

# An Investigation of the Eddy-Covariance Flux Imbalance in a Year-Long Large-Eddy Simulation of the Weather at Cabauw

Jerôme Schalkwijk<sup>1</sup> · Harm J. J. Jonker<sup>1</sup> ·  
A. Pier Siebesma<sup>1,2</sup>

Received: 10 January 2015 / Accepted: 16 February 2016 / Published online: 26 March 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The low-frequency contribution to the systematic and random sampling errors in single-tower eddy-covariance flux measurements is investigated using large-eddy simulation (LES). We use a continuous LES integration that covers a full year of realistic weather conditions over Cabauw, the Netherlands, and emulate eddy-covariance measurements. We focus on the daytime flux imbalance, when the turbulent flux is sufficiently resolved. Averaged over the year, daytime single-tower eddy-covariance flux measurements lead to a significant systematic underestimation of the turbulent flux. This underestimation depends on the averaging period and measurement height. For a 3600-s averaging period at 16-m height, the systematic underestimation reduces to a few percent, but for 900-s averaged tall-tower measurements at 100-m height, the fluxes are systematically underestimated by over 20 %. The year-long dataset facilitates an investigation into the environmental conditions that influence the eddy-covariance flux imbalance. The imbalance problem is found to vary widely from day to day, strongly dependent on the flow regime. In general, the imbalance problem reduces with increased mean wind speed, but days having the largest imbalance (over twice the average) are characterized by roll vortices that occur for average wind speeds, typically having a boundary-layer height ( $z_i$ ) to Obukhov length ( $L$ ) ratio of  $10 < -z_i/L < 100$ .

**Keywords** Eddy-covariance method · Imbalance problem · Large-eddy simulation · Spatial representativeness of fluxes

---

✉ Jérôme Schalkwijk  
J.Schalkwijk@tudelft.nl

<sup>1</sup> Department of Geoscience and Remote Sensing, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

<sup>2</sup> R&D Section on Weather & Climate Modeling, Royal Netherlands Meteorological Institute, PO Box 201, 3730 AE De Bilt, The Netherlands

## 1 Introduction

Turbulent fluxes, responsible for the vertical transport of heat, humidity, momentum and trace gases through the atmosphere and between surface and atmosphere, are among the most important boundary-layer processes that have to be modelled in numerical weather and climate models. The magnitude and partitioning of available energy over the sensible and latent fluxes are key parameters in determining the properties and height of the boundary layer. They also strongly influence the formation and behaviour of convective clouds, which form one of the largest uncertainties in climate projections (e.g., [Stephens 2005](#)).

During the past decades, the eddy-covariance (EC) method has become a widespread method to measure turbulent fluxes, and in this study we investigate observations from a single-tower EC flux measurement station. Fundamentally, there are two main limitations to such a measurement station. The first limitation is related to the finite averaging period over which fluxes are determined. Statistical convergence requires sufficient independent samples. Since the measurement station is stationary, the averaging period determines the sampling size of eddies that pass over the station (driven by mean flow). With weak mean winds, the station may sample mostly coherent eddies, i.e., may gather samples that are insufficiently statistically independent ([Mahrt 1998](#)). These large-scale coherent eddies are a result of the self-organization of turbulence (e.g., [Schmidt and Schumann 1989](#)). The sampling problem related to these eddies may be amplified in cases where the eddies elongate in the downwind direction to form roll vortices ([LeMone 1973](#)).

Secondly, a single-tower measurement is representative only of a limited area. The ‘field of view’ or ‘footprint’ of tower EC measurements is limited ([Schmid 1994](#)), which is especially problematic if the surface is heterogeneous, causing a secondary circulation that is missed by tower observations ([Desjardins et al. 1997](#)). Besides these two limitations, there are also a number of potential practical difficulties with EC flux towers, related to, for instance, tower placement and sensor alignment (see e.g., [Mahrt 1998, 2010](#), for an overview).

In this study, however, we limit ourselves only to the first point, that is, the sampling errors that occur due to averaging over a finite period of time. One may expect that these sampling errors average out if enough data are available. It turns out, however, that they also cause a systematic error, or bias. For instance, [Lenschow et al. \(1994\)](#) analytically showed that flux measurements for a finite averaging period always possess a negative bias (i.e., underestimation) related to the low-frequency flux contribution.

A related problem in measurement campaigns concerning surface fluxes is that the measured sum of sensible and latent heat fluxes does not match the available energy (e.g., [Leuning et al. 2012](#)). Several previous studies (e.g., [Twine et al. 2000](#); [Foken 2008](#)) have suggested that the closure problem in the observed surface energy budget is larger than what can be explained by experimental uncertainties. The effect of the negligence of storage terms (e.g., [Jacobs et al. 2008](#)) as well as sampling errors in the EC technique (e.g., [Foken et al. 2006](#); [Hendricks Franssen et al. 2010](#)) have been suggested to contribute to the closure problem.

[Kanda et al. \(2004\)](#) recognized that EC problems related to the turbulent flow can be isolated and investigated in a large-eddy simulation (LES). The advantage of LES in this context is twofold, since in LES (1) EC fluxes can be determined without measurement errors, and (2) the full three-dimensional turbulent field, and hence the most representative (‘true’) flux, is known and can be compared with the EC flux. Using LES, [Kanda et al. \(2004\)](#) emulated EC observations in a clear convective boundary layer, and found a systematic underestimation in single-tower EC fluxes ranging between roughly 5 and 25 % (at a measurement height of 100 m) due to turbulent organized structures, depending on wind speed. They recommended

the use of horizontally distributed observational networks to tackle the problem. [Steinfeld et al. \(2007\)](#) continued their investigation at a higher spatial resolution and closer to the surface, and concluded that whereas EC imbalance due to turbulent self-organization is always significant, it has a strong height dependence. They find that the imbalance decreases to roughly 5 % below approximately 10 m. The flux imbalance problem is further conceptualized in a LES study by [Huang et al. \(2008\)](#), who decompose the EC imbalance into bottom-up and top-down components. This allows them to investigate the imbalance as a function of non-dimensionalized numbers related to velocity scales and the source location.

LES studies, however, are typically limited to highly idealized cases, in (near-) steady state, without a diurnal cycle and in the absence of mesoscale motions. Although such an idealization often clarifies the investigation, the selected idealized case is not necessarily representative of the mean conditions. Indeed, such an investigation might not even include those cases in which the problem is most pronounced ([Neggers et al. 2012](#)). Here, we attempt to remedy this caveat by simulating an especially large range of conditions. By utilizing the graphics processing unit (GPU), we were able to significantly speed up our LES integrations, resulting in a GPU-resident atmospheric LES (GALES, see [Schalkwijk et al. 2012](#)). The simulation speed of GALES allowed us to perform a continuous LES integration of the actual weather conditions at Cabauw, the Netherlands, for the full year 2012 ([Schalkwijk et al. 2015](#)). The resulting dataset, denoted YOGA-2012 (Year of GALES), allows us to consider the EC imbalance problem in a more realistic, statistical sense than is possible in an idealized case study.

The location was chosen to correspond to the Cabauw Experimental Site for Atmospheric Research that allows direct comparison between measurements and observations. The YOGA dataset is introduced in [Schalkwijk et al. \(2015\)](#), which includes a comparison between simulated and observed surface fluxes. However, given the uncertainty in observed fluxes and the inherent lack of information on the systematic error, the observed fluxes are not considered here. This study is therefore limited to the sampling errors within the LES framework, which are known exactly. The varying conditions that are simulated should complement previous case studies by extending the range of simulated conditions. This may provide an idea of the order of magnitude of the EC flux imbalance that one might expect when averaging over a full year. Moreover, through considering the EC flux imbalance throughout the year, we are able to investigate those environmental conditions (e.g., wind speed or stability) that most influence the imbalance.

## 2 Eddy Covariance Fluxes

The theoretical framework of EC fluxes in a LES is described extensively in [Kanda et al. \(2004\)](#) and [Steinfeld et al. \(2007\)](#), therefore we will only provide a short summary. Consider a conserved variable  $\phi = \phi(x, y, z, t)$ , and denote the spatial mean by  $\langle \phi \rangle$  and the temporal mean by  $\bar{\phi}$ , defined as,

$$\bar{\phi}(x, y, z) = \frac{1}{T} \int_T \phi(x, y, z, t) dt, \quad (1)$$

and,

$$\langle \phi \rangle(z, t) = \frac{1}{A} \iint_A \phi(x, y, z, t) dx dy, \quad (2)$$

where  $T$  is the EC averaging period and  $A$  is the LES surface area. Fluctuations from the temporal and spatial mean are denoted by  $\phi'$  and  $\phi''$ , respectively, i.e.,

$$\phi'(x, y, z, t) = \phi(x, y, z, t) - \bar{\phi}(x, y, z), \quad (3)$$

$$\phi''(x, y, z, t) = \phi(x, y, z, t) - \langle \phi \rangle(z, t). \quad (4)$$

At any instance of time, the spatially-averaged vertical flux  $\langle w''\phi'' \rangle$  describes the turbulent transport over the full domain. Since the LES used here (see next section) employs periodic boundary conditions, i.e.,  $\langle w \rangle = 0$ , we find that  $\langle w''\phi'' \rangle = \langle w\phi \rangle$ .

