SMOS retrieval algorithm: trials and tribulations

Yann H. Kerr, Philippe Waldteufel, Philippe Richaume, Arnaud Mialon, François Cabot, Maria-José Escorihuela, Jean-Pierre Wigneron, Paolo Ferrazzoli, Steven Delward, Ali Mahmoodi

SMOS retrieval Algorithm: Trials and tribulations

Yann H. Kerr[1], Philippe Waldteufel[2], Philippe Richaume[1], Arnaud Mialon[1], François Cabot[1], Maria José Escorihuela[1], Jean Pierre Wigneron[3], Paolo Ferrazzoli[4], Steven Delward[5] and Ali Mahmoodi[6]

1. CESBIO, 18 avenue Edouard Belin 31401 Toulouse CEDEX France. 2. IPSL /SA Verrieres le Buisson France ; 3. INRA EPHYSE Bordeaux France;4. Tor Vergata University Roma Italy; 5. ESA ESTEC Noordwijk The Netherlands;6. Arrays Systems Toronto Canada

Abstract This contribution summarizes prominent features of the Level 2 algorithm aimed at processing land surface geophysical quantities from the ESA-led SMOS mission. It emphasizes the soil moisture retrieval and describes the decision tree built in order to select appropriate retrieval configurations. The expected performance is illustrated by preliminary results of the algorithm validation.

Introduction

In the ESA-led Soil Moisture and Ocean Salinity (SMOS) mission [Kerr et al 2001]), the principle of soil moisture retrieval consists in minimizing a cost function which sums weighted squared differences between SMOS-measured and computed brightness temperatures, using for the latter a direct radiative model (see L-MEB description by Saleh et al in this conference [Saleh et al 2007]). The cost function also includes difference terms for the retrieved quantities, with weights which reflect a priori uncertainties on these quantities.

This approach is innovative for remote sensing of land surfaces. It must be refined to account both for heterogeneities which are bound to affect soil et/or vegetation properties over the SMOS pixel, and for the auxiliary parameters which enter the radiative model.

The L-MEB model applies to uniform (all properties identical over the whole pixel) vegetated soil. However it is necessary to extend the retrieval zone to mixed (inhomogeneous) pixels because pure vegetated soil SMOS pixels are quite scarce. Hence the algorithm introduces land cover auxiliary information, from which are built default radiative contributions for other covers i.e. radiometric fractions. Additional features such as topography, as well as time dependent soil cover (frost, snow) need to be considered.

Whenever soil moisture (SM) cannot be retrieved, or when it cannot be retrieved with significant accuracy since the vegetated soil fraction is too small, we aim at making however the best use of SMOS data. Hence a decision tree (DT) is introduced in order to select the best approach, depending on the nature of the scene. For cases where the soil dielectric constant is the most relevant surface parameter, the so called modified cardioid model [Waldteufel et al 2004] extracts optimally the information available from radiometric data.

The SM retrieval uncertainty depends both on input uncertainties and observation conditions, e.g. incidence angle coverage (function of swath abscissa), and data quality (evidenced by level 1 flags). Hence the retrieval must be matched to the potential of radiometric data in order to maximize the retrieved information, while complying with the SMOS requirement.

In particular, many parameters have to be provided for by auxiliary information, which is not retrieved in the general case. Some of this information can however be improved through retrieving it from SMOS data themselves under specific, favorable conditions. The corresponding retrieval scenarios are again selected through a "decision tree" procedure.

Then, improved auxiliary information should be made available to the retrieval as either fixed values or initial values and constraints for retrieval. This calls for updated "current" files which will supply values overriding default values, when available from past retrievals. A particularly important case in this respect concerns vegetation optical thickness and the so called dual step retrieval.

In summary the retrieval algorithm includes

- Selecting a fraction where quantities are to be retrieved, as well as the radiative model to be used depending on the scene. Accounting for variations of the radiometric fraction with incidence angle is probably necessary.
- Defining the parameters to be retrieved and the input uncertainties, depending on available radiometric data, the quality of input auxiliary information, the range of input a priori values.
- Updating and maintaining current files for using the output of previous retrievals.

This presentation discusses these issues and presents first results of the ongoing calibration of the level 2 soil moisture retrieval algorithm.

Preprocessing

In the Level 1 (L1) L1a processing step, visibilities provided by the radiometric interferometer on board SMOS are first calibrated. Next (L1b step) they are, for each pre-integration period, reconstructed into angular fields of brightness temperatures TB over the alias free Field-of-View (FOV). While the full polarization vector can be obtained, only the dual polarization mode is described in this contribution. Finally (L1c step), the fields collected over each sequential FOV image or "snapshot" are reorganized in such a way that every TB values corresponding to a given geographic location are grouped together for the whole range of available incidence angles. This L1c data set is the input for the L2 processor.

