International Transactions on Artificial Intelligence (ITALIC)
Vol. 4, No. 2, May, 2026, pp. 159−170
E-ISSN: 2963-1939 | P-ISSN: 2963-6086, DOI:10.33050

Reliable Machine Learning Models for Energy Optimization in Smart Green Cities

Ignatius Agus Supriyono[1], Mochamad Heru Riza Chakim[2], Henry Zainarthur[3], Dimas Aditya Prabowo[4*]
[1]Faculty of Computer Science, Satya Wacana Christian University, Indonesia
[2]Faculty of Economics and Business, Universitas Raharja, Indonesia
[3]Department of Information System, Eduaward Incorporation, United Kingdom
[4]Faculty of Economics and Business, Universitas Raharja, Indonesia
[1]982022014@student.uksw.edu, [2]heru.riza@raharja.info, [3]henry.zthur7@eduaward.co.uk, [4]dimas.prabowo@Raharja.info
***Corresponding Author**

Article Info

Article history:

Submission February 23, 2026
Revised March 11, 2026
Accepted May 2, 2026
Published May 14, 2026

Keywords:

Machine Learning
Energy Optimization
Smart Green Cities
Reliability Analysis
Sustainable Energy Management

ABSTRACT

Rapid urbanization and increasing energy consumption have intensified the need for intelligent approaches that support sustainable and efficient energy management in smart green cities. **This study** investigates the effectiveness of machine learning models in improving energy demand forecasting and energy optimization through a reliability-oriented evaluation framework. **The research** utilizes a real-world smart city energy consumption dataset comprising 17,520 hourly observations collected between January 2022 and December 2023. Three machine learning models, namely Random Forest, Support Vector Machine (SVM), and Long Short-Term Memory (LSTM), were developed and evaluated using 30 independent execution runs. Model performance was assessed through Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination (R^2), reliability analysis, and interpretability consistency measurements. **The results** demonstrate that LSTM achieved the best predictive performance with an MAE of 0.31, RMSE of 0.45, and R^2 of 0.93, outperforming Random Forest and SVM across all evaluation metrics. Furthermore, LSTM exhibited the highest reliability score of 0.912 and superior explanation stability, indicating robust and consistent performance under repeated executions. The forecasting outputs were integrated into an energy optimization framework, resulting in reductions in peak energy loads and overall electricity consumption. **These findings** confirm that reliable and explainable machine learning models can support adaptive, data-driven energy management strategies capable of enhancing operational efficiency and sustainability in urban environments. The proposed framework contributes to the development of trustworthy intelligent systems for smart green cities and supports the achievement of sustainable development objectives related to clean energy, sustainable communities, and climate action.

DOI: https://doi.org/10.33050/italic.v4i2.1092

1. INTRODUCTION

The rapid expansion of urban populations and the continuous growth of industrial and residential activities have caused a dramatic increase in global energy consumption, making cities one of the largest contributors to carbon emissions and environmental degradation [1]. Urban areas require large amounts of electricity for transportation, housing, industrial activity, public facilities, and digital infrastructure [2]. This condition places strong pressure on existing energy systems and increases the urgency of developing intelligent, sustainable, and adaptive urban energy management approaches. This global challenge is closely aligned with the United Nations Sustainable Development Goals (SDGs), particularly SDG 7 on affordable and clean energy, SDG 11 on sustainable cities and communities, and SDG 13 on climate action [3]. These goals emphasize the importance of clean, efficient, and resilient urban energy systems that can respond to the dynamic demands of growing populations. Traditional energy management systems in many cities are often static, reactive, and inefficient because they rely heavily on historical averages rather than real-time data to regulate energy distribution and consumption [4]. As a result, conventional approaches often struggle to adapt to dynamic urban environments where energy demand changes rapidly due to weather conditions, population behavior, mobility patterns, and economic activity. In response to these challenges, the concept of smart green cities has emerged as an integrated urban development model that combines digital technologies, renewable energy, sustainability principles, and intelligent infrastructure [5]. Smart green cities aim to reduce environmental impacts while improving quality of life and operational efficiency. However, many smart city initiatives still face important limitations in predicting energy demand accurately and optimizing energy usage in a scalable and adaptive manner. Machine learning has recently gained significant attention as a transformative tool capable of analyzing large-scale urban data and generating predictive insights that support intelligent decision-making [6]. By learning complex patterns from historical and real-time data, machine learning models can forecast energy consumption, detect anomalies, and optimize energy distribution with higher accuracy than rule-based or traditional statistical approaches. Various studies have demonstrated the potential of machine learning techniques such as Random Forest, SVM, and deep learning architectures in energy demand forecasting, load balancing, and energy efficiency optimization [7].

Nevertheless, existing research often focuses on isolated components of urban energy systems, such as individual buildings, microgrids, or specific renewable energy sources. Many models are also developed using limited datasets or evaluated mainly based on prediction accuracy, without sufficient attention to reliability, robustness, reproducibility, and interpretability [8]. In this study, three constructs are explicitly distinguished to avoid terminological ambiguity: reliability refers to the consistency of model performance metrics across repeated independent executions under identical conditions; robustness refers to the model's capacity to maintain acceptable performance under data variation, noise, or distributional shift; and reproducibility refers to the ability of independent researchers to obtain comparable results by following the same dataset preparation, model configuration, and evaluation procedures. In the context of smart green cities, high prediction accuracy alone is not sufficient because urban energy decisions affect infrastructure planning, operational costs, environmental sustainability, and public policy [9]. Therefore, machine learning models for urban energy management must be accurate, stable, explainable, and reproducible. Although previous studies have examined machine learning for energy forecasting, many of them emphasize predictive accuracy without explicitly addressing reliability, reproducibility, and interpretability as essential requirements for AI-based decision-support systems [10].

