2 Echo state networks

Introducing ESN Jaeger (Jaeger and Haas, 2004) composed the network as an architecture of three layers, such as input, hidden (reservoir) and output layers. The input vector $x(t)$ is connected to the hidden layer by the weight matrix W_{in}. The hidden layer is represented as a dynamical reservoir of random weights that are connected through a matrix W_{res}. The output vector is represented by a matrix W_{out} that consists of an output vector $\mu(t)$. This structure allows output neurons to fed back signals to hidden layers by W_{sb}. If K represents the number of input neurons, then the parameters of the input vector x at the time t would be $x(t) = [x_1(t), x_2(t), \dots x_K(t)]^T$. Similarly, if L and M represent the number of neurons in the reservoir $\sigma(t)$ and output $\mu(t)$ vectors at the time t, would be $\sigma(t) = [\sigma_1(t), \sigma_2(t), \dots \sigma_L(t)]^T$ and $\mu(t) = [\mu_1(t), \mu_2(t), \dots \mu_M(t)]^T$, respectively. Then the update function of the internal reservoir is:

$$\sigma(t+1) = tanh(W_{res}\sigma(t) + W_{in}x(t) + W_{sb}\mu(t)) \quad (2)$$

After applying activation function $tanh$, in most cases it is a hyperbolic tangent function (Wang et al., 2019), the output vector is:

$$\mu(t+1) = I(W_{out}[\sigma(t+1) + x(t+1) + \mu(t)] + b_{in}) \quad (3)$$

where b_{in} denotes the bias and I represents the identity function. Since internal signals of the reservoir are linearly correlated, the obvious choice of update function is the linear regression method.

The performance of ESN is based on three important parameters, such as spectral radius ρ, reservoir scale r and connectivity rate α. First, spectral radios ρ evaluate with the largest absolute value λ_{max} among the eigenvalues of the internal weight matrix W_{res}. To ensure ESN to have echo state property, the λ_{max} and spectral radios ρ should be less than 1. In that case, the input and past reservoir information will vanish with time (Wang et al., 2019). Second, the reservoir scale r refers to the number of neurons in the reservoir. This directly impacts the performance of ESN, since it defines the memory capacity of the network and reservoirs' exponential relationship through the number of training samples. A large reservoir scale offers an advantage in learning for complex systems, otherwise, it leads to overfitting. Thus, requiring adjust the number r through a trial-and-error method for stronger generalization ability of ESN. Third, the connectivity rate α denotes the connection condition of neurons in the reservoir. In case when there is no connection among the neurons, this leads to a lack of memory

and results in loss of reservoir state. On the other hand, the large value of connectivity ratio α results in difficulty to decode the reservoir state (Hu et al., 2021).

2.4 Physics informed ESN

Recently, PIML gained popularity among scholars due to its accuracy for stock price prediction, trajectory estimation and traffic jam forecast (Barreau et al., 2021). The novel approach was firstly introduced by (Lee and Kang, 1990), who incorporated neural networks to solve partial differential equations (PDEs) back in 1989. Later, (González-García et al., 1998) suggested employing ANN without a constraint of Runge Katta schemes for time-stepping algorithms. A more recent paper (Raissi et al., 2019) applied deep learning for PIML to solve inverse and forward problems in PDEs. The weather forecasting problems are described by dynamical PDEs. Among the dynamical systems, the chaotic differential equations are the ones that are difficult to learn. The reason is the rate of divergence in their tangent space. Where the tangent space is a representation of feedback values in adjoint methods. Therefore, in the chaotic dynamical system, they diverge exponentially. In the case of employing the Lorenz equation, the neural network will result in infinite derivatives. Scholars investigating addressed this issue by using advanced adjoints (Dandekar et al., 2020). A more recent approach suggested by (Doan et al., 2019), utilizes the ESN. The main problem is to structure projection from multidimensional reservoir state W_{res} to actual output state W_{out}. Physics knowledge incorporated into the architecture of ESN allows fixing the behaviour of the reservoir (Figure 3). This eliminates the necessity of back-propagation and adjoints and transforms the learning problem into a least square regression to fit the W_{out}.

