Journal of the Meteorological Society of Japan

EARLY ONLINE RELEASE

The DOI for this manuscript is
DOI:10.2151/jmsj.2025-026
J-STAGE Advance published date: May 2, 2025
The final manuscript after publication will replace the preliminary version at the above DOI once it is available.

Re-Examination of Errors and Exclusion Range, and Gap-filling for Surface Flux Estimation using the Bowen Ratio Method

Shigenori HAGINOYA

Meteorological Research Institute, Tsukuba, Japan

and

Yasuhiro KAWABATA

Meteorological Research Institute, Tsukuba, Japan

April 9, 2025

Corresponding author: Shigenori Haginoya, Meteorological Research Institute,
1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan.
E-mail: shaginoy@mri-jma.go.jp

Abstract

This paper presents a method for estimating turbulent fluxes using the Bowen ratio method. We propose a data exclusion criterion for the Bowen ratio method, defining an exclusion range for the Bowen ratios (B) as $B_- < B < B_+$. To determine B_- and B_+, we evaluated the flux difference between the Bowen ratio method and the eddy covariance method and suggested $B_- = -2.0$ and $B_+ = -0.6$. We also propose an interpolation method to handle missing data based on the exclusion criteria by interpolating $B/(1+B)$ or $1/(1+B)$. By implementing these approaches, we establish a comprehensive and practical framework for estimating turbulent fluxes with the Bowen ratio method. We expect that these results will enhance the accuracy and reliability of flux estimates in various environmental and climate studies.

Keywords Bowen ratio method; surface flux; gap-filling; exclusion criteria; spike-error

1. Introduction

The surface energy balance is crucial for understanding the mechanisms of atmosphere−surface interactions. In recent years, turbulent flux observations using the eddy covariance method (ECM) have become widespread (FLUXNET: Baldocchi et al. 2001; Kaimal and Finnigan 1994; Foken 2017). Long-term datasets are essential for these observations. Several operational considerations and issues have been identified with the ECM (Allen et al. 2011; Cook and Sullivan 2020). In addition, the instrument's probe surface can become covered with a film of water during precipitation or condensation. Since this water can cause noise problems, countermeasures are required (Campbell 2013, 2015). It is therefore important to complement the ECM with measurements by other methods to minimize data gaps.

In this paper, we apply the Bowen ratio method to address these challenges. The Bowen ratio method has been widely used to estimate sensible and latent heat fluxes using time-averaged values, such as two-level air temperature and humidity measurements, net radiation, and soil heat flux (Stull 1988; Garratt 1994; Arya 2001; Allen et al. 2011). It is crucial to efficiently exclude inappropriate data when applying the Bowen ratio method.

Ohmura (1982) proposed a systematic approach for determining data exclusion criteria that uses objective indicators. Subsequent studies have used Ohmura's method as a foundation. Payero et al. (2003) compiled guidelines for detecting erroneous Bowen ratio data. Foken (2017) summarized the knowledge and the limitations associated with the Bowen ratio method. The Bowen ratio method has the main advantage of introduction to a ratio of sensible heat (H) to latent heat (lE), the Bowen ratio ($B = H/lE$) so that the diffusion coefficient does not need to be determined. However, the Bowen ratio near $B = -1$ can cause errors (spike-error) in H and lE. In principle, this is an unavoidable problem because the factor "$1/(1 + B)$" appears when the flux is determined by the Bowen ratio method (e.g., Stull 1988). Previous studies have proposed direct exclusion ranges for data near $B = -1$ (Unland et al. 1996; Halliwell and Rouse 1989; Tanner et al. 1987). Halliwell and Rouse (1989) obtained exclusion ranges based on a comparison between the aerodynamic method and the Bowen ratio method. Perez et al. (1999) suggested a method to analytically determine the exclusion range near $B = -1$ based on sensor accuracy and differences in observed temperature and vapor pressure. In this paper, the exclusion range near $B = -1$ is re-evaluated using the ECM and a new exclusion range is proposed.

The exclusion of data reduces the amount of data available for analysis, and the excluded data appear as missing data. Such missing values in time-series data need to be filled with appropriate values to construct a long-term flux dataset. It is therefore necessary to develop gap-filling methods using available data to minimize data loss. This study applies linear interpolation for simplicity.

Considering these issues, this paper presents a coherent and practical approach to data exclusion and interpolation for missing data when the Bowen ratio method is used.

Section 2 provides an overview of the Bowen ratio method and the criteria for data exclusion. Section 3 details the observation method and data processing procedures. Section 4 presents verification criteria and the interpolation method for gap-filling. Section 5 reports the application results. Section 6 compares this method with other surface flux estimation approaches. Section 7 provides a summary.

2. Method

2.1 Description of Method

The Bowen ratio method, which was originally proposed by Bowen (1926), has been widely used to estimate turbulent fluxes. This method uses temperature (T_1, T_2) and specific humidity (q_1, q_2) differences ($\Delta T = T_2 - T_1, \Delta q = q_2 - q_1$) measured at two heights (z_1 and z_2, where $z_2 > z_1$)

within the surface layer, along with the net radiation (Rn) and soil heat flux into the soil (G), to determine the sensible heat (H) and latent heat (lE) fluxes (e.g., Garratt 1994).

The Bowen ratio (B) is defined as the ratio of the sensible heat to latent heat fluxes:

$$B = \frac{H}{lE}. \tag{1}$$

The expressions for H and lE using diffusivities are as follows:

$$H = -K_h C_p \rho \frac{\Delta T}{\Delta z}, \tag{2a}$$

$$lE = -K_v l \rho \frac{\Delta q}{\Delta z}. \tag{2b}$$

Here, K_h and K_v represent the thermal and water vapor diffusivities, respectively. C_p is the specific heat capacity of air, ρ is the density of air, l is the latent heat of vaporization, E is the evaporation rate, and $\Delta z = z_2 - z_1$. Positive H and lE indicate upward fluxes. The relationship between flux and gradient, as in Eqs. (2a) and (2b), is based on K-theory (Stull 1988).

Assuming that transport processes of heat and water vapor are similar within the surface layer, we equated the thermal and water vapor diffusivities. Substituting Eqs. (2a) and (2b) into Eq. (1) to obtain Eq. (3):

$$B = \frac{C_p}{l} \frac{\Delta T}{\Delta q}. \tag{3}$$

Thus the Bowen ratio can be calculated from the differences in temperature and specific humidity. In contrast, we assumed that the surface energy balance equation is given by

$$Rn - G = H + lE. \tag{4}$$

Positive Rn and G indicate downward fluxes. The left-hand side of Eq. (4) defines the available energy at the surface (Garratt 1994).

By combining Eqs. (1) and (4), we can derived expressions for H and lE in terms of the available energy $Rn - G$ and the Bowen ratio:

$$H = (Rn - G)\frac{B}{1+B}, \tag{5a}$$

$$lE = (Rn - G)\frac{1}{1+B}. \tag{5b}$$

Applying Eq. (3) to Eqs. (5a) and (5b), we can calculate H and lE from the temperature difference ΔT and the specific humidity difference Δq at the two heights.

2.2 Considerations for Using the Bowen Ratio Method

a. When $|\Delta T|$ and $|\Delta q|$ are small

Based on the error analysis of Eq. (3), it follows that:

$$\left|\frac{\delta B}{B}\right| \leq \left|\frac{\delta \Delta T}{\Delta T}\right| + \left|\frac{\delta \Delta q}{\Delta q}\right|, \tag{6}$$

where, the symbol δ is indicates an error (Bevington and Robinson 2003). If the differences $|\Delta T|$ and $|\Delta q|$ at the two heights are less than the instrument accuracy, $|\delta\Delta T|$ and $|\delta\Delta q|$, respectively, the error in the Bowen ratio itself becomes significant. However, when $|\Delta T| < |\delta\Delta T|$ and $|\Delta q| \geq |\delta\Delta q|$, the impact of $|\Delta T|$ on the Bowen ratio (B) is negligible in practice (Allen et al. 2011). This conclusion follows from the fact that as $\Delta T \to 0$, Eqs. (3), (5a), and (5b) imply that $B \to 0$, $H \to 0$, and $lE \to Rn - G$. Similarly, when $|\Delta q| < |\delta\Delta q|$ and $|\Delta T| \geq |\delta\Delta T|$, if $\Delta q \to 0$ so that $|\Delta q| << |\delta\Delta q|$, the reciprocal of Eq. (3) approaches $1/B \to 0$. Consequently, Eqs. (5a) and (5b) give $H \to Rn - G$ and $lE \to 0$. There is hence no problem in applying Eqs. (5a) and (5b). However, when both ΔT and Δq approach zero at the same time, Eq. (3) becomes indeterminate, and Eqs. (5a) and (5b) become inapplicable. In such cases, flux data are considered to be missing.

b. When B is Close to -1

The error analysis of Eqs. (5a) and (5b), leads to:

$$\left|\frac{\delta H}{H}\right| \leq \left|\frac{\delta(Rn-G)}{Rn-G}\right| + \left|\frac{\delta B}{B}\right| + \left|\frac{\delta B}{1+B}\right|, \tag{7a}$$

$$\left|\frac{\delta lE}{lE}\right| \leq \left|\frac{\delta(Rn-G)}{Rn-G}\right| + \left|\frac{\delta B}{1+B}\right|. \tag{7b}$$

When the denominator $|1 + B|$ in Eqs. (5a) and (5b) is small, (i.e., when B is close to -1), the estimation errors of sensible heat and latent heat fluxes in Eqs. (7a) and (7b) become significant because of the factor $1/(1 + B)$.

The condition $B \approx -1$ occurs when $H \approx -lE$ and, based on Eq. (4), when $Rn \approx G$. Here, we consider specific cases in which this situation occurs. Under calm or light wind conditions, the absolute values of sensible and latent heat fluxes are small within the range of measurement error ($H \approx -lE \approx 0$). Conversely, in other situations, sensible heat is transported to the surface to compensate for the latent heat loss to the atmosphere due to evaporation ($H \approx -lE \neq 0$). This condition often occurs under Foehn conditions, particularly in strong wind events (Ohmura 1982; Hayashi et al. 2005; Hofsteenge et al. 2022). In the latter case, the sensible and latent heat fluxes are approximately balanced at a finite value. However, it is evident from Eqs. (7a) and (7b) that the fluxes obtained under these conditions have large errors.

Various techniques have been proposed to objectively address the above limitations and challenges inherent in the Bowen ratio method. Inspired by previous studies, we established criteria for excluding data.

c. Data Exclusion Criteria

As mentioned in Section 2.2a, we did not consider excluding data based on instrument accuracy. If either the temperature difference ΔT or the specific humidity difference Δq, but not both, approached zero, we used Eqs. (3), (5a), and (5b) to estimate H and lE. If both ΔT and Δq approached zero simultaneously, we treated the flux data as missing. We established two criteria for excluding data.

Criterion I: Condition for sign consistency of observed values related to K-theory (Ohmura 1982, Problem 2).

It follows from Eqs. (2a), (2b), and (4) that:

$$Rn - G = \frac{K_v l \rho}{\Delta z} \left(\frac{C_p}{l} \frac{K_h}{K_v} \Delta T + \Delta q \right). \tag{8}$$

Both sides of Eq. (8) should have the same sign. Considered that $K_v > 0$ and $K_h/K_v = 1$, the data exclusion criterion is following:

$$(Rn - G) \times \left(\frac{C_p}{l} \Delta T + \Delta q \right) < 0. \tag{9}$$

Criterion II: Exclusion of singular points in the Bowen ratio method (Ohmura 1982, Problem 3).

$B = -1$ is a singularity where the denominators of Eqs. (5a) and (5b) become zero. In practice, because B is derived from observational data that include errors, the singularity affects the calculations even in the vicinity of $B = -1$. Therefore, we excluded data near $B = -1$.

Data exclusion criteria can be established by directly defining a finite range for the Bowen ratio. Proposed ranges for exclusion include $B < -0.75$ (Ortega-Farias et al. 1996), $-1.3 < B < -0.7$ (Unland et al. 1996), $-1.4 \leq B \leq -0.7$ (Halliwell and Rouse 1989), and $-1.25 < B < -0.75$ (Tanner et al. 1987; Foken 2017). In this study, we defined the lower limit B_- and upper limit B_+ for the exclusion range when $B_- < B < B_+$. In Section 4.2

we determine the optimal combination of B_{-} and B_{+} that minimizes the flux difference between the Bowen ratio method and the ECM.

3. Observations

3.1 Field and Sensors

We applied the proposed method to field data collected at the Meteorological Research Institute (36° 03'17.5"N, 140° 07'24.2"E). Figure 1 shows a panoramic view of the observation site from Google Maps. The circle and square mark the locations of the eddy covariance system and the Bowen ratio system, respectively. They are about 15 m apart. The observation field is described in detail in Mouri et al. (2019). The field is 200 m long (north–south) and 100 m wide (west–east). A wooded area lies along the western border of the observation field. The surface condition is nearly homogeneous up to 100 m of the northern border. Scattered trees (≤ 10 m tall) are present between 100 and 300 m north, and the wooded area extends further beyond 300 m north. To measure surface fluxes from the ECM, we installed a three-dimensional ultrasonic anemometer CSAT3 (Campbell 2015) and a gas analyzer EC150 (Campbell 2013), facing the north, at the center of the observation field. The sensor height was set at $z = 1.8$ m. On the other hand, the sensors of the Bowen ratio system were set at $z_1 =$

0.4 m and $z_2 = 2.5$ m. The observation tower was located 5 m west of the Bowen ratio system and measured wind speed at 7.5 m height (Fig. 2a). Table 1 summarizes the instruments used for observation in this research. The original 10-minute data were averaged to 30-minute intervals for further analysis. All fluxes in this study were calculated at 30 minutes intervals. The observational dataset covered two periods:

Period 1: 20 November 2018 to 26 April 2019

Period 2: 20 November 2019 to 26 April 2020.

The grass height remained below 10 cm during both periods. Data from these periods were selected during seasons with minimal influence of vegetation to focus on the methodology of exclusion criteria. Note that data from Period 1 were used in Section 4. Verification of Criteria, Section 5. Resuts, and Section 6. Discussion, whereas data from Period 2 were used in Section 5. Results.

Fig. 1

Fig. 2

Table 1

Figure 3 shows a wind rose for the observation period. The wind rose from the 7.5 m tower anemometer (Fig. 3a) and the 1.8 m ultrasonic anemometer (Fig. 3b) differ. The wind rose at 7.5 m indicates wind observation from the west of the site. On the other hand, the wind rose at 1.8 m is dominated by NNW winds. The NNW to NE sector remains open, with grassland over 100 m. Winds from these sectors accounted for 48.8%.

