ESA CONTRACT REPORT

Contract Report to the European Space Agency

Tech Note - Phase-I - Final Report

J. Muñoz Sabater,
P. de Rosnay, A. Fouilloux,
M. Dahoui, L. Isaksen,
C. Albergel, I. Mallas,
T. Wilhelmsson

Technical Note - Phase-I - Final Report
ESA/ESRIN Contract 20244/07/I-LG

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

ECMWF

Series: ECMWF ESA Project Report Series

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

Contact: library@ecmwf.int

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

Contract Report to the European Space Agency

Tech Note - Phase-I - Final Report

Authors: J. Muñoz Sabater,
P. de Rosnay, A. Fouilloux,
M. Dahoui, L. Isaksen,
C. Albergel, I. Mallas
T. Wilhelmsson
Technical Note - Phase-I - Final Report
ESA/ESRIN Contract 20244/07/I-LG

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, Berkshire, UK

January 2013

	Name	Company
First version prepared by (January 2013)	J. Muñoz Sabater	ECMWF
Quality Visa	E. Källén	ECMWF
Application Authorized by		ESA/ESTEC

Distribution list:

ESA/ESRIN
Luc Govaert
Susanne Mecklenburg
ESA ESRIN Documentation Desk

SERCO
Raffaele Crapolicchio

ESA/ESTEC
Tania Casal
Matthias Drusch
Klaus Scipal

ECMWF
HR
Division & Section Heads

Contents

Abstract

Contracted by the European Space Agency (ESA), the European Centre for Medium-Range Weather Forecasts (ECMWF) is involved in global monitoring and data assimilation of the Soil Moisture and Ocean Salinity (SMOS) mission data. For the first time, a new innovative remote sensing technique based on radiometric aperture synthesis is used to observe soil moisture over continental surfaces and ocean salinity over oceans. Since SMOS was launched in November 2009, ECMWF has been monitoring SMOS brightness temperatures, and in near real time since November 2010. Some recent technical developments have also made it possible to incorporate SMOS data within the ECMWF Simplified Extended Kalman Filter (SEKF) for the analysis of soil moisture.
This is the final report of the ESA contract 20244/07/I-LG, which is the phase-I of the monitoring-assimilation study of SMOS brightness temperatures at ECMWF. The objective is to provide a summary of the main achievements reached in phase-I. An extended description of the technical implementation of SMOS data in the ECMWF SEKF scheme is provided too.

1 Introduction

The launch of the Soil Moisture and Ocean Salinity (SMOS) satellite of the European Space Agency (ESA) opened the door to use a new type of satellite data very sensitive to soil moisture for numerical weather prediction applications. The European Centre for Medium-Range Weather Forecasts (ECMWF) has developed an operational chain which makes it possible to process SMOS observed brightness temperature in Near Real Time (NRT) and compare it with a model equivalent. This process has been very challenging, given the very particular characteristics of the Microwave Imaging Radiometer using Aperture Synthesis (MIRAS) instrument on board the SMOS platform. Indeed, this is the first time that a 2D-interferometric radiometer is used to provide information of the amount of water stored in the most shallow layer of continental surfaces.
The objective of this report is to provide a summary of the main technical milestones reached in phase-I of this study. They enable the ECMWF system the possibility to monitor SMOS brightness temperatures in NRT, and assimilate this data in the Simplified Extended Kalman Filter (SEKF) for the analysis of soil moisture.

Table 1 provides a list of the workpackages and deliverables of Phase-I; In MS1TN-P1 [**?**] the Global Emission model used to simulate SMOS brightness temperatures was described. The structure developed within the Integrated Forecasting System (IFS), used as interface between SMOS data and the IFS, is extensively described in MS1TN-P2 [**?**]. Details about the operational pre-processing chain, the collocation software development and the offline monitoring suite is given in MS2TN, parts 1/2/3 [**?**], respectively. A first analysis of the monitoring statistics and an extensive description of the monitoring products are provided in [**?**]. The analysis of the statistics during the period November 2010- November 2011 is reported in [**?**]. In this report a summary of these activities is given, and an overview of the technical implementation for the assimilation of SMOS data in the SEKF is provided too.

2 Global Emission Model

One of the main components necessary to monitor and assimilate SMOS Level-1-based data is the forward model operator, which is able to bring the observations and a model equivalent of the observation to the same space for comparison purposes. In this context ECMWF developed the Community Microwave Emission Modelling Platform (CMEM), as the ECMWF forward operator for low-frequency passive microwave brightness temperatures from 1 to 20 GHz. The modularity of this code makes it specially suitable for implementation within the IFS. Four different modules for the soil, vegetation, snow and atmospheric microwave emission are

Workpackage name	Deliverable
MS1TN-P1	Global Surface Emission Model
MS1TN-P2	IFS Interface
MS2TN-P1	Collocation Software Development
MS2TN-P2	Operational Pre-Processing Chain
MS2TN-P3	Offline monitoring suite development
MR1	Continuous Monitoring Report - Part I
MR2	Continuous Monitoring Report - Part II

Table 1: Summary of the workpackages and corresponding deliverables of Phase-I, of the ESA-ECMWF study on SMOS monitoring and data assimilation.

used in CMEM. In [**?**] a wide overview of the CMEM main physical parameterisations and other related technical documentation is provided. Relevant results compiled from three different intercomparison studies are included too, using L-band observations from the NASA Skylab mission in 1973-1974 [**?**], in situ L-band observations of the SMOSREX (Soil Monitoring Of the Soil Reservoir Experiment) site in South-West France [**?**], and C-band observations provided by the Advance Microwave Scanning Radiometer on Earth Observing System (AMSR-E) on the NASA's AQUA satellite over the AMMA area in West Africa [**?**]. As an example, Fig. 1 represents the time-latitude diagram of the horizontally polarised brightness temperatures at C-band from AMSR-E and using 8 different land surface models coupled to CMEM, under the ALMIP-MEM framework [**?**]. A wet patch shown by AMSR-E is well captured by the eight models coupled to CMEM, but the amplitude is either overestimated or underestimated. However it emphasizes the good agreement between the observations and the model-based simulations. These studies validate the skill of CMEM to accurately represent the soil emission under different conditions. It also proposes the most adequate parameterisations to be used for each component of the soil contributing to low frequencies microwave emission. At present, CMEM is interfaced with the IFS [**?**], and is providing the first-guess for comparison to SMOS observations at the time of the observations.

ECMWF has also developed a website with lot of complementary information about the CMEM model. All the different versions of the code are freely available to the entire scientific community at the following website: `http://www.ecmwf.int/research/ESA_projects/SMOS/cmem/cmem_index.html`.

3 Acquisition and data pre-screening

The SMOS NRT products are processed at the European Space Astronomy Centre (ESAC) in Madrid (Spain), just a few minutes after the Earth's surface is sensed by the SMOS instrument. ECMWF pull this data from the Data Processing Ground Segment (DPGS) in Binary Universal Form for the Representation of meteorological data (BUFR) format. Then a small format conversion is operationally carried out at ECMWF to produce a BUFR version compatible with the IFS software, just before the data is stored in the internal ECMWF File Storage system (ECFS).
SMOS data is then fetched from the archive for monitoring and/or assimilation purposes. The first important processing step consist on pre-screening the data to avoid redundant or corrupted data to enter the IFS. In the pre-screening task it is checked that all the data contain the crucial information (latitude, longitude, date, etc.) and that the data is not corrupted. The physical value of each individual observation is also checked at this step. Data thinning is performed at this stage too. In the case of SMOS, thinning is a crucial step given the large volume of data. In [**?**] the evolution of the thinning approach used at different cycles of the IFS is described, until the current one implemented in cycle 37R1. This adopts a very flexible approach in regards to

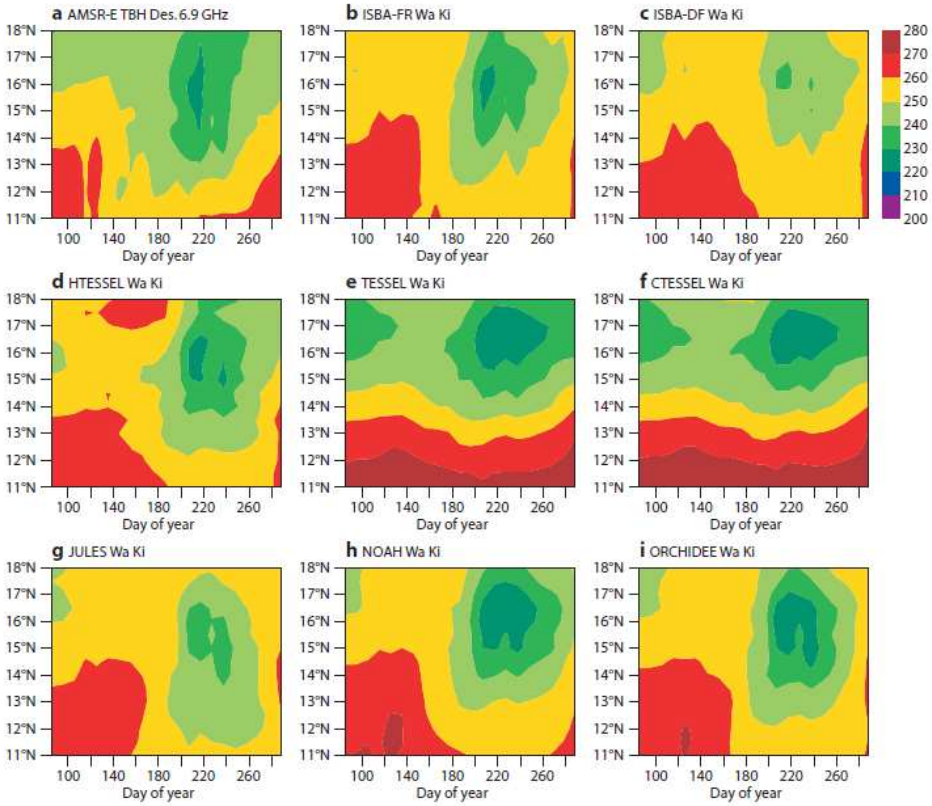

Figure 1: Time-latitude diagram of the horizontally polarised brightness temperature observed by AMSR-E and simulated by ALMIP-MEM.

the configuration of the observations (number of incidence angles, polarisations, field of view, etc.) and steered by an unique namelist. More details about the acquisition and pre-processing chain can be found in [**?**] and [**?**]. They correspond to the left box of Fig. 2. Note that the pre-processing of SMOS data in the IFS is performed in several processors in parallel, being the first type of satellite data using this configuration in the IFS.

