Intraseasonal and interseasonal applicability of a neural network model for real-time estimation of the number of air exchanges per hour of a naturally ventilated greenhouse

Ryo MATSUDA\(^\d\), Kota HAYANO, Takashi KAWASHIMA and Kazuhiro FUJWARA

( Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1–1–1 Yayoi, Bunkyo, Tokyo 113–8657, Japan )

Abstract

Neural network (NN) models with environmental data and the extent of ventilator openings as inputs have the potential to estimate the number of air exchanges per hour \(N\) in real time of a naturally ventilated greenhouse. In this study, the intraseasonal and interseasonal applicability of an NN model was verified: whether the model trained in a specific period can be applied to different periods of the same and other seasons. First, the effect of data collection periods for model training and test within the same season on the estimation accuracy of \(N\) was examined. The estimation accuracy was lowered even though the model was applied to a period immediately following that used for model training. Adjusting the training dataset so that the relative distribution of the temperature difference inside and outside the greenhouse \(\Delta T\) approaches the relative distribution of the test dataset improves the estimation accuracy slightly. However, when the model was applied to interseasonal data, such training data adjustments did not improve the estimation accuracy. This indicates that the NN model needs to be further improved for practical use to estimate \(N\) of naturally ventilated greenhouses.

Key words: \(\text{CO}_2\) concentration, Covariate shift, Environment control, Natural ventilation

1. Introduction

The number of air exchanges per hour \(N\) is used to quantify the extent of ventilation of a greenhouse. It is known that the \(N\) of a naturally ventilated greenhouse is dynamically affected by the current internal and external environment and the extent of ventilator openings. If \(N\) is estimated in real time, we can calculate the exchange rate of \(\text{CO}_2\), water vapor, and energy/enthalpy between the greenhouse and the outside in real time via ventilation. This can contribute to the cost-effective prompt environmental control, such as \(\text{CO}_2\) enrichment, even when the ventilators are open. However, unlike a greenhouse with forced ventilation, there is no practical method for estimating \(N\) of a naturally ventilated greenhouse in real time with practically enough accuracy.

There are several methods to calculate \(N\) of a naturally ventilated greenhouse (Boulard, 2006). The methods include the tracer gas method with \(\text{CO}_2\) (Okada and Takakura, 1973; Nederhoff et al., 1985; Boulard and Draoui, 1995), the energy balance method (Kozai et al., 1980; Fernandez and Bailey, 1993), the water balance method (Takakura et al., 2009, 2017), and the methods based on aerodynamic (Kozai et al., 1980; Sase et al., 1980; Boulard and Baille, 1995; Boulard et al., 1996; Kittas et al., 1997) or neural network (Seginer et al., 1994; Boulard et al., 1997) models. Among them, we have recently shown the potential availability of the NN models with internal and external environmental data and the extent of ventilator openings as the model inputs for real-time \(N\) estimation for a naturally ventilated uncropped greenhouse (Matsuda et al., 2019). Using data collected in an uncropped experimental greenhouse in the winter, NN models were trained and validated according to 10-fold cross-validation. The NN models tended to exhibit a higher accuracy in estimating \(N\) of the greenhouse than the aerodynamic and empirical models. On the other hand, several unignorable points remained uninvestigated in that study. One of them was the applicability of the model to a period different from that used for model training and validation. If real-time \(N\) estimation based on an NN model is intended to be used for a target greenhouse throughout the year, the model must first be trained with data collected over a short period in an uncropped season. Then, the model can be applied to the same greenhouse with plants to estimate \(N\) in real time.

In this study, we tested the intraseasonal and interseasonal applicability of an NN model: whether the model trained in a specific period can be applied to different periods of the same and other seasons. First, the intraseasonal applicability was examined. The NN model was trained with and tested against different periods within the same season (winter). Second, the training data for the NN model were selected and the entire training dataset was adjusted to improve the estimation accuracy. The training data were sampled so that the relative distribution of the dataset “resembles” that of a target test dataset, and its effect on improving the accuracy of estimation was verified. Third, we investigated the interseasonal applicability of the NN model trained with a dataset with or without relative distribution adjustment. The results suggested that the NN model can be...
applied to different periods in the same season after adjusting the training dataset. However, there are still issues to be addressed to apply the model to the other seasons.

