# An Uncertainty-Aware Measurement and Verification Framework for Building Energy Savings Using Cross-Validated Regression and Symbolic Regression via Genetic Programming

**Hazem Hassan Ali Sayed**
Energy Management Department, Luminous Energy Solutions, Calgary, Canada
Renewable Energy Department, New Cairo Technological University, Egypt

***Abstract.*** Quantifying energy savings from building retrofits is complicated by uncertainty in baseline model form, since agreement between two conventionally linear tools may simply reflect a shared assumption rather than genuine model adequacy. The main objective of this study is to develop and validate a general, uncertainty-aware measurement and verification (M&V) framework that resolves this ambiguity through three-way cross-validation, demonstrated via a lighting retrofit case study in a commercial building in London, Ontario, Canada. This objective was achieved by analyzing three years (2021–2023) of monthly electricity consumption alongside Cooling Degree Days (CDD) using linear regression aligned with IPMVP Option C, applying Cumulative Summation (CUSUM) analysis to track savings persistence, and cross-validating the baseline model across three independent methods: commercial M&V software (RETScreen Expert), custom numerical regression (MATLAB), and a structure-free evolutionary search (Symbolic Regression via Genetic Programming, GP) that imposes no predetermined functional form. The most important results are that all three methods converged on the same linear CDD-based baseline model ($R^2$ = 0.9762–0.9765, $\Delta R^2 \leq 0.0003$), confirming that this agreement reflects genuine model adequacy rather than shared linear bias. The case study yielded verified annual savings of 232.056 kWh (27.846.72 CAD; 23.57 $tCO_2$), with CUSUM confirming persistence, though the high fit obtained reflects this facility's specific consumption profile and is addressed as a limitation. The significance of the obtained results lies in demonstrating a replicable, uncertainty-quantified cross-validation methodology that practitioners can apply beyond the specific case examined.
***Keywords*:** Symbolic regression, genetic programming, cooling degree days, cumulative sum tracking, cross-platform validation, emission verification.

**DOI: https://doi.org/10.52254/1857-0070.2026.4-72.06**
**UDC: 004.8:620.9:628.9:519.6**

**O arhitectură de măsurare și verificare a economiilor de energie în clădiri, care ține cont de incertitudine, bazată pe regresia cu validare încrucișată și regresia simbolică prin programare genetică**

**Hazem Hassan Ali Sayed**
Departamentul de Management al Energiei, Luminous Energy Solutions, Calgary, Canada
Departamentul de Energie Regenerabilă, Universitatea Tehnologică din New Cairo, Egipt

***Rezumat.*** Cuantificarea economiilor de energie rezultate în urma modernizării clădirilor este complicată de incertitudinea asociată formei modelului de referință (*baseline*), întrucât concordanța dintre două instrumente bazate pe modele liniare convenționale poate reflecta mai degrabă utilizarea unei ipoteze comune decât adecvarea reală a modelului. Obiectivul principal al acestui studiu este dezvoltarea și validarea unui cadru general de măsurare și verificare (M&V), care să țină cont de incertitudine și să elimine această ambiguitate printr-o metodă de validare încrucișată tripartită. Abordarea este ilustrată printr-un studiu de caz privind modernizarea sistemului de iluminat într-o clădire comercială din London, Ontario, Canada. Pentru atingerea acestui obiectiv, au fost analizate datele lunare privind consumul de energie electrică din perioada 2021–2023, corelate cu indicele *Cooling Degree Days* (CDD). Analiza a inclus aplicarea regresiei liniare, în conformitate cu Opțiunea C a Protocolului Internațional de Măsurare și Verificare a Performanței (IPMVP), utilizarea analizei sumelor cumulate (CUSUM) pentru monitorizarea persistenței economiilor și validarea încrucișată a modelului de referință prin trei metode independente: software-ul comercial de M&V RETScreen Expert, regresia numerică personalizată implementată în MATLAB și o metodă de căutare evolutivă independentă de o structură prestabilită, bazată pe regresia simbolică prin programare genetică (GP), care nu impune o formă funcțională predeterminată. Rezultatele principale arată că toate cele trei metode au condus la același model liniar de referință bazat pe CDD, cu valori ale coeficientului de determinare R2 cuprinse între 0.9762 și 0.9765 și o diferență maximă $\Delta$R2 de 0.0003. Importanța rezultatelor constă în demonstrarea unei metodologii reproductibile de validare încrucișată, concepute pentru a lua în considerare incertitudinea și care poate fi aplicată de specialiști în contexte diferite, dincolo de cazul particular analizat.
***Cuvinte-cheie*:** regresie simbolică, programare genetică, grade-zile de răcire, monitorizare prin sume cumulate, validare cross-platformă, verificarea emisiilor.



© Hazem Hassan Ali Sayed, 2026

# Методика измерения и верификации энергосбережения в зданиях с учетом неопределенности, использующая регрессию с перекрестной проверкой и символьную регрессию на основе генетического программирования

**Хазем Хассан Али Сайед**
Отдел управления энергопотреблением, Luminous Energy Solutions, Калгари, Канада
Отдел возобновляемой энергетики, Технологический университет Нового Каира, Египет

***Аннотация.*** Оценка экономии энергии при модернизации зданий осложняется неопределенностью в отношении вида базовой модели, поскольку совпадение результатов двух инструментов, использующих линейные зависимости, может отражать лишь общую априорную предпосылку, а не реальную адекватность модели. Основная цель данного исследования — разработка и валидация универсальной методики измерения и верификации (M&V), учитывающей неопределенность и устраняющей указанную неоднозначность посредством перекрестной проверки по трем направлениям; эффективность подхода продемонстрирована на примере модернизации системы освещения в коммерческом здании (г. Лондон, провинция Онтарио, Канада). Для достижения этой цели был проведен анализ данных об ежемесячном потреблении электроэнергии и показателях «градусо-суток охлаждения» (CDD) за три года (2021–2023 гг.) с использованием линейной регрессии (в соответствии с опцией C протокола IPMVP); также применялся метод кумулятивных сумм (CUSUM) для отслеживания устойчивости экономии, а базовая модель проходила перекрестную проверку тремя независимыми способами: с помощью специализированного программного обеспечения (RETScreen Expert), путем построения собственной регрессионной модели (в среде MATLAB) и посредством эволюционного поиска без жесткой структуры (символьная регрессия на основе генетического программирования — GP), не навязывающего заранее определенный вид функции. Ключевым результатом стало совпадение всех трех методов на одной и той же линейной модели, зависящей от показателя CDD ($R^2 = 0.9762–0.9765$, $\Delta R^2 \leq 0.0003$); это подтверждает, что согласованность результатов свидетельствует о реальной адекватности модели, а не об общей систематической ошибке, свойственной линейным методам. Значимость работы заключается в демонстрации воспроизводимой методики перекрестной проверки с количественной оценкой неопределенности, которую специалисты-практики могут применять и за пределами рассмотренного конкретного случая.

***Ключевые слова*:** символьная регрессия, генетическое программирование, градусо-сутки охлаждения, отслеживание кумулятивных сумм, кроссплатформенная проверка, верификация выбросов.

## INTRODUCTION

Unlike prior M&V studies that validate baseline energy models using a single regression tool or a single fixed functional form, this study is, to the authors' knowledge, the first to cross-validate a building energy savings baseline across three fully independent computational approaches — commercial M&V software (RETScreen Expert), custom numerical regression (MATLAB), and a structure-free evolutionary search (Symbolic Regression via Genetic Programming) — thereby removing the possibility that agreement between conventional tools merely reflects a shared assumption of linearity. This three-way validation, compared to the two-tool comparisons typical in the literature, constitutes the central novel contribution of this work.

Buildings account for approximately 30% of global energy consumption and a significant portion of greenhouse gas (GHG) emissions, making energy efficiency improvements critical for climate change mitigation [1]-[3]. In Canada, where harsh winters and growing cooling demands create substantial energy burdens, commercial and institutional buildings represent a key opportunity for energy conservation. Heating Degree Days (HDD) and Cooling Degree Days (CDD) serve as fundamental metrics for quantifying temperature-driven energy consumption, with CDD becoming increasingly important as urban cooling demands rise [4]-[7]. While the facility's cold climate records significantly higher HDD than CDD annually, CDD is selected as the sole independent variable given that the lighting retrofit directly reduces internal heat gains and cooling-related electricity consumption, a choice statistically validated by $R^2 = 0.9762$ across the full dataset. Despite widespread implementation of Energy Conservation Measures (ECMs) such as lighting retrofits, HVAC upgrades, and building envelope improvements, a critical challenge persists: accurately quantifying and verifying the actual energy savings achieved. The International Performance Measurement and Verification Protocol (IPMVP) provides standardized guidelines for measurement and verification (M&V), yet uncertainties in baseline modeling, weather normalization, and measurement accuracy can cause reported savings to deviate from actual performance by approximately 10% [8]-[10]. This uncertainty undermines confidence in energy efficiency investments, complicates access to financial incentives, and limits the scalability of retrofit programs.