The EC flux determined using a single measurement tower is not based on spatial but on temporal deviations, and can be emulated in the LES by determining  $w'\phi'$  at a given location in the domain. The EC flux imbalance, denoted by  $I_\phi$ , is then given by the difference with the true flux,

$$I_\phi(x, y, z) = \overline{w'\phi'}(x, y, z) - \overline{\langle w''\phi'' \rangle}(z). \quad (5)$$

In LES, only the large eddies (substantially larger than grid spacing) are explicitly resolved, and smaller eddies are parametrized. In the equations above,  $\phi'$  and  $\phi''$  denote *resolved* fluctuations from the temporal and spatial main, respectively. The total turbulent vertical flux  $\Phi_\phi$  of variable  $\phi$  is modeled as the sum of the resolved and sub-filter-scale flux. Sub-filter-scale motions in GALEs are treated through fluxes based on eddy diffusivity, parametrized as a function of sub-filter-scale turbulent kinetic energy (TKE). The total vertical flux becomes,

$$\Phi_\phi(z) = \overline{\langle w''\phi'' \rangle} + \left\langle -K \frac{\partial \phi}{\partial z} \right\rangle, \quad (6)$$

where  $K = c\lambda e^{1/2}$  is the eddy diffusivity, where  $c$  is a constant,  $\lambda$  is a length scale related to the resolution and  $e$  is the sub-filter-scale TKE. The model is described in detail in Heus et al. (2010). By considering only the resolved imbalance, all sub-filter-scale flux loss is neglected (i.e., it is effectively assumed that all sub-filter-scale fluxes are captured in the observations). However, since the flux imbalance problem is associated with large time and length scales, the effect of this assumption is anticipated to be small.

We define the fractional imbalance

$$F_\phi = \frac{I_\phi}{\Phi_\phi}, \quad (7)$$

which represents the imbalance as a fraction of the total flux. This choice is consistent with the interpretation that we assume all sub-filter scale contributions to fluxes are captured without errors. Note that for simplicity of notation, the above equations describe only the time dependence within a time interval of length  $T$ . That is, over a full year, the EC flux  $w'\phi'$ , and thus also the imbalance  $I_\phi$  and  $F_\phi$  retain a time-dependency over different intervals  $T$ .

The storage of the full turbulent three-dimensional fields at high time resolution was prohibitive, given the length of the simulation. Therefore, the time series were stored at four locations, spread evenly throughout the domain. Temporal fluctuations were calculated on the basis of the time series at these locations. Note that averaging over these locations is performed at the very last step of the calculation, and therefore the number of towers does not affect the results other than improving statistical convergence with respect to considering a single time series.

In order for our results to be comparable to observations, we precede the EC computations by the the linear detrending technique described in e.g., Steinfeld et al. (2007). Instead of

calculating the flux based on deviations from the mean, we consider deviations from a linear fit to the data: instead of Eq. 3 we employ,

$$\phi' = \phi - \tilde{\phi}, \quad (8)$$

where  $\tilde{\phi} = a + bt$  is the least-squares linear fit to  $\phi$  within an interval  $-T/2 < t < T/2$ . This technique is typically employed in observational analysis with the intent of removing ‘drift’ of sonic anemometers and to remove slow variations that do not contribute to a net flux (e.g., daily cycle) from the data. However, doing so will also inevitably remove large-scale motions that do contribute to the flux.

We can understand the effect of the filtering method by considering the data in the spectral domain. Spectrally, the EC method can be approximated as a filtering operation. If we denote the temporal Fourier transformation of  $\phi$  as  $\hat{\phi}(f)$ , dependent on frequency  $f$ , its cospectrum with vertical velocity  $w$  can be written as

$$E_{w\phi}(f) = \frac{1}{2} (\hat{w}^* \hat{\phi} + \hat{w} \hat{\phi}^*), \quad (9)$$

where the asterisk denotes the complex conjugate. Lee et al. (2005) show that the EC flux determination is the result of the application of a filter  $\hat{h}(f)$  to the cospectrum, given by,

$$\hat{h}(f) = \begin{cases} \frac{\sin^2(\pi f T)}{(\pi f T)^2} & \text{Mean-only removal} \\ 1 - \left[ \frac{\sin^2(\pi f T)}{(\pi f T)^2} - 3 \frac{\left( \frac{\sin(\pi f T)}{(\pi f T)} - \cos(\pi f T) \right)^2}{(\pi f T)^2} \right] & \text{Linear detrending} \end{cases} \quad (10)$$

Essentially,  $\hat{h}(f)$  is a high-pass filter that removes the low frequency contributions to  $\overline{w'\phi'}$ . The high-pass filter is narrower for linear detrending, since that procedure removes a greater portion of the larger scales. We study the influence of the filtering method on our results in Sect. 4.3.

In every EC measurement, sampling errors lead to a non-zero imbalance  $I_\phi$ . We might typically distinguish *random* sampling errors from *systematic* sampling errors, such that random sampling errors are scattered randomly around the actual value. They may be relevant to flux measurements on short time scales, but average out on the long term. On the other hand, the filter functions in Eq. 10 fall off around  $f = 1/T$ , which implies that any non-zero mean flux at larger time scales is underestimated. This creates systematic sampling errors, which represent a bias to the mean sampled value.

We can estimate the value of the systematic error by investigating the mean value of the imbalance  $I_\phi$  over the year-long dataset. The random errors are estimated from the standard deviation  $\sigma_{I_\phi}$ . We also construct the probability density function (p.d.f.) of  $I_\phi$ , which provides information on both the systematic and random errors.

In the absence of heterogeneity, any systematic error in the EC fluxes must be due to losses in the low-frequency flux contribution in statistically homogeneous turbulence due to the effect of the high-pass filtering of Eq. 10. Therefore, the magnitude of the fractional imbalance is related to the relative contribution of the low frequencies in  $E_{w\phi}(f)$ . Jonker et al. (1999) and Roode et al. (2004) have shown that the spectra of scalar variables may attain increasingly large low-frequency contributions, if given time to develop. This may hint at an increased relative contribution of the large scales to the vertical flux, if the scalar spectra at low frequencies are correlated with vertical motions.

Several corrections exist for systematic frequency losses associated with EC observations. The most popular of these use the standard spectra of [Kaimal et al. \(1972\)](#) to calculate an estimated loss (see e.g., [Moore 1986](#); [Bosveld et al. 1999](#)). In this study, we follow the approach of [Kanda et al. \(2004\)](#) and [Steinfeld et al. \(2007\)](#) and study sampling errors without such corrections. This allows us to identify influences that affect the sampling bias directly instead of identifying influences that affect the performance of the corrections.

### 3 Large-Eddy Simulation Set-Up

The analyses are based on the YOGA dataset described in [Schalkwijk et al. \(2015\)](#), which is publicly available (see [Schalkwijk et al. 2014](#)). In summary, two separate simulations were performed, YOGA-2012 and YOGA-HR-2012, each spanning a full year, driven by realistic forcings and tendencies taken from a regional weather model ([Meijgaard et al. 2008](#)).

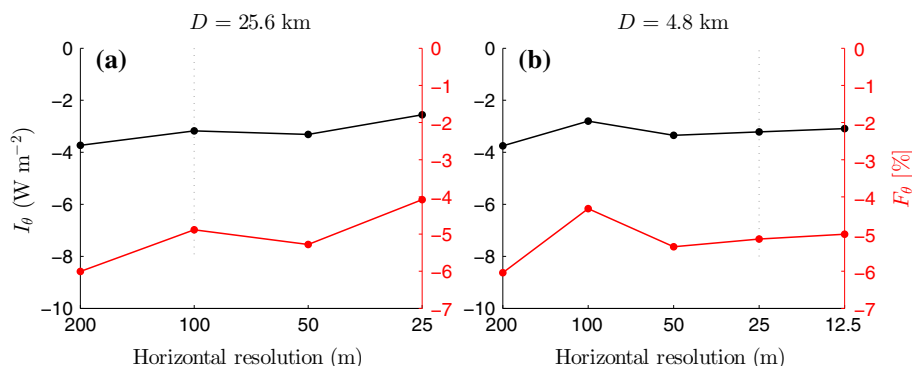
The two variants are identical, except for their spatial domain and resolution. YOGA-2012 uses a domain of  $25 \times 25 \text{ km}^2$  at 100-m grid spacing in the horizontal directions. Vertically, the grid spacing increases exponentially from 30 to 70 m at 25-km altitude. YOGA-HR spans a  $4.8 \times 4.8 \text{ km}^2$  domain, using 25-m horizontal grid spacing. The vertical grid spacing increases from 8 m at the surface to 40 m at the top of the domain, which lies at 3.6 km.

Both set-ups represent a compromise between the computational cost of a year-long integration and their ability to resolve physical phenomena. The relatively large domain of YOGA was chosen to be able to represent larger-scale motions and to prevent large-scale convective phenomena from experiencing the domain limits. On the other hand, the relatively high resolution of YOGA-HR was chosen to better resolve turbulent fluxes near the surface.

The YOGA runs are statistically horizontally homogeneous (i.e., all equations that are solved are horizontally homogeneous) and have periodic boundary conditions. The advection of turbulent fields through the domain is discretized using a second-order scheme, described in [Heus et al. \(2010\)](#). The large-scale forcings and tendencies are applied horizontally homogeneously. A land-surface model ([Schalkwijk et al. 2015](#)) calculates the evolution of the soil temperature and humidity for four layers beneath the surface using constant soil conductivities and solving the surface energy balance every timestep. The bottom boundary condition is provided by using the deep temperature and humidity provided by the regional weather model. The magnitude of the surface fluxes is determined by the difference between the horizontally-averaged soil skin layer and the horizontally-averaged turbulent fields at the first model level, in order to create horizontally homogeneous fluxes, consistent with the homogeneous set-up of the simulations. The distribution of the energy between sensible and latent heat fluxes is calculated on the basis of dry and wet ‘resistances’ for the surface layer, as described in [Viterbo and Beljaars \(1995\)](#). Effects of surface inhomogeneities or large-scale gradients are not represented in the LES run and are not considered.

