Centre for Geo-Information

Thesis Report GIRS-2004-11

Evaluation of satellite soil moisture retrieval algorithms using AMSR-E data

Ruud Hurkmans

March 2004

WAGENINGEN UNIVERSITY
WAGENINGEN UR

ALTERRA
WAGENINGEN UR

Evaluation of satellite soil moisture retrieval algorithms using AMSR-E data

R.T.W.L. Hurkmans
January 2004
Wageningen University
K075-704
Thesis GIS & Remote Sensing

Supervisors:
dr. ir. Z. Su[1]
dr. ir. J. G. P. W. Clevers[2]
dr. T. J. Jackson[3]

[1] Alterra, Wageningen UR, Wageningen
[2] Centre for Geo-information, Wageningen University
[3] Hydrology and Remote Sensing Laboratory, Agricultural Research Service, US Department of Agriculture

Preface

You are about to read the result of a half year's work at the Hydrology and Remote Sensing Laboratory in Beltsville, Maryland, USA. It was a great experience for me to spend this time in the USA and I learnt a lot about culture, the life there and especially about hydrology and remote sensing, which after all was what it was all about. There are some people that made it possible for me to do this and that's why I would like to thank them here: first of all my supervisor Bob Su who was willing to guide my thesis and used his American contacts. Also I'd like to thank Jan Clevers for all his tips and remarks to keep everything suitable for a university thesis and Tom Jackson, who gave me the possibility to go to the USA and always had the patience to answer my questions and help me achieving the goals. Also all other people at the lab, especially Ann Hsu who never got tired of helping me out when I got stuck, Donna Geiman and Laura O'Hare for the practical things they helped me with in the beginning. Last but not least I'd like to thank all other people at USDA, NASA and all the people at the SMEX03 experiment in Chickasha, Oklahoma who made sure I had a great time everywhere I came!

Ruud Hurkmans

Summary

For a better understanding of how human activities influence environmental processes like climate change, desertification and ice melting it is necessary to monitor environmental changes. This requires long-term spatially distributed datasets of the key variables in the interaction between land surface and the atmosphere. Of these key variables, soil moisture is the only one that is not yet observed over extensive areas and time periods. In order to obtain such a long-term spatially distributed soil moisture data set space borne passive microwave remote sensing can be a powerful tool, therefore various types of sensors of this kind have been proposed. Of these sensors, the recently launched Advanced Microwave Scanning Radiometer (AMSR) holds the biggest promise due to its relatively high spatial resolution and the low frequencies it uses. To retrieve soil moisture content from AMSR various retrieval algorithms have been proposed, of which three are compared in this study. One of them, the Jackson algorithm is based on an inversion of the radiative transfer equation. The other two, the de Jeu algorithm and the Wen algorithm, are based on an iterative solution of the radiative transfer equation.

The Jackson algorithm uses ancillary data to take into account vegetation conditions and estimates surface temperature from high frequency brightness temperatures. The Jeu algorithm computes surface temperature in the same way and solves the transfer equation simultaneously for vegetation optical depth and soil moisture. The Wen algorithm estimates vegetation conditions from brightness temperature information and solves the transfer equation for soil moisture and surface temperature.

To compare the algorithms they were applied to two datasets: the Mongolian dataset contains data from measuring stations in Mongolia over a three month period in the summer of 2002, the SMEX02 dataset contains data of a measuring campaign in Iowa, USA in June and July 2002. These two datasets were selected because of the different vegetation conditions: in Mongolia vegetation consists of very sparse, thin grasses, while in Iowa a much denser crop of mainly corn and alfalfa was present. Since in Iowa the problem of Radio Frequency Interference (RFI) exists, two sets of retrievals were performed here: retrievals with AMSR data were performed using higher, less sensitive to RFI, frequencies, and in addition retrievals were performed using the Polarimetric Scanning Radiometer (PSR) which is airborne with a higher resolution and less sensitive to RFI. In these retrievals, the same frequencies as in the Mongolian dataset were used.

For the Mongolian dataset estimated soil moisture contents of the three algorithms were in rather good accordance with each other, but not with observations, due to differences between the sensing depth of the radiation and the observation depth of the measuring station. Also, the soil in Mongolia contained a high portion of rock, which is known to reduce the sensitivity to soil moisture content. Since the vegetation density changed over the time period and these changes were presented only to the Jackson algorithm as ancillary data, this algorithm performed significantly better towards the end of the retrieval period where vegetation density increased. The Wen algorithm appeared to show the smallest errors, but this was due to a different reaction to rainfall events, where brightness temperatures showed large drops. In general, the Jackson algorithm performed the best. For the SMEX02 dataset with both AMSR and PSR data, roughly the same pattern could be seen since, again, the vegetation

conditions changed strongly over the time period, although all vegetation layers were much higher than for the Mongolian dataset. The results for the Wen algorithm using the PSR data were relatively poor, due to the large amount of area and frequency dependent variables in the algorithm that have to be validated thoroughly before every new application, for which time lacked in this study.

Also surface temperature was retrieved. The Jackson and de Jeu algorithms used high-frequency brightness temperatures for this goal: this worked well for the Mongolian dataset where vegetation is low; in areas with higher vegetation such as Iowa the temperature was underestimated. The Wen algorithm solved the radiative transfer equation for surface temperature: here surface temperature was consequently overestimated, because the parameters mentioned earlier were not well enough validated for an optimal surface temperature retrieval.

Concluding, it can be said that it is hard to apply soil moisture retrieval algorithms at a global scale because often either a large amount of ancillary data or many assumptions are necessary. Datasets with a sufficient enough spatial and temporal resolution are not always available and algorithms with many assumptions and parameters to set are very difficult to validate at large scales. The de Jeu algorithm hardly requires any validation and requires very little ancillary data, so it is relatively easy to apply at large (global) scales. However, over longer periods, when vegetation conditions change, it is difficult to perform a reliable retrieval without using ancillary vegetation data. The Jackson algorithm, which used this ancillary data, performed significantly better than the de Jeu and Wen algorithms when comparing the algorithms under different conditions.

Table of Contents

1 Introduction

1.1 Background

Currently there is a large interest in the interactions between human activities and climate change. There is a strong indication that human activities are disturbing the natural climate, causing acceleration in global warming, sea level rise, deforestation and desertification (de Jeu, 2003). In order to monitor these environmental changes there is a need for long-term spatially distributed data sets of the key variables in the interaction between land surface and the atmosphere: soil moisture, soil temperature, vegetation and precipitation. Soil temperature, vegetation and precipitation are all currently observed using satellite observations. Soil moisture is the only one that is not currently observed over large areas and over an extended period using remote sensing (Bindlish et al., 2003). Soil moisture is considered the key variable in hydrology because it is the switch that controls the proportion of rainfall that percolates, runs off or evaporates from the land. It is the life giving substance for vegetation and it integrates precipitation and evaporation over longer periods, thus introducing an element of memory in the atmosphere/land system (SMEX03, 2003).

1.2 Problem description

There are no validated large-scale, long-term databases containing soil moisture and there are large discrepancies among the results of different land surface hydrology models forced with the same data (Bindlish et al., 2003; Jackson et al., 1999; Njoku et al., 2003). For both hydrologic and climatologic research creation of a validated, large-scale, long-term database is a necessity. Considering the coarse resolutions global climate models are running on, space borne microwave remote sensing can be used to observe soil moisture at large scales. It can directly measure the dielectric properties which are strongly dependent on the liquid water content, and due to the long wavelength it is not sensitive to scattering of cloud particles which makes microwaves useful in all weather conditions, both day and night since there is no dependence on solar radiation (Koike et al., 2000; Njoku et al., 2003). Only vegetation can influence microwave radiation and 'mask' soil moisture content (Jackson et al., 1982).

Several space borne sensors have been developed for this purpose, including the Scanning Multichannel Microwave Radiometer (SMMR), Special Sensor Microwave/Imager (SSM/I), Tropical Rainfall Measuring Mission (TRMM) / Microwave Imager (TMI) and the Advanced Microwave Scanning Radiometer (AMSR). Being launched in the middle of 2002, the AMSR is recognized to be the most suitable for retrieval of soil moisture in regions with low vegetation cover due to the low frequency (and thus long wavelength). Also it provides one of the best spatial resolutions of current multi-frequency radiometers from space due to its large antenna, varying from 5 km at the smallest to 50 km at the largest wavelength (Jackson and Hsu, 2001; Jackson et al., 2002; Koike et al., 2000; Njoku et al., 2000). For more information about AMSR see the next chapter. Several algorithms to retrieve soil moisture from these microwave sensors' data have been proposed but not yet extensively evaluated due to lack of datasets for this purpose (Jackson et al.,

2002). Also for older sensors like the already mentioned TRMM/TMI and SMMR algorithms have been developed that can be adapted to be used with AMSR data.

1.3 Research objectives

In this work, three of these algorithms will be evaluated and compared. Where needed, algorithms will be adapted for use with AMSR data. Of these algorithms, the one developed by Jackson (1993) was developed for use with AMSR-data. A second one, by de Jeu (2003) was originally applied to SMMR data. The last one, developed by Wen et al. (2003) was originally developed for TRMM/TMI. Validation datasets that are used for this purpose are the Mongolia Match Up Dataset, data from several soil moisture observation locations in eastern Mongolia in the summer of 2002 (Kaihatsu, 2003), and soil moisture observations from Soil Moisture Experiments 2002, in the summer of 2002 in Iowa, USA (SMEX02, 2002). More detailed information about these algorithms and datasets is provided in Chapter 2.

1.4 Research questions

The research questions that can be formulated are:

- Is it possible with the proposed algorithms to provide reliable soil moisture estimations at a global scale?
- Which of the algorithms provides the most reliable estimations?
- Which of the algorithms is the most suitable for global application?
- Are the validation datasets used in this study suitable for validation and evaluation of the proposed algorithms?

1.5 Report setup

In the second chapter of this report an overview shall be given about some background theory: general microwave remote sensing, the AMSR sensor, the retrieval algorithms that were evaluated and the datasets that were used for the validation. After this, in Chapter 3 more information will be provided of the used methodologies: how and with which parameter values have the algorithms been applied to the datasets and how was the comparison realized. The results of the comparison are described in Chapter 4. In Chapter 5 they are discussed and finally in Chapter 6 conclusions are drawn and the questions asked in section 1.4 are answered.

2 Theoretical background

2.1 Microwave remote sensing and soil moisture

The main characteristics of electromagnetic radiation are wavelength, polarization, amplitude and phase. Here we examine microwave radiation, which has relatively long wavelengths from 1 cm up to 1 m. In Table 1 an overview can be found of nomenclature, wavelengths and frequencies for microwave radiation.

Wave polarization is caused by the electrical field that is always present in electromagnetic waves. This field rotates in a plane perpendicular to the propagation direction and can be considered the vector sum of a horizontal and a vertical component. The resulting polarization depends on the phase difference between the two components. It can be elliptical, circular or linear. Depending on the sense of rotation of the electric field, elliptical and circular polarized waves can be right- or left-handed; i.e. respectively rotating clockwise or counter-clockwise. The passive microwave systems that will be used in this study only use linearly -horizontally and/or vertically- polarized radiation (H or V).

Table 1: Nomenclature of microwave bands and corresponding wavelengths and frequencies.

Band designation	Wavelength [cm]	Frequency [GHz]
Ka	0.75 - 1.11	27 – 40
K	1.11 - 1.67	18 – 27
Ku	1.67 – 2.50	12 – 18
X	2.50 – 3.75	8 – 12
C	3.75 – 7.50	4 – 8
S	7.50 – 15.0	2 – 4
L	15.0 – 30.0	1 – 2
P	30.0 – 100	0.3 – 1

For the determination of the amount of soil moisture, the amplitude or field strength of radiation is very important. The amplitude of reflected or emitted waves decreases by a certain factor, depending on polarization, incidence angle of the radiation and material properties like vegetation cover, soil moisture and surface roughness. The factor by which the amplitude is attenuated is called the Fresnel coefficient. Using the Fresnel equations, it is possible to derive e.g. soil moisture from the Fresnel coefficient, assuming incidence angle and polarization are known.

The fourth and last electromagnetic characteristic to be considered is phase, which requires multiple sensors and is only used in interferometric measurements (Hoekman, 2000).

In Figure 1 the basic geometry of passive microwave remote sensing is shown: the satellite moves over the earth surface following the flight direction while at certain intervals the surface is scanned in the perpendicular direction, the scan-direction. Each ellipse along the scan line in the figure represents a so-called footprint, one measurement.

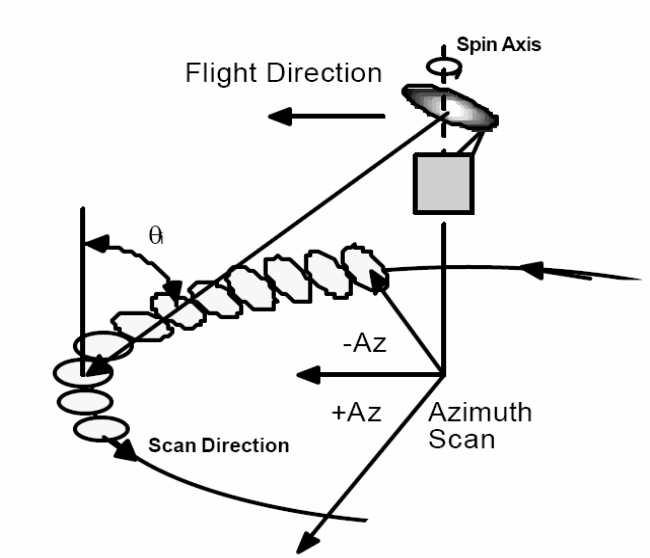


Figure 1: Geometry of passive microwave remote sensing.

The distance perpendicular to the flight direction over which the scan line extends is called the swath width. To obtain a detectable signal the energy over a large area has to be collected which results in large footprints and a low spatial resolution. In general, spatial resolution decreases with frequency as altitude increases. For satellite designs at L band, this might results in a footprint as large as 50 km. For more information on basic principles of microwave remote sensing, see Dane and Clarke Topp (2002), Hoekman (2000) and Lillesand and Kiefer (1999).

The advantages of using passive microwave remote sensing for soil moisture mapping are that observations can be made under conditions of cloud cover, are independent of solar illumination and can be made at any time in day or night, as opposed to instruments that operate in the visible or infra-red parts of the electromagnetic spectrum. However, remote sensing cannot replace ground-based methods for providing high-quality data; its advantage is in large-scale mapping such as continental or global scales. It was recognized early in research that instruments operating at low frequencies (< 6 GHz) provide the best soil moisture information: there is less influence of atmosphere and vegetation, instruments respond to a deeper soil layer, and there is a higher sensitivity to soil water content (Dane and Clarke Topp, 2002).

2.2 The AQUA/AMSR-E sensor

Two versions of the Advanced Microwave Scanning Radiometer (AMSR) were launched in 2002. AMSR was launched on the Advanced Earth Observing Satellite (ADEOS-II), by NASDA. NASA's Earth Observing System (EOS) Aqua satellite included a modified version, AMSR-E. Differences between the sensors include a slightly larger antenna for AMSR, resulting in a slightly higher spatial resolution and a swath of 1600 instead of 1445 km, and the time of overpass. Aqua's overpass time is 1.30 am and pm local time, while the overpass of ADEOS-II is at 10.30 am and pm. AMSR has also additional bands at 50 and 52 GHz with vertical polarization. An

overview of the characteristics of AMSR-E is shown in Table 2. AMSR-E coverage is global between 89.24º N and 89.24º S and the sampling interval at the earth's surface is 10 km for all channels except the 89 GHz channel, which has a sampling interval of 5 km. The distance between the scan lines (see Figure 1) also is about 10 km. Temporal coverage exists since June 2, 2002. The viewing angle of both sensors is a constant 55º. Due to the small vegetation penetration capacity at C and X band, which are most sensitive to soil moisture content (Dane and Clarke Topp, 2002), AMSR is anticipated to provide soil moisture information in regions of low vegetation cover, i.e. less than 1 kg/m^2 vegetation water content. More details about AMSR-E and AMSR can be found at http://www.ghcc.msfc.nasa.gov/AMSR/, http://nsidc.org/data/docs/daac/ae_l2a_tbs.gd.html and http://adeos2.hq.nasda.go.jp/shosai_amsr_e.htm., SMEX03 Experiment Plan (2003) and in Kawanishi et al. (2003).

Table 2: Characteristics of AMSR-E.

Frequency [GHz]	Polarization	Mean Horizontal Resolution [km]	Swath [km]	Footprint size [km]
6.925	V, H	56	1445	75 x 43
10.65	V, H	38	1445	51 x 30
18.70	V, H	21	1445	27 x 16
23.80	V, H	24	1445	31 x 18
36.50	V, H	12	1445	14 x 8
89.00	V, H	5.4	1445	6 x 4

2.3 From satellite measurements to soil moisture

There are two approaches to retrieve Volumetric Soil Moisture (*VSM*) from Brightness Temperature (T_b) data. One of them is an inversion of the forward model, which is the prediction of T_b from soil properties using radiative transfer theory. These algorithms thus invert the equations and estimate soil properties such as *VSM* from T_b. The second approach solves the forward model iteratively, optimizing soil properties to minimize the difference between predicted and measured T_b. Several algorithms of this latter kind have been proposed (de Jeu, 2003; Njoku & Li, 1999; Wen et al., 2003; Koike et al., 2000; Njoku et al., 2000; Njoku et al., 2003). There are five steps involved in extracting *VSM* using passive microwave remote sensing, which are the following:

- Normalizing brightness temperature to emissivity
- Removing the effects of vegetation
- Accounting for the effects of surface roughness
- Relating the emissivity measurement to soil dielectric properties
- Relating the dielectric properties to *VSM*

In this study, three algorithms are compared. The algorithm developed by Jackson (1993), from now called the Jackson algorithm, is based on an inversion of the forward model. The algorithm developed by de Jeu (2003), the de Jeu algorithm, solves the forward model iteratively. The third model, by Wen et al. (2003), the Wen algorithm, is also based on the iterative approach, but in a different way as is explained later. The following sections provide descriptions of each.

2.3.1 The Jackson algorithm

For a vegetated site, the T_B value measured is a combination of the radiation emitted by the vegetation and the radiation emitted by the underlying soil as modified by the vegetation. This is described by the radiative transfer equation (Jackson, 1993):

$$T_B = \Gamma \bullet e_r \bullet T_S + (1-\varpi)T_C(1-\Gamma) + (1-e_r)(1-\varpi)T_C(1-\Gamma)\Gamma \qquad (1)$$

Where T_B is the brightness temperature, T_S and T_C are temperatures of soil and canopy respectively, ω is the single scattering albedo, e_r is the surface emissivity and Γ is the transmissivity of the canopy. At microwave wavelengths, the single scattering albedo is very small. When this term is set to zero and if it is assumed that the physical temperatures of soil and canopy are nearly the same, Equation 1 reduces to:

$$e = (\frac{T_B}{T_S}) = 1-(1-e_r)\Gamma^2 \qquad (2)$$

Where e is the emissivity of the surface covered with vegetation. Inverting Equation 2 gives e_r expressed in T_B and T_S:

$$e_r = 1-\frac{(1-(T_B/T_S)}{\Gamma^2} \qquad (3)$$

The surface temperature (T_S) is estimated using the brightness temperature from the vertically polarized 37 GHz band according to the following equation (De Jeu, 2003):

$$T_s = 0.861 * T_{B,37V} + 52.550 \qquad (4)$$

The transmissivity of the canopy Γ can be described by:

$$\Gamma_{(p)} = \exp(\frac{-\tau_{(p)}}{\cos(\theta)}) \qquad (5)$$

Where p is polarization, τ is the optical depth of the canopy and θ is the incidence angle of the radiation. τ is determined by:

$$\tau = b * VWC \qquad (6)$$

Where b is a vegetation parameter, dependent on land use and frequency, and VWC is the vegetation water content in [kg m^{-2}]. The vegetation water content can be obtained from the Normalized Difference Vegetation Index (NDVI) data (Clevers, 2001), converted to VWC by means of linear relationships:

$$\begin{aligned} VWC &= 2.0 * NDVI && ,0.36 < NDVI \le 0.50 \\ VWC &= 2.5 * NDVI && ,0.20 < NDVI \le 0.36 \end{aligned} \qquad (7)$$

$$VWC = 3.0 * NDVI \qquad ,0.00 < NDVI \leq 0.20$$

The surface emissivity e_r is corrected for the influence of surface roughness s according to (Choudhurry et al., 1979), giving the smooth surface emissivity e_s:

$$e_s = 1 + (e_r - 1)\exp(h\cos^2\theta) \qquad (8)$$

Where h is a roughness parameter ($=4s^2k^2$) proportional to the root mean square (RMS) height variations of the soil surface, and $k = 2\pi/\lambda$ (with λ is wavelength). With the inverted Fresnel equation for horizontally polarized radiation the dielectric constant k is computed from the smooth surface reflectivity R_s ($=1-e_s$):

$$k = \sin^2\theta + \left[\cos\theta\left(\frac{-1-\sqrt{R_S}}{(\sqrt{R_S})-1}\right)\right]^2 \qquad (9)$$

From k the *VSM* is computed using the dielectric mixing model of Wang and Schmugge (1980). An alternative method for this mixing model is provided by Hallikainen et al. (1985). An advantage of this method is that it is also useful for X-band, while the Wang-Schmugge model is only valid for L- and C-band. Both models require soil texture information and are described in Appendix D. An overview of the computational process of the Jackson algorithm is shown in Figure 2.

Before brightness temperature values are used in the algorithm, they are screened for rainfall using the screening algorithm demonstrated in Ferraro et al. (1997), and on land use: retrieval is only possible in areas with low vegetation content, so e.g. forests are left out. Only clear pixels with retrievable land use types are used in the algorithm (Bindlish et al., 2003).

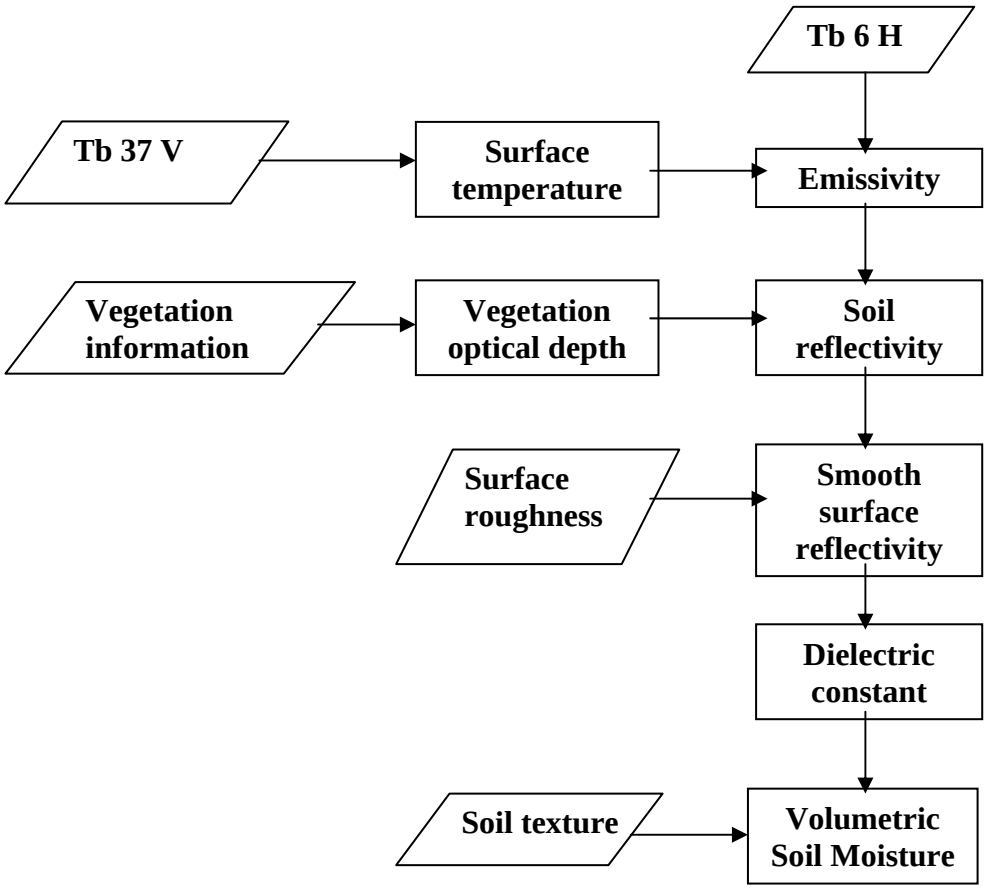


Figure 2: Flowchart showing the computational process of the Jackson algorithm.