This study is motivated by a fundamental problem: existing machine learning evaluations for urban energy management predominantly rely on single-run accuracy metrics derived from limited datasets, without sufficiently addressing whether models produce stable, reproducible, and interpretable outputs under repeated execution conditions. This limitation is consequential because urban energy decisions directly affect infrastructure planning, public resource allocation, and long-term sustainability commitments. Without reliable and consistent model behavior, deploying machine learning in high-impact smart city governance contexts introduces operational and accountability risks that undermine institutional trust and policy credibility. This study addresses this problem by evaluating machine learning models through a reliability-oriented experimental design comprising 30 independent execution runs per model applied to a real-world smart city energy dataset of 17,520 hourly observations spanning January 2022 to December 2023. The main contributions of this study are threefold: it develops a scalable machine learning framework for smart green city energy forecasting [11]; it compares the performance of Random Forest, SVM, and LSTM using real-world urban energy datasets under identical experimental conditions; and it integrates reliability, reproducibility, preprocessing transparency, and interpretability discussion into the evaluation of AI-based energy decision support [12, 13].

2. LITERATURE REVIEW

Machine learning has become a foundational technology in urban energy management due to its capacity to process large-scale, high-dimensional, and nonlinear data that conventional statistical models such as ARIMA and linear regression cannot adequately capture [14, 15]. Algorithms including Random Forest, Gradient Boosting, Support Vector Machine, and deep neural networks have demonstrated strong performance in both short-term and long-term urban energy demand forecasting [16]. These models learn from smart meters, IoT sensors, weather monitoring platforms, and historical consumption records, delivering predictive insights that support load balancing, peak demand management, and more efficient resource allocation for grid operators, energy providers, and policymakers [17, 18]. Within smart green city frameworks, energy efficiency is a central pillar that directly reduces carbon emissions, lowers operational costs, and strengthens urban infrastructure resilience [19]. Smart green cities rely on real-time monitoring and data-driven systems to respond dynamically to demand fluctuations driven by temperature extremes, working-hour cycles, and mobility patterns [9]. Machine learning transforms raw urban data into actionable forecasting outputs that support energy scheduling, demand-side management, and sustainability planning aligned with SDG 11 and SDG 13 [20]. Among deep learning architectures, Long Short-Term Memory (LSTM) networks are particularly well-suited to urban energy forecasting because they capture long-range temporal dependencies in sequential consumption series [21, 22]. Random Forest offers robust nonlinear pattern recognition with resistance to overfitting, while Support Vector Machine provides effective kernel-based regression in high-dimensional input spaces [23, 24]. Nonetheless, model selection must not be grounded solely in prediction accuracy. Recent studies establish that machine learning systems deployed in high-impact contexts must additionally satisfy reliability, robustness, and reproducibility requirements [25].

2.1. Reliability-Oriented Evaluation, Research Gaps, and Positioning

Recent reliability and stability-focused AI evaluation studies demonstrate that robust model assessment must incorporate repeated execution testing, explanation consistency analysis, and transparent methodological reporting [26, 27]. Explanation stability research further establishes that interpretable AI outputs must remain consistent across retraining or repeated evaluation, as attribution variability directly undermines governance oversight and institutional accountability [28]. Despite these advances, existing reliability-oriented frameworks remain situated in general AI or explainable AI contexts and have not been systematically applied to smart green city energy forecasting. Table 1 positions the proposed framework against these two categories of prior work to delineate the novelty of the present study.

Table 1. Comparison with Recent Reliability and Stability-Focused AI Evaluation Studies

Study Focus	**Main Evaluation Emphasis**	**Difference from the Proposed Framework**
Reliability-focused AI evaluation [26, 27]	Repeated execution, performance consistency, robustness, and reproducibility	Focus on general AI reliability without integrating energy demand forecasting, smart green city decision support, or governance alignment
Explanation stability studies [28]	Stability of interpretability outputs and consistency of model explanations	Emphasize attribution consistency in general AI contexts but do not connect explanation stability with urban energy optimization or smart city governance
Proposed framework	Predictive accuracy, 30-run reliability quantification, preprocessing transparency, interpretability consistency, and governance alignment	Integrates all dimensions within one applied framework specifically designed for smart green city energy efficiency optimization

As shown in Table 1, the proposed framework advances prior work by unifying forecasting performance and reliability-oriented assessment within a single applied framework targeting smart green city contexts. Several research gaps motivate this contribution. Many existing studies focus on isolated subsystems such as individual buildings or microgrids rather than city-wide networks, train models on static datasets that limit real-time adaptability, and report accuracy metrics without documenting repeated execution settings, normalization techniques, or interpretability stability [29, 30]. Limited attention has also been given to integrating machine learning outputs with governance and policy decision-making systems, reducing both reproducibility and practical value [31]. The present study addresses these gaps through a unified, scalable, and reliability-oriented framework grounded in transparent experimental design and structured governance alignment.

3. RESEARCH METHODOLOGY

Figure 1 illustrates the overall methodological framework of this study, which proceeds through four sequential stages: data acquisition and preprocessing, machine learning model development, repeated-execution evaluation, and governance-oriented decision-support assessment. Each stage feeds directly into the next, ensuring that forecasting quality is validated before outputs are integrated into the energy optimization decision layer.

