Embedded Self Organizing Systems (Vol 11. No 9. 2024) (pp.14-22)

Embedded Selforganizing Systems

Issue Topic: "Artificial Intelligence & Machine Learning in Computer Vision Applications "

Machine Learning Approaches for Wind Speed Prediction: A Case Study in Ajman, UAE

Kais Muhammed Fasel
University of Bolton
School of Civil Engineering and Built Environment
kk1res@bolton.ac.uk

Peter Farrell
University of Bolton
School of Civil Engineering and Built Environment
p.farrell@bolton.ac.uk

Abdul Salam Darwish
University of Bolton
School of Civil Engineering and Built Environment
a.darwish@bolton.ac.uk

***Abstract*—This study presents a novel methodology for error-correcting publicly available NASA wind data and making it more accurate for location-specific wind resource assessments for the emirate of Ajman, UAE. The approach integrates maximum climate and environmental variables, data-driven techniques and machine learning algorithms, addressing the inherent errors in publicly available NASA satellite data. The study establishes error correction factors for satellite-derived wind speed data, enhancing the dependability of wind speed data for sustainable wind resource assessment and power production forecasting. The findings of this study have significant implications for wind energy industry stakeholders and the government for decision-making and sustainability initiatives.**

Keywords—Machine learning, renewable energy, wind resource assessment, data analysis, meteorological variables, sustainability, Ajman, UAE

I. INTRODUCTION & BACKGROUND

The impact of climate change has influenced the world to shift towards sustainable energy sources that have tremendously accelerated wind power development. Accurate wind speed prediction and resource assessment are crucial for successfully planning and managing wind energy projects. Machine learning (ML) techniques were tested for site-specific wind speed prediction to enhance resource assessment, leveraging their ability to identify patterns, capture non-linear relationships, and handle large datasets [1, 2]. As per [3] various ML algorithms, namely the Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Machine (GBM) algorithm [3] got promising results in wind speed prediction and resource assessment [4, 5, 6].These algorithms can utilise wind speed data, other climate parameters, and environmental parameters to develop predictive models for predicting future wind speeds and estimating the site's wind energy.

However, the correctness of wind speed estimates mainly depends on bias-free availability of input data [7]

Satellite-derived wind speed data, such as those provided by NASA, offer a valuable resource for regions with limited ground-based measurements[8]. Nevertheless, satellite-derived data often have lower accuracy than ground-based measurements due to the limitations of remote sensing techniques [9]. Correction Factor Analysis (CFA) is used to correct the errors and biases of NASA's publicly available satellite wind speed data [10].

Starting from 2010, the Emirate of Ajman is in the forefront of striving to implement a renewable energy mix [11]. Given its coastal location and exposure to strong winds, wind energy presents a promising opportunity for diversifying Ajman's energy mix. However, successful wind energy development in Ajman requires site-specific accurate and reliable wind resource assessments that consider the region's unique environmental characteristics and challenges with data availability.

Recent studies by [12,13] have highlighted the need for site-specific, error-corrected wind resource assessments in the emirate of Ajman. Building upon these findings, the current study aims to investigate the application of ML techniques for wind speed prediction and data correction to support the site-specific wind resource assessment and development of sustainable wind energy systems in the Emirate.

II. OBJECTIVES OF THE STUDY

The main objectives of this study are to:

- Evaluate the performance of Random Forest, Support Vector Machine, and Gradient Boosting Machine algorithms for wind speed prediction in Ajman using wind data and relevant meteorological variables collected from the Ajman X ground station and NASA satellite source.
- Assess the quality and suitability of NASA satellite-derived wind data for sustainable wind energy development in the emirate and investigate the application of CFA for data correction.
- Develop a comprehensive wind resource assessment framework for Ajman that integrates ML-based wind speed prediction, data correction techniques, and site-specific environmental characteristics.
- Identify future research directions and challenges in applying ML for wind resource assessment.
- By addressing these objectives, this study intends to deliver results to the wind energy community,

increase knowledge on ML applications in wind energy, and support the sustainable development of wind power in the emirate of Ajman and the broader UAE region.

III. Literature Review

The increasing requirement for sustainable power has opened the way for increased interest in wind power as a sustainable and renewable source [14]. Accurate wind resource assessment is crucial for efficiently planning and operating wind energy projects [15]. However, the inherent variability of wind poses challenges in predicting wind speeds and estimating wind power potential [16]. To address these challenges, researchers have turned to machine learning (ML) techniques, which have shown promise in improving the accuracy of wind speed predictions and resource assessments [17].

ML focuses on developing algorithms and models that can learn and improve their performance based on data [18]. ML algorithms are divided into supervised learning, unsupervised learning, and reinforcement learning algorithms [19]. This literature review explores ML techniques used for wind speed prediction, focusing on their application in wind resource assessment. Additionally, we discuss the importance of data quality and explore methods for correcting and enhancing wind speed data, specifically from satellite-derived sources like NASA.

A. *Machine Learning Techniques for Wind Speed Prediction*

1) *Random Forest*

Random Forest (RF) is an ensemble learning method that combines multiple decision trees to make predictions [20]. It has been widely used in wind speed [20]prediction due to its ability to handle high-dimensional data and capture non-linear relationships [4]. [4] utilized RF to predict hourly wind speed in Tunisia and discovered that it performed better than other machine learning models, such as Support Vector Regression (SVR) and Artificial Neural Networks (ANN).

2) *Support Vector Machine*

Support Vector Machine (SVM) is a popular ML algorithm that has been extensively used for wind speed prediction [21]. SVM takes the input data and transforms it into a high-dimensional feature space, then identifies the best hyperplane that maximizes the margin between distinct classes[22]. [5] employed the SVR algorithm for short-term wind speed forecasting in China and demonstrated its superiority over autoregressive integrated moving average (ARIMA) models.