According to the footprint analysis, surface conditions up to 100 m upwind influence more than 70% of ECM-derived fluxes (see Appendix A). Radiation temperatures remained relatively uniform across the field after weeding (Ichinose and Pan 2023). Fig. 3

3.2 Data Processing

a. Correction of Temperature and Humidity Sensor Bias

We used an independent thermometer and hygrometer to measure differences in temperature and humidity. To determine the bias between the two sensors, we obtained comparative observations by installing the temperature and humidity sensors in a ventilated enclosure at the same height for approximately one week every two months (Fig. 2b). We defined the biases for temperature and relative humidity as $\Delta\widetilde{T}(\widetilde{T_1}) = \widetilde{T_1} - T_2$ and $\Delta\widetilde{rh}(\widetilde{rh_1}) = \widetilde{rh_1} - rh_2$, respectively. Here, $\widetilde{T_1}$ represents the temperature at height z_1 before bias correction, and $\widetilde{rh_1}$ represents the corresponding relative humidity. The baiases are assumed to take the following form:

$$\Delta\widetilde{T}(\widetilde{T_1}) = a \times \widetilde{T_1} + b, \tag{10a}$$

$$\Delta\widetilde{rh}(\widetilde{rh_1}) = c \times \widetilde{rh_1} + d, \tag{10b}$$

where a, b, c, and d are regression coefficients (see Appendix B). After bias correction, the temperature and relative humidity values are given by

$$T_1 = \widetilde{T}_1 - \Delta\widetilde{T}(\widetilde{T}_1), \tag{11a}$$

$$rh_1 = \widetilde{rh}_1 - \Delta\widetilde{rh}(\widetilde{rh}_1), \tag{11b}$$

respectively. The specific humidity values q_1 and q_2 were calculated using the corrected temperature and relative humidity values (T_1, rh_1) and (T_2, rh_2), respectively. We then obtained the temperature difference ΔT and the specific humidity difference Δq after bias correction.

b. Net Radiation

We calculated the net radiation Rn using the following equation:

$$Rn = S^{\downarrow} - S^{\uparrow} + L^{\downarrow} - L^{\uparrow}, \tag{12}$$

where $S^{\downarrow}$ is the incoming shortwave radiation, $S^{\uparrow}$ is the reflected shortwave radiation, $L^{\downarrow}$ is the incoming longwave (infrared) radiation, and $L^{\uparrow}$ is the outgoing longwave (infrared) radiation.

c. Soil Heat Flux

The surface soil heat flux G consists of two components, as described by Garratt (1994, p. 116), Arya (2001, Chapter 4), and Foken (2017, p. 19):

$$G = G_{2cm} + S_{0-2cm}, \tag{13}$$

where G_z is the measured soil heat flux at depth z, and S_{0-2cm} is the heat storage in the 0−2 cm soil layer. We calculated the heat storage as follows:

$$S_{0-2cm} = \overline{c_g \rho_g} \frac{\Delta T_g(0-2cm)}{\Delta t} \Delta z(2cm), \tag{14}$$

where, $\Delta T_g(0-2cm)/\Delta t$ is the average change rate of soil temperature between 0 and 2 cm depth, and $\Delta z(2cm)$ is the thickness of the soil layer (2 cm in this case). The surface temperature (at a depth of 0 cm) was determined using radiation temperature measurements from the infrared radiometer and the radiation thermometer (see Table 1). We calculated the soil heat capacity $c_g \rho_g$ which accounts for the difference in soil heat flux between the 2 cm and 10 cm depths and the change rate of temperature between 2 cm and 10 cm:

$$c_g \rho_g = \frac{G_{2cm} - G_{10cm}}{\frac{\Delta T_g(2-10cm)}{\Delta t} \times \Delta z(8cm)}. \tag{15}$$

Although $c_g \rho_g$ generally varies with soil moisture θ, it is not appropriate in this study to express $c_g \rho_g$ as a function of θ, since the coefficient of determination (R^2) of the regression line between $c_g \rho_g$ and θ is 0.207. The variation of $c_g \rho_g$ averaged over 10 days within the period is ±7.4%, i.e. the variation of S_{0-2cm} (denoted as ΔS_{0-2cm}) due to the variation of $c_g \rho_g$ is ±7.4%. From the observed data, $|S_{0-2cm}/G| < 26.4\%$, and therefore, $|\Delta S_{0-2cm}/G| < 2.0\%$. For simplicity, we used the average value of $c_g \rho_g$, assuming it remains constant over the analysis period in this study. The value $\overline{c_g \rho_g}$ was obtained by averaging Eq. (15) over the measurement period and applied in Eq. (14).

d. Eddy Covariance Data

The data sampling frequency of the ECM was 10 Hz, with data recorded at 30-minutes interval. Sensible and latent heat fluxes were used only when the relative humidity difference between the VAISALA hygrometer (1.5 m) and the gas analyzer was less than 5%, since condensation on the sensor of gas analyzer could lead to erroneous readings of specific humidity. Condensation occurred in 98% of cases when the relative humidity exceeded 85%. Sensible heat flux from the ECM used to determine the exclusion range. Latent heat flux was excluded, as condensation on the sensor element could cause data loss.

4. Verification of Criteria

4.1 Dependence of Available Energy on Stability

We used only the data for Period 1 in Section 4. Figure 4 shows the relationship between available energy $Rn - G$ and atmospheric stability Ri (i.e. Richardson number) which is defined as:

$$Ri = \frac{g}{T} \frac{\Delta T}{U^2} \Delta z. \tag{16}$$

Here, Ri is similar to the bulk Richardson number (Arya 2001), g is the gravitational acceleration, T is the air temperature (in Kelvin) at a height of 1.5 m, ΔT is the temperature difference used in the Bowen ratio calculation, U is the wind speed at a height of 7.5 m, and $\Delta z = z_2 - z_1 = 2.1$ m.

When $Ri < 0$, most of the available energy $Rn - G$ is positive, i.e., the turbulent transport $H + lE$ is positive. In contrast, when $Ri > 0$, the available energy $Rn - G$ is both positive and negative, and we found that the average values of $Rn - G$ for each interval of Ri in the range $Ri > 0.001$ were -11 to -14 W m^{-2}. Furthermore, the upper limits of the error bar were also negative in the range of $Ri > 0.1$. For $Ri > 0.1$, most of the $Rn - G$ was in the negative range. Fig. 4

4.2 Determination of the Values of B_- and B_+

To determine Criterion II, we examined how the difference between the sensible heat flux from the Bowen ratio method and that from the ECM (*bias* and $DIFF$) varies with the exclusion range. We calculated *bias* and

$DIFF$ using the following equations:

$$bias = \frac{1}{n}\sum_{i=1}^{n} \Delta_i, \tag{17a}$$

$$DIFF = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {\Delta_i}^2}, \tag{17b}$$

$$\Delta_i = H_i(Bowen) - H_i(eddy), \tag{17c}$$

where $H_i(Bowen)$ is the sensible heat flux calculated using the Bowen ratio method under Criterion I and II with different range of the Bowen ratio, and $H_i(eddy)$ is the sensible heat flux calculated using the ECM. We calculated Δ_i when data were available for both $H_i(Bowen)$ and $H_i(eddy)$, where i is the serial number of the existing data and n is the total number of Δ_i. We also defined two additional parameters as follows: n_1 is the number of spike data with $|\Delta_i| > 40$ W m^{-2}, and n_0 is the number of data points where both Criteria I and II were applied. In the following, data under Criteria I and II will be referred to as valid data. Therefore, n and n_0 are both valid data, and $n_1 \leq n \leq n_0$ from the definition above.

As shown in Fig. 4, when the atmospheric condition is stable ($Ri > 0$), the majority of $Rn - G$ is below 50 W m^{-2} in case of $B \sim -1$, $Rn - G \sim 0$ (Section 2.2b), and $B \sim -1$ occurs from the evening until the next morning (Ohmura 1982; Tanner et al. 1987; Halliwell and Rouse 1989; Unland et al. 1996). Therefore, we focus on the data for $Rn - G < 50$ W m^{-2} to determine the range of the Bowen ratio B_- and B_+.

Figures 5a and 5b show the relationship between the *bias* and B_+ for various values of B_-. The bias decreases monotonically for both B_- and B_+. Figures 5c and 5d show the relationship between *DIFF* and B_+. It is apparent that larger B_- correspond to higher *DIFF*. Regardless of the value of B_-, the *DIFF* has a minimum value around $B_+ = -0.6$ (hereafter, Criterion 1). Figures 5e and 5f show the dependence of n_0/N on B_+. Larger B_- becomes larger n_0/N, and larger B_+ smaller n_0/N. Figures 5a, 5c, and 5e represent data under all wind directions, while Figs. 5b, 5d and 5f represent data under NNW−NE wind directions. There is no noticeable difference in the data between the two groups of wind direction. The exclusion ranges were determined based on the conditions that the *bias* is close to zero, the *DIFF* is as small as possible, and the n_0/N is as large as possible. These conditions give $B_- = -2.2$ to -2.0 (Criterion 2). Fig. 5

Figure 6 represents the frequency distribution of the number of spikes n_1 by Bowen ratio category. Data are plotted for two groups of wind direction: one is all wind directions, and the other is only NNW−NE winds. There were no significant differences between the two groups. The frequency of spikes is high for $-1.8 \leq B \leq -0.6$. Therefore, it is appropriate to set $B_- \leq -1.8$ (Criterion 3). As a result of Criteria 1 to 3, the exclusion range was determined as $(B_-, B_+) = (-2.0, -0.6)$. Fig. 6

Figure 7 shows the relationship between the Bowen ratio and surface fluxes without the application of Criterion II. H and lE are obtained using the Bowen ratio method. When the heat balance equation is satisfied, it is evident that in the vicinity of $B \approx -1$, the fluxes of both H and lE exhibit large absolute errors (spikes). However, no such spikes are observed in the available energy $Rn - G$. The H and lE spikes may therefore not represent actual fluxes (see Appendix C). Here, we applied Criterion II, as indicated by the red dashed lines in Fig. 7, with $(B_-, B_+) = (-2.0, -0.6)$. The spikes drastically decreased outside the red dashed lines. Fig. 7

Figure 8 shows the classification of data based on the application of Criteria I and II. For this figure, there was no restriction on the value of $Rn - G$. The data are divided into three categories: "I" represents data satisfying only Criterion I, "II" represents data satisfying only Criterion II, and "I and II" represents data satisfying both Criteria I and II. The proportion of category "II" underwent the most significant change as the B_- value was varied. By applying the data exclusion criteria, it was possible to calculate the fluxes using more than 80% of the data. Fig. 8

4.3 Gap-filling of Missing Data

In this study, we considered the interpolation of data excluded due to Criteria I and II. Surface flux estimation is examined in the cases where observational data other than those obtained using the Bowen ratio method cannot be used. The interpolation method should satisfy the energy balance equation, Eq. (4). We considered several interpolation methods:

[1] Interpolation of the Bowen ratio.

[2] Interpolation of sensible heat H and latent heat lE.

[3] Interpolation of temperature differences ΔT and specific humidity differences Δq.

[4] Interpolation of the coefficients in Eqs. (5a) and (5b), such as $B/(1+B)$ or $1/(1+B)$ (hereafter referred to as "B-factor"). Note that $B/(1+B)$ interpolation is equivalent to $1/(1+B)$.

To satisfy the energy balance equation, we used interpolation method [2] to recalculate the Bowen ratio using Eq. (1) based on the interpolated values. Similarly, we used interpolation method [3] to recalculate the Bowen ratio using Eq. (3) based on the interpolated values. Using these interpolated Bowen ratios, we calculated the sensible heat flux and latent heat flux with Eqs. (5a) and (5b), respectively.

Methods [1] to [3] resulted in a large amount of data that met the exclusion criteria (B_-, B_+), and their use caused significant spikes in the fluxes. Method [4] provided reasonable results that satisfied Criterion II. We therefore adopted B-factor interpolation in this study (see Appendix D).

We used B-factor interpolation to fill in the missing data from the extracted data under the condition $(B_{-}, B_{+}) = (-2.0, -0.6)$. We considered interpolations ranging from a 6-point interpolation (30 min $\times$ 6 = 3 h) to a 12-point (30 min $\times$ 12 = 6 h). For example, a 6-point interpolation means that intervals with six or fewer consecutive missing data points would be interpolated, while intervals with more than six missing data points would not be interpolated. The percentage of the original data retained was 84.5% (n_0 = 3883) (Table 2(a)). However, the percentages of data coverage with the 6-point and 12-point interpolations were 94.2% (n_0 = 5008) and 97.5% (n_0 = 5291), respectively. Interpolation resulted in a larger *mean* and *DIFF*, but a slightly smaller standard deviation (SD).

Figure 9 shows the frequency distribution of $Rn - G$ by class when the 12-point interpolation was performed. A total of 97.9% of the data fell within the condition $Rn - G < 50$ W m^{-2}. To investigate whether the average values of H for $Rn - G < 50$ W m^{-2} changed with the number of interpolations over the entire period, we performed a statistical test. If we assumed equal variances between pairs of samples from Table 2(b), we found no significant difference in the *mean* values between the 6-point and 12-point interpolations at a Type I error rate of 0.05. That result did not change even when we assumed unequal variances.

Table 2

Fig. 9

5. Results

To validate the appropriateness of the Bowen ratio data exclusion and interpolation methods, we compared data from November 2019 to April 2020 (Period 2) with the ECM. Figure 10 shows an example of time-series of the fluxes. The fluxes calculated by the Bowen ratio method and the interpolated fluxes (12-point interpolation) did not show large spikes and were in good agreement with fluxes by the ECM. As shown in Fig. 9, more than 97% of the interpolated data due to Criteria I and II occurred when $Rn - G < 50$ W m^{-2}. Therefore, the data exclusion and interpolation locations were concentrated in the evening, night, and morning hours.

