Implementation of SMOS data monitoring in the ECMWF Integrated Forecasting System

Joaquín Muñoz Sabater, Anne Fouilloux and Patricia de Rosnay

ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
Joaquin.Munoz@ecmwf.int

ABSTRACT

The recent launch of the Soil Moisture and Ocean Salinity (SMOS) satellite of the European Space Agency opens the way to use a new type of satellite data very sensitive to soil moisture and ocean salinity. The European Centre for Medium-Range Weather Forecasts (ECMWF) has developed an operational chain in order to monitor this new type of remotely sensed data and incorporate it in their surface analyses. SMOS data will be assimilated in the ECMWF Integrated Forecast System and it is expected to have an impact in the weather forecast at short-medium range. Previous to assimilation, the quality of the data has to be assessed. This can be done through monitoring activities, which is the main purpose of this paper. Monitoring is a routinely task performed with all satellite data, and among others it makes it possible to localize temporal (or spatial) bias or drifts in the real data, thus providing almost near real time reports to the calibration teams which can act consequently. In this paper the implementation of SMOS data in the Integrated Forecast System of ECMWF is discussed. Special emphasis is given to soil moisture. The system was developed using a simulated file for the Level-1C data processor and it has been tested with the first flow of available NRT Level-1C brightness temperatures from the commissioning phase. Some preliminary results are presented in this paper.

1 Introduction

The successful launch of the Soil Moisture and Ocean Salinity (SMOS) satellite of the European Space Agency [9] is already providing an unprecedented new source of remotely sensed data sensitive to soil moisture in land and salinity in the oceans. Soil moisture has been extensively identified as a critical land variable due to its strong influence in the exchanges of water, energy and carbon fluxes at the interface between the soil, vegetation and the lowest level of the atmosphere. On the other hand, sea surface salinity is needed to know the water density which in turn has a strong influence on oceanic circulations.

While a good estimate of salinity over time in ocean re-analysis is important to address seasonal forecast and climate change ([8]), a good estimation of soil moisture has a direct impact on precipitation and air temperature predictability at short and medium ranges [6, 5]. Remotely sensed microwave observations from 1 to 10 GHz are used to obtain information about water content of a shallow near surface layer [15]. In this microwave region, attenuation from clouds and vegetation is smaller than a higher frequency [16]. For example, C-band (5.255 GHz) active microwave data from the Advanced Scatterometer instrument (ASCAT) on-board of the MetOp platform in synergy with time series of the European Remote Sensing (ERS) 1/2 data is currently being used by the Vienna University of Technology to produce a global soil moisture map product [1]. One of the main problems of active observations is that they are very sensitive to local soil roughness changes. Comparatively, factors such as vegetation or soil roughness are less significant in passive remote sensing. Satellite missions carrying passive instruments in microwaves low-frequencies have demonstrated the potential to obtain information about soil moisture [12, 3] . However, the information they provide is limited to the most superficial soil layer and to areas of low to moderate vegetation [11].

Passive L-band measurements were identified as the most suitable ones for soil moisture retrievals [9]. Nonetheless, the heavy cost and technological challenge of arranging a large antenna in L-band (the satellite antenna size is directly proportional to the squared wavelength [17]) in a single platform has prevented an earlier spatial L-band mission. For SMOS, an antenna of 8 m is required to comply with the spatial resolution requirements of the mission. In SMOS this problem is overcome by applying the interferometric technique. Instead of one large antenna, sixty nine little receivers installed in three Y-shaped arms collect the radiation emitted by the Earth's surface between 1400 and 1427 MHz. The phase difference measured between the individual receivers makes it possible to reconstruct an image with the science requirements, i.e., volumetric soil moisture with 4% accuracy and spatial resolution 40-50 km and salinity in open waters with 0.1 psu accuracy for a 10-30 day average and an open ocean area of 200 km x 200 km [9].

