EUMETSAT ECMWF

EUMETSAT Contract Report

RFQ/21/1383948

Contract Report to EUMETSAT

Assessment of the Arctic Weather Satellite in NWP

Final Report for EUMETSAT Contract RFQ/21/1383948

David I. Duncan, Niels Bormann, Mohamed Dahoui, Marijana Crepulja

September 2025

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 Arctic Weather Satellite (AWS) is an ESA mission launched on August 16th, 2024. AWS is notable for its small size, rapid development from concept to launch, and for providing the first observations of sub-millimetre wavelength channels for operational meteorology. AWS represents the first meteorological mission led by a space agency to manifest so-called "new space" concepts of increased miniaturisation, swift deployment, and independent platforms that enable relatively low-cost constellations of passive microwave (MW) radiometers. The payload of AWS is a newly developed MW instrument, which carries traditional MW sounding channels in the 50 GHz oxygen and 183 GHz humidity bands in addition to sub-mm humidity-sounding channels near 325 GHz. In addition to its stand-alone significance, AWS is also a pathfinder mission for the proposed EPS-Sterna constellation of radiometers that EUMETSAT is considering for launch in the late 2020's to augment the coverage of EPS-SG MW instruments.

This report assesses the performance of the AWS radiometer via comparison with equivalents from the ECMWF model, the Integrated Forecasting System (IFS). The vicarious approach to calibration and validation (cal/val) analysis takes the IFS as a calibration reference, permitting analysis of AWS radiometric performance in the context of similar instruments (e.g. AMSU-A and MHS) over a wide range of geophysical conditions. It is applied here in the all-sky assimilation framework, in which model clouds are simulated, but scenes with known model biases such as optically thick clouds, sea-ice, and difficult surfaces are screened out, depending on each channel's sensitivity. The driving philosophy is to maximise the data sample whilst ensuring a balanced statistical sample.

The work herein builds upon two previous EUMETSAT-funded projects, both in scientific terms and on a technical level. The first was a study on the potential impact of an EPS-Sterna constellation of radiometers, which led to several developments that underlie the assimilation of AWS and considered the potential assimilation of sub-mm channels for the first time (Lean *et al.*, 2023; Lean and Bormann, 2024). Second was a study on Metop-SG instruments MWI and ICI in which the method for NWP-based cal/val in the all-sky framework was developed (Duncan *et al.*, 2024b). This previous work set the stage for rapid evaluation of AWS and its unique observations.

Monitoring capabilities for AWS were prepared in advance of the launch and the NWP-based analysis permitted swift feedback of instrument performance to ESA and EUMETSAT during the commissioning phase. Our analysis showed that data quality from AWS is generally good, with biases and noise performance comparable to AMSU-A and MHS but short of ATMS when assessed over broadly similar spatial scales. Of the performance requirements that can be adequately assessed with an NWP model, not all have been met in our estimation, most notably inter-channel biases between channels in feedhorn 1 and those of other horns. Some of these may be possible to address through updates in the level 1 processing (Kangas 2025, pers. communication), though for assimilation purposes most of these outstanding bias differences are adequately handled by the variational bias correction scheme.

Assimilation follows the all-sky method as applied to heritage MW sounders. Of the 19 channels on AWS, here we assimilate 50 GHz channels 4-7 (equivalent to AMSU-A 5-8), 183 GHz channels 11-15 (equivalent to ATMS 18-22), and three 325 GHz channels (16-18). Four months of assimilation trials indicate a positive impact on short- to medium-range forecasts, particularly for humidity and winds. Z500 RMSE is improved by about 1% out to day 3 in the SH and by about 0.5% in the NH at day 2. Short-range forecasts of humidity and winds are improved, and there is a small improvement in tropospheric temperature as evidenced by radiosonde and ATMS background departures. Addition of the new 325 GHz (i.e. sub-mm) channels shows some benefit on top of the other assimilated channels on AWS. This band has similar water vapour sensitivity as 183 GHz but greater sensitivity to cirrus clouds,

and inclusion of these channels slightly improves short-range forecasts of humidity, particularly at higher latitudes. Fuller exploitation of the cloud information in this band warrants investigation in the future.

The microwave sounder on AWS is a valuable addition to the global observing system and represents the first small-satellite radiometer whose performance and stability meet the requirements of operational NWP. The significant benefit to forecasts from AWS assimilation comes partly from its unique orbital crossing time that complements larger backbone platforms, and this bodes well for the proposed EPS-Sterna constellation. Due to the positive impacts seen in forecast skill from assimilation of its radiances, AWS has been included in the ECMWF operational assimilation from July 2025.

1 Introduction

The assimilation of microwave (MW) radiances benefits global forecast skill significantly (Bormann *et al.*, 2019), and are also a leading observation type for driving forecast skill in regional Numerical Weather Prediction (NWP) systems (Randriamampianina *et al.*, 2021). Furthermore, despite the existence of several high quality MW sensors already in orbit, there is continued improvement in forecast skill realised by assimilation of additional sounders. This has been demonstrated both with real observational data and simulated observations, in observing system experiments (OSEs) and analyses using the Ensemble of Data Assimilations (EDA) methods, respectively (Duncan *et al.*, 2021; Lean *et al.*, 2025).

As technology for satellites and radiometers has evolved, increased miniaturisation has led to a new generation of meteorological satellites that are small enough to allow consideration of cost-effective constellations. Compared to the previous paradigm of bus- or car-sized platforms with as many as a dozen instruments, satellites as small as a shoebox can carry a single sensor whose performance may be comparable to a heritage instrument that was one or two orders of magnitude larger. Instruments such as TROPICS[1] and TEMPEST have flown on "CubeSat" platforms; for example, each TROPICS satellite is a 3U CubeSat, meaning that its dimensions are 30x10x10 cm. In the case of the Arctic Weather Satellite (AWS), it is significantly larger than a CubeSat but much smaller than traditional weather satellites and could be classified as a "SmallSat" or microsatellite[2]. At 125 kg, AWS is roughly the mass of a Vespa (Voosen, 2024) with the volume of a washing machine when stowed for launch, albeit with a wider wingspan once its solar panels are deployed. AWS serves as a precursor satellite to the EPS-Sterna constellation proposed by EUMETSAT, a constellation with six such satellites distributed over three polar-orbiting planes.

Part of the new paradigm for meteorological satellite mission design is to allow for more rapid development of technology. This is intended to leverage new technology more quickly than in the past. There have previously been cases in which the period between mission approval and launch was be so long that newly launched instruments had been designed a decade or more in the past, with on-orbit technology thus lagging many years behind the state of the art. However, rapid deployment of new technologies does come with increased risk. Despite improvements in satellite technology, there has not yet been a SmallSat or CubeSat that is assimilated operationally at ECMWF with the exception of the relatively simple GNSS radio occultation receivers. AWS aims to shift the paradigm for the type of mission that is suitable for assimilation in NWP models.

The quality of radiance data required for assimilation in NWP models is very high. The observation geolocation needs to be known very precisely, which requires high pointing accuracy and orbital stability. Radiometric characteristics need to be well described in terms of central frequencies, bandwidths, spectral response, and calibration over a range of scene temperatures. Stability of calibration in time and as a function of orbital position are also important. The lifetime of an instrument in orbit is also a key concern for operational assimilation, with significant resources needed to implement and test a new sensor in an NWP model before it can be considered for operational use. To be blunt, if the lifetime of an instrument is too short, it will not be worth the effort to undergo the months of development and testing required for operational assimilation. All of these are factors that will influence the usage of smaller satellite missions in future years within the NWP community. This report is not intended to address all of these elements, but it is worth stating that this is the context in which AWS evaluation should be considered.

This report focuses on the radiometric performance and initial NWP impact of AWS. Given the status

[1] Acronyms can be found in the glossary here: C

[2] https://www.nasa.gov/what-are-smallsats-and-cubesats/

of AWS as the first operational small satellite led by a space agency, this assessment is done with an eye to the future, especially regarding the new channels and importance of orbit for AWS. The proposed EPS-Sterna constellation from EUMETSAT would maintain a constellation of AWS-type radiometers in orbits designed to complement larger missions like EPS-SG, and thus the on-orbit performance of AWS is important also in the context of being a pathfinder for this proposed constellation.

2 Sensor characteristics and requirements

2.1 The AWS Microwave Radiometer

The AWS instrument[3] observes at 19 channels between about 50 and 330 GHz. The physical dimensions of the instrument are approximately 39x66x54 cm (Albers *et al.*, 2024), with the satellite a little less than one cubic metre when stowed for launch. As shown in Table 1, there are four groups of channels on AWS, representing separate feedhorns on the instrument. The 50 GHz temperature sounding channels are on the first feedhorn, the 89 GHz window channel has its own feedhorn, the 183 GHz humidity sounding channels comprise the third feedhorn along with the 165.5 GHz window channel, and the four 325 GHz sub-mm channels are on the fourth feedhorn.

Like the MW sounders on the operational Metop series of satellites from EUMETSAT (AMSU-A and MHS), the AWS radiometer scans in a cross-track pattern. The scan mechanism observes from nadir out to 54.5 degrees on both sides, with the scan direction perpendicular to the ground track of the satellite. The full scan period is 1.1906 seconds (0.84 Hz, 50.4 rpm). Integration time is 2.5 ms per sample, with 145 measurements taken across the scan (Eriksson *et al.*, 2025).

Central frequencies and bandwidths for AWS are quite typical for heritage MW sounders. Many of the central frequencies are identical to those of channels on AMSU-A, MHS, ATMS, and others. The 183 GHz humidity sounding channels are however single-banded channels, whereas many previous humidity sounders have employed dual passbands to symmetrically sample the 183.311 GHz water vapour absorption line; their weighting functions are however nearly identical to those of dual-banded channels. The 325 GHz humidity sounding channels all feature dual passbands, sampling on either side of the spectral absorption line at 325.15 GHz. For an example of the weighting functions of horns 3 and 4, see Fig. 1 in Camplani *et al.* (2024).

The noise specifications are also nominally similar to those of previous MW sounders. As with Metop-SG, the Noise Equivalent Differential Temperature (NEDT) requirements are specified in terms of the integrated noise over the -3dB Field of View (FOV).In the following we will refer to this as the "footprint" NEDT. At nadir, the footprint on the ground (i.e. FOV) is roughly a circle with diameter given by the final column of Table 1. The footprint grows in size as a function of the scan angle, with the largest FOVs at scan edge. AWS observations are spatially over-sampled for horns 1 and 2, leading to an NEDT for the actual observation sample that is significantly larger than the specified footprint NEDT. In this report, we will hence differentiate between "sample" and "footprint" quantities. Sample NEDTs are expected to be close to the specified footprint values for horns 3 and 4 and significantly larger for horns 1 and 2 (see Table 1 and Fig. 15 in Albers *et al.* (2023)). The ratio of sample to footprint NEDT is a function of the integration time to cover the -3dB footprint (T_{int3dB}) and the integration time per sample (T_{int} =2.5 ms), and is channel-dependent (i) due to the different footprint sizes:

[3]Officially the Microwave Radiometer for AWS mission, but here referred to as the AWS instrument for brevity. https://space.oscar.wmo.int/instruments/view/mwr_aws

Table 1: Specifications for centre frequencies, bandwidths, footprint NEDT, dynamic range, and field of view (FOV) size from the EPS Sterna EURD. All channels have quasi-vertical polarisation (QV-pol). The FOV is defined as the average diameter of the -3dB ellipse at nadir. The NEDT specified is the maximum noise level for the footprint size, rather than sample NEDT (see text). Bandwidths are absolute maximum limits on a -3dB level. So-called AWS channel names are also given (see Eriksson et al. (2025)), with the first number indicating the feedhorn and the second the channel number within the feedhorn. True FOV sizes for the feedhorns at nadir are around 32, 18, 10, and 11 km, respectively.

Ch. #	**AWS Ch.**	*Frequency [GHz]*	*Bandwidth [MHz]*	*NEDT*	*Dynamic range [K]*		*FOV*
					Low	*High*	
1	11	50.3	180	0.6 K	100	310	40 km
2	12	52.80	400	0.4 K	100	295	40 km
3	13	53.246	300	0.4 K	100	285	40 km
4	14	53.596	370	0.4 K	100	280	40 km
5	15	54.40	400	0.4 K	100	260	40 km
6	16	54.94	400	0.4 K	100	250	40 km
7	17	55.50	330	0.5 K	100	240	40 km
8	18	57.290334	330	0.6 K	100	245	40 km
9	21	89.0	4000	0.3 K	80	315	20 km
10	31	165.5	2800	0.6 K	80	315	10 km
11	32	176.311	2000	0.7 K	80	310	10 km
12	33	178.811	2000	0.7 K	80	305	10 km
13	34	180.311	1000	1.0 K	80	300	10 km
14	35	181.511	1000	1.0 K	80	300	10 km
15	36	182.311	500	1.3 K	80	295	10 km
16	41	325.15 ± 1.2	2x800	1.7 K	80	300	10 km
17	42	325.15 ± 2.4	2x1200	1.4 K	80	300	10 km
18	43	325.15 ± 4.1	2x1800	1.2 K	80	300	10 km
19	44	325.15 ± 6.6	2x2800	1.0 K	80	300	10 km

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

Exact -3dB integration times are not published for AWS channels, but following the nominal integration times of Eriksson *et al.* (2025) yields sample to footprint NEDT ratios as found in Table 2.

In contrast to heritage instrument designs, the size constraints of the AWS radiometer led to some key differences in the footprint alignment on the ground that ultimately have ramifications for how the data are processed prior to data assimilation. As discussed by Albers *et al.* (2023), MW radiometers that require aligned beams usually face a design trade-off between limiting the number of bands and the size/complexity of quasioptics needed. For AWS, rather than focus on alignment of beams, the feedhorns are arranged adjacent to each other in a feed cluster, making the quasi-optics much more compact. This leads to footprints that are asymmetric and not co-located: “Since the feedhorns are not co-aligned on the focal axis of the reflector, they introduce asymmetries and other imperfections in several aspects of the quasi-optical design, which will limit its performance compared to other solutions (e.g., MWS).

Table 2: Sample to footprint NEDT ratios for each feedhorn of AWS, given nominal -3dB integration times found in the top row and sample integration time of 2.50 ms.

	Horn 1	**Horn 2**	**Horn 3**	**Horn 4**
T_{int} [ms]	7.5	5	2.5	2.5
NEDT ratio	1.73	1.41	1.00	1.00

Furthermore, the beams diverge on the ground track and measurements taken at a given moment will not be of the same location for each band" (Albers *et al.*, 2023). As seen in their figures 12-15, the antenna response is relatively elliptical at the -3dB level but akin to a fried egg at the outer contours of response, with response varying as a function of scan angle. Even at nadir, the beam centres are tens of kilometres apart, with the 50 and 183 GHz feedhorns nearly 100 km apart. The geolocations of each channel must be accounted for in the observation pre-processing. A co-located observation vector allows for more efficient radiative transfer, quality control that depends upon use of co-located channels, observation error modelling, and so on.

The feed cluster design leads to scan widths and locations on the ground that are quite different for channels of different feedhorns, as seen in Fig. 1. This shows that the scan width and edge of scan are quite different between the feedhorns, with the maximum and minimum zenith angles also feedhorn-dependent. For use in data assimilation, the footprint geolocations and different zenith angles do require attention. The IFS code can efficiently process co-located observations vectors with a common geolocation, and so it simplifies the observation processing within the IFS to have all channels co-located prior to ingest. The method for this will be discussed in Sec. 4.1 and is similar to horn matching for the combined MWI and ICI on Metop-SG (Duncan *et al.*, 2024b).

Due to these differences in feedhorn pointing, the swath width of AWS differs as a function of feedhorn as seen in Table 3 and illustrated in Fig. 2, where the approximate FOVs are shown for every fifth scan in an overpass of the Mediterranean. All AWS horns take 145 samples per scan, one every 0.78 degrees. There is a spacing between scans of about 8.5 km, ensuring that there are overlapping footprints and thus continuous coverage within the swath. The lower frequency channels have the largest swath width of approximately 2500 km, akin to ATMS, whereas the higher frequency channels have a narrower swath width of approximately 2000 km. These values are in the range of heritage MW sounders such as AMSU-A, MHS, and ATMS. The scan characteristics of AWS given here assume a nominal orbit altitude of about 600 km, whereas the altitude of AWS is close to 605 km as of May 2025.

One last aspect worth mentioning here is that AWS scans further from side to side than most heritage radiometers (Table 3). AWS has a maximum scan angle of 54.5 degrees, slightly larger than that of ATMS (52.725), AMSU-A (48.95), and MHS (49.44), but not much larger than MWHS-2 (53.35). This larger range of scan angles leads to a wider swath, but naturally exhibits larger zenith angles for the radiative transfer calculations and longer slant paths. Maximum zenith angles depend on the feedhorn, but can be as large as 65 to 74 degrees, with the largest angles seen for the 89 GHz channel. The large angular range of AWS observations is a consideration for radiometric biases, as both antenna pattern corrections and forward model errors could be significant. Such biases have been an issue for some instruments in the past, with outer scan positions discarded from operational assimilation for some MWHS-2 channels (maximum zenith angle of 66 degrees) and all AMSU-A channels, for example, whereas ATMS radiances are used up to a zenith angle of 64 degrees.

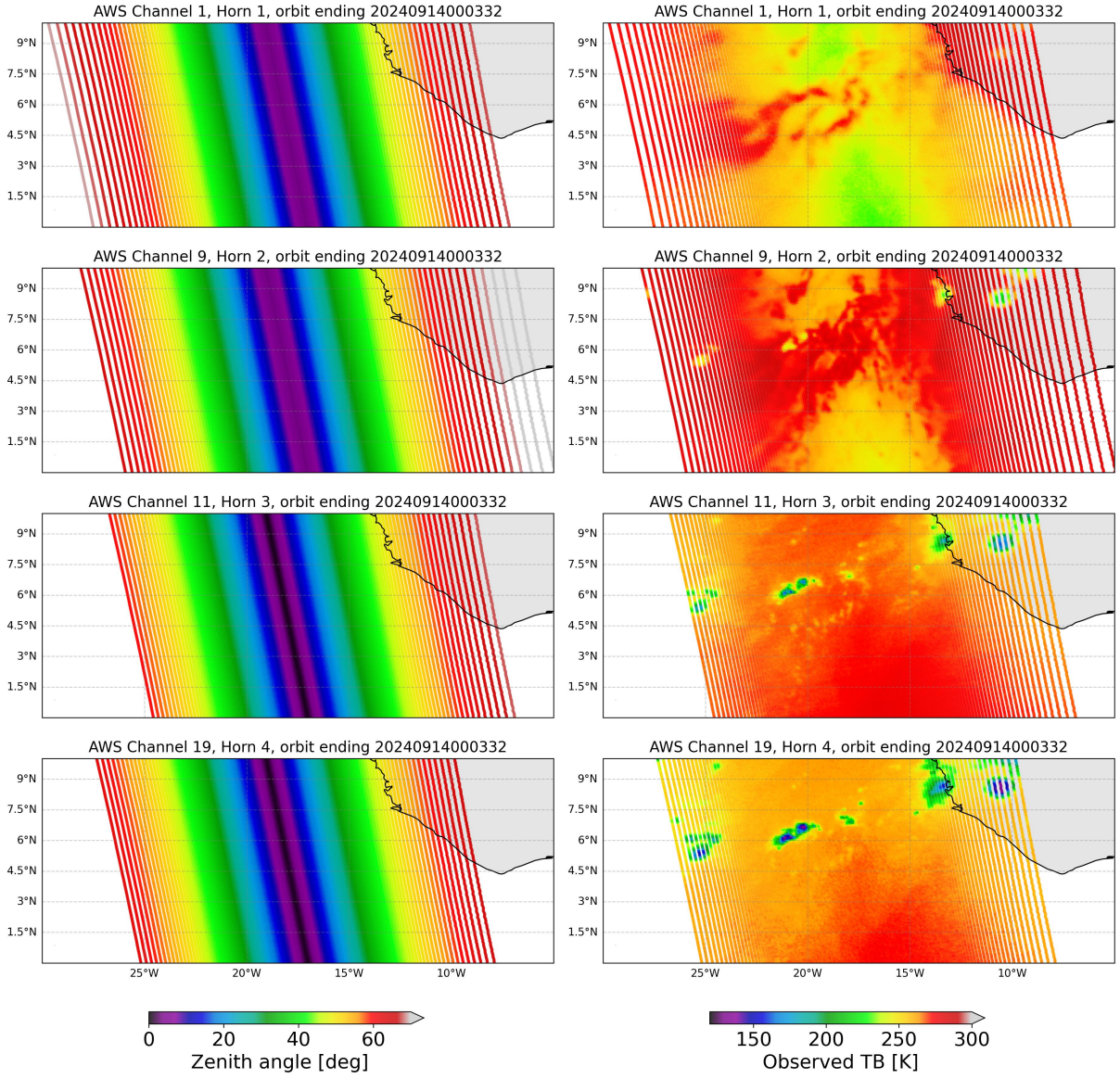

Figure 1: AWS data in the central Atlantic by West Africa, shown for one channel from each AWS feedhorn. Zenith angles are shown on the left with observed brightness temperatures (TB) on the right.

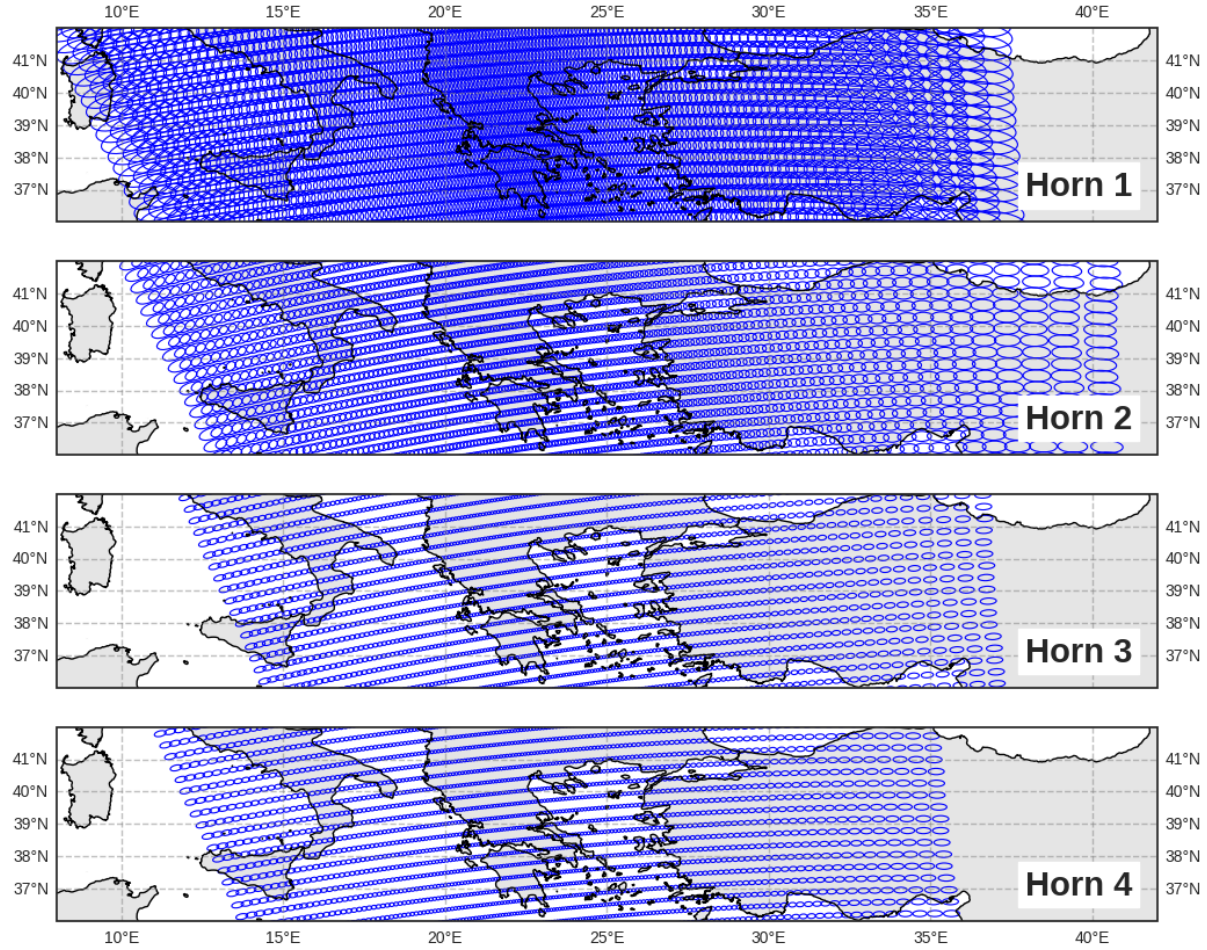

Figure 2: Approximate FOV sizes for each feedhorn, from feedhorn 1 on the top panel to 4 in the bottom. All scan positions are shown but only every 5th scan is plotted for clarity.

Table 3: Scan characteristics for AWS, separated by horn, and for selected currently operational microwave sounders. Swath widths will depend on the nominal satellite altitude and are thus approximate. AWS maximum scan angles do not account for the per-horn contribution that arises from the feedhorn geometry (see Albers et al. (2023)). Swath widths, scan spacing, and altitude are given in km; scan angles and maximum zenith angles are given in degrees.

	Altitude	*Swath width*	*Samples*	*Scan spacing*	*Max scan angle*	*Max zenith*
AWS horn 1	600	2404 - 2581	145	9	54.5	69.7
AWS horn 2	600	2282 - 2441	145	9	54.5	73.3
AWS horn 3	600	1933 - 2059	145	9	54.5	68.3
AWS horn 4	600	1945 - 2071	145	9	54.5	65.4
AMSU-A	830	2250	30	48	48.95	57.5
MHS	830	2180	90	16	49.44	59.1
ATMS	824	2500	96	16	52.725	64.1

2.2 EUMETSAT user requirements for AWS

- **Sterna-EURD-00050** The maximum absolute shift of any Sterna channel centre frequency shall be ± 1 MHz for channels 1-8, ± 130 MHz for channel 9, ± 3 MHz for 10-15, and ± 10 for 16-19
- **Sterna-EURD-00060** Sterna channel bandwidths shall be according to the Table 1
- **Sterna-EURD-00080** Sterna channel shape resolution knowledge shall be bandwidth divided by 100 as a minimum until channel response is down to -40 dB relative to channel maximum response
- **Sterna-EURD-00100** Sterna radiometric sensitivity (NEDT) shall be according to Table 1. This requirement applies to all scene temperatures as defined in the table.
- **Sterna-EURD-00110** The Sterna radiometric bias shall be less than 1 K for all channels and for the Sterna Earth Scene Dynamic Range as defined in Table 1
- **Sterna-EURD-00120** The orbit stability shall be such that variations of the radiometric bias of the measured Sterna brightness temperature during any single orbit shall be < 0.20 K over Sterna Earth Scene Dynamic Range as defined in Table 1
- **Sterna-EURD-00130** Inter-channel radiometric bias differences between brightness temperatures of the same Sterna spatial sample shall be less than 0.5 K over Sterna Earth Scene Dynamic Range as defined in Table 1
- **Sterna-EURD-00140** Radiometric bias differences between brightness temperatures of the same Sterna spectral channel at different spatial samples shall be less than 0.3 K over Sterna Earth Scene Dynamic Range as defined in Table 1
- **Sterna-EURD-00025** The performance of the Level 1b data derived from instruments of the constellation shall be characterised by cross-calibration against a reference, throughout the mission lifetime

User requirements are listed for AWS in the End User Requirements Document (EURD) (see appendix A). Here we list the requirements from the Sterna EURD relevant for evaluation in this report, reworded slightly as needed to fit the report. An item's inclusion in this list does not mean that it can be comprehensively evaluated by NWP-based validation, but at least that NWP-based evaluation can aid in assessing the requirement. In addition to these requirements, there are others in the EURD regarding aspects such as polarisation that are not analysed here.

