**Response to Reviewer A:**

Thank you for your generous and constructive comments and suggestions. In the following, your original comments are in italic.

*The use of CRMs or LES as point of comparison to SCMs and as driven by the identical forcing datasets as SCMs could be made more prominent in the present manuscript. In particular, I think in Section 1.1, they should be discussed saying that when they’re driven by the same forcing that they can be used as a source of information on some quantities that observations cannot yet give.*

We have revised the paper to follow your suggestion. It is in Section 1.1, lines 48-51.

*I also think that there should be some more explicit reference to the use of SCM by GCSS/GASS – rather than only references to the GCSS/GASS cases that happen to use ARM datasets. In particular, it would be worth mentioning that one can create forcing datasets focused on a particular sub-component of the atmosphere – as was done in numerous boundary layer cases of the former GCSS boundary layer cloud working group (e.g. BOMEX, ATEX, DYCOMS, etc.). The forcing for these cases can be idealized from the observations, and indeed one of these cases is the ARM Shallow Cumulus case (used in the Lenderink and Suselj references of this paper) – which used the constrained variational analysis as only one part of a forcing assembled from different sources of information.*

Thank you for the suggestion. We have added a paragraph in Section 1.2, lines 106-115.

*Finally, it would be worth mentioning somewhere that SCMs with prescribed forcing have limitations. For example, SCMs/CRMs driven in the way described here are somewhat less appropriate for deep convection cases when the convection propagates into the domain or middle-latitude cyclones with differential advection, whereas they are more appropriate for boundary layer cloud conditions under conditions of subsidence, where the break between the large-scale and small-scale is very clear. One might also mention the use of the weak-temperature gradient SCM/CRM modeling of tropical convection (e.g. Sobel et al. JAS 2001, Zhu and Sobel QJ 2012) to get around problems of not having interactive large-scale forcing.*

These are excellent points. We have revised the paper by adding a new section of 1.9 to discuss the limitations on page 24-25.

*Detailed Minor comments*

*Line 10. “The chapter also” is an incomplete sentence.*

Corrected

*Line 76. Lee et al. is not in the reference list.*

Corrected

*Lines 86 – 88 and Lines 93 – 96 say the same thing and thus are redundant.*

Redundancy removed.

*Lines 102 – 104. The sentence beginning with “Because” is not complete.*

Corrected.

*Line 118. “vapor” not “vaper”.*

Corrected.

*Line 165. I would say “reduced” not “removed”.*

Corrected.

*Line 211. “Barnes” not “Barns”*

Corrected.

*Line 243. The term with ql needs to be multiplied by L in order for the units to be consistent. This mistake was also present in Zhang et al. 2001*

Corrected. This was a typo in the original paper. Thank you for pointing this out.

*Line 283. “are” not “ar”*

Corrected.

*Line 292. “Hume” not “Humme”; “Davies” not “Davis”*

Corrected.

*Lines 301-309. This might be slightly rewritten to indicate that these boundary sites and wind profilers have not and are not always available. For example, ARM decommissioned the boundary facilities in 2010. I’m also not sure of the current status of the NOAA wind profilers.*

Revised as you suggested.

*Line 334. I think that currently MWR is only available at the Central Facility and that the other stations to which you refer were at decommissioned boundary facilities.* Done. Thank you.

*Line 377. MC3E was in “2011” not “2012”*

Corrected.

*Line 408. Guichard is not in the reference list.*

Added now.

*Line 431. Lenderink is not in the reference list.*

Added now.

*Line 440-441. This is not a complete sentence.*

Corrected.

*Line 443. For MPACE, only the Morrison paper used the constrained variational analysis derived forcings. The Klein paper created their own forcing from ECMWF analysis data.*

The Klein paper is now removed.

*Line 459. Evaluation should also include observations – not just CRMs and LES.*

Corrected. Thank you!

**Response to Reviewer B**

Thank you for your generous and constructive comments and suggestions. In the following, your original comments are in italic.

*Two topics that are not discussed, which I think deserve mention, are (a) the limitations of SCMs, and (b) the consequences of using the constrained variational analysis forcing for what can and cannot be learned.*

*Regarding (a): SCMs have become an invaluable tool for evaluating the effects of different parameterization assumptions (and by the way, perhaps the contribution of SCMs and ARM to GCSS - now GASS - should be mentioned somewhere). However, we have found out over the years that they are not perfect predictors of what one will experience when one tries to port one's insights from an SCM into the parent GCM, because of how feedbacks with the dynamics can sometimes change things. Done GCSS, GCM relevance limitation,*

To follow your suggestion, we have added a paragraph in Section 1.2, lines 106-115, to describe GCSS and GASS and their relationships with ARM. We have revised the paper by adding a new section of 1.9 to discuss the limitations on page 24-25.

*Regarding (b): What one can learn from an SCM using the constrained variational analysis forcing depends on how one uses it to drive the SCM. Many SCM studies just apply the forcing directly. In precipitating regimes, the observed precipitation is the strongest constraint on the variational analysis, resulting in an adjustment of the grid-scale vertical velocity to be consistent with the precipitation. Most SCMs, given that vertical velocity, will produce approximately the correct precipitation. Thus, an SCM run that way cannot provide information about the fidelity of a model's representation of precipitation processes. It can provide information on other things (clouds, heating and moistening profile errors, etc.), so it is useful, but for other reasons. I find that a large segment of the community still does not understand this. Done in the limitation section a new section is added*

These are now described in the new Section 1.9. Thank you for the excellent points.

