ENERGY ASSESSMENT OF URBAN BUILDINGS BASED ON GEOGRAPHIC INFORMATION SYSTEM

Wei Tian,[1–2*] Chuanqi Zhu,[1] Yunliang Liu[3], Baoquan Yin[4], Jiaxin Shi[1]

ABSTRACT

Urban building energy analysis has attracted more attention as the population living in cities increases as does the associated energy consumption in urban environments. This paper proposes a systematic bottom-up method to conduct energy analysis and assess energy saving potentials by combining dynamic engineering-based energy models, machine learning models, and global sensitivity analysis within the GIS (Geographic Information System) environment for large-scale urban buildings. This method includes five steps: database construction of building parameters, automation of creating building models at the GIS environment, construction of machine learning models for building energy assessment, sensitivity analysis for choosing energy saving measures, and GIS visual evaluation of energy saving schemes. Campus buildings in Tianjin (China) are used as a case study to demonstrate the application of the method proposed in this research. The results indicate that the method proposed here can provide reliable and fast analysis to evaluate the energy performance of urban buildings and determine effective energy saving measures to reduce energy consumption of urban buildings. Moreover, the GIS-based analysis is very useful to both create energy models of buildings and display energy analysis results for urban buildings.

KEYWORDS

urban buildings, energy model, machine learning model, Geographic Information System (GIS), sensitivity analysis

INTRODUCTION

Buildings in cities are responsible for around 70% of total primary energy use [1]. Hence, more recent research has focused on the energy analysis of urban buildings to deal with the increasing energy use and associated carbon emissions in urban environments [2–4]. Urban building energy modelling can be categorized into two types: top-down and bottom-up [5]. The top-down analysis investigates urban buildings from total energy use in a region and then

1. College of Mechanical Engineering, Tianjin University of Science and Technology, Tianjin, China, 300222 (*Corresponding email:tjtianjin@126.com)
2. Tianjin International Joint Research and Development Center of Low-Carbon Green Process Equipment, Tianjin 300222, China
3.School of Mechanical Engineering, Tongji University, Shanghai, China, 200092
4.Tianjin Architecture Design Institute, Tianjin, China, 300074

explores energy patterns at a smaller spatial scale. By contrast, the bottom-up method is from the archetype or individual buildings where the energy characteristics of these buildings are explored and then summed to the total region for energy consumption. The bottom-up approach can be further divided into statistical and engineering models for urban building energy analysis [6]. The statistical bottom-up approach explores urban energy use from energy bills or surveys and the associated explanatory variables, while engineering models create engineering-based building energy models by considering the detailed heat transfer occurring in buildings. Hence, the engineering-based bottom-up approaches at urban scale have the capability of simulating the detailed effects of various energy saving measures. Li et al. [7] developed an engineering-based, bottom-up energy model to simulate heating and cooling use in Chongqing, China. The results indicated that it is not feasible to implement central space heating similar to northern China, which would result in a significant increase in carbon dioxide in the residential sector of Chongqing. Tian et al. [8] describe an automated process of extracting information from a high-resolution database to create building energy models in a GIS (Geographic Information System) environment. They implemented this process to explore energy characteristics for the Westminster area of London. Hong et al. [9] developed a web-based computing environment to simulate energy use for city-scale buildings. This computing environment is named as the CityBES (City Building Energy Saver) based on the international open data standard, CityGML (City Geography Markup Language). Sun et al. [10] propose an integrated spatial analysis computing environment to explore the characteristics of energy performance of urban buildings. This computing environment has been used to assess energy patterns of London residential and non-residential sectors. Pasichnyi et al. [11] propose an urban building modelling approach based on rich datasets to develop various archetypes depending on the purposes of urban analysis. They apply this approach for two case studies: building retrofitting and electric heating in Stockholm, Sweden. Wei et al. [12] compare the predictive performance of six machine learning models in creating reliable and fast energy models for urban buildings. The SVM (support vector machine) models perform the best for predicting both gas and electricity use in London residential buildings. Chen and Hong [13] compare three zoning methods in urban-scale building energy modelling. The results indicate that the zoning methods have significant impact on energy use for urban building simulation. Hu et al. [14] developed an engineering-based bottom up model to predict the energy use of urban residential buildings in China. They found that centralized cooling is not a good choice for sustainable cooling in the urban residential sector. Tian et al. [15] implemented two global sensitivity analysis methods (standardized regression coefficients and multivariate adaptive regression splines) to determine the key variables influencing energy use for London school buildings. This previous research has made significant progress on the methods and applications for urban building energy analysis. However, there is a lack of research on a holistic engineering-based approach for urban building energy analysis by combining several emerging techniques, including GIS, machine learning, and global sensitivity analysis in an integrated computing environment. The combination of these techniques would provide a more convenient and robust computing environment for the energy evaluation of urban buildings by taking advantage of both engineering-based and machine learning energy models.

