Dear reviewer,

Thank you for your comments. Please find below a first reply to your specific comments.

Comment:

The introduction section lacks a description of the scientific context as it does not provide a description of related work that has already been performed at other NWP centers, internationally. For example, by saying "new approach" (on line 76), presumably this refers to being new at ECMWF and not within the international field of NWP. As an example, an early study by Garand et al (2004, Journal of Applied Meteorology and Climatology) described the Canadian system that already used a gridded skin temperature field for assimilating radiance observations (one of the new innovations described in this manuscript). They then evaluated the impact of including the background-error correlation between skin and air temperature, which in some respects goes beyond what is described in this manuscript.

Reply:

In the revised version, we will expand the review of relevant previous literature and present more clearly what developments are new in the ECMWF/IFS system and what is new or not in the presented methodology compared to the methodology of other NWP centers.

Comment:

The design of the experiments does not allow for a systematic evaluation of how the different modifications affect the results. As already mentioned, in some situations the specified standard deviation of skin temperature appears to have been changed in the new formulation in combination with the use of a common gridded skin temperature field for all instruments and introducing both spatial and temporal background-error correlations. This makes it difficult to draw any conclusions about the impact of the changes over sea ice, where it is noted that the standard deviation is significantly changed. This aspect should be more fully explored.
Reply:

The study presented here extends a previous one published as an ECMWF technical memorandum (Massart et al. 2020). In this previous study, we presented the impact of both spatial and temporal background error correlations in the SKTACV approach. We also showed that due to neglecting the spatial and temporal correlations in the skin temperature background error in the TOVSCV approach, the background error standard deviation needs to be inflated in this approach. In contrast, when the background error correlations are taken into account for the SKTACV approach, the background error standard deviations are not inflated. This makes it difficult to have similar standard deviations in the two approaches as the true inflation factor needed to compare them is unknown and would be complicated to compute (see Appendix of Massart et al. 2020).

We wanted the configuration of control experiment to be as close as possible to the configuration of the operational analysis. In particular, we kept the value of 7.5K for the background error standard deviation over sea-ice. This value accounts not only for random errors but also for the large systematic errors present in polar regions. For the SKTACV approach the value is around 1.5K (using the times 3 factor) and represents only the random part of the errors. This is indeed significantly lower than for the control, even if we consider the effective standard deviation (from taking into account the spatial correlations). We chose to nevertheless select this value and to work in parallel on the reduction of the systematic errors (not part of this study). This means that it may be indeed difficult to draw any conclusions about the impact of the changes over sea-ice, but in the meantime we have a reference for future improvement of the SKTACV approach there.

Comment:

The reduced amplitude of the analysis increments with the new approach are assumed to be related to the additional constraints imposed on skin temperature increments in the new approach. These constraints result from the addition of spatial and temporal correlations and also the use of a single 2D skin temperature field for all microwave instruments and another for all infrared instruments. I would think that some additional analysis of the results could provide concrete evidence of this and possibly suggest if any of these differences are more important than the others. For example, the increments obtained with the two approaches could be compared (at observation locations) in terms of their spatial, temporal and between instrument variability. Showing a much higher variability when using the previous approach, either in space, time or between instruments, would support the assertion that one or more of the added constraints cause reduced increments. The authors could then also comment on if the higher variability is more or less physically realistic.

Reply:

The increments obtained with the two approaches have been compared in terms of their spatial variability (not shown) but not their temporal and in between instrument variability. We will compute these quantities and present them either in the final reply to the reviewer or in the revised version of the paper.