

# Downwelling longwave radiation and sensible heat flux observations are critical for surface temperature and emissivity estimation from flux tower data

Gitanjali Thakur<sup>1,\*</sup>, Stanislaus J. Schymanski<sup>1,\*</sup>, Ivonne Trebs<sup>1</sup>, Mauro Sulis<sup>1</sup>, and Kaniska Mallick<sup>1</sup>

<sup>1</sup>Environmental Sensing and Modelling unit (ENVISION), Environmental Research and Innovation department (ERIN), Luxembourg Institute of Science and Technology (LIST), Belvaux, Luxembourg

\*gitanjali.thakur90@gmail.com

\*stanislaus.schymanski@list.lu

## ABSTRACT

Land surface temperature (LST) is a preeminent state variable that controls the energy and water exchange between the Earth's surface and the atmosphere. At the landscape-scale, LST is derived from thermal infrared radiance measured using space-borne radiometers. At the plot-scale, the flux tower recorded longwave radiation components are inverted to retrieve LST. Since the down-welling longwave component was not measured routinely until recently, usually only the up-welling longwave component is used for the plot-scale LST retrieval. However, we found that ignoring reflected down-welling longwave radiation for plot-scale LST estimations can lead to substantial error. This also has important implications for estimating the correct surface emissivity using flux tower measurements, which is needed for plot-scale LST retrievals. The present study proposes a new method for plot-scale emissivity and LST estimation and addresses in detail the consequences of omitting down-welling longwave radiation as frequently done in the literature. Our analysis uses ten eddy covariance sites with different land cover types and found that the LST values obtained using both up-welling and down-welling longwave radiation components are 0.5 to 1.5 K lower than estimates using only up-welling longwave radiation. Furthermore, the proposed method helps identify inconsistencies between plot-scale radiometric and aerodynamic measurements, likely due to footprint mismatch between measurement approaches. We also found that such inconsistencies can be removed by slight corrections to the up-welling longwave component and subsequent energy balance closure, resulting in realistic estimates of surface emissivity and consistent relationships between energy fluxes and surface-air temperature differences. The correspondence between plot-scale LST and landscape-scale LST depends on site-specific characteristics, such as canopy density, sensor locations and viewing angles. Here we also quantify the uncertainty in plot-scale LST estimates due to uncertainty in tower-based measurements using the different methods. The results of this work have significant implications for the combined use of aerodynamic and radiometric measurements to understand the interactions and feedbacks between LST and surface-atmosphere exchange processes.

## Introduction

The effects of global change are reflected in land surface temperature (LST) anomalies and their interannual variability<sup>1</sup>. It controls the magnitude and variability of the surface energy balance (SEB) components and simultaneously gets modulated by the SEB partitioning<sup>2,3</sup>. LST contains imprints of surface moisture and is extremely sensitive to evaporative cooling, which makes it a preeminent variable for studying evaporation and surface-atmosphere exchange<sup>4-6</sup>. It directly affects the amount of emitted longwave radiation and influences the saturation vapor pressure at the surface that drives latent heat flux. Thus, the ecohydrological functioning and carbon-water coupling are largely controlled by the surface temperature of the soil-vegetation system<sup>7</sup>. The availability of an extensive network of eddy covariance measurements (FLUXNET) allows us to understand the interactions and feedbacks between the surface-atmosphere exchange processes such as evaporation, transpiration, and its control by the atmosphere and vegetation at the diurnal time scale. However, the unavailability of direct LST measurements at the same scale hinders a detailed understanding of the interactions and feedbacks between LST and surface-atmosphere exchange processes, which is of utmost importance to the climate modeling community<sup>8</sup>.

Inversion of the longwave radiation in FLUXNET data to obtain LST has gained popularity in recent years. LST estimation depends on the emissivity of the underlying surface<sup>9</sup>, which is not available as routine measurement. Therefore, estimating in-situ LST is not straightforward due to the involvement of two unknowns (LST and emissivity) inside one measurement

variable (up-welling longwave radiation). To circumvent this challenge, we conducted simultaneous retrievals of LST and emissivity by exploiting the longwave radiation components in conjunction with associated SEB flux measurements<sup>10,11</sup>.

The SEB components can be sub-divided into radiative components (often lumped in net radiation,  $R_{net}$ ) and thermodynamic components, including sensible, latent and ground heat flux ( $H$ ,  $LE$ ,  $G$  respectively):

$$R_{net} = H + LE + G \quad (1)$$

The instantaneous value of LST is the result of interplay between the net radiation at the surface, ground heat flux ( $G$ ), sensible heat flux ( $H$ ) and latent heat flux ( $LE$ )<sup>12</sup>. Thus, LST can also be used for the estimation of  $H$ <sup>13</sup> and  $LE$ <sup>14</sup> between the surface and the atmosphere. LST provides the lower-boundary condition in SEB models for diagnostic estimates of  $LE$  and is highly relevant for drought monitoring<sup>2,5,15</sup>. As the surface-to-air temperature difference drives the exchange of sensible heat between surface and atmosphere, all components of Eq. (1) depend on the LST.

Net radiation ( $R_{net}$ ) can be sub-divided into down-welling and up-welling components<sup>16</sup> as shown below:

$$R_{net} = R_{sdown} + R_{ldwn} - R_{sref} - R_{lref} - R_{lem} \quad (2)$$

Only a fraction of solar top-of-the-atmosphere radiation reaches the Earth's surface, as some is reflected back to space by clouds, some is absorbed by the atmosphere and emitted later as longwave radiation. Reflected shortwave in Eq. (2) is expressed as  $R_{sref} = \alpha R_{sdown}$ , while reflected longwave is represented as  $R_{lref} = \alpha R_{ldwn}$ , where  $\alpha$  is the surface albedo. The emitted longwave radiation as a function of surface temperature ( $T_s$ ) and surface emissivity ( $\epsilon$ ) is given by Stefan-Boltzmann (SB) equation<sup>17</sup>

$$R_{lem} = \epsilon \sigma T_s^4 \quad (3)$$

where  $\sigma$  ( $Wm^{-2}K^{-4}$ ) is the SB constant,  $\epsilon$  is the surface emissivity ranging between 0 and 1, and  $T_s$  (K) is the LST. For a land surface, emissivity depends on soil type, vegetation cover, soil moisture, soil chemistry, roughness, spectral wavelength, temperature and view angle<sup>18</sup>.

The emitted and down-welling longwave radiance are measured at given angle within its instantaneous field of view (fov) by a downward facing sensor relatively close to the surface (a few meters for an eddy covariance tower). The radiation received by a pyrgeometer or infrared sensor is a combination of the radiation emitted ( $R_{lem}$ ) and reflected ( $R_{lref}$ ) by the surfaces in its fov as shown in Eq. (4):

$$R_{lup} = R_{lem} + R_{lref} \quad (4)$$

Substitution of Eq. (3) into Eq. (4) and replacing  $\alpha$  as  $1 - \epsilon$ ,  $R_{lup}$  becomes a function of emissivity, surface temperature and down-welling longwave radiation:

$$R_{lup} = \epsilon \sigma T_s^4 + (1 - \epsilon) R_{ldwn} \quad (5)$$

Eq. (5) is then solved for LST as a function of measured longwave and known surface emissivity:

$$T_s = \sqrt[4]{\frac{R_{ldwn}}{\sigma} - \frac{R_{ldwn}}{\epsilon \sigma} + \frac{R_{lup}}{\epsilon \sigma}} \quad (6)$$

In order to invert LST as shown in Eq. (6),  $\epsilon$  values are required. However, radiometers at eddy covariance sites (ECS) do not measure spectral bands separately to deduce emissivity directly. Therefore, we will deduce site specific  $\epsilon$  from observations of air temperature ( $T_a$ ), measured longwave ( $R_{ldwn}$ ,  $R_{lup}$ ) and sensible heat flux ( $H$ )<sup>19</sup>. In analogy to Ohm's law, the linear relationship between  $H$  and  $\Delta T$  can be expressed mathematically as:

$$H = m(T_s - T_a) \quad (7)$$

where  $m$  ( $ms^{-1}$ ) is a proportionality constant (defined as  $m = \rho C_p / r_a$  and broadly referred to as heat transfer coefficient) and depends on surface characteristics and micro-meteorology<sup>17</sup>,  $T_a$  (K) is the temperature of the air measured at a reference height above the surface,  $C_p$  ( $Jkg^{-1}K^{-1}$ ) is the specific heat capacity of air,  $\rho$  ( $kgm^{-3}$ ) is the air-density, and  $r_a$  ( $sm^{-1}$ ) is the total resistance to heat transport from surface to the atmosphere. It is evident from Eq. (7) that for  $T_s - T_a = 0$ ,  $H$  will be zero. This boundary condition and the linear relationship between  $H$  and  $\Delta T$  is used to estimate  $\epsilon$ <sup>10,19</sup>. Another approach for plot-scale  $\epsilon$  estimation filters the data where  $H$  is close to zero, substitutes  $T_s$  in Eq. (5) by  $T_a$  and solves for  $\epsilon$ <sup>11</sup>.

However, due to surface heterogeneity, sparse canopies are prone to footprint mismatch between the aerodynamic (flux tower) footprint and radiometric (hemispherical) footprint<sup>20-22</sup>, where the aerodynamic footprint represents the area contributing

to measured sensible heat flux, while the radiometric footprint is dominated by the surface below the sensor at a nadir viewing angle, contributing to the measured longwave radiation (used for  $T_s$  estimation). This can result in a different boundary condition i.e. at  $\Delta T = 0$ ,  $H \neq 0$  as expressed in Eq. (8):

$$H = m(T_s - T_a) + c \quad (8)$$

where  $H$  is representative of the sensible heat flux from the eddy covariance tower footprint,  $T_s$  is representative of all the radiating surfaces in the radiometric sensor's view, and  $c$  is interpreted as the  $H$  from surfaces in the aerodynamic footprint that are not seen by the radiometer.

