EUMETSAT Satellite Application Facility on
Support to Operational Hydrology and Water Management
http://hsaf.meteoam.it/

Algorithm Theoretical Baseline Document (ATBD) Metop ASCAT Soil Moisture CDR and offline products

Climate Data Record (CDR) and offline products

Revision History

Revision	Date	Author(s)	Description
0.1	2013/08/06	S. Hahn	First draft. Adopted content from H25 ATBD.
0.2	2016/01/19	S. Hahn	Update annex, add product table list.
0.3	2016/07/07	S. Hahn	Modifications according to review item discrepancies (RIDs) from the data record readiness review (DRR) 2016.
0.4	2016/10/12	S. Hahn	Modifications according to review from EUMETSAT Secretariat.
0.5	2017/06/20	S. Hahn, T. Melzer	Update H SAF logo, add Climate Data Record (CDR) and offline product description, modified description of flags, update time series example, update NetCDF file example.
0.6	2017/09/15	S. Hahn	Add H SAF URL to title page.
0.7	2018/09/06	S. Hahn	Update sections radiometric calibration, azimuth correction coefficients, soil moisture computation and limitation and caveats.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ASAR Advanced Synthetic Aperture Radar (on Envisat)

ASAR GM ASAR Global Monitoring

ASCAT Advanced Scatterometer

ATBD Algorithm Theoretical Baseline Document

BUFR Binary Universal Form for the Representation of meteorological data

DORIS Doppler Orbitography and Radiopositioning Integrated by Satellite (on Envisat)

ECMWF European Centre for Medium-range Weather Forecasts

Envisat Environmental Satellite

ERS European Remote-sensing Satellite (1 and 2)

ESA European Space Agency

EUM Short for EUMETSAT

EUMETCast EUMETSAT's Broadcast System for Environment Data

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites

FTP File Transfer Protocol

H SAF SAF on Support to Operational Hydrology and Water Management

Météo France National Meteorological Service of France

Metop Meteorological Operational Platform

NRT Near Real-Time

NWP Near Weather Prediction

PRD Product Requirements Document

PUM Product User Manual

PVR Product Validation Report

SAF Satellite Application Facility

SAR Synthetic Aperture Radar

SRTM Shuttle Radar Topography Mission

SZF Sigma Zero Full resolution

SZO Sigma Zero Operational (25 km spatial sampling)

SZR Sigma Zero Research (12.5 km spatial sampling)

TU Wien Technische Universität Wien (Vienna University of Technology)

WARP Soil Water Retrieval Package

WARP H WARP Hydrology

WARP NRT WARP Near Real-Time

ZAMG Zentralanstalt für Meteorologie und Geodynamic (National Meteorological Service of Austria)

1. Executive summary

The Algorithm Theoretical Baseline Document (ATBD) provides a detailed description of the algorithm used to retrieve surface soil moisture information. The so-called TU Wien surface soil moisture retrieval algorithm (TUW-SMR) is a physically motivated change detection method. The first realization of the algorithm was based on ERS-1/2 scatterometer data [4–7]. Later, the same approach was successfully transferred to the Advanced Scatterometer (ASCAT) onboard the series of Metop satellites [8–10]. The TUW-SMR is implemented within a software package called soil WAter Retrieval Package (WARP). WARP is implemented in the programming language Python.

2. Introduction

2.1. 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 Metop ASCAT soil moisture Climate Data Records (CDR).

2.2. Targeted audience

This document mainly targets:

1. Remote sensing experts interested in the retrieval and error characterization of satellite soil moisture 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.

2.3. Related H SAF surface soil moisture products

In the framework of the H SAF project several soil moisture products, with different timeliness (e.g. NRT, offline, data records), spatial resolution, format (e.g. time series, swath orbit geometry) or the representation of the water content in various soil layers (e.g. surface, root-zone), are generated on a regular basis and distributed to users. A list of all available soil moisture products, as well as other H SAF products (such as precipitation or snow) can be looked up on the H SAF website[1]. The following Table 2.1 gives an overview of the surface soil moisture CDR and CDR extensions products, which are related to this ATBD.

3. ASCAT on-board Metop

3.1. Introduction

The European contribution within the framework of the Initial Joint Polar System (IJPS) is the EUMETSAT Polar System (EPS). The space segment of the EPS programme is envisaged to contain three sun-synchronous Meteorological Operational Platforms (Metop-A, Metop-B and

[1] http://hsaf.meteoam.it/.

Table 2.1: List of Surface Soil Moisture Climate Data Reocrds (SSM CDR) and their extensions (SSM CDR-EXT) related to this ATBD.

H<id> CDR product name	H<id> CDR extension (EXT) product name
H25 Metop ASCAT SSM CDR2014	H108 Metop ASCAT SSM CDR2014-EXT
H109 Metop ASCAT SSM CDR2015	H110 Metop ASCAT SSM CDR2015-EXT
H111 Metop ASCAT SSM CDR2016	H112 Metop ASCAT SSM CDR2016-EXT
H113 Metop ASCAT SSM CDR2017	H114 Metop ASCAT SSM CDR2017-EXT
H115 Metop ASCAT SSM CDR2018	H116 Metop ASCAT SSM CDR2018-EXT
H117 Metop ASCAT SSM CDR2019	H118 Metop ASCAT SSM CDR2019-EXT
H119 Metop ASCAT SSM CDR2020	H120 Metop ASCAT SSM CDR2020-EXT
H121 Metop ASCAT SSM CDR2021	-

Metop-C) jointly developed by ESA and EUMETSAT. Each satellite has a nominal lifetime in orbit of about 5 years, with a planned 6 month overlap between consecutive satellites. The first satellite Metop-A, was launched on 19 October 2006, Metop-B was launched on 17 September 2012. Metop-C is planned to be launched approximately in 2018. The Advanced Scatterometer (ASCAT) is one of the instruments carried on-board the series of Metop satellites.

3.2. Instrument description

ASCAT is a real aperture radar system operating in C-band (VV polarization) and provides day- and night-time measurement capability, which is almost unaffected by cloud cover. The instrument measures the Normalized Radar Cross Section (NRCS), or so-called backscatter coefficient σ° expressed in m^2m^{-2} or dB, from the Earth's surface. The design and performance specification is based on the experience of the Scatterometer flown on the ERS-1 and ERS-2 satellite. It can basically work in two different modes: Measurement or Calibration. The Measurement Mode is the only mode that generates science data for users. The instrument uses a pulse compression method called "chirp", at which long pulses with linear frequency modulation were generated at a carrier frequency of 5.255 GHz. After receiving and de-chirping of the ground echoes, a Fourier transform is applied in order to relate different frequencies in the signal to slant range distances. A pre-processing of the noise and echo measurements is done already on board to reduce the data rate to the ground stations, e.g. various averaging takes place reducing the raw data by a factor of approximately 25. Within each pulse repetition interval, after all pulse echoes have been decayed, the contribution of the thermal noise is monitored in order to perform a measurement noise subtraction during the ground processing. A side effect of on board processing is a degree of spatial correlation between different and within received echoes, but this is taken into account later on by the Level 1b processing. The radar pulse repetition frequency (PRF) is approximately 28.26 Hz, which yields to 4.71 Hz for the beam pulse repetition frequency in the sequence fore-, mid- and aft beam [11].

The Calibration Mode is used during external calibration campaigns (29 days every 13 months), when the platform passes over three different ground transponders located in central Turkey. The objective of such calibration campaigns is to verify that the backscatter measurements from a target is correct for the whole incidence angle range. In other words, the absolute and relative calibration will be monitored and checked. A stable radar cross section and an accurately known

position on the Earth's surface are the basic and important requirements for each transponder. After receiving a pulse from ASCAT, the active transponder send a delayed pulse back, which in turn can be recorded by ASCAT. A comparison between the localised transponder position and the expected data allows to estimate three depointing angles and gain correction values for each antenna. An east-west distribution with a displacement of approximately 150 km, will allow to obtain transponder measurements from most of the incidence angles for each antenna beam. This high sampling is necessary in order to reconstruct the characteristics of each antenna gain pattern for the whole incidence angle range and achieve an appropriate relative calibration. Such a calibration set up allows to establish a reference system in order to evaluate and monitor the instrument performance on regular basis [12].

3.3. Geometry and coverage

The Metop satellites are flying in a sun-synchronous orbit (29 day repeat cycle) with an ascending/descending node at 9:30 p.m./a.m. and a minimum orbit height of 822 km. This implies that the satellite will make a little more than 14 orbits per day. The advanced measurement geometry of ASCAT allows twice the coverage, compared to its predecessors flying on ERS-1 and ERS-2. Thus, the daily global coverage increased from 41% to 82%. The spatial resolution was also improved from 50 km to 25-34 km, whilst maintaining the same radiometric accuracy compared to the ERS-1/2 scatterometer.

The advanced global coverage capability is accomplished by two sets of three fan-beam antennas, compared to only one set at the ERS satellites. The fan-beams are arranged broadside and $\pm 45°$ of broadside, thus, allowing to observe three azimuthal directions in each of its two 550 km swaths (see Figure 3.1). The swaths are separated from the satellite ground track by about 360 km for the minimum orbit height.

Each point on the Earth's surface that falls within one of the two swaths, will be seen by all three antennas and a so-called $sigma°$ triplet can be observed. The incidence angles for those two antennas which are perpendicular to the flight direction intersect the surface between 25° and 53.3°, whereas the other four antennas have an incidence angle range from 33.7° to 64.5° [1]. The multi-angle viewing capability of ASCAT is needed for an effective and precise geophysical retrieval of surface parameters (e.g. wind speed and direction over ocean, vegetation conditions over land). Otherwise remaining ambiguities hamper or render impossible a successful retrieval of such parameters.

