Technical Memo

ECMWF
European Centre for Medium-Range Weather Forecasts

894

Improved surface treatment for all-sky microwave observations

Alan J. Geer, Katrin Lonitz, David I. Duncan, Niels Bormann
(Research Department)

April 2022

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/publications/

Contact: library@ecmwf.int

Abstract

Cycle 48r1 of the Integrated Forecasting System (IFS) contains numerous improvements to the surface representation for all-sky microwave observations. For the first time, mid-frequency microwave window channels (in the region of 37 GHz, 89 GHz and 150 GHz) are assimilated over land surfaces, bringing sensitivity to water cloud, rain and snow in the atmosphere. This makes use of a new technique to extrapolate surface emissivity retrievals (made at lower frequencies) up to the frequencies of the assimilated channels. Further, and recognising that at least 20% of all observations can be mixed scenes (combinations of water, land and sea-ice), the treatment of mixed surfaces has been improved. In particular, the emissivity retrieval and extrapolation is now restricted to the land and sea-ice part of the scene, while the water part is represented with a successful physically-based model. Where possible, this also uses a land-sea mask tailored to the field of view of the observational data, rather than relying on the model land-sea mask. There is also more precise screening of sea-ice and surface snow, so that observations can now be assimilated at higher latitudes, in colder seasons, and closer to coastlines. Overall, this package increases data usage by between 4% and 33% in most actively assimilated channels, with the biggest effects in the northern hemisphere (NH), in winter, and in polar regions. This improves forecasts in the NH higher latitudes out to around day 3. This work is a precursor to further developments aiming to realise a full 'all-sky all-surface' assimilation of microwave radiances in the IFS.

1 Introduction

The direct assimilation of satellite radiances is one of the major advances in weather forecasting in recent decades. This was first achieved during the late 1990s (e.g. Eyre *et al.*, 2020) but with a stringent approach to quality control that excluded observations with sensitivity to the non-ocean surface, as well as those areas affected by clouds and precipitation. The development of dynamic surface emissivity retrievals (e.g. Karbou *et al.*, 2006, 2008; Krzeminski *et al.*, 2009; Baordo and Geer, 2016) and surface emissivity atlases (e.g. Aires *et al.*, 2011) subsequently opened up the assimilation of observations over land and sea-ice surfaces, but only for those channels with limited surface sensitivity. The development of all-sky radiance assimilation (e.g. Bauer *et al.*, 2010; Geer *et al.*, 2014, 2017b) similarly opened up the use of observations sensitive to cloud and precipitation. Assimilating radiances over land and sea-ice, and globally in all-sky conditions, led to significant improvements in forecast quality (Bormann *et al.*, 2017; Geer *et al.*, 2017b). A first benefit was to use satellite data in situations that are poorly constrained by other observations, such as remote land, polar areas (Lawrence *et al.*, 2019) or locations with cloud (McNally, 2002). A second benefit was to directly influence geophysical variables, such as water vapour, clouds and precipitation, that were previously thought to be of marginal benefit (e.g. Bengtsson and Hodges, 2005) and difficult to assimilate (Errico *et al.*, 2007). All-sky radiance observations brought very different geophysical sensitivities compared to the temperature-sounding observations that were initially developed, and this was a great advantage, bringing new information on wind fields (Geer *et al.*, 2014), influencing smaller scale features than the temperature-sensitive radiances (Bonavita *et al.*, 2020) and directly improving severe weather events such as tropical cyclones (Duncan *et al.*, 2022). But these developments are incomplete: observations with stronger sensitivity to land or sea-ice surfaces are still not assimilated, and sensitivities to the geophysical variables of the surface, such as skin temperature, soil moisture and sea-ice concentration, are still not used directly. The past successes, and the possibility that inferring surface variables directly from radiances could help improve predictability beyond the medium-range (de Rosnay *et al.*, 2022), motivate further developments in a strategy that is now dubbed 'all-sky all-surface assimilation'.

As part of this strategy, the current memo describes improvements in the representation of the surface for all-sky microwave observations, along with the extension to new situations (such as mixed scenes) and to the use of more strongly surface-sensitive channels over land surfaces, particularly the mid-frequency microwave imaging channels (e.g. 37 GHz, 89 / 92 GHz, 150 / 166 GHz). At cycle 47r3, the all-sky assimilation framework is applied to microwave cross-track sounders AMSU-A, MHS and MWHS-2, to conically scanning microwave imagers AMSR2 and MWRI, and to conical sounder-imagers SSMIS and GMI[1]. Some of these sensors are assimilated from more than one satellite so, as of October 2021, a total of 17 sensors are assimilated in the all-sky framework. As much as possible, this framework treats all sensors in a standardised and generic way, based on the physical details of each satellite channel, minimising the need for sensor and instrument-driven configuration. This means that scientific developments can be rolled out to all applicable sensors in one go, making it more likely that those developments will provide noticeable impact on the forecasts, and making it worthwhile to work on fine details that might have little impact when applied to a single sensor. An example is the extension of the microwave humidity sounder channels to coastal and lake areas, which gave significant improvements in forecast scores (Weston *et al.*, 2017).

In the same spirit, the current work combines many different extensions and improvements, with some of the larger changes being the use of surface-sensitive channels over coast and lake areas, extension of imager channels into ice-free polar oceans, and greater use of surface-sensitive channels over high latitude land. The biggest single change in the current work is to start assimilating microwave imager channels over land surfaces, following the development of the atlas-constrained extrapolation method for using dynamic emissivity retrievals (Lonitz *et al.*, 2022). Note that all this work is based on a physical assumption of the surface as a specular reflector described by a surface emissivity and a frequency-independent skin temperature, despite known problems in this approach (e.g. Bormann *et al.*, 2017). This work explores developments that are feasible within the current paradigm, leaving an improved physical representation of the surface radiative transfer for the future. Section 2 gives further background and motivation, and Sect. 3 gives more details of the existing configuration of all-sky microwave observations. Sections 4 and 5 detail the new methods. Results and conclusion are in Sects. 6 and 7.

2 Benefits and challenges of assimilating microwave window channels over more difficult surfaces

Putting aside the challenges of modelling land, snow and sea-ice surface radiative transfer for a moment, a fundamental question is whether microwave window channels give any worthwhile sensitivity to the atmosphere above these surfaces. The utility of window channels over oceans is driven by the strong radiative contrast between the low-emissivity ocean surface and thermal emission from the lower atmosphere (e.g. from water vapour, cloud and rain). Land, snow and sea-ice surfaces are generally high-emissivity, and the thermal radiative contrast is typically small. Figure 1 (taken from Fig. 8 of Geer *et al.*, 2021) illustrates the impact of different hydrometeor types in a theoretical slab cloud over an upwelling brightness temperature (TB) of either 100 K or 280 K, representing ocean and land respectively. In this example, the cloud is of a similar temperature to the upwelling brightness temperature over land, so the thermal contrast is minimal. Hence, the main signals over 'land' are generated by scattering, which decreases the brightness temperature: the impact of scattering from rain is strongest around 50 GHz and scattering from frozen particles (snow, graupel, ice cloud) increases rapidly above this frequency.

A more comprehensive estimate of the hydrometeor signal over land is given in Fig. 2a, based on simula-

[1] Satellite and sensor acronyms will not be spelled out in this work, but can be found at e.g. https://space.oscar.wmo.int

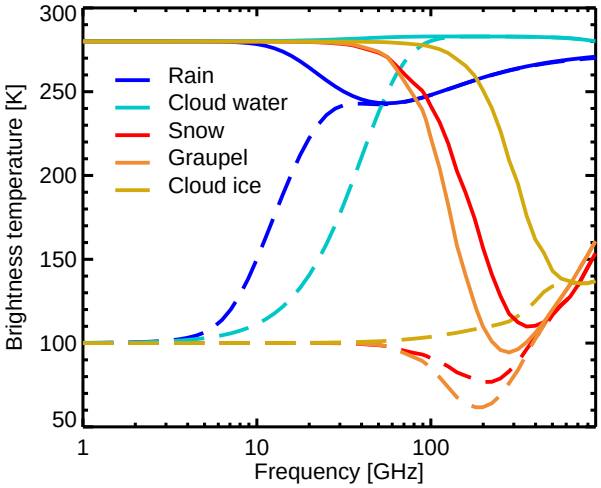

Figure 1: Reproduction of Figure 8 from Geer *et al.* (2021), illustrating the impact of different hydrometeor types over radiatively cold and warm surfaces, using an idealised model. Cloud-top brightness temperatures are simulated using a two-stream radiative transfer approximation for a uniform slab of hydrometeor that is 2 km thick and has a water content of 1×10^{-3} kg m^{-3}. The upwelling brightness temperature below the hydrometeor slab is either 280 K (solid lines) or 100 K (dashed lines), representing land or ocean surfaces. The slab temperature is 253 K for snow, graupel and cloud ice, and 283 K for rain and cloud water).

tions of AMSR2 observations at 89 GHz, v-polarised (channel 89v). The difference between the all-sky and clear-sky simulated TB shows the effect of hydrometeors on the brightness temperature. At high latitudes (for example Greenland, and parts of Siberia), there is at least a 5 K increase in TB, indicating thermal emission over a relatively cold surface, likely from liquid water cloud. TBs over land are more typically reduced by at least 5 K in many areas, and up to 40 K in places (off the scale of the figure, not shown). This is the scattering signal from larger frozen particles and rain that was illustrated theoretically in Fig. 1. The examples in the figure are seen in tropical convection areas (e.g. equatorial areas affected by the ITCZ, in South America and Africa and the maritime continent), monsoon areas (e.g. India and SE Asia) and summertime convection over midlatitude land areas (e.g. central North America, Siberia).

The window channels at 89 GHz are somewhat below the peak frequencies for the snow scattering signal, but the rain scattering signal is near its maximum (Fig. 1). Hence, these mid- and low-frequency window channels can be sensitive to snow and graupel particles in the full column of the atmosphere, and sensitive to rain too, which is typically found in the lowest few km. Some of the scattering signal from snow and graupel is already assimilated from 183 GHz water vapour sounding channels over land surfaces (e.g. Geer *et al.*, 2014) but due to the clear-sky weighting function and the intensity of the scattering, this signal mainly comes from the tops of the convection (e.g. Hong *et al.*, 2005; Geer, 2021a). Ground-based radar also sees rain and snow, but its coverage is limited, and its assimilation at ECMWF is restricted to the contiguous USA (Lopez, 2013; Geer *et al.*, 2017a). Hence, these window channels bring completely new information in global weather forecasting.

A further motivation for using low-frequency channels over land is given by Liu *et al.* (2021), who focused on the 24 GHz humidity channels. Although the sensitivity to vertically integrated water vapour is overall close to zero in these channels, this is partly due to a compensation of sensitivities, with positive sensitivity to water vapour in the lowermost atmosphere (below 800 hPa) and negative sensitivity

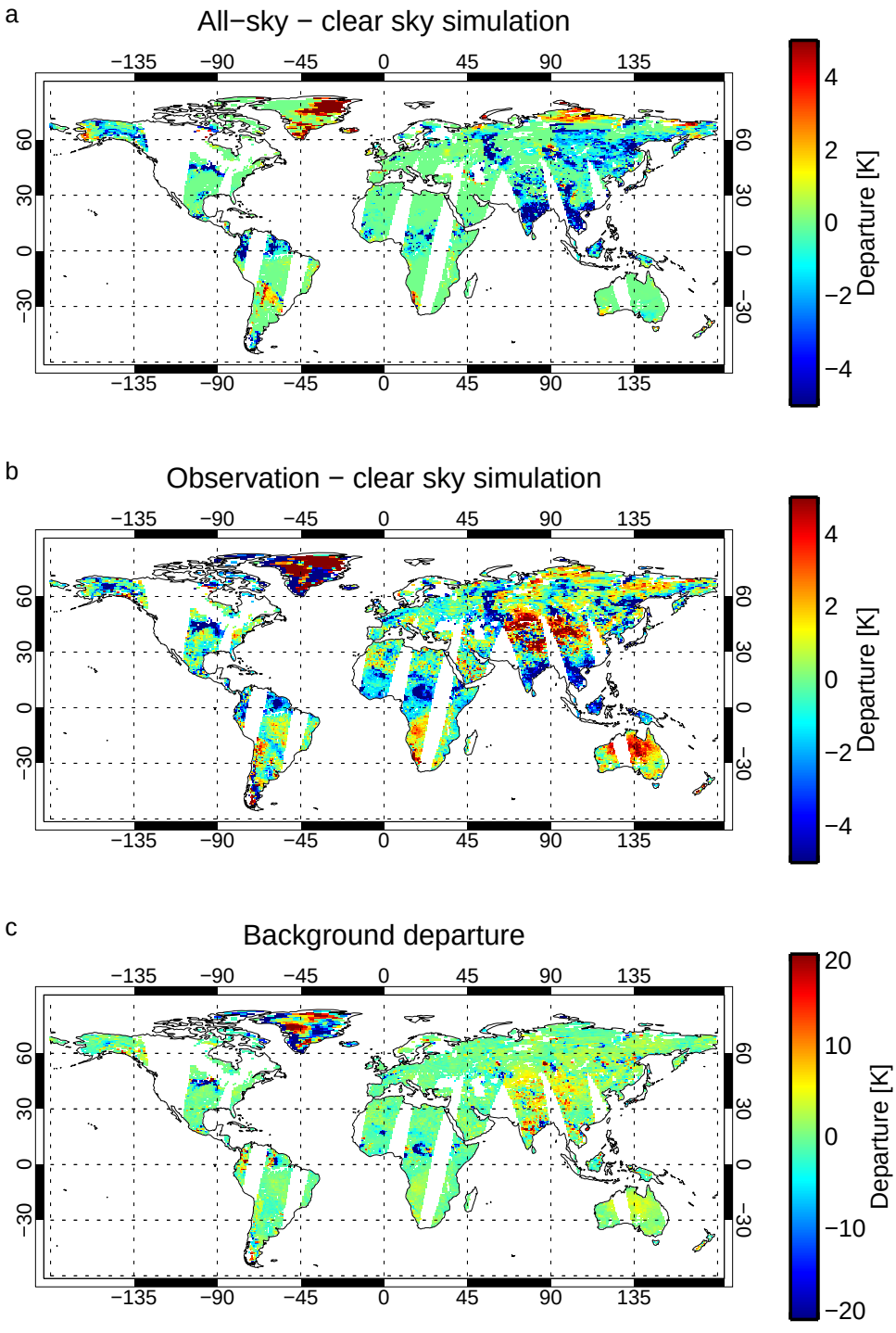

Figure 2: Departures for all AMSR2 89v data over land available to the 00 UTC analysis on 1st July 2019: (a) Simulated hydrometeor signal (cloud effect); (b) Observation minus clear-sky simulated TB; (c) All-sky background departure, using a broader contour range.

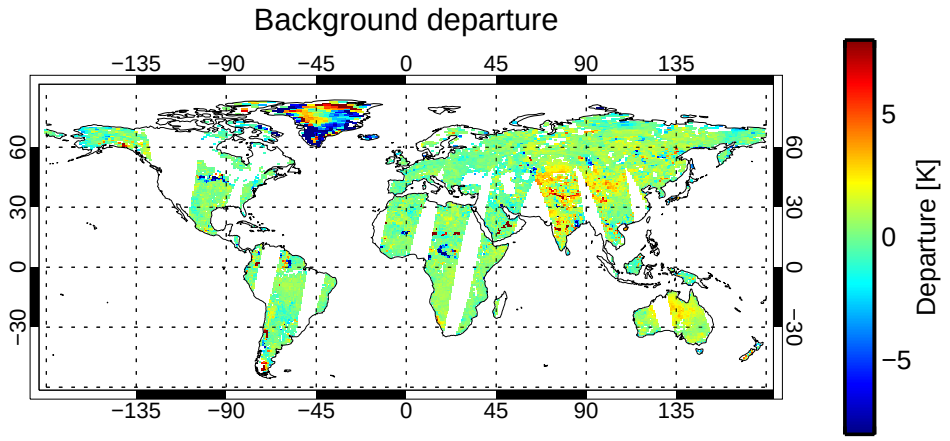

Figure 3: Background departures for AMSR2 37v data over land in the 00 UTC analysis on 1st July 2019.

higher up. If the upper-level moisture is constrained by 183 GHz sounding channels, then information on humidity in the lowermost atmosphere can be extracted from the 24 GHz channels over land. Hence, there is also strong interest in using window channels for their humidity information.

Figure 2b highlights some of the challenges of using window channels over land. Assuming that the background humidity errors are relatively low (and relatively invisible in a single channel), the difference between observed all-sky brightness temperatures and the clear-sky background simulation is a combination of the observed hydrometeor signal and the error in the surface representation. The rain and snow scattering signal from ITCZ convection, monsoon areas and summertime midlatitude convection is still clear, and quite similar to the simulated picture in panel a. However, the signal is often present over a much wider area than in the simulations, particularly over equatorial Africa. This is a known problem: the IFS seems to have trouble simulating sufficient convection over land surfaces, and this is visible in both microwave and IR radiances (e.g. Geer *et al.*, 2019; Geer, 2021b). Compared to panel a, panel b also has some large new areas of positive departures, for example over Greenland, the Tibetan plateau and northern Australia. Most likely this comes from the difficulty in representing the surface over snow, ice and dry soils (Bormann *et al.*, 2017, and Sect. 5.3 later).

Figure 2c shows the all-sky background departure on a broader contour range than used in the rest of the figure. The all-sky background departure highlights positive and negative departures that are likely due to displaced or poorly forecast precipitation features. Some of these departures could provide useful corrections to the location and extent of convection in the background, but in some areas the errors coming from the representation of the surface (as identified from panel b) are nearly the same magnitude, for example over Australia. Hence, the challenge of using window channels over land needs to be addressed through better representation of the surface (Sect. 5), and failing that, through appropriate quality control and observation error modelling.

Figure 3 shows the background departures in the AMSR2 37 GHz v-polarised channel (37v). Here, the signal from convection is reduced, but the signal from poorly represented surfaces (such as over the Tibetan plateau) is not much smaller than at 89 GHz. Lonitz *et al.* (2022) did not test 37 GHz over land, justified by this relatively small signal from hydrometeors. However, 37 GHz is included in the current work with a view to future developments. First, it is a stepping stone to using the 24 GHz channel over land, and hence unique information content on boundary layer moisture (Liu *et al.*, 2021). Second, it is

the channel with perhaps the purest sensitivity to rain, since the snow scattering signal is typically very small at this frequency (Fig. 1). With appropriate QC and observation error modelling, the data can be assimilated with little risk. Even if the amount of useful atmospheric signal that can be extracted is small to start with, it will benefit from any future improvements in the surface representation. Further, by its presence in the system, it will challenge us to do better.

3 Current usage and overview of methods

As mentioned in the introduction, the all-sky framework includes 17 different sensor and satellite combinations as of October 2021. This encompasses humidity sounding, temperature sounding and window channels. Table 1 summarises the sensors according to their channel availability and usage. For brevity it excludes channels with stratospheric sensitivities, but these are available on MWHS-2, AMSU-A and SSMIS (they are assimilated from MWHS-2 and AMSU-A but not SSMIS). The channel ID is a combination of the central frequency in GHz, typically but not always rounded down, and sometimes a polarisation identifier, and sometimes the frequencies of any sidebands. The polarisation identifier has been omitted from sounding channels (those around 50, 118 and 183 GHz) since no instrument has dual polarisation sounding channels, and although the polarisation varies across the sensors, it is more compact to ignore it in this presentation. In some cases, nearby frequencies are combined into one channel ID. For example, GMI has dual-polarisation channels at 166 GHz, SSMIS and MWHS-2 have a channel at 150 GHz and MHS has one at 157 GHz. However, all of these channels have similar properties. The table gives the maximum possible channel usage; on some satellites, mainly due to broken channels, usage may be reduced. The horizontal lines break the channels into logical sets, starting with low-frequency window channels and then the temperature-sounding channels in the 50 - 60 GHz oxygen line complex. The mid to high-frequency window channels around 90 GHz and 150 GHz are split by the 118.75 GHz oxygen line, which is used for channels on MWHS-2. Finally, the 183 GHz channels are for humidity sounding. The main exclusion in this work is ATMS, which still uses the clear-sky approach and follows a different path through the IFS, so it is not touched here.

Table 2 lists the satellite and sensor combinations being assimilated. Further instrument and satellite details (and all acronyms) are given on the WMO OSCAR web page (https://space.oscar.wmo.int). There is vast scope and heterogeneity across the sensors listed in Tables 1 and 2; even more so given that some satellites have reduced channel sets due to sensor deterioration (see e.g Duncan *et al.*, 2022). To tame this complexity, the all-sky microwave processing is based as much as possible on the physical characteristics of each generic channel (e.g. the channel ID in Table 1) and the availability of channels with valid observations, rather than making decisions based on a specific sensor or satellite ID. One tool for simplification is a rough classification of the channel's vertical sensitivity, given as the 'type' in Table 1.

The current work describes a large number of changes to the all-sky framework that are made according to this generic approach. The heterogeneity of the sensors and satellites means that results are channel, sensor and satellite dependent. Based on the channel ID, Table 3 gives an overview of current usage by surface type, and it highlights changes made in the current work. The table is valid for sensors with full capabilities. Due to missing channels, some sensors do not benefit as broadly. To summarise the main themes:

- Low and mid frequency imager channels (19v - 37v, 89/92v) were previously used over ocean only, excluding high latitude areas. This work adds ocean scenes at high latitudes, by using a new sea-ice screening based on 10 GHz channels where available (Sect. 4.2). At 37v and 89/92v, they

Table 1: Channel availability and usage of sensors in the all-sky framework, as of October 2021. Where a channel exists on a sensor and is used actively, the channel number is given without brackets. Channel numbers with brackets are not actively assimilated; these are typically used for quality control and error modelling. The channel ID contains the relevant frequency in GHz and possibly the polarisation (v=vertical; h=horizontal) and possibly the sidebands in GHz (e.g. ±7). The type ID roughly describes the vertical sensitivity of the channels: 0=surface sensitive; 1=tropospheric; 2=upper troposphere/lower stratosphere (UTLS); 3=stratosphere. Further details are in the text.

Channel ID	Type	AMSR2	MWRI	GMI	SSMIS	MHS	MWHS-2	AMSU-A
6v	0	(1)						
6h	0	(2)						
7v	0	(3)						
7h	0	(4)						
10v	0	(5)	(1)	(1)				
10h	0	(6)	(2)	(2)				
19v	0	7	3	3	13			
19h	0	8	4	4	12			
22/24v	0	9	5	5	13			(1)
22/24h	0	10	6		14			
31/37v	0	11	7	6	16			(2)
37h	0	(12)	(8)	(7)	(15)			
50.3	0				(1)			(3)
52.8	0				(2)			(4)
53.6	1				(3)			5
54.4	1				(4)			6
54.9	1							7
55.5	2				(5)			8
89/92v	1	13	9	8	17	(1)	(1)	(15)
89/92h	1	(14)	(10)	(9)	(18)			
118±5	0						(9)	
118±3	0						(8)	
118±2.5	0						7	
118±1.1	1						6	
118±0.8	2						5	
150/157/166v	0			10		(2)	(10)	
150/166h	0			11	8			
183±7	0			13	9	5	15	
183±4.5	0						14	
183±3	1			12	10	4	13	
183±1.8	1						12	
183±1.1	1				11	3	11	

Table 2: Satellites assimilated for each sensor, as of November 2021. Citations covering the all-sky implementation of these sensors in the IFS are also given.

Sensor	Satellite	Citation
AMSR2	GCOM-W2	Kazumori *et al.* (2016)
MWRI	FY-3D	Bormann *et al.* (2021)
GMI	GPM	Lean *et al.* (2017)
SSMIS	DMSP F-17; DMSP F-18	Bauer *et al.* (2010); Geer *et al.* (2014)
MHS	NOAA-19; Metop-A; Metop-B; Metop-C	Geer *et al.* (2014)
MWHS-2	FY-3C; FY-3D	Lawrence *et al.* (2018)
AMSU-A	NOAA-15, NOAA-18; NOAA-19; Metop-A; Metop-B; Metop-C	Duncan *et al.* (2022)

Table 3: Current usage (✓ = actively assimilated; × = not assimilated) and changes made in the current work ('Added'), as a function of surface type.

Channel ID	Ocean $\leq 60°$	Ocean $> 60°$	Ocean coast	Sea-ice	Land $\leq 60°$	Land $> 60°$	Snow	Mixed land
19v	✓	Added	✓	×	×	×	×	×
19h	✓	Added	✓	×	×	×	×	×
22/24v	✓	Added	✓	×	×	×	×	×
22/24h	✓	Added	✓	×	×	×	×	×
31/37v	✓	Added	✓	×	Added	Added	×	Added
53.6	✓	✓	✓	✓	✓	✓	✓	✓
54.4	✓	✓	✓	✓	✓	✓	✓	✓
54.9	✓	✓	✓	✓	✓	✓	✓	✓
55.5	✓	✓	✓	✓	✓	✓	✓	✓
89/92v	✓	Added	✓	×	Added	Added	×	Added
118±2.5	✓	×	Added	×	×	×	×	×
118±1.1	✓	✓	✓	×	✓	✓	✓	Added
118±0.8	✓	✓	✓	×	✓	✓	✓	Added
150/157/166v	✓	×	✓	×	Added	×	×	Added
150/166h	✓	×	✓	×	Added	×	×	Added
183±7	✓	Added	Added	×	✓	Added	×	Added
183±4.5	✓	Added	Added	×	✓	Added	×	Added
183±3	✓	✓	✓	✓	✓	✓	✓	Added
183±1.8	✓	✓	✓	✓	✓	✓	✓	Added
183±1.1	✓	✓	✓	✓	✓	✓	✓	Added

also gain land scenes, but excluding snow (Sect. 5).

