Hydrol. Earth Syst. Sci. Discuss., 7, 7007–7057, 2010 www.hydrol-earth-syst-sci-discuss.net/7/7007/2010/ doi:10.5194/hessd-7-7007-2010 © Author(s) 2010. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Strategies for validating and directions for employing SMOS data, in the Cal-Val project SWEX (3275) for wetlands

W. Marczewski¹, J. Slominski¹, E. Slominska¹, B. Usowicz², J. Usowicz³, S. Romanov⁴, O. Maryskevych⁵, J. Nastula¹, and J. Zawadzki⁶

¹Space Research Center (SRC PAS), Bartycka 18a 00-716, Warsaw, Poland
²Institute of Agrophysics (IA PAS), Lublin, Poland
³University of Nicolaus Copernicus (UMK), Torun, Poland
⁴Geo-Information Systems (GIS NASB), Minsk, Belarus
⁵Institute of Ecology for Carpathians (IEC NASU), Lviv, Ukraine
⁶Warsaw Technical University Env. Eng. Dpt. (PW WIS), Warsaw, Poland

Received: 23 July 2010 - Accepted: 30 July 2010 - Published: 17 September 2010

Correspondence to: W. Marczewski (wmar@cbk.waw.pl)

Published by Copernicus Publications on behalf of the European Geosciences Union.



Abstract

The paper tries interpreting how the method of SMOS observations realizes managing the problem of large scales and the target heterogeneity by means of employing the polarization angular signature. Land surface target on the Earth is naturally heterogeneous in its continuity of physical and biophysical properties. Soil moisture (SM) retrieval from SMOS data requires using the model CMEM to determine relations between the temperature brightness and water related properties and conditions, which are anchored to the ground by auxiliary data. SM retrieval must start from the conditions at least approaching physical reality. SMOS performs the data fusion in NRT
10 (Nearly Real Time) in a very specific way, what is a new quality added to EO (Earth Observations). The paper demonstrates several effects of employing the SM retrievals from L1C data. Authors explain how they validate few selected test sites in Poland, and come to conclusions on choosing a strategy focused on validating single sites. Finally, they come to an understanding that SM retrieval is an advanced statistical method re-

quiring good referencing to ground based physical conditions in large scales, worth confronting the shallow water content obtained from SMOS to that assessments of the total water content on continental scales, which available from effects of gravitational missions.

1 Introduction

- A proper strategy is a very sensitive subject of choice in methodology, being dependent on particular goals and levels of data under validation, and possible compromising practical constraints. In general, it must be determined in order to not loosing the relevance of observations to large spatio-temporal scales, when the subject is large, complex and combine heterogeneous entities under different physical processes. The inhorant unity of control and temporal demains must be respected. SMOS has already
- ²⁵ inherent unity of spatial and temporal domains must be respected. SMOS has already determined possible strategies by the principle of a radiometric interferometer, and



by a use of polarizations. The paper aims choosing a proper strategy for validations, respecting the principle of SMOS, and taking into account limited abilities for ground based observations. The aim remains in a proper recognition of necessary tools and needs for the project realized in Poland, and taking such possible directions, which are well determined by the existing knowledge on statistics and methodologies for man-5 aging data bases in meteorology, hydrology, ecology, and remote sensing techniques. In particular, some possible directions have been fostered and pointed out by a use of GLDAS data, and other collective data bases predicted for integrating hydrological data globally. GLDAS does it by means of four hydrological models, keeping consistency between water and energy related parameters, in a column. One preferred is 10 the NOAH model, since it manages the water and radiation transfer for layered media in a column. The GLDAS base, and SMOS observations are provided in global scales. They provide data on large spatial distributions of physical properties. Effects of mass transport processes acting horizontally, are finally expressed by the spatial

- ¹⁵ composition of observed elements, but are also in a special care about sensing them consistently over large areas, by means of particular instrumental methods organizing the image in the field of view. In optical sensing spectral variables, that methods are based on scanning, while SMOS does it completely differently taking care about coherence between component variables instrumentally. That care is for the same purposes
- of keeping relevance which motivates modelling processes of mass transport spatially, but in SMOS it comes with the observational principle of instruments, and determines unusual or new requirements on methods for employing data. It makes that sometimes the purposes of SMOS observations are understood for being esoteric and specific to the researches in global scales. However, SMOS only reveals in a new way a common
- fundamental need of synergy between necessary data sources, on large scales, and attempts employing accessible real potential existing in different remote sensing techniques. Modern strategies in observing must attempt real problems multidisciplinary, on multi paths, not only on the level of operating the mission, but also on the level of understanding complexity of observations and employing effects. SMOS is a very good



example showing, how water relates and organizes that what is observed on Earth, by means of differents techniques, and methods of interpreting effects. It enforces integrating sciences and techniques deeply, from physical fundamentals, through observational principles, up to estimating effects in agreement to advanced mathematical and in particular, to advanced statistical methods.

2 Validating SMOS data in Poland

5

Commonly, a strategy for validating observations on regional scales assumes employing airborne observations for wide coverage in some transects monitored on the ground, and/or establishing a sufficiently dense grid of reference monitoring stations.

- ¹⁰ The project SWEX (Soil Water and Energy eXchange) for SMOS Cal-Val program, engages several partnering institutes in Poland, which agreed on cooperating, keep their own local stations and perform own research programs according to their scientific profiles in particular domains. Partners are spread over the country and can't form a highly integrated grid. Considering the Fig. 1, one can see locations of the test sites, in Poland. The dilemma of establishing sparse or dense grids for monitoring in validation
 - purposes, is a common trouble for managing in other Cal-Val projects, see Jackson (2010).

IA PAS, Lublin, is the partner in SWEX, where a broad experience in determining spatial patterns of physical properties of soils by statistical means, has been devel oped and accumulated through last decades. Therefore, taking the large scale goals seemed being natural and expected simply in extending scales. It was expected that aggregation and/or disaggregation techniques, can assist catching a correspondence to SMOS data, well enough. However, the problem of discrepancy of scales, between all that what is available on the ground, to that what is available from space, appeared not only in scales. Some supervising principle is needed yet, to couple the spatial and temporal domains. The assist comes with the principle of SMOS, sampling the instant SM state in many shots, taken once per 3 days, under conditions for



the radiation transfer equation in layered media, and with the SMOS data order, keeping consistency of observations between many pixels over large snapshot area, due to the principle of interferometric imaging. This technology of performing observations from the orbit, determines possible strategies on the ground. For long term measures,

- the strategies must respect temporal evolution cumulatively, including available knowledge on the environment from independent sources, while for short term validation of the instant SM state, the strategies must involve available characteristics possible for taking from the ground, and respecting radiation transfer rules. In a consequence, that what is observed in temporal, respects large area, and that what is in trends, finds a
- steady maintenance repeated in long term cycling of observations. Physics related to the observation principle, respects statistical demands deeply in principles involving both spatial and temporal domains, respecting small and large ranges of scales. Small ranges in the spatial domain are supported by the polarization components related to the wavelength (20 cm) and the spatial target heterogeneity. Large spatial ranges are
- ¹⁵ supported by the interferometric order of taking data. Small ranges in the time domain are respected by the relevance of sensing by frequent shots, taken in every second when the target is in view (approximately 230 shots for a target pass), and in possibly the most quiet time of a daily cycle, in the time about the sunrise. Large ranges in the time domain are supported by the revisiting time not being longer than three to four
- ²⁰ days in low or medium latitudes. That is a very consequent and sustainable strategy of observations, not employing preferences on focusing high resolution or the radiometric precision. Therefore, one may conclude that ground based monitoring, should be also similarly focused more on relevance, than on the coverage by dense grids or the precision in measurement techniques. The ground based monitoring should be focused on
- statistical relevance to physics in a heterogeneous subject, by means of proper modeling, and on the areal ground based measures kept in representativeness of large areas under characteristics. The means available on the ground may be different from those on the orbit, but the scope of activities on the ground should comply that undertaken on the orbit. Things are not so simple that one could require only high precision and



extensive modeling from ground data. Neither precise measurements, or enormously great data volumes collected in dense grids on the ground, cannot save relevance to the water in heterogeneous targets, when the involved spatial and temporal scales are not joint in meso scales. Some limited support of the ground strategies is possible by

