Application of artificial neural network for natural ventilation schemes to control operable windows

Thanyalak Srisamranrungruang\textsuperscript{a,}\textsuperscript{*}, Kyosuke Hiyama\textsuperscript{b}

\textsuperscript{a} JSPS International Research Fellow, Meiji University, Kawasaki, Kanagawa, 214-8571, Japan
\textsuperscript{b} School of Science and Technology, Meiji University, Kawasaki, Kanagawa, 214-8571, Japan

ARTICLE INFO

Keywords:
Natural ventilation
Window control
Stack
Artificial neural network
Machine learning
Airflow

ABSTRACT

An artificial neural network (ANN) has been broadly developed as a design tool in various application scenarios in building sectors. One of the most important perspectives in building fields is human comfort. Various control strategies of natural ventilation schemes exist to maintain good air quality in buildings. Nevertheless, this study presented a novel strategy by applying a simple ANN to predict the trends of indoor air temperature and determine the operation status of operable windows. Building simulations had been conducted to train, test, and validate the ANN model. The ANN model has one hidden layer and performs training using the Levenberg-Marquardt algorithm. The nodes in the hidden layer were varied to configure the best-fitting model. The best structure of the ANN model in this study is the model with one hidden layer and 20 nodes. This study compares the performance of the ANN model when adopting the differential data set exhibiting better performance in predicting the indoor air temperature increase or decrease than that of the raw data. The prediction precision between the simulation and the ANN model when adopting the differential data is higher than that of raw data by 18%. This study discovered a new simple method and verified that a simple control strategy has been achieved by predicting the window operations using the increase or decrease in indoor temperatures via the ANN application.

1. Introduction

Artificial intelligence (AI) is becoming more prevalent in several applications that are adopted for various purposes. Its definitions are harnessed and referenced in several ways. In a simple definition by an online dictionary [1], AI is a branch of computer science that deals with the simulation of intelligent behavior in computers; in addition, it is the ability of a machine to imitate human intelligence. In 2004, John McCarthy [2] defined AI as “the science and engineering of creating intelligent machines, especially intelligent computer programs, which is similar to the task of using computers to understand human intelligence, but not confined to biologically observable methods.” AI is broadly categorized into two branches: expert systems and artificial neural networks (ANNs) [3]. An expert system is a logic program that employs computers to make decisions, while an ANN is implemented on computers that are not programmed to perform a specific task. In addition, it is a simulation of human biological brains comprising interconnections of collected data.

In the building sector, AI is adopted to enhance building performance, increase energy efficiency, reduce energy consumption, and improve the comfort and satisfaction of occupants. Kalogirou et al., (2001) [3], presented the various applications of ANNs in renewable energy problems. Another review work, an assertion paper [4], presented the potential of ANNs as design tools in many areas of building energy fields.

Kalogirou et al., (2000) [5], attempted to generate a training data set by first adopting a simulation program (ZID), and then developing ANNs to model the thermal behavior of a building and predict the energy consumption in a passive solar building. The heat load of buildings was estimated by adopting a backpropagation neural network developed on a MATLAB toolbox and then compared to the results of a conventional calculation. Accordingly, comparison results exhibited high accuracy [6].

In a recent review [7] related to the modeling and prediction of building energy consumption, the adopted methods include engineering, statistical, and artificial intelligence methods. However, the most widely used artificial intelligence methods are ANNs.

Kalogirou et al. (1999) [8], developed an ANN to estimate the energy extracted from solar domestic water heating systems (SDWHs) and the

\textsuperscript{*} Corresponding author.
E-mail address: thanyalak.s@meiji.ac.jp (T. Srisamranrungruang).

https://doi.org/10.1016/j.heliyon.2022.e11817
Received 26 April 2022; Received in revised form 13 June 2022; Accepted 15 November 2022
2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
temperature in stored water. Their objective was to adopt the developed ANN as a tool for estimating the performance of SDWHs.

The ANN presented in another study [9] was used to model the thermal behavior of a building’s space and predict its indoor temperature. For this application, four inputs were adopted: outdoor temperature, solar irradiance, indoor temperature, and heating system control signal. The proposed ANN was trained by the Levenberg-Marquardt algorithm. The activation function adopted with the input and output neurons was linear; however, in the hidden neurons, it took a sigmoid form. Natural ventilation is a crucial parameter in creating a conducive indoor air environment. However, it is difficult to operate the building with a naturally ventilated scheme and complex. To predict indoor air environments, ANNs have been developed by several researchers to predict and achieve conducive environments.

