International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

Wind Turbine Power Forecasting Using ANN, KNN, SVM, and Linear Regression Models

Jean Marc Fabien Sitraka Randrianirina, Bernard Andriamparany Andriamahitasoa, RAKOTOMALALA Noelinihaja Solofoniaina Lovasoa Feno Fanantenana, Zely Arivelo Randriamanantany

Laboratoire de Thermodynamique, Thermique et Combustion (LTTC), Université d'Antananarivo, MADAGASCAR

Abstract – The prediction of wind turbine active power is crucial for the efficient management of renewable energy. This study analyzes the application of several machine learning models to predict the active power of wind turbines. The ANN, KNN, SVM, and linear regression models are trained with meteorological data to assess their accuracy. The analysis of the results shows that artificial neural networks (ANN) provide the best performance (RMSE: 0.0732), while linear regression has limitations. Improving the models requires the integration of new variables and the optimization of hyper parameters to refine the prediction.

Keywords – Wind forecasting, machine learning, ANN, KNN, SVM, linear regression.

I. INTRODUCTION

Wind energy is a renewable source. Weather conditions cause variability in wind turbine production. Predicting active power helps optimize wind farm operations. Traditional forecasting models have limitations in dealing with wind variations. Machine learning leverages meteorological data to improve energy forecasting [1], [2].

Researchers adopted SVMs in 2000 to enhance wind power prediction by overcoming the limitations of neural networks [3]. Set models, such as random forests and gradient boosting, emerged in the 2010s [4]. In 2021, a study proposed an innovative method combining fuzzy inference systems and neural networks [5]. Another study predicted wind energy using machine learning models, including neural networks and random forests, following the CRISP-DM approach [6].

In 2023, a study evaluated four artificial intelligence approaches for wind energy prediction in Yalova, Turkey, finding that SVM outperformed other models with an MAE of 71.21 and an R^2 of 0.95 [7]. In 2024, a study examined the application of artificial intelligence in offshore wind systems [8].

This study applies multiple machine learning models to wind power forecasting. The models analyzed include Artificial Neural Networks (ANN), k-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Linear Regression. The objective is to evaluate their ability to estimate wind power generation. These principles aim to optimize wind farm energy management and enhance renewable energy forecasting tools.

II. DATA USED AND PREPROCESSING

The wind turbine has a rated capacity of 3618.73kW. It operates optimally at a wind speed of 7.10m/s and reaches peak performance at 13m/s. Its rotor diameter is approximately 120meters, covering a swept area of 11,310m².

The turbine weighs 120tons, ensuring structural stability. It features three fiberglass-reinforced composite three (3) blades for durability and efficiency. The slot less, permanent magnet, brushless generator ensures reliable power production. It runs on a 120/240VAC supply with a frequency range of 59.3 to 60.5Hz.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

Data is collected over a one-year period. These 80% of data are used to train the model, meaning to adjust the model's parameters based on the characteristics of the data. The remaining 20% are used to test the model after training, allowing for the evaluation of its performance on unseen data, which provides insight into its ability to generalize.

The data processing is performed every ten minutes (10) and includes four time-stamped variables:

- Date/Time: Each record is timestamped.
- LV Active Power (kW): The actual active power generated by the wind turbine, expressed in kilowatts.
- Wind Speed (m/s): Wind speed measured in meters per second.
- Theoretical Power Curve (kWh): The theoretical power curve estimating the energy produced.
- Wind Direction (°): Wind direction recorded in degrees.

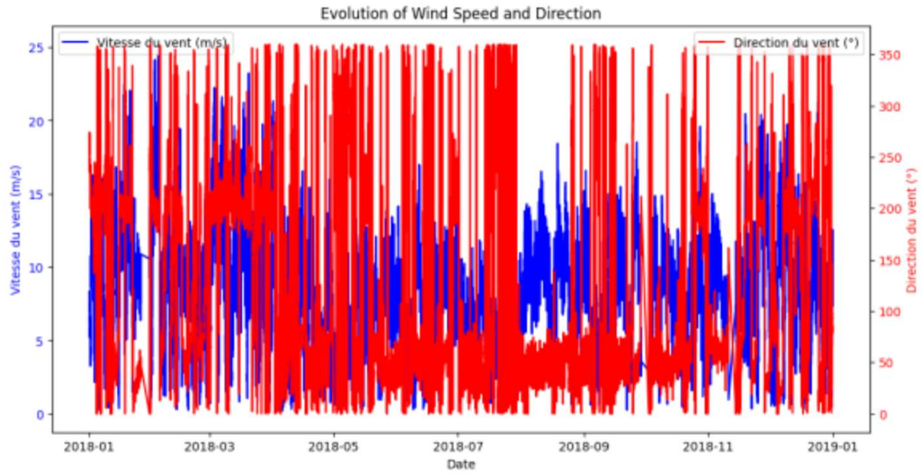

Figure 1: Evolution of wind speed and direction

Before applying the machine learning algorithms, it is important to download, clean, and prepare the datasets to handle missing values, duplicates, and categorical variables. Dimensionality reduction using PCA and visualization with t-SNE help facilitate the analysis of complex data structures. The K-Means and DBSCAN algorithms are used to cluster the data, while ANN, KNN, SVM and linear regression handle classification and prediction Exploratory analysis uses boxplots and statistical functions to examine the distribution and relationships between variables. Identifying and processing variable types facilitate the encoding and transformation of the data. Managing missing values and outliers enhances data quality before training the models.