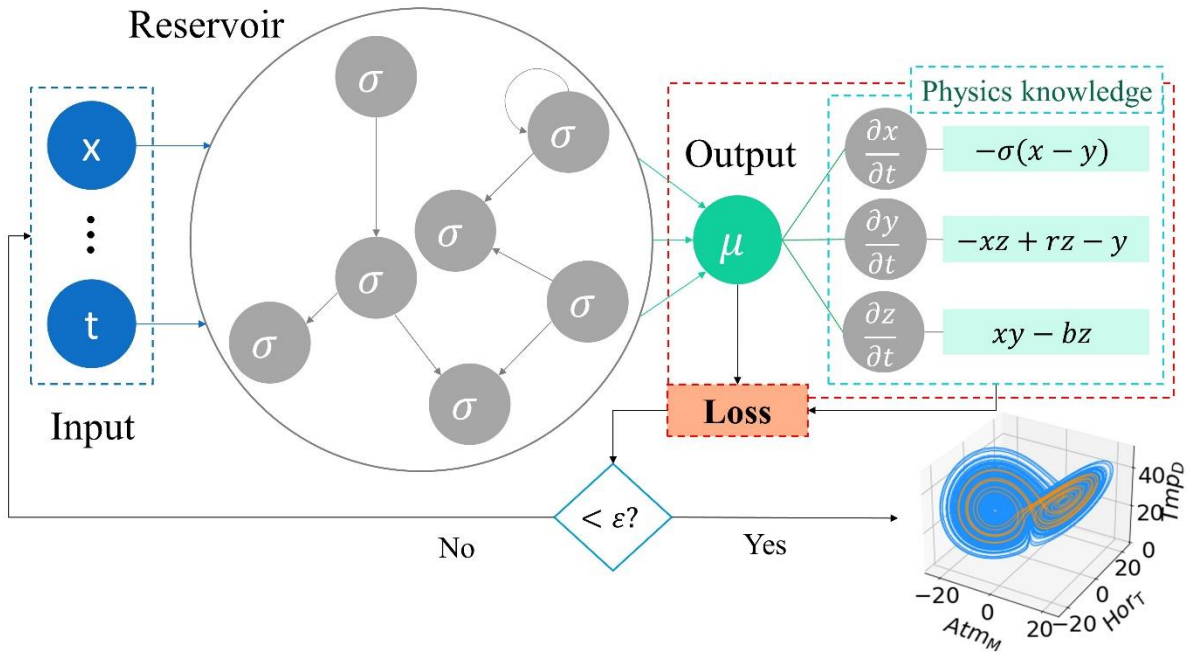

Figure 3. Physics informed Echo State Networks

The hybrid approach of the physics model and ESN employed in this study is based on the open-source algorithm provided by (Racca and Magri, 2021). The physical knowledge describing the system's chaotic behaviour is embedded by governing equations through input functions $PI[x(t)]$. This allows the model to gain information regarding the output trajectory of the future time step. The parameters of the updated output vector should be $\mu(t) = [\mu_1(t+1); 1; PI(x(t))]^T$. The open and closed loops are two types of configurations used to run the ESN. This study utilizes an open loop is used during training of the model and a closed loop is used for validation and testing (Racca and Magri, 2021). First, to meet the echo state condition the washout interval is introduced for independent initial reservoir state, $x(t) = 0$. Then, W_{out} is trained by minimizing the mean square error (MSE) of $\mu(t)$ and $x(t)$

$$MSE = \frac{1}{N_S N_M} \sum_i^{N_S} \left\| \mu(t) - PI(x(t)) \right\|^2 \quad (4)$$

where N_S and N_M denotes the number of samples and number of output neurons. Then, the $\|...\|$ is the least-square minimization problem. Generally, the regression is used for a linear system, whereas this problem is no more linear. Therefore, a stochastic output vector is incorporated into the reservoir fed-back matrix. The equation for ridge regression is presented as follows:

$$(W_{sb}\sigma(t)^T + \beta I)W_{out} = W_{sb}PI\big(x(t)\big)^T \quad (5)$$

where the horizontal concatenation of reservoir update state is represented as $W_{sb} \in R^{N_L N_S}$ and $x(t)^T \in R^{N_M N_S}$. Moreover, β and I denote the Tikhonov regularization parameter (Racca and Magri, 2021) and identity matrix, respectively. Finally, the validation and testing utilize the closed-loop configuration by introducing the washout interval. This allows the model to update the reservoir by sending output $\mu(t)$ back as an input in time-series prediction. As a result, the model is able to evolve.

3. Wind speed forecasting model

The wind speed forecasting model was based on the Gazanjyk wind speed dataset collected during 2020 and 2021. This section provides a dataset, decomposition analysis and prediction model for time-series analysis. The computation was carried on the Python open-source software and executed on a personal computer with an AMD Ryzen-5-3500U 2.10 GHz CPU and 4.00 GB of RAM.

3.1 Dataset and site description

Based on the finding of (Bahrami et al., 2019), this study evaluates wind speed performance for Gazanjyk. The author analysed the eighteen locations for potential wind energy development in Turkmenistan. Gazanjyk is located in the western region of Balkan with coordinates: 39.3 North latitude and 55.5 East longitude. According to references (Hu et al., 2021; Wang et al., 2019), this study provides four seasonal groups with a dataset in the range of 600 to 700 sample points each. Figure 4 depicts the seasonal directional distribution of wind speed in Gazanjyk ("National Committee of Hydrometeorology," 2021). It can be observed that autumn and spring shows the highest capacity compared to the performance of the rest data. For further analysis, each seasonal data was run through decomposition analysis.

