# THE RETRIEVAL OF SOIL MOISTURE FROM ENVISAT/ASAR DATA

*Simonetta Paloscia, Paolo Pampaloni, Simone Pettinato, Pietro Poggi and Emanuele Santi*

Institute of Applied Physics 'Carrara', CNR-IFAc, 50127 Firenze, Italy; s.paloscia(at)ifac.cnr.it

**ABSTRACT**

In order to verify the possibility of estimating soil parameters by means of SAR (Synthetic Aperture Radar) images at C-band, an experimental campaign was carried out in two agricultural areas in Italy (Alessandria and Montespertoli) during ENVISAT/ASAR passages and the results obtained after a preliminary data analysis are described in this paper. First of all, a classification of the Alessandria area, obtained by combining HH and HV polarizations of a SAR C-band image collected on 7 November 2003, was performed with the identification of five surface types. A preliminary investigation on the sensitivity of the backscattering coefficient $\sigma^o$ in HH polarization to the ground measurements of soil moisture was carried out and confirmed a certain correlation between these two parameters. Finally, a statistical algorithm for the retrieval of soil moisture has been tested on the Montespertoli area. This method, based on the Bayes theorem, made it possible to retrieve five classes of soil moisture content, with a mean error of less than 10%.

**Keywords:** Soil moisture, radar backscattering coefficient, ENVISAT ASAR

## INTRODUCTION

Soil moisture plays a critical role in the surface energy balance at the soil-atmosphere interface and is a key state variable that influences the redistribution of the radiant energy and the runoff generation and percolation of water in soil. We know that local measurements of soil moisture content (*SMC*) are strongly affected by spatial variability, besides being time-consuming and expensive. Moreover, the use of hydrological models for extending the forecast of soil moisture over larger areas is not easy, and depends on the homogeneity of the selected areas and the information available about them (soil properties, i.e. hydraulic characteristics, and permeability, together with meteorological and climatological data, etc.). The possibility of measuring soil moisture on a large scale from satellite sensors, with complete and frequent coverage of the Earth's surface is therefore extremely attractive.

The sensitivity to *SMC* of the radar backscattering coefficient $\sigma°$, measured at low microwave frequencies is a well-known phenomenon, already investigated by many scientists. Indeed, research activities carried out worldwide in the past have demonstrated that sensors operating in the low-frequency portion of the microwave electromagnetic spectrum (P- to L-band) are able to measure the moisture of a soil layer, the depth of which depends on soil characteristics and moisture profile, and is of the order of some tenths of the wavelength. The most significant information was obtained by combining different frequencies, polarizations, and incidence angles (1,2,3). Unfortunately, P- and L-bands are not still available from current satellite sensors, which, moreover, operate in a single-frequency band. Thus, in this paper, the research for the retrieval of soil moisture has been focused on the potentials of C-band, which is operational on ERS-2, RADARSAT, and ENVISAT satellites. The radar signal at C-band is still sensitive to *SMC*, but it is significantly influenced by vegetation and surface roughness, so that the estimation of spatial variations of moisture with the accuracy requested in many applications is still rather problematic, and needs the use of correcting procedures.

In this paper we present a statistical algorithm based on the Bayes theorem for retrieving *SMC* from the images of the ENVISAT Advanced Synthetic Aperture Radar (ASAR), collected on two agricultural areas in Italy: Alessandria, in the North-west, and Montespertoli, in Central Italy (see Figure 1). Reference values of the backscattering coefficient $\sigma^o$, required by the Bayesian approach, have been generated through the Integral Equation Model (IEM) by Fung (4) and the results compared with ground data collected on both test areas during the ENVISAT acquisitions, showing a general agreement between measurements and estimates of the soil moisture content.