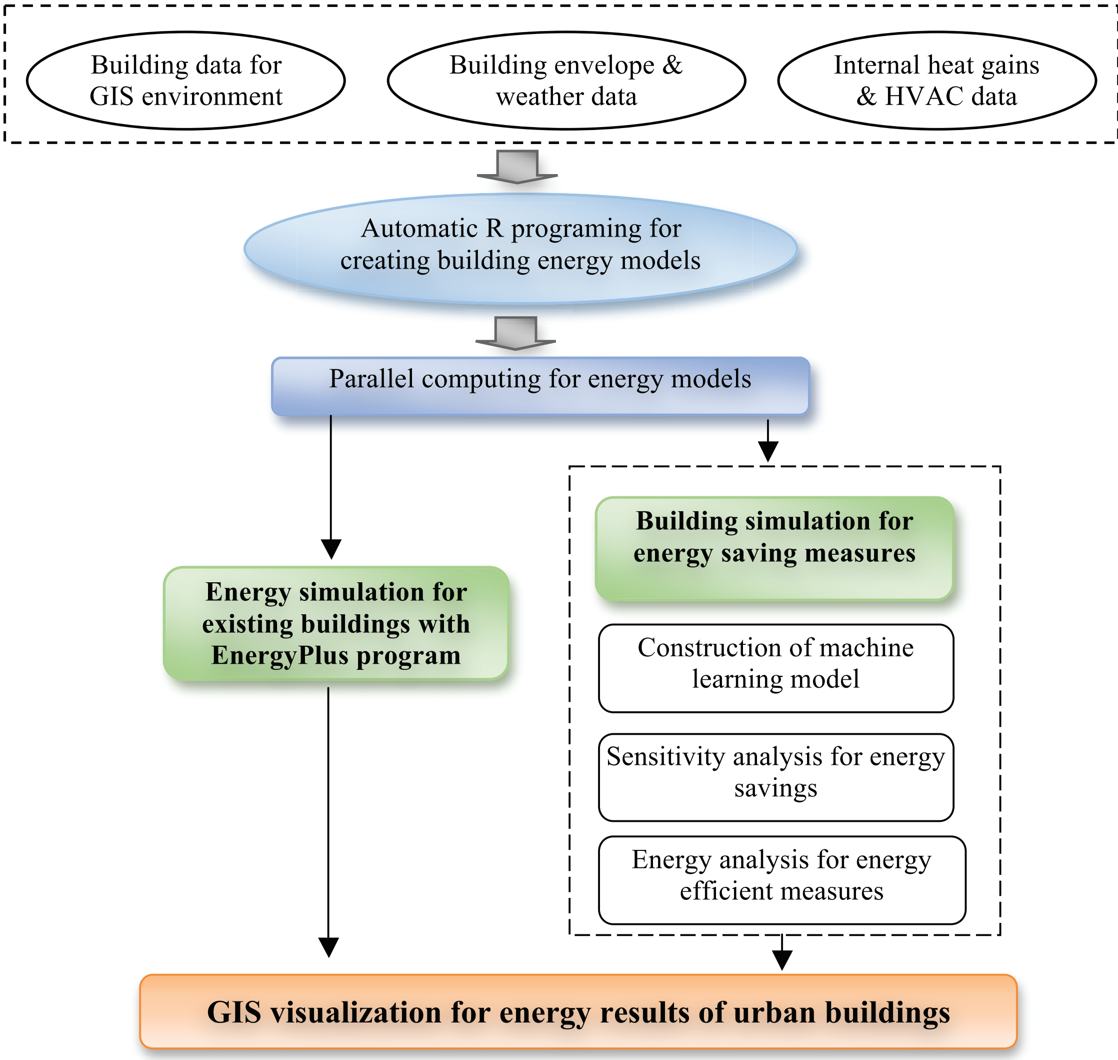
Therefore, this paper proposes a systematic engineering-based bottom-up method to conduct energy analysis of urban buildings in an integrated computing environment. The main features of this method include the automation of construction and energy models based on the GIS data, engineering-based dynamic energy simulation, machine learning techniques, sensitivity analysis, and GIS-based visualization of energy saving results. Energy analysis for urban environments usually involves a large number of building energy models, which requires the automation of creating energy models for these buildings. The GIS environment, which contains location and 2D coordinate data of buildings, can be used to create 3D building geometry for building energy models. The engineering-based dynamic energy models are usually required for urban building assessment to provide an accurate analysis for the effects of various energy saving measures. These engineering-based energy models, however, are computationally expensive for a large number of buildings. Hence, machine learning technique can create reliable and fast-running statistical energy models for various purposes of energy evaluation (for example, sensitivity, uncertainty or optimization analysis). The machine learning models in this research have two aims: global sensitivity analysis and energy saving computation. Global sensitivity analysis is an effective technique to determine energy saving measures in urban buildings by considering the nonlinear and interaction relationships between input and output variables [16]. After obtaining effective energy conservation measures, these learning models are used to compute energy saving results by implementing these energy saving measures. The GIS environment again is used to visualize the change of energy use in these urban buildings. The campus buildings in Tianjin are used to demonstrate the application of this method for urban building analysis.

