Responses to the reviewers.

We thank the reviewers for their time and effort. We tried to revise the paper following the spirit and the letter of their reviews. Overall, the reviewers point out that most of the findings reported in the paper are not new, and we revised the manuscript to expose that. Specifically, we refer to most of the papers suggested by the reviewers and put our results in the context of previous modeling results. We do believe, however, that our results are worth publishing as they expose problems with the application of the off-the-shelf community mesoscale model to the problem at hand, and report several sensitivity studies (mostly unsuccessful, unfortunately) performed to improve the simulation. Such a study should have its place in the ACP's VOCALS Special Issue collection.

Below we respond to the specific points, presenting first the original reviewers' comments in italics and following with our responses.

Reviewer 1 comments:

1) It is not clear why the simulations were conducted only for a short period of time
(42-hour) and with a particular selection of 12-13 November as the simulation period considering that the VOCALS-REx is a month-long campaign with profile measurements available from both dropsondes (BAe-146) and radiosondes (Ron Brown Research Vessel). It is likely that WRF simulations have a larger bias in the simulated MBL heights at a particular day.

The revised manuscript includes additional simulations, extending those in the original submission. The added simulations feature a larger domain and extended simulation period by selecting earlier starting dates, see Table 1 for details. There are previous publications discussing performance of various numerical models for the whole VOCALS campaign and it was not our intention to repeat those. We looked at the model performance from the case study point of view applying several subgrid-scale parameterizations and focusing on a specific day. The fact that the aircraft encountered POCs at the western edge of its track motivated our selection, as now explained in text.

2) Use of the initial and boundary conditions from GFS was concluded to be one of the two main reasons that led to the low bias in simulated MBL heights in WRF. I think this is a simple technical problem. The simulations should be rerun with a longer spinup time and with a *bigger outer domain (especially, extend the domain sufficiently along the direction of inflow boundaries). Note that it is not computationally expensive to extend the outer domain.*

We ran additional simulations with either a larger domain and longer simulation time, or with only a longer simulation time. Results from these are now included in the discussion. Extending the domain and the simulation time improves solution in some profile locations (shown in Fig. 1), especially far from the shore (see Fig.6a). The results do not give enough evidence, however, that the solution is significantly improved over the entire domain. Thus we believe the problem remains. We stress this aspect in the discussion.

3) It is not clear what you mean by 'the scheme confusing the cloud base change of
the vertical temperature and moisture gradients with the change at the boundary-layer inversion'. In the case of using YSU scheme, clouds tend to form above the diagnosed PBL heights, which has been noted by other researchers. The MBL height should be diagnosed offline from the simulated thermodynamically profiles. When using the YSU PBL scheme in WRF, the direct model outputs of PBL heights are not the traditional defined PBL heights; they are the heights where the Richardson number is 0 in a model vertical column.

We meant that the cloud-base change of the vertical temperature gradient from the dry-adiabatic within the sub-cloud layer to the moist-adiabatic within the cloud was interpreted by the scheme as the boundary-layer (BL) inversion. This height was outputted by the model and used in our comparison. We do not know why the reviewer objects to such a methodology. To our understanding, BL scheme should force BL to be close to being well-mixed (in the sense of conserved moist variables) up to the BL height. If BL height is selected near the cloud base rather than the cloud top, this is clearly a problem. To weaken the statement in question, we rephrase this sentence to read: "… the boundary layer depth used in the boundary-layer scheme is typically close to the height of the cloud base, rather than cloud top." Since the surface fluxes are distributed within BL, this is likely the reason why BL does not grow to the observed height even with observed SSTs.

4) It is not clear whether the simulation results (low PBL biases) are consistent with those in the current literature using WRF/WRF-Chem. Very few existing modeling or observational studies in the current literature (such as those from VOCALS-REx studies) were compared with the results from this study.

We added relevant references to several papers in the revised manuscript. It should be pointed out, however, that those studies typically considered time-and space-averaged BL characteristics.