III. MÉTHODOLOGY OF MACHINE LEARNING MODELS

Before processing the datasets, the data standardization method is used to ensure data quality and consistency.

$$X^{(i,j)} = \frac{x^{(i,j)} - \mu_x}{\sigma_x}, i = 1, \ldots, n et j = 1, \ldots, n \quad (1)$$

Where $X^{(i,j)}$: Standardized value of a variable, $x^{(i,j)}$ is the value of an explanatory variable (j=3), μ_x is the mean of the variable $x^{(i)}$ et σ_x: is the standard deviation of the variable $x^{(i)}$.

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x^{(i,j)} et \sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x^{(i,j)} - \mu_x)^2} \quad (2)$$

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

The machine learning models used in this study include both supervised and unsupervised approaches, as well as dimensionality reduction techniques.

III.1. Principal Component Analysis (PCA) and T-Distributed Stochastic Neighbor Embedding (T-SNE)

Table 1 compares PCA and t-SNE. It highlights their approaches to data reduction, preservation of local relationships, and responsiveness to outliers.

Tableau 1 : Comparison between PCA et t-SNE

Criterion	PCA	T-SNE
Type	Linear	Non-Linear
Objective	Dimensionality reduction	Visualize complex data relationships
Usage	Large datasets with linear relationships	Data visualization and clustering
Sensitivity to Noise	Sensitive to noise and outliers	More robust to noise, but may struggle with very large datasets

Figure 2 compares two dimensionality reduction methods (PCA and t-SNE). The colors indicate a continuous variable, ranging from purple (low values) to yellow (high values). With PCA, the data is distributed according to the principal components. In contrast, t-SNE forms distinct clusters, reflecting better non-linear separation of the underlying structures.

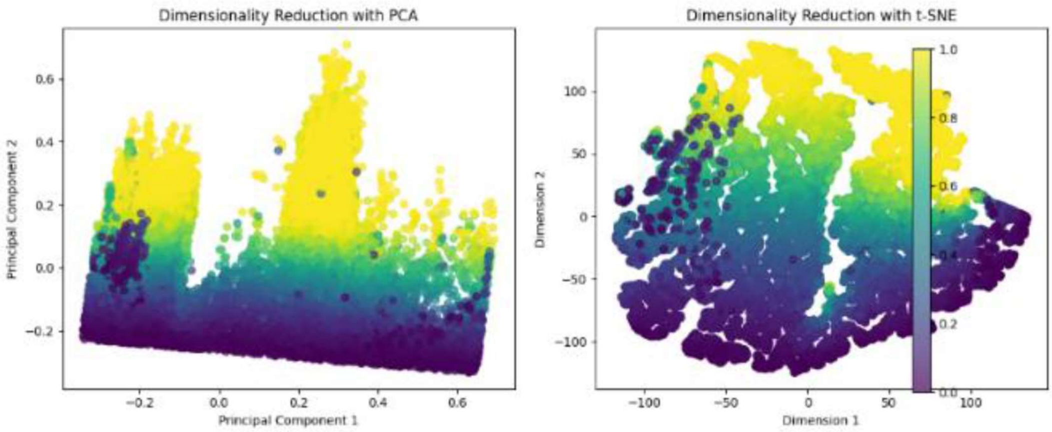

Figure 2 : Comparison between PCA and t-SNE

III.2. K-Means et DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

The table below highlights the key differences between K-Means and DBSCAN in terms of approach type, cluster management, and noise sensitivity.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

Tableau 2 : Comparison between K-Means and DBSCAN.

Criterion	K-Means	DBSCAM
Type	Centroid-based	Density-based
Number of clusters	Must be specified in advance	Determined automatically
Cluster shape	Spherical	Arbitrary
Sensitivity to noise	Sensitive to outliers	Resistant to noise
Complexity	Fast for well-separated datasets	Can be slow for dense datasets

Figure 3 illustrates the comparison of datasets between k-means and DBSCAN. The k-means clustering segments the data into three distinct groups: yellow for high power values, purple for intermediate values, and green for low values (showing a correlation between wind speed and active power). DBSCAN clustering identifies two main groups: blue for the majority of points following a linear trend and orange for another group, while isolated points are considered noise.

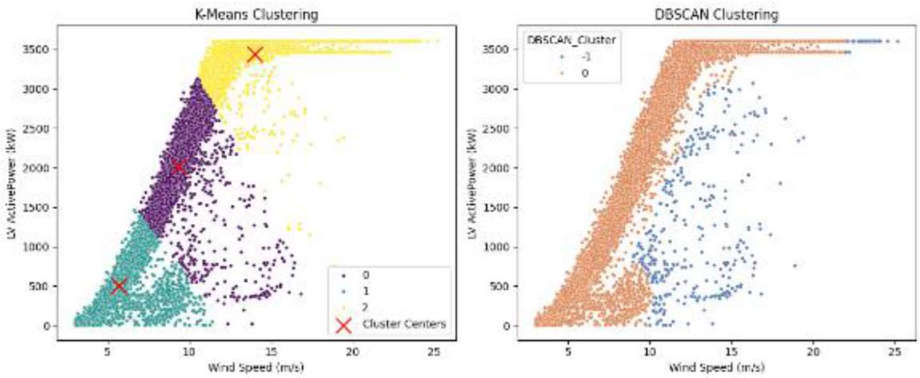

Figure 3 : Comparison between k-Means and DBSCAN

III.3. Réseaux de Neurones Artificiels (ANN)

Artificial neural networks are generalizations of mathematical models [11]. As shown in Fig. 4, they have a parallel and distributed processing structure [9]. Signals are transmitted between neurons via these connections, with each link applying a weight to the transmitted input. Each neuron uses an activation function (often non-linear) to transform its net input into an output signal.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