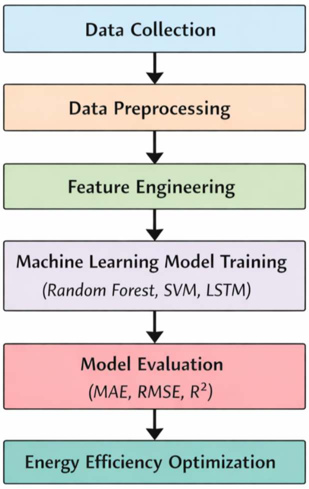

Figure 1. Research Methodology Framework: Sequential Flow from Data Acquisition to Energy Optimization Decision Support

As illustrated in Figure 1, the preprocessing stage determines the quality and comparability of model inputs; the model training and evaluation stage generates and validates predictive outputs through a 30-run repeated-execution protocol; and the decision-support stage connects forecasting results to urban energy policy requirements through structured governance alignment. This directional structure ensures the proposed framework is not a purely academic exercise but is designed to deliver actionable and trustworthy outputs for smart green city implementation [32, 33].

3.1. Dataset, Preprocessing, and Experimental Design

The dataset comprises 17,520 real-world hourly electricity consumption observations collected from publicly available smart city energy repositories and smart meter systems, spanning January 2022 to December 2023 [34]. Input features include hourly electricity consumption (kWh) as the target variable, weather variables (ambient temperature in °C and relative humidity in %), and temporal features (hour of day, day of week, and a binary weekday or weekend indicator). To preserve temporal ordering, the dataset was partitioned chronologically into 80% training (14,016 observations) and 20% testing (3,504 observations) without shuffling. These three data source types were selected because they represent diverse and complementary decision-support domains in smart green city energy management: smart-meter records capture operational consumption patterns, weather data reflect environmental drivers of heating and cooling demand, and temporal features encode behavioral and scheduling patterns [35]. Table 2 summarizes the key dataset attributes.

Table 2. Dataset Characteristics and Preprocessing Summary

Attribute	Description
Time Span	January 2022 to December 2023
Total Records	17,520 hourly observations

Training Partition	14,016 observations (80%, chronological)
Testing Partition	3,504 observations (20%, chronological)
Target Variable	Hourly electricity consumption (kWh)
Input Features	Temperature (°C), humidity (%), hour, day of week, weekday/weekend, 24-hr rolling mean
Missing Values	<0.5%; treated by linear interpolation
Outlier Handling	IQR-based winsorization at 1st and 99th percentiles
Normalization	Min-Max scaling to $[0, 1]$ for all continuous features
Imbalance Handling	Demand-quantile analysis; sample weighting applied
Feature Engineering	Sine-cosine cyclical encoding of hour and day; 24-hr rolling mean consumption

As shown in Table 2, preprocessing was standardized across all models to ensure comparability. Missing values were handled by linear interpolation to preserve temporal continuity; outliers were winsorized rather than discarded to retain valid extreme observations; continuous features were Min-Max normalized to $[0, 1]$, which is essential for gradient-sensitive models such as LSTM and SVM; and cyclical temporal features were encoded using sine-cosine transformations to avoid ordinal misrepresentation [36]. A 24-hour rolling mean was computed as an additional feature to capture short-term consumption momentum.

To strengthen methodological reliability and reproducibility, each model was executed 30 independent times using fixed random seeds from 1 to 30, with final results reported as averages across all repetitions [37]. This repeated-execution design isolates performance variability attributable to stochastic initialization rather than data or configuration differences. Consistent with the definitions introduced in Section 2, reliability is quantified as the normalized inverse coefficient of variation of MAE, RMSE, and R^2 across the 30 runs; robustness reflects performance stability under data variation and preprocessing conditions; and reproducibility is ensured by full documentation of all procedures described in this section [38].

3.2. Model Configurations and Evaluation Metrics

Three machine learning models representing distinct algorithmic paradigms were implemented and compared to provide a balanced evaluation across conventional and deep learning approaches. Random Forest (RF) employs an ensemble of 200 decision trees (maximum depth 10, minimum samples per split 5), offering robust nonlinear pattern recognition and resistance to overfitting through bagging. SVM uses an RBF kernel ($C = 10$, $\varepsilon = 0.1$) for kernel-based regression in high-dimensional feature spaces. LSTM networks employ two stacked layers (64 and 32 units), dropout regularization at 0.2, and are trained for 100 epochs using the Adam optimizer (learning rate $= 0.001$), making them architecturally suited to capturing long-range temporal dependencies in sequential hourly consumption data. Hyperparameters for all models were determined through five-fold cross-validation on the training partition prior to final training. All three models were equipped with SHAP post-hoc explainability modules to enable interpretability consistency analysis across repeated runs. Table 3 summarizes the configurations.

