Soil moisture retrieval over irrigated grassland using X-band SAR data

Mohammad El Hajj, Nicolas Baghdadi, Mehrez Zribi, Gilles Belaud, B. Cheviron, Dominique Courault, F. Charron

El Hajj M., Baghdadi N., Zribi M., Belaud G., Cheviron B., Courault D., and Charron F., 2016. Soil moisture retrieval over irrigated grassland using X-band SAR data. *Remote Sensing of Environment, vol. 176, pp. 202-218, doi:* http://dx.doi.org/10.1016/j.rse.2016.01.027.

Soil moisture retrieval over irrigated grassland using X-band SAR data

Mohammad El Hajj [1,]*, Nicolas Baghdadi [1], Mehrez Zribi [2], Gilles Belaud [3], Bruno Cheviron [4], Dominique Courault [5], François Charron [3]

[1] IRSTEA, UMR TETIS, 500 rue François Breton, 34093 Montpellier cedex 5, France

[2] CNRS, CESBIO, 18 av. Edouard Belin, bpi 2801, 31401 Toulouse cedex 9, France

[3] SupAgro, UMR G-EAU, 2 place Pierre Viala, 34060 Montpellier, France

[4] IRSTEA, UMR G-EAU, 361 rue François Breton, 34196 Montpellier cedex 5, France

[5] INRA, UMR 1114 EMMAH, Domaine St. Paul, 84914, Avignon, France

Abstract: The aim of this study was to develop an inversion approach to estimate surface soil moisture from X-band SAR data over irrigated grassland areas. This approach simulates a coupling scenario between Synthetic Aperture Radar (SAR) and optical images through the Water Cloud Model (WCM). A time series of SAR (TerraSAR-X and COSMO-SkyMed) and optical (SPOT 4/5 and LANDSAT 7/8) images were acquired over an irrigated grassland region in southeastern France.

An inversion technique based on multi-layer perceptron neural networks (NNs) was used to invert the Water Cloud Model (WCM) for soil moisture estimation. Three inversion configurations based on SAR and optical images were defined: (1) HH polarization, (2) HV polarization, and (3) both HH and HV polarizations, all with one vegetation descriptor derived from optical data. The investigated vegetation descriptors were the Normalized Difference Vegetation Index "NDVI", Leaf Area Index "LAI", Fraction of Absorbed Photosynthetically Active Radiation "FAPAR", and the Fractional vegetation COVER "FCOVER". These vegetation descriptors were derived from optical images. For the three inversion configurations, the NNs were trained and validated using a noisy synthetic dataset generated by the WCM for a wide range of soil moisture and vegetation descriptor values. The trained NNs were then validated from a real dataset composed of X-band SAR backscattering coefficients and vegetation descriptor derived from optical images. The use of X-band SAR measurements in HH polarization (in addition to one vegetation descriptor derived from optical images) yields more precise results on soil moisture (M_v) estimates. In the case of NDVI derived from optical images as the vegetation descriptor, the Root Mean Square Error on M_v estimates was 3.6 Vol.% for NDVI values between 0.45 and 0.75, and 6.1 Vol.% for NDVI between 0.75 and 0.90. Similar results were obtained regardless of the other vegetation descriptor used.

Keywords: grassland; TerraSAR-X; COSMO-SkyMED; neural networks; inversion; soil moisture; vegetation indices

1. Introduction

Monitoring the spatio-temporal evolution of soil moisture over irrigated grassland areas is of crucial importance for effective irrigation and crop management (Allen et al., 1998; Brereton and Hope-Cawdery, 1988; Hong et al., 2013; Leenhardt et al., 2004; Merot et al., 2008). In situ sensors to measure soil moisture are costly and provide only local information. Thus, these sensors are not sufficient for monitoring the soil moisture in huge irrigated grassland areas because the soil moisture presents large heterogeneities due to environmental characteristics and irrigation practices. SAR (Synthetic Aperture Radar) data have shown great potential to provide spatially distributed surface soil moisture measurements over bare and vegetated soil (Aubert et al., 2011; Baghdadi et al., 2012a; Gherboudj et al., 2011; Paloscia et al., 2008, 2013; Prevot et al., 1993; Santi et al., 2013). Due to their ability to operate in all weather conditions, SAR sensors offer the opportunity to monitor and quantify the surface soil moisture at a large scale with high spatial and temporal resolution.

SAR remote sensing was widely and primarily used to estimate the soil moisture and surface roughness. Over bare soil (or soil with little vegetation cover) the estimation of soil moisture was performed using either a physical (e.g the Integral Equation Model, Fung et al., 1992) or statistical (e.g Dubois and Oh models, Dubois et al., 1995; Oh, 2004) model in an inversion scheme. In contrast to physical models, statistical models need to be calibrated using in situ measurement and SAR observation acquired over the study area. Moreover, the use of statistical models is limited to the ranges of data variation used for calibration. Most of the studies used radar data in the X- and C-bands to estimate the soil moisture of bare soil and have shown good results, with an accuracy between 3 and 6 Vol.% (Aubert et al., 2011; Baghdadi et al., 2012a; Srivastava et al., 2003, 2009; Zribi et al., 2005).

The presence of vegetation cover complicates soil moisture retrieval from SAR data because vegetation canopy not only introduces two-way attenuation in SAR backscatter from soil, but also contributes its own backscatter (He et al., 2014; Srivastava et al., 2011). Most studies used the Water Cloud Model (WCM) in an inversion scheme for soil moisture estimation over areas with vegetated cover. In the WCM the total reflected radar signal is modeled as a function of the vegetation and soil contribution. The vegetation contribution, direct scattering and attenuation, is computed mainly using one biophysical parameter representing the vegetation effect. This biophysical parameter could be estimated from optical data. Therefore, it is important to combine SAR and optical data for operational mapping of soil moisture over areas covered by vegetation (Fieuzal et al., 2011; He et al., 2014; Hosseini and Saradjian, 2011; Notarnicola et al., 2006; Prakash et al., 2012). Currently, the high temporal repetitiveness of X-band (at least one day in case of TSX and CSK) and optical (between 16 and 26 days for Landsat-7/8 and SPOT-4/5 data, respectively) data makes the combined use of SAR and optical data for soil and vegetation parameter monitoring more reliable in near real time.

Optical data have shown a great potential to estimate biophysical parameters of vegetation. These parameters can be derived from optical data using physical and statistical models. Physical models (e.g PROSAIL, and SAFY) invert the vegetation spectral reflectance and estimate the biophysical parameters of the vegetation (Botha et al., 2010; Ceccato et al., 2001; Darvishzadeh et al., 2008; Fieuzal et al., 2011). Most statistical models are based on direct relationships between the Normalized Differential Vegetation Index (NDVI) and the measured biophysical parameters of vegetation, such as the LAI of wheat, grasslands, maize, corn and rice (Baret and Guerif, 2006; Baret et al., 2007; Bsaibes et al., 2009; Courault et al., 2008, 2010)

The possibility of retrieving soil parameters from vegetated surfaces was widely investigated using C-band configurations, whereas few studies were carried out using X-band data. Hajj *et al.* (2014) showed that the radar signal penetration depth in the X-band (incidence about 30°) is high, even in dense grass cover (HVE "Vegetation Height" about 1m, BIO "Biomass" up to 3.9 kg/m^2). These results encourage the use of X-band with medium angle (about 30°) in both HH and HV polarizations for soil moisture estimates over grassland. For C and X-bands SAR data, studies showed that it is possible to estimate the soil moisture with accuracy from 2 to 8 Vol.% (RMSE "Root Mean Square Error") (Gherboudj et al., 2011; He et al., 2014; Notarnicola et al., 2006; Prévot et al., 1993; Sikdar and Cumming, 2004; Wang et al., 2011; Yang et al., 2012; Yu and Zhao, 2011; Zribi et al., 2011).

The aim of this study is to evaluate the potential of X-band SAR data combined with optical data to estimate soil moisture over irrigated grassland areas located in southeastern France. An approach based on the inversion of the WCM using multi-layer neural networks (NNs) was developed.This approach relies on four main steps: (1) parameterize the WCM, (2) simulate learning the SAR synthetic dataset, (3) train the neural networks according to three inversion configurations using a part of the synthetic dataset, and finally (4) apply the trained NNs on synthetic and real datasets to validate the inversion approach. In this paper, section 2 presents the study areas and the ground-truth measurements performed in situ. Section 3 describes the methodology. The results are shown in section 4. Finally, section 5 presents the principal conclusions.

2. Study area and in situ measurements

2.1 Study area

The study area, named "Domaine de Merle", is an experimental farm located in southeastern France (centered at 43.64° N, 5.00° E). Its extent is approximately 400 hectares, among which 150 hectares are irrigated grassland for hay production (Figure 1). The produced hay is high-value with a Certified Origin Product label (COP) thanks to the specific environmental conditions and conventional irrigation guidelines.

The climate is Mediterranean with a rainy season between September and November. The average cumulative rainfall collected at the study site reached 457.5 mm in 2013, and in general varies between 350 mm and 800 mm over the past 20 years (Courault et al., 2010). The mean air temperature is approximately 8°C and 24°C during winter and summer, respectively (Courault et al., 2010). The in situ measured evaporation rate (potential evapotranspiration) can reach 10 mm/day during the summer due to high temperatures associated with dry and windy conditions. Meteorological instruments installed in the study area allow for recording hourly temperature and precipitation.

The topsoil texture of irrigated plots is stony loam (15% to 20% pebbles) with the depth varying from 30 cm to 80 cm, depending on the plot age (between 10 years and 3 centuries) (Bottraud et al., 1984; Mérot, 2007). The soil is always very smooth thanks to regular irrigation (approximately every 10 days) by gravity. Moreover, the soil has a moderate retention capacity, with concentrated vegetation roots in the upper 30 cm (Merot et al., 2008).

Plots were leveled with a very gentle slope to allow irrigation by gravity (border irrigation). Irrigation is applied by means of canals which bring water to the highest extremities of the plots.

Each plot is irrigated every 10 days on average from April to September. Plots are harvested three times a year, in May, July, and September.

Figure 1. Location of the study site (Domaine du Merle). Black polygons delineate training irrigated grassland plots where ground measurements were made.

2.2 SAR Images

Twenty three X-band SAR images were acquired by the COSMO-SkyMed (CSK) and TerraSAR-X (TSX) sensors between April and October 2013. All SAR images are in dual-polarization mode (HH and HV) with incidence angles between 28.3° and 32.5° (Table 1). Moreover, TSX and CSK images are in Stripmap (pixel spacing of 3 m) and Stripmap Pingpong (pixel spacing of 8 m) imaging modes, respectively.

Radiometric calibration of SAR images was performed using algorithms developed by the German Aerospace Center (DLR) and the Italian Space Agency (ASI). The radiometric calibration transforms the digital number of each pixel (DN_i) to a radar backscattering coefficient

($\sigma_i°$). For the seven TSX MGD (Multi Look Ground Range Detected), the radiometric calibrations were performed according to the following equation (Eineder et al., 2008):

$$\sigma_i° = Ks \cdot DN_i^2 \cdot \sin(\theta) - NESZ \quad \text{(Eq. 1)}$$

where Ks is the calibration constant, θ is the reference incidence angle, and NESZ is the Noise Equivalent Sigma Zero.

For the sixteen CSK images, $\sigma_i°$ was computed from the DN_i using the following equation:

$$\sigma_i^{\circ} = DN_i^{\,2} \cdot \frac{1}{K \cdot F^2} \cdot \sin(\theta)\, R_{ref}^{2 \cdot R_{exp}} \quad \text{(Eq. 2)}$$

where R_{ref} is the reference slant range, R_{exp} is the reference slant range exponent, K is the calibration constant, and F is the rescaling factor.

Values of parameters given in Equations 1 and 2 are given in the metadata associated with each TSX and CSK image. The $\sigma_i°$ were then averaged for each grassland plot and converted to the decibel scale according to the following equation:

$$\sigma^{o}_{dB} = 10 \cdot \log_{10} (\sum \sigma_i°) \quad \text{(Eq. 3)}$$

The number of looks used to generate a pixel spacing of 3 m x 3 m is one look in both the range and the azimuth. However, to generate a pixel spacing of 8 m x 8 m, the number of looks is one look in the range and four in the azimuth. The radar image pixel count in the training plots is between 521 and 1686 pixels for the CSK images, and between 3425 and 11320 for the TSX images.

For training plots, a comparison was performed between the backscattering coefficients (in both HH and HV polarizations) derived from one TSX and one CSK image, both acquired on the same day (08/07/2013) with about 40 minutes time interval. For such time interval the soil and vegetation conditions remain unchanged. For both HH and HV polarizations, results showed unbiased comparison with low Root Mean Square Error (RMSE ~ 0.4 dB), low Mean Absolute Percentage Error (MAPE < 5 %) and high correlation coefficient (R^2~0.9).

2.3 Optical Images

Thirty optical images were acquired by SPOT-4, SPOT-5, LANDSAT-7 and LANDSAT-8 between April and October 2013 at dates very close to the SAR images (Table 1). The calibration of optical images includes correction for atmospheric effects and ortho-rectification. SPOT-4 images were calibrated by the CESBIO (Centre d'Etudes Spatiales de la BIOsphère) in the framework of the Take 5 experiment *(http://www.cesbio.ups-tlse.fr/)*. Atmospheric correction of SPOT-4 images was performed according to the method described in the study of Hagolle *et al.* (2008). SPOT-5 and LANDSAT-8 were corrected for atmospheric effects using the Simplified Method of Atmospheric Correction (SMAC) (Rahman and Dedieu, 1994). The SMAC model transforms the TOA reflectance (Top Of Atmosphere) to an atmospherically corrected reflectance. Input data to the SMAC model, the Aerosol Optical Thickness (AOT) at 550 nm, the water vapor content (g/m^2), and Ozone, were obtained from the AERONET (AErosol Robotic NETwork) website *(http://aeronet.gsfc.nasa.gov/)*. LANDSAT-7 images, already corrected for atmospheric effects, were downloaded directly from the website of the USGS *(http://earthexplorer.usgs.gov/)*. The atmospheric correction of LANDSAT-7 images were carried out by NASA (National Aeronautics and Space Administration) by applying the 6S (Second Simulation of a Satellite Signal in the Solar Spectrum) radiative transfer model data as described by Masek et *al.* (2013) . Finally, LANDSAT-7/8 images were already ortho-rectified, whereas SPOT-5 images were ortho-rectified using the terrain correction module implemented in the ERDAS imaging software. The optical image pixel count in the training plots is between 39 and 108 for LANDSAT images, and between 79 and 223 for SPOT images.

The NDVI was computed from the optical images. Then, NDVI pixel values were averaged for each plot. For all training plots, a comparison was performed between NDVI derived from images acquired by different sensors (LANDSAT-7/8, SPOT-4/5) with time interval less than four days. Results showed unbiased comparison with low RMSE (≤ 0.04), low MAPE (< 5%), and good correlation coefficient (R^2 between 0.70 and 0.98). Thus, NDVI derived from different sensors were comparable.

Table 1. Acquisition dates of SAR and optical images (in 2013). Ground measurements are soil moisture and roughness, LAI, FAPAR, FCOVER, BIO, VWC, and HVE (described in section below).

