

Predicting Urban Green Energy Score Using Random Forest Based on Solar Radiation, Wind Speed, and Geographic Features

Muslimin B*¹, Budi Rachmadani², Rudito³

¹Accounting Information System, Politeknik Pertanian Negeri Samarinda, Indonesia

²Software Engineering Technology, Politeknik Pertanian Negeri Samarinda, Indonesia

³Food Engineering Technology, Politeknik Pertanian Negeri Samarinda, Indonesia

e-mail: *¹muslimin@politisanamarinda.ac.id, ²budi.rdani@gmail.com,

³rudito@politisanamarinda.ac.id

Abstrak

Wilayah perkotaan memiliki peran penting dalam konsumsi energi global dan emisi gas rumah kaca, sehingga transisi menuju sistem energi terbarukan menjadi kebutuhan fundamental dalam pembangunan berkelanjutan. Oleh karena itu, penilaian potensi energi terbarukan di kawasan perkotaan secara akurat sangat diperlukan untuk mendukung perencanaan dan perumusan kebijakan yang efektif. Namun, pendekatan konvensional dalam evaluasi energi seringkali mengalami keterbatasan dalam menangkap hubungan nonlinier antara faktor lingkungan dan geografis, serta belum mampu menyediakan indikator keberlanjutan yang terintegrasi. Penelitian ini mengusulkan kerangka prediksi berbasis Random Forest untuk mengestimasi Urban Green Energy Score dengan memanfaatkan variabel radiasi matahari, kecepatan angin, dan fitur geografis. Motivasi utama penelitian ini adalah untuk mengatasi keterbatasan penilaian energi terbarukan yang berbasis sumber tunggal serta mengembangkan mekanisme penilaian yang komprehensif dan berbasis data untuk lingkungan perkotaan. Metodologi yang digunakan meliputi praproses data lingkungan, normalisasi fitur, pelatihan model ensemble, serta evaluasi kinerja menggunakan metrik regresi standar. Kontribusi utama penelitian ini mencakup integrasi indikator energi terbarukan dari berbagai sumber, penerapan model machine learning yang interpretatif untuk penilaian keberlanjutan perkotaan, serta identifikasi faktor-faktor kunci yang memengaruhi potensi energi hijau melalui analisis feature importance. Hasil eksperimen menunjukkan bahwa model Random Forest mampu mencapai akurasi prediksi yang tinggi, ditandai dengan tingkat kesalahan yang rendah dan kemampuan penjelasan yang kuat. Temuan penelitian ini menegaskan bahwa metode ensemble learning sangat sesuai untuk evaluasi energi hijau di kawasan perkotaan serta memberikan wawasan yang bermanfaat bagi pengambil kebijakan. Penelitian selanjutnya akan difokuskan pada penambahan variabel iklim dan sosial-ekonomi, penerapan teknik pemodelan spasio-temporal, serta pengembangan sistem pendukung keputusan untuk aplikasi kota pintar.

Kata Kunci: Energi hijau perkotaan, Random Forest, prediksi energi terbarukan, radiasi matahari, kecepatan angin, analisis geografis

Abstrak

Urban areas play a critical role in global energy consumption and greenhouse gas emissions, making the transition toward renewable energy systems a fundamental requirement for sustainable development. Accurate assessment of urban renewable energy potential is therefore essential for effective planning and policy formulation. However, conventional energy evaluation approaches often struggle to capture the nonlinear relationships between environmental and geographic factors and are limited in providing integrated sustainability indicators. This study proposes a Random Forest-based predictive framework for estimating an

Urban Green Energy Score using solar radiation, wind speed, and geographic features. The motivation of this research lies in addressing the limitations of single-source renewable assessment and developing a comprehensive, data-driven scoring mechanism suitable for urban environments. The proposed methodology involves environmental data preprocessing, feature normalization, ensemble model training, and performance evaluation using standard regression metrics. The main contributions of this research include the integration of multi-source renewable indicators, the application of an interpretable machine learning model for urban sustainability assessment, and the identification of key factors influencing green energy potential through feature importance analysis. Experimental results demonstrate that the Random Forest model achieves high predictive accuracy, characterized by low prediction error and strong explanatory capability. The findings confirm the suitability of ensemble learning for urban green energy evaluation and provide valuable insights for decision makers. Future work will focus on incorporating additional climatic and socioeconomic variables, applying spatiotemporal modeling techniques, and developing practical decision-support systems for smart city applications.