2. Materials and Methods

2.1 Greenhouse specifications

The experimental greenhouse used in this study was in the head office of Seiwa Co., Ltd. in Shimotsuke, Tochigi, Japan (36°22’ N). A multi-span Venlo-type greenhouse was divided into eight single-span compartments and one compartment was used for the measurements. The roof covering material was the ethylene-tetrafluoroethylene film (F-CLEANTM Clear, AGC Green-Tech Co., Ltd., Tokyo, Japan). For the outer side wall, acrylic multi-layer panels and polycarbonate panels were used for the lower (up to 2 m from the ground) and the remaining upper portions, respectively. The inside compartments were partitioned with figure and transparent glass for the lower and upper portion, respectively, and were spatially isolated from each other. The compartment was oriented northeast and southwest. The length, width, eave height, and ridge height of the compartment were approximately 160 m² and 1,100 m³, respectively. There were ventilators on each side of the roof and the ends of the ventilator openings were independently adjusted based on the levels of environmental factors inside and outside the greenhouse using a greenhouse environmental measurement system (Priva Connex; Priva B.V., De Lier, The Netherlands). On the southwest wall, there were four fans with louvers. Fans were off and louvers were open during the experiment. The entire floor inside the greenhouse was covered with concrete or impermeable mulch. There were no crop plants in the compartment during measurements.

2.2 Measurements

Air temperature inside \( (T_i) \) and outside \( (T_o) \) the greenhouse and solar irradiance outside the greenhouse \( (I) \) were measured and recorded by the greenhouse environmental measurement system. The extent of the windward \( (R_w) \) and leeward \( (R_l) \) roof ventilator openings was evaluated as the relative vertical height of the openings to the maximum height (1.16 m). The wind speed \( (U) \) and direction outside the greenhouse were measured with an anemometer (PGWS-100; Gill Instruments Limited, Hampshire, UK). The wind speed components orthogonal and parallel to the ridgepole \( (U_x \) and \( U_y \), respectively) were calculated. To calculate \( N \), pure CO₂ gas was released into the greenhouse at a constant rate of 10 L min⁻¹ \( (S) \) at 20°C, 101 kPa using a mass flow controller (3655; KOFLOC Corp., Kyoto, Japan) throughout the daytime. The outlet of gas tube for pure CO₂ was fixed in front of an oscillating fan that was placed at the center of the greenhouse compartment so as to homogenize CO₂ concentration in it. The concentration of CO₂ inside \( (C_i) \) and outside \( (C_o) \) the greenhouse was measured using two nondispersive infrared CO₂ analyzers (ZFP9; Fuji Electric Co., Ltd., Kanagawa, Japan). The air inside and outside the greenhouse was passed through a thermoelectric dehumidifier (DH-209C-1-R; KELK Ltd., Kanagawa, Japan) and supplied to the analyzers at a constant flow rate of 0.5 L min⁻¹ \( (20°C, 101 kPa) \) using mass flow meters (3810DS-V; KOFLOC Corp.). The data were recorded using a data logger (GL-220; Graphtec Corp., Yokohama, Japan). The environmental data were recorded every 10 s, except for \( I, R_w, \) and \( R_l \) which were recorded every 5 min.

The experiments were conducted in the winter, summer, and fall of 2019 (Table 1). We excluded all-day data, including the hours when \( T_i \) in summer exceeded 50°C, as it was in the non-operating range of the mass flow controller for pure CO₂ gas released into the greenhouse.

2.3 Calculation of the number of air exchanges per hour

The \( N \) of a greenhouse is defined by the following differential equation:

\[
\frac{dC_i}{dt} = N(C_o - C_i) + \frac{S - P_s}{V}
\]

where \( P_s \) is the amount of net photosynthetic CO₂ absorption of all crops in the greenhouse per unit of time. Assuming that \( P_s = 0 \) for an uncropped greenhouse,

\[
\frac{dC_i}{dt} = N(C_o - C_i) + \frac{S}{V}
\]

To solve Eq. 2 to obtain \( N \), the differential coefficient of the left differential term at time \( t = t_s \) [min] was approximated as:

\[
\frac{dC_i}{dt}(t_s) = \frac{\bar{C}_i(t_{s+1}) - \bar{C}_i(t_s)}{(t_{s+1} - t_s)} = \bar{C}_i(t_s) - \bar{C}_i(t_{s+1})
\]

where \( \bar{C}_i(t_s) \) is the centered moving average of \( C_i \) between \( t_s - 15 \) and \( t_s + 15 \).

