

Employing Hybrid GPR-KNN with Diverse Acquisition Functions for the Prediction of Current Consumption in Aeronautical Ground Lighting Systems

Wan Mohammed Rais Jamaludin*, Wan Mazlina Wan Mohamed and Nik Hakimi Nik Ali

Abstract— Numerous studies have explored runway visibility prediction. However, relatively few have focused on predicting nominal current output based on visibility, time, and acquisition functions, as effective management of aeronautical lighting is crucial for optimizing energy consumption and reducing environmental impact. This study decisively targets optimizing nominal current output of runway edge lights at Subang Airport using advanced Machine Learning (ML) techniques, leveraging a range of acquisition functions and comprehensive meteorological data. A hybrid combination of Gaussian Process Regression, which is particularly effective in accounting for data-specific uncertainties and non-linear relationships, and Nearest Neighbor Classifiers are utilized to accurately predict the nominal current output based on predicted visibility and time. Both 5-fold cross-validation and holdout validation were performed to ensure robust evaluation. Model performance is assessed using key metrics such as accuracy, precision, recall, and F1 score. The findings explicitly demonstrate that the Expected Improvement (EI) acquisition function outperformed others, which is the most accurate in predicting the nominal current output in both validation methods by achieving 99.62% accuracy. In conclusion, this study presents a groundbreaking approach to predicting and improving nominal output current for runway edge lights by applying the EI acquisition function.

Index Terms— Acquisition Function, Aeronautical Ground Lighting, GPR, KNN, Meteorological Data, Nominal Current Output

I. INTRODUCTION

The absence of standardized guidelines regarding nominal

This manuscript is submitted on 13th May 2025, revised on 24th June 2025, accepted on 3rd July 2025 and published on 31st October 2025. Wan Mohammed Rais Jamaludin is with the Faculty of Mechanical Engineering, Universiti Teknologi MARA, Shah Alam (2022283062@student.uitm.edu.my). Assoc. Prof. Dr. Wan Mazlina Wan Mohamed is with the Malaysia Institute of Transport and Faculty of Mechanical Engineering, Universiti Teknologi MARA, Shah Alam (wmazlina@uitm.edu.my). Ir. Dr. Nik Hakimi Nik Ismail is with the Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam (hakimiali@uitm.edu.my).

*Corresponding author
Email address: 2022283062@student.uitm.edu.my

1985-5389/© 2023 The Authors. Published by UiTM Press. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

current output in airport operations has emerged as a significant issue in the rapidly evolving aviation field.

Effective monitoring and control of Aeronautical Ground Lighting (AGL) systems are vital to ensure that these lights do not operate at full power unnecessarily [1]. This practice is crucial, as excessive lighting increases energy waste and operational costs. Adjusting the intensity of lights within an aerodrome lighting system is essential for maintaining safety and efficiency, particularly in response to varying environmental conditions. Factors such as background brightness, which can fluctuate dramatically from day to night, and visibility conditions, impacted by weather or time of day, play a key role in determining the optimal lighting intensity [2]. Each country should take the initiative to develop and refine its intensity-setting procedures as recommended by the International Civil Aviation Organization (ICAO) to enhance energy efficiency and safety [3]. By doing so, they can ensure that lighting levels are appropriately calibrated, ultimately supporting better operational practices in airport environments while reducing unnecessary energy consumption.

Most research has focused on commercial spaces and Airport Terminal Buildings (ATBs). A Machine Learning (ML) model is vital for predicting energy consumption in airport buildings [4]. ATBs consume a significant amount of electricity for lighting, HVAC systems, and electronic equipment, as mentioned by [5], posing challenges to energy management and sustainability. Additionally, energy consumption varies with weather conditions and passenger numbers, complicating accurate predictions. However, runway lighting has unique energy needs, have been overlooked.

A recent study emphasizes that airports are major energy consumers, particularly in ATB, where the demand for HVAC and lighting peaks [6]. To tackle this challenge, implementing standardized energy management systems is essential for monitoring and reducing energy use across airports. One promising approach is adopting advanced ML models, which can significantly enhance energy efficiency [7]. These systems facilitate real-time adjustments to lighting based on occupancy levels and weather conditions, thus optimizing energy consumption. Moreover, airports can leverage ML to analyze historical data to predict and optimize energy usage [8]. This

capability allows for continuous monitoring and the ability to make immediate adjustments, enabling informed decision-making regarding energy efficiency upgrades and overall management strategies.

External factors, such as air passenger traffic and weather conditions, are crucial in shaping airport energy consumption patterns [9]. As energy costs continue to rise and environmental concerns grow, it has become increasingly important to implement practices that effectively reduce energy use while maintaining operational efficiency. Various ML techniques have been employed to predict energy consumption. These include hybrids of knowledge-based and ML models, as well as specific algorithms such as Random Forest, XGBoost, LightGBM, Support Vector Regression, Long Short-Term Memory networks, and different types of Artificial Neural Networks (ANN), including ANN-LR, ANN-RR, ANN-RF, ANN-GB, ANN-DT, and ANN-ELM [10]–[13]. Utilizing these advanced methodologies can improve energy management and contribute to sustainable operations.

The literature shows that MLs can be used in many existing studies concentrating on ATB and general buildings. However, runway lighting has received significantly less attention regarding energy efficiency improvements. To address this gap, automation is essential for predicting lighting usage based on real-time factors, such as time of day and weather conditions [1], [14]. Effective management of AGL systems is critical in ensuring safety and operational efficiency within airport environments. Unfortunately, many current lighting systems operate on fixed schedules or rely on manual controls. This reliance often results in suboptimal lighting that fails to adapt to real-time environmental conditions. The lack of automation in lighting systems can lead to excessive energy consumption.