Plot-scale estimation of  $\epsilon$  and LST using observed  $H$ ,  $T_a$ ,  $R_{lup}$  and  $R_{ldwn}$  as described above and in the Methods section, may be prone to substantial uncertainty. It is unclear how uncertainties in observed fluxes propagate into the uncertainty of estimated LST and  $\epsilon$ . By design, infrared thermal (IRT) sensors only measure up-welling infrared radiance and therefore cannot explicitly account for the amount of reflected down-welling infrared radiation in the signal. For a long time, down-welling longwave ( $R_{ldwn}$ ) was not routinely observed at ECS<sup>23</sup> and was also considered to be the most poorly quantified component of the radiation budget<sup>24</sup>. Therefore, the second term in Eq. (5) is commonly omitted, arguing that  $\epsilon \approx 1$ , and hence, Eq. (5) is simplified to Eq. (3)<sup>25</sup>:

$$R_{lup} \approx \epsilon \sigma T_s^4 \quad (9)$$

Eq. (9) can be solved for  $T_s$  to yield what we will term the "short equation" (seq) for  $T_s$ :

$$T_s \approx \sqrt[4]{\frac{R_{lup}}{\epsilon \sigma}} \quad (10)$$

Note that the above derivation is actually flawed, as the second term of Eq. (5) was omitted arguing that  $\epsilon \approx 1$ , and yet  $\epsilon$  was retained in the first part of the equation. Nevertheless, even with the availability of down-welling longwave measurements<sup>26</sup>, the use of Eq. (9) is still a common practice<sup>9,25</sup>. This gives rise to the question if the short equation (Eq. (10)) is adequate to estimate LST from ground-based measurements. In the remainder of this paper, we will refer to LST obtained using the long equation (Eq. (6)) as  $T_{leq}$  and to LST obtained using the short equation (Eq. (10)) as  $T_{seq}$ .

To better understand and improve approaches of plot-scale LST estimation, the present study addresses the following research questions:

1. Can we obtain an adequate estimate of plot-scale LST while neglecting the reflected down-welling longwave radiation?
2. Does the estimation of plot-scale  $\epsilon$  based on observed sensible heat flux ( $H$ ) have an advantage over satellite-derived  $\epsilon$  for plot-scale LST estimation?
3. How much uncertainty is introduced in plot-scale LST and  $\epsilon$  due to uncertainty in measured EC fluxes?

To answer these questions, we analysed data for ten eddy covariance sites in different biomes and climates (see Table 2). Plot-scale broadband monthly emissivity ( $\epsilon_{plot}$ ) was derived using observed  $H$  and estimated  $\Delta T$  as proposed by Holmes et al.<sup>10</sup>. Plot-scale LST was estimated using Eq. (6) and Eq. (10)  $\epsilon_{plot}$  landscape-scale emissivity ( $\epsilon_{MODIS}$ ). Estimated LST was compared with MODIS LST (TERRA satellite-sensed) for the times of satellite overpass. Uncertainty in  $\epsilon_{plot}$  and LST due to uncertainty in observed fluxes was calculated using SOBOL based uncertainty analysis (SALib)<sup>27</sup>. See the Methods section for more details.

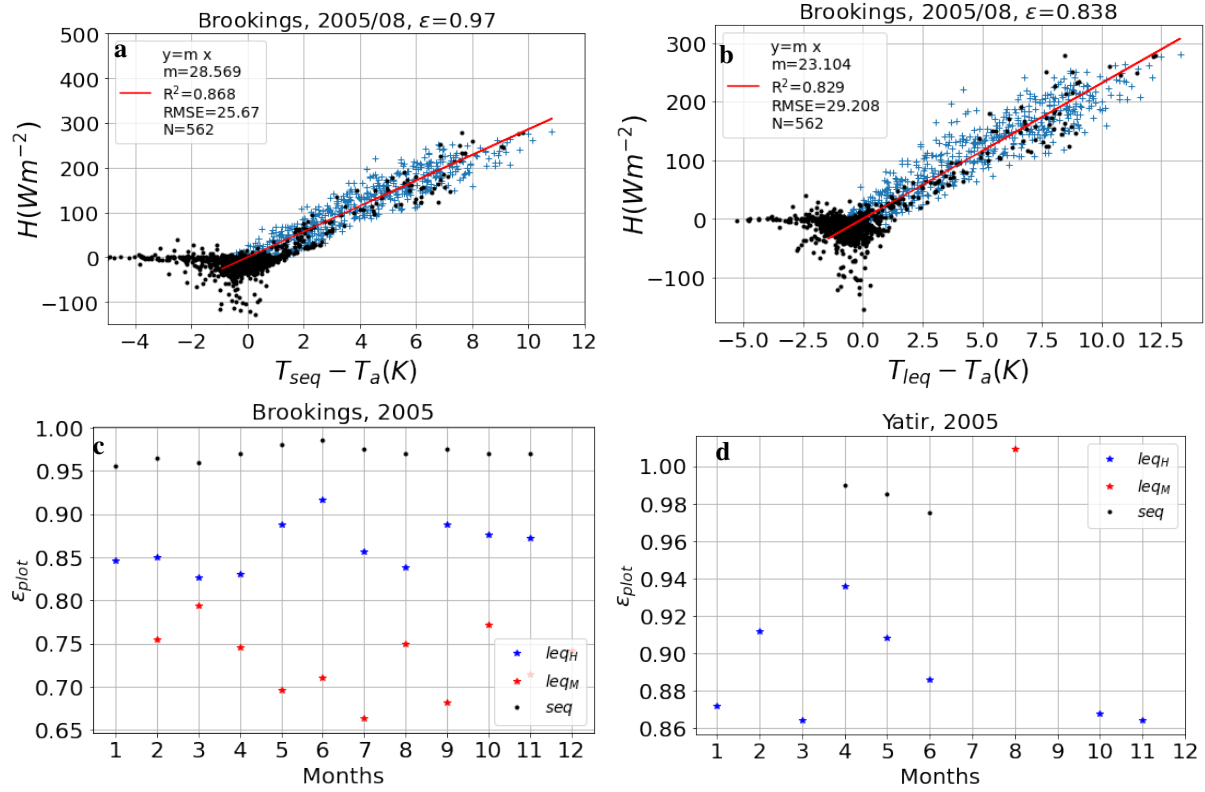
## 1 Results

### 1.1 Plot-scale $\epsilon$ using long and short equation

Following the method proposed by Holmes et al.<sup>10,19</sup>, plot-scale monthly  $\epsilon$  was estimated at the study sites by fitting  $\epsilon$  to minimise the root mean square error (RMSE) of the regression between  $H$  and  $T_s - T_a$  (see SI Figure 3). In Fig. 1a, c, and d, we used the original data and reproduced Figs. 2a, 3C, and 3Q from Holmes et al. (2009)<sup>10</sup> to validate our interpretation of their approach using the short equation (Eq. (10)). We noted only marginal differences between the two results based on the short equation, which are likely due to different fitting algorithms. The replication of the  $H(\Delta T)$  plot using the long equation (Eq. (6)) with the same data is given in Fig. 1b and the monthly  $\epsilon$  values are shown in Fig. 1c, d, indicated by blue stars. The retrieved LST values were slightly higher when using Eq. (6) (compare a and b in Fig. 1). The use of the long equation (Eq. (6)) resulted in substantially (10%) lower values of  $\epsilon$  as compared to the values estimated by Holmes et al.<sup>10</sup> for the common study sites (Brookings, Fig. 1c and Yatir, Fig. 1d). The reduction in  $\epsilon$  can be attributed to the sensitivity of the two equations to

the emissivity. As shown in the SI (Fig. 4), the  $T_s$  estimation using the short equation is more sensitive to  $\epsilon$  than for the long equation, thus even a small reduction in  $\epsilon$  can lead to a large increase in the  $T_s$  (to minimise RMSE).

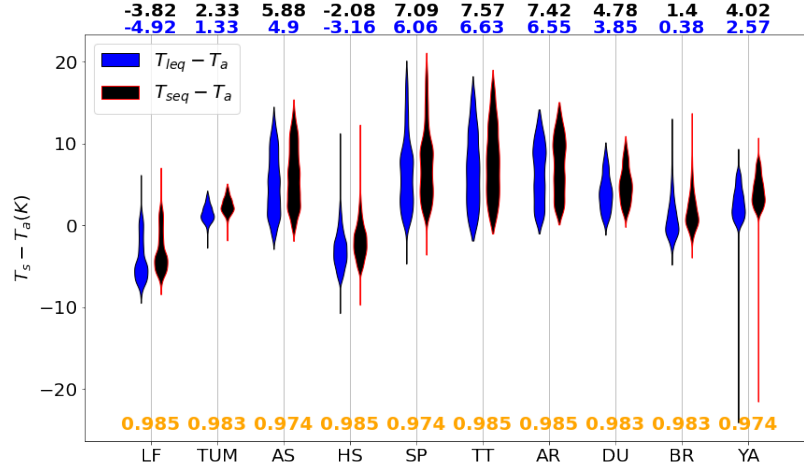
Another approach for plot-scale  $\epsilon$  estimation (Maes et al. (2019)<sup>11</sup>) in combination with Eq. 6) resulted in even lower  $\epsilon$  values for Brookings, as shown in Fig. 1c (red stars), whereas at Yatir, this approach gave an  $\epsilon$  value higher than 1 (red star in Fig. 1d). Note that the long equation also yielded  $H(\Delta T)$  relationship for more months at Yatir Forest (blue star) than the short equation (black dots) as shown in Fig. 1d, as it resulted in achieving a strong correlation between  $H$  and  $\Delta T$  (see Methods section for details). The pattern of lower  $\epsilon$  and higher LST using the long equation compared to the short equation was confirmed for all the ten sites used in the present study (SI Table 2).



**Figure 1.** Reproduction of analysis presented in Figs. 2a, 3C, and Q in Holmes et al. (2009)<sup>10</sup>. (a) Sensible heat ( $H$ ) vs.  $\Delta T = T_{seq} - T_a$  based on the short equation ( $T_{seq}$ , Eq. (10)); (b)  $H$  vs.  $\Delta T$  based on the long equation ( $T_{leq}$ , Eq. 6). Both show data for August 2005 at Brookings. Blue crosses represent data points satisfying the filtering criteria, while black dots represent points not considered in the analysis.  $N$  is the number of blue crosses used for regression (red line),  $m$  is the slope of regression,  $RMSE$  is the root mean square error and  $R^2$  is the square of the coefficient of determination. The fitted  $\epsilon$  value is reported in the title. (c) Optimised  $\epsilon$  values at Brookings obtained for the months where  $R^2 > 0.5$  using the short equation (Eq. 9, black dots) and long equation (Eq. 6, blue stars), and  $\epsilon$  obtained using the approach of Maes et. al (2019)<sup>11</sup> (red stars). (d) Same as (c), but for Yatir Forest, see table 2 for site descriptions.