*Figure 1: The two test areas selected for the experiment (Landsat image, copyright Eurimage, courtesy of Planetek Italia).*

**DESCRIPTION OF THE EXPERIMENT**

Two agricultural areas have been selected in Italy for performing the experiments: Alessandria in Northern Italy and Montespertoli in Tuscany. Three ground campaigns were carried out in 2003: one in the Alessandria area in November 6-7, 2003 and two in Montespertoli area on June 11 and November 20, 2003. In Alessandria, ground measurements were carried out in two areas along the Scrivia river: Castelnuovo Scrivia, at the confluence between the Scrivia and Po rivers, and the Borbero sub-basin, a small affluent of Scrivia, close to the border of the Liguria region. Ground measurements included: soil moisture, by means of a Time Domain Reflectometry (TDR) probe and gravimetric samples, collected as a reference calibration of TDR; roughness measurements by means of a needle profilometer and some vegetation parameters (plant height, density, leaf number).

Some photos of the fields were also gathered. Crops present in the areas were: wheat, fodder crops, alfalfa, corn and sugarbeet. *SMC* values were rather high in November 2003 (higher than 35%). A series of ENVISAT/ASAR images were acquired from these areas (see Table 1).

As an example, Figure 2 shows an ENVISAT composite image collected in November of the Alessandria area, with HH/HV polarizations. From the figure, a preliminary land classification is feasible: black zones point out the presence of water (rivers, lakes), light yellow pixels correspond to urban areas, green to forest or dense vegetation (mainly present, in fact, along the rivers), brown to bare, smooth fields, and red to bare, rough fields.

*Table 1: List of the ENVISAT/ASAR images collected on the Italian test areas (IMG = Image Mode Geocoded Ellipsoid Image; IMP = Image Mode Precision Image; APP = Alternating Polarization Mode Precision Image).*

| Test sites | Dates | Product | Polarization |
|---|---|---|---|
| ALESSANDRIA | December 8, 2002 | IMG | VV |
| | July 6, 2003 | IMP | VV |
| | August 29, 2003 | APP | HH/VV |
| | November 7, 2003 | APP | HH/HV |
| MONTESPERTOLI | June 11, 2003 | APP | HH/HV |
| | November 20, 2003 | APP | HH/HV |

