# The influence of idealized surface heterogeneity on virtual turbulent flux measurements

Frederik De Roo[1] and Matthias Mauder[1,2]

[1] Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology (KIT), Kreuzeckbahnstrasse 19, 82467 Garmisch-Partenkirchen, Germany
[2]Institute of Geography and Geoecology (IfGG), Karlsruhe Institute of Technology (KIT), Kaiserstrasse 12, 76131 Karlsruhe, Germany

*Correspondence to:* Frederik De Roo (frederik.deroo@kit.edu)

**Abstract.**

The imbalance of the surface energy budget in eddy-covariance measurements is still an unsolved problem. A possible cause is the presence of land surface heterogeneity. The influence of surface heterogeneities on the atmospheric boundary layer has been intensively investigated since two decades. Previous studies found that heterogeneities of the boundary-layer scale or

5    larger are most effective in influencing the boundary layer turbulence. Subsequent large-eddy simulations showed that also the turbulent fluxes are influenced by large-scale organized structures in the boundary layer. However, the precise influence of the surface characteristics on the energy imbalance of measurements in the surface layer and its partitioning is still unknown. To investigate the influence of surface variables on all the components of the flux budget under convective conditions, we set up a systematic parameter study by means of large-eddy simulation. For the study we use a virtual control volume approach,

10    which allows the determination of advection by the mean flow, flux-divergence and storage terms of the energy budget at the virtual measurement site, in addition to the standard turbulent flux. We focus on the heterogeneity of the surface fluxes and keep the topography flat. The surface fluxes vary locally in intensity and these patches have different length scales. Intensity and length scales can vary for the two horizontal dimensions but follow an idealized chessboard pattern. Our main focus lies on heterogeneities of length scales of the kilometer scale, and length scales of one order of magnitude smaller. For hetero-

15    geneities of these two types, we investigate the average response of the fluxes at a number of virtual towers, when varying the heterogeneity length within the length scale and when varying the contrast between the different patches. For each simulation, virtual measurement towers were positioned at functionally different positions (e.g. downdraft region, updraft region, at border between domains, etc.). Furthermore, we seek correlators for the energy balance ratio and the energy residual in the simulations. Besides the expected correlation with measurable atmospheric quantities such as the friction velocity, boundary-layer

20    depth and temperature and moisture gradients, we have also found an unexpected correlation with the temperature difference between sonic temperature and surface temperature.

*Copyright statement.* TEXT

# 1 Introduction

## 1.1 The role of landscape heterogeneity in the energy balance closure problem

The interpretation of the turbulent fluxes of latent and sensible heat at the Earth's surface still suffers from the unresolved energy balance closure problem of the eddy covariance (EC) measurement technique. That is, the measured turbulent fluxes do not

5 close the available energy at the earth's surface (e.g., Foken, 2008; Leuning et al., 2012). There is an ongoing debate whether the missing energy can perhaps be solely described by additional missing terms related to energy conversion and storage or that the imbalance is a consequence of measurement errors in the velocity measurement due to flow distortion from the sonic anemometer pins. With respect to flow distortion, Horst et al. (2015) quoted an error of maximal 5%, but Kochendorfer et al. (2012) and Frank et al. (2013) claimed an error up to 15%. In response to the 15% error, one of us (Mauder, 2013) has pointed

10 out some counterevidence, and a recent modeling study by Huq et al. (2017) on flow distortion did not find evidence for such large errors either. In short, it is unlikely that the previously mentioned issues can explain the fact that very different sites around the world often exhibit an imbalance of more than 20% (e.g., Wilson et al., 2002; Hendricks-Franssen et al., 2010; Stoy et al., 2013).

In fact, the studies by Mauder et al. (2007) and Stoy et al. (2013) have shown that a common property among sites that

15 do not close the energy balance, is a more pronounced surface heterogeneity on the landscape-scale. This motivates us to investigate the energy balance closure problem in the context of landscape heterogeneity. Moreover, Stoy et al. (2013) also found a good correlation between the friction velocity ($u_*$) and the energy balance closure. This result was reproduced by Eder et al. (2015b) by means of a study combining Doppler wind LiDAR and EC tower data. The same correlation has also been noticed in a recent year-long large-eddy simulation (LES) by Schalkwijk et al. (2016) and in an idealized LES study

20 by Inagaki et al. (2006). In addition, the study of Eder et al. (2015b) could relate the energy balance residual to the mean gradients in the lower boundary-layer, thereby providing more evidence for the connection between the energy imbalance and the presence of quasi-stationary structures in the boundary-layer. These circulations typically arise in heterogeneous terrain but may also develop over a completely homogeneous surface to a lesser extent, depending on the atmospheric stability regime, due to self-organization.

## 25 1.2 The influence of landscape heterogeneity on the boundary-layer structure

The influence of heterogeneous landscapes on properties of the atmospheric boundary-layer has already been investigated for a few decades with numerical models, primarily large-eddy simulation. We will summarize a few results that are relevant to the non-closure of the energy balance. Avissar and Chen (1993) obtained significant mesoscale fluxes tied to the terrain heterogeneity. These mesoscale fluxes are carried by the vertical wind of the meso-scale circulations, however, they are not

30 present at the ground level. Raupach and Finnigan (1995) also found that surface heterogeneity induces boundary-layer motions, nevertheless the area-averaged properties, including the fluxes, were not significantly influenced by the heterogeneity or the circulation. At the first glance, both statements appear in conflict with a generic influence of the landscape heterogeneity around a measurement site on the energy balance closure.

On the other hand, Shen and Leclerc (1995) found that the horizontally averaged variances and covariances were influenced by land surface heterogeneity with scales smaller than the boundary-layer depth. This was also confirmed by Raasch and Harbusch (2001). This apparent contradiction can be explained by the fact the resolution of these models was coarse due to computational restrictions at that time, which has a few implications. Firstly, from continuity we indeed expect no vertical
5    meso-scale transport by advection with the mean flow at the lowest grid point, since w=0, but horizontal flux-divergence plays a role, too. Secondly, we should keep in mind that areally averaging over sufficiently large distances represents a form of spatial filtering due to the coarse resolution. Steinfeld et al. (2007) argued that a spatial filtering method will yield energy balance closure, whereas single-tower temporal averaging of the sensible heat flux signal in heterogeneous domain suffers from low-frequency contributions due to the shifted co-spectrum.
10   In summary, the previously mentioned studies showed that landscape hetereogeneity can induce mesoscale motions in the boundary-layer, especially for heterogeneity of length scales larger than the boundary-layer height. By using a large-eddy simulation model coupled to a land-surface-scheme, Patton et al. (2005) investigated striplike heterogeneities between 2 and 30 km. They found that the heterogeneities with length scales of 4 to 9 km were the most influential in altering the structure of the boundary-layer. A similar coupled model approach was used by Brunsell et al. (2011) to study three heterogeneity
15   scales (approximately $10^{-1}z_i$, $z_i$, $10 z_i$, with $z_i$ the boundary-layer height). They found that only in the surface layer the length scale of the heterogeneity affected the spectral signature of the turbulent heat fluxes, and signals appeared blended in the mixed layer. Still, for the heterogeneity length of $10z_i$, secondary circulations arising from surface heterogeneity that extend through the whole boundary-layer were found. Furthermore, Brunsell et al. (2011) found that the partitioning between latent and sensible heat was affected by the scale of heterogeneity as the simulations for the intermediate scales led to a higher
20   Bowen ratio. Since the intermediate scales (of scale $z_i$) appear more heterogeneous than the small or the large scales, this points toward the dominant influence of the sensible heat flux. Charuchittipan et al. (2014) also suggested to ascribe a larger fraction of the residual to the sensible heat flux than to the latent heat flux. The influence of synthetic surface heterogeneity on the Bowen ratio was also investigated by Friedrich et al. (2000) who found a non-linear response of the aggregated Bowen-ratio on the underlying land-surface distribution. Bünzli and Schmid (1998) investigated idealized heterogeneity by means of
25   a two-dimensional $E - \epsilon$ model and found good correspondence with an analytical averaging scheme based on the context of a numerical blending height.