## 1.2 Landscape-scale vs plot-scale estimates of $\epsilon$ and LST

At each site, LST was estimated using both the short equation ( $T_{seq}$ , Eq. 10) and the long equation ( $T_{leq}$ , Eq. 6). In the first step, tower-based longwave radiation and landscape-scale broadband  $\epsilon$  from MODIS spectral  $\epsilon$  ( $\epsilon_{MODIS}$ , Eq. 12) was used. The yearly daytime surface-to-air temperature difference for each study site is estimated and shown in Fig. 2. At all sites, Eq. (10) resulted in higher day-time plot-scale  $T_s$  estimates as compared to Eq. (6), when using  $\epsilon_{MODIS}$ , with the medians of surface-to-air temperature differences ( $\Delta T$ ) differing by 0.8 to 1.5 K (Fig. 2). The difference in  $\Delta T$  using the two equations is highest at the water limited sites, e.g. AS and YA. Note that for two sites (LF and HS), the median values of daytime  $\Delta T$  are negative. Comparison of plot-scale LST estimated using  $\epsilon_{MODIS}$  at satellite overpass time with landscape-scale LST ( $T_{MODIS}$ ) revealed strong correlations at most of study sites but systematically lower plot-scale LST (Fig. 3a, b). Use of  $\epsilon_{plot}$  for LST estimation ( $T_{seq}$  and  $T_{leq}$ ) resulted in substantial reduction of the bias as shown in Fig. 3c, d. This trend in bias reduction was similar at other sites (Table SI2 for details). The minimum bias is found at TUM, a closed canopy (eucalyptus forest) and the highest bias was obtained at LF and HS, heterogeneous ecosystems with sparse canopies (woodland savanna). However, for



**Figure 2.** Yearly distributions of half-hourly surface-to-air temperature differences ( $\Delta T = T_s - T_a$ ) for a representative year at each site. LST was calculated using the short equation (Eq. (10)) or long equation (Eq. (6)) with landscape-scale emissivity ( $\epsilon_{MODIS}$ ). The median values of  $\Delta T$  are shown at the top of the plot and the  $\epsilon_{MODIS}$  used for the  $T_s$  retrieval are shown at the bottom in orange. See Table 2 for site abbreviations.

some sites, weak correlation between satellite-derived and local LST estimates were also evident (at DU,  $R^2$  was reduced from 0.8 to 0.4, see Table SI2). The low correlation between MODIS LST and plot-scale LST can be due to various reasons, such as differences in sensor types, viewing angles and distance between the sensors and sources, e.g. requiring atmospheric correction for satellite-based sensors. Also, using plot-scale  $\epsilon$  for LST estimation resulted in positive  $T_s - T_a$  at LF and HS as shown in SI Figure 1 in comparison to Fig. 2.

### 1.3 Plot-scale $\epsilon$ estimation using long equation with intercept

In order to account for the possibility of bias between radiometric and aerodynamic measurements (e.g. due to footprint mismatch of measuring devices or instrument bias), we also fitted Eq. 8, i.e. a relationship allowing for an intercept in the linear fit between  $H$  and  $\Delta T$  (instead of forcing it through zero as in Fig. 1) for plot-scale  $\epsilon$  estimation. As shown in Fig. 4, the plot-scale  $\epsilon$  values resulting from this approach ( $H = m\Delta T + c$ ) were substantially closer to the landscape-scale  $\epsilon$  values compared with the approach without intercept ( $H = m\Delta T$ ), as shown in Table 1. However, comparison of the resulting plot-scale LST with landscape-scale LST values revealed an increase in bias at most sites compared to the LST obtained using  $\epsilon_{plot}$  without an intercept (Table 1). The median values of the resulting intercept ranged from  $-24$  to  $+258 \text{ W m}^{-2}$ , with the highest intercept values at Howard Springs (amounting to 70% of the maximum observed  $H$  at this site). The minimum value of intercept was obtained at Tumbumba (5% of the maximum observed  $H$ ). Note, that if we assumed just a slight under-estimation of up-welling longwave radiation by  $40 \text{ W m}^{-2}$  at Howard Springs (ca. 8% of observed  $R_{lup}$ ), the intercept was reduced from 294 (Fig. 4c) to  $17 \text{ W m}^{-2}$  (Fig. 5a) without change in other regression parameters ( $m$ , RMSE,  $R^2$ ). In this study, we did not apply any energy balance closure scheme, as a Bowen ratio closure, although resulting in higher  $R^2$  values at HS, also led to even greater intercept ( $c$ ) (Fig. 5b). Interestingly, adding  $40 \text{ W m}^{-2}$  to the measured up-welling longwave radiation and subsequent energy balance closure largely removes the intercept and at the same time increases  $R^2$ , as shown in SI Figure 6. Also, the bias between MODIS and plot-scale LST is reduced from  $-10.66 \text{ K}$  (Table 1) to  $4.01 \text{ K}$  by adding  $40 \text{ W m}^{-2}$  and closing the the energy balance.

### 1.4 Uncertainty in plot-scale $\epsilon$ and LST

Each of the observed input variables used for the estimation of plot-scale  $\epsilon$  and LST has an associated uncertainty. Here we present exemplary results for Alice Springs, which showed the highest correlation between plot-scale and landscape-scale LST estimations (Table 1). The uncertainty in plot-scale  $\epsilon$  estimated using Eq. 6 ('leq') and Eq. 7 (i.e. without intercept in  $H(\Delta T)$ ) was mainly in the range of  $\pm 0.02$  to  $\pm 0.05$ , with a maximum of  $\pm 0.2$  if outliers are included (blue color in Fig. 6a). The short equation (Eq. 10, 'seq') resulted in a very narrow range of  $\epsilon$  values between 0.94 and 0.99 throughout the year, with very small uncertainty (around  $\pm 0.01$ , black boxes in Fig. 6a). Interestingly, the differences in  $\epsilon$  uncertainty did not propagate into differences in LST uncertainty, which were around  $\pm 0.2 \text{ K}$  at the hourly scale for each equation if plot-scale emissivity was used (blue boxes in Fig. 6b and black boxes in Fig. 6c). In fact, if landscape-scale values of  $\epsilon$  were used, the LST uncertainty was even bigger ( $\pm 0.5 \text{ K}$ , orange boxes in Fig. 6b and c). The comparison of uncertainty in  $\epsilon_{plot}$  using short equation and the



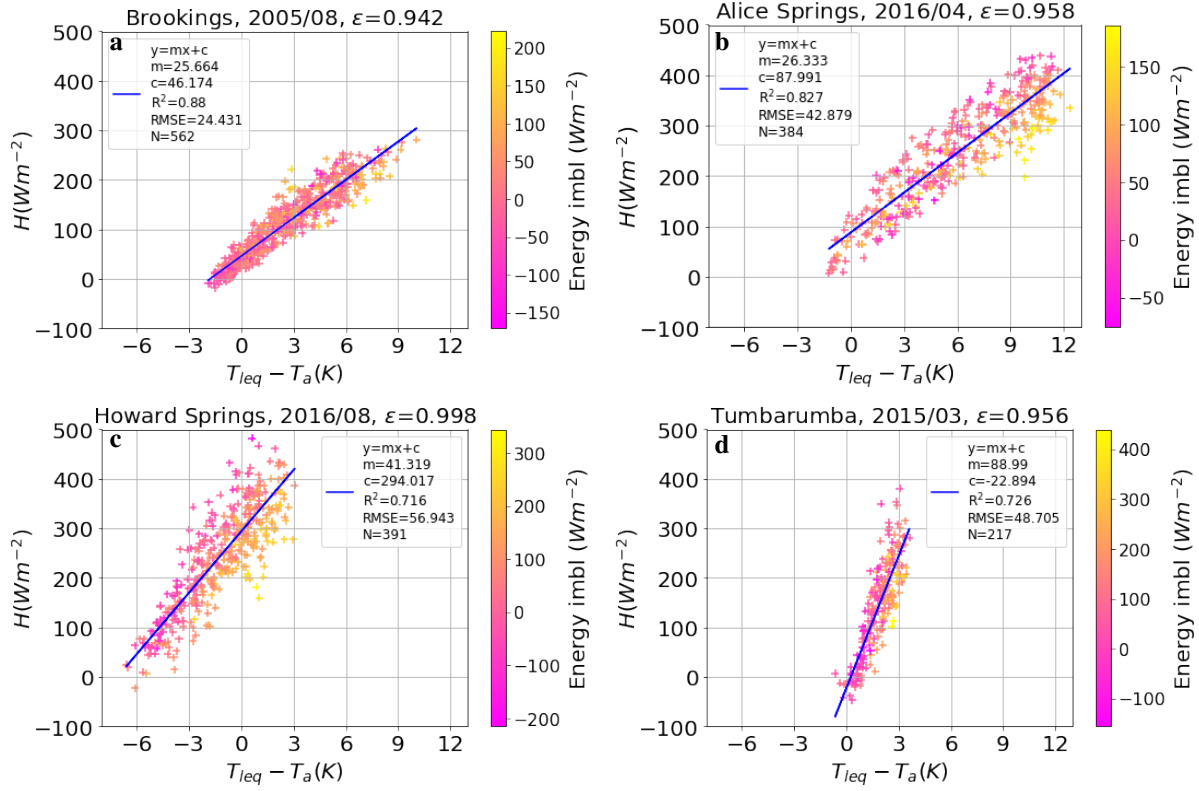


**Figure 3.** Landscape-scale LST ( $T_{MODIS}$  derived from MOD11A1) vs. plot-scale LST at Alice Springs for 2016-2018. **(a)**  $T_{seq}$  based on short equation (Eq. 10) and satellite-derived (MODIS) broadband emissivity; **(b)** Same as (a), but  $T_{leq}$  based on long equation (Eq. 6); **(c)**  $T_{seq}$  based on short equation (Eq. 10) and monthly plot-scale emissivity; **(d)** Same as (c), but  $T_{leq}$  based on long equation (Eq. 6). Bias is mean  $T_{seq} - T_{MODIS}$ , N is the number of daily overpasses of MODIS between 2016 and 2018, c is the intercept, m the slope, RMSE is the root mean square error and  $R^2$  is coefficient of determination. At each site, LST was estimated using both the short equation ( $T_{seq}$ , Eq. 10) and the long equation ( $T_{leq}$ , Eq. 6). In a first step, we used satellite-derived landscape-scale broadband emissivity from MODIS ( $\epsilon_{MODIS}$ , Eq. 12) for estimating plot-scale LST from tower-based longwave measurements, and compared these with landscape-scale LST extracted from MODIS LST dataset ( $T_{MODIS}$ )

