We appreciate the constructive comments and suggestions by Reviewer 2. Point-by-point responses to the Reviewer comments are provided below (with Reviewer comments in *italics*).

In the, well written, no comments on the English that is used, paper the authors only look at the resolved vertical velocity of the AROME model, which has a resolution of 2.5 km, and therefore does not employ a deep convection parameterization. However, a shallow convection parameterization, one that should represent non precipitating cumulus and stratocumulus, is active in this model. It is an EDMF (Eddy Diffusion Mass Flux) scheme that represents the largest thermals in the boundary layer through a Mass flux scheme (the Kain Fritsch shallow cumulus parameterization) and the small scale eddies through a TKE scheme (Cuxart et al, 2000).
The authors are looking at all clouds together and do not distinguish between cumulus clouds (very small scale) and stratocumulus/stratus clouds (all represented by the parameterization) or large scale clouds. Especially the parameterized clouds have their own dynamics embedded in the parameterizations, with their own vertical velocity calculations for determination of the cloud dimensions. Somehow this should be incorporated into the diagnostics, as the conclusion that AROME is not representing the vertical velocities in the correct way may not be true when the subgrid scale contribution present in the parameterization is added to the resolved part.
Note that this cannot be done by simply adding a TKE term due to the fact that AROME works with a EDMF-type of scheme. This causes the TKE to be very small at the top of the parameterized boundary layer, where the turbulence is dominated by the large eddies represented by the mass flux part of the scheme. This of course depends on the weather situation, with the situation above being representative for the surface forced boundary layer during the day in Spring and Summer.

This comment has raised the need to clarify the motivation behind this paper. An extra paragraph has been included in the slightly re-written Introduction as well as in section 4.1, together with additional sentences in the Conclusions (section 6). These state that we are investigating the ability of AROME to resolve the vertical velocity fields at cloud base, for the purpose of taking data from the operational AROME forecasts to derive vertical velocity parameterizations for larger-scale models. This is why the parameterized component of vertical velocity from the Eddy Diffusivity Mass Flux scheme (EDMF, or EDKF in the case of AROME due to the Kain-Fritsch shallow cumulus formulation) was not considered in this study.

The parameterized subgrid updraft velocity embedded inside the EDKF-scheme is calculated essentially by vertical integration of the buoyancy force in convective boundary layer, also taking into account entrainment and detrainment processes. We do acknowledge the importance of the EDKF-scheme in representing the non-local mixing by the largest eddies and the implications of this process to boundary layer and cloud development. The shallow convective updraft parameterization could also be used for aerosol-cloud interactions in cumulus clouds, if an aerosol activation scheme were implemented in AROME (however, for stratiform cloud layers, other solutions would be needed). This is pointed out in the revised manuscript (section 6). However, for developing vertical velocity parameterizations for larger-scale models we are interested in the resolved circulations.

At the request by Reviewer 1, we now show in the revised manuscript (Figs. 5c and 5f, section 5.1.1) the impact of adding a TKE term to the resolved vertical velocities from the model. We found that, when the TKE term is added to the grid-scale vertical velocity in AROME, the modelled σ_w is similar to the observed σ_w, but only below an altitude of about 1-2 km (within the model boundary layer); at higher altitudes, the TKE-term is relatively small, as anticipated by Reviewer 2.

Our response to the Reviewer's comment on cloud classification is the same as the response we gave to Reviewer 1. We have included an additional figure (Figure 6) and an additional subsection 5.1.2 regarding this analysis. Figure 6 shows the statistics of vertical velocity for three subclasses, separated by cloud geometrical thickness (0-250 m, 250-500 m and > 500 m). Here we see a tendency for mean values (Fig. 6a) of the cloud base vertical velocity simulated by AROME to increase with cloud geometrical thickness while the observed mean vertical velocity becomes more negative for the deepest clouds (further discussion on this result follows below). For σ_w (Fig. 6b), AROME also shows an increasing tendency with increasing cloud depth, while in the observations a considerable increase in σ_w is only seen for the deepest clouds.

