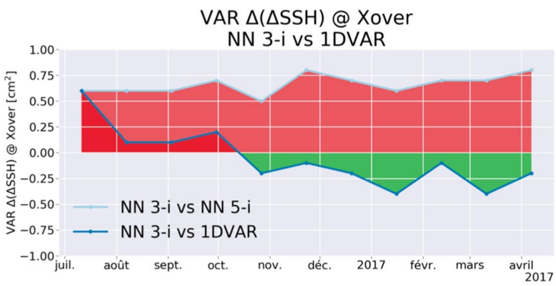

Comparison of the AMTROC 1D-VAR wet tropospheric correction (WTC) against Sentinel-3 standard WTC products. See section 5.6 in this report for further details.

EUMETSAT contract

EUM/CO/19/4600002223/EO

Altimeter 1D-VAR Tropospheric Correction (AMTROC)

Product Validation and Evolution Report

(Deliverables D-10, D-13)

Berlin, 30. March 2020

Document control sheet	
Project	Altimeter 1D-VAR Tropospheric Correction (AMTROC)
Client	EUMETSAT Eumetsat-Allee 1 64295 Darmstadt, Germany http://www.eumetsat.int
Main contractor	Informus GmbH Brehmestr. 50 13187 Berlin, Germany http://www.informus.de
Other project partners	Fluctus SAS 54, Promenade des Lices 81800 Rabastens, France http://satobsfluctus.eu/
Title	Validation Report
Status	Released
Authors	Ralf Bennartz (VU), Frank Fell (INF), Bruno Picard (FLU)
Distribution	AMTROC consortium, EUMETSAT consortium
Version control	2019-12-02 V0.60, initial draft 2020-03-24 V0.90, final draft 2020-03-30 V1.00, initial version

Table of contents

List of tables

List of figures

Summary

A newly implemented 1D-VAR scheme (AMTROC 1D-VAR) to retrieve total column water vapour (TCWV) and wet tropospheric correction (WTC) over the global ice-free ocean has been applied to observations made by the Microwave Radiometer (MWR) onboard Sentinel-3A (S3A) for the period 15th June 2016 to 15th April 2017. This report presents and analyses the validation efforts of the corresponding AMTROC 1D-VAR TCWV and WTC retrievals.

(1) *A novel bias correction approach for MWR-like radiometers has been established allowing for the correction of remaining biases in instrument calibration as well as for the correction of biases related to the surface emissivity model.*

(2) *The bias correction has been evaluated for MWR and independently also for the Advanced Technology Microwave Sounder (ATMS). Operational bias-monitoring is available for the latter from the European Centre for Medium-Range Weather Forecasts (ECWMF), so that the AMTS bias correction results obtained herein can be compared to their operational ECMWF counterparts:*

   a) *The AMTROC bias correction for ATMS is in line with operational bias monitoring from ECWMF.*

   b) *MWR on S3A shows generally higher biases than ATMS, suggesting room for further improving MWR's calibration.*

   c) *Comparing retrievals with and without the application of prior bias correction, it becomes clear that bias correction is crucial to achieve good retrieval accuracy.*

(3) *An analysis of failed AMTROC 1D-VAR retrievals was performed:*

   a) *About 2.6 % of 1D-VAR retrievals over ocean have failed but had concomitant ANN retrievals from the operational S3A processing.*

   b) *The vast majority of those failed 1D-VAR retrievals are associated with thick precipitating clouds.*

   c) *While ANN retrievals for those cases exist, these retrievals show unrealistically large TCWV values.*

(4) *1D-VAR TCWV and WTC retrievals at Level-2 (i.e. from individual observations) were assessed against the S3A operational (ANN) retrievals and collocated ECMWF Reanalysis (ERA) results:*

   a) *1D-VAR and ANN produce comparable results in terms of bias and RMSE deviations with respect to ERA.*

   b) *Particular attention was paid to the impact of auxiliary input parameters to the 1D-VAR retrieval, including sea surface temperature, surface pressure, and the mean atmospheric temperature. If those auxiliary input parameters were taken from climatology instead of observations, the 1D-VAR retrieval results deteriorated, but the performance was still slightly better than that of the ANN.*

   c) *The 1D-VAR retrieval showed lesser dependency on surface wind speed than did the ANN retrievals.*

(5) AMTROC 1D-VAR and S3A ANN TCWV retrievals were validated at Level 3 (i.e. spatially and temporally averaged) against the "Merged Total Precipitable Water 1-deg Monthly Climate Product" [1] *from Remote Sensing System (RSS) and gridded ERA-Interim and ERA-5 data.*

a) The 1D-VAR and ANN retrievals produce comparable results in terms of bias and RMSE deviations with respect to ERA at monthly mean scales.

b) The differences among TCWV climatologies are larger than the differences between the 1D-VAR and ANN retrievals.

(6) 1D-VAR WTC retrievals were validated against the operational S3A (ANN) WTC retrievals using crossover analysis:

a) WTC performances of 1D-VAR and ANN are comparable, with the 1D-VAR performing slightly better than the ANN at the end of the time period covered by this study.

b) The 5-input ANN performance is slightly inferior to that of the 3-input ANN. We believe this to be caused by the use of climatologies for central input parameters of the former.

c) Referring to the above, the 1D-VAR retrieval also degrades if climatological input is used for auxiliary parameters.

In summary, for all validation exercises reported herein, the 1D-VAR retrieval approach performs as good as or slightly better than the ANN retrieval. Furthermore, the 1D-VAR retrieval provides two additional significant advantages over statistical retrievals:

- *It allows for realistic and physically traceable uncertainty estimates associated with all parameters. This allows for a better characterization of retrievals and their accuracy.*
- *It allows to clearly separate different types of input parameters (observations, background, first guess, and first guess) with respect to their impact on the retrieval.*

While the input parameters can obviously also be fed into statistical retrievals, the latter do not provide the opportunity to clearly separate these parameters in terms of their impact. Thus, the dependency of statistical retrievals on auxiliary input parameters is difficult to disentangle from other types of retrieval errors.

The recommendations arising from the AMTROC validation activities can be summarised as follows:

(1) The AMTROC 1D-VAR retrieval scheme should be considered as a candidate algorithm for future operational retrievals of TCWV and WTC.

(2) Regardless of which retrieval is used (1D-VAR or ANN), a thorough bias-correction of the underlying brightness temperatures is crucial to the success of any retrieval.

(3) Precipitation screening is crucial for the improvement of WTC and TCWV estimates under precipitating conditions. The addition of high-frequency channels (>80 GHz) will be beneficial in that respect.

[1] See: http://www.remss.com/measurements/atmospheric-water-vapor/tpw-1-deg-product/

1 Introduction

1.1 Purpose of the document

The purpose of the present *"Product Validation and Evolution Report"* is to describe the validation efforts undertaken in the context of the AMTROC study, during which a newly implemented 1D-VAR based retrieval scheme to derive TCWV and WTC over the ice-free global oceans has been applied to ten months (June 2016 to April 2017) of top-of-atmosphere brightness temperature observations from the Sentinel-3A Microwave Radiometer (MWR).