#### 3.1 Resolution Dependence

To appreciate the numerical differences between YOGA and YOGA-HR datasets, we first perform a convergence test as resolution and domain are varied. Since the full YOGA datasets are far too computationally expensive to perform convergence tests, we choose to study the numerical convergence in an idealized, dry boundary layer. The simulation is driven by a surface heat flux of  $73 \text{ W m}^{-2}$ , roughly representative of the average magnitude found in the YOGA datasets, and with  $2 \text{ m s}^{-1}$  geostrophic wind speed, described in the Appendix. The geostrophic wind speed is chosen to be relatively low, which is anticipated to result in



**Fig. 1** Convergence of the EC flux imbalance of the sensible heat flux for increasing resolution at  $z = 100$  m. The case of 100-m grid spacing in panel **a** corresponds to the grid spacing set-up of YOGA-2012, the 25-m case in panel **b** to YOGA-HR, as indicated by vertical dotted lines. Vertical grid spacing changes with the same factor as horizontal grid spacing. The absolute imbalance is shown in black, fractional imbalance is shown in red. The sub-filter scale contribution to the total flux decreases from roughly 25 % to below 1 % as the resolution varies from 200 to 12.5 m

a relatively large imbalance. The fluxes are diagnosed for four intervals of  $T = 3600$  s at  $z = 60$  m, between 12 and 16 h into the integration, and then integrated. During this time, the boundary-layer height grows roughly from 900 to 1000 m.

Figure 1 shows the mean EC flux imbalance as diagnosed from the LES for varying resolution and domain. The horizontal grid spacing is varied between 200 and 25 m (12.5 m for a domain of 4.8 km). The vertical grid spacing is varied such that the grid cell aspect ratio is similar for all runs. This exercise is performed twice, once for a domain of  $25.6 \times 25.6 \text{ km}^2$  as in YOGA, and once for a 4.8-km domain as in YOGA-HR. Therefore, the 100-m case in panel (a) corresponds to the grid of YOGA and the 25-m case in panel (b) to YOGA-HR.

Figure 1 shows that the EC flux imbalance is quite robust for the given simulations, even though the sub-filter-scale contribution to the true flux ranges between 25 and 0.8 % for resolutions between 200 and 12.5 m. This provides further confidence that, as expected from Eq. 10, the sampling errors occur in the low-frequency component of the flow, which is the component best resolved by the LES.

### 3.2 Analysis Method

Since Fig. 1 suggests that the domain size has less influence on the EC flux imbalance than does the grid spacing, in the following we determine the imbalance in the YOGA-HR dataset. Furthermore, the improved vertical resolution of YOGA-HR allows us to study the imbalance closer to the surface.

The nighttime boundary layer often remains under-resolved in YOGA-HR (Schalkwijk et al. 2015). Since the evaluation of EC fluxes in LES requires sufficiently resolved vertical transport, we here focus on the daytime EC errors. Note that Hendricks Franssen et al. (2010) have shown that, while the relative imbalance peaks at night, the largest absolute contribution to the flux imbalance problem occurs during the day (in unstable conditions), due to the larger magnitude of daytime surface fluxes. Therefore we limit ourselves to the instances where the area-averaged resolved flux is larger than  $10 \text{ W m}^{-2}$ . In practice, this implies the elimination of all nighttime fluxes, as well as that of instances of stable conditions during daytime (which occasionally occur during winter). Due to averaging effects, the results presented below are



biased toward the summer season, since that season is generally characterized by the largest fluxes.

We focus on averaging times and heights that are roughly representative of a typical small-tower observational set-up. Specifically, we concentrate on averaging periods of 15, 30 and 60 min. Furthermore, we focus on sampling errors at heights of 16, 58 and 101 m, which coincide with vertical grid points of the LES to avoid the need to interpolate, which potentially affects accuracy. We urge caution in interpreting the results at 16 m, since the sub-filter-scale contribution is especially large at this level (see Sect. 4). However, since the sampling bias is expected to be caused by large time and length scales (which are best resolved), these results should still be sufficient to provide a rough estimate of the actual sampling bias.

The fluxes that are considered are the latent heat flux  $L_v E = \rho L_v \overline{w'q'_l}$  and the sensible heat flux  $H = \rho c_p \overline{w'\theta'_l}$ , where  $\rho$  is the density of dry air,  $c_p$  is the specific heat capacity and  $L_v$  is the latent heat of vaporization. Thus, we use the prognostic variables in GALES for the imbalance calculations, that is  $\phi = \{\theta_l, q_l\}$ , the liquid water potential temperature and total specific humidity, respectively. The main difference between  $\theta_l$  and  $q_l$ , in the context of imbalance calculations, is that the value of  $\theta_l$  directly influences buoyancy and therefore  $\theta_l$  is an ‘active’ field (i.e., it influences the dynamics). Humidity  $q_l$ , on the other hand, has a much weaker impact on buoyancy and can therefore roughly be regarded as representative of a passive scalar. Therefore, whereas the imbalance of  $\theta_l$  may be characteristic for this variable, the imbalance of  $q_l$  is expected to be similar to the transport of other passive scalars.

## 4 Year-Averaged Imbalance Results

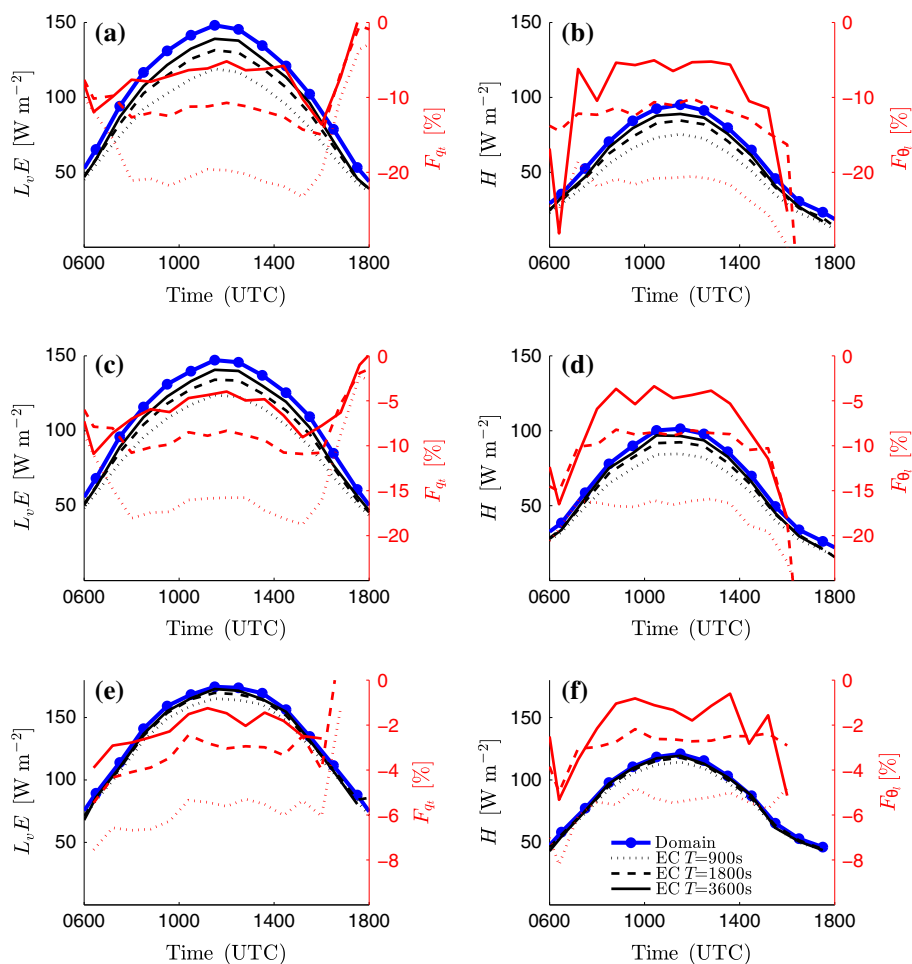
### 4.1 Mean Imbalance

Figure 2 shows the year-averaged diurnal cycle of sensible and latent heat fluxes at heights of 16, 58 and 101 m. The EC fluxes, evaluated using averaging periods of 900, 1800 and 3600 s, are shown in black with solid, dashed and dotted lines, respectively. They should be compared with the “true” flux depicted in blue. The EC fluxes are constructed by removing the average imbalance  $\overline{I_\phi}$  from the true flux (to account for the sub-filter-scale contribution). The fractional imbalance  $F$  is also shown and is depicted in red. Note that the fractional imbalance is calculated as  $F_\phi = \overline{I_\phi} / \overline{\Phi_\phi}$ , i.e., the ratio between the averages instead of the average ratio. The left panels show the latent heat flux  $L_v E$ , the right panels show the sensible heat flux  $H$ . Note that, although theoretically these fluxes are representative for the transport of passive and active scalars, respectively, Fig. 2 shows that the relative imbalance for both variables is very similar.

Since the fluxes are averaged over a full year of LES data, the mean imbalance in Fig. 2 represents the systematic error in the EC fluxes. Figure 2 thus confirms the presence of a significant systematic error, also after averaging over an entire year comprising varying weather conditions.