3) *Gradient Boosting Machine*

Gradient Boosting Machine (GBM) is an ensemble learning method that combines weak learners. typically, decision trees, to create a strong predictive model [23]. GBM iteratively trains new models to minimise the residual errors of the previous models, thus improving the overall prediction accuracy [24]. [6] applied GBM for short-term wind speed forecasting in China and found that it outperformed other ML models such as RF and SVR.

B. *Data Quality and Correction Methods*

1) *Importance of Wind Data Quality*

The accuracy of wind speed predictions heavily relies on the quality of the input data [25]. Wind speed data can be obtained from various sources, including ground-based measurements, meteorological masts, and satellite-derived datasets[26] However, these datasets often contain errors, inconsistencies, and missing values, which can significantly impact the validity of ML models [27] . Therefore, it is crucial to pre-process and clean the wind speed data before using it for prediction tasks.

2) *NASA Satellite-Derived Wind Speed Data*

NASA provides a valuable free source of satellite-derived wind data that are usually used for wind resource assessment in regions where ground-based measurements are scarce [28]. However, satellite-derived wind speed data often has lower accuracy compared to ground-based measurements due to the inherent limitations of remote sensing techniques[29].

3) *Correction Factor Analysis*

Correction Factor Analysis (CFA) is a technique for adjusting satellite-derived wind speed data based on ground-based measurements. [10]. [30] Applied CFA to correct NASA satellite-derived wind speed data in Europe and reported significant improvements in the accuracy of wind resource assessments.

C. *Wind Resource Assessment in Ajman, UAE*

[12] conducted a preliminary investigation into wind speed variations between two data sources in the Emirate of Ajman, UAE. The [31] study compared wind speed data from two different sources: the European Centre for Medium-Range Weather Forecasts (ECMWF), Reanalysis version 5 (ERA5), and NASA Satellite Power Data for an offshore and an onshore location in Ajman. The study found substantial variation in wind speed frequency distributions among the two data sources, with ERA5 showing lower wind speeds than the NASA data. The findings highlight the importance of using site-specific, error-corrected wind data for accurate wind resource assessments.

[13] reviewed Ajman's wind energy potential and possible use for various applications, mainly desalination, street lighting, and building-integrated wind turbines. The study estimated that the Emirate's available wind power has the potential to provide 20% of its total electricity requirement, which can very well support Ajman's efforts towards sustainable development.

D. *Future Directions and Challenges*

Despite the advancements in ML techniques for wind speed prediction and data correction methods, future research still has challenges and opportunities. One area of interest is the integration of physics-based models with ML approaches to improve the interpretability and generalizability of wind speed predictions. [7]. Another challenge is the lack of error-free, economically feasible access to wind speed data in many regions, particularly developing countries [32]. Therefore, more extensive data collection campaigns and collaborative efforts to share wind speed datasets are needed to develop a sustainable wind energy mix.

E. Conclusion

This literature review highlights the significant progress made in using ML techniques for wind speed prediction and the importance of data quality and correction methods. ML algorithms such as Random Forest, Support Vector Machine, and Gradient Boosting Machine improved the quality of wind modelling. Additionally, correction methods like Correction Factor Analysis have effectively enhanced the accuracy of satellite-derived wind speed data, particularly from NASA. The studies by [12,13] emphasise the need for site-specific, error-corrected wind resource assessments in the Emirate of Ajman, UAE. The current research aims to fill the gaps in Ajman in meeting the requirements for developing wind power by examining applications of ML techniques for wind speed prediction and data correction. However, further research is required to fully address the Emirate's specific challenges and optimise the integration of wind energy into the existing power infrastructure.

IV. METHODOLOGY

The methodology employed in this study encompasses a systematic approach designed to investigate wind speed prediction methodologies and their application in renewable energy research. Leveraging machine learning techniques, this methodology aims to advance wind speed prediction accuracy while addressing the unique challenges posed by wind resource assessment in complex urban environments.

A. Data Preparation

The data preparation phase involves collecting and processing raw wind speed data from Ajman X station, a selected ground station for the case study, and supplementary data from NASA's satellite power data source. The study utilised the above-specified ground-based station wind data at Ajman X station and meteorological/climatic variables and site-specific environmental data (Please see Fig. 1), which is collected from Ajman Municipality & Planning Department meteorological data and satellite data parameters from NASA's publicly available power database. The collected data underwent pre-processing to address missing or incomplete values and standardised formats, ensuring consistency and reliability. Quality control measures are applied to validate data integrity and identify and mitigate discrepancies or outliers.

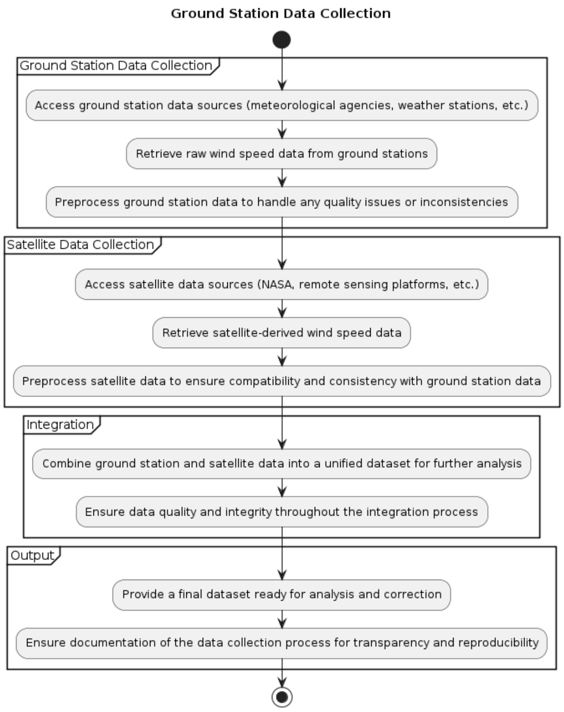

Fig. 1. Data Collection Workflow from Ground Stations and Satellite Sources

B. Data Preprocessing

Following data preparation, the processed dataset underwent feature engineering to extract meaningful insights into wind speed dynamics. Additional features were derived from existing variables using mathematical calculations, aggregation functions, and time-series analysis techniques. These engineered features, serving as proxies for atmospheric dynamics, topographical influences, and temporal variations, improved the performance of predictive models and enhanced understanding of wind resource characteristics.