Table 3. Machine Learning Model Configurations

Model	Algorithmic Rationale	Key Hyperparameters
Random Forest	Ensemble bagging for nonlinear demand prediction; robust to overfitting	200 trees; max depth 10; min samples split 5
Support Vector Machine	Kernel-based regression for high-dimensional feature spaces	RBF kernel; $C = 10$; $\varepsilon = 0.1$
Long Short-Term Memory	Recurrent architecture for sequential temporal dependency modeling	2 layers (64, 32 units); dropout 0.2; Adam lr $= 0.001$; 100 epochs

As presented in Table 3, the three models collectively span tree-based ensemble, kernel-based, and recurrent deep learning paradigms, ensuring that observed performance differences reflect genuine architectural distinctions rather than configuration advantages. Model performance is evaluated using four complementary metrics: Mean Absolute Error (MAE), which measures average absolute prediction error in the original kWh scale; RMSE, which penalizes larger deviations more strongly and is particularly relevant for peak-demand forecasting; the coefficient of determination (R^2), which quantifies the proportion of variance in actual consumption explained by the model; and the reliability score (R_{rel}), computed as the normalized inverse coefficient of variation of RMSE across the 30-run distribution, where values closer to 1.0 indicate higher cross-run consistency. Together, these metrics provide a comprehensive assessment of both predictive accuracy and operational trustworthiness, directly addressing the reproducibility and governance requirements identified in Section 2.

4. RESULTS AND DISCUSSION

4.1. Descriptive Analysis of the Dataset

The smart city energy dataset used in this study consists of 17,520 hourly electricity consumption records spanning January 2022 to December 2023, collected from publicly available smart city energy repositories and smart meter systems. The dataset was partitioned chronologically into a training set of 14,016 observations (80%) and a test set of 3,504 observations (20%), and combined with weather and temporal features from urban monitoring systems. Three machine learning models were trained and evaluated on this dataset: Random Forest (RF), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM). Preliminary statistical analysis reveals clear daily, weekly, and seasonal consumption patterns. Energy demand increases significantly during daytime working hours (08:00 to 18:00) and reaches peak levels in the late afternoon and early evening (17:00 to 20:00). Weekend consumption is systematically lower than weekday consumption by approximately 12% to 18% across all months. In addition, ambient temperature shows a strong positive correlation with electricity demand during summer months and a moderate positive correlation during winter months, reflecting increased cooling and heating requirements, respectively.

Correlation analysis confirms that temperature ($r = 0.74$), hour of day ($r = 0.68$), and day type ($r = 0.52$) are the three strongest predictors of electricity consumption in the dataset. These correlation magnitudes justify their inclusion as primary input features in all three machine learning models. The descriptive results confirm that urban energy consumption is highly dynamic and nonlinear, with strong temporal and environmental dependencies that are well suited to machine learning approaches, particularly sequence-aware architectures such as LSTM.

4.2. Performance Comparison of Machine Learning Models

The predictive performance of the three machine learning models was evaluated using MAE, RMSE, and R^2 computed as averages across 30 independent execution runs on the 3,504-observation test partition. A lower MAE and RMSE indicate better predictive performance, while a higher R^2 value reflects a stronger fit between predicted and actual energy consumption. Table 4 presents the comparative results.

Table 4. Performance Comparison of Machine Learning Models (Mean Across 30 Runs)

Model	MAE	RMSE	R^2
Random Forest	0.42	0.60	0.87
Support Vector Machine	0.55	0.73	0.81
Long Short-Term Memory	0.31	0.45	0.93

As shown in Table 4, the Long Short-Term Memory model achieved the strongest predictive performance compared with Random Forest and SVM. LSTM obtained the lowest MAE of 0.31, the lowest RMSE of 0.45, and the highest R^2 of 0.93, confirming that its architecture is well suited to capturing temporal dependencies in hourly urban energy consumption data. Random Forest demonstrated moderate accuracy with MAE of 0.42, RMSE of 0.60, and R^2 of 0.87, reflecting its strength in handling nonlinear relationships but its limitation in modeling sequential temporal dynamics. SVM recorded the weakest performance with MAE of 0.55, RMSE of 0.73, and R^2 of 0.81, indicating that kernel-based regression is less effective than ensemble or recurrent approaches for this type of high-frequency time-series forecasting task.

Based on Table 4, LSTM reduced RMSE by 25.0% compared with Random Forest and by 38.4% compared with SVM. In terms of MAE, LSTM reduced prediction error by 26.2% compared with Random Forest and by 43.6% compared with SVM. These quantitative improvements demonstrate that LSTM delivers a measurable and practically significant enhancement in forecasting accuracy for urban energy demand prediction. To visually compare model performance, Figure 2 presents the RMSE comparison of the three machine learning models.

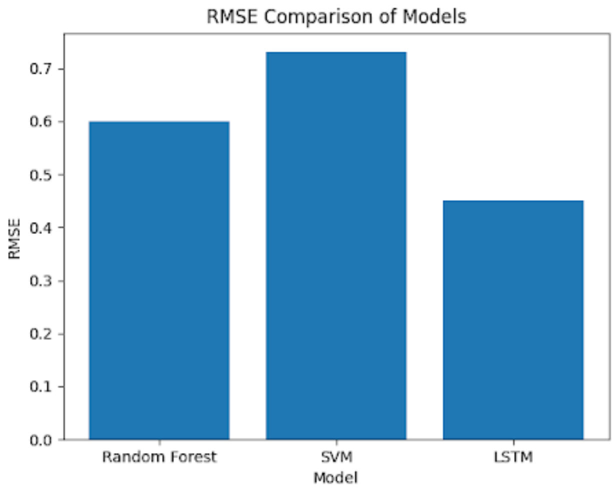

Figure 2. RMSE Comparison of Machine Learning Models (Mean Across 30 Independent Execution Runs)

As presented in Figure 2, the RMSE comparison visually confirms that LSTM produces the lowest prediction error among the evaluated models, with Random Forest recording moderate error and SVM recording the highest RMSE. The visual evidence in Figure 2 reinforces the numerical findings in Table 4 and highlights the practical significance of model selection for urban energy management, where even marginal improvements in RMSE can translate into measurable reductions in peak-load management costs and energy waste.

