Research Article

Development of Regression Models considering Time-Lag and Aerosols for Predicting Heating Loads in Buildings

Hong Soo Lim and Gon Kim

Department of Architectural Engineering, Kyung Hee University, Yongin 446-701, Republic of Korea

Correspondence should be addressed to Gon Kim; gonkim@khu.ac.kr

Received 1 August 2017; Revised 2 November 2017; Accepted 23 November 2017; Published 1 February 2018

Academic Editor: Gianmarco de Felice

Copyright © 2018 Hong Soo Lim and Gon Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Building automation systems is becoming more vital, especially in regard to reduced building energy consumption. However, the accuracy of such systems in calculating building thermal loads is limited as they are unable to predict future thermal loads based on prevailing environmental factors. The current paper therefore seeks to improve the understanding of the interactions between outdoor meteorological data and building energy consumption through a statistical analysis. Using weather data collected by the Korean Meteorological Agency (KMA) over a period of three years (2011–2014), prediction models that are able to predict heating thermal loads considering the time-lag phenomenon are developed. In addition, the study develops different prediction models for buildings of different sizes. The results confirm the existence of the time-lag phenomenon: the heating load experienced by a building at a given time is better explained by a regression model developed using the climatic conditions that existed two hours before. As such, conventional building simulation programs must endeavor to include time-lag as well as Aerosol Optical Depth (AOD) data as important factors in the prediction of building heating loads.

1. Introduction

Energy efficiency is of particular importance, especially in the building sector. The building sector consumes a significant amount of energy in comparison with other sectors [1]. As such, reducing the energy consumed by buildings has been at the core interest of building-related research.

Over the years, various energy-saving systems have been developed and implemented. Research from multiple disciplines, such as IT (information technology) and mechanical engineering, are being utilized to improve the energy performance of buildings through building automation. For simplification, building automation can be divided into three categories: the building management system (BMS), the security system (SS), and the energy management system (EMS). The EMS is more closely associated with the energy-consuming units of the building, such as the heating, ventilating, and air conditioning (HVAC) systems and systems that require mechanical operations, such as elevators, automated doors, and so forth.

The likes of such systems are essential in predicting the energy consumed by buildings in that they collect the energy usage patterns of a given building, which can be used later in the calculation of the total energy consumed by the building [2]. These systems are also helpful in understanding complex relationships between building occupants and their environments [3].

Although systems that assist in the prediction of building energy performance already exist, such as the EMS and other related building simulation software, the process of accurately predicting the amount of energy consumed by a given space is rather complex [4]. It requires an extensive understanding of certain factors that influence how a building behaves. An example of such a factor is the outdoor climatic conditions. There is a plethora of studies that deal with the complex relationships between buildings and physical factors [5]. However, studies that deal with outdoor climatic conditions and their influence on building energy consumption are rather insufficient. The current study seeks to fill this gap and improve the understanding of the

interaction between the prevailing outdoor elements and building energy consumption. In addition, this study seeks to present ways in which the newly found relationship between outdoor climatic conditions and building energy performance can be used to predict building heating thermal loads. Furthermore, this study considers aerosol weather data that were not previously taken into account when predicting building energy using computer simulation programs.

2. Energy Prediction in Buildings

Achieving significant reductions in the energy consumed by buildings requires knowledge about the thermal behavior of a given building. With the knowledge regarding building behavior and how a building interacts with its surroundings, the energy consumed by a building can then be predicted and subsequently reduced. Traditionally, building energy consumption could be predicted through mathematical equations. To do this, factors related to building thermal loads and their influence on the thermal performance of the building had to be extensively studied. Nowadays, new and simpler methods are constantly being developed to assist building professionals in their endeavors to optimize building designs and improve building energy performance [6]. Additionally, the estimation of building energy consumption has become a key factor in documenting energy consumed by buildings. However, it is not a simple task to estimate building energy consumption, as it involves various factors that are associated with the characteristics of buildings, such as the climatic conditions, occupant schedule, HVAC systems, and so forth [7].

Recently, scientific studies have been focusing on human behavior and its influence on energy consumption by buildings. Erickson et al. alluded to the issue of modern buildings being conditioned by relying on the assumption of maximum occupancy and, thus, conditioning spaces unnecessarily. Subsequently, energy is utilized where it is not needed. In addition, the study presented a method by which energy occupancy models can be developed using a wireless sensor network integrated into building HVAC systems to ensure that spaces are not conditioned without occupancy [8]. Furthermore, Ahn and Park [9] explained how poor consideration of occupant behavior might lead to inconsistencies between the actual building energy data and the modeled building energy data. In the study, multiple approaches to optimize the occupant schedule data in building modelling were also discussed. Virote [10] also stresses the importance of considering occupant behavior and space occupancy in building energy modelling. Different occupant schedules are likely to lead to different energy usage patterns. In addition, Virote demonstrates the applicability of stochastic models to estimate occupant behavior and space occupancy.

Various statistical modelling techniques are also playing a significant role in building modelling. A number of studies have developed models that either predict the amount of energy consumed by buildings or models that improve the accuracy of the building energy prediction software that

already exists. For instance, a study by Ahmad et al. [11] has demonstrated how building electric energy can be predicted using artificial intelligence methods. His study discusses two common approaches used in deep learning-related predictions: support vector machine (SVM) and artificial neural network (ANN). In addition, the study shines light on new data-handling techniques being developed to improve building modelling and building behavior prediction. Some of these methods include least square support vector machine (LSSVM) and group method of data handling (GMDH). Ahmad concludes that there are advantages and disadvantages associated with each model, and it is thus difficult to pinpoint which of the modelling techniques is superior. Deep learning techniques have also been used to study other aspects of building modelling such as occupant behavior. Zhao et al. [12] developed a data-mining model that learns the behavior of occupants based on the office appliance usage. Grey-box models have also been predominantly used in studying short-term behavior of buildings [13].

Li et al. [14] also introduced an advanced deep learning approach to increase accuracy in building energy consumption predictions. The introduced method is a combination of two approaches: stacked autoencoders (SAEs) and extreme learning machine (ELM). The SAE is used in studying the building's energy usage patterns whereas the ELM is used as the prediction learning tool. The proposed approach shows better performance when compared with other deep learning techniques such as support vector regression (SVR), generalized radial basis function neural network (GRBFNN) and multiple linear regression (MLR). Sensor-based machine learning techniques are also widely applied in forecasting building energy consumption. Such techniques require less input data than conventional statistical methods and are thus less complex [15]. Fan et al. [16] discuss the application of data-mining techniques in developing a set of predictive models that estimate next-day energy consumption and peak power demand. The developed ensemble models show a much higher performance in comparison with single-base models.

Among statistical approaches used in building energy estimation, regression analysis has shown significant outcomes. Regression analysis is widely adopted because it is easy to use and provides reasonably accurate results. Braun et al. used regression models to predict the future electricity and gas consumption of a supermarket in northern England, obtaining theoretically accurate results [17]. Similarly, Tso and Yau also developed regression models to predict the electricity consumed by buildings [18]. Lam et al. further highlighted the usefulness of regression analysis to predict building behavior in the early design stage of buildings [19]. The accuracy of regression analysis in predicting the behavior of buildings has been so far outstanding. Regression models have been used to estimate the heating energy demand in buildings, obtaining accurate results with R-values as high as 0.987 [20] and 0.90 [21]. The accuracy of regression models to predict building energy consumption is constantly being improved through continuous research. For instance, Fumo and Rafe Biswas [22] point out the

TABLE 1: Weather data and aerosol data considered in this study.

Name	Units	Used by EnergyPlus
Dry bulb temperature	°C	Provided by KMA
Wet bulb temperature	°C	Calculated by equation
Dew point temperature	°C	Provided by KMA
Humidity ratio	g/m ³	Provided by KMA
Relative humidity	%	Provided by KMA
Atmospheric pressure	Pa	Provided by KMA
Extraterrestrial horizontal radiation	Wh/m ²	Calculated by equation
Extraterrestrial direct normal radiation	Wh/m ²	Calculated by equation
Solar altitude	Degrees	Calculated by equation
Horizontal infrared radiation intensity from the sky	Wh/m ²	Calculated by equation
Global horizontal radiation	Wh/m ²	Provided by KMA
Direct normal radiation	Wh/m ²	Calculated by equation
Diffuse horizontal radiation	Wh/m ²	Calculated by equation
Wind direction	Degrees	Provided by KMA
Wind speed	m/s	Provided by KMA
Total sky cover	Tenths	Provided by KMA
Visibility	km	Provided by KMA
Precipitable water	mm	Measured by AERONET
Aerosol optical depth	Thousandths	Measured by AERONET

FIGURE 1: Time-lag phenomenon.

importance of the time interval of measured data in regard to the quality of produced regression models. Similarly, the current paper seeks to utilize regression models to improve accuracy in the prediction of building thermal heating loads by considering the time-lag phenomenon and a new climatic factor, aerosols. Both of these factors have not been studied extensively in regard to building energy modelling through regression analysis.

3. Methodology

3.1. *Weather Data and Energy Predictions in Buildings.* Predicting the amount of energy used by a given

space requires a number of factors. One of the most important factors is the weather data. Previous studies [23] have shown indoor climatic conditions and outdoor climatic conditions to be correlated in a directly proportional manner. As such, existing outdoor climatic conditions have a significant influence on how a building occupant behaves; for instance, if it is cold outside, the occupant is more likely to activate the heating system. Many energy prediction software programs have methods in which information about the climatic conditions of an area is entered into the program for analysis. In cases where real-time data from weather stations are unavailable, weather data can be obtained from online databases. Some types of weather data are

TABLE 2: DOE commercial reference buildings.

