Heliyon 10 (2024) e34807

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Adaptive machine learning for forecasting in wind energy: A dynamic, multi-algorithmic approach for short and long-term predictions

Mutaz AlShafeey *, Csaba Csaki

Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Budapest, Fővám tér 13-15, H-1093, Hungary

A R T I C L E I N F O

Keywords:
Wind energy forecasting
Machine learning techniques
Dynamic hybrid model
Comparative evaluation method
Short and long term forecasting

A B S T R A C T

This study elucidates the formulation and validation of a dynamic hybrid model for wind energy forecasting, with a particular emphasis on its capability for both short-term and long-term predictive accuracy. The model is predicated on the assimilation of time-series data from past wind energy generation and employs a triad of machine learning algorithms: Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbors (K-NN). Empirical data, harvested from a 2 MW grid-connected wind turbine, served as the basis for the training and validation phases. A comparative evaluation methodology was devised to scrutinize the performance of each constituent algorithm across a diverse array of metrics. This evaluation framework facilitated the identification of individual algorithmic limitations, which were subsequently mitigated through the implementation of a dynamic switching mechanism within the hybrid model. This innovative feature enables the model to adaptively select the most efficacious forecasting technique based on historical performance data. The hybrid model demonstrated superior forecasting accuracy in both, short-term energy forecasts at 15-min intervals over a day, and in broad, long-term. It recorded a Normalized Mean Absolute Error (NMAE) of 5.54 %, which is notably lower than the NMAE range of 5.65 %–9.22 % observed in other tested models, and significantly better than the average NMAE found in the literature, which spans from 6.73 % to 10.07 %. Such versatility renders it invaluable for grid operators and wind farm management, aiding in both operational and strategic planning. The study's findings not only contribute to the existing body of knowledge in renewable energy forecasting but also suggest the hybrid model's broader applicability in various other predictive analytics domains.

1. Introduction

1.1. Background

Harnessing wind energy offers a promising avenue for both sustainable economic growth and environmental conservation. However, the stochastic nature of wind patterns presents a formidable challenge in the seamless integration and operational management of wind energy units within electrical grids. This issue becomes particularly acute given the large generation capacities often

* Corresponding author. Budapest, Fővám tér 13-15, H-1093, Hungary.
E-mail addresses: Mutaz.AlShafeey@uni-corvinus.hu (M. AlShafeey), Csaki.Csaba@uni-corvinus.hu (C. Csaki).

https://doi.org/10.1016/j.heliyon.2024.e34807
Received 21 September 2023; Received in revised form 15 July 2024; Accepted 17 July 2024
Available online 25 July 2024
2405-8440/

associated with wind farms [1]. Even minor fluctuations in wind speed and other meteorological variables can disproportionately affect grid stability, making the task of managing these units exceedingly complex.

To address this challenge, precise forecasting of wind energy production has emerged as a pivotal strategy. Accurate forecasts are indispensable for maintaining grid stability and are typically segmented into various time horizons such as short-term (up to 24 h), medium-term (several days), and long-term (weeks to years) [2–4]. While there is no universally accepted criterion for categorizing these time horizons, effective forecasting equips grid operators with invaluable insights [5]. These insights are particularly useful for real-time and 'day-ahead' energy market trading, thereby reducing imbalance costs and penalties [6].

Short-term forecasting is particularly important as it allows for the optimization of renewable assets bidding in day-ahead and intraday markets, leading to significant cost reductions [7]. Further, accurate short-term forecasts can help grid operators better manage the integration of these renewable resources into the power grid, ensuring a stable and reliable electricity supply [7]. Moreover, short-term forecasting also plays a key role in the financial viability of wind energy projects. Providing accurate predictions of wind energy production allows energy producers to effectively participate in electricity markets. This can lead to improved profitability for wind energy projects and stimulate further investments in wind energy technologies [8]. Long-term forecasting, on the other hand, is crucial for strategic planning and operation [9]. It aids in identifying regulatory changes, technological advancements, and understanding the increasing integration of energy markets [10]. Long-term forecasts are essential for making strategic decisions about investments, market-entry, and capacity build-out [10].

As specifying short and long-term forecasting can be relative, defining intervals by the number of points may offer increased scientific objectivity in certain fields like the electricity spot market. However, this approach may not always be superior to time-based intervals. Most importantly, energy sector operations, especially those involving power grids, are normally organized around time-based schedules [11]. These systems require forecasts that directly map onto time interval schedules. Furthermore, in the context of wind energy forecasting, aligning intervals with the inherent temporal cycles of weather such as seasonal shifts provides a crucial advantage [12]. Time-based intervals correspond to the natural occurrences and fluctuations in wind patterns, which are essential for accurate predictions [13]. Additionally, the integration of wind forecasts into broader weather prediction models, which are also based on time intervals, necessitates a uniform time-based approach to ensure coherence and compatibility. Beyond the realm of technical and scientific utility, time intervals are inherently more communicable and understandable to a broad audience, which is a non-trivial factor in the dissemination and application of wind forecasts. Thus, while point-based intervals can enhance precision in certain analyses, the multifaceted nature of wind forecasting and its applications justifies the continued use of time intervals in this domain.

Expanding on the intricacies of wind energy forecasting, a multitude of approaches have been developed. These approaches primarily fall into three core domains: physical models grounded in fundamental physical principles, statistical models that leverage historical data for future projections, and Machine Learning (ML) or Artificial Intelligence (AI) models that utilize computational algorithms to learn from data and generate forecasts [14]. Within the machine learning domain, algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and k-Nearest Neighbors (KNN) have gained prominence. These algorithms offer varying degrees of accuracy and computational efficiency and are often tailored to meet specific forecasting challenges [15]. While each of these machine learning techniques offers distinct advantages, they also present inherent limitations when deployed independently. To address these limitations, some studies have explored the use of hybrid models that combine multiple algorithms [16]. These hybrid models can be broadly categorized into static and dynamic types. Static models are pre-configured with a fixed combination of algorithms and lack the adaptability to respond to real-time changes. While effective in certain scenarios, their rigidity can be sub-optimal for addressing the complex and variable factors influencing wind energy production. In contrast, dynamic models are designed to adapt to fluctuating conditions by dynamically toggling between different algorithms. Despite their potential for greater accuracy and adaptability, dynamically adaptive hybrid models are notably scarce in existing literature. This represents a significant research gap, as such models could offer a more versatile and adaptive framework for wind energy forecasting, enhancing the reliability and accuracy of predictions crucial for both short-term and long-term planning in renewable energy systems. Hence, this article focuses on the development and examination of dynamic hybrid and machine learning models for wind energy forecasting. These models aim to offer enhanced adaptability and precision in predicting wind energy outputs, addressing the limitations of current forecasting methods.

1.2. Literature

Recent trend shows a growing number of various applications that are based on AI technologies, and this applies to wind forecasting as well [17]. A vast number of researchers are improving intelligent technologies to accurately predict wind speed and power. Sideratos and Hatziargyriou [18] used machine learning methods for short-term wind power forecasting. They applied a combination of fuzzy logic and neural network techniques to forecast wind farm power output. The authors stated that the results can be used effectively for the operational planning of wind farms 1–48 h ahead. Rahmani et al. [19] proposed a hybrid system that consists of two meta-heuristic techniques under the category of swarm intelligence to forecast the energy output of a wind farm. The empirical results indicate that the proposed technique can estimate the output wind power based on the wind speed and ambient temperature with acceptable accuracy. Najeebullah et al. [20] proposed a wind power prediction system that uses a combination of machine learning techniques for feature selection and regression. The authors concluded that the proposed model performs better than the existing prediction models in terms of performance measures and can be used as an effective wind power prediction model. Chi et al. [21] studied the performance of direct and iterative methods for multi-step ahead wind speed forecasting. Three machine learning methods including linear regression, multi-layer perceptron, and support vector machine were developed. The results show that neither direct nor iterative forecasting can outperform the other in terms of all the error measures. Wang et al. [22] proposed hybrid models utilizing various ML techniques such as Support Vector Regression (SVR) with seasonal index adjustment (SIA) and Elman recurrent neural network (ERNN). The hybrid models were applied in three different sites in China and predicted the behaviors of daily wind in a reasonable way. Yao Zhang and Jianxue Wang [23] developed a combination of the KNN and the kernel density estimator (KDE) method for probabilistic wind power forecasting. The suggested approach showed good forecasting performance.

Building upon the established foundations of statistical and machine learning techniques in wind energy forecasting, more recent studies further advance the field by addressing specific challenges and incorporating innovative methodologies. For instance, in a recent study by Sharma et al. [24], the authors delved into improving wind energy forecast accuracy through hierarchical reconciliation methods that account for temporal correlation. This research contributed to the field by addressing the variability and uncertainty inherent in wind energy, which poses challenges for power system operations. They proposed an approach that aligns forecasts at varying temporal scales, enhancing decision-making processes in energy management. Another research by Moreno et al., highlighted the integration of machine learning with decomposition techniques such as Singular Spectral Analysis (SSA) and Variational Mode Decomposition (VMD) to improve forecasting accuracy [25]. The authors demonstrated how this combined approach can significantly mitigate forecast errors, thus offering a more reliable basis for operational and maintenance scheduling in wind farms.

Moreover, Junior et al. addressed improving wind energy forecast through optimal reconciliation methods [26]. This approach capitalizes on the hierarchical structure of wind energy data across different geographical locations and forecast horizons, integrating temporal correlation to enhance prediction precision. The study emphasized the complexity of wind energy patterns due to varying conditions such as wind speed, direction, and other meteorological factors, making accurate forecasting crucial for effective energy management and grid integration. The need for accurate forecasts was also addressed by Hao Chen by developing a method for wind model downscaling that employs statistical regression techniques [27]. Chen's methodology, which emphasized the integration of topographic and meteorological data, demonstrated a significant reduction in predictive errors compared to existing models. Da Silva et al. also employed integration and combination of models, outlining a method that merges multi-stage signal decomposition techniques — specifically variational mode decomposition (VMD) and singular spectrum analysis (SSA) — with a stacking-ensemble learning approach, aiming to boost the accuracy of short-term, multi-step wind speed forecasts [28].

By addressing limitations in the forecasting capabilities of basic models and traditional multi-step-ahead output strategies, Yun Wang et al. developed an improved Wavenet network for multi-step-ahead predictions of wind speed and power [29]. This model integrates two Wavenet networks with an encoder-decoder framework enhanced by a multi-head self-attention mechanism. The use of teacher forcing as the multi-step-ahead output strategy marks an improvement over existing methodologies.

