ESA Contract Report

SMOS ESL contract 4000130567/20/I-BG

Contract Report to the European Space Agency

Quarter 3 2022: Operations Service Report

Authors: Pete Weston and Patricia de Rosnay
Contract officer: Raffaele Crapolicchio

October 2022

Series: ECMWF - ESA Contract Report

A full list of ECMWF Publications can be found on our web site under:

http://www.ecmwf.int/publications/

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

Abbreviations

BUFR ....Binary Universal Form for the Representation of meteorological data
CESBIO ....Centre d'Etudes Spatiales de la Biosphère
DPGS ....Data Processing Ground Segment
ECFS ....ECMWF's File Storage system
ECMWF ....European Centre for Medium-range Weather Forecasts
ESA.......................European Space Agency
ESAC ....European Space Astronomy Centre
ESL ....Expert Support Laboratory
FTP ....File Transfer Protocol
MIRAS ....Microwave Imaging Radiometer using Aperture Synthesis
NetCDF ....Network Common Data Form
NN.........................Neural Network
NRT ....Near Real Time
NWP ....Numerical Weather Prediction
SAPP.....................Scalable Acquisition and Pre-Processing system
SEKF.....................Simplified Extended Kalman Filter
SMOS ....Soil Moisture and Ocean Salinity

1. Introduction

This document summarises the production and dissemination status of the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) neural network (NN) nominal soil moisture product for the third quarter of 2022. The NN nominal product is produced at the European Centre for Medium-range Weather Forecasts (ECMWF) and it processes raw SMOS BUFR files within 30 minutes of their arrival via the Scalable Acquisition and Pre-Processing system (SAPP). The SMOS BUFR files should be available to ECMWF less than 165 minutes from the initial observation time and the NN product NetCDF files should be delivered to ESA less than 240 minutes from the initial observation time in the corresponding source BUFR file. Statistics of the production and timeliness of the delivered product are presented, reasons for the lack of completeness and/or failure to meet the timeliness deadline are given and corrective actions (if possible) are described in this report.

2. Quarterly statistics of completeness and timeliness of the SMOS NN product

Figure 1 shows the time series of daily file completeness and timeliness as defined by files that are delivered to ESA within 240 minutes of the initial observation time in the corresponding input BUFR file. The percentages are calculated by dividing the total time covered in the output files by the 24 hours in any single day. For example, for a single day if there are 30 BUFR files covering 48 minutes of data each and 1 file is not produced and 1 file is delivered late then the completeness percentage is 96.67% and the timeliness percentage is 93.33%. The time series covers the third quarter of 2022, 1st July 2022 to 30th September 2022. The data shows that for the vast majority of days the completeness is 100% or very close to 100% and the timeliness is greater than 90%. An explanation of the periods where completeness drops below 95% and timeliness drops below 80% can be found in section 3.

Table 1 shows the monthly and entire quarter mean statistics of completeness and timeliness. The completeness is above 99% for all months and the entire quarter average is 99.9%. The timeliness is 92% or above for all months and the entire quarter average is 94.6%.

Month	Completeness	Timeliness
July	100.0%	94.9%
August	99.7%	92.1%
September	100.0%	96.9%
Quarter	**99.9%**	**94.6%**

Table 1: Monthly mean statistics of completeness and timeliness of SMOS NN nominal soil moisture product delivery

