Common response for all three reviewers

Although the three referees found the findings and methods of this research interesting, they have a common point in their evaluations concerning the originality of the research. We plan to stress the new insights brought by the paper and here we take the opportunity to give a short summary that will be incorporated in the paper (mainly in the introduction and conclusions).

Our main aim was to propose and apply a new and relatively simple method of assessing the influence of large-scale forcings on a canonical boundary layer influenced by the proximity of complex terrain. Therefore, the gravity of this research lies in the application and discussion of the proposed method. We study systematically the role of subsidence and advection of heat and moisture in governing a prototypical convective boundary layer, normally driven by surface and entrainment turbulent fluxes. By combining observations of boundary-layer height and thermodynamic variables with a 0-order model based on mixed-layer theory, we are able to quantify magnitudes of subsidence and advection consistent with the boundary-layer development. Note, as Figure 1 shows, that in our assumption of subsidence we represent in one single value all the meso- and synoptic downward vertical motions. To our knowledge this is the first time that this quantification is done using such a conceptual model, especially applied on a data set with such a large variety of observational methods of the boundary-layer height. We also think that our results are complementary to estimated subsidence by large-scale mesoscale or synoptic models or by treating upper air measurements, with the additional advantage that our method enables us to disentangle the different contributions to the budgets of heat and moisture. To further verify these estimates and the validity of the canonical boundary layer, we employ a turbulence-resolving code (large-eddy simulation) using the estimated values of subsidence and advection, and we compare the second-order turbulent statistics with surface and aircraft observations.

The main novelty of this method lies in the fact that one does not have to rely on elaborate observational constructions (simultaneous soundings), or three-dimensional large-scale numerical models to estimate values for the subsidence and advection affecting the local boundary layer. These current methods introduce a host of new issues to deal with, whilst our proposed method sticks closely to the observations.

Instead of merely doing an academic case study, we use a comprehensive observational dataset to verify our findings. As explained in the paper, although subsidence was present in majority of the days analyzed by BLLAST, IOP5 is the situation in a set of days with larger influence by subsidence. By selecting this case, we can further test the validity of our method in estimating large-scale downward motions near complex terrain subsidence near complex terrain. As such, our findings will be very useful in an ongoing comparison based on the same day using mesoscale models (e.g., WRF, MESO-NH) (Jiménez et al., 2014).

From the reviewers responses, we concluded that we should be clearer on the motivation for the study and its scope. Subsidence and advection are processes that are difficult to estimate, yet have a large influence on the boundary layer. We will therefore rewrite part of the introduction, giving more attention to our research questions and leaving out distracting sidelines.

In short, more emphasis will be placed on the novelty and practical use of our proposed/applied method for inferring large scale influences on the boundary layer. Already from the introduction, a clearer separation will be made between the inference of subsidence by combining the conceptual model with observations and the verification part, where we compare model results (MXL and LES) with observations. We have also modified Figure 1 to be more precise on the location of the prototypical convective boundary layer and the influence of shear, subsidence and large-scale horizontal advection of heat and moisture. We plan to create a new section after the introduction to introduce the BLLAST experiment and its connected problems and questions.

References

Jiménez M. A., Angevine W., Bazile E., Couvreaux F., Cuxart J. Pino. D and Sastre (2014) An intercomparison of mesoscale simulations during the Boundary Layer Afternoon and Sunset Turbulence (BLLAST) experimental field. 21st Symposium on Boundary Layers and Turbulence, American Meteorological Society, Leeds, United Kingdom.

Reviewer #1

In addition to our specific answers, please see our common response to the three reviewers.

Question: Why did the authors not examine a longer time series of the advective terms? I have no great love of 'golden day science.'

Response: The one-month BLLAST campaign was indeed characterized by diurnal boundary layers with different boundary layer growth rate (see Section 4.1.2 and the classification of the diurnal evolution of the PBL depth by Lothon et al., 2014). However all of them have a non-neglectable component of the large scale downward vertical motions, and we select one case with the largest subsidence as the most extreme days in which the contrast between local and large-scale forcing is larger. We agree with the referee that this is a single day, but by focusing in one day we can go deeper in the analysis and in the interpretation of the comprehensive data set. The analysis of the advective transport with mesoscale models is still under research and will be shown in another paper in which the methodology explained here in length can be further applied.