Despite the numerous contributions to wind energy forecasting, there remains a broad spectrum of research gaps, particularly in the crafting of models that span various forecasting horizons. A review of solar and wind energy found that very few researchers studied and developed models to predict wind energy for different horizons i.e. long and short-term [30]. Furthermore, the existing literature on dynamic energy forecasting models is notably sparse. Among the limited studies that have ventured into this domain, the focus has predominantly been on short-term, near-future predictions, often restricted to just a few time steps ahead [31]. Remarkably, there is a conspicuous absence of research that extends these dynamic hybrid models to a comprehensive 24-h forecasting horizon. This gap in the literature underscores the need for more expansive and temporally inclusive hybrid dynamic models for energy forecasting.

1.3. Contributions

The novelty of this study is highlighted by two primary and interconnected contributions. Firstly, a pioneering dynamic hybrid model for wind energy forecasting is introduced. This model uniquely adapts in real-time, with a variety of machine learning techniques — namely ANN, SVM, and K-NN — being selectively employed based on the immediate data inputs. This enhances adaptability and significantly boosts predictive accuracy for both immediate, short-term forecasts (up to 24 h ahead in 15-min segments) and broader, long-term energy production projections. A considerable evolution from traditional static hybrid models is marked by this approach, offering a new level of flexibility and precision in predictions.

Secondly, a notable void in existing research is effectively filled by the proposed model, providing a forecasting tool adept across diverse temporal dimensions. In contrast to conventional models that are limited to either short-range or long-range predictions, this innovative hybrid model excels across both spectra, making it exceptionally valuable to grid operators, especially within the stringent confines of the EU's regulatory demands. The focus of the study is honed in on real-time wind energy data, with the recognition of the direct impact of environmental conditions on wind power generation. By the model being intentionally designed to require minimal data inputs, not only its operation is streamlined but this also results in a broader applicability under various forecasting conditions. This is particularly advantageous in contexts where supplementary weather information might be unreliable or not forthcoming. By variables like temperature, humidity, and air pressure being consciously omitted, the core innovation of the model is spotlighted: its robust performance, relying solely on wind energy time-series data, is demonstrated. Furthermore, comparative analyses between this hybrid model and conventional standalone methods are included, all within the EU's strict regulatory framework for grid management, which mandates precise 24-h-ahead forecasts at 15-min intervals [32,33]. This not only underscores the model's relevance but also its compliance and utility in real-world applications.

This manuscript is organized as follows: Section 1 provides an introduction to the topic. Section 2 elaborates on the methodologies employed, including data collection approaches and the specifics of machine learning and hybrid model development. Section 3 presents the empirical findings of the study. Section 4 delves into a comprehensive analysis of these findings. The final section concludes the paper, summarizing key outcomes along with their broader implications and potential directions for future research.

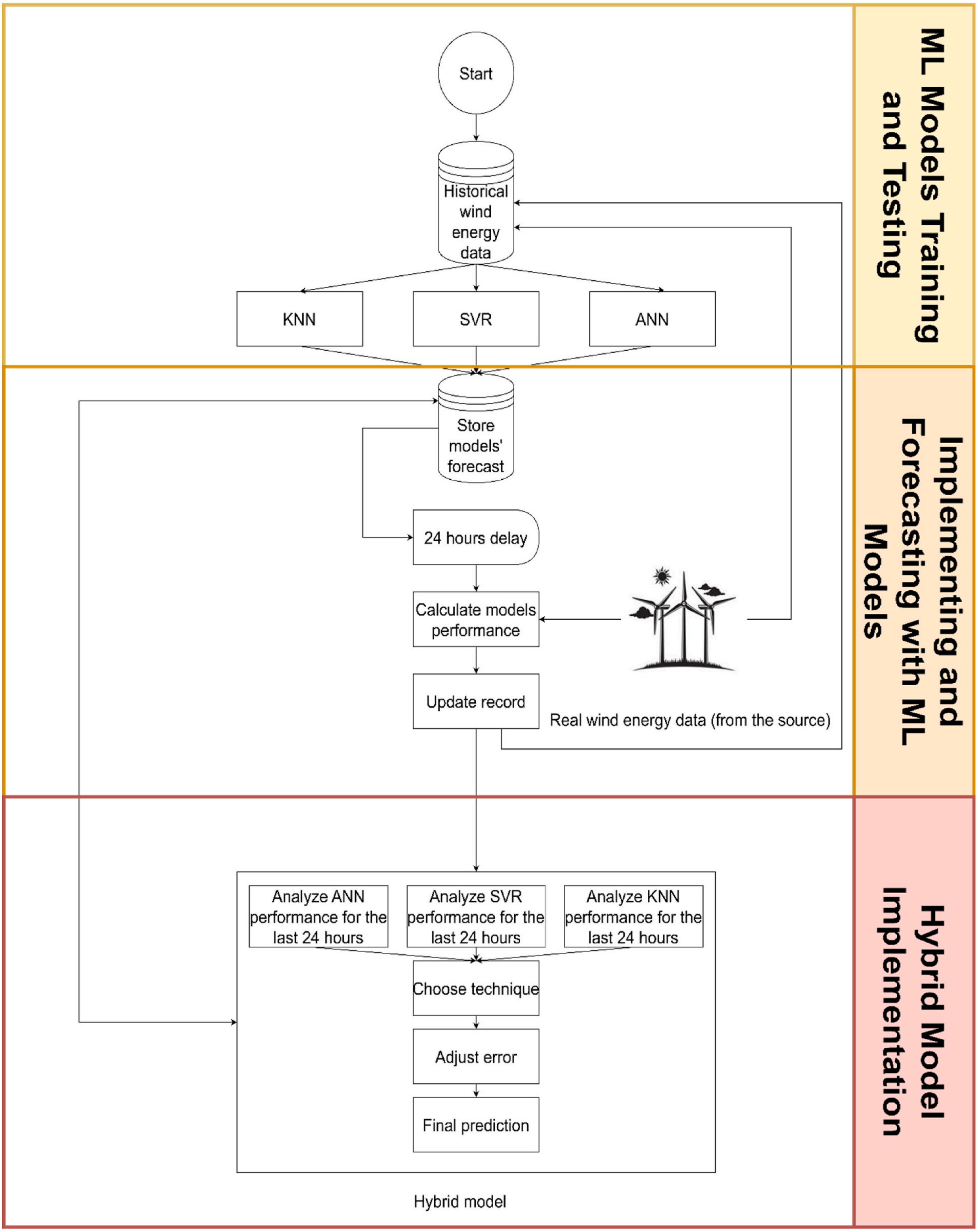

Fig. 1. General overview of the methodology flowchart.

2. Methods

Fig. 1 presents a schematic representation of the methodology used in this study. Note, that specific methodologies deployed will be clarified further in their corresponding sections. The methodology starts with the collection of historical wind energy data, which serves as the foundational input for the subsequent processes. This data is then used in the "ML Models Training and Testing" stage, where three different machine learning models - K-Nearest Neighbors (KNN), Support Vector Regression (SVR), and Artificial Neural Networks (ANN) - are trained and tested to evaluate their forecasting capabilities. Following model training, the process moves to the "Implementing and Forecasting with ML" stage. Here, forecasts from each trained model are stored and a 24-h delay is applied before the models' performance is calculated and recorded. This step ensures that there is a standardized time frame for comparing the models' predictive accuracy based on the European Union (EU) regulation. In the final stage, "Hybrid Model Implementation," the performance of each ML model (ANN, SVR, KNN) is analyzed over the last 24 h. Based on this performance analysis, the most appropriate technique is selected. Any errors in prediction are adjusted, leading to the final prediction output. This hybrid approach leverages the strengths of each individual model to produce a more accurate and robust wind energy forecast.

The methodology employed in this study involved utilizing MATLAB for modeling, encompassing the implementation of three ML models as well as a hybrid model. MATLAB was selected due to its suitability for handling complex numerical computations and facilitating model development. The ML models, including KNN, SVR, and ANN, were trained and tested within the MATLAB environment to assess their forecasting capabilities. Concurrently, data collection was facilitated by retrieving information from a wind turbine Application Programming Interface (API), which allowed for the extraction of historical wind energy data. This data was organized into an Excel spreadsheet and subsequently linked to Python for data completeness checks and statistical analyses. Python's capabilities in data manipulation and statistical testing enabled a thorough examination of the dataset, ensuring its integrity and readiness for analysis.

2.1. Data collection

In this study, empirical data on wind energy generation from a 2 MW wind turbine situated in Hungary, provided by E.ON Hungary, are utilized. The dataset encompasses the period from May 1, 2019, to June 13, 2020, offering insights into the dynamics of wind energy production within the Hungarian context. Data were systematically recorded at 15-min intervals.

The singular focus on wind energy generation, measured in megawatt-hours (MWh), serves to streamline the study. This deliberate concentration on a single variable not only facilitates an in-depth examination of the fluctuations and patterns specific to wind energy generation in Hungary but also simplifies the overall process of data collection and reduces the complexity involved in model training. This approach significantly aids in the development of efficient, yet simplified forecasting models tailored to the unique environmental and operational conditions encountered in Hungary.

The data collected about wind energy generation presents intriguing characteristics as can be seen in Table 1. The average (mean) generation stands at 126.07 MWh, indicating the typical output level across the observed period. However, the median value, at 60.1 MWh, reveals that half of the observed values fall below this mark, suggesting a wide dispersion in generation amounts. The mode, or the most frequently observed wind energy generation, is 0.0 MWh, pointing to periods of inactivity or minimal wind conditions. The range of energy generation spans from 0.0 MWh to a maximum of 582.8 MWh, indicating significant variability in generation capacity, which is typical in wind energy due to its dependence on wind speed and environmental factors. The 25th percentile, at 3.8 MWh, and the 75th percentile, at 203.3 MWh, highlighting the skewed nature of wind energy generation data. This skewness is quantified at 1.15, confirming that the distribution of wind energy generation is not symmetrical and leans towards lower values. Nonetheless, the kurtosis value of 0.086 indicates a relatively flat distribution, suggesting that extreme values are not as prevalent as one might expect from a skewed dataset.

2.2. Performance measures

The accuracy of the forecasting models can be evaluated using Mean Absolute Error (MAE), Normalized Mean Absolute Error

Table 1
Statistical summary of wind energy generation.

Statistic	Wind Energy Generation (MWh)
Mean	126.07
Mode	0.0
Min	0.0
Max	582.8
Standard Deviation	149.52
Variance	22356.99
Range	582.8
25th Percentile	3.8
75th Percentile	203.3
Skewness	1.15695
Kurtosis	0.086331

(NMAE), Mean Squared Error (MSE), Coefficient of Determination (COD), and Error (ε). Equations 1-5 in Table 2 summarize the evaluation methods used in this study [34], where n is the number of observations, *Pinst* is the installed wind farm capacity, *xt* and *yt* are the observed (real) and the forecasted output power values at time *t*, respectively, and $\hat{x}$ is the average of the observed values. For higher modeling accuracy, MAE, NMAE, MSE, and ε indices should be closer to zero but the COD value should be closer to 1.