As a Numerical Weather Prediction (NWP) centre, the European Centre for Medium-Range Weather Forecasts (ECMWF) is receiving a near real time (NRT) product (Level-1C brightness temperatures), which is automatically recovered from the SMOS Data Processing Ground Segment (DPGS). To take fully advantage of this NRT product, ECMWF has implemented this new data type in the Integrated Forecast System (IFS). This is a challenging task due to several reasons: a) the multi-angular and multi-polarized characteristics of the instrument measuring principle provides an unique dataset which has to be treated independently from any other source of satellite data, b) the high angular resolution of the data provides a very large set of daily data which cannot be completely ingested in the NWP system. Although the ultimate objective is to study the impact of SMOS data in the weather prediction, the quality of the data has to be demonstrated first. A possible validated way is to routinely monitor the data at global scale. This makes it possible to report preliminary strengths and weaknesses on first SMOS data and contribute at key decision points during the commissioning phase. In this paper a review of the main steps involved in the design of the SMOS data implementation in the IFS is presented. The developed chain is tested with the first available data sets of real SMOS data.

2 Data product used at ECMWF

The product used at ECMWF is the NRT Level-1C brightness temperatures. SMOS Level-1C products constitute reprocessed Level-1b and they differ with the latter in that they are geographically sorted swath-based maps of brightness temperature. The geolocated product received at ECMWF is arranged in an equal area grid system called ISEA 4H9 (Icosahedron Snyder Equal Area grid with Aperture 4 at resolution 9) [14]. For this grid, the centre of the cell grids are at equal distance of 15 km over land, with a standard deviation of 0.9 km. Over oceans the grid has coarser resolution, that is half of the resolution over land.

The data is organized in messages. Each message corresponds to a snapshot where the integration time is 1.2 seconds, as this is the time in which all correlations of a single scene are measured. Within each snapshot a number of subsets are found which are referencing all the sensed point of the ISEA grid. In average, each snapshot contains around 4800 subsets over land if the instrument runs at dual-polarisation mode. At this running mode, data set records are generated alternatively each 1.2 seconds at horizontal (H) and vertical (V) polarisation. In full-polarisation mode all receivers in the three arms are in the same polarisation for the first integration, whereas in the next integration the receivers in an arm switch the polarisation and two data set records are generated, thus doubling the information per snapshot.

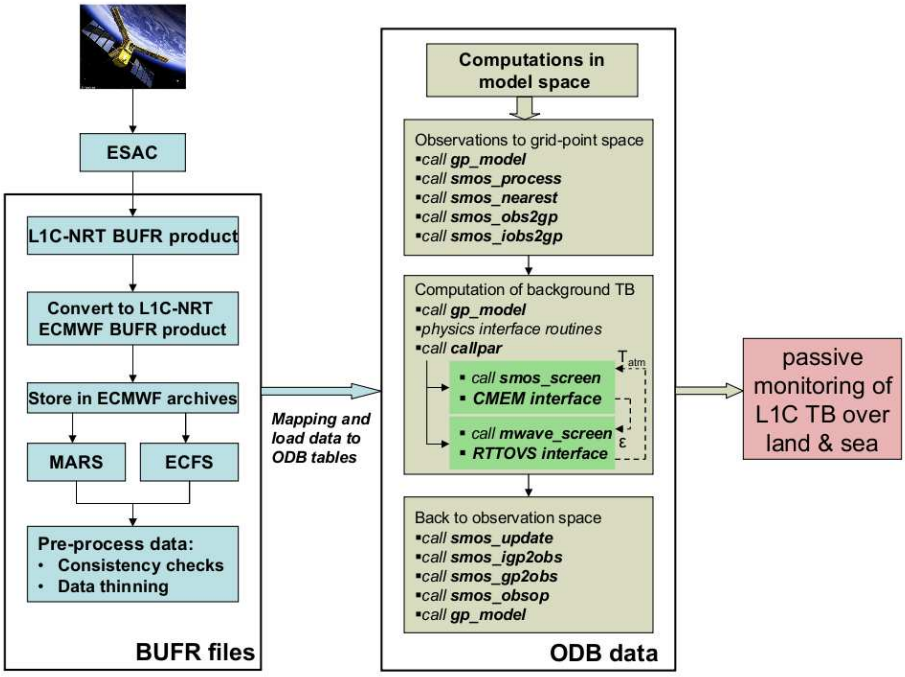

Figure 1: General organigram representing how SMOS data is implemented in the Integrated Forecast System. The left box implies pre-processing of raw NRT Level-1C SMOS data. The middle box represents implementations in model space. The right box represents operational monitoring products.