Although the above findings indicate that surface heterogeneity at scales of boundary-layer depth and larger can couple to the full boundary layer, surface heterogeneity at scales considerably smaller than the boundary-layer height appears to be blended, as observed by Raupach and Finnigan (1995). Furthermore, Avissar and Schmidt (1998) found that under a mild
30   background wind, the influence of surface heterogeneity is quickly destroyed, in accordance with the findings of Hechtel et al. (1990). However, Maronga and Raasch (2013), who performed LES simulations for the response of the convective boundary layer in realistic heterogeneous terrain, advised that sufficient time and ensemble averaging is needed to extract the heterogeneity-induced signal, and they concluded that the upstream surface conditions can still influence the boundary-layer properties under light winds. Albertson and Parlange (1999) showed that blending of the surface heterogeneity appears even
35   under convective conditions, except for very large heterogeneities. However, Suehring and Raasch (2013) suggest that the

blending of the surface follows from insufficient averaging. Therefore an apparent blending does not necessarily imply that small-scale surface heterogeneity could not have an influence on the energy budget at the surface. However, if smaller scales are indeed completely blended in the mixed layer and therefore do not lead to circulations that involve the full boundary-layer, then we cannot expect non-surface layer properties (say, bulk gradients in the mixed layer or entrainment parameters) to correlate

5 well with the energy balance residual. Though, even in the blended case small scale heterogeneity could still influence the surface energy budget through motions in the surface layer when the latter survive half-hour averaging. Indeed, for suburban terrain Schmid et al. (1990) noted significant differences in energy balance ratios at scales of $10^2 - 10^3 \, \mathrm{m}$, presumably due to micro-advection between the patches of different surface type.

### 1.3  Scope of this paper

10 Acknowledging the connection between the energy imbalance and quasi-stationary flow on the one hand, and quasi-stationary flow and surface heterogeneity on the other hand, we will investigate the effect of surface heterogeneity on the energy balance closure problem in this work. To this end, we will study a series of synthetic idealized landscapes that consist of a chess-board pattern of surface fluxes with different amplitude and different wavelength in the $x$ and the $y$ direction. We will quantify the average influence on virtual tower data, and perform principal component analysis to link the energy balance ratio with surface

15 characteristics, boundary-layer properties and turbulence statistics. As hinted to in Brunsell et al. (2011), in heterogeneous terrain the sensible heat flux appears more important for the imbalance at the intermediate length scales considered in their work, and we shall therefore focus on simulations that are practically dry (we have added a very small moisture flux). In addition, as both the lack of closure and the strength of the secondary circulations are most pronounced for strongly convective conditions, we will likewise focus on strong to free convection. Despite the issues related to blending, we will focus on

20 heterogeneity of length scales between $10^2 - 10^3 \, \mathrm{m}$ since for these scales the energy imbalance is most pronounced. The intermediate scales of $\mathrm{O}(10^3 \, \mathrm{m})$ are of the order of the boundary-layer depth under typical convective conditions for mid-latitude afternoons, whereas the smaller scales of $\mathrm{O}(10^2 \, \mathrm{m})$ are of the order of the surface-layer height. To keep the terminology more general than typical convection for mid-latitude afternoons, we will refer to them as heterogeneity of kilometer scale and hectometer scale, however. According to the classification of Orlanski (1975) these length scales are at the lower end of the

25 meso-gamma-scale and at the upper end of the micro-alpha-scale, respectively.

Previous investigations with LES on the energy budget had been limited to more regular terrain with at least one homogeneous dimension, see the works of e.g. Kanda et al. (2004), Inagaki et al. (2006), Steinfeld et al. (2007), or Huang et al. (2008). Typically, the storage term was subtracted from the surface flux and only the vertical components of the energy balance were considered: i.e. the turbulent flux and a meso-scale flux (i.e. vertical advection) arising from turbulent organized structures

30 (TOS) or heterogeneity-induced meso-scale motions (TMC). On the contrary, we will also analyze the contribution of the storage flux to the energy imbalance explicitly. Furthermore, the results presented there hold for the domain-averaged imbalance, and the method used is limited to heterogeneous terrain with at least one homogeneous dimension. In this work, however, we can extend the analysis of the energy budget to a full budget of the turbulent fluxes, by including additional terms stemming from horizontal advection by the mean flow. We take full account of all horizontal and vertical energy balance components

with a so-called control volume approach, as in Finnigan et al. (2003), Wang (2010), and Eder et al. (2015a). As such, a study of two-dimensional heterogeneous domain becomes possible.

Let us stress again the research questions of this paper. The first aim is to investigate the average influence on virtual flux measurements of land surface heterogeneity in the form of a variable surface heat flux, for a given length scale of the heterogeneity. We focus on length scales of the order of kilometer, and also on length scales of the order of hectometers. The second aim is to correlate the simulated energy balance ratio to various observables that can be obtained from the simulation output and that are also measurable in a realistic setting.

## 2 Methods

For our simulations we have made use of the LES model PALM (Maronga et al., 2015). More precisely, we ran our simulations with PALM version 3.9. Relevant parameters of the simulation setup are summarized in Table 1, the grid spacing is $10\,\mathrm{m}$. The simulations are driven by a spatially variable surface sensible heat flux, the variation of which is controlled by a few parameters. More precisely, the surface sensible heat flux at each surface point $(x, y)$ is determined by

$$H(x,y) = (1 + A_x \mathbb{1}(x/L_x))(1 + A_y \mathbb{1}(y/L_y)) H_0, \tag{1}$$

where $\mathbb{1}$ is a antisymmetric periodic function with period equal to 2, and alternating between $-1$ and $1$, i.e.

$$\mathbb{1}(x) = \sin(\pi x)/|\sin(\pi x)|. \tag{2}$$

The amplitudes of the two-dimensional surface heat flux is given by $A_x$ and $A_y$ and the periods by $L_x$ and $L_y$. $H_0$ is the average surface heat flux. In Fig. 1 we plot an example of a synthetic surface heat flux as in (1). The main aim of this parameter study is to find out the response of virtual towers when the surface parameters are varied. The latter are the length scales $L_x$ and $L_y$ and the amplitudes $A_x$ and $A_y$. For determining the average behavior under the varying surface fluxes, no weighing is applied to a particular configuration of the parameters, all amplitudes and length scales under consideration are treated equally. In Tables 2–3 we have summarized the range of the parameters that determine the landscape heterogeneity. The boundary conditions of the simulations are periodic in the lateral dimensions. For the velocity we have Dirichlet conditions at the top and the bottom. For potential temperature and humidity we have Neumann conditions at the lower boundary (given by the surface fluxes) and at the top boundary (given by the lapse rate at initialization). The top of the domain is situated within a stable inversion layer, which prevents that the turbulence within the boundary-layer is influenced by the vertical domain size. For the vertical velocity we have added an additional small subsidence term in the equations to achieve quasi-stationary conditions for the data analysis, to overcome the destabilizing influence of the surface heat flux. The data is extracted for four hours after two hours of spin-up time. For each hour a data point is collected by averaging over virtual measurements sampled at every second. As our focus lies on the influence of the surface characteristics, we concentrate in the present study on the wind circulations purely generated by the surface heat flux, without complicating the analysis with additional synoptic drivers such as a geostrophic wind. In the lateral dimensions the domain is about 3 to 5 times the boundary-layer depth.

