Response to Review 2

Dear Authors, This study used the aerial thermal imaging to detect the turbulence characteristics near the ground surface. It is compared with the ground-based sonic anemometer measurement in an EC site. Although the time-sequential TIR imaging has already been used for the same purpose, the use of UAS to capture wider area has not been accomplished before and its knowledge is useful. – Thank You, we deeply appreciate this positive response. While the current study is rather experimental, we hope that it lays down the foundations for a new generation of turbulence studies. Meanwhile, there are some points to be clarified especially in the data analysis to keep the generality of the discussion. – We did our best to address the points You have raised.

Comments: L91 Delete ",." - Done
L134 Please explain what is the Structural Similarity Index. Also, how are the RMSE and SNR defined in this process? – Thank You for the question – detailed explanations have been added as new Appendix C.
L181 Please explain why the authors selected 150 m in this process. – The document must have been incorrectly presented on your screen; the text actually reads "at the spatial scales of 1-50 m". So the wavelet analysis was done for those spatial scales. However, the explanation is now added, such a range was chosen based on the assumption that the maximum eddy size well represented by a TIR image cover"ing ca. 300x400 m would be roughly 100 m."
L182, L185 Please show how sensitive these parameters (14 m and +-3.5 m) for thelatter discussion (i.e. the ratio of the length and width of the isolated structures, Figs.11and 12). – Thank You for raising this important point. Indeed, they introduce some sensitivity in the processing and need to be treated with care. The wavelet decomposition scale is more sensitive than the filtering threshold. The following comments were added: "It must be noted that the wavelet transform scale is a sensitive parameter requiring adjustment to the scale of the dominant eddies; an excessively small scale value would lead to improper fractionation of the eddies, while a scale value which is too high would result in grouping where the eddies are apparently separate.", and "The threshold for this filtering operation should also be chosen with care, as the slopes separating positive and negative wavelet regions can be steep (see the effect of the ±3.5 threshold in Fig. 3c)."
L210 Please describe how large the interrogation area in meters, and also the timeincrement to derive the velocity in sec. – The resolution of the georeferenced images is 1m/pix, therefore the interrogation area was 100 m; as per the specifics of the PIV method, the sequence is divided into "frames", i.e. pairs of images, and the wind field is computed for each frame before they are averaged for the given periods of interest – as a result, one can say that the time increment is 2s.
P212 Please describe the mean height of the roughness elements (vegetations) of theobservation area. – Description added. It is mainly the sedges that create nearly all of the roughness at the site; they grow to the mean height of about 0.25m.
P219 How is the flux footprint used in the latter analysis and/or discussion? – explanation has been added. EC footprint was used solely for reconciling the UAG surface temperatures with the sonic temperature.
L285 Is this FFT analysis applied for the time series of the surface temperature at a certain point in the images, and later it is averaged horizontally? – This is exactly correct, that's the ways the UAS FFT spectrum has been obtained. An explanation has been added in the end of section 2.2.6: "FFT was also applied to the thermal sequences in the temporal domain thus: first, FFT was performed on individual pixel time series, and those pixel-wise spectra were averaging to yield a single FFT spectrum of a flight."
Is there any reasonwhy the two spectra in Fig,7 are different at the low frequency region? – I think this is the region where the spectra are not representative of the turbulent fluctuation as they are calculated from short (<=20 min) records. The difference increases at frequencies lower than 1e-4 Hz, to judge from the figure, which corresponds to a period comparable to the length of the record.

Another possibility lies in the contribution of poorly understood artifacts in the thermal data which the present methods failed to eliminate; those should be addressed in the future studies. However, I don't have any good answer as to what these may be.

Are there -1power law region (e.g. Drobinski et al. 2004) both in the spectra of EC Ts and UAVTg? – we did specifically attempt to detect the -1 power law, given its importance for the interpretation of turbulence origin, but the evidence remained inconclusive. Maybe the large-scale scturcture of the ABL during the flights did not favor the turbulent organization which leads to the -1 power law relationship.

L308 "The relatively small..." It is difficult to understand this sentence just from thecorresponding figures (Fig.8a,b). – I would say 5-50 m based on visual inspection. In any case, the main point here is the contrast between the different regimes, which is rather apparent from the images of Fig.8.

L318 "Wall effects at the forest edge..." This is not certain yet from the snapshots of the temperature anomaly. It should be evaluated, for example, after ensemble or temporal average to extract the effect of the heterogeneous roughness. – Actually, the forest edge is most pronounced in the temperature standard deviation (Figure 6). I believe that this approach for visualizing the step-change in surface roughness is analogous to what You are proposing.

Figure 8 Is there any extra process to obtain these velocity vectors after the image correlation calculation? Please describe details about it if there are any (i.e. smoothing, averaging, handling of the error vectors, etc.). – Done (added in 2.2.7). Also, please describe how the result of PIV calculation is sensitive to the accuracy of the image registration and/or georefer-encing. – This is an important question, now answered in 2.2.7.