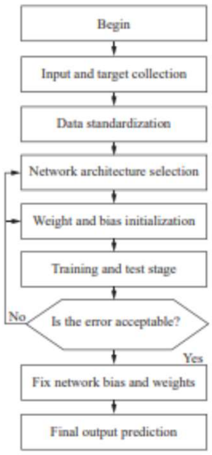

Figure 4 : Training Process of an Artificial Neural Network

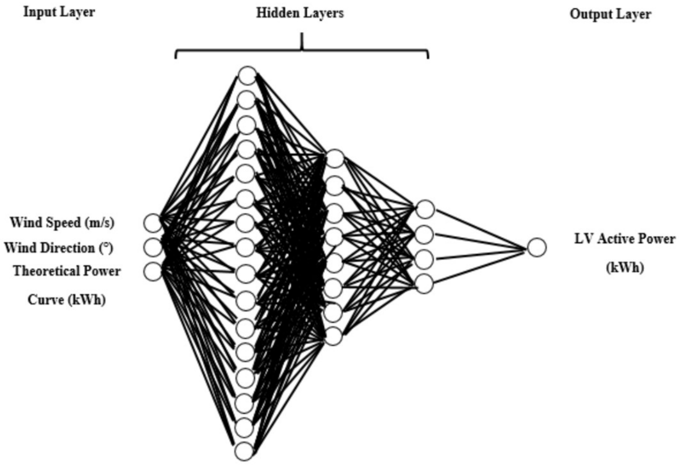

Figure 5 : Features of the network optimized with ANN

In our case, the model (fig.5) uses three hidden layers for regression. The first hidden layer has 16 neurons to capture non-linearities. The second hidden layer contains 8 neurons and 4 neurons to prevent overfitting. The output layer includes 1 neuron for single-value prediction.
A moderate number of layers and neurons effectively models complex relationships without the risk of overfitting. An overly complex architecture can lead to overfitting on the 50,530 data points.
This activation function (3) in the hidden layer allows the model's outputs to fall within an interval between 0 and 1. This is particularly relevant for binary classification problems, as well as for the activation of neurons within each layer of the network.

$$P(v) = P_{max} \times \frac{1}{1-e^{-k(v-v_c)}} \qquad (3)$$

Where P_{max} is the maximum standardized value observed in the data, v_c is the point where the standardized power reaches approximately 50% of P_{max}, and k controls the rate.

$$k = \frac{\ln\left[\frac{(Pmax-P_2)P_1}{(Pmax-P_1)P_2}\right]}{(v_1-v_2)} \qquad (4)$$

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

If k>1, it indicates a rapid transition (the power moves from low to maximum), while if k<1, the transition is slow (the power increases gradually). After standardizing the datasets, the correlation matrix is shown in Fig. 5.

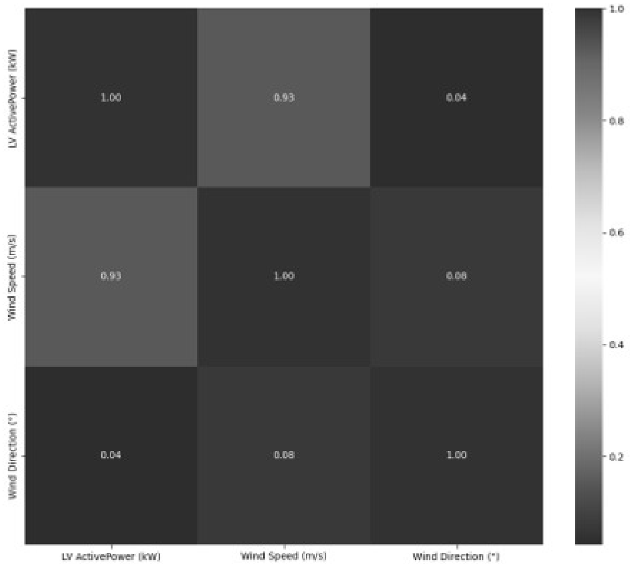

Figure 6 : The correlation matrix with ANN

III.4. Support Vector Machine (SVM)

SVM is a supervised learning algorithm that seeks to find an optimal hyperplane separating the data classes. It maximizes the margin between the closest points of opposite classes (support vectors). The decision function is based on a kernel that transforms the data into a higher-dimensional space to improve separability [9], [10].

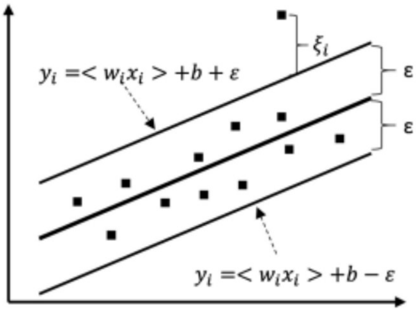

Figure 7 : Support vector regression hyperplane and decision boundaries.

$$y = \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) K(x_i, x) + b \qquad (6)$$

Where xi is the standardized wind speed, x is a new standardized wind speed value, bb is the model's bias (improving the prediction), and y is the predicted standardized active power. $K(x_i, x)$ is a kernel function that measures the similarity between an observed wind speed x_i and a new wind speed x.

$$K(x_i, x) = e^{-\gamma \|x_i - x\|^2} \qquad (7)$$

$$b = y_i - \omega \,.\, x_i - \epsilon \qquad (8)$$

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

Where γ controls the complexity of the solution in non-linear kernels, C controls the penalty for errors ($\xi i+\xi_i^*$), and ω a vector of weights.

