ESA Climate Change Initiative Phase II - Soil Moisture

Algorithm Theoretical Baseline Document (ATBD) D2.1 Version 04.4

Active Soil Moisture Retrievals

12 November 2018

Prepared by

Earth Observation Data Centre for Water Resources Monitoring (EODC) GmbH

in cooperation with

TU Wien, GeoVille, ETH Zürich, TRANSMISSIVITY, AWST, FMI, UCC and NILU

This document forms part 2 (of 4) of deliverable D2.1 Algorithm Theoretical Baseline Document (ATBD) and was compiled for the ESA Climate Change Initiative Phase 2 Soil Moisture Project (ESA Contract No.4000112226/14/I-NB). For more information on the CCI programme of the European Space Agency (ESA) see http://www.esa-cci.org/.

Number of pages: 63

Authors:	D. Chung, W. Dorigo, C. Reimer, S. Hahn. T. Melzer. C. Paulik., M. Vreugdenhil, W. Wagner, R. Kidd		
Circulation:	Not for public release Project Internal ONLY until review completed by Science Lead		
Release	Date	Details	Editor
0.1	02/12/2015	Updated to match ESA CCI SM v02.2 release	D. Chung
0.2	04/12/2015	Editing + typesetting	W Dorigo
0.3	30/11/2016	Update according to ESA CCI SM v03.2 release	C. Reimer
3.2	01/12/2016	Editing + typesetting	W Dorigo
3.2	24/02/2017	Minor update	D. Chung
3.3	10/11/2017	Update to product version 03.3	T. Scanlon
4.2	15/01/2018	Update to product version 04.2	T. Scanlon
4.3	17/04/2018	Update to product version 04.3 (not publically released)	T. Scanlon
4.4	12/11/2018	Update to product version 04.4 (data extension)	P. Buttinger T. Scanlon

For any clarifications please contact Wouter Dorigo (wouter.dorigo@geo.tuwien.ac.at).

Project Partners

Prime Contractor **Scientific Lead**	**EODC**, Earth Observation Data Centre for Water Resources Monitoring (Austria)
Project Management	**GeoVille**, GeoVille Information Systems GmbH (Austria)
System Engineering Partners	**AWST**, Angewandte Wissenschaft Software und Technologie GmbH (Austria)
Earth Observation Partners	**TU Wien**, Vienna University of Technology (Austria) **TMS**, Transmissivity, (The Netherlands) **FMI**, Finnish Meteorological Institute, (Finland) **UCC**, University College Cork, (Ireland)
Climate Research Partners	**ETH**, Institute for Atmospheric and Climate Science, (Switzerland) **NILU**,Norsk institutt for luftforskning (Norway)

Table of Content

List of figures

List of Tables

Definitions, acronyms and abbreviations

AMI	Active Microwave Instrument
AMSR-E	Advanced Microwave Scanning Radiometer-Earth Observing System
AMSU	Advanced Microwave Sounding Unit
ASAR	Advanced Synthetic Aperture Radar
ASCAT	Advanced Scatterometer (Metop)
CCI	Climate Change Initiative
CEOP	Coordinated Energy and Water Cycle Observations Project
CMORPH	Morphing Method of the Climate Prediction Centre
CPC	Climate Prediction Centre
DARD	Data Access Requirement Document
DMSP	Defence Meteorological Satellite Program
DTED	Digital Terrain Elevation Model
EASE	Equal-Area Scalable Earth
ECV	Essential Climate Variable
ENVISAT	Environmental Satellite
EO	Earth Observation
ERA-40	ECMWF ReAnalysis 40 data set
ERS	European Remote Sensing Satellite (ESA)
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FTP	File Transfer Protocol
GIMMS	Global Inventory Modeling and Mapping Studies
GLDAS	Global Land Data Assimilation System
GLWD	Global Lakes and Wetlands Database (GSPC/University of Kassel)
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GRACE	Gravity Recovery And Climate Experiment
GSWP	Global Soil Wetness Project
ISMN	International Soil Moisture Network
ITRDB	International Tree-Ring Data Bank
JAXA	(Japan Aerospace Exploration Agency)
JPL	Jet Propulsion Laboratory (NASA)
METOP	Meteorological Operational Satellite (EUMETSAT)
NASA	National Aeronautics and Space Administration
NIMA	National Imagery and Mapping Agency

NOAA   National Oceanic and Atmospheric Administration

NSIDC   National Snow and Ice Data Center (radlab)

NWS   National Weather Service (NOAA)

SAR   Synthetic Aperture Radar

SCAT   Scatterometer

SMAP   Soil Moisture Active and Passive mission

SMMR   Scanning Multichannel Microwave Radiometer

SMOS   Soil Moisture and Ocean Salinity (ESA)

SOW   Statement of Work

SSM   Surface Soil Moisture

SSM/I   Special Sensor Microwave Imager

TDR   Time Domain Reflectometry

TMI   TRMM Microwave Imager

TRMM   Tropical Rainfall Measuring Mission

TWS   Terrestrial Water Storage

USGS   United States Geological Survey

VIC   Variable Infiltration Capacity

VOD   Vegetation Optical Depth

WACMOS   Water Cycle Multimission Observation Strategy

WindSat   WindSat Radiometer

List of symbols

θ	Incidence angle (degree), generic
$\theta_{i,b}, \theta_{j,b}$	Observed incidence angle of beam $b \in \{f, m, a\}$ (fore-, mid-, aft-beam) of i-th record in the time series of the current GPI or of j-th record in the time series of the current calibration target
$\varphi, \varphi_{i,b}$	Azimuth angle (degree), generic and observed
$\sigma^0, \sigma^0_{i,b}$	Radar cross-section, backscattering coefficient $\left(\frac{m^2}{m^2}\ \text{or}\ dB\right)$, generic and observed
t, t_i	Time, generic and observed
$d = doy(t), d_i$	Day of year, $d \in \mathbb{N}, 1 \le d \le 366$, as function of t (t_i)
$mon(t)$	Month including the year of time t
$\sigma^0(\theta, d)$	Backscatter, modelled as function of incidence angle, with the model depending on the day of year d (i.e., d indexes one instance of the model class)
$\sigma^0_{i,b}(\theta_i)$	Observed backscatter, represented in terms of the model
$\sigma'(\theta, d)$	First derivative of $\sigma^0(\theta, d)$
$\sigma'(\theta_{ref}, d)$	First derivative ('slope') at reference angle, parameter array
$\sigma''(\theta, d)$	Second derivative of $\sigma^0(\theta, d)$
$\sigma''(\theta_{ref}, d)$	Second derivative ('curvature') at reference angle, parameter array
$\overline{\sigma^0_i}(\theta_{ref})$	Normalised backscatter at reference angle, averaged over the beams, of the i-th record in the time series
θ_{dry}	Dry crossover angle
$\sigma^{dry}(\theta_{ref}, d)$	Dry reference at reference angle, parameter array
θ_{wet}	Wet crossover angle
$\sigma^{wet}(\theta_{ref}, d)$	Wet reference at reference angle, parameter array
$\sigma^0{}_{ref}(\theta)$	Backscatter calibration reference, modelled as function of incidence angle
$C_{intra}(\theta)$	Intra-Calibration coefficient, modelled as function of incidence angle
$\sigma^0_{intra,i,b}$	Intra-calibrated backscatter coefficient, observed

1 Executive Summary

The Algorithm Theoretical Baseline Document (ATBD) provides a detailed description of the algorithms that are used within the CCI Soil Moisture ECV production system. The ECV production system has initially been developed within CCI Phase 1 and is continuously being updated within phase 2 to reflect the current state of the science driving the system. The ATBD is, by its nature, rather in-depth, and in order to facilitate frequent updates, and to provide a more manageable document to the reader the ATBD is provided as four distinct documents. These documents consist of an ATBD for the active retrieval, an ATBD for the passive retrieval and an ATBD for the merging process. An overriding document (part 1) provides an executive summary sets the ATBD documents within framework for the CCI project and the ECV production system.

Section 5 of this document outlines the processing steps involved in the active product retrieval. The algorithms that underlie these processing steps are presented in Section 6 and form the basis of the change detection model developed at the Vienna University of Technology (TU Wien). Section 7 identifies the shortcomings of the active product retrieval, and the scientific advances that are currently being investigated to address these are described in Section 8. These include the latest results from research into level-1 inter-calibration biases between scatterometer missions, improved vegetation modelling, and the improved resampling of scatterometer measurements to the Discrete Global Grid (DGG).

2 Change log

2.1 Current version 04.4

This document forms deliverable 2.1 of CCI Phase 2 and provides an update for the ESA CCI SM 04.4 product released on 12th November 2018. At version 04.4, there are no changes to the algorithm for generating the active L2 data product.

2.2 Pre v04.4

The dataset and corresponding ATBD versions are summarised in the executive summary of the ATBD. Further information can be found in the changelog provided with the data and the relevant documentation.

3   Introduction

3.1   Purpose of the Document

The Algorithm Theoretical Baseline Document (ATBD) is intended to provide a detailed description of the scientific background and theoretical justification for the algorithms used to produce the ECV soil moisture data sets. Furthermore, it describes the scientific advances and algorithmic improvements which are made within the CCI project.

3.2   Targeted Audience

The primary audience for this document are:

1. Remote sensing experts interested in the retrieval and error characterisation of soil moisture from active microwave data sets.
2. Users of the remotely sensed soil moisture data sets who want to obtain a more in-depth understanding of the algorithms and sources of error.

4 Reference Documents

The following references are of relevance to this document. Within the document, for the sake of clarity it has been sometimes necessary to provide sections of quoted texts taken from referenced documents, rather than just providing a reference to the document. In these cases, texts are "*presented in quotes as italic text*".

[RD-1]	ESA Climate Change Initiative Phase 1, Statement of Work for Soil Moisture and Ice Sheets, European Space Agency, EOEP-STRI-EOPS-SW-11-0001.
[RD-2]	Technical Proposal (Part 3) in response to ESA Climate Change Initiative Phase 1 ESRIN/AO/1-6782/11/I-NB, Vienna University of Technology.
[RD-3]	W. Dorigo, R. Kidd, R. De Jeu, S. Seneviratne, H. Mittelbach, J. Pulliainen, W.A. Lahoz, N. Dwyer, B. Barrett, Eva Haas, W. Wagner.ESA CCI Soil Moisture Data Access Requirements Document, v1.2
[RD-4]	Wagner, W., W. Dorigo, R. De Jeu, D. Fernandez, J. Benveniste, E. Haas, M. Ertl (2012) Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture, Proceedings of the ISPRS Congress 2012, Melbourne, Australia, August 25-September 1, 2012.
[RD-5]	Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences, 15, 425-436, doi:10.5194/hess-15-425-2011
[RD-6]	Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M., Wagner, W., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280-297, doi: 10.1016/j.rse.2012.03.014.
[RD-7]	D. Chung, R.A.M de Jeu, W. Dorigo, S. Hahn. T. Melzer. R.M. Parinussa, C. Paulik. C. Reimer, M. Vreugdenhil, W. Wagner (2012). ESA CCI Soil Moisture. Algorithm Theoretical Baseline Document (ATBD), Version 0, 30th April 2012.
[RD-8]	D. Chung, R.A.M de Jeu, W. Dorigo, S. Hahn. T. Melzer. R.M. Parinussa, C. Paulik. C. Reimer, M. Vreugdenhil, W. Wagner (2013), Algorithm Theoretical Baseline Document (ATBD), Version 1.0, release 1.0, 14th February 2013.
[RD-9]	D. Chung, W. Dorigo, S. Hahn. T. Melzer. C. Paulik. C. Reimer, M. Vreugdenhil, W. Wagner, R. Kidd (2013). Soil Moisture Retrieval from Active Microwave

	Sensors: Algorithm Theoretical Baseline Document (ATBD) Version 2.0, Release 0.3, 15th October 2014.
[RD-10]	D. Chung, R.A.M. de Jeu, W. Dorigo, S. Hahn. T. Melzer. R.M. Parinussa, C. Paulik. C. Reimer, M. Vreugdenhil, W. Wagner, R. Kidd (2013). ESA CCI Soil Moisture – Soil Moisture Retrieval from Passive Microwave Sensors: Algorithm Theoretical Baseline Document (ATBD), Version 2.0, Release 0.4, 15th October 2014.
[RD-11]	D. Chung, W. Dorigo, S. Hahn. T. Melzer. C. Paulik. C. Reimer, M. Vreugdenhil, W. Wagner, R. Kidd (2013). ESA CCI Soil Moisture – ECV production, Fusion of Soil Moisture Products: Algorithm Theoretical Baseline Document (ATBD) Version 2.0 Release 0.5, 15th October 2014.
[RD-12]	R. Kidd, D. Chung, W. Dorigo, R. De Jeu (2013). ESA CCI Soil Moisture – Detailed Processing Model (DPM), Version 1.2, 26th November 2013.
[RD-13]	D. Chung R. Kidd, W. Dorigo (2013) ESA CCI Soil Moisture – System Prototype Description (SPD), Version 1.0, 4th July 2013.
[RD-14]	R. Kidd, D. Chung, W. Dorigo, R. De Jeu (2012). Input/ Output Data Definition Document (IODD), Version 1.0, 21st December 2012.

5   Methodological description on the retrieval of soil moisture from active microwave sensors

5.1   Principle of the products

The concept of the Level 2 surface soil moisture retrieval model developed at the Vienna University of Technology (TU Wien) for use with C-band scatterometers is a physically motivated change detection method. The first realisation of the concept was based on ERS 1/2 wind scatterometer (AMI-WS) data sets (Wagner et al. 1999a; Wagner et al. 1999b; Wagner et al. 1999c) and later the approach was successfully transferred to the Advanced Scatterometer (ASCAT) data on board the METOP-A satellite (Bartalis et al. 2007; Naeimi et al. 2008; Naeimi et al. 2009). The soil moisture retrieval algorithm is implemented within a software package called Soil Water Retrieval Package (WARP).

The TU Wien change detection algorithm is, from a mathematical point of view, less complex than a radiative transfer model and can be inverted analytically. Therefore soil moisture can be estimated directly from the scatterometer measurements without the need for iterative adjustment process. Because of this it is also quite straight forward to perform an error propagation to estimate the retrieval error for each land surface pixel (Naeimi et al. 2009). A disadvantage of the change detection model is that it is a lumped representation of the measurement process. Therefore, the different contributions to the observed total backscatter from the soil, vegetation, and soil-vegetation-interaction effects cannot be separated as would be the case for a radiative transfer modelling approach. It also means that it is necessary to calibrate its model parameters using long backscatter time series to implicitly account for variations in land cover, surface roughness, and many other effects. The basic assumptions of the TU Wien change detection model are:

1. The relationship between the backscattering coefficient σ^0 expressed in decibels (dB) and the surface soil moisture content is linear.
2. The backscattering coefficient σ^0 depends strongly on the incidence angle θ. The relationship σ^0 - θ is characteristic of the roughness conditions and the land cover, but is not affected by changes in the soil moisture content.
3. At the spatial scale of the scatterometer measurements roughness and land cover are stable in time.
4. When vegetation grows, backscatter may decrease or increase, depending on whether the attenuation of the soil contribution is more important than the enhanced contribution from the vegetation canopy, or vice versa. Because the relative magnitude of these effects depends upon the incidence angle, the curve σ^0 - θ changes with vegetation phenology over the year. This effect can be exploited to correct for the impact of vegetation phenology in the soil moisture retrieval by

assuming that there are distinct incidence angles θ_{dry} and θ_{wet}, where the backscattering coefficient σ^0 is stable despite seasonal changes in above ground vegetation biomass for dry and wet conditions.

5. Vegetation phenology influences σ^0 on a seasonal scale. Local short-term fluctuations are suppressed at the scale of the scatterometer measurements.

Overall, the results obtained in experimental validation studies, for both ERS-1/2 AMI-WS and METOP-A ASCAT, have shown that these assumptions are in general appropriate. Also, they have received support from theoretical studies. For example, the linearity assumption (point 1 above) has been held to be in contradiction to the Integral Equation Model (IEM) that suggests a non-linear relationship. But, as recently demonstrated by Zribi (personal communication), this contradiction disappears when IEM is coupled with an air-soil transition model as first proposed for the SMOS retrieval (Schneeberger et al. 2004). With respect to seasonal vegetation effects, a recent study by Crow et al. (2010) shows that the change detection model is better able to describe the soil moisture retrieval skill over a larger range of incidence angles than the widely used Cloud Model (Attema and Ulaby 1978) in combination with the IEM (Fung 1994; Hsieh et al. 1997). But of course, there are instances where one or more of these assumptions break down. Currently the biggest problem appears to be that under extremely dry conditions, as might be found in deserts or semi-arid environments during the dry season, backscatter decreases when the soil becomes slightly wet.

5.2 Overview of processing steps

In the software package WARP the TU Wien change detection model is applied to ERS 1/2 AMI-WS and METOP-A ASCAT measurements via a sequence of processing steps (see Figure 1 for an overview):

1. **Resampling of data:** The scatterometer measurements in orbit geometry are resampled to a fixed Discrete Global Grid (DGG), called WARP 5 grid.
2. **Sensor intra-calibration:** Resampled backscatter values are calibrated to a defined calibration reference, in order to correct for temporal emerging variations of the σ^0 backscatter coefficient caused by instrument related anomalies.
3. **Azimuthal normalisation:** Backscatter values are normalised in terms of their acquisition azimuth angle, based on look-up tables with long-term mean values.
4. **Estimate noise of σ^0:** Estimate the standard deviation (ESD) of σ^0 due to instrument noise, speckle and residual azimuthal effects based on the measurements of the fore- and aft antennas.
5. **Model incidence angle dependency:** Determine the mean annual cycle of the incidence angle behaviour of σ^0 by making use of the fact that the scatterometer

provides instantaneous measurements at two different incidence angles. The incidence angle dependency is described by a second order polynomial determined by the slope and the curvature. The slope and the curvature show a distinct annual cycle, determined by vegetation growth and decay. Slope and curvature parameters are determined by fitting a first degree polynomial to each group of local slope values. The results are the first and second derivatives of backscatter at 40° for each day of year. The final slope and curvature values are the result of averaging these derivative values over several periods with different duration (14-84 days). Corresponding noise values are also calculated.