One other problem lies in the interpretation of the observations and the assumption that clouds are only present with positive (upward) vertical velocities. It has been demonstrated (Heus and Jonker, 2008) that the updraft of cumulus clouds is surrounded by a subsiding shell still containing cloud water, where the interaction with the environment takes place. This means that the velocities can be quite negative for cumulus clouds and it all depends on how the cloud moves over the observation site which part of the cloud is visible for the doppler radar.

Our method of sampling does not make any assumptions about clouds only existing in regions with updrafts. We agree with the Reviewer that clouds can exist in both up and downdrafts, and that the Doppler radar may only sample a descending portion of a single cloud. Given that the slice of the cloud that the Doppler radar observes for each cloud passing overhead is random, and that we have a large dataset, we would expect this to provide an unbiased sample of cumulus vertical velocities. However, the screening procedure then has the potential to introduce an artificial bias, if the sampling statistics are skewed towards particular profiles (e.g. descending edges of cumulus clouds rather than ascending precipitating cores). For shallow clouds the sampling statistics appear to be essentially unbiased as shown in Fig. 2a where no profiles were discarded. Our new Figure 6a shows that the two shallow cloud classes (both < 500 m deep) have a negative mean vertical velocity (-0.2 m s^{-1}), which is similar to that seen in the overall sampling (Fig. 5). Therefore we feel this negative value does not arise from sampling statistics, but from retrieval artefacts. However, for the deep clouds (> 500 m, Fig. 6a), we may be conditionally sampling only the cloud edges by discarding the precipitating profiles (presumably the profiles with an updraft in cumulus-type clouds). We think this may be responsible for the increase in the mean vertical velocity (from -0.2 to -0.5 m s^{-1} at worst) for the clouds thicker than 500 m.

Further, the very nice case of January 2008 was a situation under the influence of a high pressure area with large scale subsidence, probably causing the average vertical velocity to be negative.

High-pressure areas are accompanied by large-scale subsidence. We would expect the model to be reasonably accurate in predicting the large-scale vertical motion (at least in a domain-averaged sense). However, the typical magnitude of the large-scale vertical motion in such situations is on the order of a few cm s^{-1}, which is rather small compared to the observed mean vertical velocity of about 20-30 cm s^{-1}. We still suspect that the main reason for this discrepancy is the artefacts in the Doppler radar retrievals, as no such bias is seen in the model (which is consistently close to zero).

SPECIFIC COMMENTS

Why is there a difference in the size of the model domain for the SGP and Lindenberg experiments and why do you only use the 3-hourly output instead of much higher resolution (in time) output? Now you have the possibility that the fixed timing of the output causes some waves travelling through the domain artificially influencing the model results, which can be less when e.g. hourly output is used.

A smaller domain size was selected for Lindenberg simply to reduce the computational costs. The results shown for Lindenberg are extracted from AROME runs for a longer time period, used for other studies besides this work.

The data were stored at 3-hour (rather than 1-hour) resolution to reduce the data volume (which is even now much larger for the model than for the observations). We tested the impact of temporal sampling by computing statistics for data sampled at even lower temporal resolution (e.g., every 6 hours, separately for 00, 06, 12 and 18 UTC, and 03, 09, 15 and 21 UTC). The statistics of vertical velocity differed very little between these subsets. Based on these analyses, we are confident that use of higher temporal resolution for model output would not change our results substantially.

Also how do you distinguish between large scale vertical motion and the motion being caused by the buoyancy underneath the clouds?

