EUMETSAT ECMWF

EUMETSAT Contract Report

EUM/CO/22/4600002673/SDM

Contract Report to EUMETSAT

Vicarious calibration monitoring for MWI and ICI using NWP fields

Final Report for EUMETSAT contract no. EUM/CO/22/4600002673/SDM

David I. Duncan, Alan J. Geer, Niels Bormann, Mohamed Dahoui

May 2024

Series: EUMETSAT/ECMWF Contract Report Series

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

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, UK

Contents

Executive summary

The MicroWave Imager (MWI) and Ice Cloud Imager (ICI) on board Metop-SG-B will provide a wealth of information to constrain our knowledge of the global hydrological cycle, from ice particles in the upper troposphere to the heaviest tropical precipitation. These instruments will represent the first microwave imagers launched by Europe, building upon designs that have underpinned global estimates of precipitation since the 1980s, but crucially adding coverage of the currently unexploited "sub-mm" part of the spectrum by observing above 300 GHz. These new frequency bands will provide much greater sensitivity to atmospheric ice than ever before. Beyond observing hydrometeors in all forms, the combined capabilities of MWI and ICI span sensitivities to surface properties, water vapour, atmospheric temperature, and even ozone. Because of this highly complementary information across the spectrum, the two sensors can be combined into a single super-sensor designated "MWIICI" and this reflects how the data will be used at ECMWF, both scientifically and technically. The large range of physical sensitivities and cross-spectral sensitivity is the great strength of MWIICI, but it represents a significant challenge to traditional calibration and validation activities ("cal/val") for new instruments.

In this study, a method is developed for analysing the in-orbit performance of MWIICI in terms of biases. It is based on comparing the observations against modelled radiances from a state-of-the-art numerical weather prediction (NWP) model. By examining departures (observations minus model background, or O-B) against the IFS as a reference, we can analyse sensor performance everywhere on the globe—millions of observations each day—and readily compare to similar sensors, as well as potentially other observations. The monitoring is underpinned by the "all-sky" approach for assimilating microwave radiances that has been pioneered at ECMWF and uses the RTTOV-SCATT radiative transfer model to include the radiative effects of clouds and precipitation. The latest version of RTTOV-SCATT is used, capitalising on numerous radiative transfer developments in recent years such as the SURFEM-Ocean emissivity model, updated gas spectroscopy for the sub-mm, and improved scattering properties of ice particles. By including hydrometeors in the forward model and considering each channel's surface and cloud sensitivity for every scene, this method retains observations that might be thrown away by a traditional "clear-sky" sampling approach whilst yielding a more balanced data sample for cal/val. The all-sky sampling approach seeks to maximise the data available for cal/val whilst removing most scenes with known model biases, namely thicker clouds and surfaces such as sea-ice and most land.

The sampling method relies on a symmetric approach for screening out radiatively significant cloud signals, checking for the presence of cloud in both the model and the observation. This is accomplished using the cloud impact (CI) parameter, defined by the difference in brightness temperature (TB) between a clear scene and the TBs that are observed and modelled. Additionally, a channel-specific surface-to-space transmittance is checked to determine if an observation has too much surface sensitivity to be included in the cal/val sample. Because CI and surface-to-space transmittance are location- and channel-specific, the main cal/val sample ("stringent" data selection) is maximised based on each channel's specific sensitivities to the surface and hydrometeors; a "dynamic" sample adds data over land for surface-sensitive channels to better capture the dynamic range of MWIICI. In addition to these channel-based samples for cal/val, a "unified" sample is defined using CI at key wavelengths that considers the sensitivities of all channels together to define a common sample. Thus three different data samples are defined for cal/val, applied equally to MWIICI and microwave imagers currently assimilated at ECMWF like GMI, SSMIS, and AMSR2.

Results are split into two sections: application of the method to two test orbits of MWIICI and application to current microwave imagers. The former clarifies the expected sensitivities of MWIICI and shows the data flow through the IFS. By applying the method to current sensors, we can assess its ability to detect

known sensor biases. For example, the analysis clearly shows orbital biases for F17 SSMIS, positive overall biases for AMSR2 channels, and generally small biases for GMI, each of which is well known in the cal/val literature. In addition, the method points to some bias structures that were not previously clear such as apparent scene-dependent biases for several channels, most notably 150 GHz on SSMIS. These results and the consistency of geographic patterns between similar channels on different sensors give confidence in the method, in that the cal/val sample exhibits similar patterns of atmospheric variability despite distinct orbits and differing central frequencies of matched channels. Between the three data samples defined here, between 20 and 70% of the total data are used for cal/val analysis, depending on the channel and the sample chosen. For the channel-specific samples, more data are retained for purely sounding frequencies such as the 50 and 118 GHz sounding complexes, with less data retained for frequencies like 89 GHz that have strong surface and cloud sensitivity.

Performance of the MWI and ICI instruments will be monitored in near-real-time after launch, leveraging the statistical and graphical tools developed at ECMWF for monitoring the global observing system. This will be the first time that such NWP-based monitoring is tailor-made to comprehensively evaluate instrument specifications, included as an integral part of the cal/val activities for MWI and ICI. Thus cal/val requirements such as inter-channel and intra-scan bias characteristics can be assessed almost immediately via an ECMWF-hosted website. Of particular interest is the direct comparison with equivalent channels on the reference-quality instrument GMI, permitting a type of “double difference” analysis of channel biases using the ECMWF model as a transfer standard. This facility will provide EUMETSAT and the global meteorological community with valuable information for assessing radiances from the MWI and ICI instruments. The technical developments of this project lay the groundwork for assimilation of MWI and ICI, aiding early operational exploitation of these missions.

1 Introduction

The EUMETSAT Polar System (EPS) Second Generation programme consists of six Metop Second Generation (Metop-SG) satellites. Three of these satellites, Metop-SG-B1, -2, and -3, will each hold two passive microwave radiometers for observing the Earth's water cycle in an unprecedented way. The MicroWave Imager (MWI) and Ice Cloud Imager (ICI) together will cover the microwave and sub-mm spectrum from 18 to 664 GHz. The "sub-mm" is the part of the spectrum above 300 GHz where wavelengths are shorter than 1 mm, and ICI will be one of the first spaceborne sensors observing this spectral region.

Microwave imagers like MWI and ICI are powerful because of their large range of sensitivities: to surface properties, profiles of moisture and temperature, cloud presence and type, and even properties of precipitation. This range of sensitivities is their strength, but it makes these sensors challenging from a cal/val perspective. To ignore any of these sensitivities can lead to misinterpretation of the other signals. That is why this study is predicated upon "all-sky" simulations of modelled scenes with observed radiances, simulating all constituent radiative effects rather than just thermal emission from the surface and atmospheric gases.

At ECMWF, the all-sky approach to assimilation of microwave radiances has proven that treating the complex radiative signatures of emission and scattering from hydrometeors is effective at producing a more balanced distribution when comparing observations with the model. This is in contrast to a "clear-sky" approach in which observed scenes with cloud contamination are discarded and only the clear-air components of the atmosphere are modelled, namely temperature and water vapour. This key difference is a crucial consideration for the cal/val strategy of instruments with sensitivity to hydrometeors, particularly for ICI as it has greater sensitivity to thin clouds than any previous microwave imager.

NWP systems have been used to evaluate current and future satellite sensors in various contexts regarding calibration (Newman *et al.*, 2020). For example, in the microwave part of the spectrum, NWP fields have been used to assess and develop corrections for sounders (Lu *et al.*, 2015), imagers (Geer *et al.*, 2010), and combined sounder/imager sensors (Bell *et al.*, 2008). Due to the combined heft of millions of data points analysed using high-quality model fields with an accurate forward model, this type of analysis can identify calibration aspects down to the hundredths of a degree for relative biases of a pure sounding channel (Bormann *et al.*, 2023); and while analysis of surface- and cloud-sensitive channels such as those on an imager is not that precise, the wealth of data still permits otherwise hidden signals to be identified (Scanlon *et al.*, 2023). For example, all-sky NWP-based analysis of sounders and imagers has been crucial for understanding the data quality from the FY-3 satellites (Lawrence *et al.*, 2017; Lawrence *et al.*, 2018; Steele *et al.*, 2023).

This study builds upon previous work that has shown the power of using NWP fields for cal/val, and leverages the state of the art in numerical modelling, radiative transfer, and assimilation methodology to prepare for the assessment of Europe's first operational microwave imager and sub-millimetre sounder. There is a particular emphasis here on the data selection decisions required to provide a large but unbiased cal/val data sample. Tools to assess specific instrument requirements are developed and analysed, both for the imagers on Metop-SG-B and currently orbiting microwave imagers. In particular, vicarious calibration against NWP-simulated observations is a formal part of the EUMETSAT cal/val plans for MWI and ICI (see Appendix A) and in this project, ECMWF has developed the statistical monitoring framework to support this.

Table 1: MWI channel frequencies and polarisations from MWI-L1B-PGS. NEDT and bias are requirements from the EURD (Table 18) that will be evaluated using the tools derived in this project. The NEDT specification is for footprint rather than sample NEDT (see text). Bandwidths are maximum half-power bandwidths, except for channels 1-6 which are -20dB power bandwidths.

Ch. No.	*Frequency [GHz]*	*Bandwidth [MHz]*	*Pol.*	*NEDT*	*Bias*	*Dynamic range [K]*		*FOV*
						Low	High	
1,2	18.7	200	V, H	0.8 K	1.0 K	80	335	50 km
3,4	23.8	400	V, H	0.7 K	1.0 K	80	335	50 km
5,6	31.4	200	V, H	0.9 K	1.0 K	80	335	30 km
7,8	50.3	180	V, H	1.1 K	1.0 K	100	320	30 km
9,10	52.70	180	V, H	1.1 K	1.0 K	100	320	30 km
11,12	53.24	400	V, H	1.1 K	1.0 K	100	320	30 km
13,14	53.75	400	V, H	1.1 K	1.0 K	100	320	30 km
15,16	89.0	4000	V, H	1.1 K	1.0 K	80	335	10 km
17	118.7503 ± 3.2	2x500	V	1.3 K	1.0 K	80	320	10 km
18	118.7503 ± 2.1	2x400	V	1.3 K	1.0 K	80	320	10 km
19	118.7503 ± 1.4	2x400	V	1.3 K	1.0 K	80	320	10 km
20	118.7503 ± 1.2	2x400	V	1.3 K	1.0 K	80	320	10 km
21	165.5 ± 0.725	2x1350	V	1.2 K	1.0 K	80	335	10 km
22	183.31 ± 7.0	2x2000	V	1.3 K	1.0 K	80	320	10 km
23	183.31 ± 6.1	2x1500	V	1.2 K	1.0 K	80	320	10 km
24	183.31 ± 4.9	2x1500	V	1.2 K	1.0 K	80	320	10 km
25	183.31 ± 3.4	2x1500	V	1.2 K	1.0 K	80	320	10 km
26	183.31 ± 2.0	2x1500	V	1.3 K	1.0 K	80	320	10 km

2 Sensor characteristics and requirements

MWI and ICI hold a combined 39 channels with 29 distinct frequency bands. All frequencies are measured at vertical polarisation (V), and some frequencies are measured at both vertical and horizontal polarisations (V and H). Channel numbers are given in Table 1 for MWI and Table 2 for ICI. Also given are specifications for the Noise Equivalent Differential Temperature (NEDT), a measure of the random noise when viewing a known target, and the Field of View (FOV), defined as the half-power beam width of the antenna pattern on the Earth's surface. A peculiarity of the MWI and ICI mission requirements is that both the NEDT and FOV requirements refer to a notional "footprint" that is based on the combination multiple raw measurement "samples" as further explained next. The footprint (i.e. FOV) requirements for both MWI and ICI are defined as arithmetic average of the footprint's two axes at 3dB, as the FOV is elliptical with a ratio of roughly 3:5 (user requirements MWI-05200 and ICI-06200, Sec. 2.1 and 2.2).

To be clear, throughout this report the following convention is used, in line with the EURD documentation:

- **Sample** is the raw measurement, of which there are 1394 per scan on MWI and 782 on ICI in the EUMETSAT test data
- **Footprint** is made up of several raw measurements due to over-sampling; used interchangeably in the text with FOV

The difference between sample and footprint NEDT is crucial to understand for MWI and ICI due to their

Table 2: ICI channel frequencies and polarisations from tables 19 and 20 of the EURD. The NEDT and FOV size specifications are for footprint rather than sample (see text).

Ch. No.	*Frequency [GHz]*	*Bandwidth [GHz]*	*Pol.*	*NEDT*	*Bias*	*FOV*
1	183.31 ± 7.0	2x2.0	V	0.8 K	1.0 K	16 km
2	183.31 ± 3.4	2x1.5	V	0.8 K	1.0 K	16 km
3	183.31 ± 2.0	2x1.5	V	0.8 K	1.0 K	16 km
4,5	243.2 ± 2.5	2x3.0	V, H	0.7 K	1.5 K	16 km
6	325.15 ± 9.5	2x3.0	V	1.2 K	1.5 K	16 km
7	325.15 ± 3.5	2x2.4	V	1.3 K	1.5 K	16 km
8	325.15 ± 1.5	2x1.6	V	1.5 K	1.5 K	16 km
9	448 ± 7.2	2x3.0	V	1.4 K	1.5 K	16 km
10	448 ± 3.0	2x2.0	V	1.6 K	1.5 K	16 km
11	448 ± 1.4	2x1.2	V	2.0 K	1.5 K	16 km
12,13	664 ± 4.2	2x5.0	V, H	1.6 K	1.5 K	16 km

spatial over-sampling. In Equation 1 the sample and footprint NEDT values are related by a function of the integration time of the sensor (0.394 ms for MWI, 0.663[1] ms for ICI) and the integration time to cover the 3dB footprints as defined in the specifications. Integration times for the 3dB footprints are given for MWI and ICI channels in Table 3. The specification NEDT values are for the footprint, as are NEDT values given in Level 1 data files. Thus to convert from footprint values in L1B back to those of the individual samples (i.e. original observation NEDT), one needs to scale $NEDT_{footprint}$ by one over the square root of the ratio of these integration times. For example, if MWI channel 10 shows NEDT of 0.80 K in the L1B BUFR file, this 3dB footprint NEDT converts to a sample NEDT of $0.8/\sqrt{0.394/4.253} = 2.63$ K. Or for ICI channel 3, a 3dB footprint NEDT of 0.71 K converts to $0.71/\sqrt{0.663/2.627} = 1.41$ K. The ratio between 3dB NEDT and sample NEDT is given in Table 3 for every MWI and ICI channel, with values between about 4.65 and 1.68. In the equation below, subscript i is for each channel and *int* stands for integration.

$$NEDT_{footprint}(i) = \sqrt{T_{int}/T_{int3dB}(i)}NEDT_{sample}(i) \tag{1}$$

Or rewritten to give the ratio between footprint and sample NEDT as given in the following table:

$$Ratio(i) = 1/\sqrt{T_{int}/T_{int3dB}(i)} = NEDT_{sample}(i)/NEDT_{footprint}(i) \tag{2}$$

For some perspective on the MWI and ICI specifications, we can compare the sensor characteristics briefly to other sensors. The noise characteristics of MWI and ICI are expected to be comparable to similar-calibre sensors such as GMI, AMSR2, and MHS. The FOV sizes are also comparable, though the smallest FOV sizes for MWI (10 km) are larger than GMI and AMSR2 due to a smaller reflector and also a higher orbital altitude—GMI on GPM flies at 407[2] km and AMSR2 on GCOM-W at 700 km, whereas Metop-SG is expected to orbit at 835 km. FOVs for 183 GHz channels on MWI are smaller than those of MHS, which is 16 km at nadir and larger off-nadir.

Both MWI and ICI are conically-scanning radiometers and thus have a fixed view angle across the scan,

[1] 0.663 ms in the test data; 0.640 ms for the PFM.

[2] The GPM orbit was boosted during 7-8 November 2023 from roughly 405 to 435 km altitude to extend mission life. Values in this study reflect the spacecraft's original altitude.

Table 3: Integration times given in milliseconds (ms) for 3dB footprint for each MWI and ICI channel (T_{int3dB}), measured from the first Proto Flight Model (PFM) of EPS-SG, from satellite B1. Values are given for MWI first (1-26) and then ICI. The ratios describe the sample to 3dB NEDT values. In other words, to convert from the specified 3dB footprint NEDT to sample NEDT, multiply by these values (as defined in Eq. 2).

MWI	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
T_{int3dB}	8.525	8.475	8.170	8.407	5.209	5.212	4.323	4.328	4.187	4.253	4.173	4.239	4.132
Ratio	4.65	4.64	4.55	4.62	3.64	3.64	3.31	3.31	3.26	3.29	3.24	3.27	2.14
MWI	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
T_{int3dB}	4.217	1.805	1.786	1.688	1.670	1.671	1.671	1.339	1.122	1.124	1.132	1.134	1.127
Ratio	3.27	2.14	2.13	2.07	2.06	2.06	2.06	1.84	1.69	1.69	1.70	1.70	1.69
ICI	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
T_{int3dB}	2.632	2.637	2.627	2.579	2.563	2.080	2.087	2.082	1.872	1.873	1.874	2.776	2.596
Ratio	1.99	1.99	1.99	1.97	1.97	1.77	1.77	1.77	1.68	1.68	1.68	2.05	1.98

Table 4: Scan characteristics for MWI, ICI, and three current microwave imagers.

	MWI	**ICI**	**SSMIS**	**GMI**	**AMSR2**
Scan width (km)	1700	1700	1700	850	1450
Samples per scan	1394	782	60	221	243
Distance between scans (km)	9	9	12.5	13.4	10

along with a fixed field of view size. This is an advantage for window channels as they can maximise the difference of vertically- vs. horizontally-polarised (V vs. H) radiation by viewing near the Brewster angle across the entire scan, with the emissivity over sea at H typically about half that of V. These fixed observation parameters also facilitate superobbing, which is described later. This scanning strategy follows a long lineage of microwave imagers going back several decades.

One scan is approximately 1700 km wide for both sensors, with 1394 samples per scan on MWI and 782 on ICI (Table 4)[3]. Each scan line is spaced about 9 km apart, consistent with a scanning rate of 45 scans per minute (one rotation every 1.333 s) and 7 km/s platform velocity. With FOV sizes of 10 km and larger, all MWI and ICI channels have overlapped FOVs across-track and most channels (MWI 1-16 and ICI 1-13) have significant overlap along-track as well. This heavy over-sampling permits convolution techniques to remap the FOVs to target sizes for purposes of noise reduction or greater viewing consistency for retrievals (Rydberg and Eriksson, 2019).

Early microwave radiometers (e.g. SSMI) were designed to point at the same geolocation point with all channels (Hollinger *et al.*, 1990; Berg *et al.*, 2013), so that all channels would share a latitude and longitude for the FOV's centre, despite FOV sizes that depend on frequency. Sensors with a large spectral range such as SSMIS have separate feedhorns for separate channel groups (Kunkee *et al.*, 2008) and GMI has two sets of feedhorns for the low- and high-frequency channel sets, with significantly different viewing angles (Draper *et al.*, 2015). In contrast, the AMSR2 radiometer has six feedhorns for its 7 different frequency bands, but these are arrayed in such a way along the direction of rotation that almost all FOVs are aligned within about 5km on the ground and share a common incidence angle (Maeda *et al.*,

[3]Slightly larger numbers of samples per scan are expected for the real data; the numbers 1394 and 782 are used in this project reflecting the specifications in place when the EUMETSAT test data was generated. But for example the actual ICI PFM instrument intended for the first launch has 810 samples per scan.

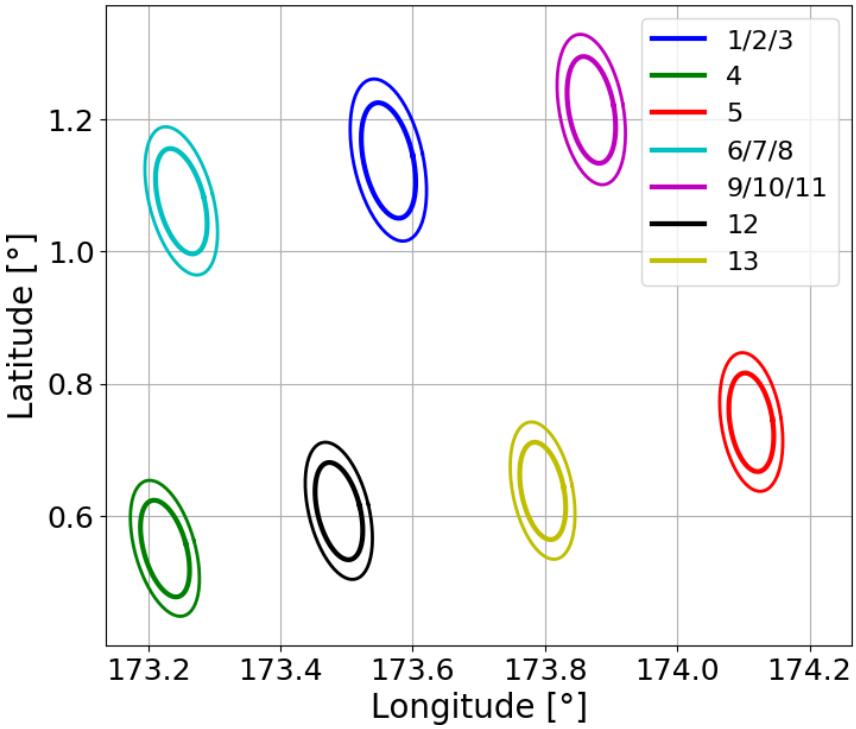


Figure 1: Taken from Eriksson et al. (2020) (their Fig. 2), showing geolocations and FOV sizes for the 7 ICI feedhorns at a given observation vector: "Instantaneous ICI footprints. The inner and outer contours represent the -3 and -6 dB level of normalised antenna patterns. The assumed sensor position is 6.9 S, 175.3 E at an altitude of 824.5 km." Reproduced under the Creative Commons Attribution 4.0 License from https://doi.org/10.5194/amt-13-53-2020.

2016, their Fig. 2). MWI and ICI were designed with separate feedhorns that point at slightly different angles from the spacecraft and thus the FOVs are not co-located and the viewing geometry is different for each horn at a given sample and scan number. This results in observations that are staggered on the ground as shown in Fig. 1 for ICI and Fig. 2 for MWI, requiring that several geolocations and angles are provided in L1 files—one for each feedhorn. The footprints are thus separated by as much as roughly 100 km on the ground for ICI, with roughly 200 km between the first and fourth MWI horns' footprints; these separations are along-scan (or across-track considering the spacecraft's direction of travel), leading to edges of the scan with areas of no overlap between channels on the same instrument.

2.1 EUMETSAT user requirements to be evaluated using vicarious cal/val: MWI

A main purpose of the vicarious cal/val system is to help validate the EUMETSAT user requirements listed in Table 1, along with more general monitoring of the quality and long-term stability of the instruments. The following are relevant requirements extracted from the EPS-SG End User Requirements Document (EURD, appendix A) for MWI:

- **MWI-05060** The dynamic range of the MWI shall cover the spectral radiances in terms of brightness temperatures as given in Table 1.
- **MWI-05065** All MWI requirements shall be met over the dynamic ranges given in Table 1

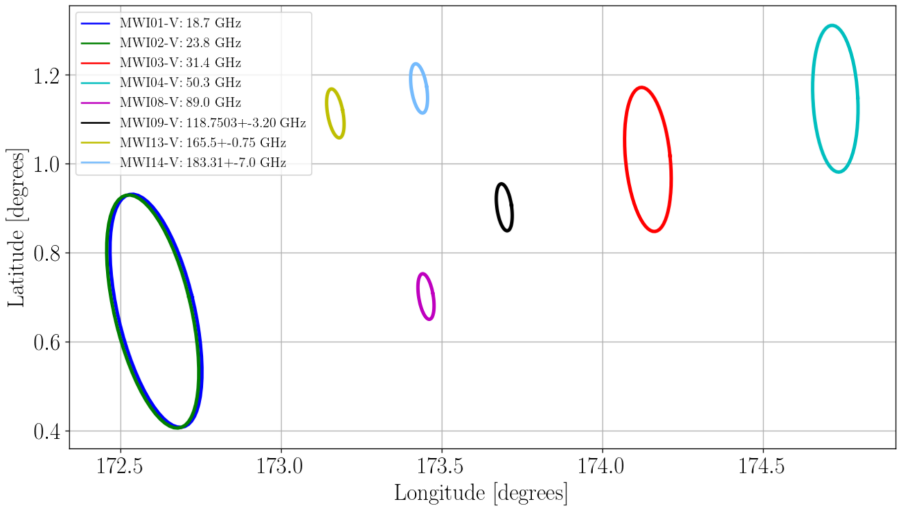


Figure 2: Taken from Rydberg and Eriksson (2019) (their Fig. 3), showing geolocations and FOV sizes (-3 dB) for the 8 MWI feedhorns at a given observation vector. As with Fig. 1, the same satellite position is assumed.

- **MWI-05070** The MWI radiometric sensitivity requirement is the NEDT in Table 1, representing one standard deviation of noise. The noise contribution of the calibration process is included.
- **MWI-05080** The MWI radiometric bias shall be better than those listed in Table 1.
- **MWI-05090** Orbit stability: Variations of radiometric biases of the measured MWI brightness temperature during an orbit shall be < 0.6 K.
- **MWI-05100** Lifetime stability: Variations of the running average of radiometric biases over any one orbit of the measured MWI brightness temperature shall be < 0.25 K throughout the mission lifetime.
- **MWI-05110** Inter-channel radiometric bias differences between brightness temperatures of the same MWI spatial sample shall be < 0.6 K.
- **MWI-05120** Inter-footprint radiometric bias differences between brightness temperatures of the same MWI spectral channel shall be < 0.4 K.

2.2 EUMETSAT user requirements to be evaluated using vicarious cal/val: ICI

As in the previous section, below are relevant requirements extracted from the EURD for ICI:

- **ICI-06060** The dynamic range of the ICI shall cover the spectral radiances in terms of brightness temperatures covering the range from 130 K to 300 K.
- **ICI-06065** All ICI requirements shall be met over the dynamic range given in ICI-06060.
- **ICI-06070** The ICI radiometric sensitivity requirement is given in the fifth column of Table 2. The values listed are to be understood as one standard deviation. The noise contribution of the calibration process is included.
- **ICI-06080** The ICI radiometric bias shall be better than those listed in Table 2.

- **ICI-06090** Orbit stability: Variations of radiometric biases of the measured ICI brightness temperature during an orbit shall be < 1.1 K.
- **ICI-06100** Lifetime stability: Variations of the running average of radiometric biases over any one orbit of the measured ICI brightness temperature shall be < 0.6 K throughout the mission lifetime.
- **ICI-06110** Inter-channel radiometric bias differences between brightness temperatures of the same ICI spatial sample shall be < 0.6 K.
- **ICI-06120** Inter-footprint radiometric bias differences between brightness temperatures of the same ICI spectral channel shall be < 0.6 K.

3 Background

3.1 Radiative transfer

In the microwave and sub-mm part of the spectrum, radiances are primarily sensitive to gaseous emission from oxygen and water vapour, along with emission and scattering from hydrometeors with sizes spanning several orders of magnitude—from cloud drops and cirrus to rain drops, snowflakes, hail, and graupel—roughly $10\mu m$ to 1cm. Generally speaking, lower frequencies experience less gas absorption and rarely observe scattering from hydrometeors, whereas higher frequencies observe greater optical depths, caused by absorption from cloud water and water vapour continuum absorption as well as scattering from rain and ice particles. Strong absorption bands exist for water vapour at 22.235 and 183.31 GHz, with strong oxygen bands existing between 50 to 70 GHz and a weaker one at 118.75 GHz. Further water vapour absorption features occur through the sub-mm range and ICI will make use of bands around 325 GHz and 448 GHz. Traditionally, sensors that primarily sample the strong absorption bands have been called "sounders" from their ability to provide vertical information for atmospheric profiles, whereas sensors primarily sampling the window regions of the microwave spectrum are called "imagers". As noted by Eriksson *et al.* (2020), MWI and ICI are both termed as imagers but MWI in particular holds a mixture of sounding and imaging channels, further narrowing the traditional separation imagers and sounders; ICI will also provide extensive profile information of water vapour, especially in drier profiles (see Eriksson *et al.* (2020) their Fig. 8).

Water vapour continuum absorption increases quasi-linearly with frequency. Thus there is greater water vapour profile sensitivity from a "window" channel at 165 GHz than say 31 GHz. This effect becomes more pronounced with window channels on ICI due to their higher frequencies. The 243 GHz channels on ICI are considered window channels, which is why they measure both V- and H-polarisation, but their sensitivity to water vapour means that they can act as sounding channels especially in the tropics, perhaps in a similar way as the 150-166 GHz channels on SSMIS and GMI, which are assimilated in lower latitudes for this reason (Lonitz and Geer, 2020). See Fig. 3 for an example of gas absorption effects visible on frequencies up to 1000 GHz (covering the frequency range of ICI), with locations of ICI frequencies indicated.

Cloud impacts on Brightness Temperature (TB) tend to increase with frequency and vertically-integrated ice mass or Ice Water Path (IWP). The magnitude of scattering signals from higher-frequency ICI channels is several times larger than those at currently-available frequencies in space (i.e. 190 GHz and below) (Geer *et al.*, 2021). These scattering signals are significantly larger than most expected radiative transfer errors (Barlakas *et al.*, 2022) and should provide sensitivity to relatively thin ice clouds (Eriksson *et al.*,

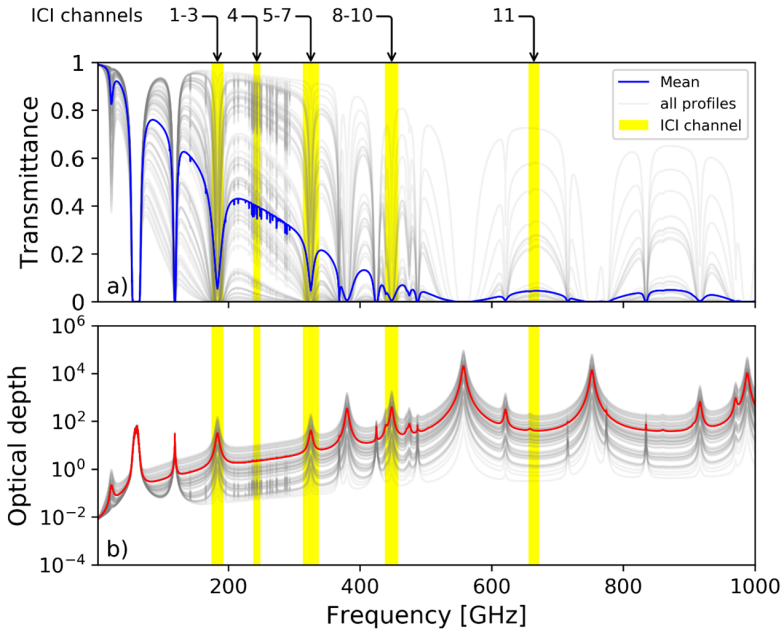


Figure 3: Taken from https://nwp-saf.eumetsat.int/site/featured-articles/microwave-and-sub-millimetre-spectroscopy-for-metop-second-generation/: "Total surface to space a) transmittance and b) optical depth for 83 diverse atmospheric profiles (grey lines) and their mean (bold lines), produced by the AMSUTRAN line-by-line model. The coverage of ICI channels are shaded yellow."

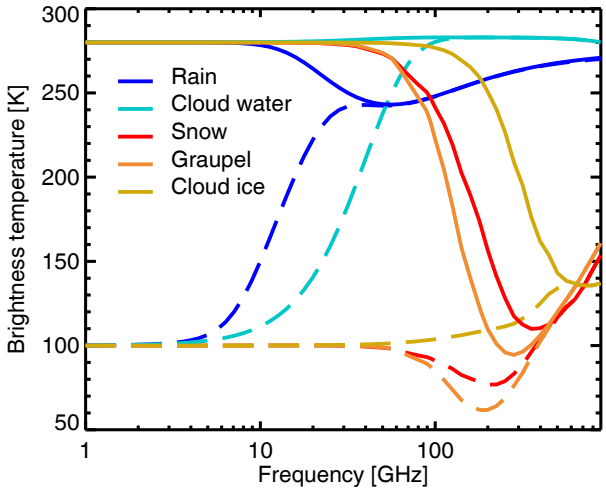


Figure 4: Taken from Geer et al. (2021) (their Fig. 8): "Cloud-top brightness temperatures simulated from uniform slabs composed of one of the default hydrometeor types...present in a 2km thick layer with a water content of 1e-3 kgm^{-3}. The cloud temperature is 253K if frozen (snow, graupel or cloud ice), or 283K if melted (rain, cloud water). Upwelling brightness temperature below the cloud is 280K (solid lines) or 100K (dashed lines)." Reproduced under the Creative Commons Attribution 4.0 License from: https://doi.org/10.5194/gmd-14-7497-2021

2020). Figure 4 shows the radiative effect of a theoretical cloud of various species as a function of frequency. This nicely demonstrates how sensitivity to liquid species typically increases with frequency up to mid microwave frequencies, whilst sensitivity to frozen species only picks up at higher frequencies, with cloud ice in particular largely invisible below about 200 GHz.

Modelling the scattering from frozen hydrometeors is a real challenge at microwave frequencies, as it is a multi-dimensional problem not well-constrained by current models (Geer, 2021). Scattering signals are sensitive to the number concentration and mass of particles in a volume, but also the particle shapes of the scatterers. The dielectric properties of water and ice change with temperature, and a volume of air can hold mixtures of hydrometeors in liquid, frozen, and supercooled states. It is thus a challenge to adequately model the radiative effects of clouds and precipitation, particularly on higher frequency channels. Concerning calibration and validation of MWI and ICI, cloud-affected scenes will often be removed from the population of radiances analysed, a topic discussed further in the section on the scientific method (Sec. 6). Such is the case in this study, where radiances with significant cloud impacts are screened out when considering the calibration.

Lastly, the frequencies measured by ICI also have some non-negligible sensitivity to atmospheric ozone. Whereas microwave bands up to 183 GHz typically have low sensitivity to ozone ($< 0.1K$), the highest-frequency channels on ICI have ozone sensitivity of 1 K or more (Mattioli *et al.*, 2019). As stratospheric ozone concentrations are variable in time and space, ozone effects are a non-negligible source of radiative transfer error for ICI. If the model does not have an accurate distribution of ozone then this would also be a possible source of background error. Variable ozone from the model is not included in the results presented here, as the provided test data did not consider variable ozone. See Appendix B for discussion of ozone's impact on the radiative transfer for MWI and ICI frequencies. Full forward modelling with variable ozone is expected to be in place in the IFS before the launch of MWI and ICI.

In this project, the forward model used in the IFS is RTTOV-SCATT v13.2, which will be part of the IFS cycle expected for release in 2024, Cycle 49r1. This version includes improved ice particle scattering (Geer *et al.*, 2021) and support for a new surface emissivity model that supports sub-mm frequencies (see following section). In addition, a scheme now exists within RTTOV-SCATT to parametrize the effect of polarised scattering from oriented ice particles that is used in the simulation of radiances (Barlakas *et al.*, 2021). The absorption coefficients used for MWI and ICI follow the latest spectroscopy (Turner *et al.*, 2022), including updated water vapour absorption, variable ozone, and use of measured spectral response functions[4] where available for ICI (Emma Turner, personal communication). These advances mean that simulation of MWI and ICI radiances employs the state of the art for fast radiative transfer calculation for microwave and sub-mm radiation.

3.2 Surface emissivity

Surface emissivity is a key concern for MWI, as a majority of MWI channels are sensitive to the surface in most atmospheric conditions. In contrast, the strong water vapour absorption of many ICI channels means that these are less sensitive to the surface emissivity model; however, ICI channels 1, 4, and 5 will have surface sensitivity in most situations, and in a dry atmosphere such as for mid-latitude winter scenes, channels 2, 6, and 7 could also have significant surface sensitivity (Eriksson *et al.*, 2020, their Fig. 4). It is worth stressing that even for channels with limited surface sensitivity (i.e. sounding channels), significant emissivity errors could manifest in O-B monitoring and result in erroneous atmospheric increments in

[4]RTTOV coefficient file with combined MWI and ICI produced by NWP-SAF on March 6, 2023, version 13.2. This is not yet officially available from the NWP-SAF (as of early 2024), and some centre frequencies were incorrect in this version relative to the PFM (52.7 vs. 52.61 GHz for channels 9 and 10, 165.5±0.725 rather than ±0.75.

data assimilation (English, 2008; Bormann, 2022), particularly in clear atmospheric conditions. This is most pronounced over snow and sea-ice surfaces where departures can be several Kelvin (see Figs. 8 and 9 in Bormann (2022)), arising from a poor representation of surface temperature and emissivity. These areas are hence excluded from analysis in this study.