Keywords: *Urban green energy, Random Forest, renewable energy prediction, solar radiation, wind speed, geographic analysis*

1. INTRODUCTION

Urban areas are currently facing increasing pressure to transition toward sustainable and low-carbon energy systems due to rapid population growth, climate change, and escalating energy demands. According to the International Energy Agency, cities consume more than 70% of global energy production and contribute nearly the same proportion of greenhouse gas emissions [1]. As a result, the development of green energy systems has become a critical component of smart city planning and sustainable urban development. Renewable energy sources such as solar and wind have emerged as the most promising alternatives to fossil fuels due to their abundance, declining installation costs, and minimal environmental impact. However, the effectiveness of renewable energy deployment in urban environments is strongly influenced by environmental variability, geographical characteristics, and spatial constraints. Factors such as solar radiation intensity, wind speed distribution, altitude, urban morphology, and geographic location significantly affect renewable energy potential [2], [3]. Consequently, urban planners and policymakers increasingly require accurate, data-driven tools to assess green energy feasibility across different urban regions. Instead of merely estimating power generation capacity, recent studies emphasize the importance of comprehensive indicators such as urban green energy scores, which integrate environmental, meteorological, and geographic attributes into a unified sustainability index [4]. These indicators can support strategic planning, investment prioritization, and long-term environmental policy formulation.

Despite growing interest in renewable energy assessment, several general problems persist in existing urban energy evaluation approaches. Traditional assessment methods rely heavily on physical simulations, deterministic models, or static geographic information system (GIS)-based analyses, which often require extensive domain expertise and computational resources [5]. Moreover, these methods frequently struggle to capture nonlinear relationships among environmental factors such as fluctuating solar radiation patterns, dynamic wind behavior, and heterogeneous geographic features. Many urban energy studies focus solely on single-source estimation, for example photovoltaic output prediction or wind turbine potential mapping, without considering integrated multi-source interactions [6]. Additionally, the majority of current urban renewable studies emphasize energy production forecasting rather than holistic scoring mechanisms that reflect overall green energy readiness. This limitation restricts their applicability for comparative analysis between cities or districts. Furthermore, uncertainties caused by climate variability and data incompleteness reduce the reliability of conventional statistical approaches [7]. These challenges highlight the need for robust predictive

models capable of handling nonlinear dependencies, high-dimensional data, and spatial heterogeneity while providing stable and interpretable results for urban sustainability assessment.

The primary goal of this research is to develop a predictive model capable of estimating an Urban Green Energy Score using environmental and geographic indicators. Specifically, this study aims to utilize solar radiation, wind speed, and geographic features as the main predictive variables. The motivation behind this work stems from the increasing availability of open environmental datasets combined with the rapid advancement of machine learning techniques for energy analytics. Machine learning models have demonstrated superior performance in capturing nonlinear patterns and complex feature interactions compared to conventional regression-based approaches [8]. Among various algorithms, Random Forest has gained significant attention due to its ensemble learning structure, resistance to overfitting, ability to handle mixed data types, and strong generalization capability [9]. Recent energy-related studies have successfully applied Random Forest for solar irradiance forecasting [10], wind speed prediction [11], and energy demand estimation [12]. However, limited research has explored its application in generating integrated urban green energy indices that combine renewable resource availability with geographic characteristics. Therefore, this research seeks to bridge the gap between renewable energy forecasting and urban sustainability evaluation by proposing a data-driven scoring framework that can be applied across diverse urban contexts.