In this study, we do not make this distinction. Again, we point out that at the background of this study lies the question of how to parameterize the aerosol activation. In current state-of-the-art aerosol activation schemes, cloud base vertical velocity is taken as a proxy of the supersaturation production by adiabatic cooling (Fountoukis and Nenes, 2005; Abdul-Razzak & Ghan, 2000). Therefore, in the current manuscript, our interest lies in the scale differences between the observations and the model, and not so much on the physical factors causing vertical motion. We acknowledge that distinguishing these physical factors might be of more interest for a study more focused on cloud dynamics.

Also the temporal resolution of the boundaries is quite coarse.

About the temporal resolution of the boundaries we, in principle, agree with the Reviewer. For instance, Termonia et al. (2009) suggest that 3h coupling frequency would be sufficient for NWP purposes in most of the cases. They estimate that with 3h coupling the maximum potential mean sea level pressure error (48h simulation), due to boundary interpolation, is between 1.4 - 11.5 hPa (Termonia et al. , 2009, Fig. 3). However, the low end of the potential error estimate does not differ much from the values with 6h coupling: 2 - 16.8 hPa. Moreover, the high end of the errors scale represents the "Storm-of-the-century" type of very rare events. Therefore, we believe that 6h coupling frequency with short forecasts (+12h) does not a have large detrimental effect on our results compared with 3h coupling frequency.

Normal vertical velocities in high pressure situation usually on the order of a few cm/s, so negative vertical velocities of 0.4 m/s are very strong! This may be caused by (standing) waves or wave-like phenomena, as there is a significant height difference of more than 100 metres in 100 km. A strong wind from the west would then mean strong subsiding motions, even close to the surface. This is another reason why I would like to see more differentiation into different cases.

On the suggestion of the Reviewer we repeated the analysis with an additional classification where we eliminated cases with strong (> 15 m s^{-1}) horizontal winds from the west sampled at a height of about 2 km. For the case referred by the Reviewer (-0.4 m s^{-1} mean velocity, SGP in January, please see the attached Figures 1 and 2 with observed mean vertical velocities below), eliminating strong westerly winds does not produce any decisive improvement in the mean of the observed vertical velocity (note that the number of points decreases significantly when the westerly winds are filtered out, which is also the reason why a velocity threshold as high as 15 m s^{-1} was used). If a consistent subsidence caused by orography would contribute the observed mean velocity, we would again expect to see a rather similar effect in the model data.