3.4. Product processing

Global data products are often classified into different levels according to their processing progress. In case of ASCAT the different product levels can be summarized as follows:

- Level 0: Unprocessed raw instrument data from the spacecraft. These are transmitted to the ground stations in binary form.
- Level 1a: Reformatted raw data together with already computed supplementary information (radiometric and geometric calibration) for the subsequent processing.
- Level 1b: Calibrated, georeferenced and quality controlled backscatter coefficients in full resolution or spatially averaged. It includes also ancillary engineering and auxiliary data.

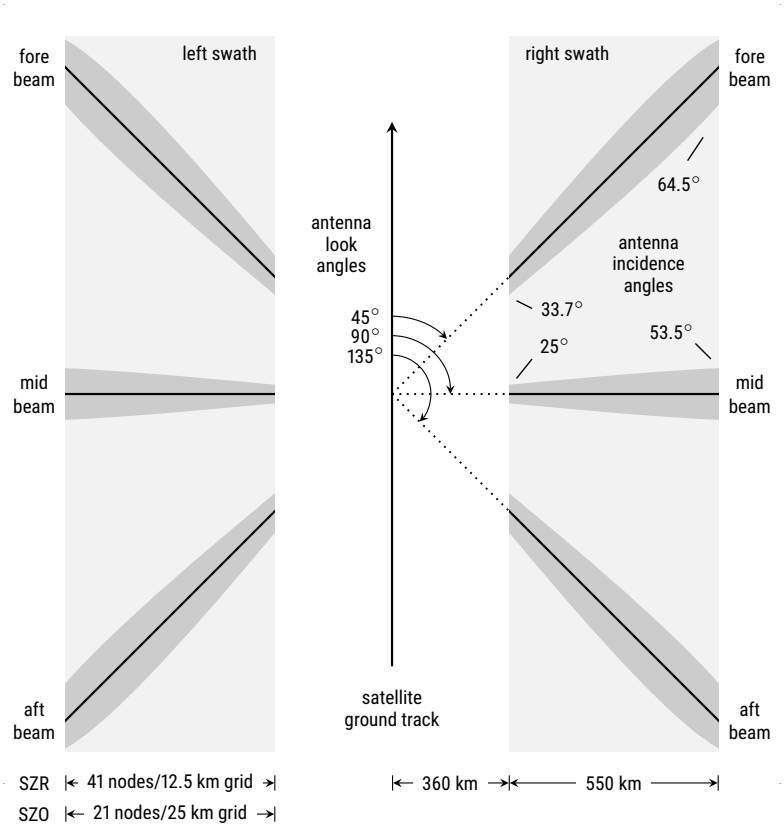

Figure 3.1: ASCAT swath geometry for a Metop minimum orbit height (822 km). The dimensions are symmetric with respect to the satellite ground track [1].

- Level 2: Geo-referenced measurements converted to geophysical parameters, at the same spatial and temporal resolution as the Level 1b data.
- Level 3: Geophysical products derived from Level 2 data, which are either re-sampled or gridded points.

The raw data arriving at the ground processing facility are passed into the Level 1 ground processor (see Figure 3.2). This data driven processing chain generates radiometrically calibrated Level 1b backscatter coefficients. It also includes an external calibration processing chain in order to support the localization and normalisation process of the measurement power echoes. Before satellite launch, it is only possible to estimate parameters characterising the expected instrument performance, since some of them depend on in-flight conditions. But once the satellite is his orbit, these parameters can be assessed during a calibration campaign. The first intermediate product generated just before the swath node generation and spatial averaging steps, is called Level 1b full resolution product. The main geophysical parameter in this product is the normalized backscatter coefficient σ°. Along each antenna beam 256 σ° values are projected on the Earth's surface. The footprint size is about 10 $\times$ 20 km of various shapes and orientations, depending on the Doppler pattern over the surface [13]. The radiometric accuracy and the inter-beam radiometric stability is expected to be less then 0.5 dB peak-to-peak, whereas the georeferencing

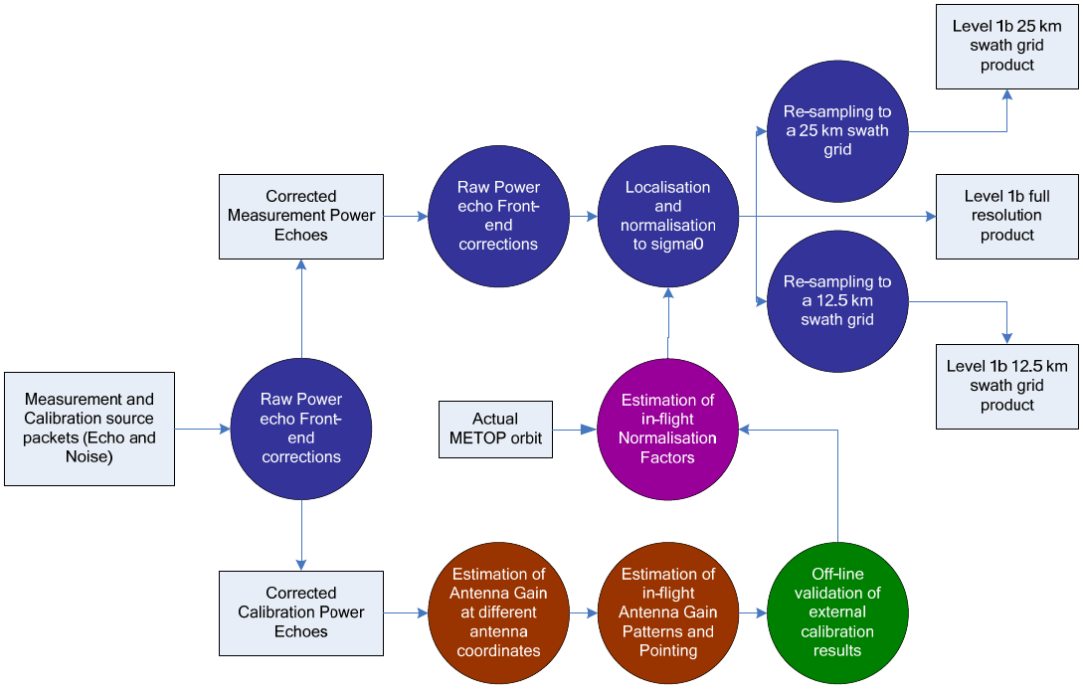

Figure 3.2: Functional overview of the ASCAT Level 1 ground processing chain [2].

accuracy is about 4 km. In addition a number of quality flags are computed associated with every individual σ° sample along each antenna beam.

The other two products generated within the Level 1b processing are averaged σ° values at two different spatial grids. A two-dimensional Hamming window function centered at every grid node is used in both cases for spatial filtering. This weighting function is based on a cosine function, which will basically attenuate the contribution of values with increasing distance. The window width determines which observations will contribute to the weighted result. Observations beyond the window width are disregarded. The spatial resolution is defined as twice the distance from the center to where the window function reaches 50% of its peak intensity (full width at half maximum, FWHM). This spatial averaging is used in the along- and across-track direction, with the objective to generate a set of σ° triplets for each grid node of each swath at the desired radiometric resolution. The first product generated with this approach is called the nominal product and has a spatial resolution of 50 km for each grid point. The grid spacing is 25 km with 21 nodes at each swath per line. The higher resolution product has a spatial resolution ranging from 25 to 34 km. This variability is the result of a trade-off between spatial resolution and the desired radiometric accuracy, which lead to an alternating Hamming window width across the swath for the different beams. In this case the grid spacing is only 12.5 km, with 41 nodes at each swath per line.

4. Soil moisture retrieval algorithm

The TU Wien soil moisture retrieval algorithm (TUW-SMR) 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 an iterative adjustment process. As a result, it is also quite straight forward to perform an error propagation estimating the retrieval error for each location on the land surface [9]. However, a disadvantage of TUW-SMR is the lumped representation of the measurement process. 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 radiative transfer modelling approaches. It also means that it is necessary to calibrate its model parameters using long backscatter time series to implicitly account for land cover, surface roughness, and other effects.

4.1. Model assumptions and requirements

The basic assumptions of the TU Wien soil moisture retrieval algorithm (TUW-SMR) are:

1. The relationship between the backscattering coefficient σ° expressed in decibels (dB) and the surface soil moisture content is linear.
2. The backscattering coefficient σ° depends strongly on the incidence angle θ. The exact functional form of the dependency is mainly determined by surface roughness and land cover, but is not affected by changes in the soil moisture content.
3. At the spatial scale of the scatterometer measurements (25-50 km) 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 on the incidence angle, the curve $\sigma^\circ(\theta)$ 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_d and θ_w) where the backscattering coefficient is stable (in the sense that it is not affected by seasonal changes in above ground vegetation biomass) for dry and wet conditions.
5. Vegetation phenology influences backscatter σ° on a seasonal scale. Local short-term fluctuations are suppressed at the scale of the scatterometer measurements.

In order to be able to apply TUW-SMR the following requirements needs to be fulfilled:

1. Regular temporal revisit time of 1-3 days,
2. Multi-angle backscatter measurements (approx. 20°-60°) from 2-3 beams,
3. Sufficiently long time period ($>$ 2 years).

If all of these requirements are met, it is possible to derive robust model parameters, such as the characterization of the incidence angle dependency of backscatter or the estimation of the dry and wet reference.

4.2. TU Wien soil moisture retrieval algorithm (TUW-SMR)

The following subsections describe the individual steps in the TU Wien soil moisture retrieval algorithm (TUW-SMR).