Recent M&V research has explored increasingly sophisticated statistical and machine-learning

approaches, each with distinct limitations. Gaussian process models [11] and Bayesian approaches [12] offer strong uncertainty quantification but demand substantial computational resources and specialized statistical expertise, limiting practical deployment. Deep reinforcement learning [13] shows promise for HVAC optimization but requires extensive training data and lacks the interpretability that M&V transparency demands. Simplified performance-indicator models [14] address complexity concerns but risk oversimplifying building dynamics. Energy management system studies [15] and policy-focused research [16]-[17] provide valuable regulatory and organizational context but often disconnect from practical implementation. Benchmarking [18] and audit-based methodologies [19]-[20] contribute process documentation but tend to oversimplify site-specific variability. Measurement-uncertainty research from manufacturing contexts [21] and optimal energy management studies [22]-[23] offer useful principles but rely on assumptions that may not generalize to building operation. Smart-technology and edge-computing frameworks [24]-[25] present innovative concepts without sufficient practical validation across diverse environments. Notably absent from this literature is the use of structure-free, evolutionary search methods — such as Symbolic Regression via Genetic Programming — as an independent cross-check on baseline model form, despite their established use in other fields for discovering governing relationships without imposing a predetermined functional structure.

The current literature reveals three critical gaps: (1) limited validation of M&V software tools against independent methods for accuracy assessment; (2) insufficient practical guidance for implementing IPMVP protocols using accessible, replicable tools; and (3) inadequate quantification of model-form uncertainty — specifically, the untested assumption that agreement between two conventionally linear tools (e.g., RETScreen and MATLAB) is sufficient evidence that a linear model is the correct choice, rather than an artifact of both tools sharing the same underlying assumption.

This study addresses these gaps by developing and validating a comprehensive, uncertainty-aware M&V framework that integrates parametric regression analysis, CUSUM tracking, and — critically — a third, structure-free validation layer using Symbolic Regression via Genetic Programming, demonstrated through a three-year case study (2021–2023) of a lighting retrofit in a commercial building in London, Ontario, Canada. The research makes four key contributions:

- Validated baseline modeling methodology, using linear regression with CDD-based weather normalization, achieving high model confidence ($R^2 > 0.97$) across multiple time periods.
- Three-way cross-validation framework, comparing RETScreen Expert, independent MATLAB analysis, and an unconstrained Genetic Programming search to quantify model-form uncertainty beyond what two-tool comparisons can establish.
- Practical complexity assessment, evaluating the trade-off between model complexity (linear vs. polynomial fits up to 10th degree, and unrestricted symbolic forms) and practical accuracy.
- Complete uncertainty quantification, aligned with IPMVP requirements, demonstrating that a structure-free search independently rediscovers the same linear relationship rather than favoring added complexity.

This article is organized as follows: Section 2 presents the Theoretical Framework, covering the counterfactual M&V method, the IPMVP guidelines, policy and measures for building energy efficiency, general principles of lighting retrofits, the CDD-based regression formulation, and the Symbolic Regression via Genetic Programming formalism used as an independent, structure-free validation method. Section 3 presents the Case Study, describing the building, the data collection and preparation process, and a qualitative demonstration of the retrofit's effect on consumption. Section 4 presents the Results and Discussion, reporting the RETScreen–MATLAB–GP cross-validation findings and a discussion of the study's limitations. Section 5 concludes the article. This examination is conducted while adhering to the IPMVP options and considering Canadian conditions relevant to these analyses, including the application of RETScreen for effective energy savings evaluations.

## THEORETICAL FRAMEWORK

### *A. Counterfactual M&V Method*

The counterfactual method quantifies energy savings by comparing observed post-intervention outcomes to a predicted baseline representing the absence of implemented measures. Most M&V methodologies use this approach to determine avoided consumption—the difference between actual post-retrofit use and adjusted baseline (Fig. 1)—extending beyond electricity to water and

natural gas. The fundamental equation for energy savings calculation is described as follows:

$$ES_t = \left[E_{b,adj}(t) - E_{meas}(t)\right] \pm RA_t \quad (1)$$

Energy savings $ES_t$ at time t equals adjusted baseline consumption $E_{b,adj}(t)$ minus measured post-retrofit consumption $E_{meas}(t)$, plus routine adjustments for non-routine events $RA_t$. The baseline period requires: (1) adequate data points (typically 12 months minimum), (2) measured consumption data, and (3) independent variables like weather parameters. For weather-dependent facilities, the CDD-based baseline model is expressed as:

$$E_b = \beta_0 + \beta_1 . CDD + \varepsilon \quad (2)$$

Where $E_b$ is the monthly electricity baseline energy consumption (kWh), $\beta_0$ is the intercept (base load consumption independent of weather), $\beta_1$ is the slope coefficient (weather-dependent consumption rate), $CDD$ represents cooling degree days (°C.days), and $\varepsilon$ is the error term. The adjusted baseline for the reporting period becomes:

$$E_{b,adj}(t) = \beta_0 + \beta_1 . CDD_r(t) \quad (3)$$

Here $CDD_r(t)$ represents actual cooling degree days during the reporting period. ECM refers to actions or technologies reducing energy usage, including LED lighting, high-performance HVAC, enhanced insulation, and equipment upgrades implemented to lower costs and carbon emissions.

The cumulative energy savings CUSUM over the reporting period are calculated as:

$$CUSUM = \sum_{t=1}^{n} ES_t = \sum_{t=1}^{n} \left[E_{b,adj}(t) - E_{meas}(t)\right] \quad (4)$$

Where n represents the number of reporting periods (months or years). The associated cost savings CS are determined by:

$$CS = \text{Cumulative Savings} \times \text{Unit Energy Cost} \quad (5)$$

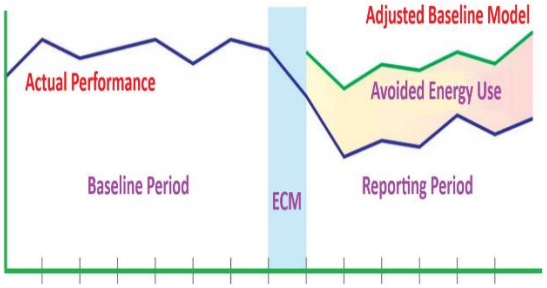



**Fig. 1. Counterfactual method [9].**

The coefficient of determination ($R^2$) assesses baseline model accuracy:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (6)$$

Actual energy consumption $y_i = E_{meas}(t)$ is compared to predicted consumption $y_i$ from the baseline model, and mean consumption $\bar{y}$ using residual $SS_{res}$ and total sums $SS_{tot}$ of squares to calculate $R^2$. Per IPMVP guidelines, $R^2 \geq 0.75$ indicates acceptable model fit, while $R^2 \geq 0.9$ represents excellent fit.

*B. IPMVP*

M&V methods provide systematic, transparent approaches to quantify energy savings with verifiable accuracy. IPMVP, developed by the Efficiency Valuation Organization, serves as the global standard for efficiency verification across 40+ countries [3], defining four options (Table 1): A (key parameter measurement), B (all parameter measurement), C (whole facility analysis via utility data and regression), and D (calibrated simulation). IPMVP is critical for Canadian buildings, providing standardized methodology, flexible options, regulatory alignment with NECB and provincial codes, compatibility with incentive programs, support for GHG reduction targets, and professional development through organizations like AEE [7]. IPMVP emphasizes uncertainty quantification, recognizing M&V calculations can deviate from actual savings by ~10% due to measurement errors and modeling limitations [3].

*C. Policy and Measures of Building Energy Efficiency*

Canadian building energy efficiency policies establish comprehensive frameworks to enhance sustainability and reduce operational costs. The NECB mandates minimum performance standards for building envelope, HVAC, lighting power density, and electrical systems, with provincial/municipal jurisdictions often exceeding these minimums to meet 30-40% GHG reduction

targets by 2030.

Financial incentives drive implementation through utility rebates, federal grants (Investing in Canada Infrastructure Program), and provincial initiatives in Ontario and Alberta. Mandatory energy audits identify savings potential for incentive eligibility. Key measures include LED retrofits with advanced controls (30-75% reductions), high-efficiency HVAC with heat recovery, envelope improvements, renewable energy integration, and smart building technologies. LEED certification fosters sustainable practices and market differentiation [6]. These coordinated policies deliver economic, environmental, and social benefits, positioning commercial retrofits as strategic investments supporting national climate commitments.

