

## ***Interactive comment on “Variations in air and ground temperature and the POM model: results from the Northern Hemisphere” by R. N. Harris***

### **Anonymous Referee #1**

Received and published: 2 April 2007

This paper is primarily a theoretical investigation that evaluates the skill of the POM model. The author does a good job of making this point clear, and it certainly seems like he has a useful model. The author also provides a "real world" example applying and comparing the model to actual surface air temperatures (SAT) from a Northern Hemisphere data set. However, I think there are some important limitations to this real world example/application, as I outline below. These issues should be resolved prior to publications, and/or adequately toned down.

### **Major Points:**

Are the coarse  $5^\circ \times 5^\circ$  Jones and Moberg (2003) SAT data adequate for this application? From Figure 6a it appears as though many grid cells only have one (or a couple)

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

of boreholes. How can one point-measurement be representative of such a large area? Similarly, some grid cells contain a lot of boreholes, yet all of these sites are "assigned" the same SAT value (or rather, the borehole profiles are averaged, and thus a grid cell that contains one borehole carries equal weight as one that contains 30)? It seems this removes a lot of the variability in the observational records. There are much finer-resolution observations available (though maybe not extending as far back in time). For instance, the CRU TS 2.1 data go back to 1900, and are at a 0.5 degree resolution. Could the author comment on this? Perhaps he is confident that, given the long time scales employed here, that the low-frequency variations from boreholes are indeed comparable to such large SAT grid cell-wide values?

Is a "linear fit" (p. 348, line 13-14) adequate for the very non-linear SAT time series? Figure 6b would certainly imply *no*. Since  $0.8^{\circ}\text{C}$  therefore does not actually represent the SAT variations from 1856-2001, is it coincidence that the model reproduces this same value? Are they both (SAT and the model) incorrect?

Could statistical significance be provided for the power spectrum depicted in Figure 7? There are simple ways to calculate a median red- or white-noise background, and analogous confidence levels (see, e.g., the MTM-SSA toolkit). I think this would add some useful information as to the particular 20+ year spectral peak described on p. 349, line 24.

I think I disagree with the notion that the impact of snow cover on ground temperatures is small (though it may be on air temperatures?). It is of course fine if that is a potential conclusion of this work, yet there is a vast literature suggesting otherwise. Furthermore, this conclusion is probably scale-dependent. Is the finding here that snow cover is not important on *low-frequency variations*?

While this paper certainly reflects very worthwhile theoretical/numerical simulation exercises, I am not sure if I accept the comparisons with Northern Hemisphere data. Given that very [too?] coarse SAT observations are used, the averaging employed re-

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

duces much of the available variability, and then comparing (and reproducing) a potentially incorrect linear SAT trend (when the data themselves are not linear) – certainly some strong caveats are needed, if not improved methodologies. I therefore do not think that Conclusion 3 is supported by this analysis.

### **Minor/Technical Points:**

I think the abstract needs a lot of work, as it is currently written for someone who has already read the paper, as opposed to providing a summary of the paper:

–"POM" is not defined until line 8 of the abstract – this should be done when the acronym is first used (line 1).

–The abstract also refers to "the misfit" (line 6), yet this has not been discussed or introduced.

–Line 10 states that "The POM is significant as long as..." I am not sure what this means. Does it refer to the "mean" or the "model" – the "pre-observation mean" is "important" or the POM model is "skillful" as long as...? Or somehow significant in a "statistically significant" sense?

–What are "these tests" that line 11 refers to?

Line 8 on p. 348: Since temperature measurements cannot, at the same time, be at least as shallow as 100 m (implying 100 m or shallower) and at least as deep (i.e. deeper) than 200 m, does the author mean ..."include *both* temperature measurements at least...?"

---

Interactive comment on Clim. Past Discuss., 3, 337, 2007.