ESA Contract Report

SMOS-E contract 4000125399/18/I-BG

Contract Report to the European Space Agency

Multi-year SMOS brightness temperature monitoring

Authors: Pete Weston and Patricia de Rosnay
Contract officer: Matthias Drusch

July 2022

Abbreviations

BUFR ....Binary Universal Form for the Representation of meteorological data
CDF ....Cumulative Distribution Function
CMEM ....Community Microwave Emissivity Model
DJF ....December, January, February
ECFS ....ECMWF's File Storage system
ECMWF ....European Centre for Medium-range Weather Forecasts
ESA......................European Space Agency
ESL ....Expert Support Laboratory
IFS ....Integrated Forecast System
JJA ....June, July, August
L1 ....Level 1
MARS ....Meteorological Archival and Retrieval System
MIRAS ....Microwave Imaging Radiometer using Aperture Synthesis
NH ....Northern hemisphere
NRT ....Near Real Time
NWP ....Numerical Weather Prediction
ODB2 ....Observation DataBase
RFI.......................Radio Frequency Interference
SAPP....................Scalable Acquisition and Pre-Processing system
SH ....Southern hemisphere
SMOS ....Soil Moisture and Ocean Salinity
Tb ....Brightness temperature
UTC ....Coordinated Universal Time

1. Introduction

This document summarises the technical setup, production and results from 11 years of NWP-based monitoring of brightness temperatures from the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) satellite. The observations are compared against a stable NWP baseline provided from the ERA5 reanalysis and transformed into observation space using the Community Microwave Emissivity Model (CMEM).

2. Monitoring configuration

In this section the configuration of the SMOS brightness temperature monitoring system is introduced including the specific experimental setup used to produce the multi-year monitoring results.

2.1. v7.24 SMOS L1 data

The Microwave Imaging Radiometer using Aperture Synthesis (MIRAS) instrument onboard the ESA SMOS satellite measures microwave radiances at L-band (1.41GHz) frequencies and a spatial resolution of between 35km and 50km.

The observations used in this study are the re-processed v7.24 SMOS level 1 (L1) brightness temperatures (Tbs). Full details of this dataset are available in the SMOS level 1 and auxiliary data products specifications (see section 7 for the link). Binary Universal Form for the Representation of meteorological data (BUFR) files containing the observations were acquired at ECMWF for the period 12/01/2010 to 31/12/2020. The SMOS instrument measures at incidence angles between 0° and 65°. Before being processed through the ECMWF Integrated Forecast System (IFS) the observations are pre-processed which involves some preliminary quality control (see section 2.4.1 for details) and averaging and thinning over incidence angles. Here we only process observations with incidence angles between 39.5° and 40.5° to reduce the large number of available observations and make the monitoring computationally affordable. In addition, a noise filtering using a 1° binning angle is applied to the observations, see Muñoz-Sabater et al. (2014) for details. Observations at both V and H polarisations in the antenna reference frame are used.

2.2. ERA5 reanalysis

The ERA5 reanalysis (Hersbach et al., 2020) provides a consistent and thorough description of the meteorology of Earth's atmosphere and surface from 1950 to the present day. It uses the scientific configuration of the ECMWF IFS at cycle 41r2 (IFS Documentation CY41R2) which was operational between March and November 2016. It is run at T_L639 resolution which corresponds to a model grid point spacing of ~30km.

In this study the ERA5 reanalysis is used as a stable, consistent and high-quality NWP-based background to compare the SMOS observations against. This contrasts with the ECMWF operational NWP forecast fields which are not consistent over long periods of time due to frequent upgrades to the configuration of the IFS. When looking at SMOS observations over an 11-year period, this would result in multiple discontinuities and changes in the time-series, for example when there are significant

changes to the land surface model physics or monitoring quality control procedures, which is why ERA5 is used instead.

2.3.   CMEM

The Community Microwave Emission Model (CMEM) (de Rosnay et al., 2009) is used as an observation operator to transform NWP fields into simulated brightness temperatures. Over land, it uses a combination of vegetation, soil, and snow models and parametrisations to calculate an accurate emissivity and effective temperature from the input NWP model parameters. The output emissivity and effective temperature allow for variable penetration depths depending on the microwave frequency, soil moisture, and soil temperature, which depend on soil texture and type. In addition, the atmospheric radiative term is calculated, which allows for top of atmosphere brightness temperature to be simulated. For more details of the exact configuration used for CMEM, see de Rosnay et al. (2020).

The CMEM outputs are simulated Tbs in the Earth frame of reference and these are then converted into the SMOS antenna frame of reference using the geometric and Faraday rotation angles provided with each observation in the BUFR files.

2.4.   Operational monitoring system

SMOS observations are monitored operationally using the operational NWP forecast (transformed into observations space by the CMEM observation operator) as the background to compare against. This system runs two 12-hour model trajectories from 09-21UTC and 21-09UTC every 24 hours. Figure 1 shows a schematic for the monitoring run from 09-21UTC.
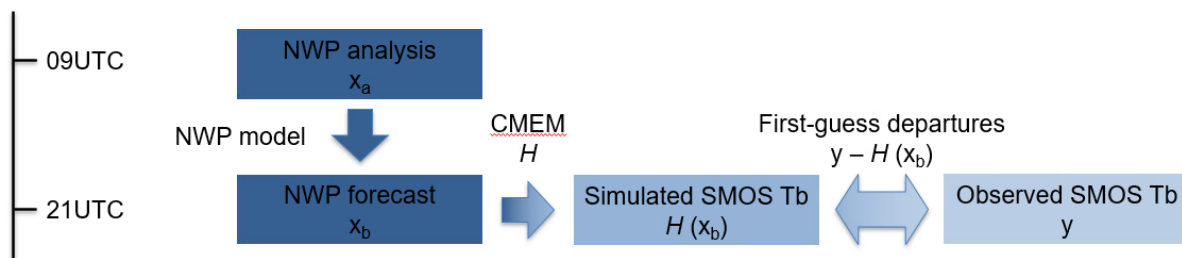

Figure 1: Schematic showing the calculation procedure for the first-guess departures.

The observations are then compared to their model equivalents at the nearest time and location. The observation minus background values are known as first-guess (or background) departures and statistics of this quantity are plotted, displayed online (https://www.ecmwf.int/en/forecasts/quality-our-forecasts/monitoring/smos-monitoring) and analysed. This constitutes the observation monitoring system and many millions of meteorological observations are monitored in this way at ECMWF.

In this study, the same framework as the operational monitoring system is used although, as explained in section 2.2, fields from the ERA5 reanalysis provide the background instead of the operational NWP forecast. Also, the system is run at the same resolution as ERA5, T_L639 (~30km grid spacing), to avoid any unnecessary interpolation and as it is a reasonably close match to the SMOS field of view (FOV) resolution of 35-50km.

The model trajectories used for the monitoring make use of the ECMWF IFS at cycle 47r2 (IFS Documentation CY47R1 - identical to CY47R2). This includes recent upgrades to the quality control procedures used for the SMOS monitoring as described fully in Weston et al. (2021) and briefly below.

2.4.1. Quality control

Data quality information is supplied with each observation in the BUFR files. A series of bits in the SMOS BUFR flag table (de Rosnay et al., 2021) are set to indicate any problems with the associated data. Table 1 shows only those bits which are used and the meaning of each one.

At ECMWF, before the monitoring runs, a pre-screening program is run to remove observations which are known to contain anomalous data or cannot be handled successfully by the monitoring system. Currently, if bit number 3 of the SMOS information flags in the SMOS input BUFR files is set then this indicates that the observation is in the border between the exclusive alias-free zone and the extended alias-free zone and any observations with this bit set are not processed any further. In addition, bit number 7 indicates that data has been affected by solar reflection and any observations with this bit set are not processed any further. Any observations which have any of bit numbers 1, 4 or 9 set are affected by radio frequency interference (RFI) and, although these are processed by the monitoring system, there are sets of monitoring plots in section 4 which filter the data according to these flags.

Bit number	Meaning
1	Pixel is affected by RFI effects as identified in the AUX_RFILST or it has exceeded the BT thresholds
3	Pixel is close to the border delimiting the Extended Alias free zone or to the unit circle replicas borders.
4	Measurement is affected by the tails of a point source RFI as identified in the AUX RFI list (tail width is dependent on the RFI expected BT, from each snapshot measurements, corresponding to 0.16 of the radius of the RFI circle flagged)
7	Pixel is located in a zone where Sun reflection has been detected
9	Measurement is affected by RFI effects as identified in the AUX_RFI list whose contribution generates a contamination in Brightness Temperature above 30K in the corresponding polarization

Table 1: SMOS information flags from the flag table (code 025144) as part of the SMOS NRT product specification. Only those flags which are used at ECMWF are shown.

After the pre-screening the remaining SMOS observations are read into the IFS and further screening procedures are undertaken to avoid areas where the observations cannot be accurately modelled by CMEM, for example over snow-covered surfaces. These are summarised in table 2. In addition to the checks in table 2 each observation location is classified as land if the model land-sea mask value in the collocated grid point is greater than 0.95 and sea if the land-sea mask value is less than 0.01. In this study only observations over land are monitored.