4.2.1. Re-sampling to a Discrete Global Grid (DGG)

The original Level 1b backscatter measurements are provided in orbit geometry, i.e. the instrument acquisition geometry (see Figure 3.1). In principle, all Level 1b backscatter products can be re-sampled on a fixed Earth grid, but the Level 1b SZR product is used as input for the computation of the soil moisture data records. Level 1b backscatter measurements are re-sampled on a fixed Earth grid creating a time series for each defined grid point. The fixed Earth grid used for this purpose is called WARP 5 grid and has been developed by TU Wien. The WARP 5 represents a Discrete Global Grid (DGG) and contains 3264391 grid points (GP) with an equal spacing of 12.5 km in longitude and latitude. Each GP has a unique grid point identifier (GPI). The WARP 5 is discussed in more detail in [14].

A Hamming window function is used to compute the weighted average of the nearest backscatter observations, defined as

$$w\left(x\right) = 0.54 + 0.46 \cdot \cos\left(2\pi \cdot \frac{\delta x}{d}\right) \tag{1}$$

whereby δx denotes the distance between the actual GP on the fixed Earth grid and the orbit grid points within a search radius 18 km (i.e. $d = 36$ km). The Hamming window has been chosen for interpolation, because it is also used in the creation of the spatial averaged Level 1b backscatter product. The result of the re-sampling step is a time series for each GP containing a list of time stamps represented as Julian dates, backscatter observation for each beam σ_b°, the incidence angle of each acquisition θ_b and the azimuth angle of each acquisition ϕ_b. The subscript $b \in \{f, m, a\}$ is used to distinguish between the fore, mid and aft beam.

4.2.2. Radiometric calibration

The radiometric calibration aims to monitor and correct for potential calibration drifts or biases. It is performed by utilizing natural calibration targets on the Earth's surface with a temporally stable, spatial homogeneous and isotropic backscatter response (e.g. Tropical Rainforests). For this reason, a backscatter calibration model was introduced based on the model from Long and Skouson [15] and adopted for the measurement geometry of Metop ASCAT. The model allows to compute intra-calibration coefficients, which incorporates any arbitrary performance anomalies related to the instrument [3]. In case of a perfectly calibrated instrument, the intra-calibration coefficient vanishes.

The selection of extended area land-targets for the calibration of ASCAT is based on backscatter characteristics indicating spatial, azimuthal and temporal (long-term) stability. Targets with a temporal stable backscatter coefficient are assumed to be unaffected by time-depended geophysical processes controlling the scattering properties of the target. The temporal long-term variability is determined by computing the standard deviation of the incidence-angle-normalized backscatter time series and the azimuthal anisotropy by comparing backscatter characteristics from all azimuth directions. Grid points with less than 0.4 dB temporal variability and 0.2 dB azimuthal anisotropy have been selected and used in a mean-shift clustering method to deter-

mine the spatial variability of the complete target area. Targets with less than 0.25 dB are finally used to determine monthly calibration correction parameters.

Analyses of the backscatter characteristics from the natural calibration targets shows that a second-order polynomial centered at 40° incidence angle is able to describe the incidence angle dependency of backscatter (see Eqn. 2). The polynomial coefficients (B_T^p) are determined by an ordinary least square estimation for all beams and their azimuth directions (i.e. Ascending/Descending + Fore/Mid/Aft beam = 12 configurations).

$$\sigma_T^\circ(\Theta) = \sum_{p=0}^{2} B_T^p(40) \cdot (\theta - 40)^p \tag{2}$$

The backscatter calibration reference σ_T° constitutes the time invariant backscatter response of a calibration target T. Deviations from σ_T° are anomaly estimates incorporated in the final calibration coefficients C.

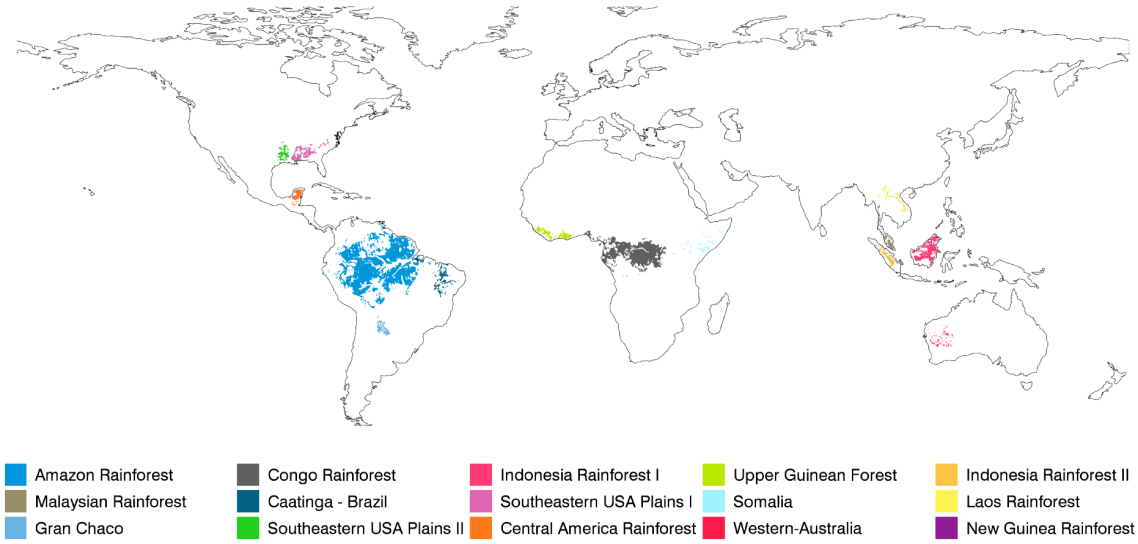

Figure 4.1: Selected backscatter calibration targets [3].

$$C_T(t, \theta, b) = \sigma^\circ(t, \theta, b) - \sigma_T^\circ \tag{3}$$

The target calibration coefficients C_T are derived for each antenna beam (b $\in \{f, m, a\}$) and orbit overpass. Numerous calibration targets (see Figure 4.1) are used for a robust determination of the calibration coefficients C. These coefficients are inferred monthly for each antenna beam by fitting a linear regression model into C_T. The last step for the radiometric calibration consists of subtracting the calibration coefficients C from the actual backscatter measurements in order to get a consistent backscatter time series.

$$\sigma_C^\circ(t, \theta, b) = \sigma^\circ(t, \theta, b) - C(t, \theta, b) \tag{4}$$

4.2.3. Azimuth correction coefficients

The backscattering coefficient changes as a function of the observed surface properties, most importantly surface dielectric properties, roughness and vegetation. However, the overall magnitude also strongly depends on the measurement geometry, i.e. the incidence angle θ as well as the azimuth angle ϕ.

In some dedicated regions azimuth effects on the backscattering coefficient are very strong (e.g. mountainous areas, sandy desert regions). This target anisotropy is accounted for by applying a polynomial correction term to the backscatter signal as suggested by [16]. 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 orbit direction (ascending or descending), resulting in twelve different azimuth configurations (i.e. ϕ_i with $i \in [1, 12]$) for Metop ASCAT. For each of these configurations the incidence angle dependency of backscatter is modeled as a second order polynomial, where θ_r represents the reference incidence angle.

$$p_i := a_i \cdot (\theta - \theta_r)^2 + b_i \cdot (\theta - \theta_r) + c_i \tag{5}$$

The coefficients of these twelve polynomials (a_i, b_i, c_i) are determined by fitting the second order polynomial to all observations falling into the respective configuration category (e.g. fit through all measurements from the left fore beam in ascending direction). Furthermore, an overall model is fitted to all observations, resulting in a total of $3 \times 13 = 39$ coefficients.

$$p_0 := a_0 \cdot (\theta - \theta_r)^2 + b_0 \cdot (\theta - \theta_r) + c_0 \tag{6}$$

The following Equation 7 describes the computation of the azimuth correction coefficients (A_i, B_i, C_i). A difference of the polynomial coefficients between the twelve individual acquisition geometries and the reference (using all measurements) leads to a new polynomial, which is used to determine the correction for the backscatter measurements. This way, the individual incidence angle dependency of backscatter are adjusted to a mean relationship canceling static azimuth effects.

$$\begin{aligned} pc_i :=& (a_i - a_0) \cdot (\theta - \theta_r)^2 + (b_i - b_0) \cdot (\theta - \theta_r) + (c_i - c_0) \\ :=& A_i \cdot (\theta - \theta_r)^2 + B_i \cdot (\theta - \theta_r) + C_i \end{aligned} \tag{7}$$

The static azimuth correction is applied to the backscatter time series for each configuration individually:

$$\hat{\sigma}_i^{\circ}(t) = \sigma_i^{\circ}(t) + A_i * (\theta - \theta_r)^2 + B_i * (\theta - \theta_r) + C_i \tag{8}$$

where $\hat{\sigma}_i^{\circ}(t)$ represents the corrected backscatter time series for each configuration i. All following processing steps are using the corrected backscatter time series as input and formulas are written without using the $\hat{}$ nomenclature.