Within the domain, we have positioned nine virtual control volumes. These control volumes are located at functionally different positions with respect to the surface heterogeneity, as can be seen in Figure 1. Four of them are located at the centers of the patches, four others are located on the borders between the patches, and one is located at the crossing of the four patches. The four at the center are positioned in a site that is homogeneous at the site scale, but heterogeneous at the landscape level. The

5   virtual towers that are located at the borders of the patches are positioned at a site which is not homogeneous at the site level. For every control volume around a virtual tower the size is $5 \times 5$ grid points in the horizontal and 5 grid points in the vertical, representing a cube of $(50\,\mathrm{m})^3$. The limits of the control volume are set on the staggered vector grid. The implementation of the energy balance calculation for the control volumes follows the method described in Eder et al. (2015a), which incorporates the approach suggested by Wang (2010). We briefly summarize the main equation, obtained in two steps; first by spatially

10   averaging over the control volume, and then by additional temporally averaging over 1 hour intervals:

$$\left\langle \overline{H} \right\rangle = \left\langle \overline{w'\theta'} \right\rangle + \sum_{s=1}^{4} \left\langle \overline{\mathbf{v}'_\perp \theta'} \right\rangle_s + \left\langle \overline{w} \right\rangle \left\langle \overline{\theta} \right\rangle + \sum_{s=1}^{4} \left\langle \overline{\mathbf{v}}_\perp \right\rangle_s \left\langle \overline{\theta} \right\rangle_s + \left\langle \delta \overline{w}\, \delta \overline{\theta} \right\rangle + \sum_{s=1}^{4} \left\langle \delta \overline{\mathbf{v}}_\perp\, \delta \overline{\theta} \right\rangle_s + \left\langle \int \overline{\frac{\partial \theta}{\partial t}} \mathrm{d}z \right\rangle. \qquad (3)$$

Here $H$ denotes the surface heat flux, $x$, $y$ and $z$ are the Cartesian coordinates, $w$ the wind component in $z$ direction, $\theta$ the potential temperature, $\mathbf{v}_\perp$ the velocity vector perpendicular to the lateral faces in the $xz$- or $yz$-planes, which are indicated by "$s$" during the summation over the 4 lateral faces. The angular brackets indicate the spatial average over a face of the cube,

15   either lateral ("s"), top or ground surface, and the $\delta$ are the corresponding spatial fluctuations. An overbar indicates a temporal average and the primes the corresponding temporal fluctuations. The term on the left-hand side of the equation is the "true" surface heat flux, whereas the terms of the right-hand side denote the eddy-covariance flux at the top of the control volume, the horizontal flux divergence, the vertical and horizontal advection by the mean flow, the vertical and horizontal dispersive fluxes (Belcher et al., 2012) and the storage of $\theta$ in the control volume. For a pictorial representation of the above formula we refer

20   to Eder et al. (2015a). A positive sign for the directional fluxes means that they point outward of the control volume. Where possible, the Gauß-Ostrogradski theorem has been used to reformulate a divergence within the control volume as a surface term.

From a control volume point of view the net fluxes through the faces are what balances the storage term inside the volume, and in this manner advection effects are automatically included in the energy balance of the volume. Of course, in analogy with

25   measurements, the fluctuations at the top face yield the "virtually measured" turbulent heat flux: first the temporal correlations are calculated, then a spatial average over the upper face of the volume is calculated. The latter average improves the statistical significance of the virtual measurement. Although the subgrid fluxes become small at the height of the control volume, we nevertheless include the vertical component of the subgrid flux into the turbulent heat flux. In this manner we can also capture the highest-frequency correlations. Real data from measurement towers is usually sampled up to $10 - 50\,\mathrm{Hz}$, whereas for

30   computational efficiency our simulation advances with a time step of one second, i.e. our simulated data is obtained at $1$ Hz. A higher sampling frequency would not resolve the turbulence better, as the resolution of the latter is limited by the grid spacing. The part of the total turbulent flux that is not captured directly by the resolved turbulent flux by 1-Hertz sampling is transported by the subgrid turbulent flux. For the advective components we have made a distinction between advection due to the mean flow versus advection due to the horizontal flux-divergence. In complex terrain we do not know a well-defined choice

of reference for the base temperature, in contrast to the base temperature in homogeneous terrain that appeared in Webb et al. (1980). Therefore we have avoided introducing a base temperature altogether by adding up the advection by the mean flow components. The virtual measurement height is quite high, but this is due to the vertical resolution and the need for sufficient grid points in the vertical direction to suppress the influence of the subgrid-fluxes whence the turbulence becomes sufficiently

5     resolved.

For investigating the response of the virtual tower measurements to the changes in the parameters, and for investigating the correlation between the measured variables, a principle component analysis (PCA) is applied. We will present many of our results in correlation biplots introduced by Gabriel (1971, 1978). For a pedagogical description of biplots see e.g. Greenacre (2010). Correlation biplots offer a picture of the relationship between interdependent variables. The variables are projected on

10    the first principal components and these vectors are plotted. The inner product between the variable vectors (hence, the product of their length and the cosine of the angle) directly measures their correlation. The scree plots related to the biplots express how much variance is captured by the principal components. The variance explained is a measure for the goodness of the fit by the first principal components only. The better the variable can be explained by the first two principal components, the longer the length of the vector of the variable in a two-dimensional correlation biplot, which is at most unity.

## 15    3    Results and discussion

### 3.1    Virtual tower measurements for landscape heterogeneity of kilometer scale

In Fig. 2, we look at the response of the towers with respect to their location, corresponding to the simulations summarized in Table 2. The towers are ordered according to the available energy at their location. For each tower we have plotted the energy balance residual (available energy minus the turbulent flux), the advection component from the mean flow, the flux-divergence

20    and the storage flux, all normalized by the available energy at the respective tower, with the plot on the left collecting the towers located in the centers of the patches and the plot on the right collecting the towers located at the borders of the patches. The normalized turbulent flux is effectively the energy balance ratio (EBR), but we plot the non-closure $(1 - \text{EBR})$, i.e. the normalized energy balance residual, as the latter's magnitude is of the same size as the remaining components. The normalized fluxes in Fig. 2 are also averaged for all the available data points of the respective tower. That is, we averaged over the data

25    with different time stamps and also over all cases corresponding to the intermediate length scales formed by the choice of surface flux amplitude and length scale. The error bars on the normalized fluxes denote the spread on the virtual measurements of each tower. The spread is naturally quite large since at each tower, different amplitudes for the surface heat flux pattern are considered. The abscissa is the available energy at the tower, but scaled by the median available energy of the nine towers.

To analyze the tower response in more detail, we have separated the towers at the centers from those at the borders. We

30    notice that most towers show the typical underestimation of the energy balance, except for the tower located at the warmest spot where there is an updraft. In fact, the closed energy balance for the tower in the warm patch is similar to a result in Eder et al. (2015a) where the energy balance was closed for the site with a pronounced updraft. The residual clearly depends on the location of the tower: towers located at the centers of the patches are located in a more homogeneous environment and they

exhibit remarkably smaller residuals, as expected. Towers at the borders have up to 10% more imbalance than the adjacent towers in the center. The tower on the corner of the four patches has the worst closure of only 69%. For towers located at the centers, it is evident that the tower sites are locally homogeneous but there is still a clear imbalance. As a consistency check we note that the similar towers (the two towers in the center of the patches with same surface heating; the two sets of two towers

5 on the borders between patches of similar surface heating) behave similarly.

