Comment on tc-2020-377
Anonymous Referee #1

Snow cover has a strong influence on surface energy balance but does not always uniformly blanket the ground. There have therefore been many papers proposing parameterizations for fractional snow-covered area in surface energy balance models, often very simple and based on limited observations. Ideally, a seasonal snow cover parameterization will account for terrain influences, the scale of the model cells and hysteresis between accumulating and melting snow covers. Helbig et al. build on their valuable earlier work to present such a parameterization and evaluate it with several extensive observed datasets. There is good work here, but it is very hard work for the reader; I have read the paper three times and am still struggling. I think that the descriptions, the evaluations and the algorithm itself need to be substantially simplified. We are referred to Helbig et al. (2015b, 2020) for details of the algorithm, and it is actually impossible to tell what is being done here without reading those papers. Brief explanations of how c, d, μ and ξ are calculated should be given. The appendix will be essential (but not quite sufficient) for anyone wishing to implement this algorithm in another model, and the schematic in Figure 1 should be moved to that appendix (the figure is not fully comprehensible just from material presented in the main text). For readers wanting to get an overview of the method, I suggest that an alternative Figure 1 showing typical modelled fSCA behaviour over a season would be better (this is more or less done in Figure 7, but without explaining why the models differ in the ways that they do).

Six different performance measures are presented with very little consideration of what aspects of performance they measure and a lack of context for what could be considered a good performance. The Kolmogorov-Smirnov test statistic implies a significance test and the Q-Q plot statistic suggests a comparison of distribution shapes that are never presented. Cut this down to a set of measures that are meaningfully used to measure performance and to communicate information.

The Niu and Yang (2007) fSCA parameterization can be implemented in one line of code and includes hysteresis to some extent through snow density. Just the pseudocode for the algorithm presented here requires 32 non-comment lines and contains many apparently ad hoc design decisions: what is the significance of 14 days for new snow accumulation?
how flat does a flat cell have to be? why use the flat parameterization for new snow in mountains rather than any other value? Considering uncertainties in observations revealed when different datasets overlap, errors in the modelled mass balance and ad hoc decisions, is the complexity justified? Tables 3 and 4 and Figure 7 are not very convincing in this regard.

If a revised paper is submitted and is more comprehensible, I still expect to have major scientific comments. I look forward to being able to get into that level of detail.