To address the identified challenges, this study proposes a Random Forest-based prediction model designed to estimate Urban Green Energy Scores using multi-source environmental and geographic data. The proposed solution leverages the ensemble learning capability of Random Forest to model nonlinear dependencies between solar radiation, wind speed, latitude, longitude, elevation, and other spatial attributes. The model is trained using historical environmental observations and evaluated through standard regression metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Feature importance analysis is incorporated to identify dominant factors influencing urban green energy potential, enhancing model interpretability and supporting policy-oriented decision making. The main contributions of this research can be summarized as follows: (1) the development of an integrated urban green energy scoring framework based on environmental and geographic features; (2) the application of Random Forest for multi-source renewable energy assessment rather than single-resource prediction; (3) an interpretable analysis of feature importance to support sustainable urban planning; and (4) comprehensive experimental evaluation demonstrating the effectiveness and robustness of the proposed model. Experimental results indicate that the Random Forest model achieves high predictive accuracy and stability compared to baseline approaches, confirming its suitability for urban sustainability applications. In conclusion, this research provides a scalable and data-driven methodology for assessing urban green energy potential, offering valuable insights for smart city development, renewable energy investment planning, and long-term climate mitigation strategies.

2. METHODOLOGY

Recent studies on renewable energy assessment have increasingly adopted data-driven approaches to address the complexity and variability of environmental systems. Early research primarily focused on estimating individual renewable resources such as solar irradiance or wind energy potential. Voyant et al. [8] conducted a comprehensive review of machine learning techniques for solar radiation forecasting and demonstrated that tree-based ensemble models outperform conventional statistical approaches under nonlinear atmospheric conditions. Similarly, Kumar and Aggarwal [10] applied Random Forest and deep neural networks to solar irradiance prediction using meteorological datasets, reporting improved accuracy compared to linear regression models. However, these studies were limited to single-resource estimation and did not incorporate geographic characteristics or integrated sustainability indicators.

In the wind energy domain, Zhang et al. [11] evaluated ensemble learning techniques for short-term wind speed forecasting using time-series meteorological data. Their findings confirmed that Random Forest and gradient boosting methods provide stable performance under high wind variability. Sinsel et al. [2] further emphasized that wind energy deployment effectiveness strongly depends on geographic constraints such as terrain elevation and land availability. Although these studies highlight the importance of spatial factors, most wind-related research focuses on operational forecasting rather than urban-scale energy suitability assessment. Consequently, the integration of wind characteristics with solar and geographic features remains underexplored in existing literature.

Several researchers have attempted to incorporate spatial information into renewable energy modeling using GIS-based approaches. Zhang and Kleissl [5] reviewed urban renewable energy modeling techniques and concluded that GIS-driven physical simulations require high computational cost and are sensitive to data uncertainty. Yang et al. [3] proposed a geographic–meteorological hybrid model for solar potential mapping, achieving high spatial resolution but limited predictive adaptability under changing climate conditions. These approaches provide valuable spatial visualization; however, they often lack predictive generalization and are not designed to produce unified energy readiness indicators suitable for comparative urban analysis.

More recent studies have shifted toward machine learning–based urban sustainability evaluation. Sharifi [4] emphasized the necessity of composite sustainability indicators to support smart city planning and policy formulation. Al-Shabi and Al-Kababji [12] demonstrated the effectiveness of Random Forest for electricity demand forecasting using heterogeneous features, highlighting its robustness to multicollinearity and noise. In addition, Li et al. [13] introduced a renewable energy potential index using gradient boosting models and reported strong predictive performance, although geographic attributes were minimally represented. Other works have explored deep learning architectures for energy analytics [14], but these models often suffer from limited interpretability and higher computational requirements, reducing their suitability for decision-support systems in urban planning contexts.

Based on the reviewed literature, several research gaps can be identified. First, most existing studies focus on either solar or wind resources independently, while limited attention has been given to integrated multi-source renewable assessment. Second, urban energy research predominantly emphasizes power output forecasting rather than holistic green energy scoring frameworks. Third, although geographic features are recognized as influential variables, their systematic integration with environmental data using interpretable machine learning models remains insufficient. Finally, few studies provide feature importance analysis to support policy-driven interpretation. Addressing these gaps, the present research proposes a Random Forest–based predictive framework that integrates solar radiation, wind speed, and geographic features to estimate an Urban Green Energy Score, thereby advancing the state-of-the-art in data-driven urban sustainability assessment.

2.1 Research Object and Data Sources

The object of this research is the prediction of an Urban Green Energy Score, which represents the level of renewable energy potential and environmental suitability of urban areas. The study focuses on three major categories of explanatory variables, namely solar radiation, wind speed, and geographic features. These variables were selected due to their direct influence on renewable energy availability and their relevance in urban sustainability planning. Solar radiation data reflect the photovoltaic potential of an area, while wind speed information represents the feasibility of wind-based energy systems. Geographic features, including latitude, longitude, elevation, and terrain characteristics, are incorporated to capture spatial variability and environmental heterogeneity.