Screening reason	Threshold for rejection
Extreme values	Measured Tb less than 50K or greater than 340K
Snow	Model snow depth greater than 1cm
Frozen surfaces	Model 2 metre temperature less than 273K
Sea-ice	Model sea-ice concentration greater than 1%
Coasts	Model land-sea mask values between 0.01 and 0.95

Table 2: Quality control applied to SMOS observations within the IFS

Figure 2 shows the geographical distribution of the surface type and quality control applied. As expected, most observations at high latitudes are screened out by the snow, frozen surface and sea ice checks. There are RFI detections over the middle East, Eastern Europe and large parts of Asia. Very few observations are screened out by the simple extreme value check. Around all the coasts, inland lakes and small islands observations are screened out by the coast check.
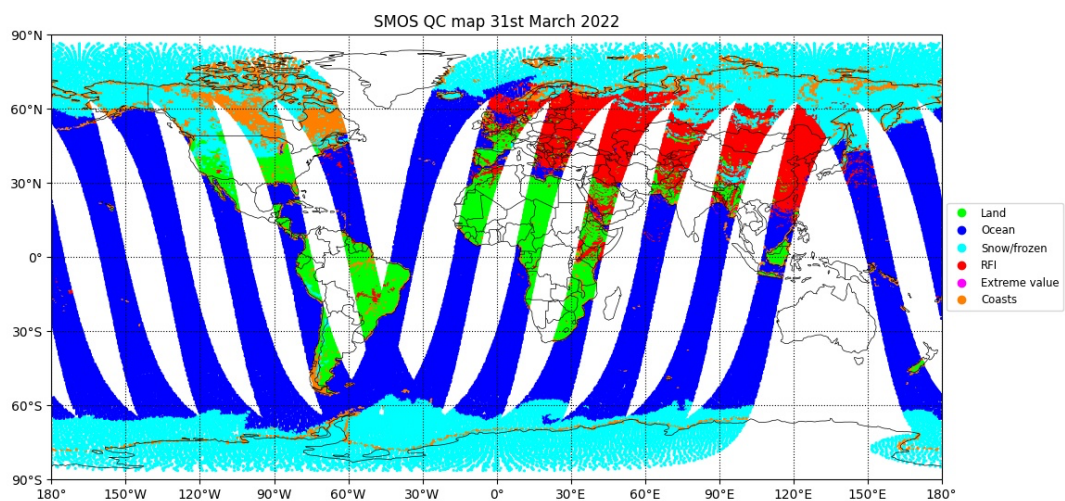

Figure 2: Map showing SMOS observations classified by surface type (land: green; sea: blue) and quality control rejection reason (extreme value: magenta; snow, frozen ground or sea ice: cyan; RFI: red; coasts: orange) for data between 09:00 and 21:00 UTC on 31st March 2022.

2.5. Experiment setup

The ECMWF assimilation system has been run in "monitoring only" mode to obtain monitoring results faster than in full assimilation experiments. This means that the system has been modified to read in the ERA5 reanalysis fields at 09 and 21UTC each day to provide the background for the 12-hour model trajectory to which the observations are compared.

For pragmatic reasons 11 separate experiments were run in parallel covering each of the 11 years separately. This produces the same results as would have been obtained running a single experiment because each 12-hour period is independent of any other 12-hour period with the externally produced ERA5 reanalysis providing the background rather than a forecast generated within the experiment itself.

The first-guess departures are calculated and stored in Observation DataBase (ODB2) files along with other metadata (e.g. latitude, longitude, time, polarisation etc.) which are archived to the ECMWF Meteorological Archival and Retrieval System (MARS) for later analysis and processing.

3. Bias correction

Most data assimilation methods assume that observations are unbiased prior to being assimilated which usually means some form of bias correction needs to be applied. Biases in the first-guess departures can come from a range of different sources including the calibration of the observations themselves, the observation operator and the NWP model. In land surface modelling, biases can be very heterogeneous in both space and time and can depend on many different parameters of the land surface model. This is visible in gridded maps of the mean of first-guess departures, see figure 3.

Figure 3 shows that there are large geographical variations in the biases with positive biases over Russia, Eastern China and South America and negative biases over the Sahara, middle East and Western Australia. It also shows variations in biases between December, January, February (DJF) and June, July, August (JJA), for example over the Sahel, Western Europe, central US and Eastern Australia. This shows there are spatial but also temporal variations in the biases meaning that any bias correction needs to take both aspects into account.

3.1. CDF-matching approach

The bias correction method used is cumulative distribution function matching (CDF-matching) (Scipal et al., 2008) and follows a very similar methodology to that of de Rosnay et al. (2020). The difference here is that the CDF-matching parameters are applied in observation space rather than in model space. CDF-matching aims to match the first and second statistical moments of the cumulative distribution functions (CDFs) of the observed and simulated Tbs at each observation location. The equations below are used to correct the observed Tbs.

$$T_{B\,SMOS}{}^{*} = A + B \,.\, T_{B\,SMOS}$$

$$A = <T_{B\,CMEM}> - <T_{B\,SMOS}> * Stdev(T_{B\,CMEM}) / Stdev(T_{B\,SMOS})$$

$$B = Stdev(T_{B\,CMEM}) / Stdev(T_{B\,SMOS})$$

The training sample period used to calculate the CDF-matching parameters A and B is taken from the middle of the SMOS lifetime, 2014 to 2017. This period was chosen to reduce the effects of RFI which was larger in the early years and have also grown in recent years. The CDF-matching parameters are calculated separately for each month of the year with a three-month moving sample. For example, the parameters for July are calculated using data from June, July, August from 2014 to 2017. This should enable the CDF-matching approach to correct for seasonal biases.
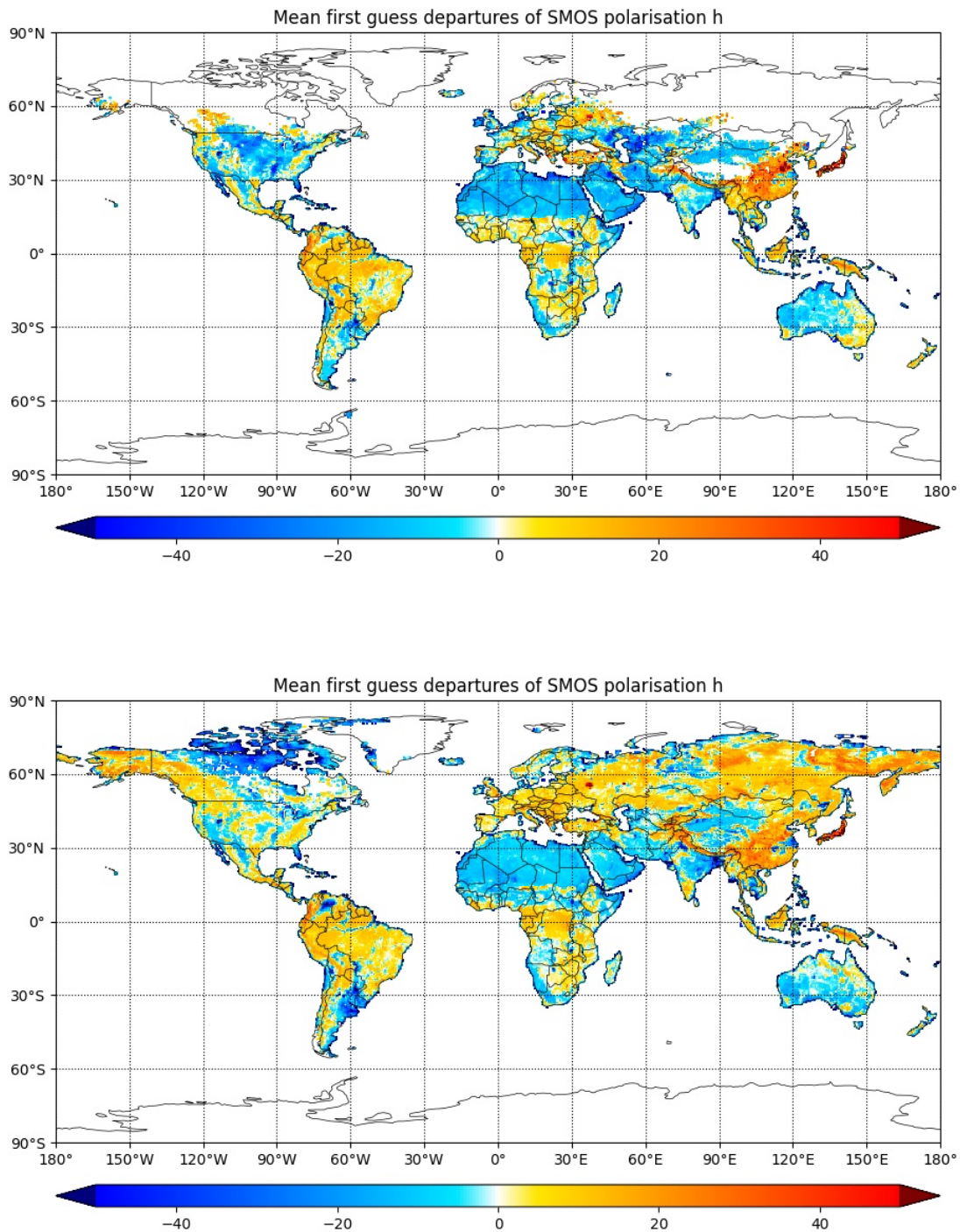
Figure 3: Gridded map of mean SMOS first-guess departures at H polarisation and 40° incidence angle. The sample of data is from 2014 to 2017 with the upper panel showing data for December, January, February and the lower panel showing data for June, July, August.

