Multi-sensor analyses of the skin temperature for the assimilation of satellite radiances in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS, cycle 47R1)

Sebastien Massart[1], Niels Bormann[1], Massimo Bonavita[1], and Cristina Lupu[1]

[1]ECMWF Shinfield Park Reading RG2 9AX United Kingdom

Correspondence: S. Massart (sebastien.massart@ecmwf.int)

Abstract. The assimilation of clear-sky radiance in the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric analysis relies on the clear-sky radiances observation operator. Some of these radiances have frequencies that make them sensitive to both the surface and atmosphere. Because the atmospheric and surface analyses are currently not strongly coupled, a specific treatment of the surface is required. The observation operator expects in particular, a skin temperature value

5   at the observation location and time, together with the profiles of the atmospheric variables along the viewing path. This skin temperature is added to the control variable and optimised simultaneously with all the atmospheric variables to produce optimal simulated radiances.

We present two approaches to add the skin temperature to the control variable. In the current TOVSCV approach, a series of skin temperature value per observation location is added to the control variable. Effectively, in the optimisation process, the

10   skin temperature acts as a sink variable in observation space and is uncoupled from the skin temperature at other locations. In the novel SKTACV approach, two-dimensional skin temperature fields are added to the control variable. All clear-sky radiances then participate in the optimisation of these two-dimensional fields and the analysis produces temporally and spatially consistent skin temperature fields.

We compare the two approaches over two seasons of three months each. Overall, there is a neutral impact of the new

15   approach on the analysis and forecast. Besides, there are some evidences that the contribution of the sub-surface layers should be represented in the new approach for the skin temperature associated with the microwave instruments.

1   Introduction

Data assimilation methods used in Numerical Weather Prediction (NWP) exploit all the available observations by computing

20   the difference between them and the model state at the observation time and location. This later is obtained by first integrating in time the model state and then transforming it into an observation-equivalent by the so-called observation operator. The

observation operator is dependent on the observation type. For most of the in-situ observations, the observation operator is a simple time and space interpolation. For other observation types, the operator can be much more complex. For instance, a fast and accurate radiative transfer model to simulate observations is required for satellite observations. Furthermore, the
25   complexity of the observation operator increases in an all-sky radiance assimilation context, where cloud and precipitation absorption and scattering have to be explicitly modelled.

In this paper, we focus on the observation operator used to assimilate clear-sky radiances. Radiances represent the vast majority in numbers of the data that are currently assimilated in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS). Although other satellite measurements make important contributions to the
30   ECMWF forecasting skill, microwave and infrared soundings are among the most important contributors when it comes to improving the forecast skill (McNally, 2015; Bormann et al., 2019). This makes the observation operator for these soundings of utmost importance.

For radiance assimilation, the observation operator in the IFS contains a radiative transfer component that integrates the radiative transfer equation from the outputs of the forecast model along the viewing path. This requires model profiles of
35   pressure, temperature, specific humidity, cloud properties and atmospheric composition, and, for the surface, surface emissivity, skin temperature, surface pressure, 2m-temperature, 2m-humidity and 10m-wind. The sensitivity to the various layers of the atmosphere and to the surface in the radiative transfer equation depends on the frequency of the measurement. Observing radiances at various frequencies therefore provide indirect information on the atmospheric variables on diverse atmospheric layers and on the surface variables (English, 1999).

40   In the IFS, this information on model atmospheric variables and model surface variables can not be used together because the atmospheric analysis and the surface analysis are only coupled during the first guess forecast step, not during the analysis update itself (weak coupling, de Rosnay et al. (2014); Browne et al. (2019)). This is problematic because an inaccurate surface emissivity or skin temperature in the observation operator could lead to an inaccurate analysis of the temperature profile or could mislead the cloud detection scheme for example (English, 2008). To bypass this issue, the spectral channels currently
45   entering the atmospheric analysis are carefully selected to avoid those which are very sensitive to the surface, and at the same time to keep those which have a significant positive forecast impact in the ECMWF system (Bormann et al., 2017).

To still be able to assimilate the remaining radiances with frequencies sensitive to the surface, the observation operator should use a surface emissivity and skin temperature that are as accurate as possible. To get the best possible surface emissivity for each surface type, instrument geometry and observation wavelength, different methods are used depending on the surface
50   type. Over the ocean, the emissivity can be accurately computed for calm waters (English and Hewison, 1998). In the IFS we use a fast parameterised model that accounts for surface wind and produces an ocean surface emissivity with a good accuracy (Kazumori and English, 2015). In the absence of snow, land emissivity varies very little, temporally or spatially. An emissivity database is then used for the infrared observations, instead of using directly emissivity models that are less accurate and may present large uncertainties and biases. For the microwave observations, surface emissivity is retrieved from window channels
55   using the model emissivity and skin temperature as background values (Karbou et al., 2006). This combined approach is a

compromise between an accurate surface emissivity and a limited computational cost. Still, some biases remain especially over the land and sea-ice.

To get the best possible skin temperature, skin temperature values were added to the IFS 4D-Var control vector for the atmospheric analysis in the early 90s. This inclusion allows to adjust the skin temperature during the atmospheric analysis

60 and circumvents the de-coupling between the atmospheric and surface analyses. It was developed when the radiances from the TOVS (TIROS-N (Television Infra-Red Observation Satellite) Operational Vertical Sounder) sounder were first directly assimilated (Eyre et al., 1992). Since then, this technique is referred to as TOVS Control Variable (TOVSCV, ECMWF, 2019) and it is used in the assimilation of clear-sky radiances. All the radiances associated with a particular field of view of a given instrument share the same value of skin temperature in the radiative transfer code used in the IFS. This gives a value of skin

65 temperature per field of view per instrument. The TOVSCV approach consists of adding all these separate values of skin temperature to the control vector. Then, the individual skin temperatures are optimised in the 4D-Var minimisation together with the other model physical variables. In this framework, each satellite sounding produces independent estimates of skin temperature, independently to its spatial scale and spectral characteristic.

When assimilating a group of radiances within the same field of view, the model state along the viewing path is constrained

70 by these radiance observations, but also by all the other measurements available within a volume centred around the field of view and with a size proportional to the background error local de-correlation lengths. By contrast, in the TOVSCV approach, the skin temperature is only constrained by the radiances within the same field of view and therefore is free to be adjusted within the constraint of its background error and within the constraint of the sensitivity of each radiance frequency to the skin temperature. Because of this lack of constraint by other surrounding measurements, the skin temperature adjustment likely

75 compensates for other errors in the background state, thus reducing the accuracy of the atmospheric analysis.

A new approach is proposed here as a substitute for TOVSCV. The general idea is that the skin temperature value provided to the radiance observation operator is derived from a two-dimensional skin temperature field interpolated at the observation time and location. This two-dimensional field is then introduced in the IFS control vector instead of independent skin temperature values, following the augmented control variable developments reported in Massart (2018). As a result, the skin temperature

80 of each field of view depends on the analysis of the two-dimensional field which is consistent with all the radiances and other observations, and consistent in space. This adds constraint on the skin temperature used in the observation operator and we want to assess here how this impacts the use of satellite radiance data in the IFS analysis.

In section 2, we present in more detail the radiance assimilation and the two approaches for the analysis of the skin temperature used in the radiance observation operator. In particular, we explain why only a single two-dimensional skin temperature

85 field is not sufficient in the SKTACV approach and how we introduce additional fields. These fields allow representing the time variation of the skin temperature and the different spectral channels of the instruments. Then, we detail in section 3 the chosen configuration of the background errors of the skin temperature fields added to the control vector in the new approach. We finally present in section 4 the results from two experiments, one for each approach, and discuss the differences.

2 Skin temperature in the assimilation of radiance

90 The skin temperature is introduced in the observation operator used in the IFS in order to assimilate radiances for the atmospheric analysis. We present here how the skin temperature is also included in the state vector and analysed in observation space. We then present a novel approach that adds several skin temperature fields in model space to the control vector. We focus hereafter only on the assimilation of the radiance measurements sensitive to the surface. To simplify the mathematical formalism, we keep in the next section only these radiance measurements and we ignore all the other type of measurements.

95 ### 2.1 Radiance assimilation

Let us introduce $\boldsymbol{y}_i^o$, an observed set of radiances at various frequencies but all belonging to the same field of view and measured by the same instrument, and for which i is the index of the observation in the observation vector $\boldsymbol{y}^o$. To assimilate these observed radiances, one has to first compare them to equivalent radiances derived from the model variables (or simulated radiances). These model-equivalent radiances are obtained with the observation operator H_i applied to the atmospheric model

100 state vector $\boldsymbol{x}_i^{atm}$ and to the surface state vector $\boldsymbol{x}_i^{sfc}$, both at the time of the observation,

$$\boldsymbol{y}_i = H_i\left(\boldsymbol{x}_i^{atm}, \boldsymbol{x}_i^{sfc}\right). \tag{1}$$

The atmospheric model state vector contains the three-dimensional distribution of the atmospheric model variables and the surface state vector contains the two-dimensional distribution of the surface model variable at the observation time. In practice, the part of the observation operator that integrates the radiative transfer equation requires only the profile of the atmospheric

105 model variable along the viewing path and the surface variables at the observation location. The observation operator H_i is then decomposed into the combination of two operators: (i) $H_{I,i}$ that performs the interpolation of the state vectors to produce profiles of the model variables along the viewing path and values of the model surface variables at the observation location, and (ii) $H_{R,i}$ that performs the radiative transfer computation to convert the interpolated model state into radiance at the required frequencies. In the IFS, this operation is performed by the Radiative Transfer for TOVS (RTTOV, Matricardi et al.,

110 2004; Saunders et al., 2018) code. Using this decomposition of the observation operator, the model equivalent to a radiance is expressed as

$$\boldsymbol{y}_i = H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i^{atm}\right), H_{I,i}\left(\boldsymbol{x}_i^{sfc}\right)\right). \tag{2}$$

For the atmospheric analysis, the surface variables are not updated, except for the skin temperature in the TOVSCV approach. We thus keep from all the surface variables $\boldsymbol{x}_i^{sfc}$, only the skin temperature at the observation location τ_i from $H_{I,i}\left(\boldsymbol{x}_i^{sfc}\right)$.

115 The observation operator equation becomes

$$\boldsymbol{y}_i = H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i^{atm}\right), \tau_i\right). \tag{3}$$

In the IFS, the skin temperature τ_i used in the observation operator is the skin temperature from the first guess extracted at the observation time and interpolated in space using $H_{I,i}$. The model state at the observation time $\boldsymbol{x}_i^{atm}$ comes from the model

state at the initial time $\boldsymbol{x}_0^{atm}$ and integrated by the model $\mathcal{M}_{i,0}$ between the initial time and the time of the observation $\boldsymbol{y}_i^o$,

$$\boldsymbol{x}_i^{atm} = \mathcal{M}_{i,0}\left(\boldsymbol{x}_0^{atm}\right),\tag{4}$$

120

Unless specified otherwise, the term skin temperature refers hereafter to the skin temperature used in the radiances observation operator and not the model skin temperature.