2.3.2 The de Jeu algorithm

The methodology of de Jeu is also based on the radiative transfer equation, Equations 1 and 5. The equation is solved in a forward, iterative manner for dielectric constant k and the optical depth of the vegetation τ. This means that some other variables have to be estimated independently: surface temperature is calculated according to equation 4, and vegetation temperature is assumed equal to the surface temperature. For the single scattering albedo ω a constant value of 0.06 is assumed. Surface roughness is assumed to be negligible in this algorithm and thus not accounted for. The unique part of this algorithm is in the fact that the radiative transfer equation is solved for k and τ simultaneously.

Another component of this algorithm is the use of the Microwave Polarization Difference Index (MPDI). This is used in an attempt to remove the dependence on the physical temperature of the emitting layer:

$$MPDI = \frac{T_B(v) - T_B(h)}{T_B(v) + T_B(h)} \qquad (10)$$

At the wavelength that is used, about 6 GHz, the MPDI contains information on both canopy properties (τ) and soil dielectric properties. An empirical relationship between τ and the MPDI was developed based on simulations:

$$\tau = C_1 \ln(MPDI)^3 + C_2 \ln(MPDI)^2 + C_3 \ln(MPDI) + C_4 \qquad (11)$$

Parameters *C1*, *C2*, *C3* and *C4* are a function of *k* and can be defined by the relationship:

$$C_j = M_{j,1} k^N + M_{j,2} k^{(N-1)} + \ldots + M_{j.N} k + M_{j,(N+1)} \qquad (12)$$

Where *j* is the coefficient number (1 to 4), *N* is the degree of the polynomial and *M* refers to the polynomial constants that are given in Table 3.

The last term in Equation 1 that is not yet expressed in *k* is the emissivity e_s. This is done by means of the Fresnel equation for horizontally polarized radiation:

$$e_s(\theta) = 1 - \left[\frac{\cos(\theta) - \sqrt{k - \sin^2(\theta)}}{\cos(\theta) + \sqrt{k - \sin^2(\theta)}} \right]^2 \qquad (13)$$

Since H-polarization has the greatest sensitivity to soil moisture, the radiative transfer equation (Equation 1) is solved using H-polarized T_B. The complete radiative transfer equation is expressed in terms of *k* and can be solved using a non-linear iterative procedure, by optimizing *k*. After convergence of the equation, soil moisture content is calculated from *k* using the Wang-Schmugge mixing model (Wang and Schmugge, 1980). An overview of the computational process of the de Jeu algorithm is shown in Figure 3. For more information about this algorithm, see de Jeu (2003).

Table 3: Polynomial parameters that describe the relation between the absolute value of the dielectric constant *k* and the fitting parameters C_1, C_2, C_3, C_4 in the de Jeu algorithm.

j	1	2	3	4
$M_{j,1}$	$3.9629 \cdot 10^{-12}$	$5.9932 \cdot 10^{-11}$	$2.4073 \cdot 10^{-10}$	$1.9334 \cdot 10^{-9}$
$M_{j,2}$	$-5.0037 \cdot 10^{-10}$	$-7.6059 \cdot 10^{-9}$	$-3.0573 \cdot 10^{-8}$	$-2.4698 \cdot 10^{-7}$
$M_{j,3}$	$2.5814 \cdot 10^{-8}$	$3.9687 \cdot 10^{-7}$	$1.6003 \cdot 10^{-6}$	$1.3092 \cdot 10^{-5}$
$M_{j,4}$	$-7.0276 \cdot 10^{-7}$	$-1.1018 \cdot 10^{-5}$	$-4.4697 \cdot 10^{-5}$	$-3.7305 \cdot 10^{-4}$
$M_{j,5}$	$1.1020 \cdot 10^{-5}$	$1.7748 \cdot 10^{-4}$	$7.2593 \cdot 10^{-4}$	$6.1918 \cdot 10^{-3}$
$M_{j,6}$	$-1.0600 \cdot 10^{-4}$	$-1.7392 \cdot 10^{-3}$	$-7.1565 \cdot 10^{-3}$	$-6.0774 \cdot 10^{-2}$
$M_{j,7}$	$7.1840 \cdot 10^{-4}$	$1.1381 \cdot 10^{-2}$	$4.7158 \cdot 10^{-2}$	$3.4518 \cdot 10^{-1}$
$M_{j,8}$	$-1.0976 \cdot 10^{-3}$	$-1.0594 \cdot 10^{-2}$	$-3.2233 \cdot 10^{-1}$	-1.1777

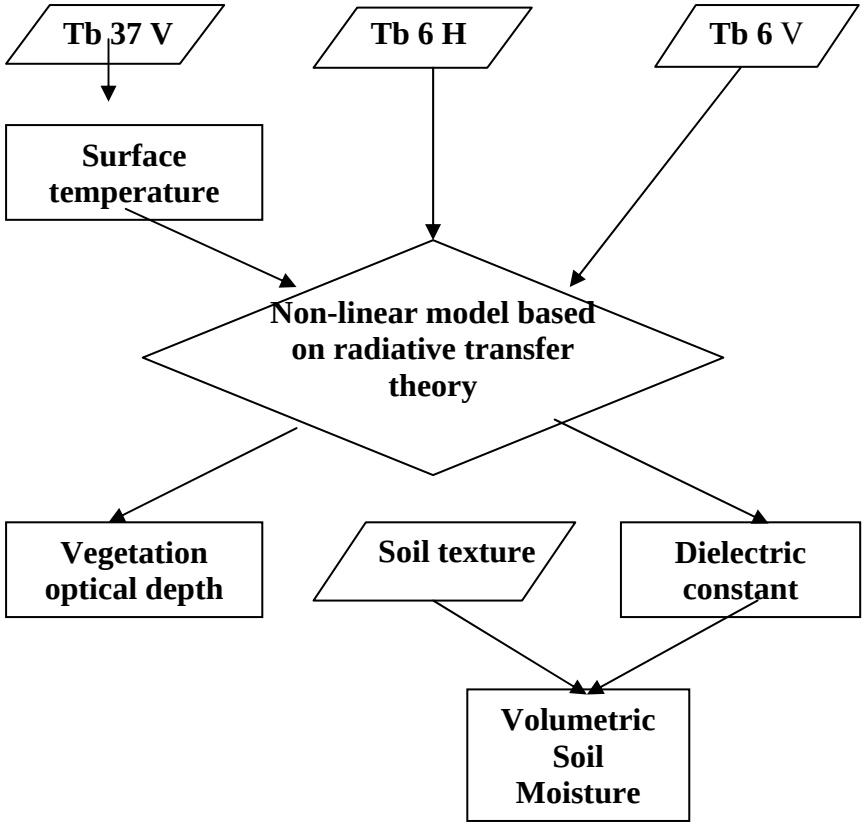


Figure 3: Flowchart showing the computational process of the De Jeu algorithm.

2.3.3 The Wen algorithm

This algorithm is based on a similar but slightly differently formulated radiative transfer equation (Wen et al., 2003):

$$T_{Bp}(\tau,\theta) = (1+Lr_p)(1-L)(1-\omega)T_V + (1-r_p)LT_S \qquad (14)$$

where T_{Bp} is the brightness temperature at p polarization, which refers to H or V. T_V and T_S are vegetation temperature and soil temperature respectively, r_p is soil reflectivity at p polarization, L is canopy attenuation given by $L = exp(-\tau sec(\theta))$. τ is the vegetation opacity or optical thickness, which depends on the canopy extinction coefficient and height, ω is the single scattering albedo of the vegetation layer. When the radiance that is emitted downward from the vegetation layer and reflected by the surface layer is neglected –this term is usually very small- equation 14 can be reduced to:

$$T_{Bp}(\tau,\theta) = (1-\omega)(1-L)T_V + (1-r_p)LT_S \qquad (15)$$

The canopy attenuation L is estimated by deriving the Leaf Area Index (*LAI*) from the Normalized Difference Polarization Index (*NDPI*):

$$LAI = \frac{\nu * \cos(\theta) * \sqrt{\lambda}}{\kappa} \ln\left[\frac{NDPI(0,\theta)}{NDPI(\tau,\theta)}\right] \qquad (16)$$

where v is an empirical correlation factor between plant water content and *LAI*, λ is wavelength, κ is a vegetation type dependent factor. Typical values for v are 3.3 for corn and 1.0 for natural grass. For κ typical values are 0.1 $m^{1/2}$ for sunflower and 0.4 $m^{1/2}$ for alfalfa. *NDPI* can be expressed as follows: (*NDPI(0,θ)* is the *NDPI* for bare soil.)

$$NDPI(\tau,\theta) = \frac{T_{B,v} - T_{B,h}}{T_{B,v} + T_{B,h}} \tag{17}$$

where $T_{B,v}$ is vertically polarized brightness temperature and $T_{B,h}$ is horizontally polarized brightness temperature. A linear relationship between τ and *LAI* and thus a logarithmic relationship between L and *LAI* is assumed.

For the single scattering albedo ω constant values are assumed based on van de Griend, (1996): 0.05 for vertically polarized radiation and 0.06 for horizontally polarized radiation. Now enough parameters in the radiative transfer equation are known, a system of 3 equations is set up using brightness temperatures of 3 channels to obtain the value of three unknowns: reflectivity for horizontal radiation, reflectivity for vertical radiation and surface temperature.

Subsequently, using the vertical and horizontal reflectivity, another system of 2 equations is set up to correct for the influence of surface roughness, using the results of the first equation system as input. The two unknowns are the corrected values for respectively horizontal and vertical reflectivity. The equations are based on Choudhurry, (1979):

$$\Gamma_p = \left[(1-Q)\Gamma_{sp} + Q\Gamma_{sq}\right]\exp(-k^2\sigma^2\cos^2\theta) \tag{18}$$

where Γ_p is reflectivity of the rough surface and Γ_{sp} is reflectivity of the smooth surface, p and q are orthogonal polarizations (h and v), k is the wavenumber (which is expressed by $k = 2\pi/\lambda$, σ is the standard deviation of the surface height and Q can be determined as expression:

$$Q = 0.35\left[1 - \exp(-0.66\sigma^2 f)\right] \tag{19}$$

where f is the frequency of the microwave radiation beam in GHz. From the resulting horizontal reflectivity soil moisture is derived according to an approximately linear relationship (Yang et al., 2000):

$$\Gamma_{sh} = a * M_v + b \tag{20}$$

where a and b are constants at the given soil texture, frequency and incidence angle and M_v is volumetric soil moisture content. An overview of the computational process of the Wen algorithm is shown in Figure 4 and for more information about this algorithm see Wen et al. (2003).

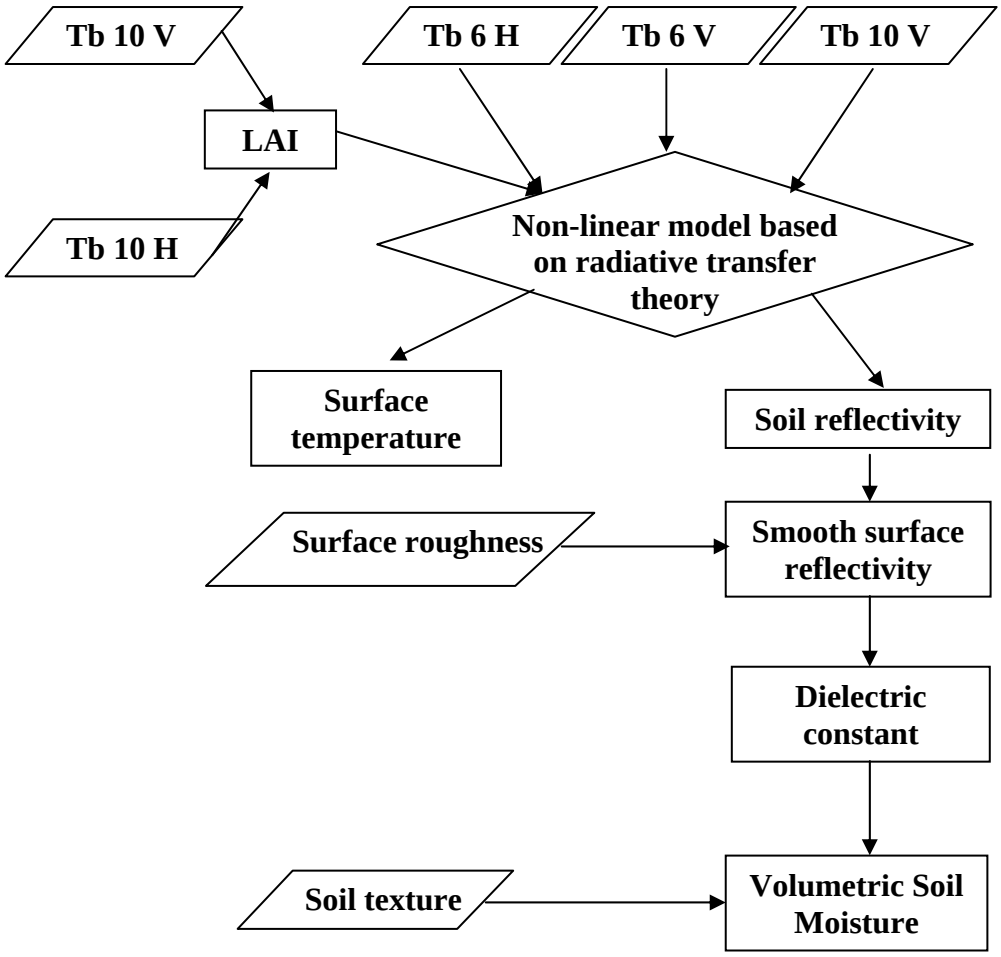


Figure 4: Flowchart showing the computational process of the Wen algorithm.

2.4 Validation datasets

Two data sets were used to compare and validate the retrieval algorithms described in the previous sections, one from Mongolia and the other from Iowa.

2.4.1 Mongolia Match up dataset

In this dataset, provided by NASDA (now JAXA), AMSR-data and ground measurements were processed and georeferenced into a grid of 26 x 26 pixels between 45.0° and 47.5° N, and 105.5° and 108.0° E. The spatial resolution of the grid is 0.1°. The location of the area in Mongolia is shown in Figure 5. The period covered was three months, between July 1 and September 21 2002. Some gaps were present in the data, especially during the beginning of August. In addition to this gridded data, the same data was provided as point type data. A file was provided for each satellite overpass and each observation site containing the AMSR brightness temperatures in all of its frequencies and the observed soil temperature and soil moisture at 3 and 10

cm depth (for the 12 hours before and after the overpass). These observations were taken at in situ stations, which provide continuous measurement of soil moisture content (Kaihatsu, 2003). A picture of such an observation station can be found in Appendix B.

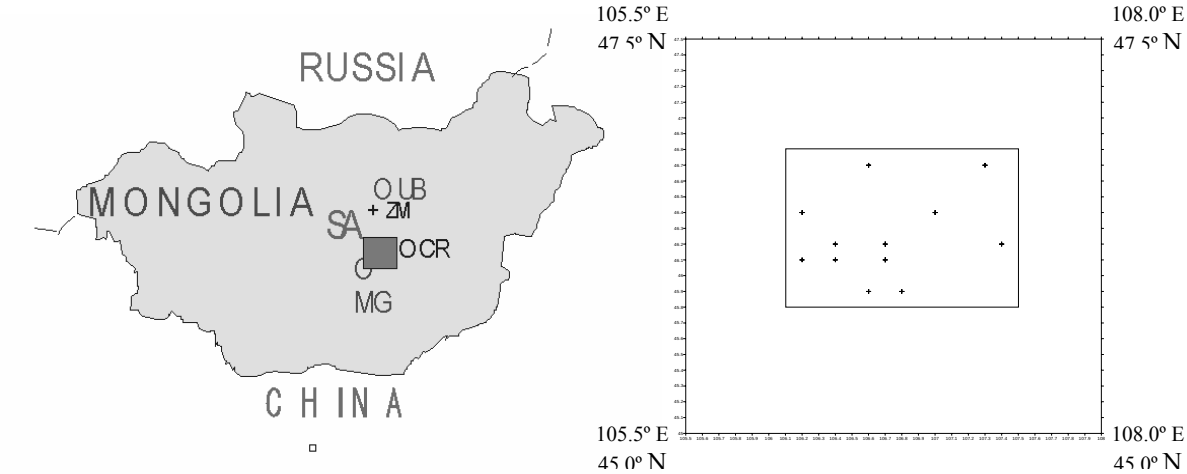


Figure 5: Location of the study area (left; the dark square) and the observation points in the study area (right). Satellite observations are averaged over the smaller box around the observation points.

For the validation of the algorithms, the soil moisture was retrieved over the whole 2.5° by 2.5° grid using the gridded brightness temperatures. The in situ observation locations were centered in a small part of the larger match-up (see Figure 5). It was not considered valid to compare a retrieval over the whole area with the relatively small observed area. Therefore, a validation box was defined around the observation locations and only pixels within this box were used for our analyses. This box was located between 45.8° and 46.8° N, and 106.1° and 107.5° E. The retrieved soil moisture was then averaged over this smaller area and statistically compared with the averaged observed values. The locations of observation points and the box mentioned above are also shown in Figure 5. Observed values were available at 3 and 10 cm depth, but considering the penetration depth of microwave radiation at the used frequency (which is assumed to be less than 3 cm, namely in the order of 1 - 1.5 cm), only the observations at 3 cm depth were used.

2.4.2 Soil Moisture experiments 2002

Soil Moisture Experiments 2002 took place in an area around Ames, Iowa, from June 25 to July 12, 2002. This area was selected to provide a dataset for the development and verification of alternative soil moisture retrieval algorithms under significant biomass levels associated with agricultural crops (SMEX02, 2002). The location of the research area in latitude/longitude coordinates and in UTM coordinates is shown in Figure 6.

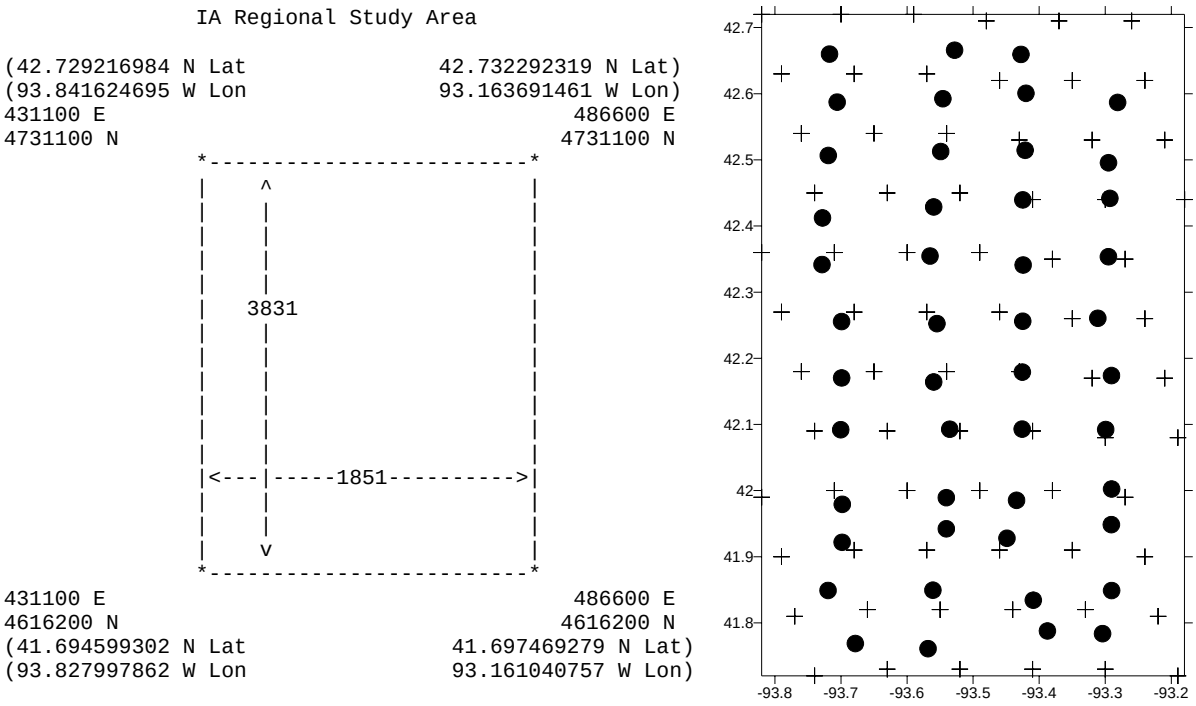


Figure 6: The SMEX02 study area. Coordinates are in UTM Zone 15, WGS84, in meters and in WGS84 Lat/Lon. These coordinates represent the pixel centers of each corner pixel. The coordinate system is designed for representation with 30 m pixels in UTM, WGS84. Ancillary data is provided, formatted in the resulting 1851 x 3831 matrix (*VWC* and Land cover). In the right plot the distribution of observation points (dots) and AMSR-E footprints (crosses) is shown.

Nearly 95 % of the region is used for row crop agriculture. Corn and soybean are grown on approximately 80 % of the row crop acreage with more than 50 % corn, 40-45 % soybean and the remaining 5-10 % in forage and grains.

Regional observation sites were selected to provide representative coverage over an area large enough to include several AMSR sized footprints. Selection of the size was based on factors like geographic distribution, travel time, balance of corn and soy bean, soil types and permission to use. At these sites soil moisture was sampled to provide a reliable estimate of the volumetric soil moisture content mean and variance within a single satellite passive microwave footprint (~ 50 km) at the overpass time. A grid of 47 sites covering approximately 50 x 100 km was sampled once every day. The primary measurement at each site was the 0-6 cm dielectric constant measured by a device called Theta Probe. The Theta Probes were calibrated using site-specific calibration equations based on supplemental gravimetric soil moisture and bulk density sampling. This was done by extracting a sample of the 0-1 cm and 1-6 cm soil layers using a coring tool. Soil temperature was also measured at these 47 sites using a hand-held infrared thermometer for surface temperature and temperature probes for soil temperatures at 1, 5 and 10 cm depths. Observed data used for validation in this study are the Theta Probe measurements over the top 6 cm of the profile, and soil temperature at 1 cm depth (SMEX02, 2002).

