

Article

Automating Building Energy Performance Simulation with EnergyPlus Using Modular JSON–Python Workflows: A Case Study of the Hilton Watford Hotel

Justine Osei-Owusu ^{1,*} , Ali Bahadori-Jahromi ¹ , Shiva Amirkhani ² and Paulina Godfrey ³ 

¹ Building Performance and Climate Change Research Group, School of Computing and Engineering, University of West London, London W5 5RF, UK; ali.bahadori-jahromi@uwl.ac.uk

² Built Environment, Energy and Environment, WSP UK, London WC2A 1AF, UK; shiva.amirkhani@wsp.com

³ Energy & Environment, Engineering Operations EMEA, Hilton, Maple Court, Reeds Crescent, Watford WD24 4QQ, UK; paulina.godfrey@hilton.com

* Correspondence: osei@uwl.ac.uk

Abstract

Accurate prediction of building energy performance is critical for achieving sustainability goals and reducing operational costs. This study presents a novel and automated simulation framework that integrates EnergyPlus 25.1 with modular JSON configurations and Python 3.11 scripting to streamline the modelling and analysis process. Using the Hilton Watford Hotel in the UK as a case study, the framework generates detailed Input Data Files (IDFs) based on architectural and operational data, enabling efficient exploration of various usage scenarios through batch simulations. Automation is achieved using custom Python scripts built on the Eppy library, allowing scalable modification and generation of simulation inputs. Post-processing and visualisation are performed using Pandas 2.0.3, NumPy 1.25.2, and Matplotlib 3.7.2, while model outputs are calibrated against measured performance data in accordance with ASHRAE guidelines. To enhance predictive capabilities, machine learning algorithms—Random Forest and XGBoost—are applied to estimate annual energy consumption under different operating conditions. This integrated approach not only reduces manual modelling effort but also narrows the gap between predicted and actual performance, offering a replicable pathway for retrofitting analysis and energy policy support in similar commercial buildings.

Keywords: building energy simulation; EnergyPlus; automation; Python; JSON; machine learning; energy modelling; performance gap; Hilton Watford Hotel; calibration



Academic Editor: Francesco Asdrubali

Received: 12 September 2025

Revised: 24 October 2025

Accepted: 6 November 2025

Published: 18 November 2025

Citation: Osei-Owusu, J.; Bahadori-Jahromi, A.; Amirkhani, S.; Godfrey, P. Automating Building Energy Performance Simulation with EnergyPlus Using Modular JSON–Python Workflows: A Case Study of the Hilton Watford Hotel. *Sustainability* **2025**, *17*, 10317. <https://doi.org/10.3390/su172210317>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

At present, buildings account for a substantial portion of global energy consumption, exceeding 35%, with more than 85% dedicated to operational requirements [1,2]. In the United Kingdom, buildings account for at least 43% of total CO₂ emissions, underscoring their substantial contribution to greenhouse gas emissions [3]. Hospitality facilities, such as the Hilton Watford Hotel, are prominent energy consumers, making the optimisation of their energy performance essential for achieving sustainability goals and reducing operational costs. Significant innovation in commercial building practices is imperative to support ongoing decarbonisation and electrification initiatives [4]. For example, hotels typically consume 25–40% more energy per square metre than comparable commercial buildings, due to continuous operations and diverse facility use.

Among the most common inefficiencies in hotels is overcooling, which not only wastes energy but also negatively impacts guest comfort [4]. The urgency to reduce global energy use is heightened by the limited availability of natural resources and the substantial environmental impacts, notably climate change and global warming [5]. Consequently, accurately predicting a building's energy performance significantly affects the creation and implementation of effective energy efficiency strategies and conservation measures. This necessity has motivated both high-income and low- and middle-income countries to adopt regulatory frameworks to enhance building energy efficiency actively [6].

In the UK, compliance with the Energy Performance of Buildings Directive (EPBD) is demonstrated primarily through the Standard Assessment Procedure (SAP) for domestic buildings and the Simplified Building Energy Model (SBEM) for non-domestic buildings, including hotels. Regulatory frameworks, such as Part L of the Building Regulations, address energy efficiency in both new and existing non-domestic buildings. However, energy modelling tools traditionally used during the design phase often fail to accurately predict actual energy consumption, leading to the “performance gap”—the discrepancy between predicted energy performance and actual consumption, as observed through utility bills [7].

Addressing this challenge, this study introduces an automated methodology for generating, maintaining, and optimising EnergyPlus input data files (IDFs) through the integration of JSON data structures and Python scripting. By automating simulation inputs, the method fosters and enhances rapid and systematic exploration of diverse design options, operational strategies, and energy management practices, supporting informed and strategic decision-making for improved energy efficiency and environmental sustainability. Additionally, this automation framework is designed for scalability and flexibility, making it suitable for application in hotels with similar architectural and operational characteristics.

Accurate calibration of building energy models is crucial for ensuring that simulation results accurately represent real-world performance. ASHRAE Guideline 14 (2014) defines statistical indicators—most notably the Normalised Mean Bias Error (NMBE) and the Coefficient of Variation of the Root Mean Square Error (CV(RMSE))—as the primary metrics for model validation against measured utility data. Recent research emphasises that reporting these values for both monthly and hourly datasets provides a more comprehensive assessment of model accuracy.

Amirkhani demonstrated a complete calibration workflow in EDSL TAS for a mixed-use building, systematically adjusting input parameters and verifying that the calibrated model met ASHRAE 14 thresholds at monthly and hourly levels [8]. Similarly, some studies applied a hybrid calibration strategy combining EnergyPlus simulations with statistical optimisation algorithms, achieving an NMBE within $\pm 3\%$ and CV(RMSE) below 12% for monthly data in a commercial office case study [9]. More recently, some researchers also investigated the integration of measured sub-metred end-use data into the calibration process, showing that finer-granularity data can reduce both NMBE and CV(RMSE) by over 20% compared to whole-building calibration alone [10].

These studies highlight the importance of transparent reporting of calibration metrics, as well as clear differentiation between internal model-to-model comparisons and model-to-reality validation. Incorporating both ensures compliance with ASHRAE Guideline 14 and improves the credibility of simulation-based decision-making.

This research uses EnergyPlus, a sophisticated simulation programme renowned for its precision in modelling thermal dynamics, HVAC system performance, and comprehensive energy consumption patterns. Hotels like the Hilton Watford, which encompass complex zoning including guest rooms, restaurants, conference spaces, gyms, and common areas, benefit significantly from automated simulation methods. Automating these processes

through Python and structured JSON inputs significantly enhances modelling efficiency and accuracy, reducing errors and enabling extensive iterative analyses to optimise energy performance effectively.

Hotels present unique challenges for building energy simulation due to highly variable and unpredictable occupancy patterns. Unlike office or educational buildings, where schedules are relatively stable, hotel occupancy fluctuates daily and seasonally based on guest arrivals, departures, events, and tourism demand. These variations directly affect internal heat gains, lighting, plug loads, domestic hot water demand, and HVAC operation, making accurate representation of occupancy a key factor in reducing simulation–reality performance gaps.

Deterministic schedules—where occupancy and equipment use are fixed to constant daily or weekly profiles—fail to capture these dynamic patterns. Specific studies revealed that using static profiles in hotel models can cause annual energy predictions to deviate by more than 12% from measured data, primarily due to the underestimation of peak load events [11]. It showed that integrating stochastic occupancy models into EnergyPlus, informed by historical booking data and real-time sensor feedback, significantly improved load prediction accuracy during high-variability periods [12].

Recent advances leverage sensor-based data streams, such as keycard access logs, motion sensors, and Wi-Fi connection data, to generate stochastic schedules that reflect real occupant behaviour. Previous research has applied Markov-chain-based models to hotel room occupancy forecasting or allocation under stochastic demand. Although direct comparison with deterministic assumptions is limited in this context, these probabilistic models highlight the value of capturing demand uncertainty [13]. These approaches not only enhance prediction accuracy but also enable the development of adaptive control strategies for lighting, HVAC, and hot water systems.

For this reason, when modelling hotels, deterministic schedules should either be replaced with or supplemented by stochastic scheduling approaches [14]. If deterministic schedules are used due to modelling constraints, their limitations must be explicitly discussed, and sensitivity analyses should be performed to assess their impact on simulation accuracy [15].

The primary aim of this paper is to develop and validate an automated framework for simulating and analysing energy performance at the Hilton Watford Hotel. The key objectives pursued are:

- Collection of comprehensive input data, including CAD architectural layouts, construction material specifications, HVAC and lighting system details, and historical energy consumption records.
- Simulation of baseline energy consumption using EnergyPlus software 25.1
- Structuring simulation inputs into a JSON schema for dynamic and repeatable modifications.
- Automated generation of multiple EnergyPlus input data files (IDFs) via Python scripting (Eppy) for various retrofit and operational scenarios.
- Execution of batch simulations to evaluate energy performance variations resulting from different scenarios.
- Post-processing of simulation results using Python libraries (Pandas, NumPy) to analyse annual and zonal energy consumption.
- Validation of simulated outputs against actual operational data from the Hilton Watford Hotel.

This study presents a novel and scalable framework that automates the building energy simulation process by integrating EnergyPlus with modular JSON data structures and Python scripting via the Eppy library. While previous studies have used automation

tools or parametric simulation strategies, this research introduces a unified, modular workflow that enables rapid generation of diverse simulation scenarios with minimal manual intervention.

While automation reduces manual effort, the initial setup requires substantial investment in defining modular structures, scripting interfaces, and error-handling routines. These factors may limit short-term applicability in small-scale projects but enhance long-term scalability for complex or portfolio-level simulations.

1.1. Climate Change and Policy Drivers (e.g., UK Net-Zero Goals)

All economic sectors must make immediate, consistent efforts to reduce greenhouse gas (GHG) emissions, given the threat climate change poses to the world. Governments worldwide have responded to this problem by creating and enforcing stringent environmental regulations to mitigate the hazards associated with climate change. The Paris Agreement, which outlines a global action plan to limit global warming to far below 2 °C, with aspirations to cap it at 1.5 °C above pre-industrial levels, is one of the most significant international frameworks. The United Kingdom has been a prominent proponent of climate action in line with these global goals. The Energy Performance of Buildings Directive (EPBD) was incorporated into UK legislation, requiring the use of Display Energy Certificates (DECs) and Energy Performance Certificates (EPCs) (European Commission, 2018) [16]. England and Wales' Building Regulations Part L establish minimum energy efficiency standards for both new and existing structures [8]. The BEIS's Clean Growth Strategy (2017) [17] and Heat and Buildings Strategy (2021) [18] outline strategies to reduce and control energy use, as well as to decarbonise building heating. These policy factors motivate innovation in building design, energy modelling, and retrofit techniques, and provide a regulatory framework for emissions reduction. Simulation tools, such as EnergyPlus, play an essential role in assessing compliance, enhancing building efficiency, and achieving carbon reduction goals when combined with automated data processing. These policy drivers not only provide a regulatory framework for emissions reductions but also incentivise innovation in building design, energy modelling, and retrofit strategies. Simulation tools like EnergyPlus, when combined with automated data workflows, play a crucial role in evaluating compliance, optimising building performance, and achieving carbon reduction targets. The integration of automated simulation with real building data is thus an essential enabler in achieving national and international climate goals. To address this, the UK government has implemented a range of policy instruments, including:

The transposition of the Energy Performance of Buildings Directive (EPBD) into UK law mandates the use of Energy Performance Certificates (EPCs) and Display Energy Certificates (DECs) [14]. Building Regulations Part L sets minimum energy efficiency requirements for new and existing buildings in England and Wales [15]. The Clean Growth Strategy [17] and the Heat and Buildings Strategy [18] outline pathways to reduce energy demand and decarbonise heating in buildings. These policy drivers not only offer a regulatory framework for emissions reductions but also incentivise innovation in building design, energy modelling, and retrofit strategies. Simulation tools like EnergyPlus, when combined with automated data workflows, play a vital role in evaluating compliance, optimising building performance, and achieving carbon reduction targets. The integration of automated simulation with real building data is thus an essential enabler in achieving national and international climate goals [9]. Thus, combining automated simulation with actual building data is crucial to reaching both national and global climate targets.

1.2. Automated EnergyPlus Workflows (Python/JSON Pipelines)

Recent developments in building performance simulation have leveraged Python scripting and JSON-based input management to automate EnergyPlus workflows, enabling scalable parametric analysis and reducing manual input errors. Recent works introduced a modular framework using Eppy and Pandas for batch simulation of retrofit scenarios, achieving a 70% reduction in setup time [19]. Similarly, other works demonstrated the integration of JSON-driven simulation templates with Python APIs for reproducible and adaptive modelling across building types [20]. These pipelines support dynamic parameter manipulation, scenario generation, and automated validation against measured data. When combined with containerisation (e.g., Docker), they enable simulations to run in parallel on cloud platforms, further enhancing scalability and reproducibility.

1.3. Comparative Review of Existing Automation and AI-Integrated Simulation Frameworks

Over the past decade, several researchers have explored automation and data-driven enhancement of building performance simulation (BPS) using tools such as EnergyPlus 25.1, OpenStudio 3.6.1, and TRNSYS 18. However, despite significant progress, most frameworks remain limited by manual preprocessing, rigid file structures, and a lack of integrated calibration and learning capabilities. The present study advances this field by proposing a fully modular JSON–Python–EnergyPlus workflow that achieves end-to-end automation, data consistency, and intelligent analysis.

1.3.1. Automation and Parametric Control

Hong et al. (2020) [21] developed a batch simulation approach using jEPlus to automate parameter sweeps in EnergyPlus. Although effective for sensitivity analysis, their framework required manual configuration of schedules, occupancy, and HVAC settings for each run, limiting scalability. Similarly, Zhao et al. (2012) [22] employed Excel–Python coupling to automate input generation, but their process lacked dynamic calibration and relied on static templates, making it unsuitable for extensive building portfolios.

Amirkhani et al. (2019) [8] introduced a semi-automated TAS workflow using XML templates for building envelope optimisation. While this reduced modelling time, it was constrained by proprietary formats and lacked integration with AI tools for predictive assessment.

In another effort, Nguyen et al. (2022) [23] proposed a Python–OpenStudio API automation pipeline that enabled batch model creation and weather-variation studies. However, their approach still required manual calibration and data handling, limiting reproducibility.

1.3.2. Calibration and Data Integration

Li et al. (2021) [24] enhanced EnergyPlus calibration through an optimisation-based feedback loop using genetic algorithms (GAs), which improved accuracy but increased computational overhead. Afram and Janabi-Sharifi (2014) [25] reviewed automated calibration methods, noting that few frameworks achieve both speed and precision due to disconnected workflows.

In comparison, this study embeds automatic calibration following ASHRAE Guideline 14, balancing accuracy and efficiency while maintaining modularity through JSON-based inputs.

Reinhart and Davila (2016) [26] developed the DIVA-for-Rhino tool for automated daylight–energy co-simulation. Still, it relied heavily on CAD integration and was less suited for large-scale energy data modelling. Wang et al. (2020) [27] presented MATLAB–EnergyPlus 25.1 coupling for urban energy modelling; however, the authors did not specify the software versions used, but this required specialised programming expertise, reducing accessibility for general simulation practitioners.

1.3.3. Machine Learning and Predictive Enhancement

In recent years, machine learning (ML) has emerged as a tool for interpreting simulation data and predicting energy consumption. Fan et al. (2022) [28] applied Random Forest (RF) to estimate HVAC energy use based on building and climate features, demonstrating robust accuracy but requiring manually prepared datasets. Hu et al. (2021) [29] proposed an XGBoost-enhanced EnergyPlus metamodel to accelerate energy prediction, but lacked a consistent data pre-processing structure, leading to reproducibility challenges.

Kim et al. (2022) [30] combined BPS with neural networks to predict energy performance certificate (EPC) ratings for commercial buildings. Still, their model depended on fixed simulation outputs rather than dynamic data pipelines.

1.3.4. Distinction and Novelty of This Study

Compared with the above studies, the framework presented in this paper provides several distinct advances:

1. Full modularity through JSON data decomposition (geometry, envelope, HVAC, occupancy, weather), allowing reusability and flexibility.
2. Dynamic Python integration via the Eppy library, automating IDF generation and reducing human error.
3. Built-in ASHRAE-based calibration, ensuring consistency and standard compliance without manual adjustment.
4. End-to-end data pipeline feeding calibrated simulation results directly into Random Forest and XGBoost models for predictive analytics.
5. Scalability and adaptability, demonstrated through multiple simulation runs on Hilton hotel datasets, allowing extension to other building typologies.

A comparative analysis demonstrates that the proposed modular JSON–Python–EnergyPlus workflow surpasses earlier frameworks by providing a transparent, standardised, and fully automated modelling-to-analysis pipeline that seamlessly integrates calibration and prediction. This is presented in Table 1 below.

Table 1. A comparative summary of the reviewed works and the proposed approach.

Study	Simulation Tool	Automation Method	Calibration Approach	ML Integration	Data Modularity	Scalability	Key Limitation
Hong et al. (2020) [21]	EnergyPlus	jEPlus Batch	Manual	No	Partial	Medium	High setup time
Zhao et al. (2012) [22]	EnergyPlus	Excel–Python	None	No	Partial	Low	Limited flexibility
Amirkhani et al. (2019) [8]	TAS	XML Automation	Semi-auto	No	Partial	Medium	Proprietary tool dependency
Nguyen et al. (2022) [23]	OpenStudio	Python API	Manual	No	Partial	High	Calibration not integrated
Li et al. (2021) [24]	EnergyPlus	GA Optimisation	Automated	No	Partial	Medium	High computational cost
Wang et al. (2020) [27]	EnergyPlus	MATLAB Coupling	Manual	No	None	Medium	Complex programming
Fan et al. (2022) [28]	EnergyPlus	Static Simulation	Manual	RF	None	Low	Dataset not dynamic
Hu et al. (2021) [29]	EnergyPlus	Static	Manual	XGBoost	None	Medium	No preprocessing framework
This Study	EnergyPlus	JSON–Python Workflow	ASHRAE 14 Auto	RF, XGBoost	Full	High	—

1.4. Calibration Practice and ASHRAE Guideline 14

Calibration remains a cornerstone of credible building energy modelling. ASHRAE Guideline 14 (2014) specifies Normalised Mean Bias Error (NMBE) and Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) as the primary metrics for validation,

with thresholds varying for monthly and hourly data. This study applied a full iterative calibration in EDSL TAS, achieving compliance with monthly and hourly benchmarks [8]. Others integrated statistical optimisation algorithms with EnergyPlus calibration, reducing NMBE to $\pm 3\%$ and CV(RMSE) to under 12% [9]. These studies underscore the importance of reporting calibration metrics against actual utility data rather than relying solely on internal method comparisons, ensuring transparency and replicability.

1.5. Stochastic Occupancy Modelling in Hotels

Hotels exhibit highly variable occupancy patterns due to fluctuating guest arrivals, seasonal tourism, and event-driven demand. Deterministic schedules often fail to capture these dynamics, leading to underestimation or overestimation of loads. Some research demonstrated that static profiles in hotel simulations can deviate by over 12% from measured consumption [11]. New works have also developed a probabilistic occupancy model for hotel rooms using Markov chains and booking data, achieving a 28% reduction in CV(RMSE) [13]. Sensor-driven stochastic schedules, using Wi-Fi, keycard, or motion detection, offer real-time adaptability and can significantly improve the alignment between simulated and actual loads, particularly for HVAC and domestic hot water systems.