Table 4 links the EURD requirements to specific evaluations with departure-based analysis from the IFS. In this report, EURD requirements regarding instrument noise (00100), and bias (00110, 00120, 00130, 00140) will be addressed specifically in Sec. 7. A discussion of the spectral response requirements (00050, 00060, 00080) lies in Sec. 7.4. An additional aspect of cal/val that is not covered by the EURD but important in NWP and reanalysis is the temporal stability of AWS observations. This will also be discussed in the results section.

Table 4: AWS (EPS-Sterna) EURD requirements linked with departure-based analysis targets. The titles are short descriptions of the requirement for convenience—see Sec. 2.2 for full requirement descriptions.

Requirement(s)	*Topic*	*Analysis*
00050/60	Bandwidth	Analyse sounding channels for gross airmass biases indicative of spectral shifts; if needed, consider modified RTTOV coefficients to verify
00080	SRF shape	As with bandwidths above, gross violations as manifested in biases will be investigated
00100	Noise	Gross violations of NEDT will be assessed from std(O-B) global mean time series for channels 2-8; effective NEDT can be estimated roughly following Lean *et al.* (2022b)
00110	Radiometric bias	Evaluate with mean O-B over a period of desired length, by channel and by scene temperature
00120	Orbital stability	Monitored over multi-day periods with Hovmöller diagrams showing orbital angle vs. time for mean O-B; intra-orbit stability can be examined as a function of scan line for shorter periods as needed
00130	Inter-channel bias	This is assessed against the IFS with reference to other instruments where appropriate, over a suitably long time period
00140	Inter-sample bias	The IFS here is the reference, using a special experiment in which super-obbing (spatial averaging) is not performed
00025	Cross-calibration	Cross-calibration is against the IFS predominantly; double-differences against ATOVS and ATMS are possible following the channel pairings given in Table 5

3 Relevant aspects of all-sky assimilation in the IFS

In this section, relevant background information is provided that connects to cal/val monitoring and assessment of AWS. This includes a description of all-sky radiance use at ECMWF and some details of the radiative transfer modelling. The treatment of AWS radiances in the IFS follows directly from developments made for the EPS-Sterna project as described by Lean *et al.* (2023).

3.1 Cloud indicators

The Integrated Forecasting System (IFS) produces operational weather forecasts at the European Centre for Medium-range Weather Forecasts (ECMWF). As of late 2024, the operational version of the IFS is Cycle 49r1. This state-of-the-art NWP system assimilates tens of millions of observations per day to optimally determine the initial conditions for the forecast model. This is done via incremental 4-dimensional variational data assimilation, or 4D-Var.

In the IFS, most microwave radiances pass through the "all-sky" system in the form of Brightness Temperature (TB) using the Rayleigh-Jeans approximation. "All-sky" means that MW radiances are simulated and assimilated in clear, cloudy and rainy conditions. This includes emission and scattering of radiation from clouds and precipitation. All-sky assimilation was first applied to conically-scanning microwave imagers (Bauer *et al.*, 2010) and then expanded to humidity and temperature sounders (Geer *et al.*, 2014; Duncan *et al.*, 2022b).

Central to the all-sky approach is modelling of the total error between observations and model, with errors typically increasing in cloudy and precipitating scenes. A so-called "symmetric" cloud amount is used in the observation error model (Geer and Bauer, 2011). Observation errors are scaled as a function of this cloud amount, with larger errors assigned in cloudy or precipitating conditions. The observation errors are inflated as a function of a cloud proxy (C), which can be defined in different ways. The "symmetric" element simply means that observed (obs) and modelled (background, B) cloud amount are given equal weight when calculating C_{sym}:

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

For each instrument and channel, C needs to be defined. This is chosen as something observable (i.e. directly related to the measurements) and correlated with increasing total error as manifest in $std(O-B)$. This choice of C can vary as a function of surface type, as some observed quantities such as Liquid Water Path (LWP) are available or useful only over one surface type. For temperature sounders like AMSU-A, three main cloud proxies are used in the IFS including LWP, Scattering Index (SI), and Cloud Impact (CI). As AWS lacks the 23 and 31 GHz channels needed to calculate LWP, we will focus on *SI* and *CI* here. *SI* exploits a relatively simple relationship in the MW spectrum, where hydrometeors cause larger scattering signals (TB depressions) as a function of frequency. It is hence defined as a difference between brightness temperatures at two frequencies, with larger values indicating heavier precipitation:

$$C_{SI} = TB_{low} - TB_{high} \tag{3}$$

In Eq. 3 above, TB_{low} and TB_{high} signify lower and higher frequencies such as 23 and 89 GHz on AMSU-A, or 89 and 157 GHz on MHS. SI is thus a flexible metric that can be defined for different channel sets.

For AWS, the only C_{SI} utilised here is the difference between 89 and 166 GHz, following the example of other humidity sounders used in the IFS.

Whereas *SI* is a function of hydrometeor scattering, the definition of *CI* accounts for the impact of cloud and precipitation on departures by comparing to a clear-sky simulation (B_{clr}), on a per-channel basis:

$$CI = |(O - BiasCorr) - B_{clr}|/2 + |B - B_{clr}|/2 \tag{4}$$

In this formulation, Eq. 4 follows the symmetric concept by comparing the observation and the model background (with clouds and precipitation, B) against the clear-sky simulation (B_{clr}) that lacks emission and scattering from hydrometeors. *CI* is quite flexible because it can be calculated for any channel. There are prior examples of using *CI* for observation error modelling (Okamoto *et al.*, 2014; Duncan *et al.*, 2022a; Lean *et al.*, 2022a) as well as analysing cloud-affected observations (Duncan *et al.*, 2024b). Note that there is a bias-correction term included (*BiasCorr*) to account for systematic differences between observations and model background.

The flexibility of *CI* as a cloud proxy is appealing for use with AWS, particularly because AWS lacks the lower frequency (23 and 31 GHz) channels that are used by AMSU-A operationally for both LWP over sea and SI over land (Duncan *et al.*, 2022b). In fact, one AMSU-A instrument with broken channels, Metop-B, now uses *CI* from channel 4 to assimilate channels 5 and 6 over land, whereas they were excluded from assimilation over land prior to Cycle 49r1. As described by Lean *et al.* (2022a), a *CI*-based cloud proxy was shown to be as effective as LWP for modelling AMSU-A observation errors. *CI* is an effective, channel-specific metric for screening of observations affected by clouds, as will be discussed later in the Sec. 5.

3.2 Radiative transfer

Radiative transfer (RT) modelling is a key element for both cal/val activities and assimilation of AWS. Most of the frequencies observed by AWS are very well modelled and characterised by NWP models. Particularly in the 50-60 and 183 GHz absorption features, the radiative transfer modelling is mature thanks to decades of development and assimilation. Modelling of sub-mm radiation is much less mature, however, and could be considered a key uncertainty in the evaluation and exploitation of AWS. EUMETSAT has recognised the improvement of sub-mm RT as a priority for the effective exploitation of AWS and the upcoming ICI sensor, with a great deal of work going into elements such as spectroscopy, scattering, and emissivity modelling (e.g. Turner *et al.*, 2022; Barlakas *et al.*, 2022a; Kilic *et al.*, 2023).

The observation operator used in the IFS is RTTOV-SCATT (Geer *et al.*, 2021). In this report, the IFS uses RTTOV-SCATT version 13.2. Model equivalents at the observation location are calculated for all observation geolocations and represent the model fields at that point. They are not averaged to match the sensor footprint, but rather represent the model "effective resolution" at that specific point. The RTTOV coefficients for AWS include emission from well-mixed gases oxygen and nitrogen, plus variable species water vapour and ozone. As with other sounders used in the IFS, AWS forward modelling treats slant-path radiative transfer following Bormann (2017).

Variable ozone in the microwave radiance forward model is now included in the IFS as of Cycle 49r1 specifically to support sub-mm channels such as those on AWS and the upcoming ICI (Duncan *et al.*, 2024b). This is because ozone sensitivity generally increases with frequency (Turner *et al.*, 2022). For lower microwave frequencies the sensitivity to ozone is on the order of hundredths of a degree and thus an insignificant source of forward model error, whereas sub-mm channels have non-negligible sensitivity to

ozone concentrations on the order of tenths of a degree (see Eriksson *et al.* (2025), their Table 4). Hence the AWS radiative transfer coefficients for RTTOV contain variable ozone, and the IFS ozone field is used in forward model calculations. AWS radiances will not influence the ozone analysis, however, as TL/AD sensitivity is currently not switched on.

Another key element for accurate RT simulation is knowledge of the spectral response functions (SRFs) of the instrument's passbands (English *et al.*, 2020; Chen *et al.*, 2021). Public availability of measured SRF information has increasingly become a followed "best practice" from space agencies for MW radiometers, and the lab-measured SRFs for AWS are thankfully available (see Eriksson *et al.* (2025), their Sec. 4.1). The measured SRFs have been accounted for in the RTTOV coefficient file used in this work to simulate AWS radiances, and are publicly available via the NWP-SAF website (`https://nwp-saf.eumetsat.int/downloads/rtcoef_info/mw_srf/rtcoef_aws_1_aws_srf.html`). Use of lab-measured SRFs are important not only for eliminating a potential source of simulation bias, but also for assimilation of sounder radiances, as the vertical sensitivity of each channel is a function of its SRF and thus the increment in temperature or humidity is more accurately placed if using the correct SRF information.

For surface emissivity, the IFS now uses the SURFEM-Ocean model over sea (Kilic *et al.*, 2023; Geer *et al.*, 2024), which was chosen as the new default ocean emissivity model for RTTOV partly because of its support for sub-mm frequencies like those on AWS. Over land, the emissivity will be dynamically retrieved using the most surface-sensitive channels (Baordo and Geer, 2016). This means using the 50.3 and 89 GHz channels primarily to retrieve emissivities for the temperature and water vapour sounding channels, respectively. The 166 GHz channel can be used for emissivity retrieval in some cases where the atmosphere is not too optically thick, such as over sea ice and high orography. The 325 GHz channels follow the emissivity assignments of equivalent 183 GHz ones, but this may be worth revisiting in the future.

Lastly, one of the key areas of interest for AWS is the scattering signal from ice hydrometeors in the sub-mm channels. Preparations for the sub-mm era have focused heavily on the scattering properties of these frequencies, with substantial work going into both reference and fast radiative transfer models (e.g. Geer *et al.*, 2021; Barlakas *et al.*, 2022a). The first data from AWS are therefore highly anticipated as a real-world validation of the scattering properties that have been implemented in radiative transfer models, including but not limited to ice particle habits, particle size distributions, and effects of oriented particles. For example, the significance of oriented particle effects on the polarisation signatures at 325 GHz can be investigated with the AWS cross-track scan, despite not measuring at multiple polarisations (Barlakas *et al.*, 2022b; Wu *et al.*, 2024). The fidelity of RTTOV-SCATT with respect to sub-mm radiances in areas of cloud and precipitation is an important consideration regarding their possible assimilation (see Sec. 8).

3.3 Similar current instruments

AWS cal/val activities can benefit from the fact that most of its channels have close analogues on currently-flying radiometers. Other than the sub-mm channels of AWS and the 53.246 GHz channel, all channels have at least one comparable channel on an operational, cross-track radiometer. In this section, the similar instruments that will be used for comparison are described. These instruments and their homologous channel numbers are given in Table 5.

The Advanced Technology Microwave Sounder (ATMS) is certainly the best similar radiometer for comparison to AWS, as it features analogous frequencies for 14 of 19 AWS channels. ATMS is a temperature and humidity sounder, currently found on three operational platforms launched by NOAA: SNPP,

Table 5: AWS channels and matching channel numbers on other cross-track instruments. Channel numbers with an asterisk (*) constitute significantly different frequency matches (e.g. 157 vs. 166 GHz or single vs. double passband). Parentheses () indicate a different polarisation than AWS for a window channel.

Centre Freq	AWS	ATMS	AMSU-A	MHS
50.3	**1**	(3)	3	
52.8	**2**	(5)	4	
53.24	**3**			
53.596	**4**	6	5	
54.94	**5**	7	6	
54.4	**6**	8	7	
55.5	**7**	9	8	
57.29	**8**	10	9	
89.0	**9**	16	15	1
165.5	**10**	(17)		2*
183±7	**11**	18		5*
183±4.5	**12**	19		
183±3	**13**	20		4
183±1.8	**14**	21		
183±1	**15**	22		3
325±1.2	**16**			
325±2.4	**17**			
325±4.1	**18**			
325±6.6	**19**			

NOAA-20, and NOAA-21. ATMS is not quite as spatially oversampled as AWS, with 96 observations per scan (see Table 3) and a relatively wide range of scan angles as mentioned in Sec. 2.1. As of 2025, ATMS remains in "clear-sky" assimilation in the ECMWF operational system, following Bormann *et al.* (2013). In this project, ATMS is considered in the all-sky framework, using a prototype system that is expected to become operational in the near future. It is treated as other humidity sounders with 50 km superobbing, following Duncan *et al.* (2024a); this means that sampling for ATMS is broadly consistent with that of AWS after super-obbing.

From the previous generation of operational sounders, ATOVS, the AMSU-A and MHS instruments remain key components of the global observing system and drive significant forecast error reduction (Duncan *et al.*, 2021). In May 2025, there are still 5 AMSU-A and 3 MHS radiometers operational on a mixture of NOAA and MetOp satellites[4]. AMSU-A has a large footprint (48x48 km at nadir) and a much longer scan period than AWS (8s versus 1.19s), which has enabled low-noise performance that is excellent for NWP without necessitating spatial averaging. The spatial sampling of AMSU-A is not nearly as dense as AWS, with 30 observations across the scan and scans lying 48 km apart. MHS samples three times more densely than AMSU-A, with 90 observations per scan and scans 16 km apart, and is thus closer to AWS sampling but still about half as dense. In ECMWF operations, MHS is superobbed to 50 km following Duncan *et al.* (2024a), whereas AMSU-A is not superobbed.

As of late 2024, the AWS orbit has an Equator crossing time (ECT) of roughly 22:35 local time on its ascending node (LTAN). It is important to consider the local time of observations in the context of

[4]The old POES series from NOAA (NOAA-15, -18, and -19) concluded their operational mission in mid June of 2025, leaving only Metop-B and -C with functional AMSU-A and MHS instruments.

Table 6: Satellites holding passive microwave sounders assimilated by the IFS as of January 2025. ECT given for ascending node.

Satellite(s)	*Sensor(s)*	*Launch*	*ECT*
NOAA-15, -18, -19	AMSU-A, MHS	1998, 2005, 2009	19:15; 22:45; 21:35
Metop-B, -C	AMSU-A, MHS	2012, 2018	09:30
SNPP, NOAA-20, -21	ATMS	2011, 2017, 2023	13:30

observing system experiments, as the impact in the assimilation system may depend on how this overlaps or complements the orbits of other instruments (Lean *et al.*, 2023; Steele *et al.*, 2023). The ECTs of satellites with MW sounders assimilated in the IFS are given in Table 6, current as of January 2025.

4 Processing of observations

4.1 Data flow and pre-processing

Radiance data ingested by the IFS is typically in BUFR format, converted from its original format before ingest if necessary. In the case of AWS, ECMWF receives level 1 (L1B) data from ESA via EUMETSAT in NetCDF format, which is then converted to BUFR format internally. The data then undergo certain pre-processing steps prior to being used in the data assimilation system, where the data format used is ODB, an ECMWF-specific format. For most microwave instruments assimilated in the IFS, the pre-processing stages include some basic quality control, averaging into superobs, and conversion from BUFR to ODB.

Similar to the preparations made for Metop-SG microwave instruments MWI and ICI (Duncan *et al.*, 2024b), AWS has several feedhorns that point at different geolocations, and this needs to be dealt with in the data pre-processing. To handle this, we first perform spatial averaging on a common grid and then merge those observations into a single, co-located observation vector by leveraging the common grid. This means performing the spatial averaging (superobbing) independently on each feedhorn. The separate files are then merged back together to create a 19-channel, co-located vector of TBs at a chosen superob resolution. Zenith angles are averaged together and stored on a per-channel basis, as these are needed as input to RTTOV for producing realistic simulations because the zenith angles are significantly different between feedhorns (see Fig. 1). Some fields such as azimuth angle and various bit-fields are not averaged together as part of the superobbing. Importantly for radiative transfer calculations, per-channel zenith angles are supported by RTTOV and IFS, so these are passed through the pre-processing chain.

To perform the spatial averaging and horn matching, a change of BUFR formats is needed to go from per-horn geolocations to a common geolocation. For details on this, see Appendix C. It is the second, custom BUFR sequence that is used for input to the IFS. This sequence includes extra fields that are filled by the superobbing procedure, namely the number of observations per superob and the standard deviation of observations in the superob. These values are passed through IFS and could potentially be used for quality control or observation error modelling in the future, but are currently used only for diagnostic purposes. However, it is interesting to see how many observation geolocations lie within each superob grid box, as this shows densely AWS samples. As seen in Fig. 3, there are about 35-40 observations per superob near nadir when using a 50 km superob size. We will come back to this oversampling later for

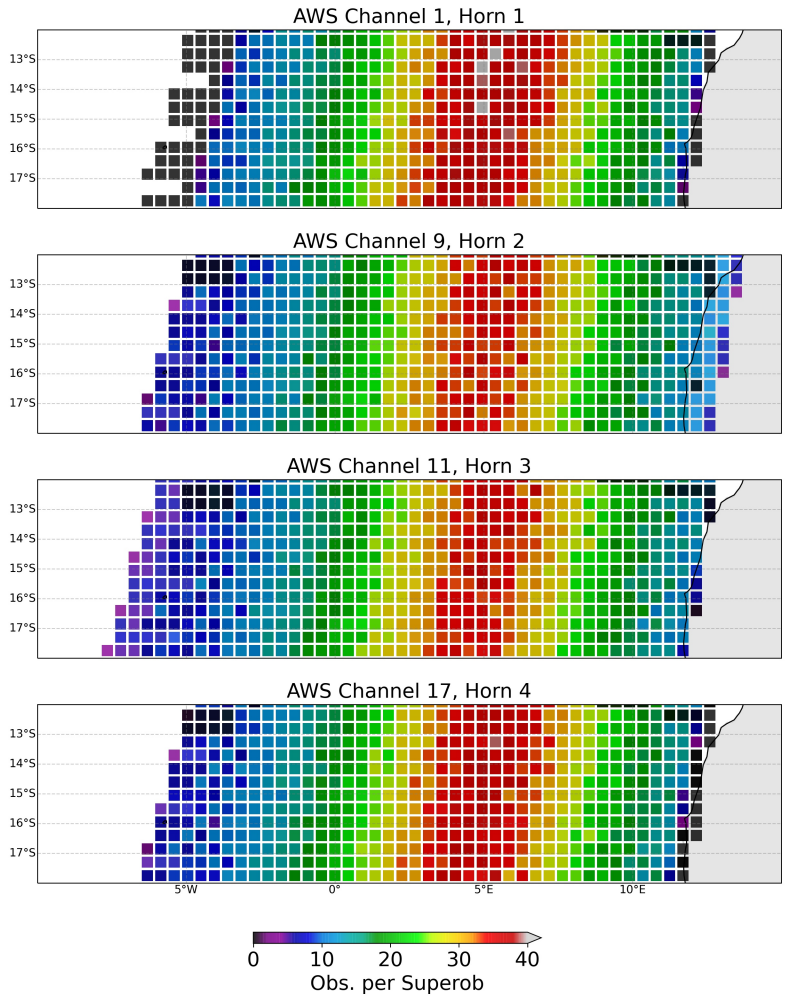

Figure 3: Observations per superob for AWS data.

horn 1, as the number of observations averaged together is important for the effective noise level of the superobs.

In contrast to the 3x3 averaging applied to ATMS radiances in the IFS (Bormann *et al.*, 2013), which was also used in the EPS-Sterna study by Lean *et al.* (2023), the grid-based superobbing applied here for AWS follows that of humidity sounders presented in Duncan *et al.* (2024a). The different approach of grid-based superobbing is useful for AWS due to the non-co-located feedhorns, as a common grid enables exact matching of feedhorns after averaging has been applied. The EPS-Sterna study did not have the full information of sensor geometry and thus assumed co-located feedhorns. However, the concept is the same behind both approaches, in that both forms of averaging are able to decrease the effective noise for temperature sounding channels that has been a crucial aspect for ATMS assimilation, and they also reduce representation error in the assimilation, particularly relevant for humidity-sensitive channels. In addition, superobbing AWS has the added benefit of reducing data volume by approximately 20 times. This is a combined effect of reducing the number of geolocations from the horn matching and the averaging of observations themselves.

Lastly, the BUFR file with all channels co-located on the Gaussian grid needs to be converted to ODB format for ingest to the IFS. This is a standard step that is done for all MW radiance observations. As many fields as possible are passed from BUFR to ODB, including instrument temperatures, solar angles, NEDT values, quality flags, and so on.

4.2 Processing in the IFS

The primary parameter discussed in this study is the observation (O) minus model background (B), or $O-B$, defined in brightness temperatures (i.e. in Kelvin). Means and standard deviations of $O-B$ are helpful for diagnostic evaluation of observations as well as for analysis of model behaviour. In NWP, $O-B$ is typically analysed after bias correction, as a variational bias correction (hereafter VarBC) scheme removes bias from radiance observations prior to assimilation (Dee, 2004). For cal/val purposes, most departure statistics discussed will use $O-B$ without bias correction (D), rather than after bias correction (D_{BC}), where BC is the bias correction offset applied by the data assimilation system to each observation:

$$D = O - B \tag{5}$$

$$D_{BC} = (O - B) - BC \tag{6}$$

Rather than assuming that the IFS (coupled with RTTOV-SCATT) is unbiased, using the IFS as a transfer standard readily permits direct comparison to other sensors. Background departures ($O-B$) are thus available for all channels both with and without bias correction applied after the IFS has run the background trajectory for a given assimilation cycle. The ODB file also contains a multitude of other information in addition to the data input from the BUFR, such as model orography and land fraction, sea-ice concentration, skin temperature, column water vapour, and so on. The orbit angle is also calculated in the IFS, defined relative to the equator (see Fig. 4 following Bormann *et al.* (2023)). All of these parameters can be used for downstream analysis of the departure statistics.

For cal/val purposes as well as some screening decisions within the IFS, we separate channels into window and sounder channels, shown in Table 7. This is to draw a simple distinction between channels with significant surface sensitivity and those without. The window channels used for dynamic emissivity retrieval are also given in the table.

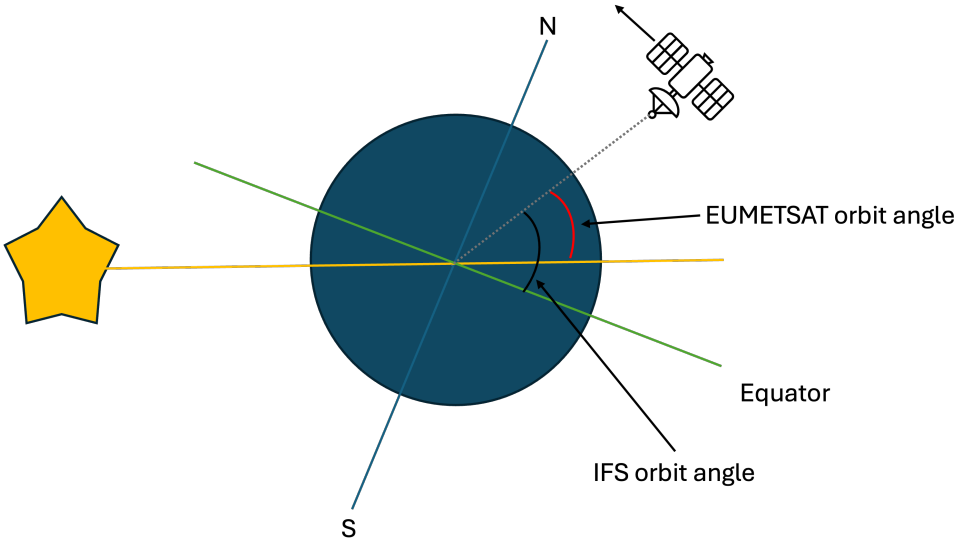

Figure 4: Cartoon of how the IFS and ESA/EUMETSAT orbital angles are defined (taken from Duncan et al. (2024b)). The latter are passed through from the L1B NetCDF input data, defined in relation to the solar beta plane.

Table 7: AWS channel numbers and centre frequencies with their all-sky code indicator(s) if given, the dynamic emissivity channel used, and comparable channels from other cross-track sounders. Channel types are given as window (W) or sounder (S). Feedhorn numbers are given per sensor.

Ch. #	*Freq. [GHz]*	*Horn*	*Type*	*Emis.*
1	50.3	1	W	50V
2	52.80	1	W	50V
3	53.246	1	W	50V
4	53.596	1	S	50V
5	54.40	1	S	50V
6	54.94	1	S	50V
7	55.50	1	S	50V
8	57.290334	1	S	50V
9	89.0	2	W	89V
10	165.5	3	W	89V/166V
11	176.311	3	W	89V/166V
12	178.811	3	W	89V/166V
13	180.311	3	S	89V/166V
14	181.511	3	S	89V/166V
15	182.311	3	S	89V/166V
16	325.15 ± 1.2	4	S	89V/166V
17	325.15 ± 2.4	4	S	89V/166V
18	325.15 ± 4.1	4	S	89V/166V
19	325.15 ± 6.6	4	W	89V/166V

5 Data selection for cal/val

In this section, the rationale and method of data selection for cal/val analysis are described. Following the data processing described in the prior section, in which L1 data are converted, ingested, and departures are calculated, a module within the all-sky code determines the eligibility of an observation to enter the different cal/val data samples. Next we describe the elements used for data selection, followed by the criteria and the outcomes of this selection.

As defined earlier in Eq. 4, the cloud impact metric is a way to quantify the impact of hydrometeors on the background departure, symmetrically balancing observed and modelled clouds. It is important to note that this is done on a per-channel basis, and thus the different sensitivities to cloud and precipitation are taken into account automatically. In addition to *CI* being a useful proxy for observation error modelling, it is similarly a helpful diagnostic to screen out significantly cloud-affected scenes.

Secondly, the sensitivity to the Earth's surface is an important factor in screening. Again, this takes into account the channel sensitivity for each scene, rather than treating all scenes and channels the same way. Therefore the criteria uses the clear-sky surface to space transmittance, hereafter τ, which is an output of RTTOV. This accounts for differences in atmospheric opacity and model surface altitude, given as a number between 0 and 1 (0 signifies no surface sensitivity).

Lastly, we divide AWS into two classes for cal/val analysis: window and sounding channels (given in Table 7). Channels whose average τ value is larger than 0.02 are considered window channels. The sounding channels' surface sensitivity is so small that even large errors in simulated emission from the surface, for example due to poor emissivity or model skin temperature, will not significantly affect the $O-B$. For example, even a 5 K error in surface emission contribution should result in a 0.1 K or smaller error in $O-B$. Note that for humidity-sounding channels, τ is mainly dependent on column water vapour.

It is important to note that beyond this division by τ, the split into window and sounder channels is to some degree a separation of cloud sensitivity as well. This is mainly because lower-peaking sounder channels will penetrate deeper into the troposphere where most clouds and precipitation exist. With this in mind, we can set tighter limits on the amount of cloud impact that is permitted into the cal/val sample whilst retaining a large fraction of the total number of observations for analysis. For example, a high-peaking humidity sounding channel can have a stricter check on *CI* than a low-peaking one without significantly altering the mean behaviour. As a corollary, the signal to noise ratio of temperature sounding channels is smaller than humidity channels, and therefore a 2 K cloud signal is quite large at say 54 or 55 GHz but relatively small at 178 GHz, meaning that a tighter check on *CI* is preferred for the 50 GHz channels that are not surface-sensitive.