The CDF-matching parameters are calculated on a 0.5°x0.5° grid and applied to observations using a nearest neighbour interpolation. A 0.25°x0.25° grid was also tried (which would have better matched

the SMOS FOV and model resolution) but the number of samples for each grid point was not sufficient and led to large variability in the parameters between neighbouring grid points. This should enable the CDF-matching approach to correct for geographical biases.

The CDF-matching parameters are calculated separately for H and V polarisations and only use data which passes all quality control procedures as described in section 2.4.1.

3.2. Sources of residual bias and higher variability

As shown in Figure 3 there are significant geographically varying biases over different areas of the globe. One of the advantages of assessing such a long time series of observation minus model statistics is that the persistence and variability of such residual biases can be ascertained and this gives information on the potential sources of such biases.

Figure 4 shows that the standard deviation of first-guess departures also varies significantly over different areas of the globe as well as seasonally. In particular, there is an area around the Sahel in central Africa that has higher standard deviations of first-guess departures in the JJA period than the DJF period. The same region also has variable biases between the two seasons, with Figure 3 showing negative biases in JJA and more positive biases in DJF. The opposite signals can be seen in the far North of Australia.

The source of the higher variability and different biases is uncertain although the periods with the higher variability do coincide with the wet seasons in these two areas (i.e. JJA for central Africa, DJF for Northern Australia). To some extent, higher variability in soil moisture (and hence SMOS Tbs) would be expected from larger precipitation amounts but this would be expected to affect both the simulated and observed Tbs. It's possible that if the model precipitation is in the wrong location that this would lead to larger variability in the first-guess departures and hence the higher standard deviations seen.

The more negative biases in the wet season mean that the simulated values are too high which, if the error is due to the underlying NWP model, suggests that the model soil moisture is too dry, potentially due to insufficient precipitation compared to reality, or the model vegetation is overestimated compared to reality. However, it is not as simple as that because the NWP model output is converted to Tbs using the observation operator, CMEM, and there could be sub-optimalities in CMEM that contribute to these differences. For instance, there will be more vegetation during the wet season which will affect the leaf area index and the simple way that CMEM accounts for vegetation water content is a large source of error in the simulated Tbs These errors in the vegetation parametrisations used in CMEM could lead to the residual biases seen in figure 3.

Another notable feature is the positive bias over Russia in JJA (not visible in DJF due to snow and/or frozen ground). This could be due to the NWP model soil moisture being too wet, possibly due to too much snow in the Northern hemisphere winter which then melts and percolates into the soil in spring and summer. However, it could also be due to errors in vegetation due to rapid growth during the short growing season in that part of the world.
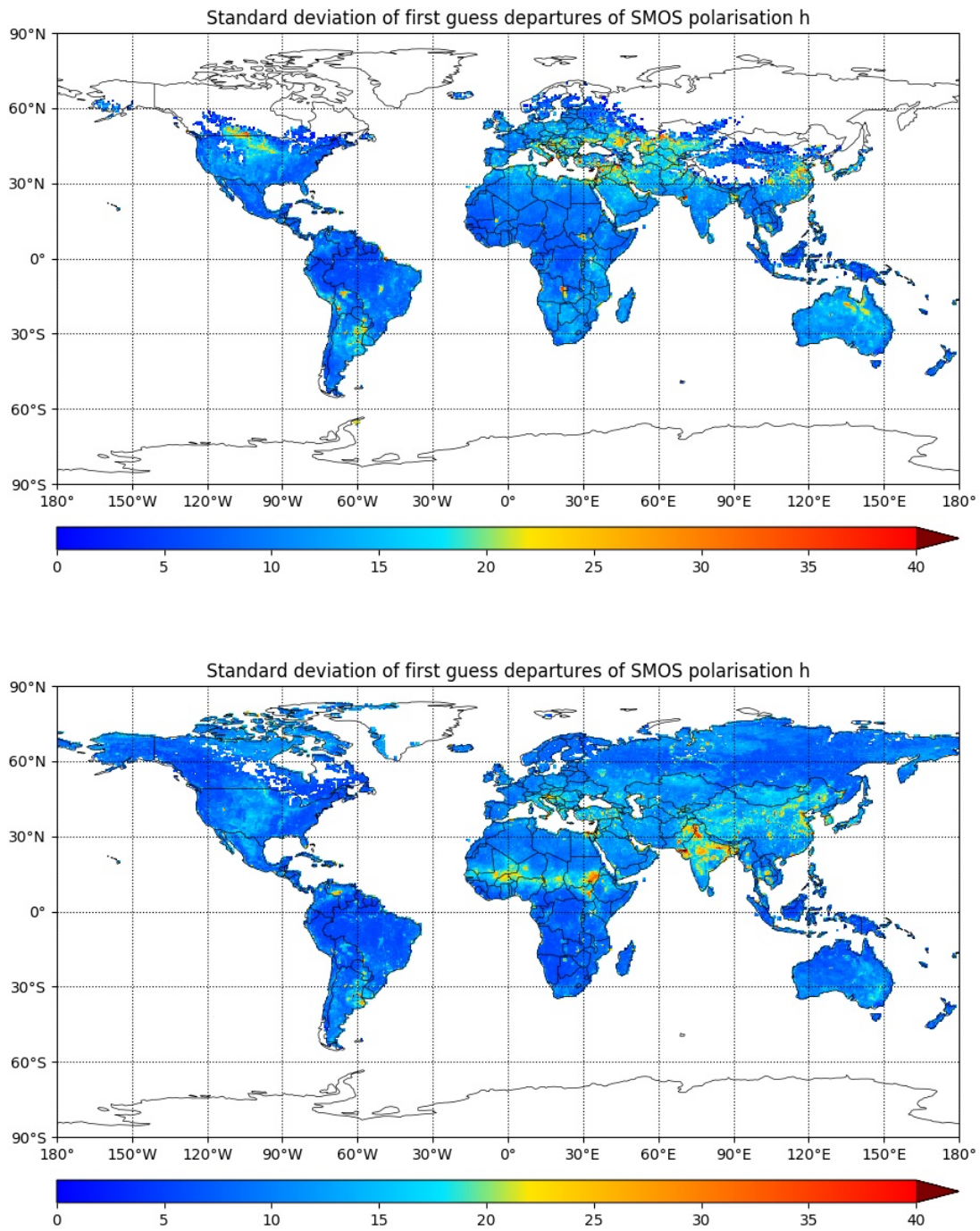
Figure 4: Gridded map of standard deviation of SMOS first-guess departures at H polarisation and 40° incidence angle. The sample of data is from 2014 to 2017 with the upper panel showing data for December, January, February and the lower panel showing data for June, July, August.

For now, the CDF-matching will correct the majority of these biases by construction, but confronting models with observations over such a long period can produce powerful diagnostics which can be used to diagnose and then improve or correct model and/or observation operator issues.

4.   Results

In this section the results of the multi-year monitoring will be presented and key features analysed.
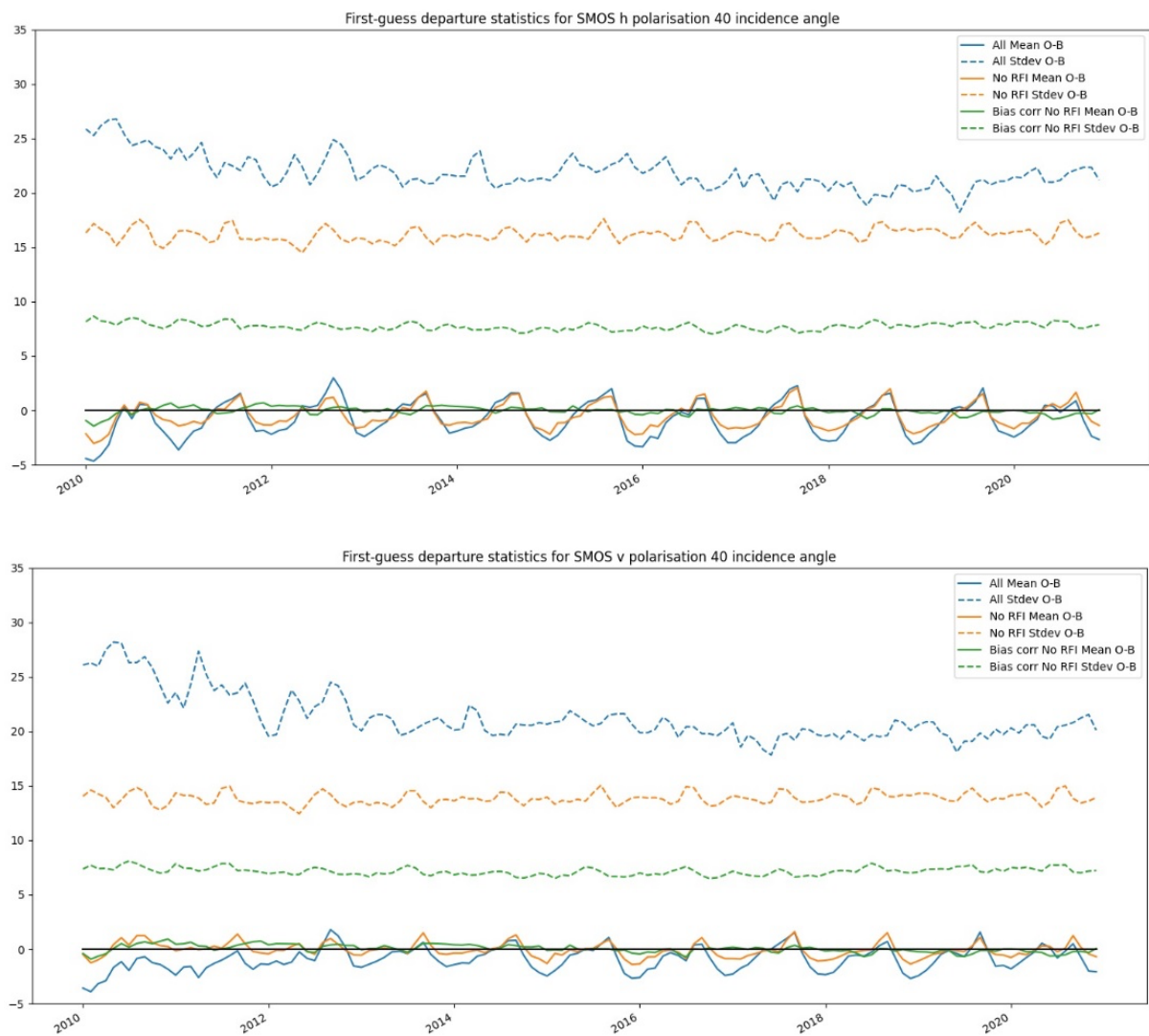
4.1.   Global statistics