As the residual is formed by the sum of advection by the mean flow, storage and flux-divergence, we now turn our attention to these flux components. It turns out that primarily the advection by the mean flow determines the different residuals, but that the flux-divergence has to be taken into account as well for the full picture. In addition, the storage flux also plays a role, but its signature is independent on the location of the tower, and it is always small, which is understandable for our type of

10 surface conditions: there is only a storage flux due to the heating of the air inside the control volumes. For different towers the allocation of the residual to advection by the mean flow versus flux-divergence varies. At first the behavior of the flux-divergence appears irregular. Let us however take a closer look in Fig. 4, where the flux-divergence, resp. advection by the mean flow, is plotted against the energy balance ratio. In the left panel of Fig. 4 we note that the normalized flux-divergence correlates rather well to the normalized turbulent flux, when we look at their average behavior at each tower. For the individual

15 data points the correlation is nevertheless scattered (not plotted). It is somewhat remarkable that both the towers at the center and those at the borders exhibit a similar average behavior. Indeed, the linear regression is very satisfactory when fitting the $B$-type towers and the $C$-type towers together. We could have made two separate fits, one for each tower type as in Figure 2, but with only 3 or 4 towers of different functionality a linear regression through those 3 or 4 points would carry less meaning than considering all 9 virtual towers together. If we repeat this linear regression for the advection by the mean flow versus the

20 energy balance ratio we see that the linear correlation fits even better (Fig. 3 right panel) but that it has opposite slope. We had expected that the sum of both components correlates very well with the energy balance ratio, since the storage is small and constant, but it is an interesting result that also flux-divergence and advection separately correlate well with energy balance ratio and consequently also with each other.

Finally, we want to remark that, due to computational constraints, the virtual measurement height in our simulations lies at

25 50 meters, which is an order of magnitude larger than the typical tower height over short vegetation with comparable surface roughness. This means that our findings for virtual EC towers cannot be directly transferred to real eddy-covariance towers. Other LES studies of the energy balance closure point towards a larger imbalance at higher z-levels, e.g. Steinfeld et al. (2007), Huang et al. (2008), and Schalkwijk et al. (2016). It remains an open question if we can scale the measurement height (as long as it is in the constant flux layer) with the boundary-layer depth and the scale of the heterogeneity.

30 ## 3.2 Virtual tower measurements for landscape heterogeneity of hectometer scale

In Fig. 4 we repeat the foregoing analysis for the hectometer landscape heterogeneity of one order of magnitude smaller. The difference between the towers is much less pronounced here compared to the kilometer scale. Furthermore, the towers in the center of the patches even behave in the opposite manner when the kilometer and hectometer scales are compared. Indeed, for the hectometer scales the cooler patches have a smaller residual, hence better energy balance closure, up to even a mean over-

closure for the tower in the coolest patch, whereas the energy balance at the hottest patch is not closed. Another example of the opposite behavior is shown by the flux-divergence. In Fig. 4 it is positively correlated with the normalized residual, and in Fig. 5 we notice that the flux-divergence is now indeed anti-correlated with the EBR. The advection by the mean flow is again anti-correlated with the EBR, as it was for the kilometer scale. The storage is again roughly constant for all towers. The likely

5 cause for the different behavior between the two scales of heterogeneity would be the blending of the hectometer landscape heterogeneity due to the virtual tower heights of 50 meter. For the surface heterogeneity of $O(10^2\,\mathrm{m})$ the flux footprint of each of the towers can cover several of the surface patches, regardless of the type of tower.

### 3.3 Correlations with the energy balance ratio

We now turn our attention to the possible connection between the energy balance ratio, the different flux contributions, and

10 other measured variables like the friction velocity. We investigate the correlations for the datasets of Tables 2–3 in separate biplots (Fig. 6). We made one restriction, which is to limit the boundary-layer depth to values larger than 1 km, thereby excluding about 8% of the data, in order to obtain a better representation of the boundary-layer depth within the biplots (when the boundary-layer depths smaller than 1 km are included, the correlation deteriorates). Within the subspace spanned by the first two principal components, the temperature difference between surface and measurement height has the highest correlation to

15 the EBR. The temperature difference between the surface and the measurement height can be considered as proportional to the temperature gradient in the surface layer. We did not find a similarly good agreement of EBR with the temperature gradient from the surface with respect to the center of the mixed layer. In addition, friction velocity and boundary-layer depth cluster together separately, as do the normalized flux-divergence and advection. The former clustering can be explained because in our scenario a higher boundary-layer height arises from more vigorous surface heating, which also leads to stronger circulations patterns,

20 and hence larger $u_*$. Though advection and flux-divergence correlate well with EBR, they cannot be measured independently and therefore cannot be used as independent predictors, different from the temperature difference.

From the scree plots in Fig. 7 that depict the fraction of explained variance, it can be seen that there is still some scatter in the data and the explanatory strength of the two most important principal components is not very large (a combined 61% and 63% respectively). This is also the reason why we have limited the number of variables that are highlighted in the biplots to

25 only 6, as more variables always introduced additional variance, and the explanatory power of the most important principal components decreased further. The biplots represent the correlation between the individual (hourly) data points, whereas the regression lines in Figs. 3 and 5 were obtained from averaging those individual data points over similar tower locations. It is somewhat puzzling that the individual hourly data points leading to the biplots do not show such a good correlation. As we have learned from Fig. 3, there can be a good fit between the parameter-averages of two variables, e.g. normalized flux-divergence

30 and energy balance ratio average, despite the fact that the individual data points do not correlate as well. This highlights the importance of testing parameterizations for the energy balance closure problem on the level of a data ensemble, instead of parameterizing on the level of the individual hourly measurements.

## 4 Conclusions

In this work, we have investigated the effect of idealized surface heterogeneity on the components of the surface energy budget measured at virtual measurement towers, by means of large-eddy simulation. We have found that, for heterogeneities with length scale of order kilometer, there is a clear relation between the energy budget components and the location of the tower with respect to the patches of surface heterogeneity. For heterogeneities with length scales on one magnitude smaller, the response of the different towers appears to depend to a lesser extent on their respective location. Towers located at the borders between patches with different surface heat flux have worse closure than towers located in the center of a patch. Although storage terms are not negligible, the size of the residual depends mostly on the advection and flux-divergence terms. Remarkably, flux-divergence and advection by the mean flow separately correlate very well with the energy balance ratio. For the kilometer scale heterogeneities, advection by the mean flow and flux-divergence behave opposite, but they are positively correlated for hectometer scale heterogeneities. We did not find a strong correlation between the friction velocity and energy balance ratio, but this could be due to the limited range of $u_*$ as we have mainly investigated strong to free convection. However, we found an unexpected correlation between the energy balance ratio and the difference between sonic temperature and surface temperature. Both can be measured readily by typical micrometeorological stations. Therefore, this difference could be a promising predictor for a potential energy balance closure correction. Nevertheless, the virtual measurement height remains a critical issue and further investigations are needed with more realistic measurements heights to confirm such a relationship.

