Soil moisture retrieval through a merging of multi-temporal L-band SAR data and hydrologic modelling

F. Mattia¹, G. Satalino¹, V. R. N. Pauwels², and A. Loew³

¹Consiglio Nazionale delle Ricerche, Istituto di Studi sui Sistemi Intelligenti per l’Automazione (ISSIA), Bari, Italy
²Ghent University, Laboratory of Hydrology and Water Management (LHWM), Ghent, Belgium
³Max-Planck-Institute for Meteorology, The Land in the Earth System, Hamburg, Germany

Received: 1 October 2008 – Published in Hydrol. Earth Syst. Sci. Discuss.: 3 December 2008
Revised: 26 February 2009 – Accepted: 27 February 2009 – Published: 13 March 2009

Abstract. The objective of the study is to investigate the potential of retrieving superficial soil moisture content ($m_v$) from multi-temporal L-band synthetic aperture radar (SAR) data and hydrologic modelling. The study focuses on assessing the performances of an L-band SAR retrieval algorithm intended for agricultural areas and for watershed spatial scales (e.g. from 100 to 10 000 km²). The algorithm transforms temporal series of L-band SAR data into soil moisture contents by using a constrained minimization technique integrating a priori information on soil parameters. The rationale of the approach consists of exploiting soil moisture predictions, obtained at coarse spatial resolution (e.g. 15–30 km²) by point scale hydrologic models (or by simplified estimators), as a priori information for the SAR retrieval algorithm that provides soil moisture maps at high spatial resolution (e.g. 0.01 km²) by point scale hydrologic models (or by simplified estimators), as a priori information for the SAR retrieval algorithm that provides soil moisture maps at high spatial resolution (e.g. 0.01 km²). In the present form, the retrieval algorithm applies to cereal fields and has been assessed on simulated and experimental data. The latter were acquired by the airborne E-SAR system during the AgriSAR campaign carried out over the Demmin site (Northern Germany) in 2006. Results indicate that the retrieval algorithm always improves the a priori information on soil moisture content though the improvement may be marginal when the accuracy of prior $m_v$ estimates is better than 5%.

1 Introduction

The monitoring of the spatial and temporal distribution of soil moisture content ($m_v$) is of major importance for a better understanding of the water cycle on land surfaces with an impact on several applications ranging from drought and flood prediction (e.g. Hong and Kainay, 1996; Pauwels et al., 2002) to meteorology (Betts et al., 1996) and agriculture (Bastiaanssen et al., 2005). Due to the high sensitivity to soil moisture content (e.g. Du et al., 2000), microwave remote sensing holds a great deal of potential for the retrieval of $m_v$. In fact, considerable progresses have been made on the use of microwave remote sensing systems (e.g. Jackson et al., 1993; Kerr et al., 2001; Njoku et al., 2003; Wagner et al., 2003; Shi et al., 2006; Loew, 2008) to measure superficial soil moisture content at coarse spatial resolution (e.g. 15–50 km²). On the contrary, the use of SAR data for the retrieval of soil moisture maps at high spatial resolution (e.g. 2.5–90 × 10⁻³ km²) has been generally limited up to date and no operational algorithm is yet available, while numerous research approaches exist (for a review see Moran et al., 2004). An important part of the limitations to monitor superficial soil moisture contents by means of SAR systems, is due to the fact that the observed backscatter significantly depends not only on soil roughness, soil moisture and plant water content but also on crop structure. As a consequence, there generally exist many combinations of surface parameters mapping the same SAR observable, so the retrieved “optimal” solution (e.g. most probable or minimum root mean square (rms) error) may be characterized by poor accuracy (Satalino et al., 2002). This problem can be tackled by introducing a priori information about the surface parameters and using multi-temporal SAR data (Mattia et al., 2006).

In this context, the objective of this paper is to assess an algorithm for the retrieval, at high spatial resolution, of superficial soil moisture content underlying agricultural crops from multi-temporal L-band SAR data and hydrologic...
modelling. The higher penetration of L-band SAR signal into the canopy, with respect to shorter wavelengths such as C- or X-bands, reduces the sensitivity to vegetation constituents and is expected to improve the SAR capability to monitor soil moisture content. In particular, for cereal crops it is possible to disregard the interaction between L-band SAR signal and crop canopy, at least at HH polarization (Mattia et al., 2007). For this reason, the presented algorithm focuses on soil moisture retrieval of cereal fields.

The rationale of the approach consists of exploiting soil moisture predictions, obtained at coarse spatial resolution by point scale hydrologic models (or by simplified \( m_v \) estimators), as a priori information for the SAR retrieval algorithm. An important aspect for the study is also to obtain indications about the errors affecting the modelling of prior soil moisture predictions. The latter may arise from several factors including incorrect meteorological forcing and model approximations. For this reason, more than one source of meteorological data and two hydrologic models, namely the TOPMODEL-based land-atmosphere transfer scheme (TOPLATS) (Famiglietti and Wood, 1994) and the Process Oriented Multiscale EvapoTranspiration model (PROMET) (Mauser and Schdlich, 1998; Mauser and Bach, 2008), have been employed. In addition, in order to assess the potential of simplified empirical approaches as proxy of soil moisture predictions the use of the Antecedent Precipitation Index (API) (e.g. Crow and Zhan, 2007) has also been investigated.