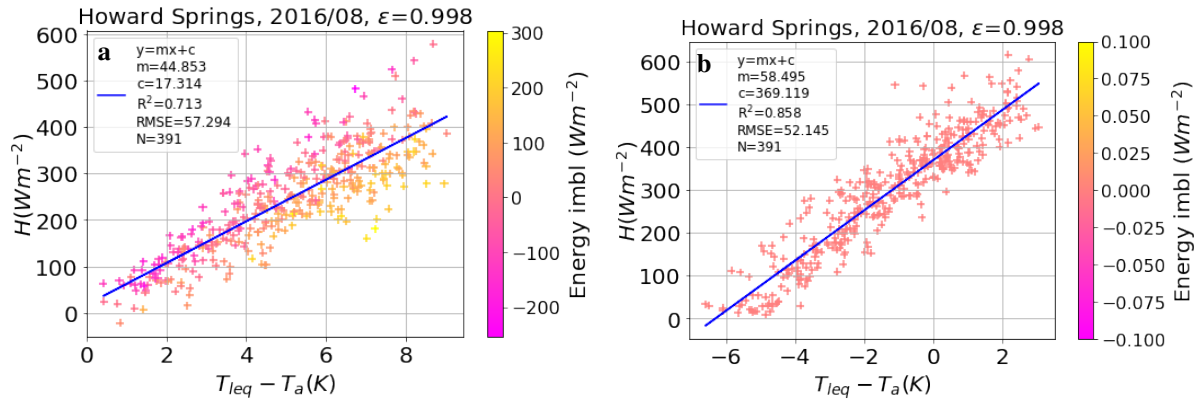
long equation with and without intercept is shown in SI Figure 7. The uncertainty in hourly  $T_s - T_a$  more than doubles if an intercept is allowed during the estimation of  $\epsilon$  (SI Figure 7b).

## 2 Discussion

Our analysis revealed a fundamental flaw in the commonly used short equation (Eq. (10)) for estimating plot-scale LST and  $\epsilon_{plot}$ , as it does not produce the same results as the long equation (Eq. (6)) even with high values of  $\epsilon_{MODIS}$ . In fact, the short equation strongly over-estimates the sensitivity of LST to  $\epsilon$  (SI Figure 4), as it neglects the fact that low emissivity results in a greater fraction of reflected longwave in the sensor signal (compare Eq. (10) and (6)). The sensitivity of the long equation (Eq. (6)) to  $\epsilon$  is driven by the contrast between  $R_{lup}$  and  $R_{ldwn}$ , whereas for the short equation (Eq. (10)), it is only driven by observed  $R_{lup}$  (SI Figure 2). For instance, an error of 0.01 in  $\epsilon$  at a water-limited site (e.g. AS) can cause an error of 0.17 K using Eq. (6) and 0.79 K using Eq. (10) respectively (SI Figure 4). This means that small errors in  $\epsilon$  can result in large differences in LST when using the short equation, or conversely, unrealistic LST values can conveniently be rectified by slightly changing the  $\epsilon$  value. This is illustrated e.g. in Fig. 6, where estimation of  $\epsilon_{plot}$  resulted in similar LST values between the short and long equations, but with vastly different  $\epsilon$  values and much greater uncertainty in estimated  $\epsilon$  using the long equation compared to the short equation. Considering that the short equation ignores an important component of longwave radiation, it must be concluded that in this case, it achieves seemingly the right results for the wrong reasons. The reduced sensitivity of the long equation (Eq. (6)) to  $\epsilon$  is of advantage for plot-scale LST estimation, since  $\epsilon_{plot}$  is usually unknown and therefore used as an approximate value<sup>9</sup>. However, when using the long equation in conjunction with plot-scale  $H$  measurements to estimate  $\epsilon_{plot}$ , we obtained unrealistically low  $\epsilon$  values at some sites (e.g. HS and LF, Table 1) in comparison to previously



**Figure 4.** Sensible heat flux as a function of surface-to-air temperature difference based on Eq. (8) ( $H = m(T_s - T_a) + c$ ).  $\epsilon$  was fitted to minimise RMSE of a robust linear regression. The title of the plot contains site, year, month and the fitted  $\epsilon$ -value. The legend correspond to Fig. 1. The colour code indicates the degree of energy imbalance of each data point (i.e.  $R_{net} - H - LE - G$ ).



**Figure 5.** Sensible heat flux as a function of surface-to-air temperature difference based on Eq. (8) ( $H = m(T_s - T_a) + c$ ). Same analysis and legends as in Fig. 4(c), but (a) After adding  $40 \text{ Wm}^{-2}$  to measured  $R_{lup}$ , and (b) after closing the energy imbalance using a Bowen ratio closure scheme.

reported  $\epsilon$  values for a soil-vegetation system<sup>28,29</sup>. This strong bias in plot-scale  $\epsilon$  estimates was largely removed if  $H(\Delta T)$  linearfit was allowed to have an intercept (Table 1, plot-scale  $\epsilon$ ). The intercept (i.e.  $\Delta T \neq 0$  at  $H = 0$ ) could be caused by combining measurements coming from instruments (radiometer, eddy covariance system) with different footprints<sup>21</sup>. The mismatch of source areas becomes important if the surface underlying the instruments has heterogeneous land cover. Although "footprint awareness" is often omitted at ECS under the assumption of homogeneity<sup>20</sup>, in patchy vegetation, the radiometer can be "seeing" a different vegetation fraction than that contributing to EC measurements, meaning that  $H \neq 0$  at  $\Delta T = 0$ . This problem was not detected by Holmes et al.<sup>10</sup>, as the short equation (Eq. (10)) was used, and due to its high sensitivity to  $\epsilon$  (SI Figure 4(a)) even a small reduction in  $\epsilon$  corrected the offset in  $H(\Delta T)$  (Fig. 1a). In contrast, when repeating the same analysis

Sites	Landscape-scale $\varepsilon$			Plot-scale $\varepsilon$ $H = m\Delta T$			Plot-scale $\varepsilon$ $H = m\Delta T + c$			
	$\varepsilon_{land}$	$R^2$	bias (K)	$\varepsilon_{plot}$	$R^2$	bias (K)	$\varepsilon_{plot}$	$R^2$	bias (K)	c ( $Wm^{-2}$ )
SP	0.974	0.81	-4.61	0.85	0.82	-1.91	0.92	0.774	-2.563	18.12
AS	0.974	0.93	-6.24	0.82	0.93	-1.92	0.993	0.915	-4.884	72.46
TT	0.974	0.57	-8.30	0.80	0.52	-4.02	0.939	0.521	-7.466	58.70
HS	0.985	0.16	-9.90	0.6	0.22	-2.47	0.949	0.18	-10.45	237.29
LF	0.985	0.41	-11.0	0.6	0.41	-2.57	0.968	0.378	-11.47	258
AR	0.985	0.27	-3.51	0.960	0.252	-2.98	0.996	0.27	-3.567	14.72
DU	0.985	0.81	4.61	0.985	0.425	-3.926	0.994	0.405	-4.603	-8.11
TUM	0.983	0.84	-2.10	0.97	0.89	-1.93	0.955	0.85	-1.696	-24.24
BR	0.983	0.937	-0.195	0.82	0.895	2.72	0.919	0.906	1.662	17.72
YA	0.974	0.855	-3.45	0.93	0.793	-0.582	0.873	0.826	0.073	-22.95

**Table 1.** Correspondence between daytime landscape-scale LST ( $T_{MODIS}$ ) and plot-scale LST ( $T_s$ ) (estimated at TERRA time of pass), using different emissivity estimates. The emissivity values used to retrieve plot-scale LST is either taken from MODIS ( $\epsilon_{land}$ ), or derived flux tower data ( $\epsilon_{plot}$ ), using Eq. 7 ( $H = m\Delta T$ ) or Eq. 8 ( $H = m\Delta T + c$ ). The reported  $\epsilon_{plot}$  and intercept ( $c$ ) are median values over all months for each site. Bias is defined as the mean of  $T_s - T_{MODIS}$ ,  $R^2$  is the coefficient of determination between plot-scale LST in comparison to landscape-scale LST. The site acronyms are explained in Table 2.