- The AMSU-A 50 - 60 GHz temperature sounding channels were already assimilated in almost all conditions, and the current developments do not much change the number of observations being assimilated. However, there are complex regional changes in the scene usage. The most important changes affecting AMSU-A are the treatment of surface emissivity and screening in mixed scenes (Sect. 4.1). The main increases are in fractional ice scenes, affecting 53.6 and 54.4 GHz channels. There are also some losses due to more severe quality control over mixed scenes.

- The upper peaking 118 GHz channels on MWHS-2 (118±1.1 and upwards) were already assimilated over most ocean and land surfaces, but not sea-ice. The surface-sensitive channel 118±2.5 was only assimilated over oceans excluding high-latitude areas. The current work increases available data due to the changes in the treatment of mixed scenes (Sect. 4.1).

- The 150/157/166 GHz channels are only assimilated within latitudes of 45°, to restrict their usage to high humidity areas where the surface sensitivity is small. Hence changes relating to higher latitudes have no effect. However, for the microwave imagers, land scenes are added for the first time (Sect. 5). These channels are only assimilated from microwave imagers, not sounders.

- The low-peaking (surface-sensitive) 183 GHz channels were already assimilated over land and ocean but excluding high latitudes. High latitude ocean scenes are added where a 10 GHz channel is available for sea-ice detection (Sect. 4.2); high latitude snow-free land is added for all sensors (Sect. 4.3). Mixed ocean and land scenes are added for MHS in similar ways as for MWHS-2 (Sect. 4.1).

The developments were broken down into two stages during development and testing. Stage 1 covers the changes affecting mainly window channels, and is described in Sect. 4. Stage 2 covers the addition of microwave imager channels over land surfaces, and is described in Sect. 5. Stage 2 also includes the upgrade of GMI 183 GHz usage to the same standard as other humidity sounders (Sect. 5.4, with gains over land, sea-ice, snow, and high-latitude oceans); it also includes changes to the SSMIS 183 GHz error model (Sect. 5.5). These stages are mainly relevant to the results from experimental testing described in Sect. 6, which also summarises utilisation changes more precisely as a function of sensor. The following sections take illustrative examples from the experimental testing, but detailed descriptions of this testing (e.g. dates, resolution) are left until Sect. 6.

Further information on the configuration of the all-sky microwave assimilation in the IFS can be found in preceding articles. Geer *et al.* (2017b) gives an overview of historical all-sky microwave assimilation developments along with a summary of the impact of microwave window and humidity-sounding channels. These channels (when also including ATMS) currently give around 18% of all impact from the global observing system on the short-range forecast quality, as measured by forecast sensitivity to observation impact (FSOI). Duncan *et al.* (2022) describes the recent move of AMSU-A into the all-sky processing stream at ECMWF. The microwave temperature sounding channels (including ATMS) account for another 18% of short-range forecast impact. Hence microwave sensors are currently the dominant observation type in the IFS. Further descriptions of the all-sky framework, including superobbing and thinning, quality control, and observation error modelling, are given by Bauer *et al.* (2010); Geer and Bauer (2010, 2011); details of the all-sky microwave observation operator (RTTOV-SCATT) are given by Bauer *et al.* (2006); Geer and Baordo (2014); Geer *et al.* (2021). The existing all-sky treatment of land surfaces with a dynamic surface emissivity retrieval is described by Baordo and Geer (2016). Specific papers and tech. memos cover the addition of different sensors to the all-sky framework; these are listed in Table 2.

4 Methods for stage 1: better treatment of surface-sensitive channels

4.1 Treatment of mixed surfaces

4.1.1 Introduction

Until recently, mixed scenes have been treated with caution, due to the very different radiative characteristics of ocean, land and sea-ice, and particularly because the emissivity of land and sea-ice surfaces is variable and poorly known. The effective radiating temperature of ocean can be as low as 90 K, whereas for land or sea-ice it is typically higher than 260 K. When attempting to physically model a window channel radiance, it is necessary to know the exact fraction of ocean, land and sea-ice in the scene. The proportion of mixed scenes is larger than is often appreciated, and can be as much as 33%, but a representative figure might be 22% (see Appendix A).

For surface-sensitive channels, mixed scenes have usually been rejected, so that only pure ocean or pure land scenes have been used until now. From the first assimilation of SSM/I 1D-Var total column water vapour at ECMWF (Gérard and Saunders, 1999), a typical definition of 'pure ocean' has been a land fraction less than 0.01. This threshold gives a maximum simulation error of around 2 K when the surface emissivity is modelled as if the whole scene is ocean, ignoring the small land component. When direct all-sky assimilation was implemented (Bauer *et al.*, 2010) the maximum land fraction check was applied at every inner loop resolution used in 4D-Var, and a more stringent threshold was introduced in some places, rejecting observations with land fractions greater than 0.0. The aim of checking the land fraction at every resolution was to avoid having inconsistent surface types at different stages of the assimilation process. However, this very cautious approach creates a halo of rejected data out to around 100 km around coasts and islands, depending on the inner loop resolutions. With the extension of all-sky assimilation to land surfaces (Karbou *et al.*, 2008; Geer *et al.*, 2014; Baordo and Geer, 2016) a minimum land fraction of 0.95 was permitted, and again this was checked at all resolutions, creating a halo of rejected data around coastlines and lakes. Scenes with land fractions between 0.01 and 0.95 were labelled 'coast' and discarded. This excludes many observations over land, especially at high latitudes in the northern hemisphere.

Two more recent developments improved the situation, at the expense of further complicating the processing. The first development was that sounding channels with little sensitivity to the surface (e.g. channels 3 and 4 of MHS) in situations with a land fraction greater than 0.01 can be treated as land, rather than 'coast', and actively assimilated (Weston *et al.*, 2017). Land scenes use the dynamic emissivity retrieval approach, which is agnostic to the amount of land or water in the scene (e.g. Baordo and Geer, 2016). One possible issue with mixed scenes in this case is the observation error modelling, since the symmetric cloud predictors are necessarily different between land and ocean. For example, MHS uses different flavours of scattering index (SI, Geer *et al.*, 2014). The current solution is to use the land error model over mixed scenes, with no obvious issues so far.

The second development covers the microwave imagers SSMIS, AMSR2 and GMI. Since cycle 46r1, the land-sea mask for these sensors has been calculated for their exact field of view (FOV), based on a high-resolution (5 km) land sea mask, which is convolved with a 2D Gaussian model of the FOV. This is based on the 10 GHz channel characteristics for GMI and AMSR2, and on the 19 GHZ channels in the case of SSMIS. The per-FOV land-sea masks are then averaged to form the superob land-sea mask. For the sensors with the exact FOV land fraction, the checks through the minimisation on the model land fraction are removed. For safety, the maximum permissible land fraction was reduced from 0.01 to 0.00 through much of the all-sky processing. Using the FOV-appropriate land fraction for screening

has allowed microwave imager observations to be assimilated much closer to coastlines than previously when the screening relied on the model land fraction instead.

This accumulation of developments has led to a complex, error-prone treatment of different surface types in the all-sky system. Two examples are given here because, when corrected in the current work, there have been locally large improvements. The first affects AMSU-A channels with land fractions greater than 0.0 but less than 0.01 which are treated in the ocean processing path, but until now they have used a crude estimate of the land emissivity as the surface emissivity. An ocean emissivity should really have been used. Correcting this simple mistake had a locally large impact on assimilated AMSU-A 53.6 GHz data (channel 5, Sect. 6.2). Second, an 'on-off' approach was applied to sea-ice areas in non-assimilated channels, i.e. those that did not use the more modern dynamic emissivity approach over sea-ice. The FASTEM (FAST microwave Emissivity Model, Kazumori and English, 2015) ocean emissivity was used for scenes with sea-ice fractions less than 0.5, and an emissivity of 0.7 was used for sea-ice fractions greater than this. This matters because some of the non-assimilated channels are used for deriving observation error and for QC. For example the non-assimilated 52.8 channel (ch. 4) is used for departure-based screening of the some of the active AMSU-A channels. Improving the surface emissivity modelling for mixed sea-ice scenes in the current work has widespread effects on the active AMSU-A channels (see Sect. 4.1.3). In this work it is hoped to simulate all channels over all surfaces with as much accuracy as possible to support current usage and future developments.

4.1.2 New approach

Given the complex and fragile current state of the all-sky treatment of mixed scenes, this work has two aims: first, to simplify and rationalise the various approaches and second, to represent mixed surfaces as well as possible within the existing paradigm of dynamic emissivity retrieval. This approach should be able to cope with any mixed scene, so for example the coastal screeening has been dropped for scenes with land fraction greater than 0.01. This extends the Weston *et al.* (2017) approach to surface-sensitive channels, so that mixed land-ocean scenes are lumped into the 'land' processing, and treated in a similar way to land. The checks on the model land fraction at inner loop resolutions have been mostly dropped, given that the FOV-appropriate land fraction is now available for some sensors (where it is not available, e.g. imagers like MWRI and sounders like MWHS-2, these still retain checks on the inner loop model land fractions). These screening changes add significant data in areas such as Canada, Siberia and Scandinavia in channels such as MHS channel 5, its lowest-peaking humidity sounding channel. This threshold for 'land' rejection and/or processing is now mostly standardised on a land fraction of 0.01, which affects AMSU-A for example, which was previously rejected for scenes with land fractions between 0.01 and 0.5.

An issue with the emissivity retrieval approach is how to extrapolate across frequencies. Typically, the retrieval at one frequency is expected to be valid at another, which is most likely to work in the 50 GHz temperature channels, where the sounding channels are close in frequency to the window channels used for the emissivity retrieval. However, for the 183 GHz channels over land surfaces, emissivity is retrieved around 89 GHz, making the approach more prone to errors over surfaces with strong spectral variability in emissivity. One surface that has a strong but well-known variability with frequency is water. Hence mixed scenes with large water fractions should not be extrapolated as constant in frequency. One main idea in the current work is to use FASTEM (Kazumori and English, 2015) to physically model the part of the scene that is known to be water. Further, although the 'coastal' class of data has been removed, a new class of 'liminal' scenes is introduced, representing mixtures of land, ocean and sea-ice.

The new approach starts by establishing the combined land and sea-ice fraction, l_{landice}, noting that the

sum of sea-ice and land fractions is not guaranteed to add to 1, but the land-sea mask is bounded by 1:

$$l_{\text{landice}} = \min(l_{\text{land}} + l_{\text{seaice}}, 1.0) \quad (1)$$

$$l'_{\text{seaice}} = l_{\text{landice}} - l_{\text{land}} \quad (2)$$

$$l'_{\text{water}} = 1.0 - l_{\text{landice}} \quad (3)$$

$$\therefore l'_{\text{water}} + l'_{\text{seaice}} + l_{\text{land}} = l'_{\text{water}} + l_{\text{landice}} = 1.0 \quad (4)$$

Here, l_{land} is the land fraction, which is set to the FOV-appropriate land fraction for GMI, SSMIS and AMSR2, and to the IFS land-sea mask for other sensors. The IFS sea-ice fraction l_{seaice} comes from the OCEAN5 analysis (Browne *et al.*, 2019, prior to cycle 45r1, the OSTIA product was used directly). Where the sea-ice and land fractions sum to more than 1, priority is given to the land fraction, so that a corrected sea-ice fraction l'_{seaice} and ocean fraction l'_{water} are used. The new 'liminal' category for mixed scenes is defined by $0.01 < l_{\text{landice}} < 0.95$.

The dynamic emissivity retrieval (Baordo and Geer, 2016) estimates a whole-scene emissivity $e_{\text{retrieved}}(i)$ where i represents the channel. Since we have a reasonably accurate estimate of the ocean surface emissivity from FASTEM, $e_{\text{water}}(i)$, this can be used to constrain the emissivity of the water part of the scene, so that the retrieval is now used only to infer the emissivity of the land and sea-ice part of the scene:

$$e_{\text{landice}}(i) = \frac{1}{l_{\text{landice}}} \left(e_{\text{retrieved}}(i) - l'_{\text{water}} e_{\text{water}}(i) \right) \quad (5)$$

This is done for 'liminal' scenes only. The resulting part-scene emissivity retrieval $e_{\text{landice}}(i)$ is checked against the relevant existing thresholds (for example over land the minimum and maximum permissible emissivities are 0.55 and 1.0; Baordo and Geer, 2016). If this check fails, the retrieval approach is abandoned. Over surfaces that are more homogeneously land and ice ($l_{\text{landice}}(i) \geq 0.95$) a check is also made against the TELSEM2 emissivity atlas value (Tool to Estimate Land Surface Emissivity from Microwave to Submillimeter Waves Aires *et al.*, 2011), but for liminal scenes this is not done, since the atlas may not reflect the true mix of surface types. Broadly, the new treatment of liminal scenes imposes a stricter test of the retrieved emissivity than before, particularly where scenes are dominated by water. As will be seen, the amount of data that gets rejected depends on the quality of the land and sea-ice fractions. Note that the new approach is most appropriate when the FOV-appropriate land fraction is available. If this is not the case, interpolated model values will be used, even though they may not accurately represent the actual land fraction in the FOV, and this may negatively affect the emissivity assignment. As a result, more data is retained with sensors that have the FOV-appropriate land fraction, but for sensors where this has not been implemented, particularly AMSU-A, there can be a loss of data instead.

For liminal scenes with a successful retrieval of $e_{\text{landice}}(i)$ (i.e. within the emissivity bounds mentioned above) the land and sea-ice emissivity can be extrapolated to another channel j using the constant-in-frequency approximation. However, the water surface emissivity in that channel j should use the modelled value for the relevant channel, $e_{\text{water}}(j)$. Hence the whole-scene emissivity in the other channel is

$$e(j) = l_{\text{landice}} e_{\text{landice}}(i) + l'_{\text{water}} e_{\text{water}}(j). \quad (6)$$

For all other land or sea-ice scenes (i.e. $l_{\text{landice}} \geq 0.95$) the standard approach is followed, with no attempt to adjust for any small amount of water surface in the scene:

$$e(j) = e_{\text{retrieved}}(i). \quad (7)$$

Figure 4 shows the importance of treating water surfaces carefully when extrapolating the emissivity in frequency. This example shows the extrapolation from 50.3 GHz (AMSU-A channel 3) to 89 GHz

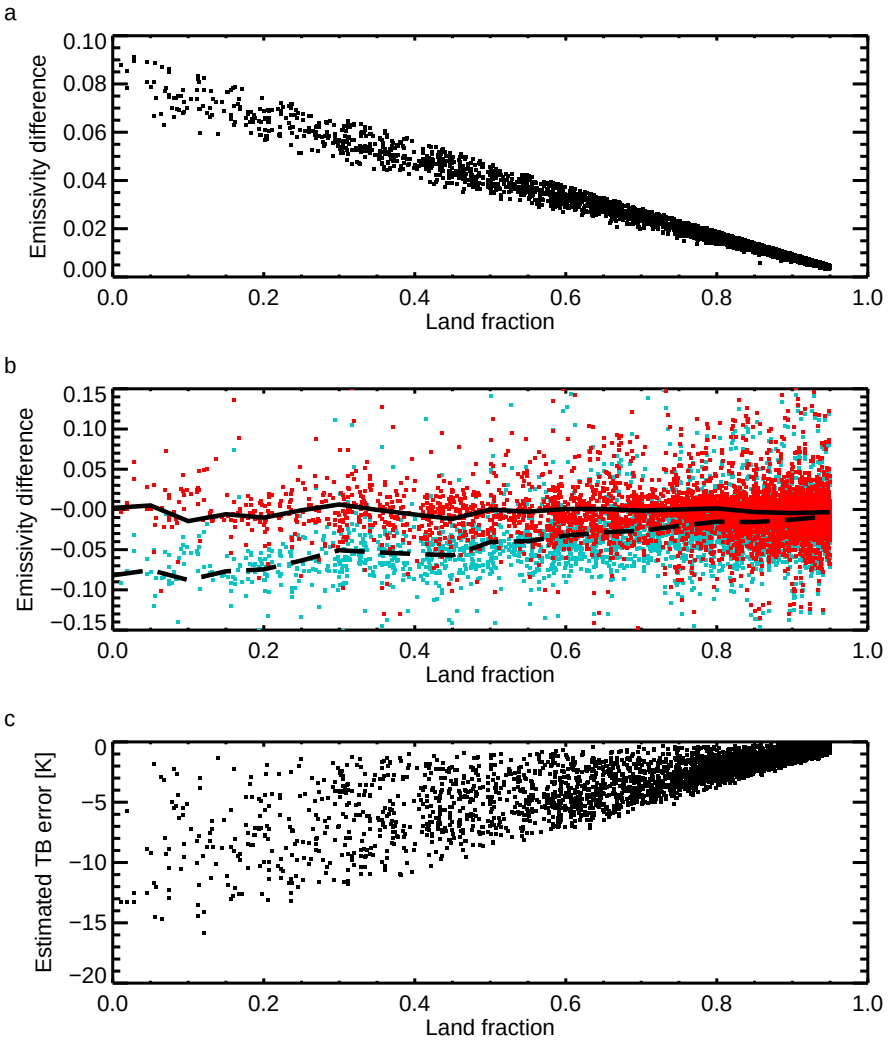

Figure 4: Scatter plots of (a) Emissivity difference between using Eq. 6 to correctly model the emissivity of the water part of the 89 GHz AMSU-A (ch. 15) scene, compared to using the 50.3 GHz (ch. 3) emissivity retrieval directly (Eq. 7); (b) Difference between the previous two emissivity options and the retrieved surface emissivity at 89 GHz (ch. 15): red, solid black line uses Eq. 6; blue, dashed black line uses Eq. 7, where the lines are the binned means; (c) Estimated increase in error in brightness temperature simulations due to the differences in panel a. Sample is based on all AMSU-A observations with a valid 50.3 GHz and 89 GHz emissivity retrieval, for mixed scenes ($0.01 < l_{\text{land}} < 0.95$) excluding areas with sea-ice. Observations are those available at the 00 UTC analysis on 20th June 2019.

(AMSU-A channel 15). The ocean surface emissivity can be up to 0.1 higher at 89 GHz than at 50.3 GHz; hence the emissivity for a mixed scene varies too, dependent on the land fraction (panel a). The dynamic emissivity retrieval at 89 GHz, where available, can be used to validate this approach (panel b). This shows that Eq. 6 is unbiased compared to the retrievals, whereas a constant-in-frequency extrapolation from 50.3 GHz to 89 GHz (Eq. 7) generates a significant low bias in the surface emissivity. The effect of this low emissivity bias on the simulated brightness temperature can be estimated in using the formula of English (2008), simplified to

$$dTB = T_{\rm skin}\tau^2\delta e, \tag{8}$$

where τ is the clear-sky surface-to-space transmittance in the 89 GHz channel, $T_{\rm skin}$ is the skin temperature (assumed identical to the atmosphere temperature) and δe is the emissivity error. Panel c shows that the error in simulating the 89 GHz brightness temperatures can be as much as -15 K. Hence, it is clearly worth using FASTEM to model the water part of a scene correctly.

Going back to the new framework for mixed scenes, in the case that the retrieval from the window channel (either $e_{\rm landice}(i)$ or $e_{\rm retrieved}(i)$) was not successful, then there is the possible use of the TELSEM2 atlas (Aires *et al.*, 2011) as backup, where available, but this is not done for liminal scenes, or for sea-ice or snow scenes (since the snow and sea-ice locations can be very different from one year to the next). Failing this, a universal backstop surface emissivity is provided:

$$e_{\rm backstop}(j) = l_{\rm land} \times 0.95 + l'_{\rm seaice} \times 0.8 + l'_{\rm water} e_{\rm water}(j). \tag{9}$$

If a channel has very small surface sensitivity (channels falling under the 'middle atmosphere' category; for example AMSU-A channel 9 and upwards) then it will be assimilated even if the backstop emissivity is being used. However, the assumed emissivities of 0.95 and 0.8 for land and sea-ice are far from universally valid. Especially over snow and some types of sea-ice, they could be in error by 20 - 40%. Hence, all other channels with the backstop emissivity are rejected, but the backstop provides a physically-justifiable estimate for computing a background departure for monitoring purposes; previously this was not even attempted in some cases.

The backstop of Eq. 9 is currently very crude, but it is only applied in channels where the surface has little or no influence on the observations. However, it is a framework that will be built on in subsequent developments. An obvious improvement would be to use land and sea-ice emissivity estimates that vary with frequency, polarisation and situation, possibly using physical, empirical or hybrid models to do so (see remarks in the conclusion). With these developments, if they are accurate enough, it could eventually replace the dynamic emissivity framework for all channels.

4.1.3 Examples of impact on AMSU-A

Figure 5 further explores the effect of the new framework with a sample of observations from AMSU-A channel ID 52.8, i.e. channel 4 (see Table 1). This channel is not assimilated, due to its partial surface sensitivity, but it illustrates the influence of the surface representation on the background departures. In the control setup (panel a) observations with land fractions $0.0 < l_{\rm land} \leq 0.95$ and in the surface-sensitive category (see Table 1) defaulted to a poor estimate of the 100% land surface emissivity. The mainly positive departures for land fractions between 0.8 and 0.95 show that the estimated land surface emissivity is too low (it is around 0.87[2], whereas the atlas and retrievals give around 0.94, not shown). The departures are more balanced above land fractions of 0.95 because these use the dynamic surface

[2]The land emissivity in the all-sky path defaults to the FASTEM estimate, but this uses a set of land surface coefficients of unknown origin, and FASTEM land emissivities using recommended coefficients would have been more accurate

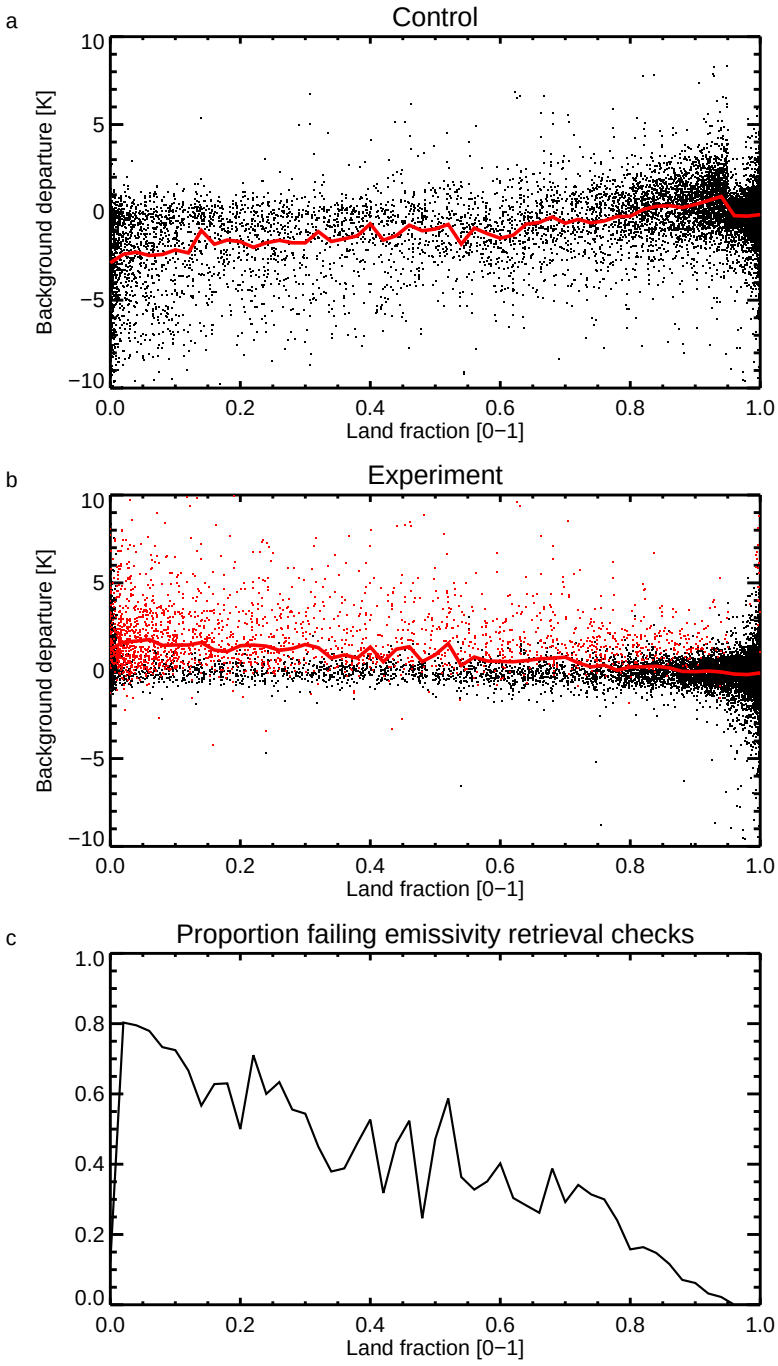

Figure 5: Scatter plots of AMSU-A channel 4 background departures as a function of land fraction for (a) control and (b) experiment using the improved emissivity processing described in this section. The red line shows the mean of all points on panels a and b respectively. (c) shows the proportion of scenes in the experiment that are eliminated due to emissivity retrieval checks, and hence that use the backstop emissivity; these scenes are coloured red in panel b and are included along with all other points on the panel in the red line (the mean). The sample is all mixed land scenes ($0.0 < l_{\text{land}} < 1.0$) based on all Metop-A,B and -C AMSU-A observations available to the 00 UTC analysis on 20th June 2019.

emissivity retrieval. For small land fractions, the assumption of 100% land was also problematic, leading to negative departures on average. For land fractions close to zero, departures could be as large as -10 K in this channel.