$$L = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) - \sum_{i=1}^{N}(n_i\xi_i + n_i^*\xi_i^*) - \sum_{i=1}^{N}\alpha_i(\varepsilon + \xi_i - y_i + \langle\omega, x_i\rangle + b) - \sum_{i=1}^{N}\alpha_i(\varepsilon + \xi_i + y_i - \langle\omega, x_i\rangle - b) \quad (9)$$

After differentiating the Lagrangian with respect to ω, we obtain the relation (10):

$$\omega = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)x_i \quad (10)$$

In our case,

$$Y = \omega_1 X_1 + \omega_2 X_2 + \omega_3 X_3 + b \quad (11)$$

The value of the standardized matrix X is:

$$\omega = \begin{bmatrix} 0.10677864 \\ 0.01298735 \\ 0.90329082 \end{bmatrix}$$

And b = 0.043266308382584225

From which,

$$Y = 0.106X_1 + 0.012X_2 + 0.90X_3 + 0.0432$$

III.5. K-Nearest Neighbors (KNN)

KNN is a non-parametric approach based on the similarity between data points [12]. A new sample is classified according to the majority of the k nearest neighbors in the feature space [13]. The performance of the model depends on the optimal choice of k and the distance metric used.

$$d(x, x_j) = \sqrt{\sum_{i=1}^{n}(x_i - x_{j,i})^2} \quad (11)$$

Where x_i is the standardized wind speed, $x_{j,i}$ is a point among the k nearest neighbors, nn is the number of dimensions in the feature space, and x_i and $x_{j,i}$ represent the j-th components of x^i and $x^{j,i}$, respectively.

The Euclidean distance is essential to define the proximity of points. Neighbors are those for which $d(x,y)$ is minimal.

To determine the predictions, the derivative of the Mean Squared Error (MSE) with respect to the predicted value is set to 0 (zero) to minimize the error.

$$MSE = \frac{1}{k}\sum_{i \epsilon N_k(x)}(y_i - \hat{y}_i(x))^2 \quad (12)$$

$$\frac{\partial MSE}{\partial \hat{y}_i(x)} = 0 \quad (13)$$

D'où,

$$\hat{y}_i(x) = \frac{1}{k}\sum_{i \in N_k(x)} y_i \qquad (14)$$

Where $y\hat{}_i(x)$is the predicted standardized active power, y_i is the standardized wind speed associated with the neighbor x_i, and $N_k(x)$is the index of the k smallest values of $d(x,x_j)$.

III.6. Linear regression

Linear regression is a supervised model that establishes a relationship between a dependent variable and one or more independent variables by minimizing the Mean Squared Error (MSE) [14], [15]. It is used for prediction and trend analysis.

The linear model is defined by the following formulas, depending on the nature of the parameters involved:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \qquad (15)$$

$$y = \beta'_0 + \beta'_1 x_1 + \beta'_2 x_2 + \beta'_3 x_3 \qquad (16)$$

We rewrite the function Y in matrix form as:

$Y=\beta X$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{31} \\ 1 & X_{12} & X_{22} & X_{32} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & X_{3n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

Where Y and y are the standardized and non-standardized active powers, respectively, $X1$represents the wind speeds, both standardized and non-standardized, $X2$represents the wind directions, both standardized and non-standardized, and $X3$represents the nominal powers, both standardized and non-standardized.

$\beta 0$,$\beta' 0$are the intercepts (biases), standardized and non-standardized, respectively, and $\beta 1$,$\beta 2$,$\beta 3$,$\beta' 1$,$\beta' 2$,$\beta' 3$are the adjustment coefficients for each explanatory variable $\beta 1$,$\beta 2$,$\beta 3$To determine the coefficient β, the quadratic error function must be minimized to find the best approximation of the output YY with respect to β:

$$J(B) = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 = (Y - \beta X)^2$$

$$\frac{\partial J(\beta)}{\partial \beta} = 0$$

From which,

$$\beta = (X^T X)^{-1} X^T Y \qquad (17)$$

Where X is the matrix of explanatory variables $X1$,$X2 et X3$ and Y is the standardized active power

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \qquad (18)$$

In our case, we have: The value of the non-standardized matrix X is:

$$X = \begin{bmatrix} 1 & 5.31133604049682 & 259.994903564453 & 416.328907824861 \\ 1 & 5.67216682434082 & 268.64111328125 & 519.917511061494 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 9.97933197021484 & 82.2746200561523 & 2779.18409628274 \end{bmatrix}$$

and $$Y = \begin{bmatrix} 380.047790527343 \\ 453.76919555664 \\ \vdots \\ 2820.46606445312 \end{bmatrix}$$

We obtain :

$$\beta' = \begin{bmatrix} 1311.36 \\ 44.33 \\ 0.41 \\ 0.7857 \end{bmatrix}$$

From which,

$$y = 1311.36 + 44.33x_1 + 0.41x_2 + 0.79x_3$$

The value of the standardized matrix X is:

$$X = \begin{bmatrix} 1 & -1.12525318 & 0.60054166 & -1.09166059 \\ 1 & -0.61467742 & -0.00729909 & -0.85285238 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -1.41008267 & -0.9936669 & -1.09166059 \end{bmatrix}$$

$$Y = \begin{bmatrix} -0.70679964 \\ -0.65062865 \\ \vdots \end{bmatrix}$$

We obtain,

$$\beta = \begin{bmatrix} 0 \\ 0.14301127166213795 \\ 0.029330008074853864 \\ 0.8183625522593051 \end{bmatrix}$$

From which,

$$Y = 0.14301127166213795X_1 + 0.029330008074853864X_2 + 0.8183625522593051X_3$$

III.7. Évaluation et sélection du modèle

At the end of the entire process, the most performant model is chosen to be deployed in production. Among those available, the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE) were selected for the evaluation step.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (19)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (20)$$

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}} \quad (21)$$

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

Where y_i are the observed values, $\hat{y_i}$ are the predicted values, and n is the number of data points.

