Agricultural and Forest Meteorology 236 (2017) 87–99

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Evaluation of eddy covariance latent heat fluxes with independent lysimeter and sapflow estimates in a Mediterranean savannah ecosystem

Oscar Perez-Priego[a,*], Tarek S. El-Madany[a], Mirco Migliavacca[a], Andrew S. Kowalski[b], Martin Jung[a], Arnaud Carrara[c], Olaf Kolle[a], M. Pilar Martín[d], Javier Pacheco-Labrador[d], Gerardo Moreno[e], Markus Reichstein[a]

[a] Max Planck Institute for Biogeochemistry, Jena, Germany
[b] Departamento de Física Aplicada, Universidad de Granada, Granada, Spain
[c] Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM), Valencia, Spain
[d] Environmental Remote Sensing and Spectroscopy Laboratory (SpecLab), Institute of Economics, Geography and Demography (IEGD-CCHS), Spanish National Research Council (CSIC), Spain
[e] Forest Research Group, Universidad de Extremadura, Plasencia, Spain

ARTICLE INFO

Article history:
Received 5 August 2016
Received in revised form 9 December 2016
Accepted 11 January 2017
Available online 23 January 2017

Keywords:
Evapotranspiration
Eddy covariance
Measurement errors
Lysimeter
Energy balance closure
Random forest
Savannah ecosystem

ABSTRACT

We evaluated the underlying causes of differences between latent heat (LE) fluxes measured with two enclosed-path eddy covariance systems (EC) at two measurement levels and independent estimates in an open oak-tree grass savannah over almost one year. Estimates of LE of the well-stablished underlying grass by replicated weighable tension-controlled lysimiters (LE_{Lys}) provided a robust baseline against which to compare EC LE measured at 1.6 m above ground ($LE_{\_1.6}$). Similarly and at the ecosystem level, LE up-scaled using independent measurements ($LE_{upscaled}$ = sap flow + lysimeter) was benchmarked with 3 EC-derived LE estimates: 1) LE measured by a EC tower at 15 m above ground ($LE_{\_15}$), 2) $LE_{\_15}$ adjusted to close the energy balance by using the Bowen ratio method ($LE_{Bowen} = (R_n - G)/(1 + \beta)$), and 3) LE derived from the energy budget residual ($LE_{residual} = R_n - G - H_{\_15}$). The sensitivity of EC LE to the correction method applied (i.e. corrections for low-pass filtering effects on water vapor fluctuations and the so-called angle-of-attack correction) and its impact on the energy balance closure (EBC) were also evaluated.

Comparison of EC LE between 1.6 m- and 15 m-heights showed that grass dominated annual evaporative loss from 69 to 87% depending upon the spectral correction method applied. Results revealed substantial underestimation of $LE_{\_1.6}$ (up to 35%) compared to LE_{Lys}, which mostly occurred during the growing season. However those differences were remarkably lower when likening $LE_{\_15}$ versus $LE_{upscaled}$ (14%) suggesting that the dampening of the water vapor fluctuations due to low-pass filtering effects is more pronounced near the surface. Interestingly, a diagnostic evaluation of the errors with a random forest model showed that differences followed quite structured patterns and were associated with certain atmospheric conditions: turbulent mixing deficiencies and or stable atmospheric stratification. In addition, the model showed that differences increased with increasing relative humidity (RH) and soil moisture. Our results revealed that the degree of EBC is highly sensitive to the flux correction method applied, in particular when correcting for flow distortion effects. Typically, turbulent fluxes fell below the measured available energy (slope 0.92) but the slope switched abruptly when the angle-of-attack correction was applied (slope 1.07). Consistent with the EBC, independent LE estimates matched well with LE_{Bowen} and the EBC gap decreased when $LE_{upscaled}$ was used (slope 0.96). The use of independent estimates of LE together with machine learning methods are proposed as a powerful means to diagnose the complexity behind LE errors and give insights into the energy imbalance problem. In addition to inherent randomness of EC LE data, accounting for uncertainties associated with the appropriateness of the correction method applied is highly recommended.

* Corresponding author at: Biosphere-Atmosphere Interactions and Experimentation Group, Biogeochemical Integration Department, Max Planck Institute for Biogeochemistry, Jena, Germany.
E-mail addresses: opriego@bgc-jena.mpg.de, oscarperezpriego@gmail.com (O. Perez-Priego).
http://dx.doi.org/10.1016/j.agrformet.2017.01.009
0168-1923/

Nomenclature

A_{trunk}	The average cross sectional area of sapwood for trees within the half-hourly determined tower footprint
EC	Eddy covariance
ϵ	The relative error in EC LE, error estimated as indicated below: $\epsilon_{1.6\_M}$ calculated as (LE_{Lys} – $LE_{\_1.6\_M}$)/LE_{Lys}; $\epsilon_{1.6\_F}$ calculated as (LE_{Lys} – $LE_{\_1.6\_F}$)/LE_{Lys}; ϵ_{15_M} calculated as ($LE_{upscaled}$ – $LE_{\_15\_M}$)/$LE_{upscaled}$; ϵ_{15_I} calculated as ($LE_{upscaled}$ –$LE_{\_15\_I}$)/$LE_{upscaled}$
$H_{\_15}$	The sensible heat flux measured by the EC tower 15 m above ground
LE_{Bowen}	Estimated latent heat as $R_n - G/(1+\beta)$, where β is the Bowen ratio calculated as $H_{\_15}/LE_{\_15\_M}$
$LE_{\_1.6}$	Latent heat measured by the 1.6 m EC tower: $LE_{\_1.6\_M}$, including the spectral correction of Moncrieff et al. (1997); $LE_{\_1.6\_F}$, including the spectral correction of Fratini et al. (2012); $LE_{\_1.6\_filter}$, gap-filled $LE_{\_1.6\_F}$ when $\epsilon_{1.6\text{-}F} > 30\%$
$LE_{\_15}$	Latent heat measured by the 15 m EC tower: $LE_{\_15\text{-}M}$ including the spectral correction of Moncrieff et al. (1997); $LE_{\_15\text{-}I}$, including the spectral correction of Ibrom et al. (2007); $LE_{\_15\text{-}I\_N}$, including the spectral correction of Ibrom et al. (2007) and angle-of-attack correction (Nakai et al., 2006); $LE_{\_15\text{-}I\_filter}$, gap-filled $LE_{\_15\text{-}I}$ when $\epsilon_{15\_F} > 30\%$
LE_{Lys}	Latent heat exchange by the understory layer from replicated lysimeters (n = 6)
$LE_{residuals}$	Latent heat exchange determined from the energy budget residual ($R_n - G - H_{\_15}$)
LE_{RF}	Modelled latent heat exchange via RF
LE_{sap}	Estimated stand transpiration and expressed in energy terms ($W\,m^{-2}$)
$LE_{upscaled}$	Estimated LE via aggregation of LE_{Lys} and LE_{sap}
$LE_{ensemble}$	Estimates computed with an ensemble of $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residuals}$, LE_{Bowen}
RF	The random forest model
RH	Air relative humidity
θ_s	Scaled soil water content at 10 cm depth ($\theta_s = SWC/SWC_{max}$)
$Td_{footprint}$	Is the tree density in the footprint area
u.	The friction velocity
z/L	Monin–Obukhov dimensionless stability parameter

1. Introduction

The main causes for the observed lack of energy surface balance closure – the mismatch between turbulent energy fluxes (latent heat, LE, and sensible heat, H) and the available energy (net radiation, (R_n) ground heat flux (G) and changes in heat storage) – are currently under debate (e.g. Foken et al., 2011). This inconsistency casts doubt on the accuracy of eddy covariance (EC) data and further evaluation of measurement errors using independent methods is highly desirable (Mamadou et al., 2016; Soubie et al., 2016). Among micrometeorological methods, EC is widely used in global long-term observation networks such as FLUXNET (Baldocchi et al., 2001), and provides measurements of both LE and H, as well as other trace gas fluxes (e.g. CO_2, and CH_4) over plant canopies. During recent decades, many studies have analyzed the observed gap in the energy balance closure (EBC) across sites with contrasted characteristics and environmental conditions (Barr et al., 2006; Foken et al., 2010; Franssen et al., 2010; Lee and Black, 1993; Moderow et al., 2009; Oncley et al., 2007; Stoy et al., 2006; Wilson et al., 2002). Such studies have shown that the sum of EC-derived H and LE systematically falls below the measured available energy.

Reasonable efforts have been made to study the underlying errors of turbulent fluxes that might explain such inconsistencies (Foken et al., 2011). On one hand, mismatches between radiometric and turbulent flux footprints or errors in available energy estimates have been shown to be minor (<10%) compared to the widely observed gap (10–30%; Stoy et al., 2013) and cannot fully explain the lack of EBC (Foken et al., 2010; Twine et al., 2000). Therefore, factors related to the processing steps of EC flux calculation and corrections, turbulence statistics, atmospheric stability, storage, advection and other issues related to instrumental set up, and site characteristics have been identified as possible causes of bias in EC data (Aubinet et al., 2000; Finnigan et al., 2003; Foken et al., 2006; Horst et al., 2015; Leuning et al., 2012; Mamadou et al., 2016; Mauder et al., 2010; Mauder and Foken, 2005; Van Der Molen et al., 2004). For example, the lack of EBC has been shown to be modest under highly turbulent conditions, but increases markedly when turbulence is limited (Amiro, 2009; Barr et al., 2006; Chávez et al., 2009; Franssen et al., 2010; Oliphant et al., 2004; Stoy et al., 2013; Wilson et al., 2002). This suggests that the role of the friction velocity (u_*) or stability parameters should not be only considered for screening nocturnal CO_2 turbulent exchange errors but also for EC-derived H and LE as well (Stoy et al., 2006). Further causes of bias in EC data have been associated with measurement errors of vertical wind velocity due to flow distortion effects by non-orthogonal anemometer types (e.g. Gill type), which might result in inaccurate measurements of H and LE (Van Der Molen et al., 2004). Also, it has been shown that discrepancies in the EBC are reduced by applying the so-called angle-of-attack correction (Frank et al., 2013; Horst et al., 2015; Kochendorfer et al., 2012; Nakai et al., 2006). However, some concerns have been raised regarding this correction and it remains a matter of debate (Mauder, 2013; Kochendorfer et al., 2013). These facts, among others (e.g, the storage term; Leuning et al., 2012), highlight important uncertainties, particularly when EC data are widely used to evaluate or parameterize terrestrial biosphere or hydrological models, or to derive ecosystem functional properties such as water use efficiency or evaporative fraction, that assume complete EBC (Jaeger and Kessler, 1997). One open question is the degree to which the energy balance residual can be i) equally attributed to measurement errors in LE and H, or mostly assigned to either ii) LE or iii) H (e.g. Wohlfahrt et al., 2009).