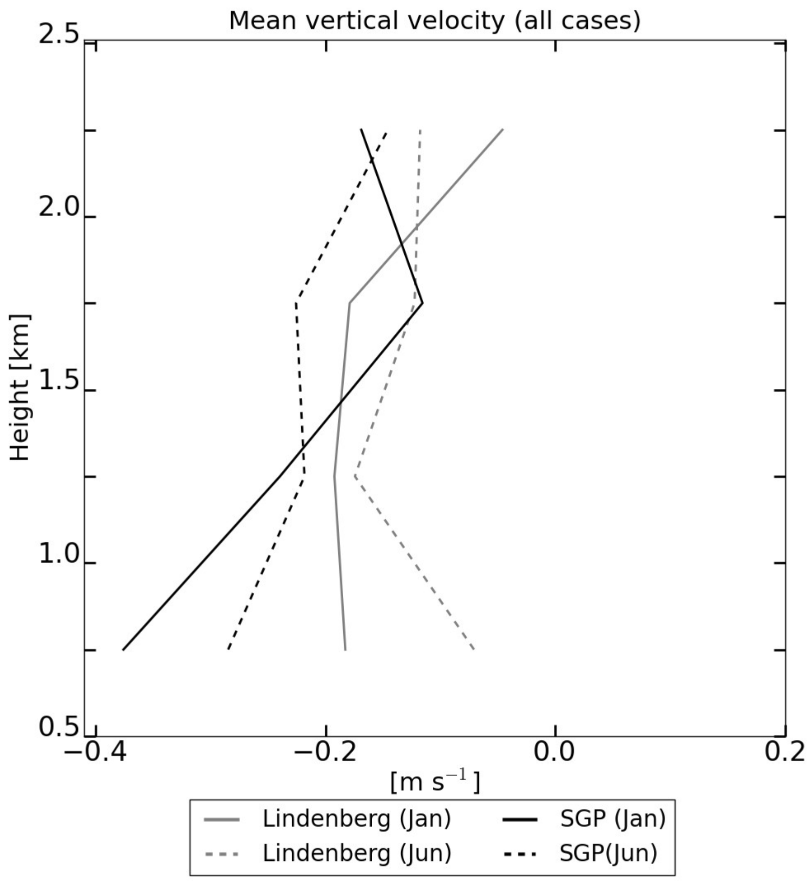

Figure 1: Observed mean vertical velocities for Lindenberg (grey) and SGP (black) in January (solid) and June (dashed) for all data (similar to the profiles shown in Figures 5a and 5d).

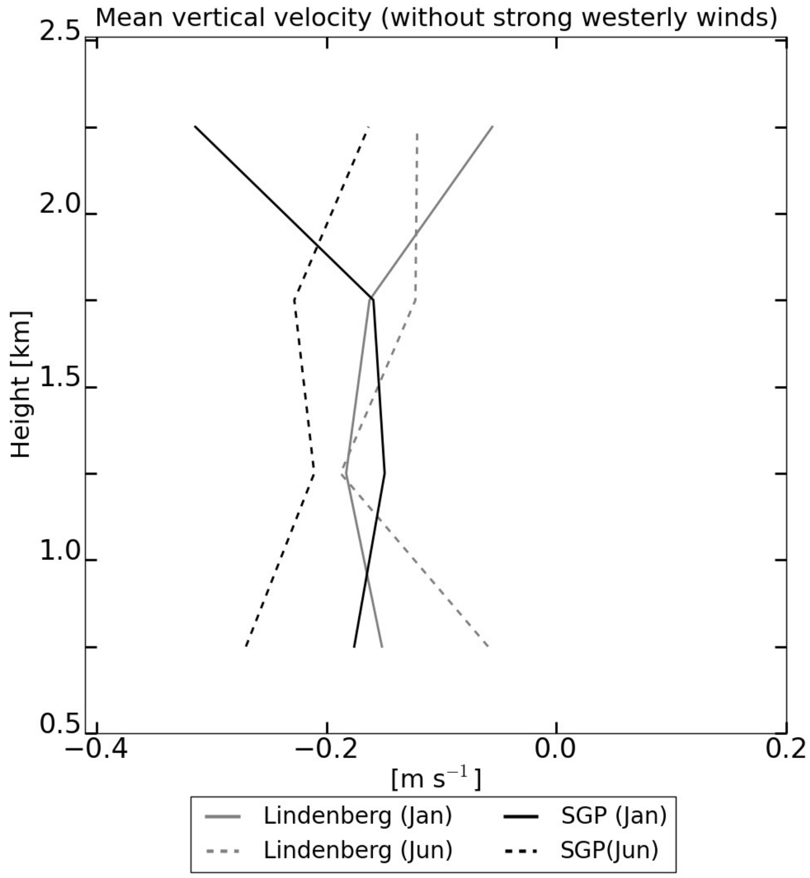

Figure 2: Observed mean vertical velocities for Lindenberg (grey) and SGP (black) in January (solid) and June (dashed), where cases with strong (> 15 m s^{-1}) horizontal wind from the west have been filtered out.

References

Abdul-Razzak, H. and Ghan, S. J.: Parameterization of aerosol activations 2. Multiple aerosol types. J Geophys. Res., 105(D5), doi:10.1029/1999JD901161, 2000

Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models . J. Geophys. Res., 110, D11212, doi:10.1029/2004JD005591, 2005

Termonia, P., A. Deckmyn and R.Hamdi: Study of the lateral boundary condition temporal resolution problem and a proposed solution by means of boundary error restart. Mon. Wea. Rev, 137, 3551 – 3566, 2009