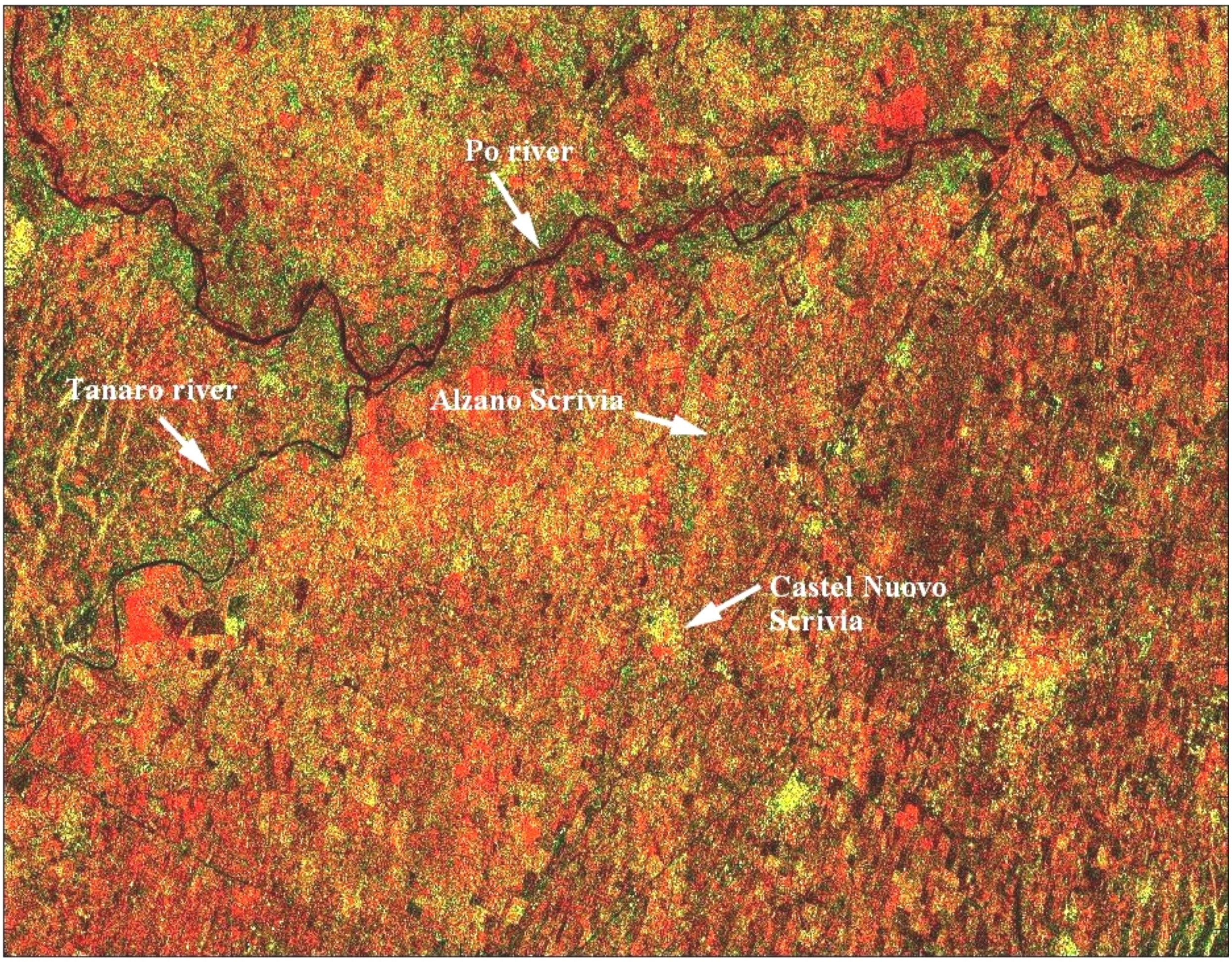


*Figure 2: ENVISAT/ASAR composite RG image in APP (HH, HV polarization) acquired on November 2003 of the Alessandria area. R = HH polarization, G = HV polarization. Colours roughly correspond to different surface types: yellow = urban areas; black = water bodies; red = rough bare soils; brown = smooth bare soils; green = dense vegetation, forests.*

## SOIL MOISTURE RETRIEVAL

In order to compare $\sigma°$ with data collected on the ground, ASAR data have been geocoded with a regional map of the site (scale 1:10000), so that the correct areas, where the ground measurements were carried out, have been selected with the pixel precision.

Direct correlation between $\sigma°$ in HH polarization and *SMC* is shown in Figure 3. Data refer to the Alessandria area in November 2003 when most fields were bare soils. The spread of data is significant, as it can be observed from the value of correlation coefficient ($r$= 0.7), probably due to the surface roughness of bare soils; most of them were in fact ploughed. The result is in any case comparable to those obtained in the past with similar data sets (e.g. (3,5). Another reason for the rather scant correlation between $\sigma°$ and the *SMC* may be attributed to the different sampling principles of TDR and radar. TDR integrates over a soil layer of several centimetres, whereas the radar supposedly investigates only a soil surface layer of a few centimetres. Moreover, in November 2003, the soil surface was dryer than the underlying layers, due to the evaporation of the bare worked surfaces.