## Appendix A: Example of the heterogeneity length scale of a field site

Even though the focus of this study is on virtual flux measurements, we can look at an example of a real EC measurement site to make a qualitative comparison of these virtual tower measurements with real tower measurements. In a first approximation, the heterogeneity of the landscape around a measurement site can be characterized by the dominant length scale of a suitable surface variable. In Eder et al. (2014), the dominant length scales corresponding to a few sites belonging to the TERENO measurement network (Zacharias et al., 2011) were computed from the Fourier spectrum of the surface roughness. The site of them with the least pronounced topography, the site Fendt, has an effective length scale close to 3 km and a mean EBR of 0.77, which is a typical value for the energy balance ratio (Stoy et al., 2013). The location of the measurement tower in Fendt would correspond to a tower of the central type, and due to its location in the meadow with lower albedo than the forest or the small built-up area, we would assign it to the central tower of the cooler patch. However, the Fourier spectrum of the sensible heat flux may differ from that of the surface roughness. Moreover, the Fourier spectrum for the TERENO site in Fendt exhibits an additional local maximum in its Fourier spectrum of the surface roughness, at 600 m (Eder, pers. comm.). Additionally, it should be noted that even a simplified version of the landscape heterogeneity of Fendt would appear primarily striplike, in contrast the synthetic chessboard pattern here. The EC tower of Fendt is located in a large north-south oriented meadow which is flanked by two forests further away to the west and the east. Despite these apparent differences between our idealised

simulations and the real situation at the Fendt site, the EBR of 0.77 is comparable to the EBR of the virtual towers investigated here for the kilometer heterogeneity.

*Competing interests.* There are no competing interests.

*Acknowledgements.* This work was conducted within the Helmholtz Young Investigators Group "Capturing all relevant scales of biosphere-atmosphere exchange — the enigmatic energy balance closure problem", which is funded by the Helmholtz-Association through the President's Initiative and Networking Fund, and by KIT. We thank the PALM group at Leibniz University Hannover for their open-source PALM code and their support.

## References

Albertson, J. and Parlange, M.: Natural integration of scalar fluxes from complex terrain, Adv Water Res, 23, 239–252, 1999.

Avissar, R. and Chen, F.: Development and analysis of prognostic equations for mesoscale kinetic energy and mesoscale (subgrid-scale) fluxes for large scale atmospheric models, J Atmos Sci, 50, 3751–3774, 1993.

5   Avissar, R. and Schmidt, T.: An Evaluation of the Scale at which Ground-Surface Heat Flux Patchiness Affects the Convective Boundary Layer Using Large-Eddy Simulations, J Atmos Sci, 55, 2666–2689, 1998.

Belcher, S., Harman, I., and Finnigan, J.: The Wind in the Willows: Flows in Forest Canopies in Complex Terrain, Annu. Rev. Fluid Mech., 44, 479–504, 2012.

Brunsell, N., Mechem, D., and Anderson, M.: Surface heterogeneity impacts on boundary layer dynamics via energy balance partitioning,
10      Atmos Chem Phys, 11, 3403–3416, 2011.

Bünzli, D. and Schmid, H.: The influence of surface texture on regionally aggregated evaporation and energy partitioning, J Atmos Sci, 55, 961–972, 1998.

Charuchittipan, D., Babel, W., Mauder, M., Leps, J., and Foken, T.: Extension of the averaging time in eddy-covariance measurements and its effect on the energy balance closure, Bound-Lay Meteorol, 152, 303–327, 2014.

15   Eder, F., De Roo, F., Kohnert, K., Desjardins, R., Schmid, H., and Mauder, M.: Evaluation of Two Energy Balance Closure Parametrizations, Bound-Lay Meteorol, 151, 195–219, 2014.

Eder, F., De Roo, F., Rotenberg, E., Yakir, D., Schmid, H., and Mauder, M.: Secondary circulations at a solitary forest surrounded by semi-arid shrubland and its impact on eddy-covariance measurements, Agr Forest Meteorol, 211-212, 115–127, 2015a.

Eder, F., Schmidt, M., Damian, T., Träumner, K., and Mauder, M.: Mesoscale Eddies Affect Near-Surface Turbulence Exchange: Evidence
20      from Lidar and Tower Measurements, J Appl Meteorol Clim, 54, 189–206, 2015b.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., and Cleugh, H. A.: A re-evaluation of long-term flux measurement techniques part 1: Averaging and coordinate rotation, Bound-Lay Meteorol, 107, 1–48, 2003.

Foken, T.: The energy balance closure problem: an overview, Ecol Appl, 18, 1351–1367, 2008.

Frank, J. M., Massman, W. J., and Ewers, B. E.: Underestimates of sensible heat flux due to vertical velocity measurement errors in non-
25      orthogonal sonic anemometers, Agr. Forest Meteorol., 171, 72–81, 2013.

Friedrich, K., Mölders, N., and Tetzlaff, G.: On the influence of surface heterogeneity on the Bowen-ratio: A theoretical case study, Theor Appl Clim, 65, 181–196, 2000.

Gabriel, K. R.: The biplot-graphic display of matrices with application to principal component analysis, Biometrika, 58, 453–467, 1971.

Gabriel, K. R.: Analysis of meteorological data by means of canonical decomposition and biplots, J. Appl. Meteorol, 11, 1072–1077, 1978.

30   Greenacre, M.: Biplots in Practice, fundacion BVBA, 2010.

Hechtel, L., Moeng, C., and Stull, R.: The effects of nonhomogeneous surface fluxes on the convective boundary layer: a case study using large-eddy simulation, J Atmos Sci, 47, 1721–1741, 1990.

Hendricks-Franssen, H., Stöckli, R., Lehner, I., Rotenberg, E., and Seneviratne, S.: Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations, Agr Forest Meteorol, 150, 1553–1567, 2010.

35   Horst, T., Semmer, S., and MacLean, G.: Correction of a non-orthogonal and three-component sonic anemometer for flow distortion by transducer shadowing, Bound-Lay Meteorol, 155, 371–395, 2015.

Huang, J., Lee, X., and Patton, E.: A modelling study of flux imbalance and the influence of entrainment in the convective boundary layer, Bound-Lay Meteorol, 127, 273–292, 2008.

Huq, S., De Roo, F., Foken, T., and Mauder, M.: Evaluation of probe-induced flow distortion of Campbell CSAT3 sonic anemometers by numerical simulation, Boundary-Lay Meteorology, in press, doi:10.1007/s10546-017-0264-z, 2017.

5   Inagaki, A., Letzel, M. O., Raasch, S., and Kanda, M.: Impact of surface heterogeneity on energy imbalance, J Meteorol Soc Jpn, 84, 187–198, 2006.

Kanda, M., Inagaki, A., Letzel, M. O., Raasch, S., and Watanabe, T.: LES study of the energy imbalance problem with eddy covariance fluxes, Bound-Lay Meteorol, 110, 381–404, 2004.

Kochendorfer, J., Meyers, T. P., Frank, J., Massman, W. J., and Heuer, M. W.: How well can we measure the vertical wind speed? Implications

10   for fluxes of energy and mass, Bound-Lay Meteorol, 145, 383–398, 2012.

Leuning, R., van Gorsel, E., Massman, W. J., and Isaac, P. R.: Reflections on the surface energy imbalance problem, Agr. Forest Meteorol., 156, 2012.

Maronga, B. and Raasch, S.: Large-eddy simulations of surface heterogeneity effects on the convective boundary layer during the LITFASS-2003 experiment, Bound-Lay Meteorol, 146, 17–44, 2013.

15   Maronga, B., Gryschka, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, M., Sühring, M., and Raasch, S.: The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives, Geosci. Model Dev., 8, 1539–1637, 2015.