5) The 'arguable POC' mentioned in this study looks more like cloud-free regions. WRF does not consider aerosol gradients, and assumptions were made in the microphysics schemes, such as a constant droplet/aerosol number concentration. Even if a 3 km resolution (9 km results were used actually for this analysis) is used, it would be surprising to see the co-existence of POC embedded in generally cloudy regions. It is not clear what the purpose is to study the cloud-free regions in the model simulations given that there seems to be large discrepancies in the simulated and observed cloud fields. Those clear spots were called 'POC' for a few times (including in Figs. 11, 14, and 15), which I think should be consistently called 'cloud-free regions'.

We replaced references to POC with "cloud-free regions" in the discussion of model results. We do not think aerosol gradients are the only possible explanation of POC formation. For instance, the observational study by Allen et al. (2011) suggests that a POC can form through the interactions between propagating inertia-gravity waves and boundary-layer cloudiness. We agree with the reviewer that the resolution is too coarse and microphysics too simple in the current study, and that high-resolution LES studies (e.g., of the type reported in Wang et al. ACP 2010) should be considered in the future.

Minor comments:

a) Page 25522 Lines 6-9: The PBL heights observed during the VOCALS-REx are
about 1000-2000 m, so the 343-m resolution near the MBL top (in the 34-layer simulation) is apparently too coarse. It is surprising that with 81 and 121 vertical levels, the vertical grid spacings are only 120 and 81 m near the cloud top, respectively. It would make more sense to only increase model vertical resolutions for the lowest 2-3 km.

We describe the way vertical levels are set in the second paragraph of the section 2. We agree that perhaps a different distribution of vertical levels in high-vertical-resolution simulations would be desirable. However, this would require a new set of model simulations, an effort not possible at this time. Despite the fact that vertical resolution for 36 levels run is coarse, we do not see significant differences compared to 81/121 levels runs as documented in Fig. 2 of the revised manuscript. This indicates that the problem with the BL height prediction is only partially related to the vertical resolution, and the combination of the BL scheme and inflow boundary conditions is the main culprit.

b) Pages 25522-25523: I recommend the authors provide references for the descriptions of different schemes used in WRF.

Key references describing parameterizations used in the paper are now provided.

c) Page 25541, Fig. 2: In panel e, those horizontal lines (black) below 0.5 km in height are confusing.

This figure was replaced with the new one, without horizontal lines.

*e) In Figs. 3-6, how the measurement variability is defined? The variability in the figure is difficult to interpret with observed values. I recommend the variability to be plotted as observed values 2*sigma/variability).*

As described in the revised manuscript, the variability is defined using measurements sampled up to 30 minutes before a given profile (either a partial profile or a horizontal leg as shown in fig. 1.b using a gray color). This is explained just before the 'Results' section. What the reviewer suggests is possible only for horizontal legs where several data points at a given height are available.

f) A lot of statements are descriptive rather than quantitative, such as in Page 25525
Lines 25-26, 'too moist', 'colder', etc.

The statements are qualitative because of the variability of the differences in height. The quantitative information is provided in graphical form in the figures. We tried to add quantitative information in discussion of some of the figures.

g) In some panels of Fig. 7, a very different vertical range (y-axis) is used, is there a
reason for that?

We removed this figure from the revised manuscript

h) In Fig. 9, I am not sure why erroneous satellite data were used in this Figure. I
would suggest filter the data with a solar zenith angle threshold. As you mentioned,
'the large scale pattern does not move significantly in space : : :', maybe in this figure

only presenting a temporal-averaged field would be sufficient rather than 4 panels with every 15-minute outputs.

This is now Fig. 8. We removed the panel with observations affected by the low solar zenith angle. Temporal averaging would prevent us from showing time evolution, which we later refer to when showing the model results in Fig. 10.

j) I would recommend remove Fig. 10 since it was only used to show that the large scale pattern does not change in space, which apparently is already reflected in Fig. 9.

