

Response to referee comments

Referee 1

Suggestion	Action
I do, however, find the text to be a bit too 'wordy' with the descriptions of the processes involved taking up too much text and being hard to follow.	Introduction and conclusions simplified and shortened. The main body otherwise stands largely unchanged; while we are happy to accept and respond to further criticism of this section, I hope we can respond appropriately to more specific critique in the next phase of review.
I suggest that the authors present their conceptual model of the relevant components (eg source = [turbulent flux] + [turbulent flux lost as advection] + [storage] + [storage lost as advection]) so that it is easier to present these concepts in the text.	We are not able to quantify the suggested terms using the method outlined in the paper. All we are able to ascertain is the partitioning of the source term between advection, turbulent flux and storage. This probably underscores that the original manuscript did not provide sufficient clarity in explaining the approach. Again, hopefully this is something that can be amended in full review.

5

Referee

2

2

Suggestion	Action
I recommend to change in the y axis of several figures (e.g. figure 14, micromol C into micromol CO2). Reason: the mole unit should be referred to molecules, and C alone is not a molecule.	All instances amended in Figures
I also recommend to change the somewhat confusing term 'mixing ratio' into 'dry molar fraction', if that was the unit the Authors wanted to use.	All instances amended in text

10

# Interactions between nocturnal turbulent flux, storage and advection at an ‘ideal’ eucalypt woodland site

Ian D. McHugh<sup>1</sup>, Jason Beringer<sup>2</sup>, Shaun C. Cunningham<sup>3,6</sup>, Patrick J. Baker<sup>4</sup>, Timothy R. Cavagnaro<sup>5</sup>,  
Ralph Mac Nally<sup>6</sup>, Ross M. Thompson<sup>6</sup>

<sup>1</sup>School of Earth, Atmosphere and Environment, Monash University, Melbourne, 3800, Australia

<sup>2</sup>School of Earth and Environment, University of Western Australia, Perth, 6907, Australia

<sup>3</sup>School of Life and Environmental Sciences, Deakin University, Melbourne, 3125, Australia

<sup>4</sup>Forest dynamics laboratory, University of Melbourne, Melbourne, 3052, Australia

<sup>5</sup>Waite Research Institute and School of Agriculture, Food and Wine, University of Adelaide, Adelaide, 5005, Australia

<sup>6</sup>Institute for Applied Ecology, University of Canberra, Canberra, 2617, Australia

*Correspondence to:* Ian D. McHugh (ian.mchugh@monash.edu)

## Abstract.

While the eddy covariance technique has become an important technique for estimating long-term ecosystem carbon balance, under certain conditions the measured turbulent flux of carbon at a given height above an ecosystem does not represent the true surface flux. Profile systems have been deployed to measure periodic storage of carbon below the measurement height, but have not been widely adopted. This is most likely due to the additional expense and complexity, and possibly also the perception – given that net storage over intervals exceeding 24 hours is generally negligible – that these measurements are not particularly important. In this study, we used a three year record of net ecosystem exchange of carbon and simultaneous measurements of carbon storage to ascertain the relative contributions of turbulent carbon flux, storage and advection (calculated as a residual quantity) to the nocturnal carbon balance, and to quantify the effect of neglecting storage. The conditions at the site are in relative terms highly favourable for eddy covariance measurements, yet we found a substantial contribution (~40%) of advection to nocturnal turbulent flux underestimation. The most likely mechanism for advection is cooling-induced drainage flows, the effects of which were observed in the storage measurements. The remaining ~60% of flux underestimation was due to storage of carbon. We also showed that substantial underestimation of carbon uptake (approximately 80 gC m<sup>-2</sup> a<sup>-1</sup>, or 25% of annual carbon uptake) arose when standard methods of nocturnal flux correction were implemented in the absence of storage estimates. These biases were much larger than quantifiable uncertainties in the data. Neglect of storage also distorted the relationships between the carbon exchange processes (respiration and photosynthesis) and their key controls (light and temperature, respectively). We conclude that addition of storage measurements to eddy covariance sites with all but the lowest measurement heights should be a high priority for the flux measurement community.

## **1 Introduction**

Over the past 2 decades, eddy covariance measurements have been widely adopted as a tool for aggregate flux measurement (Baldocchi, 2003), and there are now over 650 operational monitoring sites registered with the international flux network (Fluxnet: fluxnet.ornl.gov). Within the Australian regional network (OzFlux: www.ozflux.org.au), there are 29 active sites (Beringer et al., 2016, this issue). The use of the eddy covariance technique allows continuous automated monitoring of mass and energy fluxes, and long-term multi-site datasets have yielded valuable ecological insights in recent years (Baldocchi, 2008).

It has long been documented that eddy covariance measurements are prone to underestimation of the true surface efflux of carbon at night recognised (e.g. Goulden et al., 1996b). The key processes associated with this underestimation are storage and advection (Aubinet et al., 2012). In the former process, carbon may be stored below the measurement height under calm conditions. However, since the CO<sub>2</sub> mole fraction must be approximately preserved over longer time scales (e.g. 24 hours +), carbon respired and accumulated nocturnally is generally released with the initiation of buoyancy-generated turbulence following sunrise, such that the net storage of CO<sub>2</sub> is zero.

Advection involves mean transfer of carbon due to the development of horizontal and vertical gradients in scalar fields; the primary process by which this occurs is the initiation of gravity-driven drainage currents on sloping terrain (Aubinet, 2008). In contrast to storage, this mechanism generally results in a net loss of carbon from the observing system. Since substantial respiration occurs nocturnally, this loss causes a selective systematic error towards carbon uptake for cumulative carbon budgets at time scales > 24 hours. These drainage flows have been observed to occur on slopes of < 1° (Aubinet et al., 2003; Staebler and Fitzjarrald, 2004). Most non-agricultural ecosystem measurement sites are on sloped terrain because historically, flat, arable land has been cleared for agriculture. Drainage-induced advection is therefore thought to occur to varying extent at most sites.

The storage term can be calculated using relatively simple instrumentation that measures CO<sub>2</sub> concentrations along a vertical profile between the eddy covariance instrumentation and the ground (Yang et al., 2007). In contrast, attempts to measure advection have involved deployment of complex instrumentation and delivered highly uncertain results. Thus indirect approaches to data correction have been devised, the most common of which is the identification of a threshold below which the nocturnal turbulent carbon flux declines with turbulent activity (as expressed by friction velocity [ $u_*$ ]) (Goulden et al., 1996b). Since there should be no relationship between  $u_*$  and ecosystem respiration, this decline is interpreted as an increase in the storage and advection terms at the expense of the turbulent flux. Data below this threshold (herein  $u_{*th}$ ) are discarded and replaced using functional relationships between known physical respiratory drivers (primarily temperature) and nocturnal carbon fluxes.

The  $u_*$  correction has been criticised on theoretical and practical grounds (Aubinet et al., 2012; Aubinet and Feigenwinter, 2010; Van Gorsel et al., 2007), but remains the most widely adopted approach to nocturnal data correction. It should be applied to the sum of the measured turbulent flux and storage terms (Papale, 2006), but is quite commonly applied to measurements of turbulent flux in isolation, rather than the sum of the turbulent flux and storage terms. This is because only 10-30% of sites have deployed profile systems to measure storage (Papale, pers. comm., 23/11/2015). In Australia, only four of the 29 active sites have profile measurements, whereas  $\geq 15$  sites have canopies of sufficient height to warrant them (using measurement height  $> 3$  m as a threshold for requirement).

While the net storage over time is approximately zero, it is generally positive nocturnally as carbon accumulates below the measurement height. In the morning, the sign reverses due to two processes: 1) the turbulent transfer of the accumulated carbon upwards through the measurement plane, and; 2) photosynthetic carbon uptake by the canopy. Thus neglect of storage means that nocturnal respiratory carbon release is underestimated, but this is balanced by underestimation of morning ecosystem photosynthetic carbon uptake.

However, when the nocturnal  $u_*$  correction is applied, it implicitly accounts for both storage and advection. Since there is no corresponding correction for the reversal of the storage term after sunrise (i.e. a negative change in storage balancing the vertical turbulent transfer of the nocturnally accumulated carbon), the requirement for the storage term to be approximately zero is violated, and the nocturnally respired carbon is effectively counted twice (Aubinet et al., 2002). This unavoidably biases measurements towards net carbon efflux, and also affects the apparent relationship between ecosystem carbon fluxes and climatic controls.

Given the number of sites that do not have profile systems, it is thus important to quantify the effects of failing to measure storage. In this study, we use a three year-record of carbon exchange (including storage measurements) for an Australian eucalypt woodland to investigate the interaction between nocturnal turbulent flux, storage and advection. We devise a simple method to infer the magnitude of advection and in turn quantify the apportionment of the nocturnal ecosystem carbon source between turbulent flux, storage and advection. We quantify the biases in annual carbon exchange that arise from neglecting the storage term and discuss its effects on interpretation of carbon fluxes in the context of climatic drivers. Given the significant additional investment and complexity associated with the construction and deployment of profile systems alongside eddy covariance systems, it might be argued that the incurred bias of neglecting storage could be ignored if it is small relative to other measurement uncertainties. We therefore also propagate the errors associated with determination of  $u_{*th}$ , random measurement error and imputation (gap-filling) error to annual estimates of net carbon exchange, and assess their magnitude relative to biases due to neglect of storage.

Over the past 2 decades, eddy covariance measurements have been widely adopted as a tool for aggregate flux measurement (Baldocechi, 2003), and there are now over 650 operational monitoring sites registered with the international flux network (Fluxnet: fluxnet.ornl.gov). Within the Australian regional network (OzFlux: www.ozflux.org.au), there are 29 active sites (Beringer et al., 2016, this issue). The use of the eddy covariance technique allows continuous automated monitoring of mass and energy fluxes, and long-term multi-site datasets have yielded valuable ecological insights in recent years (Baldocechi, 2008).

The eddy covariance technique requires a number of implicit simplifying assumptions, perhaps the most crucial of which is that the vertical flux across the horizontal measurement plane at a given height above the surface approximates the true environmental source / sink of the measured entity. Only under a limited range of atmospheric (vigorous turbulence and approximate statistical stationarity of vector and scalar quantities) and surface conditions (level terrain and upwind horizontal homogeneity) is this assumption strictly met (Kaimal and Finnigan, 1994). When these conditions are not met, the turbulent flux at the measurement height may be unrepresentative of surface exchange, as storage of the entity below the measurement height or mean transport (advection) associated with the development of horizontal or vertical gradients in the scalar fields (assuming negligible flux divergence) occurs (Aubinet et al., 2012). This situation primarily arises nocturnally when surface radiative cooling causes the collapse of buoyancy-generated turbulence. It has long been recognised (e.g. Goulden et al., 1996b) that this is particularly problematic for the measurement of ecosystem carbon exchanges because there is substantial nocturnal respiratory CO<sub>2</sub> efflux from soil and vegetation. If this efflux is underestimated, then the net carbon exchange over aggregating periods > 24 hours is correspondingly biased towards surface uptake.

While the sum of turbulent flux and storage may be sufficient to characterise the surface flux at some sites (e.g. Dolman et al., 2002), nocturnal advection is thought to occur to varying extents at most flux sites (Aubinet, 2008), primarily via horizontal gravity-induced terrain drainage flows that occur on slopes of as little as 1° (Aubinet et al., 2003; Staebler and Fitzjarrald, 2004). Moreover, length rather than gradient of slope may be more critical for the initiation of such flows (Finnigan, 2008). Most non-agricultural ecosystem measurement sites are on sloped terrain because historically, flat, arable land has been cleared for agriculture. Drainage-induced horizontal advection is thought to be among the most important source of systematic nocturnal measurement error (Aubinet et al., 2012).

Up to the present, attempts to measure advection directly have been complex and resource-intensive and have yielded highly uncertain results (Aubinet et al., 2010; Feigenwinter et al., 2008; Leuning et al., 2008). Thus indirect approaches to nocturnal data correction are generally applied. Most common is the identification of a threshold of turbulence (using friction velocity—herein  $u_*$ —as a measure of mean turbulent activity) below which the nocturnal turbulent carbon flux declines with  $u_*$  (Goulden et al., 1996b). Given that there should be no functional relationship between ecosystem respiration and  $u_*$ , this decline—in the absence of corresponding declines in the factors that control respiration—is interpreted as an increase in the

storage and advection terms at the expense of the turbulent flux. Data below this threshold (herein  $u_{\lambda}$ ) are discarded and replaced using functional relationships between known physical respiratory drivers (primarily temperature) and nocturnal carbon fluxes, optimised for periods when  $u_{\lambda} > u_{\lambda th}$  (herein this approach is referred to as the  $u_{\lambda}$ -correction).