The absolute and relative values for the imbalance are summarized in Table 1. The imbalance depends on the averaging period  $T$  and the averaging height  $z$ ; for an averaging period of 900 s, the year-averaged daytime imbalance ranges from 9 W m<sup>-2</sup> (6%) at 16 m to over 20 W m<sup>-2</sup> (20%) at 100 m for latent heat fluxes. The results improve for an averaging period of 3600 s, where the imbalance decreases to approximately 6% (6 W m<sup>-2</sup>) at  $z = 100$  m. Close to the surface, for  $z = 16$  m, the imbalance for  $T = 3600$  s is reduced to only 1–2% (1–2 W m<sup>-2</sup>) for both latent and sensible heat fluxes. However, the average sub-filter-scale





**Fig. 2** Year-averaged diurnal cycle of spatial and EC turbulent fluxes of latent ( $E$ , left) and sensible ( $H$ , right) heat in YOGA-HR-2012, shown at three 'tower heights'  $z = 101$  m (panels **a**, **b**),  $58$  m (**c**, **d**) and  $16$  m (**e**, **f**). 'True' (domain-averaged) fluxes are depicted with a thick blue line, EC fluxes are shown for averaging periods  $T = 900$  s,  $1800$  s and  $3600$  s using dotted, dash-dotted and dashed lines, respectively. Relative imbalance  $F$  is shown in red

contribution to the true flux is very large at this height (roughly 60 %). The resolution is not sufficient to properly resolve the turbulent processes here, especially so since during the year the boundary layer is often driven by processes other than surface convection (i.e., large-scale conditions driving the boundary layer through shear) that are harder to resolve in a LES.

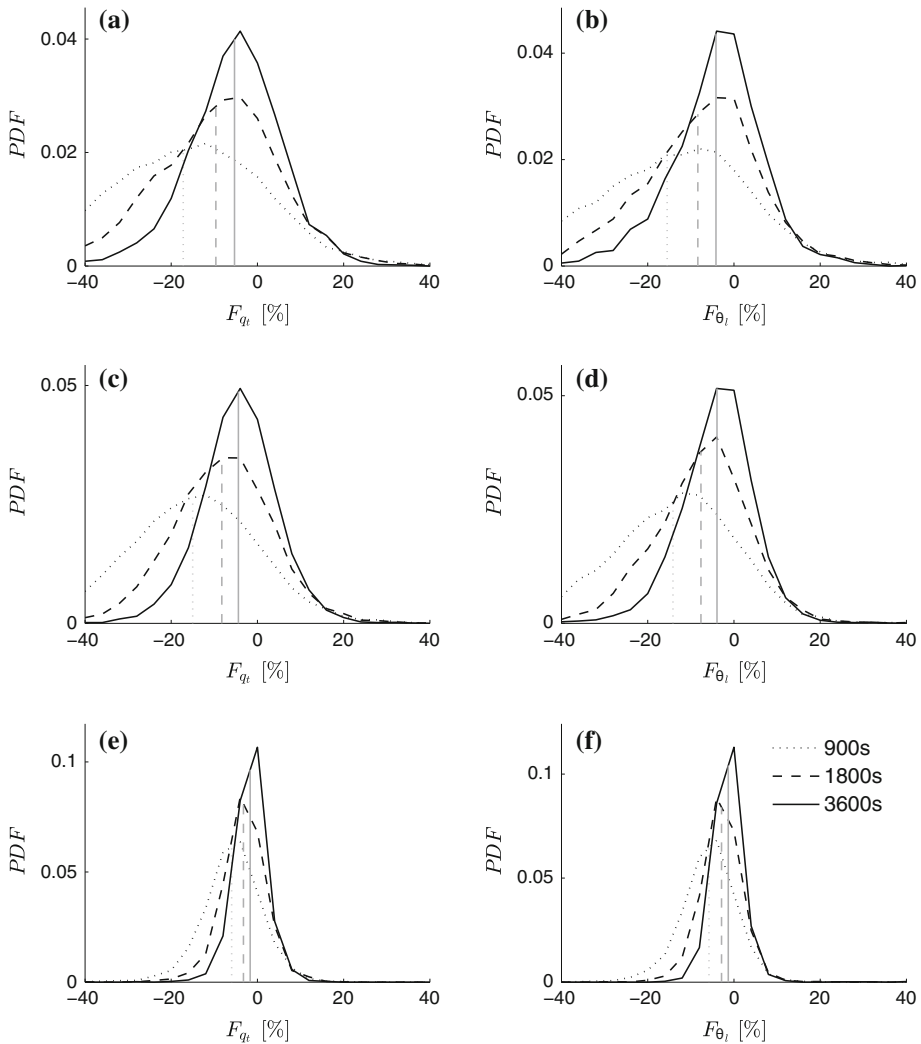
## 4.2 Imbalance Spread

The mean imbalance discussed above is a measure of the systematic error, or bias, of the EC flux. In cases where one is interested in fluxes on short time scales, the random error may be equally important. In order to investigate the random error of the EC flux, the use p.d.f. of the imbalance is informative. Figure 3 shows the use p.d.f. of the fractional imbalance  $F$  for identical averaging periods and heights as in Fig. 2; latent and sensible heat fluxes are shown in the left and right panels, respectively, as before.

Table 1 True fluxes and EC flux imbalance

Latent heat									
z (m)	True flux (W m <sup>-2</sup> )	SFS (%)	$\overline{I_{q_t}}$ (W m <sup>-2</sup> ), for T =			$\overline{F_{q_t}}$ (%), for T =			
			3600 s (%)	1800 s (%)	900 s (%)	3600 s (%)	1800 s (%)	900 s (%)	
16	146.8	62.1	-2.36	-4.56	-8.748	-1.6	-3.1	-5.8	
58	110.4	15.1	-5.17	-9.91	-18.07	-4.7	-9.0	-16.4	
101	104.0	7.1	-6.04	-11.56	-20.86	-5.9	-11.2	-20.1	
Sensible heat									
z (m)	True flux (W m <sup>-2</sup> )	SFS (%)	$\overline{I_{\theta_1}}$ (W m <sup>-2</sup> ), for T =			$\overline{F_{\theta_1}}$ (%), for T =			
			3600 s (%)	1800 s (%)	900 s (%)	3600 s (%)	1800 s (%)	900 s (%)	
16	98.2	59.9	-1.18	-2.59	-5.20	-1.2	-2.6	-5.3	
58	77.1	10.9	-3.26	-6.66	-12.78	-4.3	-8.7	-16.6	
101	72.4	3.9	-3.90	-8.05	-15.36	-5.4	-11.1	-21.2	

Averaging is performed over all instances where the absolute area-averaged resolved flux is larger than 10 W m<sup>-2</sup>



**Fig. 3** Year-averaged probability density function (*PDF*) of the daytime EC fractional flux imbalance, for different averaging periods and averaging heights;  $z = 101$  m for panels **a** and **b**,  $z = 58$  m for panels **c** and **d** and  $z = 16$  m for panels **e** and **f**. *Left panels* show the latent heat fluxes, *right panels* the sensible heat fluxes

Figure 3 shows that in general, the mean of the imbalance correlates well with its spread: instances with a relatively large mean imbalance (i.e., at higher altitudes or for shorter averaging periods) also feature a relatively wide p.d.f. Hence, at these instances not only is the mean (systematic) error large, but so is the chance of encountering a very large imbalance in a given measurement interval. This is consistent with expectations based on the theoretical framework of Lenschow et al. (1994), who show that both random and systematic errors scale with the turbulent time scale.

Note that the combined result of an increased standard deviation and an increased magnitude of mean imbalance for shorter averaging periods is a pronounced increase in probabilities of underestimating the true flux, while the probabilities of overestimation remain roughly

**Table 2** Imbalance statistics for  $T = 1800$  s

$z$ (m)	Latent heat flux			Sensible heat flux		
	$\sigma_F$ (%)	Probability (%)		$\sigma_F$ (%)	Probability (%)	
		$F < -10\%$	$F < -20\%$		$F < -10\%$	$F < -20\%$
16	6.4	7.9	0.5	6.5	6.4	0.1
58	14.2	42.8	11.7	20.4	38.1	10.4
101	17.1	46.0	19.6	19.3	41.2	16.6

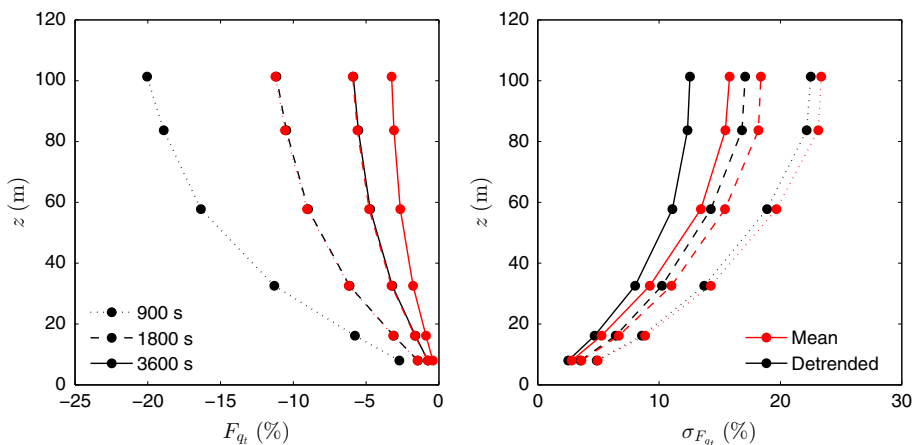
Averaging is performed over all instances where the absolute area-averaged resolved flux is larger than  $10 \text{ W m}^{-2}$

similar. This can be ascribed to the fact that the two effects add for negative imbalances, but partially cancel for positive imbalances. As a result, the chance to underestimate the true flux significantly increases for shorter averaging times, as shown in Table 2, which tabulates the chance of underestimation by more than 10 and 20 %, along with the standard deviation of the flux imbalance, for  $T = 1800$  s. It emphasizes the need for longer duration measurements especially for tall-tower experiments, as the chance to underestimate the true flux by more 20 % can be as large as one in five for latent heat fluxes at  $z = 100$  m.

The standard deviation for both  $F_{qt}$  and  $F_{\theta_l}$  varies from approximately 5 % at 16 m with 1-h averaging to 23 % at 100 m with 900-s averaging. Note also that the shapes of the imbalance p.d.f. are relatively symmetrical: the bias seems mainly related to the location of the centre of the p.d.f. not its skewness.