- ⁵ means of airborne observations. When not accessible, then finally the satellite borne means of SMOS must be used. A broad statistical experience in determining spatial patterns of particular physical properties on the ground, occurred not sufficiently helpful in validating, because it was too narrow for managing the heterogeneity in space of the target, and in involved environmental processes. Only particular properties were
- investigated separately. The more, the statistical experience occurred not sufficiently clearly related to the principles of retrieving SM from SMOS data, and methods of modeling (CMEM), which are in use for the SM retrieval from SMOS data. SMOS already engaged quite advanced statistical means forced by the nature of the polarization state characteristics, and an interferometric principle of surveying. Therefore, our indepen-
- dent attempts, not employing and not respecting SM retrieval on the base of the radiation transfer, occurred only helpful for developing skills in managing ancillary data from other missions, estimating vegetation indexes, and developing new ways of concluding. That way, our knowledge in soil sciences and statistics occurred helpful, but could not assist fundamentally in validation. Aggregation and disaggregation methods, different
- ²⁰ ways of modeling particular physical properties, and independent attempts to spatial hydrological modeling, are still possible for matching and including them to CMEM, when they can find proper places in the radiation transfer equations if necessary, and when they are designed for conditioning the SM retrieval from the SMOS, by means of L-MEB, or other similar algorithm, but not for substituting SMOS observations. Valida-
- tion of SMOS data by simple comparing the ground based physical property of SM, to the same SM value obtained from the orbit, is possible but is not applicable to other sites and cannot assist improving the SM retrieval from SMOS data. Concluding that was in fact elaborating a proper understanding of a desired strategy. When the project can't afford on employing the airborne supporting means, and can't provide a dense



grid of monitoring on the ground, then it must take such strategy, which follow the line pointed out by CMEM and L-MEB.

3 General methodology

In general, the methodology is ordered by the nature of data and is common to all sorts of ecosystems and capabilities of representing them in CMEM. The scope is put firstly on wetland sites in Poland, because expected effects of validation should correspond strongly to high water content within reach vegetation canopy. First the following actions were needed:

- to provide a representation of the environmental characteristics of selected test areas in Poland in terms of the CMEM model parameters.
- to apply CMEM for retrieving SM for a single SMOS pixel,
- to perform a comparative study on synthetic temperature brightness values with all collected SMOS shots,
- to introduce necessary corrections to the environmental characteristics and retrieve a final SM value within the error estimation.

One can say that the main purpose is in determining the estimations being representative to the entire pixel area of about 32 km in the diameter. Particular single measurements are available on the ground within nearly arbitrary precision, but are only specific to the test site and can't be representative the entire area, for example 32 km wide.

²⁰ The first approach was creating a synthetic data set for CMEM to describe a particular test site, as wide around the monitoring station as possible. Then the analysis employing the L-MEB retrieval procedure is repeated for many pixels around. The selected test sites were covered by 42 pixel subsections, available from SMOS in many satellite passes in a selected time interval. That allows on determining spatial and



10

15

temporal evolution in SM terms, and other parameters conditioning the SM retrieval. Exercising single pixel areas in a particular context, provides some evaluation of the consistency before aiming a map. It was realized, that currently our experience allows on processing full L1C snapshots covering the entire country, under managing it

- ⁵ in DGG system, but only selected single pixels served in the test sites on the ground shall become validated. In our national conditions, more important constraits are not in a limited potential for extensive measurements on the ground, but in inconsistency due to RFI contamination, and other reasons in data conditioning successful and repeatable retrievals. There are about 10 RFI sources spread around the country. A
- ¹⁰ proper corrective action for clearing RFI pollution, has been open with some preliminary successes. Two RFI sources have been recently cleared. Other remain and contaminate large area bringing non-physical effects in retrievals. SM is not the only one environmental parameter possible for retrieving.

The SM retrieval using RTE and the $\tau - \omega$ method described in (ATBD (2010), de Rosnay (2008), Wigneron (2007), Holmes (2006)), requires employing the fitting procedures L-MEB, with a use of synthetic modeling TB in CMEM. In effect of fitting the synthetic TB response from the L-MEB to the real response of the instrument in satellite data, one obtains the best estimate of SM and a series of other synthetic parameters estimated. The procedure of fitting is performed by varying several parameters conditioning the model. The effect of fitting is sensitive on the order of varying particular

parameters under the fitting. The ground based data determining the model, enable obtaing only the starting estimates, and the effect of fitting supplies the same parameter data corrected, that is how these parameters are "seen" by the satellite instrument.

Since finally, the interests are in seasonal changes, one needs scoping on effects in the time interval of about three months, for detecting changes in correspondence to the annual evolution for selected variable values. The SMOS data is currently delivered nearly with daily rate, for the morning, about 06:00 local time, and for evening shots, about 17:00–18:00. However, only an every third day contains FOV (Filed Of View) passing a test site centrally. More frequent passes are more or less aside to the



test site. There are sufficiently frequent data for checking consistency between daily shots. Currently we cannot clearly state reasons for which there are differences between morning and evening shots, which could be justified by a history of the day. It seems that the evening passes indicate TB values overestimated.

- Recently, since July 2010, L2 data for fixed test sites is delivered systematically, and a comparison between our own SM retrieved values to that in L2 data, is possible. Example data L2 for few selected test sites are presented on the Figs. 2 and 3. One can see that quite much of pixels in L2 data is filled by no data values. The retrieval was unsuccessful. Therefore, not a comparison of SM from the retrieved L1C data to
- ¹⁰ L2 data, is scoped for discussing now, but the conditions for retrieving. The area of Polesie/Podlasie (Fig. 2) is not contaminated by RFIs, but the area of Wikrowo (Fig. 3) is. Unsuccessful retrieval indicates that probably the auxiliary data, which determine starting estimates for the procedure L-MEB, or similar used by SMOS production system, may require correcting, or the process of retrieval should be supervised differently
- than it was. This is in scope of the SMOS data production, and of the validating teams, on independent basis. To explain why and how it is possible, that the retrieval may be unsuccessful, one needs describing how the CMEM model may be employed, and what means the supervision of L-MEB under retrieving. We cannot explain only what disturbs our successful retrievals from L1C data, on our own exercises.

20 **4 CMEM**

25

The Community Microwave Emission Model (called CMEM), was developed at ECMWF. It is based on the Radiation Transfer Equation (RTE). Fundamental definitions are given in the document ATBD (2010) and deeply developed descriptions are discussed by (Holmes, 2006). Currently only a brief summary is provided in order to explain the way of our proceeding. The model requires an input data set representing the pixel content by physical and environmental properties, parameters, soil compounds, fractions etc., and supplies the product of TB (temperature brightness) values in two



polarization components H, V, properly to the coordinate system on the ground. The product is a synthetic value of TB, possibly observed by MIRAS SMOS. Therefore, the values TB_H and TB_V , are assumed possibly representing the heterogeneous content of the pixel by effective TB Fresnel components. It means that one agrees on substituting a particular heterogeneous target, by a three layer media: the ground, the vegetation canopy, and the atmosphere, by the values of TB_H and TB_V . That media emit radiation up-welling, but also transform it by down-welling components, within each having its own emission and the temperature brightness. The received or synthesized TB_P

values must behave the radiation transfer equation (RTE).