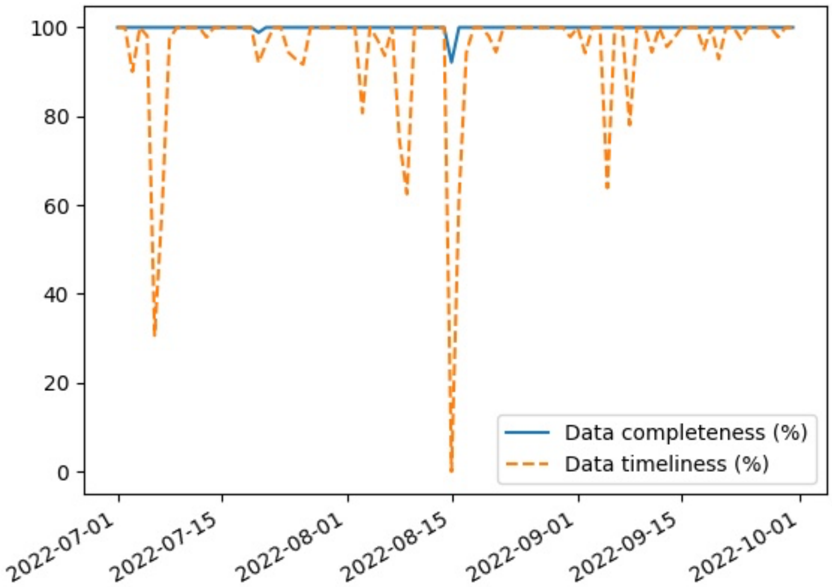

Figure 1: Daily SMOS NN nominal soil moisture production completeness and delivery timeliness percentages (see text for how these are calculated) for the third quarter of 2022: 1st July to 30th September 2022

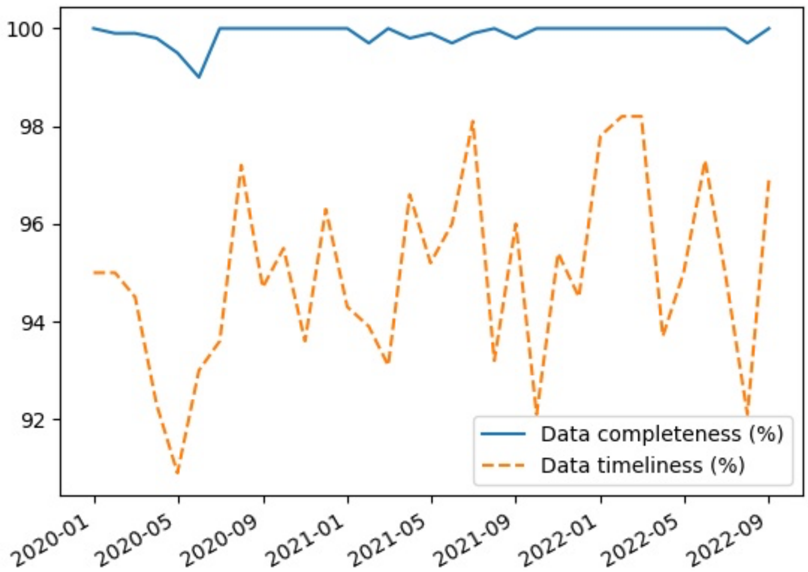

Figure 2: Monthly SMOS NN nominal soil moisture production completeness and delivery timeliness percentages (see text for how these are calculated) for the period January 2020 to September 2022

Figure 2 shows the monthly statistics of completeness and timeliness since January 2020 and shows that the completeness and timeliness have remained fairly constant in quarter 3 of 2022 compared to quarter 2 of 2022.

3. Operational anomalies in this quarter

Figure 1 shows that there was one day where completeness dropped below 95% this quarter. This was on 15th August where the completeness dropped to 92.1% which represents a single BUFR file for a full SMOS orbit not being processed. This instance was caused by a delay to the delivery of the BUFR files. There was one BUFR file covering data between 08:57 UTC and 10:37 UTC which was only received at ECMWF after midnight on the 17th August 2022. The SMOS NN processor has a feature that allows it to catch up on older files but only goes back one previous day, in this case to the 16th August. Because this file contained data from the 15th August and was delayed by over 36 hours, the processor did not process this file. This is the first time such an event has taken place in the past 3 years and so is a very rare event. Also, the purpose of the SMOS NN processor is to provide data in NRT, so a delay of 36 hours in producing the product means it would no longer be considered NRT and therefore, in this case, the product would be of limited value. There are some other days where the percentage drops very slightly below 100% and these are due to a small number of input SMOS BUFR files containing only ocean points. When the neural network processor encounters such a file it skips the file because the neural network product is only validly produced over land.