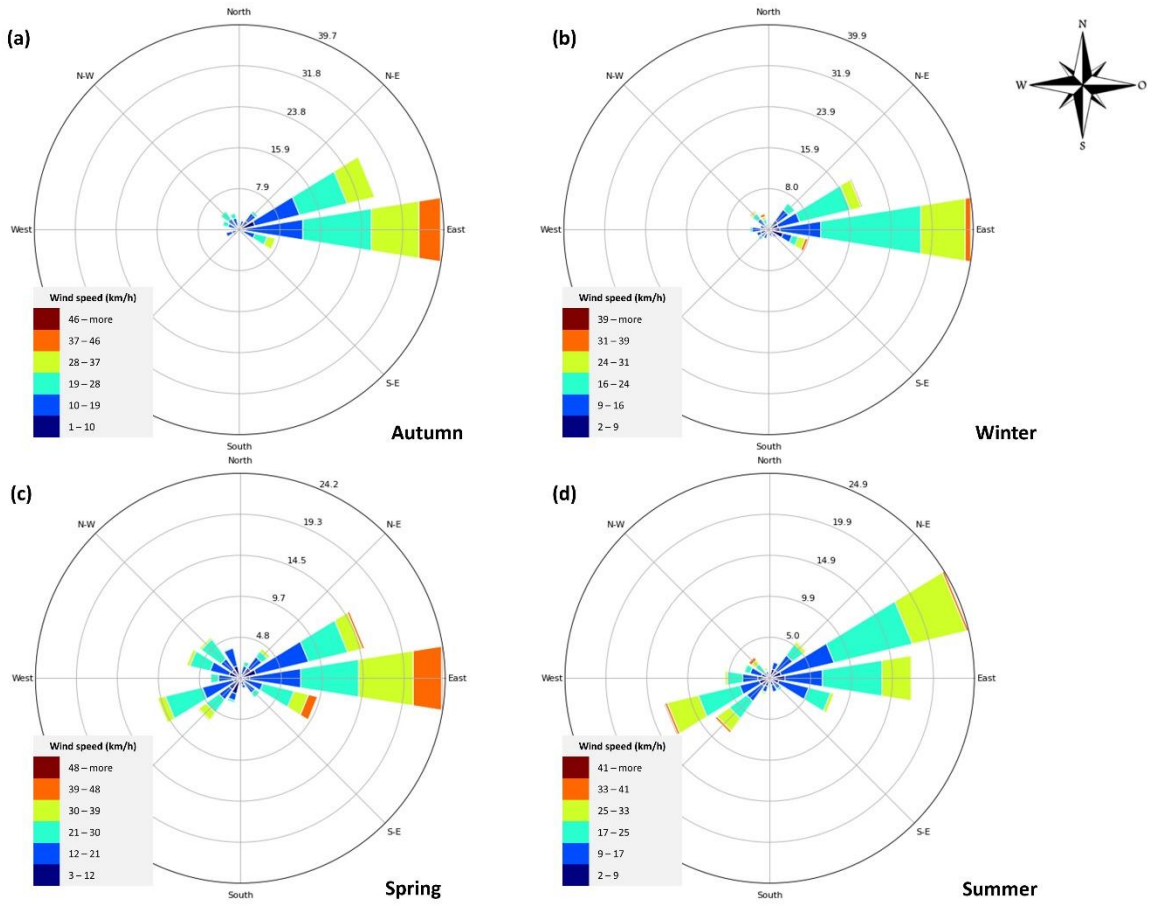

Figure 4. Directional distribution of winds in Gazanjyk

3.2 Data pre-processing

The WT is a common decomposition approach used for wind speed sequence data (Wang et al., 2019). Therefore, this study employs the WT with Daubechies function due to its advantages in providing a balanced presentation of smoothness and wavelength. The final algorithm decomposes wind speed data into low and high pass filters. This helps to understand original wind speed performance through low-frequency approximate sub-signal and several high-frequency detail sub-signals. Figure 5 presents the results of WT that serve as a basis for sequent steps in the wind speed forecasting model.

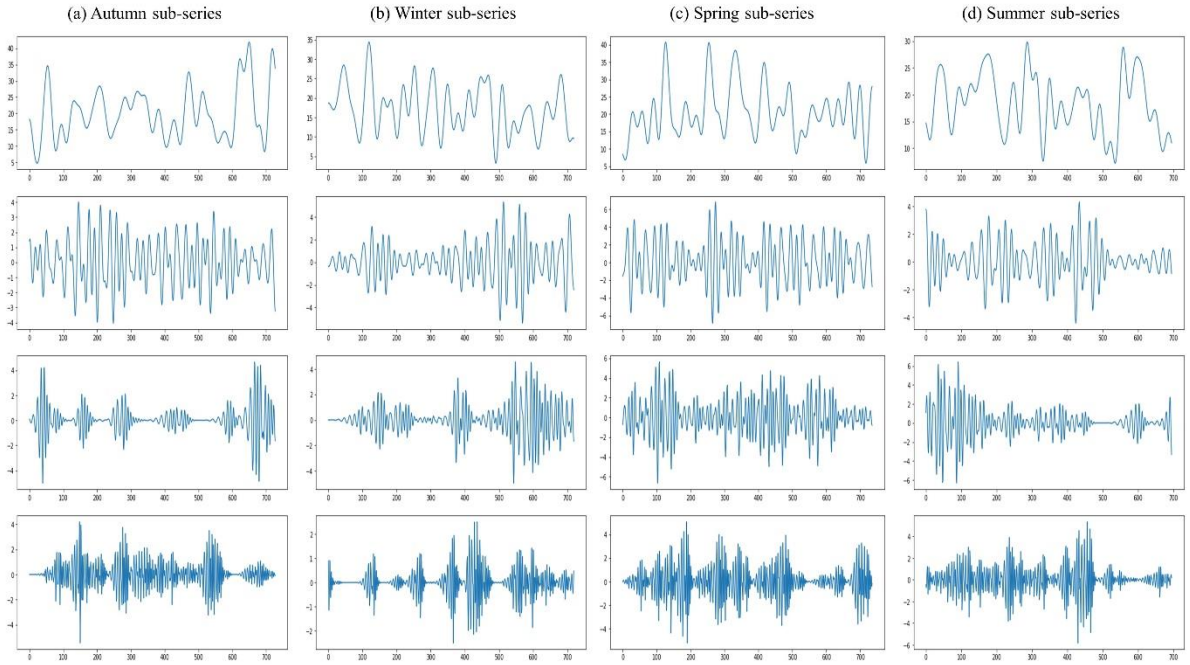

Figure 5. Figures (a) to (d) separately depict WT decomposition of wind speed data