- 5 The  $u_{\lambda}$ -correction has been criticised on theoretical and practical grounds (Aubinet et al., 2012; Aubinet and Feigenwinter, 2010; Van Gorsel et al., 2007), but remains the most widely adopted approach to nocturnal data correction. The  $u_{\lambda}$  correction should be applied to the sum of the measured turbulent flux and storage terms (Papale, 2006), yet only 10–30% of sites have deployed profile systems to measure storage (Papale, pers. comm., 23/11/2015). In Australia, only four of the 29 active sites have profile measurements, whereas  $\geq 15$  sites have canopies of sufficient height to warrant them (using measurement height  $> 3$  m as a threshold for requirement). The storage term (which is proportional to the time rate of change of  $\text{CO}_2$ -carbon density) is approximately zero over the long term (e.g. annually) because turbulent mixing ensures the ambient  $\text{CO}_2$ -mixing ratio is (neglecting anthropogenic addition) approximately stable over time. As such, it may appear that neglect of the storage term would not affect annual carbon balances. However, this is not the case, as discussed below.
- 15 The storage term is positive at night because the  $\text{CO}_2$ -mixing ratio in the layer between the eddy covariance measurement level and the ground rises as respired carbon accumulates due to the decline of turbulent transfer. Following sunrise, the sign of the storage term reverses. This occurs through two processes. First, insolation-driven surface heating re-establishes buoyancy-generated turbulent activity, and so, assuming horizontal homogeneity, drives vertical turbulent transfer of nocturnally respired carbon upwards through the measurement plane, thereby reducing the  $\text{CO}_2$ -mixing ratio below the plane. The second process is net photosynthetic drawdown of carbon. For several hours after sunrise, the storage term reflects the superimposed effects of both processes, but it is the first process that is problematic when the  $u_{\lambda}$ -correction is applied without storage measurements.
- 20

- Since the  $u_{\lambda}$ -correction removes all nocturnal data where the turbulent flux declines as a function of  $u_{\lambda}$ , it is implicitly correcting for the combined effects of nocturnal advection *and* storage. However, the turbulent transfer of nocturnally respired carbon after sunrise is embedded in the turbulent flux signal, so is counted twice (Aubinet et al., 2002). However, mass conservation dictates that the morning turbulent release of the nocturnally respired carbon must be balanced by an equivalent change in storage. Therefore, the sum of the turbulent flux *and* storage terms is required to approximate the real time carbon source / sink. This does not imply that the  $u_{\lambda}$ -correction should be avoided at sites where storage measurements are not available because it is necessary to correct for nocturnal advective carbon losses. The amount of bias following application of the  $u_{\lambda}$ -correction at such sites depends on the relative effects of storage and advection on the nocturnal carbon balance. This implies that if the effects of advection dominate in reducing nocturnal turbulent carbon flux, then the correction is advisable because most of the carbon that would otherwise be ventilated in the morning has been removed; however, if the dominant process is storage, the correction should be avoided.
- 25
- 30

There is a need to understand the implications of neglecting the storage term given the number of sites that do not have profile systems. These implications include the likely long-term biases in ecosystem / atmosphere carbon exchanges but also the distortion of relationships between carbon exchange and key environmental controls of exchange processes. Moreover, it is important to quantify the contributions of advection and storage to nocturnal flux underestimation to understand the implications of applying the  $u_{st}$  correction at a given site. In this study, we used a three-year record of carbon exchange (including storage measurements) for an Australian eucalypt woodland to investigate the interaction between nocturnal turbulent flux, storage and advection. We devised a simple method to infer the influence of advection and we quantify the biases in annual carbon exchange that arise from neglecting the storage term and the complications that arise with respect to ecophysiological interpretation of flux data. Given the substantial cost of profile systems (the system described subsequently cost approximately \$17000AUD), it might be argued that the incurred bias of neglecting storage could be ignored if small relative to other measurement uncertainties. We therefore also propagate the errors associated with determination of  $u_{st}$ , random measurement error and imputation (gap-filling) error to annual estimates of net carbon exchange, and assess their magnitude relative to biases due to neglect of storage.

## 12 Methods

### 2.1 Site description

The site became operational in December 2011 (see Table 1 for site characteristics). It was established as part of a project investigating the concurrent effects of catchment reforestation on biodiversity, carbon sequestration and stream water yields. From an eddy covariance perspective, the site is considered 'ideal' in that it is relatively flat and homogeneous within the footprint area, and the canopy is very open (leaf area index  $\approx 1$ ). While the tower is situated on a slope of approximate south-easterly aspect (Figure 1), the slope is generally less than  $1^\circ$ . Vegetation is relatively homogeneous, consisting of a sparse eucalypt overstorey (dominant species is *E. microcarpa*, stand leaf area index [LAI]  $\approx 1$ ; see Table 1) and a sparse shrub understorey. The open canopy reduces the potential for strong decoupling of the sub-canopy space from the overlying air, and underestimation of respiration by above-canopy eddy covariance systems relative to direct chamber measurements correspondingly decreases with declining leaf area (Speckman et al., 2014). Mean annual temperature for 2012-2014 was  $15.9^\circ\text{C}$  (minimum and maximum temperatures of  $-3.1^\circ\text{C}$  and  $45.0^\circ\text{C}$ ). Average annual rainfall for the nearest long-term rainfall measurement site (Mangalore Airport; Bureau of Meteorology station ID 088109) is 560 mm (1971-2000 average).

### 2.2 Instrumentation

The eddy covariance (herein EC) method was used to measure  $\text{CO}_2$  fluxes at 36 m (mean [ $\pm$ SD] canopy height was only 15.3 [ $\pm 6.4$ ] m, but emergent individual trees were up to 30 m). The EC method requires fast-response instrumentation to

measure simultaneous variations in scalar (here carbon) and vector (3D wind velocities) quantities. A Campbell Scientific (Logan, USA) CSAT3 sonic anemometer (Table 2) was used to measure wind velocities and a Licor (Lincoln, USA) LI7500 infra-red gas analyser (herein IRGA) to measure CO<sub>2</sub> and H<sub>2</sub>O vapour ~~mixing-ratio~~ mole fractions. EC data were logged at 10 Hz (post-processing is described below), and 30-minute averages for radiant and subsurface energy fluxes and standard meteorology (temperature, humidity, wind speed and direction, rainfall, barometric pressure) were also logged. All data were transferred telemetrically to a central server at Monash University, Clayton, Australia.

A custom-built profile system using a Licor LI840 IRGA measured changes in CO<sub>2</sub> storage below the EC measurement height. The system consisted of an array of 6 gas intakes (configured logarithmically in the vertical - 0.5, 2, 4, 8, 16, 36 m) connected to a sampling system (Figure 2) via tubing (of equal length – 18m – for all heights; enclosure mounted at 18 m). A KNF Neuberger (Freiberg, Germany) NMP 850.1.2 vacuum pump drew air from all levels through a common manifold with sample and exhaust chambers. A bank of 12V SMC (Tokyo, Japan) VO307 solenoid valves switched each of the gas lines sequentially to a sampling loop (flow rate = 0.5lpm) consisting of a gas analyser (Licor LI840) and Alicat Scientific (Tucson, USA) mass flow meter, while the remaining lines bypassed the loop and were exhausted to the atmosphere. Dwell time for each level was 20 s (first 15 s for flushing of the manifold, average of last 5 s logged), translating to a measurement cycle of 2 minutes.

### 2.3 Data Processing and Analysis

Post-processing of EC data (including quality assurance and quality control) was undertaken using OzFluxQC, a software package developed by the OzFlux community (primarily P. Isaac, the data manager of OzFlux) in the Python programming language (Isaac et al., 2016, this issue). The Python programming language was used for all subsequent data analysis. Thirty-minute fluxes were calculated from the 10 Hz data using block averaging (Moncrieff et al., 2004). Corrections applied to the raw data included: 2-D coordinate rotation (Lee et al., 2004), in which the coordinate frame is rotated to force first the mean cross-wind, then second the mean vertical wind, to zero over the measurement period; frequency attenuation corrections, which account for loss of covariance associated with high frequency cut-off, sensor separation and path averaging of then instrumentation (Massman and Clement, 2004), and; density corrections associated with the effects of surface sensible and latent heat fluxes (Webb et al., 1980). Data were processed to level 4. In OzFluxQC, this represents the point at which all QA / QC except  $u_*$  correction has been applied to flux data, and meteorological data have undergone QA / QC and gap-filling (Isaac et al., 2016).

We identified the  $u_*$  threshold ( $u_{*th}$ ) using change-point detection (CPD) following Barr et al. (2013). The time series was divided into multiple temperature-stratified samples, and a two-phase linear regression model was fitted to all possible change points (change in carbon exchange as a function of  $u_*$ ) for each sample. If the change point that minimises the sum of



squares error shows statistically significant improvement over a null model (no change point), the change point (i.e.  $u_{*th}$ ) is retained (subject to additional quality control criteria, as described by Barr et al., 2013). A bootstrapping procedure (in which the data for each year were randomly sampled with replacement 1000 times; see Papale, 2006) was used to generate a probability distribution for  $u_{*th}$ , the mean and 95% confidence interval of which provides a best estimate and uncertainty interval for  $u_{*th}$ .

The rate of change of carbon storage was calculated from the difference between quasi-instantaneous (2-minute) vertical concentration profiles at the tower at the beginning and end of the flux averaging period (Finnigan, 2006). We adopted the approach for storage calculation of Yang et al. (2007):

$$\left(\frac{\Delta C}{\Delta t}\right)_{k=1} \times z_{k=1} + \sum_{k=2}^n \left\{ \left[\left(\frac{\Delta C}{\Delta t}\right)_k + \left(\frac{\Delta C}{\Delta t}\right)_{k-1}\right] \times \frac{z_k - z_{k-1}}{2} \right\} \quad (1)$$

Here,  $\Delta C / \Delta t$  is the time rate of change of carbon molar density ( $\mu\text{mol C m}^{-3}$ ),  $k$  is the profile level,  $z$  is height above the surface (m) and  $n$  is the number of profile levels. **Mixing-ratio** **Mole fraction** reported by the IRGA was converted to carbon molar density using the ideal gas law (temperature measurements were drawn from instruments co-located with the air inlet for each profile level, whereas pressure measurements were drawn from ground level). The 2-minute period preceding the beginning and the end of the 30-minute period (e.g. 1128-1130 and 1158-1200 for the 1130-1200 period) were used to calculate  $\Delta C$ . Given that the pump draw was simultaneously divided across 6 lines, there was a lag > 1 minute, so that the sampling was approximately temporally centred on the half hour. The average of the time derivative for two levels ( $k$  and  $k-1$ ) is the best estimate of the time derivative for the *layer* that has  $k$  and  $k-1$  as its upper and lower boundaries, except for the lower layer, for which it is assumed that  $\Delta C / \Delta t$  for  $k = 1$  is representative for the layer. Layers were scaled according to the layer thickness and the storage term represented the sum over all layers.

## 2.4 Analytical framework

We assess the carbon balance in the familiar context of a notional control volume with an orthogonal coordinate system (Finnigan et al., 2003), the mass balance of which (neglecting horizontal turbulent flux divergence) is:

$$NEE = \underbrace{\overline{w'c'}}_I(h_m) + \underbrace{\int_0^{h_m} \frac{\partial \bar{c}(z)}{\partial t} dz}_{II} + \underbrace{\int_0^{h_m} \left( \bar{u}(z) \frac{\partial \bar{c}(z)}{\partial x} + \bar{v}(z) \frac{\partial \bar{c}(z)}{\partial y} \right) dz}_{III} + \underbrace{\int_0^{h_m} \left( \bar{w}(z) \frac{\partial \bar{c}(z)}{\partial z} \right) dz}_{IV} \quad (2)$$

Here, NEE (net ecosystem exchange of carbon) is the true source term, term  $I$  is the turbulent flux across the upper horizontal plane of the control volume at instrument height  $h_m$  ( $\mathbf{w}$  is the vertical velocity, and overbar and prime denote mean and quasi-instantaneous fluctuation from mean, respectively), term  $II$  is the storage term integrated over finite time period ( $t$ ) and control volume depth ( $z$ ), term  $III$  is the sum of the advection components in the horizontal dimensions ( $x$  and  $y$ , with corresponding vectors  $\mathbf{u}$  and  $\mathbf{v}$ ) and term  $IV$  the vertical advection. We adopted the standard micrometeorological convention suggested by Chapin et al. (2006) in which NEE is positive (negative) when the net transfer of carbon is from ecosystem to atmosphere (atmosphere to ecosystem). In the following text, equation 2 is simplified to:

$$NEE = F_c + S_c + Av_c + Ah_c \quad (3)$$

Here,  $F_c$  and  $S_c$  are the turbulent flux and storage terms, and  $Av_c$  and  $Av_h$  are the vertical and horizontal advection terms, respectively. During the day, when turbulence is well-developed, the turbulent flux ( $F_c$ ) is generally the dominant term, but at night, the other terms may become dominant under weak mixing. Following the identification of  $u_* > u_{*th}$ , nocturnal data were rejected where  $u_* < u_{*th}$ . While nocturnal advection was not measured, it was inferred as the residual of the terms in equation