Figure 1 also shows that there are several days in the past three months where the timeliness drops significantly below 80%, namely 6th July, 7th July, 8th August, 9th August, 15th August, 16th August, 5th September and 8th September, where it drops to 30.7%, 59.2%, 74.4%, 62.5%, 0.0%, 61.2%, 63.9% and 78.0% respectively. Most of these significant drops were caused by ESA delays to the delivery of the BUFR files due to a degraded near-real time (NRT) dissemination service. On 8th and 9th August the delay was due to a MIRAS CCU reset. On 15th August, 16th August and 5th September this was due to network issues at ESAC. These events are out of ECMWF's control, so no corrective action can be taken to stop these events happening in the future.

On the 6th and 7th July the delay was due to a change made at ESAC to only accept relative directory paths when retrieving the BUFR files. At ECMWF we were using absolute directory paths which meant the retrieval of the BUFR files were delayed. The problem was investigated quickly, and the change was reverted. Later in the quarter, on 2nd August, ECMWF changed the acquisition system to use relative paths and the original change was reinstated with no adverse effects this time.

On the 8th September the delay was due to a backlog in the availability of the ECMWF File Storage system (ECFS) used at ECMWF to archive the products, as forewarned in the 2022 quarter 2 report (Weston & de Rosnay, 2022). From 8th September to 30th September the ECMWF data handling system (DHS) started to be moved from Reading to Bologna. Thanks to work carried out over the past year to make the SMOS NN processor more robust, there was only a very short and unavoidable drop off in the performance of the SMOS NN processor linked to this event. From 9th September onwards the SMOS NN processor performed nominally despite the continuation of the DHS move.

4. Comparisons between the ESA nominal and ECMWF assimilation neural network products

In this section the retrieved soil moisture from both the nominal neural network product delivered to ESA and the assimilation neural network product used at ECMWF will be compared. The month chosen for the comparison is August 2022 as this is the middle month of the 3rd quarter.

