Journal: Atmospheric Chemistry and Physics

Manuscript number: ACPD-2015-409

Title: Observed spatial variability of boundary-layer turbulence over flat, heterogeneous terrain

Authors: V. Maurer, N. Kalthoff, A. Wieser, M. Kohler and M. Mauder

General comment

This study discusses the variability of the vertical profiles of the air vertical velocity variance observed into the convective boundary layer within a small area of about 3 km horizontal length-scale. It is based on the measurements made during 6 selected fair-weather days by 5 Doppler lidar systems installed at three different sites, which were around 3 km apart. The goal is to determine how much of the observed variability of the vertical velocity variance can be due to the small scale heterogeneity within the 3-km side triangle made by those three sites. This is an interesting issue for which the litterature does need some more answers, and which can be quite well addressed here.

The manuscript is well-written, well presented, the data analysis is based on a nice dataset, and is made rigorously in several aspects (error analysis especially). However, this analysis misses several important hypotheses for the interpretation of the data to actually investigate this issue as much as it could. The starting hypothesis is put into question at the end. This was actually expected, and the analysis could be very interesting and publishable if the starting hypothesis was different and if the analysis was pushed further.

I believe that this study can be worth publishing, but only after major revision of the data analysis.

Main comments

- One of the most important point here that seems to be missed in the analysis is that when the authors are considering 1h-samples of the lidar measurements (for the calculation of the variance), they are considering turbulence structures (or thermals) passing through the lidar as they are advected by the mean wind during this one hour. This corresponds to length scales that are 5 to 10 times larger than 3 km, in the case of the windspeeds observed here (Table 2). For a 5 ms^{-1} mean windspeed for example, it is a horizontal scale of 18 km that will be represented by the turbulent moment calculated from the one-hour sample. This sample length is very much larger than the surface heterogeneity scale that the authors are considering (about 100 m). So if the lidars were located really on the same heterogeneous area (let us say a 50 km by 50 km square of surface heterogeneity of scale 100 m like that shown within the triangle of Fig.1), but on different fields as they are here, we would expect them to show a very similar turbulence profile, except very close to surface.

(That is the reason why, when the lidars are aligned with the wind, the authors do observe very similar time series and statistics, but with a delay of a few hundred of seconds, which corresponds to the time it takes the structure to move from one site to the other with the mean wind.)

Overall, this means that:

1. the basic hypothesis that the measurements of the lidars are independent as long as they are 2 km apart cannot be right
2. the authors should consider a larger area to have an idea of the surfaces and general area that are contributing to the turbulence observed with the lidars
3. the authors should also consider that the larger the wind, the more structures they take into account in their samples, i.e. the more statistics they have into their computed turbulent moments

- Another miss is the consideration of the wind profile, and effect of wind shear. This is not at all discussed, but it can be very important to understand the variability observed from one place to the other, and from one day to the other (wind is 8 to 12 ms^{-1} on some cases, which is quite moderate). The wind will increase statistics, but will also increase the shear production. The authors seem to have the possibility to estimate the wind shear close to surface and at the top of the CBL (from soundings at one site, and maybe from the lidars if VADs were made on the same selected days).

  For point 2 above, and looking at the area at larger scale (see Fig. below), one can see that the triangle made by the three lidars is located in an area with heterogeneities of large scale. Especially, one can see a 42 km long forest to the southwest of the area, a large coal mine to the north of Hambach and another coal mine to the west of Wasserwerk, and also a few villages around. Depending on the wind, those surfaces will significantly contribute to the observed turbulence statistics in the experimental area, and will also potentially induce a change of wind profile (and shear production) from one site to the other. The large presence of areas of small scale crop fields like shown in the considered triangle is also obvious from this larger scale map.

  For point 2 above, the authors could use the area-averaged flux as they did in their current study, but not only over the small triangle made by the three sites: what is the effective area (scale) to be considered in the area-averaged calculation of the normalizing convective scale, in order to minimize the scatter of the day-to-day (and site-to site) variability (of the profiles, or of the maximum normalize variance) ? How do the results change with increasing height ? What is the influence of the wind profile ?

(Note that ideally, an analysis similar to a surface footprint analysis would be very enriching here, but I understand it could correspond to a too large additional analysis. However, even without using a footprint-type analysis, considering various (larger) scales of the area over which the authors are calculating the area-averaged flux, and considering the effect of wind in some way, should help a lot in the understanding and improving the article.)