3 Methodology

3.1 Ancillary Datasets

The algorithms require different amounts and kinds of input data. The Jackson algorithm requires land cover data to filter out 'bad' pixels, vegetation cover data to compute optical depth and soil texture information to calculate the soil moisture content from the dielectric properties. The de Jeu algorithm uses the soil texture information as well, but does not require a land-cover based filter and attempts to retrieve vegetation properties directly. The Wen algorithm requires soil moisture observations to calibrate the soil texture coefficients in Equation 20 and attempts to retrieve vegetation properties directly from brightness temperature data as well. To ensure a proper comparison between the algorithms however, the same pixels should be used for all retrievals. Therefore, the land cover and rainfall filters in the Jackson algorithm were also applied to the de Jeu and Wen algorithms.

For land cover information in the Mongolian area, the global 8-km scale product from the University of Maryland (http://glcf.umiacs.umd.edu) was used. Soil texture information came from Reynolds et al. (2000). The following land cover types were considered for retrieval:

1. wooded grasslands/shrubs
2. closed bushlands or shrublands
3. open shrublands
4. grasses
5. bare
6. mosses and lichens

A change that was made to the Jackson algorithm was the replacement of the NDVI with the Enhanced Vegetation Index (EVI). This is developed by the Terrestrial Biophysics and Remote Sensing Lab (2003) and derived from MODIS data. This data appeared to be in better agreement with ground vegetation observations, because the images were real-time, as the NDVI originally used was averaged over 10 years. EVI can be described as follows:

$$EVI = G * [\frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + C_1 * \rho_{RED} - C_2 * \rho_{BLUE} + L}] \qquad (21)$$

Where G is a gain factor, L is a canopy background brightness correction factor, C_1 is an atmosphere resistance red correction coefficient and C_2 is an atmosphere resistance blue correction coefficient. C_1 and C_2 are meant to correct for the influence of aerosols. To derive *EVI*, a value for G of 2.5 is used; L is 1, C_1 is 6 and C_2 is 7.5.

In the three-month period used for the comparison, *EVI* data was available at six dates, spread over the period. For the other dates, *EVI* was linearly interpolated between these six days. In Figure 7 the *VWC*, which is directly computed from *EVI* by means of a set of linear relations (Equation 7) for the period of the experiment in Mongolia is shown. Since the multipliers in Equation 7 are arbitrarily and based on

experience, Equation 7 was used to compute *VWC* from *EVI* as well, assuming the differences between *NDVI* and *EVI* very small.

Also a change was made to the Wen algorithm: since the coefficients in Equation 20 are difficult and unpractical to determine for both the Mongolian and the SMEX02 datasets, instead of Equation 20 the dielectric constant was calculated from the smooth surface reflectivity for horizontally polarized radiation using the inverted Fresnel equation (Equation 9) and subsequently the *VSM* was calculated from the dielectric constant using the Wang-Schmugge dielectric mixing model, using the same soil texture datasets as the Jackson and de Jeu algorithms.

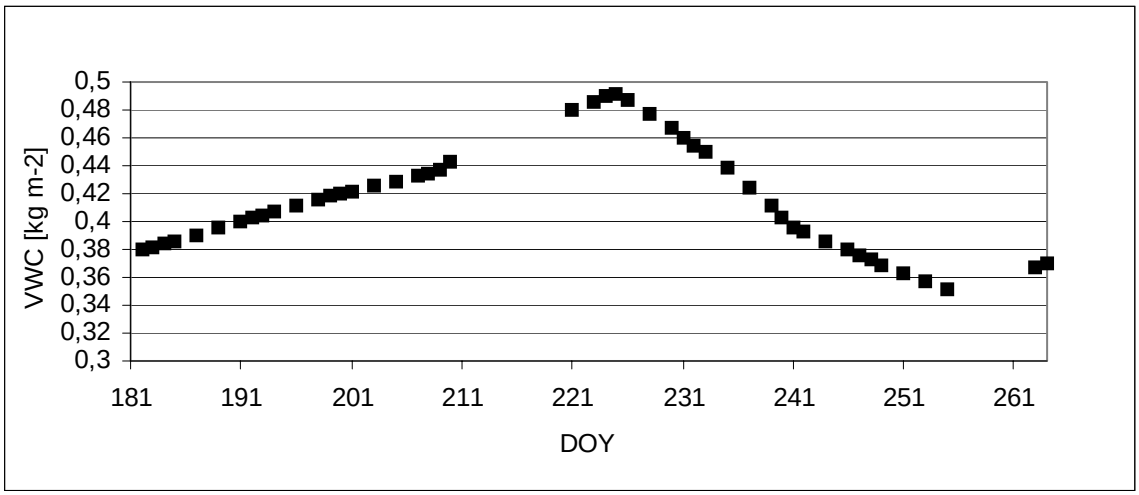


Figure 7: Average *VWC* as derived from *EVI* for the Jackson algorithm during the time period of the Mongolia experiment.

For the SMEX02 dataset, a large amount of ancillary data for the specific location and period was available. Of these data, the *VWC* and Land cover data were used. *VWC* data was available over the whole area with a 30 m pixel resolution (see Figure 6). The average *VWC* over the area during the experiment period is shown in Figure 8. It was estimated using Landsat TM data, available at 5 dates during SMEX02. For the process of deriving *VWC* see Appendix C.

Land cover data was used in all algorithms. During SMEX02 the area was classified using Landsat TM data. For classification procedures see Appendix C. As a result, the following classes (also with a resolution of 30m, Figure 5) were used:

0	Unclassified
1	Alfalfa
2	Corn
3	Grass
4	Soybean
5	Trees
6	Urban
7	Water
10	Overlaid Roads

Of these classes, classes 1, 2, 3 and 4 were considered to be retrievable. Pixels containing other classes were filtered out.

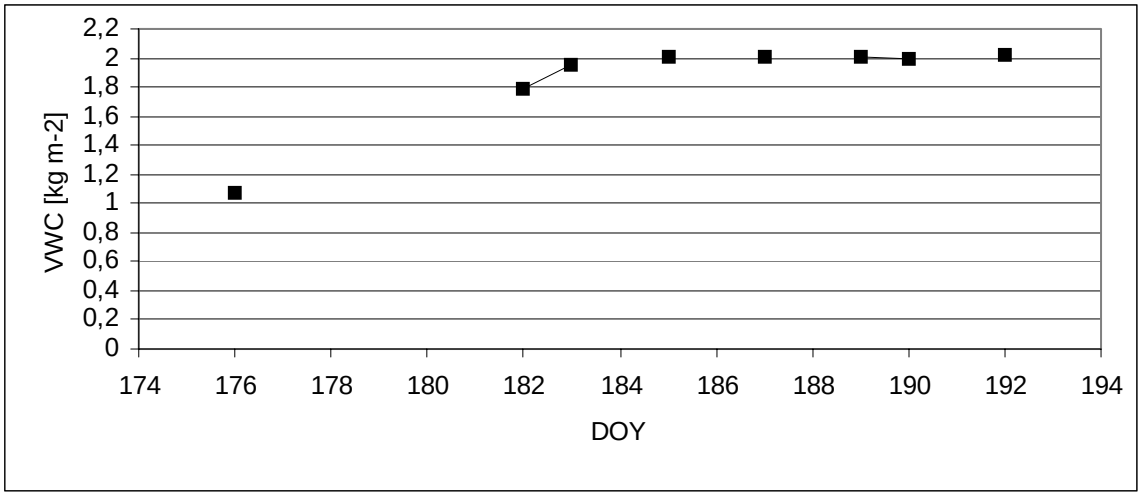


Figure 8: Average *VWC* as derived from the SMEX02 ancillary dataset for the Jackson algorithm during the experiment period of SMEX02.

3.2 Data quality check

A data quality check was performed on the Mongolian data set. Several kinds of analyses and plots were produced (Appendix A). The features observed in the quality check were taken into account in the rest of the procedures. The first check was to see whether there was any connection between the (observed) soil moisture contents and the horizontally polarized brightness temperatures at lower frequencies (6 and 10 GHz), which are supposed to be the most sensitive to changes in soil moisture content. This was done by plotting moisture content and brightness temperature against each other in scatter plots. These graphs are shown in Figures A.1 and A.2 of Appendix A, for all the twelve observation locations in the area, for brightness temperatures of both ascending and descending orbits. Since generally a trend seemed to exist of increasing soil moisture with increasing brightness temperature, which is the opposite of what's to be expected, additional analyses were performed.

Firstly, the temporal patterns of both observed soil moisture and brightness temperatures were compared (Figures A.3 and A.4 of Appendix A). It appeared that rainfall events (large drops in brightness temperature) corresponded with sharp increases in soil moisture content. After producing the same plots for surface temperature (Figure A.6), it was concluded that both observed and satellite data are in good agreement from temporal respect. Secondly, to eliminate the influence of surface temperature, brightness temperatures are normalized to emissivities (Figure A.5) and lastly, since the Mongolian dataset was already georeferenced, the spatial pattern of brightness temperature of the provided dataset is compared to the original AMSR-E data, resampled to the same 26 x 26 matrix as the provided dataset. Original and provided data appeared to be in good agreement from a spatial point of view. It was concluded that the brightness temperatures were provided for correct area and period (no mistakes were made). Therefore, the features observed here have to be taken into account during the interpretation of the results.

3.3 Running the algorithms for the Mongolian dataset

For the Mongolian dataset, the already processed and georeferenced brightness temperatures (see previous section) were used as input. For the actual soil moisture computation brightness temperatures at the lowest possible frequencies were used, i.e. 6.9 GHz horizontally polarized for the Jackson algorithm and 6.9 GHz dual-polarized for the de Jeu algorithm. The Wen algorithm used 6.9 GHz dual-polarized, 10.65 GHz vertically polarized and in addition to estimate the LAI 10.65 GHz dual-polarized. The rainfall- and data quality filters require additional bands, namely T_B at 18, 23, 36 and 89 GHz. Since the T_B-value measured at the satellite is composed of the brightness temperature of the surface and the atmosphere, a correction factor for atmospheric effects was introduced. All T_B-values used in the retrievals were subtracted with 3 Kelvin, since sky brightness is constant for most atmospheric conditions (Jackson, 1993).

For the Jackson algorithm, values of the *h*- and *b*-parameters are determined by means of an optimization run where different combinations of parameter values were evaluated. The combination producing the lowest statistical errors (see section 4.5) was selected and used in the retrievals. In the de Jeu algorithm the single scattering albedo was assumed to be constant at a value of 0.06. For computation of soil moisture content from dielectric constants the Wang-Schmugge dielectric mixing model (Wang and Schmugge, 1980) was used in all algorithms. Besides the roughness parameter *h*, which is assumed to be 0.2, the Wen algorithm requires some additional parameter settings. Some are defined in chapter 2.3.3 (Equation 16), other parameters are factors to account for the difference in transmissivity and single scattering albedo between the two frequencies that are used in the retrieval. The factor with which the transmissivity for 6 GHz is multiplied to obtain transmissivity for 10 GHz is 0.5 and a similar factor for single scattering albedo is 0.9. To calculate vegetation optical depth in the Wen algorithm the LAI is multiplied by a factor *k*, similar to the *b*-parameter in the Jackson algorithm. For the Mongolian dataset this factor is assumed to be 0.3. These three multipliers in the Wen algorithm are dependent on vegetation conditions and frequency and their values are estimated based on experience and trial and error. As is already mentioned, the *LAI* is estimated from T_B at 10.65 GHz. The estimation method is based on linear regression and it appeared to be difficult to obtain reasonable values: probably due to lack of pixels in the images *LAI* values differed significantly between various days. Retrievals were performed using both estimated and a constant value of *LAI*. The retrievals are compared statistically in Table 4 and graphically in the upper two plots of Figure 9. The constant value assumed for the Mongolian dataset was 0.5. This seemed a reasonable value considering the pictures of the landscape in Appendix B. As can be seen in Table 4 errors were significantly lower for the constant value of *LAI*. In Figure 9 it can be seen that with the estimated *LAI VSM* is constantly and significantly overestimated. Therefore, the retrievals for the Mongolian dataset were performed using a constant value of *LAI*. For a more detailed explanation of the values in the Table 4, see section 3.5.

Each algorithm was used to produce output information in terms of computed soil moisture, soil temperature, and, for the de Jeu algorithm, also optical depth for each of the 676 pixels for each day, as well as averaged over the pixels in the box described in section 3.2. Since overpass times of the satellite are at 1:30 AM and 1:30 PM, all computations were subject to diurnal cycles, so temperatures and *VSM* differed at

these times. Therefore, all analyses were separated in ascending and descending orbits.

Table 4: Statistics for the Wen algorithm retrievals for both datasets: using estimated *LAI* by T_B 10 GHz, or a constant value of 0.5 for the Mongolian dataset and 1.5 for the SMEX02 dataset.

Mongolia dataset	**Ascending**		**Descending**	
	Estimated LAI	**LAI = 0.5**	**Estimated LAI**	**LAI = 0.5**
Bias *VSM*	*5.0881*	*0.3803*	*6.4589*	*0.8995*
SEE *VSM*	*6.6554*	*2.9482*	*6.6554*	*2.8369*
% of days with error < 3 %	*29.8*	*78.7*	*20.9*	*72.1*
SMEX02 dataset	**Ascending**		**Descending**	
	Estimated LAI	**LAI = 1.5**	**Estimated LAI**	**LAI = 1.5**
Bias *VSM*	*-6.2592*	*-4.9727*	*1.1806*	*0.3156*
SEE *VSM*	*9.0527*	*8.4073*	*9.6089*	*4.5683*
% of days with error < 3 %	*25.0*	*37.5*	*62.5*	*62.5*

3.4 Running the algorithms for the SMEX02 dataset

In populated areas, there can be sources of Radio Frequency Interference (RFI) at the same frequency as the channels that are used by AMSR (Li et al., 2003). The effects of this RFI can be considerable. While this problem was relatively small in Mongolia due to the sparse population, in Iowa this problem was significant.

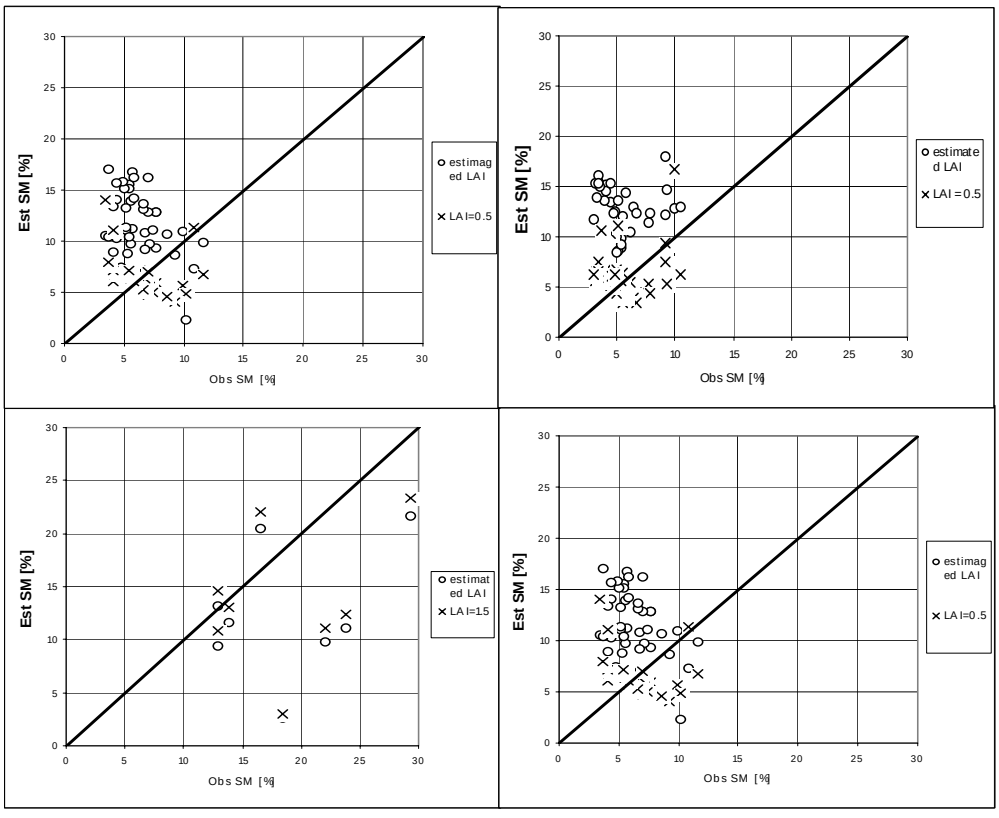


Figure 9: Retrieval results of the Wen algorithm for the Mongolian dataset in the upper two plots and the SMEX02 dataset in the lower two plots. Ascending orbits are left and descending orbits are right, for both estimated *LAI* using T_B 10 GHz and a constant LAI (0.5 for Mongolia and 1.5 for SMEX02).

Therefore, the algorithms could not be used in a straightforward manner with using C-band. As a result, brightness temperatures at the slightly higher X-band were used. The bands that will be used for each algorithm are 10 GHz horizontally polarized for the Jackson algorithm and 10 GHz dual polarized for the de Jeu algorithm. For the Wen algorithm, 10 GHz dual polarized and 18 GHz vertically polarized were used. Additionally, the rainfall and data quality filters required T_B values at 18, 23, 36 and 89 GHz. For the land use filter, the high resolution land cover dataset of the SMEX02 dataset was used, as well as the high-resolution *VWC* data for the Jackson algorithm. For soil texture however, the same (global) dataset was used as for the Mongolian dataset.
The surface roughness parameter *h* in the Jackson and Wen algorithms was assumed to be (similar to the Mongolian dataset) 0.2. In the de Jeu algorithm *h* was not taken into account. Since the VWC-values increase significantly over the experiment period (Figure 8), two values for the *b*-parameter were selected: in case the *VWC* was higher than 1.5 kg m^{-2} a value of 0.25 was used, if the *VWC* was lower than or equal to 1.5 kg m^{-2} a value of 0.35 was used. The Wang-Schmugge mixing model to compute soil moisture from the dielectric constant that was used in the Mongolia computations is only valid for L and C bands (Wang and Schmugge, 1980). Therefore, an alternative model, developed by Hallikainen (Hallikainen et al., 1985) was used. This model also used soil texture information and the dielectric constant to compute soil moisture content and it is described in Appendix D. As a correction for atmospheric effects, 3 Kelvin was subtracted from all T_B-values used in the retrievals in ascending orbits. For descending orbits, the atmosphere was assumed to be cooler, so effects would be smaller and here assumed negligible. Again, the Wen algorithm required some additional parameter settings, which are different from the Mongolian dataset due to different frequencies that are used and different vegetation conditions. Similarly to the Mongolian dataset, the regression-based estimated *LAI* and a constant value were compared, of which the results can be found in Table 4 and the lower two plots of Figure 9. Due to the higher amount of vegetation, as can be seen in the pictures in Appendix B, this time a *LAI* value of 1.5 was assumed. Although the differences were much smaller compared to the Mongolian dataset (Figure 9), statistical errors were slightly lower as can be seen in Table 4. Therefore, again the constant value is used in the retrievals. The multiplier to obtain vegetation optical depth from *LAI* is higher than in the Mongolia retrievals, namely 0.35. Factors to account for differences in transmissivity and single scattering albedo caused by frequency were both 0.9.

During the SMEX02 experiment, an aircraft sensor that was similar to AMSR-E was used as well: the aircraft based Polarimetric Scanning Radiometer (PSR). For details on PSR, see Jackson et al. (2002). Due to the smaller scale PSR operates on, it is less sensitive to RFI; therefore the C-band of PSR could be used. PSR data were available in a grid of 800 x 800 m, but since surface temperature was only available at AMSR-E footprint scale (AMSR-E 37 GHz V band), the PSR pixels were averaged to AMSR-E footprints. Also, PSR overpasses only occurred during daytime, so only ascending AMSR overpasses were used. Parameters that depend on frequency were changed with respect to the retrievals using AMSR data for the Jackson algorithm: instead of *b*-values of 0.35 (*VWC* < 1.5 kg m^{-2}) and 0.25 (*VWC* > 1.5 kg m^{-2}), values of respectively 0.33 and 0.23 were used. Changes in *h* were assumed negligible, so *h* remained 0.2 in both the Jackson and Wen algorithms. In the de Jeu algorithm there were no parameters to be changed: the polynomials in Table 3 are frequency independent. The Hallikainen mixing model is frequency dependent and therefore its

parameters were changed according to Table D.1 in Appendix D. Since this is again a different combination of frequencies and vegetation conditions, some parameters in the Wen algorithm were changed as well: although the *LAI* is, similar to retrievals with AMSR data, 1.5, the vegetation multiplier *k*, which is frequency dependent, is 0.55. The other multipliers to account for frequency differences are both set at 0.9. These values are set, as already mentioned in the previous section, based on experience and trial and error. An extensive optimalization of the parameters for all conditions in this study was not possible due to lack of time and the fact that the Wen algorithm, as it was used in this study, was not very computationally efficient.

Since this time no ready-made matrix was available, retrieval is performed on a footprint base, where those footprints were selected that fall in the study area. This was roughly between latitudes 41.6º and 42.7º N, and longitudes 93.2º and 93.8º W, for precise coordinates see Figure 5. The latitude/longitude coordinates of each pixel were converted to the corresponding pixel in the *VWC* and land cover matrices (section 4.1). After retrieval, the *VSM* and T_S were averaged over the mentioned footprints and compared to the average of the observations in the area. Over all orbits (separated for ascending and descending), *Bias* and *SEE* were computed for *VSM* and T_S.

3.5 Statistical assessment

After running all algorithms output information consisted of a file for each day and each orbit containing retrieved *VSM*, T_S and vegetation parameters. In addition, each algorithm produced two files (ascending and descending orbits) containing average values of these parameters over the pixels in the box described in section 3.1 for the Mongolian dataset, or over the research area for SMEX02. Over these averages, the following statistical quantities were computed over time: the Standard Error of Estimate (*SEE*) and the *Bias*. The latter can be described as follows:

$$BIAS = \frac{\sum_{i=1}^{N}(EST_i - OBS_i)}{N} \quad . \tag{22}$$

Where *EST* is the retrieved value, *OBS* is the observed value and *N* is the number of observations. The *SEE* can be described by:

$$SEE = \sqrt{\frac{\sum_{i=1}^{N}(EST_i - OBS_i)^2}{N}} \tag{23}$$

A third quantity that was computed was the percentage of observations where the difference between estimated and observed values is smaller than 3 %. All three quantities can be found for each retrieval in Tables 5 and 6. Besides these statistical quantities, for each algorithm temporal plots (observed and estimated average *VSM* and T_S over time) and scatter plots (observed versus estimated) were produced.

4 Results

In this chapter, the results of the retrievals with all algorithms are described for both datasets. First, the results of the Mongolian dataset are discussed.

4.1 Mongolian dataset

In this section, the results of the analyses for the Mongolian dataset are described. First, parameters *h* and *b* in the Jackson algorithm were optimized as described in the previous section. Graphical results can be found in Figure 10, where it appears that the results differ slightly between ascending and descending orbits, but overall the optimal value for the *b*-parameter seems to be 0.3: a little over 0.3 for ascending and a little under 0.3 for descending orbits. For this value, the *SEE* is lowest and the *Bias* is closest to zero. For the *h*-parameter, the optimal value for ascending orbits seems to be lower for descending orbits as well compared to ascending orbits, but overall the best value seems to be 0.2. These values were used in the remaining retrievals.