5.1 Channel-specific selection

The main cal/val selection criteria used in this study is channel-specific. This allows balancing channel- and scene-specific sensitivity with maximising the data sample.

The channel-specific data selection criteria are found in Table 8, with AWS split into window and sounder channels (as per Table 7). As described above, these categories are broad and should not be over-interpreted as to their true atmospheric sensitivity. The table has a column for "stringent" calibration assessment, which is the main cal/val sample that we analyse in this study. In addition, there is a more relaxed criteria designed to better assess the full dynamic range of the instrument, "dynamic"; this is intended to include land impacts for lower frequency window channels, for example, as ocean-only data

Table 8: Channel-based data selection criteria for AWS channel groups, separated into stringent and dynamic criteria. Here "land" refers to the model land fraction. Any land scenes used must be snow-free as currently determined for similar channels (see Geer et al. (2022)) and exclude high altitude land. Window channels' stringent criteria requires model SST $> 277\,K$.

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$< 1\%$, $CI < (1.5 * \sigma_{clr})$	$\tau < 0.10$, $CI < 2K$

selection permits only a fraction of the total dynamic range to be assessed. Note that the "dynamic" sample still does not cover the full dynamic range, as that would require heavily cloud-affected scenes to be included, but these remain excluded as the representation errors associated with these scenes are considered too large.

The window channel threshold values used in this study follow those of Duncan *et al.* (2024b), as these have previously been tested for current microwave instruments like GMI. However, the sounding channels on AWS require some further consideration to ensure that the data sample is optimal for the purposes of cal/val at these frequency bands. Justification of the *CI* limits for sounding channels is expanded next.

The *CI* threshold of 0.5 K used in the previous study is too restrictive for the humidity sounding channels if a realistic standard deviation is desired, i.e. to indicate relative noise performance, as the background error even in clear-sky is often larger than 0.5 K. This means that some scenes will be filtered out due to too large errors in the background rather than cloud-signals, due to the first term in Eq. 4 being a difference of the clear-sky B and the observation. Thus every scene in which the clear-sky background deviates by more than 1.0 K from the (bias-corrected) observation will be removed by a $CI < 0.5K$ check, regardless of whether there is cloud present. This may not matter for analysing mean biases, but for assessing the relative noisiness of humidity sounding channels, this artificially truncated distribution does not provide useful information.

To explore the importance of *CI* settings on departure statistics, Table 9 presents several static *CI* thresholds for the same data sample, focussing on a selection of sounding channels. If we first focus on the temperature sounding channels 4 to 6, it is clear that the different *CI* thresholds have no impact on the diagnosed global mean bias. Even with no cloud screening at all, the diagnosed global bias is within 0.01 K at these frequencies. However, especially for the lower-peaking channels, the tighter *CI* thresholds change the standard deviations considerably, which is an important way to implicitly measure instrument noise in NWP monitoring; this matters almost not at all for higher peaking channels such as 6-8 that have little cloud sensitivity, but is a key consideration for channel 4 as seen here. For humidity-sensitive channels, again it is clear that global mean biases are not particularly sensitive to the choice of threshold value, though having no cloud screening at all does cause model biases to affect the mean. But again the standard deviations are sensitive to the choice of *CI* threshold. An overly restrictive *CI* threshold of 0.5 leads to $std(O-B)$ of effectively 0.5 at 183 and 325 GHz frequencies, yielding no information on instrument quality. In contrast, larger thresholds allow more data into the sample, yield approximately the same diagnosed bias, and provide some insight on instrument noise, though the majority of this signal remains dominated by background humidity errors.

Following this investigation, we decided to use diagnosed clear-sky observation errors to determine the appropriate *CI* threshold for sounding channels. As seen in Table 8, the threshold is set to $1.5 * \sigma_{clr}$ in this study, based on the observation error assigned in clear skies (σ_{clr}). Full definition of the clear-sky observation errors can be found later in Sec. 6, but σ_{clr} is roughly 0.2 to 0.4 K for channels 4 to 7 and

Table 9: Sensitivity of global departure statistics to different CI thresholds for scenes meeting the stringent screening criteria ($\tau < 0.02$ or over sea), shown for selected AWS sounding channels. Means are given ($avg(D)$) with standard deviations of bias-corrected departures ($std(D_{BC})$) in parentheses. Data are from the 0Z LWDA cycle, 12th February, 2025. Note that these data are prior to a L1 calibration update from ESA in March, and biases thus differ from those presented elsewhere in the report.

Ch	CI = 0.5	CI = 1.0	CI = 1.5	CI = 2.0	CI = None
4	-2.05 (0.255)	-2.06 (0.292)	-2.06 (0.317)	-2.06 (0.336)	-2.05 (0.395)
5	-1.69 (0.236)	-1.69 (0.243)	-1.69 (0.245)	-1.69 (0.247)	-1.69 (0.251)
6	-1.75 (0.228)	-1.75 (0.231)	-1.75 (0.232)	-1.75 (0.232)	-1.75 (0.232)
13	-1.59 (0.500)	-1.62 (0.809)	-1.65 (0.992)	-1.66 (1.117)	-1.57 (3.381)
15	-1.36 (0.517)	-1.40 (0.870)	-1.43 (1.081)	-1.44 (1.219)	-1.46 (2.066)
16	-0.56 (0.518)	-0.59 (0.860)	-0.60 (1.071)	-0.60 (1.218)	-0.41 (3.950)
18	-0.62 (0.496)	-0.67 (0.794)	-0.70 (0.980)	-0.69 (1.117)	0.07 (7.140)

1.4 to 2.4 K for 183 and 325 GHz channels.

5.2 Unified data selection

In addition to the channel-specific criteria given above, there are cal/val analyses that require a homogeneous sample across all channels. For example, inter-channel biases should be examined with a consistent data sample. For these cases, a "unified" data selection is created that can be applied evenly across all AWS channels. This will necessarily be a smaller data sample than that of most channel-specific selections, as significant cloud biases need to be removed from even the most cloud-sensitive channels. Also, as a direct consequence of hewing to the most sensitive channels' selection criteria, the unified data sample is limited to ocean-only.

The unified criteria, as defined by Duncan *et al.* (2024b), uses a combination of *CI* at several key frequency bands. This essentially expands on the channel-specific method as given above by selecting *CI* at channels with specific sensitivities which we want to use for screening. These can be considered representative or especially cloud-sensitive frequency bands. For AWS these channels and their primary sensitivities are:

- 50V (Ch 1) – precipitation and surface
- 89V (Ch 9) – liquid cloud
- 166V (Ch 10) – liquid and frozen hydrometeors
- 325±6 (Ch 19) – frozen hydrometeors

For each channel, $CI < 2K$ must be met for the scene to remain in the unified cal/val sample. Crucially for AWS, each of these channels needs valid data for the scene to be considered, which removes many observations near the edge of the swath because all four feedhorns need to sufficiently overlap for a point to be considered. Due to this criteria's simplicity, it can be applied in the same manner to other sensors on the basis of which frequency bands they have. For instance, ATMS has 50V, 89V, and 166V channels (but no sub-mm channels), so its unified sample depends only on *CI* at the three bands it has available.

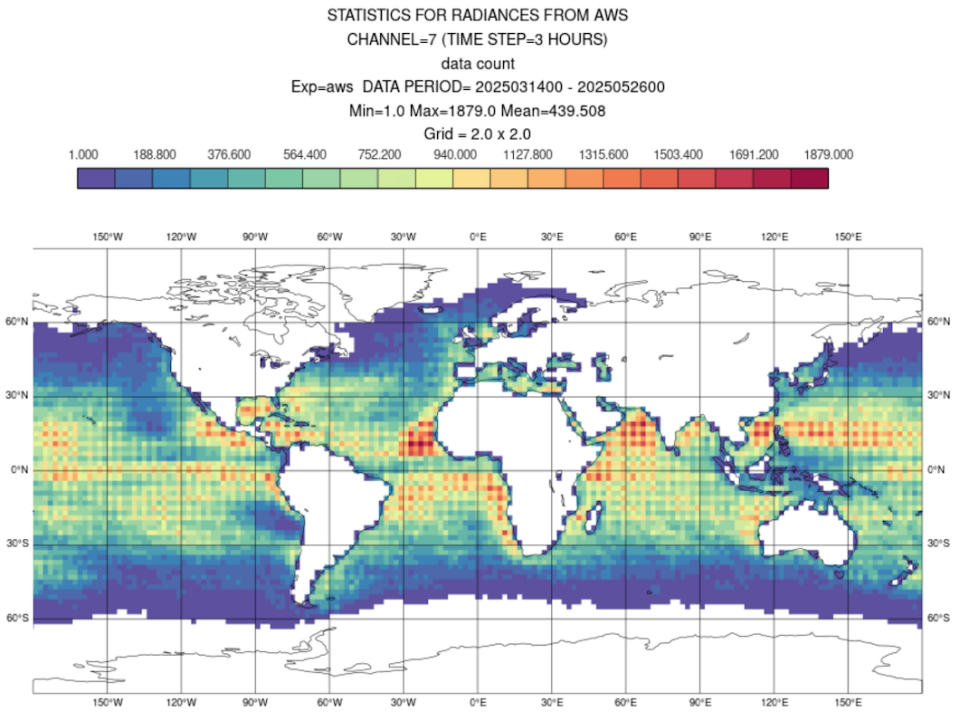

Figure 5: Data counts in the unified selection for AWS, covering mid March to late May of 2025. Some apparent pixelation is the result of a mismatch between the 2x2 degree grid and the 50 km Gaussian grid used for observation processing.

However, the geographical sampling could differ between instruments, for example between AWS and ATMS if there are greater model biases at 325 GHz, so this should be kept in mind when interpreting results using the unified sample.

It is worth examining the geographical distribution of the unified data sample, as this will be relevant in the analysis later. Figure 5 shows the total data counts on a 2x2 degree global grid for a sample AWS channel, but of course the counts are identical for all channels by design. This plot shows that the vast majority of data points exist in the tropics within areas of widespread subsidence. The *CI* thresholds applied across four frequency bands should lead to an effective removal of most clouds of any kind, and indeed the population remaining appears to reflect this.

5.3 Outcome of data selection

Before exploring the results, it is worth considering the volumes of data that are retained for cal/val analysis from the above-mentioned approaches, and illustrating what these data samples look like. Taking a single LWDA cycle as an example, here we will examine the spatial distribution of observations that inform the cal/val analysis, and what fraction of the original data are retained for this purpose.

To illustrate the channel-specific screening for the cal/val analysis, Figures 6 to 9 show swath-level departure maps for four representative AWS channels—temperature and humidity sounding channels, one window channel (166 GHz), and one sub-mm channel. Shown are panels for all data and just the stringent cal/val sample. To emphasise the differences, departures are shown with relatively short colour scales.

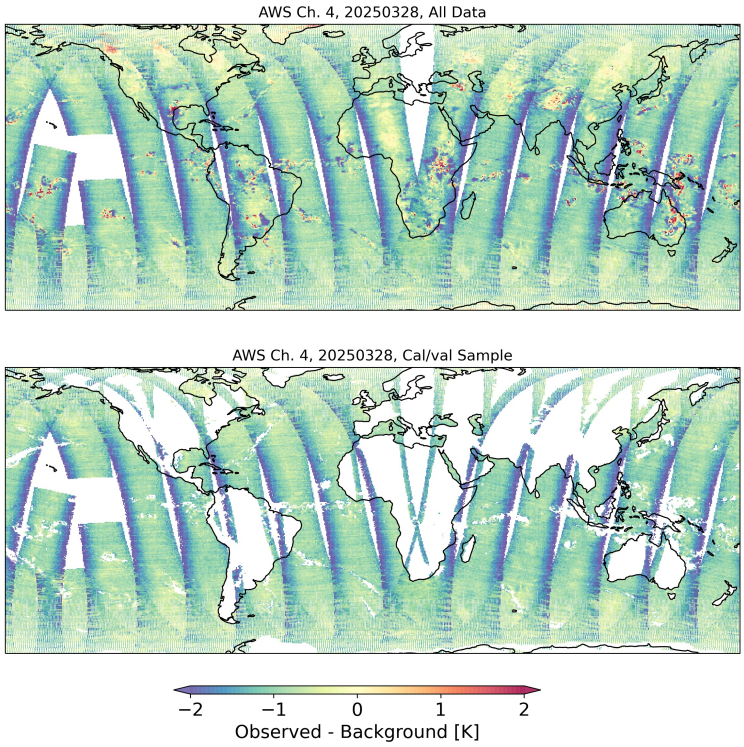

Figure 6: Departures for AWS channel 4 (53.596 GHz) on 28th March, 2025, with no bias correction applied. All data (top) and the stringent cal/val sample (bottom) are shown.

The stringent criteria screens out most clear examples of cloud and surface contamination in the channel 4 departures map whilst retaining a good fraction (but not all) of the biased outer scan positions (Fig. 6). Not much data is retained over land for channel 4 due to its non-negligible surface sensitivity at lower zenith angles, but some data at higher incidence angles are kept. Moving up to channel 10 at 166 GHz (Fig. 7), this window channel has much less data kept overall in the stringent sample. Again the large cloud signals in the tropics and frontal regions are largely removed, as are a large fraction of cloudy scenes near the poles where CAOs (cold-air outbreaks; see Lonitz and Geer (2015)) are common. For channel 15 at 182.31 GHz, which is mainly sensitive to upper-tropospheric humidity and some convective clouds, the *CI* threshold is not very tight and thus many humidity displacement features are retained in the stringent cal/val sample (Fig. 8). Lastly for the lowest-peaking sub-mm channel, 19, the criteria removes a lot of data including large cloud signals in tropical and frontal regions (Fig. 9).

Especially for the higher frequencies with greater cloud and humidity sensitivity, it is clear that there is no magic threshold to remove all areas of potential model bias in cloud, and that features of humidity displacement in the model background would be even harder to remove. But the criteria as defined do appear to yield relatively balanced populations with a mixture of positive and negative features. Ultimately what we want is to remove the clear examples of model bias and feed as large a data sample into our analysis to get robust statistics of instrument performance. Although the eye can pick out regions in which slightly stricter screening might be warranted, this balance seems reasonable in a qualitative sense.

Lastly, Table 10 provides the numbers of observations relating to the previous figures. Shown in the right column are totals for each channel in terms of available superobs. Note that these are not identical across feedhorns, as the swaths cover different areas (see Fig. 2), though the totals for horns 1 and 2 are quite

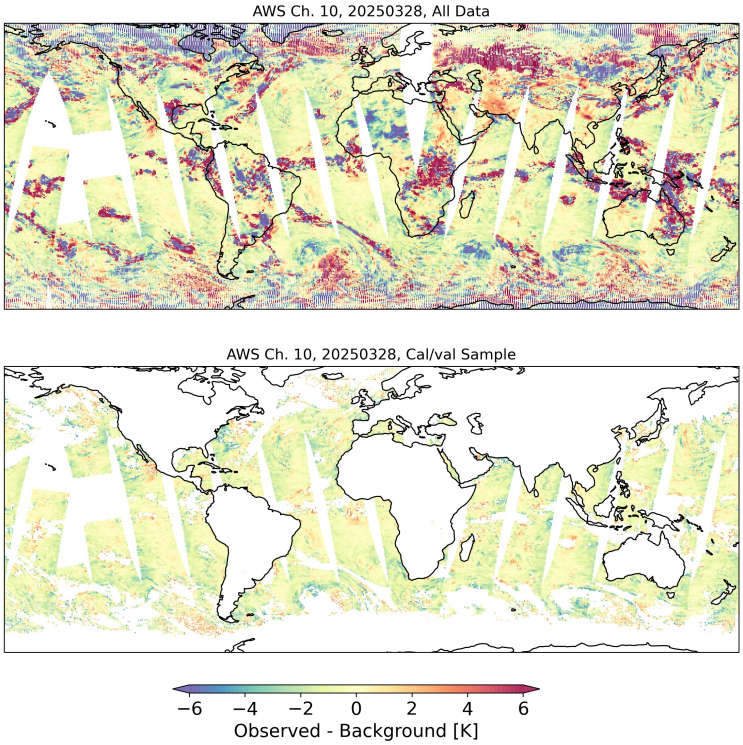

Figure 7: As the previous figure, but for AWS channel 10 (165.5 GHz).

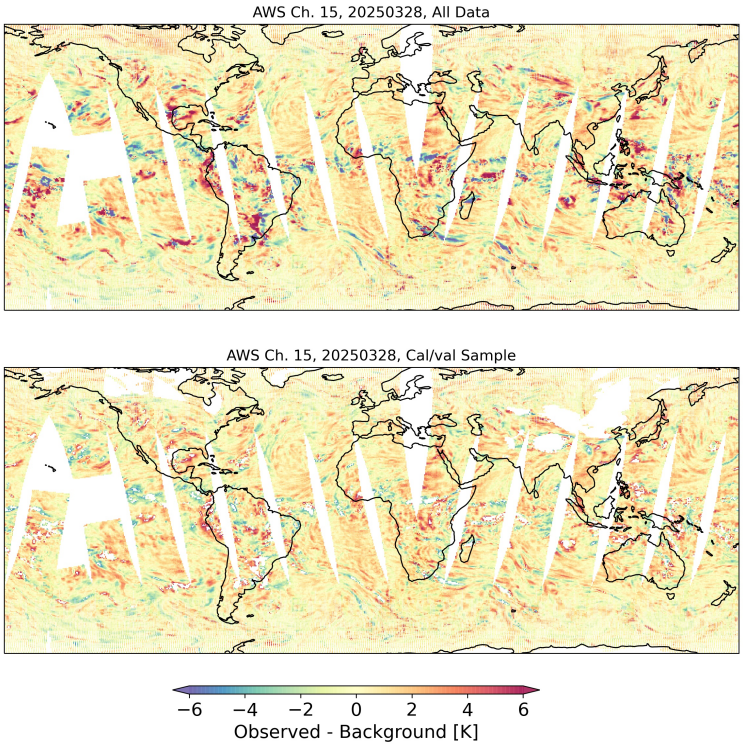

Figure 8: As the previous figure, but for AWS channel 15 (182.31 GHz).

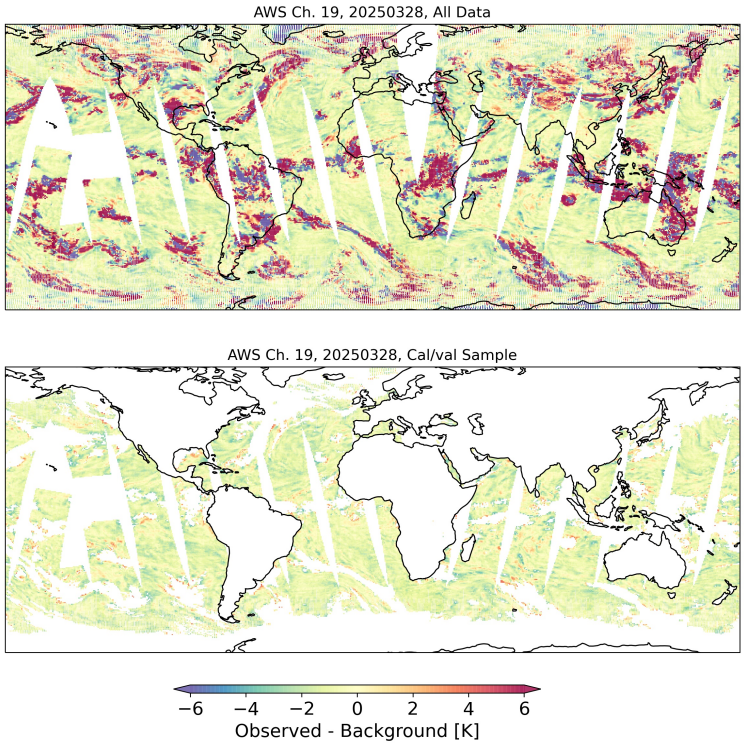

Figure 9: As the previous figure, but for AWS channel 19 ($325 \pm 6.6 GHz$).

Table 10: Total data points (superobs) available in one LWDA cycle and those in each cal/val sample, separated by channel, for March 28th 0Z, 2025. Percentages of the total are given in parentheses.

Ch.	Stringent (%)	Dynamic (%)	Unified (%)	Total
1	83181 (*27.7*)	111700 (*37.2*)	32422 (*10.8*)	300192
2	129038 (*43.0*)	161449 (*53.8*)	32422 (*10.8*)	300192
3	130986 (*43.6*)	164037 (*54.6*)	32422 (*10.8*)	300192
4	196160 (*65.3*)	205746 (*68.5*)	32422 (*10.8*)	300192
5	265002 (*88.3*)	291847 (*97.2*)	32422 (*10.8*)	300192
6	293360 (*97.7*)	299689 (*99.8*)	32422 (*10.8*)	300192
7	290856 (*96.9*)	299458 (*99.8*)	32422 (*10.8*)	300192
8	299662 (*99.8*)	299662 (*99.8*)	32422 (*10.8*)	300192
9	49925 (*16.4*)	72303 (*23.8*)	32422 (*10.7*)	304302
10	81351 (*31.3*)	102225 (*39.4*)	32422 (*12.5*)	259767
11	95279 (*36.7*)	119306 (*45.9*)	32422 (*12.5*)	259767
12	99216 (*38.2*)	124839 (*48.1*)	32422 (*12.5*)	259767
13	195961 (*75.4*)	193541 (*74.5*)	32422 (*12.5*)	259767
14	210149 (*80.9*)	217694 (*83.8*)	32422 (*12.5*)	259767
15	219440 (*84.5*)	229590 (*88.4*)	32422 (*12.5*)	259767
16	207294 (*80.9*)	217284 (*84.8*)	32422 (*12.6*)	256324
17	190182 (*74.2*)	197377 (*77.0*)	32422 (*12.6*)	256324
18	171735 (*67.0*)	166685 (*65.0*)	32422 (*12.6*)	256324
19	82324 (*32.1*)	102949 (*40.2*)	32422 (*12.6*)	256324

similar, as are the totals for horns 3 and 4. The unified criteria provides the same set of scenes across all channels, and thus these are the same across all channels, though the percentages differ due to the totals not being identical between channels. The stringent and dynamic criteria yield nearly 100 percent of superobs at high-peaking horn 1 channels, or as low as roughly a quarter of total superobs for window channels like 50.3 or 89 GHz.

6 Definition of observation errors

As stated in Sec. 3.1, the all-sky assimilation approach requires cloud proxies to facilitate scaling of observation errors with cloud amount. And as described in the previous section, clear-sky observation errors are used as part of the cal/val selection criteria for sounding channels. It is thus worthwhile elaborating on how these are defined, and especially the augmentation to the traditional all-sky error model required for noise-dominated channels on AWS horn 1.

The all-sky assimilation method for AWS uses three different cloud proxies for the three feedhorns of assimilated channel sets (the 89 GHz channel from feedhorn 2 is not assimilated here). The same proxy is used for both land and sea scenes for all AWS channels. These are listed below:

- Horn 1: Cloud impact (CI) at channel 2 ($52.8GHz$)
- Horn 3: Scattering index (SI) of channels 9 and 10 ($89-165GHz$)
- Horn 4: Cloud impact (CI) at channel 19 ($325\pm6.6GHz$)

The use of these cloud proxies follows a mixture of previous methods used operationally in the IFS and in preparatory work for EPS-Sterna (Lean and Bormann, 2024). Here we will focus only on observation errors for the channels that are assimilated in this study, namely channel numbers 4 to 7 and 11 to 18.

The main determinant of the observation error model is the behaviour of $std(O-B)$ as a function of the cloud proxy, following Geer and Bauer (2011). These relationships are determined for the population of assimilated data points over a given surface type, either land or sea. Table 11 lists the clear and cloudy tie points of the all-sky observation error model for each channel (σ_{clr} and σ_{cld}), covering both sea and land surfaces. The maximum cloud proxy (C_{cld}) is also given, with the minimum cloud proxy being zero in all models here. To visualise the observation error models, Figure 10 shows these alongside standard deviations of bias-corrected departures as a function of the cloud proxy. The shapes and magnitudes of these error models are similar to those of other instruments used in the ECMWF all-sky system (e.g. Steele *et al.*, 2023).

The superobbing procedure that is applied to AWS data leads to lower effective radiometric noise in the middle of the swath than at its edges. In Duncan *et al.* (2024a) it was determined to be relatively unimportant to account for this effect for humidity sounding channels at 183 GHz, where the normalised standard deviation of $O-B$ would vary at most 10-15% between nadir and scan edge. However, for temperature sounding channels where the radiometric noise is the dominant component of the standard deviation of background departures in clear-sky regions, this should to be accounted for in the observation error model.

From Geer and Bauer (2011), observation errors in all-sky assimilation are best understood as a representation of total error, which for humidity-sensitive radiances is typically dominated by so-called "representation error" resulting from mismatches in represented scales and processes in the forecast model as

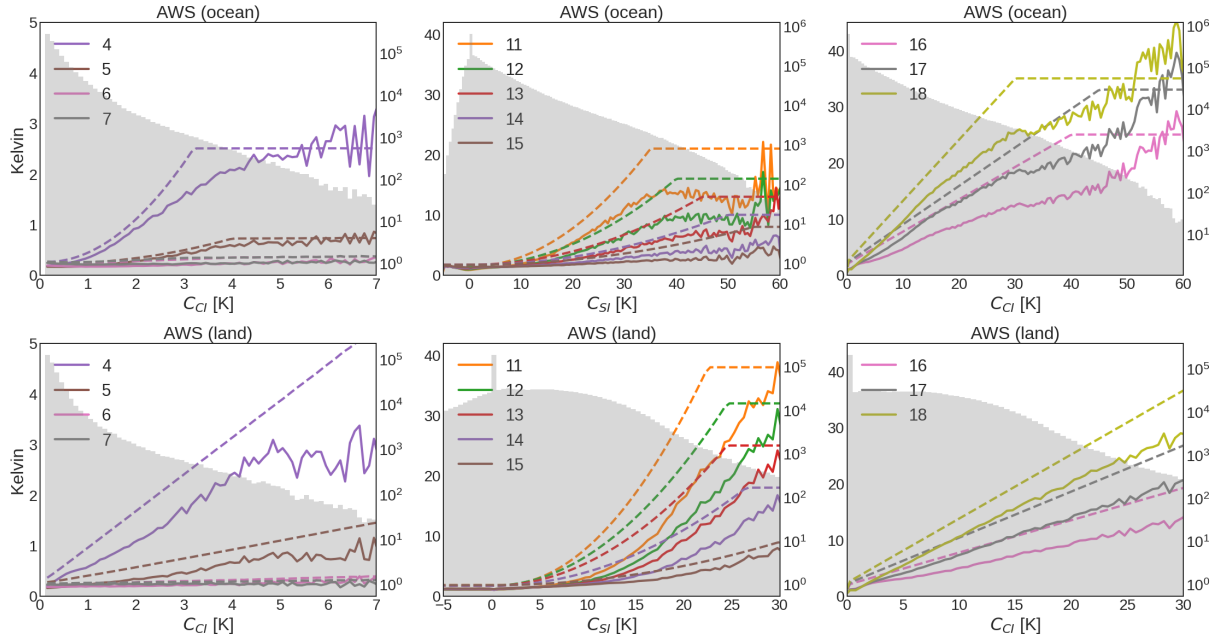

Figure 10: Observation error models in dashed lines for selected channels from horn 1 (left panels), horn 3 (middle), and horn 4 (right panels) over sea (top) and land (bottom). Standard deviations of background departures for assimilated data are shown in solid lines as a function of binned cloud proxy. The population size of each cloud proxy bin is shown in the grey bars. Statistics cover 1st to 23rd July, 2025, using assimilated data after quality control.

well as errors in the observation operator. For temperature-sounding channels of AWS, the varying size of the instrument noise contribution is modelled through an extra term ($f(\theta)$) that is dependent on the zenith angle (in radians). In principle, the term could be formulated in terms of the sample NEDT and the number of samples used in the superob, but for pragmatic reasons this is not pursued here. Instead, a fit to the standard deviation of background departures as a function of the zenith angle (θ) is used. This σ_{zen} term is scaled as a function of $f_{scale}(\theta)$ and added in quadrature to the original formulation of the observation error (i.e. a total error with cloudy and clear-sky contributions, $\sigma_{clr+cld}$), which may be cloudy or clear.

$$\sigma_{total} = \sqrt{\sigma^2_{clr+cld} + (\sigma_{zen}/f(\theta))^2} \tag{7}$$

$$f_{scale}(\theta) = 0.1 + 0.9exp(-0.8\theta^2) \tag{8}$$

Table 11 contains the values for σ_{clr}, σ_{cld}, and the new term σ_{zen}. It is worth emphasising that the σ_{zen} term does not necessarily have physical meaning, but should be related to general NEDT characteristics as these become more evident at higher zenith angles where superobs have fewer observations. For example, we know from pre-launch measurements that channel 7 is noisier than channels 4 and 5 (Eriksson *et al.*, 2025), and this is reflected here, but there is not a one to one correspondence between this term and on-orbit NEDT.