An application of machine learning (ML) with reinforcement learning to control indoor air temperature and humidity was achieved in Miami and Los Angeles throughout the year [10]. The predictive control model, which adopts ML on a natural ventilation application, significantly improved the building performance. The research developed ML to create a model predictive control and control ventilation systems with four different models, [11]. The baseline building was set under four different climate zones in the USA: Atlanta, Los Angeles, San Francisco, and Seattle. The neural network was the best method for controlling all phases under three different modes (natural ventilation (NV), air-conditioning (AC), and transition between NV and AC). An ANN model was applied to predict the natural ventilation rate in bedrooms at nighttime [12]. The 24 apartments in the cities of Chongqing and Chengdu, China, were selected and obtained data from them as a data set. The developed ANN model exhibited high accuracy in predicting the quality of natural ventilation in the buildings. This research suggested that engineers can evaluate indoor air quality by using this model as a guideline. Another research had validated the ANN model feasibility for predicting air velocity and distribution by comparing and training with computational fluid dynamics (CFD) [13]. This study exhibited the variations of pre-data processing that significantly impact the ANN model. Another work [14], adopted the CFD model to generate the database for training and validating an ANN model. The objective of this study was to create a meta-model and setup guidelines by optimizing window designs, to achieve thermal comfort in naturally ventilated buildings. In several buildings, the natural ventilation rate (NVR) was also designed as an index to calculate the amount of airflow and identify indoor air quality. Predictive models for estimating NVR were developed and investigated using eight ML models [15]. The data set of this study were collected from an existing office building in Daejeon, Republic of Korea. In a temperate climate such as the UK, heat loss occurred during the heating season in several buildings that employed operable windows for natural ventilation. An artificial intelligence-powered camera was used to develop ML with a deep learning network and control strategies for window operations, to reduce heat loss by identifying the windows operating in actual time, [16].

As stated by several literature reviews, the integrations of ML techniques such as ANN are beneficial in controlling engineered natural ventilation and in other fields such as chemical engineering, [17]. However, training an ANN usually requires big data, and data is limited in a practical situation. A crucial challenge in the application of ANNs in building fields is the amount of training data required to generate an optimal model. The data pre-processing based on fundamentals of physics for the ML model was developed under data sparsity in the building operation, [18].

The general operations of naturally ventilated buildings are required to be carefully designated and controlled. Nevertheless, the most prevalent problem among occupants is the noise generated during natural air ventilation. Another problem is the accuracy of control strategies defined in natural ventilation. A parameter commonly used in control strategies is the instant air temperature. If there is a big data set of an indoor thermal environment, the indoor air temperature has been possibly predicted by ML. According to the facts, a representative barrier is the difficulty in assembling a sufficient training data set to directly estimate the indoor thermal environment. An indoor thermal environment created by engineered natural ventilation principally depends on unique individual building features. Therefore, it is difficult to obtain a sufficient data set while the interval of operating the target building is underestimated. Therefore, an approach has to be analyzed to introduce significant information as an obtained data set for controlling engineered natural ventilation from a single target building with a limited amount of data. A building needs a few control actions to maintain a comfortable indoor temperature when the indoor temperature exceeds or drops below the target temperature. The opening or closing of natural ventilation windows can be defined as a simple control strategy. The window operation, whether opened or closed with a consequent increase or decrease in indoor temperature, can be exploited to determine the required action to take. However, under the condition that the amount of data is insufficient in predicting the indoor air temperature itself, the window operation remains beneficial. Nevertheless, control of the opening size has not always been linear to the indoor thermal environment because the indoor thermal environment with the engineered natural ventilation mode has mainly been influenced by weather conditions and the building’s use patterns, which are related to internal heat generations. Hence, in addition to weather data, the prediction requires data on the building’s use patterns. Based on the conditions of engineered natural ventilation, proportional relationships exist between the indoor and outdoor air temperatures as internal heat generators. Under these contexts, the information on the increase or drop in outdoor temperature, as the internal heat generation increases or decreases, must be sufficiently dominant to predict between two choices: increase or decrease in indoor temperatures.

