

Predictive Power-Demand Modelling for Adaptive Street Lighting Using Random Forest Regression

Aditya Krishan Goel

The Shri Ram School, Aravali
Gurugram, Haryana, India
aditya.goel@outlook.in

Abstract

Street lighting is one of the most energy-intensive municipal services, largely because most installations run at a fixed static output throughout the night regardless of conditions. This study asks a simple question: can streetlight-power demand be predicted from weather, traffic, and temporal context, and can such a model drive an adaptive-dimming controller capable of reducing consumption by a significant figure? Using four months of feeder-level data from twelve smart-lighting controllers in Delhi (30 November 2024 to 31 March 2025; approximately 143,800 fifteen-minute records), that can lead to algebraic deduction being removed and using only contextual features under a strict chronological eighty-twenty train-test split, a Random Forest predicts active power with a mean absolute error of about one kilowatt and an R-squared of 0.895, outperforming static, gradient-boosted, linear, and hour-mean baselines. Adding weather and traffic improves prediction only modestly. The model is then coupled to a luminance-anchored dimming controller whose floor equals the ratio of two road-lighting class luminances, and a luminance check verifies that road luminance never falls below the class minimum. On the held-out month, the controller reduces operating-hour power by about twenty-nine percent, a saving reproduced almost exactly when the same policy is applied to actual metered power, while remaining fully luminance-compliant. The value of the added data lies not in sharpening prediction but in enabling demand-aware control.

Keywords

Adaptive street lighting, power-demand prediction, Random Forest regression, demand-responsive dimming, and energy efficiency.

1. Introduction

In recent years urban electricity demand has risen sharply with economic and population growth, placing public lighting as one of the largest controllable loads that a municipality manages (Pasolini et al. 2024). Conventionally, these installations are static: the lamps switch on at dusk, holding a fixed output till dawn, irrespective of any factor such as traffic, weather, or pedestrians, which wastes energy during the various low-activity hours at night. The reported figures for the share of municipal electricity bills attributable to public lighting vary greatly with city size and electrification, and the often-quoted forty-to-sixty percent range (Pasolini et al. 2024) should be taken as an upper-bound estimate for those municipalities that are lighting-heavy rather than as a universal constant; we therefore treat it as indicative.

A recurring ambiguity in this literature is the conflation of two distinct problems. The first is prediction: the estimation, from observable conditions, of how much power a lamp will draw. The second is control: deciding how bright a lamp should be set depending on the time of day. A model that predicts demand under the current control regime does not, by itself, prescribe the dimming actions, and much of the prior work blurs this important distinction.

Therefore, the research question is twofold: (i) to what extent can feeder-level streetlight power be predicted from contextual features alone, once all the features that trivially determine the target are removed; and (ii) can the resulting model be used, within a safety floor constraint expressed in photometric terms, to create a dimming policy capable of

demonstrating energy-use reduction. Our working thesis is that contextual prediction is accurate enough to support scheduling and control, but that the contextual features add little to point accuracy beyond the day-night signal, so their value is seen in control rather than in regression.

1.1 Objectives

This paper makes the following contributions:

1. It diagnoses and removes a data-leakage pathway and establishes a performance baseline under a strict chronological train-test split.
2. It quantifies, by ablation, the marginal value of weather and traffic features for power prediction.
3. It couples the demand model to a luminance-anchored dimming controller in which the model gates operation and supplies the demand baseline, closing the prediction-to-control gap.
4. It anchors the dimming floor for a road-lighting class and verifies, by a luminance check, that the dimmed floor never falls below the class safety minimum.
5. It reports an energy saving that is validated out-of-sample against actual metered power and presented as a simulation that awaits field deployment.

The scope is deliberately limited to a single four-month, cool-season dataset of two Delhi localities, and to a simulated controller rather than a hardware deployment; these limitations are stated explicitly and revisited in the discussion.