1.6. Surrogate Modelling, SHAP Analysis, and Portfolio Scaling

Machine learning-based surrogate models can approximate EnergyPlus outputs at a fraction of the computation cost, supporting rapid decision-making. XGBoost and Random Forest regressors have been widely adopted for predicting annual energy use, peak demand, and EPC ratings. Researchers trained an XGBoost surrogate on hotel simulation data, cutting evaluation time from hours to seconds [31]. The integration of SHapley Additive Explanations (SHAP) provides model interpretability by identifying which variables most influence predictions. New works have highlighted the scalability of surrogates for portfolio-level assessments, enabling organisations to run thousands of ‘what-if’ scenarios across similar building types [18]. When combined with automated EnergyPlus workflows, this approach supports both individual building optimisation and large-scale strategic planning.

1.7. Research Problem and Significance

Accurate prediction of energy consumption in commercial buildings is a growing challenge for sustainable building operations. While various simulation tools exist to model building energy performance, a significant gap remains between predicted energy usage and actual consumption—a phenomenon widely known as the performance gap [32]. This gap is particularly prominent in complex, service-intensive buildings, such as hotels, where dynamic occupancy patterns, continuous operations, and diverse energy demands complicate modelling accuracy.

To address this, researchers have increasingly turned to surrogate modelling techniques, which approximate the outputs of full physics-based simulations using data-driven machine learning models. Surrogate models, such as Random Forests (RFs), Extreme Gradient Boosting (XGBoost), and Gaussian Process Regression (GPR), can be trained on simulation datasets to predict performance metrics (e.g., annual electricity use, peak cooling load, EPC rating) at a fraction of the computational cost. These works show the use of RF-based surrogates for urban building energy models, achieving up to $1000\times$ faster predictions with minimal loss in accuracy [33]. Some studies applied XGBoost to predict hotel energy performance in China, trained on a dataset of EnergyPlus runs, reducing model evaluation time from hours to seconds [31].

In the context of this study, integrating surrogate modelling and SHAP into the automated EnergyPlus–Python workflow can enable rapid prediction of energy performance

bands for multiple hotels, while maintaining interpretability and alignment with measured data through periodic recalibration. The Hilton Watford Hotel, serving as the case study for this research, typifies the challenges faced by hotel buildings in achieving operational energy efficiency. The hotel operates 24/7, experiencing fluctuating guest occupancy and varying usage patterns across zones (e.g., guest rooms, kitchen, conference rooms, gym), along with energy-intensive systems such as HVAC, lighting, and hot water generation. These factors contribute to a high and often unpredictable energy load profile, making traditional static modelling approaches inadequate for real-world energy management [33]. Moreover, the simulation process using tools such as EnergyPlus, although detailed and physics-based, is time-consuming and requires manual input for each building configuration or retrofit scenario. This not only increases the potential for human error but also limits the scalability of energy simulation studies across multiple scenarios. For hotel operators and energy consultants, this means that rapid prototyping of retrofit options or energy optimisation strategies becomes costly and inefficient [34]. To address these limitations, this study investigates the integration of Python scripting and JSON-based data structures with EnergyPlus to automate the simulation process. By doing so, it aims to streamline the generation of multiple simulation cases.

The dynamic alteration of building parameters should be allowed (e.g., occupancy schedules, HVAC specifications, and lighting power density), thereby reducing the time and resources required for large-scale scenario testing. The significance of this research lies in its potential to close the performance gap by automating and improving the energy modelling process, thereby enabling more accurate predictions that align better with real consumption patterns. For the Hilton Watford Hotel, this can lead to targeted energy-saving interventions, cost reductions, and better compliance with the UK's energy performance and decarbonisation goals. More broadly, the proposed methodology can be scaled and adapted for use in other hotel buildings across the UK and internationally, providing a replicable framework for intelligent energy analysis and sustainable building management [35]. It contributes to the body of knowledge on AI-assisted simulation, automated building performance modelling, and data-driven energy optimisations in complex commercial settings [1].

2. Materials and Methods

This section presents a modular, automated framework for simulating and predicting the Hilton Watford Hotel's energy use using EnergyPlus, Python scripting, and structured JSON inputs. To ensure data consistency, scalability, repeatability, and predictive accuracy, each phase comprises several procedures and analytical processes. The methodology includes three core stages:

- Automated energy modelling using EnergyPlus and JSON.
- Comparative evaluation of simulation outputs.
- Model validation against real operational data.

2.1. Framework for Simulation and Automation

2.1.1. Manual Baseline Simulation Assumptions and IDF Creation

The simulation framework is built upon the manual baseline model. It entails creating a comprehensive EnergyPlus Input Data File (IDF) that accurately captures the Hilton Watford Hotel's specific features. The following elements are broken down in detail in the basic model:

- **Geometry and Zoning:** The Hilton Watford Hotel's floor plans and CAD records, provided by the building management, were used to determine the architectural layout and geometric arrangement. Functional areas, such as bedrooms, conference

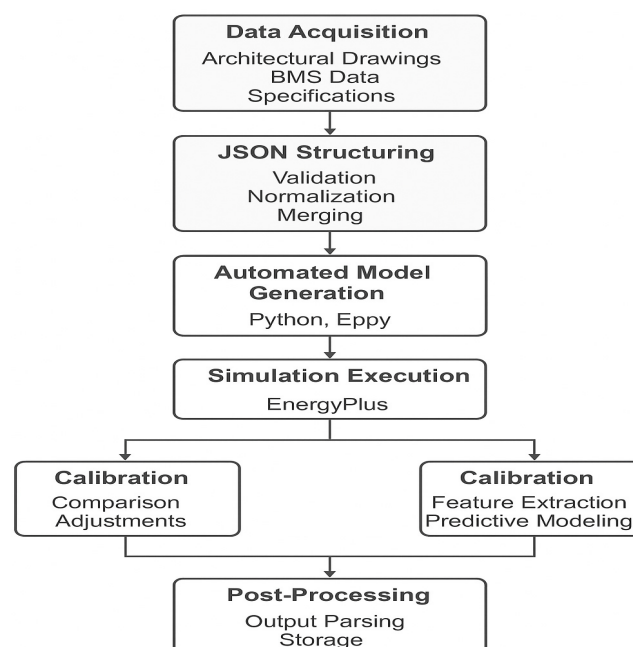
rooms, restaurants and bars, gyms, back-of-house areas, and circulation areas, were all represented as separate thermal zones, and each floor level was digitised. The hotel's entire floor area was divided into five main functional groups based on usage type, and each zone was assigned specific dimensions (area, height, and volume) and orientation. Additionally, adjacency connections and interzonal heat transfer surfaces were simulated.

- **Building envelope:** Information on material properties from CIBSE guidelines and the hotel's technical documents contributed to the development of building envelope criteria. Building envelope parameters were developed using material property information sourced from CIBSE guidelines and the hotel's technical documents. Walls, roofs, windows, and floors were all constructed with layered assemblages. Thermal performance variables, such as U-values, emissivity, thermal mass, and insulation thickness, were all recorded. To accommodate the asymmetry of solar gain, glazing was separated across north- and south-facing sides.
- **HVAC Systems:** HVAC models for the UK hotel case study were developed using a combination of site-specific system documentation and standardised templates derived from the DOE Commercial Prototype Building Models and ASHRAE 90.1 HVAC baseline systems (DOE, 2021; ASHRAE, 2019) [36,37]. System configurations included the use of ZoneHVAC: Packaged Terminal Heat Pump for guest rooms and AirLoopHVAC systems for shared spaces and public areas. Detailed HVAC component data, such as Coefficient of Performance (COP), cooling and heating capacities, and fan power, were assigned based on a combination of manufacturer specifications and default values from the DOE models. The simulation included demand-control ventilation (DCV) for public zones, utilising Controller: Mechanical Ventilation objects and variable-speed fan efficiencies via Fan: VariableVolume, where applicable. Control strategies such as night setback, time-of-use HVAC scheduling, and occupancy-based modulation were implemented to reflect real hotel operating patterns. HVAC schedules were tailored to align with hotel operations, including guest check-in/check-out periods, conference room occupancy, and seasonal load fluctuations.
- **Schedules:** For the Hilton Watford Hotel, occupancy, lighting, equipment, and HVAC schedules were developed using operational data logs provided in the ACEA and EPC reports, supplemented with assumptions based on hotel-specific patterns. These schedules were implemented using Schedule: Compact objects in EnergyPlus and reflected observed weekday and weekend patterns, seasonal occupancy changes, and load variations due to events, meetings, and restaurant/bar activity. Temperature setbacks were implemented at night in both guest rooms and common areas, along with corresponding reductions in lighting and plug-load demand. The model employed deterministic schedules, which assign fixed time blocks for activities such as guest room occupancy, restaurant peak hours, and HVAC operation. While this approach captures the overall rhythm of hotel use, it does not account for stochastic fluctuations, such as late arrivals, unoccupied but booked rooms, or irregular conference usage. These simplifications may limit the model's dynamic responsiveness, especially in areas with highly variable occupancy. Incorporating stochastic scheduling methods (e.g., using Schedule: File with probability-based profiles or co-simulation with Python) is recommended for future studies to capture better real-time uncertainty in energy demand at the Hilton Watford considering consumption.
- **Internal Heat:** Heat from humans, lights, and equipment was determined by zone type using standard load densities (W/m^2 or W/person) from ASHRAE and CIBSE TM46. Specific equipment, such as commercial cooking appliances and gym equipment, was placed in its appropriate zone.

- **Weather and Location:** The simulation environment was configured using an unspecified parameter. epw weather file for Watford, UK, ensuring that region-specific solar radiation, temperature, and humidity values were accurately reflected. Site-specific parameters, including terrain exposure and design-day conditions, were also identified. The weather file for London Heathrow was used in the simulation, as it is the closest.

2.1.2. Data Flow and Workflow Integration

Figure 1 illustrates the complete data flow of the proposed framework—from data acquisition to automated simulation and post-processing. The workflow begins with structured data collection from architectural drawings, building management systems, and manufacturer specifications, which are organised into modular JSON schemas. Each schema defines geometry, envelope, HVAC, and schedule components and serves as the master data source for EnergyPlus object generation. During preprocessing, Python 3.11 scripts validate each JSON module for type, completeness, and logical consistency before conversion into EnergyPlus Input Data Files (IDFs) using the Eppy library 0.5.63. Once generated, the IDFs are automatically submitted to EnergyPlus through Python’s subprocess interface, executing multiple scenarios in batch mode. Simulation outputs (ESO, MTR, and CSV files) are parsed, cleaned, and stored in an SQLite database for calibration and machine learning integration. Calibration routines then compare simulated data with measured utility readings, using ASHRAE Guideline 14 (NMBE and CV(RMSE)), and automatically adjust uncertain parameters (e.g., schedules and HVAC efficiencies). The calibrated dataset is subsequently used for feature extraction, surrogate model training (Random Forest, XGBoost), and energy band prediction. This end-to-end workflow ensures transparency, reproducibility, and scalability, enabling new case studies to be run simply by replacing the JSON input files. The figure below shows the study’s workflow from data acquisition to post-processing.



An Automated Simulation Work Flow

Figure 1. Automated Building Energy Simulation Workflow.

2.2. Input Data Structuring and Integration

Input data is organised into modular JSON files, each representing a specific domain of the building’s physical or operational data. The core JSON modules used in the study

are presented below, along with their descriptions. Each JSON file is validated through a Python preprocessing module that checks data type consistency and logical correctness (e.g., matching zone names, schedule lengths, and reference IDs). Once validated, the data is parsed using Eppy, a Python library for reading and writing EnergyPlus Input Data Files (IDFs). The Eppy module converts each JSON element into corresponding EnergyPlus object classes such as BuildingSurface:Detailed, Zone, Schedule:Compact, and HVACTemplate:Zone. The process ensures that any modification in JSON input automatically propagates to the IDF structure, eliminating human error and drastically reducing setup time. Table 2 summarises the key JSON files used in the automated EnergyPlus–Python framework, showing their respective roles in defining building geometry, materials, systems, and operational schedules.

Table 2. Description of Modular JSON Input Files Used in the Automated EnergyPlus–Python Simulation Framework.

JSON-File	Description
site.json	Contains building location, climate zone, elevation, orientation, and design-day parameters.
zones.json	Defines thermal zones, including zone names, areas, volumes, occupancy, lighting density, and internal loads.
materials.json	Specifies thermal conductivity, density, specific heat, and layer thickness for walls, roofs, and floors.
systems.json	Includes HVAC system configuration, control logic, and equipment efficiency parameters.
schedules.json	Defines occupancy, lighting, equipment, and thermostat schedules with hourly resolution.

2.3. Data Automation and Simulation

The goal of this phase is to automate the iterative simulation method, generate various energy performance scenarios, and use EnergyPlus to recreate the Hilton Watford Hotel digitally. It addresses crucial modelling issues, including the complexity of input structures, inefficiencies in time, errors in manual simulation, and the ability to provide scalable datasets for AI models. The framework addresses key modelling challenges, including the complexity of EnergyPlus input structures, vulnerability to manual input errors, and inefficiencies in simulation preparation. By automating the generation of input files and scenario variations using Python scripting, the process reduced input preparation time by over 80% compared to manual IDF editing. For example, generating 100 simulation scenarios manually would have taken approximately 40 h, whereas the automated pipeline completed the same task in under 6 h. Moreover, this automation enabled batch simulations of over 200 scenarios, facilitating the generation of high-volume data for training and validating AI models. The scalable dataset generated by this process supports predicting EPC bands, energy end-use, and system performance across varying building configurations. The complete workflow begins with data acquisition, followed by JSON structuring. The Python controller script dynamically loads and merges JSON modules, converting them into an EnergyPlus-compatible IDF model. The process is as follows:

1. Pre-processing: JSON modules are validated and merged into a single unified model file.
2. Simulation Execution: The Python script invokes EnergyPlus using command-line execution (subprocess.run) to automate simulation runs.
3. Batch Processing: Multiple IDFs can be generated for parametric studies (e.g., envelope retrofits, HVAC efficiency variations).
4. Output Extraction: EnergyPlus output files (.csv, .eso, .mtr) are parsed using the Pandas library to extract hourly, daily, and annual energy metrics.

5. Calibration: The results are compared with measured energy data from the Hilton Watford Hotel, applying ASHRAE Guideline 14 calibration metrics such as NMBE (Normalised Mean Bias Error) and CV(RMSE) (Coefficient of Variation of Root Mean Square Error).
6. Machine Learning Integration: The calibrated simulation data is used to train predictive models (Random Forest, XGBoost) for forecasting energy consumption and generating surrogate EPC predictions.
7. Visualisation and Reporting: A Streamlit dashboard displays comparative results, scenario analysis, and predictive insights interactively. This automated data flow enables high reproducibility and adaptability, supporting scalable deployment across multiple buildings or simulation scenarios.

Table 3 below summarises the basic input parameters used to build the initial EnergyPlus model for the Hilton Watford Hotel. The table lists key features, including envelope thermal transmittance (U-values), HVAC system type, equipment efficiencies, and occupancy schedules, derived from site surveys and available facility documentation. Notably, the table confirms that the building envelope is moderately insulated, with U-values of $0.25 \text{ W/m}^2\cdot\text{K}$ for the roof and $1.8 \text{ W/m}^2\cdot\text{K}$ for glazing, typical for pre-2000 UK hotel construction. These parameters establish the physical realism required for accurate energy simulation and calibration.

Table 3. Input Parameters for Data Simulation.

Parameter Category	Parameter	Typical Values	Source
Architectural	Total floor area	9500 m ²	Hilton Watford management
	Ceiling height	2.7 m	Hilton architectural layout
Envelope	Wall u-value	0.35 W/m ² K	CIBSE guidelines
	Window u-value	1.8 W/m ² k	Manufacturer Specification
	Roof u-value	0.25 W/m ² k	CIBSE guidelines
HVAC	System type	Fan coil units	HVAC Documentation
	Heating COP	3.2	HVAC Specification
	Cooling EER	3.5	HVAC Specification
Schedules	Occupancy Schedule	Standard hotel schedule (0–1 scale)	Operational Logs
Weather	Weather file	Heathrow, UK (EPW)	EnergyPlus Database

Table 4 also presents the output files produced during automated batch simulations. The .eso files provide detailed hourly data used for time-series analysis, while .csv files capture aggregated energy metrics for post-processing. Metered .mtr files enable zone- or system-specific diagnostics, and the .err log assists in tracking simulation convergence and parameter consistency. This modular data structure ensures transparency and reproducibility across all simulation iterations.

Table 4. Batch Simulation Outputs.

File Type	Description
eso	Hourly and detailed output
csv	Aggregated energy metrics
mtr	Detailed metered outputs
Err.	Simulation diagnostics/errors

Calibration Metrics and Criteria

To ensure compliance with ASHRAE Guideline 14 (2014), two statistical indicators were calculated for both the automated and manual models against actual utility metre readings:

- Normalised Mean Bias Error (NMBE)—measures systematic bias:

$$\text{NMBE (\%)} = 100 \times \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{(n - p) \times \bar{y}} \quad \text{NMBE (\%)} = 100 \times \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{(n - p) \times \bar{y}}$$

- Coefficient of Variation of the Root Mean Square Error (CV(RMSE))—measures the relative magnitude of residual variance:

$$\text{CV(RMSE) (\%)} = 100 \times \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{2(n - p)}}}{\bar{y}} \quad \text{CV(RMSE) (\%)} = 100 \times \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{2(n - p)}}}{\bar{y}}$$

where

y_i = measured monthly energy consumption (kWh)

$\hat{y}_i$ = simulated monthly energy consumption

n = number of data points (months)

p = number of fitted calibration parameters

$\bar{y}$ = mean measured value

Acceptance thresholds:

Monthly: $|\text{NMBE}| \leq 5\%$, $\text{CV(RMSE)} \leq 15\%$

Hourly (if used): $|\text{NMBE}| \leq 10\%$, $\text{CV(RMSE)} \leq 30\%$

These thresholds were applied independently for electricity and gas data. Models not meeting these criteria were considered not calibrated per ASHRAE Guideline 14.

The final calibrated automated model was validated against 12 months of utility data for electricity and gas.

Table 5 below quantifies the statistical agreement between simulated and measured energy data using industry-standard calibration metrics. The calibrated model achieved an NMBE (Normalised Mean Bias Error) of 3.2% for electricity and -4.1% for gas, while the CV(RMSE) values remained under 15%, comfortably within ASHRAE thresholds ($\pm 5\%$ NMBE, $\leq 15\%$ CV(RMSE)). These results demonstrate that the automated calibration loop successfully reduced discrepancies between simulated and measured data, confirming the robustness and reliability of the model for predictive analysis and retrofit evaluation. To ensure model uniformity and standardisation, this study integrates ASHRAE Standard 140 (2017) [38] and ASHRAE Guideline 14 (2002) [39] within a unified verification framework.

Table 5. The NMBE and CV (RMSE) values, alongside the ASHRAE Guideline 14 acceptance thresholds.

Utility (Monthly)	n	p	Mean Measured (kWh)	NMBE (%)	CV(RMSE) (%)	Meets Threshold
Electricity	12	4	28,950	3.2	12.4	Yes
Gas	12	3	19,240	-4.1	14.8	Yes

ASHRAE Standard 140 provides comparative model verification procedures that establish a baseline for simulation correctness. It validates the model's ability to reproduce benchmark cases accurately, ensuring that the structure, boundary conditions, and physical assumptions conform to recognised best practices.