U.S. Department of Energy commercial reference building models of the national building stock		
Small office	Medium office	Large office

even derived from developed mathematical equations. However, weather data obtained using methods other than actual data from weather stations have been known to cause inconsistencies in the predicted energy results. In fact, the Lawrence Berkeley National Laboratory (LBNL) attributes the differences between the actual amount of energy consumed by a building and the amount of energy predicted by simulation software to be due to inaccuracies in the weather data used [24]. Due to the amount of climatic conditions used in simulation software to predict energy used by a building, there have been attempts [25] to use statistical methods such as the Taguchi method to sieve out the least important climatic factors and thus reduce the number of input parameters that describe weather conditions. In a similar study, Kapetanakis et al. [26] used data from Building Energy Management System (BEMS) to examine and subsequently provide appropriate input weather variables necessary for predicting building heating thermal loads. Different energy prediction software also uses different formats of weather data [27]. This is mainly because such software was developed using different algorithms and might not be able to process a weather data file with a different format. Consequently, the current study converted the format of the original data (data obtained from KMA) into a format that can be read by EnergyPlus, “epw.” Table 1 shows the weather data used within this study.

3.2. The Time-Lag Phenomenon. The time-lag phenomenon is being studied in the current paper to address the inconsistencies brought upon by the usage of weather data in building energy predictions. Theoretically, it is usually assumed that indoor climatic conditions are a direct representation of outdoor climatic conditions, especially in passive buildings that lack mechanical systems to adjust indoor environmental conditions. In reality, however, outdoor climatic conditions at a given time are not accurately matched by indoor climatic conditions at that same

time. The term “time-lag” is being used to represent the delay that occurs between peak energy usage by a given building and peak outdoor climatic conditions. Figure 1 gives an example of the time-lag phenomenon. For May 13, 2011, the peak radiation from the sun, peak temperature, and peak cooling load occurred at approximately 1 pm, 3 pm, and 4 pm, respectively. Theoretically, the peak cooling load should occur at the same time as the peak outdoor temperature, but in reality, there was a delay of 1 hour between the time the peak temperature occurred and the time the peak cooling load occurred. Moon and Kim [28] discussed in detail the issue of time-lag in buildings.

3.3. Reference Building. The current study used the U.S. Department of Energy (DOE) commercial buildings as case models for analysis. The DOE reference models are standard models used for energy performance assessments in building-related studies [29]. The reference building models incorporate a total of 16 building types and 16 different geographical locations in the U.S. They represent approximately 60% of the characteristics usually found within commercial buildings [30]. Table 2 shows the DOE reference buildings used.

Furthermore, ASHRAE specifies building material according to different climatic conditions. South Korea, which is the base area of this study, falls under ASHRAE climatic zones 4A and 4B; the building materials used are in line with ASHRAE standard 90.1–2004 to match the climatic conditions of the location under study. It is worth noting that the regression equations determined through this study might need revisions for buildings made of different materials.

3.4. Research Process. As mentioned in the previous sections of this study, the current paper seeks to develop heating load prediction models while considering the time-lag phenomenon. Multiple regression modelling is used to study the

FIGURE 2: Study methodology workflow.

TABLE 3: DOE small office model summary considering time-lag.

Model summary ^a										
Model	R	R square	Adjusted R square	Std. error of the estimate	R square change	Change statistics F change	df1	df2	Sig. F change	Durbin-Watson
Time-lag0	0.534 ^b	0.285	0.285	7463851.345	0.285	737.844	11	20,340	0.000	0.799
Time-lag1	0.542 ^b	0.294	0.294	7417381.535	0.294	770.449	11	20,339	0.000	0.834
Time-lag2	0.592 ^b	0.350	0.350	7116200.437	0.350	996.953	11	20,338	0.000	0.778

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

TABLE 4: ANOVA for a DOE small building depending on time-lag.

ANOVA ^a						
Model		Sum of squares	df	Mean square	F	Sig.
Time-lag0	Regression	452150586586186300.000	11	41104598780562392.000	737.844	0.000 ^b
	Residual	1133122624135237120.000	20,340	55709076899470.850	—	—
	Total	1585273210721423360.000	20,351	—	—	—
Time-lag1	Regression	466270598690733180.000	11	42388236244612096.000	770.449	0.000 ^b
	Residual	1119001925868610430.000	20,339	55017548840582.650	—	—
	Total	1585272524559343620.000	20,350	—	—	—
Time-lag2	Regression	555346313497669060.000	11	50486028499788096.000	996.953	0.000 ^b
	Residual	1029922597401856130.000	20,338	50640308653842.860	—	—
	Total	1585268910899525120.000	20,349	—	—	—

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

relationship between the outdoor climatic conditions and heating load in buildings of three different sizes while considering time-lag. Ideally, the purpose is to find which

weather data best describes the heating load at a given time. To do that, EnergyPlus is made to output hourly data in three phases. First, when the time-lag is zero (time-lag0), meaning

TABLE 5: Regression coefficients for a DOE small office building depending on time-lag.

		Coefficients ^a			<i>t</i>	Sig.	Collinearity statistics	
Model		Unstandardized coefficients		Standardized coefficients			Tolerance	VIF
		<i>B</i>	Std. error	Beta				
Time-lag0	(Constant)	-173421890.552	11828102.396	—	-14.662	0.000	—	—
	Dry bulb temperatures	-577018.579	43,428.001	-0.540	-13.287	0.000	0.021	47.057
	Dew point temperatures	-185925.394	46,080.321	-0.195	-4.035	0.000	0.015	66.370
	Relative humidity	27,650.459	12,842.837	0.061	2.153	0.031	0.044	22.624
	Atmospheric station pressure	1668.517	114.145	0.117	14.618	0.000	0.553	1.809
	Horizontal infrared radiation	35,821.811	1622.471	1.031	22.079	0.000	0.016	62.031
	Global horizontal radiation	-6515.011	1386.208	-0.144	-4.700	0.000	0.038	26.628
	Direct normal radiation	4723.975	747.304	0.113	6.321	0.000	0.109	9.146
	Diffuse radiation	11,820.058	1778.320	0.128	6.647	0.000	0.095	10.484
	Wind speed	-294501.046	40,293.443	-0.050	-7.309	0.000	0.753	1.328
	Total sky cover	-2154444.294	95,068.697	-0.957	-22.662	0.000	0.020	50.797
	Visibility	123613.977	102564.923	0.009	1.205	0.228	0.669	1.495
Time-lag1	(Constant)	-181234643.381	11755825.707	—	-15.417	0.000	—	—
	Dry bulb temperatures	-614138.021	43,157.633	-0.575	-14.230	0.000	0.021	47.050
	Dew point temperatures	-68945.791	45,794.132	-0.072	-1.506	0.132	0.015	66.368
	Relative humidity	-1938.780	12,763.123	-0.004	-0.152	0.879	0.044	22.625
	Atmospheric station pressure	1780.718	113.449	0.124	15.696	0.000	0.553	1.809
	Horizontal infrared radiation	29,704.944	1612.435	0.855	18.422	0.000	0.016	62.032
	Global horizontal radiation	-24780.315	1377.584	-0.547	-17.988	0.000	0.038	26.627
	Direct normal radiation	17,040.575	742.657	0.409	22.945	0.000	0.109	9.146
	Diffuse radiation	32,792.426	1767.248	0.354	18.556	0.000	0.095	10.484
	Wind speed	-348289.011	40,047.290	-0.059	-8.697	0.000	0.753	1.328
	Total sky cover	-1728577.821	94,481.810	-0.768	-18.295	0.000	0.020	50.800
	Visibility	-64173.273	101930.905	-0.005	-0.630	0.529	0.669	1.495
Time-lag2	(Constant)	-179888432.974	11279313.576	—	-15.949	0.000	—	—
	Dry bulb temperatures	-685748.939	41,405.258	-0.642	-16.562	0.000	0.021	47.042
	Dew point temperatures	86,402.645	43,935.195	0.091	1.967	0.049	0.015	66.365
	Relative humidity	-47343.593	12,245.089	-0.104	-3.866	0.000	0.044	22.625
	Atmospheric station pressure	1811.848	108.852	0.127	16.645	0.000	0.553	1.810
	Horizontal infrared radiation	24,711.629	1547.029	0.711	15.974	0.000	0.016	62.033
	Global horizontal radiation	-49815.842	1321.653	-1.099	-37.692	0.000	0.038	26.627
	Direct normal radiation	34,307.818	712.510	0.823	48.151	0.000	0.109	9.146
	Diffuse radiation	61,552.875	1695.489	0.664	36.304	0.000	0.095	10.484
	Wind speed	-371456.480	38,421.767	-0.063	-9.668	0.000	0.753	1.328
	Total sky cover	-1367478.673	90,650.584	-0.608	-15.085	0.000	0.020	50.803
	Visibility	-544550.045	97,794.601	-0.038	-5.568	0.000	0.669	1.495

^aDependent variable: heating load of a DOE small office building.

TABLE 6: Model summary for a DOE small office building depending on different models.