2.2 Current formulation

The idea behind the TOVSCV approach is to adjust τ_i together with the atmospheric variables. With this aim, each τ_i is renamed $x_{p,i}$ in order to be consistent with the notation of the state vector. All the scalar values $x_{p,i}$ are concatenated into the vector $\boldsymbol{x}_p$, which is then added to the atmospheric state vector and leads to the augmented state vector

125

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_p \end{pmatrix}.\tag{5}$$

Note that the superscript atm had been removed from the atmospheric state vector $\boldsymbol{x}_0$ for simplification. With these notations, the observation operator of Eq. (3) becomes

$$\boldsymbol{y}_i = H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i\right), x_{p,i}\right),\tag{6}$$

130

with $\boldsymbol{x}_i$ still computed from Eq. (4) and with

$$x_{p,i} = H_{I,i}\left(\boldsymbol{x}_p\right),\tag{7}$$

There are fundamental differences between the two components of the augmented state vector of Eq. (5). First, $\boldsymbol{x}_0$ is a vector containing the atmospheric model variables defined on the model grid and at the initial time. On the other hand, $\boldsymbol{x}_p$ is a vector containing a surface variable at the observation locations and at the observation times. For these reasons, we identify $\boldsymbol{x}_p$ as defined in observation space while $\boldsymbol{x}_0$ is defined in model space and at initial time. Both components are optimised simultaneously during the minimisation of the 4D-Var cost function. In the current formulation, the background errors between the model state vector $\boldsymbol{x}_0$ and the interpolated skin temperatures in observation space $\boldsymbol{x}_p$ are assumed to be uncorrelated. The background error covariance matrix is block diagonal and can be split in two: $\mathbf{B}$ associated with the background state $\boldsymbol{x}_0^b$ and $\mathbf{B}_p$ associated with the background state $\boldsymbol{x}_p^b$ of $\boldsymbol{x}_p$ for which each of component i is derived from the background values of the skin temperature,

135

140

$$x_{p,i}^b = \tau_i.\tag{8}$$

The 4D-Var cost function $J(\boldsymbol{x}_0, \boldsymbol{x}_p)$ is then

$$
\begin{aligned}
J(\boldsymbol{x}_0, \boldsymbol{x}_p) &= J^b(\boldsymbol{x}_0, \boldsymbol{x}_p) + J^o(\boldsymbol{x}_0, \boldsymbol{x}_p) \\
&= \frac{1}{2}\left[\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right] \\
&+ \frac{1}{2}\left[\boldsymbol{x}_p - \boldsymbol{x}_p^b\right]^T \mathbf{B}_p^{-1}\left[\boldsymbol{x}_p - \boldsymbol{x}_p^b\right] \\
&+ \frac{1}{2}\sum_i \left[\boldsymbol{y}_i^o - H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i\right), x_{p,i}\right)\right]^T \\
&\quad \mathbf{R}_i^{-1}\left[\boldsymbol{y}_i^o - H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i\right), x_{p,i}\right)\right],
\end{aligned}
\tag{9}
$$

where $\mathbf{R}_i^{-1}$ is the observation error covariance matrix for the observation $\boldsymbol{y}_i^o$.

From Eq. (9), we can compute the gradient of the part of the cost function related to the radiance observations with respect to each component $x_{p,i}$ of the vector $\boldsymbol{x}_p$,

$$
\frac{\partial J^o(\boldsymbol{x}_0, \boldsymbol{x}_p)}{\partial x_{p,i}} = H_{R,i}^T \mathbf{R}_i^{-1}\left[\boldsymbol{y}_i^o - H_{R,i}\left(H_{I,i}\left(\boldsymbol{x}_i\right), x_{p,i}\right)\right],
\tag{10}
$$

where $\mathbf{H}_{R,i}^T$ is the adjoint of $H_{R,i}$ with respect to $x_{p,i}$. It is clear from this expression that the only observations that will directly contribute to the update of $x_{p,i}$ are $\boldsymbol{y}_i^o$, i.e. all the radiances within the same field of view of the same instrument.

Because the skin temperature values are independent, the matrix $\mathbf{B}_p$ is diagonal. The diagonal is filled with different values for different surface types, i.e., land, ocean, sea-ice. This means that the observation $\boldsymbol{y}_i^o$ will bring information on $x_{p,i}$ as shown by Eq. (10), and only on $x_{p,i}$ because of the lack of cross-covariances in $\mathbf{B}_p$. There is thus no spatial consistency among all the analysed values of the skin temperatures $x_{p,i}$ in observation space and each value $x_{p,i}$ is analysed independently from one another.

The minimisation of the 4D-Var cost function in the IFS follows the incremental approach proposed by Courtier et al. (1994). This is an iterative process in which the 4D-Var cost function is successively linearised around a first guess. The first guess is chosen as the analysis from the previous iteration or the background state for the first iteration. The TOVSCV implementation follows the incremental approach and a new first guess of $\boldsymbol{x}_p^g$ is computed at every outer loop iteration. The iterative process also provides a new first guess of the skin temperature from the surface analysis, but this is currently not used for the radiance assimilation.

To summarise the TOVSCV approach, the assimilation of a set of radiance observations from the same field of view $\boldsymbol{y}_i^o$ will adjust the three-dimensional atmospheric variables within the domain defined by the local background error correlation length-scale. All the other observations within this domain also constrain the atmospheric variables. On the other hand, the assimilation will adjust the skin temperature $x_{p,i}$ local to the field of view only and will have no constrain from the surrounding observations.

Currently, in the IFS, the radiance observations are thinned to a resolution of around $125\,\mathrm{km}$ and concatenated into $30\,\mathrm{min}$ time slots. Over the land, the skin temperature can be spatially very heterogeneous and can change quickly in time. Under these conditions, and if the characteristic length scale of the spatial heterogeneity is lower than $125\,\mathrm{km}$ and the characteristic time

scale is under $30\,\mathrm{min}$, the TOVSCV approach could be sufficient. For other situations that may occur over the ocean where the skin temperature is more homogeneous and varies slowly in time, we believe that the TOVSCV approach could be improved by constraining $x_{p,i}$ with surrounding radiance observations in space and time. We expect that these additional constraints could be beneficial for the assimilation of the radiance observations.

2.3 New formulation

It would be possible to constrain neighbouring values of $\boldsymbol{x}_p$ in the current TOVSCV formulation by adding correlations in the associated background errors. But this would require to build a correlation model in space and in time on the unstructured grid formed by the current set of observation locations and times, grid that would change at each assimilation cycle. Even if this is feasible, we prefer instead a new formulation for TOVSCV that analyses the skin temperature used in the radiative transfer equation directly in model space, where the background error correlations are easier to compute.

2.3.1 Skin temperature analysis in model space

The new formulation, first documented in Massart et al. (2020), is also based on Eq. (3), but the skin temperature value τ_i is replaced by an interpolated value from a two-dimensional field $\boldsymbol{x}_\alpha$. We choose this field to have the same horizontal dimension as $\boldsymbol{x}_0$ and we use the same spatial interpolation operator $H_{I,i}$, so τ_i becomes $H_{I,i}(\boldsymbol{x}_\alpha)$ and Eq. (3) becomes

$$
\begin{aligned}
\boldsymbol{y}_i &= H_{R,i}\left(H_{I,i}(\boldsymbol{x}_i)\,H_{I,i}(\boldsymbol{x}_\alpha)\right) \\
&= H_{R,i} \circ H_{I,i}(\boldsymbol{x}_i, \boldsymbol{x}_\alpha).
\end{aligned}
\tag{11}
$$

As in the previous method, the vector $\boldsymbol{x}_\alpha$ is added to the state vector $\boldsymbol{x}$ to form the augmented state vector

$$
\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_\alpha \end{pmatrix}.
\tag{12}
$$

The initial results obtained with the introduction of $\boldsymbol{x}_\alpha$ were encouraging but we encountered some degradations detailed in Massart et al. (2020). We associated these degradations to two plausible causes. First, the skin temperature may not have the same meaning for different instruments as the depth of the surface layer that contributes to the observed radiation depends on the used spectral band. Secondly, we were not accounting for the time evolution of the skin temperature within the assimilation window. Therefore, we extended $\boldsymbol{x}_\alpha$ to address these two issues.

2.3.2 Spectral band dependency

The ground depth down to which a radiance observation is sensitive to depends on the frequency of its measurement (Prigent et al., 1999). This depth defines the skin temperature for this particular radiance observation, which makes the skin temperature dependent on the spectral band of the observation. For this reason, we decided to have a different skin temperature field for the two separate spectral bands used in the radiative transfer code RTTOV: microwave (*mw*) and infrared (*ir*). Thus, we introduced

two separate two-dimensional fields in $\boldsymbol{x}_\alpha$ instead of one,

$$\boldsymbol{x}_\alpha = \begin{pmatrix} \boldsymbol{x}_\alpha^{mw} \\ \boldsymbol{x}_\alpha^{ir} \end{pmatrix}, \tag{13}$$

where $\boldsymbol{x}_\alpha^{mw}$ and $\boldsymbol{x}_\alpha^{ir}$ are respectively the skin temperature field associated with the microwave and infrared instruments. For
205   simplicity, we will hereafter refer to these two fields as microwave and infrared skin temperature.

2.3.3   Time dependency

Over the land, the amplitude of the skin temperature diurnal cycle can be considerable, and the difference between the minimum and maximum skin temperature during the day can reach $30\,\mathrm{K}$ over the desert areas (Pinker et al., 2007). Over the ocean, the amplitude of the surface temperature diurnal cycle is usually small and less than $1\,\mathrm{K}$. Nevertheless, the amplitude can be
210   larger during particular events. For example, daily increases between $5\,\mathrm{K}$ and $7\,\mathrm{K}$ were observed by independent satellite measurements of ocean surface temperature (Gentemann et al., 2008). For these reasons, the fields in $\boldsymbol{x}_\alpha$ should also evolve in time within the 12-hour assimilation window.

