

Article

Prediction Model Based on an Artificial Neural Network for User-Based Building Energy Consumption in South Korea

Seunghui Lee ¹, Sungwon Jung ^{1,*} and Jaewook Lee ²

¹ Department of Architecture, Sejong University, Seoul 05006, Korea; softbutter25@sju.ac.kr

² Department of Architectural Engineering, Sejong University, Seoul 05006, Korea; jaewook@sejong.ac.kr

* Correspondence: swjung@sejong.ac.kr; Tel.: +82-02-3408-3289

Received: 20 January 2019; Accepted: 13 February 2019; Published: 15 February 2019



Abstract: The evaluation of building energy consumption is heavily based on building characteristics and thus often deviates from the true consumption. Consequently, user-based estimation of building energy consumption is necessary because the actual consumption is greatly affected by user characteristics and activities. This work aims to examine the variation in energy consumption as a function of user activities within the same building, and to employ an artificial neural network (ANN) to predict user-based energy consumption. The study exploited the actual 24-h schedules of 5240 single-person households and computed the respective energy consumption using EnergyPlus V 8.8.0 software. The calculated values were clustered according to gender, age, occupation, income, educational level, and occupancy period and the difference among them was analyzed. The simulation results showed that for single-person households in Korea, females used more energy than males did, and the difference increased with age. Furthermore, unemployed and low-income individuals consumed more energy whereas consumption was inversely proportional to the educational level. Energy consumption increased with the occupancy period. Based on the simulation results and six user characteristics, the ANN model indicated a correlation between user characteristics and energy usage. This study analyzed the differences in energy usage depending on user activity and characteristics that affect building energy consumption.

Keywords: artificial neural network; big data; energy-performance gap; building energy prediction; building user activity; single-person household; Korean household energy consumption

1. Introduction

Buildings are responsible for a high percentage of CO₂ emissions in cities as well as 40% of the total energy consumption worldwide [1]. Many countries have passed environmental laws and policies to increase the energy efficiency of new buildings and to promote green remodeling of existing buildings for improved efficiency. Subsequently, building energy performance assessments are conducted according to national standards and performance ratings must be disclosed during real estate transactions. In Europe, for instance, all countries registered in Energy Performance of Building Directives (EPBD) since 2009 are required to disclose the building energy performance rating to the market. A similar policy has been adopted by the city of Seoul, Korea since 2013 and is gradually expanding to other cities to raise awareness and encourage the development of energy-efficient buildings, green remodeling revitalization.

The conventional approach to the evaluation of energy consumption often yields inaccurate estimates. Majcen et al. observed that the total amount of energy used by the occupants of buildings with a high energy performance rating was higher than the estimated value [2]. Researchers refer to

this discrepancy between actual and predicted energy usage as the “energy–performance gap” [3]. The fact that current building energy performance assessments are based heavily on the physical building characteristics rather than the actual characteristics and activities of building occupants is responsible for introducing the energy–performance gap. According to Al-Zubaidy and Kaddory, building energy consumption may vary by factors of 0.5–2.8, depending on the user [4].

The building energy performance rating is publicly disclosed in an effort to promote the voluntary revitalization of the energy performance of buildings. Nonetheless, the provision of accurate building information to users is challenging because of the “Energy-performance Gap.” Moreover, even though the building performance enhancement before and after green remodeling can be assessed objectively, the energy consumption variation between two or more occupants is difficult to evaluate. However, occupant characteristics and activities must be considered along with building characteristics when predicting the energy consumption to enhance the accuracy of energy information provided to buyers and residents.

This study analyzes the influence of users on the energy usage of buildings and proposes the possibility of predicting energy usage considering users' characteristics and activities. This study combines energy simulations and an artificial neural network (ANN) to examine the variation in building energy consumption with respect to user activities and characteristics within the same building. A model for energy consumption forecasting is then constructed based on the analysis.

2. Literature Review

2.1. Effects of Occupant Characteristics on Energy Consumption

Various researchers have demonstrated that occupants affect the building energy consumption the most, and they are also the primary cause of the “Energy-performance Gap” [5–8]. Several previous studies have investigated the variation in energy consumption as a function of the actual occupant characteristics and activities. For example, Van den Brom et al. analyzed information from 1.4 million households and observed a variation in energy consumption according to income and age [3], whereas Guerra-Santin showed energy usage to be directly proportional to education level [9]. Schipper et al. and Noh demonstrated variation in energy consumption as a function of the occupancy period because of different lifestyles [10,11], whereas Rätty observed that the energy consumption varied with respect to the gender of the occupant [12]. Jones et al. reported that population socioeconomic characteristics affect the use of indoor electrical energy [13].

The differences in the aforementioned demographic, social, and economic characteristics of the occupants give rise to distinct daily activities. These activities determine user-dependent indoor energy usage in terms of cooling, heating, lighting, the use of appliances which eventually results in various energy consumption profiles. In this work, the energy usage is computed as a function of the activities of the occupant, and its correlation with the demographic, social, and economic characteristics of the occupants is examined to study the variations in user-based energy consumption.

2.2. Building Energy Performance Assessment Methods

Two building energy performance assessment methods exist. The asset rating (AR) method computes energy consumption by taking into account the physical characteristics of buildings whereas the operational rating (OR) method is based on the actual energy usage of a building [14].

The AR method determines the energy consumption with simulation tools that consider the floor plan, installation, and construction materials of a building when evaluating the energy performance of a building. Although the AR method allows for objective performance assessment based on building characteristics, the variation in consumption as a function of occupant activities cannot be easily examined. A number of tools are available to evaluate the dynamic building load by taking into account the standard building schedule and the number of occupants, but they cannot consider the indoor energy usage and usage hours accurately. The ECO2 software, which is used widely in

Korea for building performance assessment, cannot take the operational schedule of the building into account [15].

The OR method evaluates energy consumption by importing the actual data pertaining to building energy usage over a time period from utility bills or companies. However, consistent building use over a significant time period is necessary because the evaluation is based on actual measurements. Patterns must be identified from large data sets or probabilistic forecasting is necessary to analyze the correlation between the energy consumption of a particular building and user activities and characteristics. Data collection, however, typically requires a prolonged measurement period and the prediction of random user activities poses clear limitations [14,15].

Both of the aforementioned methods are restrictive in terms of obtaining variations in user-based energy consumption. This work exploits the AR method, which takes the detailed building characteristics into account for simulations, to complement the drawbacks of the OR method, which takes the actual user activities into consideration. The simulations that were conducted as part of the present study used EnergyPlus V 8.8.0, which is a dynamic simulation tool capable of reflecting the indoor energy usage and occupant characteristics with sufficient accuracy. In addition, we controlled the physical variables in the same building and studied the changes in user-based energy consumption using actual activity data from 5240 residents. This work contributes to the ultimate reduction in building energy consumption by making it easier to identify and predict user-based building energy consumption for the purpose of real estate transactions or building remodeling.

3. Materials and Methods

3.1. Research Method and Procedure

The present study utilizes data from the “Korean Time Using Survey” provided by Statistics Korea. The data relate to the activities and characteristics of 5240 users and were used to investigate variations in user-based energy consumption [16]. In addition, EnergyPlus V 8.8.0 software and ANN, a multi-variate machine-learning method, were employed to construct a prediction model encompassing both user activities and characteristics. The study was conducted as illustrated in Figure 1.