3. Nocturnally, NEE is equivalent to ecosystem respiration (herein ER). While ER is unknown when  $u_* < u_{*th}$ , it can be estimated ( $\widehat{ER}$ ) using an empirical model, the parameters of which are optimised for periods in which the sum of turbulent flux and storage approximates ER (*i.e.* when  $u_* > u_{*th}$ ). Equation 3 thus becomes:

$$\widehat{ER} - F_c - S_c = Av_c + Ah_c \quad (4)$$

## 2.5 Imputation

Carbon flux and temperature data were used to optimise the parameters of an empirical temperature response function (optimisation used the Levenberg Marquardt algorithm implemented in the Python Scipy package) that was then used to estimate ER for  $u_* < u_{*th}$  (and for subsequent gap-filling). The model was an Arrhenius-style function proposed by Lloyd and Taylor (1994):

$$ER = rb e^{E_o \left( \frac{1}{T_{ref} - T_0} - \frac{1}{T - T_0} \right)} \quad (5)$$

Here,  $rb$  is the reference respiration at a reference temperature ( $T_{ref}$ ),  $E_o$  is an activation energy parameter that controls temperature sensitivity, and  $T_0$  is the temperature at which metabolic activity approaches zero.  $T_0$  and  $T_{ref}$  are fixed at

–46.02 °C and 10 °C, respectively (Lloyd and Taylor 1994); the unconstrained version of the function is overparameterised (Reichstein et al., 2005; Richardson and Hollinger, 2005). We fitted  $E_o$  annually and then derived  $rb$  for successive 15-day windows (5-day step with linear interpolation to generate a continuous time series) (Reichstein et al., 2005). This avoids confounding of diurnal and seasonal temperature responses, and allows the model to capture low-frequency variation in ER associated with variables not explicitly represented in the model (*e.g.* soil moisture and substrate availability). We used temperature measured at the EC height (36 m) as the respiratory driver because it was found to have the lowest RMSE (relative to soil – which had the highest - and 0.5, 2, 4, 8 and 16 m air temperatures; data not shown).

Gap-filling was also required for daytime to assess the effects of nocturnal data treatment on annual NEE. We used a Michaelis Menten-type rectangular hyperbolic model (Ruimy et al., 1995) of modified form (Falge et al., 2001) to estimate NEE, where ER was calculated from equation 5 using daytime temperatures in conjunction with nocturnally-derived parameter estimates,:

$$NEE = \frac{\alpha Q}{1 - Q/2000 + \alpha Q/\beta} + ER. \quad (6)$$

Here,  $\alpha$  is the initial slope of the photosynthetic light response,  $Q$  is photosynthetic photon flux density, and  $\beta$  is photosynthetic capacity at 2000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ . The same window size, window step and interpolation procedure was used as for the nocturnal fitting of the respiration model. We adopted the additional light-response model criterion in which  $A_{opt}$  is modified to include a non-linear scaling factor to account for the effects of vapour pressure deficit (VPD) on stomatal conductance (Lasslop et al., 2010):

$$\beta = \begin{cases} \beta_0 e^{(-k(VPD - VPD_0))}, & VPD > VPD_0, \\ \beta_0, & VPD < VPD_0 \end{cases} \quad (7)$$

Here,  $VPD_0$  is a threshold value above which stomatal conductance becomes sensitive to VPD, and  $k$  is a fitted parameter defining the  $\beta$  response to VPD.

## 2.6 Uncertainty estimation

We quantified sources of uncertainty in the data arising from random measurement and model error. We calculated random error from a daily differencing procedure (Hollinger and Richardson, 2005). When differences in critical drivers are sufficiently small (<35  $\text{Wm}^{-2}$  for insolation, <3 °C for air temperature, and < 1  $\text{m s}^{-1}$  for wind speed), differences between

NEE data pairs separated by 24 hours were considered to represent random error. Since random error in EC data is heteroschedastic and follows the Laplacian distribution, we calculated the standard deviation of the error ( $\sigma[\delta]$ ) for  $j$  samples in  $i$   $u_{*}$  bins as:

$$\sigma(\theta)_i = \sqrt{2} \frac{1}{n_i} \sum_{j=1}^{n_i} |\theta_{i,j} - \bar{\theta}_i|. \quad (8)$$

$\sigma[\delta]$  was regressed on the flux magnitude to derive a linear relationship that was in turn used to estimate  $\sigma[\delta]$  for each datum. Monte Carlo simulations were used to translate these estimates to annual uncertainty, whereby for each of  $10^4$  simulations, estimates of random error for all observational data were aggregated over one year. This yielded a normal distribution of uncertainties, the  $2\sigma$  bounds of which were taken as the annual uncertainty.

With respect to model error, we followed Keith *et al.* (2009) in which, for day and night conditions, a sub-sample of  $10^3$  records with observational data was randomly selected from the annual dataset. A proportion of the observational data in the subsample equal to the observed proportion of data missing in the annual time series was then replaced with model estimates, and the summed difference between the complete observational and gap-filled subsamples was calculated and expressed as a proportion of the observational sum. This was repeated  $10^4$  times, and the  $2\sigma$  bounds of the proportional error was calculated, and then applied to the annual sum to produce an absolute annual model error.

To combine uncertainties, we assumed independence of the random and model estimates and sum in quadrature:

$$\varepsilon_{tot} = \sqrt{\varepsilon_r^2 + \varepsilon_m^2}, \quad (9)$$

Here,  $\varepsilon_{tot}$ ,  $\varepsilon_r$  and  $\varepsilon_m$  are combined total, random and model uncertainty, respectively.

Other unquantified sources of error may also be present. Barr et al. (2013) argued that one of the key sources of uncertainty in annual NEE is the estimation of  $u_{*th}$ . We used their bootstrapping approach to derive robust confidence intervals for  $u_{*th}$ , and we included estimates of the effect on annual NEE of setting  $u_{*th}$  to the upper and lower bounds of the 95% confidence interval for  $u_{*th}$ .

## 23 Results and Discussion

### 3.1 Contribution of mass balance components to nocturnal carbon dynamics

A nocturnal relationship was observed between friction velocity ( $u_*$ ) and  $F_c$  (Figure 3), with  $F_c$  declining quasi-linearly below  $u_{*th} = 0.42 \text{ m s}^{-1}$  and approaching zero at zero turbulence. There was little seasonal variation in  $u_{*th}$  (data not shown), and the estimates and uncertainty bounds for all years were similar (Table 3); we conservatively used the highest estimate across all years. Given that the primary abiotic respiration controls, temperature and soil moisture, showed no relationship with  $u_*$  except at  $u_* < 0.08 \text{ m s}^{-1}$  which is linked to declines in temperature, we interpreted this as an increase in the non-turbulent terms of the mass balance.

The decline in  $F_c$  with declining  $u_*$  was accompanied by a corresponding increase in  $S_c$  because as turbulence is progressively suppressed, carbon is expected (excluding advective losses) to be increasingly stored below the measurement height. The strong sensitivity of  $\text{CO}_2$  accumulation to  $u_{*th}$  below the measurement height was observed in raw time series data under varying  $u_*$  (Figure 4); the  $\text{CO}_2$  ~~mixing-ratio~~<sup>mole fraction</sup> responded very sensitively on variations in  $u_*$ , and the effect of  $u_*$  crossing  $u_{*th}$  is evident even in the raw data.

However,  $S_c$  was not the only important additional term in the nocturnal mass balance. A  $u_*$  dependent decline in  $F_c + S_c$  was also evident (Figure 5) below a threshold of  $u_* = 0.32 \text{ m s}^{-1}$ , although this was inherently more uncertain due to the much higher random error in storage relative to the EC measurements (Table 3). There is no plausible explanation for  $u_*$ -dependent declines in the respiratory carbon source (again, excluding changes in relevant controls), and so we inferred that this represents carbon losses associated with the remaining mass balance terms (*i.e.* advection). This can be quantitatively estimated as a residual following equation 4. Note that parameter optimisation of the temperature response function used  $F_c + S_c$  as the target variable because  $S_c$  is observed to be non-zero when  $u_* \gg u_{*th}$  (see Figure 3, and subsequent discussion).

The terms in equation 4 are plotted as a function of  $u_*$  in Figure 5. The inferred advection estimate increased rapidly below  $u_*$  threshold  $= 0.32 \text{ m s}^{-1}$ , and was comparable to  $S_c$  at the lowest  $u_*$  values. This indicates that under the calmest conditions,  $F_c$  approached zero, and approximately half of the carbon respired by the ecosystem was stored below the measurement height while the remainder was advected away. Integrated over the interval  $0 < u_* < u_{*th}$ ,  $S_c$  accounted for 61% of the difference between  $F_c$  and  $\widehat{ER}$ , with the other 39% attributed to the advection components ( $Av_c + Ah_c$ ). This indicates that even on very flat terrain, the nocturnal advection term is significant.

### 3.2 Inferred advection mechanisms

Given that  $Av_c + Ah_c$  is inferred, it is not possible to assess the relative contributions of the two components to the mass balance. However, gravity-induced drainage flow is a common mechanism for horizontal advective carbon losses. These flows are expected to be most readily identifiable from changes in  $S_c$  because drainage flows are initiated under stable conditions when turbulence is suppressed.  $F_c$  is expected to be small and  $S_c$  correspondingly large (prior to the initiation of drainage flows) under these conditions.

Drainage flows are also expected to have a vertical spatial fingerprint. Specifically, their onset may result in changes in the magnitude of  $S_c$  primarily in the lower layers of the control volume. At the study site, mean canopy height was  $15.3 \pm 6.4$  m (SD; Table 1). Given that drainage flows generally confined to the trunk space below the canopy (Aubinet et al., 2003), and conservatively assuming that the canopy comprises the upper 30% of tree height, drainage flows may be confined to depths of  $<10$  m, which are comparable to commonly reported values (Goulden et al., 2006; Mahrt et al., 2001). The impact on the carbon balance may nonetheless be large because as much as 70% of carbon is sourced from the soil in temperate forest ecosystems (Goulden et al., 1996a; Janssens et al., 2001; Law et al., 1999).

$S_c$  for the individual layers is presented in Figure 6. There was a clear decline in  $S_c$  for all layers below 8 m when  $u_*$  was less than approximately  $0.25 \text{ m s}^{-1}$ . In contrast, storage in the higher 8-16 m and 16-36 m layers continued to increase near linearly. We hypothesise that this indicates the onset of drainage flows at low levels under stable conditions, causing horizontal advective (i.e.  $Ah_c$ ) losses of carbon from the lower layers of the control volume. The ongoing increases in storage in the 8-16 and 16-36 m layers may indicate that the carbon source in these layers originated primarily from the vegetation than from upward transfer from lower layers. In the interval between the  $u_*$  thresholds for  $F_c$  alone and  $F_c + S_c$  (i.e.  $0.32 \leq u_* \leq 0.42 \text{ m s}^{-1}$ ),  $Av_c + Ah_c$  was not significantly different to zero (see Figure 5). The linear relationship between each of the lower layers (0-0.5, 0.5-2, 2-4 and 4-8 m) and the mean 8-36 m layer in this interval  $0.32 \leq u_* \leq 0.42 \text{ m s}^{-1}$  can be used to extrapolate the expected rate of change for those layers in the absence of advection when  $u_* < 0.31 \text{ m s}^{-1}$ .

Extrapolation of this linear relationship to conditions where  $u_* < 0.32$  provides an estimate of the expected magnitude of  $S_c$  in the absence of advection. If drainage flows are the primary advective mechanism, then the sum of  $F_c$  and the linearly adjusted storage term should approximate  $\bar{E}R$ . The correction to the storage in the 0-8 m layers when  $u_* < 0.31 \text{ m s}^{-1}$  increased the 0-36 m storage term such that for the interval  $0 < u_* < 0.31$ , the mass balance was approximately closed because  $\bar{E}R - F_c \approx S_c$  to within the uncertainty (95% confidence interval) of the bin means over this interval (Figure 7). This indicates that the decline in  $S_c$  at lower layers was of approximately the same magnitude as the inferred advection, consistent with the presence of low-level drainage flows removing carbon from the control volume.

Our approach is subject to substantial uncertainty because the assumption that the linear relationship between levels holds for declining  $u_*$  may not be correct. Direct observation of the proposed advective mechanism (terrain-induced drainage flows) would provide clarification of the mechanisms driving the nocturnal C dynamics at the site. While wind speed and direction were measured for all levels, the instrumentation lacked the resolution to detect the weak winds that generally characterise drainage flows in moderate terrain (typically less than 0.5 m/s; Aubinet et al., 2003; Goulden et al., 2006; Mahrt et al., 2001) (typically less than 0.5 m s<sup>-1</sup>; Aubinet et al., 2003; Goulden et al., 2006; Mahrt et al., 2001). Nonetheless, the available data are consistent with a primary advective mechanism of terrain drainage flows.