As each of these measures has a different scale, and different high and low, it is hard to compare the model's performance under different scales. Thus, all of the mentioned measures were normalized to have a value between 0 and 1, where 0 represents the lowest errors (best possible performance) and 1 the highest errors (worst possible performance) as shown in Equation (6):

$$\text{M normalized} = \frac{(M - M\,minimum)}{(M\,maximum - M\,minimum)} \qquad \text{Equation 6}$$

Where *M* represents the values of each performance measure. *M minimum* and *M maximum* are the minimum and maximum observed values of the performance measures across all datasets, respectively. For all measures, higher values mean worse performance, but for COD it is the opposite. So in order to make COD in the same scale, where 0 is the best possible performance Equation (7) was applied, where Scaled COD is the adjusted COD with 0 representing the best possible performance and 1 representing the worst possible performance:

$$\text{Scaled COD} = |1 - COD| \qquad \text{Equation 7}$$

To evaluate each ML technique, all the 3 performance measures were used. The 3 measures were used to create a measure score (MS). So, for each method, the average value of the 3 normalized measures was computed and used as an indicator for the method's performance, as can be found in Equation (8). The MS was used by the algorithm to evaluate each method tested.

$$\text{MS} = \frac{(MAE\ normalized + MSE\ normalized + Scaled\ COD)}{3} \qquad \text{Equation 8}$$

In the context of the hybrid model, employing a single performance measure presents significant limitations due to the complexity and multifaceted nature of forecasting accuracy. Specifically, relying solely on MAE might overlook model precision in scenarios of high variance, whereas MSE can disproportionately penalize larger errors, potentially skewing model selection. Similarly, utilizing COD alone may not fully capture the predictive nuances across different scales and distributions of data. In response to these challenges, the MS emerges as a crucial evaluation tool, ingeniously integrating MAE, MSE, and Scaled COD into a unified metric. This amalgamation permits a holistic assessment, balancing considerations of average error magnitude, error distribution, and proportion of variance explained by the model. Such integration is particularly imperative for the hybrid model, where diverse methodologies converge and singular measures may fail to adequately represent performance across varying contexts. Thus, MS serves as an essential mechanism, ensuring that the chosen model not only addresses individual shortcomings of each measure but also aligns with the comprehensive accuracy demands of wind energy forecasting.

To evaluate the long-term forecasting abilities in terms of the total sum of energy, the total amount of energy generated by the wind turbine was calculated as the total sum of wind energy generated each time step as in Equation (9), where *TEG* is total energy generated by wind turbine, and *Energy at each time step* refers to the amount of energy produced by the wind turbine every 15 min.

$$TEG = \sum (Energy\ at\ each\ time\ step) \qquad \text{Equation 9}$$

Similarly, the total amount of energy forecasted by each model (TEG_M) was calculated as the total sum of wind energy forecasted by each model for each time step, as shown in Equation (10), where *Wind energy forecasted by model M at each time step* refers to the predicted amount of energy output from the wind turbine, as estimated by model M, for every 15-min interval:

$$TEG_M = \sum (Wind\ energy\ forecasted\ by\ model\ M\ at\ each\ time\ step) \qquad \text{Equation 10}$$

The Total energy generation error for model *M* is then calculated as in Equation (11):

$$Total\ energy\ generation\ error\ for\ model\ M = \frac{TEG_M - TEG}{TEG} * 100\% \qquad \text{Equation 11}$$

It should be noted that equations 9–11 were utilized in two ways. Firstly, they were used to compute the total energy generation

Table 2
Evaluation criteria of wind energy forecasting.

Performance measure	Expression	Equation number	Purpose
Mean Absolute Error (MAE)	$\frac{1}{n} * \sum_{i=1}^{n} \lvert xt - yt \rvert$	Equation 1	Reflects overall level of errors
Normalized Mean Absolute Error (NMAE)	$\frac{1}{n} * \sum_{i=1}^{n} \left\lvert \frac{xt - yt}{Pinst} \right\rvert$	Equation 2	Dimensionless comparison of overall level for errors
Mean Squared Error (MSE)	$\frac{1}{n} * \sum_{i=1}^{n} (xt - yt)^2$	Equation 3	Reveals contribution of positive and negative errors
Coefficient of Determination (R^2)	$\frac{\sum_{i=1}^{n} (xt - yt)^2}{\sum_{i=1}^{n} (xt - \hat{x})^2}$	Equation 4	Describe how well a model fits the data
Error (ε)	$xt - yt$	Equation 5	Describes the difference between forecasted and actual value

error by summing the energy forecasted by each model across all timesteps during the testing period. Secondly, the equations were used to compute the daily error for each model by summing the forecasted energy across daily timesteps during the testing period.

2.3. Machine learning models

Although ML wind energy forecasting models demonstrate strong capabilities and perform well, the limitations of AI and machine learning forecasting models are highlighted by the following factors, as explained by Makridakis et al. [35].

- their conclusions are based on a few, or even a single time series, raising questions about the statistical significance of the results and their generalization;
- the methods are evaluated for short-term forecasting horizons, often one step ahead, not considering medium and long-term ones;
- no benchmarks are used to compare the accuracy of ML methods versus alternative ones.

To circumvent prevalent limitations, this study employed full-year time-series data for day-ahead wind power forecasting (96 steps ahead). Three distinct machine learning models—ANN, SVM, and KNN—were developed and subsequently integrated into a hybrid model. The hybrid model's performance was rigorously benchmarked and comparatively analyzed against the individual ANN, SVM, and KNN models across both short and long-term forecasting horizons.

2.3.1. ANN forecasting model

ANN is a network of "neurons" that are arranged in a layered structure. Fig. 2 shows a simple diagram of ANN where input variables arrive from the bottom, while the forecasted variable(s) (output) appears at the top layer. An ANN also includes one or more hidden layers and hidden neurons.

Most neural network approaches to the problem of forecasting use a Multi-Layer Feed-Forward Neural Network (MLFFNN). In MLFFNN, each layer receives its inputs from the previous layer, so the outputs of a certain layer are the inputs for the next one. The inputs of the neurons in each layer are combined using a weighted linear combination, this weighted sum (ws) can be seen in Equation (12), where $w1, w2, \ldots, wn$ are the weights of the input data $v1, v2, \ldots, vn$ for the neuron, and b is neuron's bias. Fig. 3 shows the flow of information in an artificial neuron.

$$ws = w_1 \times v_1 + w_2 \times v_2 + \ldots + w_n \times v_n + b \qquad \text{Equation 12}$$

Then a nonlinear function ϕ modifies the neuron's weighted input sum (ws) before producing the output X, as can be seen in Equation (13).

$$X = \phi(ws) \qquad \text{Equation 13}$$

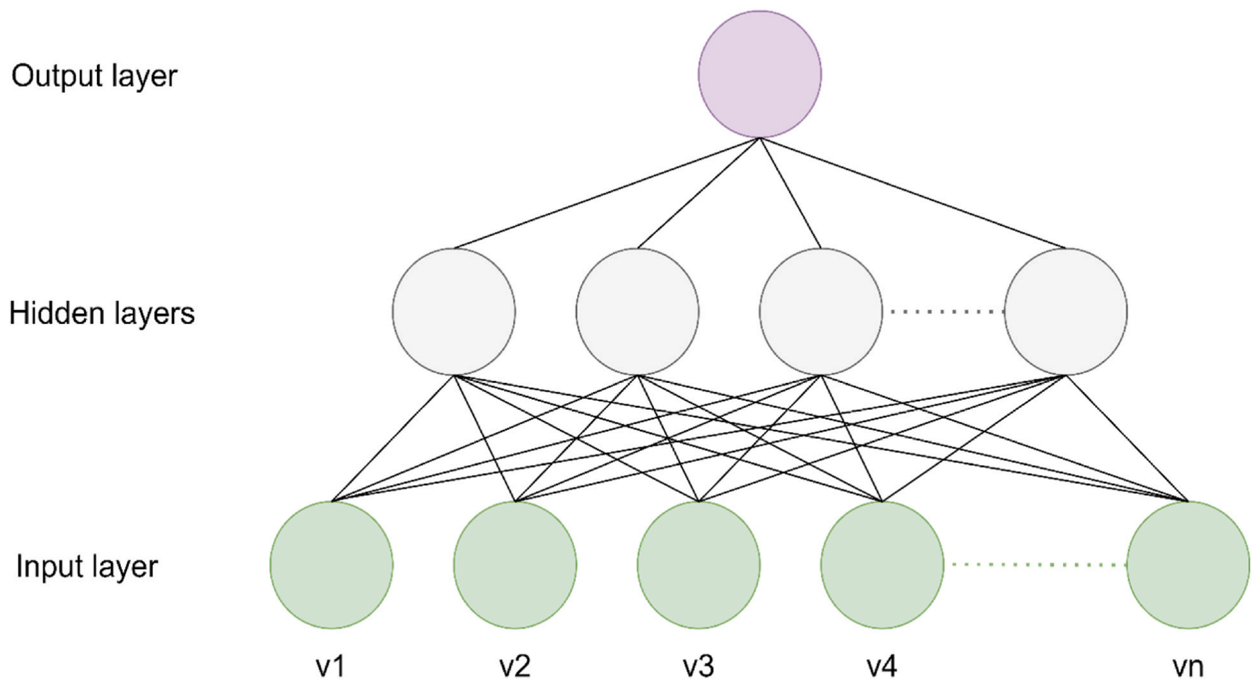

Fig. 2. A neural network with n inputs and one output.