2. Literature Review

Approaches to adaptive street lighting span a wide range, including rule-based dimming, sensor-threshold control, and learning-based optimisation, including random-forest models with permutation-importance feature selection developed for short-term load forecasting (Huang et al. 2016). Field and simulation studies consistently report meaningful and substantial energy savings, but under heterogeneous methods, reporting conventions, and baseline values, which complicates direct comparison. Large-scale, long-term implementation governed by the European EN 13201 standard and its Italian extension UNI 11248 reports savings of about forty-eight percent in a measured municipal pilot, and an even higher saving of up to seventy percent when the lighting class is reduced by several levels in low-traffic conditions (Pasolini et al. 2024; European Committee for Standardization 2015). Real-time hybrid architectures that dim based on traffic and weather over low-power radio report savings of about fifty-five percent (Takruri et al. 2023), and traffic-model-driven simulation studies report figures ranging from thirty to sixty-six percent. From these studies, two observations follow. First, headline savings in this field are often achieved by simpler, sensor- or scheduler-based methods (Pasolini et al. 2024; Takruri et al. 2023), so a learning-based method must justify itself on grounds other than raw savings — for example robustness, principled safety bounding, or transferability. Second, several widely cited results are derived from simulations rather than from metered deployment, a distinction this paper carefully preserves. Table 1 compares and situates the current study against prior representative work (Pasolini et al. 2024; Takruri et al. 2023; Godara 2024).

Table 1. Comparison of the current study against prior representative work

Study (representative)	Method	Setting / data	Reported saving	Limitation relevant here
Context-adaptive lighting, large-scale (2024)	Class-reduction dimming under EN 13201 / UNI 11248	Measured municipal pilots (e.g., 878 lamps)	~48% measured; up to 70% by class reduction	Requires real-time sensing; standard-specific
Real-time hybrid control (2023)	Traffic + weather dimming over NB-IoT	Experimental deployment	Up to ~55%	Hardware-specific; limited statistical evaluation
Traffic-model dimming (review)	Greenshields traffic model, SUMO simulation	Washington DC / Montgomery County	Up to ~66%	Simulation only
Environment-based adaptive control	PWM dimming on traffic sensing	SUMO / OMNeT++ simulation	Up to ~30% vs state of the art	Simulation only
This study	Leakage-controlled RF demand model +	12 feeders, Delhi, 4 months, metered	~29% operating-hour (validated	Single season; typical traffic; not yet field-metered

	luminance-anchored dimming		out-of-sample, simulated control)	
--	-------------------------------	--	--------------------------------------	--

3. Methods

3.1 Feature set and leakage control

A sanity check confirmed that the target is reconstructed algebraically by the electrical channels. The sum of the three per-phase active powers reproduces total active power at an R-squared of 0.9998, and apparent power multiplied by the magnitude of the power factor reproduces it at 0.9985; both follow from the identity that active power equals voltage times current times power factor. A model given these electrical channels learns arithmetic, not context, so an R-squared near unity is trivial; this is a recognised pitfall, in which information unavailable at prediction time leaks into the features (Kaufman et al. 2012). All thirty-eight such electrical features were removed. The retained predictors are the external weather variables, two traffic variables, cyclic encodings of hour-of-day and day-of-year (sine and cosine), a binary precipitation flag derived from rain, one-hot encodings of controller identity and locality, and a day-of-week index and weekend flag. A complete feature list with the rationale for each engineering decision is provided in the Appendix.

3.2 Chronological split and demand model

Since the data form a time series, a strict chronological split was used: the earliest eighty percent of readings (30 November 2024 to 7 March 2025) for training, and the final twenty percent (7 March to 31 March) for testing, with no shuffling, so that every test timestamp follows every training timestamp. This avoids the leakage that a random eighty/twenty split introduces when adjacent readings from the same night are divided between train and test. Hyperparameters were tuned by grid search under a time-series cross-validation that preserves ordering within each fold. The predictor is a Random Forest (Breiman 2001) of three hundred trees, regularised to a maximum depth of sixteen with a minimum of five samples per leaf to reduce the gap between train and test.

3.3 Baselines

The model is compared against four baselines on the same held-out month: an hour-of-day mean, a gradient-boosted-tree regressor, a static rated-power schedule, and an ordinary linear regression on the same features.