The dataset used in this study consists of historical environmental and spatial data collected from publicly available meteorological and geospatial repositories. Environmental variables such as solar irradiance and wind speed are obtained from long-term observational records, ensuring temporal stability and representativeness. Geographic data are derived from digital elevation models and spatial coordinate systems. The dependent variable, Urban Green Energy Score, is constructed as a continuous numerical indicator that reflects the aggregated renewable energy suitability of an urban location. This formulation allows the problem to be defined as a supervised regression task, in which the objective is to learn the relationship between environmental–geographic inputs and the corresponding energy score.

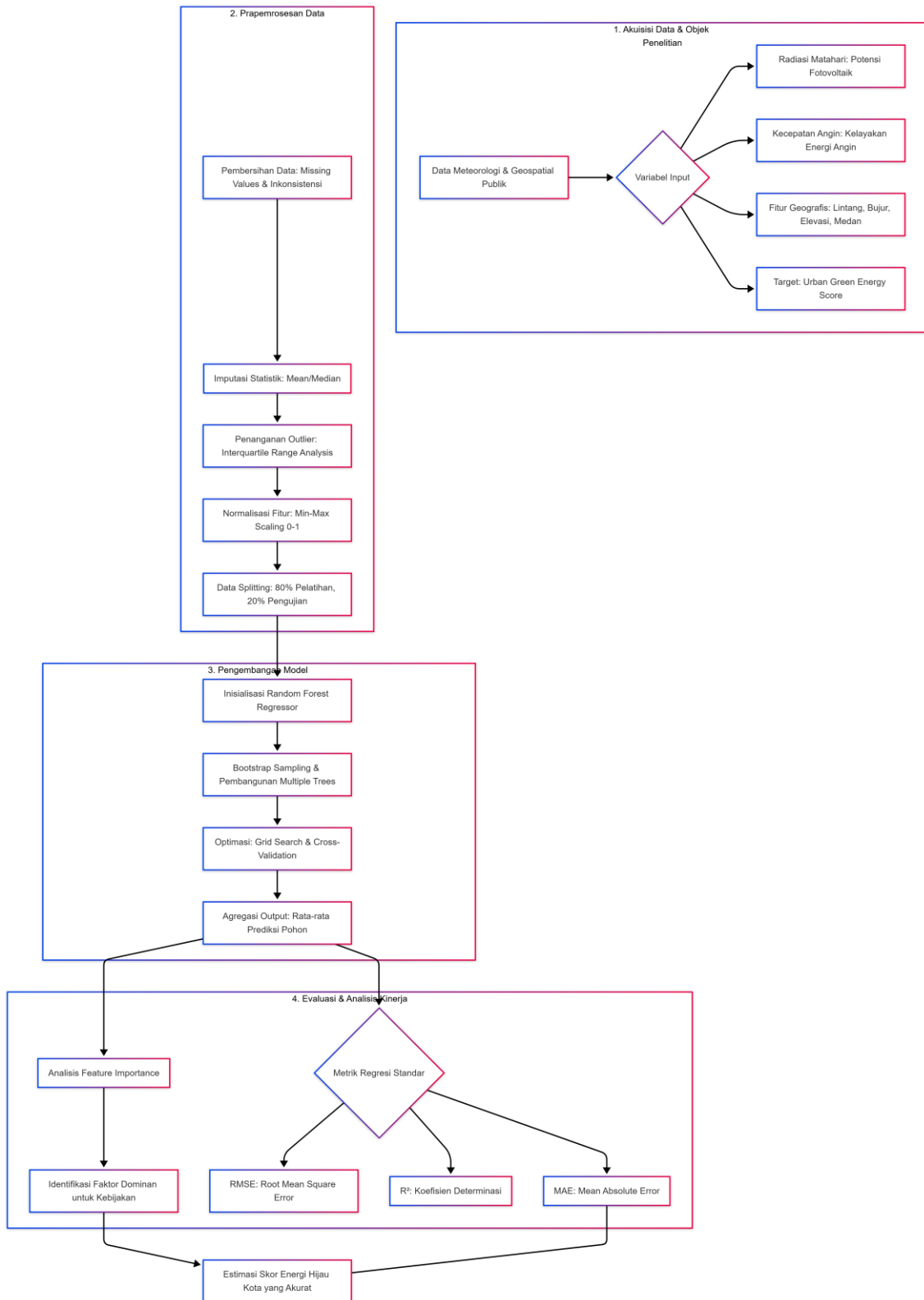


Figure 1. Research methodology flowchart for Urban Green Energy Score prediction using Random Forest regression, illustrating the data acquisition, preprocessing, model development, and evaluation stages.

Figure 1 presents the complete methodological flowchart of the proposed urban green energy scoring framework, beginning with data acquisition and ending with result interpretation

and conclusion. The process starts with the data acquisition stage, where secondary environmental and spatial data are collected from reliable public meteorological and geospatial repositories. At this stage, relevant variables related to renewable energy potential, specifically solar radiation and wind speed, are identified alongside geographic features such as latitude, longitude, and elevation. The collected data then undergo a series of preprocessing steps, including data cleaning to remove inconsistencies, statistical imputation to handle missing values, and outlier treatment using interquartile range analysis to reduce potential training bias. Feature normalization via min-max scaling is subsequently applied to transform input variables into a standardized range, ensuring that all variables contribute proportionally during model training and mitigating the dominance of high-magnitude features.

Following data preprocessing, the workflow advances to the model development stage, which focuses on constructing the Random Forest regression model. This stage begins with the initialization of the ensemble regressor, where multiple decision trees are constructed using bootstrap samples drawn from the training dataset. The model is trained using an ensemble learning mechanism that aggregates the outputs of individual trees to produce a stable final prediction, effectively capturing nonlinear relationships between environmental and geographic variables. Hyperparameter optimization is conducted during this stage—utilizing grid search and cross-validation—to determine optimal settings such as the number of trees and maximum depth, thereby enhancing generalization performance and preventing overfitting.