Model summary ^a										
Model	<i>R</i>	<i>R</i> square	Adjusted <i>R</i> square	Std. error of the estimate	<i>R</i> square change	Change statistics			Sig. <i>F</i> change	Durbin-Watson
						<i>F</i> change	df1	df2		
Modified model ^b	0.592 ^b	0.350	0.350	7116200.437	0.350	996.953	11	20,338	0.000	0.778
New model ^c	0.422 ^c	0.178	0.176	3742573.281	0.178	130.924	8	4845	0.000	1.509
Limited model ^d	0.513 ^d	0.264	0.263	7575212.077	0.264	1213.784	6	20,343	0.000	0.767

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures; ^cpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures; ^dpredictors: (constant), wind speed, dry bulb temperatures, relative humidity, global horizontal radiation, atmospheric station pressure, and dew point temperatures.

TABLE 7: ANOVA for a DOE small office building depending on different models.

ANOVA ^a						
Model		Sum of squares	df	Mean square	F	Sig.
Modified model ^b	Regression	555346313497669060.000	11	50486028499788096.000	996.953	0.000 ^b
	Residual	1029922597401856130.000	20,338	50640308653842.860	—	—
	Total	1585268910899525120.000	20,349	—	—	—
New model ^c	Regression	14670632319938900.000	8	1833829039992363.000	130.924	0.000 ^c
	Residual	67863211318381720.000	4845	14006854761275.896	—	—
	Total	82533843638320624.000	4853	—	—	—
Limited model ^d	Regression	417909494247049020.000	6	69651582374508168.000	1213.784	0.000 ^d
	Residual	1167359416652476160.000	20,343	57383838010739.625	—	—
	Total	1585268910899525120.000	20,349	—	—	—

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, relative humidity, atmospheric station pressure, diffuse radiation, dew point temperatures, global horizontal radiation, and horizontal infrared radiation; ^dpredictors: (constant), wind speed, dry bulb temperatures, relative humidity, global horizontal radiation, atmospheric station pressure, and dew point temperatures.

TABLE 8: Residual statistics for a DOE small office building depending on different models.

Residual statistics ^a						
		Minimum	Maximum	Mean	Std. deviation	N
Modified model ^b	Predicted value	−16410367.00	28590780.00	3786390.10	5224087.118	20,350
	Residual	−20649606.000	64699612.000	0.000	7114276.785	20,350
	Std. predicted value	−3.866	4.748	0.000	1.000	20,350
	Std. residual	−2.902	9.092	0.000	1.000	20,350
New model ^c	Predicted value	−2313629.00	13152829.00	1123572.72	1738678.448	4854
	Residual	−10689199.000	46347212.000	0.000	3739487.258	4854
	Std. predicted value	−1.977	6.919	0.000	1.000	4854
	Std. residual	−2.856	12.384	0.000	0.999	4854
Limited model ^d	Predicted value	−8359116.00	17571524.00	3786390.10	4531787.978	20,350
	Residual	−15785886.000	70738864.000	0.000	7574095.201	20,350
	Std. predicted value	−2.680	3.042	0.000	1.000	20,350
	Std. residual	−2.084	9.338	0.000	1.000	20,350

^aDependent variable: heating load of a DOE small office building. ^bPredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, relative humidity, atmospheric station pressure, diffuse radiation, dew point temperatures, global horizontal radiation, and horizontal infrared radiation; ^dpredictors: (constant), wind speed, dry bulb temperatures, relative humidity, global horizontal radiation, atmospheric station pressure, and dew point temperatures.

that the output results from EnergyPlus are for a time that is in sync with the time of the weather data used. Second, when the time-lag is one (time-lag1), meaning that the output results from EnergyPlus are for a time that is one hour ahead of the time of the weather data used. Finally, when the time-lag is two (time-lag2), meaning that the output results from EnergyPlus are for a time that is two hours ahead of the time of the weather data used. This procedure is done for each building size (small, medium, and large DOE reference buildings). Figure 2 below explains the study procedure.

4. Results

4.1. Multiple Regression Analysis for the Prediction of the Heating Load for a Small Office considering Time-Lag. To estimate the heating load of a small office building, 20,352 weather data points were used. In addition, the influence of outdoor climatic conditions on the heating load was assessed through multiple regression analysis. Without considering

time-lag, the heating load values extracted from EnergyPlus after one hour and two hours were matched separately by hour during the analysis. The results show R-values of the time-lag0 model, time-lag1 model, and time-lag2 model to be 0.534, 0.542, and 0.592, respectively. In addition, the R-square values of time-lag0, time-lag1, and time-lag2 were 0.285, 0.294, and 0.350, respectively; this implies that the regression model for the DOE small office has a higher explanatory or predictive power when using a time-lag of two hours. Table 3 shows the results of the small office model summary considering time-lag.

The *F*-variations for all the developed models were higher than 700, and the *p* values were all less than 0.001, thus confirming the statistical significance of the models. Furthermore, the Durbin–Watson test was conducted, and the results were 0.799, 0.834, and 0.778 for the time-lag0, time-lag1, and time-lag2 models, respectively; this means that among the three developed models, the time-lag1 model has a higher estimating power as it has a higher R-value and Durbin–Watson value.

TABLE 9: Regression coefficients for a DOE small office building depending on different models.

Coefficients ^a								
Model	Unstandardized coefficients		Standardized coefficients	<i>t</i>	Sig.	Collinearity statistics		
	<i>B</i>	Std. error	Beta			Tolerance	VIF	
Modified model	(Constant)	−179888432.974	11279313.576	—	−15.949	0.000	—	—
	Dry bulb temperatures	−685748.939	41,405.258	−0.642	−16.562	0.000	0.021	47.042
	Dew point temperatures	86,402.645	43,935.195	0.091	1.967	0.049	0.015	66.365
	Relative humidity	−47343.593	12,245.089	−0.104	−3.866	0.000	0.044	22.625
	Atmospheric station pressure	1811.848	108.852	0.127	16.645	0.000	0.553	1.810
	Horizontal infrared radiation	24,711.629	1547.029	0.711	15.974	0.000	0.016	62.033
	Global horizontal radiation	−49815.842	1321.653	−1.099	−37.692	0.000	0.038	26.627
	Direct normal radiation	34,307.818	712.510	0.823	48.151	0.000	0.109	9.146
	Diffuse radiation	61,552.875	1695.489	0.664	36.304	0.000	0.095	10.484
	Wind speed	−371456.480	38,421.767	−0.063	−9.668	0.000	0.753	1.328
	Total sky cover	−1367478.673	90,650.584	−0.608	−15.085	0.000	0.020	50.803
	Visibility	−544550.045	97,794.601	−0.038	−5.568	0.000	0.669	1.495
New model	(Constant)	−84468362.750	11178380.426	—	−7.556	0.000	—	—
	Dew point temperatures	−153900.636	11,782.451	−0.332	−13.062	0.000	0.263	3.800
	Atmospheric station pressure	820.654	109.775	0.123	7.476	0.000	0.626	1.599
	Horizontal infrared radiation	7302.493	1652.577	0.423	4.419	0.000	0.019	54.035
	Global horizontal radiation	−28329.843	5547.379	−0.340	−5.107	0.000	0.038	26.094
	Direct normal radiation	22,573.966	1829.080	0.430	12.342	0.000	0.140	7.149
	Diffuse radiation	31,803.161	6859.159	0.208	4.637	0.000	0.084	11.876
	Total sky cover	−380750.599	94,783.568	−0.349	−4.017	0.000	0.023	44.375
	AOD	−215562.688	100939.898	−0.029	−2.136	0.033	0.945	1.059
Limited model	(Constant)	−198730470.861	11744538.154	—	−16.921	0.000	—	—
	Dry bulb temperatures	−717138.604	42,245.667	−0.671	−16.975	0.000	0.023	43.216
	Dew point temperatures	296105.476	45,462.975	0.310	6.513	0.000	0.016	62.710
	Relative humidity	−90189.068	12,260.853	−0.198	−7.356	0.000	0.050	20.018
	Atmospheric station pressure	2096.977	113.658	0.146	18.450	0.000	0.574	1.741
	Global horizontal radiation	8298.520	340.007	0.183	24.407	0.000	0.643	1.555
	Wind speed	−241986.775	40,515.794	−0.041	−5.973	0.000	0.767	1.303

^aDependent variable: heating load of a DOE small office building.

The results of the ANOVA analysis shown in Table 4 indicate a *p* value less than 0.05 for all the models; thus, all the models are assumed to be statistically significant. The time-lag2 model showed the largest variance, whereas the time-lag0 model showed the largest variance of residue. Additionally, the *F*-value belonging to the time-lag2 model was the highest among the three developed models, thus indicating that the time-lag2 model has the best fit.

Table 5 shows the tolerance limit, significance of probability, and *t*-value of the variables for each model. For accuracy, if a high goodness-of-fit is found for the independent variables, the *t*-value and *p* value should meet certain criteria: *t*-value ≥ 1.96 and *p* < 0.05. In addition, to avoid issues related to multicollinearity, the tolerance limit should be 0.1 or higher; variables that do not satisfy these conditions are usually omitted from the analysis to improve the accuracy of the developed models.

Consequently, visibility was omitted from the time-lag0 model, whereas the dew point temperature, relative

humidity, and visibility were also removed from the time-lag1 model; these variables are thus absent from the final models.

As indicated by the results after taking into account the *t*-value, significance of the probability, and multicollinearity issues, the time-lag2 model was found to be the most suitable for a DOE small office building; this model is defined as the “Modified Model” and is presented in Table 6. Secondly, another model is developed based on all the variables considered in the time-lag2 model, AOD, and precipitable water values; this model is defined as the “New Model,” and it is also presented in Table 6. Additionally, a third model based on only those variables received from the Korea Meteorological Administration (KMA) and which can be input by the user into EnergyPlus was developed considering a time-lag of 2 hours; this third model is defined as the “Limited Model” and is shown in Table 6.