3.4 Operating-schedule characterisation

To characterise the system under control, the daily turn-on and turn-off times were estimated. From December to March, the evening turn-on drifts about seventy minutes later, closely tracking the seasonal sunset shift, and the morning turn-off drifts earlier with sunrise. This pattern shows no response to same-day cloud cover, indicating that a pre-programmed astronomical schedule is followed rather than live light sensing. Once on, these lamps draw an almost constant power, confirming that the present installation performs no intensity dimming; this is the headroom the controller targets.

3.5 Closed-loop control architecture

The predictor and the controller are linked as follows. The operating window is decided by the model's forecast of demand: a lamp is treated as operating when predicted power exceeds a threshold, so the predictor inherits the dusk-to-dawn logic from the radiation, identity, and time features it uses. During operation, the dimming policy sets the intensity between a floor and full output as a function of demand and safety conditions: with rising traffic congestion, intensity rises from the floor to full output; it is boosted to full output when atmospheric visibility falls into the haze range; and it is reduced to the floor in dense fog, on the principle that strong light backscatters and does not aid visibility. The controlled power is the predicted baseline scaled by an intensity-to-power map that includes a fixed driver overhead of ten percent so that savings are conservative. The saved energy is the difference between the predicted baseline and the controlled power. The predictor therefore contributes the operating-window forecast and the demand baseline that make the schedule deployable from a forecast; the magnitude of the savings is determined by the safety-bounded control policy. Here, condition-responsive control denotes a schedule and dimming policy driven by forecastable context and a deterministic traffic profile, rather than a closed-loop response to real-time vehicle counts, which is left to future work.

3.6 Luminance-class anchoring and safety check

Illumination standards specify a minimum maintained luminance, in candela per square metre, for each road-lighting class, not an absolute power in kilowatts; the conversion depends on fixture, mounting height, and road geometry.

Accordingly, the dimming floor is expressed as a luminance ratio rather than a fixed power. Taking a design class of M3 (1.00 candela per square metre at full output) and a permitted night-minimum class of M5 (0.50 candela per square metre) from the EN 13201 / CIE 115 ladder, and noting that road-surface luminance scales with luminous flux, the floor is their ratio, fifty percent (European Committee for Standardization 2015). A luminance check then computes the revised luminance at every operating hour as intensity times design-class luminance and verifies that it never falls below the class minimum; since intensity is bounded below by the floor, compliance is guaranteed by construction and reported explicitly. For simplicity, a single common class is assumed for both localities; in practice the class, and hence the floor, should differ by road type. The Indian standard IS 1944 provides the national analogue, expressed in lux by road group, and the same ratio method applies to those values (Bureau of Indian Standards 2010).

3.7 Energy-savings estimation and ethics

On the held-out month, two saving figures are reported: a planned saving computed from the model's forecast baseline, and a revised saving computed by applying the identical policy to the actual metered power. Agreement between them confirms that the forecast-driven plan is achievable on unseen data. This study uses secondary feeder-level data and public environmental and traffic data; it involves no human participation, and no personally identifiable information is used.

4. Data Collection

The dataset comprises readings from twelve feeder-level smart-lighting controllers in Delhi, India, logged at approximately fifteen-minute intervals from 30 November 2024 to 31 March 2025 (Terra-Byte.ai 2024). Seven of these controllers serve a residential locality (Laxmi Nagar) and five a commercial locality (Mayur Vihar). After removing records with missing timestamps or targets, approximately 143,800 readings remain. Each record reports the target — total active power in kilowatts — together with electrical channels (per-phase and aggregate voltage, current, active, reactive, and apparent power, power factor, frequency, energy, and run-hour) and contextual channels. The magnitude of the per-controller load (medians of twelve to eighteen kilowatts) indicates that each unit meters a street segment of multiple luminaires rather than a single lamp.

The feeder power and electrical readings serve as the primary sensor data; the on-board temperature and humidity channels were disregarded as they were inoperative (constant zero value). Environmental context was instead obtained from a public weather server (Open-Meteo hourly analysis for the site coordinates: temperature, relative humidity, rain, cloud cover, visibility, sunshine duration, and shortwave solar radiation) (Open-Meteo 2024), and traffic context from a Delhi congestion profile (a typical weekday-by-hour congestion index and a monthly level) (TomTom 2024).