3 Implementation

The introduction of a new type of satellite data in a NWP model is a challenging task. Even more so when the data comes from a new measuring technique which has never been tested before. This is the case of the SMOS research mission, being the second Earth's Explorer mission of the ESA Living Planet Programme. In this section the main steps involved in the implementation of SMOS data in the ECMWF IFS are addressed. The implementation of SMOS data in the IFS can be classified in two big sets of tasks: Data pre-screening and model space computations.

In the first group of tasks, all the data is checked to be a consistent and valid set of values in the same format as the data is received: the Binary Universal Form for the Representation of meteorological data (BUFR) format. All the data that goes through the first system of filters is transformed in a format acceptable to the IFS structure, the so called Observational Data Base (ODB) format, and implementations in model space are performed in preparation for comparison with a model simulation (see Fig.1).

3.1 Data pre-screening

Firstly, NRT raw data (see section 2) processed at the European Space Astronomy Centre (ESAC) in Madrid (Spain) is retrieved from the DPGS and slightly modified to feed the pre-screening tasks. The resulting product is subsequently stored in ECMWF archive systems. SMOS data is extracted from these archives to be used for consistency and quality checks. These checks are the following:

1. Generic checks: files which fail to contain crucial header information are rejected: it is checked that files are encoded in BUFR format, date and time are complete, geographic coordinates are

not missing and instrument data corresponds to SMOS data,

2. The validity of data is checked; a) individual observations are checked to be in a correct geographical position and b) brightness temperatures are checked to be in the range of physically reasonable values (not lower than 50 K and not greater than 350 K),

3. Data is thinned to reduce the size of SMOS files as input for the IFS.

Data thinning is a critical step in so far as it selects which data from the original files will be monitored, but also which data has a potential input to correct the soil moisture state and the ocean salinity value through an assimilation experiment. Thinning is a mandatory step in order to avoid any redundant data in the assimilation system and also due to the large volume of data contained in a single orbit. The reason is because for each grid point of the ISEA grid typically one hundred records of brightness temperatures between 0 and 65 degrees are provided. That means that if the instrument operates normally, more than 1 Gb of data would be provided for a timeslot of 6h in dual-polarisation mode, whereas this quantity can be doubled if the instrument operates in full-polarisation mode. This amount of data cannot all be introduced in the IFS just for a single satellite instrument, taking into account that many other satellite data are used simultaneously. Thinning is therefore mandatory. Data thinning can be done in many different ways. For monitoring purposes the thinned subset of observations should keep the same statistical characteristics as the original set. Several experiments investigated the optimal configuration of the thinning exercise (not shown). The conclusion was that a simple filter which keeps only one out of 10 subsets of a BUFR message was the best adapted for monitoring purposes as 1) the data volume was reduced by $\sim$ 90% but still keeping the statistical signature of the original data set and 2) there is still the possibility of using any incidence angle for monitoring or assimilation experiments. At the end of the pre-screening step, a reduced, consistent set of SMOS data keeping most of the angular information is remaining for monitoring purposes.

3.2 Model space computations

The implementation of SMOS data monitoring can basically be done either in observation or in model space. Whether in observation or model space, surface heterogeneities need to be accounted for soil moisture retrieval. In the case of SMOS, the implementation in model space adds several advantages mainly in view of a future comparison with model simulations and assimilation experiments: 1/ all the background fields necessary to simulate brightness temperatures at the top of the atmosphere are available in model grid point space. Thus, it avoids interpolating physical quantities to observation location; 2/ other satellite data sensitive to soil moisture, as AMSR-E data in C-band, is also available in model space, making it possible a comparison with other satellite data.