METHODS

Figure 1 shows the flow diagram used in this research. The procedure of urban building analysis can be divided into five steps. The first step is to create a database for urban-scale building energy analysis, including weather data, GIS data, thermal performance of building envelopes, internal heat gains, and HVAC systems. The second step is to construct dynamic building energy models based on the database created in this first step. The automation of this step using computing languages is necessary since a large number of energy models are required at the urban-scale analysis. The third step is to create machine learning models for replacing time-consuming engineering-based energy models to assess the energy performance of urban buildings by conducting sensitivity analysis and computing energy savings. The fourth step is to run a sensitivity analysis based on the machine learning models created in the last step to determine effective energy saving measures to improve energy performance for urban environments. The fifth step is to display the results of the energy analysis computed based on the sensitivity analysis in the fourth step and learning models created in the third step for urban buildings in the GIS environment.

The main computing language used in this research is R environment [17], including the automation of creating building energy models, the construction of machine learning models, and the implementation of sensitivity analysis. The text function in R environment is used to create EnergyPlus models for building energy simulation based on the database constructed in the first step as shown in Figure 1. The EnergyPlus program [18], developed by the USA DOE (Department of Energy), has features to use ASCII text files as input data files (IDF) for building energy models, which can be easily edited using the R language. The R caret package [19] is used to create machine learning models in this study. The R sensitivity package [20] is used to conduct sensitivity analysis to choose effective energy saving measures.

FIGURE 1. Flow diagram of the research used in this paper.



CONSTRUCTION OF BUILDING ENERGY MODELS BASED ON GIS ENVIRONMENT

This section describes the first three steps as discussed in the section "Methods," including the construction of a database for energy analysis of urban buildings, the automation of creating dynamic energy models, and the construction of machine learning models.

