

***"December 2001 sample image generated using ERS-2 MWR brightness temperature data"***

Image from "CEOS International Directory Network Portals, ERS-1/2 Microwave Sounder Micro Brightness Temperature Product".

Download: 2012-12-28.

*ESA contract*  
**4000109537/13/I-AM**

*ERS/Envisat MWR  
Recalibration and Water  
Vapour FDR Generation  
(EMiR)*



**EMiR Version 1.1  
Level-2 and -3  
product generation**

**Algorithm  
theoretical basis  
document**

(DLV-EXT-07B)

*Berlin, 14. May 2019*

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

<b>Document control sheet</b>																	
Project	ERS/Envisat MWR Recalibration and Water Vapour FDR Generation (EMiR)																
Client	<p>European Space Agency, ESRIN Via Galileo Galilei, Casella Postale 64 00044 Frascati, Italy <a href="http://www.esa.int">http://www.esa.int</a></p> 																
Main contractor	<p>Informus GmbH Brehmestr. 50 13187 Berlin, Germany <a href="http://www.informus.de">http://www.informus.de</a></p> 																
Title	EMiR Version 1.1 Level-2 and Level-3 product generation - Algorithm Technical Basis Document																
Status	Released																
Authors	Ralf Bennartz (VU), Martin Stengel (DWD), Oliver Sus (DWD), Heidrun Höschen (DWD), Frank Fell (INF)																
Distribution	EMiR consortium, ESA																
Version control	<table border="0"> <tr> <td>2015-04-15</td> <td>V0.90, first draft</td> </tr> <tr> <td>2015-04-15</td> <td>V0.91, revised draft</td> </tr> <tr> <td>2015-05-06</td> <td>V1.00, initial version</td> </tr> <tr> <td>2015-10-06</td> <td>V1.10, "FCDR" and "TCDR" replaced by "FDR" and "TDR"</td> </tr> <tr> <td>2016-04-19</td> <td>V2.00, section on Level-3 products added, full editorial revision</td> </tr> <tr> <td>2016-06-02</td> <td>V2.10, further feedback from ESA considered</td> </tr> <tr> <td>2017-02-20</td> <td>V2.11, metadata harmonised</td> </tr> <tr> <td>2019-05-14</td> <td>V2.20, adapted to new EMiR version 1.1</td> </tr> </table>	2015-04-15	V0.90, first draft	2015-04-15	V0.91, revised draft	2015-05-06	V1.00, initial version	2015-10-06	V1.10, "FCDR" and "TCDR" replaced by "FDR" and "TDR"	2016-04-19	V2.00, section on Level-3 products added, full editorial revision	2016-06-02	V2.10, further feedback from ESA considered	2017-02-20	V2.11, metadata harmonised	2019-05-14	V2.20, adapted to new EMiR version 1.1
2015-04-15	V0.90, first draft																
2015-04-15	V0.91, revised draft																
2015-05-06	V1.00, initial version																
2015-10-06	V1.10, "FCDR" and "TCDR" replaced by "FDR" and "TDR"																
2016-04-19	V2.00, section on Level-3 products added, full editorial revision																
2016-06-02	V2.10, further feedback from ESA considered																
2017-02-20	V2.11, metadata harmonised																
2019-05-14	V2.20, adapted to new EMiR version 1.1																

## Table of contents

<b>1</b>	<b>Introduction.....</b>	<b>6</b>
1.1	Purpose and structure of this document.....	6
1.2	Scientific background.....	6
1.3	Water vapour retrieval using passive microwave observations.....	7
1.4	Implementation overview.....	9
1.5	Acronyms and abbreviations.....	10
<b>2</b>	<b>Data sources.....</b>	<b>12</b>
<b>3</b>	<b>Total column water vapour retrieval.....</b>	<b>14</b>
3.1	Theoretical baseline.....	14
3.1.1	Radiative transfer.....	14
3.1.2	The MWR 1D-VAR retrieval scheme.....	15
3.1.3	Variational assimilation technique.....	15
3.1.4	Accounting for clouds.....	17
3.1.5	Minimization technique.....	17
3.2	Practical application.....	17
3.2.1	MWR (Microwave Radiometer).....	18
3.2.2	Retrieval schemes adaptation made for ERS-1/2 and Envisat MWR.....	18
3.2.3	Background profile.....	18
3.2.4	Rain flag.....	19
3.2.5	Spatial coverage.....	19
3.2.6	Quality control.....	19
<b>4</b>	<b>Inter-calibration and bias correction.....</b>	<b>20</b>
4.1	Method.....	20
4.2	Bias analysis.....	22
4.3	Sensitivity to choice of background profile.....	25
<b>5</b>	<b>Calculation of wet tropospheric delay.....</b>	<b>28</b>
5.1	General overview.....	28
5.2	Dry delay.....	29
5.3	Wet delay.....	29
<b>6</b>	<b>L2 data output.....</b>	<b>31</b>
6.1	Fill values.....	32
<b>7</b>	<b>EMiR Level-3 data.....</b>	<b>33</b>
7.1	Level-2 pre-screening.....	33
7.2	Level-3 product generation.....	33

7.3	Level-3 file structure .....	33
<b>8</b>	<b>Summary .....</b>	<b>35</b>
<b>9</b>	<b>References.....</b>	<b>36</b>

## List of tables

Table 1: Processing of ERS-1 and ERS-2 MWR brightness temperatures for generation of EMiR L1B dataset. Source: MWR calibration assessment [DLV-EXT-06]. .....	12
Table 2: Quality flags of the MWR 1D-VAR output.....	19
Table 3: Mean optimal bias correction values and regression slopes for the all instruments and time periods. The values given here correspond to the straight lines and dashed lines in Figure 8. A negative bias correction value means the observations are warmer than the simulations, thus need to be corrected downwards.....	24
Table 4: Retrieval statistics for four different 1D-VAR configurations. Corresponding scatterplots are shown in Figure 9. Total number of retrievals was 35,584. Reported bias and RMSE values are for valid only for the fraction of retrievals with cost function lower than 5.....	27
Table 5: Contents of EMiR V1.1 Level-2 netCDF files. Variables shown in green background have been newly added to V1.1.....	31
Table 6: Contents of EMiR V1.1 Level-3 netCDF files. Coordinates refer to the grid-box centre. ....	33

## List of figures

Figure 1: Zenith transmittance in the microwave spectral domain at different humidity levels (left) and for different absorbers / emitters (right) [Petty, 2006].....	8
Figure 2: Principal relationship between MWR-observed brightness temperatures and LWP and WVP (=TCWV), assuming a surface wind speed of 8 m/s. The left two panels show brightness temperatures as function of LWP and WVP. The right panel shows the inverse problem, i.e. LWP and WVP as function of the two observed brightness temperatures. To illustrate the effect of surface wind speed, a second grid (blue) is overlaid, calculated for a surface wind speed of 1 m/s. ....	8
Figure 3: Flow chart of the EMiR processor. External input data are marked grey, EMiR Level-2 and -3 products are marked orange, and EMiR processing software is marked green. The corresponding sections in this document are also indicated. ....	9
Figure 4: MWR brightness temperatures used as input for EMiR processing. Top: 23.8 GHz channel for ERS-1 (gold), ERS-2 (red), and Envisat (turquoise). Bottom: As above, but for the 36.5 GHz channel. ....	13

Figure 5: Microwave attenuation of water vapour and oxygen [*Liebe, 1985*]..... 15

Figure 6: Example of histograms of retrieved LWP and fitted Gaussian for different bias correction values at 36 GHz. The example is shows for Envisat, January 2001. The bias at correction at 23 GHz is set fixed to -3 K. .... 21

Figure 7: The plot in the upper panel shows contours of the LWP bias as a function of the bias correction values for 23 GHz (x-axis) and 36 GHz (y-axis). The plot in the lower panel shows the corresponding TCWV biases. Labelled isolines of both TCWV bias and LWP bias are overlaid in both plots. Data shown are for January 2001..... 22

Figure 8: Optimal bias correction values for the entire MWR time series shown as coloured lines with filled circles marking monthly values. Also shown are the temporal mean values for each channel and instrument (straight lines) and a regression (dashed lines). For ERS-2, two separate fits were performed, one corresponding to the period before the 23 GHz gain drop and another one for the period after. The blue arrow highlights the day the drop happened (6/26/1996). A negative bias correction value means that the observations are warmer than the simulations, thus need to be corrected downwards. .... 23

Figure 9: Scatterplots of retrievals obtained for four different 1D-VAR configurations. Total number of retrievals was 35,584. Reported values are for valid retrievals with cost function lower than 5. Corresponding statistics are listed in --- Remainder of page intentionally left blank--- ..... 26

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

# 1 Introduction

## 1.1 Purpose and structure of this document

This Algorithm Technical Basis Document (ATBD) describes the theoretical basis of total column water vapour (TCWV), Cloud Liquid Water Path (LWP), and Wet Tropospheric Correction (WTC) retrieval from Microwave Radiometer (MWR) observations on-board the ERS-1/2 and Envisat satellites, as developed and applied in the ERS/Envisat MWR recalibration (EMiR) project.

The ATBD is structured as follows:

- Section 1 describes the scientific context of the EMiR project.
- Section 2 introduces the MWR brightness temperatures on which the EMiR processing is built.
- Section 3 outlines the TCWV and LWP retrievals. The TCWV retrieval algorithm is based on an optimal estimation framework [Deblonde, 2001], which has also been used in the context of the ESA DUE GlobVapour project [GlobVapour, 2012].
- In Section 0, a novel approach for inter-calibrating microwave radiometers is described.
- In Section 5, the retrieval for Wet Tropospheric Correction (WTC) is laid out.
- Section 6 provides a detailed description of the EMiR Level-2 data sets produced from individual orbits.
- Section 7 describes generation and file structure of temporally and spatially gridded Level-3 EMiR data products.

This document is an updated version of the original EMiR ATBD and considers the changes made between the original EMiR data record (V1.0) and the revised version (V1.1). While the retrieval algorithm itself is unchanged, the new EMiR data record (V1.1) provides additional auxiliary and diagnostic information. These changes are reflected mainly in sections 6 and 7. In addition, a few minor editorial issues were corrected throughout the document.

## 1.2 Scientific background

ESA's altimetry missions are at the heart of significant progress on oceanography. The combined coverage of high-quality observations by ERS-1, ERS-2, and Envisat spans over more than 20 years from 1991 to 2012. During this period, improvements in instrument data processing as well as orbit and geophysical corrections allowed reaching an accuracy/sensitivity of 1 cm on instantaneous sea surface height (SSH) measurements and demonstrated the capability to observe a 3 mm/year sea level rise [Ablain et al., 2009].

A major source of uncertainty for radar altimetry is the wet tropospheric correction (WTC) taking into account the reduction of the speed of light in the atmosphere due to the presence of water vapour. The spatial and temporal variability of water vapour is such that an instantaneous estimation of its impact is needed. To provide the observations required for the WTC is the primary role of the nadir looking Microwave Radiometer (MWR) embedded into the altimetry missions on board ERS-1, ERS-2,

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

and Envisat.<sup>1</sup> In this context, constraints on accuracy, sensitivity, and long term stability of the atmospheric water vapour observations are particularly strong since altimetry missions require a precision better than 1 cm in WTC (RMS) [Eymard *et al.*, 2005] and a temporal stability better than 1mm/year [Ablain *et al.*, 2009]. Note that a total column water vapour (TCWV) contribution of 1 kg/m<sup>2</sup> is equivalent to a WTC of about 6.4 mm.