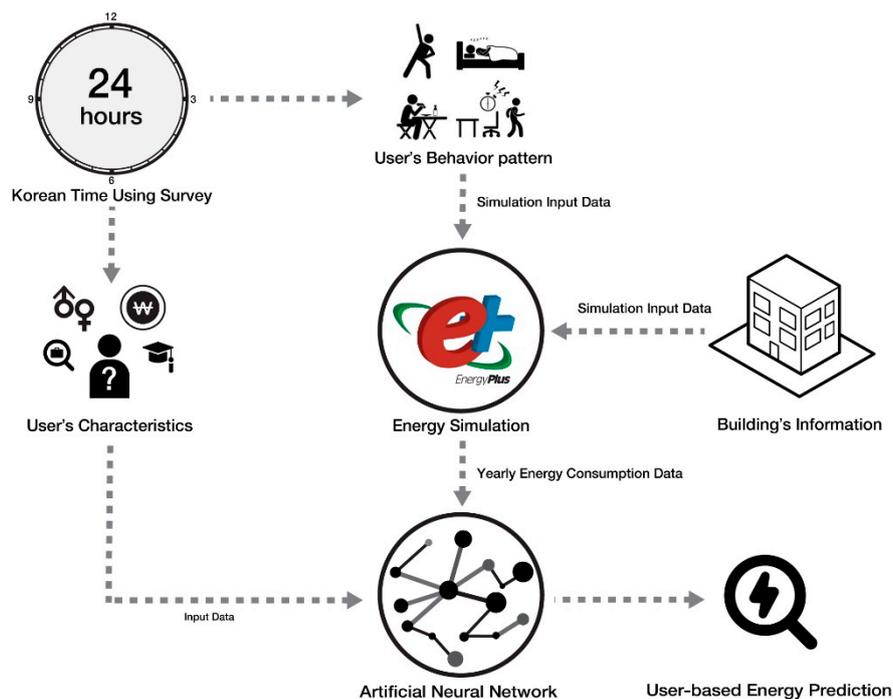


Figure 1. Flow chart depicting research process.

EnergyPlus is a dynamic simulation program that is used to compute energy consumption as a function of user activities in the same building. The data referenced in the simulation were the daily activity schedules of 5240 single-person households in Korea from the “2014 Korean Time Using Survey” compiled by Statistics Korea. The annual energy consumption was calculated based on user activity schedules and was then clustered according to the following user characteristics: income, age, level of education, occupancy time, gender, and occupation. This approach enabled us to examine variations in the energy consumption with respect to these six distinct characteristics. Subsequently, the validated user characteristics and simulated annual energy consumption were used to devise a model for energy consumption forecasting based on an ANN machine-learning approach as a function of the six user characteristics.

3.2. 2014 Korean Time Using Survey

This work utilized single-person household data from “2014 Korean Time Using Survey” conducted by Statistics Korea [16]. The number of single-person households is increasing rapidly worldwide due to aging, changes in perception. Therefore, governments must perceive single-person households as a major type of household and devise concrete policies and energy reduction plans accordingly. Hence, this study focused on single-person households and analyzed the energy consumption according to their characteristics and activities.

The survey data included approximately 12,000 Korean households with members of age 10 or older and can be divided largely into two sets. Table 1 lists the gender, age, occupation, income, educational level, as the characteristics of each household member and Table 2 presents the 24-h activity log of each household member recorded at 10-min intervals. The single-person household data comprised 5240 people across 800 cities in the Republic of Korea.

Table 1. Household information of the Korean Time-Using Survey [16].

Classification	Details	Literature Review
Number	Household ID	
Gender		Rätý (2010)
Age		Van den Brom et al. (2018)
Job	Job status	Noh (2013)
Educational Level		Guerra-Santin (2010)
Income Level		Paula et al. (2018)
Marital Status		
Care	Senior, Handicapped	
Resident	Kind, Area, Ownership	

The daily activity log was categorized according to the codes in Table 2. Nine activity categories were used and subcodes were assigned to identify particular activities. A total of 144 activity codes were recorded by dividing 1440 min in 10 min intervals. Activity codes can be referenced to determine the occupancy period, the use of appliances pertaining to particular activities.

Table 2. Behavior codes of the Korean Time-Using Survey [16].

	Activity	Details
A	Personal	Sleep, Eat, Individual Hygiene
B	Work	
C	Learning	School, Internet Lectures, Private Educational
D	Home Management	Cleaning, Washing, Shopping, Public Office
E	Family Care	
F	Meeting	Religion, Volunteer Work
G	Social & Leisure	Date, Media, Exercise, Rest
H	Move	Commute, Personal
Z	Other	Unclassified Act

4. Simulation Modeling of the Typical Living Environment in Korea

EnergyPlus is a building user-based simulation tool developed by the U.S. Department of Energy and was employed for the present study. EnergyPlus not only accounts for the physical characteristics of a building, but it also considers user schedules to estimate the energy consumption of respective buildings. Numerous previous studies employed EnergyPlus to perform building user-based energy simulations and the tool has been demonstrated to yield highly accurate simulation results based on user activities.

With the rapid industrialization and urbanization of Korea and the growing housing shortage in cities, construction of apartments has increased rapidly. As a result, typical apartments are currently of the residential type in Korea [17]. This work considered a representative Korean residential building type. According to the “2016 Korean Housing Survey,” the average gross floor area of a typical single-person household in Korea is 48.5 m²; hence, a household with an effective living area of 41.39 m² was considered in this study. Figure 2 depicts the considered building, Happy House, promoted by the Korea Land and Housing Corporation [18].



Figure 2. (a) Typical housing in the Republic of Korea [18]; (b) Typical housing layout for a one-person household [18].

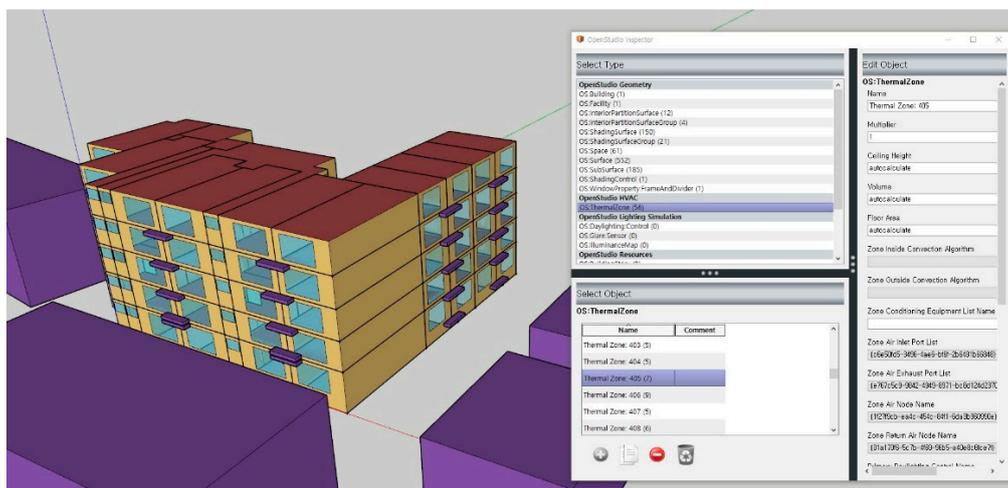


Figure 3. Building modeling for simulation.

Open Studio, which is provided as a Google Sketch-up plugin, was used to model the building [19]. The open studio version used in this study was 2.4.0, sketch-up version 2017. The building was modeled and entered as Figure 3. The entire building was modeled to account for thermal loads between households and a household located on a middle floor was analyzed herein. The input

parameters were the spatial properties and those of the building materials that were obtained from the floor plan and energy-savings plan in addition to the installation capacity.

Geographical conditions of the simulation were focused on those of Seoul, the capital of Korea. As shown in Figure 4, we used "Seoul's standard weather data" from the National Weather Service [20]. Seoul is geographically located in a mid-latitude temperate climate with four clear seasons. It is cold and dry in the winter while, in the summer, there is sultry weather. The average annual temperature in Seoul is 10–15 °C; it is above 30 °C in July–August, which is the warmest period, and temperatures below freezing occur in December–January, which is the coldest period. Approximately 61% of precipitation is distributed in the summer. The weather data for Seoul were converted into an EPW file and imported accordingly [20].