We chose to expand the state vector $\boldsymbol{x}_\alpha$ to one skin temperature field per hour and per instrument type. For our 12-hour assimilation window, we then have 13 fields per instrument type,

$$215 \quad \boldsymbol{x}_\alpha^{mw} = \begin{pmatrix} \boldsymbol{x}_\alpha^{mw}\left(t_0\right) \\ \vdots \\ \boldsymbol{x}_\alpha^{mw}\left(t_{12}\right) \end{pmatrix}, \quad \boldsymbol{x}_\alpha^{ir} = \begin{pmatrix} \boldsymbol{x}_\alpha^{ir}\left(t_0\right) \\ \vdots \\ \boldsymbol{x}_\alpha^{ir}\left(t_{12}\right) \end{pmatrix}. \tag{14}$$

This means that now $\boldsymbol{x}_\alpha$ contains $2 \times 13 = 26$ two-dimensional fields. We thus introduced a spectral band and time selection operator $H_{S,i}$ in the radiance observation operator. The spectral band selector picks the microwave or infrared field depending on the instrument spectral band. The time selection operator then picks from $\boldsymbol{x}_\alpha^{mw}$ or $\boldsymbol{x}_\alpha^{ir}$ the two closest fields in time from the observation time and performs a linear interpolation in time between these two fields. With this selector operator, the model
220   equivalent to radiance observations is

$$\boldsymbol{y}_i = H_{R,i} \circ H_{I,i}\big(\boldsymbol{x}_i, H_{S,i}\left(\boldsymbol{x}_\alpha\right)\big) \qquad . \tag{15}$$

2.3.4   Cost function and its gradient

As a first step, we chose to use the same background for both $\boldsymbol{x}_\alpha^{mv}$ and $\boldsymbol{x}_\alpha^{ir}$, and we chose the model hourly skin temperature fields from the short-range forecast that we concatenated into $\boldsymbol{x}_\alpha^{b}$. We therefore ignore the spectral dependence of the skin
225   temperature in our background constraint at this stage. With $\mathbf{B}_\alpha$ the background error covariances associated with $\boldsymbol{x}_\alpha^{b}$, and

omitting the observations that are not radiances, the cost function of the proposed method is

$$
\begin{aligned}
J\left(\boldsymbol{x}_0, \boldsymbol{x}_\alpha\right) = & \frac{1}{2}\left[\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right] \\
& + \frac{1}{2}\left[\boldsymbol{x}_\alpha - \boldsymbol{x}_\alpha^b\right]^T \mathbf{B}_\alpha^{-1}\left[\boldsymbol{x}_\alpha - \boldsymbol{x}_\alpha^b\right] \\
& + \frac{1}{2}\sum_i \left[\boldsymbol{y}_i^o - H_{R,i} \circ H_{I,i}\left(\boldsymbol{x}_i, H_{S,i}\left(\boldsymbol{x}_\alpha\right)\right)\right]^T \\
& \quad \mathbf{R}_i^{-1}\left[\boldsymbol{y}_i^o - H_{R,i} \circ H_{I,i}\left(\boldsymbol{x}_i, H_{S,i}\left(\boldsymbol{x}_\alpha\right)\right)\right],
\end{aligned}
\tag{16}
$$

and its gradient with respect to $\boldsymbol{x}_\alpha$ is

$$
\begin{aligned}
\frac{\partial J\left(\boldsymbol{x}_0, \boldsymbol{x}_\alpha\right)}{\partial \boldsymbol{x}_\alpha} = & \mathbf{B}_\alpha^{-1}\left[\boldsymbol{x}_\alpha - \boldsymbol{x}_\alpha^b\right] \\
& + \sum_i H_{S,i}^T H_{I,i}^T H_{R,i}^T \mathbf{R}_i^{-1} \\
& \left[\boldsymbol{y}_i^o - H_{R,i} \circ H_{I,i}\left(\boldsymbol{x}_i, H_{S,i}\left(\boldsymbol{x}_\alpha\right)\right)\right],
\end{aligned}
\tag{17}
$$

where $\mathbf{H}_{S,i}^T$ and $\mathbf{H}_{I,i}^T$ are respectively the adjoint of $H_{S,i}$ and $H_{I,i}$ with respect to $\boldsymbol{x}_\alpha$. From these expressions, it is clear that each radiance observation inside the assimilation window will contribute to update the vector $\boldsymbol{x}_\alpha$.

For each radiance observation, the information from the observation space toward the model space is propagated through $\mathbf{H}_{I,i}^T$ and then $\mathbf{H}_{S,i}^T$. If, for example, $H_{I,i}$ is a bi-linear interpolation then $\mathbf{H}_{I,i}^T$ will propagate back the information in the four surrounding model grid cells of the observation location and then either to the microwave or infrared field. All other radiance observations for which the field of view is within these grid cells will also provide information on the skin temperature of these grid cells and not only on the grid cell where they are observed as for the TOVSCV approach. The information is then further spread in space through the background error covariances $\mathbf{B}_\alpha^{-1}$.

To improve the condition number of the cost function and to accelerate the convergence, the state vector of Eq. (12) is transformed into the control vector χ, such that

$$
\begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_\alpha \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_0^b \\ \boldsymbol{x}_\alpha^b \end{pmatrix} + \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_\alpha \end{pmatrix} \chi.
\tag{18}
$$

2.3.5 Background errors

We assume here that there is no cross-correlation in the background errors of the skin temperature fields between the two fields $\boldsymbol{x}_\alpha^{mw}$ and $\boldsymbol{x}_\alpha^{ir}$. This may appear to be a strong assumption, especially because they share the same background, but it reflects that we have at this stage little knowledge of these cross-correlations. The background error covariance matrix is then block-diagonal, each block representing a spectral band,

$$
\mathbf{B}_\alpha = \begin{pmatrix} \mathbf{B}_\alpha^{mw} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_\alpha^{ir} \end{pmatrix}.
\tag{19}
$$

Both $\mathbf{B}_\alpha^{mw}$ and $\mathbf{B}_\alpha^{ir}$ contain spatial and temporal covariances between the skin temperature fields.

2.3.6 Other considerations

In the absence of a model to propagate in time the fields in $\boldsymbol{x}_\alpha$, our strategy is effectively a hybrid between a 4D-Var for most
255 variables and a 4D-EnVar for the skin temperature fields (Liu et al., 2008).

There is no technical difference between the skin temperature fields $\boldsymbol{x}_\alpha$ and any other field of the control vector. Therefore, no specific developments were needed for the incremental formulation of the IFS 4D-Var.

The forecast model integrates in time the 4D-Var analysis state to produce the background for the next cycle. Here, we do not propagate in time the analysis of $\boldsymbol{x}_\alpha$. Instead, we use the new background skin temperature from the model as the background
260 $\boldsymbol{x}_\alpha^b$ for the next assimilation cycle.

To summarise, we have introduced a new variable $\boldsymbol{x}_\alpha$ in the 4D-Var state vector that combines two-dimensional fields in the model space representing the skin temperature for a specific spectral band and time. These fields are used in the observation operator for the clear-sky radiance assimilation instead of the discrete values of the current TOVSCV approach. We refer to the new approach as SKTACV (SKin Temperature in the Augmented Control Vector). In this new approach, the transformation of
265 model variables into radiance observation equivalents is based on the same skin temperature field for all observations in the same spectral band and, as a consequence, all radiance observations are used to optimise this field.

3 Background errors for the SKTACV formulation

As new components of the state vector, the new fields associated with the skin temperature require the specification of their background error covariance matrix. Similarly to other variables of the state vector, we decompose the background error
270 covariances of the new fields into the background error standard deviations and the background error correlations, that is based on the wavelet formulation (Fisher and Andersson, 2001; Massart et al., 2020).

Differently to the other variables of the state vector, we have one field per hour (for each spectral band). This means that on top of the spatial correlations of the background errors, we also have temporal correlations that we should account for.

3.1 Standard deviation

275 We saw from Eq. (9) that the TOVSCV formulation requires values for the standard deviation of the skin temperature background errors at the observation locations. For this approach, the background error standard deviation is constant over the ocean and sea-ice, with respective values of 1 K and 7.5 K. These values have been largely empirically derived based on analysis performance. Over the land, the errors can be very heterogeneous and can change rapidly during the day, synoptically and seasonally. To have situation-dependent background error standard deviations there, we make use of hourly estimates of the
280 model skin temperature background error standard deviation which are part of the 50-members ensemble of data assimilations (EDA, Bonavita et al., 2012). The TOVSCV approach uses these EDA-based estimates interpolated to the observation location and time. These values were found to be not large enough and an additional 1.5K is added over the land. This value was derived based on trial and error (Cristina Lupu, ECMWF, personal communication).

For the new SKTACV formulation, we follow the implementation of the other variables of the state vector. We need for each of the additional two-dimensional fields a standard deviation map defined on a horizontal reduced gaussian grid corresponding to a T159 triangular spectral truncation in spectral space. We chose to have the same standard deviations for the two spectral bands and we compute the associated maps with the hourly background error standard deviation of the model skin temperature based on the same EDA as used for the TOVSCV approach (Massart et al., 2020). Differently from the TOVSCV approach we use these flow-dependent background error standard deviations for the skin temperature not only over the land, but also over the ocean and sea-ice.

We demonstrated in (Massart et al., 2020) that the introduction of background error correlation in the SKTACV approach results in an effective background error standard deviation larger than the specified standard deviation. For that reason, the values of the standard deviation are not inflated over land as for the TOVSCV approach.

Over the sea-ice, the value of standard deviation derived from the EDA is of the order of $0.5\,\mathrm{K}$. In the meantime, the TOVSCV approach has a fixed value of $7.5\,\mathrm{K}$ which accounts for the large uncertainty of the skin temperature over the sea-ice. Preliminary results showed that the consequence of this large difference in standard deviation was much larger skin temperature increments in the TOVSCV experiment than in the SKTACV experiment.

Insufficient variance in the EDA for skin temperature over the sea-ice is likely related to unaccounted sources of uncertainty in the sea-ice model. Specific perturbations over the sea-ice in the EDA are available in the IFS but not activated in the operational EDA. Recent EDA experiments with these perturbations activated were carried out in a configuration similar to the one of this paper (Philip Browne, ECMWF, personal communication). Preliminary results show that with such perturbations, there is an increase of the skin temperature background error standard deviation by a factor of about 3. We therefore decided to artificially scale the skin temperature background error over the sea-ice by applying a factor 3 to the standard deviation derived from the EDA.

3.2 Spatial and temporal correlations

The implementation of the spatial and temporal correlations follows the wavelet formulation described in Fisher and Andersson (2001). The main difference is that the vertical covariance matrices are replaced by temporal covariance matrices. For a given wavelet index, a temporal covariance matrix is defined on every point of a horizontal grid associated with this wavelet. The spatial covariance are generated by making the temporal covariance matrix dependent on the wavelet index.

Building such a covariance model requires the estimation of the local spatial correlation length-scale and the estimation of the temporal correlation. Both are computed for every assimilation cycle with the hourly fields of the model skin temperature short-range forecast from the 50 members EDA which makes them flow-dependent. For each time and each grid point, we compute the local time-scale of a gaussian function that best fits the time correlation computed at this time with all the fields at different times. The local spatial correlation length-scale is computed for each time with the HybridDiag software (Ménétrier and Auligné, 2015).

The diagnosed horizontal correlation length-scale was found to be of the order of 100 km over the land and sea-ice and of the order of 300 km over the ocean. The data thinning resolution being around 125 Km, we should expect the biggest impact of the introduction of horizontal correlation in the SKTACV approach over the oceans.