Figure 5: Time series showing global mean (solid lines) and standard deviation (dashed lines) of SMOS first-guess departures for January 2010 to December 2020. The samples of data shown are all data (blue); data after RFI screening (orange) and data after the bias correction has been applied (green). The upper panel shows H polarisation data and the lower panel shows V polarisation data. The statistics plotted are monthly averages over the 11-year period.

Figure 5 shows that there is a decreasing trend in the global standard deviation of first-guess departures for all data between 2010 and 2013. This is due to lobbying efforts to reduce RFI sources during the early years of the SMOS mission.

The presence of RFI sources at L-band frequencies was not expected to be a problem before the SMOS launch because SMOS measures in a protected band of the radio frequency spectrum (1.4 - 1.427 GHz). However, the RFI sources, from satellite TV signals, military radar and wireless security cameras amongst others, were discovered when looking at the first data from SMOS soon after launch. When SMOS measures an RFI source the observed Tb will be very different to the simulated Tb from the model (because the model has no knowledge of RFI sources), which translates into larger and more variable first-guess departures. This leads to larger standard deviation of first-guess departures when looking at larger data samples.

The first main approach to reduce the effect of the RFI on the SMOS measurements was to lobby national governments to enforce the protection of the radio frequency band that SMOS measures in and this led to a large number of RFI sources being switched off. This is what is driving the reducing trend of the global standard deviation of first-guess departures between 2010 and 2013. After the apparent success of this early lobbying effort the time series shows that the standard deviation of first-guess departures remains approximately constant from 2013 onwards. This is due to the emergence of new RFI sources being balanced by continued lobbying efforts.

The second main approach to reduce the effect of RFI is to use algorithms and databases of RFI source locations to screen RFI contaminated data. The results from the data sample after this screening is applied are shown in the orange lines on the time series. The main result here is that the global standard deviation of first-guess departures is significantly reduced from ~20-25K for the "all" data sample to ~13-17K for the "RFI screened" data sample. After the screening, the global first-guess departure statistics are stable throughout the period, no longer exhibiting the decreasing trend in the early years. This shows that the RFI screening is performing consistently throughout the period despite some RFI sources being turned off and new RFI sources emerging.

Both the "all" and "RFI screened" data samples show significant seasonally varying global biases in the time series with the magnitude of these global biases varying by ~4K for H polarisation and ~3K for V polarisation. Figure 5 shows that the bias corrected data in green almost completely removes the seasonally varying biases with the global mean first-guess departures being very close to zero after the bias correction has been applied. In addition, there is a significant reduction in the global standard deviation of first-guess departures. When calculating the global standard deviation, the effect of the geographical variability of the biases will result in larger global standard deviations, which are consequently reduced when the geographically varying biases, shown in figure 3, are corrected.

The fact that the difference between the standard deviation of first-guess departures before and after bias correction is stable throughout the period shows that, despite the bias correction only being trained on data from 2014 to 2017, it is performing consistently for the data throughout the 2010 to 2020 period. The overall results and trends for the H and V polarisation data are largely similar, albeit with slightly

different specific values for the statistics themselves (e.g. standard deviation of first-guess departures are ~8K for H polarisation and ~7.5K for V polarisation).

Figure 6 shows time series split up by hemisphere. The same decreasing trend in the early years is visible in the Northern hemisphere (NH) but not in the Southern hemisphere (SH) and this is expected because the vast majority of RFI sources at SMOS frequencies are found in the NH. This is also backed up when looking at the differences between the "all" and "RFI screened" data samples which are large in the NH but small in the SH.

Notably the global standard deviation of first-guess departures for the two hemispheres are remarkably similar after the RFI screening at ~13-17K before bias correction and 6-9K after bias correction. This is more evidence that the RFI screening is doing a good job of removing the effects of RFI on the data.
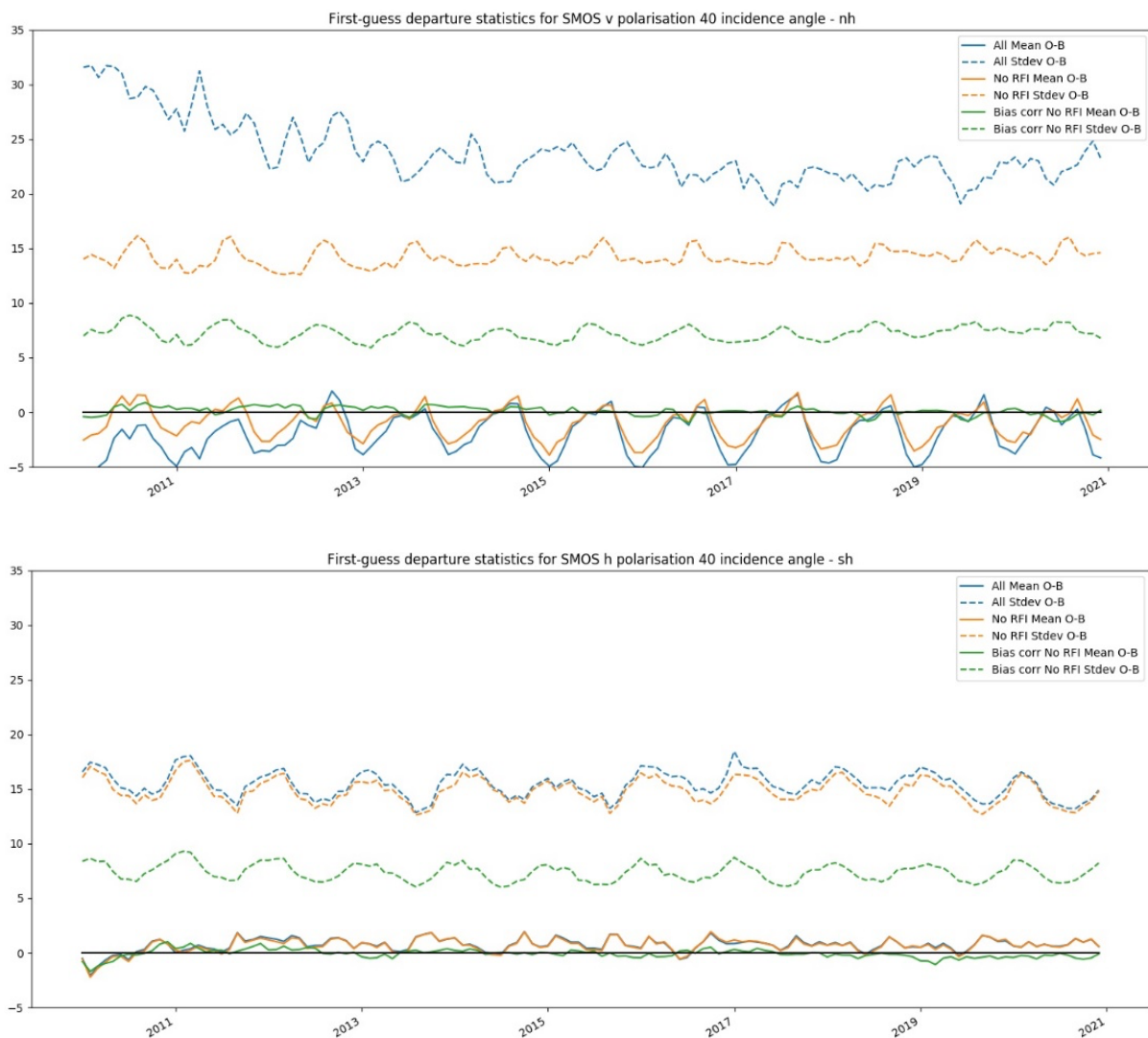

Figure 6: As Figure 5 but for only H polarisation observations from the Northern hemisphere (upper panel) and Southern hemisphere (lower panel).