L335 "the EC WS was higher..." This is interesting since the movement of the surfacetemperature structures seems to be associated with the convective thermal structuresin this observation, which probably move faster than the bottom air whose speed ismeasured by EC (z=3m, below RSL) if the mean wind profile follow the typical log-law plus MOS function. Please explain why EC WS is faster. Some discussion were seenin Garai et al (2013) and Inagaki et al. (2013).

- This is a good point, thank You for raising it. Due to the uncertainty in the PIV process, the PIV "flow" velocity has a random (and possibly systematic?) uncertainty which can be estimated at 30%. It is also important to bear in mind that the specific input for the PIV was the wavelet transform at the scale of 5m, hence the PIV output shows the velocity field of the smaller eddies 5-10 m in size. Some previous research (this is now summarized in the updated Introduction) indeed found coherent structures to advect faster than the mean wind near the ground, but in the case of smaller eddies which are attached to the surface and well-coupled with the ground roughness, can well be advected at roughly the mean wind speed measured by the EC.

Figure 9 Are the periods of the lower wavelet power, which are the majority of theentire period, corresponding to the quiescent period as in Fig.8b? Please describe what happens in it. – Precisely so, the low wavelet powers (the bluish colours) in Fig. 9 correspond to such "quiescent periods". Large, well-defined structures contribute the most to the wavelet powers (especially at scales approaching the limit of 50 m), so when they were absent, the wavelet power dropped. A clarifying sentence has been added to the discussion of Fig.9.

L365 Probably, the spectral power s at 128m and 10m are selected due to the FOV and the resolution of the observation. Are they representing the entire spectral shapes? Please describe, for example, they are within the energy containing range or the inertial subrange if those wavenumber spectra follow the ordinal spectral shape of turbulence. – You are absolutely correct. 10 m is the limit is the eddy size which is well-represented in the measurements, and relevant for the coherent structure discussion; 128 m is the largest scale that still fits in the FOV. Undoubtedly, there are larger coherent structures (i.e. VLSMs) some kilometers long which fall outside the domain achievable in the current experiment. However, VLSMs are not anymore "eddies" in the classical sense of the word, and I think it would be fair to say that the scales represented in the current UAS experiment do illustrate the eddies on all relevant scales.

Regarding the eddy scales: the 128 m-scale structures have a characteristic length scale of ~10^{-2} Hz, i.e. correspond to the energy containing subrange (Fig. 7). Consequently, the smaller eddies (somewhere under 100 m in size) fall in the inertial subrange. This has been added to the discussion.

L427 "...were contemporaneous with..." Does this mean that 5-min average is notenough long relative to the time scale of the large coherent structures?

- This is a difficult question, but I'm grateful that You have raised it as it leads to some interesting discussion. In general, I have to agree that the 5 min averaging may not be the best approach as it leads to the loss of low-frequency contributions. It is mainly the shortness of the data set used in this study that prompted the division into 5 min periods, and had the data been more extensive, we would have used 30 min averages., Many similar studies use 5 min averages and probably suffer from similar issues. The simplest "back of the envelope" estimate whether the 5 min averaging is valid is by looking at the scalograms in Fig.9, dividing the period into four or two parts (for the short fourth flight). It appears that 5-min periods include several cycles of "intense" and "quiet" turbulence in the flights 3 and 4, a little less in the flight 1, and even less not in the flight 2. Perhaps longer averaging times would reduce the scatter in the relationships such as in Fig. 13 – this remains to be seen in a future study.

An extra comment. This study is motivated to examine the applicability of the TIR imag-ing for the surface heat flux measurement as written in the entire of the manuscript. Itis also obviously written in the last section. Besides, there is no direct comparison between the ground-based sensible heat flux and the TIR images. Therefore, I recommend to add the data of the sensible heat flux together with that of TIR (e.g. showtogether with Fig.9,12,13).

- Thank You for the suggestion. I have added the panels with the kinematic sensible heat flux in Figure 13, which seemed most suitable for these data. It seems that it would be wrong to expectat a simple link between the EC fluxes and the surface temperature, as the relationship is rather scattered, although there is a definite positive slope. 30-min averaging instead of 5-min may help eliminate some of this scatter (as discussed above), while a larger number of flights will further increase the R^2.

References: Drobinski P, Carlotti P, Newsom RK, Banta RM, Foster RC,RedelspergerJ-L (2014) The Structure of the Near-Neutral Atmospheric Surface Layer.J Atmos Sci 61(6), 699–714.

Garai A, Pardyjak E, Steeneveld G-J, Kleissl J (2013)Surface Temperature and Surface-Layer Turbulence in a Convective Boundary Layer.Boundary-Layer Meteorol, 148, 51–72.

Inagaki A, Kanda M, Onomura S, KumemuraH (2013) Thermal Image Velocimetry. Boundary-Layer Meteorol, 149, 1–18.