4 Observational Data Base and collocation software

In order for the IFS to be able to handle SMOS data in an efficient way (and for any other source of data), the pre-screened dataset is mapped into an Observational Data Base (ODB). This database contains all the information about each observational register and about the model equivalent. The IFS is then ready to perform other tasks in model space. The first of all of them is collocating the observations to the model grid at the required model resolution, using the nearest neighbour technique. This implementing approach resembles that one used for all-sky radiances for AMSR-E and SSMI data [**?**] (currently, the all-sky radiances follow a different implementing approach in the IFS [**?**]). However, a major difference is that the number of SMOS observations found per timeslot in model space is notably larger than for that of other microwave sensors. This has been a major obstacle in the operational implementation and therefore a substantial re-structuration of the SMOS managing routines was undertaken in order to efficiently collocate and process SMOS observations in model space. Details about these tasks were given in section 5 of [**?**] and in Part II of [**?**]. They correspond to the right box of Fig. 2.

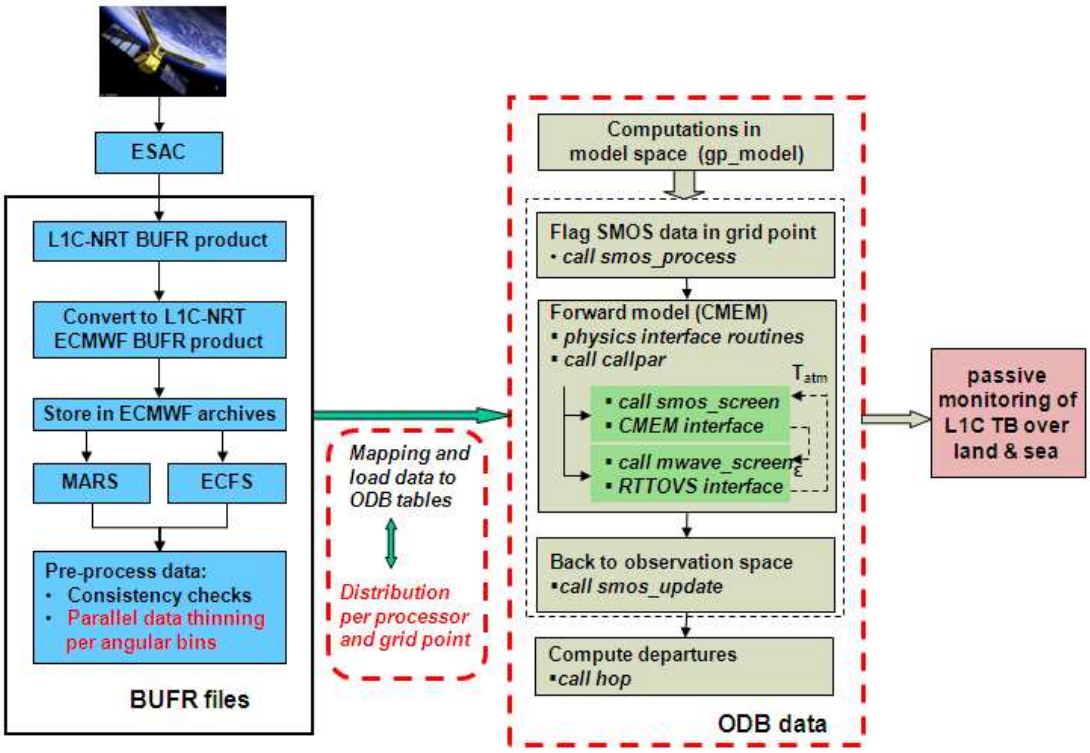

Figure 2: Organigram of the SMOS offline monitoring chain developed at ECMWF.

5 SMOS offline monitoring suite

The previous technical developments summarized in sections 2, 3 and 4 made it possible to achieve one of the main milestones with SMOS data at ECMWF: the monitoring of SMOS data in NRT, both for land and oceans, at different incidence angles (multiples of 10) and two polarisation states (XX, YY). This, as well as it is done for any other source of satellite data used at ECMWF, consists on computing a series of temporal and spatial statistics with the most recent data, providing information about the observations, the model equivalent of the observations and the first-guess departures. These statistics are updated daily or weekly depending on the statistical product used and they are a robust method to identify systematic differences between model values and observations, or deficiencies in the model or the observations.

In [**?**] and [**?**] the technical details of the implementation and a description of the statistical products are given. In the former, some examples of results obtained with the monitoring suite are shown. Time series of area averages are currently being also produced in NRT for several validation sites distributed, mainly, between Europe and United States. In the first monitoring report [**?**] the location of these sites, and a few examples with results, are provided. There is also a preliminary assessment of the main reasons for large first-guess departures. In [**?**] a more exhaustive report covering the first year of Near-Real Time monitoring (Nov-2010/Nov-2011) is provided. Fig.3 shows the angular distribution of the mean bias from August to November 2011 at global scale, for the North and South hemisphere, and for XX and YY polarisations. For the XX polarisation, the bias show quite different behaviour depending on the hemisphere, reflecting the important weight of the snow covered areas in the computation of bias. At this polarisation, the trend is increasing bias (in absolute value) with increasing the incidence angle, being in most of the cases maximum at 60 degrees. The behaviour of the YY polarisation is different and bias are maximum around 30 degrees incidence angle. At this polarisation the snow influence in the soil emission is lower than for the XX polarisation. The colour bar in these figures represents the number of observations for each level of bias. It is observed that there is also a significant amount

of bias with high values, which is mainly due to the RFI effect, but this is not the only reason as explained in [**?**].

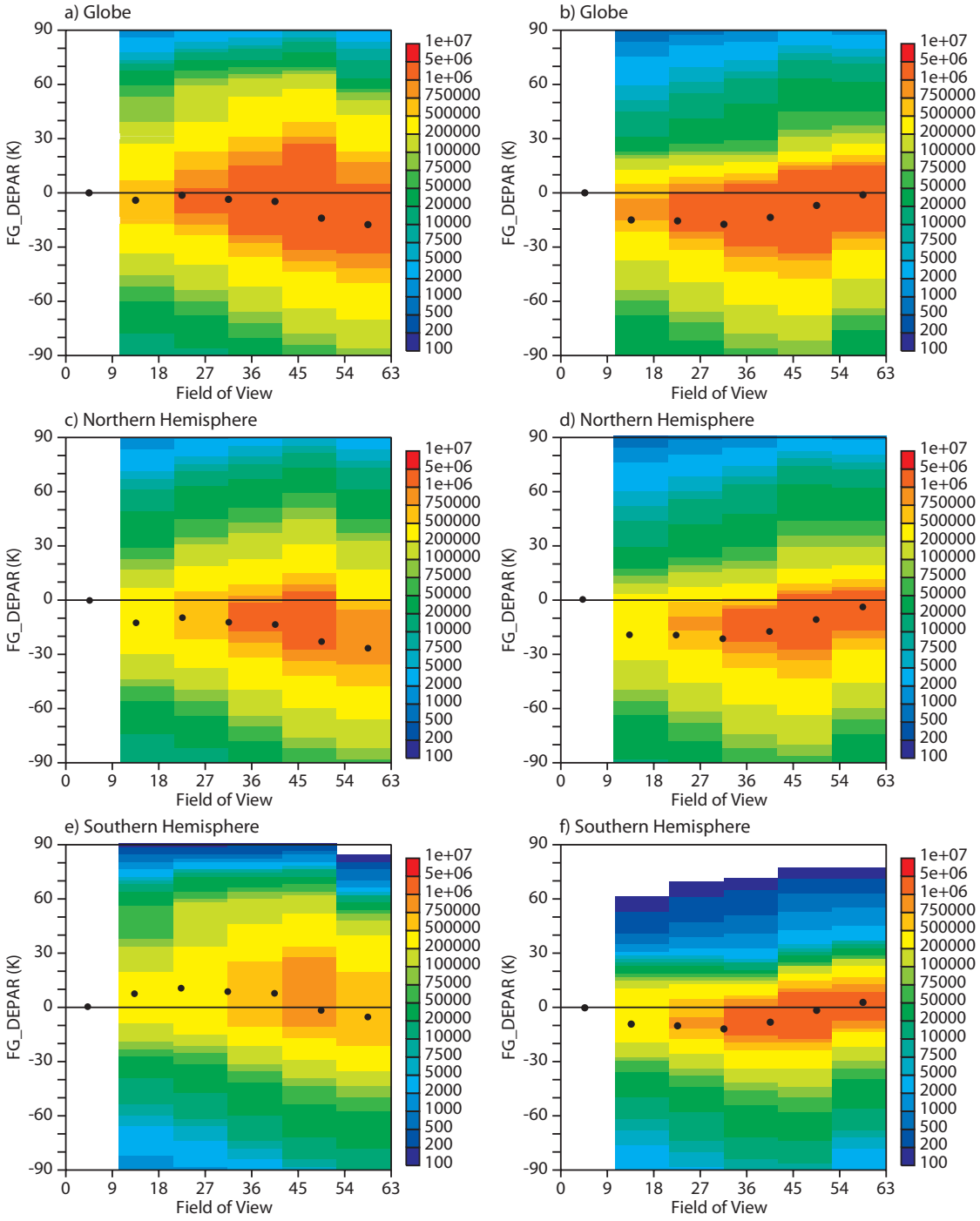

Figure 3: Mean bias as a function of the incidence angle for XX polarisation (left column) and YY polarisation (right column). The period considered spans from 1 August to 31 October 2011. Only continental surfaces are considered in these figures.