2.4 NN model

The NN model used to estimate \( N \) was the same as that reported by Matsuda et al. (2019), with some exceptions. Measured variables used for the input layer (independent variables or the feature vector) were \( U_x, U_y, R_w, R_l, T_i, T_o, \) and \( I \). These were selected because an NN model with these variables and the wind direction as the input layer showed the highest \( N \) estimation accuracy in our previous study (Matsuda et al., 2019). As the effect of wind direction on \( N \) was not necessarily large (Matsuda et al., 2019), it was removed here. The input layer consisted of these variables measured at the same time as that for \( N \) estimation and those measured at 1, 2, 3, 4, and 5 min before the time for \( N \) estimation, assuming that the effects of environmental conditions on \( N \) can be delayed (Matsuda et al., 2019). For \( I, R_w, \) and \( R_l \), linear interpolation was used to estimate the 1-min data from the 5-min data. The output layer (or dependent variable) was \( N \). There were two hidden layers between the input and output layers, and each of the two hidden layers had 20 units. The hyperparameters used were the same as those reported by Matsuda et al. (2019), except that the batch size was 1,024. The NN model was trained with mini-batch gradient descent using backpropagation.

Table 1. Dates of measurements of environmental factors.

| Season   | Date (number of days)               | Number of data |
|----------|-------------------------------------|----------------|
| Winter   | Jan 23, 25–29, 31–Feb 4, 6–8, and 10 (15 days) | 18,213         |
| Summer   | Aug 12, 21, 22, 24–Sept 3, 5, 7–9, 13–15, 18–20 (24 days) | 25,589         |
| Fall     | Oct 30–Nov 6, 8–11 (12 days)       | 17,004         |
2.5 Intraseasonal applicability of NN model

The data measured between January 23 and February 4 were randomly divided into two datasets. The sizes were 75% and 25% of the original dataset, and the former and latter datasets were used as training and validation datasets, respectively. This validation dataset was referred to as the “simultaneous” validation dataset. The data measured between February 6 and 10 were also used as a test dataset, which was referred to as the “non-simultaneous” test dataset. This indicated that the data were acquired in a different period from that for the training dataset. Root mean square error (RMSE) and the coefficient for determination \( r^2 \) for the 1:1 relationship between the measured and the estimated \( N \) were compared.

2.6 Sampling of training data to adjust the relative frequency distribution of temperature difference

The relative frequency distribution of temperature difference inside and outside the greenhouse \((T_i - T_o)\), hereafter referred to as \( \Delta T \) in the dataset used to train the NN model was adjusted to approximate that in the test dataset based on the following steps:

(i) Two feature vectors \((U_{c_1}, U_{c_2}, R_{c_1}, R_{c_2}, T_i, T_o, \) and \( I \)) having the maximum \( \Delta T (\Delta T_{\text{max}}) \) and minimum \( \Delta T (\Delta T_{\text{min}}) \) values in the test dataset, respectively, were selected. The difference between \( \Delta T_{\text{max}} \) and \( \Delta T_{\text{min}} \) \((\Delta T_{\text{max}} - \Delta T_{\text{min}})\) was divided into 500 equal intervals, \( d_i = \{ k = 1, 2, \ldots, 500 \} \):

\[
\{ d_i | \Delta T_{\text{max}} + (k-1) \frac{\Delta T_{\text{max}} - \Delta T_{\text{min}}}{500} \leq d_i < \Delta T_{\text{max}} + k \frac{\Delta T_{\text{max}} - \Delta T_{\text{min}}}{500} \}
\]

(ii) The number of feature vectors with \( \Delta T \) in each \( d_i \) of the test dataset \( c_i \) was counted. Then, the relative frequency distribution of \( \Delta T, p_i = c_i / \sum_{k=1}^{500} c_i \) \((0 \leq p_i \leq 1)\) was obtained for the test dataset.