Presumably, option ii) might prevail if we consider the loss of high frequency eddies, which may cause underestimation of LE of up to 10% of the annual values (Wilson et al., 2000). Attenuation of water vapor fluctuations is well recognized in closed path EC systems (Fratini et al., 2012; Ibrom et al., 2007; Mammarella et al., 2009; Massman, 2000; Runkle et al., 2012) and can reduce the EBC by up to 19% (Su et al., 2004). Spectral analysis provides a means to check the quality of EC LE (Baldocchi and Meyers, 1991). Different analytical and empirical spectral correction methods have been proposed to correct for the attenuation of the true water vapor flux (Fratini et al., 2012; Horst, 1997; Ibrom et al., 2007; Moncrieff et al., 1997). Although the strengths and weakness among methods have been discussed (Massman and Lee, 2002), we lack quantitative comparison among methods (Su et al., 2004). Whilst all correction methods are prone to biases (Massman and Lee, 2002), inconsistencies in EC LE are often observed when comparing with independent approaches or models (Allen et al., 2011; Meiresonne et al., 2003; Twine et al., 2000; Wohlfahrt et al., 2010), and derivations of LE from the energy balance equation via either the residual ($LE_{residual} = R_n\text{-}G\text{-}H$) or the Bowen ratio [$LE_{Bowen} = R_n - G/(1+\beta)$] have been applied (Amiro, 2009; Castellví and Snyder, 2010; Chávez et al., 2009; Falge et al., 2005; Garratt, 1984; Twine et al., 2000;

Wilson et al., 2002; Wohlfahrt et al., 2010). Obviously, the reliability of $LE_{residual}$ is based on the supposition that biases in H are minor compared to those in LE. This agrees with the theoretical consideration of most of the spectral correction methods, which use the heat spectrum as a reference to correct for the attenuation of the water vapor signal (Fratini et al., 2012). Here we focus on bias in EC LE, which can be quantitatively evaluated by direct comparison with independent and non-micrometeorological approaches.

Physiological and hydrological approaches allow for independent assessment of EC LE. Among physiological approaches (Kool et al., 2014), sap-flow measurements are most often used in combination with EC (Berbigier et al., 1996; Cammalleri et al., 2013; Hogg et al., 1997; Paco et al., 2009; Wilson et al., 2001). However, comparing results is not straightforward due to temporal and spatial resolution mismatches among the techniques. Moreover, in complex ecosystems such as tree-grass savannah, additional challenges arise from i) difficulties in scaling up sap-velocities from single point-measurements to tree stands and their representativeness within the EC flux "footprint" (Oishi et al., 2008; Wilson et al., 2000), and ii) the difficulty in measuring LE and its relative contribution by understory EC (Baldocchi and Meyers, 1991). Due to the latest developments in precision lysimetry, weighable lysimeters have been used as a benchmark for LE methods, including EC (Castellví and Snyder, 2010; Chávez et al., 2009; Evett et al., 2012; Holmes, 1984). However, comparisons should be handled carefully, particularly for heterogeneous surfaces due to the fact that lysimetric measurements of LE usually embrace only a minor portion of EC flux "footprint".

Here we will evaluate the consistency of ecosystem and understory LE by EC in a typical sparse oak-tree grass Mediterranean savannah ecosystem, by synchronous, combined measurements via classical EC, understory EC, sap-flow, replicated lysimeters, and modelling approaches based both on Random Forests and LE estimates derived from the energy balance equation. More specifically, we aim to:

a) Quantify measurement errors in EC LE and its components via a multi-technique approach. Half-hourly LE from the understory grass using six lysimeters (LE_{Lys}) will be used to evaluate LE measured by an understory EC tower installed at 1.6 m above the ground ($LE_{\_1.6}$). Simultaneously and at the ecosystem scale, the degree of agreement between total LE measured by 15 m-high EC tower ($LE_{\_15}$) will be also evaluated with independent estimates via lysimeter and sap flow ($LE_{upscaled}$) along with EC-derived LE that adjusts LE to be consistent with the EBC.
b) Evaluate underlying causes of errors by using machine-learning methods (i.e. Random Forest) as a more comprehensive means for identifying patterns and drivers that yield discrepancies between EC LE and independent methods.
c) Evaluate the sensitivity of EC LE to the method of correction for i) spectral loss (Moncrieff et al., 1997; Fratini et al., 2012; Ibrom et al., 2007) and ii) angle-of-attack (Nakai et al., 2006).

2. Material and methods

2.1. Site description

This study was carried out in a Mediterranean tree-grass savannah in Spain (39°56′ 24.68" N, 5°45′50.27" W; Majadas de Tietar, Caceres) for almost one year (2015). The site is characterized by a mean annual temperature of 16 °C, mean annual precipitation of ca.700 mm, falling mostly from November until May, and by a prolonged dry summer. Similar to most Mediterranean savannahs, low-intensity grazing (<0.3 cows ha^{-1}) is the main land use at the site. The site is characterized by low density oak tree cover (mostly *Quercus Ilex* (L.), ~22 trees ha^{-1}) and dominated by a grass stratum. The diameter at the breast high (DBH) of the sampled trees ranges from 30 to 75 cm (mean DBH = 46.86 cm). The herbaceous stratum is composed by native annual species of the three main functional plant forms (grasses, forbs and legumes), whose fractional cover varies seasonally according to their phenological status (Perez-Priego et al., 2015), with important inter-annual variations related to the onset of the dry period. Overall, leaf area measurements of the herbaceous stratum characterized the growing season phenology as peaking early in May with mean plant area index values at the peak of the season of 2 m^2 m^{-2} and achieving senescence by the end of May (plant area index ~0.4 m^2 m^{-2}, Perez-Priego et al., 2015). The soil is classified as Abruptic Luvisol (IUSS Working Group WRB, 2015) and originates from Pliocene–Miocene alluvial deposits. The upper limit of the clay horizon is found at a depth between 30 and 100 cm. The texture in the upper horizons is sandy (80% sand, 9% clay, 11% silt).

2.2. Instrumentation

2.2.1. Eddy covariance

The equipment to measure ecosystem scale fluxes of $LE_{\_15}$, $H_{\_15}$ and R_n at 15 m above ground consisted of a sonic anemometer (SA-Gill R3-50; Gill Instruments Limited, Lymington, UK), enclosed-path infrared gas analyzer (LI-7200, LI-COR Biosciences Inc., Lincoln, NE, USA) and net radiometer (CNR4, Kipp&Zonen, Delft, the Netherlands). Measurements took place from mid-February to end December 2015. The CNR4 incorporates an optimal integrated ventilation unit to guarantee efficient air-flow and minimize unwanted disturbances on both long-wave and short-wave measurements due to dew or dust deposition on the domes and windows. About 10 m north of the EC tower, an EC tower was installed with the same ultrasonic anemometer and gas analyzer models at 1.6 m above ground in an open space to measure turbulent fluxes of $LE_{\_1.6}$ and $H_{\_1.6}$ from the well-established understory grass. EC raw data – including the three dimensional wind velocities (u,v,w in m/s), sonic temperature (K), and dry CO_2/H_2O mixing ratios – were collected at 20 Hz. Gases were sampled through 1 m long sampling tubes with 4.5 mm of inner diameter with a nominal flow rate of 15 L min^{-1}. The intake lines were sheltered with an insulation system plus a heater to avoid water condensation. Sonic temperature was internally corrected for cross wind sensitivity and data were processed with EddyPro software (version 5.2.0, LI-COR Biosciences Inc., Lincoln, NE, USA) according to the following settings:

1) Spike removal according to Vickers and Mahrt (1997) with linear interpolation for the detected spikes, mean was removed from time series, time lags between time series were determined for 10% relative humidity intervals via an automatic time lag optimization procedure. Coordinate rotation for tilt correction was applied according to the planar fit method proposed by (Wilczak et al., 2001),
2) Both $LE_{\_1.6}$ and $LE_{\_15}$ were determined by applying two different spectral correction methods:

i) the analytical cospectra correction method proposed by Moncrieff for low and high-pass filtering effect (Moncrieff et al., 2004; Moncrieff et al., 1997). For distinction referred to as $LE_{\_1.6\_M}$ and $LE_{\_15\_M}$ hereafter.
ii) an *in-situ* correction that considers the dependency of attenuation of water vapor fluctuations on relative humidity (Fratini et al., 2012; Ibrom et al., 2007). Shortly, those methods correct for the attenuation of water vapor fluctuations by the tendency of water vapor to adsorb to the inside of the sampling tube (low-

pass filtering effects). Considering the suitability of the method depending on surface roughness and following recommendations, the correction of Fratini et al. (2012) was applied to correct $LE_{\_1.6}$– considered a smooth surface – and that of Ibrom et al. (2007) was applied to $LE_{\_15}$ due to suitability of the latter over rough surfaces. For distinction hereafter, we will refer these to as $LE_{\_1.6\_F}$ and $LE_{\_15\_I}$, respectively. According to Fratini et al. (2012), high quality spectra were achieved by not considering those spectra with LE or H fluxes lower than 20 W m^{-2} in the ensemble averaging period. In addition, we considered fluxes as low-quality data when spectral correction factors were higher than 3 or lower than −3.

Quality flags of the calculated fluxes were assigned according to Mauder and Foken (2005) and Vickers and Mahrt (1997). For further analysis, data flagged as 2 or measured during rain events were rejected and removed from the dataset. In addition, EC data from critical sectors with probable flow distortion (150°–190°) – identified with the angle of attack – were also discarded from analysis. Turbulent parameters such as Monin-Obukhov stability (z/L), wind friction velocity (u_*) were also calculated. Data with low turbulent mixing were identified and discarded from analysis according to a standard u_* filtering criterion. Night-time CO_2 data were classified in different season/temperature subclasses and the u_* thresholds calculated using the moving point test method (Papale et al., 2006). The procedure was applied separately at both towers and u_* thresholds of 0.074 and 0.14 m s^{-1} were used as the best estimates for filtering low-quality $LE_{\_1.6}$ and $LE_{\_15}$ data, respectively.

The LE storage term was estimated from the differences in the vertical integral of the dry molar water vapor fraction over the averaging time (Finnigan 2006; Kowalski, 2008). Profile integration was estimated from the spline fit of the dry molar water vapor fractions measured at 7 measurement levels (0.1, 0.5, 1, 2, 5, 9, 15 m), which accurately characterized the whole 15 m-height profile. The temperature profile was inferred from temperatures measured at 15 and 2 m heights, while no changes in pressure was assumed over the profile. We used thetypical profiling system, which consisted of a pump circulating air at 15 L min^{-1} through a sample circuit spatially distributed over the profile and flushed a residual flow of 1 L min^{-1} through an infrared gas analyzer (IRGA LI-840, Lincoln, NE, USA) to measure the water vapor mole fraction at 1 Hz The missing LE flux by storage was measured in two nearby towers and a mean of the two estimates was used.

Eight replicated heat-flux plates (Heatflux Ramco HP3, McVan Instruments, Mulgrave, Victoria, Australia) were installed – four under trees and four under sky – at 2 cm depth. Soil temperature and moisture probes (Delta-ML3, Delta-T Devices Ltd, Burwell Cambridge, UK) were installed near the heat-flux plates at 5, 10, 20, 30, 50, and 100 cm depth. The soil heat flux (G) was calculated as a weighted mean of the 8 sensors according to the 11% canopy cover. Each EC tower was equipped with a hygro-thermometer (Thies Clima, Göttingen, Germany) in a ventilated shelter to measure air temperature and humidity at appropriate height.

2.2.2. Lower boundary layer-controlled weighing-lysimeters

Three hermetically sealed polyethylene-high-density container stations were installed underground within the fetch of the tall EC tower. Each station contained 2 weighable lysimeters of the temperature- and tension-controlled type. An intact soil monolith of around 2500 kg was packed into each lysimeter vessel using a novel lysimeter soil retriever technique (Reth et al., 2007). This original lysimeter type includes a lower-boundary controlled system and provides reliable estimates of LE with high-precision while minimizing soil disturbances. The lower-boundary controlled system enables control of both soil matrix tension and temperature to mirror the actual values in the surrounding soil. Briefly, a series of porous ceramic bars connect the capillary system of the soil at the bottom of the lysimeter to a pressure-regulated, airtight water tank. The pressure inside the tank is adjusted according to the current soil matrix tension measured by a reference tensiometer (Tensio 160, VKWA 100 Tipping counter, Umwelt-Geräte-Technik GmbH, Müncheberg, Germany), installed horizontally at 1 m depth in the surrounding soil beside each station. For comparison, another one was installed inside each lysimeter at the same depth. In addition, a heat exchanger system maintained the temperature of the bottom layer of the soil column in the lysimeter equal to that of the surrounding soil. Each lysimeter vessel was made of a stainless steel cylinder with cross sectional area of 1.0 m^2 and 1.2 m depth and covered with a temperature-insulated sheet to keep natural tension and temperature gradients along the soil monolith profile. Soil moisture and temperature (UMP-1, Umwelt-Geräte-Technik GmbH, Müncheberg, Germany) at 10, 30, 75, and 100 cm were measured inside each lysimeter.