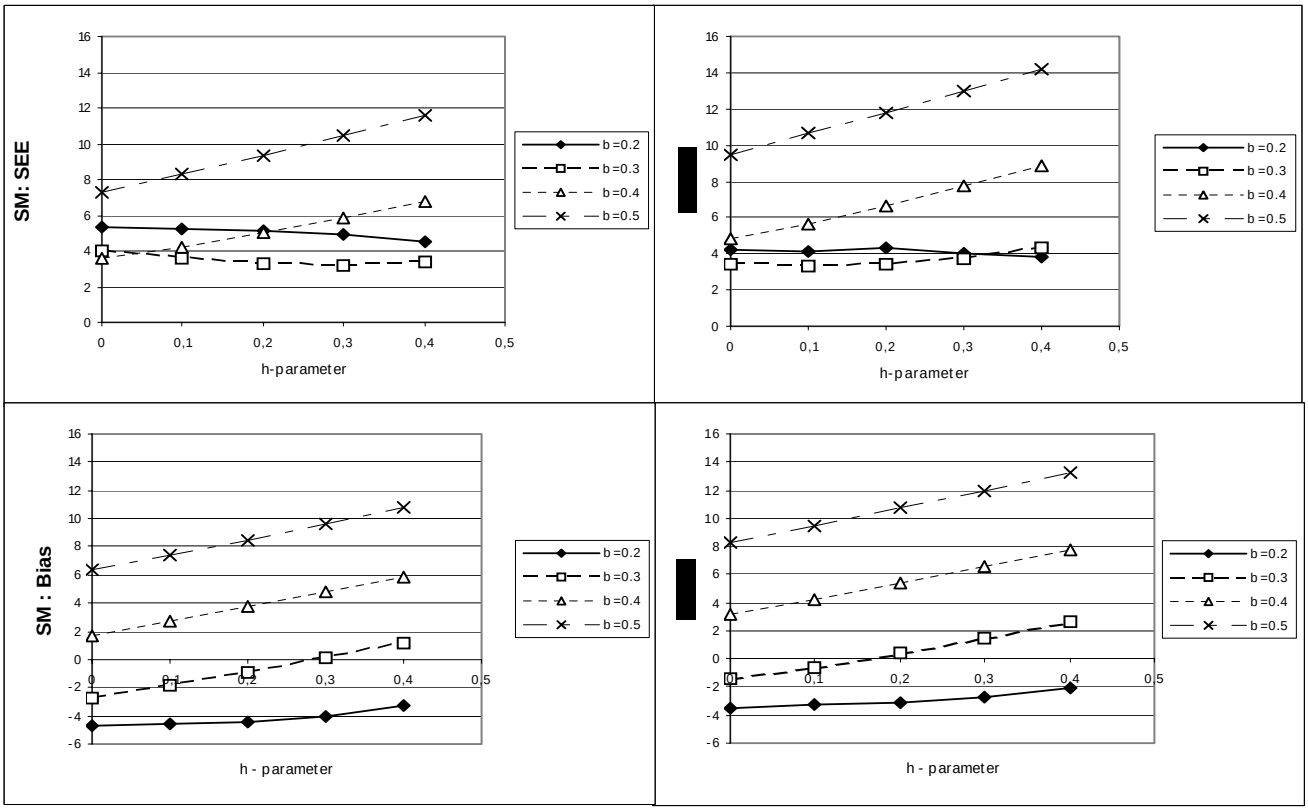


Figure 10: Optimization of the *b*- and *h*-parameters for the Jackson algorithm and the Mongolian dataset. In the upper graphs *SEE*s for ascending (left) and descending (right) orbits are shown and in the lower graphs the *Bias* for ascending (left) and descending (right) orbits.

Statistical results of all retrievals can be found in Table 5. As follows from Equations 22 and 23, negative values for *Bias* mean that *VSM* was on average underestimated, positive values indicate overestimation. It appeared that all algorithms on average overestimated the *VSM*, except for the Jackson algorithm at ascending orbits, where *VSM* was slightly underestimated. In Figure 11 the distribution of the errors can be

seen: it appears that at the majority of the days for all algorithms *VSM* was slightly underestimated (max. about 5 percent), while at only a few days *VSM* was strongly overestimated (sometimes up to about 10 percent). The amount of days at which the difference between observed and estimated *VSM* was smaller than 3 % is on average about 70 % (Table 5; somewhat higher for the Wen algorithm, somewhat lower for the de Jeu algorithm), resulting in very reasonable *SEE*s for all algorithms: all *SEE*s were between 2.5 and 4.5 %. Especially *SEE*s for the Wen algorithm are very low with only a little over 2 %. Results were about the same for ascending and descending orbits. An explanation for the phenomenon that on a small number of days a large overestimation occurred can be found in Figure 12, where the observed and estimated *VSM* are shown over time for all algorithms. Here it can be seen that large peaks in estimated *VSM* occurred: for ascending orbits about 3 times during the 3-month period, for descending orbits about 5 times. These peaks are the overestimated points in Figure 11. It seems that overall the estimated *VSM* was in the same order of magnitude as the observed values, but when sudden changes occurred, estimated and observed *VSM* reacted differently: estimated *VSM* values rose and dropped in a very short time (typically one or two days), while in the observed *VSM* the results of such an event lasted much longer. The underestimations in Figure 11 correspond with the days immediately after the peaks. This can also explain the relatively big difference between estimations and observations in the first 10 days of the simulation: probably a similar sudden rise just before the start of the observed period caused higher observed levels of soil moisture. The temporal pattern described above can be seen in Figure 12 for all algorithms.

Table 5: Summarizing statistics for the Mongolian dataset.

Mongolia Match Up	**Jackson**		**de Jeu**		**Wen**	
	Ascending	**Descending**	**Ascending**	**Descending**	**Ascending**	**Descending**
Bias *VSM*	*-0.9144*	*0.3605*	*1.6498*	*0.8072*	*0.3803*	*0.8995*
SEE *VSM*	*3.2350*	*3.3799*	*4.1958*	*4.1643*	*2.9482*	*2.8369*
% of days with error < 3% *VSM*	*70.2*	*74.4*	*63.8*	*69.8*	*78.7*	*72.1*
Bias T_S	*0.0577*	*-3.1113*	*0.0577*	*-3.1112*	*13.7744*	*9.2084*
SEE T_S	*1.5964*	*3.2149*	*1.5964*	*3.2148*	*13.8178*	*9.2717*
% of days with error < 3 K T_S	95.7	37.2	95.7	37.2	0	0

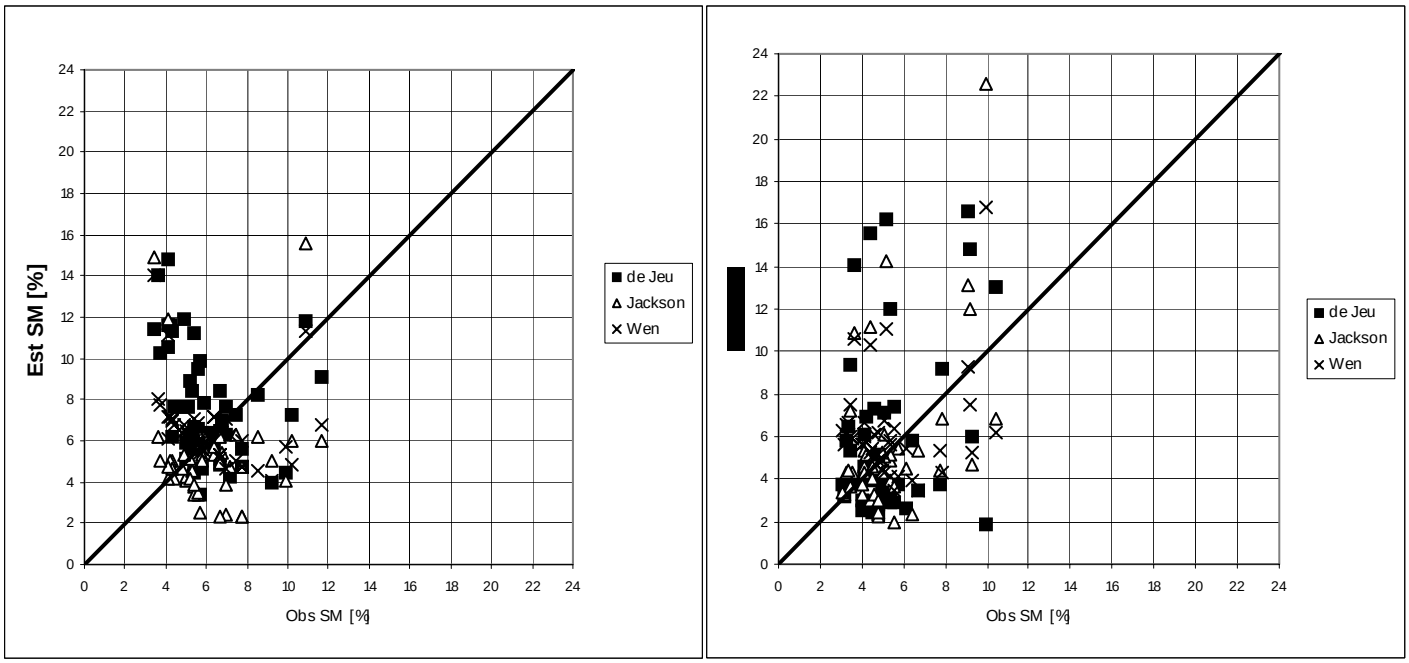


Figure 11: Scatter plots of observed versus estimated *VSM* for the Mongolian dataset and all algorithms, ascending (left) and descending (right) orbits.

As can be seen in Figure 13, the sudden changes in estimated *VSM* correspond with the brightness temperature input data: drops there also occurred in a very short period and correspond to the rises in estimated *VSM*. The anomaly between estimated and observed data can partly be explained by the fact that observations took place at a depth of 3 cm and the sampling depth of the 6 GHz radiation that was used is only in the order of about 1.25 cm. The closer to the surface, the faster the *VSM* reacts to rainfall events and evaporation. After a rainfall event the *VSM* at 1.25 cm decreases faster than the *VSM* at a depth of 3 cm. Also, as can be seen at the photographs in Appendix B, the Mongolian soil contains a lot of stones and rocks. It is known that the presence of large amounts of rocks in the soil will reduce the sensitivity to *VSM* for C-band data (Jackson et al., 1992).

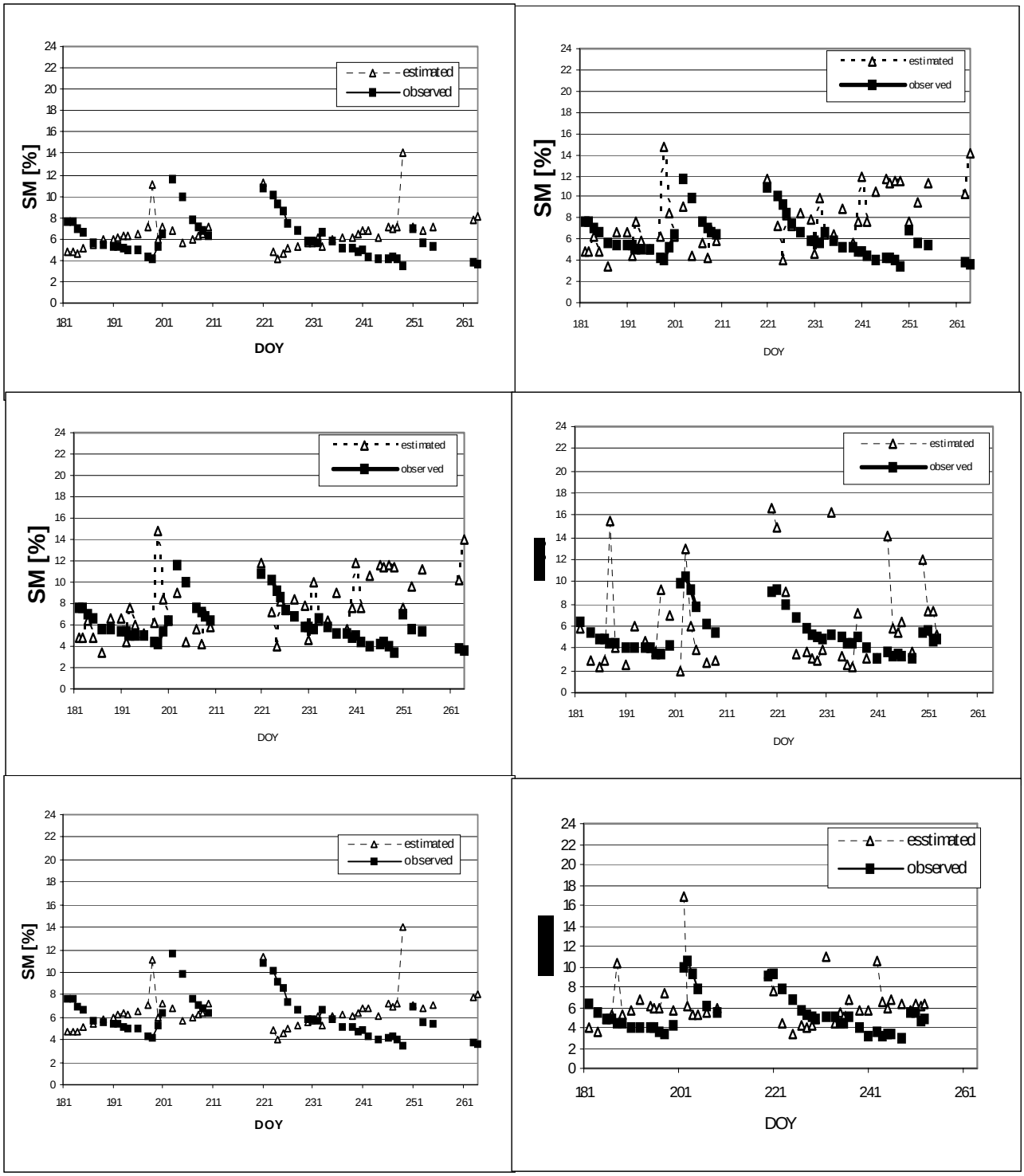


Figure 12: Observed and estimated *VSM* over time for all algorithms for the Mongolian dataset, ascending and descending orbits. From top to bottom Jackson, de Jeu and Wen algorithms, ascending orbits left and descending orbits right. Triangles are estimated, squares observed *VSM*.

All algorithms showed a roughly similar pattern, but in the last ±30 days of the period, estimated *VSM* for the de Jeu and Wen algorithms was significantly higher than the estimations for the Jackson algorithm and the observed *VSM*, especially for ascending orbits. As can be seen in Figure 13, this rise in *VSM* corresponds to a slight decrease in T_B over the same period, which is indeed more distinct in the ascending orbits. A reason for this occurring only in the de Jeu and Wen algorithms can be found in the

vegetation data: the *VWC*, which was indirectly presented to the Jackson algorithm as ancillary data (Figure 7), rose the first part of the period, then it dropped strongly during the last 30 days, from a value of about 0.5 kg m^{-2} to 0.36 kg m^{-2}. This drop in *VWC* compensated the lower T_B values in the Jackson algorithm, so the estimated *VSM* remained at the same level. The other algorithms lacked this ancillary data and computed both *VSM* and vegetation optical depth from T_B. The resulting optical depth for the Jeu algorithm can be found in Figure 14 and stayed more or less at the same level (about 0.2) throughout the period, except for some negative values that occur at days with excessively high or low T_B-values (Figure 13), presumably caused by rainfall events. Optical depth as computed by the Wen algorithm stayed at one level due to the constant value of *LAI* that was assumed. Due to this lack of compensation by vegetation information, *VSM* reacted to the drop in T_B causing a higher *VSM*-level for the rest of the period. A second difference between the two algorithms was that the estimated *VSM* by the Jackson algorithm changed rather 'smoothly' over time, while the estimated *VSM* by de Jeu fluctuated more abruptly. The Wen algorithm shows a bit of both: for ascending orbits estimated *VSM* varies rather smoothly, for descending orbits changes are rather abruptly. This may partly be explained by the ancillary vegetation data as well: in the Jackson algorithm *VSM* was not only dependent on T_B values but also on the, much more gradually changing, *VWC*. This had a smoothening effect on estimated *VSM*, while the estimated *VSM* in the de Jeu algorithm responded directly to every change in T_B, which also fluctuated rather abruptly, see Figure 13. The Wen algorithm used T_B values at two frequencies as opposed to the de Jeu algorithm, which also has a smoothening effect. From Figure 12 it also appears that the low *SEE* and *Bias* for the Wen algorithm is mainly caused by the lower value of the 'peaks' compared to the Jackson algorithm. Overall, apart from the peaks, differences between observed and estimated values are smallest for the Jackson algorithm.

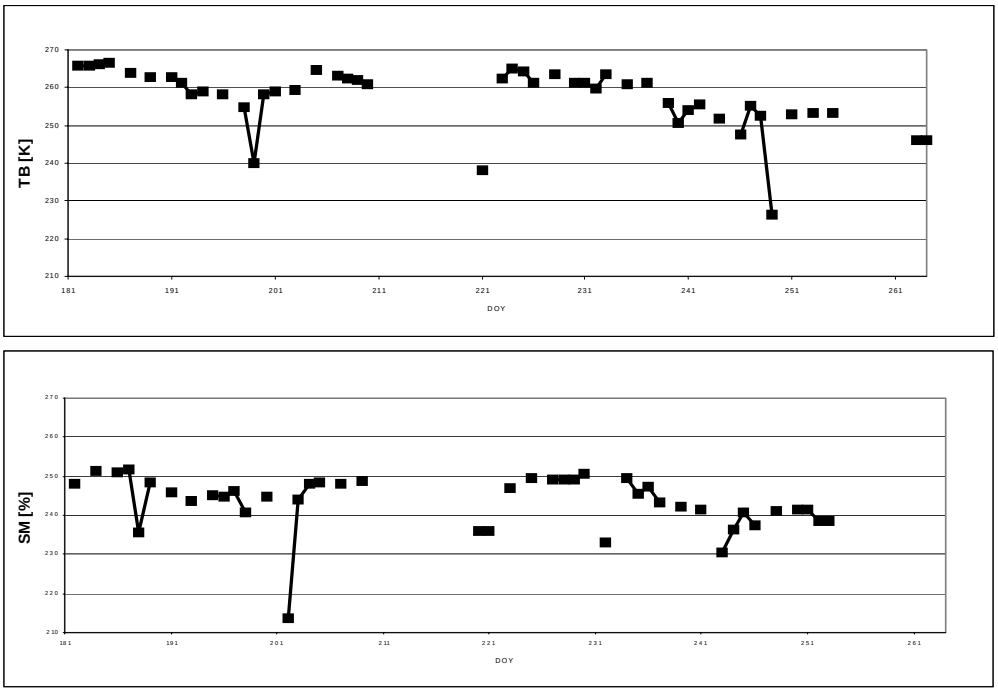


Figure 13: H-polarized brightness temperature over the Mongolian experiment period for both ascending (top) and descending (bottom) orbits.

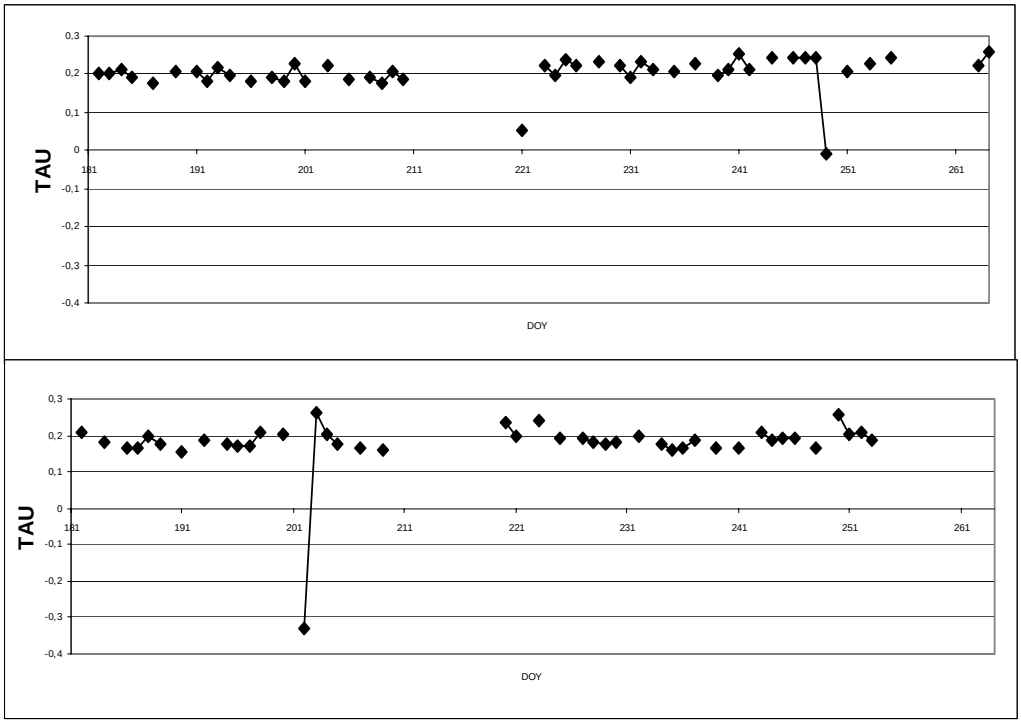


Figure 14: Vegetation optical depth over the Mongolian experiment period as computed by the de Jeu algorithm for both ascending (top) and descending

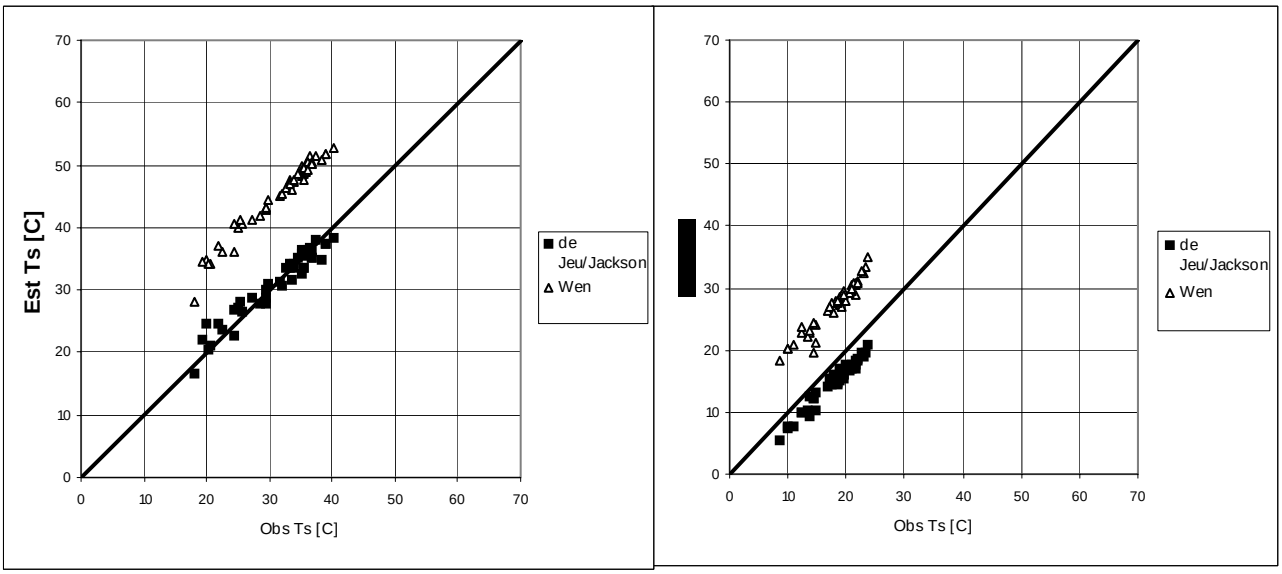


Figure 15: Observed versus estimated surface temperature for the Mongolian dataset as computed by Equation 4 (blocks) and by the Wen algorithm (triangles) for both ascending (left) and descending (right) orbits.