A multiple regression analysis was performed to predict the heating thermal load. As indicated by the obtained

TABLE 10: Model summary for a DOE medium office building depending on time-lag.

Model	Model summary ^a									
	<i>R</i>	<i>R</i> square	Adjusted <i>R</i> square	Std. error of the estimate	<i>R</i> square change	Change statistics <i>F</i> change	df1	df2	Sig. <i>F</i> change	Durbin–Watson
Time-lag0	0.537 ^b	0.289	0.288	12945018.511	0.289	750.955	11	20,340	0.000	0.706
Time-lag1	0.537 ^b	0.289	0.288	12947428.028	0.289	750.070	11	20,339	0.000	0.731
Time-lag2	0.579 ^b	0.335	0.335	12518253.739	0.335	931.428	11	20,338	0.000	0.677

^aDependent variable: heating load (medium); ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

TABLE 11: Residual summary for a DOE medium office building depending on time-lag.

ANOVA ^a						
Model		Sum of squares	df	Mean square	<i>F</i>	Sig.
Time-lag0	Regression	1384242517670826240.000	11	125840228879166016.000	750.955	0.000 ^b
	Residual	3408445076552884200.000	20,340	167573504255304.030	—	—
	Total	4792687594223710200.000	20,351	—	—	—
Time-lag1	Regression	1383124549966582780.000	11	125738595451507520.000	750.070	0.000 ^b
	Residual	3409546418166475800.000	20,339	167635892529941.300	—	—
	Total	4792670968133058600.000	20,350	—	—	—
Time-lag2	Regression	1605570375500446210.000	11	145960943227313184.000	931.428	0.000 ^b
	Residual	3187100390118337000.000	20,338	156706676670190.620	—	—
	Total	4792670765618783200.000	20,349	—	—	—

^aDependent variable: heating load of DOE medium office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

results, the “New Model” showed the minimum *R*-squared (0.422) value and the highest Durbin–Watson value (1.509); in this case, a Durbin–Watson value above 1.3 indicates a weak correlation between the variables. In addition, although the *R*-squared values for the “Modified Model” and “Limited Model” are higher than that of the “New Model,” multicollinearity issues were found for the two models.

The *F*-variation of the “New Model” was shown to be higher than 170 at $p < 0.001$. Furthermore, the variance of the regression model and residual were smaller than those of both the “Modified Model” and “Limited Model.” The *F*-value was shown to be highest in the “Limited Model.” Additionally, the results predicted by the “New Model” exhibited the smallest mean and standard deviation, as well as the smallest difference between the maximum and minimum predictions (Tables 7–9). A regression equation for the “New Model” can be drafted based on the final variables listed in Table 6.

4.2. Multiple Regression Analysis for the Prediction of Heating Load for a Medium Office considering Time-Lag. For a DOE medium office building, 20,352 weather data points were used to estimate the heating load. In addition, the influence of the outdoor climatic conditions on the heating load was assessed through multiple regression analysis. Without considering time-lag, heating load values extracted from

EnergyPlus after one hour and two hours were matched separately by hour during the analysis. The results of the analysis together with the input variables describing the climatic conditions are shown in Tables 10–13. The results show the *R*-values of the time-lag0 model, time-lag1 model, and time-lag2 model to be 0.537, 0.537, and 0.579, respectively. In addition, the *R*-square values of time-lag0, time-lag1, and time-lag2 were 0.289, 0.289, and 0.335, respectively; this implies that the regression model for the DOE medium office has a higher explanatory or predictive power when using a time-lag of two hours. Table 10 shows the model summaries of a DOE medium office building considering time-lag.

The *F*-variation for all the developed models was higher than 750, and the p values were all less than 0.001, thus confirming the statistical significance of the models. Furthermore, the Durbin–Watson test was conducted, and the results were 0.706, 0.731, and 0.677 for the time-lag0 model, time-lag1 model, and time-lag2 model, respectively; this means that among the three developed models, the time-lag1 model has the highest estimating power as it had the highest *R*-value and Durbin–Watson value, as shown in Table 11.

The results of the ANOVA analysis indicate a p value less than 0.05 for all models, and thus, all the models are assumed to be statistically significant. The time-lag2 model

TABLE 12: Model summary for a DOE medium office building depending on time-lag.

		Coefficients ^a		t	Sig.	Collinearity statistics	
Model	Unstandardized coefficients		Standardized coefficients				
	B	Std. Error	Beta			Tolerance	VIF
Time-lag0	(Constant)	-232365034.136	20514208.737	—	-11.327	0.000	—
	Dry bulb temperatures	-1239414.048	75,319.865	-0.667	-16.455	0.000	0.021
	Dew point temperatures	-354626.266	79,919.948	-0.214	-4.437	0.000	0.015
	Relative humidity	66,371.259	22,274.126	0.084	2.980	0.003	0.044
	Atmospheric station pressure	2116.094	197.968	0.085	10.689	0.000	0.553
	Horizontal infrared radiation	99,198.849	2813.952	1.642	35.252	0.000	0.016
	Global horizontal radiation	-3376.020	2404.186	-0.043	-1.404	0.160	-0.038
	Direct normal radiation	788.013	1296.095	0.011	0.608	0.543	0.109
	Diffuse radiation	13,477.568	3084.249	0.084	4.370	0.000	0.095
	Wind speed	-479633.798	69,883.408	-0.047	-6.863	0.000	0.753
	Total sky cover	-6135755.540	164883.515	-1.568	-37.213	0.000	0.020
	Visibility	408798.727	177884.682	0.017	2.298	0.022	0.669
Time-lag1	(Constant)	-242363953.699	20520409.598	—	-11.811	0.000	—
	Dry bulb temperatures	-1289408.649	75,333.909	-0.694	-17.116	0.000	0.021
	Dew point temperatures	-155279.237	79,936.057	-0.094	-1.943	0.052	0.015
	Relative humidity	16,763.249	22,278.700	0.021	0.752	0.452	0.044
	Atmospheric station pressure	2277.502	198.032	0.091	11.501	0.000	0.553
	Horizontal infrared radiation	87,070.308	2814.591	1.441	30.935	0.000	0.016
	Global horizontal radiation	-39853.596	2404.644	-0.506	-16.574	0.000	0.038
	Direct normal radiation	25,968.009	1296.347	0.358	20.032	0.000	0.109
	Diffuse radiation	54,012.154	3084.824	0.335	17.509	0.000	0.095
	Wind speed	-616322.819	69,904.642	-0.060	-8.817	0.000	0.753
	Total sky cover	-5285555.158	164922.949	-1.351	-32.049	0.000	0.020
	Visibility	35,564.622	177925.734	0.001	0.200	0.842	0.669
Time-lag2	(Constant)	-242862773.124	19841671.213	—	-12.240	0.000	—
	Dew point temperatures	102726.330	77,287.300	0.062	1.329	0.184	0.015
	Relative humidity	-57775.571	21,540.586	-0.073	-2.682	0.007	0.044
	Atmospheric station pressure	2360.095	191.484	0.095	12.325	0.000	0.553
	Horizontal infrared radiation	76,361.042	2721.411	1.264	28.059	0.000	0.016
	Global horizontal radiation	-82461.104	2324.947	-1.047	-35.468	0.000	0.038
	Direct normal radiation	56,837.692	1253.391	0.784	45.347	0.000	0.109
	Diffuse radiation	101106.744	2982.570	0.628	33.899	0.000	0.095
	Total sky cover	-4510822.624	159465.297	-1.153	-28.287	0.000	0.020
	Dry bulb temperatures	-1387805.281	72,836.836	-0.747	-19.054	0.000	0.021
	Wind speed	-677958.805	67,588.517	-0.066	-10.031	0.000	0.753
	Visibility	-799488.056	172032.483	-0.032	-4.647	0.000	0.669

^aDependent variable: heating load of a DOE medium office building.

TABLE 13: Residual statistics for a DOE medium office building depending on time-lag.

		Residual statistics ^a				
		Minimum	Maximum	Mean	Std. deviation	N
Time-lag0 ^b	Predicted value	-59638412.00	34621280.00	5299548.66	8247327.016	20,352
	Residual	-31375230.000	142983696.000	0.000	12941519.557	20,352
	Std. predicted value	-7.874	3.555	0.000	1.000	20,352
	Std. residual	-2.424	11.045	0.000	1.000	20,352
Time-lag1 ^b	Predicted value	-53244212.00	40414716.00	5299348.31	8244198.466	20,351
	Residual	-28412308.000	145224064.000	0.000	12943928.250	20,351
	Std. predicted value	-7.101	4.259	0.000	1.000	20,351
	Std. residual	-2.194	11.216	0.000	1.000	20,351
Time-lag2 ^b	Predicted value	-47442988.00	48511468.00	5299326.20	8882662.010	20,350
	Residual	-35604016.000	136248016.000	0.000	12514869.803	20,350
	Std. predicted value	-5.938	4.865	0.000	1.000	20,350
	Std. residual	-2.844	10.884	0.000	1.000	20,350

^aDependent variable: heating load of a DOE medium office building. ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

TABLE 14: Model summary for a DOE medium office building depending on different models.