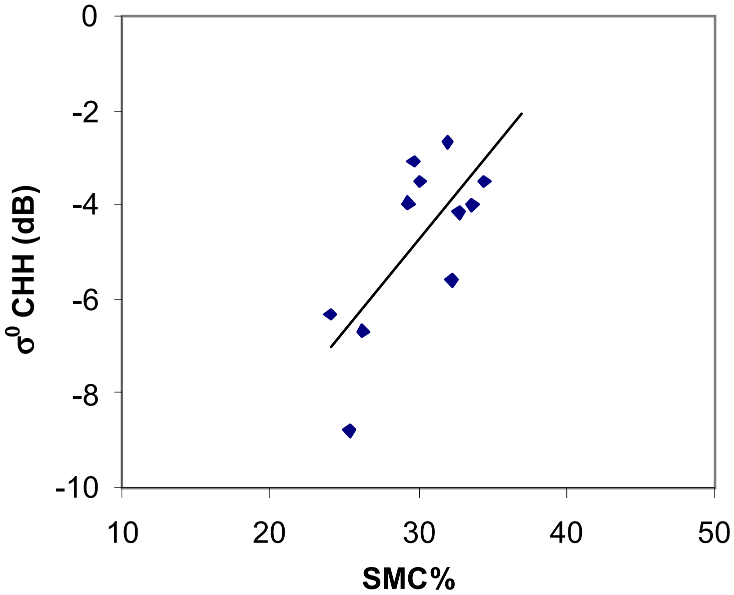


*Figure 3: $\sigma^0$ (C band, HH polarization) as a function of the soil moisture content SMC. The regression equation is: $\sigma°$=0.38* SMC *- 16.2, with a correlation coefficient* r*=0.71.*

## BAYES APPROACH

The Bayesian approach was applied to the retrieval of soil parameters from $\sigma°$ at C band in HH and VV polarization. In particular, the attention was focused on the retrieval of the dielectric constant of soil from which the moisture content can be estimated by means of the Dobson *et al.* model (6). Applying Bayes theorem, the conditional density function $P(\varepsilon,s,lc\,|\,\sigma^o_{\mathrm{VV}},\sigma^o_{\mathrm{HH}})$, which represents the probability to have these values of ground parameters once given the measured $\sigma°$, can be expressed as:

$$P(\varepsilon,s,lc\,|\,\sigma^o_{\mathrm{VV}},\sigma^o_{\mathrm{HH}}) = \frac{P_{prior}(\varepsilon,s,lc)\; P_{post}(\sigma^o_{VV},\sigma^o_{HH}\;\varepsilon,s,lc)}{P(\sigma^o_{VV},\sigma^o_{HH})} \quad , \qquad (1)$$

where: $\varepsilon$ is the dielectric constant of the soil,
$s$ is the standard deviation of the soil heights, i.e. a surface roughness index,
$lc$ is the soil correlation length,
$P(\sigma^o_{VV},\sigma^o_{HH})$ is the joint density of $\sigma^o_{VV}$ and $\sigma^o_{HH}$.

$P_{prior}(\varepsilon,s,lc)$ includes all the *a priori* information about these parameters, such as estimates based on data from other instruments. In cases where no estimate is available, it is possible to apply the 'principle of indifference', by associating each of these parameters a uniform density function in the interval of possible values.

$P_{post}(\sigma^o_{VV},\sigma^o_{HH}\ \varepsilon,s,lc)$ represents the probability to measure $\sigma^o_{VV}$ and $\sigma^o_{HH}$ once given $\varepsilon$, $s$, $lc$. This function can be expressed in a more convenient form considering the radar measurements as theoretical values affected by a multiplicative noise, modelled by means of two random variables independent of the soil parameters $\varepsilon$, $s$ and $lc$, as follows:

$$\sigma^o_{VV,meas} = R_1\sigma^o_{VV,theo} \tag{2a}$$

$$\sigma^o_{HH,meas} = R_2\sigma^o_{HH,theo}\ , \tag{2b}$$

where $\sigma^o_{VV,meas}$ and $\sigma^o_{HH,meas}$ are the backscattering coefficients retrieved by the SAR sensor; $\sigma^o_{VV,theo}$ and $\sigma^o_{HH,theo}$ are theoretical backscattering coefficients computed by the IEM model and $R_1$, $R_2$ are two random variables representing the measurement noise, for both VV and HH polarizations. With this assumption, $P_{post}(\sigma^o_{VV},\sigma^o_{HH}\ \varepsilon,s,lc)$ can be expressed as:

$$P_{post}(\sigma^o_{VV},\sigma^o_{HH} \mid \varepsilon,s,lc) = \frac{1}{\sigma^o_{VV}\cdot\sigma^o_{HH}}P(R_1,R_2) \tag{3}$$