Fig.1 shows a general organigram about how Level-1C SMOS data received at ECMWF is implemented in the IFS. Those observations which survive the pre-screening filters go (roughly) through a two-step phase:

1. Observations are brought to model grid-point space at the required resolution using the nearest neighbour technique. Along with it a flag mask containing information of the grid point (mainly whether it contains an observation) is created. The number of observations which will be monitored depends on the model grid resolution and the distance limit parameter. At T799 ($\sim$ 25 Km) SMOS observations within the distance limit are found for all grid points (not shown),

2. Observations in model grid-point can potentially be coupled with background values simulated with a forward model operator and thus, obtaining the innovation vector as input for an assimilation scheme,

Table 1: Number of observations rejected for 6 hours of data during the early quality checks phase.

Date	snapshots	subsets	rejections	% rejected
28-11-2009	17940	28203176	147185	0.52
20-12-2009	17592	28739029	58967	0.21
16-01-2010	15347	24322415	34386	0.14

Forward computation will be carried out in model grid-point space following the approach of [7]. To this end, the passive microwave forward model operator CMEM (Community Microwave Emission Model (CMEM), [4, 2]) will be applied. Additionally, the RTTOVS (Radiative Transfer for the advanced TIROS Operational Vertical Sounder)-CMEM interface will also be implemented (see Fig.1). RTTOVS is the fast radiative transfer model used at ECMWF and elsewhere for the assimilation of nadir microwave or infra-red radiances [10]. The RTTOVS version used is version 9. Information can be exchanged between both codes. RTTOVS provides the atmospheric contribution to the total brightness temperature at the top of the atmosphere for frequencies larger than 10 GHz. In turn, CMEM provides the surface emissivity to RTTOVS, thus a more realistic emissivity for the lowest atmospheric level.

4 Results with real data

The chain described in the previous section was tested with a series of real data corresponding to the two first months and a half of the commissioning phase. At this early stage, testing on the different components of the system (instrument, data server, data processor, etc.) is frequent and data is not always available. Hence, the data analysed in this paper was selected mainly based on availability. The comparison of these data sets makes it possible to observe a clear evolution on the quality of the data still in a premature phase. The dates of the data sets are the following:

- Data set 1): 28^{th} November 2009,
- Data set 2): 20^{th} December 2009,
- Data set 3): 16^{th} January 2010.

4.1 pre-screening results

The different quality checks listed in section 3.1 were tested with data sets 1), 2) and 3). Fig.2 shows the number of individual brightness temperatures values rejected as a function of the first 18000 snapshots. This corresponds to the first 6 hours of collected data for these days. Only snapshots with less than 5000 subsets are shown because they correspond to pure H or V polarisation integrations. Cross-correlated polarisations are not shown here since they are not available for all files. This figure clearly shows how the number of rejected radiances is maximum for the 28^{th} of November, when still no calibration in the data was applied, whereas they are significantly reduced the 20^{th} of December and 16^{th} of January. In table 1 a quantitative comparison between the three data sets is presented. It shows how the quality of the data is best in January 2010, with only 0.14% observations rejected after the first group of quality checks. Although this percent is still small in November, it is relatively more significant than in December and January.

It was also checked that the collocation of SMOS observations to a model grid is accurate and that only one observation is associated to a grid point, independently for each polarisation and for several model

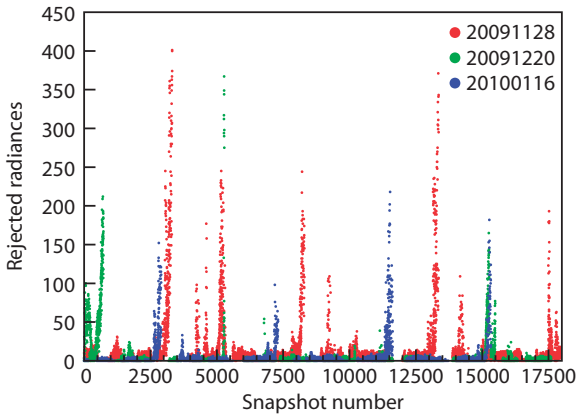

Figure 2: Number of rejected observations at the pre-screening tasks for 6h of Level-1C SMOS data.

grid resolutions (not shown).