10 4.1 Fresnel radiation transfer in layered media

The assumption on Fresnel conditions is a necessary simplification, corresponding to the way of managing alternating polarization regimes claimed DOUBLE, representing responses in H,V directions from flat media layers. SMOS employs the regime of FULL polarization on regular basis, but the regime DOUBLE may be considered sufficiently well for that part of FULL, which may determine SM values and most of conditioning parameters. The cross polarization components, postponed for a while, are predicted for other purposes in more advanced error estimations. RTE describes the temperature brightness $T_{B,toa}$ in two polarizations TB_P H or V, for the Top Of Atmosphere (toa), in a general form as the following

²⁰ $T_{B,toa}^{P} = T_{B,au}^{P} + \exp(-\tau_{atm}^{P})T_{B,tov}^{P}$

15

25

where $T_{B,au}$ is for up-welling radiation; and $T_{B,tov}$ is for the component from the top of vegetation canopy, under the polarization P and within the optical thickness of the atmosphere τ_{atm}^{P} , determined also for two polarizations. However, one needs modeling more, that is how is the soil visible through the obscuring canopy, in terms of TB. To describe that, a formulation must be developed on the next layer, what is called for the $\tau - \omega$ model based on the simplified solution of RTE. Assuming that the vegetation is



(1)

represented by a single scattering layer, above a rough ground surface, the $T_{\rm B,toa}$ is usually written as:

$$T_{\rm B,toa}^{P} = T_{\rm B,au}^{P} + \exp(-\tau_{\rm atm}^{P})T_{\rm B,tov}^{P},$$

 $T_{B,tov}^{P} =$ 5 $T_{B,soil}^{P} \exp(-\tau_{veg}^{P}) +$ $T_{B,veg}^{P} (1 + r_{r}^{P} \exp(-\tau_{veg}^{P})) +$ $T_{B,ad} r_{r}^{P} \exp(-2\tau_{veg}^{P})$

where $T_{B,au}$ and $T_{B,ad}$ are the up- and down-welling atmospheric temperature brightness; τ_{atm} means the atmospheric opacity; τ_{veg} means the vegetation opacity; and ω - is the single scattering albedo; T_{eff} is the effective temperature of the surface matter and the canopy temperature (T_c), which is usually considered being equal to either the T_{eff} or the air temperature; r_r^P and e^P are the surface reflectivity and emissivity, respectively, with simple relation between them $r_r^P = 1 - e^P$. The reflectivity of a non-smooth surface r^P can be described by a semi-empirical approach based on three roughness parameters h_s , Q_s and N_s :

$$r_{\mathcal{P}}(\theta) = \left[(1 - Q_s) r_{\mathcal{P}}^*(\theta) + Q_s r_{\mathcal{P}}^*(\theta) \right] \exp(-h_s \cos^{N_s}(\theta)),$$

where θ is the incidence angle and r_{ρ}^{*} the soil reectivity of a plane surface, which depends on the soil dielectric permittivity and the incidence angle.

In CMEM, the aggregated $T_{B,toa}$ consists of components related to emissions from individual tiles, which are: soil, vegetation and atmosphere, and are expressed with following relations:

- temperature brightness from the soil

$$T_{\rm B,soil} = e^P T_{\rm eff}$$

(2)

(3)

(4)

(5)

- temperature brightness from the vegetation

 $T_{\rm B,veg} = T_{\rm c}(1 - \omega^{P})(1 - \exp(-\tau_{\rm veg}^{P}))$

Parametrization of the effective temperature proposed by Choudhury et al. (1982) is applied in CMEM, and fulfills following formula:

 $T_{\text{eff}} = T_{\text{deep}} - (T_{\text{deep}} - T_{\text{surf}})C$

with T_{deep} and T_{surf} the soil temperature at specified depth and surface soil temperature. Wigneron et al. (2001) proposed for L-band to include in parameter *C* dependency of soil moisture and coefficients *b* and *w*0 what leads to formula

 $C = (sm/w_0)^b.$

The Fresnel equations relate the reflectivity of a smooth soil, r^* , to the soil dielectric constant ε , and the latter is a function of the moisture content at the surface amongst other soil properties. Following CMEM implementation, the Dobson model (Dobson et al., 1985) was used in this study to retrieve soil dielectric constant.

Considering a vegetation layer and $\tau - \omega$ approach (Mo et al. 1981), vegetation optical thickness is computed with semi empirical approach (Jackson and Schmugge, 1991)

$$\tau_{\text{veg},p} = b \frac{\text{VWC}}{\cos\theta} \tag{9}$$

where *b* and VWC are the vegetation structure parameter and the vegetation water content, respectively. The more precise formula for $\tau_{\text{veg},p}$ was introduced by Wigneron:

²⁰
$$\tau_{\text{veg},p} = \tau_{\text{nadir}} \left(\cos^2 \theta + t t_p \sin^2 \theta \right) \frac{1}{\cos \theta}$$
 (10)

with two individual cases: for a low vegetation layer $\tau_{\text{nadir}} = b^{\circ}\text{LAI} + b^{\circ}$, and for a high vegetation layer $\tau_{\text{nadir}} = b^{\circ}$, tt_{ρ} represents the angular effects on vegetation optical thickness for each polarization and vegetation type.



(6)

(7)

(8)

Since, the soil parameters are considered important, the formula from Kirdyashev is also tested. In this approach parametrization expresses by means of: the vegetation optical thickness as a function of the wave number *k* (between 1 GHz and 7.5 GHz), the dielectric constant (imaginary part) of saline water $\varepsilon_{sw}^{"}$, the volumetric water content VWC, the incidence angle θ , the water density ρ_{water} and a vegetation structure parameter a_{geo} . All that leads to the formula

$$\tau_{\rm veg} = a_{\rm geo} k \frac{\rm VWC}{\rho_{\rm water}} \varepsilon_{\rm sw}^{"} \frac{1}{\cos\theta}.$$

5

Our early attempts to using CMEM, were simply minded, performed without real SMOS data yet, and were static exercises, in order to organize the data necessary for validations, and to try generating TB. When we processed many pixels, then an image was composed, and that way we could see how TB varies across the spatial domain, according to the input data.

CMEM needs a series of input data components, and SMOS also needs similar the same data, in the auxiliary data. That data is for generating a synthetic value TB by the model, but not for determining SM, from SMOS data. If SMOS data need similar AUX data, it does not mean the same as in modeling. SMOS needs it for converting TB measured in L1C, to SM determined in the level L2. In effect of conversion SM value is determined by TB, but not by AUX data, though these data plays an important role of drawing the starting estimates closely to the state converging to final values of SM,

- ²⁰ corresponding to measured TB. The L2 converter works also like CMEM, but in the opposite direction. However, the converter is not an inverted CMEM exactly. CMEM is not directly invertible, though it can be inverted for a particular input data set, in limited ranges of variables. CMEM is designed for all conditions possible globally, that is for 219 types of ecosystems, what is recognized by the content of input data. It can be
- inverted for particular types of the environment, and satisfy applications in acceptably broad ranges of a land use classes, but cannot be shaped inversely the same way for all types globally. Concluding, the retrieval of a particular variable or parameter, requires CMEM like model, or another, working in two directions. That enables searching



(11)

the best fit of the model output to the TB from the orbit, by common methods of the curve fitting and under a defined cost function, with all relevant conditioning of TB by many parameters. The criteria on convergence must be defined, and a strategy for variational changes for particular variables must be decided. One may design better