Mauder, M.: A Comment on "How well can we measure the vertical wind speed? Implications for fluxes of energy and mass" by Kochendorfer et al, Bound-Lay Meteorol, 147, 329–335, 2013.

20   Mauder, M., Desjardins, R., and MacPherson, I.: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem, J Geophys Res, 112, D13 112, 2007.

Orlanski, I.: A rational subdivision of scales for atmospheric processes, Bull. Am. Meteorol. Soc., 56, 527–530, 1975.

Patton, E., Sullivan, P., and Moeng, C.: The influence of idealized heterogeneity on wet and dry planetary boundary layers coupled to the land surface, J Atmos Sci, 62, 2078–2097, 2005.

25   Raasch, S. and Harbusch, G.: An analysis of secondary circulations and their effects caused by small-scale surface inhomogeneities using large-eddy simulation, Bound-Lay Meteorol, 101, 31–59, 2001.

Raupach, M. R. and Finnigan, J. J.: Scale issues in boundary-layer meteorology: surface energy balances in heterogeneous terrain, Hydr Proc, 9, 589–612, 1995.

Schalkwijk, J., Jonker, H. J. J., and Siebesma, A. P.: An Investigation of the Eddy-Covariance Flux Imbalance in a Year-Long Large-Eddy

30   Simulation of the Weather at Cabauw, Bound-Lay Meteorol, 160, 17–39, 2016.

Schmid, H., Cleugh, H., Grimmond, C., and Oke, T.: Spatial variability of energy fluxes in suburban terrain, Bound-Lay Meteorol, 54, 249–276, 1990.

Shen, S. and Leclerc, M.: How large must surface inhomogeneities be before they influence the convective boundary layer structure? A case study, Q J R Meteorol Soc, 121, 1209–1228, 1995.

35   Steinfeld, G., Letzel, M., Raasch, S., Kanda, M., and Inagaki, A.: Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study, Bound-Lay Meteorol, 123, 77–98, 2007.

Stoy, P., Mauder, M., Foken, T., Marcolla, B., Boegh, E., Ibrom, A., Arain, M., Arneth, A., Aurela, M., Bernhofer, C., Cescatti, A., Dellwik, E., Duce, P., Gianelle, D., van Gorsel, E., Kiely, G., Knohl, A., Margolis, H., McCaughey, J., Merbold, L., Montagnani, L., Papale,

D., and Reichstein, M.: A data-driven analysis of energy balance closure across FLUXNET research sites: The role of landscape-scale heterogeneity, Agr. Forest Meteorol., 171-172, 137–152, 2013.

Suehring, M. and Raasch, S.: Heterogeneity-Induced Heat-Flux Patterns in the Convective Boundary Layer: Can they be Detected from Observations and is There a Blending Height?—A Large-Eddy Simulation Study for the LITFASS-2003 Experiment, Bound-Lay Meteorol, 148, 309–333, 2013.

Wang, W.: The influence of topography on single-tower-based carbon flux measurements under unstable conditions: a modeling perspective, Theor Appl Clim, 99, 125–138, 2010.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Quart. J. R. Met. Soc., 106, 85–100, 1980.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, R., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Agr Forest Meteorol, 113, 223–243, 2002.