The retrieval algorithm has been assessed on multi-temporal L-band SAR data acquired by the German Aerospace Centre (DLR) E-SAR system during the AgriSAR 2006 campaign (Hajnsek et al., 2008). However, the algorithm has been developed with a view to the possible future use of data acquired by the L-band spaceborne Phased Array type L-band Synthetic Aperture Radar (PalSAR) system at the highest repetition time (i.e. default acquisition mode). In this respect, despite the fact that the E-SAR airborne system acquired fully polarimetric L-band SAR data, the presented algorithm exploits only single polarized HH multi-temporal SAR data.

In the next sections, the AgriSAR 2006 data set, the retrieval algorithm and the approach adopted for the modelling of prior soil moisture values are described. Then, the experimental assessment of the retrieval algorithm is discussed and conclusions are summarized.

2 The experimental data set

The ground and SAR data analyzed in this study were collected during the AgriSAR campaign conducted over the Demmin agricultural site, in Mecklenburg-Western Pomerania (Northern Germany), from April to July 2006 (Hajnsek et al., 2008). The campaign was funded by the European Space Agency (ESA), coordinated by DLR and included the participation of 16 European Institutes. The experiment encompassed multi-temporal airborne and spaceborne SAR and optical acquisitions together with extensive in situ measurements of bio-physical parameters. The principal objective of the campaign was to assess the impact of the future ESA Sentinel-1 and -2 missions for land applications and to provide a well documented database to investigate the bio-physical parameter retrieval. In the following sections, a short summary of the data set is reported, more details can be found in (Hajnsek et al., 2008).

2.1 In situ data

The Demmin site is an agricultural area characterized by an average annual rainfall of approximately 489 mm and an average temperature ranging between 18° in July and 1° in January. The study area, extending over approximately 25 km² nearby the Goermin village (53.98 N, 13.25 E), is cultivated mainly with winter wheat, winter barley, maize, winter rape and sugar beet. From 19 April through July 26, in situ measurements of volumetric soil moisture content and fresh biomass were collected, roughly every week, over two winter wheat fields (namely field 230 and 250) and two winter barely fields (namely field 440 and 450), all of which larger than 5 ha. Figure 1 shows a land use map of the study area on which the location of the investigated fields is also identified. In total 44 observations (4 fields \( \times \) 11 dates) have been considered in the analysis. In addition, on field 250 there was a ground station with Time Domain Reflectometry (TDR) probes continuously measuring soil moisture content at five different depths, a Bowen Ratio Energy Balance (BREB) station and a Large Aperture Scintillometer (LAS) (a detailed description of these stations is given in Pauwels et al., 2008). Figure 2 shows the temporal behavior of in situ soil moisture measurements for the above-mentioned four cereal fields and also the continuous TDR observations at 0–9 cm. It is worth emphasizing that the study area is characterized by an almost flat topography (i.e. altitude variations within ±60 m),
due to which the variability of soil moisture content recorded across the site, per each date, is generally limited within 4–6% [m$^3$ m$^{-3}$].

2.2 SAR data

A time series of 11 geocoded and coregistered L-band SAR images acquired, from April to July 2006, by the airborne E-SAR system along the West-East track have been used in the analysis. Data were acquired at incidence angles ranging between $25^\circ$ and $55^\circ$ and processed by DLR (Hajnsek et al., 2008). The range and azimuth spatial resolution of the geocoded products is 2 m and 4.5 m, respectively. The pixel spacing is 2 m $\times$ 2 m and the radiometric accuracy is better than 2 dB (Scheiber et al., 2007).

In order to better understand the extent to which it is possible to disregard the interaction between L-band SAR signal and wheat canopy, an assessment of the sensitivity of L-band backscatter to surface parameters has been carried out.

Figures 3 and 4 show the sensitivity of L-band backscatter to soil moisture content and fresh biomass, respectively. The data refer to the entire experimental campaign and were acquired over field 230. The sensitivity to $m_v$ is better at HH than at VV polarization and better for fairly dry than wet soils. In average, there is an increment of approximately 2 dB at HH polarization per 5 vol. % increments in soil moisture content. However, there is also an important scatter of HH and VV backscatter, which is probably partly due to calibration errors (error bars equal to $\pm 1$ dB) and partly to changes in surface conditions. In particular, Fig. 4 shows that at H polarization there is a negligible interaction with the crop canopy as almost no correlation is found between the HH backscattering and the fresh biomass sampled on field 230. While Fig. 3 shows that the backscatter increases in average by approximately 7 dB when the soil moisture increases from 5 to 35%, a strong increase in the biomass leads to an almost zero increase in the backscatter (the contribution from the
ground corresponds to the lowest biomass value in Fig. 4). In other words, the wheat canopy has only a very minor impact on the HH backscatter, which then has been modelled as simply due to surface scattering. This approximation is in agreement with previous studies (e.g. Touré’ et al., 1994), however future work is required to better assess its validity. Indeed, while numerous recent studies have addressed the scattering of winter wheat at C-band (e.g. Cookmartin, et al., 2000; Marliani et al., 2002; Mattia et al., 2003a; Piccardi et al., 2003; Della Vecchia et al., 2006), relatively little work has been carried out at L-band. The surface scattering model adopted in this study is the Integral Equation Model (IEM) (Fung and Chen, 1992), which is an asymptotic model developed to bridge the gap between the Small Perturbation Method (SPM) model and the Kirchhoff approximation (KA). Ulaby et al. (1982), thus covering a wide range of roughness conditions particularly at L-band. From an electromagnetic point of view, the IEM essentially is a second iteration of the iterative Kirchhoff approximation (Liszka and McCoy, 1982). One drawback of this approach is that the conditions for the convergence of the iterative series are not known a priori. It is worth noting that the IEM model was built to predict both single and multiple scattering contributions to surface scattering. It was expected to predict well both co and cross polarized components over a quite wide range of roughness parameters. However, some of the assumptions made in the IEM development have been subsequently recognized as simplistic by the same original authors (for a critical review of the IEM see Alvarez–Perez, 2001). An improved version of IEM was released in Hsieh et al. (1997), a further version was published in Chen et al. (2000). The expressions of cross-polarised scattering coefficient have been continuously amended until recently (Chen et al., 2003). However, the expressions of co-polarized backscattering coefficient (i.e. single scattering contribution) have not changed with respect to the original IEM. It is for this reason that in this paper the expressions of the original IEM model are used.