Every lysimeter rests on a weighing system –load triangle– consisting of three precision shear-stress cells (Model 3510, Stainless Steel Shear Beam Load Cell, VPG Transducers, Heilbronn, Germany) mounted on a stainless steel supporter. The cells were previously calibrated with known mass and were able to detect small changes in the lysimeter weight of up to 10 g m^{-2} ground surface (0.01 mm). Water seepage was measured with a tipping counter (VKWA 100 Tipping counter, Umwelt-Geräte-Technik GmbH, Müncheberg, Germany). The lysimeter weight was registered every 1 min and recorded as a 15 min mean value. Once correcting for changes in weight via the LBC and water seepage periods and periods with precipitation excluded, evaporative loss (expressed as positive values) or mass gain by dew fall (expressed as negative values) was calculated then as the differences in mass from the previous time interval (15 min). For comparison with eddy covariance system, evapotranspiration rates were averaged to 30 min and converted to latent heat (hereafter referred to as LE_{Lys}, W m^{-2}).

An automatic quality control procedure based on examination of half-hourly data in time and space was developed to detect anomalies in the lysimeter data. To detect outliers in the time series of LE_{Lys}, smoothing spline functions were applied to half-hourly time series over three-day intervals. The residuals of the smoothing splines and their standard deviations (σ) were computed. Residuals larger than 3σ were flagged as highly questionable data (flag = 1). A second filter criterion used the median absolute deviation of half-hourly LE_{Lys} of unflagged data from the 6 lysimeters (MAD_{Lys}). Accordingly, data for which $|LE_{Lys}\text{-median}(LE_{Lys})|/MAD_{Lys} > 2$ were also considered as outliers and flagged as 2. Finally, the ensemble of half-hourly LE from the 6 lysimeters (neither flagged as 1 nor 2) were used to calculate the median and the standard deviation, which were considered as an integrative quantity of understory LE_{Lys} and its uncertainty.

2.2.3. Sap flow

Tree transpiration was estimated by the heat ratio method (SFM1 Sap Flow Meter, ICT International) in six trees within the footprint of the tall EC tower. The trees were selected in accordance with stem diameter at breast height (DBH) to cover the range of the distribution found within the footprint of the EC tower. The heat ratio method has been proven to measure low and reverse sap flow rates in a wide range of species, dimensions and environments (Burgess et al., 2001). The sensor consisted of two temperature probes and a heater (35 mm long and 1.3 mm in diameter). The temperature probes were inserted equidistant (0.5 cm) from the heater. For robust estimation of radial sap velocity profile, each temperature probe contains two thermocouple junctions to sample two points dispersed through the sapwood (7.5 mm and 22.5 mm from the needle tip). Heat pulses at the two sampling points, powered with 20 J, were set at half-hourly intervals and sap velocities

were computed from the ratio of the temperature rise of the upper to lower sensor. This ratio was computed as the mean of 14 ratios measured between 60 and 100 s after each heat pulse.

In post-processing, corrections for probe misalignment – when applicable – and wounding effects, sap velocity determination and sap flow conversion were applied according to Burgess et al. (2001). To integrate the sap velocity profile, we assumed constant sap velocity within each annulus and sap flow density was calculated as the sap velocity (v_j) weighted by the cross-sectional area of sapwood (A_j) most closely associated with the two point temperature measurements. Total sapwood area was derived from an empirical relationship between tree sapwood area ($A = \Sigma A_j$) and DBH determined by destructive sampling and using a footprint model (Section S1.1 of the supplementary material). The DBH of all trees (273 individuals) inside the flux footprint area of the EC system at 15 m above ground were sampled and georeferenced. Stand tree transpiration was then estimated as:

$$LE_{sap} = \frac{\sum_{i}^{n} \frac{\sum_{j}^{m} v_j \times A_j}{\sum_{j}^{m} A_j}}{n} \times A_{trunk} \times Td_{footprint} \quad (1)$$

where A_{trunk} represents the average cross-sectional area of sapwood for trees within the half-hourly determined tower footprint and $Td_{footprint}$ is the tree density (more details on the calculation of A_{trunk} and $Td_{footprint}$ in S1.2 of the supplementary material). Similar to lysimeter data, sap flow (kg m^{-2} h^{-1}) was converted to latent heat (LE_{sap}, W m^{-2}).

2.3. Diagnostic evaluation of EC LE errors using random forest model

An RF algorithm for regression, implemented in the R package *randomForest* (Liaw and Wiener, 2002) was used to investigate underlying causes of observed errors in both EC $LE_{\_15}$ and $LE_{\_1.6}$. The relative error in LE (ϵ) was defined as:

$$\varepsilon = \frac{[LE] - LE_{eddy}}{[LE]}, \quad (2)$$

where [LE] refers to LE_{Lys} or $LE_{upscaled}$ when comparing either $LE_{\_1.6}$ or $LE_{\_15}$, respectively. More specifically, we trained two RF models at the respective heights with

i $\epsilon_{\_1.6}$ calculated as $(LE_{Lys} - LE_{\_1.6})/LE_{Lys}$ and
ii $\epsilon_{\_15}$ as $(LE_{upscaled} - LE_{\_15})/LE_{upscaled}$.

Note that $LE_{upscaled}$ is the sum of LE_{Lys} and LE_{sap} estimated independently. Here, direct evaporation by tree interception is neglected due to the low tree canopy cover. For training, we used their corresponding predictors (except for soil moisture, which was obviously shared). Explicitly, scaled soil water content ($\theta_s = SWC/SWC_{max}$), air relative humidity (RH), stability parameter (z/L) and friction velocity (u_*) were used as main explanatory variables.

$$\varepsilon = f(u_*, z/L, RH, \theta_s). \quad (3)$$

We evaluated the variable importance of each run via a permutation-based test. The variable importance quantifies the influence of a given variable in the model performance. This variable is expressed in percentage and is calculated as the difference of the error prediction of the model (% IncMSE) before and after permutation, then divided by its mean standard error (Liaw and Wiener, 2002). The marginal effect of each explanatory variable on the average response of the model was also evaluated in detail via the shape of the partial dependence plot (S2 of the supplementary material).

In the evaluation, we made additional analyses to assess the impact of using different spectral corrections on both $\epsilon_{\_1.6}$ and $\epsilon_{\_15}$. Accordingly, we trained four model versions at the two measurement levels:

i) $\epsilon_{1.6\_M}$ calculated as $(LE_{Lys} - LE_{\_1.6\_M})/LE_{Lys}$ was used to train the "$\epsilon_{1.6\_M}$ RF model", ii) $\epsilon_{1.6\_F} = (LE_{Lys} - LE_{\_1.6\_F})/LE_{Lys}$, which explicitly accounts for low-pass filtering effects (Fratini et al., 2012), for the "$\epsilon_{1.6\_F}$ RF model".

iii and iv) Likewise, "ϵ_{15_M} RF model" was trained with ϵ_{15_M} (Moncrieff et al., 1997) calculated as $(LE_{upscaled} - LE_{\_15\_M})/LE_{upscaled}$ and "ϵ_{15_I} RF model" with $\epsilon_{15\_I} = (LE_{upscaled} - LE_{\_15\_I})/LE_{upscaled}$, which also explicitly accounts for the effect of the damping of the water vapor fluctuations (Ibrom et al., 2007).

2.4. Evaluating correction methods

Short- and long-term variations among EC-derived estimates of LE were evaluated against independent methods at their respective measurement level. In order to evaluate the impact of different estimates of LE on the annual and seasonal budget, gaps due to missing or bad-quality data (i.e. quality check, precipitation; 11% of the data) were filled using the MDS gap filling algorithm (Reichstein et al., 2005) implemented in the R package *REddyProc* (http://r-forge.r-project.org/projects/reddyproc/). The algorithm uses either the look-up table method or the simple interpolation technique of mean diurnal variation depending on the size of the gap (Reichstein et al., 2005). We also investigated whether the mismatch between EC-derived LE and independent methods is minimized over the long term by using the MDS gap filling algorithm. Specifically, we filtered EC LE below an acceptable threshold in terms of the discrepancy with independent methods ($\epsilon_{1.6\_F}$ or $\epsilon_{15\_I} < 30\%$), the rest was classified as *"bad-quality"* data and gap-filled according to Reichstein et al. (2005). Hereafter, those estimates are referred to as $LE_{\_1.6\_F\_filter}$ and $LE_{\_15\_I\_filter}$, respectively for both measurement levels.

Finally, we compared the results with two of the most widely used methods to correct EC LE for the inconsistency with the lack of EBC: i) the residual method, and ii) the Bowen ratio method. While the residual method attributes the residual error of the energy balance equation entirely to LE ($LE_{residual} = R_n - G - H_{\_15}$), the Bowen ratio method (Twine et al., 2000) relies on the use of the Bowen ratio (β) defined as $H_{\_15}/LE_{\_15}$ and corrects LE as $LE_{Bowen} = (R_n - G)/(1 + \beta)$. In addition and instead of correcting LE for the lack of EBC, we used the RF model. This approach enabled us to better characterize the propagation of errors in LE by explicitly accounting for the underlying environmental drivers (Eq. (3)). Alternatively to the relative error (ϵ) in Eq. (3), we targeted directly on $LE_{upscaled}$ ("*LE RF model*") and $LE_{\_15\_M}$ was included as an additional predictor instead. In addition, we evaluated the sensitivity of the degree of EBC when flow distortion effects on LE and H were accounted for by applying the angle-of-attack correction as proposed by Nakai et al. (2006). This analysis was also performed with those estimates computed with the ensemble of corrections methods applied (i.e. $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residulas}$, LE_{Bowen}).

All calculations and statistical analyses were computed using the R statistics toolbox (R Development Core Team, 2010).

2.5. Uncertainty estimations

When only single point measurements were available such EC LE, uncertainties in half-hourly data were estimated using the MDS gap-filling algorithm, which assign errors (σ_i) based on the look-up table approach (Reichstein et al., 2005). Briefly, the method

estimates uncertainties via the variance of a set of observed LE with similar meteorological and phenological conditions. Assuming independence, uncertainties were propagated and estimated in annual sum as:

$$\sigma_{Eddy} = \sqrt{\sum_{1}^{n} \sigma_i^2} \quad (4)$$

where n is the number of observations.

On the other hand, when a number of replicates are available as in the case of the independent estimates (i.e. LE_{Lys}, LE_{sap}), σ_i was computed based on the variance of the number of samples. Accordingly, uncertainties of $LE_{upscaled}$ were calculated by weighting for the dominance of the flux as:

$$\sigma_{upscaled} = \sqrt{\frac{LE_{Lys}}{LE_{Lys+}LE_{sap}} \sum_{1}^{n} \sigma(LE_{Lys})_i^2 + \frac{LE_{sap}}{LE_{Lys+}LE_{sap}} \sum_{1}^{n} \sigma(LE_{sap})_i^2} \quad (5)$$

Notice that EC and independent uncertainties are not fully comparable by definition, being the first related the random error and the latter to the spatial variability in LE_{Lys} and LE_{sap} estimates.