In addition, the T_S was evaluated: the Jackson- and de Jeu algorithm used Equation 4 for that purpose, in the Wen algorithm the non-linear equation system was solved for both *VSM* and T_S. The results are shown in Figure 15 and the statistics are summarized in Table 5. Especially for ascending orbits in the Jackson and de Jeu algorithms observed and estimated values appear to be very close: the *SEE* was only about 1.5 K and 96% of the days were within 3 K of the observed value. For descending orbits however, all estimations were structurally below the observed values, by about 2 to 3 Kelvin; almost 40 % of the days were within 3 K of the observations. The *SEE* for descending orbits was almost completely explained by the *Bias*, while the *SEE* for ascending orbits was relatively large considering the very small *Bias*. This may be explained by the fact that ascending orbits correspond to the afternoon (1:30 pm), when local differences can be larger than at night, when the descending orbits take place, e.g. in or out the shadow of the observation station. The

estimated T_S values as obtained by the Wen algorithm were structurally higher than observed values, as can be seen in Figure 15. During daytime T_S was overestimated by more than 13 K, during nighttime descending overpasses by about 9 K. Although the Bias was very high, the *SEE* is hardly higher, indicating almost no deviation around the average. No clear physical explanation exists for the large *Bias*, probably it is caused by the parameter settings, which are chosen to obtain reasonable *VSM* results. With other parameter settings, a reasonable T_S value may be found, causing less reliable *VSM* values. As already mentioned, an extensive optimization could not be performed. As in the Jackson and de Jeu algorithms T_S was underestimated for descending orbits, the overestimation in the Wen algorithm is smaller for descending orbits.

The structural underestimation for the descending orbits may, similar to *VSM*, be explained by the observation depth: T_S was estimated using T_B values at 37 GHz; this corresponds to a sampling depth of about 2 mm, while the observation took place at a depth of 3 cm. The closer to the surface, the faster the soil warms up and cools down compared to deeper layers. By the time of the descending overpass (1.30 am), the surface (at 2 mm depth) has cooled down faster than the soil at 3 cm, resulting in a lower estimated T_S compared to the observed T_S. This would suggest a slightly overestimated surface temperature during daytime (ascending orbits) for the Jackson and de Jeu algorithms, which can indeed be seen in Figure 15, albeit less clear than the underestimation for descending orbits.

4.2 SMEX02

In this section, first the results are discussed of the retrievals with AMSR-E X-band for the SMEX02 dataset, after that the results of the retrievals with PSR C-band will be described. In Table 6, the summarizing statistics of all retrievals can be found. In Figure 16 the observed and estimated *VSM* are plotted for the retrievals with the AMSR-E T_B data, for all three algorithms and both ascending and descending orbits. From Table 6 it appears that for this dataset the *VSM* was mostly underestimated, as opposed to the Mongolian dataset. For the Jackson algorithm, the *Bias* was relatively small for both orbits (less than 1 %), but the *SEE* was relatively high, over 5 %. For descending orbits the algorithm performed slightly better, as can be derived from the amount of days where the error was less than 3 % as well: 50 % for descending orbits, 12.5 % for ascending orbits. The *Bias* was relatively small because *VSM* was almost equally often over- and underestimated, as can be seen in Figure 16 as well. For the de Jeu algorithm, the difference between descending and ascending orbits was bigger, which was mainly caused by 3 days with ascending orbits that heavily underestimated *VSM*. Three similar points cause a high negative bias and SEE in the Wen algorithm. An explanation for this phenomenon follows later on. These three points caused a high negative *Bias* (-6.5%) and a very large *SEE* (more than 9%). Apart from these 3 days, the algorithm didn't perform worse than for descending orbits: the number of days with errors smaller than 3 % was even higher than for descending orbits. The Wen algorithm also shows a large difference in

Table 6: Summarizing statistics for the SMEX02 dataset.

SMEX02	Jackson		de Jeu		Wen	
	Ascending	**Descending**	**Ascending**	**Descending**	**Ascending**	**Descending**
Bias *VSM*	*-0.9989*	*-0.4636*	*-6.5661*	*-2.2169*	*-4.9727*	*0.3156*
SEE *VSM*	*5.4405*	*5.0948*	*9.1403*	*5.1203*	*8.4073*	*4.5683*
% of days with error < 3% *VSM*	*12.5*	*50.0*	*25.0*	*20.0*	*37.5*	*50.0*
Bias T_S	*-8.9115*	*-13.2631*	*-8.9345*	*-13.2751*	*11.2350*	*17.4507*
SEE T_S	*9.4148*	*15.6898*	*9.4272*	*15.6998*	*17.5059*	*19.4590*
% of days with error < 3 K T_S	*0*	*0*	*0*	*0*	*0*	*0*
	PSR C-band (ascending)	**Descending without days 180 & 188***	**PSR C-band (ascending)**	**Descending without days 180 & 188 ***	**PSR C-band (ascending)**	**Descending without days 180 & 188 ***
Bias *VSM*	*-2.1385*	*-0.5865*	*1.5972*	*-4.3236*	*4.3288*	*0.6134*
SEE *VSM*	*3.6606*	*3.8066*	*5.4919*	*4.7647*	*12.7106*	*5.0777*
% of days with error < 3% *VSM*	*50.0*	*62.5*	*33.3*	*25.0*	*33.3*	*37.5*
Bias T_S	*-7.7906*	*-15.4666*	*-7.7997*	*-15.4816*	*4.0460*	*15.0778*
SEE T_S	*8.3488*	*16.1595*	*8.3599*	*16.1716*	*6.7101*	*15.7983*
% of days with error < 3 K T_S	*0*	*0*	*0*	*0*	*16.7*	*0*

** these two days had in the used dataset exceptionally high or low brightness temperature values. In another datasets of the same satellite they didn't occur at all. Therefore here these days are omitted from the statistics' computation.*

performance between ascending and descending orbits. For descending orbits *SEE* is good: less than 5%. As opposed to the other algorithms, here *VSM* is a little overestimated. For ascending orbits however, *VSM* is as much underestimated as with the other algorithms. As can be seen in Table 6, statistics were also computed for descending orbits in the case when 2 days were omitted, because of excessive high or low T_B-values at those days. This is illustrated in Figure 17: day number 180 had a 30 K higher T_B-value than the other days and day number 188 was 10 K lower. For the Jackson and de Jeu algorithms *SEE* values decreased compared to the situation with those two days, especially for the Jackson algorithm, while the *Bias*-es increased for both algorithms. In the Wen algorithm the bias increased, but also the *SEE*. An explanation for this can be found in Figure 18. Here it appears that at day numbers 180 and 188 *VSM* was overestimated by the de Jeu algorithm, while at other days *VSM* was underestimated. Leaving out these days thus resulted in a higher (negative) *Bias*. Since in the Jacksonalgorithm days where *VSM* was underestimated and overestimated were more evenly

spread, the *Bias* didn't change significantly. The *SEE* decreased because of the large errors at the omitted days compared to the other days. For the Wen algorithm

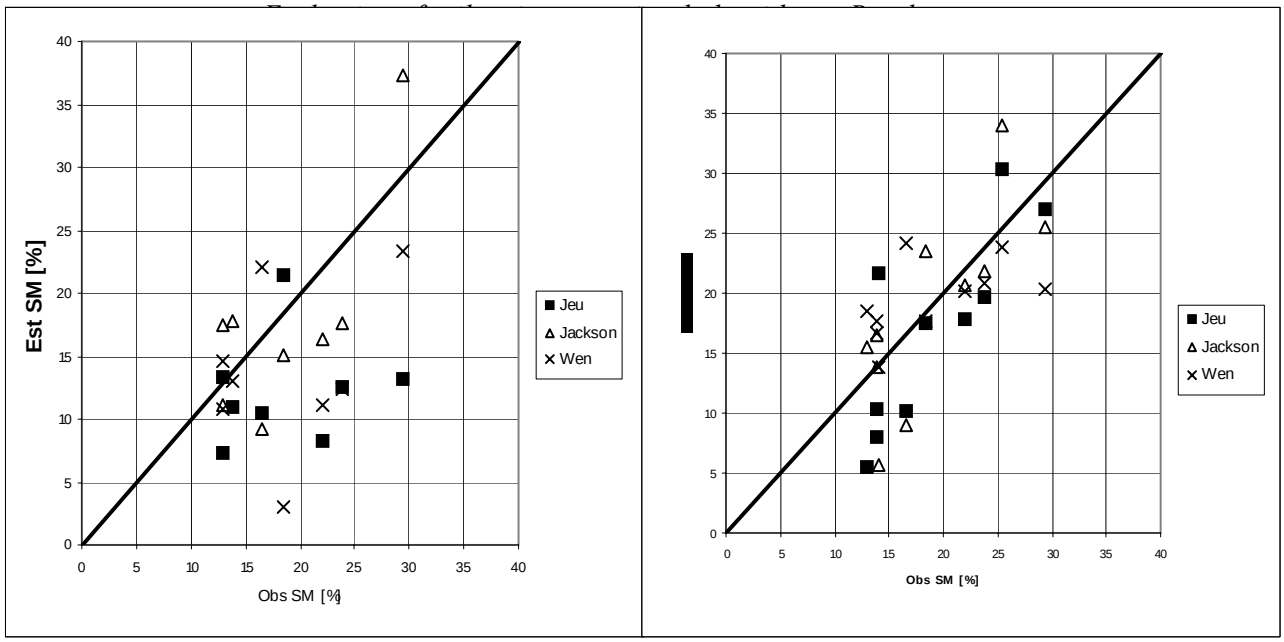


Figure 16: Scatter plots of observed versus estimated *VSM* for the SMEX02 dataset and all algorithms, ascending (left) and descending orbits (right). AMSR-E T_B values are used.

however, the errors were smallest at the two omitted days. Therefore, both *Bias* and *SEE* increased with an equal amount.

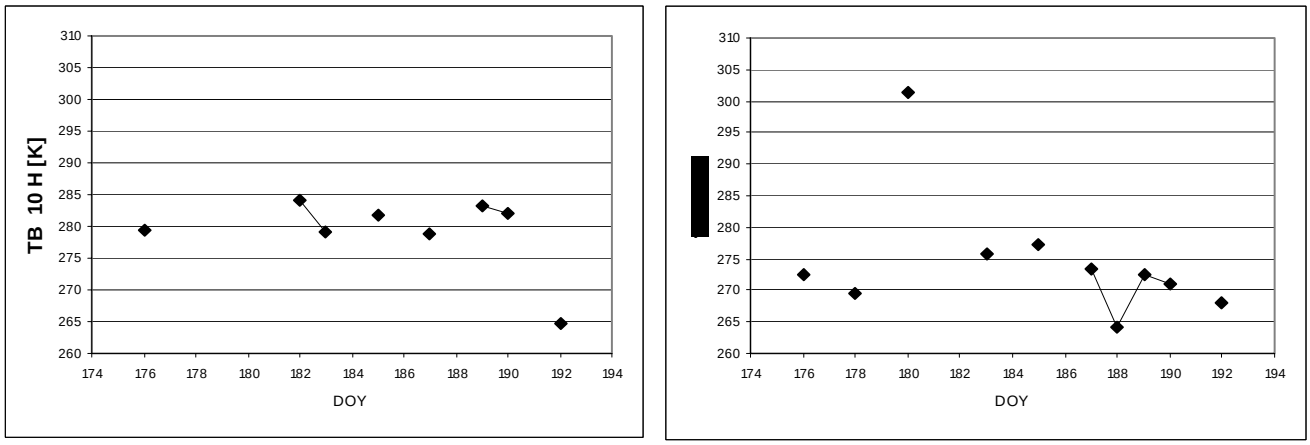


Figure 17: H-polarized brightness temperature over the SMEX02 experiment period, for ascending and descending orbits.

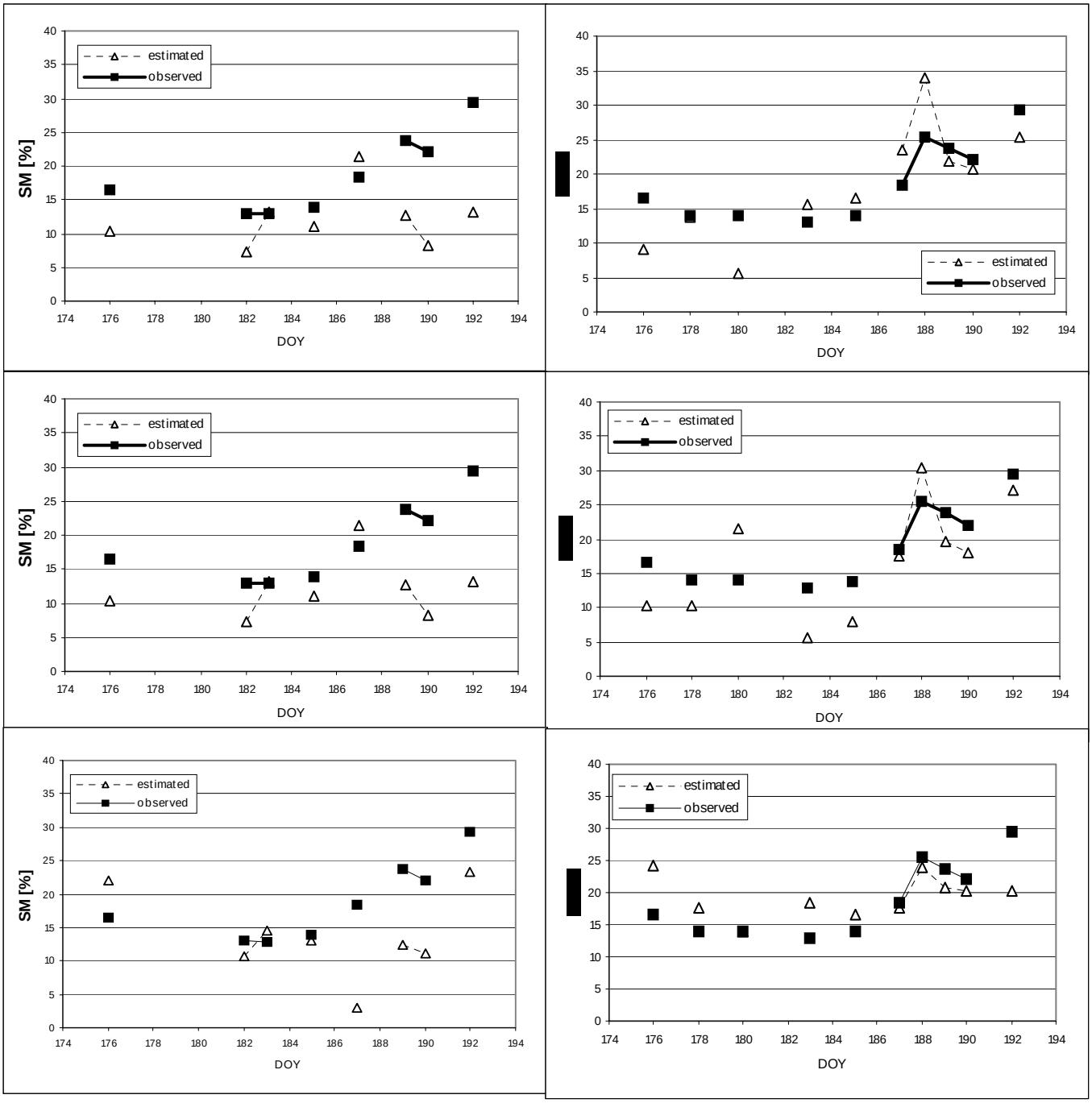


Figure 18: Observed and estimated *VSM* over time for all algorithms for the SMEX02 dataset, ascending and descending orbits. From top to bottom the Jackson, de Jeu and Wen algorithms, ascending left, descending right. Triangles are estimated, squares observed *VSM*.

The three days causing very big errors in the de Jeu and Wen algorithms all occurred in the end of the retrieval period as can be seen in Figure 18. This cannot be explained using T_B-values because T_B also drops towards the end of the retrieval, but problems occurred in this part of the retrieval due to the dense vegetation layer: *VWC* values were about 2 kg m^{-2} (Figure 8), which is higher than values at which retrieval is considered possible by AMSR-E (see section 1.2). Early in the period *VWC*-values were a little over 1 kg m^{-2}, causing less vegetation related problems in the retrieval. In the Jackson algorithm this phenomenon is less clear because similar to the Mongolian

dataset, by the *VWC* information provided to the Jackson algorithm vegetation effects are taken into account.

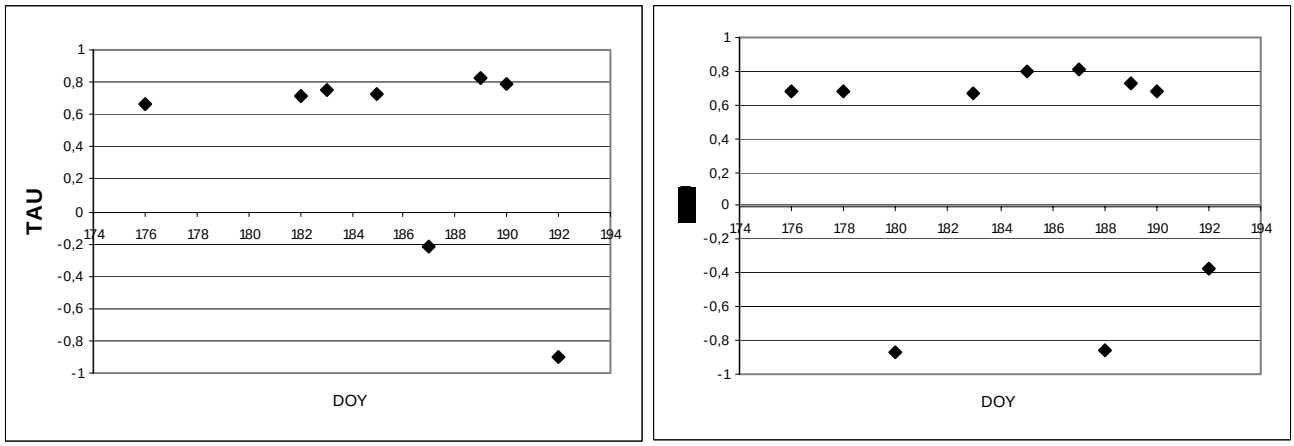


Figure 19: Vegetation optical depth over the SMEX02 experiment period as computed by the de Jeu algorithm for both ascending (left) and descending (right) orbits.

All algorithms performed significantly better for the descending orbits compared to the ascending orbits, while observations took place during daytime and thus close to the ascending overpass. Also due to this fact, the error introduced by T_S (also measured during daytime) was smaller for ascending overpasses (see Table 6). One explanation for this could be in the atmospheric effects: although a correction of 3 K for the ascending orbits was assumed, the atmospheric brightness maybe was higher than this value due to higher temperatures causing an overestimation of the T_B of the surface and subsequently an underestimation of *VSM*. Another explanation is in the dense canopy layer: T_S was estimated from T_B at 37 GHz, which has a very small penetration depth and thus measured the canopy temperature rather than T_S. In daytime when the sun is shining, the difference between these temperatures is bigger than at night, causing an underestimation of T_S, an overestimation of the soil emissivity and subsequently an underestimation of *VSM*. In the results for the de Jeu and Wen algorithms this effect was more distinctly visible because, similar to the Mongolian dataset, the *VWC*, which increased strongly over the experiment period, was presented to the Jackson algorithm as ancillary data. This caused the *VSM* for the Jackson algorithm to increase, while for the Jeu algorithm, where computed optical depth values remained rather constant (see Figure 19), *VSM* was underestimated throughout the experiment period and, due to vegetation issues mentioned earlier, especially in the latter part of the simulation. The Wen algorithm performed similar due to the constant *LAI* value that was assumed. The optical depth as computed by the de Jeu algorithm (Figure 19) contained some days where the optical depth turned out to be negative. Similar to the Mongolian dataset, this was caused by excessively high or low T_B-values (such as rainfall events, or when the soil seems locally saturated). Note that at both day numbers 180 and 188 (of which the descending orbits were omitted) the optical depth was negative.

As can be seen in Figure 20 and Table 6, for the Jackson and de Jeu algorithms T_S was underestimated for both ascending and descending orbits, while in the Wen algorithm T_S was, similar to the Mongolian dataset, consequently overestimated. Since temperature observations took place in the afternoon, *Bias* and *SEE* of T_S for

descending orbits were much higher (about 5 K) than for ascending orbits, since temperatures during nighttime are much lower. For ascending orbits, the underestimation was about 8 K, which was explained before: T_S was estimated from T_B at 37 GHz, which measured the temperature of the canopy more than that of the soil (which was observed). Especially during daytime, the temperature of the canopy is lower than the soil temperature, causing a large underestimation of T_S and thus also introducing an error to the *VSM* retrieval. The Wen algorithm structurally overestimates T_S, predicting roughly the same values for both ascending and descending orbits (Figure 20), causing a high overestimation especially for descending orbits with a bias of more than 17 K. As explained before for the Mongolian dataset these large errors for the Wen algorithm are caused by the parameter settings: the algorithm included a large number of guess values; other settings could have lead to better T_S estimations and probably poorer *VSM* estimations. An extensive optimalization for both *VSM* and T_S was not possible, as is explained before as well.

Next, the results of retrievals using C-band from PSR are discussed. Statistical results can be found in Table 6, observed and retrieved *VSM* values are plotted in Figures 21 and 22.

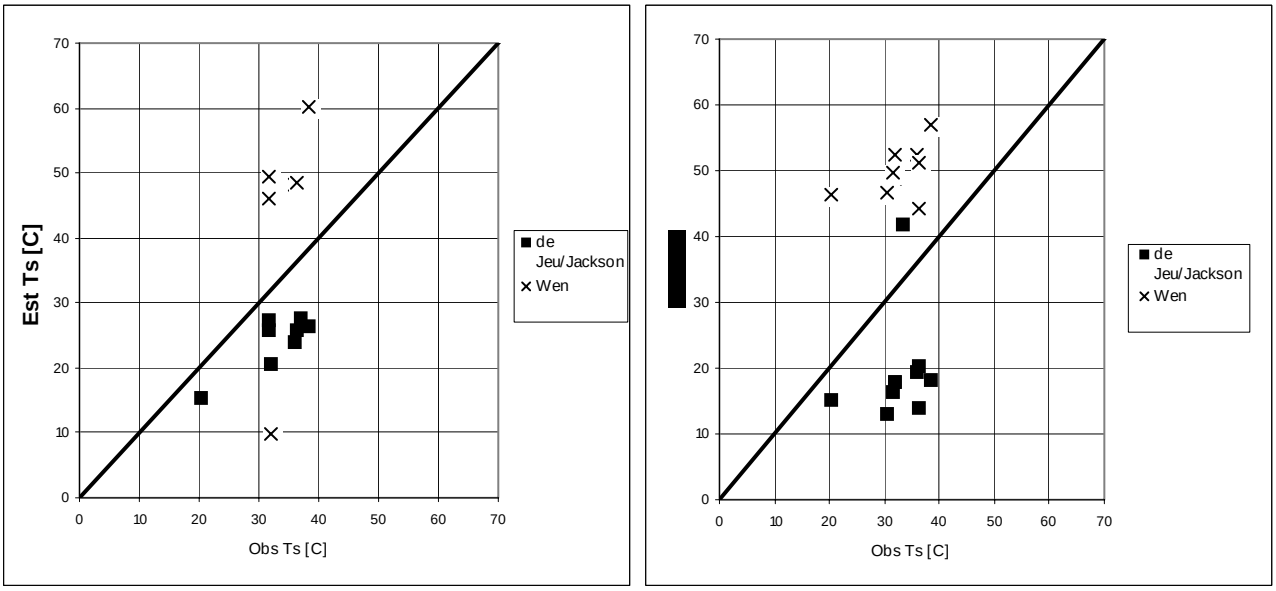


Figure 20: Observed versus estimated T_S for the SMEX02 dataset, for ascending and descending orbits.