The effect of the bias correction is also similar across both hemispheres although there is a slight seasonal cycle in the standard deviation of first-guess departures with larger values in the summer for the relevant hemisphere (e.g, JJA for NH and DJF for SH). This is partially driven by the higher variability in standard deviations during the tropical wet seasons as discussed in section 3.2.

The global first-guess departure statistics have also been partitioned into observations from the ascending or descending node of the SMOS satellite's orbit. It was found that the standard deviations of first-guess departures for the different nodes were very similar but there was a significant ascending minus descending bias of ~2K for H polarisation and ~1.75K for V polarisation. These biases could come from the instrument due to the viewing geometry, or there could be systematic diurnal model errors relating to the local solar time which is 06:00 on the ascending node and 18:00 on the descending node. At present, the node of the orbit is not included as a predictor in the bias correction, but given these non-negligible biases, this could be included in the future.
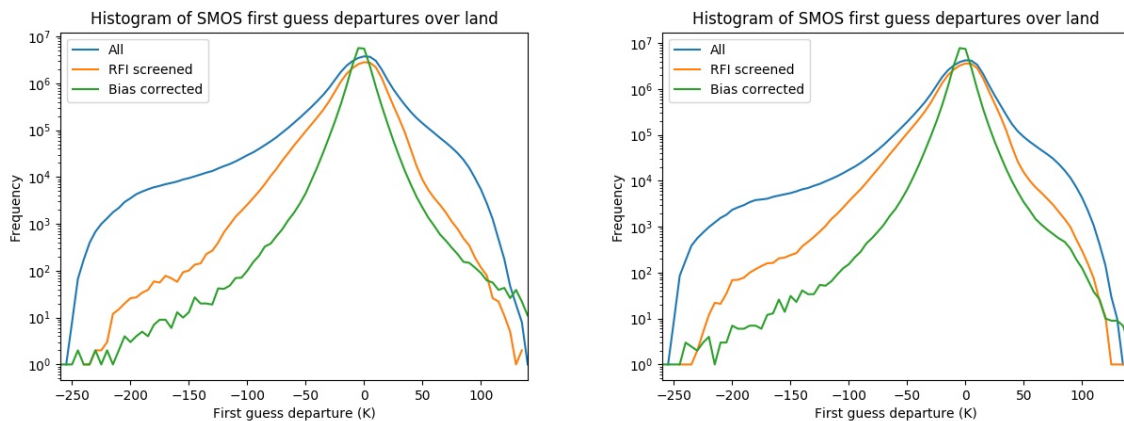

Figure 7: Histograms showing the distribution of SMOS first-guess departures of H polarisation for the all (blue), RFI screened (orange) and bias corrected (green) data samples for 2010 (left) and 2020 (right)

Figure 7 shows that the first-guess departure distributions for the "all" data sample have very large tails which are largely due to RFI and are signficantly reduced by the RFI screening. This reduction is larger in 2010 than in 2020 due to the larger number of severe RFI sources in the early years of the SMOS mission. The bias correction also reduces the tails and makes the distribution more symmetric and Gaussian meaning that the data after bias correction are more suitable for data assimilation. Results for V polarisation (not shown) are very similar.

4.2. Regional statistics

Figure 8 shows that there are significant seasonal cycles in the mean first-guess departures throughout the period. For example, the larger positive biases between 40°N and 70°N in the NH summer are driven by the biases over Russia discussed in section 3.2 and appear at a consistent time and magnitude every year. There are also seasonal cycles of smaller magnitude over the tropics and SH, some of which were also discussed in section 3.2.

After the bias correction has been applied the magnitude of the residual biases is very small, in comparison, and the seasonally varying regional biases have been almost completely removed. This is a good result which justifies the use of the seasonal component of the CDF-matching approach because a temporally constant correction wouldn't be able to correct the features highlighted above. In the residual biases after bias correction there is still a slight positive bias in the NH in 2010 but this is not present for any of the subsequent years so could relate to residual RFI or the exceptionally warm temperatures in Russia during that summer. The only other notable signal is a residual bias which switches sign on a yearly timescale at around 10°N and another one of opposite phase at 40°S. It is not clear what is driving this but the magnitude is relatively small so it is of little concern.
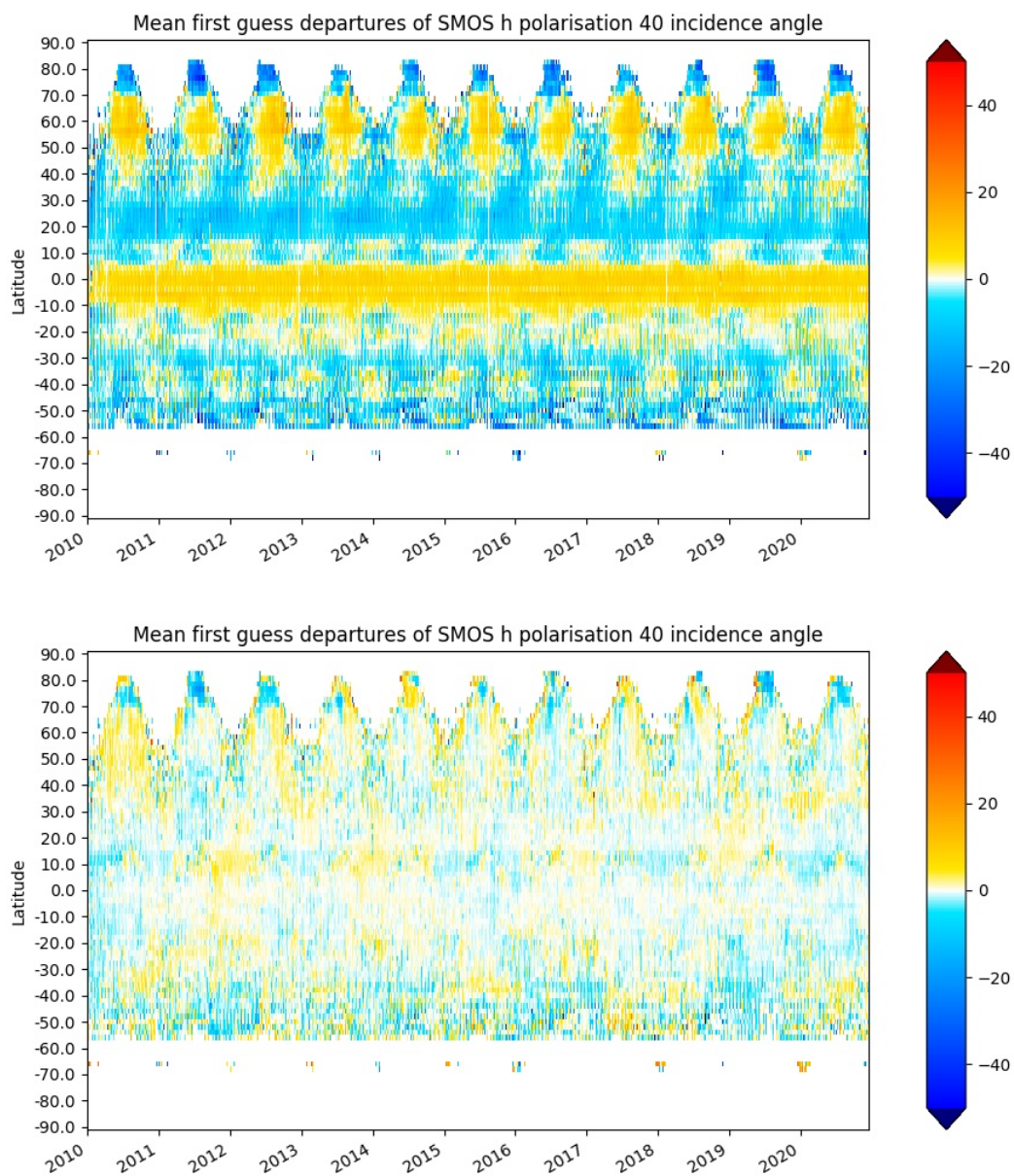

Figure 8: Hovmöller plots of mean SMOS first-guess departures for H polarisation before (upper panel) and after (lower panel) bias correction.