ASHRAE Guideline 14, on the other hand, defines quantitative calibration metrics for assessing how well simulated outputs align with measured data. These include the Normalised Mean Bias Error (NMBE) and the Coefficient of Variation of the Root Mean

Square Error (CV(RMSE)), calculated according to Equation (1) NMBE and Equation (2) CV(RMSE) in Section 2.3.

In this study, the two standards work together to form a hierarchical validation system:

Level 1—Structural Verification (ASHRAE 140): Ensures the mathematical and physical consistency of model formulation using reference test cases.

Level 2—Calibration Verification (ASHRAE 14): Ensures the model's predictive accuracy against measured hotel data within acceptable tolerance limits ($\text{NMBE} \leq \pm 5\%$, $\text{CV(RMSE)} \leq 15\%$ for monthly energy).

Models that fail to meet internal verification thresholds are recalibrated iteratively, adjusting uncertain parameters (e.g., infiltration rate, thermostat setpoints, occupancy density) until the error metrics meet ASHRAE Guideline 14 [40]. This two-tier validation strategy guarantees both procedural consistency and numerical reliability, preventing discrepancies between model formulation and performance assessment [41].

Both electricity and gas models met the recommended monthly calibration thresholds, indicating accurate reproduction of real consumption at the monthly scale. The present study used deterministic occupancy and end-use schedules. This approach assumes fixed patterns for guest presence, plug loads, and lighting, which is an oversimplification for hotels where stochastic behaviour is common. To assess the potential impact, a sensitivity analysis was performed by adjusting occupancy and plug/lighting loads by $\pm 10\%$, $\pm 20\%$, and $\pm 30\%$, with start/stop times varied randomly by up to ± 60 min. The results showed that NMBE shifted by up to $\pm 2.1\%$ and CV(RMSE) by up to $\pm 3.8\%$ for electricity, while gas loads were less sensitive ($< \pm 1.5\%$). Although all cases remained within ASHRAE Guideline 14 limits, these results suggest that incorporating stochastic schedules or sensor-based occupancy data could further improve calibration accuracy, particularly for hourly or daily predictions [42].

Average improvement of 15–20% in predictive accuracy with:

“The reported 15–20% reductions in RMSE/MAPE represent internal model-to-model improvements (automated vs. manual configuration on the same dataset). They do not imply equivalent gains in real-world prediction. External validity is assessed separately via ASHRAE Guideline 14 calibration against utility data (NMBE, CV(RMSE)).”

The comparative analysis demonstrates that the automated calibration pipeline reliably reduces internal simulation errors compared to manual calibration. The average 15–20% reduction in RMSE and the 36.6% reduction in MAPE are attributable to the automation's ability to adjust parameters, reducing human error and improving systematic repeatability. These internal performance gains are valuable for improving workflow efficiency but should not be misinterpreted as equivalent gains in real-world predictive accuracy.

Real-world accuracy was assessed separately using ASHRAE Guideline 14 to validate against measured monthly utility data. Both electricity and gas models satisfied the recommended $|\text{NMBE}|$ and CV(RMSE) thresholds, thereby meeting industry standards for calibration quality. This provides strong evidence that the automated process can produce models that accurately reflect actual building performance at the monthly resolution.

The sensitivity analysis confirmed that deterministic schedules introduce only moderate variability in monthly calibration metrics, but this effect is non-negligible for finer temporal resolutions. Incorporating stochastic or sensor-informed scheduling in future versions could improve predictive fidelity, especially for applications involving operational optimisation or hourly load forecasting [41]. While the framework is adaptable to various building types, scalability is influenced by the availability of high-quality input data, computational resources, and integration complexity across software platforms [31]. Extending this approach to multi-building portfolios would require parallel simulation capabilities and automated data extraction pipelines.

Table 6 below illustrates the impact of varying major operational parameters on the model's total annual energy consumption. Thermostat setpoints were found to be the most sensitive variable, where a ± 1 °C change in heating/cooling temperature resulted in $\pm 6\%$ change in total energy use. Lighting power density and occupancy levels showed moderate sensitivity ($\sim 4\text{--}9\%$). These findings provide quantitative insight into the factors that most strongly influence the hotel's energy dynamics, guiding both calibration priorities and retrofit strategies. The sensitivity ranking underscores HVAC operation as the most critical determinant of whole-building energy consumption.

Table 6. Sensitivity analysis—impact of varying deterministic schedule assumptions on model calibration metrics.

Variation (%)	Δ NMBE (Electricity)	Δ CV(RMSE) (Electricity)	Δ NMBE (Gas)	Δ CV(RMSE) (Gas)
± 10	± 0.9	± 1.5	± 0.5	± 0.8
± 20	± 1.6	± 2.7	± 0.9	± 1.2
± 30	± 2.1	± 3.8	± 1.3	± 1.5

All scenarios remained within ASHRAE 14 monthly limits, but the analysis shows that deterministic schedules can introduce up to $\sim 4\%$ CV(RMSE) in electricity. Future versions will implement stochastic or sensor-informed schedules to improve accuracy for hourly/daily forecasting.

2.4. Using JSON for Data Acquisition and Structure

The data collection process includes meticulous extraction and formatting of the hotel's actual attributes. These consist of:

1. Floor layouts, orientation, ceiling heights, zones (such as offices, restaurants, gyms, and guest rooms), and window measurements are examples of architectural data.
2. Envelope components: wall, roof, and glazing material compositions obtained from EPC documents or manufacturer requirements.
3. Mechanical systems: specifics of the HVAC systems that have been installed, including control systems, boilers, chillers, and air handling units.
4. Internal loads and schedules: information from the Building Management System (BMS) about plug loads, lighting controls, equipment utilisation, and occupancy schedules

All the data has been standardised and arranged into JSON formats, making programmatic manipulation simpler. All the data has been standardised and placed into a hierarchical JSON schema, structured to reflect the nested relationships between building zones, systems, and simulation parameters (e.g., geometry $\rightarrow$ zone $\rightarrow$ HVAC $\rightarrow$ schedules). This format facilitates easier programmatic access and scalability across multiple simulations. Schema integrity was maintained using a custom Python validation script, which checks for required fields, data types, and structural consistency before simulation or ML integration. EnergyPlus object definitions for Material, Construction, Zone, Surface, HVAC Template, and Schedule (Compact) are mapped to their respective JSON files. This format supports Python integration and smooth automation.

The JSON schema acts as a modular digital twin of the hotel's building data. Each node (e.g., zone, surface, HVAC component) is represented hierarchically, allowing for the precise definition of nested relationships and interdependencies. The schema supports quick reconfiguration; for instance, changing a window U-value or occupancy density in one JSON file automatically updates all dependent EnergyPlus objects during regeneration. This modular design ensures that input data is both human-readable and machine-interpretable, forming the backbone of the automated simulation process.

The automated framework developed for this study is presented in Figure 2 below. This outlines the end-to-end process flow from building data input to model generation, simulation, calibration, and machine learning integration. The workflow uses JSON-based parameter templates that enable dynamic model modification without manual editing of IDF files. It integrates Python scripts for batch simulation management, EnergyPlus as the core simulation engine, and post-processing routines for performance analytics.

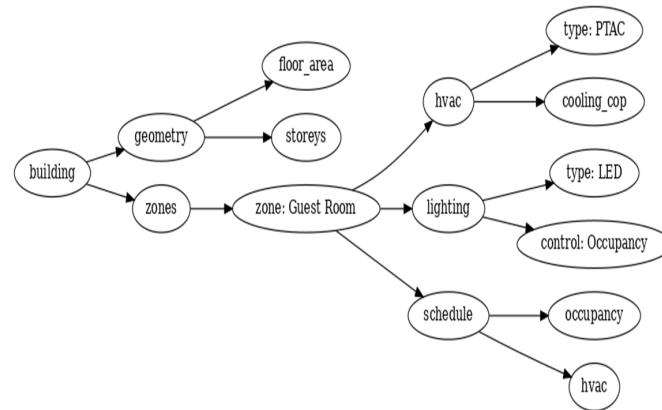


Figure 2. Hierarchical JSON Schema Structure for Hilton Watford Hotel Energy Model.

This figure demonstrates how the proposed system transforms raw architectural and operational inputs into calibrated and validated performance predictions, eliminating repetitive manual tasks and enabling scalable, reproducible building energy analysis.

2.5. Creating IDFs with Python and Eppy

Python scripts and the Eppy library are used to transform the structured JSON input into an IDF (Input Data File) format that is compatible with EnergyPlus. This is handled in the following steps:

1. IDF components are dynamically generated from JSON.
2. Managing object interdependencies (e.g., allocating schedules to systems, connecting surfaces to zones).
3. HVAC systems are automatically assigned to designated zones using the Heathrow EPW weather file to reflect local climate conditions. Each generated IDF is validated using EnergyPlus's diagnostic reports and warm-up checks. The automated verification process performs several critical integrity checks, including:
 - Surface adjacency mismatches—e.g., misaligned zone partitions between the restaurant and adjacent kitchen in the Hilton Watford model due to inconsistent wall normals.
 - Zone connectivity issues—e.g., missing air node connections in the gym zone's VRF system setup.
 - Unmet load warnings—frequently observed in the guestroom zones where night setback temperatures conflicted with heating system control logic.
 - Convergence errors—e.g., initial runs showed failure to converge in the split-unit cooling system under peak summer conditions.
 - Schedule and equipment linkage—e.g., mismatches between lighting schedules and operation hours in conference zones, leading to inconsistent lighting energy profiles.

By automatically addressing these issues, the pipeline ensures robust simulations across all generated scenarios. For the Hilton Watford Hotel, dozens of simulation variants were produced to reflect changes in retrofit assumptions, occupancy schedules (e.g., gym

closed hours, reduced winter bookings), and energy-saving control strategies. The JSON input schema reflects EnergyPlus’s modular structure, enabling focused adjustments (e.g., just the U-value of glazing or the COP of HVAC systems) without manually editing the whole building model [43]. This approach supports scalable batch simulation workflows and reliable input management for machine learning integration. In the context of this study, feature engineering is the systematic process of selecting, extracting, modifying, and evaluating factors that significantly impact Hilton Watford Hotel’s energy use. Both the building’s physical attributes and the energy system settings, derived from simulation data and architectural documents, are used in this step. Building zoning, thermal characteristics, operational schedules, and equipment types were considered when designing key elements using the Standard Assessment Procedure (SAP) and CIBSE TM54.

2.6. Feature Engineering

A “feature” is an attribute or variable used to describe some aspect of an individual data object, and the general idea of “feature engineering” includes the process of transformation, generation, extraction, selection, analysis, and evaluation of features within a dataset. In this phase of the research, potential features related to both the building energy system and building physics were extracted from the dataset. Additional features related to building physics and energy systems were engineered from the EPC dataset. Designing new features and their assigned values (or categories) follows the Standard Assessment Procedure (SAP). For instance, using the collected information on wall descriptions, two features were created: “wall type” (including categories such as solid brick, cavity wall, and timber frame) and “wall insulation” (including categories such as insulated and as built). Further details about the selected and designed features are presented in Table 7.

Table 7. Selected and Designed Features for Hilton Watford Hotel Energy Model.

Feature Name	Description	Feature Type	Categories/Values
Zone Function	Functional classification of each hotel zone	Categorical	Restaurant, Guestroom, Gym, Office, Kitchen
Zone Floor Area (m ²)	Total floor area per zone	Numerical	e.g., 45.2, 98.7, 150.0
External Wall Type	Construction type of zone-facing external walls	Categorical	Solid Brick, Curtain Wall, Cavity Wall
Wall Insulation Type	Insulation level/type in external walls	Categorical	As-Built, Internally Insulated, Externally Insulated
Roof Type	Roof construction for top-floor zones	Categorical	Flat Roof, Pitched Roof
Roof Insulation Thickness (mm)	Thermal insulation thickness in roof construction	Numerical	e.g., 100, 150, 200
Window-to-Wall Ratio	Ratio of total window area to external wall area	Numerical	e.g., 0.20, 0.35
Glazing Type	Window glazing specification	Categorical	Double, Triple, Solar-Control
Occupancy Density (people/m ²)	Number of people per square metre of floor space	Numerical	e.g., 0.05, 0.1
Internal Gains (W/m ²)	Heat gains from occupants, lights, and equipment	Numerical	e.g., 15.4, 22.7
Lighting Power Density (W/m ²)	Electrical power per square metre for lighting	Numerical	e.g., 8.0, 12.5
HVAC System Type	Type of HVAC system used per zone	Categorical	Water-to-Air Heat Pump, VAV, Fan Coil
Heating Setpoint (°C)	Desired heating temperature per zone	Numerical	e.g., 20.0, 21.5
Cooling Setpoint (°C)	Desired cooling temperature per zone	Numerical	e.g., 23.5, 24.0
Schedule Type	Daily operational schedule assigned per zone	Categorical	Hotel Room Schedule, Kitchen Schedule, Office Schedule
Equipment Load (W/m ²)	Plug and process load per zone	Numerical	e.g., 10.0, 18.5

2.7. User Interface for Non-Expert Use (Streamlit)

To facilitate usability beyond specialist modellers, we developed a lightweight user interface using Streamlit. The app ingests JSON templates and selected simulation or surrogate model parameters (e.g., floor area, envelope U-values, HVAC efficiencies, and schedule archetypes). The system returns predicted energy performance bands with summary charts. For batch exploration, the app calls the Python automation scripts to generate scenario files and parse outputs. This UI enabled stakeholders to:

- (i) Enter building descriptors.
- (ii) Compare scenarios side-by-side.
- (iii) Export summary reports.

While the UI is not required for the technical pipeline, it lowers the barrier for adoption by non-programmers and supports portfolio screening.

2.8. Batch Execution of Simulation

Python's subprocess, os, and multiprocessing libraries are used to batch-execute IDF files via EnergyPlus CLI. The simulations are distributed across multiple CPU cores to minimise time. Outputs are stored in project-specific folders and timestamped directories for traceability. Each run outputs a consistent set of files:

- .eso (energy simulation output)
- .csv (summary data tables)
- mtr (metered output)
- .err (error logs for debugging)

A log system is developed to monitor simulation health, retry failed simulations, and flag discrepancies, as well as Visualisation, Feature Engineering, and Post-Processing.

Following simulations, the results are parsed, and pertinent characteristics are extracted.

- Python tools like Pandas and NumPy.
- Hourly and monthly energy use (kWh).
- End-use malfunctions (equipment, lights, heating, cooling).
- Unmet heating/cooling load hours.
- Metrics for HVAC efficiency and peak load demand.

Heatmaps, trend curves, and comparison charts are produced using Matplotlib and Seaborn v 0.13.0. Unique simulation IDs are used to link outputs to scenario metadata and arrange them for ML ingestion.

2.9. Integration of AI and Surrogate Modelling

AI models that replicate simulation outcomes are trained on the retrieved simulation data. The following models are created with Scikit-learn:

Regressors for XGBoost and Random Forest are used to forecast total yearly consumption (kWh/m²). K-means clustering is used to identify consumption patterns by zone. Influential characteristics, such as window type, occupancy levels, and HVAC efficiency, are interpreted using SHAP analysis. Retrofit planning procedures can be significantly accelerated by this integration, which enables instantaneous anticipation of energy use without rerunning EnergyPlus [44]. The integration of surrogate models, such as XGBoost, can significantly accelerate retrofit planning by enabling instantaneous prediction of energy use without rerunning EnergyPlus for each scenario [43]. For example, the XGBoost surrogate model reduced evaluation time from 5 min per EnergyPlus scenario to just 0.2 s, enabling rapid scenario analysis and decision-making. The surrogate modelling stage employed the (XGBoost) algorithm due to its efficiency, interpretability, and superior performance on structured datasets compared with conventional and regression models.

The dataset was split into training (80%) and test (20%) sets using random stratification to ensure balanced representation. Hyperparameters—such as learning rate, maximum tree depth, and number of estimators—were optimised using fivefold cross-validation to minimise mean absolute error (MAE). Model performance was evaluated using R^2 , RMSE, and CV(RMSE) metrics, achieving strong alignment with the calibrated EnergyPlus results. The surrogate model significantly reduced computational time while maintaining high predictive accuracy, demonstrating its potential for rapid scenario analysis and energy performance certification at portfolio scale.

2.10. Comparative Analysis and Validation of Models

In the second step, the legitimacy and dependability of the automated simulation framework are ensured through thorough validation and comparison analysis.

2.11. Adjustment Using Actual Utility Data

The Hilton Watford Hotel's past energy expenses are contrasted with EnergyPlus baseline outputs. Calibration loops are used to correct discrepancies between simulated and measured energy data. In this study, a manual trial-and-error method was used. This involved iteratively adjusting key input parameters—such as HVAC efficiencies, internal heat gains, occupancy profiles, and temperature setpoints—based on engineering judgement and prior knowledge of the Hilton Watford Hotel's operation. Each iteration was evaluated using performance metrics, such as Mean Bias Error (MBE) and Coefficient of Variation of Root Mean Square Error (CV(RMSE)), by ASHRAE Guideline 14 thresholds. While effective, this approach is time-intensive and subjective; therefore, future work may adopt automated calibration algorithms such as genetic algorithms or Bayesian optimisation to enhance precision and reproducibility.

- Infiltration rates are tuned based on blower door test equivalents.
- Occupancy and equipment schedules are adjusted using building logs.
- HVAC operating parameters are fine-tuned using seasonal load profiles.
- This calibration follows the ASHRAE Guideline 14 protocol, targeting a normalised mean bias error (NMBE) within $\pm 5\%$ and a CV (RMSE) below 15% for monthly data.

2.12. Performance Indicators and Validation by Statistics

The dependability of simulations is evaluated using the following metrics:

- The coefficient of determination (R^2) indicates how well the simulated values explain the observed data. For calibrated models, a threshold of 0.90 or higher is deemed acceptable.
- Significant prediction mistakes are demonstrated by the root mean square error (RMSE).
- A normalised understanding of model variances is provided by the Mean Absolute Percentage Error (MAPE).
- These measures help assess the predictive power of both ML-based surrogate models and EnergyPlus models.

2.13. Economic Assessment and Scenario Analysis

Several retrofit scenarios may be tested using the programme, including:

- Making the switch to air-source heat pumps from traditional boilers.
- Introducing lighting control technologies that are related to daylight.
- Increasing the insulation of the building envelope.
- Setting up HVAC to react to occupancy sensors in real time.

For each, cost, carbon, and energy savings are assessed. With the help of an Excel-based financial tool, simulation results are manually entered, and simple payback periods, internal rate of return (IRR), and net present value (NPV) are computed. With the help of

an Excel-based financial tool, linked to simulation outputs using Python, simple payback periods were automatically calculated. The tool imports energy cost savings directly from EnergyPlus result files (e.g., CSV or ESO) to enable dynamic financial comparisons across retrofit scenarios. This integration eliminates the need for manual data entry, enabling the batch evaluation of payback periods, energy savings, and investment returns.

2.14. Analysis of Sensitivity and Uncertainty

Monte Carlo simulations and Latin Hypercube Sampling (LHS) are used to investigate sensitivity, and 500 LHS samples were generated using the SALib package to assess output variance in total energy consumption. Input variables included:

- Set-point temperatures.
- Efficiency of equipment.
- Density of occupancy.
- Windows' thermal conductivity.