	April						May						Jun												July								
	14	17	19	24	25	30	03	04	11	14	22	27	03	04	06	10	11	12	13	14	18	26	28	30	05	08	12	14	16	19	22	29	30
TSX			X			X				X	X															X							X
CSK															X	X	X			X		X				X	X		X				
SPOT-4 & 5	X			X				X		X						X			X		X			X	X								
LANDSAT-7 & 8		X	`		X		X		X			X		X				X					X					X			X		X
In situ measurement			X			X	X			X	X		X		X	X	X			X		X				X	X		X	X		X	X

	August											September								October			
	01	09	13	15	20	21	22	23	26	29	31	02	03	04	10	16	22	24	01	04	06	11	16
TSX																			X				
CSK	X	X							X	X		X			X					X			X
SPOT-4 & 5	X				X												X				X	X	
LANDSAT-7 & 8				X				X			X					X		X					
In situ measurement	X	X	X	X		X	X		X	X		X	X	X	X				X	X	X		X

2.4 In situ measurements

In situ campaigns were conducted simultaneously with SAR acquisitions to collect ground-truthed measurements of soil and vegetation parameters in twelve training plots (plots completely flooded or under harvest were not considered). These plots are well levelled and have enough size to be considered as sampling unit (Patel and Srivastava, 2013). The dimension of sampled plot ranges between 2.9 ha and 8.80 ha.

2.4.1 Soil moisture and roughness

Due to the high irrigation frequency and evapotranspiration rates, soil moisture measurements were performed close in time (within a window of 2 hours) to the satellite overpass. For each training plot, twenty five to thirty measurements of volumetric soil moisture approximately evenly distributed in space (on average every 20 m) were conducted in the top 5 cm of soil by means of a calibrated TDR (Time Domain Reflectometry) probe. Soil moisture was measured in the top 5 cm of soil because the radar penetration depth is assumed to be a few centimeters in the X-band (Ulaby et al., 1986). The soil moisture of each plot was represented by the mean of all soil moisture measurements performed in that plot, except when high spatial variability of soil moisture was observed. This variability is the result to current or recent (few hours before) irrigation events. In this case, many homogenous sub-plots were defined using hand-held GPS (brand: GARMIN, model: OREGON 550, location precision < 2m). The soil moisture was approximately 12 Vol.% when the plot was not supplied by water (irrigation or rainfalls) for 10 days during the summer, and it reached approximately 45 Vol.% approximately 10 hours after irrigation ended. The standard deviation of soil moisture measurements within a plot was between 1 and 5 Vol.%.

Soil roughness measurements were conducted only once because soil roughness remains stable, using a needle profile-meter (total length of 1 m, and needle spacing of 2 cm). Ten roughness profiles (five parallels and five perpendiculars to SAR's line of sight) were recorded for each plot couple of days after the third harvest, when the vegetation was very short. The root mean square height (*Hrms*) which represents the vertical scale of roughness, and the correlation length (L), representing the horizontal scale, were derived by processing the roughness profile. The individual autocorrelation functions are averaged, to produce a mean autocorrelation function representing each training plot (exponential function). Then, this mean autocorrelation function was used to derive H*rms* and L. The *Hrms* values varied between 0.35 and 0.55 cm, and the correlation length (*L*) ranged from 2.00 to 4.60 cm.

2.4.2 Vegetation parameters

Additionally, in situ measurement of vegetation parameters were performed to estimate the fresh Biomass (BIO), Vegetation Water Content (VWC), Vegetation Height (HVE), leaf area index (LAI), Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), and Fractional vegetation COVER (FCOVER). The vegetation characteristics within each plot are relatively homogeneous. To determine the BIO, two vegetation samples over a 50 cm x 50 cm square were clipped using shears at the center of each plot, and then weighed (wet biomass). Later, these samples were dried at 70°C for three days to calculate the VWC (VWC = wet biomass – dry biomass). The VWC is well correlated to the BIO (VWC = 0.80 BIO, R^2=0.99), it increases as BIO increases (i.e. growing season progresses). A poor correlation was found between VWC and soil moisture. Twenty measurements of vegetation heights were performed for each plot (the standard deviation of HVE measurements within a plot was between 5 and 10 cm). Finally, twenty to thirty hemispherical images were acquired for each plot by means of a fish eye lens. These photos were processed using CanEye software *(http://www6.paca.inra.fr/can-eye)* to estimate the LAI, FAPAR and FCOVER. Figure 2 showed photos for plots at different vegetation growth stage. For HVE, LAI, FAPAR, and FCOVER the measurements location within each plot were approximately evenly distributed in space (on average every 20 m). All vegetation measurements within each plot were averaged to provide a mean value for each plot. Figure 2 showed photos for plots at different vegetation growth stage.

In our study site, HVE reaches a value between 80 cm and 120 cm (BIO ~ 4.2 kg/m^2, LAI ~ 5 m^2/m^2) five to seven days before harvest. About ten days after harvest, the HVE reaches a value of about 30 cm (BIO ~0.80 kg/m^2, LAI ~2.5 m^2/m^2).

	Training plot 2e (Jun 10, 2013) BIO= 0.89 kg/m^2 HVE = 0.25 m LAI = 1.01 m^2/m^2
	Training plot 1l (April 19, 2013) BIO= 1.90 kg/m^2 HVE = 0.50 cm LAI = 3.98 m^2/m^2
	Training plot 1m (May 14, 2013) BIO= 3.56 kg/m^2 HVE = 1.13 m LAI = 4.71 m^2/m^2

Figure 2. Ground-based photographs of study sites illustrating variations in grass growth stages along with in situ measurements.

The *in situ* campaigns, frequently performed along each of the three growth cycles, capture full range of soil moisture and vegetation conditions (Table 2). Table 2 shows the number of sampled plots that correspond to soil and vegetation conditions.

Table 2 : The number of sampled plots for each soil moisture and vegetation conditions

	Low ($HVE \leq 25$ cm)	Medium ($25 < HVE \leq 60$ cm)	High ($HVE > 60$ cm)
Low ($M_v \leq 20$ Vol.%)	10	17	13
Medium ($20 < M_v \leq 30$ Vol.%)	19	40	20
High ($M_v > 30$ Vol.%)	20	21	18
	Low ($VWC \leq 0.7$ kg/m^2)	Medium ($0.7 < VWC \leq 1.3$ kg/m^2)	High ($VWC > 1.3$ kg/m^2)
Low ($M_v \leq 20$ Vol.%)	12	13	15
Medium ($20 < M_v \leq 30$ Vol.%)	30	26	23
High ($M_v > 30$ Vol.%)	20	22	17

3. Methods

3.1 Radar signal modeling

In this study, the Water Cloud model (WCM), developed by Attema and Ulaby (1978), was used for modeling the total backscattered radar signal according to soil moisture and vegetation parameters. This semi-empirical model is widely used over soil with vegetation cover because it can be easily performed in an inversion scheme to estimate soil moisture and vegetation parameters (Gherboudj et al., 2011; Prevot et al., 1993; De Roo et al., 2001; Sikdar and Cumming, 2004; Soon-Koo Kweon et al., 2012; Wang et al., 2011; Yang et al., 2012; Yu and Zhao, 2011; Zribi et al., 2011). The significant variables in the WCM are the medium height and dielectric cylinder density (Attema and Ulaby, 1978). The latter was assumed to be proportional to the volumetric water content of the canopy. Very few studies have compared different vegetation parameters to define the optimal one for use in the WCM. Champion (1991) and Champion and Guyot (1991) found that the LAI (m^2/m^2) better represents the wheat canopy in the WCM than the VWC per unit volume (kg/m^3). Said *et al.* (2012) compared the use of LAI (m^2/m^2), VWC (kg/m^2), and HVE and found that the use of LAI as the vegetation descriptor allows the accurate simulation of the vegetation volume contribution (sugarcane, cherry, rice, and grassland).

In this context, the WCM represents the total backscattered radar signal (σ^0_{tot}) in linear scale as a sum of the direct vegetation contribution (σ^0_{veg}) and soil contribution attenuated by the vegetation volume ($T^2\, \sigma^0_{sol}$).

$$\sigma^0_{tot} = \sigma^0_{veg} + T^2\, \sigma^0_{sol} \qquad \text{(Eq. 4)}$$

$$\sigma^0_{veg} = A.V_1.\cos\theta\ (1 - T^2) \qquad \text{(Eq. 5)}$$

$$T^2 = \mathrm{Exp}\,(-2.B.V_2.\sec\theta) \qquad \text{(Eq. 6)}$$

$$\sigma^0_{sol} = C(\theta)\, \exp\,(D.M_v) \qquad \text{(Eq. 7)}$$

Where:

- V_1 and V_2 are vegetation descriptors (BIO (kg/m^2), VWC (kg/m^2), HVE (m), LAI (m^2/m^2), FAPAR, FCOVER, and NDVI)
- θ is the radar incidence angle
- A and B are parameters that depend on the canopy descriptors and radar configurations
- T^2 is the two way attenuation
- C is dependent on the roughness and incidence angle
- D is the sensitivity of the radar signal to volumetric soil moisture in the case of bare soils, which is dependent on radar configurations
- M_v is the volumetric soil moisture (expressed in Vol.%).

3.2 Soil moisture retrieval

In this study, soil moisture was estimated by means of multi-layer perceptron neural networks (NNs). The Levenberg-Marquardt optimization algorithm (Marquardt, 1963) was used to train the NNs. The NNs architecture is composed of three layers: input, one hidden, and output. The NNs have a two dimensional input vector when using one polarization (HH or HV) in addition to one vegetation descriptor. Using two polarizations (HH and HV) in addition to one vegetation descriptor, the NNs have a three dimensional input vector. The one dimensional output vector contains soil moisture. The numbers of neurons associated with the hidden layer was determined by training the NNs using different numbers of neurons. 20 hidden neurons provided accurate estimates of reference parameters (Baghdadi et al., 2012a; Chai et al., 2009). Sigmoidal and linear transfer functions were associated with the hidden and output layer, respectively. These functions allow non-linear transformations from input to output (Del Frate and Solimini, 2004; Del Frate et al., 2003; Paloscia et al., 2008). To study the performance of the inversion approach, the NNs were trained and validated on the synthetic datasets.

A Synthetic dataset of SAR backscatter data was generated from the parameterized WCM to be used in the procedures leading to the estimation of soil moisture by means of the neural networks (NNs) technique. The parameterized WCM is able to simulate the backscattering coefficients at both HH and HV polarizations using the volumetric soil moisture, one vegetation descriptor, and incidence angle values as input variables. Only parameters easily estimated from optical images such as NDVI, LAI, FAPAR and FCOVER were considered in the synthetic datasets generation. Indeed, only few studies showed that the optical data could be used for estimating the biomass, vegetation water content, and the vegetation height. Four synthetic datasets have thus been generated using NDVI, LAI, FAPAR and FCOVER as vegetation descriptors (V1 and V2 in equations 4 and 5) to evaluate the most adequate vegetation descriptor for vegetation layer characterization in the WCM, and to open a perspective for future works based on SAR and optical data coupling. Indeed, several studies have developed methods to correct atmospheric effects in optical images, allowing the accurate estimation of the NDVI (Agapiou et al., 2011; Masek et al., 2013; Rahman and Dedieu, 1994; Saastamoinen, 1972; Vermote et al., 2002). Regarding the other vegetation descriptors, many studies have developed methods to estimate LAI, FAPAR, and FCOVER from optical images (Baret and Guyot, 1991; Bsaibes et al., 2009; Carlson and Ripley, 1997; Carlson et al., 1994; Claverie et al., 2013; Courault et al., 2008; Darvishzadeh et al., 2008b; Duveiller et al., 2011; Fensholt et al., 2004; Guerschman et al., 2009; Li et al., 2014; North, 2002). In addition, in the framework of our study, LAI, FAPAR, and FCOVER of our studied grassland were derived from optical images

(SPOT-4, SPOT-5, LANDSAT-7, LANDSAT-8) using the BV-NNET (Biophysical Variables Neural NETwork) tool developed based on algorithms proposed by Baret *et al.* (2007) and then optimized through the SIRRIMED project (http://www.sirrimed.org/index.php). A comparison was performed between the LAI, FAPAR and FCOVER derived from BV-NNET (using optical images) and those derived from hemispherical photos. Results showed unbiased estimations of LAI, FAPAR, and FCOVER by the BV-NNET. Moreover, the BV-NNET estimates the LAI with an RMSE of 0.66 m^2/m^2 and an RRMSE (as well as MAPE) around 29%. For FAPAR and FCOVER, an RMSE around 0.13 and an RRMSE (as well as MAPE) around 19% were obtained.

The synthetic dataset based on NDVI as the vegetation descriptor comprises 80 elements (8 x 10, Table 3). Each element of the dataset contains radar signals in HH and HV polarizations for a given NDVI and volumetric soil moisture (Table 3). Moreover, synthetic dataset based on LAI and FAPAR (as well as FCOVER) comprised 248 (8 x 31, table 3) and 168 (8 x 21, table 3) elements, respectively.

Table 3. The minimum, maximum, and step values of WCM inputs.

Parameter	Min value	Max value	Step	Total elements
NDVI	0.45	0.90	0.05	10
LAI (m^2/m^2)	0.0001	6	0.20	31
FAPAR	0.0001	1	0.05	21
FCOVER	0.0001	1	0.05	21
M_v (Vol.%)	10	45	5	8

To make WCM simulations more realistic, uncertainties of SAR measurements were added to the simulated radar response. The uncertainties range is between 0.6 and 1 dB for CSK and TSX sensors (Agenzia Spaziale Italiana, 2007; Coletta et al., 2007; Iorio et al., 2010; Schwerdt et al., 2008; Torre et al., 2011). Thus, we considered two absolute uncertainties values (±0.75, and ±1.00 dB) to be added to the simulated radar response. Moreover, relative uncertainties were added on our reference vegetation descriptor values (NDVI, LAI, FAPAR, and FCOVER) to handle the associated uncertainty. For NDVI, Simoniello *et al.* (2004) reported a relative uncertainty of approximately 8% on NDVI values estimated from AVHRR (Advanced Very High Resolution Radiometer) calibrated data over pasture and cultivated areas. El Hajj *et al.* (2008) found that the relative uncertainty on NDVI computed from SPOT-5 surface reflectance data over sugarcane fields is approximately 13%. For the other vegetation descriptors, studies showed for crop canopies (corn, grass, sunflower, maize, wheat, rapeseed and sunflower) relative uncertainty between 10% and 30% for LAI, and between 5% and 20% for FAPAR and FCOVER (Bsaibes et al., 2009; Claverie et al., 2013; Courault et al., 2008; Duveiller et al., 2011; North, 2002). In addition, the uncertainty on the vegetation descriptor estimates depends on crop type (Bsaibes et al., 2009; Claverie et al., 2013). Moreover, the comparison between derived LAI, FAPAR, and FCOVER from our optical images with ground-truthed measurements yields a relative RMSE (Root Mean Square Error) of 29.12, 19.24, and 18.14%, respectively. Therefore, in our study we considered a relative additive noise of 15, 30, and 20% on the NDVI, LAI, and FAPAR (as well as FCOVER), respectively.