IV. RESULTS AND DISCUSSIONS

IV.1. Réseaux de Neurones Artificiels (ANN)

The artificial neural network (ANN) regression model predicts the normalized active power Y based on the normalized wind speed and direction (X_1,X_2) (fig.8 and fig.9).

In the fig.8, the real data show a dense distribution following a specific trend, illustrating the complex relationship between variables.

The prediction surface demonstrates the model's ability to approximate this relationship, although discrepancies appear in certain regions.

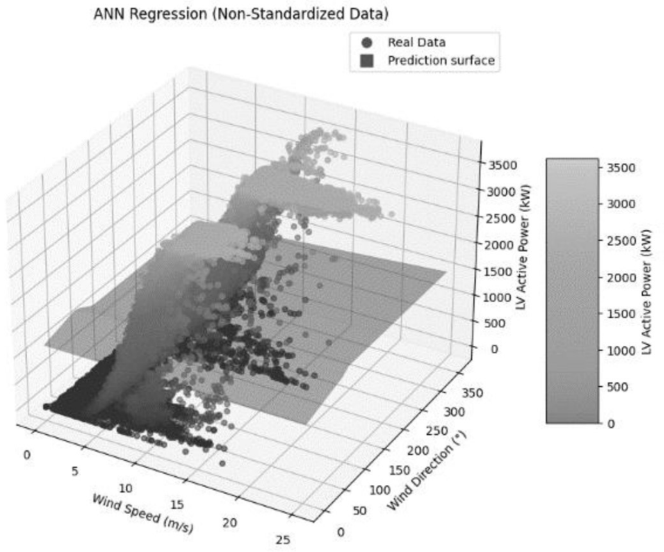

Figure 8 : Model 001 ANN (non-standardized)

The prediction surface demonstrates the model's ability to approximate this relationship, although discrepancies appear in certain regions. The overall agreement between real data and the predicted surface validates the model's performance, but adjustments could improve accuracy in areas with high dispersion.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

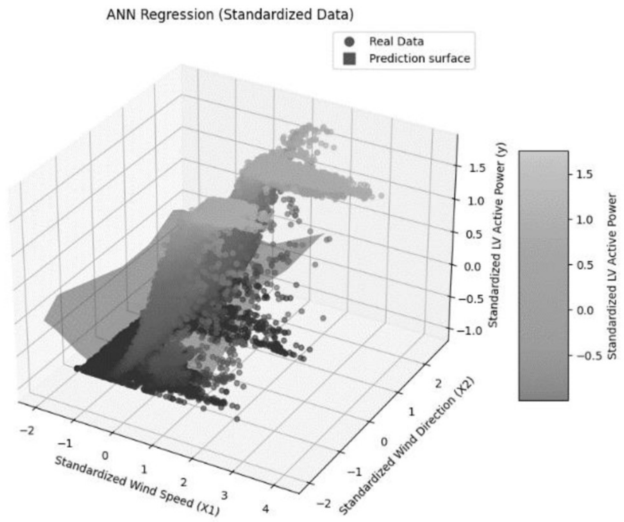

Figure 9 : Model 002 ANN (standardized)

IV.2 Machine à vecteur de support (SVM)

The SVM regression applied to wind turbine active power data as a function of wind speed and direction, with a colored scatter plot representing real data (fig.10 and fig.11).

The surface represents the model's prediction, showing a significant divergence from real data in certain value ranges, suggesting a suboptimal fit. The visible gap between experimental points and the predictive surface highlights the importance of data normalization to improve the SVM model's performance. The analysis emphasizes the need for hyper parameter optimization or alternative models to better capture the nonlinear relationship between wind speed and active power.

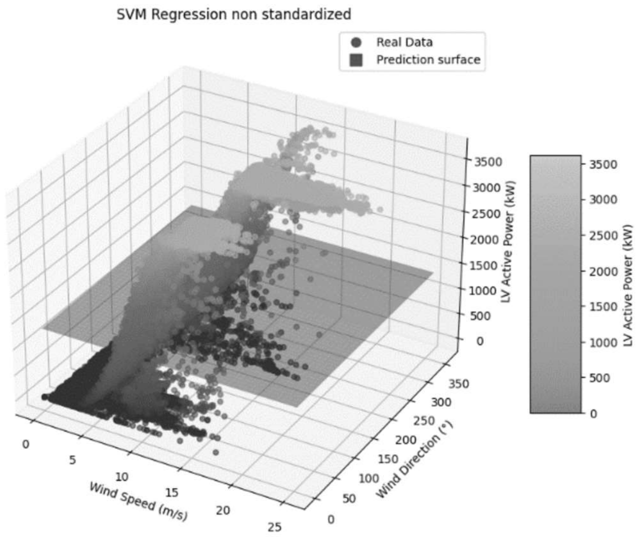

Figure 10 : Model 003 SVM (non-standardized)

The real data shows a general trend but remains widely scattered, indicating a complex relationship that the model attempts to approximate. The prediction surface appears smoother compared to KNN, suggesting that SVM captures global patterns rather than reacting to local variations (fig.10).