Data pre-processing involved removing outliers to maintain the dataset's integrity. As the NASA satellite data was gathered at a height of 10 meters above ground level, a threshold of 12 m/s was chosen to address potential differences between ground-level measurements and satellite-derived data, following recommendations from previous research. Missing values in the satellite-derived wind speed dataset were addressed using KNN (K-Nearest Neighbors) imputation with a neighbourhood size of 5.

C. Feature Engineering

A feature importance analysis identified the significance of various meteorological variables in predicting wind speed at the Ajman X location Fig. 2. Wind direction emerged as the most influential feature, followed by air pressure, relative humidity, temperature, and air density. Although exhibiting minimal importance, roughness length was included for comprehensiveness. All variables contributed substantially to wind speed estimations. Feature engineering enhanced the dataset's richness and predictive power by deriving new features from existing variables, capturing meaningful insights into wind speed dynamics. Techniques employed included mathematical calculations, aggregation functions, and time-series analysis to compute wind speed differences, identify temporal trends, and extract recurring patterns.

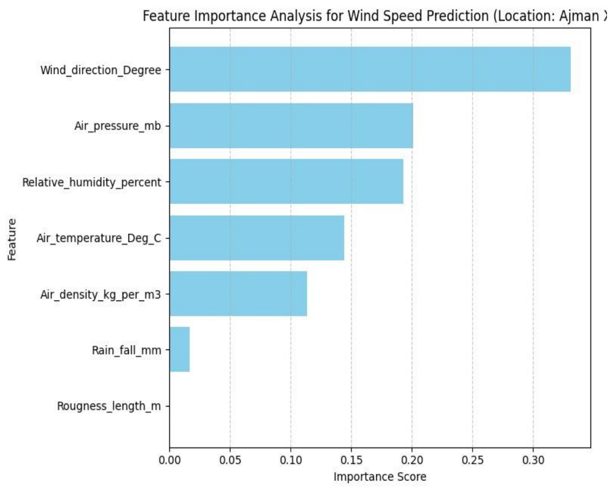

Fig. 2. Feature Importance Analysis for Wind Speed Prediction

D. Model Training and Evaluation

In this phase, the collected and pre-processed data were utilised to train machine learning models for wind speed prediction. In line with the methodology from [33] Three distinct algorithms, such as Random Forest, Support Vector Machine (SVM), and Gradient Boosting Machine (GBM), have been implemented. The dataset was split into two sets, one for training and the other for testing, and cross-validation techniques were implemented to validate the models' performance. The key assessment criteria were to estimate Mean Absolute Error (MAE) and R-squared, to check the efficiency of the chosen models.

E. Correction Factor Determination

Correction factors were determined through a comparative analysis between ground station measurements and satellite data to refine the accuracy and reliability of NASA satellite-derived wind speed data. Advanced statistical methods, including multiple linear regression and error correction methods, were used to estimate accurate correction factors to minimise discrepancies and enhance the fidelity of the NASA satellite data.

F. Error Correction

After determining correction factors, error correction procedures were implemented to refine NASA's satellite-derived wind speed dataset. The derived correction factors were applied to adjust the satellite data to align it more closely with ground-truth measurements. The effectiveness of these error correction techniques was validated against ground-truth data, ensuring the accuracy and reliability of the corrected dataset so that it can be used for site-specific sustainable wind resource assessment and further design of the wind energy mix.

G. Validation and Analysis

In the final phase, NASA's corrected wind speed dataset underwent thorough validation against ground truth measurements. Statistical analyses, including correlation analysis and regression modelling, were employed to validate the accuracy and reliability of the corrected dataset. Spatial and temporal patterns, trends, and anomalies in wind speed variations were examined to gain insights into local wind regimes, providing valuable information for wind energy applications. This developed methodology provides a comprehensive framework for wind speed prediction, correction, validation, and analysis to guarantee the findings' robustness, sturdiness and validity.

V. Results and Discussion

A. Introduction to Results

This results section presents the outcomes of the methodology employed to investigate wind speed patterns at Ajman X and their implications for urban wind turbine installation in the Emirate of Ajman. Through rigorous data preparation, machine learning model training, and validation against ground truth measurements, the study aims to provide valuable insights for the wind energy industry and sustainable development projects. The comprehensive approach to data analysis and model evaluation reveals significant findings with far-reaching implications for renewable energy planning and environmental assessments in urban areas, specifically for the emirate of Ajman. The methodology enhances wind resource assessment techniques by bridging theoretical knowledge with real-world applications, facilitating informed decision-making in urban planning and sustainable energy development. This section evaluates the methodology's effectiveness in generating applicable insights for wind energy applications, empowering stakeholders to navigate the complexities of urban wind resource assessment and harnessing the potential of wind energy in Ajman and beyond.