Fig. 5b shows the departures computed with the new approach to surface emissivities for mixed scenes. Above land fractions of 0.8, departures are now mostly unbiased. Towards small land fractions, the departures now tend positive, due to scenes where the backstop emissivity is used (Eq. 9; dots on Fig. 5b shown in red). This is likely due to the FOV of AMSU-A, which is around 48 km, and is hence significantly larger than the model grid (in this case, 25 km, and operationally it is 8-9 km). A small land fraction typically indicates a coastal area, meaning there is likely to be larger amounts of land close by, so the true land fraction in the field of view would generally be larger. In coastal areas generally, even where $l_{\text{land}} = 0.0$ (figure not shown) there are some unexpectedly large positive background departures from AMSU-A, further suggesting a problem of land contamination coming from the large field of view size. However, all the scenes that use the backstop emissivity are removed by quality control. The remaining scenes (black points) use the extrapolated emissivity retrieval (Eq. 6 or 7) and give unbiased departures. These are the additional coastal and mixed scenes that would be available for assimilation if channel 4 were assimilated.

Fig. 5c shows the proportion of mixed scenes where the backstop emissivity is used, or equivalently, where the retrieval $e_{\text{landice}}(i)$ from Eq. 5 failed the bounds checks. The proportion of failures increases as the land fraction gets smaller. This shows how constraining the emissivity of the water part of the scene (as in Eq. 5) gives a new and quite strict test of the retrieval, and of the consistency of the assumed and actual land (or sea-ice) fraction. For sensors such as AMSR2, which have the FOV-appropriate land fraction, the proportion of failures is smaller (max. 60%, not shown) and even when the backstop emissivity is used, the land fraction is more appropriate so the background departures are much less biased. Hence it would be important in future developments to extend the FOV-appropriate land fraction to all sensors including AMSU-A.

Figure 6 shows the locations of AMSU-A liminal scenes, along with the status of the emissivity retrieval QC check which, if failed, rejects all surface-sensitive channels along with tropospheric and tropopause sounding channels, up to AMSU-A channel 8 (see Table 1). Liminal areas are often found at high northern latitudes, in parts of Canada, Scandinavia and Siberia where lakes are frequent in the landscape. They are also found around the coastlines, islands, and along major rivers such as the Amazon. Over ocean, liminal scenes are areas of fractional sea-ice. Sea-ice areas were already assimilated from AMSU-A, so the blue areas (failed emissivity QC) are often scenes that were previously assimilated and are now lost. As with land areas, using FASTEM to constrain the emissivity of the ocean part of the scene imposes a stricter test of the emssivity retrieval (and of the correctness of the assumed sea-ice fraction). Many of the areas now lost to QC are around the sea-ice edge, where it is likely the IFS sea-ice fraction is inconsistent with the true fraction in the FOV (see Sec. 4.2). There is also a large patch of blue extending deep into the Antarctic sea-ice near the Greenwich meridian. Here, the retrieved whole-scene surface emissivity is close to 1, and yet the sea-ice fraction is around 0.95 or less. From Eq. 5 this situation would generate emissivities in the sea-ice part of the scene that are greater than 1, and hence trigger the QC. This means that the IFS sea-ice fraction and the retrieved surface emissivity are physically inconsistent in these regions.

Figure 7 further summarises the results, again focusing on the 52.8 GHz channel (AMSU-A channel 4) which is not yet assimilated. Panel a shows the RMS background departure of samples binned over July 2019 in the new setup (see Sect. 6 for details of the relevant experiments). This figure shows the results without any of the usual screening. RMS errors are high around coasts, islands and lakes (so there is no need to draw the coastlines on this figure). RMS errors are high in some parts of the sea-

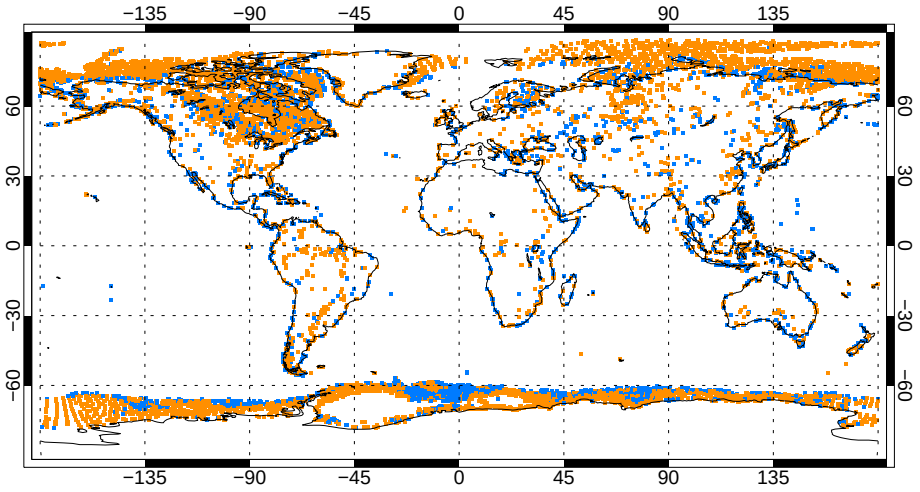

Figure 6: Locations of liminal scenes ($0.01 < \min(1.0, l_{\text{land}} + l_{\text{seaice}}) \leq 0.95$) based on all Metop-A,B and -C AMSU-A observations available to the 00 UTC analysis on 20th June 2019. Orange spots have valid surface emissivity retrievals; blue spots do not. These scenes make up 12% of all available Metop-A,B and -C data. With a broader definition of mixed scenes ($0.0 < \min(1.0, l_{\text{land}} + l_{\text{seaice}}) < 1.0$) 27% of all scenes would be mixed (not shown).

ice (e.g. around Antarctica) and snow and ice-covered areas (e.g. Greenland). Errors are also high through the ITCZ, monsoon regions and other convective areas due to cloud and precipitation errors in the background (Sect. 2). The new approach is generally much better than the old approach, with some exceptions. Around the coastlines, islands, and lake areas, there has been a significant improvement in the fit to observations (reduction in RMS background departures, panel b) mainly due to the replacement of the old approach that assigned 100% land emissivities in many cases (Fig. 5). The region of the sea-ice edge in the Arctic and Antarctic also shows improvement, mainly due to the replacement of the 'on-off' approach to fractional sea-ice. There are also some areas where convection errors are randomly different between experiments (e.g. ITCZ, monsoon areas). These areas should be ignored. The areas where the RMS gets worse are those where the emissivity retrieval QC has become stricter (Fig. 6) and hence these are areas where the backstop surface emissivity is now being used.

The most visible degradation in Figure 7b is the Antarctic sea-ice, where a previous fixed backstop emissivity of 0.7 for sea-ice fraction greater than 0.5 has been replaced by Eq. 9. The new backstop uses 0.8 for the sea-ice emissivity, and though this is a better estimate generally, it generates larger background departures in these cases. The stricter QC also results in some cases where the backstop emissivity is now used rather than the dynamic emissivity retrieval, but where the retrieval would have given better results. For the future, it may be possibly to identify situations where the emissivity retrieval is inadequate but still better than the backstop. However, the Antarctic sea-ice regions have always been tricky to deal with for AMSU-A, and they have always been removed by a special additional QC for the lowest-peaking assimilated AMSU-A channel, 53.6 GHz (channel 5). The best future solution would be to find a better way of modelling the sea-ice emissivity.

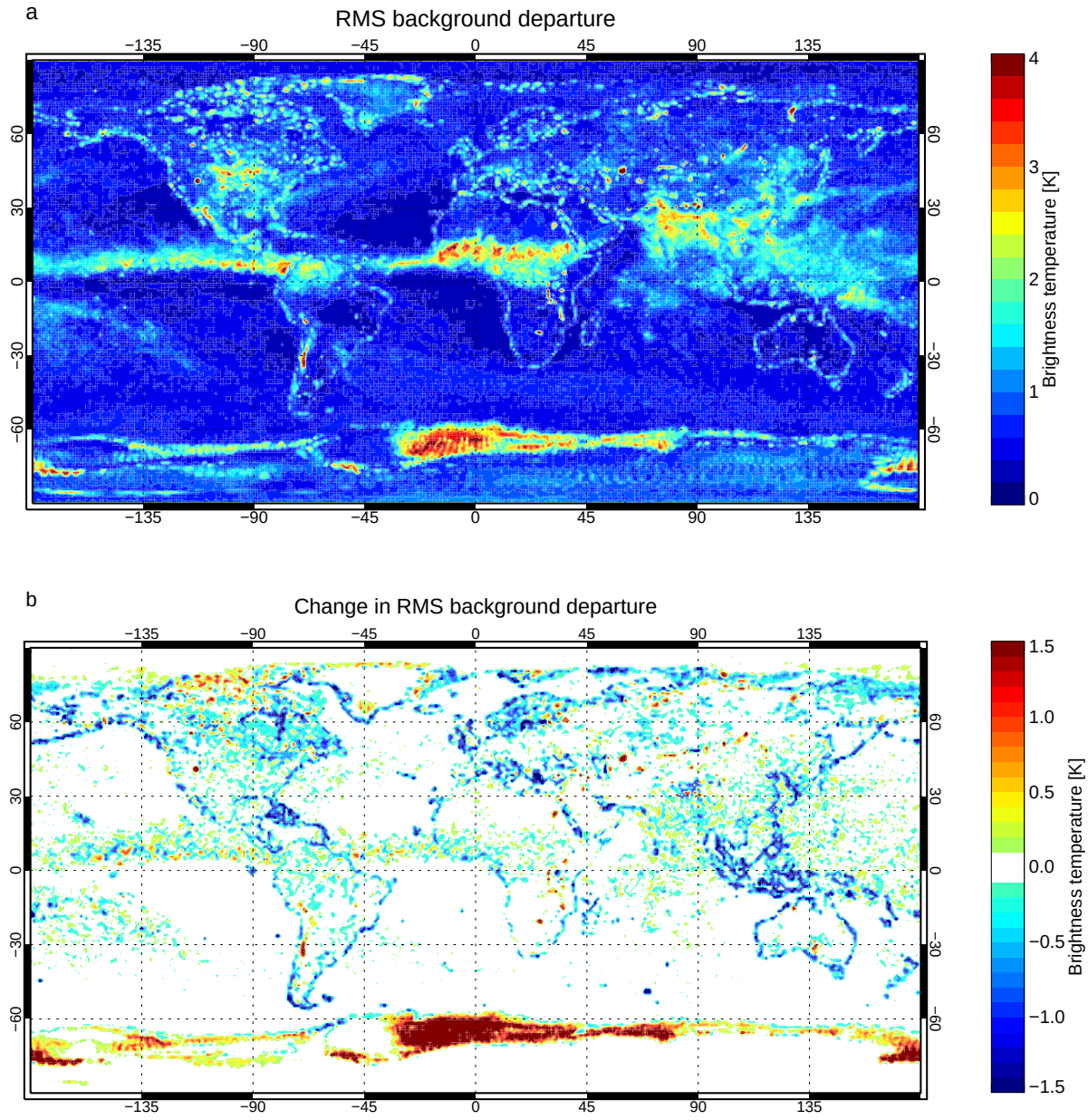

Figure 7: For July 2019, for AMSU-A 52.8 GHz (Ch. 4) (a) the RMS error of background departures; (b) the difference in RMS error between the new approach and the control, in 1° bins, including all AMSU-A observations without any screening.

4.2 Initial sea-ice retrieval for quality control

As mentioned above, the sea-ice concentration in the IFS comes from the OCEAN5 analysis (Browne *et al.*, 2019) which assimilates OSTIA L4 sea-ice maps (Donlon *et al.*, 2012; Good *et al.*, 2020) that use OSI-SAF retrievals as input (Lavergne *et al.*, 2019). OSI-SAF retrievals are based on the same microwave observations that are assimilated as radiances at ECMWF, but they use a simpler representation of the problem, based on a tie-point method to infer the sea-ice fraction. The long-term strategy at ECMWF is to infer the sea-ice fraction directly from the microwave radiances already assimilated at ECMWF. This is to benefit from the more optimal use of observations possible in a coupled assimilation system, and to be able to use the most up-to-date observations (OSTIA could be as much as 48 h out of date when used directly in the IFS Baordo and Geer, 2015; de Rosnay *et al.*, 2022). Thanks to the presence of a sea-ice model to propagate information in time and space, the IFS sea-ice fields are now more accurate (Browne *et al.*, 2019) than when OSTIA was used directly (e.g. Baordo and Geer, 2015). However, the IFS sea-ice is still not accurate enough to use for screening surface-sensitive microwave observations; if it were used as the sole screening criterion then some very large departures, coming from undetected sea-ice, would be visible in the vicinity of the sea-ice edge (see Fig. 10 later).

The current approach over sea-ice is to assimilate channels with sensitivity to the troposphere and upwards, whereas window channels are rejected (precisely, channel types 1–3 and 0 on Table 1). Over scenes defined as ocean ($l_{\text{land}} < 0.01$), if the IFS sea-ice fraction is greater than zero ($l_{\text{seaice}} > 0.0$) then the sea-ice dynamic emissivity retrieval is triggered (e.g. Di Tomaso *et al.*, 2013; Geer *et al.*, 2014). Where the sea-ice fraction is zero, the scene emisivity is always treated as being pure ocean, using FASTEM (note that scenes with $l_{\text{land}} >= 0.01$ are treated through a separate land path with dynamic emissivity retrieval, so coastal sea-ice is implicitly handled in the land path, and is not further discussed in this section). As mentioned above, there is still some chance of the observations being contaminated by sea-ice, due to inaccuracies in the IFS sea-ice fraction. Hence, additional screening has previously been performed for window channels, by excluding ocean observations at latitudes greater than 60° latitude and/or with skin temperatures less than 274.0 K. The main goal for the sea-ice retrieval in 48r1 has been to remove this overly-broad latitude-based check on window channels, which rejects large amounts of potentially useful data in the high-latitude oceans. To do this requires an observation-based check for sea-ice.

A simple retrieval of sea-ice fraction has been introduced for any sensor with a 10 GHz h-polarised channel (e.g. GMI, AMSR2, MWRI, see Table 1). 10 GHz is chosen because of the relatively small atmospheric signal at these frequencies, and the more limited variation of sea-ice surface emissivity, compared to higher frequencies where volume scattering from snow becomes important. Also, since its field of view is larger than that of higher frequencies, it is more sensitive in spatial terms too. Finally, the h-polarised channel is preferred over the v-polarised because of the much stronger contrast between ocean and sea-ice emissivity. As in the new backstop surface emissivity (Eq. 9), the sea-ice surface emissivity is assumed to be $e_{\text{ice}} = 0.8$. Then the sea-ice fraction can be estimated from the dynamic emissivity retrieval $e_{\text{retrieved}}(i)$ and the FASTEM estimate for the ocean water surface emissivity $e_{\text{water}}(i)$ where i=10h is the index of the 10 GHz h-polarised channel:

$$l^*_{\text{seaice}} = \frac{e_{\text{retrieved}}(10\text{h}) - e_{\text{water}}(10\text{h})}{e_{\text{ice}} - e_{\text{water}}(10\text{h})} \tag{10}$$

To do so, the dynamic emissivity retrieval is also applied over ocean surfaces for the first time. Given the potential for mis-identifying heavy precipitation as sea-ice, this retrieval is only carried out with sea surface temperatures less than 278.0 K and for latitudes greater than 40°. The retrieval is protected against very small denominators and against emissivities outside of the range 0 to 1. Note that in case of

melt ponds on the surface of the sea-ice, these will be identified as ocean, but from a radiative transfer point of view, that is the correct thing to do.

Figure 8 illustrates the sea-ice retrieval in comparison to the IFS sea-ice fraction. Both in Arctic summer (top panels) and Antarctic winter (bottom panels) the sea-ice locations are very similar between the IFS and the retrievals. In mixed scenes in Arctic summer, the retrieval clearly generates higher sea-ice fractions than those seen in the IFS: for example, along the Greenland E coast, and in Hudson Bay, the IFS has fractions around 0.5 - 0.7 whereas the retrieval gives fractions around 1.0. This might suggest that the assumed sea-ice surface emissivity of 0.8 is too low. However, given $e_{\mathrm{water}}(10\mathrm{h}) \simeq 0.4$ and $e_{\mathrm{retrieved}}(10\mathrm{h}) \geq 0.8$ it would not be possible to generate sea-ice fractions of 0.5 in Eq. 10 without unphysical $e_{\mathrm{ice}} \geq 1.2$. This suggests either that the IFS sea-ice fraction is too low in those areas, or that the emissivity retrieval $e_{\mathrm{retrieved}}(10\mathrm{h})$ is too high, perhaps due to incorrect physical assumptions about the sea-ice characteristics.

Figure 9 summarises the agreement between IFS and retrieved sea-ice fraction on a scatter plot. Where the IFS sea-ice fraction is high, retrieved sea-ice fractions go as high as 1.4, particularly in the Antarctic winter. These areas mainly correspond to the sea-ice interior, and it seems likely that the true sea-ice fraction is close to 1 in these areas. The retrieved sea-ice fraction could be reduced within physical bounds by assuming $e_{\mathrm{ice}} \simeq 1.0$; alternatively the emissivity retrieval could be too high due to physical inconsistencies in the retrieval. However, for sea-ice fractions below 0.5, the cloud of points is not excessively displaced compared to the 1:1 line. A simplistic interpretation would be that the assumed sea-ice surface emissivity of 0.8 is better tuned to detecting the sea-ice edge in areas of discontinuous sea-ice, and is less suited to the interior regions. These issues will be examined in much more depth in future. For the smallest sea-ice fractions, the discrepancies between the IFS sea-ice fraction and the retrieval are likely to be real. The retrieved sea-ice is more up to date than the OSTIA data assimilated in OCEAN5, and also the retrievals are more precisely tailored to the field of view of the observations.

o In any case, for the present work, the main point of this initial sea-ice retrieval at cycle 48r1 is screening. Any ocean scene within the area previously subject to blanket latitude screening (latitudes greater than 60°) is now rejected if $l^{*}_{\mathrm{seaice}} > 0.05$; scenes where the sea-ice retrieval fails are also rejected. The areas where the IFS sea-ice mask is non-zero ($l_{\mathrm{seaice}} > 0.0$) are also rejected, as before. The results are explored in Fig. 10. In the top panel, this shows the background departures of all new observations (ocean observations at latitudes greater than 60°) that could be used if the IFS sea-ice fraction was precise enough to use for sea-ice screening. However, around the edges of the main sea-ice areas (compare to Fig. 8a) there are positive departures greater than 10 K. These positive departures are consistent with assuming an ocean surface when the scene actually contains some sea-ice, and hence has higher surface emissivity. The lower panel shows that the new rejection criterion, based on retrieved sea-ice, eliminates all these scenes while barely reducing the number of new observations that can be assimilated. A few remaining spots with large positive departures (e.g. just off the coast of Norway) are in areas that are unlikely to be affected by sea-ice, and hence are most likely explained by insufficient liquid cloud or rain in the background forecast. Hence, the new sea-ice QC appears precise, and the overall result is to provide large amounts of new data in the window channels in high-latitude oceans. The relative success of the simple sea-ice retrieval inspires further improvements, with the goal of assimilating the retrieved sea-ice fraction in the IFS ocean analysis.

4.3 Land surfaces

A 60° latitude check was previously applied over land surfaces for the lowest humidity sounding channel (e.g. the 183±7 GHz). However, this was unnecessarily stringent seeing as the ‘land snow’ check is so

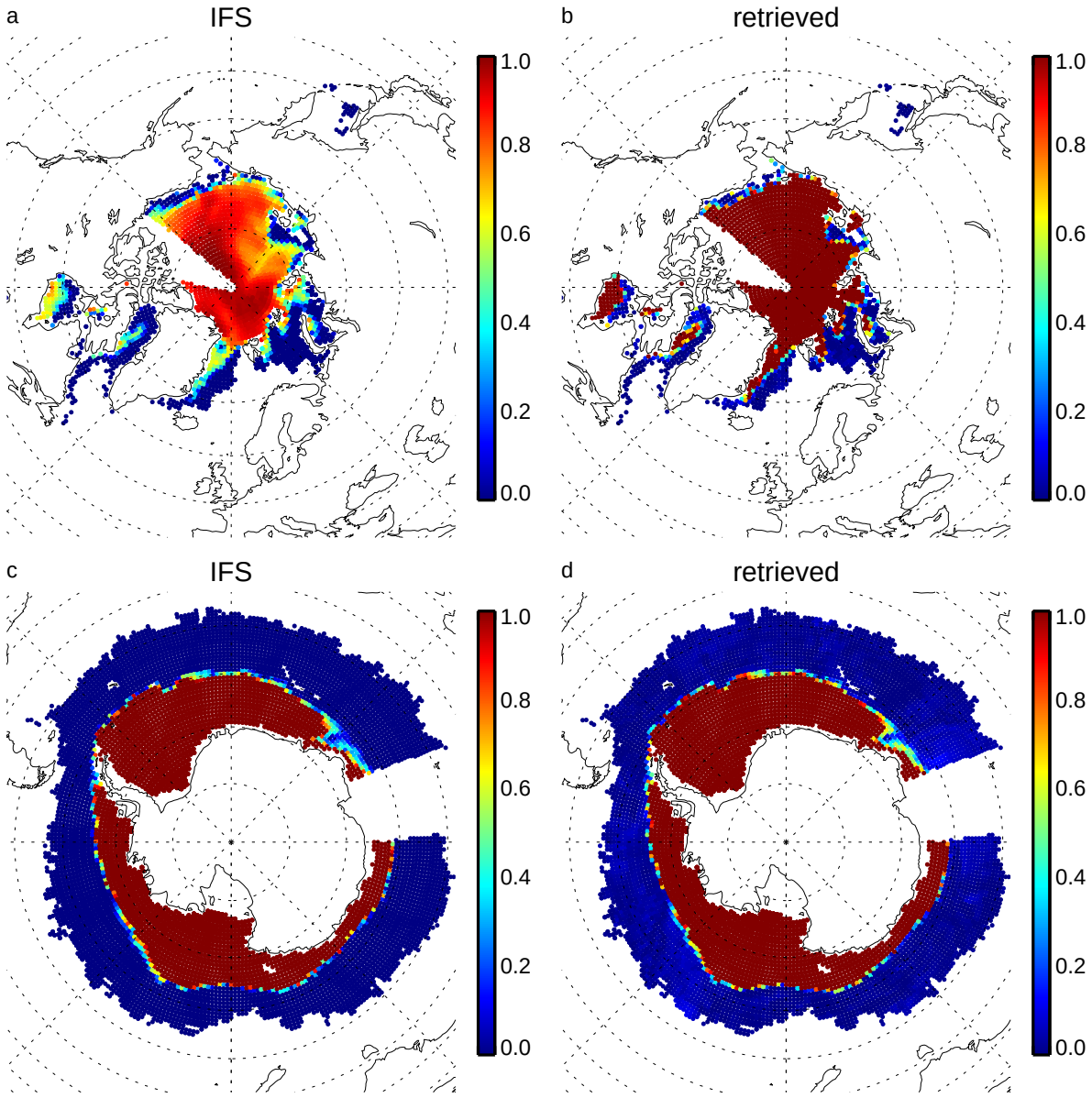

Figure 8: Sea-ice fraction (a/c) from the IFS and (b/d) retrieved via Eq. 10 from AMSR2 observations, selecting ocean areas ($l_{\text{land}} < 0.01$) with latitudes greater than 40° and skin temperatures below 278.0 K. Sea-ice fractions greater than 1 are present in the retrievals but are displayed as if 1 on the colour scale. Observations are those available at the 00 UTC analysis on 1st July 2019.