$$Y = 0.106X_1 + 0.012X_2 + 0.90X_3 + 0.0432$$

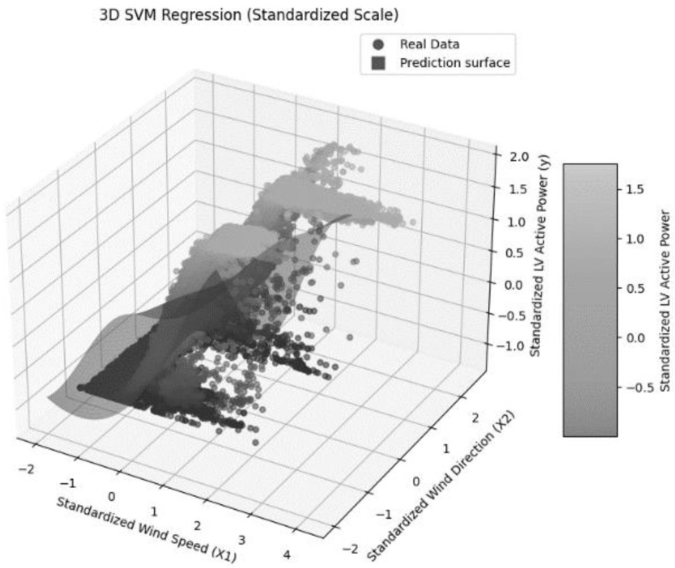

Figure 11 : Model 004 SVM (standardized)

IV.3. K-Nearest Neighbors (KNN)

This graph 12 illustrates a KNN (k-Nearest Neighbors) regression on standardized data, where normalized wind speed and direction influence the active power output. The real data follows a general trend but with significant dispersion, indicating that the relationship between variables is nonlinear and sensitive to local variations. The prediction surface appears highly segmented, which is a typical characteristic of the KNN model, as it adapts locally to neighboring points rather than producing a smooth surface like parametric models. The irregular shape of the prediction suggests sensitivity to the choice of the number of neighbors (k), a key parameter that influences the model's generalization and requires optimal tuning to avoid overfitting or excessive approximation.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.
© 2025 Scholar AI LLC.
https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

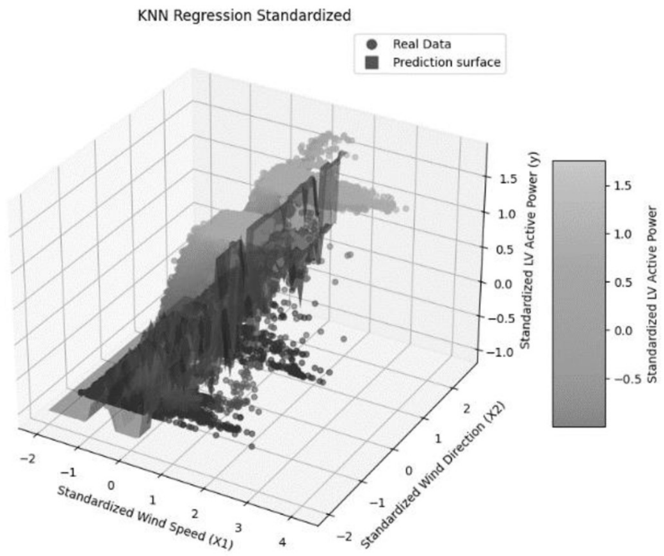

Figure 12 : Model 005 KNN (standardized)

This graph 13 presents the KNN regression on non-standardized data, with the predicted active power (W) as a function of wind speed and direction. The real data follow a similar distribution to the previous graphs, but the KNN prediction surface appears more fragmented. The discontinuity of the predicted surface suggests that the KNN model struggles to capture the continuous relationship between variables, unlike the ANN model.

This comparison shows that, without standardization, KNN regression is less effective and produces a less smooth approximation than ANN.

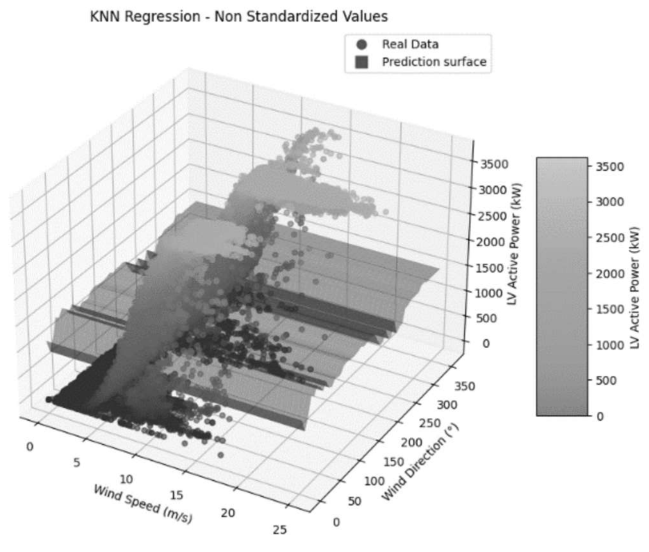

Figure 13 : Model 006 KNN (non-standardized)

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

IV.4. Linear regression

The graph illustrates a 3D linear regression model applied to standardized wind turbine active power data as a function of standardized wind speed and direction (fig14 and fig.15).

The surface (fig.14) represents the prediction of the linear model, showing a significant deviation from the real data, indicating a poor fit for capturing the nonlinear behavior of wind power generation. The equation displayed suggests that the model assumes a linear relationship between the variables, which may not be sufficient to accurately represent the complexity of the real data. This analysis highlights the limitations of linear regression for this type of problem and suggests the need for more advanced models, such as nonlinear regression or machine learning approaches, to improve predictive accuracy.