6. **Normalisation of backscatter measurements:** Extrapolate all σ^0 taken over the entire incidence angle range to a reference angle of $\theta_{ref} = 40°$ and calculate the average $\sigma^0(\theta_{ref})$ based on the backscatter triplet.
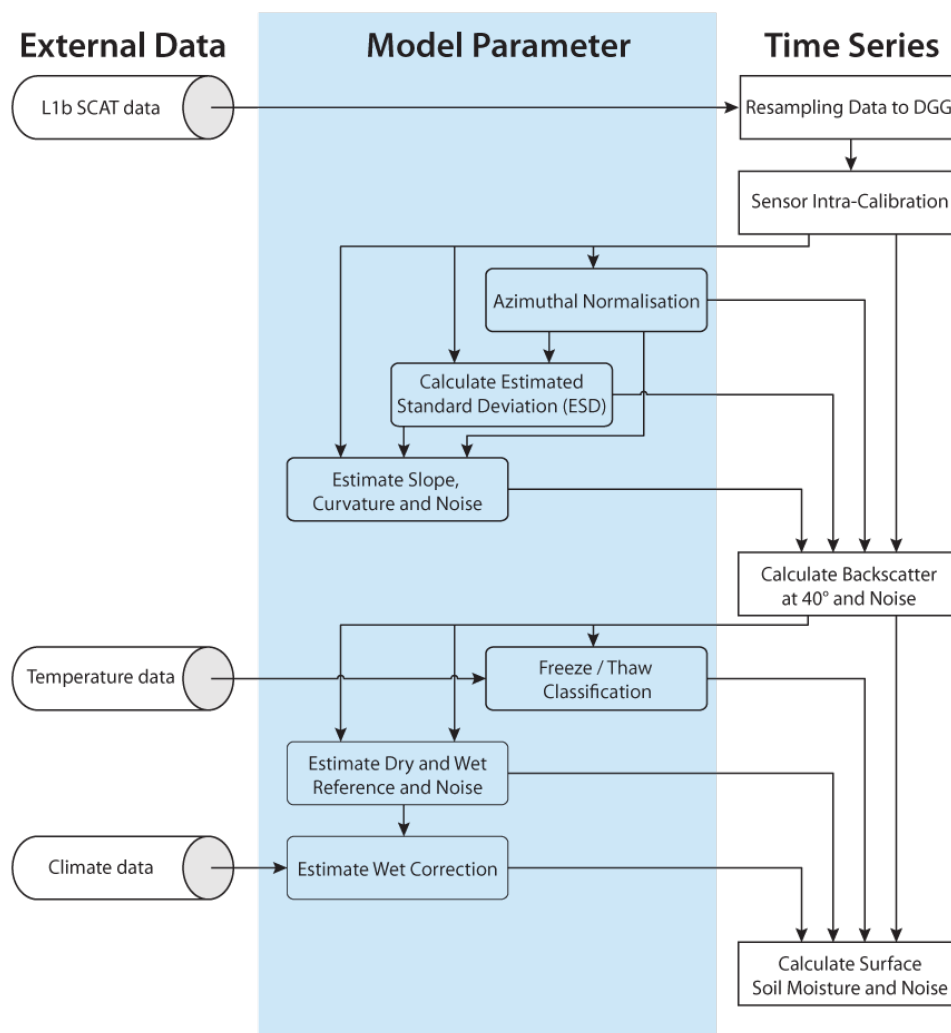
Figure 1: Overview of the processing steps in WARP 5.6

7. **Estimate noise of $\sigma^0(\theta_{ref})$:** Based on the rules of error propagation the estimated standard deviation of $\sigma^0(\theta_{ref})$ is calculated.

8. **Detect frost and snow conditions:** Use decision tree trained by historical temperature data to detect freeze/thawing events

9. **Determine dry and wet references:** After σ^0 has been normalised with respect to the incidence angle, vegetation phenology effects and $\sigma^0(\theta_{ref})$ outliers have been removed, dry and wet soil backscatter reference curves, $\sigma^{dry}(\theta_{ref}, d)$ and $\sigma^{wet}(\theta_{ref}, d)$ are determined. These maximum and minimum $\sigma^0(\theta_{ref})$ are determined by statistical methods of noise analysis.

10. **Wet reference correction:** In dry climates the wet reference estimation can be biased given that there may never be enough rainfall to thoroughly wet the soil surface layer (Wagner and Scipal 2000). To correct biased $\sigma^{wet}(\theta_{ref}, d)$ in such dry climates, Koeppen climate classification data (Kottek et al. 2006) is used in conjunction with the sensitivity to soil moisture (defined in turn as the difference between the dry and wet parameters derived in the previous step).

11. **Calculate surface soil moisture:** Calculate the surface soil moisture by comparing $\sigma^0(\theta_{ref})$ to the seasonally varying dry and wet reference values.

12. **Estimate retrieval error of surface soil moisture:** Calculate the estimated standard deviation of the surface soil moisture by rules of error propagation.

6 Description of Algorithms

This section presents the algorithms that underlie the processing steps presented in section 5.2, and forms the core of this algorithm theoretical baseline document (ATBD).

6.1 Resampling

The task of resampling is to interpolate L1b Scatterometer measurements, given in the orbit grid, to a fixed Earth grid. For this purpose a Discrete Global Grid (DGG) has been developed by TU Wien and is called WARP 5 grid. The WARP 5 grid contains 3264391 grid points with an equal spacing of 12.5 km in longitude and latitude. Each of the grid points is identified by a unique grid point index (GPI). The result of the resampling is a time series of interpolated measurements at each GPI over land (Figure 2).
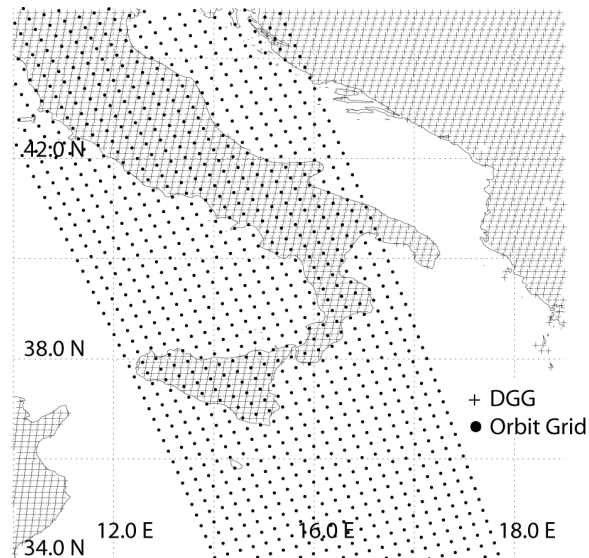
Figure 2: Orbit grid (dots) and WARP 5 grid (crosses) over Italy.

The geometry of the AMI-WS and ASCAT instruments is described in Figure 3 where the three satellite beams are indicated as fore, mid and aft beam. For each point in the orbit grid, all GPIs within an 18 km radius are determined by a nearest neighbour search, from which the interpolated values for the backscatter sigma naught σ^0 for each of the three beams (and other attributes such as incidence angle) are obtained as weighted average, with weighting coefficients computed according to the Hamming window function:

$$w(x) = 0.54 + 0.46 \cos\left(2\pi \frac{\delta x}{r}\right)$$

Eqn. 6-1

whereby δx denotes the distance between the actual GPI and the orbit grid point, and r the diameter of the search radius. We chose the Hamming window function for interpolation, because it is also used in the creation of the L1b product. Also two other window functions

are currently supported, namely Lanczos and Inverse Distance Weighting, but the relative merits of the respective functions have yet to be evaluated.

The result of the resampling step is, for each land GPI, a time series ts_{gpi}, containing N_{gpi} records

$$ts_{gpi}[i] = \sigma^0_{i,b}, \theta_{i,b}, \varphi_{i,b}, t_i, \qquad 1 \leq i \leq N_{gpi} \qquad \text{Eqn. 6-2}$$

each consisting of a time stamp t_i and measurement triples for backscatter $\sigma^0_{i,b}$, incidence angle $\theta_{i,b}$ and azimuth angle $\phi_{i,b}$. The subscript $b \in \{f, m, a\}$ distinguishes between the fore, mid-, and aft-beam. Note that in the processing chain described below, the time-series are processed for each GPI separately.

6.2   Sensor intra-calibration

During the mission lifetime of a scatterometer in space, numerous disturbances may influence the overall sensor performance and accordingly affect the accuracy of the normalised radar cross section σ^0 if disregarded. Space and satellite agencies, such as ESA and EUMETSAT, are routinely monitoring the scatterometer performance in order to correct for such sensor-related performance variations. A radiometric calibration methodology for European C-band scatterometer missions was developed at TU Wien to ensure consistent backscatter observations of these scatterometer missions (Reimer 2014). Sensor intra-calibration aims to support already established calibration efforts, undertaken by the operating Space Agency, with the objective to monitor and correct for residual scatterometer performance anomalies. The radiometric calibration is performed by utilizing natural calibration targets on the Earth's surface, presumed to exhibit a temporally stable, spatially homogeneous, and azimuthally isotropic backscatter response over an extended area. With reference to these backscatter properties, a backscatter calibration model, (Eqn. 5-3), has been introduced for sensor intra-calibration. The calibration model was adopted from Long and Skouson (1996) with respect to the measurement geometry of ERS AMI-WS. Backscatter coefficients $\sigma^0(t, \theta, \phi_b)$ observed for a calibration target T are composed of the backscatter coefficient $\sigma^0_T(\theta)$ of the calibration target, the intra-calibration coefficient $C_{IA}(t, \theta, \phi_b)$ and sensor noise ε. The azimuth angle ϕ_b denotes a specific antenna beam b of the fan-beam scatterometer, determined as a discrete azimuth angle resulting from the chosen orbit and the antenna mounting with respect to the satellite ground track.

$$\sigma^0(t, \theta, \varphi_b) = \sigma^0_T(\theta) + C_{IA}(t, \theta, \varphi_b) + \varepsilon \qquad \text{Eqn. 6-3}$$

Because of the postulated characteristics of the employed calibration targets, the backscatter coefficient $\sigma^0_T(\theta)$ of a specific calibration target T can be defined as a function of the incidence angle θ exclusively. The intra-calibration coefficient $C_{IA}(t, \theta, \phi_b)$ incorporates any arbitrary performance anomalies related to the instrument, accounting for variations in individual

antenna characteristics, sensor component degradations or any other anomalies influencing the calibration level of the scatterometer. In the case of a perfectly calibrated instrument, the intra-calibration coefficient $C_{IA}(t, \theta, \phi_b)$ vanishes, resulting in observations $\sigma^0(t, \theta, \phi_b)$ deviating from the target backscatter coefficient, $\sigma^0_T(\theta)$, by the additive instrument noise term ε. Furthermore, instrument noise ε is treated as white Gaussian noise with zero mean in the calibration model. Assuming a perfectly calibrated scatterometer for the time being, an estimate of the unknown backscatter coefficient $\sigma^0_T(\theta)$ can be determined for each calibration target as a function of incidence angle θ by averaging a sufficient number of observations. Analyses of the backscatter characteristics of the used natural calibration targets indicated that the backscatter–/incidence angle dependency of targets can be adequately modeled by a 2-order-polynomial function centered at 40 degrees incidence angle as stated in the following equation (Reimer 2014).

$$\sigma^0_T(\theta) = B^0_T(40°) + \sum_{p=1}^{2} B^p_T(40°) * (\theta - 40°)^p \qquad \text{Eqn. 6-4}$$

Polynomial coefficients of the calibration target backscatter model, i.e., the calibration reference, are determined by an ordinary least square estimation with respect to the extracted data. Separate calibration references are determined for ascending and descending orbit overpasses of the scatterometers because of known systematic differences in the recorded backscatter coefficient (Bartalis et al. 2006). The derived backscatter calibration reference, $\sigma^0_T(\theta)$, constitutes the time invariant backscatter response of a calibration target T. Hence, deviations of the recorded backscatter coefficient $\sigma^0(t, \theta, \phi_b)$ to the calibration reference $\sigma^0_T(\theta)$ are held to give estimates of calibration anomalies incorporated in the calibration coefficient $C_{IA}(t, \theta, \phi_b)$. In the case of the European C-band scatterometers, calibration anomalies can affect particular antenna beams or the entire scatterometer system. Consequently, intra-calibration coefficients are determined for each scatterometer antenna beam ϕ_b separately. Eqn. 5-3 can be solved with respect to the intra-calibration coefficient, resulting in realizations of $C_{IA}(t, \theta, \phi_b)$ affected by additive instrument noise ε per calibration target T as stated in the following.

$$C_{IA,T}(t, \theta, \varphi_b) = \sigma^0(t, \theta, \varphi_b) - \sigma^0_T(\theta) \qquad \text{Eqn. 6-5}$$

Calibration target specific intra-calibration coefficients $C_{IA, T}(t, \theta, \phi_b)$ are deduced for each antenna beam of the scatterometer separately, discriminating between ascending and descending orbit overpasses by utilizing the corresponding calibration target reference $\sigma^0_T(\theta)$. The intra-calibration coefficient is exclusively an attribute of the scatterometer and consequently independent of the calibration target T used for determination. Therefore, the presented intra-calibration approach makes use of numerous calibration targets for a robust determination of the scatterometer related intra-calibration coefficients. Calibration coefficients $C_{IA}(t, \theta, \phi_b)$ are inferred for each antenna beam, per month, as a function of the

incidence angle θ by fitting a straight line through the target specific coefficients $C_{IA, T}(t, θ, φ_b)$ of all employed calibration targets T. Finally, calibrated backscatter observations are derived by subtracting the intra-calibration coefficient C_{IA} from the observed backscatter coefficient $σ^0(t, θ, φ_b)$ to achieve a consistent calibration level of the scatterometer over time.

$$σ_{IA}^0(t, θ, φ_b) = σ^0(t, θ, φ_b) - C_{IA,T}(t, θ, φ_b)$$ *Eqn. 6-6*
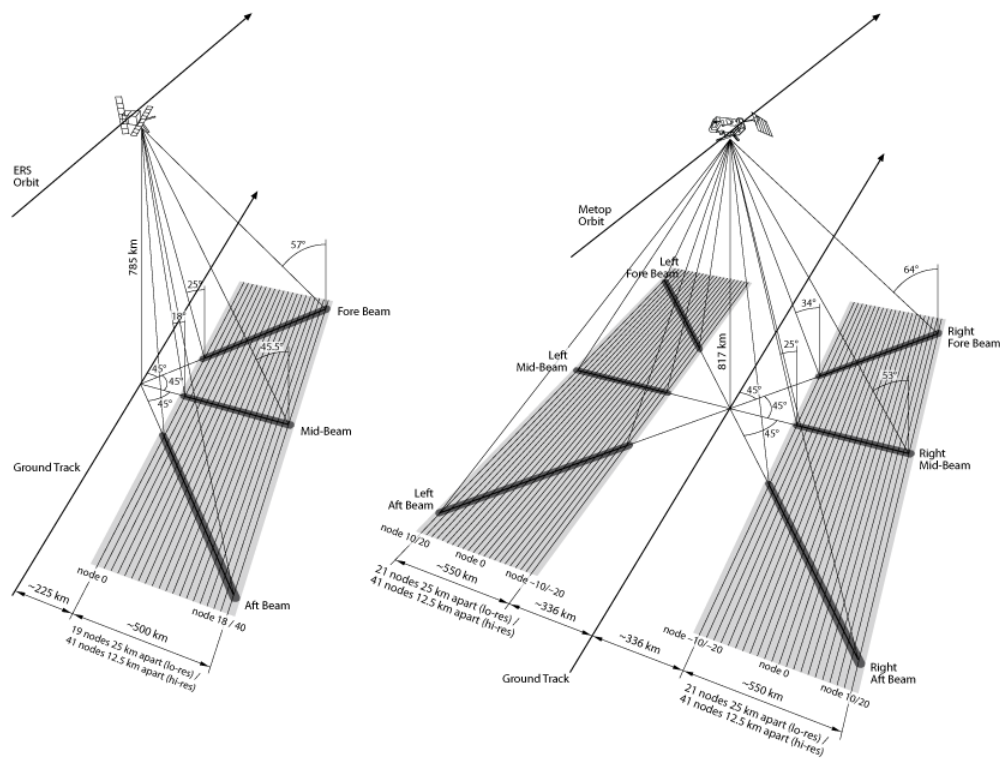

Figure 3: ERS-1/2 AMI-WS and Metop-A ASCAT geometry, introducing swaths, beams and nodes.

6.3   Azimuthal Normalisation

In some regions backscatter σ^0 varies strongly with azimuth or look angle, an effect known as azimuthally anisotropy. These azimuthal effects are accounted for by applying a polynomial correction term to the backscatter values. In this step, the coefficients of the polynomials are computed from the backscatter time series.

In general, the azimuth angle under which a location is seen depends on the beam (fore-, mid- or aft-beam), the swath (left or right) and the satellite direction (ascending or descending), resulting in 6 azimuth configurations for AMI-WS and 12 azimuth configurations for ASCAT respectively. For each of these configurations c, the σ^0 - θ dependency is modelled as a second order polynomial $p_c(\theta)$. The coefficients of these polynomials are determined by fitting the model to all observations falling into the respective configuration category. Furthermore, an overall model $p_o(\theta)$ is fitted to all observations, resulting in a total of 3 x 13 = 39 parameters for ASCAT and 3 x 7 = 21 parameters for AMI-WS respectively.

During the subsequent steps, a correction bias is applied to each backscatter value $\sigma^0_{i,b}$, depending on its azimuthal configuration:

$$\sigma^0_{i,b} \leftarrow \sigma^0_{i,b} + p_o(\theta_{i,b}) - p_c(\theta_{i,b})$$

Eqn. 6-7

This approach has been suggested, and is justified and described in more detail in Bartalis et al. (2006).