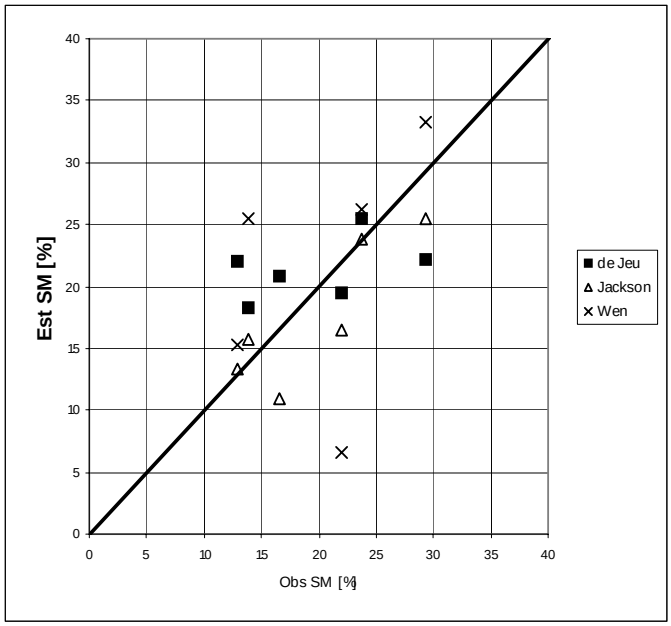


Figure 21: Observed versus estimated VSM for the SMEX02 dataset using PSR C-band data.

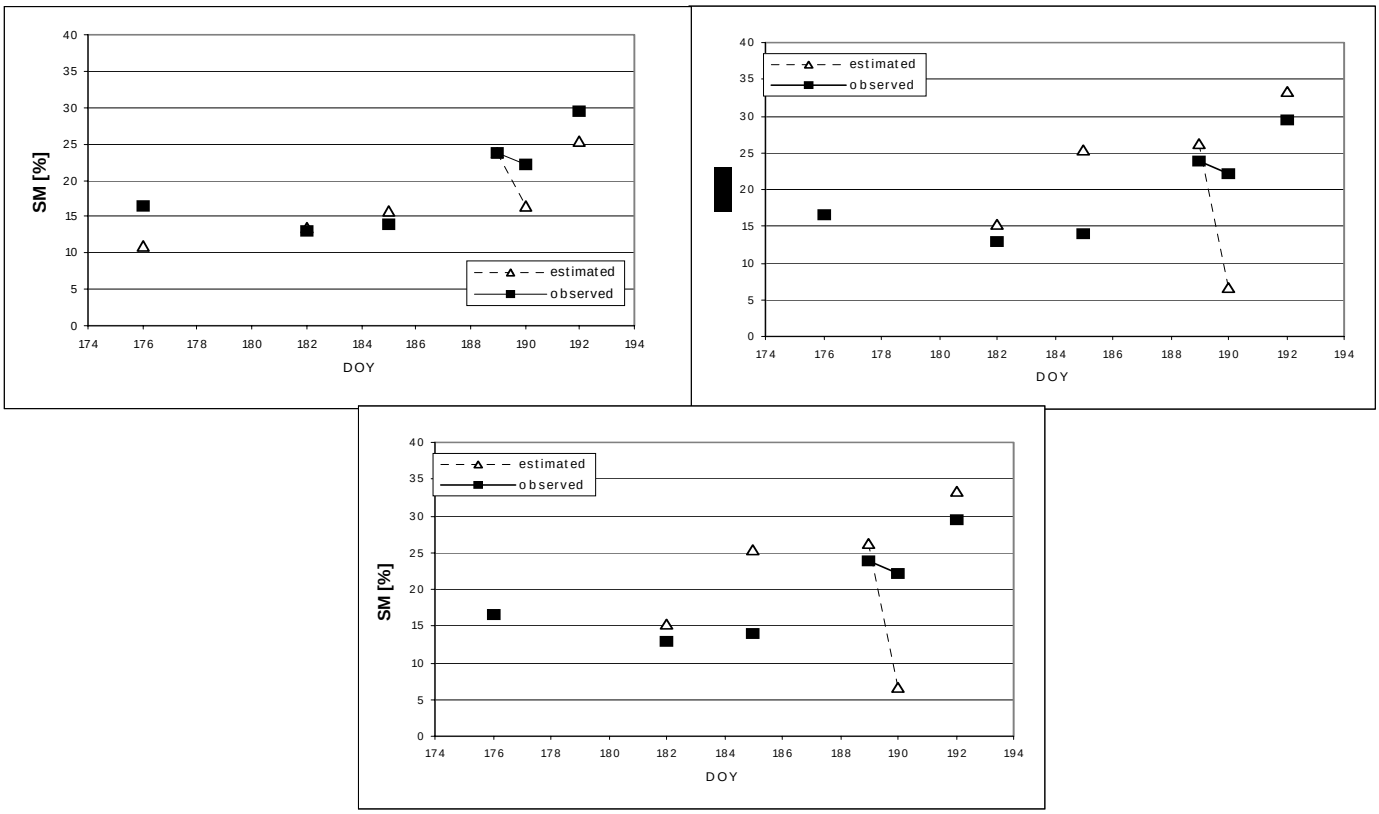


Figure 22: Observed and estimated VSM for theSMEX02 dataset with PSR C-band T_B, for the Jackson algorithm (top left), the de Jeu algorithm (top right) and the Wen algorithm (bottom).

The PSR C-band T_B-values (Figure 23) appeared to be on average 3 to 4 Kelvin lower

than the AMSR X-band T_B-values (Figure 17). This is due to differences in dielectric-, vegetation- and roughness effects between C- and X-band and in correspondence with reported differences between PSR C-band and X-band (Jackson et al., 2002). The

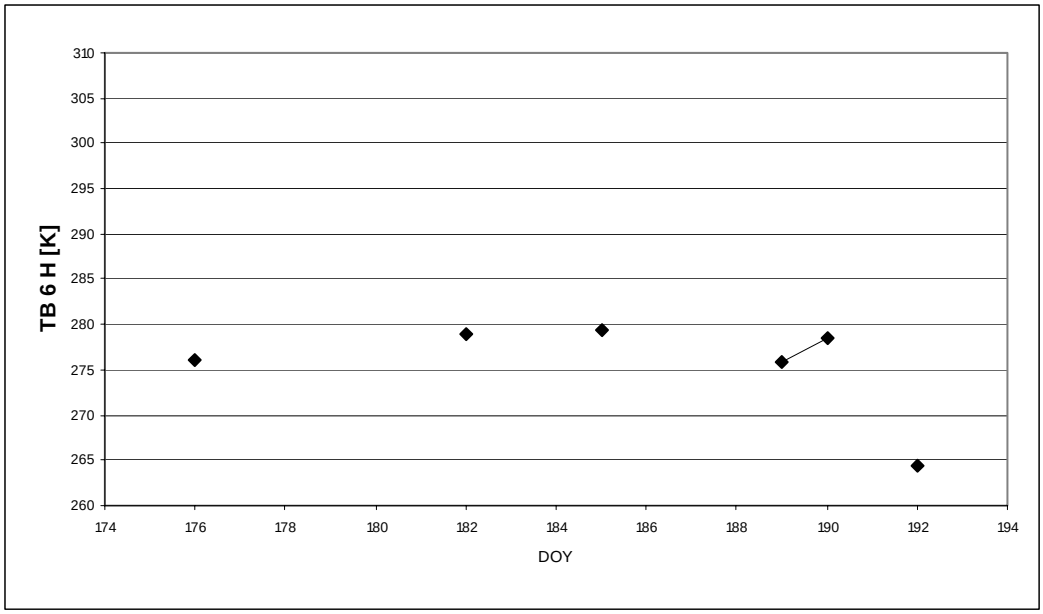


Figure 23: H-polarized T_B values of PSR for the SMEX02 experiment period.

Jackson algorithm results were in good accordance with observations with a *SEE* of 3.6 %.

At some days however, *VSM* was underestimated resulting in a relatively high negative *Bias* and only 50 % of the days had an error of less than 3 % (Table 6). These days mainly occurred in the latter part of the period when due to the dense canopy layer the problems occurred that were described earlier, since PSR overpasses occurred in daytime and were thus comparable with ascending AMSR overpasses. The de Jeu algorithm mainly overestimated *VSM* which can be explained by lower T_B-values compared to the X-band. Since the algorithm does not have frequency dependent parameters that can be adjusted, the lower T_B values were not compensated. Nevertheless, the algorithm performed better than for the AMSR-data with a *SEE* of 5.5. Also the *Bias* was relatively small, because in the latter part of the retrieval *VSM* was underestimated, compensating the earlier overestimation. This underestimation was caused by the dense canopy layer as was explained earlier. For the Wen algorithm especially the *SEE* is very high with 12 %, the *Bias* is relatively small because over- and underestimations compensate each other. As can be seen in Figure 22, there are mainly 2 points that cause the large errors: at day number 176 the estimated *VSM* is off the scale with approximately 45 %. At day number 190 *VSM* is underestimated by about 15 %. For the other days the errors are relatively small. An explanation is again that the parameter settings should be determined more carefully for changing vegetation and radiation conditions.

In Figure 23 the PSR C-band T_B-values are shown: no anomalies are present and the pattern of the estimated *VSM* is in accordance with the pattern showing in the T_B. Surface temperature as computed by the Wen algorithm is surprisingly well in accordance with observed values compared to the AMSR-E retrieval (Table 6 and

Figure 24): a *SEE* of less than 7 K for PSR and a *SEE* of 17 K for AMSR-E. Again, this is the result of the parameter settings: while the *VSM* results were relatively poor, results for T_S are better. For optimal results of the Wen algorithm a complicated and extensive optimalization should be performed for all changing frequencies, vegetation and soil conditions, taking *VSM* as well as T_S into account.

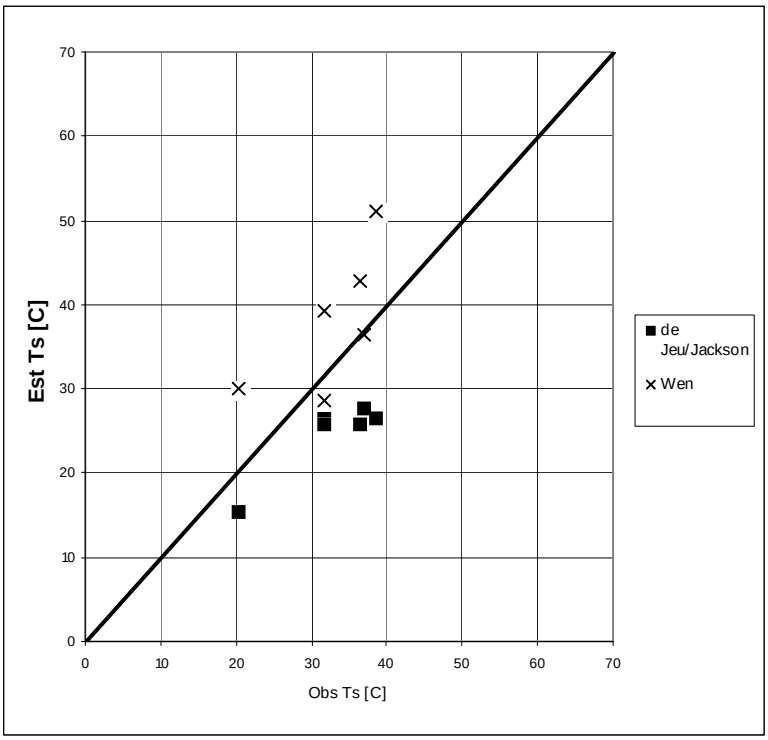


Figure 24: Surface temperature as computed by Equation 4 for the de Jeu and Jackson algorithms and the Wen algorithm for the SMEX02, using PSR C-band data (T_S for de Jeu and Jackson are the same as in Figure 18).

5 Discussion

In this study three soil moisture retrieval algorithms have been compared; one, the Jackson algorithm, is based on an inversion of the forward radiative transfer model and uses ancillary datasets to account for vegetation effects, while the remaining two, the de Jeu and Wen algorithm, solve the radiative transfer equation iteratively and do not need ancillary data to account for vegetation. After validation for the specific site, all algorithms yielded reasonable results, although the amount of validation that is required differed strongly between the algorithms.

Especially the Jackson and Wen algorithms contain many components that have to be carefully selected and validated, which makes it difficult to apply at a global scale. For example, one has to decide at which frequencies and polarization the brightness temperature is measured. In general, the lower the frequency, the larger the sensitivity to soil moisture content. In populated areas however, radio traffic interfering with the low-frequency radiation causes considerable problems, making it necessary to attempt to retrieve soil moisture using higher frequencies. A second choice that has to be made is the selection of the dielectric mixing model. In this study only two have been used, but there are more possibilities and each of them has its advantages and limitations: the Wang-Schmugge model for example is only valid for L and C-band retrievals, while the Hallikainen model is valid for other frequencies but consists of a rather simplistic empirical approach, where the effects of porosity and temperature are neglected.

Thirdly, vegetation relationships have to be set up. Often ancillary data is provided in the form of vegetation indices such as *NDVI* or *EVI*. There are several possible relations between the index and *VWC*: in this study, a set of linear equations is used (Equation 7), but there are also examples of quadratic functions (Jackson et al., 1999). Also the choice which index to use is important: in the Mongolian dataset *NDVI* (averaged over 10 years) and real-time *EVI* have been evaluated; even though *EVI* was only available at six days, retrievals using *EVI* yielded significantly better results. Even when *VWC* is directly presented as ancillary data, such as in the SMEX02 dataset, the problem of selecting a proper value for *b* remains, which has a very large influence in areas with a dense vegetation cover. Another parameter for which a value has to be selected, the *h* parameter, has a smaller influence because it is associated with the (rather large) incidence angle as is given by Equation 8. In the Wen algorithm less ancillary datasets are necessary, but there is a large amount of other parameters involved that have to be set, concerning the derivation of vegetation conditions from T_B-values (or assuming constant conditions as was done in this study) and accounting for the differences that occur because T_B-values of different frequencies are used. A factor that makes this algorithm more complicated is the fact that every parameter setting effects both *VSM* and T_S results, causing the necessity for a very extensive optimalization.

Apart from the problems mentioned above, there are also site related problems: in the SMEX02 area there is for example a dense vegetation cover with a *VWC* of about 2 kg m^{-2}, causing considerable problems in the retrieval. In Mongolia, there was the problem of the high amount of rocks in the soil, affecting dielectric properties of the soil. Thus, the datasets used here did not provide optimal conditions for soil moisture

retrieval, but they did provide very different conditions under which a retrieval algorithm should be able to retrieve when applied to a global scale.

6 Conclusions & Recommendations

In this concluding chapter, the questions that were formulated in Chapter 1 are answered and some recommendations for future research are made.

Is it possible to provide reliable estimations of soil moisture content using the algorithms proposed in this study?
When validated for a specific area and frequencies all three algorithms can provide reasonable results with standard errors of estimation of a bit less than or equal to 5 %. The amount of validation that is required, however, differs significantly per algorithm.

Which of the algorithms provides the most reliable results?
Although the de Jeu and Wen algorithms in this study yield reasonable results, the estimated *VSM* values fluctuate more in time, and do not react to changes in vegetation cover, resulting in higher overall errors compared to the Jackson algorithm. The only input for the de Jeu and Wen algorithms consists of satellite brightness temperature, while the Jackson algorithm makes use of vegetation information as well. Therefore the Jackson algorithm responds to vegetation changes as well, causing the rapid fluctuations occuring in the de Jeu algorithm to be smoothened, resulting in a better accordance with observed data. Especially in areas and time periods where vegetation cover changes significantly, such as in the SMEX02 dataset, ancillary data appears to be a necessity. Without that, observations and estimations can differ significantly. Besides, when comparing the results of different datasets and orbits for each algorithm, it appears that for the Jackson algorithm these differences are by far smallest. Especially in the SMEX02 dataset, there were large differences between ascending and descending orbits for de Jeu and Wen algorithms. This indicates a bigger robustness for the Jackson algorithm: when validated properly this algorithm is capable of providing equally reliable results in different conditions, while the de Jeu algorithm is hard to validate and the Wen algorithm requires a validation that is too complicated.

Which of the proposed algorithms is most suitable for application at a global scale?
It can be said that ancillary data makes it more difficult to apply an algorithm like the Jackson algorithm at a global scale: proper datasets with a resolution that is high enough are not always available, and the kind of dataset to be used can differ per area. In addition, the parameters that are often associated with ancillary data (e.g. *b*) can differ per area. Although the Wen algorithm does not use much ancillary data it contains a lot of area- and frequency dependent parameters which makes it difficult to apply globally as well. An algorithm like the de Jeu algorithm, which does not need much ancillary data and contains hardly any area- or frequency dependent parameters is much easier to apply globally.

Are the datasets that are used suitable for validation and evaluation of the three algorithms?
As was already mentioned, in the research area in Mongolia vegetation was very sparse but the soil contained many rocks, which affects the sensitivity of microwave radiation to soil moisture content. On the other hand, in the SMEX02 area a very dense vegetation was present, especially during the latter part of the retrieval period causing problems in the retrieval. In the Mongolian dataset the measuring devices

measured at depths that were deeper than the sensing depth of the microwave radiation, which is also causing an error between estimated and observed values. Concluding, the datasets used here did not provide optimal conditions, but they did provide very different conditions in terms of soil and vegetation. Since the goal is to apply the algorithms to a global scale and since optimal conditions are sparse the combination of these datasets provided a good test case for the algorithms.

Finally, some recommendations for future research are made. As is already mentioned many different soil moisture retrieval algorithms have been proposed, of which only three were compared in this study. In future research other algorithms can be taken into account as well. Since ancillary data can be very important for reliable retrievals, existing datasets containing soil texture, *NDVI* or *VWC* can be improved or newly created with a global coverage and a high resolution. Another aspect that can be and is constantly improved is the satellite sensor; at the moment the most promising passive microwave sensor is AMSR-E, but more sophisticated sensors are currently being developed, providing higher spatial resolution and lower frequency (e.g. L-band). It will be a great challenge to develop reliable, global scale retrieval algorithms for new satellites. Also, since the eventual goal is to obtain global estimations of soil moisture content, algorithms should be evaluated under a variety of conditions and in many different areas, because when an algorithm works for a dataset, that is not a guarantee it will work in other cases as well.

7 References

Bindlish, R. T. J. Jackson, E. Wood, H. Gao, P. Starks, D. Bosch and V. Lakshmi, 2003, "Soil moisture estimates from TRMM Microwave Imager observations over the Southern United States", Remote Sensing of Environment, 85:507-515.

Choudhurry, B. J., T. J. Schmugge, A. Chang and R. W. Newton, 1979, "Effect of surface roughness on the microwave emission for soils", Journal of Geophysical Research, 84:5699–5706.

Clevers, J., 'Vegetation indices and red edge index', 2001, reader Remote Sensing Optical, K075-222, Laboratory of Geo-Information Science and Remote Sensing, Wageneningen University.

Dane, J. H. and G. Clarke Topp, 2002, 'Methods of soil analysis, Part 4: Physical methods', Soil Science Society of America, Inc., Madison, Wisconsin.

De Jeu, R. A. M., 2003, "Retrieval of land Surface Parameters using Passive Microwave
Remote Sensing", Thesis Vrije Universiteit Amsterdam.

Ferraro, R. F., Grody, N. C., and Marks, G., 1994, "Effects of surface conditions on rain identification using DMSP-SSM/I", Remote Sensing Reviews, 11:195-209.

Griend, van de, A. A., M. Owe, J. de Ruiter and B.T. Gouweleeuw, 1996, "Measurement and Behaviour of Dual-Polarization Vegetation Optical Depth and Single Scattering Albedo at 1.4- and 5-GHz Microwave Frequencies", IEEE Transactions on Geoscience and Remote Sensing, 34(4):957-965.

Hallikainen, M., F. T. Ulaby, M. C. Dobson, and M. El-Rayes, 1985, "Microwave dielectric behavior of wet soil - Part I: Empirical models and experimental observations", IEEE Transactions on Geoscience and Remote Sensing, 23(1):25-34.

Hoekman, D. H., "Introduction to remote sensing physics and radar", 2000, Lecture notes K150-327, Wageningen University.

Jackson, T. J., 1993, "Measuring large scale surface soil moisture using passive microwave remote sensing", Hydrological Processes, 7:139-152.

Jackson, T. J., A. J. Gasiewski, A. Oldak, M. Klein, E. G. Njoku, A. Yevgrafov, S. Christiani and R. Bindlish, 2002, "Soil moisture retrieval using the C-band polarimetric scanning radiometer during the Southern Great Plains 1999 experiment", IEEE Transactions on Geoscience and Remote Sensing, 40(10): 2151–2161.

Jackson, T. J. and A. Y. Hsu, 2001, "Soil moisture and TRMM microwave imager relationships in the Southern Great Plains 1999 (SGP99) experiment", IEEE Transactions on Geoscience and Remote Sensing, 39: 1632-1642.

Jackson, T. J., A. Y. Hsu and P. E. O'Neill, 2002, "Surface soil moisture retrieval and mapping using high-frequency microwave satellite observations in the Southern Great Plains", Journal of Hydrometeorology, 3:688–699.

Jackson, T. J., K. G. Kostov, S. S. Saatchi, 1992, "Rock fraction effects on the interpretation of microwave emission from soils", IEEE Transactions on Geoscience and Remote Sensing, 30(3): 610-616.

Jackson, T. J., D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. D. Isham and M. Haken, 1999, 'Moisture Mapping at Regional Scales using Microwave Radiometry: The Southern Great Plains Hydrology Experiment", IEEE Transactions on Geoscience and Remote Sensing, 37(5): 2136-2151.

Jackson, T. J., T. J. Schmugge and J. R. Wang, 1982, "Passive microwave sensing of soil moisture under vegetation canopies", Water Resources Research, 18(4):1137-1142.

Kaihatsu, I., 2003, 'Match-up data for AMSR/Soil moisture', Hiroshima University.

Kawanishi, T., T. Sezai, Y. Ito, K. Imaoka, T. Takeshima, Y. Ishido, A. Shibata, M. Miura, H. Inahata and R. W. Spencer, 2003, "The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's Contribution to the EOS for Global Engergy and Water Cycle Studies", IEEE Transactions on Geoscience and Remote Sensing, 41(2):184-193.

Koike, T., E. G. Njoku, T. J. Jackson and S. Paloscia, 2000, "Soil moisture algorithm development and validation for the ADEOS-II/AMSR", IGARSS 2000, IEEE, 1253-1255.

Li, L., E. Njoku, E. Im, P. Chang and K. St.Germain, 2003, 'A preliminary survey of Radio-Frequency Interference over the U.S. in Aqua AMSR-E data', to be published in IEEE Transactions in Geoscience and Remote Sensing, http://www.ghcc.msfc.nasa.gov/AMSR/.

Lillesand, T. and Kiefer R., 1999, "Remote sensing and image interpretation", 4th edition, John Wiley and Sons, New York.

Njoku, E. G., T. J. Jackson, V. Lakshmi, T. K. Chan and S. V. Nghiem, 2003, "Soil moisture retrieval from AMSR-E", IEEE Transactions on Geoscience and Remote Sensing, 41(2):215-229.