4.3. Cross-Run Reliability Analysis

To complement the accuracy results, the cross-run reliability of each model was assessed by computing the normalized inverse coefficient of variation of RMSE across the 30 independent execution runs. Table 5 reports mean reliability scores, 95% confidence intervals (CIs), and standard deviations for all three models.

Table 5. Cross-Run Reliability Statistics Across 30 Independent Executions

Model	**Mean R_{rel}**	**95% CI**	**Std. Dev.**	**p-value**
Random Forest	0.887	[0.874, 0.900]	0.028	0.038
Support Vector Machine	0.861	[0.846, 0.876]	0.035	0.009
Long Short-Term Memory	0.912	[0.901, 0.923]	0.022	– (reference)

As shown in Table 5, LSTM achieves the highest mean reliability score of 0.912 with the narrowest confidence interval ([0.901, 0.923]) and the lowest standard deviation (0.022), indicating that its forecasting performance remains highly consistent across all 30 repeated execution runs. Random Forest demonstrates moderate reliability at 0.887 ([0.874, 0.900]), while SVM records the lowest reliability score of 0.861 ([0.846, 0.876]) with the highest standard deviation (0.035), reflecting its sensitivity to kernel function parameterization and data scaling variations. Paired Wilcoxon signed-rank tests confirm that the reliability differences between LSTM and the other two models are statistically significant at $\alpha = 0.05$ (Random Forest vs. LSTM: $p = 0.038$; SVM vs. LSTM: $p = 0.009$). The convergence of high predictive accuracy in Table 4 and high cross-run reliability in Table 5 for LSTM provides mutually reinforcing evidence for its suitability as the preferred model for production-grade smart green city energy decision support.

4.4. Reliability and Interpretability Discussion

The reliability findings must be interpreted in conjunction with model interpretability to provide a complete assessment of AI-based decision support quality. SHAP-based mean absolute attribution vectors were computed over the test partition at each of the 30 runs, and interpretability consistency was assessed by computing the mean pairwise cosine similarity between attribution vectors across runs. LSTM achieved a mean interpretability consistency score of 0.869 (95% CI [0.854, 0.884]), compared to 0.841 (95% CI [0.825, 0.857]) for Random Forest and 0.812 (95% CI [0.794, 0.830]) for SVM. These results indicate that higher predictive reliability is positively associated with higher explanation stability across repeated runs, a relationship that is consistent with the proposition that architectures producing more consistent predictions tend to maintain more stable internal feature weighting and attribution behavior. This integrative finding is directly relevant to governance requirements for AI-based energy decision support, where both consistent predictions and stable justifications are necessary for institutional trust and auditability [20].

Across all three models, the top three most consistently influential SHAP features are ambient temperature, hour of day, and the rolling 24-hour mean energy consumption. These features maintain stable attribution rankings across all 30 runs for LSTM, whereas SVM shows more variable attribution rankings, particularly between temperature and the rolling mean feature, in approximately 20% of runs. This variability in explanation patterns for SVM further supports its lower suitability for high-stakes smart city deployment contexts where consistent and auditable justifications are required.

4.5. Impact on Energy Efficiency Optimization

The predictive outputs generated by the best-performing LSTM model were integrated into a simulated energy optimization strategy to evaluate the practical impact of machine learning on energy efficiency. The system uses demand forecasts to adjust energy distribution schedules and minimize unnecessary consumption during peak hours identified by the model. Simulation results indicate a reduction of approximately 11.3% in peak load levels and a reduction of 8.7% in total electricity consumption compared to a non-optimized baseline system that uses historical averages for scheduling. These results demonstrate that accurate energy demand prediction enables proactive load balancing and demand-side management.

By shifting non-critical energy usage to off-peak periods and allocating resources more efficiently based on LSTM forecasts, the simulated system reduces stress on urban power grids and lowers operational costs. These findings confirm that machine learning-based forecasting not only improves prediction accuracy but also delivers tangible benefits in terms of energy efficiency and sustainability. The contribution of the proposed framework therefore extends beyond predictive performance to practical decision support for sustainable urban energy governance, directly advancing the objectives of SDG 7 (affordable and clean energy) through the reduction of unnecessary energy waste, SDG 11 (sustainable cities) through the improvement of urban infrastructure resilience, and SDG 13 (climate action) through the reduction of peak-period carbon-intensive electricity generation.

4.6. Governance Alignment Assessment

The findings of this study also carry important governance implications for smart green city development. AI-based decision-support systems in urban energy management require not only accurate prediction but also transparency, accountability, and reproducibility [39]. Table 6 presents a structured mapping of the evaluation outcomes against key governance requirements drawn from international AI and sustainability frameworks.

