TECHNOLOGICAL HETEROGENEITY IN BUILDING STOCK ENERGY MODELLING: EVIDENCE OF IMPORTANCE FROM JAPAN

Rihoko Shinohara[1], Xukang Zhang[1], Yohei Yamaguchi[1]
Masato Miyata[2], Hideaki Uchida[1], Yoshiyuki Shimoda[1]

[1]Osaka University, Osaka, Japan
[2]National Institute for Land and Infrastructure Management, Japan

ABSTRACT

Bottom-up physics-based building stock energy models (BSEMs) have been developed to quantify the energy demand, performance, and carbon dioxide emissions of building stocks to support policymaking. The literature on building stock energy modelling lacks a methodology to manage the heterogeneity in the adopted technologies and their dynamics. In this study, a method is proposed that integrates sample buildings clustered by building specifications, with the development of statistical models used for analysing the building stock according to the adopted technologies. Clustering also enables representative specifications of developed clusters in building archetypes or reference building models to be determined. The resultant changes in energy demand can then be quantified such that the developed BSEM represents building stock heterogeneity in adopted technologies and their dynamics. The method is demonstrated using Japanese office building stock and building energy code compliance evaluation data submitted for newly constructed buildings in 2018. Statistical analysis showed that Japanese office buildings are technologically heterogeneous, even among those constructed in the same year, which implies potential errors in the vintage-based technology adoption assumptions widely used in BSEMs. Additionally, the analysis indicates that future technological changes may not be uniform and responses to policy interventions may vary significantly between clusters. The proposed method is theoretically capable of addressing reality, which will be verified in future studies.

Keywords: building stock energy model; building energy code; building envelope design; building equipment design; cluster analysis.

INTRODUCTION

In the context of climate change mitigation, reducing carbon dioxide emissions from building stocks has become urgent. National-level top-down historical analyses can estimate the aggregated reduction due to past energy efficiency improvements in the building sector but cannot quantify the individual contributions from specific technologies, making it difficult to ensure the achievement of mid/long-term reduction targets. Bottom-up physics-based building stock energy models (BSEMs) can address this issue as the building stock energy demand is quantified based on physics-based simulations. In BSEMs for commercial building stock, energy demand is generally quantified as the sum of the energy consumption of building stock segments, which are the product of the total floor area and energy use intensity of each segment. Many models use building archetypes, also called reference building models (RBMs), to quantify energy use intensity (Mata et al., 2014). The consideration of building stock segments regarding adopted technologies allows BSEMs to quantify the resulting reductions. Sandberg et al. (2021) demonstrated this approach by assuming the dissemination of zero-emission buildings in Norway.

In existing BSEMs, the number of building stock segments for each adopted technology is generally small, because only major combinations of technologies are considered, such as building insulation performance improvement, heat source electrification, and

corresponding conventional technologies. More importantly, many models assume that the adopted technologies differ according to the construction period in the modelling of building stock composition in adopted technologies. However, the current need for evidence-based policymaking for climate change mitigation requires a more detailed consideration of adopted technologies. First, several mitigation technologies were considered. Second, short-term changes in the building stock should be captured to conduct annual or biannual evaluations of building stock emissions to ensure mid- and long-term reduction targets. Previous studies on BSEMs lack a methodology to deal with technological heterogeneity and short-term changes. For residential building stock, Nägeli et al. (2020) developed an agent-based synthetic building stock modelling approach and showed that the approach is effective in representing the heterogeneity and dynamics of building stock characteristics over time. However, this method has not been applied to commercial building stocks.

In this study a modelling method is proposed to represent the technological heterogeneity and dynamics of the building stock in the adopted technologies. This study is the first to show how newly constructed Japanese office buildings are heterogeneous in adopted technologies, even among buildings constructed during the same period, based on building data used for the building energy code compliance evaluation in 2018. We then propose a method to incorporate building stock heterogeneity and short-term changes in the BSEM of the Japanese commercial building stock.