Model summary ^a										
Model	<i>R</i>	<i>R</i> square	Adjusted <i>R</i> square	Std. error of the estimate	<i>R</i> square change	Change statistics			Sig. <i>F</i> change	Durbin-Watson
						<i>F</i> change	df1	df2		
Modified model ^b	0.568 ^b	0.323	0.323	12629176.660	0.323	970.983	10	20,339	0.000	0.671
New model ^c	0.458 ^c	0.210	0.209	6510053.800	0.210	159.315	8	4798	0.000	1.227
Limited model ^d	0.485 ^d	0.235	0.235	13421600.463	0.235	894.764	7	20,342	0.000	0.655

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dew point temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, atmospheric station pressure, dew point temperatures, dry bulb temperatures, and horizontal infrared radiation; ^dpredictors: (constant), total sky cover, wind speed, dry bulb temperatures, global horizontal radiation, relative humidity, atmospheric station pressure, and dew point temperatures.

TABLE 15: ANOVA for a DOE medium office building depending on different models.

ANOVA ^a						
Model		Sum of squares	df	Mean square	<i>F</i>	Sig.
Modified model ^b	Regression	1548679524700758530.000	10	154867952470075808.000	970.983	0.000 ^b
	Residual	3243991240918024700.000	20,339	159496103098383.620	—	—
	Total	4792670765618783200.000	20,349	—	—	—
New model ^c	Regression	54015103645277112.000	8	6751887955659638.000	159.315	0.000 ^c
	Residual	203343080683039488.000	4798	42380800475831.500	—	—
	Total	257358184328316608.000	4806	—	—	—
Limited model ^d	Regression	1128275925296639100.000	7	161182275042376992.000	894.764	0.000 ^d
	Residual	3664394840322144300.000	20,342	180139358977590.400	—	—
	Total	4792670765618783200.000	20,349	—	—	—

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dew point temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, atmospheric station pressure, dew point temperatures, dry bulb temperatures, and horizontal infrared radiation; ^dpredictors: (constant), total sky cover, wind speed, dry bulb temperatures, global horizontal radiation, relative humidity, atmospheric station pressure, and dew point temperatures.

TABLE 16: Residual statistics for a DOE medium office building depending on different models.

Residual statistics ^a						
		Minimum	Maximum	Mean	Std. deviation	<i>N</i>
Modified model ^b	Predicted value	−41989976.00	47291244.00	5299326.20	8723871.147	20,350
	Residual	−34259568.000	135437520.000	0.000	12626073.134	20,350
	Std. predicted value	−5.421	4.813	0.000	1.000	20,350
	Std. residual	−2.713	10.724	0.000	1.000	20,350
New model ^c	Predicted value	−5426438.00	19135408.00	1919966.69	3352476.356	4807
	Residual	−19135408.000	73064632.000	0.000	6504633.271	4807
	Std. predicted value	−2.191	5.135	0.000	1.000	4807
	Std. residual	−2.939	11.223	0.000	0.999	4807
Limited model ^d	Predicted value	−15162106.00	27670688.00	5299326.20	7446224.483	20,350
	Residual	−22803048.000	150226688.000	0.000	13419291.767	20,350
	Std. predicted value	−2.748	3.004	0.000	1.000	20,350
	Std. residual	−1.699	11.193	0.000	1.000	20,350

^aDependent variable: heating load of a DOE medium office building; ^bPredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dew point temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, atmospheric station pressure, dew point temperatures, dry bulb temperatures, horizontal infrared radiation; ^dpredictors: (constant), total sky cover, wind speed, dry bulb temperatures, global horizontal radiation, relative humidity, atmospheric station pressure, dew point temperatures.

TABLE 17: Regression coefficients for a DOE medium office building depending on different models.

		Coefficients ^a						
Model		Unstandardized coefficients		Standardized coefficients	<i>t</i>	Sig.	Collinearity statistics	
		<i>B</i>	Std. error	Beta			Tolerance	VIF
Modified model	(Constant)	−332178078.900	19450841.631	—	−17.078	0.000	—	—
	Dew point temperatures	−1278243.171	27,075.391	−0.770	−47.211	0.000	0.125	8.002
	Relative humidity	318029.341	8735.744	0.402	36.406	0.000	0.274	3.656
	Atmospheric station pressure	2937.887	190.743	0.118	15.402	0.000	0.567	1.764
	Horizontal infrared radiation	66,114.532	2691.386	1.094	24.565	0.000	0.017	59.610
	Global horizontal radiation	−90530.015	2306.310	−1.149	−39.253	0.000	0.039	25.744
	Direct normal radiation	59,029.740	1259.159	0.814	46.880	0.000	0.110	9.069
	Diffuse radiation	109882.074	2972.907	0.682	36.961	0.000	0.098	10.234
	Total sky cover	−3968922.089	158298.925	−1.014	−25.072	0.000	0.020	49.187
	Wind speed	−640957.692	68,159.260	−0.062	−9.404	0.000	0.754	1.327
	Visibility	−340103.073	171843.921	−0.014	−1.979	0.048	0.682	1.466
New model	(Constant)	−137086417.229	19542748.049	—	−7.015	0.000	—	—
	Dew point temperatures	−112436.116	27,307.607	−0.136	−4.117	0.000	0.151	6.625
	Atmospheric station pressure	1264.005	192.177	0.107	6.577	0.000	0.623	1.604
	Horizontal infrared radiation	47,644.241	3375.835	1.556	14.113	0.000	0.014	73.809
	Direct normal radiation	16,462.981	1213.513	0.178	13.566	0.000	0.962	1.040
	Total sky cover	−2787320.969	196312.532	−1.437	−14.198	0.000	0.016	62.171
	Dry bulb temperatures	−581774.230	37,805.690	−0.586	−15.389	0.000	0.114	8.794
	AOD	−566944.126	182564.249	−0.042	−3.105	0.002	0.879	1.137
	Precipitable water	1076401.110	240985.736	0.080	4.467	0.000	0.512	1.955
Limited model	(Constant)	−263437804.945	21012890.768	—	−12.537	0.000	—	—
	Dry bulb temperatures	−1285938.756	75,197.392	−0.692	−17.101	0.000	0.023	43.618
	Dew point temperatures	582520.586	80,798.698	0.351	7.210	0.000	0.016	63.097
	Relative humidity	−152490.019	22,022.183	−0.193	−6.924	0.000	0.049	20.572
	Atmospheric station pressure	2831.572	203.680	0.114	13.902	0.000	0.561	1.781
	Global horizontal radiation	13,177.043	602.849	0.167	21.858	0.000	0.642	1.557
	Wind speed	−445559.152	71,983.529	−0.043	−6.190	0.000	0.763	1.310
	Total sky cover	−153324.249	27,559.024	−0.039	−5.563	0.000	0.758	1.320

^aDependent variable: heating load of a DOE medium office building.

TABLE 18: Model summary for a DOE large office building depending on time-lag.

Model	Model summary ^a									
	<i>R</i>	<i>R</i> square	Adjusted <i>R</i> square	Std. error of the estimate	<i>R</i> square change	Change statistics			Sig. <i>F</i> change	Durbin–Watson
						<i>F</i> change	df1	df2		
Time-lag0	0.514 ^b	0.264	0.264	172903173.811	0.264	663.260	11	20,340	0.000	0.732
Time-lag1	0.516 ^b	0.266	0.266	172668699.008	0.266	670.178	11	20,339	0.000	0.761
Time-lag2	0.579 ^b	0.335	0.335	12518253.739	0.335	931.428	11	20,338	0.000	0.677

^aDependent variable: heating load of a DOE large office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

TABLE 19: ANOVA for a DOE large building depending on time-lag.

ANOVA ^a						
Model		Sum of squares	df	Mean square	F	Sig.
Time-lag0 ^b	Regression	218113322176116700000.000	11	19828483834192430000.000	663.260	0.000 ^b
	Residual	608074622833789500000.000	20,340	29895507513952288.000	—	—
	Total	826187945009906200000.000	20,351	—	—	—
Time-lag1 ^b	Regression	219791177256880500000.000	11	19981016114261864000.000	670.178	0.000 ^b
	Residual	606396700934593400000.000	20,339	29814479617217832.000	—	—
	Total	826187878191473900000.000	20,350	—	—	—
Time-lag2 ^b	Regression	1605570375500446210.000	11	145960943227313184.000	931.428	0.000 ^b
	Residual	3187100390118337000.000	20,338	156706676670190.620	—	—
	Total	4792670765618783200.000	20,349	—	—	—

^aDependent variable: heating load of DOE large office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, horizontal infrared radiation, and dew point temperatures.

showed the largest variance with the smallest variance in the residuals. In addition, the F-value belonging to the time-lag2 model was the highest among the three developed models and thus indicated that the time-lag2 model has the best fit.

Table 12 shows the tolerance limit, significance of the probability, and *t*-value of the variables for each model. For accuracy, if a high goodness-of-fit is found for the independent variables, the *t*-value and *p* value should meet certain criteria: *t*-value ≥ 1.96 and *p* < 0.05. In addition, to avoid issues related to multicollinearity, the tolerance limit should be 0.1 or higher; variables that do not satisfy these conditions are usually omitted from the analysis to improve the accuracy of the developed models.

Since the time-lag phenomenon was not taken into consideration in the time-lag0 model, the most important variables when it comes to thermal conditions, such as global horizontal radiation and direct normal radiation, failed to satisfy the conditions. The dew point temperature and relative humidity in the time-lag1 model and the dew point temperature in the time-lag2 model were the variables found to be rejected from each model. These variables are thus absent from the final models. Table 13 shows residual statistics for a DOE medium office building considering time-lag.