This study introduces a novel approach centered on two key components to fulfill the objective of developing an automated AGL system:

- 1) Improving the Gaussian Process Regression – K Nearest Neighbor (GPR-KNN) model by incorporating various acquisition functions, including Expected Improvement (EI), Expected Improvement Plus (EIP), Expected Improvement Per Second (EIPS), Expected Improvement Per Second Plus (EIPSP), and Lower Confidence Bound (LCB). These functions enhance the model's predictive performance and sampling efficiency.

- 2) Evaluating the chosen acquisition function's effectiveness and integrating it into a comprehensive ML framework to build a robust model that accurately predicts nominal current output for practical applications.

The research paper consists of Section I, which explains the previous studies, problem statements, objectives, scope of work, and novelty. The rest of the paper is organized as follows: Section II presents the methodology used to achieve the objectives of this study, Section III describes the results and discussion, and Section IV concludes the overall study.

II. METHODOLOGY

The methodology for this study is shown in Fig. 1, and details are explained below.

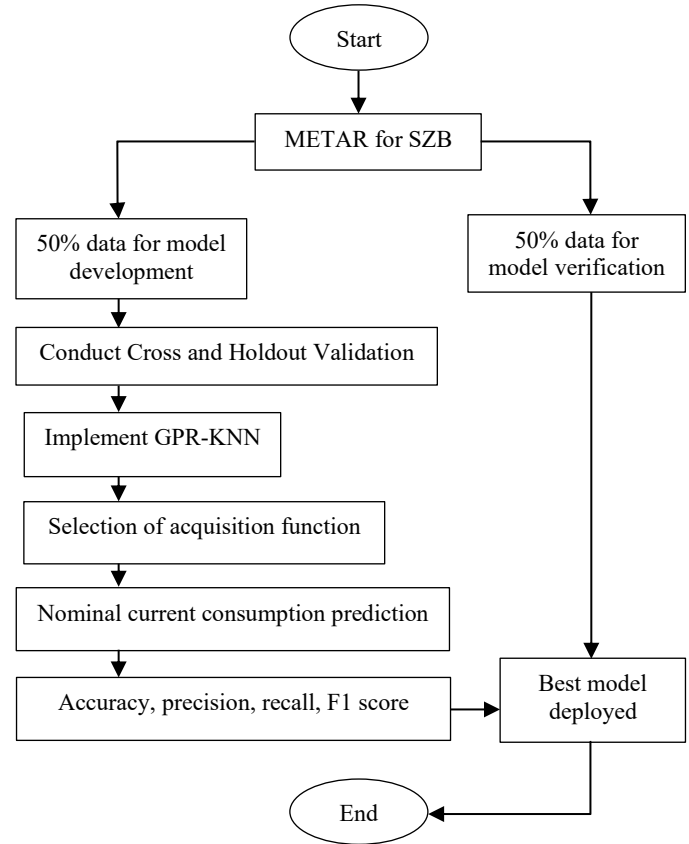


Fig. 1. Methodology for developing GPR-KNN model

A. Study Area

Subang Airport, or SZB, is located in Subang, Selangor, Malaysia, about 15 kilometers west of Kuala Lumpur. It has a favorable climate year-round, making it a good option for domestic and regional flights. Daily Meteorological Aerodrome Report (METAR) weather data from the National Oceanic and Atmospheric Administration (NOAA), was collected from January 1 to December 31, 2023, totaling 8,760 data points [15]. As stated by NOAA, METAR serves as the global standard coding system for hourly surface weather reports, highlighting its essential function in providing accurate weather data. Half of this data was used to train and validate a model, while the other half was used to verify it.

Meteorological factors such as daily air temperature, dew point temperature, wind direction, wind speed, and pressure were used in the calculations. Details on the collected data are in Table I and were important for predicting runway visibility through modeling. To strengthen the simulation, data such as “VRB” for variable wind direction was replaced with the most common value, 0 [16]. This change helps convert unclear data into a format that ML algorithms can efficiently work with. Additionally, visibility values over 10,000 meters were labeled as 10 to create a uniform reference for analysis. These

improvements help the model perform better and increase the reliability of the research results.

TABLE I. INFORMATION REGARDING THE VARIABLES USED IN THIS STUDY

Features	Unit
Hourly Temperature (T)	Degree Celsius (°C)
Hourly Dew-point Temperature (Td)	Degree Celsius (°C)
Difference between Hourly Temperature and Dew-point Temperature (T-Td)	Degree Celsius (°C)
Hourly Wind Speed (WS)	Knot (kt)
Hourly Wind Direction (WD)	Degree (°)
Hourly Pressure (P)	Hectopascal (hPa)
Visibility (V)	Kilometer (km)
Current Consumption (CC)	Ampere (A)

B. Cross and Holdout Validation

A cross-validation approach was used to create each algorithm, regardless of the data split for training and assessment [17], [18]. In this study, the folds were configured to 5, with 20% allocated for validation and 80% for training. This configuration aimed to mitigate overfitting and to ensure optimal model performance by avoiding an imbalanced dataset distribution [19]. An example of the stacking structure with the fold set to 5 is shown in Fig. 2.

This study also considered the holdout validation method. The holdout validation method allocates 25% of the total dataset for validation, while the remaining 75% is utilized for training the model [20], [21]. This methodology is particularly suitable for large datasets. It is vital to ensure that the model's performance is evaluated accurately and devoid of any bias that may arise from the training data [22].