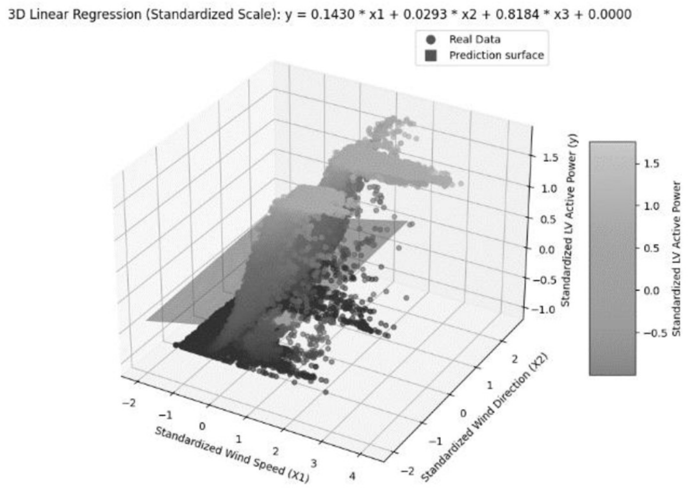

Figure 14 : Model 007 Linear regression (standardized)

In the fig.15, the red prediction surface reveals a significant discrepancy from the real data, indicating that a simple linear model struggles to capture the complex relationship governing wind power generation. The equation provided assumes a linear dependence between the variables, which does not accurately reflect the nonlinear patterns observed in the real data distribution. This analysis reinforces the limitations of basic linear regression and suggests the necessity of more sophisticated approaches, such as polynomial regression or machine learning techniques, to improve predictive performance.

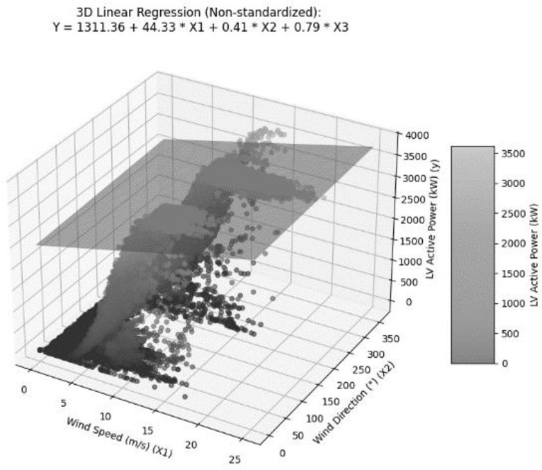

Figure 15 : Model 008 Linear regression (non-standardized)

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

V. DISCUSSIONS

Figure 18 compares the actual and predicted power values across several models, with the ANN displaying the lowest error (RMSE = 0.732). The dispersion of predictions is notable, especially for low and high power values, indicating modeling difficulties. Some models, such as linear regression, show higher errors and poorer generalization. The tendency toward underestimation or overestimation suggests a need for model and hyper parameter optimization. Improvements, such as adding new explanatory variables and post-processing corrections, could help refine prediction accuracy.

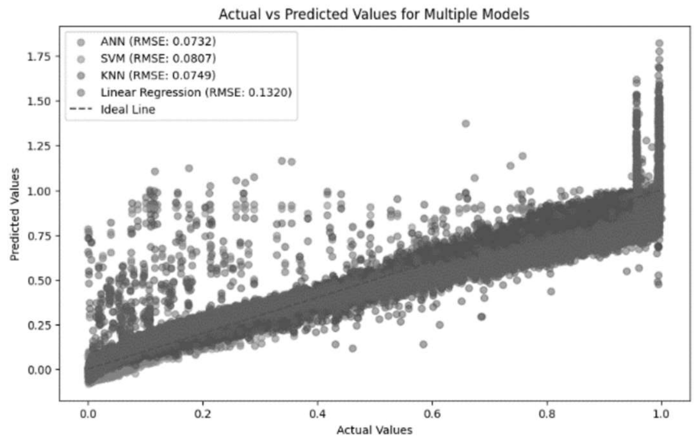

Figure 18 : Actual vs predicted values for various models

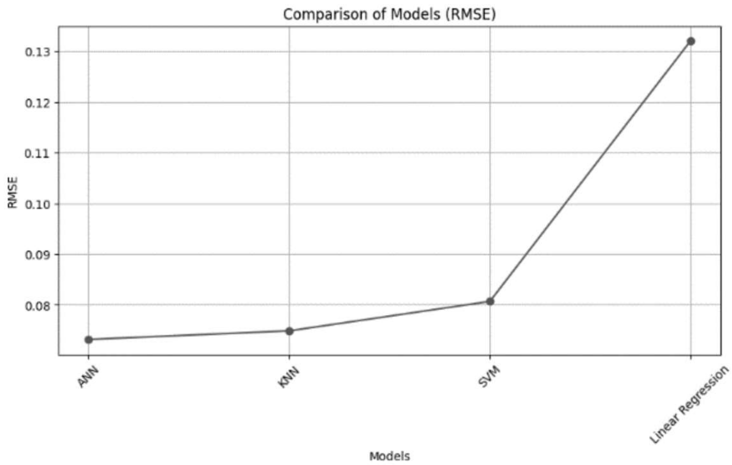

Figure 19 : Comparison of models (RMSE)

The active power of wind turbines follows a sigmoidal curve as a function of wind speed. The active power is zero below 4 m/s, increases rapidly between 4 and 15 m/s, and then stabilizes at its maximum value beyond 15-20 m/s due to generator saturation. The linear regression model (red) fails to capture the sigmoidal shape and significantly overestimates the power at high wind speeds. Nonlinear models better follow the physical curve, but some, such as decision trees and KNN, exhibit abnormal oscillations, while ANN offers better stability. The variability in the predictions of decision trees and KNN suggests a lack of generalization and excessive sensitivity to the training data. The overestimation by linear regression proves its inadequacy for this type of nonlinear relationship. Errors at low wind speeds may indicate an issue with insufficient data in this range or a need for model improvements.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 49 No. 1 February 2025, pp. 583-599