As indicated by the results after taking into account the *t*-value, significance of the probability, and multicollinearity issues, the time-lag2 model was found to be the most suitable for a DOE medium office building; this model is defined as the “Modified Model” and is presented in Table 14. Secondly, another model is developed based on all the variables considered in the time-lag2 model, AOD, and precipitable water values; this model is defined as the “New Model,” and it is also presented in Table 14. Additionally, a third model based on only those variables received from the Korea Meteorological Administration (KMA) and which can be input by the user into EnergyPlus was developed considering a time-lag of 2 hours; this third model is defined as the “Limited Model.”

Multiple regression analysis was performed to predict the heating thermal load. As indicated by the obtained results, the “New Model” showed the minimum *R*-squared

(0.422) value and the highest Durbin–Watson value (1.509); in this case, a Durbin–Watson value above 1.3 indicates a weak correlation between the variables. In addition, although the *R*-squared values for the “Modified Model” and “Limited Model” are higher than that of the “New Model,” multicollinearity issues were found for the two models.

The “Modified Model” is a good explanation of the dependent variable, as it exhibits the highest *R*-values. The variables considered in the “Modified Model” also exhibit positive correlations, as indicated by the Durbin–Watson value (0.671). The “Limited Model” indicated *R*-values lower than those of the “Modified Model” but higher than those of the “New Model” by a value of 0.03. The “New Model” showed the lowest *R*-value (0.458) and thus is an indication of the lowest explanatory power.

Tables 15–17 show the statistical results of the new regression model dealing with AOD. As shown in the table, the *F*-variation of the model was higher than 159 at *p* < 0.001. Its variance and residual are slightly lower than that of the “Modified Model.” In addition, the *F*-values of the “Modified Model” were shown to be the highest of all three models. The “Limited Model” indicated the largest difference between the maximum and minimum predictions, as well as the largest mean.

4.3. Multiple Regression Analysis for the Prediction of the Heating Load for a Large Office considering Time-Lag. For a DOE large office building, 20,352 weather data points were used to estimate the heating load. In addition, the influence of the outdoor climatic conditions on the heating load was assessed through multiple regression analysis. Without considering time-lag, heating load values extracted from EnergyPlus after one hour and two hours were matched separately by hour during the analysis. The results of the analysis together with the input variables describing the climatic conditions are shown in Tables 18–21. The results show *R*-values of the time-lag0 model, time-lag1 model, and time-lag2 model to be 0.514, 0.516, and 0.579, respectively. In addition, the *R*-square values of time-lag0, time-lag1, and time-lag2 were 0.264, 0.266, and 0.335, respectively; this implies that the regression model for the DOE large office

TABLE 20: Regression coefficients for a DOE large office building depending on time-lag.

		Coefficients ^a						
Model	Unstandardized coefficients		Standardized coefficients	<i>t</i>	Sig.	Collinearity statistics		
	<i>B</i>	Std. error	Beta			Tolerance	VIF	
Time-lag0	(Constant)	−273311537.015	274002837.132	—	−9.975	0.000	—	—
	Dry bulb temperatures	−15671733.697	1006027.437	−0.643	−15.578	0.000	0.021	47.057
	Dew point temperatures	−4753374.820	1067469.519	−0.218	−4.453	0.000	0.015	66.370
	Relative humidity	990564.362	297509.579	0.095	3.330	0.001	0.044	22.624
	Atmospheric station pressure	24,491.862	2644.212	0.075	9.262	0.000	0.553	1.809
	Horizontal infrared radiation	1277515.990	37,585.214	1.610	33.990	0.000	0.016	62.031
	Global horizontal radiation	−33027.580	32,112.069	−0.032	−1.029	0.304	0.038	26.628
	Direct normal radiation	9862.429	17,311.599	0.010	0.570	0.569	0.109	9.146
	Diffuse radiation	163741.885	41,195.501	0.077	3.975	0.000	0.095	10.484
	Wind speed	−5511093.363	933414.112	−0.041	−5.904	0.000	0.753	1.328
	Total sky cover	−78732270.417	2202305.310	−1.533	−35.750	0.000	0.020	50.797
	Visibility	6728944.425	2375958.447	0.021	2.832	0.005	0.669	1.495
Time-lag1	(Constant)	−2823375262.525	273663033.374	—	−10.317	0.000	—	—
	Dry bulb temperatures	−16209199.274	1004663.472	−0.665	−16.134	0.000	0.021	47.050
	Dew point temperatures	−2245663.075	1066038.360	−0.103	−2.107	0.035	0.015	66.368
	Relative humidity	362476.854	297111.842	0.035	1.220	0.222	0.044	22.625
	Atmospheric station pressure	26,181.850	2640.978	0.080	9.914	0.000	0.553	1.809
	Horizontal infrared radiation	1119195.915	37,535.772	1.411	29.817	0.000	0.016	62.032
	Global horizontal radiation	−538804.813	32,068.672	−0.521	−16.802	0.000	0.038	26.627
	Direct normal radiation	355624.778	17,288.268	0.374	20.570	0.000	0.109	9.146
	Diffuse radiation	718224.310	41,139.637	0.340	17.458	0.000	0.095	10.484
	Wind speed	−7333997.059	932258.024	−0.054	−7.867	0.000	0.753	1.328
	Total sky cover	−67661397.473	2199435.361	−1.317	−30.763	0.000	0.020	50.800
	Visibility	1918837.055	2372842.304	0.006	0.809	0.419	0.669	1.495
Time-lag2	(Constant)	−242862773.124	19841671.213	—	−12.240	0.000	—	—
	Dew point temperatures	102726.330	77,287.300	0.062	1.329	0.184	0.015	66.365
	Atmospheric station pressure	2360.095	191.484	0.095	12.325	0.000	0.553	1.810
	Horizontal infrared radiation	76,361.042	2721.411	1.264	28.059	0.000	0.016	62.033
	Direct normal radiation	56,837.692	1253.391	0.784	45.347	0.000	0.109	9.146
	Total sky cover	−4510822.624	159465.297	−1.153	−28.287	0.000	0.020	50.803
	Dry bulb temperatures	−1387805.281	72,836.836	−0.747	−19.054	0.000	0.021	47.042
	Relative humidity	−57775.571	21,540.586	−0.073	−2.682	0.007	0.044	22.625
	Global horizontal radiation	−82461.104	2324.947	−1.047	−35.468	0.000	0.038	26.627
	Diffuse radiation	101106.744	2982.570	0.628	33.899	0.000	0.095	10.484
	Wind speed	−677958.805	67,588.517	−0.066	−10.031	0.000	0.753	1.328
	Visibility	−799488.056	172032.483	−0.032	−4.647	0.000	0.669	1.495

^aDependent variable: heating load of a DOE large office building.

has a higher explanatory or predictive power when using a time-lag of two hours.

All regression models developed in the study show *F*-variations higher than 660 at $p < 0.001$. In addition, the time-lag2 model indicated the highest *R*-value (0.931) amongst the three models. The Durbin–Watson test showed values of 0.732, 0.761, and 0.677 for the time-lag1, time-lag2, and time-lag3 models, respectively.

The results of the ANOVA analysis indicate a *p* value less than 0.05 for all the models, and thus, all the models are assumed to be statistically significant. The time-lag2 model

showed the largest variance with the smallest variance in the residuals. In addition, the *F*-value belonging to the time-lag2 model was the highest among the three developed models and thus indicated that the time-lag2 model has the best fit.

Table 20 shows the tolerance limit, significance of the probability, and *t*-value of the variables for each model. For accuracy, if a high goodness-of-fit is found for the independent variables, the *t*-value and *p* value should meet certain criteria: $t\text{-value} \geq 1.96$ and $p < 0.05$. In addition, to avoid issues related to multicollinearity, the tolerance limit should be 0.1 or higher; variables that do not satisfy these

TABLE 21: Residual statistics for a DOE large office building depending on time-lag.

		Residual statistics ^a				
		Minimum	Maximum	Mean	Std. deviation	<i>N</i>
Time-lag0	Predicted value	-767601856.00	432252768.00	63956011.14	103525710.371	20,352
	Residual	-392874016.000	1979168896.000	0.000	172856439.204	20,352
	Std. predicted value	-8.032	3.558	0.000	1.000	20,352
	Std. residual	-2.272	11.447	0.000	1.000	20,352
Time-lag1	Predicted value	-684187840.00	516718464.00	63956412.79	103925691.005	20,351
	Residual	-356752608.000	2007745792.000	0.000	172622025.484	20,351
	Std. predicted value	-7.199	4.357	0.000	1.000	20,351
	Std. residual	-2.066	11.628	0.000	1.000	20,351
Time-lag2	Predicted value	-47442988.00	48511468.00	5299326.20	8882662.010	20,350
	Residual	-35604016.000	136248016.000	0.000	12514869.803	20,350
	Std. predicted value	-5.938	4.865	0.000	1.000	20,350
	Std. residual	-2.844	10.884	0.000	1.000	20,350

^aDependent variable: heating load of a DOE large office building.

TABLE 22: Model summary for a DOE large office building depending on different models.

Model summary ^a										
Model	<i>R</i>	<i>R</i> square	Adjusted <i>R</i> square	Std. error of the estimate	<i>R</i> square change	Change statistics			Sig. <i>F</i> change	Durbin-Watson
						<i>F</i> change	df1	df2		
Modified model ^b	0.579 ^b	0.335	0.335	12518489.662	0.335	1024.355	10	20,339	0.000	0.678
New model ^c	0.458 ^c	0.210	0.209	6510053.800	0.210	159.315	8	4798	0.000	1.227
Limited model ^d	0.462 ^d	0.214	0.213	178702667.526	0.214	789.878	7	20,342	0.000	0.684

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, atmospheric station pressure, dew point temperatures, dry bulb temperatures, and horizontal infrared radiation; ^dpredictors: (constant), total sky cover, wind speed, dry bulb temperatures, global horizontal radiation, relative humidity, atmospheric station pressure, and dew point temperatures.