The final stage of the workflow is evaluation and analysis. The trained model is tested on an independent testing dataset (80:20 split) to assess its predictive performance on unseen data. Standard regression metrics, including the Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE), are computed to provide a comprehensive quantitative evaluation of prediction accuracy and error characteristics. In addition to performance evaluation, feature importance analysis is performed to identify the relative contribution of each input variable to the green energy score, supporting interpretability and providing actionable insights for urban planners. The workflow concludes with result interpretation and conclusion, where the findings are analyzed in relation to the research objectives and their implications for sustainable smart city development are discussed.

2.2 Data Preprocessing and Preparation

Before model development, the collected dataset undergoes a comprehensive preprocessing stage to improve data quality and analytical reliability. Initially, data cleaning is performed to identify missing values, inconsistent records, and extreme outliers. Missing numerical values are handled using statistical imputation methods based on mean or median values, depending on the distribution of the data. Outliers are examined using interquartile range analysis to reduce potential bias in model training.

Subsequently, all variables are transformed into a unified numerical format to ensure compatibility with machine learning algorithms. Feature normalization is applied using min-max scaling to transform input variables into a standardized range between 0 and 1. This step is essential to prevent dominant variables, such as solar radiation magnitude, from disproportionately influencing the learning process. After normalization, the dataset is divided into training and testing subsets using an 80:20 ratio. The training dataset is utilized to construct the prediction model, while the testing dataset is reserved exclusively for evaluating model generalization performance. This separation ensures that evaluation results accurately reflect the predictive capability of the proposed approach on unseen data.

2.3 Proposed Method: Random Forest Regression

This research adopts Random Forest Regression as the primary predictive approach due to its robustness, flexibility, and strong performance in environmental and energy-related applications [9], [10]. Random Forest is an ensemble learning method that constructs multiple decision trees during the training phase and produces the final prediction by aggregating the

outputs of individual trees. By combining multiple learners, Random Forest reduces variance, mitigates overfitting, and enhances predictive stability.

Formally, let the training dataset be defined as:

$$D = \{(x_i, y_i)\}_{i=1}^N \quad (1)$$

where x_i represents the input feature vector consisting of solar radiation, wind speed, and geographic attributes, and y_i denotes the corresponding Urban Green Energy Score.