Based on this hypothesis, this research investigates the significance of conducting the training data set with time series differentials by comparing them to the direct usage of the data set. Another objective is to exploit the increase or decrease of indoor temperatures to predict the status of window operations.

2. Methodology

2.1. Simulation setting for creating data set to train ANN model

In this study, a virtual building [19] is used for simulation, a general office building that has been reliably used to simulate the potential performances of Japanese buildings, as presented in Figure 1. However,
in this case study, the height of all windows has been modified from the reference model adapted to a present ordinary feature that typically utilizes enlarged openings. The model has seven floors, and the floor plan of each floor, except the 3rd floor designated for studying natural ventilation, has been illustrated (Figure 2). This study defines opening and vent to allow EnergyPlus is auto-generating natural flow. The 3rd floor is selected because to create stack effect, the selected floor should be the floor that has negative airflow pressure to draw airflow in. The floor plan of Office 1, which was used to assess natural ventilation flow is illustrated in Figure 3. The material and window properties used to construct the models are written in Tables 1 and 2. To calculate the thermal heat transfer of the analyzed floor (3rd floor), it was assumed that there is no heat exchange between adjacent floors (2nd and 4th floors). Void spaces in the building were merged into Office 1, to become a stack for ventilating

Figure 2. Floor plan of 1st–7th Fl (except 3rd Floor).

Figure 3. Floor plan of 3rd floor and flow path of natural ventilation.
air from the openings of the low level through the top floor (penthouse) where the installed vent faces four major directions. The natural ventilated pathways of the analyzed zone comprise seven windows facing the south and connected to the stack vent. The opening area of all windows in Office 1 is approximately 3.4 m². The vent area of the penthouse is approximately 3.64 m², while the discharge coefficient is 0.52. The wind pressure coefficients are defined in accordance with, [20]. In addition, the neutral pressure level around two-thirds of the building height is 22 m.

The indoor air temperatures under the natural ventilation condition are determined on the Design Builder software, [21]. The weather file adopted in this study contains data obtained from the Japan Meteorological Agency Weather Data of Tokyo, Japan, from the year 2016–2019. The natural ventilation is extensively operated during the months of April and May, every day between 7 am and 8 pm. Outside this specified time, the mechanical ventilation systems are turned on at 0.5 ac/h, to complete the 24-h ventilation of this building. According to virus pandemics, improving ventilation can help prevent virus particles from accumulating interior, [22]. The easiest method is open windows. If it is possible to open windows safely, intake of fresh air can help clear out the virus indoors. In Figure 3, this study defines a specific configuration example by setting two of seven windows always opened as the fixed-opening windows. The other windows are on a mixed-mode operation as operable windows that have been opened when the outdoor temperature is over 20 °C.

For the internal heat load in the Office 1 zone, the operation schedules are based on the standard Japanese building simulation. The occupancy density is 0.1 people/m², except on the third floor, which has an occupancy density of 0.08 people/m². The occupancy schedule is presented in Figure 4.

During custom schedules, the office equipment has a power density of 12 W/m², as shown in Figure 5, except on the third floor, which has a power density of 8 W/m². The lighting power density in the occupied zone has constantly been set at 7 W/m², and its schedule is presented in Table 3.

### 2.2. ML approaches with ANN

This study attempts to conduct ML applied with an ANN model to a naturally ventilated building in Japan. The platform that has been adopted in developing ANN is MATLAB with the deep learning toolbox, [23]. A feedforward neural network is the simplest ANN that transfers data from input layers, passes through hidden layers, then finally to output layers. A fitnet function fitting neural network with a hidden layer size is illustrated in Figure 6. Most of the neural network application has been trained by Levenberg-Marquardt because it performed the best among other algorithms, [24]. Accordingly, this study trained the neural network by using the Levenberg-Marquardt algorithm.

#### 2.2.1. Design of ANN and data set preparation

As aforementioned, this study adopts the Levenberg-Marquardt algorithm to train the ANN model. The hyperbolic tangent sigmoid transfer function is used as the activation function in the hidden layer, and the linear activation function is used in the output layer. The number of hidden layers and nodes was determined by trial and error. The learning rate was 0.01, and the number of epochs was 1000. The training was stopped when the validation loss stopped decreasing.