The diagnosed correlation time-scale was found to be of the order of mostly over 24 h over the ocean and between 2 h and 12 h over the land and sea-ice. We can then expect to resolve the diurnal cycle over the land and to have a smooth field in time over the ocean.

The local correlation time-scale and spatial correlation length-scale are diagnosed from the ensemble of model skin temperature short-range forecast. We are using them for both the microwave and infrared skin temperature fields as we do not have yet separate short-range forecast of these fields in the EDA.

4 Results

4.1 Experiments

To assess the differences between the current approach for the skin temperature analysis in the clear-sky radiance observation operator (TOVSCV) and the proposed approach SKTACV, we ran two parallel experiments, one for each approach. For both experiments, we have a northern hemisphere summer period: July to September 2019, referred to as JAS 2019; and a northern hemisphere winter period: January to March 2020, referred to as JFM 2020. Each period is preceded by a spin-up of a few weeks. These experiments are based on the IFS cycle CY47R1 and RTTOV version 12.2 (Saunders et al., 2018).

4.1.1 Clear-sky radiance observations

The instruments providing clear-sky radiance observations and assimilated in the IFS cycle CY47R1 are presented in Tab 1. We discussed in the introduction that we have a careful selection of the channels to avoid those which are very sensitive to the surface. For the geostationary instruments, we use only the water vapour channels, which are rather weakly sensitive to the surface. The data over a grid cell where the model orography is higher than 1.5 km are also rejected. For the other infrared instruments, a situation dependent screening is applied to identify channels strongly sensitive to the surface emission over land, and these channels are blacklisted. We do use near-window channels sensitive to the surface over the ocean but not over sea-ice. Nonetheless, these are used only when the relevant channels are cloud free which excludes a large number of them.

A detailed usage of surface-sensitive microwave channels is presented in Table 1 of Bormann et al. (2017). For this study, we focus on the AMSU-A instruments which have channels near the 50 GHz oxygen absorption line, and on the MWHS instruments which have channels near the 183 GHz water vapour absorption line. The ATMS instruments have the channels for both frequencies.

There are more channels with stronger sensitivity to the surface from the infrared sounders. On the other hand, there are more microwave sounders and there is no rejection due to cloud-contamination for those. All this affects the number of observations sensitive to the surface for both type of sounders.

Table 1. List of instruments used in clear-sky radiance assimilation.

Category	Instrument	Payload
Microwave	AMSU-A	AQUA
		METOP-A
		METOP-B
		NOAA 15
		NOAA 18
		NOAA 19
	ATMS	NOAA 20
		NPP
	MWHS	FY-3B
Geostationary	IMAGER	GOES 15
	ABI	GOES 16
	AHI	Himawari 8
	SEVIRI	METEOSAT 8
	SEVIRI	METEOSAT 11
Hyper-spectral	IASI	METOP-A
		METOP-B
	AIRS	AQUA
	CRIS	NOAA 20

4.1.2 Resolution

The incremental formulation of the IFS 4D-Var means that each inner loop has a different resolution. Here, the outer loop has a TCo 399 horizontal grid (or spatial resolution of about 25 km, Malardel et al., 2016) and a vertical grid with 137 vertical

350 levels. The inner loop horizontal grids are successively TL 95 / TL 159 / TL 255 (or spatial resolution of about 200 km, 125 km and 80 km respectively), with the same vertical grid as the outer loop.

4.2 Skin temperature analysis

The key difference between the two approaches is the skin temperature analysis. We thus start by inspecting the skin temperature analysis in both experiments. First, we focus on the two-dimensional fields provided by the SKTACV experiment. Then,

355 we compare the skin temperature analyses from the two experiments in the observation space, which is the space where the TOVSCV experiment provides its skin temperature analysis.
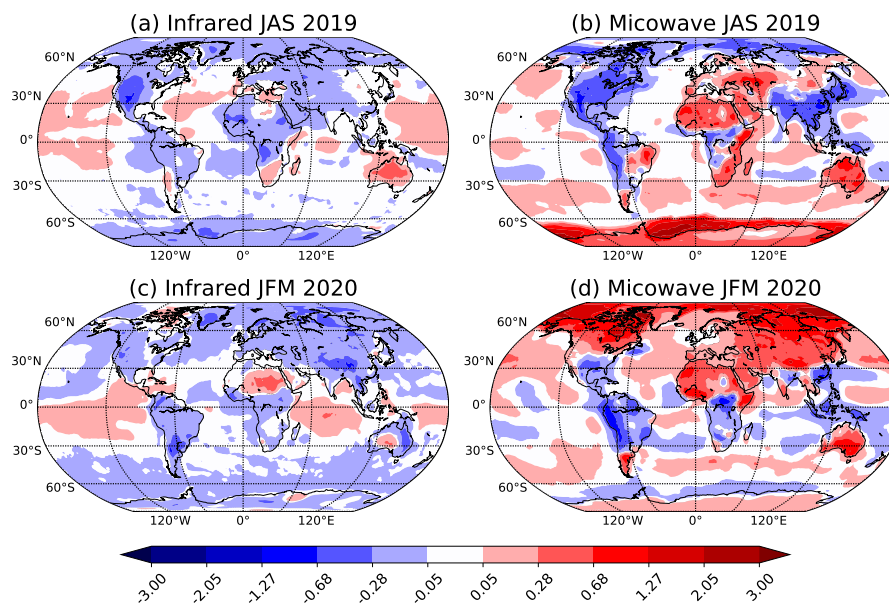
Figure 1. Time average of the skin temperature increment (in K) from the SKTACV experiment: (top): JAS 2019, (bottom): JFM 2020, (left): infrared field, (right): microwave field.

4.2.1 Model space

The SKTACV approach starts from the model skin temperature as a background. The 4D-Var produces the analysis increment which is the optimal adjustment to the background and the new batch of observations. The temporal mean of the analysis

360 increment is a proxy of the background bias. We computed the temporal mean of the analysis increment for the two periods and the two skin temperature fields for the SKTACV experiment (Fig. 1).

For the infrared field, the mean value is mainly under 0.3 K in absolute value. The difference between the two seasons is minimal. It appears that the model skin temperature provides a relatively unbiased background for the infrared skin temperature field. As discussed in section 4.1.1, most of the actively assimilated infrared radiances have relatively weak surface-sensitivity

365 over land and sea-ice. This limits the increments in these regions.

For the microwave field, there is a contrast between the ocean on one hand and the land and sea-ice on the other hand. Over the ocean, the absolute value of the mean is mostly under 0.3 K. Elsewhere, where the SKTACV approach makes use of surface sensitive microwave channels, the mean can reach absolute values of more than 3 K. Moreover, there are large differences between the two seasons. For example, there is a change of sign in the mean over sea-ice in Antarctica or over

370 Canada. The larger differences over land likely reflect that microwave frequencies can be sensitive to deeper sub-surface layers, and this is not captured by using the model skin temperature as background field. In addition, the specular assumption currently used in the radiative-transfer calculations for microwave frequencies has been found to be sub-optimal over snow and sea-ice areas, whereas diffuse, Lambertian reflection may be more appropriate (Karbou et al., 2006; Bormann et al., 2017). The
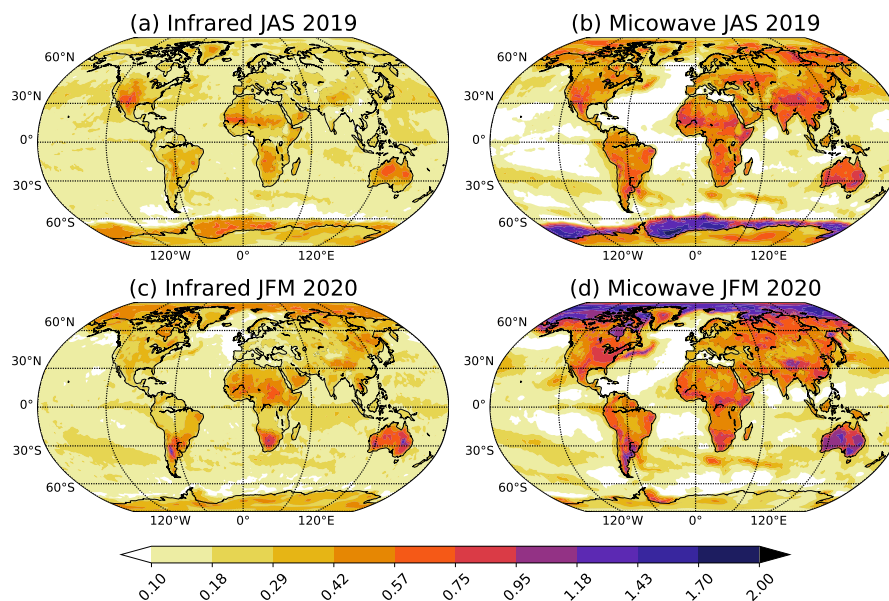
Figure 2. Standard deviation of the skin temperature increment (in K) from the SKTACV experiment: (top): JAS 2019, (bottom): JFM 2020, (left): infrared field, (right): microwave field.

specular assumption leads to viewing-angle dependent biases, and these are partly compensated for through increments in the

375 skin temperature. This likely contributes to the larger mean increments over Canada and Northern Asia in winter.

From the time series of the analysis increment, we also computed the temporal standard deviation for the two periods and the two skin temperature fields (Fig. 2). This quantity is a proxy of the analysis activity, the larger the standard deviation, the larger the activity. The pattern of the standard deviation is similar for the two fields: the lowest values are found over the ocean (mostly under 0.3 K) and the largest values over the sea-ice (up to 2 K). This is expected as the standard deviation of the

380 background errors is lower over the ocean and the uncertainty is large over the sea-ice.

Over the ocean, the analysis is slightly more active for the infrared field than for the for microwave field. For both, the activity is larger inside the subtropical gyres limited by the ocean currents, and the activity is larger in the region of the Gulf stream current along the east American coast.