3.3 Forecasting results and analysis

In this section, the results and simulation steps of RNN is explained to illustrate its advantages and shortcoming for computing wind speed data. First, the result of WT was set as a two-dimensional NumPy array with 600-700 observations, consisting of 16 input sub-signals and 4 output signals. The training of RNN was divided into sub-sequences with batch size 256. This was done to eliminate complexity for training the entire sequence and run the model with optimal CPU performance. Moreover, to cover the risk of overfitting the model, the performance of training is monitored by its weight after each epoch. In other words, the validation is done by monitoring the performance of the dataset, to stop the training in case it gets worse. Next, the gated recurrent unit (GRU) serves as the first layer of the network and requires a batch for an arbitrarily long sequence. The state size is 512 outputs for each time step. However, to match the required output vector with 4 signals, the fully connected dense layer is added to the model. After, using the scaler object the output signals were scaled between 0 and 1. To take this further, the sigmoid activation function is employed to hidden layer to limit the output of RNN to be scaled between 0 and 1 as well. On the other hand, considering the limitation of negative approximation in the sigmoid function, the last layer uses a linear activation function to take on arbitrary values. The last step is setting the lost function. The mean square error (MSE) is used to minimize the loss in matching the model's output with the original data. Consequently, the Keras model was compiled with an RMSprop optimizer. The reduce learning rate for the call back function was set for 1e-4 and patience of 0 epochs with factor 0.1. the RNN model in this study consists of 20 epochs with 100 steps per epoch.

Figure 6 depicts the results of prediction analysis in autumn, winter, spring, and summer, respectively. The blue line corresponds to the original wind speed observation, and the orange line represents the forecast. The training consists of a train and warmup period that let the model learn the dynamic behaviour. The MSE is employed to match the difference between these lines as close as possible. The vertical line represents the time step of the model, where at each time step the model proceeds with experience of previous steps. Thus, the warmup period helps to ignore initial steps to minimize the noise that may mislead the model. Figure 6 illustrates that the model has learned daily oscillations of wind speed. However, it frequently misrepresents the peaks of the original distribution. Therefore, the model is capable to mimic the wind speed swings in general, however, limited to match the unexpected peaks. In conclusion, the model is limited for accuracy given the wind speed input signals.