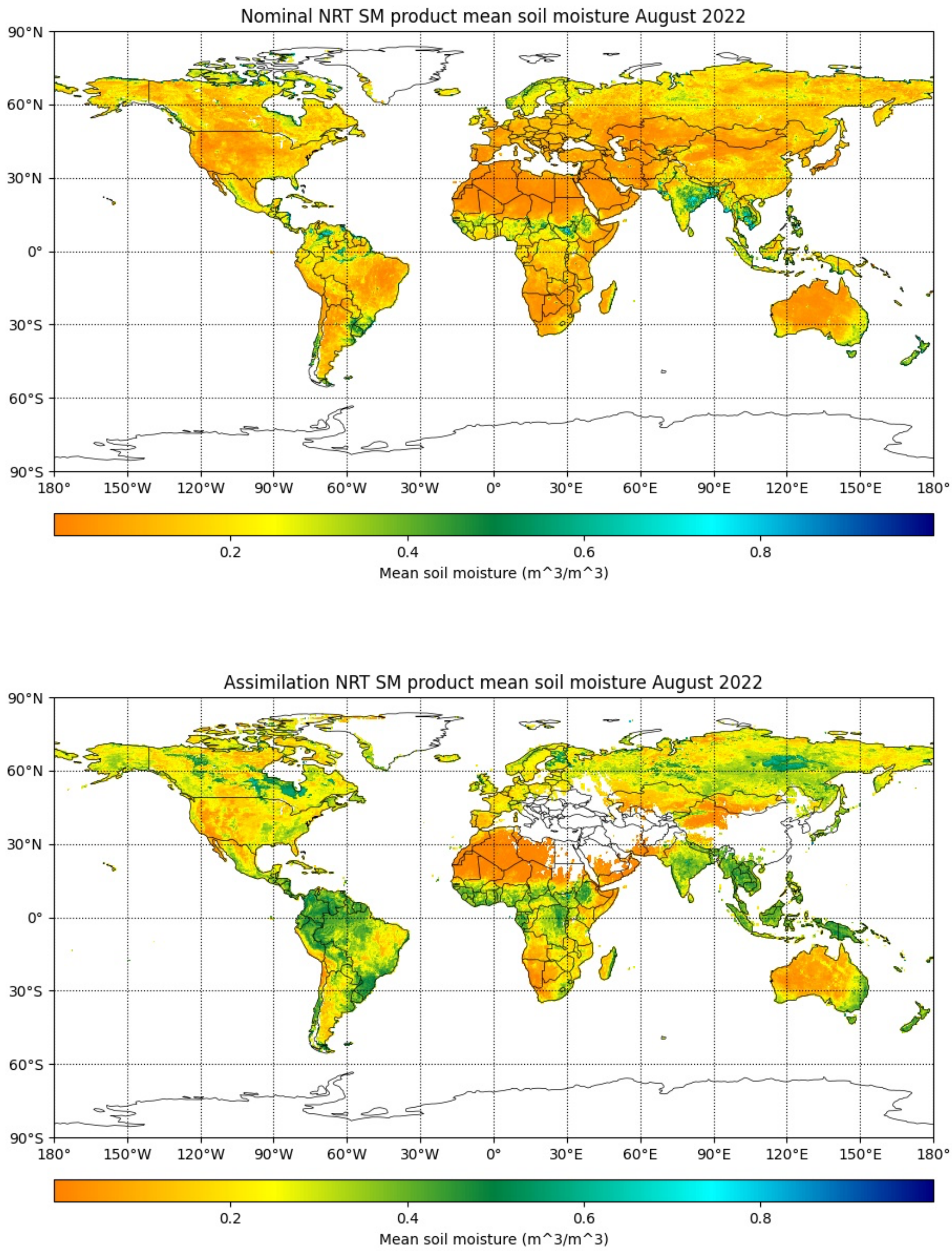

Figure 3: Mean retrieved soil moisture (m^3/m^3) for August 2022 for the nominal NRT product (upper) and assimilation NRT product (lower)

Figure 3 shows that data is missing over China and the Middle East for the ECMWF assimilation product due to extensive radio frequency interference (RFI) in the SMOS brightness temperatures over those regions. These areas are not missing for the ESA nominal product due to a different use of RFI flags in the training of the nominal and assimilation products. It has been decided to leave the nominal and assimilation products as they are until the next re-training when the use of the RFI flags in the training will be re-assessed and made more consistent between the two products.

Figure 3 also shows that the two products have significant mean differences with the ECMWF assimilation soil moisture product generally moister than the ESA nominal product in August 2022. The maps show that the differences are largest in the tropics (over South America, central Africa and the maritime continent in particular) and the Northern high latitudes (Siberia and Northern Canada). The products are in better agreement over Europe, the US as well as in arid regions. The differences are due to the different datasets which the two neural networks are trained on and are consistent with what is seen in July and September 2022 as well as other months throughout the year. The nominal ESA product is trained on historical values of SMOS level 2 soil moisture whereas the ECMWF assimilation product is trained on the ECMWF model soil moisture. These datasets have different characteristics and represent different soil depths which lead to the differences in figure 3. The SMOS level 2 soil moisture represents the top most 2-3cm of soil whereas the ECMWF model soil moisture represents the top most 7cm of soil.