3.2.1 Ocean emissivity

Emissivity of the ocean surface is primarily a function of frequency, incidence angle, temperature, and wind speed in the microwave part of the spectrum (Wilheit, 1979; Wentz, 1983). Other factors like relative wind direction (from the line of sight) and sea surface salinity also have a small impact on the emissivity, depending on the frequency (Meissner and Wentz, 2012; Reul *et al.*, 2009). The surface emissivity of a still ocean scene can be approximated with a dielectric model alone. Different parameterizations are used to approximate the effect from larger waves and the sea foam that forms at higher wind speeds, typically about 10 m/s and above.

In the IFS, microwave emissivities over sea have been calculated via the Fast ocean microwave emissivity model (FASTEM) model embedded within RTTOV-SCATT for many years. The latest version of this emissivity model (FASTEM-6) has been in operational use since 2015 (Kazumori and English, 2015). However, FASTEM is limited in frequency space to an upper limit of 200 GHz, and this is insufficient for the ICI frequencies. One effort to expand ocean microwave emissivity models in frequency space is the Tool to Estimate Sea-Surface Emissivity from Microwaves to sub-Millimeter waves (TESSEM2) model (Prigent *et al.*, 2017), which applied a neural network model to support emissivity calculation from 10 to 700 GHz; TESSEM2 is based upon FASTEM at frequencies below 200 GHz but follows a physical model at higher frequencies. TESSEM2 has been an option for use in RTTOV since version 12, but it will be dropped from RTTOV version 14 and hence is not a good choice here.

The reason for dropping TESSEM2 is that a more recent international effort has focused on producing a reference-quality model of ocean emissivity across a wide range of frequencies (English *et al.*, 2020). As part of this project, the reference-quality Passive and Active Reference Microwave to Infrared Ocean (PARMIO) model spans 1 to 1000 GHz (Dinnat *et al.*, 2023). A fast version of the PARMIO model was developed for use with fast codes such as RTTOV, named SURFEM-ocean (Kilic *et al.*, 2023, referred to simply as SURFEM from here). The SURFEM model has been packaged with RTTOV since version 13.2, which was released in late 2022. This new version of RTTOV has been tested in the IFS and thus the SURFEM model is now available for use in the current project. An advantage of the SURFEM model is that development is ongoing, whereas no path for future upgrades exists with the current FASTEM model. In-depth comparison of SURFEM with FASTEM-6 in the IFS (Geer *et al.*, 2024) suggested that wind-dependent biases were improved for lower frequencies, whereas there were larger O-B biases particularly in the higher-frequency window channels (37 - 166 GHz) and larger biases in particular for mid-frequency H-pol channels. However, in most cases the mean biases were dealt with sufficiently by the variational bias correction, and the overall impact in the IFS was a small improvement in short-range forecast quality. Hence, SURFEM is the ocean emissivity model used by the IFS starting with Cycle 49r1, the model version which this project is based upon.

3.2.2 Land, snow, and sea-ice emissivity

The emissivity over land and frozen surfaces is highly variable in the microwave part of the spectrum. The spectral signature depends upon many factors for frozen surfaces, including the grain size of snow and the age of sea-ice; for land surfaces aspects such as vegetation and soil moisture can impact the

emissivity. Multiple surfaces can exist within a single microwave FOV containing bare ice, land, and/or ice covered with snow of different depths, melted pools, and so on. This is already a complex topic in the relatively well-studied frequency range of current microwave sensors. For frequencies above 190 GHz, there are few studies to go on. Wang *et al.* (2017) extended a previous climatology of microwave emissivities based on the old SSMI sensor up to and beyond 200 GHz by extrapolation. Munchak *et al.* (2020) retrieved emissivities from GMI over various surfaces using a five-year data set, though this extended only up to 166 GHz. Furthermore, surface temperature from model analyses or short-range forecasts may contain very significant biases of perhaps 10 or 20 K in extreme cases (Bormann, 2022). A fundamental issue is that surface penetration depths vary across the spectrum and so the effective emitting temperature of the surface can vary widely. Thus, surface emission errors can be caused by errors in the assumed emissivity or the surface temperature.

In this study, a dynamic emissivity retrieval is used to solve for the surface emissivity over all non-ocean surfaces. This approach follows work to assimilate surface-sensitive microwave observations for non-ocean scenes by Baordo and Geer (2016) and Geer *et al.* (2022). Dynamic emissivity retrieval is more flexible than an atlas-based approach and typically more accurate, though in the case of clearly erroneous retrievals an atlas value from TELSEM2 (Wang *et al.*, 2017) is used as the backup. For mixed scenes that are part-ocean, the fraction that is ocean is fixed using the ocean emissivity model (i.e. SURFEM), with only the remaining fraction retrievable by the dynamic emissivity retrieval; this serves to better constrain the retrieval and leads to smaller overall background departures (see Sec. 4 in Geer *et al.* (2022)). For the dynamic emissivity retrieval, the skin temperature from the model is fixed (i.e. the location-specific temperature from the model is not modified by the retrieval) and the emissivity is retrieved from an appropriate window channel. This is then used to specify the emissivity for assimilating higher-frequency channels. For most purposes, the land emissivity is extrapolated as constant in frequency. However, for some vertically polarised microwave imager channels, the frequency-dependence of an atlas is imposed (Geer *et al.*, 2022; Lonitz *et al.*, 2022). For the 183 GHz sounding channels, a constant-in-frequency extrapolation is used, with dynamic emissivity retrievals around 90 GHz used to specify the land surface emissivity over warmer surfaces, and 150 GHz over colder surfaces (likely snow-covered surfaces and sea ice). The initial approach for ICI will exactly follow the treatment of 183 GHz channels, including the constant-in-frequency extrapolation and the use of 90 GHz and 150 GHz dynamic retrievals, though there is future scope to consider 243 GHz for dynamic emissivity retrievals over colder surfaces. The dynamic emissivity retrieval has the potential downside of aliasing atmospheric signal into the emissivity estimate. There can thus be a trade-off between using the surface-sensitive channel with the cleanest surface signal and using a channel close in frequency space to the channel of interest.

3.3 The Integrated Forecasting System and all-sky assimilation

The Integrated Forecasting System (IFS) produces operational weather forecasts at the European Centre for Medium-range Weather Forecasts (ECMWF). The operational version of the IFS has been Cycle 48r1 since June 2023 (ECMWF, 2023), with a planned upgrade to Cycle 49r1 in the second half of 2024. This state-of-the-art NWP model ingests tens of millions of observations every day to provide the most accurate possible initial conditions for the forecast model.

Most passive microwave radiance data passes through the "all-sky" system in the IFS, that is, the data are used in clear, cloudy and rainy conditions, with model clouds and precipitation included in the radiative transfer. All-sky assimilation of microwave radiances was developed at ECMWF and first applied to conically-scanning microwave imagers (Geer and Bauer, 2010; Bauer *et al.*, 2010). It has since been extended to microwave sounders for humidity (Geer *et al.*, 2014; Lawrence *et al.*, 2018) and temperature

(Duncan *et al.*, 2022b) in the IFS whilst also proliferating at other centres (Geer *et al.*, 2018). As of mid 2023, there are 4 imagers, 5 humidity sounders, and 5 temperature sounders assimilated operationally in the IFS via the all-sky framework using direct radiance assimilation.

All-sky assimilation is predicated upon use of a "symmetric" cloud amount to scale observation errors (Geer and Bauer, 2011). This recognises that observation errors are larger in cloudy or precipitating conditions due to issues in forward modelling and representation. Rather than discarding these more problematic observations, the observation errors are inflated as a linear function of some cloud proxy (C). The symmetric nature of the observation error model simply means that observed (obs) and modelled (background, B) cloud amount are weighted equally when calculating C_{sym}:

$$C_{sym} = (C_{obs} + C_B)/2 \tag{3}$$

For microwave imagers like SSMIS and GMI, there are separate cloud proxies used over sea and land due to the different sensitivities. These are referred to as the 37 GHz polarisation difference (P37) over sea and the scattering index (SI) over land, where V and H refer to vertically- and horizontally-polarised radiation, respectively, and clear refers to clear-sky radiative transfer:

$$C_{P37} = 1 - (TB_V - TB_H)/(TB_{V,clear} - TB_{H,clear}) \tag{4}$$

$$C_{SI} = TB_{91GHz} - TB_{150GHz} \tag{5}$$

In the symmetric approach, the quantities TB can be obtained either from observations or simulations as required, except for $TB_{V,clear}$ and $TB_{H,clear}$, which are always obtained from simulations. The scattering index here is defined using 91 and 150 GHz channels (a la SSMIS) but can be computed using two window channels at nearby frequencies (GMI with 89 and 166 GHz channels), lower frequencies (AMSU-A with 23 and 89 GHz), or potentially higher frequencies as well. The key is to use two channels with differential sensitivity to scattering so that C_{SI} increases with cloud and precipitation—this is accomplished because scattering sensitivity generally increases with frequency. It is possible that the higher frequencies observed by ICI will prove useful in this context, as a scattering index that employs the 243 GHz window channel should be better able to identify cirrus clouds, for example. However, scattering indices may start to become problematic above 183 GHz, especially for thicker clouds and precipitation, due to the transition from primarily scattering to absorption (and thus increased brightness temperatures due to cloud) which occurs typically between 183 and 500 GHz (Geer *et al.*, 2021, also Fig. 4 here).

Model equivalents for TB at the observation location are provided by the observation operator RTTOV-SCATT, including contributions from clouds and precipitation. The clear-sky equivalents used in Eq. 4 consider only emission from the surface and gases in the atmosphere, i.e. no emission or scattering from hydrometeors. Model equivalents are calculated at the observation location and represent the model fields at that point—they are not currently averaged to match the sensor footprint. This is not currently considered a significant shortcoming in the IFS for all-sky assimilation because the model "effective resolution" even at the current operational model resolution of 9 km is on order 30 to 40 km, as the effective resolution of fields like cloud and precipitation is at least three times larger than the model grid resolution (Klaver *et al.*, 2020). However, special treatment is necessary for mixed-surface scenes. Here, the FOV-integrated land sea mask is used, not the land sea mask of the model grid point. The difficulty of simulating radiances from coastal scenes (i.e. land fraction of 1 to 99%) with sufficient fidelity for calibration purposes means that these will be excluded from all analysis in this project.

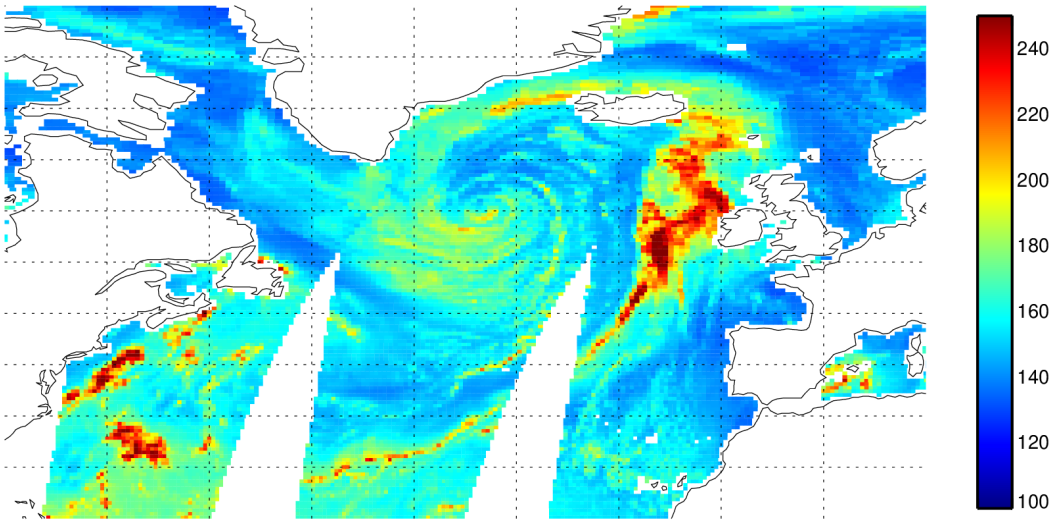


Figure 5: Observed TBs after superobbing for the 37H channel on AMSR2 (similar to MWI channel 6), as observed October 23, 2021. Only superobs over sea are shown, with units in Kelvin.

Most microwave radiances are averaged prior to ingest in the IFS, with "super observations" or superobs created in the pre-processing of radiances. As of Cycle 49r1 (the model version used in this study), microwave imager radiances are superobbed to 40 km on a Gaussian grid. An example of 40 km superobs from an imager channel is given in Fig. 5, showing observed TBs from the 37H channel from AMSR2, approximating how MWI channel 6 will appear after pre-processing. The superobbing procedure calculates a simple average of all TBs whose geolocation midpoint falls within the grid box. The zenith and azimuth angle are typically taken from the observation nearest to the grid box centre, with the superob latitude and longitude given by the grid box midpoint. Superobbing acts to homogenise all-sky observations, decrease representation error, reduce geolocation differences between channels, and beat down sensor noise (e.g. Duncan *et al.*, 2023). All valid superobs have departures calculated as they are processed through the all-sky data path.

3.4 Similar current sensors

MWI and ICI will introduce new capabilities for Earth observation, but they build upon designs and concepts that have been in operational use for many years. This is a significant advantage for calibration and validation, as existing knowledge can be brought to bear and existing sensors can be used for testing and prototyping prior to launch. Table 5 provides names and characteristics for satellites that hold similar sensors, then Table 6 holds details on these sensors themselves.

As both MWI and ICI are conically-scanning radiometers, it makes sense to use other conically-scanning radiometers as the primary point of comparison; thus currently assimilated imagers SSMIS, AMSR2, and GMI are all useful for comparison to MWI. GMI and SSMIS are particularly useful for comparison as they hold many of the same channels found on MWI and cover roughly the same spectral range. The one downside to comparing with GMI and SSMIS is that they lack the 118 GHz channels found on MWI; for these the best point of comparison is MWHS-2, a cross-track sensor but still the only operational radiometer observing around the 118 GHz absorption line. As no sub-mm frequencies are currently

Table 5: Satellites holding similar passive microwave sensors. Acronym definitions are available in the glossary. Equator crossing times are given for all polar-orbiting satellites, with an A for ascending or D for descending node. Only currently functional satellites are included (as of 2023).

Platform	*Satellite(s)*	*Sensor(s)*	*Launch*	*Altitude [km]*	*ECT*
DMSP	F-17, -18	SSMIS	2006, 2009	848	06:40 D, 04:50 D
FY-3	FY-3D	MWRI	2017	836	13:30 A
FY-3	FY-3D, -3E	MWHS-2	2017, 2021	836	13:30 A, 05:40 D
GCOM-W	GCOM-W	AMSR2	2012	700	13:30 A
GPM	GPM Core	GMI	2014	407	-
TROPICS	TROPICS-01	TMS	2021	550	13:30 A

Table 6: Similar passive microwave sensors currently in use.

Sensor	*Freq. Range [GHz]*	*Scanning*	*Channels*
AMSR2	6.925 - 89	Conical	16[5]
GMI	10.65 - 190	Conical	13
MWRI	10.65 - 89	Conical	10
MWHS-2	89.0 - 190	Cross-track	15
SSMIS	19.35 - 190	Conical	24
TMS	91 - 205	Cross-track	12

flown on any operational platform, it is harder to find useful proxies for ICI channels. The 183 GHz channels on ICI (1-3) have counterparts on GMI and SSMIS, but the only channel currently measuring above 200 GHz is on the TROPICS TMS at 205 GHz.

4 MWI and ICI processing in the IFS

4.1 Data flow

All observational data ingested by the IFS needs to be in BUFR format and if necessary is converted from its original format before ingest. The data then undergo certain pre-processing procedures prior to being used in data assimilation, where the data format is changed to ODB. For most microwave imagers, the pre-procesing stages include some basic quality control, averaging into superobs, and conversion from BUFR to ODB.

[5] AMSR2 has two 89 GHz scans, A and B.

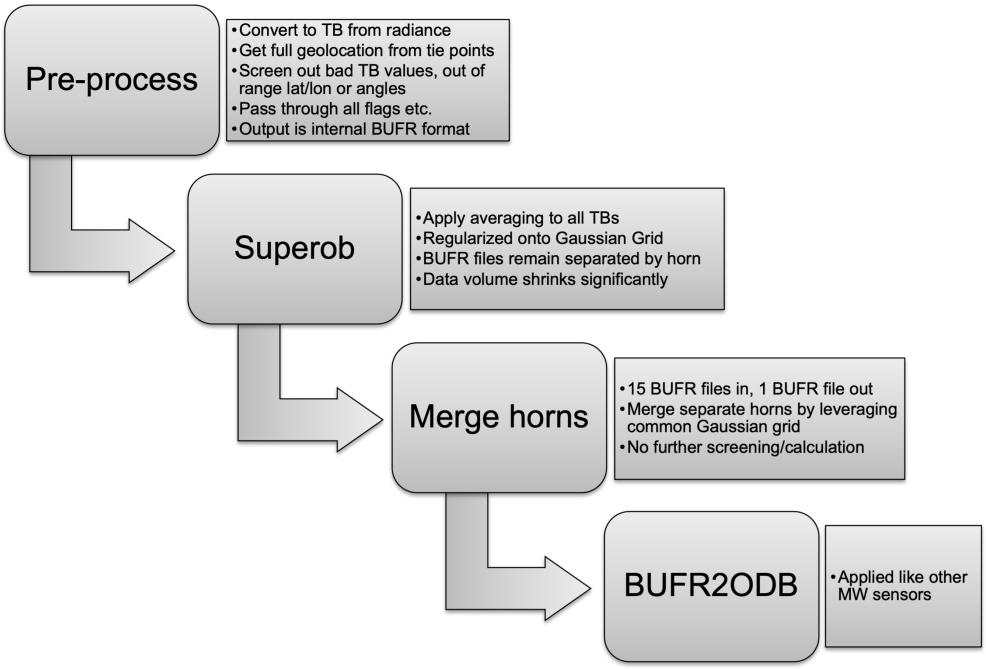


Figure 6: Pre-processing data flow for MWI and ICI radiances prior to ingest in the IFS. L1B radiances from EUMETSAT are read in and converted to a local BUFR sequence, superobbed, merged into a super-sensor BUFR file that contains data from both sensors, then converted to the ODB file format expected by the IFS.

A few peculiarities of MWI and ICI design and processing require modifications to this typical data flow. First, MWI and ICI have feedhorns that point at different geolocations (Fig. 1), and the IFS assumes that channels from the same sensor are viewing the same point; related to this point, geolocations and geometry information are provided on a per-horn basis. Second, due to the heavy oversampling of these sensors and the large number of channels, geolocation and geometry information is provided at specific "tie points" along the scan so as to decrease total data volume. Third, the measurements are provided in radiance space ($Wm^{-2}sr^{-1}(cm^{-1})^{-1}$) whereas the IFS expects brightness temperatures in Kelvin. Last, for purposes of scientific synergy in the IFS, MWI and ICI will be combined into the super-sensor MWIICI, with all 39 channels in one file. With these aspects in mind, the IFS pre-processing for MWI and ICI L1B data needs to include conversion from radiance to TB (detailed in the relevant product format specifications, see Appendix A), provide geolocation and geometry information for every channel at every sample across the scan, and homogenise the measurements to the same geolocation for processing in the IFS. This pre-processing should produce a combined MWIICI ODB file to the IFS that contains all 39 channels in TB space with common geolocations, whilst passing through all relevant flagging and geometry information from L1B. The IFS can then treat MWIICI as a single sensor.

Compared to other microwave BUFR data, an extra pre-processing routine is required at the outset to handle two of these aspects: conversion to TB space and interpolation of geolocation from tie points to all samples across the scan. This *preprocess.py* script uses the ECMWF-developed ecCodes[6] software package to decode and encode BUFR files in Python. As part of this step, an internal "L1B2" BUFR sequence is defined for the output format that is in TB rather than radiance space (see Appendix C for BUFR sequences); thus the first step of the data flow reads in L1B BUFR and outputs a L1B2 BUFR file with most fields unchanged but TB in place of radiance and geolocations present for every sample. This

[6] https://confluence.ecmwf.int/display/ECC/ecCodes+Home

is the first step in the pre-processing data flow for MWI and ICI data (see Fig. 6). Initial quality control can be done at this stage, for instance in the case of unphysical TBs ($50 > TB$ or $330 < TB$) or scans marked as poor quality.

The internal BUFR sequence contains positions for all 39 channels of the combined MWIICI, flexibly defined via the "delayed replication factor" allowed by BUFR. The pre-processing script converts the relevant fields but also splits each MWI and ICI BUFR file into several files—one for each feedhorn. Thus the output for this pre-processing step is 15 total files from two input files. Key parameters that vary on a horn-to-horn basis such as zenith angle and land fraction are now provided on a per-channel basis to ensure this information is passed through properly. The entries for all channels other than those of the horn in question are left missing for now. This permits the per-horn data to be combined at a later stage whilst using the same BUFR sequence.

Next is the procedure of superobbing (see Sec. 3.3), where TBs are averaged together on a Gaussian grid at 40 km resolution (N256 Gaussian grid[7]). This needs to be done on a per-horn basis because each feedhorn has separate geolocations and the superobbing code expects vectors of observations with a common geolocation. Superobbing is done in *bufr_grid_screen.F90* (hereafter BGS). A major benefit of the superobbing procedure is that it regularises the grid of the observations, making it simpler to combine data from each feedhorn at the next step. BGS ingests the L1B2-formatted BUFR data and outputs BUFR files with the same sequence. The other upside of superobbing is the significant reduction in data volume. One 40 km superob should contain data from at least 3 scans and roughly 18 samples per scan from ICI (about 33 from MWI), so there is a reduction in data points of roughly 50 times for ICI and 100 times from MWI due to superobbing alone. In terms of data volume, this results in a reduction of only about one order of magnitude, due to the horns' data remaining in separate files at this stage, which limits data compression.

The superobbing procedure homogenises the BUFR data onto a known Gaussian grid at 40 km, which permits exact geolocation matching of data from the individual horns. After superobbing, the merge task takes BUFR data from all 15 horns as input and outputs a single BUFR file with all 39 channels combined at a single geolocation each—what we can call L1R format (following Maeda *et al.* (2016), with "R" for resampled), but noting that it uses the same BUFR sequence as L1B2. No further screening is done at this step. Due to peculiarities of BUFR files, care needs to be taken to merge the disparate messages and subsets of each horn into a single set of messages and subsets. The horn matching script uses sets of lat/lon pairs, first finding all valid data from horn 1 (i.e. per message and subset) and filling in the missing channels on top of the horn 1 BUFR file. As a pragmatic choice, only observation vectors with a valid TB from either channel 1 or 7 is passed through the horn matching task, as these are the left- and right-most aligned MWI horns (see Fig. 2). Hence the same procedure is repeated for channel 7, with the unique set appended on to the channel 1 BUFR file. This choice is necessary due to a combination of the restrictive BUFR file format and the differently-aligned MWI and ICI feedhorns, but permits >99% of all observation locations through the horn matching task. This final BUFR file is over 100 times smaller than the total size of all L1B2 BUFR files that went into it, now helped by the common geolocations that allow significant data compression. This move to common geolocations results in about 20-fold reduction of data volume.

Lastly, the BUFR file containing all channels on the Gaussian grid needs to be converted to ODB format for ingest to the IFS (via the task bufr2odb). This is a step common for all microwave sensors and follows the same procedure for MWIICI. One notable element of the BUFR to ODB conversion is that ODB and IFS support per-channel zenith angles, but some variables are read into ODB as one value for all channels. Land fraction, azimuth angle, and solar angles are all read into ODB from channel 1

[7] https://confluence.ecmwf.int/display/UDOC/Gaussian+grids

(18V) to represent all channels. This is by necessity, as it depends what is supported in the all-sky code at present, but as the full information exists in the input BUFR file it could be fed through fully in the future. It is also best to pass through the land fraction information from the largest FOV (i.e. 18 GHz), as this provides the greatest sensitivity to coastal contamination for the purposes of data selection.

4.2 Processing in the IFS

Following the pre-processing steps described above, the data from MWI and ICI are now ready for input to the IFS as MWIICI. Modelled fields from a short-range forecast from the previous analysis constitute the model background. An operator interpolates model fields to the observation locations and then the forward model RTTOV-SCATT is called, producing model equivalents in TB space. These equivalents of the model background in observation space (B) should contain all radiometrically significant elements of the model including oxygen, water vapour, surface emission, and various hydrometeors from clouds and precipitation.

The primary parameter used for monitoring in this study is the difference between observations (O) and the model background (B), or $O-B$, defined in Kelvin. The mean and standard deviation of $O-B$ provide helpful information for diagnostic evaluation of observations as well as diagnosing model shortcomings such as misplaced cloud in the short-range forecast. $O-B$ is typically examined after bias correction, as a variational bias correction (hereafter VarBC) scheme is used to remove bias from most observations processed in the IFS prior to assimilation (Dee, 2004). In this study, most departure statistics will use $O-B$ without bias correction (D), rather than after bias correction (D_{BC}), where BC is defined as the bias correction offset applied by the IFS to each observation:

$$D = O - B \tag{6}$$

$$D_{BC} = (O - B) - BC \tag{7}$$

This means that we are implicitly assuming the IFS to be unbiased, but this is simply a pragmatic choice and allows direct comparison to other sensors, effectively allowing the IFS and a consistent data selection criteria to act as a standard for comparison. Comparison between sensors for calibration purposes will thus use departures prior to bias correction, D, though bias-corrected departures will also be available.

The output of the IFS in observation space is added to the input ODB file so that these "observation feedback" files from the IFS contain dozens of fields. In addition to calculating the $O-B$ and passing on crucial information such as latitude and longitude, zenith and azimuth angles, scan position, observation time, orbital angle (see Fig. 7), orography, and land fraction, the all-sky code also passes various model fields (from the 12 hour background forecast, and in some cases also from analysis) to the output ODB that can be used in downstream analysis. These include fields such as 10 m wind speed, skin temperature, TCWV, sea-ice fraction, columnar rain, snow, cloud water, and cloud ice, surface-to-space transmittance in clear-sky (τ), rain rate, cloud fraction, and many others. As all of these fields exist in the output ODB that is archived into MARS, these can be used in subsequent analysis of the calibration. The ODB contains all superobs that are input to the IFS.

Two key output parameters are worth explaining further. First, the cloud proxy (see Sec. 3.3) is computed for every all-sky observation for the purpose of determining the observation error to assign (Geer and Bauer, 2011). This is channel- and surface-dependent to permit channels with different cloud sensitivities to use separate proxies such as P37 or SI, as one cloud proxy may perform better over a given surface than another. As an initial implementation, P37 and SI are the cloud proxies used for MWIICI (depending on the surface type); it is conceivable that new proxies could be developed using ICI channels for example,

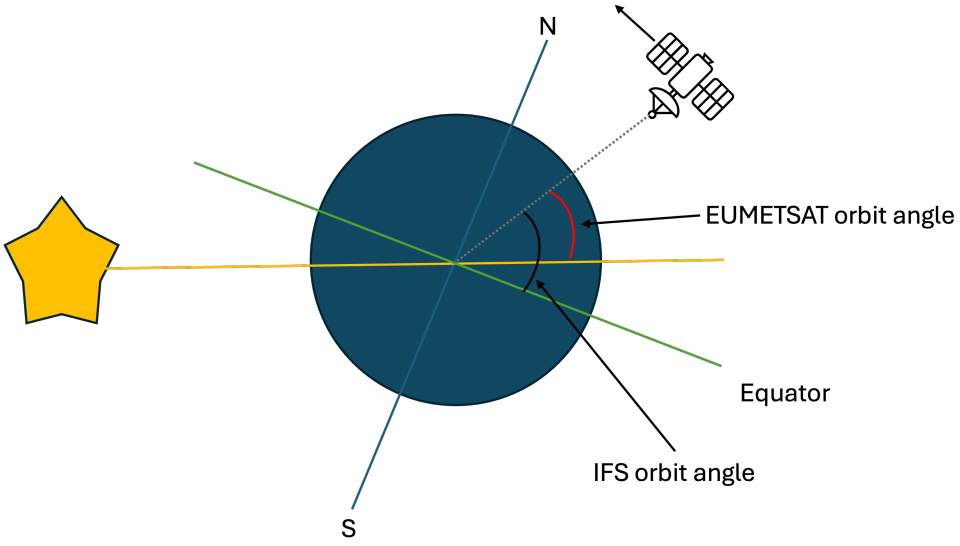


Figure 7: Cartoon of how the IFS and EUMETSAT orbital angles are defined.

but that is outside the scope of this project as it is only important for assimilation. For the purposes of this study, the cloud proxies for observation error can also be helpful diagnostic indicators of cloud contamination in either the observation or the model background. The proxies that are computed for use in the symmetric error model are also stored in the ODB for further use. However, in the end neither P37 or SI cloud proxies were used for data selection in the current work but rather the cloud impact parameters defined in Sec. 6.1. Second, the all-sky code produces various flags (contained within the bitfield *datum_tbflag*) that can also be used for diagnosing cloud or other potential issues with the observation, such as cold sector contamination (Cold air outbreak (CAO), see Lonitz and Geer (2015)), a failed emissivity retrieval, cold skin temperature, and many others. These flags can also be helpful in determining a suitable population for compiling statistics. For example, a good population for calibration purposes may be clear-sky, ice-free ocean where SST is within a certain range—this is easily determined from the IFS output and can be further stratified by parameters such as orbital angle, wind speed, and so on.

The orbital angle in the IFS takes a "circular" orbit and calculates an angle between the equator, the centre of the earth, and the intersection of the ascending node of the orbit and the equatorial plane (Bormann *et al.*, 2023), visualised by the cartoon in Fig. 7. The IFS angle is 0 degrees at the equator for the satellite's ascending node and then -180 at the equator for the descending node. For conically scanning radiometers, there is an offset between orbital angle and spacecraft position from using observation geolocations to calculate the orbital angle. Notably, this definition of orbital angle differs from that provided by EUMETSAT in the L1B data, which uses the sun's ecliptic plane as the reference (starting at zero when the satellite crosses the solar ecliptic plane northbound). Both orbital angles are passed through to ODB for later analysis.

4.2.1 Channel treatment in the IFS

The all-sky framework relies upon shorter channel indicators (i.e. *53H* or *37V* for an H-pol 53 GHz or V-pol 37 GHz channel) to allow similar treatment of similar channels across various sensors. These don't need to share the same central frequency, but rather share general radiative properties. For example,

the 31.4V channel on AMSU-A is referred to as *37V* in the all-sky code because it exists in the same atmospheric window as 37 GHz and thus warrants similar treatment within the IFS for screening and so on. In Table 7, all 39 channels of the combined MWIICI super-sensor are given alongside their code indicators and channels from other sensors that share the same indicator. This is intended for use as a reference when comparing channels between sensors.

Some of the all-sky indicators are approximate for MWIICI, due to being implemented earlier for a different sensor (e.g. 37V for the 31.4V channel). Most indicators used for ICI are new for this implementation, plus the 183PM6.1 is unique for MWI. Note that the ICI 183 GHz channels require their own all-sky indicators, as the code expects only one per sensor; thus ICI has 183PM7I compared to 183PM7 for MWI. Channels from MWHS-2, AMSU-A, and ATMS are given as a comparison for some 50, 118, and 183 GHz channels, but these are not used for quantitative comparison due to their different scanning geometry.

5 Demonstration of the MWIICI processing

Three test orbits of simulated MWI and ICI radiances were produced by EUMETSAT for use by the broader community to prepare for future Metop-SG data. These data provide an invaluable opportunity to test the data flow described above. Two of the orbits of test data are from September 12th, 2007, with the other orbit from February 23rd, 2008. In this report, the focus is on orbits 4655 and 4656, which covers 8:43 to 12:01 Z on September 12th. The beginning of orbit 4655 crosses one of the long-window cutoff times for 4D-Var, which run from 9-21 Z and 21-9Z for 12Z and 0Z cycles respectively, hence the analysis here focuses on the 12Z cycle which contains most of 4655 and all of 4656.

The test data preparation is detailed in Rydberg and Eriksson (2019). It relies upon ERA5 (Hersbach *et al.*, 2020) for atmospheric fields and makes necessary assumptions about surface emissivities and hydrometeors to permit forward modelling of all MWI and ICI channels, in addition to using all geometric knowledge of the sensors as known at that time. Of relevance here is that ERA5 does not archive convective hydrometeor concentrations, so the atmospheric profiles used by Rydberg and Eriksson (2019) had to extrapolate from convective precipitation rates to water and ice concentrations aloft. There are numerous scientific differences in the modelling of the test data and the forward model in this study, and these are key to understanding and interpretation of the study's results—these are discussed in detail in Sec. 5.3. The test data contains added noise in line with expectations of channel NEDT (Rydberg and Eriksson, 2019, their section 7.2.1).

This section details the end-to-end processing of the MWIICI test data, from L1B BUFR to ODBs output from the IFS.

5.1 Before and after pre-processing

The un-averaged BUFR data can be seen in Fig. 8 for a channel each from MWI and ICI, showing one granule over the Iberian peninsula. Here the dot size in the figure is not representative whatsoever of the FOV, but rather was chosen to overlap as little as possible with neighbouring dots. The figures give a good sense of how densely MWI and ICI sample. The two frequencies plotted exist at the extreme ends of the MWIICI spectral range and thus demonstrate very different sensitivities—18 GHz offers primarily surface sensitivity whilst 664 GHz primarily sees water vapour and clouds.

One feature to note is that the swath edges can be quite different for separate feedhorns, for example

Table 7: Combined MWIICI super-sensor channel numbers and centre frequencies with their all-sky code indicator(s) if given, the dynamic emissivity channel used, and comparable channels from other sensors. Asterisks (*) mark where the comparison is for V-pol channels only; plus-minus (±) mark where the comparison is for H-pol channels only. For the dynamic emissivity channel, if an extrapolation is used following Lonitz et al. (2022) this is notated with a $. Channel types are given as window (W) or sounder (S). Feedhorn numbers are given per sensor. Mean zenith angle from test data given for illustrative purposes.

Ch. #	*Freq. [GHz]*	*Horn*	*Type*	*Zen. Ang.*	*All-sky code*	*Emis.*	*Similar channels*
1, 2	18.7	1	W	53.1	19V/H	19V/H	GMI 3,4; SSMIS 13,12
3, 4	23.8	2	W	53.1	23V/H	19V$^{\$}$/H	GMI 5*; SSMIS 14*
5, 6	31.4	3	W	52.9	37V/H	19V$^{\$}$/H	GMI 6,7; SSMIS 16,15
7, 8	50.3	4	W	53.2	50V/H	50V/H	AMSU-A 3*; SSMIS 1$^{\pm}$
9, 10	52.70	4	W	53.2	52V/H	50V/H	AMSU-A 4*; SSMIS 2$^{\pm}$
11, 12	53.24	4	S	53.2	53V/H	50V/H	AMSU-A 5$^{\pm}$; SSMIS 3$^{\pm}$
13, 14	53.75	4	S	53.2	54V/H	50V/H	AMSU-A 6
15, 16	89.0	5	W	52.3	89V/H	19V$^{\$}$/H	GMI 8,9; SSMIS 17,18
17	118.7503 ± 3.2	6	W	52.8	-	89V	MWHS2 8
18	118.7503 ± 2.1	6	W	52.8	-	89V	MWHS2 7
19	118.7503 ± 1.4	6	S	52.8	-	89V	MWHS2 6
20	118.7503 ± 1.2	6	S	52.8	-	89V	MWHS2 6
21	165.5 ± 0.725	7	W	54.2	166V	19V$^{\$}$	GMI 10
22	183.31 ± 7.0	8	W	54.1	183PM7	89V	SSMIS 9; GMI 13
23	183.31 ± 6.1	8	W	54.1	183PM6	89V	-
24	183.31 ± 4.9	8	S	54.1	183PM4	89V	MWHS2 14; ATMS 19
25	183.31 ± 3.4	8	S	54.1	183PM3	89V	GMI 12; SSMIS 10
26	183.31 ± 2.0	8	S	54.1	183PM2	89V	ATMS 21; MWHS2 12
27	183.31 ± 7.0	1	W	53.8	183PM7I	89V	SSMIS 9; GMI 13
28	183.31 ± 3.4	1	S	53.8	183PM3I	89V	GMI 12; SSMIS 10
29	183.31 ± 2.0	1	S	53.8	183PM2I	89V	ATMS 21; MWHS2 12
30	243.2 ± 2.5	2	W	51.8	243V	89V	-
31	243.2 ± 2.5	3	W	51.8	243H	89H	-
32	325.15 ± 9.5	4	W	53.9	325PM9	89V	-
33	325.15 ± 3.5	4	S	53.9	325PM4	89V	-
34	325.15 ± 1.5	4	S	53.9	325PM1	89V	-
35	448 ± 7.2	5	S	53.9	448PM7	89V	-
36	448 ± 3.0	5	S	53.9	448PM3	89V	-
37	448 ± 1.4	5	S	53.9	448PM1	89V	-
38	664 ± 4.2	6	S	51.8	664V	89V	-
39	664 ± 4.2	7	S	51.6	664H	89H	-

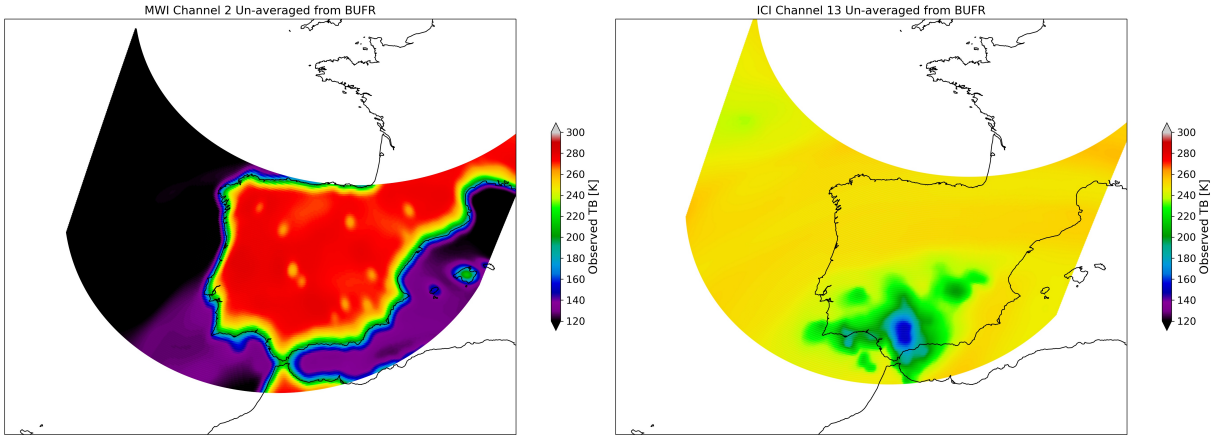


Figure 8: Un-averaged TBs from MWI 18H and ICI 664H over Spain from test orbit 4656.

when comparing the coverage of the Balearic islands on the right side of Fig. 8. To examine this more closely, we can view the edges of the scan for a selection of MWIICI horns as seen in Fig. 9. Here the scan edge observations of MWI 18 GHz (horn 1) are observed to be roughly 100 km away from two of the horns of ICI (243 and 664 GHz, horns 3 and 7). This is a limitation of the instrument geometry that needs to be kept in mind for analysis of several channels from one instrument or especially all channels on MWIICI. The total swath width of MWIICI that contains the full 39-channel observation vector is thus narrower than that of any individual channel.