3. Results and discussion

3.1. Seasonal dynamics of LE, surface energy components and relationship with environmental variables

In this Mediterranean climate, net radiation (R_n) and ground heat fluxes (G) both increased over the growing period in spring, peaked around June and decreased until reaching their lowest values late in the fall (Fig. 1a). In 2015, the seasonal time course of $LE_{\_15}$ was highly dynamic: $LE_{\_15}$ peaked during the growing period (Fig. 1b), and despite high evaporative demand it decreased later on over the dry summer (but see second peak around June after a rare summer rain, Fig. 1c). This seasonal pattern was largely modulated by the scaled volumetric soil water content (θ_s) of the upper 10 cm (Fig. 1d). While the highest $LE_{\_15}$ values were found in spring, the peak of $H_{\_15}$ occurred during summer (Fig. 1b). $LE_{\_15}$ tended to decrease as long as the top soil layer was drying, (see the drop in both $LE_{\_15}$ just after May, Fig. 1b). During this period, most of the available energy was partitioned into $H_{\_15}$ (>70%, Fig. 1a, b). Similar patterns and magnitude of $LE_{\_1.6}$ as measured by understory EC tower and LE_{Lys} were observed (Fig. 2).

3.2. Comparing EC LE at 1.6 m above ground with replicated lysimeters

In general, the weekly average of LE_{Lys} was higher than $LE_{\_1.6}$, independently of the spectral correction applied (Fig. 2a). As a result, the annual sum of $LE_{\_1.6\_M}$ was 35% lower than that measured by lysimeters (Table 1). This difference substantially decreased when high-frequency losses in the water vapor fluctuations were accounted for by applying in situ spectral loss correction methods (25%). Interestingly, that difference was even smaller with $LE_{\_1.6\_F\_filter}$ (12%), when $LE_{\_1.6\_F}$ data with $\epsilon_{1.6\_F} >$ 30% was filtered and gap filled. Although with certain fluctuations, the relative errors had a tendency to be high – around 40% under wet conditions (around March or July after the rain event, Fig. 2b) and minor over the very dry period (around September). Consistently, $\epsilon_{1.6\_F}$ was smaller than $\epsilon_{1.6\_M}$ and suggests that correcting for low-pass filtering effects considerably reduces the mismatch between EC LE and lysimeter. Overall, substantial scatter in both $\epsilon_{1.6}$ was observed during the fall, probably due the large numbers of rain events during this period.

At the half-hourly time scale, discrepancies between LE_{Lys} and $LE_{\_1.6\_F}$ were observed under certain conditions. Illustrative, diurnal time courses of half-hourly LE_{Lys} and $LE_{\_1.6\_F}$ are presented for

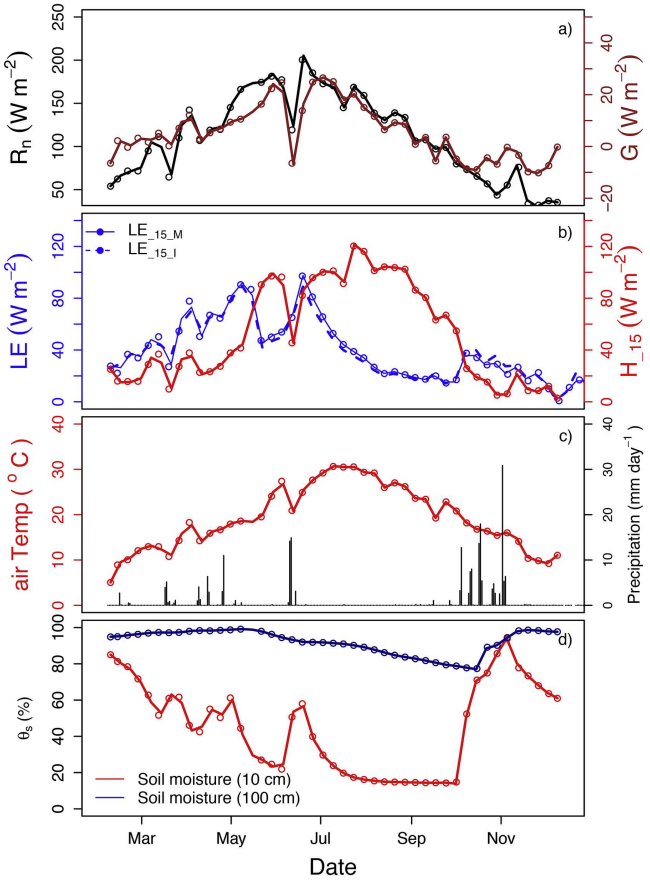

Fig. 1. Seasonal variation of energy balance components and environmental variables. Data represent weekly means and are presented from mid-February to end December 2015.

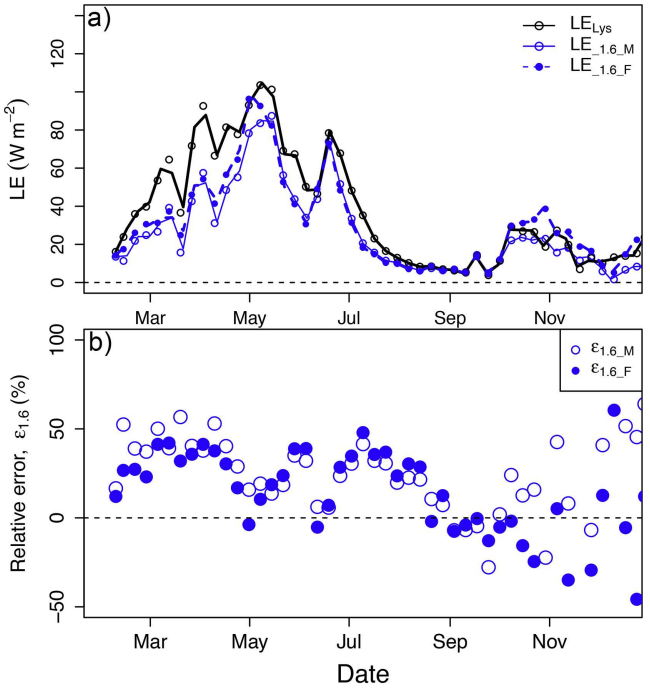

Fig. 2. (a) Seasonal time course of weekly mean understory latent heat (LE) fluxes by eddy covariance (EC) and lysimeters. Differences between EC LE using analytical versus in situ spectral loss correction methods proposed by Moncrieff et al. (1997) and Fratini et al. (2012), respectively are presented. (b) Seasonal trend of relative difference (ϵ, %) between LE estimates by EC and using lysimiters. Data are presented from mid-February to end December 2015.

Table 1
Annual sum of LE measured by EC and expressed as evapotranspiration (ET, mm year^{-1}) at both 15 m-height (ET $_{15\_M}$, ET $_{15\_I}$, ET $_{15\_I\_filter}$; subscripts M and I refer to their respective spectral correction method Moncrieff et al. (1997) and Ibrom et al. (2007)) and 1.6m-height above ground (ET $_{1.6\_M}$, ET $_{1.6\_F}$, ET $_{1.6\_F\_filter}$;, subscript M and F refers to their respective spectral correction method (Moncrieff et al. (1997) and Fratini et al. (2012)). Values are also reported for independent measurements taken also at both levels with replicated lysimeters (ET_{Lys}) and up-scaled ET ($ET_{upscaled} = ET_{Lys} + ET_{sap}$), modelled ET_{RF}, and EC-derived ET; $ET_{residual}$ estimated as the sum of net radiation (R_n), ground soil heat flux (G) and sensible heat flux (H_15) and ET_{Bowen} as $(R_n - G)/(1 + \beta)$, where the Bowen ratio is defined as $H_{\_15}/LE_{\_15\_M}$. $LE_{ensemble}$ computed from the ensemble of $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residual}$, and LE_{Bowen} is also included. Uncertainties were computed according to Section 2.5, noting that errors in EC-derived LE and independent LE are not comparable. For the acronyms refers to 'Nomenclature' section.

Ecosystem ET

ET_15_M (mm year^{-1})	ET_15_I (mm year^{-1})	ET_15_I_filter (mm year^{-1})	ET_15_I_N (mm year^{-1})	ETupscaled (mm year^{-1})	ET_RF (mm year^{-1})	ETBowen (mm year^{-1})	ETresidual (mm year^{-1})	ETensemble (mm year^{-1})
466 ± 3	455 ± 2	497 ± 2	552 ± 3	554 ± 36	597 ± 1	580 ± 3	691 ± 3	549 ± 43

Understory ET

ET_1.6_M (mm year^{-1})	ET_1.6_F (mm year^{-1})	ET_1.6_F_filter (mm year^{-1})	ETLys (mm year^{-1})
323 ± 2	372 ± 2	435 ± 2	497 ± 39

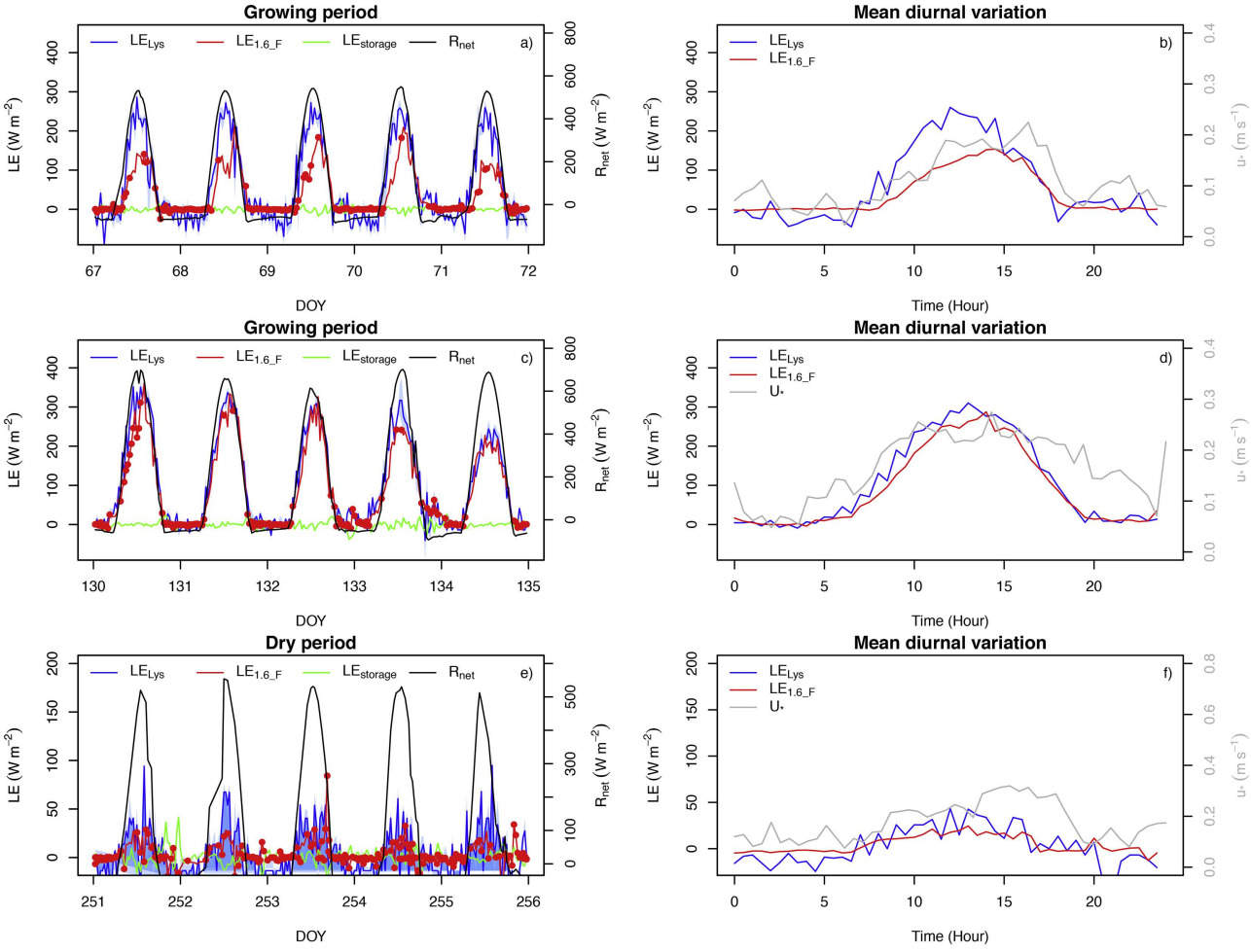

Fig. 3. Diurnal time course of half-hourly values of understory latent heat (LE) fluxes by eddy covariance (EC) and lysimeter, net radiation (R_{net}), and flux storage ($LE_{storage}$) for two-week periods in spring (Fig. 3a, c) and one week during the dry season (Fig. 3e). Filled circles represent low-quality data for EC. Respective mean diurnal variations and the friction velocity (u_*) are also presented (Fig. 3b, d, f). Shade areas in LE_{Lys} represent the error (σ) computed from the 6 lysimeter stations.