Figure 10 shows that there were days with little or no error and days with a large error, especially in the latent heat fluxes during the day. Weather conditions during this period included heavy precipitation of 30 mm on April 1, clear skies from April 2 to 4, a southerly wind direction and dry air on April 4, and a significant decrease in soil moisture during the day. These conditions resulted in rapid drying state that may have prevented a uniform progression of surface drying throughout the day. Moreover, the wind direction was not from NNW−NE, which would be the ideal footprint condition. These factors may have caused the large difference in latent heat fluxes between the Bowen ratio method and the ECM. Fig. 10

Figure 11 represents the summary of the procedure for surface flux estimation. We can see the spike errors by the Bowen ratio method in Fig. 11a. In contrast, the spike errors are eliminated due to Criteria I and II (Fig. 11b). Figure 11c shows the interpolation results for the removed data. Compared to the ECM, no data deviate significantly, indicating that the interpolation method is appropriate. The slope in Fig. 11d, is slightly closer to 1 than that in Fig. 11b. The coefficient of determination R^2 is slightly smaller, however it can be considered almost the same. Table 3 summarizes validation for Period 1 and Period 2. The number of spikes in the original data is drastically reduced by data applied Criteria I and II ("in" in Table 3). In Period 1, 790 data points were excluded, and more than 88% of these data (700 data points) were recovered as interpolated values; in Period 2, more than 84% (740 data points) were recovered as interpolated values. While errors also appear in the interpolated values, the number of errors is smaller than that in the original data. In general, there is no significant difference between Period 1 and Period 2. Therefore, it is confirmed that the present method is also effective for other years. Fig. 11

The slope shown in Figs. 11b and 11d is not near 1, but about 0.85. Table 3
Comparing Fig. 11b with Fig. 11d, the improvement in the slope due to the addition of interpolation data (Fig. 11c) is negligible. The slope (for both "in" and "in+int") obtained only for daytime data ($Rn - G \geq 50$ W m^{-2}) is about 0.87 (figure not shown). On the other hand, as shown in Fig.

9, more than 97% of the interpolated data (data applied Criteria I and II) appear under the condition of $Rn - G < 50$ W m^{-2}. Therefore, the slope is mainly determined by the daytime data, and the deviation of the slope from 1 is not due to the application of Criteria I and II. This slope indicates that the sensible heat flux measured by the ECM tends to be lower than that of the Bowen ratio method. One of the possible factors is the use of data from all wind directions, including data from other than the ideal footprint directions (NNW−NE). If we extract the data from NNW−NE and draw a figure similar to Fig. 11b, the slope of the regression line is as follows (figures not shown.):

Period : all wind directions → NNW−NE

Period 1: 0.809 → 0.878

Period 2: 0.849 → 0.873

We also checked the dependence of the surface energy imbalance ratio $(H(eddy) + lE(eddy))/(Rn - G)$ (Sun et al. 2021) on the wind direction.

Period : all wind directions → NNW−NE

Period 1: 0.785 → 0.848

Period 2: 0.796 → 0.824

The slope of the sensible heat flux improves, approaching 1 in both periods in better footprint conditions. The surface energy imbalance ratio also improves under the ideal footprint condition. The above two results are consistent.

The data exclusion and interpolation methods used in this study focused mostly on periods with small temporal variations in nighttime fluxes. We consequently observed no significant differences in *mean* values between the maximum 6-point (3 h) and 12-point (6 h) interpolations (see Table 2a).

6. Discussion

6.1 Comparison of Exclusion Criteria

The analytical method of determination for the exclusion criteria with sensor accuracy is a different approach from the present method. Using Eq. (6),

$$|\delta B| \leq |B| \left(\left| \frac{\delta \Delta T}{\Delta T} \right| + \left| \frac{\delta \Delta q}{\Delta q} \right| \right), \tag{18}$$

Expressing the Bowen ratio error $|\delta B|$ as ε, we substitute Eq. (3) into Eq. (18),

$$\varepsilon \leq \frac{\frac{C_p}{l} \delta \Delta T + |B| \, \delta \Delta q}{|\Delta q|}. \tag{19}$$

According to Eq. (19), the Bowen ratio error ε, depends on thermometer and hygrometer accuracy, specific humidity difference, and the Bowen ratio itself. When $B = -1$ in Eq. (19), the error range of B is as follows (Perez et al. 1999):

$$-1 - \varepsilon < B < -1 + \varepsilon. \tag{20}$$

$$\varepsilon \leq \frac{\frac{C_p}{l}\delta\Delta T + \delta\Delta q}{|\Delta q|} \equiv \frac{\delta E_t}{|\Delta q|} \tag{21}$$

where,

$$\delta E_t = \frac{C_p}{l}\delta\Delta T + \delta\Delta q. \tag{22}$$

Applying these expressions of the Bowen ratio exclusion range, we excluded spike data using $\delta\Delta T = \sigma(\Delta T)$ and $\delta\Delta q = \sigma(\Delta q)$ obtained from the instrumental calibration in this study. $\sigma(\Delta T)$ is the standard deviation of the corrected temperature difference when the thermometers are installed at the same height. $\sigma(\Delta q)$ is the standard deviation for the specific humidity. (see Appendix B).

Fig. 12 shows the relationship between Δq and B. The hyperbolic region ($-1 - \varepsilon$ and $-1 + \varepsilon$) except for the $(\Delta q, B) = (0, -1)$ is the data extraction region in Perez et al. (1999), and the linear region (B_+ and B_-) except for the $(\Delta q, B) = (0, -1)$ is the data extraction region in this study. In Perez et al. (1999), the value of the extraction region changes with Δq. In the range $\Delta q < 0$ (evaporation conditions), there is a large amount of data in the region bounded by $-1-\varepsilon$ and B_+ in the second quadrant and the region bounded by $-1+\varepsilon$ and B_- in the third quadrant. Data in these regions are used in the Perez et al. (1999) criterion, however the data are excluded under the Criterion II. On the other hand, for $\Delta q > 0$ (condensation conditions), there is no difference in the amount of data between the two criteria. Fig. 12

Table 4 shows the results of Perez et al. (1999) and the present analysis. The analyzed data satisfy $Rn - G < 50$ W m^{-2}. Note that n in Perez et al. (1999) is the total number of Δ_i when $H(Bowen)$ is used with Perez's criterion instead of Criterion II. In the "in" rows, the number of valid data n in the present analysis is 265 fewer than in Perez et al. (1999). However, the spike ratio n_1/n in the present analysis is less than half, and the *DIFF* is smaller, compared with Perez et al. (1999). In other words, the Perez et al. (1999) method is not sufficient to eliminate spike data. When the interpolation method [4] is applied, the difference in the number of valid data n in the "in+int" row is 48, which is drastically smaller compared to 265 in the "in" row. Furthermore, the spike ratio n_1/n and the *DIFF* in the present analysis are smaller than those in Perez et al. (1999). Table 4

In Section 4.2, the exclusion conditions of the Bowen ratio were determined by comparing the fluxes obtained from the ECM:

$$-2 < B < -0.6. \tag{23}$$

Now consider the following case for Eq. (21):

$$\varepsilon \leq \delta E_t / |\Delta q| \sim 0.3 \tag{24}$$

From Eq. (20):

$$-1.3 < B < -0.7. \tag{25}$$

This error range for B is the same as that of Unland et al. (1996). Similarly:

$$\varepsilon \leq \delta E_t / |\Delta q| \sim 0.4 \Rightarrow -1.4 < B < -0.6. \tag{26}$$

$$\varepsilon \leq \delta E_t / |\Delta q| \sim 1 \Rightarrow -2 < B < 0. \tag{27}$$

The upper value in Eq. (26) and the lower value in Eq. (27) are the same as those in Eq. (23), respectively. The difference in the values of $\delta E_t / |\Delta q|$ may be due to the fact that the upper and lower limits were determined separately using the ECM. Halliwell and Rouse (1989) proposed $-1.4 \leq B \leq -0.7$, based on a comparison with aerodynamic methods from profile observations and the use of existing shear functions. The corresponding value of $\delta E_t / |\Delta q|$ are ~ 0.4 for the lower limit and ~ 0.3 for the upper limit. Therefore, the difference in $\delta E_t / |\Delta q|$ corresponding to the upper and lower limits may arise from the reason that (B_-, B_+) was determined by comparing the Bowen ratio method with another independent method.

6.2 Universality of Results

This section discusses the universality of the method to obtain exclusion criteria for Bowen ratio data near $B = -1$. In this paper, we used data under flat condition after weeding, from November to April. We also checked whether the criteria are valid for other years at the same site.

a. Temperature Dependence

The Bowen ratio (Eq. (3)) uses the difference between specific humidity (i.e. water vapor pressure) at two heights. Water vapor pressure is an exponential function of temperature, which is nonlinear. Therefore, the Bowen ratio remains temperature dependent. Fig. 13a shows frequency of data for the temperature dependence of the Bowen ratio under all wind directions without applying Criteria I or II. The data are divided into two groups, one for low temperatures (-7 °C $< T_1 \leq 9$ °C) and the other for high temperatures (9 °C $< T_1 <$ 24 °C). The number of data points for low temperatures is 4,430 and for high temperatures is 3,345. The frequency distribution between two groups is clearly different. However, since errors occur under conditions near $B = -1$, the divergent condition (Eqs. (5a) and (5b)) is fixed near $B = -1$. The frequency distribution of the number of spike occurrences (the number of data points with $|\Delta| > 40$ W m^{-2}) was obtained for the two groups as shown in Fig. 13b. The frequencies were normalized by the total number of spikes for each temperature group. The number of spikes for low and high temperatures is 108 and 117, respectively. The frequency distributions for the two temperature groups were found to be equal by the chi-square test. Therefore that (B_-, B_+) is independent of temperature. Fig. 13

b. Height Difference Dependence

In this study, the Bowen ratios were determined at heights of 0.4 m and 2.5 m. The heights of 0.4 m and 2.5 m are within the surface boundary layer after weeding. The height ratio is approximately $2.5/0.4 = 6.2$. Foken et al. (1997) recommend a height ratio of $4-8$ or higher, because the height ratio decreases, the accuracy of the temperature difference and specific humidity difference at the two heights becomes less precise, and the reliability of the Bowen ratio decreases. If the ratio of the two heights is above the threshold and within the surface boundary layer, the results could be the same even if other height differences are used to determine the Bowen ratio, since the conditions for a constant flux layer are satisfied.

c. Location Dependence

The exclusion criteria were obtained by comparison with the ECM. The ECM is highly dependent on surface homogeneity and footprint. The ECM measures at a height of 1.8 m, while the Bowen ratio method measures at heights of $z_1 = 0.4$ m and $z_2 = 2.5$ m. The footprint of the Bowen ratio method is nearly equal to that of the ECM when installed at the geometric mean of the two measured heights ($\sqrt{z_1 z_2}$) (Stannard 1997). Therefore, in the case of this study, the footprint of the Bowen ratio method is contained within that of the ECM (Appendix A). When the footprint of the ECM is under homogeneous surface conditions, the effect of the surface conditions on the measurement site is the same for the ECM and the Bowen ratio method. Therefore, it is sufficient for the measurement sites of the Bowen ratio method and ECM to be under homogeneous surface conditions. If this condition is not satisfied, it is quite possible that the fluxes from the ECM and Bowen ratio methods will differ and consequently affect the (B_-, B_+) values.

6.3 Comparison with Bulk Method

Another method of gap-filling for missing data is the bulk method (Wang et al. 2006). This section discusses comparisons of the present interpolation method with the bulk method. Since the wind of 7.5 m tower anemometer is influenced by the forest, the wind of 1.8 m ultrasonic anemometer ($U_{1.8m}$) was used for the bulk method. The temperature difference ($\Delta T_b = T_s - T$) was the difference between the ground surface temperature ($T_s = T_g(0cm)$) measured by an infrared thermometer and the air temperature (T) at 1.5

m. Two types of sensible heat fluxes, one estimated by the Bowen ratio method and the other observed by the ECM, were used to obtain the bulk coefficient. When the Bowen ratio method for sensible heat flux was used, the bulk coefficient was calculated.

$$C_h = \frac{H(Bowen)}{C_p \rho U_{1.8m} \Delta T_b} \tag{28}$$

Bulk coefficients were determined separately for unstable $\Delta T_b \geq 0$ and stable $\Delta T_b < 0$ conditions. The following conditions were imposed on the data used in the calculation of Eq. (28): $|\Delta T_b| \geq 1°\mathrm{C}$, $|H(Bowen)| \geq 10$ W m^{-2}, wind directions from NNW−NE, and both Criterion I and II. The surface fluxes were calculated using the following equation.

$$H(bulk) = C_p \rho C_h U_{1.8m} \Delta T_b \tag{29a}$$

$$lE(bulk) = Rn - G - H(bulk) \tag{29b}$$

Figure 14 shows the dependence of the bulk coefficient on wind speed. Figure 15 represents comparison of the fluxes obtained by the bulk method with those by the ECM for the missing data. Table 5 summarized verification of gap-filling. Note that “bulk(B)” in Table 5 is the result used in Eq. (28) and “bulk(e)” is the result when $H(eddy)$ is used instead of $H(Bowen)$ in Eq. (28). The number of interpolations in the present method is 12-point or less (see Section 4.3), thus the number of recovered data n is slightly higher for the bulk method. The number of data points and $bias$ are better for the bulk method. The $DIFF$ are about the same. However, for the present method, the number of spikes is fewer and the R^2 is slightly higher (Table 5a). A comparison of gap-filled data with the ECM data is shown in Table 5b. There is no significant difference in the coefficients of the regression equation, $bias$, $DIFF$, and R^2 for three method ("bulk(B)", "bulk(e)", and "present"). The bulk method and the Bowen ratio method use different surface temperatures: as shown in Table 1, the bulk method uses a radiation thermometer, while the Bowen ratio method uses an infrared radiometer. The radiation thermometer measures the temperature in a narrow field of view (0.14 m diameter at a 2 m distance), while the infrared radiometer measures infrared radiation from a hemisphere (2π steradian). The wavelength range of the infrared radiation is also different between the two sensors ($8-16$ μm for radiation thermometer and $4-50$ μm for infrared radiometer, respectively). Because of these differences, it is difficult to judge which gap-filling method is better based on the degree of difference between Figs. 15a, b, and c. While the bulk method requires prior determination of bulk coefficients and additional measurements of surface radiative temperature and wind speed during its application, the present method employs the B-factor and requires fewer input variables, making it a simpler and more practical alternative in field applications.

Fig. 14

Fig. 15

Table 5

7. Summary

In this analysis, we presented two exclusion criteria for the Bowen ratio method. Criterion I concerned the sign consistency of observed values according to K-theory, whereas Criterion II focused on the exclusion of singular points in the Bowen ratio method. In the latter case, we discussed the determination of exclusion range for the Bowen ratio. The number of valid data points, $bias$, *DIFF*, and the number of spike errors were calculated using the sensible heat flux estimated by the ECM and the Bowen ratio method. The Bowen ratio range of spike occurrence was revealed from the frequency distribution of the number of spikes per Bowen ratio class. B_+ was found to be about -0.6 at the lowest *DIFF*, while B_- was obtained with as many valid data points as possible, with bias as close to zero as possible, and with reference to the frequency distribution of the number of spikes. We therefore adopted the optimal combination of $(B_-, B_+) = (-2.0, -0.6)$. It is possible that the exclusion criteria (B_-, B_+) may change due to site dependency. However, there is no dependence on temperature or height difference.