*Table 1*

IPMVP measurement and verification option

| IPMVP Option | Option A | Option B | Option C | Option D |
|---|---|---|---|---|
| Measurement Approach | Partial measurement of key parameters; remaining parameters estimated | Full measurement of all relevant parameters | Regression analysis of utility billing data with normalization | Computer-based energy modeling calibrated to measured data |
| Data Requirements | Short-term or spot measurements; engineering estimates | Continuous or long-term measurements of energy use | Monthly or interval utility data; weather data (HDD/CDD) | Detailed building geometry, systems, schedules, calibration data |
| Typical Applications | Lighting retrofits, motor replacements, simple ECMs | HVAC upgrades, variable-speed drives, process improvements | Whole-building retrofits, multiple ECMs | New buildings, major renovations, complex systems |
| Advantages | Low cost; simple implementation; quick assessment | High accuracy; reduced uncertainty | Captures interactive effects; uses existing data | Applicable when baseline data are unavailable |
| Limitations | Higher uncertainty due to estimated parameters | Higher cost; increased instrumentation and data handling | Savings attribution may be less granular | High modeling effort; requires expert knowledge |

### *D. Lighting Retrofits: General Principles*

Building lighting retrofits in Canadian commercial and institutional buildings enhance energy efficiency, indoor environmental quality, and operational costs. Retrofits begin with audits identifying inefficiencies and outdated technologies. LED transition is most effective, reducing consumption up to 75% versus incandescent/fluorescent fixtures, with ~25,000-hour lifespan versus 1,000 hours [5]. Advanced controls—occupancy sensors, daylight harvesting, dimming—add 30-50% savings by activating lights only when needed. Compliance with NECB lighting power density requirements ensures projects meet or exceed efficiency standards. Benefits are substantial: (1) 30-50% energy savings with 2-5 year payback periods, (2) enhanced comfort and 5-10% productivity gains, (3) reduced maintenance from extended LED lifespan, and (4) decreased GHG emissions—50,000 kWh saved annually yields ~30 tons $CO_2$ reduction. Successful retrofits across Canadian public buildings, schools, healthcare facilities, and offices demonstrate proven frameworks. As shown in Table 2, lighting retrofits deliver robust financial returns while advancing Canada's climate action objectives.

*Table 2*

Impacts of lighting retrofits in Canadian commercial buildings

| Aspect | Conventional Lighting | LED Retrofits with Controls | Typical Impact |
|---|---|---|---|

| Fixture Type | Incandescent / Fluorescent | LED luminaires | — |
|---|---|---|---|
| Power Reduction | — | 30–75% lower wattage | Significant energy reduction |
| Lamp Lifetime | ~1,000–10,000 h | ~25,000–50,000 h | Reduced maintenance |
| Lighting Controls | Manual switching | Occupancy sensors, daylight harvesting, dimming | Additional 30–50% savings |
| Annual Energy Savings | Baseline | 30–50% | Lower utility bills |
| Simple Payback | — | 2–5 years | Financially attractive |
| Maintenance Cost | High (frequent replacement) | Low | O&M savings |
| GHG Emission Reductions | Baseline | ~30 $tCO_2$ per 50,000 kWh saved | Environmental benefit |
| Code Compliance | Often non-compliant | NECB-compliant or better | Regulatory alignment |

*E. CDD-Based Regression Modeling*

The weather normalization is as follows:

$$E_{norm} = E_{meas} \frac{CDD_{typical}}{CDD_{actual}} \quad (7)$$

Where $E_{norm}$ is normalized energy consumption (kWh), $E_{meas}$ is actual measured consumption (kWh), $CDD_{typical}$ is long-term average CDD for the location, and $CDD_{actual}$ is actual CDD during measurement period. The monthly CDD billing period aggregation can be calculated as follows:

$$CDD_{monthly} = \sum_{i=1}^{n} \left[ T_{avg,i} - T_{base} \right] \; T_{avg,i} > T_{base} \quad (8)$$

Where $T_{avg,\,i}$ represents average daily temperature (°C), and $T_{base}$ acts as base temperature (typically 18°C for cooling). A comprehensive analytical framework quantifies energy savings and validates model accuracy through four complementary techniques: time series visualization for consumption patterns, linear regression to establish the energy-CDD relationship and baseline models, CUSUM analysis for progressive savings tracking, and M&V charting for net savings per IPMVP standards. Model-form uncertainty between two regression platforms can be quantified as:

$$\Delta R^2 = \left| R^2_{RETScreen} - R^2_{MATLAB} \right| \quad (9)$$

The polynomial regression models are given:

$$E(CDD) = \begin{cases} \beta_0 + \beta_1.CDD + \varepsilon \\ \beta_0 + \beta_1.CDD + \beta_2.CDD^2 + \varepsilon \\ \beta_0 + \beta_1.CDD + \beta_2.CDD^2 + \beta_3.CDD^3 + \varepsilon \\ \beta_0 + \sum_{i=1}^{n} \beta_i.CDD^i + \varepsilon \end{cases} \quad (10)$$

Where E(CDD) is predicted energy consumption (kWh), and $\beta_i$ = regression coefficients.

*F. Symbolic Regression via Genetic Programming*

This study further validates the baseline modeling approach using Symbolic Regression via Genetic Programming (GP), an independent, structure-free search method. Unlike regression's fixed functional form, GP evolves structure and coefficients together, using expression trees under selection, crossover, and mutation to minimize error against the electricity-CDD data, determining whether non-linearity or linearity better explains the dependency. GP searches the unrestricted expression space built from function set F and terminal set T:

$$\hat{E}(CDD) = \Phi(CDD), \Phi \in S(F,T) \quad (11)$$

Where $\Phi$ is a candidate expression (tree), and $S(F,T)$ denotes the space of all syntactically valid trees composable from:

$$F = \left\{ +, -, \times, \div, \sqrt{.}, \ln., \exp(.) \right\}, T = \{CDD, c\} \quad (12)$$

With $c$ an ephemeral random constant (ERC). Fitness function is F: each candidate $\Phi_i$ is evaluated against the n measured observations using:

$$RMSE\left(\Phi_i\right)=\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left[E_{meas}\left(t\right)-\Phi_i\left(CDD_t\right)\right]^2} \quad (13)$$

$$MAE\left(\Phi_i\right)=\frac{1}{n}\sum_{t=1}^{n}\left|E_{meas}\left(t\right)-\Phi_i\left(CDD_t\right)\right| \quad (14)$$

$$R^2\left(\Phi_i\right)=1-\frac{\sum_{t=1}^{n}\left[E_{meas}\left(t\right)-\Phi_i\left(CDD_t\right)\right]^2}{\sum_{t=1}^{n}\left[E_{meas}\left(t\right)-\overline{E_{meas}\left(t\right)}\right]^2} \quad (15)$$

Selection (tournament): the fittest individual among $k$ randomly drawn candidates are selected as a parent:

$$\Phi_{parent}=arg\min_{\Phi_i\in(\Phi_1,\ldots,\Phi_k)}RMSE\left(\Phi_i\right) \quad (16)$$

Genetic operators: new individuals for generation $g+1$ are produced from generation $g$ according to:

$$\Phi^{(g+1)}=\begin{cases}Crossover\left(\Phi_{p1},\Phi_{p2}\right), with\ probability\ p_c\\ Mutation\left(\Phi_p\right), with\ probability\ p_m\\ \Phi_p, with\ probability\ 1-p_c-p_m\end{cases} \quad (17)$$

Where crossover exchanges randomly selected subtrees between two parents, and mutation replaces a randomly selected subtree with a newly generated random subtree. Convergence criterion: the evolutionary process terminates after $G$ generations, returning the best overall model:

$$\Phi_{best}=\arg\min_{g\in(1,\ldots,G)}\min_{i\in(1,\ldots,P)}RMSE\left(\Phi_i^{(g)}\right) \quad (18)$$

Where $P$ is the population size.

## CASE STUDY

### *1. Building Information*

This study applies RETScreen Expert to analyze 3 years of electricity data from a London Airport, Ontario commercial building, evaluating performance following a February 2022 lighting retrofit. The baseline period (Jan 2021–Dec 2021) and reporting period (Jan 2022–Dec 2023) enable linear regression modeling correlating consumption with CDD, quantifying energy savings, cost reductions, and GHG emission reductions for stakeholders.

Cross-validation via RETScreen Expert and MATLAB ensures robust, verifiable results supporting investment decisions and a replicable framework for Canadian retrofits—capturing whole-facility effects, particularly reduced cooling loads from decreased LED internal heat gains. Although lighting retrofits are commonly evaluated using IPMVP Option A, this study applies IPMVP Option C to capture whole-facility and interactive effects using utility billing data and weather-normalized regression (Table 3). Lighting retrofits offer 30–75% energy reductions via LED technology and controls. M&V focuses on fixture power, operating hours, and coincident peak operation [4].

*Table 3*

Measurement and verification approach for the lighting retrofit based on IPMVP option C

| Item | Description | Source / Method |
|---|---|---|
| IPMVP option | Option C – Whole facility measurement and verification | IPMVP [8] |
| ECM type | LED lighting retrofit | Project documentation |
| Measurement boundary | Entire building electricity consumption | Utility billing data |
| Analysis period | 36 months (Jan 2021 – Dec 2023) | Utility records |
| Baseline period | Pre-retrofit period (Jan 2021 – Dec 2021) | Utility data |
| Reporting period | Post-retrofit period (Jan 2022 – Dec 2023) | Utility data |
| Energy data | Monthly electricity consumption (kWh) | Utility bills |
| Independent variable | Cooling Degree Days (CDD) | NASA / RETScreen |
| Baseline model | Linear regression of energy vs. CDD | RETScreen & MATLAB |
| Weather normalization | Adjusted baseline using actual CDD | IPMVP Option C |
| Savings calculation | Difference between adjusted baseline and actual consumption | Counterfactual method |
| Performance tracking | CUSUM analysis | RETScreen & MATLAB |
| HVAC interactive | Captured implicitly at whole-facility level | Option C characteristic |
| Uncertainty treatment | Cross-validation (RETScreen vs. MATLAB), $R^2$ analysis | This study |

Power is directly measured; hours are established via energy surveys, forming a consistent baseline. Key assumptions: constant pre-retrofit hours, direct power measurement, exclusion of HVAC effects, and unchanged lighting levels. The study examines a London, Ontario office building (15,290 ft², mixed-use), retrofitted February 2022, with the retrofit's effect on consumption demonstrated qualitatively in Table 6.