a two-weeks period in spring (Fig. 3a,c), the period with largest discrepancies, and one week during the dry season (Fig. 3e). Diurnal time trends of R_{net} and the storage LE term are also included. Respective mean diurnal variation of LE_{Lys} and $LE_{\_1.6\_F}$ (LE corrected for low-pass filtering effects) as well as u_* of those weeks are also presented (Fig. 3b, d, f). The degree of agreement between LE_{Lys} and $LE_{\_1.6\_F}$ varied largely over the growing season, alternating days with considerable differences (Fig. 3a) with periods having a remarkable agreement (Fig. 3b). Over the diurnal time course, whilst $LE_{\_1.6\_F}$ was systematically lower than LE_{Lys} in the morning and around noon, better agreement was observed in the late afternoon (Fig. 3a) with higher turbulence conditions (Fig. 3b). Contrary to LE_{Lys}, which apparently scaled with R_{net} as long as energy was the main limiting factor of LE, the degree of correlation between R_{net} and $LE_{\_1.6\_F}$ varied among consecutive days (Fig. 3a). Occasionally, those biases disappeared and both LE_{Lys} and $LE_{\_1.6\_F}$ matched well during most of the week (Fig. 3c-d). The high density of low-quality night-time $LE_{\_1.6\_F}$ data (red dots in Fig. 3a, c, e) reflects the low per-

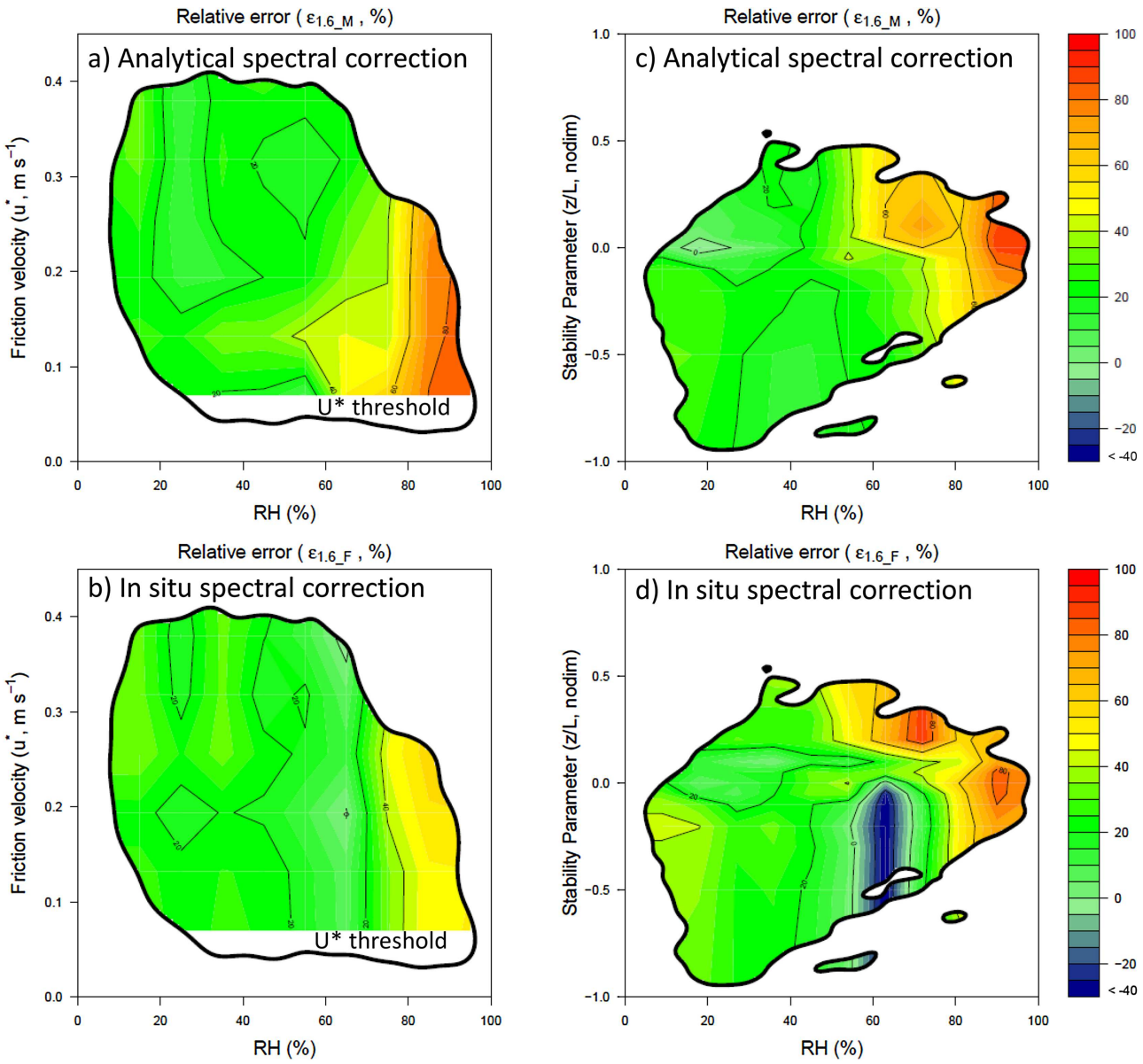

Fig. 4. Sensitivity analysis of $\epsilon_{1.6\text{-}M}$ and $\epsilon_{1.6\text{-}F}$ using RF model by varying (a, c) u_* and RH and fixing θ_s as 39.43% and z/L as −0.05, and (b, d) varying z/L and RH whilst fixing θ_s and u_* as 39.43%vol. and 0.18 ms^{-1}, respectively. Fixed values were estimated from their respective mean. The black line represents the area containing the 95% of the data defined by the Kernel distribution estimate at 95%.

formance of EC at night and the consequent low ability to measure negative LE fluxes when dewfall occurs over calm nights as shown by night-time negative values in March (Fig. 3a) with the lysimeter. In fact, dew preceded calm mornings coinciding with days with the largest discrepancies (Fig. 3b) as compared to nights without dew (Fig. 3d). Under dry conditions, closer agreement was observed between $LE_{\_1.6\_F}$ and LE_{Lys} throughout the week, but variations were small due to water limitation (Fig. 3e, f).

When available energy is mostly used to evaporate dew, heat and mass transport do little to actuate turbulence (i.e., the buoyancy flux is weak). Note that dewfall at night was a consideration particularly over calm nights during the growing season at our site. During wet mornings, the magnitude and impact of such errors is not negligible since $LE_{\_1.6}$ was highly proportional to available energy. Therefore, conditions with turbulent mixing deficiencies do not preclude high evapotranspiration rates as long as LE reaches equilibrium evaporation and the degree of uncoupling with changing environmental conditions is low (Jarvis and McNaughton, 1986).

Previous studies have shown that the spatial distribution of sources and sinks results in turbulence intermittency and measurement errors (Baldocchi and Meyers, 1991). For example, radiation gradients on the ground due to tree shadows, which bring about horizontal advection of energy, might cause imprecision and errors in turbulent fluxes. Although this effect is well described in semi-arid sites (Raz-Yaseef et al., 2010; Tuzet et al., 1997), here we have shown that most of the observed inconsistencies in $LE_{\_1.6}$ occurred during periods with low convective transport and discrepancies followed quite structured patterns. Although the site is characterized by low tree density with ca. 22 trees ha^{-1} on average (11% of canopy fraction in the daytime footprint area), strong radiation gradients causing flux divergences could explain the discrepancies observed under very dry conditions. On an hourly basis, better agreement was normally observed late in the afternoon when a lower radiation gradient is expected. Obviously, these conditions hold as long as issues related to stationarity and low wind velocities among