Figure 9 shows that the standard deviation of first-guess departures are also significantly reduced after the bias correction is applied. This effect appears to be homogeneous throughout the period and at all latitude bands. The most notable seasonal cycle is either side of the Equator at 10°N and 20°S with the increased values relating to the tropical wet season discussed in section 3.2. There are also slightly larger values at 50°N in 2010 which could, again, be related to the hot temperatures over Russia and a slight increase in the standard deviation of first-guess departures around 35°N since 2018 due to new RFI sources over China that are not completely removed by the screening.
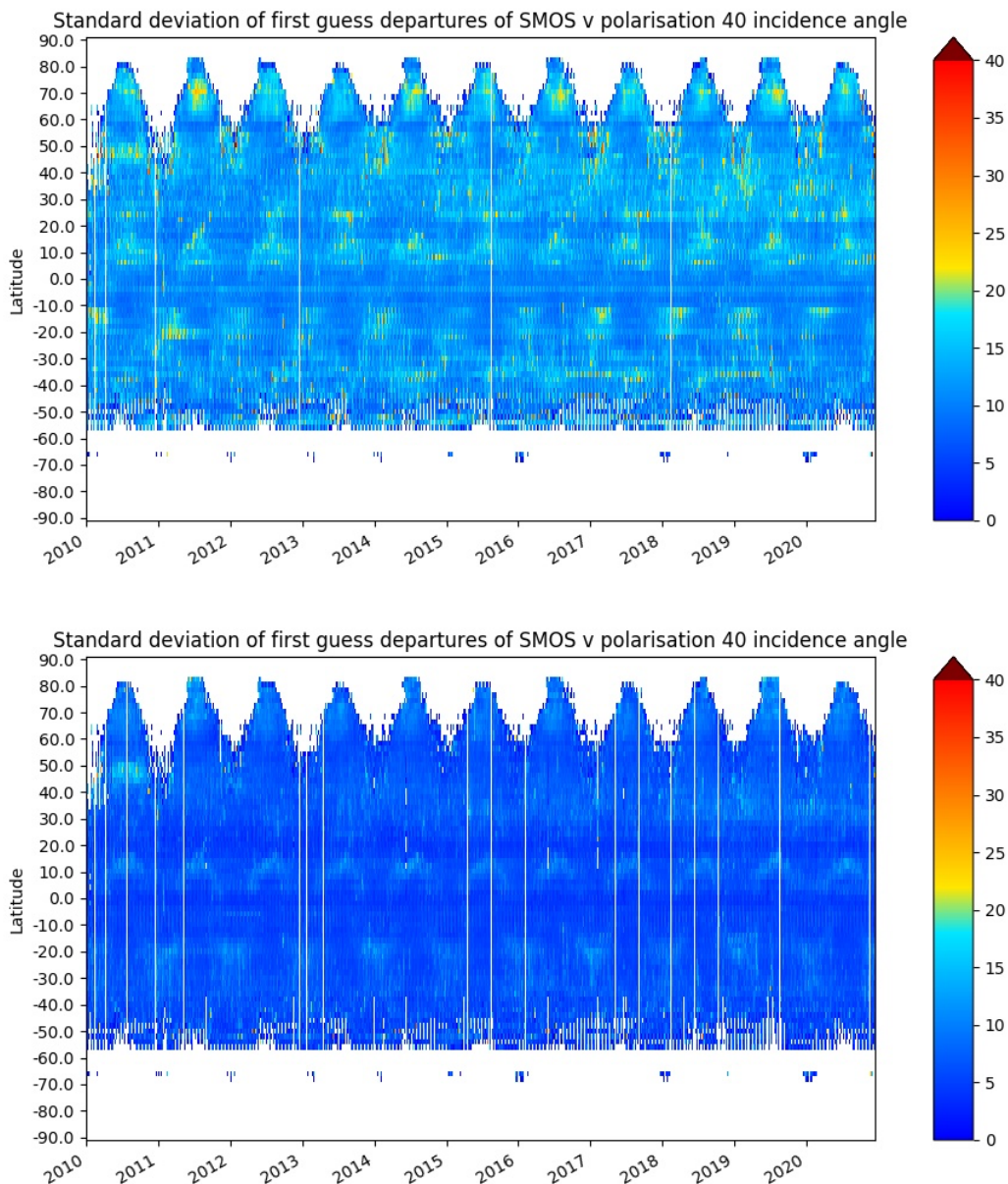

Figure 9: Hovmöller plots of standard deviation of SMOS first-guess departures for V polarisation before (upper panel) and after (lower panel) bias correction
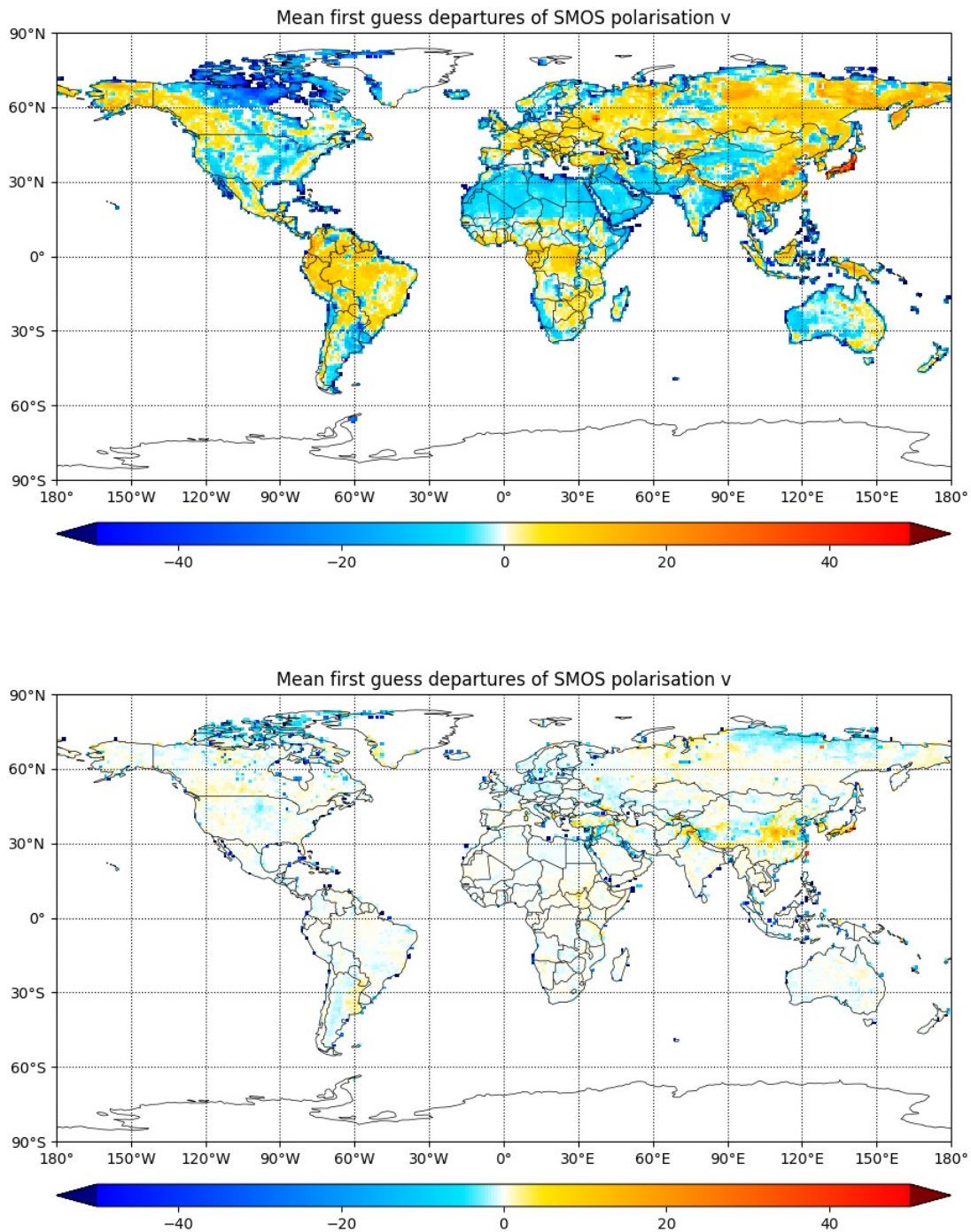
Figure 10: Gridded maps of mean SMOS first-guess departures for V polarisation before (upper panel) and after (lower panel) bias correction. Data is accumulated for the entire 2010-2020 period.

Figure 10 shows that the vast majority of geographical biases are removed by the CDF-matching. The only significant exceptions to this are over parts of the middle East, China and Japan where residual RFI is potentially still contaminating the data despite the RFI screening. There are also residual biases after the bias correction over the far North of Siberia and Canada. These areas are covered by snow and