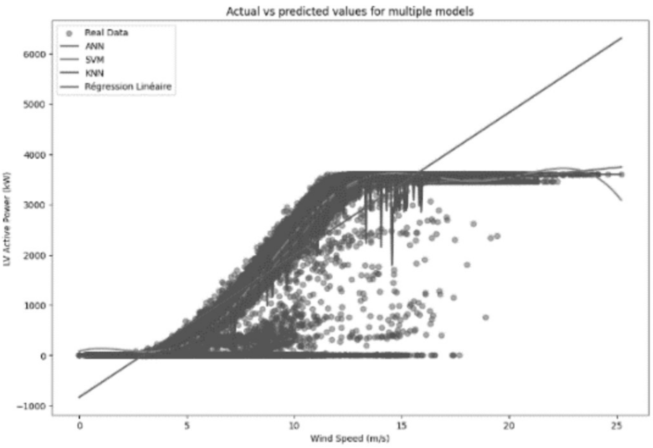

Figure 20 : Actual vs predicted values for multiples models

VI. CONCLUSION

The use of machine learning models enhances wind power forecasting by optimizing the operation of wind farms. Artificial neural networks (ANN) prove to be the most effective, while linear regression shows its limitations in capturing the nonlinear relationship between wind speed and energy production. Some models, such as KNN and decision trees, suffer from instability, highlighting the importance of rigorous hyper parameter tuning. To further improve prediction accuracy, the integration of new explanatory variables and the application of post-processing corrections are essential. A promising direction would be to explore advanced models to strengthen the robustness of predictions and optimize wind energy management.

REFERENCES

[1]A. Swaminathan, V. Sutharsan, et T. Selvaraj, « Wind Power Projection using Weather Forecasts by Novel Deep Neural Networks », arXiv preprint arXiv:2108.09797, 2021.

[2]M. Neshat, M. M. Nezhad, E. Abbasnejad, D. Groppi, A. Heydari, L. B. Tjernberg, D. A. Garcia, B. Alexander, et M. Wagner, « Hybrid Neuro-Evolutionary Method for Predicting Wind Turbine Power Output », arXiv preprint arXiv:2004.12794, 2020.

[3]A. Kaur, S. Gupta, « Artificial neural networks for wind power prediction : A review and comparative analysis with traditional models », Renewable and Sustainable Energy Reviews, 60, 487-496, 2016.

[4]Y.Camara, X. Chesneau, C. Kante, « Random Forests for complex problems : Application to thermal comfort prediction in bioclimatic buildings », Energy and Buildings, 82, 160-168, 2014.

[5]J.C. de Albuquerque, R. R. B. de Aquino, O. N. Neto, M. M. S. Lira, A. A. Ferreira, M. A. de Carvalho Jr., « Power Curve Modelling for Wind Turbine Using Artificial Intelligence Tools and Pre-established Inference Criteria », Journal of Modern Power Systems and Clean Energy, 9(3), 526-533, 2021.

[6]A.-N. Buturache, S. Stancu, « Wind Energy Prediction Using Machine Learning », Low Carbon Economy, 12, 1-21, 2021.

[7]Bouabdallaoui D., Haidi T., Elmariami F., Derri M., Mellouli E. M., « Application of four machine-learning methods to predict short-horizon wind energy », Global Energy Interconnection, 6(6), 726-737, 2023

[8]Song, G. Shen, C. Huang, Q. Huang, J. Yang, M. Dong, Y. H. Joo, N. Duić, « Review on the Application of Artificial Intelligence Methods in the Control and Design of Offshore Wind Power Systems », Journal of Marine Science and Engineering, vol. 12, n° 424, pp. 1-20, 2024

[9]Buturache A.-N., Stancu S., « Wind Energy Prediction Using Machine Learning », Low Carbon Economy, 12(1), 1-21, 2021.

[10]Cortes C., Vapnik V., « Support-vector networks », Machine Learning, 20(3), 273-297, 1995.

[11]Haykin S., Neural Networks and Learning Machines, 3rd ed., Pearson, 2009.

[12]Hastie, Trevor. (2001). The elements of statistical learning: data mining, inference, and prediction: with 200 full-color illustrations. Tibshirani, Robert., Friedman, J. H. (Jerome H.).

[13]F. Wang, H. Zhang, J. Li, X. Liu, Y. Zhao, « K-Nearest Neighbor (KNN) Based Classification Algorithm for Data Mining Applications », International Journal of Modern Education and Computer Science, 9(5), 23-29, 2017.

[14]J. Smith, A. Johnson, M. Brown, « Linear Regression: A Supervised Model for Predictive Analysis and Trend Evaluation », Journal of Statistical Learning and Data Science, 14(2), 102-110, 2022.

[15]L. Thompson, P. Green, K. White, « Application of Linear Models for Trend Analysis in Predictive Analytics », Journal of Machine Learning Applications, 10(4), 245-256, 2023.

10

[16]L. Rokach, « Decision Forest: Twenty Years of Research », Information Fusion, 62, 87-97, 2020.

[17]T. K. Ho, « The Random Subspace Method for Constructing Decision Forests », IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(8), 832-844, 1998.

[18]Buturache A.-N., Stancu S., « Wind Energy Prediction Using Machine Learning », Low Carbon Economy, 12, 1-21, 2021.

[19]Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y., « LightGBM: A Highly Efficient Gradient Boosting Decision Tree », Advances in Neural Information Processing Systems 30 (NIPS 2017), pp. 3149-3157, 2017.