METHOD

Data Used for Analysis

This study used data from building energy code compliance applications. The Building Energy Efficiency Act requires building owners to submit the results of their evaluations of the energy performance of constructed and renovated buildings. Energy performance was measured using the building energy index (BEI), which is calculated as the ratio of the design primary energy consumption to the reference primary energy consumption which is the energy consumption assuming standard building specifications as of 2013 that were used when the building energy code compliance system was established in 2016. To comply with the building energy code, the BEI $\leq$ 1.0 for buildings with a total floor area $\geq$ 300 m². Although there is no compliance requirement for the insulation performance, the building perimeter annual load index (BPI) was used as an indicator. The BPI is calculated as the ratio of the designed thermal load in the perimeter zones to that of the reference insulation performance. The government decided to tighten the code effective April 2024 from BEI ≤ 1.0 to BEI ≤ 0.75–0.85 for buildings with a total floor area $\geq$ 2,000 m². By 2026, it will be extended to buildings with a total floor area $\geq$ 300 m². The BSEM should represent a change in technology adoption owing to code enhancement.

This study used data from newly constructed office buildings evaluated in 2018 using the model-building method[1]. The indicators calculated using the model-building method were denoted as BEIm and BPIm. The data contained BEI values for air conditioning (BEIm/AC), water heating (BEIm/HW), lighting (BEIm/L), and ventilation systems (BEIm/V). The data also included basic building characteristics, such as total floor area, floor composition for different floor usages, regional classifications, and design specifications used to quantify the performance indicators. The regional classification followed the meteorological divisions of the Building Energy Efficiency Act.

For data preprocessing, 21 outliers beyond the three standard deviations of BEIm were excluded, and the remaining samples were standardised. Of the data used for the analysis, 818 office buildings met the enhanced code, whereas 335 did not. A detailed analysis was presented by Shinohara et al. (2023).

Analysis Method

To elucidate the heterogeneity in technology adoption, the sample buildings were classified based on BEIm/AC, BEIm/HW, BEIm/L, BEIm/V) and BPIm. Principal

[1] Both indicators are evaluated using the standard input method or the model building method by a program released by the National Institute for Land and Infrastructure Management. In 2018, 6% of applications used the standard input method, while the rest used the model building method.

component analysis was not a suitable analysis tool because of the lack of high correlations among the variables, therefore cluster analysis using the Ward method was applied for classification.

Next, a logistic regression analysis was performed to develop a multinomial logistic regression model that calculated the probability of belonging to each cluster developed in the previous step. Table 1 lists the predictor variables used in the analysis. The variables for the total floor area, window area ratio, and external wall area per unit floor area (shape factor) were standardised. Variables were selected using a stepwise method from the predictor variables listed in Table 1. The probability of belonging to each cluster p_k, is expressed in Equation (1), where a_{kj} is the partial regression coefficient (j = 1 to n, where n is the number of predictor variables x_j) and b_k is a constant.

$$p_k = \frac{1}{1+exp\left\{-\left(\sum_{j=1}^{n} a_{kj}x_j+b_k\right)\right\}} \quad (1)$$

The predictor variables in the model include building attributes, such as location and scale. The regression model obtained can be used to decompose the total floor area into those belonging to the identified clusters, considering the building stock composition of the building attributes in the BSEM application.

each cluster. CL1 exhibited high performance across insulation and design specifications. CL2 had a lower performance in ventilation, CL3 in air conditioning and insulation, CL4 in lighting, and CL5 in hot water supply. CL1 and CL5 accounted for approximately 65% of the samples. There was no cluster with low performance for all variables. The variation in the variables is attributed to the distribution of the design specifications. Figure 3 shows the cumulative frequency distribution of the design specifications for each cluster, with graphs showing the cooling coefficient of performance (COP), cooling capacity (Cap), heating COP, heating capacity, hot water supply efficiency, lighting power consumption, ventilation equipment power consumption, window U-values, and wall U-values.

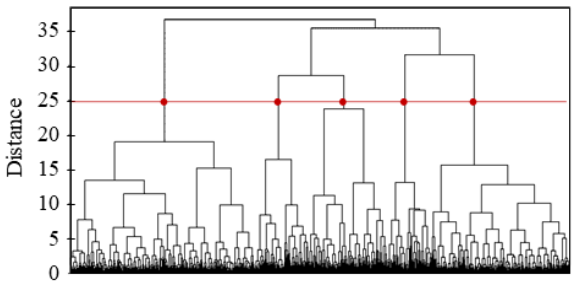

Figure 1: Dendrogram (Office)