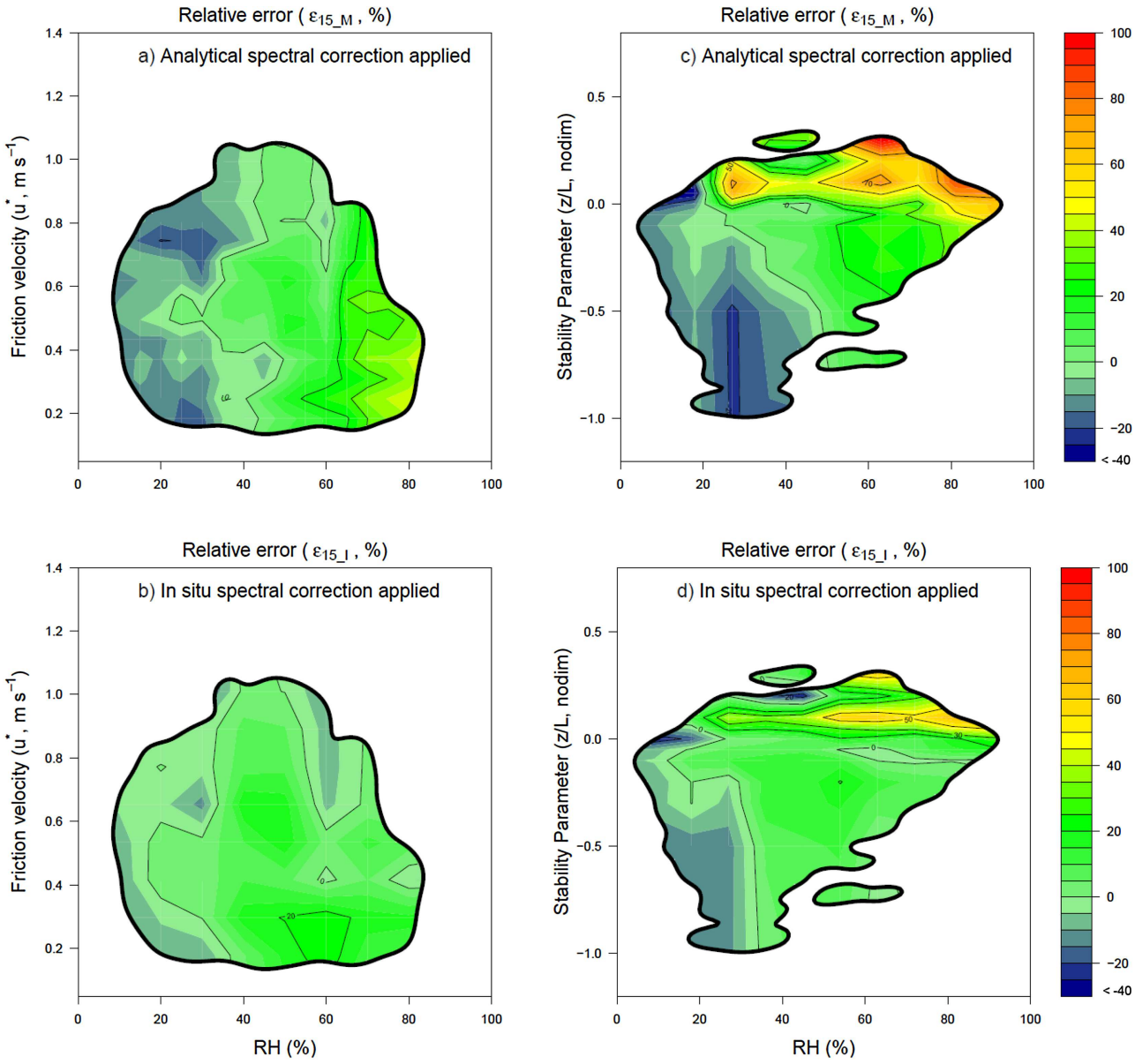

Fig. 5. Sensitivity analysis of ϵ_{15_M} and ϵ_{15_I} using RF model by varying (a, c) u_* and RH and fixing θ_s as 45% and z/L as −0.07, and (b, d) varying z/L and RH whilst fixing θ_s and u_* as 45%vol. and 0.48 ms^{-1}, respectively. Fixed values were estimated from their respective mean. Note that each panel only includes results from representative conditions. The black line represents the area containing the 95% of the data defined by the Kernel distribution estimate at 95%.

other limiting environmental conditions constrain the ability of EC to measure LE under canopies.

3.3. *Diagnosing underlying causes of ϵ using the random forest model*

The response of understory $\epsilon_{1.6\_M}$ to two pairs of selected predictors (i.e. RH with u_* and RH with z/L) was evaluated with the RF model (Fig. 4). The response of the $\epsilon_{1.6\_M}$ RF model to variations in RH, z/L and u_* showed that the greatest discrepancies occurred with high RH (Fig. 4a) or mid to high-RH-stable atmospheric stratification conditions (Fig. 4c). Under certain conditions (i.e. unstable and moderate-low LE), minor differences were observed just above the estimated u_* threshold (Fig. 4a, b). This suggests that the night time u_* filtering – a commonly used method that largely depends on the respiration-temperature relationship under night-time conditions – is an effective means of reducing bias errors in EC LE. As for CO_2 fluxes, identifying periods with turbulent deficiencies is critical for accurate estimates of EC LE. Here we have demonstrated that for LE the problem considerably extends to daytime, particularly during wet mornings with low heat transport, when significant moisture was present in the near-ground and plant surfaces.

Although with certain distinctions upon atmospheric conditions, $\epsilon_{1.6\_F}$ was slightly reduced as compared to $\epsilon_{1.6\_M}$ in particular during conditions with mid to high RH-stable conditions (Fig. 4a, b). However, the response of the $\epsilon_{1.6\_F}$ RF model to RH and z/L exhibited similar patterns and errors persisted even though the damping of water vapor fluctuations by low-pass filtering effects was explicitly accounted for (Fig. 4d). Considering that bias errors in LE_{Lys} were minor, the analysis with $\epsilon_{1.6\_F}$ RF model predicted that LE can be accurately estimated by the understory EC under sufficiently turbulent conditions and RH lower than 80%. However, large differences were observed under stable stratification conditions, like those conditions during wet mornings with low heat transport, when significant moisture was present in the near-ground and plant surfaces.

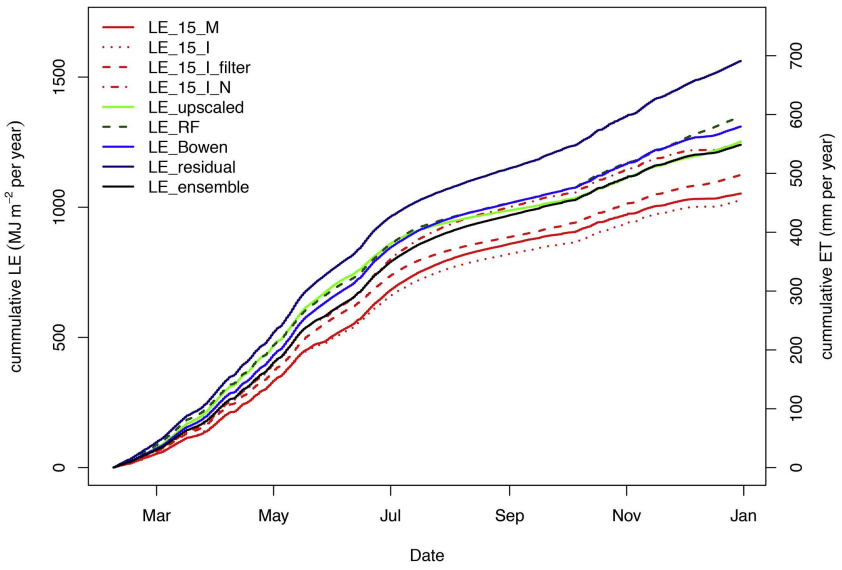

Fig. 6. Seasonal time course of cumulative EC LE according to the means of correcting for spectral loss ($LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_filter}$) and flow distortion ($LE_{\_15\_I\_N}$), independent up-scaled LE ($LE_{upscaled} = LE_{Lys} + LE_{sap}$), modelled LE_{RF}, and those EC-derived LE; $LE_{residual}$ estimated as the sum of net radiation (R_n), ground soil heat flux (G) and sensible heat flux (H_15) and LE_{Bowen} as $(R_n - G)/(1 + \beta)$, where the Bowen ratio is defined as $H_{\_15}/LE_{\_15\_M}$. An EC-derived LE computed with an ensemble of $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residual}$, and LE_{Bowen} is also presented. Data are presented from mid-February to end December 2015. For the acronyms refers to 'Nomenclature' section). Cumulative LE expressed as evapotranspiration (ET, mm $year^{-1}$) is also depicted.

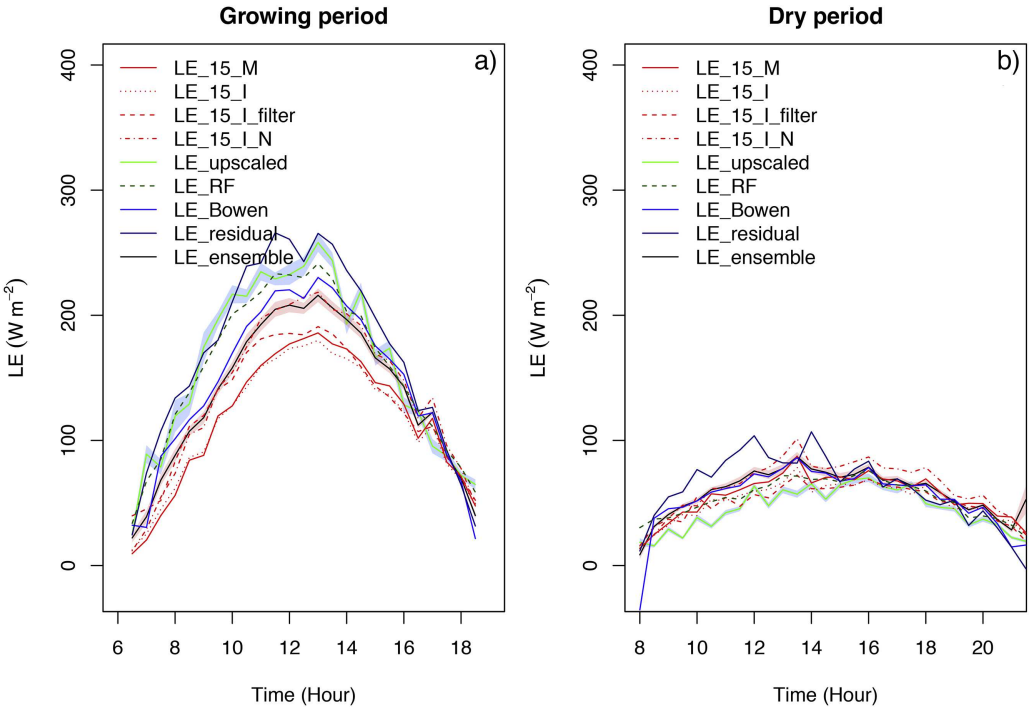

Fig. 7. Mean diurnal variation of EC LE according to the means of correcting for spectral loss ($LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_filter}$) and flow distortion ($LE_{\_15\_I\_N}$), independent up-scaled LE ($LE_{upscaled} = LE_{Lys} + LE_{sap}$), modelled LE_{RF}, and those EC-derived LE; $LE_{residual}$ estimated as the sum of net radiation (R_n), ground soil heat flux (G) and sensible heat flux (H_15) and LE_{Bowen} as $(R_n - G)/(1 + \beta)$, where the Bowen ratio is defined as $H_{\_15}/LE_{\_15\_M}$ and $LE_{ensemble}$ computed as the mean of $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residual}$, and LE_{Bowen} for (a) growing season, the period with largest discrepancies (March–June 2015), and (b) dry period (August-September 2015). For the acronyms refers to 'Nomenclature' section). Shadow areas represent the propagated error (σ).

The differences of applying analytical versus in situ spectral loss correction methods on the distribution of ϵ_{15} according to atmospheric conditions are explicitly shown in Fig. 5. Interestingly, at 15 m-height above ground, the patterns shown by both ϵ_{15_M} and ϵ_{15_I} RF models revealed minor differences as compared to those shown by both $\epsilon_{1.6\_M}$ and $\epsilon_{1.6\_F}$ RF models. Contrary to the understory tower, $LE_{\_15\_M}$ had a tendency to over-estimate $LE_{upscaled}$ in particular during dry conditions. However, those features disappeared when the in situ spectral correction was applied. On the other hand, $LE_{\_15\_I}$ apparently reduced the effect of low-pass filtering effects, particularly under nearly neutral and stable conditions (compare Fig. 5c and d). Therefore, once the attenuation of water vapor fluctuations due to water adsorption processes was explicitly accounted for, the largest ϵ_{15_I} were mostly restricted to medium/high RH and stable stratification conditions (Fig. 5d). An overview of the frequency of occurrence of relative errors and their