### 4.3 Filter Method

As explained in Sect. 2, the EC flux in Figs. 2 and 3 were constructed using a linear detrending filter. To investigate the effects of linear detrending, Fig. 4 compares the fractional imbalance for specific humidity flux  $F_{qt}$  with and without linear detrending of the data, i.e., using Eqs. 8 and 3, respectively.



**Fig. 4** The year-averaged dependence of daytime fractional imbalance  $F$  of the latent heat flux to averaging height  $z$ , for different filter methods and averaging period  $T$ . Red lines indicate no filtering (subtract the mean only), and black lines result after linear detrending

Figure 4 shows that the EC flux imbalance is significantly larger for the linearly detrended signal, consistent with [Steinfeld et al. \(2007\)](#), and with expectations based on Eq. 10 for a positive  $E_{w\phi}$ . In fact, the average fractional imbalance for detrended data for  $T = \{3600 \text{ s}, 1800 \text{ s}\}$  is strikingly close to the imbalance for unfiltered (mean-removal only) data for  $T = \{1800 \text{ s}, 900 \text{ s}\}$ , respectively. This implies that the linear detrending method effectively halves the averaging period  $T$ . Note that removing the mean for  $0 < t < T/2$  and the mean for  $T/2 < t < T$  effectively removes, on average, the first-order effect of the  $0 < t < T$  trend, which may explain the equivalence of detrending over a period  $T$  and averaging over a period  $T/2$ . This effect becomes noticeable only with sufficient statistics: it becomes apparent when analyzing time series longer than 100 h.

Whereas the systematic error increases due to linear detrending, the standard deviation (related to the random error) decreases. The fact that the standard deviation is significantly larger for unfiltered data than for detrended data implies that the largest scales are responsible for a significant contribution to the random error, by removing these scales the random error is reduced at the cost of an increased mean error. The large contribution of the large scales to the random error can be explained by the fact that these scales inherently are also those most poorly sampled, due to their large decorrelation lengths. Hence, in situations where the absolute uncertainty must be diminished and only a relatively short time-frame is available for measurement, linear detrending might be the best option. The same holds in cases where one can correct for the bias.

It is also important to realize that the above analyses cover only measurements without error. As a result, the procedure of linear detrending removes a physical trend in the observed quantity, and thus removes a portion of the actual flux. In the event that the measured series is polluted with an additional non-physical trend (e.g., an artifact of the measurement set-up, for instance, drift of a sonic anemometer), linear detrending may be unavoidable ([Kroon et al. 2010](#)).

In the following we continue to employ the linear detrending method, with the notion that, on average, the approximate relation between the linear detrended EC flux and the unfiltered EC flux is known.

## 5 Dependence on Environmental Conditions

In this section we investigate the environmental factors that influence the magnitude of systematic flux-sampling errors. Given that the systematic errors originate in the large time scales (Eq. 10), we focus on variables that are likely to affect these scales.

The time scales observed from a steady tower can be subdivided into two main components. The first is related to the speed at which turbulent eddies are advected over the tower, which is governed by the mean wind speed. The scales of the turbulent eddies themselves represent the second component.

### 5.1 Wind-Speed Dependence

[Kanda et al. \(2004\)](#) investigated the effect of the mean wind speed on the EC flux imbalance. They argued that as the mean wind speed increases, more eddies pass the measurement station, improving the statistical stability and reducing the flux imbalance. Indeed, [Kanda et al. \(2004\)](#) find that the relative imbalance decreases from 20 % to less than 4 % as the geostrophic wind speed increases in magnitude from 0 to  $4 \text{ m s}^{-1}$  (at 100-m height,  $T = 3600 \text{ s}$ ). On the other hand, [Steinfeld et al. \(2007\)](#) described how high wind speeds can fundamentally alter

the convection pattern, and with that the turbulent time scales, significantly. They find that an increased wind speed induces a change to roll-like convection plumes that align in rows parallel to the mean flow, significantly increasing the EC flux imbalance. This increase occurs since measurement towers in such situations are often located in either a warm updraught or a cold downdraught during the full averaging period.

In order to quantify the wind-speed dependence of the systematic error in the EC flux in YOGA-HR, we have correlated imbalance with horizontal wind speed  $U = \sqrt{u^2 + v^2}$ , taken at the, same height as the EC flux measurement. The instances of imbalance at similar wind speeds (bins of  $1 \text{ m s}^{-1}$  are taken) are then averaged. Figure 5 shows the results for the fractional imbalance in the latent heat flux  $L_v E$ ; results for the sensible heat flux (not shown) are similar. On average, the imbalance decreases with wind speed for all averaging periods and height; the imbalance strongly decreases as the wind speed increases up to  $3 \text{ m s}^{-1}$ . Then, between 3 and  $8 \text{ m s}^{-1}$  (3 and  $6 \text{ m s}^{-1}$  at 16-m height) the imbalance seems to plateau or even increase with wind speed, only to decrease again for higher wind speed.

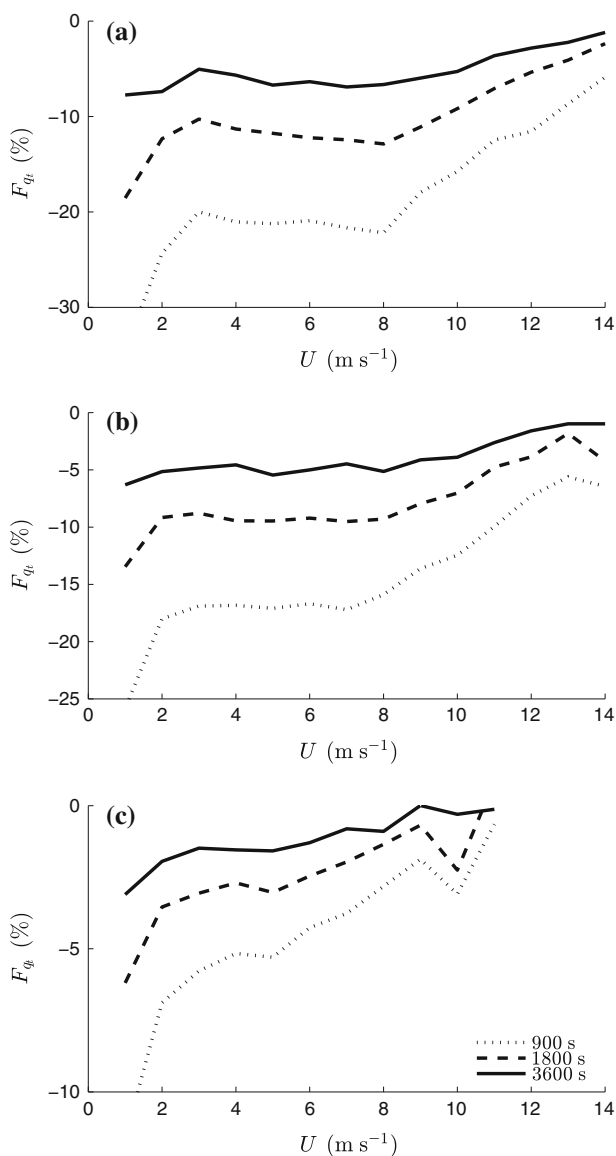
Therefore, it is likely that the effects described by Kanda et al. (2004) and by Steinfeld et al. (2007) both occur. The initial decrease of imbalance with wind speed may be the same effect as that reported by Kanda et al. (2004): so long as the flow regime remains similar, the imbalance decreases with wind speed. However, the magnitude of the imbalance does not decrease further for intermediate wind speeds, which might indicate a change in flow regime. On average, we find a net decrease of imbalance with wind speed, suggesting that the effect of improving statistical stability with increasing wind speed dominates in this dataset.

Note that the effect that wind speed has on the imbalance may be influenced by the (surface) homogeneity of the current YOGA set-up. Inagaki et al. (2006) have shown that surface heterogeneity stimulates the formation of large-scale structures when large horizontal differences in the surface heat flux occur. Moreover, surface heterogeneity may bind turbulent organized structures to the surface structure, rendering their location steady in time, potentially further increasing the magnitude of the imbalance. However, these structures will be destroyed for higher wind speeds, potentially resulting in a larger difference between low-wind and high-wind conditions.

## 5.2 Extreme Imbalance

To further our understanding of the conditions that promote imbalance, we attempt to characterize the days that result in an especially large EC flux imbalance. To this end, we calculate the average imbalance per day, again considering only those instances having a resolved flux that exceeds  $10 \text{ W m}^{-2}$ , and then select the days that rank in the top 10 % in terms of relative imbalance  $F$  (in latent heat fluxes) for  $z = 100 \text{ m}$  and  $T = 1800 \text{ s}$ . The resulting set spans 36 days in which, on average, the fractional daytime imbalance is  $F_{qt} \approx -23 \%$ , against  $-10 \%$  over the full year. We will refer to this sub-set of days as the top-imbalance days.