- or worse algorithm on iterations, but some strategy must be chosen. Success in the fitting depends on many analytical formulas in many elementary models engaged by CMEM, but depends even more on the starting state of iterations. Therefore it is safer varying a single variable chosen for searching a local minimum of the cost, than putting on variations many or all parameters determining a current state of the cost and the
- ¹⁰ product TB. The supervision is performed by setting constraints on particular ranges for varied parameters, and eventual choosing the parameters intended for varying. That is known, that some procedures are are able managing the search with several variables, but it is not recommended, in general. The control on ranges for variations is just the supervision, strictly related to the ground based knowledge. The process may be
- performed automatically but within well determined conditions for starting it, and this is a function of the AUX data. One should not take the understanding that the ground data are postponed, put on variation freely and all effects are gained incidentally. That would be completely wrong understanding of the multi variable functions. Particular variables involved may be not fully independent, and that happens in real targets, but
- ²⁰ the sense of CMEM is just in that the variables contributing to the final product of BT, are ordered in a hierarchy proposed in CMEM, and driven by respecting RTE, under assumed simplifications to Fresnel conditions, to DOUBLE polarizations, to the principle of substituting a heterogeneous rough layers by Fresnel layers under conditioning by means of the "roughness", " τ ", " ω ", etc. Few demonstration examples collected and
- ²⁵ displayed on the Fig. 4 can provide an insight to typical relations between TB products, in two polarization components TB_H, TB_V versus the incident angle and the dielectric constant values Real and Imag. This set of the TB function patterns, show possible behavior of TB components versus the incident angle, as a preferential choice among possible characteristics, if they are sought for dependence on the dielectric



components. This exercises are given to express that the retrieval by fitting the best approach to real data, can be driven under supervising by an essential knowledge on modeling the contribution from the dielectric constant. Similar plot can be created for other contributing parameters, but the success of supervision strongly depends on the

- ⁵ choice order among the parameters. Some parameters like " τ ", " ω ", roughness combine contributions bringing much of the dependence from other parameters like the dielectric constant, the soil moisture, etc., but the order of these contributions is in the hierarchy of these parameters and should not have been undertaken freely. It creates the need of elaborating some well defined strategy, respecting relationships between
- elementary physical models in CMEM, and checking efficiency of the decision tree, for performing processing and the supervision. Most of that necessary conditioning is already done and prepared in CMEM, but a user is free in elaborating his own tree of decisions and checking it for a desired class of ecosystems in the land cover.

4.2 Retrieving parameters and cost function

¹⁵ The retrieval of model parameters and surface properties is based on the optimization of the CMEM modeled values with collected measurements. The inversion technique uses a modification of least squares cost function given by formula:

$$CF = \frac{\sum (TB_{\theta,P}^m - TB_{\theta,P})^2}{\sigma_{TB}^2} + \frac{\sum (p - p_i^{\text{ini}})^2}{\sigma_p^2},$$
(12)

where subscripts *P*, θ denote dependency on polarization and incidence angle, ²⁰ TB^{*m*}_{θ,P}, TB_{θ,P} refer to temperature brightness, respectively measured and modeled, *p* refers to retrieved parameter, and *p*ⁱⁿⁱ its initial value, and with variances σ_p , σ_{TB} . For a given time pixel of interest (ID point), a set of TB measurements was available (five angles and two polarizations), from which a single parameter was retrieved. To determine the best fit parameters a numerical algorithm of the Levenberg-Marquardt

²⁵ type optimization method was used.

The process of fitting is necessary, because the product from the model is obtained essentially on artificial way, by means of combining analytical function or formulas, driven by input data deterministically. Products from the model are contiguous including first derivatives, while the orbital instrument is driven by a stochastic input and data values are spread, especially under splitting on polarization components, when each component reacts on linear polarizations differently. Circular polarization components are in essential not represented by CMEM. Real data must be spread, therefore fitting must be done in an iterative process.

Expected TB vales for H, V are not constant, because the instrument moves along the orbit. All what is received at antennas is under moving, under varying incident angles sweeping a wide range of angles, and additionally under possibly large integrating constant, comparable a second. The instrument performs about 230 shots (3 shots form a polarization snapshot), while it still "looks" at the same pixel. This is the desired observational feature, but it accounts that in that time quite much happens in the target and it is not static. The data values are still varying and the spread of data is large.

One must fit a model to data for estimating its values.

5

We can't dare describing the L2 converter because it is on the SMOS side, and is still under development. However, the " $\tau - \omega$ " scheme for expressing RTE is the same, and the same are other component models for determining physical compounds on

- the ground. Therefore, SMOS needs the auxiliary data, and the same CMEM, but possibly taken from different sources. Most critical ground parameters in auxiliary data for CMEM may be generated on the ground, from the best knowledge on the test site, while those less critical may be even the same as the auxiliary for SMOS, if they only occur sufficient for validation. One can perform validation better, locally in a pixel, than SMOS can do converting TB to SM L2 globally.
 - 4.3 Exercises on fitting to real SMOS L1C data

Real SMOS data L1C level were used for exercising the procedure of L-MEB, driving the process by the formulas on TB derived from the CMEM model for two wetland



areas Biebrza and Polesie. The process searched for the best fits evaluated by the cost function from Eq. (12), with results shown on Figs. 5 and 6.

Input data for each test area was determined very roughly on the base of the ground data in single test sites, without developing disaggregation of a particular DGG pixel

- ⁵ by statistical analysis, yet. The retrieved values were for the real part of the dielectric constant at Biebrza, is reasonable, but for the wetland Polesie not. The retrieved value occurred negative. It means that the starting point data could be distant for the state corresponding to reality. That exercises were only for determining procedures of processing. Corresponding data on the retrieved dielectric constants, in bot real and imaginary components from SMOS L2 are realistic in both areas, and repeatable in
- In aginary components from SMOS L2 are realistic in both areas, and repeatable in several passes, and available in maps for nearly all DGG pixels in these subsections. This project just started with the definition of procedures for processing. However, we means the results encouraging.

4.4 Practical approaches

- ¹⁵ CMEM is a modular software package computing emission, in essential for Fresnel components, combining several variables from elementary physical models into one complex response representing microwave emission from soil for smooth targets. The product is expressed in TB for horizontal and vertical components, properly to nadir in the coordinate system on the ground. It is a kind of a collective model of physics
- in soil, related to emission. It serves physical reference analogue measures of TB, that are finally measured by SMOS from the orbit. However, that is not an analogue substitute of the instrument and its principle. CMEM allows referring the ground based knowledge, to the data from space. It organizes input data in some hierarchy necessary for using the radiation transfer equation and elementary models. Output product
- from CMEM is only comparable to TB in L1C data. The opportunity of comparing TB in L1C data to TB known on the ground, is enhanced by the opportunity of investigating how the data ground contributes to SM retrieved from L1C data, which parameters are more or less sensitive, and in what ranges of their values. SMOS instrument works in



FULL polarization. CMEM is constrained to only two polarization components H,V, at nadir. Here are subtle but very important differences for employing SMOS data in coherent and incoherent polarization components. The first conclusions on a necessary strategy in validations came with the real data from SMOS. CMEM products behave deterministically on input parameters, while L1C data supplies stochastic responses. It was the first challenge for accepting consequences that data from the instrument must be employed for a real value, while CMEM model and even the best ground based input data may only assist in consuming this value.

Another aspect of taking a proper strategy for validation is that CMEM supplies its products for a particular pixel, which may be an arbitrarily small or large pixel like for SMOS. It can work for a stream of many pixels of data but must be compared to a particular response of the MIRAS instrument. A user can organize validating in many sub-pixels but the data in L1C is organized in a way being fixed and ordered by DGG (Digital Global Grid) system. It means that a task of organizing data in the image is

- ¹⁵ put beyond CMEM. That creates a secondary challenge for organizing rules for dis aggregating or aggregating data, and overcoming a preliminary trouble how to manage a problem of the gap between scales for all activities on the ground, and the data from SMOS. The trouble is rather common and universal for employing large scales, but a use of DGG system forces a specific nature of non-raster data of L1C level,
- and demand new skills for transferring results between DGG and raster data. SMOS created that trouble, but also fostered tools and assisting solutions. One can mean that it is a way of forcing new thinking and managing data in large scales, or in global scales in particular. Tools like BEAM, NEST, SMOS Viewer and many others introduce new practices being not known before.