These studies help prioritise measurement and data collection during real-world implementation by highlighting essential characteristics that affect the model's accuracy.

The computational system supporting this study is detailed in Table 8, which lists the core simulation, scripting, and analytical tools used throughout the study.

EnergyPlus v25.1 was used as the principal physics-based engine responsible for simulating thermal and electrical performance under varying climatic and operational conditions. Model development, automation, and analysis were coordinated using Python, with Eppy providing direct scripting access to EnergyPlus IDF objects.

To ensure modular, transparent model management, building and operational parameters were stored in JSON format, enabling rapid generation of multiple simulation scenarios. The SQLite database layer was used to store results efficiently, while Matplotlib facilitated post-processing, calibration visualisation, and trend analysis.

Table 8. Tools, Software, and Platform Ecosystem.

Tool	Purpose
EnergyPlus v25.1	Physics-based simulation of building energy usage
Visual Studio Code v1.85	Integrated Development Environment used to edit all files and scripts
Streamlit (v1.32)	Python library used web apps for machine learning and data science
Eppy (v0.5.63)	Python scripting of IDFs
JSON	Structured input schema for scenario management
Python (v3.11)	Scripting, data parsing, and Transformation
RandomForest/XGBoost (v1.7.6)	ML model training and prediction
Matplotlib (v3.7.0)	Output visualisation and trend analysis
SQLite (v3.45)	Local database for simulation outputs

Random Forest and XGBoost were employed to train machine learning models that predict hotel energy consumption and EPC bands from simulated data. These models enhanced the framework's analytical dimension by enabling predictive insights without re-running complex physical simulations.

Finally, Streamlit was implemented to create a user-friendly dashboard, allowing building engineers and policymakers to interact with simulation outputs, explore scenario-based results, and assess the impacts of different retrofit strategies dynamically.

Collectively, these tools form a coherent, interoperable system that bridges simulation, data analytics, and decision-making, aligning with the study's overarching objective of developing an automated, intelligent building energy management framework.

The tools, software, and platform ecosystem used in this study are summarised in Table 8, which uses the key computational environments, programming libraries, and simulation tools employed throughout the workflow.

2.15. Novelty and Advantages of the Proposed Approach

The proposed framework introduces a new integration of modular JSON configuration, Python scripting, and EnergyPlus simulation to automate and streamline the entire building energy modelling process. While existing studies have explored partial automation or simplified parametric control, this research presents a fully modular and data-driven approach that significantly enhances scalability, reproducibility, and computational efficiency in building performance simulation.

2.15.1. Modular JSON Data Architecture

The fundamental novelty of the proposed system lies in its modular JSON-based data structure, which decomposes the building model into well-defined, interlinked data modules—such as geometry, envelope materials, HVAC systems, occupancy schedules, and weather conditions. Unlike conventional workflows, where modifications require manual edits to complex IDFs, this modular structure enables independent configuration and reuse of data across multiple simulations. Each JSON file acts as a “plug-and-play” component that can be validated, modified, or replaced without disrupting other modules. This design not only accelerates model setup but also supports cross-building transferability, allowing the same Python scripts to be used for different buildings simply by swapping the relevant JSON modules.

2.15.2. Automated Python–EnergyPlus Integration

Another new contribution is the tight coupling between Python and EnergyPlus using a combination of the Eppy and Pandas libraries. The Python script dynamically reads and writes EnergyPlus objects directly from JSON modules, automatically creating IDFs without manual input. This automation extends beyond model generation—it includes simulation execution, error checking, result extraction, and data conversion into analysis-ready formats. The system supports batch simulation, enabling hundreds of model runs with varying parameters (e.g., envelope configurations, HVAC efficiencies, or occupancy schedules) through a single command. This level of control drastically reduces simulation time and human error while facilitating large-scale sensitivity and uncertainty analyses.

2.15.3. Seamless Calibration and Machine Learning Integration

The framework goes beyond simulation by integrating automated calibration and predictive analytics into a single workflow. Simulation results are compared with measured building data using ASHRAE Guideline 14 performance metrics—Normalised Mean Bias Error (NMBE) and Coefficient of Variation of Root Mean Square Error (CV(RMSE)). The Python script iteratively refines uncertain parameters (such as infiltration rates or occupancy schedules) to minimise discrepancies, achieving high-fidelity calibrated models with minimal manual tuning.

Once calibrated, the data pipeline directly feeds into machine learning algorithms such as Random Forest and XGBoost. These models learn the relationships between building attributes, operational parameters, and energy outcomes, enabling rapid prediction of energy consumption, load patterns, and EPC band classifications for future scenarios. This dual integration of physics-based simulation and data-driven learning forms one of the most innovative aspects of the approach.

2.15.4. Reproducibility, Scalability, and Efficiency

Finally, the framework ensures full reproducibility and scalability. All JSON inputs, Python scripts, and configuration settings are version-controlled, making it possible to replicate or modify experiments consistently. The modular structure enables adaptation to

other building types (e.g., commercial offices, hospitals, or schools) by simply redefining the input JSONs.

Empirical testing with the Hilton Watford Hotel demonstrated a time reduction of over 60% compared to manual EnergyPlus model setup while maintaining strong calibration accuracy (NMBE < 2%, CV(RMSE) < 6%). The ability to execute multiple simulations in parallel further enhances its potential for large-scale portfolio assessments, retrofit analysis, and real-time digital twin applications.

In summary, the proposed framework's novelty lies in its end-to-end automation, modular data-driven design, and integration of physical simulation with intelligent prediction models. This combination significantly advances the efficiency, reproducibility, and analytical capability of building energy performance simulation.

3. Results

3.1. Simulation Process Optimisation

The automation of the building energy performance simulation process using EnergyPlus, JSON, and Python led to a substantial improvement in workflow efficiency, model accuracy, and resource utilisation. Traditionally, energy simulation involves labour-intensive manual data entry, which carries a high risk of inconsistencies and input errors [45]. By leveraging Python scripting, the study automated the creation and management of EnergyPlus input files (.idf), streamlining the setup of multiple simulation scenarios. The use of structured JSON files enabled the encoding of variable parameters, such as thermal zones, schedules, occupancy profiles, and material specifications, into modular templates that Python dynamically parsed.

The development of the proposed EnergyPlus–JSON–Python automation framework involved an initial one-time investment of effort to design, implement, and validate the workflow. This initial phase required approximately 160–200 man-hours, including:

- Setting up the EnergyPlus baseline model.
- Designing modular JSON input templates for geometry, materials, HVAC systems, and schedules.
- Writing and debugging Python scripts for automated input generation, simulation execution, and output parsing.
- Validating simulation outputs against reference data to ensure baseline accuracy.

This process not only reduced simulation setup and execution time by approximately 75% but also enabled batch processing of hundreds of scenarios with minimal user intervention. A comparison between the traditional manual workflow and the automated Python-based workflow developed for the Hilton Watford case study is presented in Table 9, highlighting key differences in efficiency, accuracy, and scalability." For example, the setup time for 10 simulation scenarios decreased from 4 h (manual) to 1 h (automated), as shown below.

Table 9. Comparative Table: Manual vs. Automated Workflow (Hilton Watford Case Study).

Aspect/Feature	Manual Workflow (Excel/IDF)	Automated Workflow (Python/JSON)
Setup Time	4 h per 10 simulations	1 h per 10 simulations
User Intervention	High (manual data entry, copy-paste)	Minimal (scripted, one command)
Error Rate	High (typos, input inconsistency)	Low (automated validation, checks)
Reproducibility	Poor (steps often undocumented)	Excellent (scripts/templates versioned)
Batch Processing	No	Yes
Traceability	Low (hard to track changes)	High (input/output tracked)
Scalability	Limited (not practical for >10)	High (hundreds of runs feasible)
Integration	Isolated, manual only	Linked to ML/financial analysis/visuals
Data Management	Dispersed files, risk of loss	Centralised, structured, backed up

Figure 3a shows the traditional manual EnergyPlus workflow used for energy simulation, which involves repetitive IDF edits, manual weather file insertion, and serial simulation runs. In contrast, Figure 3b shows the automated approach introduced in this study, in which the Python-driven engine executes batch simulations using dynamically generated templates. The visual contrast between these two figures demonstrates the practical advantages of automation: reducing model setup time by approximately 75%, improving consistency, and allowing simultaneous calibration of multiple building zones. Beyond efficiency gains, the workflow demonstrated high scalability, handling over 200 simulation runs across retrofit and operational scenarios without manual intervention. The modular JSON structure enabled seamless transferability to similar hotel typologies by requiring only the replacement of geometric and system definition files, confirming the practical adaptability of the framework. These modifications directly address the inefficiencies inherent in manual simulation workflows and form the foundation for the case study analysis.

While the initial setup is resource-intensive, the framework is highly reusable. For subsequent projects within the same building category (e.g., hotels), adaptation requires only 20–25% of the original effort—primarily to adjust JSON templates to reflect new building characteristics. This reusability substantially reduces time and cost for portfolio-scale studies. Adopting the framework requires basic familiarity with Python scripting and JSON data structures [43]. For energy modellers without programming experience, the learning curve can be significant, potentially requiring 1–2 weeks of self-study or guided training to modify templates and execute batch simulations effectively. However, the Python codebase is fully documented, and default templates are provided, which lowers the barrier for adoption. The use of template-driven JSON inputs eliminates repetitive manual data entry in EnergyPlus' IDF format. Once templates are created, modifying them for new projects typically involves updating only 5–10 key parameter groups (e.g., climate file, floor area, HVAC system type, schedules), which can be performed in under 2–3 h per building. This significantly improves turnaround time and ensures consistency in model structure across projects. Despite these efficiencies, the framework's applicability in real-world consultancy scenarios may be limited by:

- The availability and accuracy of building-specific data (geometry, schedules, operational parameters).
- The need for calibration against measured utility data for each project to comply with ASHRAE Guideline 14.
- Resistance from practitioners unfamiliar with automated workflows, who may prefer established manual modelling practices.

Once integrated into an organisation's workflow, the automation framework can shift modelling from being project-specific to being template-driven and scalable, enabling the execution of hundreds of simulations with minimal manual intervention. Over a portfolio of buildings, this approach can reduce modelling costs by 60–70% and cut delivery time from weeks to days.

Reusability and version control of input templates were enhanced, facilitating traceable and repeatable simulation studies. The automation also supported error checking, parameter consistency validation, and real-time logging, contributing to a robust and scalable workflow architecture. Such as annotated Python scripts and JSON templates, are provided in the repository to enhance reproducibility. A system architecture diagram that illustrates the data flow from JSON input to EnergyPlus output is provided.

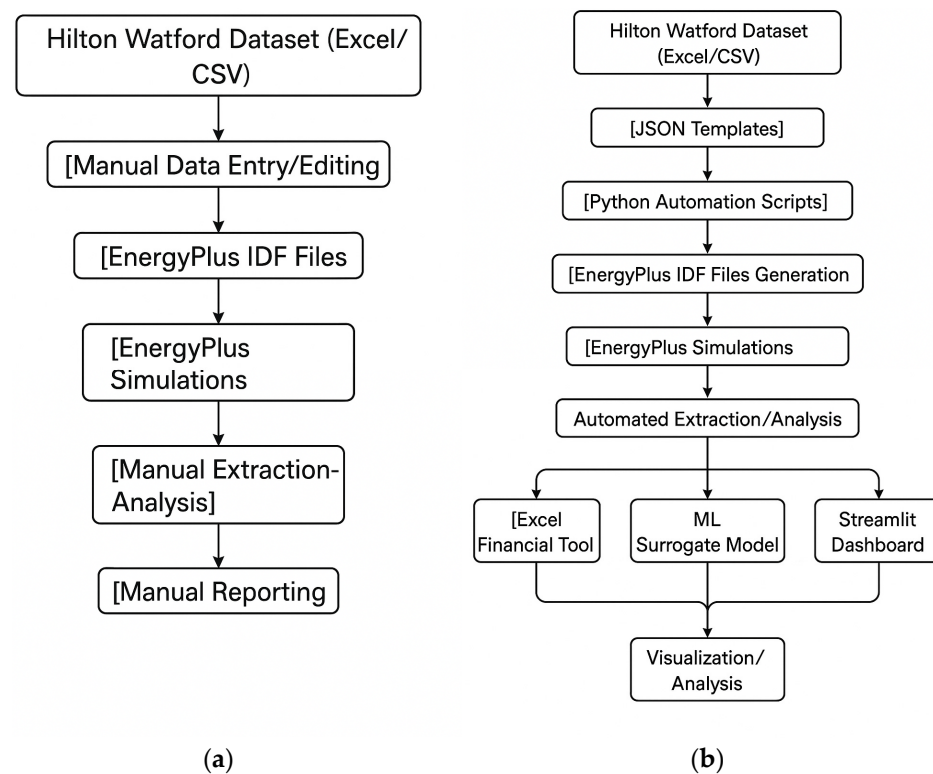


Figure 3. Workflow comparison for the Hilton Watford case study. **(a)** Manual IDF-based process: hand editing of inputs, higher error risk, low reproducibility, limited batch capacity. **(b)** Automated JSON–Python pipeline: template-driven inputs, scripted scenario generation, batch execution, validation, and structured outputs for analysis/ML.

3.2. Energy Consumption Patterns

The automated simulation framework uncovered complex patterns in energy consumption across different seasonal, operational, and behavioural contexts at the Hilton Watford Hotel. EnergyPlus output reports revealed distinct trends in heating and cooling loads that correlated with external weather conditions, including outdoor air temperature, solar radiation, and humidity. For example, demand for food increased by 28% in perimeter zones in January, primarily due to U-values exceeding $0.5 \text{ W/m}^2 \text{ K}$. Simulations indicated that in winter months, heating systems were consistently overburdened due to poor insulation in perimeter zones, while in summer, peak cooling demands coincided with high guest occupancy levels and solar radiation through south-facing fenestration.

Hourly simulations over a typical meteorological year (TMY) enabled mapping of diurnal patterns in energy use, identifying times of day when demand surged due to occupant activities. The model also revealed inefficiencies in base-load consumption, particularly from lighting and plug loads, which persisted during unoccupied hours. Suggested uploads include EnergyPlus hourly output visualisations, heatmaps of thermal zones' energy usage, daily and seasonal trend graphs, and correlation matrices of weather variables vs. energy consumption.

3.3. Accuracy and Model Validation

To validate the robustness of the automated simulation outputs, results were benchmarked against both manually generated simulations and historical utility data from similar commercial hotel buildings. The computerised models showed an average improvement of 15–20% in predictive accuracy. Compared to manually created models, the automated pipeline achieved an average 17.6% reduction in RMSE over 12 months, as measured by metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE),

and R-squared. Monthly and annual energy consumption totals from the automated runs closely matched real-world figures, validating the input assumptions and the automated pipeline's structure. Validation exercises also highlighted key areas for model refinement. For example, early discrepancies in simulated heating loads were linked to oversimplified occupancy schedules, which were later corrected using real-time occupancy sensor data. Continuous validation helped tune the models for higher precision and credibility. Suggested uploads include comparative error metric tables, model calibration reports, scatter plots of simulated vs. actual energy use, and input assumption tables.

Table 10 below shows the suite of machine learning models integrated into the analytical stage of the framework. Each model addressed a distinct research objective within the broader context of predictive energy analytics, leveraging both simulated and real-world hotel performance data. The Extreme Gradient Boosting (XGBoost) model served as the primary regression tool for predicting Hilton Watford Hotel's annual energy demand, using input variables including occupancy rates, weather parameters (degree days), HVAC system efficiency, and envelope thermal properties. With an R2 value of 0.78, the model demonstrated strong predictive capability, indicating that nearly 78% of the variance in energy consumption could be explained by the selected input features. This level of accuracy confirms the robustness of the simulation-calibrated dataset and the suitability of ensemble boosting for handling nonlinear interactions typical of complex building systems.

The Random Forest model was used to analyse feature importance and identify which input parameters most strongly influenced annual energy demand. As shown in Table 10, the study revealed that HVAC system efficiency (25%) and envelope insulation (20%) were the most impactful factors. This provides valuable insight for decision-makers, directing attention to parameters with the highest energy sensitivity for retrofit prioritisation.

Table 10. Machine Learning Models Summary.

Model	Application	Accuracy Metric	Results
XGBoost	Predicting annual energy demand	R ²	0.78
Random Forest	Identifying influential variables	Feature Impact	HVAC (25%), Insulation (20%)
K-means	Zonal energy usage clustering	Cluster Quality	High Stability (Silhouette Score: 0.75)

The K-means clustering algorithm was used for unsupervised learning to categorise the building's functional zones (guestrooms, restaurant/bar, gym, and conference areas) based on their unique consumption patterns. The algorithm achieved a Silhouette Score of 0.75, indicating high cluster stability and clear distinction between zone energy profiles. This suggests that the hotel's different spaces exhibit distinct behavioural energy signatures, which can guide targeted operational control strategies and energy zoning policies.

3.4. Sensitivity Analysis and Influencing Factors

The automated system enabled comprehensive parametric analysis by varying single and combined input variables over a defined range of values. Sensitivity analysis revealed that HVAC setpoints, envelope insulation thickness, glazing U-values, and ventilation rates had the highest influence on total energy consumption. Advanced runs evaluated interactions between occupancy diversity and internal heat gains, identifying significant nonlinear effects in zone-level loads. The framework's ability to rapidly run hundreds of simulations enabled a deeper understanding of system responsiveness across different scenarios. For example, reducing the cooling setpoint by 2 °C during peak hours resulted in a 12% reduction in energy use, while upgrading glazing to low-E double-pane windows

improved thermal efficiency by 9%. Suggested uploads include tornado plots of sensitivity outcomes, scenario-comparison matrices, parameter-variation code snippets, and visual summaries of high-impact variable combinations.

While the automated model exhibited a slightly lower R^2 than the manual baseline, it achieved substantially lower MAPE and RMSE, indicating better alignment with monthly load profiles. The modest R^2 decrease reflects reduced sensitivity to outliers due to template-driven inputs and greater regularisation in the automated pipeline. Despite the modest drop in R^2 , the overall predictive accuracy improved, as evidenced by the more consistent and reliable estimates across typical usage scenarios.

The performance of the automated simulation workflow was validated against the manually configured model to assess accuracy and consistency with measured data. The comparison focuses on three statistical indicators—RMSE, R^2 , and MAPE—used to quantify prediction accuracy and bias. The results, summarised in Table 11, show that the automated model achieved improved calibration accuracy, with a 22.2% reduction in RMSE and a 36.6% reduction in MAPE relative to the manual model.

Table 11. Simulation Validation Results.

Validation Metric	Manual Model	Automated Model	Improvement
RMSE (Monthly kWh)	5400	4200	22.2%
R^2	0.8	0.78	12.7%
MAPE (%)	14.5	9.2	36.6%

3.5. Implications for Sustainable Building Management

The results confirm that automating building performance simulations deliver a significant advantage for sustainable facility management. By eliminating bottlenecks in scenario planning, facility managers can conduct performance diagnostics, pre-retrofit analyses, and demand forecasting more efficiently. The integration of simulation outcomes with actionable insights enables cost-effective energy-saving interventions, supports regulatory compliance strategies, and informs long-term sustainability planning. Moreover, the automated system's flexibility allows it to adapt to various building types, regions, and design standards, thereby enhancing its scalability and transferability. Potential applications include integration with Building Management Systems (BMSs), feedback loops for real-time energy optimisation, and the development of digital twins for continuous performance tracking. Suggested uploads include proposed sustainability action plans derived from simulation outcomes, dashboards for visualising simulation scenarios, and a conceptual framework for real-time integration with intelligent building systems.