5. Results and Discussion

5.1 Numerical Results

As established in Section 3, the algebraic identities reproduce the target almost perfectly (R-squared of 0.9998 for the sum of the phase powers), confirming that near-unity accuracy reflects arithmetic modelling rather than context-awareness; all such features are excluded from the model reported here.

On the held-out month, the Random Forest achieves a mean absolute error of about 1.0 kilowatt, a root-mean-square error of about 2.5 kilowatts, and an R-squared of about 0.895, outperforming every baseline, as summarised in Table 2.

Table 2. Held-out (chronological) test performance against baselines.

Model	MAE (kW)	RMSE (kW)	R ²
Static rated power	19.41	20.84	—
Hour-of-day mean	3.79	5.54	0.466
Linear regression	3.77	4.53	0.644
Gradient-boosted trees	0.98	2.72	0.872
Random Forest (this study)	1.01	2.46	0.895

Regularisation reduces the train-to-test mean-absolute-error ratio to about 1.4 (training error near 0.7 kilowatt against test error near 1.0 kilowatt), down from roughly 1.9 in an unregularised model, indicating mild residual overfitting with acceptable generalisation, since the test error matched the cross-validated error. The mean absolute percentage error is inappropriate here because half of all records are daytime values near zero, which makes the percentage error diverge; restricted to operating hours, it is eleven percent. The primary metrics reported are therefore mean absolute and root-mean-square errors in kilowatts.

Re-fitting with and without the external features, holding the model fixed, improves mean absolute error by about two percent, root-mean-square error by about four percent, and R-squared by about 0.009. The added features therefore deliver only a modest gain, concentrated in the harder high-power cases rather than the median, and are best described as confirming the demand signal rather than transforming prediction accuracy.

5.2 Graphical Results

A permutation importance on the test set ranks solar radiation first, followed by hour-of-day, several controller identifiers, sunshine duration, and traffic congestion. Visibility, cloud, temperature, rain, humidity, grid frequency, and the precipitation flag carry essentially no importance. The leading rank of solar radiation is a collinearity artefact: in a model without weather, the hour-of-day cosine alone carries the day-night signal, and adding radiation — which also encodes whether it is dark — merely redistributes that importance. Traffic congestion correlates equally with power and with hour-of-day, so its apparent value is largely a proxy for time-of-day. None of the contextual features are leaky, as all are exogenous to the meter.

Figure 1 presents the held-out distribution of predicted versus actual power and the mean absolute error stratified by hour band and locality. The error is concentrated in the evening peak (eighteen to twenty-one hours), where power is highest and most variable, and is far smaller during the daytime-off and deep-night periods; the commercial locality shows slightly higher peak-hour error than the residential one.