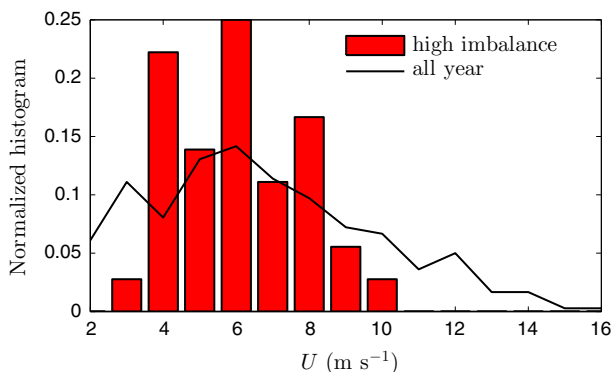
Figure 6 shows the probability density distribution of wind speed  $U$  (at  $z = 100 \text{ m}$ ) over all days in YOGA-HR, together with a histogram of the top-imbalance days. The top-imbalance days are neither especially windy nor are they windless, as may have been expected based on Fig. 5. Instead, they are characterized by medium wind speeds of roughly  $4\text{--}8 \text{ m s}^{-1}$ . Note that these wind speeds roughly span the range in which the mean imbalance plateaus in Fig. 5, implying that in this range of wind speeds, the effect of the top-imbalance days is compensated by the effect of days with lesser imbalance, for otherwise the imbalance would have reached a minimum here. This reinforces the suggestion that wind speed is not the sole controlling factor.



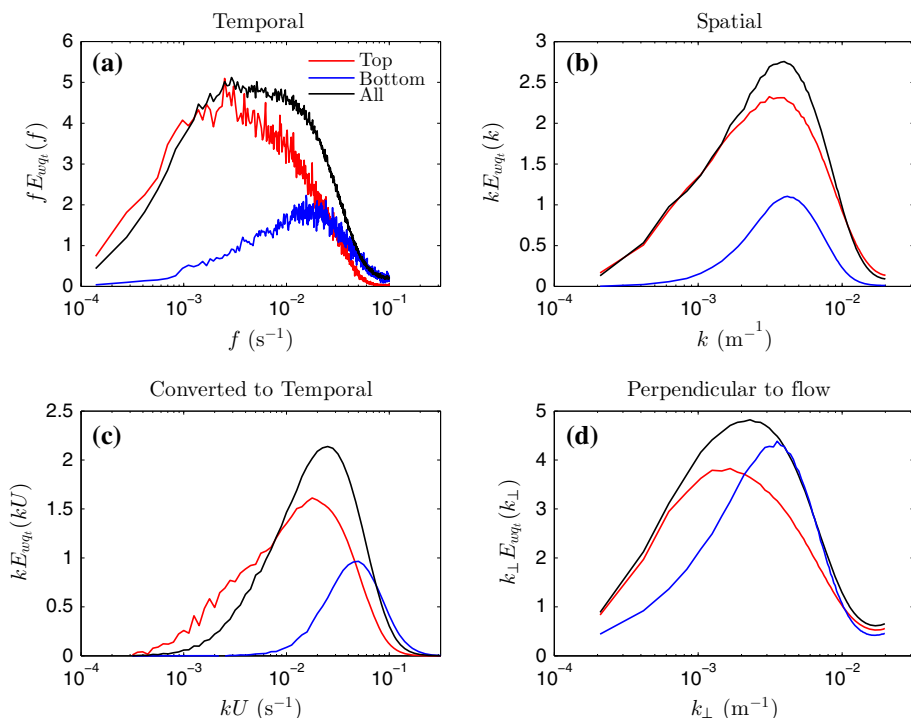
**Fig. 5** The year-averaged dependence of imbalance of the latent heat flux on the horizontal wind magnitude  $U = \sqrt{u^2 + v^2}$ .  $U$  is taken at the height of the measurement;  $z = 101$  m for (a),  $z = 58$  m for (b) and  $z = 16$  m for (c). Note that there are no instances with  $U > 12 \text{ m s}^{-1}$  for  $z = 16$  m. The imbalance is shown for averaging periods  $T = 900$  s,  $1800$  s and  $3600$  s

The cospectrum  $E_{wq_t}(f)$  is shown in Fig. 7a, where the year-averaged cospectrum in black is compared with the top-imbalance selection in red. Additionally, a bottom-imbalance selection is made analogous to the top-selection to identify the bottom 10 % imbalance days. The average cospectrum of this selection is shown in blue. Note that the cospectrum so plotted retains its surface area (i.e., the turbulent flux) intact under the logarithmic axis:  $\int E df = \int f E d \log f$ .





**Fig. 6** Normalized histogram of horizontal wind speed in bins of  $2 \text{ m s}^{-1}$  width, for the full year of data in black and for the top-imbalance days in red, at a height of 100 m



**Fig. 7** Spectral analysis of the latent heat fluxes over the full year of YOGA-HR in black, the top-imbalance days in red and the bottom-imbalance days in blue. Shown are **a** the temporal cospectra, **b** the two-dimensional spatial cospectra, **c** the spatial cospectra, converted to temporal cospectra using the mean horizontal wind  $U$  and **d** the one-dimensional spatial cospectrum in the direction perpendicular to  $U$ . All data are taken at a height of 100 m

As expected, the top-imbalance days show a cospectrum having a larger relative contribution of low frequencies (large time scales) than the average cospectrum, while the bottom-imbalance days are centered around the high frequencies. This confirms that the large imbalance in the top-imbalance days is related to a larger contribution of large time

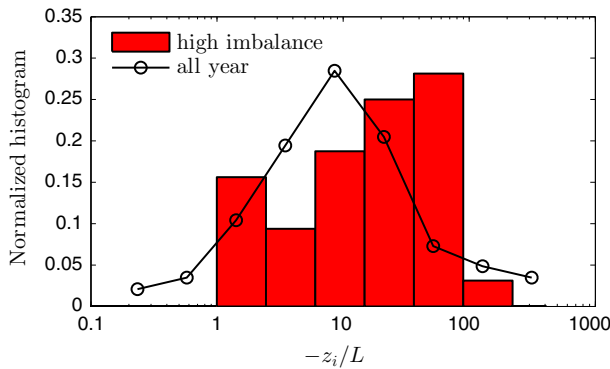
scales in the turbulent fields. Although we might intuitively extend this notion to the idea that top-imbalance days are characterized by large turbulent structures, and thus also large *spatial* scales, this seems to be contradicted by panel b of the same figure, which shows the spatial cospectrum  $E_{wq_t}(k)$ , where  $k = \sqrt{k_x^2 + k_y^2}$  is the wavenumber. This shows that the spatial spectrum in top-imbalance days does not significantly differ from the year-averaged spectrum, that is, the contribution of the larger spatial scales in the spectrum is not significantly larger.

Hence, the large imbalance is caused by the presence of large time scales, but no spatial scales stand out in a horizontally isotropic analysis. In panel c, the spatial spectrum is converted into a temporal spectrum by estimating a temporal scale using the average wind speed:  $\tilde{f} = kU$ . This construction of temporal scales based on the spatial spectrum shows that the large temporal scales are, in fact, resolved turbulent motions and not due to a transition in time that is enforced through external forces on the LES (by the regional weather model, e.g., the diurnal cycle). Thus, the larger temporal scales are a result of the manner in which the spatial scales are advected past a tower.

Figure 7d shows the cospectrum  $E_{wq_t}(k_\perp)$ , where the spectrum is calculated in lines perpendicular to the mean wind direction (i.e.,  $k_\perp$  is the wavenumber in the direction perpendicular to  $U$ ), and averaged along the wind direction. As a result, this spectrum shows the *spatial* scales that are perpendicular to the wind direction. Panel d shows that the bulk of these scales is significantly larger in the top-imbalance selection, and smaller in the bottom-imbalance selection. Comparison of panels b and d shows that, whereas the length scales of the top-imbalance days do not seem larger than average when studied in a horizontally isotropic manner, they are significantly larger than average when the data are first aligned with wind direction. In particular, the difference of the scale at which the energy peaks between panels b and d is much larger for the top-imbalance days than for the average. This shows that the *spatial* structures of the boundary layer in top-imbalance days are different from those on an average day. The alignment of turbulent structures with the mean flow is often referred to as the formation of ‘roll vortices’. Roll vortices were identified as a cause for an increased sampling bias in [Steinfeld et al. \(2007\)](#), who showed that in these circumstances, an observer is more likely to be in an updraught or downdraught for a complete measurement period  $T$ , resulting in a significant underestimation of turbulent transport.

Although it might seem more natural to study the scale of roll vortices in the along-wind direction, our LES domain is not sufficiently large to allow such a study, limiting the analysis to the cross-wind direction. Since roll vortices are essentially the alignment of the boundary-layer structure with the mean flow, one would expect to see larger scales also in the cross-wind cospectrum. To explain this notion, assume that the cross-wind scale is roughly characterized by the boundary-layer depth  $z_i$ , such that a scale analysis along the cross-wind direction would identify precisely this scale due to the consistent ‘angle of attack’ of the analysis. In a horizontally isotropic boundary layer with length scale  $z_i$ , however, any analysis will intersect the turbulent structures in a random matter, finding scales between zero and  $z_i$  such that the average scale is significantly smaller than  $z_i$ . Therefore, the difference between panels b and d, emphasizing the difference between an unaligned and a horizontally isotropic analysis, is evidence for the alignment of the turbulent structures with the mean flow.

Several previous studies (e.g., [Moeng and Sullivan 1994](#); [Khanna and Brasseur 1998](#)) have shown that the flow regime of the boundary layer in general, and the formation of roll vortices in particular, may be characterized by the stability parameter  $-z_i/L$ , where  $z_i$  is the boundary-layer depth and  $L$  is the Obukhov length,



**Fig. 8** Normalized histogram of  $-z_i/L$  in bins of exponentially increasing size, for the full year of data in black and for the top-imbalance days in red

$$L = -\frac{u_\star^3 \langle \theta_v \rangle}{\kappa g \langle w' \theta_v' \rangle_s}. \quad (11)$$

For  $-z_i/L \rightarrow \infty$ , the boundary layer is completely driven by buoyancy, and large scales form without a preferential direction. For  $-z_i/L \rightarrow 0$ , the boundary layer is shear-driven, and the turbulent scales are generally relatively small. For some intermediate values of  $-z_i/L$ , the larger scales that form due to buoyancy may align with the wind.