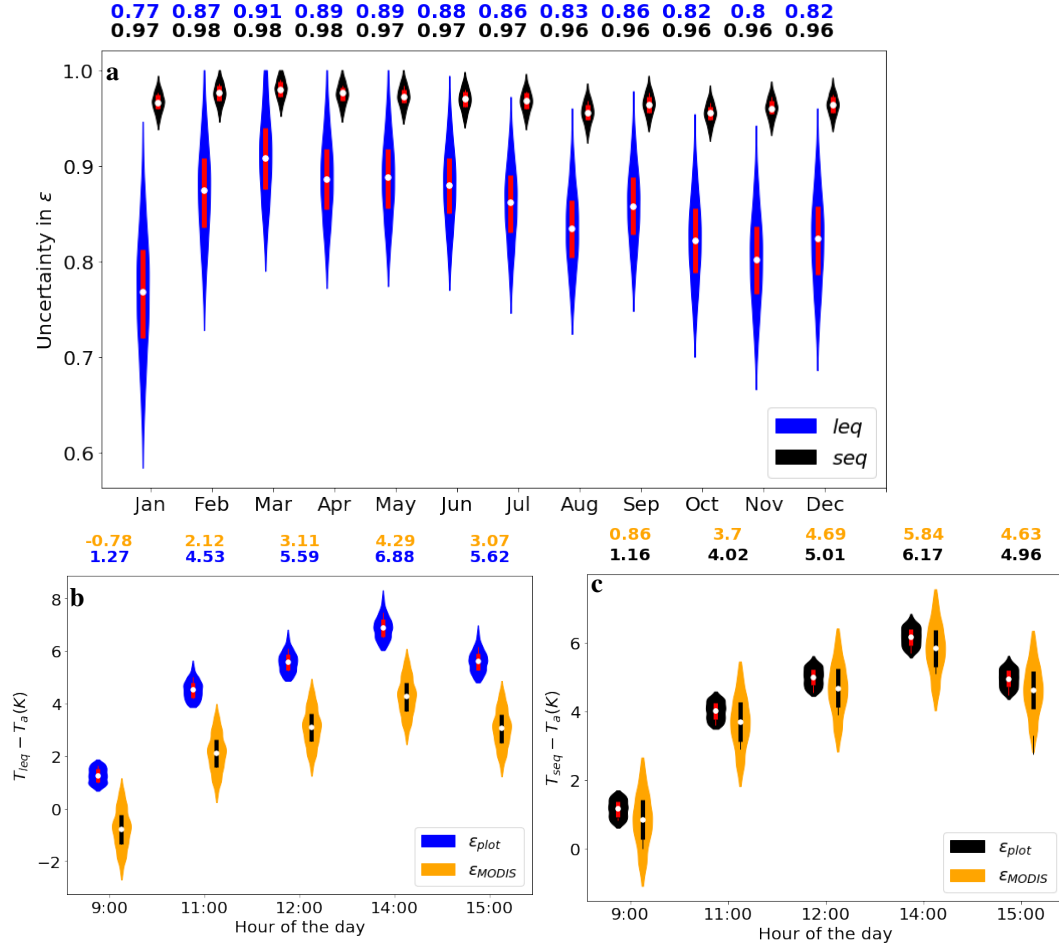
using the long equation (Eq. (6)), a larger reduction in  $\epsilon$  is required to remove the intercept, resulting in lower  $\epsilon$  (Fig. 1b). By allowing an intercept in the  $H(\Delta T)$  linear fit, we implicitly account for the possibility of a footprint mismatch or instrument bias in the data. This small change in methodology enables us to detect such problems by inspecting the value of the intercept ( $c$ ). Considering the aerodynamic footprint to be larger than the radiometric footprint<sup>20,21</sup>, a positive intercept can be interpreted as the  $H$  from the aerodynamic footprint which is not seen by the radiometer.

The intercept was very high for the sites HS and LF (Table. 1). A close inspection of the  $H(\Delta T)$  plots at these sites (SI Figure1) revealed negative day-time  $T_s - T_a$  (Fig. 2), which may suggest an underestimation of  $R_{lup}$ . While testing this hypothesis at HS (having the highest intercept, Fig. 4c) we found that adding roughly  $40 Wm^{-2}$  (approx. 8% of observed  $R_{lup}$ , Fig. 5 in observed  $R_{lup}$  led to significant reduction in the intercept from  $294 Wm^{-2}$  (Fig. 4c) to  $17 Wm^{-2}$  and positive day-time  $T_s - T_a$  (Fig. (5a)). The other linear regression parameters ( $m$ ,  $R^2$ , RMSE) were not affected (compare Fig. 5a and Fig. 4c). The hemispherical view of the radiometers looking down at a heterogeneous canopy makes it possible that they "see" more tree crowns and less soil than the area contributing to the eddy covariance footprint. This could lead to an underestimation of  $R_{lup}$ , and an underestimation by  $30 - 40 Wm^{-2}$  would be equivalent to approximately 5-10% of the observed flux, which is within the range of a typical energy imbalance found at this site. Previous studies have found a dependence of footprint mismatch on wind direction<sup>20-22</sup>, but we did not find a significant relation between monthly intercept and dominant wind direction at Howard Springs.

Surface heterogeneity has also been recognized as one of the potential causes for the lack of energy balance closure observed at most ECS at diurnal scales<sup>30,31</sup>. However, in our analysis the use of an energy balance closure scheme (based on the Bowen ratio) led to much lower values of  $\epsilon_{plot}$  using Holmes approach with the long equation and without intercept. In contrast, if an intercept was allowed, energy balance closure led to an increase in positive intercept (Fig. 5b). Perhaps this is the reason why other studies on plot-scale  $\epsilon$  estimation have also used the observed fluxes without correction<sup>10,11,19</sup>. Other energy balance closure schemes add the missing energy to  $H$  in water limited ecosystems<sup>32</sup>, or to  $LE$  in energy limited ecosystems<sup>33</sup>. However, our analysis suggests that the footprint mismatch may cause a small bias in the up-welling longwave radiation measurements that is not accounted in any conventional energy balance closure approaches. When we added  $35 Wm^{-2}$  (instead of  $40 Wm^{-2}$ , see Fig. 5a) to the measured up-welling longwave radiation and subsequently closed the energy balance at HS site (which had the largest  $H(\Delta T)$  intercept), we largely removed the intercept and at the same time obtained realistic  $\epsilon$  values and an increased  $R^2$  (SI Figure 6). In addition, the bias between MODIS LST and plot-scale LST at HS was reduced by 6.4 K (SI Fig.6b) compared to using up-welling longwave without correction.

When estimating plot-scale LST using  $\epsilon_{MODIS}$  values, we found at many sites with a sparse canopy strongly negative bias in comparison to MODIS LST, which is in agreement with previous studies where the bias for sparse canopies reached up to  $12 K$ <sup>34</sup>. The MODIS overpass can have a large off-Nadir viewing angle, which would lead to an elongated foot-print<sup>35</sup> and therefore, a different distribution of bare soil and vegetated areas compared to the mostly Nadir viewing angle of the tower-mounted sensor. The difference in footprint and viewing angles between the tower mounted pyrgeometers and MODIS radiometers could also be the reason for bias between the two LST estimates. Plot-scale LST estimates based on plot-scale





**Figure 6.** Uncertainty in plot-scale estimations of  $\epsilon$  and surface-air temperature differences ( $T_s - T_a$ ) at Alice Springs (AS), based on Eq. 7 (no intercept in  $H(\Delta T)$ ). Monthly values of  $\epsilon$  shown for 2017 and hourly  $T_s - T_a$  for 15 August 2017. (a) Uncertainty in monthly  $\epsilon_{plot}$  due to uncertainty in  $H$ ,  $R_{lup}$ ,  $R_{ldw}$  and  $T_a$ , using Eq. 5 ('leq', blue) and Eq. 9 ('seq', black). (b) Hourly uncertainty in  $T_s - T_a$  on 15 July based on Eq. 5, due to uncertainty in  $R_{lup}$ ,  $R_{ldw}$  and  $T_a$  when landscape-scale emissivity is used ( $\epsilon_{MODIS}$ , orange) or due to uncertainty in  $H$ ,  $R_{lup}$ ,  $R_{ldw}$  and  $T_a$  when  $\epsilon_{plot}$  is used (blue). (c) Same as Panel b, but based on Eq. (9).

$\epsilon$  using a linear  $H(\Delta T)$  fit without an intercept largely reduced this bias between plot-scale and MODIS LST (Table 1) and also reduced the uncertainty in diurnal LST (Fig. 6b, c) in comparison to the use of  $\epsilon_{MODIS}$ . However, the resulting plot-scale  $\epsilon$  values were unrealistically low at some sites (Table 1, center). In contrast, allowing an intercept ( $H = m * \Delta T + c$ ) in  $\epsilon_{plot}$  estimation resulted in more realistic  $\epsilon$  values at these sites, but very large intercept values (over  $200 \text{ Wm}^{-2}$  at some sites), indicating that the plot-scale LST values cannot be used in combination with the observed aerodynamic fluxes at these sites, as strongly positive  $H$  at 0 surface-air temperature difference is physically inconsistent (Fig. 4c). In addition, this approach increased the bias between plot-scale and MODIS LST at most of the study sites (Table 1). Note that the correspondence between landscape-scale LST and plot-scale LST can vary strongly between sites, depending on canopy densities and viewing angles (tower vs. satellite)<sup>35</sup>, sensor installation height and position, and sensor types<sup>21</sup>. At the sites with the largest intercept values, we found that an assumed bias in up-welling longwave radiation by only 6-9% would largely remove the intercept and also reduce the bias between MODIS and plot-scale LST (Fig. 5a, SI Figure 6(b)). A detailed analysis of such bias and potential correction approaches is beyond the scope of this study. Given that the fit of a linear model without intercept is statistically questionable in general<sup>36</sup>, and the fact that such a fit resulted in unrealistically low values of  $\epsilon$  at some sites, we conclude that fitting a model with intercept is the more robust approach, and that a significant intercept should be used as a red flag for the utility of the data for estimation of plot-scale LST.

Note that the fluxes observed at ECS are representative of the composite signal from both, soil and vegetation, which typically have a different range of surface temperatures and emissivities<sup>37</sup>. The  $\epsilon$  of soil strongly depends on soil moisture content<sup>38</sup>, whereas the  $\epsilon$  of a canopy depends on its structural attributes and leaf area index, the latter of which can vary

strongly at the seasonal scale<sup>39</sup>. For example, the laboratory-measured directional  $\epsilon$  for various canopy elements (bark, leaf and its arrangement, stem wood) ranged between 0.9 to 1<sup>40</sup>. Laboratory measurements of thermal infrared reflectance spectra suggest that the  $\epsilon$  uncertainty due to structural unknowns, such as leaf orientation, is more significant than the differences in leaf component emissivity among plant species<sup>29</sup>. Consequently, it is clear that the  $\epsilon$  of a surface is a function of many factors and detailed study of all these factors is out of scope of the present study. Derivation of landscape-scale broadband emissivity ( $\epsilon_{MODIS}$ ) from narrowband spectral emissivity is a first-order approximation for capturing the integrated effects of land cover from MODIS spectral bands<sup>37</sup>, whereas the derivation of  $\epsilon_{plot}$  from EC flux data provides an independent alternative for ecosystem-scale  $\epsilon$  estimation. The study could also be extended for urban settings with eddy covariance measurements and known anthropogenic heat components. However,  $\epsilon$  of urban surfaces is complicated by other factors e.g variability in building materials to morphology to thermal anisotropy, which adds another source of uncertainty to  $\epsilon$  estimates<sup>33</sup>.