TABLE 23: ANOVA for a DOE large office building depending on different models.

ANOVA ^a						
Model		Sum of squares	df	Mean square	<i>F</i>	Sig.
Modified model ^b	Regression	1605293531717734140.000	10	160529353171773408.000	1024.355	0.000 ^b
	Residual	3187377233901048800.000	20,339	156712583406315.400	—	—
	Total	4792670765618783200.000	20,349	—	—	—
New model ^c	Regression	54015103645277112.000	8	6751887955659638.000	159.315	0.000 ^c
	Residual	203343080683039488.000	4798	42380800475831.500	—	—
	Total	257358184328316608.000	4806	—	—	—
Limited model ^d	Regression	176571232917599750000.000	7	25224461845371390000.000	789.878	0.000 ^d
	Residual	649614515656888500000.000	20,342	31934643381028828.000	—	—
	Total	826185748574488200000.000	20,349	—	—	—

^aDependent variable: heating load of a DOE small office building; ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, atmospheric station pressure, dew point temperatures, dry bulb temperatures, and horizontal infrared radiation; ^dpredictors: (constant), total sky cover, wind speed, dry bulb temperatures, global horizontal radiation, relative humidity, atmospheric station pressure, and dew point temperatures.

conditions are usually omitted from the analysis to improve the accuracy of the developed models.

Since the time-lag phenomenon was not taken into consideration in the time-lag0 model, the most important

variables when it comes to thermal conditions, such as global horizontal radiation and direct normal radiation, failed to satisfy the criteria. In addition, the dew point temperature and relative humidity in the time-lag1 model and the dew

TABLE 24: Residual statistics for a DOE large office building depending on different models.

		Residual statistics ^a				
		Minimum	Maximum	Mean	Std. Deviation	N
Modified model ^b	Predicted value	-47685040.00	48544072.00	5299326.20	8881896.171	20,350
	Residual	-35572376.000	136128544.000	0.000	12515413.336	20,350
	Std. predicted value	-5.965	4.869	0.000	1.000	20,350
	Std. residual	-2.842	10.874	0.000	1.000	20,350
New model ^c	Predicted value	-5426438.00	19135408.00	1919966.69	3352476.356	4807
	Residual	-19135408.000	73064632.000	0.000	6504633.271	4807
	Std. predicted value	-2.191	5.135	0.000	1.000	4807
	Std. residual	-2.939	11.223	0.000	0.999	4807
Limited model ^d	Predicted value	-198986816.00	346896704.00	63958680.44	93151196.759	20,350
	Residual	-280478208.000	2066232704.000	0.000	178671928.270	20,350
	Std. predicted value	-2.823	3.037	0.000	1.000	20,350
	Std. residual	-1.570	11.562	0.000	1.000	20,350

^aDependent variable: heating load of a DOE large office building. ^bpredictors: (constant), visibility, diffuse radiation, atmospheric station pressure, wind speed, total sky cover, dry bulb temperatures, direct normal radiation, relative humidity, global horizontal radiation, and horizontal infrared radiation, and dew point temperatures; ^cpredictors: (constant), precipitable water, direct normal radiation, total sky cover, AOD, relative humidity, atmospheric station pressure, diffuse radiation, dew point temperatures, global horizontal radiation, and horizontal infrared radiation; ^dpredictors: (constant), wind speed, dry bulb temperatures, relative humidity, global horizontal radiation, atmospheric station pressure, and dew point temperatures.

TABLE 25: Regression coefficients for a DOE large office building depending on different models.

		Coefficients ^a						
Model		Unstandardized coefficients		Standardized coefficients	<i>t</i>	Sig.	Collinearity statistics	
		<i>B</i>	Std. error	Beta			Tolerance	VIF
Modified Model	(Constant)	-247363320.740	19550986.572	—	-12.652	0.000	—	—
	Atmospheric station pressure	2380.032	190.899	0.096	12.467	0.000	0.556	1.799
	Horizontal infrared radiation	76,728.709	2707.367	1.270	28.341	0.000	0.016	61.392
	Direct normal radiation	56,924.641	1251.706	0.785	45.478	0.000	0.110	9.121
	Total sky cover	-4535356.153	158396.421	-1.159	-28.633	0.000	0.020	50.122
	Dry bulb temperatures	-1297018.379	25,292.663	-0.698	-51.280	0.000	0.176	5.672
	Relative humidity	-30575.346	6723.627	-0.039	-4.547	0.000	0.454	2.204
	Global horizontal radiation	-82908.168	2300.532	-1.052	-36.039	0.000	0.038	26.070
	Diffuse radiation	101684.144	2950.819	0.631	34.460	0.000	0.097	10.261
	Wind speed	-678986.494	67,585.368	-0.066	-10.046	0.000	0.753	1.328
	Visibility	-773429.426	170914.893	-0.031	-4.525	0.000	0.678	1.476
New Model	(Constant)	-137086417.229	19542748.049	—	-7.015	0.000	—	—
	Atmospheric station pressure	1264.005	192.177	0.107	6.577	0.000	0.623	1.604
	Horizontal infrared radiation	47,644.241	3375.835	1.556	14.113	0.000	0.014	73.809
	Direct normal radiation	16,462.981	1213.513	0.178	13.566	0.000	0.962	1.040
	Total sky cover	-2787320.969	196312.532	-1.437	-14.198	0.000	0.016	62.171
	Dry bulb temperatures	-581774.230	37,805.690	-0.586	-15.389	0.000	0.114	8.794
	Dew point temperatures	-112436.116	27,307.607	-0.136	-4.117	0.000	0.151	6.625
	AOD	-566944.126	182564.249	-0.042	-3.105	0.002	0.879	1.137
	Precipitable water	1076401.110	240985.736	0.080	4.467	0.000	0.512	1.955
Limited Model	(Constant)	-3092395121.054	279777336.777	—	-11.053	0.000	—	—
	Dry bulb temperatures	-16496447.837	1001219.983	-0.677	-16.476	0.000	0.023	43.618
	Dew point temperatures	7565123.922	1075798.887	0.347	7.032	0.000	0.016	63.097
	Relative humidity	-1931850.757	293215.611	-0.186	-6.588	0.000	0.049	20.572
	Atmospheric station pressure	33,401.664	2711.907	0.102	12.317	0.000	0.561	1.781
	Global horizontal radiation	172387.218	8026.661	0.167	21.477	0.000	0.642	1.557
	Wind speed	-4842190.418	958428.818	-0.036	-5.052	0.000	0.763	1.310
	Total sky cover	-1842155.224	366936.200	-0.036	-5.020	0.000	0.758	1.320

^aDependent variable: heating load of a DOE large office building.

TABLE 26: Multiple regression equations using new variables based on building size.

Building Types		Linear Regression Model
Original variables only	DOE small office (time-lag0)	$Y = -179888432.974 + (-685748.939)$ $X_1 + (86,402.645)X_2 + (-47343.593)X_3 + (1811.848)$ $X_4 + (24,711.629)X_5 + (-49815.842)X_6 + (34,307.818)$ $X_7 + (61,552.875)X_8 + (-371456.480)$ $X_9 + (-1367478.673)X_{10} + (-544550.045)X_{11}$
		$Y = -242862773.124 + (-57775.571)X_2 + (2360.095)$ $X_3 + (76,361.042)X_4 + (-82461.104)X_5 + (56,837.692)$ $X_6 + (101106.744)X_7 + (-4510822.624)$ $X_8 + (-1387805.281)X_9 + (-677958.805)$ $X_{10} + (-799488.056)X_{11}$
		$Y = -242862773.124 + (2360.095)X_2 + (76,361.042)$ $X_3 + (56,837.692)X_4 + (-4510822.624)$ $X_5 + (-1387805.281)X_6 + (-57775.571)$ $X_7 + (-82461.104)X_8 + (101106.744)$ $X_9 + (-677958.805)X_{10} + (-799488.056)X_{11}$
	DOE medium office (time-lag1)	$Y = -84468362.750 + (-153900.636)X_2 + (820.654)$ $X_4 + (7302.493)X_5 + (-28329.843)X_6 + (22,573.966)$ $X_7 + (31,803.161)X_8 + (-380750.599)$ $X_{10} + (-215562.688)X_{12}$
		$Y = -137086417.229 + (-581774.230)$ $X_1 + (-112436.116)X_2 + (1264.005)X_4 + (47,644.241)$ $X_5 + (16,462.981)X_7 + (-2787320.969)$ $X_{10} + (-566944.126)X_{12} + (1076401.110)X_{13}$
		$Y = -137086417.229 + (-581774.230)$ $X_1 + (-112436.116)X_2 + (1264.005)X_4 + (47,644.241)$ $X_5 + (16,462.981)X_7 + (-2787320.969)$ $X_{10} + (-566944.126)X_{12} + (1076401.110)X_{13}$
With new variables	DOE large office (time-lag2)	$Y = -84468362.750 + (-153900.636)X_2 + (820.654)$ $X_4 + (7302.493)X_5 + (-28329.843)X_6 + (22,573.966)$ $X_7 + (31,803.161)X_8 + (-380750.599)$ $X_{10} + (-215562.688)X_{12}$
		$Y = -137086417.229 + (-581774.230)$ $X_1 + (-112436.116)X_2 + (1264.005)X_4 + (47,644.241)$ $X_5 + (16,462.981)X_7 + (-2787320.969)$ $X_{10} + (-566944.126)X_{12} + (1076401.110)X_{13}$
		$Y = -137086417.229 + (-581774.230)$ $X_1 + (-112436.116)X_2 + (1264.005)X_4 + (47,644.241)$ $X_5 + (16,462.981)X_7 + (-2787320.969)$ $X_{10} + (-566944.126)X_{12} + (1076401.110)X_{13}$

X_1 = dry bulb temperatures, X_2 = dew point temperatures, X_3 = relative humidity, X_4 = atmospheric station pressure, X_5 = horizontal infrared radiation, X_6 = global horizontal radiation, X_7 = direct normal radiation, X_8 = diffuse radiation, X_9 = wind speed, X_{10} = total sky cover, X_{11} = visibility, X_{12} = AOD, X_{13} = precipitable water, and Y = heating load.

point temperature in the time-lag2 model did not meet the criteria and thus were omitted from the final model.