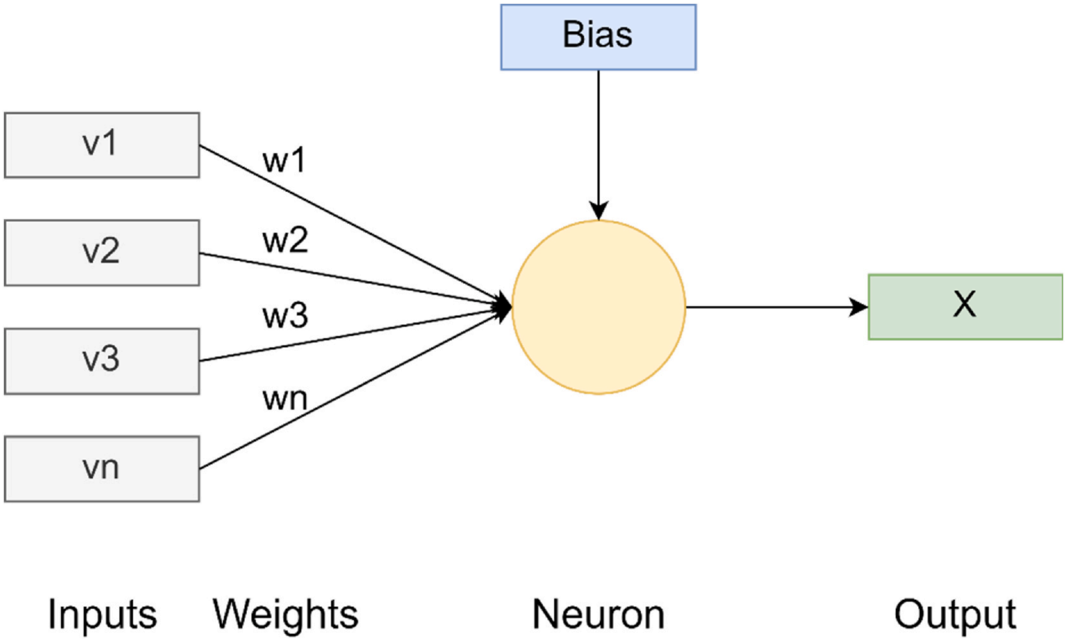

Fig. 3. Flow of information in an artificial neuron.

There are different transfer functions that might be used, yet, the sigmoid transfer function is among the most used [36]. Equation (14) shows the sigmoid transfer function, while Equation (15) shows the Tansig transfer function which was used instead of the sigmoid in case of negative values are found.

$$\phi(ws) = \frac{1}{1 + e^{-ws}} \quad \text{Equation 14}$$

$$\phi(ws) = \frac{1 - e^{-2ws}}{1 + e^{-2ws}} \quad \text{Equation 15}$$

The training process of a neural network involves passing the training data forward through the network, and then adjusting the weights of each neuron based on the error. This process is known as an epoch, and it is repeated until the weights are optimized or a specific number of epochs is reached. This research utilizes a fully connected Multilayer Feedforward Neural Network (MLFFNN) with one hidden layer. The number of hidden neurons is also an important parameter in the ANN as it affects the network's ability to generate a function that accurately represents the underlying problem. To prevent over-fitting of the training set, the number of hidden neurons was set to 33 % (one-third) of the number of inputs [37]. Table 3 shows the ANN parameters used in this study.

2.3.2. SVM forecasting model

SVM regression is a nonparametric supervised learning technique as it relies on kernel functions. Unlike other methods and techniques where regression models try to minimize the errors (squared error normally) between the real and predicted values, Support Vector Regression (SVR) tries to minimize the coefficients (W) as can be seen in Equation (16) [38] – or, more precisely, tries to minimize the norm of the coefficient vector. Thus, in SVR the error is represented by constraints where the absolute error is set less than or equal to a specified margin called the maximum error (ϵ), as can be seen in Equation (17) [38]. In other words, SVR tries to find the best-fit line, such that this fit is the hyperplane that has the maximum number of points.

Table 3
ANN parameters.

Parameter	Description	Value
Number of inputs	Number of input data variables	96
Number of outputs	Number of output forecasted variables	1
Number of hidden neurons	Number of hidden neurons hyperparameter	32
Maximum epochs	Maximum number of training iterations before training is stopped	1000
Maximum training time	Maximum time in seconds before training is stopped	Unlimited
Performance goal	The minimum target value of MSE	0

$$Minimize : \frac{1}{2}\|w\| \quad \text{Equation 16}$$

$$Constraints : |y_i - w_i x_i| \le \epsilon \quad \text{Equation 17}$$

Sometimes, some values might fall outside the specified margin (ϵ), these values are called slack variables and are donated by ξ. Slack variables have the potential to exist, so it has to be minimized. Thus, equations (16) and (17) become as in equations (18) and (19) [38]. Note that C in Equation (18) is an additional hyperparameter, as C increases, the tolerance for points outside of ϵ increases. As C moves towards 0, the tolerance approaches 0 and the equation collapses into the simplified (infeasible) one.

$$Minimize : \frac{1}{2}\|w\| + C\sum_{i=1}^{n}|\xi_i| \quad \text{Equation 18}$$

$$Constraints : |y_i - w_i x_i| \le \epsilon + |\xi_i| \quad \text{Equation 19}$$

In non-linear SVR, kernel functions (K) transform the data into a higher dimensional feature space to make it possible to perform the linear separation as shown in Equation (20), where b is the bias and a and a^* are the Lagrange multipliers (or dual variables) and are nonnegative real numbers [38].

$$y = \sum_{i=1}^{n}(a_i - a_i^*).K\langle x_i, x\rangle + b \quad \text{Equation 20}$$

In this work, the Gaussian radial basis kernel function was used as can be seen in Equation (21), where $\|X - X'\|^2$ is the squared Euclidean distance between the two feature vectors X and X'. σ is a free hyperparameter.

$$K(X, X') = \exp\left(-\frac{\|X - X'\|^2}{2\sigma^2}\right) \quad \text{Equation 21}$$

In this work, the SVR model was built as described in equations (16)–(21), and the maximum iteration limit was set to 1000.

2.3.3. KNN forecasting model

K-nearest Neighbors Regression or KNN is also a non-parametric method for prediction. KNN uses feature similarity to predict the values of any new data points. This means that the new point is assigned based on how closely it resembles the points in the training set. The main steps of forecasting wind power using KNN technique can be boiled down to the following three steps.

1. Calculating the distance between the new point and each training point.
2. Based on the closest distance (which is calculated in step 1), the closest k data points are selected.
3. The average of the k data points is the prediction for the new point.

Many methods can be used to calculate the distance between new and training points, yet, Euclidean distance, Manhattan distance, and Mahalanobis distance are the most commonly used [23]. In this work, the Euclidean distance was utilized as can be seen in Equation (22), where k_dimensions represents the number of dimensions or features in the input data points.

$$Euclidean\ distance = \sqrt{\sum_{i=1}^{k\_dimensions}(x_i - y_i)^2} \quad \text{Equation 22}$$

In this work, the KNN regression model described above, the model considers the 4 nearest neighbors and uses the Euclidean distance approach to predict the output value.

2.3.4. Hybrid forecasting model

In this research, the proposed hybrid model aims to leverage the benefits of three different prediction techniques, namely ANN, SVR, and KNN. The hybrid model is designed to adaptively select the most appropriate forecasting technique (i.e., ANN, SVR, or KNN) based on external variables, such as weather conditions. Further, the hybrid model has an error correction function to improve the forecasting accuracy as will be seen.

The implementation and validation of the proposed hybrid model are systematically approached through three foundational tasks, reflecting the aim to harness the predictive strengths of ANN, SVR, and KNN while addressing their individual limitations. Initially, the performance of each method is quantified using the Measure Score (MS) during the performance evaluation phase, with assessments being conducted retrospectively over a 24-h timeframe. The model that demonstrates the lowest MS value, indicative of superior predictive accuracy, is then selected for subsequent real-time forecasting applications.

Following the identification of the most accurate algorithm, the selection phase for model forecasting ensues. In this phase, the chosen algorithm is deployed for forecasting wind energy generation for the forthcoming 24-h period. This strategic selection is based on the historical performance of the models under different environmental conditions, ensuring that the most reliable algorithm is utilized under the current weather scenarios.

The final phase (error correction and refinement) plays a crucial role in the validation process. Here, adjustments are made to the forecast outputs generated by the selected model, aiming to reduce discrepancies and enhance the fidelity of the predictions. This systematic application of the error correction mechanism is designed to refine the accuracy and reliability of the forecasts, providing a robust and reliable tool for wind energy forecasting. Through this comprehensive approach, the hybrid model's validation is conducted in a rigorous manner, promising significant improvements in the accuracy and reliability of wind energy predictions.

The workflow of the proposed hybrid model is illustrated in Fig. 1. It starts by retrieving the historical data. The data is then used to train and evaluate the three models, and the most accurate model is selected based on its past performance. The selected model is then used to perform the forecast.

As the model is chosen to do a future forecast based on its past performance, some errors are expected. However, these potential errors can be reduced. Thus, for the selected model (the model which is selected to perform the next 24 h forecast) the average past errors for each time step are calculated according to Equation (23). The term $(t - 24 * n)$ represents the time steps $24, 48, 72, \ldots\ldots n$ hours before the prediction starts.

$$\textit{Average error for each time step } (ae_t) = \frac{1}{n} * \sum_{i=1}^{n} \left(\mathrm{x}_{(t-24*n)} - \mathrm{y}_{(t-24*n)} \right) \qquad \text{Equation 23}$$

Then the final prediction will be the sum of the original prediction done by the selected model (y_t) and the ae_t of the model as can be seen in Equation (24).

$$\text{Final prediction}_t = \mathrm{y}_t + \mathrm{ae}_t \qquad \text{Equation 24}$$

The dynamic nature of the proposed hybrid model stands out through its capacity to integrate and adaptively leverage the unique strengths of three distinct forecasting techniques: ANN, SVR, and KNN. This adaptability is rooted in the model's innovative design, which selects the optimal forecasting technique based on real-time environmental variables and past performance metrics. This selection mechanism ensures that the most accurate and situationally appropriate model is applied to each new forecasting scenario. Furthermore, the integration of an error correction function significantly enhances the model's predictive accuracy by systematically addressing and reducing potential forecast discrepancies. The dynamic adaptability, coupled with the continuous refinement process, points out the hybrid model's capability to provide more reliable and precise wind energy forecasts domain.

3. Results

After building the models and the evaluation methods based on the mathematical background provided in the previous sections, the models were utilized to perform a 24-h wind energy forecasting with 15-min resolution. All models were trained and tested for 406 days, starting on the 5th of May 2019, till the 13th of June 2020. The training period lasts for 389 days (5th of May 2019, to 28th of May 2020), while the testing period lasts for 16 days (29th to 13th of June 2020). It is important to note that the forecasting resolution and horizon were set based on regulatory requirements for grid operators.