3 The Retrieval algorithm

The proposed algorithm transforms a temporal series of L-band SAR data, acquired at HH polarization and low-medium incidence angles (approximatively 20°–40°) over cereal fields, into soil moisture values. According to the above-reported sensitivity analysis, at L-band and HH polarization, there is a reduced sensitivity of backscatter to the fresh biomass of cereal fields. On the contrary, the most important contribution to HH backscatter comes from the soil and its moisture variations. As a consequence, the adopted approach disregards the presence of vegetation and inverts the IEM surface scattering model by using a constrained optimization technique, which integrates a priori information on soil parameters (such as vertical surface roughness and soil moisture content, see later on Sect. 5) to obtain robust and accurate estimates of soil moisture content (Mattia et al., 2006). More precisely, the technique minimizes the following cost function:

\[ C = \frac{1}{N} \sum_{n=1}^{N} \frac{|(\sigma_0)_n - F_M(\theta, \lambda, p_{m=1,M})|^2}{(\Delta^s(\sigma_0)_n)^2} + \frac{1}{M} \sum_{m=1}^{M} \frac{|p_m - \hat{p}_m|^2}{(\Delta^l(p_m))^2} \]  

(1)

where \( N \) is the number of \( \sigma_0 \) observations, \( F(\cdot) \) is the IEM model (depending on the SAR incidence angle and wavelength, i.e. \( \theta \) and \( \lambda \), and on the \( M \) surface parameters \( p_m \), \( \hat{p}_m \) are the a priori estimates of surface parameters, \( \Delta^s(\sigma_0) \) includes the backscatter calibration, statistical and model errors and \( \Delta^l(p) \) is the error affecting the prior estimates of surface parameters. The latter basically consist of the surface height standard deviation (\( s \)), the surface autocorrelation function (ACF), the surface correlation length (\( l \)) and the soil relative dielectric constant (\( \varepsilon' \)), which in turn depends on the soil moisture content and on the soil texture composition. An exponential shape for the ACF has been assumed because past studies (e.g. Mattia et al., 1997, 2003b) showed that this is the shape most often observed in field measurements. To relate the soil dielectric constant to the volumetric soil moisture content, the empirical expression derived by Hallikainen et al. (1985) has been employed. This expression models the soil dielectric constant as a second order polynomial in \( m_v \), which can be analytically inverted. In order to obtain estimates of soil moisture content, the algorithm firstly estimates the soil dielectric constant, and then uses the inverted empirical expression of Hallikainen to derive the soil moisture content. To simplify, it will be assumed that \( M=3 \) and \( (p_{m=1,M})=(s, l, m_v) \). In the implementation of the algorithm the \( N \) \( \sigma_0 \) observations have been obtained by using \( N \) multi-temporal L-band, HH polarization, E-SAR backscattering. The use of multi-temporal data is beneficial for the accuracy of the retrieved soil moisture content under the condition that the surface roughness remains almost constant during the time-span (\( T \)) of the \( N \) acquisitions. For instance, for a temporal series of \( N \) images, disregarding the presence of vegetation, the number of surface parameters to be estimated is \( N+2 \) (\( N \) soil moisture values and 2 surface roughness parameters, namely \( s \) and the correlation length \( l \)). For \( N \) equals to 1 there is the worst ratio (i.e. 1/3) between independent measurements and parameters to be estimated (highly inaccurate retrieval). Whereas for \( N \) large the ratio tends to 1 (highly accurate retrieval). In order to minimize (1) an iterative efficient approach based on the Generalized Reduced Gradient Method (Lasdon et al., 1978) was employed.