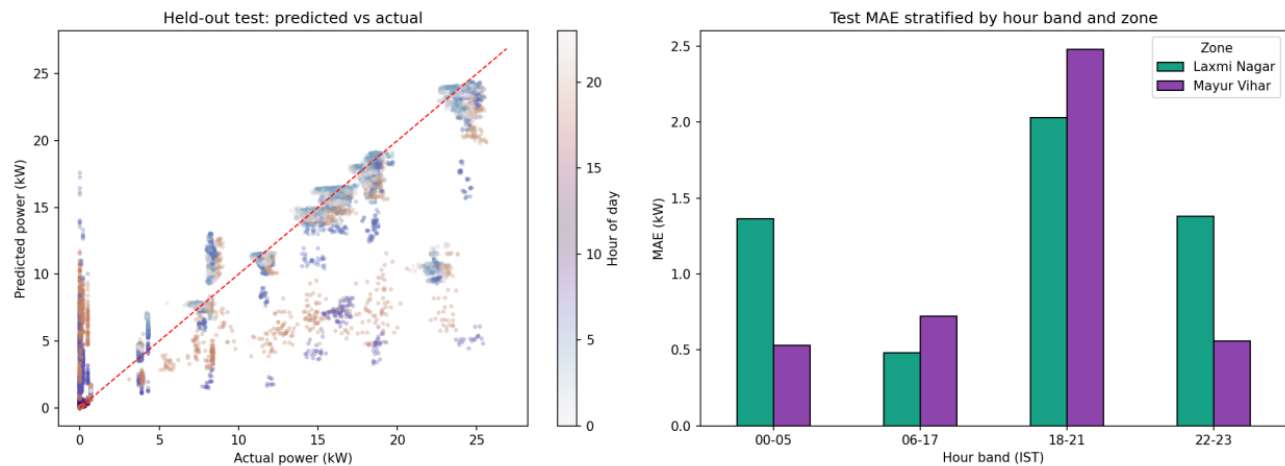


Figure 1. Held-out predicted vs actual power (left) and test MAE stratified by hour band and zone (right).

With the M3-to-M5 anchoring (a fifty-percent floor), the controller dims most in the empty small hours and holds full output during the evening traffic peak. During the held-out month, the simulated controller achieved an estimated operating-hour saving of approximately 28.9 percent using the forecast baseline. Applying the same dimming policy to the actual metered power also gave 28.9 percent to one decimal place, providing a retrospective validation of the control policy on unseen data. The mean luminosity during operating hours is sixty-eight percent of the design output — an average luminosity reduction of about thirty-two percent — reaching the fifty-percent floor in the deep-night low-activity window. The luminance check is satisfied at every operating hour: the minimum seen road luminance is exactly 0.50 candela per square metre, equal to and never below the M5 class minimum, giving one-hundred-percent compliance. Figures 2 and 3 show the closed loop over a representative four-day window and summarise the savings and luminance compliance.

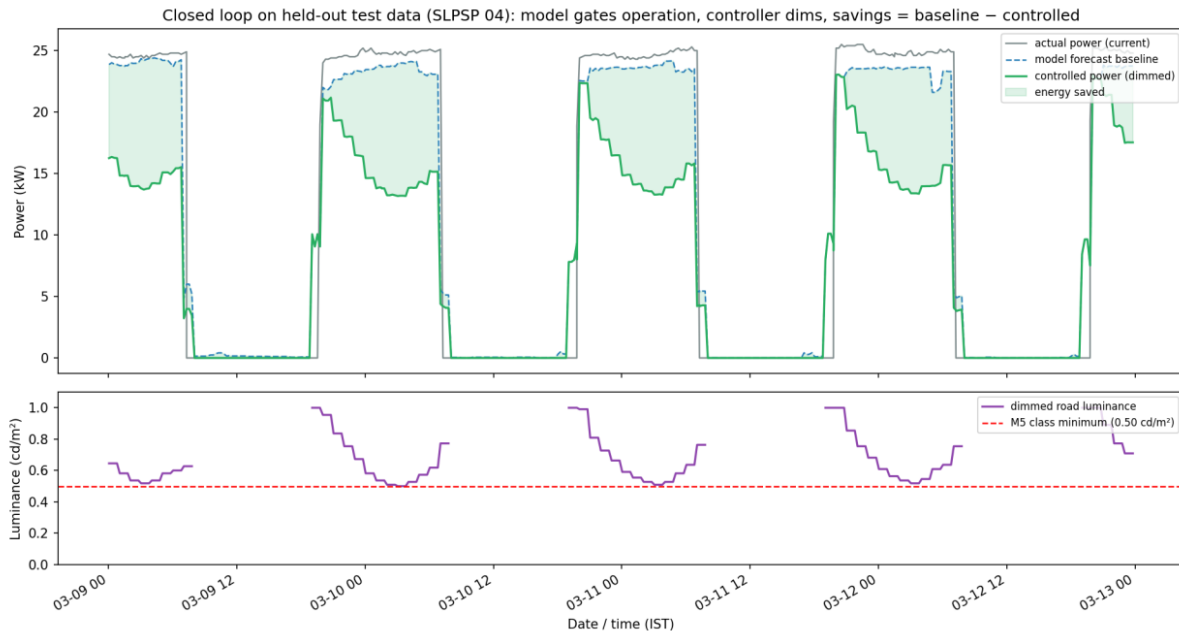


Figure 2. Closed loop on held-out data: the model forecast (dashed) sets the baseline, the controller dims (green), and the shaded area is energy saved; the lower panel shows luminance held at or above the M5 minimum (red).