(iii) In the training dataset, the feature vectors with \( \Delta T \) within each \( d_i \) were defined as the set \( X_i \). The number of elements of \( X_i \) was \( n_i \), and the total number of feature vectors of the entire training dataset was \( n = \sum_{i=1}^{500} n_i \). For the adjusted training dataset, the feature vectors were randomly sampled from each \( X_i \) without replacement so that the number of samples became \( n \cdot p_i \) for each \( d_i \). If \( n_i < n \cdot p_i \) for a given \( d_i \), the number of samples became \( n_i \). If \( p_i = 0 \) or \( n_i = 0 \) \((\text{i.e., } X_i \text{ was the empty set})\), the number of samples became 0.

(iv) Finally, the relative frequency distribution of \( \Delta T \) of the adjusted training dataset was obtained, approximating that of the test dataset.

2.7 Intersesonal applicability of NN model

Training and test datasets were taken from winter, summer, and fall with or without adjustments of the training datasets (Table 1). RMSE and \( r^2 \) between the measured and the estimated \( N \) for the test datasets were compared.

3. Results and Discussion

Table 2 shows the summary of environmental conditions inside and outside the greenhouse and daily mean number of air exchanges of the greenhouse during the experimental periods for three seasons.

| Table 2. Daily means or integral of the levels of environmental elements inside and outside the greenhouse and daily mean number of air exchanges of the greenhouse during the experimental periods for three seasons. |
|---------------------------------|----------------|----------------|----------------|----------------|
| Daily mean or integral inside the greenhouse | Winter \(^{a}\) | Summer \(^{b}\) | Fall \(^{c}\) |
| Daily mean wind speed outside the greenhouse [m s \(^{-1}\)] | 1.7 ± 0.96 | 1.3 ± 0.39 | 1.5 ± 0.39 | 1.1 ± 0.44 | 1.2 ± 0.19 | 0.8 ± 0.19 |
| Daily integral of solar irradiance outside the greenhouse [MJ m \(^{-2}\)] | 11.4 ± 3.15 | – | 13.7 ± 0.39 | – | 11.5 ± 2.87 | – |
| Daily mean air temperature outside the greenhouse [°C] | 4.3 ± 5.72 | 3.7 ± 2.9 | 26.5 ± 2.93 | 23.0 ± 2.31 | 15.4 ± 1.39 | 11.3 ± 1.75 |
| Daily mean air temperature inside the greenhouse [°C] | 14.4 ± 5.84 | 12.5 ± 3.44 | 34.9 ± 3.57 | 25.9 ± 1.99 | 25.6 ± 1.6 | 14.2 ± 1.56 |
| Daily mean CO\(_2\) concentration outside the greenhouse [\(\mu\)mol mol \(^{-1}\)] | 515 ± 38.8 | 522 ± 25.7 | 414 ± 11.1 | 433 ± 17.6 | 431 ± 7.7 | 454 ± 16.6 |
| Daily mean CO\(_2\) concentration inside the greenhouse [\(\mu\)mol mol \(^{-1}\)] | 710 ± 98.1 | 655 ± 49.6 | 689 ± 85.1 | 474 ± 28.4 | 670 ± 45.2 | 610 ± 71.6 |
| Daily mean number of air exchanges of the greenhouse [h \(^{-1}\)] | 1.4 ± 0.55 | 0.9 ± 0.53 | 2.0 ± 0.50 | 0.5 ± 0.13 | 2.1 ± 0.68 | 0.3 ± 0.15 |

\(^{a}\) Means ± standard deviations among days

\(^{b}\) Daily mean or integral of 1-min intervals was calculated for daytime and nighttime in each day, respectively, and then averaged during the experimental period.
To simplify calculations, we assume a steady state of $dC_i/dt = 0$. Then, by solving Eq.1 for $P_n$,

$$P_n = S + NV(C_a - C_i)$$

(5)

is obtained. Considering the propagation of uncertainty, the relationship between standard deviation of uncertainty for $P_n$ quantification, $\sigma_{P_n}$ [mol h$^{-1}$], and that for $N$ estimation, $\sigma_N$ [h$^{-1}$], is given as follows:

$$\sigma_{P_n} = |V(C_a - C_i)| \cdot \sigma_N$$

(6)

