ISSN 0003-701X, Applied Solar Energy, 2022, Vol. 58, No. 2, pp. 291–296.

Application of the Clustering Method to the Determination of Typical Days of Thermal Loads of Buildings in Uzbekistan

A. S. Halimov[a], *, J. S. Akhatov[a], Z. S. Iskandarov[b], and N. M. Nazarova[c]

[a] *Physical–Technical Institute NGO Physics–Sun, Academy of Sciences of the Republic of Uzbekistan, Tashkent, 100084 Uzbekistan*

[b] *Tashkent State Agrarian University, Kibray region, Tashkent oblast, 100140 Uzbekistan*

[c] *Bukhara State University, Bukhara, 705018 Uzbekistan*

**e-mail: akbar.halimov@rwth-aachen.de*

Received April 11, 2022revised May 5, 2022accepted June 1, 2022

Abstract—Optimal design, sizing, and operation of building energy systems is challenging due to the variety of generation and storage devices available and the high-resolution input data needed to account for seasonal and diurnal fluctuations in thermal loads of buildings. A common measure to reduce the size and complexity of a problem is to group requirements into representative periods. In this study, in order to simplify the problem of optimizing building envelopes and integrating various energy generators operating on renewable energy sources on an annual scale, a clustering method of *k*-means of hourly thermal load of a building is proposed. In this study, for the first time, the typical days of thermal loads for heating and cooling a building are determined with the optimal planning of one or another reconstruction measure. For further research, there is a new opportunity to identify typical days of thermal demand in order to determine the thermal performance of buildings and introduce new measures for energy planning reconstruction in them in 13 regions of Uzbekistan with different levels of thermal insulation and integration of various energy generators operating on renewable energy sources.

Keywords: clustering, *k*-means, thermal load, buildings Case 600 ASHRAE 140, model validation

DOI: 10.3103/S0003701X22020086

INTRODUCTION

By 2050, the global energy demand is expected to double [1]. The energy balance of the building sector is a major consumer of energy, which is mainly due to the large heat loads for cooling and heating in order to achieve optimal thermal performance and an acceptable level of thermal comfort in buildings. However, the construction sector is considered one of the most attractive sectors of the economy, where there is an opportunity to reduce energy consumption due to progress and the use of energy efficient technologies and building materials [2]. Moreover, the sustainable development strategy considers that the key trend is not only to increase the share of renewable energy sources (RESs) in this sector, but also to reconstruct inefficient existing buildings [3] with the introduction of passive solar systems in buildings [4].

At present, in Uzbekistan, the specific energy consumption in buildings varies from 150 to 690 kWh/m^2 [5, 6] depending on the level of thermal protection. However, this indicator, in turn, is 2–3 times higher than in the developed countries of the world [7]. It should be noted that in Uzbekistan there is an opportunity to reduce heat loads for heating residential buildings using RESs (the technical potential is 179 million toe) [8]. In particular, the introduction of passive solar heating systems in buildings [9–14] and optimal design can reduce the heat load for cooling by up to 54% and for heating by up to 87% in buildings [15, 16].

It should be noted that the optimal design of reconstruction is a complex task, which includes thermal loads in buildings and energy generated by various generators. In addition, there is a relationship between the energy supply systems of buildings and heat consumption, which among other factors depends on passive building measures [3]. According to the analyses of the authors [17], the optimal design of the energy reconstruction of residential buildings was performed with 40% of the studies being aimed at optimal designs for building envelopes, while only 20% of them were aimed at optimizing the geometric shape of buildings.

Most studies that optimize building energy systems use precise optimization algorithms such as mixed integer linear programming (MILP) [18]. For example, in the studies of the authors [19, 20], the size and operation of the energy supply system of buildings were optimized based on MILP models, while in studies [21–24] the combined optimal design and operation of the energy supply system were considered, and this approach was applied on an urban scale [25–28]. In these studies, heat loads in buildings are considered to be fixed time series, which are calculated before optimization and are constant during optimization.