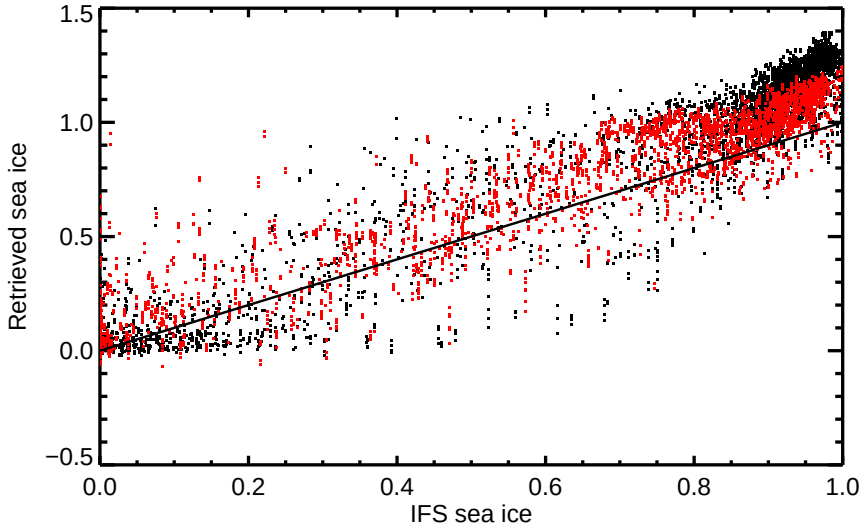

Figure 9: Scatter plot of IFS and retrieved sea-ice fraction, for the same sample as Fig. 8, based on the 00 UTC analysis on 1st July 2019. Red spots indicate Arctic locations; black spots Antarctic. The solid diagonal line is the 1:1 line.

broad: any scene with skin temperature less than 278.0 K is treated as possibly containing snow, which also leads to the rejection of the data. Hence, for 48r1, the latitude check was removed completely over land. This adds significant amounts of new observations in the warmer months, particularly over Siberia and Canada.

Since the work of Geer and Baordo (2014), scenes with excess convection over land surfaces ('Excess SI') have been screened out, based on the poor representation of deep convection over land in the IFS and in RTTOV. The exact criteria has been to reject any scene with symmetric scattering index (SI) greater than 20 K. However, RTTOV v13.0 gives improved representation of convection over land in the midlatitudes (Geer, 2021b; Geer *et al.*, 2021), so the excess SI check has been dropped. This adds a small fraction of extra, very cloudy, observations over land for sensors such as MWHS-2 and MHS, which see slightly larger standard deviations as a result.

4.4 Changes specific to AMSU-A

A number of changes have been made that only affect AMSU-A, broadly aiming at further standardisation of the all-sky framework by removing special-purpose processing. One change is to start providing bias corrections for AMSU-A channels 1,2 and 15 by assimilating them passively (i.e. for VarBC spinup purposes, but not otherwise affecting the analysis). This causes small changes in the background departures in these channels, which gives very minor downstream effects on the active channels 5-14, because the observation error models and QC and depend in various ways on channels 1,2 and 15. Note that the changes to processing of mixed scenes (Sect. 4.1) affect AMSU-A active channels directly, but also indirectly through the use of nominally 'passive' channels in screening and observation error modelling. Another important change is that previously coastal regions were removed for $0.01 \leq l_{\text{land}} \leq 0.5$ for AMSU-A channels 5 and 6. Since fractional land-sea areas are now provided with more physically robust surface emissivity estimates, these observations are now allowed to be assimilated. Other small tidying of the AMSU-A denylist (formerly known as blacklist) has been carried out, but without any

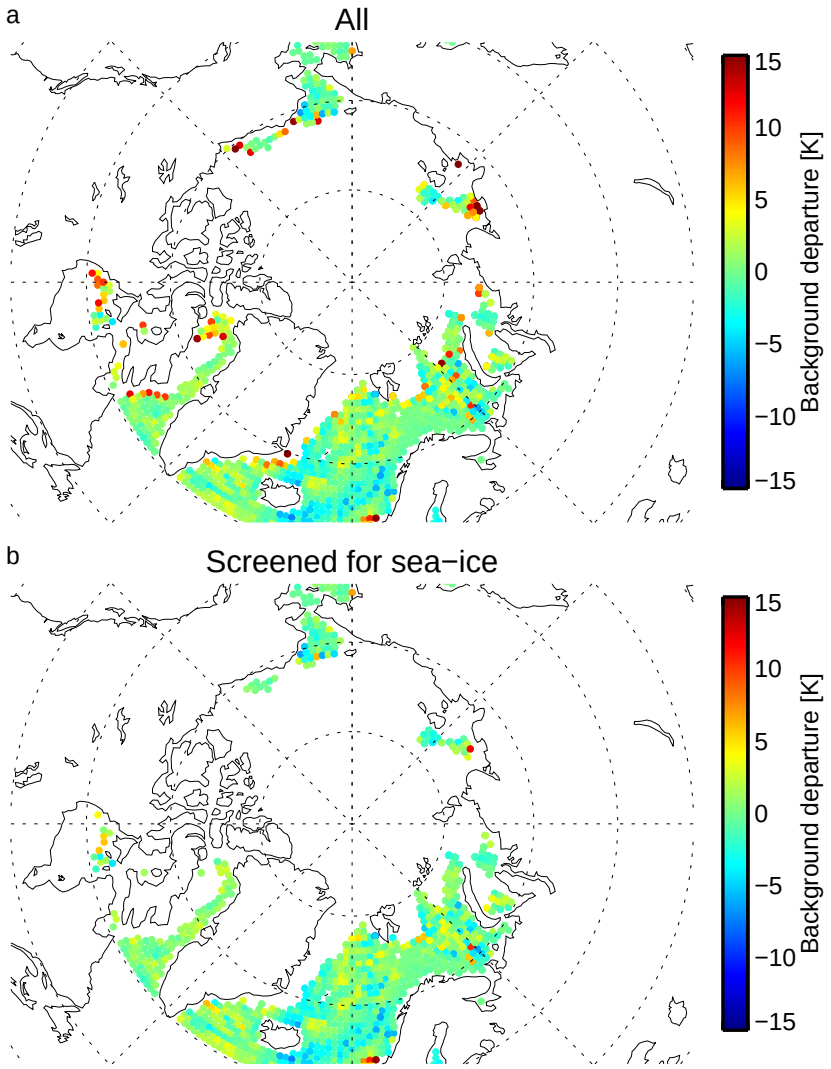

Figure 10: Background departures at 19v for all ocean areas, eliminating areas with IFS sea-ice greater than zero ($l_{\text{seaice}} > 0.0$), and in panel (b), also excluding areas with retrieved sea-ice greater than 0.05 ($l^*_{\text{seaice}} > 0.05$). The sample of observations is otherwise the same as in Fig. 8, based on AMSR2 observations in the 00 UTC analysis on 1st July 2019.

active effects (see below). The overall effects of the changes to AMSU-A processing are explored in Sect. 6.

4.5 Passive changes

Passively monitored channels and observations are also affected by the cycle 48r1 upgrade. One development has been to provide land and sea-ice support for all all-sky sensors, including passively-monitored sensors like Windsat and historical sensors like TMI, by removing all relevant sensor-dependent switches within the code; this has no effect on active data utilisation, but should allow future reanalyses to use the most up-to-date configuration. Observation Database (ODB) archiving has been extended to include soil moisture, soil temperature, LAI, and snow density as possible drivers for developing land bias models, QC, emissivity models or errors; this uses additional surface model fields available in the observation operator following the work of Hirahara *et al.* (2020). A lot of standardisation has been done, for example to move sensor-dependent checks out of the denylist and to replace them with generic, physically-based screening criteria within the all-sky code path (e.g. AMSU-A channel 5 sea-ice and snow checks for Antarctica). Further, denylist checks which duplicate QC checks within the code have been removed entirely (e.g. MHS ch. 2 and 5 snow and land QC).

5 Methods for stage 2: window channels over land surfaces

5.1 Atlas-constrained emissivity extrapolation

To enable the assimilation of microwave imager channels over land, Lonitz *et al.* (2022) developed a new approach that extends the dynamic emissivity retrieval used in the assimilation of surface-sensitive microwave sounding channels. The idea is to retrieve the emissivity at a frequency less affected by cloud and rain, and to adjust the retrieved emissivity to the required frequency using the frequency-dependence of the TELSEM2 atlas. The work of Lonitz *et al.* (2022) showed promising results at an earlier cycle of the IFS; the current work extends the approach further, and starts with a quick recap of the emissvity extrapolation.

The atlas-based emissivity adjustment uses the emissivity delta from the TELSEM2 atlas, that is $e_{\mathrm{TELSEM2}}(j) - e_{\mathrm{TELSEM2}}(i)$ where j is the channel to which the emissivity is applied, and i is the channel where the emissivity retrieval is made:

$$e(j) = e_{\mathrm{retrieval}}(i) + e_{\mathrm{TELSEM2}}(j) - e_{\mathrm{TELSEM2}}(i). \tag{11}$$

Figure 11 motivates the atlas adjustment. Panel a shows the difference in dynamic surface emissivity retrievals between the 89v and 19v channels of AMSR2. In some areas associated with convection and frozen hydrometeors, the all-sky emissivity retrieval gives a strong decrease in retrieved emissivity with frequency, often greater than the -0.2 limit of the colour scale. Examples of this are most visible over North America, over parts of equatorial America and Africa. This is due to scattering signals from large frozen particles, rather than true surface emissivity differences. Scattering is rarely apparent at 19 GHz, but it results in a large drop in TB at 89 GHz, which is aliased into the emissivity retrieval if the model does not have a corresponding representation of the convection (Baordo and Geer, 2016). Large frequency gradients in both directions are also seen over Greenland and Antarctica, where they depend on the microstructure of the snow and ice surfaces. There are also spots with positive emissivity

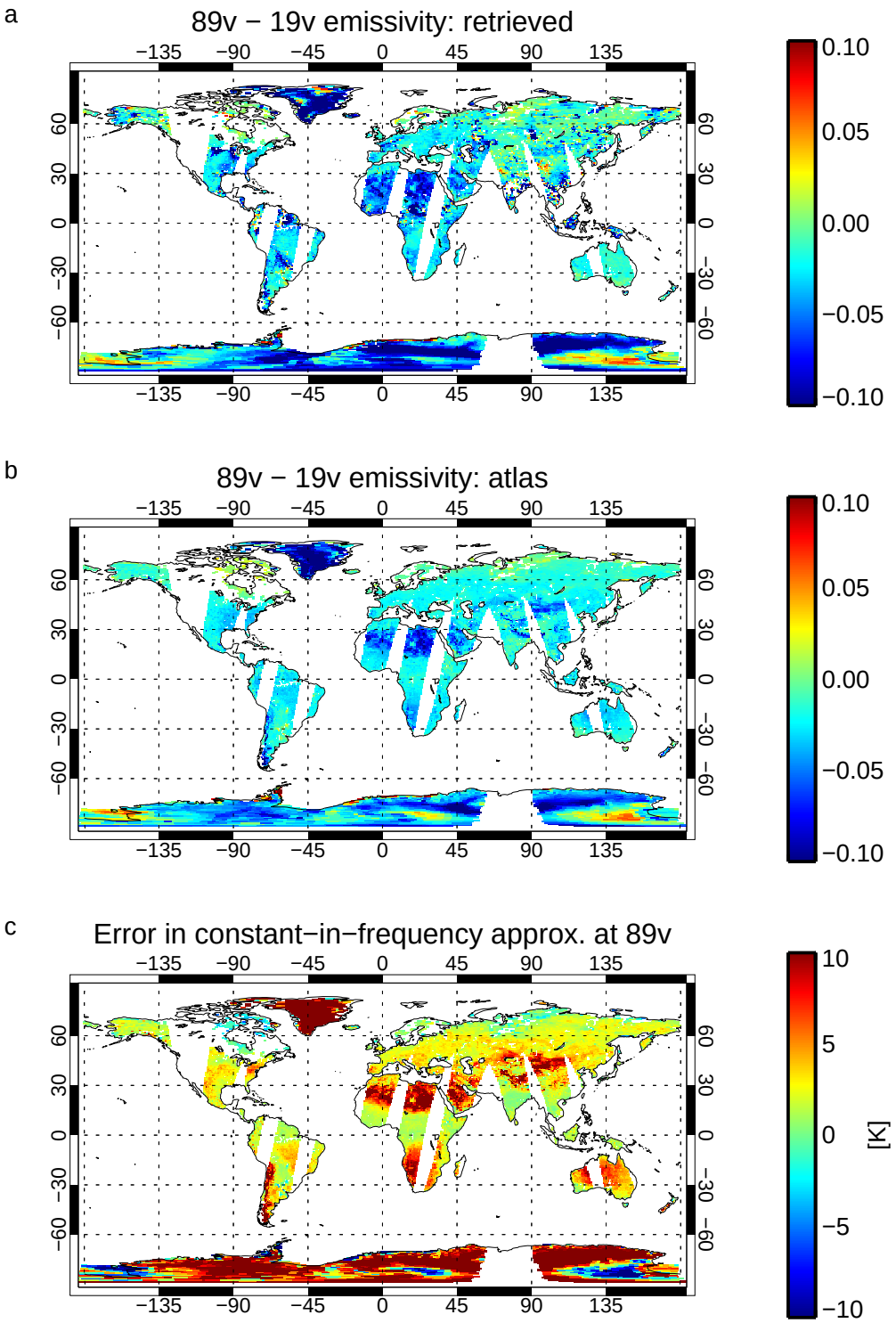

Figure 11: Difference in surface emissivity between 89v and 19v, based on (a) dynamic emissivity retrievals and (b) TELSEM2 atlas; (c) brightness temperature error when using the constant-in-frequency extrapolation from 19v to 89v, compared to using the atlas adjustment. Data from AMSR2, selecting land areas ($l_{\text{land}} > 0.95$) available at the 00 UTC analysis on 1st July 2019.

deltas reaching 0.05 over Canada and Asia, likely associated with snow cover, and possibly explained by wet snow (a water surface would also have a positive delta) or the presence of water on the surface more generally. Finally, dry soil areas such as the Sahara, southern Africa and Australia also have large changes with frequency that are likely to be associated with variations in penetration depth.

Figure 11b shows the atlas-derived emissivity delta, which does not show any contamination from hydrometeor scattering, since the TELSEM2 atlas was based on cloud-screened retrievals. It also shows fewer of the positive deltas, apart from parts of Canada and northern Siberia that might have a significant amount of surface water. The atlas-adjustment technique allows us to use the 19v dynamic emissivity retrieval in order to follow real-time changes in surface properties, but to apply a sensible adjustment to account for the frequency-variation of surface emissivity. As done earlier in Fig. 4, panel c uses the formula of English (2008, Eq. 8 here) to estimate the impact on 89v simulated brightness temperatures of using the constant-in-frequency approximation. This is the difference between taking the retrieved emissivity from 19v without adjustment, rather than the atlas-adjusted approach (Eq. 11), and so δe here is given by the atlas delta between 19v and 89v. The potential error exceeds 10 K in dry areas, where the skin temperature is inappropriate, and also in Greenland and Antartica. Most of these areas will be screened out in any case (see later) but even in areas that are assimilated, such as parts of the continental USA, the error reaches 5 K.

In practice, the atlas adjustment is only applied to vertically polarised channels, since Lonitz *et al.* (2022) saw that the the delta emissivities in TELSEM2 for horizontally polarised channels were less consistent with the frequency dependence of emissivity retrievals in the IFS. In any case, mid-frequency h-polarised channels are not part of the microwave imager assimilation over ocean. Hence the retrieval channel i in Eq. 11 is 19v and the assimilated channel is 37v, 89/92v and 150/166v wherever available (see Tables 1 and 3). The 150/166h channels are also activated over land in this development, but using the constant-in-frequency approach, with the 89/92 GHz h-pol channel as the retrieval channel, following the existing approach for 183 GHz channels over land. The 37h and 89/92h channels are not assimilated in any circumstances, but it is important to simulate these channels well, since 89/92h is used in some observation error models (e.g. SSMIS, which has h-polarised high frequency channels) and 37h and 89/92h are used in the snow screening for the atlas-adjusted retrievals (Sect. 5.3 below). Hence the 37h and 89/92h channels use the 19h dynamic emissivity retrieval with the constant-in-frequency approach.

The treatment of mixed scenes affected by possible snow or sea-ice is less advanced with this method. For $0.01 < l_{\mathrm{land}} < 0.95$, any mixed scene with low temperature (< 278.0 K), high latitude ($> 60^{\circ}$), or sea-ice ($l_{\mathrm{seaice}} > 0.0$) is removed. For the mixed scenes that are retained, there is no attempt to use FASTEM to control the frequency-dependence of the emissivity of the water part of the scene; however, the quality control of the emissivity retrieval is applied as for all 'liminal' scenes, based on the Eq. 5 retrieval of the land and sea-ice portion of the scene.

Using the atlas-constrained emissivity extrapolation, Lonitz *et al.* (2022) demonstrated the successful assimilation of microwave imager channels at 89 GHz and 150/166 GHz, focusing on GMI and SSMIS. However, they also noted issues with the bias correction, such that adding observations over land affected the bias correction of existing observations over oceans, and QC and observation error models were not fully developed. The current work develops the method further, and we report here on refinements that allowed activation in cycle 48r1. New QC and observation error models have been developed; the approach is upgraded for consistency with the developments of Sect. 4, and it is extended to AMSR2 and to the 37 GHz channels in order to be able to use as much data as possible, and to support future use of the rain and water vapour information over land (as discussed in Sect. 2). For economy of experimentation, these developments are combined with the activation of the GMI sounding channels over land and sea-ice, a capability that was not part of the initial implementation of GMI (Lean *et al.*, 2017).

Table 4: Bias predictors for GMI, AMSR2 and SSMIS. ω is a function of the scan position that emulates the scan angle of a cross-track sensor. x = Active for GMI, AMSR2 and SSMIS; * = Active for SSMIS only. Brackets indicate the predictors that have been added in the current work.

Predictor	24 GHz and below	37v, 37h, 89/92v, 89/92h	150/166v/h, 183 GHz
constant offset	x	x	x
l_{land}		(x)	
T_{skin}	x	x	
$l_{\text{land}}T_{\text{skin}}$		(x)	
TCWV	x	x	
10m windspeed	x	x	
ω	x	x	x
ω^2	x	x	x
ω^3	x	x	x
ω^4	x	x	x
1000-300 hPa thickness			x
200-50 hPa thickness			x
50-5 hPa thickness			x
10-1 hPa thickness			x
ascending-descending bias			*

5.2 Bias correction over land surfaces

Bias correction is provided in the IFS by VarBC (Dee, 2004) which simultaneously optimises the 4D-Var analysis and the observational bias correction. The initial approach of Lonitz *et al.* (2022) was to use the product of 19 GHz surface emissivity and skin temperature, $e(19\text{v})T_{\text{skin}}$, as a replacement for the T_{skin} bias predictor for the 89v window channel. However, despite the emissivity allowing some separation between land and ocean bias corrections, that approach resulted in significant changes in the bias correction over ocean. This also led to unwanted changes in analysed humidity over ocean that were some of the largest impacts of the whole experiment. The current work aims to ensure that window channels can be introduced over land without affecting the use of the observations over ocean.

Initial exploration showed that biases could be different by several Kelvin between land and ocean surfaces, and there is a different skin temperature dependence of the bias over land compared to ocean. As discussed later, a likely cause for these biases would be instrument error. Hence, the new VarBC bias predictors for the relevant window channels are the land fraction, l_{land}, and the product of this with the skin temperature, $l_{\text{land}}T_{\text{skin}}$. This provides an offset between the land and ocean bias correction, as well as allowing different skin-temperature dependencies (since the skin temperature, unmodified by the land fraction, remains a predictor). This should better isolate the bias correction over land from that over ocean. The emissivity-based predictor of Lonitz *et al.* (2022) has not been used here. Table 4 summarises the bias predictors used for the microwave imagers.

Figure 12 illustrates the effectiveness of the new bias correction in AMSR2 channel 89v. Before bias correction (panel a) the populations of departures are quite different over land and sea, and there is a small increase in bias as a function of skin temperature in the land population. Over ocean, there is no obvious skin temperature dependence. After bias correction (panel b), using the two additional predictors, the ocean and land populations are centred around zero and the skin temperature dependence has been removed. The land population is also clustered quite close to the zero line, suggesting that no

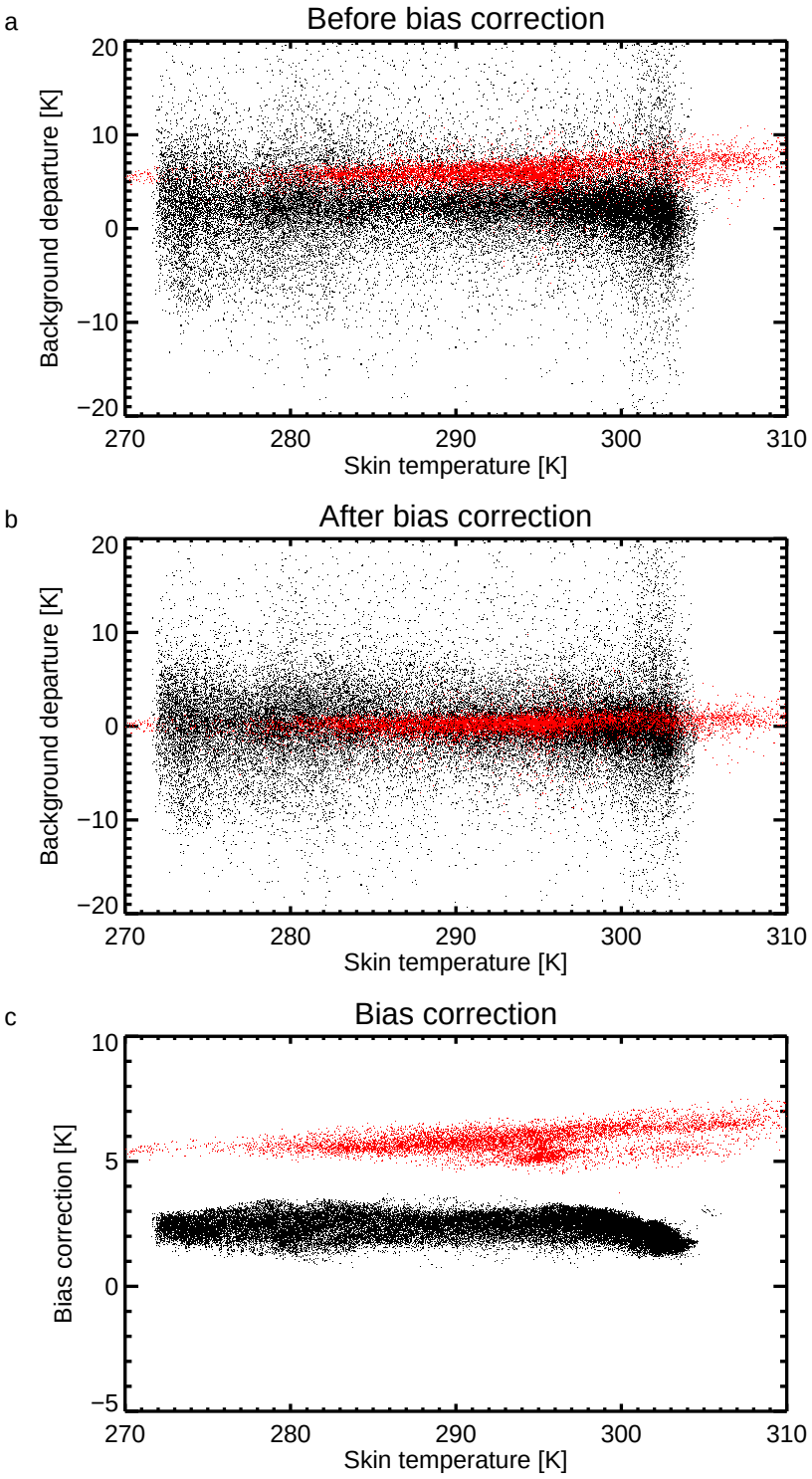

Figure 12: Scatter plots of AMSR2 89v background departures (a) before and (b) after bias correction, and (c) the bias correction itself. Points in black are over ocean (excluding sea-ice) and in red are over land ($l_{\text{land}} > 0.95$) after applying the quality control of Sec. 5.3. The sample of data comes from the 00 UTC analysis on 1st July 2019

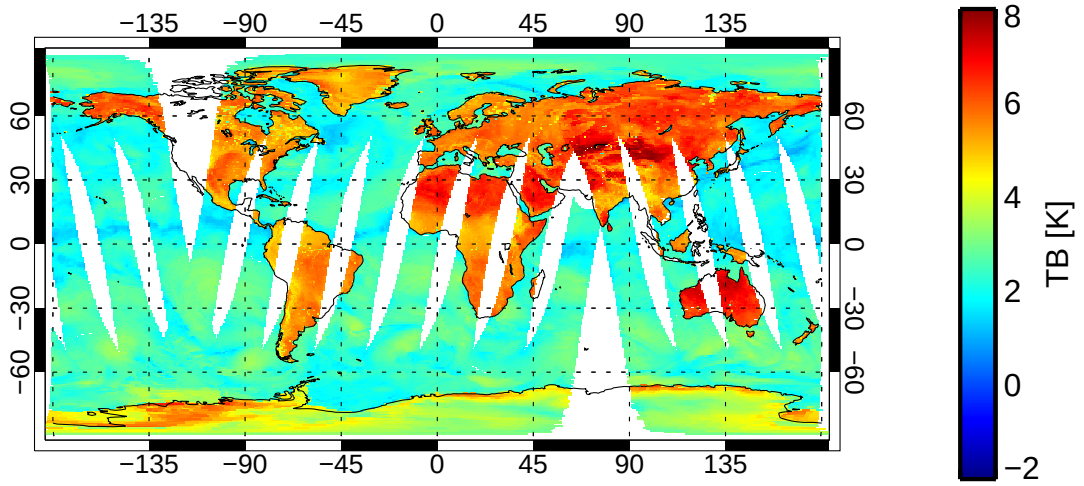

Figure 13: Bias correction generated for all AMSR2 89v data available to the 00 UTC analysis on 1st July 2019.

additional bias predictors are needed. Panel c shows the bias correction itself: over land, this ranges from about 5 K to 7 K as a function of skin temperature, whereas over ocean it is around 2 K.