Zacharias, S., Bogena, H., Samaniego, S., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J., Papen, H., Priesack, E., Schmid, H., Steinbrecher, R., Rosenbaum, U., Teutsch, F., and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, Vadose Zo J, 10, 955, 2011.
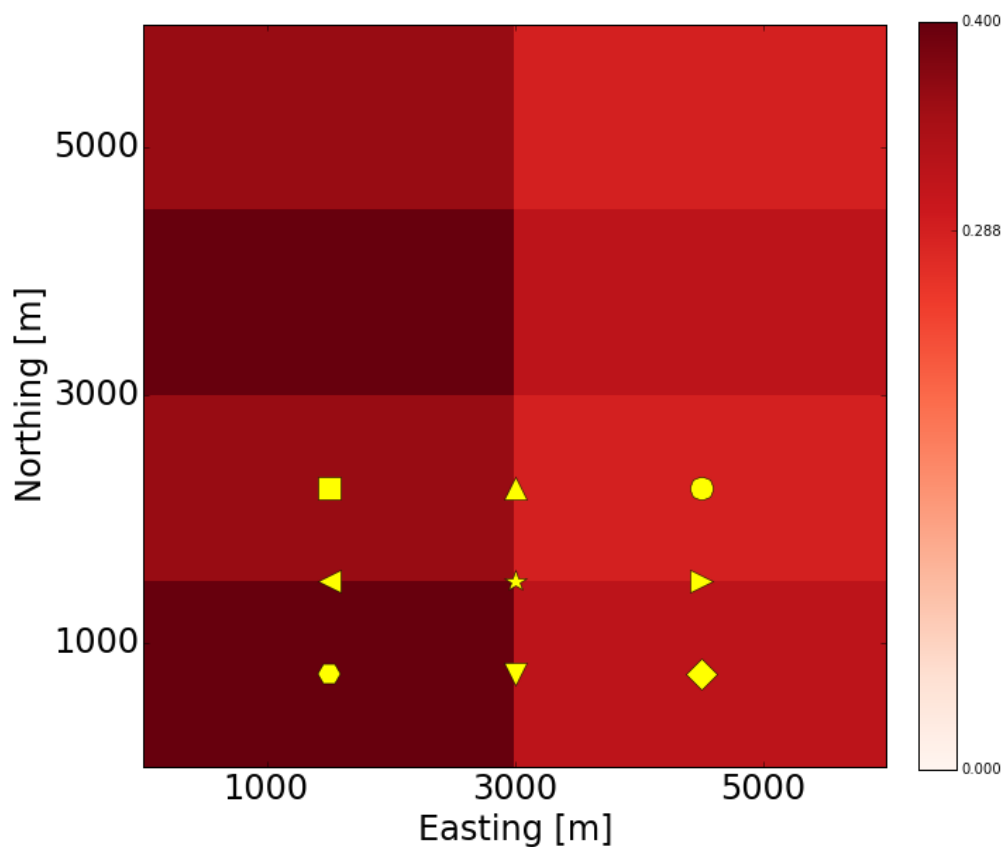
**Figure 1.** Fixed location of the virtual towers for the kilometer scale heterogeneity. The surface heat flux pattern of this example corresponds to $L_x = 3000\,\mathrm{m}$, $L_y = 1500\,\mathrm{m}$. Please note that all the control volumes have the same shape of $5 \times 5 \times 5$ grid points, the symbols are only to distinguish the different types of towers. For the hectometer scale heterogeneity, the towers are located at the similar positions in or in between the patches, only the patches are smaller. The towers fall into two classes: those located at the center of the patches and those located at the borders.
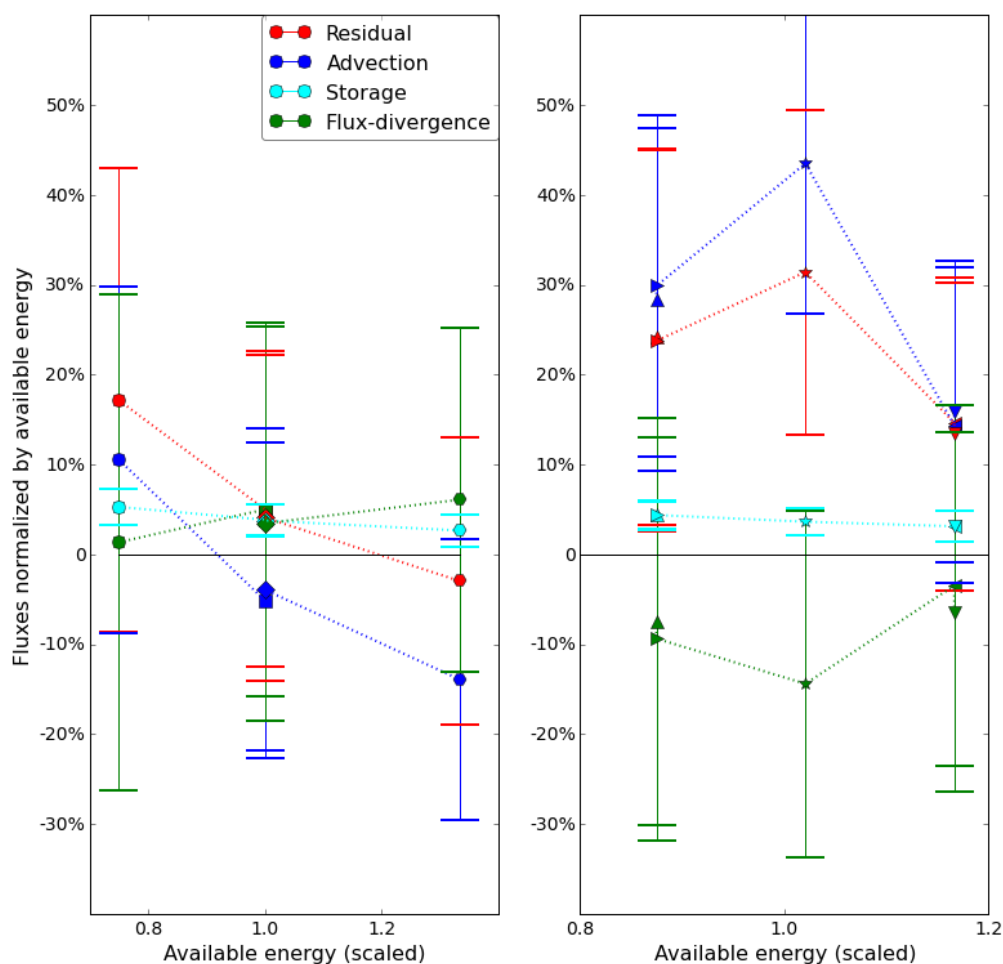
**Figure 2.** Control volume fluxes as a function of available energy (scaled by the median value) for kilometer scale landscape heterogeneity. The fluxes are normalized by the available energy at their respective location, in our setup this means normalization by the surface flux. Please note we have plotted the non-closure (normalized energy balance residual) instead of the energy balance ratio EBR (normalized turbulent flux). Left panel: towers at centers. Right panel: towers at borders. For the symbols, see Fig. 1. The error bars denote the spread over the different cases of surface heterogeneity within the suite of kilometer scale surface heterogeneity. See text for further discussion.
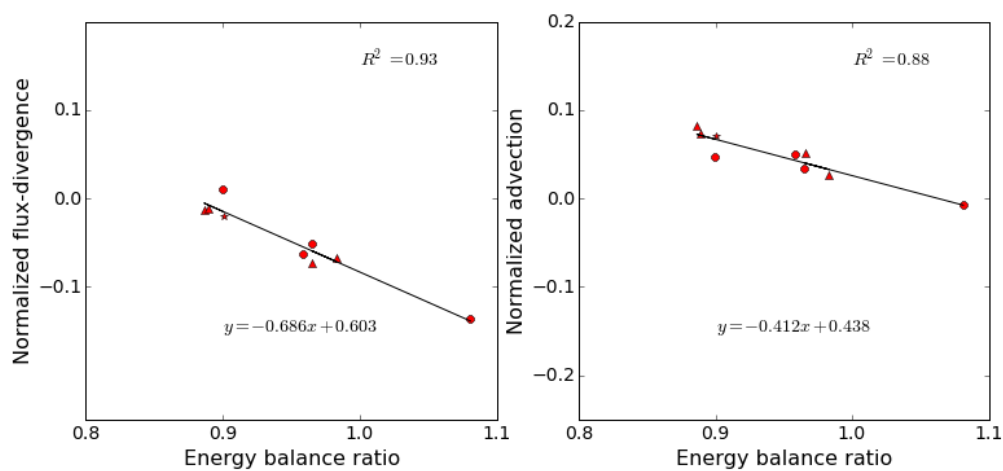
**Figure 3.** Correlation between flux-divergence and EBR for kilometer scale heterogeneity (left panel); correlation between advection and EBR for kilometer scale heterogeneity (right panel)
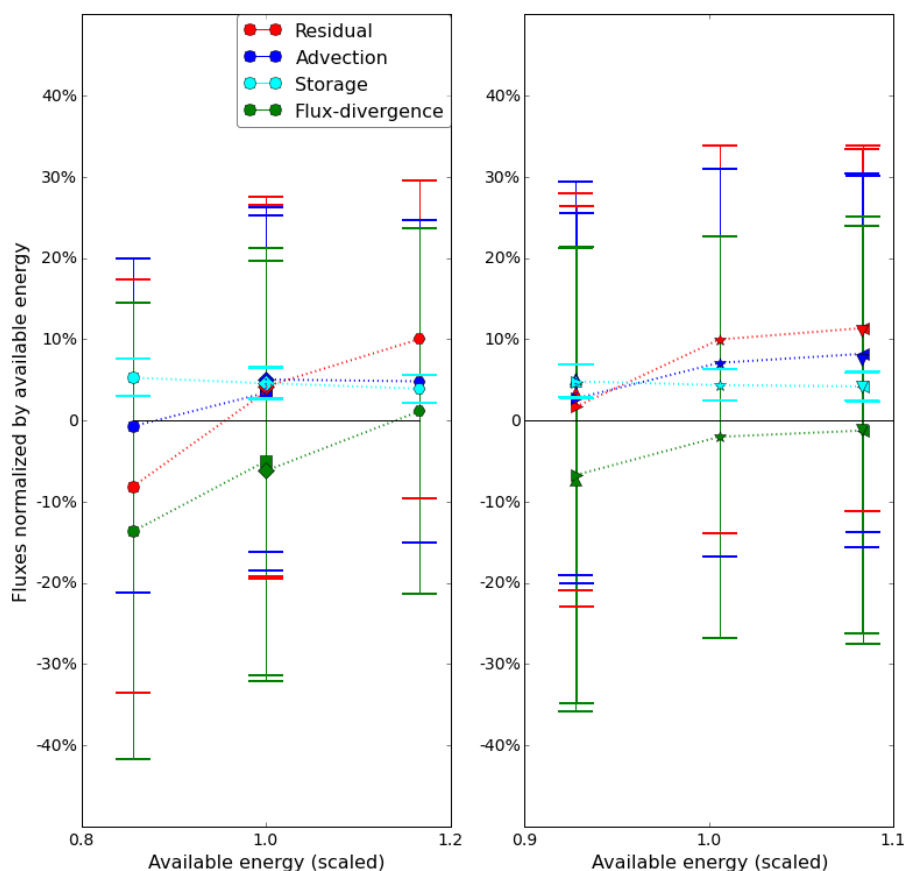
**Figure 4.** Control volume fluxes as a function of available energy (scaled by the median value) for hectometer scale landscape heterogeneity. The fluxes are normalized by the available energy at their respective location, in our setup this means normalization by the surface flux. Please note we have plotted the non-closure (normalized energy balance residual) instead of the energy balance ratio EBR (normalized turbulent flux). Left panel: towers at centers. Right panel: towers at borders. For the symbols, see Fig. 1. The error bars denote the spread over the different cases of surface heterogeneity within the suite of hectometer scale surface heterogeneity. See text for further discussion.
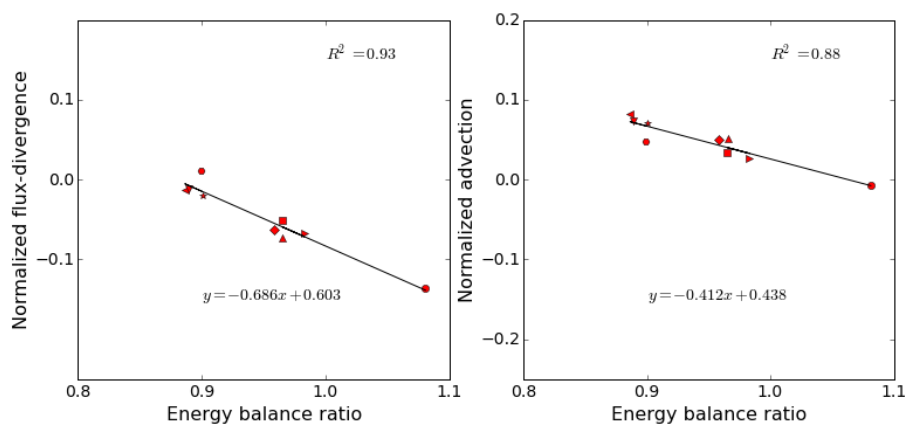
**Figure 5.** Correlation between flux-divergence and EBR for kilometer scale heterogeneity (left panel); correlation between advection and EBR for kilometer scale heterogeneity (right panel)
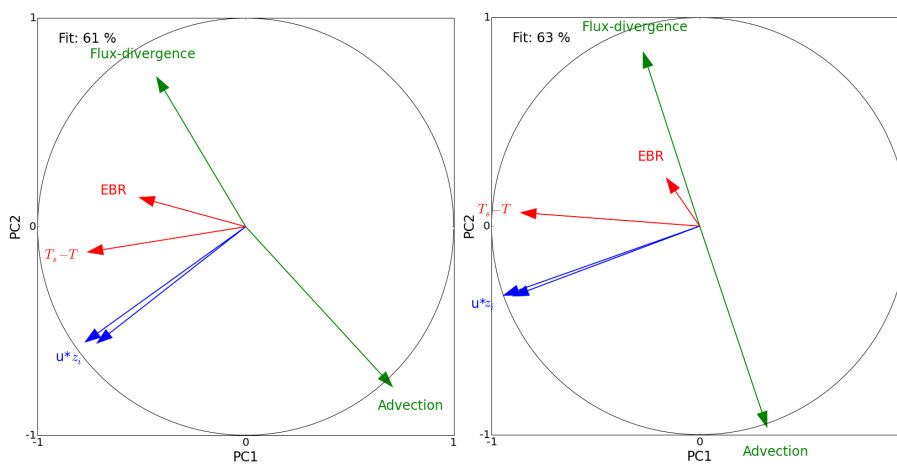
**Figure 6.** Correlation biplots for the kilometer scale heterogeneity (left panel) and the hectometer scale heterogeneity (right panel)
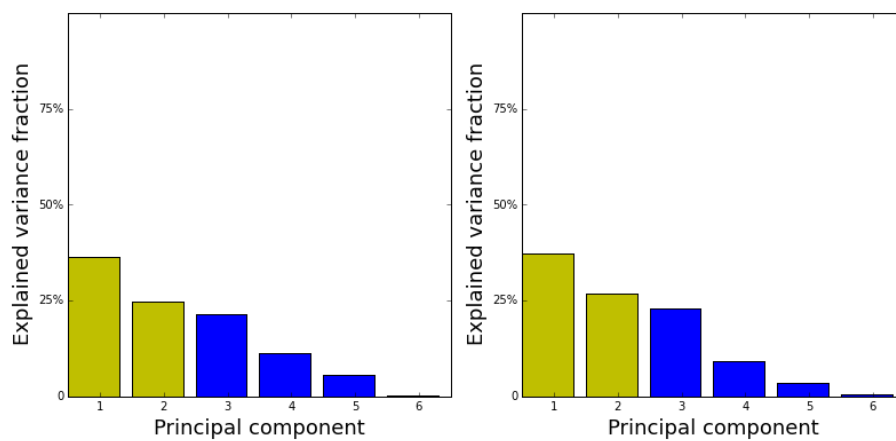
**Figure 7.** Scree plots of the biplot analyses of Fig. 6 with all 6 principal components. Left panel: order kilometer length scale. Right panel: order hectometer length scale. As far as the explained variance is concerned, both orders of length scale behave very similarly.