Q: This is an extremely cautious paper—the authors do not try to demonstrate whether or not their findings are new or useful. They set up a small challenge for their models, meet their low expectations and then come up the excessively general conclusions.

R: We seek to introduce a method that is able to estimate the influence of large-scale motions on a local boundary layer without further knowledge of the boundary conditions. The challenge lies in finding the correct values for subsidence and advection. Using our models, we then check if the found values hold for other variables than those we used to set up our case. The novelty is therefore not in the model results as such, but in the modeling itself. We use a real-life case with a very high observational density to constrain and evaluate our results. As mentioned in our general comment, we will modify the introduction and conclusions to stress the novelty and show where we deviate from explored paths.

Q: The abstract ends with: "We conclude that the prototypical CBL can still be used as a valid representation of the boundary-layer dynamics near regions characterized by complex topography and small-scale surface heterogeneity, provided that surface- and large-scale forcings are well characterized." Why do they select a single day from their field project and proudly label it "representative"? How can a day without clouds be called representative? The authors have to provide some context. Is it never cloudy in Spain? This is relevant, because folks are not going to be surprised that a box model works fine when, well.., when the box model works!

R: As stated in the general BLLAST article (Lothon et al., 2014), this day is representative for the IOP4-6 characterized by large downward vertical motions (ensemble 1, pp 10946, section 4.1.3) in the campaign. To increase the knowledge of this phenomenon, the first step is to stay close to what we know conceptually and to extend to the role of large-scale forcing. We agree with the reviewer that the presence of clouds can disturb the boundary layer growth, but we consider that for cloudless conditions this day is representative.

Q: What is the 'news' here? Why are the many authors surprised by the fact that "...this pattern suggests that not only synoptic scales exert their influence on the boundary layer, but also mesoscale circulations." The authors should be more forceful in pointing out their new and surprising results. If everything to say is in accord with what we already expected, cannot we say that in a sentence. They key here is determining what situations are amenable to this modeling approach now often these occur, and whether or not this effort may sometime lead to improved forecasts of surface conditions, pollutant dispersal or local climate.

R: We agree that the novelty of our research has been understated, whilst focusing on results that meet expectations. Therefore, the abstract, introduction and conclusions shall be rewritten to give more emphasis to the method we used to infer the large-scale motions. In our modeling section, we make a clearer separation between step one: finding the values for subsidence and advection using our boundary layer model, and step two: verifying our results with a LES model and observational data.

<u>Reviewer #2</u>

General comment

As explained in the common response, our results and methods are complementary to a study of the same IOP situation by a mesoscale model, particularly aimed at obtaining an independent estimation of the vertical velocity subsidence evolution. Note that our approach bridges between the local scale observed by the measurements and the larger scale of mesoscale models. In fact, our results will be used for the interpretation and analysis of the intercomparison of mesoscale models (WRF, MESO-NH) currently conducted by other BLLAST researchers (Jiménez et al., 2014). Based on these preliminary results, we have noticed that the mesoscale models results (vertical profiles of potential temperature and specific humidity) disagree with the measurements and therefore they cannot reproduce typical features of the BLLAST boundary layers. Model results show a cold and moist bias during convective conditions. Therefore, there calculations of the advective contributions of heat and moisture should also be taken as a qualitative estimation of the advection, but not as the absolute truth since they depend on the physical parameterizations and the horizontal and vertical resolutions. In conclusion, we think that our findings are an independent contribution to those that will be presented and discussed in the mesoscale modelling intercomparison.

Moreover, other aspects related to turbulence variables and their evolution will not be treated in detail in the mesoscale intercomparison. Our main goal is to study in detail the validity of the prototypical CBL in a region close to the Pyrenees. These results can be useful for further interpretation of data or other model results. In that respect we disagree with the referee that the paper would be speculative as we finish with the concrete conclusion that this boundary layer archetype is valid, provided that the large-scale forcings are well estimated. For this, the combination of different instrumental methods to measure the boundary-layer height is crucial. Our results represent a bottom-up estimation of the active large scale forcings and are as such stand-alone results that can be useful for comparisons of the estimations by mesoscale models or ECMWF (see Figure 7).

Q: In the introduction, there is a lot of emphasis on the CBL-SBL transition that is not so much addressed later in the paper. Keep the introduction relatively short and focused on the specific research questions for this paper and include a short review on previous studies of CBL near complex topography.