We removed this figure.

k) Fig. 11. The y-axis is very confusing. Suggest using two y-axes with one for the
LWP/PWP and the other for the height.

We left the plot unchanged after trying some different options. The key information in the panels is the co-evolution of several model variables. We understand that reading specific values of the variables may be difficult, but we do not see any better way to compactly present the data.

l) 'LWP' is typically defined as vertical integration of rainwater and cloud water, however, in this study only cloud water was used in integration. This should be pointed out in the manuscript.

In the revised manuscript, we specified in the first paragraph of section 4.2 that LWP is the vertical integral of the cloud water content alone.

m) Page 25531, Lines 14-16: It seems to too speculative to say 'drizzle evaporation
initializes subsidence'.

We see nothing wrong with such a statement: drizzle evaporation is a source of negative buoyancy and in principle can locally drive descending motion.

n) In Fig. 12, the purple lines as described in the caption seem to be missing and the
contour lines do not have labels. What is the purpose of plotting lines with theta=299.5 K? The wind speed legend is missing too.

We apologize for the error. The purple line was shown in a previous version of the figure and the caption was not modified. The wind speed scale is now explained in the caption.

o) Page 25533, Lines 4-9: The connection between the horizontal temperature and moisture gradients and entrainment events is a stretch. The larger variations near the inversion layer are most likely related to the variations of MBL heights.

We agree with this comment. The figure and accompanying discussion were removed from the manuscript.

p) Page 25533, Lines 11-17: In analyzing the evolution, advection was not mentioned. The horizontal wind was blowing mostly from southeast (the wind speeds legend is missing in Fig. 12). Assuming 10 m s -1 wind speed, the cloudy or cloud-free air will be advected roughly a horizontal distance of 36 km in an hour (9 km in 15 min). This might have an impact on the evolution analysis depending on the actual wind speed.

Advection does play some role as illustrated by Fig. 10, where the cloud-free region expands, but it moves from NE to SW. The revised figure also shows that the wind speed is ~6 ms^{-1}, so advection is less important than the reviewer suggests. In addition to advection, clouds also evolve in time. Unfortunately, we did not find a better way to present diagnostics other than by columns with locations fixed in space. We do not think this has significant impact on mostly qualitative results concerning the cloud-free regions.

q) Page 25535, Lines 18-27: It is a stretch to connect your results with that of Allen et al. (2011). There are a lot more relevant papers during the VOCALS. Are you indicating that the simulation results reflect the transition from closed to open cells?

We state at the end of the first paragraph of the conclusion section that our results might have some relevance to POC formation. We now refer to more previous modeling studies to put our results in the context of past investigations.

l) Page 25536, Lines 4-6: I do not think you need LES models to capture gravity waves/lower troposphere waves.

The key point is that one needs to capture the coupling between gravity waves and resolved boundary-layer circulations. One needs LES model for that.

Reviewer 2 comments:

1. The finding that WRF underestimate the MBL depth (by a factor of 2) over the SE Pacific is not new at all, but little referenced in this manuscript. At the end of this review I provided a list of works –many of the in ACP VOCALS special issue- noting this problem, not only in WRF but in many other regional models. The authors spend significant text and 6 figures showing this problem (probably figs. 3 and 6 will make the job): : :.reducing the length of this section (4.1) is needed. On the other hand, the authors doesn't comment that the model bias tend to reduce offshore.

Previous studies suggested by the reviewer are now appropriately referenced in the text and our results are put in the context of those studies. We did eliminate two of the figures referenced by the reviewer, but then we added additional figure to present additional simulations, as suggested by other reviewers. The discussion in section 4.1 was modified to reflect that.

2. The authors then try different model configurations (changes in vertical levels, grid
spacing and many different physical parametrizations) without getting any significant
improvement. Next, the authors present a conjecture: a problem with the initialization
based on GFS. I agree that GFS represents the MBL rather poorly and this is a good
point but I'd like to see something more concrete. Is here where I strongly suggest further modelling work: use a longer simulation, with an initialization several days before of your target time so the WRF can depart from the GFS initial condition. Alternative, the authors could modify the initial conditions so as to better represent the lower troposphere: : :I understand that this could be difficult but definitely worth trying.

