

I would like to see some quantitative analyses of how big this improvement is; 3SOM sounds a conceptually exciting proposition when in a traditional 2-component situation, but this needs to be explicitly demonstrated.

The quantitative analyses of the 2-source model in Bewley et al. (2010) were added to Table 2 for convenience and to facilitate the discussion showing model improvements. A more thorough comparison of the performance of the two models is now in Section 3.4. Please note that the reference height for wind speed was modified (snowdepth was not previously subtracted). This was not described in the previous version of the manuscript but as this was found to have an effect on model results, the description of the resistances was added to the revised manuscript.

For example, a short additional experiment testing the impact of gridbox size to evaluate the influence of boundary line location on fluxes would be welcomed.

Thank you for your suggestion which led to the addition of a new Section (4.1) in the revised manuscript which describes a sensitivity analysis of standard deviation of snow depth to gridbox resolution. We found that modelled standard deviation of snow depth is highly dependent on gridbox resolution and that, as a consequence, larger gridboxes would fail to represent many of the processes explicitly. A short experiment investigating the impact of gridbox size on turbulent fluxes was conducted and summarized in the first paragraph of Section 4.3.

Minor comments:

P 288, ln 7 – please cite the key relevant studies (of the 100 available) rather than rely on the pers comm.

Done.

P 232, ln 8 – can the difference to JULES albedo be stated briefly to explain why this has been changed?

The model description has been expanded and the albedo parametrization is now described more thoroughly. As a result of this expansion, this specific reference to JULES was not deemed relevant and was removed.

P 232, ln 20 – quantify how much closer modelled SWE and depth are to measurements in 2004 than 2003.

This was removed from the revised manuscript because model results changed following the modifications described above.

P 233, ln 27 – ‘perform well enough’. What is well enough? Can a quantitative threshold be provided for this assertion?

This section now provides a more thorough comparison with the 2-source model and of the improvements that the addition of a 3rd source add to representation of processes. The concluding sentence of this section was changed and the terms “well enough” were removed.

P 234, ln 17 – why did you choose a 8 m grid – please justify briefly?

An explanation is now provided in Section 4.1

P 234, ln 18 – what is the resolution of LiDAR data?

An explanation is now provided in Section 4.1

P 235, ln 11 – was the WIA – plateau wind speed difference higher or lower?

Corrected.

P 236, ln 3 – rewrite to say ‘there are some large errors’.

The sentence was re-written.

P 236, In 24 – are these ‘errors’ just enhanced uncertainty during melt as a result of increased spatial variability?

This was clarified in the text.

P 237, In 2 – you have not currently shown that the models have been able to capture evolution of broad spatial patterns. Need to demonstrate this or re-write.

A figure of measured vs. modelled snow depth along the transect with the longest persisting snow cover was added to provide some information that spatial averaging in Figure 7 (previously 6) does not clearly convey. We believe that the addition of this Figure, Section 4.1 and clarification in Section 4.3 have now demonstrated this statement.

P 241, In 28 – in relation to my first main point above, I would suggest that the known limitation (and improvement resulting from this study) need to be explicitly demonstrated here through comparison between 3SOM and the two-source model.

Done; see answer above.

Table 1 – although relatively intuitive, please state the units in the table.

Done (in caption).

Table 2 - please state units (presume meters?).

Done (in caption).

Fig 5 – ‘a’ and ‘b’ are not visible on the plots.

Changed.

Are the peaks in the measured data missing prior to April 30? Why?

An explanation is now provided in the caption.