6 Implementation of SMOS data in the SEKF

Although SMOS brightness temperatures have been monitored in Near Real Time thanks to the technical benchmarks reached and described in [**?**], [**?**] and [**?**], the use of these data within the Simplified Extended Kalman Filter (SEKF) was also a subject of challenging technical developments. The main objective was to develop the structure necessary to accommodate SMOS data in the ECMWF version of the SEKF, and make it compatible with the monitoring suite and other data used for the soil moisture analysis (remote sensed data and screen level variables).

6.1 Technical developments

Integrating SMOS data in the SEKF has been a very challenging technical task, partly because the assimilation of SMOS data needs information provided by some of the tasks performed by the ODB and the monitoring suite. This involves interacting and make compatible two spaces which are nearly independent (atmospheric 4DVAR space where the atmospheric variables are analysed, and SEKF space where soil moisture is analysed). The 4DVAR structure is governed by high-level routines and it runs per model time step. In this space most of the observations are distributed between processors, stored and later retrieved per time slot. Decisions about which (previously thinned and quality controlled) SMOS observations are assimilated, and flagged as active observations, are undertaken in 4DVAR space per model time step. It is at this stage that the ODB for SMOS is filled with information which will be needed in the SEKF. This will permit to match the time stamp of the observations flagged as active with the model time step, and the location of the observation with model grid-point location. Without this information all the SMOS data in the SEKF would be accepted for assimilation.
On the other hand, the SEKF runs independently of the atmospheric analysis and it does not run per model time step. The SEKF equations are solved per model grid point (and per chunks of grid points) using parallel directives. Firstly, the SEKF generates perturbed runs (one per analysed variable) which are generated by forcing small perturbations of the analysed variables. For each perturbed run, a new simulated brightness temperature is computed by CMEM, and only for those observations which were flagged as active in 4DVAR space. The new model equivalents are temporally stored and later retrieved in the SEKF to compute the Jacobians. The Jacobians evaluate the sensitivity of the model to small perturbations of the state vector. In the ECMWF SEKF, the Jacobians are computed in finite differences in order to linearise the observation operator. Finally, the gain and the increments are computed to adjust the predicted value of soil moisture.

Based on the above explanation, the implementation strategy to ensure the compatibility between 4DVAR and SEKF spaces (for the assimilation of SMOS brightness temperatures) implied the following main steps:

1. Within the atmospheric model first integration (the 4DVAR atmospheric analysis is an iterative process involving several model integrations at different model resolutions) creating a new column in the ODB containing information of the model time step at which each observation belongs.
2. In the SEKF, for each perturbed run, re-evaluate which observations are active and for each of them compute a model equivalent with the L-band forward operator.
3. Store the perturbed simulations at the time of the observations, and retrieve this information in the SEKF for the soil moisture analysis.

The main disadvantage of this strategy is that active observations (potentially assimilated in the SEKF) have to be re-computed again for each perturbed run. However, the associated extra cost of re-evaluating active

observations for each run is negligible. Another implementation strategy was tested too, but with the risk of increasing the complexity of the code. Therefore, the above mentioned strategy was chosen for implementation.

Figs. 4 and 5 show in a schematic way, how the 4D-VAR and SEKF spaces are coupled to enable the assimilation of SMOS data. Fig. 4 summarizes the main steps of the SEKF flow for SMOS data. They run per grid point and in a sequential manner:

1. Initialization of soil moisture analysis,
2. Opening of ODB to retrieve information of the SMOS observations, model equivalents, time step and location,
3. Perturbation runs and storage of simulated perturbed variables,
4. Collocation of observations location with model grid and model time step,
5. Filling of observation vector,
6. Filling of first-guess vector,
7. Retrieval of sensitivity to perturbations,
8. Jacobian computation,
9. Gain computation,
10. Computation of increments at analysis time

In Fig. 5 the corresponding routines for each of the previous tasks are detailed. It is observed that the coupling between both spaces is produced at the high-level managing routines for the SEKF. The organigram of Fig. 6 shows a whole picture of the implementation of SMOS data for both, monitoring and assimilation purposes. For each experiment, all the different tasks necessary to run the monitoring or assimilation of SMOS data are controlled by the Supervisor Monitor Scheduler (SMS). In Fig. 7 a snapshot of a SMS including all the tasks processing SMOS data is shown. It is observed that the preparation of the data is done in the *obs* family, the computation of the first-guess in the first *ifstraj* task, and the assimilation of SMOS data in the *sekf*.

6.2 Main technical challenges in the implementation

The implementation of SMOS data as being part of the soil moisture analysis was a technical challenge. All the technical developments were started in CY36R4, and later on all the routines were transferred to the newer cycles CY37s. This is important in order to benefit from the latest model developments and for being up to date with the scripting tasks. However, in CY37s a systematic failure of all experiments handling SMOS data took place, very difficult to trace back. The ”SMOS failure” in cycle 37R1 onwards was due to a cleaning aiming to remove duplicated code. The removed code did direct memory copy of observation data to model space arrays in the case that the data was already on the right processor. Instead all observation data was message passed to model space. Most observation types in IFS are randomly distributed over processors, but the SMOS observations are co-located with the model grid. So although the cleaning was fine for other data types, for SMOS it introduced an unbalanced memory requirement for message passing to the point of crashing the tasks with the most SMOS observations. Testing to understand and solve this problem was long and slow and it required a significant amount of time.

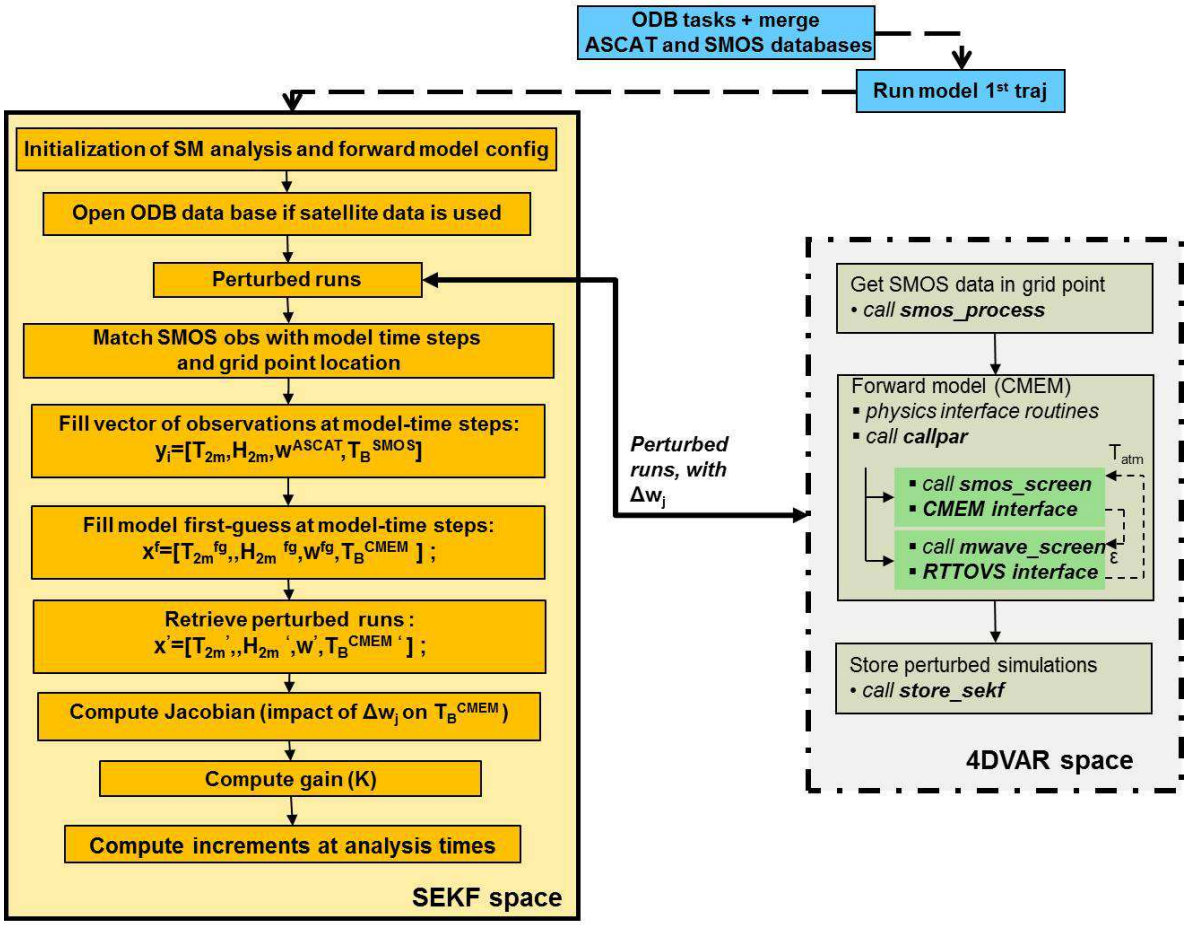

Figure 4: Organigram of the sequential steps of the SEKF (left box). The perturbed runs for the Jacobian computation, requires of the interaccion with the 4DVAR space.