3.1 Numerical assessment of the algorithm performances

To characterize the performances of the developed retrieval algorithm a simulation study was carried out. A synthetic
Table 1. Mean values of synthetic ground data

<table>
<thead>
<tr>
<th>Soil parameters</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
</tr>
<tr>
<td>Surface height std (cm)</td>
<td>1.2</td>
</tr>
<tr>
<td>Correlation length (cm)</td>
<td>15.0</td>
</tr>
<tr>
<td>Real part of $\epsilon_r$</td>
<td>6.0</td>
</tr>
</tbody>
</table>

The data set of ground data was built simulating three different acquisition dates (i.e. $N=3$), Table 1 reports the average values of the considered surface parameters. Then, the IEM model was employed to obtain the backscatter values at L-band, HH polarization and $23^\circ$ incidence, associated to the surface parameters of Table 1. In order to simulate the presence of measurement errors (including radiometric, statistical and model errors) a zero-mean Gaussian noise with increasing standard deviation (std) (ranging from 0.5 to 1.5 dB) has been superposed to the IEM predictions. A priori information for the retrieval algorithm have been obtained by perturbing the surface parameters reported in Table 1 with a zero mean Gaussian noise with increasing std (ranging from 10 to 30% of the total variability range of surface parameters). It should be emphasized that the simulated a priori information still represents an ideal unbiased case (the error was at zero mean). Finally the retrieval algorithm has been applied to the synthetic data set and the results have been analyzed. A necessary condition that should be always fulfilled by the algorithm is that the final error, computed as the rms error between retrieved and observed soil moisture values (i.e. $\Delta^f(m_v)$), is smaller than the initial error, computed as the rms error between prior and observed soil moisture values (i.e. $\Delta^i(m_v)$). Of course, the higher the ratio $\Delta^i(m_v)/\Delta^f(m_v)$ (referred to as gain ($G$)), the better the algorithm performs. For this reason, the $G$ parameter has been adopted to synthetically represent the algorithm performances in the numerical study. Figure 5 shows the gain parameter ($G$), obtained by applying the retrieval algorithm over the synthetic data set, versus the initial rms error $\Delta^i(m_v)$ for increasing values of measurement errors (referred to $\Delta^i(\sigma_0)$). Figure 5 shows that the algorithm gain increases with the initial error $\Delta^i(m_v)$ and that lower measurement errors $\Delta^i(\sigma_0)$ coincide with increasing values of $G$. In other words, if the prior information on soil moisture content is already quite good (e.g. better than 5%), the algorithm gain is expected to be marginal (i.e. $G \approx 1$) unless the measurement error is very small (e.g. less or equal to 0.50 dB). On the other hand, for $\Delta^i(m_v)$ approximately equal to 7% and $\Delta^i(\sigma_0)$ equal to 0.75 dB, the expected gain is approximately 1.3, corresponding to a final rms error $\Delta^f(m_v)$ approximately equal to 5%. The above-illustrated characteristics of the algorithm together with the fact that its output provides soil moisture maps at high resolution (e.g. 0.001 km$^2$), prompts the following didactic example on the potential of the method. Let us consider a study area of 25 km$^2$ consisting of three homogeneous sub-areas (of the same size): one fairly wet (e.g. 24%±2%), one medium wet (e.g. 17%±2%) and the third one fairly dry (e.g. 10%±2%). Over this area, the SAR retrieval algorithm is applied using as a priori information an average value of 17% (in this case the rms error between the constant guess and the true soil moisture values of the fairly wet, medium wet and dry areas is approximately 7%, 2% and 7%, respectively). Under these circumstances, the algorithm is expected to retrieve approximately the following mean and rms error values for the three classes: 24%±5%, 17%±2% and 10%±5%. Hence, the three sub-areas with different soil moisture content can be identified and separated (within 1-std). In other words, despite the gain of the retrieval algorithm may often be relatively small (mainly due to the high measurement error budget), still the asset of providing soil moisture maps at high resolution can be regarded as a valuable feature. In the following an experimental assessment of the algorithm performances will be carried out.

4 Modelling of prior soil moisture values

In order to obtain a priori information on soil moisture content, at coarse scale, the TOPLATS and PROMET hydrologic models and the API index have been exploited. The hydrologic models were applied at the point scale. It is thus not possible to assign a spatial scale to the model simulations. The point scale prior information was used for the entire test site. This is a valid assumption, because 1) the test site is relatively small, and meteorological forcing can be assumed to be homogeneous for the entire test site, and 2) the land
cover properties (winter wheat) were very homogeneous for the test site as well. When the methodology is extended for larger domains, however, a grid resolution in the order of 1 km can be recommended.

In the next subsections the three approaches are briefly described, Sect. 4.4 then illustrates a comparison between modelled and observed soil moisture values.

4.1 TOPLATS

The TOPMODEL-based land atmosphere transfer scheme (TOPLATS) model has its foundation in the concept that shallow groundwater gradients set up spatial patterns of soil moisture that influence infiltration and runoff during storm events, and evaporation and drainage between these events. The assumption is made that these gradients can be estimated from local topography, through a soil-topographic index (Sivapalan et al., 1987). From this foundation, the model was expanded to include infiltration and resistance-based evaporation processes, a surface vegetation layer, and a surface energy balance equation with an improved ground heat flux parameterization, and the effect of atmospheric stability on heat fluxes (Famiglietti and Wood, 1994; Peters-Lidard et al., 1997). The model was originally developed to simulate the surface water and energy balance for warm seasons (Famiglietti and Wood, 1994; Peters-Lidard et al., 1997). Afterwards, winter processes (frozen ground and a snow pack), an improved water and energy balance scheme for open water bodies, and a two-layer vegetation parameterization were added (Pauwels and Wood, 1999). For a detailed model description, we refer to Famiglietti and Wood (1994); Peters-Lidard et al. (1997), and Pauwels and Wood (1999). Loaiza Usuga and Pauwels (2008) list an overview of the field experiments and test sites for which the model has been applied, based on which it can be concluded that the model can adequately simulate the partitioning of the energy and mass balances into their different terms.