4.2 Preliminary assessment of NRT data

The multi-angular global maps of brightness temperatures are a very interesting product of the data monitoring implementation. This data is currently being monitored for different incidence angles (0, 10, 20, 30, 40, 50 and 60) and for both H and V polarisations. Plots are currently available at *http://www.ecmwf.int/research/ESA_pro-jects/SMOS/monitoring/smos_monitor.html*. A simple inspection of time series of this data at global scale makes it possible to observe an evolution and localize possible spatial or angular effects. This is specially important during the early stage of the commissioning phase when calibration activities are needed. Figures 3 and 4 show the brightness temperatures at 40 and 50 degrees incidence angle, respectively. Left panels correspond to H polarisation, whereas right panels show the V polarisation. Figures are shown for days 1) (top figures), 2) (middle figures) and 3) (bottom figures). Firstly, each figure shows a clear evolution on the quality of the data, from day 1) (top) to day 3) (bottom). The day in November (top) is presented as to be very noisy. This is data received within the two first weeks of the instrument switch on phase. At this stage no calibration was carried out yet, radio-frequency interference was present in many areas, geolocalisation was not accurate, the data processor was not fully operational, etc. In December a major calibration event took place and the difference in the product is quite significant when comparing top with middle figures. Improvements are present almost everywhere. The data is even better the 16*th* of January, although this needs of a closer look-up and it needs of quantitative results to confirm it. Secondly, at this stage of the commissioning phase it is important to check the correct functioning of the instrument. Days in December and January have an expected behaviour for both polarisations: brightness temperatures values getting colder with the incidence angle for H polarisation and an opposite behaviour for the V polarisation, with both displaying values within an acceptable physical range, as confirmed in the pre-screening phase. This confirms that the novel technique used in SMOS is running successfully. Finally, it is also an objective of data monitoring activities reporting on possible spatial or temporal effects on the data: 1/ For both polarisations, at 50 degrees, it is observed a thin stripe away of the main satellite track; 2/ For V polarisation, over oceans, the most outer sides of the satellite track look colder than the inner part, mainly visible at 40 degrees. 3/ There is still residual RFI over Europe, Middle-East and Asia, which is particularly straightforward to spot when the data looks very ”red” and noisy. However, in January the data look apparently of good quality over the whole America, Australia and South of Africa.