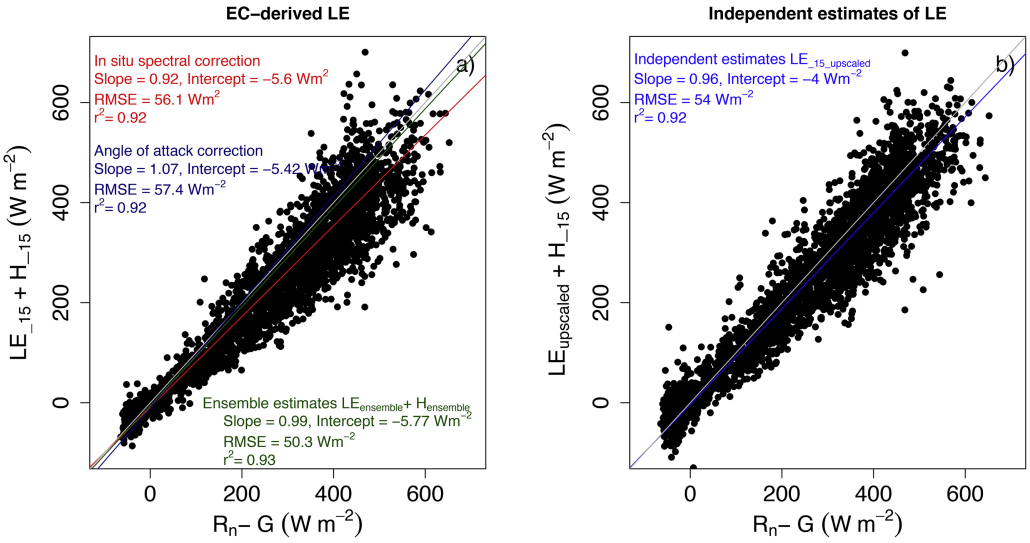

Fig. 8. The relationship between available energy (the difference between net radiation (R_n) and ground heat flux (G)) and (a) the sum of both EC-derived LE and H fluxes and b) the sum of EC-derived H and independent LE estimates ($LE_{upscaled}$). The slope in different colors from the best fit of each relationship is presented: i) in situ spectral correction applied (red color), ii) in situ spectral correction applied + angle of attack correction applied in LE and H (dark blue), iii) LE and H are computed from the ensemble of $LE_{\_15\_M}$, $LE_{\_15\_I}$, $LE_{\_15\_I\_N}$, $LE_{residual}$, and LE_{Bowen} for $LE_{ensemble}$ and $H_{\_15\_M}$, $LE_{\_15\_I\_N}$ for $H_{ensemble}$. Resulting statistics (i.e. root mean square error, RMSE, and determination coefficient, R^2) are also presented for each regression. The slope that defines the degree of energy balance closure (EBC) was calculated by using major axis regression implemented in the model II regression methods in the R package *lmodel2*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

absolute magnitudes are shown in Fig. S3.1 and Fig. S3.2, respectively in the supplementary material.

3.4. Short- and long-term impacts on LE of applying different correction methods

Irrespective of the spectral correction applied, we found that $LE_{\_15}$ had a tendency to underestimate $LE_{upscaled}$, and alone cannot fully explain the mismatch between LE by EC and independent estimates. According to the results presented in Table 1, relative differences between annual sums of EC LE and $LE_{upscaled}$ were 10% ($LE_{\_15\_I\_filter}$), 18% ($LE_{\_15\_I}$), or 16% ($LE_{\_15\_I}$) depending on the means of correcting for high-frequency losses. Those differences were clearly smaller than the discrepancies found between $LE_{\_1.6}$ and LE_{Lys} confirming that errors in LE due to high frequency loss are more pronounced near the ground. However, comparable estimates were found between $LE_{\_15\_I}$ and $LE_{upscaled}$ once the angle of attack correction was applied. Comparing EC-derived estimates that correct for the lack of EBC, LE_{Bowen} matches with $LE_{upscaled}$ and only a 1% difference was found. However, $LE_{residual}$ exceeded $LE_{upscaled}$ by 20%. Interestingly, most of the differences among methods emerged during the growing period since trends in the cumulative sums of LE held parallel over the dry period and autumn (Fig. 6), when LE was largely restricted by water limitation and later by energy, respectively.

Remarkably, $LE_{upscaled}$ slightly fell below the others in particular during the dry period when ground LE had less control and the grass had already dried out. Considering that the seasonal trends of $LE_{\_1.6}$ and LE_{Lys} were in agreement during the dry period, $LE_{upscaled}$ could have been slightly biased by LE_{sap}. However, the magnitude of the bias in LE_{sap} should be almost negligible due to the low contribution of the tree canopy to annual LE at our site. Notice that the minimum value of the ratio of daily sums of $LE_{\_1.6\_F}$ to $LE_{\_15\_I}$ was around 0.4 in summer. Interestingly, the RF LE model seemed to correct such discrepancies and its cumulative sum closely followed LE_{Bowen}.

A clear picture of the observed discrepancies among EC-derived estimates of LE and independent methods can be drawn from the mean diurnal variation over both growing and dry periods (Fig. 7). Similar to the pattern shown in the comparison between $LE_{\_1.6}$ and LE_{Lys}, a large gap was found between $LE_{\_15}$ and the rest of LE estimates during the morning and noon hours in the growing period (Fig. 7a). Contrarily and excepting the peak shown in the morning by $LE_{residual}$, discrepancies were minor during the dry season among all LE estimates (Fig. 7b). Although biases in $LE_{\_15\_I\_filter}$ (LE corrected for low-pass filtering and gap-filled when $\epsilon_{\_15\_I} > 30\%$) were slightly reduced during the mornings, large differences still remained. Obviously, $LE_{\_15\_I\_filter}$ unsatisfactorily reproduced the complete diurnal pattern of $LE_{upscaled}$, probably due to the poor performance of the gap-filling method when large gaps in the data are provided, particularly in the growing season. Interestingly, differences between LE_{Bowen} and $LE_{\_15\_I}$ were minor when the latter included angle of attack corrections. However, effects by flow distortion were higher on the H and therefore the impact of the method on energy balance closure should be assessed more comprehensively.

In fact, one of the questions left open by this study regards the extent to which the bias in EC LE explains the observed residuals of the energy equation against H. Elucidating this question is always difficult when i) biases in H and LE can be subjected to different sources of error (i.e. meteorological conditions, instrumental bias, correction methods) and its relative contribution might vary upon conditions (i.e. water availability) and ii) other causes of the lack of energy balance closure such as changes in the heat storage of air and mass, advection processes, and mismatches in the radiometric and turbulent flux footprints among others should be considered. Previous studies on EBC have found that the degree of agreement is highly sensitive to environmental conditions (Amiro, 2009; Barr et al., 2006; Chávez et al., 2009; Wilson et al., 2002). Our results suggested that considering the total gap of the energy budget (gap = Rnet-G-$LE_{\_15}$-$H_{\_15}$) of 508 MJ m^{-2} year^{-1} and that the difference between $LE_{upscaled}$ and $LE_{\_15\_I}$ reflects the portion of the gap explained by LE (199 MJ m^{-2} year^{-1}, 39% of the total gap) and the remaining by $H_{\_15}$ (309 MJ m^{-2} year^{-1}), biases in $H_{\_15}$ accounted for 61% of the total energy budget gap and consequently $LE_{residual}$ was overestimated by 20%. We can conclude that $LE_{residual}$ tends to over account LE as long as underestimates in H are not negligible and might largely explain the lack of EBC. This would explain the tendency of $LE_{residual}$ to depart from the other during the dry

period and suggest that the relative contribution of errors in H and LE causing the lack of EBC might vary with environmental conditions. Assuming LE estimates by the independent method to be true, LE_{Bowen} seems to be a more suitable correction method than $LE_{residual}$.

Consistent with the EBC, the observed lack of agreement between EC-derived turbulent fluxes and available energy (slope = 0.92, r^2 = 0.92, Fig. 8a) was to some extent attributed to errors in LE_{-15} since LE + H satisfactorily accounted for the measured available energy when using independent measurements of LE (slope = 0.96, r^2 = 0.92, Fig. 8b). Nevertheless, LE + H over-estimated the measured available energy when corrections for flow distortion effects on vertical wind speed were applied (slope = 1.07, r^2 = 0.92, Fig. 8a). In agreement with Kochendorfer et al. (2012), H was found to be affected by the angle-of-attack correction more than was LE. These results put forward that the degree of EBC is highly sensitive to the flux calculation method and, therefore, turbulent fluxes should be better defined as an ensemble of estimates according to correction methods in most common use (slope = 0.99, r^2 = 0.93, Fig. 8a).

4. Conclusions

Here we have evaluated underlying factors explaining biases in EC LE fluxes, which should be considered when EC LE is used for evaluating or parameterizing land-surface models as well as for water budget purposes. The use of independent estimates of LE together with machine learning methods are proposed as a powerful means to diagnose the causes of LE errors (ϵ) and give insights into the energy imbalance problem.

Elucidating the relative importance of each environmental predictor of ϵ is challenging due to the susceptibility of LE measurements to different sources of error, which result in complex responses. Random forest (RF) models have been shown to handle such complexity and have demonstrated that the attenuation of water vapor fluxes by EC is systematic under certain environmental conditions. As shown, such conditions are relevant in the morning when a convective layer is weakly developed and most of the available energy is used to evaporate dew. Although the use of appropriate spectral corrections is recommended to explicitly account for high frequency losses in water vapor fluctuations, our results indicated that the problem is not completely solved under stable conditions at our site. Because of the lack of consensus on the suitability of a particular correction method in EC data, the use of independent estimates of LE together with the use of machine-learning methods eased the identification of site-specific conditions that lead to measurement errors.

Acknowledgements

The authors acknowledge the Alexander von Humboldt Foundation for supporting this research with the Max-Planck Prize to MR, and the EUFAR TA project DEHESHyrE. We acknowledge city council of Majadas de Tiétar for its support. The authors acknowledge Martin Hertel from MPI-Jena, Ramon Lopez-Jimenez (CEAM), and Enrique Juarez-Alcalde from University of Extremadura for technical assistance. We also thank Andreas Ibrom for earlier discussion on the results.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.agrformet.2017.01.009.