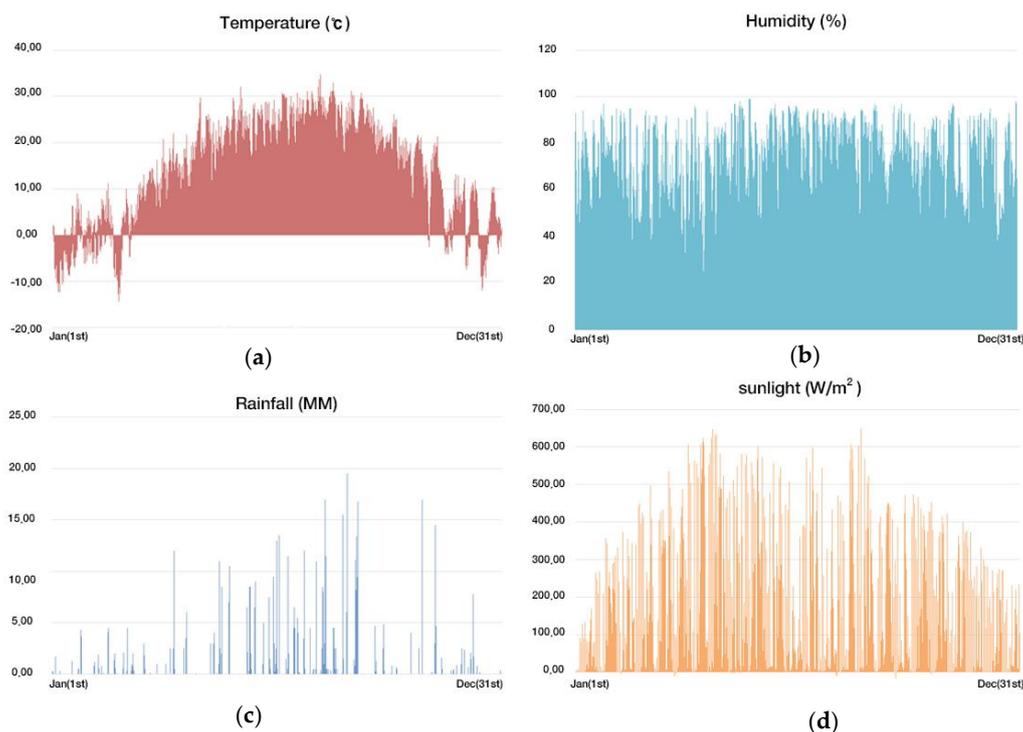


Figure 4. Standard Weather Data in Seoul, Korea [20]: annual values in Seoul for (a) temperature; (b) humidity; (c) rainfall; (d) sunlight.

The Korean population traditionally lives a comfortable life using floor heating systems called "Ondol." In modern Korean common housing, hot water piping is installed throughout the floor, and a heating system using a boiler is typical. Penetration of air-conditioning systems is substantial and in single households, wall-mounted air conditioners are often used [21]. Because the actual installation details, particularly those of the HVAC system, were not available, the Korean "Energy Saving Design Standards" were referenced to set December–March and June–September as the heating and cooling periods, respectively, and the indoor temperature was set between 20–26 °C [22]. Furthermore, the cooling system was modeled as a 2300 W wall-mounted "Ductless Air-conditioner" considering the typical HVAC requirements for a single-person household in Korea and the available installation area. On the other hand, the "Low Temperature Radiant Variable Flow" in EnergyPlus was used to model the heating system in the simulation because boilers are used for floor heating in Korea [23,24]. Information about the simulated building is summarized in Table 3.

Table 3. Simulation conditions.

Control Variables	Details
Total Area	2401.53 m ²
Simulation Area	41.39 m ²
Indoor Design Temperature	Set point: 20–26 °C
Heating System	Air Loop HVAC - Ductless Air-conditioner
Cooling System	Low-temperature Radiant Variable Flow
Weather Data	Seoul, Republic of Korea

The EnergyPlus software can estimate the energy consumption as a function of the activities of occupants by importing occupants' schedules as well as the indoor appliance capacity and usage profile. Occupant schedules were employed in the simulation to reflect the variation in energy consumption with respect to occupant activities by referencing the "2014 Korea Time Using Survey" activity codes. Indoor energy consumption was estimated based on the occupancy period and indoor activities, and the total energy used by home appliances and their operating periods were then obtained accordingly. The power consumption of home appliances was obtained from Statistics Korea's "2013 Home Appliance Power Consumption Survey" as listed in Table 4 [25]. Among the listed appliances, the refrigerator was considered to operate continuously for 365 days whereas 25 W light bulbs were considered to be installed in each room with a total power consumption of 100 W.

Table 4. Household appliances and their energy consumption [25].

Appliances	Power Consumption (W)
TV	130.6
Computer	255.9
Washer	242.8
Refrigerator	40.6
Lighting	100
Air conditioner	1200

The occupancy and operating period of each home appliance were interpreted using the activity codes and imported as in Table 5. The light was always on except when the occupant was asleep. The operating periods of other appliances were also estimated accordingly.

Table 5. Sample of daily energy schedule of household on 28th Aug 2014

Time	In & Out	TV	Computer	Lighting	Fridge	Hot Water	Cooking Stove
00:00	0	0	0	0	1	0	0
00:10	0	0	0	0	1	0	0
00:20	0	0	0	0	1	0	0
00:30	0	0	0	0	1	0	0
00:40	1	1	1	1	1	0	0
00:50	1	1	1	1	1	0	0
01:00	1	1	1	1	1	0	0
01:10	1	1	0	1	1	1	0
01:20	1	1	0	1	1	1	0
01:30	1	1	0	1	1	1	0
01:40	1	0	0	0	1	1	0
01:50	1	0	0	0	1	1	0
02:00	1	0	0	0	1	1	0
...
24:00	1	0	0	0	1	0	0

5. Modeling of an Artificial Neural Network (ANN) Based on User Information

The algorithm underlying the ANN is a machine-learning algorithm that mimics the human neural network and is used for prediction, clustering, and pattern recognition based on past and present training data. ANNs were first proposed by McCulloch and Pitts, but did not gain popularity in the early days because of the prohibitive amount of time required for training as the complexity of the model increased and as clear correlation between inputs and outputs was lacking [26]. However, Rumelhart et al. proposed the backpropagation algorithm and solved the optimal weight and bias for multi-layers with multiple nodes [27]. Methods for solving the phenomenon of information blurring, such as the ReLU function, have since been suggested, and ANN research has been revived due to improvements in computer performance. Subsequently, the problem associated with the ANN technique has been solved and improved results can be obtained by increasing the number of hidden layers existing between the input and output layers of ANN. In general, when two or more hidden layers are used, it is known as a Deep Neural Network (DNN) or Deep Learning. A DNN can classify high-level data by using iterative learning to process large amounts of data. ANNs allow the inclusion of a large number of variables, and different weights can be assigned to each variable to yield outputs that closely approximate measurements [28].

An ANN has the general structure shown in Figure 5 and is expressed as Equation (1).