Table 1 Predictors of logistic regression analysis

Item	Predictor variables
Cold	Dummy variable of cold region (the regions 5 to 8 are 0, whereas the regions 1 to 4 are 1)
Hot	Dummy variable of warm region (the regions 1 to 6 are 0, whereas the regions 7 and 8 are 1)
Total floor area	Base 10 logarithm of building total floor area [m²]
Window area rate	Standardized window area rate
Shape factor	Standardized envelope area per total floor area
Air conditioning system	Dummy variable of air conditioning system (decentralized system is 0, centralized is 1)
Bathroom	Dummy variable of bathroom (not having a bathroom is 0, having a bathroom is 1)
Parking	Dummy variable of parking (no parking is 0, parking available is 1)

RESULTS FOR OFFICE BUILDINGS

Cluster Analysis

Figure 1 shows the dendrogram from the cluster analysis, with data points on the horizontal axis and the distance between clusters on the vertical axis. The dendrogram suggested five clusters (CL1 to CL5), as the change in height between the clusters decreased significantly after six clusters.

Figure 2 presents the cumulative frequency distributions of BEIm/AC, BEIm/HW, BEIm/L, BEIm/V, and BPIm for

Distribution of BEIm of the clusters

Figure 4 shows the frequency distribution of the BEIm. The enhanced building energy code for office buildings is BEIm = 0.80. CL1 generally met the enhanced code, indicating high-energy performance. The BEIm < 0.80 for most buildings in CL2 and CL5, even though they had relatively low performances in ventilation and water heating. This is because the proportions of energy consumed for ventilation and water heating are relatively small in office buildings. The impact of the code enhancement was small in these clusters.

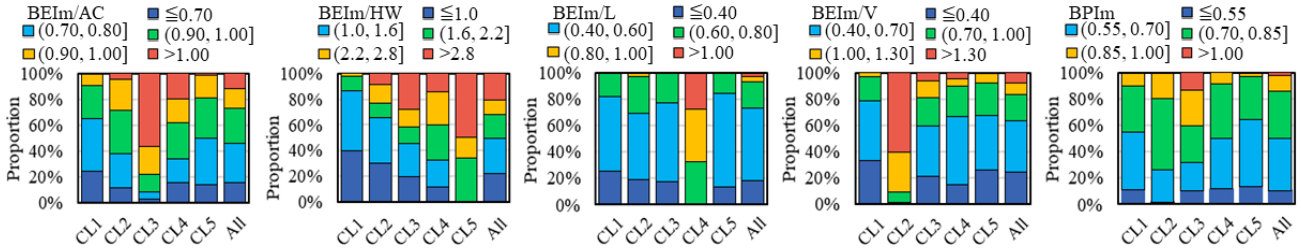

Figure 2: Distribution of BEIm of equipment and Bim (Office)