Njoku, E. G., T. Koike, T. J. Jackson and S. Paloscia, 2000, "Retrieval of soil moisture from AMSR data", Microw. Radiomet. Remote Sens. Earth's Surf. Atmosphere, 525-533.

Njoku, E. G. and L. Li, 1999, 'Retrieval of land surface parameters using passive microwave measurements at 6-18 GHz', IEEE Transactions on Geoscience and Remote Sensing, 37(1):79-93.

Reynolds, C. A., T. J. Jackson and W. J. Rawls, 2000. "Estimating water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions", Water Resources Research, 36(12):3653-3662.

Soil Moisture Experiments in 2002 (SMEX02), 2002, Experiment Plan.

Soil Moisture Experiments in 2003 (SMEX03), 2003, Experiment Plan.

Terrestrial Biophysics and Remote Sensing Lab, University of Arizona, 2003, 'MODIS –EVI', http://tbrs.arizona.edu/project/MODIS/evi.php.

Ulaby, F. T., R. K. Moore and A. K. Fung, 1986, "Microwave Remote Sensing, Active and Passive, vol.III: From Theory to Applications", Artech House, Massachusetts.

Wang, J. and T. J. Schmugge, 1980, "An empirical model for the complex dielectric permittivity of soil as a function of water content", IEEE Transactions on Geoscience and Remote Sensing, GE 18:288-295.

Wen, J., Z. Su and Y. Ma, 2003, "Determination of land surface temperature and soil moisture from Tropical Rainfall Measuring Mission/Microwave Imager remote sensing data", Journal of Geophysical research, 108(D2)), 10.1029/2002JD002176.

Yang, D., F. T. Ulaby and M. C. Dobson, 2000. "Sensitivity to soil moisture by active and passive microwave sensors", IEEE Transactions on Geoscience and Remote Sensing, 38(1):105-114.

Appendix A Data quality check

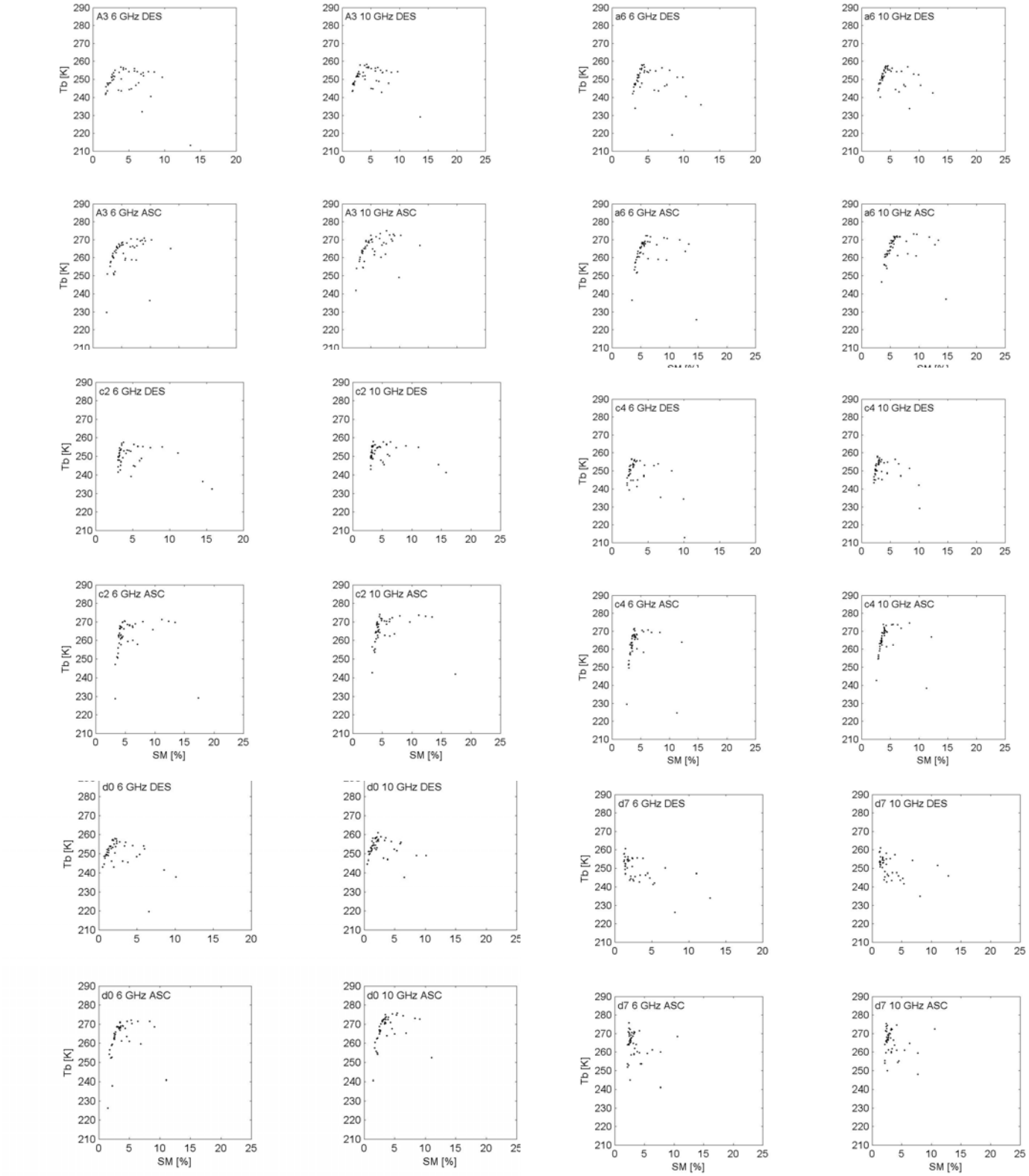


Figure A.1: Obs. VSM versus T_B 6 and 10 GHz, ascending and descending orbits for the first 6 obervation points.

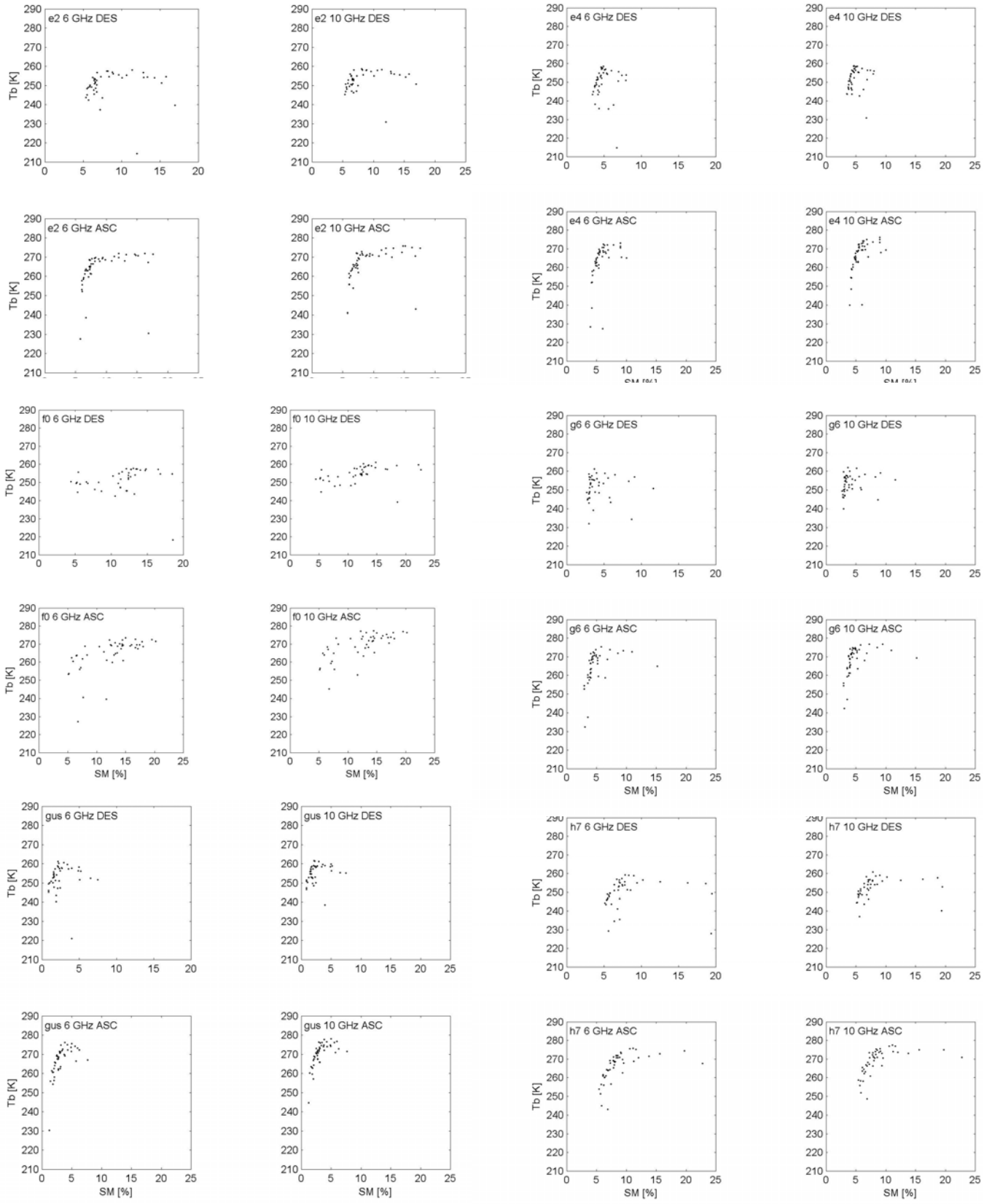


Figure A.2: Obs. VSM versus T_B 6 and 10 GHz, ascending and descending orbits for the latter 6 obervation points.

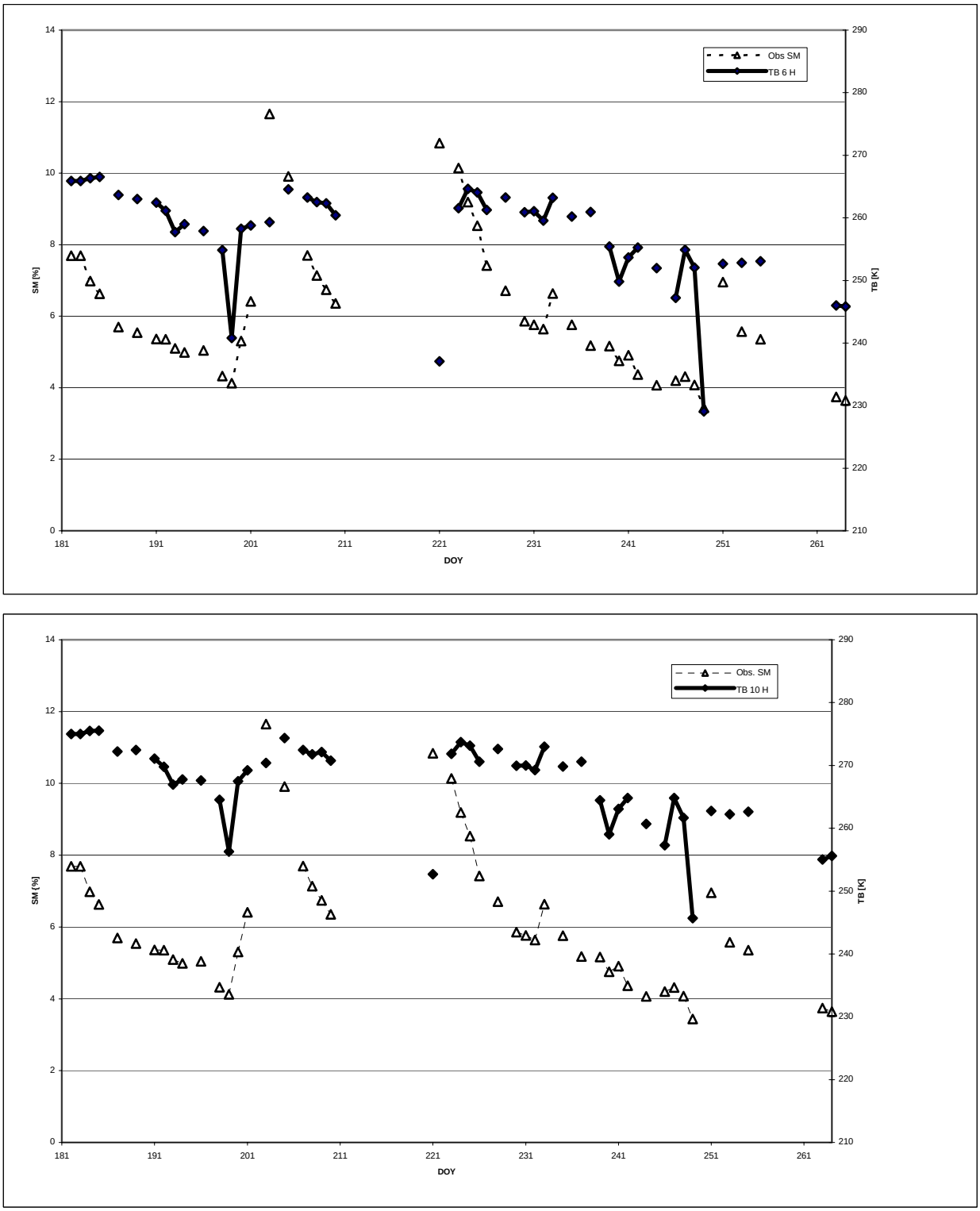
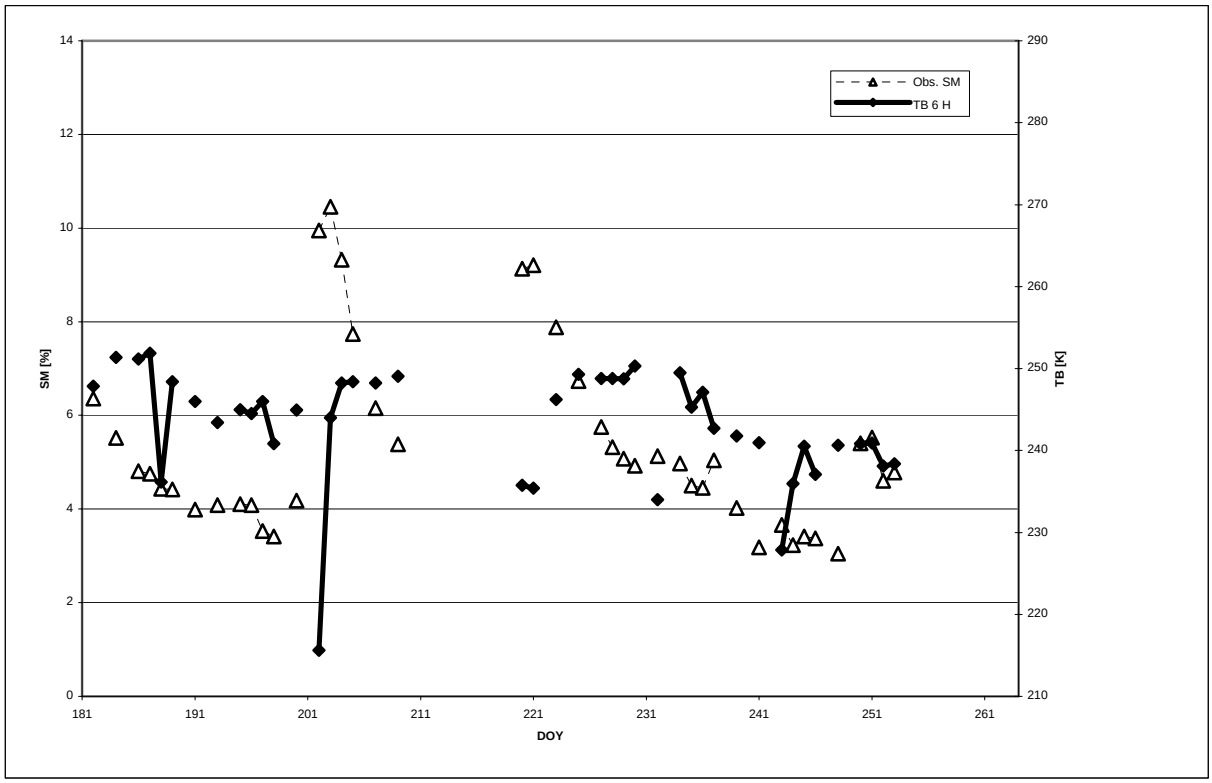


Figure A.3: Average T_B 6 H, TB 10 H and Observed Soil moisture over time, ascending orbits.



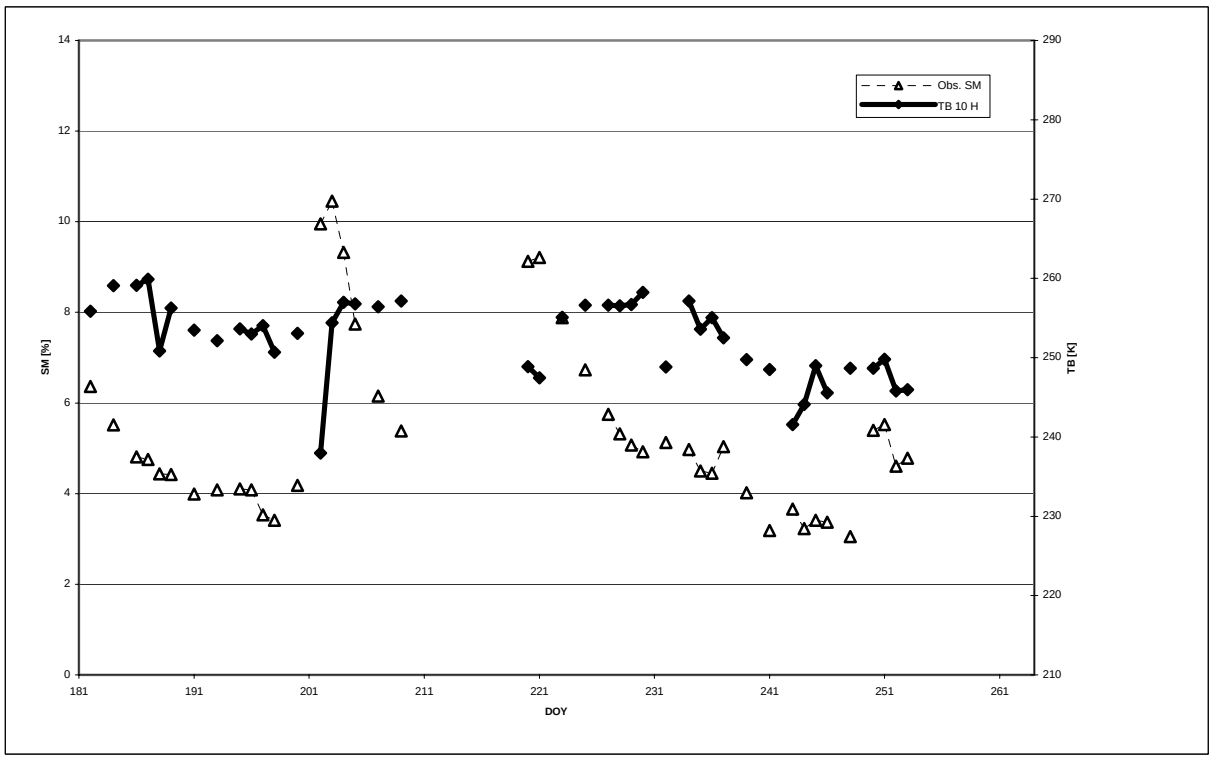


Figure A.4: Average TB 6 H, TB 10 H and Observed Soil moisture over time, descending orbits.

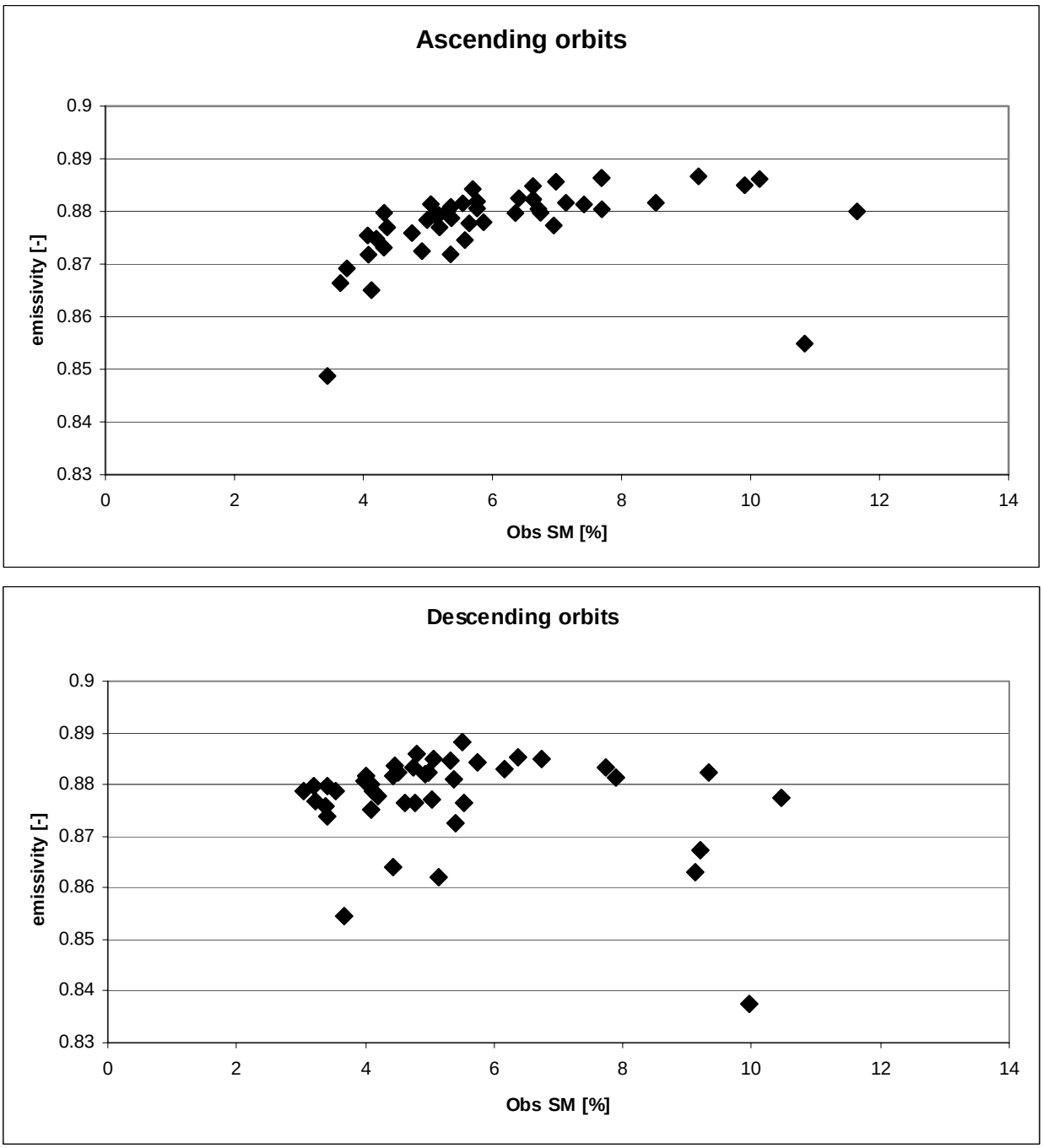


Figure A.5: Relationship between emissivity (ratio of brightness temperature TB and surface temperature TS), and observed soil moisture content for ascending and descending orbits.