Water vapour is a highly important climate variable in its own right. The atmospheric water vapour feedback is believed to be the strongest feedback mechanism in climate change, approximately doubling the direct warming impact of increased CO<sub>2</sub> forcing [Cess *et al.*, 1990; IPCC-AR4, 2007]. Trends in the amount of columnar water vapour have been reported by various groups. In particular, over the oceans, a strong trend in TCWV has been observed [Trenberth *et al.*, 2005]. TCWV also appears to be a key factor regulating tropical precipitation [Bretherton *et al.*, 2004].

The MWR instruments on-board ERS-1, ERS-2, and Envisat are based on very similar architectures and have measured water vapour over the ocean between 1991 and 2012. MWR is a nadir-looking passive microwave radiometer. Its two channels located at 23.8 GHz and 36.5 GHz allow for the simultaneous retrieval of TCWV and cloud liquid water path (LWP) as outlined in Section 4.1. Through the REAPER (REprocessing of Altimeter Products for ERS) project<sup>2</sup>, ESA has provided significant efforts to produce a consistent and inter-calibrated time series of MWR observations for climate studies. The efforts described herein are based on the REAPER dataset. However, over the course of the EMiR activities it became clear that additional inter-calibration and bias correction efforts were needed, which are part of the current document as well.

The focus of the EMiR activities lies on the generation of a thematic data record (TDR) for water vapour, which can be used as an independent data record for climate studies. As a secondary product WTC is provided as well.

### 1.3 Water vapour retrieval using passive microwave observations

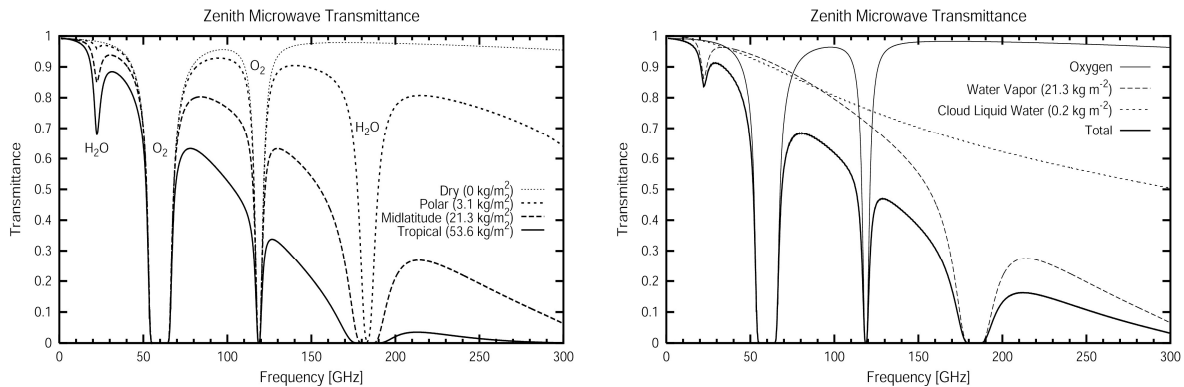
Passive microwave imagers have a long history for the retrieval of total column water vapour by measuring radiation close to the 22.231 GHz water vapour absorption line. Figure 1 shows the atmospheric zenith transmittance in the microwave region for different humidity levels and absorbing/emitting species. In the left panel, the sensitivity of microwave measurements to the total columnar water vapour is apparent both in the line centres, such as the absorption line centred around 22.231 GHz, as well as in the continuum in between the individual absorption/emission lines. The right panel shows the transmittance in the same spectral range, indicating the individual contributions to the transmittance by water vapour, oxygen, and a liquid cloud.

Clouds are semi-transparent throughout the microwave region and their transmittance does not exhibit line structures. It is possible to distinguish the signals due to water vapour and clouds by using channels at different water vapour optical depths, such as the Envisat MWR channels at 23.8 GHz and

<sup>1</sup> A microwave radiometer with very similar characteristics is flown on the Sentinel-3 series of satellites.

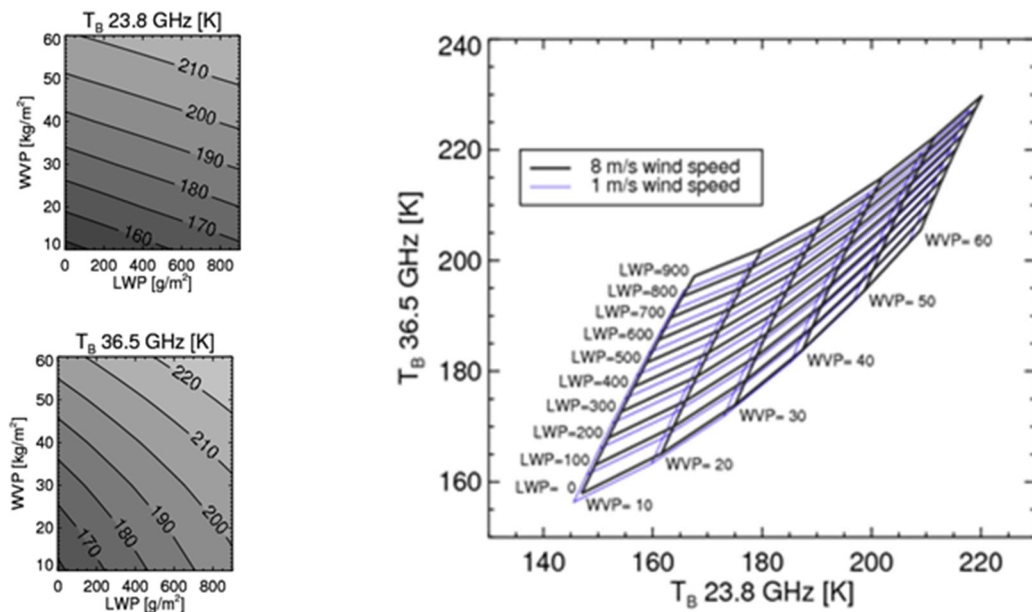
<sup>2</sup> <https://earth.esa.int/web/sppa/activities/multi-sensors-timeseries/reaper/>

36.5 GHz. Passive microwave measurements can therefore be used for the retrieval of water vapour under all-sky conditions with the exception of heavily precipitating situations.



**Figure 1: Zenith transmittance in the microwave spectral domain at different humidity levels (left) and for different absorbers / emitters (right) [Petty, 2006].**

However, due to the high and highly variable emissivity of land surfaces, this technique can only be applied over the oceans. The microwave emissivity of the ocean surface is a function of temperature, salinity and surface roughness, but is generally small, providing a good background for water vapour retrievals.



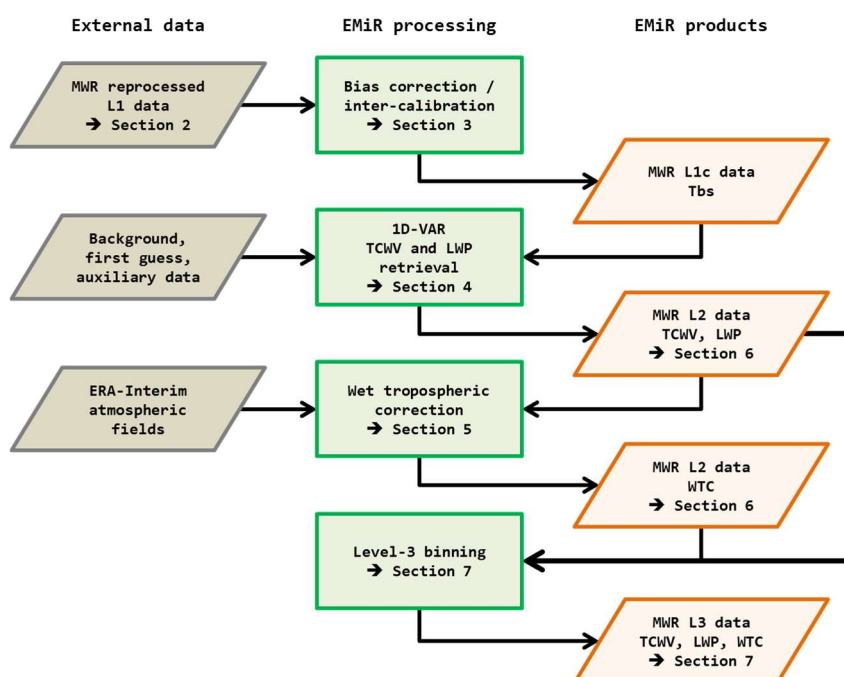
**Figure 2: Principal relationship between MWR-observed brightness temperatures and LWP and WVP (=TCWV), assuming a surface wind speed of 8 m/s. The left two panels show brightness temperatures as function of LWP and WVP. The right panel shows the inverse problem, i.e. LWP and WVP as function of the two observed brightness temperatures. To illustrate the effect of surface wind speed, a second grid (blue) is overlaid, calculated for a surface wind speed of 1 m/s.**



Figure 2 shows contour lines of equal brightness temperature in both Envisat MWR channels as a function of liquid water path and TCWV (termed water vapour path (WVP) in the figure). In both channels, the measured brightness temperature is a function of water vapour as well as liquid water, however with differing sensitivities to the two parameters. This is reflected in the lower plot, showing contour lines of equal amounts of TCWV and liquid water, with the x- and y-axis representing the brightness temperatures at 23.8 GHz and 36.5 GHz, respectively. The influence of the sea surface roughness (using wind speed as a proxy) is shown as well in order to indicate the impact of the resulting sea surface emissivity on the top-of-the-atmosphere brightness temperature.

## 1.4 Implementation overview

The EMiR TCWV retrieval is based on a 1D-VAR scheme initially developed at ECMWF by *Phalippou [1996]* with a focus on microwave observations from SSMIS and AMSU. It was extended by *Deblonde and English [2001]* towards a stand-alone scheme applicable to SSM/I, SSMIS, and AMSU. In the context of the ESA DUE GlobVapour project, the TCWV retrieval scheme has been optimized for SSM/I, including the usage of the CMSAF SSM/I Fundamental Data Record (FDR), and later on adopted to MWR observations.



**Figure 3: Flow chart of the EMiR processor. External input data are marked grey, EMiR Level-2 and -3 products are marked orange, and EMiR processing software is marked green. The corresponding sections in this document are also indicated.**

Within the EMiR project, this scheme has been further improved to derive TCWV from brightness temperatures specifically from the MWR sensor family on-board ERS-1/2 and Envisat over the ice-free ocean. The best estimate of the atmospheric state, characterised through atmospheric temperature and moisture profiles as well as surface temperature and wind speed, is determined by an iterative procedure to match simulated satellite radiances with the corresponding measurements within their

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

respective uncertainties. The scheme follows optimal estimation theory considering the uncertainties in the required meteorological background information, forward modelling (radiative transfer simulations), and satellite observations. This methodology enables the provision of retrieval uncertainties that are mathematically consistent with the uncertainties of the input brightness temperatures and background fields and consistent among the retrieved variables.