### Table 1. Building model information.

| Type           | Material          | Thickness (mm) |
|----------------|-------------------|----------------|
| Exterior Wall  | Tile              | 10             |
|                | Mortar            | 25             |
|                | Concrete          | 150            |
|                | Extruded polystyrene (XPS) | 50          |
|                | Air gap           | -              |
|                | Gypsum plasterboard | 8            |
| Interior Wall  | Gypsum plasterboard | 12            |
|                | Air gap           | -              |
|                | Gypsum plasterboard | 12            |
| Floor and Ceiling | Vinyl tile      | 3              |
|                | Air gap           | -              |
|                | Concrete          | 150            |
|                | Air gap           | -              |
|                | Gypsum plasterboard | 9            |
|                | MW board          | 15             |

### Table 2. Window properties.

| Type                      | Properties       | Value | Unit |
|---------------------------|------------------|-------|------|
| Low-E Glass               | Window to wall ratio | 40 | %    |
|                           | Heat transfer coefficient, [U] | 1.6 | W/m²·K |
|                           | Solar heat gain coefficient, [SHGC] | 0.4 |         |
| Window Frame (Aluminum)   | Heat transfer coefficient, [U] | 6.67 | W/m²·K |

Figure 4. Schedule and load percentage of occupancy in office zones.
function used in a hidden layer and the linear activation function used in the output layer is selected as the activation function because these two functions are mentioned in ref. [24], that it is most used and represented with the best performance. In general, there is no exact guideline on how to define the number of nodes (N); therefore, the number of nodes (N) in a hidden layer can be configured by varying sizes to achieve the best-fit model. In this study, the number of nodes (N) was preliminarily investigated as 15, 20, 25, and 30 nodes. The network performances of the trained model have been evaluated by adopting the mean square error (MSE) to determine the best model. Smaller MSE values indicate better performance. Another index that points to the network performance of the trained model is the correlation coefficient (R-value). The R-value is an indication of the relationship between the outputs and targets. If the R-value is equal to one, this indicates that a precisely linear relationship exists. However, if the R-value is close to zero, then there is no linear relationship between outputs and targets. In developing an ANN model, another test performance is required for its clarification by applying unknown data that have not been used in the training model to test its prediction. The test performances and number of node indexes in this study are presented in terms of standard deviation (S.D.) and percentage of positive direction (PP). In this study, the correlation of indoor air temperatures alternating between predicted and simulated values has been presented in terms of PP. This term implies that the percentages of the predicted and simulated values move together in the same direction when one variable increases as the other increases.

The neural network model has been tested in this paper; however, there is a universal variety of ML algorithms. Therefore, other ML learning algorithms that aim to solve regression problems will be tested and compared. A support vector machine (SVM) is selected to compare with the ANN model because it is a good ML tool for small data sets and is efficient. Furthermore, the SVM model with a linear kernel function is comprehensive and not expensive for computation like the ANN model, [25]. The SVM model has also been developed on the identical platform by using the function \texttt{fitrsvm} with a default setting, the linear kernel function, and solving the problem by Sequential Minimal Optimization (SMO).

### 2.2.2. Data description and feature selection

The weather conditions, internal heat generation conditions, and indoor air temperature history have been adopted as the feature variables in ML modeling. The indoor temperature change and the status of temperature increase or decrease have been set as the main objective functions to be predicted. The change in air temperature between the previous and current hours has been considered to help in deciding to open or close the windows. If the temperature increases, compared to the previous hour, the windows should be opened to ventilate the room and reduce excessive heat. When the temperature decreases, the windows should be shut. The indoor air temperatures of the previous 1 h and 2 h have been used as predictors in the data set. The weather data of outdoor

| Day        | Schedule                                                                 |
|------------|---------------------------------------------------------------------------|
| Every day  | 8 am–12 pm: 100% | 12 pm–1 pm: 50% | 1 pm–7 pm: 100% | 7 pm–8 pm: 80% | 8 pm–7 am: Closed |