1.2 Context

Sea surface height (SSH) retrievals from Sentinel-3 Synthetic Aperture Radar Altimeter (SRAL) observations need to be corrected for the effects of atmospheric moisture, a process often termed as "Wet Tropospheric Correction" (WTC).

To allow for such corrections, each satellite of the Sentinel-3 series carries a dedicated instrument, the Microwave Radiometer (MWR), to determine total column water vapour (TCWV) and liquid water path (LWP) concomitantly to the altimeter observations, day and night, under cloudy and clear sky conditions. The such gained knowledge on TCWV and LWP is then further processed to provide precise information on the WTC for the area observed by the altimeter.

The TCWV and WTC products currently distributed as part of the standard Sentinel-3 product suite (AD-2, AD-3) are affected by a number of limitations:

(1) The use of an artificial neural network (ANN) based retrieval scheme (AD-4) requires re-training the algorithm in case of changing instrument biases or other variables.

(2) Rigorous uncertainty estimates cannot be provided.

(3) The incorporation of appropriate background knowledge (as e.g. provided by model forecasts or analysis) is not possible. This is particularly important if the atmospheric stratification does not adhere to the profiles used for neural network training.

1DVAR retrieval schemes are not affected by the above-mentioned limitations so that they could constitute an attractive alternative or complement to the operational ANN-based retrievals if their retrieval accuracies are at least on a comparable level.

In this context, EUMETSAT has funded the activity *"Altimeter 1D-VAR Tropospheric Correction"* (AMTROC) [AD-AMTROC-SSOW] as part of the *"Level 2 Product Evolution/Development Studies"* [AD-AMTROC-GSOW], specifically asking for the development of a 1DVAR-based TCWV retrieval scheme applicable to Sentinel-3 MWR observations as a basis for an improved WTC.

The AMTROC retrieval scheme validated herein builds on a 1DVAR approach that has been developed in the context of the ESA-funded EMiR study [EMiR-FINREP] for the very similar MWR instruments flown onboard ERS-1/2 and Envisat. This earlier approach has been improved on a number of aspects as compared to the EMiR implementation, such as updating to latest version of the supporting software packages 1D-VAR and RTTOV, implementation of an improved TB bias correction, use of sigma_0 for sea surface roughness estimation, etc.

This report does not provide a detailed description of the 1DVAR method applied herein. Please refer to the AMTROC Algorithm Theoretical Basis Document [AD-AMTROC-ATBD] for further information in this respect.

1.3 Validation strategy

The AMTROC validation strategy has been adapted to the relatively limited scope of the AMTROC study. To make best use of available resources, comparison was done against readily available reference data. The following validation steps have been performed within AMTROC:

- Comparison of bias-corrected S3A-MWR brightness temperatures against those from a potentially better calibrated instrument (ATMS).
- Comparison of AMTROC-derived TCWV and WTC against the corresponding standard Level-2 products.
- The validation efforts have been limited to a global analysis.

The following validation steps could NOT be carried out in AMTROC but should be attempted for potential follow-on activities:

- Regional cases studies targeting potential specific weaknesses of the AMTROC retrieval schemes.
- Comparison of AMTROC TCWV against TCWV derived from other space-based observations, preferably based on differing retrieval mechanisms.
- Validation of AMTROC TCWV against reference in situ (GNSS) TCWV.
- Validation of the AMTROC LWP product.

Further validation of the AMTROC products will also be done externally in the context of thematically related activities, for example:

- Companion study *"H2O from S3 SLSTR (AIRWAVES)"* in the context of the EUMETSAT *"Level 2 Product Evolution/Development Studies"* dealing with improving the S3-SLSTR TCWV retrievals.
- Validation project in the context of the Sentinel-3 Validation Team (S3VT), aiming at comparing SLSTR, OLCI, and MWR derived TCWV products.

1.4 Acronyms and abbreviations

Acronym	Description
1D-VAR	One-dimensional Variational Data Assimilation
AMTROC	Altimeter 1D-VAR Tropospheric Correction
ANN	Artificial Neural Network
ATBD	Algorithm technical basis document
ATMS	Advanced Technology Microwave Sounder
CODA-REP	Copernicus Online Data Access – REProcessed
ECMWF	European Centre for Medium-Range Weather Forecasts
EMiR	ERS/Envisat MWR Recalibration and Water Vapour FDR Generation
ERA (-5)	ECMWF Reanalysis (5th Generation)

Acronym	Description
ERS	European Remote Sensing satellite
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
GEOSS	Global Earth Observation System of Systems
GNSS	Global Navigation Satellite System
GSOW	Generic Statement of Work
JPSS	Joint Polar Satellite System
L1	Level-1 processing
LWP	Liquid water path
MWR	Microwave Radiometer
NWP	Numerical weather prediction
OLCI	Ocean and Land Colour Instrument
QA4EO	Quality Assurance for Earth Observation
QI	Quality indicator
S3	Sentinel-3
S3VT	Sentinel-3 Validation Team
SLSTR	Sea and Land Surface Temperature Radiometer
SNPP	Suomi National Polar-orbiting Partnership
SRAL	SAR Radar Altimeter
SSOW	Specific Statement of Work
SSH	Sea surface height
SSM/I	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager/Sounder
SST	Sea surface temperature
SWS	Surface wind speed
TOA	Top of atmosphere
TCWV	Total column water vapour
WTC	Wet tropospheric correction
WVP	Water vapour path

1.5 Applicable documents

[AD-AMTROC-ATBD] AMTROC Consortium, 2019: Algorithm theoretical basis document for AMTROC products (Deliverable D-7), Version 0.9, 26. November 2019.

[AD-AMTROC-GSOW] EUMETSAT, 2018: Generic Statement of Work for Level 2 Product Evolution/Development Studies, Document EUM/TSS/SOW/18/1018464, Issue v1A, 4. September 2018.

[AD-AMTROC-SSOW] EUMETSAT, 2018: Statement of Work for Altimeter 1-D VAR Tropospheric Correction, Document EUM/RSP/SOW/18/1001261, Issue v1, 18 June 2018.

[AD-QA4EO-PRNCPL] QA4EO Task Team, 2010: A Quality Assurance Framework for Earth Observation: Principles, Version 4.0, 14 January 2010, http://qa4eo.org/docs/QA4EO_Principles_v4.0.pdf.

[PDF_S3_SRAL_HANDBOOK] EUMETSAT, 2017: Sentinel-3 SRAL Marine User Handbook, Document EUM/OPS-SEN3/MAN/17/920901, Issue v1A, 12 December 2017. Available online.

2 Community guidelines to validation

The Quality assessment for Earth observation (QA4EO) framework formulates three guiding principles to enhance the quality and usefulness of Earth observation data products *[AD-QA4EO-PRNCPL]*. In the following subsections, we present these principles and analyse in how far the AMTROC data products adhere to those.

2.1 QA4EO guiding principle for data quality

All data and derived products must have associated with them a quality indicator (QI) based on documented quantitative assessment of its traceability to community agreed (ideally tied to SI) reference standards.