This approach is particularly advantageous when working with large datasets, as it facilitates a comprehensive analysis of model effectiveness. By systematically partitioning the data, a thorough assessment of the trained models can be conducted, ensuring that their accuracy and reliability are evaluated rigorously prior to deployment in real-world applications. Such meticulous validation is essential to instill confidence in the performance of the models and their generalization capabilities.

C. Selection of GPR-KNN

In this study, the selection of GPR was based on [23]. From the paper, it has shown that GPR is the most accurate model used for visibility prediction in MATLAB. As for the selection of KNN, the justification was based on the paper [24], where it shows that KNN is the most reliable model for current consumption prediction.

The proposed integrated model of GPR-KNN that incorporates various acquisition functions was developed using MATLAB R2024b, with the Regression Learner utilized for the GPR component and the Classification Learner Toolbox for the KNN component. Performance analysis was conducted through various metrics, including accuracy, precision, recall, and F1 score. MATLAB R2024b was acquired directly from www.mathworks.com, with the associated license number being 41245384. The software operates on a computer equipped with an Intel Core i5-1235U CPU, which operates at 1.30 GHz,

and has 16 GB of DDR4 RAM.

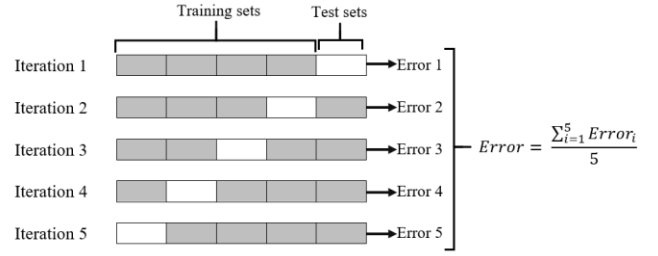


Fig. 2. Cross validation with 5-fold validation

D. Acquisition Function

1) Expected Improvement (EI)

The EI acquisition function plays a crucial role in Bayesian Optimization (BO) by guiding the sequential sampling process to find the optimum of complex objective functions [25]–[27]. The EI acquisition function is widely recognized for its essential role in BO, as it guides the sequential sampling process to approximate the complex objective function [25]–[27]. It balances exploring unknown regions with exploiting known promising samples, making it a crucial component in optimization [27], [28]. The EI acquisition function in MATLAB utilizes a surrogate model, often based on a Gaussian process, to guide the optimization process by balancing exploration and exploitation [28], [29]. The EI acquisition function has been widely used to solve practical engineering problems due to its closed-form nature, effectively reducing sampling costs and improving optimization accuracy [28], [30].

2) Plus

A key challenge is avoiding local minima when using BO to find the best solution. The acquisition function “plus” was used to address this. It adjusted their behavior to prevent over-focusing on specific areas, thus promoting exploration. The algorithm checks for excessive concentration on particular spots. It modifies its approach to explore new territories if needed, with adjustments occurring up to five times. The Exploration Ratio plays a crucial role in balancing the exploration of new areas with revisiting known ones, enhancing the efficiency of the optimization process, and guiding the algorithm toward better solutions.

3) Per Second

When finding the best solution using computers, evaluation times can vary significantly. For instance, some areas take longer to assess with Support Vector Machines. To enhance efficiency, BO strategically considers evaluation times and makes informed decisions. By focusing on evaluations that promise the greatest improvement relative to time spent, BO optimizes resources, speeding up the search process and increasing the chances of finding the best solution. This approach balances precision and speed, guiding toward optimal results.

4) Lower Confidence Bound (LCB)

The LCB acquisition function offers a refined strategy for decision-making under uncertainty in BO. It evaluates potential outcomes by focusing on options slightly below average, specifically, two standard deviations lower, acknowledging inherent prediction uncertainties. This method is akin to a cautious explorer identifying valuable opportunities while considering potential risks. It balances the excitement of exploration with careful uncertainty assessment, representing an intelligent strategy that navigates between embracing new possibilities and safeguarding against setbacks.

E. Nominal Current Consumption

The electrical power for AGL circuits, especially in series configurations, is efficiently supplied by Constant Current Regulators (CCRs). These essential devices play a crucial role in maintaining consistent light output over long distances, such as aerodrome runways. CCRs are engineered to deliver a stable current output, even amidst input voltage or load resistance fluctuations. In aviation, ensuring uniform lighting conditions is vital for pilots during critical phases like take-off, landing, and taxiing. These regulators are adept at producing a constant current output, which remains stable regardless of circuit load or voltage changes from the power source. The current consumption of CCRs is classified into five distinct steps, as summarized in Table II [31]. This classification system allows for careful selection of lighting intensity, with options ranging from 2.8 amperes to 6.6 amperes and a tolerance of ± 0.1 amperes.

TABLE II. NOMINAL CURRENT OUTPUT RANGE

Current step	Nominal output (RMS Amperes)	Nominal output (RMS Amperes)
5	6.60	6.50 - 6.70
4	5.20	5.10 - 5.30
3	4.10	4.00 - 4.20
2	3.40	3.30 - 3.50
1	2.80	2.70 - 2.90

F. Model Performance Evaluation

Since the main objective of this study is to predict the nominal current output, the performance assessment of the trained models focused solely on the final output, which is the classification model. This involved evaluating several key metrics: accuracy, precision, recall, and F1 Score [22], [32]. Each metric provides unique insights into the model's effectiveness and contributes to a comprehensive understanding of its performance.