4.2 PROMET

The physically based land surface model PROMET (Process Oriented Multiscale EvapoTranspiration model) is used in the present study to simulate the surface energy budget and exchange of water and matter within the soil-plant-atmosphere continuum. The model describes the actual evapotranspiration and water balance at different scales, ranging from point scale, to microscale and mesoscale (Mauser and Schädlích, 1998; Mauser and Bach, 2008). The model consists of a kernel which is based on five sub-modules (radiation balance, soil model, vegetation model, aerodynamic model, snow model) to simulate the actual water and energy fluxes and a spatial data modeler, which provides and organizes the spatial input data on the field-, micro and macroscale. The simulations are made on hourly basis.

PROMET solves the surface energy balance in an iterative way. The ground heat flux is estimated using a soil temperature model (Mueth, 2008). Actual evapotranspiration is simulated within PROMET using the Penman-Monteith equation (Monteith, 1965). Canopy surface resistance is simulated as a function of vegetation type using a resistance network approach (Baldocchi et al., 1987), while the soil resistance is estimated based on the approach of Eagleson (1978). A four layer soil model (0-5, 5-20, 20-65, 65-200 cm) is used to calculate soil water fluxes and soil temperature profiles. The change of volumetric soil moisture content, percolation, exfiltration, capillary rise and surface runoff are explicitly considered. The infiltration into the soil layer is described using the model of Philip (1957). The soil water retention model of Brooks and Corey (1964) is used to relate soil moisture content to soil suction head. A detailed description of the model is given by Mauser and Schädlích (1998) and Mauser and Bach (2008). A physical snow model extends PROMET to allow for simulations in cold climates (Strasser and Mauser, 2001).

PROMET simulations are based on GIS information as e.g. soil maps and land use information. Meteorological forcing data might be either provided from station networks as well as from gridded forcing fields. PROMET has been extensively validated in different geographic locations in Central Europe (Upper Rhine Valley – 10\times 10 km², Bavarian Alpine Foreland – 200 \times 100 km², Upper Danube catchment – 76 000 km², Weser catchment – 35 000 km²) using evapotranspiration measurements of micrometeorological stations at the local scale and by comparison to thermal remote sensing informations at the regional scale (Mauser and Schädlích, 1998; Ludwig and Mauser, 2000).

It provides interfaces to integrate remote sensing derived information into the model. It has been used together with optical and microwave remote sensing data to improve land surface simulations. Bach and Mauser (2003) used the model to improve crop yield prediction and surface runoff prediction by combining PROMET results with optical (Landsat-TM) and microwave (ERS) remote sensing data. Schneider (2003) used LANDSAT-TM data to determine vegetation model parameters and improve plant growth simulations. Loew et al. (2007) compared PROMET simulations at different spatial scales with soil moisture information derived from active microwave data (Loew et al., 2006), and found a good agreement between the spatial patterns of observed and simulated soil moisture at multiple scales.

4.3 Antecedent soil moisture simulation

Precipitation information is available on a regular basis from a large number of stations. Simple concepts to derive information on actual soil moisture status, based exclusively on precipitation data, have therefore been developed. One simple approach is based on the concept of the so called Antecedent Precipitation Index (API). As the API is exclusively
based on precipitation data as model input, it has been widely used in rainfall-runoff applications to parameterize the soil moisture conditions in hydrological catchments (e.g. Sittner et al., 1969; Rose, 1998; Descroix et al., 2002; Vries and Hromadka, 1993). The API for day \( i \) is defined as

\[
\text{API}_i = \gamma_i \text{API}_{i-1} + P_i
\]

(2)

where \( P_i \) is the observed precipitation [mm] on day \( i \) and \( \gamma_i \) is the corresponding API recession coefficient at that day which is used to parameterize the loss of water in the soil column due to evapotranspiration, groundwater recharge and lateral soil water fluxes.

Given some information on the antecedent precipitation, one might use the API as a prior proxy for soil moisture conditions on an operational basis as precipitation information is (at least) available in terms of short term forecasts on the global scale. However, large uncertainties in API result from uncertainties in the available precipitation information as well as in the parameterization of the corresponding recession coefficient \( \gamma \). Different approaches to parameterize \( \gamma \) have been proposed. Its value might vary in between 0.7 for dry conditions and 1.0 for wet soil conditions (Crow, 2007a). An exponential decay of the form \( \gamma = e^{-\delta} \) has been proposed, whereas the factor \( \delta \) is the inverse of the characteristic time of soil moisture depletion. Its value might be empirically calibrated or it might be parameterized using additional information like e.g. the ratio of potential evapotranspiration to maximum available soil moisture (Chodhury et al., 1993; Descroix et al., 2002). In the present study we follow the parameterization proposed by Crow (2007a) whereas the variation of \( \gamma \) is defined as

\[
\gamma_i = A + B \cos \left( 2 \pi JD / 365 \right)
\]

(3)

with the parameters \( A = 0.85 \) and \( B = 0.1 \) and \( JD \) the julian day, which is a very simple approach to roughly estimate the seasonal effects of evapotranspiration loss. The model parameters could be calibrated using available in situ soil moisture data. In order to keep the model as general as possible, no calibration of the model is done for the test site in the present study.