We ran additional simulations with increased simulation times, increased domain size and different start times. A brief discussion of these results is now included in the manuscript. In the revised manuscript, we removed profile 6, but kept the rest of the 5 profiles. For each profile, the results from different sensitivity simulations are now presented.

We did not try to modify initial conditions. This seems not as trivial as the reviewer suggests because such modifications should be consistent with the model design (for instance, maintaining model-assumed momentum and thermodynamic conservation laws). This is left for future research.

3. Section 4.2 is devoted to the simulated formation of mesoscale cloud-free regions.
I was a bit reluctant to read this part after all the negative issues on the model performance noted by the same authors (section 4.1 and first paragraph of 4.2). In any case, the model does produce cloud-free regions and the authors present some interesting diagnostic of this. The main problem in this section is the quality of the figures. The authors can do a lot improving them to guide the reader in their reasoning. For instance, given the key role of the vertical velocity, they could shade the periods with upward/downward motion in Fig. 11. In line 13 of page 25531 they said "the disappearance seems to result from strong (up to 4 cm s-1, not shown) subsidence in the model column": : :what do you mean with not shown: : :isn't that presented in Fig. 11? Likewise, the aspect ratio of Fig. 9 and the colored vectors in Fig. 12 make them difficult to interpret.

The figures were modified in response to this comment. Fig. 11 (number Fig. 9 in the revised manuscript) shows the evolution of the vertical velocity averaged across the cloud layer. Thus its maximum value is not shown in the figure. This is why we explicitly state the maximum value of the subsidence in the text. The aspect ratio of Fig. 9 (now Fig. 8) and colors in Fig. 12 (now Fig 10) have been modified.

Reviewer 3 comments:

This manuscript present results from a short simulation (42hr) of the WRF regional model over the SE Pacific. The first part of this study examines the sensitivity of the vertical profiles (but not the radiatively important cloud properties such as cloud cover and LWP) to the PBL mixing scheme used in the model. The second part examines two mesoscale clearings in the model and concludes that they are produced by patches of strong large scale subsidence.

We do show comparison between observed (by the BAE-146 aircraft) and simulated cloud properties, such as profiles of the cloud water mixing ratio. In the revised manuscript, figure 3 does show how these profiles depend on BL parameterizations. In the second part, figure 9 (8 in the revised manuscript) does show LWP for the entire computational domain in the reference run and compares it to LWP derived from GOES satellite.

Overall, I don't think this is worthy of publication because the simulations offer very little that is truly new. The analysis is rather superficial. The overly shallow stratocumulus-topped PBL problem has been known for a long time and there are solutions to the problem stemming back over a decade (e.g. UK Met Office PBL scheme of Lock et al. 2000, plus Bretherton+McCaa+Park, Mechoso and others have all worked on improvements), and there are versions of WRF that include moist physics PBL schemes (perhaps these are not yet publicly available). The second part of the study demonstrates that strong subsidence can cause cloud clearing, but again this has been known and explored for two decades (e.g. Randall and Suarez 1984, JAS). What would be new is to understand how the mesoscale patches of strong subsidence originate from (e.g. gravity wave breaking, land-atmosphere interactions), but this wasn't explored. I provide some more comments on each of the parts below.

We do not agree with the initial sentence in this comment. We do believe that our results are worth publishing as they expose problems with the application of the off-the-shelf community mesoscale model to the problem at hand, and report several sensitivity studies (mostly unsuccessful, unfortunately) performed to improve the simulation. We do refer to previous studies in the revised manuscript. We are aware of the fact that there are versions of the WRF model that include BL parameterization scheme designed to deal with the problem identified in previous studies. Such schemes are unfortunately not available in the off-the-shelf version of WRF that we used. We also would like to point out that according to the results reported in Abel et al, the Lock's scheme still produces some biases of the BL height in the UM model.