Figure 13 shows the same bias correction on a map. The land bias correction has more complex features than would be expected from the skin temperature dependence alone. This is because other bias predictors such as TCWV are active over land and ocean (Table 4). Beyond the skin temperature, the most important predictor for variations over land is the TCWV. For example, the bias correction is lower over the N tip of South America, equatorial Africa, India and SE Asia; these are ITCZ and monsoon areas of elevated moisture and precipitation. The spatial patterns of these features in the land bias correction follow the TCWV map quite closely (not shown). The reduced land bias correction in high TCWV areas is also evident on Fig. 12c, where there is a population below the main distribution (centred on (295 K, 5 K)). Most likely the TCWV dependence compensates the inability of the IFS and RTTOV-SCATT to simulate enough scattering in tropical convection (Geer and Baordo, 2014; Geer, 2021b). This effect is common to land and ocean, so there is no urgent need to better separate the bias correction; also the addition of land observations has not much affected the ocean bias correction despite there being predictors that operate over both surfaces. Future developments could attempt to better isolate the ocean and land bias correction: for example the surface windspeed predictor is intended to correct windspeed-dependent errors in the FASTEM-modelled ocean surface emissivity, so it makes no sense for it to be active over land. However, windspeeds are generally very low over land surfaces compared to the ocean (not shown), so the impact of the windspeed-dependent bias correction term is unlikely to be significant over land.

Figure 14 completes the survey of the bias correction for the new surface-sensitive channels over land. AMSR2 channel 37v has very similar bias correction to AMSR2 89v, with a separation of about 4 K between the land and ocean bias corrections. However, for GMI and SSMIS, although the land bias correction is slightly higher than the ocean bias correction, this is by less than 1 K in most cases. The skin-temperature dependence is consistent across sensors, varying by around 1 K across the range shown in the figures. The large separation between land and ocean biases is therefore a feature of the AMSR2 observations, and can be associated with the large difference in scene temperature between ocean and land. The observed separation is mostly consistent with the assessment of the GPM intercalibration work-

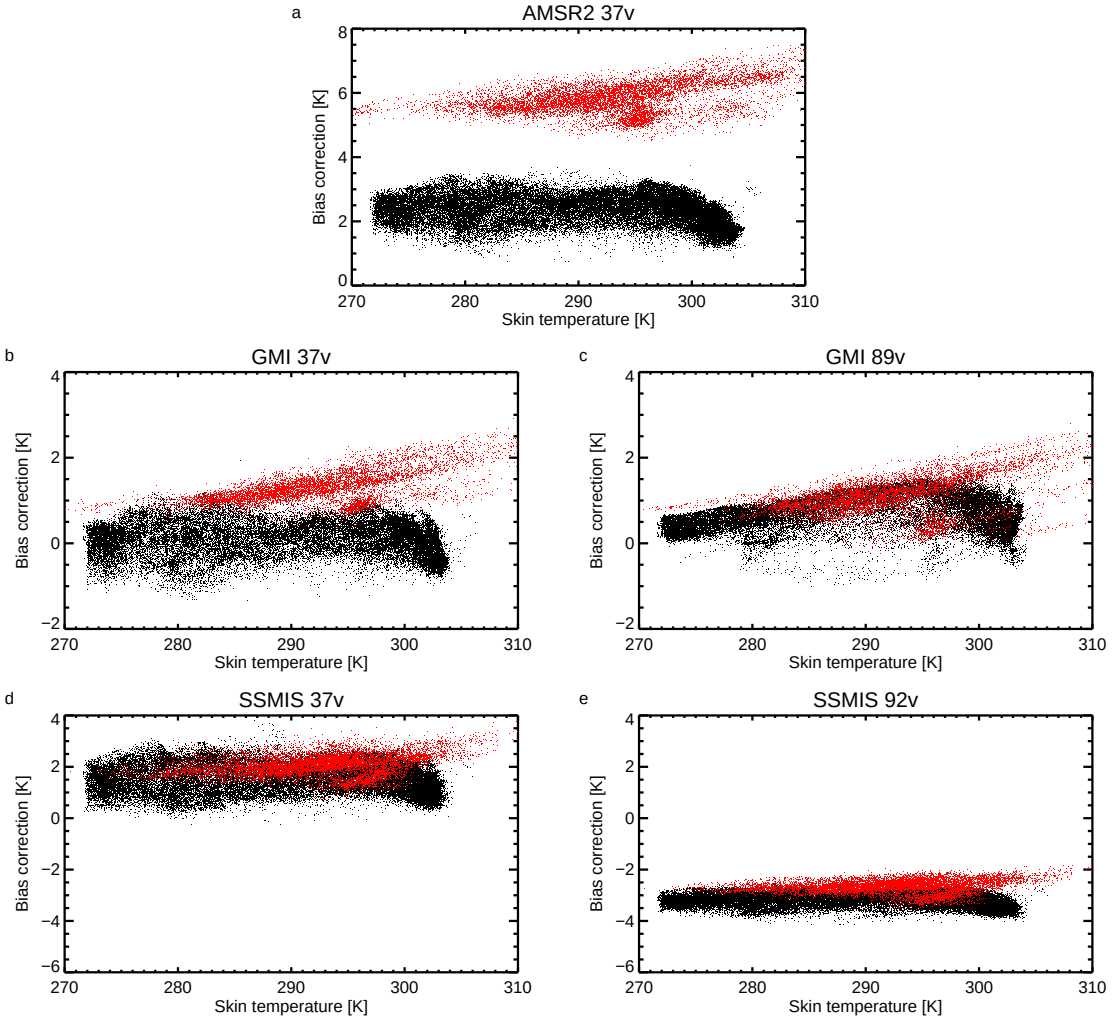

Figure 14: Scatter plots of (a) AMSR2; (b-c) GMI and (d-e) SSMIS F17 bias correction. Points in black are over ocean (excluding sea-ice) and in red are over land ($l_{\text{land}} > 0.95$) after applying the quality control of Sec. 5.3. The sample of data comes from the 00 UTC analysis on 1st July 2019.

ing group (XCAL), which also finds around 4 K difference in calibration between warm (land) and cold (ocean) scenes in most of the AMSR2 window channels (Berg *et al.*, 2016). As discussed in that work, a scene-temperature dependent bias could come from inadequate antenna pattern correction, emissive reflector issues, warm calibration target issues, or specifically for AMSR2, the nonlinearity correction (see also Gentemann and Hilburn, 2015). However, as far as assimilation in the IFS is concerned, the main thing is that VarBC is effective in correcting the bias (Fig. 12).

The skin-temperature dependence of the bias correction over land is harder to attribute, though it could also be caused by any of the instrument calibration issues discussed in the previous paragraph. However, since it is mostly consistent across sensors (Fig. 14) it is as likely to be an issue of the modelling or of the assimilation methodology. Given the complex processes involved in dynamic emissivity retrievals, the dependence on the bias correction in the emissivity retrieval channel, and the inadequacy of the IFS skin temperature as a description of the effective emitting temperature of the surface in warm (and typically dry soil) areas (e.g. Bormann *et al.*, 2017), it would be hard to diagnose the issue precisely.

The average bias correction over land varies strongly between the three sensors, from +6 K in AMSR2 to -3 K in SSMIS channel 92v. Apart from the issue with AMSR2 land-ocean contrasts, these inter-sensor differences are mainly consistent with the ocean bias corrections and consistent with bias patterns previously seen in the IFS (e.g. Lean *et al.*, 2017; Lawrence *et al.*, 2018). The inter-sensor differences are not so consistent with XCAL (Berg *et al.*, 2016), but the explanation is unknown and outside the scope of the current work.

5.3 Quality control, including snow and desert screening

In the following, we describe work to enhance the quality control for challenging surfaces, such as snow and desert regions. A first aim is to avoid the assimilation of observations in regions for which the representation of the surface contributions in the radiative transfer is currently too poor. Both the experimental work of Lonitz *et al.* (2022) and the operational assimilation of AMSU-A over land (Weston *et al.*, 2019) have shown suspicious mean changes in analysis and forecast fields over central and north Africa. It is thought that these arise due to using the model surface skin temperature as the effective emitting temperature in regions where the microwave radiation penetrates some distance into dry soils, thus leading to considerable biases in the background departures (Bormann *et al.*, 2017). We aim to remove these areas from active assimilation, until a better solution can be found.

A second aim was to be more ambitious with the coverage of the atlas-adjusted retrieval in winter. However, it is not possible to drop the snow screening completely. Early testing showed that snow-covered areas are not well-handled by the atlas-adjusted approach (Eq. 11). In one example, not shown here, in a large area of northern Siberia in June, the atlas was representative of years when there had been significant snow cover, and the frequency dependence of the atlas emissivity was characteristic of snow (dropping significantly with frequency between 19 and 89 GHz). In contrast, the actual situation in June 2019 was no snow cover, and the dynamic retrievals showed a more constant surface emissivity as a function of frequency. The result was systematic background departures of around 10 K in the affected region. A further motivation is to find something better than the pre-existing $T_{\mathrm{skin}} < 278.0\,\mathrm{K}$ 'possible snow' rejection criterion, which causes large losses of data in the northern hemisphere. The skin temperature is an imprecise way of identifying possible snow cover. Issues include the strong diurnal cycle, meaning that areas of the Sahara can be identified as 'possible snow' during the night time. Hence this work aimed to find alternative 'possible snow' criteria.

Investigation showed that a typical characteristic of dry desert soils and some snow cover is a polarised

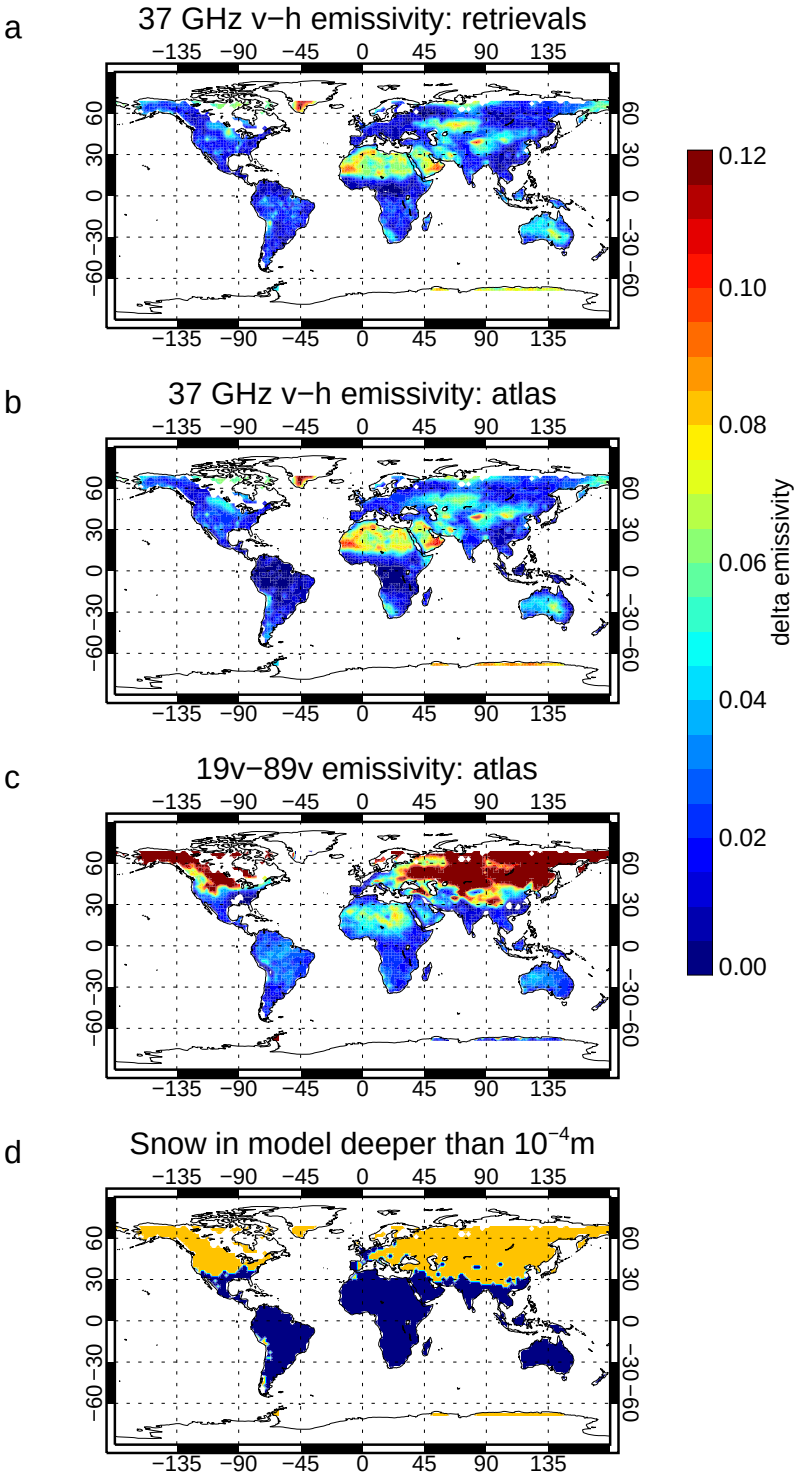

Figure 15: GMI screening parameters, averaged in 2° bins in January 2020, for land surfaces only ($l_{\text{land}} > 0.95$): (a) Difference in retrieved emissivity between 37v and 37h channels; (b) Difference in atlas emissivity between 37v and 37h channels; (c) Difference in atlas emissivity between 19v and 89v GHz channels; d) Locations where monthly average snow depth is greater than 1×10^{-4}m (orange) or less than or equal to this (blue).

surface emissivity. Figure 15 panels a and b, for January 2020, show 37v - 37h emissivity polarisation is on average higher than 0.04 over the Sahara, Arabia, the interior of Australia, all areas where dry soils or sand would be expected. There are also patches of strong polarisation in the NH extratropics that may be caused by certain types of snow cover. This polarisation is hypothesised to be a sign of volume scattering within the surface media; certainly in the atmosphere, preferential horizontal alignment of frozen particles leads to similar polarisation of optical properties (e.g. Gong and Wu, 2017; Barlakas *et al.*, 2021). However, the surface polarisation is not sufficient to detect snow cover more generally. In the atlas, potential snow cover can be identified approximately by a strong decrease in emissivity between 19 and 89 GHz, which could indicate volume scattering from mm-sized particles (Fig. 15c). The model's snow depth gives a more direct way of identifying snow in the model (Fig. 15d) although trial and error showed that a threshold as small as 1×10^{-4} m was needed to comprehensively identify snow-covered areas in the model. Based on these investigations, if any of the following checks are failed, any channel using the atlas-adjusted emissivity retrieval is rejected over land:

- scenes with 37v - 37h polarisation of the surface emissivity greater than 0.04, in either the atlas or the retrievals, are rejected as likely to contain desert or snow;
- scenes with 19v - 89v surface emissivity estimates greater than 0.04 in the atlas are rejected as 'snow in atlas';
- scenes with model snow depths greater than 1×10^{-4} m are rejected.

A final change in QC compared to Lonitz *et al.* (2022), and different also to the use of water vapour sounding channels, was to relax the orography checks on window channels. The previous approach rejected orography above 800 m in all channels. This makes sense for sounding channels, since high orography increases the chance of the surface being visible. However, since the surface is already visible in most window channels, there is less justification for an orography check. However, it was empirically found that areas with altitudes greater than around 1600 m, for example in southern Africa, tended to have systematically higher background departures, even after the snow and dry-soil QC listed above. Hence, 1600 m was adopted as the rejection threshold for window channels.

Figure 16 illustrates the effect of the QC on GMI observations in July and January. Starting with the available observations over non-mixed land surfaces ($l_{\text{land}} > 0.95$; panels a and e) it can be seen that GMI has the highest density of observations around 60° latitude, with a very similar coverage in winter and summer, as expected. The high density around 60° comes from GPM's inclined orbit. Panels b and f show the results of applying the QC described in this section. Common between July and January, there is almost complete loss of data over the Sahara, Arabia, the Australian interior, and the SW extremes of southern Africa and south America, primarily due to dry soil QC. The orographic QC removes areas of the Rockies, Andes and Himalayan region and is obviously constant with season. The big difference between July and January is the QC for snow: in July, almost all available observations can be used in the NH midlatitudes; in January, they are mostly lost, except for areas of the USA, Europe and China which were less affected by snow during the year studied.

The lower panels of Fig. 16 show the RMS of the background departure as an indicator of both possible signal and possible errors. In July, the RMS background departure is greater than around 5 K in near-equatorial regions of South America and Africa, due to the position of the ITCZ; over India and SE Asia, due to the monsoon; and in areas of the midlatitudes (such as central North America and areas of Europe and Asia), due to summertime convection. These are all signals that we would like to retain in all-sky assimilation, and indeed most signals are retained after the surface QC (compare panels c and d). The main area where signal is lost is the Sahel region, where dry soil QC appears to remove convective

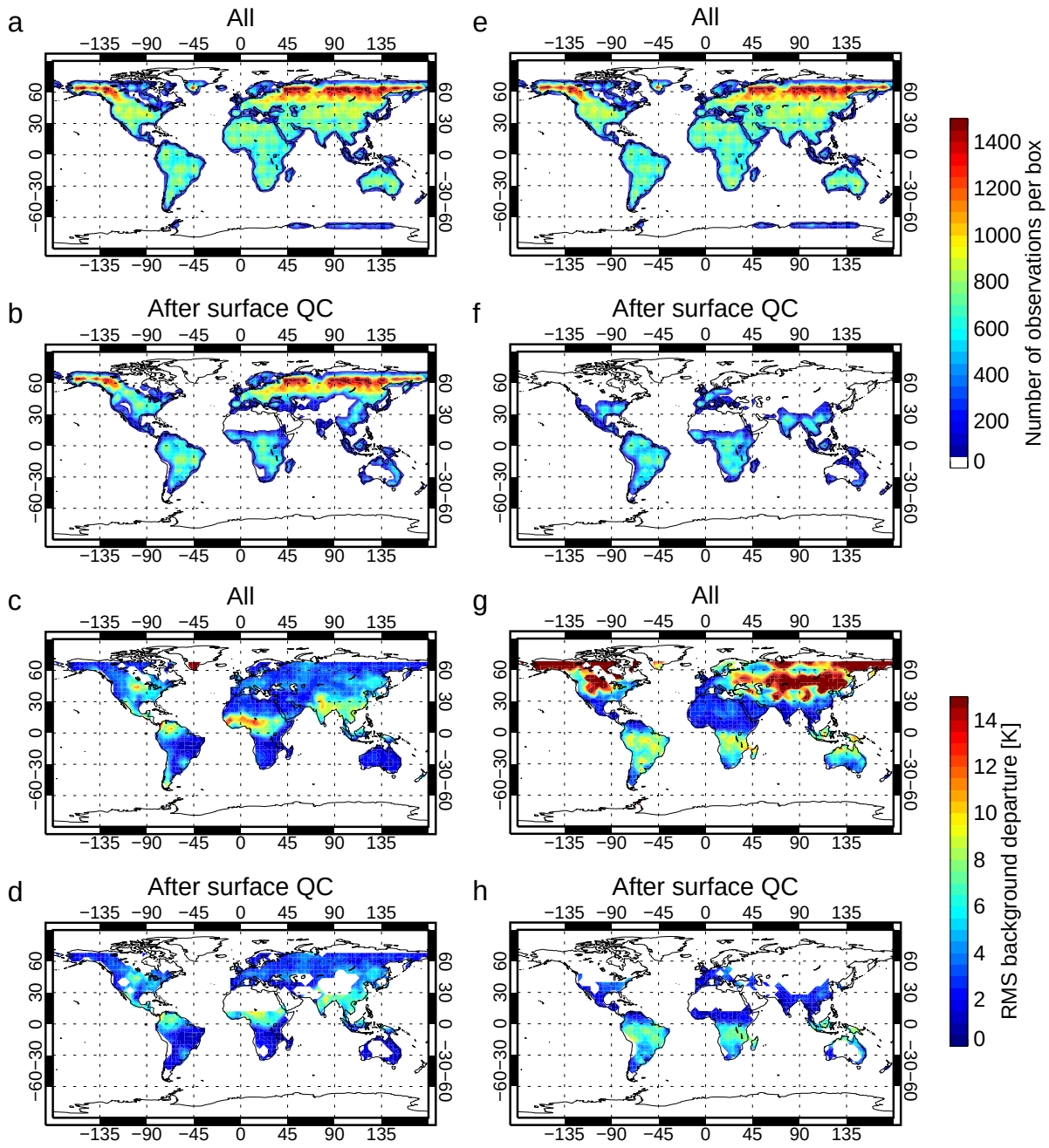

Figure 16: GMI usage in channel 89v, over land surfaces ($l_{\text{land}} > 0.95$) for July 2019 (left, a-d) and January 2020 (right, e-h). The top four panels show the number of observations and the bottom the RMS of the background departure, using bins of 4 by 4°. 'After surface QC' is the observations remaining after checking that the 19 GHz retrieval is valid, after applying surface emissivity checks (e.g. snow and desert surfaces) and after removing orography over 1600 m. White areas have minimal or no data.

scenes. In January, the RMS departures are often very large (>15 K) in snow-covered areas, where the dynamic emissivity retrieval does not work well. The snow QC is effective in removing these areas (compare panels g and h). Most of the convection signal (in the ITCZ over South America and southern Africa) is retained, however. In panel g, there are RMS departures reaching around 5 K in the Sahara region. These are highly unlikely to come from cloud or precipitation signals, so this is evidence of the surface emissivity and skin temperature problems associated with dry soils (e.g. Bormann *et al.*, 2017). However, the errors appear to be smaller than those coming from snow surfaces, and the dry soil QC eliminates a larger area than those with obviously elevated departures. Hence, there is scope to refine the dry soil QC in future (or, better, to improve the representation of the surface in these areas, so that they might be assimilated).

5.4 GMI 183 GHz over land and sea-ice

As mentioned, this work also adds GMI 183 GHz channels over land surfaces and over possible sea-ice for the first time. Following the standard approach used with sensors like MHS and SSMIS, both 183±3 and 183±7 GHz are assimilated over land, but just 183±3 is assimilated over possible snow or sea-ice. As usual, the dynamic emissivity approach is used for these channels, taking either the 89v or 166v retrieval over snow-free land and possible snow/sea-ice respectively. This ensures that processing and QC approaches applied to GMI are consistent with other similar channels. However, it means that the GMI QC is now different depending on whether constant-in-frequency or atlas-adjusted retrievals apply in a particular channel (see Fig. 17 in the next subsection; the QC for the atlas adjusted channel, 166v, is significantly more cautious). Note that use of the GMI 183±7 GHz channel over polar oceans, avoiding sea-ice, is unique new data, as this is the only sensor with both a 183±7 GHz chanel and a 10 GHz channel for sea-ice screening. However, GMI only observes up to about 67° latitude due to its inclined orbit.

5.5 New error models over land surfaces

Examination of background departure patterns over land surfaces suggests that even outside of cloud and precipitation areas, and even after quality control, the departures can be large in some places. For example at 37 GHz, there are areas with departures of around 3 K, although the majority of areas have much smaller departures. Since these departures are most likely explained by residual errors in the surface representation, they should not be assimilated with much weight. We are not expecting the atmospheric signals in these channels to be so large in clear-sky areas. In an ideal world we would develop a situation-dependent error model for the surface, but this would be a major piece of work and it is not attempted here. Instead, a rough estimate of the largest instantaneous departures in clear-sky areas over land is used to guide the choice of a constant clear-sky error in the error model. With the aim to make the normalised departure small enough that it would not strongly affect the analysis, a target of no more than 1/3 normalised departure was chosen. For the example of the 37 GHz channel, that would require a clear-sky error over land surfaces of 9 K. This means that in practice, we can only really extract information in cloudy and precipitation areas, where the departures can be much larger than this.