Zero-mean Gaussian noise with a standard deviation equal to absolute and relative uncertainties were added to the radar signal simulated by the WCM and reference vegetation descriptors, respectively. Finally, to obtain statistically significant datasets, 500 random samplings of zero-mean Gaussian noise was added to each simulated radar response and each vegetation descriptor value.

Three case studies to estimate soil moisture using X-band SAR data were evaluated:

- Case 1: Noisy radar signal at HH polarization and noisy vegetation descriptor as the inputs to NNs, and soil moisture as the target.
- Case 2: Noisy radar signal at HV polarization and noisy vegetation descriptor as the inputs to NNs, and soil moisture as the target.
- Case 3: Noisy radar signal at HH and HV polarizations and noisy vegetation descriptor as the inputs to NNs, and soil moisture as the target.

Finally, the calibrated NNs were used to invert real SAR measurements for estimation of the soil moisture. The inversion was performed according to the configurations above, but using SAR and a vegetation descriptor (LAI, FAPAR, and FCOVER) derived from optical images instead of the noisy radar signal and vegetation descriptors.

4. Results and discussions

4.1 Water Cloud Model parameterization, and modelling results

This section presents the results of the Water Cloud Model (WCM) parameterization, and shows the radar signal modelling results as a function of soil and vegetation parameters.

4.1.1 Water Cloud Model parameterization

The real dataset composed of SAR data and measurements of soil moisture and vegetation descriptors was divided into two sub-datasets. The first sub-dataset (training dataset) was used to fit the WCM model, whereas the second (validation dataset) was used to validate the soil moisture estimation of the WCM model. The training dataset contains the SAR and the ground-truthed data obtained during the three cycles for the half of training plots, whereas the validation dataset comprises the data collected for other half of plots. These two real sub-datasets contain a wide range of soil moisture (M_v) and vegetation descriptor values measured in situ (BIO, VWC, HVE, LAI, FAPAR, FCOVER) and derived from optical images (NDVI, LAI, FAPAR, and FCOVER) (Table 4). The two real sub-datasets have almost the same ranges of variation.

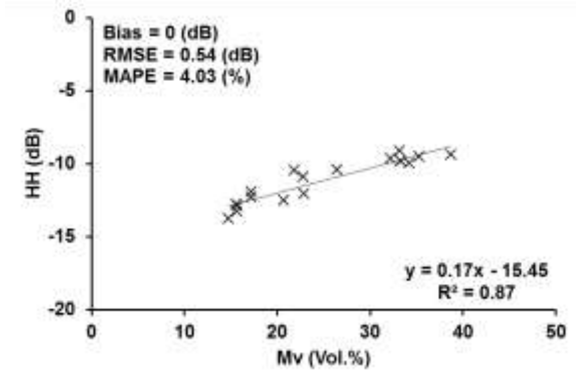
Table 4: ranges of variation of real training and validation datasets

		Min	Mean	MAX	Unit
Real training dataset	Mv	10.9	25.6	39.0	Vol.%
	In situ BIO	0.28	1.41	4.14	Kg/m^2
	In situ VWC	0.15	1.12	3.35	Kg/m^2
	In situ HVE	0.08	0.48	1.20	m
	In situ LAI	0.10	2.64	5.88	m^2/m^2
	In situ FAPAR	0.20	0.79	1.00	-
	In situ FCOVER	0.12	0.63	0.96	-
	LAI (BV-NNET)	0.20	2.63	5.04	m^2/m^2
	FAPAR (BV-NNET)	0.16	0.77	0.98	-
	FCOVER (BV-NNET)	0.16	0.66	0.96	-
	NDVI	0.47	0.73	0.88	-
Real validation dataset	Mv	14.1	27.0	47.0	Vol.%
	In situ BIO	0.30	1.31	3.46	Kg/m^2
	In situ VWC	0.03	1.02	2.87	Kg/m^2
	In situ HVE	0.08	0.45	1.15	m
	In situ LAI	0.26	2.23	4.00	m^2/m^2
	In situ FAPAR	0.20	0.73	0.93	-
	In situ FCOVER	0.09	0.57	0.88	-
	LAI (BV-NNET)	0.26	2.16	5.10	m^2/m^2
	FAPAR (BV-NNET)	0.09	0.69	0.98	-
	FCOVER (BV-NNET)	0.09	0.58	0.94	-
	NDVI	0.48	0.69	0.87	-

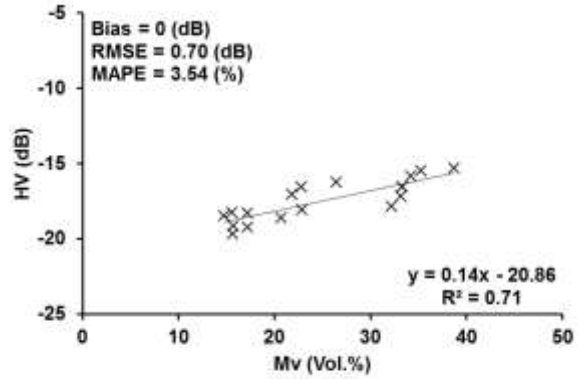
WCM parameterization consists of first estimating the sensitivity parameter D before fitting the model against ground-truthed measurements to estimate parameters A, B, and C (equations 4, 5, and 6).

- To estimate parameter D, SAR backscattering coefficients in HH and HV polarizations (dB scale) were linearly related to soil moisture (Vol.%) for 18 plots recently harvested (vegetation very short), to have the minimum vegetation effect on the backscattering coefficients (Figure 3). The slopes of these linear regressions represent the sensitivity of the backscattered radar signal to volumetric soil moisture on the dB scale (Figure 3). Results showed a good correlation between radar signal and volumetric soil moisture (R^2 = 0.87 and 0.71 for HH and HV, respectively). Moreover, results showed that the HH polarization is slightly more sensitive (0.172 dB/Vol.%) to volumetric soil moisture rather than HV (0.135 dB/Vol.%) polarization (Figure 3). In the WCM model, the sensitivity parameter D is represented on a linear scale. In linear unit, these sensitivities D_{HH} and D_{HV} are 0.03971 [m^2/m^2]/[Vol.%] and 0.03116 [m^2/m^2]/[Vol.%] for HH and HV polarizations, respectively

- A, B and C parameters were then estimated for each radar polarization and each vegetation descriptor (NDVI and ground-truthed BIO, VWC, HVE, LAI, FAPAR, and FCOVER) by minimizing the sum of squares of the differences between the simulated and measured radar signal. Therefore, the WCM was parameterized according to seven vegetation descriptors (Table 5). With A, B and C parameters, it becomes possible to predict WCM components (σ^0_{veg}, T^2, and σ^0_{sol}) and consequently the total backscattering coefficient (σ^0_{tot}) using one vegetation descriptor and the soil moisture values as inputs in the WCM.



(a)



(b)

Figure 3. Sensitivity of radar signal in both HH and HV polarization to volumetric soil moisture.

To validate the fitted WCM, a comparison was performed between the radar backscattering coefficients predicted by the mean of the parameterized WCM (using the soil moisture and ground-truthed vegetation descriptors of the real validation dataset) and the observed backscattering coefficients of the real validation dataset. Results showed that the fit of the WCM was slightly better in HH polarization than in HV polarization (Table 5). The limited correlation coefficient (R^2) is not due to difficulty of model to simulate radar data but particularly to limited range of radar data dynamic for different moisture and vegetation conditions. In addition, the quality of the fit is approximately the same for all the used vegetation descriptors with the RMSE (Root Mean Square Error) on the predicted backscattering coefficients between 0.76 and 0.86 dB in HH, and between 0.85 and 0.94 dB in HV polarization, depending on the used vegetation descriptor. Water cloud model is considered adequately fitted because the RMSE on simulated radar signal in both HH and HV polarizations is less than 1 dB, which is the same magnitude as the CSK and TSX sensors precision (Agenzia Spaziale Italiana, 2007; Coletta et al., 2007; Iorio et al., 2010; Schwerdt et al., 2008; Torre et al., 2011). Several studies used the WCM model to predict radar backscattering coefficients (Attema and Ulaby, 1978; Gherboudj et al., 2011;

Prevot et al., 1993; Ulaby et al., 1984). Attema and Ulaby, (1978) simulated the X-band backscattering coefficients for crops fields (alfalfa, corn, milo, and wheat) in HH and VV polarizations for a wide range of incidence angles (0^o-70^o) with a RMSE of simulated backscattering coefficients ranging between 1.5 and 2 dB, depending on the crop type. Ulaby *et al*, (1984) simulated the radar backscattering coefficients in the X-band (VV polarization and 50^o incidence angle) for wheat fields with a RMSE of 1.6 dB. Prevot *et al*. (1993) obtained a RMSE for wheat fields on the simulated backscattering coefficients of 1.24 and 0.72 dB in the C-band (HH, 20^o) and X-band (VV, 40^o), respectively. Gherboudj et al. (2012) predicted the backscattering coefficients in the C-band, in quad-polarization mode with a 30^o incidence angle for wheat and pea fields. The RMSE on the predicted backscattering coefficients in HH and VV polarizations was approximately 1 (for wheat) and 0.7 dB (for peas), respectively. In cross polarization, the backscattering coefficient was simulated with a RMSE of 1.2 and 0.2 dB for wheat and pea fields, respectively.

Table 5. Fit of WCM parameters for HH and HV polarizations (real validation dataset).

V1=V2	A_{HH}	B_{HH}	C_{HH}	D_{HH}	A_{HV}	B_{HV}	C_{HV}	D_{HV}	R^2_{HH} (R^2_{HV})	$RMSE_{HH}$ ($RMSE_{HV}$) (dB)
Ground-truthed BIO	0.0345	0.0995	0.0334	0.03971	0.0068	0.1850	0.0093	0.03116	0.49 (0.39)	0.85 (0.86)
Ground-truthed VWC	0.0438	0.1047	0.0324	0.039711	0.0084	0.1927	0.0088	0.03116	0.49 (0.39)	0.86 (0.86)
Ground-truthed HVE	0.1045	0.4314	0.0357	0.03971	0.0207	0.7882	0.0105	0.03116	0.52 (0.40)	0.79 (0.85)
Ground-truthed LAI	0.0205	0.0613	0.0338	0.03971	0.0041	0.0856	0.0088	0.03116	0.48 (0.29)	0.86 (0.95)
Ground-truthed FAPAR	0.0911	0.3275	0.0354	0.03971	0.0177	0.4662	0.0096	0.03116	0.47 (0.25)	0.80 (0.93)
Ground-truthed FCOVER	0.1021	0.3696	0.0355	0.03971	0.0203	0.5212	0.0095	0.03116	0.48 (0.27)	0.82 (0.94)
NDVI	0.0767	0.7944	0.0644	0.03971	0.016474	1.134	0.0221	0.03116	0.51 (0.33)	0.76 (0.93)

4.1.2 Modelling results

Modelling results obtained by using the NDVI as the vegetation descriptor in the WCM model will be presented first because (i) the best fit of water cloud model was obtained with NDVI as vegetation descriptor, and (ii) it is easier to derive NDVI from optical data than LAI, FAPAR, and FCOVER. Next, results with the LAI, FAPAR, FCOVER, BIO, VWC, and HVE as vegetation descriptors will be briefly discussed.

The WCM components ($T^2\sigma°_{sol}$ and $\sigma°_{veg}$) were simulated for wide ranges of soil moisture (M_v) and NDVI values using the WCM with the NDVI as the vegetation descriptor. For both HH and HV polarizations, the vegetation contribution ($\sigma°_{veg}$), soil contribution ($\sigma°_{sol}$), two-way attenuation (T^2), and consequently, the total backscattered signal ($\sigma°_{tot}$) were generated in a linear scale using the parameterized equations (3) to (6). NDVI values ranging from 0.45 to 0.90 were used to simulate the vegetation contribution and the two-way attenuation (V1=V2=NDVI in equations 4 and 5). In addition, the soil contribution was simulated using M_v-values ranging from 10 to 45 Vol.% (equation 6). The maximum values of NDVI and M_v correspond to the highest values derived from optical images and measured in situ, respectively.

Figure 4 shows the modelled $\sigma°_{veg}$, $T^2\sigma°_{sol}$ and $\sigma°_{tot}$ in dB units as a function of M_v using different values of NDVI (0.5, 0.7, and 0.9). In addition, the modelled $\sigma°_{veg}$, $T^2\sigma°_{sol}$ and $\sigma°_{tot}$ were also plotted according to NDVI for M_v values of 15, 20, 30 and 40 Vol.% (Figure 5).

Figure 4 shows that $\sigma°_{tot}$ in both HH and HV polarizations are always sensitive to soil moisture even for high NDVI values. The sensitivity of $\sigma°_{tot}$ to soil moisture decreases with the NDVI for NDVI between 0.45 and 0.90. For NDVI value equal to 0.50 this sensitivity is about 0.14 dB/% and 0.10dB/% for HH and HV, respectively. Moreover, for a NDVI value equal to 0.9, this sensitivity is approximately 0.08 and 0.04 dB/Vol.% in HH and HV, respectively. For each case in figure 4 statistical index were provided in table 6. Results showed that the WCM adequately simulates SAR real validation dataset observations (0 <Bias < 0.3, RMSE < 1dB, RRMSE and MAPE < 7%).

Figure 5 shows that $\sigma°_{tot}$ in both HH and HV polarization is slightly sensitive to the NDVI (for NDVI between 0.45 and 0.90). Indeed, as the vegetation grows, the decreasing soil contribution is similar to the increasing vegetation contribution. $\sigma°_{tot}$ shows slight decreases with increases in the NDVI until reaching a minimum, and starts to slightly increase. In both HH and HV polarizations, $\sigma°_{tot}$ decreases with NDVI for a NDVI lower than 0.60, 0.75, and 0.90 for M_v of 15, 20, and 30 Vol.%, respectively. However, the $\sigma°_{tot}$ in both HH and HV polarizations always decreases with NDVI (NDVI between 0.45 and 0.90) for M_V equal to 40 Vol.% due to the high soil contribution (Figures 5 d and h). This decrease of $\sigma°_{tot}$ with the NDVI is related to an increase in the attenuation of the soil contribution (T^2), which is more important than the enhanced contribution from the vegetation canopy (Balenzano et al., 2011; Brown et al., 2003; Mattia et al., 2003). Beyond these values of NDVI thresholds, $\sigma°_{tot}$ increases slightly with NDVI for M_v values between 15 and 30 Vol.%. This increase of $\sigma°_{tot}$ with NDVI results in the increase of the vegetation contribution combined with the decrease in the soil contribution. Moreover, the decrease and increase of $\sigma°_{tot}$ according to the NDVI are slightly more pronounced in HV than in HH polarization. Regarding vegetation contribution ($\sigma°_{veg}$), results showed that the modelled $\sigma°_{veg}$ in HH polarization increases from -17.7 dB for NDVI of 0.45 to -13.2 dB for NDVI of 0.90. For HV polarization, $\sigma°_{veg}$ increases from -23.5 dB to -18.8 dB for NDVI between 0.45 and 0.90. For each case in figure 5, statistical index were provided in table 7. Results showed that the WCM adequately simulates SAR real validation dataset observations ($0 < \text{Bias} < 0.7$, RMSE $\leq$ 1dB, RRMSE and MAPE < 8%).