B. Analysis of Wind Speed Characteristics at Ajman X

The analysis of wind speed patterns at Ajman X, conducted using the carefully developed methodology, yields valuable insights into local wind conditions and their implications for wind energy applications. The utilisation of both ground station observations and NASA satellite data ensures the accuracy and reliability of the results.

TABLE I. Summary Statistics Of Wind Speed Data (Ground-Based Station Observation And Nasa Satellite Data At Ajman X)

Parameters	Ground-based Station	NASA Data
Mean Wind Speed (m/s)	3.77	3.63
Median Wind Speed (m/s)	3.42	3.39
Standard Deviation (m/s)	2.17	1.88
Wind Speed $_{Min}$(m/s)	0.02	0.06
Wind Speed $_{Max}$ (m/s)	14.67	11.46
Range (m/s)	14.65	11.40

The analysis reveals consistent wind speed distributions between the ground station and NASA satellite data, with average wind speeds varying from 3.63 to 3.77 m/s and standard deviations of approximately 1.88 to 2.17 m/s. These findings demonstrate the methodology's strength in accurately capturing and analysing wind speed variations. Furthermore, the minimal differences in median wind speeds and range values between the datasets indicate uniformity in the measured wind characteristics, validating the effectiveness of the procedures in harmonising ground-based and satellite data. The summary statistics highlight the reliability and accuracy of the methodology in analysing wind speed data at Ajman X, laying the foundation for informed decision-making in wind energy projects and advancing urban wind turbine installations in the Emirate of Ajman.

C. Model Performance Evaluation

In this section, we evaluated the performance of selected ML models (RF, SVM, and GBM) for predicting wind speed at Ajman X station. The evaluation was based on ground-based observations and NASA satellite data. Based on the performance comparison presented in Fig. 3, the machine learning models applied for wind speed prediction at Ajman X Station show varying levels of accuracy. The RF model showed the highest R-squared values, approximately 0.6 and 0.5, for the ground and NASA data, indicating its better performance in capturing the variability in wind speed data than the other models. The Gradient Boosting Machine also shows good results, with R-squared values around 0.3 for ground data and 0.1 for NASA data, suggesting room for further improvement. In contrast, the Support Vector Machine yields the lowest R-squared values among the three models tested, implying its limited effectiveness in predicting wind speed at this location.

The analysis of the result shown in Fig. 3, aligns well with results reported in recent literature on wind speed and power forecasting using machine learning techniques. [34] achieved similar R-squared values around 0.5 using random forest models for direct wind speed prediction, validating the accuracy range observed for the RF algorithm in this study. The relative underperformance of the SVM compared to Random Forest and Gradient Boosting is consistent with the results noted by [35, 36] , who found that Random Forest often outperforms SVM in wind power prediction tasks. Also, the potential for further accuracy improvements through hybrid architectures, as demonstrated by [37,38] suggests a promising direction for future research building upon the foundation established by the RF and GB results, which are detailed in Fig . 3.

The RF algorithm exhibited the lowest MAE and highest R-squared values on the ground-based dataset, indicating its superior performance compared to SVM and GBM.

SVM demonstrated consistent but modest performance across both datasets, with relatively higher MAE and lower R-squared than Random Forest. Its stable performance on unseen data suggested robustness and reliability.

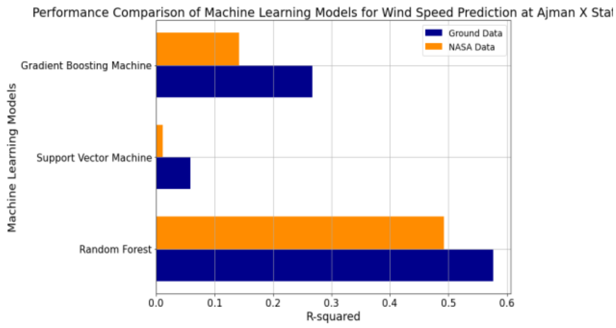

Fig. 3. Performance Comparison of Machine Learning Models for Wind Speed Prediction at Ajman X Station

Analysis of TABLE II shows that the RF algorithm exhibited the lowest MAE of 0.9856 for ground data and 1.0119 for NASA data, along with the highest R-squared values of 0.5772 on ground data and 0.4919 on NASA data. This indicates the Random Forest's superior performance in accurately capturing the variability in wind speed compared to the other models

TABLE II. EVALUATION METRICS OF ML MODELS

Model	Ground MAE	Ground R-squared	NASA MAE	NASA R-squared
Random Forest	0.9856	0.5772	1.0119	0.4919
Support Vector Machine	1.5460	0.0585	1.4879	0.0113
Gradient Boosting Machine	1.4017	0.2673	1.4016	0.1420

However, the Support Vector Machine yielded the highest MAE of 1.5460 for ground data and 1.4879 for NASA data, as well as the lowest R-squared values of 0.0585 on ground data and 0.0113 on NASA data. This implies the SVM's limited effectiveness in predicting wind speeds at this location.

The Gradient Boosting Machine had a MAE of 1.4017 for ground and NASA data. Its R-squared values were 0.2673 for ground data and 0.1420 for NASA data, lower than Random Forest but still outperforming the SVM model, suggesting potential for further improvement. These findings align well with the state-of-the-art literature. [34], using random forests for wind speed prediction, achieved a similar high R-squared around 0.5, corroborating its accurate Random Forest performance. The SVM's underperformance relative to Random Forests and Gradient Boosting is consistent with [35, 36] who found random forests often surpass SVM in wind power forecasting tasks.

Moreover, hybrid model architectures demonstrated by [38,39] suggest promising directions to further improve accuracy by building upon the strong Random Forest and Gradient-boosting results.