As indicated through the obtained results, the time-lag2 model was shown to be the most suitable for heating load prediction of the DOE reference for a large office building. After taking into consideration the criteria related to the t-value, significance probability, and multicollinearity, the dew point temperature was excluded from the analysis; this model is defined as the “Modified Model” and is presented in Table 22. Secondly, another model is developed based on all the variables considered in the time-lag2 model, AOD, and precipitable water values; this model is defined as the “New Model” and it is also presented in Table 22. Additionally, a third model based on only those variables received from the Korea Meteorological Administration (KMA) and which can be input by the user into EnergyPlus was developed considering a time-lag of 2 hours; this third model is defined as the “Limited Model” and is shown in Table 22. In other words, the third model considers all real data with the exception of the data obtained from equations. The data considered for each model are described as prediction a, prediction b, and prediction c and are shown in Table 22.

Tables 23–25 show the statistical results of the new regression model dealing with AOD. As shown in the table, the F-variation of the model was higher than 159 at $p < 0.001$. The variance and the residuals of the “New Model” are slightly lower than those of the “Modified Model.” In

addition, the F-values of the “Modified Model” were shown to be the highest of all three models. The “Limited Model” indicated the largest difference between the maximum and minimum predictions, as well as the largest mean.

5. Conclusion

The current study aimed to understand the effect of time-lag as well as Aerosol Optical Depth on the prediction of the heating load in buildings. In the process, this study also explains how heating load predictions can be improved through the consideration of time-lag. Various studies have already established that outdoor weather conditions play a significant role in determining the thermal behavior of a building; the term time-lag in this sense is used to define the phenomenon in which the time of the effect of the outdoor climatic conditions on the building heating load is misrepresented during prediction. The current study uses regression models to identify which outdoor climatic conditions, in terms of time, best explain the heating load being experienced.

Many building energy simulation programs assume that the time at which peak outdoor conditions occur is the same time the effect of these conditions is highest inside the building. However, the results obtained by the current study prove otherwise. For instance, for all the building sizes considered (DOE reference buildings for small, medium,

and large offices), it has been shown that the heating load at a given time is best described by the outdoor climatic conditions that occur two hours prior to that time. In this same manner, the developed regression models can be used to predict the heating load of a future time, for example, two hours in advance; a number of factors, such as the building materials used, can explain the reason for this. For example, the outer material of a building might absorb heat during the peak solar hours only to release it later into the building. In this case, a building simulation program would estimate the highest thermal load to be at the time of the peak solar hours, yet the effect is actually experienced later when the heat has been released from the building mass. The regression models developed here, however, can be used accurately to estimate the heating load since they consider the time-lag caused by factors such as the building mass or building size. Table 26 is a summary of the regression models that best estimate the heating load per building size. In addition, the table presents the developed models in two sets: the first set considering only the time-lag and the second set considering both time-lag and Aerosol Optical Depth data.

In conclusion, the key finding of this study is that both time-lag and Aerosol Optical Depth (AOD) data are critical factors in the estimation of building energy usage. As such, conventional building simulation programs must endeavor to include time-lag as an important factor in the prediction process. The impact of each input variable on the output can be assessed by the size of the coefficient values. For example, for the DOE small office linear regression model, total sky cover has the biggest impact on building heating loads.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] H.-x. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.
- [2] C. Buratti, P. Ricciardi, and M. Vergoni, "HVAC systems testing and check: a simplified model to predict thermal comfort conditions in moderate environments," *Applied Energy*, vol. 104, pp. 117–127, 2013.
- [3] G. S. Brager and R. de Dear, "A standard for natural ventilation," *ASHRAE Journal*, vol. 42, no. 10, pp. 21–29, 2000.
- [4] J. A. Clarke, *Energy Simulation in Building Design*, Routledge, Abingdon, UK, 2001.
- [5] P. Depecker, C. Menezes, J. Virgone, and S. Lepers, "Design of buildings shape and energetic consumption," *Building and Environment*, vol. 36, no. 5, pp. 627–635, 2001.
- [6] A. Fouquier, S. Robert, F. Suard, L. Stéphan, and A. Jay, "State of the art in building modelling and energy performances prediction: a review," *Renewable and Sustainable Energy Reviews*, vol. 23, pp. 272–288, 2013.
- [7] N. Fumo, "A review on the basics of building energy estimation," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 53–60, 2014.
- [8] V. L. Erickson, M. Á. Carreira-Perpiñán, and A. E. Cerpa, "Occupancy modeling and prediction for building energy management," *ACM Transactions on Sensor Networks (TOSN)*, vol. 10, no. 3, pp. 1–28, 2014.
- [9] K.-U. Ahn and C. S. Park, "Different occupant modeling approaches for building energy prediction," *Energy Procedia*, vol. 88, pp. 721–724, 2016.
- [10] J. Virote and R. Neves-Silva, "Stochastic models for building energy prediction based on occupant behavior assessment," *Energy and Buildings*, vol. 53, pp. 183–193, 2012.
- [11] A. S. Ahmad, M. Y. Hassan, M. P. Abdullah et al., "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 102–109, 2014.
- [12] J. Zhao, B. Lasternas, K. Poh Lam, R. Yun, and V. Loftness, "Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining," *Energy and Buildings*, vol. 82, pp. 341–355, 2014.
- [13] A. Fonti, G. Comodi, S. Pizzuti, A. Arteconi, and L. Helsén, "Low order grey-box models for short-term thermal behavior prediction in buildings," *Energy Procedia*, vol. 105, pp. 2107–2112, 2017.
- [14] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, "Building energy consumption prediction: an extreme deep learning approach," *Energies*, vol. 10, no. 10, p. 1525, 2017.
- [15] R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy," *Applied Energy*, vol. 123, pp. 168–178, 2014.
- [16] C. Fan, F. Xiao, and S. Wang, "Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques," *Applied Energy*, vol. 127, pp. 1–10, 2014.
- [17] M. R. Braun, H. Altan, and S. B. M. Beck, "Using regression analysis to predict the future energy consumption of a supermarket in the UK," *Applied Energy*, vol. 130, pp. 305–313, 2014.
- [18] G. K. F. Tso and K. K. W. Yau, "Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761–1768, 2007.
- [19] J. C. Lam, S. C. M. Hui, and A. L. S. Chan, "Regression analysis of high-rise fully air-conditioned office buildings," *Energy and Buildings*, vol. 26, no. 2, pp. 189–197, 1997.
- [20] T. Catalina, V. Iordache, and B. Caracaleanu, "Multiple regression model for fast prediction of the heating energy demand," *Energy and Buildings*, vol. 57, pp. 302–312, 2013.
- [21] N. Aghdaei, G. Kokogiannakis, D. Daly, and T. McCarthy, "Linear regression models for prediction of annual heating and cooling demand in representative Australian residential dwellings," *Energy Procedia*, vol. 121, pp. 79–86, 2017.
- [22] N. Fumo and M. A. Rafe Biswas, "Regression analysis for prediction of residential energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 332–343, 2015.
- [23] G. Y. Yun, J. H. Lee, and K. Steemers, "Extending the applicability of the adaptive comfort model to the control of air-conditioning systems," *Building and Environment*, vol. 105, pp. 13–23, 2016.
- [24] L. Wang, P. Mathew, and X. Pang, "Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building," *Energy and Buildings*, vol. 53, pp. 152–158, 2012.

- [25] S. Sholahudin and H. Han, "Simplified dynamic neural network model to predict heating load of a building using Taguchi method," *Energy*, vol. 115, pp. 1672–1678, 2016.
- [26] D.-S. Kapetanakis, E. Mangina, and D. P. Finn, "Input variable selection for thermal load predictive models of commercial buildings," *Energy and Buildings*, vol. 137, pp. 13–26, 2017.
- [27] D. B. Crawley, "Which weather data should you use for energy simulations of commercial buildings?/discussion," *Ashrae Transactions*, vol. 104, p. 498, 1998.
- [28] J. W. Moon and J.-J. Kim, "ANN-based thermal control models for residential buildings," *Building and Environment*, vol. 45, no. 7, pp. 1612–1625, 2010.
- [29] M. Deru, K. Field, D. Studer et al., "US Department of Energy commercial reference building models of the national building stock," Tech. Rep. NREL/TP-5500-46861, p. 1, NREL, Golden, CO, USA, 2011, 2011p. 1.
- [30] K. Field, M. Deru, and D. Studer, "Using DOE commercial reference buildings for simulation studies," *IBPSA-USA Journal*, vol. 4, no. 1, pp. 85–93, 2010.