Over the land and sea-ice, the analysis is less active for the infrared field than for the microwave field as expected from the

385 blacklisting of the surface sensitive channels of the infrared radiances. Still, the standard deviation is higher than over ocean which means that there is still some surface sensitivity in the remaining infrared radiances. For the microwave field, the activity is the largest over sea-ice. Over the land, the activity is larger over desert regions where the diurnal cycle of the skin temperature can be strong and around regions with high orography.
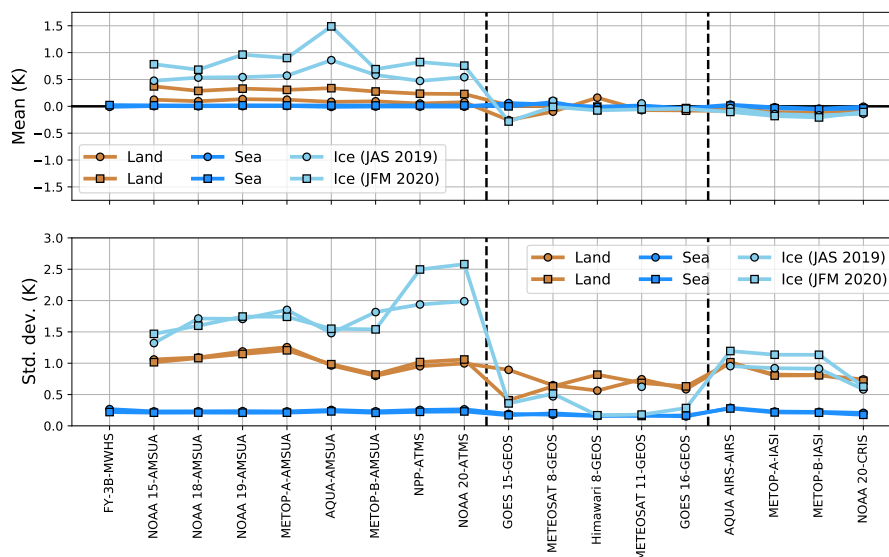
Figure 3. Mean (top) and standard deviation (bottom) of the skin temperature analysis difference between the SKTACV and the TOVSCV experiments for each instrument and for the two seasons: JAS 2019 (circle) and JFM 2020 (square). The difference is computed over the ocean, sea-ice and land (see legend). Left part of the plot: microwave instruments. Middle part: infrared instruments (geostationary). Right part: infrared instruments (polar orbiting).

4.2.2 Observation space

390 For each field of view of the radiance observations, the background and analysis values of the skin temperature are stored in the Observational DataBase (ODB, Fouilloux, 2009). For both experiments, we can extract the matching skin temperature analysis and compute the difference (SKTACV-TOVSCV) in observation space (time and location of each field of view). Figure 3 presents the mean and standard deviation of the difference by instrument and for the two seasons. The statistics are computed over three surface properties (land, ocean, sea-ice) as we saw previously that there are large differences between them.

395 Note that due to instrument problems, the three lowest-peaking channels (5-7) of AQUA AMSU-A were not used over the sea-ice. This reduces the constraint of this instrument on its skin temperature analysis in the TOVSCV approach and reduces to zero the constraint over the sea-ice.

Over the ocean, for all instruments, the mean difference is close to zero and the standard deviation is around 0.25 K which is of the same order as the mean background error standard deviation. Over the land, the mean difference is also close to zero

400 except for the microwave instruments during the JFM 2020 season, the standard deviation is between 0.5 K and 1 K, that is again of the same order as the mean background error standard deviation. This means that both approaches are similar in terms of skin temperature analysis over the land and ocean, except for microwave instruments during the JFM 2020 season over land. This difference originates largely from areas with large analysis increments over the Northern Hemisphere as discussed earlier, for which locally values can reach up to 3 K.

405   As the usage of the surface sensitive channels for the infrared instruments is limited over the sea-ice, there is not much difference between the two approaches. The standard deviation of the difference is about 1 K which is lower than the background error standard deviation. There is a bigger difference between the two approaches for the microwave instruments over the sea-ice. First, the mean difference is between 0.5 K and 1 K. This difference comes from the SKTACV experiment for which the mean increment is positive while the mean increment has values close to 0. K for the TOVSCV experiment (not shown).

410   Even if the SKTACV skin temperature analysis is quite active over the sea-ice as discussed previously, the large values of standard deviation are due to large variability of the increments in the TOVSCV experiment. This experiment has indeed a more active skin temperature analysis due to the large background error standard deviation of 7 K as described previously.

4.2.3   Outliers

With the exception of sea-ice, we can conclude that the two approaches are similar on average. We believe that the SKTACV
415 approach could be beneficial for particular situations, because the skin temperature is better constrained through several instruments at the same time, and via the spatial and temporal correlations of the background error. The TOVSCV experiment in contrast allows other errors to be aliased into skin-temperature increments. Thus, we investigated the cases for which the skin temperature analysis is significantly different between the SKTACV and TOVSCV experiments, as this could point to situations where the skin temperature in TOVSCV compensates for other large errors. We searched for the outliers defined
420 as scenes where the skin temperature analysis difference is larger than a threshold t and simultaneously the skin temperature background difference is larger than $t/2$.

    We chose the threshold to be equal to 3 times the local value of the skin temperature background error standard deviation from the TOVSCV experiment. The threshold is then situation dependent over land but constant over the ocean and sea-ice with respective values of 3 K and 21 K.

425   For the infrared instruments, we did not find any outlier with this criteria. For the microwave instruments, all the outliers are found in the northern hemisphere and most of them over the ocean or near the coastline (Fig. 4). This is linked to the usage of the background error standard deviation from the EDA over the ocean in the SKTACV experiment which allows the skin temperature increments to be larger. For instance, for the outliers located in the East coast of United States of America, the sea surface uncertainty is higher than 1 K and associated with the Gulf Stream. The uncertainty is also large in the Arctic
430 region. Looking at these outliers in more details, the larger increment in the SKTACV experiment allows the analysis to better fit channel 6 of ATMS on average (not shown).

    If we change the threshold to a lower value, we start to have outliers over the sea-ice. There, the skin temperature increment is always larger in the TOVSCV experiment, as expected from the difference in the value of the background error standard deviation. We also start to have outliers over land in the regions where the skin temperature is very sensitive to the meteorolog-
435 ical conditions. Small differences in the weather parameters between the two experiments can cause differences in model skin temperature close to or larger than the threshold in these regions. Then, the assimilation processes of both experiments were usually not able to change enough the skin temperature to bring their analyses close to each other.
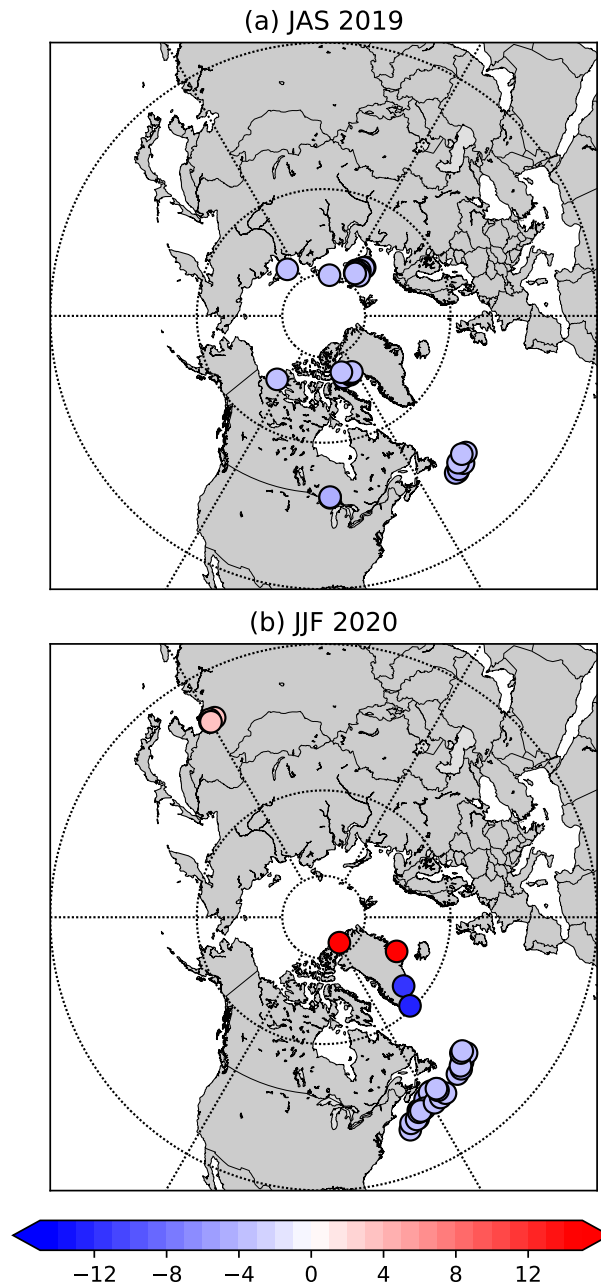
Figure 4. Difference of the skin temperature analysis in observation space (in K) between the experiment and the control when the difference exceeds 3 background error standard deviation and for NOAA-20-ATMS. Bluish colours: SKTACV values lower than TOVSCV values. Reddish colours: SKTACV values greater than TOVSCV values. Periods: (a) JAS 2019, (b) JFM 2020.
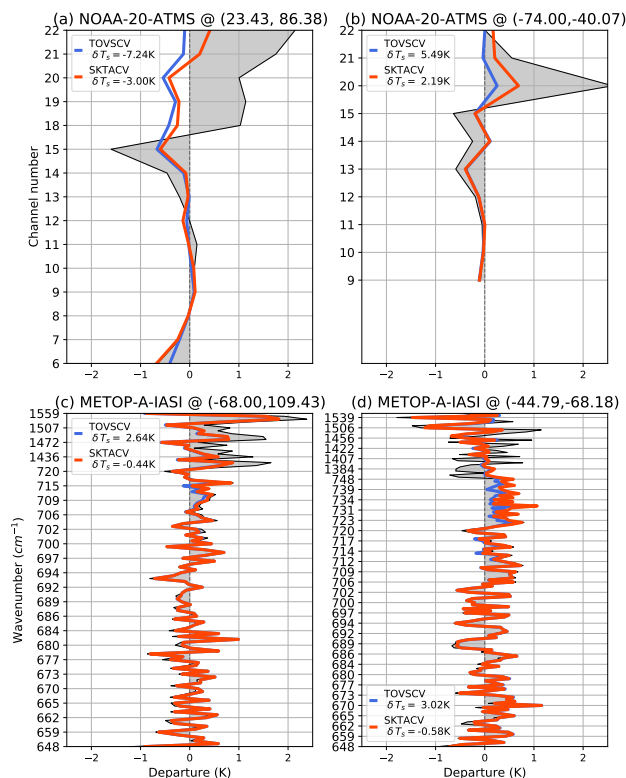
Figure 5. First guess (grey shade) and analysis (coloured lines) departures (in K) for ATMS (top) and IASI (bottom) for four particular scenes (see plot titles for the location (latitude, longitude) of the scenes). In blue: departures form the TOVSCV experiment. In red: departures form the SKTACV experiment. In the legend, δT_s is the value of the skin temperature increment.

In conclusion, we did not find any scenes for which the TOVSCV experiment seems to excessively boost the skin temperature increment to compensate for other errors.

4.2.4   Cases study

To illustrate the difference in behaviour between the two approaches, we detail here four particular scenes, two for the ATMS instrument (microwave) and two for the IASI instrument (infrared). For each scene, we present the first guess and analysis departures in radiance space (Fig. 5).

The scenes are from the very first assimilation cycle of the summer period. The background values for the atmospheric variables and for the skin temperature are the same for the two approaches. This facilitates the interpretation of the differences between them. Nonetheless, one has to be careful not to attribute the observed differences in the analysis departures only to the skin temperature analysis at the location of the scene. For a given scene, a difference in the skin temperature analysis in the

surrounding model grid cells creates a difference in the atmospheric variables in those grid cells and subsequently in the grid cell of the scene due to the spatial covariances of the atmospheric variables background error.