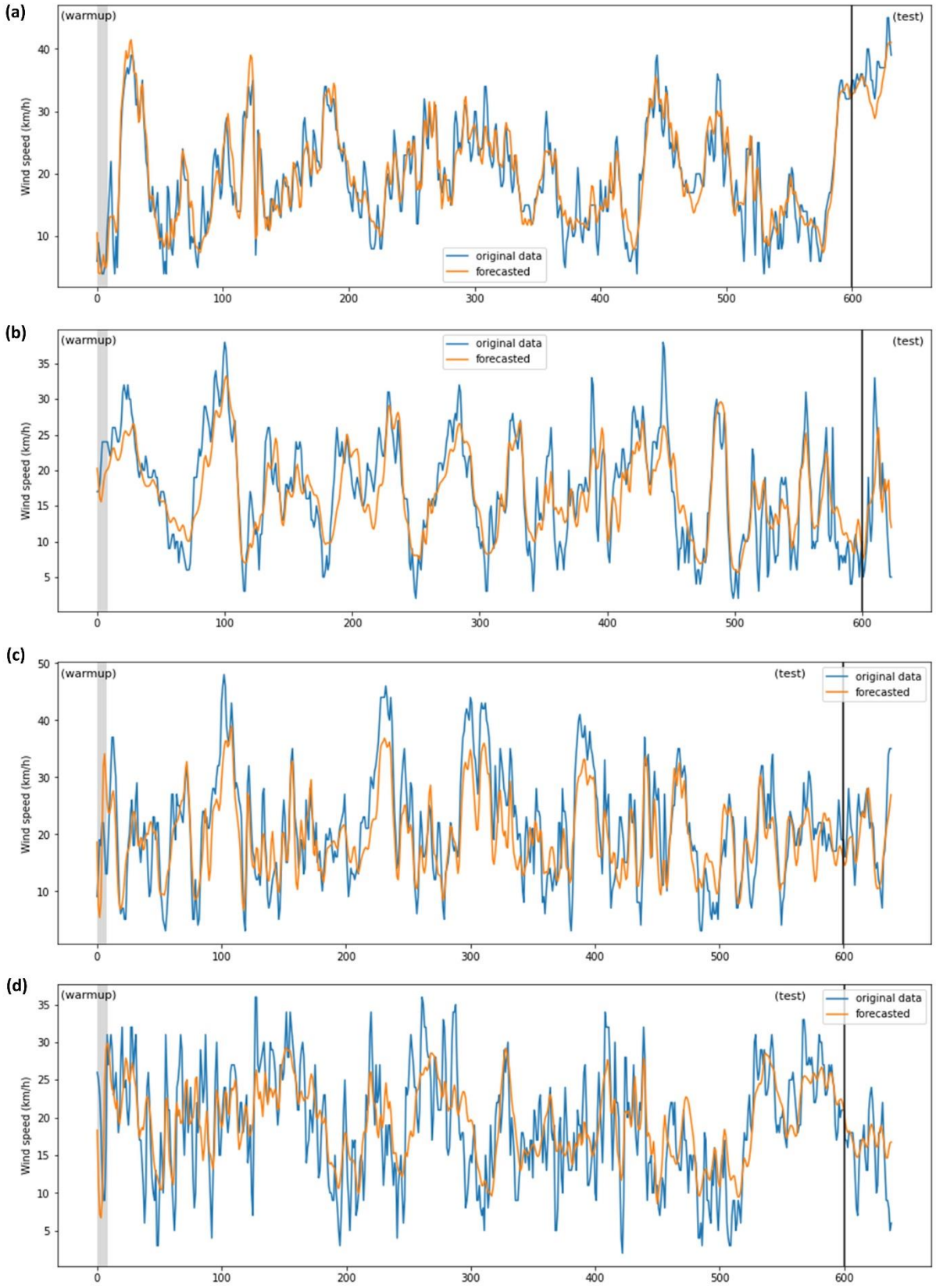

Figure 6. WT-RNN forecasting results

3.4 Discussion

In previous sub-sections, this paper provides a comprehensive performance of the WT-RNN model for wind speed forecasting framework. The chaotic behaviour of wind speed requires high accuracy of prediction analysis that

should be balanced to learn the dynamic distribution by not overfitting the model. This section discusses the performance of the proposed WT-RNN model based on the wind speed dataset collected during 2020-2021. First, the dataset consists of a 500-600 sample size that is sufficient for accurate prediction (Duan et al., 2021). Hence, it is stated that a larger dataset is always a key for better performance (Hu et al., 2021). Second, the structure of training and warmup period provides additional robustness to noise and disturbances of the proposed model. This could be observed in the general result of the WT-RNN model. Third, load checkpoints in the training dataset were implemented to eliminate the curse of dimensionality in the test dataset. Fourth, the sigmoid activation function enables the model to infer the hidden pattern. This is achieved by taking the sigmoid of each element in the hidden layer to identify the pattern. On the other hand, the proposed model has shown the following limitations. The first aspect is the long computation time. Even with a defined early stop for optimal loss in 20 epochs with 100 sub-epochs in each time-step, the proposed model has ended with a wall time of 26 min and 51 sec. Second, you cannot get convergence proof. The third aspect is the requirement for a large dataset.

One approach to overcome the previously stated shortcoming of the data-driven wind speed forecasting model is to acquire a larger dataset. It is stated (Zhang et al., 2018) that the availability of larger weather data would lead to more accurate performance. Furthermore, to describe the approximate behaviour of dynamic systems such as wind speed, it is suggested to employ minimum and maximum range prediction rather than forecasting actual value. Another approach is to enhance the model with more advanced architecture (Wang et al., 2019). There are limited studies that focus on this approach by incorporating physics information (Zhang and Zhao, 2021). The advantage of using a physics-based approach is stated in three aspects. First, is the incorporated definition of PDE into the model that do not require additional measurements. It is solely based on boundary conditions and initial conditions. Second, due to predefined conditions, the forecasting model is fast to compute. Third, the acquired solution is relatively easy and provides the error bounds. This is possible thanks to the contribution of researchers in evaluating the performance of partial differential algorithms through history (Barreau et al., 2021). Based on these propositions next section provides a weather forecasting model based on physics informed reservoir computing to learn chaotic atmospheric behaviour.

4. Atmospheric simulation

The atmospheric forecasting model was based on the Lorenz system due to its ability to describe the chaotic behaviour of the weather. This section introduces the Lorenz system and describes the architecture of the PI-ESN algorithm, its validation and sensitivity for time-series analysis. The section concludes with a discussion.

4.1 Lorenz system

Lorenz equation was put forward in 1963 by Edward Lorenz (Lorenz, 1963). The model aims to describe non-periodic flow in deterministic systems. The application of the framework is based on the natural convection system for weather forecast, which was proposed by following PDE:

$$\frac{\partial x}{\partial t} = -\sigma(x - y),\ \frac{\partial y}{\partial t} = -xz + rz - y,\ \frac{\partial z}{\partial t} = xy - bz \tag{6}$$

where $\frac{\partial x}{\partial t}$ represents the intensity of the atmospheric convection motion, $\frac{\partial y}{\partial t}$ represents a difference in horizontal temperature and $\frac{\partial z}{\partial t}$ is a departure of the temperature when atmospheric convention does not occur. Furthermore, the dimensionless parameters are σ, which denotes Prandtl number and suggested to be 10, r equal to 28 and represent Rayleigh number, b denoted region microclimate and assumed to be 8/3 (Zhang et al., 2018). These standard parameters of the Lorenz system will spawn the chaotic behaviour of the atmospheric system (Doan et al., 2019). The model is anticipated to learn this pattern and accurately describe the dynamic motion of a targeted system.

4.2 Modelling and validation settings

In this subsection, the study describes the architecture of the PI-ESN model and demonstrates its effectiveness and validity in learning the Lorenz system. The validation function is employed to identify hyperparameters. This is achieved by minimizing an MSE with respect to the spectral radius ρ and input scaling x_{in} for a fixed length of validation interval. In the study of (Racca and Magri, 2021), they proposed a chaotic version of Recycle validation (CRV) with emphasis on two main factors such as, prediction interval and signal behaviour that intended to reproduce. To elaborate this further, the authors highlight that the first objective is to predict multiple intervals in

a trajectory that expands. The second reason is an ergodic trajectory with no time dependency in the mean of the studied signal. This is explained by identical intervals generated during the process in the physical model.