4.2.4. Estimated standard deviation of backscatter

The Estimated Standard Deviation (ESD) describes the backscatter standard deviation (expressed in dB) representing a measure of noise. The ESD computation is the first step in the

error propagation of the TUW-SMR. The ESD is derived from the fore and aft beam measurements and based on the following observation: all three beams observe the same target and due to the observation geometry the fore and aft beam have the same incidence angle. Thus, as long as there are no azimuth effects, the measurements of the fore and aft beam are comparable, i.e. statistically, they are instances of the same distribution. Hence, the expectation of the difference

$$\delta := \mathrm{E}\left[\sigma_f^\circ - \sigma_a^\circ\right] = 0 \tag{9}$$

should be zero, and its variance should be twice the variance of one of the beams (i.e $\mathrm{Var}\left[\delta\right] = 2 \cdot \mathrm{Var}\left[\sigma^\circ\right]$). This can be derived using error propagation and neglecting higher order terms:

$$\mathrm{Var}\left[\delta\right] \approx \mathrm{Var}\left[\sigma_f^\circ\right] \cdot \left(\frac{\partial \delta}{\partial \sigma_f^\circ}\right)^2 + \mathrm{Var}\left[\sigma_a^\circ\right] \cdot \left(\frac{\partial \delta}{\partial \sigma_a^\circ}\right)^2 + 2 \cdot \mathrm{Cov}(\sigma_f^\circ, \sigma_a^\circ) \cdot \left(\frac{\partial \delta}{\partial \sigma_f^\circ}\right) \cdot \left(\frac{\partial \delta}{\partial \sigma_a^\circ}\right) + \ldots ,$$

whereby the third term is zero under the assumption of i.i.d. (independent and identically distributed random variables).

$$\mathrm{Var}\left[\delta\right] \approx \mathrm{Var}\left[\sigma_f^\circ\right] \cdot (1)^2 + \mathrm{Var}\left[\sigma_a^\circ\right] \cdot (-1)^2 + 0 \tag{10}$$

Assuming equal variances for the fore- and aft-beam we have

$$\mathrm{Var}\left[\delta\right] \approx \mathrm{Var}\left[\sigma_f^\circ\right] + \mathrm{Var}\left[\sigma_a^\circ\right] = 2 \cdot \mathrm{Var}\left[\sigma^\circ\right] \tag{11}$$

The previous equation can be re-written to find the final formula for the ESD

$$\mathrm{ESD}\left[\sigma^\circ\right] = \sqrt{\frac{\mathrm{Var}\left[\delta\right]}{2}} \tag{12}$$

The ESD is computed after strong outliers (defined as $\delta > Q3 + 3 \times \mathrm{IQR}$ or $\delta < Q1 - 3 \times \mathrm{IQR}$) have been removed.

4.2.5. Incidence angle normalization of backscatter

In general, backscatter measurements over land show a strong dependence on the incidence angle and a linear function (defined in the dB domain) is usually deemed sufficient to describe this relationship. The backscatter coefficient at an arbitrary incidence angle σ_θ° (in dB) can be modelled as

$$\sigma_\theta^\circ = \sigma_r^\circ + \sigma_r' \cdot (\theta - \theta_r)\,, \tag{13}$$

where θ_r represents the reference incidence angle (in degree), σ_r° the backscatter at the reference incidence angle (in dB) and σ_r' the slope (in dB/degree). For almost all land cover types this situation is valid, because of the steep decrease of backscatter with increasing incidence angle. However, in case of a considerable amount of volume scattering effects, the signal can also flatten out towards the higher end of incidence angles. In order to account for these effects, a second-order polynomial is used to describe the incidence angle dependency of backscatter in the TUW-SMR adding the curvature σ_r'' as an additional term.

$$\sigma^\circ_\theta = \sigma^\circ_r + \sigma'_r \cdot (\theta - \theta_r) + \frac{1}{2} \cdot \sigma''_r \cdot (\theta - \theta_r)^2 \tag{14}$$

From a mathematical point of view, Eqn. 14 represents a Taylor polynomial, since the function is based on the finite sum of values of the function's derivatives at a single point. Therefore, the reference incidence angle θ_r can be regarded as expansion point and the first σ'_r (dB/degree) and second derivative σ''_r (dB/degree^2) of σ°_r are referred to slope and curvature, respectively. Once these parameters are known, the Taylor polynomial can be used to estimate backscatter in the proximity of θ_r. Reversing Eqn. 14 allows to compute σ°_r from backscatter taken at an arbitrary incidence angle σ°_θ, which is also called incidence angle normalization of backscatter (Eqn. 15). In order to perform the incidence angle normalization of backscatter the slope σ'_r and curvature σ''_r parameter needs to be known.

$$\sigma^\circ_r = \sigma^\circ_\theta - \sigma'_r \cdot (\theta - \theta_r) - \frac{1}{2} \cdot \sigma''_r \cdot (\theta - \theta_r)^2 \tag{15}$$

The methodology to estimate σ'_r and σ''_r has evolved over time, but each approach is based on the local slope computation. A local slope represents an instantaneous characterization of the incidence angle dependency of backscatter, which can be estimated from the differences $\sigma^\circ_f - \sigma^\circ_m$ and $\sigma^\circ_a - \sigma^\circ_m$. This is possible because for a given measurement triplet, the fore- and aft-beam share the same incidence angle (i.e. $\theta_f = \theta_a$):

$$\sigma'_{fm} = \frac{\sigma^\circ_m - \sigma^\circ_f}{\theta_m - \theta_f}, \quad \sigma'_{am} = \frac{\sigma^\circ_m - \sigma^\circ_a}{\theta_m - \theta_a} \tag{16}$$

σ'_{fm} and σ'_{am} are estimates of the slope value at the midpoint local incidence angles

$$\theta_{fm} = \frac{\theta_m + \theta_f}{2}, \quad \theta_{am} = \frac{\theta_m + \theta_a}{2} \tag{17}$$

If it is possible to generate a large sample of local slope values evenly distributed over the entire incidence angle range, a first order approximation of σ'_θ can be used to determine σ'_r and σ''_r derived from a linear regression model.

$$\sigma'_\theta = \sigma'_r + \sigma''_r \cdot (\theta - \theta_r) \tag{18}$$

The selection and weighting of the local slope values in the regression model has changed over time in the TUW-SMR. Data density constraints and computational considerations have lead to alternative implementations. An overview of these methods can be found in [17]. The most recent method estimating σ'_r and σ''_r is based on a Kernel Smoother (KS) [18]. The KS approach offers the possibility to estimate σ'_r and σ''_r either as a climatology (i.e. 366 values, similar to previous methods) or as a time series (i.e. daily values). However, for the moment the computation of σ'_r and σ''_r is based on the climatology and the time series approach is under investigation.

In case of the climatology, the regression coefficients are estimated for each day of year d by selecting local slope values from all available years in close proximity (i.e. ± 21 days or $\lambda = 42$ days) for the day under question d_0. Within the time window, the local slope values are weighted according to their temporal distance using the Epanechnikov kernel (see [19] and Figure 4.2, chapter 6)):

$$k(d_0, d) = D\left(\frac{d - d_0}{\lambda}\right) \tag{19}$$

with

$$D(t) = \begin{cases} \frac{3}{4} \cdot (1 - t^2) & \text{if } |t| \leq 1 \\ 0 & \text{else} \end{cases} \tag{20}$$

The Epanechnikov kernel integrates to one and has finite support in range $[-\lambda, \lambda]$. The kernel width λ is a crucial parameter that determines how strongly neighboring measurements influence the final estimate: an increase in λ decreases the variance, but, at the same time, increases the bias of the estimate.

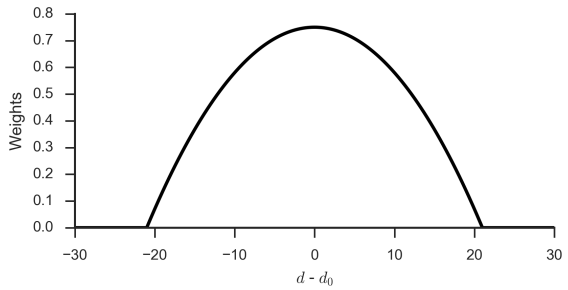

Figure 4.2: Example of the Epanechnikov kernel ($\lambda = 21$).

Normally, the kernel is a function of the independent variable (in our case θ), but since it is the goal to find the coefficients of the linear approximation (Eqn. 18) over the whole incidence angle range at a given day d_0, the kernel is a function of the day of the year. All local slope values in-between the interval $|(d - d_0)/\lambda| \leq 1$ are weighted according to their distance to the current day of year d_0 and take part in the linear regression (Eqn. 21).

Let $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^N$ denote the vectors containing θ_l and σ'_l, respectively, within the support of the kernel for day d_0. Let further $\mathbf{A} \in \mathbb{R}^{N \times 2}$ be the design matrix, whose first column is all 1, and whose second column is given by $\mathbf{x}$, and $\mathbf{W}(d_0) \in \mathbb{R}^{N \times N}$ a diagonal weight matrix, whose i-th diagonal entry is given by the value of the kernel for the i-th local slope, i.e. $\mathbf{W}(d_0)[i, i] = k(d_0, d_i)$. Then, the estimates for σ'_r and σ''_r at day d_0 are obtained as

$$\begin{pmatrix} \sigma'_r(d_0) \\ \sigma''_r(d_0) \end{pmatrix} = (\mathbf{A}^T \mathbf{W}(d_0) \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}(d_0) \mathbf{y} = \mathbf{B}\mathbf{y} \tag{21}$$

If the error covariance of the local slope values used in the fit is given by $\mathbf{\Sigma_l}$, then the error covariance of the parameters is given by

$$\mathbf{\Sigma}_{\sigma'_r, \sigma''_r} = \mathbf{B} \mathbf{\Sigma_l} \mathbf{B}^T \tag{22}$$

At the moment, it is assumed that $\mathbf{\Sigma_l}$ is homeoscedastic, i.e., the errors for all local slope values have the same variance s_l^2 and are uncorrelated: $\mathbf{\Sigma_l} = \mathbf{I} s_l^2$, and estimate s_l^2 from the residuals of the fit (in order to avoid confusion with σ for the representation of backscatter, the variance notation is s^2).

Once σ'_r and σ''_r have been computed, the incidence angle normalization can be applied to the re-sampled backscatter measurements using Eqn. 15.

The simple mathematical description of the incidence angle dependency of backscatter allows a straight forward application of error propagation. In order to estimate the error of σ_r°, only the measurement noise of σ_θ°, as well as noise estimates of σ_r' and σ_r'' are needed. Letting

$$\boldsymbol{x} = \left[\sigma_{\theta,b}^\circ, \sigma_r', \sigma_r''\right] \tag{23}$$

we can formulate the normalised backscatter as function of the observed backscatter and the slope and curvature parameters:

$$\sigma_{r,b}^\circ = f\left(\boldsymbol{x}\right) = \sigma_{\theta,b}^\circ - \sigma_r' \cdot \Delta\theta_b - \frac{1}{2} \cdot \sigma_r'' \cdot \left(\Delta\theta_b\right)^2 \tag{24}$$

with $\Delta\theta_b = \theta_b - \theta_r$.