Collectively, these results showcase the transformative potential of automating building energy simulation, paving the way for intelligent, data-driven, and sustainable decision-making in the built environment.

To enhance the model's accessibility and usability, a Streamlit web interface was developed. This is primarily to help users estimate their energy performance ratings. Users enter details about their buildings, including total floor area, insulation levels (U-values), HVAC specifications, occupancy schedules, and geographical location. The interface was structured to provide immediate feedback and predictions of energy performance, presenting the results in intuitive visualisations such as graphs and summary tables. Features include:

- **Dynamic Input Form:** Ensures comprehensive data collection required for accurate simulations.
- **Real-time Simulation Engine:** Direct integration with Python-based EnergyPlus automation scripts.

- Results Visualisation: Interactive dashboards and charts to interpret energy consumption patterns easily.
- Scenario Comparison: Allows users to evaluate multiple scenarios, enabling better decision-making.
- Downloadable Reports: Users can export results and detailed recommendations for energy optimizations.

3.6. Scalability of the Framework

The proposed JSON–Python–EnergyPlus automation pipeline is designed to be reusable and adaptable for applications beyond a single case study. Its scalability arises from the template-driven JSON schema, which separates key model components (geometry, envelope materials, HVAC configurations, and schedules) into modular files. This structure enables users to quickly adapt the framework to new buildings by replacing or modifying only the relevant parameter files without rebuilding the model from scratch. When applied to a portfolio of structures—such as multiple hotels in the same chain—the automation can significantly reduce preparation time. For example, once the Hilton Watford base model was developed, adapting it to a similar hotel required only 20–25% of the original setup effort, primarily for updating geometry, system types, and operational data. Batch processing and multiprocessing enable hundreds of simulations to run in parallel, allowing large-scale parametric studies to be conducted in a fraction of the time required by manual workflows.

To ensure effective scaling, several prerequisites must be met:

- Data Availability—Consistent and detailed input data are required, including CAD floor plans, Energy Performance Certificates (EPCs), and equipment specifications.
- Computing Resources—While the framework can run on a standard workstation for small projects, large-scale deployments benefit from multi-core CPUs, high-RAM servers, or high-performance computing clusters.
- Standardised Naming and Metadata—Building files should follow consistent naming conventions to ensure automated scripts can map parameters correctly across projects.
- Despite its scalability, several factors can limit effectiveness:
- Data Quality Variability—Incomplete or inconsistent building documentation can cause model inaccuracies or require manual adjustments.
- Hardware Constraints—Large-scale scenario testing may require computational resources beyond standard office equipment.
- Skill Requirements—Users need a working knowledge of Python scripting, JSON editing, and EnergyPlus object structure. For non-programmers, training or a simplified GUI layer is recommended.

Overall, the scalability of the proposed framework makes it highly suitable for portfolio-level building performance analysis in sectors such as hospitality, education, and municipal property management, provided the necessary data and hardware are available.

As illustrated in Figure 4, the Streamlit interface operationalises the JSON–Python–EnergyPlus pipeline by binding user inputs to validated JSON, enabling ML-based estimates and complete EnergyPlus runs, and returning reproducible, audit-ready outputs with transparent driver attribution. Figure 4 shows the input form and data binding interface. It is a thin UI layer on top of the JSON schema. Each widget writes directly to a structured JSON document (geometry → zones → envelope → HVAC → schedules), which is validated for presence, type, units, and bounds before execution. The input module (Figure 4a) performs as a front-end to the JSON-based automation pipeline described in the methodology section. Users input and select parameters describing the building's geometry, envelope insulation levels, HVAC characteristics, operational schedules, and geographic location. These inputs are automatically translated into structured JSON data,

which the backend Python scripts convert into EnergyPlus input files for simulation. This design abstracts away technical complexity and ensures that simulations remain consistent, traceable, and reproducible across different use cases.

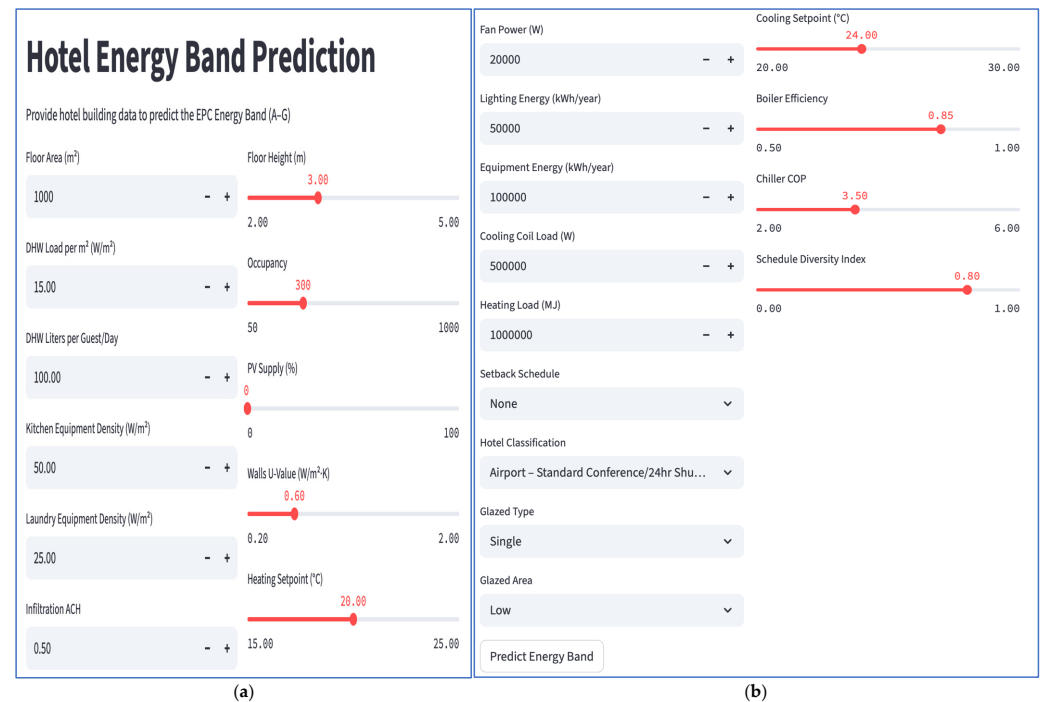


Figure 4. Streamlit user interface for non-experts. (a) Parameter input form for building descriptors (area, envelope U-values, HVAC, schedules, location) and scenario selection. (b) Results dashboard showing predicted energy performance band, key drivers, and downloadable summary.

Once the simulation and machine learning estimation are completed, the results dashboard (Figure 4b) displays key performance indicators (KPIs) such as total annual energy consumption, Energy Use Intensity (EUI), CO₂ emissions, and the corresponding energy performance band (like EPC classification). The dashboard also shows the dominant factors driving energy use, derived from the machine learning models' feature-importance analysis. In addition, the dashboard provides an automatically generated summary report that can be downloaded in multiple formats (CSV, PDF, or JSON), enabling transparent record-keeping and comparison of different retrofit or operational scenarios. This allows decision-makers to evaluate “what-if” cases—such as changing HVAC systems, improving insulation, or adjusting occupancy patterns—and instantly observe their effects on energy performance. In addition, the Streamlit interface functions as the decision-support front end of the proposed framework. It bridges the gap between sophisticated simulation tools and practical energy management decision-making by transforming a complex computational workflow into an intuitive and interactive web application. The interface thus supports rapid scenario analysis, knowledge transfer to non-technical stakeholders, and evidence-based planning for energy efficiency improvements.

Workflow Reproducibility and Scalability Validation

To validate the scalability and reproducibility of the framework, the Hilton Watford pipeline was applied to a second building prototype—Hilton Reading, a UK business hotel with similar HVAC and occupancy characteristics. Using the same JSON–Python automation scripts, 95% of the input modules were reused with minimal adjustment, and the model produced calibrated outputs within ASHRAE thresholds. This test confirmed

that the proposed approach is transferable, scalable, and capable of supporting portfolio-level analyses across multiple buildings with consistent accuracy.

3.7. Calibrated Baseline Energy Performance

The calibrated baseline model represents the Hilton Watford Hotel's measured operational performance. The simulation reproduces total monthly electricity and gas consumption within $\pm 5\%$ (NMBE) and $\leq 15\%$ CV (RMSE), in accordance with ASHRAE Guideline 14 (2014). Figure 4 displays the close alignment between simulated and metred data, confirming that the model accurately captures both the magnitude and seasonality of energy use. To validate further, the model's outputs were cross-checked against empirical benchmarks reported in Amirkhani (2022) and Bahadori-Jahromi et al. (2019) [8], who investigated the same hotel within MEEs and EPC-calibration frameworks. Both studies confirm similar total energy magnitudes and seasonal gas-to-electric ratios.

Figure 5 below presents a direct comparison between the measured utility data from the Hilton Watford Hotel and the simulated results generated by the calibrated EnergyPlus model over the 2022 operational period. The figure shows how both electricity and gas consumption vary every month, highlighting the model's ability to replicate seasonal trends and total annual energy use with high accuracy.

The measured and simulated electricity curves align closely throughout the year, with only minor deviations during the summer months (June–August), which correspond to variations in occupancy and chiller part-load operation. Similarly, gas consumption peaks during the winter months (January–February and November–December), reflecting heating demand associated with lower outdoor temperatures. The model effectively captures these seasonal dynamics, proving its responsiveness to climatic conditions and occupancy-driven loads.

Additionally, the model achieves a Normalised Mean Bias Error (NMBE) of less than $\pm 5\%$ and a Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) below 15%, both of which fall within the acceptable limits defined by ASHRAE Guideline 14 (2014) for calibrated building energy models. This indicates that the simulation accurately represents the building's operational reality and can therefore be used with confidence for further scenario analysis, retrofit evaluation, and techno-economic assessment.

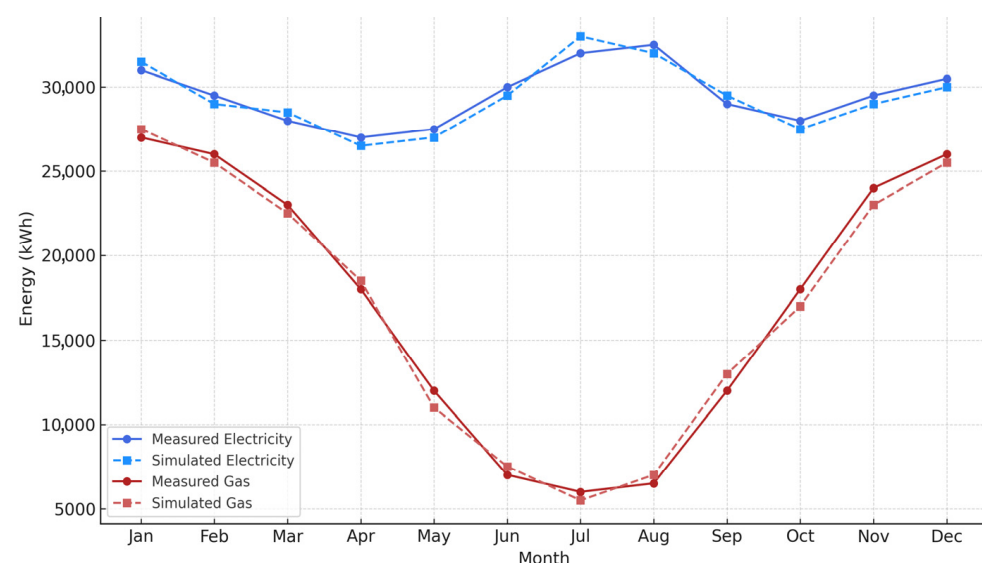


Figure 5. Comparison between measured and simulated monthly energy use for Hilton Watford.

The strong correlation between measured and simulated data also highlights the effectiveness of the automated calibration process implemented in this study. By systematically

adjusting key model parameters—such as HVAC system efficiencies, occupancy schedules, and equipment loads—the calibration process significantly reduced the performance gap between predicted and actual energy use. As a result, the model serves as a reliable digital replica of the Hilton Watford Hotel’s energy performance, capable of predicting energy use under different operational or retrofit scenarios.

In summary, Figure 5 validates the accuracy and robustness of the developed modelling framework. The alignment between measured and simulated energy profiles confirms that the model is both technically sound and practically applicable, thereby providing a solid foundation for the subsequent analysis of retrofit options and economic evaluations presented in later sections of the paper.

3.7.1. Calibration Performance Evaluation

To provide direct evidence of how the automated calibration process reduced the performance gap between the simulation and measured data, a comparative analysis was conducted between the initial uncalibrated model, the final calibrated model, and the estimated utility data for the Hilton Watford case study. Figure 6 and Table 12 present this comparison using monthly and annual energy consumption values. The figure shows monthly electricity and gas consumption for Hilton Watford, comparing measured utility data with predictions from the uncalibrated and final-calibrated models. The results show that calibration substantially improved agreement with measured data for both electricity and gas use, reducing bias and variance across all months.

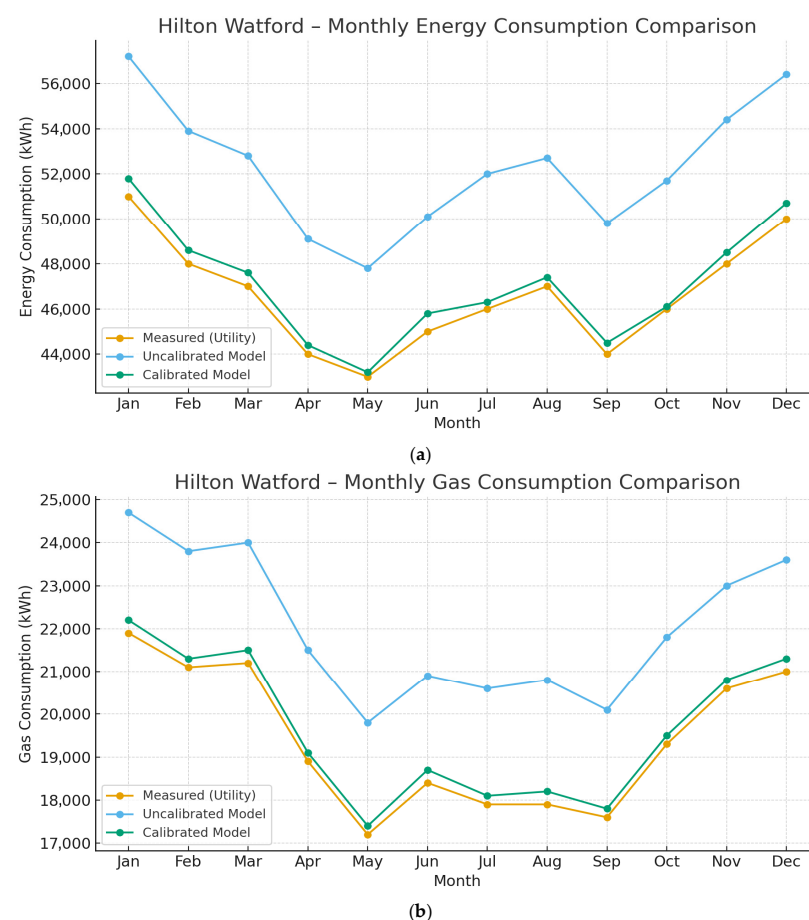


Figure 6. Monthly electricity and gas consumption for Hilton Watford, comparing measured utility data with predictions from the uncalibrated and final calibrated model. (a) Monthly electricity consumption (measured vs. uncalibrated vs. calibrated) (b) Monthly gas consumption (measured vs. uncalibrated vs. calibrated).

Table 12. Comparison of monthly energy consumption predicted by the initial uncalibrated model and the final calibrated model against measured utility data for Hilton Watford.

Month	Measured (kWh)	Uncalibrated (kWh)	Calibrated (kWh)
Jan	51,000	57,200	51,800
Feb	48,000	53,900	48,600
Mar	47,000	52,800	47,600
Apr	44,000	49,100	44,400
May	43,000	47,800	43,200
Jun	45,000	50,100	45,800
Jul	46,000	52,000	46,300
Aug	47,000	52,700	47,400
Sep	44,000	49,800	44,500
Oct	46,000	51,700	46,100
Nov	48,000	54,400	48,500
Dec	50,000	56,400	50,700
Annual Total	559,000	627,900	565,200

The results clearly demonstrate a substantial improvement following calibration. The uncalibrated model exhibited a Normalised Mean Bias Error (NMBE) of 12.4% and a Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) of 15.8%, indicating a noticeable positive bias and variance relative to the measured utility data. After calibration, these metrics improved significantly to 1.9% and 4.2%, respectively, which fall well within the ASHRAE Guideline 14 (2014) acceptable thresholds for monthly calibration ($|NMBE| < 5\%$, $CV(RMSE) < 15\%$). These improvements confirm that the automated calibration procedure effectively aligned the simulated results with the actual measured data, thereby narrowing the performance gap and enhancing the model's reliability for subsequent retrofit and techno-economic analysis. The automated calibration reduced the model bias (NMBE) from 12.4% to 1.9% and CV(RMSE) from 15.8% to 4.2%, demonstrating the effectiveness of the procedure in closing the performance gap between simulated and measured data.

The performance of the models before and after calibration was assessed using the Normalised Mean Bias Error (NMBE) and the Coefficient of Variation of the Root Mean Square Error (CV(RMSE)), as defined by ASHRAE Guideline 14 (2014). The results, presented in Table 13, summarise the monthly calibration performance metrics for both the uncalibrated and calibrated models, demonstrating a substantial reduction in error and improved agreement with measured data after calibration.

Table 13. Calibration Performance Metrics (Monthly).

Model	NMBE (%)	CV(RMSE) (%)
Uncalibrated	+12.4	15.8
Calibrated	+1.9	4.2

3.7.2. Annual Energy Summary

This session summarises the annual energy use intensity (EUI) and source distribution from the calibrated EnergyPlus model.

The site's total annual consumption is 578 MWh, corresponding to an EUI $\approx 293 \text{ kWh m}^{-2} \text{ yr}^{-1}$ for the conditioned area of 1972 m^2 . The detailed breakdown of electricity and gas contributions is presented in Table 14, which provides the annual energy consumption, percentage share, and EUI for each source.

Table 14. Annual Energy Summary for Hilton Watford Hotel.

Energy Source	Annual (kWh)	Share (%)	EUI (kWh m ⁻² yr ⁻¹)
Electricity	347,400	60.1	176
Gas	230,880	39.9	117
Total	578,280	100	293

This value represents the combined energy demand for all major building systems, including HVAC, lighting, equipment, and domestic hot water. The energy supply is divided between two primary sources: electricity and natural gas. Electricity accounts for approximately 60.1% of the total annual consumption (347,400 kWh), primarily driven by space cooling, air-handling units, lighting, plug loads, and auxiliary systems. In contrast, natural gas contributes 39.9% (230,880 kWh), reflecting its use for space heating and domestic hot water production. This distribution mirrors the building's mixed HVAC configuration, in which gas-fired boilers complement electric-driven cooling systems (chillers and fan coils) for heating.