Several authors [29–35] carried out a simultaneous multi-purpose optimization of building envelopes and the energy supply system of buildings. Asadi et al. [30], proposed an even more simplified static model based on ISO 13790, which uses heating period degree days (HPDDs). Based on this method, Fang and Xia [35] optimized the building envelope and the areas of photovoltaic panels installed on the roofs of buildings. In the previous study of the authors [36], a multi-purpose optimization of typical four-room rural residential buildings in all regions of Uzbekistan with different levels of thermal insulation and integration of various energy supply systems based on RESs (photovoltaic panels and solar collectors) and installed on the roofs of buildings was carried out for the first time. For this purpose, a simplified static building model based on HPDDs was used. However, Schütz et al. [34], noted the need to use a non-stationary thermal model of buildings when conducting multi-criteria optimization and proposed a more accurate model for optimizing building envelopes and building energy systems.

In this study, in order to simplify the problem of optimizing building envelopes and integrating various energy generators operating on RESs on an annual scale, a clustering method of *k*-means of the hourly heat load of a building is proposed. The *k*-means clustering algorithm is quite simple to implement and also very computationally efficient compared to other clustering algorithms, which may explain its popularity. The *k*-means algorithm belongs to the category of prototype-based clustering. Prototype-based clustering means that each cluster is represented by a prototype, which can be either a centroid (mean) of similar points with continuous features, or a medoid (most representative or most frequent point) in the case of categorical features. As a result, we have determined the characteristic (typical) days of the thermal loads of a building.

METHODS

Fundamentals of the k-means Method

The *k*-means algorithm is extremely easy to implement and also very computationally efficient compared to other clustering algorithms, which may explain its popularity. The *k*-means algorithm belongs to the category of prototype-based clustering. Prototype-based clustering means that each cluster is represented by a prototype, which can be either a centroid (mean) of similar points with continuous features, or a medoid (most representative or most frequent point) in the case of categorical features. Although the *k*-means method is effective for identifying spherical clusters, one disadvantage of this clustering algorithm is that number of clusters *k* must be specified in advance. Choosing the wrong value for *k* can lead to poor clustering performance.

Description of the Method

Given a set of observations (x_1, x_2, ..., x_n), where each observation is a *d*-dimensional real vector, *k*-means clustering aims to divide *n* observations into *k* ($\leq n$) sets $S = \{S_1, S_2, ..., S_k\}$ to minimize the total square deviation. The idea is to group the samples based on the similarity of their features, which can be achieved using the *k*-means algorithm that can be summarized in the following four steps:

• *k*-centroids are randomly selected from the sampling points as the initial centers of the clusters.

• Each sample is compared to the nearest centroid $\mu(j), j \in \{1, ..., k\}$.

• The centroids are moved to the center of the samples that were assigned to the center.

• Steps 2 and 3 are repeated until the cluster assignments change or a user-specified tolerance or maximum number of iterations is reached.

Now the similarity between objects is measured. For this purpose, the square of the Euclidean distance between two points *x* and *y* in *m*-dimensional space is determined:

$$d(x, y)^2 - \sum_{j=1}^{m} (x_i - y_i)^2 = x - y_2^2. \quad (1)$$

It should be noted that in this equation index *j* refers to the *j*-th dimension (column of features) of the sample points *x* and *y*. We will use superscripts *i* and *j* to denote the sample index and cluster index, respectively. Based on this Euclidean distance metric, one can describe the *k*-means algorithm as a simple optimization problem, an iterative approach to minimizing the sum squared error (SSE), which is sometimes also called cluster inertia:

$$\mathrm{SSE} = \sum_{i=1}^{n}\sum_{j=1}^{k} x^{(i)} - \mu^{(j)2}, \quad (2)$$

where $x^{(i)}$ is the data set and $\mu^{(j)}$ is the centroid of the data set. It should be noted that when the *k*-means method is applied to real data using the Euclidean distance metric, one must ensure that the datasets are measured in the same units.