In summary, our results reveal that the short equation (Eq. (10), neglecting down-welling longwave radiation) leads to biased estimates of LST and hugely over-estimated sensitivity of LST to surface emissivity. Therefore, the use of Eq. (10) is not recommended and should be replaced by Eq. (6) if down-welling longwave radiation measurements are available. At some sites, the use of Eq. (6) resulted in plot-scale LST estimates that were far below satellite-derived landscape-scale LST values, and also inconsistent with plot-scale flux data (negative surface-air temperature difference when sensible heat flux is strongly positive). In many previous studies, such bias would have been removed by slightly lowering surface emissivity ( $\epsilon$ ), but the reduced sensitivity of Eq. (6) to  $\epsilon$  would require unrealistically low values of  $\epsilon$  to remove the low-bias in LST. When estimating plot-scale  $\epsilon$  values, realistic estimates based on Eq. (6) are only possible at these sites if we include an intercept in the  $H(\Delta T)$  relationship, but this again results in very high intercept values (over 200 W m<sup>-2</sup>). However, assuming that the intercept is a consequence of a foot print mismatch between the aerodynamic and radiometric measurements, a small correction in up-welling longwave (6-9%) and subsequent energy balance closure (based on Bowen ratio) largely removed the intercept and produced realistic  $\epsilon$  values and self-consistent  $H(\Delta T)$  plots. This approach also reduced the bias between plot-scale LST and MODIS LST, although it did not improve the weak correlation between them (SI Figure 6(b)). The intercept value can be used as a criterion to check the consistency of observed data (radiometric and aerodynamic measurements) before using them in combination, as a strong intercept indicates inconsistency between observed sensible heat flux and surface-to-air temperature difference. Therefore, the proposed method of fitting a linear relation with intercept to  $H$  and  $\Delta T$  data has the potential to provide more robust benchmark data sets for model evaluation and validation at the ecosystem scale (plot-scale). Unfortunately, we are currently not in the position to predict a priori at which sites (ecosystems) a low correlation between plot-scale and landscape-scale LST should be expected.

The  $\epsilon_{plot}$  estimation by combining radiometric and aerodynamic measurements also provides a quality check on the observation correspondence at ECS. The independent estimate of LST and  $\epsilon$  using eddy co-variance measurements can be used to parameterize climate and weather prediction model at ecosystem scale. Overall, the implications of our study are of particular relevance for the research community interested in understanding the diurnal and seasonal feedback in soil-vegetation systems based on observed fluxes.

### 3 Methods

In the last two decades, plot-scale radiometric data collected at ECS have gained popularity for in-situ LST retrieval due to its high temporal resolution<sup>31,42</sup>. In addition to this, the LST estimates at plot-scale originate from a relatively homogeneous footprint in comparison to the satellite-derived LST (MODIS pixels). This section describes: (i) how to retrieve plot-scale LST and  $\epsilon$  using eddy covariance measurements, (ii) how to quantify the correspondence between plot-scale LST with MODIS LST and, (iii) how to quantify the uncertainty in plot-scale LST and  $\epsilon$ .

**Tower data:** ECS collect micro-meteorological measurements above the surface (vegetation canopy) using towers (flux tower) following common measurement protocols<sup>43</sup>. The towers are generally equipped with an instrument made up of pyrgeometers or radiometers to measure up-welling and down-welling shortwave and longwave radiation, which is further used to calculate net radiation (Eq. 2). Besides radiative fluxes, measurement at ECS also include sensible and latent heat fluxes, net carbon-dioxide exchange and a range of meteorological variables, such as air temperature ( $T_a$ ), humidity and wind speed.  $T_a$  is the air temperature measured at a reference height above the canopy. Each flux measurement is accompanied by a flagging system based on the second CarboEurope-IP QA/QC workshop<sup>44</sup>. In our current work, we use high quality available data (flag 0) as it is without atmospheric corrections. For the analysis, ten sites were selected to represent a variety of land cover types and climates (Table 2). Eight sites belong to the North Australian Tropical Transect (NATT) and two sites (Yatir Forest, Brookings) are chosen to replicate results from Holmes et.al<sup>10</sup> as shown in Table 2. Eddy covariance level 3 data is obtained from <http://data.ozflux.org.au/portal/pub/listPubCollections.aspx> for Australian sites. The data for Brookings was obtained from Ameriflux whereas the data for Yatir Forest was obtained through personal communication with Professor Yakir's lab in order to obtain the older version of the data, which was used by Holmes et al.<sup>10</sup>.

**MODIS data:** Landscape-scale emissivity and LST data (MODIS product MOD11A1) was downloaded from NASA earth

data . It is a level 3 daily LST product gridded in the sinusoidal projection at a spatial resolution of 0.928 km by 0.928 km. The daily LST pixel values in each granules (tile contains 1200 x 1200 grids in 1200 rows and 1200 columns) is retrieved by the generalized split-window algorithm under clear-sky conditions and MODIS LST values are averaged by overlapping pixels in each grid with overlapping areas as weight<sup>45</sup>. The downloaded data in hierarchical data format (hdf), were converted into tagged image file format (tiff) using a python package called PyModis<sup>46</sup>. From tiff the files are converted into csv format. The details of data extraction and conversion can be found at [https://renkulab.io/projects/gitanjali.thakur/modis\\_lst\\_fpar/](https://renkulab.io/projects/gitanjali.thakur/modis_lst_fpar/). Thermal remote sensing is used to measure spectral emissivity through four channels (28, 29, 30, 31) at wavelengths ranging between 8-12  $\mu\text{m}$ <sup>37</sup> and the system of equations is iteratively solved for a given range of wavelengths (8 - 12  $\mu\text{m}$ ) to obtain  $\epsilon$  and LST using radiative transfer models<sup>23,37,47</sup>. In the current study, dataset columns used to compare plot-scale LST are: day time daily LST and local view time. In order to obtain landscape-scale  $\epsilon$  the emissivity from bands 31 and 32 are used. These bands have more stable emissivities ranging from 0.92-1, thus used to derive broadband emissivity<sup>45</sup>.

**Table 2.** Description of study sites

Study site	Latitude, Longitude	Landcover	Time-period	Longwave sensors	Sensor installation height (m)	Altitude (m)
Sturt Plains (SP)	-17.1507, 133.3502	Mitchell Grass	2016-2019	pyrgeometers (CG-2)	4.8	230
Alice Springs (AS)	-22.2828, 133.2493	Mulga woodland, hummock grassland, river red gum forest	2016-2018	radiometer (CNR1)	12.2	606
Ti Tree East (TT)	-22.2870, 133.6400	Grassy mulga woodland, Corymbia/Triodia savanna	2016-2018	radiometer (CNR1)	9.9	553
Howard Springs (HS)	-12.4943, 131.1523	Open woodland savanna	2016-2018	pyrgeometers (CM-7B, CG-2)	23	63
Litchfield (LF)	-13.1790, 130.7945	Tropical savanna	2016-2018	radiometer (CNR4)	31	222
Adelaide River (AR)	-13.0769, 131.1178	Savanna dominated by Eucalyptus tectifica and Planchonia careya	2006-2009	pyrgeometers (CNR1)	15	90
Daly Uncleared (DU)	-14.1592, 131.3881	Woodland savanna	2016-2018	radiometer (NR-lite)	21	110
Tumbarumba (TUM)	-35.6566, 148.1517	Wet sclerophyll	2015-2018	pyrgeometers (CM3 and CG3)	70	1200
Brookings (BR)	44.352, 96.840	Cropland	2005	pyrgeometers <sup>48</sup>	NA	510 <sup>49</sup>
Yatir Forest (YF)	31.344894, 35.051922	Evergreen needleleaf forest	2005	pyrgeometers <sup>48</sup>	NA	641

**Plot-scale  $\epsilon$  and LST estimation:** LST is the “ensemble directional radiometric surface temperature”<sup>18</sup>, and can be estimated from the infrared radiance emanating from a given surface with known emissivity<sup>50</sup>. We solve Eq. (6) and Eq. (10) measured longwave and known surface emissivity (we start with maximum possible value for grey body, 0.99). This approach was initially proposed by Holmes<sup>10</sup> to estimate plot-scale  $\epsilon$  using the short equation (Eq. 10) with  $H$ ,  $R_{lup}$ ,  $T_a$ . In the present work, we have used both the long equation (Eq. 6) and short equation (Eq. 10) to estimate  $\epsilon$ . The prime variables used in the study are  $H$ ,  $R_{lup}$ ,  $R_{ldwn}$ , and,  $T_a$ , whereas the ancillary variables  $R_n$  and wind speed ( $Ws$ ) are used to filter the data for analysis.

The data filtering criteria are ( $R_n > 25 \text{ Wm}^{-2}$ ) and wind speed ( $Ws > 2 \text{ ms}^{-1}$ )<sup>10</sup>. Data ( $R_{lup}$ ,  $R_{ldwn}$ ,  $H$ ,  $T_a$ ) satisfying the filtering criteria are segregated for each month. Starting  $\epsilon$  value of 0.4 is used with observed longwave radiation ( $R_{lup}$ ,  $R_{ldwn}$ ) is used to estimate  $T_s$ ,  $T_s - T_a$  is estimated ( $T_a$  is measured). For each month, a linear regression (with or without intercept, see main text) is performed between sensible heat ( $H$ ) and surface-air temperature difference ( $T_s - T_a$ ) (Fig. 7b) using Scipy <https://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.stats.linregress.html>. This procedure is repeated using a predefined range of  $\epsilon$  from 0.4 to 0.998, step size of 0.002.