The integration of perturbed runs for the Jacobian computation requires to run several times the L-band forward operator. As mentioned previously, the information of the model time step at which the observation is assimilated needs to be stored, however this information was lost when running the forward model from the surface analysis. Some developments were then necessary in order to extend the ODB and include this information, so it could be retrieved within the surface analysis.

It is often the case to come across some bugs in the code with new implementations. For the SMOS managing routines two complex bugs were found. In particular, it was found that the time step of some SMOS observations was not exactly matching that of the model time step. This resulted in pairs [observed brightness temperature-model equivalent] at different times, often comparing an observation with a missing value. This problem was solved by using the extended ODB for SMOS. Concerning the second bug, the observations corresponding to the first model time step were missing. This was due to a complex problem of the ODB software and an incorrect array allocation.

The implementation of SMOS data in the SEKF also provided an opportunity to revise old parts of the code. Two changes are proposed:

1. The current operational SEKF for soil moisture analysis is only triggered if the water content of soil is greater than $0.01 m^3 m^{-3}$. The reason for that was to prevent the system to produce a negative value of

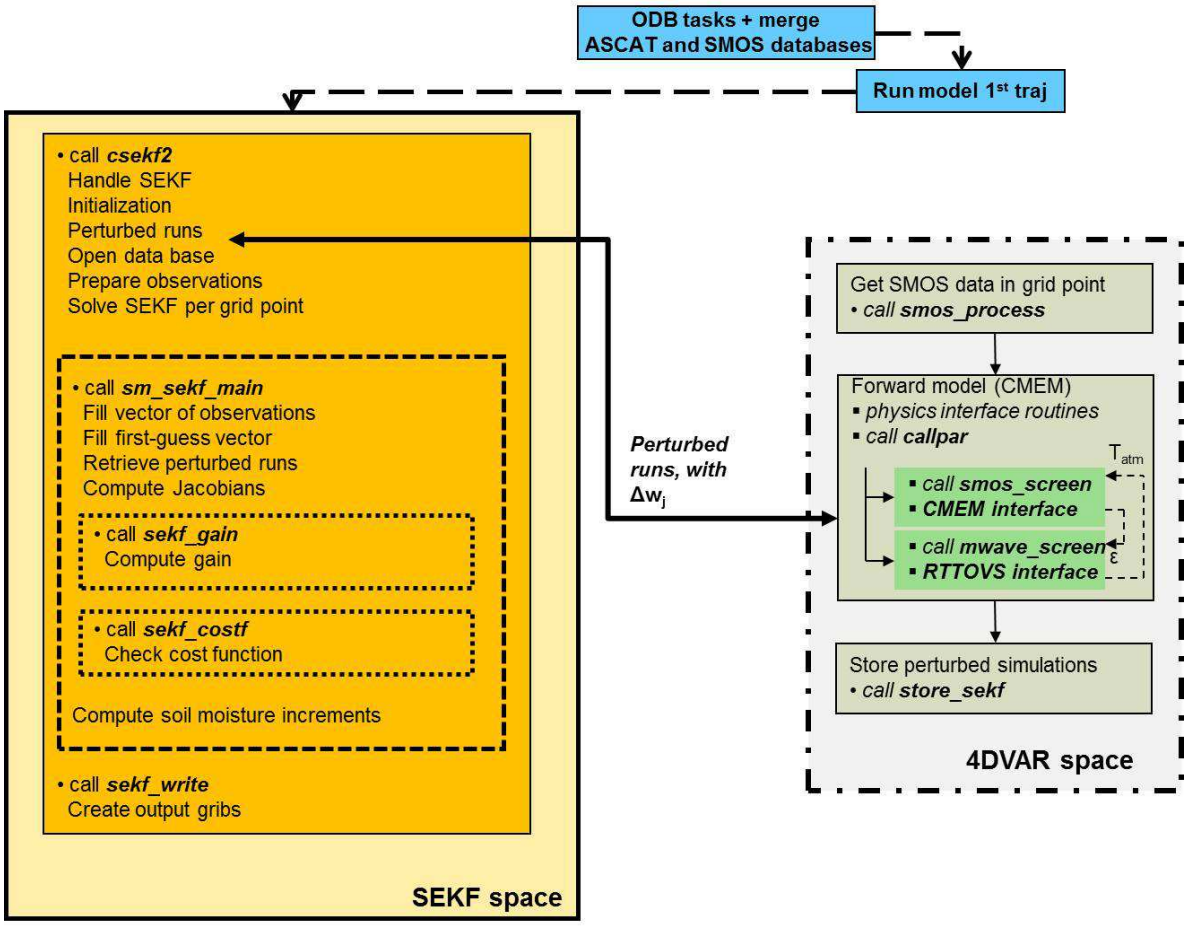

Figure 5: As Fig. 4, but this snapshot details the routines used in the IFS for the SEKF analysis.

soil moisture if a negative perturbation in the Jacobians computation was applied, in the case a chess-like perturbation method was chosen. This could happen if the size of the perturbation is larger than $0.01m^3m^{-3}$ when the soil water content is lower than $0.01m^3m^{-3}$. Hence, the analysis was prevented for being active in areas where SMOS data can potentially bring a large benefit. Currently, only positive perturbations are applied and therefore to allow analysis of soil moisture for very dry soil conditions, it is proposed to substitute this test by a land-sea mask condition; if the land surface mask indicates more than 50% of the pixel to be land, then the soil moisture analysis is triggered.

2. The size of the perturbation should be strictly the same for all grid-points and equal to the value specified in a namelist. This was not strictly true for the current implementation, and some small differences were observed. In order to get the right size of the perturbation, the unperturbed and perturbed forecasted soil moisture for the first model time step is retrieved from the SEKF and the difference associated to the size of the perturbation.

6.3 SMOS data in the SEKF

In the current operational system, the state vector of the SEKF is composed by the soil moisture of the three first layers of the operational land surface scheme H-TESSEL [**?**]. H-TESSEL predicts a value (or first-guess) of soil moisture for each grid point and layer, as well as other land surface processes such as the evolution of the surface temperature or the snow cover extension. Screen level variables (2-metre temperature and relative

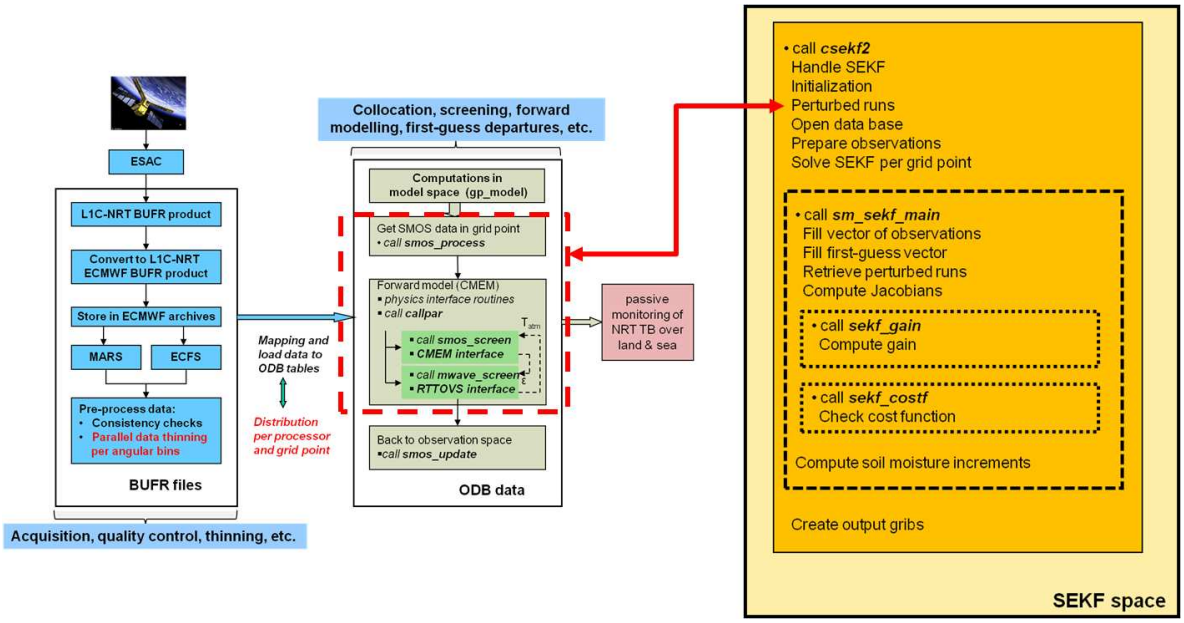

Figure 6: Scheme showing the acquisition and pre-processing (left box), monitoring (middle box) and assimilation (right box) of SMOS data in the IFS, and their relationship.