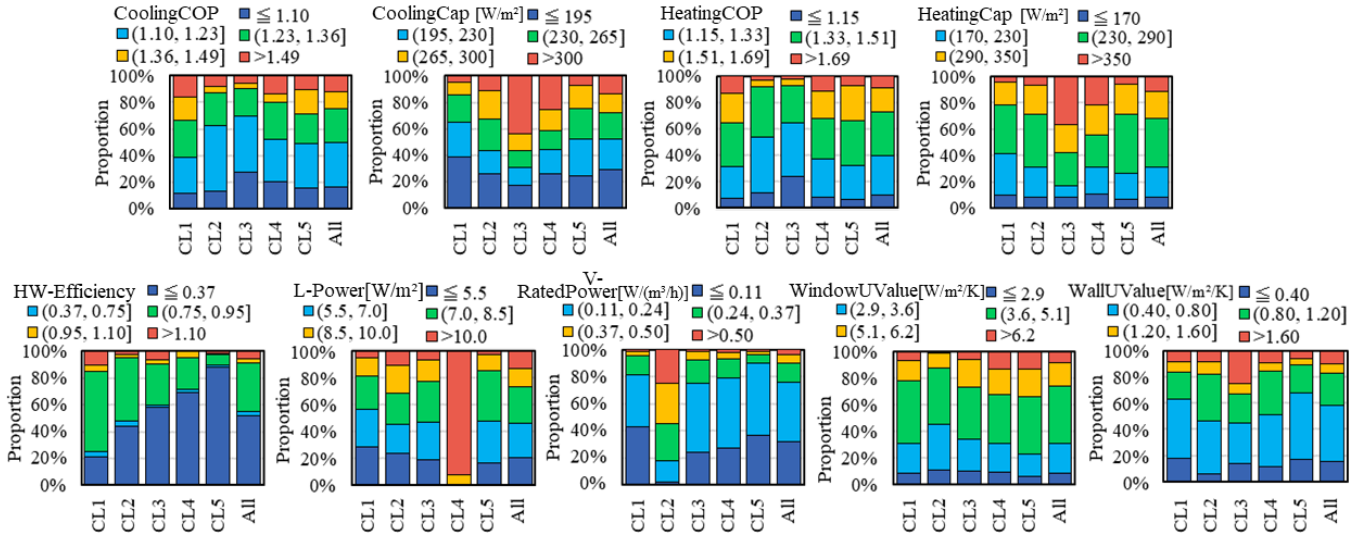

Figure 3: Cumulative frequency distribution envelope and equipment design specifications in each cluster (Office)

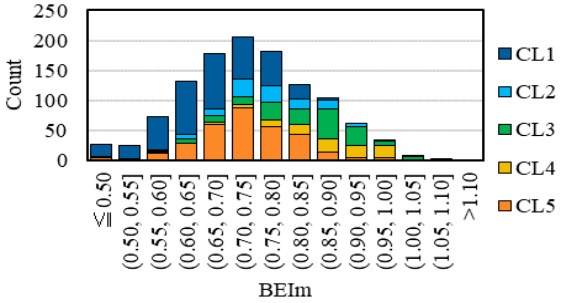

Figure 4: Distribution of BEIm (Office)

BEIm > 0.80 predominantly for CL3 and CL4, indicating a possible reclassification with the enhanced code. CL3 exhibited high BEIm/AC and BPIm values, indicating low energy performance in air conditioning and insulation. These areas therefore need improvements to satisfy enhanced codes. As shown in Figure 3, CL3 had higher heating and cooling capacities than CL1. Reducing these capacities could align the CL3 specifications with those of CL5. CL4 had a high BEIm/L, indicating lower lighting energy performance. A high percentage of lighting power consumption per unit floor area, > 10 W/m², suggests the use of non-LED lighting or designs with high lighting capacity. Improving lighting energy performance could bring CL4 closer to CL1 post codes.

Logistic Regression Analysis

Table 2 shows the predictor variables selected using the stepwise method and the logistic regression analysis coefficients. The Hosmer-Lemeshow test yielded a p-value of 0.32 ($p > 0.05$), indicating that the model fits the data used to develop it.

Figure 5 illustrates the relationship between the probabilities of cluster membership and the building attributes. For the estimation, the predictor variables listed in Table 3 were used to illustrate how membership probabilities changed with the conditions considered as predictor variables.

As shown in Figure 5, the cluster membership probabilities change significantly with location, total floor area, window area ratio, shape factor (exterior wall area/unit floor area), air-conditioning heat source type, presence of a bathroom, and presence of a parking lot. In colder areas, the probability of belonging to CL2, which had a higher power consumption per unit ventilation rate, was higher. As the total floor area increased, the probability of belonging to CL2 increased,

whereas the probabilities for CL1 and CL5 decreased. A smaller window area ratio increased the likelihood of CL5 and decreased the likelihood of CL3. A lower shape factor increased the probability of belonging to CL1 and decreased the probability of belonging to CL3. The central heat source system had a higher proportion of CL2 members than the individual systems. The presence of bathrooms and parking lots also affected the cluster membership probabilities, indicating that room use influenced equipment specifications.

the remaining buildings exhibited lower performance in one or more specifications. There were few buildings with low performance in all specifications, indicating that noncompliance is due to poor performance in certain specifications rather than overall low performance. The results also showed that buildings that do not meet the code often exhibit low-performing air-conditioning systems, insulation, or lighting. Thus, enhancing the performance of the air conditioning systems, insulation, and lighting can help achieve

Table 2 Coefficient of the developed regression model for office

Item / Cluster	CL1	CL2	CL3	CL4	CL5
Cold	0.12	0.26	0.11	-0.57	0.07
Total floor area	-0.42	0.89	0.00	-0.25	-0.23
Window area rate	0.00	0.04	0.31	-0.04	-0.32
Shape factor	-0.38	0.02	0.12	0.14	0.09
Air conditioning system	-0.04	-0.33	0.33	-0.63	-0.25
Bathroom	0.76	0.14	-0.12	-0.35	-0.43
Parking	0.13	-0.70	0.06	0.85	-0.33