R: We will move the CBL-SBL transition to a separate limited and more concise subsection at the end of the introduction. A short review on BL modeling in heterogeneous conditions will be included, although this will not be a large issue for this case, because of the length scales involved. We will also pay some attention to this topic in Section 2.3: surface conditions.

Q: Even if the presented case study fulfills all the criteria for the investigation of a proto- typical boundary layer, it would still be useful to know how the CBL evolved in the other cases. Is the selected case typical or a-typical for the CBL development in the region during BLASST?

R: As mentioned in the general BLLAST paper at section 4.1.3 (Lothon et al., 2014), this case is part of a set of 6 out of the total of 12 IOPs for the campaign. We therefore think that this case is representative for at least part of the campaign, and in particular for the situations characterized by a strong subsidence.

Reviewer #3

General comment

The referee raises two other important points in addition to the ones discussed in our common answer: (a) the role of shear (mechanical turbulence) on the development, maintenance and decaying of the convective boundary layer, in particular at the entrainment zone and (b) the role of surface heterogeneity in forming the convective boundary layer. Both points are very interesting in the analysis but in our opinion they were already covered in the original manuscript. In relation with shear, Figure 3 shows clearly that the wind shear is occurring mainly in the layer between 1800-2000 m. Consequently, the potential enhancement of the boundary layer growth due to the contribution of shear does not occur during IOP5, since the observed and modeled boundary layer reached around 600 m. We have explained this in section 2.2, but we will add a new sentence to explicitly state that shear is not important in our case study.

Related to surface heterogeneity influencing the boundary-layer development by inducing the formation of secondary circulations, we had already included Figure 4 to explicitly show the characteristic length scales around the BLLAST surface site. As the figure shows, the length scales are in average smaller than approximately 500 meter, with an upper limit of 1000 m. There is current debate (Patton et al. 2005; Maronga and Raasch, 2013, van Heerwaarden and Mellado, 2014) on how large the length scale of heterogeneity should be to form secondary circulations influencing the CBL dynamics. The studies show that the lower extreme in the length scale of the heterogeneity is 1-2 z_i, where z_i is the boundary-layer height. Therefore, and based on these studies we have assumed that surface heterogeneity has a minimal influence on the development of CBL. This was further corroborated by the observations of potential temperature collected in horizontal legs by unmanned aircrafts (Reuder, personal communication) that show that above the atmospheric surface layer (roughly 100 m above the surface) the temperature fields were already very homogeneous and well mixed. We will clarify and elaborate more on this point in Section 2.3.

Q: Abstract, 5: "Combining" really isn't an appropriate word here. The two models are only minimally "combined". When I read the abstract before the paper I got the impression that I would read about a new way to force LES with a mixed layer model (I'm not even sure what that would mean but it made me curious). After reading the paper, it is clear this isn't the case. I think the wording of this sentence was what got me.

R: We agree with the referee and we have changed the wording in the abstract. We mean that constrained by the observations, we reproduce the IOP5 by the large-eddy simulations and mixed-layer model. We also impose the same initial and boundary conditions.

Q: Abstract, 20: I would not consider the surface forcing to be "well-characterized". I wouldn't object to using the word "adequately" but well is stretching it considering the gross simplification (and lack of testing of this simplification).

R: The surface forcing relies heavily on the observations. The sinusoidal shape of the fluxes has been chosen to connect with the prototypical boundary layer that we use. We change "well characterized" into: "representative for the local boundary layer".

Q: Introduction, 19249, 15: I really don't like the way that this paragraph is formed. It misses key recent studies (e.g. see work by Nadeau et al.) and tries to diminish the observational studies as too focused on specifics. This paragraph needs to be rewritten to better reflect prior work. I think its current form is biased by the authors preconception of the uniqueness of BLLAST. For example, arguably, the study of Acevedo and Fitzjarrald is much more realistic than what is presented in this paper when it comes to evening transition. So that while I agree that they don't address large scale forcing, they still are important contributions to understanding transitions.

R: As a result of the reviewer's comments, we have rewritten this paragraph. The CBL-SBL transition receives less attention and it is placed at the end of the introduction. The BLLAST campaign will receive less attention, since the main focus of this paper is not on studying the transition, but on the dynamic evolution of the convective boundary layer before the transition. The results can be helpful for future work on transition modeling and should be seen as such. A small section on boundary-layer modeling over heterogeneous terrain will be included in the introduction. More emphasis will be put on the actual novelty of this research: the inference of large scale motions that influence the local boundary layer using a highly conceptual model.