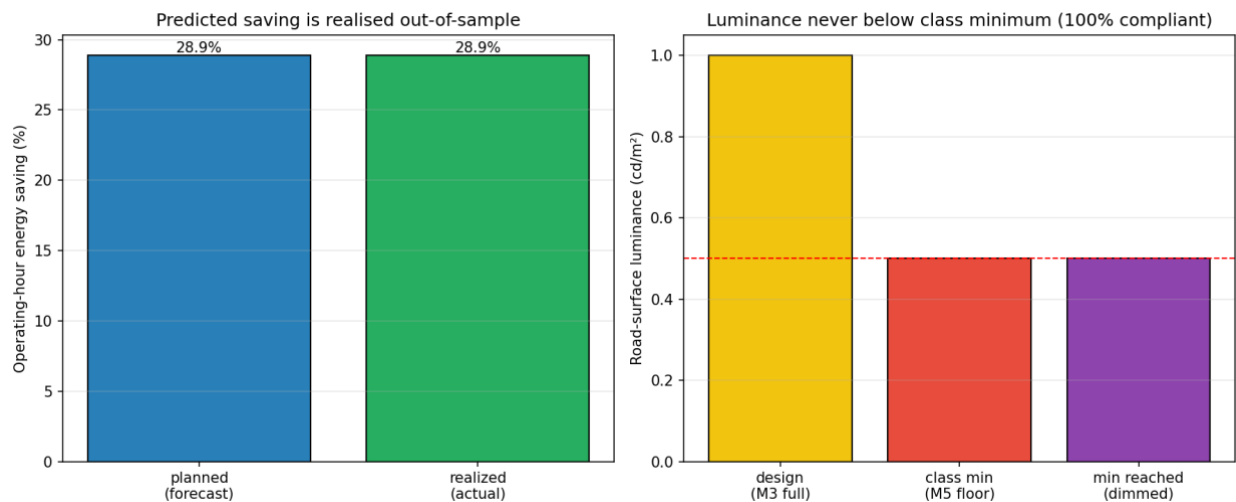


Figure 3. Planned versus revised operating-hour savings (left) and luminance compliance against the class minimum (right).

5.3 Discussion

The added weather and traffic features do not materially sharpen point prediction; their value is that they make demand explainable from forecastable, controllable inputs, which is precisely what enables scheduling and control. The role of the model in the loop is to forecast the operating window and the demand baseline, so that a schedule and an expected saving can be computed in advance from the weather-and-traffic forecast, and to confirm, through an R-squared of 0.895, that the conditions carry the demand signal that justifies condition-driven control. The magnitude of the savings is set by the safety-bounded control policy, and the agreement between planned and revised savings shows that the operating-window forecast is accurate enough that the plan is achievable.

Positioning against prior work, the present saving of about twenty-nine percent is lower than several headline figures because it is bounded by an explicit luminance floor (Pasolini et al. 2024; Takruri et al. 2023; Godara 2024), computed

conservatively with a fixed driver overhead, validated out-of-sample, and measured against the installation's current constant-output behaviour rather than an idealised always-on baseline. The contribution is therefore methodological rigour and principled safety bounding rather than a claim to maximal savings. Deployment challenges remain substantial: the controller assumes a typical traffic profile rather than real-time counts, a single weather location rather than on-pole sensing, and a linear power-to-output relationship. Plausible failure cases include the seasonally shifting dusk boundary, which a learned controller must track through solar radiation rather than clock time; rare dense-fog events, which are underrepresented in the data; and over-dimming if the floor is set without reference to the road's actual lighting class.