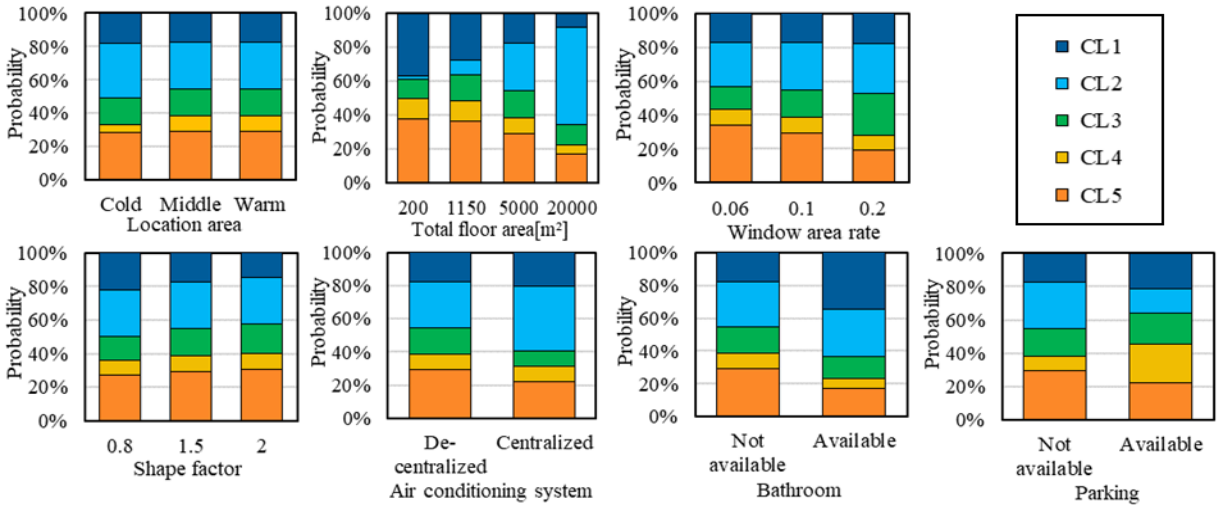

Figure 5: Membership Probability

Table 3 Standard specification for offices

Item	Predictor variables
Location area	Middle
Total floor area [m²]	5,000
Window area rate [-]	0.1
Shape factor [-]	1.5
Air conditioning system	Decentralized
Bathroom	Not available
Parking	Not available

The results showed that 40% of the newly constructed office buildings possessed high energy performance across all specifications, as indicated by CL1, whereas

compliance with the enhanced code. Based on these results, we estimated the changes resulting from the building energy code enhancement by comparing the average values between clusters with low energy performance and other clusters.

The results are listed in Table 4, focusing on CL3 and CL4, which do not satisfy the enhanced code. For CL4 (9% of buildings), lighting power should be reduced by 6.4 W/m². For CL3 (16% of buildings), heating and cooling COP needs slight improvement with reduced capacity of 60–70 W/m², and external wall U-values need to be reduced by 0.34 W/m²/K to meet the average levels of other buildings. Using the Model Building Method Input

Assistance Tool (Ver. 3.3.3) for the 2016 building energy code, we randomly selected 30% of the buildings that did not meet the enhanced code to re-evaluate whether these changes resulted in compliance with the enhanced code. Figure 6 shows that 73% of the sampled buildings satisfied the requirements of the enhanced code. Thus, code enhancement can force changes to the design specifications of these clusters to improve the performance of these systems.

enhanced code. CL4, which has a low lighting energy performance, needs to upgrade the lighting system performance. The results also indicate that future technological changes may not be uniform, and the response to building energy code enhancements differs significantly among clusters. In addition, the logistic regression analysis results showed that the proportion of each cluster differed significantly with building characteristics such as the total floor area. The

Table 4: Variation in envelope and equipment design specifications associated with enhanced codes (Office)