25 4.5 Soil sand clay compounds

To run CMEM, one needs defining a set of input data to represent the target. Gathering first necessary input data, some examples were taken from ECOCLIMAP [5] data bases, on soil compounds. This project, needed confronting the data from ECO-



CLIMAP [ECOCLIMAP (2006)] on soil compounds to the data available from official national resources. First output products from CMEM, revealed a meaningful contrast of TB, across the Eastern boundary of Poland.

That was an artificial effect only, because related only to input data not being a real observation. One cannot state more than constrained quality of data. Similar contrast appears also in the sand soil compounds in official soil maps in Poland, and repeated the same way in auxiliary data from ECOCLIMAP. The differences in the land use and the environment across the borders to the neighboring countries, seems to be not justifying it. Soil standards employ more than 100 possible soil classes. Each class has the sand compounds defined in the range of 20–50% wide, and one cannot simply convert these soil types to the desired fractions, because fractions vary with the location, morphology, etc. A supervision based on the local knowledge is needed.

Such work was done in IA PAS, Lublin for a comparison to the auxiliary data from ECOCLIMAP is shown on Figs. 8 and 9. We hope that a use of this data in validations shall clarify whether that contrast in TB repeats across the state borders, due to differences in standards in characterizing soil compounds, or not, if the auxiliary for starting a fit by iterations would occur being not so much affecting the SM retrieval.

4.6 Statistical modeling of soil properties

The project aims also employing the statistical-physical model given by Usowicz (Usow-icz, 1993, 2000), IA PAS, Lublin. The concept was developed firstly for statistical modeling other physical properties of the soil, treated as a mixture of compounds defined by particle fractions. Figure 10 gives a general view of a geometrical configuration. A sample of the mixture, is substituted by equal radius spheres, representing particular fractions. The spheres contact one another, create paths for conduction, and substitute the capacity. The idea is based on introducing a net of the conductive paths and capacitive storage elements. It corresponds to the heat conduction, and the capacitive storage of the charge representing the heat capacity. The purpose is to related



Similar idea was used for approaching the electrical capacitance and the electrical conductance, both for representing the dielectric constant by its real and imaginary components, and the same in the relationship to the water content. The spheres are only abstracts, being necessary to define starting conditions for a combinatorial play,

- ⁵ for determining what is a number of spheres for each assumed compound fraction, which can define sufficiently well the relationship between the dielectric constant and the water content, under particular soil type definition. The spheres are to allow on describing a desired characteristic in terms of the probability of contacts and conductive paths, by means of a common formula proper to the series and parallel network
- ¹⁰ capacitive and conductive elements. All three regular state phases, the content of gas, liquid and solid compounds, can be entered at input by fractions. However, the control by numbers of necessary spheres gives a desired opportunity on wide and free substituting different sorts of the soil mixture. The motivation to this modeling technique, was finally in finding a relation to the property of porosity of soils. This is a very im-
- ¹⁵ portant property which determines strongly possible water contents in soils. The aim of gaining a unified control of the mixture, is a substitute of other empirical formulas. The idea is formally similar to the concept of the " $\tau - \omega$ " method, which also enables choosing a desired synthesized angular signature of TB, from among an entire class of functions defined by resolutions of the RTE, though the radiation transfer and polariza-
- tion states, are not involved. One can see the task of fitting the CMEM characteristic to the angular signature as looking for a type of a functional pattern. The same goal was taken similarly for this model, but in a different domain, and under less complexity. The concept was firstly developed for thermal properties, what is explained in several papers by Usowicz (Usowicz, 2000). Now it is extended on dielectric properties of soil.
- A general expression on the real part of the dielectric constant value is given by the following:

$$\varepsilon_r = \frac{4\pi}{u\sum_{j^L=1}\frac{P(x_{1j},\dots,x_{kj})}{x_{1j}\varepsilon_1(T)r_j+\dots+x_{kj}\varepsilon_k(T)r_k}}$$



(13)

This expression on the dielectric constant, corresponds to analogue electric substitute networks for behaving fundamental Kirchoff's and Ohm's laws, under modeling the dependence on water by means of the number of spheres u, and the number of possible contacts L. One of elementary configuration of spheres is shown on the Fig. 11, for

- the case *u*=3 and *u*=7, when the spheres are to exemplify how all the phase states are substituted by equal spheres, differing only by numbers involved. That example was to generate possible characteristics in dependence on water content. In effect of determining a proper number of spheres, the entire range of the domain is discretized, and available types of characteristics for modeling, associated to that numbers of spheres, are shown on the Fig. 12. The clue is to provide a class of available functions, de-
- ¹⁰ are shown on the Fig. 12. The clue is to provide a class of available functions, determine them by the granulometric fractions of compounds, before final calibration to physical values.

That way any desired mixture can be defined on formal way for other modeling purposes. The model was very intensively checked in physical cases of many soil types, and was employed for supporting calibration of TDR instruments. Currently the interest is in using that statistical model for CMEM, in order to introduce and opportunity for fitting the CMEM response to the angular signature from SMOS, on the base of the porosity versus water content. The models of Dobson, or Mironov, could be checked in relation to the porosity.

We expect that coupling this model to CMEM, one can achieve new possibility on matching the optical thickness of soil τ on the base of the porosity. The area of a SMOS pixel is not intended for investigating the spatial distribution of the porosity of soils, but SMOS rather more then less senses the optical thickness and the dielectric constant. It is believed that the model can enhance capabilities for SM retrieval.

²⁵ Figure 13 displays the proof on agreement of the statistical-physical model on the dielectric constant property to few other models. That reference models do not cover Dobson or Mironov, however.



4.7 Common platform for validation

It was assumed that the project takes tools from BEAM (2010) and NEST (2010) for confronting input data and products of validation. They employ ENVI standard of binaries and necessary meta data for reading them, organized in bands. The bands can

 ⁵ be stacked, re-sampled according to needs, and put into common statistical characteristics on one platform. That way satellite products can be combined with ground data, including SMOS data. One can use other platforms, but NEST and BEAM are just created for similar purposes. The choice seems being proper to purposes, and can easily provide importing/exporting products to and from, handling the formats NETCDF and
 HDF5.

GIS tools have not been used, except needs of importing some auxiliary data from CORINE or similar sources. Figures 7, 8, and 9 are the products from data imported to NEST via NETCDF format, created and exported to other formats when needed. NEST allows on creating large complex products, and provide the control of meta data

- ¹⁵ by XML organizing files. The Table 1 is only for listing what sorts of data are drawn to NEST and what their contents is. The table is not representative to full specification of CMEM input data. Proper information is given in CMEM documentation and the documents referred. Some variable are values, some are indexes, with eventually other information on splitting classes. The order of data is necessary for satisfying a hierar-
- chical structure of the emission models in CMEM. Formally all data is delivered to input in matrix order, and the output product of CMEM, the variable TB behaves the input matrix order. NEST creates a possibility of running CMEM from its platform, employing control on the XML level. Other SMOS tools are employed independently for choosing proper SMOS pixels to be validated. That way SMOS, other satellite and ground prod-
- ²⁵ ucts can gain necessary consistency under proper geo-location, projections and other order for identification. Currently that system of tools is exercised for gaining proper skills and flexibility, employing SMOS simulations and ground data.



4.8 From local to large scales

Comparing the local area tested on the ground for the spatial SM distribution (Fig. 14), to the spatial SM pattern from ASAR (Fig. 15), one cannot state that the patterns are matched, or that the local pattern is extendible to the satellite pattern, with the areas related approximately like 1:100. Only the kernel of highly concentrated SM confirms

related approximately like 1:100. Only the kernel of highly concentrated SM confirms one detail of the ASAR image. A validation task is not looking hopeless, however. The ASAR pattern is anchored to the ground. While working on the ground, the concept on catching correspondence to large SMOS pixels (35×35 km), appeared simply as extending scales, to intermediate scales (ASAR, MERIS). First one must have spatial data for extending, and much efforts was put at the wetlands Polish Polesie, and wetlands Biebrza test sites.