### 3.3 Effects of correction methods on diurnal and annual carbon balances

The importance of the contribution of  $S_e$  to the diurnal carbon balance is presented in Figure 8a. Given that the contribution averages approximately zero over the diurnal cycle, annual NEE sums for both  $F_e$  and  $F_e + S_e$  were comparable: approximately -450, -400 and -560 gC m<sup>-2</sup> a<sup>-1</sup> for 2012, 2013 and 2014, respectively (Table 4; small differences [ $< 20$  gC m<sup>-2</sup> a<sup>-1</sup>] were observed for  $F_e$  versus  $F_e + S_e$ , reflecting small differences in parameters of functions used for gap filling). However, the diurnal dynamics were substantially changed with the addition of  $S_e$ , with increased amplitude and phase shift of peak carbon uptake (from midday – synchronous with the solar radiative peak – on average, to about 1100). Nocturnal  $u_*$  dependency of  $F_e + S_e$  indicates the presence of advection under weak turbulence. Following  $u_*$  correction for this dependency (Figure 8b), estimated annual carbon uptake was reduced by 50–75 gC m<sup>-2</sup> a<sup>-1</sup>.

Given that the majority of sites both internationally and in Australia do not have profile measurement systems, we discuss the effects of neglect of  $S_e$  because this is the *de facto* approach taken for sites without profile systems. As a secondary option, a single point storage term (herein  $S_{e\_pt}$ ) can be derived from the EC gas analyser. This will underestimate storage for taller towers, where much carbon accumulates within the control volume, and is subject to substantial error (Gu et al., 2012; Yang et al., 2007) but may nonetheless potentially reduce bias.

Summing the turbulent flux and single point estimate (i.e.  $F_e + S_{e\_pt}$ ) and applying the  $u_*$  correction, the diurnal cycle was relatively well-approximated on average (Figure 9a).  $S_{e\_pt}$  sums to zero over each year, but substantially underestimates  $S_e$  over the diurnal cycle (Figure 10). We expect to see a bias towards efflux because although the  $u_*$  correction rectifies the nocturnal underestimation of efflux, the ventilation of accumulated carbon through the turbulent flux is not balanced by a decline of corresponding magnitude in storage. However, the results differed among years (Table 4). In 2012,  $u_*$ -corrected NEE from  $F_e + S_{e\_pt}$  was comparable to the best estimate for  $F_e + S_e$  (-375 and -382 gC m<sup>-2</sup> a<sup>-1</sup> for  $F_e + S_{e\_pt}$  and  $F_e + S_e$ , respectively). In 2013 uptake was actually higher for  $F_e + S_{e\_pt}$  (-362 vs -338 gC m<sup>-2</sup> a<sup>-1</sup> for  $F_e + S_e$ ). In 2014, the value was substantially lower (-431 versus -480 gC m<sup>-2</sup> a<sup>-1</sup> for  $F_e + S_e$ ).

The inconsistency of the annual NEE estimate between years arose due to the effect of error on model parameterisation and subsequent imputation. Random error estimates (derived from equation 8) are shown in Figure 11; at low flux magnitudes the error associated with  $F_c + S_{c,pt}$  was large relative to that for  $F_c$  alone or  $F_c + S_c$ . This may be due to a mismatch in the timing of peak  $S_{c,pt}$  relative to  $S_c$ ; small day-to-day variations in this timing may – in conjunction with noise in  $S_{c,pt}$  – result in large variations in NEE estimates for a given set of environmental conditions (as essentially demonstrated in Figure 11). Large errors in point-based storage estimates were also reported by Gu et al. (2012) for a forest ecosystem. This in turn affects the parameter estimates for  $rb$  in the respiration function (equation 5) and  $\alpha$  in the light response function (equation 6) in particular; even if this translates to small relative errors in GPP and ER, the effect on NEE – as the small difference between two large exchanges – may in relative terms be substantial.

Correction of  $F_c$  yielded comparable nocturnal NEE to that estimated from  $u^*$ -corrected  $F_c + S_c$  because it accounted for the nocturnal effects of storage and advection (Figure 9b). However, the morning increase in turbulent flux mixes nocturnally respired carbon stored in the control volume up through the measurement height. In the absence of a corresponding change in storage to account for this, the respiratory flux is counted twice, and annual NEE calculated using this method was underestimated by approximately  $80 \text{ gC m}^{-2} \text{ a}^{-1}$  relative to the best estimate (Table 4). This is larger than the correction for  $F_c + S_c$  associated with advection ( $50\text{--}75 \text{ gC m}^{-2} \text{ a}^{-1}$ ), which is consistent with the previously established fact that  $S_c$  contributes more to  $F_c$  underestimation than advection. In the absence of  $S_c$  estimates, lower annual bias would be obtained for this study if the  $u^*$  correction were *not* applied.

The double-counting problem might be mitigated at sites without storage measurements if, after nocturnal data correction, the morning 'flush' period, when the accumulated carbon in the control volume is vented following the re-establishment of thermally generated turbulence after sunrise, is identified and removed. This flush is expected to be manifested as an upward spike in  $F_c$  following sunrise (e.g. Aubinet et al., 2012, Figure 5.2 therein). However, it may be difficult to identify objectively the flushing of accumulated  $\text{CO}_2$  from the control volume from the behaviour of  $F_c$  alone.

We observed no spike in  $F_c$  indicating morning venting of accumulated carbon (Figure 8a). The magnitude of the expected  $F_c$  spike depends on: (i) the quantity of carbon stored, which depends on prior respiration and advection, and; (ii) the relative timing and magnitude of increase of source / sink activity and turbulent activity. On average,  $u^*$  reached  $u_{th}^*$  for  $F_c$  approximately two hours after sunrise, whereas the storage term began to decline immediately when insolation increased (Figure 10). This indicates that carbon stored within the control volume began to be consumed by photosynthesis before efficient ventilation of the control volume was underway. The  $\text{CO}_2$  mixing-ratios of the lowest levels measured by the profile system (0.5 and 2 m) also began to decline first (Figure 12), suggesting that shrubs reach the light-compensation point earlier after sunrise than do trees probably because lower near-surface temperatures are expected to suppress respiration. The effect on storage was small since the layers represent a small proportion of the control volume.



Early-morning photosynthesis may be substantial in eucalypt-dominated ecosystems in which the characteristically pendulous (in some species up to 75% of mature leaves typically hang at angles > 80 degrees from horizontal; Pereira et al., 1987), amphistomatous leaves evolved to maximise incident radiation at low sun angles, which shifts photosynthetic activity towards periods with lower vapour pressure deficit (James and Bell, 1996). Mutual shading would partially counteract this effect at low sun angles, but this may have less effect in systems with sparse canopies such as the woodland in this study. Similarly, an Australian temperate eucalypt forest site with long-term turbulent flux and CO<sub>2</sub> profile measurements showed no morning spike in CO<sub>2</sub> efflux (Van Gorsel et al., 2007, Figure 4 therein).

Alternatively, the flush period may be defined as occurring between sunrise and the time at which the quasi-instantaneous, vertically integrated control volume mean of the CO<sub>2</sub> ~~mixing-ratiomole fraction~~ approaches its long-term temporal mean. The long term temporal mean in this study was 395 ppm CO<sub>2</sub>, within 2 ppm of the mean southern-hemisphere ambient CO<sub>2</sub> ~~mixing-ratiomole fraction~~ during 2012-2014. As long as the CO<sub>2</sub> ~~mixing-ratiomole fraction~~ in the control volume exceeds that of the overlying atmosphere, there is potential for flushing of previously respired carbon from the control volume. If it is assumed that, over a sufficient period, mean CO<sub>2</sub> ~~mixing-ratiomole fraction~~ in the control volume is in equilibrium with the overlying atmosphere, then the concentration gradient between the control volume and the overlying atmosphere will continue to facilitate upward turbulent transfer of CO<sub>2</sub> until the control volume CO<sub>2</sub> ~~mixing-ratiomole fraction~~ declines to its mean. In this study, this occurred on average at approximately 0900 local standard time compared with austral / autumnal vernal equinox sunrise times of 0622 and 0611, respectively. Ensuring that no data were included during the period where flushing of nocturnally respired carbon could potentially have occurred would require removal of data between 0600-0900.

Essentially the reverse process to that described above occurred during daytime and early evening. Negative daytime  $S_c$  after 0900 reflects ongoing photosynthetic CO<sub>2</sub> drawdown within the control volume.  $S_c$  approached zero at approximately 1600 on average, because insolation, which drives photosynthesis, is out of phase with surface heating, which drives  $u_*$  and respiration and so photosynthesis declines earlier than respiration and turbulent mixing. Therefore, NEE estimates continued to be biased throughout the day in the absence of  $S_c$ . However, this has little effect on long-term carbon balance because the daytime bias in  $F_c$  due to neglect of  $S_c$  is offset by nocturnal bias of opposite sign in the early evening. Given that the CO<sub>2</sub> ~~mixing-ratiomole fraction~~ dropped below the mean at 0900 and remained below it until 2100,  $S_c$  summed over this period must be approximately zero.

The period between sunset (austral / autumnal vernal equinox sunset times: 1830 and 1813) and 2100 is therefore the inverse of the flush period between 0600-0900, with the CO<sub>2</sub> ~~mixing-ratiomole fraction~~ in the control volume *lower* than ambient which, in the absence of respiratory carbon production, would facilitate downward turbulent transfer of carbon. The increase in the storage term after sunset (Figure 10) partly reflects the 'refilling' of the control volume reservoir following daytime

drawdown, which may contribute to the observation that nocturnal  $S_c$  reached its highest value when  $u_* > u_{*th}$ . While the effect on annual NEE may be minor, it violates the assumption in the  $u_*$ -correction approach that only the turbulent flux term is significant when  $u_* > u_{*th}$ . The implications of this for sites without storage measurements are discussed below.

### 5 3.4 Effects of neglecting carbon storage on physiological interpretation of data

From the perspective of deriving annual NEE sums, it is daytime rather than nocturnal measurements that are more critical; applying the  $u_*$  correction to either  $F_c$  or  $F_c + S_c$  resulted in similar estimates of nocturnal NEE on average. But the effects of neglecting  $S_c$  depend on time of night.  $F_c$  underestimated NEE following sunset, even where  $u_* > u_{*th}$  for  $F_c$ .

10 However, a secondary nocturnal problem is recognised in the literature: when  $u_*$  increases following extended calm periods, stored  $CO_2$  is vented from the control volume, which artificially inflates  $F_c$  relative to the true source term, the extent of which effect will depend on the importance of advection (Aubinet et al., 2012). When a  $u_*$  threshold is imposed, such periods are likely to be included in the retained data. This has the opposite effect to the early-evening effect, and is more likely to be problematic later in the evening when stable stratification and substantial storage of respired carbon is more

15 likely. Both effects were observed in the nocturnal progression of  $S_c$  for periods when  $u_* > u_{*th}$  (Figure 13):  $S_c$  was on average  $> 0$  for the first 4–6 hours after sunset and  $< 0$  afterwards. On balance, the effect in this study was to increase slightly the estimation of ER. This explains why  $S_c$  was slightly positive when  $u_* > u_{*th}$  in Figure 3, and why  $F_c$  was slightly lower than  $F_c + S_c$  at night (primarily in the early evening) in Figure 9b.

20 However, given that temperature decreases over the evening, this suggests that the slope of temperature response functions will be slightly increased for  $F_c + S_c$  versus  $F_c$  alone. Given that the optimisation procedure minimises the prediction error, this may not have a large quantitative effect averaged over the evening, but interpretation of system response to temperature is distorted. Moreover, extrapolation beyond the parameterisation domain (e.g. estimation of daytime ER) may result in substantial error because distortion of function parameters (e.g.  $E_o$  and  $rb$  in equation 5) will potentially result in systematic

25 error (because the function optimised using  $F_c$  will underestimate NEE at high temperatures). Any bias in estimated daytime ER will then necessarily propagate to estimation of GPP (commonly calculated as  $NEE - ER$ ). Because these errors are offsetting, this is not likely to have a large effect on annual NEE estimates.

Similar distortion of response to insolation occurs during the day. The addition of  $S_c$  substantially affects diurnal NEE

30 dynamics, particularly during the morning, which affects the interpretation of the controls on NEE. For example, Figure 14 shows the difference in radiation use efficiency (RUE – here simply defined as the ratio of mean NEE to mean insolation) during daylight hours for  $F_c$  alone versus  $F_c + S_c$ . RUE was higher in the early morning, and declined more sharply, when

NEE =  $F_c + S_c$ . Such declines in RUE are often associated with stomatal response to increasing VPD, and so the importance of this driver may be missed or minimised when  $S_c$  is not measured.