The data exclusion criteria proposed in this study are practical and applicable for data processing using the Bowen ratio method. We believe that they are also applicable to the modified Bowen ratio method (Liu and Foken 2001). The B-factor interpolation method provided reasonable results for gap-filling of excluded data points, and interpolation of 6−12 data points was feasible.

It is valuable to estimate surface fluxes using the method described in this study and to accumulate long-term datasets. This methodology is applicable to various observational fields under homogeneous surface conditions in climate/environmental studies and/or to the verification of numerical models of the atmospheric boundary layer.

Data Availability Statement

The observational data from the field in the Meteorological Research Institute are available from the corresponding author upon reasonable request.

Acknowledgments

We would like to thank Dr. Hideaki Mouri for his constant encouragement and useful comments, as well as our colleagues for their helpful discussions. We also sincerely thank the two anonymous reviewers for their careful reading of our manuscript and for providing numerous insightful

comments and suggestions that have significantly improved the quality of this paper.

References

Allen, R. G., L. S. Pereira, T. A. Howell, and M. E. Jensen, 2011: Evapotranspiration information reporting: I. factors governing measurement accuracy. *Agric. Water Manage.*, **98**, 899–920.

Arya, S., 2001: *Introduction to Micrometeorology (2nd ed.).* Academic Press, 415 pp.

Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans, J. Fuentes, A. Goldstein, G. Katul, B. Law, X. Lee, Y. Malhi, T. Meyers, W. Munger, W. Oechel, K. T. P. U, K. Pilegaard, H. P. Schmid, R. Valentini, S. Verma, T. Vesala, K. Wilson, and S. Wofsy, 2001: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.*, **82**, 2415–2435.

Bevington, P. R., and D. K. Robinson, 2003: *Data Reduction and Error Analysis (3rd ed.).* MacGraw-Hill, pp 320.

Bowen, I. S., 1926: The ratio of heat losses by conduction and by evaporation from any water surface. *Phys. Rev.*, **27**, 779–787.

Campbell, S., 2013: *Instruction manual (EC150 CO2 and H2O Open-Path Gas Analyzer)*. Campbell Scientific, Inc., 38 pp.

Campbell, S., 2015: *Instruction manual (CSAT3 Three Dimensional Sonic Anemometer)*. Campbell Scientific, Inc., 44 pp.

Cook, D., and R. Sullivan, 2020: *Eddy Correlation Flux Measurement System (ECOR) Instrument Handbook*. U.S. Department of Energy, 17 pp.

Foken, T., 2017: *Micrometeorology*. Springer, 362 pp.

Foken, T., S. H. Richter, and H. Muller, 1997: Zur genauigkeit der Bowen-Ratio-Methode. *Wetter und Leben*, **49**, 57–77.

Garratt, J. R., 1994: *The Atmospheric Boundary Layer*. Cambridge University Press, 336 pp.

Halliwell, D. H., and W. R. Rouse, 1989: A comparison of sensible and latent heat flux calculations using the Bowen ratio and aerodynamic methods. *J. Atmos. Oceanic Technol.*, **6**, 563–574.

Hayashi, M., T. Hirota, Y. Iwata, and I. Takayabu, 2005: Snowmelt energy balance and its relation to Foehn events in Tokachi. *J. Meteor. Soc. Japan*, **83**, 783–798.

Hofsteenge, M. G., N. J. Cullen, M. K. C. H. Reijmer, M. van den Broeke, and J. F. Orwin, 2022: The surface energy balance during Foehn events at Joyce Glacier, McMurdo Dry Valleys, Antarctica. *Cryosphere*, **16**, 5041–5059.

Ichinose, N., and Y. Pan, 2023: Numerical evaluation of outdoor thermal comfort change by a re-development of a large urban green park. In *Proceedings of Annual Meeting of Environmental Systems Research*, Volume 51, Japan Society of Civil Engineers, 107–110.

Kaimal, J. C., and J. J. Finnigan, 1994: *Atmospheric boundary layer flows -Their structure and measurement.* Oxford University Press, 289 pp.

Liu, H., and T. Foken, 2001: A modified Bowen ratio method to determine sensible and latent heat fluxes. *Meteor. Z.*, **10**, 71–80.

Mouri, H., T. Morinaga, and S. Haginoya, 2019: Unlikely existence of k_x^{-1} spectral law in wall turbulence: An observation of the atmospheric surface layer. *Phys. Fluids*, **31**, 035103.

Ohmura, A., 1982: Objective criteria for rejecting data for Bowen ratio flux calculations. *J. Appl. Meteor.*, **21**, 595–598.

Ortega-Farias, S. O., R. H. Cuenca, and M. Ek, 1996: Daytime variation of sensible heat flux estimated by the bulk aerodynamic method over a grass canopy. *Agric. For. Meteorol.*, **81**, 131–143.

Payero, J. O., C. M. U. Neale, J. L. Wright, and R. G. Allen, 2003: Guidelines for validating Bowen ratio data. *Trans. ASAE*, **46**, 1051–1060.

Perez, P. J., F. Castellvi, M. Ibanez, and J. I. Rosell, 1999: Assessment of reliability of Bowen ratio method for partitioning fluxes. *Agric. For. Meteorol.*, **97**, 141–150.

Stannard, D. I., 1997: A theoretically based determination of Bowen-ratio fetch requirements. *Boundary-Layer Meteorology*, **83**, 375–406.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Springer, 666 pp.

Sun, J., W. J. Massman, R. M. Banta, and S. P. Burns, 2021: Revisiting the surface energy imbalance. *J. Geophys. Res.: Atmos.*, **126**, https://doi.org/10.1029/2020JD034219.

Tanner, B. D., J. P. Greene, and G. E. Bingham, 1987: A Bowen ratio design for long term measurements. In *ASAE 1987 International Winter Meeting. St. Joseph, Mich.*, American Society of Agricultural Engineers, Paper No. 87–2563.

Unland, H. E., P. R. Houser, W. J. Shuttleworth, and Z. L. Yang, 1996: Surface flux measurement and modelling at a semi-arid Sonoran Desert site. *Agric. For. Meteorol.*, **82**, 119–153.

Wang, K., Z. Li, and M. Cribb, 2006: Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: A new method to determine the Priestley-Taylor parameter. *Remote Sens. Environ.*, **102**, 293–305.

List of Figures

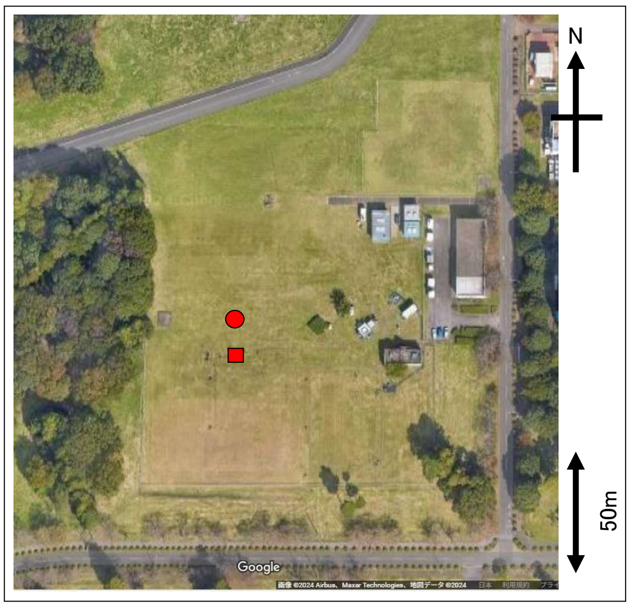


Fig. 1. Observation site from Google Maps. The circle and square indicate the eddy covariance system and the Bowen ratio system, respectively.

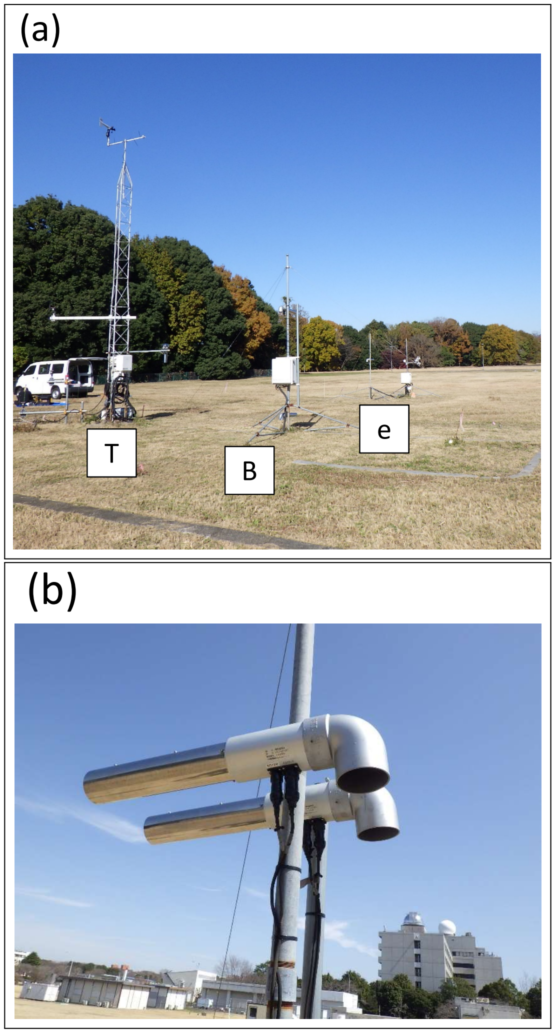


Fig. 2. (a) Observation site: Bowen ratio system (B), eddy covariance system (e), and tower (T). “e” is 15 m north of “B”, and “T” is 5 m west of “B”. (b) Comparative observations at the same height for calibration of the Bowen ratio system. The air intake faces north.

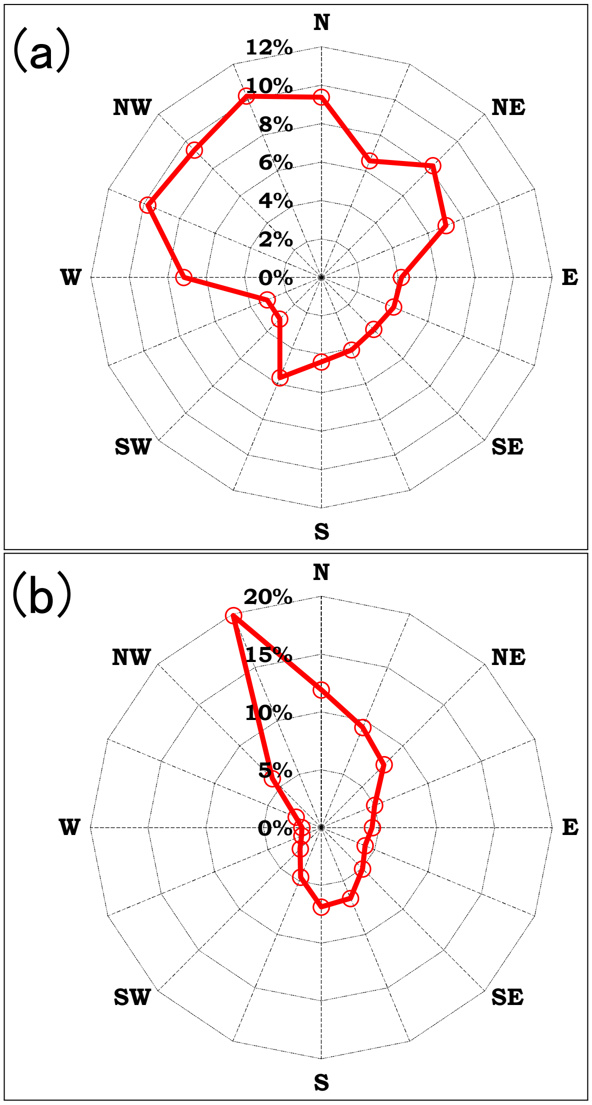


Fig. 3. Wind rose for Period 1. (a) Tower-mounted anemometer at 7.5 m height. (b) Eddy covariance system at 1.8 m height.

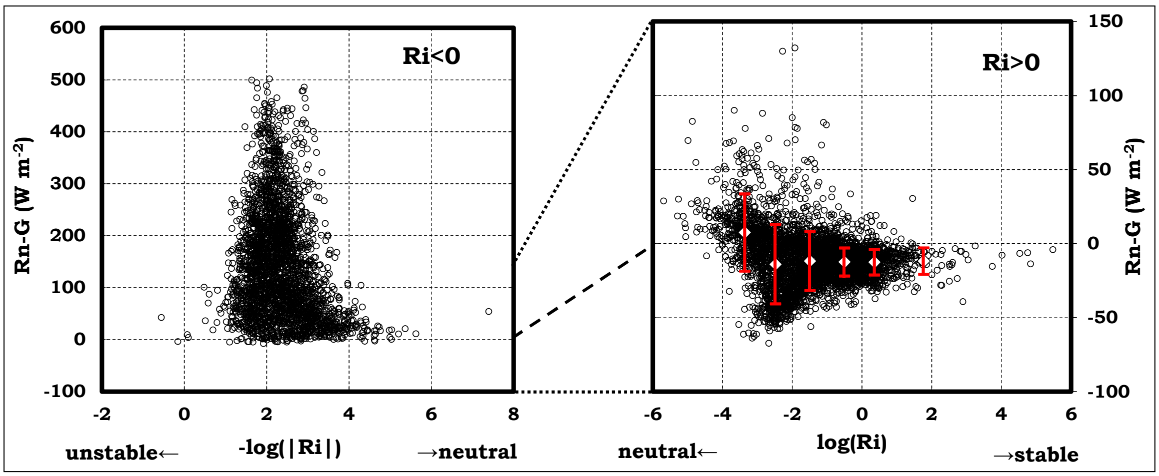


Fig. 4. Relationship between $Rn - G$ and atmospheric stability Ri in the Period 1 dataset. The left and right panels correspond to $Ri < 0$ and $Ri > 0$, respectively, with values plotted as logarithms of the absolute values of Ri. The diamonds and error bars indicate the average and standard deviation, respectively, for each Ri interval.