Having defined the probability density function of the errors $P(R_1,R_2)$, the algorithm is able to generate the “optimal estimator” for the dielectric constant of soil (7,8,9), giving as input the measured values of the $\sigma^o$ in *H* and *V* polarization:

$$\bar{\varepsilon} = \frac{\iiint\limits_{\varepsilon,s,lc} \varepsilon P_{prior}(\varepsilon,s,lc)\frac{1}{\sigma^o_{VV}\cdot\sigma^o_{HH}}P(R_1,R_2)\,\mathrm{d}\varepsilon\,\mathrm{d}s\,\mathrm{d}lc}{P(\sigma^o_{VV},\sigma^o_{HH})}\ , \tag{4}$$

where $\bar{\varepsilon}$ is the estimated dielectric constant.

Tests of the algorithm were carried out by means of a dataset of backscattering measurements collected over the Montespertoli test area during the MAC 91 and SIR-C/X-SAR missions in 1991 and 1994. In order to use the algorithm for the ENVISAT/ASAR sensor, only backscattering coefficients at C band in *HH* and *VV* polarization and at low incidence angle of 25° were considered (3). To provide the theoretical backscattering used as reference in Eqs. (2a,b), the IEM model by Fung (4) was used. This model is able to obtain backscattering coefficients for a wide range of natural soil surfaces, where the validity range is given by the condition $Ks$ <3, where $K=2\pi/\lambda$ is the wave number, $\lambda$ is the radar wavelength in centimetres and $s$ is the standard deviation in centimetres defined above. As input to the model, ground measurements of *SMC* and surface roughness (both $s$, height standard deviation of the surface, and $lc$, correlation length), simultaneously with the radar acquisitions, were used, together with additional information such as incidence angle and frequency.

To define the probability density function of the errors $P(R_1,R_2)$, a sub-dataset of 40 elements was extracted and used as reference. For each SAR measurement of this dataset, corresponding ground measurements were used as input of IEM to obtain theoretical backscattering values used in Eq. (4) to calculate mean and variance of $R_1$ and $R_2$. The shape of the PDF was assumed Gaussian, because this function fits acceptably the considered sub-dataset.

The other *a priori* information required by the method is the probability density function of the soil parameters $\varepsilon$, $s$, and $lc$. According to the values of the sub-dataset, $P_{prior}(\varepsilon,s,lc)$ was assumed uniform over the range $4 < \varepsilon < 20$, 0.5 cm< $s$ < 1.2 cm and 2 cm < $lc$ < 8 cm. Thus, the algorithm is able to generate the “optimal estimator” for the dielectric constant of soil, using as input the measured $\sigma^o$ in HH and VV polarization. From $\varepsilon$ estimates, the SMC values are obtained by inverting the Dobson *et al.* model (6).

Having characterized the statistical distribution of the errors $R_1$ and $R_2$, preliminary tests were carried out on the remaining part of the dataset, composed by several tenth of measurements. Due to the complexity of the equation which includes a triple integration operation, the computing time required by this method is of about 1 second for each point (pixel), making this algorithm unusable for real time prediction from wide SAR images. Therefore, a first simplification of the algorithm was carried out, by substituting the probability density function of the correlation length with fixed values of *lc* in the interval 1 cm < *lc* < 8 cm.