Figure 10 shows the observed TBs from six selected MWIICI channels after superobbing. Here the superobs are plotted as rectangles to approximate the 40 km Gaussian grid boxes into which they are averaged. Comparison of the first and last panels of this figure with Fig 8 demonstrates the effect of superobbing. In these panels the differing swath position between channels is again visible, with slightly different coverage at swath edge seen for channels on the same sensor. The 40 km superob size appears sufficient to resolve the feature of deep convection seen over southern Spain in the higher frequency channels, with a maximum TB depression of about 100 K.

The act of superobbing provides a welcome data compression, which of course means some loss of information. However, one way to retain some of the information lost is to track the variability and number of observations within a superob, passing this through the BUFR and into ODB. Because we have control over the BUFR format for MWIICI, it was possible to add these fields to the BUFR sequence (see Appendix C). First we can look at the number of observations averaged together for a section of the test data in Fig. 11. The number of observations per superob is relatively consistent between channels on the same sensor except near the swath edge, where the feedhorn alignment causes differences. In the middle of the swath, the number of observations per superob is roughly 110 for MWI and about 70 for ICI. Near the swath edges, MWI channels can have 400 or more observations per superob, with 180 or more for ICI. To give an idea how much this oversampling can hypothetically decrease the random noise, superobbing should decrease NEDT by a factor of about 20 at swath edge (square root of the number of observations), and about a factor of 10 in the swath's middle. For example, for MWI sounding channels with a specified NEDT of 2 K, this should yield an effective NEDT of about 0.2 K or less at a minimum, a noise performance comparable to ATMS temperature sounding channels after superobbing. Such noise performance for sounding channels could be very significant for the eventual impact of MWI sounding channels when assimilated, though if there are correlated features to the observation error, then the reduction in overall error may not be so large.

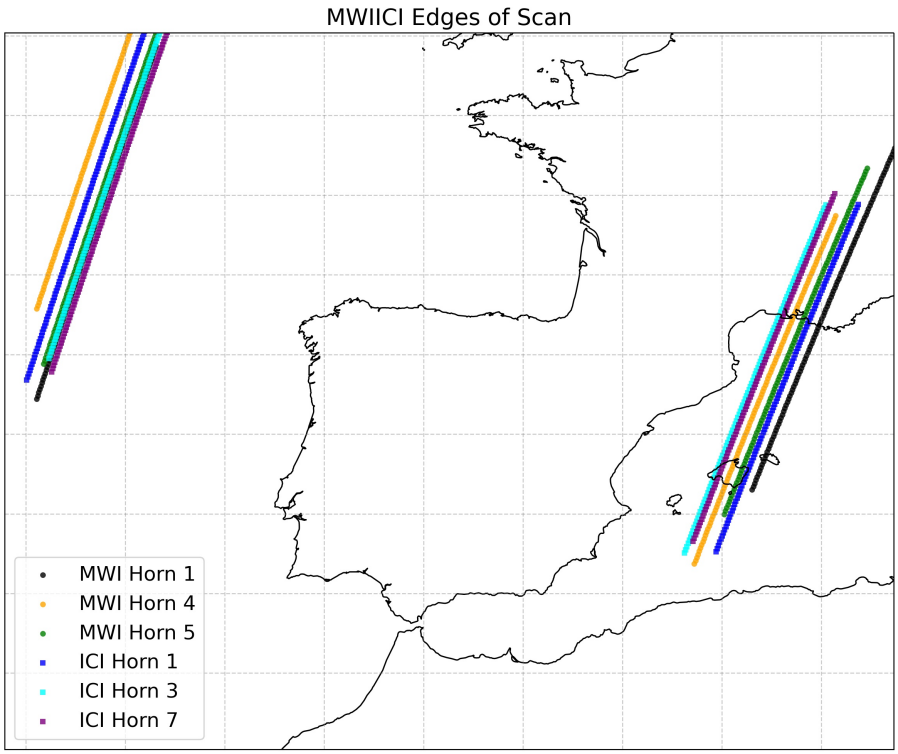


Figure 9: As Fig. 8, but showing only the first and last sample per scan for three selected horns on each instrument.

Variability of observations within each superob is captured by the standard deviation of TBs, seen in Fig. 12 for a selection of MWI channels over the same domain. The variability of observations within a superob is a new diagnostic for all-sky assimilation, and as seen in the figure, it is primarily a signal of coastlines for imaging channels. However, this metric could prove useful in identifying the variability of finer features like small-scale convection or even water vapour inhomogeneities. Especially once we have real observational data from Metop-SG, this diagnostic could prove useful to identify scenes with large representation errors or beam-filling effects in either surface or atmospheric properties. However, this diagnostic will not be used in the data selection, as will be discussed in Sec. 6.3.

5.2 Output from the IFS

5.2.1 Departures for MWIICI

For the purpose of calibration monitoring, the main output from the IFS is the background departure D, or $O-B$. After processing through the IFS for the date of the test data, September 12th 2007, we can examine the all-sky departure maps for MWIICI. The same selection of six MWIICI channels shown previously are found in Fig. 13, showing the all-sky departures over the Iberian peninsula. These maps of departures appear reasonable, in that they look like maps typically seen for a sensor like SSMIS or GMI. Coastlines show large positive and negative departures at imaging frequencies. The higher frequency channels of 183 GHz and above indicate a swirl of larger upper-level humidity in the background around a low pressure centred over southern Spain. The 243H channel shows an interesting mix of positive and negative departures near this feature, probably a combination of displacement errors and some disagreement over liquid and frozen hydrometeors. Consistent negative departures near the tip of Gibraltar

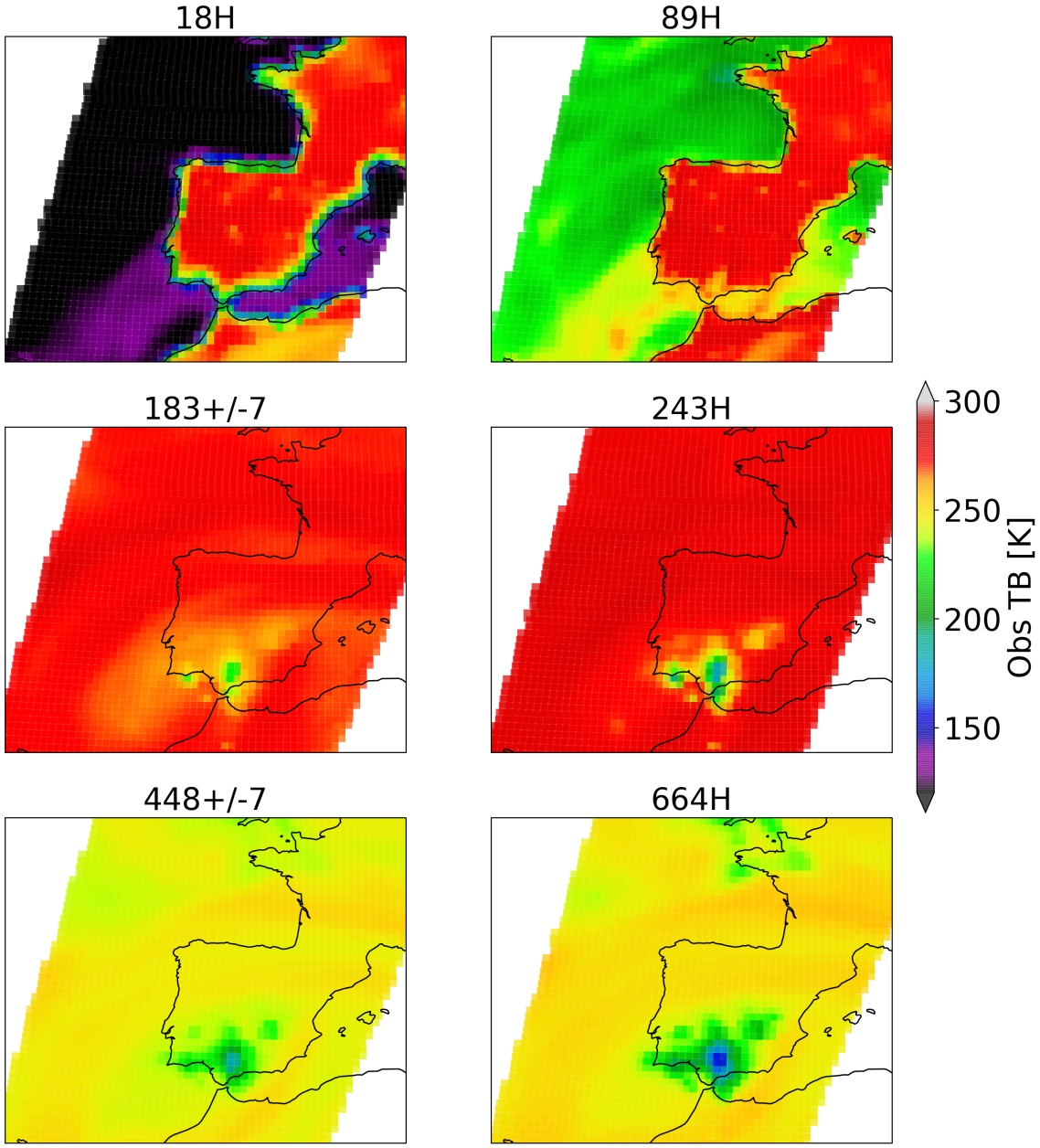


Figure 10: Superobbed TBs from six selected MWIICI channels over Spain from test orbit 4656. Here the 183 GHz channel shown is from MWI.

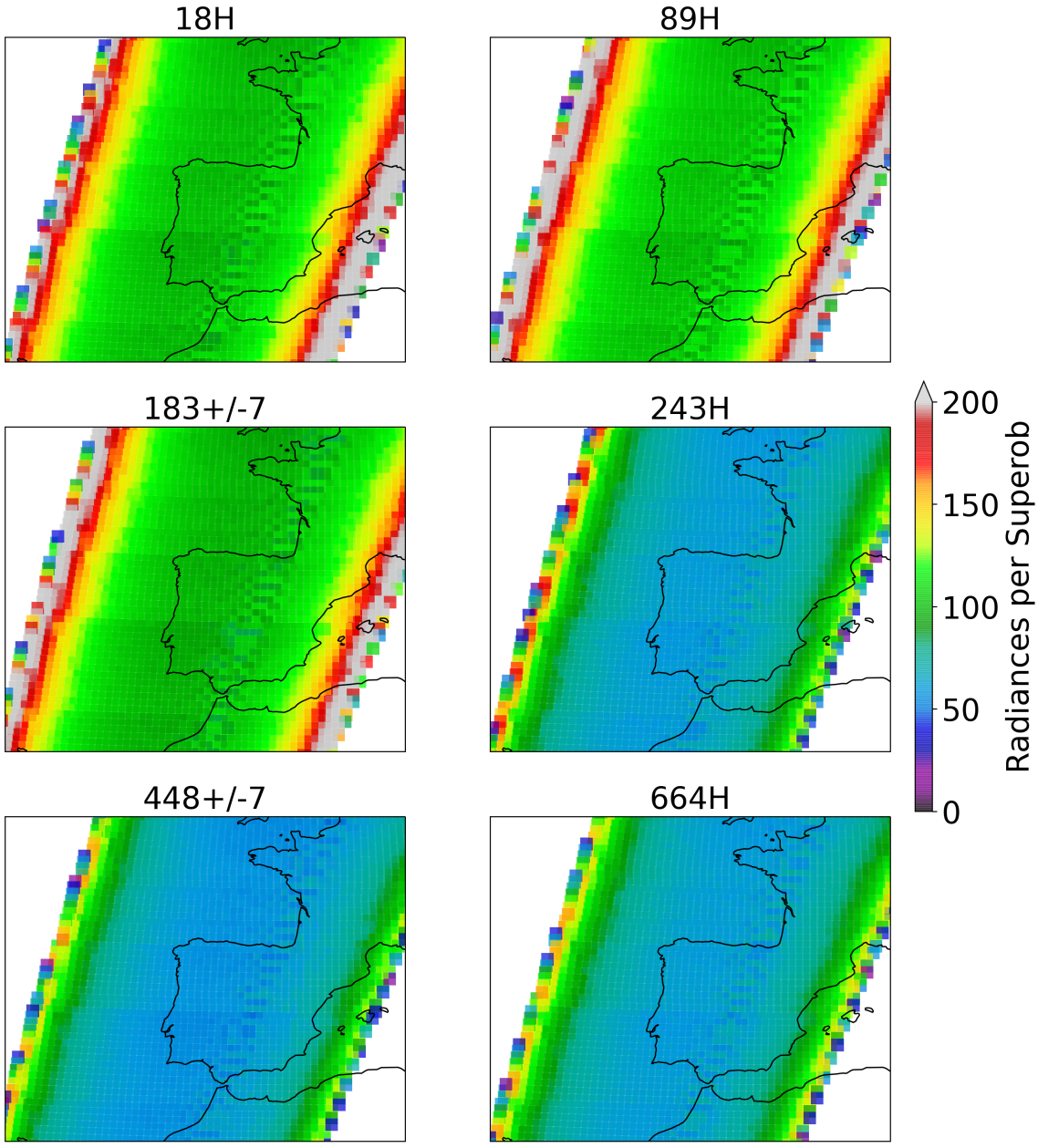


Figure 11: As Fig. 10, but showing the number of observations per superob.

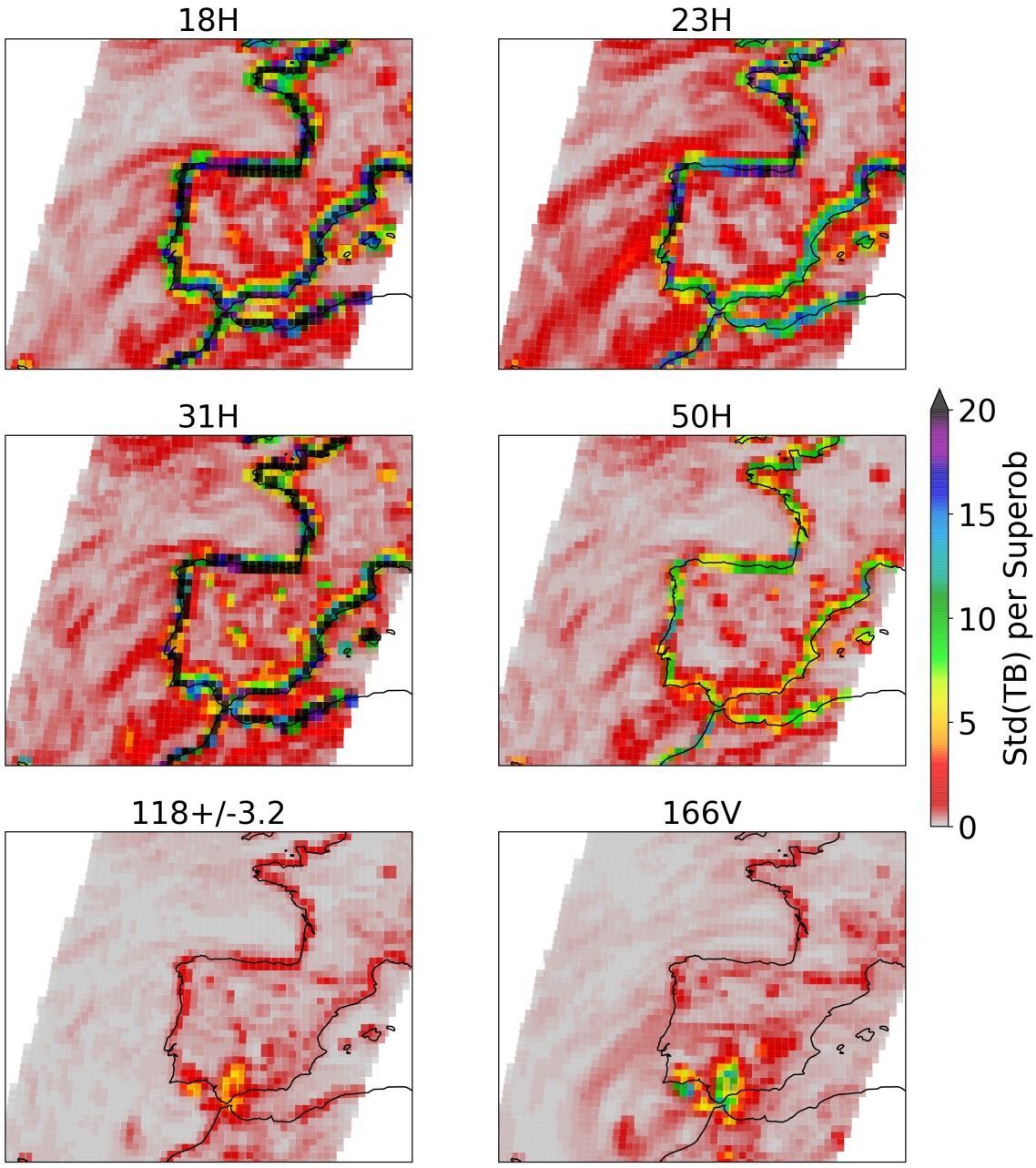


Figure 12: Standard deviation of observed TBs per superob for six selected MWI channels over Spain from test orbit 4656.

indicate more scattering in the test data than the model background.

To provide a global picture, we can examine most of two orbits together, 4655 and 4656, from the 12Z long window on Sept. 12 2007. Figure 14 shows two channels, 89H from MWI and 243H from ICI. Orbit 4655 begins at 08:43Z, before the 12Z long-window cutoff of 09:00Z, so part of this orbit exists in the 0Z window. As with the zoomed-in view above, the global maps appear quite reasonable in most respects. The 89H channel shows a good mixture of positive and negative departures as we would expect, with some systematic deviations over difficult terrain (e.g. the Kamchatka coastline and the sea-ice edge around Antarctica) and 10-20 K signals of cloud displacement in the background. The 243H departures are primarily negative in areas of cloud and precipitation, again showing more scattering present in the simulated observations than the model background, often as much as 40 K or more. The two channels shown are relatively representative of the window and sounding channels' behaviour, though 243H is a channel with mixed sensitivity, predominantly sounding in the tropics but surface-sensitive at high latitudes.

Figure 15 illustrates the geophysical properties of a region over the northwest Pacific and east Asia that will be used to explore the impact of screening criteria. There is a strong frontal structure east of Japan as well as smaller-scale convective activity over northeast China, as illustrated by the model snow, cloud water, water vapour and wind at observation locations. Comparing to the background departures for channel 27 on the same plot, it is clear that the largest O-B departures are associated with the heavy clouds and precipitation of the frontal system and convection over land. Although a primary explanation for such large O-Bs is the displacement of cloud and precipitation features between the observations and the background, the more systematic error component in these departures is a result of different scientific assumptions in the representation of cloud and precipitation between the IFS and in the creation of the simulated MWIICI data. This will be discussed further in Sec. 5.3. It is clear that these areas cannot be included as part of the cal/val sample.

The stringent, dynamic and unified screening criteria used for selecting data for cal/val will be introduced properly in Sec. 6. However, we can make an initial examination of the proposed screening criteria in terms of the departures that remain in the sample for cal/val purposes. Figure 16 uses the same scene as the previous figure but displays only departures that meet the stringent screening criteria as laid out in Sec. 6. Comparison of this figure with the previous one shows that most of the cloudy areas have been screened out as intended, whilst the other observations that remain in the sample differ from channel to channel depending on their sensitivity. From this selection of channels, 89H is the one with the least data remaining after the screening. 89H displays an overall negative bias in this scene, likely a feature of emissivity bias between TESSEM2, which was used in generating the EUMETSAT test data and SURFEM (Geer *et al.*, 2024), which is used in the IFS (further information in Sec. 5.3). The *CI* criteria that is part of the stringent screening and which will be introduced later is in this case removing data points due to the bias between surface emissivity models rather than the cloud impact.

The fraction of MWIICI test data from the global sample that is retained by the screening procedure can be seen in Fig. 17, alongside the mean departures for the different data samples. This figure provides some sense for how the screening criteria are performing for each channel. Many channels show a near-zero bias in the screened data as we would expect, with the screening appearing to work quite well for the higher frequency channels especially. Overall the screening retains about one third of the total observations, with mid-frequency sounding channels having the most observations left after screening. Window channels show the smallest fraction of observations passing the screening, as one might expect. For analysis of how well the selection criteria are performing, see Sec. 6.3.

A final example that hints at the possible performance of temperature sounding from MWI is given in

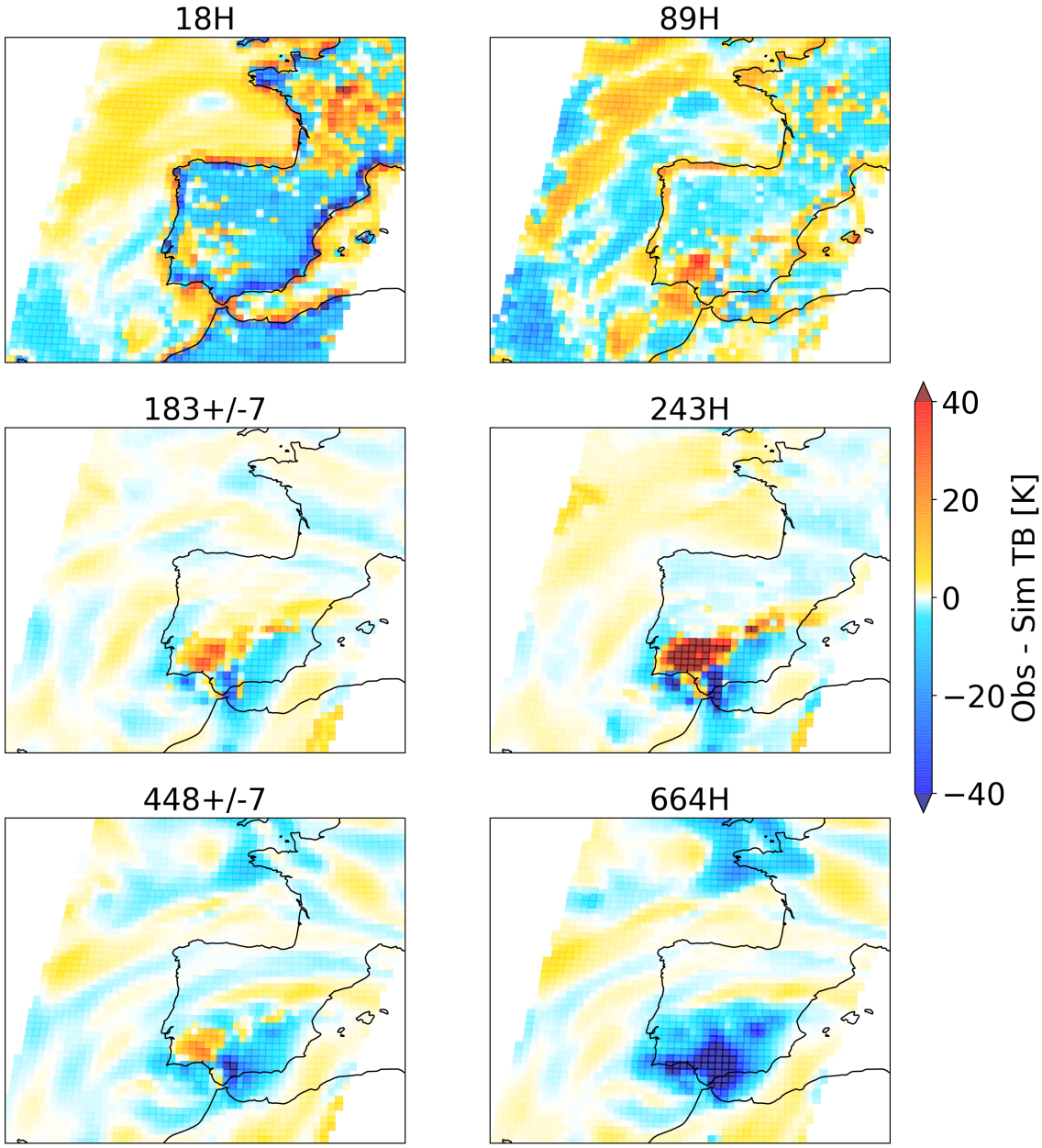


Figure 13: Background departures (O-B) for six selected MWI channels over Spain from test orbit 4656.

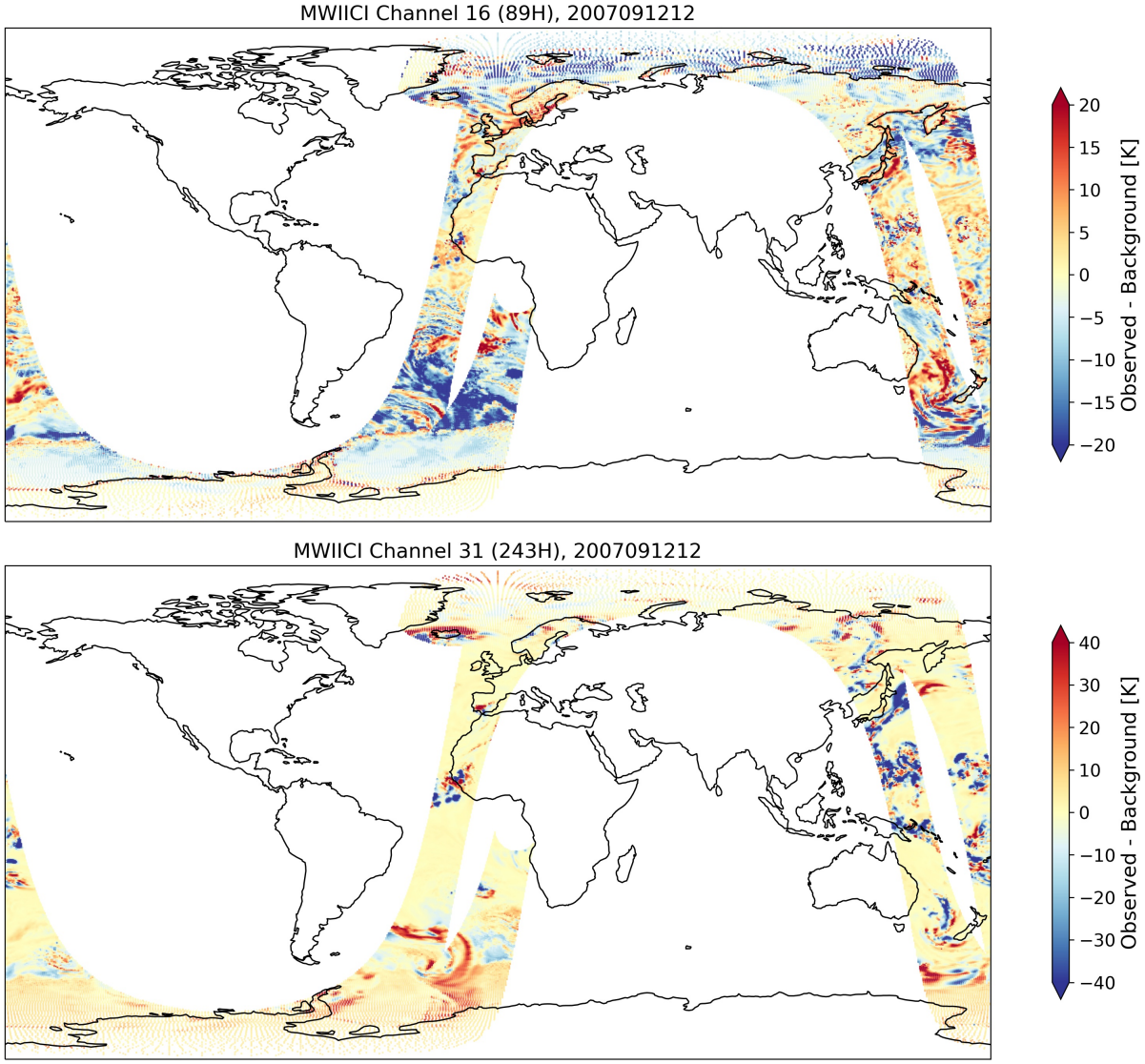


Figure 14: Background departures (O-B) from MWIICI channels 16 (89H) and 31 (243H) in the 12Z long window, 12th September 2007.

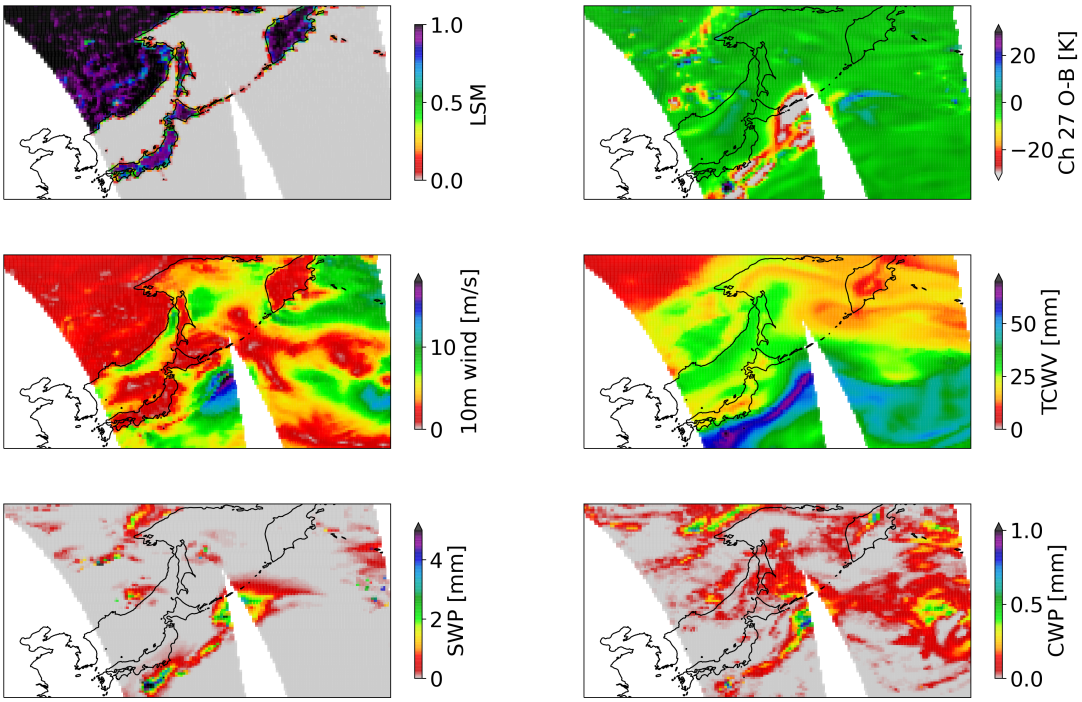


Figure 15: IFS output from test orbits 4655 and 4656 over the northwest Pacific ocean and east Asia. Channel 27 (ICI 183±7) departure is shown (top right) along with 5 ancillary fields: LSM (land-sea mask from channel 1), 10m wind speed, TCWV (total columnar water vapour), SWP (snow water path), and CWP (cloud water path)

.

Fig. 18, where MWIICI channel 13 $std(O-B)$ is shown for the two sample orbits in the 12Z long window. This particular plot shows how well the model can fit sounder channels in the 50 GHz oxygen band, with most grid boxes showing $std(O-B) < 0.2K$ and only areas of deep convection causing larger variability of departures, despite no data selection having been applied. If the data selection is limited to the unified criteria, most grid box averages exhibit std(O-B) values consistently less than 0.10 K (not shown). Despite the test data containing simulated sensor noise, the std(O-B) values are almost certainly lower than we should expect for real data, as the added noise was random white noise with no "striping" noise or other effects that can be seen for radiometers. However, the low std(O-B) values shown here point to how much the superobbing procedure decreases the sensor noise that tends to dominate std(O-B) for such frequencies.

5.2.2 Surface and vertical sensitivities of MWIICI channels

It is also useful to consider the surface sensitivity but also the variability of various sounding channels on MWI and ICI, as their atmospheric and surface sensitivities can vary considerably depending on atmospheric conditions. Figure 19 shows the clear-sky surface-to-space transmittance for a section of the test data covering most of the Atlantic Ocean with parts of both test orbits, with a selection of sounding channels from MWIICI given in the various panels. The window channels at 166 and 243 GHz have the greatest variability in surface sensitivity, with near-zero sensitivity in the deep tropics but $\tau > 0.5$ at high latitudes. The 50 GHz sounding channels exhibit much more uniform sensitivity through the orbits, with a similar story for the higher-peaking 118 GHz channels. The sub-mm channels only exhibit significant surface sensitivity at high latitudes, over high orography, and for the lowest-peaking 325 GHz channel.

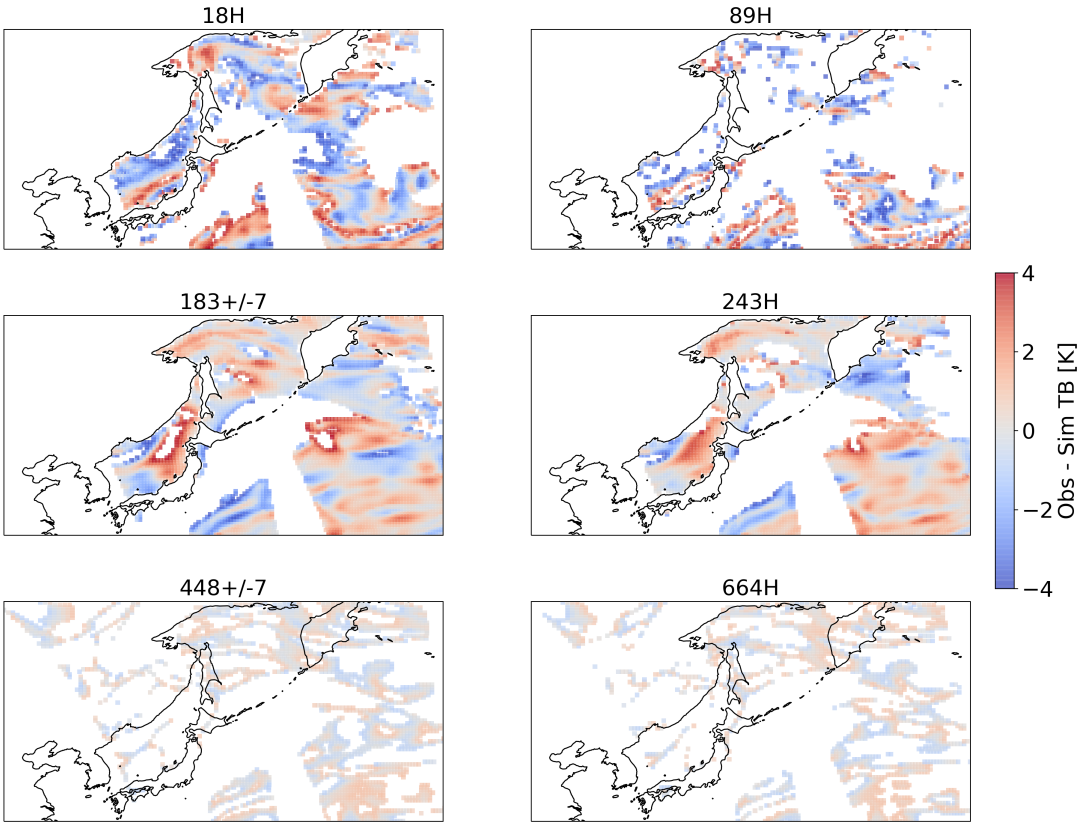


Figure 16: As Fig. 15 but for background departures that meet the stringent screening criteria.