In the EMiR application, the 1D-VAR scheme uses daily global ERA-Interim TCWV and atmospheric temperature and cloud water content profiles as well as surface fields as a-priori (background) and first guess information. The impact of the choice of background fields is studied in Section 4.3.

Furthermore, inter-calibrated and homogenized MWR brightness temperatures (Level-1B) from the ERS-1, ERS-2 and Envisat satellites are used as measurement input (see Section 0 for details on the inter-calibration). Figure 3 shows a simplified sketch of the EMiR processing flow.

## 1.5 Acronyms and abbreviations

Acronym	Description
1D-VAR	One-dimensional Variational Data Assimilation
AATSR	Advanced Along-Track Scanning Radiometer
ACE-2	Altimeter Corrected Elevations (2)
AMSU	Advanced Microwave Sounding Unit
ATBD	Algorithm technical basis document
CMSAF	Satellite Application Facility on Climate Monitoring
DUE	Data User Element
ECMWF	European Centre for Medium-Range Weather Forecasts
EMiR	ERS/Envisat MWR Recalibration and Water Vapour FDR Generation
Envisat	Environmental Satellite
ERA-Interim	Global atmospheric reanalysis from 1979 to present by ECMWF
ERS	European Remote Sensing satellite
ESA	European Space Agency
FASTEM	Fast Microwave Emissivity Model
FDR	Fundamental data record
GTOPO30	Global DEM with a grid spacing of 30 arc seconds by USGS
L1	Level-1 processing
LWP	Liquid water path
MERIS	Medium Resolution Imaging Spectrometer
MWR	Microwave Radiometer
NSIDC	National Snow and Ice Data Center
NWP	Numerical weather prediction
REAPER	Reprocessing of Altimeter Products for ERS
RTM	Radiative transfer model
RTTOV	Radiative Transfer for TOVS
RMS	Rot mean square
SAF	Satellite Application Facility
SSH	Sea surface height
SSM/I	Special Sensor Microwave/Imager

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

Acronym	Description
SSMIS	Special Sensor Microwave Imager/Sounder
SWS	Surface wind speed
TDR	Thematic data record
TCWV	Total column water vapour
USGS	United States Geological Survey
UTC	Universal Time Coordinated
WTC	Wet tropospheric correction
WVP	Water vapour path

--- Remainder of page intentionally left blank---

## 2 Data sources

The Envisat MWR brightness temperatures (v2.1b) used herein have been generated by CLS in 2014 in the frame of the Envisat MWR L1B Expert Support Laboratory (ESL) activities funded by ESA. It consists of a corrected dataset that removes the anomaly that has been observed in version 2.0.

The ERS-1 and ERS-2 MWR brightness temperatures used herein have been entirely reprocessed in the frame of the EMiR project. The so-called “first run” REAPER L1B data have been the basis for this reprocessing. Land measurements are discarded, and no specific processing is applied in coastal areas so that contamination from land may occur above coastal waters at distances of less than ca. 50 km from land.

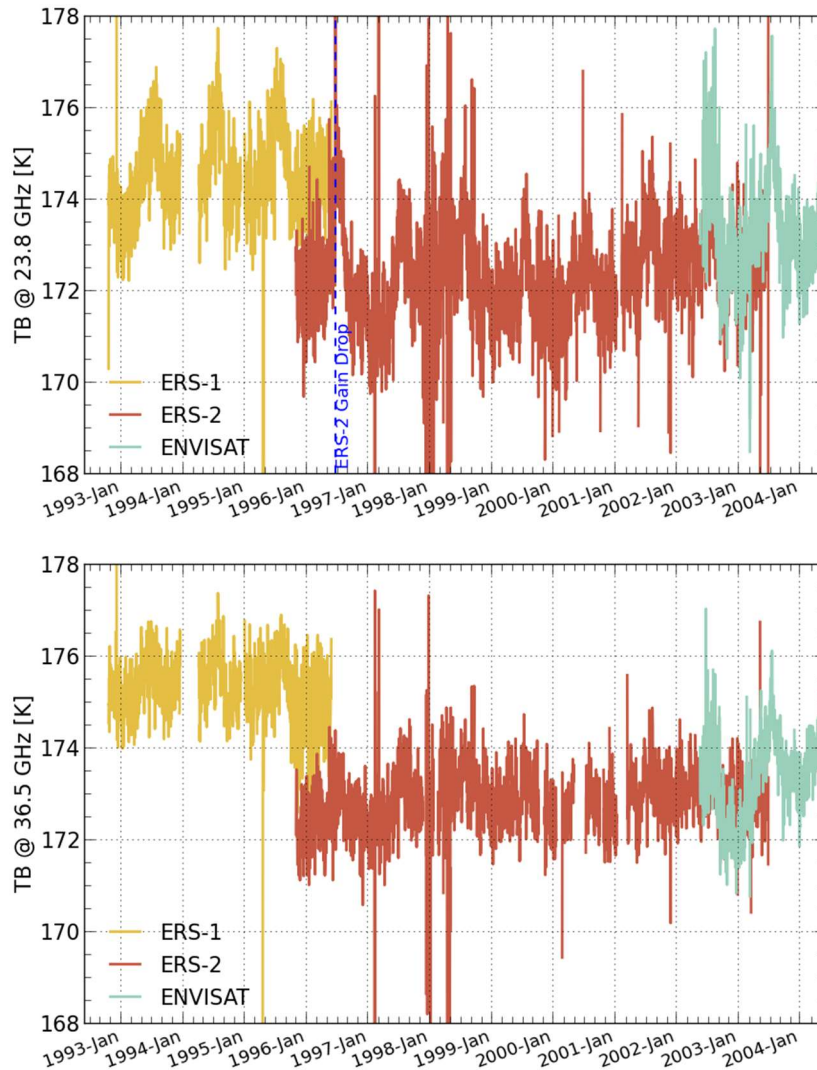
**Table 1: Processing of ERS-1 and ERS-2 MWR brightness temperatures for generation of EMiR L1B dataset. Source: MWR calibration assessment [DLV-EXT-06].**

ERS-1 and ERS-2 Tb dataset processing for EMiR	
L1B source	Basis is the REAPER “1 <sup>st</sup> run” (non-public) dataset: <ul style="list-style-type: none"> <li>➤ L0 consolidation (e.g. gap filling)</li> <li>➤ Envisat side-lobe correction algorithm applied</li> </ul>
Interpolation on altimeter time tag	Linear interpolation is applied to collocate MWR with the altimeter time tag
Surface coverage	Surface coverage is limited to ocean and sea ice; land surfaces are discarded
Coastal corrections	No specific coastal processing is applied; land contamination is possible for distances to coast $\leq$ 50 km
ERS-2 23.8 GHz channel	Corrections have been applied to correct gain drop and drift observed in the ERS-2 MWR 23.8 GHz channel after 1996-06-26

The main characteristics of the brightness temperatures used as input for the EMiR processing scheme are summarised in Table 1. Further details can be found in EMiR deliverable DLV-EXT-06.

--- Remainder of page intentionally left blank---

A graphical representation of the brightness temperatures is shown in Figure 4.



**Figure 4: MWR brightness temperatures used as input for EMiR processing. Top: 23.8 GHz channel for ERS-1 (gold), ERS-2 (red), and Envisat (turquoise). Bottom: As above, but for the 36.5 GHz channel.**

--- Remainder of page intentionally left blank---

## 3 Total column water vapour retrieval

### 3.1 Theoretical baseline

#### 3.1.1 Radiative transfer

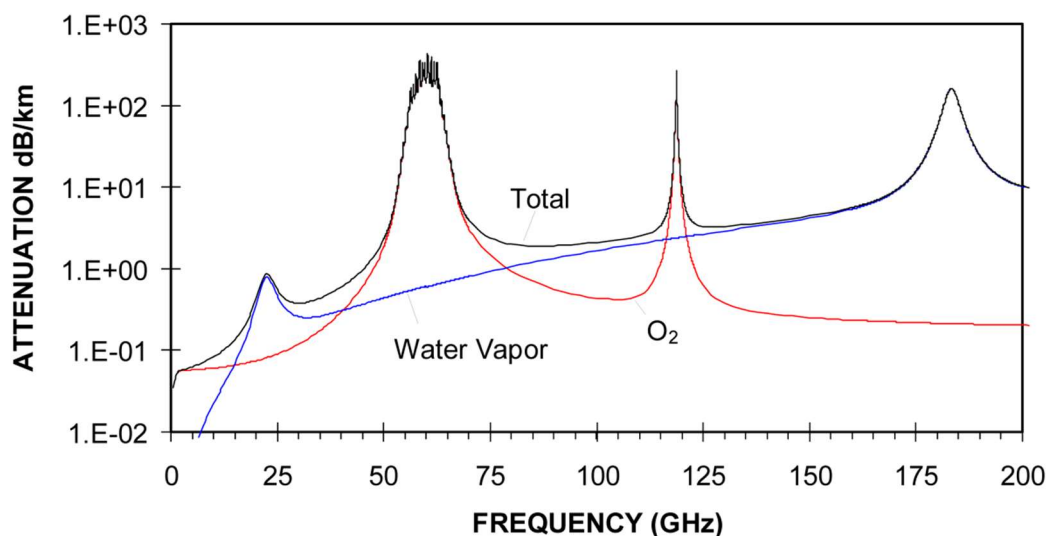
The following section about radiative transfer was adopted from CMSAFs ATBD on water vapour products [CMSAF, 2009], which uses the same underlying software package for water vapour retrievals [Deblonde, 2001]. The radiative transfer is approximated by Equation (1):

$$I_\nu = \varepsilon_\nu B_\nu(T_s) \tau_\nu^* + \int_{p_s}^0 B_\nu(T) \frac{\partial \tau_\nu}{\partial p} dp + R_\nu \tau_\nu^* \int_0^{p_s} B_\nu(T) \frac{\partial \tau_\nu}{\partial p} dp \quad (1)$$

where  $\nu$  is frequency,  $T$  is temperature,  $T_s$  is surface temperature,  $\varepsilon_\nu$  is surface emissivity,  $R_\nu$  is surface reflectivity,  $B_\nu$  is the Planck function,  $p$  is pressure,  $p_s$  is surface pressure,  $\tau_\nu$  is transmission and  $\tau_\nu^*$  is total atmospheric transmission. The three right-hand terms describe the following processes:

- The first (left-most) term describes surface emission at temperature  $T_s$  and emissivity  $\varepsilon_\nu$ , transmitted through the atmosphere with transmissivity  $\tau_\nu^*$ .
- The second term describes the upwelling radiation emitted in the atmosphere integrated from the surface to the top (in pressure coordinates). The emission  $B_\nu(T)$  is weighted by the vertical derivative of transmission as function of atmospheric pressure.
- The third term describes the downwelling atmospheric radiation that is reflected at the surface with the reflectivity  $R_\nu$  and transmitted through the atmosphere with the total transmissivity  $\tau_\nu^*$ . Note that over a rough ocean surface  $R_\nu \neq 1 - \varepsilon_\nu$  because of the effect of the slope variability of the ocean surface facets. For the particular treatment and its evolution within RTTOV, see e.g. Bormann et al. [2012].