Application Example: ASHRAE 140, CASE 600

As is known, Case 600 of the ASHRAE 140 [38] standard is a simple and well-established test case of a building, which has been verified by several building thermal efficiency modeling tools. Given this circumstance, we studied a modified version of Case 600. In addition, a comprehensive Case 600 package of the

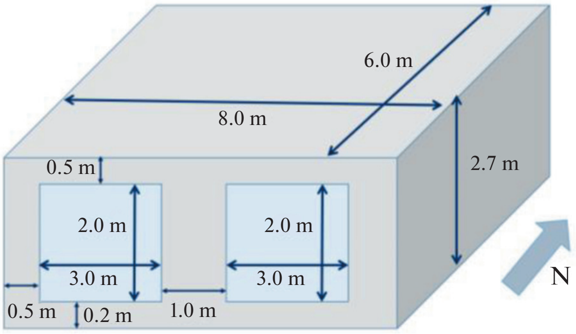

Fig. 1. Isometric south windows without shutters for Case 600 ASHRAE Standard 140.

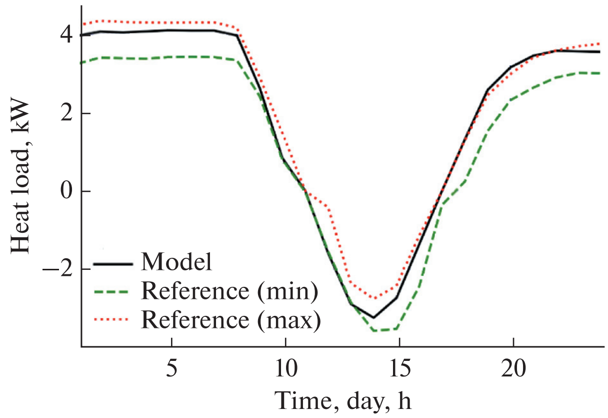

Fig. 2. Hourly heat load (heating and cooling load) of case 600 buildings on January 4 in Denver climate conditions (39.8° N, 104.9° W), Colorado, United States.

ASHRAE 140 standard developed in Modelica is available in [37, 38].

The basic test case is a light rectangular single-zone building with dimensions of 8 × 6 × 2.7 m (Fig. 1). The building has no internal partitions, has a total window area of 12 m^2 on the south wall, an internal load of 200 W (60% radiant load, 40% convective load) and a highly insulated slab used to significantly eliminate thermal coupling to the ground. The infiltration was set to 0.5 air changes per hour. The building's mechanical system is an ideal system with a 100% convective air subsystem and 100% efficiency with no duct loss and no capacity limitation. The thermostat is set to a dead zone, so heating occurs at temperatures below 20°C, and cooling occurs at temperatures above 27°C.

Calibrating the CASE 600 Model

ASHRAE Standard 140 provides eight benchmark results from the building thermal efficiency modeling tools such as DOE2, TRNSYS, and ESP-r in Denver (39.8° N, 104.9° W) climates, Colorado, United States. We calibrate the Case 600 base model against reference datasets. To this end, we quantify the deviation between the hourly heating and cooling loads (Fig. 2) from the base tested Case 600 model and the reference data for January 4 using coefficients of determination and standard deviations (RMS) (Table 1).

Here, the average value of the hourly heating and cooling load from eight ASHRAE Standard 140 reference data is used as the observed values to calculate the coefficient of determination and RMS. In Table 1, the minimum and maximum ranges of the reference data correspond to the lower and upper limits of the hourly peak heating/cooling and annual loads of the reference data. The mean value indicates the mean value of

Table 1. Determination coefficient R^2 and standard deviation for hourly and annual heating and cooling loads

Calibration criteria	Unit	Model	Reference		
			minimum value	maximum value	average value
Test for January 4					
Peak cooling capacity (–)	kW	3.2445	2.76	3.58	3.17
Deviation*	%	–	17.55	9.37	2.35
Peak heating load (+)	kW	4.1074	3.43	4.35	4
Deviation*	%	–	19.75	5.57	2.68
RMS	kW	0.1873			
R^2	–	0.9996			
One full year testing					
Annual cooling capacity (–)	kWh	6.735	6.137	7.964	6.832
Deviation*	%	–	8.87	18.24	1.44
Annual heating load (+)	kWh	4.971	4.296	5.709	5.09
Deviation*	%	–	13.57	14.84	2.39