Application of light-response function analysis to daytime data to extract either photosynthetic or respiratory parameters is problematic in the absence of storage measurements because  $F_c \neq$  NEE during most of the day. The estimation of ER (and quantum efficiency) derived from light response function analysis (e.g. Gilmanov et al., 2003; Lasslop et al., 2010) is strongly dependent on the magnitude of observed NEE when insolation is low (sunrise and sunset), and thus the effect of neglecting the storage term may be particularly distorting to these parameters (Aubinet et al., 2012).

### 3.5 Sources of uncertainty

One of the largest sources of uncertainty in annual NEE estimates is expected to derive from uncertainty in  $u_{*th}$  (Barr et al., 2013; Papale, 2006). We propagated this uncertainty to annual NEE (for both  $F_c$  and  $F_c + S_c$ ) by filtering and gap-filling the data using the lower and upper bounds of the 95% confidence interval (CI) of the normally distributed population ( $N = 10^3$ ) of  $u_{*th}$  derived from CPD (Table 3). Much larger effects were evident for the lower uncertainty bound ( $\mu - 2\sigma$ , where  $\mu$  is the best estimate for  $u_{*th}$  and  $\sigma$  is the standard deviation), which is to be expected because systematic errors in nocturnal flux measurement occur at low  $u_*$ . However, there should be no systematic variation in NEE when  $u_* > u_{*th}$ . The direct effect of the upper uncertainty bound ( $\mu + 2\sigma$ ) is expected to be minimal. While the reduction of nocturnal data availability for higher  $u_{*th}$  is expected to increase parameter imputation uncertainty, the effect here was minor, with annual NEE for  $u_{*th} = \mu$  and  $u_{*th} = \mu + 2\sigma$  differing by  $<10 \text{ gC m}^{-2} \text{ a}^{-1}$  in all years.

The uncertainty in  $u_{*th}$  was greater for  $F_c + S_c$  than for  $F_c$  alone due to the additional random error inherent in  $S_c$  (see Finnigan, 2006 for further discussion), which feeds into the change-point detection process. This propagated to larger uncertainty in the lower bound for annual NEE ( $50\text{--}75 \text{ gC m}^{-2} \text{ a}^{-1}$  compared to  $20\text{--}40 \text{ gC m}^{-2} \text{ a}^{-1}$  for  $F_c$ ), despite the effect of  $u_*$  correction for  $F_c$  was almost double that for  $F_c + S_c$ . This is because for  $F_c + S_c$ , the lower bound of the  $u_{*th}$  uncertainty is below the 1<sup>st</sup> percentile of the nocturnal data, such that the full effect of advection was propagated to annual NEE but for  $F_c$  alone, it was closer to the 40<sup>th</sup> percentile, such that only a small proportion of storage and advection occurred in the interval between  $u_{*th} = \mu$  and  $u_{*th} = \mu - 2\sigma$ .

While the lower bound uncertainty due to  $u_*$  threshold determination error was therefore of a magnitude only slightly smaller than the bias introduced by applying the  $u_*$  correction in the absence of storage measurements, the uncertainty resulted in an increase in the potential uptake of carbon, whereas the bias was of opposite sign. Thus the uncertainty ranges for  $F_c$  and for  $F_c + S_c$  do not overlap. In this sense, the uncertainty range for  $F_c$  alone is meaningful only in a very narrow sense (i.e. in terms of what has been measured as opposed to true source / sink uncertainty), and should not be reported as determining the

uncertainty in annual NEE. We have determined that the annual NEE estimate derived from  $F_c + S_c$  is not contained within the  $u_{\text{th}}$  uncertainty interval for annual NEE derived from  $F_c$  alone.

It should also be noted that the large lower-bound uncertainty in annual NEE derived from  $F_c + S_c$  is very likely overestimated. If the lack of correlation between  $u_{\text{a}}$  and  $F_c$  above  $u_{\text{th}}$  indicates that additional terms in the mass balance are negligible, then  $S_c$  should approach zero at  $u_{\text{th}}$ . Given that this is what we observed (Figure 3), and the measurement system (and associated measurement errors) for  $S_c$  is independent of that for  $F_c$ , this is an independent validation of  $u_{\text{th}}$ . It is not clear how such information might be used in the context of frequentist statistical analysis, but it strongly suggests that the uncertainty bounds for NEE that include the effects of  $u_{\text{th}}$  uncertainty presented here are unrealistically large.

The NEE uncertainty contribution of combined random and model error was small ( $<25 \text{ gC m}^{-2} \text{ a}^{-1}$ ) by comparison (Table 3),  $< 10\%$  of annual NEE for each year. Model error was greater than random uncertainty for both day and night, although the difference was larger at night. This is because the majority of nocturnal observational data was removed by the  $u_{\text{a}}$  correction (thereby reducing random error) and replaced with model data. Given that the signal:noise is lower at night, model error is expected to be relatively large. The method used calculates and compounds observation – model data differences, but the observational data already contain random error so this necessarily inflates the propagated model uncertainty above errors associated with missing driver information or systematic measurement error. This problem is partially propagated to the daytime because random error contributes to uncertainty in the nocturnally derived parameters of equation 5, which are then used to calculate the ER component of daytime NEE (equation 6). Moreover, the light-response function estimates are affected by daytime random error, which explains why model error was relatively large comparable to nocturnal values during the day despite 70-90% of the data being retained.

Annual NEE uncertainty due to random and model error was approximately 20% greater for  $F_c + S_c$  than for  $F_c$  alone, due to additional random error to that for  $F_c$  arising from the addition of  $S_c$  (Figure 11). This is largely nocturnally determined because the storage term is smaller and less variable during the daytime when fluxes are largest. The increased annual uncertainty of  $F_c + S_c$  is largely due to higher model uncertainty, which most likely reflects the propagation of random error to model uncertainty through effects on non-linear parameter estimation.

The interdependence of model and random error technically renders invalid the assumption of independence in equation 9. However, the effect is to increase rather than to decrease uncertainty, which is small. While there are methods to separate model and random components (see Dragoni et al., 2007), this generally requires co-location of two instrument arrays. However, the daily differencing procedure we used is known to overestimate error by up to a factor of 2 (Billesbach, 2011; Dragoni et al., 2007) due both to potential wind-dependent temporal variations in source / sink strength (which may

materially affect annual NEE estimates; Griebel et al., 2016) and because some signal is included in the differencing procedure.

It should be emphasised that there are numerous sources of uncertainty that have not been quantified here. Perhaps most important of these is systematic errors in the measurements themselves, which may be an extremely important source of true uncertainty (Lasslop et al., 2008). Thus the uncertainties reported here for  $F_e + S_e$  also should not be formally interpreted as total uncertainty in the true source / sink term, but as the uncertainty contributed by a subset of quantifiable errors.

### 34 Conclusions

We used a simple method to infer advection from measurements combined with a simple and widely used empirical respiration model. Even at our very flat site, approximately 40% of flux underestimation was attributable to advection. Observation of reductions in storage at lower levels (within 8m of the surface) in response to declining  $u_*$  indicate that the most likely advective mechanism is terrain drainage flows. High resolution measurements of wind directions within the control volume would be invaluable for directly detecting the presence of terrain-aligned flows, and are planned for this site.

Given that this site is relatively ideal, this follows earlier findings that drainage-induced advection is likely to affect most sites to some degree. Where the NEE time series consists of turbulent flux and storage, nocturnal correction for these effects should be made under all circumstances. The level of bias incurred in the absence of storage measurements is contingent on the contribution of advection to the nocturnal mass balance. Nocturnal  $u_*$  correction is not advisable if the dominant term in nocturnal flux underestimation is expected to be storage (as found in this study), because the biasing effect of correcting the nocturnal component without the counterbalancing daytime component is larger than correcting for the smaller advection term. At sites with severe advection problems, neglect of storage is relatively less important, because if  $CO_2$  is drained away nocturnally, the early-morning venting that causes bias following nocturnal correction is small.

But this contingency underscores the intractable nature of the problem: the relative contributions of storage and advection to the nocturnal mass balance cannot be quantitatively assessed in the absence of profile measurements. Moreover, even at sites where drainage flows are known to regularly occur at night, it is likely that shear-induced turbulence penetrates below canopy only under strong winds; the rarity of such conditions may result in the rejection of an unacceptably large number of data. Where this is the case, profile measurements are required to increase the proportion of available nocturnal data because storage increases prior to the onset of drainage flows, which only occur once the cooling air mass adjacent to the surface achieves sufficient density to overcome friction and begin to flow (Van Gorsel et al., 2007).

Storage measurements nonetheless introduce some complications for data interpretation. The additional random error in nocturnal storage measurements increases uncertainty in  $u_*$  threshold and, correspondingly, annual NEE (uncertainties due to direct random observation error and imputation error were small by comparison). But as we have argued, the lower bound uncertainty for  $u_*$  threshold is unrealistic, since the storage term on average approaches zero at the  $u_*$  threshold. This behaviour is expected if the central  $u_*$  threshold estimate from change point detection is approximately correct. Even if these uncertainties are considered accurate, when propagated to annual NEE, the resulting uncertainty intervals for the sum of turbulent flux and storage versus turbulent flux alone do not overlap. This indicates that biases are not subsumed within (quantified) uncertainties; effectively, profile measurements reduce precision and increase accuracy of annual NEE estimates.

We therefore believe that for both OzFlux and Fluxnet, the installation of profile systems for sites with trees (woodlands, forests, savannas) is extremely important to ensure that both determination of annual carbon exchange and interpretation of ecosystem processes are accurate. At the very least, the issues explored here need to be taken into consideration during data analysis, and alternative methods of estimating uncertainties at sites without profile systems need to be developed. For sites under the auspices of the Integrated Carbon Observation System (ICOS: [www.icos-ri.eu](http://www.icos-ri.eu)), profile systems are mandatory; while this is not yet the case for OzFlux and Fluxnet, for accurate estimates of annual NEE, profile systems are vital.

We used a simple method to infer advection from measurements combined with a simple and widely used empirical respiration model. Even at our very flat site, approximately 40% of flux underestimation was attributable to advection, which was consistent with previous findings that advection is likely to be important at most sites (Aubinet, 2008). While the mechanism for advection could not be confirmed, the behaviour of the profile measurements suggested horizontal advection associated with from landscape drainage flows. High resolution measurements of wind directions within the control volume would be invaluable for directly detecting the presence of terrain-aligned flows. Such measurements are planned for this site.

We demonstrated the importance of profile measurements for improved measurements of annual carbon balances. The nocturnal biases introduced by the neglect of storage measurements altogether were up to 25% of annual NEE at our study site and were larger than the actual correction applied when storage measurements were included, so that less bias would ensue if the correction were not applied. Use of the point-based storage estimate caused large errors in parameter estimation and thus inconsistent effects on gap-filled annual sums of NEE.

The level of bias from neglecting storage is likely to depend on the site, contingent on the importance of advection. At sites without profile measurements, nocturnal  $u_*$  correction is not advisable if the dominant term in nocturnal flux underestimation is expected to be storage because the effect of correcting the nocturnal component without the counterbalancing daytime component is larger than correcting for the smaller advection term. At sites with severe advection

problems, nocturnal correction is strongly recommended, and neglect of profile measurements is likely to be relatively less important. This is because if  $\text{CO}_2$  is drained away nocturnally, the early morning venting that causes bias following nocturnal correction is small.

5 This contingency underscores the intractable nature of the problem: the relative contributions of storage and advection to the nocturnal mass balance cannot be quantitatively assessed in the absence of profile measurements. Moreover, even at sites where drainage flows are known to regularly occur at night, it is likely that shear-induced turbulence penetrates below canopy only under strong winds; the rarity of such conditions may result in the rejection of an unacceptably large number of data. Where this is the case, profile measurements are required to increase the proportion of available nocturnal data because  
10 storage increases prior to the onset of drainage flows, which only occur once the cooling air mass adjacent to the surface achieves sufficient density to overcome friction and begin to flow (Van Gorsel et al., 2007).

Storage measurements introduce some complications for data interpretation. The additional random error inherent in nocturnal storage measurements meant that the lower uncertainty bound for  $u_{\text{sub}}$  was approximately zero for each year and,  
15 as such, the full effect of advection was included in the associated annual NEE estimate. However, the profile system provides an independent check of the  $u_{\text{sub}}$ , which, in this case, indicates the central estimate of  $u_{\text{sub}}$  is correct and the lower bound of the uncertainty range is unrealistic. By comparison to uncertainty in  $u_{\text{sub}}$ , random and model error uncertainties in annual NEE were small, and despite the storage term increasing the error induced by approximately 25%, in absolute terms the combined uncertainty was < 10% of annual NEE. Moreover, the reduced annual uncertainty interval obtained when  
20 storage was excluded was largely meaningless because it did not contain the actual estimate of annual NEE when storage was included.