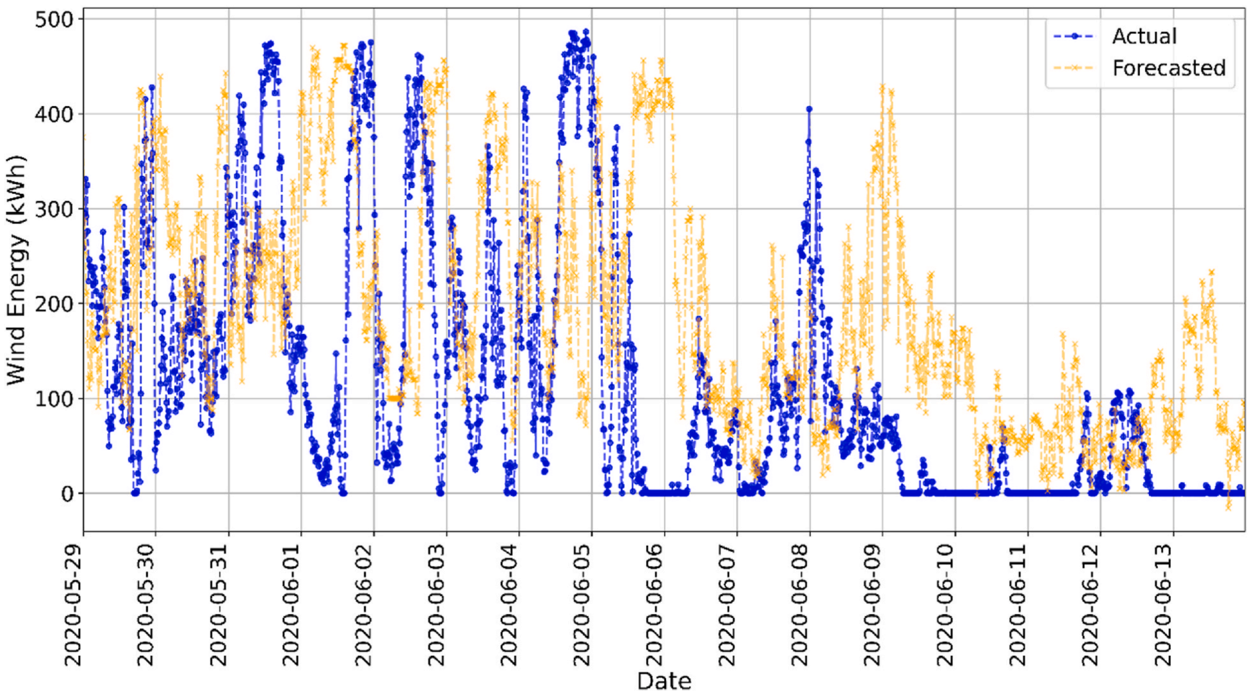

Fig. 4. Actual Vs. forecasted wind energy for the ANN model throughout the test period.

3.1. ANN forecasting model performance

Over the training period, ANN showed fair performance in predicting 24 h ahead wind energy with an MAE of 136.84 and MSE of 34669.39. After training, the testing period has taken place. As expected, ANN shows a also fair performance with 136.60 MAE, 30691.13 MSE, and 0.066 COD. Fig. 4 below shows the observed (actual) wind energy vs. the ANN forecasted output energy for the testing period.

The evidence of a memory effect can also be observed in Fig. 4. This phenomenon occurs in time series data when the model is influenced by past inputs, causing it to have a "memory" of previous events. While the memory effect can be beneficial in allowing the model to learn patterns and trends, it can also lead to errors if the model is unable to adapt to changes in the data. To address the memory effect, the hybrid model was designed to switch to the model with the best accuracy when one of the models performs better than the others with error correction ability. Understanding the impact of the memory effect is crucial for improving the accuracy and reliability of ANN forecasting models.

Although ANN shows fair abilities in 24-h energy forecasting, this model shows bad abilities in long-term forecasting in terms of the total sum of generated energy. ANN results analysis shows that the ANN model was able to predict the total amount of energy generated in the test period (16 days) with a 65.6 % total energy generation error. Yet, the average error of predicting the daily sum of energy was worse at 205 %.

3.2. SVR forecasting model performance

Over the training period, SVR showed fair performance in predicting 24 h ahead wind energy with an MAE of 151.82 and MSE of 35845.84. After training, the testing period has taken place. SVR also shows fair performance but a bit worse than the ANN with 184.42 MAE, 50951.77 MSE, and 0.037 COD. Fig. 5 shows the observed (actual) wind energy vs. the SVM forecasted output energy for the testing period. While the forecasted values attempt to mirror the general trends observed in actual wind energy, there is a significant disparity, especially during periods of low energy production. This discrepancy suggests that the forecasting SVR model may struggle with accurately predicting lower energy outputs.

Similar to ANN, SVR shows bad abilities in long-term forecasting in terms of the total sum of generated energy. SVR results analysis shows that the model was able to predict the total amount of energy generated in the test period (16 days) with a 113.5 % total energy generation error. Yet, the average error of predicting the daily sum of energy was worse at 410 %.

3.3. KNN forecasting model performance

KNN ML technique was utilized to build the KNN forecasting model as discussed in section 2.3.3. Over the training period, KNN showed fair performance in predicting 24 h ahead wind energy with an MAE of 135.09 and MSE of 35736.12. After training, the testing

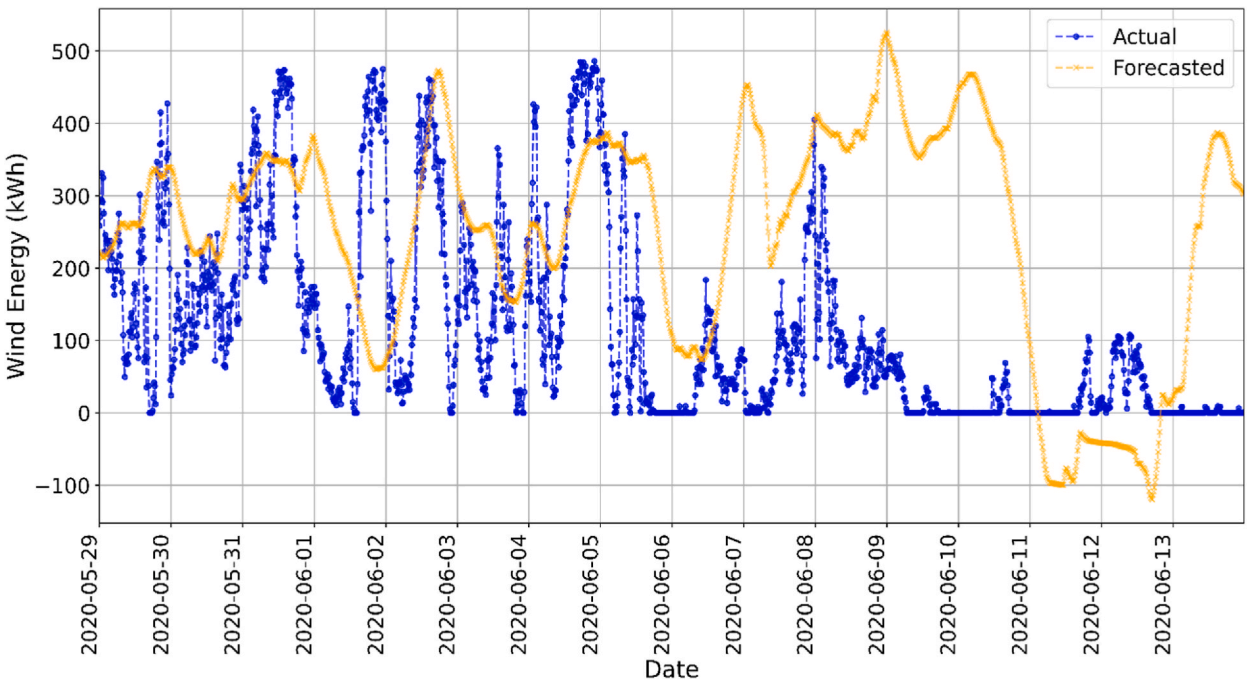

Fig. 5. Actual Vs. Forecasted wind energy for the SVR model throughout the test period.

period has taken place. KNN shows a also fair performance, better than SVR and comparable to ANN with 113.10 MAE, 25711.93 MSE, and 0.045 COD. Fig. 6 shows observed (actual) wind energy vs. the KNN forecasted output power for the testing period. Similar to the ANN time-series forecasting model, the memory effect can be observed. This effect can be seen in some forecasting models, including the KNN model, which may exhibit a noticeable pattern in its predicted values that follows the pattern of the observed ones. This is because KNN, like other non-parametric methods, uses historical observations as the basis for its predictions. Therefore, if the underlying process being modeled has a significant memory effect that may be captured in the KNN predictions.

Although KNN shows fair abilities in 24-h energy forecasting, this model shows excellent abilities in long-term forecasting in terms of the total sum of generated energy. KNN results analysis shows that the model was able to predict the total amount of energy generated in the test period (16 days) with only a 1 % total energy generation error. Yet, the average error of predicting the daily sum of energy was worse at 41 %.

3.4. Hybrid forecasting model performance

After training ANN, SVR, and KNN models, the hybrid model was then tested. The hybrid model shows a good performance in predicting 24 h ahead of wind energy. Actually, the hybrid model performed the best compared to ANN, SVR, and KNN with 110.86 MAE, 23959.95 MSE, and 0.196 COD. Fig. 7 below shows the observed (actual) wind energy vs. the ANN forecasted output energy for the testing period. Incorporating multiple models in the hybrid forecasting approach has reduced the memory effect, as observed in ANN and KNN models. However, some degree of memory effect still exists in the hybrid model, albeit at a lower level.

The hybrid model also shows good abilities in long-term forecasting in terms of the total sum of generated energy, better than ANN and SVM for example (more details in the following section), yet worse than KNN. Hybrid model results analysis shows that this model was able to predict the total amount of energy generated in the train and test period with a 39.0 % total energy generation error. Yet, the average error of predicting the daily sum of energy was worse at 47.6 %.

One noteworthy aspect of the proposed hybrid forecasting model is its ability to select the most accurate model for the next 24-h prediction. As explained in section 2.3.4, the model analyzes the past 24-h performance of the three prediction techniques and selects the model that provides the best accuracy for the upcoming 24 h. During the analysis of the model, it was observed that out of the 16 days of testing, the SVR model, which had the worst overall performance among the three models, was selected by the algorithm for the next 24 h forecast 10 times as can be seen in Fig. 8. This finding is interesting because it suggests that the KNN model, despite its poorer overall performance, has certain advantages over the other models for some specific time periods. In contrast, the KNN and ANN models were each chosen three times, indicating that they were not the preferred choices for most of the testing period. These results demonstrate that the hybrid model was successful in capturing and integrating the unique forecasting abilities of each model and using them to improve the overall forecasting accuracy. The hybrid model's ability to dynamically choose the best forecasting technique based on past performance and external variables is a significant advantage over traditional forecasting methods.