Specific comments

- *page 18012, Abstract:* The abstracts does not introduce clearly the addressed issue and main aim of the study. It does not mention where is the experiment set up.
- *page 18013, lines 12-17:* It is Taylor's hypothesis which is made here, and should be mentioned. Also stationarity of the sample is assumed. The dataset shown here, with multiple measurements close to each other, gives a very nice opportunity to visit the Taylor hypothesis, and verify when it can actually be made.
- *page 18014, lines 15-17, 'such that the measurements could be assumed to be independent':* The authors needs to clarify what they mean here, and also revise it as they found that they were not independent, or not always.
- *page 18014, lines 21-27, 'aims of this study':* To me, the points enumerated here correspond more to the different steps of the strategy toward the aim. In any case, 'aims' or 'steps', the main goal or main issue should be expressed before those stages.
- *page 18015, section 2.1:* A map of larger scale than that presented in Fig. 1 would be very useful. I needed it to think about the observations and analysis, and I think it is very important to have it in mind (see Figure below).
- *page 18015, lines 18-21: 'energy balance... '* at same (Selhausen) site ?
- *page 18016, line 13-14:* Horizontal wind profiles from lidar VAD do not seem to be discussed and used in the study. Are they ? (for the estimate in Table 2 of the mean wind in the CBL ?)
- *page 18017, section 2.1.2:* I would indicate here (rather than later) the fields in which the stations are installed, and describe their nearby environment.

Figure 1: Google-Earth picture of a larger area around the experimental site (indicated with the yellow triangle)

- *page 18018, section 2.2:* I was curious of watching the fluxes directly too, at least the sensible heat flux or buoyancy flux, which will be used later in the convective velocity calculation.

- *page 18018, line 20:* There is no clear justification of the choice of this area for the area-averaged flux. And as said before, I think the authors have a good opportunity to test the hypothesis made here for the representative flux, by making a sensitivity study to the area (size, and maybe also location) over wich the averaged flux is calculated.

- *page 18018, line 23-28:* I am surprised why the pairs are not {Ruraue, Selhausen} and {Hambach, Wasserwerk}, which is what we deduce from Fig. 2.

  Do the fluxes themselves also behave simimarly among the pairs ? What do you call 'meadow' ? It seems very different from forest to me. Maybe give a few words about it. Not that needle leaf forest can have very large sensible heat flux.

- *page 18019, section 2.3:* I guess the 5 selected days are selected among the 19 IOP days. But it is worth mentioning it (that especially means there were radiosoundings every 2 hours). Was the wind estimated from soundings or from another device (lidar VAD ?) ?

- *page 18020, lines 12-14 / page 18025, lines 3-4:* It is very nice to see the combination of measurements between the two lidars, which enables you to have a cover from 50 m to the CBL top at least.

- *page 18020, lines 15-28:* Relate energy peak to scales. It is missing here in the discussion, even if it is quantitatively addressed later in the text.

- *page 18021, 1-10:* There are several effects which are mixed here, and the discussion is missing some points. At least four points should be considered when analysing those the spectra:
  - The expected variation of the vertical velocity variance with height (smaller at top and bottom of the CBL)
  - The expected variation of the wavelength of maximum vertical velocity spectral energy (as well smaller at top and bottom of the CBL)
  - The effect of beam averaging (very small loss of energy at the smalles scales)
  - The slopes of the inertial subrange which are found to be steeper than the -5/3 law within the CBL. And this is not only due to beam averaging (the latter has a much smaller effect), but rather to coherent structures (See Lothon et al. (2007), Lothon et al. (2009), Darbieu et al. (2014)).

- *page 18021, 10-12, 'as the main aim of our investigation... this effect will be neglected below':* It is also justified by its small contribution relative to the total variance.

- *page 18021, 18-27:* Is this estimate of scales done at 600 m ? at what site ?

- *page 18022, 15-20:* If possible, give an explanation for the small difference observed (size of beam and pulse ?, ...).

- *page 18023 lines 20-21, page 18024 lines 1-5:* Yes, this is consistent with the results of Lothon et al. (2009). They found that sometimes, a layer above the CBL with significant vertical velocity variances can be seen (from gravity waves for example, as said later in the text here). The threshold on the aerosol backscatter was giving more robust results on Z_i estimate. The numerous radiosoundings should really help on validating Z_i estimates robustly here, in a systematic way.

- *page 18025 lines 10-15:* Profiles of skewness should be discussed more in this study. Lenschow et al. (2012) have shown profiles of higher-order moments of the vertical velocity in the CBL, and discussed them qualitatively in sheared and less sheared CBL. They show that the profiles of skewness are quite sensitive to the shear (or wind) and also to the resolution (of an LES) or spatial averaging (of observations), see figures 5 and 9 of Lenschow et al. (2012). It should be quite sensitive to the sample length and statistics (which can be related to mean wind in your study, as said before). The fact that Selhausen in Fig. 6c shows profiles of smaller skewness, and less marked change drop at the CBL, means that there are different conditions at that site, maybe in wind profile or in the 'quality' of the samples (homogeneity, stationarity).