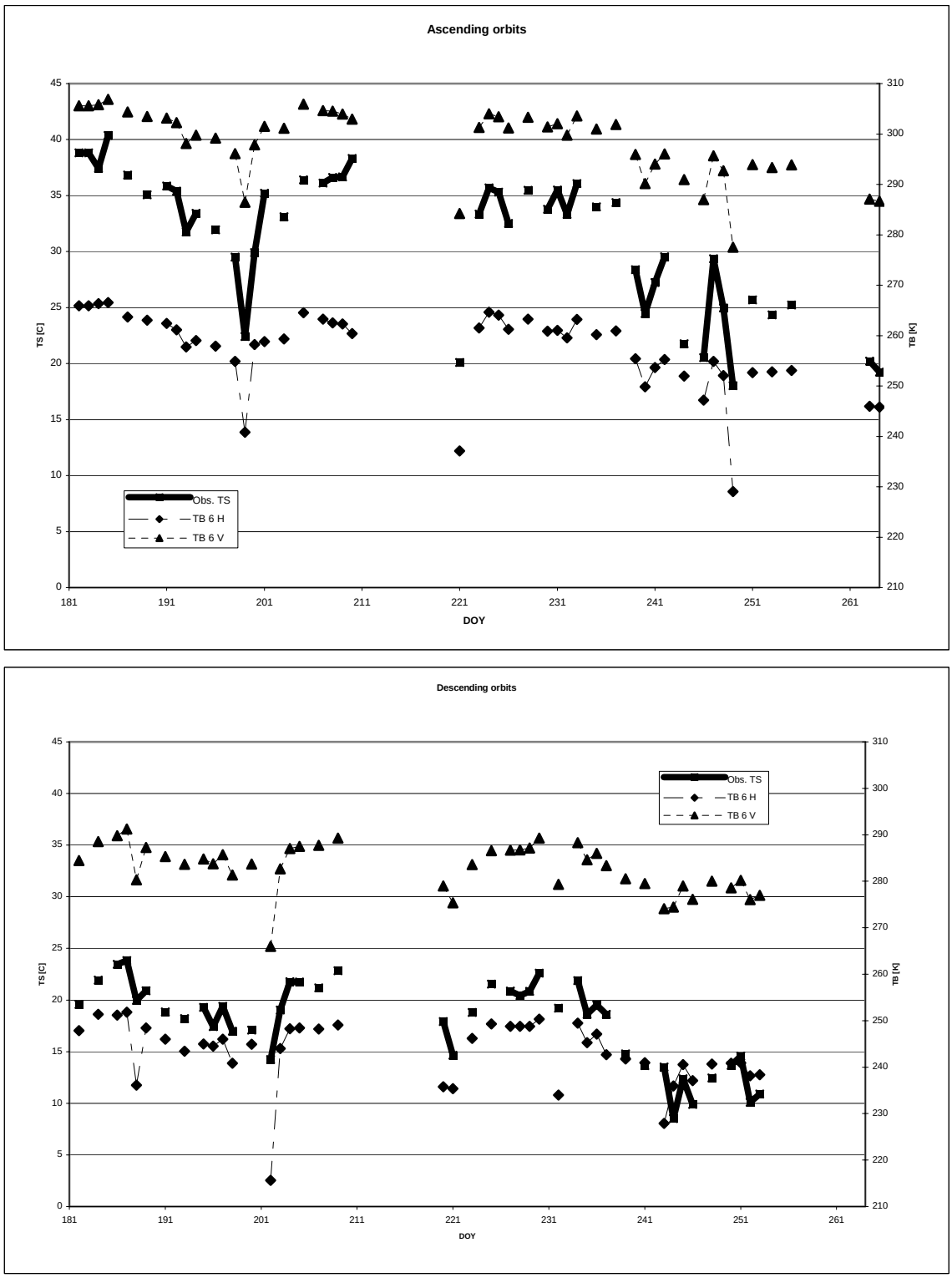


Figure A.6: Observed soil temperature at 3 cm depth and TB at 6 GHz H/V in time, for ascending and descending orbits.

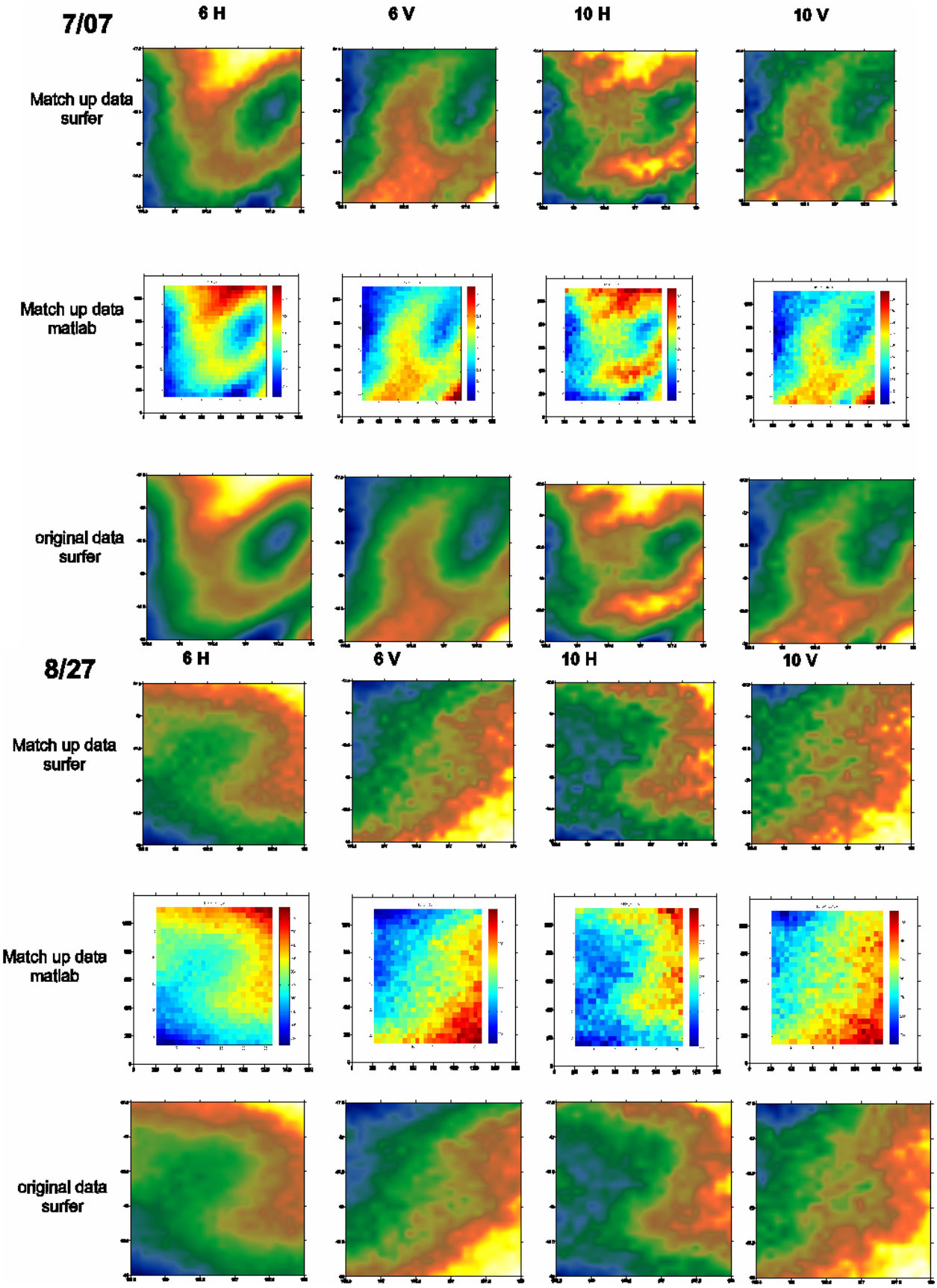
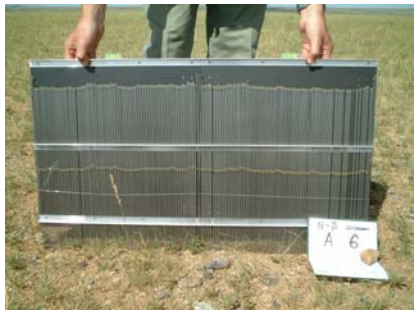


Figure A.7: Comparison of spatial patterns between match-up dataset and original AMSR footprint data.

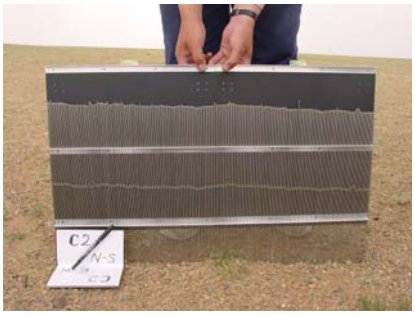
Appendix B Pictures of the research areas

A3

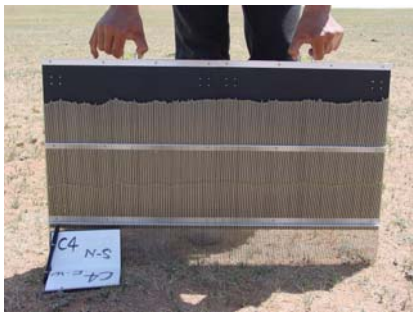
A6



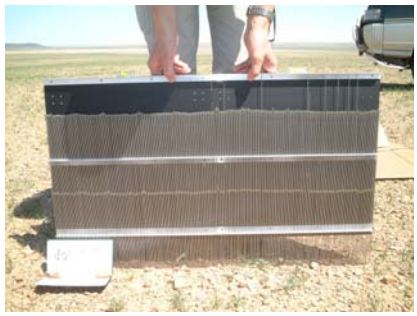
C2



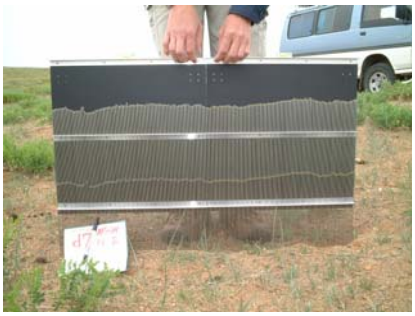
C4



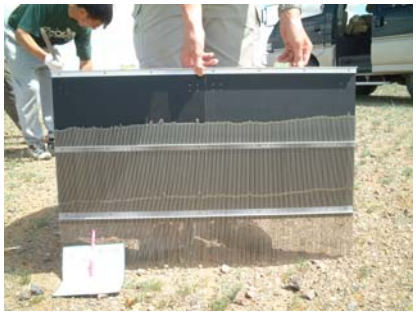
D0



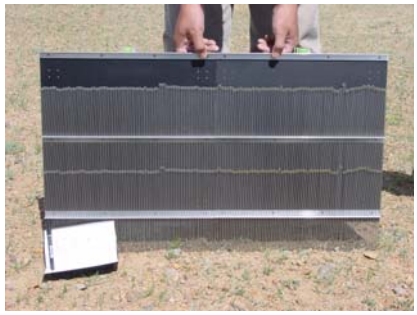
D7



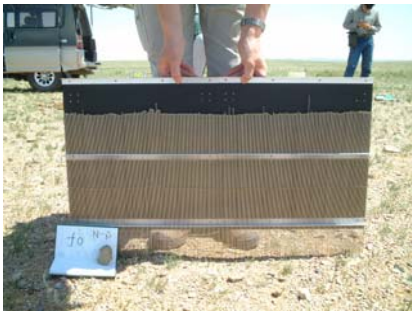
E2



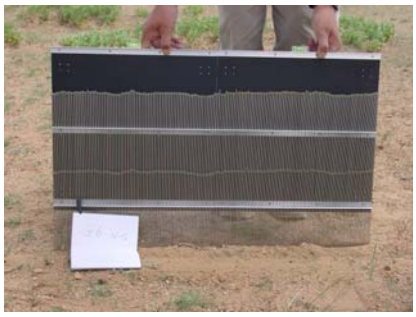
E4



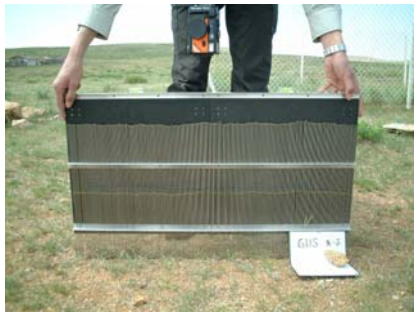
F0



G6



GUS



H7

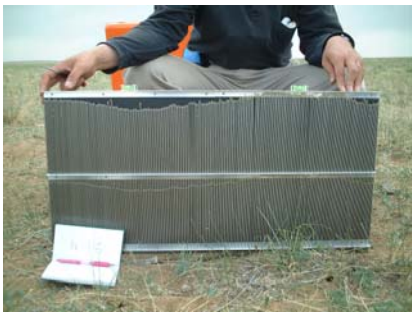


Figure B.1: Pictures near all twelve of the observation sites in the Mongolia study area.

Figure B.2: One of the 12 observation stations (top) and impression of the landscape (bottom) in the Mongolian study area.

Figure B.3: Landscape in Iowa (SMEX02) soy crop (upper) and corn crop (lower).

Appendix C Derivation of ancillary datasets for SMEX02

To begin, this analysis requires that TM bands needed to be atmospherically corrected prior to any analysis. This topic is covered in the NDVI/NDWI dataset discussion and will be assumed to be accurate for the purposes of VWC calculation.

Regional Calibration

Preliminary Regional VWC maps were generated in a similar manner to the Watershed, with some key exceptions. Three sets of scenes were available for the IA Region during the experiment, June 23, July 1, and July 17. The two scenes for each day were mosaiked together and regional maps of NDWI were generated. From these NDWI maps, VWC maps were generated using the same relationships that were developed in the WC region. From these three scenes, a linear interpolation was developed for the entire study period. To verify the accuracy of this preliminary interpolation method, the WC sampling sites
were revisited and statistics were recalculated.

Total:	Bias: -0.1883	rmse: 0.5381	R^2: 0.9150
Corn:	Bias: -0.2619	rmse: 0.6501	
Soy:	Bias: -0.0497	rmse: 0.1965	

The high Corn bias can be explained by the inability to interpolate during the later stages of the study period, when there was a precipitation event and rapid growth. This fact coupled with the tendency for NDWI to saturation at approximately 0.5 kg/m^2 results in the poor fit of this data.

A primary reason for the preliminary nature of this data is the presence of clouds on 7/17/02 which potentially influenced VWC readings for that area. Future datasets will provide a better remedy for this issue.

Procedures for Landcover Classification for SMEX02 area:

1) Collected Ground truth throughout the SMEX02 site on two separate trips (June and July)

2) Ground truth was converted to AOI's and 1/3 of each class were set aside for verification purposes

3) Landsat TM data for three dates: May 14th, July 1st, July 17th were collected and imported to ERDAS for Path 26/Row 31. This path/row covered most of SMEX02 . In addition, the southern portion of Path 26/Row 30 was added to complete the entire area.

4) The May 14th and July 17th images were referenced to the July 1st image.

5) The study area was split into two pieces due to clouds. The North and South areas were cloud free on all three dates, while the center portion (approximately at equal latitude with the City of Ames) was cloud covered on July 17th.

6) Bands 3,4,5,7 were extracted for each date.

7) For the North-South area, a 12 band image was created and for the Middle area an 8 band image was created.

8) Training classifications within AOI's used ISODATA reiterations. Means were initialized along diagonal axis and standard deviations were set to one. Maximum iterations were set to 24, convergence thresholds were set to .990

9) Signatures based on AOI's were combined into one large signature file for each area and used within supervised classification. The supervised classification used the non-parametric rule of parallelepiped and the parametric rule of Mahalanobis distance. In the process of optimizing the classification, other variations were used for the supervised classification but this was found to be the best.

10) The two areas (north-south and middle) were joined together. In the areas where the two areas overlapped the following conditions were applied. If the middle was alfalfa the merged was alfalfa. If the middle was corn and the north-south was not soybean, the merged version was corn. If the middle was corn and the north-south was soybean, the merged image was soybean. If the middle area was grass the merged was grass. If the middle area was trees, urban or water the merged version was whatever the north-south classification had been. If the middle was soybean and the north-south was grass, the merged was grass. Lastly if the middle was soybean and the north-south was not grass, the merged was soybean.

11) Accuracies were calculated using the AOI's that had not been used in the classification process.

The table is as follows:

Final merg	Ground								
Image	Alfalfa	Corn	Grass	Soybean	Trees	Urban	Water	Total	Accuracy
Alfalfa	66		3	20	1			90	73.33333
Corn	12	19291	209	379	17		10	19918	96.85209
Grass	50	97	472	93	104	4	46	866	54.50346
Soybean	136	66	63	14429	14	14	1	14723	98.00312
Trees	18	100	45	72	77		63	375	20.53333
Urban	3	2	7	20	1	2	0	35	5.714286
Water		7	5	1	6		56	75	74.66667
Total	285	19563	804	15014	220	20	176	Total	Accuracy
Accuracy	23.15789	98.60962	58.70647	96.10364	35	10	31.81818	36082	95.319

Corn and soybean accuracies are quite good. However due to the large sample size of the corn and soybeans accuracies in other land features are not as good. Small

misclassifications in the soybean or corn areas, can create larger inaccuracies in other classes.

Accuracies to the right of the chart are based on the image being the correct class for the the ground truth. For example, of 90 classified alfalfa pixels 66 of them are in the ground truth for alfalfa, for 73% accuracy. The accuracy at the bottom of the chart is the ground truth being compared to the image. For example of the 285 pixels that were deemed to be alfalfa on the ground only 66 of them were classified to be alfalfa, while 136 of them were classified as soybean.

12) Lastly, the classified image was registered to the road network provide by the Iowa Department of transportation. The road network was converted into an image where each road was 60 meters wide. This image was embedded onto the classified image to remove the speckle along the roadways and to improve the aesthetic quality of the image.

Appendix D Dielectric mixing models

The Wang-Schmugge model

In the Wang-Schmugge (Wang and Schmugge, 1980) model the dielectric constant ε of a soil-water mixture is described as:

$$\varepsilon = \theta \cdot \varepsilon_x + (P - \theta) \cdot \varepsilon_a + (1 - P) \cdot \varepsilon_r , \qquad \theta \le \theta_t$$
(D.1)

with

$$\varepsilon_x = \varepsilon_i + (\varepsilon_w - \varepsilon_i) \cdot \frac{\theta}{\theta_t} \cdot \gamma$$
(D.2)

and

$$\varepsilon = \theta_t \cdot \varepsilon_x + (\theta - \theta_t) \cdot \varepsilon_w + (P - \theta) \cdot \varepsilon_a + (1 - P) \cdot \varepsilon_r , \qquad \theta > \theta_t$$
(D.3)

with

$$\varepsilon_x = \varepsilon_i + (\varepsilon_w - \varepsilon_i) \cdot \gamma$$
(D.4)

Where θ is the volumetric water content [m^3 m^{-3}] of the soil, P the porosity of the dry soil (total volume occupied by pores per unit volume of soil), γ is an empirical parameter and θ_t is the transition moisture [m^3 m^{-3}]. ε_a, ε_w, and ε_i, in sequential order, are the dielectric constants of air, water, rock and ice. ε_x stands for the dielectric constant of initially absorbed water. The transition moisture is defined as the moisture content at which the free water phase begins to dominate the soil system and can be described as:

$$\theta_t = 0.49 \cdot WP + 0.165$$
(D.5)

Where WP is the wilting point of soil in [m^3 m^{-3}]. When WP is unknown but there is information about the particle size distribution of the soil, WP can be redetermined as:

$$WP = 0.06774 - 0.00064 \cdot SAND + 0.00478 \cdot CLAY$$
(D.6)

Where $SAND$ and $CLAY$ are the sand and clay contents in percent of dry weight of a soil. γ can be estimated by:

$$\gamma = 0.57 \cdot WP + 0.481$$
(D.7)

the complex dielectric constants for ice (ε_i), solid rock (ε_r) and air (ε_a) are *3.2+0.1i*, *5.5+0.2i* and *1+0i*, respectively. The dielectric constant for water (ε_w) is given by the Debye equation:

$$\varepsilon_w = \varepsilon_{w\infty} + \frac{\varepsilon_{w0} - \varepsilon_{w\infty}}{1 + (2 \cdot \pi \cdot t_w \cdot f)i} \quad \text{(D.8)}$$

Where $\varepsilon_{w\infty}$ is the high frequency limit of the dielectric constant of pure water (≈ 4.9), ε_{w0} is the static dielectric constant of pure water, t_w the relaxation time of pure water in seconds, and f is the electromagnetic frequency in Hz. Equation D.8 can be divided in a real and an imaginary part. The real part is defined by Ulaby et al. (1986):

$$\varepsilon'_w = \varepsilon_{w\infty} + \frac{\varepsilon_{w0} - \varepsilon_{w\infty}}{1 + (2 \cdot \pi \cdot t_w \cdot f)^2} \quad \text{(D.9)}$$

and the imaginary part as:

$$\varepsilon''_w = \frac{2 \cdot \pi \cdot t_w \cdot f \cdot (\varepsilon_{w0} - \varepsilon_{w\infty})}{1 + (2 \cdot \pi \cdot t_w \cdot f)^2} \quad \text{(D.10)}$$

The static dielectric constant of pure water is given as:

$$\varepsilon_{w0} = 88.045 - 0.4147 \cdot (T - 273.15) + 6.295 \cdot 10^{-4} \cdot (T - 273.15)^2 + 1.075 \cdot 10^{-5} \cdot (T - 273.15)^3 \quad \text{(D.11)}$$

where T is the effective temperature of the emitting layer in Kelvin. The relaxation time of pure water is:

$$2\pi t_w = 1.1109 \cdot 10^{-10} - 3.824 \cdot 10^{-12} \cdot (T - 273.15) + 6.938 \cdot 10^{-14} \cdot (T - 273.15)^2 - 5.096 \cdot 10^{-16} \cdot (T - 273.15)^3 \quad \text{(D.12)}$$

The Hallikainen model

Using an extensive database of measurements for different soil textures and frequencies, Hallikainen et al. (1985) developed empirical functions that take into account both factors.

$$\theta = \frac{-b + \sqrt{b^2 - 4(a - \varepsilon')c}}{2c} \quad \text{(D.13)}$$

Where θ is the soil moisture content and ε' is the real part of the dielectric constant of the soil-water mixture. a, b and c are described as:

$$a = a_0 + a_1(SAND) + a_2(CLAY) \quad \text{(D.14)}$$

$$b = b_0 + b_1(SAND) + b_2(CLAY) \quad \text{(D.15)}$$

$$c = c_0 + c_1(SAND) + c_2(CLAY) \quad \text{(D.16)}$$

where *SAND* and *CLAY* are the percentages of sand and clay in the soil. The coefficients *a0…c2* were derived by fitting to experimental data at specific frequencies and some values are summerized in Table D.1.

Table D.1: Coefficients for the real part of the dielectric constant from Hallikainen et al. (1985)

Frequency [GHz]	$\mathbf{a_0}$	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{b_0}$	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{c_0}$	$\mathbf{c_1}$	$\mathbf{c_2}$
1.4	*2.862*	*-0.012*	*0.001*	*3.803*	*0.462*	*-0.341*	*119.006*	*-0.5*	*0.633*
6	*1.993*	*0.002*	*0.015*	*38.06*	*-0.176*	*-0.633*	*10.72*	*1.256*	*1.522*
10	*2.502*	*-0.003*	*-0.003*	*10.101*	*0.221*	*-0.004*	*77.482*	*-0.061*	*-0.135*
18	*1.912*	*0.007*	*0.021*	*29.123*	*-0.19*	*-0.545*	*6.96*	*0.822*	*1.195*