At Polesie, particular interests were focused on thermal properties of soils and statistical modeling the physical thermal conductivity, diffusivity, heat fluxes in soil, and the topics leading to the energy balance. At Biebrza, interests were put on ET (Evap-

otranspiration) in its relations to the water and energy balance. In that light, the scale extensions seemed to matched the concept of extending scales. There was a good basis in data for taking the direction from small scales to large.

From the other side, there is a series of papers on employing pixel disaggregation methods. It seems being supported by aiming to fill the area of large pixel with the en-

- ²⁰ hanced ground supported content taken from various sources and the ground. Means for aggregating and disaggregating are similar but employed in opposite purposes. However, a sense of using one of these methods, is whether they can support keeping temporal variations and changes. The more details, is involved spatially, the more important is synchronism in monitoring. Satellite data provide nearly perfectly instan-
- taneous imaging, but with limited repeatability, often rarely available. Geostationary boards satisfy repeatability, but not resolution, for those who take disaggregation. The line of generalizing going from local to large scales, was chosen anyway as more clear, in the context of SMOS.



Thinking about large scales in terms of spatial distributions, as it is a practice on the ground, becomes misleading. Staying on the ground one experiences the heterogeneity in both scales, spatial and temporal, but can't conclude equally on taking care about consistency of scales without a concept leading how to achieve that. Simple exten-

- sions very soon fail, because become non-realistic or occur insufficient. Developing dense grids generate costs and plenty of new problems. Sparse grids are obviously insufficient. Statistics knows the methods for determining optimal sampling grids, but they are poorly applicable when the subject is too complex and known in many aspects poorly. Similarly with modeling, which of should not aim modeling all details and company the problems in their complexity. That is not abvious that the problem of large
- component processes in their complexity. That is not obvious that the problem of large scales and the problem of heterogeneity are strongly coupled, and require a common treatment. Advanced statistical methods applied simply do not assist, until the problem is not defined sufficiently.

Fortunately, it can be recognized that the method of SMOS observations, combines complexly available solutions for consistency of large scales in both spatio-temporal domains and the heterogeneity, in one common principle of interferometric imaging, under respecting the state of polarization properties. Therefore, SMOS observations and also validations of SMOS data, looks realistically.

SMOS takes snapshots over the 1000 km swath, what makes clear that the entire image is taken instantly. The same action on the ground is hardly available. Even thinking about consequences of the instant SM state, taken approximately every third day, seems being unusual. Approximately three daily cycles are missed, what is also against common expectations for keeping temporal relevance. However, images are taken in 1000 km wide snapshots, and in a number of 1 sec shots, up to 230 in number,

for a particular place remaining in the FOV (Field Of View) for 230 s, while the satellite is moving on the orbit. Three shots form a single snapshot, assigned to alternating polarization. A series of about 70 snapshots provide a varying value of TB polarization components (Dual or Full polarization) under still varying incident angle sweeping the range of approximately from nadir to 70 degree. That way the series of snapshots



realizes multi looking in a great number of looks, on the same target component, and for each pixel of a huge 1000 km wide snapshot. All pixels are imaged nearly instantly, what fulfills the postulate of synchronism in large scales and coherence of their relations. A three day revisiting, that is a three day sampling is not an essential loss for observations because the purpose is in watching long-term trends. But from the reverse side, the relevance is strongly enhanced by the multi looking in the number of 70, and wide range of incident angles from different positions.

5

Microwave methods do not possess abilities for characterizing the target matter like spectral methods, due to larger wavelengths. Spectral methods characterize the tar-

- 10 get matter in many spectral channels, using multi or even hyper spectral techniques. However, they also need multi looking techniques, and use them for enhancing observational capabilities, by real or virtual looks, but never in so great number and with so wide range of incident angles. Spectral techniques perform observations excellently, but under limits of operating on variables in the measures of intensities. SMOS ra-
- dio interferometer operates by complex values and not having the potential of spectral analysis, it has another capability of reacting on the boundary conditions in the target by phases of the received radiation, now enhanced by the polarization state analysis. This is that what is currently desired and employed by all observational currently operating microwave missions, from Earth observations, to the cosmological missions observing
- ²⁰ CMB from the most distant deep space. Currently, the polarization techniques seems to be the most powerful tool which has not been exhausted in its potential. SMOS employs it intensively.

SMOS integrates the measured emission over the pixel area, taking into account the entire existing environment even with respect to geometrical dimensions, boundaries,

²⁵ and optical thickness, respecting the wavelength. Ground measurements remain being specific to a particular site of the measurement site, while the entire surrounding environment is simply ignored, until it is not included into a statistical estimation by spatio-temporal modeling, or into some other thematic model developed over the area, like for instance a hydrological model of the catchment. Even a very good grid of the



test sites, cannot keep consistency of data so perfectly in spatio-temporal aspects, as the SMOS instrument performs it. Besides that, any great number of samples and measurements taken from the ground is still poor and less than sufficient for statistical approaches. SMOS delivers data massively.

We demonstrate a believe that the principle of SMOS provides de facto a statistical observation based not only on the physics of emission and instrumental specific, but also on statistical aspects of further processing the polarization signature. The instrumental principle was developed with with great respect of fundamental postulates statistics. If that was not an aim taken directly, it came in effect of technical principles of interferometric radiometry employing polarizations.

4.9 Catena and the water field

In the project, the test areas are not supported by a spatial hydrological model on meaningfully large scales. Partners have not been able fostering such a model. However, a spatial pattern of SM, and the water balance in some test catchment, may be a final purpose possibly supported by employing SMOS data, not being a task directly

- conditioning validation. Watching temporal variations of SM, in relation to precipitation, runoff, water retention and ET due to vegetation, requires more other data sources than only SMOS. Some test land area pushed into wet beginning conditions by the precipitation impact, is drying differently depending on morphology, physical properties
- of soil, and vegetation cover. All that should be organized in a hydrological model. Not having such model developed, one can take monitoring SM dynamics in few representative slopes, that is in selected catenas and extend the description on the surrounding area on the base of physical properties of soil, at least. The topic is known as a description of the hydraulic behavior of soil being genetic due to its physical prop-
- erties. It has been postulated by Polynov (1925), and defined as a problem of catena by Milne (1935). Since that times it was investigated by many others and now there is some experience continued in Belorus. That early works in soil sciences were focused on deterministic methods in describing soil properties, while currently the problem is in



determining SM and other properties representatively to large areas under presence of processes proceeding in a great deal stochastically. Now we want coupling that valuable heritage concept of catena and genetic properties to currently employed data bases and models for including relations and dependence on the vegetation canopy.

⁵ The experience in soil knowledge on wetlands, accumulated in Belorus, should assist very much in observational tasks of SMOS on the huge wetland areas of Polesie. The idea is to engage existing knowledge of large European wetlands to validating SMOS data, firstly by employing the concept of test catenas.

The catchment is a fundamental entity in hydrological models. Slopes are compo-¹⁰ nent entities of catchments. Each catchment is composed not only of the water paths conducting the surface runoff, but also of slopes, that are the catenas. A test catena includes at least a slope, eventually with the foot and the culmination of the hill. Its characteristic behavior in time and space, needed for hydrological modelling, includes particular distribution of soil types and soil properties, which are effects of the history of sedimentation and the land use. In most cases, the order of soil types from top to bottom, is more or less stable, frequently regular and common for the types of geological structures and landscapes.

The "water field" properties beyond the selected catena, can be determined first locally and then developed in etxtensions by scaling and parameterizing the spatial dependence in other places, on the base of a reference catena and the distribution of physical properties. Anyway, that is an attempt for applying a rudiment hydrological model, for catching the correspondence to other accessible data on the area of the entire catchment, including SMOS observations. The aim is watching spatial SM distributions in SMOS by limited necessary proofs of the properties gathered on the ground.