We believe that for both OzFlux and Fluxnet, the installation of profile systems for sites with trees (woodlands, forests, savannas) is extremely important to ensure that both determination of annual carbon exchange and interpretation of  
25 ecosystem processes are accurate. At the very least, the issues explored here need to be taken into consideration during data analysis, and alternative methods of estimating uncertainties at sites without profile systems need to be developed. For sites under the auspices of the Integrated Carbon Observation System (ICOS: [www.icos-ri.eu](http://www.icos-ri.eu)), profile systems are mandatory; while this is not yet the case for OzFlux and Fluxnet, for accurate estimates of annual NEE, profile systems are vital.

30 *Acknowledgements.* This work was made possible by funding from the Australian Research Council (ARC; Linkage Project 'More bang for your carbon buck: carbon, biodiversity and water balance consequences of whole-catchment carbon farming' [LP0990038]). IM would also like to thank Peter Isaac for his invaluable comments during development of the manuscript.

## 45 References

- Aubinet, M.: Eddy covariance CO<sub>2</sub> flux measurements in nocturnal conditions: an analysis of the problem, *Ecol. Appl.*, 18, 1368–1378, 2008.
- Aubinet, M. and Feigenwinter, C.: Direct CO<sub>2</sub> advection measurements and the night flux problem, *Agric. For. Meteorol.*, 150, 651–654, doi:10.1016/j.agrformet.2010.03.007, 2010.
- Aubinet, M., Heinesch, B. and Longdoz, B.: Estimation of the carbon sequestration by a heterogeneous forest: Night flux corrections, heterogeneity of the site and inter-annual variability, *Glob. Change Biol.*, 8(11), 1053–1071, 2002.
- Aubinet, M., Heinesch, B. and Yernaux, M.: Horizontal and Vertical CO<sub>2</sub> Advection In A Sloping Forest, *Bound.-Layer Meteorol.*, 108, 397–417, doi:10.1023/a:1024168428135, 2003.
- Aubinet, M., Feigenwinter, C., Heinesch, B., Laffineur, Q., Papale, D., Reichstein, M., Rinne, J. and Van Gorsel, E.: Nighttime Flux Correction, in *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*, edited by M. Aubinet, T. Vesala, and D. Papale, Springer Netherlands, Dordrecht. [online] Available from: <http://link.springer.com/10.1007/978-94-007-2351-1> (Accessed 12 January 2015), 2012.
- Baldocchi, D. D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, *Glob. Change Biol.*, 9, 479–492, doi:10.1046/j.1365-2486.2003.00629.x, 2003.
- Barr, A. G., Richardson, A. D., Hollinger, D. Y., Papale, D., Arain, M. A., Black, T. A., Bohrer, G., Dragoni, D., Fischer, M. L., Gu, L., Law, B. E., Margolis, H. A., McCaughey, J. H., Munger, J. W., Oechel, W. and Schaeffer, K.: Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies, *Agric. For. Meteorol.*, 171–172, 31–45, doi:10.1016/j.agrformet.2012.11.023, 2013.
- Beringer, J., Hutley, L. B., McHugh, I., Arndt, S. K., Campbell, D., Cleugh, H. A., Cleverly, J., Resco de Dios, V., Eamus, D., Evans, B., Ewenz, C., Grace, P., Griebel, A., Haverd, V., Hinko-Najera, N., Huete, A., Isaac, P., Kanniah, K., Leuning, R., Liddell, M. J., Macfarlane, C., Meyer, W., Moore, C., Pendall, E., Phillips, A., Phillips, R. L., Prober, S., Restrepo-Coupe, N., Rutledge, S., Schroder, I., Silberstein, R., Southall, P., Sun, M., Tapper, N. J., van Gorsel, E., Vote, C., Walker, J. and Wardlaw, T.: An introduction to the Australian and New Zealand flux tower network &ndash; OzFlux, *Biogeosciences Discuss.*, 1–52, doi:10.5194/bg-2016-152, 2016.
- Billesbach, D. P.: Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method, *Agric. For. Meteorol.*, 151(3), 394–405, doi:10.1016/j.agrformet.2010.12.001, 2011.
- Chapin, F. S., III, Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., Wirth, C., Aber, J. D., Cole, J. J., Goulden, M. L., Harden, J. W., Heimann, M., Howarth, R. W., Matson, P. A., McGuire, A. D., Melillo, J. M., Mooney, H. A., Neff, J. C., Houghton, R. A., Pace, M. L., Ryan, M. G., Running, S. W., Sala, O. E., Schlesinger, W. H. and Schulze, E. D.: Reconciling Carbon-Cycle Concepts, Terminology, and Methods, *Ecosystems*, 9, 1041–1050, doi:10.2307/25470403, 2006.
- Dragoni, D., Schmid, H., Grimmond, C. and Loescher, H.: Uncertainty of annual ecosystem productivity estimated using eddy covariance flux measurements, *J. Geophys. Res.-Biogeosciences*, 112, D17102, 2007.
- Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R. and Dolman, H.: Gap filling strategies for defensible annual sums of net ecosystem exchange, *Agric. For. Meteorol.*, 107, 43–69, 2001.



- Finnigan, J.: The storage term in eddy flux calculations, *Agric. For. Meteorol. Adv. Surf.-Atmosphere Exch. - Tribute Marv Wesely*, 136, 108–113, 2006.
- Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R. and Cleugh, H. A.: A Re-Evaluation of Long-Term Flux Measurement Techniques Part I: Averaging and Coordinate Rotation, *Bound.-Layer Meteorol.*, 107, 1–48, 2003.
- 5 Gilmanov, T. G., Verma, S. B., Sims, P. L., Meyers, T. P., Bradford, J. A., Burba, G. G. and Suyker, A. E.: Gross primary production and light response parameters of four Southern Plains ecosystems estimated using long-term CO<sub>2</sub>-flux tower measurements, *Glob. Biogeochem Cycles*, 17, 1071, doi:10.1029/2002gb002023, 2003.
- Goulden, M., Munger, J. W., Fan, S.-M., Daube, B. and Wofsy, S.: Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability., *Science*, 271, 1996a.
- 10 Goulden, M. L., Munger, J. W., Fan, S.-M., Daube, B. C. and Wofsy, S. C.: Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy, *Glob. Change Biol.*, 2, 169–182, doi:10.1111/j.1365-2486.1996.tb00070.x, 1996b.
- Goulden, M. L., Miller, S. D. and da Rocha, H. R.: Nocturnal cold air drainage and pooling in a tropical forest, *J. Geophys. Res.*, 111(D8), doi:10.1029/2005JD006037, 2006.
- 15 Griebel, A., Bennett, L. T., Metzen, D., Cleverly, J., Burba, G. and Arndt, S. K.: Effects of inhomogeneities within the flux footprint on the interpretation of seasonal, annual, and interannual ecosystem carbon exchange, *Agric. For. Meteorol.*, 221, 50–60, doi:10.1016/j.agrformet.2016.02.002, 2016.
- Gu, L., Massman, W. J., Leuning, R., Pallardy, S. G., Meyers, T., Hanson, P. J., Riggs, J. S., Hosman, K. P. and Yang, B.: The fundamental equation of eddy covariance and its application in flux measurements, *Agric. For. Meteorol.*, 152, 135–148, 2012.
- 20 Hollinger, D. Y. and Richardson, A. D.: Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiol.*, 25(7), 873–885, 2005.
- Isaac, P., Cleverly, J., McHugh, I., Van Gorsel, E. and Beringer, J.: OzFlux Data: Network Integration from Collection to Curation, *Biogeosciences Discuss.*, 189, 2016.
- James, S. A. and Bell, D. T.: Leaf orientation in juvenile *Eucalyptus camaldulensis*, *Aust. J. Bot.*, 44, 139–156, 1996.
- 25 Janssens, I. A., Lankreijer, H., Matteucci, G., Kowalski, A. S., Buchmann, N., Epron, D., Pilegaard, K., Kutsch, W., Longdoz, B., Grünwald, T., Montagnani, L., Dore, S., Rebmann, C., Moors, E. J., Grelle, A., Rannik, Ü., Morgenstern, K., Oltchev, S., Clement, R., Guðmundsson, J., Minerbi, S., Berbigier, P., Ibrom, A., Moncrieff, J., Aubinet, M., Bernhofer, C., Jensen, N. O., Vesala, T., Granier, A., Schulze, E. D., Lindroth, A., Dolman, A. J., Jarvis, P. G., Ceulemans, R. and Valentini, R.: Productivity overshadows temperature in determining soil and ecosystem respiration across European forests, *Glob. Change Biol.*, 7, 269–278, doi:10.1046/j.1365-2486.2001.00412.x, 2001.
- 30 Keith, H., Leuning, R., Jacobsen, K. L., Cleugh, H. A., van Gorsel, E., Raison, R. J., Medlyn, B. E., Winters, A. and Keitel, C.: Multiple measurements constrain estimates of net carbon exchange by a *Eucalyptus* forest, *Agric. For. Meteorol.*, 149(3–4), 535–558, doi:10.1016/j.agrformet.2008.10.002, 2009.
- Lasslop, G., Reichstein, M., Kattge, J. and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, *Biogeosciences Discuss.*, 5(1), 751–785, 2008.
- 35

- Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Armeth, A., Barr, A., Stoy, P. and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, *Glob. Change Biol.*, 16(1), 187–208, doi:10.1111/j.1365-2486.2009.02041.x, 2010.
- 5 Law, B. E., Ryan, M. G. and Anthoni, P. M.: Seasonal and annual respiration of a ponderosa pine ecosystem, *Glob. Change Biol.*, 5, 169–182, doi:10.1046/j.1365-2486.1999.00214.x, 1999.
- Lee, X., Finnigan, J. and Paw U, K.: Coordinate Systems and Flux Bias Error, in *Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis*, edited by X. Lee, B. Massman, and B. Law, Kluwer Academic Publishing, Dordrecht, Netherlands., 2004.
- 10 Mahrt, L., Vickers, D., Nakamura, R., Soler, M. R., Sun, J., Burns, S. and Lenschow, D. H.: Shallow drainage flows, *Bound.-Layer Meteorol.*, 101(2), 243–260, 2001.
- Massman, B. and Clement, R.: Uncertainty in eddy covariance flux estimates resulting from spectral attenuation, in *Handbook of Micrometeorology*, edited by X. Lee, B. Massman, and B. Law, pp. 67–99, Kluwer Academic Publishers, Dordrecht., 2004.
- Moncrieff, J., Clement, R., Finnigan, J. and Meyers, T.: Averaging, detrending and filtering of eddy covariance time series, in *Handbook of Micrometeorology*, edited by X. Lee, B. Massman, and B. Law, Kluwer Academic Publishers, Dordrecht., 2004.
- 15 Papale, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, [online] Available from: <http://dspace.unitus.it/handle/2067/1321> (Accessed 22 November 2013), 2006.
- Pereira, J. S., Araújo, C. C. and Borralho, N.: Crown structure of *Eucalyptus globulus* Labill. in a coppiced plantation, in *Plant Response to Stress*, pp. 521–530, Springer. [online] Available from: [http://link.springer.com/chapter/10.1007/978-3-642-70868-8\\_35](http://link.springer.com/chapter/10.1007/978-3-642-70868-8_35) (Accessed 7 March 2016), 1987.
- 20 Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, *Glob. Change Biol.*, 11, 1424–1439, doi:10.1111/j.1365-2486.2005.001002.x, 2005.
- 25 Richardson, A. D. and Hollinger, D. Y.: Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation, and Monte Carlo simulation of model and parameter uncertainty, applied to three simple models, *Agric. For. Meteorol.*, 131, 191–208, doi:10.1016/j.agrformet.2005.05.008, 2005.
- 30 Ruimy, A., Jarvis, P. and Baldocchi, D. : CO<sub>2</sub> fluxes over plant canopies and solar radiation: a review., *Adv. Ecol. Res.*, 26, 1–68, 1995.
- Speckman, H. N., Frank, J. M., Bradford, J. B., Miles, B. L., Massman, W. J., Parton, W. J. and Ryan, M. G.: Forest ecosystem respiration estimated from eddy covariance and chamber measurements under high turbulence and substantial tree mortality from bark beetles, *Glob. Change Biol.*, n/a-n/a, doi:10.1111/gcb.12731, 2014.
- 35 Staebler, R. M. and Fitzjarrald, D. R.: Observing subcanopy CO<sub>2</sub> advection, *Agric. For. Meteorol.*, 122(3–4), 139–156, doi:10.1016/j.agrformet.2003.09.011, 2004.

Van Gorsel, E. V. A., Leuning, R. A. Y., Cleugh, H. A., Keith, H. and Suni, T.: Nocturnal carbon efflux: reconciliation of eddy covariance and chamber measurements using an alternative to the  $u^*$ -threshold filtering technique, *Tellus B*, 59, 397–403, doi:10.1111/j.1600-0889.2007.00252.x, 2007.

5 Webb, E., Pearman, G. and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, *Q. J. R. Meteorol. Soc.*, 106, 85–100, 1980.