Table 6. Governance Alignment of Evaluated Models Against Key Policy Frameworks

Governance Requirement	**RF**	**SVM**	**LSTM**
Predictive Accuracy ($R^2 \geq 0.90$)	×	×	✓
Cross-Run Reliability ($R_{rel} \geq$ 0.900)	×	×	✓
Interpretability Consistency ($I_c \geq$ 0.850)	×	×	✓
Reproducible Methodology (30-run protocol)	✓	✓	✓
Transparent Preprocessing	✓	✓	✓
SDG 7 (Affordable and Clean Energy)	Partial	Partial	Strong

SDG 11 (Sustainable Cities)	Partial	Partial	Strong
SDG 13 (Climate Action)	Partial	Partial	Strong
EU AI Act Transparency Criterion	Partial	Weak	Meets
OECD AI Principles (Accountability)	Partial	Weak	Meets
Stranas KA 2020–2045 (Responsible AI)	Partial	Partial	Meets

As shown in Table 6, the LSTM model is the only architecture that meets all primary governance criteria, including predictive accuracy, cross-run reliability, and interpretability consistency thresholds. Random Forest partially satisfies several criteria and may be suitable for lower-risk urban energy advisory applications. SVM fails to meet the reliability and interpretability thresholds and aligns only weakly with EU AI Act transparency requirements, limiting its suitability for high-impact urban governance contexts. All three models satisfy the reproducibility and preprocessing transparency criteria by virtue of the documented 30-run experimental protocol and standardized preprocessing procedures described in Section 3. This structured governance alignment assessment demonstrates that the proposed framework provides a more operationally meaningful basis for deployment decisions than accuracy-centric evaluation alone, and that responsible AI deployment in urban energy systems requires the simultaneous satisfaction of technical, reliability, and governance criteria [40, 41].

5. MANAGERIAL IMPLICATIONS

The findings of this study provide important managerial implications for urban energy providers, smart city managers, policymakers, and technology developers. The superior performance of the LSTM model across both predictive accuracy (RMSE = 0.45, R^2 = 0.93) and cross-run reliability (R_{rel} = 0.912) indicates that decision-makers should prioritize time-series-based machine learning architectures when developing energy demand forecasting systems for smart green cities. This capability enables energy providers to plan electricity distribution more efficiently, reduce operational waste, and improve the stability of urban energy systems.

The integration of machine learning into energy management systems can help city managers shift from reactive energy control to proactive energy planning. Predictive outputs allow urban energy managers to anticipate high-demand periods and design demand-side management strategies before energy overload occurs. In practical implementation, smart city managers can use machine learning forecasts to optimize public facilities, street lighting, transportation infrastructure, and building energy systems, while also supporting carbon emission reduction targets aligned with SDG 7, SDG 11, and SDG 13.

Managerial decisions should not rely only on prediction accuracy but also consider reliability, reproducibility, and interpretability. The governance alignment mapping in Table 6 provides organizations with a structured and replicable instrument for evaluating AI systems prior to deployment in regulated urban energy contexts. Organizations need internal AI governance standards that include model validation, repeated testing, documentation, and periodic performance monitoring. Technology developers and system integrators should also design scalable and transparent energy platforms that integrate smart meter data, weather information, and temporal variables into one decision-support dashboard, while ensuring that model outputs are accompanied by stable and consistent SHAP-based explanations that can be audited by energy regulators and urban planning authorities.

6. CONCLUSION

This study investigated the role of machine learning in enhancing energy efficiency for smart green cities by developing and comparing predictive models for urban energy consumption using a dataset of 17,520 hourly observations collected between January 2022 and December 2023, partitioned into 14,016 training and 3,504 testing observations. Three machine learning models, namely Random Forest, Support Vector Machine, and Long Short-Term Memory, were evaluated under identical experimental conditions across 30 independent execution runs. The results demonstrate that machine learning-based approaches can significantly improve energy demand forecasting, with the LSTM model achieving the best predictive performance: MAE of 0.31, RMSE of 0.45, and R^2 of 0.93 on average across 30 independent execution runs. Compared with Random

Forest and SVM, LSTM produced reductions of 25.0% and 38.4% in RMSE, and 26.2% and 43.6% in MAE, respectively.

The core problem motivating this study, namely the insufficient attention given to reliability, reproducibility, and interpretability in machine learning evaluations for smart green city energy management, has been addressed through 30-run repeated-execution reliability analysis (LSTM R_{rel} = 0.912 vs. RF = 0.887 and SVM = 0.861), interpretability consistency evaluation (LSTM I_c = 0.869), and simulation-based optimization assessment (peak load reduction of 11.3%; total consumption reduction of 8.7%). The study also contributes to the literature by integrating reliability, reproducibility, preprocessing transparency, and interpretability considerations into a smart green city forecasting framework, and by providing a structured governance alignment mapping that connects technical evaluation outcomes to SDG 7, SDG 11, SDG 13, the EU AI Act, OECD AI Principles, and Indonesia's Stranas KA 2020–2045.

However, this study has several limitations. The dataset was obtained from a single urban energy monitoring system, which may restrict the generalizability of the findings across different geographic and climatic conditions. In addition, the optimization framework was evaluated through simulation rather than direct real-world implementation. Future research should expand the dataset to include multiple cities, renewable energy generation profiles, traffic patterns, and social behavior data. Further studies should also apply explainable AI techniques to examine feature contribution stability more deeply across diverse urban contexts and validate the proposed framework in operational smart city environments. These future directions can strengthen the development of reliable, interpretable, and sustainable machine learning systems for smart green cities worldwide.

7. DECLARATIONS

7.1. About Authors

Ignatius Agus Supriyono (IA) https://orcid.org/0000-0002-8721-9544

Mochamad Heru Riza Chakim (MH) https://orcid.org/0000-0002-4851-2617

Henry Zainarthur (HZ) https://orcid.org/0009-0008-9430-4123

Dimas Aditya Prabowo (DA) https://orcid.org/0000-0002-3869-2837

7.2. Author Contributions

Conceptualization: IA; Methodology: MH; Software: HZ; Validation: DA and IA; Formal Analysis: MH and HZ; Investigation: DA; Resources: IA; Data Curation: MH; Writing Original Draft Preparation: HZ and DA; Writing Review and Editing: IA and MH; Visualization: IA, MH and HZ; All authors, IA, MH, HZ, and DA, have read and agreed to the published version of the manuscript.