*2. Data Collection and Preparation*

A 36-month electricity dataset (Jan 2021–Dec 2023) was collected with owner authorization

from the local utility serving London, Ontario, including monthly consumption, billing dates, and commercial rates (0.10–0.13 CAD/kWh, Table 5). CDD data were extracted from NASA via RETScreen, synchronized with billing periods; supplementary parameters (temperature, solar radiation, humidity, precipitation) enabled energy-weather correlation analysis. Although Climate Zone 6A-Cold-Humid shows HDD (4,009) far exceeding CDD (1,078), CDD remains methodologically appropriate since the LED retrofit reduces internal heat gains and cooling loads—validated by $R^2 = 0.9762$ (full dataset) and $R^2 = 0.9952$ (baseline).

*Table 4*

Building characteristics for the case study facility

| Parameter | Description |
|---|---|
| Building type | Commercial / Institutional office building |
| Location | London, Ontario, Canada |
| Year of construction | ~25 years old |
| Number of floors | Two |
| Total floor area | 15,290 ft² |
| Primary occupancy | Offices, classrooms, common areas |
| Climate zone | Cold climate (Canadian context) |
| Utility provider | Local electric utility (Ontario) |
| Primary energy source | Electricity |
| Energy Conservation Measure (ECM) | LED lighting retrofit |
| ECM implementation date | February 2022 |
| Analysis boundary | Whole-building electricity consumption |
| Operating profile | Typical commercial occupancy |

*Table 5*

Electricity pricing trends prices

| Sector | Average Price (CAD per kWh) |
|---|---|
| Residential | 0.12 - 0.15 |
| Commercial | 0.10 - 0.13 |
| Industrial | 0.08 - 0.10 |
| Agricultural | 0.09 - 0.11 |
| Institutional (e.g., schools, hospitals) | 0.10 - 0.12 |
| Government | 0.11 - 0.13 |
| Electric Vehicle Charging | 0.15 - 0.20 |

The dataset was prepared using RETScreen Expert, integrating NASA weather data with building-specific parameters (coordinates, climate zone, elevation, design temperatures) for London, Ontario. A dual-platform approach was implemented: RETScreen for baseline modeling and savings quantification, and MATLAB's basic fitting tool for independent verification of regression coefficients, $R^2$, and uncertainty across polynomial degrees—strengthening confidence in results and supporting transparent uncertainty quantification per IPMVP standards.

*3. Facility Information*

RETScreen's facility characterization module consolidates building metadata (type, classification, address, manager) and benchmark parameters. Key metrics include facility size (15,290 ft²), Energy Use Intensity (kWh/ft²), and comparative benchmarks against similar commercial/institutional buildings. Automated benchmark analysis evaluates performance against industry standards, incorporating base case and proposed case EUI values, energy plans, and projected fuel savings. While facility size was manually input, all derivative parameters—benchmark ranges, intensity metrics, performance indicators—were automatically calculated based on location, climate zone, and building classification.

*4. Qualitative Evidence of Lighting Savings*

Before the full regression-based validation, Table 6 illustrates directly how the February 2022 lighting retrofit reduced consumption, independent of any statistical model. It compares selected months from pre-retrofit 2021 against post-retrofit 2023 (measured data, Table 9), chosen to hold cooling demand constant—equal, near-zero, or higher CDD post-retrofit—so that reductions cannot be attributed to milder weather.

January and February show zero CDD in both years, so cooling plays no role; the 3.5% and 3.9% reductions reflect genuine base-load savings, consistent with lower internal heat and standby draw

from the new LED fixtures. April and November show higher post-retrofit CDD (48 vs. 29; 22 vs. 19), which should raise cooling use, yet consumption still fell—indicating lighting savings outweighed added cooling demand, confirming the retrofit, not weather, drove the reduction. December shows the largest drop (16.5%), consistent with peak pre-retrofit lighting load during the year's shortest daylight period.

*5. Energy Consumption Data*

The workflow began by importing 36 months of electricity data (Jan 2021–Dec 2023) into RETScreen Expert, which calculated monthly usage (kWh/month). NASA weather data was integrated, using CDD instead of raw temperature to better capture cooling-related demand.

*Table 6*

Before/after comparison of matched-CDD months, pre-retrofit (2021) versus post-retrofit (2023)

| Month | Pre (2021) kWh | Pre CDD | Post (2023) kWh | Post CDD | Δ kWh (%) | Note |
|---|---|---|---|---|---|---|
| Jan | 147,719 | 0 | 142,609 | 0 | −5,110 (−3.5%) | Identical CDD → pure base-load drop |
| Feb | 151,876 | 0 | 146,001 | 0 | −5,875 (−3.9%) | Identical CDD → pure base-load drop |
| Apr | 155,876 | 29 | 154,705 | 48 | −1,171 (−0.8%) | Post CDD 65% higher, kWh still lower |
| Jun | 225,817 | 245 | 213,932 | 246 | −11,885 (−5.3%) | Virtually identical CDD |
| Aug | 278,402 | 414 | 273,659 | 406 | −4,743 (−1.7%) | Nearly identical CDD |
| Nov | 160,089 | 19 | 142,968 | 22 | −17,121 (−10.7%) | Post CDD slightly higher, still lower kWh |
| Dec | 142,109 | 1 | 118,721 | 0 | −23,388 (−16.5%) | Nearly identical CDD |

CDD values were aggregated per billing period for temporal alignment. An event log marking the lighting retrofit's completion (February 2022) established the baseline-reporting boundary, separating pre- and post-retrofit consumption for accurate M&V-based savings quantification. RETScreen Expert enabled scenario modeling, sensitivity analysis, and GHG calculations. For robustness, regression, CUSUM, and M&V procedures were run separately across the full dataset (36 months), pre-retrofit baseline (11 months), and post-retrofit period (25 months). Parallel MATLAB fitting (Fig. 2) independently validated RETScreen results, comparing regression coefficients, $R^2$ values across polynomial degrees (linear–10th order), and error quantification—strengthening confidence and transparency in reported savings.

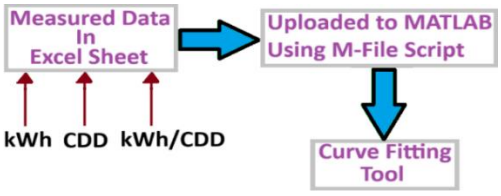



**Fig. 2. Basics fitting tool based-MATLAB m-file coding.**

The CDD variable, rather than HDD, is used because: (1) this is a lighting retrofit, not a heating upgrade—LEDs produce less waste heat, reducing cooling loads, making HDD irrelevant; (2) reduced internal heat gains from LEDs lower cooling-related electricity use in warm months, a relationship CDD captures directly.

## RESULTS AND DISCUSSION

The location of the facility according to the nearest climate data station are illustrated in Fig. 3 and Table 7. The climate data parameters based on NASA database are depicted in Fig. 4 and Table 8. The electricity bill data and other relevant information for a commercial building in London, Ontario, were analyzed to assess electricity consumption patterns over the years as illustrated in Table 9. The goal was to establish a relationship between the annual electricity usage and the CDDs for the same period, in order to evaluate the energy savings from the February 2022 lighting retrofit according to the climate data location. Energy savings were estimated by applying the theoretical framework of Section 2 to the 36-month measured dataset (2021–2023) described in Table 9, under the following two validation cases:

*A. Case 1: RETScreen Versus MATLAB Cross Validation*

A time series graph is obtained which represents a relationship between the CDD (°C·days) and energy (kWh/month) on the vertical axis, and the number of the months on the horizontal axis (time)

as illustrated in Fig. 5. The graph in Fig. 5 is obtained via RETScreen and MATLAB software to validate the accuracy and reliability of the energy savings and performance models created in RETScreen by cross-referencing them with a custom-built model in MATLAB.