The solution of the radiative transfer equation in the microwave part of the spectrum requires a description of the transmission of the atmosphere and the quantification of the surface emission. As shown in Figure 5, water vapour and oxygen are the relevant absorbers within the spectral range of MWR observations. In the radiative transfer model RTTOV used herein, gaseous absorption coefficients are calculated using a combination of different gas absorption models and spectroscopic database. For details, see Saunders et al. [2008].



**Figure 5: Microwave attenuation of water vapour and oxygen [Liebe, 1985].**

Hydrometeors do also affect the atmospheric transmission by scattering. As this process is not covered by the radiative transfer model (RTM) applied for the retrieval of TCWV, heavy precipitation events have to be filtered out *a priori* to the application of the algorithm or can be identified and removed as low-quality afterwards by applying constraints on the retrieved chi-squared error. The latter method is used within EMiR.

### 3.1.2 The MWR 1D-VAR retrieval scheme

*This section follows the NWP SAF User Guide [Deblonde, 2001]*

The MWR 1D-VAR solves for atmospheric temperature  $T$ , atmospheric water vapour  $Q$ , oceanic surface wind speed and liquid water path (LWP). The scheme requires atmospheric input profiles (background profiles) that are spatially and temporally collocated with the satellite observations and returns solution for  $T$  and  $Q$  profiles as well as LWP that optimally fit both observations and background profiles. The optimal fit is determined by the relative weight of the background error co-variances and the observation errors. The forward model applied is RTTOV 6.7, modified in a way that MWR brightness temperatures and their Jacobians can be computed. The FAST emissivity ocean model (FASTEM) Version 2.0 [Deblonde and English, 2001; English and Hewison, 1998] is used for this work. Since FASTEM-2 is included operationally only in RTTOV versions 7 and onwards, it had to be implemented into RTTOV 6.7 used for the 1D-VAR retrieval scheme developed for EMiR.

### 3.1.3 Variational assimilation technique

A variational retrieval is applied in which *a priori* or background information of the atmosphere and surface  $x^b$ , and the measurements  $y^0$  (observed brightness temperatures) are combined in a statistically optimal way (with a Bayesian analysis) to estimate the most probable atmospheric state  $x$ . The approach is common to a number of areas where non-linear inverse problems are encountered and has been described in detail by various authors [Rodgers, 1976; Tarantola and Valette, 1982; Lorenc, 1986]. Gaussian error distributions are assumed and consequently, obtaining the most

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

probable state is equivalent to minimising a cost function  $J(x)$  also referred to as a penalty function. Following the notation of *Ide et al. [1997]*,  $J(x)$  may be written as:

$$J(x) = \frac{1}{2}(x - x^b)^T B^{-1}(x - x^b) + \frac{1}{2}[y^o - H(x)]^T (E + F)^{-1}[y^o - H(x)] + J_s \quad (2)$$

where  $B$ ,  $E$ , and  $F$  represent background, instrumental, and representativeness (including errors of the forward model) error covariance matrices.  $J_s$  is a cubic function that limits the super-saturation and acts as a weak constraint [*Phalippou, 1996*]:

$$J_s = a(x - x_s)^3 \quad (3)$$

where  $x_s$  is the value of the control variable at saturation,  $H(x)$  is the forward operator that maps the control vector  $x$  into measurement space. Here,  $H(x)$  is the radiative transfer model RTTOV 6.7. The superscripts  $T$  and  $^{-1}$  denote matrix transpose and inverse respectively.

The control vector  $x$  consists of temperature (at 43 fixed RTTOV standard pressure levels), the natural logarithm of specific humidity (defined for the lowest 19 of those pressure levels) and the oceanic surface wind speed. Optionally, the liquid water path (LWP) defined below can also be added to the control vector (done within the EMiR context):

$$LWP = \frac{1}{g} \int_0^{P_s} q_L(P) dP \quad (4)$$

where  $g$  is the gravitational acceleration at the earth's surface,  $P_s$  is the surface pressure and  $q_L$  is the cloud liquid water content ( $\text{kg kg}^{-1}$ ). If LWP is not chosen as a control variable, then one solves for the natural logarithm of total water content. The total water content is defined as follows:

$$q_{total}(P) = q(P) + q_L(P) \quad (5)$$

where  $q$  is the specific humidity ( $\text{kg kg}^{-1}$ ). In general, the minimum of the cost function is found by the iterative solution of Newtonian iteration approach:

$$J''(x_n)(x_{n+1} - x_n) = -J'(x_n) \quad (6)$$

and

$$J'(x_n) \rightarrow 0 \quad (7)$$

where  $x_n$  and  $x_{n+1}$  are the  $n$ -th and  $(n+1)$ -th approximation of  $x$ ,  $J'$  and  $J''$  are the first and second derivatives of the cost function with respect to  $x$ . These are given by:

$$J'(x_n) = B^{-1}(x_n - x^b) - H'(x_n)^T (E + F)^{-1}(y^o - H(x_n)) \quad (8)$$

where  $H'(x_n)$  is the Jacobian matrix and contains the partial derivatives of  $H(x)$  with respect to  $x$ . In the linear limit,

$$J''(x_n) = B^{-1} + H'(x_n)^T (E + F)^{-1} H'(x_n) = A^{-1} \quad (9)$$

where  $A$  is the error covariance matrix of the solution if  $H(x)$  is linear.  $J''(x_n)$  is also referred to as the Hessian of the cost function.  $A$  is also called the analysis error covariance matrix. In this document,  $A$  will be referred to as the theoretical error.



ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

### 3.1.4 Accounting for clouds

To properly account for the absorption of cloud particle in the retrieval system, the liquid water path ( $LWP$ ) is included in the state vector. Thus the control vector consists of the profile of natural logarithm of specific humidity ( $\ln q$ ), the oceanic surface wind speed ( $SWS$ ) and the liquid water path.

Thus  $x = \{\ln q, SWS, LWP\}$ .

During the minimisation process of Eq. (2),  $LWP$  is allowed to vary while the cloud structure  $S(P)$  is maintained fixed. The cloud structure  $S(P)$  is defined as follows.

$$S(P) = q_L(P) / LWP \quad (10)$$

with  $q_L(P)$  as the profile of cloud liquid water.

If there is a cloud in the background profile, then  $S(P)$  is given by:

$$S(P) = q_{LB}(P) / LWP_B \quad (11)$$

where  $q_{LB}(p)$  is the background profile of cloud liquid water content and  $LWP_B$  is the liquid water path of the background profile.

If there is no cloud in the background profile, then a non-zero cloud structure is generated where the relative humidity of the background profile exceeds a pre-set threshold value (e.g. 80%). If there is still no cloud, then a non-zero cloud structure is assigned to the lowest levels of the profile. In all cases, the first guess  $LWP$  is set to  $0.1 \text{ kg m}^{-2}$ . The inclusion of the  $LWP$  in the state vector allows the retrieval to converge and retrieve meaningful water vapour values in the presence of cloud.

An alternative option is to retrieve the total water content instead of cloud liquid water and water vapour separately which follows closely the concept developed in *Blankenship et al. [2000]*. After each iteration step the total water content, as included in the control vector is split into the liquid and vapour part to be able to calculate the Jacobians with RTTOV. This option has not been tested within the EMiR project. The reader is referred to *Deblonde [2001]* for further details on both options.

### 3.1.5 Minimization technique

The Levenberg-Marquardt method was implemented for the minimisation as described in *Press et al. [1989]* (see page 523). The control vector is iteratively changed towards the most likely solution where the control vector fits best both the background information and the MWR observations, under consideration of the corresponding background and observation errors. The iteration is stopped after the gradient of the cost function fulfils a minimum limit.

## 3.2 Practical application

The 1D-VAR scheme was implemented at DWD, including the procedure of adjustment of the corresponding interfaces. This includes a collocation software that provides background and first guess data to the 1D-VAR at spatial locations of the MWR observations. It is possible to choose between ERA-Interim reanalysis data and climatologies as background data. The 1D-VAR was modified to read Level-1 files in NetCDF format and to output results in NetCDF format following the CF-1.4 standard.

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

### 3.2.1 MWR (Microwave Radiometer)

The MWR principal duty is the measurement of atmospheric humidity as supplementary information for tropospheric path correction of the radar altimeter signal, which is influenced both by the integrated atmospheric water vapour content and by liquid water. The MWR instrument on board Envisat is a heritage to its precursor instruments on board ERS-1 and ERS-2 satellites. The MWR is a dual-channel nadir-pointing Dicke-type radiometer, operating at frequencies of 23.8 and 36.5 GHz [ESA Earth Online, 2015].

### 3.2.2 Retrieval schemes adaptation made for ERS-1/2 and Envisat MWR

As previously mentioned, the 1D-VAR framework for MWR originates from developments by *Phalippou [1996]* and adjustments made for SSM/I, SSMIS and AMSU [Deblonde, 2001]. The following three pillars summarized the technical adaptations necessary for MWR:

- Defining MWR as a satellite instrument in RTTOV (RTTOV 6.7 used in this implementation).
- Composing the absorption coefficient file to be used as input for RTTOV (collection of the instrument characteristics for MWR, e.g. central wavelength, bandwidth, polarization, etc.).
- Defining MWR as one input option in 1D-VAR.

Since MWR is nadir looking only, it does not provide any polarization information. Compared to other microwave sensors, its spectral range is also limited to just two frequencies below 37 GHz. Therefore, screening observations affected by frozen hydrometeor scattering will be less effective as compared to other microwave instruments.

Further, no ice mask is provided in the MWR L-1 data. Therefore, sea ice coverage is determined from SSM/I sea ice concentration maps available from the National Snow and Ice Data Center (NSIDC)<sup>3</sup>. These maps are available daily with a spatial resolution of 25 x 25 km. Distance to land is calculated based on the ACE-2 topography<sup>4</sup> [Berry et al., 2008] and GTOPO30<sup>5</sup> available from USGS.

### 3.2.3 Background profile

1D-VAR expects a background temperature and humidity profile, which is also used as a first guess in the 1D-VAR retrievals. For this purpose, we make use of ERA-Interim reanalysis data. For each processed day, one representative ERA-Interim data file is chosen. This background data is collocated with MWR footprints. The data contains the atmospheric profiles for temperature, specific humidity and liquid water content provided on pressure levels as well as the surface values for wind speed, temperature and pressure.

As a second option, climatological profiles, can be chosen as background and first guess. However, the correspondingly large background errors might not be well represented in the currently used background error covariance matrix. We have studied the impact of different first guess/background

<sup>3</sup> SSM/I sea ice concentration maps: <http://dx.doi.org/10.5067/8GQ8LZQVLOVL>

<sup>4</sup> ACE-2 topography: <http://tethys.eaprs.cse.dmu.ac.uk/ACE2/>

<sup>5</sup> GTOPO30: <https://lta.cr.usgs.gov/GTOPO30>

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

profiles in the framework of this study (Section 4.3) and the retrieval is to a large extent independent on the exact choice of the background. Only in situations where the background deviates significantly from the true state will the retrieval fail. Such would be the case, if the true state of the atmosphere was for example tropic (e.g. TCWV = 40 kg/m<sup>2</sup>) and the background would assume a mid-latitude atmosphere (TCWV = 28 kg/m<sup>2</sup>). Since ERA reanalysis is used as background, the background will always be close enough to the true state, and extreme discrepancies will be highly unlikely in the EMiR retrievals.