Yang, B., Hanson, P. J., Riggs, J. S., Pallardy, S. G., Heuer, M., Hosman, K. P., Meyers, T. P., Wullschleger, S. D. and Gu, L.-H.: Biases of  $\text{CO}_2$  storage in eddy flux measurements in a forest pertinent to vertical configurations of a profile system and  $\text{CO}_2$  density averaging, *J. Geophys. Res.*, 112(D20), doi:10.1029/2006JD008243, 2007.

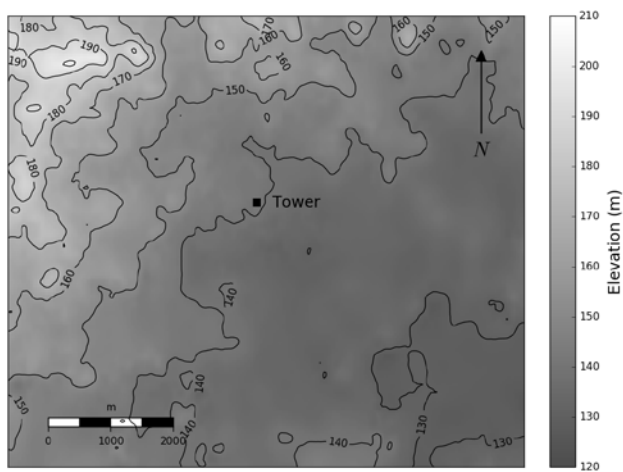


Figure 1: topography of terrain surrounding tower.

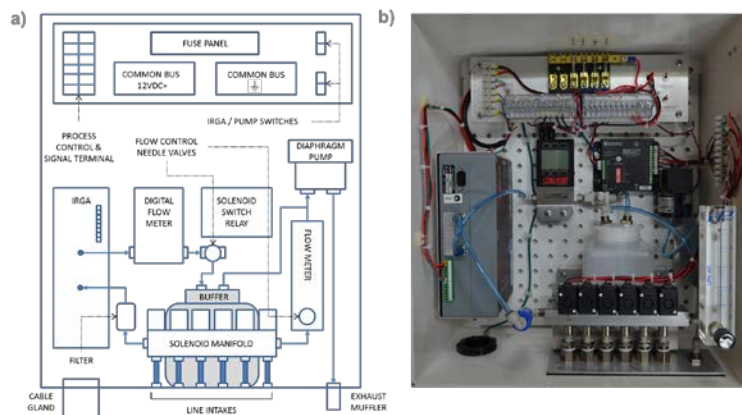


Figure 2: a) schematic, and; b) photographic layout of profile gas analysis system (note that a Licor LI820 is pictured here; an LI840 was used for measurement).

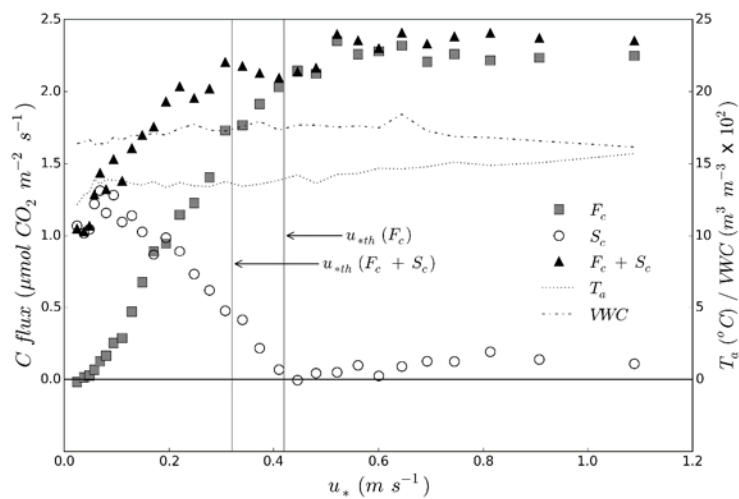


Figure 3: LH axis) dependence of mean measured nocturnal carbon mass balance components (turbulent flux [ $F_c$ ], storage [ $S_c$ ] and  $F_c + S_c$ ) on friction velocity ( $u_*$ ); RH axis) air temperature at EC instrumentation height (36m) and volumetric soil moisture content at 10cm. Vertical lines denote  $u_{*th}$  for both  $F_c$  and  $F_c + S_c$ , as labelled.

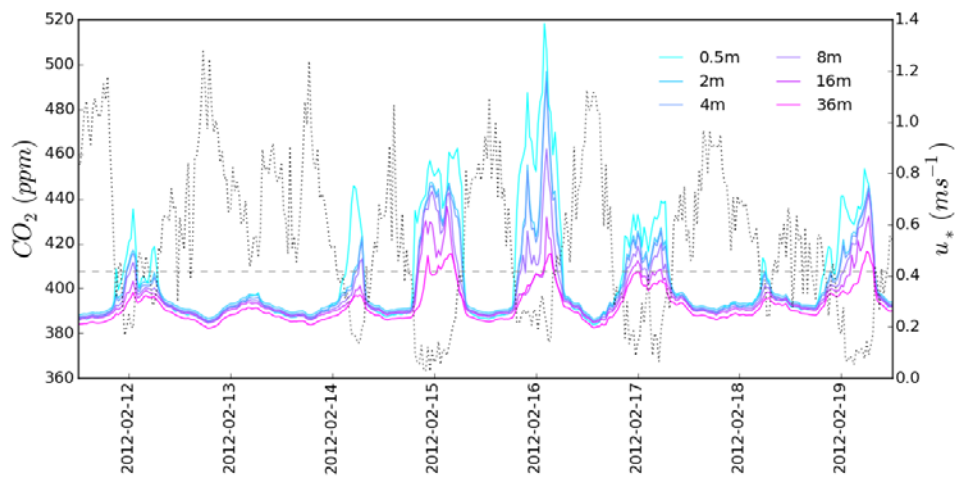


Figure 4: LH axis) profile system time series (date labels correspond to midnight) of CO<sub>2</sub> ~~mixing ratio~~ mole fraction (coloured solid lines); RH axis) time series of  $u_*$  measured at 36m (grey dotted line). Note horizontal dashed line marks  $u_{*_{crit}}$  for  $F_c$ .

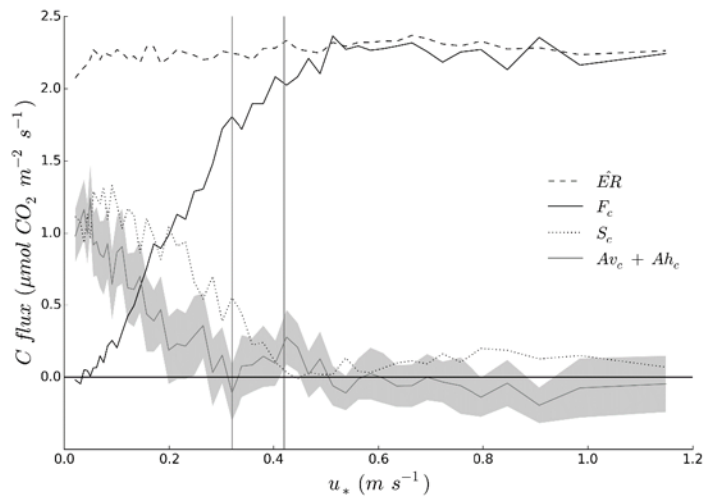


Figure 5: dependence of measured ( $F_c$ ,  $S_c$ ), model-estimated ( $\bar{E}R$ ) and inferred ( $Av_c$ ,  $Ah_c$ ) mass balance components on friction velocity ( $u_*$ ; the grey shaded area represents the 95% confidence interval for the sample bin mean of the inferred advection components). Vertical lines denote  $u_*^{\text{th}}$  for both  $F_c$  and  $F_c + S_c$ , as per Figure 3.

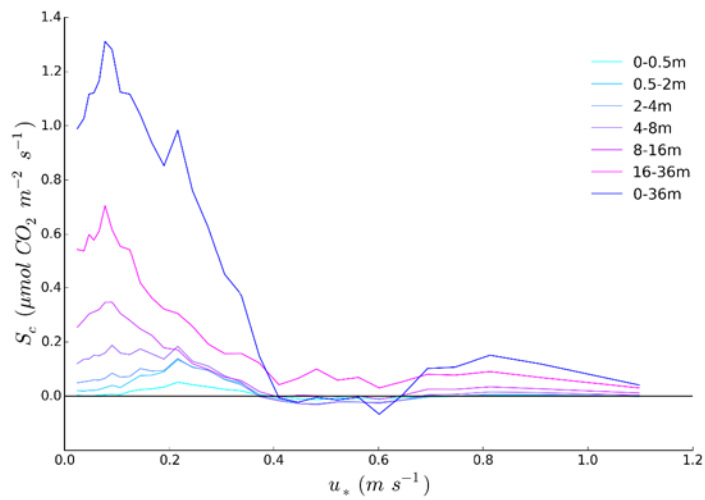


Figure 6: dependence of measured storage components on friction velocity for individual layers.



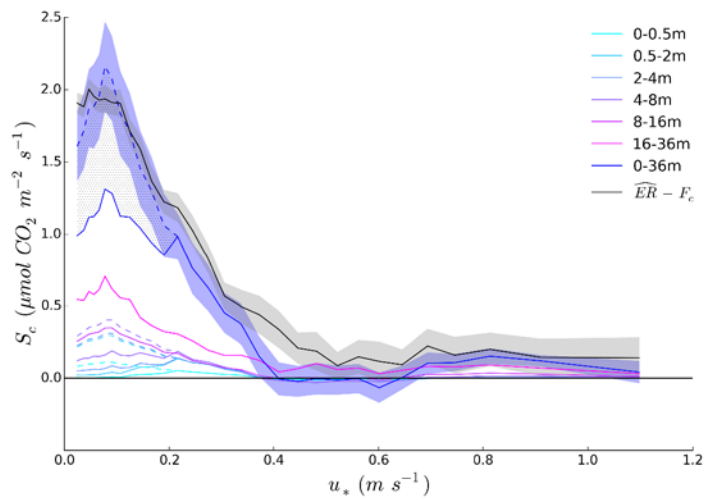


Figure 7: dependence of corrected storage components and  $R_c^{\delta} - F_c$  on friction velocity (dashed lines represent corrected storage estimates; stippled region represents difference between measured and corrected 0-36m storage; shaded regions represent 95%CI for the  $u_*$  bin mean).

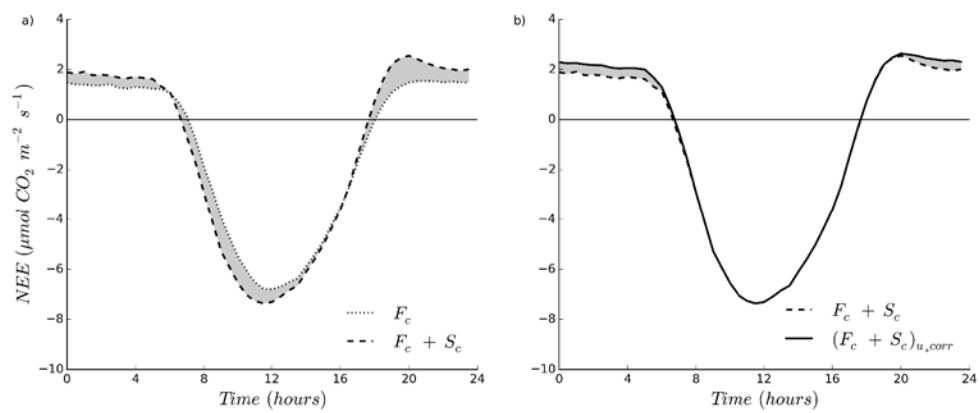


Figure 8: effects of (a) storage addition and (b)  $u_+$  correction on diurnal mean NEE dynamics.