**Table 3. Schedule and load percentage of lighting system.**

**Figure 5.** Schedule and load percentage of equipment used in office zones.

**Figure 6.** ANN structure.
parameters on time and 1 h before, which are outdoor dry-bulb temperatures, wind speed, and solar incident on windows are included in the data set. The internal loads on time and previous 1 h in the building, lighting load, equipment load, and occupants load are contained in the data set. In addition, the statuses of the windows on time and 1 h before, which are open or close, should be set at the minimum of the comfort range to cover the all comfortable range while still ventilating airflow. This reason leads to defining the comfortable indoor air temperature at 24 °C as one of the criteria for a simple control strategy for this study. So if the indoor air temperatures in this period are less than 24 °C, it has been demonstrated that there is straightforwardly no need to open any window within this period. In addition, the predicted indoor air temperature by ML modeling (both ANN and SVM) examines its accuracy by comparing it to the simulation results. For the data set preparation, training data obtain in three years, from 2016 to 2018, have been adopted, with a total of 1830 data. The data for the created ML model in the first two weeks of April 2019 have been used as a test data set with 140 data.

At the second step of this study, the prediction accuracy of the ML model that directly adopts the feature variables has been compared with the results of the improved ML model using the data differences in the feature variable histories. At this stage, a total of 1830 data in three years, from 2016 to 2018, are used for the training to create the ML model. The prediction results under the developed ML model in 2019, a total of 610 data per year, are available for ML modeling and checked by comparing the energy simulation outputs.

At the next step, the extent to which the usage of data differences contributes to the improvement has been investigated. In general, the increase in the quantity of the data set contributes to improving the prediction accuracy. The simulations by changing starting different days of the week have been made to create bigger training data because of data limitations. Then, the number of training data is increased by changing the day of the week for the first day in the energy simulation. Originally, the day of the week to start the simulation is set to Monday. By changing the setting to Tuesday, Wednesday, Thursday, and Friday, the training data increases by 5 times. In total, 3050 data per year are available, and 9150 data obtained in three years are available as the training data set. The case studies of selecting features for the ML model are presented in Table 5.

### Table 4. Descriptive statistics of the data used for the ML model.

| Predictor                     | Data type | S.D. | Mean | Min | Max | Variance |
|-------------------------------|-----------|------|------|-----|-----|----------|
| Lighting load (kW)            | Raw       | 0.44 | 2.78 | 1.47| 2.93| 0.19     |
|                               | Differential | 0.65 | 0.00 | -1.46| 1.46| 0.43     |
| Equipment load (kW)           | Raw       | 0.57 | 2.74 | 1.68| 3.35| 0.33     |
|                               | Differential | 0.33 | 0.00 | -1.34| 1.34| 0.11     |
| Occupancy load (kW)           | Raw       | 0.60 | 1.56 | 0.37| 3.36| 0.37     |
|                               | Differential | 0.52 | 0.01 | -1.64| 1.66| 0.27     |
| Outside Temperature (°C)      | Raw       | 4.30 | 20.25| 5.30| 30.13| 18.47    |
|                               | Differential | 0.99 | 0.17 | -5.35| 3.13| 0.98     |
| Indoor Temperature (°C)       | Raw       | 2.59 | 24.18| 13.61| 29.49| 6.71     |
|                               | Differential | 0.36 | 0.16 | -1.38| 1.33| 0.13     |
| Wind Speed (m/s)              | Raw       | 1.22 | 2.82 | 0.38| 7.93| 1.49     |
|                               | Differential | 0.63 | 0.12 | -2.25| 2.97| 0.40     |
| Solar on windows (kW)         | Raw       | 15.54| 24.75| 0.13| 53.54| 241.61   |
|                               | Differential | 6.48 | -1.53| -26.11| 27.19| 41.96    |
| Window status                 | Raw       | 0.00 | 1.00 | 0   | 1   |          |
|                               | Differential | 0.00 | 1.00 | 0   | 1   |          |

### Table 5. Case studies in ML model.

| Case | No. of Training Data | No. of Test Data | Data Type |
|------|----------------------|-----------------|-----------|
| 1    | 1830                 | 140             | Raw       |
| 2    | 1830                 | 140             | Differential |
| 3    | 1830                 | 610             | Raw       |
| 4    | 1830                 | 610             | Differential |
| 5    | 5490                 | 610             | Raw       |
| 6    | 9150                 | 610             | Raw       |

### Figure 7. Learning performance index of ANN for Cases 1 and 2.
The data in the first two weeks of April are used for both test data. The training and testing performance indexes are presented in Figures 7 and 8. To determine the best model for this research, all indexes have to be considered together. However, the most significant index for this research is the PP index for indicating the change in indoor air temperature to make a decision on the window status in the next step. From the results, the best test performance index is a model of Case 2 with 20 nodes at the highest PP of 100%, together with S.D., MSE, and R-value 0.1728, 0.0311, and 88.28%, respectively. The momentum update (MU) and gradient of the best ANN model in case 2 are 0.00001 and 0.1079 at epoch 13, respectively.