Database for urban-scale building energy analysis

The database used in this research is named as the Dynamic Building Energy Model Input Parameter Set (DBEMIPS) for urban buildings. As shown in Figure 1, there are three types of data: data for GIS environment, building envelope & weather data, and internal heat gains & HVAC data. The first type of data is mainly used for creating building geometry for dynamic energy models. There is no GIS data available for the campus buildings investigated in this study. Hence, the GIS data is created based on the publicly available data on the Baidu Map using the geo-referencing function to specify target coordinates in a raster image. The second type of data is related to the thermal performance of a building envelope and weather data. The typical data here include wall U-value, roof U-value, floor U-value, window U-value and window SHGC (solar heat gain coefficient). Most campus buildings studied were constructed before the introduction of the energy saving code released in 2005 in China. As a result, the typical value for these parameters are chosen based on the recommendations from the appendix of China building energy code [21]. The weather data depends on the availability of measured weather data. If there were no 8760 measured hourly data, typical meteorological data can be used to assess the typical energy performance of urban buildings. The third type of data includes internal heat gains and HVAC data, which are difficult to obtain for urban-scale buildings. Hence, the typical data set from previous studies, building standard, and research reports, are used [21–24]. These schedules for summer and winter holidays have been adjusted based on the university calendar.

Automation of creating urban-scale building energy models

This step is a key procedure for this urban-scale building energy analysis to automatically create a large number of building energy models. The coordinates of polygons at the GIS environment are exported with the well-known text format in which the data is 2D coordinates. Then these coordinates and building height are edited in the csv format to be imported in the R environment to create 3D building energy models of IDF files with typical zoning (four external and one internal zones) for building energy models. The depth of perimeter zones is 4.5 meters when zoning these campus buildings. The whole procedure in this step is written by using the R language. The R codes are updated from previous research [8] with only one zone per floor to create five zones per floor. The global geometry rules of buildings are in line with the EnergyPlus requirements. The starting corner for the surface is the lower left corner and the vertex energy direction is in a counterclockwise direction in viewing the specific surface from the outside of the given zone [18]. Figure 2 illustrates the 3D EnergyPlus energy models (68 in this case study) created using the R language. Figure 3 compares the geometry of the main building between the actual building and energy model.

FIGURE 2. Energy models of campus buildings.

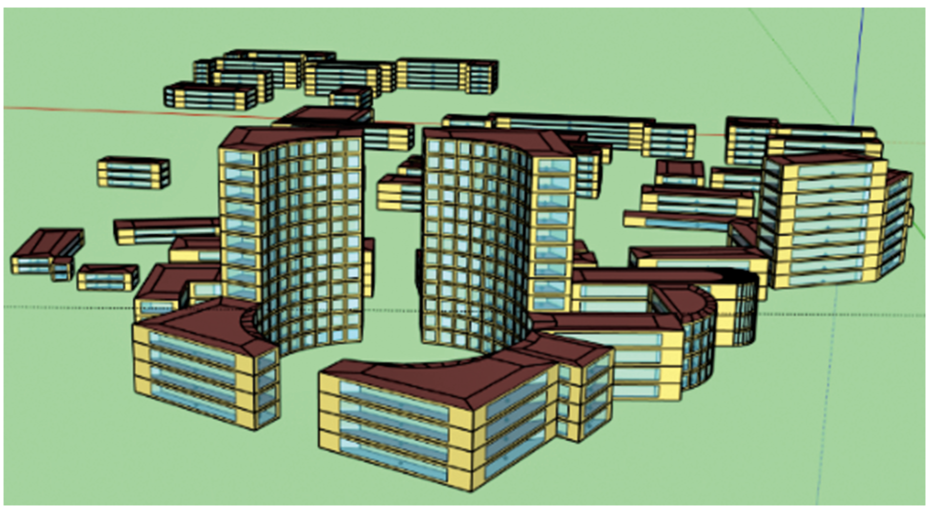
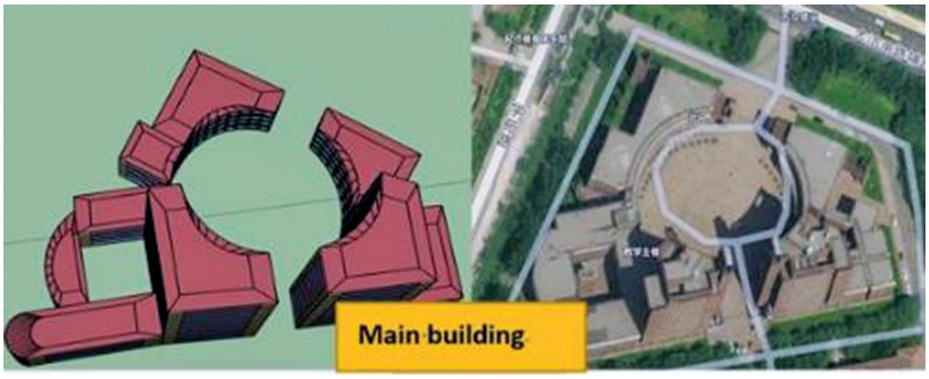