References

Allen, R.G., Pereira, L.S., Howell, T.A., Jensen, M.E., 2011. Evapotranspiration information reporting: I. Factors governing measurement accuracy. Agric. Water Manage. 98 (6), 899–920.
Amiro, B., 2009. Measuring boreal forest evapotranspiration using the energy balance residual. J. Hydrol. 366 (1–4), 112–118.
Aubinet, M., et al., 2000. Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. Adv. Ecol. Res. 30, 113–175.
Baldocchi, D.D., Meyers, T.P., 1991. Trace gas-exchange above the floor of a deciduous forest. 1. Evaporation and Co2 efflux. J. Geophys. Res. Atmos. 96 (D4), 7271–7285.
Baldocchi, D., et al., 2001. 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 (11), 2415–2434.
Barr, A.G., Morgenstern, K., Black, T.A., McCaughey, J.H., Nesic, Z., 2006. Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurement of the CO_2 flux. Agric. For. Meteorol. 140, 322–337.
Berbigier, P., et al., 1996. Transpiration of a 64-year-old maritime pine stand in Portugal. 2: Evapotranspiration and canopy stomatal conductance measured by an eddy covariance technique. Oecologia 107 (1), 43–52.
Burgess, S.S.O., et al., 2001. An improved heat pulse method to measure low and reverse rates of sap flow in woody plants. Tree Physiol. 21 (9), 589–598.
Cammalleri, C., et al., 2013. Combined use of eddy covariance and sap flow techniques for partition of ET fluxes and water stress assessment in an irrigated olive orchard. Agric. Water Manage. 120, 89–97.
Castellví, F., Snyder, R.L., 2010. A comparison between latent heat fluxes over grass using a weighing lysimeter and surface renewal analysis. J. Hydrol. 381 (3–4), 213–220.
Chávez, J.L., Howell, T.A., Copeland, K.S., 2009. Evaluating eddy covariance cotton ET measurements in an advective environment with large weighing lysimeters. Irrig. Sci. 28 (1), 35–50.
Evett, S.R., Schwartz, R.C., Howell, T.A., Louis Baumhardt, R., Copeland, K.S., 2012. Can weighing lysimeter ET represent surrounding field ET well enough to test flux station measurements of daily and sub-daily ET? Adv. Water Resour. 50 (0), 79–90.
Falge, E., et al., 2005. Comparison of surface energy exchange models with eddy flux data in forest and grassland ecosystems of Germany. Ecol. Modell. 188, 174–216.
Finnigan, J.J., Clement, R., Malhi, Y., Leuning, R., Cleugh, H.A., 2003. A re-evaluation of long-term flux measurement techniques – Part I: Averaging and coordinate rotation. Boundary Layer Meteorol. 107 (1), 1–48.
Finnigan, J., 2006. The storage term in eddy flux calculations. Agric. For. Meteorol. 136 (3–4), 108–113.
Foken, T., Wimmer, F., Mauder, M., Thomas, C., Liebethal, C., 2006. Some aspects of the energy balance closure problem. Atmos. Chem. Phys. 6, 4395–4402.
Foken, T., et al., 2010. Energy balance closure for the LITFASS-2003 experiment. Theor. Appl. Climatol. 101, 149–160.
Foken, T., et al., 2011. Results of a panel discussion about the energy balance closure correction for trace gases. Bull. Am. Meteorol. Soc. 92 (4), ES13–ES18.
Frank, J.M., Massman, W.J., Ewers, B.E., 2013. Underestimates of sensible heat flux due to vertical velocity measurement errors in non-orthogonal sonic anemometers. Agric. For. Meteorol. 171, 72–81.
Franssen, H.J.H., Stockli, R., Lehner, I., Rotenberg, E., Seneviratne, S.I., 2010. Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations. Agric. For. Meteorol. 150 (12), 1553–1567.
Fratini, G., Ibrom, A., Arriga, N., Burba, G., Papale, D., 2012. Relative humidity effects on water vapour fluxes measured with closed-path eddy-covariance systems with short sampling lines. Agric. For. Meteorol. 165, 53–63.
Garratt, J.R., 1984. Evapotranspiration From Plant Communities The measurement of evaporation by meteorological methods. Agric. Water Manage. 8 (1), 99–117.
Hogg, E.H., et al., 1997. A comparison of sap flow and eddy fluxes of water vapor from a boreal deciduous forest. J. Geophys. Res. Atmos. 102 (D24), 28929–28937.
Holmes, J.W., 1984. Measuring evapotranspiration by hydrological methods. Agric. Water Manage. 8 (1-3), 29–40.
Horst, T.W., Semmer, S.R., Maclean, G., 2015. Correction of a non-orthogonal, three-component sonic anemometer for flow distortion by transducer shadowing. Boundary Layer Meteorol. 155 (3), 371–395.
Horst, T.W., 1997. A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors. Boundary Layer Meteorol. 82 (2), 219–233.
IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.
Ibrom, A., Dellwik, E., Flyvbjerg, H., Jensen, N.O., Pilegaard, K., 2007. Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems. Agric. For. Meteorol. 147 (3–4), 140–156.
Jaeger, L., Kessler, A., 1997. Research on forest environmental influences in a changing world twenty years of heat and water balance climatology at the Hartheim pine forest, Germany. Agric. For. Meteorol. 84 (1), 25–36.
Jarvis, P.G., McNaughton, K.G., 1986. Stomatal control of transpiration: scaling up from leaf to region. In: MacFadyen, A., Ford, E.D. (Eds.), Advances in Ecological Research. Academic Press, pp. 1–49.

Kochendorfer, J., Meyers, T.P., Frank, J., Massman, W.J., Heuer, M.W., 2012. How well can we measure the vertical wind speed? Implications for fluxes of energy and mass. Boundary Layer Meteorol. 145 (2), 383–398.
Kochendorfer, J., Meyers, T.P., Frank, J.M., Massman, W.J., Heuer, M.W., 2013. Reply to the comment by mauder on how well can we measure the vertical wind speed? implications for fluxes of energy and mass. Boundary Layer Meteorol. 147 (2), 337–345.
Kool, D., et al., 2014. A review of approaches for evapotranspiration partitioning. Agric. For. Meteorol. 184 (0), 56–70.
Kowalski, A.S., 2008. Comment on the storage term in eddy flux calculations. Agric. For. Meteorol. 148 (4), 691–692.
Lee, X., Black, T.A., 1993. Atmospheric turbulence within and above a douglas-fir stand: part II: eddy fluxes of sensible heat and water vapour. Boundary Layer Meteorol. 64, 369–389.
Leuning, R., van Gorsel, E., Massman, W.J., Isaac, P.R., 2012. Reflections on the surface energy imbalance problem. Agric. For. Meteorol. 156, 65–74.
Liaw, A., Wiener, M., 2002. Classification and regression by randomforest. R. News 2 (3), 18–22.
Mamadou, O., Gourlez de la Motte, L., De Ligne, A., Heinesch, B., Aubinet, M., 2016. Sensitivity of the annual net ecosystem exchange to the cospectral model used for high frequency loss corrections at a grazed grassland site. Agric. For. Meteorol. 228, 360–369.
Mammarella, I., et al., 2009. Relative humidity effect on the high-frequency attenuation of water vapor flux measured by a closed-path eddy covariance system. J. Atmos. Ocean. Technol. 26 (9), 1856–1866.
Massman, W.J., Lee, X., 2002. Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. Agric. For. Meteorol. 113 (1–4), 121–144.
Massman, W.J., 2000. A simple method for estimating frequency response corrections for eddy covariance systems. Agric. For. Meteorol. 104 (3), 185–198.
Mauder, M., Foken, T., 2005. Processing and quality control of eddy covariance measurements. Geophys. Res. Abstr. 7 (03322) (SRef-ID: 1607-7962/gra/EGU05-A-03322).
Mauder, M., Desjardins, R.L., Pattey, E., Worth, D., 2010. An attempt to close the daytime surface energy balance using spatially-averaged flux measurements. Boundary Layer Meteorol. 136 (2), 175–191.
Mauder, M., 2013. A comment on how well can we measure the vertical wind speed? Implications for fluxes of energy and mass by Kochendorfer et al. Boundary Layer Meteorol. 147 (2), 329–335.
Meiresonne, L., et al., 2003. Water flux estimates from a Belgian Scots pine stand: a comparison of different approaches. J. Hydrol. 270 (3–4), 230–252.
Moderow, U., et al., 2009. Available energy and energy balance closure at four coniferous forest sites across Europe. Theor. Appl. Climatol. 98 (3–4), 397–412.
Moncrieff, J.B., et al., 1997. A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. J. Hydrol. 189 (1–4), 589–611.
Moncrieff, J., Clement, R., Finnigan, J., Meyers, T., 2004. Averaging, detrending, and filtering of eddy covariance time series. In: Lee, X., Massman, W., Law, B. (Eds.), Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis. Kluwer Academic, Atmospheric and Oceanographic Sciences Library, Dordrecht, pp. 7–31.
Nakai, T., van der Molen, M.K., Gash, J.H.C., Kodama, Y., 2006. Correction of sonic anemometer angle of attack errors. Agric. For. Meteorol. 136 (1-2), 19–30.
Oishi, A.C., Oren, R., Stoy, P.C., 2008. Estimating components of forest evapotranspiration: a footprint approach for scaling sap flux measurements. Agric. For. Meteorol. 148 (11), 1719–1732.
Oliphant, A.J., et al., 2004. Heat storage and energy balance fluxes for a temperate deciduous forest. Agric. For. Meteorol. 126 (3–4), 185–201.
Oncley, S.P., et al., 2007. The energy balance experiment EBEX-2000. Part I: overview and energy balance. Boundary Layer Meteorol. 123 (1), 1–28.
Paco, T.A., et al., 2009. Evapotranspiration from a Mediterranean evergreen oak savannah: the role of trees and pasture. J. Hydrol. 369 (1–2), 98–106.
Papale, D., et al., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3 (4), 571–583.
Perez-Priego, O., et al., 2015. Sun-induced chlorophyll fluorescence and photochemical reflectance index improve remote-sensing gross primary production estimates under varying nutrient availability in a typical Mediterranean savanna ecosystem. Biogeosciences 12 (21), 6351–6367.
R Development Core Team, 2010. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org.
Raz-Yaseef, N., Rotenberg, E., Yakir, D., 2010. Effects of spatial variations in soil evaporation caused by tree shading on water flux partitioning in a semi-arid pine forest. Agric. Forest Meteorol. 150 (3), 454–462.
Reichstein, M., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biol. 11 (9), 1424–1439.
Reth, S., Seyfarth, M., Gefke, O., Friedrich, H., 2007. Lysimeter Soil Retriever (LSR)—a new technique for retrieving soil from lysimeters for analysis. J. Plant Nutrit. Soil Sci. 170 (3), 345–346.
Runkle, B.R.K., Wille, C., Gažovič, M., Kutzbach, L., 2012. Attenuation correction procedures for water vapour fluxes from closed-path eddy-covariance systems. Boundary Layer Meteorol. 142 (3), 401–423.
Soubie, R., Heinesch, B., Granier, A., Aubinet, M., Vincke, C., 2016. Evapotranspiration assessment of a mixed temperate forest by four methods: eddy covariance, soil water budget, analytical and model. Agric. For. Meteorol. 228–229, 191–204.
Stoy, P.C., et al., 2006. An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration. Agric. For. Meteorol. 141 (1), 2–18.
Stoy, P.C., et al., 2013. A data-driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity. Agric. For. Meteorol. 171–172, 137–152.
Su, H.B., Schmid, H.P., Grimmond, C.S.B., Vogel, C.S., Oliphant, A.J., 2004. Spectral characteristics and correction of long-term eddy-covariance measurements over two mixed hardwood forests in non-flat terrain. Boundary Layer Meteorol. 110 (2), 213–253.
Tuzet, A., Castell, J.F., Perrier, A., Zurfluh, O., 1997. Flux heterogeneity and evapotranspiration partitioning in a sparse canopy: the fallow savanna. J. Hydrol. 189 (1–4), 482–493.
Twine, T.E., et al., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. For. Meteorol. 103 (3), 279–300.
Van Der Molen, M.K., Gash, J.H.C., Elbers, J.A., 2004. Sonic anemometer (co)sine response and flux measurement: II: The effect of introducing an angle of attack dependent calibration. Agric. For. Meteorol. 122 (1–2), 95–109.
Vickers, D., Mahrt, L., 1997. Quality control and flux sampling problems for tower and aircraft data. J. Atmos. Ocean. Technol. 14 (3), 512–526.
Wilczak, J.M., Oncley, S.P., Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. Boundary Layer Meteorol. 99 (1), 127–150.
Wilson, K.B., Hanson, P.J., Baldocchi, D.D., 2000. Factors controlling evaporation and energy partitioning beneath a deciduous forest over an annual cycle. Agric. For. Meteorol. 102 (2–3), 83–103.
Wilson, K.B., Hanson, P.J., Mulholland, P.J., Baldocchi, D.D., Wullschleger, S.D., 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance. Agric. For. Meteorol. 106 (2), 153–168.
Wilson, K., et al., 2002. Energy balance closure at FLUXNET sites. Agric. For. Meteorol. 113 (1–4), 223–243.
Wohlfahrt, G., et al., 2009. On the consequences of the energy imbalance for calculating surface conductance to water vapour. Agric. For. Meteorol. 149 (9), 1556–1559.
Wohlfahrt, G., et al., 2010. Insights from independent evapotranspiration estimates for closing the energy balance: a grassland case study. Vadose Zone J. 9 (4), 1025–1033.