*Another issue with using the forcing directly is that as model temperature and moisture errors build up, after a few days a given SCM can drift into a very different climate regime and is no longer useful as a diagnostic after that. In some cases, given the nonlinear behavior of some parameterizations, a small change in the initial condition can even cause the model to drift to a completely different state. This was noted by Hack and Pedretti (2000), which I see it is in the references but is not mentioned in the text. It should be. Thus, most SCM studies are done for a few days at best.*

*One can use the forcing in a different way, applying it and then adding a relaxation term that adjusts the SCM T and q back to the observed values over a time scale of a few hours. With this relaxation approach, the SCM never drifts too far, and now since the SCM is decoupled a bit from the precipitation effect on the forcing omega, it becomes possible (1) to look at how well the SCM simulates precipitation, and (2) to run the SCM for longer periods of time, since it is more anchored to the observed state. It is no longer as good for diagnosing heating/drying errors (although the relaxation increments contain information about that). I think it would be very valuable to readers of this chapter to understand the different ways of using the forcing; some of this is discussed by Randall and Cripe (1999), but most readers won't be aware of that.*

These are great points. We have added a paragraph (Lines 171 to 177) to describe the possible drift and how people treat the drafts. We have also added a paragraph (Lines 491-496) to describe the sensitivity of SCMs to forcing and initial condition as well as the use of ensemble forcing.

Minor edits:

*Line 11: Should be "also highlights SCM results" or "also presents highlights of SCM results" or something along those lines reworded*

The abstract is now removed because it is now required.

*Line 55: "realistic data" - there is no such thing. There are only analyses and reanalyses to initialize an NWP simulation. These are a blend of data, model parameters strongly constrained by the data and by basic constraints such as geostrophy, and other model parameters more constrained by the model parameterizations than by any data.*

This has been corrected. “realistic” is removed (line 72)

*Line 67: Computationally intensive, but not impossible, as large ensemble perturbed parameter studies in GCMs have demonstrated in recent years. Originally these were crowd-sourced, but now they can be run on any high-performance computing platform.*

We have revised the sentence (Lines 77-78).

*Lines 102-104: As written this is a clause, not a complete sentence.*

Corrected.

*Lines 113-114: It would be good to mention a specific example; the obvious one is marine stratocumulus.*

Your suggestion is taken (Line 104)

*Lines 134-135: "more often" - I agree with this statement for LES models simulating boundary layer clouds, but I don't think many (any?) CRMs doing deep precipitating convection use moist conserved variables. So "sometimes" rather than "more often."*

Suggestion taken (Line 134). Thank you!

*Line 147: Should be "following”*

Corrected. Thank you!

*Line 202: Should be "datasets" done*

Revised.

*Line 211: Should be "Barnes" done*

Corrected.

*Line 292: Should be "Hume"*

The sentence is now removed.

*Line 425: "ne" should be "one" done*

Corrected.

*Line 441: End of sentence missing*

Corrected.

*Line 443: Add "and" between the two cited papers; and their dates are 2009, not 2010, I believe. Removed the Klein reference because its forcing was constructed from ECMWF analysis*

The Klein paper is now removed.

*Line 450: Should be "Fridlind"*

Corrected.

*Line 456: Should be "of models"*

Corrected.

*Line 459: Should be "hypotheses"*

Corrected.

**Response to the Editor’s comments:**

Thank you for your generous and constructive comments and suggestions. In the following, your original comments are in italic.

*Major: The main thing missing from this article was a sense of how, and perhaps a little more motivation on why, ARM decided that SCM modeling was the approach that was needed. This story wouldn’t be complete without some background on the discussions that were made in the early days of the program: what observations were needed, the trade-offs that were made to get these observations, the challenges that the program needed to overcome in order to achieve the successes that the chapter does a good job explaining. In short, why did ARM do as it did, and what were the decision points?*

We have followed your excellent suggestion and revised the paper by adding two paragraphs at the beginning of the paper (Lines 16-38) to describe the motivation and history of the SCM strategy in ARM.

*As an example: Part of this story, in my opinion, is the formation of the SGP boundary facilities (BFs), which provided information on the boundaries of the domain. But these facilities had more than just radiosondes: there were also MWRs, AERIs, ceilometers, and other instruments. Were these instruments all useful for the forcing datasets? What else would have been nice to have there, from an observational point of view, but was not done (or possible) because of budget?*

In addition to the first two new paragraphs that give some history of the sites (Lines 29-38), we have also added a paragraph on page 15-16, Lines 314 to 320, to describe the boundary facilities and their roles.

Minor points:

*I think that a good portion of the major point above can be alleviated by the incorporation of an introduction. We aren’t using abstracts in this monograph, so you can replace that (and the outline) with a significant introduction that sets the stage for the rest of the chapter.*

Done

The sentence from line 102-104 starting “Because…” is not a complete sentence

Corrected

*The section 1.6 is very heavily oriented to the SGP, as is much of the chapter. However, as was pointed out, these forcing datasets were created for the other sites too, including some AMF campaigns. What were some of the challenges that were experienced there, and how were they overcome?*

We have revised the paper to emphasize that the section primarily describes the SGP. We also added some descriptions of forcing data at the other sites (Lines 314-316, Lines 376-377)

*The continuous forcing dataset is a very unique contribution by the ARM program. I feel that it should be make more prominent – this could be done by taking that material and making it into its own section. Be sure to indicate why the program felt that this was a good thing to do!*

The continuous forcing is now highlighted in a separate new paragraph Line 369-375.

*Fig X1: be sure to indicate that this is the ARM SGP site. And “circles” is misspelled.*

Corrected

*Pg 17, line 441 – this sentence seems to have lost something as the thought isn’t completed.*

Corrected