Rather than attempt a cloud-driven observation error model in all cases (e.g. Geer *et al.*, 2010, 2014; Duncan *et al.*, 2022), in channels where the clear-sky error is similar to the one that might be applied in cloudy scenes, it makes sense to use a constant error. Hence the following configuration is applied:

- 37v channels on GMI, AMSR2, SSMIS F17 have a flat 9K observation error to minimise the

impact of estimated O(3K) maximum surface-related errors due to skin temperature, volume scattering etc.;

- 89v channels on GMI, AMSR2, SSMIS F17 have a flat 15K observation error, due to O(5K) maximum surface-related errors;
- 166v and 166h on GMI use an SI-driven error model but with a high clear-sky error (15K) for similar reasons; maximum cloudy errors are 35.0 and 40.0 K respectively;
- 183±3 and 183±7 GHz channels on GMI use a quadratic situation-dependent error model, using 90 - 150 GHz SI, similar to the treatment of SSMIS and MHS (Geer *et al.*, 2014). For 183±3, errors vary between 2.0 and 25.0 K and for 183±7 from 3.0 to 35.0 K;
- Since 150h on SSMIS F17 uses the 89h dynamic emissivity retrieval directly, it has remaining significant O(7K) errors over more difficult land surfaces. Hence a baseline 20K error is applied in clear-sky areas, alongside an SI-based increase in areas of particularly strong SI.
- Observation errors in the 183 GHz channels of SSMIS F17 have been re-fitted, bringing them downwards and more into line with MHS and GMI. For 183±1, 183±3, and 183±7 clear-sky errors are 2.5, 2.5 and 3.0 K and cloudy-sky maximum errors are 12.0, 25.0 and 32.0 K.

For the 6 new channels assimilated over land from GMI, Fig. 17 shows PDFs of the background departure normalised by the observation error. These are not the assimilated samples, but the sample of data passing basic surface emissivity retrieval, snow, and orography checks appropriate to each channel. The sample is divided into clear and cloudy scenes, with cloud being identified if SI$>$ 0.0 K in either the observations or the background, or both. The left and right columns compare the situation in July and January, where the primary difference is that there is around 2 - 3 times more data in NH summer. In summer, data usage extends to the Arctic ocean, but in winter it is constrained mainly to the tropics and SH, due to the screening of surfaces that may have snow.

In the clear-sky sample on Fig. 17 (red) and for imaging channels (37, 89 and 166 GHz), the PDF is very narrow compared to the expected Gaussian, suggesting these observations have little weight. This is a natural result of using a high minimum error in these channels (9.0 to 15.0 K). Only the cloudy scenes (black) generate significant numbers of normalised departures greater than ±1, which would be expected to have stronger influence on the analysis. However, at 37 GHz there seems little difference between the clear and cloudy samples in winter (panel g) though in summer (panel a) there are at least some departures greater than 1, likely coming from convective situations. The PDF of cloudy observations is closest to the intended Gaussian for the high frequency channels (166 and 183 GHz), where the cloud-dependent error model has been used. Only at 183 GHz, with smaller observation error in clear skies (2.0 to 3.0 K) do the clear observations generate relatively large departures. This can be considered a cautious first approach using a new assimilation technique. Hence, at low frequencies only the bigger cloud and precipitation signals are expected to affect the analysis.

6 Results

6.1 Experimental setup

Experiments (Table 5) are based on cycle 47r1 of the IFS (ECMWF, 2020) but with a number of additional developments that went into cycles 47r3 and cycle 48r1. On top of 47r1, RTTOV has been

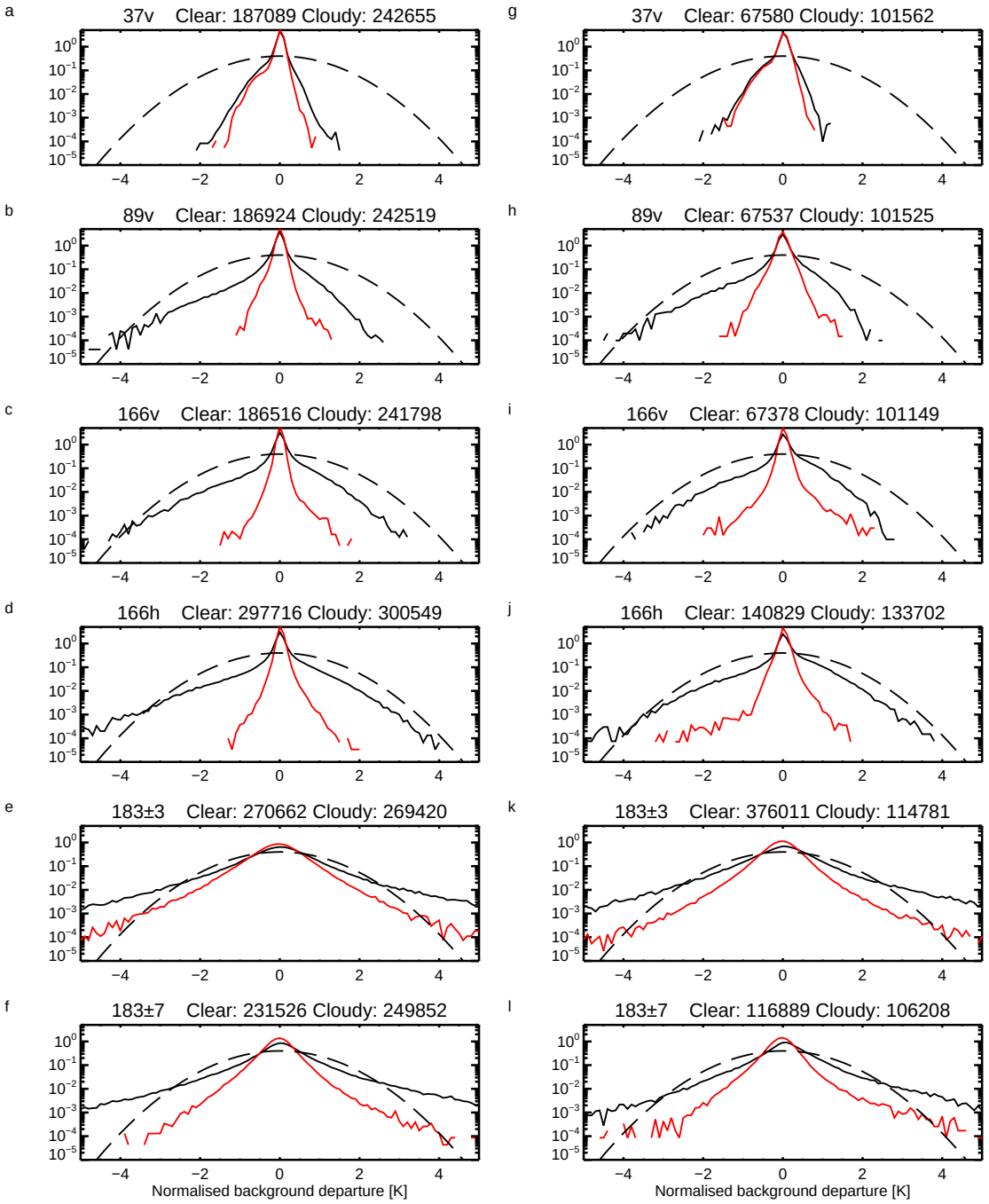

Figure 17: PDF of GMI background departure normalised by observation error, for a cloudy sample (black) or clear sample (red), over land surfaces ($l_{\text{land}} > 0.95$) for July 2019 (left, a-f) and January 2020 (right, g-l). The dashed line shows a standard Gaussian.

Table 5: Experiment details

	Summer	Winter
Dates		
Start	00 UTC 20 June 2019	00 UTC 1 Sept. 2019
Verification start	00 UTC 25 June 2019	00 UTC 6 Dec. 2019
End	12 UTC 30 Sept. 2019	12 UTC 31 March 2020
Experiment IDs		
Control	hk2k	hk2l
Stage 1	hkbx	hkby
Stage 2	hkhg	hkhh

upgraded to version 13.0 (Saunders *et al.*, 2020) including the new RTTOV-SCATT configuration (Geer, 2021b; Geer *et al.*, 2021). The all-sky configuration has been extended to include AMSU-A (Duncan *et al.*, 2022) and slant path interpolation has been added for all-sky sounders MHS, MWHS-2 and AMSU-A following Bormann (2017). Additionally, ATMS usage (in the clear-sky path) has been improved over snow and ice surfaces (Bormann, 2021). This forms the control configuration. Two sets of experiments are performed, first using the Stage 1 upgrades to surface-sensitive channels (Sect. 4). The second set adds the Stage 2 upgrades on top of Stage 1 (so it tests both together). Stage 2 additions are primarily the assimilation of microwave imager channels over land and improved use of GMI humidity sounding channels (Sect. 5). The experiments are based on a standard summer and winter test period, with a 5 day spinup period discarded from the statistics to allow VarBC to adapt to the new configuration. Experiment horizontal resolution is TCo399 (29 km) and there are 137 terrain-following (hybrid-sigma coordinate) levels in the vertical.

6.2 Changes to observation usage

Figure 18 gives an overview of the changes by channel ID, grouping together the results from all sensors (as in Table 1). The largest relative gains, of around 20 - 33%, are in channels 37v, 89/92v, 150/166v and 150/166h. Around 5 percentage points of the increase in these channels come from Stage 1 and the rest from Stage 2. Figure 19 breaks down the numbers by sensor, showing that these gains are from the sensors AMSR2, GMI and SSMIS. Figure 20, panels a - d, maps the changes in spatial coverage in a few channels, focusing on NH summer (July 2019) where land coverage is at a maximum. The Stage 1 gains come from the use of window channels over high latitude oceans, and the Stage 2 gains (which dominate) come from adding selected window channels over land. The main gaps over land in summer come from high orography and dry soil screening; in winter the usage over the NH landmasses is much reduced due to possible snow (Sect. 5.3).

The next largest gains are an increase of 15% in channel 183±7 (Fig. 18). This comes from a wide variety of developments affecting most sensors (Fig. 19). The changes, also summarised on Table 3, include the addition of GMI over land, the better use of coastal scenes, and additional usage over high latitude land. The additional data in GMI 183±7 is not shown, but it is similar to the changes in GMI 37v, shown in Fig. 20d, which has data over land surfaces for the first time, as well as over high latitude oceans. Fig. 20h shows the gains in MHS channel 183±7. Much data is added around coastlines and islands, due to improvements in using mixed scenes (Sect. 4.1). This includes the removal of the coastal screening at lower resolutions and the use of coastal scenes for the first time. Areas with many lakes, such as Canada, also gain a lot of data. Finally, the land regions poleward of 60° latitude are activated

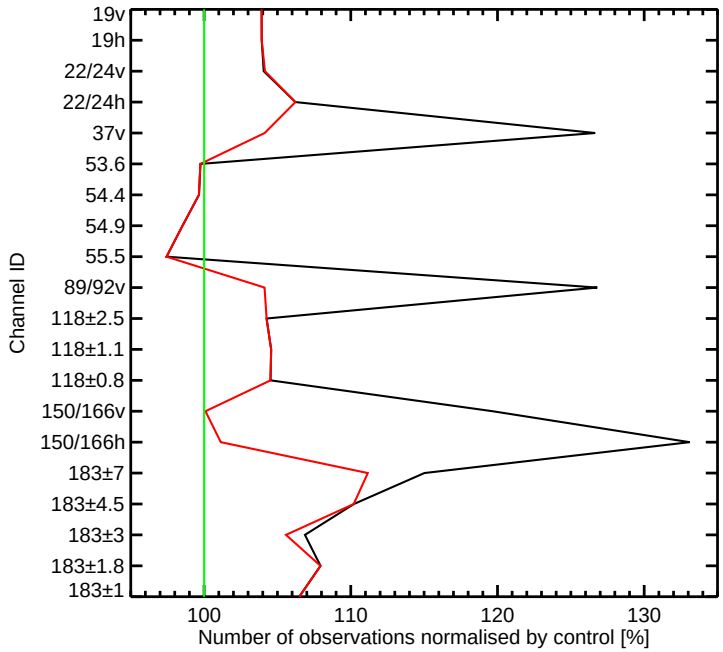

Figure 18: Relative number of observations per channel ID at control (100%, green) and in Stage 1 (red) and Stage 2 (black).

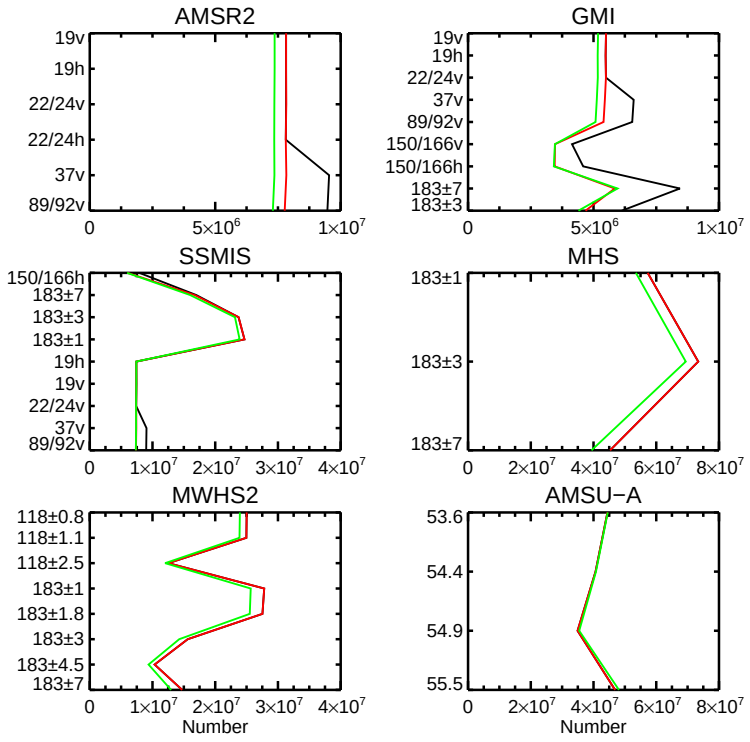

Figure 19: Number of observations per instrument and per channel ID at control (green) and in Stage 1 (red) and Stage 2 (black).

for the first time. The 183±4.5 channel, only on MWHS-2, is mostly equivalent in terms of screening to 183±7, and gains data in a similar way to MHS 183±7 (Fig. 20d) and this amounts to around 10% extra data overall (Fig. 18).

Higher peaking 183 GHz channels (183±3, 183±1.8, 183±1) gain around 6% to 8% more data. Examples are shown from GMI and MHS (Fig. 20f and j). Apart from GMI, these channels were already used over high-latitude land and ocean and sea-ice, and some coastal scenes were already added following the work of Weston *et al.* (2017). One gain is the addition of 183±3 from GMI over land and sea-ice for the first time (Fig. 20f). The other gain is because coastal scenes in sea-ice and possible snow areas were still subject to the inner-loop land fraction screening, which has only now been removed. Fig. 20j illustrates the additional coverage from that change, which mainly affects high latitudes, including the Antarctic coastline.

Gains of around 3 - 4% are seen in most remaining channels, e.g. 19v, 19h, 22/24v, 22/24h and the 118 GHz channels. For the low frequency window channels, the gains come from using ocean areas polewards of 60° latitude on sensors with a 10 GHz channel (AMSR2 and GMI). These patterns are similar to Fig. 20b and d but without the additions over land. For the 118 GHz channels, the lowest-peaking channel 118±2.5 gains in a similar way to MHS 183±7, particularly around coastlines (Fig. 20h), whereas the higher-peaking channels, including the stratospheric channels that have not been otherwise discussed here, gain in the same way as MHS 183±7, mainly around high-latitude coastlines, and for the same reasons.

The only channel IDs for which usage gets smaller, by up to 3%, are the 50 GHz oxygen sounding channels of AMSU-A (Fig. 18, 19). Figure 21 shows the changes on a map. For the two lowest sounding channels (53.6, 54.4 GHz, AMSU-A ch. 5 and 6) gains and losses are almost equally balanced. The losses, over parts of the Antarctic sea-ice, Canada and parts of Siberia and central Asia are mostly due to the stricter QC of mixed scenes, constraining the surface emissivity of the water fraction (Fig. 6). These were attributed (Sect. 4.1) to the relatively large 48 km FOV of AMSU-A and the large discrepancy between the true land fraction in the FOV and the IFS model grid-box land fraction, which are likely very different. This could probably be improved by introducing the FOV-appropriate land fraction, as is already done for the microwave imagers. The losses are slightly smaller in channel 5 because it has a specific screening of the Antarctic sea-ice; channel 6 sees slightly larger losses because the Antarctic sea-ice region is intended to be assimilated.

The gains in AMSU-A channels 53.6 and 54.4 (chs. 5 and 6) are over regions of fractional Arctic sea-ice, and come indirectly from the improved treatment of surface emissivity in channel 52.8 (AMSU-A ch. 4; Sect. 4.1, Fig. 7b). This is because, for AMSU-A only, the background departures in channel 52.8 are used for additional surface QC over sea-ice and snow. If the absolute departure is greater than 0.7 K, the tropospheric sounding channels (chs. 5-7) are rejected in these areas. One thing that does not show up on Fig. 21 is the change for channel 5 and 6, which allows mixed land scenes with land fraction 0.01 to 0.5 to be assimilated for the first time. This does add some scenes, but it is not visible on the figures, partly due to the strict QC of mixed scenes with low land fractions for AMSU-A. Further, these additions are balanced by the stricter QC of coastal scenes (Fig. 5).

AMSU-A channel 7 shows a greater loss than the lower-peaking channels (Fig. 18, 19). It does gain some scenes over the Arctic sea-ice, as for channels 5 and 6, but it loses data over land, Antarctic sea-ice and coasts. The coastal losses are clearer to see (not shown) because unlike the lower-peaking channels, there is no compensating gain from the relaxed coastal screening. Channel 8 shows the largest losses of all (Fig. 21b and d) because unlike lower peaking channels, it does not use the surface QC based on channel 4 departures. Hence the only changes are the losses due to stricter surface QC in mixed scenes

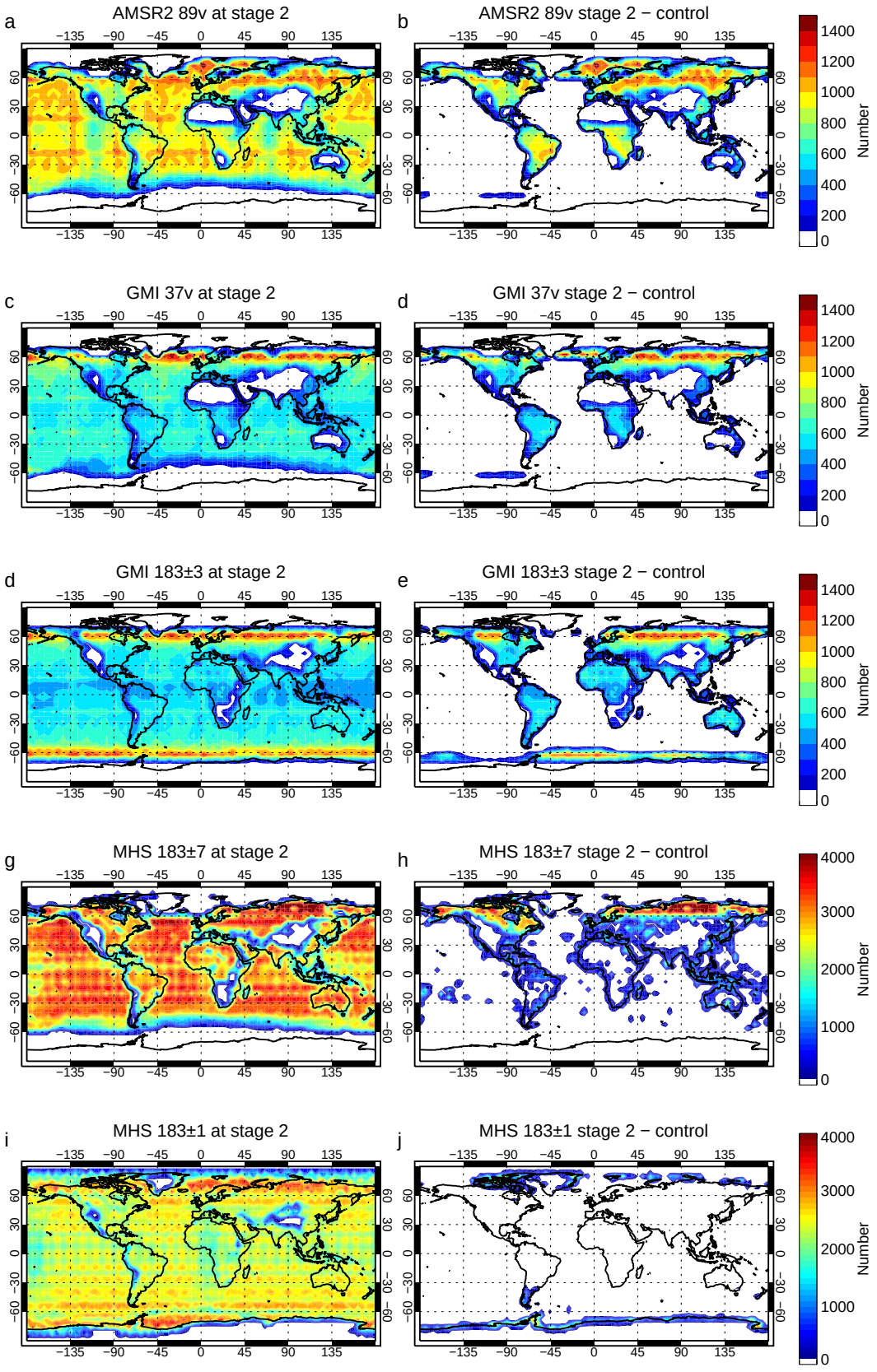

Figure 20: Number of observations (left) actively assimilated in the Stage 2 experiment; (right) the increase compared to control; per 4° latitude box, in July 2019. Selected channels and instruments.

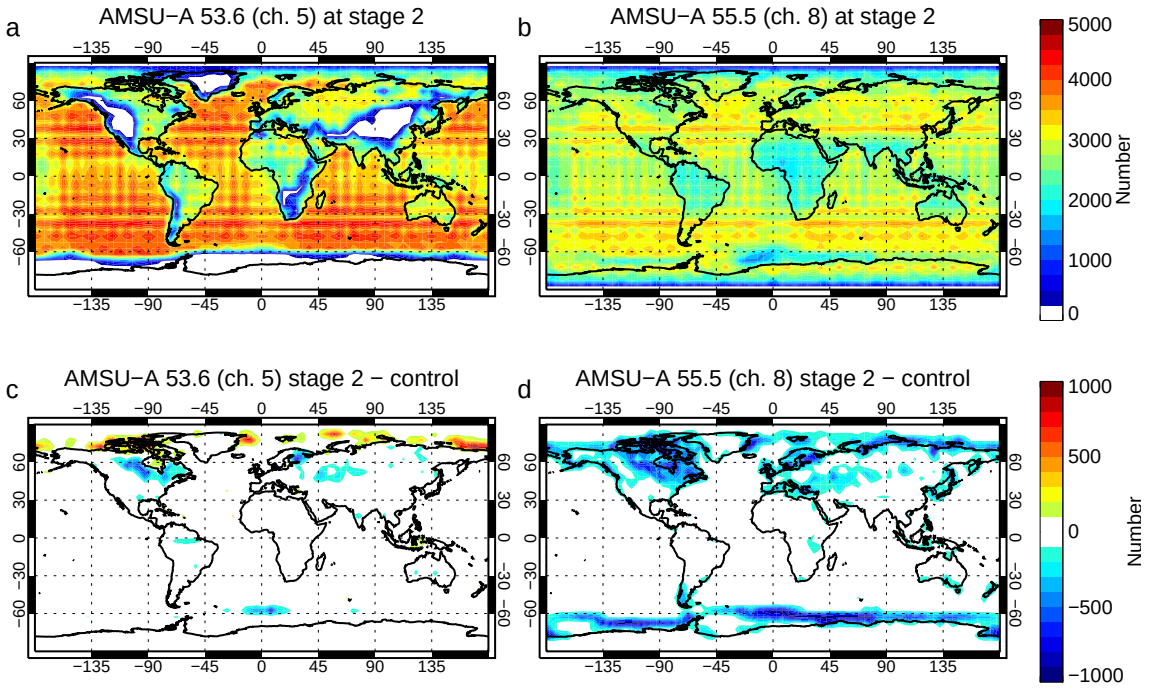

Figure 21: Number of AMSU-A observations (top) actively assimilated in the Stage 2 experiment; (bottom) the change compared to control; per 4° latitude box, in July 2019.

(panel d). It is debatable whether this is the correct behaviour, since channel 8 has almost no surface sensitivity. Additional experiments were run in which channel 8 was treated like a stratospheric channel, with no surface-related QC, but the results were inconclusive (not shown).

A final point relating to AMSU-A is the improved background departure standard deviation for the directly assimilated data, particularly in channel 5 where it is 4.5% smaller globally (Fig. 22). This comes purely from scenes with land fractions greater than 0 and less than 0.01, which are almost completely ocean but were previously, and incorrectly, given the 100% land surface emissivity (Sect. 4.1, Fig. 5). Although the number of these coastal scenes is not large, the errors have been reduced by around 2 K in places (not shown). Since this channel typically has 0.35 K background departure standard deviations, these improvements in a small fraction of coastal scenes are large enough to be able to influence the global background departures. This is a good reminder of the important surface sensitivity in AMSU-A channel 5 (e.g. English, 2008). Since channel 5 is a highly important channel for the forecast quality (being sensitive to mid-tropospheric temperature) this encourages further work to improve the surface emissivity modelling of AMSU-A.