The API modelling approach is used in the present study to provide a further prior guess on soil moisture for the SAR based soil moisture retrieval algorithm. Two sets of precipitation data (\( P \)), acquired by two weather stations located approximately 10 km apart, were used to estimate two API series. The first weather station (referred to as Goermin station) was located on the study area (nearby the Goermin village), whereas the second one (referred to as Greisfwald station) was located in the town of Greisfwald. In order to transfer the API values [mm] to volumetric surface soil moisture a linear regression between the TDR measurements, collected at 5 cm depth on field 250, and API was calculated. Table 2 reports the parameters of the linear fit, for the two computed API series, namely the API based on precipitation measured by the Goermin and Greisfwald weather stations.

<table>
<thead>
<tr>
<th>Table 2. TDR Soil moisture values versus API.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weather stations</td>
</tr>
<tr>
<td>Goermin 9.7 0.8 0.60</td>
</tr>
<tr>
<td>Greisfwald 9.6 0.7 0.65</td>
</tr>
</tbody>
</table>

4.4 Comparison of modelled and observed soil moisture values

In a previous study, Pauwels et al. (2008) have thoroughly investigated the water and energy balance for a winter wheat of the Demmin site (i.e. field 250). In particular, a remarkable agreement between the time series of TDR measurements reported in Fig. 1 and TOPLATS and PROMET predictions, i.e. an rms error better than 4%, was found. However, the objective of this section is to assess the extent to which point scale hydrologic model predictions can represent not only the temporal but also the spatial variability of soil moisture content over the Goermin study area. For this reason, the TOPLATS, PROMET and API predictions have been compared to in situ measurements of volumetric soil moisture content (sampled at a soil depth of between 5 and 10 cm) collected over four different cereal fields during the entire AgriSAR 2006 campaign (see Fig. 1). In the analysis two sets of meteorological forcing data, acquired by the weather stations located at the Goermin village and at the town of Greisfwald, were employed. For each of the aforementioned simulated data sets, Table 3 reports a comparison with the time series of \( m_v \) measured in situ. The rms error (\( \Delta^4(m_v) \)), the correlation (\( R \)) and the parameters of a linear fit between observed (i.e. X) and modelled (i.e. Y) soil moisture values are shown. In all but one case, i.e. TOPLATS (Greisfwald), the mean soil moisture values predicted by the models underestimate the observed ones (the bias ranges between 1 and 4%). The effect is more pronounced for simulations based on Goermin than Greisfwald weather data (though, in general, the impact of using meteorological data collected by a station located 10 km a part from the study area seems to be quite limited). The rms error of PROMET and TOPLATS predictions (\( \Delta^4(m_v) \)) is always better than 5%, the \( R \)-values are higher than 0.8 and the slope parameters range between 0.47 and 1.05. API predictions are affected by rms errors larger than 6.0%, while the slope and correlation parameters are lower than 0.3 and 0.55, respectively. Under these circumstances, it is confirmed that API should be regarded as a weak prior proxy for surface soil moisture conditions. Nevertheless, it is worth emphasizing that the API asset is its simplicity and the fact that it requires as input solely precipitation information. On the contrary, SVAT models, such as PROMET and TOPLATS, hold a strong potential to provide quite accurate (i.e. better than 5%) prior estimates of \( m_v \), at least at coarse resolution. However, they require significant
Table 3. Volumetric moisture content ($m_v$): hydrologic model predictions versus in situ measurements

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean $X$ (%)</th>
<th>Mean $Y$ (%)</th>
<th>$A$ (%)</th>
<th>$B$</th>
<th>$\Delta^v(m_v)$ (%)</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROMET (Goermin)</td>
<td>15.5</td>
<td>13.14</td>
<td>4.51</td>
<td>0.59</td>
<td>4.35</td>
<td>0.85</td>
</tr>
<tr>
<td>TOPLATS (Goermin)</td>
<td>15.5</td>
<td>12.82</td>
<td>5.45</td>
<td>0.47</td>
<td>4.99</td>
<td>0.80</td>
</tr>
<tr>
<td>PROMET (Greisfwald)</td>
<td>15.5</td>
<td>14.47</td>
<td>4.09</td>
<td>0.67</td>
<td>4.01</td>
<td>0.81</td>
</tr>
<tr>
<td>TOPLATS (Greisfwald)</td>
<td>15.5</td>
<td>17.36</td>
<td>1.09</td>
<td>1.05</td>
<td>4.5</td>
<td>0.86</td>
</tr>
<tr>
<td>API (Goermin)</td>
<td>15.5</td>
<td>12.68</td>
<td>8.59</td>
<td>0.26</td>
<td>6.27</td>
<td>0.54</td>
</tr>
<tr>
<td>API (Greisfwald)</td>
<td>15.5</td>
<td>12.6</td>
<td>8.05</td>
<td>0.29</td>
<td>6.58</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Table 4. Volumetric moisture content ($m_v$): perturbed hydrologic model predictions versus in situ measurements