frozen surfaces for most of the year meaning that the data sample used to calculate the CDF-matching parameters is smaller which potentially leads to slightly poorer performance than in other areas.
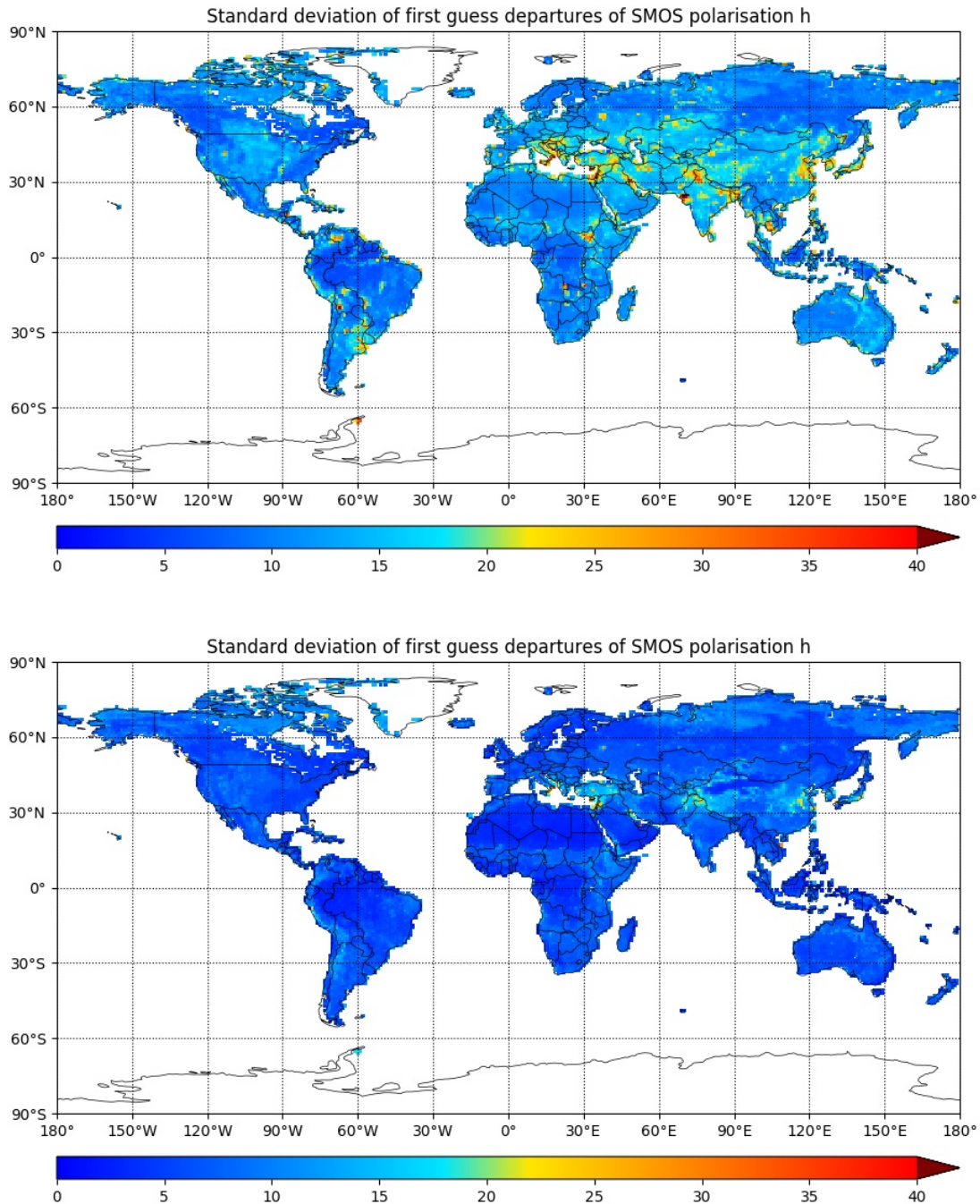

Figure 11: Gridded maps of standard deviation of SMOS first-guess departures for H polarisation before (upper panel) and after (lower panel) bias correction. Data is accumulated for the entire 2010-2020 period.

Figure 11 shows that the bias correction significantly reduces the standard deviations of first-guess departures over the whole globe, with a few exceptions. The middle East, central Asia and China are

particularly prominent, and this is potentially due to residual RFI as was seen in the residual biases in those areas too. There are also slightly larger values over the Northern high latitudes, again due to the smaller sample used to calculate the CDF-matching parameters. Finally, there are slightly larger values over central Africa and Northern Australia which relate to the possible model sub-optimalities discussed in section 3.2. In other areas the standard deviation of first-guess departures are fairly stable between 5K and 10K which corresponds well to the global average of ~8K seen in Figure 5.
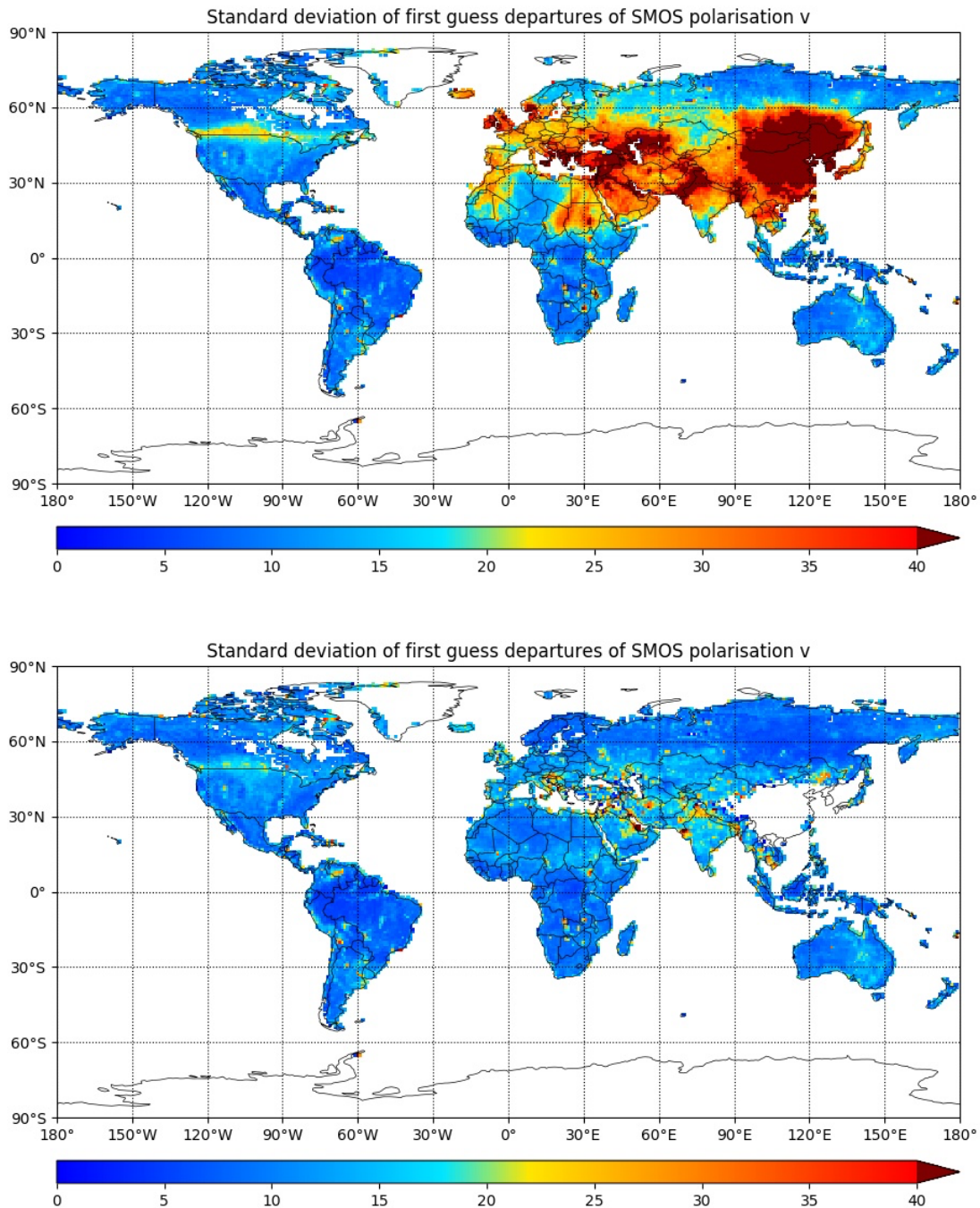

Figure 12: Gridded maps of standard deviation of SMOS first-guess departures for V polarisation in 2010, before (upper panel) and after (lower panel) RFI screening.
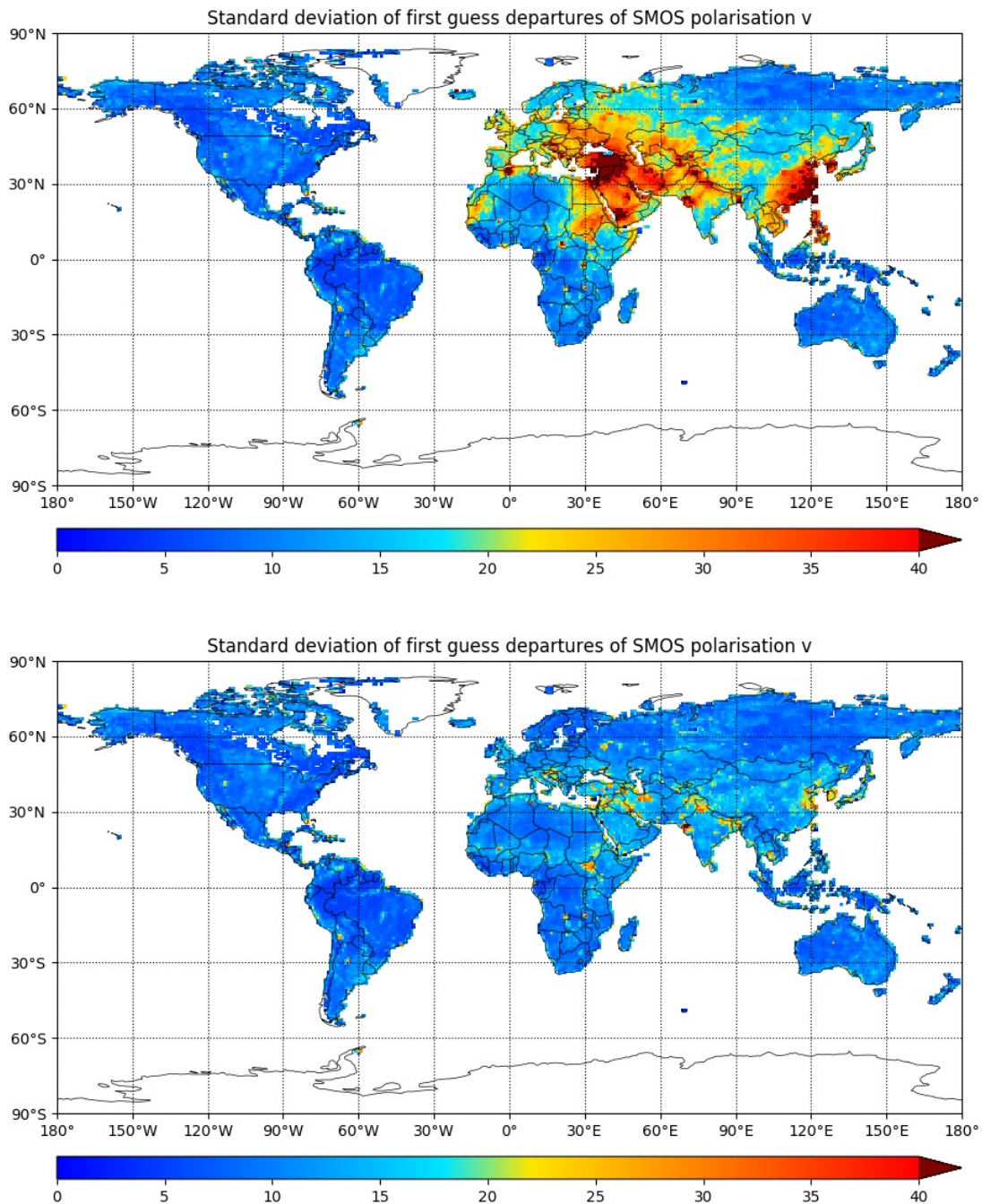
Figure 13: Gridded maps of standard deviation of SMOS first-guess departures for V polarisation in 2020, before (upper panel) and after (lower panel) RFI screening.