5.4 Validation and Limitations

The planned saving computed from the forecast baseline and the saving computed on the actual metered power agree to one decimal place (28.9 percent), and the luminance check confirms full compliance with the M5 minimum; together these provide an out-of-sample validation of the control policy. This agreement is not an independent test of prediction accuracy: the magnitude of the saving is fixed primarily by the dimming schedule rather than by the forecast, so the planned-versus-revised match confirms that the operating-window forecast and the policy transfer to unseen metered data, not that the predictor itself drives the saving. Several limitations bound the result. The dataset spans four months of the cool, dry season only, so monsoon and summer conditions and a fully representative seasonal range are unobserved, and the test month is a single seasonal slice. The traffic input is a deterministic weekday-by-hour profile rather than real-time data, so the controller is open-loop with respect to traffic (TomTom 2024; Open-Meteo 2024). Weather is drawn from a single grid point, with a thin tail of safety-relevant low-visibility events. The savings are simulations validated against actual load, not a metered field deployment with physically dimmed lamps. A single lighting class is assumed for both localities, whereas a commercial and a residential street should occupy different classes. The power-to-output model is a first-order linear approximation, and mild residual overfitting remains.

6. Conclusion

A leakage-controlled, chronologically validated model predicts feeder-level streetlight power with an R-squared of 0.895, using weather, traffic, and time alone, once the features that algebraically determine the target are removed. Adding weather and traffic improves prediction only modestly, and the apparently dominant solar-radiation feature is a daylight proxy rather than an independent signal. The model's value is seen in control: gating a luminance-anchored dimming policy, it supports a forecast-driven schedule that reduces operating-hour power by about twenty-nine percent out-of-sample, validated against actual metered power and provably bounded by the M5 luminance minimum. The objective of using a data model to reduce streetlight electricity through safe, condition-responsive dimming is thus demonstrated in simulation. The principal next steps towards full validation are a metered field pilot with real-time traffic and on-pole light sensing, reinforcement-learning control that optimises against the demand forecast directly, and an edge implementation on the controllers. The broader implication is that honest, leakage-controlled modelling, coupled to a safety-bounded controller, can turn a familiar prediction task into a deployable, accountable energy-saving tool.

References

- Akbas, D., Gurbuz, T. and Alptekin, G. I., Smart street lighting systems for sustainable smart cities: a conceptual framework, *2024 IEEE Sustainable Smart Lighting World Conference & Expo (LS:24)*, pp. 1–4, 2024.
- Breiman, L., Random Forests, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- Bureau of Indian Standards, Code of practice for lighting of public thoroughfares (IS 1944); National Lighting Code (SP 72), 2010.
- De Paz, J., Bajo, J., Rodríguez, S., Villarrubia, G. and Corchado Rodríguez, J., Intelligent system for lighting control in smart cities, *Information Sciences*, vol. 372, pp. 241–255, 2016.
- European Committee for Standardization, Road lighting — Part 2: Performance requirements (EN 13201-2:2015); see also CIE 115:2010, Lighting of roads for motor and pedestrian traffic, 2015.
- Godara, S., An IoT-based smart street lighting system to lower energy consumption, *Journal of Electrical Systems*, vol. 19, pp. 163–173, 2024.
- Huang, N., Lu, G. and Xu, D., A permutation importance-based feature selection method for short-term electricity load forecasting using random forest, *Energies*, vol. 9, no. 10, 767, 2016.
- Kaufman, S., Rosset, S., Perlich, C. and Stitelman, O., Leakage in data mining: formulation, detection, and avoidance, *ACM Transactions on Knowledge Discovery from Data*, vol. 6, no. 4, article 15, 2012.