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean $X$ (%)</th>
<th>Mean $Y$ (%)</th>
<th>$A$ (%)</th>
<th>$B$</th>
<th>$\Delta^v(m_v)$ (%)</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perturbed PROMET (Goermin)</td>
<td>15.5</td>
<td>8.16</td>
<td>-0.97</td>
<td>0.59</td>
<td>8.22</td>
<td>0.85</td>
</tr>
<tr>
<td>Perturbed TOPLATS (Goermin)</td>
<td>15.5</td>
<td>7.84</td>
<td>0.47</td>
<td>0.47</td>
<td>8.76</td>
<td>0.80</td>
</tr>
</tbody>
</table>

more information on a specific site as model input (e.g. meteorological data, soil and land cover maps, etc.). A drawback of these findings is that, according to the numerical analysis of Sect. 3.1, the $m_v$ prior predictions of PROMET and TOPLATS are too accurate to represent a stringent test-bed for the SAR-retrieval algorithm. For this reason, two further data sets, referred to as perturbed PROMET and perturbed TOPLATS, characterized by a bias and an rms error of approximately 7% and 8%, respectively, have been included in the analysis. These two perturbed data sets have been obtained by subtracting from the $m_v$ predictions of the PROMET and TOPLATS models a constant value of approximately 5% (more precisely 4.98%). This choice was aimed at obtaining two data sets affected by biases and rms errors higher than those obtained by means of API but still characterized by high correlations with the $m_v$ in situ measurements. Table 4 reports information similar to Table 3 but it refers to the data sets obtained by perturbed PROMET and perturbed TOPLATS (based on meteorological data acquired at the Goermin weather station). Furthermore, Fig. 7 shows the scatterplot between the soil moisture values simulated by all the illustrated modelling approaches, namely the PROMET and TOPLATS (based on meteorological data acquired at the Goermin and Greisfwald weather stations), the API (based on meteorological data acquired at the Goermin and Greisfwald weather stations) and the perturbed PROMET and TOPLATS (based on meteorological data acquired at the Goermin weather station), and those measured in situ. Figure 7 shows that in general model predictions tend to cluster around a few discrete values whereas the in situ measurements are evenly distributed. In addition, it is observed that the model underestimation is more important for medium-high than for low $m_v$ values (similar results were found in Pauwels et al., 2009).

5 Experimental assessment of the algorithm performances

The performances of the retrieval algorithm described in Sect. 3 have been assessed on the AgriSAR 2006 data set. Figure 6 shows a flow chart of the implemented algorithm. Ancillary information concerning land cover and soil texture maps as well as the initial guess values for vertical surface roughness ($s$) and soil moisture content ($m_v$) are required. Conversely, no a priori information on the correlation length $\ell$ was used. This is because: 1) it is extremely difficult to

![Flow chart of the implemented SAR retrieval algorithm.](image-url)
Table 5. Volumetric moisture content \( (m_v) \): SAR retrieved values versus in situ measurements

<table>
<thead>
<tr>
<th>Models used to derive the a priori information</th>
<th>SAR retrieved (Y) vs. in situ measurements (X): ( Y = A + BX )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ( X(%) )</td>
</tr>
<tr>
<td>PROMET (Goermin)</td>
<td>15.5</td>
</tr>
<tr>
<td>TOPLATS (Goermin)</td>
<td>15.5</td>
</tr>
<tr>
<td>PROMET (Greisfwald)</td>
<td>15.5</td>
</tr>
<tr>
<td>TOPLATS (Greisfwald)</td>
<td>15.5</td>
</tr>
<tr>
<td>API (Goermin)</td>
<td>15.5</td>
</tr>
<tr>
<td>API (Greisfwald)</td>
<td>15.5</td>
</tr>
<tr>
<td>Perturbed PROMET (Goermin)</td>
<td>15.5</td>
</tr>
<tr>
<td>Perturbed TOPLATS (Goermin)</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Table 5 shows the comparison between SAR-retrieved \( m_v \) values and the in situ measured \( m_v \) values. The rms error \( \Delta^f (m_v) \), the correlation \( R \) and the parameters of a linear fit between observed \( m_v \) and SAR-retrieved \( m_v \) are provided for each model. The results indicate that:

- for the case of non-perturbed initial \( m_v \) guess values, the difference between the mean of observed and retrieved soil moisture values ranges between 0.8 and 2.2%, significantly smaller than the bias reported in Table 3. Whereas, in the case of perturbed PROMET and TOPLATS predictions, the bias reduces from approximately 7% to 1.6% and 4%, respectively;

- the rms error reported in Table 5 (i.e. \( \Delta^f (m_v) \)) is always smaller than the corresponding rms error reported in Table 3 and 4 (i.e. \( \Delta^f (m_v) \)). The best performances, in terms of algorithm gain, are observed when the perturbed PROMET and TOPLATS estimates are used as guess values, in these cases the retrieval algorithm achieves a \( G \) parameter of approximately 1.4;

- the \( R \) coefficient is lower (or equal) than the correspondent values shown in Table 3 and 4, when the \( m_v \) guess values are provided either by TOPLATS or PROMET models. The opposite is true when the prior estimates of \( m_v \) are obtained by means of API estimator;

- non-optimal behaviour of the algorithm is observed in the two cases of API Goermin and API Greisfwald, where the prior estimates were not only biased but also poorly correlated (i.e. \( R < 0.55 \)) with the in situ measurements.