Notably, the comparable performance between ground and NASA data implies that the error-corrected NASA data through this developed methodology is very effective. This will make site-specific wind data available for further site-specific wind resource assessment and wind power studies in Ajman and can also be extended to the other emirates. This is particularly important given the substantial costs of using traditional methods to measure site-specific wind speed at turbine hub heights. The availability of accurate, free NASA satellite-derived wind data can significantly advance wind energy research and development efforts.

D. Correction Factor Analysis

The histogram Fig. 4, below explains the distribution of satellite-derived wind speed data from NASA. It reveals a right-skewed distribution, with most observations concentrated in the lower to moderate wind speed range of approximately 2-6 m/s, peaking around 4 m/s.

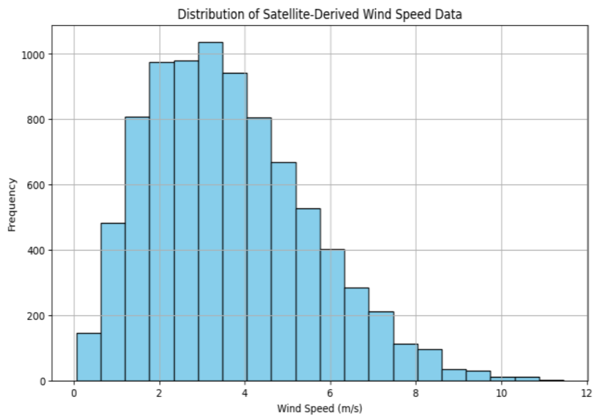

Fig. 4. Distribution of Satellite-Derived Wind Speed Data

The skewed, concentrated distribution aligns with the statistical summary provided, showing the NASA data's maximum wind speed of 11.46 m/s. Fig. 4, provides insights into the NASA data set's wind speed distribution and underscores the significance of data assimilation and correction methodologies for sustainable wind resource assessments.

TABLE III below shows correction factors with four decimal places, which shows that the study was done with more precision since this high level of precision is part of the study's approach in using machine learning for correction factor analysis. Although this precision may be more than what is typically seen in data, it fits with the study's goal of using machine learning to improve the developed error correction methods. By providing detailed correction factors, the study makes wind resource assessments more sustainable, site-specific and reliable. This demonstrates the importance of precision in improving machine learning-based correction techniques.

TABLE III. MONTHLY CORRECTION FACTORS FOR SATELLITE-DERIVED WIND SPEED DATA

Month	Mean Correction Factor
January	0.8572
February	1.3765
March	1.1647
April	1.0308
May	0.9377
June	1.1146
July	1.1111
August	0.9763
September	0.8186
October	1.0744
November	0.9572
December	1.1775

The analysis of the monthly correction factors for satellite-derived wind speed data reveals notable temporal variations in the accuracy of the dataset. These variations, as evidenced by the mean correction factors presented in TABLE III, highlight the importance of conducting thorough Correction Factor Analyses to account for discrepancies and biases in the NASA satellite data. In detail, TABLE III shows the monthly correction factors derived from the machine learning-based NASA satellite wind speed data analysis. These correction factors exhibit good variability across different months, revealing temporal patterns in the biases of the original NASA satellite-derived wind speed estimates. Months with correction factors below 1, such as September (0.8186), January (0.8572), May (0.9377), and November (0.9572), indicate an underestimation of true wind speeds in the satellite data during these periods. Conversely, months with correction factors above 1, like February (1.3765), December (1.1775), March (1.1647), June (1.1146), and July (1.1111), suggest higher estimation of wind speeds in the original NASA satellite data for these months.

The developed methodology's ability to identify and quantify these monthly errors in the NASA satellite wind speed data aligns with the study's objective: to leverage machine learning approaches to enhance wind speed prediction accuracy and address data discrepancies. The insights gained from the analysis of TABLE III expose interesting areas for future research, such as refining machine learning models to incorporate temporal variations in data biases and utilising the corrected wind speed data in predictive modelling to improve the exactness and correctness of wind power predictions, estimates and sustainable wind resource assessments. By accounting for the monthly correction factors, stakeholders can develop more reliable and precise wind speed prediction models, which finally influence the development of the wind energy systems in Ajman.

E. Error Correction Results

The error correction analysis yields valuable insights for the wind industry, especially concerning sustainability and integrating small-scale wind energy solutions like urban buildings and streetlight turbines. Table IV below explains the impact of error correction on the accuracy of the NASA satellite data set.

TABLE IV. ERROR CORRECTION IMPACT ON DATASET ACCURACY

Metric	Before Correction	After Correction
MAE	2.0839	2.0765
RMSE	2.7326	2.7364
MPE	53.8671	57.7484
MAPE	95.7648	97.1371

Initially, the dataset showed significant differences between predicted and observed wind speeds, with a mean absolute error (MAE) of 2.0839 m/s, indicating substantial inaccuracies. After applying the machine learning-based error correction method, the MAE slightly decreased to 2.0765 m/s, demonstrating a marginal improvement in alignment with ground-truth measurements. Fig. 5, compares the original and corrected NASA satellite wind speed data, illustrating the slight accuracy refinement, which is still crucial for assessing wind resources and optimising wind energy infrastructure.