Item / Cluster	Average						Variation		
	Except CL3 (84%)	Except CL4 (91%)	CL3 (16%)	CL4 (9%)	0.6<BEIm ≦0.8 (60%)	BEIm>0.8 (29%)	CL3 (16%)	CL4 (9%)	BEIm-based (29%)
Cooling COP [-]	1.34	-	1.18	1.43	1.28	1.20	0.16	0	0.08
Cooling capacity [W/m^2]	226	-	286	254	218	289	-60	0	-71
Heating COP [-]	1.51	-	1.27	1.59	1.43	1.30	0.24	0	0.13
Heating capacity [W/m^2]	260	-	330	299	249	331	-70	0	-82
Hot water efficiency [-]	-	-	0.62	0.53	0.65	0.58	0	0	0.07
Lighting power [W/m^2]	-	6.9	7.1	13.3	7.0	9.3	0	-6.4	-2.2
Ventilation rated power [W/(m^3/h)]	-	-	0.19	0.18	0.19	0.23	0	0	-0.05
Window U value [W/m^2/K]	-	-	4.22	4.34	4.28	4.23	0	0	0
Wall U value [W/m^2/K]	0.82	-	1.16	0.91	0.86	0.94	-0.34	0	-0.08

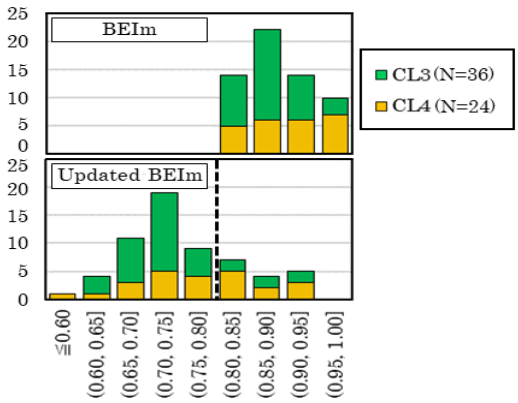

Figure 6: Distribution of BEI after updated equipment specifications

DISCUSSION

Heterogeneity in Adopted Technologies

The analysis revealed that buildings constructed in the same year exhibited technological heterogeneity, suggesting that the vintage-based assumption of the adopted technologies may include significant errors in the BSEM. CL1, characterised by high performance in all building systems, accounted for approximately 40% of all newly constructed buildings. In contrast, buildings in CL2 and CL5 had ventilation and water-heating systems with relatively lower performance. CL3, which exhibits low performance in air-conditioning systems and insulation, requires enhancements to satisfy the

recognition of clusters and potential changes due to code enhancement would contribute to improving the model performance of the BSEM by better representing the baseline development of building stock.

Based on these results, the building stock composition was estimated, as shown in Figure 7. The total floor area was determined based on Narukawa et al. (2023), and the composition was estimated by applying the developed logistic regression model to the building stock segments, considering building size and location. The legend "Before 2018" indicates those constructed before 2018, which were excluded from this analysis. CL1–CL5 correspond to clusters with locations divided into regions 1-4 (cold regions) and 5-8 (other regions).

As shown in the figure, CL1 and CL2 constitute a large portion of the new building construction. The proportions of CL3 and CL4 will account for only 4% and 1%, respectively, by 2030. This result indicates that code enhancement by 2030 may be small. Figure 8 shows the total floor area based on building size for each cluster. The proportions were significantly different from those shown in Figure 4 because Figures 7 and 8 show the composition by total floor area instead of the number of buildings shown in Figure 4. The increase in the proportion of CL2 was attributed to CL2 having a high proportion among large buildings, as shown in Figure 5,

and the composition of large buildings was large in the building stock, as shown in Figure 8.