*100% × |1 – Reference/model|

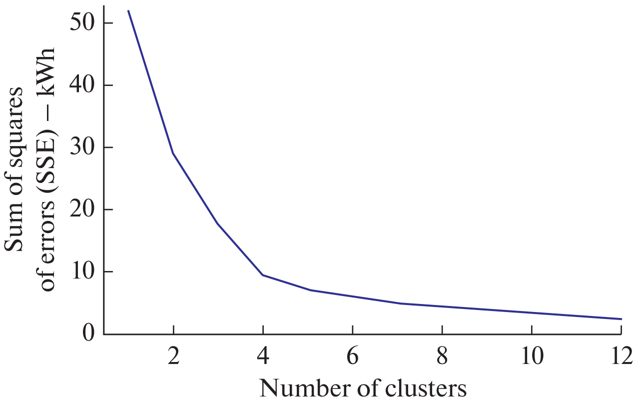

Fig. 3. Dependence of the sum of square errors of thermal loads of the Case 600 building in Tashkent for the entire year on the number of clusters.

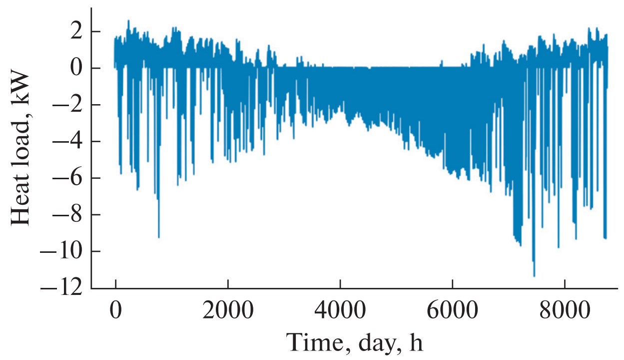

Fig. 4. Hourly heat load (heating and cooling load) of case 600 buildings in Tashkent for the entire year.

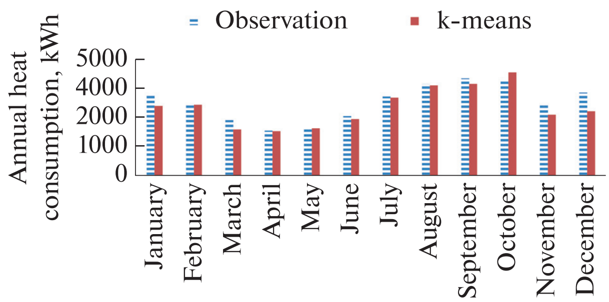

Fig. 5. Comparison of the results of monthly thermal loads and thermal loads, which were obtained by *k*-means clustering.

the corresponding values from the reference data (Fig. 2). According to the accuracy assessment method we used, the coefficient of determination and RMS for the created Case 600 model are 0.9996 and 0.1873 kW, respectively. Peak values for hourly heat loads and annual energy loads are in the minimum and maximum ranges of the reference data, while they are completely close to the average value of the reference data. Now the model is calibrated and can be used for comparative testing by changing the climate data of a particular region.

Climate Data Modification

As noted above, the proposed building model is carefully calibrated and can be used for comparative analyzes by changing the climatic data of other regions. However, it should be noted that climate data must be presented in the typical meteorological year (TMY) format with hourly time steps for the entire climate year. In this work, we studied the city of Tashkent in Uzbekistan with a continental climate with hot summers and Mediterranean influence—Dsa, according to climatic conditions, based on the Köppen—Geiger classification [18].