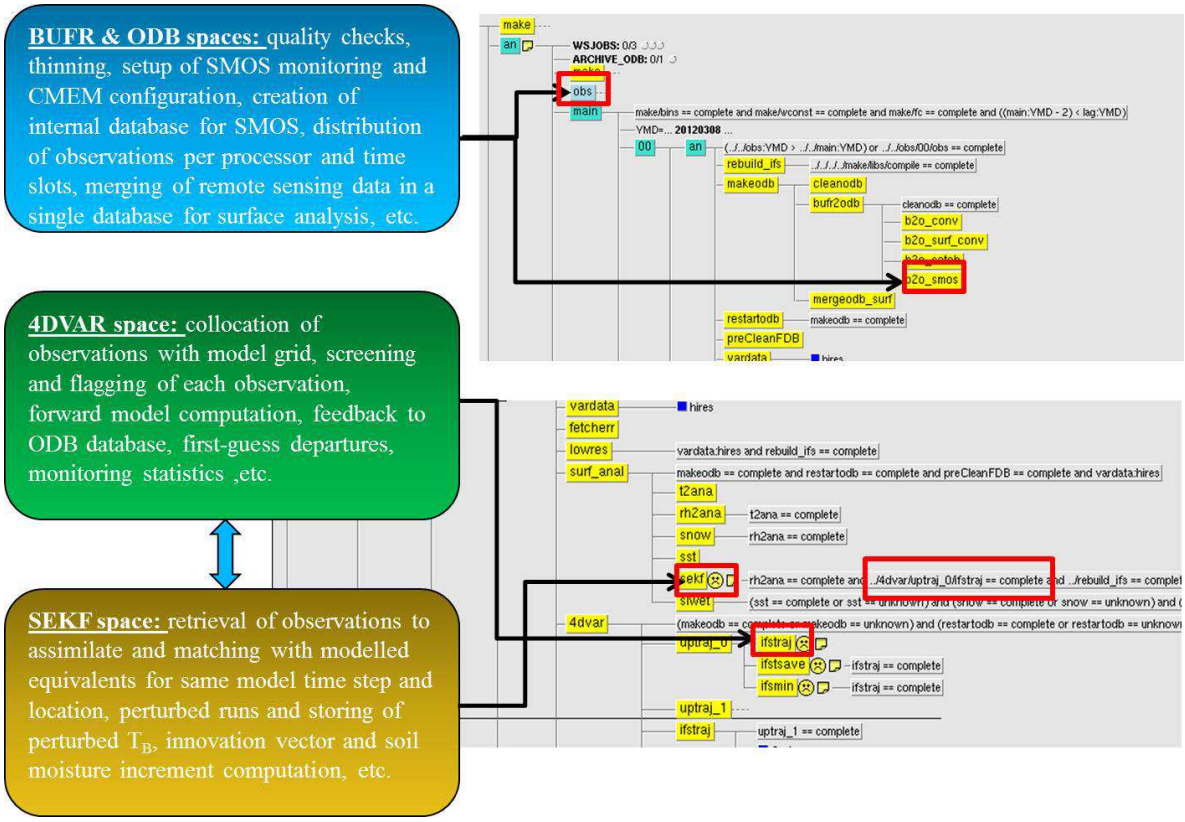

Figure 7: Schematic presentation of the different Supervisor Monitor Scheduler jobs where SMOS data is processed. Note that most of the pre-processing tasks take place within the "obs" family, the computation of the first-guess departure within the ifstraj and the assimilation of SMOS data in the sekf task of the surf_anal family.

humidity) are used operationally to adjust the soil moisture value predicted by H-TESSEL using the equations of the SEKF. Optionally, active data in C-band from the ASCAT (Advanced SCATterometer) sensor on MetOp can also be used in research mode to analyse soil moisture. The implementation of SMOS data in the SEKF adds the potential of using passive remote sensing data, more sensitive to the variations of soil moisture, to provide a more accurate value of the soil moisture analysis.

Soil moisture is updated twice per 12h cycle, at 0000 and 0600 UTC for the 2100 to 0900 UTC cycle, and at 1200 and 1800 UTC for the 0900 to 2100 cycle [**?** **?**]. At each *i* analysis time (0000, 0600, 1200 and 1800 UTC) the soil moisture of the first 3 layers of the HTESSEL land surface model is updated for each grid point according to:

$$\mathbf{x}^a_{i,j} = \mathbf{x}^b_{i,j} + \mathbf{K}[\mathbf{y}^{OBS}_i + H_i(\mathbf{x}^b)] \tag{1}$$

with j the soil layer and the a, b and OBS superscripts standing for analysis, background and observations, respectively.

The vector of observations **y** depends on the observations used to analyse soil moisture. If screen level variables and SMOS brightness temperatures (T_B) are used, then for the 2100-0900 UTC assimilation window:

$$\mathbf{y}^{OBS} = [T^{2m}_{00}, T^{2m}_{06}, RH^{2m}_{00}, RH^{2m}_{06}, T_B(\theta_1, XX), \ldots, T_B(\theta_n, XX), T_B(\theta_n, YY)] \tag{2}$$

with θ the incidence angle of the SMOS observation and n the number of SMOS observations assimilated for this grid point during the 12h assimilation window. Note that for this cycle, the screen level variables are only available at synoptic times (0000 and 0600 UTC) and for SMOS we only use the pure polarisation modes XX and YY.

The Kalman gain matrix **K** depends on the linearised observation operator $\mathbf{H}_i$, and on the background (**B**) and observations (**R**) variance-covariance error matrices according to:

$$\mathbf{K} = [\mathbf{B}^{-1} + \mathbf{H}_i^T \mathbf{R}^{-1} \mathbf{H}_i]^{-1} \mathbf{H}_i^T \mathbf{R}^{-1} \tag{3}$$

Note that the variance-covariance error matrices do not depend on the time of the analysis or the observations, but on the contrary a fixed error value for each observation is assumed during the whole assimilation window. With SMOS data in the SEKF, the observation error matrix **R** for the 2100-0900 UTC assimilation window takes, for each grid point, the following form:

$$\mathbf{R} = \begin{bmatrix}
\sigma^2_{T^{2m}_{00}} & & \cdots & \sigma_{RH^{2m}_{06}}\sigma_{T^{2m}_{00}} & \cdots & & \cdots & & \sigma_{T_B(\theta_n,YY)}\sigma_{T^{2m}_{00}} \\
& \sigma^2_{T^{2m}_{06}} & & & & & & & \vdots \\
\vdots & & \sigma^2_{RH^{2m}_{00}} & \vdots & & & & & \\
\sigma_{T^{2m}_{00}}\sigma_{RH^{2m}_{06}} & & \cdots & \sigma^2_{RH^{2m}_{06}} & & & & & \vdots \\
\vdots & & & & \sigma^2_{T_B(\theta_1,XX)} & & \cdots & & \sigma_{T_B(\theta_n,YY)}\sigma_{T_B(\theta_1,XX)} \\
& & & & & \sigma^2_{T_B(\theta_1,YY)} & & & \\
\vdots & & & & \vdots & & \ddots & & \vdots \\
& & & & & & & \sigma^2_{T_B(\theta_n,XX)} & \\
\sigma_{T^{2m}_{00}}\sigma_{T_B(\theta_n,YY)} & & \cdots & & \sigma_{T_B(\theta_1,XX)}\sigma_{T_B(\theta_n,YY)} & & \cdots & & \sigma^2_{T_B(\theta_n,YY)}
\end{bmatrix} \tag{4}$$

In the operational soil moisture analysis this matrix is a diagonal matrix, with the square root of the diagonal terms the standard deviation of the screen level variables ($\sigma_{T^{2m}} = 2$K, $\sigma_{RH^{2m}} = 10\%$). The background variance-covariance error matrix **B** is also static and diagonal, and accounts for the background error associated to the state vector.