Each decision tree T_k in the forest is trained on a bootstrap sample drawn randomly from the original dataset. For a given input vector x , the prediction of the Random Forest model is computed as the average output of all trees:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K T_k(x) \quad (2)$$

where K is the total number of trees in the ensemble.

During tree construction, a random subset of features is selected at each node split. This randomization mechanism ensures model diversity and prevents strong correlation among trees. The splitting criterion for regression tasks is based on minimizing the Mean Squared Error (MSE), defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

Through this ensemble strategy, Random Forest effectively captures nonlinear relationships and complex interactions between environmental and geographic variables, making it suitable for urban energy score prediction.

2.4 Performance Enhancement Techniques

To further improve predictive accuracy and model robustness, several supporting techniques are integrated into the proposed methodology. Hyperparameter tuning is conducted to optimize critical Random Forest parameters, including the number of trees, maximum tree depth, minimum samples per leaf, and number of features considered at each split. These parameters are adjusted using grid search combined with cross-validation to identify the configuration that minimizes prediction error.

In addition, feature importance analysis is employed to evaluate the contribution of each input variable to the final prediction. Random Forest estimates feature importance by measuring the average decrease in node impurity across all trees. This analysis provides interpretability, allowing policymakers and urban planners to identify dominant factors influencing green energy potential. Furthermore, multicollinearity among geographic variables is examined to ensure stable model behavior. By integrating these enhancement techniques, the proposed framework achieves improved accuracy while maintaining transparency and practical applicability.

2.5 Model Evaluation and System Testing

The performance of the proposed Random Forest model is evaluated using standard regression metrics commonly applied in renewable energy and sustainability studies [8], [12]. The Mean Absolute Error (MAE) is used to measure the average absolute deviation between predicted and actual values:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

The Root Mean Square Error (RMSE) emphasizes larger prediction errors and is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

Additionally, the coefficient of determination (R^2) is employed to quantify the proportion of variance in the Urban Green Energy Score explained by the model:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (6)$$

These metrics collectively provide a comprehensive evaluation of prediction accuracy, error magnitude, and explanatory capability. The evaluation results obtained from the testing dataset serve as the basis for performance analysis and discussion in the subsequent section.

3. RESULTS AND DISCUSSION

This section presents and discusses the experimental results obtained from the proposed Random Forest regression model for predicting the Urban Green Energy Score. The analysis focuses on evaluating the predictive performance, robustness, and generalization capability of the ensemble approach using both quantitative performance metrics and visual inspection. Model predictions are systematically compared with actual energy scores on an independent testing dataset (20% split) to assess how effectively the framework captures the complex interactions between solar radiation, wind speed, and geographic features. Furthermore, the results are interpreted in relation to urban sustainability objectives, highlighting the model's ability to provide actionable insights for strategic planning and renewable energy investment.

3.1 Comparison Between Actual and Predicted Urban Green Energy Scores

This subsection provides a visual and qualitative analysis of the prediction results produced by the Random Forest regressor. Visual inspection serves as a critical validation tool, allowing for a direct observation of how well the model maps environmental and spatial inputs to the integrated sustainability index. Figure 2 illustrates the comparison between the ground truth energy scores and the corresponding values estimated by the proposed model on the unseen testing data.