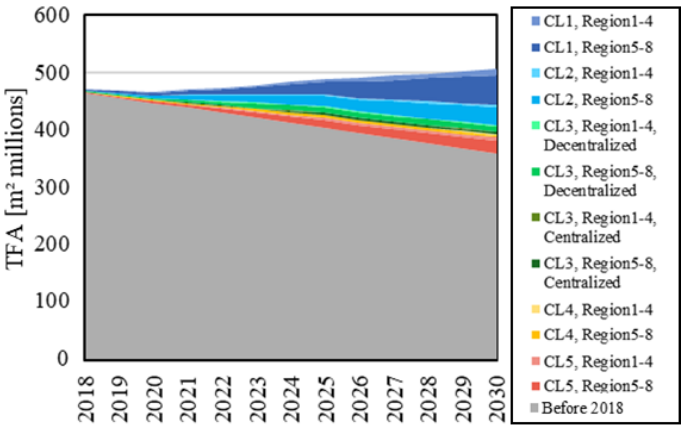

Figure 7: Transition of total floor area (TFA) in stock composition by time

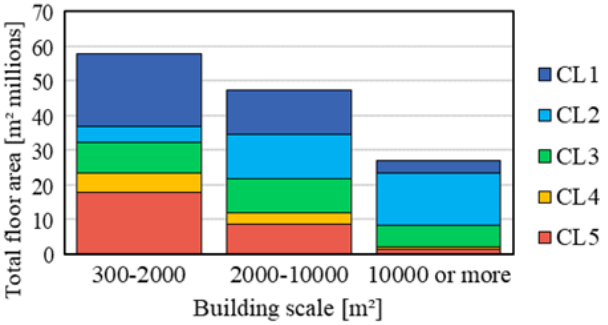

Figure 8: Composition of new buildings constructed since 2018

BSEM Application

We propose a method to apply building stock development by 2030 to the BSEM for the Japanese office building stock (Yamaguchi et al., 2022). The model considers building stock segments based on building size, location, and heating, ventilation, air conditioning and water heating system types. In addition to these elements, building stock segments were developed to consider CL1–CL5. The total floor area can be obtained by following the results shown in Figure 7. The specifications of the RBMs are listed in Table 5, considering a combination of technologies. Because cold regions exhibit unique characteristics in terms of equipment and insulation, regions 1-4 and 5-8 were treated separately. The RBMs for CL1–CL5 correspond to the clusters, representing the average specifications of each cluster. Considering the significant impact of the total floor area on the composition ratio, the ratios are shown for three categories: total floor area of 300–2,000 m², 2,000–10,000 m², and over 10,000 m².

By incorporating these combinations in adopted technologies into the RBMs, the BSEM can represent technological heterogeneity and potential changes in the code enhancement.

Table 5: Archetypes Considering Technological Combinations

Composition ratio **%:TFA300-2000m², **%:TFA2000-10000m², **%:TFA10000m² or more

		Air conditioning								Hot water supply		Lighting	Ventilation		Insulation		
		Cooling				Heating									Wall	Window	
		COP [-]	Capacity [W/m²]	Unit	System	COP [-]	Capacity [W/m²]	Unit	System	Efficiency [-]	Unit	Power [W/m²]	Power [W/(m³/h)]	Unit	U value [W/m²/K]	U value [W/m²/K]	Window area rate [-]
CL1 37%	Region 1-4	1.32	192	PAC	Decentralized	1.43	232	PAC	Decentralized	0.86	CvB	6.28	0.13	Type 2or3	0.53	3.77	0.10
25%	Region 5-8	1.26	214	PAC	Decentralized	1.45	243	PAC	Decentralized	0.85	CvB	6.82	0.13	Type 2or3	0.68	4.10	0.12
CL2 9%	Region 1-4	1.26	229	PAC	Decentralized	1.37	280	PAC	Decentralized	0.68	CvB	6.53	0.43	Type 1	0.62	3.69	0.13
31%	Region 5-8	1.19	240	PAC	Decentralized	1.31	262	PAC	Decentralized	0.66	CvB	7.59	0.39	Type 1	0.84	3.62	0.15
CL3 16%	Region 1-4	1.20	250	PAC	Decentralized	1.21	334	PAC	Decentralized	0.62	CvB	6.59	0.18	Type 2or3	0.91	3.86	0.15
16%	Region 5-8	1.18	294	PAC	Decentralized	1.31	328	PAC	Decentralized								
2%	Region 1-4	1.19	115	WCU	Centralized	0.95	377	HWG	Centralized	0.62	CvB	7.21	0.20	Type 2or3	1.22	4.35	0.15
2%	Region 5-8	1.11	225	WCU	Centralized	1.18	270	WCU	Centralized								
CL4 9%	Region 1-4	1.16	234	PAC	Decentralized	1.34	264	PAC	Decentralized	0.54	CvB	13.09	0.16	Type 2or3	0.46	3.35	0.12
6%	Region 5-8	1.16	234	PAC	Decentralized	1.34	264	PAC	Decentralized	0.53	CvB	13.35	0.18	Type 2or3	0.82	4.17	0.11
CL5 27%	Region 1-4	1.28	214	PAC	Decentralized	1.39	258	PAC	Decentralized	0.37	ELE	6.49	0.13	Type 2or3	0.53	3.83	0.10
20%	Region 5-8	1.24	228	PAC	Decentralized	1.43	259	PAC	Decentralized	0.37	ELE	7.15	0.15	Type 2or3	0.66	4.47	0.10