*Does AMTROC comply? **Yes**.*

Due to the nature of the 1DVAR retrieval scheme, all retrieved individual values are inherently assigned a corresponding uncertainty value.

2.2 QA4EO guiding principle for data management

The data product must be freely and readily available / accessible / useable in an unencumbered manner for the good of the GEOSS community, for both current and future users. This necessitates that all EO data and associated support information (metadata, processing methodologies, QA, etc.) is associated with the means to effectively implement a quality indicator. In return, the provider must be consistently acknowledged.

*Does AMTROC comply? **Potentially yes**.*

AMTROC has applied a new retrieval to a limited period of Sentinel-3A MWR observations. While we judge the retrieval as successful (see validation results presented herein) and the provided quality information as comprehensive, we consider the data products generated under AMTROC as not yet mature and complete enough to be publicly distributed (see recommendations for further improvements provided herein). Ultimately, any future publishing of the AMTROC data record is at the discretion of the project owner (i.e. EUMETSAT).

2.3 QA4EO guiding principle for documentation management and outreach

Sound and effective harmonised documentation management is needed to facilitate and enhance interoperability and achieve the objectives of consistent and traceable quality information. To enable this activity, all stakeholders must have a clear understanding of the adequacy of the information that they are accessing and using for their specific application. The evidence for this clarity should ideally be accessible through a centralised portal and should be fully traceable to its origins. The traceability and interoperability process must be understandable by any appropriately trained individual within GEOSS and efforts must be made to encourage the wider usage of information and facilitate the training of GEOSS users.

*Does AMTROC comply? **Potentially yes**.*

The above principle does currently not apply to the AMTROC data product since it has not yet been made publicly available. Notwithstanding this, the AMTROC documentation has been designed to convey adequate and concise information on product generation (through the AMTROC ATBD, *[AD-AMTROC-ATBD]*) and quality aspects (through the present Validation Report, *[AD-AMTROC-VALREP]*).

3 Validation data

3.1 The MWR instrument

The Sentinel-3 Microwave Radiometer (S3 MWR) is a two-channels noise injection microwave radiometer. Following the heritage of ERS-1, ERS-2 and Envisat, it operates at 23.8 GHz to observe atmospheric water vapour and at 36.5 GHz to record the presence of atmospheric liquid water.

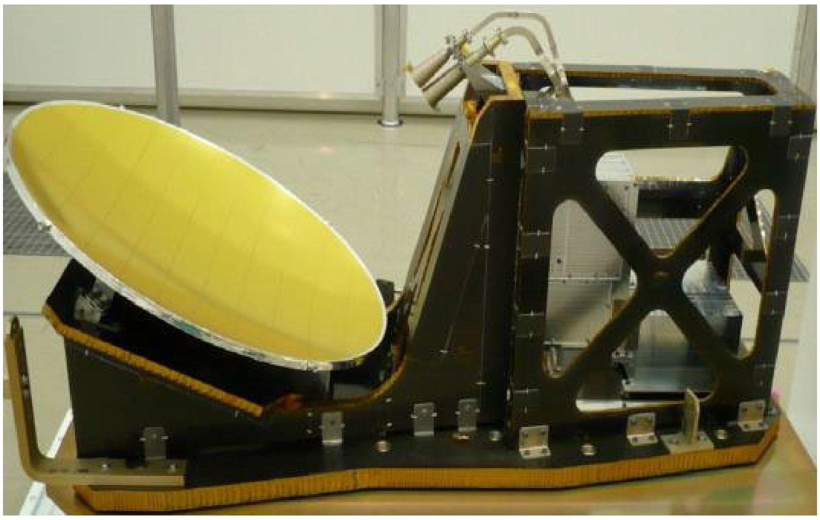

Figure 1: Photo of the MWR antenna plus two feed horns. Source: Sentinel-3 SRAL Marine User Handbook *[PDF_S3_SRAL_HANDBOOK]*.

By using feed horns that are not directly on the boresight of the antenna, the 24 km diameter footprint at 23.8 GHz is located 28 km in front of the sub-satellite point and the 18.5 km diameter footprint at 36.5 GHz is located 27 km behind. Thus, the time series of brightness temperatures at the two frequencies have to be shifted to match the spatial locations of the altimeter measurements.

MWR brightness temperatures are used to infer the amount of water vapour and liquid water in the sub-satellite atmospheric column, and to subsequently calculate the Wet Tropospheric Correction (WTC), i.e. the correction to the range, and the atmospheric attenuation (correction to σ_0) to support altimetry observations. Table 2 contains a summary of MWR instrument specifications.

Table 1: Key characteristics of the MWR flown on the Sentinel-3 series of satellites. Source: Sentinel-3 SRAL Marine User Handbook *[PDF_S3_SRAL_HANDBOOK]* and https://www.wmo-sat.info/oscar/instruments/view/348.

Parameter	23.8 GHz	36.5 GHz
Bandwidth	200 MHz	200 MHz
Integration time (typical)	152.88 ms	152.88 ms
Polarization	linear	linear
Noise (at 25 °C)	<4.4 dB (main path)	<5.1 dB (main path)
Radiometric sensitivity (main path, NIR mode)	0.29 K	0.34 K
Radiometric accuracy	<3 K	<3 K
Radiometric stability	0.6 K	0.6 K
Main reflector size (projected diameter)	0.6 m	
Scanning technique	Nadir-only viewing	
Calibration	Noise injection Dicke radiometer configuration with a separate deep space viewing sky horn to provide cold reference at 50% and 100% noise injection. Dedicated instrument calibration sensors.	
Mass / power / data rate	24.2 kg / 26 W / 5 kbps	
Utilisation period	2016-07-13 to ≥2032	

3.2 MWR observations and pre-processing

For the purpose of the AMTROC study, the Reprocessing 2 dataset has been downloaded from the Copernicus Online Data Access – REProcessed (CODA-REP) server[2] operated by EUMETSAT. Covering the period from 15th of June 2016 (cycle 5, orbit 187) to 15th April 2017 (cycle 16, orbit 300), this Level-2 Marine Product dataset has been reprocessed with the IPF version IPF-SM-2 06.12, corresponding to the S3A Processing Baseline 2.27[3].

Focusing on MWR, the main evolution included in this reprocessing compared to the initial processing is the improvement of the 5-input Neural Network wet tropospheric correction (WTC) solution (see section 3.6.2 for additional details), associated with a reduction of the standard deviation of the difference between this retrieval and the ECMWF WTC [4].

As the AMTROC 1D-VAR retrieval scheme is only applicable to the global ice-free ocean, land measurements are discarded. No specific processing is applied in coastal areas so that contamination

[2] https://codarep.eumetsat.int/#/home (registration required)

[3] Sentinel-3 STM Annual Performance Report - Year 2018, available at https://sentinel.esa.int/documents/247904/3519647/Sentinel-3-STM-Annual-Performance-Report-2018

[4] Sentinel-3 STM Product Evolution for Processing Baseline 2.24, available at https://sentinel.esa.int/documents/247904/3147059/Sentinel-3-STM-Product-Evolution-Processing-Baseline-2.24

from land may occur above coastal waters. Such potentially land-contaminated pixels are excluded from the analysis presented herein by rejecting any observation less than 100 km offshore.