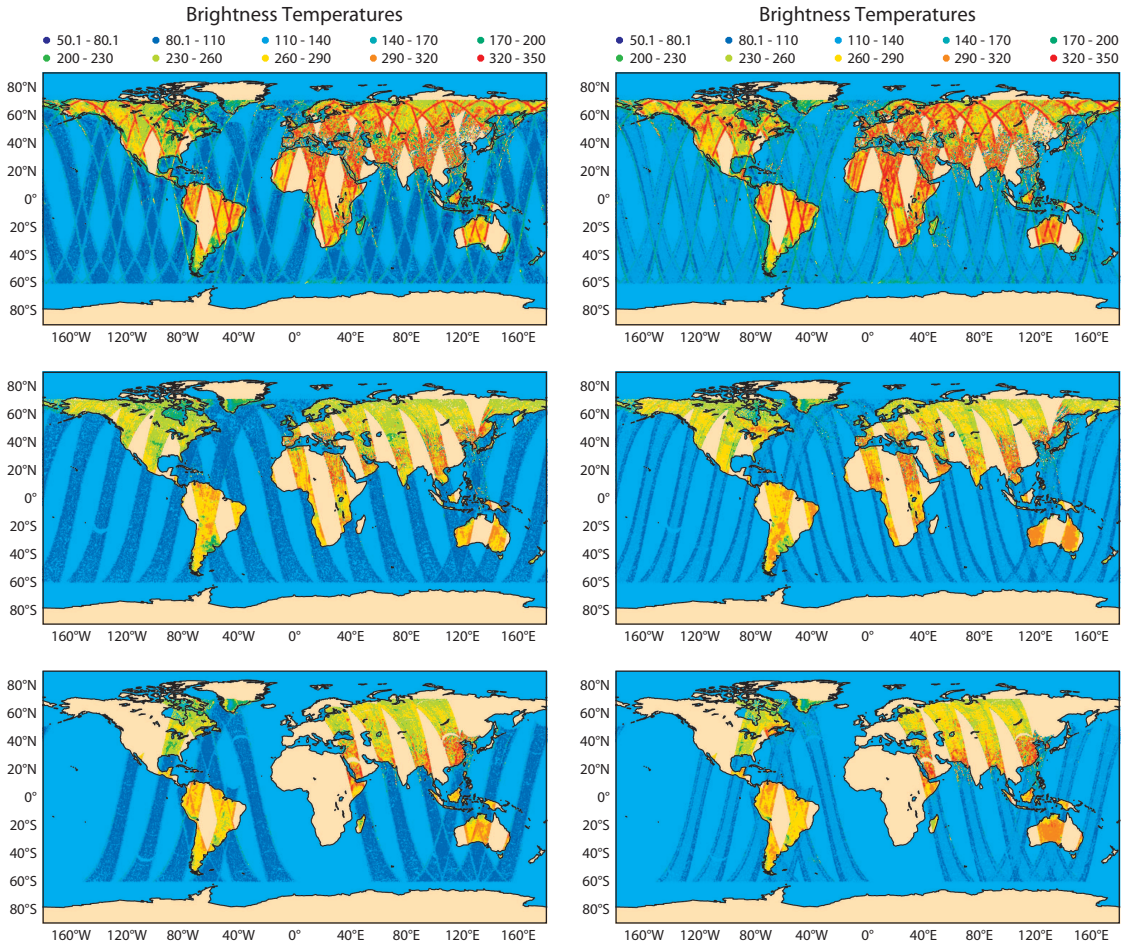

Figure 3: Brightness temperatures for real NRT Level-1C SMOS data at 40 degrees incidence angle. The left panel corresponds to H polarisation whereas the right panel is V polarisation. Figures on top are for the 28th November 2009, middle figures correspond to the 20th of December 2009 and bottom figures to the 16th of January 2010.