As a realistic condition for the greenhouse used in this study while plants are assumed to be grown therein, we assign $V = 1,100$ m$^3 \approx 45,500$ mol at 20°C, 101 kPa, $C_a = 400$ μmol mol$^{-1}$, and $C_i = 1,000$ μmol mol$^{-1}$, and

$$\sigma_{P_n} = 27.3 \sigma_N$$

(7)

is obtained. Because the standard deviation and RMSE will have the same order of magnitude, we substitute a value of 0.63 h$^{-1}$, which was observed as RMSE in the “non-simultaneous” test dataset (Fig. 1), for $\sigma_N$, and finally a $\sigma_{P_n}$ of 17.1 mol h$^{-1}$ is obtained. Next, let us assume that plants grown in the greenhouse have an average leaf net photosynthetic rate of 10 μmol m$^{-2}$ s$^{-1}$ and a leaf area index of 4 m$^2$ m$^{-3}$. Given that the floor area of the greenhouse was 160 m$^2$, $P_n$ is computed at 23.0 mol h$^{-1}$ (10 × 10$^{-8}$ × 3,600 × 4 × 160 = 23.0). Collectively, the $\sigma_{P_n}$ value of 17.1 mol h$^{-1}$ indicates that $P_n$ can often be 74% over- or under-estimated. Conversely, if one wants to quantify the $P_n$ within a lower relative error of 50% (i.e., 23.0 ± 11.5 mol h$^{-1}$) for this hypothetical greenhouse, for example, RMSE for $N$ estimation must be 0.42 h$^{-1}$ or lower, according to Eq. 7. Furthermore, if measurement uncertainties of $V$, $C_a$, and $C_i$, which have been ignored in this discussion, are taken into consideration, the required accuracy of $N$ estimation will be much higher. Thus, although the above calculations are a rough estimate in a certain case and not necessarily quantitatively strict, the estimation accuracy of the NN model applied to the

Fig. 1. The relationship between the measured numbers of air exchanges per hour ($N$) and estimated $N$ using the neural network (NN) model for the “simultaneous” validation dataset (white circles) and “non-simultaneous” test dataset (dark grey circles). The solid line is the 1:1 line. For the “simultaneous” validation dataset, RMSE = 0.48 h$^{-1}$ and $r^2 = 0.75$; for the “non-simultaneous” test dataset, RMSE = 0.63 h$^{-1}$ and $r^2 = 0.52$.

Fig. 2. Time course of the measured numbers of air exchanges per hour ($N$) and estimated $N$ using the neural network (NN) model (a, b), measured air temperature inside ($T_i$) and outside ($T_o$) the greenhouse (c, d), and measured wind speed ($U$) and solar irradiance ($I$) outside the greenhouse (e, f) between 8:00 and 17:00 on February 6 (a, c, e) and 8 (b, d, f).
“non-simultaneous” test dataset must be quite insufficient in terms of monitoring $P_a$ in a greenhouse.

Nevertheless, looking at the time course of the diurnal measurements and estimation of $N$, the trend of measured and estimated $N$ in the non-simultaneous test dataset changed daily. For example, the estimated $N$ on February 6 fluctuated, but the measured $N$ was fairly constant (Fig. 2a). On the other hand, the measured and estimated $N$ on February 8 agreed relatively well except at 11:00–12:00 and 14:00–15:00 (Fig. 2b). No clear correlation was observed in the daily environmental conditions and estimation accuracy (Fig. 2c–f). The factors underlying the variation in estimation accuracy have not yet been determined.

One of the reasons for the low estimation accuracy against a non-simultaneous test dataset can be the “covariate shift” that is common in supervised learning (Shimodaira, 2000; Sugiyama, 2006; Sugiyama et al., 2014). Covariate shift means that the distribution of covariates (or model input variables) differ between the training and test datasets, while the input-output relationship does not change. We hypothesized that the estimation accuracy can be improved if the distribution of model-related parameters in the training dataset is similar to that of the test dataset. As such parameter candidates, we adjusted the relative distribution of $U$, $T_i$, $T_o$, $\Delta T$ or $I$. The data used for training were sampled so that the distribution resembled that of the test dataset (Fig. 3, see Materials and Methods for details). When we adjusted the distribution of $\Delta T$ in the training dataset, the estimation accuracy of the NN model trained with the adjusted dataset was slightly improved for the non-simultaneous test dataset compared to the model trained with the raw dataset (Fig. 4). Adjusting the training dataset appears to improve the estimation accuracy of $N$. In practice, the relative distribution of $\Delta T$ in the test dataset is not known a priori at the time which the model was trained. Instead, the relative distribution of $\Delta T$ acquired in the target season of the previous years can be used as a reference.