FIGURE 3. Spot check of building shapes for main building.



Construction of machine learning models for building energy analysis

Energy analysis of urban buildings often involves a large number of buildings, which requires high computational cost. To expedite the computation of energy analysis, it is necessary to create fast running energy models instead of using engineering-based models [15, 25]. Five machine learning models are used here by considering the balance between model accuracy and interpretability, including MARS (multivariate adaptive regression splines), BMARS (bagging MARS), RF (random forest), SVM (support vector machine), and TGP (treed Gaussian process). Detailed information on these learning methods is available [26, 27]. The eight input parameters are considered to represent the variations of input parameters, including wall U-value (WU), ground U-value (GU), roof U-value (RU), window U-value (DU), solar heat gain coefficient of windows (SH), infiltration rate (FI), lighting power density (LH), and equipment power density (EP). The change of the values for these eight parameters is also regarded as a potential energy conservation measure when refurbishing these campus buildings. For every building, there are 100 combinations of input parameters using Latin hyper-cube sampling [28], which means 100 EnergyPlus models would be run for every building. As a result, there are 6800 EnergyPlus models since there are 68 buildings in this research. The ten-fold cross validation is used to tune the parameters in these learning models by considering both bias and variance [27]. To further validate the results from these five machine learning models, extra 6800 EnergyPlus models are created and run to compare the predictive capability of these machine learning models. The comparison results are summarized in Table 1. In terms of model accuracy, the first two models

TABLE 1. Performance comparison from five machine learning models.

Model	R^2	RMSE
MARS (multivariate adaptive regression splines)	0.998	2.081
BMARS (bagging MARS)	0.998	2.004
RF (random forest)	0.996	3.523
SVM (support vector machine)	0.956	13.794
TGP (treed Gaussian process)	0.923	15.979

are MARS and BMARS with the high R^2 (coefficient of determination) and the low RMSE (root mean square error) [19]. However, the computational cost for the BMARS model is high since a large number of regression models should be constructed for the ensemble learning models. Hence, the MARS models have been selected in this research for further sensitivity analysis and energy assessment.