The non-linear observation operator H_i maps the model state vector $\mathbf{x}$ into the observation space. In an Extended Kalman Filter, the observation operator H_i is linearised, because then the solution of the analysed state at time i is analitically obtained by 1 [**?**]. A possible way to evaluate the tangent linear observation operator matrix is in finite differences, by forcing small perturbations of each component of the state vector ($\delta\mathbf{x}_j$) and evaluating the impact of the individual perturbations on each observed variable. If, as in 2 and 4, we suppose only screen level variables and SMOS observations for the 2100-0900 UTC assimilation window, the linearised observation operator is computed as:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial T_{00}^{2m}}{\delta(sm_1)} & \frac{\partial T_{00}^{2m}}{\delta(sm_2)} & \frac{\partial T_{00}^{2m}}{\delta(sm_3)} \\ \frac{\partial T_{06}^{2m}}{\delta(sm_1)} & \frac{\partial T_{06}^{2m}}{\delta(sm_2)} & \frac{\partial T_{06}^{2m}}{\delta(sm_3)} \\ \frac{\partial RH_{00}^{2m}}{\delta(sm_1)} & \frac{\partial RH_{00}^{2m}}{\delta(sm_2)} & \frac{\partial RH_{00}^{2m}}{\delta(sm_3)} \\ \frac{\partial RH_{06}^{2m}}{\delta(sm_1)} & \frac{\partial RH_{06}^{2m}}{\delta(sm_2)} & \frac{\partial RH_{06}^{2m}}{\delta(sm_3)} \\ \frac{\partial T_B(\theta_1,XX)}{\delta(sm_1)} & \frac{\partial T_B(\theta_1,XX)}{\delta(sm_2)} & \frac{\partial T_B(\theta_1,XX)}{\delta(sm_3)} \\ \frac{\partial T_B(\theta_1,YY)}{\delta(sm_1)} & \frac{\partial T_B(\theta_1,YY)}{\delta(sm_2)} & \frac{\partial T_B(\theta_1,YY)}{\delta(sm_3)} \\ \vdots & \vdots & \vdots \\ \frac{\partial T_B(\theta_n,XX)}{\delta(sm_1)} & \frac{\partial T_B(\theta_n,XX)}{\delta(sm_2)} & \frac{\partial T_B(\theta_n,XX)}{\delta(sm_3)} \\ \frac{\partial T_B(\theta_n,YY)}{\delta(sm_1)} & \frac{\partial T_B(\theta_n,YY)}{\delta(sm_2)} & \frac{\partial T_B(\theta_n,YY)}{\delta(sm_3)} \end{bmatrix} \tag{5}$$

where $\partial\mathbf{y}_t$ is $H(\mathbf{x}^b + \delta\mathbf{x}^b) - H(\mathbf{x}^b)$. After the analysis step the state vector evolves from time i to time $i+1$ according to:

$$\mathbf{x}_{i+1}^b = M_i[\mathbf{x}_i^a] \tag{6}$$

with M the non-linear forecast model.

6.4 Experiments

The implementation of SMOS data in the SEKF was technically tested by running the following three experiments:

- CTRL: This is the control experiment which assimilates only T^{2m}, RH^{2m} observations.
- SMOS40XX-YY: Bi-polarised SMOS configuration, which assimiles T^{2m}, RH^{2m} and SMOS brightness temperatures at 40 degrees incidence angle and XX and YY polarisations (40XX, 40YY),

- SMOS20-50XX: Bi-angular SMOS configuration, which assimilates T^{2m}, RH^{2m} and SMOS brightness temperatures at 20 degrees incidence angle and XX polarisation (20XX, 50XX).

The main goal of this exercise is testing that the technical implementation is working fine. The scientific impact of these experiments, the bias correction and the CMEM global calibration will be addressed in future reports corresponding to the next phase of the study.

6.4.1 Experimental setup

The observations were only assimilated over Australia (free from Radio Frequency Interference), from the 01 April 2011 00UTC to 30 April 2011 12UTC analysis cycles. This is a period of hydrological recharge for Australia, when some substantial variations in the soil water content occurs. SMOS brightness temperatures were assimilated to an approximate horizontal model spatial resolution of 40 km, being very close to the spatial resolution of SMOS observations, and avoiding in this way large undesired horizontal correlations between observations if a higher model resolution was used. The assimilated product is the standard NRT. The configuration of the model forward operator was that of the monitoring suite, based on the research works of [**?**], [**?**] and [**?**]. However, more recently the CMEM forward operator has been calibrated at global scale, and the results will be reported in [**?**] and [**?**]. With the current configuration, low sensitivity is found for some areas, mainly caused by a lack of sensitivity of the Choudhury roughness model [**?**] used in CMEM. The Jacobians are computed by increasing in $0.01 m^3 m^{-3}$ the soil water in the three first layers of the surface model and computing each term of the Jacobians in 5. In order to avoid too large, unrealistic sensitivity of the Jacobians terms, the maximum sensitivity allowed for the T^{2m} and RH^{2m} observations is set to $50 K/m^3 m^{-3}$ and $500\%/m^3 m^{-3}$, respectively, and for SMOS brightness temperatures to 250 $K/m^3 m^{-3}$. The background error matrix **B** is not cycled and is for each grid point a diagonal matrix. The standard deviation of the three layers of soil moisture is set to $0.01 m^3 m^{-3}$. For the **R** matrix the error standard deviation of the T^{2m}, RH^{2m} observations is fixed and equal to 2 K and 10 %, respectively, whereas for the brightness temperatures observations the radiometric accuracy of each individual observation is used.
These type of experiments are quite expensive to run as the land surface model is coupled to the atmospheric analysis. To reduce the computational expense, a degraded observational system was chosen for the atmospheric analysis, and consisting only of conventional observations which are able to constrain the meteorological state of the atmosphere in a reasonable accurate position. Moreover, by using a degraded observational system, the impact of assimilating SMOS observations on the forecast skill can be more clearly evaluated than if the whole ECMWF observational system is used.

6.4.2 Quality control

Each observation goes through a standard quality control process, as explained in [**?**] and [**?**], consisting on routine checks and a Radio Frequency Interference (RFI) hard filter. In order to prevent too much unnecesary data in the surface analysis, data are also thinned in the early processing stages at the resolution of the analysis. Only the closest observations to the model grid points are ingested in the IFS. At T511 this makes SMOS a much lighter dataset (see section 2.5 of [**?**]). Snow and frozen masks are applied to each observation, based on snow depths and T^{2m} forecasted fields. The top panel of Fig. 8 shows the geographical distribution of the number of observations which enter the soil moisture analysis for experiments SMOS40XX-YY (left) and SMOS20-50XX (right), after the quality control and thinning steps. The bottom panel corresponds to the histogram of number of SMOS observations in the SEKF. The histograms show that a larger number of observations are assimilated for the 40XX, 40YY configuration than for the 20XX, 50XX. For SMOS40XX-YY, the number of grid-points assimilating between 45-50 observations are maximum, in contrast to the 25-30 bin for experiment SMOS20-50XX. It is also observed, than at least 19 observations are assimilated for a grid-point in SMOS40XX-YY and

a minimum of 9 for SMOS20-50XX, since the lattest observational configuration uses observations at 20 and 50 degrees incidence angle, less numerous than 40.
Several quality checks are also applied within the SEKF. Among them the most important is the first-guess departure check, in other words, the maximum disaggrement allowed between SMOS observations and the model equivalent. In these experiments the first-guess limit is arbitrarely setup to 16K. Fig. 9 shows the number of observations rejected by the first-guess check. They are maximum in regions where the observations showed stronger variability. This is caused by a combination of a low sensitivity of the CMEM configuration used in this study for this region and period, and an approximate bias correction which does not account for the spatial variability of the observations (see 6.4.3). Further quality checks prevent too large sensitivity of the Jacobians or too large analysis increments. Almost no observations were rejected by too sensitive Jacobians.

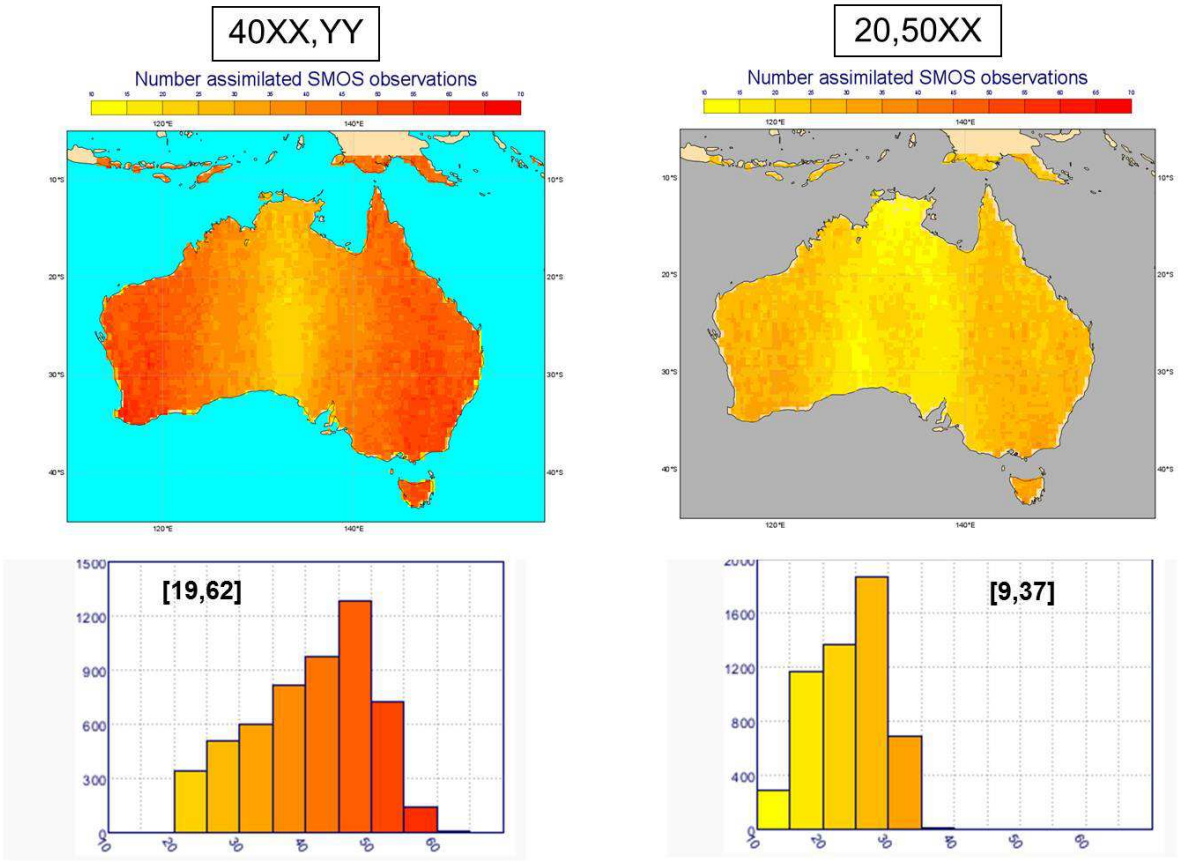

Figure 8: Geographical distribution of the number of SMOS observations in the SEKF for the soil moisture analysis (top), and the histograms (bottom). Left column corresponds to the 40XX, 40YY configuration, right column for the 20XX, 50 XX.