In summary, the experimental analysis substantially confirms the characteristics of the retrieval algorithm as illustrated in Sect. 3.1. Besides, it is worth mentioning that the algorithm showed a strong robustness versus the presence of biases in the prior estimates of \( m_v \). Whereas, its performances were significantly lowered when the prior estimates of \( m_v \) were poorly correlated to the in situ measurements.

6 Conclusions

The investigated retrieval algorithm uses prior information on soil moisture content at coarse spatial scale (e.g. \( 25 \text{ km}^2 \)) in order to transform a temporal series of 3 SAR images, acquired at L-band and HH polarization, into multi-temporal soil moisture maps at high spatial resolution (e.g. \( 0.01 \text{ km}^2 \)). In the present form, the retrieval algorithm applies to bare and cereal fields only and it has been tested for time series of SAR images acquired over a time-span of three weeks. The results of the experimental analysis, conducted over the data set acquired during the AgriSAR 2006 campaign and based on prior estimates of soil moisture content obtained by means of TOPLATS and PROMET hydrologic models and by means of the API estimator, showed that the algorithm has always a gain \( (G) \) greater than 1 thus implying that it always improves the prior information. The best performances, in

www.hydrol-earth-syst-sci.net/13/343/2009/
Fig. 7. Scatterplot between hydrologic model predictions of $m_v$ and in situ measurements. The data reported on the Y-axis are: (a) PROMET predictions (based on Goermin weather station); (b) TOPLATS predictions (based on Goermin weather station); (c) PROMET predictions (based on Greisfwald weather station); (d) TOPLATS predictions (based on Greisfwald weather station); (e) API predictions (based on Goermin weather station); (f) API predictions (based on Greisfwald weather station); (g) perturbed PROMET predictions (based on Goermin weather station); (h) perturbed TOPLATS predictions (based on Goermin weather station).
Fig. 8. Scatterplot between SAR-retrieved values and in situ measurements. The data reported on the Y-axis were obtained using as \( m_v \) guess values: (a) PROMET predictions (based on Goermin weather station); (b) TOPLATS predictions (based on Goermin weather station); (c) PROMET predictions (based on Greisfwald weather station); (d) TOPLATS predictions (based on Greisfwald weather station); (e) API predictions (based on Goermin weather station); (f) API predictions (based on Greisfwald weather station); (g) perturbed PROMET predictions (based on Goermin weather station); (h) perturbed TOPLATS predictions (based on Goermin weather station).

www.hydrol-earth-syst-sci.net/13/343/2009/  
terms of the $G$ parameter, were observed in the case of perturbed PROMET and TOPLATS predictions, for which the prior information was considerably biased but highly correlated ($R \geq 0.8$) with the in situ measurements. In these cases, the algorithm was able to reduce the bias of PROMET and TOPLATS predictions from approximately 7% to less than 2% and 4%, respectively. In addition, the rms error was reduced from approximately 8.2% to 5.6% and from 8.8% to 6.4%. Conversely, when the prior information was not only biased but also poorly correlated with the in situ measurements (as it is the case of prior information provided by the API estimator) the algorithm marginally improved the initial error. In the intermediate cases, when the prior information was highly correlated with in situ measurements and showed a relatively small bias, the algorithm reduced the bias (e.g. from approximately 4% to 2%) and marginally the rms error (e.g. from approximately 5% to 4%). Nevertheless, also in these cases it was argued that the algorithm can be quite useful in identifying areas characterized by significantly different soil moisture content within the swath area (e.g. 25 km²).

In conclusion, it is worth emphasizing that while the proposed algorithm can retrieve quite accurate multi-temporal maps of soil moisture content over agricultural sites from L-band SAR data at single polarization, its main drawback is that it requires not only prior guess values of soil moisture content at coarse scale but also updated information about crop maps (at least in terms of principal crops, e.g. broad leaves vs small stems). Therefore, it can be feasible to systematically retrieve soil moisture maps over agricultural sites, predominantly devoted to cereal cultivation, if different sources of remote sensing data are employed. For instance, a possible scenario to further validate the proposed approach could encompass PolSAR acquisitions in ScanSAR WB1 mode for soil moisture retrieval, C-band ASAR data for crop mapping and data acquired by active/passive microwave spaceborne systems (e.g. ASCAT, AMSR-E, or the MIRAS system on board the satellite platform of the forthcoming Soil Moisture and Ocean Salinity Mission) or forecasts provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) as sources of prior guess values of soil moisture content, at coarse resolution.

Acknowledgements. This work was supported by ESA-ESTEC under contracts no. 19558/06/NL/HE and 19974/06/I/LG. The authors would like to acknowledge the AgriSAR team for providing the in situ data collected during the campaign.

Edited by: W. Wagner

References

F. Mattia et al.: Soil moisture retrieval