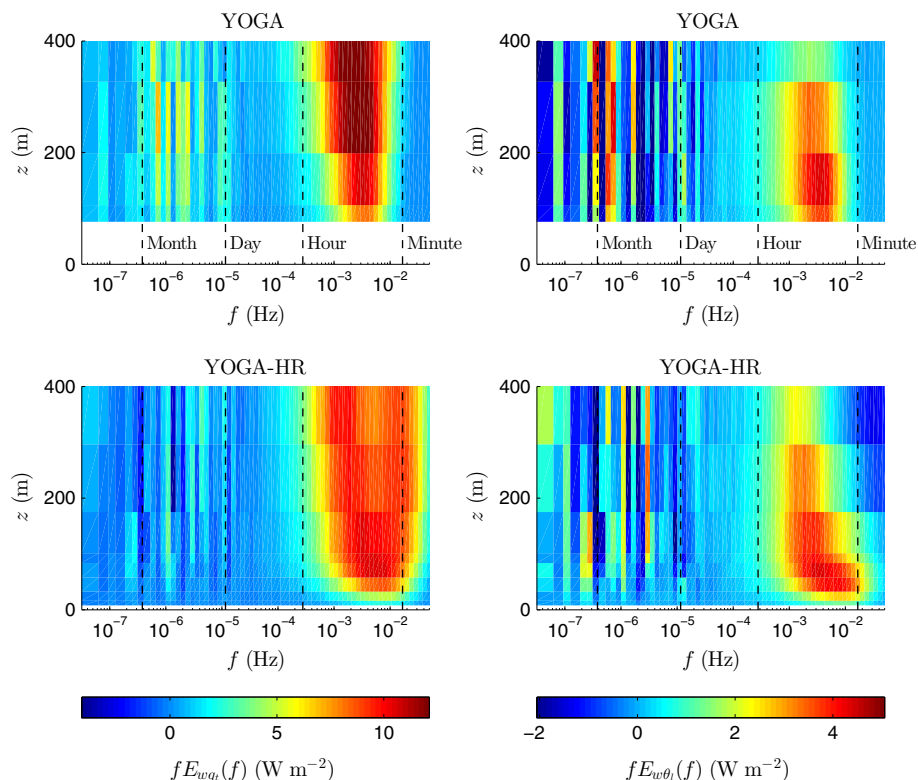
The histogram for  $-z_i/L$  in Fig. 8 shows significantly increased probabilities of high values of  $-z_i/L$  for the top-imbalance days when compared to year-averaged values. Top-imbalance days feature significantly larger values for  $-z_i/L$  than average, which confirms that  $-z_i/L$  is indicative of large imbalance. Days with a large EC flux imbalance are characterized roughly by  $10 \leq -z_i/L \leq 100$ . Although Moeng and Sullivan (1994) and Khanna and Brasseur (1998) estimate roll vortex formation is characterized by smaller values for  $-z_i/L$  (roughly  $1.5 < -z_i/L < 10$ ), these studies were limited to the clear boundary layer. Given that nearly every day is cloudy in the YOGA dataset, and that  $z_i$  is defined as the height of the steepest gradient in the buoyancy profile (Schalkwijk et al. 2015), which typically occurs above the cloud layer, some discrepancy may be attributed to cloudiness.

Note also that  $-z_i/L = (w_\star/u_\star)^3$ , where  $u_\star/w_\star$  is the dimensionless ratio with which Huang et al. (2008) propose to scale the imbalance, and who found that the imbalance increases as  $u_\star/w_\star$  decreases, which is in accordance with our findings for large  $-z_i/L$ .

Furthermore, the standard cospectra proposed by Kaimal et al. (1972) are not dependent on  $z_i$ , implying that corrections that employ these spectra (e.g., Moore 1986; Bosveld et al. 1999) are not either. Also, these corrections cannot account for preferential directions of large scales (as occurs in roll vortices), and so the effect of roll vortex formation on EC sampling errors is not easily corrected.

## 6 Time Scales of Turbulent Transport

We have emphasized the role of the time scales of turbulent transport on the quality of EC flux determination. Therefore, we now investigate the time scales of turbulent transport throughout the year by studying the Fourier transform of the full year-long time series.



**Fig. 9** Year-averaged temporal power cospectra of  $w$  and  $q_t$  (left panels) and  $w$  and  $\theta_l$  (right panels). The cospectra were averaged over exponentially increasing bin sizes to reduce noise. Colours denote the cospectral energy at a given height and frequency

Figure 9 shows the cospectra of  $w$  and  $\theta_l$  and  $w$  and  $q_t$  as a function of height. Note that the power spectra of wind velocity in YOGA-2012 show artifacts of the horizontal domain size (Schalkwijk et al. 2015), since the limited domain size limits the large-scale variance that can be simulated. Unlike the power spectra, however, the cospectra of YOGA and YOGA-HR data agree very well on the dominant scales, suggesting that the domain size does not affect cospectra as much as power spectra.

Figure 9 shows that the bulk of the net turbulent transport throughout the year is performed at time scales  $< 1$  h, although at  $z = 100$  m, a significant amount of transport is still performed at 2-h time scales. The peaks at low frequencies (day to month scale) are due to noise, caused by the inherent relatively poor sampling of the largest scales in performing a Fourier analysis.

Given the relative magnitude of daytime fluxes when compared to nighttime fluxes, the cospectra in Fig. 9 are representative mostly of the situation at daytime. Nighttime cospectra are typically limited to much smaller time scales (below 10 min, see e.g., Vickers and Mahrt 2003). As expected, the range of time scales involved in turbulent transport is limited to the smallest scales close to the surface. The time scales quickly grow as the height increases to 100 m, above which the growth is much slower. Note also the double-peak structure of the cospectrum of  $wq_t$  at  $z > 200$  m, which is due to the summing of two flow regimes. In the convective boundary layer (occurring most frequently during summer), scales quickly grow with height and the large scales dominate the transport. In the strongly stratified, shear-

driven boundary layer (occurring most frequently in winter), the turbulent transport of  $q_t$  is dominated by the smallest scales over the entire boundary layer. Hence, although the peaks of convective and stratified transport coincide at low heights, they separate at higher altitudes as the convective scales grow. These effects are only visible in YOGA-HR, since the resolution of YOGA is insufficient to resolve this difference. The heat flux shows a strongly positive convective ‘branch’, i.e., the growth of scales with height, whereas the stratified branch shows as a negative transport since heat is typically transported downwards on stably stratified days.

## 7 Discussion and Conclusions

We have extended the work of [Kanda et al. \(2004\)](#) and [Steinfeld et al. \(2007\)](#), who used LES to study the eddy-covariance (EC) flux imbalance in terms of flow properties. Where these authors studied imbalance in idealized archetypical situations, our year-long LES integrations allowed us to study the EC flux imbalance problem in a more realistic setting and to identify environmental factors that influence sampling errors in the EC flux imbalance.

We confirm the presence of a significant systematic bias in EC flux measurements for all measurement periods and measurement heights, also after averaging over a full year. Nevertheless, the problem is much smaller at lower heights and for longer averaging periods. We find that an EC averaging period of 900 s is rather short, resulting in a systematic error in the daytime flux determination of more than 5 % at 16-m height and over 15 % at 100-m height. For  $T = 3600$  s, the bias is reduced to 1–2 % at 16 m and 4–6 % at 100 m. Note that the dependency on averaging period and height was also theoretically established in [Lenschow et al. \(1994\)](#), and it would be an interesting topic for further research to investigate to what accuracy these theoretical predictions hold quantitatively. To that end, note that the full datasets of YOGA and YOGA-HR are publicly available ([Schalkwijk et al. 2014, 2015](#)).

In accordance with [Steinfeld et al. \(2007\)](#), we thus find that the low-frequency contribution to the turbulent flux is insufficient to account for the 25 % surface energy imbalance reported ([Twine et al. 2000](#)). Nevertheless, the low-frequency contribution remains an important effect to consider for tall-tower applications, whose ‘footprint’ allows such towers to be representative of a larger surface area. In this study, we have not considered the effect of spectral corrections (e.g., [Moore 1986](#); [Bosveld et al. 1999](#)).

The imbalance investigated is only a limited component of the actual imbalance. First, the emulated EC measurements in the LES are without measurement errors; second, the sampling errors that are captured are limited to the resolved range of motions of the LES. Third, the LES integrations are performed over a horizontally homogeneous surface and on a periodic domain, such that they are forced to be statistically horizontally homogeneous.

In general, we find that linear detrending of the signal before applying EC analysis significantly increases the systematic error, but reduces the random error. In particular, the systematic error in an EC measurement after linear detrending with an averaging period  $T$  is found to be, on average, equal to that of an EC measurement that does not use filtering and has an averaging period  $T/2$ . The application of linear detrending thus effectively halves the averaging period with respect to the systematic error. The random error is reduced by linear detrending, but that effect is smaller than the effect on the systematic error.

The EC flux imbalance is strongly influenced by wind speed, but in a non-uniform manner. On average, an EC flux imbalance decreases with increasing wind speed, but we find that days of especially large imbalance typically feature wind speeds  $\approx 6 \text{ m s}^{-1}$ . On these days, persistent turbulent structures arise that align with the mean flow (i.e., roll vortices),

causing the EC sampling bias to double in magnitude as compared to the yearly average. The variable  $-z_i/L$  is indicative of the occurrence of these phenomena, with instances of largest imbalance occurring for very large  $-z_i/L$ , i.e.,  $10 < -z_i/L < 100$ . The large differences in the magnitude of the EC flux imbalance from day to day emphasizes the effect of case set-up, since idealized case studies often consider only a few sets of conditions. Therefore, case-to-case differences may be sufficient to explain the differences in the effect of wind speed reported by Kanda et al. (2004) and Steinfeld et al. (2007). Furthermore, corrections for the low-frequency loss based on the standard spectra of Kaimal et al. (1972) are independent of  $z_i$  and can therefore not fully account for the extra imbalance due to roll vortices. As such, roll vortex formation may present an important complexity when observing EC fluxes, especially during relatively short measurement campaigns.