To take a couple of illustrative examples, an observation in clear skies at channel 7 over sea will have a diagnosed observation error of 0.198 K at nadir, increasing to 0.220 and then 0.327 K at 30 and 60 degrees zenith angles, respectively. In contrast, a somewhat cloudy channel 5 observation which at nadir

Table 11: Components of the all-sky observation error model for AWS channels assimilated in this study. The σ values are in Kelvin, and C_{cld} is in the units of the cloud proxies outlined above, which are also all defined in Kelvin. The last column (A) provides a priori probabilities of gross error for VarQC (see Andersson and Järvinen (1999)).

Ch	**Surface**	σ_{clr}	σ_{cld}	C_{cld}	σ_{zen}	A
4	Sea	0.2	2.5	3.2	0.1	0.20
4	Land	0.22	7.5	10	0.1	0.20
5	Sea	0.16	0.7	4	0.12	0.10
5	Land	0.18	1.8	9	0.12	0.10
6	Sea	0.15	0.3	4	0.12	0.02
6	Land	0.15	0.4	9	0.12	0.02
7	Sea	0.14	0.28	3	0.14	0.01
7	Land	0.14	0.28	10	0.14	0.01
11	Sea	1.4	21	35	0	0.50
11	Land	1.8	38	22.5	0	0.50
12	Sea	1.55	16	40	0	0.50
12	Land	1.7	32	24.5	0	0.50
13	Sea	1.7	13	45	0	0.30
13	Land	1.7	25	24.5	0	0.30
14	Sea	1.7	10	50	0	0.30
14	Land	1.7	18	26.5	0	0.30
15	Sea	1.7	8	55	0	0.30
15	Land	1.8	9	30	0	0.30
16	Sea	2	25	40	0	0.30
16	Land	2	25	40	0	0.30
17	Sea	2.2	33	45	0	0.30
17	Land	2.2	35	40	0	0.30
18	Sea	2.4	35	30	0	0.50
18	Land	2.4	48	40	0	0.50
19	Sea	3	50	30	0	0.50
19	Land	3	60	40	0	0.50

has a total observation error of 0.430 K near nadir would have a total error of 0.454 K at 30 degrees, as the noise contribution is relatively insignificant in cloudy skies due to adding in quadrature. Importantly for the cal/val analysis, the threshold for cloud impact ($CI < (1.5 * \sigma_{clr})$) is defined using the full clear-sky error, or the total of σ_{clr} with the zenith-dependent component: $\sqrt{\sigma_{clr}^2 + \sigma_{zen}^2}$. In this way, the cal/val analysis should retain noisier superobs from the scan edge.

Lastly, though not observation error parameters per se (and not discussed further here), the Variational Quality Control (VarQC) settings for AWS are given as the final column in Table 11 for completeness (Andersson and Järvinen, 1999). Here A is the a priori probability of gross error. These values follow those used for sounders such as AMSU-A and MHS, and are important for down-weighting outliers in the 4D-Var analysis that have large normalised analysis departures (see Duncan *et al.* (2022b)).

7 Cal/val results

In this section the bias, noise, and spectral characteristics of AWS radiances are analysed relative to the IFS. Comparisons are also offered for departure statistics of other MW sounders, allowing cross-evaluation without the need for co-location. Where relevant, these analyses are linked directly to the performance requirements outlined in Sec. 2.2. The following is a summary of the main characteristics found, highlighting a few topics of particular interest. Full details of the departure statistics can be found on the dedicated AWS monitoring website made available to EUMETSAT.Though not linked to cal/val requirements per se, the final subsection here looks at the radiative transfer fidelity of RTTOV-SCATT at sub-mm frequencies, as this analysis affects not only our perspective on the 325 GHz cal/val results, but also whether assimilation of these channels should be attempted in the following section.

These results generally draw from AWS L1B data after a calibration adjustment was made in the L1 processing by ESA in mid March of 2025. Unless stated otherwise, analysed departures are considered without bias correction applied and using the standard IFS processing as laid out in Sec. 4, i.e. with 50 km superobbing applied. All results come from IFS Cycle 49r1 run at Tco399 (27 km) horizontal resolution from an experiment in which AWS channels 4-7 and 11-18 are active.

7.1 AWS biases relative to IFS

Figure 11 shows the global mean biases of all AWS channels with respect to the IFS, following the three different cal/val selection criteria outlined earlier. The fraction of total observations that each cal/val sample represents is given by vertical bars, showing similar values as reported earlier for a single 12hr sample of data (Table 10). To indicate the requirement of 1 K bias accuracy (EURD-00110), black dots are shown at $\pm 1K$ in the figure. From this view, a majority of AWS channels are within 1 K of IFS, including all of the horn 3 and 4 channels. The one obvious outlier is the 89 GHz channel 9, which has a warm bias relative to IFS of a bit more than 2 K. Most of the horn 1 channels have a net bias of about -1 K, suggesting that these channels are borderline compared to the 1 K bias requirement.

A key message in this plot is that the global biases are relatively similar against IFS despite the different selection criteria applied, highlighting that there is little sensitivity to the selection criterion applied. In most cases, the global means agree within one or two tenths of a Kelvin. Unsurprisingly, the most surface-sensitive frequencies with the most restrictive screening criteria – channels 1 and 9 – see more variation between sample means, and are also the channels with the greatest variability within the stringent population. The means differ slightly between the unified and channel-dependent screening criteria,

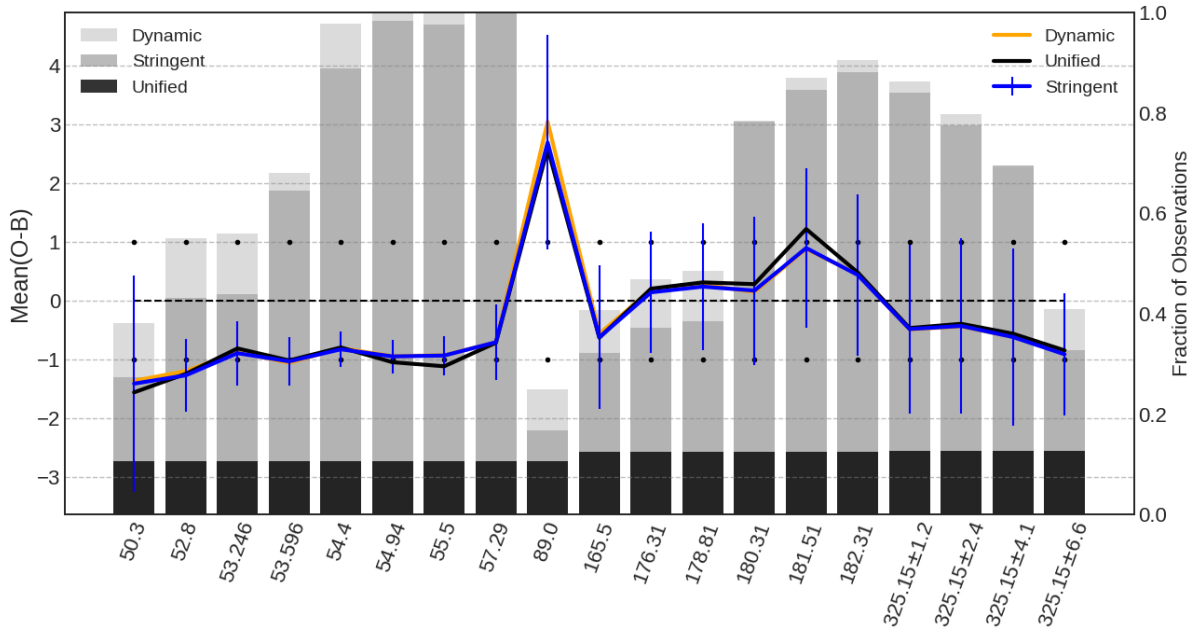

Figure 11: Mean biases against IFS for un-bias-corrected background departures. Global data from April 4 to 29, 2025. Fractions of total observations in each cal/val sample are provided in the grey bars. Standard deviations of the stringent sample are shown by vertical lines.

whereas the stringent and dynamic means tend to be very close together as the samples overlap significantly (see Table 10). The unified sample is very limited in geographic extent (Fig. 5) and so it is possible that model or scene-temperature biases in these subsidence regions of the deep tropics that are over-represented in the unified sample are somewhat skewing the mean values.

The other anomaly here that sticks out is channel 14 (181.51 GHz), which is several tenths of a degree warmer than the other 183 GHz channels in horn 3. In contrast, the never before flown 325 GHz frequencies show a relatively consistent bias structure that is within about 0.5 K, albeit with slightly larger negative biases at the lower peaking channels 18 and 19. We will return to the behaviour of channel 14 below (Sec. 7.4). Regarding Sterna-EURD-00130 (0.5 K inter-channel bias), we can examine the black line in Fig. 11 that represents the unified data sample. This performance requirement is generally met within horns, but not across the entire instrument. Considering each horn separately, the sounding channels in horn 1 (channels 3-8) meet the requirement, as do four of five 183 GHz channels and all of the 325 GHz channels. In horn 3, channels 10 and 14 are colder and warmer than the others in this horn, not at a level that is a concern for assimilation but enough to not meet the requirement. Certainly 89 GHz is well outside this rather strict 0.5 K threshold. Inter-channel biases between channels in horns 1, 3, and 4 appear to exceed the 0.5 K threshold, primarily due to the sizeable cold bias in horn 1. While biases in the background may contribute to these cold biases, comparisons of departure statistics for other instruments suggest that AWS is the odd one out.

How do the global biases of AWS compare to those of other MW sounders? This relates to the cross-calibration and inter-channel EURD requirements (00025, 00130). Figure 12 shows AWS departure means against ATMS (NOAA-21), AMSU-A (Metop-C), and MHS (Metop-C). Biases at 183 GHz are very close to those of ATMS and MHS, with channel 14 being the lone exception. AWS channel 10 at 166 GHz is on the cold side relative to ATMS and MHS, but only by about 1 K. The 89 GHz window channel is an interesting one, as all instruments examined here are a bit warmer than IFS, but AWS

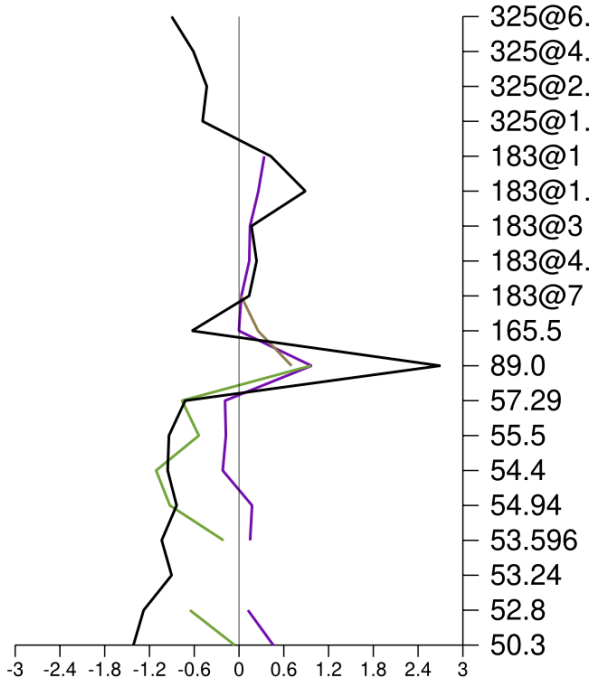

Figure 12: Mean departures for the stringent cal/val sample for AWS (black), ATMS on NOAA-21 (purple), AMSU-A on Metop-C (green) and MHS on Metop-C (gold), with the grey zero line indicating the IFS background as reference. Global data from mid March through mid May. Channel frequencies are given on the y-axis in GHz.

channel 9 is warmer still by nearly 2 K. This suggests that some of the bias observed against the IFS for the 89 GHz channel of AWS may be due to biases in model fields or the forward modelling, but there is likely a significant portion that is due to a bias in the observation. For the temperature-sounding channels, ATMS is close to un-biased globally against the IFS, in contrast to the consistent bias of around -1 K for AWS. AMSU-A on Metop-C shows a negative bias against the IFS that ranges from about -1 to 0 K for these channels. However, the present analysis includes the outer scan positions (1-3 and 28-30) which show significant negative biases, thought to be biases in the AMSU-A APC, and are not included in the assimilation. If these scan-positions are excluded, different AMSU-A instruments typically exhibit biases of within 0.5 K against the IFS for the temperature-sounding channels (see Weston and Bormann (2018), their Fig. 17). The comparison suggests that the diagnosed negative biases for horn 1 AWS channels likely reflect a bias in the observations, rather than an IFS bias.

Beyond global mean biases, maps of bias patterns are generally in line with heritage instruments (not shown), but we can also examine the bias characteristics in other dimensions, such as by a function of the instrument's scan position. Figures 13, 14, and 15 show the bias characteristics of superobs as a function of the horn 1 scan position for the stringent cal/val selection. The horn 1 channels exhibit biases near the edge of scan that VarBC struggles to adequately correct through a third-order polynomial in the scan angle. Especially for channels 4 and 5, which show increasingly negative biases near scan edge, the bias-corrected departures (dotted lines) stray from the zero line. However, biases through the middle of the scan are fairly constant and well-managed by VarBC. In contrast, the horn 3 channels around 183 GHz show an almost perfectly linear bias increase as a function of scan position. The cause of this appears to be understood as debris contamination rather than APC (see Eriksson *et al.* (2025), their section 6.3), and the resultant bias is handled well by VarBC. Lastly, the sub-mm channels of horn 4 show almost no bias variation across the scan. Considering the EURD requirement of inter-sample biases of less than 0.3 K (00140), several channels in horns 1 and 3 do not meet this threshold.

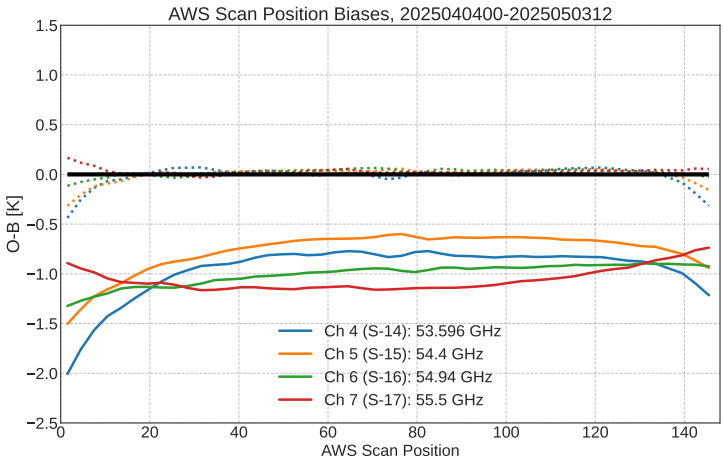

Figure 13: Scan position biases for assimilated horn 1 channels. Global data that meet the stringent cal/val criteria from April 4 to May 3, 2025. The X-axis represents the horn 1 scan position number. Dotted lines show biases after bias correction.

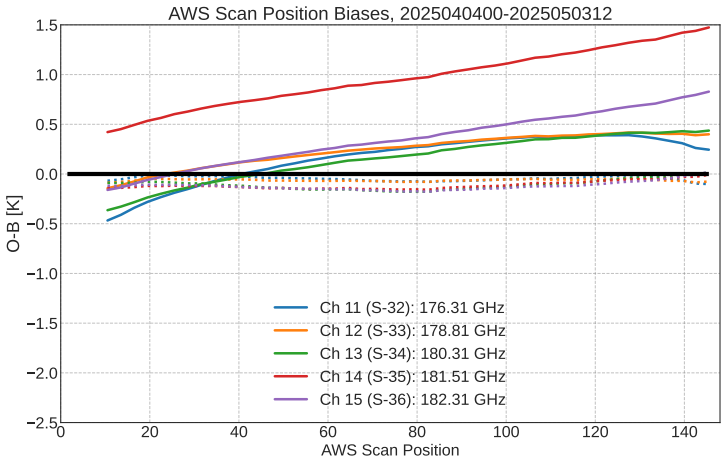

Figure 14: As Fig. 13 but for horn 3 assimilated channels.

The scan-dependent biases for horn 1 are likely due to current short-comings of the Antenna pattern correction in the level 1 processing. In this context, it is worth mentioning that a short EUMETSAT-processed sample of AWS data showed smaller overall biases for horn 1, combined with reduced scan-dependent biases (not shown). This is thought to be the result of a more complete APC. It is expected that the refined corrections will be applied in the ESA processing in due course.

In addition to the superobbed data, a special experiment was run in which AWS data underwent no superobbing; using these data we can examine the inter-sample bias (EURD 00140) and analyse the full scan position biases to see if any sharp features are present that may be obscured in the averaged data. Fig. 16 shows the un-averaged data for a more limited time period to demonstrate that the results are almost identical to Fig. 13. The lines are a little noisier due to the shorter time period analysed, but the similar structures give us confidence that assessing bias structures with the superobbed data is sufficient. It is striking how similar these two plots appear—despite differences in data pre-processing and time periods and numbers of data points considered. One curiosity to note is that there are little bumps in the mean curves that seem nearly identical between the plots, with both showing small variations in scan

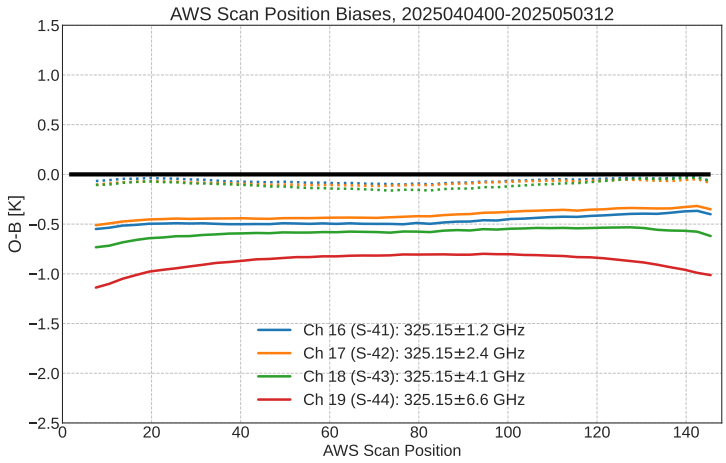

Figure 15: As Fig. 13 but for horn 4 channels.

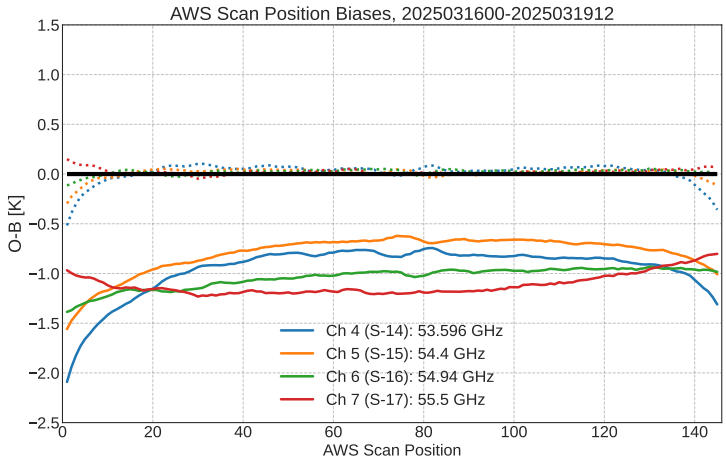

Figure 16: Scan position biases for assimilated horn 1 channels with no superobbing applied. Data from March 16 to 19, 2025.

bias of a few hundredths of a degree that seem to be real signals rather than statistical noise. This could be caused by standing waves, and might warrant further investigation.

Another way to visualise the fit between observations and model is to examine histograms of the departures. Figure 17 shows both "all data" over sea and just the stringent cal/val sample over sea. This is a nice illustration of the cal/val method and the way that outliers tend to be removed by the screening, and especially how the PDFs become more Gaussian after the screening. There is a great deal to digest in this figure, from the relatively similar shapes of 183 and 325 GHz PDFs, to the rather un-Gaussian PDFs of window channels such as 1 and 9. In most cases, the mean of the cal/val sample is within a few tenths of a degree of the full sample, but these differ the most for cloud-sensitive frequencies. The residual scan pattern bias described above is visible in the skewed PDFs of channels 3 to 5, in contrast to channels 6 to 8 which exhibit less of this effect. It is also clear how much worse the NEDT performance is for channel 8, with its PDF roughly double the width of channels 6 and 7.

As can be seen from results in the previous pages, the sounding channels of AWS generally show good agreement with both the IFS and other comparable MW sounders in a global mean sense. However, this

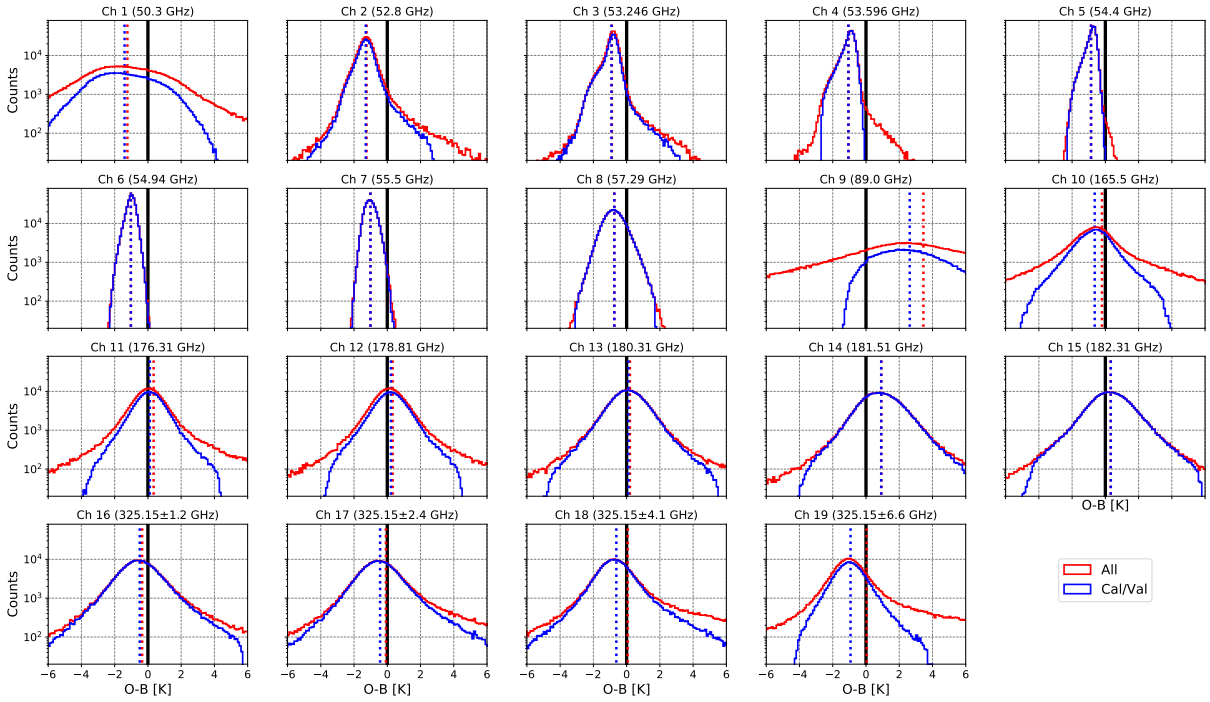

Figure 17: Histograms of un-corrected AWS departures from 28 March, 2025, data over sea-ice free ocean only. Dotted vertical lines indicate the mean of each distribution.

is not to say that there are no remaining bias structures. Here we will highlight two interesting cases in which it is difficult to disentangle potential instrument and model biases. Figure 18 shows 2D histograms of $O-B$ versus the IFS orbit angle for two channels, 7 and 14 on AWS. Previously we saw that channel 14 is warmer than its compatriot channels in horn 3 by about half a degree or more (Fig. 14), exhibiting greater inter-channel differences in bias than horns 1 or 4. When examining horn 3 biases as a function of scene temperature or orbital position, it is notable that channel 14 exhibits a greater pole to equator difference in bias than its neighbouring channels, with an amplitude of about one degree. More clearly but with a smaller amplitude, channel 7 is observed with about a 0.5 K amplitude shift in bias between equator and poles. To compare the amplitudes of biases for horn 1 and 3 AWS channels, see Fig. 19.

As interesting as these patterns appear, it is important to note that heritage sensors like ATMS exhibit similar patterns (Fig. 20). Temperature-dependent calibration biases, spectral response differences between observations and simulations, and model biases could play key roles in producing such bias structures. Generally speaking, such patterns are not a major concern in the IFS because airmass bias predictors are used in VarBC; however, recent work has shown that small orbital biases in ATMS and AMSU-A can indeed have knock-on effects (Bormann *et al.*, 2023). The orbital bias EURD requirement (00120) of < 0.2 K variations "during any single orbit" does not appear to be met by any channel in our estimation, but it is worth emphasising that such a strict orbital bias requirement is challenging to assess in absolute terms due to the potential sources of error mentioned above.