450   For the first ATMS scene which is over land, the TOVSCV experiment produces a large skin temperature increment (-7.24 K), larger than the one from the SKTACV experiment (-3. K). This allows the TOVSCV analysis to better fit the measurements for the surface-sensitive channel 6 than the first guess and than the SKTACV analysis, but results in a worse fit for the channels 18 to 20, compared to the SKTACV analysis. For the second scene which is over the sea-ice, none of the lower-tropospheric temperature-sounding channels are assimilated. Then the TOVSCV experiment produces a large skin temperature increment

455   (5.5 K) that allows its analysis to better fit the measurements for the channels 20 to 22 compared to the SKTACV analysis for which the skin temperature increment is smaller (2.2 K).

These two ATMS cases illustrate how, for the TOVSCV experiment, the increment in the skin temperature is largely driven by aiming to fit the most surface-sensitive channels present for a given field-of-view, at times leading to relatively large increments. If present, the lower-tropospheric temperature-sounding channels will be the more impacted as, for example, the prescribed

460   observation error is more than 5 times lower for channel 6 compared to channels 18 to 22. In contrast, in the SKTACV experiment, the skin temperature is also constrained by adjacent observations via the skin temperature background errors, reducing the strong response to one particular observation.

The first IASI scene is over Antarctica, while the second scene is over land. For both scenes, the skin temperature increments from the SKTACV analysis are small (respectively -0.4 K and -0.6 K). The increments from the TOVSCV analysis are larger

465   (respectively 2.6 K and 3. K), which improves the analysis-fit to observation for the wavenumbers between about $710 \, \text{cm}^{-1}$ and $750 \, \text{cm}^{-1}$ corresponding to lower tropospheric sensitive channels. For the first scene, despite the large skin temperature increment, the analysis-fit is only marginally improved, which may suggest that the skin temperature adjustment compensates for other errors.

4.3   Meteorological analysis

470   On average the skin temperature analysis is similar for the two approaches. We now assess how the two analyses compare for other variables.

4.3.1   Analysis-fit to observation

We define the analysis-fit to observations as the standard deviation of the SKTACV analysis departure (analysis minus observation) normalised by the standard deviation of the TOVSCV analysis departure. If the analysis-fit to observations is lower than

475   100 % , the SKTACV analysis is labelled as closer to the observations. Otherwise the TOVSCV analysis is labelled as closer to the observations.

We present first the analysis-fit to the radiance observations. By design, the skin temperature analysis impacts the radiance observation operator. This impact is detectable in the analysis projected in the radiance observation space through the observation operator. We present only the results for ATMS and IASI because these two examples are representative of the other

480   microwave and infrared instruments.
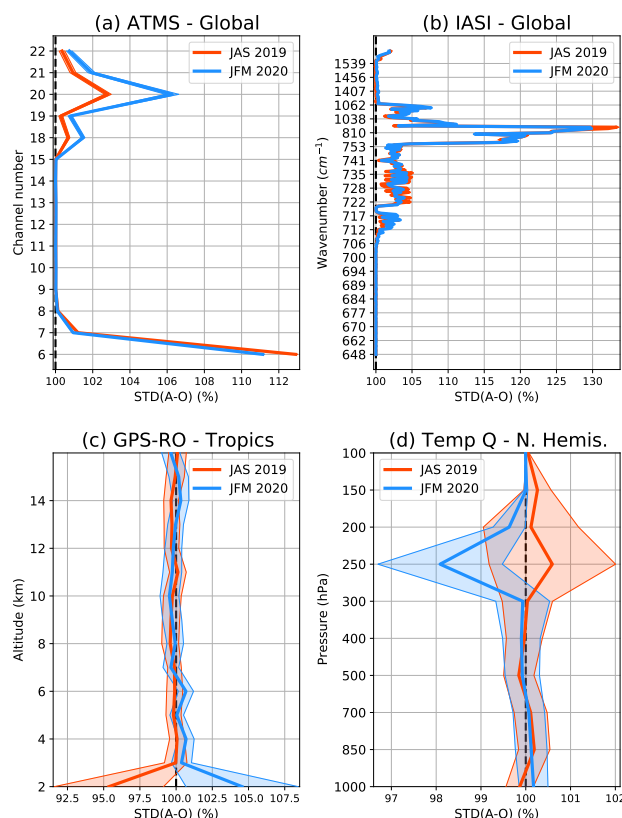
Figure 6. Normalised analysis-fit to observation in %), for various observation type (see sub-caption) and for JAS 2019 (red) and JFM 2020 (blue). The shaded area represents the variability (one standard deviation).

For ATMS, the analysis-fit to observation is increased in the SKTACV experiment compared to the TOVSCV experiment for two sets of channels (Fig. 6(a)). The increase is larger by 10% for channel 6 and by up to 6 % for channels 18 to 22 depending on the season. This is likely linked with the mechanism described in the case study. The TOVSCV experiment uses the skin temperature to better fit the measurements for channel 6 when assimilated and for channels 18 to 22 otherwise.

For IASI, the difference between the two experiments is larger than for ATMS for the channels sensitive to the surface. For the window channels around $810 \, \mathrm{cm}^{-1}$, the analysis-fit of the SKTACV experiment is about 30 % larger (Fig. 6(b)). The fit is also larger but by less than 5 % for the lower tropospheric sensitive channels (around 700-800 cm^{-1}). This is in accordance with the case studied before: the TOVSCV experiment uses the skin temperature to better fit the measurements for these channels. For the geostationary infrared, the analysis-fit to observation is also slightly larger for the SKTACV experiment, by around 0.25 % for JAS 2019 and around 0.5 % for JFM 2020 (not shown).

From these results alone it is not possible to say which of the two systems performs better. The findings may indicate that the TOVSCV experiment uses the skin-temperature to over-fit the radiance observations, either due to too large assumed background errors or possibly by aliasing scene-dependent errors with suitable spectral signatures into skin-temperature increments
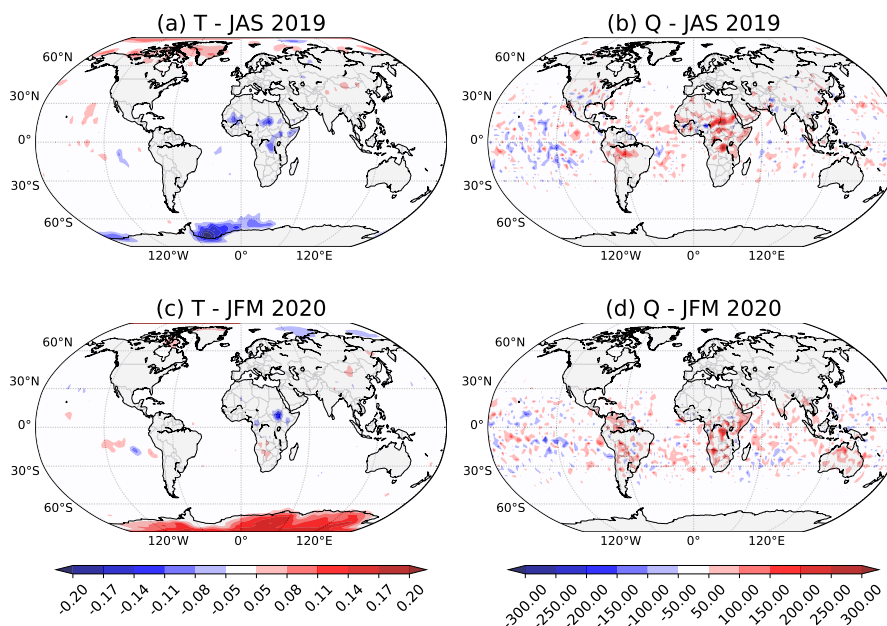
Figure 7. Time average difference between the SKTACV and the TOVSCV analyses at 850 hPa: (top): JAS 2019, (bottom): JFM 2020, (left): temperature (in K), (right): specific humidity (in $\mu g \cdot g^{-1}$).

(e.g., cloud-screening errors or emissivity/specularity errors). On the other hand, the results may indicate that the skin tem-
495 perature is too strongly constrained in the SKTACV experiment which prevents the analysis to fit the radiance observations more closely. Either way, the temperature and humidity analyses are expected to be different from the surface, up to the middle troposphere. This can be assessed with the analysis-fit to other observations.

We found that the SKTACV temperature analysis is generally slightly closer (by less than 0.5 %) to the GPS-RO, sondes and aircraft data, for JAS 2019. For JFM 2020, the results are more mixed. The largest difference was detected for the fit to the GPS-
500 RO data, in the lower troposphere and over the tropics, where there is a contrasting behaviour between the two seasons. The SKTACV temperature analysis has a closer fit by about 5 % in JAS 2019 and a larger fit by about 5 % in JFM 2020 (Fig. 6(c)). It is not clear where this difference comes from, as no such differences were found in the analysis-fit to other observations in the tropics.

The difference in humidity analysis-fit is mostly not statistically significant when compared to the sondes and aircraft data.
505 The humidity analysis-fit values are within ± 0.5 %. The exception is for the analysis-fit to sondes data in the upper tropo-sphere (around 250hPa) for JFM 2020 where the SKTACV humidity analysis is significantly closer to the observation by 2 % (Fig. 6(d)). This is the region where the ATMS channels 18 to 22 are sensitive to humidity and where the TOVSCV analysis is closer to the ATMS observations. This suggests that the TOVSCV approach uses the skin temperature to compensate for other errors in the northern hemisphere for JFM 2020, allowing the analysis to be closer to ATMS channels 18 to 22 but in the
510 meantime putting the humidity analysis further away from the radiosondes data.

4.3.2 Mean analysis difference

We saw that the analyses from the two approaches differ when compared to observations. Here, we compare directly the two analyses in physical space. We found that the mean atmospheric states are very similar in the two analyses and the few differences appear in the lower atmosphere only. There, the mean temperature analyses differ mainly by less than 0.05 K and 515 the mean humidity analyses by less than 50 μg.g⁻¹.

The largest differences for temperature are found in the polar regions. For example, at 850 hPa, when compared to the TOVSCV experiment, the SKTACV experiment is warmer in the Arctic region and cooler in the Antarctic region in JAS 2019 and the opposite in JFM 2020 (Figs. 7(a) and (b)).

We found previously that, over the Antarctic region, the mean increment for the microwave instruments is positive and large 520 in the winter season (JAS 2019) especially in the Weddell sea. There, the analysis tends to warm the atmosphere at 850 hPa. By increasing the skin temperature for the microwave instruments, the SKTACV analysis reduces the warming of the atmosphere compared to the TOVSCV analysis. This is a very limited area and it is difficult to assess which analysis is better.