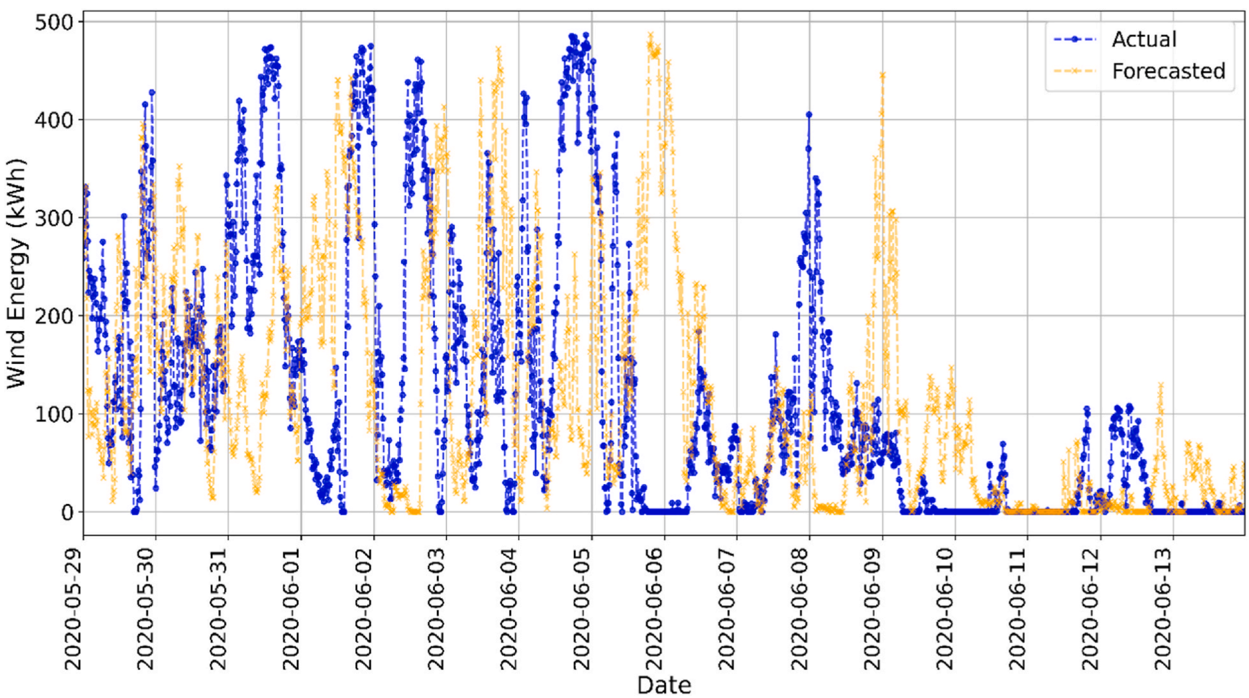

Fig. 6. Actual vs. Forecasted wind energy for the KNN model throughout the test period.

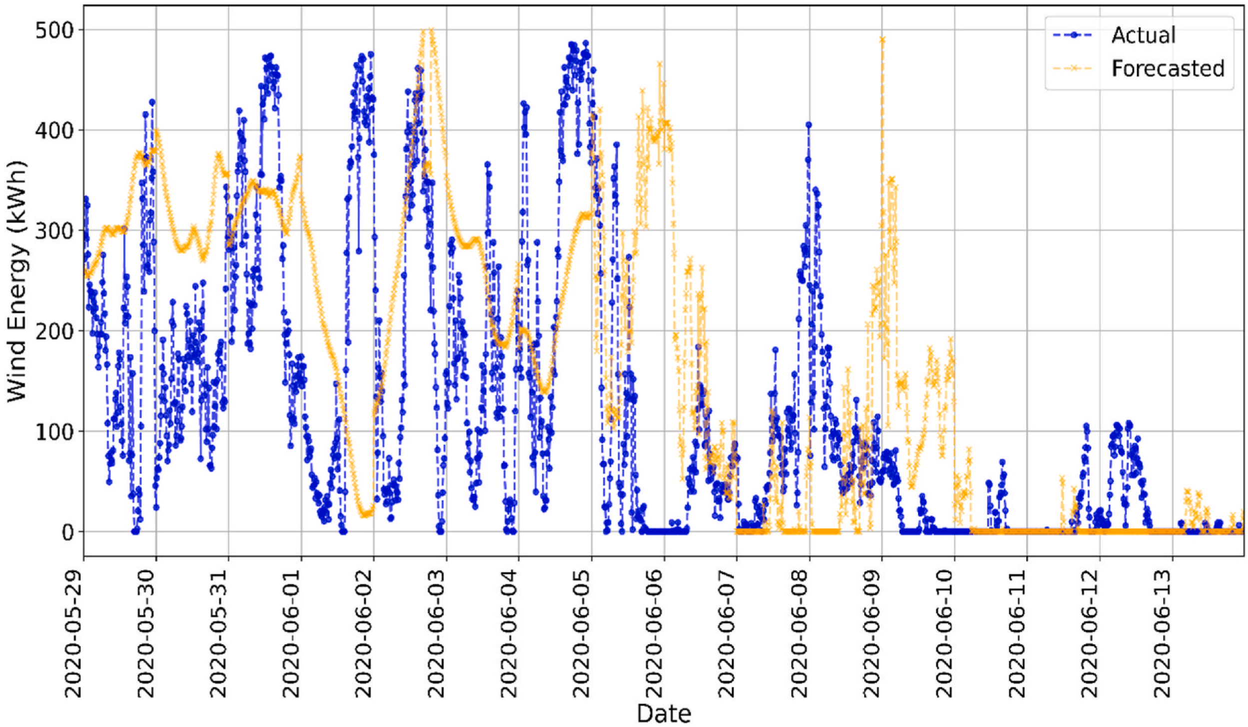

Fig. 7. Actual Vs. Forecasted wind energy for the hybrid model throughout the test period.

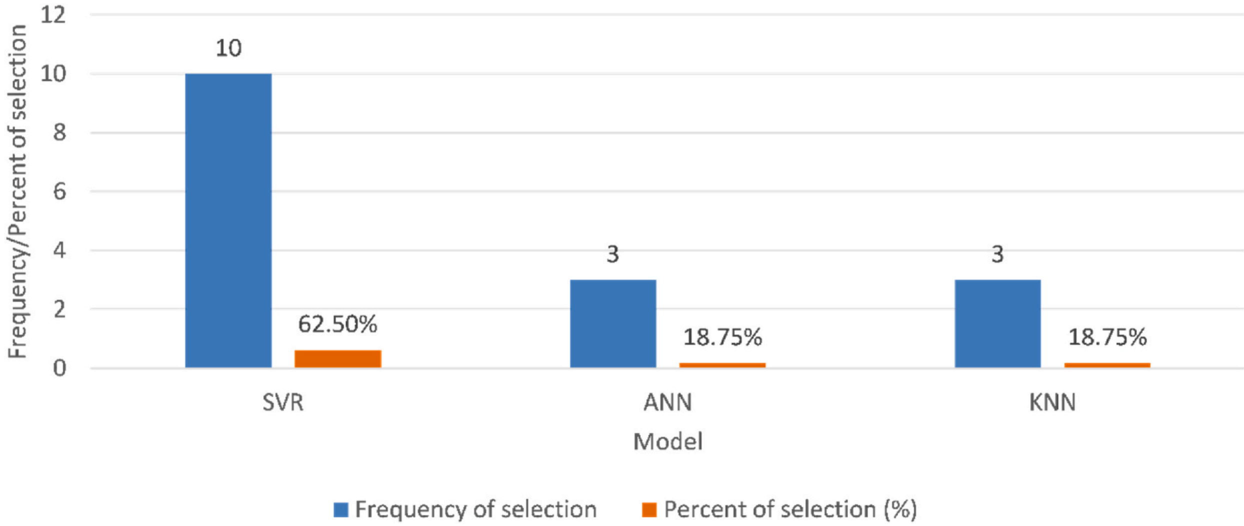

Fig. 8. The frequency and percent of techniques selected by the algorithms to perform the next 24 h' forecast.

Table 4
Performance measures comparison during the test period – short-term.

Model	Performance measure		
	NMAE	MSE	COD
ANN	6.82 %	30691.13	0.066
SVR	9.22 %	50951.77	0.037
KNN	5.65 %	25711.93	0.045
Hybrid	5.54 %	23959.95	0.196

4. Discussion

4.1. Performance comparison for short-term forecasting

The short-term wind energy forecasting abilities for day-ahead in 15-min resolution are analyzed and benchmarked. Table 4 shows the performance of the different models during the test period.

The results in Table 4 indicate that the hybrid model outperformed other models, achieving the lowest NMAE (5.54 %) and MSE (23959.95) values. The hybrid model also demonstrated a higher COD (0.196) compared to the other models. In comparison, the KNN model performed fair in terms of NMAE (5.65 %), COD (0.045), and MSE (25711.93). On the other hand, compared to the hybrid and KNN models, the ANN model had higher NMAE (6.82 %) and MSE (30691.13) values, while COD (0.066) is higher than KNN but lower than the hybrid model. The SVR model demonstrated poor performance, with the highest NMAE (9.22 %), MSE (50951.77), and the lowest COD (0.037) values.

In their review of wind forecasting, González et al. [39] noted that for a 24-h forecasting scenario, the NMAE ranges from 5.95 % for the Variational Mode Decomposition - Feed-Forward Neural Network Multi-Input Multi-Output (VMD-FFNN MIMO) model to 63.79 % for the FFNN Recursive model. This context underscores the robust performance of the hybrid model proposed in this study, which achieves a significantly lower NMAE of 5.54 %. Yousuf et al. [40] reported that the mean of wind energy forecasting NMAE spans from 6.73 % to 10.07 %. Our results fall well within these established benchmarks, affirming their validity. Notably, our hybrid model outperforms these standards, offering a valuable edge in forecasting accuracy, as evidenced by its lower NMAE value.

The performance comparison of the four models was also evaluated using the MS approach, which allows for the comparison of models based on multiple performance measures. The MS approach calculates the score of each model by assigning a value of 1 to the worst-performing model, and a value of 0 to the best-performing model as can be seen in Fig. 9. This approach is particularly useful when comparing models with different performance measures, as it allows for a straightforward ranking of the models based on their overall forecasting performance. The results of the MS analysis demonstrated that the SVR model had the worst performance among the four models, with a score of 1. On the other hand, the hybrid model was the best-performing model, with a score of 0. The MS approach allows for the comparison of models even when their performance measures have different scales. For instance, in the case of comparing ANN and KNN, KNN had better MAE and MSE values, but worse COD compared to ANN. The MS approach was able to rank KNN as a better model than ANN based on overall performance.