RESULTS AND DISCUSSION

Previously, it was pointed out that one of the disadvantages of the *k*-means clustering method was the need to specify the number of clusters (characteristic days) in advance, since the wrong choice of the number of clusters can lead to poor clustering performance. In order to determine the number of clusters, an attempt was made to find the dependence of the sum of square errors on the number of clusters.

Figure 3 shows the dependence of the sum of square errors on the number of clusters obtained by iteration. The idea of setting the number of clusters made it possible to determine the characteristic day that corresponds to the characteristic heat load of a building in each month. As can be seen from Fig. 3, it is enough to have 12 characteristic days instead of 365 days to characterize the annual heat load of a building in the climatic conditions of Tashkent. The hourly heat load of a building, which was obtained without using the *k*-means clustering method, is shown in Fig. 4.

Figure 5 shows that the annual heat load in case of using the *k*-means clustering method differs from the heat loads obtained taking into account 8760 points by only 6%. It should be noted that only 288 points were used for clustering.

Thus, the characteristic days of thermal loads in buildings for the climatic conditions of Tashkent are January 22, February 6, March 2, April 11, May 15, June 20, July 22, August 14, September 22, October 31, November 3, and December 4.

Figure 6 shows the dependence of global horizontal solar radiation on ambient temperature, where only 288 characteristic points can be used instead of 8760 points thanks to the *k*-means clustering method.

CONCLUSIONS

In this study, in order to simplify the problem of optimizing building envelopes and integrating various

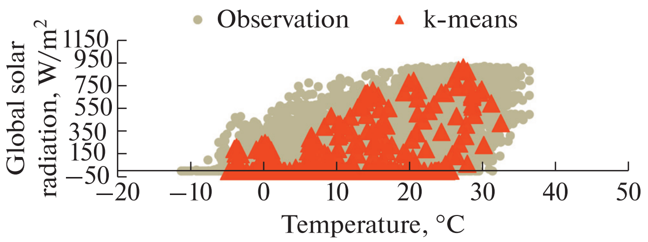

Fig. 6. Comparison of hourly characteristic days (*k*-means) and all days of global solar radiation and ambient temperature.

energy generators operating on RESs on an annual scale, the method for *k*-means clustering of the hourly heat load of a building is proposed. The *k*-means clustering algorithm is quite simple to implement and also very computationally efficient compared to other clustering algorithms. Optimal design, sizing, and operation of building energy systems is challenging due to the variety of devices available and high-resolution input data required to account for seasonal and diurnal fluctuations in building heat loads. As a rule, a measure to reduce the size and complexity of the problem is the use of grouping requirements for representative periods.

In this study, the typical days of thermal loads for heating and cooling a building were for the first time determined with the optimal planning of one or another measure for its reconstruction. For further research, a new opportunity opens to determine the typical days of the year in order to determine the thermal performance of buildings and introduce new measures for energy planning reconstruction in other regions of Uzbekistan with different levels of thermal insulation and integration of various energy generators operating on RESs.

ACKNOWLEDGMENTS

We thank colleagues from the Physical—Technical Institute NGO Physics—Sun, Academy of Sciences of the Republic of Uzbekistan for their constant support.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest. The authors declare that they do not have a conflict of interest.