The CRV is used to tune the hyperparameters by exploiting information from both open and closed loops configurations and chaotic extensions were utilized for shifting the validation intervals forward by one LP (Racca and Magri, 2021). The robustness of CRV was evaluated on the Lorenz system. The time series for the Lorenz system was generated by using forward Euler and splitting the dataset into the washout, training, validation, and test. The time step between two time-instants is $\partial t = 0.9 * 10^{-3}$ Lyapunov time (LP), where LP is ≈ 1.1 (Racca and Magri, 2021). The data range for training, validation and test are 1 to 9 LTs, 9 to 12 LTs, and 12 to 15 LTs, respectively. The parameters for PI-ESN are taken to be $r = 100$ neurons, $s = 97\%$, $\beta = 10^{-11}$ and $b_{in} = 1$ (Racca and Magri, 2021). The $\log_{10} MSE$ was used during validation to identify hyperparameters x_{in} and ρ with a given range of $[0.5,5] \times [0.1,1]$. The reason for this range is given with an intention for spectral radius ρ intends to mimic the echo state property and input scaling x_{in} to normalize the data.

Figure 7 presents the solution of PI-ESN based on the Lorenz system. The behaviour of signal distribution illustrated in the first figure consists of the strength of the atmospheric convection motion represented in the blue line, the red line represents a difference in horizontal temperature and the green line is a departure of the temperature. The second figure is the Lorenz attractor during the long run in blue and the results of forecasting are illustrated in orange.

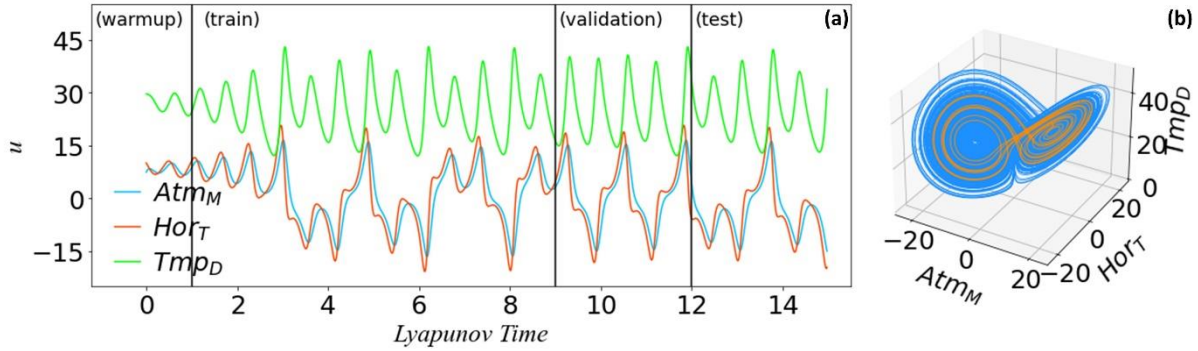

Figure 7. Atmospheric model solution. a) signal distribution b) Lorenz attractor and prediction results

4.3 Sensitivity analysis

The analysis of sensitivity was conducted to compare optimization results from validation and test datasets. (Racca and Magri, 2021) stated the outperformance of Bayesian optimization (BO) in identifying minimum MSE within the hyperparameter space of the validation set. BO computes objective function without the need for gradient information and employs Gaussian Process (GP) regression to incorporate knowledge of the entire search space. The reconstruction of the next objective function is achieved by the mean and standard deviation of GP, which was used to optimize the acquisition function of the following point in the enlarged dataset. This study employed a scikit-optimize Python library with BO based on 5×5 starting points and GP regression computed 24 points. The performance of PI-ESN was computed by MSE for both validation and test datasets. The results of the chaotic system could be observed in GP process reconstruction. Figure 8 presents a spatial illustration of MSE of GP reconstruction based on input scaling x_{in} and spectral radius ρ in validation and test datasets. The last figure represents their differences in terms of 30 grids in $\log_{10} MSE$.

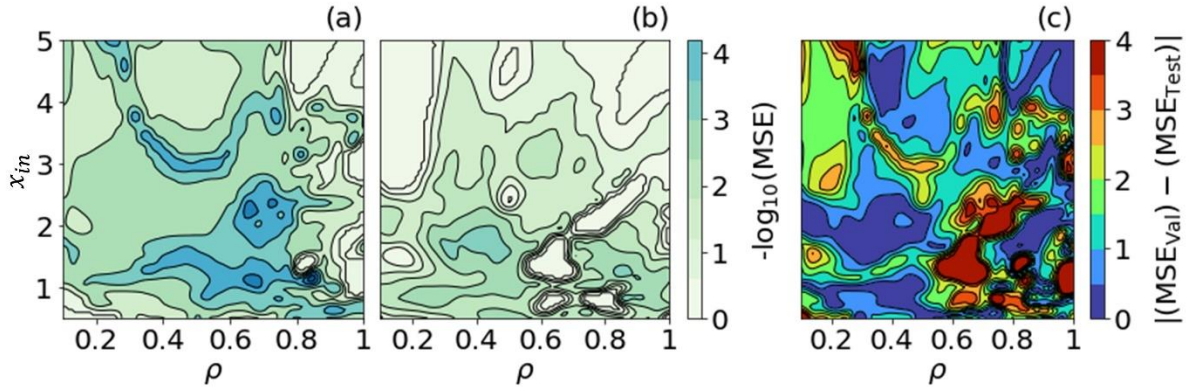

Figure 8. Gaussian Process alteration for the validation dataset in (a) and test dataset in (b), and their differences in (c)