Accuracy evaluates the model's accuracy by assessing the ratio of true outcomes (positives and negatives) to the total number of cases analyzed. In contrast, precision zeroes in on the quality of positive predictions, measuring the ratio of true positive cases to all instances identified as positive. This metric is especially crucial in scenarios where the consequences of false positives are significant.

Recall assesses how well a model can recognize all relevant

instances by measuring the ratio of true positives to the total number of positive cases, including both true and false negatives. This metric is crucial in situations where it is important to identify as many positive cases as possible, even if it means allowing for some false positives. The F1 Score is a harmonic mean that combines precision and recall, yielding a single metric that addresses both aspects. This metric is especially beneficial in scenarios with imbalanced class distributions or when one of the metrics (precision or recall) is given more importance.

The following equations as in (1) until (4) are utilized [22], where True Positives (TP) refer to instances that are correctly identified as belonging to the target class, True Negatives (TN) indicate instances accurately identified as not belonging to the target class, False Positives (FP) represent instances incorrectly classified as belonging to the target class, False Negatives (FN) describe instances that are incorrectly identified as not belonging to the target class.

$$Accuracy = \frac{TP+TN}{Total\ Predictions} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$

III. RESULTS AND DISCUSSION

Table III presents the results for each acquisition function based on cross-validation and holdout validation.

TABLE III. RESULT FOR EACH ACQUISITION FUNCTION BASED ON TYPE OF VALIDATION

Validation	Acq. Function	Accuracy	Precision	Recall	F1 Score
Cross	EIPSP	97.17%	93.90%	95.33%	94.58%
	EI	99.41%	98.18%	99.45%	98.79%
	EIP	99.41%	98.18%	99.45%	98.79%
	EIPS	95.46%	92.17%	93.46%	92.79%
	LCB	96.76%	93.09%	93.71%	93.39%
Holdout	EIPSP	96.62%	93.36%	96.55%	94.80%
	EI	99.73%	99.71%	99.79%	99.75%
	EIP	94.79%	94.65%	92.83%	93.70%
	EIPS	97.63%	97.61%	95.51%	96.50%
	LCB	94.61%	92.26%	93.38%	92.79%

A. Analysis of Cross-Validation Results

EIPSP achieves a commendable accuracy of 97.17%, with a precision of 93.90% and a recall of 95.33%, resulting in an F1 score of 94.58%. While its performance is moderate, it effectively balances time efficiency and prediction reliability. In contrast, EI stands out with an exceptional accuracy of 99.41%, precision of 98.18%, recall of 99.45%, and an impressive F1 score of 98.79%, making it ideal for highly reliable processes. EIP matches EI's metrics perfectly, boasting an identical accuracy, precision, recall, and F1 score. It is

equally beneficial for dependable cross-validation. EIPS shows moderate performance with an accuracy of 95.46%, precision of 92.17%, recall of 93.46%, and an F1 score of 92.79%. Its computational efficiency suits resource-limited contexts. Lastly, LCB has a respectable accuracy of 96.76%, with a precision of 93.09% and a recall of 93.71%, yielding an F1 score of 93.39%.

B. Analysis of Holdout Validation Results

EIPSP achieves an accuracy of 96.62%, with a precision of 93.36%, a recall of 96.55%, and an F1 score of 94.80%. It balances computational efficiency with slightly reduced predictive metrics, which is ideal for resource-optimized scenarios. In contrast, EI excels with an accuracy of 99.73%, boasting precision (99.71%), recall (99.79%), and an F1 score (99.75%), making it the top choice for high-precision applications. EIP shows commendable performance at 94.79% accuracy, precision of 94.65%, recall of 92.83%, and F1 score of 93.70%. It emphasizes robustness and is suitable for reliability-focused situations. EIPS delivers a well-rounded accuracy of 97.63%, precision of 97.61%, recall of 95.51%, and F1 score of 96.50%. It merges computational efficiency with good predictive accuracy, making it adaptable for various applications. Finally, LCB has an accuracy of 94.61%, a precision of 92.26%, a recall of 93.38%, and an F1 score of 92.79%.

C. Choosing the Best Model

The higher the accuracy, precision, recall, and F1 Score, the better the results are. The models are ranked in Table IV, based on the performance score (β) that is defined in as in (5). Higher values of β signify better model performance:

$$\beta = \text{Accuracy} + \text{Precision} + \text{Recall} + \text{F1 score} \quad (5)$$

TABLE IV. RANKING BASED ON THE PERFORMANCE SCORE

Validation	Acquisition function	Performance score
Holdout	EI	398.98%
Cross	EI	395.83%
Cross	EIP	395.83%
Holdout	EIPS	387.25%
Holdout	EIPSP	381.33%
Cross	EIPSP	380.98%
Cross	LCB	376.95%
Holdout	EIP	375.97%
Cross	EIPS	373.88%
Holdout	LCB	373.04%

Holdout validation shows exceptional values for EI, achieving a remarkable score of 398.98%. This demonstrates the method's ability to effectively produced high-quality outcomes with specific acquisition strategies. In contrast, cross-validation yields slightly lower results, with EI and EIP scoring 395.83%. While not as high as holdout validation, the consistent performance across multiple folds enhances confidence in these acquisition functions' reliability and robustness. Both methods highlight EI's strong performance, showcasing its effectiveness in balancing exploration and

exploitation, which is crucial for predicting predictive models. The findings align with previous research suggesting that EI adeptly manages the trade-off between exploring unknown areas and exploiting known promising regions [28]. EI's performance in noisy environments is noteworthy, as it maintains computational efficiency without requiring complex optimization processes, making it suitable for challenging settings.