3.3 ATMS observations and pre-processing

The Advanced Technology Microwave Sounder (ATMS) is a 22-channel cross-track scanning passive microwave radiometer flown onboard the satellites of the US Joint Polar Satellite System (JPSS). The data used here stem from the Suomi National Polar-orbiting Partnership (SNPP) satellite. In order to create a dataset comparable to the MWR dataset, the same time period as for MWR was selected (June 2016 – April 2017) and only nadir observations for the two lowest frequency channels (23.8 and 31.0 GHz) were evaluated. Land and coastal measurements were discarded in the same way as for MWR. Note, that the second ATMS channel has a slightly different spectral position than the corresponding MWR channel (31.0 GHz versus 36.5 GHz). A direct comparison of the MWR and ATMS absolute brightness temperature statistics is therefore only possible for the 23.8 GHz channel. Still, the same bias-correction method can be applied to all channels as the information content at 31 and 36.5 GHz is very similar. Due to its cross-track scanning concept, ATMS is easier to calibrate than the purely nadir-looking MWR and its calibration accuracy is well known. *Weng and Yang [2016]* find the accuracy of the ATMS to be within a range of ±0.4 K over a wide range of brightness temperatures.

3.4 Atmospheric profiles and surface parameters

ERA-Interim *[Dee et al., 2011]* is a third generation global atmospheric reanalysis provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). It improves on previous versions (e.g. ERA-40) by using an enhanced atmospheric model and assimilation system (ECMWF, 4D-VAR, 2006). The spatial resolution of the data set is approximately 80 km (T255 spectral) on 60 vertical levels from the surface up to 0.1 hPa. Among others, it assimilates significant amounts of satellite radiances, (e.g. SSM/I and SSMIS). ERA-Interim data are available since 1979 and were continuously updated through mid-2019. For this study, 6-hourly surface as well as vertically resolved moisture and temperature fields were used (downloaded from https://rda.ucar.edu/datasets/ds627.0/). Each MWR and ATMS observation was matched to the closest ERA-Interim reanalysis profile, leading to maximum time differences of 3 hours between the observations and the reanalysis. The such extracted profiles, together with the concomitant sea surface temperature and surface wind speed formed the basis of the subsequent cloud-free radiative transfer simulations.

3.5 Radiative transfer simulations

Radiative transfer simulations were carried out using the TIROS Operational Vertical Sounder Radiative Transfer (RTTOV, Version 12.2) Model *[Saunders et al., 2007; Hocking et al., 2011; Saunders et al., 2018]*, which is widely used both in the operational and research community. All physical parameterizations including absorption by gases and liquid water and rough ocean surface scattering are used unchanged from RTTOV. Radiative transfer simulations were carried out for the MWR channels at 23.8 and 36.5 GHz as well as for the corresponding ATMS channels at 23.8 and 31.0 GHz.

3.6 MWR retrievals of TCWV and WTC

3.6.1 AMTROC 1D-VAR retrieval of TCWV and WTC

The 1D-VAR approach applied in AMTROC is described in detail in *[AD-AMTROC-ATBD]*. The geophysical parameters retrieved are listed in Table 2.

Table 2: Geophysical parameters derived from observations of the MWR instrument flown onboard the Sentinel-3 series of satellites in the frame of the AMTROC project. Due to the nadir viewing observation geometry, a temporal resolution can only be specified for spatially and temporally averaged Level-3 products.

AMTROC parameter	Spatial resolution	Remark
Total column water vapor (TCWV)	Ca. 20 km	Nadir viewing instrument
Liquid water path (LWP)	Ca. 20 km	See above
Wet tropospheric correction (WTC)	Ca. 20 km	See above

3.6.2 Operational WTC product: ANN and model-based WTC

Five wet tropospheric corrections solutions are provided in the Sentinel-3 standard Level-2 Marine Products. The historical retrieval is a three-input neural network solution (3-i NN) based on the two MWR brightness temperatures (TB) and the altimeter backscattering coefficient (σ_0). Details on set-up and training of these neural networks can be found in *[Obligis et al., 2006]*.

A more elaborated solution is based on a five-input neural network (5-i NN), using two additional input parameters, Sea Surface Temperature (SST) and the atmospheric temperature lapse rate (γ_{800}) *[Obligis et al., 2009]*. Both latter parameters are offered in the form of static gridded maps derived from climatologies.

Both 3-i NN and 5-i NN solutions are applied twice, using different sources for the altimeter σ_0: first the Synthetic Aperture Radar mode (σ_0 SAR) and second the Pseudo Low-Rate Mode (σ_0 PLRM). For this study, the SAR mode version of the WTC has been used.

Finally, a fully model-based WTC is also available. Here, the WTC is computed from operational ECMWF analysis using a temporal linear interpolation between two subsequent 6 h analyses and a bilinear spatial interpolation at the location of each altimeter observation.

3.7 Uncertainty considerations

One advantage of the proposed 1D-VAR framework over other methods is that it provides *a posteriori* uncertainties for all retrieved parameters at pixel level. These uncertainties include contributions from all sources of uncertainty in the retrieval process including instrument noise, forward model error, and representativeness errors. Uncertainty estimates will be made available as part of the output data for each observation in accordance with the Q4EO recommendations.

The availability of such uncertainty estimates taking all potential error sources explicitly into account allows for the identification of retrievals meeting application-specific quality requirements.

4 Validation results

4.1 Brightness temperatures

4.1.1 Brightness temperature bias sources

In the context of retrieval studies, two different sources of bias need to be addressed:

(1) Systematic errors associated with the calibration of the passive microwave radiometer under consideration. Those biases might be caused by imperfect knowledge of the instruments' characteristics, instrument drift, or any other variables that directly affect the instrument calibration.

(2) Systematic errors in the forward radiative transfer model used, including 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.

While the second source of bias is not caused by the instrument itself, it will result in retrieval biases similar to those caused by the instrument. Therefore, any correction for the purpose of retrieval studies does not necessarily need to separate the two error sources, as long as it is capable of effectively correcting for their combined effect.

In the following, we first discuss the observed and simulated brightness temperature statistics used in this study. We then propose a method that allows for correcting the cumulative effect of the two types of biases discussed above.

Comparing ATMS results with MWR results, we further estimate the magnitude of the actual instrument calibration bias. This estimation hinges on the assumption that the absolute calibration of ATMS with respect to instrument biases is accurate. As outlined above, this assumption is justified based on absolute calibration accuracy studies of ATMS, such as the one performed by *Weng and Yang [2016]*.

4.1.2 Comparison of all-sky observed with simulated cloud-free brightness temperatures

Figure 2 and Figure 3 show comparisons between observed all-sky TBs (OBS) with simulated cloud-free TBs (SIM) for 23.8 GHz (ATMS and MWR), 31.0 GHz (ATMS) and 36 GHz (MWR). The following can be observed:

– A large tail of warm TBs exists where OBS-SIM is significantly positive, often in excess of 50 K. Those cloud liquid water affected observations are not considered in the bias correction.