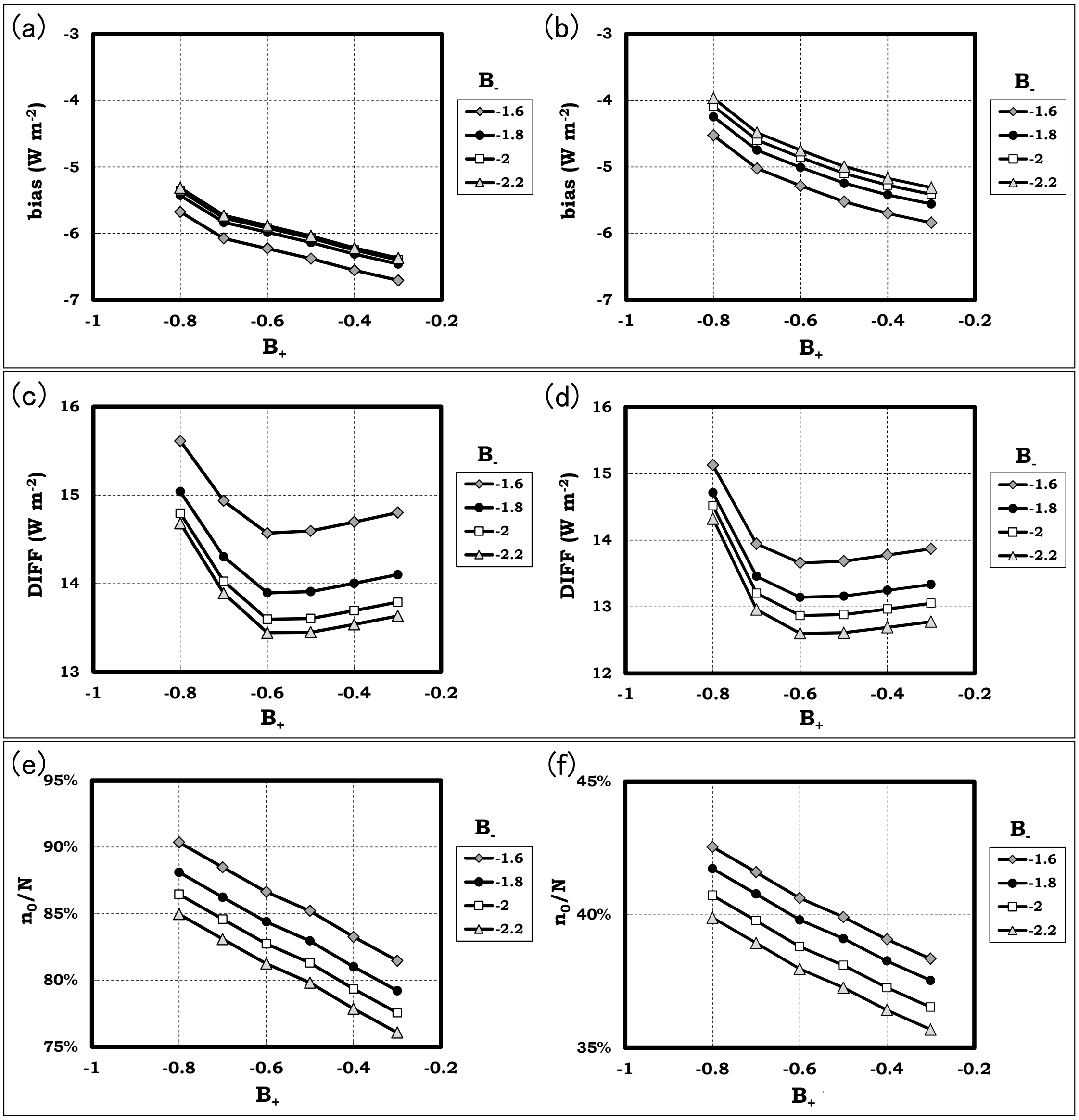


Fig. 5. (B_{-}, B_{+}) dependence of *bias*, $DIFF$, and the valid data ratio n_0/N under all wind directions (ALL) or NNW−NE wind directions (NNW−NE). (a) *bias*, ALL. (b) *bias*, NNW−NE. (c) $DIFF$, ALL. (d) $DIFF$, NNW−NE. (e) n_0/N, ALL. (f) n_0/N, NNW−NE. Data are for $Rn - G < 50$ W m^{-2} and $N = 3690$, which is the total number of data points for all wind directions, excluding those filtered by Criteria I and II.

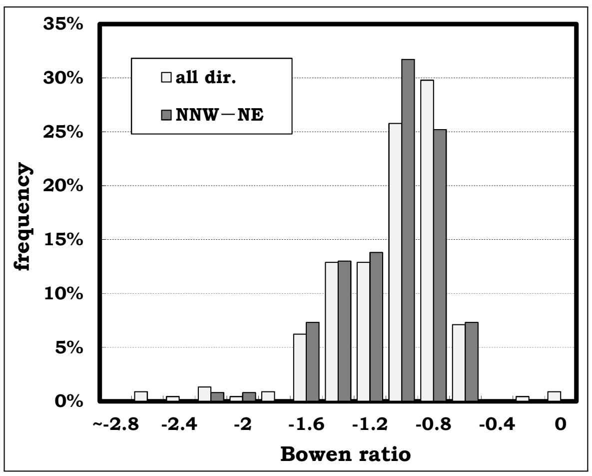


Fig. 6. Frequency distribution of the number of spikes n_1 by Bowen ratio class. Frequencies are expressed as a percentage of the total number of spikes for $-3.0 \leq B \leq 0.0$. The total number of spikes is 225 for all wind directions and 123 for NNW−NE. Data are for $Rn - G < 50$ W m^{-2}.

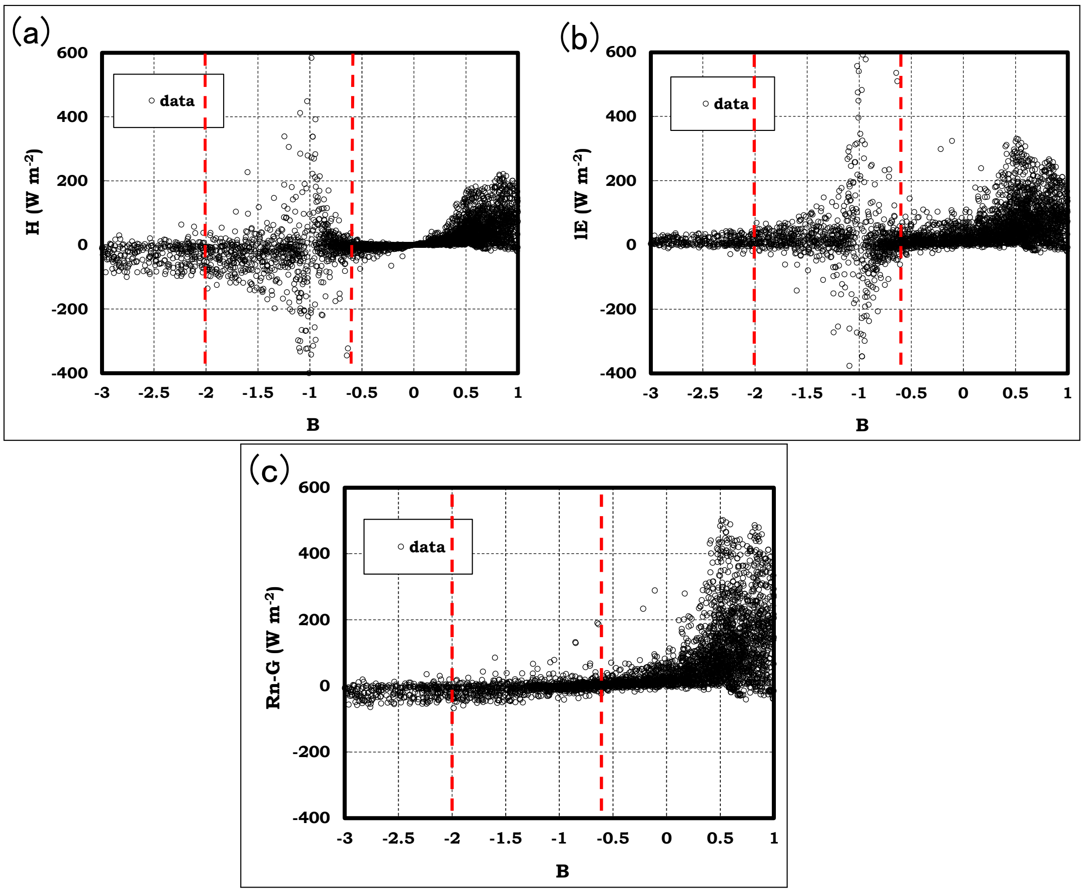


Fig. 7. Relationship between the Bowen ratio and surface fluxes. (a) sensible heat flux (H), (b) latent heat flux (lE), and (c) available energy ($Rn - G$) in the Period 1 dataset. The interval between the red dashed lines represents the exclusion range $(B_{-}, B_{+}) = (-2.0, -0.6)$. Data are under all wind directions.

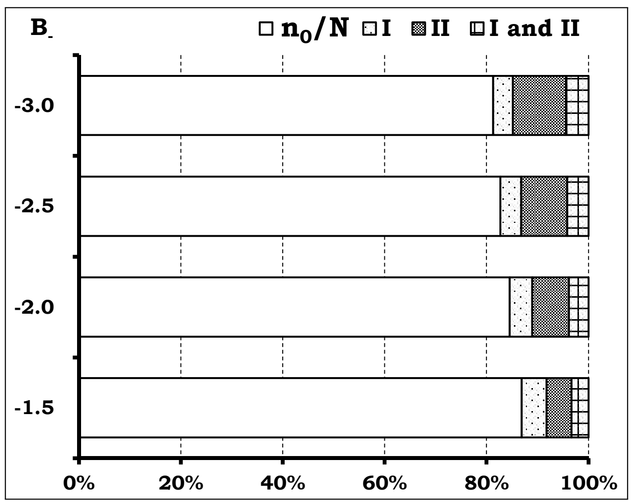


Fig. 8. Compositional classification of the data after applying exclusion Criteria I and/or II when $B_+ = -0.6$. All data are not restricted by $Rn - G$. The total number of data points is N=6690, which is set to 100%.

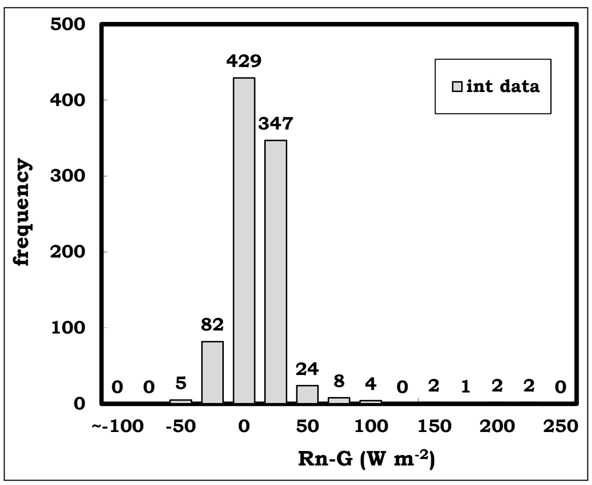


Fig. 9. Frequency distribution of available energy $Rn - G$ by class at the time of interpolation.

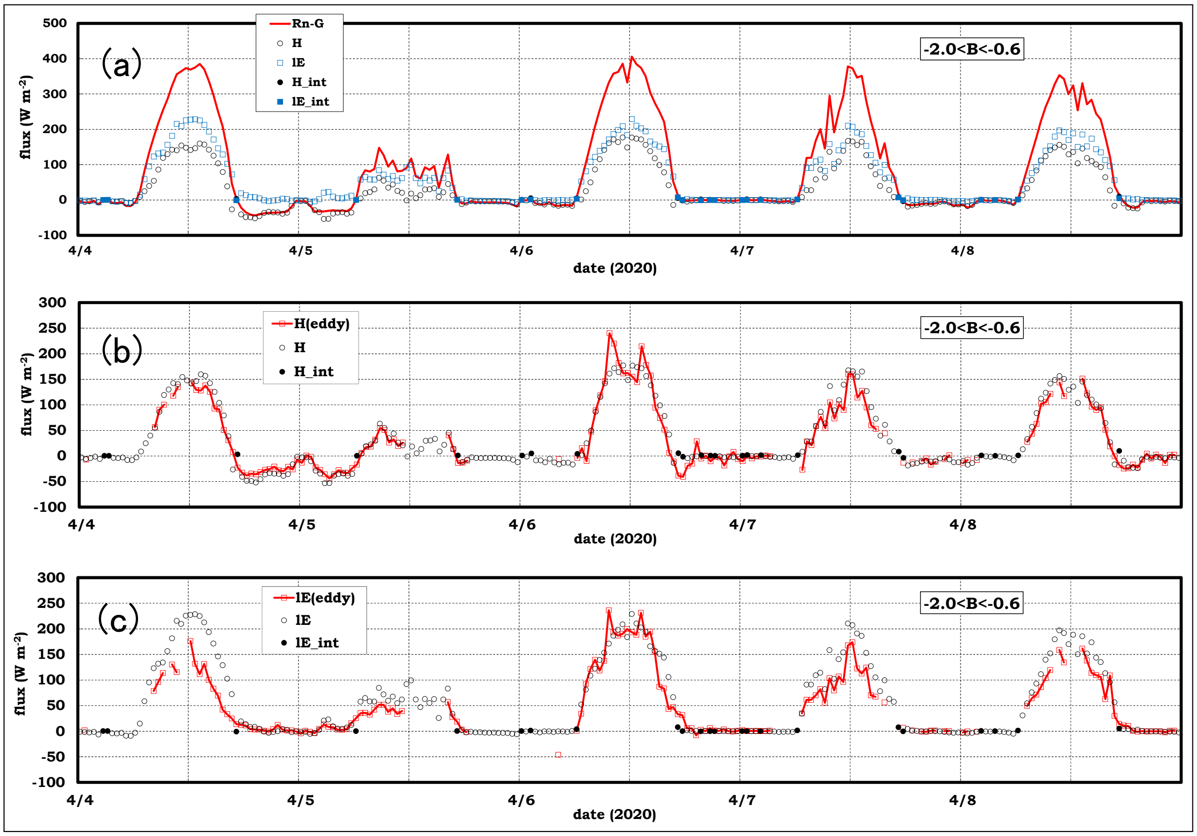


Fig. 10. Example of flux time-series in the Period 2 dataset. Extracted data (white markers) and 12-point interpolation results (filled markers). (a) Available energy ($Rn - G$), sensible heat flux (H), and latent heat flux (lE). The abbreviation "int" stands for interpolation. (b) Comparison between the Bowen ratio method and ECM for sensible heat flux. (c) Same as (b), but for latent heat flux.