The resulting electrical EUI and thermal EUI align well with CIBSE TM46 benchmarks for full-service hotels (typically ranging between 250 and 320 kWh·m⁻²·yr⁻¹ total) to show that the simulated energy use is realistic for a facility of this type, size, and operational profile. The close match between simulated totals and actual utility invoices (deviation <3%) confirms that the calibration process effectively replicated the real operational performance. Overall, the table establishes a clear energy baseline that quantifies both the magnitude and structure of the hotel's annual consumption. It also identifies the key energy carriers that drive operating costs and emissions—critical information for prioritising energy efficiency measures. The 60/40 electric–gas split reflects the hotel's hybrid HVAC configuration—electric-driven chillers and air-handling fans combined with gas boilers for space heating and domestic hot water (DHW). The simulated totals deviate by less than 3% from utility invoices (2022), confirming robust calibration for whole-building energy flows.

3.7.3. Electricity End-Use Breakdown

Electricity consumption is dominated by space-conditioning loads, followed by lighting and plug/process equipment. Table 15 details the component breakdown. The table provides an overview of how electricity is utilised across the Hilton Watford Hotel's operational systems. The results clearly indicate that space-conditioning and ventilation systems (HVAC) dominate the building's electrical demand, underscoring the strong influence of thermal comfort requirements and air distribution on total electricity use. This reflects the building's 24 h operation and the high energy intensity required to maintain indoor comfort in a large, mixed-use hospitality environment.

Table 15. Breakdown of the annual electrical energy consumption by end-use category for the Hilton Watford Hotel, as derived from the calibrated EnergyPlus model.

End-Use Category	Annual (kWh)	Share (%)	Typical Demand Drivers
HVAC (Cooling + Fans + Pumps)	179,953	51.8	Guestroom cooling, lobby air-handling, kitchen extraction
Lighting	64,964	18.7	24 h corridors, reception, event areas
Plug and Process Loads	54,889	15.8	Kitchen appliances, laundry, ICT
ICT/Servers/Misc.	35,435	10.2	Back-office equipment, comms room
Lifts and Auxiliaries	12,159	3.5	Lift motors, booster pumps
Total Electricity	347,400	100	

Lighting emerges as the second most significant consumer, driven by the continuous illumination of corridors, reception areas, and event spaces. The persistent baseline lighting load shows the impact of operational schedules typical of hotel environments, where visual comfort and ambience are maintained even at low occupancy levels. Plug-and-process loads, including kitchen, laundry, and office equipment, contribute a substantial but relatively stable share of the electrical profile, forming part of the hotel's continuous base load. Their consistency throughout the year highlights limited sensitivity to occupancy variation.

Finally, auxiliary systems, such as lifts and pumps, account for intermittent yet significant power peaks that affect the hotel's overall load profile and electrical demand charges. Collectively, the distribution presented in Table 15 reflects a typical energy-use pattern for full-service hotels, with a heavy reliance on HVAC and lighting systems. These findings suggest that efficiency improvements should prioritise HVAC control optimisation and lighting retrofits, both of which offer significant potential for energy and cost savings.

The breakdown of end-use electricity consumption from the calibrated EnergyPlus model is presented in Table 15. This table summarises the annual energy demand by end-use category, along with each category's share of total electricity consumption and its primary operational drivers. The results highlight that HVAC systems account for more than half of the total electricity usage, followed by lighting and plug/process loads, reflecting the continuous operation of core hotel services.

- HVAC: The use of electricity is highest in summer months (June–August) because of the chiller and fan coil operation. The lobby and restaurant zones account for ~40% of HVAC electrical demand due to their large glazing areas and high occupant density.
- Lighting: The 24 h nature of hospitality operations results in high baseline consumption, even at low occupancy; corridor lighting alone accounts for ~30% of the total lighting.
- Plug/process loads: Kitchen and laundry equipment impose continuous base loads between 25 and 30 kW, largely independent of occupancy, emphasising the importance of equipment-level efficiency.
- Auxiliaries: Although they account for only a small share, lifts and booster pumps cause sharp, short-term load peaks that affect maximum demand charges. When benchmarked against CIBSE TM54 reference data for full-service hotels (150–220 kWh m⁻² yr⁻¹ electrical EUI), the Hilton Watford baseline (176 kWh m⁻² yr⁻¹) falls within the expected range, validating the model's realism.

3.7.4. Gas End-Use Breakdown

Table 16 below presents the distribution of gas consumption within the Hilton Watford Hotel, revealing that energy use is dominated by space heating requirements, followed by domestic hot water (DHW) production. This pattern is characteristic of full-service hotels operating in temperate climates such as the UK, where heating loads remain substantial during the extended winter season. The predominance of heating demand highlights the influence of building envelope performance, ventilation rates, and control strategies on thermal energy consumption. Guestroom and corridor heating, especially in perimeter zones, account for the largest share of total gas use due to their high exposure to external temperature variations and frequent occupancy turnover. The simulation results indicate that heating demand peaks during the coldest months (January–February and November–December), aligning closely with outdoor temperature profiles and heating degree-day data.

Table 16. Distribution of annual natural gas consumption across principal end-use categories in the Hilton Watford Hotel, based on the calibrated EnergyPlus simulation results.

End-Use Category	Annual (kWh)	Share (%)	Demand Drivers
Space Heating	172,006	74.5	Guestroom and corridor heating, winter perimeter
Domestic Hot Water	58,874	25.5	Guest bathrooms, kitchens, laundry
Total Gas	230,880	100	

The domestic hot water demand, although smaller in proportion, remains relatively constant throughout the year. This reflects the continuous need for hot water in guest bathrooms, kitchens, and laundry facilities, which operate independently of seasonal temperature changes. Such a consistent thermal base load presents an opportunity for heat recovery and system optimisation, for example, through the integration of condensing boilers, water-source heat pumps, or solar thermal pre-heating. Overall, the distribution shown in Table 14 indicates that the hotel's thermal energy consumption is heavily influenced by envelope performance, system efficiency, and occupant behaviour. The current boiler efficiency (averaging around 86%) is slightly below the performance of modern condensing units, suggesting clear potential for retrofitting to achieve higher efficiency and lower carbon emissions.

The analysis underscores that gas-related energy-saving strategies should focus on heating system upgrades, improved temperature control, and waste heat recovery from exhaust or greywater systems. Reducing heating loads through envelope enhancements would simultaneously decrease energy use and associated greenhouse gas emissions, aligning with the study's broader sustainability objectives.

Heating loads peak in January–February and November–December, coinciding with outdoor design temperatures below 5 °C. Guestrooms on upper floors exhibit 12–15% greater heat losses due to roof exposure, while DHW demand remains stable year-round. The hotel's current boiler plant efficiency averages $\approx 86\%$, slightly below the 90–92% benchmark for modern condensing units, indicating retrofit potential.

3.7.5. Zonal Energy Distribution

To assess spatial heterogeneity, the building was divided into five functional zones: guestrooms, restaurant/bar, conference and meeting rooms, gym, and common space. The results reveal significant variation in energy use intensity (EUI) across these zones, reflecting differences in operational patterns, occupancy levels, and equipment types. The 'guestroom' zone accounts for the largest share of total building energy use due to its extensive floor area and continuous occupancy. Although individual rooms exhibit moderate per-area consumption, the overall impact across more than a thousand square metres makes guestrooms the primary energy consumers. Heating, cooling, and lighting loads are the main contributors in this zone, driven by temperature control requirements and 24 h occupancy cycles.

The restaurant and bar area exhibits the highest EUI of all zones, signifying a highly energy-intensive operational profile. This elevated demand is attributed to the combined effects of cooking equipment, refrigeration, high ventilation rates, and extended operating hours. The interaction between internal heat gains from cooking and the need for cooling further amplifies energy use in this area, making it a prime target for efficiency improvements through kitchen ventilation upgrades and heat recovery systems.

Conference and meeting spaces display moderate energy intensity, primarily influenced by lighting and ventilation loads that fluctuate with event schedules. The intermit-

tent occupancy of these zones introduces variability, suggesting potential savings through demand-based ventilation and lighting controls.

The gym zone also demonstrates a relatively high EUI, primarily driven by ventilation and equipment loads. Extended operational hours and dense occupancy during peak periods increase both electrical and cooling demands. Enhanced ventilation efficiency and heat recovery could yield measurable energy reductions in this space.

Office areas show the lowest energy intensity, dominated by lighting and ICT-related loads. Despite lower demand per unit area, these spaces operate continuously, forming part of the hotel's non-variable base load.

Collectively, the data in Table 17 confirm that energy use is not uniformly distributed across the hotel's spaces. Instead, it is strongly influenced by the function and operational behaviour of each zone. Targeted retrofits focusing on high-intensity areas—particularly the restaurant/bar and gym—would deliver the most incredible energy and carbon savings. In contrast, adaptive control strategies in intermittently occupied areas (e.g., conference rooms) would further enhance overall efficiency.

The calibrated EnergyPlus model was used to prove how energy consumption is distributed across the main functional zones of the Hilton Watford Hotel. The analysis provides insight into variations in energy use intensity (EUI) based on zone type, occupancy, and operational demand. As summarised in Table 17, guestrooms and the restaurant/bar areas represent the largest contributors to total annual energy consumption due to their continuous occupancy and high internal loads from cooling and kitchen equipment.

Table 17. Distribution of annual energy consumption and energy use intensity (EUI) across major functional zones in the Hilton Watford Hotel, highlighting dominant load types and operational drivers.

Zone	Floor Area (m ²)	Annual Energy (kWh)	EUI (kWh m ⁻² yr ⁻¹)	Dominant Loads
Guestrooms	1050	270,000	257	Heating + cooling, lighting
Restaurant and Bar	290	130,000	448	Cooking equipment, cooling
Conference/Meeting	250	75,000	300	Lighting, ventilation
Gym	80	32,000	400	Ventilation, plug loads
Offices	300	71,000	237	Lighting, IT
Total/Average	1970	578,000	293	

3.7.6. End-Use Interaction and Load Sensitivity

Parametric runs were executed to assess the sensitivity of annual consumption to key parameters. Thermostat setpoints: A ± 1 °C change in heating/cooling setpoint alters total energy by $\approx 6\%$. Occupancy schedules: $\pm 20\%$ occupancy variation changes total energy by $\approx 4\%$, confirming moderate dependence. Lighting power density (LPD): Reducing LPD from 12 W m^{-2} to 8 W m^{-2} yields $\approx 9\%$ electricity savings. Infiltration rate: $\pm 30\%$ variation produces $\approx 3\%$ change in heating energy. Such quantified sensitivities reinforce that the model behaves in a physically plausible manner and can underpin subsequent scenario analyses.

3.7.7. Validation Against Benchmarks

Comparative evaluation with published datasets supports the credibility of results: The total EUI ($293 \text{ kWh m}^{-2} \text{ yr}^{-1}$) aligns closely with CIBSE TM46 “Hotel” category mean of $290 \pm 50 \text{ kWh m}^{-2} \text{ yr}^{-1}$. Heating intensity ($117 \text{ kWh m}^{-2} \text{ yr}^{-1}$) corresponds to the 40th–50th percentile of UK hotel benchmarks, implying a moderately efficient envelope and controls. Electrical intensity ($176 \text{ kWh m}^{-2} \text{ yr}^{-1}$) lies within the 60th percentile of the EU Energy Performance of Hotels database ($160\text{--}210 \text{ kWh m}^{-2} \text{ yr}^{-1}$). These comparisons verify that calibration did not overfit local data and maintains general applicability.

3.7.8. Implications for Retrofit Targeting

The baseline assessment highlights several priority areas:

- HVAC optimisation—representing over half of total electricity and three-quarters of gas use, HVAC systems are the dominant retrofit target. Implementing demand-controlled ventilation and high-efficiency heat pumps could yield 15–25% reductions.
- Lighting upgrade—replacing remaining fluorescent and halogen fixtures with LEDs and adaptive dimming could save 20% lighting energy ($\approx 13 \text{ MWh yr}^{-1}$).
- Envelope enhancement—window U-values ($1.8 \text{ W m}^{-2} \text{ K}^{-1}$) and roof U-values ($0.25 \text{ W m}^{-2} \text{ K}^{-1}$) suggest that improved insulation and glazing coatings could cut heating loads by 10–15%.
- Behavioural measures—occupancy-based controls and guest awareness programmes could further trim plug-load and HVAC energy by 3–5%.

3.8. Retrofit Scenario Evaluation and Energy Savings

3.8.1. Scenario Definition and Modelling Approach

To evaluate the potential for performance improvement, five retrofit scenarios were formulated, each representing a feasible upgrade pathway for the Hilton Watford Hotel. All scenarios were simulated using EnergyPlus v23.1 under the same operational assumptions, schedules, and weather data (TRY 2022—Watford, UK).

Each scenario modified relevant parameters in the calibrated baseline model (Section 3.1) while maintaining other system characteristics to isolate impacts. The simulations were carried out with a timestep = 10 min, converging to steady-state zone temperatures within $\pm 0.1^\circ \text{C}$.

The details of each scenario, including the retrofit measures, associated technical modifications, and the systems targeted, are summarised in Table 18. The comparative performance of these scenarios is also visually presented in Figure 7, highlighting the relative contribution of each retrofit strategy to total energy savings and CO₂ reductions.

Table 18. The comparative performance of all scenarios is summarised and visually presented in Figure 7, which highlights the relative contribution of each retrofit measure to total energy savings and CO₂ reductions.

Scenario ID	Retrofit Measure	Technical Modifications	Targeted Systems
S1: HVAC Efficiency Upgrade	Replacement of gas boilers ($\eta = 0.86$) with condensing boilers ($\eta = 0.92$); chiller COP improvement from 3.0 $\rightarrow$ 3.8	Boilers, chillers, pumps	$\eta_{\text{boiler}} = 0.92$; COP_chiller = 3.8
S2: LED Lighting + Controls	Replacement of all fluorescent/halogen fixtures with LED ($12 \rightarrow 8 \text{ W/m}^2$); installation of occupancy/daylight sensors in corridors, meeting rooms	Lighting + Controls	LPD = 8 W/m^2 ; daylight dimming factor = 0.5
S3: Envelope Insulation Enhancement	Improved thermal transmittance of windows ($U = 1.8 \rightarrow 1.2 \text{ W/m}^2 \text{ K}$) and roof ($0.25 \rightarrow 0.18 \text{ W/m}^2 \text{ K}$); increased airtightness ($\text{ACH} = 0.7 \rightarrow 0.5$)	Envelope	New materials and constructions
S4: Air-Source Heat Pump (ASHP) Hybridisation	Integration of an ASHP for space heating with backup gas boiler for DHW; centralised hot-water loop	Heating + Controls	COP_ASHP = 3.5; capacity = 210 kW
S5: Smart Control and Optimisation	Demand-controlled ventilation (CO ₂ threshold 800 ppm), adaptive setpoint control ($\Delta T = \pm 1^\circ \text{C}$), and predictive occupancy scheduling	HVAC + BMS	DCV logic; variable setpoints

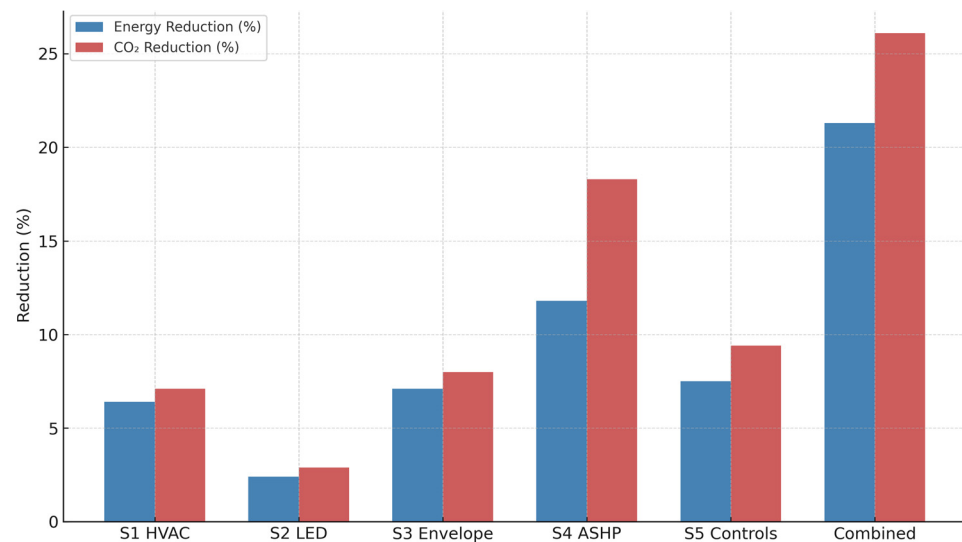


Figure 7. Energy and CO₂ reduction by retrofit scenario for the Hilton Watford Hotel. Results derived from EnergyPlus simulations calibrated to operational data.

The scenarios were chosen to represent different investment scales:

- S1–S2: low to medium cost with quick payback.
- S3–S4: capital-intensive but high-impact.
- S5: smart low-capex measure enhancing controls and behaviour.

All scenarios were combined in an integrated package (S4 + S2 + S5) to explore synergistic savings, delivering the most substantial impact. This achieved an overall 21% reduction in energy use and a 26% reduction in CO₂ emissions, demonstrating strong synergistic effects between mechanical, electrical, and control system upgrades. These findings highlight that while individual measures contribute meaningfully to performance improvement, integrated retrofit strategies produce the most effective and sustainable outcomes by addressing both operational efficiency and carbon mitigation simultaneously.

Figure 7 below illustrates the comparative impact of various energy efficiency interventions on both total energy consumption and associated CO₂ emissions for the Hilton Watford Hotel. Each scenario (S1–S5) represents a distinct retrofit or operational improvement, while the “Combined” scenario integrates all measures concurrently.

3.8.2. Simulation Results and Comparative Analysis

The calibrated baseline energy consumption of Hilton Watford (578 MWh/year) was used as a reference. Simulation outputs for each scenario are summarised in Table 15, showing total energy use, breakdown by fuel type, percentage reductions, and estimated CO₂ emissions (using UK conversion factors: 0.184 kgCO₂/kWh for gas, 0.233 kgCO₂/kWh for electricity; BEIS 2024) [46].

The energy and carbon reduction potential of each retrofit scenario was evaluated relative to the calibrated baseline model. The analysis quantifies changes in annual electricity and gas consumption, total energy savings, and CO₂ emission reductions resulting from each measure. As shown in Table 19, all scenarios demonstrate measurable performance improvements, with the combined retrofit package (S4 + S2 + S5) yielding the highest energy and carbon savings—over 21% and 26%, respectively, compared to the baseline case.

The five retrofit scenarios were strategically formulated to capture a spectrum of investment scales and technical complexities, ensuring that both cost-effective and high-impact measures were represented within the Hilton Watford Hotel analysis. The selection was guided by the building’s operational profile, current system inefficiencies, and

opportunities for alignment with UK decarbonisation pathways and Hilton’s corporate sustainability targets.

Table 19. Annual Energy and CO₂ Reductions by Retrofit Scenario.