6.4   Estimate Noise of Backscatter Measurements

This step initialises the error propagation in the algorithm. It estimates the random noise of a single beam measurement σ^0. This is based on the following observation: all three beams observe the same region (soil moisture), and the fore- and aft-beam have the same incidence angle. Thus, as long as there are no azimuthal effects, the measurements of the for- and aft-beam are comparable, i.e., statistically speaking, they are instances of the same distribution. Hence, the expectation of the difference:

$$\delta := \sigma^0_f - \sigma^0_a$$

Eqn.6-8

should be 0, and its variance should be twice the variance of one of the beams (assuming, the measurements are independent):

$$\text{var}[\delta] = 2 \, \text{var}[\sigma^0]$$

Eqn.6-9

By taking the square root and re-arranging, this gives us an estimate of the standard deviation of σ^0, which is called estimated standard deviation (ESD, see also Figure 4):

$$\text{ESD} = \text{std}[\sigma^0] = \frac{\text{std}[\delta]}{\sqrt{2}}$$

Eqn.6-10

whereby $std[\delta]$ is obtained as empirical standard deviation of δ over the whole time series.
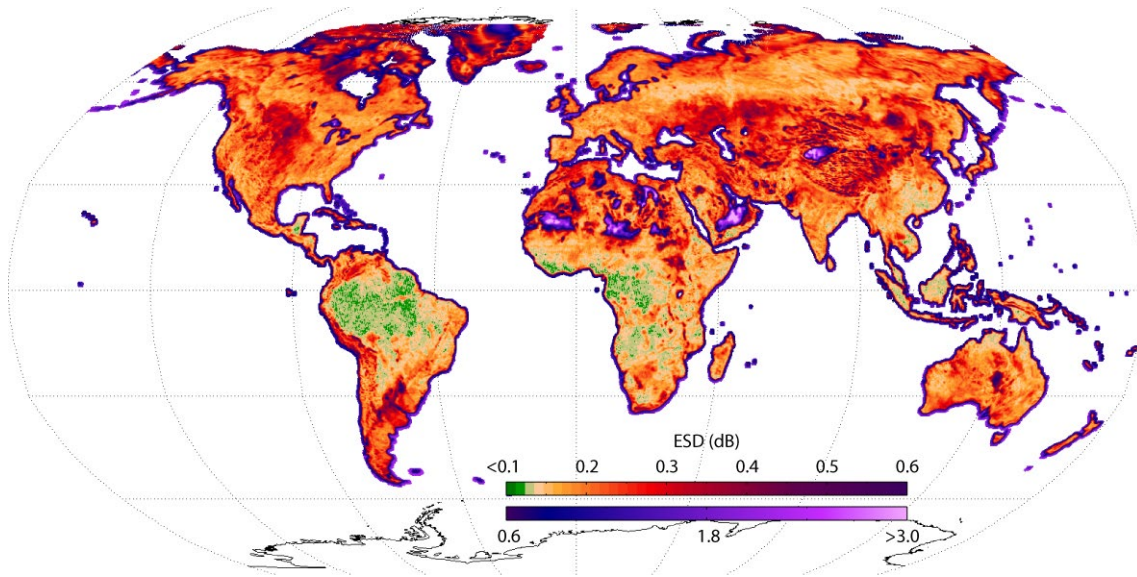
Figure 4: Global distribution of ESD.

6.5   Model Incidence Angle Dependence and Vegetation Correction

The key equation of the model expresses the observed backscatter $\sigma^0(\theta, t)$ as a function of the incidence angle θ at day d, more precisely as a second order polynomial about the reference angle $\theta_{ref} = 40^o$ (Wagner et al. 1999b):

$$\sigma^0(\theta, d) = \sigma^0(\theta_{ref}, d) + \sigma'(\theta_{ref}, d)(\theta - \theta_{ref}) + \frac{1}{2}\sigma''(\theta_{ref}, d)(\theta - \theta_{ref})^2$$

Eqn.6-11

whereby the 0th-order coefficient $\sigma^0(\theta_{ref}, d)$ is the normalised backscatter at the 40° reference incidence angle, and the 1st and 2nd order coefficients $\sigma'(\theta_{ref}, d)$ and $\sigma''(\theta_{ref}, d)$ are referred to as slope and curvature parameters (see Figure 5). Slope and curvature mediate the effect of vegetation on the functional relationship between σ^0 and θ: for sparse vegetation, the curve tends to drop off rapidly, while for fully grown vegetation, it becomes less steep, almost horizontal in the case of rain forest (Figure 5b). In the model, we assume that the vegetation state is always the same at the same day of the year, i.e. it does not change inter-annually, and is thus a function of the day-of-year d. Hence, for each GPI, there will be

366 vegetation curves, each determined by a slope/curvature pair $\sigma'(\theta_{ref}, d)$, $\sigma''(\theta_{ref}, d)$. The slope and curvature parameters, which determine, in conjunction with the incidence angle, the effect of vegetation on the backscatter, are estimated during this step.
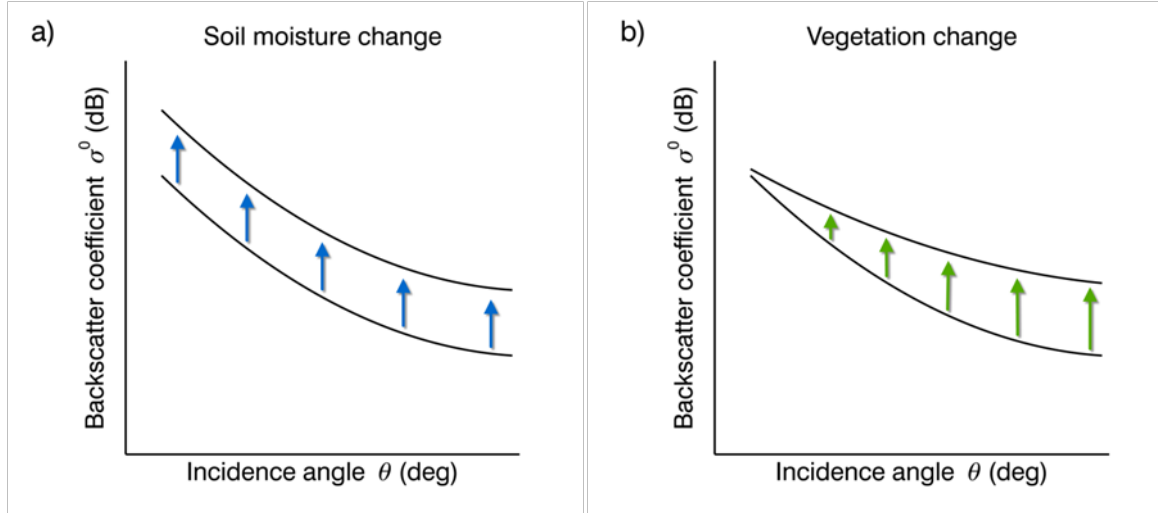
Figure 5: Backscatter as function of the incidence angle. In WARP, it is assumed that an increase in soil moisture simply shifts the curve upwards (a), while a change in vegetation affects its shape, i.e., higher order moments (b).

Slope and curvature are determined as the coefficients of a straight line fitted to the so called *local slopes*. Local slopes are estimates of the first derivative of the backscatter - incidence angle dependency, and are computed as difference quotients between fore- and mid-beam, and aft- and mid-beam, respectively:

$$\sigma'_{local}(\theta, t) = \frac{\Delta\sigma^0}{\Delta\theta} \qquad \text{Eqn.6-12}$$

To be more specific, each backscatter beam-triple $[\sigma_{i,f}, \sigma_{i,m}, \sigma_{i,a}]$ (fore-, mid-, and aft-beam measurements) taken at incidence angles $[\theta_{i,f}, \theta_{i,m}, \theta_{i,a}]$ yields two local slope estimates at day d_i:

$$\sigma'_{i,f}\left(\frac{\theta_{i,m} + \theta_{i,f}}{2}, d_i\right) = \frac{\sigma_{i,m} - \sigma_{i,f}}{\theta_{i,m} - \theta_{i,f}} \qquad \text{Eqn.6-13}$$

$$\sigma'_{i,a}\left(\frac{\theta_{i,m} + \theta_{i,a}}{2}, d_i\right) = \frac{\sigma_{i,m} - \sigma_{i,a}}{\theta_{i,m} - \theta_{i,a}} \qquad \text{Eqn.6-14}$$

These local slopes are taken as instances of the first derivative of Eqn.6-15

$$\sigma'(\theta, d) = \sigma'(\theta_{ref}, d) + \sigma''(\theta_{ref}, d)(\theta - \theta_{ref}) \qquad \text{Eqn.6-15}$$

The width τ of the time window is crucial for the quality of the estimates. A too short time-window length τ yields noisy slope $\sigma'(\theta_{ref}, d)$ and curvature $\sigma''(\theta_{ref}, d)$ estimates, while a too long window filters a remarkable part of the vegetation variation resulting in a bias.

Moreover, the time-window length τ is also dependent on the regional climate. With reference to the spatial resolution of scatterometers, it is assumed that vegetation is not changing remarkably during a less than 2-week period and the seasonal vegetation change does not take longer than 12 weeks. Simulations have been performed to quantify an optimum time window length τ for different climate regions. It was found that a time window length of 6 weeks represents a good balance between noise and bias introduced to the estimate of slope and curvature globally. Slope and curvature values are computed by employing local slope values located within the fixed time window length τ centered at the specific day-of-year $d = doy(t)$. The regression fit is performed by making use of kernel smoother method known as Local Linear Regression utilizing an Epanechnikov kernel. Slope and curvature are determined as the parameters of the local linear fit conducted over the whole range of incidence angles of the local slopes. Therefore, local slope values are assigned with weights according to their distance in time from the evaluation day determined by the Epanechnikov kernel. The error variance of the slope $var\left[\sigma'\left(\theta_{ref}, d\right)\right]$ and curvature $var\left[\sigma''\left(\theta_{ref}, d\right)\right]$ parameters is estimated by means of standard linear estimation theory. Detailed information about the estimation process of slope and curvature can be found in (Melzer 2013).
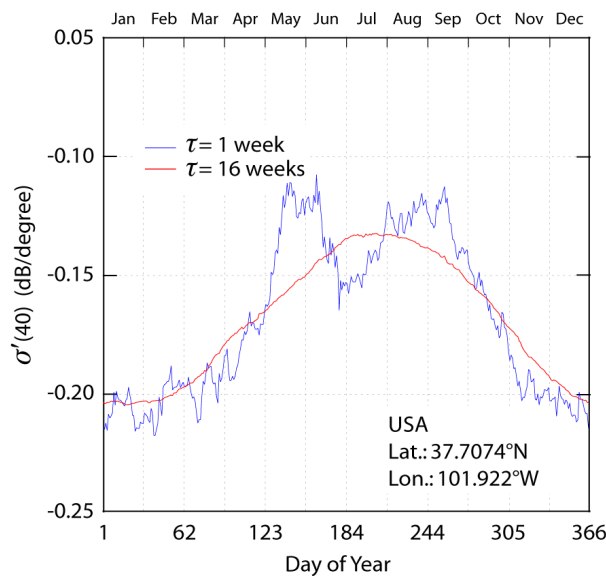
Figure 6: The effect of the time window size on the slope estimate.

6.6 Incidence Angle Normalisation of Backscatter

Backscatter measurements taken at different incidence angles are not directly comparable. Having retrieved the slope and curvature parameters, we can invert the model Eqn.6-15 in order to compute from a backscatter measurement taken at an arbitrary incidence angle the corresponding value at the reference angle. Letting

$$x = \left[\sigma_{i,b}^0(\theta_i), \sigma'(\theta_{ref}, d_i), \sigma''(\theta_{ref}, d_i)\right], \qquad \text{Eqn.6-16}$$

we get Eqn.6-17

$$f(x) = \sigma_{i,b}^0(\theta_{ref})$$
$$= \sigma_{i,b}^0(\theta_i) - \sigma'(\theta_{ref}, d_i)\Delta\theta_i - \frac{1}{2}\sigma''(\theta_{ref}, d_i)(\Delta\theta_i)^2 \qquad \text{Eqn.6-17}$$

Note that we have not included the day of year d_i as parameter of the backscatter $\sigma_{i,b}^0(\theta_i)$ for several reasons. First, in the model the backscatter for a given day is thought of a function of the incidence angle, but not of time. It does depend on time, though not in a direct functional sense, but indirectly, through d_i's effect on the slope and curvature, which it indexes. Second, the time parameter can always be retrieved from the time series via the index i, so adding it to the parameter list is redundant. Third, the notation becomes more concise. However, we must use d_i as argument to the slope and curvatures parameters, since it is used as index into these parameter arrays.

If we assume that the errors of the normalised backscatter, slope and curvature – i.e., the components of x - are uncorrelated, the covariance matrix of x is simply

$$\text{Cov}_x = I_{3x3} \left[ESD^2, \text{var}[\sigma'(\theta_{ref}, d_i)], \text{var}[\sigma'(\theta_{ref}, d_i)]\right]^T \qquad \text{Eqn.6-18}$$

The Jacobian of f is obtained as:

$$\left(\frac{\partial f}{\partial x}\right) = [1, -\Delta\theta_i, -0.5(\Delta\theta_i)^2.] \qquad \text{Eqn.6-19}$$

Thus, according to Eqn.6-35, the noise variance of the normalised backscatter for beam b is Eqn.6-20:

$$\text{var}[\sigma_{i,b}^0(\theta_{ref})] = ESD^2 + \text{var}[\sigma'(\theta_{ref}, d_i)](\Delta\theta_i)^2$$
$$+ 0.25 \ \text{var}[\sigma''(\theta_{ref}, d_i)](\Delta\theta_i)^4 \qquad \text{Eqn.6-20}$$

Finally, the three beams – now having been shifted to a common reference angle – are averaged:

$$\overline{\sigma_i^0}(\theta_{ref}) = \frac{1}{3} \sum_{b \in \{f,m,a\}} \sigma_{i,b}^0(\theta_{ref}) \qquad \text{Eqn.6-21}$$

The corresponding noise variance is given by

$$\mathrm{var}[\overline{\sigma_i^0}(\theta_{ref})] = \frac{1}{9} \sum_{b\in\{f,m,a\}} \mathrm{var}[\sigma_{i,b}^0(\theta_{ref})]$$

Eqn.6-22

As can be seen, averaging over the three beams has the effect that the variance of the noise due to instrument noise, speckle and azimuthal effects is lowered by a factor of three. It does, however, not lower the error due to the lack of fit of the slope model (Wagner 1998).

6.7   Determination of Dry and Wet References

For a given GPI, the dry $\sigma^{dry}(\theta_{ref}, d)$ and wet $\sigma^{wet}(\theta_{ref}, d)$ reference are the historically lowest and highest normalized backscatter values, respectively, measured at this location at a given day d. The dry and wet references are stored as parameter arrays indexed by the time, just as slope and curvature.

The WARP model assumes that the vegetation (i.e., backscatter-vs.-incidence angle) curves for dormant and full vegetation intersect, and that the point of intersection depends on the soil moisture conditions: the intersection points for the driest and wettest conditions are called dry and wet crossover angles, respectively (Figure 7). The wet crossover angle θ_{wet} is at 40 degrees (which is also the reference angle), while the dry crossover angle θ_{dry} is located at 25 degrees (these values have been determined empirically). The importance of the crossover angle concept lies in the fact that at the crossover angles, vegetation has no effect on backscatter (Wagner 1998).
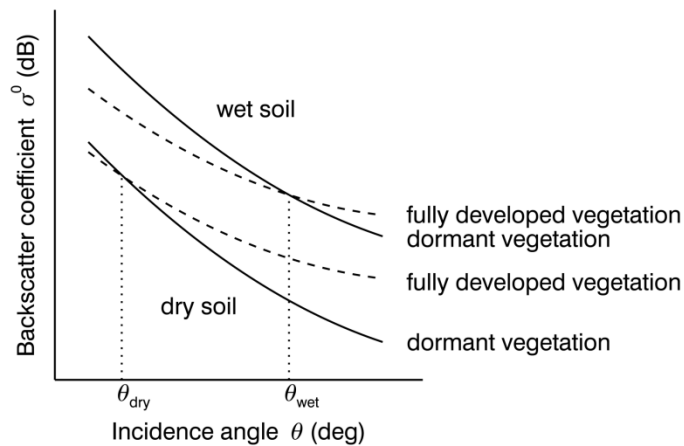

Figure 7: Cross-over angle concept for vegetation correction.

In order to determine the lowest backscatter value irrespective of the vegetation conditions, the normalised backscatter measurements are first shifted to the dry crossover angle:

$$\overline{\sigma_i^0}(\theta_{dry}) = \overline{\sigma_i^0}(\theta_{ref}) + \sigma'(\theta_{ref}, d_i)(\Delta\theta_{dry})$$
$$+ \frac{1}{2}\sigma''(\theta_{ref}, d_i)(\Delta\theta_{dry})^2,$$

Eqn.6-23

with $\Delta\theta_{dry} = (\theta_{dry} - \theta_{ref})$,, and corresponding noise estimate

$$var[\overline{\sigma_i^0}(\theta_{dry})] = var[\overline{\sigma_i^0}](\theta_{ref})] + var[\sigma'(\theta_{ref}, d_i)](\Delta\theta_{dry})^2$$
$$+ \frac{1}{4}var[\sigma''(\theta_{ref}, d_i)](\Delta\theta_{dry})^4.$$

Eqn.6-24

Note the similarity to Eqn.6-17 and Eqn.6-20, but in this case, we are not shifting from the individual incidence angle to the reference angle, but from the reference angle to the dry crossover angle.