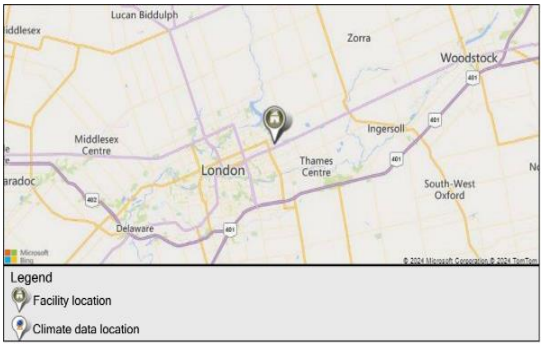



**Fig. 3. Facility legend using RETScreen software.**

*Table 7*

Location details of the facility

| | Unit | Climate data location | Facility location |
|---|---|---|---|
| Name | | Canada-Ontario-London-Airport | Canada-ON-London |
| Latitude | °N | 43.0 | 43.0 |
| Longitude | °E | -81.2 | -81.2 |
| Climate zone | | 6A-Cold-Humid | 6A-Cold-Humid |
| Elevation | m | 278 | 275 |

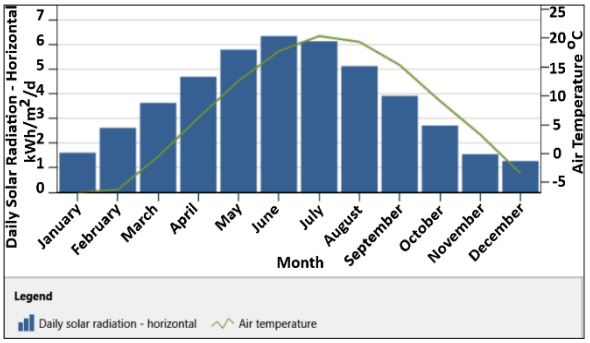


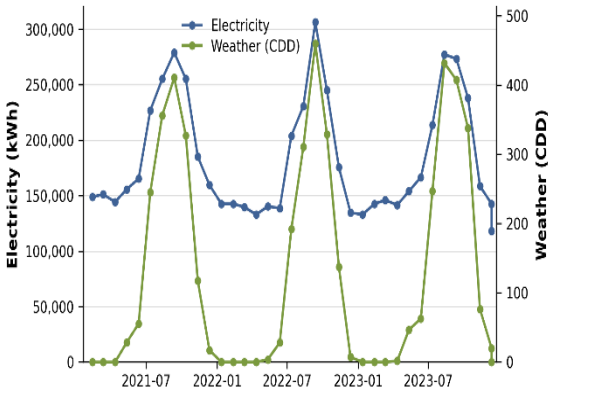

**Fig. 4. Climate data-based NASA database using RETScreen software: Daily solar radiation and air temperature versus time.**

*Table 8*

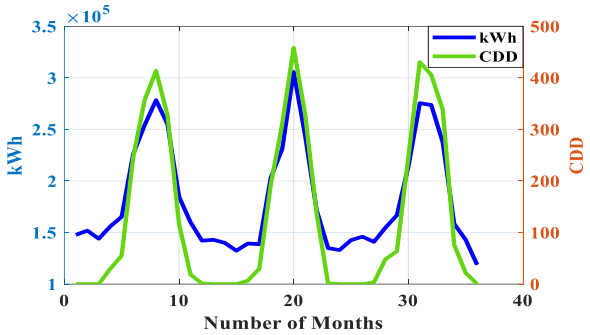
**Weather parameters-based NASA database using RETScreen software**

| Heating design temperature | | | | -15.9 | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Cooling design temperature | | | | 28.5 | | | | | |
| Earth temperature amplitude | | | | 21.4 | | | | | |
| Month | Air Temperature | Relative humidity | Precipitation | Daily solar radiation-horizontal | Atmospheric pressure | Wind speed | Earth Temperature | HDD | CDD |
| | °C | % | mm | kWh/m²/d | kPa | m/s | °C | °C-d | °C-d |
| January | -6.7 | 79.0% | 44.95 | 1.61 | 98.2 | 5.3 | -5.6 | 766 | 0 |
| February | -6.2 | 76.0% | 40.04 | 2.62 | 98.3 | 5.0 | -5.0 | 678 | 0 |
| March | -0.5 | 72.8% | 45.26 | 3.63 | 98.2 | 5.0 | -0.2 | 574 | 0 |
| April | 6.2 | 68.1% | 57.00 | 4.69 | 98.1 | 5.0 | 6.4 | 354 | 0 |
| May | 12.6 | 67.9% | 68.51 | 5.79 | 98.2 | 4.2 | 12.9 | 167 | 81 |
| June | 17.7 | 70.5% | 69.30 | 6.34 | 98.2 | 3.6 | 18.8 | 9 | 231 |
| July | 20.3 | 72.8% | 70.06 | 6.13 | 98.3 | 3.1 | 22.1 | 0 | 319 |
| August | 19.3 | 76.7% | 63.55 | 5.12 | 98.4 | 3.1 | 21.3 | 0 | 288 |
| September | 15.3 | 77.3% | 74.40 | 3.92 | 98.5 | 3.3 | 16.7 | 81 | 159 |
| October | 9.1 | 77.6% | 68.20 | 2.71 | 98.5 | 3.9 | 9.5 | 276 | 0 |
| November | 3.3 | 79.6% | 61.80 | 1.55 | 98.3 | 4.7 | 3.1 | 441 | 0 |
| December | -3.4 | 80.7% | 52.39 | 1.27 | 98.3 | 5.0 | -2.7 | 663 | 0 |
| Annual | 7.3 | 74.9% | 715.46 | 3.79 | 98.3 | 4.3 | 8.2 | 4009 | 1078 |

This graph depicted in Fig. 5 helps identify seasonal trends in energy consumption, particularly during warmer months when cooling demand increases. By correlating CDDs with electricity usage, users can assess how temperature fluctuations drive energy needs. This visualization also facilitates performance monitoring, enabling the evaluation of energy efficiency measures such as in this case lighting retrofits by observing reductions in energy use during periods of high cooling demand. Additionally, the graph can reveal anomalies, such as unexpected spikes or drops in consumption, which may indicate operational inefficiencies or system malfunctions. It also aids in benchmarking, allowing comparisons between actual consumption and expected performance based on historical data, ensuring the building operates efficiently under varying cooling requirements.

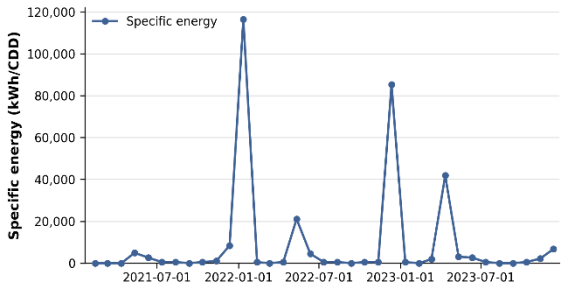



**(a)**

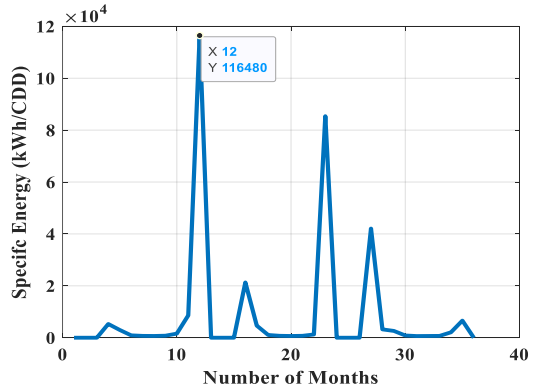



**(b)**

**Fig. 5. Time series graph using: (a) RETScreen software. (b) MATLAB software.**




**(a)**




**(b)**

**Fig. 6. Time series graph of specific energy consumption versus time using: (a) RETScreen software. (b) MATLAB software.**

*Table 9*

Measured data of energy consumption, CDD, and specific energy

| Measured Data | | | |
|---|---|---|---|
| End Date | Electricity (kWh) | Cooling Degree Days (CDD) | Specific Energy (kWh/CDD) |
| 1/13/2021 | 147719 | 0 | 0 |
| 2/13/2021 | 151876 | 0 | 0 |
| 3/15/2021 | 144003 | 0 | 0 |
| 4/14/2021 | 155876 | 29 | 5297 |
| 5/15/2021 | 165287 | 55 | 2995 |
| 6/15/2021 | 225817 | 245 | 920 |
| 7/15/2021 | 254126 | 357 | 712 |
| 8/14/2021 | 278402 | 414 | 673 |
| 9/14/2021 | 254973 | 327 | 780 |
| 10/14/2021 | 184822 | 117 | 1581 |
| 11/15/2021 | 160089 | 19 | 8602 |
| 12/15/2021 | 142109 | 1 | 116483 |
| 1/15/2022 | 142935 | 0 | 0 |
| 2/13/2022 | 139964 | 0 | 0 |
| 3/14/2022 | 132370 | 0 | 0 |
| 4/13/2022 | 139338 | 7 | 21306 |
| 5/12/2022 | 138656 | 30 | 4656 |
| 6/12/2022 | 202999 | 192 | 1055 |
| 7/12/2022 | 231160 | 310 | 745 |
| 8/14/2022 | 305790 | 458 | 667 |
| 9/15/2022 | 244017 | 329 | 741 |
| 10/15/2022 | 171816 | 138 | 1249 |
| 11/14/2022 | 134897 | 2 | 85378 |
| 12/13/2022 | 132993 | 0 | 0 |
| 1/13/2023 | 142609 | 0 | 0 |
| 2/13/2023 | 146001 | 0 | 0 |
| 3/15/2023 | 140928 | 3 | 42068 |
| 4/15/2023 | 154705 | 48 | 3214 |
| 5/15/2023 | 166711 | 64 | 2620 |
| 6/15/2023 | 213932 | 246 | 870 |
| 7/16/2023 | 275489 | 431 | 640 |
| 8/16/2023 | 273659 | 406 | 674 |
| 9/16/2023 | 237560 | 338 | 702 |
| 10/15/2023 | 158415 | 76 | 2083 |
| 11/15/2023 | 142968 | 22 | 6594 |
| 12/12/2023 | 118721 | 0 | 0 |

Ultimately, the graph provides a clear way to track energy consumption patterns and assess the effectiveness of energy-saving initiatives in response to environmental factors. The second step involves plotting specific energy consumption

(kWh/CDD) versus time, as depicted in Fig. 6, to analyze cooling system efficiency relative to outdoor temperature variations. Normalizing energy consumption to CDD reveals seasonal trends, performance issues, and the impact of energy-saving measures such as lighting retrofits by comparing pre- and post-intervention consumption.