If we assume that the errors of the normalized backscatter, slope and curvature are uncorrelated, the covariance matrix of $\boldsymbol{x}$ is simply

$$\mathrm{Cov}_x = \begin{pmatrix} \mathrm{ESD}\left[\sigma^\circ\right]^2 & 0 & 0 \\ 0 & \mathrm{Var}\left[\sigma_r'\right] & 0 \\ 0 & 0 & \mathrm{Var}\left[\sigma_r''\right] \end{pmatrix} \tag{25}$$

The Jacobian J of $f\left(\boldsymbol{x}\right)$ is obtained as

$$J = \left(\frac{\delta f}{\delta \boldsymbol{x}}\right) = \left[1, -\Delta\theta_b, -\frac{1}{2} \cdot \left(\Delta\theta_b\right)^2\right] \tag{26}$$

Thus, the noise variance of the normalized backscatter for each beam $b \in \{f, m, a\}$ is

$$\mathrm{Var}\left[\sigma_r^\circ\right]_b = \mathrm{ESD}\left[\sigma^\circ\right]^2 + \mathrm{Var}\left[\sigma_r'\right] \cdot \left(\Delta\theta_b\right)^2 + \frac{1}{4} \cdot \mathrm{Var}\left[\sigma_r''\right] \cdot \left(\Delta\theta_b\right)^4 \tag{27}$$

The three beams (now having been shifted to a common reference angle) are averaged

$$\bar{\sigma}_r^\circ = \frac{1}{3} \cdot \sum_{b \in \{f,m,a\}} \sigma_{r,b}^\circ \tag{28}$$

The corresponding variance is given by

$$\mathrm{Var}\left[\bar{\sigma}_r^\circ\right] = \frac{1}{9} \cdot \sum_{b \in \{f,m,a\}} \mathrm{Var}\left[\sigma_r^\circ\right]_b \tag{29}$$

As can be seen, averaging over the three beams has the effect that the variance of the noise is lowered by a factor of three. It does, however, not lower the error due to the lack of fit of the slope model [5].

The final result of this processing step are the climatology of σ_r' and σ_r'' and the normalized backscatter time series $\bar{\sigma}_r^\circ$ and its noise $\mathrm{Var}\left[\bar{\sigma}_r^\circ\right]$.

4.2.6. Dry and wet reference estimation

The estimation of the dry and wet reference is based upon the average of the extreme lowest (driest) and highest (wettest) normalized backscatter measurements that have ever been recorded for each location. These references are used to translate backscatter variations into changes of (relative) surface soil moisture and to correct for multiple signal disturbances, such as vegetation,

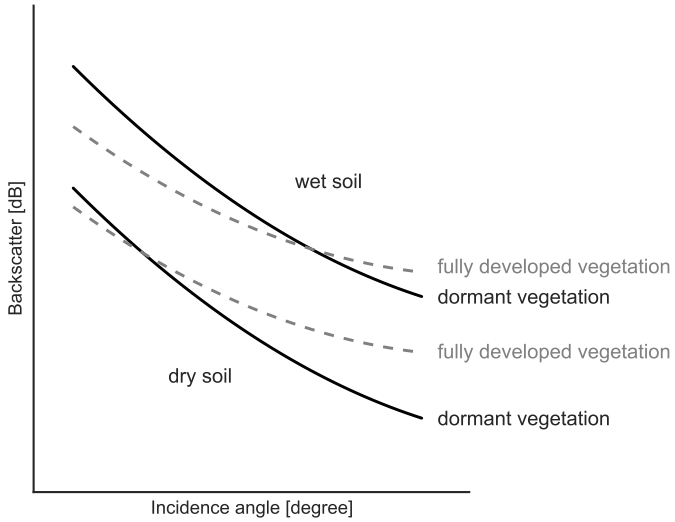

Figure 4.3: Crossover angle concept.

soil roughness and land cover. At the moment, soil roughness and land cover are assumed to be stable and described by a static component in the TUW-SMR. Vegetation, on the other hand, is assumed to remain unchanged from year to year and is therefore characterized by a seasonal component.

The vegetation correction is derived from the temporal variation in the incidence angle dependency of backscatter and based on the following physical observations: In general, the backscattering coefficient rapidly decreases with the incidence angle, with the exception of very rough surfaces [20]. The dominant scattering mechanisms of vegetation are volume scattering in the vegetation canopy and surface scattering from the underlying soil surface [21]. A distinctly different behaviour can normally be observed between volume and surface scattering. Thus, backscatter curves with the same soil moisture content, but with different vegetation states, will intersect at a certain incidence angle. At this cross-over angle vegetation effects are minimized. Figure 4.3 shows this situation for the two extreme cases of complete dry and wet soil. This, so-called cross-over angle concept, leads to the dry θ_d and wet θ_w cross-over angles. These angles are globally set to $\theta_d = 25°$ and $\theta_w = 40°$ and have been experimentally derived from previous studies [5, 6].

In order to determine the lowest/highest backscatter value irrespective of the vegetation conditions, the normalized backscatter measurements are first shifted to the dry/wet crossover angle using Equation 14.

$$\sigma_d^\circ = \bar{\sigma}_r^\circ + \sigma_r' \cdot (\Delta\theta_d) + \frac{1}{2} \cdot \sigma_r'' \cdot (\Delta\theta_d)^2 \tag{30}$$

$$\sigma_w^\circ = \bar{\sigma}_r^\circ + \sigma_r' \cdot (\Delta\theta_w) + \frac{1}{2} \cdot \sigma_r'' \cdot (\Delta\theta_w)^2 \tag{31}$$

with $\Delta\theta_d = (\theta_d - \theta_r)$ and $\Delta\theta_w = (\theta_w - \theta_r)$. Since $\theta_r = \theta_w$, no interpolation is required and Equation (31) simply is

$$\sigma_w^\circ = \bar{\sigma}_r^\circ \tag{32}$$

The noise of σ_d° and σ_w° can be computed using error propagation.

$$\mathrm{Var}\left[\sigma_d^\circ\right] = \mathrm{Var}\left[\bar{\sigma}_r^\circ\right] + \mathrm{Var}\left[\sigma_r'\right] \cdot (\Delta\theta_d)^2 + \frac{1}{4} \cdot \mathrm{Var}\left[\sigma_r''\right] \cdot (\Delta\theta_d)^4 \tag{33}$$

$$\mathrm{Var}\left[\sigma_w^\circ\right] = \mathrm{Var}\left[\bar{\sigma}_r^\circ\right] \tag{34}$$

The 10% lowest/highest backscatter measurements are selected at the dry/wet cross-over angle. Depending on the distribution of the noise, the number of selected measurements will be increased. An uncertainty range, defined as a 95% 2-sided confidence interval, is used to increase the number of low/high backscatter measurements [10]. The definition of the confidence interval considers the noise of dry and wet reference.

$$\text{Confidence Interval (dry)} = \pm 1.96 \cdot \mathrm{Var}\left[\sigma_d^\circ\right] \tag{35}$$

$$\text{Confidence Interval (wet)} = \pm 1.96 \cdot \mathrm{Var}\left[\sigma_w^\circ\right] \tag{36}$$

Once the number of extreme measurements for the lower (N_b) and upper (N_u) part is known, the dry and wet reference at the their cross-over angle are calculated as the mean value.

$$C^d = \frac{1}{N_b} \cdot \sum_{j=1}^{N_b} \sigma_{d,j}^\circ \tag{37}$$

$$C^w = \frac{1}{N_u} \cdot \sum_{j=1}^{N_u} \sigma_{w,j}^\circ \tag{38}$$

The dry and wet reference at their cross-over angles (C^d, C^w) are constant values and can only be compared to backscatter measurements under the same incidence angle. Thus, a normalization to the reference incidence angle θ_r of the dry/wet reference is required.

$$\sigma_r^d = C^d - \sigma_r' \cdot (\theta_d - \theta_r) - \frac{1}{2} \cdot \sigma_r'' \cdot (\theta_d - \theta_r)^2 \tag{39}$$

$$\sigma_r^w = C^w - \sigma_r' \cdot (\theta_w - \theta_r) - \frac{1}{2} \cdot \sigma_r'' \cdot (\theta_w - \theta_r)^2 \tag{40}$$

The noise is given by

$$\mathrm{Var}\left[\sigma_r^d\right] = \mathrm{Var}\left[C^d\right] - \mathrm{Var}\left[\sigma_r'\right] \cdot (\Delta\theta_d)^2 - \frac{1}{4} \cdot \mathrm{Var}\left[\sigma_r''\right] \cdot (\Delta\theta_d)^4 \tag{41}$$

$$\mathrm{Var}\left[\sigma_r^w\right] = \mathrm{Var}\left[C^w\right] - \mathrm{Var}\left[\sigma_r'\right] \cdot (\Delta\theta_w)^2 - \frac{1}{4} \cdot \mathrm{Var}\left[\sigma_r''\right] \cdot (\Delta\theta_w)^4 \tag{42}$$

The result of this processing step are the dry and wet reference at the reference incidence angle and their noise estimates.

Dry correction

In the current formulation of the TUW-SMR no dry correction is involved.

Wet correction

In some regions truly saturated conditions are never captured simply due to the prevailing climate. Hence, a correction needs to be applied simulating wet conditions allowing to estimate a real or better wet reference. The application of wet correction is based on an external climate data set [22], since scatterometer measurements alone are not sufficient to locate these regions. The wet correction is done in two steps: first the lowest level of the wet reference is set to -10 dB and subsequently, in regions with rarely saturated soil moisture conditions σ_r^w values are raised until a sensitivity (defined as $\sigma_r^w - \sigma_r^d$) of at least 5 dB has been reached [10].