From all the values of  $\epsilon$  that resulted in strong correlation between  $H$  and  $T_s - T_a$  (with  $R^2 > 0.5$ ), the value that achieved the smallest RMSE was chosen as the reliable emissivity value ( $\epsilon_{plot}$ ) for a given month. An illustration plot for RMSE and  $\epsilon$  is shown in SI Fig. 5. The  $\epsilon_{plot}$  is obtained for each month using both, long (Eq. (6)) and short equation (Eq.(10)) and termed as  $\epsilon_{leq}$  and  $\epsilon_{seq}$  respectively, as shown in Fig. 7b. For sites like HS and LF having a high value of intercept, we hypothesized an underestimation of  $R_{lup}$  and corrected it by adding 6 – 8% of the observed  $R_{lup}$  to the measured  $R_{lup}$ . After correcting  $R_{lup}$ , energy imbalance is closed for each month and monthly  $\epsilon_{plot}$  is obtained. Recently, another approach for plot-scale  $\epsilon$  estimation using Eq. (4) was discussed by Maes et.al (2019)<sup>11</sup>. For each month, the observed  $R_{lup}$ ,  $R_{ldwn}$ ,  $T_a$  are filtered for non rainy days using the filtering criteria ( $-2 < H < 2$ , and  $\alpha < 0.4$ ). The  $\epsilon$  is estimated by substituting  $T_s = T_a$  in Eq. (4) as shown in Eq. (11). The monthly  $\epsilon$  was represented as the median of  $\epsilon$  obtained by substituting filtered data in Eq. (11).

$$\epsilon = \frac{R_{ldwn} - R_{lup}}{R_{ldwn} - T_a^4 \sigma} \quad (11)$$

**LST comparison:** MODIS LSTs are a global reference for LST and used world-wide, also in conjunction with plot-scale flux measurements. Plot-scale LST based on optimum  $\epsilon$  for each month using a linear regression with and without intercept  $H(\Delta T)$  is compared to the MODIS daily LST (without correction for the viewing angle). For each month the  $T_s$  estimates using plot-scale  $\epsilon$  and MODIS  $\epsilon$  corresponding to the exact TERRA day-time of pass was obtained (30 minute tower data was upsampled to each minute using linear interpolation). TERRA (satellite) overpasses at local solar time between 10:30 am to 12 pm in ascending mode<sup>48</sup>. Plot-scale LST at local satellite overpass ( $T_{leq}$  and  $T_{seq}$ ) was obtained using measured longwave radiation that was interpolated to match the exact time of overpass and corresponding monthly plot-scale  $\epsilon$ . Plot-scale daily LST is compared to MODIS LST in terms of mean, bias, RMSE and  $R^2$  using a robust linear regression model (scipy stat model, [https://github.com/scipy/scipy/blob/v1.7.1/scipy/stats/\\_stats\\_mstats\\_common.py](https://github.com/scipy/scipy/blob/v1.7.1/scipy/stats/_stats_mstats_common.py)) as shown in Fig. 7a. The goodness of fit between plot-scale and landscape-scale LST was determined by looking at  $R^2$  (Fig. 7b). The bias is estimated as the mean of deviation between daily MODIS LST and ground-based  $T_s$  ( $T_{leq,seq} - T_{MODIS}$ ). In the past ground based radiometric measurements have been used for validation of MODIS LST product<sup>49</sup> whereas the we compared the plot-scale LST with MODIS LST as sanity check. See SI Table 1 for data sources and acronyms.

**General approach:** We estimate landscape-scale broadband  $\epsilon$  using MODIS spectral  $\epsilon$  as shown in Bahir et. al (2017)<sup>51</sup>.

$$\epsilon_{MODIS} = 0.4587\epsilon_{31} + 0.5414\epsilon_{32} \quad (12)$$

Tower-based longwave radiation measurement ( $R_{lup}$ ,  $R_{ldwn}$ ) passing the filtering criteria (as mentioned in plot-scale emissivity estimation) along with MODIS based  $\epsilon$  was used to invert LST using Eq. (6) and Eq. (10). The obtained plot-scale LST was compared to landscape-scale MODIS LST using a robust linear regression as mentioned above and shown in Fig. 7a.

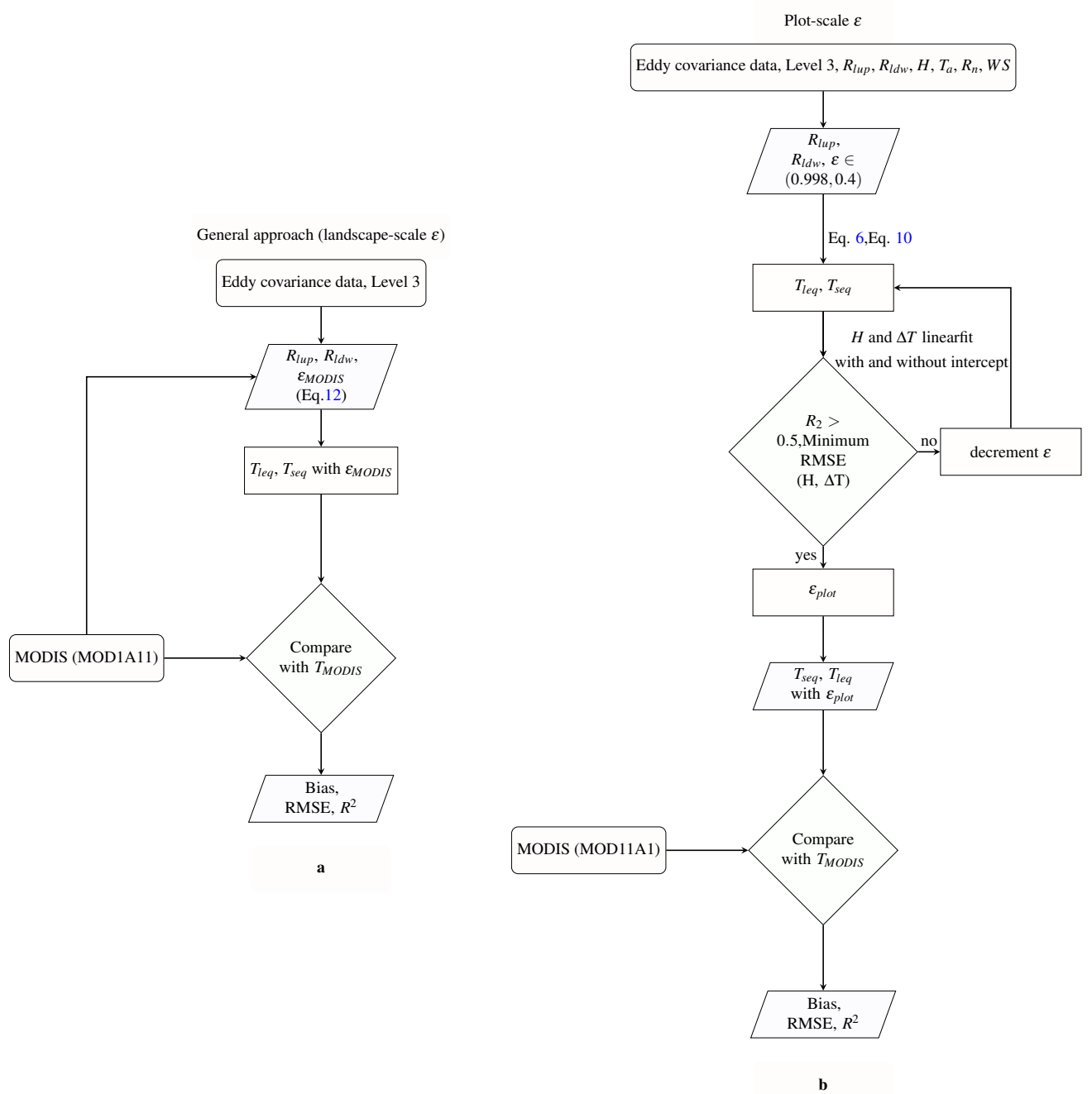
**Uncertainty estimation:** Uncertainty in plot-scale  $\epsilon$  and LST was quantified based on an assumed systematic error (caused by a potential bias in measurement devices) at the study sites. In a first step, based on the literature<sup>24,52</sup> the error bounds of each input variable ( $H$ ,  $R_{lup}$ ,  $R_{ldwn}$ ,  $T_a$ ) used for plot-scale  $\epsilon$  estimation were defined. The error bounds for  $R_{lup}$  and  $R_{ldwn}$  are -5 to 5  $\text{Wm}^{-2}$ <sup>24</sup>, for  $H$ , we used -20 to 20  $\text{Wm}^{-2}$  and for  $T_a$  we used -1 to 1  $\text{K}$ <sup>52</sup>. The error samples (perturbation) within these bounds were generated using the Saltelli sampling scheme (using python based package SALIB<sup>53</sup>). Each error sample is added to the monthly segregated measured fluxes as explained above. Observed fluxes combined with perturbed fluxes are used to estimate  $T_s$  using Eq. (10) and Eq. (6). The obtained range of diurnal  $T_s$  and observed  $T_a$  with perturbation is used to calculate the uncertainty in  $\Delta T$ ; an example for July 15 is shown in Fig. 6c. Perturbed sensible heat flux ( $H$ +sample error) and perturbed  $\Delta T$  is used to obtain  $\epsilon_{plot}$  as described above. The range of monthly  $\epsilon_{plot}$  obtained is reported as uncertainty in monthly  $\epsilon$ .

## 4 Author contributions statement

GT, SJS, KM, and IT designed the study. GT conducted the analysis and prepared the manuscript together with SJS. KM, IT and MS provided textual input, corrections, suggestions and critical discussion of the manuscript.

## 5 Code and data availability

The data and code used for this study is uploaded to zenodo.org (<https://doi.org/10.5281/zenodo.6385016>)



**Figure 7.** Schematic representation of steps followed for plot-scale LST retrieval using eddy covariance measurement (a) Landscape emissivity and longwave measurement and compared to Landscape-scale LST ( $T_{MODIS}$ ) (b) Plot-scale emissivity estimation  $\epsilon_{plot}$  using observed  $H$ ,  $R_{ldwn}$ ,  $R_{lup}$  and plot-scale LST is estimated using  $\epsilon_{plot}$  and  $\epsilon_{MODIS}$  is compared using  $R^2$ , RMSE, Bias are mentioned in Fig. (3).