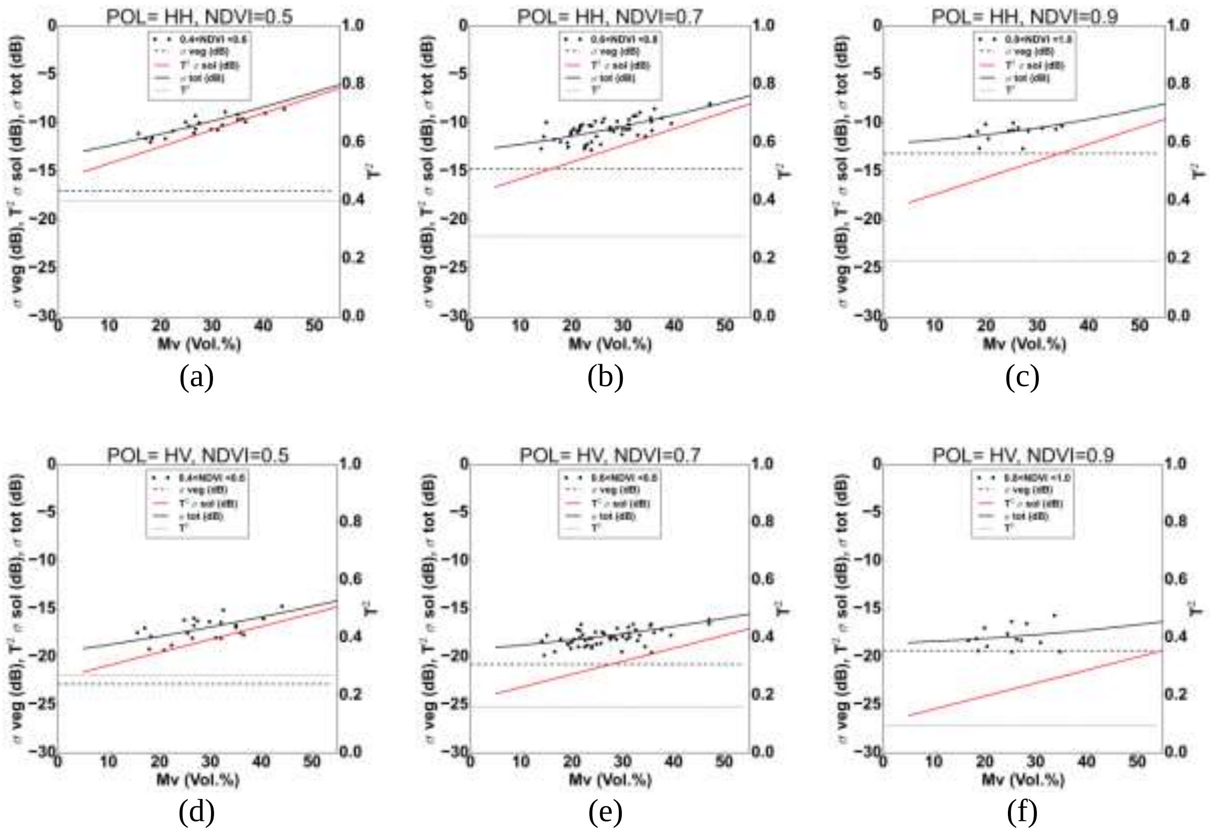


Figure 4. Behavior of WCM components ($\sigma°_{veg}$, $T^2\sigma°_{sol}$, and $\sigma°_{tot}$) in both HH and HV polarizations according to M_v. Black points represent the SAR data ($\sigma°_{tot}$: real validation dataset) associated with NDVI measurements within ± 0.1 of the NDVI used in the modelling.

Table 6: Statistical index for each case in figure 4

Case	Polarization	NDVI	Bias (dB)	RMSE (dB)	RRMSE (dB)	MAPE (dB)	R^2	Nb
Figure 4a	HH	0.50	0.3	0.6	6.0	5.4	0.71	23
Figure 4b	HH	0.70	0.0	0.9	8.0	6.7	0.45	52
Figure 4c	HH	0.90	0.1	0.8	7.0	4.8	0.12	14
Figure 4d	HV	0.50	0.1	1.0	5.7	5.1	0.30	23
Figure 4e	HV	0.70	0.2	0.8	4.3	3.3	0.26	52
Figure 4f	HV	0.90	0.1	1.1	6.4	5.5	0.03	14

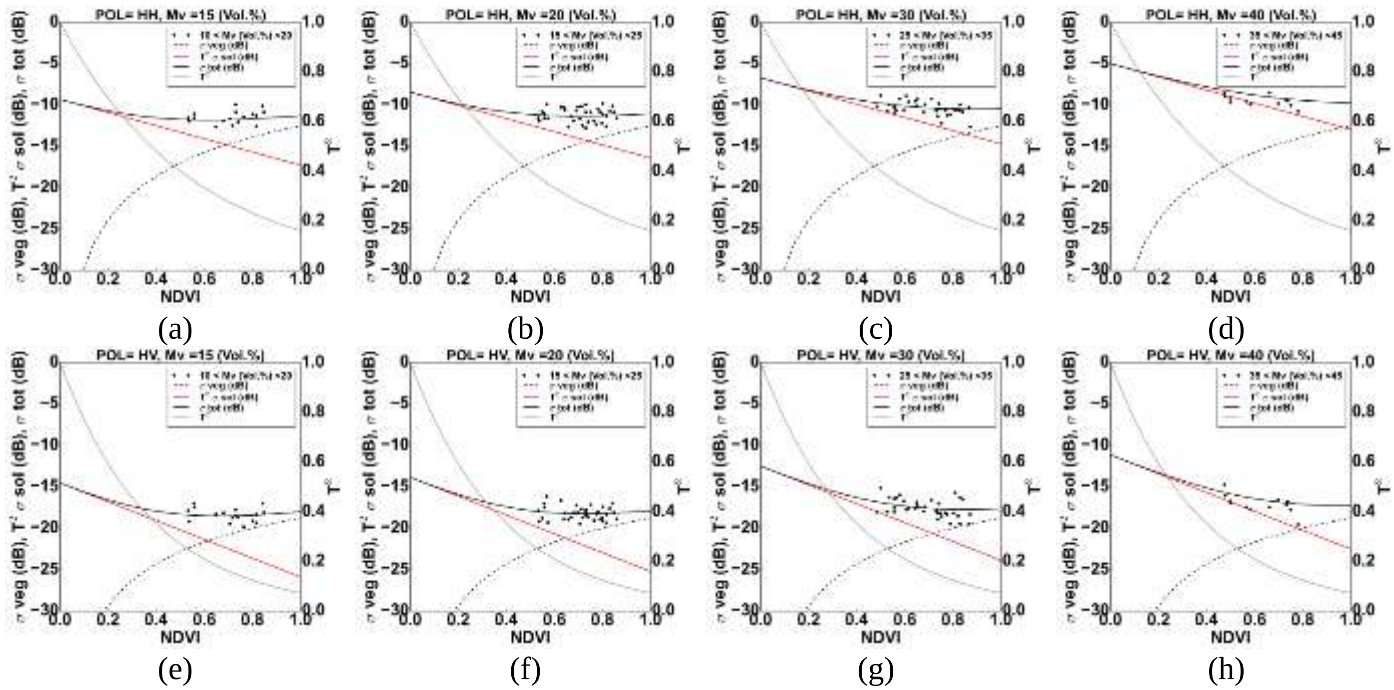


Figure 5. Behavior of WCM components ($\sigma°_{veg,}$ $T^2\sigma°_{sol}$, and $\sigma°_{tot}$) in both HH and HV polarizations according to NDVI. Black points represent the SAR data ($\sigma°_{tot}$: real validation dataset) associated with M_v measurements within ± 5 vol. % of the M_v used in the modelling.

Table 7: Statistical index for each case in figure 5

Case	Polarization	M_V (Vol.%)	Bias (dB)	RMSE (dB)	RRMSE (%)	MAPE (%)	R^2	Nb
Figure 5a	HH	15	-0.3	0.8	6.8	5.6	0.13	17
Figure 5b	HH	20	-0.1	0.9	7.7	6.7	0.00	36
Figure 5c	HH	30	0.1	0.7	7.1	5.4	0.16	37
Figure 5d	HH	40	0.6	0.8	8.7	7.6	0.32	12
Figure 5e	HV	15	-0.1	0.8	4.3	3.6	0.05	17
Figure 5f	HV	20	0.0	0.8	4.5	3.7	0.01	36
Figure 5g	HV	30	0.1	1.0	5.5	4.7	0.18	37
Figure 5h	HV	40	0.6	1.1	6.3	5.0	0.41	12

Table 8 shows NDVI thresholds from which the $T^2\sigma°_{sol}$ is dominated by $\sigma°_{veg}$ ($T^2\sigma°_{sol} < \sigma°_{veg}$). In HH polarization, these thresholds are approximately 0.69, 0.74, 0.85, 0.97 for soil moisture of

15, 20, 30 and 40 Vol.%, respectively. In HV polarizations and for M_v values of 15, 20, 30 and 40 Vol.%, $\sigma°_{veg}$ dominates $T^2\sigma°_{sol}$ for NDVI values greater than 0.62, 0.65, 0.71, and 0.79, respectively. Thus, for a given soil moisture value, the thresholds of NDVI for which the vegetation contribution dominates the soil contribution are lower in HV than in HH (Table 8).

Table 8. Threshold values of vegetation descriptors **at** which $\sigma°_{veg}$ dominates $T^2\sigma°_{sol}$ at both HH and HV polarizations. Dash symbols mean that the $\sigma°_{veg}$ is always dominated by $T^2\sigma°_{sol}$.

	HH polarization				**HV polarization**			
M_v (Vol.%)	15	20	30	40	15	20	30	40
NDVI	0.69	0.74	0.85	0.97	0.62	0.65	0.71	0.79
LAI (m^2/m^2)	4.22	4.60	5.43	-	3.69	3.94	4.47	5.05
FAPAR	0.87	0.95	-	-	0.77	0.82	0.93	-
FCOVER	0.78	0.84	0.99	-	0.68	0.72	0.82	0.92
BIO (kg/m^2)	2.55	2.77	3.28	3.85	1.95	2.07	2.34	2.64
VWC (kg/m^2)	2.20	2.40	2.84	3.35	1.70	1.82	2.06	2.32
HVE (m)	0.70	0.76	0.90	-	0.55	0.58	0.65	0.73

WCM components were also modelled using the LAI, FAPAR, FCOVER, BIO, VWC and HVE as vegetation descriptors. Similar results on the behavior of modelled total backscattered radar signal ($\sigma°_{tot}$) were obtained with all vegetation descriptors. Table 8 shows the values of the vegetation descriptors at which $\sigma°_{veg}$ dominates $T^2\sigma°_{sol}$. As an example, for soil moisture of 20 Vol.%, $\sigma°_{veg}$ in HH polarization dominates T^2 $\sigma°_{sol}$ for LAI values higher than 4.60 m^2/m^2. In addition, for some soil moisture and vegetation descriptor conditions, the vegetation contribution is always dominated by the soil contribution (dash symbol in Table 8). As an example, for soil moisture of 40 Vol.%, the vegetation contribution in HH polarization is always dominated by the soil contribution for HVE values between 0 and 1.2 m (maximum value of HVE obtained by ground-truthed measurements and used in modelling). In addition, Table 8 shows that the vegetation contribution in HV polarization dominates the soil contribution at threshold values of vegetation descriptors which are lower than those observed in HH polarization.

4.2 Soil moisture retrieval

Synthetic and real datasets were used to estimate the soil moisture for the three inversion configurations defined in section 3.2: (1) using the radar signal in HH and one vegetation descriptor, (2) using the radar signal in HV and one vegetation descriptor, and (3) using the radar signal in both HH and HV and one vegetation descriptor. The estimated soil moistures were compared to reference soil moisture values to evaluate the accuracy of the soil moisture estimates of each inversion configuration.

Before the use of neural networks for soil moisture estimation, the WCM model was numerically inverted. For some points of the synthetic and real datasets where the SAR backscattering coefficient is lower than the vegetation contribution simulated by the WCM, the direct inversion of the WCM is not numerically possible (about 10% of the datasets). Such limitation is overcome when using the NNs for both synthetic and real datasets. In addition, the Root Mean Square Error on M_v estimates was better with the NNs than using the direct inversion of the WCM (precision on M_v two times better). For these reasons, the neural networks inversion technique for soil moisture estimation was considered.

To estimate the soil moisture, neural networks were built for each inversion configuration using a part of the synthetic dataset. The quality of inversion approaches were studied using both the other part of the synthetic dataset and the real validation dataset.

4.2.1 Synthetic dataset

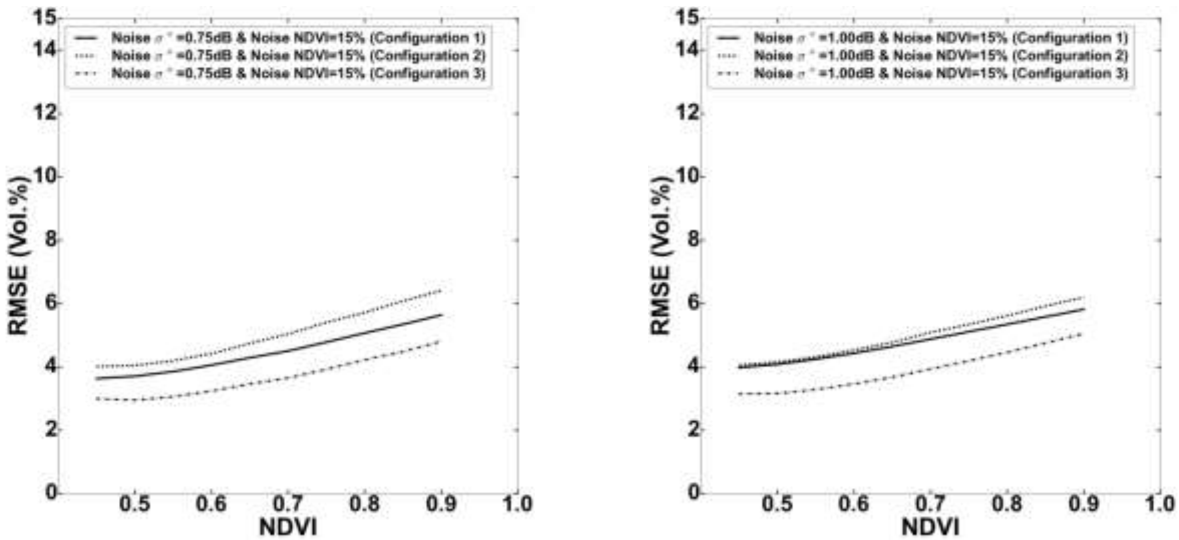
The synthetic dataset was composed of 2.10^7 elements (10 NDVI values x 8 M_v values x 500 random sampling values of the NDVI x 500 random sampling values of the simulated radar signal). According to the radiometric accuracy of the TerraSAR-X and COSMO-SkyMed signals, the radar signal simulated by the WCM model was noised using an additive Gaussian noise with zero mean and a standard deviation of 0.75 and 1 dB. The synthetic dataset was randomly divided into 80% training and 20% validation data samples. The prediction error based on a 5-fold cross-validation was estimated for each inversion configuration to assess the performance of the neural networks. Analysis of the results obtained with NDVI as the vegetation descriptor will be provided in detail whereas the results based on LAI, FAPAR, and FCOVER as the vegetation descriptors will be briefly described.