4.2.7. Surface soil moisture computation

The surface soil moisture computation is based on the assumption of a linear relationship between backscatter and surface soil moisture. Therefore, the normalized backscatter measurements are scaled between the dry and wet reference to obtain relative surface soil moisture expressed in degree of saturation ranging from 0 % (dry soil) and 100 % (saturated soil). The top soil layer is directly exposed to rain, wind, sunlight, etc. and thereby going through rapid changes in the top soil moisture conditions. Furthermore, the penetration depth of the C-band microwave signal into the soil (depending soil surface conditions) is just a few centimeter (1-2 cm), but could also be deeper in very dry soil regimes. Thus, dry or saturated soil refers to the very top surface soil moisture conditions. From a hydrological perspective, soil will never be absolutely dry and reach a minimum moisture level, however the exact amount of residual water content is usually unknown and therefore not accounted for in this retrieval.

$$sm = \frac{\sigma_r^\circ - \sigma_r^d}{\sigma_r^w - \sigma_r^d} \cdot 100 \tag{43}$$

The estimated surface soil moisture represents the topmost soil layer (< 5 cm) and can be converted into (absolute) volumetric units (m^3m^{-3}) with the help of soil porosity information.

$$\Theta = \Theta_r + p \cdot \frac{sm}{100} \tag{44}$$

where Θ is absolute soil moisture in m^3m^{-3}, p is porosity in m^3m^{-3}. If the exact amount of residual water content Θ_r in m^3m^{-3} is known, it can be used to adjust the absolute soil moisture content.

By proceeding with the error propagation, the following noise estimate of surface soil moisture and absolute surface soil moisture can be obtained:

$$\begin{aligned}\mathrm{Var}\left[sm\right] = \mathrm{Var}\left[\bar{\sigma}_r^{\circ}\right]_i \cdot \frac{1}{\left(\sigma_r^w - \sigma_r^d\right)^2} \cdot 100^2 \\ + \mathrm{Var}\left[\sigma_r^d\right]_i \cdot \left(\frac{\bar{\sigma}_r^{\circ} - \sigma_r^w}{\left(\sigma_r^w - \sigma_r^d\right)^2}\right)^2 \cdot 100^2 \\ + \mathrm{Var}\left[\sigma_r^w\right]_i \cdot \left(\frac{\bar{\sigma}_r^{\circ} - \sigma_r^d}{\left(\sigma_r^w - \sigma_r^d\right)^2}\right)^2 \cdot 100^2\end{aligned} \tag{45}$$

$$\mathrm{Var}\left[\Theta\right] = \mathrm{Var}\left[p\right] \cdot \frac{sm}{100} + \mathrm{Var}\left[sm\right] \cdot \frac{p}{100} \tag{46}$$

4.3. Side note on error propagation

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

$$\boldsymbol{\Sigma_y} = \boldsymbol{A}\boldsymbol{\Sigma_x}\boldsymbol{A^T} \tag{47}$$

whereby $\boldsymbol{A^T}$ denotes the transpose of $\boldsymbol{A}$.

If, on the other hand, f is a non-linear mapping, we first linearise it by replacing its first order Taylor approximation about the operation point $\boldsymbol{x_0}$:

$$\boldsymbol{y} = f\left(\boldsymbol{x}\right) \approx f\left(\boldsymbol{x_0}\right) + \left(\frac{\delta f}{\delta \boldsymbol{x}}\right)\left(\boldsymbol{x} - \boldsymbol{x_0}\right) \tag{48}$$

whereby $\left(\frac{\delta f}{\delta \boldsymbol{x}}\right)$ is the Jacobian J of f. Putting everything together we finally obtain

$$\boldsymbol{\Sigma_y} = \left(\frac{\delta f}{\delta \boldsymbol{x}}\right) \boldsymbol{\Sigma_x} \left(\frac{\delta f}{\delta \boldsymbol{x}}\right)^T \tag{49}$$

for the variance of $\boldsymbol{y}$ under the mapping f. Equation 49 is the workhorse of the WARP error propagation scheme.

4.4. Limitations and caveats

The TUW-SMR is applied on a global basis (i.e. for each grid point over land), but in certain situations (dense forest, snow cover, frozen soil, open water or topographic complex area dominating the satellite footprint), the retrieval of meaningful soil moisture values is not always possible . Nonetheless, soil moisture is computed and additional information (e.g. on the soil state) is provided to mask invalid soil moisture measurements. Users are advised to to use the best auxiliary data available to improve the flagging of snow, frozen soil and (temporally) standing water. Processing flags, correction flags, confidence flags and noise estimation are provided in the soil moisture product. It is important to note that the error model is not able to describe all error sources, specifically frozen soil, snow or wetland. In such cases, the noise estimation is not reliable.

Under the following conditions the TUW-SMR should be most suited:

- low to moderate vegetation regimes,
- unfrozen and no snow,
- low to moderate topographic variations
- no wetlands and coastal areas.

4.4.1. Vegetation

The influence of vegetation on the backscatter signal is currently accounted for using a seasonal correction. The so-called cross-over angle concept (section 4.2.6) allows to determine the dry/wet reference at incidence angles where the backscattering coefficient is stable despite seasonal changes in the above ground biomass. The assumption that a climatology is sufficient to represent the inter-annual cycle of vegetation phenology no longer holds true in areas with a high level of variation in seasonal vegetation patterns. Semi-arid regions or areas with large scale disruptions (e.g. fires) can therefore have a decreased quality in the representation of the seasonal cycle of soil moisture. However, short-term anomalies in soil moisture are not affected by a misinterpretation of the seasonal vegetation cycle. The impact of strongly varying inter-annual vegetation cycles has not been quantified and needs to be recognized as a limitation in the current representation of the vegetation correction.

In densely vegetated areas (such as Tropical rainforests) the C-band microwaves are not able to penetrate to the ground. Hence, a soil moisture signal from the soil surface is not contained in the backscatter measurements, which is typically characterized by a very low signal variation (signal sensitivity). However, such cases lead to an increase in the error of the the soil moisture estimate (according to Eq. 45, the rms-error is proportional to the inverse of the squared sensitivity) and can thus be identified.

4.4.2. Desert areas

Desert areas exhibit challenging conditions for the retrieval of surface soil moisture information from radar scatterometry. The backscatter signal strength and dynamics in combination with different scattering mechanisms are the main challenges. These effects can be attributed to the climate and surface conditions leading to different implications (e.g. lack of saturated soil conditions, azimuth anisotropy, high noise, volume scattering, subsurface scattering).

Desert environments are characterized by a shortage of soil moisture due to a limited amount of annual precipitation. As pointed out in section 4.2.6, this leads to the application of the wet correction, since saturated soil conditions are never captured by the scatterometer. In addition, the overall backscatter signal strength and dynamics are very low, especially sandy deserts can have a backscatter coefficient below -20 dB. Such areas are prone to be affected by azimuth anisotropy and Bragg scattering [23]. Latter is related to small scale ripples on dunes enhancing the backscatter signal for a specific range of incidence angles.

In very dry environments soil volume scattering can play an import role. The penetration depth of microwaves can be large (i.e. up to several meters), causing volume scattering and increasing the chance of hitting a subsurface bedrock. The latter can be observed especially in some parts of North Africa leading to an inverse relationship between backscatter and soil

moisture. During drying up the TUW-SMR shows an increase of soil moisture even though the real soil surface is completely dry. The reason is that the backscatter signal is coming from the subsurface when its completely dry and during wet conditions the penetration depth is to short to reach this layer.

Temporal changes in the slope σ' and curvature σ'' signal are assumed to be related to vegetation phenology. However, the local slope computation from low backscatter signal can lead to a systematic noise component wrongly interpreted as vegetation signal. First tests have shown that the computation of constant slope and curvature parameters helps to account for this problem.

4.4.3. Snow

Depending on the physical parameters of the snow layer (i.e. liquid water content, roughness of the air-snow interface, depth and layering of the snow pack, grain size and shape) different scattering mechanisms characterize the overall backscatter signal. Snow scattering phenomena are not treated in the TU Wien model and snow covered periods needs to be masked based on auxiliary information.

4.4.4. Frozen soil

At temperatures below 0°C microwave backscatter drops, because of the inability of the soil water molecules to align themselves to the external electromagnetic field. The effect of freezing is more complex in case of vegetation, because plants have adaptive mechanisms to withstand the cold winter weather. Like for snow, handling backscatter from soil freezing periods is not covered by the TU Wien model and affected time periods needs to be masked based on auxiliary information.

4.4.5. Surface water

Backscatter characteristics are primarily controlled by the roughness of the water surface, which is related to the short penetration depth of C-band microwaves ($<$ 1-2 mm). A calm water surface acts like a mirror scattering almost the complete signal into the forward direction. Waves caused by near surface winds increase the backscatter looking at up- and downwind direction, and reduces the signal normal to the wind direction. As a result, open water can have a disturbing influence on the retrieval of soil moisture, if the area covered is large compared to the overall footprint size. Locations with (temporary) standing water needs to be treated carefully.

4.4.6. Topographic complexity

A high variability of backscatter can be observed in mountainous areas which is not necessarily coupled with soil moisture changes. Hence, the surface topography directly influences the scattering behaviour, which is not considered in the TU Wien model. A reduced soil moisture retrieval performance can be expected in topographic complex areas.