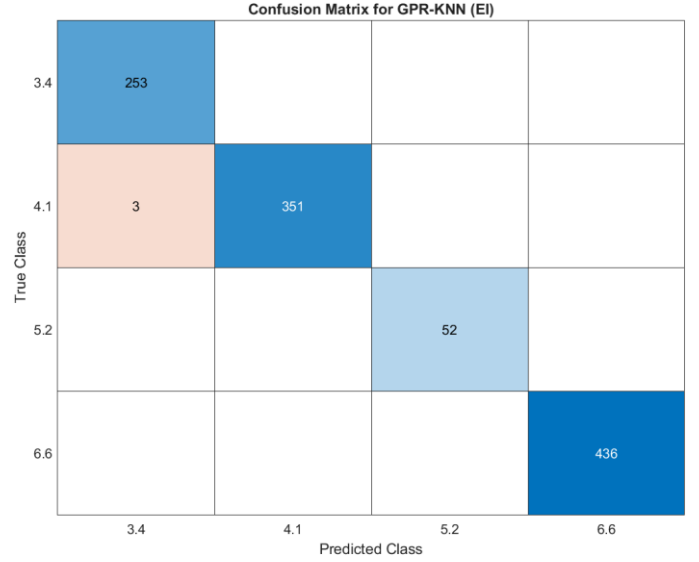


Fig. 3. Confusion matrix for GPR-KNN (EI)

The confusion matrix shown in Fig. 3 indicates that GPR-KNN model using the EI acquisition function in holdout validation shows exceptional classification performance across four nominal current output classes: 3.4, 4.1, 5.2, and 6.6. Specifically, the model achieved correct classifications of 253, 351, 52, and 436 instances for each respective class, resulting in a minimal number of misclassifications, where only three occurring between classes 4.1 and 3.4. Consequently, the overall classification accuracy is approximately 99.73%, which is consistent with the accuracy reported in the accompanying performance analysis.

Furthermore, the model demonstrated perfect recall for classes 3.4, 5.2, and 6.6, and near-perfect recall for class 4.1. This performance underscores the model's robustness in addressing both minority and majority classes, despite the inherent class imbalance present in the dataset. Notably, the model exhibited zero-error performance for the higher-demand classes 5.2 and 6.6, which are critical for ensuring operational safety, thereby highlighting its practical reliability for real-world applications. These findings underscore the GPR-KNN model's strong generalization ability, high discriminative power, and operational suitability for automated AGL systems.

The ROC curve analysis in Fig. 4 shows the GPR-KNN model utilizing the EI acquisition function indicates remarkable classification performance across all evaluated target classes. The model achieved high area under the curve (AUC) values, specifically registering 0.9982 for class 3.4, 0.9958 for class 4.1, and a perfect score of 1.0000 for both classes 5.2 and 6.6. These results affirm to the model's exceptional discriminative

capability, exhibiting minimal overlap between positive and negative class distributions.

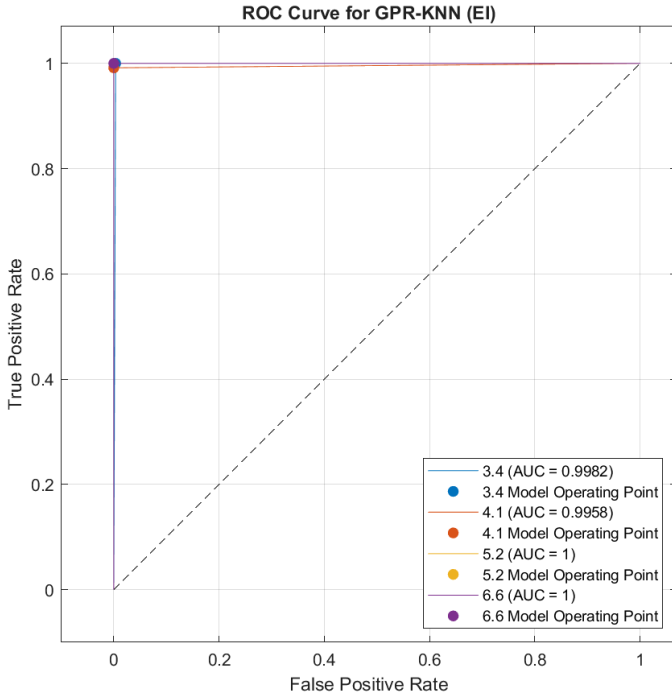


Fig. 4. ROC curve for GPR-KNN (EI)

The positioning of the model's operating points near the ideal upper-left corner of the ROC space underscores the robustness of the threshold settings. This positioning reinforces the model's ability to achieve a high true positive rate while effectively suppressing false positives. Notably, the slightly reduced AUC for class 4.1 indicates some degree of boundary ambiguity in relation to class 3.4, which correlates with minor misclassifications observed in the confusion matrix.

These findings demonstrate that the optimized model possesses outstanding sensitivity, specificity, and operational reliability. This underscores its potential as a highly effective decision-support tool for the optimization of AGL systems under varying meteorological conditions.

Fig. 5 shows the minimum classification error plot for the model that provides essential insights into the hyperparameter optimization process. A notable decline in classification error occurs during the initial five iterations, decreasing from approximately 0.008 to a stabilized plateau between 0.0025 and 0.003. This trend indicates a rapid convergence and effective exploration of the hyperparameter space.