REFERENCES

1. IEA World Energy Balances 2018. https://webstore.iea.org/world-energy-balances-2018 Accessed April 27, 2022.
2. Abbassi, F., Dimassi, N., and Dehmani, L., Energetic study of a Trombe wall system under different Tunisian building configurations, *Energy Build.,* 2014, vol. 80, pp. 302–308.
3. Kaklauskas, A., Zavadskas, E.K., and Raslanas, S., Multivariant design and multiple criteria analysis of building refurbishments, *Energy Build.,* 2005, vol. 37, pp. 361–372.
4. Hu, Z., He, W., Ji, J., and Zhang, Sh., A review on the application of Trombe wall system in buildings, *Renewable Sustainable Energy Rev.,* 2017, vol. 70. pp. 976–987.
5. *PROON—Ministerstvo ekonomiki. K ustoichivoi energii: Strategiya nizko-uglerodnogo razvitiya Respubliki Uzbekistan* (UNDP—Ministry of Economy. Toward Sustainable Energy: Low-Carbon Development Strategy of the Republic of Uzbekistan), Tashkent, 2015.
6. *Scaling-up energy efficiency in rural buildings of Uzbekistan, Analytical Report,* Tashkent: Center for Economic Research, 2016.
7. *GEF—PROON. Povyshenie energoeffektivnosti zdanii v Uzbekistane: napravleniya reform i ozhidaemye effekty* (GEF—UNDP. Improving the Energy Efficiency of Buildings in Uzbekistan: Reform Directions and Expected Effects), Tashkent, 2014.
8. *Tret'e natsional'noe soobshchenie Respubliki Uzbekistan po ramochnoi konventsii OON ob izmenenii klimata* (Third National Communication of the Republic of Uzbekistan under the UN Framework Convention on Climate Change),Tashkent: Center of Hydrometeorological Service under the Cabinet of Ministers of the Republic of Uzbekistan, 2016.
9. Avezov, R.R., Avezova, N.R., Rakhimov, E.Yu., Kuralov, M.A., and Abdukhamidov, D.U., Ray absorption ability by premises with insolation passive heating system, *Al'tern. Energ. Ekol.,* 2017, nos. 7–9, pp. 34–42.
10. Pakhaluev, V.M., Shcheklein, S.E., and Matveev, A.V., Solar system with seasonal thermal energy storage, *Al'tern. Energ. Ekol.,* 2018, nos. 1–3, pp. 17–25.
11. Sheps, R.A., Golovinskii, P.A., Shchukina, T.V., and Yaremenko, S.A., Heat flux in a passive multi-layered solar panel, *Al'tern. Energ. Ekol.,* 2018, nos. 7–9, pp. 55–62.
12. Samiev, K.A., Heat efficiency of complex translucent barrier with partially absorbing layer and airway in passive solar insolation heating systems, *Appl. Sol. Energy,* 2008, vol. 44, pp. 291–294. https://doi.org/10.3103/S0003701X08040142
13. Avezova, N.R. and Sadykov, Z.D., Influence of the thermal resistance of the collecting-accumulating wall of a passive solar heating system on the replacement factor of heat load, *Appl. Sol. Energy,* 2012, vol. 48, pp. 42–46. https://doi.org/10.3103/S0003701X12010045
14. Samiev, K.A., Simulation of thermal regime of room, heated by passive insolation solar heating systems with three-layer ventilated translucent barriers, *Appl. Sol. Energy,* 2009, vol. 45, pp. 298–302. https://doi.org/10.3103/S0003701X09040197
15. Harkouss, F., Fardoun, F., and Biwole, P.H., Passive design optimization of low energy buildings in different climates, *Energy,* 2018, vol. 165, pp. 591–613.
16. Avezova, N.R., Avezov, R.R., Samiev, K.A., et al., Integration of the Trombe Wall into rural residential buildings in climatic conditions of Uzbekistan, *Appl.*

Sol. Energy, 2021, vol. 57, pp. 333–339.
https://doi.org/10.3103/S0003701X21040022

17. Evins, R., A review of computational optimisation methods applied to sustainable building design, *Renewable Sustainable Energy Rev.,* 2013, vol. 22, pp. 230–245.

18. Schrijver, A., *Theory of Linear and Integer Programming,* Hoboken, NJ: John Wiley & Sons, 1998.

19. Zhang, D., Shah, N., and Papageorgiou, L.G., Efficient energy consumption and operation management in a smart building with microgrid, *Energy Convers. Manage.,* 2013, vol. 74, pp. 209–222.

20. Wang, C., Jiao, B., Guo, L., et al., Robust scheduling of building energy system under uncertainty, *Appl. Energy,* 2016, vol. 167, pp. 366–376.

21. Mehleri, E.D., Sarimveis, H., Markatos, N.C., and Papageorgiou, L.G., Optimal design and operation of distributed energy systems: Application to Greek residential sector, *Renewable Energy,* 2013, vol. 51, pp. 331–342.