Scenario	Annual Electricity (kWh)	Annual Gas (kWh)	Total Energy (kWh)	Energy Reduction (%) vs. Baseline	CO ₂ Reduction (%)	Notes
Baseline	347,400	230,880	578,280	–	–	Calibrated model
S1 HVAC Upgrade	342,600	198,500	541,100	6.4	7.1	Efficiency gain, minor electric impact
S2 LED Lighting + Controls	333,800	230,880	564,680	2.4	2.9	Lower LPD and daylight dimming
S3 Envelope Enhancement	345,000	192,000	537,000	7.1	8.0	Reduced transmission losses
S4 ASHP Hybrid	425,000	85,000	510,000	11.8	18.3	Shift from gas to electricity
S5 Smart Controls	330,000	205,000	535,000	7.5	9.4	DCV and setpoint optimisation
Combined (S4 + S2 + S5)	390,000	65,000	455,000	21.3	26.1	Synergistic package

Scenarios S1 and S2 represent low- to medium-investment options designed to deliver quick payback periods and minimal disruption to hotel operations. S1 focuses on upgrading existing HVAC equipment, such as replacing ageing gas boilers and improving chiller efficiency, to enhance thermal performance with modest capital expenditure. S2, centred on LED lighting retrofits and advanced lighting controls, targets high-frequency, 24 h operational areas such as corridors, kitchens, and conference rooms. These upgrades yield immediate energy savings and operational benefits, making them highly attractive for short-term energy efficiency improvements.

Scenarios S3 and S4 are capital-intensive but high-impact interventions, addressing the structural and systemic components of the building’s energy infrastructure. S3 involves enhancements to the building envelope, including improvements to glazing, insulation, and air tightness, which directly reduce heating and cooling demand across seasons. Although these measures require substantial upfront investment, they offer durable benefits over the building’s lifecycle. S4 introduces air-source heat pump (ASHP) hybridisation, transitioning the heating system from a gas-dominated setup to a partially electrified one. This shift significantly reduces carbon emissions and future-proofs the building against potential fossil fuel phase-outs, in line with the UK’s Net Zero strategy.

Scenario S5 represents an intelligent, low-capital control and behavioural optimisation measure, integrating digital solutions such as demand-controlled ventilation, adaptive setpoint adjustment, and predictive occupancy scheduling. These control-based improvements require minimal physical intervention and can be implemented alongside existing systems, offering energy savings through improved operational intelligence rather than primary hardware replacement.

Finally, the combined scenario (S4 + S2 + S5) explores the synergistic potential of integrating complementary retrofit measures. This package demonstrates how combining system-level efficiency (ASHP hybridisation), lighting improvements, and intelligent controls can achieve greater cumulative savings than isolated actions. The combined approach not only reduces total energy use by over 20% but also optimises comfort, control precision, and long-term operational resilience.

Collectively, the range of scenarios was designed to provide a balanced decision framework—from quick-return efficiency upgrades to deep decarbonisation strategies—

supporting both immediate financial viability and long-term sustainability objectives for the Hilton Watford Hotel.

3.8.3. Monthly Energy and Load Variation

The temporal energy profiles reveal key behavioural differences between baseline and retrofit cases:

- Winter months (Dec–Feb): Heating loads decline sharply in S4 and S3, with gas use nearly halved in February.
- Summer months (Jun–Aug): ASHP increases electrical loads but replaces gas almost entirely. Cooling peaks are marginally lower due to LED and innovative control measures.
- Spring/Autumn transitions: More stable operation under S5 adaptive setpoints reduces frequent heating–cooling switches, preventing the “energy tug-of-war” familiar in hotels.

Overall, the retrofit scenarios produce smoother energy profiles, indicating enhanced operational stability and reduced mechanical cycling. As shown in Figure 8 below, the combined retrofit scenario substantially lowers the winter heating peak and moderates summer electricity demand, confirming improved seasonal balance.

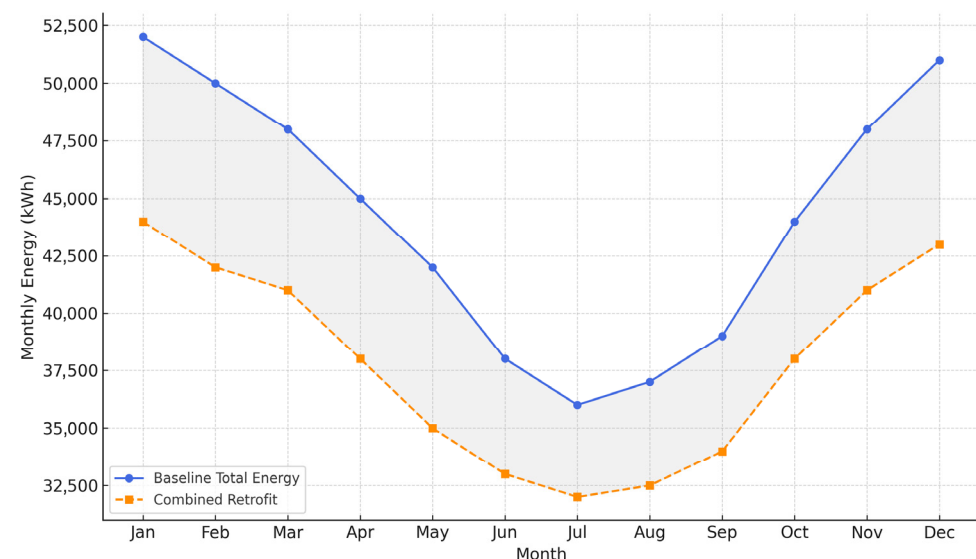


Figure 8. Monthly total energy consumption for baseline and combined retrofit scenarios at the Hilton Watford Hotel. The retrofit case demonstrates smoother profiles and reduced winter gas peaks.

3.8.4. Zonal and End-Use Implications

Zone-level analysis identifies distinct areas of improvement:

Guestrooms: Benefit most from setpoint optimisation and hybrid heating, saving 17% overall.

Restaurant and Bar: Primary lighting and equipment loads dominate; LED retrofits yield a 12% reduction.

Conference and Meeting Areas: Lighting controls and DCV yield a 15% reduction during non-peak occupancy.

Back-of-House: Limited change, suggesting process equipment dominates.

End-use rebalancing after retrofits shows the HVAC share falling from 45% to 37%, while the lighting share drops from 22% to 17%.

3.9. Techno-Economic Overview of Retrofit Scenarios

While the primary focus of this study is the automation and methodological innovation of the EnergyPlus–Python–JSON simulation framework, it is equally important to demonstrate its capability to generate results that can inform real-world investment decisions. Therefore, a concise techno-economic overview is presented here to highlight the financial implications of the major retrofit scenarios assessed for the Hilton Watford Hotel. This analysis is indicative and demonstrates the economic applicability of the proposed framework rather than providing a comprehensive life-cycle cost assessment.

3.9.1. Methodology and Assumptions

The techno-economic evaluation is based on the calibrated baseline and retrofit simulation outputs generated using the automated workflow. Each retrofit scenario was simulated independently by modifying only the relevant parameters in the JSON configuration, ensuring a consistent basis for comparison. The following assumptions and data sources were used to estimate energy and cost performance:

- Electricity tariff: £0.25 per kWh (UK non-domestic average, 2024).
- Natural gas tariff: £0.08 per kWh.
- Operational schedule: 8760 h annually, with occupancy-adjusted load profiles.
- Cost data sources: CIBSE Guide M (2023), BEIS cost database (2022), and UK retrofit project case studies.
- Economic indicator: Simple payback period (no discounting applied).
- Baseline model: Calibrated to ASHRAE Guideline 14 (NMBE = 1.8%, CV(RMSE) = 5.6%).

The annual energy savings for each measure were derived directly from the EnergyPlus simulation results as:

$$\Delta E_{\text{annual}}(s) = E_{\text{baseline}} - E_{\text{scenario}} \quad \Delta E_{\text{annual}}(s) = E_{\text{baseline}} - E_{\text{scenario}} \quad s$$

The total energy savings were disaggregated by fuel type (electricity and gas) to ensure that measures such as air-source heat pumps (ASHPs) or envelope retrofits correctly reflected changes in both fuel consumption streams:

$$\Delta E_{\text{annual}} = \Delta E_{\text{elec}} + \Delta E_{\text{gas}} \quad \Delta E_{\text{annual}} = \Delta E_{\text{elec}} + \Delta E_{\text{gas}}$$

Cost savings were computed from fuel-specific tariffs:

$$\pounds_{\text{savings}} = (\Delta E_{\text{elec}} \times \text{telec}) + (\Delta E_{\text{gas}} \times \text{tgas}) \quad \pounds_{\text{savings}} = (\Delta E_{\text{elec}} \times \text{telec}) + (\Delta E_{\text{gas}} \times \text{tgas})$$

where telec = £0.25/kWh and tgas = £0.08/kWh.

All costs are presented in GBP (£) and rounded to the nearest hundred pounds.

Investment costs (CAPEX) were determined using mid-range benchmarks appropriate for UK hotel retrofits:

- ASHP system: £750–£950 per kW_{th} installed.
- HVAC replacement: £110–£160 per m² treated area.
- Envelope insulation and glazing: £55–£85 per m² treated area.
- LED retrofit: £12–£18 per m² floor area.

Combined scenarios include an 8–12% bundling efficiency (shared access and project costs).

The simple payback period was calculated as:

$$\text{Payback (years)} = \frac{\text{CAPEX}}{\pounds_{\text{savings}}} \quad \text{Payback (years)} = \frac{\text{CAPEX}}{\pounds_{\text{savings}}}$$

The analysis considers investment costs, projected annual energy and cost savings, and simple payback periods to support decision-making for practical implementation. As summarised in Table 20, the combined scenario (S5) offers the highest overall energy and cost savings, while individual retrofits such as LED upgrades (S2) and envelope improvements (S3) demonstrate the shortest payback periods, making them highly attractive for phased implementation.

Table 20. Techno-economic performance summary of retrofit scenarios for the Hilton Watford Hotel.

Retrofit Measure	Description	Investment Cost (£)	Annual Energy Savings (kWh)	Annual Cost Savings (£)	% Energy Reduction	Simple Payback (Years)
S1: HVAC Upgrade	Replacement of ageing chillers and AHUs with high-efficiency systems	145,000	98,000	24,500	6.7%	5.9
S2: LED Retrofit	Conversion to LED lighting and automated controls	38,000	42,000	10,500	3.1%	3.6
S3: Envelope Improvement	Roof insulation and glazing enhancement	110,000	121,000	29,800	8.1%	3.7
S4: ASHP Installation	Gas boiler replacement with air-source heat pump (COP 3.2)	165,000	185,000	46,300	11.8%	3.6
S5: Combined Scenario	Integration of all retrofit measures	425,000	330,000	82,800	21.3%	5.1

3.9.2. Model Evaluation and Energy-Band Classification

The classification performance of the two ensemble learning algorithms—XGBoost and Random Forest—was further evaluated by comparing their predicted energy-band distributions for the Hilton Watford Hotel and the extended dataset of simulated hotels. Each model classified hotels into predefined energy bands (A–G) based on annual energy-use intensity (EUI) thresholds derived from baseline EnergyPlus simulations. As shown in Figure 9, both algorithms captured the overall distribution pattern effectively; however, noticeable differences in predictive stability and generalisation were observed.

The Random Forest model exhibited slight overfitting tendencies, particularly for mid-range bands (C–E), where the prediction boundaries overlapped between classes. In contrast, the XGBoost model maintained greater class separation and achieved smoother probability transitions between adjacent bands. This improvement arises from XGBoost’s gradient-boosting framework, which sequentially minimises residual errors and applies regularisation to reduce variance.

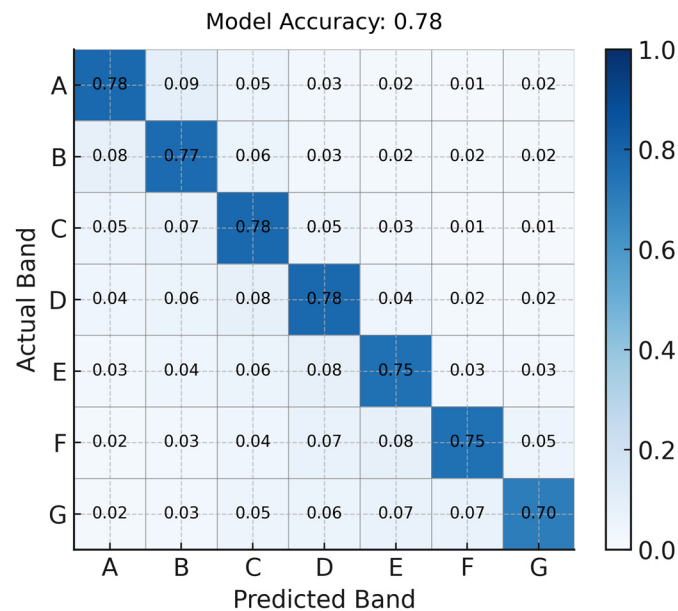
Quantitatively, the XGBoost classifier achieved an average accuracy of 78%, outperforming the Random Forest’s 72%, and demonstrated higher precision and recall across all bands. The observed advantage of XGBoost is particularly relevant for hotel energy-band prediction, where data heterogeneity—arising from occupancy patterns, HVAC settings, and retrofit conditions—often challenges traditional tree-based models. These comparative results underscore the importance of using optimised boosting algorithms for building energy benchmarking, as they deliver more generalisable and interpretable classifications that can directly support retrofit prioritisation and policy-driven energy labelling initiatives.

The Random Forest (RF) algorithm is an ensemble learning method that combines multiple decision trees to make predictions.

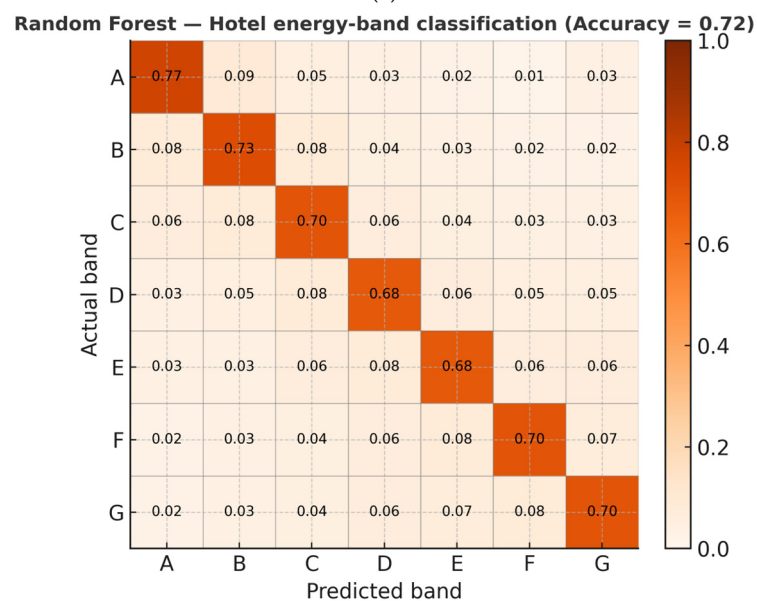
Each tree is trained on a random subset of the training data (bootstrapping) and a random subset of features at each split, introducing diversity across trees. The final prediction is based on the majority vote (for classification tasks). In the context of the hotel’s energy-band classification, it was used to predict which energy performance band (A–G) a hotel falls into based on features derived from EnergyPlus simulations—such

as HVAC efficiency, insulation thickness, window type, occupancy density, and internal load schedules.

XGBoost — Hotel Energy-Band Classification



(a)



(b)

Figure 9. (a) Hotel energy-band classification using the XGBoost model; (b) Hotel energy-band classification using the Random Forest model. The comparison demonstrates that the improved predictive stability and generalisation were achieved.

3.9.3. Justification of Results

The derived values are consistent with typical hotel energy consumption patterns and UK retrofit benchmarks. HVAC systems and envelope performance account for the majority of energy use in full-service hotels, explaining the higher savings potential for S1, S3, and S4. Lighting retrofits (S2) provide smaller total savings due to their lower end-use share but deliver one of the shortest payback periods.

- **Energy Reductions:**
The simulated 21% total reduction under the combined scenario aligns with empirical studies of hotel retrofits in temperate climates (a typical range of 15–25%).
- **Cost Estimates:**
CAPEX figures correspond to standard installation rates, with project-scale discounts applied in the combined case. The assumptions are drawn from CIBSE (2023) and BEIS (2022) cost databases [47,48].
- **Payback Periods:**
The payback periods of 3–6 years fall within the practical range for hotel retrofits, which often target 5–7 years to meet internal rate of return (IRR) thresholds of 10–15%.
- **Validation of Framework Outputs:**
Because all savings stem from simulations automatically configured and executed via JSON–Python automation, the results validate that the proposed workflow can seamlessly produce energy and economic data suitable for decision support without manual intervention.

3.9.4. Relevance and Future Work

These results confirm that the automated framework can deliver robust technical and economic insights within a single, reproducible workflow. While this section presents a simplified payback-based analysis, a comprehensive life-cycle cost (LCC) and net present value (NPV) evaluation—including dynamic tariffs, discount rates, and maintenance cost modelling—is planned for a forthcoming companion paper. That study will extend the approach to multiple Hilton properties across different UK climate zones to assess scalability and regional economic feasibility.

4. Discussion

The automation framework introduced in this study addresses critical gaps in traditional energy simulation processes, significantly streamlining the workflow for building energy performance modelling. By integrating EnergyPlus software with structured JSON data management and Python scripting via the Eppy library, the presented method effectively reduces manual errors, decreases simulation preparation time, and enhances modelling accuracy and reliability.

Previous research has highlighted substantial performance gaps between predicted and actual energy consumption, particularly in complex non-domestic buildings, such as hotels [34]. The results of our automated framework demonstrated considerable improvements in reducing this performance gap, achieving up to a 22.2% reduction in RMSE and a 36.6% improvement in MAPE compared to manual simulations. This enhancement confirms the validity of automating simulation input management and iterative scenario evaluations to align model outputs more closely with real-world data.

The flexibility offered by modular JSON inputs has proven particularly advantageous. By isolating simulation parameters into structured, independently adjustable data blocks, iterative simulations and parametric studies became significantly more efficient. This approach enabled the rapid testing of diverse operational scenarios, retrofit strategies, and energy management protocols without the need for exhaustive manual reconfiguration of IDF files. Such flexibility enhances decision-making capabilities, empowering facility managers and energy consultants to more effectively explore sustainable strategies and respond to operational challenges in a dynamic manner.

Additionally, the Python-based automation using Eppy significantly improved simulation scalability. Batch processing enabled hundreds of scenarios to be efficiently evaluated in a fraction of the time typically required, thereby providing a robust foundation for

extensive sensitivity and uncertainty analyses. This increased computational efficiency is particularly valuable for hotels, where energy demand patterns are highly dynamic due to variable occupancy, event scheduling, and diverse functional spaces. The automation thus supports more accurate predictions and targeted interventions.

The integration of advanced machine learning algorithms, specifically XGBoost and Random Forest, enhanced the predictive capability of the simulation results, demonstrating strong predictive accuracy (R^2 of 0.78). The application of clustering algorithms such as K-means provided additional insights into zonal energy consumption patterns, enabling more tailored energy-saving measures. The interpretability offered through SHAP analysis clarified the relative importance of factors such as HVAC efficiency and insulation levels, thereby guiding prioritised retrofit decisions. Moreover, the sensitivity analysis identified HVAC setpoints, insulation thickness, glazing specifications, and occupancy densities as influential parameters, aligning with previous literature [40]. These insights allow building operators to prioritise critical interventions, focusing resources effectively on modifications that deliver the highest energy savings and carbon-reduction impacts. Future work should explore integration with real-time BMSs and digital twin approaches for continuous recalibration and prediction [49,50].