  In general in the manuscript, the effect of wind and shear is not enough taken into account.

- *page 18026 lines 9-10: 'At Wasserwerk, the variance is slightly lower than at Hambach because less convective cells passed the site':* Theoretically, if the sample are representative enough (has enough statistics and homogeneity), the moment should not depend on the number of structures that passed over the site. This might mean that the samples are not long enough. Or that this specific sample is maybe less homogeneous than others. This couls also lead to larger skewness for this sample.

  Organized structures like those discussed later in the text can also lead to such kind of bias and lack or representativeness.

- *page 18027, section 4.3* I am not sure the discussion in 4.3.1 (starting line 12) is needed. The authors could directly address the w_* scaling issue in a whole. It seems to me that Fig. 12 is telling a lot by itself. Fig. 12b directly shows that the local scaling is not appropriate for scaling the maximum variance. The area-averaged scaling is more appropriate. And one question could be: can we minimize the observed scatter (due to day-to-day variability) with an optimized area-representative flux ?

  The authors can also address this question with height dependency, expecting the local scaling to be potentially more and more appropriate as we get closer to the ground. (And the sonics at surface and 30 m can help on this point as well). But this might be seen only below 50 or 40 m, that is only with in situ measurements ...

  And as said before, sensitivity to sampling representativeness could also be done, or sampling representativeness be taken into account in some way (for example by weighting the cases of most representativity).

- *page 18028,lines 12-14, 'it must be concluded that the heterogeneous surface conditions can not explain the statistically significant differences of the w variances.'* This is expected from the sample representativity discussed in main comments. The authors should also consider the surfaces around the area, and the wind, in their discussing the variability of the variance profiles with sites and days. For example, when the wind is souther-westerly, the experimental site seems to be at the lee of a 42 km long forest area, which definitely must impact the turbulence observed (both from the buoyant and the dynamic point of view). Similarly, in north-easterly flows, Hambach is in the closest to a large coal mine, which also can impact a lot the observed statistics.

- *page 18028,lines 19-21, 'It is assumed that the local diurnal cycle of the energy input as well as local differences from day to day can be taken into account better by local scaling than by averaged one.':* Isn't this contradictoray to the above conclusion ? (page 18028, lines 12-14)

- *page 18028, lines 23-23:* Why is 19 May excluded ? I find this case s a good testimony of the analysis, with smaller heterogeneity for this case deduced from the wet ground. It should help in the analysis of the most appropriate scaling, and in the general understanding even (or especially?) if it turns out to be an outlyer sometimes.

- *page 18030 line 25 to page 18031 line 6:* This is an interesting discussion to be associated with the sampling issues, for the understanding of the variability of the variance profiles. But it is not clear what scales are considered here, when talking about 'variance of thermal' and 'variance of environment'. t seems that the scales considered for the moment calculation are much smaller in this conditional analysis than those considered in your analysis.

- *page 18031, section 4.4:* It is very interesting to study the influence of sample length in this study. However, it seems to me that the authors do not explore scales larger than 1 hour. As soon as the authors are averaging variances computed over 1h samples for periods of Δt larger than 1h, the scales that are represented will be still those smaller than 1 h (30 min actually). It is a 1 h filtering. You gain more statistics and reduce random errors, since you have 5 samples in a 5 h sample, but the sampled lengthscales remain the same, and are not larger than 1h (30 min). To me, that is why the curves in Fig. 13c are nicely leveling.

  Fig. 5 though, very interesting as well, shows that the authors can still consider intervals that are larger than 1h: Most of the days shown are quite stationary during the period from 11:00 UTC to 15:00 UTC. That is the authors could consider samples of 2, 3, and 4 hour long over this period of the day.

  Filtering could still be done for all samples at a given cut off frequency, when wishing to keep smaller scales only in the computed variance (and longer samples would then increase the statistics). But I am not sure this was the goal here.

- *page 18033, lines 9-10:* Variance is smaller and skewness is smaller as well, and with a less tilted height dependency than other cases. Hypotheses of a difference in wind profiles or in sample homogeneity should be investigated, along with the role of the long and large forest to the south-west of the experimental area.

- *page 18034, section 4.5, title and discussion:* I would not call this section 'influence of wind'. The influence of the wind is almost not considered in the study. The discussion of section 4.5 is linked with the observations of very coherent measurements between two sites that are aligned with the wind. We do expect this coherency, as well as the delay of 200 s and 400 s respectively for April 20 and April 24, given the mean wind of $\sim$ 10 ms^{-1} and $\sim$ 5 ms^{-1} respectively. Even if it is actually very nice to see it so well, and to be able to quantitatively explore Taylor's hypothesis. But this section is more linked with sampling issues and analysis strategy, than with the influence of the wind profile (and wind shear) itself on the observed vertical velocity statistics.