SMOS provides its observations instantly and consistently over large scales, giving estimates being areal representatives over large pixels, in terms of BT, and finally in SM. The water measures of SMOS are believed to be deeply consistent over large areas, that is possibly being more relevant in large spatio-temporal scales than in situ



measurements. Therefore, not having a ready developed hydrological model, one may expect assisting the temporal relevance of SM in a large pixel by the ground based SM data, in few representative catenas. The concept was given generally for paedology by Romanova et al. (1976, 1981), and now is proposed for developing it in validating

SMOS data in Belarus and Ukraine, where there is a huge wetland area Polesie, spread on 700 km East, along the borders between these countries. Only the small western tip of Polesie is located in Poland, the most eastern part of Polesie, is in Russia. Firstly we want to propose a program dedicated to two TBRs (Transboundary Biosphere Reserves) laid along the eastern borders to Poland. The first one is TBR 1 Bialowieza, the other one is TBR 2 Western Polesie, depicted on the Fig. 16 by means of SRTM elevation.

5 Employing SMOS data

15

Directions for exploiting the data L2 in, follow the destination of SMOS. SMOS is predicted for large scale assessing water conditions, in spatial and temporal variations by means of quantitative measures. Therefore, we expect main advantages in regional scales, than for purposes of water management, though such may occur meaningful in maturing this kind of Earth Observations.

5.1 Shallow subsurface water resource

SMOS revisits the same site on the Earth every 3rd or 4th day, in the local morning time approx 06:00 a.m. The time of the operational pass of the satellite, is characteristic in that then the gradient of soil temperature slowly changes its sign, what means that processes of ET are just starting. They are not developed and variations of the energy exchange rate are low and slow, what occurs once per a daily cycle about the Sunrise. The time of Sunset is not slow and stable. SMOS samples TB, or SM, loosing 3 to 4 daily cycles. The daily cycle is not and objective for observations. The physical



terrain and its environment performs like a factor of integration over the 3 to 4 cycles, because physical matter accumulates all precipitation, runoff, retention and ET. SMOS samples the accumulated effects in the most quiet time. However, SMOS performs with a limited precision. Using CMEM, one can not that all models are simple and ⁵ get input data in relatively few classes. Among the consequences of that, is a loss of precision due to generalizing measures, and one of them is the optical thickness. It is believed that the optical thickness is approximately 5 cm. Only 5 cm thick soil layer with the unsaturated water content can be estimated. The diversity of soil properties, vegetation canopy, and necessary generalizing by data, make that thickness of this 5 cm
10 layer cannot be sure. However, most of the error contribution can be expected from

- the diversities in spatial distributions. A quiet interval of sampling the site minimizes errors in temporal contributions. Uncertainty in optical thickness, makes that absolute measures of the water content is biased. If one cannot surely extract the bias then can employ increments or decrements, the differences between subsequent samples. That
- ¹⁵ way the absolute measure of the water volume is not improved, but the water volume increased or decreased can be get know better. From one sample to another, that what is responsible for the observed changes is mainly an accumulated effect of all processes moving, loosing or gaining water. Similar line of concluding is employed in interpreting data from gravitational missions. Extrema of water concentration flow over Northern Europe in monthly data of anomalies, and the scale of SMOS should reveal
- correspondence in surface observations.

5.2 GRACE Gravitational Data

Gravitational missions provide results of the nonuniform gravitation, in spatial and temporary variations, by means of the thickness hypothetical equivalent water layer laying on the geoide, in (cm). They employ GRACE data, and share a common expectation that, next important sources of water contribution to the prediction, may come from the estimation of the unsaturated water layer in soil. The gravitational mission do not recognize whether the water resources are stored deeply under ground or are in shallow



layers. SMOS can provide some base to that recognition. Temporal behavior of the surface water resources is dependent on the ground water. SMOS data interpretation may need such information on large scales as the background.

The choice of different data sources shows that such anomalies move on the conti-⁵ nent, and in some periods appears over the region of Carpatian mountains. Our test sites at Western Polesie in Poland, are located between that extremal contours. In SRC PAS, there is the group of researchers working on prediction of the Earth poles nutation.

But also such data on the ground water bound in large catchments is very important for purposes of investigating climate constituents, and even for assessing trends important in smaller regional scales. Figure 17 shows annual trends in monthly anomalies of the total water stored in the catchments of selected great rivers in the region of Poland and eastern neighboring countries. That data is believed being in good agreement to common hydrological observations, even if such works were not elaborated in Poland.

¹⁵ Therefore we plan undertaking a use of interpreted data accessible publicly for catching correspondence to SMOS data interpretations.

6 Conclusions

In ground based activities for validating SMOS, we take the strategy of extending scales from small to large. The aim is to support a validation of a single SMOS test site pixel.

- ²⁰ When the SM retrieval from ground data and other auxiliary data when needed, is achieved properly, then the correspondence of ground based characteristics to this particular DGG pixel from L1C data appears behaving fitting the TB components well, under variations of optical thickness (τ), albedo (ω) or emissivity/reflectivity, roughness, dielectric constant and SM. We assume that successful fitting of the CMEM TB prod-
- ²⁵ ucts to L1C, by means of L-MEB, should be available, providing these parameters from effects of the retrieval for comparing at least by acceptable ranges, and in a number



of 3 to 4 parameters mentioned above. If one of then cannot obey properly under retrieving, is out of acceptable ranges, then the retrieval is unsuccessful, and requires a supervision on reasons. They may be in RFI glints, or other biasing conditions. The comparison of these parameters to the ground measured values, including SM and the estimated error, are effects of validating. When the starting point characteristics of

5

- the input data for L-MEB and CMEM is relevant to the environment in the pixel, and the data quality in L1C is sufficiently good, then the procedure can be applied to other surrounding L1C pixel data with slight changes supervised by the user knowing the surrounding terrain. That way effects of relevant validation on one pixel monitored directly
- on the ground, are extended on interpretations of few or tens of pixels around, without the need of extending all monitoring on larger area. The same, but a bit differently is done in the process of retrieving SM in L2 data, by the the SMOS data production system. The difference is that SMOS uses the auxiliary data on global scales, and do not monitor it directly everywhere but in selected test sites in SVRT Cal/Val program.
- ¹⁵ Currently the data in level L2 is not officially released yet, and it contains several parameters obtained under retrieving substituted by constant values, and/or with many pixels with no data retrieved. This a process of tuning the procedure and models for retrieving SM, and other parameters of environmental characteristics being relevant to water condition. SMOS is for observing SM on lands but it equally provides other op-
- ²⁰ portunities on assessing the environmental state. That opportunities are open for the end user. One can also conclude that SMOS develops new paths for integrating many kinds of other Earth Observation data and means. It is not an exaggerated expectation, that effects may appear not only in aimed SM global assessments in relationships to the climate change, but also in developing new methods of managing data bases
- ²⁵ and data fusion. Other else conclusion is that the observation principles of SMOS forced a broad use of advanced statistical methods, necessary for managing with the problems of large scales and the heterogeneity of the observed target, under keeping the relevance to the water. Another direction for a use of SMOS data is catching a correspondence of unsaturated water resources from SMOS to the results available



from public interpretations of the ground water resources, assessed on the base of gravitational missions.

Acknowledgements. The project was conducted in a part, in the frames of the program PECS realized between ESA and partners from Poland, and in another ground based part by national grants N305 046 31/1707, and N305 107 32/3865. The EO data from ENVISTA ASAR, and SMOS data were provided by the ESA EOPI Cat-1 project SMOS Cal/ Val 3275.