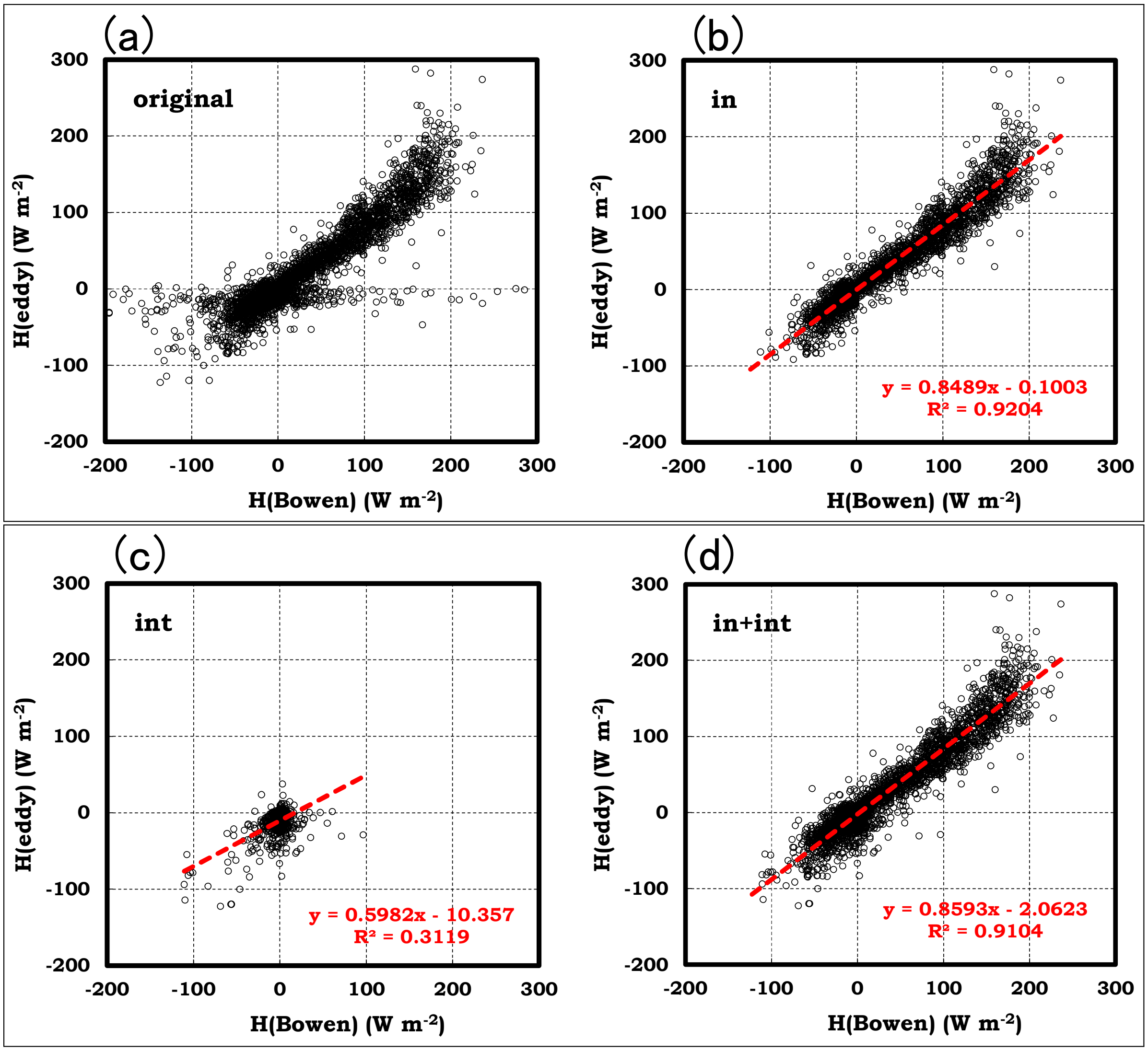


Fig. 11. Summary of procedure. (a) Original data. (b) Data with applied Criteria I and II (“in” data). (c) Interpolated data (“int” data). (d) Sum of “in” and “int” data. Red dashed line and equation are the regression line and function, respectively. R^2 is coefficient of determination. Period 2 dataset.

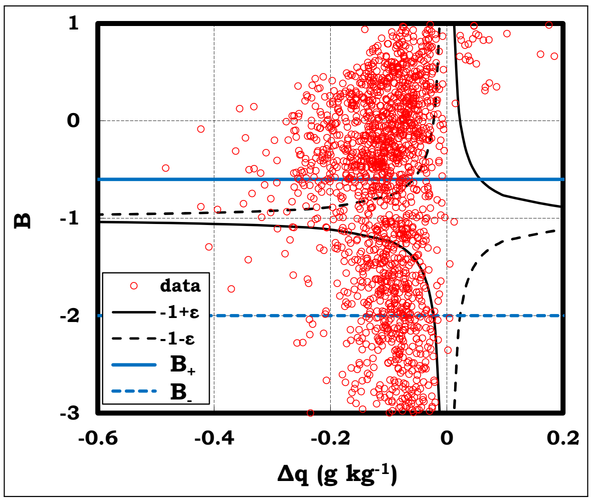


Fig. 12. Relationship between specific humidity difference (Δq) and the Bowen ratio (B). The hyperbola represents $B = -1 \pm \varepsilon$ in Eq. (20), while the straight lines represent $B_+ = -0.6$ and $B_- = -2.0$. Circles indicate the observed data from this study: Period 1, Criterion I applied, $Rn - G < 50$ W m^{-2}.

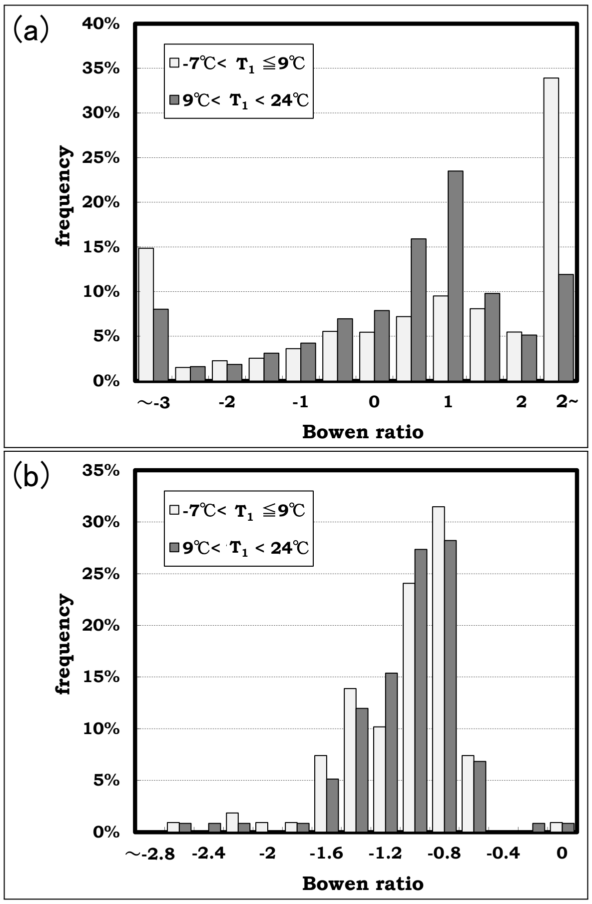
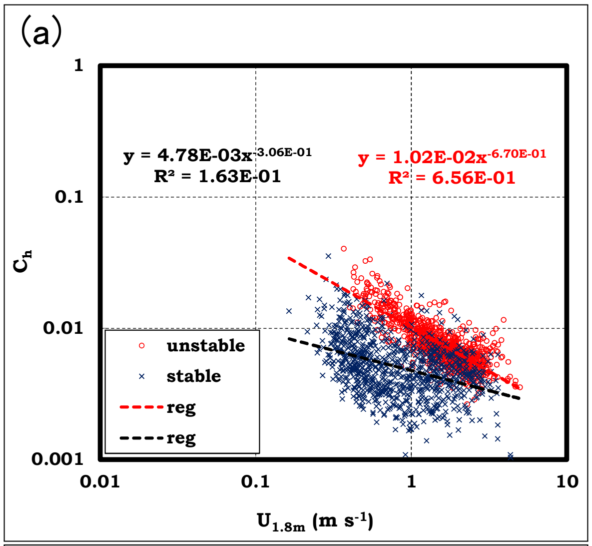


Fig. 13. (a) Temperature dependence of the Bowen ratio. Data under all wind directions without applying Criteria I and II were used. The number of data points for low and high temperatures is 4430 and 3345, respectively. (b) Frequency of data for temperature dependence of spikes under all wind directions. The number of spikes for low and high temperatures is 108 and 117, respectively.



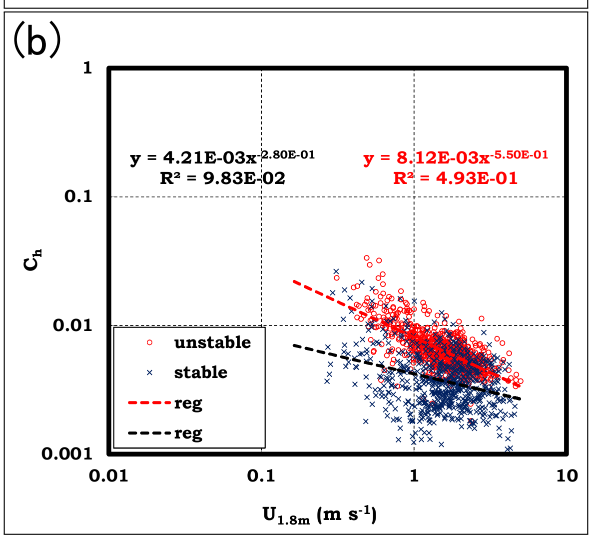


Fig. 14. Dependence of the bulk coefficient (C_h) on wind speed. (a) C_h is calculated using Eq. (28). (b) Same as (a), but C_h is calculated using Eq. (28) but $H(eddy)$ is used instead of $H(Bowen)$. Dashed line and equation are the regression line and function, respectively. R^2 is coefficient of determination.

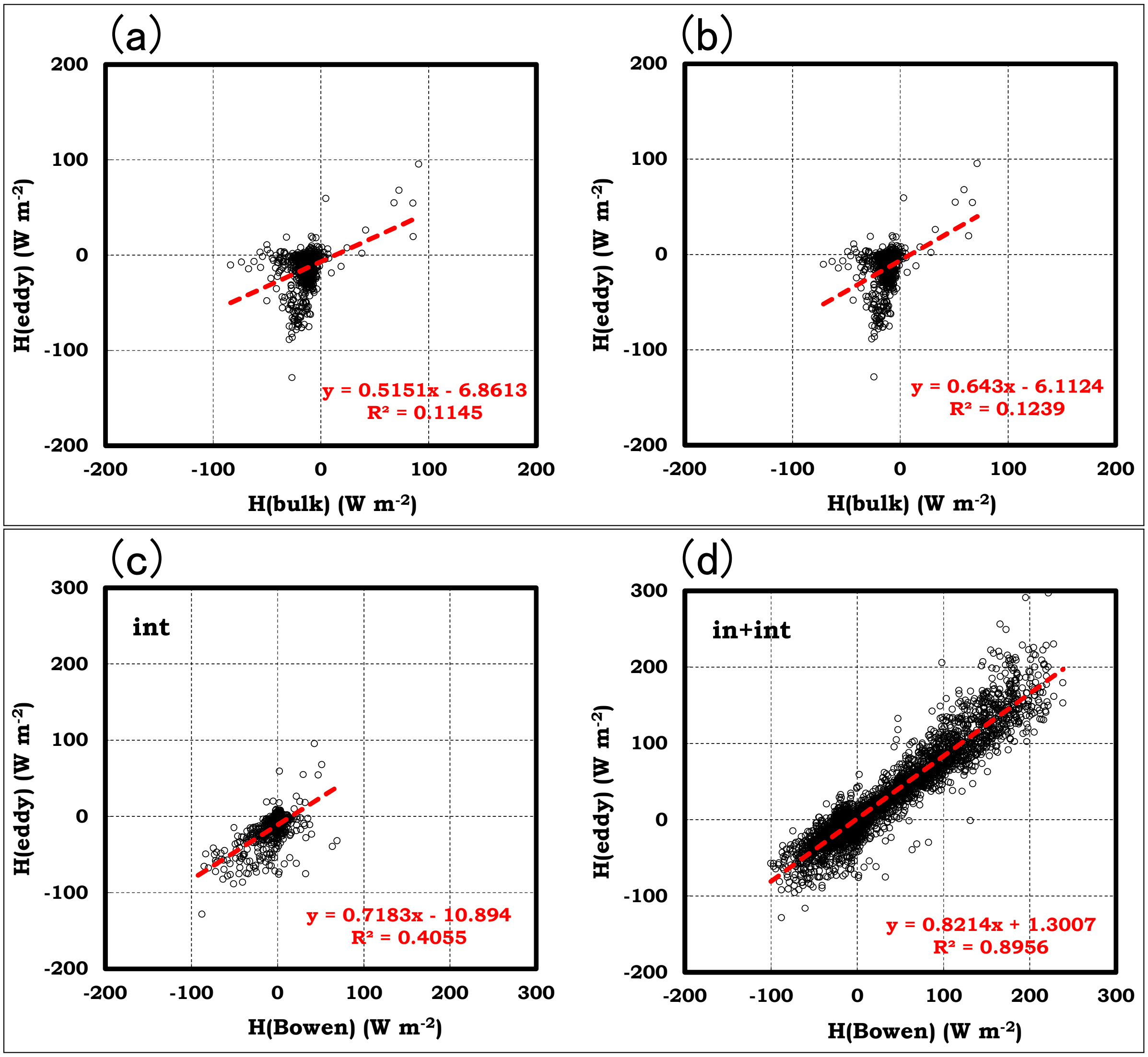


Fig. 15. Comparison of gap-filling data with observed values from ECM. (a) $H(bulk)$ in (a) is calculated using the bulk coefficient obtained from Eq. (28). (b) $H(bulk)$ in (b) is calculated using the bulk coefficient obtained by substituting $H(eddy)$ into Eq. (28). (c) Interpolated data in this study ("int" data). (d) Sum of "in", which is data applied Criteria I and II, and "int" data. Both (c) and (d) correspond to Fig. 11c and Fig. 11d respectively, except for Period 1.