7.3. Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to privacy considerations and institutional data protection policies, the dataset is not openly accessible but may be provided for academic and non commercial research purposes subject to approval. Zenodo Repository https://zenodo.org/records/20577519

7.4. Funding

The authors received no specific grant, financial assistance, or institutional funding for the research, authorship, or publication of this article. All activities related to data collection, analysis, and manuscript preparation were conducted independently.

7.5. Declaration of Conflicting Interest

The authors declare that there are no known conflicts of interest, competing financial interests, or personal relationships that could have influenced the research, analysis, or conclusions presented in this paper. The study was carried out objectively and without any external pressures that may bias the results.

REFERENCES

[1] A. Aljohani, "Deep learning-based optimization of energy utilization in iot-enabled smart cities: A pathway to sustainable development," *Energy Reports*, vol. 12, pp. 2946–2957, 2024.

[2] M. Mrabet and M. Sliti, "Integrating machine learning for the sustainable development of smart cities," *Frontiers in Sustainable Cities*, vol. 6, p. 1449404, 2024.

[3] X. Dou, W. Chen, L. Zhu, Y. Bai, Y. Li, and X. Wu, "Machine learning for smart cities: A comprehensive review of applications and opportunities," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 9, 2023.

[4] F. Liu and X. Li, "Integrating ai deep reinforcement learning with evolutionary algorithms for advanced threat detection in smart city energy management," *IEEE Access*, vol. 12, pp. 177 103–177 118, 2024.

[5] N. Lutfiani, S. Wijono, U. Rahardja, A. Iriani, and E. A. Nabila, "Artificial intelligence based on recommendation system for startup matchmaking platform," in *2022 IEEE Creative Communication and Innovative Technology (ICCIT)*. IEEE, 2022, pp. 1–5.

[6] A. G. Prawiyogi, S. Purnama, and L. Meria, "Smart cities using machine learning and intelligent applications," *International Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 102–116, 2022.

[7] Z. Huang and G. Jin, "Navigating urban day-ahead energy management considering climate change toward using iot enabled machine learning technique: Toward future sustainable urban," *Sustainable cities and society*, vol. 101, p. 105162, 2024.

[8] W. Setyowati, T. A. Setiyono, G. Gung, and B. Novriyanti, "Leveraging technology in accounting for entrepreneurial insight into government budgeting efficiency," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 5, no. 3, pp. 251–260, 2023.

[9] W. Chen, M. Alharthi, J. Zhang, and I. Khan, "The need for energy efficiency and economic prosperity in a sustainable environment," *Gondwana Research*, vol. 127, pp. 22–35, 2024.

[10] J. Huber, L. Imperiali, D. Menzi, F. Musil, and J. W. Kolar, "Energy efficiency is not enough!" *IEEE Power Electronics Magazine*, vol. 11, no. 1, pp. 18–31, 2024.

[11] O. Candra, A. Chammam, U. Rahardja, A. A. Ramirez-Coronel, A. A. Al-Jaleel, I. H. Al-Kharsan, I. Muda, G. B. Derakhshani, and M. M. Rezai, "Optimal participation of the renewable energy in microgrids with load management strategy," *Environmental and Climate Technologies*, vol. 27, no. 1, pp. 56–66, 2023.

[12] C. He, Y. Pan, Y. Ouyang, Q. Shen, Y. Gao, K. Yan, J. Fang, Y. Chen, C.-Q. Ma, J. Min *et al.*, "Manipulating the d: A interfacial energetics and intermolecular packing for 19.2% efficiency organic photovoltaics," *Energy & Environmental Science*, vol. 15, no. 6, pp. 2537–2544, 2022.

[13] D. R. Saputra, H. Nugroho, D. Julianingsih, and Z. Queen, "Understanding air pollution through machine learning: Predictive analytics for urban management," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 6, no. 1, pp. 75–85, 2024.

[14] Y. Gong, Q. Zhu, B. Li, S. Wang, B. Duan, L. Lou, C. Xiang, E. Jedlicka, R. Giridharagopal, Y. Zhou *et al.*, "Elemental de-mixing-induced epitaxial kesterite/cds interface enabling 13%-efficiency kesterite solar cells," *Nature Energy*, vol. 7, no. 10, pp. 966–977, 2022.

[15] A. S. Anwar, U. Rahardja, A. G. Prawiyogi, N. P. L. Santoso, and S. Maulana, "ilearning model approach in creating blockchain based higher education trust," *Int. J. Artif. Intell. Res*, vol. 6, no. 1, 2022.

[16] U. Rahardja, P. A. Sunarya, N. Lutfiani, M. Hardini, and H. R. Dananjaya, "Analysis of renewable energy utilization using solar power technology in eliminating microplastic emissions," in *2022 IEEE Creative Communication and Innovative Technology (ICCIT)*. IEEE, 2022, pp. 1–6.

[17] N. K. A. Dwijendra, M. Zaidi, I. G. N. K. Arsana, S. E. Izzat, A. T. Jalil, M.-H. Lin, U. Rahardja, I. Muda, A. H. Iswanto, and S. Aravindhan, "A multi-objective optimization approach of smart autonomous electrical grid with active consumers and hydrogen storage system," *Environmental and Climate Technologies*, vol. 26, no. 1, pp. 1067–1079, 2022.