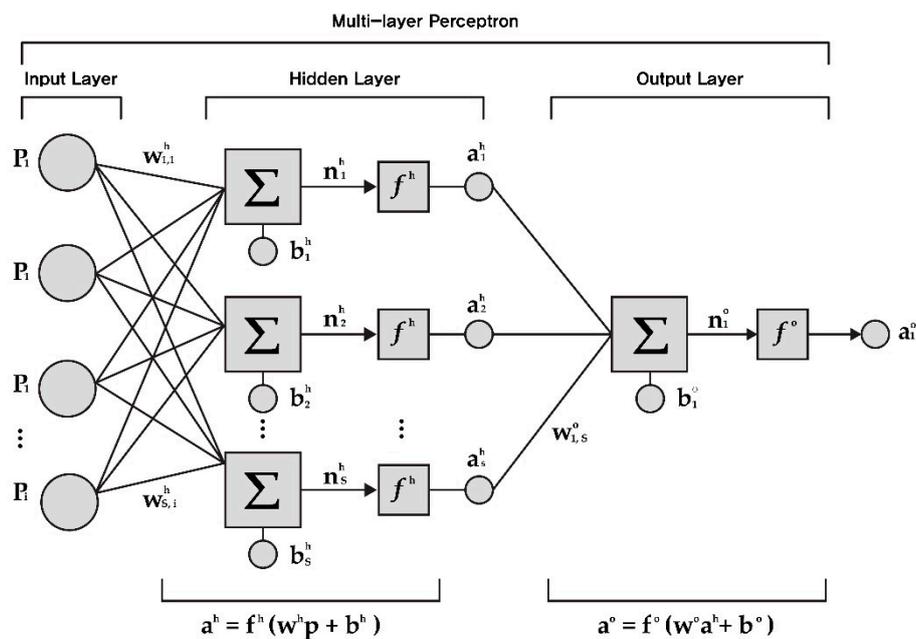


Figure 5. Structure of ANN model.

$$n_k^h = \sum_{j=1}^R w_{kj}^h p_j + b_k^h, \quad k = 1 \text{ to } S. \quad (1)$$

where R is the number of input variables and S is the number of hidden neurons. Further, p is the input variable, b^h is the bias of the hidden layer, and w^h is the weight. The calculated value is used as input for an activation function. The input to the ANN is processed to obtain the output by modifying the weight sum of the values from the previous layer by using the activation function. In general, previous studies used the sigmoid function as an activation function. However, the function exhibits the gradient vanishing phenomenon in which existing information converges to zero as the neural network expands. In addition, the sigmoid function requires additional computing time being an exponential function. In an effort to address this challenge, the nonlinear ReLU function was proposed. Figure 6 shows the sigmoid function and Relu function [29]. This function does not lose information

because it outputs the input value without any modification when it exceeds the threshold value, and the calculation speed is fast with a simple gradient value of 0 or 1. As a result, the performance of ANN increases remarkably with the ReLU function, and this approach was employed in numerous studies and is also used in the present work [30].

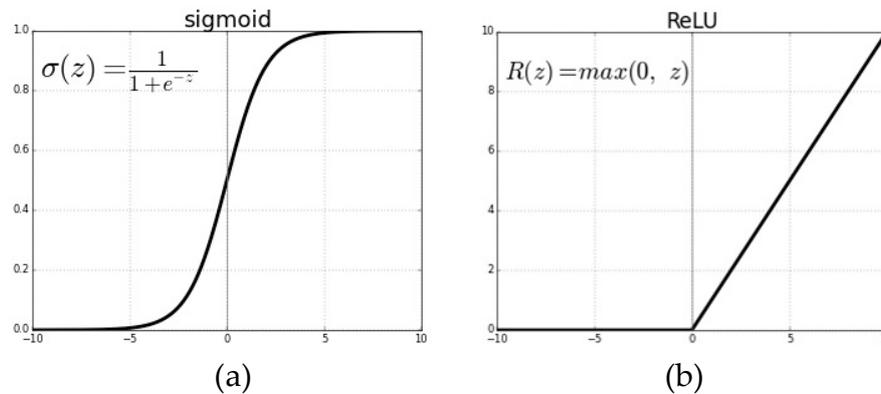


Figure 6. Comparison of two commonly used activation functions [29]: (a) Sigmoid; (b) ReLU.

In civil engineering and construction, ANN has been used consistently to predict the performance of structures. Yeh predicted the strength of concrete using the components of concrete [31]. Mata predicted the condition of a concrete dam by using environmental factors [32]. In addition, Cascardi predicted the strength of a concrete column, and proceeded to the wall shear strength [33,34]. Abambres et al. carried out the load prediction of an I-Section steel beam [35]. ANNs have been employed extensively in recent building load prediction studies owing to their advantages and improved algorithm. Factors related to building loads were typically given as inputs to forecast the actual load. Dong et al. performed energy prediction using the material properties of a wood office as data [36], whereas Kang forecasted the cooling load as a function of ambient and indoor temperatures [37]. Azadeh et al. considered environmental and economic factors to predict building energy consumption [38]. Martellotta et al. performed energy estimation with high accuracy with EnergyPlus and ANN using a user's representative schedule [39].

Previous studies employed ANNs to predict the building load based on the physical characteristics, environmental factors, or building use schedule. Sena et al. proposed that the behavior and characteristics of users and physical elements of buildings should be predicted through ANN but their approach could not be implemented [40]. The present study employed the demographic, social, and economic characteristics of building residents in an ANN to predict user-based energy consumption.

The user activity-based energy simulation results and user characteristics were used to construct a model to forecast user-based energy consumption. In this study, the ANN was implemented using NN Toolbox in MATLAB R2018b. Among the available user information, demographic, social, and economic characteristics such as age, income, gender, level of education, occupation, and occupancy period were provided as inputs and the energy consumption was obtained as the simulation output. The model was formulated as illustrated in Figure 7 and accepted a total of six inputs and yielded energy consumption as the sole output.

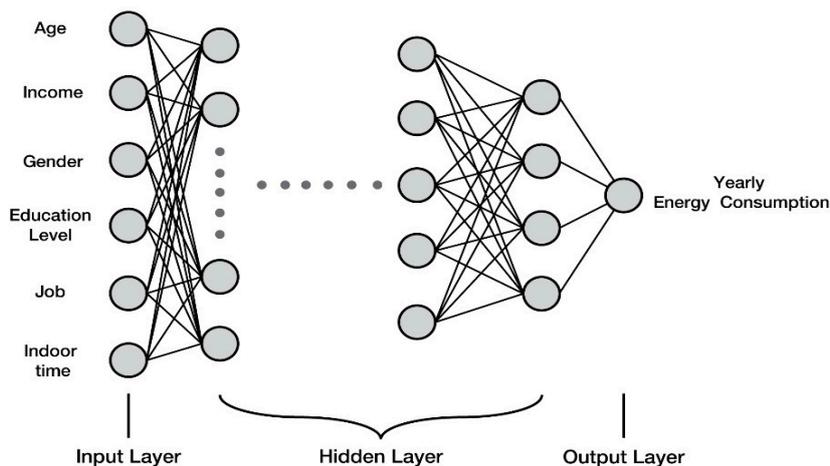


Figure 7. Model structure of the yearly energy prediction model

When using an ANN, there is no clear criterion for determining the number of hidden nodes and layers. The appropriate number of hidden nodes and layers is typically selected by varying the number of hidden nodes in the course of learning, and the optimal number of hidden nodes and layers is determined based on the best prediction performance. Huang and Foo found that it is desirable to determine the number of concealed neurons within $2n + 1$ when the number of input variables is n . In this study, six user characteristics are used as input variables. Among them, Job has 10 nominal variables and has 14 inputs in total. Therefore, in this study it is necessary to determine the number of nodes within a total of 29 [41]. In this study, 6 different conditions were set, as listed in Table 6, and were used to construct 6 models with distinct learning rates, number of hidden layers, and nodes. The Levenberg-Marquardt backpropagation algorithm was used as a network training function for learning ANN parameters [42].

Table 6. Running conditions for the ANN

No.	Activation Function	Learning Rate	Number of Layer	Number of Neuron in Layer
Network 1	ReLU	0.01	1	18
Network 2	ReLU	0.01	1	29
Network 3	ReLU	0.01	2	10–19
Network 4	ReLU	0.01	2	14–15
Network 5	ReLU	0.01	3	9–10–10
Network 6	ReLU	0.01	4	5–7–9–8

The ANN model uses given data for training and is then cross-validated with a new set of data. In this study, 70% of the energy simulation results were used for learning, 15% for model validation and 15% for model testing. The number of learning repetitions was set to 200, and training terminated when the number of epochs reached the maximum number or the mean squared error (MSE) continuously increased up to six times.