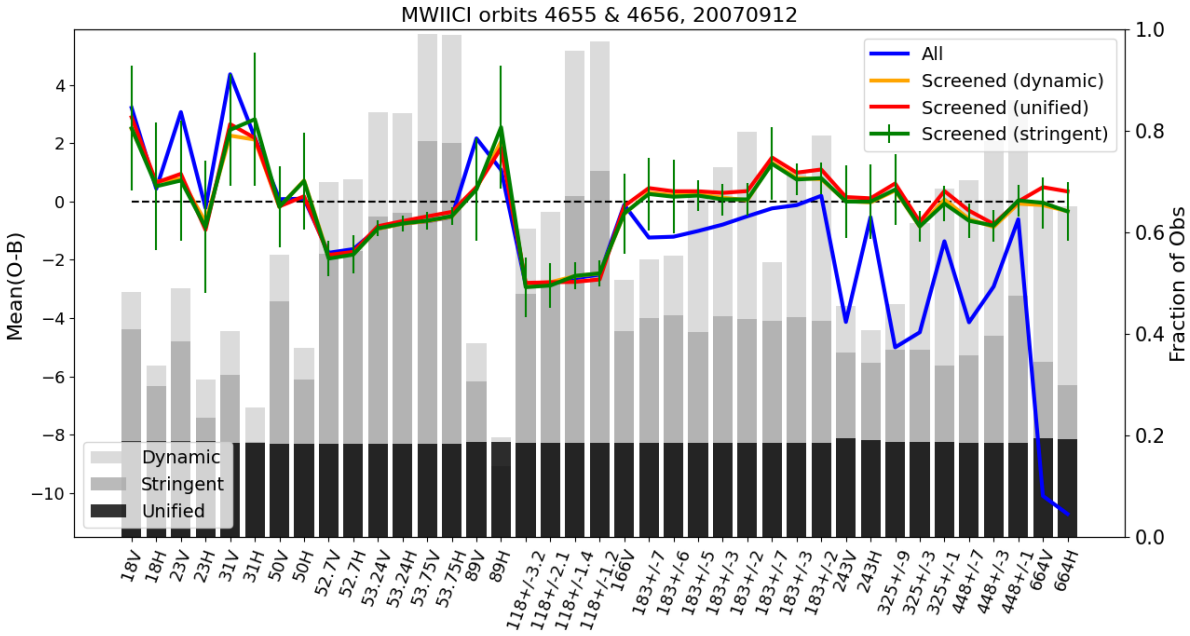


Figure 17: The mean departure per channel for all data (blue) and the screened samples using the dynamic (orange), stringent (green), and unified (red) criteria, given for all MWIICI channels in the 12Z long-window. The fraction of total observations in the sample is given in horizontal bars, with light grey for dynamic, dark grey for stringent, and black for unified. Vertical bars for std(O-B) per channel are given for the stringent data points.

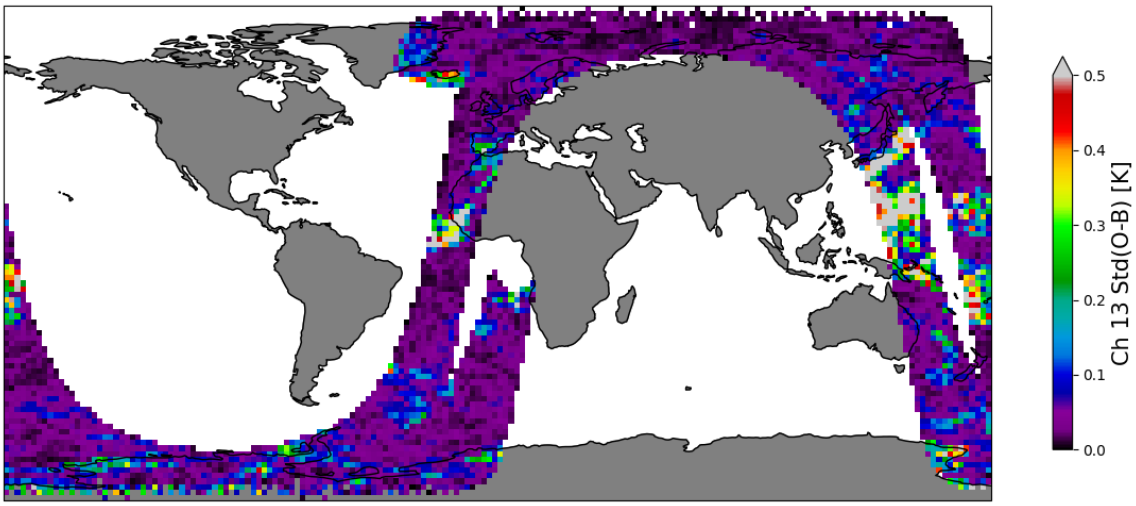


Figure 18: Gridded global statistics for the MWIICI test data, here showing channel 13 (53.75V) std$(O-B)$ for the two orbits available in the 12Z long window. Values shown are in Kelvin, with gridding at 2 degrees.

In the same vein, Fig. 20 shows the same set of channels but shows the peak of the weighting function, indicating the pressure level of greatest TB sensitivity for the MWIICI sounding channels. The weighting function peak is defined considering all-sky radiative transfer, and so the peak sensitivity may be due to cloud signals or gaseous emission signals. For example, the 118 ± 1.2 and 118 ± 1.4 GHz channels predominantly sense the upper troposphere with a weighting function peak around 300 hPa, but they also have strong cloud sensitivity that picks up mid- and lower-tropospheric clouds. In contrast, the 50 GHz channels have more consistent peak sensitivities due to weaker cloud interactions that are mostly restricted to liquid cloud in the lower troposphere. The sub-mm channels display a range of vertical sensitivities, with weighting functions peaks of about 200-400 hPa in the deep tropics all the way down to lower tropospheric sensitivity in dry air ahead of a frontal boundary in the Southern Ocean. This set of plots hints at the profiling capability that is expected from ICI (Eriksson *et al.*, 2020).

5.3 Scientific differences in forward modelling

Forward modelling of the test data (i.e. synthetic observations) was performed with different assumptions than those of the IFS background. In this section these are spelled out to aid discussion of the differences in bias between the test data observations and the model background. Here we set aside the (expected) minor differences between an IFS background at Cycle 49r1 (initialised from the operational analysis in 2007) and the ERA5 analysis used by Rydberg and Eriksson (2019) to create the test data.

One likely area of forward model biases between the test data and the IFS (c.f. Fig. 17) is seen in the 50 and 118 GHz sounding complexes. Perhaps the most probable cause of these biases is the different spectroscopy used between RTTOV and ARTS, namely AMSUTRAN versus the Rosenkranz absorption model for atmospheric oxygen (Patrick Eriksson, pers. comm.). However, SRF differences in the assumptions of the forward models could also play a role, especially as the test data simulations make a monochromatic assumption (i.e. ignoring sidebands) that could cause up to 1.8K bias in clear scenes for some channels (Rydberg and Eriksson, 2019). In contrast, the IFS simulations use RTTOV coefficients that are based upon the lab-measured SRFs for ICI and for MWI the coefficients are based on top hat functions representing the current SRF specifications, including all sidebands that compose the channel

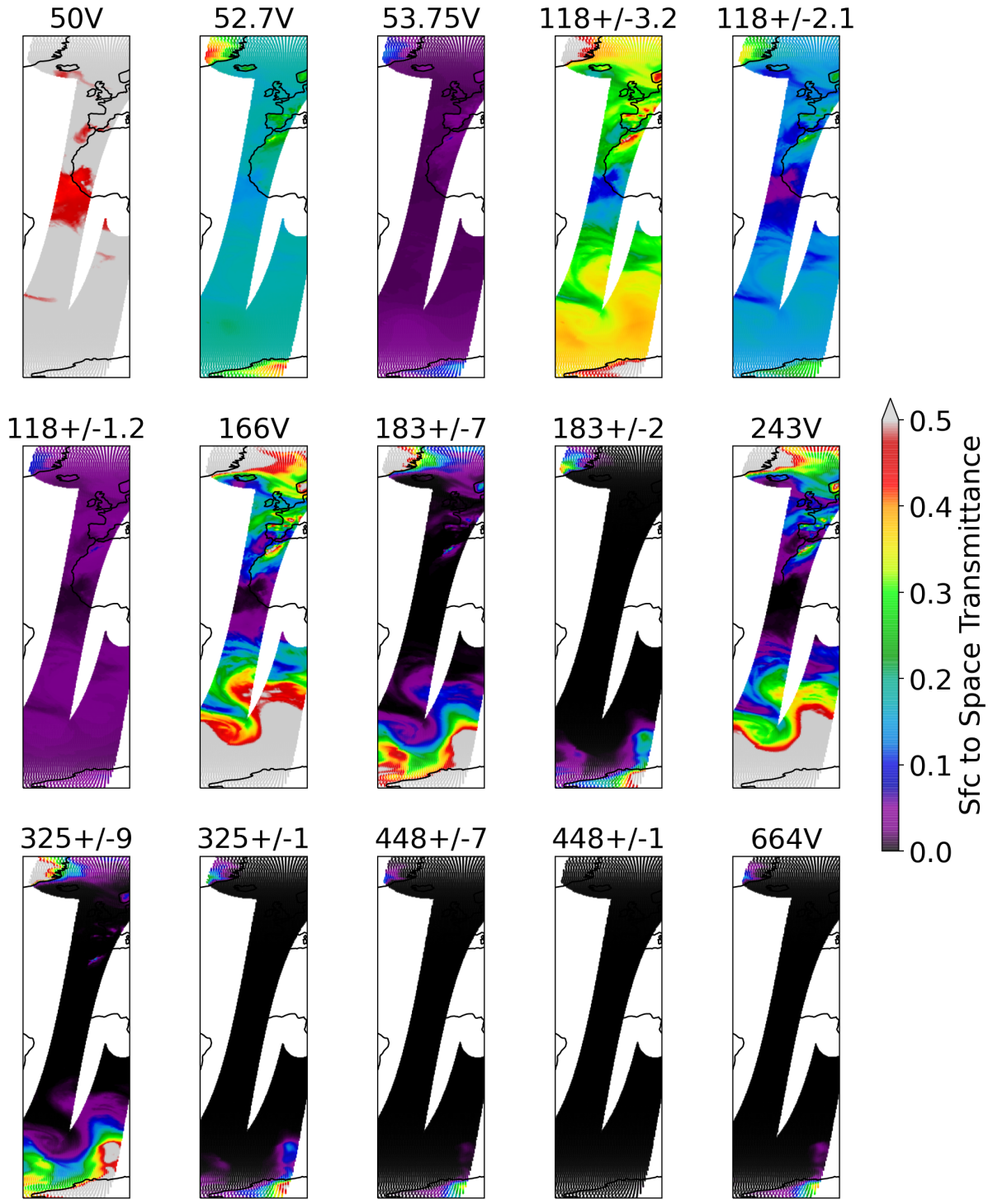


Figure 19: The clear-sky surface to space transmittance (τ) for a selection of MWI and ICI channels, as seen by RTTOV-SCATT for part of the test data orbits.

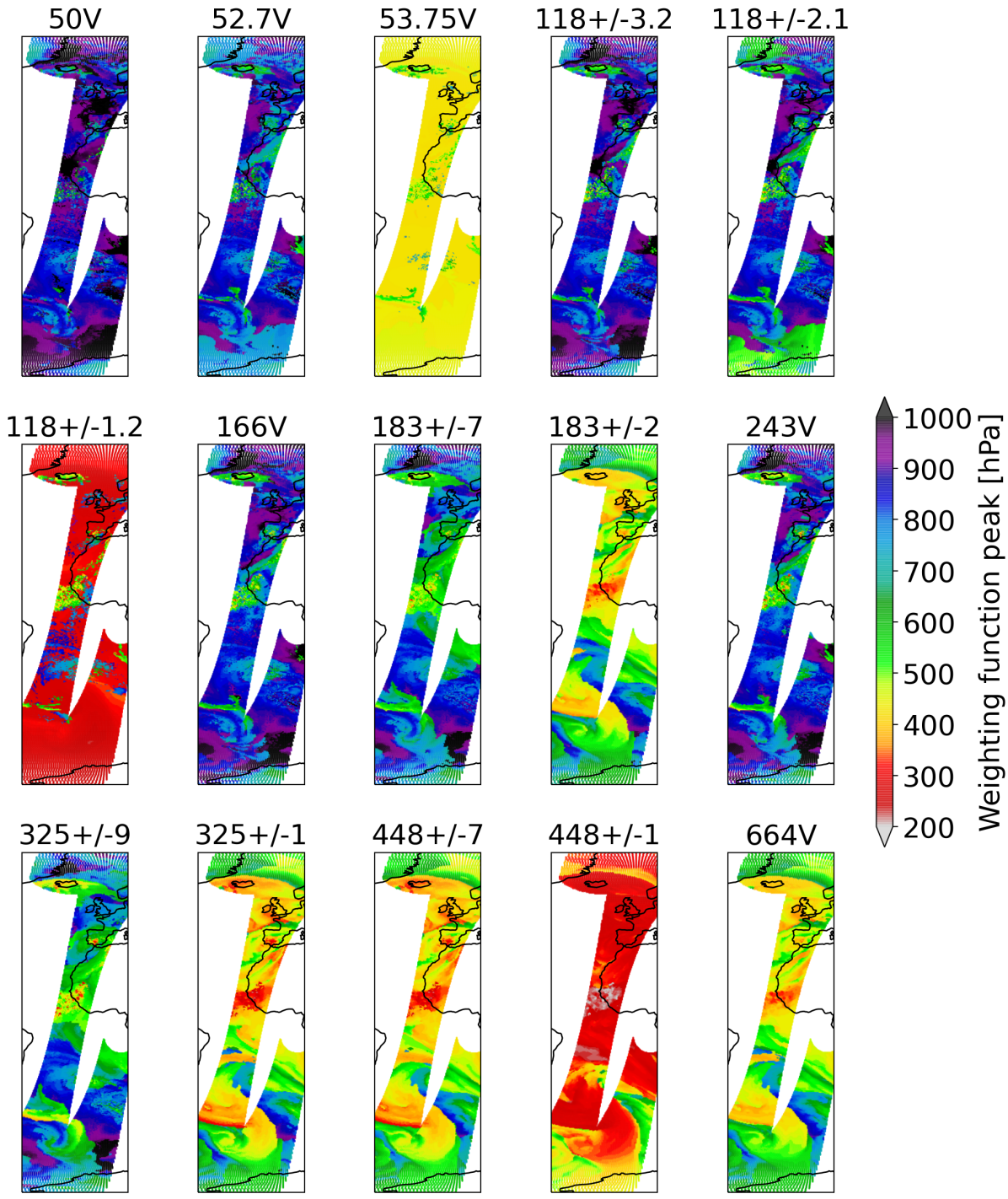


Figure 20: The peak of the weighting functions for a selection of MWI and ICI channels, as seen by RTTOV-SCATT for part of the test data orbits.

(the MWI coefficients will be further updated to reflect the measured SRFs once available).

The sea surface emissivity in the test data was modelled with TESSEM2, whereas the IFS now uses SURFEM. As TESSEM2 was based upon FASTEM and Geer *et al.* (2024) analysed the impact of changing from FASTEM to SURFEM on all-sky radiances, that study can be instructive in this case although it is not a direct comparison. They found that while the switch to SURFEM generally improved biases for low frequency channels and especially for high wind speeds, the mean bias for mid-frequency microwave channels did appear to be worse over a range of wind speeds. Biases changed by up to around 2 K in some channels, notably 37H, 89H and 150H.

For the sub-mm channels below 664 GHz, the main difference between IFS and the test data appears to lie in scattering signals from areas of cloud and precipitation. There are several key assumptions made in the forward modelling of scattering processes that could contribute:

- Convective mass concentrations
- Particle size distribution (PSD) and particle shape
- Scattering solver

As the test data used ERA5 for atmospheric data and as of yet the ECMWF reanalyses do not archive convective mass profiles for rain and snow, these have been recreated using a mapping between surface convection precipitation rate and the database of ECMWF convective mass flux profiles. The conversion from flux to water content uses a 1 m/s fall speed, whereas the IFS uses a slightly more complex approach (size-dependent fall speeds) and the implied fall speed is closer to 1.5 - 2 m/s. That would suggest the test data is be based on higher convective mass than the IFS simulations.

Regarding microphysical assumptions, there is a single ice hydrometeor category used to generate the test data (cloud ice + large-scale snow + convective snow). This is modelled in ARTS using the same microphysical assumptions that the IFS uses for large-scale snow, i.e. Field *et al.* (2007) PSD with the ARTS large plate aggregate (LPA). In contrast, the IFS models the cloud ice category separately using a PSD that implies much smaller particles and with a particle shape (ARTS large column aggregate) that generates very little scattering. These choices were tuned against SSMIS 183 GHz observations, which show that cloud ice, e.g. cirrus, is nearly invisible at these frequencies. This is likely a main explanation for the differences in the higher ICI channels. For convective snow, the IFS models this separately as well, using Field *et al.* (2007) PSD plus a particle shape giving slightly more scattering compared to the ARTS LPA (instead the ARTS column). While that may seem to suggest the IFS should provide more convective scattering, in the IFS the convective and large-scale ice mass are put in two separate PSDs, whereas in the test data the mass is summed into one PSD. When the mass is combined into one PSD, this will generate larger particle sizes compared to splitting into two PSDs, so this effect likely also contributes to much deeper scattering signatures in the test data.

Lastly, ARTS and RTTOV-SCATT utilise different scattering solvers. It is known that the delta-Eddington solver in RTTOV-SCATT sometimes generates insufficient scattering in high optical depths, particularly for high forward scattering, high single scattering albedo situations, compared to ARTS, which uses a more sophisticated discrete ordinates solver. This has been studied in detail (Barlakas *et al.*, 2022, see their Fig. 5), and could be a contributing factor for discrepancies in the scattering magnitudes. This could be a key driver of the discrepancies seen at 664 GHz. However, the performance at 664 GHz in particular appears anomalous, with large and systematic differences between RTTOV and the test data; at present it is not known why, but it appears that RTTOV is underestimating the scattering expected at 664 GHz (e.g. relative to Geer *et al.* (2021)).

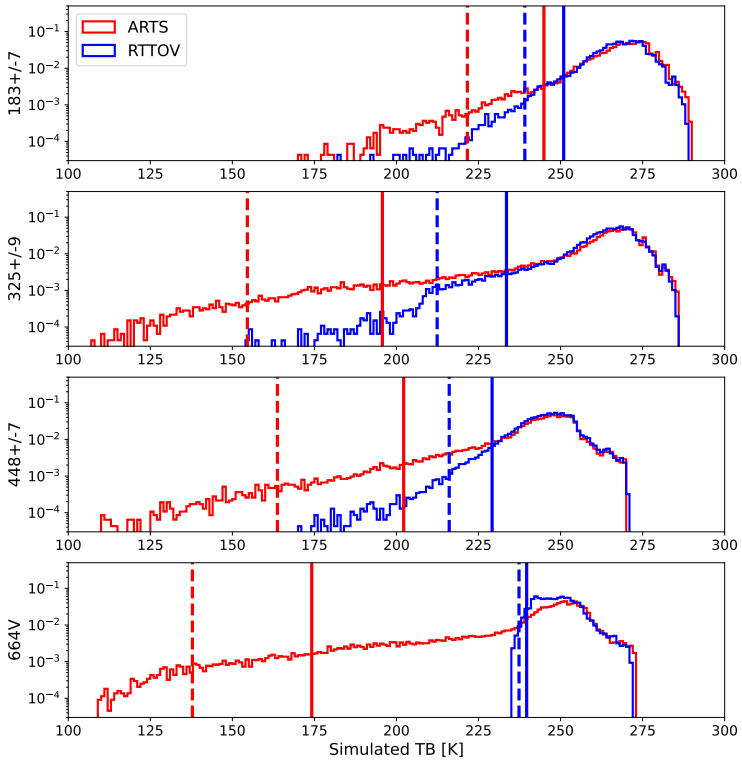


Figure 21: PDFs of simulated TBs from ARTS (i.e. test data) and RTTOV-SCATT (i.e. IFS) for all ice-free observations over sea from the two test orbits, shown for MWIICI channels 27 (183±7), 32 (325±9), 35 (448±7), and 38 (664V). Vertical lines indicate the 5% (solid) and 1% (dashed) quantiles of the PDFs, the scattering quantiles.

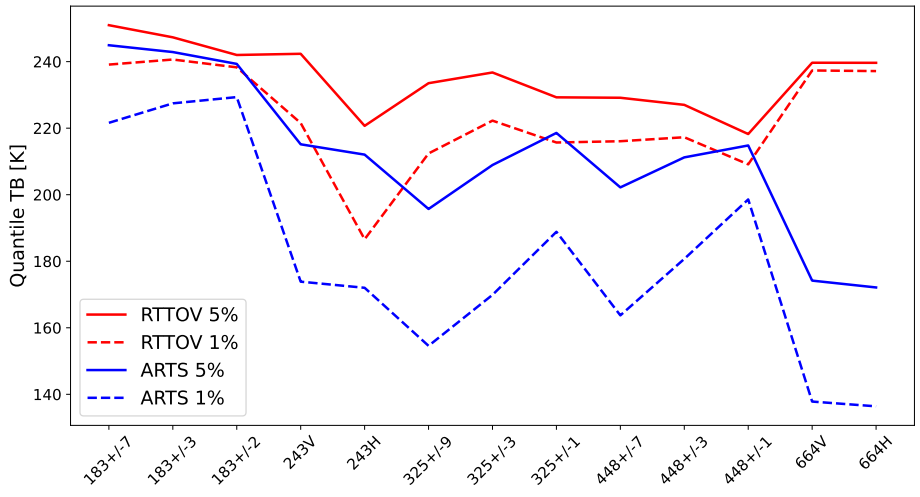


Figure 22: Scattering quantiles of 5% and 1% for all ice-free ocean observations in the two orbits of test data, as visualised in Fig. 21. All ICI channels are shown (MWIICI channels 27-39).

To quantify the scattering differences discussed, here we introduce "scattering quantiles" from PDFs of all simulated TBs from ice-free ocean scenes from both ARTS (the EUMETSAT test data simulations) and RTTOV-SCATT (the IFS). The scattering quantile is defined as the first X% of the PDF of TBs, so the 1% scattering quantile is the TB representing the dividing line between the lowest 1% and highest 99% of TBs. This is visualised in Fig. 21 for four ICI channels.

Figure 22 provides the scattering quantiles as a function of frequency, showing all the channels from the ICI instrument. This indicates that whilst the ARTS simulations are systematically colder (i.e. showing greater scattering), the difference is more pronounced at we go up in frequency and the PDF is also much wider. The 1% scattering quantile is roughly 10-15 K lower for ARTS at 183 GHz channels, but this gap generally grows with frequency, with a 60 K gap at 325±9 and about 100 K at 664 GHz. This type of analysis could prove useful to gauge the realism of scattering simulations once observational sub-mm data first become available. Although RTTOV is simulating far less scattering than ARTS, we believe this is due to the different scattering-related assumptions being made, as explained above. However there may be a real issue with the quality of the RTTOV 664 GHz simulations in cloudy areas, given the almost total lack of a scattering signature compared to ARTS at this frequency (Fig. 21), which will require further scientific attention.

The differences noted here highlight some of the key remaining uncertainties in radiative transfer calculations for sub-mm frequencies. A great deal of work has gone into reducing radiative transfer uncertainties to prepare for ICI (e.g. Geer *et al.*, 2021; Turner *et al.*, 2022), and this first look at ICI channels with RTTOV-SCATT has proven to be valuable practical experience. Some aspects such as the sensitivity to microphysical and spectroscopic assumptions may require real data for further progress to be made.

5.4 Addressing scan-dependent biases without superobbing

Although superobbing has numerous benefits for all-sky assimilation, from a calibration/validation perspective there is a concern that the averaging process may hide subtle or sharp bias patterns that could be present in the un-averaged radiances. The solution to this problem appears simple—run the un-averaged radiances through the IFS. From the perspective of the flow chart in Fig. 6, we could simply skip the superobbing step and the horn matching step (as no common geolocations exist), although as will be explained below, this is a substantial technical challenge. The specific target of the no superobbing framework in this work is to validate requirements on inter-footprint radiometric bias, specifically its variation with scan position. This in particular could exhibit very sharp jumps and features that could be smoothed out by superobbing over multiple scan positions.

Passing MWIICI data through the IFS without any superobbing applied is a distinct technical challenge. For example, the act of superobbing decreases the number of MWI radiances by about two orders of magnitude, and thus instead of forward modelling say 10 million radiances in one long window DA cycle, we have 1 billion. Processing a billion observations is not typically a problem for the IFS, as the observation processing architecture has been built to handle many billions of observations. The real challenge here is that RTTOV and the file formats of BUFR and ODB are built to run fast and compress efficiently for multiple observations at a single geolocation, such as for a hyperspectral sensor with thousands of channels at a single geolocation. In contrast, the raw/un-averaged MWIICI data provide 15 different geolocations across 39 channels.

It is worth emphasising the scope of this problem. A single orbit of MWIICI data with normal superobbing applied (i.e. averaged to 40 km superobs) yields about 1.8m radiances to the IFS at 60k geolocations; with no superobbing applied, a single orbit has about 210m radiances at 66m geolocations. Even

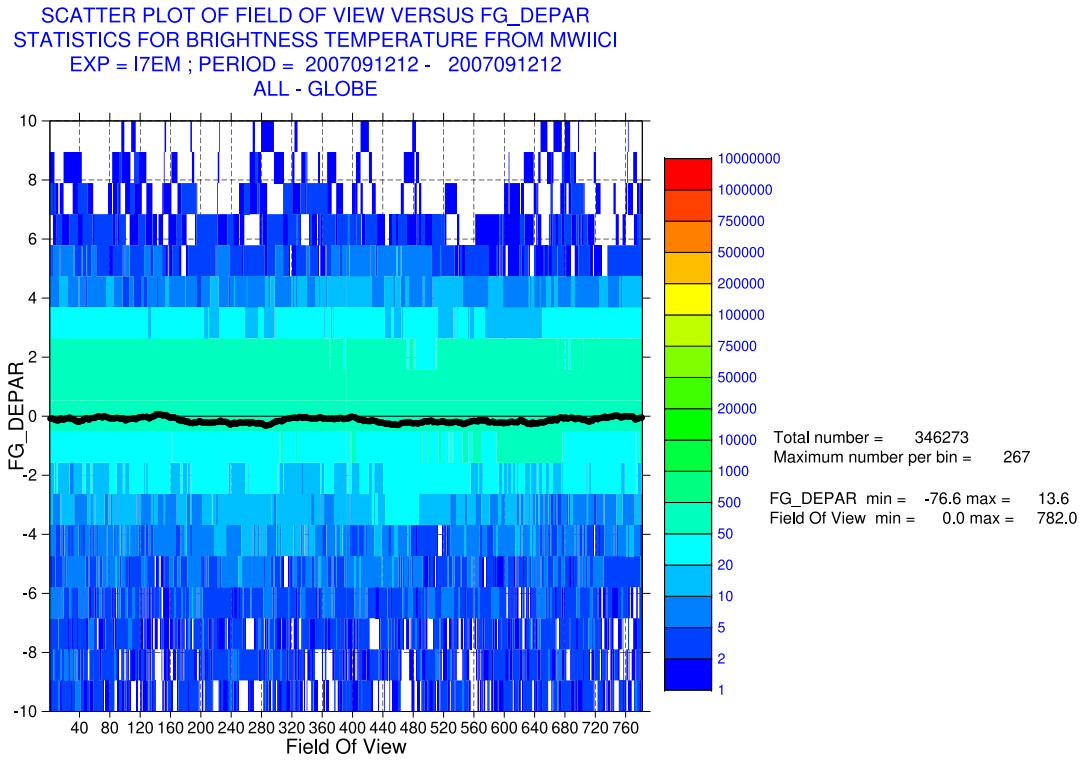


Figure 23: For the No Superob experiment, the histogram of sample vs. O-B ("field of view" and "FG − DEPAR" in the IFS) is given in the colours. Mean O-B per sample is shown in the black line. This constitutes all available data in the 12Z window from channel 37 (448±1.4 GHz).

with modifications to bufr2odb to shrink data volumes as much as possible, one full orbit of un-averaged MWIICI data yielded a 22GB ODB going into the IFS. For comparison, all orbits from 12 hours of IASI data on one Metop platform is about 5GB after pre-screening (albeit for only 420 of its channels, and one of four FOVs). A full 12hr assimilation window with un-averaged MWIICI data would yield over a billion radiances at hundreds of millions of separate geolocations; in contrast, IASI has about 160k geolocations, and thus the data compress more easily. Needless to say, these data volumes cause stress on numerous aspects of the IFS data processing.

Due to these technical challenges, it is necessary to perform some scan line thinning to provide un-averaged departure statistics at all samples across the scan. Thus for this test, 1 in 20 scan lines were retained by modifying the pre-processing script. An example of scan position biases for MWIICI test data can be seen in Fig. 23, given for channel 37. As this shows the simulated observations, there are of course no "real" issues in the sensor performance as a function of scan position. The less cloud-affected channels such as 13 or 20 show a nearly perfect straight line even for all data, whereas the window and sub-mm channels exhibit greater noise because of the limited sample from two orbits. This type of analysis should benefit significantly from having multiple days of data to reduce the random noise caused by cloud effects and sampling variability. This particular figure shows all available data, but the screened radiances would be used to provide a cleaner sample. Note that per-channel screening is available for the un-superobbed radiances, but the unified screening is not due to its reliance on other channels being co-located.

Mission requirements MWI-05120 and ICI-06120 can be assessed using the output from this special experiment, concerning inter-footprint radiometric bias differences. In addition, 2D histograms and

Hovmöller-style plots of bias as a function of sample number are available on the monitoring website as part of the routine monitoring, i.e. for superobbed radiances in near-real time.

6 Data selections tailored to cal/val

In this section on the scientific method of the study, the data selection rationale for cal/val analysis is described.

The population of radiance data used for calibration and validation activities is a crucial aspect of using NWP fields for vicarious monitoring of radiometric performance. As Fig. 24 demonstrates, the frequencies observed from MWI have quite variable surface to space transmittance, as assessed by IFS fields in conjunction with RTTOV-SCATT; this figure uses a nearby SSMIS channel (where available) to work as a proxy for selected MWI channels. The figure gives a sense of which channels have strong surface sensitivity and would thus benefit from strict screening to remove difficult surfaces from the data selection, for example.

Further, Fig. 25 provides a measure of how much impact cloud emission and scattering have on the radiative transfer of the same MWI channels. This shows the all-sky TB effect, i.e. the net effect of including emission and scattering from hydrometeors, (the difference of all-sky and clear-sky RTTOV simulations of the model background, $B_{Ch} - B_{Ch,clr}$, the second half of CI as given in Eq. 8) again using two days of data from SSMIS as an MWI proxy. This series of plots indicates some valuable aspects about the data selection. First, most MWI channels have cloud sensitivity that is non-negligible over a large part of the globe; cloud signals on the order of 1-5 K are pervasive even for a high-peaking channel such as MWI-26. Second, mid-latitude oceans typically exhibit cloud signals ranging from zero up to tens of Kelvin. Purely cloud-free areas of the globe are relatively rare at many of these frequencies, which is why analysis of samples with relatively low cloud impacts (but not strictly "cloud-free") is a key part of the data selection strategy.

To aid with identifying a reasonable selection for purposes of assessing instrument calibration, it is helpful to define a cloud impact metric, CI_{Ch}, per channel and for every observation (and accounting for bias via $BiasCorr$, as will be discussed below). This uses the observation and the model equivalent from the background interpolated to the observation location, both considering all-sky conditions (i.e. with clouds and precipitation, B) and without (B_{clr}):

$$CI_{Ch} = |(O_{Ch} - BiasCorr_{Ch}) - B_{Ch,clr}|/2 + |B_{Ch} - B_{Ch,clr}|/2 \tag{8}$$

This definition follows the symmetric cloud predictor concept in some sense, equally weighting cloud in the observation and the model, effectively yielding information on whether hydrometeors are present in either the observation or the model background, and increasing with increasing cloud impact. CI should thus be a reliable metric for ascertaining the suitability of an observed radiance for analysis, ensuring the observation has cloud impacts below a certain threshold. The CI definition follows the cloud effect average of Okamoto *et al.* (2014) and the predictor for window channel assimilation of AMSU-A used by Duncan *et al.* (2022a) among others, so this is not unique to this study. Crucially, CI is channel-dependent by definition and thus it is sensitive to the true radiative impact of cloud for each sensor frequency. This is a significant advantage over metrics such as $P37$ or SI that may be applicable to some channels' sensitivities but not all. For this reason CI and surface to space transmittance τ constitute the main selection criteria, and $P37$ and SI, although initially considered for use in the cal/val data selection, were ultimately abandoned in favour of CI.

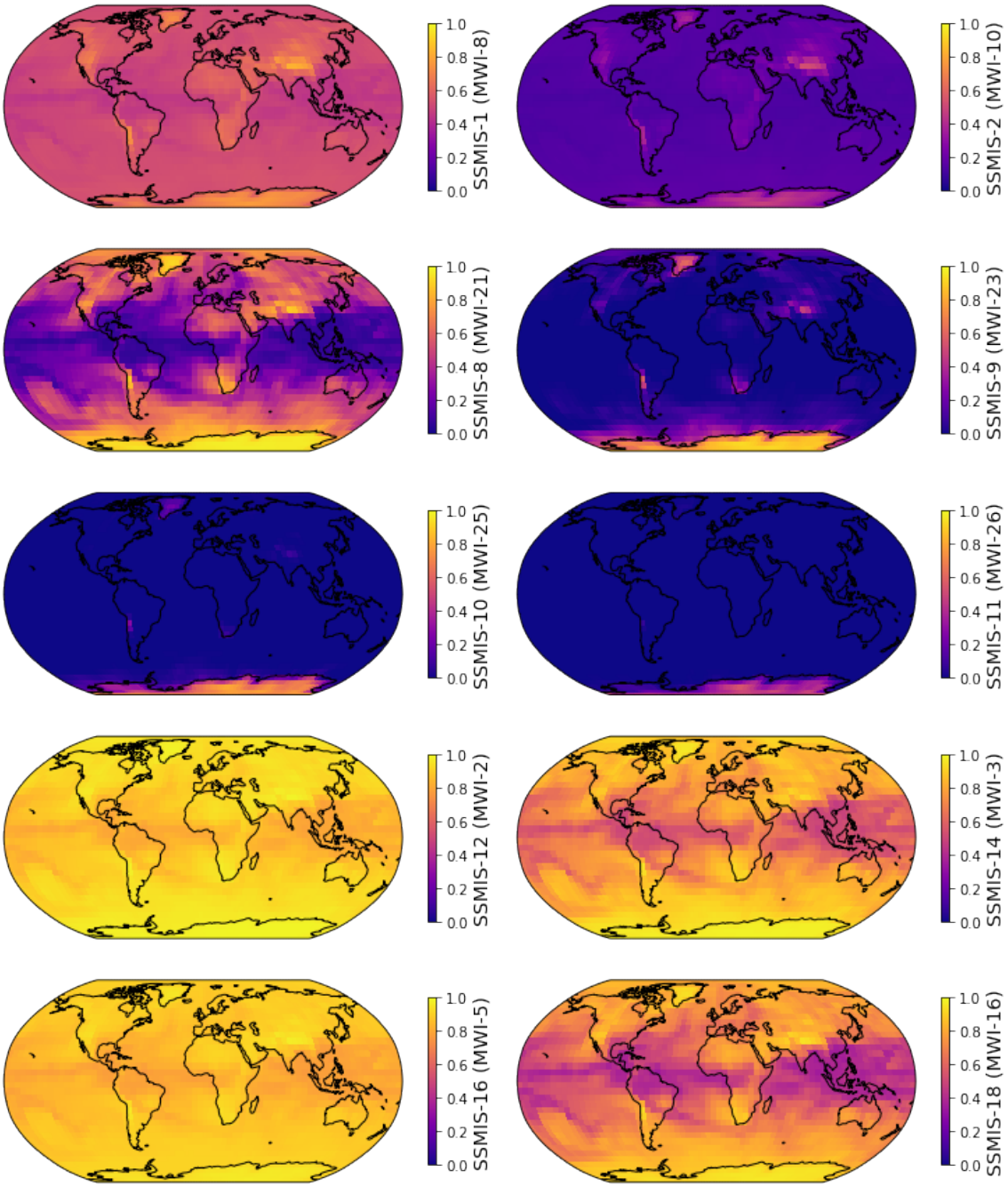


Figure 24: Mean surface to space transmittance (i.e. τ) as seen by SSMIS F17 channels that approximate several of those on MWI. Data taken from June 2-3, 2022 with no pre-selection of all-sky radiances and averaged on a 5-degree grid. This is unit-less, with 1.0 indicating 100% sensitivity to the surface. Note that this is clear-sky τ, so attenuation from clouds and precipitation is not considered here.