In the comparison between the performance of the ANN model and SVM model, as in Table 6. For the raw data set, the percentage change of case 1 in MSE of ANN is lower than SVM by approximately 1.3%. The percentage change of case 1 in the R-value of ANN is higher than SVM by approximately 0.1%. The percentage change of case 1 in the S.D. value of ANN is higher than SVM by approximately 14%. In addition, the percentage change of case 1 in PP of ANN is higher than SVM by approximately 67%. For the differential data set, the percentage change of case 2 in MSE of ANN is higher than SVM by approximately 1.2%. The percentage change of case 2 in the R-value of ANN is higher than SVM by approximately 0.4%. The percentage change of case 2 in the S.D. value of ANN is lower than SVM by approximately 9%. In addition, the percentage change of case 2 in PP of ANN is higher than SVM by approximately 1%.

Even the learning index performances of SVM when using differential data exhibit slightly better performance than the ANN model; however, the test indexes’ performances of ANN are better than those of SVM. Additionally, the most significant index for this study is the PP index, as mentioned above. Therefore, the performance of the ANN model outperformed the SVM model in this study which is similar to research about predicting water footprint that discovered that the ANN performance is better than the SVM model, [27].

The results above demonstrate and confirm that there is no need to open the windows under the condition of indoor temperatures lower than 24 °C because it will be too cold if it is opened. The harmonization between the ANN and simulation outputs of air temperature that change gradually are presented in Figure 9.

In Figure 9, the results of indoor air temperature trends can indicate how to plan the status of the operable window. Regardless of the temperature trends (increase (positive value) or decrease (negative value) of the indoor air), if the indoor air temperature is less than 24 °C, the...
Operable windows will be closed – assigned the binary number as 0. If the trends of the indoor air temperatures increase (positive value) or decrease (negative value), but the current air temperature is less than 24°C, then the operable windows will not change from their previous status – assigned in the graph as 0.5. If the trends of the indoor air temperatures increase (positive value) along with the air temperature greater than 24°C, then the operable windows will be opened – assigned a binary number as 1.

Consequently, the data set with differential predictors is sufficient in predicting the operating status of windows and indoor air temperatures for a larger data set in the future by adopting one hidden layer of the neural network model with 20 nodes. The learning performance and testing performance indexes of Cases 3 to 6 are presented in Figures 10 and 11, respectively.

The next step is the results with larger test data - 610 data of Cases 3 and 4. The learning performance indexes of Case 3 are MSE and R-value of 0.0330 and 99.83%, respectively. The test performance indexes are S.D. and PP values of 0.1786 and 78.33%, respectively. When using the differential test data of Case 4, the learning performance indexes are MSE of 0.0345 and the R-value of 88.41%. The test performance indexes are S.D. and PP values of 0.1666 and 99.18%.

The last analysis has evaluated the sizing of the data set for the ML model. In this section, the increasing amounts of raw data set in Cases 5 and 6 are 5490 and 9150, respectively, which have an identical size of test data at 610. The learning performance indexes of Case 5 are MSE and R-values of 0.0240 and 99.85%, respectively. The test performance indexes are S.D. and PP values of 0.1850 and 80.79%, respectively. When using the data of Case 6, the learning performance indexes are MSE and R-value of 0.0216 and 99.86%, respectively. The test performance indexes are S.D. and PP of 0.1772 and 78.98%, respectively. This implies that when more training raw data sets are employed, the ML model exhibits better performance.