The Root Mean Square Error (RMSE), the Relative Root Mean Square Error (RRMSE), the Mean Absolute Percentage Error (MAPE), the associated mean deviation (bias = estimated M_v - reference M_v), and the correlation coefficient (R^2) were used to evaluate the performance of each inversion configuration. Table 9 presents statistical indexes (RMSE, RRMSE, MAPE, bias, and R^2) on M_v estimates computed from the validation dataset for reference M_v between 10 and 45 Vol.% and NDVI values between 0.45 and 0.90. Table 9 shows that the RMSE (as well as RRMSE, and MAPE) on M_v estimates is lower with HH polarization than with HV polarization (configuration 1 in comparison to configuration 2, Table 9). For a noise condition on the radar signal of ±0.75 dB, the RMSE is 4.5 Vol.% (RRMSE and MAPE about 17 %) with HH compared to 5.1 Vol.% (RRMSE and MAPE 19 %) with HV. In addition, results showed that the use of both HH and HV (in addition to the NDVI, configuration 3) slightly decreases the RMSE on M_v estimates (lower than 1 Vol.%). With configuration 3, the RMSE on M_v estimates reaches 3.7 Vol.% (RRMSE and MAPE about 14%) for a noise on the modeled radar signal of ±0.75 (Table 9). Table 9 also shows that the RMSE on M_v increases when the noise added to the modeled radar signal increases. This increase is approximately 1 Vol.% when the noise on the radar signal increases from ±0.75 dB to ±1.00 dB (Table 9). Finally, Table 9 also shows that the three inversion configurations provide un-biased M_v estimates and significant correlation coefficient (R^2 between 0.77 and 0.90).

Table 9. Statistical indexes on M_v estimates according to the three inversion configurations (RMSE (Vol.%) | RRMSE (%) | MAPE (%) | bias (Vol.%) | R^2). Configuration 1 uses HH and NDVI, configuration 2 uses HV and NDVI, and configuration 3 uses HH, HV and NDVI. Relative noise of the NDVI=15%.

	Noise on σ^0_{tot}: ±0.75 dB	**Noise on σ^0_{tot}: ±1.00 dB**
Configuration 1 (HH and NDVI)	4.5|16.5|17.1|0.0|0.85	5.5|19.8|21.0|0.0|0.78
Configuration 2 (HV and NDVI)	5.1|18.5|19.2|0.0|0.81	5.7|20.7|21.8|0.0|0.77
Configuration 3 (HH, HV and NDVI)	3.7|13.6|13.7|0.0|0.90	4.5|16.2|16.7|0.0|0.85

Figure 6 illustrates the RMSE evolution of M_v estimates as a function of NDVI for values between 0.45 and 0.90 for each inversion configuration. For each value of NDVI, statistics were calculated using all M_v values. The results showed that the RMSE of M_v estimates increases with NDVI for all inversion configurations. As an example, in configuration 3 (HH, HV and NDVI), the RMSE of soil moisture estimates increases from 3.0 Vol.% for NDVI of 0.45 to 4.8 Vol.% for a NDVI of 0.9 for a noise condition on the radar signal of ± 0.75 dB (Figure 6a). The results showed that for a given NDVI value between 0.45 and 0.90, the RMSE is in same order in configurations 1 and 2 (configuration 1 is slightly better than configuration 2) (Figure 6). In addition, results obtained with HH were worse than those obtained with HH and HV.



(a) (b)

Figure 6. Evolution of RMSE **of** M_v estimates according to the three inversion configurations as a function of NDVI for noise conditions on the modeled radar signal of ±0.75 dB (a), and ± 1 dB (b).

Moreover, the performances of neural networks for estimating soil moisture were analyzed according to NDVI for given M_v values (Figure 7). The results showed that the relative RMSE (RRMSE=RMSE/ M_v) of M_v estimates increases with the NDVI for the three inversion configurations. Indeed, as the vegetation grows (i.e., increasing NDVI values) the soil contribution decreases and the backscattering coefficients become less sensitive to soil moisture.

In addition, for a given NDVI between 0.45 and 0.90 the RRMSE decreases when M_v increases (Figure 7) because for a given NDVI value the soil contribution is more important for high than for low soil moisture conditions, and consequently, the errors on M_v estimates decrease when M_v increases. As an example, in configuration 1 (HH and NDVI), for a NDVI of 0.75 (LAI about 3 m^2/m^2), the RRMSE values are approximately 28.3, 20.0, 16.3, and 12.0% for reference M_v of 15, 20, 30 and 40 Vol.%, respectively. For low M_v (lower than 20 Vol.%), the RRMSE increases significantly with NDVI for high NDVI values (higher than 0.75, LAI about 3 m^2/m^2) in comparison to the RRMSE observed for higher M_v values (higher than 20 Vol.%). As an example, in configuration 3 (HH, HV and NDVI), the RRMSE on M_v estimates increases for M_v of 15 Vol.% and noise condition on the simulated radar signal of 0.75 dB from approximately 21% for NDVI=0.45 to 30% for NDVI=0.90. This increase in the RRMSE is only approximately 5% for M_v of 30 Vol.% (RRMSE increases from approximately 11% for NDVI=0.45 to 16% for NDVI=0.90) (Figure 7).

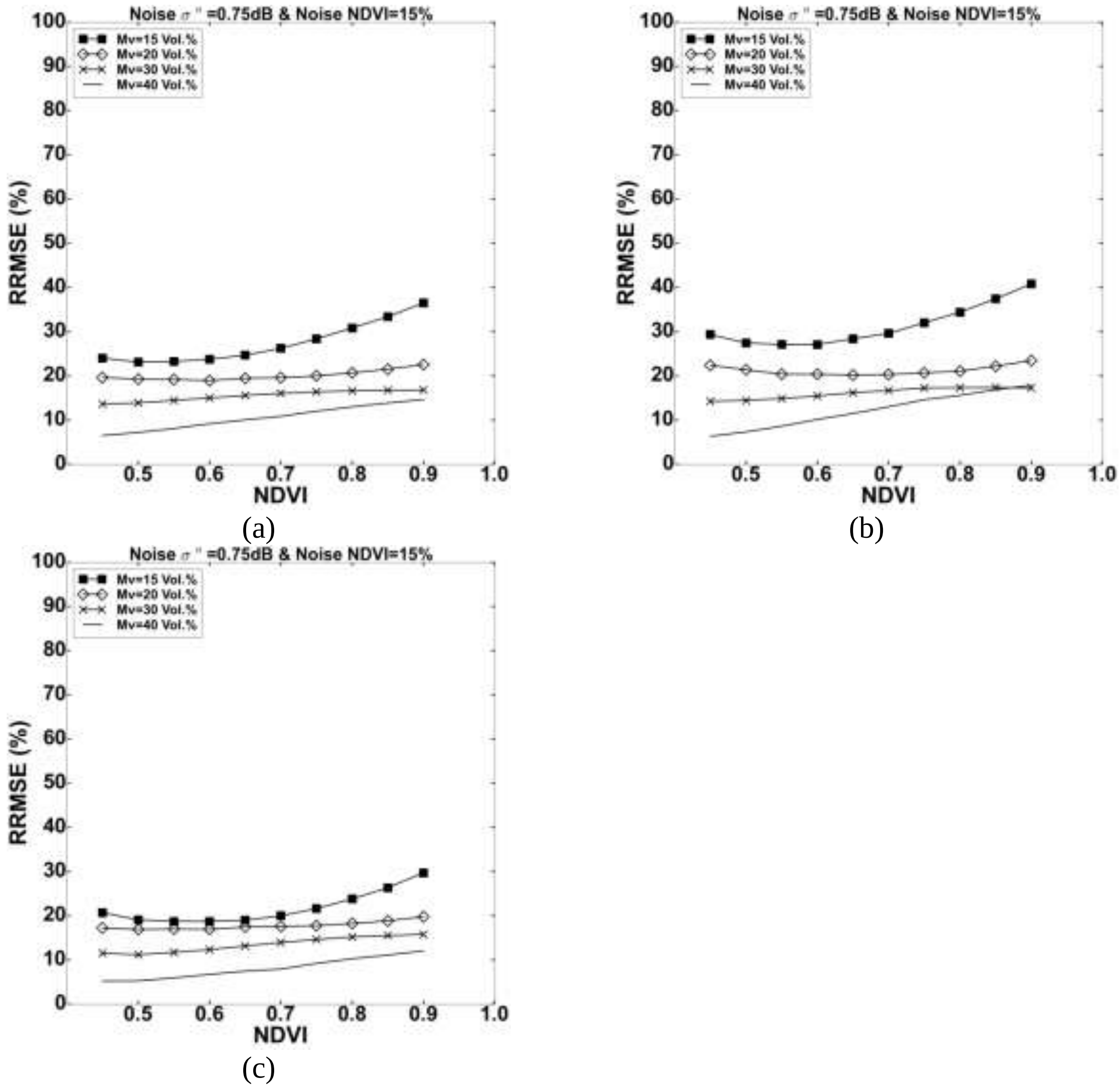


Figure 7. Evolution of the relative RMSE (in percent) **of** M_v estimates (RRMSE=RMSE/ M_v) according to NDVI and M_v. (a) configuration 1: HH and NDVI, (b) configuration 2: HV and NDVI, and (c) configuration 3: HH, HV and NDVI.

The difference between the estimated and reference M_v were also analyzed as a function of NDVI using for each NDVI and all M_v values (Figure 8). For a given NDVI between 0.45 and 0.90, the bias on M_v estimates is similar for radar signal noise of ±0.75 and ±1 dB. The results showed a slight underestimation (lower than approximately 1 Vol.%) of M_v estimates for NDVI values between 0.60 (LAI about 1m^2/m^2) and 0.90 (LAI about 6 m^2/m^2). In addition, a slight overestimation of M_v is observed for a NDVI lower than 0.60 (lower than approximately 1 Vol. %).

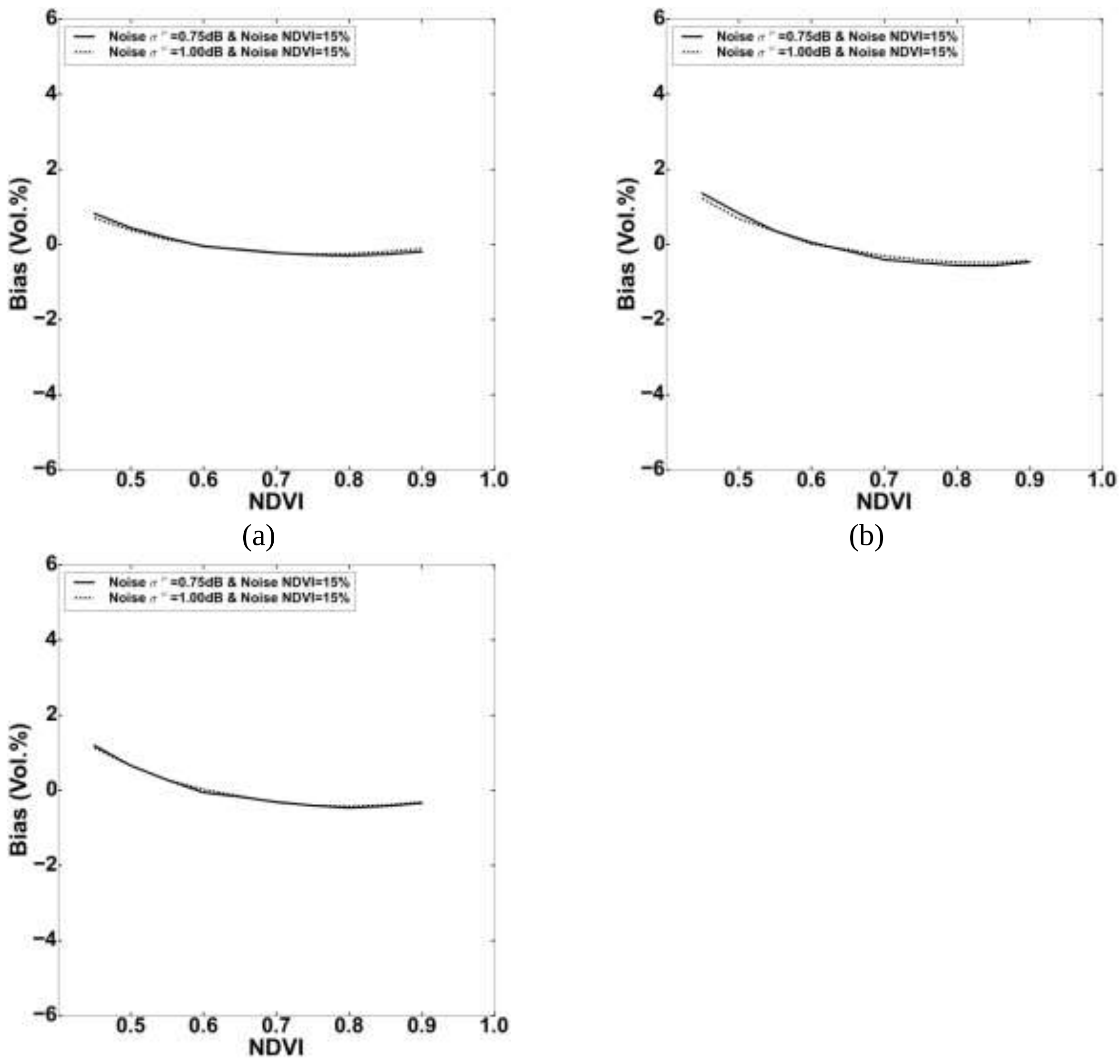
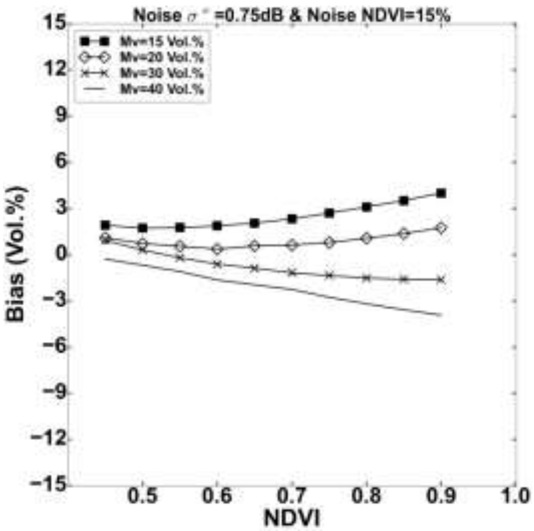
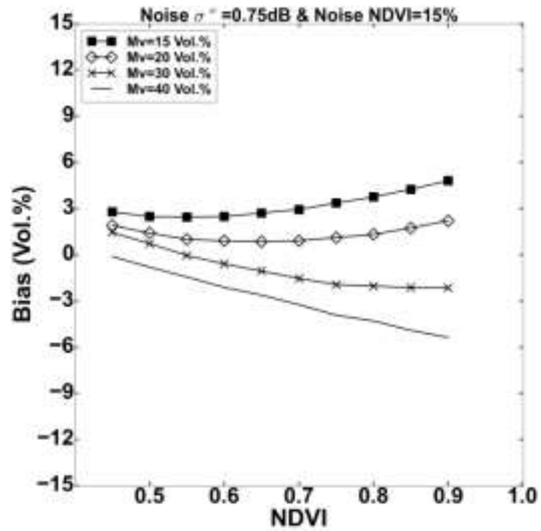


Figure 8. Evolution of the bias (estimated M_v – reference M_v) **of** M_v estimates according to NDVI values. (a) Inversion configuration 1, (b) inversion configuration 2, and (c) inversion configuration 3.