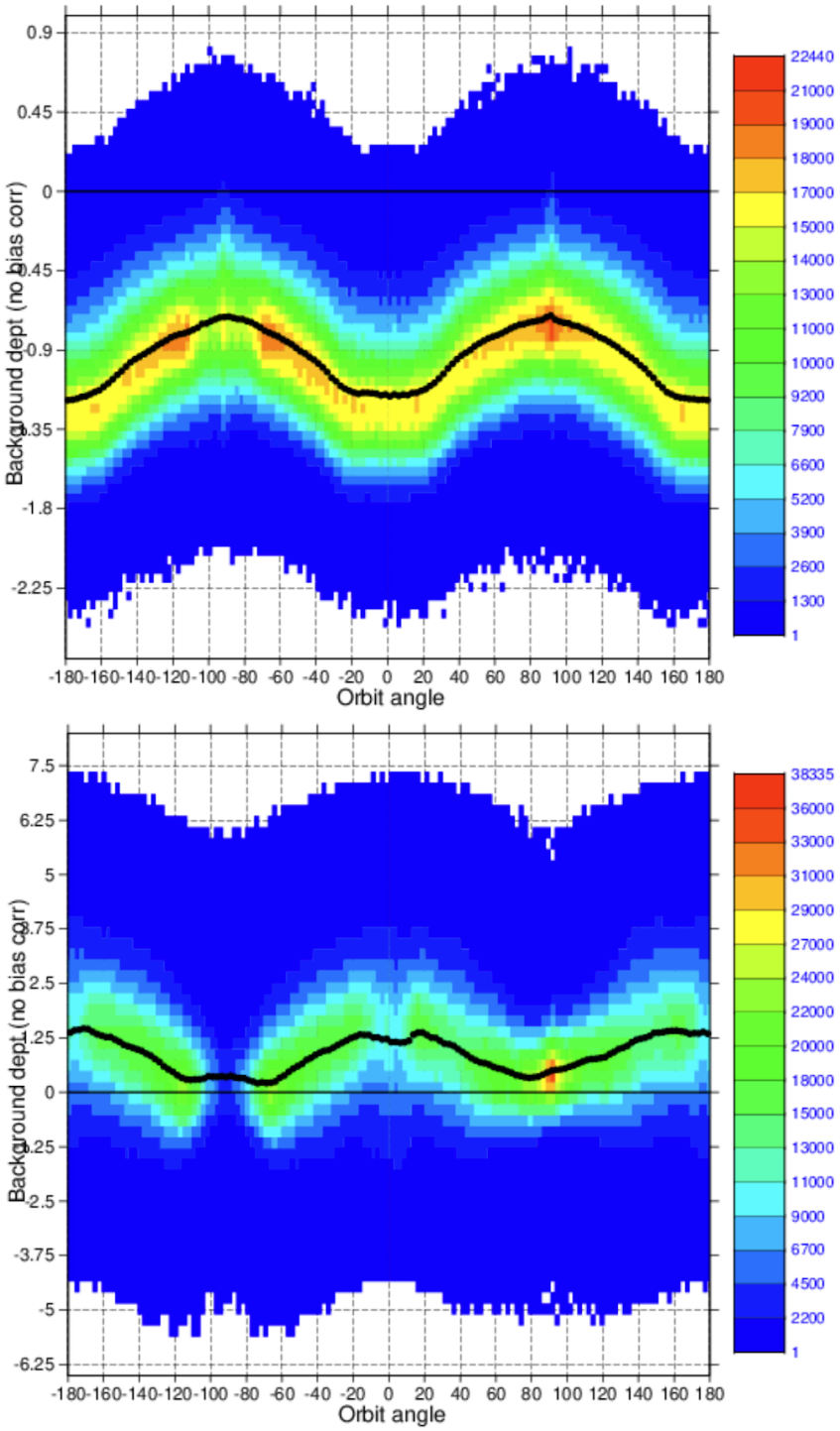

Figure 18: 2D histograms of $O-B$ for channels 7 (top) and 14 (bottom) against IFS orbital angle, using the stringent cal/val sample from mid March to mid May. Averages for each orbital angle bin are given by black dots. Note the different y-axes.

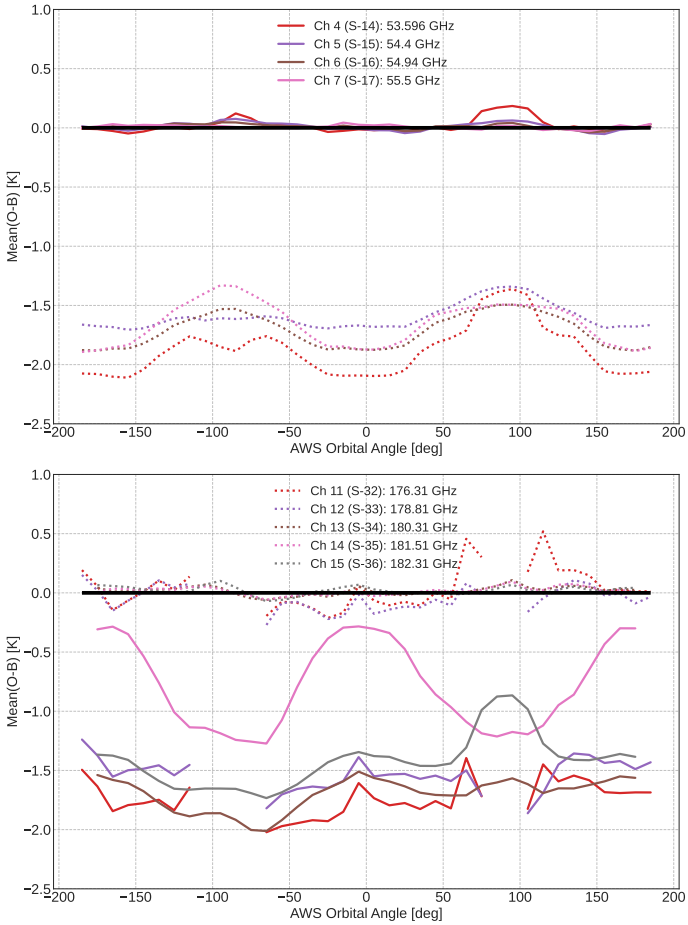

Figure 19: Biases as a function of orbital angle for selected channels from AWS horn 1 (top) and horn 3 (bottom). Data from late December 2024, stringent cal/val sample. Note that these data were prior to a calibration update from ESA and thus the biases differ from those elsewhere in the report.

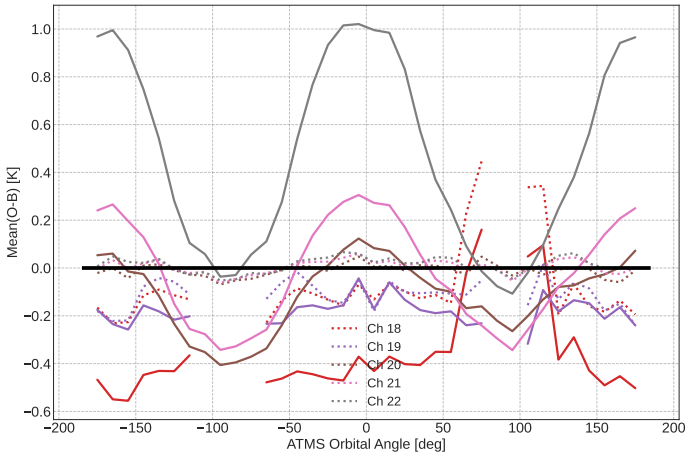

Figure 20: As Fig. 19 but for NOAA-20 ATMS 183 GHz channels.

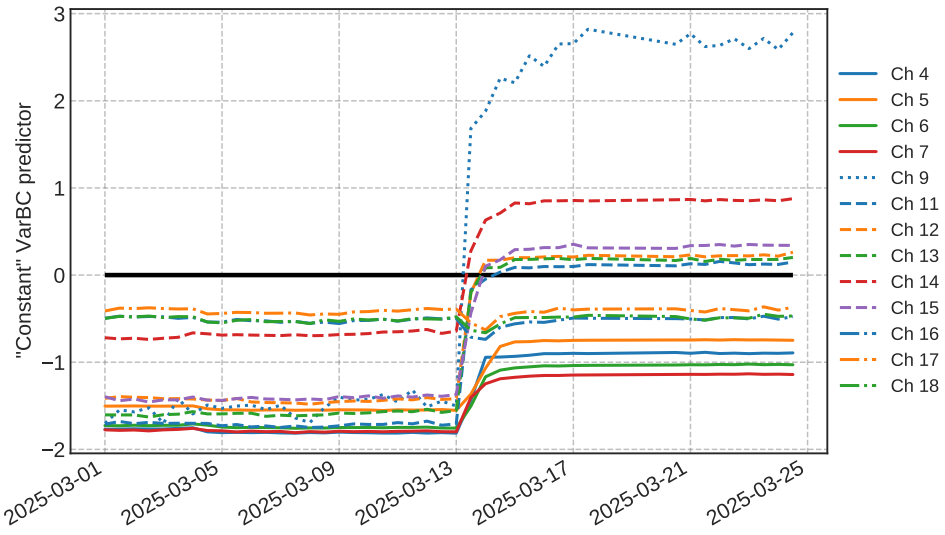

Figure 21: Time series of the VarBC constant bias predictor for assimilated AWS channels and channel 9, March 2025. Note the L1 calibration update from ESA was applied in mid March.

7.2 Bias variation

There are two aspects of calibration performance that are important for NWP but not explicitly captured in the EURD: the day-to-day temporal stability of biases and the temperature dependence of biases[5]. Although VarBC adjusts automatically, it takes some time to adjust and thus temporally varying biases can be a significant issue for operational assimilation. Figure 21 provides a view of the calibration stability by showing the time series of the VarBC constant predictor, equivalent to the mean channel bias relative to IFS that is not airmass- or angle-dependent. This plot contains the main L1 calibration update from ESA that occurred in mid March, but shows that biases were stable before this update in processing and have been stable since then. Note that VarBC needs a few days to adjust fully to the new calibration, and thus these days were removed from active assimilation in the assimilation trials. NEDT performance and other bias characteristics have also been stable in time but are not shown here.

For biases as a function of scene temperature, Fig. 22 shows 2D histograms from the monitoring website for channels 5 and 12. In both of these plots, there is considerable spread outside the main bullseye of the histogram, but where it is well-populated there are relatively small but consistent increasing biases as a function of scene temperature. This signal is consistent across all sounding channels, with channels 5 and 12 being representative. As with the orbital biases, some or most of this temperature dependence is removed by VarBC in the assimilation (not shown), but it would be good to understand this source of bias better, including how it interacts with the orbital bias noted above.

[5]Short-term stability is covered by 00120 on a per-orbit basis. Temperature dependence is implicitly covered by 00110, in that the overall bias has to be within 1K over the full Earth Scene Dynamic Range.

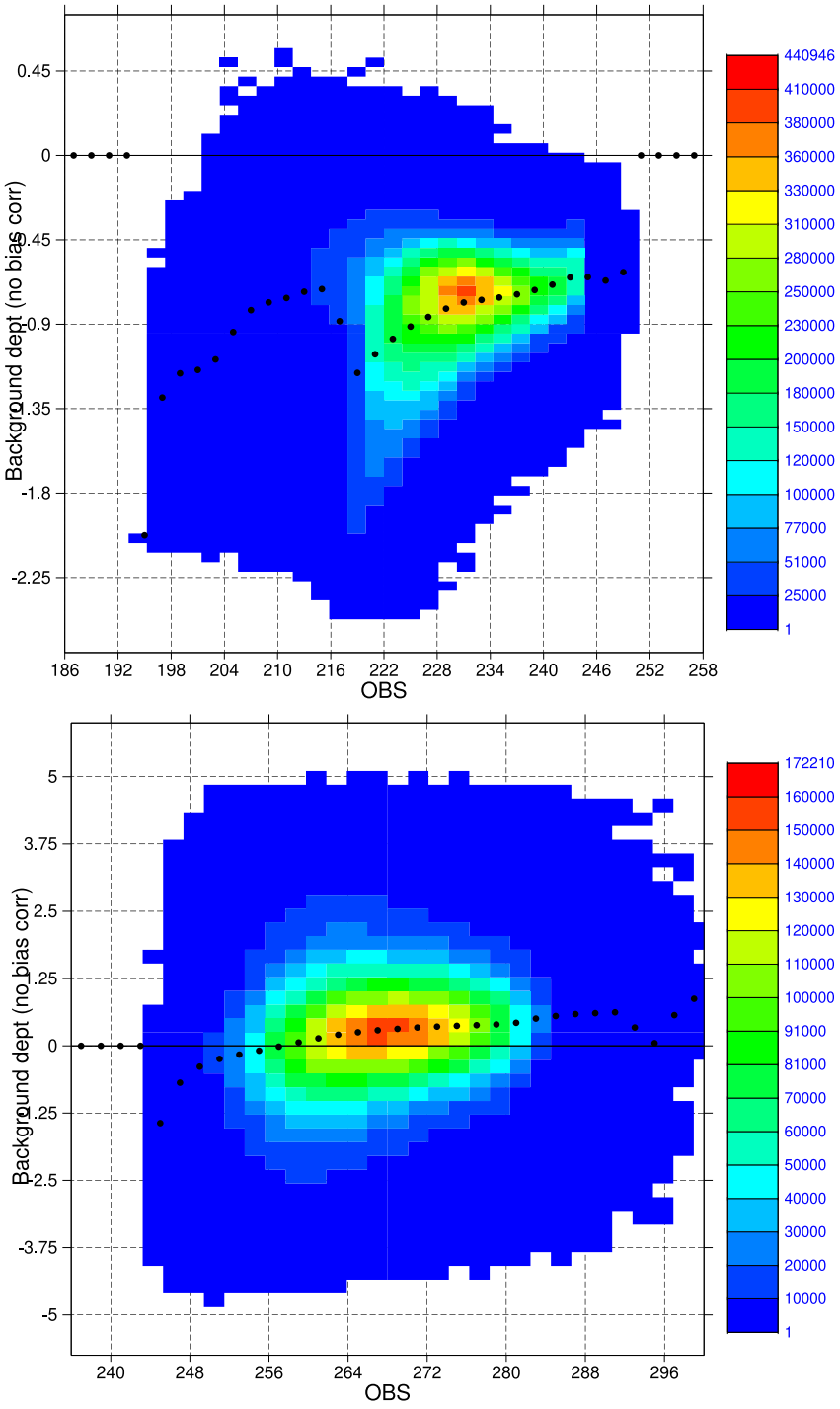

Figure 22: As Fig. 18 but for channels 5 (top) and 12 (bottom) as a function of observed TB in Kelvin.

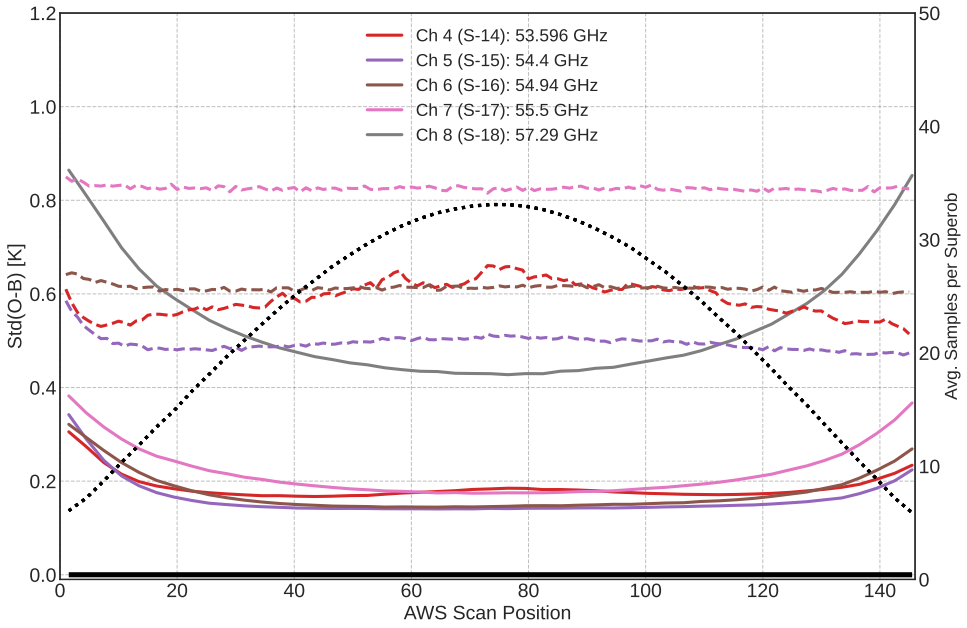

Figure 23: Standard deviations of bias-corrected departures ($std(D_{BC})$) for superobbed data at 50 km resolution (solid lines) and un-superobbed data (dashed lines) for selected horn 1 channels. The superobbed data are from April 4 to May 3, 2025; un-superobbed data are from March 16 to 19. Superobbed data have the stringent criteria applied, whereas un-superobbed data have a relaxed CI threshold due to their greater variability. The dotted black line shows the average number of observations per superob.

7.3 Radiometric noise performance

Due to the low background errors of temperature-sensitive channels in the 50-60 GHz band, NWP models can act as a check on not only radiometric bias but also radiometric noise, with background errors of order 0.1 K (Bell *et al.*, 2008; Bormann *et al.*, 2013). As a case in point, some channels of ATMS on NOAA-20 and -21 (7-9, equivalent to AWS 5-7) achieve $std(O-B)$ of 0.10 K or less in ECMWF operations after 3x3 superobbing, which is only achievable from a combination of low radiometric noise and small background errors in temperature. Analysis of $std(O-B)$ can thus be used to identify effective instrument noise at temperature sounding channels; in contrast, humidity background errors are larger (of order 1 K), and we will therefore restrict the analysis here to horn 1 performance.

Figure 23 shows $std(O-B)$ as a function of scan position for AWS channels 4 to 8, for un-superobbed data (dashed line) and superobbed data (solid), again using the stringent screening criteria[6]. The dotted line represents the average number of observations per superob, indicating that horn 1 channels typically have about 30 to 35 AWS observations that make up one 50 km superob in the middle of the swath. Channels 6 and 7 have very little sensitivity to clouds, and thus the variability of departures is almost entirely decided by how much the superobbing procedure decreases the effective radiometric noise. In contrast, there is a slight increase in channel 4 departure variability near nadir, as this frequency is more sensitive to clouds in the lower troposphere. It was known pre-launch that channel 8 was out of spec with respect to its NEDT (Eriksson *et al.*, 2025), but it is included here as an example of how NWP-based monitoring can identify an out-of-spec channel immediately, and the clear impact of superobbing on

[6]For un-superobbed data, the *CI* criteria is far too restrictive as it throws away a large fraction of observations from noise alone. Thus the *CI* threshold was removed in this case, as especially for channel 8 any reasonable value for *CI* was still removing non-cloudy data.

Table 12: Footprint NEDT specifications from the EURD and estimates from $std(O-B)$. Note that background error contributions have not been removed from $std(O-B)$, and thus these are upper estimates of effective radiometric noise from the NWP method.

	Std(O-B)	NEDT spec.
4	0.32	0.40
5	0.28	0.40
6	0.35	0.40
7	0.48	0.50
8	1.15	0.60

effective noise (un-superobbed channel 8 is off the plot at around 2.0 K).

Also shown in the figure are dashed lines indicating the un-superobbed $std(O-B)$. These values are quite directly related to instrument radiometric noise, and indeed the values seen here are relatively close to pre-launch NEDT values after converting from the provided -3dB NEDT to sample NEDT (a factor of roughly 1.73, see Table 2). For example, pre-launch NEDT for channel 7 was given as 0.86 K (0.50 K at -3dB; Eriksson *et al.*, 2025) and here we realise a standard deviation of about 0.8 K; channel 5 exhibits the lowest radiometric noise in our analysis, with about 0.5 K standard deviation that is better than expected after a pre-launch measurement of 0.64 K (0.37 at -3dB) NEDT. It is interesting to note that Eriksson *et al.* (2025) also reports in-orbit NEDT estimates that are lower than on-ground NEDT values for most channels, and indeed the analysis here points to NEDT performance that lies between their in-orbit and on-ground estimates for channels 4 to 7. It is a pleasant surprise that channels 4 and 5 especially outperform their NEDT specifications, as these are impactful channels in all-sky assimilation when radiometric noise is low (Duncan *et al.*, 2022b; Lean *et al.*, 2023).

By converting from sample to footprint NEDT (Table 2), we can connect these results back to the NEDT requirements for horn 1 listed in Table 1. With the exception of channel 8, which was known to be outside of spec pre-launch and is confirmed to be so here, the footprint NEDT values achieved by AWS horn 1 appear to lie within the requirements. The un-superobbed $std(O-B)$ in the figure leads to diagnosed footprint NEDT estimates of 0.32, 0.28, 0.35, 0.48 K for channels 4 to 7 in order. Compared to the specified NEDT (see Table 12), each of these channels appear to be inside the requirements. The IFS-based estimated NEDT has several caveats, including that background error contributions are not entirely negligible and that cloud effects have not been removed here, so this is intended as a qualitative check on the NEDT requirement.

We can compare the relative noise performance of AWS horn 1 channels to their counterparts on ATMS and AMSU-A (see Table 5), with the caveat that AMSU-A and ATMS have very different integration times and footprint sizes as discussed in Sec. 3.3. The focus is in the middle of the swath where superobs are well-populated and the scan angles are similar enough between instruments[7]. Figure 24 allows direct comparison in noise performance of AWS channels 4 to 7 against AMSU-A and ATMS equivalents, with the same cal/val sampling applied to each. For AWS, these lines are the same as in the previous figure, simply zoomed in. Note that to zoom in, we have lost sight of two out-of-spec (and no longer assimilated) AMSU-A channels, Metop-B 6 and Metop-C 8, as well as AWS channel 8. From these plots it is clear that ATMS possesses lower radiometric noise than either AMSU-A or AWS, and that NOAA-21 ATMS exhibits lower noise than its predecessor on SNPP. Near the swath's centre, AWS effective noise after superobbing is on par or better than the AMSU-A instruments on Metop-B and -C; and indeed

[7]The outer-most scan positions represent rather different scan angles for each instrument, see Table 3.

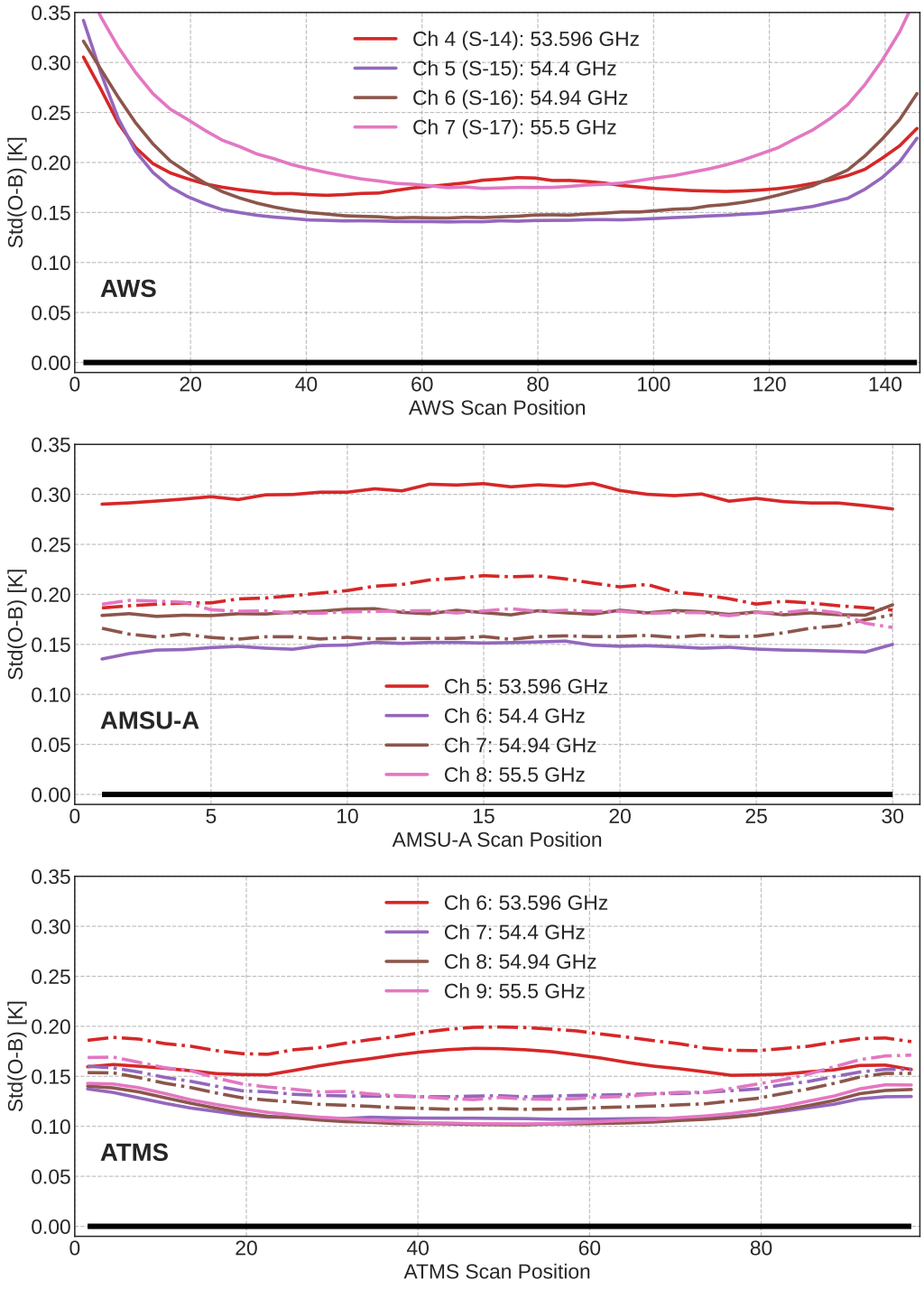

Figure 24: Standard deviations of $O-B$ after bias correction for the stringent cal/val sample for AWS (top), AMSU-A (middle), and ATMS (bottom). For AMSU-A, Metop-B (dash-dot) and Metop-C (solid) are shown; for ATMS, SNPP (dash-dot) and NOAA-21 (solid) are shown. Analysis covers global data from April 4 to 11, 2025 for ATMS and AMSU-A, and April 4 to May 3 for AWS.

AWS channel 4 achieves $std(O-B)$ comparable to ATMS channel 6 as well. Though there are subtle differences from channel to channel, and each AMSU-A or ATMS has slightly different performance, we can conclude that AWS radiometric performance after superobbing lies roughly between AMSU-A and ATMS. The comparison breaks down at outer scan positions, where AWS observes at higher incidence angles and few observations exist in a given superob for AWS.

Lastly, an aspect of noise performance that has affected some heritage sounders is striping, also known as pink or "1/f" noise (Bormann *et al.*, 2013; Yang *et al.*, 2022). In this paper, we do not quantify striping for AWS. Rather, striping was analysed early on to not be a significant concern for AWS, judged by analysis of departure statistics for un-superobbed data. However, this does not mean that there is no striping for AWS. To look at striping qualitatively in comparison to instruments with known striping issues, Fig. 25 provides global snapshots of background departures (after bias correction) for one 12 hr period in August 2025. Here we take one typical example each in the temperature and humidity bands. Significant striping is visible in Metop-C MHS channel 3, particularly in the clear skies of the subtropics. Note that this is a known feature of Metop-C MHS, whereas other MHS instruments showed considerably less striping. In contrast, the equivalent channel on AWS (15) does not show any clear striping in the departure map. From this snapshot alone, it is clear that AWS has lower radiometric noise than this particular MHS channel. For channel 9 on SNPP ATMS, an upper tropospheric temperature channel, some striping is noticeable in the departures due to the zoomed-in colour scale; the equivalent channel on AWS (7) does exhibit some striping, but the main feature noticeable here is the sub-optimal APC. A more quantitative analysis of striping for AWS could be considered in future work.

7.4 Spectral performance

Simulations of AWS radiances in the IFS use measured spectral response functions, as translated into RTTOV coefficients. As almost all MW sounders used in the IFS currently use idealised ("boxcar") SRFs, this is already a significant advantage for the simulation and assimilation of AWS. However, the possibility exists that either the on-ground SRF measurements were imperfect or that the true SRFs have shifted somehow since the measurements were taken. In either of these cases, analysis of deparature statistics should be able to shed light on gross errors in the expected spectral performance of AWS; but as stated by Lu and Bell (2014), it is difficult to separate spectral and radiometric errors. For example, apparent orbital or scene temperature biases in departures for sounding channels could be indicative of spectral mismatches between observation and simulation, as the weighting function may be placed higher or lower in the simulation than the channel is truly sensitive to.

The requirements of spectral performance for AWS (Sec. 2.2) provide strict limits on passband centre frequency shifts. Figure 26 translates the magnitudes of those passband shift requirements into top of atmosphere TB differences as simulated by RTTOV for a range of clear-sky profiles. The bottom panel shows the same analysis but for AWS with measured versus boxcar SRFs. The magnitudes of realised shifts in TB space are quite different in the two plots, with small spectral shifts leading to possibly unmeasurably small changes in the mean and variability of simulated TBs for most AWS channels. In contrast, simply using the measured SRFs for AWS led to mean differences of several tenths of a Kelvin for a few AWS channels, although for some frequency bands the measured SRF is close enough to the specification that its use in simulation makes little difference. From this analysis it is evident that small spectral shifts in AWS passbands would be undetectable in most circumstances, but it also underlines how important SRF knowledge is in accurate simulation of these bands.