6. Results and Discussion

6.1. Simulation Result

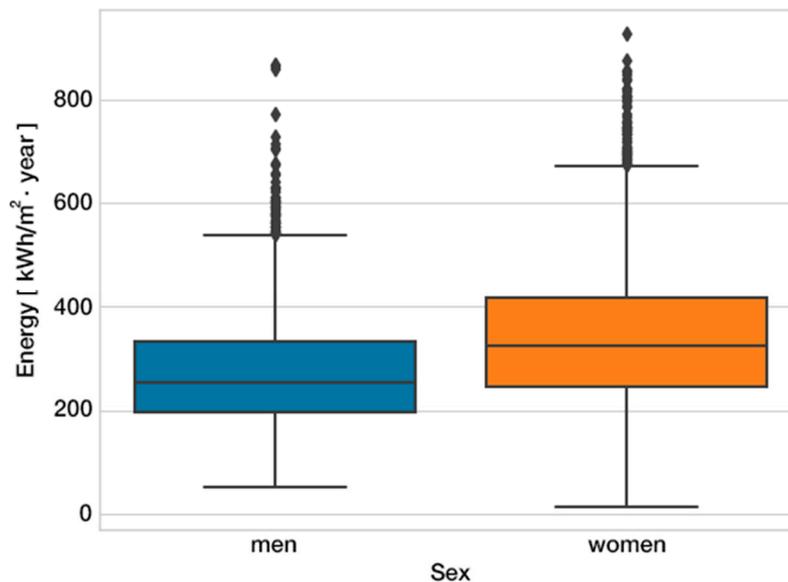
Simulation was performed based on information of 5240 single-person Korean households to estimate the annual building energy consumption. A total of 48 households with a zero occupancy period were excluded from the simulation and thus the energy consumption for a total of 5192 residents was obtained as summarized in Table 7. The results varied by approximately 20 times from 46.27 kWh/m²·year to 926.38 kWh/m²·year depending on the building users.

Table 7. Simulation results

Item	Value
Total	5240
Outlier	48
Extraction	5192
Min	46.27 (kWh/m ² ·year)
Max	926.38 (kWh/m ² ·year)
Mean	314.62 (kWh/m ² ·year)
Median	292.919 (kWh/m ² ·year)

6.2. Energy Consumption by Sex and Age

The building energy consumption varies depending on user gender and age. As shown in Figure 8, women feature a wider distribution than men and exhibit higher relative energy usage. The median, i.e., the usage at which 50% of the total population resides, is 274.66 kWh/m²·year for males and 341.01 kWh/m²·year for females. Figure 9 depicts the energy consumption by age and shows that consumption increases with age. The maximum value is found in the bin containing people in their 70s, and the lowest value was observed for those in their 30s. The average occupancy period for residents in their 30s is 13.33 h, which is the lowest of all age groups, whereas those in their 70s featured the longest average occupancy period of 18.34 h. Because the difference in occupancy period depends on the extent of economic activity, the younger the occupant, the less energy is used in the house. On the other hand, elderly individuals tend to spend more time at home because of retirement or aging and thus higher energy consumption is observed within this age group.

**Figure 8.** Comparison of energy consumption by sex.

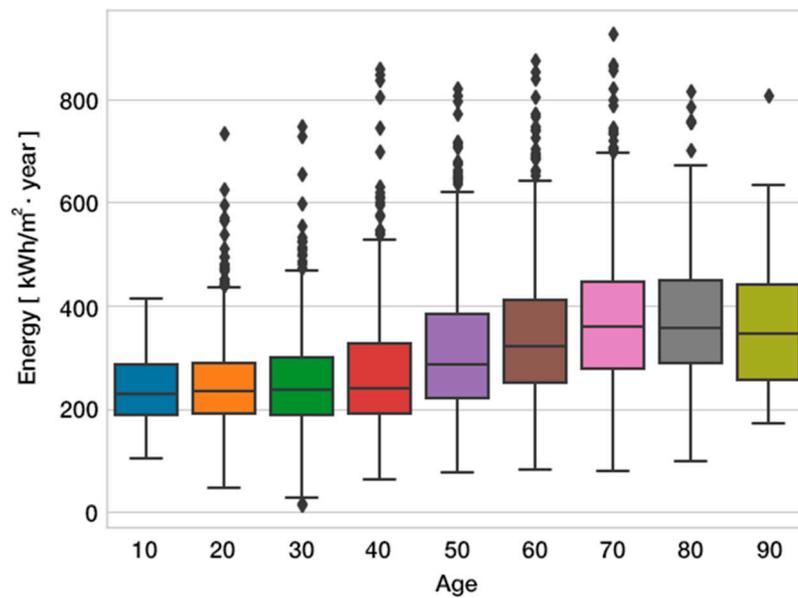


Figure 9. Comparison of energy consumption by age

6.3. Energy Consumption by Occupation and Income

Figure 10 shows That the group with analysis of the energy consumption by occupational status. The analysis of the energy consumption by occupation revealed that the unemployed group showed the highest energy usage because students were also included in this group and more frequently engage in indoor activities. The distribution of energy consumption was high for simple manual labor and service work, and the lowest distribution was determined for occupants working in management. Although the distribution in energy usage varies slightly depending on the occupation, all except for the unemployed group showed a distribution between 200 and 300 kWh/m²·year.

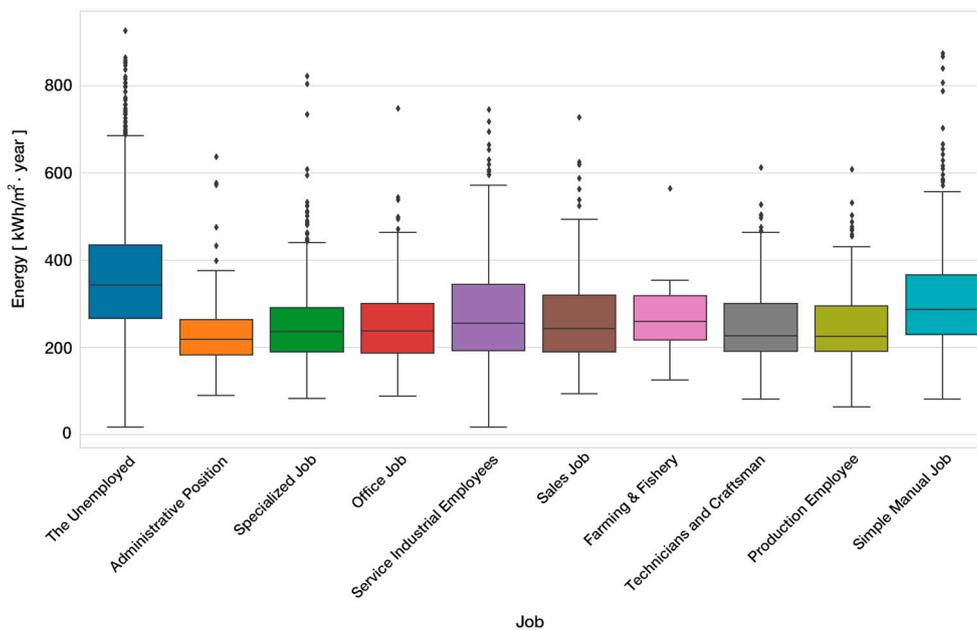


Figure 10. Comparison of energy consumption according to occupational status

Figure 11 shows that the group with a monthly average income of 0–1 million won consumes the most energy. This corresponds to the results in Figure 9. In the case of the unemployed, the number of elderly and retired people aged 60 or over accounted for 1605 out of 2135 or 75.17%. As a result,

elderly individuals without an occupation because of aging or retirement constitute a large portion of the low-income class. Energy consumption also varied as a function of the level of education.