Due to the small size, in statistical terms, of the dataset considered, the results were not sufficient t as an evaluation the suitability of the algorithm; however, they show that *lc* is a key factor for the convergence of the method, since the error introduced from a bad estimate strongly affects the retrieval. Best results were obtained by imposing *lc* = 4 cm, a value close to the average of measured *lc* over the whole sub-dataset.

In order to obtain a statistically significant dataset, and to better evaluate the Bayesian algorithm, the Montespertoli data were increased adding 3000 simulated values of backscattering coefficient with the same statistical characterization of the ground measurements. Simulated values of backscattering were obtained by means of Eq. (2), by multiplying the outputs of IEM for random variables having the same distribution, mean and variance of the $R_1$ and $R_2$ noise functions above described and derived from the sub-dataset of 40 elements.

These simulated values were added to the data and used as input for testing the Bayesian algorithm. Although with a large spread of data, the retrieved SMC values followed the same trend of ground measurements, as shown in Figure 4. The obtained regression equation is: *SMCe*=1.33 *SMCm* - 6.68, with a correlation coefficient *r* = 0.78.

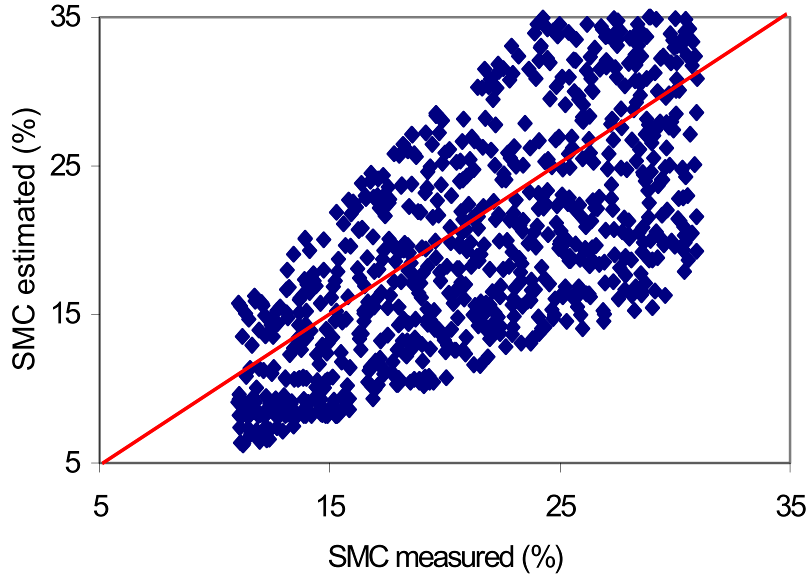


*Figure 4: Comparison between the soil moisture values estimated with the Bayesian algorithm and the soil moisture measured on ground.*

In order to reduce the spread, *SMC* values were grouped in five classes: for each class, averaged values of measurements and corresponding estimates have been calculated and compared. The result is shown in the histogram of Figure 5 (10). We can note that all the classes were correctly retrieved, although the method slightly overestimates the lower values of *SMC* (< 10%) and underestimates the higher ones: however, the mean values of SMC for each class were retrieved with an error ranging between 1 and 10%.