One further point of note is that all the experiments have, by necessity, been carried out at lower resolution (Tco399/29 km, with coarsest inner-loop resolution of 125 km) rather than the operational resolution of Tco1279 (8-9 km with coarsest inner loop resolution of 80 km). As noted by Weston *et al.* (2017), the results will be slightly different in the full-resolution context. For example, the high resolution configuration would not see so much benefit from the removal of the inner-loop land fraction checks, since fewer observations would have been lost with an 80 km inner loop resolution than with 125 km. Higher resolution experimentation has been run, but not in the same configurations as used in this study, and only for a short period. These experiments are not analysed in detail here, but they show that the higher resolution configuration has no effect on results from AMSR2, GMI and SSMIS (which use the FOV-appropriate land fraction, not the model land fraction). For most MHS and MWHS2 channels

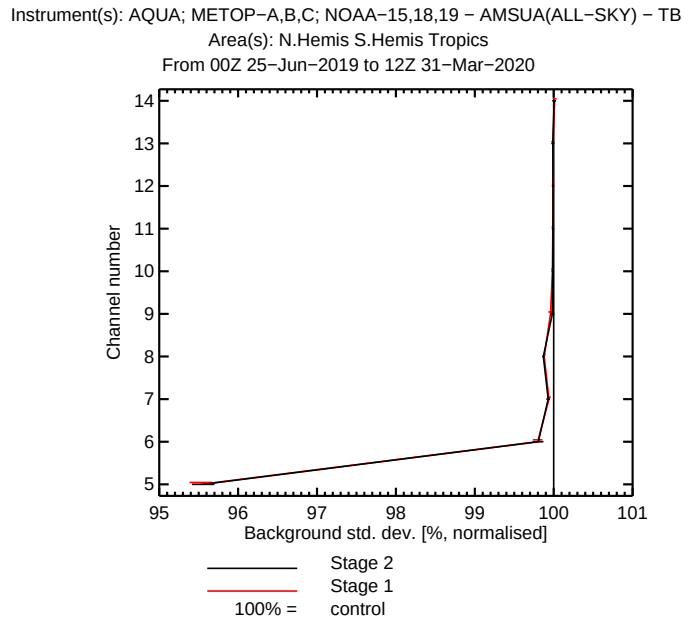

Figure 22: Relative change in standard deviation of AMSU-A background departures at control (100%) and in Stage 1 (red) and Stage 2 (black). The number of observations in each sample differs by up to 2.5% according to the changes in usage shown in Fig. 19.

the gains in observation numbers are 20 - 40% smaller, likely mainly due to the inner-loop resolution differences. However, since the gains in these sensors were already fairly small (Fig. 19) this does not much affect the headline results of this study. Further, for AMSU-A the loss of data in channels 54.9 and 55.5 (7 and 8) is smaller, which is not yet explained but is outside of the scope of this report to investigate. Overall, testing in the high-resolution operational context suggests that the current results are robust.

6.3 Changes in forecast and analysis

Figure 23 shows zonal cross-sections of the normalised change in RMS forecast error from T+12 to T+96, comparing all developments to the Control (Stage 2 includes all Stage 1 developments). These results combine both seasons of testing (Table 5). Both geopotential height and vector wind errors are reduced in the troposphere out to around day 3 (T+72), mainly polewards of 60°N. Improvements also seem to propagate to the NH stratosphere, with possibly longer duration of impact. The statistical significance testing in the stratosphere is likely to be less reliable (Geer, 2016) so this result should be treated with caution, but it has appeared in other testing too (not shown). The contributions to the reduction in RMS error come from reductions in both the mean and standard deviation of error (not shown). The impact on hemispheric scores, however, is not significant (not shown).

There are also some degradations south of 75°S in Fig. 23 which are significant mainly at day 3 (T+72). However, anything below around 700 hPa should be ignored as it is an artifact in extrapolating data onto sub-surface pressure levels. Also, these changes must have been caused indirectly by changes in observation usage around the Antarctic coast and sea-ice, since this work did not affect the usage over Antarctica itself. This points to a need to assimilate more tropospheric satellite data over Antarctica, which has the twin challenges of high altitude and a frozen surface, and is unlikely to be particularly

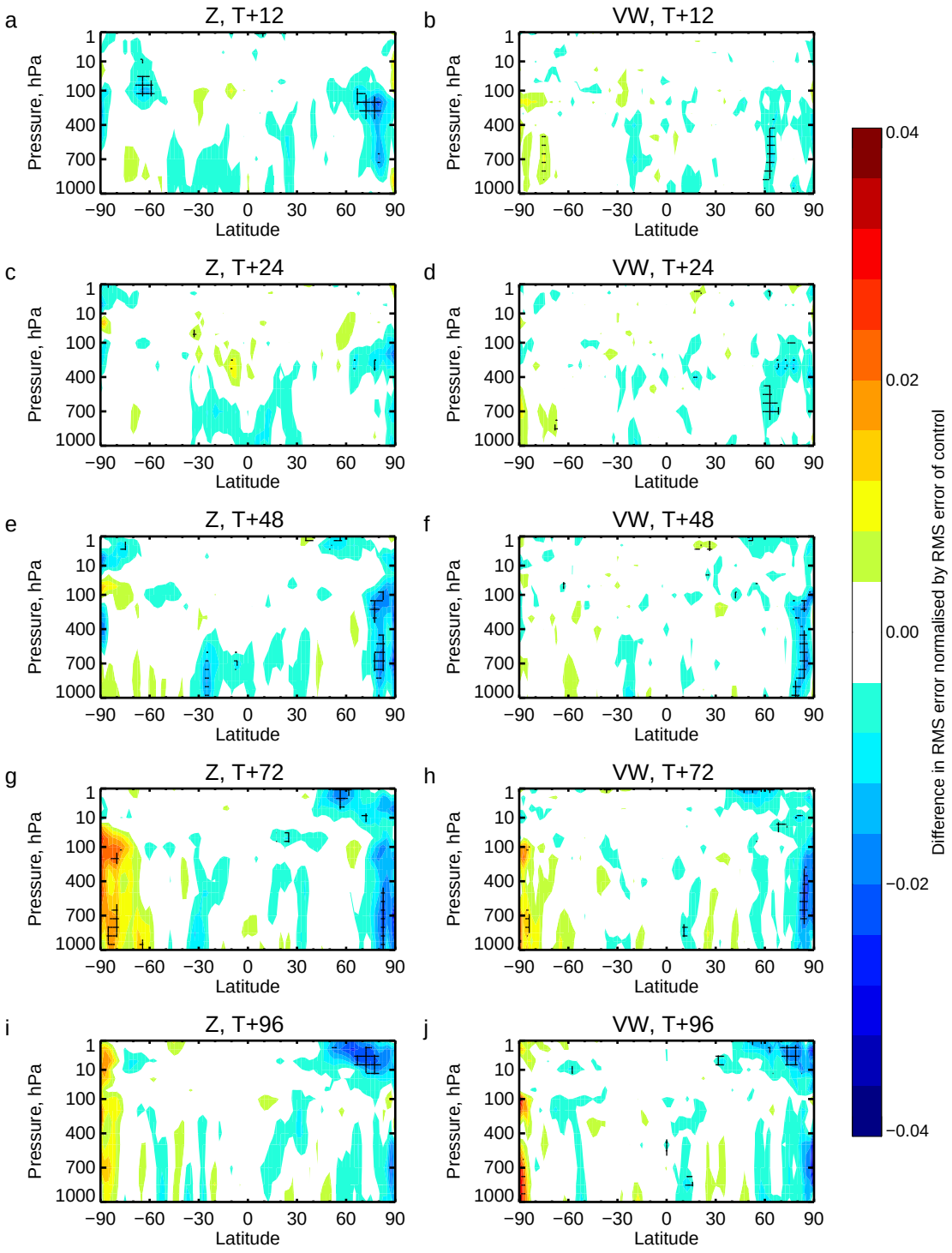

Figure 23: Normalised change in RMS error in Stage 2 forecasts (Z=geopotential height, VW=vector wind) compared to control, based on 390 to 428 samples, with cross-hatching indicating statistical significance to 95% using a Šidak correction for 20 independent tests (Geer, 2016). Verified against own-analysis. Remember that Stage 2 includes microwave imager assimilation over land surfaces as well as the Stage 1 changes.

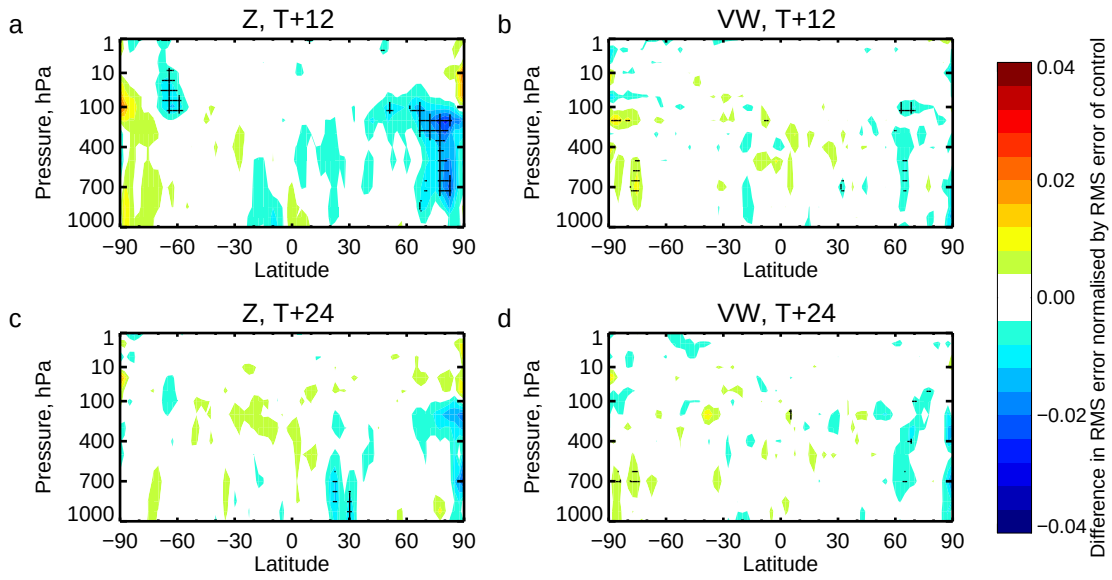

Figure 24: Normalised change in RMS error in Stage 1 forecasts compared to control; other details as Fig. 23.

well-constrained by observations at the moment. But to sum up, the main significant degradations are at 100 - 200 hPa at around 80°S at day 3, and are much outweighed by the improvements in the NH.

To attempt to separate the impact of the different stages of development, Fig. 24 shows the change in RMS error at T+12 and T+24 from the Stage 1 experiments compared to the control. At Stage 1 the impact is similar, perhaps smaller in the vector wind statistics. Broadly, it seems that both stages of development have contributed to the improvement in scores, but there is not statistical power to properly distinguish the two stages.

Figure 25 shows the impact on the relative humidity RMS error. This is elevated in the own-analysis scores, primarily over ocean, polewards of 60°, where a lot of extra window channel data has been added (these changes all come from Stage 1). The increased RMS error in relative humidity in the short-range forecasts is a typical signal of the assimilation of microwave imager observations, when measured by own-analysis verification (e.g. Geer and Bauer, 2010). It relates both to increased increment size and also to small mean changes in the humidity and temperature of the lowermost troposphere (not shown). However, the observation-based statistics show that these changes are beneficial, both globally and over the N pole (see below). Also the signal in own-analysis scores is mostly gone by day 3. Hence, this signal in the RMS relative humidity error (and in related fields and metrics) is not thought to be a problem.

Figure 26 summarises the global change in quality of the background (12 h) forecast as measured by the fit to other observations in the analysis. The clearest changes are in the fit to ATMS (panel a), with up to 0.1% improvements in fit to tropospheric and lowermost stratosphere channels (channels 6 – 10, channel IDs 53.6 to 55.5) and up to around 0.2% improvements in 183 GHz humidity sounding channels (18 – 22, channel IDs 183±7 to 183±1). These channels are very similar to those affected in the current work, but still the most likely explanation for improved fit to ATMS is a real improvement in moisture and temperature forecasts. The improvements in humidity channels come not just at high latitudes but also in the tropics (not shown) and clearly both Stage 1 and Stage 2 contribute to the improvement. AIRS (panel b) measures hyperspectral infrared radiances, and it also shows significant improvements in the fit to humidity-sounding channels around 1300 – 1800 cm^{-1}, confirming the improvements in the

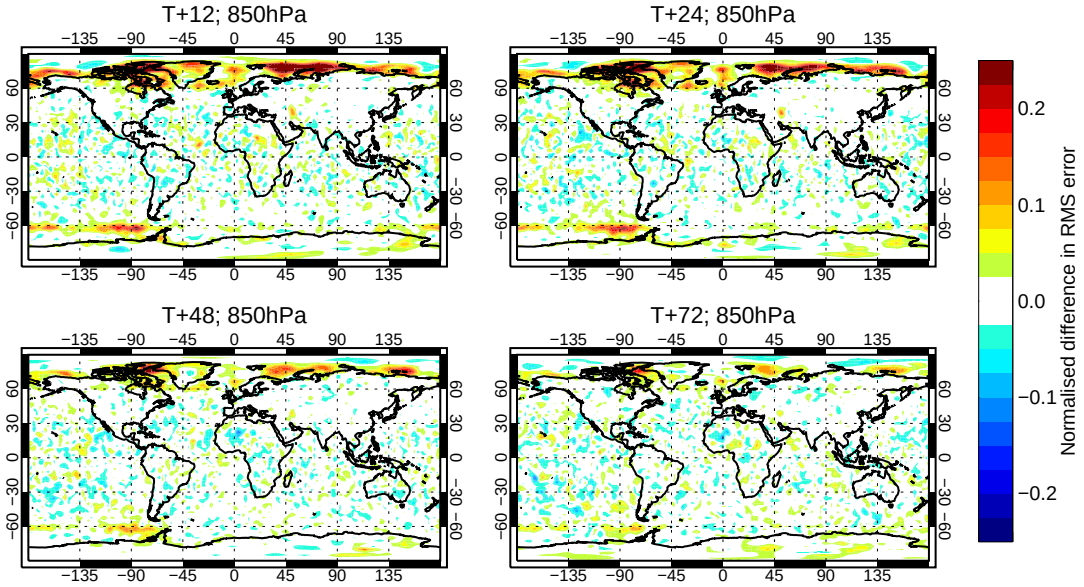

Figure 25: Normalised change in global RMS relative humidity error in Stage 2 forecasts compared to control at 850 hPa. No statistical significance testing is applied; other details as Fig. 23.

humidity forecast. Panel c shows the impact on fits to humidity radiosonde (TEMP-Q), with significant improvements in moisture at 700 hPa, that come mainly from the NH areas (not shown). Finally, panel d shows the impact on fits to various surface observations. Significant impacts are seen in ocean surface wind fields (moored buoy and scatterometer winds; MBUOY and ASCAT) which is consistent with the addition of window channels over the polar oceans, which have direct sensitivity to the surface windspeed through the ocean surface emissivity, as well as indirect sensitivity through water vapour, rain and cloud. The relative humidity surface station measurements (SYNOP RH2M) also improve; it is not clear whether over ocean or land or both. Fits to other observations do not show any significant changes on these global metrics (not shown).

Figure 27 shows the change in background fit to observations over the Arctic region; the selection is different from the previous figure because statistics for the polar regions are only automatically computed for these observation types. The fit to satellite-derived atmospheric motion vectors (AMVs, panel a) is not significantly changed, but the fit to in-situ wind measurements (panel b) is improved by around 0.5%. It is not clear where this difference comes from, but it could be associated with different geographical sampling of the sensors or the inherent difficulty of generating AMVs over polar regions. The impact on radiosonde humidity fits is mostly neutral except at 500 hPa (panel c) but there is a slight increase in low-level humidity bias (1000 - 850 hPa, not shown) consistent with a low-level drying seen in the analysis and short-range forecast around the Siberian sector of the Arctic ocean (not shown). This bias change contributes to the increased short-range RMS errors in relative humidity mentioned earlier, but this limitation is outweighed by beneficial impacts on short-range wind and temperature standard deviations in the Arctic.

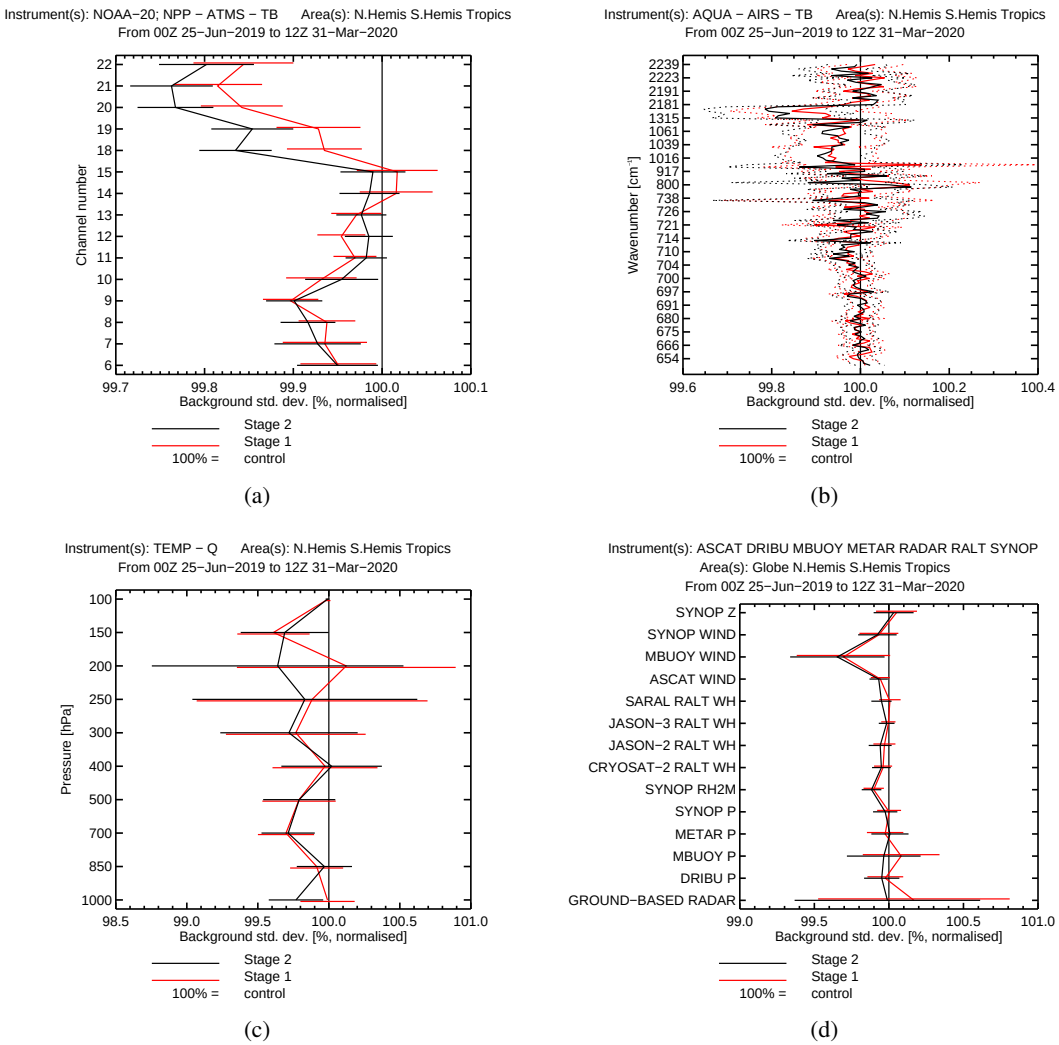

Figure 26: Relative change in global standard deviation of background departures at control (100%) and in Stage 1 (red) and Stage 2 (black).

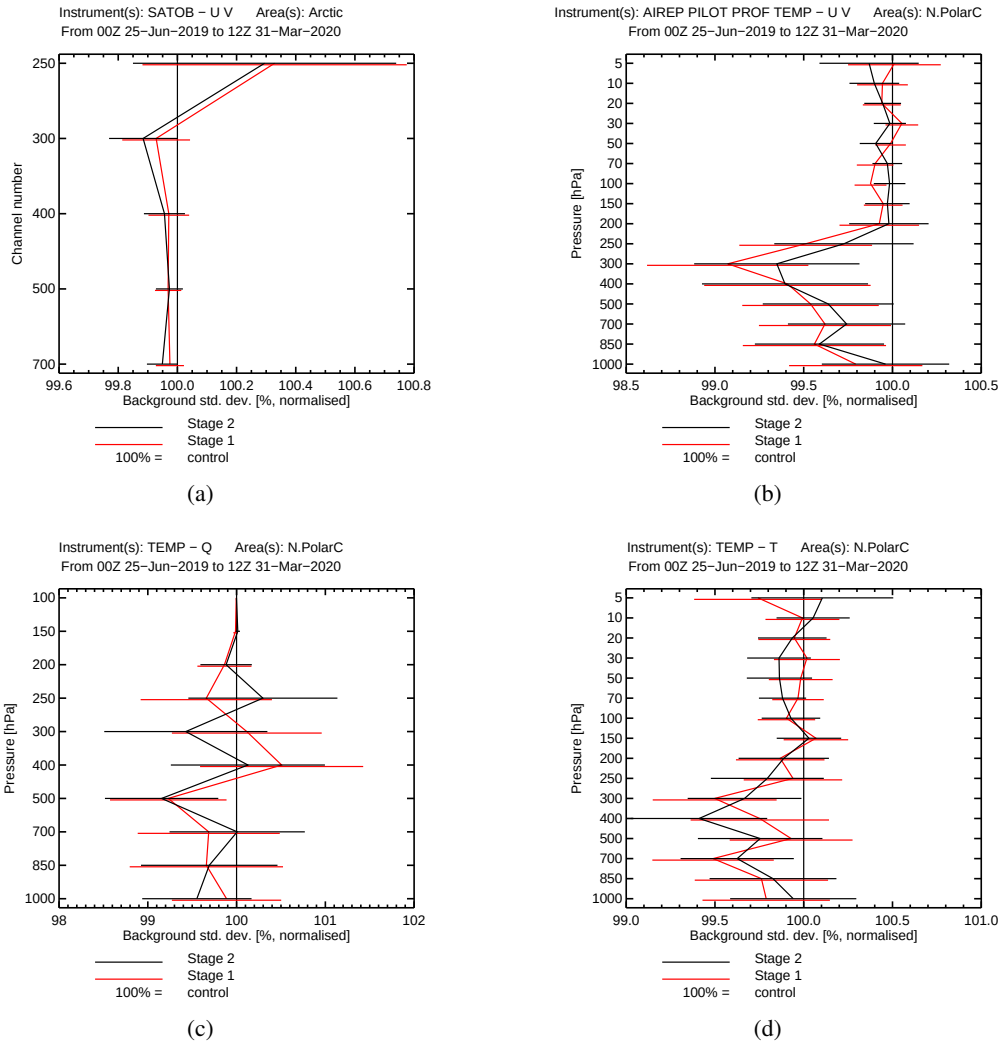

Figure 27: Relative change in standard deviation of background departures at control (100%) and in Stage 1 (red) and Stage 2 (black), for the Arctic area. This is defined as north of 60°N in the case of plots marked 'Arctic' and north of 66.5°N in case of 'NPolarC'. The observation types included are atmospheric motion vectors (referred to as SATOB on the figure) and aircraft, profilers, pilot balloons and radiosondes (AIREP, PILOT, PROF and TEMP on the figure).

7 Conclusion

The aim of this work has been to increase utilisation of all-sky microwave radiances over difficult surfaces such as land, snow, sea-ice, or mixtures of all of these and water. It has also been an aim to improve the processing as a baseline for future 'all-sky all-surface' developments. Although there is a relatively low radiometric contrast between land or sea-ice surfaces and the thermal emission from the atmosphere, there are clear scattering signals from rain, snow and graupel in window channels from 37 GHz to 166 GHz that can be assimilated in an all-sky framework. There are also thermal emission signals from rain and cloud water over cooler, high latitude surfaces, and water vapour information can also be exploited (Liu *et al.*, 2021, though the 24 GHz channels still await development). In addition, mixed surfaces are between 14 and 33% of all observations (Appendix A) and hence deserve better treatment: previously much of the data from mixed scenes was rejected. The following main developments helped improve utilisation over land, snow, sea-ice and mixed surfaces:

Mixed surfaces: A broad set of developments has tried to improve utilisation of mixed surfaces, standardising on a land fraction of 0.01 to switch between ocean and land processing, and stopping the rejection of data just because a surface is mixed. Now the QC depends on the quality of the dynamic emissivity retrieval and on its consistency with the land and sea-ice fractions. Throughout the code, the surface emissivity is now treated as being a fractional mix of ocean, sea-ice and land surface emissivities. The fractional ocean or water part of the scene is now always constrained by the relevant FASTEM ocean emissivity, which gives the much stricter QC of the dynamic emissivity retrieval that is now targeted to the land and/or sea-ice part of the scene. When extrapolating emissivity retrievals from one frequency to another, the correct modelling of the surface emissivity of the ocean part of the scene reduces errors in brightness temperature of up to 15 K for the most surface-sensitive channels. Large improvements in fit to observations are seen around coastlines, islands, lakes and mixed sea-ice areas, and many more observations can be assimilated as a result. The main drawback is that stricter QC reduces the amount of AMSU-A data available, particularly over Antarctic sea-ice regions, though only by a few percent.