**Acknowledgments** This work would not have been possible without the KNMI parametrization testbed (KPT), and are thankful for the continuous efforts of the KNMI team responsible for maintaining this project. Of particular value has been the work by Roel Neggers who initiated the testbed, by Erik van Meijgaard for generating and providing the advective forcings from the regional atmospheric climate model (RACMO), and by Bert van Uft for continuously maintaining the testbed. We acknowledge PRACE for awarding us access to the CURIE hybrid nodes based in France at TGCC, owned by GENCI and operated by CEA. Last, but not least, the authors gratefully thank three anonymous reviewers for their suggestions to improve the manuscript.

**Editor's note** I leave the reader to interpret the undefined acronyms RACMO, PRACE, CURIE, TGCC, GENCI, and CEA. This is an excellent example of their overuse and assumption that such acronyms are well-known! JRG.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix: Idealized CBL Set-Up

We tested the LES resolution dependence of the eddy-covariance flux determination in an idealized dry convective boundary layer. This was done by initializing the simulation with a profile

$$\theta(t = 0) = \theta^{\text{ref}} + \Gamma z \quad (12)$$

with  $\theta^{\text{ref}} = 290 \text{ K}$  and  $\Gamma = 6 \text{ K km}^{-1}$ . The boundary layer is cooled through a constant radiative cooling of  $6 \text{ K day}^{-1}$  over the entire depth. This compensates a constant subsidence of  $w_s = 0.01 \text{ m s}^{-1}$  such that the profile of  $\theta$  above the boundary layer remains constant in time. The surface pressure is set to  $1 \times 10^5 \text{ Pa}$ , and the surface heat flux is set to  $\langle w''\theta'' \rangle_0 = 0.06 \text{ K m s}^{-1}$ . A geostrophic wind speed of  $u_g = 2 \text{ m s}^{-1}$  is imposed, constant in height.

This set-up has the advantage that a steady-state solution is guaranteed, and a minimum of external length scales is introduced a priori. Further details of this type of LES set-up are provided in Schalkwijk et al. (2013). At initialization, the simulations are perturbed with a random (white noise) perturbation that has an amplitude of 0.1 K to facilitate the initial development of turbulence. Simulations are performed with varying grid discretization to resemble the YOGA and YOGA-HR datasets as closely as possible.

## References

- Bosveld FC, van der Vliet JG, Monna WAA (1999) The KNMI Garderen experiment: micro-meteorological observations 1988–1989. Technical report, 'KNMI', 53 pp

- de Roode SR, Duynkerke PG, Jonker HJJ (2004) Large-eddy simulation: how large is large enough? *J Atmos Sci* 61(4):403–421
- Desjardins RL, MacPherson JJ, Mahrt L, Schuepp P, Pattey E, Neumann H, Baldocchi D, Wofsy S, Fitzjarrald D, McCaughey H, Joiner DW (1997) Scaling up flux measurements for the boreal forest using aircraft-tower combinations. *J Geophys Res* 102(D24):29,125
- Foken T (2008) The energy balance closure problem: an overview. *Ecol Appl* 18(6):1351–1367
- Foken T, Wimmer F, Mauder M, Thomas C, Liebethal C (2006) Some aspects of the energy balance closure problem. *Atmos Chem Phys* 6(12):4395–4402
- Hendricks Franssen HJ, Stöckli R, Lehner I, Rotenberg E, Seneviratne SI (2010) Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations. *Agric For Meteorol* 150(12):1553–1567
- Heus T, van Heerwaarden C, Jonker H, Pier Siebesma A, Axelsen S, van den Dries K, Geoffroy O, Moene A, Pino D, de Roode S, de Arellano J (2010) Formulation of the Dutch Atmospheric Large-Eddy Simulation (DALES) and overview of its applications. *Geosci Model Dev* 3(2):415–444
- Huang J, Lee X, Patton EG (2008) A modelling study of flux Imbalance and the influence of entrainment in the convective boundary layer. *Boundary-Layer Meteorol* 127(2):273–292
- Inagaki A, Letzel MO, Raasch S, Kanda M (2006) The impact of the surface heterogeneity on the energy imbalance problem using LES. *J Meteorol Soc Jpn* 84(1):187–198
- Jacobs AFG, Heusinkveld BG, Holtslag AAM (2008) Towards closing the surface energy budget of a mid-latitude grassland. *Boundary-Layer Meteorol* 126(1):125–136
- Jonker HJJ, Duynkerke PG, Cuijpers JWM (1999) Mesoscale fluctuations in scalars generated by boundary layer convection. *J Atmos Sci* 56(5):801–808
- Kaimal J, Wyngaard J, Izumi Y, Coté O (1972) Spectral characteristics of surface-layer turbulence. *Q J R Meteorol Soc* 98(417):563–589
- Kanda M, Inagaki A, Letzel MO, Raasch S, Watanabe T (2004) LES study of the energy imbalance problem with eddy covariance fluxes. *Boundary-Layer Meteorol* 110(3):381–404
- Khanna S, Brasseur JG (1998) Three-dimensional buoyancy- and shear-induced local structure of the atmospheric boundary layer. *J Atmos Sci* 55(5):710–743
- Kroon PS, Hensen A, Jonker HJJ, Ouwersloot HG, Vermeulen AT, Bosveld FC (2010) Uncertainties in eddy covariance flux measurements assessed from CH<sub>4</sub> and N<sub>2</sub>O observations. *Agric For Meteorol* 150(6):806–816
- Lee X, Massman W, Law B (2005) Handbook of micrometeorology, atmospheric and oceanographic sciences library, vol 29. Kluwer Academic Publishers, Dordrecht, 250 pp
- LeMone MA (1973) The structure and dynamics of horizontal roll vortices in the planetary boundary layer. *J Atmos Sci* 30(6):1077–1091
- Lenschow DH, Mann J, Kristensen L (1994) How long is long enough when measuring fluxes and other turbulence statistics? *J Atmos Ocean Technol* 11(3):661–673
- Leuning R, van Gorsel E, Massman WJ, Isaac PR (2012) Reflections on the surface energy imbalance problem. *Agric For Meteorol* 156:65–74
- Mahrt L (1998) Flux sampling errors for aircraft and towers. *J Atmos Ocean Technol* 15(2):416–429
- Mahrt L (2010) Computing turbulent fluxes near the surface: needed improvements. *Agric For Meteorol* 150(4):501–509
- Moeng CH, Sullivan PP (1994) A comparison of shear- and buoyancy-driven planetary boundary layer flows. *J Atmos Sci* 51(7):999–1022
- Moore CJ (1986) Frequency response corrections for eddy correlation systems. *Boundary-Layer Meteorol* 37(1–2):17–35
- Neggess RAJ, Siebesma AP, Heus T (2012) Continuous single-column model evaluation at a permanent meteorological supersite. *Bull Am Meteorol Soc* 93(9):1389–1400. doi:[10.1175/BAMS-D-11-00162.1](https://doi.org/10.1175/BAMS-D-11-00162.1)
- Schalkwijk J, Griffith EJ, Post FH, Jonker HJJ (2012) High-performance simulations of turbulent clouds on a desktop PC: exploiting the GPU. *Bull Am Meteorol Soc* 93(3):307–314
- Schalkwijk J, Jonker HJ, Siebesma A (2013) Simple solutions to steady-state cumulus regimes in the convective boundary layer. *J Atmos Sci* 70(11):3656–3672
- Schalkwijk J, Jonker H, Siebesma A, Bosveld F (2014) A year-long large-eddy simulation of the weather over Cabauw. *Mon Weather Rev*. doi:[10.4121/uuid:d48e2ebb-386e-4bb1-9331-955b3ef3c293](https://doi.org/10.4121/uuid:d48e2ebb-386e-4bb1-9331-955b3ef3c293)
- Schalkwijk J, Jonker HJJ, Siebesma AP, Bosveld FC (2015) A year-long large-eddy simulation of the weather over Cabauw: an overview. *Mon Weather Rev* 143(3):828–844
- Schmid HP (1994) Source areas for scalars and scalar fluxes. *Boundary-Layer Meteorol* 67(3):293–318
- Schmidt H, Schumann U (1989) Coherent structure of the convective boundary layer derived from large-eddy simulations. *J Fluid Mech* 200(1):511–562



- Steinfeld G, Letzel MO, Raasch S, Kanda M, Inagaki A (2007) Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study. *Boundary-Layer Meteorol* 123(1):77–98
- Stephens GL (2005) Cloud feedbacks in the climate system: a critical review. *J Clim* 18(2):237–273
- Twine TE, Kustas WP, Norman JM, Cook DR, Houser PR, Meyers TP, Prueger JH, Starks PJ, Wesely ML (2000) Correcting eddy-covariance flux underestimates over a grassland. *Agric For Meteorol* 103(3):279–300
- van Meijgaard E, van Ulft LH, de Berg WJ, Bosveld FC, den Hurk BJJM, Lenderink G, Siebesma AP (2008) The KNMI regional atmospheric climate model RACMO version 2.1. Technical report, KNMI, Postbus 201, 3730 AE, De Bilt, The Netherlands, 43 pp
- Vickers D, Mahrt L (2003) The cospectral gap and turbulent flux calculations. *J Atmos Ocean Technol* 20(5):660–672
- Viterbo P, Beljaars ACM (1995) An improved land surface parameterization scheme in the ECMWF model and its validation. *J Clim* 8(11):2716–2748