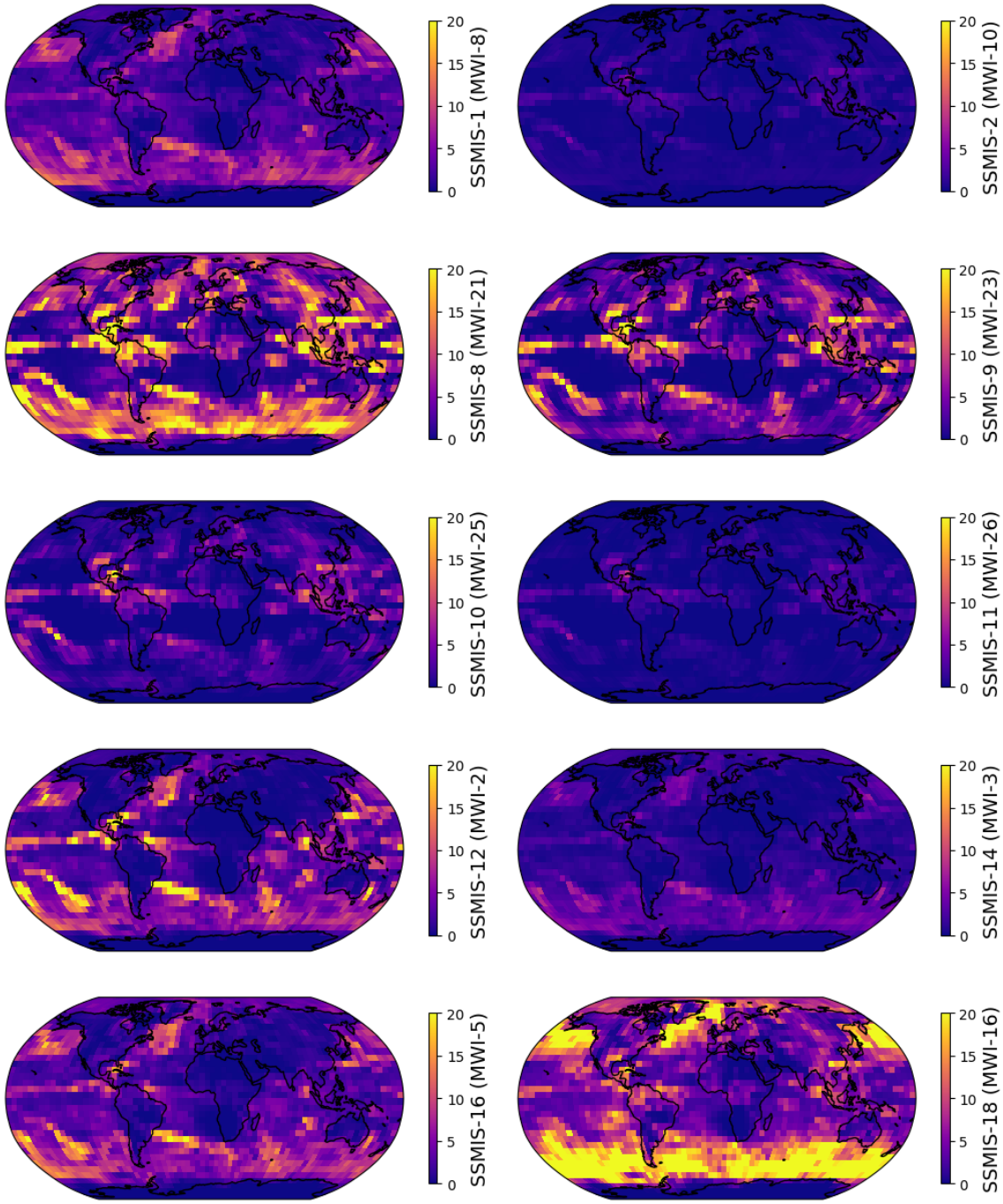


Figure 25: RMS of the all-sky impact on TB (in Kelvin), i.e. the radiative impact of cloud and precipitation from RTTOV-SCATT. Data selection is the same as Fig. 24.

A key aspect of the *CI* definition is that it relies upon differences between B, B_{clr}, and O. If any of these terms has a systematic bias against the others, this poses a problem for the method. For example, some AMSR2 channels exhibit a -5 K global bias, and if we use a *CI* threshold of 2 K to define our sample for cal/val, most of the actual cloud-free data from these channels will be discarded. Ideally the observed and modelled radiances would be wholly unbiased so that *CI* can function as an effective metric for gauging cloud impact. For pragmatic reasons, in this project the forward model is assumed to be unbiased. The IFS solves for the bias correction that needs to be applied to the observations to remove most suspected causes of bias between the model and the observations, some of which are sensor-related and some of which are forward model biases. The bias-corrected observations are therefore used in the calculation of *CI*, with the bias correction *BiasCorr* a function of channel and subtracted from the observed TB as shown above.

Note that the following subsections cover the stringent, dynamic, and unified data selections, of which the first two are channel-specific, and the last is applied over all channels. The monitoring website also provide selections "All" for all available data, i.e. no selection criteria are applied, and "Used", which signifies assimilated data. The used data selection is not meaningful in the case of MWIICI, but can be useful for other sensors as this selection includes other forms of quality control that come from the data assimilation process such as variational quality control.

6.1 Channel-specific data selection

The channel-specific data selection criteria are found in Table 8, with MWI and ICI channels divided up into window and sounder channels (as per Table 7). The channel categories are broadly defined by their mean τ (clear-sky surface to space transmittance) as modelled by RTTOV and follow the definitions used within the all-sky code to separate surface-sensitive, tropospheric sounding, and higher-peaking channels (see Geer *et al.* (2022), their Table 1). To be a pure sounding channel, τ needs to be generally less than 0.02, whereas window channels broadly have average τ values larger than 0.02. There is significant regional variation in τ particularly for higher frequency channels in atmospheric windows. The table includes columns for "stringent" calibration assessment and a more relaxed one to better assess the full dynamic range, "dynamic." The dynamic criteria should allow for more of the channel's full dynamic range to be assessed, including land impacts for lower frequency window channels for example, as ocean-only data selection permits only a small fraction of the total dynamic range to be assessed; compare Fig. 26 with Fig. 27 to see the difference in spread for O-B versus observed TB for ocean vs. land scenes using SSMIS data. For this channel, analogous to MWI-22 and ICI-1, there is significant spread of O-B at higher observed TB values that comes primarily from cloud and precipitation effects. However, we can see over land that a large number of data points exist at lower observed TBs that would be excluded if we considered an ocean-only selection of scenes, and the total dynamic range of the channel would be poorly sampled if considering clear-sky ocean scenes only. In practice, the lowest brightness temperatures over land at the higher frequencies typically come from volume scattering within the snowpack, in other words from areas where the forward modelling is highly uncertain and which would have to be removed from any cal/val sample. Hence, there is typically considerable tension between the goals of extending the dynamic range and retaining only scenes where the forward modelling is accurate. Hence the dynamic data selection in practice has to remove snow-affected areas. However, at lower frequencies than shown in Fig. 27, snow-free land tends to be be radiatively warmer than the ocean surface and these scenes can be successfully used to extend the dynamic range included in the sample. Note that τ is also used in the screening for sounding channels, since these channels will often be considerably more sensitive to the surface in certain conditions, such as a very dry atmospheric profile.

Table 8: The primary channel-based data selection criteria for MWIICI channel groups, separated into stringent and dynamic criteria. Here "land" refers to the FOV-specific land fraction at 18 GHz. Any land scenes used must be snow-free as currently determined for similar all-sky channels and exclude high orography (see Geer et al. (2022), their Sec. 4.3). Furthermore, window channels' stringent criteria requires SST $>$ 277 K and excludes the Great Lakes as well as Caspian and Aral Sea regions.

Channel Group	**Stringent**	**Dynamic**
Window channels	land$<1\%$, $CI < 2K$, $SST > 277K$	land$<1\%$ or land$>99\%$, $CI < 2K$
Sounder channels	$\tau < 0.02$ or land$<1\%$, $CI < 0.5K$	$\tau < 0.10$, $CI < 2K$

The stringent criteria are designed to screen out most cloud-affected scenes and most scenes with problematic surface emissivity characterisation. The dynamic criteria are designed to provide a larger data sample that better samples the dynamic range of the channel whilst avoiding the most biased scenes. Sounder channels have a tighter *CI* threshold for the stringent criteria as cloud signals are typically smaller for sounder frequencies, particularly in the 50 GHz band. For both stringent and dynamic data selections, sea-ice in the model is screened out for both sets of channels. To be conservative, there is a further SST threshold of 277 K included for window channels' stringent criteria to catch possible sea-ice that is not present in the model. This also helps to homogenise the sampling between GMI and the polar-orbiting sensors. Additionally, the Great Lakes and the area containing the Caspian and Aral seas are screened out from the window channels' stringent criteria as these regions have long been excluded from all-sky window channel assimilation. The main difference between the stringent and dynamic criteria is the addition of snow-free land scenes for window channels and the addition of scenes with low but greater surface sensitivity for sounding channels. The more relaxed dynamic criteria aims to provide a larger statistical sample to better assess the calibration characteristics over as large a portion of the full dynamic range as possible.

The *CI* thresholds chosen here are a compromise between allowing more data into the cal/val sample and excluding areas of cloud biases in the model. The emphasis was to find thresholds where the bias against the IFS was not especially sensitive to the choice of threshold; in other words, the relative bias should not be much different if using $CI < 2.5$ or $CI < 1.5$, though of course std(O-B) will be larger with a more lenient *CI* threshold. The use of the a *CI* threshold does mean that there is an implicit "first-guess check", as a clear scene in the model means that $(O-B) > 4.0$ is automatically discarded for window channels (i.e. $CI < 2.0$), and $(O-B) > 1.0$ is discarded for sounder channels ($CI < 0.5$). This rather strict first-guess check for sounder channels is more justified for 50 GHz channels where background errors are of order 0.1K than those at higher frequencies where background errors in clear-sky are closer to 1.0K. The decision to treat window and sounder channels in a combined way does simplify the interpretation when comparing similar but not identical channels within or between sensors, but this could be refined in the future.

The motivation for the τ-based selection criteria for sounding channels comes from English (2008), where the importance of surface emissivity errors on sounding channels was scrutinised. That paper showed that even small radiometric contributions from the surface (e.g. $\tau = 0.05$) could cause errors in top of atmosphere TBs, particularly when considering sounding channels where tenths of a Kelvin are non-negligible errors. Figure 1 from English (2008) showed that for $\tau < 0.02$ there is almost no sensitivity to emissivity errors, and even a 4K skin temperature error results in about 0.08K error in simulated TB. This was thus the transmittance threshold used for sounding frequencies in this study. The *tau*-based threshold is also situation-specific, ensuring that optically thicker scenes such as in the deep tropics for sub-mm channels are not screened out where surface emissivity errors are not at all significant.

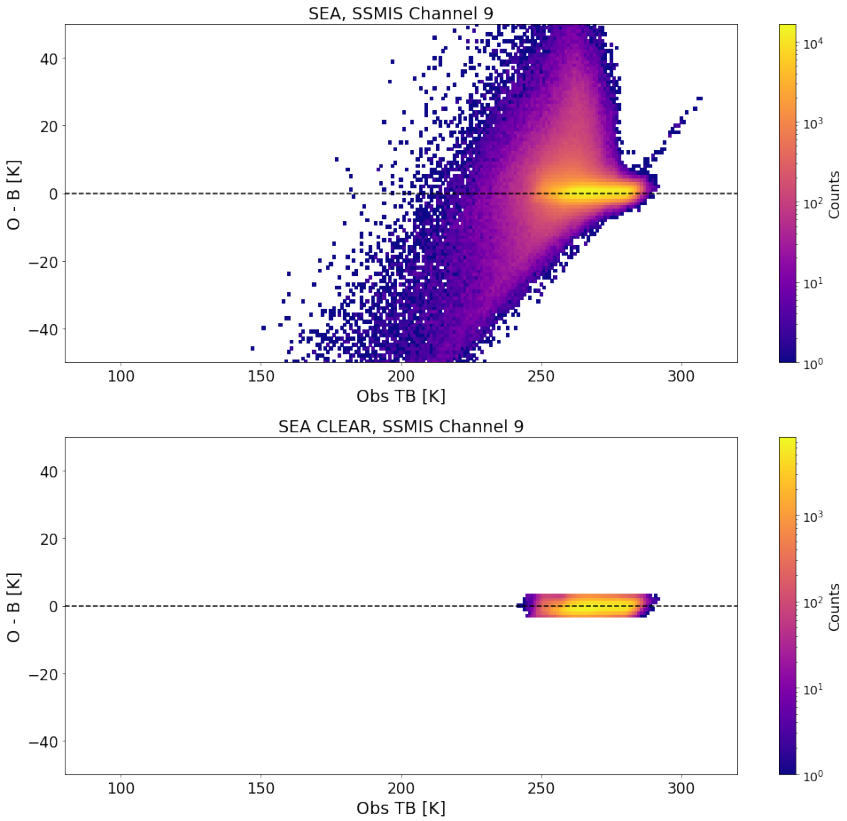


Figure 26: SSMIS channel 9 O-B vs. observed TB over sea (land fraction $< 1\%$). On top there is no data selection, i.e. a full all-sky sample, whereas on the bottom we use the $CI < 2K$ screening criteria.

This argument also holds for high-peaking 183 GHz channels (MWI-26 for example) that seldom see the surface.

6.2 Unified data selection

For assessment of requirements MWI-05110 and ICI-06110, a homogeneous sample is required to analyse inter-channel radiometric bias differences. This means that a selection criteria is needed that is applied evenly across all channels, irrespective of their particular sensitivities—a unified selection criteria. The unified selection should be as free of cloud and surface effects as possible for all channels. It is necessary to assiduously screen out areas of known model bias for any channel on the instrument—as such, cold air outbreak regions[8] are removed from the sample. This extra selection criteria is not deemed necessary for the channel-based screening, as CAO regions should be removed by checks on CI.

The difficulty in choosing a unified sample is that the channels on MWI and ICI have vastly different sensitivities. Furthermore, to compare between different sensors it is ideal to have criteria that can be applied similarly to all imagers. Initially, testing of different unified criteria for this project was centred on C_{P37}, as this is the backbone all-sky cloud predictor for MW imagers in the IFS. This is however not a good predictor for most ICI channels, and attempts to use a combination of C_{P37} with other predictors such as C_{SI} produced a biased overall sample by being too restrictive. Instead of using these traditional

[8]CAO regions are defined as scenes over sea with $TCWV < 8.0 kgm^{-2}$ or $(\theta_{700hPa} - \theta_{surface}) < 12.0K$ where θ is the potential temperature. See Lonitz and Geer (2015) for more detail.

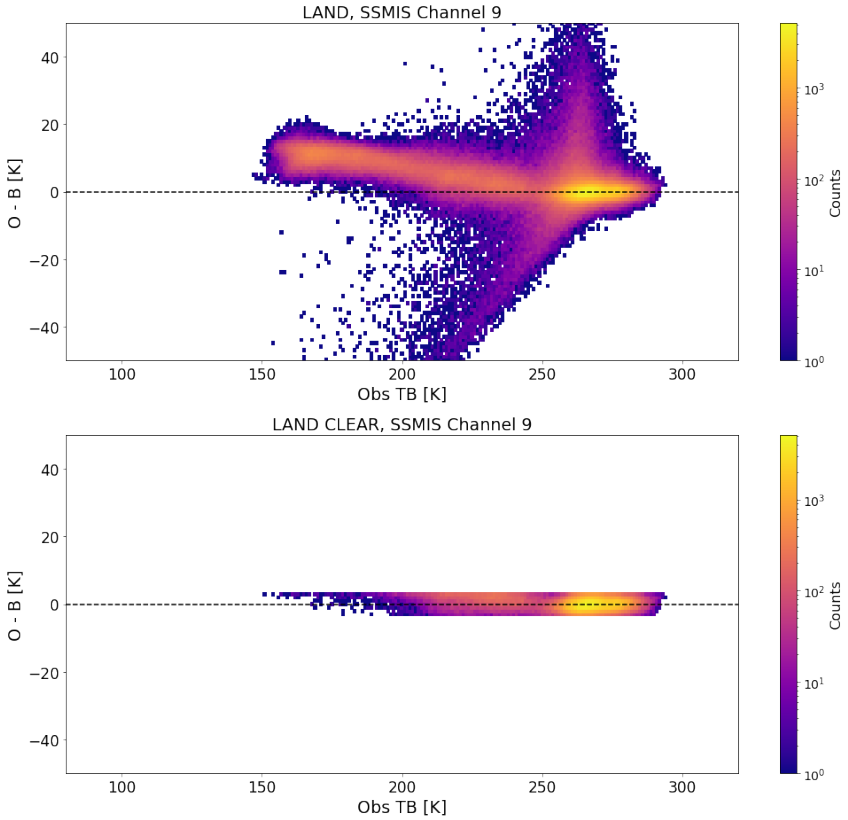


Figure 27: SSMIS channel 9 O-B vs. observed TB over land (land fraction $>$ 1%). On top there is no data selection, i.e. a full all-sky sample including mixed surfaces, coasts, lakes, snow, etc., whereas on the bottom we use the CI $<$ 2K screening criteria.

Table 9: The unified data criteria for different types of MW imagers. For all imagers, scenes need to be land< 0.01*, sea-ice free, SST*$> 277K$*, and avoid CAO regions.*

Sensor	Unified Criteria
MWIICI	$CI_{19V} < 2, CI_{89V} < 2, CI_{166V} < 2, CI_{243V} < 2, CI_{664V} < 2$
GMI	$CI_{19V} < 2, CI_{89V} < 2, CI_{166V} < 2$
SSMIS	$CI_{19V} < 2, CI_{89V} < 2, CI_{166H} < 2$
AMSR2	$CI_{19V} < 2, CI_{89V} < 2$

all-sky cloud predictors, the unified criteria uses a combination of *CI* at several key frequency bands. This essentially expands on the method defined above by selecting *CI* at channels with specific sensitivities which we want to use for screening. For MWIICI these channels and their primary sensitivities are:

- 19V – precipitation and surface
- 89V – liquid cloud
- 166V[9] – liquid and frozen hydrometeors
- 243V – frozen hydrometeors
- 664V – cirrus clouds

For each of these bands, *CI* is checked as it is for the channel-specific criteria, i.e. $CI < 2K$. If the threshold is met for all channels that are available, then these data points are included in the unified data sample. Due to the simplicity of this criteria, it can be applied in the same manner to other sensors on the basis of which frequency bands they have. For instance, GMI has 19V, 89V, and 166V channels (but no 243V or 664V), so its unified sample depends only on *CI* at the three bands it has available. Note that V-pol channels are preferred in the selection criteria due to known emissivity biases for H-pol channels at mid frequencies (Geer *et al.*, 2024), but not all current imagers have a 166V, so 166H is used instead if that is available.

Table 9 shows the unified selection criteria for different MW imagers. Regardless of the sensor, the criteria requires land-free, ice-free scenes and avoids regions designated as cold-air outbreaks as these are known to be biased in the IFS for imager channels (Forbes *et al.*, 2016). In addition, SSTs colder than 277K are excluded from the sample to avoid regions with fractional sea-ice that may not be represented in the model sea-ice field. The range of criteria applied for MWIICI is intended to cover the full range of cloud sensitivities from sounding and window channels on the two sensors, from precipitation to cirrus clouds.

6.3 Preliminary analysis of data selection

In this section, some targeted analysis is performed to investigate the efficacy of the data selection criteria previously presented. It is of course preliminary due to the use of simulated test data for MWIICI. Realistic performance of the data selection is provided in the following section when applied to current instruments.

[9] If 166V not available, use 166H (e.g. SSMIS)

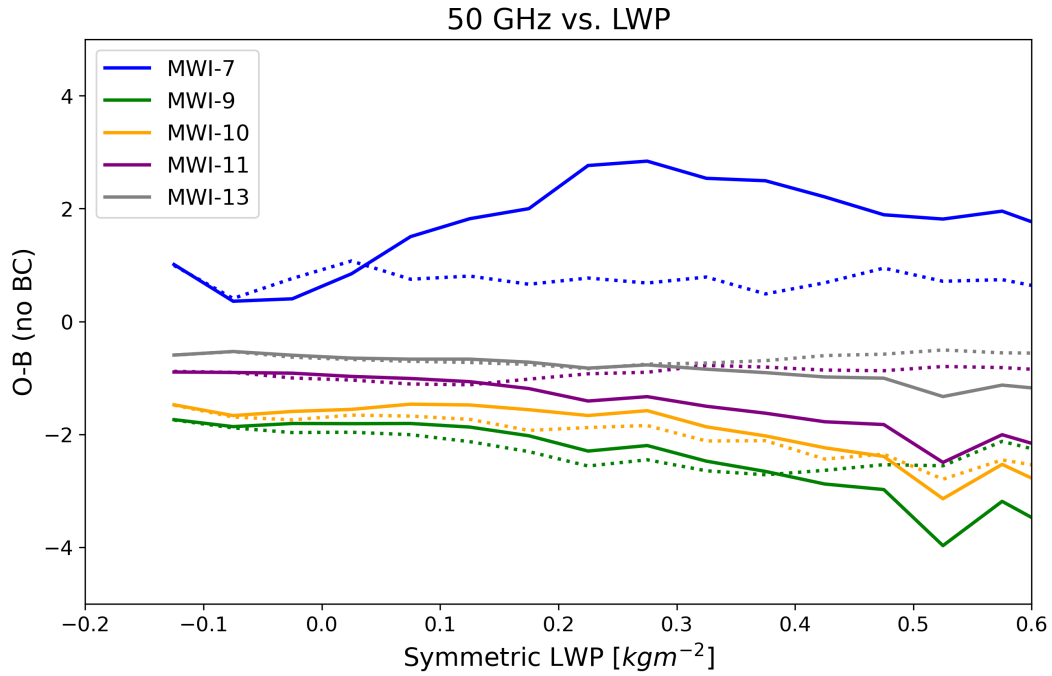


Figure 28: O-B binned as a function of symmetric LWP (C_{LWP}) for V-pol 50 GHz channels on MWI. The selections shown are all ice-free data points over sea (solid) and stringent criteria (dotted). Data from 12Z long window only. Negative LWP values are an artefact of the regression used to derive LWP, see Duncan et al. (2022b).

Figure 28 shows the dependence of mean analysed biases at a selection of 50GHz channels on the symmetric Liquid Water Path (LWP). This demonstrates that the negative O-B bias seen in global statistics does have a dependency on the observed and background liquid cloud amount, but that even for LWP values near zero this negative bias persists. The application of different data selection criteria largely mitigates this LWP dependence, particularly for the more conservative criteria (i.e. stringent and unified), with the stringent criteria shown in the figure as having a more limited dependence of analysed bias on a scene's symmetric LWP. The contrast of positive bias at channel 7 and negative bias at higher-peaking 50 GHz channels indicates the differing signs of bias between the test data and IFS from aspects of surface emissivity and atmospheric emission.

For a selection of MWIICI channels, Fig. 29 shows how the mean O-B is related to the variability of observed radiances within a superob. Larger departures are often associated with more variable superobs, though this signal saturates for the more scattering-affected channels 30 and 39. There is significant variability between channels as to whether std(TB) is a good predictor of larger magnitude departures. The imager channels 1 and 16 show little dependence between O-B and std(TB), apart from larger positive departures for channel 1. Particularly after data selection as seen in the bottom panel, channel 16 shows a flat curve. For some sounding channels, std(TB) could be a useful discriminant in addition to the current data selection criteria. For example, channels 11 and 39 have quite different atmospheric sensitivities, but both channels show a correlation between larger negative departures and more variable superobs. A further view is found in Fig. 30, showing histograms of selected MWIICI channels binned by superob variability. This shows that the data selection criteria removes a large fraction of the more variable superobs and retains the most homogeneous superobs without explicitly needing an extra criteria to do so. It may be worth considering std(TB) as a rough quality control instrument for data selection, but it is clear that a more targeted use of std(TB) would require consideration on a per-channel basis; indeed this may

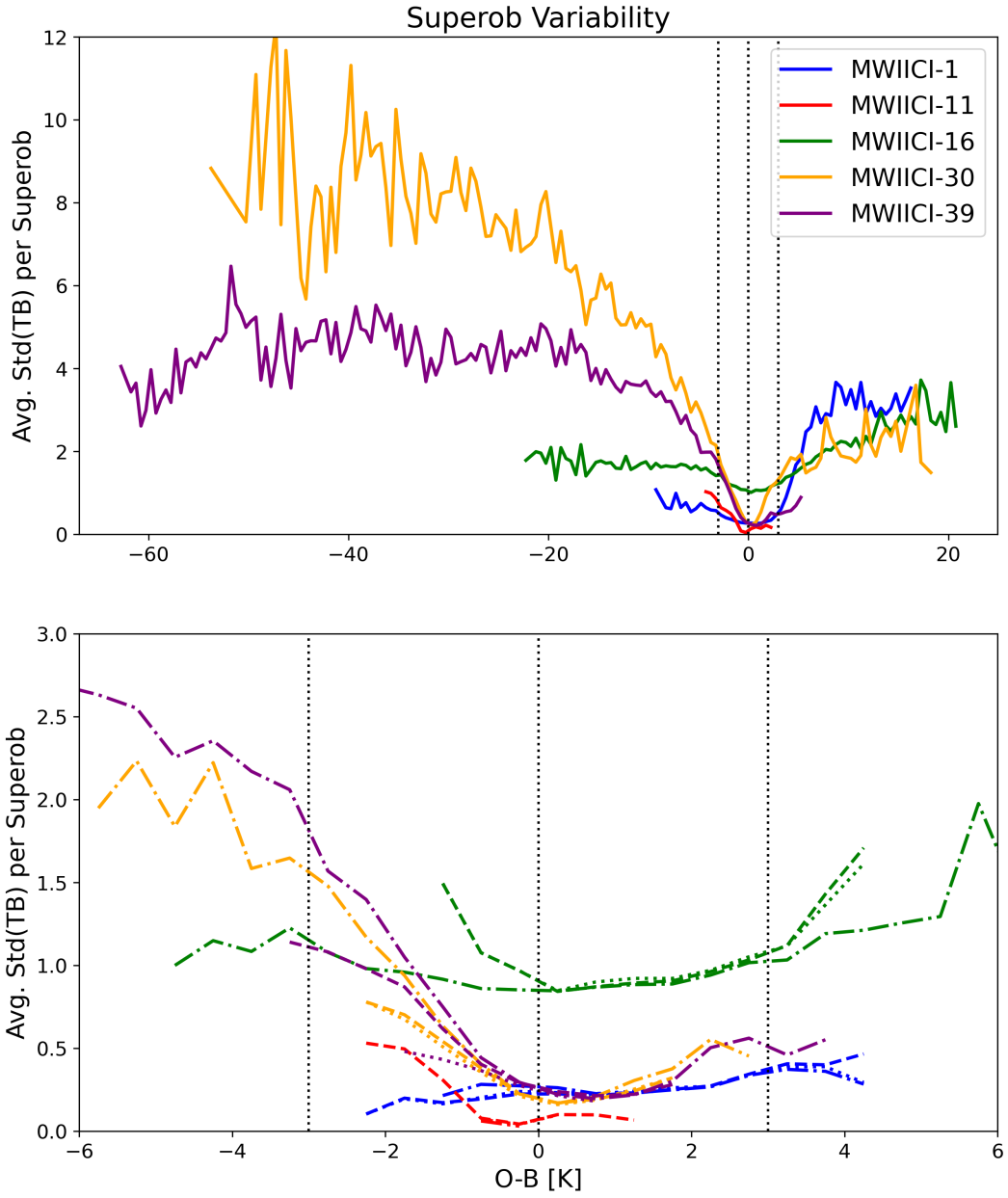


Figure 29: Standard deviation of superob (std(TB)) as a function of O-B. These comprise MWIICI data from the two test orbits on Sept 12 2007, all data over ice-free ocean scenes. In the top panel are shown all data (solid), while in the bottom panel there is a zoomed-in view showing the data selection criteria for dynamic (dashed), stringent (dotted), and unified (dot-dash). Note that because these data are ocean-only, the dotted line is identical to the dashed line for most channels and thus not visible.

prove more useful as a metric for quality control when applied to real observed radiances with sensor noise. This has been left for future work, and would only be applicable to MWIICI at first, as this is the only sensor in the IFS that has std(TB) stored in ODB.

Further consideration of the data selection's efficacy is considered in the section on existing sensors (Sec. 8), as this type of analysis is more suited to real observational data than the simulated test data.

7 Framework to support cal/val

7.1 Monitoring data quality via Obstat

Obstat is the operational software for calculating and plotting statistics for all observations monitored and assimilated in the IFS. Operational monitoring figures are updated automatically as part of the ECMWF operations and are publicly available. The monitoring website introduced in the next subsection follows the format of the general observation monitoring suite run operationally at ECMWF, which covers satellite radiances as well as conventional data, ocean data, and more. The monitoring page of ecCharts for all observations can be found here: https://charts.ecmwf.int/catalogue/packages/obstat/

This study uses the Obstat capabilities for calculation of departure statistics and production of figures. The Obstat software aggregates observational statistics across all instruments that are processed through the IFS. This includes tracking the number of observations monitored and used in the assimilation by every sensor on every platform, with various departure statistics calculated and collated at the end of every cycle in an experiment. These include global statistics as well as pre-defined regions such as the northern and southern hemisphere extratropics, tropics, Europe, and so on. This wealth of statistical information is archived for every cycle of an IFS experiment, permitting downstream analysis with either online or off-line plotting tools. Obstat contains a suite of plotting tools for maps, time series, Hovmöllers, scatter plots, line plots, and more.

There is a standard set of monitoring plots that is produced by Obstat for all operational suites at ECMWF. This capability can be leveraged by any IFS experiment, using the statistics files routinely output during the archive tasks of a running experiment. In this project, a custom set of Obstat plotting scripts has been developed to display statistics tailored to cal/val analysis of MWIICI.

7.2 Website (including user guide)

A website has been created for analysing departure statistics from MWIICI. On the website, the user can subset the total number of plots with a few different options, provided along the left side of the screen. One can select the instrument, for instance AMSR2, GMI, SSMIS, or MWIICI. Although bias correction estimates for MWIICI test data have limited validity due to the limited sample of two orbits (i.e. VarBC has not had time to spin up[10]), it is possible to view the statistics for all instruments with and without bias correction applied. The full set of plot types includes the following, with associated requirements in parentheses:

[10]To simulate the spin-up of bias correction coefficients and test that everything was working as expected, the experiment was submitted three times, each time using the VarBC coefficients from the previous run. This was a quick way to get somewhat reasonable bias correction for testing, but it is not a real substitute for several days of full data coverage to spin up VarBC.

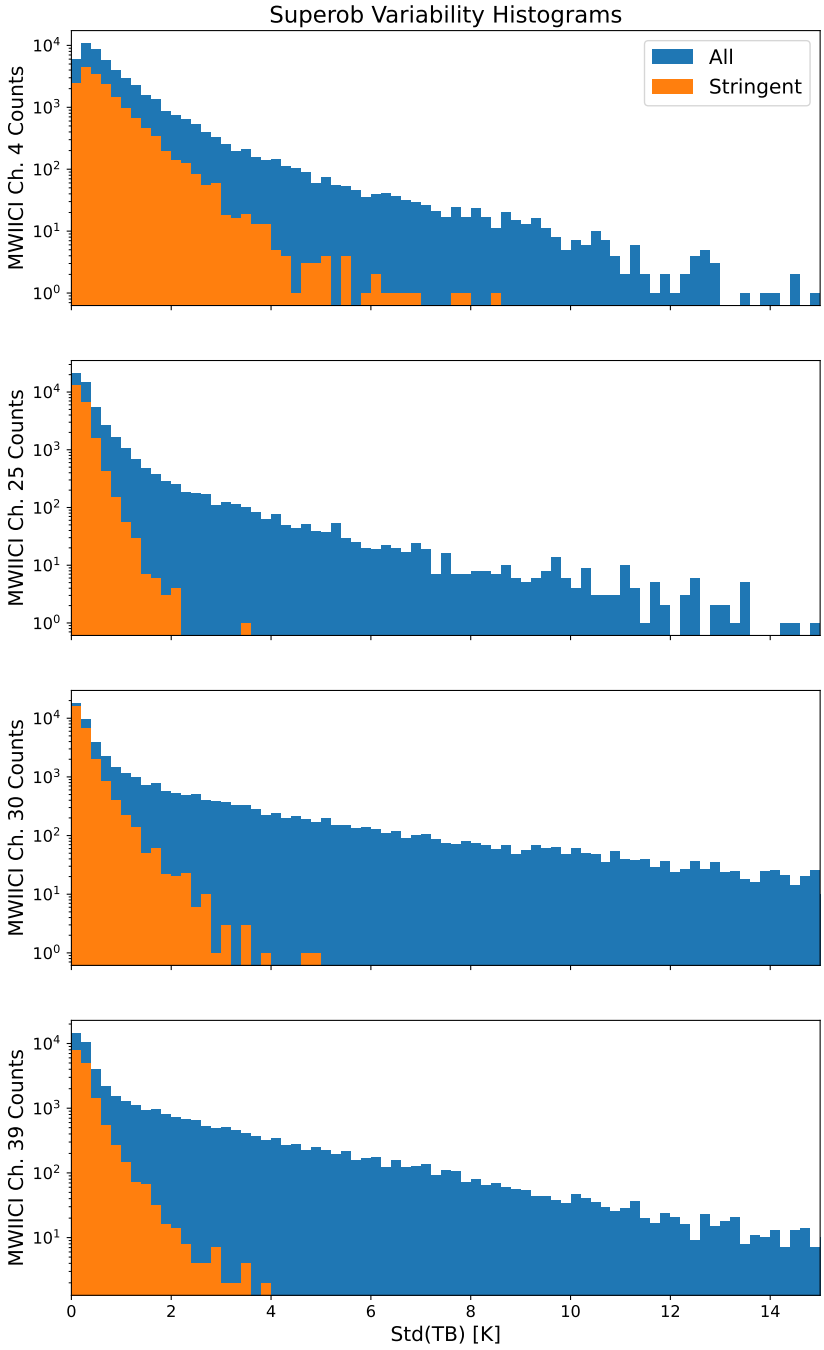


Figure 30: Histograms of selected MWIICI channels binned by std(TB) per superob, covering the same data as the previous figure.

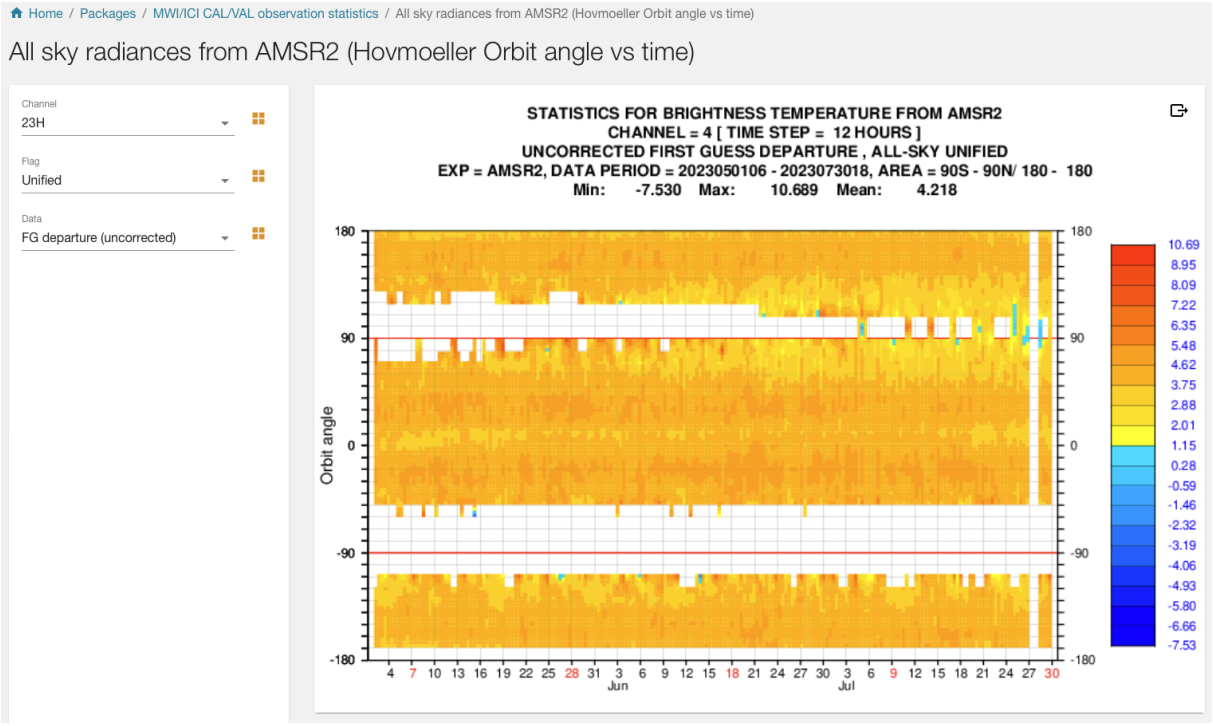


Figure 31: Screenshot of the Hovmöller plot type, with time on the x-axis and the orbital angle on the y-axis. Here the AMSR2 23H channel is shown (corresponding to MWIICI channel 4) with the unified data selection criteria applied.