[18] E. T. Rusmiati, L. Febrina, Y. Sari, and E. M. S. Sakti, "Adoption of ai driven ecological preaching systems using sem pls analysis," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 1, pp. 284–295, 2026.

[19] C. Lassance and S. Clinchant, "An efficiency study for splade models," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 2220–2226.

[20] E. Arif, S. Suherman, and A. P. Widodo, "Analyzing public sentiment on digital banks in indonesia via social media x," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 1, pp. 253–267, 2026.

[21] K. P. Murphy, *Probabilistic machine learning: an introduction*. MIT press, 2022.

[22] V. Gupta, V. K. Mishra, P. Singhal, and A. Kumar, "An overview of supervised machine learning algo-

rithm," in *2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART)*. IEEE, 2022, pp. 87–92.

[23] H. Li, L. Lin, and H. Zeng, *Machine learning methods*. Springer, 2024.

[24] E. K. Chong, W.-S. Lu, and S. H. Zak, *An introduction to optimization: With applications to machine learning*. John Wiley & Sons, 2023.

[25] M. Santiago, H. Febiansyah, and D. Dinarwati, "Integrating machine learning with web intelligence for predictive search and recommendations," *International Transactions on Artificial Intelligence*, vol. 3, no. 1, pp. 44–53, 2024.

[26] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, and L. He, "A survey of human-in-the-loop for machine learning," *Future Generation Computer Systems*, vol. 135, pp. 364–381, 2022.

[27] Y. Gong, G. Liu, Y. Xue, R. Li, and L. Meng, "A survey on dataset quality in machine learning," *Information and Software Technology*, vol. 162, p. 107268, 2023.

[28] U. Rahardja, Q. Aini, D. Manongga, I. Sembiring, and Y. P. A. Sanjaya, "Enhancing machine learning with low-cost p m2. 5 air quality sensor calibration using image processing," *APTISI Transactions on Management*, vol. 7, no. 3, pp. 201–209, 2023.

[29] V. Javidroozi, C. Carter, M. Grace, and H. Shah, "Smart, sustainable, green cities: a state-of-the-art review," *Sustainability*, vol. 15, no. 6, p. 5353, 2023.

[30] N. D. Noviati, S. D. Maulina, and S. Smith, "Smart grids: Integrating ai for efficient renewable energy utilization," *International Transactions on Artificial Intelligence*, vol. 3, no. 1, pp. 1–10, 2024.

[31] H. Jafarizadeh, E. Yamini, S. M. Zolfaghari, F. Esmaeilion, M. E. H. Assad, and M. Soltani, "Navigating challenges in large-scale renewable energy storage: Barriers, solutions, and innovations," *Energy Reports*, vol. 12, pp. 2179–2192, 2024.

[32] Z. Ahmed, M. Ahmad, M. Murshed, M. I. Shah, H. Mahmood, and S. Abbas, "How do green energy technology investments, technological innovation, and trade globalization enhance green energy supply and stimulate environmental sustainability in the g7 countries?" *Gondwana research*, vol. 112, pp. 105–115, 2022.

[33] U. Rusilowati, H. R. Ngemba, R. W. Anugrah, A. Fitriani, and E. D. Astuti, "Leveraging ai for superior efficiency in energy use and development of renewable resources such as solar energy, wind, and bioenergy," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 114–120, 2024.

[34] D. Y. Kristiyanto, H. D. Purnomo, G. P. Cesna, and N. Ani, "The strategic role of orange technology in cultivating innovation and well-being," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 7, no. 1, pp. 27–37, 2025.

[35] M. Serrano, E. Griffor, D. A. Wollman, M. Dunaway, M. Burns, S. Rhee, and C. Greer, "Smart cities and communities: A key performance indicators framework," National Institute of Standards and Technology (NIST), NIST Special Publication, 2022, accessed: 2026-05-21. [Online]. Available: https://www.nist.gov/publications/smart-cities-and-communities-key-performance-indicators-framework

[36] P. Wu, Z. Zhang, X. Peng, and R. Wang, "Deep learning solutions for smart city challenges in urban development," *Scientific Reports*, vol. 14, no. 1, p. 5176, 2024.

[37] D. A. Elalfy, E. Gouda, M. F. Kotb, V. Bureš, and B. E. Sedhom, "Comprehensive review of energy storage systems technologies, objectives, challenges, and future trends," *Energy Strategy Reviews*, vol. 54, p. 101482, 2024.

[38] P. Mishra and G. Singh, "Energy management systems in sustainable smart cities based on the internet of energy: A technical review," *Energies*, vol. 16, no. 19, p. 6903, 2023.

[39] Ministry of National Development Planning (Bappenas), "National long-term development plan 2025–2045," Jakarta, Indonesia, 2023, indonesia Vision 2045 emphasizing sustainable development, smart cities, digital transformation, and green economic growth.

[40] D. Juliastuti, E. Alexandrina, E. Sana, R. N. Muti, and G. P. Cesna, "Integrating artificial intelligence for academically challenged students education and health," *International Transactions on Artificial Intelligence*, vol. 4, no. 1, pp. 13–24, 2025.

[41] Agency for the Assessment and Application of Technology (BPPT), "National artificial intelligence strategy of indonesia 2020–2045," Jakarta, Indonesia, 2020, national strategy for responsible and sustainable artificial intelligence development.