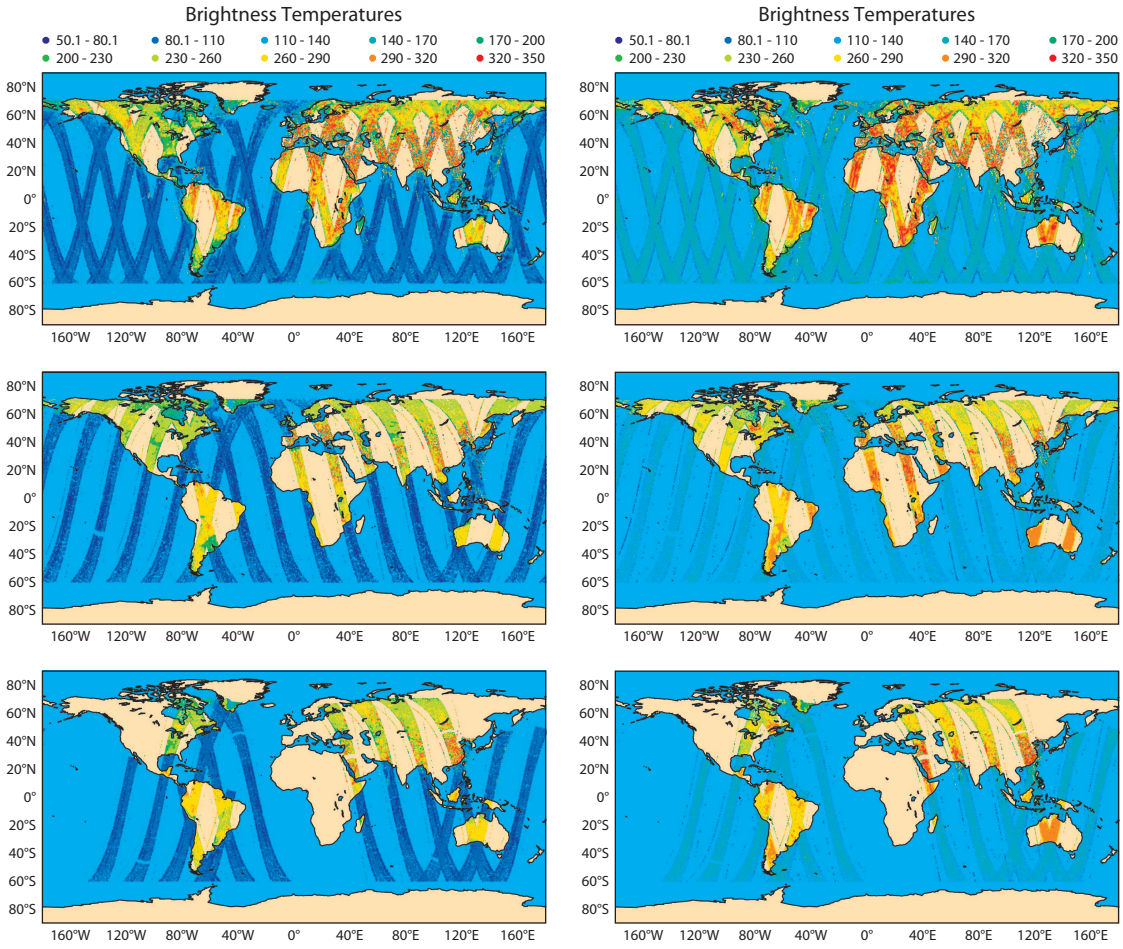

Figure 4: Brightness temperatures for real NRT Level-1C SMOS data at 50 degrees incidence angle. The left panel corresponds to H polarisation whereas the right panel is V polarisation. Figures on top are for the 28th November 2009, middle figures correspond to the 20th of December 2009 and bottom figures to the 16th of January 2010.

5 Summary and Perspectives

The chain described in this paper constitutes the first step towards a full operational monitoring of SMOS data in NRT and the assimilation of SMOS data for soil moisture and ocean salinity analyses. The main steps followed to implement a new satellite data type quantity in a complex system as the ECMWF IFS were described. This is a challenging task for SMOS, since the individual attributes of the data are very different to any other satellite data. The steps addressed in this paper demonstrate that the chain developed at ECMWF runs successfully to fulfil data monitoring requirements, as it is the quality check of each individual observation that enters the system and the collocation of the observations with an atmospheric grid, where first-guess departures can be computed.

The data is still in an early premature stage and calibration and adjustments are still being carried out. Thus, rather than providing an exhaustive analysis of the data, this paper is more oriented towards demonstrating that the SMOS data monitoring chain is efficient. In this paper several NRT files have been tested and a clear enhancement in the quality of the data is observed. There are still a few issues that the scientific teams have to cope with before the instrument enters in a fully operational phase. For example, RFI is still present in Europe, North Africa and Asia and for high incidence angles a side-effect in the satellite track is observed over oceans.

Although the thinning filter selected in the implementation phase is quite simple, at this stage it is considered to be efficient for assimilation purposes. Future activities will investigate more deeply an optimal system to remove noise and sources of RFI from the input data as well as to develop a thinning filter which complies with assimilation requirements. After the quality of the data will be demonstrated, assimilation experiments of SMOS data will be carried out. The main input for the Extended Kalman Filter land surface assimilation scheme at ECMWF are the first-guess departures, i.e. the difference between the observation and the model equivalent. In this sense, the monitoring will be completed by providing the modelled value in the model grid point, where first-guess departures will be obtained. The modelled value will be provided by interfacing the CMEM microwave emission model in the IFS and by using a modular configuration based on the studies of [4], [2] and [13]. Daily, weekly and monthly statistics will be provided for the raw brightness temperature values and the first-guess departures. Based on these statistic reports, information about the observation and model error will be obtained. Finally, the data will be put through another quality control system based on background departure values and they will input the assimilation scheme.

Acknowledgements

This work is funded under the ESA-ESRIN contract number 20244/07/I-LG. The authors would like to thank Milan Dragosavac, Ioannis Mallas and Alfred Hofstadler to support the production of pre-processed data within the NRT chain. Also, special thanks to Lars Isaksen for his constant advice.

References

[1] Z. Bartalis, W. Wagner, V. Naeimi, S. Hasenauer, K. Scipal, H. Bonekamp, J. Figa, and C. Anderson. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.*, 34, 2007. L20401, doi:10.1029/2007GL031088.