- Time series of area averages (MWI-05100, ICI-06100)
- Hovmöller diagrams of orbital angle against time (MWI-05090, ICI-06090)
- Maps of time-mean statistics
- Area-averaged statistics for several channels shown together, either for one or multiple sensors on the same figure (MWI-05080, ICI-06080; MWI-05110, ICI-06110)
- 2D histograms for assessing bias as a function of a predictor (MWI-05120, ICI-06120)

Two types of time series plots are available, namely a standard time series of area averages and a Hovmöller-style plot showing possible orbital angle dependence of departures. Figure 31 shows departures as a function of orbital angle, here showing AMSR2 23H data from the unified data selection (data selections are defined in Sec. 6). In this case, no data are available over the poles (near 90 and -90 orbital angle, the north and south poles, respectively), with a relatively constant positive mean bias as a function of time and orbital angle. The standard global time-series plot type is exemplified in Fig. 32, with the 183±7 channel from GMI shown over a three month period.

Maps of gridded observations are available at a 2 degree resolution globally. An example of this is given by Fig. 33, showing the mean departure for the 37V channel on GMI for a month of data in 2021. As this channel is relatively unbiased against the IFS, the main signals seen here are over land and areas of probable sea-ice.

For each instrument the user can see area averages of channel statistics provided in the same figure. These use the MWIICI channel indicators to represent each microwave imager in a common context, utilising

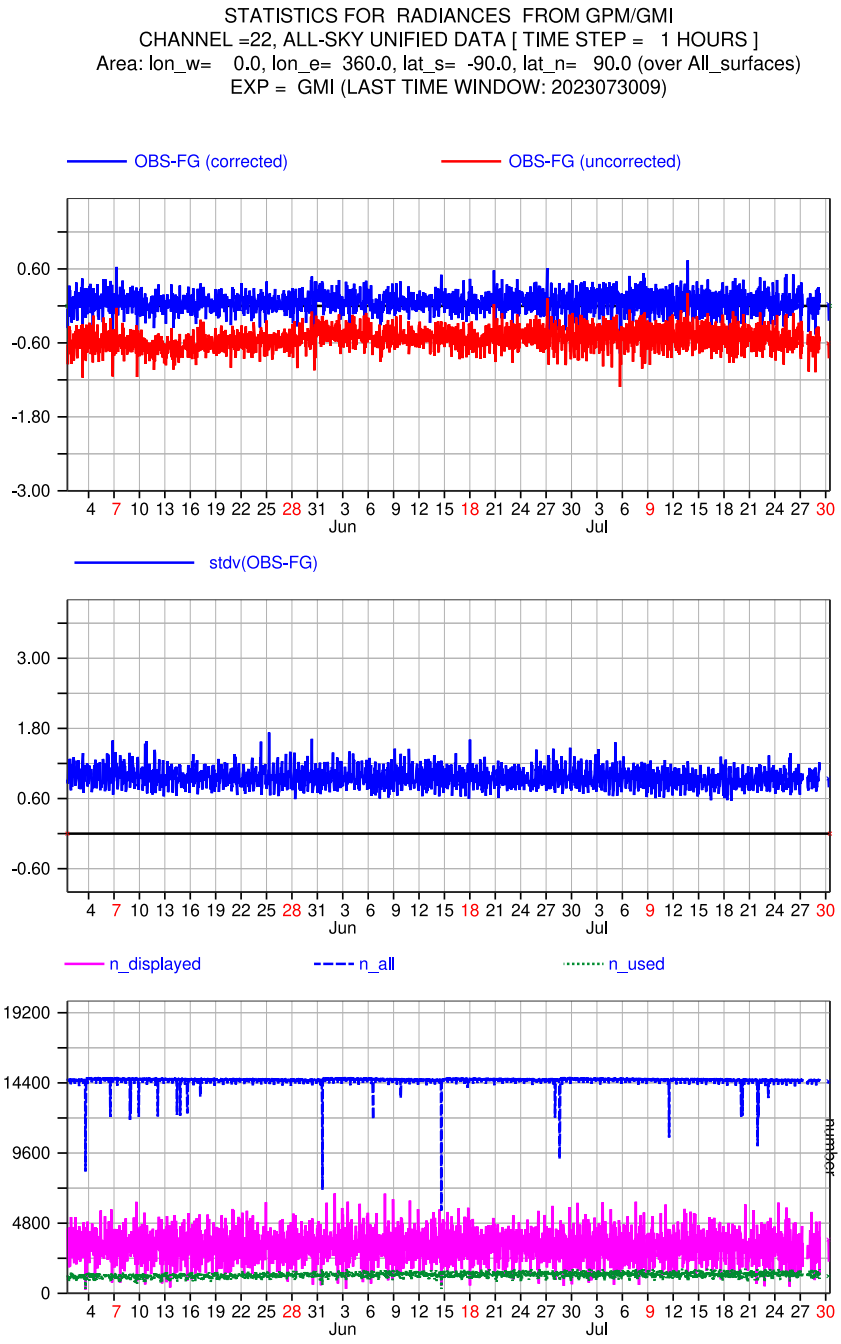


Figure 32: Example of the time series plot type, with time on the x-axis and panels showing mean O-B (top), std(O-B) (middle), and data counts (bottom). This example shows the 183±7 channel from GMI (N.B. mapped to MWIICI channel 22) with the unified selection criteria applied. In the bottom panel, counts are given for all, used, and selected (i.e. unified) data.

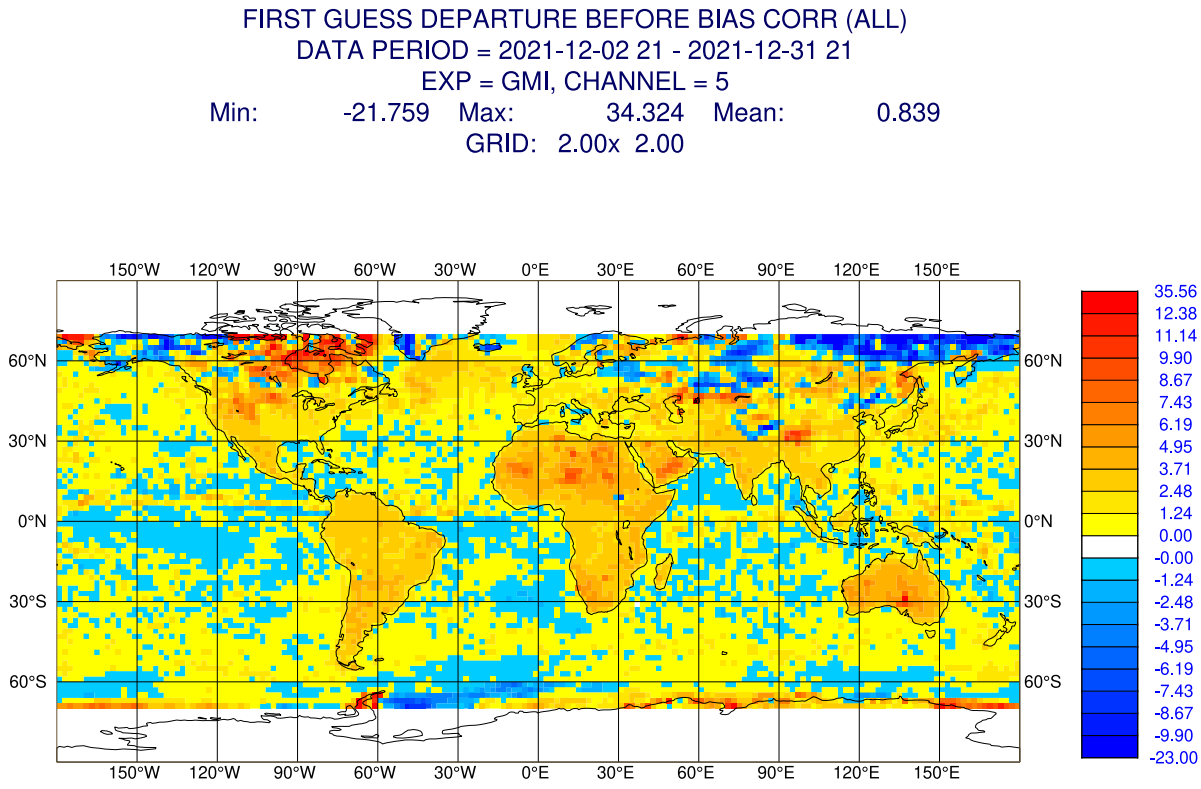


Figure 33: Map of gridded global departure statistics, shown for GMI channel 37V (N.B. MWIICI channel 5) for one month of all available data.

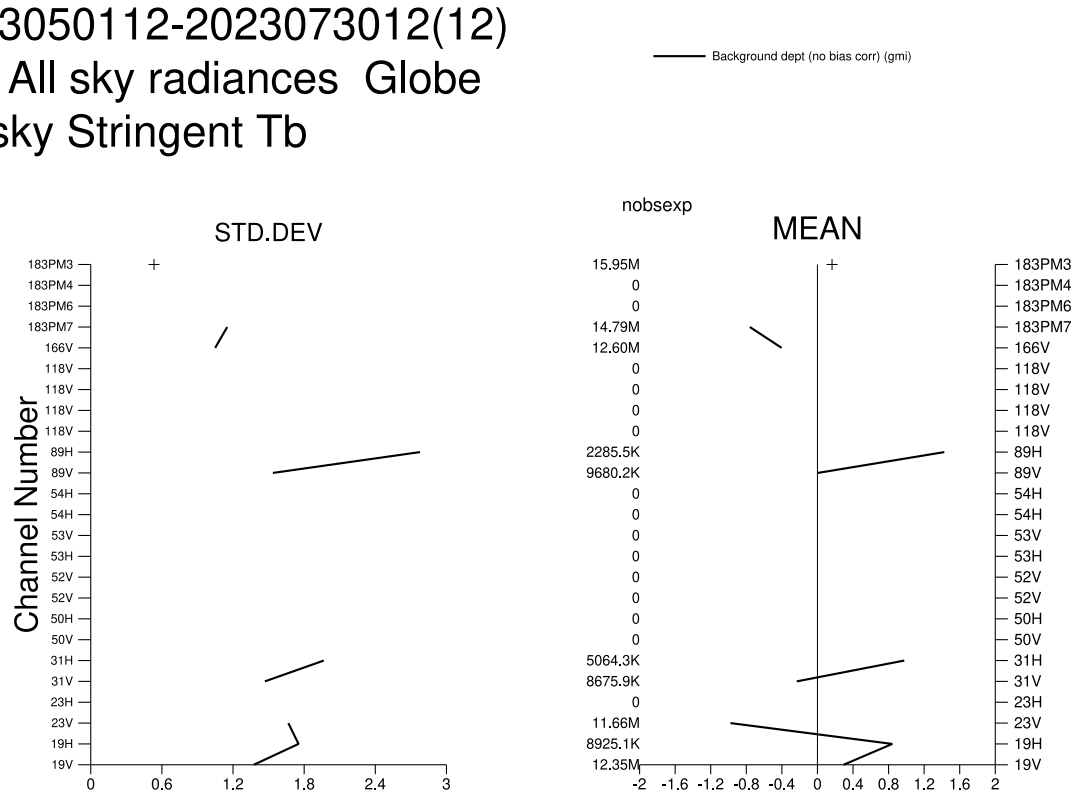


Figure 34: Global data using the stringent data selection, given for GMI. The y-axis shows a subset of MWIICI channel indicators, with the panels showing the standard deviation of O-B (left) and mean O-B (right), each with no bias applied. The number of observations in the sample for each channel is given along the y-axis on the right panel (K for thousand, M for million).

the channel mapping from Table 7. Figure 34 is an example of global data for the GMI sensor. Here the standard deviation of un-bias-corrected departures is on the left with mean departures on the right. For this type of plot, different regions can be selected: Global (90N-90S), Northern Hemisphere (20-90N), Tropics (20S-20N), or Southern Hemisphere (20-90S). Different data selection criteria are available in the drop-down menu (All, Used, Dynamic, Stringent, Unified), but just the stringent selection is shown in the figure. Consecutively available channels are shown with black lines, with grey crosses showing the other channels' data points.

2D histogram plots are available for a range of parameters, providing departures as a function of total columnar water vapour (TCWV), 10m wind speed, sample number across the scan, orbit angle, and scene (observed) TB. Figure 35 is an example of assimilated (i.e. used) AMSR2 data points at the 23V channel as a function of TCWV over a period of one month. The user can make different data selections, choose the channel, and choose whether the statistics shown are for bias-corrected departures or not. For each bin, the mean departure is marked with a black dot.

One type of plot combines data from multiple sensors together. Figure 36 provides departure statistics for three microwave imagers together using the MWIICI channel set (y-axis) as the basis for comparison. This plot shows that GMI has the lowest mean biases against the IFS for most channels, with both SSMIS and AMSR2 showing significant biases prior to bias correction, particularly for the lower frequency

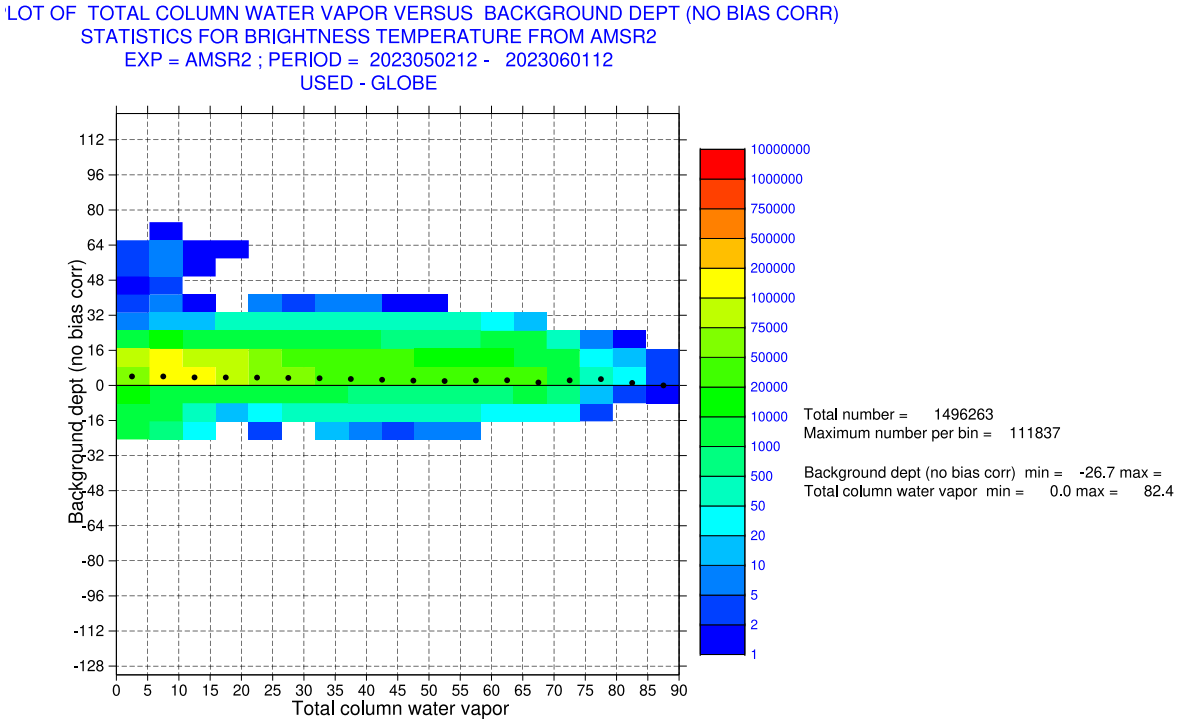


Figure 35: 2D histogram plot of the AMSR2 23V channel un-bias-corrected departures (y-axis) against TCWV (x-axis) for the selection of assimilated (i.e. used) radiances only. The coloured squares show the histogram of data counts, with black dots showing the mean value.

channels[11]. Due to these strong biases, the standard deviations are also elevated for these channels.

Lastly, the functionality to download summary statistics is available via the website. Beneath most figures (excepting the Hovmöller FOV vs. time and map plots) there is a generic title and then a *wget* command to obtain the statistics ASCII file that was used to create the plot. This should permit the user to execute the text on a command line and obtain the summary statistics. The ASCII files provide counts, means, and standard deviations per 12 hour assimilation cycle or per bin, allowing further aggregation of statistics as the user wishes. For example, the combination of counts and standard deviations could be used for significance testing of time series values.

7.3 Analysis of requirements

With reference to the EURD requirements listed for MWI and ICI in Sec. 2, this section describes whether and how these requirements can be assessed using the plot types generated in this project. The list combines those from MWI and ICI with the requirements paraphrased as needed and the analysis in italics. Note that some of these requirements are included only for information and are not a main goal of the project: for example NEDT and dynamic range requirements. Only for a subset of these requirements is the vicarious calibration using NWP fields an intended part of the relevant cal/val plans.

- **MWI-05060, ICI-06060** The dynamic range of the MWIICI shall cover the spectral radiances in terms of brightness temperatures as given in Tables 1, 2. *This cannot be quantified directly using*

[11] SSMIS 50 and 52 GHz channels are considered broken; they have very large biases and are thus not assimilated in the IFS

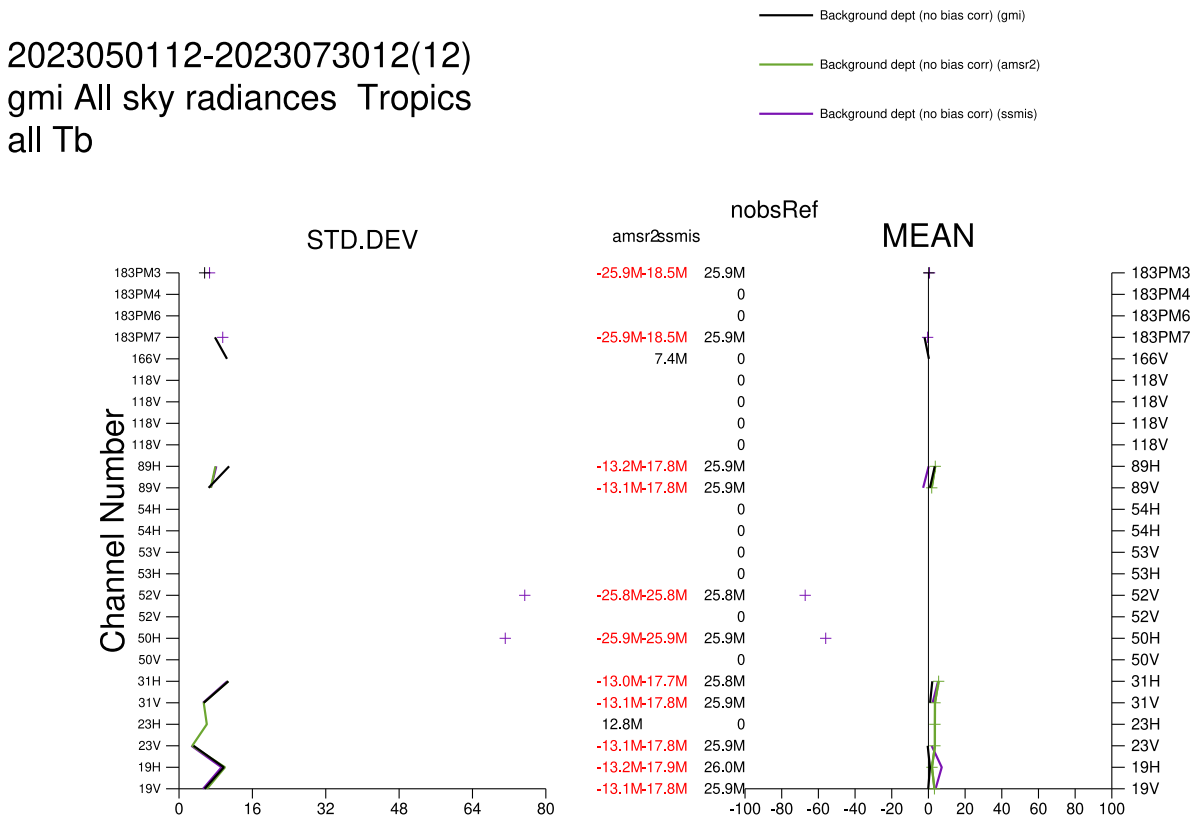


Figure 36: Example of inter-channel comparison for GMI (black), AMSR2 (green), and SSMIS (purple). Here all data over the tropics (20N-20S) was selected. The y-axis is given in terms of MWIICI channel indicators. Channel counts are shown in the middle, with SSMIS as the reference.

departure analysis, but scatter plots such as Fig. 27 can be used qualitatively.

- **MWI-05065, ICI-06065** All MWIICI requirements shall be met over the dynamic ranges given in Tables 1, 2. *As with the previous requirement, this may be addressed qualitatively with some of the analysis using the "dynamic range" data selection, but this will be a challenge for certain channels and especially low and high TBs in particular.*

- **MWI-05070, ICI-06070** The MWIICI radiometric sensitivity requirement is the NEDT in Tables 1, 2. *The NEDT requirements can be assessed via $std(O-B)$ as shown for example in Fig. 37. However, this is really only applicable for channels where sensor noise dominates the departure signals, such as 50 GHz sounding channels in clear scenes. Gross violations of NEDT requirements should be noticeable for most channels using $std(O-B)$ analysis using a stringent data selection, and insight on relative NEDT can be gained via comparison with other sensors, as in Fig. 36.*

- **MWI-05080, ICI-06080** The MWIICI radiometric bias shall be better than those listed in Tables 1, 2 *This can be assessed relative to the IFS in the context of other sensors. There is no guarantee on the accuracy of the IFS, but it is useful as a transfer standard in particular for other sensors that are already well simulated; this is more of a challenge for sub-mm frequencies, where the accuracy of the background and the forward model are not as well known. This is assessed in Fig. 37 against the IFS and can be inter-compared to other sensors using plots like Fig. 36.*

- **MWI-05090, ICI-06090** Orbit stability: Variations of radiometric biases of the measured MWIICI brightness temperature during an orbit shall be < 0.6 K / < 1.1 K. *The orbital angle shown in Fig. 7 is used in a Hovmöller-style plot to assess orbit stability over time (as in Fig. 31). Both the IFS-defined and EUMETSAT orbit angles are passed through the IFS, with Hovmöller-style plots for each to show bias against orbit angle as a function of time. 2D histograms of biases as a function of both orbit angles are also provided, covering the most recent month of data.*

- **MWI-05100, ICI-06100** Lifetime stability: Variations of the running average of radiometric biases over any one orbit of the measured MWIICI brightness temperature shall be < 0.25 K / < 0.6 K throughout the mission lifetime. *This requirement will be assessed with the stringent data selection, showing the global mean (per channel) as a function of time. Means can be calculated per-orbit or per-day, with the caveat that per-orbit data samples could be too restrictive for reasonable statistics particularly for the stringent selection criteria. Assessment is done using the orbit-time Hovmöllers as illustrated in Fig. 7 as well as the regional time means as a function of day or orbit number as illustrated in Fig. 32. Any changes in the IFS itself (e.g. from a cycle upgrade) could manifest as discontinuities in the time series, but these should be clear from cross-checking with other sensors..*

- **MWI-05110, ICI-06110** Inter-channel radiometric bias differences between brightness temperatures of the same MWIICI spatial sample shall be < 0.6 K. *This is assessed by plots such as Fig. 37, however the statistics should be calculated using the same data selection across all channels. This is accomplished using the unified data selection.*

- **MWI-05120, ICI-06120** Inter-footprint radiometric bias differences between brightness temperatures of the same MWIICI spectral channel shall be < 0.4 K / < 0.6 K. *To assess this exactly requires a special experiment with data passed through without superobbing so that individual scan positions can be analysed (Sec. 5.4). Biases as a function of scan position ("scan bias") are assessed using a special time-mean scatter plot of the same scan position biases (Fig. 23). It is also possible to learn a lot about scan bias from the O-B bias as a function of the mean scan position of*

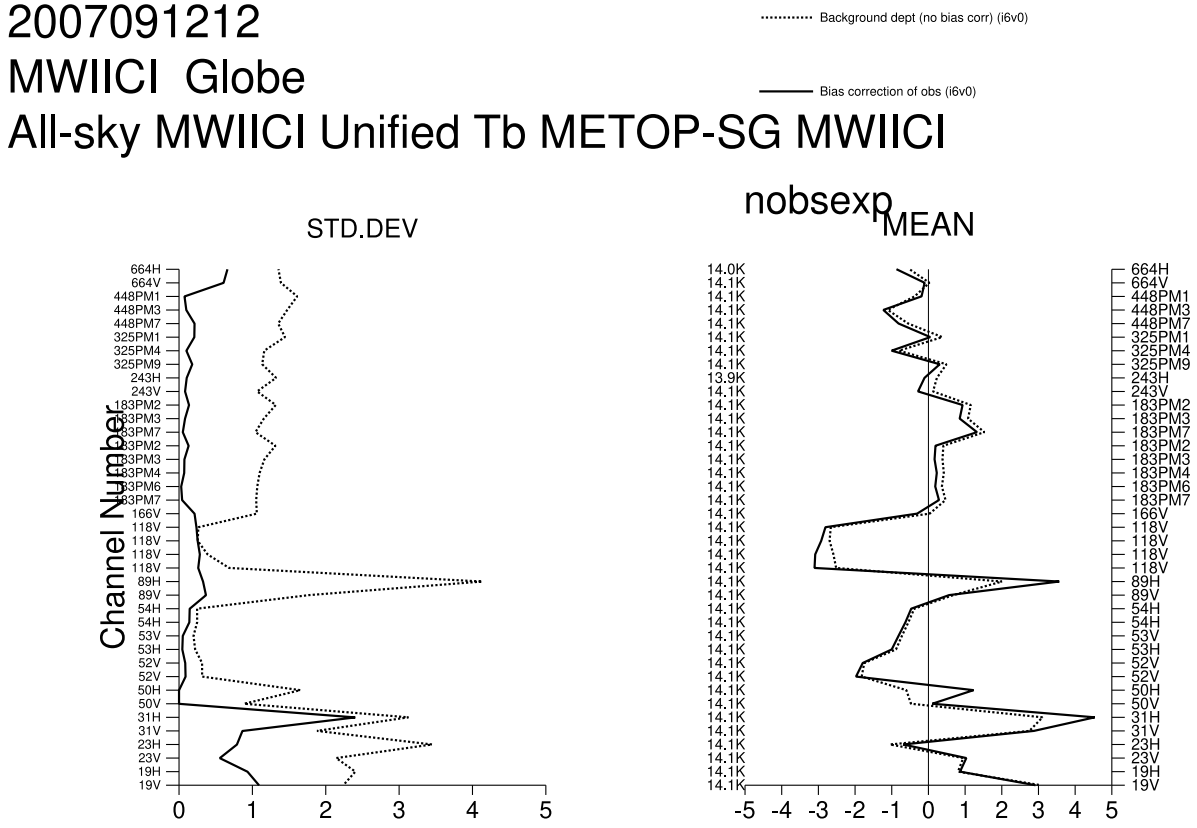


Figure 37: Global data from MWIICI using the unified data selection, 12Z cycle on 12th September 2007. Standard deviation of O-B (dotted) and bias correction (solid) in the left panel; average departure without bias correction (dotted) and averaged bias correction (solid) in the right panel.

the superobs. These plots will have the advantage of being generated routinely, in order to provide continuous monitoring of the scan bias, albeit with some smoothing over scan position compared to the special non-superobbed experiments. For superobbed radiances, the bias as a function of scan position can be assessed regularly (per 12hr cycle) via 2D histograms and Hovmöller-style plots.

Figure 37 shows an example of global statistics for the unified data sample from MWIICI, the means and standard deviations of departures as plotted by Obstat. Note that the number of observations is nearly the same for all channels due to the unified data selection being used, differing only due to the edges of the scan where not all channels have valid data.

8 Results for current sensors

In this section, the data selection criteria are applied to current MW imagers GMI, SSMIS on DMSP F17, and AMSR2. Following the results is an analysis of the data selection's efficacy. Further plots supporting the analysis are available on the demonstration website.

8.1 Comparison of GMI, SSMIS, and AMSR2

To illustrate the three different selection criteria (dynamic, stringent, unified) as applied to real data, these are compared across Figs. 38 and 39 for the 166V channel on GMI, corresponding to MWI channel 21. This is a channel with mixed imaging and sounding capabilities, with almost no surface sensitivity in the tropics and 50% surface to space transmittance in dry profiles at high latitudes (Fig. 19), though it is defined as a window channel in the IFS. As seen in the panel with all data considered, mean biases can easily reach 5-10 K in areas of sea-ice or snow, with generally positive biases in the deep tropics and storm track regions. The dynamic criteria shrinks the colour scale significantly and produces a mean global bias of about -0.3 K, though this includes regions of considerable bias over land, particularly over likely snow cover in Siberia and Alaska, with negative biases over the Sahara. The unified and stringent criteria remove these problematic regions over land, though the global mean bias is almost the same. The stringent selection contains some smaller areas of slight positive biases between 0.0-0.5 K in the deep tropics and Southern Ocean, likely representing weak cloud biases that were not entirely screened out. Outside these regions the picture is quite consistent, with a mean bias between GMI 166V and the IFS of about -0.2 to -0.7 K. The unified data selection shows a global mean bias that is nearly identical, but with more limited sampling. In fact, the fraction of data retained for analysis varies widely, from 31m total observations, 17.8m remain in the dynamic population, 12.7m for stringent, and only 7.0m for unified.

Figure 40 shows the inter-sensor means and standard deviations of departures for the stringent and unified selections, comparing the three MW imagers currently assimilated in the IFS. Seeing these plots together also lets us compare the stringent and unified selection criteria themselves with real-world data. For the stringent criteria, GMI exhibits a mean bias within -1 to 1 K for most channels, and, of the three sensors, the lowest std(O-B) for all channels except 89H. AMSR2 exhibits fairly consistent positive biases of about 3 to 5 K. F17 SSMIS channels exhibit more variable biases, from strong positive biases at lower frequencies to a -4 K bias at 89V, whereas std(O-B) follows equivalent GMI channels closely, albeit with greater variability at 183 GHz. These results are broadly in line with Lean *et al.* (2017, their Fig. 3), with the caveats that we are using SURFEM here rather than FASTEM, and that GMI L1 products saw a change in the spillover correction in 2017 that affected lower frequencies by as much as 1K. For the unified data selection, most channel biases are nearly identical to those of the stringent selection. The main channels that show any discrepancy are the mid-frequency horizontally-polarised channels, 31H and 89H; this is because the unified criteria checks vertically-polarised channels for *CI*, allowing more data through the screening. These channels are known to have the worst bias characteristics for surface emissivity modelling (Geer *et al.*, 2024), and the bias correction scheme struggles to spin up to a reasonable state because these data are not assimilated. Overall, comparison of these plots shows good consistency of the bias characteristics of imager channels whether using the stringent or the unified selection for cal/val.

8.2 Analysis of data selection for current sensors

A good selection criteria for cal/val should ideally show consistency in the geographical distribution of departure statistics between equivalent channels on different instruments. Furthermore, we would expect a good selection criteria to reveal instrument biases that have been noted previously, with the caveat that instrument behaviour can indeed change over time and may have done so in intervening years. A corollary of this is that a good criteria is expected to show near-zero, geographically consistent bias in all GMI channels, in light of its use as a calibration standard for the GPM constellation (Berg *et al.*, 2016) and the extraordinary care taken to ensure its precise calibration (Wentz and Draper, 2016).

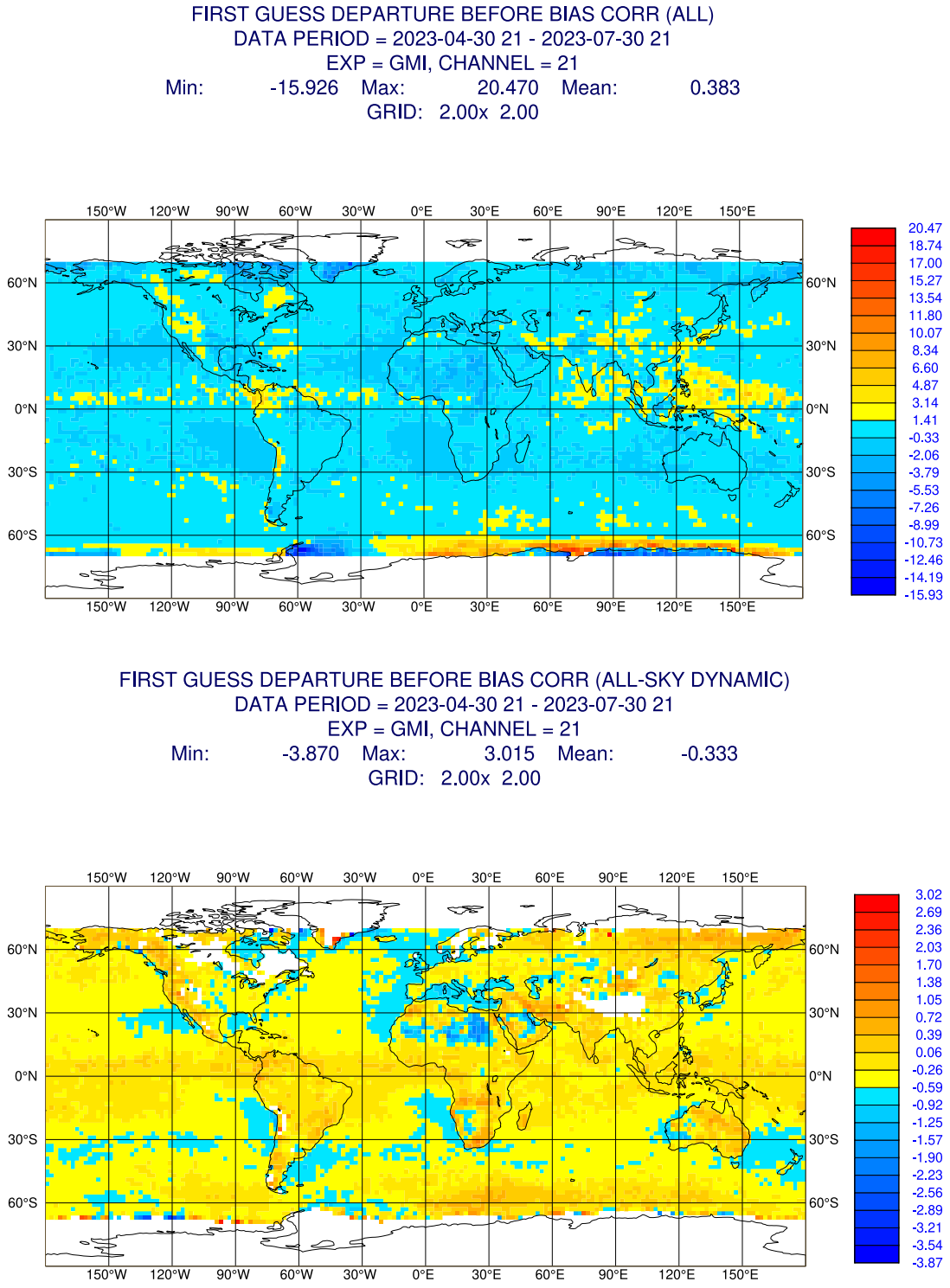
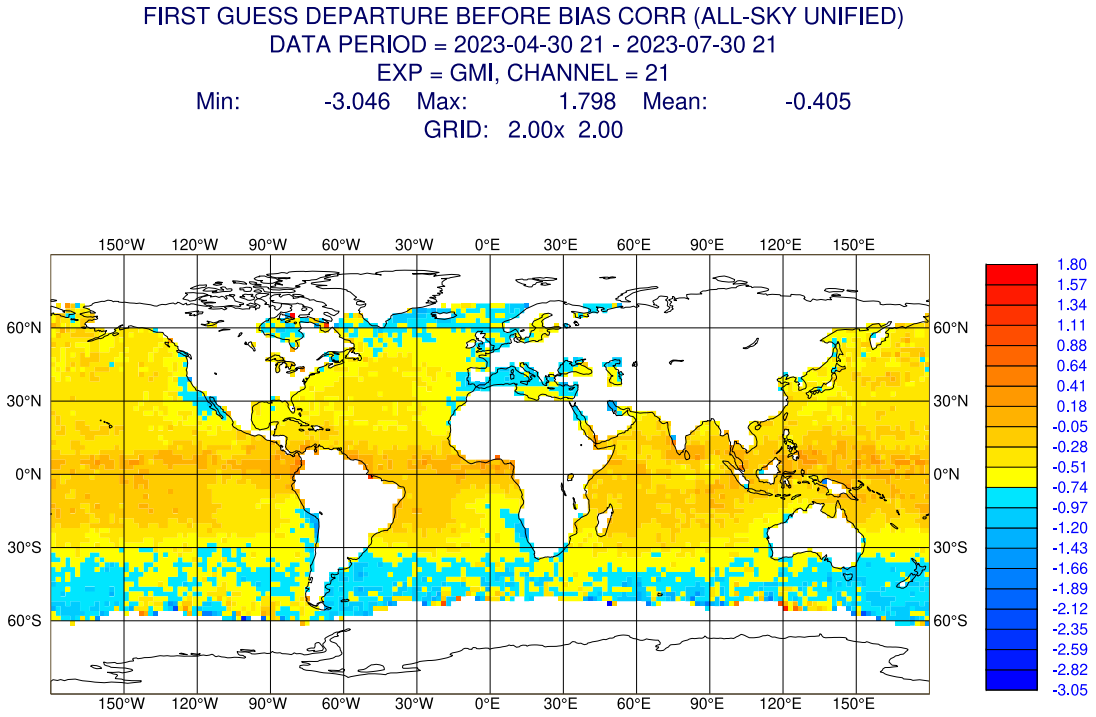


Figure 38: GMI 166V channel departures without bias correction applied, averaged over three months (May to July 2023) on a 2x2 degree global grid. The top panel shows all data whilst the bottom panel shows the dynamic data selection.