– One can see that most of the data lie on a line parallel, but offset, against the zero difference (OBS-SIM) line, *i.e.* the 'zero-bias line'. Figure 3 zooms in on that feature and compares ATMS with MWR brightness temperatures for 23.8 GHz.

– Cloud-free observations fall around the zero line with some noise caused by the instrument. One can see from Figure 2 (a) and (c) that ATMS TB biases lie closer to the zero bias line.

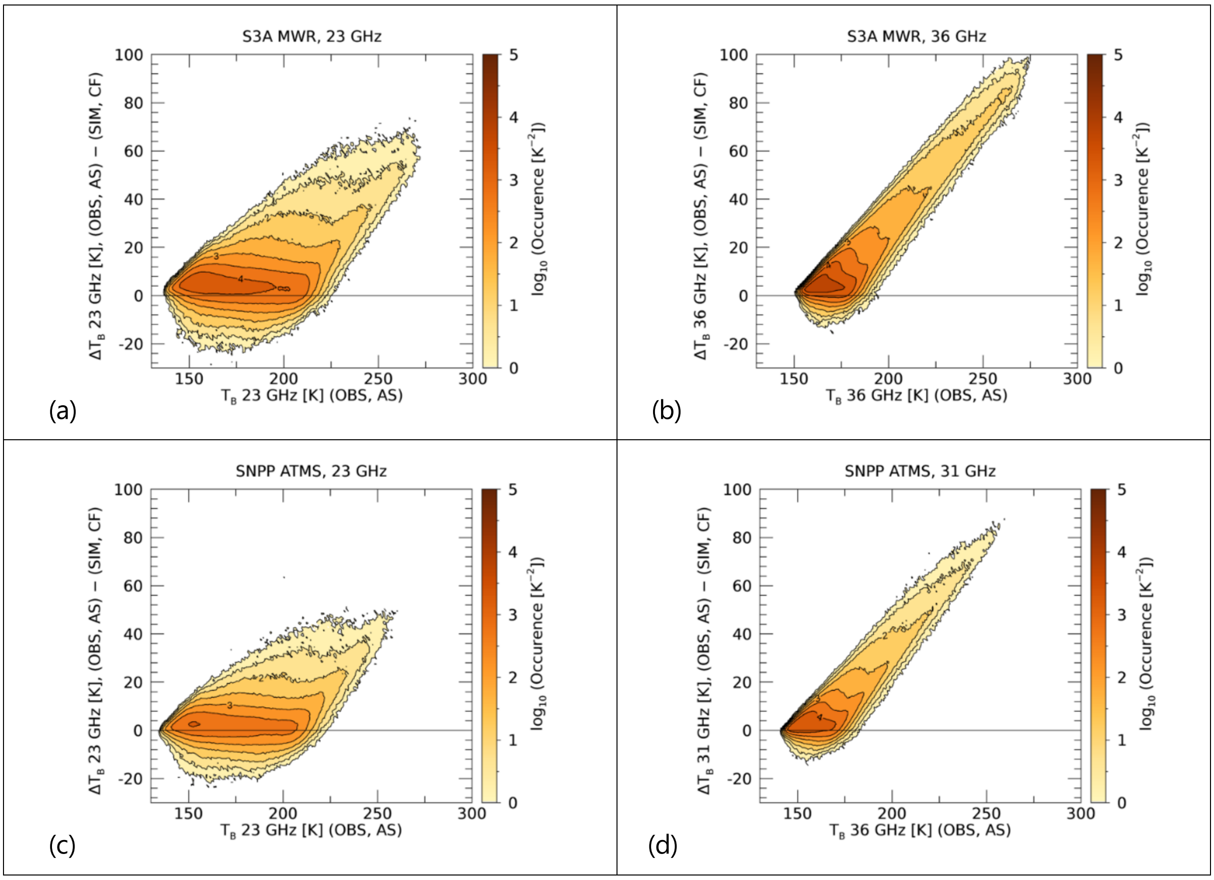

Figure 2: Two-dimensional histograms of observed all-sky minus simulated cloud-free brightness temperatures as function observed all-sky brightness temperatures for (a) S3A MWR 23 GHz, (b) S3A MWR 36 GHz, (c) SNPP ATMS 23 GHz, and (d) SNPP ATMS 31 GHZ. Acronyms in axis labels: 'OBS': observed, 'SIM': simulated, 'AS': all-sky, 'CF': cloud-free. Both datasets are for the open ocean, span the time period June 2016 until April 2017, cover latitudes between 55° N and 55 °S, and are at least 100 km offshore. The total number of observations is 14 million for MWR and 4.3 million for ATMS.

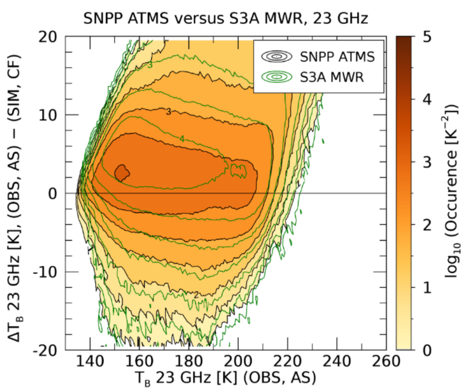

Figure 3: Same as Figure 2 (c) but focussing on the areas around zero bias. In addition to SNPP ATMS, the S3A MWR 23 GHz isolines are shown in green (see Figure 2 (a)).