- Khan, M. A., Rahman, M. T. and Ahmed, S., Energy-efficient intelligent street lighting system using IoT and adaptive control, *Sustainable Cities and Society*, vol. 45, pp. 401–411, 2019.
- Liu, J., Zhang, H. and Chen, Y., IoT-based smart street lighting system with distributed intelligence, *IET Smart Cities*, vol. 1, pp. 105–114, 2019.
- Open-Meteo, Historical Forecast API, Available: <https://open-meteo.com/en/docs/historical-forecast-api>, Accessed November 30, 2024.
- Pasolini, G., Toppan, P., Toppan, A., Bandiera, R., Mirabella, M., Zabini, F., Bonata, D. and Andrisano, O., Comprehensive assessment of context-adaptive street lighting: technical aspects, economic insights, and measurements from large-scale, long-term implementations, *Sensors*, vol. 24, pp. 5942, 2024.
- Takruri, M., Thulasingham, K. P., Attia, H., Omar, A., Altunaiji, A. and Almaeeni, S., Design and implementation of a real-time street light dimming system based on a hybrid control architecture, *International Journal of Distributed Sensor Networks*, vol. 2023, pp. 1–14, 2023.
- Terra-Byte.ai, AI-powered operational intelligence (data platform), Available: <https://terra-byte.ai/>, Accessed 2024.
- TomTom, Maps and location technology, Available: <https://www.tomtom.com/>, Accessed 2024.
- Wang, L., Li, X. and Zhao, Z., Machine-learning-enabled smart street lighting systems: design and implementation, *Sensors*, vol. 23, 2023.

Biography

Aditya Krishan Goel is a high school researcher and innovator at The Shri Ram School, Aravali, India. His work focuses on creating IOT based intelligent energy systems for sustainable usage. He founded AeiX, an industrial sustainability platform designed to help small & medium-sized enterprises monitor energy consumption, identify inefficiencies, and reduce carbon emissions through data-driven dashboards and analytics. He was selected as a Finalist at the IRIS National Science Fair (Regeneron ISEF affiliated) for his project on sustainable energy systems aligned with UN Sustainable Development Goal 12: Responsible Consumption and Production. He also won Sigma XI 2026 Student Research Showcase Interdisciplinary Award in Engineering Category for research on IoT-based Industrial Energy Monitoring & Carbon Visibility. He authored a research paper titled *AI and Machine Learning Driven Intelligent Streetlight System for Sustainable Energy Efficiency and Smart Urban Management*, which proposes adaptive lighting systems that adjust illumination based on traffic density, environmental conditions, and time-based patterns to reduce energy consumption while maintaining public safety. Aditya has conducted mathematical research on chaotic systems through his paper *Spreadsheet-Based Coupling of Henon and Logistic Maps*, where he developed a model coupling chaotic maps using variable coupling factors to study synchronization and complex system behavior. The research was published in the *International Journal of Mathematics Trends and Technology*. Aditya has contributed to mechanical engineering innovation. He designed an automatic thread-pulling machine to separate flat-knitted garment components, improving manufacturing efficiency and worker safety. A patent application for this system was filed in February 2025. Aditya has gained industry exposure through technical internships at TerraTech (Dubai), Prakriti Creations, and Vandana Solvex, where he worked on energy systems, machine automation, and industrial process optimization.

Appendix: Feature List and Engineering Rationale

The model uses contextual and temporal features only; all electrical channels were removed as leakage. Identifiers are one-hot encoded; time is cyclically encoded.

Appendix Table. Retained features and the rationale for each (electrical channels excluded as leakage).

Feature	Type / encoding	Rationale
Hour-of-day (sin, cos)	Cyclic	Captures the dominant day–night demand cycle without a 23→0 discontinuity.
Day-of-year (sin, cos)	Cyclic	Captures seasonal drift in dusk/dawn within the observed window.
Day-of-week; weekend flag	Integer / binary	Captures weekly activity patterns.
Solar radiation; sunshine duration	Numeric (weather)	Daylight proxies; strongly associated with on/off, though collinear with hour.

Temperature; relative humidity	Numeric (weather)	Tested for independent effect; found weak after time and radiation.
Rain; precipitation flag; cloud cover	Numeric / binary (weather)	Weather variation; low predictive value but retained for completeness and the fog rule.
Visibility	Numeric (weather)	Used by the controller's safety overlay; negligible for power prediction.
Traffic congestion; monthly level	Numeric (traffic)	Activity proxy; collinear with time-of-day.
Controller identity; locality	One-hot	Captures per-feeder connected load and road type.

Code and data availability. The analysis code and processed dataset are available at <https://github.com/adityakgoel/StreetLighting.git>