Still over the Antarctic region but for the summer season (JFM 2020), the mean increment for the microwave instruments is close to zero and the skin temperature analysis is not very active compare to the one from the TOVSCV experiment. Both 525 analyses tends to warm the atmosphere there. By having a more active skin temperature analysis the TOVSCV experiment reduces the warming of the atmosphere compared to the SKTACV analysis. This difference makes the SKTACV temperature analysis slightly closer to the radiosonde measurements than the TOVSCV one (not shown).

There are also some differences locally in the mean temperature analysis in the tropical region of Africa. This could be linked to errors in the emissivity that are propagated to the skin temperature in the TOVSCV experiment. There is also a moistening 530 of the tropics on average in the SKTACV experiment compared to the TOVSCV experiment (Figs. 7(c) and (d)). This makes also the SKTACV humidity analysis slightly closer to the radiosonde measurements than the TOVSCV one (not shown).

The discrepancy in the temporal standard deviation of the analysis between the two experiments informs us of the activity of the analysis. For the temperature, we found that the difference is negligible with maximum values of 0.05 K detected in the polar regions. For the humidity, the difference is also small. The largest difference is found in the tropics, close to the surface, 535 where the SKTACV experiment has a lower standard deviation by 20 μg.g⁻¹ maximum.

To summarise, the two analyses are very similar on average, without an important difference between the two seasons. The main differences are located in the polar regions where the temperature in the lowest model levels is warmer during the summer season in the SKTACV analysis. There are other differences in the tropics, close to the surface too, where the SKTACV analysis is slightly moister and slightly less active.

4.4 Short range forecast

The atmospheric analysis serves as an initial condition for the short-range forecast that is used as the first guess for the next assimilation cycle. This first guess is compared to all available observations to compute the first guess departure. The standard deviation of the first guess departure gives information on the quality of the first guess. We compute here the normalised first-
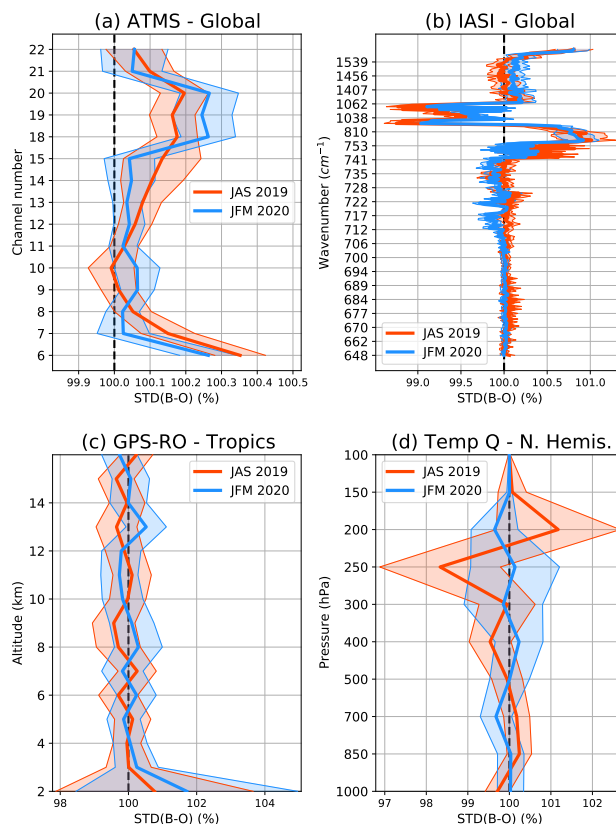
Figure 8. Normalised first-guess-fit to observation in %), for various observation type (see sub-caption) and for JAS 2019 (red) and JFM 2020 (blue). The shaded area represents the variability (one standard deviation).

guess-fit to observation which is the standard deviation of the first guess departure of the SKTACV experiment normalised

545 by standard deviation of the first guess departure from the TOVSCV experiment . For a value lower (greater) than 100 %, the SKTACV first-guess is labelled as better (worse) compared to the relevant observations.

For the ATMS instrument the SKTACV first-guess is slightly worse by up to 0.3 % for the channels sensitive to the surface (Fig. 8(a)). The only exception is for the northern hemisphere in JFM 2020 for which the first-guess is better by 0.1 % (not shown). Combined with the finding that the SKTACV analysis-fit to ATMS observations is larger than the TOVSCV analysis-

550 fit, the result suggests that, on average, the skin temperature may be too strongly constrained in SKTACV for the microwave instruments.

For the IASI instrument, there is a dipole in the normalised first guess departure (Fig. 8(b)). The first guess is better for the ozone channels by up to 1 % and worse for the window channels by up to 1 % too. This phenomenon has already been observed and is under investigation (Cristina Lupu, ECMWF, personal communication). Here, we can only conclude that even
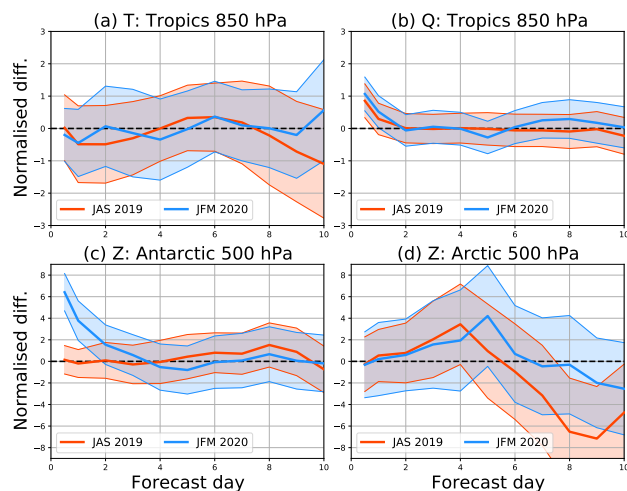
Figure 9. Normalised difference in root mean square error (full line, in %) between the SKTACV and the TOVSCV experiments, when compared to the operational analyses, and associated uncertainty (coloured shape). Red: JAS 2019. Blue: JFM 2020. See the titles for the parameter and region plotted.

if the SKTACV approach largely affects the analysis-fit to hyper spectral infrared measurements, it is neutral when it comes to the first guess fit.

When compared to other instruments, the difference in the standard deviation of the first guess fit is mostly not statistically significant as illustrated with the GPS-RO data in Fig. 8(c). The only statistically significant difference comes from the comparison against humidity measurements from radiosondes at $250\,\mathrm{hPa}$ in the Northern Hemisphere where the SKTACV experiment is better for JAS 2019 (Fig. 8(d)). This is the region where we saw a better fit for the SKTACV experiment in the analysis-fit to humidity, but for JFM 2020.

4.5 Forecast scores

For assessing the forecast quality at up to day 10, we use here the operational analysis as the reference to compute the forecast error. The operational analysis is assumed to be our best knowledge of the global atmospheric state at a given time. Because the operational analysis has a TCo 1279 horizontal grid (spatial resolution of about $9\,\mathrm{km}$), we do not expect it to favour the TOVSCV experiment even if it uses the TOVSCV approach.

The metric we are using is the normalised difference of the root mean square error (RMSE) of the forecast error for forecast lead times between 1 and 10 days (Geer, 2016). If the normalised difference is negative (positive) for a given lead time, the forecast from the SKTACV experiment is labelled as better (worse) than the one from the TOVSCV experiment for this lead time. If the $0\,\%$ line is outside the uncertainty range, the result is statistically significant. Otherwise it is not. These scores are computed by region and for each parameter.

We focus here on the Tropics and on the polar regions, below $60°$S and above $60°$N, where the largest differences in the analysis were observed. In the tropics at $850\,\mathrm{hPa}$, the normalised difference of the RMSE is within $0.5\,\%$ for humidity, and within $1\,\%$ for temperature, for all forecast ranges and without any statically significant signal (Figs. 9(a) and (b)).

575 For the polar regions, we discuss the normalised difference of the RMSE for the geopotential (Z) at $500\,\mathrm{hPa}$ as it summarises the quality of the forecast from the surface up to $500\,\mathrm{hPa}$ (Figs.9(c) and (d)). The first statistically significant difference is in Antarctica, in winter and up to day 3. The forecast error is up to $6\,\%$ larger for the SKTACV experiment. This is also confirmed by the verification against observations but then the increase in error is between $1\,\%$ and $2\,\%$. This increase comes primarily from a larger mean error in the SKTACV experiment linked with a sightly warmer analysis as discussed previously. The

580 second statistically significant difference is over the Arctic where the SKTACV forecast is better by $3\,\%$ to $5\,\%$ after day 7. This improvement is confirmed by the comparison with observations, but without statistical significance.

5 Conclusions

The vertical sensitivity of a radiance observation to the various layers of the atmosphere depends on the frequency of the measurement. Assimilating radiance observations in the IFS at various frequencies provides information on the atmospheric

585 variables on diverse atmospheric layers. Some of the assimilated frequencies make the clear-sky radiance sensitive to the surface variables, and to the skin temperature in particular. In that case, during the assimilation process, the skin temperature is adjusted together with all the atmospheric variables in order to produce a simulated radiance that best fits the observed radiance. This is achieved by extending the atmospheric 4D-Var control vector with a set of skin temperature values defined in observation space for each field of view. This approach is known as TOVS control variable (or TOVSCV).

590 Note that in the work presented here, the skin temperature refers to the skin temperature from the 4D-Var control vector and that is used in the radiance observation operator. The background values for the skin temperature are derived from the model skin temperature. Afterwards, the retrieved skin temperature and the model skin temperature are disconnected.

The TOVSCV approach is sub-optimal because without spatial or temporal constraints, each individual skin temperature value is adjusted independently. The adjustment may alias real atmospheric information into skin temperature due, for example,

595 to issues with the cloud screening. This could be improved by constraining the skin temperature with all available radiance observations as it is the case for the atmospheric variables. One solution would be to define spatial and temporal correlations in the background error for the skin temperature. These correlations would need to be constructed in observation space which is a complex task technically challenging. Instead, we proposed in this paper a new approach to the optimisation of the skin temperature that allows defining these spatial and temporal correlations in model space similarly to all other variables of the

600 analysis state vector.

For this new approach, referred to as SKTACV, a set of two-dimensional skin temperature fields defined in model space are added to the analysis control vector in the IFS 4D-Var. These fields are divided into two categories corresponding to two separate spectral bands (microwave and infrared). For each category, we also introduce hourly fields to account for the time variation of the skin temperature within the assimilation window.

605   Some surfaces have a greater uncertainty than others due to the nature of the terrain and the uncertainty depends on the time of day that the satellite overpasses. Hence, it is important to communicate the appropriate skin temperature background errors which control the adjustment of the surface by the satellite for both approaches. This is achieved in the IFS by using the ensemble of background model skin temperature from a 50 member EDA. In the TOVSCV approach, the flow-dependent background errors are used over land only. In the SKTACV approach, they are used on all surface. Moreover, the members of

610   the EDA are used to compute the spatial and temporal correlations of the skin temperature background errors which add some constraint on the skin temperature fields, that is not present in the current TOVSCV approach.