List of Tables

Table 1. List of observational instruments.

<table>
<tr><th>Meteorological element</th><th>Instrument Type and Model number</th><th>Height/Depth</th><th>sampling and recording</th><th>Symbol</th><th colspan="2">flux</th></tr>
<tr><td>Difference of temperature and humidity</td><td>VAISALA Humidity & temperature probe HMP155</td><td>$z_1 = 0.4$ m
$z_2 = 2.5$ m</td><td rowspan="8">0.2 Hz sampling 10 minutes recording</td><td>ΔT and Δq</td><td></td><td rowspan="6">$H(Bowen)$ and $lE(Bowen)$</td></tr>
<tr><td>Short wave radiation</td><td>HukuseFlux pyranometer HF-SR20</td><td>1.5 m</td><td>$S\downarrow$ and $S\uparrow$</td><td rowspan="2">Rn</td></tr>
<tr><td>Infrared radiation</td><td>HukuseFlux pyradiometer HF-IR20</td><td>1.5 m</td><td>$L\downarrow$ and $L\uparrow$</td></tr>
<tr><td>Soil heat flux</td><td>HukseFlux soil heat plate HF-HFP01</td><td>2 and 10 cm (depth)</td><td>G_z</td><td rowspan="3">G</td></tr>
<tr><td>Soil temperature</td><td>Thermocouple</td><td>1, 2, … , 10 cm (depth)</td><td rowspan="2">T_g</td></tr>
<tr><td>Surface temperature</td><td>Radiation thermometer TASCO THI-303N</td><td>0 cm</td></tr>
<tr><td>Air temperature and humidity</td><td>VAISALA Humidity & temperature probe HMP155</td><td>1.5 m</td><td>T</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Wind speed and direction</td><td>Young Wind Monitor CYG-5103</td><td>7.5 m</td><td>U</td></tr>
<tr><td>Wind and temperature fluctuation</td><td>Campbell CSAT3 3-Dimensional Sonic Anemometer</td><td>1.8 m</td><td rowspan="2">10 Hz sampling 30 minutes recording</td><td rowspan="2"></td><td colspan="2" rowspan="2">$H(eddy)$ and $lE(eddy)$</td></tr>
<tr><td>Water vapor fluctuation</td><td>Campbell EC150 CO_2 and H_2O Open-Path Gas Analyzer</td><td>1.8 m</td></tr>
</table>

Table 2. Number of interpolations and mean sensible heat fluxes. (a) n_0, *mean*, standard deviation (SD), and the difference between the Bowen ratio method and the ECM ($DIFF$) by the number of interpolations (int. No.) for the period $Rn - G < 50$ W m^{-2}. The $DIFF$ is calculated using interpolation. (b) Tests of *means* for different numbers of interpolations. Null hypothesis: The *means* of the two samples are equal. Alternative hypothesis: The *means* of the two samples are different. The values in the table represent the P-value (Type I error rate) for the null hypothesis.

(a)

	H (W m^{-2})			
int. No.	n_0	*mean*	*SD*	*DIFF*
0	3883	-13.43	16.93	13.43
≦6	5008	-11.57	16.80	14.16
≦8	5147	-11.28	16.74	14.16
≦10	5222	-11.21	16.75	14.33
≦12	5291	-11.05	16.70	14.35

(b)

	P-value			
int. No.	≦6	≦8	≦10	≦12
≦6	1.000	0.381	0.278	0.120
≦8		1.000	0.835	0.498
≦10			1.000	0.638
≦12				1.000

Table 3. Summary of validation: number of data (n), bias ($bias$), difference between the Bowen ratio method and the ECM ($DIFF$), number of spikes (n_1), coefficient of determination (R^2), and corresponding figure symbols. Data in column * are not restricted to $Rn - G$; all other columns are $Rn - G < 50$ W m^{-2}. In method column, "in", "int", and "in+int" represent data applied Criteria I and II, interpolated data, and sum of "in" and "int", respectively. [4] indicates the interpolation method [4].

method	n	$bias$	$DIFF$	n_1	n_1/n	R^2	R^2 *	Fig. 15 *	Period
original	3690	---	---	252	6.83%	---	---	---	1
in	2900	−6.8	13.4	31	1.07%	0.582	0.908	---	
int [4]	700	9.2	17.7	30	4.29%	0.420	0.406	(c)	
in+int [4]	3600	−3.7	14.3	61	1.69%	0.460	0.896	(d)	

method	n	$bias$	$DIFF$	n_1	n_1/n	R^2	R^2 *	Fig. 11 *	Period
original	2853	---	---	222	7.78%	---	---	(a)	2
in	1977	−2.5	11.2	13	0.66%	0.644	0.920	(b)	
int [4]	740	8.8	16.9	32	4.32%	0.374	0.312	(c)	
in+int [4]	2717	0.6	13.0	45	1.66%	0.518	0.910	(d)	

Table 4. Results of Perez et al. (1999) and the present analysis: number of valid data (n), bias ($bias$), difference between the Bowen ratio method and the ECM ($DIFF$), number of spikes (n_1) and coefficient of determination (R^2). In row, "in" represents data applied Criteria I and II, and "in+int" represents sum of "in" and "int", which is interpolated data. The difference in n is 265 for the "in" row and 48 for the "in+int" row. Period 1 data and $Rn - G < 50$ W m^{-2}.

	method	n	$bias$	$DIFF$	n_1	n_1/n	R^2
in	Perez et al.	3165	−7.8	15.3	72	2.27%	0.584
	Present	2900	−6.8	13.4	31	1.07%	0.582
in+int	Perez et al.	3648	−5.9	15.7	95	2.60%	0.506
	Present	3600	−3.7	14.3	61	1.69%	0.460

Table 5. (a) Summary of the gap-filling method: number of data (n), bias ($bias$), difference between the Bowen ratio method and the ECM ($DIFF$), number of spikes (n_1), coefficient of determination (R^2), and corresponding figure numbers. (b) Summary of the coefficients of the regression line obtained using all data after gap-filling: slope (a) and intercept (b). "bulk(B)" is the result used in Eq. (28). "bulk(e)" is the result when $H(eddy)$ is used instead of $H(Bowen)$ in Eq. (28). Period 1.

(a)

method	n	$bias$	$DIFF$	n_1	n_1/n	R^2	Fig. 15
bulk(B)	790	0.3	18.2	50	6.33%	0.115	(a)
bulk(e)	790	−1.5	18.2	48	6.08%	0.124	(b)
present	700	9.2	17.7	30	4.29%	0.406	(c)

(b)

method	n	$bias$	$DIFF$	a	b	R^2	Fig. 15
bulk(B)	5466	−1.1	20.0	0.816	2.36	0.895	---
bulk(e)	5466	−1.3	19.9	0.818	2.12	0.895	---
present	5376	2.4	20.0	0.821	1.30	0.896	(d)

Appendix A. Footprint

Footprints were calculated using the analytical model of Hsieh et al. (2000). This model includes atmospheric stability (L) and surface roughness z_0. L is the Monin-Obukhov length.

$$S(x) = \begin{cases} 0 & x < 0, \\ S_0 & x \geq 0. \end{cases} \tag{A1}$$

$$F(x, z_m) = \int_{-\infty}^{x} S(x) f(x, z_m) dx \tag{A2}$$

$f(x, z_m)$ is the footprint function, $F(x, z_m)$ is the scalar flux, $S(x)$ is the source intensity, z_m is the measured height, and x is the axis of the mean wind direction (where positive values indicate the upwind direction).

$F(x, z_m)$ is the percentage of upwind surface flux from x to 0 observed at sensor height z_m. $F(x, z_m)/S_0$ indicates the normalized flux. For example, considering the neutral state in Fig. A1a, $F(100\text{m}, 1.8\text{m})/S_0 = 0.75$ means that 75% of the total flux measured at a height of 1.8 m comes from the ground surface between 0 m and 100 m upwind. Fig. 1

References

Hsieh, C., G. Katul and T. Chi, 2000: An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows, *Adv. Water Resour*, **23**, 756-772.

Appendix B. Instrumental Calibration Methods

To determine the bias between the two sensors, the temperature and humidity sensors are placed in a ventilated enclosure at the same height and comparative observations are made. The biases for temperature and relative humidity are defined as $\Delta\widetilde{T}(\widetilde{T_1}) = \widetilde{T_1} - T_2$ and $\Delta\widetilde{rh}(\widetilde{rh_1}) = \widetilde{rh_1} - rh_2$, respectively, where $\widetilde{T_1}$ is the temperature before bias correction, and $\widetilde{rh_1}$ is the corresponding relative humidity. T_2 and rh_2 are taken as standard values.

A linear function is assumed for the instrumental difference:

$$\Delta\widetilde{T}(\widetilde{T_1}) = a \times \widetilde{T_1} + b, \tag{A3a}$$

$$\Delta\widetilde{rh}(\widetilde{rh_1}) = c \times \widetilde{rh_1} + d. \tag{A3b}$$

The regression coefficients $a, b, c,$ and d are found by the least squares method. Figure A2 shows an example of a comparison of a thermometer and a hygrometer at the same height. Fig. 2

The temperature and relative humidity after correction for instrumental differences are as follows:

$$T_1 = \widetilde{T_1} - \Delta\widetilde{T}(\widetilde{T_1}), \tag{A4a}$$

$$rh_1 = \widetilde{rh_1} - \Delta\widetilde{rh}(\widetilde{rh_1}). \tag{A4b}$$

The specific humidity q_1 is calculated from the corrected (T_1, rh_1). The variation of the difference for each sensor after instrumental correction is shown in Fig. A3. Fig. 3

Table A1 summarizes the results of the instrumental calibration, where $\Delta T = T_1 - T_2$ is the average corrected temperature difference between two thermometers installed at the same height and $\sigma(\Delta T)$ is its standard deviation. Similarly, specific humidity is expressed as $\Delta q = q_1 - q_2$ with its standard deviation $\sigma(\Delta q)$. The average over the entire period of instrumental correction is $\sigma(\Delta T) <\sim 0.033$°C and $\sigma(\Delta q) <\sim 0.014$ g kg^{-1}. This corresponds to $(C_p/l)\sigma(\Delta T) <\sim 0.014$ g kg^{-1}. Therefore, the errors in temperature and humidity differences with respect to flux are comparable. In the text, the instrument accuracy is set to $\delta\Delta T = \sigma(\Delta T)$ and $\delta\Delta q = \sigma(\Delta q)$. Table 1

Appendix C. Behavior of the Data for $B \sim -1$ and Strong Wind Conditions

The behavior of the data for $B \sim -1$ described in Sections 2.2b and 4.2 is confirmed, especially under high wind conditions. First, we extract the data with $|Rn - G| < 20$ W m^{-2} and $U > 4$ m s^{-1}. Next, we exclude the data with $|\Delta| > 40$ W m^{-2} to remove spikes. Data inside the lower-left dashed region in Fig. A4a represent the data with $B \sim -1$ and $U > 4$ m s^{-1} except spikes. Figure A4b shows the relationship between H and lE under the above conditions. The wind speed conditions in this case were

6.3 m s^{-1} > U > 4 m s^{-1}. Wind speed conditions were below 5.1 m s^{-1} when relatively large fluxes consistent with the ECM were observed.

In the field data, there are indeed data consistent with the eddy covariance method (ECM), with large absolute values of negative sensible heat flux and positive latent heat flux under strong winds with $B \sim -1$. If the wind speed conditions are stronger, the magnitude of the flux would be larger than the value of 60 W m^{-2} observed in Fig. A4b. Fig. 4

Appendix D. Interpolation

Four interpolation methods ([1]−[4]) were considered to recover as much of the excluded data as possible.

[1] Interpolation of the Bowen ratio.
[2] Interpolation of sensible heat H and latent heat lE.
[3] Interpolation of temperature differences ΔT and specific humidity differences Δq.
[4] Interpolation of "B-factor", such as $B/(1+B)$ or $1/(1+B)$. Note that $B/(1+B)$ interpolation is equivalent to $1/(1+B)$.

Figure A5 shows an example of a time-series where the four interpolation methods were applied. Large spikes occurred in [1]−[3], while no spikes were observed in [4]. Fig. 5

Table A2 summarizes the number of spikes generated by each interpolation method. The total number of interpolated data points is 700 (the same as "int [4]" in Period 1 of Table 3). n_1/n is the ratio of the number of spikes to the total number of interpolated data points, and n_2/n is the ratio of the number of data that return to the excluded range ($-2 < B < -0.6$) after interpolation to the number of interpolated data points. In [1]−[3], the percentage of data in the excluded range after interpolation exceeds 30%. In [4], the percentage is zero. The percentage of spikes is at least 13% for [1]−[3] and 4.3% for [4]. The percentage of spikes after interpolation is reduced by 1/4 or 1/3 for [4] compared to the other interpolation methods. Table 2

Figure A6 shows the frequency distribution by category for the four spike data sets shown in Table A2. In [4], the spikes are $|\Delta| < 120$ W m^{-2}, while in [1]−[3], there are many spikes of 200 W m^{-2} or more. Fig. 6

List of Figures

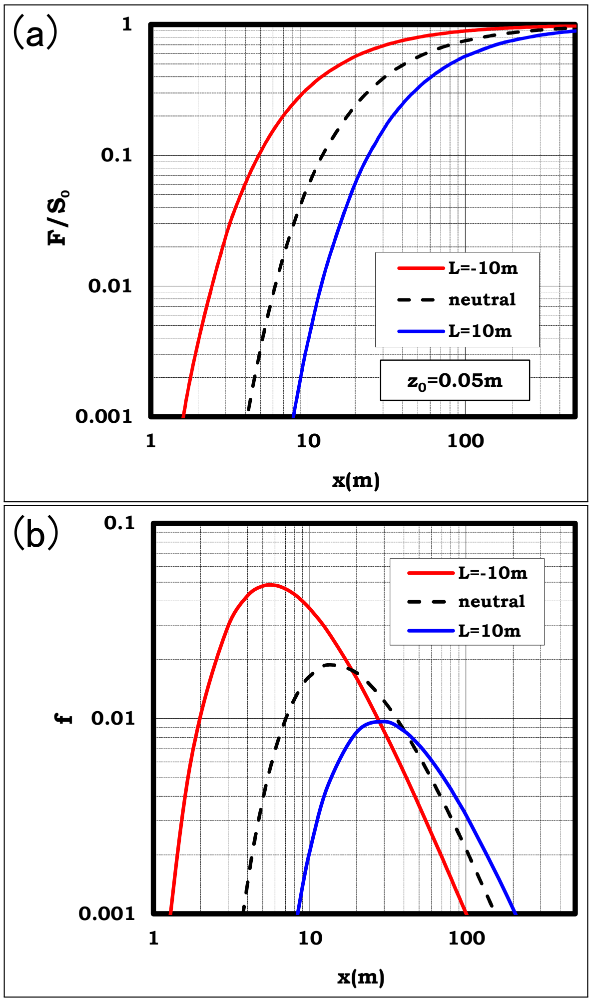


Fig. A1. Footprint analysis. (a) Normalized flux $F(x, z_m)/S_0$. (b) Footprint function $f(x, z_m)$. L is the Monin-Obukhov length, and z_0 is surface roughness.