Q: Introduction, 19250, 25: I think this needs to be a more complete review of diurnal cycles and forcing. While I agree that simulations over land with subsidence are rare, several papers have used diurnally varying forcing (large scale advection, geostrophic forcing etc) and looked at its influence. A short (and probably not exhaustive list) includes, Basu et al., JAMC 2008 (Wangara); Kumar et al., JAMC, 2010 (GABLS2 case), Edwards et al., 2014 (GABLS3), Rizza et al., Meteorol. Appl., 2013, Holtslag et al., BAMS, 2013.

R: In the revised introduction, we will add some of the mentioned references to put in a better perspective our study.

Q: Section 2.2: I find it very strange that shear is completely ignored in this study. It is well established that shear plays an important role in CBL dynamics (e.g. Conzemius and Fedorovich, JAS 2006) and especially in entrainment. Ignoring this by saying wind speeds are small seems problematic. Some sensitivity needs to be demonstrated with one of the models. I see no justification for this choice and it may lead to improper conclusions on the impact of large scale forcing (for example level of subsidence based on BL height in the iterative determination). The authors have decided to exclude what is (arguably) the most important large-scale forcing! What makes this even stranger to me is that they do include large scale advection. To me this implies (at least) shear at the boundary layer top which should enhance entrainment and at the surface which would enhance mixing. What makes this more puzzling (the lack of inclusion) is that almost all of the deviations with the measurements are explained away by saying "well we didn't include mean wind".

R: As can be seen in figure 3, and was stated at the beginning of this response, the directional shear zone is located at an altitude higher than the boundary layer depth during the entire day. Wind speeds were relatively low and uniform throughout the boundary layer. As extra check, we have performed a mixed-layer simulation including a 4 m/s wind speed and confirmed that this does not influence the boundary layer evolution strongly.

Q: Section 2.3: Along the same lines as the fact that mean shear (at any level) is ignored, I find it puzzling that surface heterogeneity was ignored in this study. From my understanding (and the description in this paper), this site is locally highly heterogeneous (deviations in near surface fluxes of 100%, see figure 4 and 5). It is well established that surface heterogeneity plays an important role in the dynamics of the CBL across a wide range of scales (e.g., Letzel and Raasch, JAS 2003; Suhring et al., BLM, 2014; Suhring and Raasch, BLM 2013; Maronga and Raasch, 2013; to name a few of many from just one group). In addition, much of this work shows an enhanced impact under weak wind conditions with the blending height concept (ie homogenization by turbulence) all but failing (e.g Letzel and Raasch, 2003). One of the impacts is highly localized boundary layer heights (Maronga and Raasch, 2013). Given this large body of research (I've only mentioned a fraction), what is the justification for ignoring heterogeneity? To me, ignoring this without giving an indication of the impact of ignoring it makes the comparisons with the field experiments highly suspect, especially given the "tuning" used to find the large scale forcing. It also brings into question the comparisons with ground data for evening decay where one might expect highly local behavior that may not average linearly as assumed here. What is interesting is that the authors actually acknowledge this and present evidence that heterogeneity has some impact (fig 9) yet they don't follow-up on it.

R: The length scales involved in this case are smaller than what is considered to influence the boundary layer dynamics. Within the boundary layer itself, under the blending height, small circulations driven by heterogeneities with length scales smaller than 1000 m will still exist, but these do not influence the entire boundary layer. We do acknowledge that this can be of influence on the TKE at the surface, as these small circulations are not included in the LES.

Q: Section 3.2, 19257, 1: Again I find the phase "combining these models" to be somewhat of a misrepresentation of what was done. You could say combining the output (or results) or using the two models.

R: We have changed this phrase to "combining the output of these models".

Q: Section 3.3, 19259, 3-5: When I look at figure 3, I see what appears to be directional shear and (at least at 16:44) what I would call moderate and not insignificant winds in the boundary layer. This would be my conclusion from the figure if I don't read the text description. Why should I believe that these are insignificant? At least some sensitivity (or at least scaling or a reference or something) needs to be given as justification.

R: The shear zone is well above the boundary layer. The wind speeds can be characterized as light to moderate. This has been changed in the text. As we mentioned before, we did not find evidence that the inclusion of wind significantly changed the results. To stay close to the prototypical boundary layer model we have chosen to keep wind speed out.