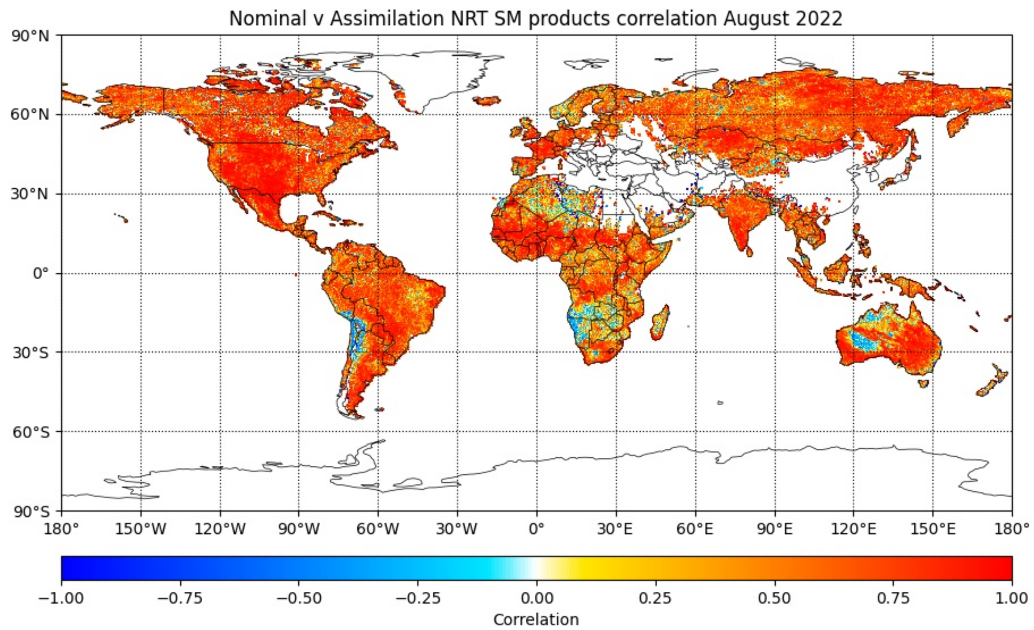

Figure 4: Correlation between the ESA nominal neural network product and the ECMWF assimilation neural network product in August 2022

Figure 4 shows that the two products have the strongest correlations in South America, Eastern Australia as well as the central US and Western Europe. There are moderate correlations in the remainder of the

Northern mid-latitudes and tropics with the weakest (and sometimes negative) correlations over arid regions such as the Sahara desert, Namib desert, the Andes and Western Australia.

5. Investigation into the SMOS NN product assimilation at ECMWF

The performance of the assimilation of the SMOS NN product into the ECMWF land data assimilation has been investigated over the last few months.

Figure 5 shows that the vast majority (~90%) of available SMOS NN soil moisture observations are rejected within the assimilation system. This leaves around 10% of available observations to be assimilated. The main reason for these rejections is a tight check on the background departures. The background departures are the differences between the observations, in this case the SMOS NN derived soil moisture, and the model background soil moisture. If the absolute background departure is greater than $0.1m^3/m^3$ then the observation is rejected. It is worth noting that this ratio of only using ~10% of observations is fairly typical of other types of satellite observations used in both the atmospheric and land data assimilation systems.

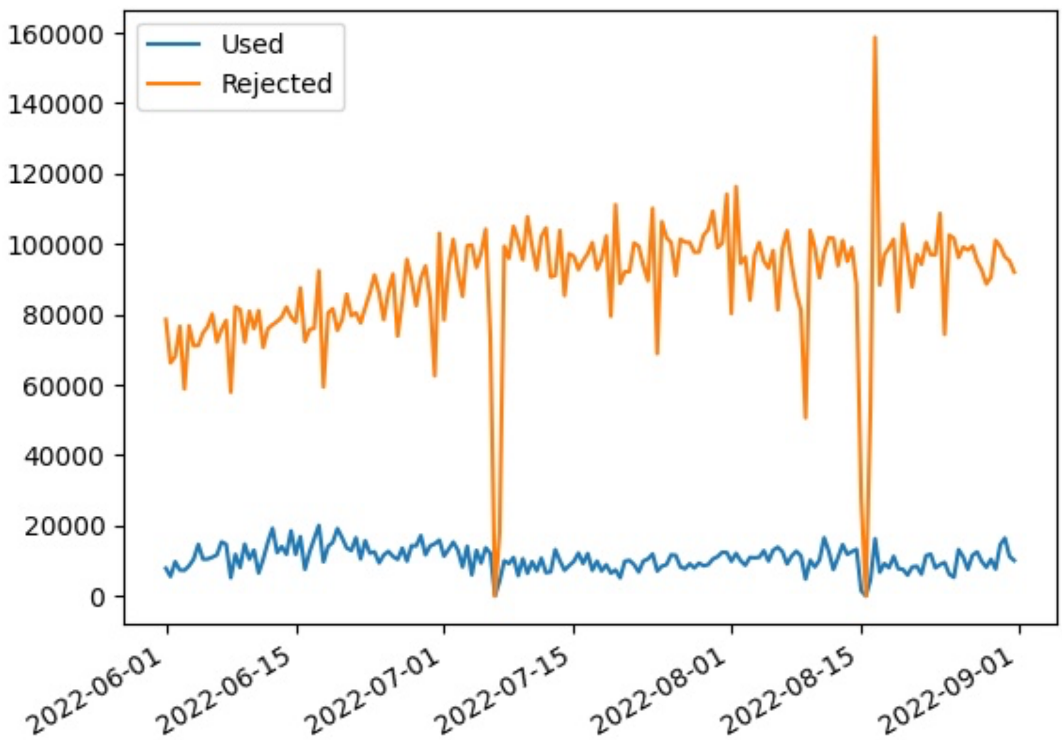

Figure 5: Time series of used (blue) and rejected (orange) numbers of SMOS NN observations per 12-hour cycle from 1st June 2022 until 31st August 2022