Figure 12 and 13 show the changing effects of RFI on the SMOS first-guess departures between 2010 and 2020. In 2010, before RFI screening is applied, there are very large areas of very high standard deviations of first-guess departures over most of China, central Asia, the middle East, Europe, Northern Africa, and the Canada-US border. Comparing to the data after the RFI screening has been applied this screening does a decent job of significantly reducing the standard deviation of first-guess departures of the remaining data. The downside being that a lot of data is lost, including pretty much all the data over

China as shown by the blank area on the map. There are also residual affects shown by the cyan colours particularly visible over the Canada-US border, Western Europe and central Asia showing that the screening is not perfect.

In 2020, before screening, there are still significant areas of contamination, but they are less severe and cover smaller areas than in 2010. The main areas now affected are South-Eastern Asia, the middle East and South-Eastern Europe. In particular, the RFI effects over Western Europe and the Canada-US border have been significantly reduced thanks to lobbying efforts. After screening, the situation is better still with much smaller regions where the residual effects of the RFI are visible in the standard deviations of first-guess departures. There are still a few hot spots but the fact that we retain at least some data over China is an improvement on the situation in 2010.

5.   Conclusions

Overall, the results show that the SMOS data quality is stable over the whole 2010-2020 period once the effects of RFI have been largely removed from the data sample. This contrasts with the results of de Rosnay et al. (2020), which concluded that the data quality improved between 2010 and 2016. It appears this trend was due to a reduction in the number and severity of RFI sources combined with insufficient RFI screening in v5.05 of the SMOS L1 Tb product.

Results here show that the RFI screening applied to v7.24 of the SMOS L1 Tb product is much more effective at removing RFI contaminated observations than in previous versions. This is highlighted by the decreasing trend in the standard deviations of first-guess departures between 2010 and 2013 being present before the RFI screening but disappearing after the RFI screening is applied. The presence of this trend shows the significant effect of the lobbying effort. The lobbying effort continues to the present day and although it hasn't had much of an effect on the global first-guess departure statistics since 2013, it has stopped any increasing trends due to new RFI sources which have emerged in recent years. The lobbying effort also means fewer SMOS observations have to be screened out for RFI and larger samples can be used to analyse soil moisture and in monitoring applications over a wider area, for example in 2010 none of the SMOS data over China was useable whereas in 2020 there are many data points which pass the RFI screening.

One of the major scientific advancements of the work here is to update the bias correction using the same method as de Rosnay et al. (2020), but now enabling this to be applied in observation space and not being tied to a specific model grid resolution. This makes it more flexible and suitable for use as part of the direct assimilation of SMOS L1 Tbs. The performance of the bias correction is stable throughout the entire 2010-2020 period despite only being trained on data from the 2014-2017 period. It effectively corrects the geographically varying biases over almost the entire globe as well as significant seasonally varying biases. Finally, the ascending-descending biases were assessed and, although small when compared to the geographical and seasonal biases, they aren't negligible. These biases could be corrected in a future version of the bias correction.

One of the major advantages of comparing observations to simulated model equivalent values over such a long period is to identify and investigate possible sub-optimalities or issues with the underlying NWP model and observation operator. A number of these areas have been identified such as over tropical

Africa and Northern Australia during the wet season and Russia during the Northern hemisphere summer. Assimilated data such as SMOS is only able to correct random errors in the NWP model whereas most of the errors discovered here have been systematic errors, which will need to be corrected by developments to the NWP model and observation operator themselves.

The improved RFI screening, quality control and bias correction developments mean that revisiting the direct assimilation of SMOS L1 Tbs into the ECMWF land surface data assimilation system is viable with the aim to improve on the previous results of Muñoz-Sabater et al. (2019). Even though SMOS is far beyond its initially planned mission duration it continues to perform at a very stable level. The stability of the data quality for SMOS over a 10+ year period potentially means it could become very useful for looking at climate trends in soil moisture in future years. In the context of the Copernicus Climate Change Service, the direct assimilation of SMOS L1 Tbs could be implemented in a future reanalysis which would enable the exploitation of data from the entire SMOS lifetime, contributing to climate trend assessments in the future.

6. Acknowledgements

7. References

de Rosnay, P.; Drusch, M.; Boone, A.; Balsamo, G.; Decharme, B.; Harris, P.; Kerr, Y.; Pellarin, T.; Polcher, J.; Wigneron, J.-P. AMMA Land Surface Model Intercomparison Experiment coupled to the Community Microwave Emission Model: ALMIP-MEM. J. Geophys. Res. Space Phys. 2009, 114.

de Rosnay, P., M. Dragosavac, M. Drusch, A. Gutiérrez, M. Rodríguez López, N. Wright, J. Muñoz Sabater, Raffaele Crapolicchio: SMOS NRT BUFR specification, 25/05/2021. SMOS_NRT_BUFR_ECMWF_v1.9 (esa.int)

de Rosnay P. and P. Weston: SMOS long-term assessment based on re-analyses: strategy and work plan", ESA contract report. ESA SMOS-E contract 4000125399/18/I-BG, December 2019

de Rosnay, P., J. Muñoz-Sabater, C. Albergel, L. Isaksen, S. English, M. Drusch, J.-P. Wigneron: SMOS brightness temperature forward modelling and long term monitoring at ECMWF. Remote Sens. Environ., 237 (2020): 111424. https://doi.org/10.1016/j.rse.2019.111424

Hersbach, H, Bell, B, Berrisford, P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc. 2020; 146: 1999– 2049. https://doi.org/10.1002/qj.3803

IFS Documentation CY41R2 (
https://www.ecmwf.int/en/publications/search/?solrsort=sort_label%20asc&secondary_title=%22IFS
%20Documentation%20CY41R2%22) Accessed on 21/06/2022

IFS Documentation CY47R1
(https://www.ecmwf.int/en/publications/search/IFS%20Documentation%20CY47R1?authors=ECMW
F&year=2020&f%5B0%5D=sm_biblio_type%3ABook%20Chapter) Accessed on 21/06/2022

Muñoz-Sabater, J., P. de Rosnay, C. Jiménez, L. Isaksen and C. Albergel, "SMOS Brightness
Temperature Angular Noise: Characterization, Filtering, and Validation," in IEEE Transactions on
Geoscience and Remote Sensing, vol. 52, no. 9, pp. 5827-5839, Sept. 2014, doi:
10.1109/TGRS.2013.2293200.

Muñoz-Sabater, J, Lawrence, H, Albergel, C, et al. Assimilation of SMOS brightness temperatures in
the ECMWF Integrated Forecasting System. Q J R Meteorol Soc. 2019; 145: 2524– 2548.
https://doi.org/10.1002/qj.3577

Scipal, K, Drusch, M, Wagner, W. Assimilation of a ERS scatterometer derived soil moisture index in
the ECMWF numerical weather prediction system. Adv. Water Resour. (2008),
https://doi.org/10.1016/j.advwatres.2008.04.013

SMOS Level 1 and Auxiliary Data Products Specifications
(https://earth.esa.int/eogateway/documents/20142/37627/SMOS-L1-Aux-Data-Product-
Specification.pdf) Accessed on 21/06/2022

Weston, P., and P. de Rosnay. "SMOS Brightness Temperature Monitoring Quality Control Review
and Enhancements" Remote Sensing 13, no. 20 (2021): -. doi: 10.3390/rs13204081