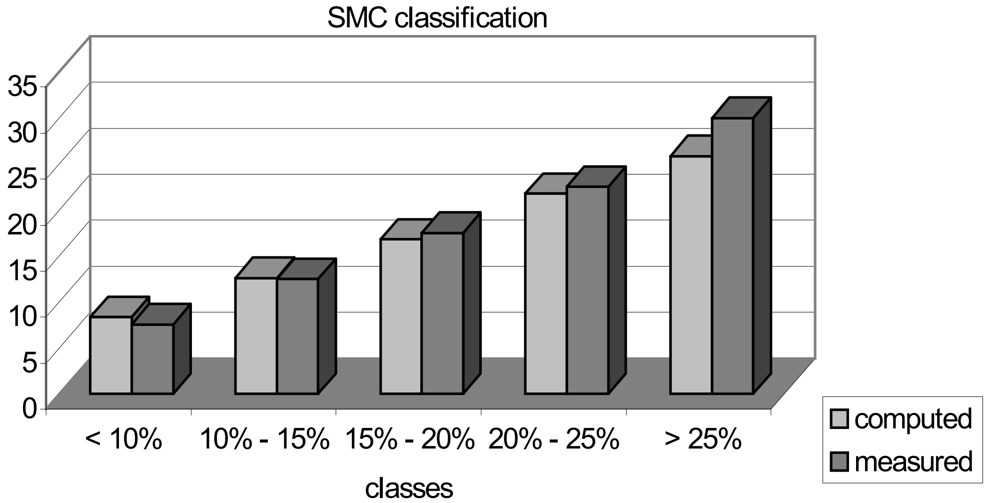


*Figure 5: Comparison between five classes of the soil moisture measured on ground and computed with the Bayes algorithm (10).*

## CONCLUSIONS

The analysis of SAR data collected on both Alessandria and Montespertoli areas, in Italy, confirmed a rather good sensitivity of C-band radar backscattering coefficient both to the surface characteristics and to soil moisture content. The direct comparison between $\sigma^o$ in HH polarization, and the on-ground measured soil moisture showed a general agreement between the two parameters, although the obtained correlation coefficient is not very high.

The use of a statistical algorithm, based on the Bayes theorem, in spite of a large spread in the data, made it possible to retrieve five classes of soil moisture from SAR images, with a mean error of less than 10%

## ACKNOWLEDGEMENTS

This work was partially supported by the EC project FloodMan (EVK1-2001-00237), the Italian Space Agency (ASI), and the ESA ENVISAT programme.

## REFERENCES

1 Dubois P C, J J van Zyl & T Engman, 1995. Measuring Soil Moisture with Imaging Radars. I.E.E.E. Transactions on Geoscience and Remote Sensing, 33, 896-904

2 Benallegue M, O Taconet, D Vidal-Madjar & Normand, 1995. The use of radar backscattering signals for measuring soil moisture and surface roughness. Remote Sensing of Environment, 53, 61-68

3 Macelloni G, S Paloscia, P Pampaloni, S Sigismondi, P de Matthæis, P Ferrazzoli, G Schiavon & D Solimini, 1999. The SIR-C/X-SAR experiment on Montespertoli: sensitivity to hydrological parameters. International Journal of Remote Sensing, 20(13), 2597-2612

4 Fung A K, 1994. Microwave Scattering and Emission Model and their Application (Artec House, Norwood, MA) 573 pp.: chapter 6

5 Macelloni G, S Paloscia, P Pampaloni, R Ruisi, M Dechambre, R Valentin, A Chanzy & J-P Wigneron ,2002. Active and passive microwave measurements for the characterization of soils and crops. Agronomie, 22, 581-586

6 Dobson M C, F T Ulaby, M Hallikainen & M El Rayes, 1985. Microwave dielectric behaviour of wet soil – part II: 4 component dielectric mixing models. IEEE Transaction on Geoscience and Remote Sensing, GE-23, 35-46

7 Haddad Z S, P C Dubois & J J van Zyl, 1996. Bayesian estimation of soil parameters from radar backscatter data. IEEE Transaction on Geoscience and Remote Sensing, 34(1), 76-82

8 Notarnicola C & F Posa, 2001. Bayesian fusion of active and passive microwave data for estimating bare soil water content. Proc. of International Geoscience and Remote Sensing Symposium IGARSS 2001, Vol 3, 1167-1169

9 Notarnicola C & F Posa, 2002. Extraction of soil parameters: two case studies using Bayesian fusion of multiple sources data. Geoscience and Remote Sensing Symposium IGARSS 2002, Vol 2, 905-907

10 Santi E, 2004. Soil moisture estimates from SAR data by using a Bayesian approach. Internal Report IFAC-CNR, RR/OST/01/04, ISSN 1120-2823, 1-8