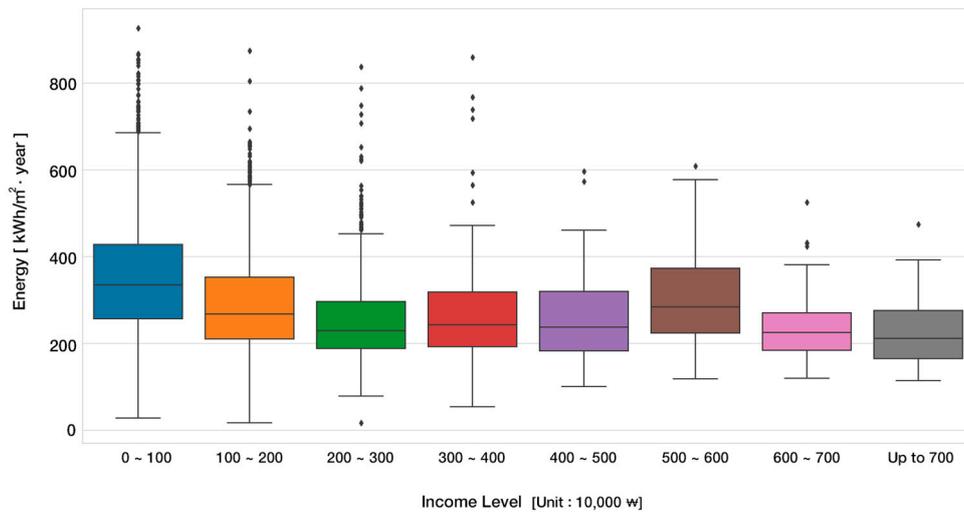


Figure 11. Comparison of energy consumption as a function of income

6.4. Energy Consumption According to the Educational Level

According to the results in Figure 12, the energy consumption decreases as the educational level increases. The group of elementary school graduates features the largest distribution in energy consumption with a median of 366.12 kWh/m²·year whereas individuals with a Master’s degree exhibit the smallest distribution with a median of 258.04 kWh/m²·year. The group of elementary school graduates achieved peak energy consumption whereas that of associate and four-year college graduates was the lowest. The differences in energy consumption among the considered groups became insignificant above college level.

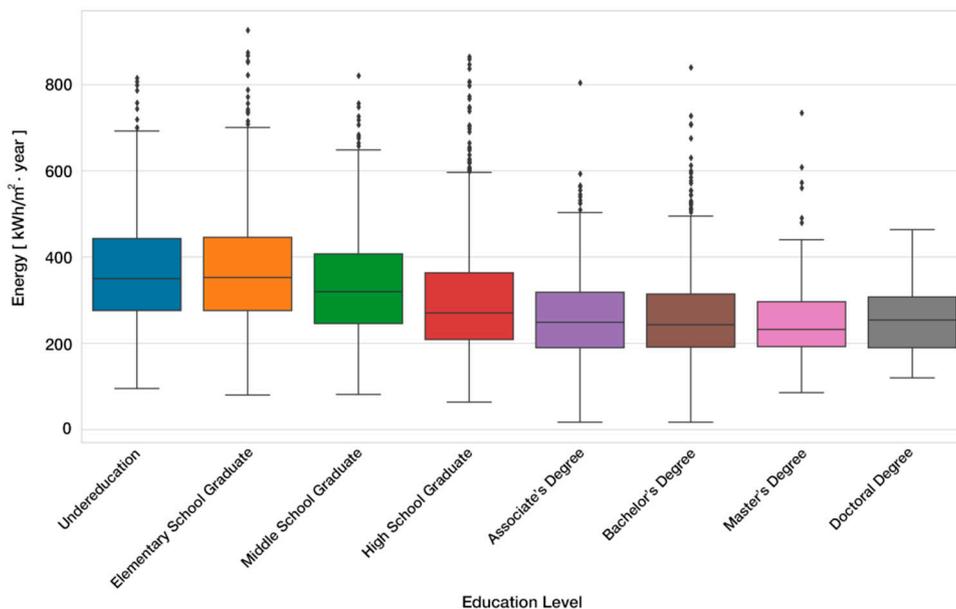


Figure 12. Comparison of energy consumption according to level of education

6.5. Energy Consumption According to the Occupancy Period

The energy consumption is directly proportional to the occupancy period. As shown in Figure 13, the longer the occupancy period, the higher the energy consumption is. This is attributed to the

high usage rates of indoor heating, cooling, lighting, and equipment as the occupancy time increases. The energy consumption is observed to be high within a period of 360 min (6 h) or less because many of the 52 users with an occupancy period of 6 h or less used home appliances such as computers and TV within that time frame. The indoor energy consumption increases in proportion to the occupancy period, but the indoor energy consumption may be higher per occupancy period depending on the user’s activity pattern.

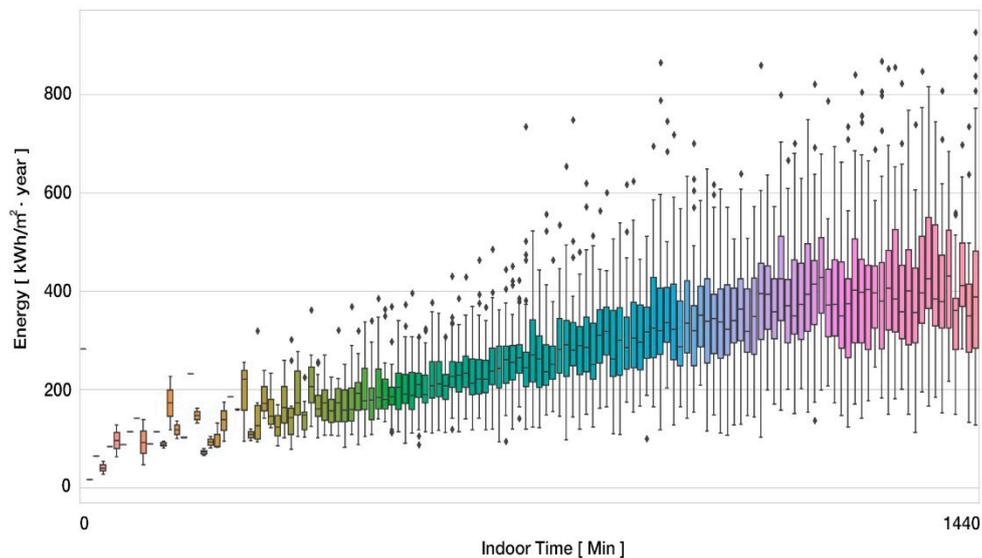


Figure 13. Comparison of energy consumption according to time spent indoors

6.6. Energy Consumption Prediction

Through simulation results, we analyzed the differences in energy usage according to six user characteristics. The ANN energy prediction model was implemented with simulation results and six user characteristics by the method presented in Section 5. Table 8 shows the datasets used in ANN, which should use continuous data as input data. Of the six user characteristics, Job has 10 nominal variables. Jobs are coded 1 and 0 for use as input data. Traditional metrics such as MSE and regression coefficient R were used to evaluate ANN performance. To show the evaluation result of ANN performance, 10 samples are extracted and the predicted value and the error rate are presented here.

Table 8. Sample of data set

Input Data											Output Data			
Age	Income	Gender	Education Level	Job							Indoor Time	Energy Consumption		
42	450	0	6	1	0	0	0	0	0	0	0	0	57	182.0869
83	50	0	1	0	0	0	0	0	0	0	0	0	123	395.9636
33	350	1	3	0	1	0	0	0	0	0	0	0	65	232.7648
29	250	1	5	0	1	0	0	0	0	0	0	0	51	149.0521
30	250	1	5	0	1	0	0	0	0	0	0	0	62	103.771
47	50	0	3	0	0	0	0	0	1	0	0	0	144	185.6166
56	150	0	2	0	0	0	0	0	0	0	0	1	63	334.8878
28	250	1	2	0	0	0	1	0	0	0	0	0	68	188.5984
70	150	0	1	0	0	0	0	0	0	0	0	1	82	239.1679

Table 9 shows the results of the ANN model executed under five conditions. Network 4 showed the best results, and the R-value of training was the highest for Network 1. However, the values of Validation and Test were low. The results of the five networks were the poorest for Network 6.