## 6 Acknowledgements

We would like to thank Dr. Maik Renner for pointing us to the work by Holmes et al. and Dan Yakir's lab for providing Yatir Forest data and helpful discussions. We are also grateful to Michael Roderick, Thomas Foken, Jason Beringer, Lindsay Hutley for insightful discussions and Remko Nijzink for his help in RENKU. This work is supported by the Luxembourg National Research Fund (FNR) ATTRACT programme (WAVE, A16/SR/11254288). Mauro Sulis acknowledges the financial support of the FNR CORE programme (CAPACITY, C19/SR/13652816).



## References

1. Rowell, D. P. A scenario of european climate change for the late twenty-first century: seasonal means and interannual variability. *Clim. Dyn.* **25**, 837–849 (2005).
2. Mallick, K. *et al.* Reintroducing radiometric surface temperature into the penman-monteith formulation. *Water Resour. Res.* **51**, 6214–6243 (2015).
3. Timmermans, W. J., Kustas, W. P., Anderson, M. C. & French, A. N. An intercomparison of the surface energy balance algorithm for land (sebal) and the two-source energy balance (tseb) modeling schemes. *Remote. Sens. Environ.* **108**, 369–384 (2007).
4. Mallick, K. *et al.* A critical evaluation on the role of aerodynamic and canopy–surface conductance parameterization in sebal and svat models for simulating evapotranspiration: A case study in the upper biebrza national park wetland in poland. *Water* **10**, 1753 (2018).
5. Trebs, I. *et al.* The role of aerodynamic resistance in thermal remote sensing-based evapotranspiration models. *Remote. Sens. Environ.* **264**, 112602 (2021).
6. Kustas, W. & Anderson, M. Advances in thermal infrared remote sensing for land surface modeling. *Agric. For. Meteorol.* **149**, 2071–2081 (2009).
7. Still, C. J. *et al.* Imaging canopy temperature: shedding (thermal) light on ecosystem processes. *New Phytol.* **230**, 1746–1753 (2021).
8. Migliavacca, M. *et al.* The three major axes of terrestrial ecosystem function. *Nature* 1–5 (2021).
9. Mallick, K. *et al.* Bridging thermal infrared sensing and physically-based evapotranspiration modeling: From theoretical implementation to validation across an aridity gradient in australian ecosystems. *Water Resour. Res.* **54**, 3409–3435 (2018).
10. Holmes, T., De Jeu, R., Owe, M. & Dolman, A. Land surface temperature from ka band (37 ghz) passive microwave observations. *J. Geophys. Res. Atmospheres* **114** (2009).
11. Maes, W. H., Gentile, P., Verhoest, N. E. & Miralles, D. G. Potential evaporation at eddy-covariance sites across the globe. *Hydrol. Earth Syst. Sci.* **23**, 925–948 (2019).
12. Wang, K. & Dickinson, R. E. Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses. *Rev. Geophys.* **51**, 150–185 (2013).
13. Sun, J. & Mahrt, L. Relationship of surface heat flux to microscale temperature variations: Application to boreas. *Boundary-Layer Meteorol.* **76**, 291–301 (1995).
14. Jacob, D. *et al.* A comprehensive model inter-comparison study investigating the water budget during the baltex-pidcap period. *Meteorol. Atmospheric Phys.* **77**, 19–43 (2001).
15. Mallick, K. *et al.* Canopy-scale biophysical controls of transpiration and evaporation in the amazon basin. *Hydrol. Earth Syst. Sci.* **20**, 4237–4264 (2016).
16. Verma, M. *et al.* Global surface net-radiation at 5 km from modis terra. *Remote. Sens.* **8**, 739 (2016).
17. Lhomme, J.-P., Katerji, N., Perrier, A. & Bertolini, J.-M. Radiative surface temperature and convective flux calculation over crop canopies. *Boundary-Layer Meteorol.* **43**, 383–392 (1988).
18. Norman, J. M. & Becker, F. Terminology in thermal infrared remote sensing of natural surfaces. *Agric. For. Meteorol.* **77**, 153–166 (1995).
19. Holmes, T. R., Hain, C. R., Anderson, M. C. & Crow, W. T. Cloud tolerance of remote-sensing technologies to measure land surface temperature. *Hydrol. Earth Syst. Sci.* **20**, 3263–3275 (2016).
20. Chu, H. *et al.* Representativeness of eddy-covariance flux footprints for areas surrounding ameriflux sites. *Agric. For. Meteorol.* **301**, 108350 (2021).
21. Marcolla, B. & Cescatti, A. Geometry of the hemispherical radiometric footprint over plant canopies. *Theor. Appl. Climatol.* **134**, 981–990 (2018).
22. Morillas, L. *et al.* Using radiometric surface temperature for surface energy flux estimation in mediterranean drylands from a two-source perspective. *Remote. Sens. Environ.* **136**, 234–246 (2013).
23. Wang, K. & Liang, S. Evaluation of aster and modis land surface temperature and emissivity products using long-term surface longwave radiation observations at surfrad sites. *Remote. Sens. Environ.* **113**, 1556–1565 (2009).

24. Trenberth, K. E. & Fasullo, J. T. Tracking earth's energy: From el niño to global warming. *Surv. geophysics* **33**, 413–426 (2012).
25. Crago, R. D. & Qualls, R. J. Use of land surface temperature to estimate surface energy fluxes: Contributions of wilfried brutsaert and collaborators. *Water Resour. Res.* **50**, 3396–3408 (2014).
26. Stephens, G. L. Review of atmospheric radiation: 1991–1994. *Rev. Geophys.* **33**, 785–794 (1995).
27. Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X. & De Gonçalves, L. G. G. A fully multiple-criteria implementation of the sobol' method for parameter sensitivity analysis. *J. Geophys. Res. Atmospheres* **117** (2012).
28. Sugita, M. & Brutsaert, W. Optimal measurement strategy for surface temperature to determine sensible heat flux from anisothermal vegetation. *Water resources research* **32**, 2129–2134 (1996).
29. Snyder, W. C., Wan, Z., Zhang, Y. & Feng, Y.-Z. Classification-based emissivity for land surface temperature measurement from space. *Int. J. Remote. Sens.* **19**, 2753–2774 (1998).
30. Wilson, K. *et al.* Energy balance closure at fluxnet sites. *Agric. For. Meteorol.* **113**, 223–243 (2002).
31. Stoy, P. C. *et al.* A data-driven analysis of energy balance closure across fluxnet research sites: The role of landscape scale heterogeneity. *Agric. forest meteorology* **171**, 137–152 (2013).
32. Twine, T. E. *et al.* Correcting eddy-covariance flux underestimates over a grassland. *Agric. forest meteorology* **103**, 279–300 (2000).
33. Chakraborty, T. *et al.* Biases in model-simulated surface energy fluxes during the indian monsoon onset period. *Boundary-Layer Meteorol.* **170**, 323–348 (2019).
34. Guillevic, P. *et al.* Land surface temperature product validation best practice protocol. version 1.1. *Best Pract. for Satell. Land Prod. Valid.* 60 (2018).
35. Margulis, S. A., Liu, Y. & Baldo, E. A joint landsat-and modis-based reanalysis approach for midlatitude montane seasonal snow characterization. *Front. Earth Sci.* 272 (2019).
36. Eisenhauer, J. G. Regression through the origin. *Teach. statistics* **25**, 76–80 (2003).
37. Jin, M. & Liang, S. An improved land surface emissivity parameter for land surface models using global remote sensing observations. *J. Clim.* **19**, 2867–2881 (2006).
38. Mira, M., Valor, E., Boluda, R., Caselles, V. & Coll, C. Influence of soil water content on the thermal infrared emissivity of bare soils: Implication for land surface temperature determination. *J. Geophys. Res. Earth Surf.* **112** (2007).
39. Chen, C. Determining the leaf emissivity of three crops by infrared thermometry. *Sensors* **15**, 11387–11401 (2015).
40. Vishnevetsky, I., Rotenberg, E., Kribus, A. & Yakir, D. Method for accurate measurement of infrared emissivity for opaque low-reflectance materials. *Appl. optics* **58**, 4599–4609 (2019).
41. Chakraborty, T., Lee, X., Ermida, S. & Zhan, W. On the land emissivity assumption and landsat-derived surface urban heat islands: A global analysis. *Remote. Sens. Environ.* **265**, 112682 (2021).
42. Cullen, N. J., Mölg, T., Kaser, G., Steffen, K. & Hardy, D. R. Energy-balance model validation on the top of kilimanjaro, tanzania, using eddy covariance data. *Annals Glaciol.* **46**, 227–233 (2007).
43. Baldocchi, D. *et al.* Fluxnet: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Am. Meteorol. Soc.* **82**, 2415–2434 (2001).
44. Gilberto, P. *et al.* The fluxnet2015 dataset and the oneflux processing pipeline for eddy covariance data. *Sci. Data* (2020).
45. Wan, Z. Collection-5 modis land surface temperature products users' guide. *ICESSE, Univ. California, Santa Barbar.* (2007).
46. Delucchi, L. pymodis: from satellite to gis maps. In *FOSS4G 2014* (2014).
47. Hulley, G. C., Hughes, C. G. & Hook, S. J. Quantifying uncertainties in land surface temperature and emissivity retrievals from aster and modis thermal infrared data. *J. Geophys. Res. Atmospheres* **117** (2012).
48. Guillevic, P. *et al.* Land surface temperature product validation best practice protocol. version 1.0. *Best Pract. for Satell. Land Prod. Valid.* 60 (2017).
49. Wang, W., Liang, S. & Meyers, T. Validating modis land surface temperature products using long-term nighttime ground measurements. *Remote. Sens. Environ.* **112**, 623–635 (2008).

50. Kustas, W. P., Anderson, M. C., Norman, J. M. & Li, F. Utility of radiometric–aerodynamic temperature relations for heat flux estimation. *Boundary-Layer Meteorol.* **122**, 167–187 (2007).
51. Bahir, M. *et al.* Evaluation and aggregation properties of thermal infra-red-based evapotranspiration algorithms from 100 m to the km scale over a semi-arid irrigated agricultural area. *Remote. Sens.* **9**, 1178 (2017).
52. Foken, T. The energy balance closure problem: an overview. *Ecol. Appl.* **18**, 1351–1367 (2008).
53. Saltelli, A., Albrecht, D., Tarantola, S. & Ferretti, F. A new sample-based algorithms to compute the total sensitivity index. *arXiv preprint arXiv:1703.05799* (2017).