4.2. Performance comparison for long-term forecasting

Fig. 10 illustrates the daily sum of wind energy predictions from various models against the actual energy generation over the testing period. Here, the KNN and hybrid models stand out for their efficacy in long-term energy prediction, despite their daily performance disparities. Also, the hybrid model demonstrates a commendable balance in its predictions, closely mirroring actual energy outputs more consistently than SVR and ANN models, especially during fluctuating energy generation periods. This subtly underscores the hybrid model's potential as a versatile forecasting tool.

Fig. 11 presents a subtle comparison of long-term wind energy forecasting errors by each model. Remarkably, the KNN model shows the least total energy generation error at a mere 1 %, indicating its strength in long-term prediction. However, the narrative shifts when considering the hybrid model, which, with a total energy generation error of 39.0 %, substantially outperforms the ANN and SVR models. Moreover, in terms of the average daily sum of energy error, the KNN model maintains the lowest error rate at 41 %. However, the Hybrid model emerges as a close second with an error rate of 47.60 %. This is a significant achievement, especially when contrasted with the significantly higher errors observed for the ANN and SVR models, at 205 % and 410 % respectively.

These analyses suggest that while the KNN model excels in the domain of long-term forecasting, the Hybrid model represents a robust and flexible alternative, offering a balanced approach to both daily and long-term energy prediction challenges. The evidence suggests the Hybrid model as a commendable option for those seeking a middle ground between the extremes of daily accuracy and long-term foresight in wind energy forecasting. This balance is particularly relevant against the backdrop of the inherent complexities and uncertainties highlighted in the existing literature on long-term wind energy forecasting. Comparative analysis within the literature on long-term wind energy forecasting presents a challenge due to the scarcity of studies focusing on extended periods [41]. Even when such studies are available, aligning research with equivalent forecast horizons and resolutions remains a rarity. Nonetheless, the literature that does exist suggests a broad variance in forecasting errors for long-term forecasting. Reported values for error can start as low as 30 % and escalate substantially [40,41], reflecting the inherent complexities and uncertainties associated with forecasting wind energy over longer durations.

All in all, even though KNN shows very good long-term forecasting abilities, its short-term performance was fair. The hybrid model inherited good prediction abilities from each model, thus the suggested hybrid model shows very good abilities in both long and short-term prediction. As a result, it can be utilized in long and short-term wind energy forecasting while maintaining good forecasting accuracy.

One limitation of the study lies in the duration of the data collection and testing period, which spanned 406 days between 2019 and 2020. While this timeframe provided valuable insights into the performance of ANN, SVR, KNN, and the hybrid model in both short-term and long-term wind energy forecasting, it may not be sufficiently extensive to capture the full range of seasonal and annual variations that could affect the models' accuracy. Longer training and testing periods could potentially offer a more comprehensive understanding of the models' robustness and adaptability to fluctuating weather patterns and energy demands. Therefore, future studies could benefit from extending the data collection and testing period to better assess the long-term reliability and generalizability

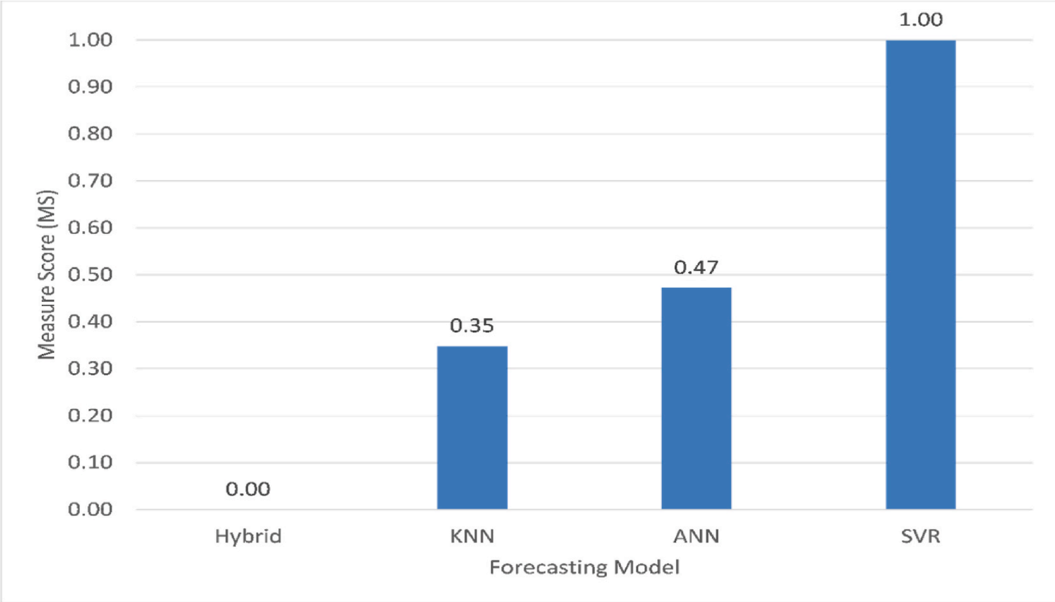

Fig. 9. MS for each forecasting model during the test period.

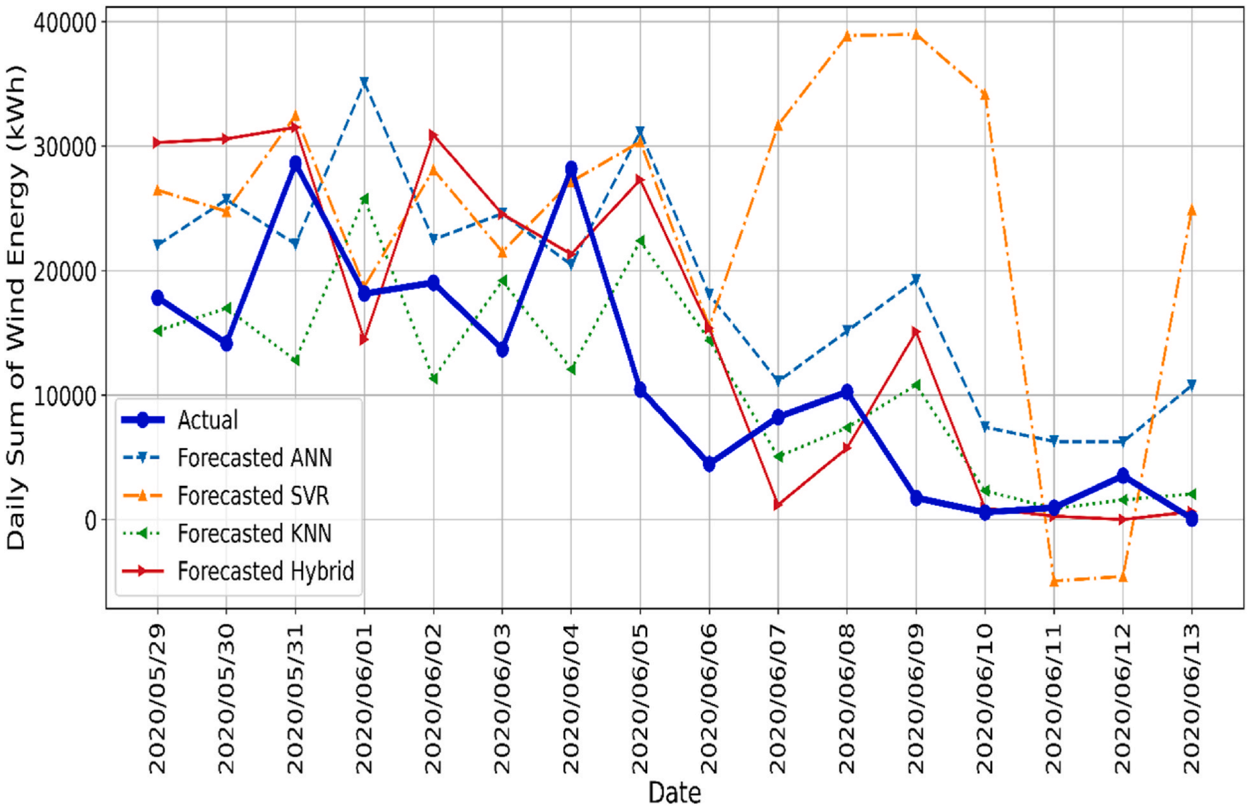

Fig. 10. Long-term forecasting during the test period for each model.

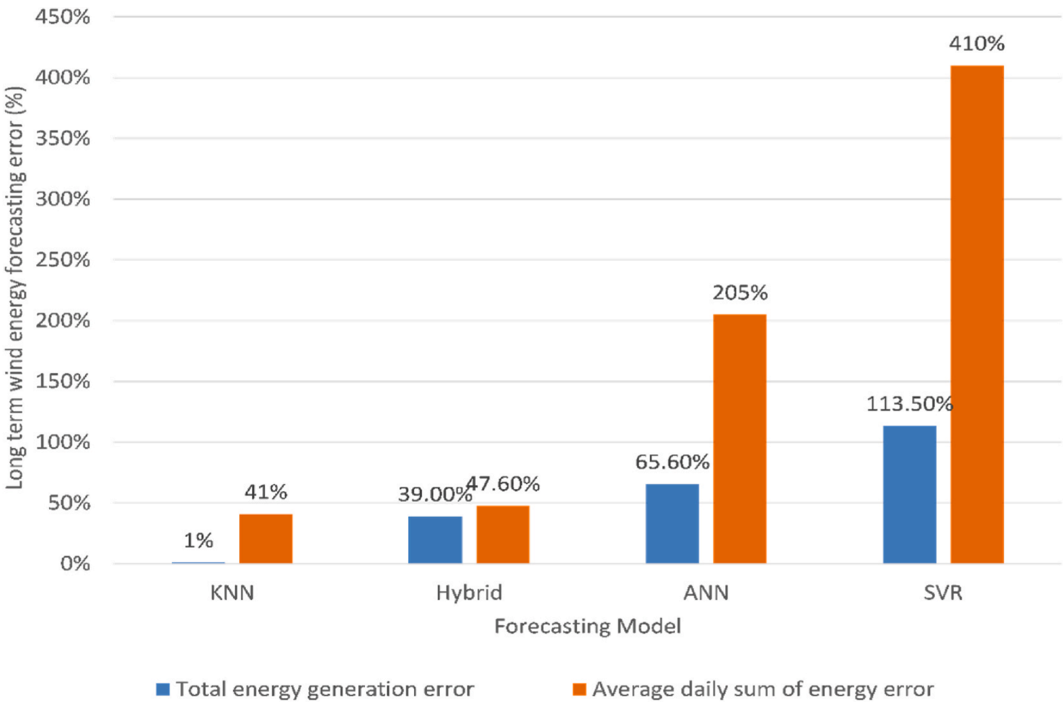

Fig. 11. Long-term forecasting error for each model.

of the dynamic forecasting models developed.