4.2 Data selection for bias correction

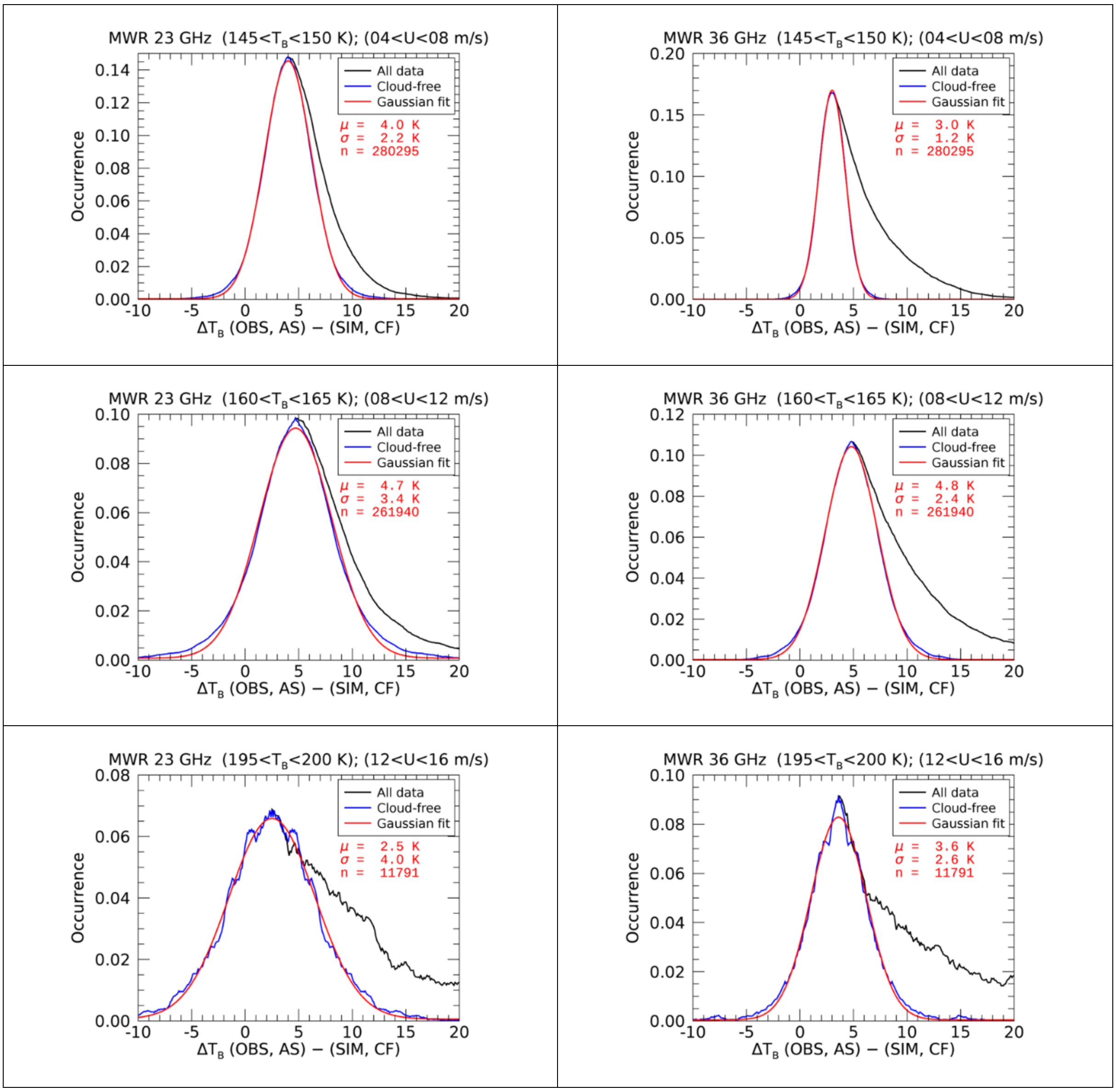

Figure 4: Histograms of differences between all-sky observed TBs and cloud-free simulated TBs for several T_B/u_{10} data slices. The black curves show the observations, the blue curves show all observations to the left of the peak of the histogram and the same observations mirrored to the right. The red curve represents the Gaussian fit to the corresponding blue curve.

We use the feature around the zero-bias line described above to determine the bias for each instrument as follows:

- We subdivide the dataset into 5 K TB intervals between 135 K and 200 K for 23.8 GHz and between 150 K and 175 K for 31.0/36.5 GHz. We further subdivide the dataset into 4 m/s wind speed intervals from collocated altimeter data.

- For these data slices, we derive histograms of OBS-SIM, examples of which are shown in Figure 4 as black curves.

- We next identify the peak of the histogram which for the TB ranges selected coincides with cloud-free data. We then postulate that all values to the left of this peak is cloud-free.
- The right-hand side of the peak of the histogram is increasingly influenced by cloudy situations, as can be identified by the super-Gaussian tail of the black curves.
- Therefore, we mirror the left-hand side of the histogram onto the righthand side (blue curves).
- We next fit a Gaussian to the blue curve (red curve). The mean value of the Gaussian is the bias derived for the considered data slice of TB and wind speed.
- We derive the bias for all TB and wind speed ranges for all considered frequencies.
- We then fit a function of the form:

$$\Delta T_B = a_0 + a_1 T_B + a_2 u_{10} + a_3 u_{10}^2 \qquad (1.1)$$

  where ΔT_B represents the brightness temperature bias, T_B the absolute brightness temperature, and u_{10} the wind speed ten meter above the sea surface.
- The coefficients $a_0, \ldots, a_3$ of the fitting function are given in Table 3. The bias correction according to Eq. (1.1) is also visualised in Figure 5.
- This function allows for the calculation of the TB bias depending on wind speed and absolute TB value, capturing the different types of biases discussed above in Section 4.1.

Table 3: Coefficients to calculate the bias correction according to Equation (1.1).

Instrument	Frequency	a0	a1	a2	a3
MWR	23.8	7.21019	-0.0281476	0.149281	0.00831931
MWR	36.5	3.03748	-0.00972400	0.321376	0.000516656
ATMS	23.8	6.30137	-0.0340420	0.206424	-0.00594865
ATMS	31.0	4.03781	-0.0208938	0.194046	0.00342436

4.3 Bias correction for cloud-free cases

Using the methodology described above in section 4.2, we have derived brightness temperature biases ΔT_B for both MWR and ATMS for the 23.8 GHz, 31.0 GHz, and 36.5 GHz channels (see Figure 5). The following can be observed:

- ATMS TB biases are in general smaller than those of MWR. The ATMS calibration appears thus to be more in line with the forward model than does the MWR calibration.
- TB biases show a strong dependency on wind speed in both datasets. This bias is likely associated with forward modelling errors.
- The ATMS biases obtained here are generally in good agreement with those reported e.g. in ECMWF's operational bias monitoring, an example of which is shown in Figure 6.