Table 9. The performances of the ANNs with different condition in hidden layer.

No.	R Value of Training	R Value of Validation	R Value of Test	MSE	Terminated Epoch
Network 1	0.64739	0.5805	0.6463	$10^4 \times 1.0513$	9th
Network 2	0.63229	0.59547	0.6153	$10^4 \times 1.0429$	5th
Network 3	0.63544	0.60723	0.60865	$10^4 \times 1.0637$	6th
Network 4	0.64312	0.62886	0.63089	$10^4 \times 1.0356$	13th
Network 5	0.63775	0.63885	0.59859	$10^4 \times 1.0855$	7th
Network 6	0.6155	0.62816	0.61952	$10^4 \times 1.0929$	10th

The performance of Network 4’s energy prediction model is shown in Figure 14. The regression R-values of the training, validation, and test data were 0.64312, 0.62886, and 0.63089, respectively. The correlation coefficient shows that the energy usage and the six user characteristics were strongly positively correlated. The MSE was 1.0356×10^4 , and the error distribution shows that more than 70% of the total data had a value between -100 and 100 . As a result, the conditions of Network 4 of two hidden layers, node configuration 14–15, and learning rate 0.01 were most appropriate among the networks with five different conditions.

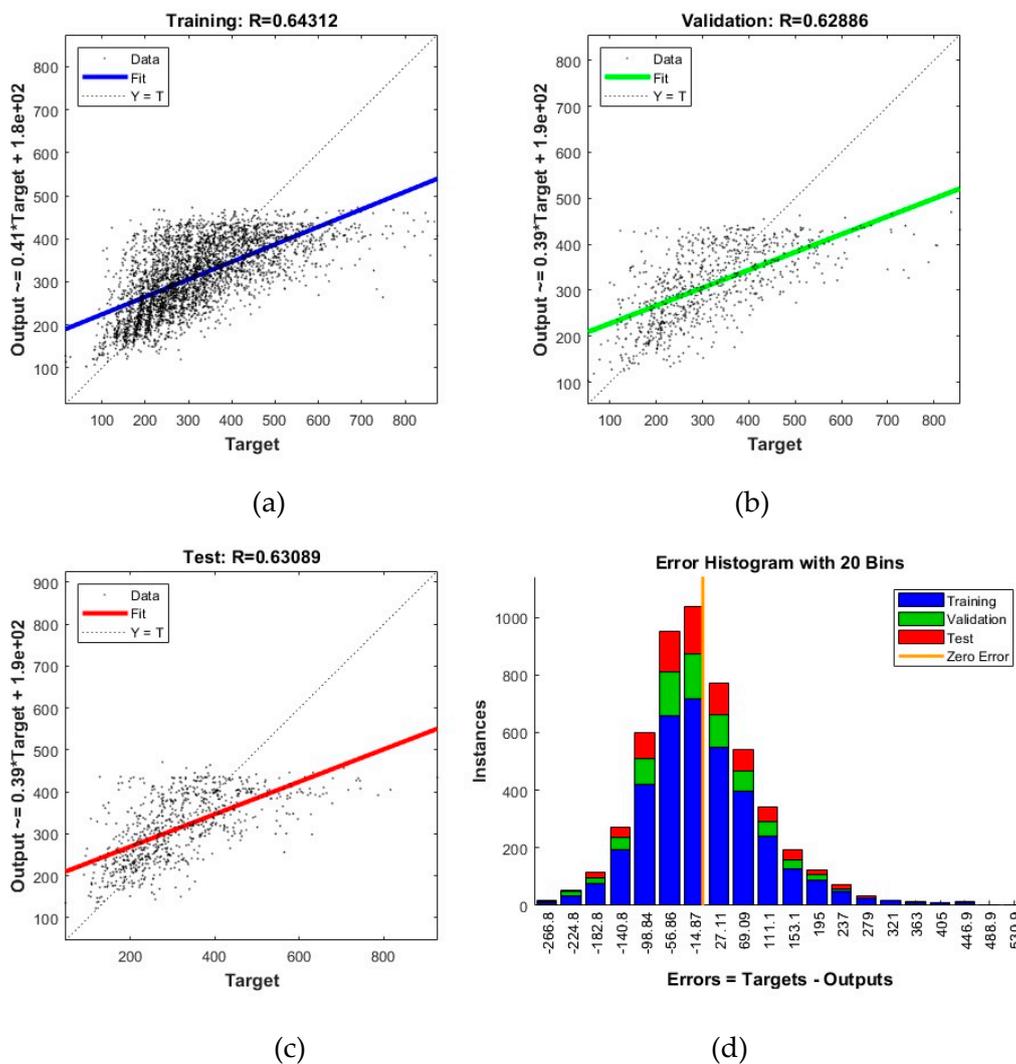


Figure 14. Training result of Network 4: (a) regression result from training data; (b) regression result from validation data; (c) regression result from test data; (d) histogram of the distribution of absolute errors.

7. Conclusions

In Korea, previous work on energy prediction using an ANN relied on building characteristics or environmental conditions whereas the present work was based on different user activities in a particular building as well as demographic, social, and economic characteristics. In this study, we tried to improve the accuracy by utilizing actual user activity and characteristic data and we drew six major user variables through simulation analysis.

This study involved the estimation and analysis of the user-based energy consumption in a building according to six user characteristics, namely gender, age, occupation, income, level of education, and occupancy period. The demographic, social, and economic characteristics and actual activity data of 5192 single-person households in Korea were used. In this study, the degree of differences in the influence of user characteristics could be more accurately compared because simulations were conducted by considering user activity within the same building, i.e., within a controlled physical environment. The six user characteristics were used as input to train the ANN model, which produced the simulated annual energy consumption of each individual as the output. As a result, the six user characteristics and energy usage were correlated with an R-value of 0.6 or more, and the model with an MSE of 1.0203×10^4 showed the best result.

The simulation results indicated that our approach enabled us to successfully analyze the difference in energy usage according to six user characteristics. However, the predictive performance of the Neural Network Predictive Model, which is based on six characteristics, is somewhat lower than that of the previous studies based on physical characteristics. The reason for the low prediction rate is that some parts of the model depend on the number of layers and neurons in the neural network, but the energy usage is considered to be a result of the physical and environmental factors of the building in addition to the user characteristics. The results of the study show that predicting the energy usage by only considering the user characteristics may have an influence on the energy consumption although the prediction rate is limited. However, the limitations are that the data used for ANN learning are simulation-based. If large amounts of energy information and user variable data were constructed for previous users in new or currently used buildings, the proposed study method using actual samples would allow users to compare energy use in a more sophisticated manner.

The process and results of this study can help to achieve a more accurate energy performance evaluation by considering all the users together when predicting the energy consumption of a building and it can serve as a basis to facilitate the transmission of energy information to users in energy management.

Author Contributions: S.L. proposed the main idea and performed the modeling, simulation, and data analysis and wrote the original draft; J.L. reviewed and edited the papers; S.J. reviewed the papers and oversaw the study. All authors contributed via discussions and participated in writing the manuscript.

Funding: This research was funded by the Ministry of Land, Infrastructure and Transport of the Korean government.