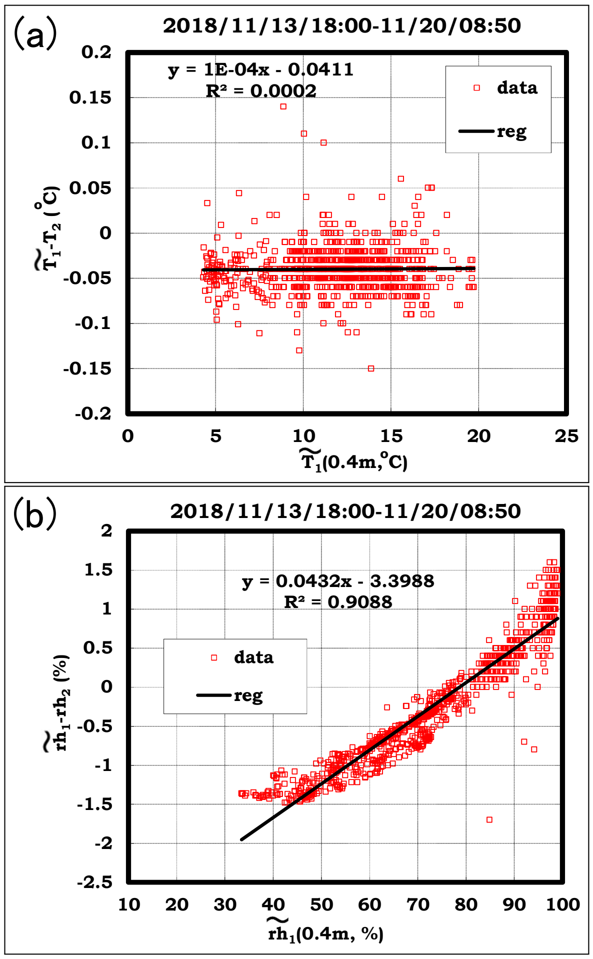


Fig. A2. Comparison of sensors installed at the same height. (a) Temperature sensor. (b) Hygrometer. Solid line and equation are the regression line and function, respectively. R^2 is coefficient of determination.

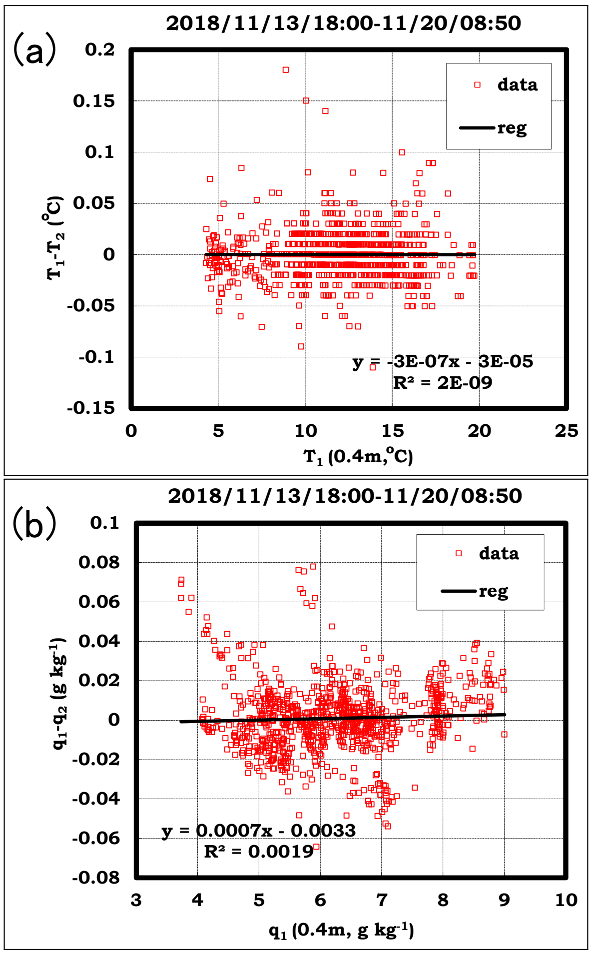


Fig. A3. Same as Fig. A2, but after correction of instrumental difference.

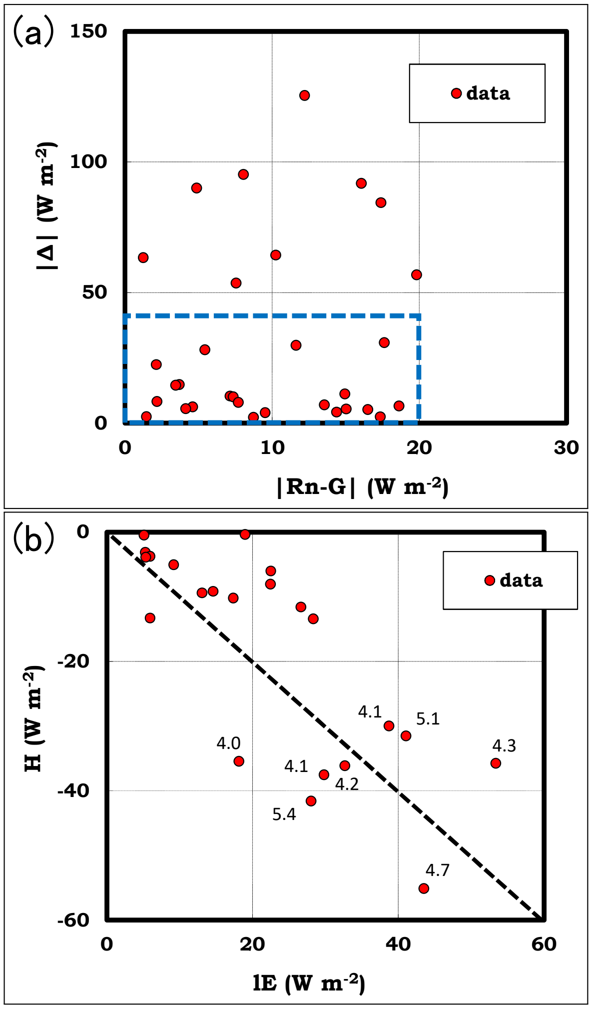


Fig. A4. (a) Relationship between $|Rn - G|$ and $|\Delta|$ under the condition of $|Rn - G| < 20$ W m^{-2} and 6.3 m s$^{-1} > U > 4$ m s^{-1}. (b) Relationship between H and lE for data within the dashed lines (i.e. $|\Delta| < 40$ W m^{-2}) in (a). Data labels indicate wind speed (m s^{-1}).

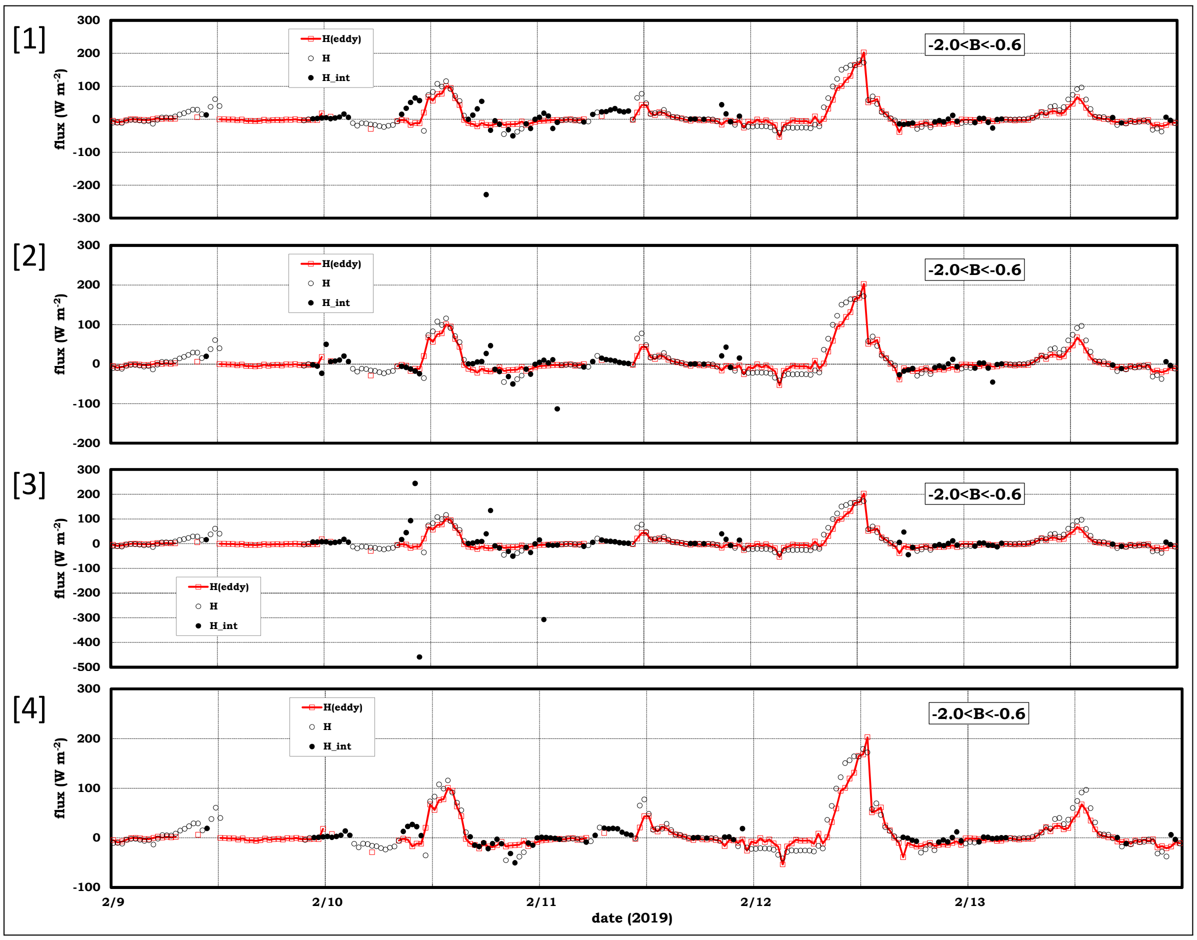


Fig. A5. Time-series example of sensible heat flux for each interpolation method [1]−[4].

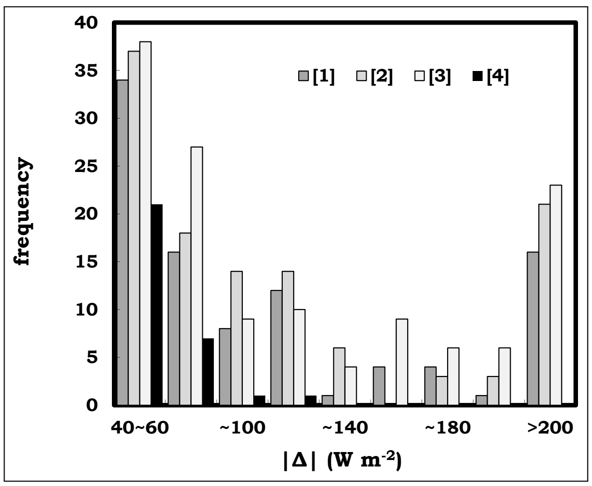


Fig. A6. Frequency distribution of spikes by class. [1]−[4] represent the interpolation methods.

List of Tables

Table A1. Summary of instrumental calibration. ΔT is the average corrected temperature difference between two temperature sensors installed at the same height, and $\sigma(\Delta T)$ is its standard deviation. Similarly, specific humidity is represented by Δq and $\sigma(\Delta q)$.

calibration period	T_1 °C	$\sigma(T_1)$ °C	q_1 g kg^{-1}	$\sigma(q_1)$ g kg^{-1}	calibration coefficient				sensor replacement
					ΔT		Δrh		
					a	b °C	c	d %	
2018/11/13–11/20	12.05	3.22	10.12	1.78	0.0001	−0.0411	0.0432	−3.3988	T_1&rh_1
2019/01/18–01/24	4.06	4.30	2.41	0.55	0.0013	0.0332	0.0040	0.6872	T_2&rh_2
2019/03/12–03/18	7.44	4.54	3.70	1.18	−0.0013	0.0398	0.0046	0.0326	
2019/05/17–05/28	20.97	4.93	9.68	2.02	0.0005	0.0297	0.0068	−0.4005	
average	11.13	4.25	6.48	1.38					

calibration period	after correction			
	$\Delta T=T_1-T_2$ °C	$\sigma(\Delta T)$ °C	$\Delta q=q_1-q_2$ g kg^{-1}	$\sigma(\Delta q)$ g kg^{-1}
2018/11/13–11/20	0.000	0.024	0.001	0.017
2019/01/18–01/24	0.000	0.031	−0.001	0.011
2019/03/12–03/18	0.000	0.047	−0.002	0.011
2019/05/17–05/28	0.001	0.030	−0.001	0.019
average	0.000	0.033	−0.001	0.014

Table A2. Interpolation results. Note: $\Delta = H(Bowen) - H(eddy)$. $|\Delta| > 40$ W m^{-2} is defined as a spike. Data for Period 1 and $Rn - G < 50$ W m^{-2}. n is the number of interpolated data points, and n_1 is the number of spikes. For $-2 < B < -0.6$ column, n_2 is the number of data points in the exclusion range.

				$-2 < B < -0.6$	
int method	data No. n	spike No. n_1	spike ratio n_1/n	data No. n_2	ratio n_2/n
[1]	700	96	13.7%	238	34.0%
[2]	700	116	16.6%	219	31.3%
[3]	700	132	18.9%	258	36.9%
[4]	700	30	4.3%	0	0.0%