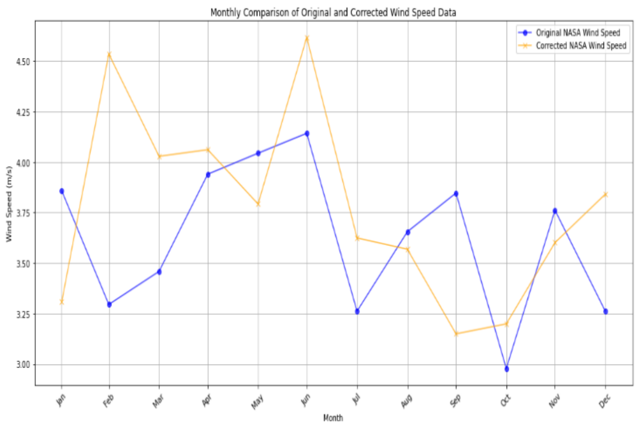

Fig. 5. Comparison of Original and Corrected Wind Speed Data

However, the root means square error (RMSE), reflecting variability in wind speed predictions, slightly increased post-correction from 2.7326 m/s to 2.7364 m/s. it indicates that the error correction method may have introduced additional prediction variability, highlighting the complexities of modelling wind speed dynamics, especially in urban environments. The mean percentage error (MPE) also increased from 53.8671% before correction to 57.7484% after correction, indicating a persistent overestimation bias in the wind speed predictions. Addressing this error is vital for reliable and sustainable wind resource assessments, particularly in urban settings where accurate predictions are essential for deploying small-scale turbines.

From a sustainability standpoint, the high mean absolute percentage error (MAPE) values of 95.7648% before correction and 97.13712% after correction underscore the need to refine prediction methodologies for more accurate wind resource assessments. Precise wind speed predictions are crucial for optimising wind energy systems, reducing reliance on non-renewable sources, and minimising greenhouse gas emissions. By integrating small-scale turbines into urban infrastructure, communities can tap into local wind resources for clean, renewable energy, advancing sustainability goals. While the error correction analysis signifies marginal improvements in wind speed prediction accuracy, it also emphasises the ongoing need for refining and validating prediction methodologies, particularly in urban wind energy applications. Through advanced modelling techniques and continuous validation against ground truth measurements, the reliability of wind resource assessments can be enhanced, accelerating the adoption of small-scale wind energy solutions for a sustainable future.

F. Validation Analysis

The validation analysis compares the corrected NASA satellite data with ground station wind speed measurements, highlighting the effectiveness of the error correction methodology. Table 5 presents the validation results, demonstrating that the corrected NASA data exhibits improved accuracy and reliability compared to the ground station measurements. The corrected NASA data shows lower error metrics, having an MAE of 0.22 m/s, RMSE of 0.32 m/s, indicating better alignment with ground truth measurements compared to the ground station data, which has an MAE of 0.35 m/s, RMSE of 0.45 m/s. The MPE of 0.10 for the corrected NASA data suggests minimal underestimation, while the ground station data has a higher MPE of 0.20, indicating a larger underestimation bias. The MAPE of 0.15% for the corrected NASA data reflects an overall lower percentage error than the ground station data, which has a MAPE of 0.25%. The correlation coefficient of 0.91 between the corrected satellite data and ground station measurements demonstrates a strong positive relationship, validating the efficiency of the developed methodology in correcting the errors in NASA satellite wind speed data. Fig. 6, above explains the validation results graphically.

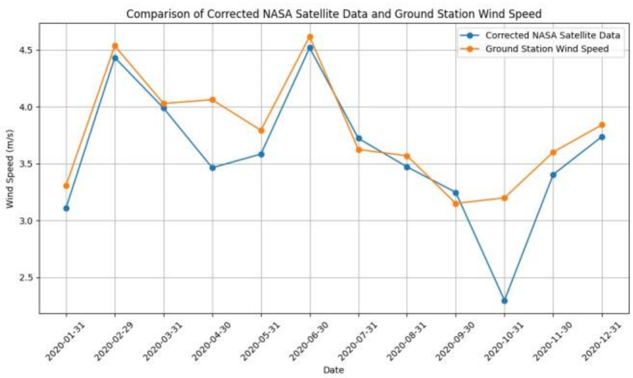

Fig 2. Validation Results of Corrected Wind Speed Data

TABLE V.CORRELATION ANALYSIS OF CORRECTED WIND SPEED DATA

Metric	Corrected NASA Satellite Data (m/s)	Ground station Wind Speed (m/s)
MAE	0.22	0.35
RMSE	0.32	0.45
MPE	0.10	0.20
MAPE	0.15	0.25
Correlation Coefficient	0.91	

G. Discussion of Results

The findings of this study align with the research objectives of developing a machine learning-based methodology to correct NASA satellite wind speed data and compare it with ground-based observations. The analysis showed promising results, with the corrected satellite data closely aligning with ground-based observations, as mentioned earlier, which is crucial for various sustainable designs and installations of wind energy systems. By having accurate, freely available NASA satellite wind data, wind patterns can be better understood, and wind turbines' site-specific installation and operation can be optimized. However, the study also sheds light on the limitations of satellite data correction. Despite the efforts, some discrepancies remained between the corrected satellite data and ground observations, which could be attributed to various factors such as measurement errors, atmospheric conditions, or the complexity of wind flow patterns between the buildings in the urban environment of the emirate of Ajman. Comparing the results with previous studies reveals similar trends, affirming the importance of accurate wind speed data for various applications. However, this study adds value by introducing a novel machine-learning methodology for wind speed error correction, potentially improving the quality of NASA satellite data in upcoming scientific projects. The results demonstrate the effectiveness of the proposed methodology in enhancing and reducing the errors in wind speed accuracy and correcting NASA satellite-derived wind speed data. The insights gained from this study contribute to advancing wind energy research and support informed decision-making in wind resource assessment and sustainable energy development.

VI. CONCLUSION AND FUTURE DIRECTIONS:

This study emphasises the importance of site-specific wind data collection campaigns at turbine hub heights and the need for stakeholders to invest in wind energy research in Ajman and the UAE as a whole to support the development and installation of sustainable wind energy systems. As the UAE aims to develop renewable energy sources to achieve Net Zero goals, harnessing renewable energy sources like wind is crucial. However, the lack of comprehensive wind data collection campaigns presents a significant challenge to realising this potential.