4.4 Discussion

This section provided PI-ESN as a prediction approach to model the chaotic behaviour of the atmospheric system. Incorporated physics informed model and reservoir computing algorithm demonstrates prediction accuracy and robustness that has been justified with analysis of sensitivity. The forecasting results suggest that the model is capable of managing physical phenomena of atmospheric motion, changes in temperature and relative factors influencing wind speed and wind power operation scenarios. Since the proposed hybrid method is a combination of physics informed model and machine learning technique, the accurate performance of PI-ESN can be appointed to seven main advantages inherited from both approaches. First, the accuracy of the model is independent of the amount of dataset. Either a small or large dataset is sufficient for accurate prediction. Second, the model is robust to noise and disturbances. Third, there is no curse of dimensionality in a reservoir. The model was tested on a high dimension Lorenz system (Champion et al., 2019). Fourth, due to the implementation of ESN, the model is capable to infer the hidden pattern. Fifth, the GP reconstruction of CRV validation and testing allows performing model tuning. Sixth, conventional RNN would require large computational time, hence the reservoir computing approach deals with it by directly learning the output with reservoir update function. The computation cost on a personal computer was 11 min and 1 sec. Seventh is the challenge of convergence proof for the optimization, whether it is an optimal solution for studied PDE. (Racca and Magri, 2021) illustrated the approximate convergence with MSE for asymptotic values in RCV validation for the Lorenz system. Therefore, adopting BO for hyperparameters results in this model achieving less stochastic output and more fast convergence.

In summary, the novel PI-ESN model overcomes the previously stated challenges of conventional PIML approaches. The proposed hybrid method demonstrates accurate performance and robustness for chaotic system forecasting that outperform conventional stand-alone wind speed prediction models such as physical, statistical and machine learning algorithms.

5. Conclusion

Considering increasing global energy consumption and the anticipated rise of energy demand, there is a need for renewable energy sources to support sustainable advancement in major energy suppliers. Wind power is considered an environmentally sustainable source of renewable energy with less significant attention given to its utilization. The study focuses on wind power due to the potential of solar and wind power sources that is overshadowed by the exploitation of oil and gas resources in Turkmenistan, which is the major energy supplier in the Central Asian market. Based on the findings of previous studies (Bahrami et al., 2019), this research contributes with a case study conducted in Gazanjyk. The paper proposes a weather forecasting method by providing a data-driven approach for wind speed and PIML for the atmospheric model. The first model is based on seasonal wind speed data collected during 2020-2021. The WT is applied to decompose raw data into several sub-series. After the noise-free data was used as input in the RNN model. The results of the hybrid method suggest its ability to learn the general distribution of the dynamic system, however, it was limited to match unexpected peaks for accurate wind speed prediction. This leads to the second model implemented to forecast the chaotic behaviour of atmospheric motion. The Lorenz system was incorporated in PI-ESN to define boundary conditions and to control the chaotic extension of a three-dimensional model. The BO optimization of hyperparameters in CRV for validation and testing dataset illustrates the accuracy and robustness of this model. This implies the

success of the PI-ESN model in extracting the nonlinear behaviour of atmospheric systems and potential in practical use for wind power forecasting.

References:

Bahrami, A., Teimourian, A., Okoye, C.O., Khosravi, N., 2019. Assessing the feasibility of wind energy as a power source in Turkmenistan; a major opportunity for Central Asia's energy market. Energy 183, 415–427. https://doi.org/10.1016/j.energy.2019.06.108

Barreau, M., Liu, J., Johansson, K.H., 2021. Learning-based State Reconstruction for a Scalar Hyperbolic PDE under noisy Lagrangian Sensing. Proc. Mach. Learn. Res.

Bollt, E., 2021. On explaining the surprising success of reservoir computing forecaster of chaos? The universal machine learning dynamical system with contrast to VAR and DMD. Chaos An Interdiscip. J. Nonlinear Sci. 31, 013108. https://doi.org/10.1063/5.0024890

Champion, K., Lusch, B., Kutz, J.N., Brunton, S.L., 2019. Data-driven discovery of coordinates and governing equations. Proc. Natl. Acad. Sci. 116, 22445–22451. https://doi.org/10.1073/PNAS.1906995116

Cui, Yanbin, Huang, C., Cui, Yanping, 2019. A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm. Environ. Sci. Pollut. Res. 27, 7353–7365. https://doi.org/10.1007/S11356-019-07402-1

Dandekar, R., Rackauckas, C., Barbastathis, G., 2020. A Machine Learning-Aided Global Diagnostic and Comparative Tool to Assess Effect of Quarantine Control in COVID-19 Spread. Patterns 1, 100145. https://doi.org/10.1016/J.PATTER.2020.100145

Doan, N.A.K., Polifke, W., Magri, L., 2019. Physics-Informed Echo State Networks for Chaotic Systems Forecasting. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11539 LNCS, 192–198. https://doi.org/10.1007/978-3-030-22747-0_15

Duan, Jikai, Zuo, H., Bai, Y., Duan, Jizheng, Chang, M., Chen, B., 2021. Short-term wind speed forecasting using recurrent neural networks with error correction. Energy 217, 119397. https://doi.org/10.1016/J.ENERGY.2020.119397

Elsaraiti, M., Merabet, A., 2021. Application of Long-Short-Term-Memory Recurrent Neural Networks to Forecast Wind Speed. Appl. Sci. 11, 2387. https://doi.org/10.3390/APP11052387