From the resulting empirical distribution, the average of the $M = N_{gpi} * 2.5\%$ smallest values is used as an estimate of the lowest backscatter value at the dry crossover angle:

$$\sigma^{dry}(\theta_{dry}) = \frac{1}{M}\sum_{i=1}^{M}\overline{\sigma_{\Pi\{i\}}^0}(\theta_{dry}),$$

Eqn.6-25

whereby Π is a permutation that sorts the timeseries in ascending order w.r.t. the backscatter values. Since the normalised backscatter values have different noise variances (depending on the day and incidence angle of acquisition), there exists no simple general expression for the noise variance of the average, but we have (assuming the noise contributions of the measurements are uncorrelated):

$$var[\sigma^{dry}(\theta_{dry})] = \frac{1}{M^2}\sum_{i=1}^{M}var[\sigma_{\Pi\{i\}}^0(\theta_{dry})].$$

Eqn.6-26

Finally, for each day t, $\sigma^{dry}(\theta_{dry})$ has to be shifted back to the reference angle along its corresponding vegetation curve, in order to obtain $\sigma^{dry}(\theta_{ref}, d)$:

$$\sigma^{dry}(\theta_{ref}, d) = \sigma^{dry}(\theta_{dry}) - \sigma'(\theta_{ref}, d)(\Delta\theta_{dry})$$
$$- \frac{1}{2}\sigma''(\theta_{ref}, d)(\Delta\theta_{dry})^2$$

Eqn.6-27

The noise is given by

$$var[\sigma^{dry}(\theta_{ref}, d)] = var[\sigma^{dry}(\theta_{dry})] + var[\sigma'(\theta_{ref}, d)(\Delta\theta_{dry})^2]$$
$$+ \frac{1}{4}var[\sigma''(\theta_{ref}, d)](\Delta\theta_{dry})^4$$

Eqn.6-28

This is the final estimate of the noise variance for the dry reference. The estimates for the wet reference

$$\sigma^{wet}(\theta_{ref}, t) = \sigma^{wet}(\theta_{wet}) - \sigma'(\theta_{ref}, d)(\Delta\theta_{wet}) \\ - \frac{1}{2}\sigma''(\theta_{ref}, d)(\Delta\theta_{wet})^2$$

Eqn.6-29

(where $\Delta\theta_{wet} = \theta_{wet} - \theta_{ref}$ and its corresponding noise

$$var[\sigma^{wet}(\theta_{ref}, d)] \\ = var[\sigma^{wet}(\theta_{wet})] + var[\sigma'(\theta_{ref}, d)(\Delta\theta_{wet})^2] \\ + \frac{1}{4}var[\sigma''(\theta_{ref}, d)](\Delta\theta_{wet})^4$$

Eqn.6-30

are obtained in a completely analogue fashion, but instead of the 2.5 % lowest values at θ_{dry}, the 2.5 % highest values have to be averaged at the wet crossover angle θ_{wet} in order to compute $\sigma^{wet}(\theta_{wet})$.

It is worth mentioning that due to the selection of the cross-over angles, which are fixed at 25° for σ^0_{dry} and 40° for σ^0_{wet} globally, the dry reference is changing over time, whereas $\sigma^{wet}(\theta_{ref}, d)$ is constant (i.e., it does not depend on the day). This is because the wet crossover angle is equal to the reference angle, and thus $\Delta\theta_{wet} \equiv 0$ (see Figure 8). A global map of abovementioned references is given in Figure 9.
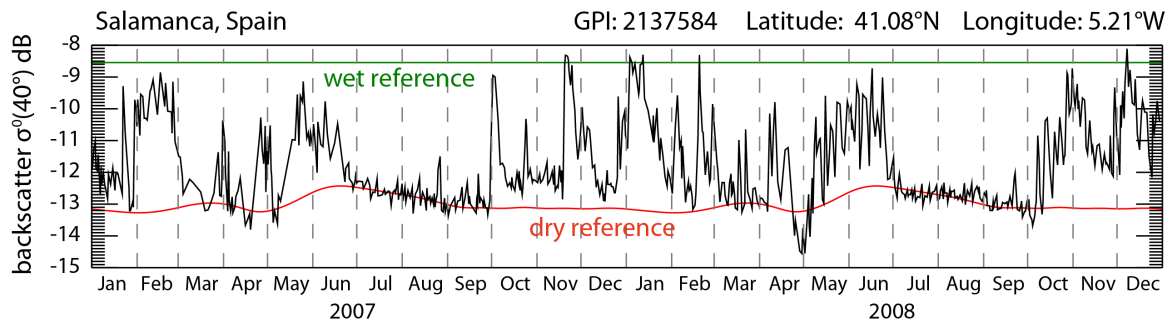

Figure 8: Example of the dry and wet reference characteristics at a GPI near Salamanca, Spain.
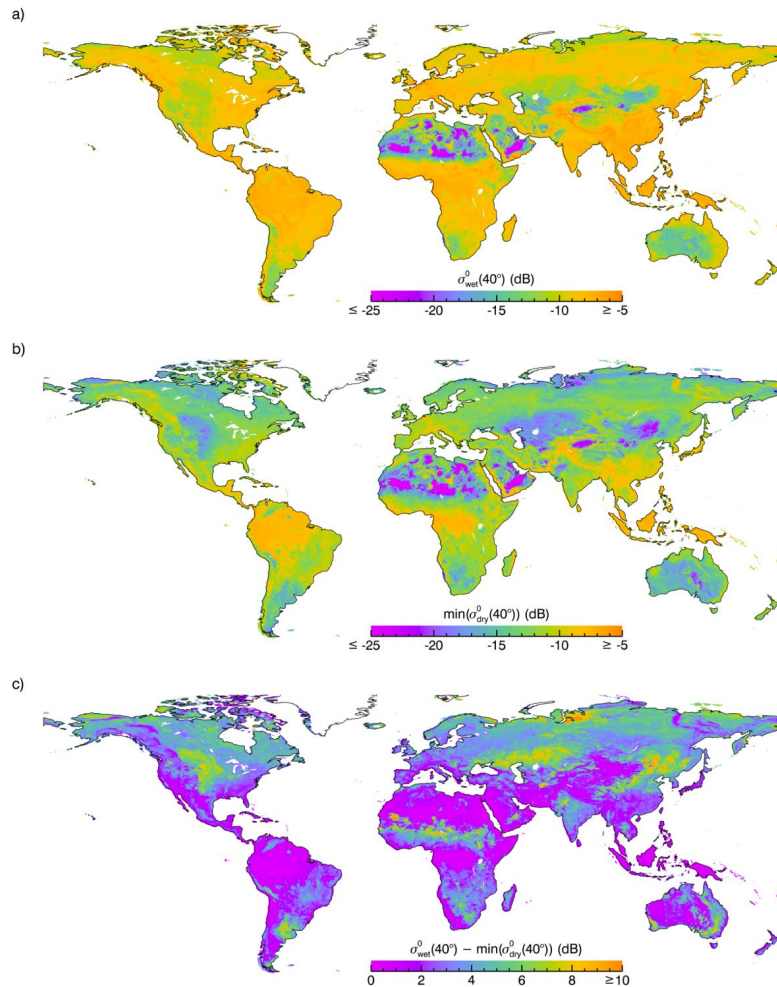
Figure 9: Example of wet backscatter reference (a), lowest dry backscatter reference (b) and sensitivity (c) derived from METOP-A ASCAT (2007-2008).

6.8   Wet reference correction

It is possible that a region has never been captured in a truly saturated condition, which could be simply due to the fact that there were none, or that it did not occur during a satellite overpass. Thus, the assumption that the highest measured backscatter value represents a saturated condition is not valid. In order to correct for the first issue a so-called wet reference correction will be applied in affected regions. However, it is not possible to identify those regions relying only on Scatterometer measurements. Therefore, an external climate classification dataset will be used (Kottek et al. 2006). The utilisation of the wet reference correction is done in 2 steps: first the lowest level of the wet reference is set to -10 dB, globally. Subsequently, in regions with rarely saturated soil moisture conditions (predominantly dry and hot climate zones) the wet reference is further raised until the sensitivity reaches at least 5 dB (see Figure 10).
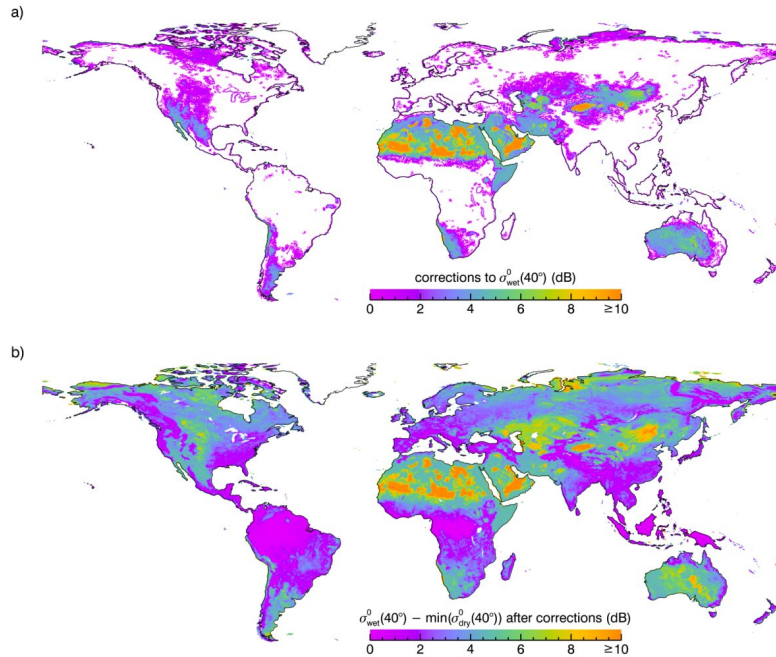
Figure 10: Wet correction (a) and its effect on sensitivity (b) globally.

6.9   Soil Moisture Calculation

The surface soil moisture detection algorithm is a change detection algorithm which basically compares the observed normalised backscatter to the highest (wettest) and lowest (driest) values ever observed at the grid point at day t. Under the assumption of a linear relationship between the backscatter in dB and surface soil moisture, the latter can be estimated as (Wagner 1998):

$$
\begin{aligned}
ssm(i) &= \frac{\overline{\sigma_i^0}(\theta_{ref}) - \sigma^{dry}(\theta_{ref}, d_i)}{\sigma^{wet}(\theta_{ref}, d_i) - \sigma^{dry}(\theta_{ref}, d_i)} . 100 \\
&= \frac{\overline{\sigma_i^0}(\theta_{ref}) - \sigma^{dry}(\theta_{ref}, d_i)}{sens} . 100.
\end{aligned}
$$

Eqn.6-31

Note that $ssm(i)$ is expressed in percent. The difference between wet and dry reference in the denominator is known as *sensitivity* (*sens*).

By proceeding along the lines of the derivation of Eqn.6-20, we obtain the following noise estimate for the soil moisture

$$
\begin{aligned}
var[ssm(i)] = &\frac{var\left[\overline{\sigma_i^0}(\theta_{ref})\right]100^2}{sens^2} \\
&+ var[\sigma^{dry}(\theta_{ref}, t)]\left(\frac{\overline{\sigma_i^0}(\theta_{ref}, d_i) - \sigma^{wet}(\theta_{ref}, d_i)}{sens^2}\right)^2 . 100^2 \\
&+ var[\sigma^{wet}(\theta_{ref}, t)]\left(\frac{\overline{\sigma_i^0}(\theta_{ref}, d_i) - \sigma^{dry}(\theta_{ref}, d_i)}{sens^2}\right)^2 . 100^2
\end{aligned}
$$

Eqn.6-32

6.10 Error Propagation

Let $x = [x_1, \dots, x_p]$ be a p-dimensional observation vector. x is assumed to be an instance of a p-dimensional random variable, with known covariance matrix Σ_x. We are interested in how the covariance transforms under a mapping $: \mathbb{R}^p \rightarrow \mathbb{R}^q, y = f(x)$, transforms, i.e., given x and f, we would like to know the covariance of y, Σ_y. If f is a linear mapping of the form $y = Ax + b, A \in \mathbb{R}^{qxp}, b \in \mathbb{R}^q$, then the covariance transforms like

$$\Sigma_y = A\Sigma_x A^T,$$

Eqn.6-33

whereby A^T denotes the transpose of A.

If, on the other hand, f is a non-linear mapping, we first linearise it by replacing it by its first order Taylor approximation about the operation point x_o:

$$y = f(x) \approx f(x_0) + \left(\frac{\partial f}{\partial x}\right)(x - x_o)$$

Eqn.6-34

whereby $\left(\frac{\partial f}{\partial x}\right)$ is the Jacobian of f.

Putting everything together, we finally obtain Eqn.6-35

$$\Sigma_y = \left(\frac{\partial f}{\partial x}\right)\Sigma_x\left(\frac{\partial f}{\partial x}\right)^T$$

Eqn.6-35

for the variance of y under the mapping f. Eqn.6-35 is the workhorse of the WARP error propagation scheme.

7 Known Limitations

7.1 Computation of Slope/Curvature Parameters

Another line of investigation we shall start is the feasibility of estimating slope/curvature values separately for each year (instead of inter-annually). This was clearly not possible for ERS, but even for ASCAT with its increased temporal resolution the question remains if under these conditions enough data remains within the time window to yield reliable parameter estimates. An interesting aspect would be the possibility to detect slope-outliers (e.g., due to extraordinary events like flash-floods or drought) by comparing the slope estimates for the same day across several years. These outliers would be removed before computing an inter-annually averaged slope parameter which will be used in all cases where a full processing is either deemed too costly (periodic time series append without reprocessing of parameters at TU Wien) or simply not possible (WARP NRT running at EUMETSAT). First investigations concentrating on slope/curvature estimates separately for each year can be found in chapter 8.4.

7.2 Dry and Wet Crossover Angles

The crossover angle concept, already discussed in chapter 6.7, states that at the dry and wet crossover angles, vegetation has no effect on backscatter (Wagner 1998). These crossover angles have been determined empirically based on four study areas (Iberian Peninsula, Ukraine, Mali, and Canadian Prairies). Nevertheless, the empirically determined dry and wet crossover angles are used on a global scale in the surface soil moisture retrieval model. A known limitation of the global use of these crossover angles is that depending on the vegetation type - or more precisely the evolution of biomass of a specific vegetation type -, crossover angles may vary across the globe, which is not yet considered in the model. Furthermore, for some regions on the Earth's surface the crossover angle concept may not be applicable, in particular regions without vegetation cover (i.e., deserts).

7.3 Backscatter in Arid Regions

In arid regions or more specifically in desert environments it appears that the dry reference shows seasonal variations, which are assumed to reflect vegetation phenology. But this cannot be true for desert environments which are characterised by very limited or no vegetation at all. In principle, seasonal variations of the dry reference are desirable to account for backscatter changes induced by vegetation referred to as vegetation correction. Vegetation correction is based upon changes in the slope parameter, which can be also observed in desert environment. These variations seem to have a big impact particularly in areas with very low

backscatter. Hence, it needs to be clarified whether it is a real physical process, noise or something else reflected in the slope parameter.

8 Scientific Advances under Investigation

8.1 Inter-Calibration of Backscatter Data Records

In order to directly compare Level 2 surface soil moisture values retrieved from the ERS-1/2 AMI-WS and MetOp-A/B ASCAT it is a pre-condition that these instruments have more or less exactly the same Level 1 calibration. Unfortunately, this is not yet the case owing to the fact that individual instrument generations underwent a somewhat different calibration procedure: While ASCAT is calibrated using active ground-based transponders only (Wilson et al. 2010), the ERS AMI-WS was calibrated using a combination of transponder measurements and vicarious calibration using Amazon Rainforest as a natural target (Lecomte and Wagner 1998). However, similarities in the instrument technical design of ERS-1/2 AMI-WS and MetOp-A/B ASCAT encourage a Level 1 data inter-calibration of these scatterometer missions towards a long-term consistent data record. Various efforts have been undertaken to examine potential differences in the Level 1 calibration of ERS-1/2 AMI-WS, and MetOp-A/B ASCAT (Bartalis 2009; Elyouncha and Neyt 2013a, b; Talone et al. 2012). Bartalis (2009) inferred inter-calibration biases of ERS-2 AMI-WS and MetOp-A ASCAT by performing a backscatter collocation using observations of the years 2007 – 2008. This study indicates a systematic incidence angle dependent bias between the two instruments with a magnitude of about 0.22 dB (see Figure 11). Elyouncha and Neyt (2013a) investigated calibration differences between ERS-1 and ERS-2 AMI-WS as well as differences in the calibration levels of ERS-2 AMI-WS and MetOp-A ASCAT.
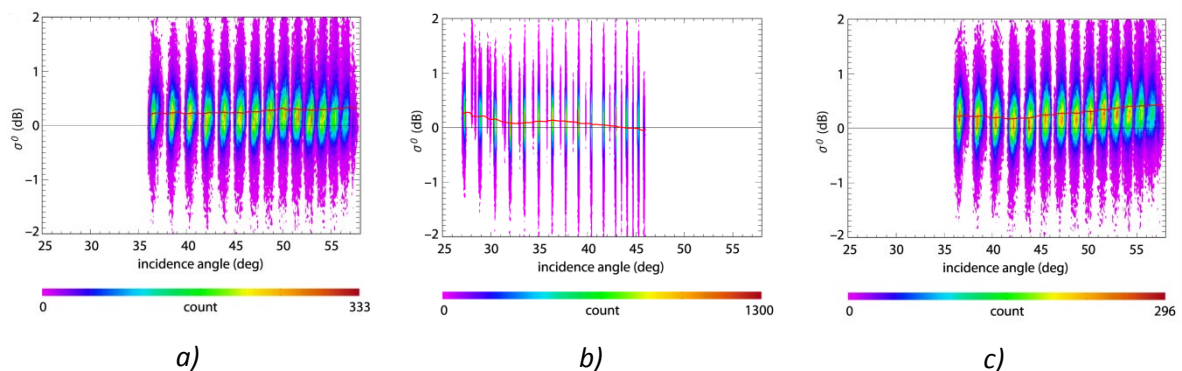
Figure 11: Level 1 inter-calibration biases between ERS-2 AMI-WS and MetOp-A ASCAT as a function of incidence angle a) Fore-beam b) Mid-beam c) Aft-beam [from Bartalis (2009)]

Furthermore, a number of diverse inter-calibration methodologies, covering backscatter collocation and model based (rainforest, ocean and sea ice) inter-calibration, were conducted and resulting inter-calibration biases of the individual approaches were compared. Examined inter-calibration methodologies reveal small discrepancies in the magnitude of the resultant Level 1 calibration biases (see Figure 12 and Figure 13), but the different methods provide relatively consistent inter-calibration results. Additionally, it was found that model based

inter-calibration methodologies like rainforest, ocean and sea ice inter-calibration performed more accurately, in terms of their standard deviation, than inter-calibration utilising collocated backscatter measurements. However, inter-calibration biases estimated by collocation between ERS-2 AMI-WS and MetOp-A ASCAT found by Elyouncha and Neyt (2013a) confirm the results discovered Bartalis (2009). Nevertheless, the assumption of neglecting backscatter differences in time, space and measurement geometry is questionable in terms of backscatter collocation, because of the larger biases and larger standard deviations observed with respect to model-based methods.