The consistent alignment between the estimated and observed minimum classification errors throughout the optimization process substantiates the robustness and predictive reliability of the surrogate model employed within the Bayesian framework. Importantly, the optimal hyperparameter configuration was identified early, specifically around the fourth or fifth iteration, resulting in minimized computational overhead while maximizing model performance. The findings of this study shows that the optimal hyperparameter

configuration for the model occurs when the number of neighbors is two, the distance metric employed is the Jaccard, and the distance weighting method utilized is the squared inverse. These parameters collectively enhance the model's performance and reliability in analysis.

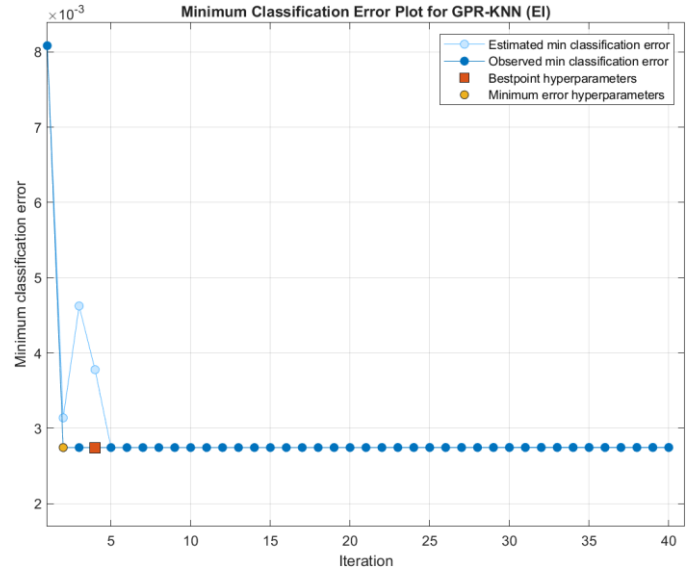


Fig. 5. Minimum classification error plot for GPR-KNN (EI)

These findings underscore the significance of BO in fine-tuning complex ML models. Not only does it facilitate high classification accuracy, but it also ensures operational efficiency, attributes that are essential for the deployment of predictive systems in real-world AGL systems.

IV. VERIFICATION OF MODEL

To assess the model's robustness and ability to generalize, the model was rigorously evaluated using the unseen fifty percent of the collected dataset, comprising 4,380 data points in total, as shown in Fig. 1. The findings from this comprehensive verification phase are meticulously detailed in Table V, highlighting the model's performance and reliability.

TABLE V. VERIFICATION RESULTS

Validation	Acquisition function	Verification
Holdout	EI	99.62%

A verification score of 99.62% highlights the model's exceptional performance on unseen data, underscoring its ability to make highly accurate predictions. This remarkable score illustrates that most predictions generated for the verification set are correct, showing that the model really understands the complicated patterns found in the data. Such an impressive achievement also suggests minimal variation in performance across diverse subsets, indicating a well-rounded and reliable model. This level of accuracy mitigates concerns regarding overfitting and reinforces the confidence in the model's capability to generalize effectively to new, unseen data. It reflects that the model is not merely memorizing the

training examples but genuinely capturing the relationships embedded within the data distribution, showcasing its potential for real-world applications.

To further analyze the impact of the model, a detailed analysis on the total current consumption for every hour has been conducted, as shown in Table VI. For 22 out of 24 hours, the total predicted current consumption matches the measured values, yielding a difference of 0 A. These findings indicate that the model has effectively captured the underlying patterns and temporal variations in current consumption, achieving near-perfect predictive accuracy. Only at 0500 and 0700 hours are there deviations between the total predicted and actual values. At 0500, the model underpredicts by 10.5 A, which is good for lowering the current consumption. However, at 0700, the model overpredicts by 28.8 A. These minor deviations could be due to transient fluctuations or anomalies in the current consumption that the model, trained on the general patterns, did not capture perfectly.

TABLE VI. ANALYSIS OF TOTAL CURRENT CONSUMPTION BASED ON TIME

Time	Actual Total Current Consumption (A)	GPR-KNN Total Current Prediction (A)	Difference (A)
0:00	681.5	681.5	0
1:00	645.1	645.1	0
2:00	572.4	572.4	0
3:00	504.1	504.1	0
4:00	534	534	0
5:00	596.2	585.7	10.5
6:00	629.9	629.9	0
7:00	746.7	775.5	-28.8
8:00	1010.8	1010.8	0
9:00	952	952	0
10:00	1026.2	1026.2	0
11:00	1077.2	1077.2	0
12:00	999.8	999.8	0
13:00	1036.4	1036.4	0
14:00	1263.4	1263.4	0
15:00	1087	1087	0
16:00	1120.4	1120.4	0
17:00	1065.2	1065.2	0
18:00	1152.2	1152.2	0
19:00	1045.6	1045.6	0
20:00	833.4	833.4	0
21:00	781.7	781.7	0
22:00	631.9	631.9	0
23:00	793.6	793.6	0

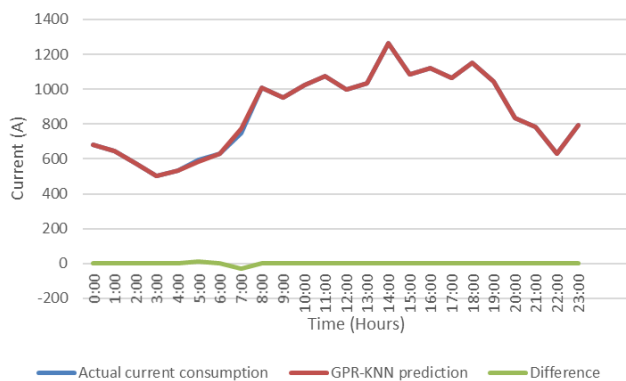


Fig. 5. Comparison chart for actual total current consumption and predicted total current consumption from 0:00 Until 23:00 during verification phase.