6.4.3 Bias correction

An approximate bias correction method was applied to each observation. It is based on removing to each observation the mean bias computed for the whole of Australia and month of April. This approach makes the hypothesis that each grid point is affected by the same mean bias during the whole period of investigation. In average the mean bias behave quite stable for April, however significant regional differences were found. In Table 2 the mean bias obtained for both experiments using SMOS data is shown. In the four cases, the model overestimates the observations. Figs. 10 and 11 show the bias for each polarisation and incidence angle before and after bias correction, respectively. In general, bias are reduced and are closer to zero for grid points with a

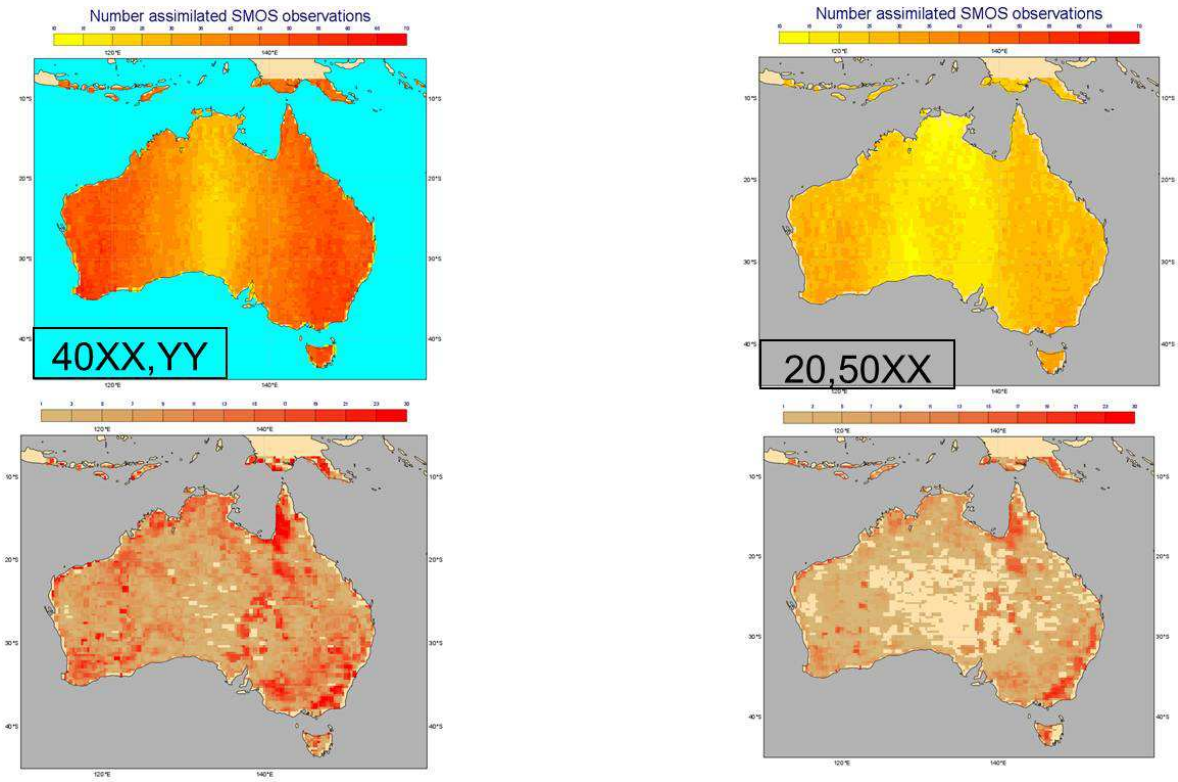

Figure 9: Geographical distribution of the number of observations rejected by the first-guess check for the 40XX, 40YY SMOS observational configuration (left) and for the 20XX, 50XX (right).

mean bias close to the mean bias of Table 2. However, since this is not a scheme based on a grid to grid point analysis, areas with initial large bias still are affected with strong bias after applying this approximate method.

Table 2: Mean bias for Australia in April-2011, using the CMEM configuration of the monitoring suite.

Mean Bias	**20XX**	**40XX**	**40YY**	**50XX**
SMOS40XX-YY	-	-19K	-21K	-
SMOS20-50XX	-15K	-	-	-32K

6.4.4 Analysis increments

Fig. 12 shows the soil moisture increments for the three first layers and for the three experiments. These increments are integrated values over the depth of each layer. Left column corresponds to the control experiment (without SMOS data), middle column for the bipolarised configuration (40XX, 40YY) and right column for the biangular configuration (20XX, 50XX). The main differences are found for the top layer, where the remote sensing data is sensitive, although the general patterns are very similar. Without SMOS data, soil moisture is adjusted only by the screen level variables, which contains an indirect information of soil moisture. In this case most of the time the analysis adds water. This picture changes by using SMOS information, more directly linked to soil moisture. In this case, some drying is produced in large areas of Australia which were not produced in the control run. It is also interesting that over the North of Australia (which was very wet at this time of the year) the analysis adds some water by using SMOS information, in contrast to the control experiment. As we go deeper into the soil, the link with remote sensing data is weaker, which makes relatives increments on the second and third layer smaller. The increments for the third layer are very similar for the three experiments.

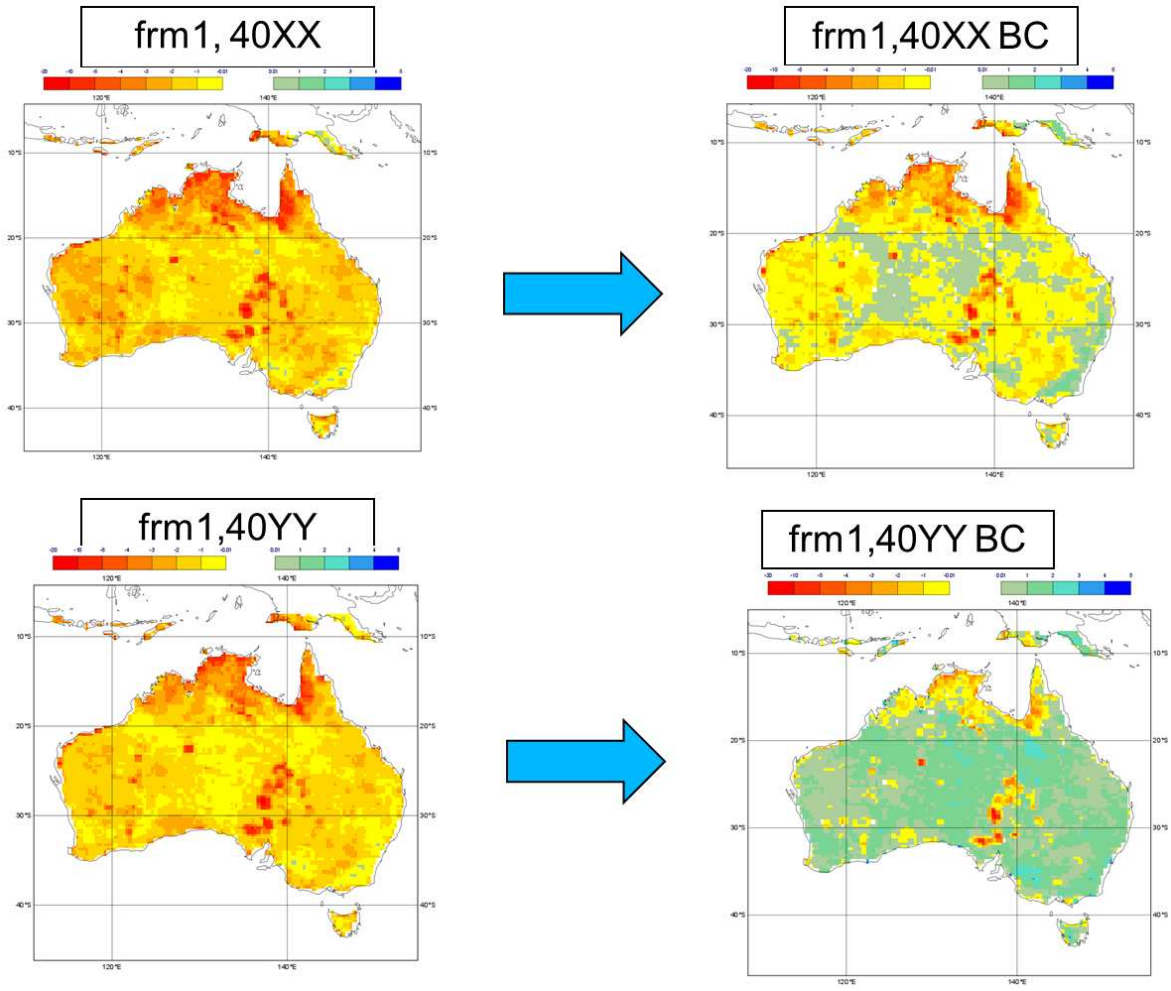

Figure 10: Mean bias over the whole of April before bias correction (left column) and after bias correction (right column).