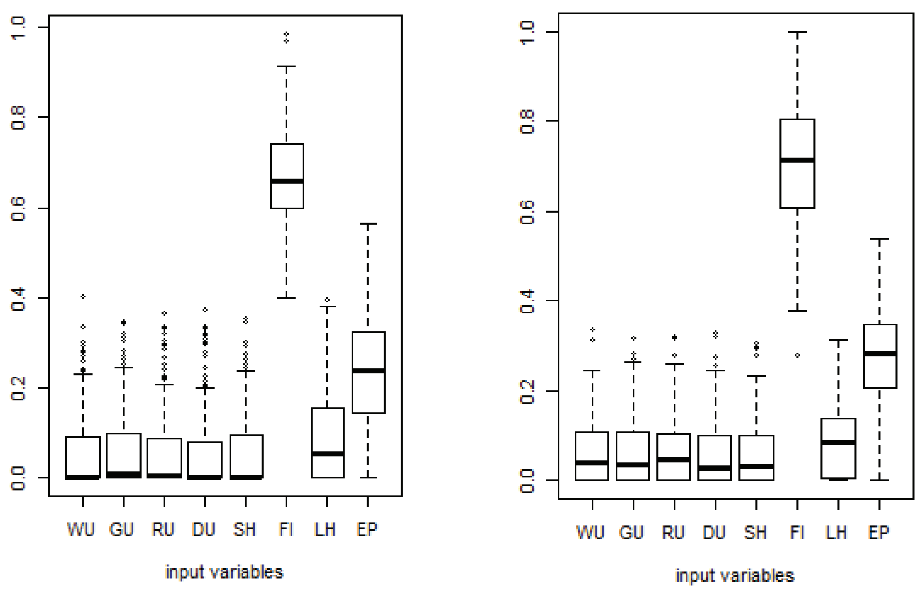
VISUALIZATION OF ENERGY SAVING RESULTS

This section presents the results from the fourth and fifth steps for sensitivity analysis and GIS-based visualization, respectively.

Results for sensitivity analysis

Sensitivity analysis can be used to prioritize energy saving measures in buildings [29]. The meta-modelling sensitivity method is a two-step method to create the machine learning models based on the data set firstly and then use this learning model to conduct variance-based sensitivity analysis [30]. Hence, the commonly used two sensitivity indicators from variance-based methods are also available to consider the influence of single variable (named as main effect) and include the second-order or higher-order interactions with other variables (named as total effect) [31]. The results of sensitivity analysis of annual heating energy for office buildings are shown in Figure 4. The first two important factors are infiltration rate (FI) and equipment heat gains (EP), which account for over 90% of output variations. The remaining factors have only marginal effects on energy use for this building. Hence, more attention should be paid on these two factors to improve energy performance of office buildings in this campus. More sensitivity results on other types of buildings are available in Liu [22]. Based on these sensitivity analysis

FIGURE 4. Global sensitivity analysis results of annual heating energy use in an office building of this campus.



(a) Main effect (c) Total effect

results and previous studies [32, 33], four energy saving measures are determined: (1) 20% reduction of equipment power density; (2) 20% reduction of lighting power density; (3) install daylight sensors; (4) 40% reduction of infiltration rate. Most of these campus buildings are old and leaky, which leads to high infiltration rates through poorly sealed window panes & doors. Hence, these joints should be upgraded to increase air tightness of these old buildings in order to reduce heating energy use. New lightings or electric equipment usually have high energy efficiency standards, and then it is recommended to replace older teaching and office equipment in these buildings, which can effectively reduce electricity use. Moreover, the daylight sensors by dimming the lighting can be installed to further reduce electricity use.

Visualization of GIS-based results

Based on the results of sensitivity analysis in the last subsection, energy saving results are computed using the machine learning models (MARS in this case study) for all the 68 buildings as shown in Figure 5 and Figure 6. Note that the height of buildings in these figures are not actual height, but energy use intensity. The energy use intensity for heating and cooling energy use has been reduced significantly (around 20% in this case study) as can be seen from the change of color and building height in Figure 5 and Figure 6. The infiltration rate in this case study is the only dominant factor affecting heating energy use, while the equipment and lighting heat gains are two important factors influencing cooling energy use. Hence, the improvement of air tightness and implementation of more efficient equipment or lighting in these old buildings would markedly improve energy performance for this campus.

The method proposed here can be applied to different spatial scales depending on the availability of urban data. If the building number is below hundreds of buildings, it is easy to

FIGURE 5. Comparison of annual heating energy (kWh/m^2) between existing buildings and energy saving measures for campus buildings.

(a) Existing buildings

(b) Energy saving measures

FIGURE 6. Comparison of annual cooling energy (kWh/m^2) between existing buildings and energy saving measures for campus buildings.

(a) Existing buildings

(b) Energy saving measures

manipulate the data, especially the construction of GIS data. Difficulties arise, however, when an attempt is made to implement this method for a larger spatial scale without the GIS data available. One solution is to simplify building geometry in the GIS environment, which can significantly expedite the process of construction of GIS data [8]. For a larger scale, it usually means the explanatory data becomes increasingly unavailable. Bayesian computation can be used to make full use of prior knowledge of building parameters at urban environments to infer unknown parameters. Another important issue is to properly consider uncertainty of energy use for the urban environment, which usually requires thousands or higher simulation runs for building energy models. The method proposed here is suitable to conduct uncertainty analysis of urban buildings using the machine learning models created in the process as illustrated in Figure 1. To this end, the R statistical environment used in this research have a number of packages (for example, EasyABC [34] and BRugs [35]) that can be used for uncertainty analysis and Bayesian model calibration.