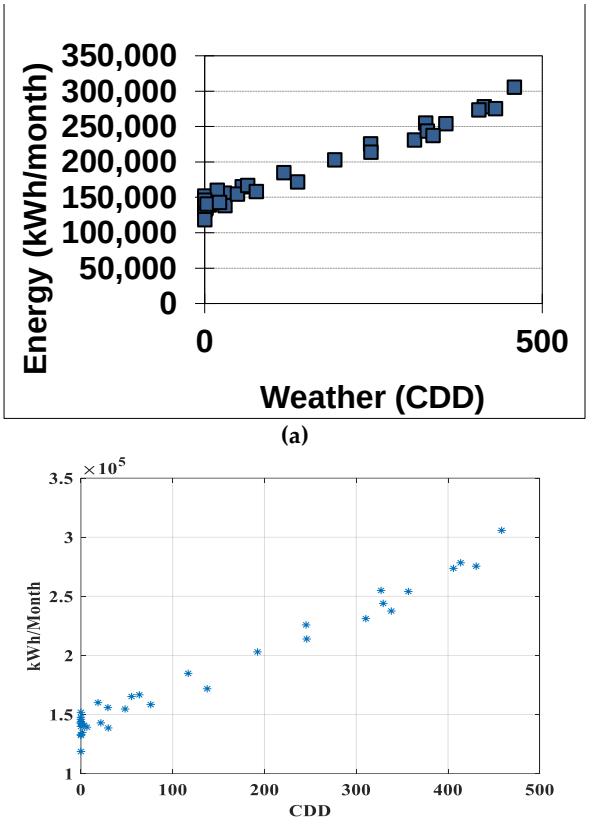



**Fig. 7. Actual energy consumption versus CDD graph using: (a) Excel sheet. (b) MATLAB software.**

Anomalies such as unexpected spikes may indicate inefficiencies or operational problems.

RETScreen and MATLAB are both used to estimate specific energy, assessing model uncertainty and potential error sources. The third step plots energy consumption (kWh/month) versus CDD, as illustrated in Fig. 7, analyzing the relationship between cooling demand and energy usage through regression analysis. Here, energy consumption serves as the dependent variable while CDD acts as the independent variable, reflecting temperature-driven cooling demand across the entire study period. Regression analysis quantifies the relationship between CDD and energy consumption, enabling a predictive model that estimates usage based on temperature variations. This identifies how cooling demand impacts energy efficiency, with Fig. 7 prepared via Excel prior to RETScreen and MATLAB validation. The fourth step establishes the baseline model using linear regression, as depicted in Fig. 8. The resulting $R^2$ value of 0.9762 indicates an excellent model fit, confirming a very strong correlation between CDD as the independent variable and monthly energy consumption as the dependent variable.

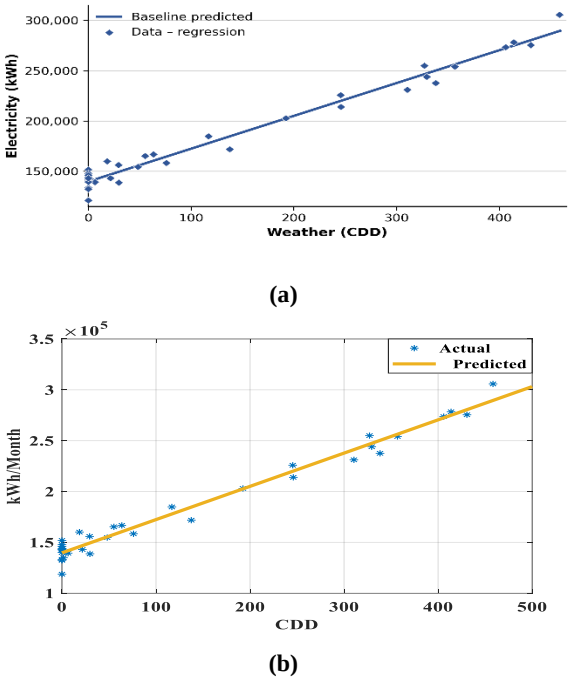



**Fig. 8. Baseline model using linear regression analysis using: (a) RETScreen software. (b) MATLAB software.**

*Table 10*

Regression results for the 3 years data

| | RETScreen | MATLAB | $\Delta R^2$ |
|---|---|---|---|
| Eq. | Y=326.9687*X+ 139685.9384 | Y=326.9669*X+ 139687.2334 | - |
| $R^2$ | 0.9762 | 0.976211 | 0.000011 |

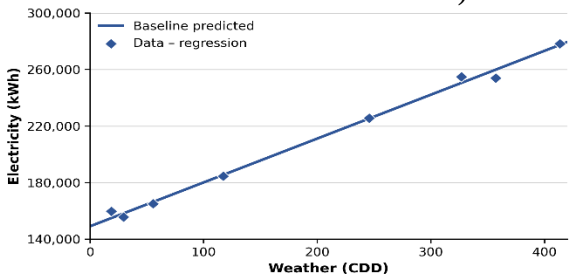
*Table 11*

Different types of fit for $R^2$ in MATLAB software

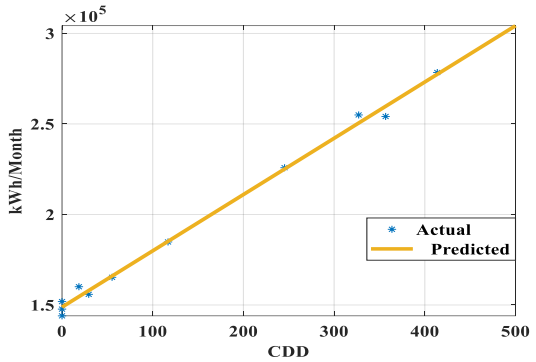
| Types of fit | $R^2$ |
|---|---|
| Linear | 0.9762 |
| Quadratic | 0.9773 |
| Cubic | 0.9786 |
| 4th degree polynomial | 0.9788 |
| 5th degree polynomial | 0.9794 |
| 6th degree polynomial | 0.9794 |
| 7th degree polynomial | 0.9794 |
| 8th degree polynomial | 0.9802 |
| 9th degree polynomial | 0.9805 |
| 10th degree polynomial | 0.9805 |

The comparison illustrated in Table 10 between the regression results from RETScreen and MATLAB

demonstrates a high degree of alignment, with both tools producing nearly identical linear regression equations and $R^2$ values (0.9762 in RETScreen and 0.976211 in MATLAB).




**(a)**




**(b)**

**Fig. 9. Baseline model using linear regression analysis during the 11 months (pre-retrofit case) using: (a) RETScreen software. (b) MATLAB software.**

This minor difference in $\Delta R^2$ values suggests that both methods provide almost equivalent results when applying a linear fit, confirming the accuracy of RETScreen's linear regression relative to MATLAB's model.

Table 11 presents results for different polynomial degrees tested in MATLAB, showing that higher-degree polynomials slightly improve $R^2$, reaching up to 0.9805 with a 9th or 10th degree polynomial.

However, the improvement becomes minimal beyond a certain degree, confirming that the linear model remains a practical and reliable fit, as supported by the similarity in $R^2$ values across models.

The fifth step involves CUSUM analysis based on the pre-retrofit interval to determine energy savings during the first year of new lighting operation. As shown in Fig. 9, $R^2$ equals 0.9952, indicating an excellent model fit.

Tables 11 and 12 further confirm that the linear model is highly effective for pre-retrofit analysis, as higher-order polynomials provide no significantly better fit, validating RETScreen's linear approach while demonstrating MATLAB's flexibility for exploring complex alternatives.