[2] P. de Rosnay, M. Drusch, A. Boone, G. Balsamo, B. Decharme, P. Harris, Y. Kerr, T. Pellarin, J. Polcher, and J.-P. Wigneron. AMMA Land Surface Model Intercomparison Experiment coupled

to the Community Microwave Emission Model: ALMIP-MEM. *J. Geophys. Res.*, 114, 2009. D05108, doi:10.1029/2008JD010724.

[3] C. S. Draper, J.-F. Mahfouf, and J. P. Walker. An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *J. Geophys. Res.*, 114, 2009. D20104, doi:10.1029/2008JD011650.

[4] M. Drusch, T. Holmes, P. de Rosnay, and G. Balsamo. Comparing ERA-40 based L-band brightness temperatures with Skylab observations: A calibration / validation study using the Community Microwave Emission Model. *Journal of Hydrometeorology*, 10:213–226, 2009. doi: 10.1175/2008JHM964.1.

[5] M. Drusch and P. Viterbo. Assimilation of screen-level variables in ECMWF's Integrated Forecast System: A study on the impact of the forecast quality and analyzed soil moisture. *Mon. Wea. Rev.*, 135(2):300–314, 2007. doi - 10.1175/MWR3309.1.

[6] L. Ferranti and P. Viterbo. The european summer of 2003: Sensitivity to soil water initial conditions. *J. Climate*, 19:3659–3680, 2003.

[7] A.J. Geer, P. Bauer, and C.W. ODell. A revised cloud overlap scheme for fast microwave radiative transfer. *J. Appl. Meteor.*, (D18102, doi:10.1029/2005JD006691), 2009. early online release available at http://ams.allenpress.com.

[8] K. Haines, J. Blower, J-P. Drecourt, C. Liu, A. Vidard, I. Astin, and X. Zhou. Salinity assimilation using S(T) relationships. *Mon. Wea. Rev.*, 134(3):759–771, 2006.

[9] Y. Kerr, P. Waldteufel, J.-P. Wigneron, J.-M. Martinuzzi, J. Font, and M. Berger. Soil moisture retrieval from space: The soil moisture and ocean salinity (SMOS) mission. *IEEE Trans. Geosc. Remote Sens.*, 39(8):1729–1735, 2001.

[10] M. Matricardi, F. Chevallier, and J.-N. Thpaut. An improved general fast radiative transfer model for the assimilation of radiance observations. *Quart J. Roy. Meteor. Soc.*, 130:153–173, 2004.

[11] E. Njoku and S. K. Chan. Vegetation and surface roughness effects on AMSR-E land observations. *Remote Sens. of Environ.*, 100:190–199, 2006.

[12] R.H. Reichle, R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku andD.B. McLaughlin, and D.Entekhabi. Comparison and assimilation of global soil moisture retrievals from the advanced microwave scanning radiometer for the earth observing system (AMSR-E) and the scanning multichannel microwave radiometer (SMMR). *J. Geophys. Res.*, 112, 2007. D09108, doi:10.1029/2006JD008033.

[13] J.M. Sabater, P.de Rosnay, and G.Balsamo. Sensitivity of L-band NWP forward modelling to soil roughness. *Int. J. Remote Sens.*, 2010. in review.

[14] K. Sahr, D. White, , and A. J. Kimerling. Geodesic discrete global grid systems cartography. *Cartography and Geographic Information Science*, 30(2):121–134, 2003.

[15] T.J. Schmugge. Remote sensing of soil moisture: Recent advances. *IEEE Trans. Geosc. Remote Sens.*, GE21:334–344, 1983.

[16] F. T. Ulaby, R. K. Moore, and A. K. Fung. *Microwave Remote Sensing: Active and Passive. Vol. I*, page 56 pp. Addison-Wesley, 1981.

[17] F. T. Ulaby, R. K. Moore, and A. K. Fung. *Microwave Remote Sensing, Active and Passive. Vol. II*, chapter Physical mechanisms and empirical models for scattering and emission, pages 816–921. Artech House, Boston, MA, USA, 1982.