Next, we applied this method to interseasonal data. In contrast to the intraseasonal data (Fig. 4), estimation accuracy was not improved by relative distribution adjustment (Table 3). Therefore, relative distribution adjustments worked for intraseasonal data, but not for interseasonal data.

In this work, the number of data points or the size of data used for model training may be small (Table 1), relative to the input-layer dimension of 42. This could lead, at least in part, to the observed lower accuracy of $N$ estimation through insufficient model generalization. Indeed, an overfitting might occur in such parameter candidates.

![Fig. 3. The relative distributions of the difference between inside and outside air temperatures ($\Delta T$). (a) Raw training (black bars) and test (grey bars) datasets. (b) Adjusted training (black bars) and test (grey bars) datasets. The training dataset was adjusted so that the distribution resembled that for the test dataset.](image)

![Fig. 4. The relationship between the measured numbers of air exchanges per hour ($N$) and estimated $N$ using the neural network (NN) model trained with the raw dataset (dark grey circles) and the adjusted dataset (light grey circles). The solid line is the 1:1 line. For the raw training dataset, RMSE = 0.63 h$^{-1}$ and $r^2 = 0.52$; for the adjusted training dataset, RMSE = 0.51 h$^{-1}$ and $r^2 = 0.66$.](image)

Table 3. Root mean square errors (RMSEs) and the coefficients of determination ($r^2$) for the relationship between measured and estimated $N$ using a neural network (NN) model. The NN model was trained with raw or adjusted training data and validated, using different combinations of data collected in three seasons.

| Validation data | RMSE for raw/adjusted training data [h$^{-1}$] | $r^2$ for raw/adjusted training data |
|-----------------|-----------------------------------------------|-------------------------------------|
| Winter          | –                                             | 0.07 / 0.05                          |
| Summer          | 1.51 / 2.39                                   | 0.01 / 0.17                          |
| Fall            | 0.94 / 1.00                                   | 0.29 / 0.32                          |

*The training data was adjusted so that the relative frequency distribution of the temperature difference approximated to that of the validation data.*
our NN model; estimated N fluctuated throughout the daytime in February 6, while measured N did not (Fig. 2a). Although a clear correlation with $T_\text{p}$, $T_\text{w}$, $U$ or $I$ was not apparently found (Fig. 2c, e), a small fluctuation in one or more input variables or in their covariances might lead to the fluctuation of estimated $N$. On the other hand, measured and estimated $N$ agreed better in February 8 (Fig. 2b), although $T_\text{p}$, $T_\text{w}$, $U$, and $I$ fluctuated more (Fig. 2d, f) than those in February 6. Thus, the reason for the possible overfitting was not identified. We selected the limited data size for model training, considering the year-round greenhouse production in reality; data collection for model training must be carried out in a cropless greenhouse between crop seasons, and the period should be limited in general to a month or shorter. To improve the estimation accuracy with the limited data size, application of feature extraction or feature reduction techniques such as the principal component analysis to model inputs might contribute.

Nevertheless, the estimation accuracy of the NN model in the case of interseasonal applications was too low (Table 3), strongly suggesting that we need to employ a completely different approach to improve it, in addition to the input variable adjustments. Possible solutions toward the interseasonal NN model application includes continuous model correction while the model is running in a cropped greenhouse. Although it may be considered somewhat unrealistic, we can consider a miniature model greenhouse without plants, which is built alongside the target greenhouse and equipped with environment measurement and control system similar to that used for the target greenhouse. The vacant model greenhouse is operated similarly to the target greenhouse where crops are cultivated, while actual $N$ is continuously measured for the model greenhouse. By utilizing such data obtained from the model greenhouse, the NN model under operation in the target greenhouse could dynamically be corrected.