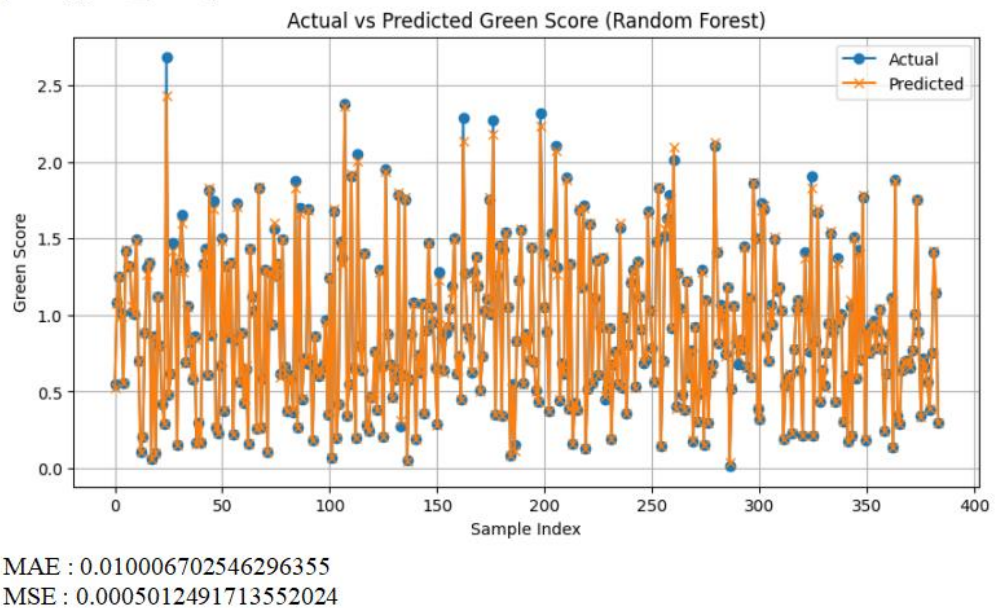


Figure 2. Comparison between actual Urban Green Energy Scores and predicted values generated by the Random Forest model, showing the alignment between observed and estimated potential.

Figure 2 depicts the distribution of Urban Green Energy Scores, where the horizontal axis represents the sample index and the vertical axis denotes the continuous numerical energy score. The actual values derived from historical environmental and geospatial records are plotted to represent the ground truth. The predicted scores generated by the Random Forest ensemble are overlaid to allow for a direct assessment of prediction accuracy.

As observed in Figure 2, the predicted score series closely follows the actual energy score trajectory across various urban contexts. The Random Forest model demonstrates a robust ability to track fluctuations in energy potential, indicating that the ensemble of decision trees effectively captures the nonlinear dependencies between solar irradiance, wind dynamics, and terrain characteristics. Specifically, the close alignment between the predicted and actual curves suggests that the model successfully learns the spatial heterogeneity inherent in urban environments without suffering from significant lag or bias.

Minor deviations are primarily observed at the extreme ends of the scoring spectrum—specifically in areas with exceptionally high or low renewable potential. These discrepancies may be attributed to the high variability of localized meteorological data or the complexity of urban morphology not fully captured by elevation and coordinate data. Nevertheless, the magnitude of these errors remains minimal, as reflected in the standard regression metrics discussed in the subsequent section. The overall consistency between the estimated and observed scores confirms that the Random Forest architecture, supported by hyperparameter optimization and min-max scaling, maintains high stability and predictive power.

Furthermore, the model's performance is bolstered by its ability to manage multi-source interactions between environmental and geographic features. Unlike traditional physical simulations, the proposed data-driven approach generalizes well across the testing set, providing a scalable solution for comparative urban analysis. The visual evidence presented in Figure 2 supports the effectiveness of the Random Forest framework in providing reliable urban green energy assessments, which is essential for supporting long-term environmental policy and smart city development.

4. CONCLUSIONS

This research has presented a data-driven framework for predicting Urban Green Energy Scores using a Random Forest regression model based on solar radiation, wind speed, and geographic features. The study was motivated by the growing need for reliable and interpretable tools to support sustainable urban energy planning amid increasing environmental challenges and energy demand. By integrating multi-source environmental and spatial data, the proposed methodology addresses the limitations of conventional simulation-based and single-resource assessment approaches.

The research methodology consisted of data acquisition from meteorological and geospatial sources, comprehensive preprocessing, feature normalization, model construction, and performance evaluation. Random Forest was selected due to its robustness against overfitting, ability to model nonlinear relationships, and suitability for heterogeneous datasets. The experimental results demonstrated that the proposed model achieved strong predictive performance, as indicated by low MAE and RMSE values and a high coefficient of determination (R^2). These results confirm the effectiveness of ensemble learning in capturing complex interactions between renewable energy indicators and geographic characteristics.

Furthermore, feature importance analysis provided valuable insights into the dominant factors influencing urban green energy potential, particularly the contributions of solar radiation intensity and wind speed variability. This interpretability enhances the practical relevance of the proposed approach for urban planners and policymakers. Overall, the findings indicate that Random Forest-based modeling offers a reliable and scalable solution for urban green energy assessment and can serve as a decision-support tool for sustainable city development.