### 3.2.4 Rain flag

If scattering information is available, the 1D-VAR is only applied in non-precipitating situations which are filtered using an internal scattering index module which rejects cases with strong scattering suggesting large particle of rain or ice. This step is necessary since scattering is not accounted for in the radiative transfer model. If scattering information is not available (as is the case for MWR), this screening cannot be applied. However, the final value of the cost function  $J$  will provide some information on the quality of the retrieval and will flagging cost function values  $J > 5$  proves an effective way to flag out heavy precipitation.

### 3.2.5 Spatial coverage

The algorithm is only applied to footprints completely filled with ocean surfaces. Land and ice surfaces are not processed due to difficulties in the provision of proper surface emissivity. Also, partially land- or ice-covered footprints are omitted.

### 3.2.6 Quality control

A quality flag exists in the MWR 1D-VAR output, adopting the values shown in Table 2. The pixel-based quality check is applied within the 1D-VAR. The valid range for TCWV is set to 0.1 to 90 kg/m<sup>2</sup>.

**Table 2: Quality flags of the MWR 1D-VAR output.**

Flag value	Meaning
0	Data not in valid range
1	Good value
2	Data not taken over ocean or brightness temperature is not correct
99	Default

--- Remainder of page intentionally left blank---

## 4 Inter-calibration and bias correction

Retrievals of geophysical variables using physical models and optimal estimation procedures require the elements of the observation vector (here: observed brightness temperatures) to be unbiased compared to the forward model applied to the true state of the atmosphere (here: simulated brightness temperatures).

Comparing first guess simulations with observations includes contributions from the following error sources:

- Representativeness of the first guess state vector (in our case ERA-Interim analyses) of the actual observations (e.g. surface winds, representation of clouds),
- Spatial and temporal colocation errors between first guess simulations and observations,
- Calibration biases / errors of the considered instruments (MWRs in our case),
- Systematic errors and uncertainties in the surface emissivity model,
- Systematic errors and uncertainties in spectroscopy of liquid water absorption, dry air absorption, and water vapour absorption,
- Impact of precipitation contamination and precipitation-ice scattering not accounted for in forward model.

While the first two above items have a significant impact on the values reported here, they only play a secondary role for the retrieval accuracy. The latter four items of the above list, while having smaller contributions to the overall bias, are of crucial importance to the accuracy and long-term stability of a Fundamental Data Record (FDR). These are addressed in an empirical bias-correction and inter-calibration scheme as outlined below.

### 4.1 Method

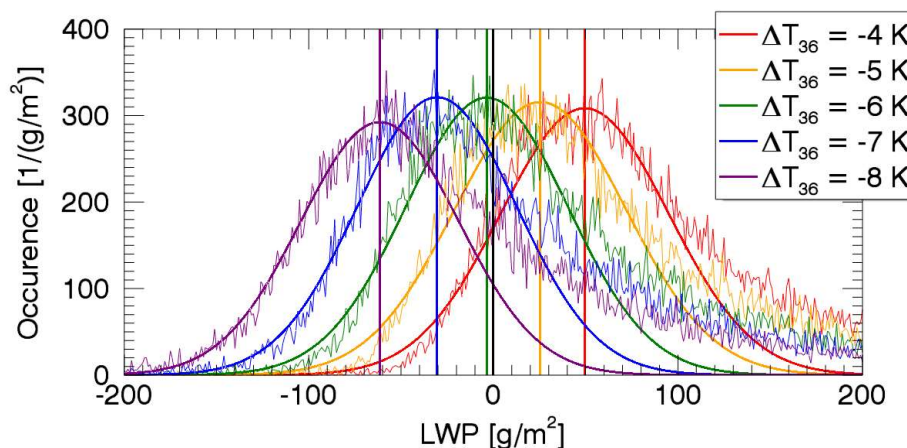
The bias correction proposed here relies on two main assumptions that are used to inter-calibrate the different MWRs:

- The globally averaged TCWV from ERA-Interim is considered reasonably accurate in terms of its absolute value to provide a reference against which to gauge the average brightness temperature biases of the MWR time series. Note, that we do not make any claims about the long-term stability of the ERA-Interim TCWV or any trends and discontinuities of the dataset. The only assumption we make is that ERA-Interim TCWV on a globally and monthly averaged basis is accurate to within, say,  $\pm 2 \text{ kg/m}^2$ . This assumption is justified from intercomparison efforts such as from *Schröder et al. [2013]*.
- The second assumption is that histograms of instantaneous retrieved LWP must show a significant fraction of negative values, corresponding to measurement noise around zero LWP for cloud-free situations. For typical bias-free optimal estimation retrievals, LWP for cloud-free cases is centred around  $0 \text{ g/m}^2$  with a standard deviation of about  $30 \text{ g/m}^2$  [*Bennartz et al., 2010; Greenwald, 2009*].

These two constraints can be used to find an optimal bias correction for both channels of each instrument in the following way:

1. For each instrument and month, randomly sub-select a certain amount of observations out of all observations available. We randomly selected 4% of the total number of observations for each month, representing the actual distribution of the observations for that month.
2. For these 4 %, perform a series of retrievals for different biases at 23 GHz and 36 GHz. We ran retrievals for bias values running from -8 K to 0 K in steps of 1 K for both channels, so that in total  $9 \times 9 = 81$  retrievals were performed on the selected sub-sets. Positive bias values were initially simulated too, but all instruments showed only negative biases, so that simulations with positive bias values were eliminated from the investigation.
3. Identify bias values for 23 and 36 GHz where the difference between background (ERA-Interim) and retrieved TCWV is smallest and the LWP additionally shows the Gaussian behaviour around zero. This will be the optimal bias value for this particular month.
4. Repeat step 3.) for all months and find optimal average bias values. The result is a bias time series for each channel at monthly resolution.

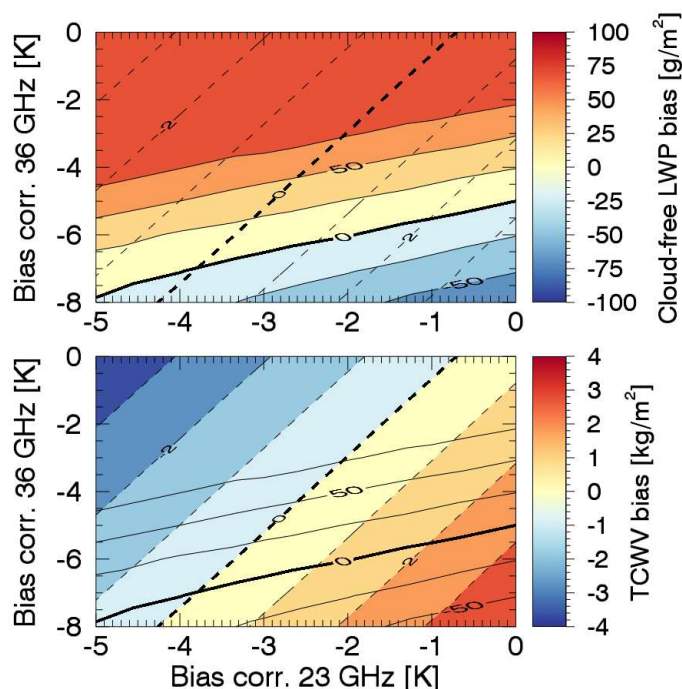
Figure 6 and Figure 7 highlight some key methodological issues related to the method. In particular, Figure 6 shows how the histogram of LWP shifts as the bias at 36 GHz varies. It also shows the histograms fitted to the retrieved LWP histograms. This fit was performed on the part of the histogram left of its peak, assuming that all values left of the peak correspond to cloud-free scenes. The skewed super-Gaussian distribution to the right of the peak corresponds to actual clouds. In the particular case shown in Figure 6 the best bias value for 36 GHz would be close to -6 K, thereby centring the histogram on zero as outlined above.



**Figure 6: Example of histograms of retrieved LWP and fitted Gaussian for different bias correction values at 36 GHz. The example is shows for Envisat, January 2001. The bias at correction at 23 GHz is set fixed to -3 K.**

Figure 7 shows isolines of TCWV and LWP histogram biases for a full set of monthly retrievals and all combinations of 23 GHz and 36 GHz biases. The optimal set of biases can now be inferred from this histogram as the intersect between the zero TCWV bias isoline (thick, solid) and the zero LWP histogram bias line (thick, dashed) and is located near (-4 K, -7K). The same analysis was performed for all months and instruments. The results are shown in Figure 8 and discussed in the next section.

From Figure 7 it is also noteworthy that the sensitivity of TCWV to biases in 23 GHz brightness temperatures is roughly 1 kg/m<sup>2</sup> per 1 K bias. The sensitivity of LWP to biases at 36 GHz is roughly 25 g/m<sup>2</sup> per 1 K bias. Note that both variables also exhibit sensitivity to the other frequency although the sensitivity is somewhat smaller, as expected e.g. from Figure 2.



**Figure 7: The plot in the upper panel shows contours of the LWP bias as a function of the bias correction values for 23 GHz (x-axis) and 36 GHz (y-axis). The plot in the lower panel shows the corresponding TCWV biases. Labelled isolines of both TCWV bias and LWP bias are overlaid in both plots. Data shown are for January 2001.**

## 4.2 Bias analysis

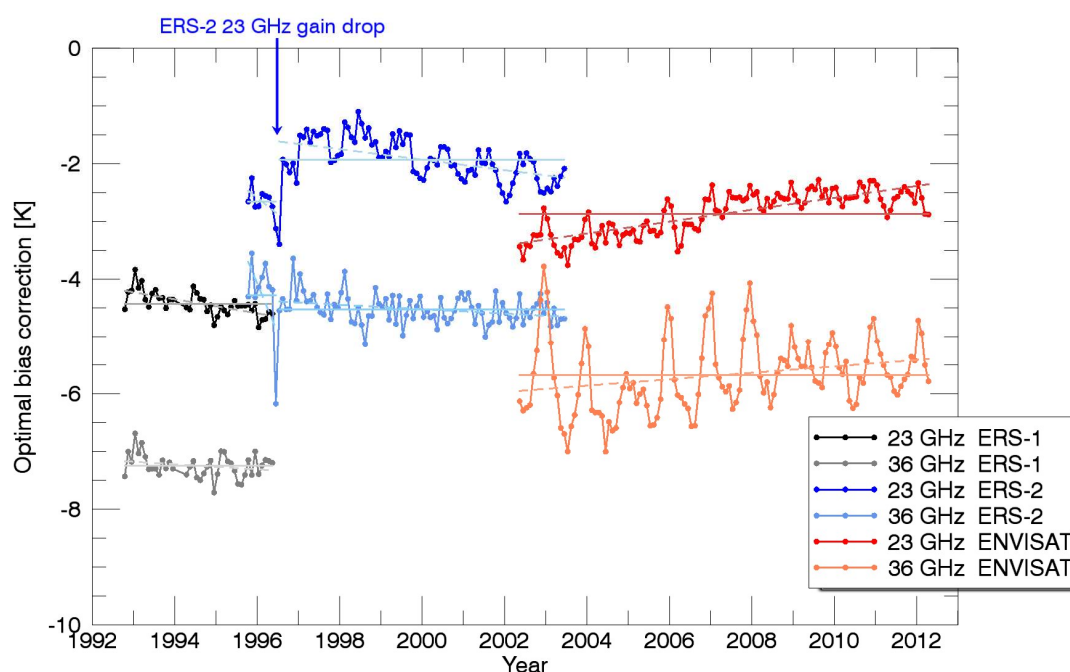
Figure 8 shows the outcome of the above-described bias analysis for all instruments and channels. Mean bias values as well as slopes are also listed in Table 3. A negative bias correction value means that the observations are warmer than the simulations, thus the observations need to be corrected downwards. The following observations can be made:

- Biases correction values at 23 GHz range between about -4 K for ERS-1 and -2 K for ERS-2 with Envisat being in between these two values.
- ERS-2 23 GHz shows a significant decrease in bias exactly at the time of the gain drop (indicated by blue arrow in Figure 7) and a downward trend in bias (slope) for the period afterwards. This strong drop prompted us to separate ERS-2 in a pre- and post-gain time period for which performed a separate analysis (listed in Table 3).
- Envisat shows a slight upward slope possibly over the first half of its lifetime or over its entire lifetime.