Q:Section 4.2: In this section (and throughout) the authors should be careful to state that this clearly shows that other processes play a role besides the mean surface forcing, explicitly this is all that is shown. Drawing other conclusions to me is stretching things.

R: We will revise this section, mentioning clearly all the processes that drive the mixed-layer diurnal variability of potential temperature and specific humidity, the large-scale advection of moisture in particular.

Q: Section 5, 19263: The statement that "the differences in the entrainment velocities . . . are due to slight different methods used" is a little misleading. It makes it sound like the difference that is being discussed is the estimation used to develop the figure. When it is really the differences between the models that is reflected. The DALES model doesn't prescribe the entrainment flux as a function of the surface flux while the mixed layer model does. I would say this difference is critical difference and not the different estimation methods. I think it is more appropriate to say that the differences are a result of the explicit representation of entrainment in DALES (I think this is what is meant in the text but I had to read a few times to convince myself of that).

R: We remove "slightly" and added a sentence stating that in DALES, the entrainment process is represented explicitly, though with a contribution of the sub-grid scale till 30-40%.

Q: Section 5, 19264, 6-13: A second explanation for the scatter in the specific moisture measurements at 60 m could be changing wind direction (see fig 3) combined with surface heterogeneity (see figs 4 and 5).

R: We do not find that the wind direction at 60 m changes significantly during the period under investigation. See also Figure 6 in Lothon et al., (2014). It can be indeed certain influence of the "small" scale heterogeneity, but we use this measurement height since this is the one with the largest footprint.

Q: Section 5.2.2, 19269: I don't understand why DALES integrated values across the entire boundary layer are compared to average surface observations? Why not simply compare DALES TKE at the surface with the observations? I know that much of the previous work has used BL integrated TKE but recent experiments have explored the near surface and can be used as a model (e.g., Nadeau et al.). Also please state if the TKE is the total or the resolved TKE (a simple technicality but important enough to be stated).

R: We changed this to the lower 10% of DALES only. We chose to use BL-integrated values because LES results near the surface are more dependent on the sub-grid parameterization used to represent the small scale motions. However, for the sake of a fair comparison, we changed this to the lower levels. TKE is total (including sub-grid), we have included this in the manuscript.

Q: Section 5.2.2: It should also be noted that the exclusion of mean wind likely plays a role in the too fast decay in the models. Basically the models miss the likely (due to orography) reduction of the mean wind and a shift in the mean wind direction in the evening. I would expect this mean BL wind behavior to impact the tke (especially near the surface) by first acting like a source by producing tke through mean shear production and then dying off (as the winds shift) resulting in a rapid decay. This is what I see in the figures. This is brought up at the end but minimized. To me this is further evidence that really this study is almost as "academic" as past studies.

R: The inclusion of wind in modeling the TKE makes a difference in the comparison to the observations. As already stated in the manuscript, mechanical generation of turbulence is a likely cause for the higher levels of TKE in the observations, along with small scale secondary circulations induced by the heterogeneous surface. We have added extra emphasis to this in section 5.2.2. The decrease of the mean wind in the early evening is a good explanation for the fast decay we see in the observations, we have included this in the article. The main topic of this paragraph, is the difference between the two cases as presented before. This leads to our recommendation of having a good representation of the boundary layer before doing TKE modeling. This approach might be called academic, albeit with an applied intent. For a more detailed analysis of the TKE decay during the BLLAST-campaign, we refer to Darbieu et al. 2014.

Q: Section 5.2.2 and Figure 14: Note that all the symbols used in figure 14 don't match the text.

R: We changed the text to match with the figure.

Q: Section 5.2.2, 19270: In previous studies, the TKE decay rate is quantified. Why not do that here so that the differences can be quantified between the different cases. This would also facilitate a wider comparison with other studies (although introduce the problem of decay of the integrated or single level value).

R: We have chosen not to include this, as our modeling study has not been exhaustive in finding all terms related to the TKE decay. Therefore, if we give a decay rate, this might be mistaken as a full scale modeling exercise including all relevant processes. Our main interest lies in the inclusion of large scale motions and their effect on the boundary layer dynamics and subsequent influence during the transition period. However, since the referee raises a relevant point, we will include a reference to the research of Darbieu et al. (2015) (to be submitted to this special issue) since they treat in detail the evolution of TKE on time and on height.