The relevance of the Randall-Suarez study to our results is highly questionable. Most importantly, they considered equilibrium states of the highly idealized mixed layer model with moist physics, whereas our solutions represent dynamically evolving transient states.

We realize that understanding the origin of subsidence regions (clearly involved with the development of cloud clearings, see Fig. 10) would be an interesting and a worthwhile effort. However, this was not our goal at the onset of this project. The goal, as explained in the introduction, was to compare case-study-type simulations (using an off-the-shelf mesoscale model) to VOCALS observations and to investigate sensitivity of model solutions to subgrid-scale parameterizations.

We agree that the origin of the mesoscale subsidence patches is an interesting question. However, as our focus is on clouds observed in the field, we think that equally interesting question is the interaction between mesoscale subsidence patches with resolved (rather than parameterized as in the current study) convective BL circulations. We plan to perform such LES simulations in the future.

Part 1: The sensitivity to the PBL schemes available in the WRF is found to make a major difference to the simulations, but all three PBL schemes tested produce a PBL that is too shallow and poorly mixed. This is a well-known problem with large scale models in regions of marine stratocumulus (e.g. Bretherton et al. 2004, BAMS), and has even been examined across a whole suite of regional and global models in this very region (Wyant et al. 2010, ACP), which the authors seem to be unaware of. It is conjectured that some of the PBL height underestimation is associated with initial conditions (which are derived from a model that itself has a rather poor quality PBL scheme), while some clearly comes from the poor representation of mixing in the PBL schemes (none of which were designed with much consideration of important processes in the marine PBL). This conjecture could be, but wasn't, tested by comparing against a simulation initialized with a better analysis (e.g. ERA Interim), and so I was left wondering what new knowledge has been created from the first part of the study.

With the model horizontal grid length of 9 km (3 km in inner domain), we ran WRF with significantly higher horizontal resolution that models discussed by Wyant et al. 2010 (we refer to this paper in the revised manuscript). This aspect seems to have an insignificant effect on model solutions. To our knowledge, none of the previous papers addressed the problems with larger-scale models (providing boundary conditions) as one of the possible explanations of the differences in model solutions and observations. Following the comment, we did try ERA Interim as a source of initial and boundary conditions, but the solutions did not improve. This was mostly because of similar problems with the lower-tropospheric profiles in the ERA interim data. This aspect is now mentioned in the paper. Overall, we believe that the problem is generic and most likely results from the lack of observations over the oceans constraining the large-scale analyses model.

Part 2: The model clearings are clearly related to mesoscale subsidence patches and not to processes that cause POCs in the real atmosphere. So there seems to me to be little point in discussing POCs at all since the model features are completely unrelated. I wanted to know whether the subsidence patches are created internally in the model (e.g. from land-atmosphere interactions) or whether they are memory of the initial conditions, but this was not explored. The most *interesting figure in the manuscript in my view (Fig. 12) shows intriguing NW/SE striations in the vertical velocity field, but the authors seem confused about what these are (gravity waves or not?). Looking more closely at their origin would be interesting.*

To our knowledge, the problem of POC formation in the real atmosphere remains unexplained because both dynamical processes (e.g. gravity waves, Allen et al., in preparation) and microphysical processes (drizzle, aerosol removal, etc) were argued to be essential in the past. As suggested by one of the reviewers and to avoid direct references to POCs, we now refer to the structures simulated by the model as cloud-free regions.

We believe subsidence patches were created by the model and they do not come from the initial conditions. This is because spatial resolution of the GFS analysis is only 1 degree and there are very few (if any) mesoscale features in it.

We do not know what the reason is for the mesoscale vertical velocity pattern shown in Fig. 12 (Fig. 10 in the revised manuscript) which is produced by the model; neither do we know what their origins. This aspect has to remain as one of the topics for follow-up studies.