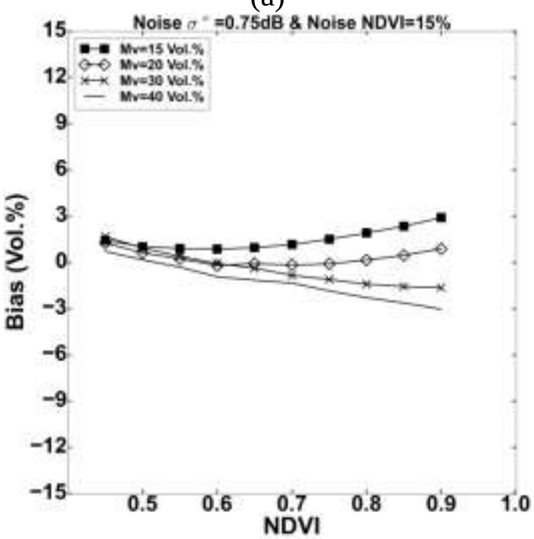
Figure 9 shows the evolution of bias on M_v estimates obtained for the three inversion configurations as a function of the NDVI for some M_v values (15, 20, 30 and 40 Vol.%). For NDVI values lower than approximately 0.65 (LAI about 1.5 m^2/m^2), the bias on M_v estimates is lower than 1.5 Vol.% for M_v between 15 and 40 Vol.%, in the case of configurations 1 and 3. For the inversion configuration 2, the bias reaches 5.4 Vol.% (for M_v= 40 Vol.%). In addition, results showed that the bias increased when the NDVI increased, regardless of the M_v values. This increase was mainly observed when the NDVI was greater than 0.75 (LAI about 3 m^2/m^2) for low M_v values (Figure 9). An overestimation of M_v estimates is mainly observed for M_v values lower than 20 Vol.%, while an underestimation is mainly observed for M_v values higher than 30 Vol.%. Figure 9 also showed that for a given M_v, the bias is lower for configurations 1 and 3. The bias reaches 3.5 Vol.% for configurations 1 and 3 compared to 5 Vol.% for configuration 2 for NDVI = 0.9 and M_v =15 Vol.%. Figure 10 shows an example of box plots calculated for the inversion of configuration 3 and some NDVI values (0.6, 0.7, 0.8 and 0.9).



(a)



(b)



(c)

Figure 9. Evolution of the **bias** (estimated M_v – reference M_v) on M_v estimates according to NDVI and M_v values for noise on the modeled radar signal of 0.75 dB. (a) configuration 1, (b) configuration 2, and (c) configuration 3.

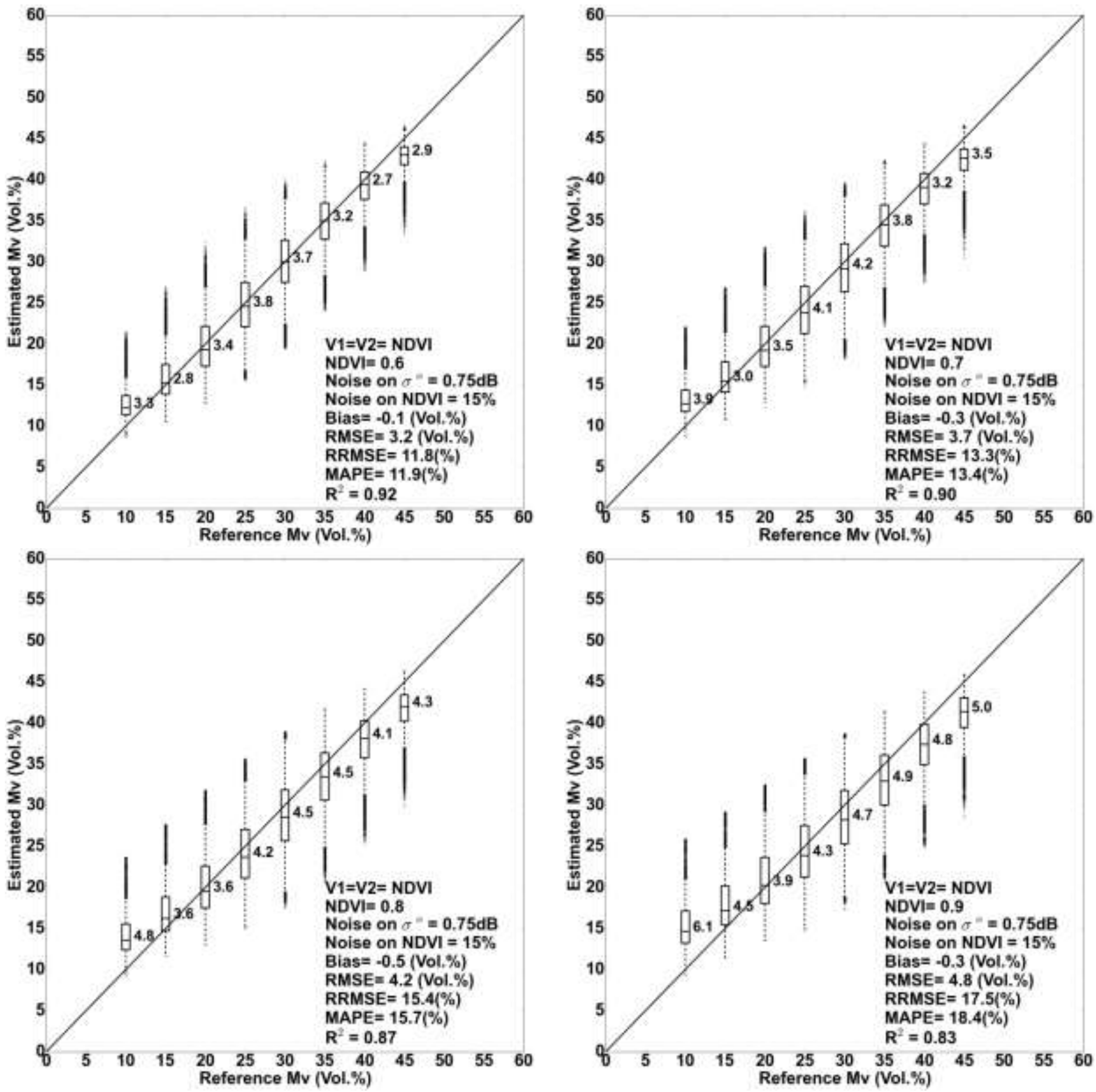


Figure 10. Box plots of M_v estimates retrieved from the synthetic dataset. Neural networks were trained and validated according to configuration 3 (using HH, HV and NDVI). Noise on the modeled radar signal is ±0.75 dB, and noise on NDVI is 15% of the NDVI value. Values to the right of the box plots represent the RMSE on M_v estimates for a given reference M_v.

Moreover, 5-fold cross-validation was used to predict errors on M_v estimates for each inversion configuration performed using the synthetic dataset with LAI, FAPAR, and FCOVER as vegetation descriptors. Table 10 shows statistics (RMSE, RRMSE, MAPE, bias, and R^2) on M_v estimates computed from the validation dataset for reference M_v values between 10 and 45 Vol.% and a LAI between 0 and 6 and FAPAR (as well as FCOER) between 0 and 1. The results show that regardless of the vegetation descriptor used, the RMSE on M_v estimates is lower using HH compared to HV polarization (configuration 1 in comparison to configuration 2). In addition, the use of HH and HV polarizations slightly decreases the RMSE on M_v estimates. Table 10 also shows that the RMSE on M_v estimates increases approximately 1 Vol.% when noise added to the radar signal increases. For each inversion configuration and for a given noise condition on the modeled radar signal, the RMSE on M_v estimates is in the same order with the use of NDVI, LAI, FAPAR, or FCOVER as a vegetation descriptor (Table 10). Finally, the results showed that whatever the vegetation descriptor used, the three inversion configurations provide un-biased M_v estimates.

Table 10. RMSE and Bias on M_v estimates according to the three inversion configurations (RMSE (Vol.%) | RRMSE (%) | MAPE (%) | bias (Vol.%) | R^2). Configuration 1 uses HH and vegetation descriptor, configuration 2 uses HV and vegetation descriptor, and configuration 3 uses HH, HV and vegetation descriptor.

	Noise on σ^0_{tot} : ±0.75 dB	**Noise on σ^0_{tot} : ±1.00 dB**
V1=V2=NDVI Relative noise = 15 %		
Configuration 1	4.5|16.5|17.1|0.0|0.85	5.5|19.8|21.0|0.0|0.78
Configuration 2	5.1|18.5|19.2|0.0|0.81	5.7|20.7|21.8|0.0|0.77
Configuration 3	3.7|13.6|13.7|0.0|0.90	4.5|16.2|16.7|0.0|0.85
V1=V2=LAI Relative noise = 30 %		
Configuration 1	5.6|20.5|20.6|0.0|0.76	6.7|24.5|25.4|0.0|0.65
Configuration 2	7.1|26.0|26.9|0.0|0.61	8.1|29.3|31.2|0.0|0.50
Configuration 3	5.2|0.0|18.9|18.8|0.79	5.8|21.1|21.3|0.0|0.74
V1=V2=FAPAR Relative noise = 20 %		
Configuration 1	5.2|18.9|18.8|0.0|0.79	6.4|23.1|24.1|0.0|0.69
Configuration 2	6.3|22.8|23.3|0.0|0.70	7.3|26.7|28.0|0.0|0.59
Configuration 3	4.4|16.0|15.7|0.0|0.85	5.4|19.7|19.9|0.0|0.78
V1=V2=FCOVER Relative noise = 20 %		
Configuration 1	5.2|18.7|18.8|0.0|0.80	6.5|23.8|24.4|0.0|0.67
Configuration 2	7.1|25.7|26.7|0.0|0.62	7.8|28.3|30.0|0.0|0.54
Configuration 3	4.7|16.9|16.8|0.0|0.84	5.7|20.7|20.9|0.0|0.75

4.2.2 Real dataset

The capacity of the developed Neural Networks (NNs) to correctly estimate the soil moisture was then tested using the real dataset. The NNs applied to the real validation dataset are those which have been trained and validated using the synthetic dataset. NDVI, LAI, FAPAR and FCOVER derived from optical images were used as the input vegetation descriptors for the trained NNs. Inversion results obtained with the NDVI derived from optical images as the vegetation descriptor will be provided in detail, whereas the results based on the LAI, FAPAR, FCOVER derived from optical images as the vegetation descriptor will be briefly described.

First, statistics (RMSE, RRMSE, MAPE, bias, R^2) on M_v estimates were also computed for all NDVI observations (Table 11). Slightly better statistics were observed with the noise on a modeled radar signal of ±1.00 dB. With the noise of ±1.00 dB, the RMSE is 4.5, 6.0 and 5.5 Vol.% in configuration 1, 2 and 3, respectively. Moreover, a slight underestimation (about -0.1 Vol.%) was observed in configuration 1 for the noise conditions of ±0.75 dB and ±1.00 dB. For configurations 2 and 3, an underestimation of M_v estimates was observed (about -1.4 Vol.% in configuration 2 and -1 Vol.% in configuration 3).

Table 11. statics on M_v estimates according to the three inversion configurations (RMSE (Vol.%) | RRMSE (%) | MAPE (%) | bias Vol.% | R^2 | samples). Configuration 1 uses HH and NDVI, configuration 2 uses HV and NDVI, and configuration 3 uses HH, HV and NDVI. Relative noise on the NDVI=15%. Real SAR measurements and the LAI derived from optical images were used to estimate M_v.

	Noise on σ^0_{tot} : ±0.75 dB	**Noise on σ^0_{tot} : ±1.00 dB**
	NDVI = [0.45-0.90]	**NDVI = [0.45-0.90]**
Configuration 1 (HH and NDVI)	4.9|18.4|16.4|-0.1|0.60|93	4.5|17.0|15.5|-0.1|0.63|93
Configuration 2 (HV and NDVI)	6.8|25.7|23.1|-1.3|0.37|93	6.0|22.6|19.8|-1.3|0.42|93
Configuration 3 (HH, HV and NDVI)	6.2|23.5|21.2|-0.8|0.49|93	5.5|20.5|18.0|-0.9|0.53|93

Next, the statistics were computed from the real dataset of validation for NDVI classes of 0.05 (NDVI was derived from optical images are between 0.45 and 0.9). The results showed that the RMSE on M_v estimates was in the same order for NDVI classes between 0.45 and 0.75 (LAI about $3m^2/m^2$) on the one hand (difference lower than 1 Vol.%), and on the other hand for NDVI classes between 0.75 (LAI about 3 m^2/m^2) and 0.90 (LAI about 6 m^2/m^2). Therefore, the results of M_v estimates were presented for two classes of NDVI: NDVI lower and higher than 0.75 (Table 12). The comparison between estimated M_v and M_v ground-truthed measurements is given in Figures 11 and 12. RMSE and bias on M_v estimates are lower with the noise condition on the modeled radar signal of ±1 dB.

RMSE of 3.6 (RRMSE and MAPE about 12%), 5.4 (RRMSE and MAPE about 18%), and 4.4 (RRMSE and MAPE about 15%) Vol.% were observed for configurations 1, 2 and 3, respectively, in the case of a NDVI lower than 0.75 and for modeled radar signal noise of ±1 dB

(Table 12, Figure 11). For a NDVI higher than 0.75, the RMSE on M_v estimates is 6.1 (RRMSE and MAPE about 24%), 7.1 (RRMSE and MAPE about 28%) and 7.3 (RRMSE and MAPE about 29%) Vol.%, respectively, for configurations 1, 2 and 3 and for the noise on the modeled radar signal of ±1 dB (Table 12, Figure 11). Moreover, results showed that for a NDVI < 0.75 the trained NNs provide M_v estimates with slight bias (0.2, -1.7, and -0.9 Vol.% in configurations 1, 2 and 3, respectively) (Table 12, Figure 11). For a NDVI > 0.75, an slight bias (between -1 and 0.1 Vol.%) was observed for the noise on the radar signal of ±1 dB, with the lower value for the inversion using HH and NDVI (0.1 Vol.%) (Table 12, Figure 11).

Table 12. RMSE and bias on M_v estimates according to the three inversion configurations (RMSE (Vol.%) | RRMSE (%) | MAPE (%) | bias Vol.% | R^2 | samples). Configuration 1 uses HH and NDVI, configuration 2 uses HV and NDVI, and configuration 3 uses HH, HV and NDVI. Relative noise on the NDVI=15%. Real SAR measurements and NDVI derived from optical images were used to estimate M_v.