- Biases at 36 GHz are comparably stable over time for both ERS-1 and ERS-2, i.e. the regression slopes are very small.
- Envisat shows a strong annual cycle in bias at 36 GHz, which diminishes somewhat after 2008. It also shows a similar trend as it does for 23 GHz before 2007. There was no explanation for this behaviour at the time of writing.

We note once more that the bias values given here are based on optimal comparison between retrieved versus background TCWV as well as constraints made on the histogram of retrieved LWP. While these constraints are physically reasonable and justifiable on average, it is not advisable to perform monthly bias corrections based on the individual values derived for that particular month. This would by example of TCWV likely result in an over-fitting of the results to the ERA-Interim time series.



**Figure 8: Optimal bias correction values for the entire MWR time series shown as coloured lines with filled circles marking monthly values. Also shown are the temporal mean values for each channel and instrument (straight lines) and a regression (dashed lines). For ERS-2, two separate fits were performed, one corresponding to the period before the 23 GHz gain drop and another one for the period after. The blue arrow highlights the day the drop happened (6/26/1996). A negative bias correction value means that the observations are warmer than the simulations, thus need to be corrected downwards.**

However, biases between different instruments observed in the overlap period can be clearly attributed to the instrument calibration, as the background is identical. Similarly, the mean bias values reported in Table 3 provide for a reasonable way of addressing the error sources outlined in Section 4.1. A first order bias correction is therefore proposed through subtracting the bias values in Table 3 from the observations.

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

We note here that biases in the order of -2 K to -5 K (simulations too warm compared to observations) are also reported by ECMWF for monitoring of AMSU-A against their operational forecasting system<sup>6</sup>. It is therefore likely that ERS-2 is calibrated to within the absolute calibration accuracy of 3 K stated for the instrument. On the other hand, ERS-1 and Envisat both show much larger biases, which might be indicative of remaining calibration issues.

**Table 3: Mean optimal bias correction values and regression slopes for the all instruments and time periods. The values given here correspond to the straight lines and dashed lines in Figure 8. A negative bias correction value means the observations are warmer than the simulations, thus need to be corrected downwards.**

Instrument	Period	Mean		Regression slope (*)		Regression offset (*)	
		23 GHz	36 GHz	23 GHz	36 GHz	23 GHz	36 GHz
		[K]	[K]	[K/yr]	[K/yr]	[K]	[K]
ERS-1	10/1992 – 06/1996	-4.42	-7.25	-0.12	-0.04	-3.86	-7.05
ERS-2	10/1995 – 06/1996	-2.66	-4.28	-0.57	-1.72	+0.83	+6.24
ERS-2	07/1996 – 06/2003	-1.93	-4.52	-0.09	-0.04	-1.02	-4.14
Envisat	05/2002 – 04/2012	-2.87	-5.68	+0.10	+0.06	-4.65	-6.65

(\*) The regression bias is calculated using  $\text{bias\_corr}(t) = \text{slope} * t + \text{offset}$ , where  $t$  is the decimal year since 1990. For example, July 2, 1991 is day 183 in the year 1991 and therefore corresponds to a value of  $t=1.5$ .

As pointed out above, no assumptions about the long-term stability of ERA-Interim should be made in this analysis. The regressions listed in Table 3 therefore cannot conclusively be interpreted as either being caused by natural variability in TCWV or as being caused by instrument drifts.

While a conclusive statement of the origins of the regression slopes and related trends cannot be made, it is interesting to relate the slopes back to the aforementioned retrieval sensitivities. Ignoring the relatively short ERS-2 period before the gain drop, the regression slopes found in Table 3 show values between -0.12 K/yr and +0.1 K/yr. Factoring in the sensitivities of the retrieval, the retrieved TCWV trends observed in global TCWV between the constant bias correction and the regression are expected to be different by roughly -1.2 kg/(m<sup>2</sup> decade) to + 1.0 kg/(m<sup>2</sup> decade). These numbers provide important bounds on how large observed TCWV trends have to be in order to be considered real, if the current dataset is being used.

In order to assess which of the two bias corrections is more realistic, one either could try identify and correct for remaining calibration issues in the individual radiometer's calibration. In addition, one could perform two-way comparisons with other instruments, such as AMSU-A, which might provide an

<sup>6</sup> [http://old.ecmwf.int/products/forecasts/d/charts/monitoring/satellite/mwimg/amsua\\_allsky/](http://old.ecmwf.int/products/forecasts/d/charts/monitoring/satellite/mwimg/amsua_allsky/)



ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

independent reference for the temporal stability at least for ERS-2 and Envisat. For the purpose of EMiR the decision was made to implement the linear correction corresponding to the two last columns of Table 3.

### 4.3 Sensitivity to choice of background profile

In order to understand the sensitivity of the retrievals to different choices of background state vectors  $x_b$  and background error covariance matrices  $B$ , a series of retrieval tests was performed on a single day (2011/12/01) of Envisat MWR observations. The following four tests were run:

1. ERA\_OLD: ERA-Interim background with tight constraints on water vapour and liquid water background error covariance.
2. ERA\_NEW: ERA-Interim background, but with less tight constraints on water vapour and liquid water. Constraints were relaxed by a factor of two in water vapour and five in cloud liquid water.
3. FIXED\_OLD: A fixed mid-latitude summer atmospheric background profile was used for all retrievals. The same tight constraints on error covariance matrices as in 1. were used.
4. FIXED\_NEW: The same fixed background profile as in 3., but with less tight constraints on background error covariance as in 2.

The fixed background was simulated using a mid-latitude standard atmosphere as first guess and background for all retrievals. Only surface wind speed and sea surface temperature were still used from ERA-Interim for the climate background. The FIXED scenario represents an extreme case of a climatological background profile in which the background is kept fixed regardless of location and season. This extreme case has been chosen because it allows studying the algorithm performance under a most restrictive scenario with a fixed background.

The modified background error covariance matrix for tests 2 and 4 was implemented expanding the background standard deviation for cloud liquid water from  $0.2 \text{ kg/m}^2$  to  $1.0 \text{ kg/m}^2$  and in addition by multiplying the  $\ln Q$  sub-matrix of the background error covariance by a factor of two. For all cases we counted retrievals as valid, if, after convergence, the final cost function value was lower than five, i.e. at maximum 2.5 times larger than the expectation value of the cost function for valid retrievals.

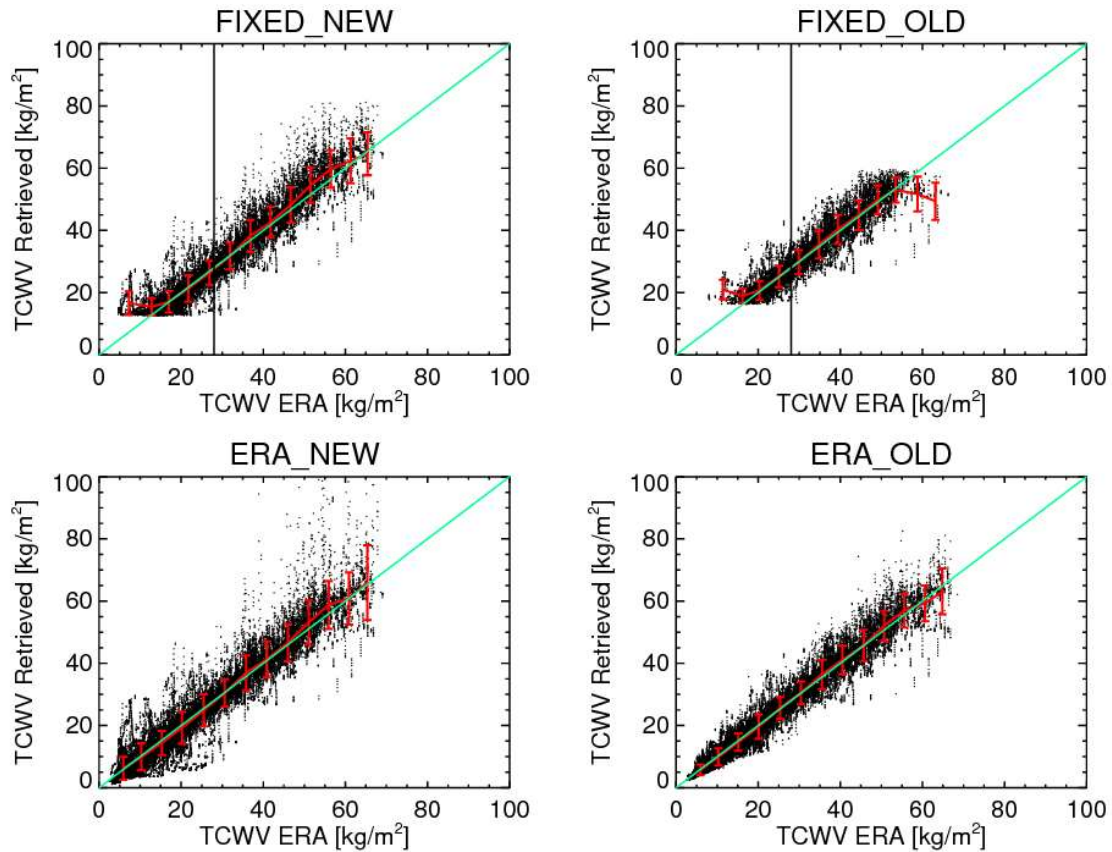
The results of these tests are summarized in --- *Remainder of page intentionally left blank*---

Table 4 and Figure 9. Both FIXED retrievals show critical deficiencies. Neither the original nor the modified settings allow for a good fit using just one constant climate-like profile as background. In both FIXED cases only about 50% of the retrievals actually converge and large biases occur both at the high and low end of TCWV. These issues can be mitigated by further increasing the  $\ln Q$  sub-matrix of the background error covariance by a factor of 10 instead of two and by increasing also the number of iterations in the minimization process from a current upper limit of five to 40. However, even with these newly revised parameters the number of converged profiles remains lower than for the ERA background.