The TB values, as well as estimated radiometric uncertainties, are available for each node of the fixed SMOS discrete global grid (DGG); the ISEA9 equi-area grid [Sahr et al 2003], with spacing close to 15km, has been selected for this purpose. Depending on the abscissa across the FOV, the number of TB data pairs (SMOS views) may reach up to about 80.

The initial L2 preprocessing includes:

- Filtering out data pairs for which the spatial resolution does not comply with SMOS requirements for SM;
- Accounting for flags raised at L1 level through either eliminating the data or increasing the uncertainty. TB exceeding realistic values may be due to RFI; they are eliminated and their detection contributes to update a RFI map.
- Rotating the TB from antenna level to surface TH & TV components ([Claassen & Fung 74], [Waldteufel & Caudal 2002]); simultaneously, the Faraday rotation is corrected using auxiliary data. The covariance matrix for (TH, TV) pairs is computed as well.

Decision tree: stage 1

Since the SMOS requirement for SM apply to vegetated soil covered by a non-existent to moderately thick vegetation cover over smooth terrain, it is necessary to characterize the land cover over each SMOS pixel.

A SMOS pixel is actually defined by the area surrounding each DGG node where the weighting function (WEF) resulting from the synthetic directional gain pattern, to be applied to every upwelling radiating contribution, takes non negligible values. Although the WEF depends on incidence angle, a mean WEF can be defined: its half-maximum width is close to 40 km over the FOV, and its broadest extent is estimated to be 123 km.

Over square 123 km sized working areas (WA) specific of each SMOS grid node, the land cover has been characterized for land surfaces. To this end, the ECOCLIMAP classification (reference [Masson et al 2003]) has been used, after the ECOCLIMAP classes have been merged into aggregated cover classes (ALC) : vegetated soil (the "nominal" target), forest, fresh or saline open water, wetlands, barren soil, ice, urban areas. Each ALC fraction is first computed with a spatial resolution close to 4km, before weighting by WEF values and summing the elementary fractions.

Since it is estimated that about 6% of SMOS pixels consist of pure vegetated soil, heterogeneous pixels must be considered. The retrieval strategy will consist in considering for nominal SM retrieval working areas where the nominal fraction is larger than 40% (that is about 51% of the nodes) and introducing default radiative contributions for other fractions present in the pixel.

Similarly, the SM will be retrieved using the L-MEB over WA dominated by forest cover, whenever the forest fraction exceeds 60%.

For water surfaces (open water and wetlands), a specific direct model [Klein and Swift, 1977] is used. For other targets as well as strongly heterogeneous pixels, the cardioid model [Waldteufel et al, 2003] provides information about the surface equivalent dielectric constant. The same choice applies to non permanent covers: frost as well as anticipated wet or mixed snow cases.

Therefore the first stage of the decision tree (DT1) selects, according to the nature of the SMOS pixel cover, the fraction where geophysical quantities will be retrieved and the model to be used.

Topography is taken care of through the use of a global index map. The rationale is to classify the surface as per the impact of topography on the emitted brightness temperature. For flat terrain or gently undulating surfaces the impact is negligible and topography is not accounted for. For high topography, the signal is too complex to enable satisfactory retrieval. The medium topography the retrieval is attempted but the user is warned through a flag that the retrieval might be wrong. It is hoped that after some use, models will be developed to correct this last case. To provide this information a global map is produced once for all using a high resolution DEM and contextual analysis and slope distribution factor. This map provides, for each node, a flagging value: 0 if the node can be considered as flat, 2 if it is very rough and 1 for the intermediate cases.

Decision tree: stage 2

Once DT1 has been applied, it remains to decide which parameters are going to be retrieved, as well as their initial values and uncertainties.

In the nominal (vegetated soil) or forest retrieval cases, SM is obviously to be retrieved; similarly, in other cases, a parameter corresponding to the dielectric constant or nadir reflectivity is selected. However it is possible in addition to retrieve other parameters used for modeling e.g. small scale roughness or properties of the vegetation layer; this will however increase the resulting retrieval uncertainty over SM, by an amount which depends on the initial constraints set upon additional parameters.

Therefore the exact retrieval configuration should depend both on the reliability of auxiliary parameters estimated from external data, and on the obligation to comply with the SMOS SM requirement. In this latter respect, what matters most is the information content of SMOS data, which in turn depends on the number of TB views: the broader the range of available incidence angles, the most ambitious may be the retrieval configuration. Another significant parameter is the expected equivalent optical thickness□ of the vegetation cover.

Therefore a number of options are prepared for the retrieval configuration. The parameters which stipulate these options are stored in a file, allowing further tuning after the launch.

One of the most promising features of SMOS is the possibility to implement a dual step processing, i.e. to use a τ value retrieved from a previous SMOS visit as initial value, together with a strong a priori constraint. To this end, a "current" τ map is built from SMOS retrievals and regularly updated.

Performance figure over vegetated soil

The performance of SMOS is described as its ability to meet the SM requirement over the largest possible range of SM and τ values. A convenient way to visualize this is to represent, over a (SM, τ) diagram, the fraction of the swath where the requirement is met. This is illustrated by Figure 1a-e, taken from the initial validation steps of the L2 SM retrieval prototype.