a)   *b)*   *c)*

Figure 12: Level 1 calibration biases between ERS-1 and ERS-2 AMI-WS Collocation (red), ocean (blue), sea ice (black) and rainforest (green) a) Fore-beam   b) Mid-beam   c) Aft-beam [from Elyouncha and Neyt (2013a)]
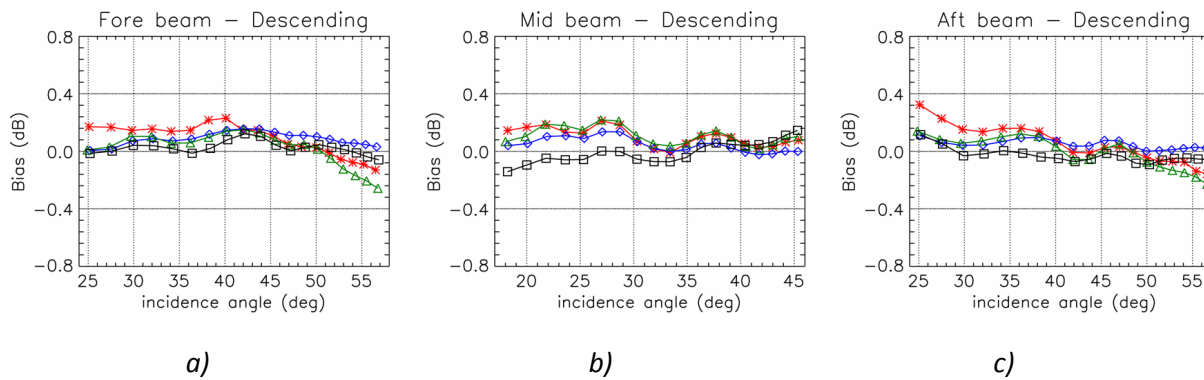

a)   *b)*   *c)*
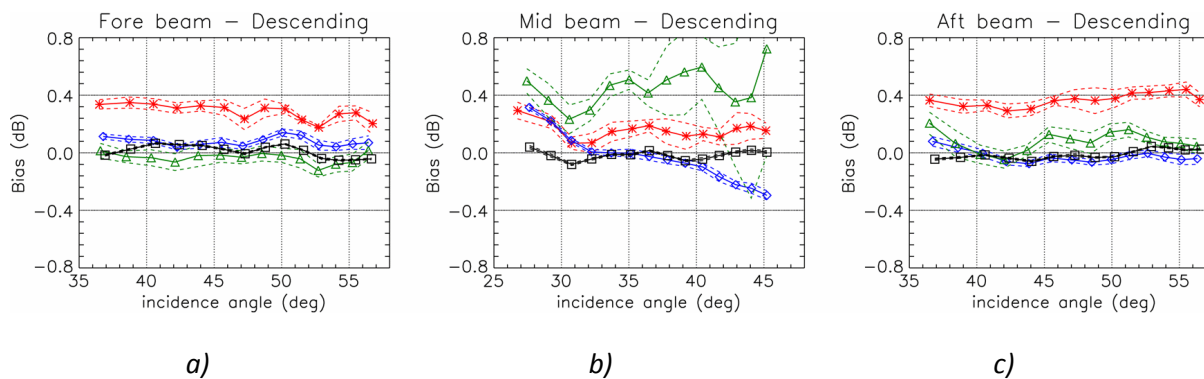
Figure 13: Level 1 calibration biases between ERS-2 AMI-WS and MetOp-A ASCAT. Collocation (red), ocean (blue), sea ice (black) and rainforest (green). a) Fore-beam   b) Mid-beam   c) Aft-beam [from Elyouncha and Neyt (2013a)]

Consolidated and accurate magnitudes of Level 1 inter-calibration biases are not yet well enough known for the considered scatterometer missions. Therefore, additional studies need to be carried out to quantify potential inter-calibration biases between diverse scatterometers. It is foreseen to concentrate on model based inter-calibration methodologies in order to achieve a consistent calibrated Level 1 backscatter observations of ERS-1/2 AMI-WS and MetOp-A/B ASCAT. In case that this activity will be successful it will be possible to merge these data sets already at the backscatter measurement level. Therefore, the objective

is to merge consistent calibrated Level 1 backscatter observations of various scatterometer missions towards a uniform Level 2 surface soil moisture retrieval, rather than utilising a statistical properties matching of different Level 2 soil moisture data sets through CDF matching (as was done within the WACMOS and CCI phase I projects). A rigorous determination of potential inter-calibration biases between various scatterometer missions is supposed to be inferred by utilising a set of different inter-calibration methodologies simultaneously. Accordingly, causes of discrepancies in the bias magnitudes between various inter-calibration approaches have to be investigated in depth. Furthermore, the combined use of different inter-calibration methodologies will demand an obvious calibration information selection scheme. Such a scheme is needed to derive highly accurate calibration biases with respect to the alignment of calibration levels of diverse scatterometer missions in accordance to defined calibration requirements. Each inter-calibration methodology will provide valuable calibration information, but with reference to the postulated calibration requirements, individual methodologies need to be selected to meet the claimed requirements. As can be envisioned, an important task is to define and formulate calibration requirements envisaged to achieve a certain accuracy of the final Level 2 soil moisture. A minimum accuracy requirement for global soil moisture data is postulated of 4 vol%, as specified for the SMOS mission. However, soil moisture retrievals from scatterometers do not allow estimations in absolute manner, but the requirement can be approximately converted to a percentage index as derived for scatterometer observations. Scatterometer retrievals are capturing soil moisture conditions ranging from completely dry (0 vol%) to saturated soil moisture conditions (app. 50 vol%). As a consequence, the postulated accuracy can be translated to a value of 8%. Inter-calibration biases and their direct effect on the Level 2 surface soil moisture retrieval from scatterometer data was analysed by Hahn et al. (2012), highlighting that a bias of 0.22 dB between two instruments will cause soil moisture deviations ranging from 4% to 16% (see Figure 14). Accordingly, to achieve the postulated soil moisture accuracy of 4 vol% or respectively 8% index globally, inter-calibration biases between two scatterometers have to be smaller than app. 0.1 dB.
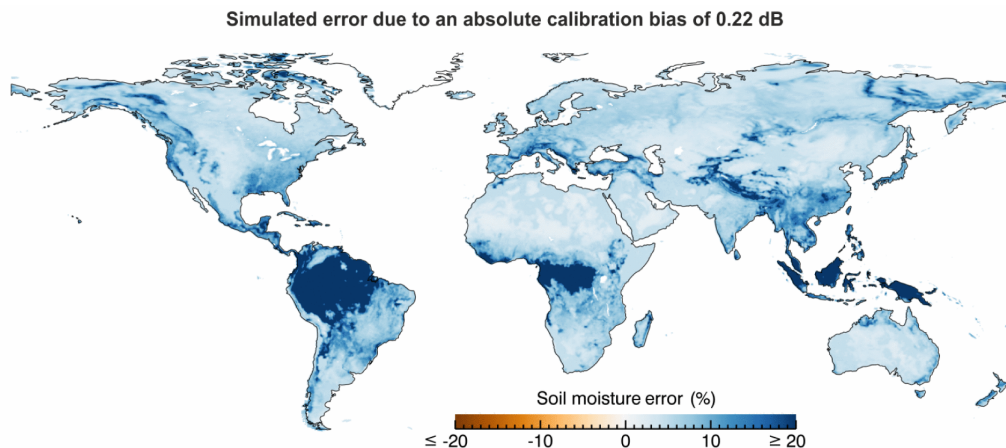
Figure 14: Effect of an inter-calibration bias of 0.22 dB on the final surface soil moisture retrieval [from Hahn et al. (2012)]

Furthermore, to achieve highly accurate surface soil moisture retrievals, Level 1 backscatter observations of different scatterometer missions have to be acquired at almost the same spatial resolution. Hahn 2012 discovered errors in soil moisture ranging from appr. -6% - 6% globally, because of a spatial resolution mismatch between ERS-2 AMI-WS (50 km) and MetOp-A ASCAT (25 km) Level 1 backscatter observations. This source of error can be neglected if all used Level 1 scatterometer products do have the same spatial resolution.

Reimer (2014) introduced a model based inter-calibration methodology which accounts for temporal calibration biases within a specific scatterometer mission and subsequently considers temporal invariant inter-calibration biases between various scatterometer missions. This approach does not make use of various inter-calibration methodologies, but employs a number of natural calibration targets (rainforests) supposed to result in a more robust estimation of inter-calibration biases. Inter-calibration biases are determined by a calibration information selection scheme taking into account the standard deviations of each individual calibration target and the calibration model. Results of this inter-calibration methodology are provided in Figure 15 for ERS-2 AMI-WS and MetOp-A ASCAT and in Figure 16 for MetOp-A and MetOp-B ASCAT. This scatterometer calibration methodology may serve as a starting point for further studies, incorporating different model based inter-calibration methodologies and an obvious calibration information selection scheme. As can be seen in Figure 15 and Figure 16, inter-calibration biases are modelled as a 1-order polynomial function with respect to the incidence angle. Modelled inter-calibration biases of ERS-2 AMI-WS and MetOp-A ASCAT are ranging from 0.39 dB to -0.08 dB, highlighting the need for a proper instrument inter-calibration to achieve soil moisture accuracies of 4 vol%.
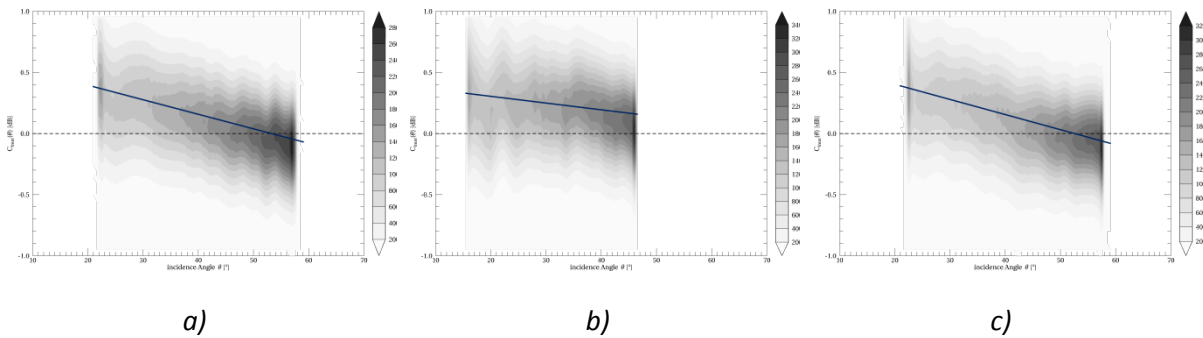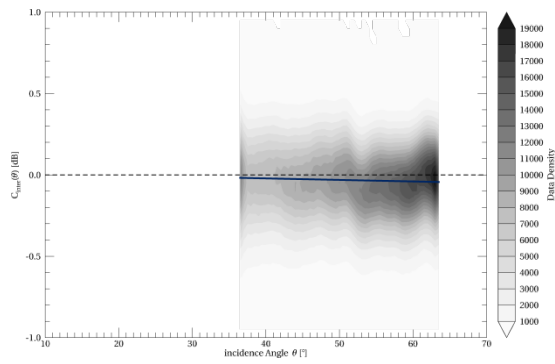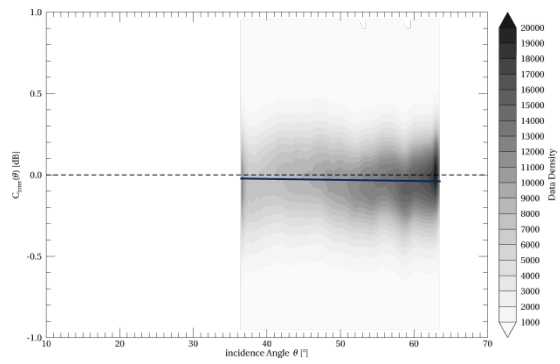
Figure 15: Level 1 inter-calibration biases between ERS-2 SCAT and MetOp-A ASCAT. Data density plot (greyscales) and estimated bias (blue line) a) Fore-beam   b) Mid-beam   c) Aft-beam [from Reimer (2014)]

1-order polynomial functions representing the inter-calibration bias are not capable to describe calibration deficiencies at specific incidence angles, but are applicable to correct for possible biases and first-order deviations of the backscatter versus incidence angle behaviour. Moreover, residual calibration insufficiencies were encountered for MetOp-A ASCAT depicted in oscillating alterations with respect to the incidence angle (Figure 15-b). Such residual calibration artefacts were not discovered for ERS AMI-WS, which foster the suggestion that the cause of these oscillations may be related to the exclusive use of ground-based transponders for calibration. Using ground-based transponders for calibration will provide calibration information of the antenna gain pattern at distinct locations. In the ERS scatterometer era, transponder calibration campaigns were supported by relative calibration campaigns over Amazon Rainforest to fine-tune the antenna gain pattern. However, the cause of these oscillating alterations (wiggles) needs to be investigated for MetOp-A and B ASCAT to derive consistent calibrated Level 1 backscatter records.

Calibration biases are estimated within a restricted range of the scatterometer measurement space, depending on the utilised inter-calibration methodology and/or calibration targets. Therefore, one research question which arises is if calibration biases derived within a restricted range of the measurement space could accurately represent calibration biases across the entire measurement space of the considered scatterometers. The relationship may be constant but could also be linear or any other function with respect to the magnitude of the backscatter observations (see Figure 17), which will require further research and encourage the use of different inter-calibration methodologies.
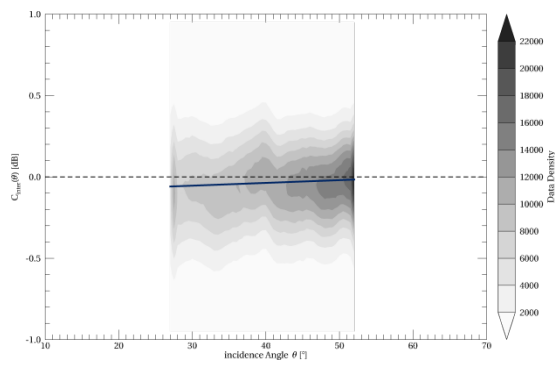
Figure 16: Level 1 inter-calibration biases between MetOp-A and MetOp-B ASCAT. Data density plot (greyscales) and estimated bias (blue line) Fore-beam a) left swath d) right swath Mid-beam b) left swath e) right swath   Aft-beam c) left swath f) right swath [from Reimer (2014)]
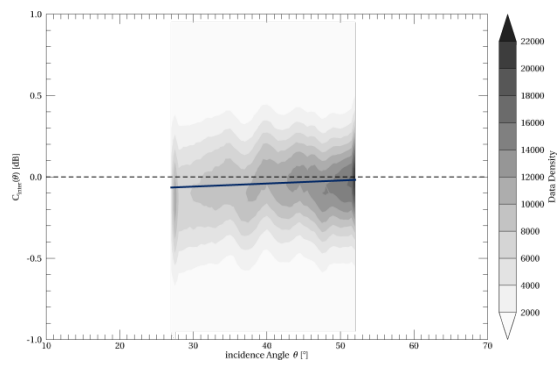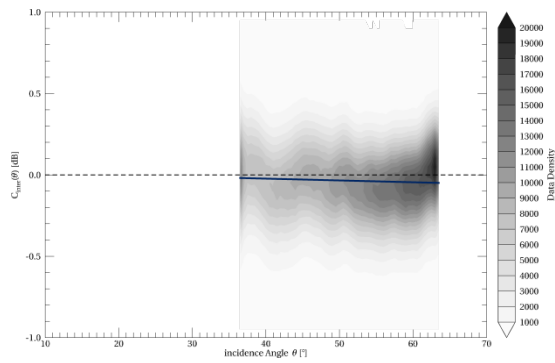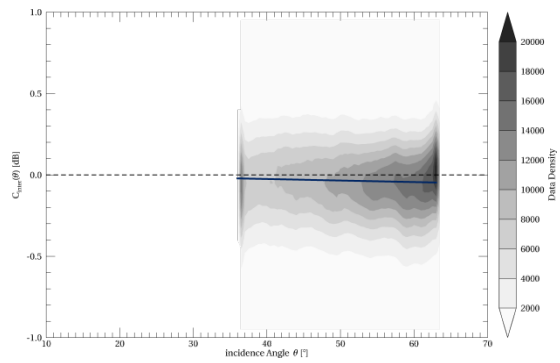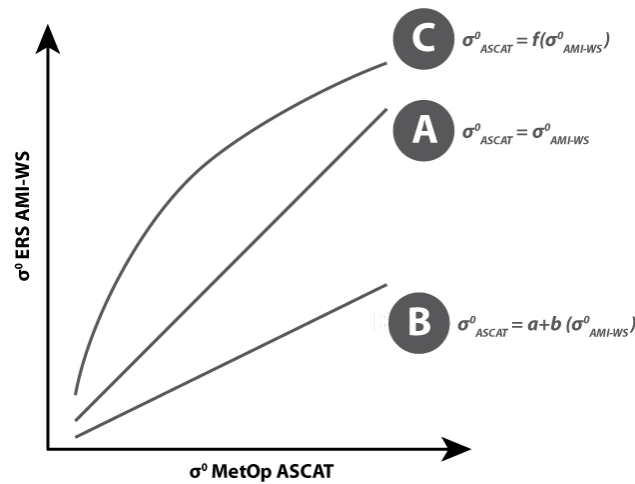
Figure 17: Possible relationship between MetOp ASCAT and ERS AMI-WS backscatter.

8.2 Estimation of Diurnal Variability

It has been noted that the backscatter measurements and, consequently, the Level 2 (L2) surface soil moisture retrievals from satellite platforms, although not dependent on temperature, show in some regions a difference between morning (i.e., day or sun-lit) and evening (i.e., night or dark) acquisitions(Friesen et al. 2012; Friesen et al. 2007). Currently, it is not clear if these observed diurnal differences are due to changes in the instrument between ascending or descending passes (e.g. due to the strong temperature differences in the sun-lit or dark orbital phases), shortcomings in the retrieval algorithm (e.g. neglecting diurnal differences in vegetation water content), or if these are just a natural expression of diurnal patterns of the surface soil moisture content.

We will try to identify underlying reasons for diurnal differences by comparing satellite ascending and descending orbit soil moisture retrievals. ASCAT measurements are performed for descending orbits (equator crossing 09:30, local time) and ascending orbits (equator crossing 21:30, local time). To assess the impact of observation time on soil moisture retrieval we will split the processing of ASCAT based on the TU Wien retrieval method as shown in Figure 18.