Given these results, no further analysis is performed on the EURD requirements related to spectral response (00050, 00060, 00080). However, the magnitudes of these potential shifts in bias should be kept

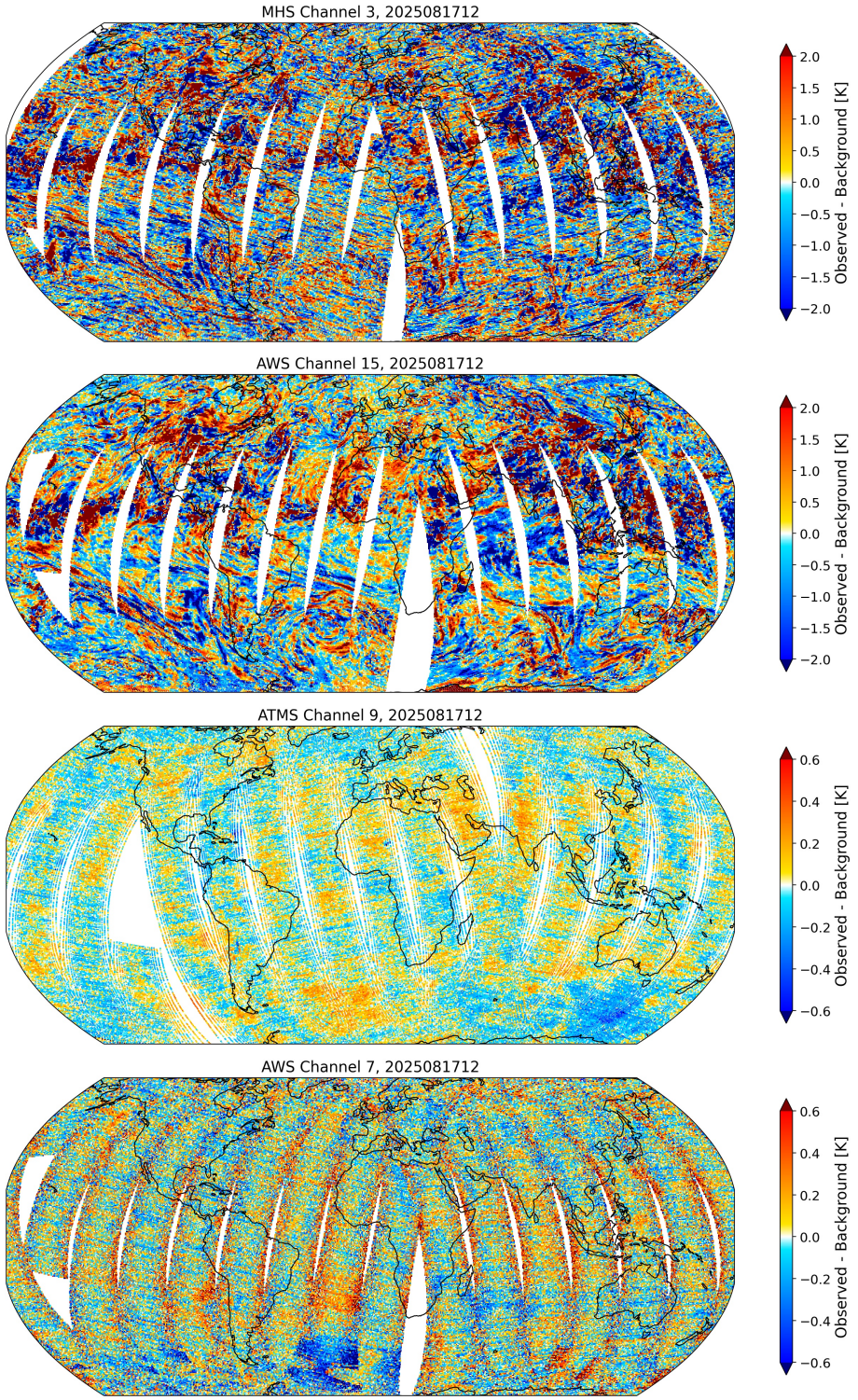

Figure 25: Bias-corrected $O-B$ for all data in the 12Z cycle on 17th August, 2025. From top: Metop-C MHS channel 3, AWS channel 15, SNPP ATMS channel 9, AWS channel 7. Note that here ATMS is from the operational clear-sky system, using 3x3 superobbing, whereas MHS and AWS are shown with 50 km superobs.

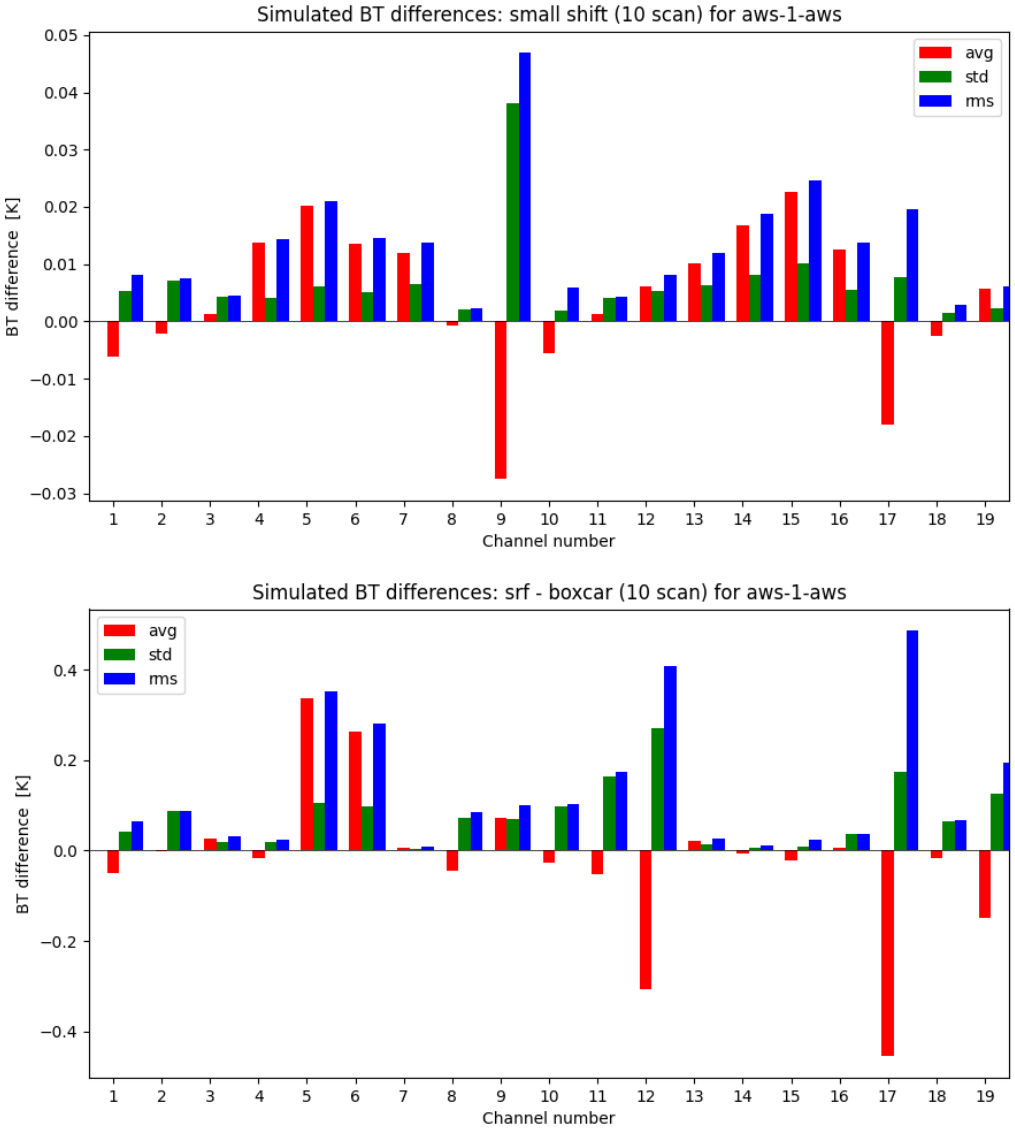

Figure 26: Effect of a small shift in bandpass centre frequency, as defined in Table 4, as simulated by RTTOV for a range of clear-sky atmospheric profiles (top panel). In the bottom panel, the same is shown for comparing measured versus idealised/specification SRFs for AWS. Note the different y-axis scales. Plots courtesy of Emma Turner.

Table 13: Restated requirements from the EURD (with numbers from the EURD in parentheses), a simplified assessment (met, partially met, not met, or unclear), and comments on this assessment.

Requirement	**Assessment**	**Comments**
Bandwidths and centre frequencies as specified (50/60)	Unclear	No indication of significant frequency shifts, though the ability to determine MHz-level shifts is quite limited
SRF shape known down to −40 dB of maximum (80)	Unclear	Analysed implicitly by use of measured SRF-based RTTOV coefficients, which showed no significant signs of deviation other than possibly channel 14
Radiometric noise (100)	Met	With exception of channel 8, horn 1 meets NEDT requirements; for other horns, NWP errors are too large to allow for quantitative assessment
Radiometric bias $<$ 1.0 K (110)	Partially met	Most channels within 1.0 K bias of reference, with channel 9 not meeting this threshold
Orbital stability $<$ 0.20 K for any single orbit (120)	Unclear	Several channels exhibit orbital bias characteristics, but these cannot be disentangled from model biases and exist in other instruments as well
Inter-channel biases $<$ 0.5 K (130)	Not met	Though largely met within horns, inter-channel biases range from roughly -1 to 3 K
Inter-sample biases $<$ 0.3 K (140)	Not met	Other than horn 4, most feedhorns exhibit biases across the scan that exceed this threshold, particularly at high scan angles for horn 1; an improved APC should allow horn 1 to reach compliance
Cross-calibration against a reference (025)	Met	The IFS is used as a stable reference, and vicarious comparison with equivalent instruments for all channels except those on horn 4

in mind when considering the biases identified in Sec. 7.1.

7.5 Cal/val summary

Linked to the EURD requirements (Sec. 2.2), Table 13 attempts to simplify the above discussion into a summary of the degree to which these requirements are met in our analysis.

8 First analysis of sub-mm departures

The sub-mm channels on AWS are an exciting addition to the global observing system, providing new sensitivity to frozen hydrometeors and previewing the even greater sensitivity expected from ICI on Metop-SG. However, for both cal/val and assimilation, these channels will only be particularly useful if we are able to simulate their properties with sufficient fidelity in radiative transfer models. This section truly just scratches the surface of this topic, and is not intended to be an exhaustive analysis. The intention is to share early results whilst highlighting the opportunities for future exploration and exploitation that exist for these novel radiance data.

8.1 Scattering signals at 325 GHz

Going back to early attempts to assimilate all-sky MW radiances, the mismatch between observed and simulated cloud signals has been a distinct challenge (Geer *et al.*, 2017). As scattering simulation capabilities have improved, it became feasible to assimilate 183 GHz channels in all-sky conditions (Geer and Baordo, 2014). To gauge the performance of RTTOV-SCATT at 325 GHz, and thereby decide whether it is worth trying to assimilate these channels in all-sky conditions, we need to look at the fidelity of simulations versus observations in the context of how well simulations perform at 183 GHz. As was seen earlier in Fig. 17, there was not a great deal of difference between PDFs of departures at 183 and 325 GHz channels with matched weighting functions. But here we want to focus on the larger scattering signals and the extremes of the distribution.

Figure 27 shows two sets of matched 183 and 325 GHz channels: 12 and 19 (low-peaking) plus 15 and 16 (high-peaking). We would expect that channel 19 features the largest TB depressions of any channel on AWS, and indeed the PDF of observed TBs has several points below 125 K, even at 50 km superob resolution. The IFS is able to simulate TBs that reach nearly as low, despite the 27 km nominal model resolution which translates to an effective cloud resolution of perhaps 80 km. Comparing channels 12 and 19, the mismatches in PDFs at channel 19 are quite similar and maybe even less severe than those at channel 12 when it comes to the highly scattering scenes. The main concerning feature in these top panels from an assimilation perspective is the middle of the PDF, where lighter scattering scenes are much more common in the model than in reality. The picture is similar for the high-peaking pair of 15 and 16, with the model able to produce scattering signals that nearly approach the largest ones in the observations, and hardly doing worse at 325 GHz than at 183 GHz. The main contrast for this channel pair is that for channel 15, the background PDF is typically not generating enough scattering, whereas at 325 GHz there is more nuance to the mismatch—not enough heavily scattering scenes but too many lightly scattering scenes. As these channels do not sense deeply into the troposphere, this likely signals a model cloud bias of too plentiful cirrus. This model bias has been noted previously in other observations and warrants further investigation with this new information from 325 GHz.

As predicted by electromagnetic theory, TB depressions caused by ice scattering will eventually saturate at higher frequencies (i.e. more ice does not lead to lower TB); higher frequencies will become more absorbing, akin to infrared radiances, and thus deep convective cores may appear warmer at sub-mm frequencies than traditional MW frequencies like 89 or 183 GHz (Geer *et al.*, 2021). To investigate this, the relative scattering signals between 183 and 325 GHz are compared in Fig. 28 for channels 12 and 19. In this case, superobbing is done at 20 km resolution, increasing the sample size whilst still allowing exact matchups between horns 3 and 4. Along the 1-to-1 line there is a high density of points as expected in clear skies, and then the distribution drops below the 1-to-1 line as 325 GHz exhibits larger scattering signals for the same scene. However, in what we could assume are the most intense deep convective

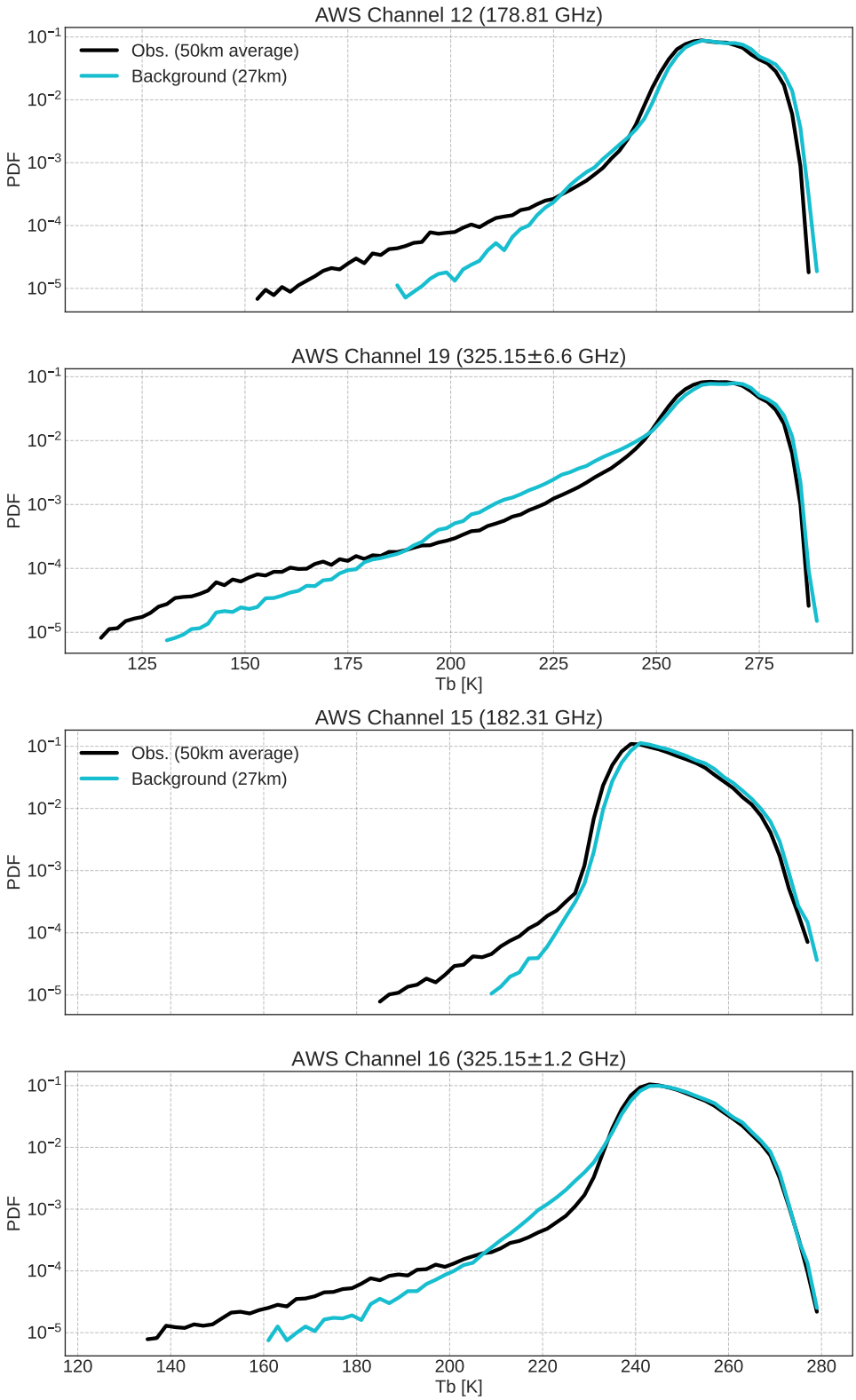

Figure 27: PDFs of observed and simulated TBs from AWS at channels 12 and 19 (top two panels), then 15 and 16 (bottom two panels). Global data over sea with no cal/val screening applied, 28-30 December, 2024. Observations are at 50 km superob resolution and the model horizontal resolution is 27 km.

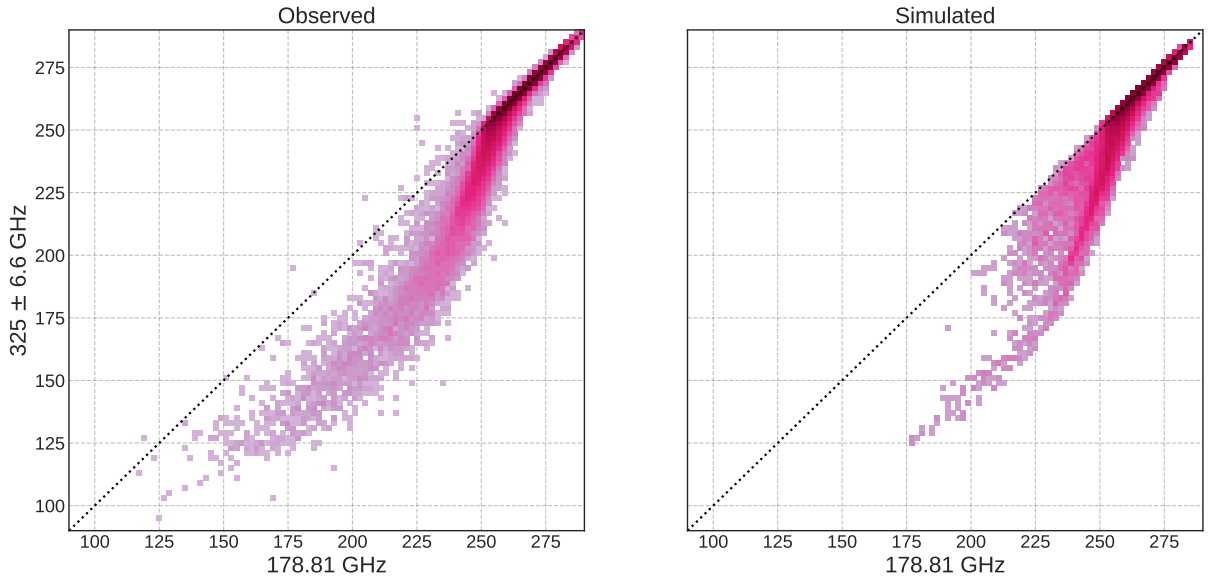

Figure 28: Histograms of observed (left) and simulated (right) TBs from AWS at channels 12 and 19, which have roughly equivalent weighting functions. Frozen surfaces (surface temperature less than 278 K) and model sea-ice are excluded. Data averaged at n512 superob (20km) to permit matchups and limit smearing from larger superob resolution. Global data from 28-30 December, 2024.

cores with TB depressions of well over 100 K, the observations show that the depressions start to come back together to similar values. The simulations in the right panel do capture some of the key features in the observed 2D histogram, but do not have cold enough TBs at either frequency.

Overall, this comparison shows us that RTTOV-SCATT is capable of simulating 325 GHz roughly as well as 183 Ghz, indicating that these channels should be worth testing in an assimilation context. There are clearly some elements of the microphysical assumptions in the radiative transfer and the model physics that could be investigated further with this new information in hand.

8.2 Case study: Typhoon Yinxing

Some provisional early data from AWS in November of 2024 perfectly captured Typhoon Yinxing in the middle of its swath, with the typhoon located south of Hong Kong as a very powerful tropical cyclone. Rather than examining statistics of observations and simulations, this offers a nice visual example to see what 325 GHz can show us and how well the model can simulate this already.

Figure 29 shows AWS channel 19 observations over the typhoon at full resolution (i.e. not superobbed) compared with model simulations from the high-resolution (9 km) model, initialised from the operational analysis. Firstly, the full resolution of AWS produces a stunning view of the eye of this tropical cyclone, with a very well defined eyewall and staggering differences in TBs within a short distance. The eye of the cyclone is about 270 K, whereas roughly 50 km away there are TBs down to about 120 K, a scattering signal of roughly 150 K. The high-resolution IFS is able to produce a cyclone quite close to the observed location with a very strong circulation as it was rapidly intensifying, even displaying isolated convection to the east of cyclone. The deepest simulated scattering signals indeed approach those of the observations. The message from this case study closely matches the conclusions reached in the above analysis, in that IFS can produce very low TBs at 325 GHz but not quite as many, whereas

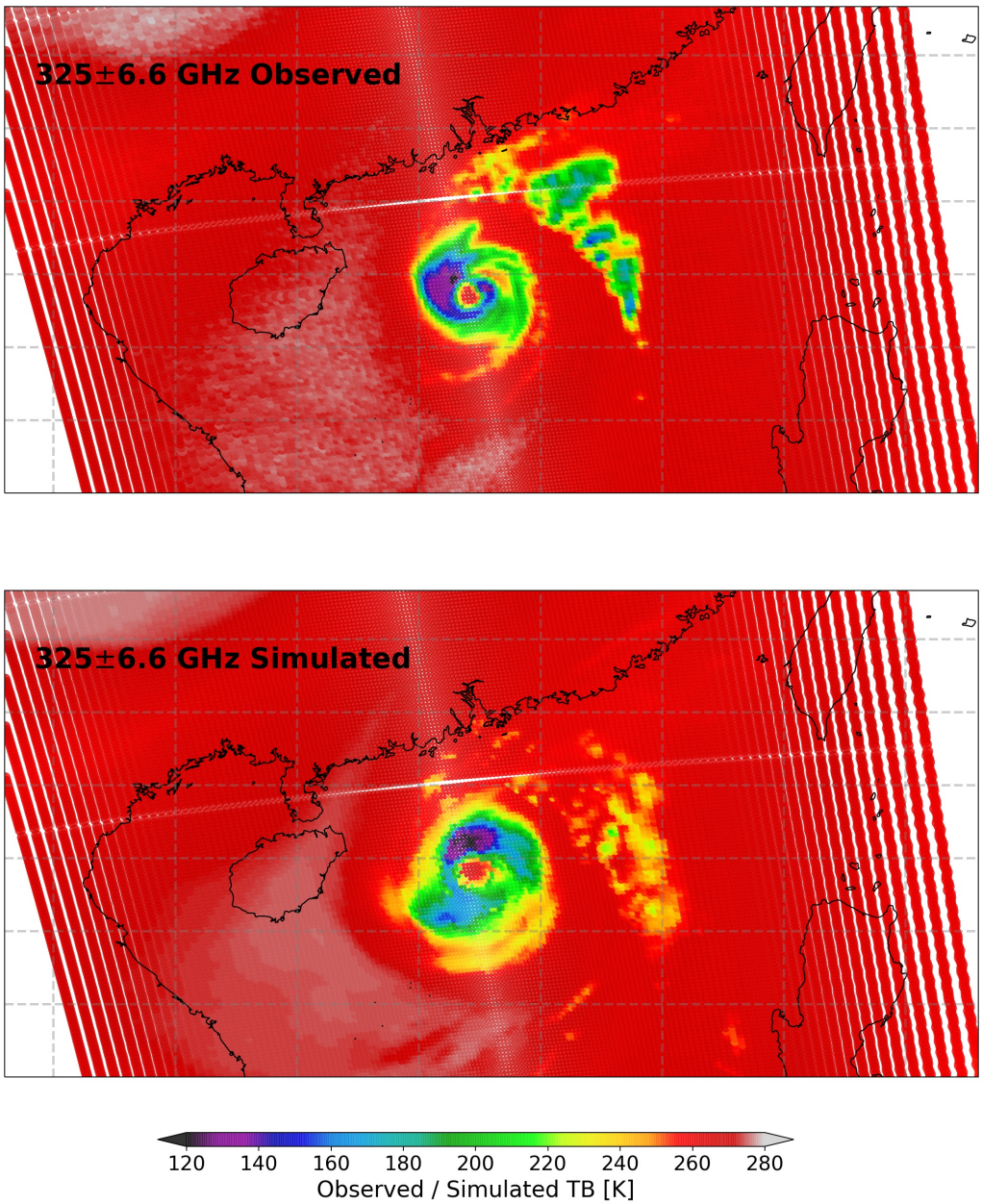

Figure 29: Observed (top) and simulated (bottom) AWS channel 19 (325 ± 6.6 GHz) over Typhoon Yinxing, 9th of November, 2024. The IFS simulation is from the 9 km (HRES) model background, initialised from the operational analysis.

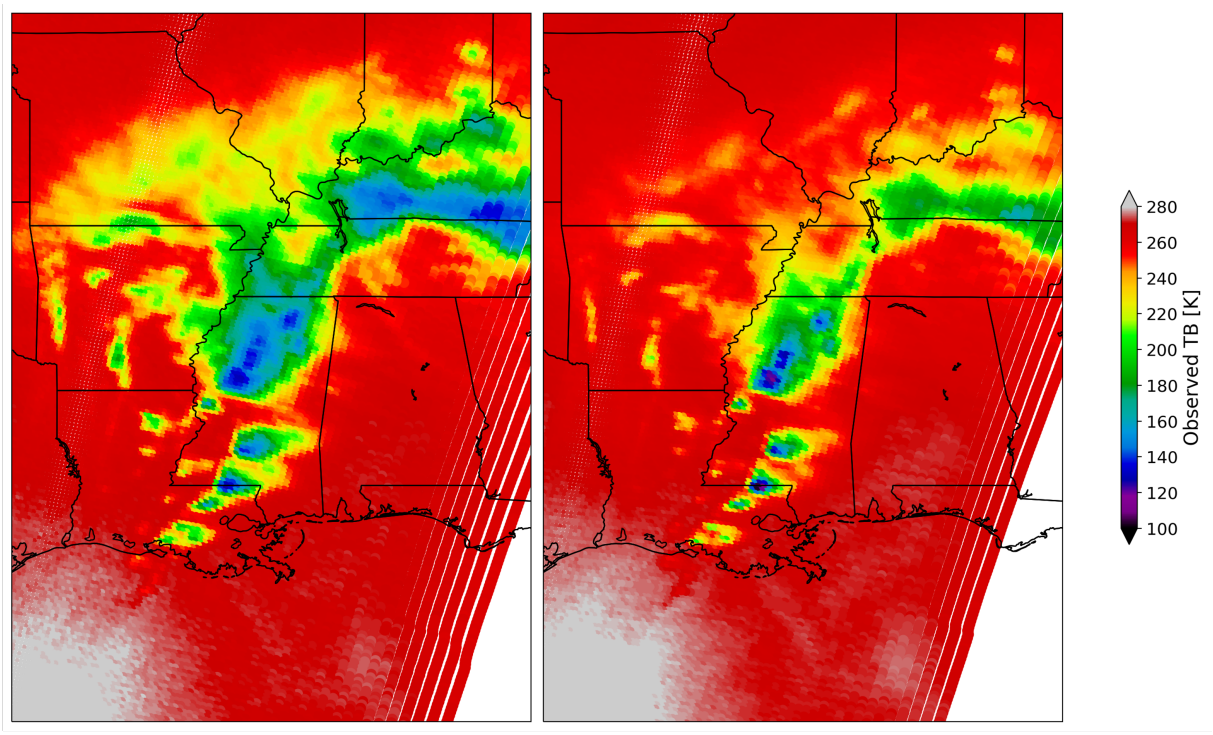

Figure 30: Observed TB from AWS in 12Z LWDA cycle, 15th March, 2025 over the southeastern United States. Channel 19 (left) and 12 (right) are shown, contrasting 325 and 183 GHz scattering signals. These are full-resolution radiances with no superobbing applied.

it produces too many medium scattering signals. The cirrus shield appears too large in the simulations, compared with a tighter upper-level cloud field in the observations.