Although these results show an influence of SMOS data in the soil moisture analysis, they have to be taken with caution, as these experiment have as objective only to demonstrate the technical approach developed. Many of the variables and parameters influencing the SEKF are not calibrated in these experiments. In the next contract 4000101703/10/NL/FF/fk, the impact of assimilating SMOS data in the soil moisture analysis (including a calibrated system, forward model and observations bias corrected) will be addressed.

7 Summary

This document reports on the technical milestones reached and implemented in the IFS, which have permitted to attain two of the main objectives of this study: the monitoring of SMOS brightness temperatures in NRT, and the incorporation of SMOS data as a part of the ECMWF soil moisture analysis. The offline monitoring suite monitors in a statistical way the spatial and temporal behaviour of SMOS observations, the model equivalent and the first-guess departures. Each day hundreds of new plots, incorporating the last observations, are made freely available to the entire scientific community. On the other hand, the ECMWF SEKF has now new features, which permits the assimilation of only screen level variables, or only active and passive remote sensing

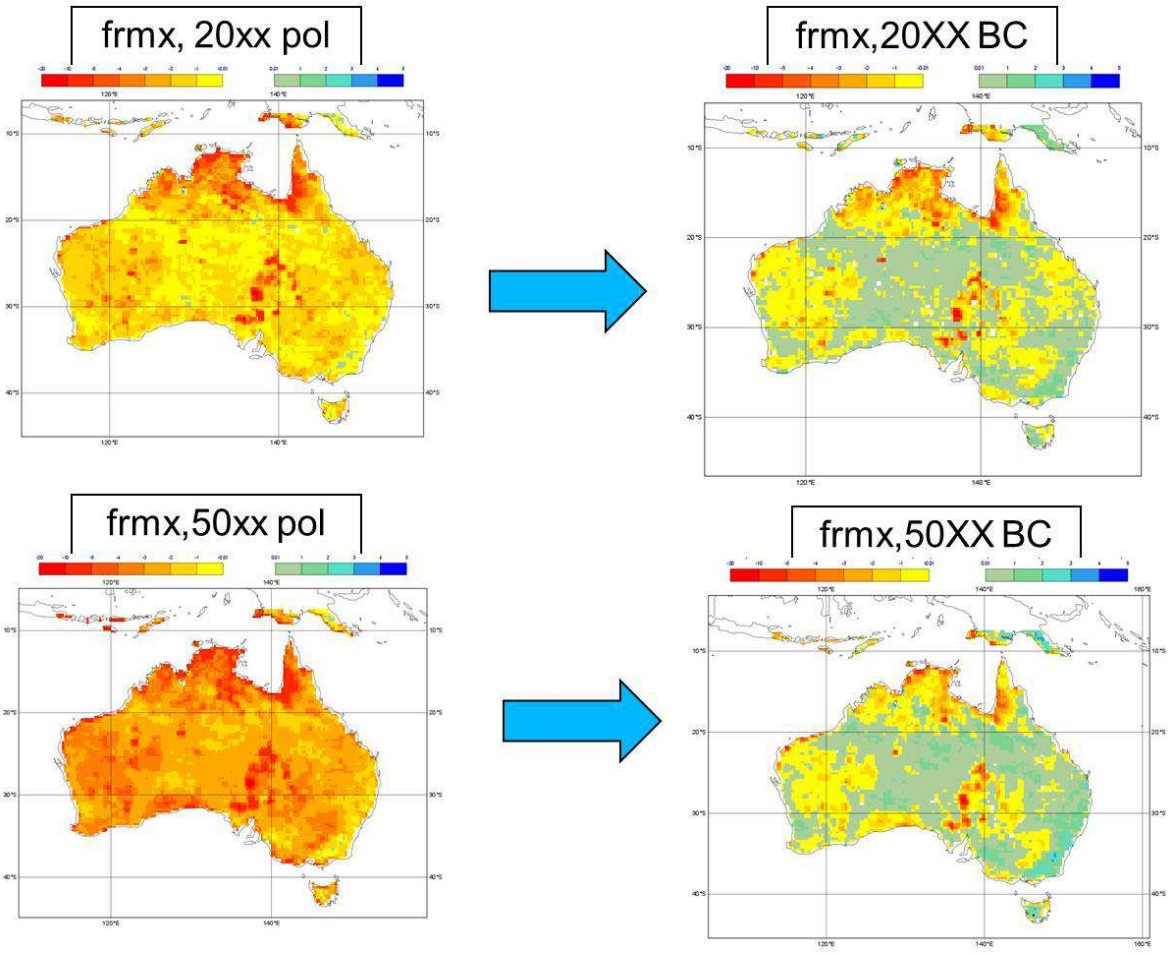

Figure 11: As Fig. 10 but for the SMOS biangular configuration.

data, or any combination of them, for the analysis of soil moisture. To this end a new database was created, opening the door to accommodate future satellite data in the SEKF, such as data from the future Soil Moisture Active Passive mission or the projected future continuation of the SMOS mission.

The experiments and results presented in this document are not intended to provide robust scientific results, on the contrary they are mainly intended to test the technical developments described along this document and to demonstrate that the system is ready to assimilate SMOS data. Other essential activities forming part of a well calibrated data assimilation system, such as a point-by-point bias correction system, or a global calibration of the observation operator, along with experiments intended to study the influence of assimilating SMOS data in the surface and atmospheric fields, will be addressed in the next contract of this study.

Acknowledgements

This work is funded under the ESA-ESRIN contract number 20244/07/I-LG. Acknowledgements to Matthias Drusch and Susanne Mecklenburg, both ESA staff, for their contribution in the project definition. Also thanks to many ECMWF staff which directly or indirectly have contributed to this work.

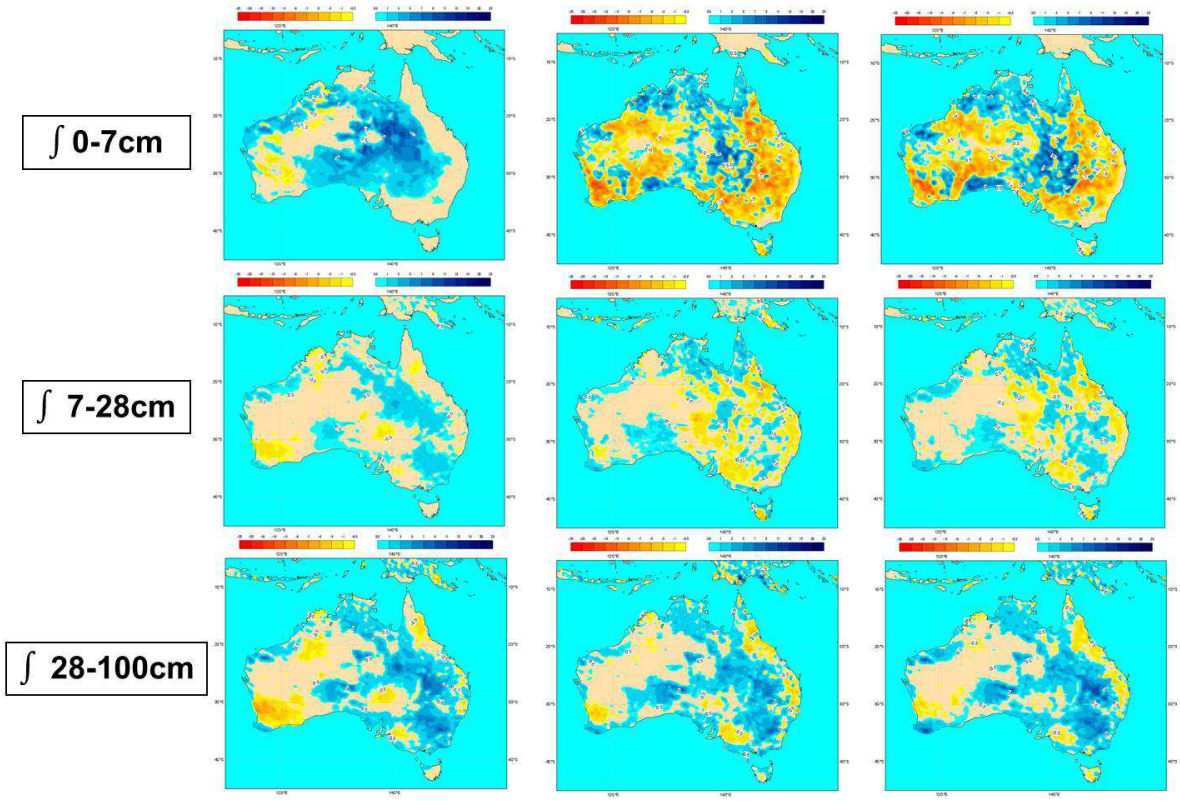

Figure 12: Analysis increments integrated over the first soil layers; 0-7 cm (top row), 7-28 cm (middle row) and 28-100 cm (bottom row). Left column is for the control experiment, middle column for the bipolarised SMOS configuration and right column for the biangular configuration.