In the comparison between the application of raw data (Case 3) and differential data (Case 4), as in Figure 10, the percentage change in MSE...
of Case 3 is lower than Case 4 by approximately 4%. The percentage change in the R-value of Case 3 is higher than Case 4 by approximately 12%. The percentage change in the S.D. value of Case 3 is higher than Case 4 by approximately 7%. In addition, the percentage change in PP of Case 3 is lower than Case 4 by approximately 20%. Even the learning index performance of Case 3 exhibits a better performance than Case 4; however, the test index performances of Case 4 are better than those of Case 4. Similar to Cases 5 and 6, the learning performance of both cases has been improved better than the case in which the differential data set (Case 4) was adopted; however, the test index performance of Case 4 has still better than other cases.

Based on the results above, the applications of the raw data exhibit a lower performance to predict the change in the indoor air temperature. This implies that the differential data set exhibits significantly higher performances in terms of the prediction operating status of windows and the change in air temperatures. The changes in the indoor air temperature from the ANN and simulations have been presented in Figure 12. This result is crucial in deciding the boarding of natural ventilation through windows in any building. The use of the differential data set exhibits high performance in predicting the change in the indoor air temperature and helps to predict the window status beforehand.

4. Conclusion

In this research, a feedforward ANN is developed to predict the indoor air temperature and define the status of operable windows for a naturally ventilated building in Japan. This study is preliminary research that boards of natural ventilation in the office building during the Spring season when the building is under moderate temperatures with no storms. This study presented a simple control strategy for natural ventilation, exploiting the increase or decrease in indoor air temperatures to predict the status of the operable windows. Most predictive models require big data; however, information from actual buildings such as newly operated or performed buildings has insufficient data. This study employs a relevant data set that can be obtained from most buildings and general weather information. The predictors in this study are indoor temperature, outdoor temperature, wind speed, solar incident on windows, window status, lighting load, occupancy load, and equipment load. This study determines the prediction performances of adopting the raw and differential data sets. The results obtained from the building simulation were used to train, test, and validate the ANN model.

In the preliminary study, the best ANN model structure for predicting the trends of indoor air temperature was the ANN model with one hidden layer and 20 nodes by adopting differential data and training model with the Levenberg-Marquardt algorithm. The learning indexes (MSE and R-value) obtained using raw data exhibited higher performances than the case with differential data by approximately 13%. However, the most important objective of this predictive research is to adopt the increase or decrease in indoor air temperatures to predict the window status in advance. From this point, the PP testing index has the most priority for analysis. The PP and S.D. of using the differential data are 100% and 0.1728 in the preliminary study, respectively.

The predictive preliminary result from the ANN model was beneficial in predicting how to set the operable windows by including the current indoor air temperature for configuring the window status. In the case of enlarging the test data set, the predictive model for configuring the window exhibits a negligible of approximately 1%. The use of raw data in the ANN model exhibited an optimal performance in the case with a big data set. The research demonstrated that the adoption of the differential data set with the time series exhibited significantly higher performances than the case with the raw data, which includes a small amount of data. This study presented the advantages of using differential data for prediction and solving the problem of insufficient information from an existing building. In addition to using the differential data set, the outcome of predicting the increase and decrease in indoor temperatures used in the control of window status beforehand can be adopted as guidelines.

This study was conducted for a specific building under specific parameters. This study verified that it is an optimal start to conduct a novel approach to applying the simple ANN for natural ventilation control. However, this method will be applicable to modify another building under different seasons or different climate zones in future works.

Disclosure

The authors declare that we have no conflict of interest, competing financial interests, or personal relationships that could have appeared to influence the work reported in this paper.

Declarations

Author contribution statement

Thanyalak Srisamranrungruang, Kyosuke Hiyama: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This work was supported by the International Research Fellow of Japan Society for the Promotion of Science (Postdoctoral Fellowships for Research in Japan (Standard)), Grants-in-aid for Scientific Research JSPS KAKENHI (JP20F20062), the School of Science and Technology, and the Organization for the Strategic Coordination of Research and Intellectual Properties of Meiji University.