8.3 Case study: Supercells

In mid March of 2025, several destructive supercells ripped through the southeastern United States, leading to multiple strong tornados and several fatalities. Figure 30 shows the difference in sensitivity observed by AWS 325 and 183 GHz channels, again comparing 19 and 12. In this case we compare only the observations, to show the sensitivity benefits that are realised from the addition of sub-mm radiances.

The total area in which cloud signals are obvious is much larger at 325 GHz, with clear scattering from clouds that extends to the middle of Missouri and southern Illinois. At 183 GHz there are some areas in which thicker cirrus may be present, but it seems that 183 GHz is mostly picking up the precipitation signals alone. Many of the precipitation features also show a larger signal at 325 GHz, for example the horizontal band in the northeast of the figure that spans the Tennessee and Kentucky border. There is more variability seen at 325 GHz, with TBs as low as about 130 K, contrasted with more uniform TBs at 183 GHz of about 200 K.

Returning to the point of Fig. 28 above, the type of scene in which we might expect the sub-mm scattering signal to be less than that of a lower frequency could be one with very extreme convection. This is indeed what we appear to witness in the convective cores captured by AWS in this scene, with colder TBs seen in the convective cores near the Louisiana/Mississippi border at 183 GHz. These supercells were indeed extreme, and this is not the type of behaviour that global models will be able to simulate. But digging

further into case studies such as this could help to constrain appropriate choices for scattering simulations in the future.

9 Assimilation results

In this section, results are presented for the first attempts to assimilate AWS in the ECMWF data assimilation system. The period covered begins January 1, 2025, shortly after an early stream of L1 data was made available by ESA to early adopters. AWS L1B data were made public by ESA via EUMETCAST in mid April, 2025. Due to the calibration update from ESA in Mid March, AWS data were set to passive in the assimilation for a few days to allow VarBC to adjust.

The experiments described herein use the ECMWF model version operational as of mid 2025, that is IFS Cycle 49r1[8]. Experiments were run at Tco399 horizontal model resolution (approximately 27 km) with 137 vertical levels, using incremental 4D-Var data assimilation with a final inner loop resolution of 80 km.

The main results in this section are from a set of experiments in which the POES instruments (AMSU-A on NOAA-15, -18, -19; MHS on NOAA-19) are denied, but otherwise featuring a full observing system as used in ECMWF operational analyses. The reason for this is that POES were switched off by NOAA in mid June 2025, and thus a POES denial experiment provides a realistic baseline to which AWS could be added in an operational context later in 2025 and beyond. Parallel experiments were run in which POES data were not denied, and ultimately the medium-range forecast impact was quite similar to that presented here, so those results are omitted for the sake of brevity.

Unless stated otherwise, the addition of AWS means that channels 4-7 and 11-18 are all assimilated following the observation error models and settings described in previous sections of the report. This means that three sub-mm channels are assimilated from AWS. For all AWS channels, they are assimilated over both land and sea with limited exceptions for sea-ice, frozen surfaces, high orography, and so on, depending on the channel. In the final section (9.4), we will explore the contribution of the 325 GHz channels to these impacts in a reduced observing system to accentuate the relatively subtle signals.

9.1 Impact on short-range forecasts

The best way to judge short-range forecast impact is by analysing the change in background fits to independent observations. Figure 31 shows the normalised change in $std(O-B)$ for six different observation types spanning temperature, humidity, and wind observations. Here the 100% line represents the control experiment (i.e. without AWS assimilated), and thus points to the left of this indicate an improved fit between the observations and the model background after adding AWS to the assimilation, i.e. an improved short-range forecast. AWS clearly improves the short-range forecasts of humidity, with decreased $std(O-B)$ seen for ATMS humidity channels 18-22, IASI's humidity-sensitive wavenumbers of 1367 to 1994, water vapour channels from GOES-18 and Meteosat-9, and humidity channels on GMI around 183 GHz. There are also some small but significant improvements in fits to wind observations as seen in panel (c). The impact on short-range temperature forecasts is relatively small, with an overall neutral change seen in hyperspectral infrared instruments like IASI but some slight improvements seen for radiosonde temperature and ATMS channels 6 and 7.

[8] https://www.ecmwf.int/en/publications/ifs-documentation

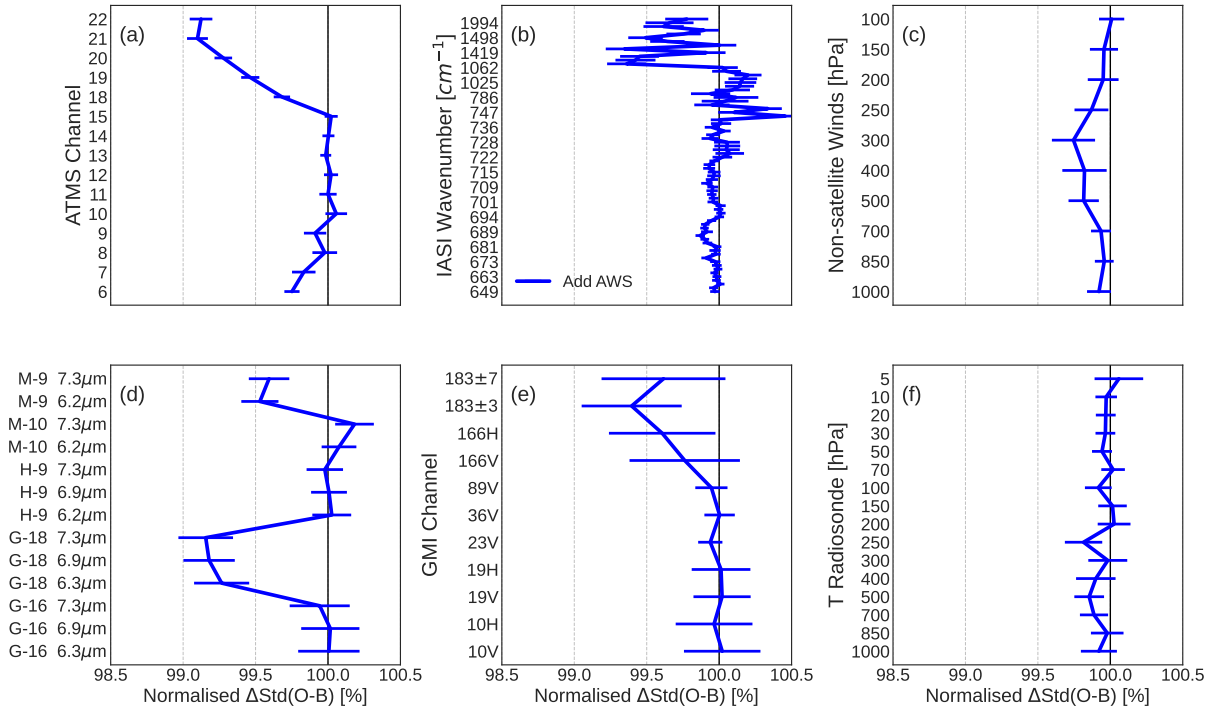

Figure 31: Changes in std$(O-B)$ for independent observations from ATMS (a), IASI (b), non-satellite winds (c), geostationary infrared radiances (d), GMI (e), and temperature radiosondes (f). Confidence intervals for 95% statistical significance are given in horizontal bars. Global data from Jan 1 to May 3, 2025, showing the impact of assimilating AWS.

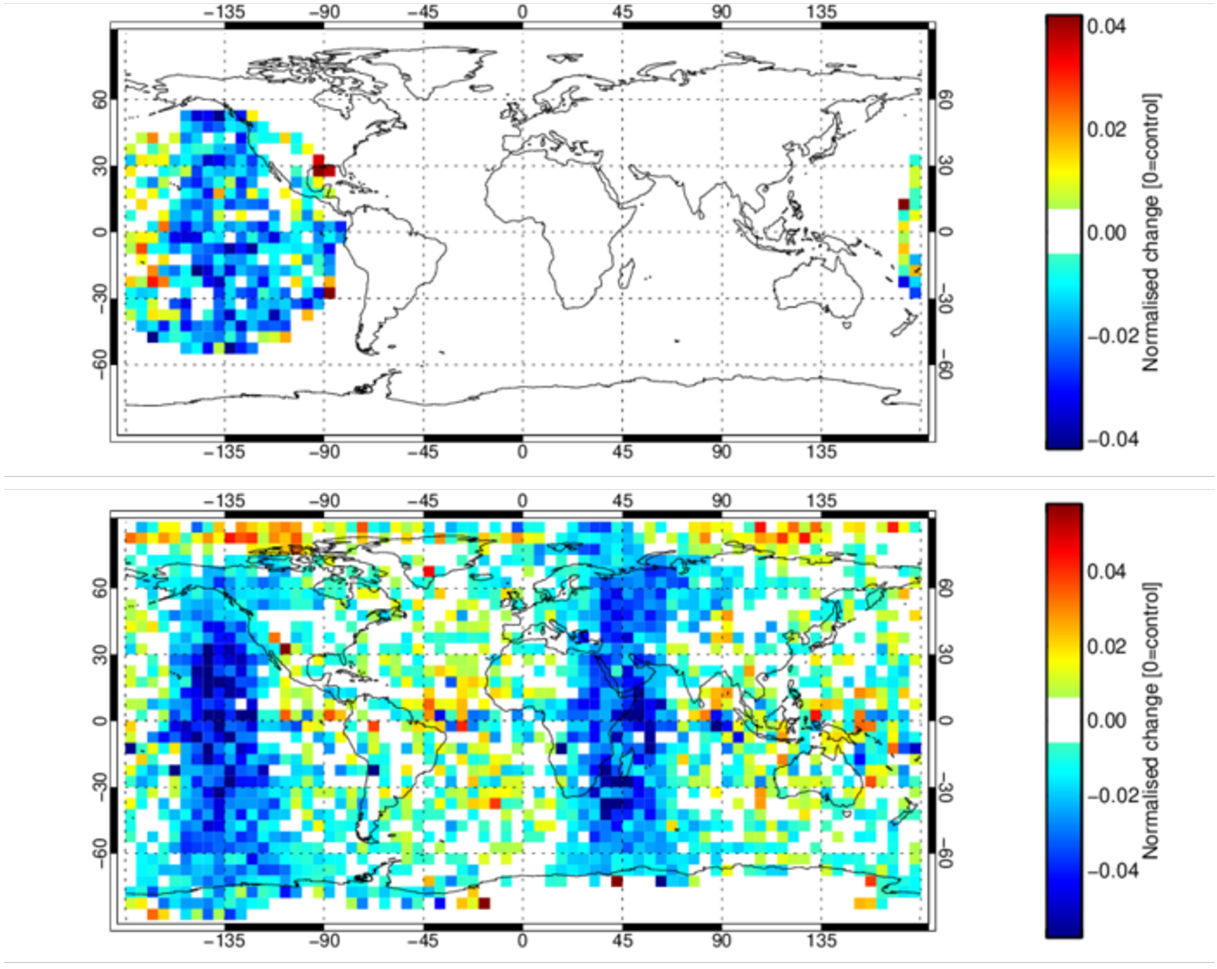

Figure 32: Normalised changes in $std(O-B)$ for GOES-18 7.3µm band (top) and ATMS channel 20 (bottom), shown averaged on a 5 degree global grid. Global data from Jan 1 to May 3, 2025 showing the effect of assimilating AWS, where blue is a positive impact.

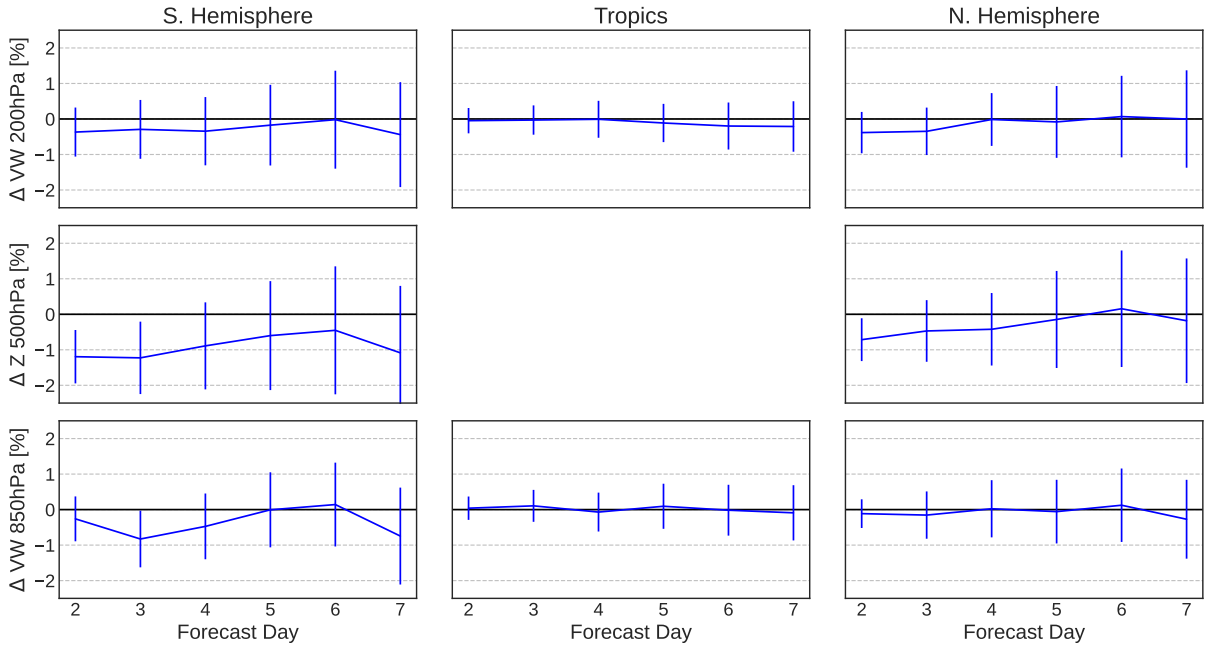

Figure 33: Changes in RMSE for forecast scores at day 2 to 7 caused by assimilation of AWS. Shown are vector wind at 200 hPa (top), geopotential height at 500 hPa (middle), and vector wind at 850 hPa (bottom) for the southern hemisphere (90S to 20S), tropics (20S to 20N), and northern hemisphere (20N to 90N). Vertical lines indicate 95% confidence intervals following Geer (2016), with verification against the operational analysis. Geopotential scores in the tropics are not informative and hence not included.

Not every signal is positive, with slight degradations seen for example in Meteosat-10 7.3μm, but the general picture here is consistent with activations of previous MW sounders on heritage platforms. Due to the large number of humidity-sensitive channels on AWS (five at 183 GHz plus three active at 325 GHz), it is not surprising to see that the main impact from AWS is on humidity. An interesting feature here is that $std(O-B)$ for GOES-18 and Meteosat-9 are quite significantly decreased but impacts are neutral for the others in the "geo ring." This is investigated in Fig. 32, showing mid-tropospheric channels sensitive to humidity from GOES-18 and ATMS, with strong improvements in background fits to observations in two distinct longitude bands. This appears directly linked to the orbital crossing time of AWS (10:35 LTAN), which provides extra constraints for humidity late in the 12 hr assimilation window in a timeslot that is otherwise poorly populated at the longitude bands these two geostationary satellites oversee. It is hypothesised that if AWS were in a well-populated orbital crossing time such as 9:30 or 13:30, the impacts would be more muted than what we realise with the 10:35 LTAN.

9.2 Impact in medium range

Whereas fits to observations are best for measuring short-range forecast impact, to analyse the medium-range impact we examine analysis-based forecast verification. In this case we will use verification against the operational ECMWF analysis, but the results look quite similar beyond day 1 if using own-analysis verification. Here we are examining just over 4 months of experimentation, which is borderline for truly determining medium-range forecast impacts beyond the null hypothesis for impacts of the expected magnitude (Geer, 2016), but appears sufficient in this case as indicated by the significance testing at the 95% confidence level.

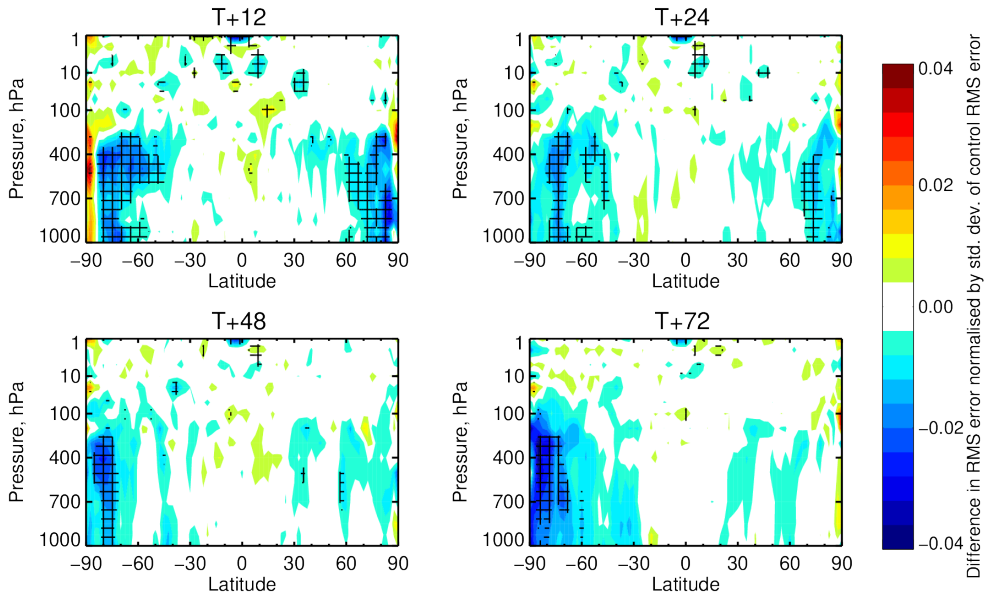

Figure 34: Change in vector wind RMSE when adding AWS to the assimilation, compared to a control with a full observing system but POES instruments removed. Period from Jan 1 to May 3, 2025. Verification against the operational analysis given at 12, 24, 48, and 72 hrs forecast lead time. Hatching for statistical significance at the 95% confidence level follows Geer (2016).

Figure 33 presents the changes in RMSE for winds at two levels (850 and 200 hPa) and geopotential height at 500 hPa. Significant improvements in forecast skill can be seen in the Southern Hemisphere through day 3, and a small but significant improvement in Z500 in the Northern Hemisphere through day 2. These impacts are roughly in line with the impacts seen from adding another MWHS-2 into the IFS (Duncan and Bormann, 2020; Steele *et al.*, 2023), but with a slightly larger impact on temperature. Impacts are comparable but in general larger than those of adding NOAA-19 (one AMSU-A and one MHS) to the POES-denial control experiment (not shown). This indicates that AWS is able to positively affect medium-range forecast skill in a similar way as heritage MW sounders assimilated in the IFS.

To examine the geographic distribution of these impacts, Fig. 34 shows the normalised change RMSE for winds at four different lead times from short- to medium-range. These plots indicate that the majority of the forecast improvements from AWS are manifested at high latitudes in both hemispheres. From the free troposphere down to the surface, winds are improved by about 1 to 3% at short-range near the poles and some of this signal lasts through to day 2 and 3 forecasts. Similar signals are witnessed in verification of humidity forecasts (not shown).

9.3 Forecast sensitivity to observation impact (FSOI)

A further way to quantify the impact of assimilated observations is through the adjoint-based Forecast Sensitivity to Observation Impact or FSOI (Cardinali, 2009). FSOI estimates whether each observation reduced or increased forecast errors at a 24 hr lead time based on analysis verification with respect to a dry total energy norm. A short FSOI experiment was run for the month of January 2025 to examine the relative impact of AWS in the overall observing system. Here we will only present the relative percentage of total FSOI for assimilated radiance instruments, and this is necessarily dependent on the rest of the global observing system. Note that FSOI statistics give only broad estimates of short-range forecast

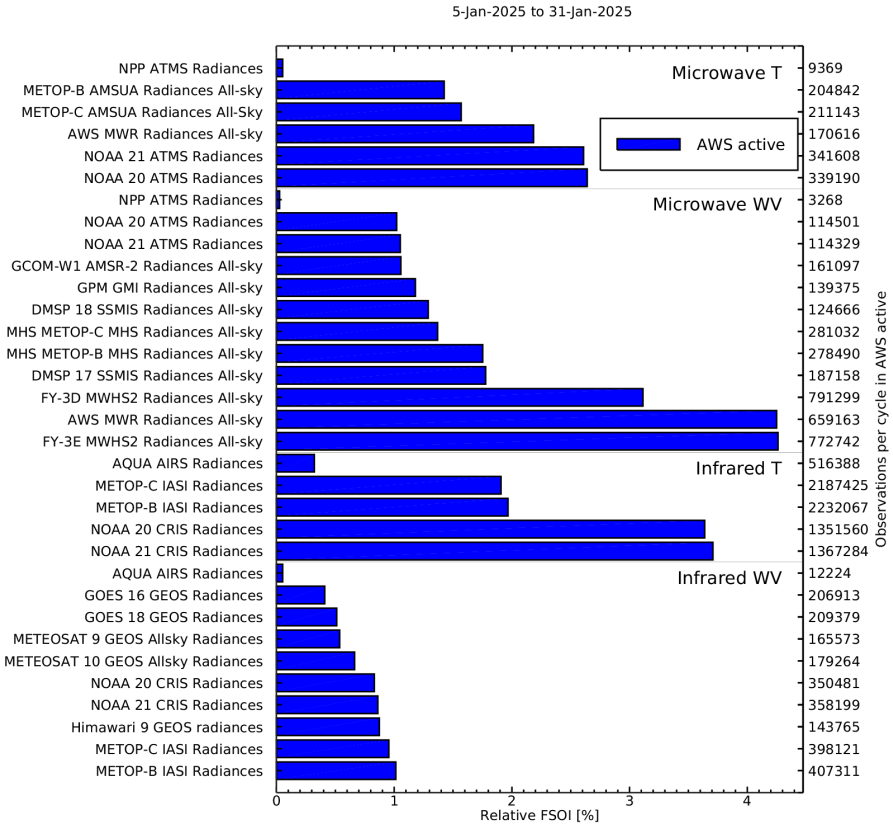

Figure 35: FSOI per satellite instrument from which radiances are assimilated, January 2025. Note that this experiment denied the POES satellites (NOAA-15, -18, and -19) though they were assimilated operationally; also note that SNPP had an extended data outage during this period.

impacts of contributing observations in the presence of the full observing system used, and limitations of FSOI and its complementarity to Observing System Experiments are further discussed by Healy *et al.* (2024).

Figure 35 plots the relative percentage of total FSOI that is attributed to each satellite instrument from which radiances are assimilated. As a convention, here FSOI is split up by temperature and humidity sounding channels, and thus AWS channels 4 to 7 are treated under "Microwave T" and channels 11 to 18 under "Microwave WV" for water vapour sensitivity. For the temperature sounders, AWS ranks between AMSU-A and ATMS, though it is worth noting that SNPP suffered an extended data outage that covered most of this time period. Amongst humidity-sensitive MW instruments (including MW imagers), AWS is one of the most impactful MW instruments assimilated in this time period according to FSOI. Like the current top FSOI instrument, MWHS2 on FY-3E, AWS features a unique orbital crossing time and has a full suite of humidity sounding channels, likely contributing to this result[9].

In terms of the spatial distribution of FSOI for AWS, this is shown on a map in Fig. 36. It is interesting to compare this map with the change in ATMS $std(O-B)$ shown in Fig. 32, as the longitudinal bands centred near 135W and 45E are again the regions of strongest impact according to FSOI. This figure also shows that the majority of impact from AWS is at higher latitudes, consistent with the analysis-based forecast verification seen above. Though not shown here, this pattern of strong FSOI results in discrete

[9] 118 GHz channels from MWHS2 are treated as humidity-sensitive frequencies in this analysis.

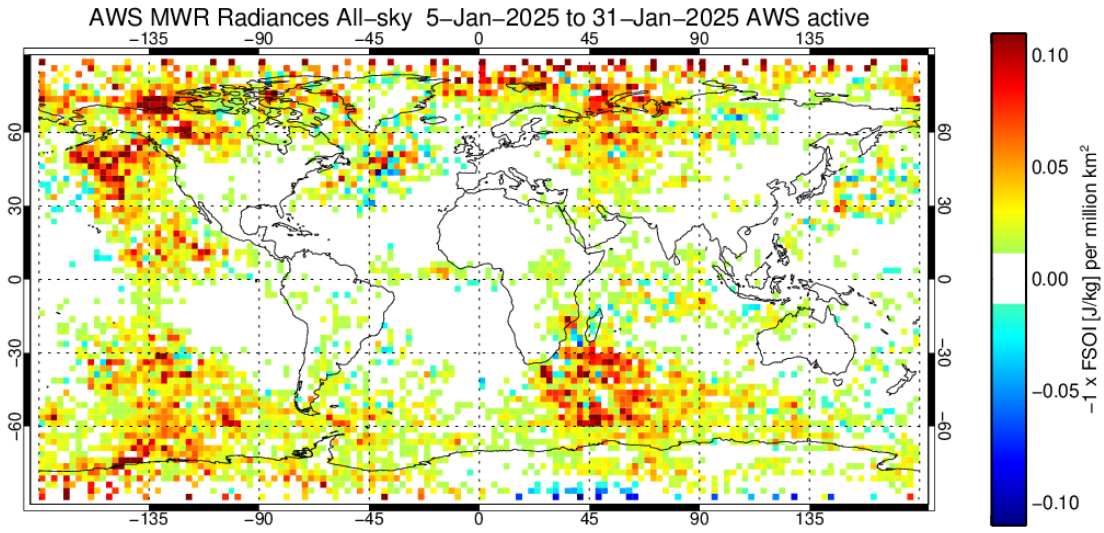

Figure 36: FSOI from all AWS channels integrated together, January 2025, shown on a 5 degree global grid. Red colours indicate good impact.

longitudinal bands is also seen for FY-3E MWHS-2, which also has a unique orbit relative to other MW sounders assimilated in the IFS.