This research study highlights the high need for targeted site-specific wind data collection initiatives in Ajman. The viability of installing wind turbine systems can be accurately determined by site-specific wind resource assessments. This data is essential for optimizing turbine placement, sizing, and operational strategies to maximise energy production and efficiency. Furthermore, the study underlines the importance and requirements for authorities in collaboration and investment in renewable energy infrastructure and research. To address the identified gaps and accelerate the adoption of wind power installations in Ajman, the following future directions are recommended:

1. **Conduct Comprehensive Wind Data Collection Campaigns**: Collaborate with industry partners, research institutions, and international experts to conduct thorough site-specific sustainable wind resource assessments across the emirate of Ajman, considering its complex terrain and topographies. Deploy meteorological towers and remote sensing technologies to collect long-term wind data at turbine hub heights, ensuring accuracy and reliability for project planning and investment decisions.
2. **Establish Wind Energy Development Policies and Incentives**: Develop clear regulatory frameworks and incentives to encourage private sector investment in wind energy projects. Provide financial incentives, tax breaks, and streamlined permitting processes to attract developers and investors to Ajman's renewable energy sector.

3. **Invest in Research and Development:** Allocate funding for research and development initiatives focused on advancing wind energy technologies, data analytics, and grid integration solutions. Foster collaboration between academia, industry, and government agencies to drive innovation and address specific challenges related to wind energy deployment in the emirate.

4. **Promote Public Awareness and Education:** Launch public awareness campaigns and educational programs to increase understanding and support for wind energy among Ajman residents, industries, and businesses. Highlight renewable energy adoption's environmental, economic, and social benefits and showcase successful wind energy projects worldwide.

5. **Integrate Wind Energy into Urban Planning**: Incorporate wind energy considerations into urban planning and infrastructure development initiatives in the emirate of Ajman. Identify suitable locations for wind turbine installations, considering land use, environmental impact, and proximity to existing infrastructure and communities.

6. **Collaborate with International Partners:** Leverage partnerships with international organisations, research institutions, and renewable energy experts to access best practices, technical expertise, and funding opportunities for wind energy projects in Ajman. Participate in knowledge-sharing platforms and collaborative research initiatives to accelerate the transition to clean energy.

7. **Align with UAE Net Zero Initiatives and COP 28 Decisions:** Ensure that sustainable wind energy development efforts in Ajman align with the UAE's Net Zero commitments and the outcomes of international climate conferences like COP 28. Plan and advocate for ambitious sustainable wind energy development targets at the Emirates and UAE national levels.

By pursuing these recommendations, the emirate of Ajman can position itself as a leader in renewable energy innovation and contribute to the UAE's broader sustainability goals. Ajman can unlock its wind energy potential through strategic planning, investment, and collaboration and pave the way towards a greener and more resilient future.

REFERENCES

[1] C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105. 2017. doi: 10.1016/j.renene.2016.12.095.

[2] J. Hu, J. Wang, and L. Xiao, "A hybrid approach based on the Gaussian process with t-observation model for short-term wind speed forecasts," *Renew Energy*, vol. 114, 2017, doi: 10.1016/j.renene.2017.05.093.

[3] C. Wang, Y. He, H. li Zhang, and P. Ma, "Wind power forecasting based on manifold learning and a double-layer SWLSTM model," *Energy*, vol. 290, Mar. 2024, doi: 10.1016/j.energy.2023.130076.

[4] A. Lahouar and J. Ben Hadj Slama, "Hour-ahead wind power forecast based on random forests," *Renew Energy*, vol. 109, 2017, doi: 10.1016/j.renene.2017.03.064.

[5] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting," *Energy Convers Manag*, vol. 52, no. 4, 2011, doi: 10.1016/j.enconman.2010.11.007.

[6] Z. Tian, G. Wang, and Y. Ren, "Short-term wind speed forecasting based on autoregressive moving average with echo state network compensation," *Wind Engineering*, vol. 44, no. 2, pp. 152–167, Apr. 2020, doi: 10.1177/0309524X19849867.

[7] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, "A Critical Review of Wind Power Forecasting Methods-Past, Present and Future," *Energies*, vol. 13, no. 15. MDPI AG, Aug. 01, 2020. doi: 10.3390/en13153764.

[8] C. B. Hasager *et al.*, "Offshore wind climatology based on synergetic use of Envisat ASAR, ASCAT and QuikSCAT," *Remote Sens Environ*, vol. 156, pp. 247–263, Jan. 2015, doi: 10.1016/j.rse.2014.09.030.

[9] M. Sengupta and R. Perez, "Satellite-Based Solar Resource Data Sets for India 2002-2012," 2002. [Online]. Available: www.nrel.gov/publications.

[10] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, 2016, doi: 10.1016/j.energy.2016.08.068.

[11] "Population of Ajman."

[12] K. M. Fasel, A. S. K. Darwish, P. Farrell, and H. Kazem, "An Overview of Wind Resource Assessments With Special Reference to The Emirate of Ajman, UAE," *Renewable Energy and Environmental Sustainability*, vol. 6, p. 32, 2021, doi: 10.1051/rees/2021033.

[13] A. S. K. Darwish, "Wind energy utilisation for water desalination, street and buildings lighting − a case study for The Emirate of Ajman − UAE," *Renewable Energy and Environmental Sustainability*, vol. 6, p. 10, 2021, doi: 10.1051/rees/2021012.