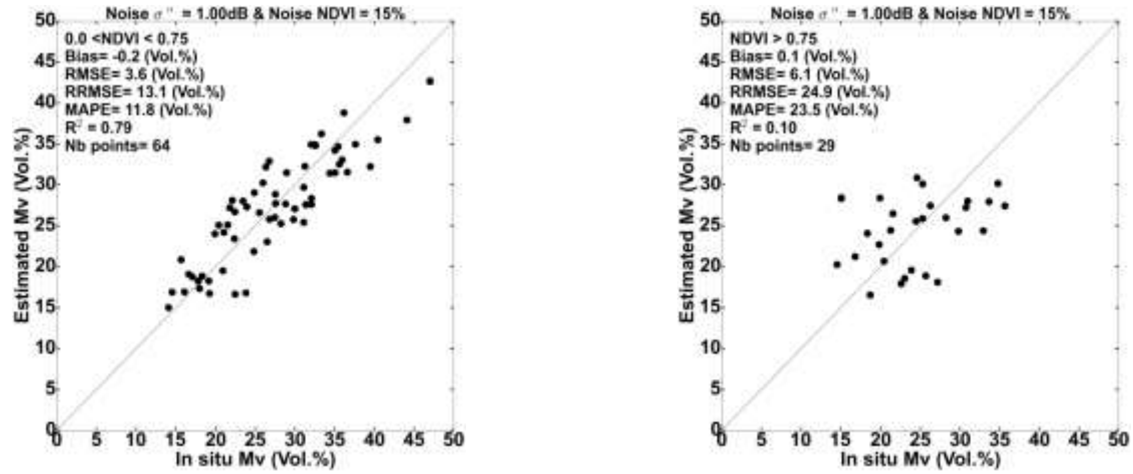
<table>
<tr><th rowspan="2"></th><th colspan="2">Noise on σ^0_{tot} : ±0.75 dB</th><th colspan="2">Noise on σ^0_{tot} : ±1.00 dB</th></tr>
<tr><th>NDVI < 0.75</th><th>NDVI > 0.75</th><th>NDVI < 0.75</th><th>NDVI > 0.75</th></tr>
<tr><td>Configuration 1</td><td>3.8|13.9|12.4 |0.0|0.77|64</td><td>6.6|27.3|25.3 |-0.3|0.07|29</td><td>3.6|13.1|11.8 |-0.2|0.79|64</td><td>6.1|24.9|23.5 |0.1|0.10|29</td></tr>
<tr><td>Configuration 2</td><td>6.0|21.7|19.8 |-1.5|0.52|64</td><td>8.4|34.5|30.6| -0.8|0.04|29</td><td>5.4|19.7|17.0| -1.7|0.56|64</td><td>7.1|29.2|26.1 |-0.5|0.07|29</td></tr>
<tr><td>Configuration 3</td><td>5.0|18.2|16.8 |-0.7|0.67|64</td><td>8.3|34.2|30.9 |-1.1|0.04|29</td><td>4.4|15.8|13.9 |-0.9|0.71|64</td><td>7.3|30.1|26.9 |-1.0|0.06|29</td></tr>
</table>

Moreover, the SAR real validation dataset was inverted to estimate soil moisture by means of trained NNs with the use of each of the vegetation descriptors derived from optical images LAI, FAPAR, and FCOVER). Table 13 shows the (RMSE, RRMSE, MAPE, bias, R^2) on M_v estimates in the three inversion configurations for two classes of NDVI: NDVI lower and higher than 0.75 (LAI about $3m^2/m^2$). The results showed that the RMSE (as well as RRMSE, MAPE) on M_v estimates are almost similar, regardless of which vegetation descriptors derived from optical images were used (NDVI, LAI, FAPAR, or FCOVER) (Table 13).

In conclusion, the use of HH polarization in addition to a vegetation descriptor derived from optical images (Configuration 1) provides a better estimation of the soil moisture with a RMSE approximately 4.5 and 7.0 Vol.% for a NDVI lower and higher than 0.75 (LAI about 3 m^2/m^2), respectively. The use of HV in addition to HH slightly lowers the precision of M_v estimates.

Table 13. Statics on M_v estimates according to the three inversion configurations (RMSE (Vol.%) | RRMSE (%) | MAPE (%) | bias Vol.% | R^2 | samples). Configuration 1 uses HH and NDVI, configuration 2 uses HV and NDVI, and configuration 3 uses HH, HV and NDVI. Real SAR measurements, and NDVI, LAI, FAPAR and FCOVER derived from optical images were used to estimate M_v.

<table>
<tr><th rowspan="2"></th><th colspan="2">Noise on σ^0_{tot} : ±0.75 dB</th><th colspan="2">Noise on σ^0_{tot} : ±1.00 dB</th></tr>
<tr><th>NDVI <0.75</th><th>NDVI > 0.75</th><th>NDVI < 0.75</th><th>NDVI> 0.75</th></tr>
<tr><td colspan="5">V1=V2=NDVI
Relative noise = 15 %</td></tr>
<tr><td>Configuration 1</td><td>3.8|13.9|12.4
|0.0|0.77|64</td><td>6.6|27.3|25.3
|-0.3|0.07|29</td><td>3.6|13.1|11.8
|-0.2|0.79|64</td><td>6.1|24.9|23.5
|0.1|0.10|29</td></tr>
<tr><td>Configuration 2</td><td>6.0|21.7|19.8
|-1.5|0.52|64</td><td>8.4|34.5|30.6
|-0.8|0.04|29</td><td>5.4|19.7|17.0
|-1.7|0.56|64</td><td>7.1|29.2|26.1
|-0.5|0.07|29</td></tr>
<tr><td>Configuration 3</td><td>5.0|18.2|16.8
|-0.7|0.67|64</td><td>8.3|34.2|30.9
|-1.1|0.04|29</td><td>4.4|15.8|13.9
|-0.9|0.71|64</td><td>7.3|30.1|26.9
|-1.0|0.06|29</td></tr>
<tr><td colspan="5">V1=V2=LAI
Relative noise = 30 %</td></tr>
<tr><td>Configuration 1</td><td>4.7|17.1|15.9
|-0.0|0.65|64</td><td>7.3|29.7|27.2
|-1.5|0.02|29</td><td>4.5|16.3|15.3
|0.6|0.67|64</td><td>7.5|30.6|28.9
|0.3|0.00|29</td></tr>
<tr><td>Configuration 2</td><td>7.5|27.1|23.8
|-1.1|0.36|64</td><td>10.0|41.0|34.9
|-3.2|0.00|29</td><td>7.1|25.8|22.2
|-1.1|0.35|64</td><td>9.0|36.8|31.4
|-2.5|0.00|29</td></tr>
<tr><td>Configuration 3</td><td>5.6|20.1|17.3
|-0.9|0.57|64</td><td>8.4|34.5|30.5
|-2.5|0.00|29</td><td>5.7|20.7|17.7
|-0.5|0.55|64</td><td>8.7|35.7|31.1
|-2.2|0.00|29</td></tr>
<tr><td colspan="5">V1=V2=FAPAR
Relative noise = 20 %</td></tr>
<tr><td>Configuration 1</td><td>5.0|18.1|16.2
|0.5|0.63|64</td><td>7.9|32.6|30.3
|-0.7|0.00|29</td><td>4.9|17.8|16.6
|1.2|0.63|64</td><td>7.4|30.4|29.0
|0.7|0.00|29</td></tr>
<tr><td>Configuration 2</td><td>8.1|29.2|25.8
|-0.0|0.34|64</td><td>10.9|44.6|39.3
|-3.1|0.00|29</td><td>7.2|26.2|22.4
|-0.1|0.34|64</td><td>9.1|37.2|32.5
|-1.7|0.00|29</td></tr>
<tr><td>Configuration 3</td><td>6.4|23.3|20.5
|0.4|0.52|64</td><td>9.5|38.9|34.2
|-2.4|0.00|29</td><td>6.2|22.4|19.5
|0.9|0.51|64</td><td>8.8|36.1|32.4
|-1.3|0.01|29</td></tr>
<tr><td colspan="5">V1=V2=FCOVER
Relative noise = 20 %</td></tr>
<tr><td>Configuration 1</td><td>5.1|18.6|16.5
|0.8|0.62|64</td><td>8.0|33.0|30.7
|-0.7|0.01|29</td><td>5.0|18.3|17.1
|0.9|0.62|64</td><td>6.8|27.5|25.2
|-0.4|0.03|29</td></tr>
<tr><td>Configuration 2</td><td>7.6|27.5|23.7
|-0.6|0.34|64</td><td>10.0|40.9|35.1
|-3.3|0.01|29</td><td>7.2|25.9|21.8
|-0.7|0.34|64</td><td>9.1|37.2|31.9
|-2.5|0.01|29</td></tr>
<tr><td>Configuration 3</td><td>6.0|21.6|19.0
|0.3|0.55|64</td><td>9.2|37.6|32.7
|-2.5|0.01|29</td><td>5.9|21.4|18.2
|0.2|0.54|64</td><td>8.4|34.5|30.1
|-1.9|0.01|29</td></tr>
</table>



(a) (b)

Figure 11. Retrieved soil moisture using configuration 1 versus ground-truthed measurements for NDVI lower and higher than 0.75 (a, and b respectively). Noise on radar signal = ±1dB. Bias = estimated M_v - reference M_v.

5. Conclusion

Inversion results of the synthetic dataset showed that the best M_v estimates were obtained with the use of the X-band radar signal in HH polarization or in using both HH and HV polarizations, in addition to one vegetation descriptor derived from optical images. However, the use of HV in addition to one vegetation descriptor derived from optical images degrades the precision on M_v estimates. Moreover, results showed that the RMSE on M_v estimates is slightly sensitive to additive noise on the modelled radar signal. The RMSE increases approximately 1 Vol.% when the noise of the radar signal increases from ±0.75 dB to ±1.00 dB. For all NDVI values, the RMSE on M_v estimates (M_v between 10 and 45 Vol.%) was approximately 5.0 Vol.% (RRMSE and MAPE about 19 %) in configurations 1 and 3. Similar values of the RMSE (as well as RRMSE and MAPE) on M_v estimates were obtained with the use of LAI, FAPAR, and FCOVER as the vegetation descriptor. The accuracy of M_v estimates degrades (i.e., an increase in the RMSE, RRMSE, and MAPE) with vegetation growth (i.e., an increase in the NDVI). As an example, in configuration 3 (HH, HV and NDVI), the RMSE on M_v estimates increases from

3.6 Vol.% (RRMSE about 13%) for NDVI of 0.45 to 5.7 Vol.% (RRMSE about 21 %) for a NDVI of 0.9.

From the real validation dataset (53% of the real dataset), the soil moisture estimation using the X-band SAR data in addition to one vegetation descriptor derived from optical images allows better results with HH polarization than with HV or both HH and HV. With HH and NDVI information derived from optical images, the accuracy on the soil moisture estimation was 3.6 Vol.% (RRMSE and MAPE about 13%) for NDVI between 0 and 0.75 (LAI about 3 m^2/m^2) and 6.1 Vol.% (RRMSE and MAPE about 25%) when the NDVI of the grassland was between 0.75 and 0.9 (LAI about 6 m^2/m^2). Similar results were obtained regardless the vegetation descriptor used.

With the arrival of new satellites, such as SENTINEL-1A (launched on 3 April 2014), in addition to future satellites SENTINEL-1B, SENTINEL-2A (optical sensor), and SENTINEL-2B, it will be possible to obtain SAR (C-band) and optical remote sensing data covering global areas with high spatial and temporal resolutions (2 days with 2 SENTINEL-1 satellites, and 5 days for 2 SENTINEL-2 satellites at 10 m spatial resolution). Combining SENTINEL-1 data with optical images (SENTINEL-2, LANDSAT-7/8) will allow more precise estimation of M_v because the radar signal penetration depth into vegetation cover is higher in the C-band compared to the X-band. This work is in the context of preparing for SENTINEL 1 and 2 missions.

This study demonstrated that the use of NNs technique to invert X-band SAR backscattering coefficients allows the estimation of soil moisture with acceptable accuracy (RMSE of 3.6 Vol.% for a NDVI lower than 0.75). Current remote sensing sensors (optical and SAR) and those available in the near future (spatial resolution better than 10 m) will allow the estimation of soil moisture at a field scale with high temporal resolution (better than weekly). Vegetation biophysical parameters (i.e., LAI) and soil moisture that can be derived from optical and SAR images could be useful to calibrate crop models for better irrigation management and crop growth monitoring. Indeed, combining optical and SAR data would enhance the relevance of remote sensing data for water and crop monitoring.

Acknowledgements

This research was supported by the French Space Study Center (CNES, DAR 2014 TOSCA) and the Islamic Development Bank (PhD Scholarship of M. Mohammad El Hajj). Field experiments were carried out within the SicMed-Crau program. The CSK images used in this analysis were supported by public funds received in the GEOSUD framework, a project (ANR-10-EQPX-20) of the "Investissements d'Avenir" program managed by the French National Research Agency. The authors wish to thank the German Space Agency (DLR) for kindly providing the TSX images under proposal HYD0007. We also wish to thank the EMMAH unit (INRA) for providing meteorological data and the technical teams of ASI and DLR for providing answers regarding the performances of CSK and TSX. Finally, we would like to thank Olivier Hagolle (CNES-CESBIO: Centre National d'Etudes Spatiales-Centre d'Etudes Spatiales de la BIOsphère) for providing calibrated SPOT 4 images in the framework of Take 5 experiments.

References

Agapiou, A., Hadjimitsis, D.G., Papoutsa, C., Alexakis, D.D., and Papadavid, G. (2011). The importance of accounting for atmospheric effects in the application of NDVI and interpretation of satellite imagery supporting archaeological research: the case studies of Palaepaphos and Nea Paphos sites in Cyprus. Remote Sens. *3*, 2605–2629.

Agenzia Spaziale Italiana (2007). COSMO-SkyMed System Description & User Guide.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO Rome *300*, 6541.

Attema, E.P.W., and Ulaby, F.T. (1978). Vegetation modeled as a water cloud. Radio Sci. *13*, 357–364.

Aubert, M., Baghdadi, N., Zribi, M., Douaoui, A., Loumagne, C., Baup, F., El Hajj, M., and Garrigues, S. (2011). Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust. Remote Sens. Environ. *115*, 1801–1810.

Baghdadi, N., Cresson, R., El Hajj, M., Ludwig, R., and La Jeunesse, I. (2012). Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. Hydrol. Earth Syst. Sci. *16*, 1607–1621.

Balenzano, A., Mattia, F., Satalino, G., and Davidson, M. (2011). Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops. Sel. Top. Appl. Earth Obs. Remote Sens. IEEE J. Of *4*, 439–450.

Baret, F., and Guerif, M. (2006). Remote detection and quantification of plant stress: opportunities remote sensing observations. In Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology, (ELSEVIER SCIENCE INC 360 PARK AVE SOUTH, NEW YORK, NY 10010-1710 USA), pp. S148–S148.

Baret, F., and Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sens. Environ. *35*, 161–173.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., and Samain, O. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. Remote Sens. Environ. *110*, 275–286.

Botha, E.J., Leblon, B., Zebarth, B.J., and Watmough, J. (2010). Non-destructive estimation of wheat leaf chlorophyll content from hyperspectral measurements through analytical model inversion. Int. J. Remote Sens. *31*, 1679–1697.