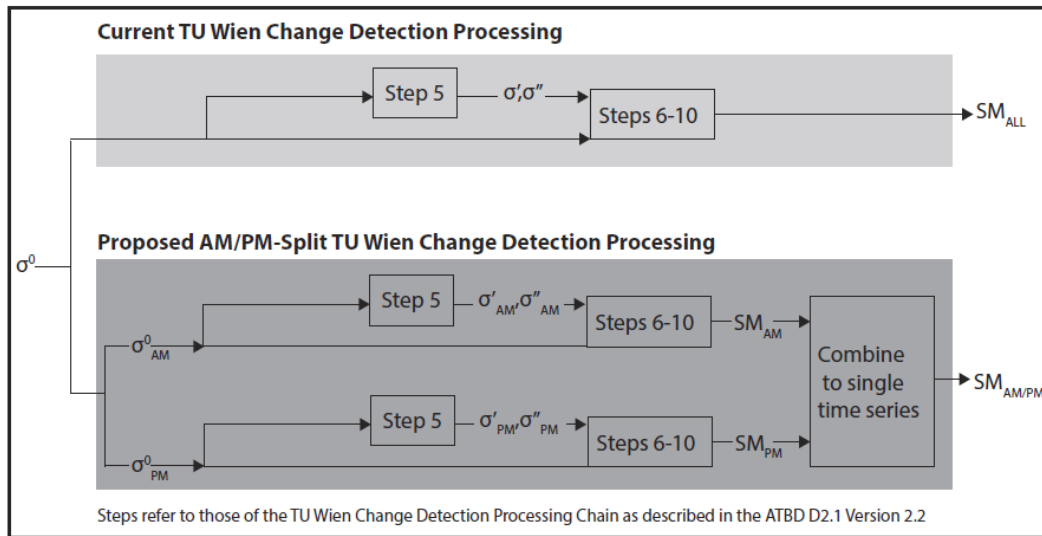*Figure 18: Proposed Possible relationship between MetOp ASCAT and ERS AMI-WS backscatter.*

8.3   Improved Modelling of Volume Scattering in Soils

It has long been noted that backscatter measurements over desert areas and semi-arid environments during a long dry spell exhibit an unusual behaviour that may lead to a situation where soil moisture from scatterometers is often less accurate than radiometer retrievals (Gruhier et al. 2009; Wagner et al. 2007). Due to the lack of high-quality reference data in arid and semi-arid environments, it has not yet been possible to determine the physical reasons for this phenomenon. Various hypotheses exist, including the occurrence of volume scattering from deeper soil layers and/or the presence of bedrock close to the surface.

An illustration of the aforementioned situation is given in Figure 19. It shows two rain events clearly depicted by SMOS (Kerr 2012) and soil moisture modelled by the NOAH Global Land Data Assimilation System (GLDAS) (Rodell et al. 2004). However, METOP-A ASCAT soil moisture surprisingly decreases at the two events and interestingly starts to increase during the drying-out afterwards. This time series from a grid point located in the An Nafud Desert, which is the northernmost desert on the Arabian Peninsula, is an example for a previously identified region affected by unusual backscatter behaviour. A preliminary attempt to look at such phenomena globally has been by investigating the Pearson correlation coefficient between soil moisture from METOP-A ASCAT and GLDAS, as shown in Figure 20. At a first examination it seems that mostly desert areas and arid environments exhibit a negative relationship. But for the interpretation it is also important to consider the time period which has been used for the computation. In this case two years (2010 and 2011) determine the Pearson correlation coefficient, which means that if unusual backscatter behaviour may have occurred within a smaller time period, an overall stronger signal characteristic can have overshadowed such events. In order to get around this shortcoming a next step will be

calculating the Pearson correlation coefficient for certain time windows, which will allow determining the point in time for such unusual backscatter events. This will limit the variety of influencing factors and can offer an understanding of the dominant physical process. Nevertheless, it must be ensured that the results are significant and meaningful, since the Pearson correlation coefficient is not robust and sensitive to outliers.

A special challenge in desert areas has always been the low signal to noise (SNR) ratio, since rain events are very rare and/or sporadic. Hence, backscatter dynamics are usually very low and prone to be affected by azimuthally anisotropy, speckle and instrument noise (see Figure 21). In addition, the signal intensity is sometimes also very close to the measurability limit of the instrument, which is in the order of -30 dB (see Figure 22). As a result, change detection in these areas is a difficult task from the very beginning. Furthermore, it has been noted that the quality of the model parameters, which are derived during the retrieval of METOP-A ASCAT surface soil moisture, is not as reliable as desired. For example, the estimation of the wet reference is problematic due to the fact that saturated soil conditions cannot always be captured, simply because there were none at all or not during a satellite overpass. This problem has already been addressed by using a wet correction based upon a climate classification map (Kottek et al. 2006). During this process dry areas will be identified and the wet reference will be raised to a backscatter value simulating the signal acquired under saturated soil conditions. This approach is not fully perfect yet and has still room for improvements. For example, discrepancies could have been detected in Spain and Turkey, where first investigations suggest that this can be attributed to the wet correction (see Figure 23).

There is also evidence that another model parameter indicates inconsistencies in dry environments. It appears that the dry reference shows seasonal variations, which are assumed to reflect vegetation phenology (see Figure 24). However, this cannot be true for desert environments with very limited or no vegetation at all. In principle the vegetation correction is based upon changes in the slope parameter, which can be also observed in desert environment (see Figure 25). These variations seem to have a big impact particularly in areas with very low backscatter (< -25 dB). Hence, it needs to be clarified whether it is a real physical process, noise or something else reflected in the slope parameter.

All these preliminary findings suggest that the characteristics of backscatter should be explored in more depth in very dry environments. In this context it would be also very beneficial looking into historic and on-going scatterometer missions, if similar backscatter characteristics can be observed. In future, it is also planned to improve the estimation of the dry and wet reference in order to increase the accuracy of the final surface soil moisture estimates. Regarding validation, it should be taken into account that the model dataset GLDAS has limitations and therefore other datasets shall be used to support the validations as well. However, as mentioned in the beginning, high-quality datasets are hardly available and

therefore a cross-comparison between datasets from different sources is even more important. Future work is also planned on using different methods in detecting and quantifying unusual backscatter characteristics. At present, a first approach goes into the direction of local linear regression (Cleveland and Devlin 1988), which allows analysing and decomposing time series characteristics.
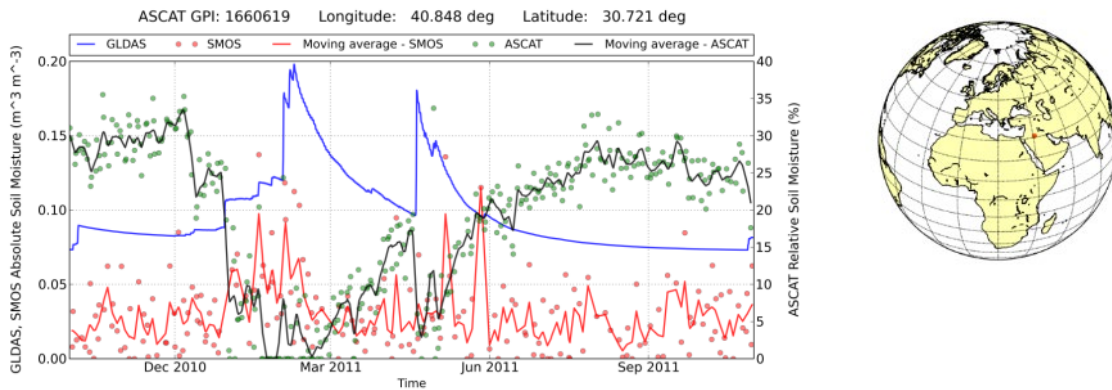
Figure 19: Soil moisture times series from GLDAS (blue), SMOS (red) and METOP-A ASCAT (black) from a grid point located in the An Nafud Desert. Two rain events are clearly visible in the GLDAS and SMOS data set in February/March 2011 and May/June 2011. However, METOP-A ASCAT shows almost an inverted behaviour.
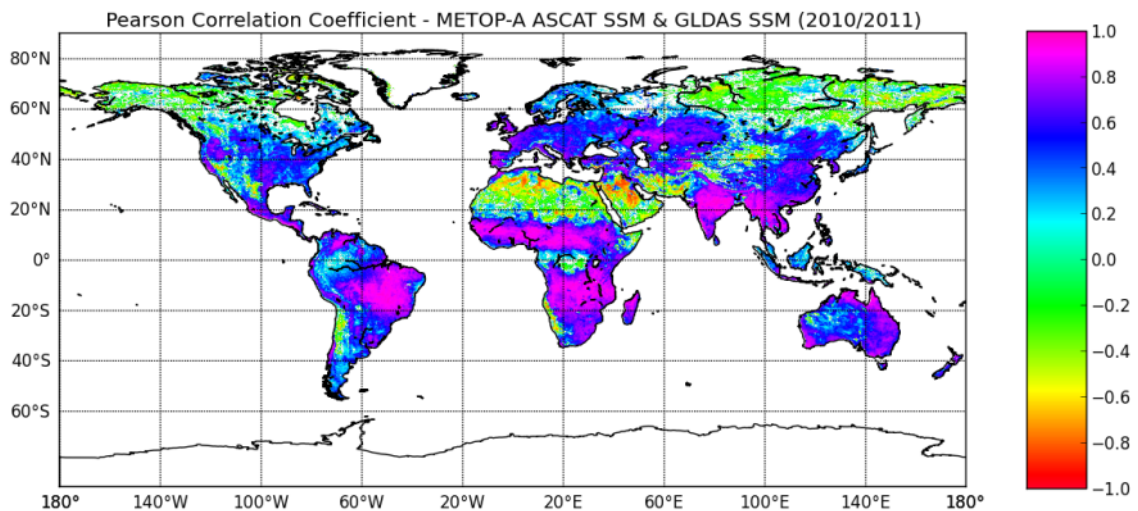
Figure 20: Pearson correlation coefficient between METOP-A ASCAT surface soil moisture and GLDAS modelled soil moisture from the first layer (0-10cm) for the time period 2010/2011. The underlying grid is regular with a pixel spacing of 0.25°. A temporal matching and spatial nearest neighbour search has been performed before computing the correlation coefficient. Grid points with a p-value below 0.05 have been masked.
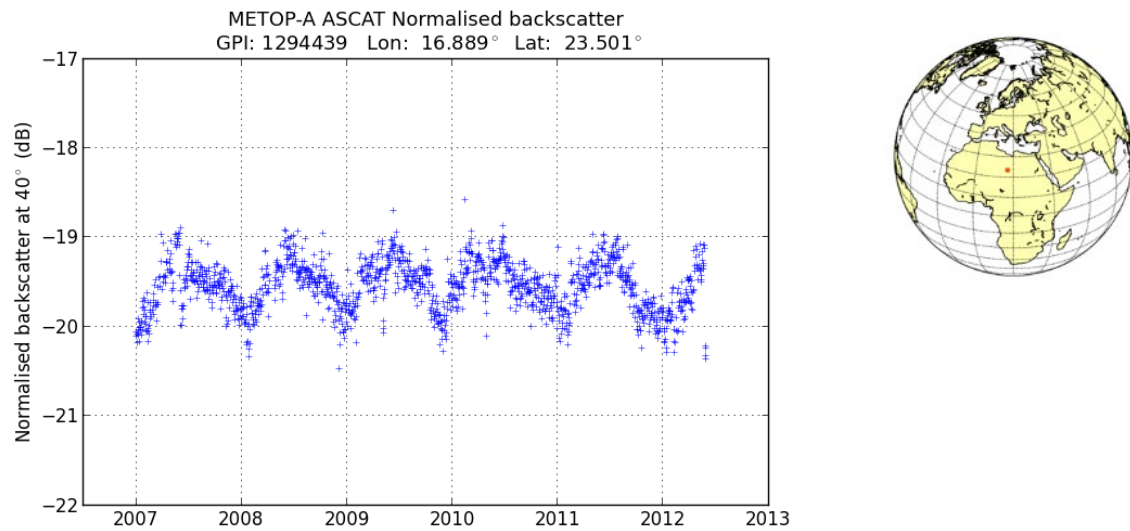
Figure 21: Normalised backscatter at 40° incidence angle from METOP-A ASCAT in the Sahara (Libya). A yearly cycle can be observed, but the overall signal variation is only in the order of 1dB.
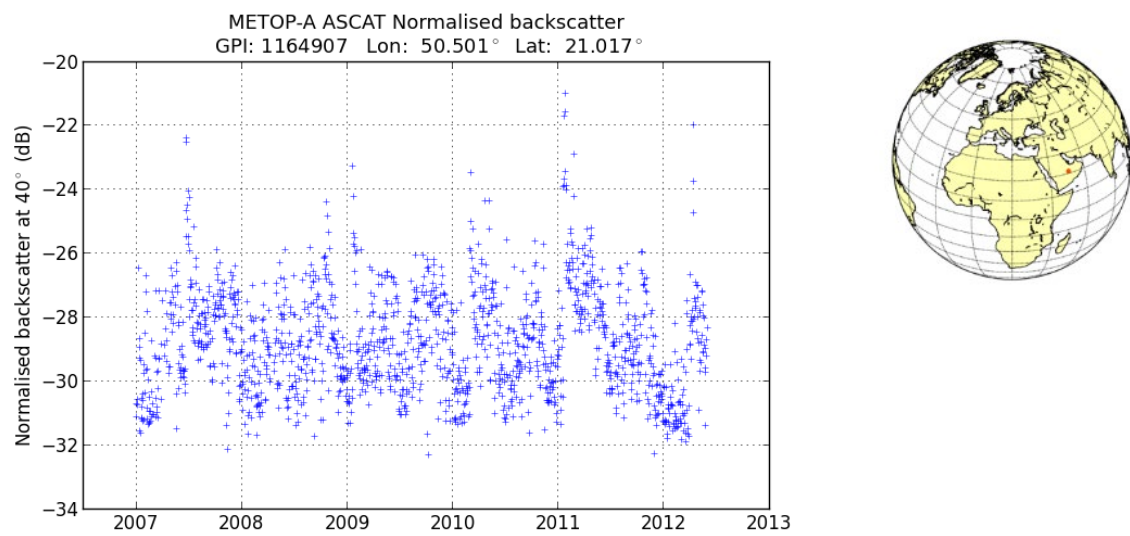
Figure 22: Normalised backscatter at 40° incidence angle from METOP-A ASCAT in the Rub' al Khali Desert (Saudi Arabia). Backscatter is very low and prone to be affected by azimuthally anisotropy, speckle and noise.
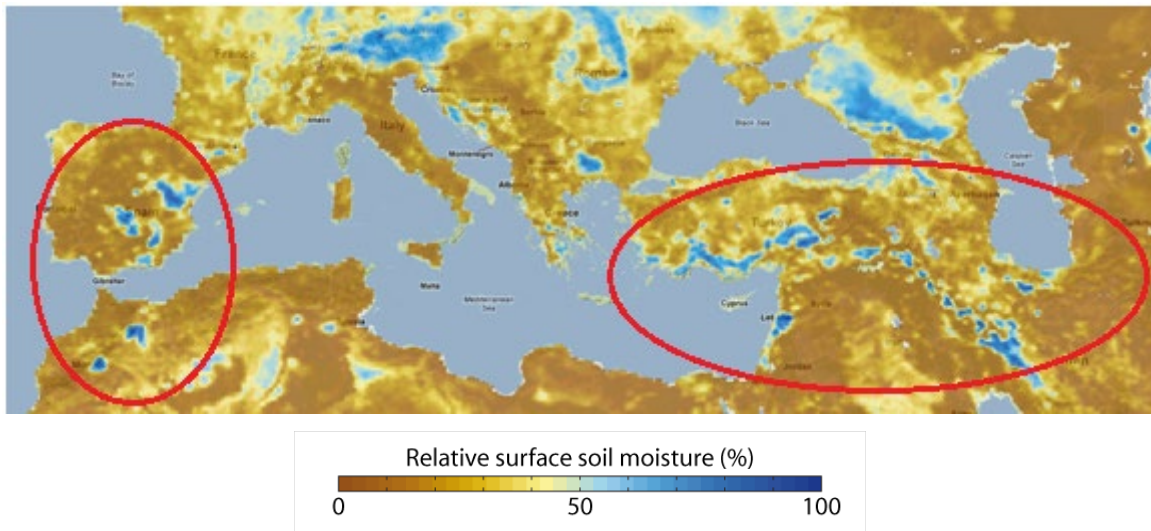
Figure 23: The map shows METOP-A ASCAT soil water index (SWI) T=20 overlaid on Google Maps from August 2012. The two marked areas show unexpected wet soil conditions, where first qualitative analyses suggest that these patterns can be attributed to an inaccurate wet correction.
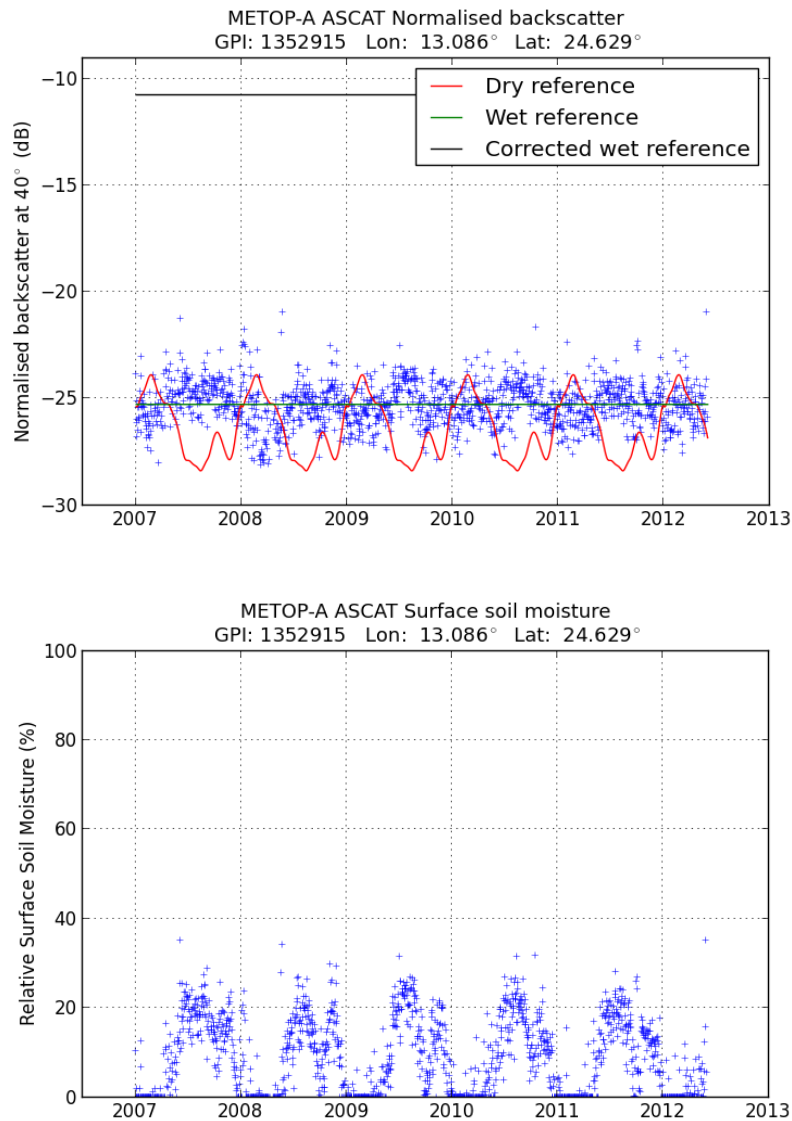
Figure 24: The upper time series illustrates METOP-A ASCAT normalised backscatter at 40° incidence angle, as well as the dry (red), wet (green) and wet corrected (black) reference. The dry reference shows a seasonal variation, which typically should account for vegetation influence. But there should be no vegetation influence in very arid environment. However, as a result the dry reference modulates an incorrect signal onto the final surface soil moisture values shown in the bottom time series.
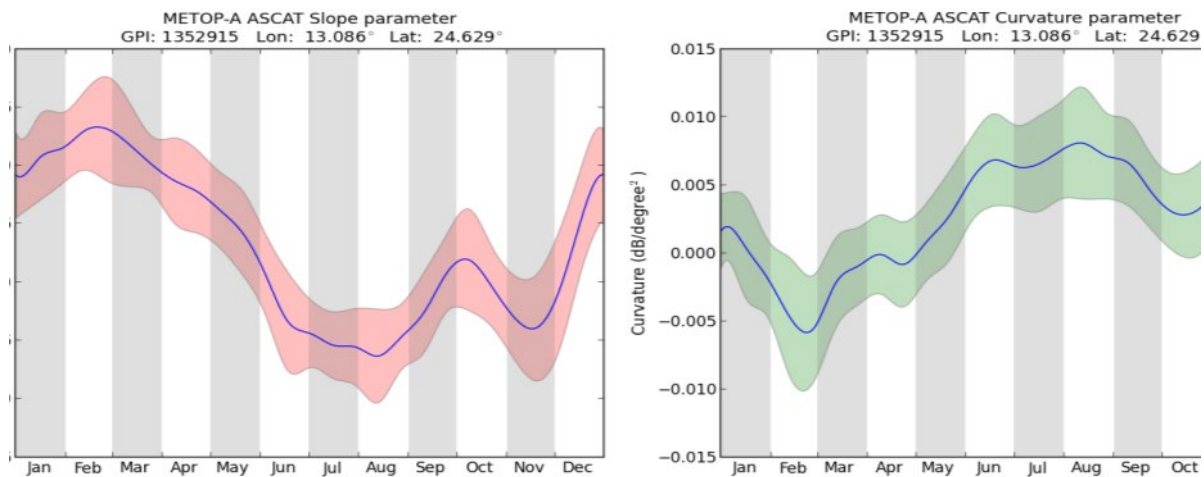
Figure 25: The slope (left) and curvature (right) parameter from the same grid point as shown in Figure 24. Since the incidence angle - backscatter behaviour is characterised by a second order polynomial, the slope and curvature parameters are defined as the first and second derivative of backscatter from the incidence angle. Changes of the slope and curvature parameter are normally associated with vegetation changes; however, in this dry environment a different phenomenon must be responsible. Additionally, the uncertainty range of each parameter is indicated by a coloured tube around the curve.