4.4.7. Long-term land cover changes

Long-term land cover changes are not treated in the TU Wien model at the moment. For example, if urbanization or desertification take place on a larger scale, the long-term backscatter

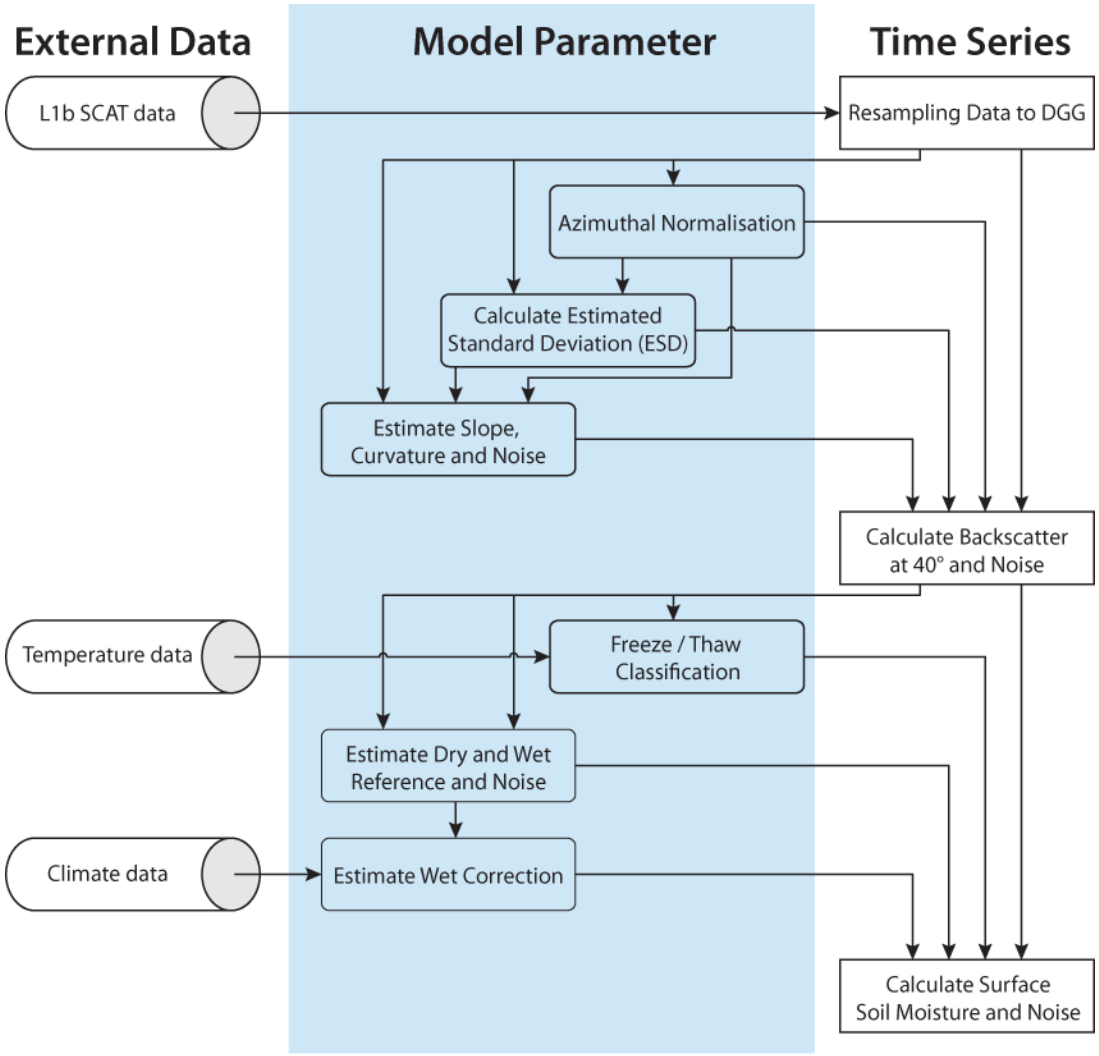

Figure 4.4: Overview of the WARP processing steps.

signal can be influenced.

4.5. Software implementation

The TU Wien soil moisture retrieval algorithm (TUW-SMR) is implemented in a software package called soil WAter Retrieval Package (WARP). The WARP software computes surface soil moisture time series and empirical model parameters. A flowchart of the WARP processing chain is shown in Figure 4.4.

4.6. Processing modes

The WARP software can operate in two modes: CDR production and CDR extension (i.e. offline products extending a CDR). The difference between the two modes is that in the latter case no new empirical model parameters are computed and soil moisture is generated based on pre-existing model parameters.

4.7. Input data and auxiliary information

The input data and auxiliary information is described in detail in the Product User Manual (PUM) [24].

4.8. Output data

The following parameters are part of output data. A more detailed list with data type and data format can be found in the Product User Manual (PUM) [24].

4.8.1. Surface soil moisture and noise

The surface soil moisture estimate represents the topmost soil layer (< 5 cm) and is given in degree of saturation, ranging from 0% (dry) to 100% (wet). Degree of saturation expresses the water volume present in the soil relative to the pore volume and can be converted into (absolute) volumetric units m^3m^{-3} with the help of soil porosity information (see Equation 44)

4.8.2. Surface State Flag (SSF)

The surface state flag (SSF) indicates the surface conditions: unknown, unfrozen, frozen, temporary (snow-)melting/water on the surface or permanent ice. The flag should be used to filter invalid soil moisture observations, since a screening has not been performed in advance. This way, users have full control of the masking and can decide on their own in a borderline case (e.g. during freeze/thaw transition periods). If land surface temperature data is available in the study area, it is recommended to combine this information with the SSF.

The retrieval of SSF is based on a logistic regression function and decision trees using temperature and backscatter data [25]. The SSF represents only the top soil layer and performs best during summer and winter periods. During transition periods and in areas with less frequent freezing the quality of the SSF deteriorates. In the latter case the relationship between negative temperature and backscatter can be no longer accurately modeled.

4.8.3. Processing, correction and confidence flags

The processing, correction and confidence flags indicate various conditions of interest advising the user on the quality and validity of the soil moisture observations. The flags provide an initial assistance on the usability and shall not prevent the usage of external data sets for masking soil moisture observations.

The processing flag explains the reason why a soil moisture value is set to Not a Number (NaN). The correction flag indicates that the soil moisture value has been modified and the confidence flag provides advise on the validity of the soil moisture observations. More information about the flag meanings can be found in the Product User Manual (PUM) [24].

5. References

[1] J. Figa-Saldana, J. J. W. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen, "The Advanced Scatterometer (ASCAT) on the Meteorological Operational (MetOp) Platform: A follow on for European Wind Scatterometers," *Canadian Journal of Remote Sensing*, vol. 28, no. 3, pp. 404–412, 2002.

[2] "ASCAT Product Guide," Tech. Rep. Doc. No: EUM/OPS-EPS/MAN/04/0028, v5, 2015.

[3] C. Reimer, "Calibration of space-borne Scatterometers: Towards a consistent climate data record for Soil Moisture Retrieval," Austria, 2014. [Online]. Available: http://repositum.tuwien.ac.at/urn:nbn:at:at-ubtuw:1-70509

[4] W. Wagner, G. Lemoine, M. Borgeaud, and H. Rott, "A study of vegetation cover effects on ERS scatterometer data," vol. 37, no. 2II, pp. 938–948.

[5] W. Wagner, G. Lemoine, and H. Rott, "A method for estimating soil moisture from ERS scatterometer and soil data," vol. 70, no. 2, pp. 191–207.

[6] W. Wagner, J. Noll, M. Borgeaud, and H. Rott, "Monitoring soil moisture over the canadian prairies with the ERS scatterometer," vol. 37, pp. 206–216.

[7] K. Scipal, W. Wagner, M. Trommler, and K. Naumann, "The global soil moisture archive 1992-2000 from ERS scatterometer data: first results," vol. 3. IEEE, pp. 1399–1401.

[8] Z. Bartalis, W. Wagner, V. Naeimi, S. Hasenauer, K. Scipal, H. Bonekamp, J. Figa, and C. Anderson, "Initial soil moisture retrievals from the METOP-a advanced scatterometer (ASCAT)," vol. 34.

[9] V. Naeimi, "Model improvements and error characterization for global ERS and METOP scatterometer soil moisture data," PhD dissertation.

[10] V. Naeimi, K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner, "An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations," vol. 47, no. 7, pp. 1999–2013.

[11] R. V. Gelsthorpe, E. Schied, and J. J. W. Wilson, "ASCAT - MetOp's Advanced Scatterometer," *ESA Bull. ISSN 0376-4265*, vol. 102, pp. 19–27, 2000.

[12] J. J. W. Wilson, C. Anderson, M. A. Baker, H. Bonekamp, J. F. Saldana, R. G. Dyer, J. A. Lerch, G. Kayal, R. V. Gelsthorpe, M. A. Brown, E. Schied, S. Schutz-Munz, F. Rostan, E. W. Pritchard, N. G. Wright, D. King, and U. Onel, "Radiometric Calibration of the Advanced Wind Scatterometer Radar ASCAT Carried Onboard the METOP-A Satellite," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 3236–3255, 2010.

[13] R. D. Lindsley, C. Anderson, J. Figa-Saldana, and D. G. Long, "A Parameterized ASCAT Measurement Spatial Response Function," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–10, 2016.

[14] "WARP 5 Grid," Tech. Rep. v0.3, 2013.

[15] D. Long and G. Skouson, "Calibration of spaceborne scatterometers using tropical rain forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 2, pp. 413–424, Mar. 1996.

[16] Z. Bartalis, K. Scipal, and W. Wagner, "Azimuthal anisotropy of scatterometer measurements over land," vol. 44, no. 8, pp. 2083–2092.

[17] S. Hahn, C. Reimer, M. Vreugdenhil, T. Melzer, and W. Wagner, "Dynamic Characterization of the Incidence Angle Dependency of Backscatter Using Metop ASCAT," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 3, p. 12, 2016.