Data availability statement

No data was used for the research described in the article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] https://www.merriam-webster.com/dictionary/artificial%20intelligence. (Accessed 14 October 2021).
[2] John. McCarthy, What is artificial intelligence?, 1998.
[3] S.A. Kalogirou, Artificial neural networks in renewable energy systems applications: a review, Renew. Sustain. Energy Rev. 5.4 (2001) 373–401.
[4] S.A. Kalogirou, Artificial neural networks in energy applications in buildings, Int. J. Low Carbon Technol. 1.3 (2006) 201–216.
[5] S.A. Kalogirou, M. Bojic, Artificial neural networks for the prediction of the energy consumption of a passive solar building, Energy 25.5 (2000) 479–491.
[6] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Software 40.5 (2009) 356–362.
[7] H.X. Zhao, F. Magoole, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16.6 (2012) 3586–3592.
[8] S.A. Kalogirou, S. Pantelis, A. Dentsoras, Modeling of solar domestic water heating systems using artificial neural networks, Sol. Energy 65.6 (1999) 335–342.
[9] M.M. Gouda, Sean Danaher, C.P. Underwood, Application of an artificial neural network for modelling the thermal dynamics of a building’s space and its heating system, Math. Comput. Model. Dyn. Syst. 8.3 (2002) 333–344.
[10] Y. Chen, L.K. Norford, H.W. Samuelson, A. Maikewi, Optimal control of HVAC and window systems for natural ventilation through reinforcement learning, Energy Build. 169 (2018) 195–205.
[11] J. Chen, G. Auguste, X. Song, Lighted-weighted model predictive control for hybrid ventilation operation based on clusters of neural network models, Autom. Constr. 89 (2018) 250–265.
[12] X. Dai, J. Liu, X. Zhang, W. Chen, An artificial neural network model using outdoor environmental parameters and residential building characteristics for predicting the nighttime natural ventilation effect, Build. Environ. 159 (2019) 106139.
Q. Zhou, R. Ooka, Influence of data preprocessing on neural network performance for reproducing CFD simulations of non-isothermal indoor airflow distribution, Energy Build. 230 (2021) 110525.

G.M. Stavrakakis, P.I. Zervas, H. Sarimveis, N.C. Markatos, Optimization of window-openings design for thermal comfort in naturally ventilated buildings, Appl. Math. Model. 36.1 (2012) 193–211.

H. Park, D.Y. Park, Comparative analysis on predictability of natural ventilation rate based on machine learning algorithms, Build. Environ. 195 (2021) 107744.

P.W. Tien, S. Wei, T. Liu, J. Calautit, J. Darkwa, C. Wood, A deep learning approach towards the detection and recognition of opening of windows for effective management of building ventilation heat losses and reducing space heating demand, Renew. Energy 177 (2021) 603–625.

H. Li, Z. Zhang, Z.Z. Zhao, Data-mining for processes in chemistry, materials, and engineering, Processes 7 (3) (2019) 151.

K. Hiyama, Y. Omokaka, Operation of climate-adaptive building shells utilizing machine learning under sparse data conditions, J. Build. Eng. 43 (2021), 103027.

SHASE, Guideline of Test Procedure for the Evaluation of Building Energy Simulation Tool, The Society of Heating,Air-Condit. Sanit. Engineers of Japan, 2016 (in Japanese).

T. Srisamranrungruang, K. Hiyama, Heliyon 8 (2022) e11817.

Design Builder Software Ltd, Design builder, Retrieved, http://www.designbuilder.co.uk/. (Accessed 10 June 2021).

Retrieved, https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/improving-ventilation-home.html. (Accessed 23 April 2022).

MATLAB®, Retrieved, https://www.mathworks.com/products/deep-learning.html. (Accessed 2 February 2022).

Roman, N.D., Bre, F., Fachinotti, V.D. and Lamberts, R., Application and characterization of metamodels based on artificial neural networks for building performance simulation: a systematic review, Energy Build. 217 (2020): 109972.

T. Østergård, R.L. Jensen, S.E. Maagaard, A comparison of six metamodeling techniques applied to building performance simulations, Appl. Energy 211 (2018) 89–103.

M. Indraganti, R. Ooka, H.B. Rijal, Thermal comfort and acceptability in offices in Japan and India: a comparative analysis, 2014 AKITA technical papers of annual meeting the society of heating, Air-Condit. Sanit. Eng. Japan 6 (2014) (2014) 17–20.

S. Meni, E. Aziz, A.P.A. Majeed, M. Malek, The prediction of blue water footprint at Semambu water treatment plant by means of Artificial Neural Networks (ANN) and Support Vector Machine (SVM) models, Phys. Chem. Earth, Parts A/B/C.123 (2021) 103052.