8.4 Better Vegetation Modelling

Backscatter depends not only on a) soil moisture, but also on b) the incidence angle between the beam and the surface and c) the vegetation state. In order to retrieve soil moisture estimates from backscatter, we must account for the effects of varying incidence angle and vegetation state have to be accounted for. In WARP, this is done by modelling the dependency between backscatter and the incidence angle as a second order polynomial for each day of the year (the vegetation state is assumed to be a function of the day of the year). The linear and quadratic coefficients of the polynomial are referred to as slope and curvature parameter, respectively; they are required for both incidence angle normalisation and vegetation correction (items b) and c) above), and are thus among the most important WARP parameters.

The process for estimating the slope and curvature parameters has been extensively investigated during the last years. Estimates are now obtained using a kernel smoother for regression (previously, standard regression on a subset followed by spline interpolation was used). In addition to the "aggregated" parameters computed as an average over all available years, yearly slopes (and curvatures) are now computed. These parameters have a higher variance than their aggregated counterparts, but can give valuable insight into the vegetation dynamics of a single year. Preliminary results show that there can be significant differences between the averaged and the yearly slope curves (Figure 26), and the question is, whether

these yearly curves will result in better soil moisture estimates. A comparison of the yearly slope curve for a test site in Texas, USA, against FAPAR and VOD (optical depth) shows a good correlation, about 0.75, suggesting that the yearly slope contains information about the yearly vegetation dynamics (Figure 27).
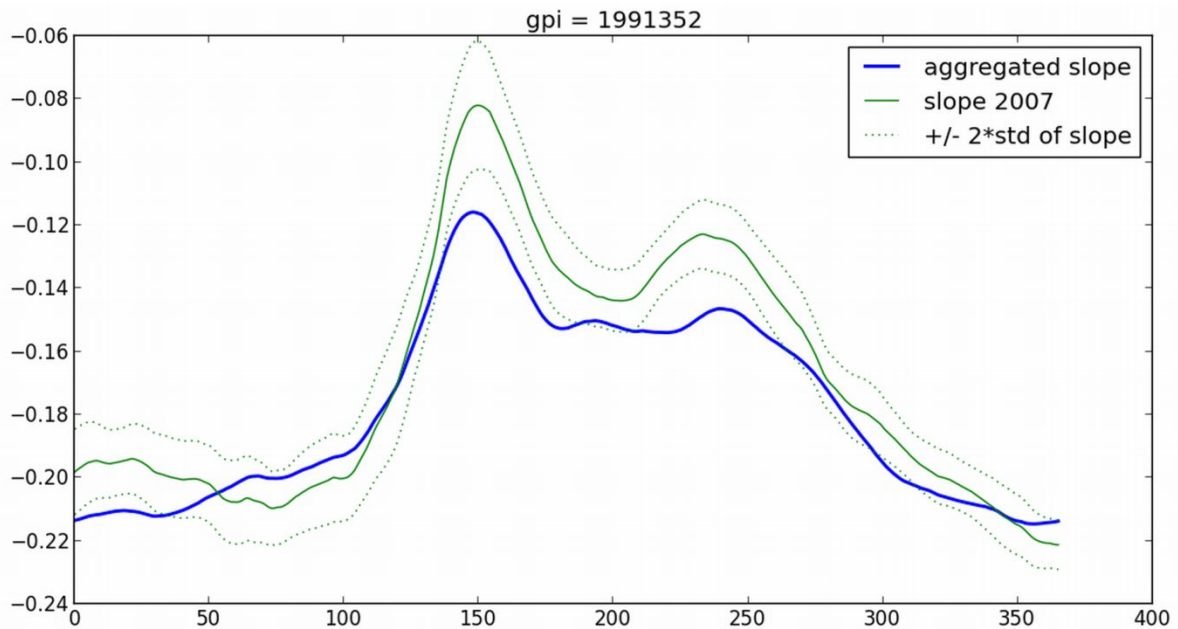
*Figure 26 The blue curve is the aggregated (averaged) slope as function of time (day of year) for an agricultural area in Kansas, USA. The green curve shows the yearly slope for 2007, the dotted green line shows the "confidence band" (+/- 2*standard deviation). The curves differ significantly during several periods, in particular around the two peaks.*
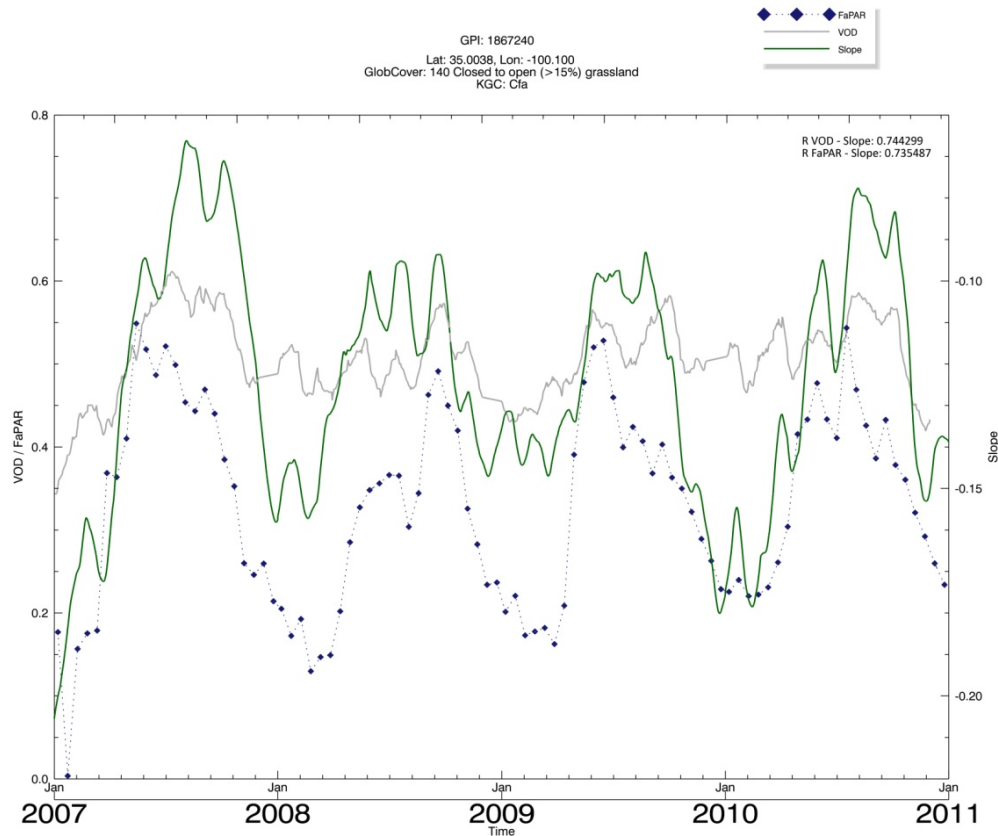
Figure 27 Comparison of the yearly slope, FAPAR and VOD for a site in Texas, USA.

If the inter-annual changes between slopes are as large as indicated by our preliminary experiments, the use of yearly slope values is expected to yield substantial improvements in the vegetation correction step, and consequently the quality of retrieved soil moisture. However, the new algorithm has not been extensively validated yet, and this will have to be done as part of the proposed project.

Up to now, the new approach has been tested with ASCAT only. It should, in principle, also be applicable to ERS, but since ERS has coarser temporal resolution than ASCAT, ERS-based estimates of the slope/curvature parameters will have a higher uncertainty (variance). This could be compensated by increasing the width of the smoothing kernel, which reduces the variance at the cost of an increased estimation bias, so there is a trade-off.

8.5 Masking of Snow and Frost Conditions

Two different approaches were tested to improve the masking of snow and frost conditions. The first called SSFv2 is based on the SSF developed by and already described in the ATBD v.1.0. It is a simplification of the SSF which works through making use of the snow cover and frozen land surface flags (climatological flags) to resolve ambiguities in the backscatter/temperature relationship. Of the existing 3 decision trees the second is no longer necessary in SSFv2 and the other 2 were simplified.
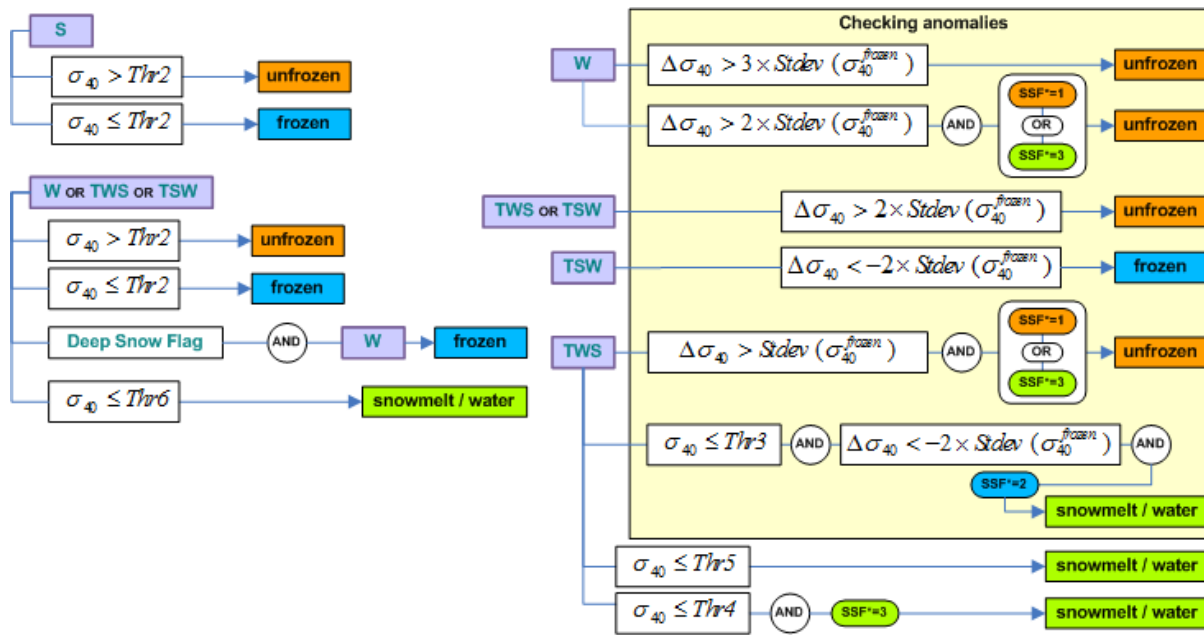
W: winter, S: summer
TWS: transition from winter to summer (60 days between W and S), TSW: transition from summer to winter (60 days between S and W)
SSF*:Previous Surface State Flag 0:unknown, 1:unfrozen, 2:frozen, 3:snowmelt / water on the surface

Figure 28: Simplified Decision Tree 1

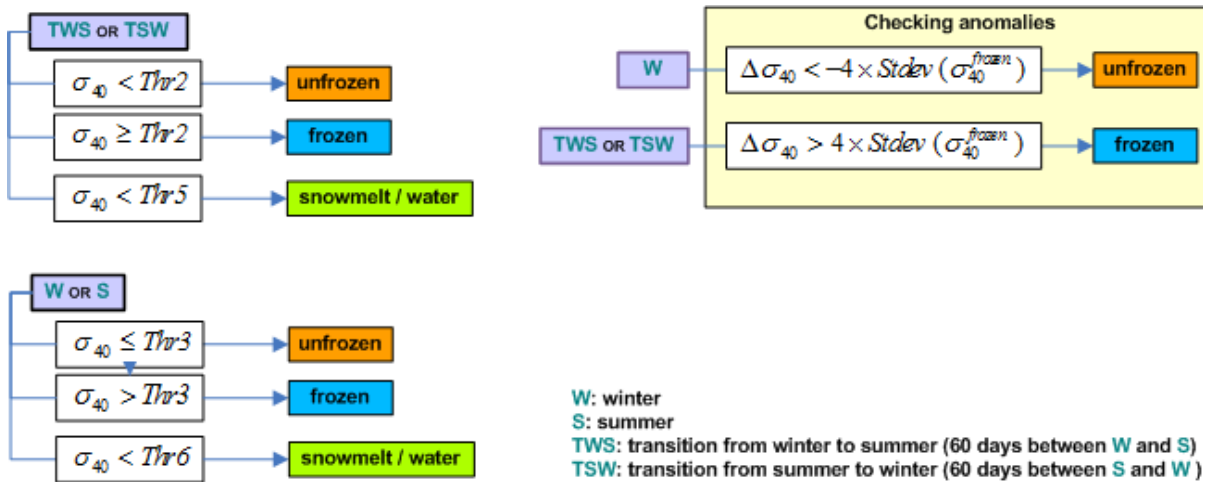*Figure 29: Simplified Decision Tree 3*

The output of the remaining 2 decision trees is checked for consistency by the climatological flags. If the probability of snow or frozen surface is higher than 95% no unfrozen flags are permitted and if they are less than 5% then no frozen flags are allowed.

If the freeze/thaw threshold cannot be defined the surface state is determined through the climatological flags as explained in the ATBD v.1. §6.2.3.7.

The second called HMM FT is based on the probabilistic fusion approach developed by (Zwieback 2012). It uses a Hidden Markov Model whose transition probabilities are determined by ERA-Interim air temperature whereas the emission probabilities are determined by the backscatter. Outputs of this algorithm are not fixed flags but probabilities for the 3 states frozen, unfrozen and thawed. The state with the highest probability is assigned to the observation.

Classifying the uncertainty of a flag describing the surface state is difficult. Nonetheless a quality measure was introduced for both methods. HMM FT produces a probability for the favoured state which can be used directly to indicate the certainty of the flag. In the quality flag this probability is saved as a value between 50 and 100.

For SSFv2 the quality of the masking is done quantitative through assigning one of two classes based on which parts of the algorithm had to be used to find a SSF. If a flag was found only through the decision trees, meaning the backscatter measurements were able to distinguish between frozen and unfrozen, then value 100 is assigned. On the other hand, if it had to be determined with the help of the climatological flags, value 200 is used.

8.5.1 Validation

Since neither global F/T products nor in situ measurements of the freeze/thaw state through e.g. frost tubes were available, temperature had to be used as a proxy for the freeze/thaw state of the soil. To ensure comparability with Naeimi et al. (2012) we used also WMO Air Temperature and GLDAS Surface Temperature from the years 2007 and 2008. ERA Interim data is not used for validation since it is used by the HMM FT algorithm.