Figure 1a shows the SM retrieval uncertainty (DQX) averaged over a ± swath abscissa away from the track: this swath ensures a maximum 3-days single orbit revisit interval. As anticipated from previous retrieval studies (SRS reference), the DQX increases broadly with both SM (above 0.2) and τ. Figure 1b & 1c show the performance figures when requesting that the requirement is met either everywhere, or on the average, across the swath.

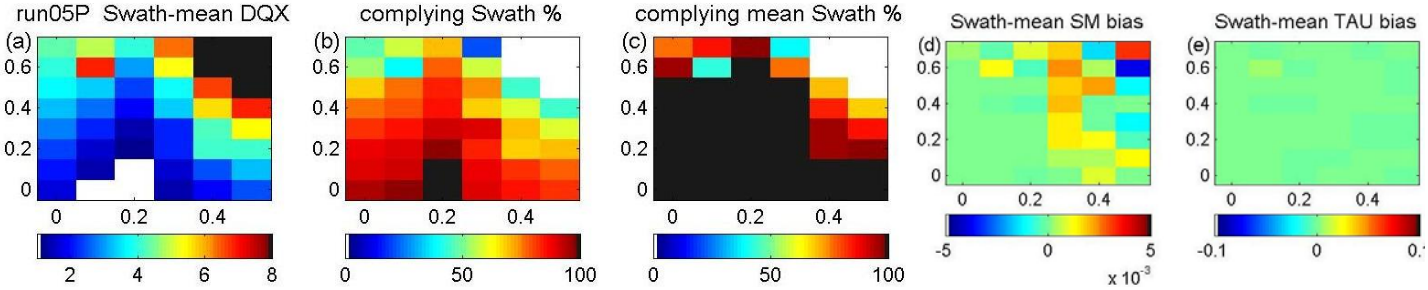

Figure 1: Example of performance map in the (SM, τ) plane

Figures 1d and 1e show the retrieval biases. While ideally they should be non existent, some bias is created because the TB measurement for orthogonal polarisation do not coincide in time, which results into slighting different values for the rotation and incidence angles. Although the observed biases are very small in the compliance zone, it is planned to improve the algorithm in order to remove them.

As seen from figures 1b & c, the fully compliant (black colored) area is rather narrow. The dual step processing, not shown here because the L2 prototype was not yet tested for it at time of writing, is expected to improve very significantly the performance. However, there will be clearly a limit in terms of optical thickness, which will occur at lower τ values for high SM values.

Conclusions

The most consuming part of the L2 SM algorithm is due to the computation of fractions for land cover. Due to the heterogeneity of land surfaces, this cannot be avoided inasmuch as estimates of SM should be obtained for the maximum possible number of cases where vegetated soil is present in the SMOS pixel.

In addition, the algorithm includes a large number of branching options, aiming at make the best of SMOS data in every case. However, every threshold for deciding retrieval option is stored in a separate file, in such a way that tuning the options along the spacecraft lifetime is very easy.

References

Claassen J.P. and A. K. Fung, "The Recovery of Polarized apparent Temperature Distributions of Flat Scenes from Antenna Temperature Measurements," IEEE Trans.Antennas Propagat., vol. AP-22, pp. 433-442, 1974.

Kerr Y.H., Waldteufel P., Wigneron J.P., Martinuzzi J.M., Font J., and Berger M., 'Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission', IEEE Trans. Geosci. Remote Sensing, vol. 39, 1729-1735, 2001.

Klein L.A. and C. T. Swift, "An improved model for the dielectric constant of sea water at microwave frequencies," IEEE Trans Antennas and Propagation, vol. AP-25, pp. 104-111., 1977.

Masson V., J.-L. Champeau, F. Chauvin, C. Meriguet, and R. Lacaze, "A global data base of land surface parameters at 1 km resolution in meteorological and climate models," Journal of Climate, vol. 16, pp. 1261-1282, 2003.

Sahr, K., White, D., Kimerling, A.J., "Geodesic Discrete Global Grid systems", Cartography and Geographic Information Science, 30, 121-134, 2003.

Saleh K., J.P. Wigneron, P. de Rosnay, M.J. Escorihuela, Y. Kerr, J.C. Calvet, M. Schwank, P. Waldteufel, "Estimates of surface soil moisture in prairies using L-band passive microwaves ", IGARSS 2007 Barcelona Spain.

Waldteufel P., J.L. Vergely and C. Cot, " A modified cardioid model for processing multiangular radiometric observations", IEEE Trans. Geosci. Remote Sens., 42, 1059-1063, 2004

Waldteufel P. and G. Caudal, "Off-axis Radiometric Measurements : Application to Interferometric Antenna Designs," IEEE Trans. Geoscience and Remote Sensing, vol. 40, pp. 1435-1439, 2002.