**Table 1.** Parameters of the LES configuration

| Quantity | Unit | Value |
|---|---|---|
| Number of grid points | (-) | 600 x 600 x 240 |
| Spatial resolution (dx, dy, dz) | (m) | 10.0, 10.0, 10.0 |
| Domain size | ($m^3$) | 6,000 x 6,000 x 2,400 |
| Temporal resolution | (s) | 1.0 |
| Spin-up time | (s) | 7,200.0 |
| Data capture | (s) | 7,200.0 - 14,400.0 |
| Averaging interval | (s) | 3,600.0 |
| Size of the control volume | ($m^3$) | 50.0 x 50.0 x 50.0 |
| Approximate walltime for one simulation | (core-hours) | 5500 |
| Roughness length | (m) | 0.1 |
| Surface moisture flux | ($kg\ kg^{-1}\ m\ s^{-1}$) | $5e^{-6}$ |

**Table 2.** Parameters of the landscape heterogeneity (kilometer scale)

| Strong to free convection | | $6 \times 6 \times 2 \times 2$ cases |
|---|---|---|
| Average surface heat flux | (K m s$^{-1}$) | 0.25 |
| Amplitude $x$ ($A_x$) | (K m s$^{-1}$) | 0.0 ; 0.1 ; 0.2 ; 0.3 ; 0.4 ; 0.5 |
| Amplitude $y$ ($A_y$) | (K m s$^{-1}$) | 0.0 ; 0.1 ; 0.2 ; 0.3 ; 0.4 ; 0.5 |
| Length scale $x$ ($L_x$) | (m) | 1,500.0 ; 3,000.0 |
| Length scale $y$ ($L_y$) | (m) | 1,500.0 ; 3,000.0 |
| Boundary-layer height | (km) | From 1.4 to 2.2 |
| Obukhov length | (m) | From -36.1 to -0.04 (average -3.96) |

**Table 3.** Parameters of the landscape heterogeneity (hectometer scale)

| Strong to free convection | | $6 \times 6 \times 2 \times 2$ cases |
| --- | --- | --- |
| Average surface heat flux | (K m s$^{-1}$) | 0.25 |
| Amplitude $x$ ($A_x$) | (K m s$^{-1}$) | 0.0 ; 0.1 ; 0.2 ; 0.3 ; 0.4 ; 0.5 |
| Amplitude $y$ ($A_y$) | (K m s$^{-1}$) | 0.0 ; 0.1 ; 0.2 ; 0.3 ; 0.4 ; 0.5 |
| Length scale $x$ ($L_x$) | (m) | 200.0 ; 400.0 |
| Length scale $y$ ($L_y$) | (m) | 200.0 ; 400.0 |
| Boundary-layer height | (km) | From 1.5 to 2.2 |
| Obukhov length | (m) | From -53.8 to -0.01 (average -5.75) |