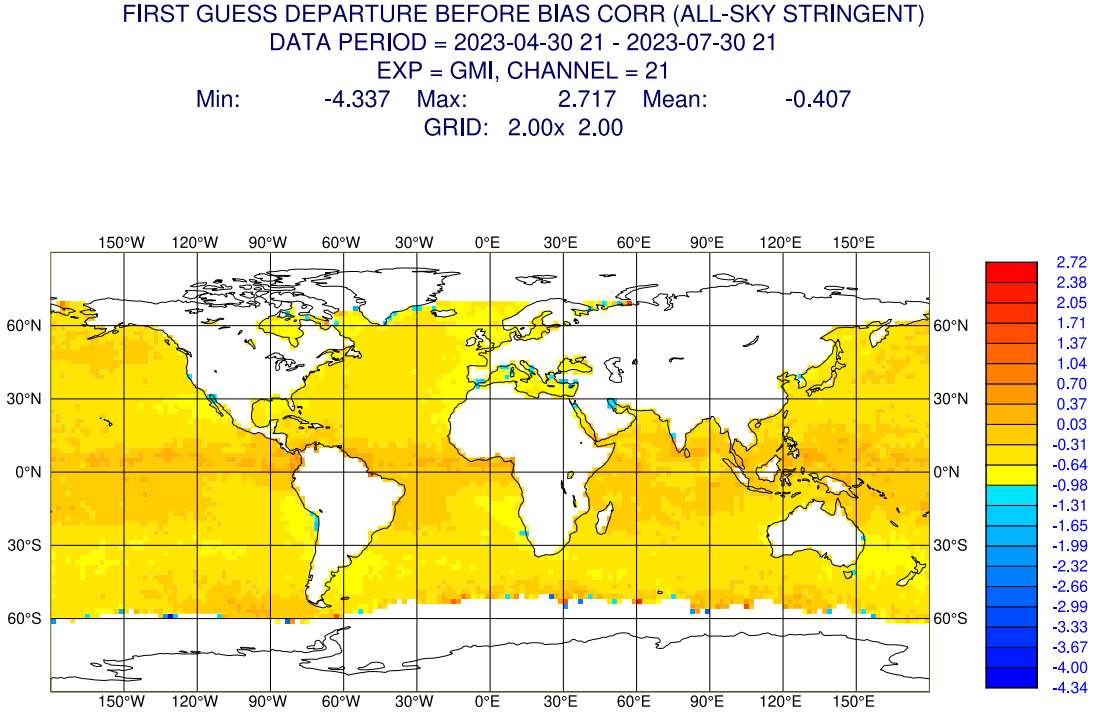


Figure 39: As Fig. 38, but the top panel shows the unified data selection whilst the bottom panel shows the stringent data selection.

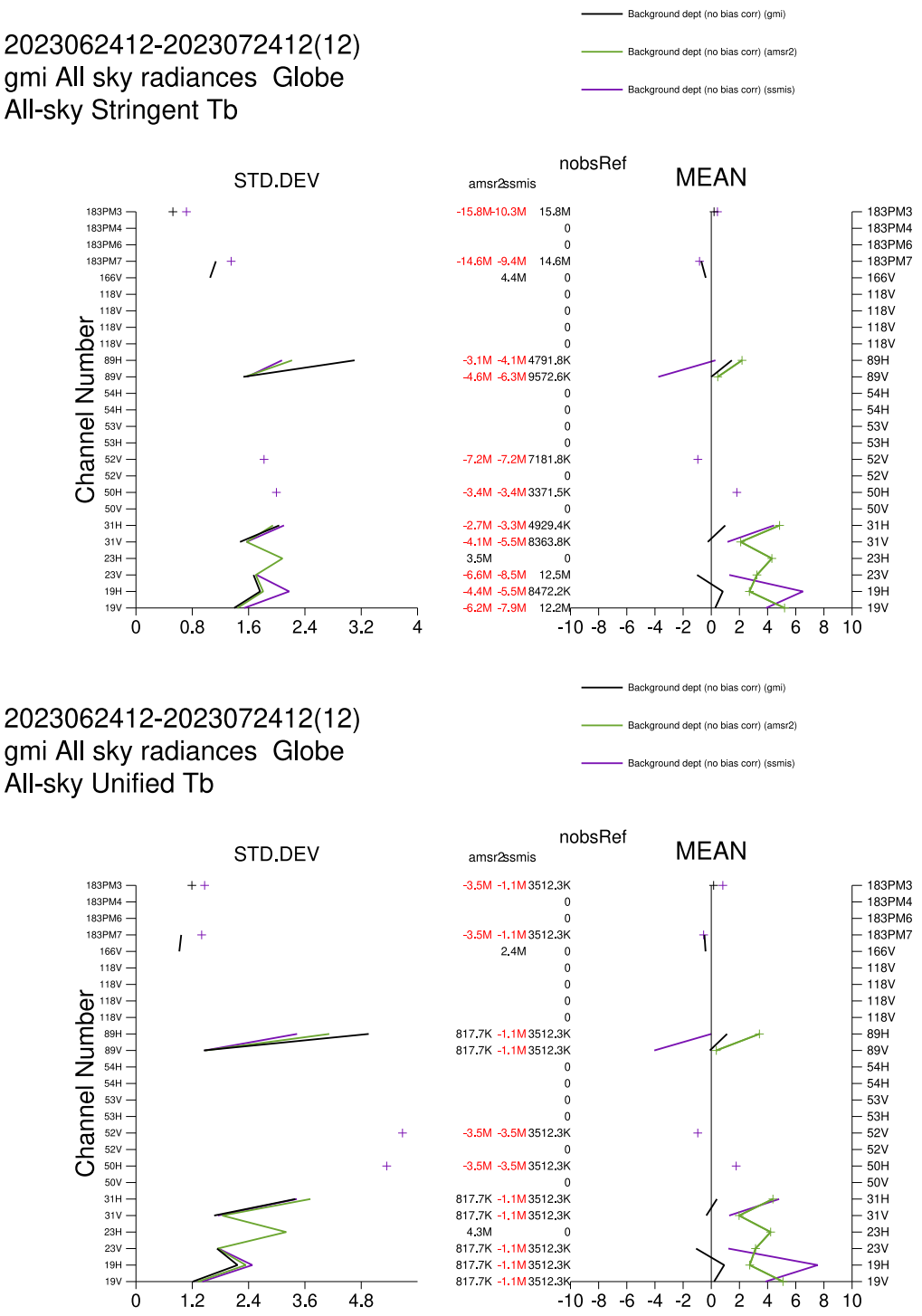


Figure 40: Inter-channel comparison for GMI (black), AMSR2 (green), and SSMIS (purple). Stringent data selection (top) and unified data selection (bottom) are shown. The y-axis is given in terms of MWIICI channel indicators. Channel counts are shown in the middle, with SSMIS as the reference.

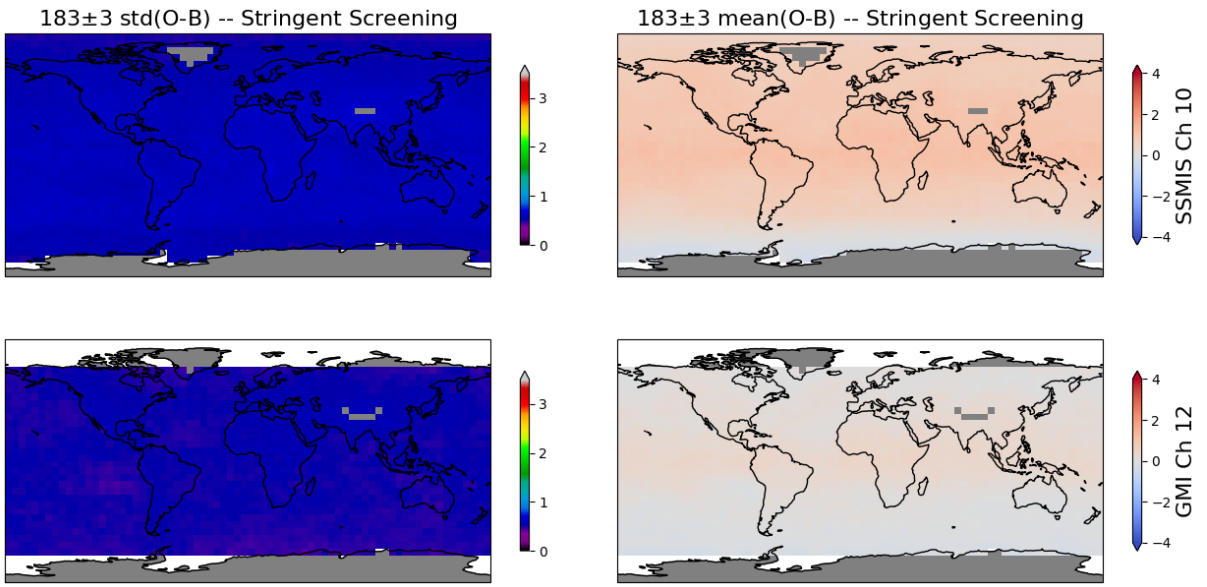


Figure 41: Stringent data selection for 183±3 channels on F17 SSMIS (top) and GMI (bottom), showing std(O-B) on the left and mean(O-B) on the right. Data are from May 2023 and averaged on a 5x5 degree grid.

Starting with the best approximation for ICI channels, Fig. 41 shows an upper-tropospheric humidity channel common to GMI and SSMIS. The std(O-B) after screening is quite low for both sensors, approaching 0.5 K. The mean biases after screening are remarkably homogeneous in the zonal direction, with variations for GMI just a few tenths of a Kelvin. SSMIS however exhibits a clear bias of about 2 K at low latitudes over the Southern Ocean. We can rule out this being a cloud bias, as it is not present in the GMI data with the same screening applied, and further analysis of biases as a function of orbital angle show that this is an orbital bias (not shown). This type of orbital bias is well-known for the SSMIS instrument, with seasonally-varying biases due to solar illumination that often manifest at low southern latitudes (Bell *et al.*, 2008; Scanlon *et al.*, 2023).

We examined the 166V channel on GMI earlier to illustrate data selection for a channel with mixed imaging and sounding capabilities. Here we use the 166H channel on GMI, as it is better matched with the sole channel in that atmospheric window on SSMIS, 150H. This comparison is found in Fig. 42, showing rather different maps of std(O-B) between the two. GMI exhibits lower variability of 166H departures in the stratocumulus regions but also at higher latitudes. Examination of the mean maps again shows GMI with a near-zero, relatively homogeneous bias against the IFS. SSMIS shows a strong dependence on latitude. Further analysis suggests that orbital biases are not at play here, whereas the bias appears correlated with tropospheric depth and/or SST. However, we can likely rule out strong surface emissivity biases here, as again we have a similar GMI channel with the same criteria applied. Although 150 and 166 GHz do have slightly different characteristics, with slightly greater water vapour and scattering sensitivity at 166 GHz, this appears to be an instrument bias for SSMIS that depends on scene temperature.

Next we examine a traditional imager frequency, the lowest measured on MWIICI. Fig. 43 shows the 19H channel from AMSR2 as well as SSMIS and GMI. The maps of departure standard deviations are quite uniform between the three sensors. Again the mean map for GMI is spatially homogeneous and close to zero, though there is perhaps a small scene dependence in the bias pattern, with positive values more likely in areas of high SSTs. This could be slight sensor bias, or could be due to emissivity model

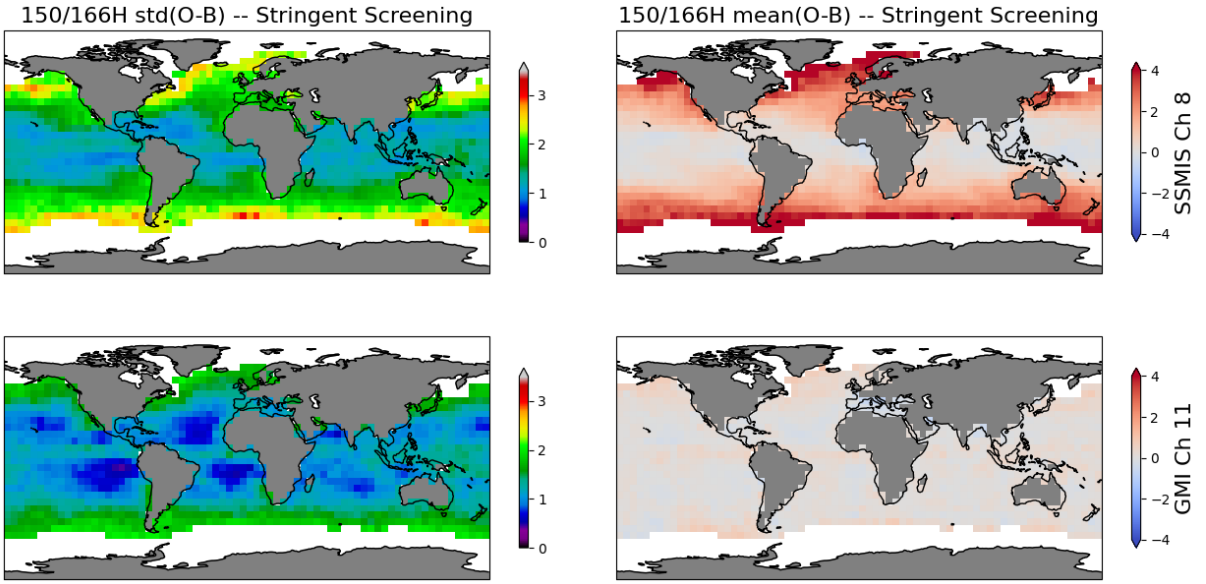


Figure 42: As Fig. 41 but showing the 150H channels on F17 SSMIS (top) and the 166H channel on GMI (bottom). Note that MWI holds a 166V rather than 166H channel, but this is the best comparison possible from these sensors.

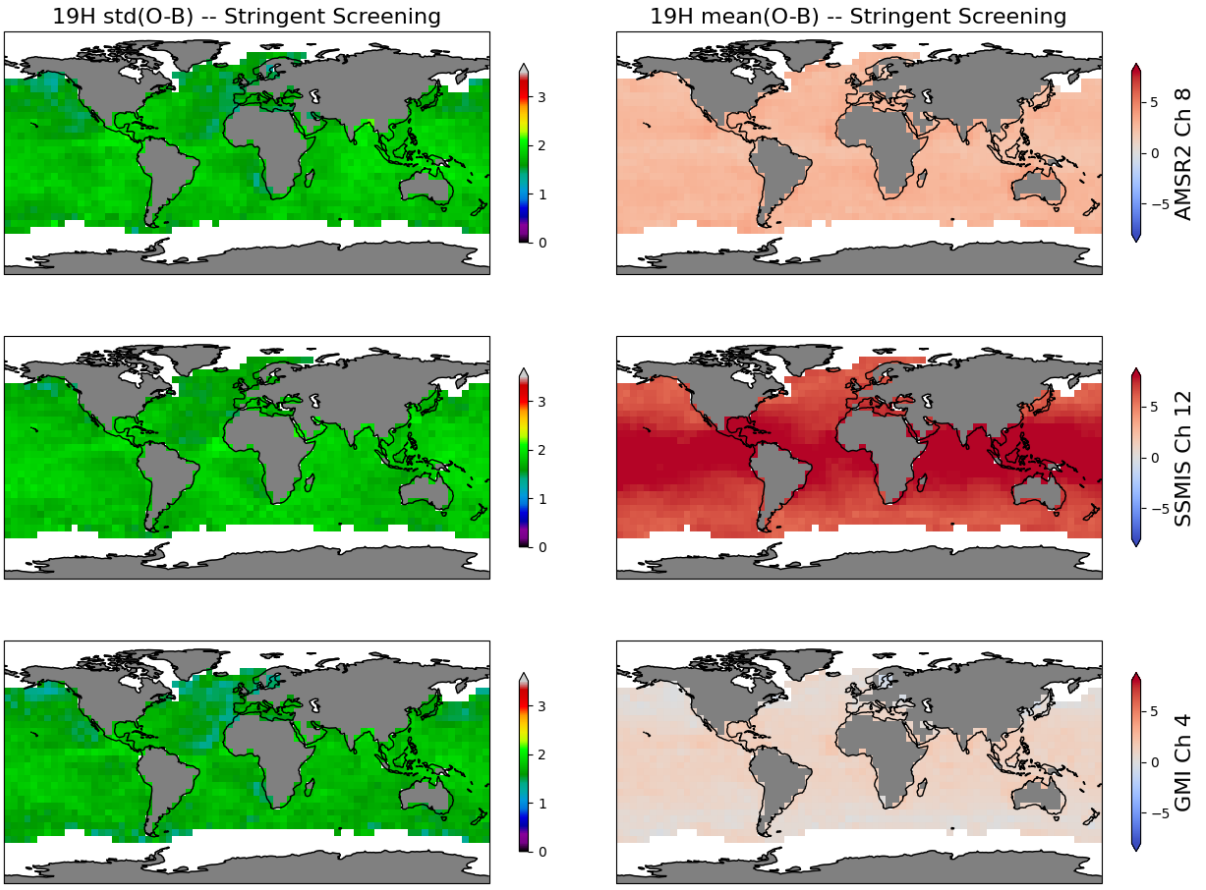


Figure 43: Stringent data selection for the 19H channels on AMSR2 (top), F17 SSMIS (middle), and GMI (bottom), showing std(O-B) on the left and mean(O-B) on the right. Data are from May 2023 and averaged on a 5x5 degree grid.

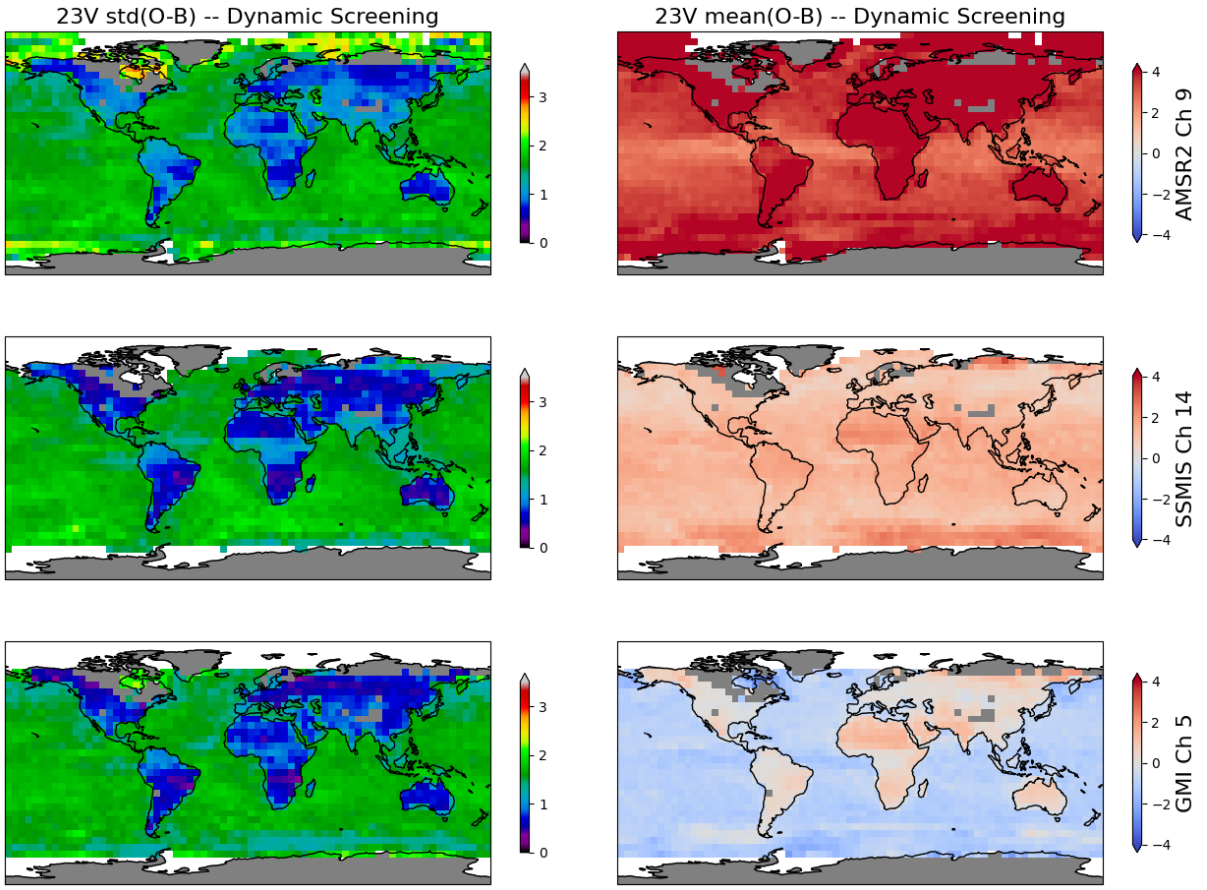


Figure 44: Dynamic data selection for the 23V channels on AMSR2 (top), F17 SSMIS (middle), and GMI (bottom), showing std(O-B) on the left and mean(O-B) on the right. Data are from May 2023 and averaged on a 5x5 degree grid.

bias or a number of other factors. AMSR2 in the top row also shows a quite flat global pattern, with a mean bias against the IFS of about 3.0 K. SSMIS 19H exhibits a larger overall bias of about 6.5 K globally, but also with more significant scene dependence than that of GMI or AMSR2.

Lastly, to show an application of the dynamic data selection, we can look at the 23V channel. Figure 44 shows that for the dynamic data selection, there are more regions filled in on the map, including areas of sea-ice and 100% land cover, with only the coldest and highest altitude land surfaces screened out. Note that AMSR2 has different sea-ice screening than SSMIS due to its additional low-frequency channels that permit its own sea-ice retrieval. For all three sensors, std(O-B) is actually lower over land, likely caused by use of the nearby 19V channel for the dynamic emissivity retrieval and cloud signals being generally smaller over land (Lonitz *et al.*, 2022, their Fig. 3). In terms of bias at 23V compared to the IFS, F17 SSMIS appears the most uniform over the globe, with a mean bias of about 1 K. GMI and AMSR2 both show significant land/sea contrast, with land departures being about 2 K warmer than ones over sea. This type of contrast is seen for various channels on each sensor, though interestingly the sign and magnitudes of these land/sea contrasts in bias are quite variable. For example, 19V departures are typically colder over land for each sensor, whilst 37 and 89 GHz departures are typically warmer (not shown). In other words, if there were a systematic skin temperature bias in the IFS for a given region, we might expect similar patterns between channels at nearby frequencies, but rather there appear to be various factors at play that likely include scene temperature-related biases. However, it is worth emphasising that O-B

biases over land are difficult to interpret due to interactions with the dynamic emissivity retrieval, and thus the results over non-ocean surfaces should be treated with caution.

The analysis in this section has focussed on the stringent data selection for a few common channels on existing MW imagers, showing that the channel-based criteria are able to distil instrument biases against the model on the order of 1 K or less from a month of data. The criteria appear to work consistently between similar channels on different sensors. Relative biases against the IFS are more difficult to interpret when including land scenes (as in the dynamic criteria) because the total range of observed TBs is much larger and thus possible scene-dependent biases are more pronounced, among other factors.

9 Conclusions

In this study, we have developed a framework for assessing and monitoring the calibration of the MWI and ICI sensors aboard Metop-SG using the power of a state-of-the-art NWP model. It relies upon the capabilities developed for all-sky assimilation at ECMWF, which have been honed and expanded over the years and applied to sensors on currently orbiting platforms that share many similarities with MWI and ICI. This framework will be made available by ECMWF to EUMETSAT on a best-efforts basis during and after the cal/val period for Metop-SG-B in order for EUMETSAT to validate cal/val requirements that are based on NWP output, and to provide a broader overview of the calibration status of the new sensors.

To analyse the calibration aspects of MWI and ICI with the IFS, several technical challenges needed to be overcome. First, we needed to create a bespoke pre-processing data flow for MWIICI, more complex than that of any other radiometer used in the ECMWF system due to the numerous unique features of the MWI and ICI instruments. This included significant effort to keep data volumes at a manageable level and parallelizing processes as much as possible. The all-sky code in the IFS was modified to accommodate the large set of channels provided by MWI and ICI, including the dual polarisations at 50GHz sounding channels and new frequencies in the sub-mm. Treating both instruments together as a super-sensor (i.e. MWIICI) permits usage of lower frequency channels to inform the surface properties for channels that may barely see the surface, and so forth. There will be future work needed to optimise aspects such as surface emissivity retrieval and orography screening for sub-mm channels in particular, but these aspects may need to wait for real data.

The all-sky cal/val sampling strategy developed in this study focused on two types of sampling: channel-based and unified. For most cal/val applications, the stricter channel-based sampling (i.e. stringent) should be preferred, as the sampling strategy is tailored to each channel's sensitivities to hydrometeors and provides a strictly limited but large data population for analysis. The looser channel-based sampling (i.e. dynamic) is intended more for interest as this will contain regions of more significant model bias, but could be useful for identifying larger temperature-dependent biases in a channel's dynamic range. The sensor-wide unified criteria is more restrictive by construction and is intended mainly for assessing inter-channel bias characteristics.

Analysis of the data selection for all-sky cal/val monitoring showed that for current sensors AMSR2, SSMIS, and GMI we can see some instrument biases that have previously been identified either in evaluations versus NWP fields or other cal/val methods, giving confidence in the approach. For instance, most AMSR2 lower frequency channels have a significant warm bias of about 3-5K (as seen in e.g. Berg *et al.* (2016)). As expected, most GMI channels exhibit a mean bias against the IFS of within about 1K (as seen in Lean *et al.* (2017)). Maps of mean departures show a relatively uniform spatial pattern for GMI channels, indicative of most model bias being effectively screened out by the selection criteria as we assume

that GMI has little scene-dependent or cross-polarisation bias (Wentz and Draper, 2016); the exception is for 37H and 89H channels due to pervasive emissivity model bias. In contrast, the same selection criteria shows significant orbital biases in equivalent SSMIS channels that have been noted elsewhere in the literature (e.g. Bell *et al.*, 2008; Booton *et al.*, 2014). This analysis also indicated a significant scene-dependent bias for the SSMIS 150H channel relative to GMI and the IFS, a bias perhaps first noted in this study but congruent with earlier difficulties to assimilate this channel in the IFS (Lonitz and Geer, 2020). The data selection criteria are however not perfect, and in particular the apparent mid-frequency bias of the SURFEM model for H-pol channels severely limits the cal/val sample for 37H and 89H. The mean bias of about -1 K between GMI 23V and the IFS should be investigated further, as this is outside the 0.25 K absolute calibration standard expected from GMI (Wentz and Draper, 2016) and a little outside the 0.8 K accuracy reported by Draper *et al.* (2015); this is roughly the magnitude of the mean change for this channel between V4 and V5 GMI L1B calibration[12], but this bias appears anomalous compared to the other GMI channels.

Some lessons have been learnt from this first hands-on experience with MWI and ICI data that may benefit future mission design:

1. Dual-polarisation capability at 166 GHz should be considered for future missions. This would help with the dynamic emissivity retrieval over land and sea-ice for higher frequency channels. Furthermore, there are polarisation signals observed at 166 GHz that are linked to oriented ice particles (Gong and Wu, 2017; Barlakas *et al.*, 2021), and thus dual-polarisation capabilities at several microwave and sub-mm frequencies could prove a useful constraint of ice microphysics.

2. There are large gaps on each side of the swath where not all channels are available (Fig. 9). This is due to the geometry of the sensors, with the placement of feedhorns leading to roughly 100 km on each side of the swath without a full observation vector of 39 channels. This leads to a narrower useful part of the swath, as some channels are needed for the assimilation of others (e.g. 50 or 89 GHz needed for land emissivity retrievals). Particularly because wider usable swaths have proven important for the total NWP impact of a sensor (e.g. comparing ATMS with AMSU-A), this should be considered for the geometry of future imagers. If it is possible to lay out the fields of view of the different feedhorns to make all channels available across the swath, this could significantly increase the impact of the instrument.

3. The large data volumes and heavy spatial oversampling of MWI and ICI will be a challenge for some future users of this data—in this project it became clear how much superobbing (or some other form of remapping) can reduce data volumes (see Sec. 5.4). As was done for AMSR2 (Maeda *et al.*, 2016), EUMETSAT should consider adding a "Level 1R" product in which TBs have been remapped to a common footprint and geolocation. With smaller data volumes and fields, this should greatly facilitate uptake of the these instruments by the broader community, such as in university research groups and in weather centres with less expertise in satellite data. Hence a level 1R product could significantly increase the total users of the data. At the same time, the sophisticated and comprehensive information provided in the Level 1B radiances is still of crucial importance in many areas: to support the current cal/val project, enabling the diagnosis of any potential instrument problems, supporting the most accurate application-specific remapping of the data (such as done in the superobbing at ECMWF), and enabling maximum future exploitation of the data.

[12] https://www.eorc.jaxa.jp/GPM/doc/product/GPM_Product_List_V5_20171004.pdf. N.B. GPM products are at V7 as of late 2023, but L1 radiances are unchanged since V5.

From the cal/val standpoint, lessons have also been learnt in this project. The symmetric screening for cloud impact using all-sky simulated radiances is argued to provide a more balanced sample for cal/val than screening out observed cloud alone. This symmetric screening strategy was indeed found to be robust using *CI*, providing consistent results between sensors. Second, one of the more ambitious aims of this project was to consider cloud-affected radiances for cal/val analysis, but there was not time to do this. Instead, the scenes with significant radiative impacts from cloud and precipitation are screened out of the cal/val sample in a similar manner to previous studies, even for the dynamic range criteria. Possible inclusion of the full range of cloudy radiances should be considered in future studies, particularly for sub-mm frequencies; this would help to provide a fuller sampling of the dynamic range, particularly for frequencies that do not see the surface. However, such analysis needs to be targeted, as it could be helpful for bias characteristics as a function of scene temperature but not for assessing sensor noise. Third, as seen in analysis of the dynamic range sample, it remains challenging to interpret instrument biases over land using NWP fields. This is confounded by the interaction of apparent biases with the dynamic emissivity retrieval, and further work will be needed to untangle these interactions to truly nail down the warm end calibration bias of window channels. Lastly, the comparisons of current sensors have proven how helpful it is to have a measurements from a reference-calibre sensor like GMI available.

The eventual assimilation of MWIICI radiances in the IFS has benefited from the technical work accomplished in this project. It is expected that most of the currently assimilated channels on MWI that are common to other MW imagers and sounders will be ready for assimilation early in the lifetime of Metop-SG-B, depending of course on the data quality. Observation error models and quality control will need fine-tuning once real data are available, but this is standard for all new all-sky sensors in the IFS. Depending on the overall fit between sub-mm radiances and the IFS simulations, it may be possible to assimilate some of the sub-mm channels early in the mission lifetime as well, but it remains to be seen which observation error model or models suit these channels best. Furthermore, full utilisation of MWI-ICI in the assimilation will certainly require consideration of inter-channel error correlations, particularly for sub-mm channels but also for the eight channels centred at 183 GHz. These topics were outside the scope of this project and will be key to unlocking the full potential impact of ICI in particular.

This report shows how the performance of novel sensors such as MWI and ICI can be evaluated against a high-quality NWP background within an all-sky assimilation framework. Furthermore, the preparations for Metop-SG data made in this project should ensure that the IFS is in a good position to achieve positive NWP impact from the MWI and ICI sensors early in their operational lifetime.

A Appendix: Applicable EUMETSAT internal documents

This appendix lists the relevant EUMETSAT internal documents for this study:

- EPS-SG End User Requirements Document, EUM/PEPS/REQ/09/0151
- MWI Calibration and Validation Plan, EUM/LEO-EPSSG/PLN/14/776068, v1F, 9 Nov 2020, 130 pp.
- ICI Calibration and Validation Plan, EUM/LEO-EPSSG/PLN/14/776069, v1D, 20 Nov 2020, 160 pp.
- EPS-SG MWI Level 1B Product Format Specification, EUM/LEO-EPSSG/SPE/14/767115, v4, 5 Sept 2022, 129 pp.
- EPS-SG ICI Level 1B Product Format Specification, EUM/LEO-EPSSG/SPE/14/771723, v4, 5 Sept 2022, 112 pp.

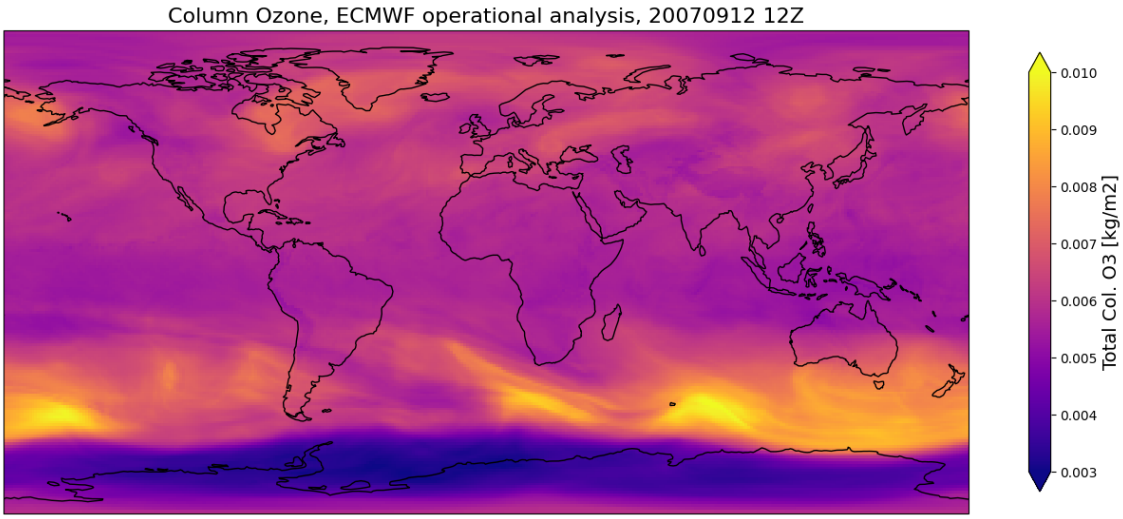


Figure 45: Column-integrated ozone from the operational ECMWF analysis on 12th Sept., 2007.

B Appendix: Ozone Impact

Whereas traditional microwave channels are not significantly affected by emission from atmospheric ozone, it is known that atmospheric ozone is not negligible at the higher frequencies of ICI (Mattioli *et al.*, 2019). To examine this, 3D fields of ozone are passed through the IFS into RTTOV to compare the effect of using ozone fields from the model versus the default profile of ozone from RTTOV (Matricardi, 2008). This analysis is possible because the RTTOV coefficient file for MWIICI (as provided by the NWP-SAF) includes ozone as a variable species; it is a fixed species for all other RTTOV microwave sensors currently assimilated. This analysis is thus the first to examine the effect of using variable ozone profiles on microwave and sub-mm radiances in the IFS. Figure 45 shows the ECMWF operational analysis of total column ozone on the day of the EUMETSAT test data for Metop-SG, September 12th 2007[13].

For frequencies of 118 GHz and below, the impact of variable ozone can be considered negligible with a mean difference of less than 0.01K. Figure 46 shows the mean global impact on a select set of MWIICI channels from the test data observation locations in the 12Z LWDA cycle. The differences shown are from matched points in two experiments, one using the RTTOV default ozone and one with variable ozone from the IFS used as input to RTTOV. The model background is identical in both experiments. The channels with the largest impact from using variable ozone are those at 664 GHz, though it is clear that most sub-mm channels experience signals in the tenths of a degree.