From Fig. 5, it can be observed that the total current consumption values fluctuate in a manner that is typical for energy usage over a day, with lower total consumption in the early morning hours, rising gradually through the morning, peaking during midday and early afternoon, and then tapering off towards the evening. For instance, at midnight (0000), the total consumption is 681.5 A, followed by a steady decrease to 504.1 A by 0300 hour, and then a gradual increase, reaching a peak of 1263.4 A at 1400 hour. Such trends suggest that the underlying processes driving current consumption are likely tied to daily human activity patterns and operational schedules of industrial or commercial equipment. From the figure, it can be deduced that the GPR-KNN hybrid model with EI acquisition function has an impressive overall performance, with a near-perfect alignment between actual measurements and predicted values across most hours of the day. The model's exceptional accuracy, demonstrated by zero differences for most of the time points, speaks to the effectiveness of the hybrid approach in capturing both global trends and local variations in current consumption. There are minor deviations at 0500 and 0700 hours that could be improved. However, these differences are small and do not significantly affect overall performance.

V. CONCLUSION

This study presents an innovative algorithm that significantly enhances AGL systems through the integration of GPR and KNN with EI acquisition function. This robust combination not only improves predictive accuracy but also optimizes lighting performance by factoring in critical elements such as time of day and prevailing weather conditions, as articulated in reference [1]. Remarkably, this approach attained a verification accuracy of 99.62%, underscoring the efficacy of the proposed model. The effective implementation of the EI function plays a pivotal role in balancing exploration and exploitation using BO, thus enabling the model to discern complex patterns within extensive datasets. Moreover, continuous monitoring and iterative model enhancement are paramount to ensuring consistent and reliable performance across varying environmental contexts. This research signifies a substantial advancement in predicting current consumption within AGL systems, effectively mitigating power waste and laying a strong foundation for future advancements in ML. The hybrid GPR-KNN model, improved by the EI acquisition function, has demonstrated significant potential for real-time energy monitoring and load forecasting. Nonetheless, ongoing research remains essential to adapt to the dynamic nature of operational conditions. Ultimately, the findings of this study are critical in boosting advancements in predictive analytics and operational efficiency, for a smarter and more sustainable energy future.

ACKNOWLEDGMENT

This study is sponsored by the Public Service Department of Malaysia and the Civil Aviation Authority of Malaysia, to which the authors are very grateful. We would also like to thank the local airport operator for supporting and participating in this study.