Figure 6 shows that there are significant local biases between the observations and the model background which exceed $0.1m^3/m^3$ in many areas of the globe. This explains why a large number of observations are rejected by the background departure check, which is protecting the analysis from the influence of the worst of these biases.

It is well known that land surface models have significant regional biases and it is possible that using a globally trained NN is not able to reproduce or correct for these biases sufficiently. In addition, the NN is trained using a single set of training data that covers multiple years and therefore accumulates many different geophysical and seasonal conditions together. The soil moisture varies hugely in certain areas of the globe with distinct wet and dry seasons and this can lead to significant model biases where, for example, the model precipitation is not able to represent reality or where deficiencies in the model processes at the land-atmosphere interface lead to biases in the soil moisture. These seasonal biases are not represented in the SMOS NN because a single set of weights is calculated for the entire period of the training data.

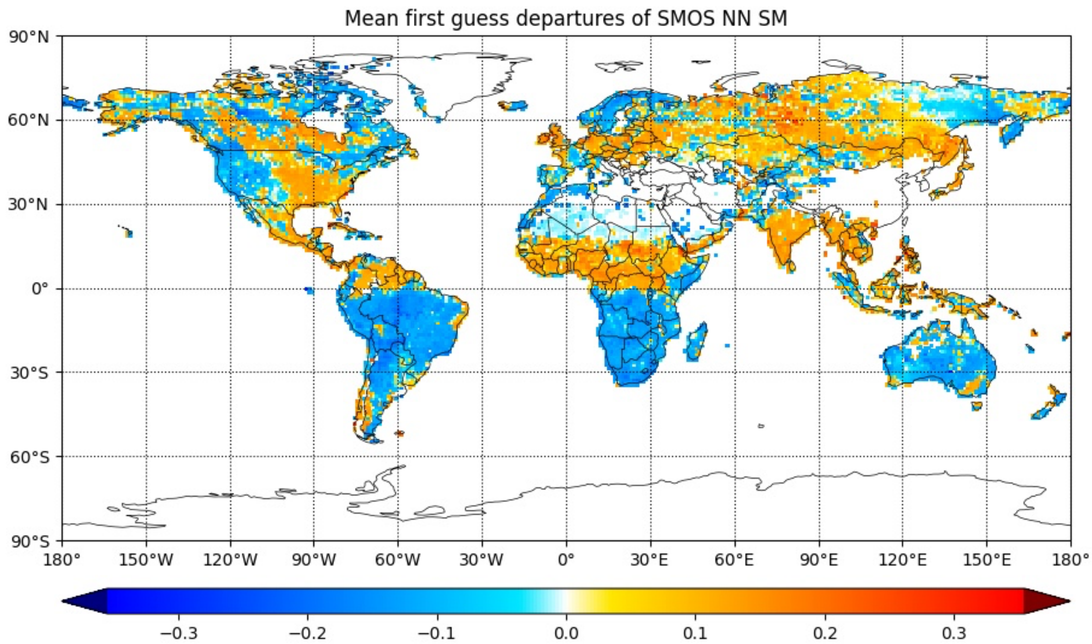

Figure 6: Gridded map of the mean background departures for the SMOS NN observations against the ECMWF model background soil moisture before quality control. Statistics are accumulated between 3rd June 2020 and 31st August 2020

If one, or both, of these reasons are to blame for the significant biases in the background departures then there are possible solutions to correct the bias and therefore allow more data to be assimilated. One solution would be to train a more localised NN to help correct the geographical regional biases. This could be done by adding longitude and latitude to the inputs of the neural network or simply by training separate NNs for each model grid point. One of the challenges associated with this approach would be to use a large enough sample to represent all possible geophysical conditions at every grid point. The aim of this approach would be to correct the biases at source.