Additionally, soil temperature measurements, taken in a depth of up to 5cm, from 14 in situ networks where used. This data was obtained from the ISMN Dorigo et al. (2011). A total of 577 stations had data available in the period 2007-2011. Since several stations used multiple sensors 748 temperature time series of varying length could be used.

The 3 algorithms all produce a flag that indicates either frozen, unfrozen, melting or unknown conditions. For the validation efforts the thawing flag was treated as unfrozen. The freeze/thaw states were compared to the different temperature datasets using a simple threshold at 0°C according to *Table 1*.

Temperature / Flag	Frozen	Unfrozen
< 0°C	Correct (true positive)	Incorrect (false negative)
> 0°C	Incorrect (false positive)	Correct (true negative)

Table 1: Classification of Errors

Table 2 shows that the SSF algorithm has worse agreement than the other two algorithms with both datasets and in all timespans. HMM FT performs better than SSF v2 when compared to the WMO stations except for summer when both show similar agreement. When compared to GLDAS NOAH soil temperature the 2 algorithms show similar agreement in winter, summer and autumn (TSW) but not in spring (TWS). This suggests that the SSF v2 algorithm would need refinement when dealing with thawing situations.

	WMO			GLDAS NOAH		
	SSF	SSFv2	HMM FT	SSF	SSFv2	HMM FT
Winter	77,87	84,12	87,45	89,52	93,77	93,50
TWS	71,67	81,57	85,77	68,72	75,66	82,70
Summer	92,32	98,62	98,64	91,56	97,35	97,88
TSW	73,41	86,71	88,68	74,16	86,21	86,41
Overall	81,93	90,26	92,09	84,24	90,75	91,94

Table 2: Agreement of different algorithms with temperature datasets

Comparison with the in-situ soil temperature data was only considered on a yearly basis and not split up into seasons because the dataset did not have complete temporal coverage for all stations. So splitting it up would possibly introduce a bias towards one algorithm. The

algorithm by Naeimi et al. (2012) depends on a fit of a logistic function into the relationship of temperature and backscatter. *Table 2* shows the agreement with the in situ soil temperature data split into stations where the SSF algorithm worked, meaning that a logistic function could be fit (l) and where it did not work (nl). It also provides values over all stations for SSFv2 and HMM FT.

	SSF(l)	SSFv2(l)	HMM(l)	SSFv2(nl)	HMM(nl)	SSFv2	HMM
flag frozen - T < 0°C	12,76	14,23	12,96	9,97	7,62	11,87	10,01
flag unfrozen - T > 0°C	61,52	71,88	73,76	67,2	74,54	69,29	74,19
flag frozen - T > 0°C	17,66	11,7	9,82	19,64	12,31	16,1	11,2
flag unfrozen - T < 0°C	2,73	2,19	3,46	3,18	5,53	2,74	4,61
Unknown	5,33	0	0	0	0	0	0
Agreement	74,28	86,11	86,72	77,17	82,16	81,16	84,2

Table 3: Agreement in percentage for different algorithms. (l) indicates results over points were a logistic function could be fitted and the SSF algorithm works normally. (nl) are the results for points were no logistic function could be fitted indicating a weak relationship between backscatter and temperature. Results without brackets are the average of all stations.

Real spatial comparison could only be done with the GLDAS NOAH dataset. Figures 26 – 28 show the differences between the agreements to the temperature data. In most regions SSF performs worse than both SSFv2 and HMM FT. The SSF algorithm works very well in northern latitudes where none of the other algorithms provides big improvements. This is not the case for temperate regions where both HMMFT and SSFv2 agree better with GLDAS NOAH soil temperature than SSF. The difference between the 2 new algorithms is not very big except for spring.
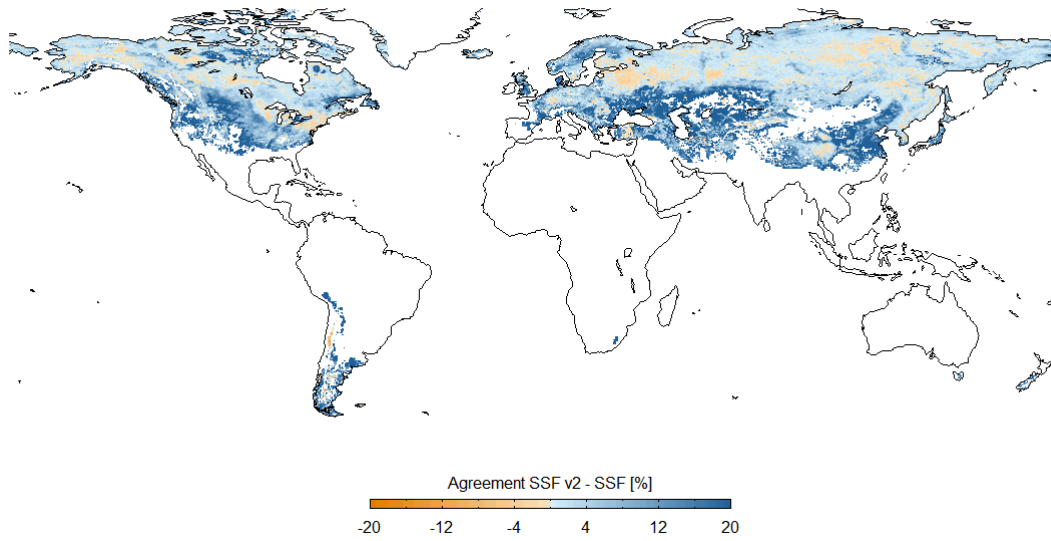
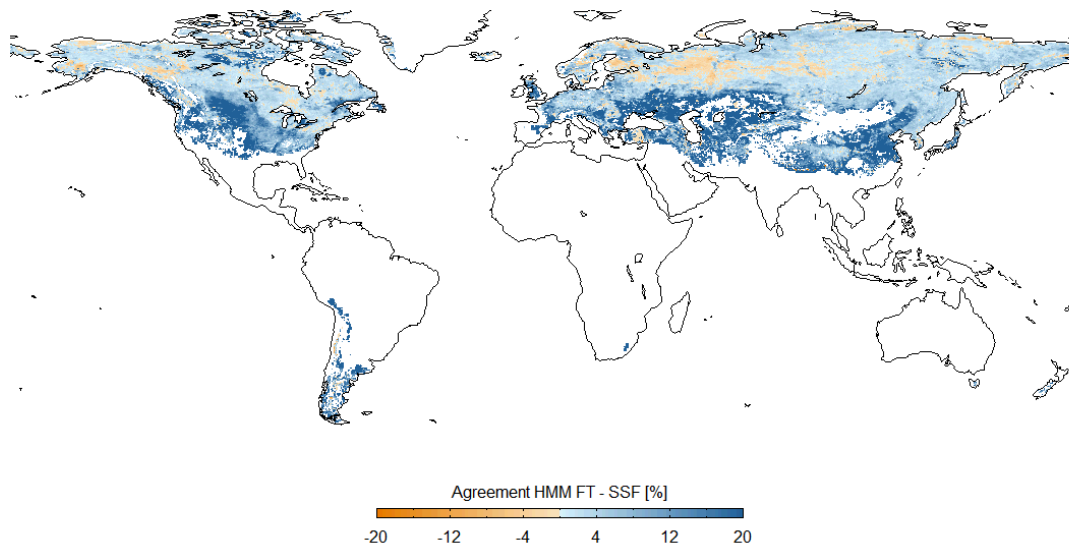*Figure 30: Agreement with GLDAS NOAH soil temperature. SSFv2 – SSF*

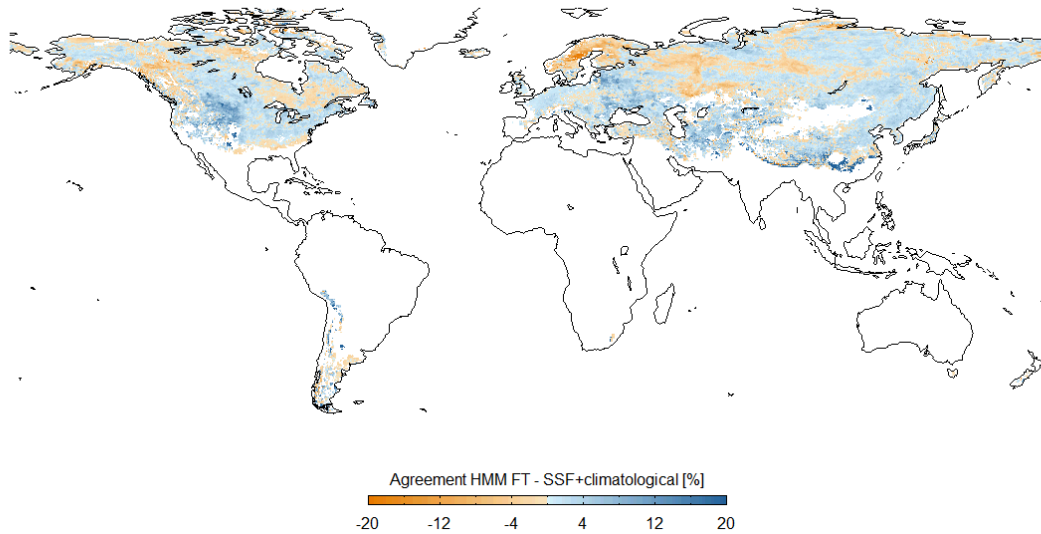*Figure 31: Agreement with GLDAS NOAH soil temperature. HMM FT – SSF*

Figure 32: Agreement with GLDAS NOAH soil temperature. HMM FT - SSFv2

Generally HMM FT is the preferred method since it delivers probabilistic information and not only flags. It is planned to implement this method in the current WARP processing package. This will require rewriting parts of the code and improving the performance of the current implementation. In the meantime SSFv2 will be used since it is already implemented in a way that is compatible to the WARP framework.

8.6 Improved Resampling of Active Data to the DGG in WARP 5.5

The task of resampling is to interpolate L1 scatterometer measurements given in the orbit grid to a fixed Discrete Global Grid (DGG), on which all further WARP-processing takes place. The resampling requires the retrieval of the nearest neighbours (NN) of each orbit point, which is a time consuming operation. In WARP 5.4, NN-candidates within a given radius were computed based on the known properties of the DGG. However, the WARP 5.4 implementation of the NN-search suffers from several problems, namely it

- misses NNs, especially in high latitudes,
- cannot deal with the discontinuity at +180/-180 degrees longitude,
- is relatively slow, and, most important,
- is specific to the DGG, i.e., it can handle no other grids,

whereby the last shortcoming is also the most severe one.

In WARP 5.5, the NN-search has been completely rewritten and is now based on a KD-tree. A KD-tree is a data structure that organises multi-dimensional data in such a way that area searches or NN-searches can be done very efficiently; it can be seen as a generalisation of the

well-known binary search tree to more than one dimension. Because an IDL-implementation would have been too slow, the KD-tree was implemented in C++, and compiled into a DLL which is called from IDL via the IDL-C-bridge. Each point is represented by a four-dimensional feature vector, containing the sine and cosine of its longitude and latitude, respectively. This way, the grid representation becomes continuous, and the data are spread more evenly over the feature space, resulting in better KD-tree performance.

To summarise: the KD-tree based implementation is approximately 20% faster than its WARP 5.4 predecessor, is not plagued by discontinuities and returns all NNs within a given radius (no misses). Also, the approach is completely general, i.e., it works with any grid, not just the DGG.

9 References

Attema, E., & Ulaby, F. (1978). Vegetation modeled as water cloud. *Radio Science, 13*, 357-364

Bartalis, Z. (2009). ERS-ASCAT Backscatter and Soil Moisture Intercomparison - First Results. In, *ASCAT Soil Moisture Working Note*: Institute of Photogrammetry and Remote Sensing, Vienna University of Technology

Bartalis, Z., Scipal, K., & Wagner, W. (2006). Azimuthal anisotropy of scatterometer measurements over land. *IEEE Transactions on Geoscience and Remote Sensing, 44*, 2083-2092

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., & Anderson, C. (2007). Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophy. Res. Lett, 34*, L20401

Cleveland, W.S., & Devlin, S.J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association, 83*, 596

Crow, W.T., Wagner, W., & Naeimi, V. (2010). The impact of radar incidence angle on soil moisture retrieval skill. *IEEE Geoscience and Remote Sensing Letters, 7*, 501-505

Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., & Jackson, T. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences, 15*, 1675-1698

Elyouncha, A., & Neyt, X. (2013a). C-Band Satellite Scatterometer Intercalibration. *IEEE Transactions on Geoscience and Remote Sensing, 51*, 1478-1491

Elyouncha, A., & Neyt, X. (2013b). Inter-calibration of Metop-A and Metop-B scatterometers using ocean measurements. In C.R. Bostater, S.P. Mertikas, & X. Neyt (Eds.) (p. 888806)

Friesen, J., Steele-Dunne, S.C., & van de Giesen, N. (2012). Diurnal Differences in Global ERS Scatterometer Backscatter Observations of the Land Surface. *Geoscience and Remote Sensing, IEEE Transactions on, 50*, 2595-2602

Friesen, J., Winsemius, H.C., Beck, R., Scipal, K., Wagner, W., & De Giesen, N.V. (2007). Spatial and seasonal patterns of diurnal differences in ERS scatterometer soil moisture data in the volta basin, West Africa. In, *IAHS Symposium on Remote Sensing for Environmental Monitoring and Change Detection - 24th General Assembly of the International Union of Geodesy and Geophysics (IUGG)* (pp. 47-55). Perugia, Italy

Fung, A.K. (1994). *Microwave scattering and emission models and their applications.* Boston: Artech House

Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., Mougin, E., Njoku, E., Timouk, F., Wagner, W., & Zribi, M. (2009). Soil moisture active and passive microwave

products: intercomparison and evaluation over a Sahelian site. *Hydrology and Earth System Sciences Discussions, 6*, 5303-5339

Hahn, S., Melzer, T., & Wagner, W. (2012). Error Assessment of the Initial Near Real-Time METOP ASCAT Surface Soil Moisture Product. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 2556-2565

Hsieh, C.-Y., Fung, A.K., Nesti, G., Sieber, A.J., & Coppo, P. (1997). A further study of the IEM surface scattering model. *IEEE Transaction on Geoscience and Remote Sensing, 35*, 901-909

Kerr, Y. (2012). UPDATE on the SMOS products versions. In

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*, 259-263

Lecomte, P., & Wagner, W. (1998). ERS Wind Scatterometer Commissioning and in-flight Calibration. In, *Workshop on Emerging Scatterometer Applications: From Research to Operations* (pp. 261-270). ESTEC, Noordwijk, The Netherlands, 5-7 October 1998: European Space Agency

Long, D.G., & Skouson, G.B. (1996). Calibration of spaceborne scatterometers using tropical rain forests. *IEEE Transactions on Geoscience and Remote Sensing, 34*, 413-424

Melzer, T. (2013). Vegetation Modelling in WARP 6.0. In

Naeimi, V., Bartalis, Z., & Wagner, W. (2008). ASCAT soil moisture: Data quality and consistency with the ERS heritage. *Journal Of Hydrometeorology*, submitted

Naeimi, V., Paulik, C., Bartsch, A., Wagner, W., Kidd, R.A., Park, S.-E., Elger, K., & Boike, J. (2012). ASCAT Surface State Flag (SSF): Extracting Information on Surface Freeze/Thaw Conditions From Backscatter Data Using an Empirical Threshold-Analysis Algorithm. *IEEE Transactions on Geoscience and Remote Sensing, in press*

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., & Wagner, W. (2009). An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. *IEEE Transaction on Geoscience and Remote Sensing, 47*, 1999-2013

Reimer, C. (2014). Calibration of space-borne Scatterometers: Towards a consistent climate data record for Soil Moisture Retrieval. In, *Departement of Geodesy and Geoinformation, Research Group Remote Sensing* (p. 118). Vienna: Vienna University of Technology

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society, 85*, 381-394

Schneeberger, K., Schwank, M., Stamm, C., de Rosnay, P., Matzler, C., & Fluhler, H. (2004). Topsoil structure influencing soil water retrieval by microwave radiometry. *Vadose Zone Journal, 3*, 1169-1179

Talone, M., Crapolicchio, R., De Chiara, G., Neyt, X., Elyouncha, A., De Miguel, L.S., Davies, G., & Bojkov, B. (2012). Cross-calibration of ERS-1 and ERS-2 wind scatterometers; Towards a homogeneous 20-year-long wind vector monitoring of the earth. In (pp. 4606-4609): IEEE

Wagner, W. (1998). Soil Moisture Retrieval from ERS Scatterometer Data. In, *Institute for Photogrammetry and Remote Sensing*. Vienna: Technical University of Vienna

Wagner, W., Lemoine, G., Borgeaud, M., & Rott, H. (1999a). A Study of Vegetation Cover Effects on ERS Scatterometer Data. *IEEE Transactions on Geoscience and Remote Sensing, 37*, 938-948

Wagner, W., Lemoine, G., & Rott, H. (1999b). A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment, 70*, 191-207

Wagner, W., Naeimi, V., Scipal, K., de Jeu, R., & Martinez-Fernandez, J. (2007). Soil moisture from operational meteorological satellites. *Hydrogeology Journal, 15*, 121-131

Wagner, W., Noll, J., Borgeaud, M., & Rott, H. (1999c). Monitoring soil Moisture over the Canadian Prairies with the ERS Scatterometer. *IEEE Trans. Geosci. Rem. Sens., 37*, 206-216

Wagner, W., & Scipal, K. (2000). Large-Scale Soil Moisture Mapping in western Africa using the ERS Scatterometer. *IEEE Trans. Geosci. Rem. Sens., 38*, 1777-1782

Wilson, J.J.W., Anderson, C., Baker, M.A., Bonekamp, H., Figa Saldaña, J., Dyer, R.G., Lerch, J.A., Kayal, G., Gelsthorpe, R.V., Brown, M.A., Schied, E., Schutz-Munz, S., Rostan, F., Pritchard, E.W., Wright, N.G., King, D., & Önel, Ü. (2010). Radiometric calibration of the Advanced Wind Scatterometer Radar ASCAT carried on-board the METOP-A satellite. *IEEE Transaction on Geoscience and Remote Sensing, 48*, 3236-3255

Zwieback, S.a.B., A. and Melzer, T. and Wagner, W. (2012). Probabilistic Fusion of Ku - and C-band Scatterometer Data for Determining the Freeze/Thaw State. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 2583 -2594