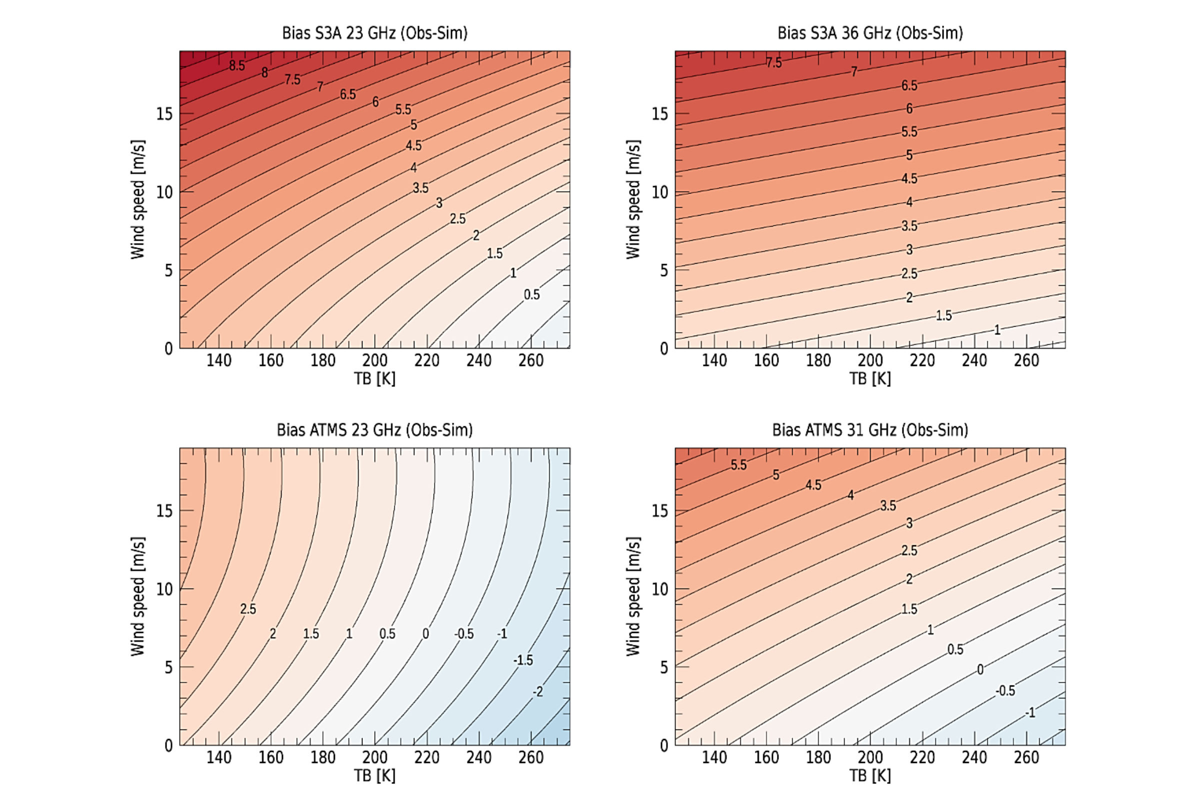

Figure 5: Brightness temperature biases for ATMS and MWR channels as derived from the method described in Section 4.3. These biases will be subtracted from the observed TBs in the retrievals discussed below.

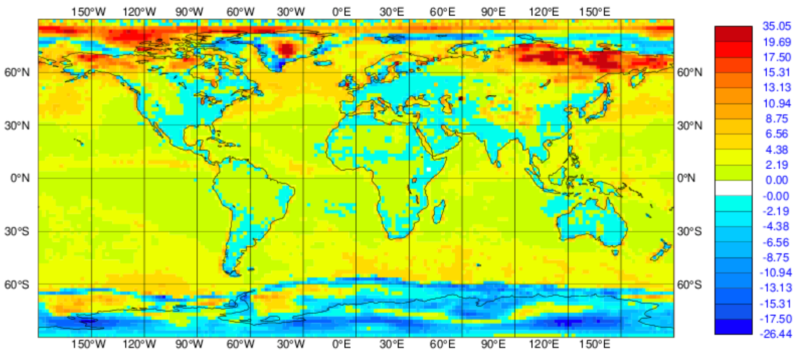

Figure 6: Example of ECMWF operational cloud-free bias monitoring[5] of ATMS 23.8 GHz from November 2019.

[5] https://www.ecmwf.int/en/forecasts/charts/obstat/?facets=Parameter,Radiances%3BData%20type,Microwave%20radiances

4.4 Initial assessment of brightness temperature bias correction

Figure 7 and Figure 8 show the results of the 1D-VAR TCWV retrieval approach (3-inputs ANN retrieval) applied to a subset of 8,000 randomly selected MWR data points with and without the brightness temperature bias correction applied:

- Without the bias correction, the 1D-VAR TCWV retrievals are biased high by about 2 kg/m^2 relative to the corresponding ERA reanalysis TCWV values.
- Once the bias-correction is applied, the 1D-VAR retrieval performs as good as the ANN retrieval and exhibits a slightly lower windspeed dependency as the ANN.

These results are shown to highlight the importance of the bias-correction. A full validation of the 1D-VAR TCWV retrieval with bias-correction applied is discussed in Section 5.

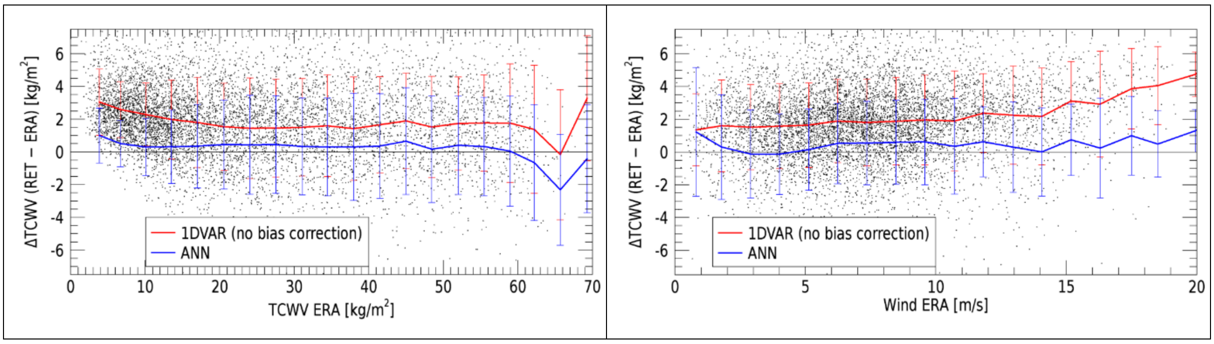

Figure 7: Difference between 1D-VAR and ERA-Interim reanalysis (red) as well as ANN and ERA-Interim reanalysis (blue) as a function of absolute TB (left) and wind speed (right) for a subset of 8,000 randomly selected MWR data points. No brightness temperature bias correction is applied for the 1D-VAR retrievals. This figure compares to Figure 8, where the bias correction is applied prior to the 1D-VAR retrieval.

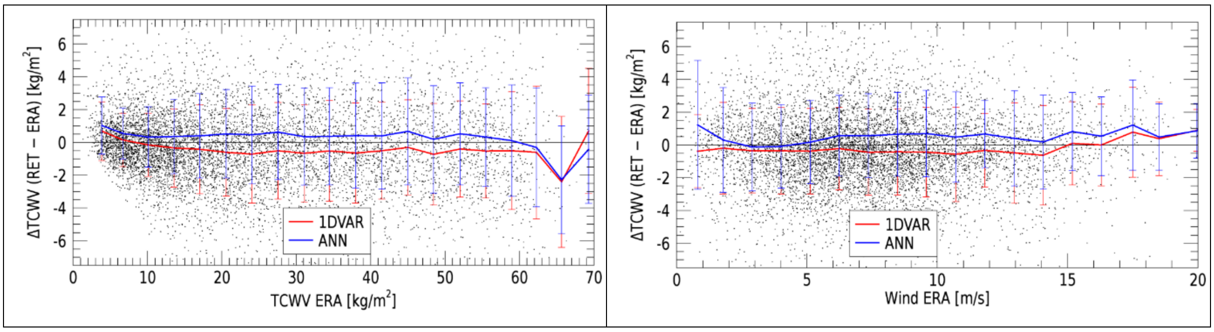

Figure 8: Same as Figure 7, but with the brightness temperature bias correction applied for the 1D-VAR retrieval.

5 Retrieval performance

This section evaluates the retrieval performance of the entire dataset.