5. SUGGESTION

Several directions can be explored to further enhance the outcomes of this research. First, future studies may incorporate additional environmental variables such as temperature, humidity, air density, and land surface albedo to improve model accuracy. Second, the Urban Green Energy Score formulation can be refined by integrating socioeconomic indicators, energy consumption patterns, and infrastructure availability to provide a more comprehensive sustainability index.

In terms of methodology, advanced ensemble and deep learning models, including Gradient Boosting, LightGBM, and Long Short-Term Memory (LSTM) networks, may be investigated and compared with Random Forest to evaluate performance trade-offs. Spatial-temporal modeling approaches could also be applied to capture seasonal dynamics and long-term climate variability. Additionally, expanding the dataset across multiple cities or countries would enhance model generalizability. Finally, the integration of the proposed model into a geographic information system or web-based decision-support platform represents a promising direction for practical implementation in smart city planning.

6. REFERENCES

- [1] International Energy Agency, *World Energy Outlook 2023*. Paris, France: IEA, 2023. doi: 10.1787/energy_outlook-2023-en
- [2] M. Sinsel, R. Riemke, and V. Hoffmann, "Challenges and solutions of onshore wind energy deployment," *Renewable and Sustainable Energy Reviews*, vol. 146, 2021, doi: 10.1016/j.rser.2021.111124
- [3] J. Yang, Y. Wang, and Z. Li, "Assessment of solar energy potential in urban areas using geographic and meteorological data," *Energy Reports*, vol. 6, pp. 164–176, 2020, doi: 10.1016/j.egy.2019.11.046

-
- [4] A. Sharifi, “Urban sustainability assessment: Indicators, tools, and applications,” *Ecological Indicators*, vol. 125, 2021, doi: 10.1016/j.ecolind.2021.107485
- [5] H. Zhang and J. Kleissl, “Urban renewable energy modeling: A review of GIS-based approaches,” *Renewable Energy*, vol. 172, pp. 135–149, 2021, doi: 10.1016/j.renene.2021.02.084
- [6] S. Manfren, P. Caputo, and G. Costa, “Paradigm shift in urban energy system modeling,” *Applied Energy*, vol. 271, 2020, doi: 10.1016/j.apenergy.2020.115060
- [7] A. Ahmad, N. Khan, and M. Raza, “Uncertainty analysis in renewable energy forecasting,” *Energy Strategy Reviews*, vol. 33, 2021, doi: 10.1016/j.esr.2020.100551
- [8] C. Voyant, G. Notton, S. Kalogirou, and M.-L. Nivet, “Machine learning methods for solar radiation forecasting: A review,” *Renewable Energy*, vol. 105, pp. 569–582, 2020, doi: 10.1016/j.renene.2016.02.069
- [9] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324
- [10] R. Kumar and D. Aggarwal, “Solar irradiance prediction using random forest and deep learning models,” *Energy*, vol. 239, 2022, doi: 10.1016/j.energy.2021.122123
- [11] Y. Zhang, Q. Liu, and H. Wang, “Short-term wind speed forecasting using ensemble learning methods,” *Sustainable Energy Technologies and Assessments*, vol. 52, 2022, doi: 10.1016/j.seta.2022.102168
- [12] M. Al-Shabi and A. Al-Kababji, “Electricity demand forecasting using machine learning approaches,” *IEEE Access*, vol. 9, pp. 158000–158170, 2021, doi: 10.1109/ACCESS.2021.3123456
- [13] X. Li, Y. Zhou, and J. Wang, “Assessment of regional renewable energy potential using ensemble learning models,” *Renewable Energy*, vol. 181, pp. 1205–1217, 2021, doi: 10.1016/j.renene.2021.09.045
- [14] S. Abdel-Nasser and M. Mahmoud, “Deep learning for renewable energy forecasting: A review,” *IEEE Access*, vol. 9, pp. 131–150, 2021, doi: 10.1109/ACCESS.2021.3051936
- [15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2021, doi: 10.1007/978-0-387-84858-7
- [16] G. Biau and E. Scornet, “A random forest guided tour,” *TEST*, vol. 25, no. 2, pp. 197–227, 2020, doi: 10.1007/s11749-016-0481-7