*Table 12*

Regression results for 11 month before ECM (pre-retrofit case)

| | RETScreen | MATLAB | $\Delta R^2$ |
|---|---|---|---|
| Eq. | Y=310.8369*X+ 148829.1626 | Y=310.8331*X+ 148830.6142 | - |
| $R^2$ | 0.9952 | 0.9952 | 0 |

*Table 13*

Different types of fit versus $R^2$ in MATLAB during pre-retrofit interval.

| Types of fit | $R^2$ |
|---|---|
| Linear | 0.9952 |
| Quadratic | 0.9952 |
| Cubic | 0.9952 |
| $4^{th}$ degree polynomial | 0.9953 |
| $5^{th}$ degree polynomial | 0.9960 |
| $6^{th}$ degree polynomial | 0.9966 |
| $7^{th}$ degree polynomial | 0.9978 |
| $8^{th}$ degree polynomial | 0.9988 |
| $9^{th}$ degree polynomial | 0.9987 |
| $10^{th}$ degree polynomial | 0.9943 |

The CUSUM analysis graph, depicted in Fig. 10, evaluates changes in energy consumption over time following energy efficiency measures, using cumulative energy usage as the dependent variable and time as the independent variable. It complements regression analysis by comparing actual cumulative consumption to predicted values, highlighting deviations from expected performance. This enables trend visualization, effectiveness assessment of conservation measures, and informed decision-making regarding energy management strategies. From Fig. 10, a negative CUSUM curve indicates that actual energy consumption is consistently lower than predicted, confirming that energy efficiency interventions have been effective in reducing usage. This reflects improved energy management performance, leading to significant cost savings through reduced energy bills, while highlighting further optimization opportunities. Fig. 10 is estimated using both RETScreen and MATLAB to verify model robustness, ensuring key indicators such as $R^2$ values and energy savings align across both platforms, providing confidence in the results. Combining the M&V chart with cumulative sum analysis enables a thorough assessment of the lighting retrofit's first-year savings and confirms their persistence over time. As shown in Fig. 11, net savings reached 232,056 kWh, equivalent to 27,846.72 CAD, calculated by measuring the

difference between actual post-retrofit and predicted baseline consumption. Consequently, gross annual $CO_2$ emission reductions amount to 23.57 $tCO_2$, based on Ontario's emission factor of 0.1016 $tCO_2$/MWh.

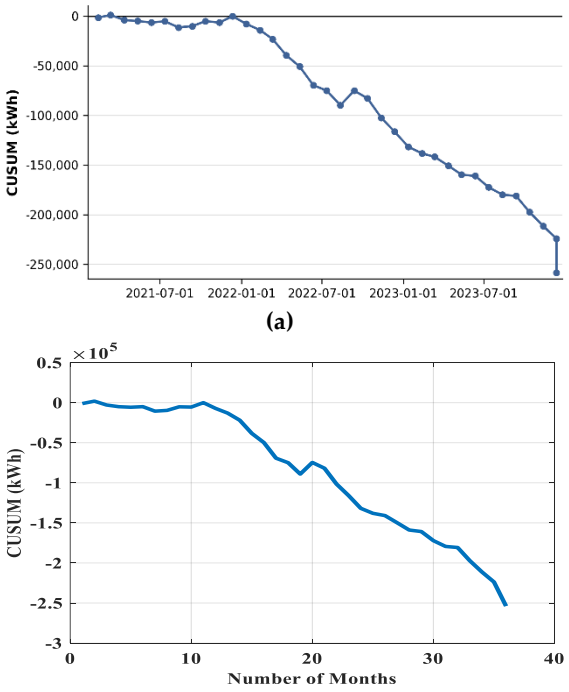



**Fig. 10. CUSUM analysis graph based on baseline model of pre-retrofit case using RETScreen software using: (a) RETScreen software. (b) MATLAB software.**

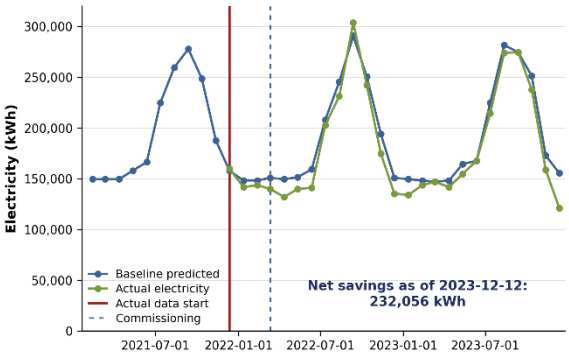



**Fig. 11. M&V chart using RETScreen software.**

### *B. Case 2: Symbolic Regression Via Genetic Programming Results*

To independently validate the baseline model established through RETScreen Expert and MATLAB's basic fitting tool, a GP algorithm was applied to the same 36-month dataset (Table 9), using a population of 60 individuals evolved over 1000 generations under the function set and operators defined in equations (11)–(18).The convergence curve demonstrates that the best-individual RMSE stabilizes by approximately generation 250, after an initial rapid descent over the first ~50 generations, indicating reliable convergence rather than premature stagnation, as illustrated in Fig. 12 and Fig. 13, respectively.

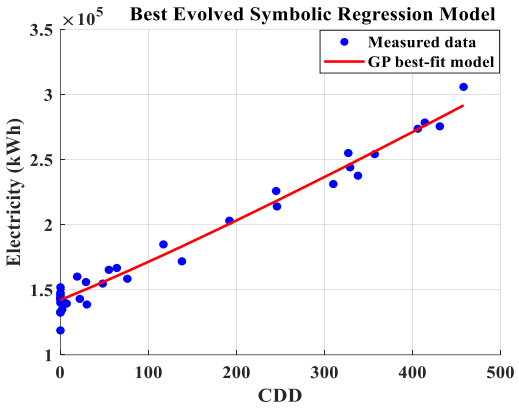



**Fig. 12. GP-derived symbolic regression model of electricity consumption vs. CDD.**

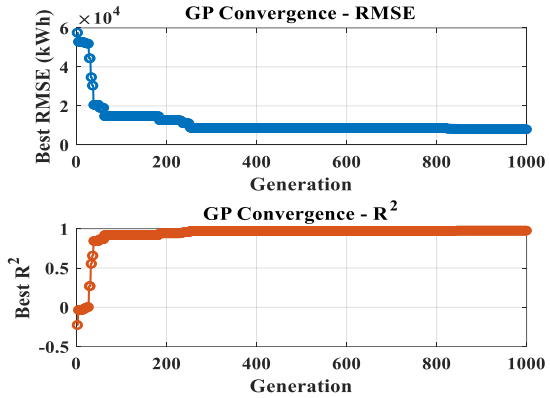



**Fig. 13. Convergence of the GP search over generations.**

The final evolved model achieved $R^2$ = 0.9765, RMSE = 7970.7 kWh, and MAE = 6404.9 kWh (Table 15), closely matching both RETScreen's and MATLAB's baseline regressions and statistically indistinguishable from their $R^2$ values (0.9762 and 0.976211, respectively). The resulting model disagreement, $\Delta R^2_{GP\text{-}RETScreen}$ =0.0003, is of the same order of magnitude as the RETScreen–MATLAB disagreement reported in Table 10 ($\Delta R^2$ =0.000011), confirming that GP—despite searching an unrestricted, structure-free space including non-linear operators such as — independently rediscovers the same linear relationship rather than favoring a more complex form. Furthermore, repeating the GP search across 15 independent runs (each with max Gen=1000) yielded a mean $R^2$ =0.9765 with standard deviation ±0.0000005, demonstrating that the evolutionary process reliably converges to a similar fit quality across runs. Collectively, this three-way cross-

validation—RETScreen, MATLAB, and GP—provides strong, independent confirmation that the linear CDD-based baseline model is not only statistically adequate but also the naturally optimal structure for this dataset, reinforcing the study's central finding that model simplicity does not compromise M&V reliability. Tables 13–15 summarize the genetic programming setup and outcomes. Table 14 lists the GP algorithm parameters used to evolve the linear-restricted model (population of 60, 1000 generations, tournament selection, arithmetic crossover, Gaussian mutation).

*Table 14*

GP Algorithm Parameters (Linear-Restricted Model, MATLAB Implementation)

| Parameter | Value |
|---|---|
| Population size $P$ | 60 |
| No. of Gen. $G$ | 1000 |
| Chromosome form | Expression tree (variable-length) $F = \{+,-,\times,\div,\sqrt{.}, \ln., \exp(.)\}$ $T = \{CDD, c\}$ |
| min/max initial tree depth | (2/5) |
| max tree depth after operations | (7) |
| Selection method | Tournament selection |
| Tournament size $k$ | 3 |
| Crossover prob. $p_c$ | 0.70 |
| Crossover type | Arithmetic (blend) crossover: $child = \alpha p_1 + (1-\alpha) p_2$, $\alpha \sim U(0,1)$ |
| Mutation prob. $p_m$ | 0.20 |
| Reproduction prob. | 0.10 |
| Elitism | Best individual preserved each generation |
| Fitness function | RMSE (minimization) |

Table 15 compares the resulting GP equation against the RETScreen and MATLAB baselines, showing an $R^2$ of 0.9765 and a $\Delta R^2$ of only 0.0003 relative to RETScreen — on the same order as the RETScreen–MATLAB discrepancy itself. Table 16 reports the GP model's error metrics (RMSE ≈ 7970.7 kWh, MAE ≈ 6404.9 kWh) and confirms convergence within 200–300 generations across 15 independent runs with negligible variance ($R^2$ = 0.9765 ± 0.0000005).