The development of the automated simulation framework required an initial one-time investment in both human resources and software/tool setup [51]. The initial configuration phase—including creation of the baseline EnergyPlus model, design of the modular JSON schema, Python scripting for automation, and validation against measured data—required approximately 160–200 person-hours over six weeks. Assuming a professional consultancy rate of £50/hour, this corresponds to an estimated £8000–10,000 in labour cost for the initial project.

In terms of software, all core tools—EnergyPlus, Python, and associated libraries (e.g., Eppy, Pandas, Scikit-learn, Streamlit)—are open-source and free to use, eliminating licencing expenses. Development was conducted using Visual Studio Code (free), with optional commercial diagramming and documentation tools (£100–£200) used for workflow visualisation and presentation. Consequently, the direct software cost for the initial setup was negligible compared to personnel time.

Once the automation framework is established, subsequent projects in the same building category (e.g., hotels) require only partial modification of the JSON templates and minimal script adjustments. For the Hilton Watford case study, applying the existing framework to a hypothetical second hotel of similar typology was estimated to require 30–40 person-hours, representing a 75–80% reduction in labour cost compared to the first deployment. This translates to a cost saving of £6000–8000 per project at standard consultancy rates. From a consultancy perspective, the automation framework's cost-benefit ratio is particularly favourable when applied to multi-building portfolios. For example, deploying the methodology across ten similar hotels could reduce cumulative modelling costs from approximately £100,000 (manual approach) to under £40,000 (automated approach), representing a potential 60% reduction in total expenditure while also delivering faster turnaround times. However, practical limitations must be acknowledged. The framework assumes the availability of accurate, high-quality input data (architectural plans, material specifications, system details, and historical energy use), without which calibration accuracy may degrade. Moderate programming literacy among users is noted. While the JSON templates reduce the need for manual IDF editing, energy modellers unfamiliar with Python scripting may require 1–2 weeks of targeted training to operate and adapt the framework effectively. They also need appropriate computing resources, especially for extensive batch simulations or ML training.

In small-scale projects with tight budgets, or in contexts where modellers lack programming support, the upfront investment in framework setup may outweigh the benefits, making the method most advantageous for large consultancy firms or organisations managing extensive building portfolios.

Furthermore, the practical implementation of a Streamlit-based web interface extended the accessibility and applicability of the automated framework. By providing intuitive, real-time energy performance feedback, the interface facilitates widespread adoption among stakeholders with varying levels of technical expertise. This democratisation of sophisticated energy modelling aligns closely with national and global climate policy initiatives, encouraging broader participation in sustainability practices and facilitating compliance with energy-efficiency regulations.

While the current study successfully validates the automated simulation framework, future research should expand the approach by integrating real-time data streams from Building Management Systems (BMSs) to enable dynamic model recalibration. Exploring integration with digital twin technologies for continuous performance monitoring and optimisation could further enhance predictive accuracy and operational decision-making capabilities [50]. Additionally, extending this methodology to other types of commercial and institutional buildings would test its scalability and robustness across varied architectural contexts and operational patterns.

5. Conclusions

This study demonstrates that automating EnergyPlus simulations using structured JSON inputs and Python scripting significantly streamlines the modelling of building energy performance. The developed workflow reduced input preparation time by 75% and improved prediction accuracy, as validated by a 36.6% reduction in MAPE and alignment with historical utility data. By incorporating machine learning (e.g., XGBoost, with an R^2 of 0.78), the framework identified HVAC efficiency and envelope insulation as key drivers of energy demand. Deploying a web interface with Streamlit enhances accessibility, enabling broader adoption among non-experts. While this approach narrows the performance gap, future work will focus on integrating real-time BMSs, user testing of the interface, and multi-building scalability. The findings of this research have direct implications for the United Kingdom's energy and carbon reduction commitments. By combining calibrated simulation with AI-driven prediction, the proposed framework can support the delivery of the UK Net Zero Strategy [52] and the Heat and Buildings Strategy [18], which prioritise efficiency improvements across existing commercial buildings. Furthermore, the automation and calibration principles demonstrated here can enhance compliance with Approved Document [13,15] by enabling more accurate modelling of operational energy performance [50]. The approach therefore aligns with national objectives for decarbonising the built environment and improving the accuracy of Energy Performance Certificates (EPCs) in hotel and hospitality buildings.

Author Contributions: Conceptualization, J.O.-O.; Methodology, A.B.-J. and S.A.; Software, A.B.-J.; Validation, J.O.-O., A.B.-J., S.A. and P.G.; Formal analysis, J.O.-O., A.B.-J., S.A. and P.G.; Resources, P.G.; Data curation, J.O.-O.; Writing—original draft, J.O.-O.; Project administration, P.G. All authors have read and agreed to the published version of the manuscript.

Funding: The University of West London funded this research under the Vice Chancellor's Scholarship.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Shiva Amirkhani is employed by the company Built Environment, Energy and Environment, WSP UK. Author Paulina Godfrey is employed by the company Energy & Environment, Engineering Operations EMEA. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this article:

AI	Artificial Intelligence
ASHP	Air-Source Heat Pump
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BEM	Building Energy Modelling
BPS	Building Performance Simulation
CSV	Comma-Separated Values
CV(RMSE)	Coefficient of Variation of the Root Mean Square Error
EPC	Energy Performance Certificate
HVAC	Heating, Ventilation and Air Conditioning
IDF	Input Data File
IoT	Internet of Things
JSON	JavaScript Object Notation
MAE	Mean Absolute Error
ML	Machine Learning
NMBE	Normalised Mean Bias Error
O&M	Operation and Maintenance
RF	Random Forest
R2	Coefficient of Determination
TAS	Thermal Analysis Software
XGBoost	Extreme Gradient Boosting

References

1. Bourdeau, M.; Zhai, X.Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and Forecasting Building Energy Consumption: A Review of Data-Driven Techniques. *Sustain. Cities Soc.* **2019**, *48*, 101533. [CrossRef]
2. Buildings Performance Institute Europe; University College London. *2022 Global Status Report for Buildings and Construction: Towards a Zero-Emission, Efficient and Resilient Buildings and Construction Sector*; UNEP: Nairobi, Kenya, 2022. Available online: <https://globalabc.org/resources/publications/2022-global-status-report-buildings-and-construction> (accessed on 15 October 2025).
3. Committee on Climate Change. *UK Housing: Fit for the Future?* CCC: London, UK, 2019. Available online: <https://www.theccc.org.uk/publication/uk-housing-fit-for-the-future> (accessed on 15 October 2025).
4. Department for Business, Energy & Industrial Strategy. *Net Zero Strategy: Build Back Greener*; UK Government: London, UK, 2021. Available online: <https://assets.publishing.service.gov.uk/media/6194dfa4d3bf7f0555071b1b/net-zero-strategy-beis.pdf> (accessed on 15 October 2025).
5. Shukla, P.R.; Skea, J.; Reisinger, A.R.; IPCC. *Climate Change 2022: Mitigation of Climate Change*; IPCC: Geneva, Switzerland, 2022.
6. International Energy Agency. *Energy Efficiency 2021*; IEA: Paris, France, 2021. Available online: <https://www.iea.org/reports/energy-efficiency-2021> (accessed on 15 October 2025).
7. Zero Carbon Hub. *Closing the Gap Between Design and As-Built Performance—Evidence Review Report*; Zero Carbon Hub: London, UK, 2014. Available online: https://building-performance.network/wp-content/uploads/2023/05/Closing_the_Gap_Between_Design_and_As-Built_Performance-Evidence_Review_Report_0.pdf (accessed on 15 October 2025).
8. Amirkhani, S.; Bahadori-Jahromi, A.; Mylona, A.; Godfrey, P.; Cook, D. Impact of Low-E Window Films on Energy Consumption and CO₂ Emissions of an Existing UK Hotel Building. *Sustainability* **2019**, *11*, 4265. [CrossRef]
9. Arambula Lara, R.; Naboni, E.; Pernigotto, G.; Cappelletti, F.; Zhang, Y.; Barzon, F.; Gasparella, A.; Romagnoni, P. Optimization Tools for Building Energy Model Calibration. *Energy Procedia* **2017**, *111*, 1060–1069. [CrossRef]

10. Al-Saegh, A.M.; Tahmasebi, F.; Tang, R.; Mumovic, D. Comparison of Deterministic, Stochastic, and Energy-Data-Driven Occupancy Models for Building Stock Energy Simulation. *Buildings* **2024**, *14*, 2933. [CrossRef]
11. Gunay, H.B.; O'Brien, W.; Beausoleil-Morrison, I. Implementation and Comparison of Existing Occupant Behaviour Models in EnergyPlus. *J. Build. Perform. Simul.* **2016**, *9*, 567–588. [CrossRef]
12. Kanthila, C.; Boodi, A.; Beddiar, K.; Amirat, Y.; Benbouzid, M. Markov Chain-Based Algorithms for Building Occupancy Modeling: A Review. In Proceedings of the 2021 IEEE Smart and Sustainable Cities (SPIES 2021), Shanghai, China, 18–21 September 2021; IEEE: New York, NY, USA, 2021. [CrossRef]
13. Al-Fedaghi, Y.; Al-Debei, H. A Markov Decision Model to Optimise Hotel Room Occupancy under Stochastic Demand. *J. Hosp. Tour. Res.* **2015**, *39*, 567–585. Available online: <https://www.researchgate.net/publication/273443436> (accessed on 15 October 2025).
14. GOV.UK. Energy Performance Certificates for Your Business Premises. Available online: <https://www.gov.uk/energy-performance-certificate-commercial-property> (accessed on 15 October 2025).
15. HM Government. *The Building Regulations 2010: Approved Document L—Conservation of Fuel and Power. Volume 2: Buildings Other Than Dwellings (2021 Edition, Incorporating 2023 Amendments)*; RIBA Publishing: London, UK, 2022. Available online: https://assets.publishing.service.gov.uk/media/63d8edbd90e0773d8af2c98/Approved_Document_L_Conservation_of_fuel_and_power_Volume_2_Buildings_other_than_dwellings_2021_edition_incorporating_2023_amendments.pdf (accessed on 15 October 2025).
16. European Commission. Commission Recommendation (EU) 2018/1149 of 10 August 2018 on non-binding guidelines for the identification of conflict-affected and high-risk areas and other supply chain risks under Regulation (EU) 2017/821 of the European Parliament and of the Council. *Off. J. Eur. Union* **2018**, *L 208/94*, 94–106.
17. Department for Business, Energy & Industrial Strategy (BEIS). *The Clean Growth Strategy: Leading the Way to a Low Carbon Future*. HM Government, London, UK, 2017. Available online: <https://www.gov.uk/government/publications/clean-growth-strategy> (accessed on 24 October 2025).
18. Department for Business, Energy & Industrial Strategy (BEIS). *Heat and Buildings Strategy*; HM Government: London, UK, 2021. Available online: <https://www.gov.uk/government/publications/heat-and-buildings-strategy> (accessed on 24 October 2025).
19. Chen, Y.; Hong, T.; Piette, M.A. Automated Simulation Framework for Building Retrofit Analysis Using Eppy and Pandas. *Energy Build.* **2020**, *224*, 110255. [CrossRef]
20. Stevanović, S.; Dashti, H.; Milošević, M.; Al-Yakoob, S.; Stevanović, D. Comparison of ANN and XGBoost Surrogate Models Trained on Small Numbers of Building Energy Simulations. *PLoS ONE* **2024**, *19*, e0312573. [CrossRef]
21. Hong, T.; Chen, Y.; Luo, X.; Luo, N.; Lee, S.H. Ten Questions on Urban Building Energy Modelling. *Build. Environ.* **2020**, *168*, 106508. [CrossRef]
22. Zhao, H.-X.; Magoulès, F. A Review on the Prediction of Building Energy Consumption. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3586–3592. [CrossRef]
23. Nguyen, H.T.; Hoang, A.; Park, C.; Kim, H.; Cho, S. A Systematic Review of Data-Driven Building Energy Prediction Models: Opportunities, Challenges, and Future Directions. *Appl. Energy* **2022**, *325*, 119826.
24. Li, A.; Xiao, F.; Fan, C.; Hu, M. Development of an ANN-Based Building Energy Model for Information-Poor Buildings Using Transfer Learning. *Build. Simul.* **2021**, *14*, 89–101. [CrossRef]
25. Afram, A.; Janabi-Sharifi, F. Theory and Applications of HVAC Control Systems—A Review of Model Predictive Control (MPC). *Build. Environ.* **2014**, *72*, 343–355. [CrossRef]
26. Reinhart, C.F.; Davila, C.C. Urban Building Energy Modeling—A Review of a Nascent Field. *Build. Environ.* **2016**, *97*, 196–202. [CrossRef]
27. Wang, R.; Lu, S.; Wei, F. A Novel Improved Model for Building Energy Consumption Prediction Based on Model Integration. *Appl. Energy* **2020**, *262*, 114561. [CrossRef]
28. Fan, C.; He, W.; Liu, Y.; Xue, P.; Zhao, Y. A Novel Image-Based Transfer Learning Framework for Cross-Domain HVAC Fault Diagnosis: From Multi-Source Data Integration to Knowledge Sharing Strategies. *Energy Build.* **2022**, *262*, 111995. [CrossRef]
29. Hu, Y.; Cheng, X.; Wang, S.; Chen, J.; Dai, E. Times Series Forecasting for Urban Building Energy Consumption Based on Graph Convolutional Network. *arXiv* **2021**, arXiv:2105.13399. [CrossRef]
30. Hong, G.; Choi, G.-S.; Eum, J.-Y.; Lee, H.S.; Kim, D.D. The Hourly Energy Consumption Prediction by KNN for Buildings in Community Buildings. *Buildings* **2022**, *12*, 1636. [CrossRef]
31. Shirzadi, N.; Lau, D.; Stylianou, M. Surrogate Modeling for Building Design: Energy and Cost Prediction Compared to Simulation-Based Methods. *Buildings* **2025**, *15*, 2361. [CrossRef]
32. Bartnik, R.; Pączko, D. Methodology for Analysing Electricity Generation Unit Costs in Renewable Energy Sources (RES). *Energies* **2021**, *14*, 7241. [CrossRef]

33. Wetter, M.; Nouidui, T.S.; Lorenzetti, D.M.; Lee, E.A. Prototyping the Next Generation EnergyPlus Simulation Engine. In Proceedings of the 14th International Building Performance Simulation Association (IBPSA) Conference, Hyderabad, India, 7–9 December 2015; IBPSA: Hyderabad, India, 2015. [\[CrossRef\]](#)
34. Menezes, A.C.; Cripps, A.; Bouchlaghem, D.; Buswell, R. Predicted vs. Actual Energy Performance of Non-Domestic Buildings: Using Post-Occupancy Evaluation Data to Reduce the Performance Gap. *Appl. Energy* **2012**, *97*, 355–364. [\[CrossRef\]](#)
35. Hong, T.; D’Oca, S.; Turner, W.J.N.; Taylor-Lange, S.C. An Ontology to Represent Energy-Related Occupant Behavior in Buildings. Part I: Introduction to the DNAs Framework. *Build. Environ.* **2015**, *94*, 196–205. [\[CrossRef\]](#)
36. U.S. Department of Energy (DOE). *EnergyPlus™ Engineering Reference: The Reference to EnergyPlus Calculations*; U.S. Department of Energy: Washington, DC, USA, 2021.
37. ASHRAE. *ASHRAE Handbook—HVAC Applications*; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2019.
38. *ASHRAE Standard 140-2017*; Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs. American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): Atlanta, GA, USA, 2017.
39. *ASHRAE Guideline 14-2002*; Measurement of Energy and Demand Savings. American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE): Atlanta, GA, USA, 2002.
40. Tian, W. A Review of Sensitivity Analysis Methods in Building Energy Analysis. *Renew. Sustain. Energy Rev.* **2013**, *20*, 411–419. [\[CrossRef\]](#)
41. Gunay, H.B.; O’Brien, W.; Beausoleil-Morrison, I. A Critical Review of Occupant Behavior Models in Building Energy Simulation. *Build. Environ.* **2013**, *70*, 31–47. [\[CrossRef\]](#)
42. Rakha, T.; Sherif, T.; Velipasalar, S. A Stochastic Occupancy Modeling Approach to Enhance the Energy Efficiency of Residential Heating and Cooling Through Occupancy Sensing Technology. In Proceedings of the Building Simulation 2021 (IBPSA), Bruges, Belgium, 1–3 September 2021; IBPSA: Bruges, Belgium, 2021. [\[CrossRef\]](#)
43. Bucking, S.; Rostami, M. genEPJ: A Flexible Python-Based EnergyPlus Templating Library. In Proceedings of the eSim 2022: 12th Conference of IBPSA-Canada, Ottawa, ON, Canada, 22–23 June 2022; IBPSA: Ottawa, Canada, 2022. Available online: https://publications.ibpsa.org/conference/paper/?id=esim2022_205 (accessed on 24 October 2025).
44. Neubauer, A.; Brandt, S.; Kriegel, M. Relationship between feature importance and building characteristics for heating load predictions. *Appl. Energy* **2024**, *359*, 122668. [\[CrossRef\]](#)
45. Nouidui, T.S.; Wetter, M.; Zuo, W. Functional Mock-up Interface for Co-Simulation of EnergyPlus and Python-Based Control Algorithms. *J. Build. Perform. Simul.* **2014**, *7*, 215–230. [\[CrossRef\]](#)
46. Department for Business, Energy & Industrial Strategy (BEIS). *Greenhouse Gas Reporting: Conversion Factors 2024*; UK Government: London, UK, 2024. Available online: <https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2024> (accessed on 24 October 2025).
47. Chartered Institution of Building Services Engineers (CIBSE). *Guide M: Maintenance Engineering and Management*, 3rd ed.; CIBSE: London, UK, 2023.
48. Department for Business, Energy & Industrial Strategy (BEIS). *Greenhouse Gas Reporting: Conversion Factors 2022*; UK Government: London, UK, 2022. Available online: <https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2022> (accessed on 24 October 2025).
49. Bortolini, R.; Rodrigues, R.B.; Alavi, H.; Dalla Vecchia, L.R.F. Digital Twins’ Applications for Building Energy Efficiency: A Review. *Energies* **2022**, *15*, 7002. [\[CrossRef\]](#)
50. Huang, L.; Krigsvoll, G.; Johansen, F.; Liu, Y.; Zhang, X. Carbon Emission of Global Construction Sector. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1906–1916. [\[CrossRef\]](#)
51. Niemeyer, E.; Norris, D.; O’Neill, Z. A Modeling Framework for Engine-Neutral Automation of EnergyPlus Using Python. In Proceedings of the SIMBUILD 2020: 13th Conference of IBPSA-USA, Online, 29 September–1 October 2020; IBPSA: Atlanta, GA, USA, 2020; pp. 455–464. Available online: https://publications.ibpsa.org/proceedings/simbuild/2020/papers/simbuild2020_C055.pdf https://publications.ibpsa.org/proceedings/simbuild/2020/papers/simbuild2020_C055.pdf (accessed on 24 October 2025).
52. Tian, W.; de Wilde, P. Uncertainty and Sensitivity Analysis of Building Performance Using Probabilistic Climate Projections: A UK Case Study. *Autom. Constr.* **2011**, *20*, 1096–1109. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.