CONCLUSIONS

This paper proposes a holistic method to assess the energy performance of urban buildings by combining engineering-based energy models, machine learning energy models, global sensitivity analysis, and data visualization in a GIS environment. The GIS environment can be used to directly create engineering-based EnergyPlus models for a large number of buildings for the urban environment. The machine learning models can provide fast and reliable energy prediction by replacing computationally expensive EnergyPlus models for urban buildings. The global sensitivity analysis can identify effective energy saving measures for these urban buildings. Campus buildings in Tianjin are used to demonstrate the application of this method. The results indicate that this GIS-based analysis method can effectively provide the characteristics of energy performance for urban buildings and also prioritize energy saving measures. This paper investigates the urban area at the community levels, up to hundreds of buildings. For a larger spatial scale, the method proposed here is still applicable although the data collection for buildings require much more effort.

ACKNOWLEDGEMENT

This research was supported by the National Natural Science Foundation of China (No. 51778416) and the Key Projects of Philosophy and Social Sciences Research, Ministry of Education (China) "Research on Green Design in Sustainable Development" (contract No. 16JZDH014, approval No. 16JZD014).

REFERENCES

[1] T. Hong, J. Langevin, K. Sun, Building simulation: Ten challenges, *Building Simulation*, 11 (2018) 871–898.

[2] W. Tian, Y. Liu, Y. Heo, D. Yan, Z. Li, J. An, S. Yang, Relative importance of factors influencing building energy in urban environment, *Energy*, 111 (2016) 237–250.

[3] W. Li, Y. Zhou, K. Cetin, J. Eom, Y. Wang, G. Chen, X. Zhang, Modeling urban building energy use: A review of modeling approaches and procedures, *Energy*, 141 (2017) 2445–2457.

[4] E. Zalamea-León, J. Mena-Campos, A. Barragán-Escandón, D. Parra-González, P. Méndez-Santos, Urban photovoltaic potential of inclined roofing for buildings in heritage centers in equatorial areas, *Journal of Green Building*, 13 (2018) 45–69.

[5] C.F. Reinhart, C. Cerezo Davila, Urban building energy modeling—A review of a nascent field, *Building and Environment*, 97 (2016) 196–202.

[6] W. Tian, R. Choudhary, *Energy use of buildings at urban scale: a case study of London school buildings*, in: IBPSA 2011, November 14–16, Sydney, 2011.

[7] X. Li, R. Yao, W. Yu, X. Meng, M. Liu, A. Short, B. Li, Low carbon heating and cooling of residential buildings in cities in the hot summer and cold winter zone—A bottom-up engineering stock modeling approach, *Journal of Cleaner Production*, 220 (2019) 271–288.

[8] W. Tian, A. Rysanek, R. Choudhary, Y. Heo, *High Resolution Energy Simulations at City Scale*, in: IBPSA Building Simulation Conference (BS2015), December 7–9, 2015, Hyderabad, India., 2015.

[9] T. Hong, Y. Chen, S.H. Lee, M.A. Piette, *CityBES: A web-based platform to support city-scale building energy efficiency*, in: 5th International Urban Computing Workshop, San Francisco, USA, August, 2016.

[10] Y. Sun, E. Silva, W. Tian, R. Choudhary, H. Leng, An Integrated Spatial Analysis Computer Environment for Urban-Building Energy in Cities, *Sustainability*, 10 (2018) 4235.

[11] O. Pasichnyi, J. Wallin, O. Kordas, Data-driven building archetypes for urban building energy modelling, *Energy*, 181 (2019) 360–377.

[12] L. Wei, W. Tian, E.A. Silva, R. Choudhary, Q. Meng, S. Yang, Comparative Study on Machine Learning for Urban Building Energy Analysis, *Procedia Engineering*, 121 (2015) 285–292.