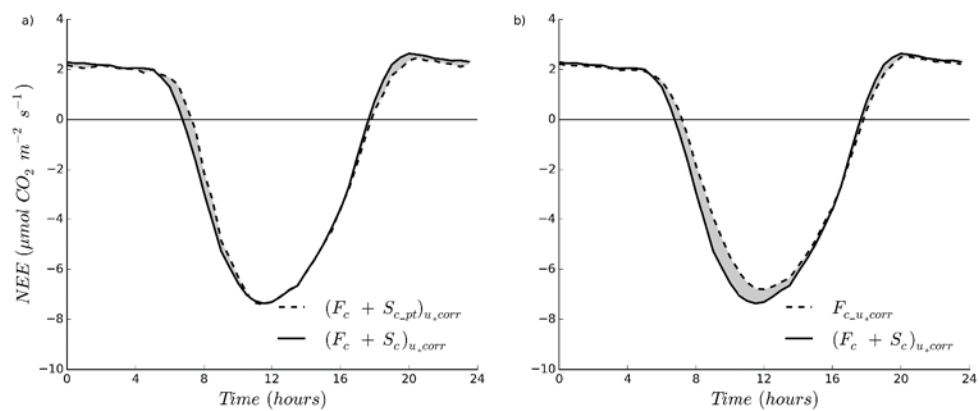
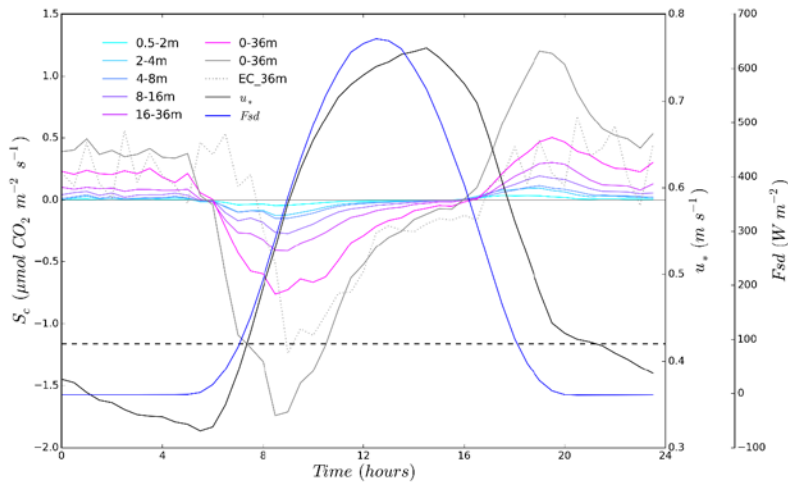


Figure 9: effects of (a) using storage term derived from eddy covariance IRGA, and (b) neglecting storage, on diurnal mean NEE dynamics.



5 Figure 10: mean diurnal cycle of aggregate and component  $S_c$  (LH axis; horizontal solid line represents mean  $S_c$ ) and  $u_*$  (RH axis; horizontal dashed line represents change point-derived nocturnal  $u_*$  threshold); ~~vertical dotted lines represent approximate sunrise and sunset times (using insolation threshold =  $5 \text{ W m}^{-2}$ ; note that day length as indicated by insolation is slightly greater than 12 hours slight asymmetry between day and night length is due to missing data for several months during winter / spring 2013, thereby slightly biasing the data towards longer photoperiod).~~

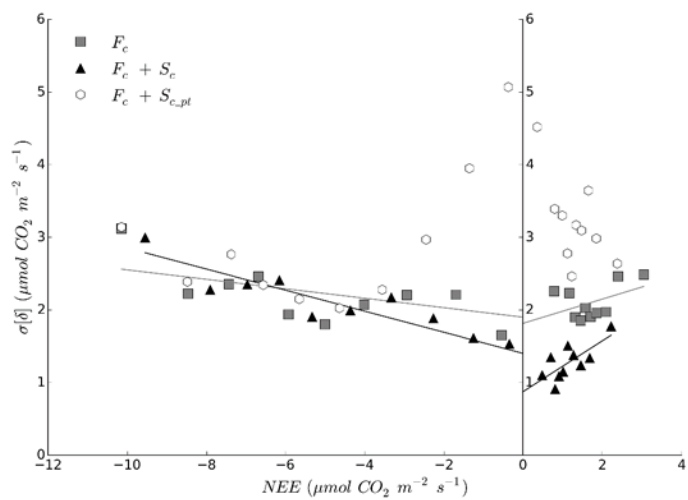


Figure 11: standard deviation of estimated random error ( $\sigma[\delta]$ ) as a function of flux magnitude for turbulent flux ( $F_c$ ), turbulent flux plus profile-based storage estimate ( $F_c + S_c$ ) and turbulent flux plus point-based storage estimate ( $F_c + S_{c,pt}$ ).

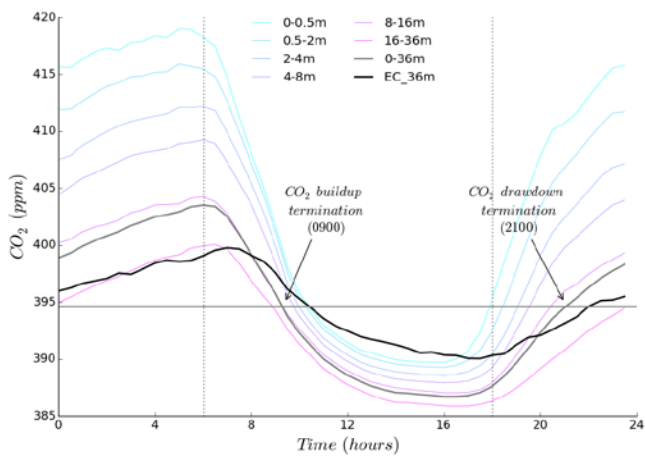


Figure 12: mean diurnal cycle of  $\text{CO}_2$  mixing-ratiomole fraction measured by profile system for given layers (coloured lines; grey line represents 0-36m height-integrated control volume mean) and by eddy covariance infra-red gas analyser at 36m (black line). Note that due to minor discrepancies between instruments, mixing-ratiomole fractions for the eddy covariance IRGA are here baselined to the height-integrated profile IRGA multi-annual mean (394.6ppm).

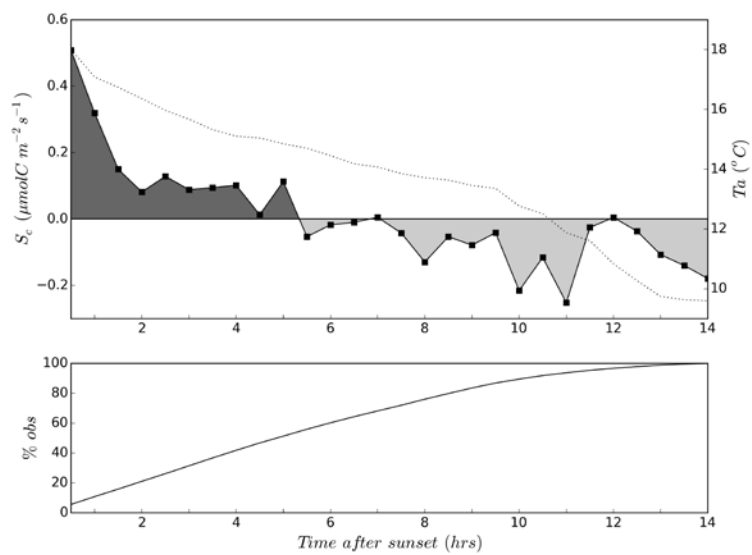


Figure 13: dependence of  $S_e$  (including only data where  $u_* > u_{*th}$ ) on time after sunset (upper panel; dotted line is air temperature); cumulative percentage of total nocturnal  $S_e$  observations (lower panel).

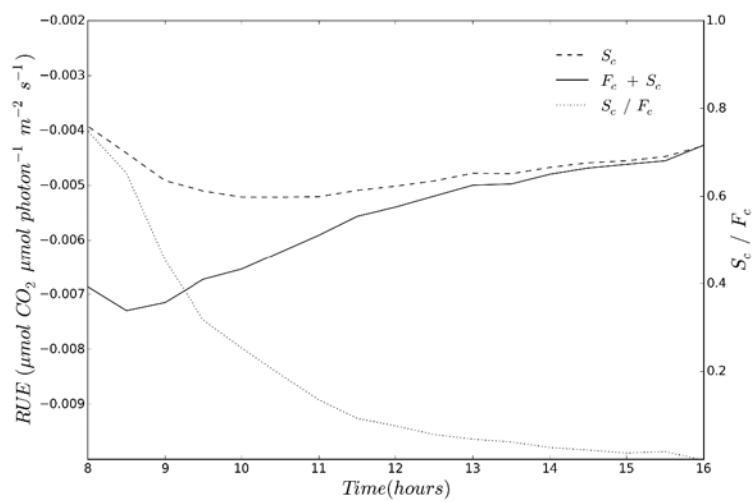


Figure 14: LH axis) effects of addition of  $F_e$  to  $S_e$  on radiation use efficiency (RUE); RH axis)  $S_e$  as a proportion of  $F_e$ .



Table 1: site characteristics

Latitude, longitude (° dec.)	-36.673215, 145.029247
Slope (°)	<1
Aspect	N/A
Dominant overstorey (>90%*) species	<i>Eucalyptus microcarpa</i>
Dominant understorey (>90%*) species	<i>Cassinia arculeata</i>
Mean canopy height ±SD (m)	15.3±6.4
Leaf area index (m <sup>2</sup> m <sup>-2</sup> )	~1.1
Mean annual temperature (°C)	15.9
Mean annual precipitation (mm)**: long term (1971-2000)	560

\* By biomass (although also by number of individuals in the case of the overstorey)  
\*\* From nearest long-term rainfall measurement site (Mangalore Airport; Bureau of Meteorology station ID 088109)

Table 2: site instrumentation

Measurement	Instrument	Manufacturer
Wind vectors / virtual temperature	CSAT3	Campbell Scientific Instruments
Radiation components	CNR4	Kipp and Zonen
CO <sub>2</sub> <del>mixing-ratio</del> mole fraction (eddy covariance)	LI7500	Licor Biosciences
CO <sub>2</sub> <del>mixing-ratio</del> mole fraction (profile)	LI840	Licor Biosciences
Temperature / humidity	HMP45C	Vaisala
Wind speed / direction (profile)	Wind Sentry Set	RM Young
Barometric pressure	PTB110	Vaisala
Volumetric soil water content	CS616	Campbell Scientific Instruments
Soil heat flux	HFP01	Hukseflux
Soil temperature	TCAV	Campbell Scientific Instruments
Data logging	CR3000	Campbell Scientific Instruments

Table 3: lower 95%CI bound ( $\mu - 2\sigma$ ), mean ( $\mu$ ), and upper 95%CI bound ( $\mu + 2\sigma$ ) of Gaussian PDF of  $u^*_{th}$  (derived from change point detection of bootstrapped samples - see Methods), data percentile (i.e. percentage data excluded for each  $u^*_{th}$ ) and resulting imputed annual estimate of NEE. Note: i)  $\mu - 2\sigma$  set to zero if  $< 0$  (e.g.  $F_c + S_c$  in 2013); ii) respiration and light response function analysis could not find a solution for  $F_c + S_c$  in 2013 when  $u^* = 0.73$  (insufficient data for robust statistical fit).

Year	Condition	F <sub>c</sub>			F <sub>c</sub> + S <sub>c</sub>		
		u* (m s <sup>-1</sup> )	Data percentile	Annual NEE (gC m <sup>-2</sup> a <sup>-1</sup> )	u* (±95% CI)	Data percentile	Annual NEE (gC m <sup>-2</sup> a <sup>-1</sup> )
2012	$\mu - 2\sigma$	0.26	47	-323.1	0.01	<1	-450.6
	$\mu$	0.39	60	-299.0	0.30	50	-381.5
	$\mu + 2\sigma$	0.52	72	-304.7	0.59	77	-387.1
2013	$\mu - 2\sigma$	0.19	38	-285.2	0	0	-364.1
	$\mu$	0.40	61	-250.4	0.32	53	-313.3
	$\mu + 2\sigma$	0.61	79	-251.4	0.73	87	-
2014	$\mu - 2\sigma$	0.23	43	-433.4	0.02	<1	-552.7
	$\mu$	0.42	63	-395.2	0.32	53	-483.2
	$\mu + 2\sigma$	0.61	79	-398.8	0.62	80	-481.9

Table 4: gap-filled annual NEE (gC m<sup>-2</sup> a<sup>-1</sup>) for 2012-2014 obtained following different data treatment (F<sub>c</sub>: turbulent flux only; F<sub>c,u\*<sub>corr</sub></sub>: turbulent flux with low u\* conditions removed; F<sub>c</sub> + S<sub>c,pt</sub>: summed turbulent flux and point-based storage estimate; (F<sub>c</sub> + S<sub>c,pt</sub>)<sub>u\*<sub>corr</sub></sub>: summed turbulent flux and point-based storage estimate with low u\* conditions removed; F<sub>c,S<sub>c</sub></sub>: summed turbulent flux and profile-based storage estimate; (F<sub>c</sub> + S<sub>c</sub>)<sub>u\*<sub>corr</sub></sub>: summed turbulent flux and profile-based storage estimate with low u\* conditions removed).

Year	F <sub>c</sub>	F <sub>c,u*<sub>corr</sub></sub>	F <sub>c</sub> + S <sub>c,pt</sub>	(F <sub>c</sub> + S <sub>c,pt</sub> ) <sub>u*<sub>corr</sub></sub>	F <sub>c</sub> + S <sub>c</sub>	(F <sub>c</sub> + S <sub>c</sub> ) <sub>u*<sub>corr</sub></sub>
2012	-463	-301	-489	-375	-447	-382
2013	-403	-267	-492	-362	-388	-338
2014	-572	-393	-586	-431	-551	-480