22. Wakui, T. and Yokoyama, R., Optimal structural design of residential cogeneration systems in consideration of their operating restrictions, *Energy,* 2014, vol. 64, pp. 719–733.

23. Harb, H., Reinhardt, J., Streblow, R., and Müller, D., MIP approach for designing heating systems in residential buildings and neighbourhoods, *J. Build. Perform. Simul.,* 2015, vol. 9, pp. 316–330.

24. Renaldi, R., Kiprakis, A., and Friedrich, D., An optimisation framework for thermal energy storage integration in a residential heat pump heating system, *Appl. Energy,* 2017, vol. 186, pp. 520–529.

25. Zhang, D., Evangelisti, S., Lettieri, P., and Papageorgiou, L.G., Optimal design of CHP based microgrids: multi-objective optimisation and life cycle assessment, *Energy,* 2015, vol. 85, pp. 181–193.

26. Weber, C. and Shah, N., Optimisation based design of a district energy system for an eco-town in the United Kingdom, *Energy,* 2011, vol. 36, pp. 1292–1308.

27. Yang, Y., Zhang, S., and Xiao, Y., Optimal design of distributed energy resource systems coupled with energy distribution networks, *Energy,* 2015, vol. 85, pp. 433–448.

28. Wouters, C., Fraga, E.S., and James, A.M., An energy integrated, multi-microgrid, MILP (mixed-integer linear programming) approach for residential distributed energy system planning—a South Australian case-study, *Energy,* 2015, vol. 85, pp. 30–44.

29. Ashouri, A., Fux, S.S., Benz, M.J., and Guzzella, L., Optimal design and operation of building services using mixed-integer linear programming techniques, *Energy,* 2013, vol. 59, pp. 365–376.

30. Asadi, E., da Silva, M.G., Antunes, C.H., and Dias, L., Multi-objective optimization for building retrofit strategies: a model and an application, *Energy Build.,* 2012, vol. 44, pp. 81–87.

31. Stadler, M., Groissböck, M., Cardoso, G., and Marnay, C., Optimizing distributed energy resources and building retrofits with the strategic DER-CAModel, *Appl. Energy,* 2014, vol. 132, pp. 557–567.

32. Evins, R., Multi-level optimization of building design, energy system sizing and operation, *Energy,* 2015, vol. 90, pp. 1775–1789.

33. Wolisz, H., Block, P., Streblow, R., and Müller, D., Dynamic activation of structural thermal mass in a multi-zonal building with due regard to thermal comfort, in *14th Conf. of Int. Building Performance Simulation Association,* Hyderabad, India, 2015, pp. 1291–1297.

34. Schütz, T., Schiffer, L., Harb, H., et al., Optimal design of energy conversion units and envelopes for residential building retrofits using a comprehensive MILP model, *Appl. Energy,* 2017, vol. 185, pp. 1–15.

35. Fan, Y. and Xia, X., A multi-objective optimization model for energy-efficiency building envelope retrofitting plan with rooftop PV system installation and maintenance, *Appl. Energy,* 2017, vol. 189, pp. 327–335.

36. Halimov, A., Nürenberg, M., Müller, D., et al., Multi-objective optimization of complex measures on supplying energy to rural residential buildings in Uzbekistan using renewable energy sources, *Appl. Sol. Energy,* 2020, vol. 56, pp. 137–148.
https://doi.org/10.3103/S0003701X20020073

37. Halimov, A., Lauster, M., and Müller, D., Validation and integration of a latent heat storage model into building envelopes of a high-order building model for Modelica library AixLib, *Energy Build.,* 2019, vol. 202, p. 109336.

38. Halimov, A., Lauster, M., and Müller, D., Development and validation of PCMs models integrated into the high-order building model of Modelica library—AixLib, *Proc. 16th IBPSA Conf.,* Rome, Italy, September 2–4, 2019, vol. 7, pp. 4698–4705.

Translated by L. Solovyova

SPELL: 1. OK