REFERENCES

- [1] A. Derraa, N. Ouaaline, and B. Nassiri, "A new airfield lighting system network architecture," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 4, 2024, doi: 10.11591/ijece.v14i4.pp3607-3615.
- [2] International Civil Aviation Organization (ICAO), *Aerodrome Design and Operations*, Ninth Edit., vol. I, no. Ninth Edition. ICAO, 2022.
- [3] International Civil Aviation Organization (ICAO), *Aerodrome Design Manual Part 4: Visual Aids*. 2006.
- [4] I. Metsä-Eerola, J. Pulkkinen, O. Niemitalo, and O. Koskela, "On hourly forecasting heating energy consumption of HVAC with recurrent neural networks," *Energies*, vol. 15, no. 14, 2022, doi: 10.3390/en15145084.
- [5] G. Vontzos, V. Laitsos, and D. Bargiotas, "Data-driven airport multi-step very short-term load forecasting," 14th Int. Conf. Information, *Intell. Syst. Appl. IISA 2023*, 2023, doi: 10.1109/IISA59645.2023.10345889.
- [6] N. Papagrigroriou, G. Palantzas, and D. Nalmpantis, "Airports and environmental sustainability: a review," *E3S Web Conf.*, vol. 436, 2023, doi: 10.1051/e3sconf/202343611006.
- [7] H. Amini, K. Alanne, and R. Kosonen, "Building simulation in adaptive training of machine learning models," *Autom. Constr.*, vol. 165, 2024, doi: 10.1016/j.autcon.2024.105564.
- [8] Z. Wang, J. Liu, Y. Zhang, H. Yuan, R. Zhang, and R. S. Srinivasan, "Practical issues in implementing machine-learning models for building energy efficiency: Moving beyond obstacles," *Renew. Sustain. Energy Rev.*, vol. 143, 2021, doi: 10.1016/j.rser.2021.110929.
- [9] R. Djimasbe, S. Gyamfi, C. D. Iweh, and B. N. Ribar, "Development of an ARIMAX model for forecasting airport electricity consumption in Accra-Ghana: The role of weather and air passenger traffic," *e-Prime - Adv. Electr. Eng. Electron. Energy*, vol. 9, 2024, doi: 10.1016/j.prime.2024.100691.
- [10] X. Chen, T. Guo, M. Kriegel, and P. Geyer, "A hybrid-model forecasting framework for reducing the building energy performance gap," *Adv. Eng. Informatics*, vol. 52, 2022, doi: 10.1016/j.aei.2022.101627.
- [11] E. Sarmas, A. Forouli, V. Marinakis, and H. Doukas, "Baseline energy modeling for improved measurement and verification through the use of ensemble artificial intelligence models," *Inf. Sci. (N.Y.)*, vol. 654, 2024, doi: 10.1016/j.ins.2023.119879.
- [12] A. Hussien, A. Maksoud, A. Al-Dahhan, A. Abdeen, and T. Baker, "Machine learning model for predicting long-term energy consumption in buildings," *Discov. Internet Things*, vol. 5, 2025, doi: 10.1007/s43926-025-00115-7.
- [13] K. Jraida, A. Mourid, Y. El Mghouchi, C. Haidar, and M. El Alami, "Comparative study of hybrid machine learning models to predict the energy consumption of buildings enhanced with PCM: A case study in Morocco," *Appl. Chem. Eng.*, vol. 8, 2025, doi: 10.59429/ace.v8i1.5593.
- [14] W. M. R. Jamaludin, W. M. W. Mohamed, N. H. N. Ali, and N. A. M. Isa, "Impact of incandescent light and LED on electricity fee and carbon emission cost at an airport in Malaysia," *Int. Conf. Power Eng. App.* 2023, doi: 10.1109/ICPEA56918.2023.10093154.
- [15] "Aviation Weather Center." <https://aviationweather.gov/data/metar/> (accessed Jan. 1, 2023).
- [16] N. Penov and G. Guerova, "Sofia Airport Visibility Estimation with Two Machine-Learning Techniques," *Remote Sens.*, vol. 15, 2023, doi: 10.3390/rs15194799.
- [17] S. Cornejo-Bueno, D. Casillas-Pérez, L. Comejo-Bueno, M. I. Chidean, A. J. Caamaño, J. Sanz-Justo, C. Casanova-Mateo and S. Salcedo-Sanz, "Persistence analysis and prediction of low-visibility events at Valladolid Airport, Spain," *Symmetry*, vol. 12, 2020, doi: 10.3390/sym12061045.
- [18] F. M. Cordeiro, G. B. França, F. L. de A. Neto, and I. Gultepe, "Visibility and ceiling nowcasting using artificial intelligence techniques for aviation applications," *Atmosphere (Basel)*, vol. 12, 2021, doi: 10.3390/atmos12121657.
- [19] P. K. Kanti, P. Sharma, B. Koneru, P. Banerjee, and K. D. Jayan, "Thermophysical profile of graphene oxide and mxene hybrid nanofluids for sustainable energy applications: Model prediction with a Bayesian optimized neural network with k-cross fold validation," *FlatChem*, vol. 39, 2023, doi: 10.1016/j.flatc.2023.100501.
- [20] M. Zubair and Y. Akram, "Enhancement in the safety and reliability of Pressurized Water reactors using Machine Learning approach," *Ann. Nucl. Energy*, vol. 201, 2024, doi: 10.1016/j.anucene.2024.110448.
- [21] M. Zubair and Y. Akram, "Utilizing MATLAB machine learning models to categorize transient events in a nuclear power plant using generic pressurized water reactor simulator," *Nucl. Eng. Des.*, vol. 415, 2023, doi: 10.1016/j.nucengdes.2023.112698.
- [22] F. Mohammadi, H. Teiri, Y. Hajizadeh, A. Abdolnnejad, and A. Ebrahimi, "Prediction of atmospheric PM2.5 level by machine learning techniques in Isfahan, Iran," *Sci. Rep.*, vol. 14, 2024, doi: 10.1038/s41598-024-52617-z.
- [23] W. M. R. Bin Jamaludin, W. M. B. Wan Mohamed, N. H. Bin Nik Ali, and N. A. Binti Mohd Isa, "Utilizing Advanced Regression Techniques to Forecast Visibility at Subang and Langkawi International Airport," *New Trends Civ. Aviat.*, 2024, doi: 10.23919/NTCA60572.2024.10517824.
- [24] W. M. R. Jamaludin, N. H. N. Ali, S. Member, and W. M. W. M., "Optimizing Nominal Current Output for Aeronautical Ground Lighting Using Machine Learning and Meteorological Data," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3427402.
- [25] H. Zhou, X. Ma, and M. B. Blaschko, "A Corrected Expected Improvement Acquisition Function Under Noisy Observations," in *Proceedings of Machine Learning Research*, 2023.
- [26] V. Nguyen, S. Gupta, S. Rana, C. Li, and S. Venkatesh, "Regret for expected improvement over the best-observed value and stopping condition," *J. Mach. Learn. Res.*, vol. 77, 2017.
- [27] S. Ament, S. Daulton, D. Eriksson, M. Balandat, and E. Bakshy, "Unexpected Improvements to Expected Improvement for Bayesian Optimization," in *Advances in Neural Information Processing Systems*, 2023.
- [28] Z. Chen, S. Mak, and C. F. J. Wu, "A Hierarchical Expected Improvement Method for Bayesian Optimization," *J. Am. Stat. Assoc.*, vol. 119, no. 546, 2024, doi: 10.1080/01621459.2023.2210803.
- [29] C. Yan, H. Du, E. Kang, D. Mi, H. Liu, and Y. You, "AVEI-BO: an efficient Bayesian optimization using adaptively varied expected improvement," *Struct. Multidiscip. Optim.*, vol. 65, 2022, doi: 10.1007/s00158-022-03256-3.
- [30] L. Zhang, G. Jin, T. Liu, and R. Zhang, "Generalized hierarchical expected improvement method based on black-box functions of adaptive search strategy," *Appl. Math. Model.*, vol. 106, 2022, doi: 10.1016/j.apm.2021.12.041.
- [31] International Civil Aviation Organization (ICAO), *Aerodrome Design Manual Part 5: Electrical System*. 2006.
- [32] O. M. Almorabea, T. J. S. Khanzada, M. A. Aslam, F. A. Hendi, and A. M. Almorabea, "IoT Network-Based Intrusion Detection Framework: A Solution to Process Ping Floods Originating from Embedded Devices," *IEEE Access*, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3327061.