9.4 Extra impact from 325 GHz

The 325 GHz channel suite is a key novel aspect of AWS, and users may wonder whether there is additional benefit from these channels on top of the humidity sounding information content already available from the 183 GHz channels. Previous work suggests that the 325 GHz channels on their own can provide a similar impact as a comparable set of 183 GHz channels, whereas additional benefit from adding these channels on top of the 50 and 183 GHz sounding channels may be more difficult to achieve and confined to limited benefits for humidity (Lean and Bormann, 2024). In this subsection, we are interested in the additional assimilation benefit achievable from including the 325 GHz channels, rather than comparing the relative impacts of the 183 and 325 GHz channel suites. Thus to isolate the signals of 325 GHz assimilation and hopefully yield a larger signal to noise for the impact, some extra experiments were run in which humidity-sounders and related channels (183, 166, and 118 GHz) for other passive MW instruments were removed from assimilation. The Control for this set of experiments excludes MHS and MWHS-2 entirely and also removes the 166–183 GHz channels from GMI, SSMIS, and ATMS. The Control does include AWS, assimilating channels 4-7 and 11-15. The AWS 183 GHz channels were left in active use because part of the hypothesised benefit of 325 GHz is in combination with the 183 GHz channels, to better differentiate between hydrometeor and humidity signals (e.g. Kaur *et al.*, 2021). Otherwise, the Control uses the full observing system. Two experiments were run: one in which AWS channels 16-18 were added, and one in which AWS channels 16-19 were added. Experiments ran from January through early May.

As before, we can investigate the short-range forecast impact by analysing the background fits to independent observations. Figure 37 shows a similar set of observation types as before, but now including AMVs rather than GMI. We can see that the sub-mm channels from AWS provide some distinct benefits

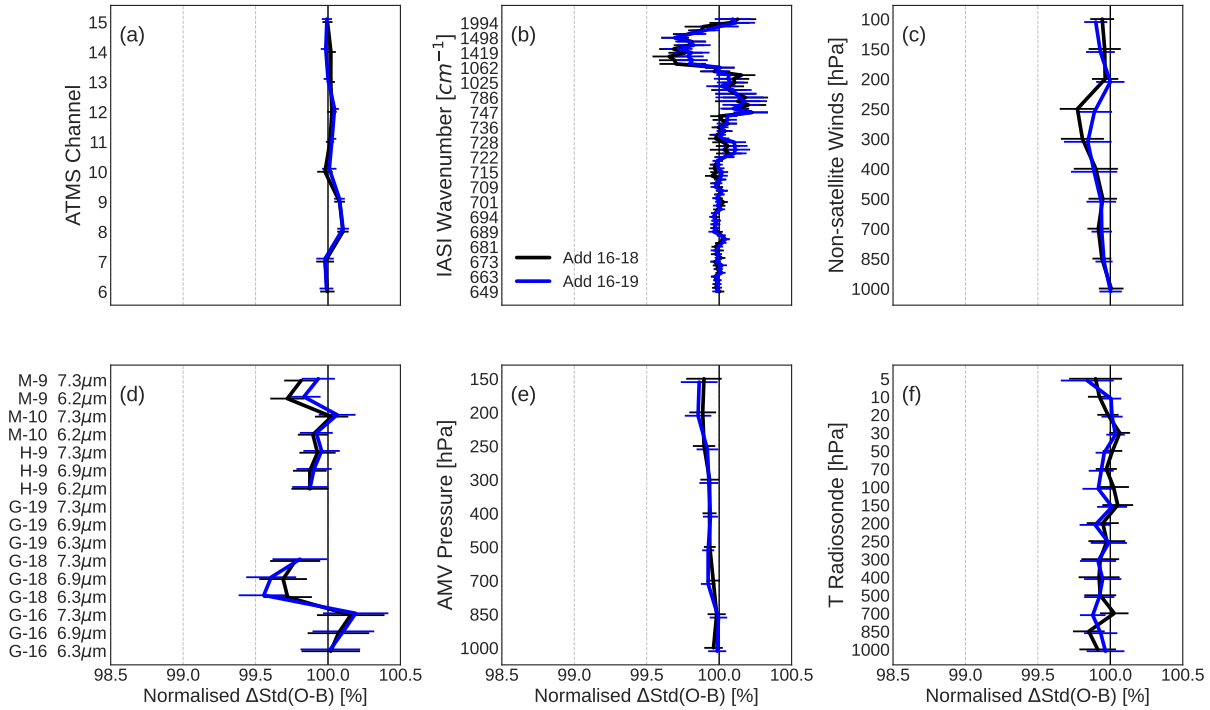

Figure 37: Change in $std(O-B)$ for independent observations caused by assimilation of AWS 325 GHz channels from a baseline with no humidity sounders and AWS 4-7 and 11-15 assimilated. Observation platforms are ATMS (a), IASI (b), non-satellite winds (c), geostationary infrared radiances (d), AMVs (e), and radiosonde temperature (f). Global data from Jan 1 to 30 April, 2025.

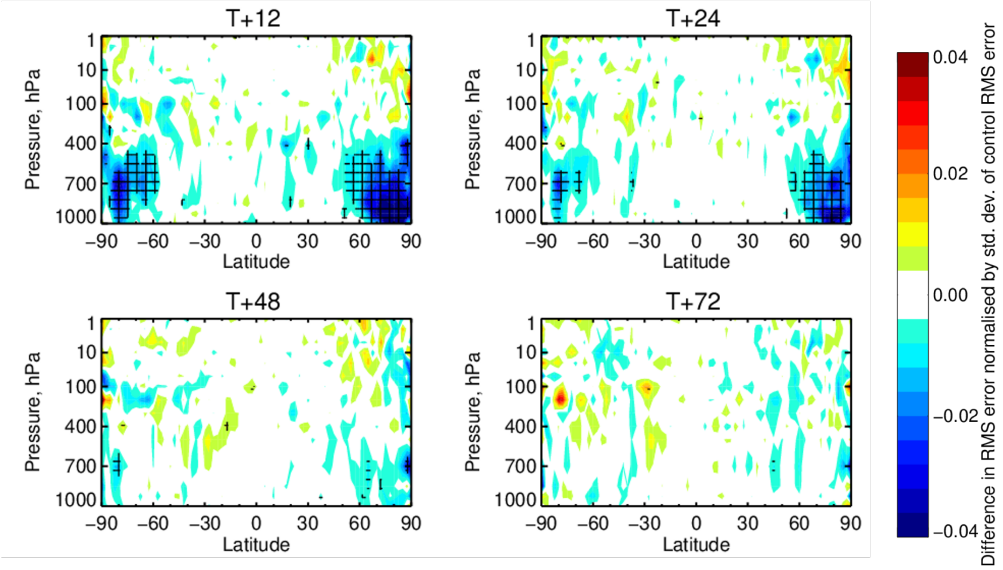

Figure 38: As Fig. 34 but for RMSE of relative humidity, showing the impact of AWS channels 16-18.

in this depleted observing system. Beyond the 50 and 183 GHz channels from AWS, the 325 GHz channels further improve the model background humidity, as seen in geostationary infrared radiances and IASI higher wavenumbers. Some very small but statistically significant improvements are seen in fits to radiosonde temperature, which is an unexpected signal of approximately 0.1% reduced $std(O-B)$ from the Northern Hemisphere in boreal winter and spring. One of the stronger signals seen here is in wind observations, with small but consistent improvements seen in AMVs and non-satellite winds. There is a small but significant degradation seen for ATMS channel 8 and 9 fits, perhaps caused by small shifts in the mean analysis state in the upper troposphere, but it is unclear where this signal comes from, as AMSU-A fits do not show the same pattern (not shown). Little benefit is realised from adding AWS channel 19, though it also does not cause any clear degradations.

Bearing in mind that this is a depleted observing system and the impact is hence larger than we should expect in a full observing system, we can also look at analysis-based verification as before. Most of the medium-range impact was neutral in this experimentation. However there are some signals of impact at high latitudes from the 325 GHz channels. Figure 38 shows the change in RMSE for relative humidity at four lead times from adding channels 16 to 18. It is striking that the benefit from 325 GHz appears to be mainly at high latitudes, and that some fraction of the total benefit seen earlier from the full AWS over the poles is indeed coming from the sub-mm channels. It is not yet clear the mechanism for these positive effects, as it could be from humidity sensitivity alone, a 4D-Var tracer effect from cirrus clouds themselves, or 4D-Var managing to better discriminate humidity and cloud signals when using 183 and 325 GHz in concert, or even an improved cloud analysis. Results for other geophysical variables are similar to the ones for relative humidity, but weaker (not shown). Overall, the finding of added benefit for relative humidity from adding the 325 GHz channels is broadly consistent with the simulation results reported in Lean and Bormann (2024) for EPS-Sterna. While the underlying observing system and the number of satellites added are different in the two studies and hence results cannot be compared directly, the qualitative consistency is a positive finding.

A great deal of further work will be needed to better understand and fully realise the assimilation potential of sub-mm channels, but as a first foray into the topic with real data, this short analysis is encouraging. It

appears that 4D-Var is already able to extract some extra information from 325 GHz channels in addition to what it wrings out of the 183 GHz channels from AWS.

10 Conclusions

This report has described the sounding instrument on board ESA's Arctic Weather Satellite, assessed its radiometric performance via comparison with a state-of-the-art NWP model, and presented results from initial attempts to assimilate these radiances in the ECMWF data assimilation system.

The NWP-based vicarious cal/val analysis has shown that AWS is a high-performing radiometer despite its small size, with biases and noise performance generally in line with heritage MW sounders such as AMSU-A and MHS. That said, some diagnosed biases appear to lie outside the rather strict specifications of the EURD, particularly for some horn 1 channels that do not meet specific inter-channel and inter-sample requirements, though standard bias correction models appear mostly adequate to correct the observed biases. Two channels' NEDT are outside the specification as was known pre-launch (Eriksson *et al.*, 2025). It is notable that the 325 GHz channels have excellent bias characteristics in our assessment despite this being a newly observed part of the spectrum. Most AWS channels exhibit close inter-calibration within a given feedhorn, though some calibration differences exist between different feedhorns that are not exhibited by a heritage instrument like ATMS. AWS biases display good temporal stability, which allows the assimilation system to correct the main biases in most cases. As of June 2025, there is work ongoing to improve the diagnosed horn 1 scan pattern biases and thus these might be improved in the L1 processing and mitigate horn 1 bias structures described herein.

It is not currently possible to assess every parameter of interest for cal/val with departure-based analysis (see Table 13 in Sec. 7.5 for a summary). For example, gross orbital biases can be quickly identified using the NWP model reference, but more subtle biases on the order of tenths of a degree can be hard to disentangle from model biases and radiative transfer errors. It is also challenging to use the NWP model to assess the full dynamic range of AWS, as the coldest observed scenes are typically ones with significant model uncertainty such as deep convection or frozen surfaces. Even the "dynamic range" cal/val sample defined here is not sufficient to capture the very large dynamic range defined in the EURD of roughly 80 to 300 K for most channels; these extremes are very rarely or never observed in most channels. This is an aspect that could be investigated further in the future, for example using scenes where both model and observations agree that a scene is fully cloudy.

Despite the above-named caveats, there is real power in the departure-based method to analyse vast amounts of data quickly with comparison to other instruments without the need for co-location. In the context of future small-satellite development and rapid feedback to data providers, the departure-based analysis developed in this project proved itself an invaluable tool for quickly pinpointing and analysing radiometric behaviour of AWS. Feedback was provided to ESA and EUMETSAT within a few days of receiving the initial on-orbit L1 data from AWS, and the near-real-time monitoring facility helped to inform responses to situations such as data outages and calibration changes. This type of facility will be crucially important for the EPS-Sterna constellation and also bodes well for upcoming EUMETSAT launches of MWS and MWI/ICI on EPS-SG.

Assimilation trials with AWS data clearly show the benefit of assimilating more MW sounder radiances, in line with previous studies using both real and simulated data. The results presented here indicate that the unique orbital crossing time of AWS is a factor in its impact in the assimilation system, again in line with earlier studies and encouraging for the range of orbits proposed as part of the EPS-Sterna constellation.

The analysis of sub-mm channels on AWS was also a focus of this report, with relatively good agreement seen between RTTOV-SCATT simulations and AWS observations. Inclusion of these channels in the assimilation on top of the 50 and 183 GHz sounding suites does show some positive impact in our experimentation, a finding congruent with Lean and Bormann (2024) but not a foregone conclusion when dealing with real observations. Years of radiative transfer developments preparing for sub-mm radiances have led to these new frequencies being more straightforward to assimilate than might have been expected. More can be done to optimise the use of sub-mm radiances in the future, but this initial work on assimilating sub-mm radiances is very promising for the upcoming launch of ICI.

AWS is the first radiometer on a small satellite with radiometric performance and stability that is sufficient for assimilation in operational NWP. As such, ECMWF began assimilating AWS radiances operationally in July 2025.

A Appendix: Applicable EUMETSAT internal documents

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

- EPS Sterna End User Requirements Document: EUM/AWS/REQ/20/1207183
- Considerations on User Requirements for Additional Mini-Microwave Sounders: EUM/STG/75/19/DOC/18

B Appendix: Online monitoring

After AWS was put into ECMWF operations in early July 2025, AWS data monitoring has been routinely available on a public website: https://charts.ecmwf.int/catalogue/packages/obstat/. The publicly available plots provide several variables including departures before and after bias correction, with time series, maps, and Hovmöller plots (as a function of latitude and scan position) to choose from.

C Appendix: BUFR Sequences

Here the internally used BUFR sequence for level 1 AWS data is shown alongside the official one from the WMO, given in Table 14. First the WMO-approved sequence for AWS is shown (310078). Second is the custom-defined sequence for internal use, featuring one geolocation for all channels and a few different BUFR entries. Note that the WMO BUFR sequence for AWS is also used for the TROPICS mission.

[10] Full element is (201132,202129,012065,202000,201000) to increase precision from 0.1 to 0.01 K.

Table 14: BUFR element sequences used in this study. Descriptors are given in shorthand; please refer to official documentation for full names (e.g. https://confluence.ecmwf.int/display/ECC/ plus "WMO%3DXX+sequence+table" where XX is the WMO version number (e.g. 40)).

WMO Sequence	**(310078)**	**Custom Sequence**	
Entry	*Descriptor*	*Entry*	*Descriptor*
001007	sat ID	001007	sat ID
001016	sat sub ID	001016	sat sub ID
002019	sat instrument	002019	sat instrument
002020	sat classification	002020	sat classification
001033	gen centre	001033	gen centre
001034	gen subcentre	001034	gen subcentre
301011	YYYYMMDD	301011	YYYYMMDD
301013	HHMMSS	301013	HHMMSS
005040	orbit #	005040	orbit #
201132	Increase bit width (+4)	201132	Increase bit width (+4)
005041	scan #	005041	scan #
201000	(end)	201000	(end)
005043	FOV #	005043	FOV #
033079	granule quality flags	033079	granule quality flags
033080	scan quality flags	033080	scan quality flags
033078	geolocation quality	033078	geolocation quality
		025182	L1 flag
		025084	orbit angle
007002	altitude	007002	altitude
102003	(repeat 2 fields 3x)	102003	(repeat 2 fields 3x)
008097	inst temp method	008097	inst temp method
012164	inst temp	012164	inst temp
117000	(17 fields to be repeated)	005001	lat
031001	*delayed rep. factor*	006001	lon
005001	lat	119000	19 fields to be repeated
006001	lon	031001	delayed rep. factor
007024	sat zenith	005042	channel #
005021	bearing/azimuth	002153	centre freq.
007025	solar zenith	002154	bandwidth
005022	solar azimuth	002104	polarization
109000	(9 fields to be repeated)	007024	sat zenith
031001	*delayed rep. factor*	005021	bearing/azimuth
005042	channel #	007025	solar zenith
002153	centre freq.	005022	solar azimuth
002154	bandwidth	12066	antenna temp
002104	polarization	012163	TB
012066	antenna temp.	012158	nedt cold
012163	TB	012159	nedt warm
012158	nedt cold	033094	cal. QC flags
		033002	quality information
012159	nedt warm	012065[10]	Std(TB)
033094	cal. QC flags	008049	N_obs averaged

Acknowledgements

It has been a pleasure to work on early evaluation of AWS data with colleagues from around Europe. Thanks especially to members of the MWS/AWS science advisory group at EUMETSAT, Ville Kangas at ESA, Patrick Eriksson at Chalmers, Héléne Dumas and Philippe Chambon at Meteo France, Adam Dybbroe at SMHI, Anders Emrich at Omnisys, and numerous others for informative discussions and exchange of information. Thanks to EUMETSAT for supporting this work, in particular Christophe Accadia and Sabatino Di Michele for directing this project, and Tim Hewison for feedback on the report. Thanks as well to Alan Geer and Emma Turner for discussions and help on radiative transfer aspects, and to Peter Lean for technical assistance with observation processing. Lastly, thanks to Tony McNally for reviewing the manuscript.

Glossary

AMSU-A Advanced Microwave Sounding Unit-A. 3, 6

AMVs Atmospheric motion vectors. 59

APC Antenna pattern correction. 34

ATMS Advanced Technology Microwave Sounder. 15

ATOVS Advanced TIROS Operational Vertical Sounder. 16

AWS Arctic Weather Satellite. 3, 5

BUFR Binary Universal Form for the Representation of meteorological data. 17, 64

CI Cloud Impact. 13

ECMWF European Centre for Medium-range Weather Forecasts. 13

ECT Equator crossing time. 16

EPS-SG EUMETSAT Polar System Second Generation. 3

ESA European Space Agency. 3

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites. 3, 6

EURD End User Requirements Document. 11

FOV Field of View. 6

ICI Ice Cloud Imager. 8

IFS Integrated Forecasting System. 3, 13

LTAN Local time ascending node. 16

LWP Liquid Water Path. 13

Metop-SG Metop Second Generation. 6

MHS Microwave Humidity Sounder. 3, 6

MWI MicroWave Imager. 8

NEDT Noise Equivalent Differential Temperature. 6

NWP Numerical Weather Prediction. 5

ODB Observational DataBase. 17

POES Polar-orbiting Operational Environmental Satellite. 16, 53

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

SI Scattering Index. 13

SURFEM-Ocean Surface Fast Emissivity Model for Ocean. 15

TB Brightness Temperature. 13

TEMPEST Temporal Experiment for Storms and Tropical Systems. 5

TROPICS Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats. 5

References

Albers, R., Emrich, A. and Murk, A. (2023). Antenna design for the Arctic Weather Satellite microwave sounder. *IEEE Open J. Antennas Propag.*, **4**, 686–694, doi:10.1109/OJAP.2023.3295390, URL https://ieeexplore.ieee.org/abstract/document/10184100.

Albers, R., Plüss, T., Eggimann, L. and Murk, A. (2024). Development of the onboard calibration target for the Arctic Weather Satellite. *IEEE Trans. Geosci. Remote Sens.*, **62**, 1–8, doi:10.1109/TGRS.2024.3436517, URL https://ieeexplore.ieee.org/abstract/document/10620343.

Andersson, E. and Järvinen, H. (1999). Variational quality control. *Quart. J. Roy. Meteor. Soc.*, **125**(554), 697–722, doi:10.1002/qj.49712555416, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712555416.

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. (2022a). 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. (2022b). Cloud particle orientation and polarisation for cross-track microwave sensors in RTTOV. *Technical Report NWP AVS22 01*, NWP SAF Visiting Scientist Report, URL https://nwp-saf.eumetsat.int/publications/vs_reports/nwpsaf-ec-vs-061.pdf.

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.

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

Bormann, N., Fouilloux, A. and Bell, W. (2013). Evaluation and assimilation of ATMS data in the ECMWF system. *J. Geophys. Res. Atmos.*, **118**(23), 12,970–12,980, doi:10.1002/2013JD020325, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020325.

Bormann, N., Lawrence, H. and Farnan, J. (2019). Global observing system experiments in the ECMWF assimilation system. *Technical Report 839*, ECMWF Tech. Memo., doi:10.21957/sr184iyz, URL https://www.ecmwf.int/node/18859.

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-observations

Camplani, A., Sanò, P., Casella, D., Panegrossi, G. and Battaglia, A. (2024). Arctic Weather Satellite sensitivity to supercooled liquid water in snowfall conditions. *Rem. Sens.*, **16**(22), doi:10.3390/rs16224164, URL https://www.mdpi.com/2072-4292/16/22/4164.

Cardinali, C. (2009). Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.*, **135**(638), 239–250, doi:10.1002/qj.366, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.366.

Chen, H., Han, W., Wang, H., Pan, C., An, D., Gu, S. and Zhang, P. (2021). Why and how does the actual spectral response matter for microwave radiance assimilation? *Geophys. Res. Lett.*, **48**(8), e2020GL092306, doi:https://doi.org/10.1029/2020GL092306, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL092306, e2020GL092306 2020GL092306.

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.

Duncan, D. and Bormann, N. (2020). On the addition of microwave sounders and NWP skill, including assessment of FY-3D sounders. *Technical Report 55*, EUMETSAT/ECMWF Fellowship Programme Research Report, doi:10.21957/iimhohmqt, URL https://www.ecmwf.int/node/19760.

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. (2024a). Superobbing and thinning scales for all-sky humidity sounder assimilation. *Mon. Weather Rev.*, **152**(8), 1821 – 1837, doi:10.1175/MWR-D-24-0020.1, URL https://journals.ametsoc.org/view/journals/mwre/152/8/MWR-D-24-0020.1.xml.

Duncan, D. I., Bormann, N. and Hólm, E. (2021). On the addition of microwave sounders and numerical weather prediction skill. *Quart. J. Roy. Meteor. Soc.*, **147**(740), 3703–3718, doi:10.1002/qj.4149, URL https://doi.org/10.1002/qj.4149.

Duncan, D. I., Geer, A., Bormann, N. and Dahoui, M. (2024b). Vicarious calibration monitoring of MWI and ICI using NWP fields. *Technical report*, EUMETSAT Contract Report, doi:10.21957/7c2d18d2e1, URL https://www.ecmwf.int/en/elibrary/81563-vicarious-calibration-monitoring-mwi-and-ici-using-nwp-fields.

English, S., Duncan, D. and Turner, E. (2020). Assessment of value of provision of spectral response functions from CGMS agencies. In *CGMS-48 WG II*, URL https://www.researchgate.net/publication/351358361_Assessment_of_Value_of_Provision_of_Spectral_Response_Functions_from_CGMS_Agencies.

Eriksson, P., Emrich, A., Kempe, K., Riesbeck, J., Aljarosha, A., Auriacombe, O., Kugelberg, J., Hekma, E., Albers, R., Murk, A., Møller Pedersen, S., John, L., Stake, J., McEvoy, P., Rydberg, B., Dybbroe, A., Thoss, A., Canestri, A., Accadia, C., Colucci, P., Gherardi, D. and Kangas, V. (2025). The Arctic Weather Satellite radiometer. *Atmos. Meas. Tech.*, **18**(18), 4709–4729, doi:10.5194/amt-18-4709-2025, URL https://amt.copernicus.org/articles/18/4709/2025/.

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

Geer, A. J. (2016). Significance of changes in medium-range forecast scores. *Tellus A*, **68**(1), 30229, doi:10.3402/tellusa.v68.30229, URL https://doi.org/10.3402/tellusa.v68.30229.

Geer, A. J., Ahlgrimm, M., Bechtold, P., Bonavita, M., Bormann, N., English, S., Fielding, M., Forbes, R., Hogan, R., Hólm, E., Janiskova, M., Lonitz, K., Lopez, P., Matricardi, M., Sandu, I. and Weston, P. (2017). Assimilating observations sensitive to cloud and precipitation. *Technical Report 815*, ECMWF Tech. Memo., Shinfield Park, Reading, doi:10.21957/sz7cr1dym, URL https://www.ecmwf.int/node/17718.

Geer, A. J. and Baordo, F. (2014). Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech.*, **7**(6), 1839–1860, doi:10.5194/amtd-7-1749-2014, URL https://amt.copernicus.org/articles/7/1839/2014/.

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. (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., 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.

Healy, S., Bormann, N., Geer, A., Holm, E., Ingleby, B., Lean, K., Lonitz, K. and Lupu, C. (2024). Methods for assessing the impact of current and future components of the global observing system. *Technical Report 916*, ECMWF Tech. Memo., Reading, doi:10.21957/2f240fe55f, URL https://www.ecmwf.int/en/elibrary/81556-methods-assessing-impact-current-and-future-components-global-observing

Kaur, I., Eriksson, P., Pfreundschuh, S. and Duncan, D. I. (2021). Can machine learning correct microwave humidity radiances for the influence of clouds? *Atmos. Meas. Tech.*, **14**(4), 2957–2979, doi:10.5194/amt-14-2957-2021, URL https://amt.copernicus.org/articles/14/2957/2021/.

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.

Lean, K. and Bormann, N. (2024). Evaluation of the EPS-Sterna 325 GHz channels in the Ensemble of Data Assimilations. *Technical report*, ECMWF, doi:10.21957/f53d05c057, URL https://www.ecmwf.int/en/elibrary/81561-evaluation-eps-sterna-325-ghz-channels-ensemble-data-assimilations.

Lean, K., Bormann, N. and Healy, S. (2022a). Calibration of EDA spread and adaptation of the observation error model. *Technical Report WP-2000*, ECMWF, doi:10.21957/1auh0nztg, URL https://www.ecmwf.int/node/20302.

Lean, K., Bormann, N. and Healy, S. (2023). Evaluation of initial future EPS-Sterna constellations with 50 and 183 GHz. *Technical report*, ECMWF, doi:doi.org/10.21957/0a695fcc39, URL https://www.ecmwf.int/en/elibrary/81511-task-11-evaluation-initial-future-eps-sterna-constellations-50-and-183-gh

Lean, K., Bormann, N., Healy, S. and English, S. (2022b). Final report: Study to assess earth observation with small satellites and their prospects for future global numerical weather prediction. *Technical report*, ESA Contract Report, doi:10.21957/kp7z1sn1n, URL https://www.ecmwf.int/en/elibrary/81324-final-report-study-assess-earth-observation-small-satellites-and-their-pr

Lean, K., Bormann, N., Healy, S., English, S., Schüttemeyer, D. and Drusch, M. (2025). Assessing forecast benefits of future constellations of microwave sounders on small satellites using an ensemble of data assimilations. *Quart. J. Roy. Meteor. Soc.*, **151**(768), e4939, doi:https://doi.org/10.1002/qj.4939, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4939.

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.

Lu, Q. and Bell, W. (2014). Characterizing channel center frequencies in AMSU-A and MSU microwave sounding instruments. *J. Atmos. Oceanic Technol.*, **31**(8), 1713 – 1732, doi:10.1175/JTECH-D-13-00136.1, URL https://journals.ametsoc.org/view/journals/atot/31/8/jtech-d-13-00136_1.xml.

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.

Randriamampianina, R., Bormann, N., Køltzow, M. A. Ø., Lawrence, H., Sandu, I. and Wang, Z. Q. (2021). Relative impact of observations on a regional Arctic numerical weather prediction system.

Quart. J. Roy. Meteor. Soc., **147**(737), 2212–2232, doi:https://doi.org/10.1002/qj.4018, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4018.

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-th

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.

Voosen, P. (2024). Small, nimble weather satellites join traditional behemoths. *Science*, doi:10.1126/science.adr3312, URL https://doi.org/10.1126/science.adr3312.

Weston, P. and Bormann, N. (2018). Enhancements to the assimilation of ATMS at ECMWF: Observation error update and addition of NOAA-20. *Technical Report 48*, EUMETSAT/ECMWF Fellowship Programme Research Report, URL https://www.ecmwf.int/node/18744.

Wu, D. L., Gong, J., Deal, W. R., Gaines, W., Cooke, C. M., De Amici, G., Pantina, P., Liu, Y., Yang, P., Eriksson, P. and Bennartz, R. (2024). Remote sensing of ice cloud properties with millimeter and submillimeter-wave polarimetry. *IEEE J. Microwaves*, pp. 1–11, doi:10.1109/JMW.2024.3487758, URL https://ieeexplore.ieee.org/document/10758756.

Yang, J. X., You, Y., Blackwell, W., Misra, S. and Kroodsma, R. A. (2022). Quantifying and characterizing striping of Microwave Humidity Sounder with observation and simulation. *IEEE T. Geosci. Remote Sens.*, **60**, 1–13, doi:10.1109/TGRS.2021.3132560.