As can be seen in Figures 47 and 48, there is significant regional variability in the sign and magnitude of the difference in simulated TB between the two experiments, with the patterns matching those of analysed column ozone seen in Fig. 45. The first set of plots shows three frequencies within the microwave range. Interestingly, of the frequencies currently assimilated in the IFS, 166 GHz is shows the largest effect from simulating variable ozone in the radiative transfer. If the size of this signal is genuine, then it should be possible to test the veracity of this impact on currently available observations from GMI and ATMS (i.e. compare O-Bs of GMI channels 10 and 11, and ATMS channel 16, with and without variable ozone). The impact at all 183 GHz channels is quite small and could be considered negligible for sensors such as

[13]Coincidentally, the date of the test data was near the peak of the southern hemisphere's ozone hole in 2007. Thus the signals simulated here are expected to be on the more extreme end of possible behaviour.

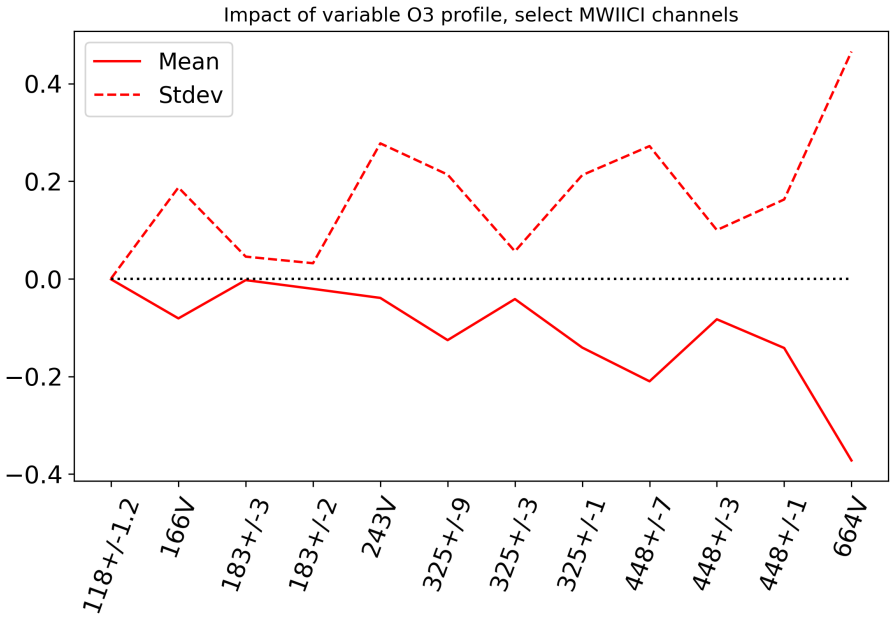


Figure 46: Impact on select MWIICI channels from using variable ozone from the IFS instead of the default RTTOV profile of ozone. These include all data from the two test orbits in the 12Z cycle.

MHS, with differences mostly in the hundredths of a Kelvin. At 243 GHz there is a non-negligible impact at higher latitudes, with more moderate impacts in the tropics where this frequency is primarily sensitive to upper tropospheric water vapour. In the sub-mm channels shown, there is a consistent regional pattern for the 325, 448, and 664 GHz channel complexes. The magnitude of the impact varies between and within these channel groups, with the largest effects seen at 664 GHz.

It is clear that assimilation of sub-mm channels without accounting for variations in ozone profiles would cause significant regional biases in increments, varying in time and space, as these patterns would not map perfectly onto current bias correction parameters. Ozone sensitivity should be considered in the radiative transfer of all sub-mm sensors, particularly ICI but also including the upcoming Arctic Weather Satellite, expected to launch in 2024. Future work will investigate whether it is beneficial for simulation and assimilation of observed microwave radiances (e.g. at 166 GHz) to consider variable ozone profiles in the radiative transfer simulation. Although variable ozone was not considered in the simulation of MWIICI radiances in this study, this was only due to the test data neglecting ozone sensitivity. Variable ozone should be included in the radiative transfer simulations used for the operational monitoring of MWIICI calibration performance.

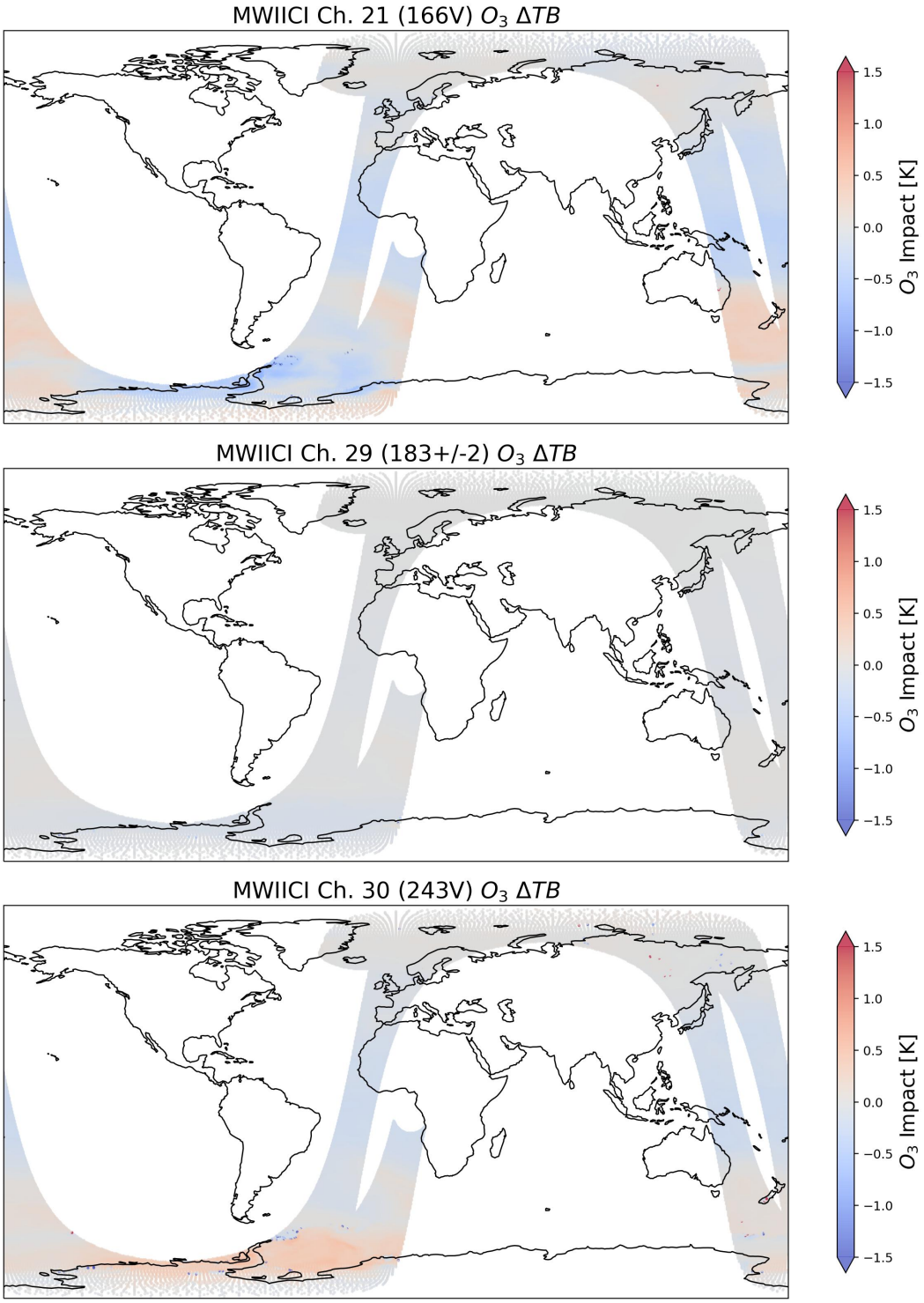


Figure 47: Impact on select MWIICI channels (as noted) from using variable ozone from the IFS instead of the default RTTOV profile of ozone, 12Z cycle on 20070912.

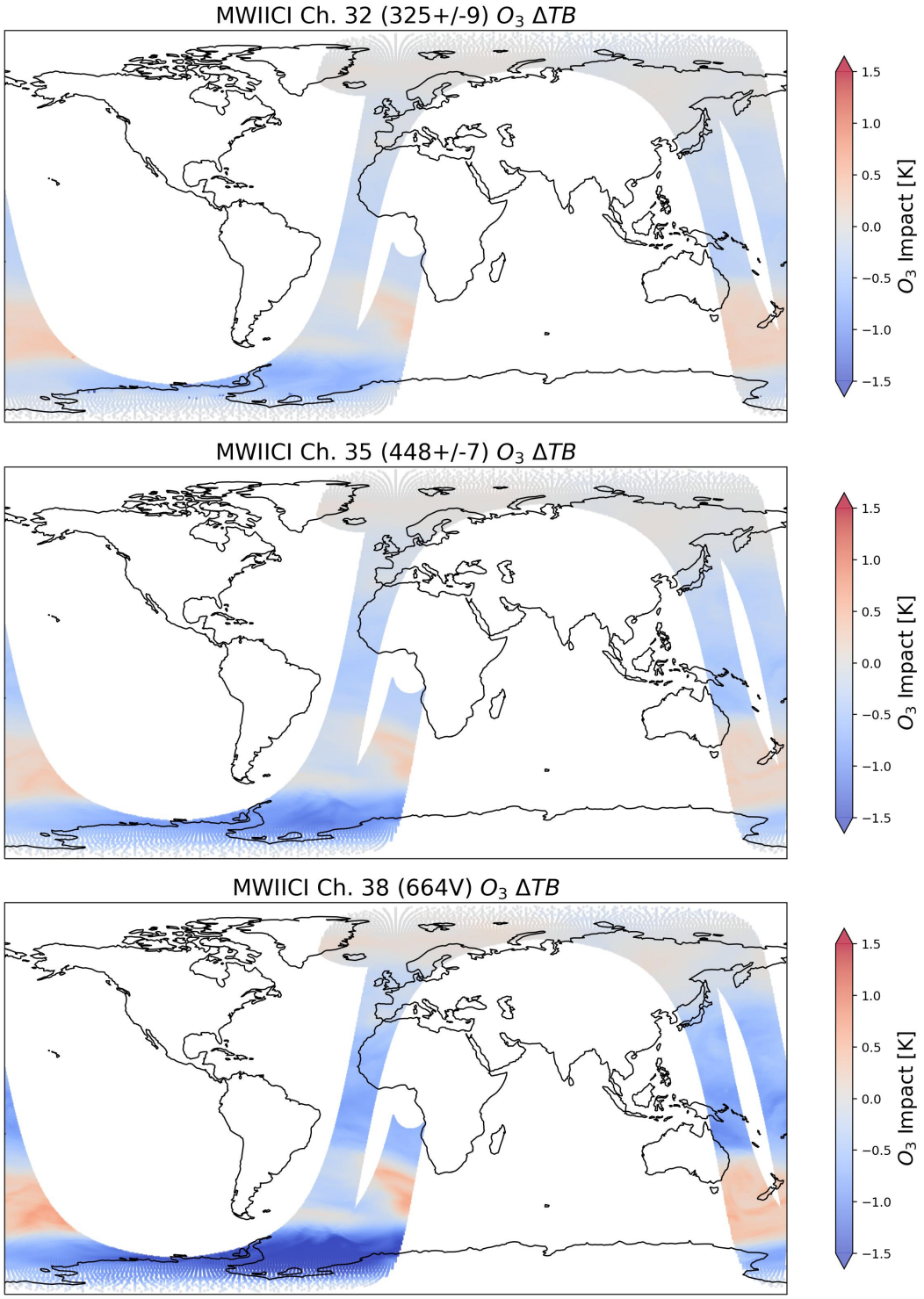


Figure 48: Impact on select MWIICI channels (as noted) from using variable ozone from the IFS instead of the default RTTOV profile of ozone, 12Z cycle on 20070912.

C Appendix: BUFR Sequences

Here the BUFR sequences for level 1 MWI and ICI data are given for reference. First are the WMO-approved sequences for ICI (310080) and MWI (310081). Second is the ECMWF-defined sequence for internal use; the sequence number 310199 is nominal and will change (it is not registered with the WMO). In each case, the numbers have been grouped to make them more readable.

310080 =

[001033,001034,001007,002019,005040,025061,201133,005041,201000,301011,301012,207003,004006,207000,202126,007001,202000,005063,
005064,005066,005077,006043,006041,006042,008098,012101,008098,012101,008098,012101,008098,012101,008098,007074,005076,201124,
202127,025084,202000,201000,033100,033102,033104,033105,113007,005078,005001,006001,007024,005021,007025,005022,007073,201132,
202129,021166,202000,201000,
117003,005042,202131,002153,002154,202000,025077,025078,033101,033103,002104,201131,202136,014045,202000,201000,012158,012159,
117002,005042,202131,002153,002154,202000,025077,025078,033101,033103,002104,201132,202136,014045,202000,201000,012158,012159,
117003,005042,202131,002153,002154,202000,025077,025078,033101,033103,002104,201133,202136,014045,202000,201000,012158,012159,
117003,005042,202131,002153,002154,202000,025077,025078,033101,033103,002104,201133,202136,014045,202000,201000,012158,012159,

117002,005042,202131,002153,002154,202000,025077,025078,033101,033103,002104,201131,202135,014045,202000,201000,012158,012159]

310081 =

[001033,001034,001007,002019,005040,025061,201133,005041,201000,301011,301012,207003,004006,207000,202126,007001,202000,005063,
005064,005066,005077,006043,006041,006042,008098,012101,008098,012101,008098,012101,008098,012101,008098,007074,005076,201124,
202127,025084,202000,201000,033095,033097,033106,033099,113008,005078,005001,006001,007024,005021,007025,005022,007073,201132,
202129,021166,202000,201000,
117002,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201131,202138,014045,202000,201000,012158,012159,
117002,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201132,202138,014045,202000,201000,012158,012159,
117002,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201132,202138,014045,202000,201000,012158,012159,
117008,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201131,202137,014045,202000,201000,012158,012159,
117002,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201132,202137,014045,202000,201000,012158,012159,
117004,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201133,202137,014045,202000,201000,012158,012159,
117001,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201131,202136,014045,202000,201000,012158,012159,
117005,005042,202131,002153,002154,202000,025077,025078,033096,033098,002104,201131,202136,014045,202000,201000,012158,012159]

Below is the BUFR sequence derived for processing in the IFS, covering the combined super-sensor MWIICI (see Sec. 4.1). This sequence uses a delayed replication factor (39 for number of channels) that is not specified in the sequence itself but indicated by the 031001 operator element.

310199 =

[001033, 001034, 001007, 002019, 005040, 025061, 201133, 005041, 201000, 301011, 301012, 207003, 004006, 207000, 202126, 007001, 202000,
005063, 005064, 005066, 005077, 006043, 006041, 006042,
008098,012101,008098,012101,008098,012101,008098,012101,008098,012101,008098,012101,008098,012101,008098,012101, 008098,
007074,005076,201124,202127,025084,202000,201000, 033095,033097,033106,033099, 033100,033102,033104,033105,
201131,005043,201000, 005001,006001, 007073,
128000, 031001,
005042, 202131,002153,002154,202000,025077,025078,033096,033098,033101,033103,002104,

012163, 201132,202129,012065,202000,201000, 201130,008049,201000, 012158,012159, 007024,005021,007025,005022,021166]

Glossary

AMSR2 Advanced Microwave Scanning Radiometer 2. 7, 17, 18

AMSU-A Advanced Microwave Sounding Unit-A. 16, 23

ATMS Advanced Technology Microwave Sounder. 23

BUFR Binary Universal Form for the Representation of meteorological data. 18, 77

CAO Cold air outbreak. 22, 49

DMSP Defense Meteorological Satellite Program (USA). 18

ECMWF European Centre for Medium-range Weather Forecasts. 15

EPS EUMETSAT Polar System. 5

EURD End User Requirements Document. 9

FASTEM Fast ocean microwave emissivity model. 14

FOV Field of View. 6

FY-3 Feng-Yun 3. 18

GCOM-W Global Change Observation Mission for Water. 7, 18

GMI Global Precipitation Measurement (GPM) Microwave Imager. 7, 11, 18

GPM Global Precipitation Measurement. 7, 18

ICI Ice Cloud Imager. 5

IFS Integrated Forecasting System. 15

IWP Ice Water Path. 11

LWP Liquid Water Path. 50

MARS Meteorological Archival and Retrieval System (at ECMWF). 21

Metop-SG Metop Second Generation. 5

MWHS-2 MicroWave Humidity Sounder 2. 18, 23

MWI MicroWave Imager. 5

MWRI MicroWave Radiation Imager. 18

NEDT Noise Equivalent Differential Temperature. 6

NWP-SAF Numerical Weather Prediction Satellite Application Facility. 73

ODB Observational DataBase. 18

P37 37 GHz polarisation difference. 21

PARMIO Passive and Active Reference Microwave to Infrared Ocean. 14

PFM Proto Flight Model. 7, 8

RTTOV-SCATT Radiative transfer for TOVS microwave scattering package. 14, 16

SI Scattering Index. 21

SRF Spectral Response Function. 35

SSMI Special Sensor Microwave Imager. 8, 15

SSMIS Special Sensor Microwave Imager Sounder. 11, 18

SST Sea surface temperature. 22

TB Brightness Temperature. 11

TCWV Total column water vapour. 21

TESSEM2 Tool to Estimate Sea-Surface Emissivity from Microwaves to sub-Millimeter waves. 14

TMS TROPICS Microwave Sounder. 18

TROPICS Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats. 18

References

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. Meteor. Soc.*, **142**(700), 2854–2866, doi:10.1002/qj.2873, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2873.

Barlakas, V., Galligani, V. S., Geer, A. J. and Eriksson, P. (2022). On the accuracy of RTTOV-SCATT for radiative transfer at all-sky microwave and submillimeter frequencies. *J. Quant. Spectrosc. Radiat. Transfer*, **283**, 108137, doi:https://doi.org/10.1016/j.jqsrt.2022.108137, URL https://www.sciencedirect.com/science/article/pii/S0022407322000747.

Barlakas, V., Geer, A. J. and Eriksson, P. (2021). Introducing hydrometeor orientation into all-sky microwave and submillimeter assimilation. *Atmos. Meas. Tech.*, **14**(5), 3427–3447, doi:10.5194/amt-14-3427-2021, URL https://amt.copernicus.org/articles/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. Meteor. Soc.*, **136**(652), 1868–1885, doi:10.1002/qj.659, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.659.

Bell, W., English, S. J., Candy, B., Atkinson, N., Hilton, F., Baker, N., Swadley, S. D., Campbell, W. F., Bormann, N., Kelly, G. and Kazumori, M. (2008). The assimilation of SSMIS radiances in numerical weather prediction models. *IEEE T. Geosci. Remote Sens.*, **46**(4), 884–900, doi:10.1109/TGRS.2008.917335, URL https://ieeexplore.ieee.org/abstract/document/4468717.

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., Wilheit, T. and Yang, J. X. (2016). Intercalibration of the GPM microwave radiometer constellation. *J. Atmos. Oceanic Technol.*, **33**(12), 2639–2654, doi:10.1175/JTECH-D-16-0100.1, URL https://journals.ametsoc.org/view/journals/atot/33/12/jtech-d-16-0100_1.xml.

Berg, W., Sapiano, M. R. P., Horsman, J. and Kummerow, C. (2013). Improved geolocation and earth incidence angle information for a fundamental climate data record of the SSM/I sensors. *IEEE T. Geosci. Remote Sens.*, doi:10.1109/TGRS.2012.2199761, URL https://ieeexplore.ieee.org/document/6239586.

Booton, A., Bell, W. and Atkinson, N. (2014). An improved bias correction for SSMIS. In *Proceedings of the International TOVS Study Conference ITSC-19*, URL https://www-cdn.eumetsat.int/files/2020-04/pdf_conf_p_s8_01_booton_v.pdf.

Bormann, N. (2022). Accounting for Lambertian reflection in the assimilation of microwave sounding radiances over snow and sea-ice. *Quart. J. Roy. Meteor. Soc.*, **148**(747), 2796–2813, doi:https://doi.org/10.1002/qj.4337, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4337.

Bormann, N., Magnusson, L., Duncan, D. and Dahoui, M. (2023). Characterisation and correction of orbital biases in AMSU-A and ATMS observations in the ECMWF system. *Technical Report 912*, ECMWF Tech. Memo., doi:10.21957/d281dc221a, URL https://www.ecmwf.int/en/elibrary/81520-characterisation-and-correction-orbital-biases-amsu-and-atms-observatio

Dee, D. P. (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, ECMWF, Reading, UK, URL https://www.ecmwf.int/sites/default/files/elibrary/2004/74143-variational-bias-correction-radiance-data-ecmwf-system_0.pdf.

Dinnat, E., English, S., Prigent, C., Kilic, L., Anguelova, M., Newman, S., Meissner, T., Boutin, J., Stoffelen, A., Yueh, S., Johnson, B., Weng, F. and Jimenez, C. (2023). PARMIO: A reference quality model for ocean surface emissivity and backscatter from the microwave to the infrared. *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-23-0023.1, URL https://journals.ametsoc.org/view/journals/bams/aop/BAMS-D-23-0023.1/BAMS-D-23-0023.1.xml.

Draper, D. W., Newell, D. A., Wentz, F. J., Krimchansky, S. and Skofronick-Jackson, G. M. (2015). The Global Precipitation Measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Top. Appl. Rem. Sens.*, **8**(7), 3452–3462, doi:10.1109/JSTARS.2015.2403303, URL https://ieeexplore.ieee.org/document/7052302.

Duncan, D. I., Bormann, N. and Geer, A. J. (2022a). All-sky assimilation of AMSU-A window channels. *Technical Report 59*, EUMETSAT/ECMWF Fellowship Programme Research Report, Shinfield Park, Reading, doi:10.21957/daefm16p8, URL https://www.ecmwf.int/node/20457.

Duncan, D. I., Bormann, N., Geer, A. J. and Weston, P. (2022b). Assimilation of AMSU-A in all-sky conditions. *Mon. Weather Rev.*, **150**(5), 1023 – 1041, doi:10.1175/MWR-D-21-0273.1, URL https://doi.org/10.1175/MWR-D-21-0273.1.

Duncan, D. I., Bormann, N., Geer, A. J. and Weston, P. (2023). Superobbing and finer thinning for all-sky humidity sounder assimilation. *Technical Report 905*, ECMWF Tech. Memo., doi:10.21957/5c3b9c8d9f, URL https://www.ecmwf.int/en/elibrary/81345-superobbing-and-finer-thinning-all-sky-humidity-sounder-assimilation.

ECMWF (2023). *IFS Documentation CY48R1 - Part II: Data Assimilation*, chapter 2, p. 108. 2, ECMWF, doi:10.21957/a744f32e74, URL https://www.ecmwf.int/en/elibrary/81368-ifs-documentation-cy48r1-part-ii-data-assimilation.

English, S., Prigent, C., Johnson, B., Yueh, S., Dinnat, E., Boutin, J., Newman, S., Anguelova, M., Meissner, T., Kazumori, M., Weng, F., Supply, A., Kilic, L., Bettenhausen, M., Stoffelen, A. and Accadia, C. (2020). Reference-quality emission and backscatter modeling for the ocean. *Bull. Amer. Meteor. Soc.*, **101**(10), E1593 – E1601, doi:10.1175/BAMS-D-20-0085.1, URL https://journals.ametsoc.org/view/journals/bams/101/10/bamsD200085.xml.

English, S. J. (2008). The importance of accurate skin temperature in assimilating radiances from satellite sounding instruments. *IEEE T. Geosci. Remote Sens.*, **46**(2), 403–408, doi:10.1109/TGRS.2007.902413, URL https://ieeexplore.ieee.org/document/6239586.

Eriksson, P., Rydberg, B., Mattioli, V., Thoss, A., Accadia, C., Klein, U. and Buehler, S. A. (2020). Towards an operational Ice Cloud Imager (ICI) retrieval product. *Atmos. Meas. Tech.*, **13**(1), 53–71, doi:10.5194/amt-13-53-2020, URL https://amt.copernicus.org/articles/13/53/2020/.

Field, P. R., Heymsfield, A. J. and Bansemer, A. (2007). Snow Size Distribution Parameterization for Midlatitude and Tropical Ice Clouds. *J. Atmos. Sci.*, **64**(12), 4346–4365, doi:10.1175/2007jas2344.1, URL https://doi.org/10.1175/2007JAS2344.1.

Forbes, R., Geer, A. J., Lonitz, K. and Ahlgrimm, M. (2016). Reducing systematic errors in cold-air outbreaks. *ECMWF Newsletter*, pp. 17–22, doi:10.21957/s41h7q7l, URL https://www.ecmwf.int/node/17261.

Geer, A., Lupu, C., Duncan, D., Bormann, N. and English, S. (2024). SURFEM-ocean microwave surface emissivity evaluated. *Technical Report 915*, ECMWF Tech. Memo., doi:10.21957/0af49d82e2, URL https://www.ecmwf.int/en/elibrary/81550-surfem-ocean-microwave-surface-emissivity-evaluated.

Geer, A. J. (2021). Physical characteristics of frozen hydrometeors inferred with parameter estimation. *Atmos. Meas. Tech.*, **14**(8), 5369–5395, doi:10.5194/amt-14-5369-2021, URL https://doi.org/10.5194/amt-14-5369-2021.

Geer, A. J., Baordo, F., Bormann, N. and English, S. (2014). All-sky assimilation of microwave humidity sounders. *Technical Report 741*, ECMWF Tech. Memo., doi:10.21957/obosmx154, URL https://www.ecmwf.int/node/9507.

Geer, A. J. and Bauer, P. (2010). Enhanced use of all-sky microwave observations sensitive to water vapour, cloud and precipitation. *Technical Report 20*, EUMETSAT/ECMWF Fellowship Programme Research Report, URL https://www.ecmwf.int/node/9506, also published as ECMWF Technical Memorandum No.620.

Geer, A. J. and Bauer, P. (2011). Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**(661), 2024–2037, doi:10.1002/qj.830, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.830.

Geer, A. J., Bauer, P. and Bormann, N. (2010). Solar biases in microwave imager observations assimilated at ECMWF. *IEEE T. Geosci. Remote Sens.*, **48**(6), 2660–2669, doi:10.1109/TGRS.2010.2040186, URL https://ieeexplore.ieee.org/document/5424018.

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. *Geosci. Model Dev.*, **14**(12), 7497–7526, doi:10.5194/gmd-14-7497-2021, URL https://gmd.copernicus.org/articles/14/7497/2021/.

Geer, A. J., Lonitz, K., Duncan, D. I. and Bormann, N. (2022). Improved surface treatment for all-sky microwave observations. *Technical Report 894*, ECMWF Tech. Memo., Shinfield Park, Reading, doi:10.21957/zi7q6hau, URL https://www.ecmwf.int/node/20337.

Geer, A. J., Lonitz, K., Weston, P., Kazumori, M., Okamoto, K., Zhu, Y., Liu, E. H., Collard, A., Bell, W., Migliorini, S., Chambon, P., Fourrié, N., Kim, M.-J., Köpken-Watts, C. and Schraff, C. (2018). All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**(713), 1191–1217, doi:10.1002/qj.3202, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3202.

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, doi:10.5194/acp-17-2741-2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy,

S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. (2020). The era5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**(730), 1999–2049, doi:10.1002/qj.3803, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803.

Hollinger, J., Peirce, J. and Poe, G. (1990). SSM/I instrument evaluation. *IEEE T. Geosci. Remote Sens.*, **28**(5), 781–790, doi:10.1109/36.58964, URL https://ieeexplore.ieee.org/document/58964.

Kazumori, M. and English, S. J. (2015). Use of the ocean surface wind direction signal in microwave radiance assimilation. *Quart. J. Roy. Meteor. Soc.*, **141**(689), 1354–1375, doi:10.1002/qj.2445, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2445.

Kilic, L., Prigent, C., Jimenez, C., Turner, E., Hocking, J., English, S., Meissner, T. and Dinnat, E. (2023). Development of the SURface Fast Emissivity Model for Ocean (SURFEM-Ocean) based on the PARMIO radiative transfer model. *Earth Space Sci.*, **10**(11), e2022EA002785, doi:10.1029/2022EA002785, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022EA002785.

Klaver, R., Haarsma, R., Vidale, P. L. and Hazeleger, W. (2020). Effective resolution in high resolution global atmospheric models for climate studies. *Atmos. Sci. Lett.*, **21**(4), e952, doi:10.1002/asl.952, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.952.

Kunkee, D. B., Poe, G. A., Boucher, D. J., Swadley, S. D., Hong, Y., Wessel, J. E. and Uliana, E. A. (2008). Design and evaluation of the first Special Sensor Microwave Imager/Sounder. *IEEE T. Geosci. Remote Sens.*, doi:10.1109/TGRS.2008.917980, URL https://ieeexplore.ieee.org/abstract/document/4475705.

Lawrence, H., Bormann, N., Geer, A. J., Lu, Q. and English, S. J. (2018). Evaluation and assimilation of the microwave sounder MWHS-2 onboard FY-3C in the ECMWF numerical weather prediction system. *IEEE T. Geosci. Remote Sens.*, **56**(6), 3333–3349, doi:10.1109/TGRS.2018.2798292, URL https://doi.org/10.1109/TGRS.2018.2798292.

Lawrence, H., Carminati, F., Bell, W., Bormann, N., Newman, S., Atkinson, N., Geer, A., Migliorini, S., Lu, Q. and Chen, K. (2017). An evaluation of FY-3C MWRI and assessment of the long-term quality of FY-3C MWHS-2 at ECMWF and the Met Office. *Technical Report 798*, ECMWF Tech. Memo., doi:10.21957/lhuph6fb3, URL https://www.ecmwf.int/node/17206.

Lean, P., Geer, A. and Lonitz, K. (2017). Assimilation of Global Precipitation Mission (GPM) Microwave Imager (GMI) in all-sky conditions. *Technical Report 799*, ECMWF Tech. Memo., doi:10.21957/8orc7sn33, URL https://www.ecmwf.int/node/17174.

Lonitz, K. and Geer, A. J. (2015). New screening of cold-air outbreak regions used in 4D-Var all-sky assimilation. *Technical Report 35*, EUMETSAT/ECMWF Fellowship Programme Research Report, Shinfield Park, Reading, URL https://www.ecmwf.int/node/10777.

Lonitz, K. and Geer, A. J. (2020). Reducing the drying effect through a water vapour correction to the all-sky error model. *Technical Report 53*, EUMETSAT/ECMWF Fellowship Programme Research Report, doi:10.21957/qmy8utbgb, URL https://www.ecmwf.int/node/19528.

Lonitz, K., Geer, A. J. and Bormann, N. (2022). Towards assimilating surface sensitive microwave channels over land. *Technical Report 58*, EUMETSAT/ECMWF Fellowship Programme Research Report, Shinfield Park, Reading, doi:10.21957/gdwqzfns, URL https://www.ecmwf.int/node/20334.

Lu, Q., Lawrence, H., Bormann, N., English, S., Lean, K., Atkinson, N., Bell, W. and Carminati, F. (2015). An evaluation of FY-3C satellite data quality at ECMWF and the Met Office. *Technical Report 767*, ECMWF Tech. Memo., doi:10.21957/3l7g9nqun, URL https://www.ecmwf.int/node/14692.

Maeda, T., Taniguchi, Y. and Imaoka, K. (2016). GCOM-W1 AMSR2 level 1R product: Dataset of brightness temperature modified using the antenna pattern matching technique. *IEEE T. Geosci. Remote Sens.*, **54**(2), 770–782, doi:10.1109/TGRS.2015.2465170.

Matricardi, M. (2008). The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database. *Technical Report 564*, ECMWF Research Dept. Tech. Memo., URL https://www.ecmwf.int/sites/default/files/elibrary/2008/11040-generation-rttov-regression-coefficients-iasi-and-airs-using-new-profil pdf.

Mattioli, V., Accadia, C., Prigent, C., Crewell, S., Geer, A., Eriksson, P., Fox, S., Pardo, J. R., Mlawer, E. J., Cadeddu, M., Bremer, M., Breuck, C. D., Smette, A., Cimini, D., Turner, E., Mech, M., Marzano, F. S., Brunel, P., Vidot, J., Bennartz, R., Wehr, T., Michele, S. D. and John, V. O. (2019). Atmospheric gas absorption knowledge in the submillimeter: Modeling, field measurements, and uncertainty quantification. *Bull. Amer. Meteor. Soc.*, **100**(12), ES291 – ES295, doi:10.1175/BAMS-D-19-0074.1, URL https://journals.ametsoc.org/view/journals/bams/100/12/bams-d-19-0074.1.xml.

Meissner, T. and Wentz, F. J. (2012). The emissivity of the ocean surface between 6 and 90 GHz a large range of wind speeds and earth incidence angles. *IEEE T. Geosci. Remote Sens.*, doi:10.1109/TGRS.2011.2179662, URL https://ieeexplore.ieee.org/document/6145646.

Munchak, S. J., Ringerud, S., Brucker, L., You, Y., de Gelis, I. and Prigent, C. (2020). An active–passive microwave land surface database from GPM. *IEEE T. Geosci. Remote Sens.*, **58**(9), 6224–6242, doi:10.1109/TGRS.2020.2975477, URL https://ieeexplore.ieee.org/document/9027116.

Newman, S., Carminati, F., Lawrence, H., Bormann, N., Salonen, K. and Bell, W. (2020). Assessment of new satellite missions within the framework of numerical weather prediction. *Remote Sens.*, **12**(10), doi:10.3390/rs12101580, URL https://www.mdpi.com/2072-4292/12/10/1580.

Okamoto, K., McNally, A. P. and Bell, W. (2014). Progress towards the assimilation of all-sky infrared radiances: an evaluation of cloud effects. *Quart. J. Roy. Meteor. Soc.*, **140**(682), 1603–1614, doi:10.1002/qj.2242, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2242.

Prigent, C., Aires, F., Wang, D., Fox, S. and Harlow, C. (2017). Sea-surface emissivity parametrization from microwaves to millimetre waves. *Quart. J. Roy. Meteor. Soc.*, **143**(702), 596–605, doi:10.1002/qj.2953, URL https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.2953.

Reul, N., Saux-Picart, S., Chapron, B., Vandemark, D., Tournadre, J. and Salisbury, J. (2009). Demonstration of ocean surface salinity microwave measurements from space using AMSR-E data

over the Amazon plume. *Geophys. Res. Lett.*, **36**(13), doi:10.1029/2009GL038860, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL038860.

Rydberg, B. and Eriksson, P. (2019). Backus-Gilbert footprint matching methodology applied on MWI and ICI observations. *Technical report*, Molflow AB, URL https://www.eumetsat.int/media/47983.

Scanlon, T., Geer, A. and Bormann, N. (2023). Microwave imagers in the ECMWF-IFS: Adding further observations and improving convective anvils in the observation operator. *Technical Report 61*, EUMETSAT/ECMWF Fellowship Programme Research Report, Shinfield Park, Reading, doi:10.21957/8542a37c46, URL https://www.ecmwf.int/en/elibrary/81347-microwave-imagers-ecmwf-ifs-adding-further-observations-and-improving-cor

Steele, L., Bormann, N. and Duncan, D. I. (2023). Assimilating FY-3E MWHS-2 observations, and assessing all-sky humidity sounder thinning scales. *Technical Report 62*, EUMETSAT/ECMWF Fellowship Programme Research Report, Shinfield Park, Reading, doi:10.21957/f42a9d9542, URL https://www.ecmwf.int/en/elibrary/81525-assimilating-fy-3e-mwhs-2-obs-and-assessing-all-sky-humidity-sounder-thi

Turner, E., Fox, S., Mattioli, V. and Cimini, D. (2022). Literature review on microwave and sub-millimetre spectroscopy for MetOp second generation. *Technical report*, EUMETSAT, URL https://nwp-saf.eumetsat.int/site/download/members_docs/cdop-3_reference_documents/NWPSAF_report_submm_litrev.pdf.

Wang, D., Prigent, C., Kilic, L., Fox, S., Harlow, C., Jimenez, C., Aires, F., Grassotti, C. and Karbou, F. (2017). Surface emissivity at microwaves to millimeter waves over polar regions: Parameterization and evaluation with aircraft experiments. *J. Atmos. Oceanic Technol.*, doi:10.1175/JTECH-D-16-0188.1, URL https://journals.ametsoc.org/view/journals/atot/34/5/jtech-d-16-0188.1.xml.

Wentz, F. J. (1983). A model function for ocean microwave brightness temperatures. *J. Geophys. Res. Oceans*, **88**(C3), 1892–1908, doi:10.1029/JC088iC03p01892, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC088iC03p01892.

Wentz, F. J. and Draper, D. (2016). On-orbit absolute calibration of the Global Precipitation Measurement Microwave Imager. *J. Atmos. Oceanic Technol.*, doi:10.1175/JTECH-D-15-0212.1, URL https://journals.ametsoc.org/view/journals/atot/33/7/jtech-d-15-0212_1.xml.

Wilheit, T. T. (1979). A model for the microwave emissivity of the ocean's surface as a function of wind speed. *IEEE T. Geosci. Electronics*, **17**(4), 244–249, doi:10.1109/TGE.1979.294653, URL https://ieeexplore.ieee.org/document/4072010.