[13] Y. Chen, T. Hong, Impacts of building geometry modeling methods on the simulation results of urban building energy models, *Applied Energy*, 215 (2018) 717–735.

[14] S. Hu, D. Yan, M. Qian, Using bottom-up model to analyze cooling energy consumption in China's urban residential building, *Energy and Buildings*, (2019) 109352.

[15] W. Tian, R. Choudhary, A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London, *Energy and Buildings*, 54 (2012) 1–11.

[16] A. Mastrucci, P. Pérez-López, E. Benetto, U. Leopold, I. Blanc, Global sensitivity analysis as a support for the generation of simplified building stock energy models, *Energy and Buildings*, 149 (2017) 368–383.

[17] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. accessed 2018-12-10, in, 2017.

[18] DOE, EnergyPlus V9.0.1, October 2018, Department of Energy, USA, in, 2018.

[19] M. Kuhn, K. Johnson, *Applied predictive modeling*, Springer, 2013.

[20] B.I. Gilles Pujol, Alexandre Janon, R package sensitivity V1.16.1: sensitivity: Global Sensitivity Analysis of Model Outputs. https://CRAN.R-project.org/package=sensitivity accessed on 2019-07-07, in, 2018.
[21] MOC, GB50189-2005. *Energy Conservation Design Regulation for Public Buildings*. Ministry of Construction (MOC) of P.R.China China Planning Press (2005) (in Chinese), 2005.
[22] Y. Liu, *Energy Saving of Urban Buildings Based on 3D Geographic Information System*, in, Tianjin University of Science and Technology, Tianjin, China, 2018.
[23] MOC, GB50189-2015. *Design standard for energy efficiency of public buildings*, Ministry of Construction (MOC) of P.R.China China Planning Press (2015) (in Chinese), 2015.
[24] W. Tian, S. Yang, J. Zuo, Z. Li, Y. Liu, Relationship between built form and energy performance of office buildings in a severe cold Chinese region, *Building Simulation*, 10 (2017) 11–24.
[25] S. Das, A. Swetapadma, C. Panigrahi, A study on the application of artificial intelligence techniques for predicting the heating and cooling loads of buildings, *Journal of Green Building*, 14 (2019) 115–128.
[26] K.P. Murphy, *Machine learning: a probabilistic perspective*, MIT press, 2012.
[27] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2009.
[28] W. Tian, Y. Heo, P. de Wilde, Z. Li, D. Yan, C.S. Park, X. Feng, G. Augenbroe, A review of uncertainty analysis in building energy assessment, *Renewable and Sustainable Energy Reviews*, 93 (2018) 285–301.
[29] H. Li, S. Wang, H. Cheung, Sensitivity analysis of design parameters and optimal design for zero/low energy buildings in subtropical regions, *Applied Energy*, 228 (2018) 1280–1291.
[30] W. Tian, A review of sensitivity analysis methods in building energy analysis, *Renewable and Sustainable Energy Reviews*, 20 (2013) 411–419.
[31] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, *Global sensitivity analysis: the primer*, Wiley-Interscience, 2008.
[32] W. Tian, P.d. Wilde, Z. Li, J. Song, B. Yin, Uncertainty and Sensitivity Analysis of Energy Assessment for Office Buildings based on Dempster-Shafer Theory, *Energy Conversion and Management*, 174 (2018) 705–718.
[33] J. Song, L. Wei, Y. Sun, W. Tian, Implementation of Meta-modelling for Sensitivity Analysis in Building Energy Analysis, in: The eSim 2014 conference, May 7 to 10, Ottawa, Canada, 2014.
[34] F. Jabot, T. Faure, N. Dumoulin, EasyABC: performing efficient approximate Bayesian computation sampling schemes using R, *Methods in Ecology and Evolution*, 4 (2013) 684–687.
[35] A. Thomas, B. O'Hara, U. Ligges, S. Sturtz, Making BUGS open, *R news*, 6 (2006) 12–17.