CONCLUSION

This study proposes a methodology to capture the heterogeneity in the adopted technologies and their dynamics in building stock energy modelling using Japanese commercial building stock as an example. The Japanese case shows significant technological differences even among buildings constructed in the same period. The proposed method integrates sample building clustering by specifications and statistical analysis to develop models for building stock decomposition based on adopted technologies. By applying the quantified specifications of developed clusters in building archetypes or reference building models, the consequences of energy demand can be quantified while representing building stock heterogeneity in the adopted technologies and their dynamics. This would allow for more accurate predictions of changes in the building stock energy demand, aiding in better planning for carbon dioxide emission reductions from commercial building stocks.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 23K26260 and the National Institute for Land and Infrastructure Management (NILIM). The data used was from the Model Building Method Input Support Tool (Version 3.3.3) for the 2018 fiscal year.

REFERENCES

Ministry of the Environment (2021). Climate Change Mitigation Plan.

Nishijima, T. et al. (2023). Development of the building stock energy model for commercial building stock considering heterogeneity and yearly change in building stock. Proceeding of Building Simulation, 3540-3547.

Ministry of Land, Infrastructure, Transport and Tourism (2015). Act on the Improvement of Energy Consumption Performance of Buildings.

Ministry of Land, Infrastructure, Transport and Tourism (2021). About the Raising of Building energy codes for Large Non-Residential Buildings.

Miyata, M. and Miki, Y. (2020). Investigation on Building Envelope Design, Equipment Design and Energy Performance Evaluation Results of Commercial Buildings: Analysis of Input/Output Data of the Energy Performance Calculation Program for Japanese Building Energy Code. Technical Note of National Institute for Land and Infrastructure Management 1107 (in Japanese).

Building Research Institute (2012). Technical Information on Energy Consumption Performance of Buildings.

Miyata, M. et al. (2015). Web-based Simulation Tool for Compliance with 2013 Building energy code for Commercial Buildings in Japan. Proceeding of Building Simulation, 1766-1773.

Miyata, M., Hirakawa, S. (2020). Analysis on Building Envelope and Building Service Equipment Design Specification Using the Input and Output Data from the Calculation Program to Confirm Compliance with Building Energy Code (Part 1): Identification of the Design Specification for Newly Built Offices in Japan According to the Evaluation Result of the Building Energy Code. Journal of Environmental Engineering (Transactions of AIJ) 85(777), 859-869 (in Japanese).

Narukawa, K. et al. (2022). Analysis on Building Envelope and Building Service Equipment Design Specification Using the Input and Output Data from the Calculation Program to Confirm Compliance with Building Energy Code (Part 2): Logistic Regression-based Analysis Considering Building Size and Location. Journal of Environmental Engineering (Transactions of AIJ) 87(797), 448-459 (in Japanese).

Yamaguchi, Y. et al. (2022). Building Stock Energy Modeling Considering Building System Composition and Long-Term Change for Climate Change Mitigation of Commercial Building Stocks. Applied Energy 306.

Miyata, M. et al. (2017). Relevant Materials for 2016 Building Energy Efficiency Standard (Promulgated in January 2016), Manual for Model Building Method Input Assistance Tool. Technical Note of National Institute for Land and Infrastructure Management 974 (in Japanese).

Sandberg, N. H. et al. (2021). Large potentials for energy saving and greenhouse gas emission reductions from large-scale deployment of zero emission building technologies in a national building stock. Energy Policy 152.

Mata, É. et al. (2014). Building-stock aggregation through archetype buildings: France, Germany, Spain and the UK. Building and Environment 81, 270-282.

Nägeli, C. et al. (2020). Towards agent-based building stock modelling: Bottom-up modelling of long-term stock dynamic affecting the energy and climate impact of building stock. Energy and Building 211.

Shinohara, R. et al. (2023). Technical Changes Associated with Enhanced Energy Consumption Efficiency Standards for Non-Residential Buildings. Technical Papers of Annual Meeting the Society of Heating 8, 25-28 (in Japanese).