At the moment, the EDA does not have explicit perturbations over sea-ice. This makes the background error standard deviation of skin temperature too small in the regions covered with ice. Preliminary EDA experimentations adding specific perturbations over sea-ice suggested multiplying the standard deviation by a factor 3 there. We used this scaling factor over

615   sea-ice. Otherwise, every other background error component is computed directly from the EDA.

We performed two sets of experiments, one for each approach (TOVSCV and SKTACV). We ran them over a 3 months period in northern hemisphere summer (JAS 2019) and a 3 months period in northern hemisphere winter (JFM 2020). As expected, by constraining the skin-temperature analysis more strongly, the SKTACV analysis does not fit the radiance observations as much as the TOVSCV analysis for the channels sensitive to the surface and to the lower troposphere. Compared to other independent

620   observations, there is no statistically significant difference between the analyses.

Apart from over the sea-ice, the two estimates of skin temperature overall agree fairly well, with differences mostly below the errors assigned to the background skin temperature values. This may suggest that aliasing of other errors in the skin temperature errors is not a major issue, and the screening applied to the affected radiances performs well in this respect. Instances with larger differences between the skin temperature estimates were nevertheless found, and these could mostly be attributed to the larger

625   skin temperature background error in the SKTACV experiment over the ocean.

The difference in the treatment of the skin temperature does not impact much the analysis on average. The main difference is a slightly moister analysis in the tropics for the SKTACV experiment, and small changes of the temperature in the polar regions, with a sign depending on the season. The differences in the polar regions come mainly from the microwave instruments for which the skin temperature analysis over sea-ice is on average warmer by 0.5 to 1 K depending on the season and instrument.

630   Moreover, there is much more variability in the skin temperature analysis in the TOVSCV experiment than in the SKTACV one, which is a consequence of the high value of background error over sea-ice in the TOVSCV experiment.

Outside the polar regions, the quality of short-range forecasts for the two approaches is mostly similar, with first-guess fits for a wide range of observations showing no statistically significant change. However, the first guess fit to the microwave data is globally larger by 0.1 to 0.3 % in the SKTACV experiment, with the highest differences found in the summer hemisphere.

635   On the other hand, there is a small signal of an improvement of the humidity at around 250 hPa. This is, for example, around where the ATMS channels 18 to 22 have some sensibility. These are also channels for which the SKTACV experiment is worse.

The sensitivity of the microwave instruments to the surface is indeed deeper than the layer represented by the model skin temperature. Therefore, using the model skin temperature as a background is not optimal. This is confirmed by the time average

of the analysis increment of the microwave fields which has values up to 3 K. We are currently investigating how to use the
640 microwave skin temperature analysis from this study to have a better background and address the bias issue.

Changing the background value of skin temperature, will change the first-guess in radiance space and may change the usage
of the microwave observations. For example, our quality control tends to reject a significant portion of microwave data over
deserts or snow-covered regions because of too large biases (Lawrence et al., 2019). If we can improve our skin temperature
background field, we may be able to reduce these biases and increase the number of assimilated radiance. A better skin-
645 temperature background should also lead to better dynamic emissivity estimates, leading to further improvements.

We considered here two separate skin temperature fields. This can be further refined as different channels within a category
can have a different sensitivity to the surface. For instance, low-frequency channels are sensitive to a deeper surface layer than
high-frequency channels for the microwave instruments. For example, we could separate for ATMS the channels 6-9 (50 GHz)
to the channels 18-22 (183 GHz). Moreover, we may be able to add more channels that are currently blacklisted because they
650 were found to be problematic due to their sensitivity to the surface. For instance, for the infrared instrument, we could stop
removing channel strongly sensitive the surface over land and start using near-window channels sensitive to the surface over
the over sea-ice.

Finally, the availability of a temporally and spatially consistent analysis of skin temperature can also be seen as an attractive
by-product of the SKTACV method. These skin temperature analysis fields can drive further improvements in the estimation
655 of physically connected fields (e.g., sea surface temperature, land surface temperature, sea ice temperature).

Data availability.   The data from the TOVSCV experiment are available at https://doi.org/10.21957/2hd6-3h68 (JAS 2019) and https://doi.org/10.21957/bwf5-8929 (JFM 2020). The data from the SKTACV experiment are available at https://doi.org/10.21957/vv2s-7h56 (JAS 2019) and https://doi.org/10.21957/945j-a645 (JFM 2020).

Author contributions.   SM proposed and led this study, implemented the new approach and run the simulations. CL contributed to the design
660 of the background errors. All authors contributed to the interpretation of the results and to the writing process.

Competing interests.

The authors declare that they have no conflict of interest.

References

Bonavita, M., Isaksen, L., and Hólm, E.: On the use of EDA background error variances in the ECMWF 4D-Var, Q.J.R. Meteorol. Soc., 138,
665   1540–1559, 2012.

Bormann, N., Lupu, C., Geer, A. J., Lawrence, H., Weston, P., and English, S.: Assessment of the forecast impact of suface-sensitive microwave radiances over land and sea-ice, ECMWF Technical Memoranda, Technical Memorandum No. 804, https://doi.org/10.21957/qyh34roht, 2017.

Bormann, N., Lawrence, H., and Farnan, J.: Global observing system experiments in the ECMWF assimilation system, ECMWF Technical
670   Memoranda, Technical Memorandum No. 839, https://doi.org/10.21957/sr184iyz, 2019.

Browne, P., de Rosnay, P., Zuo, H., Bennett, A., and Dawson, A.: Weakly Coupled Ocean-Atmosphere Data Assimilation in the ECMWF NWP System, Remote Sensing, 11, 234, https://doi.org/10.3390/rs11030234, 2019.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q.J.R. Meteorol. Soc., 120, 1367–1388, 1994.

675   de Rosnay, P., Balsamo, G., Albergel, C., Muñoz Sabater, J., and Isaksen, L.: Initialisation of Land Surface Variables for Numerical Weather Prediction, Surv. Geophys., 35, 607–621, https://doi.org/10.1007/s10712-012-9207-x, 2014.

ECMWF: Part I : Observations, IFS Documentation, IFS Documentation CY46R1, 2019.

English, S. and Hewison, T.: A fast generic millimetre-wave emissivity model, Proceedings of SPIE - The International Society for Optical Engineering, 3503, https://doi.org/10.1117/12.319490, 1998.

680   English, S. J.: Estimation of Temperature and Humidity Profile Information from Microwave Radiances over Different Surface Types, Journal of Applied Meteorology, 38, 1526–1541, 1999.

English, S. J.: The Importance of Accurate Skin Temperature in Assimilating Radiances From Satellite Sounding Instruments, IEEE Transactions on Geoscience and Remote Sensing, 46, 403–408, https://doi.org/10.1109/TGRS.2007.902413, 2008.

Eyre, J., Kelly, G., McNally, A., and Andersson, E.: Assimilation of TOVS radiance information through one-dimensional variational analy-
685   sis, ECMWF Technical Memoranda, Technical Memorandum No. 187, https://doi.org/10.21957/72yxf2jsk, 1992.

Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman Filtering, ECMWF Technical Memoranda, Technical Memorandum No. 347, 2001.

Fouilloux, A.: ODB (Observational DataBase) and its usage at ECMWF, in: Twelfth Workshop on Meteorological Operational Systems, 2-6 November 2009, pp. 86–90, ECMWF, Shinfield Park, Reading, 2009.

690   Geer, A. J.: Significance of changes in medium-range forecast scores, Tellus A: Dynamic Meteorology and Oceanography, 68, 30 229, https://doi.org/10.3402/tellusa.v68.30229, 2016.

Gentemann, C. L., Minnett, P. J., Le Borgne, P., and Merchant, C. J.: Multi-satellite measurements of large diurnal warming events, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL035730, 2008.

Karbou, F., Gérard, E., and Rabier, F.: Microwave land emissivity and skin temperature for AMSU-A and -B assimilation over land, Q.J.R.
695   Meteorol. Soc., 132, 2333–2355, https://doi.org/10.1256/qj.05.216, 2006.

Kazumori, M. and English, S. J.: Use of the ocean surface wind direction signal in microwave radiance assimilation, Q.J.R. Meteorol. Soc., 141, 1354–1375, https://doi.org/10.1002/qj.2445, 2015.

Lawrence, H., Bormann, N., Sandu, I., Day, J., Farnan, J., and Bauer, P.: Use and impact of Arctic observations in the ECMWF Numerical Weather Prediction system, Q.J.R. Meteorol. Soc., 145, 3432–3454, https://doi.org/https://doi.org/10.1002/qj.3628, 2019.

700 Liu, C., Xiao, Q., and Wang, B.: An Ensemble-Based Four-Dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test, Monthly Weather Review, 136, 3363–3373, https://doi.org/10.1175/2008MWR2312.1, 2008.

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kuehnlein, C., Mozdzynski, G., Hamrud, M., and Smolarkiewicz, P.: A new grid for the IFS, ECMWF Newsletter, Meteorology, 146, 23–28, https://doi.org/10.21957/zwdu9u5i, 2016.

Massart, S.: A new hybrid formulation for the background error covariance in the IFS: Implementation aspects, ECMWF Technical Memoranda, Technical Memorandum No. 832, https://doi.org/10.21957/6gdjcd4j3, 2018.

705

Massart, S., Bormann, N., Bonavita, M., and Lupu, C.: Skin Temperature Analysis for the Assimilation of Clear-Sky Satellite Radiances, ECMWF Technical Memoranda, Technical Memorandum No. 870, https://doi.org/10.21957/goe0ads8z, 2020.

Matricardi, M., Chevallier, F., Kelly, G., and Thépaut, J.-N.: An improved general fast radiative transfer model for the assimilation of radiance observations, Q.J.R. Meteorol. Soc., 130, 153–173, https://doi.org/10.1256/qj.02.181, 2004.

710 McNally, A.: The impact of satellite data on NWP, in: Seminar on Use of Satellite Observations in Numerical Weather Prediction, 8-12 September 2014, ECMWF, Shinfield Park, Reading, 2015.

Ménétrier, B. and Auligné, T.: Optimized Localization and Hybridization to Filter Ensemble-Based Covariances, Mon. Weather. Rev., 143, 3931–3947, https://doi.org/10.1175/MWR-D-15-0057.1, 2015.

Pinker, R., Sun, D., Miller, M., and Robinson, G.: Diurnal cycle of land surface temperature in a desert encroachment zone as observed from

715 satellites, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL030186, 2007.

Prigent, C., Rossow, W. B., Matthews, E., and Marticorena, B.: Microwave radiometric signatures of different surface types in deserts, Journal of Geophysical Research: Atmospheres, 104, 12 147–12 158, https://doi.org/10.1029/1999JD900153, 1999.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C.: An update on the RTTOV fast radiative transfer model (currently at version 12), Geosci. Model Dev., 11, 2717–2737,

720 https://doi.org/10.5194/gmd-11-2717-2018, 2018.