5. Conclusion

In this paper, two focal points were emphasized: the formulation of a dynamic forecasting model and its aptitude in wind energy forecasting for both short-term and long-term scenarios. Data from a 2 MW wind turbine in Hungary, collected over a period spanning 406 days between 2019 and 2020, were used to develop four distinct machine learning models: ANN, SVR, KNN, and a hybrid model that amalgamates the three techniques. Evaluation criteria were devised to gauge the models' forecasting capabilities over varying time frames. In short-term forecasting, both KNN and ANN performed fairly well, achieving NMAE values of 5.65 % and 6.82 % respectively, while SVR reported a higher NMAE of 9.22 %. Notably, the hybrid model surpassed these individual methods with a superior NMAE of 5.54 %, also outdoing the average NMAE range of 6.73 %–10.07 % typically found in the literature. In the context of long-term forecasting, KNN stood out with a lower error rate of 41 %, compared to the less effective ANN and SVR methods. However, the hybrid model demonstrated a commendable balance in performance with a 47.6 % error rate, indicating its potential in both short- and long-term forecasting scenarios. This versatility makes the hybrid model a valuable tool for wind energy forecasting, adeptly balancing short-term precision with long-term dependability. Consequently, the hybrid model is recommended for further investigation and application in the renewable energy sector and other predictive analytics domains.

Future work could focus on exploring the impact of parameter variations on the predictive accuracy of the hybrid model. Moreover, the universality and scalability of the model can be assessed through its application across various geographical locations and differing wind farm capacities. Additionally, future investigations might delve into integrating more predictive algorithms and incorporating a wider array of environmental variables to enhance the robustness and reliability of wind energy forecasting.

Funding

This study has no funding to declare. Open Access funding provided by Hungarian Electronic Information Services National Programme (EISZ)–Corvinus University of Budapest.

Data availability statement

The authors do not have permission to share data.

Ethics

Not applicable.

CRediT authorship contribution statement

Mutaz AlShafeey: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Csaba Csaki:** Writing – review & editing, Writing – original draft, Validation, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Bilgili, A. Yasar, E.J.R. Simsek, S.E. Reviews, Offshore Wind Power Development in Europe and its Comparison with Onshore Counterpart, vol. 15, 2011, pp. 905–915.
[2] M.Q. Raza, A. Khosravi, A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings, Renew. Sustain. Energy Rev. 50 (2015) 1352–1372.
[3] R.A. de Marcos, A. Bello, J. Reneses, Electricity price forecasting in the short term hybridising fundamental and econometric modelling, Elec. Power Syst. Res. 167 (2019) 240–251.
[4] A. Mellit, A. Massi Pavan, E. Ogliari, S. Leva, V. Lughi, Advanced methods for photovoltaic output power forecasting: a review, Appl. Sci. 10 (2020) 487.
[5] S. Sobri, S. Koohi-Kamali, N.A. Rahim, Solar photovoltaic generation forecasting methods: a review, Energy Convers. Manag. 156 (2018) 459–497.
[6] A.M. Foley, P.G. Leahy, A. Marvuglia, E.J. McKeogh, Current methods and advances in forecasting of wind power generation, Renew. Energy 37 (2012) 1–8.
[7] E. Tan, S.S. Mentes, E. Unal, Y. Unal, B. Efe, B. Barutcu, et al., Short term wind energy resource prediction using WRF model for a location in western part of Turkey, J. Renew. Sustain. Energy 13 (2021).
[8] S. Sengar, X. Liu, Ensemble approach for short term load forecasting in wind energy system using hybrid algorithm, J. Ambient Intell. Hum. Comput. 11 (2020) 5297–5314.
[9] F.A. Wolak, Long-term resource adequacy in wholesale electricity markets with significant intermittent renewables, Environmental and Energy Policy and the Economy 3 (2022) 155–220.
[10] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, H. Zareipour, Energy forecasting: a review and outlook, IEEE Open Access Journal of Power and Energy 7 (2020) 376–388.
[11] J. Aghaei, M.-I. Alizadeh, Demand response in smart electricity grids equipped with renewable energy sources: a review, Renew. Sustain. Energy Rev. 18 (2013) 64–72.
[12] J.-F. Toubeau, J. Bottieau, Y. Wang, F. Vallée, Interpretable probabilistic forecasting of imbalances in renewable-dominated electricity systems, IEEE Trans. Sustain. Energy 13 (2021) 1267–1277.
[13] X. Zhang, Y. Chen, Y. Wang, R. Ding, Y. Zheng, Y. Wang, et al., Reactive voltage partitioning method for the power grid with comprehensive consideration of wind power fluctuation and uncertainty, IEEE Access 8 (2020) 124514–124525.
[14] W.-Y. Chang, A literature review of wind forecasting methods, J. Power Energy Eng. 2 (2014) 161.
[15] H. Liu, C.J.A.E. Chen, Data processing strategies in wind energy forecasting models and applications, A comprehensive review 249 (2019) 392–408.
[16] A. Mosavi, M. Salimi, S. Faizollahzadeh Ardabili, T. Rabczuk, S. Shamshirband, A.R. Varkonyi-Koczy, State of the art of machine learning models in energy systems, a systematic review, Energies 12 (2019) 1301.
[17] M. AlShafeey, C. Csaki, Tech mining analysis: renewable energy forecasting using artificial intelligence technologies. 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON), IEEE, 2022, pp. 1–5.
[18] G. Sideratos, N.D. Hatziargyriou, An advanced statistical method for wind power forecasting, IEEE Trans. Power Syst. 22 (2007) 258–265.
[19] R. Rahmani, R. Yusof, M. Seyedmahmoudian, S. Mekhilef, Hybrid technique of ant colony and particle swarm optimization for short term wind energy forecasting, J. Wind Eng. Ind. Aerod. 123 (2013) 163–170.
[20] A. Zameer, A. Khan, S.G. Javed, Machine Learning based short term wind power prediction using a hybrid learning model, Comput. Electr. Eng. 45 (2015) 122–133.
[21] Z. Chi, W. Haikun, Z. Tingting, Z. Kanjian, L. Tianhong, Comparison of two multi-step ahead forecasting mechanisms for wind speed based on machine learning models. 2015 34th Chinese Control Conference (CCC), IEEE, 2015, pp. 8183–8187.
[22] J. Wang, S. Qin, Q. Zhou, H. Jiang, Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China, Renew. Energy 76 (2015) 91–101.
[23] Y. Zhang, J. Wang, K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting, Int. J. Forecast. 32 (2016) 1074–1080.
[24] N. Sharma, R. Bhakar, P. Jain, Optimal reconciliation of hierarchical wind energy forecasts utilizing temporal correlation, Energy Convers. Manag. 301 (2024) 118053.
[25] S.R. Moreno, L.O. Seman, S.F. Stefenon, L. dos Santos Coelho, V.C. Mariani, Enhancing wind speed forecasting through synergy of machine learning, singular spectral analysis, and variational mode decomposition, Energy 292 (2024) 130493.
[26] M.Y. Junior, R.Z. Freire, L.O. Seman, S.F. Stefenon, V.C. Mariani, L. dos Santos Coelho, Optimized hybrid ensemble learning approaches applied to very short-term load forecasting, Int. J. Electr. Power Energy Syst. 155 (2024) 109579.
[27] H. Chen, A novel wind model downscaling with statistical regression and forecast for the cleaner energy, J. Clean. Prod. 434 (2024) 140217.
[28] R.G. da Silva, S.R. Moreno, M.H.D.M. Ribeiro, J.H.K. Larcher, V.C. Mariani, L. dos Santos Coelho, Multi-step short-term wind speed forecasting based on multi-stage decomposition coupled with stacking-ensemble learning approach, Int. J. Electr. Power Energy Syst. 143 (2022) 108504.
[29] Y. Wang, T. Chen, S. Zhou, F. Zhang, R. Zou, Q. Hu, An improved Wavenet network for multi-step-ahead wind energy forecasting, Energy Convers. Manag. 278 (2023) 116709.
[30] G. Alkhayat, R. Mehmood, A review and taxonomy of wind and solar energy forecasting methods based on deep learning, Energy and AI (2021) 100060.
[31] U. Singh, M. Rizwan, A systematic review on selected applications and approaches of wind energy forecasting and integration, J. Inst. Eng.: Ser. Bibliogr. 102 (2021) 1061–1078.
[32] W. Orasch, Regulatory framework for RES-E system integration in Europe-Description and analysis of different European practices, Appendix to Deliverable 8 (2009).
[33] H. Zsiborács, N. Hegedűsné Baranyai, S. Csányi, A. Vincze, G. Pintér, Economic analysis of grid-connected PV system regulations: a Hungarian case study, Electronics 8 (2019) 149.
[34] A.H. Elsheikh, S.W. Sharshir, M. Abd Elaziz, A. Kabeel, W. Guilan, Z. Haiou, Modeling of solar energy systems using artificial neural network: a comprehensive review, Sol. Energy 180 (2019) 622–639.
[35] S. Makridakis, E. Spiliotis, V. Assimakopoulos, Statistical and Machine Learning forecasting methods: concerns and ways forward, PLoS One 13 (2018) e0194889.
[36] H.K. Ghritlahre, R.K. Prasad, Application of ANN technique to predict the performance of solar collector systems-A review, Renew. Sustain. Energy Rev. 84 (2018) 75–88.

[37] M. AlShafeey, C. Csaki, The impact of input data resolution on neural network forecasting models for wind and photovoltaic energy generation using time series data, Environ. Prog. Sustain. Energy 42 (2023) e14039.
[38] M. Awad, R. Khanna, Support vector regression. Efficient Learning Machines, Springer, 2015, pp. 67–80.
[39] J. González-Sopeña, V. Pakrashi, B. Ghosh, An overview of performance evaluation metrics for short-term statistical wind power forecasting, Renew. Sustain. Energy Rev. 138 (2021) 110515.
[40] M.U. Yousuf, I. Al-Bahadly, E. Avci, Current perspective on the accuracy of deterministic wind speed and power forecasting, IEEE Access 7 (2019) 159547–159564.
[41] B. Yang, L. Zhong, J. Wang, H. Shu, X. Zhang, T. Yu, et al., State-of-the-art one-stop handbook on wind forecasting technologies: an overview of classifications, methodologies, and analysis, J. Clean. Prod. 283 (2021) 124628.