Bottraud, J.C., Bornand, M., and Servat, E. (1984). Mesures de résistivité appliquées à la cartographie en pédologie. Sci. Sol *4*, 279–294.

Brereton, A.J., and Hope-Cawdery, M. (1988). Drumlin soils: the depression of herbage yield by shallow water table depth. Ir. J. Agric. Res. 167–178.

Brown, S.C., Quegan, S., Morrison, K., Bennett, J.C., and Cookmartin, G. (2003). High-resolution measurements of scattering in wheat canopies-Implications for crop parameter retrieval. Geosci. Remote Sens. IEEE Trans. On *41*, 1602–1610.

Bsaibes, A., Courault, D., Baret, F., Weiss, M., Olioso, A., Jacob, F., Hagolle, O., Marloie, O., Bertrand, N., and Desfond, V. (2009). Albedo and LAI estimates from FORMOSAT-2 data for crop monitoring. Remote Sens. Environ. *113*, 716–729.

Carlson, T.N., and Ripley, D.A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. *62*, 241–252.

Carlson, T.N., Gillies, R.R., and Perry, E.M. (1994). A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. Remote Sens. Rev. *9*, 161–173.

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., and Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. Remote Sens. Environ. *77*, 22–33.

Chai, S.-S., Walker, J.P., Makarynskyy, O., Kuhn, M., Veenendaal, B., and West, G. (2009). Use of soil moisture variability in artificial neural network retrieval of soil moisture. Remote Sens. *2*, 166–190.

Champion, I. (1991). Etude et mise au point de modèles Semi-empiriques représentant la réponse de couverts végétaux dans le domaine hyperfréquence. Complémentarité avec le domaine optique. Thèse, Université Paris VII. France.

Champion, I., and Guyot, G. (1991). Generalized formulation for semi-empirical radar models representing crop backscattering. ESA Phys. Meas. Signat. Remote Sens. *1*, 269–272.

Claverie, M., Vermote, E.F., Weiss, M., Baret, F., Hagolle, O., and Demarez, V. (2013). Validation of coarse spatial resolution LAI and FAPAR time series over cropland in southwest France. Remote Sens. Environ. *139*, 216–230.

Coletta, A., Angino, G., Battazza, F., Caltagirone, F., Impagnatiello, F., Valentini, G., Capuzi, A., Fagioli, S., and Leonardi, R. (2007). COSMO-SkyMed program: Utilization and description of an advanced space EO dual-use asset. In Proc. Envisat Symp, pp. 23–27.

Courault, D., Bsaibes, A., Kpemlie, E., Hadria, R., Hagolle, O., Marloie, O., Hanocq, J.-F., Olioso, A., Bertrand, N., and Desfonds, V. (2008). Assessing the potentialities of FORMOSAT-2 data for water and crop monitoring at small regional scale in South-Eastern France. Sensors *8*, 3460–3481.

Courault, D., Hadria, R., Ruget, F., Olioso, A., Duchemin, B., Hagolle, O., and Dedieu, G. (2010). Combined use of FORMOSAT-2 images with a crop model for biomass and water monitoring of permanent grassland in Mediterranean region. Hydrol. Earth Syst. Sci. Discuss. *7*, 1731–1744.

Darvishzadeh, R., Skidmore, A., Schlerf, M., and Atzberger, C. (2008a). Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. Remote Sens. Environ. *112*, 2592–2604.

Darvishzadeh, R., Skidmore, A., Atzberger, C., and van Wieren, S. (2008b). Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture. Int. J. Appl. Earth Obs. Geoinformation *10*, 358–373.

Del Frate, F., and Solimini, D. (2004). On neural network algorithms for retrieving forest biomass from SAR data. Geosci. Remote Sens. IEEE Trans. On *42*, 24–34.

Del Frate, F., Ferrazzoli, P., and Schiavon, G. (2003). Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks. Remote Sens. Environ. *84*, 174–183.

De Roo, R.D., Du, Y., Ulaby, F.T., and Dobson, M.C. (2001). A semi-empirical backscattering model at L-band and C-band for a soybean canopy with soil moisture inversion. Geosci. Remote Sens. IEEE Trans. On *39*, 864–872.

Dubois, P.C., Van Zyl, J., and Engman, T. (1995). Measuring soil moisture with imaging radars. Geosci. Remote Sens. IEEE Trans. On *33*, 915–926.

Duveiller, G., Weiss, M., Baret, F., and Defourny, P. (2011). Retrieving wheat Green Area Index during the growing season from optical time series measurements based on neural network radiative transfer inversion. Remote Sens. Environ. *115*, 887–896.

Eineder, M., Fritz, T., Mittermayer, J., Roth, A., Boerner, E., and Breit, H. (2008). TerraSAR-X Ground Segment, Basic Product Specification Document. (Munich, Germany: DTIC Document, Cluster Applied Remote Sensing. Available online: https://tandemx-science.dlr.de/pdfs/TX-GS-DD-3302_Basic-Products-Specification-Document_V1.6.pdf (accessed on 22 April 2011).).

El Hajj, M., Bégué, A., Lafrance, B., Hagolle, O., Dedieu, G., and Rumeau, M. (2008). Relative radiometric normalization and atmospheric correction of a SPOT 5 time series. Sensors *8*, 2774–2791.

Fensholt, R., Sandholt, I., and Rasmussen, M.S. (2004). Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. Remote Sens. Environ. *91*, 490–507.

Fieuzal, R., Duchemin, B., Jarlan, L., Zribi, M., Baup, F., Merlin, O., Hagolle, O., and Garatuza-Payan, J. (2011). Combined use of optical and radar satellite data for the monitoring of irrigation and soil moisture of wheat crops. Hydrol. Earth Syst. Sci. *15*, 1117–1129.

Fung, A.K., Li, Z., and Chen, K.S. (1992). Backscattering from a randomly rough dielectric surface. Geosci. Remote Sens. IEEE Trans. On *30*, 356–369.

Gherboudj, I., Magagi, R., Berg, A.A., and Toth, B. (2011). Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data. Remote Sens. Environ. *115*, 33–43.

Guerschman, J.P., Hill, M.J., Renzullo, L.J., Barrett, D.J., Marks, A.S., and Botha, E.J. (2009). Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. Remote Sens. Environ. *113*, 928–945.

Hagolle, O., Dedieu, G., Mougenot, B., Debaecker, V., Duchemin, B., and Meygret, A. (2008). Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images. Remote Sens. Environ. *112*, 1689–1701.

He, B., Xing, M., and Bai, X. (2014). A Synergistic Methodology for Soil Moisture Estimation in an Alpine Prairie Using Radar and Optical Satellite Data. Remote Sens. *6*, 10966–10985.

Hong, S., Malaterre, P.-O., Belaud, G., and Dejean, C. (2013). Optimization of water distribution for open-channel irrigation networks. Press J. Hydroinformatics.

Hosseini, M., and Saradjian, M.R. (2011). Soil moisture estimation based on integration of optical and SAR images. Can. J. Remote Sens. *37*, 112–121.

Iorio, M., Mecozzi, R., and Torre, A. (2010). Cosmo SkyMed: antenna elevation pattern data evaluation. Ital. J Remote Sens *42*, 69–77.

Leenhardt, D., Trouvat, J.-L., Gonzalès, G., Pérarnaud, V., Prats, S., and Bergez, J.-E. (2004). Estimating irrigation demand for water management on a regional scale: I. ADEAUMIS, a simulation platform based on bio-decisional modelling and spatial information. Agric. Water Manag. *68*, 207–232.

Li, F., Chen, W., Zeng, Y., Zhao, Q., and Wu, B. (2014). Improving Estimates of Grassland Fractional Vegetation Cover Based on a Pixel Dichotomy Model: A Case Study in Inner Mongolia, China. Remote Sens. *6*, 4705–4722.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. *11*, 431–441.

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., and Lim, T.K. (2013). LEDAPS calibration, reflectance, atmospheric correction preprocessing code, version 2. Model Prod.

Mattia, F., Le Toan, T., Picard, G., Posa, F.I., D'Alessio, A., Notarnicola, C., Gatti, A.M., Rinaldi, M., Satalino, G., and Pasquariello, G. (2003). Multitemporal C-band radar measurements on wheat fields. Geosci. Remote Sens. IEEE Trans. On *41*, 1551–1560.

Mérot, A. (2007). Analyse et modélisation du fonctionnement biophysique et décisionnel d'un système prairial irrigué-Application aux prairies plurispécifiques de Crau en vue de l'élaboration d'un Outil d'Aide à la Décision. Thèse, Ecole nationale superieure agronomique de montpellier-AGRO M. France.

Merot, A., Wery, J., Isberie, C., and Charron, F. (2008). Response of a plurispecific permanent grassland to border irrigation regulated by tensiometers. Eur. J. Agron. *28*, 8–18.

North, P.R. (2002). Estimation of FAPAR, LAI, and vegetation fractional cover from ATSR-2 imagery. Remote Sens. Environ. *80*, 114–121.

Notarnicola, C., Angiulli, M., and Posa, F. (2006). Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas. Geosci. Remote Sens. IEEE Trans. On *44*, 925–935.

Oh, Y. (2004). Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. Geosci. Remote Sens. IEEE Trans. On *42*, 596–601.

Paloscia, S., Pampaloni, P., Pettinato, S., and Santi, E. (2008). A comparison of algorithms for retrieving soil moisture from ENVISAT/ASAR images. Geosci. Remote Sens. IEEE Trans. On *46*, 3274–3284.

Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., and Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. Remote Sens. Environ. *134*, 234–248.

Patel, P., and Srivastava, H.S. (2013). Ground truth planning for synthetic aperture radar (SAR): addressing various challenges using statistical approach. Int. J. Adv. Remote Sens. GIS Geogr. *1*, 1–17.

Peake, W.H. (1959). Interaction of electromagnetic waves with some natural surfaces. Antennas Propag. IRE Trans. On *7*, 324–329.

Prakash, R., Singh, D., and Pathak, N.P. (2012). A fusion approach to retrieve soil moisture with SAR and optical data. Sel. Top. Appl. Earth Obs. Remote Sens. IEEE J. Of *5*, 196–206.

Prévot, L., Champion, I., and Guyot, G. (1993). Estimating surface soil moisture and leaf area index of a wheat canopy using a dual-frequency (C and X bands) scatterometer. Remote Sens. Environ. *46*, 331–339.

Rahman, H., and Dedieu, G. (1994). SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Int. J. Remote Sens. *15*, 123–143.

Saastamoinen, J. (1972). Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. Use Artif. Satell. Geod. 247–251.

Said, S., Kothyari, U.C., and Arora, M.K. (2012). Vegetation effects on soil moisture estimation from ERS-2 SAR images. Hydrol. Sci. J. *57*, 517–534.

Santi, E., Paloscia, S., Pettinato, S., Notarnicola, C., Pasolli, L., and Pistocchi, A. (2013). Comparison between SAR soil moisture estimates and hydrological model simulations over the Scrivia test site. Remote Sens. *5*, 4961–4976.

Schwerdt, M., Bräutigam, B., Bachmann, M., and Döring, B. (2008). TerraSAR-X calibration results. In Synthetic Aperture Radar (EUSAR), 2008 7th European Conference on, (Graf-Zeppelin-Haus, Friedrichshafen, Germany: VDE), pp. 1–4.

Sikdar, M., and Cumming, I. (2004). A modified empirical model for soil moisture estimation in vegetated areas using SAR data. In Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International, (Anchorage, AK: IEEE), pp. 803–806.

Simoniello, T., Cuomo, V., Lanfredi, M., Lasaponara, R., and Macchiato, M. (2004). On the relevance of accurate correction and validation procedures in the analysis of AVHRR-NDVI time series for long-term monitoring. J. Geophys. Res. Atmospheres 1984–2012 *109*.

Soon-Koo Kweon, Ji-Hwan Hwang, and Yisok Oh (2012). COSMO SkyMed AO projects -soil moisture detection for vegetation fields based on a modified water-cloud model using COSMO-SkyMed SAR data. Geosci. Remote Sens. Symp. IGARSS Munich Ger. 2012 IEEE Int. 1204–1207.

Srivastava, H.S., Patel, P., Manchanda, M.L., and Adiga, S. (2003). Use of multiincidence angle RADARSAT-1 SAR data to incorporate the effect of surface roughness in soil moisture estimation. IEEE Trans. Geosci. Remote Sens. *41*, 1638–1640.

Srivastava, H.S., Patel, P., Sharma, Y., and Navalgund, R.R. (2009). Large-area soil moisture estimation using multi-incidence-angle RADARSAT-1 SAR data. Geosci. Remote Sens. IEEE Trans. On *47*, 2528–2535.

Srivastava, H.S., Patel, P., Sharma, K.P., Krishnamurthy, Y.V.N., and Dadhwal, V.K. (2011). A semi-empirical modelling approach to calculate two-way attenuation in radar backscatter from soil due to crop cover. Curr. Sci. *100*, 1871–1874.

Torre, A., Calabrese, D., and Porfilio, M. (2011). COSMO-SkyMed: Image quality achievements. In Recent Advances in Space Technologies (RAST), 2011 5th International Conference on, (Istanbul, Turkey: IEEE), pp. 861–864.

Ulaby, F.T., Allen, C.T., Eger Iii, G., and Kanemasu, E. (1984). Relating the microwave backscattering coefficient to leaf area index. Remote Sens. Environ. *14*, 113–133.

Ulaby, F.T., Moore, R.K., and Fung, A.K. (1986). Microwave Remote Sensing: Active and Passive, vol. III, Volume Scattering and Emission Theory, Advanced Systems and Applications. Inc Dedham Mass. USA Norwood Mass. Artech House 1797–1848.

Vermote, E.F., El Saleous, N.Z., and Justice, C.O. (2002). Atmospheric correction of MODIS data in the visible to middle infrared: first results. Remote Sens. Environ. *83*, 97–111.

Wang, S.G., Li, X., Han, X.J., and Jin, R. (2011). Estimation of surface soil moisture and roughness from multi-angular ASAR imagery in the Watershed Allied Telemetry Experimental Research (WATER). Hydrol. Earth Syst. Sci. *15*, 1415–1426.

Yang, G., Shi, Y., Zhao, C., and Wang, J. (2012). Estimation of soil moisture from multi-polarized SAR data over wheat coverage areas. In Agro-Geoinformatics (Agro-Geoinformatics), 2012 First International Conference on, (Shanghai, China: IEEE), pp. 1–5.

Yu, F., and Zhao, Y. (2011). A new semi-empirical model for soil moisture content retrieval by ASAR and TM data in vegetation-covered areas. Sci. China Earth Sci. *54*, 1955–1964.

Zribi, M., Baghdadi, N., Holah, N., and Fafin, O. (2005). New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion. Remote Sens. Environ. *96*, 485–496.

Zribi, M., Chahbi, A., Shabou, M., Lili-Chabaane, Z., Duchemin, B., Baghdadi, N., Amri, R., and Chehbouni, A. (2011). Soil surface moisture estimation over a semi-arid region using ENVISAT ASAR radar data for soil evaporation evaluation. Hydrol. Earth Syst. Sci. *15*.

List of figures: