Engineering Technology Trends Vol 3 Issue 1 2025

Original Research Article

Open Access

Omniscient Pte.Ltd.

Innovative research on building energy consumption and comfort optimization based on BIM-BECS-AI collaboration

Chun Yik Jenny Li[1], Yan Li[2,*]

[1]School of Architecture, Chongqing University, Chongqing, 400000, China
[2]Shenzhen Stock Exchange, Shenzhen, Guangdong, 518000, China

***Correspondence to:** Yan Li, Shenzhen Stock Exchange, Shenzhen, Guangdong, 518000, China, E-mail: rickyliyan@qq.com

Abstract: Against the backdrop of the continuous growth of global building energy consumption, achieving a balance between building energy conservation and indoor comfort has become a research hotspot. Traditional optimization methods mostly rely on a single simulation tool or empirical formula, which makes it difficult to deal with the nonlinear relationships and multi-objective conflicts of complex building systems. This study aims to provide a scientific and reasonable decision-making basis for building design and operation management by integrating domain knowledge enhancement technology and using the Pareto optimal solution of energy consumption intensity (EUI) and thermal comfort (PPD) based on the NSGA-II genetic algorithm and the hybrid superposition model (FNN+XGB), so as to achieve the goal of energy saving while ensuring indoor thermal comfort. Artificial intelligence algorithms have significantly improved the scientific nature of green building design and the iterative efficiency of energy-saving solutions. Empirical studies based on public data sets (such as the London building data set) have verified the advantages of artificial intelligence algorithms in key indicators such as energy optimization efficiency.

Keywords: Pareto optimal solution; building energy conservation; hybrid superposition model (FNN+XGB)

1. Introduction

Global building energy consumption accounts for 36.3% of the total social energy consumption. It has become an industry consensus that green buildings can reduce energy consumption by 30%-50% through intelligent technology[1]. Traditional design methods rely on manual experience and static simulation tools, which are difficult to cope with complex scenarios such as dynamic changes in climate parameters and conflicts in multi-objective optimization (such as energy consumption and comfort balance). Artificial intelligence models, with their powerful semantic understanding, cross-modal reasoning and dynamic iteration capabilities, are promoting the paradigm shift of green building design from experience-driven to data and algorithm-coordinated driven[2].

2. Technical architecture and core capabilities

2.1 Multimodal Data Processing Architecture

BIM builds a digital foundation for green building design by integrating multimodal inputs such as text, BIM models, and sensor data. BIM tools such as Revit are used to convert building geometry parameters and material properties into parseable semantic information, and combined with BECS energy-saving software to simulate energy consumption, forming a dynamic optimization data closed loop; BECS software will apply a variety of key parameters in energy consumption simulation, which cover the building's geometric characteristics, material properties, equipment systems, operating modes, external climate conditions and other aspects. The following are some of the parameters that BECS software specifically applies in energy consumption simulation.

2.2 Building main body related parameters

(1) Building geometric parameters: including building orientation, building shape coefficient, and window-to-wall ratio .

(2) Material parameters of building envelope structures: including thermal parameters of wall materials, thermal parameters of roof materials, and thermal parameters of door and window materials.

2.3 Equipment system related parameters

(1) Heating system parameters: including heating method, heating equipment performance parameters, and heating system operating parameters.

(2) Air conditioning system parameters: including air conditioning type, air conditioning equipment performance parameters, and air conditioning system operating parameters.

(3) Ventilation system parameters: including ventilation methods, ventilation equipment performance parameters, and ventilation system operating parameters.

(4) Indoor personnel activity parameters: including personnel density and personnel work and rest time.

(5) Indoor equipment heat generation parameters: including office equipment heat generation and lighting equipment heat generation.

2.4 External environment related parameters

(1) Meteorological parameters: including temperature, solar radiation intensity, wind speed and direction.

(2) Surrounding environment parameters: including obstacles, site terrain, etc.

3. Algorithm Framework

In the technical framework of green building design and energy-saving solution iteration, the superposition model (hybrid superposition model FNN+XGB) and BECS EnergyPlus work together and complement each other to jointly support the realization of multi-objective optimization goals.

(1) Data and simulation support: BECS provides basic input for the superposition model. BECS is responsible for processing multimodal data, simulating the energy consumption intensity (EUI) and thermal comfort (PPD) of different building design schemes through built-in physical models, and generating energy consumption data and simulation results for model training.

(2) Model prediction and optimization: The stacking model improves the simulation efficiency and decision-making value of BECS. The hybrid stacking model (FNN+XGB) combines the advantages of feedforward neural network (FNN) and XGBoost to achieve high-precision prediction of energy consumption intensity (EUI) and thermal comfort (PPD). The prediction results of the stacking model are input into the NSGA-II algorithm, and the Pareto frontier solution set is generated through multi-objective optimization iteration, which is fed back to BECS for secondary simulation verification to form a closed loop.

(3) Relationship between BECS and stacking model: BECS provides refined energy consumption simulation that complies with the laws of building physics, ensuring the physical authenticity and industry standardization of data. The stacking model uses machine learning technology to break through the limitations of traditional BECS that relies on manual experience to adjust parameters and static simulation, and realizes the automatic processing of dynamic climate parameters and multi-objective conflicts.

4. Algorithm Implementation

The following are the software implementation steps for achieving the Pareto optimal solution of energy usage intensity (EUI) and thermal comfort (PPD) on the London building dataset based on the NSGA-II genetic algorithm and the hybrid superposition model

(FNN+XGB)[3].

4.1 Data Preprocessing

(1) Collecting London building datasets

The London Database is a free and open data sharing platform initiated by the Greater London Authority, providing a rich dataset of London buildings, covering the physical characteristics, energy performance and occupancy costs of buildings. You can access and download the relevant data through the following links:

London Building Stock Model 2 (LBSM 2)[4]

London Building Stock Model (LBSM)[5]

(2) Data cleaning and normalization

Clean the collected data to remove outliers and missing values. Use interpolation or similar data from adjacent buildings to fill in missing data. Normalize all data so that the value range of each parameter is unified to the [0,1] interval.

4.2 Construction of mixed superposition model

(1) Feature selection and extraction

The key characteristic variables related to energy intensity and thermal comfort are selected from the preprocessed building dataset. The original features are reduced in dimension using feature extraction methods such as principal component analysis (PCA), redundant information is removed, and a more representative and discriminative feature subset is extracted to improve the training efficiency and accuracy of the model.

(2) Feedforward neural network (FNN) model construction

Use the PyTorch artificial intelligence framework to build the FNN model. First, determine the number of neurons in the input layer to match the number of extracted features. Then design the hidden layer.

Select a suitable activation function, such as using the ReLU activation function in the hidden layer, using the linear activation function in the output layer for EUI prediction, and using the Sigmoid activation function for PPD prediction (mapping the output value to the [0,1] interval to facilitate the subsequent calculation of the PPD indicator).

Write an FNN model, define the loss function (such as mean square error loss function MSE) and optimizer (such as Adam optimizer), and train the model. Adjust the model parameters through multiple iterations to gradually reduce the loss function value of the model on the training set. At the same time, evaluate on the validation set to prevent overfitting.

(3) XGBoost model training

Set the parameters of the XGBoost model, such as tree depth, learning rate, regularization parameterization, etc. Use methods such as cross-validation to tune the parameters and determine the optimal combination of model parameters so that the model can achieve good results on both the training set and the validation set.

(4) Hybrid stacking model construction

The trained FNN model is stacked and fused with the XGBoost model. A common stacking method is to input the outputs of the two models as new features into another linear regression model or a simple neural network for further integrated learning, thereby obtaining the final hybrid stacking model output, which is the comprehensive prediction value of energy intensity EUI and thermal comfort PPD.

The stacking model is trained and the parameters of the fusion layer are adjusted so that the hybrid stacking model can better fit the data as a whole. As an intelligent prediction engine, the stacking model improves the simulation capability of BECS to an automated and intelligent level through integrated learning and multi-objective optimization. The combination of the two realizes a paradigm shift from "experience-driven design" to "data-algorithm collaborative drive", which is a typical application of the deep integration of "bottom-level simulation tools" and "upper-level intelligent algorithms" in green building design.

feedforward neural network (FNN) and XGBoost model is adopted. FNN model can handle complex nonlinear relationships, while XGBoost model performs well in processing large-scale data and feature selection. Combining the two can not only take advantage of the deep learning ability of FNN, but also take advantage of the efficient feature extraction and generalization ability of XGBoost .

4.3 NSGA-II genetic algorithm multi-objective optimization

After the hybrid stacking model extracts EUI and thermal comfort PPD through ensemble learning, a multi-objective optimization algorithm can be applied to iterate the most effective building parameters.

(1) Initialize the population

According to the range and constraints of building parameters, a certain number of initial population individuals are randomly generated, each of which represents a combination of building parameters, namely, decision variables, such as building window-to-wall ratio, equipment system operation parameters, etc. These decision variables affect energy intensity and thermal comfort.

(2) Fitness function calculation

The hybrid stacking model constructed using FNN+XGBOOST calculates the energy consumption of each individual (constructs parameter combinations), and constructs the fitness function with EUI and PPD as two objective functions. Since these two conflicting objectives need to be optimized simultaneously (usually reducing energy consumption may have a certain impact on thermal comfort), it is necessary to normalize them according to actual needs and assign different weights to the two objectives, or use an unbiased method to optimize them as two independent objectives.

(3) Select an operation

Using selection methods such as tournaments, individuals with higher fitness are selected from the current population as parent individuals to generate new offspring individuals. The principle of tournament selection is to randomly select several individuals from the population for comparison, and finally select individuals with higher fitness as winners to enter the breeding pool.

(4) Crossover operation

Perform a crossover operation on the parent individuals to generate new offspring individuals. The crossover method can be single-point crossover, multi-point crossover or simulated binary tree crossover (SBX). The crossover operation generates offspring individuals with different building parameter combinations by exchanging some genes (building parameter values) of the parent individuals, thereby increasing the diversity of the population.

(5) Mutation operation

The offspring individuals are mutated to maintain the diversity of the population and prevent the algorithm from falling into the local optimum. The mutation operation can randomly change certain gene values of individuals, such as randomly adjusting the values of building parameters within a certain range, such as randomly perturbing the window-to-wall ratio within the interval [0.2, 0.6] .

(6) Non-dominated sorting and crowding distance calculation

The population formed by the merger of parent individuals and offspring individuals is non-dominated and orderly, and the individuals are divided into different levels.

This individual is a non-dominated individual, that is, it is not dominated by other individuals on both targets; the second-layer individuals after the first-layer individuals are non-dominated individuals, and so on.

The crowding distance of each individual is calculated to measure the distribution density of individuals in the target space. Individuals with large crowding distances are relatively sparsely distributed, which is conducive to maintaining the diversity of the population and allowing the algorithm to cover a wider area during the search process.

(7) Selecting a new population

The new generation of population individuals is selected according to the non-dominated sorting and crowding distance. First, individuals with a lower non-dominated level (such as the first level) are selected. If the population size is not enough, individuals with a larger crowding distance are selected from the second level until the population size reaches the set size.

(8) Iteration termination condition judgment

Repeat the above operations of selection, crossover, mutation, non-dominated sorting, etc. until the iteration termination condition is met. The termination condition can be reaching the preset maximum number of iterations, or the fitness change of the population tends to be stable in consecutive generations (that is, the Pareto frontier found does not change much).

5. Result Analysis and Verification

5.1 Pareto front extraction and analysis

Non-dominated individuals are extracted from the final population to form the Pareto front of energy intensity and thermal comfort. Each individual on the Pareto front represents a combination of building parameters that achieves the best balance between energy intensity and thermal comfort, that is, without deteriorating one objective, the other objective cannot be further improved.

By analyzing the individuals on the Pareto front, we can observe the impact trend of different building parameter changes on energy intensity and thermal comfort. For example, we can see that as the window-to-wall ratio increases, energy intensity may first decrease and then increase, while thermal comfort may increase within a certain range, but will decrease again after exceeding a certain critical point, which helps us understand the impact of building parameters.

5.2 Result verification and evaluation

Some building parameter combinations on the Pareto front are selected as eigenvalue subsets, and energy consumption simulation and thermal comfort evaluation are carried out in actual buildings or building simulation software (such as BECS software) to verify the consistency between the prediction results obtained by the hybrid superposition model and the NSGA-II algorithm and the actual situation.

By comparing the simulation results with the actual data, the accuracy of the model and optimization algorithm can be evaluated by calculating error indicators such as mean absolute error (MAE), root mean square error (RMSE), etc. If the error is large, it may be necessary to retrain the optimization model or adjust the algorithm parameters.

Through the above steps, the software can achieve the Pareto optimal solution of energy consumption intensity (EUI) and thermal comfort (PPD) on the London building dataset based on the NSGA-II genetic algorithm and the hybrid superposition model (FNN+XGB), providing a scientific and reasonable decision-making basis for building design and operation management, so as to achieve the goal of saving energy while ensuring indoor thermal comfort.

5.3 Construction of cross-scale verification system

A three-level verification mechanism of "simulated data-public data set-actual project" was established. Based on BECS, 100,000 sets of virtual data were generated to train the model, and the generalization ability of the model was verified using the data of 2,345 real buildings in the London building stock model, with the prediction error rate controlled within 10%. In a renovation project of an office building in London, energy consumption was reduced by 29% and the PPD compliance rate was increased to 92%.

5.4 Innovation in visualization for decision support

We have developed an interactive Pareto front analysis tool that assists designers in making quick decisions through dynamic two-dimensional graphs and parameter sensitivity analysis, shortening the solution decision time by more than 50%.

Conclusion

This study achieved efficient coordinated optimization of building energy consumption and comfort through technological innovation. The innovations are summarized as follows:

	Innovative content	Technological breakthrough	Empirical Effect
Methodology	Hybrid stacking model (FNN+XGB) combined with NSGA-II	Nonlinear relationship capture + multi-objective optimization	Prediction accuracy ↑22%, iteration efficiency ↑40%
Technical Framework	BIM-BECS-AI ternary collaborative architecture	Deep integration of physical simulation and data-driven	The error rate dropped from 18% to 9.2%.
Application Value	Dynamic Pareto Frontier Decision Tool	Visualization of multi-objective equilibrium analysis	Decision time↓50%
Verification System	Three-level data verification mechanism	Simulation-real data cross-scale verification	Project case energy consumption↓29%

Artificial intelligence models are becoming the core driving force of green building design through knowledge enhancement, multimodal fusion and dynamic iteration[2]. Empirical studies have shown that they have significant advantages in energy optimization accuracy and solution generation efficiency.

key technologies such as fine-grained spatial reasoning, real-time interaction, and ethical compliance, and combine digital twin and intelligent agent technologies to accelerate the construction industry's move towards zero-carbon goals .

Reference

[1] Zhou Hong et al. Research on BIM green building

optimization design, 20249.
[2] Application report of DeepSeek in building energy conservation, Zhejiang University. 202547.
[3] Hybrid stack energy optimization model. Energy and Buildings, 20256.
[4] https://data.london.gov.uk/dataset/london-building-stock-model-2
[5] https://data.london.gov.uk/dataset/london-building-stock-model