Acknowledgments: This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAITA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 18CTAP-C133446-02).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EPBD	Energy Performance of Building Directives
ReLU	Rectified Linear Unit
MSE	Mean Square Error

References

1. International Energy Agency. *Energy, Climate Change and Environment 2014*; IEA Publications: Paris, France, 2015.
2. Majcen, D.; Itard, L.; Visscher, H. Actual and theoretical gas consumption in Dutch dwellings: What causes the differences? *Energy Policy* **2013**, *61*, 460–471. [CrossRef]
3. Van den Brom, P.; Meijer, A.; Visscher, H. Performance gaps in energy consumption: Household groups and building characteristics. *Build. Res. Inf.* **2017**, *46*, 54–70. [CrossRef]
4. Kaddory Al-Zubaidy, M.S. A literature evaluation of the energy efficiency of leadership in energy and environmental design (LEED)-certified buildings. *Am. J. Civ. Eng. Archit.* **2015**, *3*, 1–7. [CrossRef]
5. Gram-Hanssen, K. Efficient technologies or user behaviour, which is the more important when reducing households' energy consumption? *Energy Effic.* **2012**, *6*, 447–457. [CrossRef]
6. Porcar, B.; Soutullo, S.; Enriquez, R.; Jiménez, M.J. Quantification of the uncertainties produced in the construction process of a building through simulation tools: A case study. *J. Build. Eng.* **2018**, *20*, 377–386. [CrossRef]
7. Tronchin, L.; Fabbri, K. Energy performance building evaluation in Mediterranean countries: Comparison between software simulations and operating rating simulation. *Energy Build.* **2008**, *40*, 1176–1187. [CrossRef]
8. Tronchin, L.; Manfren, M.; James, P.A.B. Linking design and operation performance analysis through model calibration: Parametric assessment on a Passive House building. *Energy* **2018**, *165*, 26–40. [CrossRef]
9. Guerra-Santin, O.; Itard, L. Occupants' behaviour: Determinants and effects on residential heating consumption. *Build. Res. Inf.* **2010**, *38*, 318–338. [CrossRef]
10. Jung, J.; Yi, C.; Lee, S. An integrative analysis of the factors affecting the household energy consumption in Seoul. *J. Korea Plann. Assoc.* **2015**, *50*, 75. [CrossRef]
11. Schipper, L.; Bartlett, S.; Hawk, D.; Vine, E. Linking life-styles and energy use: A matter of time? *Annu. Rev. Energy* **1989**, *14*, 273–320. [CrossRef]
12. Rätty, R.; Carlsson-Kanyama, A. Energy consumption by gender in some European countries. *Energy Policy* **2010**, *38*, 646–649. [CrossRef]
13. Jones, R.V.; Fuertes, A.; Lomas, K.J. The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renew. Sustain. Energy Rev.* **2015**, *43*, 901–917. [CrossRef]
14. Pérez-Lombard, L.; Ortiz, J.; González, R.; Maestre, I.R. A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes. *Energy Build.* **2009**, *41*, 272–278. [CrossRef]
15. Lee, M.-J.; Kim, J.-U. A study of domestic and foreign system and evaluate tools relative to building energy performance analysis and improvement plan for zero energy building design. *J. Archit. Inst. Korea Plan. Des.* **2015**, *31*, 45–52. [CrossRef]
16. Statistics Korea. *2014 Korea Time Use Survey*; Statistics Korea, Ed.; Micro Integrated Service: Seoul, Korea, 2015.
17. Kim, M.-K.; Yoon, C.-S. A study on statistical characteristics of space dimension of 60 m², 85 m² size apartment plan types in Seoul. *J. Korean Hous. Assoc.* **2010**, *21*, 53–65. [CrossRef]
18. Korea Land & Housing. Available online: <http://www.apply.lh.or.kr> (accessed on 10 October 2018).
19. Open Studio. Available online: <http://www.openstudio.net/> (accessed on 10 October 2018).
20. National Climate Data Service System. Available online: <http://sts.kma.go.kr> (accessed on 10 October 2018).
21. Ju, J.-H.; Park, J.-C.; Jeon, Y.-H.; Kim, D.-Y. A study on the usage status and satisfaction of the ventilation system installed in apartment houses. *J. Archit. Inst. Korea Plan. Des.* **2015**, *31*, 185–192. [CrossRef]
22. Go, G.-H.; Lee, S.-R.; Yoon, S.; Kim, M.-J. Optimum design of horizontal ground-coupled heat pump systems using spiral-coil-loop heat exchangers. *Appl. Energy* **2016**, *162*, 330–345. [CrossRef]
23. US DOE. *Engineering Reference with Energy Plus*; US Department of Energy: Washington, DC, USA, 2009.
24. US DOE. *Input Output Reference with Energy Plus*; US Department of Energy: Washington, DC, USA, 2009.
25. Lee, W.H.; Oh, H.S. Distribution characteristics of neophyten in Seongnagwon of scenic site no. 35. *J. Korea Soc. Plants People Environ.* **2013**, *16*, 181–185. [CrossRef]
26. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]

27. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
28. Salakhutdinov, R.; Mnih, A.; Hinton, G. Restricted Boltzmann machines for collaborative filtering. In Proceedings of the 24th International Conference on Machine Learning, Corvallis, OR, USA, 20–24 June 2007; pp. 791–798.
29. Towards Data Science. Available online: <https://towardsdatascience.com> (accessed on 10 October 2018).
30. Nair, V.; Hinton, G.E. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
31. Yeh, I.-C. Modeling of strength of high-performance concrete using artificial neural networks. *Cem. Concr. Res.* **1998**, *28*, 1797–1808. [[CrossRef](#)]
32. Mata, J. Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Eng. Struct.* **2011**, *33*, 903–910. [[CrossRef](#)]
33. Cascardi, A.; Micelli, F.; Aiello, M.A. Analytical model based on artificial neural network for masonry shear walls strengthened with FRM systems. *Compos. Part B Eng.* **2016**, *95*, 252–263. [[CrossRef](#)]
34. Cascardi, A.; Micelli, F.; Aiello, M.A. An Artificial Neural Networks model for the prediction of the compressive strength of FRP-confined concrete circular columns. *Eng. Struct.* **2017**, *140*, 199–208. [[CrossRef](#)]
35. Abambres, M.; Rajana, K.; Tsavdaridis, K.; Ribeiro, T. Neural Network-based formula for the buckling load prediction of I-section cellular steel beams. *Computers* **2019**, *8*, 2. [[CrossRef](#)]
36. Dong, Q.; Xing, K.; Zhang, H. Artificial neural network for assessment of energy consumption and cost for cross laminated timber office building in severe cold regions. *Sustainability* **2017**, *10*, 84. [[CrossRef](#)]
37. Kang, I.; Lee, K.; Lee, J.; Moon, J. Artificial neural network-based control of a variable refrigerant flow system in the cooling season. *Energies* **2018**, *11*, 1643. [[CrossRef](#)]
38. Azadeh, A.; Saberi, M.; Anvari, M.; Mohamadi, M. An integrated artificial neural network-genetic algorithm clustering ensemble for performance assessment of decision making units. *J. Intell. Manuf.* **2009**, *22*, 229–245. [[CrossRef](#)]
39. Martellotta, F.; Ayr, U.; Stefanizzi, P.; Sacchetti, A.; Riganti, G. On the use of artificial neural networks to model household energy consumptions. *Energy Procedia* **2017**, *126*, 250–257. [[CrossRef](#)]
40. Sena, B.; Zaki, S.A.; Yakub, F.; Yusoff, N.M.; Ridwan, M.K. Conceptual framework of modelling for Malaysian household electrical energy consumption using artificial neural network based on techno-socio economic approach. *Int. J. Electr. Comput. Eng.* **2018**, *8*, 1844–1853. [[CrossRef](#)]
41. Huang, W.; Foo, S. Neural network modeling of salinity variation in Apalachicola River. *Water Res.* **2002**, *36*, 356–362. [[CrossRef](#)]
42. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [[CrossRef](#)]