González-García, R., Rico-Martínez, R., Kevrekidis, I.G., 1998. Identification of distributed parameter systems: A neural net based approach. Comput. Chem. Eng. 22, S965–S968. https://doi.org/10.1016/S0098-1354(98)00191-4

Gupta, D., Natarajan, N., Berlin, M., 2021. Short-term wind speed prediction using hybrid machine learning techniques. Environ. Sci. Pollut. Res. 1–19. https://doi.org/10.1007/S11356-021-15221-6

Hu, H., Wang, L., Tao, R., 2021. Wind speed forecasting based on variational mode decomposition and improved echo state network. Renew. Energy 164, 729–751. https://doi.org/10.1016/J.RENENE.2020.09.109

Jaeger, H., Haas, H., 2004. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. Science (80-.). 304, 78–80. https://doi.org/10.1126/SCIENCE.1091277

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. https://doi.org/10.1038/s42254-021-00314-5

Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H.A., Marcus, P., Anandkumar, A., Hassanzadeh, P., Prabhat, 2021. Physics-informed machine learning: case studies for weather and climate modelling. Philos. Trans. R. Soc. A 379. https://doi.org/10.1098/RSTA.2020.0093

Kumar, D., Mathur, H.D., Bhanot, S., Bansal, R.C., 2020. Forecasting of solar and wind power using LSTM RNN

for load frequency control in isolated microgrid. Int. J. Model. Simul. 41, 311–323. https://doi.org/10.1080/02286203.2020.1767840

Lee, H., Kang, I.S., 1990. Neural algorithm for solving differential equations. J. Comput. Phys. 91, 110–131. https://doi.org/10.1016/0021-9991(90)90007-N

Liu, H., Yu, C., Wu, H., Duan, Z., Yan, G., 2020. A new hybrid ensemble deep reinforcement learning model for wind speed short term forecasting. Energy 202, 117794. https://doi.org/10.1016/J.ENERGY.2020.117794

Lorenz, E.N., 1963. Deterministic Nonperiodic Flow. Geophys. Res. Lett. 20, 130–141. https://doi.org/10.1029/2020GL089283

Natarajan, N., Vasudevan, M., Rehman, S., 2021. Evaluation of suitability of wind speed probability distribution models: a case study from Tamil Nadu, India. Environ. Sci. Pollut. Res. 1–14. https://doi.org/10.1007/S11356-021-14315-5

National Committee of Hydrometeorology [WWW Document], 2021. URL http://www.meteo.gov.tm/en/ (accessed 9.2.21).

Racca, A., Magri, L., 2021. Robust Optimization and Validation of Echo State Networks for learning chaotic dynamics. Neural Networks 142, 252–268. https://doi.org/10.1016/J.NEUNET.2021.05.004

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707. https://doi.org/10.1016/J.JCP.2018.10.045

Snoun, H., Bellakhal, G., Kanfoudi, H., Zhang, X., Chahed, J., 2019. One-way coupling of WRF with a Gaussian dispersion model: a focused fine-scale air pollution assessment on southern Mediterranean. Environ. Sci. Pollut. Res. 26, 22892–22906. https://doi.org/10.1007/S11356-019-05486-3

Srivastava, R., Bran, S.H., 2018. Impact of dynamical and microphysical schemes on black carbon prediction in a regional climate model over India. Environ. Sci. Pollut. Res. 25, 14844–14855. https://doi.org/10.1007/S11356-018-1607-0

Tian, Z., 2020. Preliminary Research of Chaotic Characteristics and Prediction of Short-Term Wind Speed Time Series. Int. J. Bifurc. Chaos 30. https://doi.org/10.1142/S021812742050176X

Wang, H., Lei, Z., Liu, Y., Peng, J., Liu, J., 2019. Echo state network based ensemble approach for wind power forecasting. Energy Convers. Manag. 201, 112188. https://doi.org/10.1016/J.ENCONMAN.2019.112188

Zhang, J., Zhao, X., 2021. Spatiotemporal wind field prediction based on physics-informed deep learning and LIDAR measurements. Appl. Energy 288, 116641. https://doi.org/10.1016/J.APENERGY.2021.116641

Zhang, X., Zhang, Z., Su, G., Tao, H., Xu, W., Hu, L., 2019. Buoyant wind-driven pollutant dispersion and recirculation behaviour in wedge-shaped roof urban street canyons. Environ. Sci. Pollut. Res. 26, 8289–8302. https://doi.org/10.1007/S11356-019-04290-3

Zhang, Y., Li, R., Zhang, J., 2021. Optimization scheme of wind energy prediction based on artificial intelligence. Environ. Sci. Pollut. Res. 28, 39966–39981. https://doi.org/10.1007/S11356-021-13516-2

Zhang, Y., Pan, G., 2020. A hybrid prediction model for forecasting wind energy resources. Environ. Sci. Pollut. Res. 27, 19428–19446. https://doi.org/10.1007/S11356-020-08452-6

Zhang, Y., Zhang, C., Gao, S., Wang, P., Xie, F., Cheng, P., Lei, S., 2018. Wind Speed Prediction Using Wavelet Decomposition Based on Lorenz Disturbance Model. IETE J. Res. 66, 635–642. https://doi.org/10.1080/03772063.2018.1512384

Zhang, Y., Zhao, Y., Kong, C., Chen, B., 2020. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. Energy Convers. Manag. 203, 112254. https://doi.org/10.1016/J.ENCONMAN.2019.112254