- *page 18035, lines 10-15 :* This is because in case of the two sites aligned with the wind, they are sampling exactly the same air mass, one site being at the lee of the second. This is not the ase when the wind is different, and especially not the case when the wind is perpendicular to the axis made by the two sites.

- *page 18035-18036 :* I am not sure the LES is very useful here. It is not used at all for the previous questions, and especially not for the issue about surface heterogeneity and representative scaling. Nor for the study of the effect of wind. It is true that rolls occur, and that they can impact very much on our interpretation of the observed turbulence statistics. But I am not sure this limited discussion based on the LES at the really end of the manuscript is appropriate.

  Also note that there is a possible mis-interpretation of the LES fields: when averaging over 1h, the organization seems emphasized possibly artificially because the structures seen in Fig. 16a have been advected at each time step along wind during the 1 h interval. Which can make those 'rolls' appear in Fig. 12b when averaging all of them. So I believe tat the averaging is making the rolls here. The band-like structures in the instantaneous field of Fig. 12a are more reliable.

- *page 18038, lines 12-13, 'as only days with buoyancy-driven turbulence have been chosen' :* This is not quite true because this study by Lenschow et al. (2012) distinguishes the most convective cases with the least convective cases, the latter being those with stronger wind. They show differences of profiles of higher moments (including variance and skewness) between the most convective and the least convective cases, both with lidar observations and LES. Strongest wind in their case is around 8-9 ms^{-1}, which is not as large as one case here with 12 ms^{-1}, but is still moderate. The study by Maurer et al. could actually be very complementary of that previous study (and pushing ine step further), with the different suggestions made before.

- *page 18048, Table 2:* what site(s) is/are considered for those estimates here ? is the integral scale calculated at 600 m or at the height of maximum variance ?
- *page 18049, Fig.1:* I noticed from google-earth that the white patches in Fig. 1 are small villages. This should be specified and not ignored in the analysis. Add a larger scale map of land-use too.
- *page 18050, Fig.2:* Change one of the green colors, because the two greens are very close to each other, this is confusing. I suggest to identify the 6 days, selected for this study. I also suggest to add buoyancy or sensible heat flux, and a time series of Zi would be interesting too.
- *page 18050 Fig.3 and page 18061 Fig. 13:* I suggest to specify the location/site, rather than the lidar model in this figure, because that is what matters here. In Fig. 3a, the layer above 1000 m should be discussed in the text.
- *page 18059, Fig.11:* Note that the variability (standard deviation) of the variance profiles is very similar to that observed by Lenschow et al. (2012).

Formal comments

- *page 18014, lines 5-8* The sentence should be separated in two sentences here, for surface case and aircraft case respectively.
- *page 18015, line 21: '(energy balance data)... were applied as well'* To be reworded.
- *page 18016, line 11: 'because lidars only partly penetrate clouds'* To be reworded.
- *page 18017, line 22-23: '2-hourly intervals'* Do you mean that soundings were launched every 2 hours ?
- *page 18018 line 1-4:* I understand that the ultrasonic and ceilometer were also installed at Hambach site. Maybe this should be more explicetly said.
- *page 18018 line 7, 'using 09:00-15:00 UTC':* 'averaged over the 09:00-15:00 UTC interval'.
- *page 18019, line 6:* Mention that Table 2 gives several characteristic variables for the 6 selected days (not only Z_i and wind).
- *page 18027, section 4.3.1:* I suggest to give the explanation of lines 2-11 in section 3.3.

References

Darbieu, C., F. Lohou, M. Lothon, J. Vilà-Guerau de Arellano, F. Couvreux, D. Durand, D. Pino, E. G. Patton, E. Nilsson, E. Blay-Carreras, and B. Gioli: 2014, 'Turbulence vertical structure of the boundary layer during the afternoon transition'. *Atmos. Chem. Phys. Discuss.* **14**, 32491–32533.

Lenschow, D. H., M. Lothon, S. D. Mayor, P. P. Sullivan, and G. Canut: 2012, 'A comparison of higher-order vertical velocity moments in the convective boundary layer from lidar with in situ measurements and large-eddy simulation'. *Boundary-Layer Meteorol.* **143**, 107–123.

Lothon, M., D. H. Lenschow, and S. Mayor: 2009, 'Doppler lidar measurements of vertical velocity spectra in the convective boundary-layer'. *Boundary-Layer Meteor.* **132**, 205–226.

Lothon, M., D. H. Lenschow, and A. Schanot: 2007, 'Status reminder report on C-130 air-motion measurements. Test of DYCOMS-II new datasets.'. *NCAR/RAF internal report* pp. https:www.eol.ucar.edurafProjectsDYCOMS–IIwind_corrections.html.