References

5

10

25

- Dobson M. C., Ulaby F. T., Hallikainen M. T. and El-Rayes M. A.: Microwave Dielectric behavior of wet soil part II: Dielectric mixing models, IEEE Trans. Geos. Remot. Sens, 23(1), 35–46, 1985.
- Mo, T., Choudhury, B. J., and Schmugge, T. J.: A Parameterization of Effective Soil Temperature for Microwave Emission, J. Geophys. Res., 87, 1301–1304, 1981.
- Jackson T. J. and Schmugge T. J.: Vegetation effects on the microwave emission of soils, Remote Sensing of Environment, 36: 203–212, 1991.
- Holmes, T. R. H.: SMOS Community Microwave Emission Model (SMOS-CMEM), Final Report for the NWP-SAF Steering Group, 2006.
 - Wigneron, J. P., Laguerre L., and Kerr Y. H.: A Simple Parmeterization of the L-band Microwave Emission from Rough Agricultural Soils, IEEE Trans. Geos. Remot. Sens, 39, 1697–1707, 2001.
- ²⁰ Wigneron, J. P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M. J., Richaume, P., Ferrazzoli, P., de Rosnay, P., et al.: L-Band M-ve Emission of the Biosphere (L-MEB) model: Results from calibration against experimental data sets over crop fields, IEEE Transactions on Geosciences and Remote Sensing, 107, 639–655, 2007.

ATBD: SMOS level 2 processor Soil moisture, SO-TN-ESL-SM-GS-0001, Issue 3.d, Date 13/06/2010.

- CMEM: ECMWF Community Microwave Emission Model, http://www.ecmwf.int/research/ESA_ projects/SMOS/cmem/cmem_index.html, 2010.
- de Rosnay, P., Drusch, M.: The Community Microwave Emission Model, ECMWF, April 2008, ver. 1.2, 1.3., 2008.

HESSD 7, 7007–7057, 2010	
SMOS Cal/Val for wetlands	
W. Marczewski et al.	
Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
14	۶I
•	•
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	

Discussion Pape

Discussion Paper

Discussion Paper

Discussion Paper

- ECOCLIMAP: http://www.cnrm.meteo.fr/gmme/PROJETS/ECOCLIMAP/page_ecoclimap.htm, last access Aug. 2010, contains corrected database from September 2006.
- Jackson, T., Cosh, M.; Bindlish, R.: SMOS Soil Moisture Validation with Dense and Sparse Networks, ESA Living Planet Symp., Bergen 2010, 098-D3, http://www.congrex.nl/10a04/, last access Aug. 2010

5

- BEAM: software BEAM v.4.8 http://www.brockmann-consult.de/cms/web/beam/software, last access Aug. 2010.
- NEST: Next ESA SAR Toolbox http://earth.esa.int/resources/softwaretools/, last access Aug. 2010.
- ¹⁰ Usowicz, B.: A method for the estimation of thermal properties of soil, Int. Agrophysics, 7(1), 27–34, 1993.
 - Usowicz, B.: Statystyczno-fizyczne modele przeplywu masy i energii w osrodku porowatym, (Statistical-physical models for mass and energy transfer in porous media), Acta Agrophysica, IA PAN, Lublin, 2000, (in polish) also available also in English from Usowicz@demeter.
- ¹⁵ ipan.lub.pl or http://users.ipan.lublin.pl/~usowicz/pdf/0275.pdf, last access Aug. 2010.
 Polynov, B. B.: Landscape and Soil, Edited by Priroda (Nature),1–3, Minsk, 1925 (in russian).
 Milne, G.: Composite Units for the Mapping of Complex Soil Association, Trans. of the Third International Congress of Soil Sciences, Oxford, London, 1935, Vol. 1.

Romanova, T. A., Pisezsky, G. A., Puchkarjova, T. N., Kapilevich, G. A.:Mathematical Model-

ing of Dynamic of humidity of Soils combinations in Belorussian Polesie, in: Application of Mathematical Methods and electronic computer in Soil Research, 110–113, Minsk, 1976 (in russian).

Romanova, T. A., Kapilevich, G.A.: Water regime as constituent of genetic characteristic of Soils, Paedology, No 12, 10–18, 1981 (in russian).

²⁵ GRACE, TELLUS: http://geoid.colorado.edu/grace/grace.php, last access: August 2010.





Fig. 1. SMOS Cal-Val test sites in Poland (white pins), and their surrounding, on the data background SM_OPER_MIR_BWLF1C_20100708T033042_20100708T042441_344_004_1.HDR data TB_H [K], in ascending half-orbits. Two test places Przylep Stary and Szymbark are not represented in SMOS data yet. New agreed or recently installed test places are shown by silver dots.





Fig. 2. SM demonstration L2 data obtained by SMOS for the test areas Polesie/Podlasie.





Fig. 3. SM demonstration L2 data obtained by SMOS for the test area Wikrowo. Large gaps of no data are deserved by RFI glints. SM values for the sea area are false, and for the coastal line are affected by biasing from the sea. Both effects may be constrained in the retrieval.





Fig. 4. Modeled TB values (ftb_toa_P(x, Re(eps), Im(eps)))) for top of the atmosphere (top), in P polarization components *H*, *V*, versus the incident angle theta (*x*), and parametrized by the dielectric constant values Re(eps) varying in the range (1-21), and Im(eps) kept constant 0.10. Arguments of ftb_toa_P(x, 1.000, 0.1000) represent: the incident angle theta, Re(eps) and Im(eps), respectively.





Fig. 5. Best modeled values of TB, in H, V components versus the incident angles, according to the data SM_OPEB_MIR_SCLF1C_20100714T031325_20100714T035109_344_072_1.HDR, for the wetland Biebrza. Retrieved values with standard errors for soil moisture (sm), optical thickness in NADIR (tau_NADIR) and real part of dielectric constant (epsr) are presented.





Fig. 6. Best modeled values of TB, in H, V components versus the incident angles, according to the data SM_OPEB_MIR_SCLF1C_20100714T031325_20100714T035109_344_070_1.HDR, for the wetland Polesie. Retrieved values with standard errors for soil moisture (sm), optical thickness in NADIR (tau_NADIR) and real part of dielectric constant (epsr) are presented.

















Fig. 9. Soil compounds of SAND according to the ground based data (1200 samples, IA PAS, Lublin), 1 km resolution. Values outside the country of Poland are more or less false extensions.





Fig. 10. A general configuration of the Usowicz model and corresponding representation by analogue electrical R and C networks, given by Usowicz (2000).











Fig. 12. Available types of the dielectric constant characteristics plotted versus the water content, for different numbers of *u*, given by Usowicz (2000).

Interactive Discussion



Fig. 13. The dielectric constant of soil as a function of volumetric water content obtained form different models (Malicki, Friedman, Topp and statistical) for two values of porosity: $\phi = 0.3$ and 0.6 [m³ m⁻³], given by Usowicz (2000).











Fig. 15. ENVISAT ASAR image, of the same area as on Fig. 14., (in the black frame), interpreted on 16 SM classes in the area around the test station (black cross), at Polesie, Kulczyn. Green classes mean forested areas.





Fig. 16. The area of the TBR 1 Bialowieza, and TBR 2 West Polesie, with few representative catena transects (in yellow) for monitoring the water content field and its temporal dynamics.







Fig. 17. Temporary monthly anomalies of the water content over Europe, for the selected epochs 2008.63, (data CSR RL04DS), taken from GRACE data [GRACE (2009)].



Discussion Paper **HESSD** 7, 7007-7057, 2010 SMOS Cal/Val for wetlands **Discussion** Paper W. Marczewski et al. **Title Page** Abstract Introduction Conclusions References Figures **Discussion** Paper **Tables** 14 Back Close Full Screen / Esc **Discussion** Paper **Printer-friendly Version** Interactive Discussion

Fig. 18. Trends of monthly anomalies of water con-tent over the selected great river catchments in Europe for: Oder, Vistula, Dnipr, Nimen, and Dvina, showing different trends, in the period from 2003 to 2009, taken from GRACE data [GRACE (2009)].