In summary, the estimation accuracy of the NN model for $N$ of a naturally ventilated greenhouse was low when the dataset used to train the model differed in time from the model test. Adjusting the relative distribution of $\Delta N$ used for training slightly improved the estimation accuracy when the model was applied at the time immediately following training in the same season. However, the estimation accuracy was not improved when the model was applied to periods in different seasons. The method proposed here can only be used to estimate $N$ if model validation and testing are close in time. A novel dataset adjustment or fundamental improvement of the estimation procedure is needed to year-round practical use.

Acknowledgments

We would like to thank Seiwa Co., Ltd. for their technical support for measurements in the experimental greenhouse. This work was financially supported in part by a grant from commissioned project study on “the research project for the future agricultural production utilizing artificial intelligence”, Ministry of Agriculture, Forestry and Fisheries.

References

Boulard T, 2006: Greenhouse natural ventilation modelling: a survey of the different approaches. Acta Horticulturae 719, 29–40.

Boulard T, Baille A, 1995: Modelling of air exchange rate in a greenhouse equipped with continuous roof vents. Journal of Agricultural Engineering Research 61, 39–48.

Boulard T, Draoui B, 1995: Natural ventilation of a greenhouse with continuous roof vents: measurements and data analysis. Journal of Agricultural Engineering Research 61, 27–36.

Boulard T, Feulillepy P, Kittas C, 1997: Natural ventilation performance of six greenhouse and tunnel types. Journal of Agricultural Engineering Research 67, 249–266.

Boulard T, Meneses JF, Mermier M et al., 1996: The mechanisms involved in the natural ventilation of greenhouses. Agricultural and Forest Meteorology 79, 61–77.

Fernandez JE, Bailey BJ, 1993: Predicting greenhouse ventilation rates. Acta Horticulturae 328, 107–114.

Kittas C, Boulard T, Papadakis G, 1997: Natural ventilation of a greenhouse with ridge and side openings: sensitivity to temperature and wind effects. Transactions of the ASAE 40, 415–425.

Kozai T, Sase S, Nara M, 1980: A modelling approach to greenhouse ventilation control. Acta Horticulturae 106, 125–136.

Matsuda R, Hayano K, Yagi S et al., 2019: Comparison of neural network models with aerodynamic and empirical models toward real-time estimation of the number of air exchanges per hour of a naturally ventilated greenhouse. Journal of Agricultural Meteorology 75, 166–172.

Nederhoff EM, van de Vooren J, Udink ten Cate AJ, 1985: A practical tracer gas method to determine ventilation in greenhouses. Journal of Agricultural Engineering Research 31, 309–319.

Okada M, Takakura T, 1973: Guide and data for greenhouse air conditioning. 3. Heat loss due to air infiltration of heated greenhouse. Journal of Agricultural Meteorology 28, 223–230.

Sase S, Kozai T, Nara M et al., 1980: Ventilation of greenhouses. I. Wind tunnel measurements of pressure and discharge coefficients for a single-span greenhouse. Journal of Agricultural Meteorology 36, 3–12 (in Japanese with English abstract and captions).

Seginer I, Boulard T, Bailey BJ, 1994: Neural network models of the greenhouse climate. Journal of Agricultural Engineering Research 59, 203–216.

Shimodaira H, 2000: Improving predictive inference under covariate shift by weighting the log-likelihood function. Journal of Statistical Planning and Inference 90, 227–244.

Sugiyama M, 2006: Supervised learning under covariate shift. The Brain & Neural Networks 13, 111–118 (in Japanese).

Sugiyama M, Yamada M, du Plessis MC et al., 2014: Learning under non-stationarity: covariate shift adaptation, class-balance change adaptation, and change detection. Journal of the Japan Statistical Society, Japanese Issue 44, 113–136 (in Japanese with English abstract).

Takakura T, Kubota C, Sase S et al., 2009: Measurement of evapotranspiration rate in a single-span greenhouse using the energy-balance equation. Biosystems Engineering 102, 298–304.

Takakura T, Sunagawa H, Tamaki M et al., 2017: In situ net photosynthesis measurement of a plant canopy in a single-span greenhouse. Journal of Advances in Agriculture 7, 1015–1020.