A more pragmatic solution would be to use a bias correction to correct the biases within the assimilation system itself. This approach is used for the vast majority of satellite data which is assimilated into the atmospheric data assimilation system (Dee, 2004). A bias correction in the form of a CDF-matching approach is also used for the assimilation of ASCAT derived soil moisture into the ECMWF land data assimilation system (Scipal et al, 2008). This existing bias correction scheme is static and requires frequent re-training following significant changes to the characteristics of the model or observations significantly change. Recently, this approach has been enhanced to include an adaptive bias correction scheme which can adjust to seasonally varying biases automatically. This approach is being tested for ASCAT and extended to the SMOS NN too. Preliminary results show that the adaptive bias correction is able to reduce the bias between the SMOS NN observations and the model soil moisture, as shown in figure 7. There are also small improvements to the forecast scores when applying this bias correction (not shown). These results are still very preliminary and further experiments to tune the adaptive bias correction are ongoing but this could be a promising short-term solution to the bias issues.

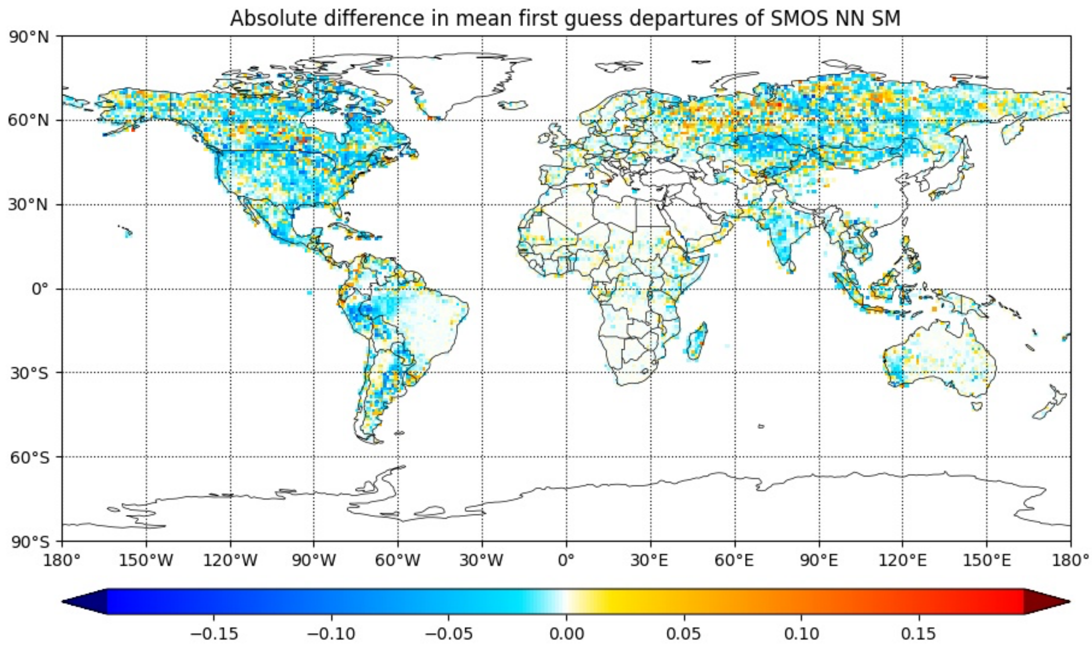

Figure 7: Absolute difference in mean background departures for the SMOS NN observations when running an experiment with an adaptive bias correction applied to the data

6. References

Dee, D., 2004. Variational bias correction of radiance data in the ECMWF system. In: ECMWF Workshop on Assimilation of high spectral resolution sounders in NWP, 28 June - 1 July 2004. ECMWF, ECMWF, Shinfield Park, Reading, pp. 97–112. URL https://www.ecmwf.int/node/8930

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

Weston, P.; de Rosnay, P. Quarter 2 2022: Operations Service Report. SMOS ESL contract 4000130567/20/I-BG.