An important finding from the FIXED cases is the relative insensitivity of the retrieval to the choice of the background. As can be seen in Figure 9, as long as the actual TCWV is less than maybe 5-10  $\text{kg/m}^2$  away from the chosen background, the retrieval will perform quite well, especially under relaxed

background error covariance conditions. This is due to the high information content of the passive microwave observations with respect to both TCWV and LWP.

We note that the use of a single global background profile is not necessarily the best choice for a climatological background. A possible compromise could consist of less stringent choices of climatological backgrounds allowing the background water vapour and temperature profile to vary with geographical position and latitude.



**Figure 9: Scatterplots of retrievals obtained for four different 1D-VAR configurations. Total number of retrievals was 35,584. Reported values are for valid retrievals with cost function lower than 5. Corresponding statistics are listed in --- Remainder of page intentionally left blank---**

**Table 4. In all four panels, the retrieved TCWV is plotted against ERA-Interim TCWV. The green line is the 1:1 line. The red error bars show the mean and standard deviation in bins of 5 kg/m<sup>2</sup>. The vertical line in the upper two plots shows the TCWV of the fixed climate background (mid-latitude summer atmosphere) used for the retrievals shown.**

--- Remainder of page intentionally left blank---

**Table 4: Retrieval statistics for four different 1D-VAR configurations. Corresponding scatterplots are shown in Figure 9. Total number of retrievals was 35,584. Reported bias and RMSE values are for valid only for the fraction of retrievals with cost function lower than 5.**

Experiment	Bias w.r.t. ERA-Interim	RMSE w.r.t. ERA-Interim	Percent retrieved	Mean $T_b$ residual after retrieval	Percent with $T_b$ residual > 1 K
	[kg/m <sup>2</sup> ]	[kg/m <sup>2</sup> ]	[%]	[K]	[%]
FIXED_NEW	0.88	4.80	68.5	0.23	0.02
FIXED_OLD	0.46	3.99	45.7	0.42	1.52
ERA_NEW	-0.01	5.06	97.9	0.07	0.91
ERA_OLD	0.00	3.53	87.2	0.41	6.24

Compared to the FIXED cases both the ERA\_NEW and ERA\_OLD case show much better results. The trade-off here lies mainly between an increased RMSE (NEW) and an increased number of profiles with large remaining  $T_b$  residuals after convergence (OLD). The ERA\_NEW case allows the 1D-VAR to find low cost solutions further away from the background water vapour profile. This will enhance the RMSE because we compare the retrieved TCWV to the background TCWV. ERA\_NEW in contrast provides tighter constraints on the background error covariance matrix, thus minimizes the RMSE better but at the cost of having a larger fraction of retrievals not converge as closely toward the observed brightness temperatures. For the particular case shown here, 6.24 % of retrievals still show a residual deviation of simulated from observed  $T_b$ s larger than 1 K.

A design choice for the final retrieved time series was therefore the extent to which it adheres to the prescribed background compared to perfectly minimizing the observed brightness temperatures. We note that in an ideal world with perfect knowledge about background and observation error covariance matrix this choice could not be made and the retrieval would provide a perfect a-posteriori estimate of the true state of the atmosphere accounting for correct background and observational information. However, as is always the case the actual retrieval will have to be tuned to some degree. In particular, one wants to minimize the risk of artefacts in the background data to affect the final TCWV time series. Such artefacts can example includes slight discontinuities in the ERA TCWV time series, at time steps where new sensors are added to the reanalysis.

With these considerations in mind we have chosen the ERA\_NEW 1D-VAR setup to be used as the basis for the full time series. The modified background error matrix allows for large deviations from the background profile, i.e. it gives stronger weight to the observations. At the same time, it provides good convergence over the entire range of variability of TCWV and allows for a high number of converged profiles and therefore provides very little sensitivity to the choice of the background profile. In particular, the choice of the background profile is uncritical, as long as it is somewhat representative for the geographical region and season.

## 5 Calculation of wet tropospheric delay

### 5.1 General overview

Radar altimeter path delay  $\Delta z$  along a path  $H$  is directly related to the real part of the refractive index of moist air  $n$ :

$$\Delta z = \int_0^H (n-1) dz \quad (12)$$

Expressing this in terms of refractivity  $N$ , with  $N$  in ppm being:

$$N = 10^6 (n-1) \quad (13)$$

we get:

$$\Delta z = 10^{-6} \int_0^H N dz \quad (14)$$

Assuming  $H$  is the satellite altitude, nadir view,  $T_v$  to be the virtual temperature, and using hydrostatic equilibrium we get:

$$\Delta z = 10^{-6} \frac{R_{AIR}}{g} \int_0^{p_{SEC}} N \frac{T_v}{p} dp \quad (15)$$

The refractivity  $N$  can be parameterized following references cited in *Mangum [2009]*:

$$N = a_d \frac{p_d}{T} + a_w \frac{e}{T} + b_w \frac{e}{T^2}$$

$$a_d : 0.776890 \frac{\text{ppm} \cdot \text{K}}{\text{Pa}} \quad (16)$$

$$a_w : 0.712952 \frac{\text{ppm} \cdot \text{K}}{\text{Pa}}$$

$$b_w : 3754.63 \frac{\text{ppm} \cdot \text{K}^2}{\text{Pa}}$$

The variable  $p$  is the total pressure and  $p_d$  represents the pressure of dry air, where the total pressure  $p = e + p_d$ , with  $e$  being the water vapour partial pressure. With these definitions we can write the total path delay  $\Delta z$  as:

$$\Delta z = 10^{-6} \frac{R_{AIR}}{g} \left[ \int_0^{p_{SFC}} a_d \frac{p_d}{T} \frac{T_v}{p} dp + \int_0^{p_{SFC}} a_w \frac{e}{T} \frac{T_v}{p} dp + \int_0^{p_{SFC}} b_w \frac{e}{T^2} \frac{T_v}{p} dp \right] \quad (17)$$

The first term in the brackets in Equation (17) can be split as follows:

$$\int_0^{p_{SFC}} \frac{p - e}{T} \frac{T_v}{p} dp = \int_0^{p_{SFC}} \frac{p}{T} \frac{T_v}{p} dp - \int_0^{p_{SFC}} \frac{e}{T} \frac{T_v}{p} dp \quad (18)$$

so that Equation (17) can be expanded to become:

$$\Delta z = 10^{-6} \frac{R_{AIR}}{g} \int_0^{p_{SFC}} a_d \frac{p}{T} \frac{T_v}{p} dp + 10^{-6} \frac{R_{AIR}}{g} \left[ \int_0^{p_{SFC}} (a_w - a_d) \frac{e}{T} \frac{T_v}{p} dp + \int_0^{p_{SFC}} b_w \frac{e}{T^2} \frac{T_v}{p} dp \right] \quad (19)$$

Dry tropospheric delay
Wet tropospheric delay

Note that  $T_v / T \approx 1$ .

## 5.2 Dry delay

Integrating the dry tropospheric part of Equation (19) yields:

$$\Delta z_d = 10^{-6} \cdot \frac{R_{AIR}}{g} \cdot a_d \cdot p_{SFC} \quad (20)$$

The dry delay is in the order of 2.3 m for a straight vertical path through the atmosphere whereas the wet tropospheric delay is only on the order of 0.4-0.5 m at maximum.

## 5.3 Wet delay

Integrating the wet tropospheric terms in Equation (19) yields:

$$\Delta z_w = 10^{-6} \frac{R_{AIR}}{g} \left[ \int_0^{p_{SFC}} (a_w - a_d) \frac{e}{T} \frac{T_v}{p} dp + \int_0^{p_{SFC}} b_w \frac{e}{T^2} \frac{T_v}{p} dp \right] \quad (21)$$

The specific humidity is defined as:

$$r = \frac{R_{H_2O}}{R_{AIR}} \frac{e}{p} \quad (22)$$

Replacing  $e/p$  accordingly with  $r$  into Equation (21) yields:

$$\Delta z_w = 10^{-6} \frac{R_{H_2O}}{g} \left[ (a_w - a_d) \int_0^{p_{SFC}} r dp + b_w \int_0^{p_{SFC}} \frac{r}{T} dp \right] \quad (23)$$

The total column water vapour (TCWV) is defined as:

$$TCWV = \frac{1}{g} \int_0^{p_{SFC}} r dp \quad (24)$$

We further define a 'water-vapour-averaged mean inverse atmospheric temperature',  $T_m$ :

$$T_m = \left( \int_0^{p_{SFC}} \frac{r}{T} dp / TCWV \right)^{-1} \quad (25)$$

With these two quantities, Equation (23) becomes:

$$\Delta z_w = 10^{-6} R_{H_2O} \left( (a_w - a_d) + \frac{b_w}{T_m} \right) \cdot TCWV = \left( A + \frac{B}{T_m} \right) \cdot TCWV \quad (26)$$

$$A : 10^{-6} \cdot R_{H_2O} \cdot (a_w - a_d) : -2.95077 \cdot 10^{-5} \quad [m / (kg / m^2)]$$

$$B : 10^{-6} \cdot R_{H_2O} \cdot b_w : 1.73276 \quad [m / (K \cdot kg / m^2)]$$

The wet tropospheric delay is in the order of 0.4-0.5 m for high atmospheric water vapour content. The wet tropospheric delay reported in the EMiR L2-files is calculated using Equation (26) with  $T_m$  being calculated from ERA-Interim via Equation (25).

## 6 L2 data output

The revised EMiR V1.1 L2 data record provides swath-based global information on a daily basis. The individual data files contain TCWV retrievals, the background equivalent, and the corresponding retrieval error. Additional data, as for example the underlying brightness temperatures as well as a quality flag, are also part of the EMiR L2 product.

In addition to the data fields already present in the previous EMiR version V1.0, the revised data record contains the following additional information for each observation:

- Satellite cycle and pass number,
- Solar zenith angle,
- Day/night/twilight flag,
- TCWV prior used in the retrieval,
- *A posteriori* retrieval uncertainty of TCWV, liquid water path (LWP) and wet tropospheric correction (WTC).

The EMiR V1.1 L2 files are organised on a per-day basis, i.e. one file contains all EMiR L2 products for one full calendar day (00:00 to 24:00 UTC). The files are provided in netCDF / CF-1.6 format and typically have a size of 4 MB per file.