[18] T. Melzer, "Vegetation Modelling in WARP 6.0," in *EUMETSAT Meteorological Satellite Conference*, Vienna, Austria, 2013.

[19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning : Data Mining, Inference, and Prediction.* New York: Springer, 2009.

[20] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive. Vol. III - Volume Scattering and Emission Theory, Advanced Systems and Applications.* Artech House, Inc.

[21] A. K. Fung, *Microwave scattering and emission models and their applications.* Artech House.

[22] M. C. Peel, B. L. Finlayson, and T. A. McMahon, "Updated world map of the Köppen-Geiger climate classification," *Hydrology and Earth System Sciences*, vol. 11, no. 5, pp. 1633–1644, oct 2007.

[23] H. Stephen and D. Long, "Microwave backscatter modeling of erg surfaces in the Sahara desert," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 238–247, Feb. 2005.

[24] "Product User Manual (PUM) Soil Moisture Data Records, Metop ASCAT Soil Moisture Time Series," Tech. Rep. Doc. No: SAF/HSAF/CDOP3/PUM, v0.7, 2018.

[25] V. Naeimi, C. Paulik, A. Bartsch, W. Wagner, R. Kidd, S.-E. Park, K. Elger, and J. Boike, "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*, 2012.

Appendices

A. Introduction to H SAF

H SAF is part of the distributed application ground segment of the "European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)". The application ground segment consists of a Central Application Facilities located at EUMETSAT Headquarters, and a network of eight "Satellite Application Facilities (SAFs)", located and managed by EUMETSAT Member States and dedicated to development and operational activities to provide satellite-derived data to support specific user communities (see Figure A.1):

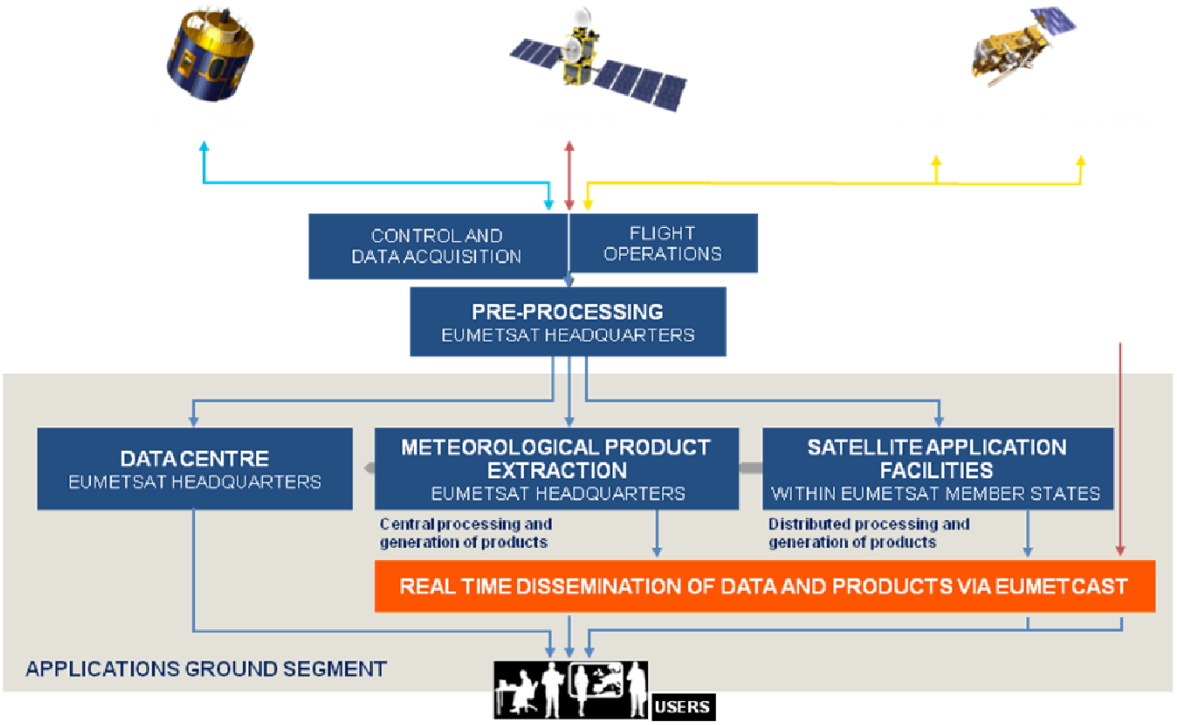

Figure A.1: Conceptual scheme of the EUMETSAT Application Ground Segment.

Figure A.2 here following depicts the composition of the EUMETSAT SAF network, with the indication of each SAF's specific theme and Leading Entity.

B. Purpose of the H SAF

The main objectives of H SAF are:

a) to provide new satellite-derived products from existing and future satellites with sufficient time and space resolution to satisfy the needs of operational hydrology, by generating, centralizing, archiving and disseminating the identified products:

- precipitation (liquid, solid, rate, accumulated);
- soil moisture (at large-scale, at local-scale, at surface, in the roots region);

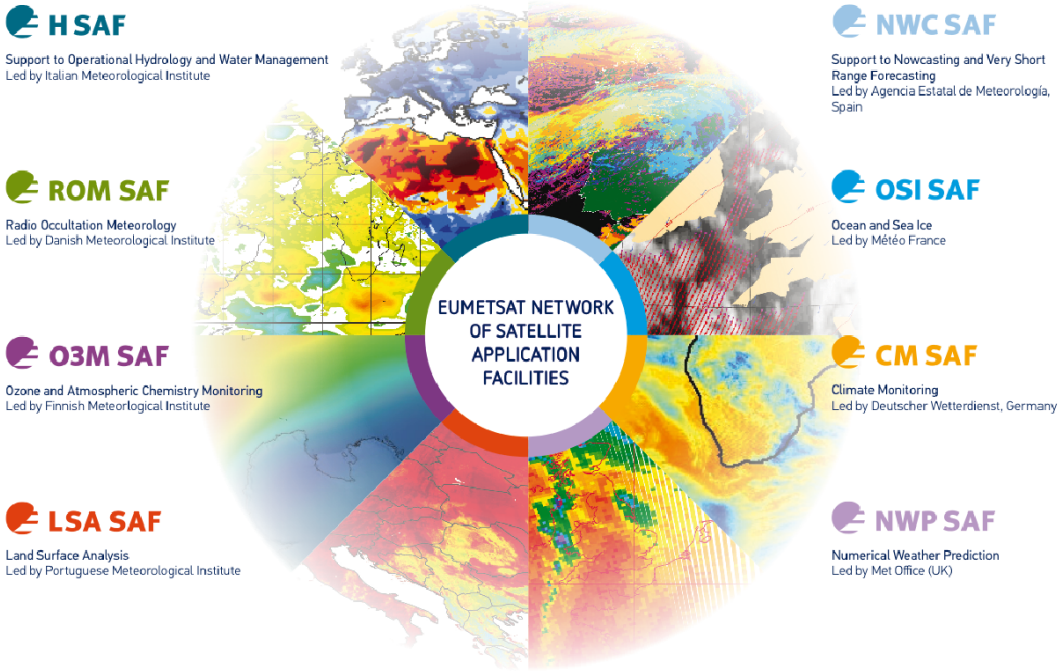

Figure A.2: Current composition of the EUMETSAT SAF Network.

- snow parameters (detection, cover, melting conditions, water equivalent);

b) to perform independent validation of the usefulness of the products for fighting against floods, landslides, avalanches, and evaluating water resources; the activity includes:

- downscaling/upscaling modelling from observed/predicted fields to basin level;
- fusion of satellite-derived measurements with data from radar and raingauge networks;
- assimilation of satellite-derived products in hydrological models;
- assessment of the impact of the new satellite-derived products on hydrological applications.

C. Products / Deliveries of the H SAF

For the full list of the Operational products delivered by H SAF, and for details on their characteristics, please see H SAF website hsaf.meteoam.it. All products are available via EUMETSAT data delivery service (EUMETCast[2]), or via ftp download; they are also published in the H SAF website[3].

All intellectual property rights of the H SAF products belong to EUMETSAT. The use of these products is granted to every interested user, free of charge. If you wish to use these products, EUMETSAT's copyright credit must be shown by displaying the words "copyright (year) EUMETSAT" on each of the products used.

[2] http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/index.html
[3] http://hsaf.meteoam.it

D. System Overview

H SAF is lead by the Italian Air Force Meteorological Service (ITAF MET) and carried on by a consortium of 21 members from 11 countries (see website: hsaf.meteoam.it for details)

Following major areas can be distinguished within the H SAF system context:

- Product generation area
- Central Services area (for data archiving, dissemination, catalogue and any other centralized services)
- Validation services area which includes Quality Monitoring/Assessment and Hydrological Impact Validation.

Products generation area is composed of 5 processing centres physically deployed in 5 different countries; these are:

- for precipitation products: ITAF CNMCA (Italy)
- for soil moisture products: ZAMG (Austria), ECMWF (UK)
- for snow products: TSMS (Turkey), FMI (Finland)

Central area provides systems for archiving and dissemination; located at ITAF CNMCA (Italy), it is interfaced with the production area through a front-end, in charge of product collecting. A central archive is aimed to the maintenance of the H SAF products; it is also located at ITAF CNMCA.

Validation services provided by H SAF consists of:

- Hydrovalidation of the products using models (hydrological impact assessment);
- Product validation (Quality Assessment and Monitoring).

Both services are based on country-specific activities such as impact studies (for hydrological study) or product validation and value assessment. Hydrovalidation service is coordinated by IMWM (Poland), whilst Quality Assessment and Monitoring service is coordinated by DPC (Italy): The Services activities are performed by experts from the national meteorological and hydrological Institutes of Austria, Belgium, Bulgaria, Finland, France, Germany, Hungary, Italy, Poland, Slovakia, Turkey, and from ECMWF.