*Table 15*

Comparison among GP vs. RETScreen vs. MATLAB

| Method | Equation | $R^2$ | $\Delta R^2$ (vs. RETScreen) |
|---|---|---|---|
| RETScreen | Y = 326.9687X + 139685.9384 | 0.9762 | — |
| MATLAB | Y = 326.9669X + 139687.2334 | 0.976211 | 0.000011 |
| GP (this study) | Electricity = (-61323.64+(((CDD/log(CDD))*sqrt((sqrt(CDD)*(CDD+187487.87))))+203080.20)) | 0.9765 | 0.0003 |

*Table 16*

GP Results

| Parameter | GP Result |
|---|---|
| RMSE (kWh) | 7970.7± 0.01 |
| MAE (kWh) | 6404.9± 0.7 |
| $R^2$ | 0.9765± 0.0000005 |
| Number of generations to convergence | ~200–300 |
| Number of independent runs | 15 |

Together, these results show that even an unconstrained, structure-free search method independently converges on the same linear CDD-based relationship identified by RETScreen and MATLAB, reinforcing the robustness of the baseline model. It is obvious in Table 15, to avoid domain errors at CDD = 0 (present in 10 of 36 monthly observations), the logarithmic operator in the GP function set F was implemented as a protected log, returning log(CDD + ε) with ε = $1\times10^{-6}$ rather than an undefined value, following standard practice in symbolic regression implementations.

### *C. Limitations*

The high $R^2$ values warrant caution rather than evidence of general validity. Three factors likely explain the strong fit: (1) 10 of 36 monthly observations (28%) recorded CDD = 0, clustering over a third of the dataset at one input value and reducing the variability to explain; (2) the sample size—36 months, only 11 pre-retrofit is modest for treating $R^2$ as generalizable rather than in-sample; and (3) CDD was the sole predictor, leaving occupancy changes and holiday schedules uncontrolled. The high $R^2$ likely reflects this facility's specific profile—a stable base load with one clean seasonal cooling swing—rather than proof that a linear model generalizes to buildings with variable loads or multiple HVAC zones. Accordingly, the study's contribution is not the $R^2$

magnitude, but the cross-validation methodology: three independent approaches—commercial M&V software, custom regression, and structure-free evolutionary search—converge on the same model form. Future work should test this across more heterogeneous buildings.

## CONCLUSION

This study developed and validated a comprehensive M&V framework for a LED lighting retrofit in a commercial building in London, Ontario, Canada. Cross-validation between RETScreen and MATLAB quantified model uncertainty as minimal ($\Delta R^2$ = 0.000011), establishing linear regression as sufficient and reliable for practical M&V. Verified annual savings of 232,056 kWh yielded 27,846.72 CAD in cost savings and 23.57 $tCO_2$ in emission reductions. Critically, Symbolic Regression based on Genetic Programming, unconstrained by any assumed functional form, independently converged on the same linear relationship ($R^2$ = 0.9765), further reducing structural uncertainty and ruling out shared linear bias as an explanation for tool agreement. This three-way validation offers a stronger, replicable basis for uncertainty-aware baseline modeling in future M&V studies.

## REFERENCES


[1] *UK Green Building Council. Climate change mitigation. 2023.* Available at: https://ukgbc.org/our-work/climate-change-mitigation/ (Accessed: 2 December 2024).

[2] Carstens, H. A Bayesian approach to energy monitoring optimization. PhD thesis. *University of Pretoria, Pretoria, South Africa*, 2017. Available at: https://repository.up.ac.za/handle/2263/63791.

[3] International Performance Measurement and Verification Protocol (IPMVP). Concepts and options for determining energy and water savings. *Washington, DC, USA: IPMVP*, 2002. Available at: https://www.evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.

[4] ASHRAE. ASHRAE handbook—fundamentals (SI edition). *ASHRAE*, 2021. Available at: https://www.ashrae.org.

[5] Carstens, H., Xia, X., Yadavalli, S. Bayesian energy measurement and verification analysis. *Energies*, 2018, vol. 11, p. 380. Available at: https://doi.org/10.3390/en11020380.

[6] Walter, T., Price, P.N., Sohn, M.D. Uncertainty estimation improves energy measurement and verification procedures. *Applied Energy*, 2014, vol. 130, pp. 230–236. Available at: https://doi.org/10.1016/j.apenergy.2014.05.030.

[7] Blum, M., Riedmiller, M. Electricity demand forecasting using Gaussian processes. *In: Proceedings of the 15th AAAI Conference on Trading Agent Design and Analysis, Washington, DC, USA*, 14–15 July 2013. Available at: https://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6431.

[8] Burkhart, M.C., Heo, Y., Zavala, V.M. Measurement and verification of building systems under uncertain data: a Gaussian process modeling approach. *Energy and Buildings*, 2014, vol. 75, pp. 189–198. Available at: https://doi.org/10.1016/j.enbuild.2014.02.020.

[9] Reddy, T.A., Claridge, D.E. Uncertainty of "measured" energy savings from statistical baseline models. *HVAC&R Research*, 2000, vol. 6, pp. 3–20. Available at: https://doi.org/10.1080/10789669.2000.10391451.

[10] Neal, R. Bayesian learning for neural networks. *Berlin/Heidelberg, Germany: Springer Science & Business Media*, 2012. Available at: https://link.springer.com/book/10.1007/978-1-4612-0745-0.

[11] Rasmussen, C.E., Williams, C.K.I. Gaussian processes for machine learning. *Cambridge, MA, USA: MIT Press*, 2006. Available at: http://www.gaussianprocess.org/gpml/.

[12] Heo, Y. Bayesian calibration of building energy models for energy retrofit decision-making under uncertainty. PhD thesis. *Georgia Institute of Technology, Atlanta, GA, USA*, 2011. Available at:https://repository.gatech.edu/entities/publication/dff58eef-1b0d-4eb6-9963-6bec903eeea4.

[13] Wang, H. et al. Energy optimization for HVAC systems in multi-VAV open offices: a deep reinforcement learning approach. *Applied Energy*, 2024, vol. 356, p. 122354. Available at: https://doi.org/10.1016/j.apenergy.2023.122354.

[14] Nakthong, V., Kubaha, K. A simplified model of energy performance indicators for sustainable energy management. *IOP Conference Series: Earth and Environmental Science*, 2020, vol. 463, no. 1, p. 012046. Available at: https://doi.org/10.1088/1755-1315/463/1/012046.

[15] Müller, M., Christian, B. Insights on the impact of energy management systems on carbon and corporate performance: an empirical analysis with data from German automotive suppliers. *Journal of Cleaner Production*, 2016, vol. 137, pp. 1–13. Available at: https://doi.org/10.1016/j.jclepro.2014.06.013.

[16] Iwaro, J., Mwasha, A. A review of building energy regulation and policy for energy conservation in developing countries. *Energy Policy*, 2010, vol. 38, no. 12, pp. 7744–7755. Available at: https://doi.org/10.1016/j.enpol.2010.08.027.

[17] Du, X., Wen, L., Wei, P., Yang, M. A systematic approach for analyzing building energy conservation and emission reduction policies based on the principle of WSR. *Energy and Buildings*, 2024, vol. 315, p. 114328. Available at:

https://doi.org/10.1016/j.enbuild.2024.114328.

[18] Li, Z., Han, Y., Xu, P. Methods for benchmarking building energy consumption against its past or intended performance: an overview. *Applied Energy*, 2014, vol. 124, pp. 325–334. Available at: https://doi.org/10.1016/j.apenergy.2014.03.020.

[19] Kumar, A., Ranjan, S., Singh, M.B.K., Kumari, P., Ramesh, L. Electrical energy audit in residential house. *Procedia Technology*, 2015, vol. 21, pp. 625–630. Available at: https://doi.org/10.1016/j.protcy.2015.10.074.

[20] Sharma, P., Salkuti, S.R., Kim, S.-C. Energy audit: types, scope, methodology and report structure. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 22, no. 1, pp. 45–52.

[23] Vandenbogaerde, L., Verbeke, S., Audenaert, A. Optimizing building energy consumption in office buildings: a review of building automation and control systems and factors influencing energy savings. *Journal of Building Engineering*, 2023, p. 107233. Available at: https://doi.org/10.1016/j.jobe.2023.107233.

Newton, S., Shirazi, A., Christensen, P. Defining and demonstrating a smart technology configuration to improve energy performance and occupant comfort in existing buildings: a conceptual framework. *International Journal of Building Pathology and Adaptation*, 2023, vol. 41, no. 1, pp. 182–200. Available at: https://doi.org/10.1108/IJBPA-02-2022-0018.

[24] Márquez-Sánchez, S. et al. Enhancing building energy management: adaptive edge computing for optimized efficiency and inhabitant comfort. *Electronics*, 2023, vol. 12, no. 19, p. 4179. Available at: https://doi.org/10.3390/electronics12194179.

Available at: https://doi.org/10.11591/ijeecs.v22.i1.pp45-52.

[21] Salamacha, D., Józwik, J. Evaluation of measurement uncertainty obtained with a tool probe on a CNC machine tool. *Manufacturing Technology*, 2023, vol. 23, no. 4, pp. 513–524. Available at: https://doi.org/10.21062/mft.2023.080.

[22] Yelisetti, S., Saini, V.K., Kumar, R., Lamba, R., Saxena, A. Optimal energy management system for residential buildings considering the time-of-use price with swarm intelligence algorithms. *Journal of Building Engineering*, 2022, vol. 59, p. 105062. Available at: https://doi.org/10.1016/j.jobe.2022.105062.

**Information about author.**



**Hazem Hassan Ali Sayed,**
Renewable Energy Department, New Cairo Technological University, Egypt. Research interests: energy management, renewable energy technologies, power quality.
E-mail: hazem.hassan@nctu.edu.eg
ORCID: https://orcid.org/0000-0003-1093-853X