Sea-ice detection: An obstacle to using window channels in the vicinity of sea-ice is the discrepancy between the IFS sea-ice fraction and the observed sea-ice fraction (which may come from deficiencies in the IFS sea-ice field, as well as from differences between the model grid and the satellite's field of view). A new observation-based sea-ice fraction retrieval has been introduced, based on the 10 GHz h-polarised channels available on GMI, AMSR2 and MWRI. This allows window channels (from 19 GHz to 183±7 GHz) to be assimilated over polar oceans for the first time, although sea-ice surfaces are still rejected. This also marks the first time the 10 GHz channels have been used in any capacity in the IFS. The sea-ice retrieval is optimised for quality control and for detection of the sea-ice edge, rather than for accurate sea-ice fractions (so it can retrieve up to 40% too high sea-ice fractions in some situations). Hence, further developments are needed before the new sea-ice information can be assimilated in the ocean component of the IFS.

Window channels over land surfaces: Using the new technique of atlas-based emissivity extrapolation (Lonitz *et al.*, 2022), selected window channels from 37 GHz to 166 GHz from GMI, SSMIS and AMSR2 are now assimilated over land surfaces. This includes an upgraded bias correction to address skin-temperature dependent biases over land, along with land-sea bias contrasts which are particularly large in AMSR2 (up to 4 K). A new set of quality control procedures is used to reject difficult surfaces such as dry desert soils and snow. Finally, observation error models have been tailored to the situation over land surfaces, where the surface emissivity is still subject to relatively large errors, and hence the 'clear-sky' observation error is now relatively large, at 9 – 20 K, depending on channel. This means that the assimilation over land primarily aims to benefit from larger cloud-affected signals. Future work will

seek to reduce the surface modelling errors and hence to extract smaller cloud and water vapour signals over land surfaces.

Other smaller changes in the current work have been the activation of GMI 183 GHz channels over land, snow and sea-ice, and changes relating to individual sensors such as AMSU-A. Table 3 and Sect. 6.2 further summarise the main changes and their impacts on the coverage of different all-sky sensors. For most channels and most sensors, coverage is increased by 4 – 33%, although AMSU-A loses up to 2.5% of scenes due to stricter surface emissivity QC, particularly over the Antarctic sea-ice. These results were obtained at lower model resolution than used in the operational system, but limited additional testing at high resolution suggests the results will be reasonably representative of the impact at high resolution. The impacts on forecasts are not large, but there is a small significant improvement to wind and geopotential height forecasts out to at least day 3 over northern hemisphere high latitudes, and there are improved background fits to independent observations such as AIRS, radiosonde humidity and surface-sensitive observations globally. And specifically in the Arctic, the background fits to in-situ wind and temperature observations are improved by around 0.5%.

Though the changes described in this work have been extensive and detailed, four main themes are worth highlighting again. First is the evolution towards treating the surface as a mixture of distinct surface types, each with a defined fraction (e.g. sea-ice fraction, land fraction) and each with potentially different radiative transfer properties. Given the often strong contrast in the radiative properties of different surface types, accurate knowledge of these fractions is a first step in being able to represent mixed scenes. Second, and as a consequence, is to start treating the satellite field of view more exactly, acknowledging that it will not exactly match up with the model grid, and in the microwave, likely it will be sensitive to many grid points at once. This is a major change in philosophy, since until now, passive sensors have been effectively treated as a point observation in global weather forecasting (though with consideration of along-path effects, e.g. Bormann, 2017; Shahabadi *et al.*, 2020). Third is that we are starting to exploit the surface information content of microwave imager observations, e.g. the use of a sea-ice retrieval for the first time. Fourth is the assimilation of microwave window channels over land surfaces which, when cycle 48r1 goes operational, will likely be a first in operational weather prediction.

The current work has highlighted some short-term issues that could be addressed while staying within the current assimilation approaches:

- The FOV-appropriate land fraction is currently only available for GMI, AMSR2 and SSMIS. It should be extended to all microwave sensors, to allow a better treatment of coastal scenes. This includes sounding instruments, and it is expected that it would particularly benefit AMSU-A, which has a large (48 km) field of view size and features several surface-sensitive channels.
- The first attempt at QC for window channels over land surfaces has been successful, but could be improved. Desert regions are clearly less problematic than snow, and a less extensive QC would be beneficial, for example to avoid rejecting convective observations over the Sahel.
- As window channels start to be more extensively used over land, snow and sea-ice surfaces, the bias correction model could be further improved, particularly by making the 10m windspeed predictor active over water surfaces only.
- The extension of window channels over land surfaces still needs to be activated for MWRI, following the example of GMI, AMSR2 and SSMIS (its omission was because it was not active in the testing done here, having only been activated mid-2020).

There are a number of larger developments that may improve the use of all-sky observations over difficult

surfaces, such as allowing skin temperature to be adjusted as a sink variable (e.g. Massart *et al.*, 2021) and a move from specular to Lambertian assumptions over snow surfaces (Bormann, 2021). However, particularly over frozen surfaces and dry soils, the penetration depth of microwave radiation is strongly variable with frequency, and volume scattering and reflection are strongly affected by the microstructure of these surface layers. Hence the idea of one 'skin' temperature or effective emitting temperature is unlikely to be helpful across the microwave and sub-mm. Similarly, the angular dependence of the surface interaction, currently represented in specular and Lambertian assumptions, is also strongly variable in space and in frequency. Hence greater progress may come from abandoning the dynamic emissivity approach, and moving to a more physical representation of radiative transfer in the surface layers, taking input from surface models of snow, ice and soil where available (e.g. Hirahara *et al.*, 2020). The main difficulty will be that current surface models do not represent all the microstructural details necessary to simulate the radiative transfer. However, this has parallels with the lack of microphysical information on clouds and precipitation, which did not stop successful all-sky assimilation. This has been dealt with using global parameter estimation (Geer *et al.*, 2014; Geer, 2021b). For all-surface assimilation, a more sophisticated and spatially- and temporally-dependent parameter estimation will probably be needed.

The benefit of the all-sky all-surface approach to atmospheric data assimilation is clear, bringing improved initial conditions particularly over land and high-latitude surfaces, where conventional observations can be sparse. However, one of the biggest goals is to infer surface information such as sea-ice fraction, soil moisture and snow cover that can be assimilated in an increasingly coupled assimilation system (de Rosnay *et al.*, 2022). It will be particularly beneficial to assimilate these quantities directly from satellite radiances, rather than from retrievals, which are a sub-optimal use of the data for many different reasons. Improved analyses of the surface should be beneficial not just to weather forecasting in the medium range but also to longer range forecasts. The developments made at cycle 48r1 are a step towards this.

Acknowledgements

During this work, Katrin Lonitz and David Duncan were funded through a EUMETSAT fellowship. Gabor Radnoti and Cristina Lupu are thanked for contributions to the development of the control experiment, and for help during this work. Stephen English is thanked for reviewing the manuscript.

References

Aires, F., Prigent, C., Bernardo, F., Jiménez, C., Saunders, R. and Brunel, P. (2011). A tool to estimate land-surface emissivities at microwave frequencies (TELSEM) for use in numerical weather prediction. *Quart. J. Roy. Meteorol. Soc.*, **137**(656), 690–699, URL https://doi.org/10.1002/qj.803.

Baordo, F. and Geer, A. (2015). Microwave surface emissivity over sea-ice. *EUMETSAT NWP-SAF visiting scientist report NWPSAF_EC_VS_026*, URL https://nwpsaf.eu/publications/vs_reports/nwpsaf-ec-vs-026.pdf.

Baordo, F. and Geer, A. J. (2016). Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval. *Quart. J. Roy. Meteorol. Soc.*, **142**, 2854–2866, URL https://doi.org/10.1002/qj.2873.

Barlakas, V., Geer, A. J. and Eriksson, P. (2021). Introducing hydrometeor orientation into all-sky millimeter/sub-millimeter assimilation. *Atmos. Meas. Tech.*, **14**, 3427—-3447, URL https://doi.org/10.5194/amt-14-3427-2021.

Bauer, P., Geer, A. J., Lopez, P. and Salmond, D. (2010). Direct 4D-Var assimilation of all-sky radiances: Part I. Implementation. *Quart. J. Roy. Meteorol. Soc.*, **136**, 1868–1885, URL https://doi.org/10.1002/qj.659.

Bauer, P., Moreau, E., Chevallier, F. and O'Keeffe, U. (2006). Multiple-scattering microwave radiative transfer for data assimilation applications. *Quart. J. Roy. Meteorol. Soc.*, **132**, 1259–1281, URL https://doi.org/10.1256/qj.05.153.

Bengtsson, L. and Hodges, K. (2005). On the impact of humidity observations in numerical weather prediction. *Tellus*, **57A**, 701–708, URL https://doi.org/10.3402/tellusa.v57i5.14734.

Berg, W., Bilanow, S., Chen, R., Datta, S., Draper, D., Ebrahimi, H., Farrar, S., Jones, W. L., Kroodsma, R., McKague, D., Payne, V., Wang, J., Wilhiet, T. and Yang, J. X. (2016). Intercalibration of the GPM microwave radiometer constellation. *Journal of Atmospheric and Oceanic Technology*, **33**(12), 2639–2654, URL https://doi.org/10.1175/JTECH-D-16-0100.1.

Bonavita, M., Geer, A. J. and Hamrud, M. (2020). All-sky microwave radiances assimilated with an ensemble Kalman filter. *Monthly Weather Review*, **148**(7), 2737–2760, URL https://doi.org/10.1175/MWR-D-19-0413.1.

Bormann, N. (2017). Slant path radiative transfer for the assimilation of sounder radiances. *Tellus A: Dynamic Meteorology and Oceanography*, **69**(1), 1272779, URL https://doi.org/10.1080/16000870.2016.1272779.

Bormann, N. (2021). Investigating the use of Lambertian reflection in the assimilation of surface-sensitive microwave sounding radiances over snow and sea-ice. *Tech. Memo. 886*, ECMWF, Reading, UK, URL https://doi.org/10.21957/hzcsrx8tt.

Bormann, N., Duncan, D., English, S., Healy, S., Lonitz, K., Chen, K., Lawrence, H. and Lu, Q. (2021). Growing operational use of FY-3 data in the ECMWF system. *Advances in Atmospheric Sciences*, pp. 1–14, URL https://doi.org/10.1007/s00376-020-0207-3.

Bormann, N., Lupu, C., Geer, A., Lawrence, H., Weston, P. and English, S. (2017). Assessment of the forecast impact of surface-sensitive microwave radiances over land and sea-ice. *Tech. Memo. 804*, ECMWF, Reading, UK, URL https://doi.org/10.21957/qyh34roht.

Browne, P. A., de Rosnay, P., Zuo, H., Bennett, A. and Dawson, A. (2019). Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system. *Remote Sensing*, **11**(3), 234, URL https://doi.org/10.3390/rs11030234.

de Rosnay, P., Browne, P., de Boisséson, E., Fairbairn, D., Hirahara, Y., Ochi, K., Schepers, D., Weston, P., Zuo, H., Alonso-Balmaseda, M., Balsamo, G., Bonavita, M., Bormann, N., Brown, A., Chrust, M., Dahoui, M., Chiara, G. D., English, S., Geer, A., Healy, S., Hersbach, H., Laloyaux, P., Magnusson, L., Massart, S., McNally, T., Pappenberger, F. and Rabier, F. (2022). Coupled assimilation at ECMWF: current status, challenges and future developments. *Quart. J. Roy. Meteorol. Soc.*, p. under review.

Dee, D. (2004). Variational bias correction of radiance data in the ECMWF system. In *ECMWF workshop proceedings: Assimilation of high spectral resolution sounders in NWP, 28 June – 1 July,*

2004, pp. 97–112, Eur. Cent. for Med. Range Weather Forecasts, Reading, UK, available from http://www.ecmwf.int.

Di Tomaso, E., Bormann, N. and English, S. (2013). Assimilation of ATOVS radiances at ECMWF: third year EUMETSAT fellowship report. *EUMETSAT/ECMWF Fellowship Programme Research Report No. 29, available from http://www.ecmwf.int*, URL https://www.ecmwf.int/node/9051.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, **116**, 140–158, URL https://doi.org/10.1016/j.rse.2010.10.017.

Duncan, D. I., Bormann, N., Geer, A. J. and Weston, P. (2022). Assimilation of AMSU-A in all-sky conditions. *Mon. Weath. Rev.*, p. in press, URL https://doi.org/10.1175/MWR-D-21-0273.1.

ECMWF (2020). IFS Documentation CY47R1. URL https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation.

English, S. J. (2008). The importance of accurate skin temperature in assimilating radiances from satellite sounding instruments. *Quart. J. Roy. Meteorol. Soc.*, **126**, 2911–2931, URL https://doi.org/10.1109/TGRS.2007.902413.

Errico, R. M., Bauer, P. and Mahfouf, J.-F. (2007). Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci.*, **64**, 3785 – 3798, URL https://doi.org/10.1175/2006JAS2044.1.

Eyre, J. R., English, S. J. and Forsythe, M. (2020). Assimilation of satellite data in numerical weather prediction. Part I: The early years. *Quart. J. Roy. Meteorol. Soc.*, **146**(726), 49–68, URL https://doi.org/10.1002/qj.3654.

Geer, A., Ahlgrimm, M., Bechtold, P., Bonavita, M., Bormann, N., English, S., Fielding, M., Forbes, R., Robin Hogan, E. H., Janisková, M., Lonitz, K., Lopez, P., Matricardi, M., Sandu, I. and Weston, P. (2017a). Assimilating observations sensitive to cloud and precipitation. *Tech. Memo. 815*, ECMWF, Reading, UK, URL https:/doi.org/10.21957/sz7cr1dym.

Geer, A. J. (2016). Significance of changes in forecast scores. *Tellus A*, **68**, 30229, URL https://doi.org/10.3402/tellusa.v68.30229.

Geer, A. J. (2021a). Learning earth system models from observations: Machine learning or data assimilation? *Phil. Trans. R. Soc. A*, **379**, 20200089, URL https://doi.org/10.1098/rsta.2020.0089.

Geer, A. J. (2021b). Physical characteristics of frozen hydrometeors inferred with parameter estimation. *Atmos. Meas. Tech.*, **14**, 5369–5395, URL https://doi.org/10.5194/amt-14-5369-2021.

Geer, A. J. and Baordo, F. (2014). Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech.*, **7**, 1839–1860, URL https:/doi.org/10.5194/amt-7-1839-2014.

Geer, A. J., Baordo, F., Bormann, N. and English, S. (2014). All-sky assimilation of microwave humidity sounders. *Tech. Memo. 741*, ECMWF, Reading, UK, URL https://doi.org/10.21957/obosmx154.

Geer, A. J., Baordo, F., Bormann, N., English, S., Kazumori, M., Lawrence, H., Lean, P., Lonitz, K. and Lupu, C. (2017b). The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quart. J. Roy. Meteorol. Soc.*, **143**, 3189–3206, URL https:/doi.org/10.1002/qj.3172.

Geer, A. J. and Bauer, P. (2010). Enhanced use of all-sky microwave observations sensitive to water vapour, cloud and precipitation. *Tech. Memo. 620*, ECMWF, Reading, UK, URL https:/doi.org/10.21957/mi79jebka.

Geer, A. J. and Bauer, P. (2011). Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteorol. Soc.*, **137**, 2024–2037, URL https://doi.org/10.1002/qj.830.

Geer, A. J., Bauer, P., Lonitz, K., Barlakas, V., Eriksson, P., Mendrok, J., Doherty, A., Hocking, J. and Chambon, P. (2021). Bulk hydrometeor optical properties for microwave and sub-millimetre radiative transfer in RTTOV-SCATT v13.0. *Geophys. Sci. Mod. Dev.*, **14**, 7497–7526, URL https://doi.org/10.5194/gmd-14-7497-2021.

Geer, A. J., Bauer, P. and Lopez, P. (2010). Direct 4D-Var assimilation of all-sky radiances: Part II. Assessment. *Quart. J. Roy. Meteorol. Soc.*, **136**, 1886–1905, URL https://doi.org/10.1002/qj.681.

Geer, A. J., Migliorini, S. and Matricardi, M. (2019). All-sky assimilation of infrared radiances sensitive to mid- and upper-tropospheric moisture and cloud. *Atmos. Meas. Tech.*, pp. 4903—4929, URL https://doi.org/10.5194/amt-12-4903-2019.

Gentemann, C. L. and Hilburn, K. A. (2015). In situ validation of sea surface temperatures from the GCOM-W1 AMSR2 RSS calibrated brightness temperatures. *Journal of Geophysical Research: Oceans*, **120**(5), 3567–3585, URL https://doi.org/10.1002/2014JC010574.

Gérard, E. and Saunders, R. (1999). Four-dimensional variational assimilation of Special Sensor Microwave / Imager total column water vapour in the ECMWF model. *Quart. J. Roy. Meteorol. Soc.*, **125**, 3077–3102, URL https://doi.org/10.1002/qj.49712556014.

Gong, J. and Wu, D. L. (2017). Microphysical properties of frozen particles inferred from global precipitation measurement (GPM) microwave imager (GMI) polarimetric measurements. *Atmos. Chem. Phys.*, **17**(4), 2741–2757, URL https://doi.org/10.5194/acp-17-2741-2017.

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J. *et al.* (2020). The current configuration of the ostia system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sensing*, **12**(4), 720, URL https://doi.org/10.3390/rs12040720.

Hirahara, Y., Rosnay, P. d. and Arduini, G. (2020). Evaluation of a microwave emissivity module for snow covered area with CMEM in the ECMWF integrated forecasting system. *Remote Sensing*, **12**(18), 2946, URL https://doi.org/10.3390/rs12182946.

Hong, G., Heygster, G., Miao, J. and Kunzi, K. (2005). Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements. *J. Geophys. Res.*, **110**, D05205, URL https://doi.org/10.1029/2004JD004949.

Karbou, F., Bauer, P., Geer, A. and Bell, W. (2008). Exploitation of microwave sounder/imager data over land surfaces in the presence of clouds and precipitation. *EUMETSAT Hydrology SAF visiting scientist report*.

Karbou, F., Gérard, É. and Rabier, F. (2006). Microwave land emissivity and skin temperature for AMSU-A and -B assimilation over land. *Quarterly Journal of the Royal Meteorological Society*, **132**(620), 2333–2355, URL https://doi.org/10.1256/qj.05.216.

Kazumori, M. and English, S. J. (2015). Use of the ocean surface wind direction signal in microwave radiance assimilation. *Quarterly Journal of the Royal Meteorological Society*, **141**(689), 1354–1375, URL https://doi.org/10.1002/qj.2445.

Kazumori, M., Geer, A. J. and English, S. J. (2016). Effects of all-sky assimilation of GCOM-W/AMSR2 radiances in the ECMWF numerical weather prediction system. *Quart. J. Roy. Meteorol. Soc.*, **142**, 721–737, URL https://doi.org/10.1002/qj.2669.

Krzeminski, B., Bormann, N., Karbou, F. and Bauer, P. (2009). Improved use of surface-sensitive microwave radiances at ECMWF. In *Proc. EUMETSAT Meteorol. Satell. Conf*, pp. 21–25.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C. *et al.* (2019). Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, **13**(1), 49–78, URL https://doi.org/10.5194/tc-13-49-2019.

Lawrence, H., Bormann, N., Geer, A., Lu, Q. and English, S. (2018). Evaluation and assimilation of the microwave sounder MWHS-2 onboard FY-3C in the ECMWF numerical weather prediction system. *IEEE Trans. Geosci. Rem. Sens.*, **56**, 3333 – 3349, URL https://doi.org/10.1109/TGRS.2018.2798292.

Lawrence, H., Bormann, N., Sandu, I., Day, J., Farnan, J. and Bauer, P. (2019). Use and impact of Arctic observations in the ECMWF numerical weather prediction system. *Quart. J. Roy. Meteorol. Soc.*, **145**(725), 3432–3454, URL https://doi.org/10.1002/qj.3628.

Lean, P., Geer, A. and Lonitz, K. (2017). Assimilation of Global Precipitation Mission (GPM) Microwave Imager (GMI) in all-sky conditions. *Tech. Memo. 799*, ECMWF, Reading, UK, URL https://doi.org/10.21957/8orc7sn33.

Liu, Q. M., Cao, C., Grassotti, C. and Lee, Y.-K. (2021). How can microwave observations at 23.8 GHz help in acquiring water vapor in the atmosphere over land? *Remote Sensing*, **13**(3), 489, URL https://doi.org/10.3390/rs13030489.

Lonitz, K., Geer, A. J. and Bormann, N. (2022). Towards assimilating microwave imager channels over land. *EUMETSAT/ECMWF Fellowship Programme Research Report 58*, ECMWF, Reading, UK, URL https://doi.org/10.21957/gdwqzfns.

Lopez, P. (2013). Experimental 4D-Var assimilation of SYNOP rain gauge data at ECMWF. *Mon. Wea. Rev.*, **141**(5), 1527–1544, URL https://doi.org/10.1175/MWR-D-12-00024.1.

Massart, S., Bormann, N., Bonavita, M. and Lupu, C. (2021). 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). *Geoscientific Model Development*, **14**(9), 5467–5485, URL https://doi.org/10.5194/gmd-14-5467-2021.

McNally, A. (2002). A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quart. J. Roy. Meteorol. Soc.*, **128**(585), 2551–2556, URL https://doi.org/10.1256/qj.01.206.

Saunders, R., Hocking, J., Turner, E., Havemann, S., Geer, A., Lupu, C., Vidot, J., Chambon, P., Köpken-Watts, C., Scheck, L., Stiller, O., Stumpf, C., Borbas, E. and Brunel, P. (2020). RTTOV-13 science and validation report. *NWP-SAF report NWPSAF-MO-TV-046*, EUMETSAT NWP-SAF, Met Office, UK, URL https://nwp-saf.eumetsat.int/site/download/documentation/rtm/docs_rttov13/rttov13_svr.pdf.

Shahabadi, M. B., Buehner, M., Aparicio, J. and Garand, L. (2020). Implementation of slant-path radiative transfer in environment canada's global deterministic weather prediction system. *Monthly Weather Review*, **148**(10), 4231–4245, URL https://doi.org/10.1175/MWR-D-20-0060.1.

Weston, P., Bormann, N., Geer, A. and Lawrence, H. (2017). Harmonisation of the usage of microwave sounder data over land, coasts, sea ice and snow: First year report. *EUMETSAT/ECMWF Fellowship Programme Research Reports 45*, ECMWF, Reading, UK, URL https://www.ecmwf.int/node/17766.

Weston, P., Geer, A. and Bormann, N. (2019). Investigations into the assimilation of AMSU-A in the presence of cloud and precipitation. *EUMETSAT/ECMWF Fellowship Programme Research Reports 50*, ECMWF, Reading, UK, URL https://doi.org/10.21957/ewahn9ce.

A Proportion of mixed surfaces

The proportion of satellite observations seeing mixed surfaces is surprisingly large, but also strongly dependent on the definitions used. Table 6 shows a range of estimates, based on the 00 UTC 1st January 2021 cycle for AMSR2 observations, and using an additional experiment based on the full operational model resolution (T1279co, which corresponds to 8 - 9 km horizontal grid box size). There are at least three factors to consider. The first is the field of view size, which for the microwave imagers is typically an 80 km by 80 km superob. The FOV-appropriate land fraction is calculated on this basis. Alternatively, the land fraction on the model grid resolution at T1279co roughly represents the highest possible resolution on current microwave instruments (which applies only to the non-superobbed (raw) data in the higher frequency channels). The second is the threshold above which a scene is considered unmixed. Here, we consider either 1.0 (which includes mixtures that would not in practice affect the radiative transfer) or 0.99 (which ignores situations where up to 1% of a scene contains some alternative surface type). The third is to define exactly which surfaces should be considered homogeneous. In the table, land is considered homogeneous, but of course it is usually a mixture of different surface types. The figures are given both ignoring sea-ice, and including it. This is because it is rare for the model's sea-ice fraction to be 1.00, so in practice much of the sea-ice area is treated as mixed; further, the sea-ice extent is variable through the year. Ignoring sea-ice, using a relaxed threshold, and for the highest FOV resolution, just 5.5% of scenes are mixed. Including sea-ice as of 1st January 2021, using a strict threshold for mixed scenes, and using the 80 km superob, 33.2% of scenes are mixed. As a headline figure in this work, 22.3% is used, representing the inclusion of sea-ice, the current relaxed ocean-land threshold

Table 6: Percentage of mixed scenes for AMSR2, using various criteria

Effective FOV size	Using 80 km superob		Using 8-9 km T1279co grid	
Thresholds	1.00	0.99	1.00	0.99
Mixed land and ocean	24.8	15.5	10.4	5.5
Mixtures of land, ocean, and/or sea-ice	33.2	22.3	21.1	13.9

of 0.01 and the 80 km superob (not so different from the AMSU-A raw 50 km resolution, so this figure is fairly representative for sounders too). Figure 6 is also worth a look: it comes to slightly different numbers using slightly different criteria.