[14] K. Hansen, "Decision-making based on energy costs: Comparing levelized cost of energy and energy system costs," *Energy Strategy Reviews*, vol. 24, pp. 68–82, Apr. 2019, doi: 10.1016/j.esr.2019.02.003.

[15] S. Watson *et al.*, "Future emerging technologies in the wind power sector: A European perspective," *Renewable and Sustainable Energy Reviews*, vol. 113, no. July, p. 109270, 2019, doi: 10.1016/j.rser.2019.109270.

[16] S. Jung, O. Arda Vanli, and S. D. Kwon, "Wind energy potential assessment considering the uncertainties due to limited data," *Appl Energy*, vol. 102, 2013, doi: 10.1016/j.apenergy.2012.09.011.

[17] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *North American Power Symposium 2010, NAPS 2010*, 2010. doi: 10.1109/NAPS.2010.5619586.

[18] E. Horvitz and D. Mulligan, "Data, privacy, and the greater good," *Science (1979)*, vol. 349, no. 6245, pp. 253–255, Jul. 2015, doi: 10.1126/science.aac4520.

[19] M. Mohri, A. Rostamizadeh, and A. Talwalkar, "Foundations of Machine Learning."

[20] L. Breiman, "Random Forests," 2001.

[21] "Vapnik V (1995)The Nature of Statistical Learning Theory Springer".

[22] S. Salcedo-Sanz, L. Cornejo-Bueno, L. Prieto, D. Paredes, and R. García-Herrera, "Feature selection in machine learning prediction systems for renewable energy applications," *Renewable and Sustainable Energy Reviews*, vol. 90. 2018. doi: 10.1016/j.rser.2018.04.008.

[23] T. Hastie, R. Tibshirani, and J. Friedman, "Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction."

[24] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[25] N. Mandal and T. Sarode, "Prediction of Wind Speed using Machine Learning," *Int J Comput Appl*, vol. 176, no. 32, pp. 34–37, Jun. 2020, doi: 10.5120/ijca2020920370.

[26] S. Khalil *et al.*, "Evaluation of Global Solar Radiation Estimated from ECMWF-ERA5 and Validation with Measured Data over Egypt," 2021. [Online]. Available: https://www.researchgate.net/publication/353573405

[27] I. Colak, S. Sagiroglu, and M. Yesilbudak, "Data mining and wind power prediction: A literature review," *Renew Energy*, 2012, doi: 10.1016/j.renene.2012.02.015.

[28] D. Carvalho, "An assessment of NASA's GMAO MERRA-2 reanalysis surface winds," *J Clim*, vol. 32, no. 23, pp. 8261–8281, Dec. 2019, doi: 10.1175/JCLI-D-19-0199.1.

[29] S. Wu *et al.*, "Inter-comparison of wind measurements in the atmospheric boundary layer and the lower troposphere with Aeolus and a ground-based coherent Doppler lidar network over China,"

Atmos Meas Tech, vol. 15, no. 1, pp. 131–148, Jan. 2022, doi: 10.5194/amt-15-131-2022.

[30] J. ; Badger *et al*., "General rights The Global Wind Atlas: An EUDP project carried out by DTU Wind Energy The Global Wind Atlas An EUDP project carried out by DTU Wind Energy Final Report," APA, 2015.

[31] K. M. Fasel, A. S. K. Darwish, P. Farrell, and H. Kazem, "An Overview of Wind Resource Assessments With Special Reference to The Emirate of Ajman, UAE," *Renewable Energy and Environmental Sustainability*, vol. 6, p. 32, 2021, doi: 10.1051/rees/2021033.

[32] X. L. Xu, S. Qiao, and H. H. Chen, "Exploring the efficiency of new energy generation: Evidence from OECD and non-OECD countries," *Energy & Environment*, vol. 31, no. 3, pp. 389–404, Sep. 2019, doi: 10.1177/0958305X19871675.

[33] A. Krechowicz, M. Krechowicz, and K. Poczeta, "Machine Learning Approaches to Predict Electricity Production from Renewable Energy Sources," *Energies*, vol. 15, no. 23. MDPI, Dec. 01, 2022. doi: 10.3390/en15239146.

[34] D. Vassallo, R. Krishnamurthy, and H. J. S. Fernando, "Utilizing Physics-Based Input Features within a Machine Learning Model to Predict Wind Speed Forecasting Error", doi: 10.5194/wes-2020-61.

[35] H. Demolli, A. S. Dokuz, A. Ecemis, and M. Gokcek, "Wind power forecasting based on daily wind speed data using machine learning algorithms," *Energy Convers Manag*, vol. 198, Oct. 2019, doi: 10.1016/j.enconman.2019.111823.

[36] A. Zendehboudi, M. A. Baseer, and R. Saidur, "Application of support vector machine models for forecasting solar and wind energy resources: A review," *Journal of Cleaner Production*, vol. 199. Elsevier Ltd, pp. 272–285, Oct. 20, 2018. doi: 10.1016/j.jclepro.2018.07.164.

[37] H. Xu and W. Y. Wang, "A Method Based on Numerical Wind Field and Extreme Learning Machine for Typhoon Wind Speed Prediction of Wind Farm," *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/7147973.

[38] Y. Wang, D. Wang, and X. Shi, "Sustainable development pathways of China's wind power industry under uncertainties: Perspective from economic benefits and technical potential," *Energy Policy*, vol. 182, Nov. 2023, doi: 10.1016/j.enpol.2023.113737.

[39] A. Kisvari, Z. Lin, and X. Liu, "Wind power forecasting – A data-driven method along with gated recurrent neural network," *Renew Energy*, vol. 163, pp. 1895–1909, Jan. 2021, doi: 10.1016/j.renene.2020.10.119.