**Table 5: Contents of EMiR V1.1 Level-2 netCDF files. Variables shown in green background have been newly added to V1.1.**

Variable	Name	Units	Description
Cycle number	cycle_number	unitless	Satellite orbit cycle for each observation
Pass number	pass_number	unitless	Satellite pass number in current orbit cycle for each observation
Time	time	days	Days since 1950-01-01 00:00:00.0
Latitude	lat	degrees N	Geographical latitude (WGS 1984), North positive
Longitude	lon	degrees E	Geographical longitude (WGS 1984), Range 0°-360°, East: 90°
Solar zenith angle	SZEN	degrees	Solar zenith angle at point of observation
Day/night/twilight flag	DNTFLAG	unitless	0: day (SZEN < 90°) 1: night (SZEN > 102°) 2: twilight (SZEN 90°-102°, definition of nautical twilight)
Prior for total column water vapour	TCWV_PRIOR	kg/m <sup>2</sup>	<i>A priori</i> value used for TCWV in RTTOV retrieval
Total column water vapour	TCWV	kg/m <sup>2</sup>	Instantaneous retrieved value
Uncertainty of TCWV	TCWV_UNC	kg/m <sup>2</sup>	<i>A posteriori</i> uncertainty of instantaneous retrieved value
Liquid water path	LWP	kg/m <sup>2</sup>	Instantaneous retrieved value

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

Variable	Name	Units	Description
Uncertainty of LWP	LWP_UNC	kg/m <sup>2</sup>	<i>A posteriori</i> uncertainty of instantaneous retrieved value
Wet tropospheric correction	WTC	m	Instantaneous retrieved value
Uncertainty of WTC	WTC_UNC	m	<i>A posteriori</i> uncertainty of instantaneous retrieved value
Cost function	cost	unitless	Value of cost function for retrieval. <i>Recommended threshold to identify valid retrievals: cost &lt; 5.</i>
Retrieval quality flag	flag	unitless	<ul style="list-style-type: none"> <li>- 1: Retrieval performed</li> <li>- 2: ERS-2 after gain drop. Retrieval performed, gain drift in 23.8 GHz channel likely</li> <li>- 3: ENVISAT initial heating period, retrieval performed</li> <li>- 98: Retrieved values out of range</li> <li>- 99: No retrieval (above sea ice or land)</li> </ul>
23 GHz brightness temperature	Tb23	K	Inter-calibrated instantaneous Tb
36 GHz brightness temperature	Tb36	K	Inter-calibrated instantaneous Tb

## 6.1 Fill values

In a number of cases, gaps in the Level-1 brightness temperatures have been filled with constant values. These are instrument dependent and amount to 323.5 K (23 GHz) and 320.5 K (36 GHz) for ERS-1, 325.2 K (23 GHz) and 324.0 K (36 GHz) for ERS-2, as well as 324.8 K (23 GHz) and 322.1 K (36 GHz) for Envisat (all values rounded to one digit).

These fill values are kept in the EMiR dataset to remain consist with the MWR Level 1 data obtained from CLS and used for product generation (see section 2 for more details). Obviously, Level-2 retrievals are not applied to these data and the corresponding retrievals are set to fill values of -999.

--- Remainder of page intentionally left blank---



ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

## 7 EMiR Level-3 data

The EMiR L2 data files are used to calculate global fields of monthly averages of TCWV, LWP, and brightness temperatures at 23 and 36 GHz on a 2°×2° as well as a 3°×3° latitude-longitude grid.

### 7.1 Level-2 pre-screening

Before calculating monthly means, filters are applied. A Level-2 data pixel is used for L3 product generation if and only if the following conditions are met:

TCWV retrieval >0.0 kg/m<sup>2</sup> AND LWP retrieval > -1.0 kg/m<sup>2</sup> AND cost function value < 5.

The condition on the cost function effectively removes heavily precipitation-contaminated pixels as well as observations with remaining sea ice or land contribution.

### 7.2 Level-3 product generation

EMiR V1.1 L3 products are generated as follows.

- (1) First, a global grid is set up according to the considered spatial resolution. Here, datasets of 2°×2° and 3°×3° latitude / longitude resolution have been established, where the 2°×2° resolution has been created to meet the specific requirements of the GEWEX Water Vapour (GVAP) intercomparison exercise (see [www.gewex-vap.org](http://www.gewex-vap.org)). The 3°×3° resolution offers better temporal coverage and reduced noise and is thus better suited for the analysis of climatic trends.
- (2) For each MWR footprint, the corresponding grid cell is determined from the latitude/longitude information. If the pre-screening conditions are met (see section 7.1 above), retrievals and auxiliary values are added to the grid box. The daily averages are then calculated as the arithmetic mean of all observations within that grid box within one day.
- (3) If for a given month more than 20 daily averages exist, the arithmetic mean of those is assigned to be the monthly mean value. This way, it is ensured that the calculated monthly mean is representative of the true conditions.

### 7.3 Level-3 file structure

The Level-3 output files contain the parameters listed in Table 6. The files are provided in netCDF/CF format and have a size of 57.6 MB for 2°×2°, and of 25.6 MB for the 3°×3° resolution product.

**Table 6: Contents of EMiR V1.1 Level-3 netCDF files. Coordinates refer to the grid-box centre.**

Variable	Name	Units	Description
Time	time	days	Days since 1950-01-01, 00:00:00.0.
Latitude	lat	degrees N	Geographical latitude (WGS 1984)
Longitude	lon	degrees E	Geographical longitude (WGS 1984)

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

Variable	Name	Units	Description
Total column water vapour	TCWV	kg/m <sup>2</sup>	Monthly mean of daily means for 2°×2° or 3°×3° grid cells
Liquid water path	LWP	kg/m <sup>2</sup>	Monthly mean of daily means for 2°×2° or 3°×3° grid cells
23 GHz brightness temperature	Tb23	K	Monthly mean of daily means for 2°×2° or 3°×3° grid cells
36 GHz brightness temperature	Tb36	K	Monthly mean of daily means for 2°×2° or 3°×3° grid cells

--- Remainder of page intentionally left blank---

## 8 Summary

A 1D-VAR retrieval scheme, which is based on *Phalippou [1996]*, *Deblonde [2001]*, and work done in the ESA DUE GlobVapour project, has been optimized and applied in the EMiR project. The scheme has been used to retrieve TCWV over the ice-free oceans from MWR satellite observations. The derived TCWV is an optimal estimate considering the provided background information and the satellite measurements with their associated errors. The derived TCWV values are on L2 satellite swath resolution. The TCWV quality is estimated to be of high standard with respect to the measurements provided. Based on the footprint data, monthly averages are provided on global, latitude/longitude fields of  $2^{\circ} \times 2^{\circ}$  or  $3^{\circ} \times 3^{\circ}$  resolution.

A number of limitations of the Level-2 data exist:

- In the current implementation, ERA-Interim fields are only used once a day for 12 UTC as background profiles. Rapid changes of atmospheric components and surface properties (e.g. diurnal cycle of surface temperature) are not accounted for. We have studied the impact of different assumptions on the background state in Section 4.3 and conclude that this limitation has only a marginal impact on data quality.
- Since MWR is nadir looking only, it does not provide any polarization information. Compared to other microwave sensors, its spectral range is also limited to frequencies below 37 GHz. Therefore, screening observations affected by frozen hydrometeor scattering will not be possible. Thus, in cases of moderate to heavy frozen hydrometeor load, such as in deep convective cores, retrieval results will likely be degraded. Here we employ a screening based on the final value of the cost function which has proven efficient in eliminating outliers.
- The algorithm relies on accurate information about surface wind speed, which drives sea surface emissivity. We are currently using the ERA-Interim surface wind speed. The impact of wind speed variability on retrieval quality can be seen in Figure 2 and is comparably small.
- A joint retrieval of surface wind speed, water vapour, and cloud liquid water based on combined radar altimeter and MWR data appears feasible and will likely have a significant positive impact on retrievals from both instruments. This cannot be addressed in the current project since emissivity model development is needed.
- The exact uncertainties in the forward modelling process are not entirely known. Only an estimation can be given which is included in the observation error covariance matrix. The observation error covariance matrix contains only values at the diagonal elements. Off-diagonal elements are currently set to zero.

ERS/Envisat MWR recalibration					
EMiR V11 L2/L3 ATBD, V2.2					

## 9 References

- Ablain, M., A. Cazenave, G. Valladeau, and S. Guinehut, 2009: A new assessment of the error budget of global mean sea level rate estimated by satellite altimetry over 1993-2008, *Ocean Science*, 5(2), 193-201.
- Berry, P. A. M., R. Smith, and J. Benveniste, 2008: ACE2: the new Global Digital Elevation Model, *IAG International Symposium on Gravity, Geoid & Earth Observation 2008*, Chania, Greece, 23-27th June 2008.
- Bormann, N., A. J. Geer, and S. J. English, 2012: Evaluation of the microwave ocean surface emissivity model FASTEM-5 in the IFS, Rep. 667, 18 pp, ECMWF, Reading, UK.
- Blankenship, C. B., A. Al-Khalaf, and T. T. Wilheit, 2000: Retrieval of Water Vapor Profiles Using SSM/T-2 and SSM/I Data. *J. Atmos. Sci.*, 57, 939–955.
- Bretherton, C. S., M. E. Peters, and L. E. Back, 2004: Relationships between water vapor path and precipitation over the tropical oceans, *Journal of Climate*, 17(7), 1517-1528.10.1175/1520-0442(2004).
- Cess, R. D., et al. (1990): Intercomparison and interpretation of climate feedback processes in 19 atmospheric general-circulation models. *JGR-Atmospheres*, 95(D10), 16601-16615.10.1029/JD095iD10p16601.
- CMSAF, 2009: ATBD for total column water vapour retrieval from SSM/I, SAF/CM/DWD/ATBD/HTW\_SSMI, 6 January 2009.
- Deblonde, G., 2001: NWP SAF user's guide: Standalone 1D-VAR scheme for the SSM/I, SSMIS and AMSU, NWPSAF-MO-UD-001 Version 1.0, 22 August 2001.
- Deblonde, G., and S. J. English, 2001: Evaluation of the FASTEM-2 fast microwave oceanic surface emissivity model. Tech. Proc. ITSC-XI Budapest, 20-26 Sept 2000, 67-78.
- English, S. J., and T. J. Hewison, 1998: Fast generic millimetre-wave emissivity model, Proc. SPIE 3503, 288.
- ESA Earth Online, 2015: <https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat/instruments/mwr>.
- Eymard, L., E. Obligis, N. Tran, F. Karbou, and M. Dedieu (2005): Long-term stability of ERS-2 and TOPEX microwave radiometer in-flight calibration, *IEEE Transactions on Geoscience and Remote Sensing*, 43(5), 1144-1158.10.1109/tgrs.2005.846129.
- GlobVapour, 2012: ATBD for L2 SSM/I and MWR, issue 3, revision 0, 19 January 2012. Available from <http://www.globvapour.info/>.
- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Met. Soc. Japan*, Ser. II, Vol. 75 (1997), No. 1B, 181-189.
- IPCC-AR4, 2007: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Edited by: S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.
- Liebe, H. J., 1985: An updated model for millimetre wave propagation in moist air. *Radio Science*, 20, 1069-1089.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 112, 1177–1194.
- Phalippou, L., 1996: Variational retrieval of humidity profile, wind speed and cloud liquid-water path with the SSM/I: Potential for numerical weather prediction. *Q. J. R. Meteor. Soc.*, 122, 327-355.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1989: Numerical Recipes in Pascal; The Art of Scientific Computing, Cambridge University Press, Cambridge.
- Rodgers, C. D., 1976: Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys.*, 14(4), 609–624.
- Saunders, R., M. Matricardi, and A. Geer, 2008: RTTOV 9.1 Users Guide, NWP SAF report, Met. Office, 57 pp.

ERS/Envisat MWR recalibration	  	 
EMiR V11 L2/L3 ATBD, V2.2		

Tarantola, A., and B. Valette, 1982: Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, 20(2), 219–232.

Trenberth, K. E.; Fasullo, J. and Smith, L. Trends and variability in column-integrated atmospheric water vapor. *Clim. Dyn.* 2005, 24, 741-758.