5.1 Identifying valid observations

The performances of the altimeter system are usually assessed for valid observations only. Here, the same validity criteria as defined in the S3MPC STM Annual Performance Report *[S3MPC-STM-APR]*[6] are applied:

- Since only ocean open observations qualify for the AMTROC retrieval, observations above sea ice are discarded based on the **`open_sea_ice`** product flag, as recommended in *[S3MPC-STM-APR]*.
- In a follow-on step, outliers over the open ocean are identified based on thresholds applied to a set of parameters and are subsequently discarded as well. Details can be found in Table 7 of *[S3MPC-STM-APR]*.

In addition to the criteria listed above, it was ensured that all observations are located at least 100 km offshore.

5.2 1D-VAR retrieval failures

5.2.1 Physical interpretation of failed retrievals

After identifying valid observations as described under section 5.1, still about 2.6% of 1D-VAR retrievals fail. These failures occur mostly under precipitating conditions with high LWP above ca. 500 g/m^2, while only a marginal fraction of failed retrievals occurs for precipitation-free scenes (Figure 9).

For reference, we list some typical LWP values for different meteorological conditions:

- Non-precipitating stratocumulus cloud: ~100 g/m^2.
- Onset of precipitation: ~250 g/m^2.
- Frontal precipitation event, about 6 km deep: ~1000 g/m^2.
- Deep convection: up to several 1000 g/m^2.

Figure 9 shows that about 2.6 % of MWR retrievals are affected by precipitation and do not lead to consistent retrieval results between the ANN and the 1D-VAR, the latter flagging 'conservatively' more retrievals as missing. In this context, it is instructive to evaluate the altimeter's own rain flag (**`'rain_flag_01_ku'`**), the idea being that cases which are already flagged as rain-affected by the altimeter and hence not used anyway, will not be adversely affected by any potential artefacts in MWR retrievals.

[6] S3MPC STM Annual Performance Report - Year 2018, S3MPC.CLS.APR.004, 28/02/2019, 1rev0

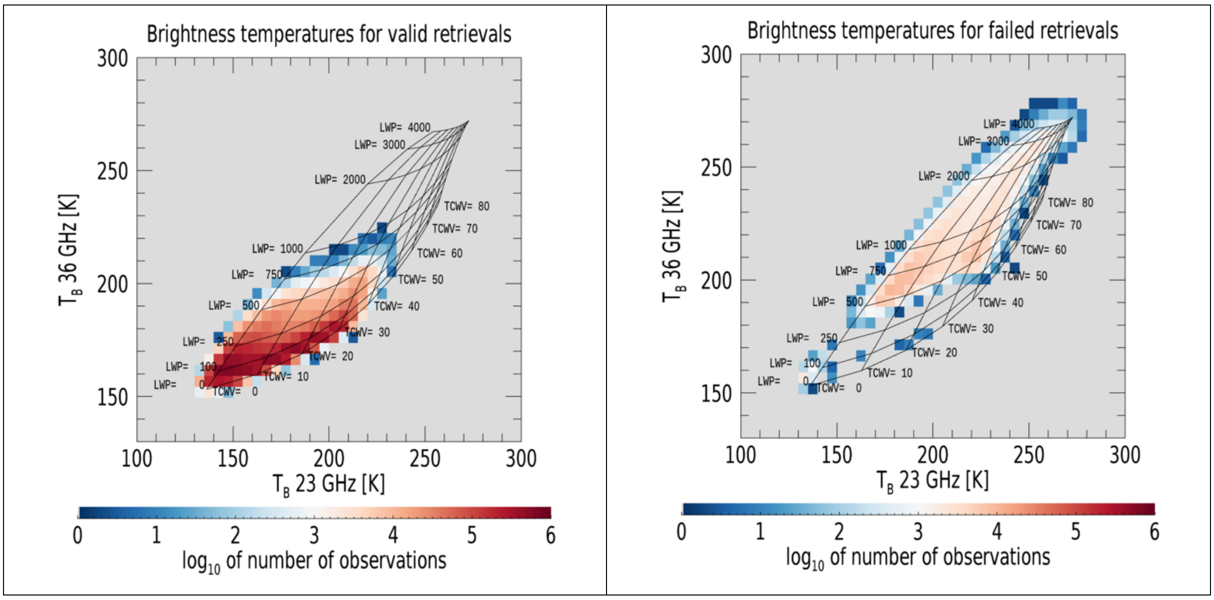

Figure 9: Probability density functions (PDFs) of TB23 and TB36 for valid (left) and failed 1D-VAR retrievals with concomitant ANN retrievals (right). On top of the data density plots, a grid of LWP (g/m^2) and TCWV values (kg/m^2) is plotted for orientation. Data are shown for the entire 14 million retrievals, of which 97.4% were valid and 2.6% failed. Of the 2.6% failed retrievals, 98.8% had TB36 larger than 180 K, *i.e.* the isolated blue spots at low TB values in the right panel make up for only about 0.03% of the total dataset. Thus, the vast majority of failed retrievals occurs under precipitating conditions with high LWP.

Table 4 shows the percentage results of altimeter flags for both the 'valid retrievals' and 'failed retrievals' reported in Figure 9. One can see that while for the valid retrievals the vast majority is flagged as no rain by the altimeter, the altimeter rain flag does not allow to identify with high skill cases, where the MWR the retrievals fail. In fact, 87 % of the failed retrievals are not identified by the altimeter.

Table 4: Altimeter rain flag ('rain_flag_01_ku') distribution corresponding to both 'valid retrievals' and 'failed retrievals' reported in Figure 9.

Rain flag value	Explanation of rain flag as provided in the S3A netCDF files	[%] of valid retrievals	[%] of failed retrievals
0	no_rain	99.53	87.30
1	rain	0.15	10.83
2	high_rain_probability_from_altimeter	0.00	0.00
3	high_probability_of_no_rain_from_altimeter	0.01	0.02
4	ambiguous_situation_possibility_of_ice	0.30	1.84

The key conclusions here are as follows:

– In 2.6 % of the total cases investigated herein, the more conservative 1D-VAR does not provide retrievals while the standard ANN methods still do.

- The above results further indicate that ANN retrievals in those 2.6% of cases are biased because of precipitation contamination. We assume that the ANN is extending retrievals beyond its training range.
- In about 2.3 % of the cases (that is 87 % of the total 2.6 %, see Table 4), this precipitation contamination is also not identified by the altimeter.
- It is therefore likely that the naïve application of ANN-derived WTCs to altimeter data will lead to biases in corrected SSH estimates under those precipitating conditions.

5.2.2 Geographic distribution and impact of failed 1D-VAR retrievals

Figure 10 shows the spatial distribution of ANN-derived TCWV values for cases where the 1D-VAR retrieval failed. It also shows the bias between the ANN-derived TCWV and the ERA-5 TCWV. For these cases, the global average bias of the ANN-retrieved TCWV (ANN-ERA) is 4.37 kg/m^2 and the corresponding RMSE is 6.12 kg/m^2, which is unusually large.

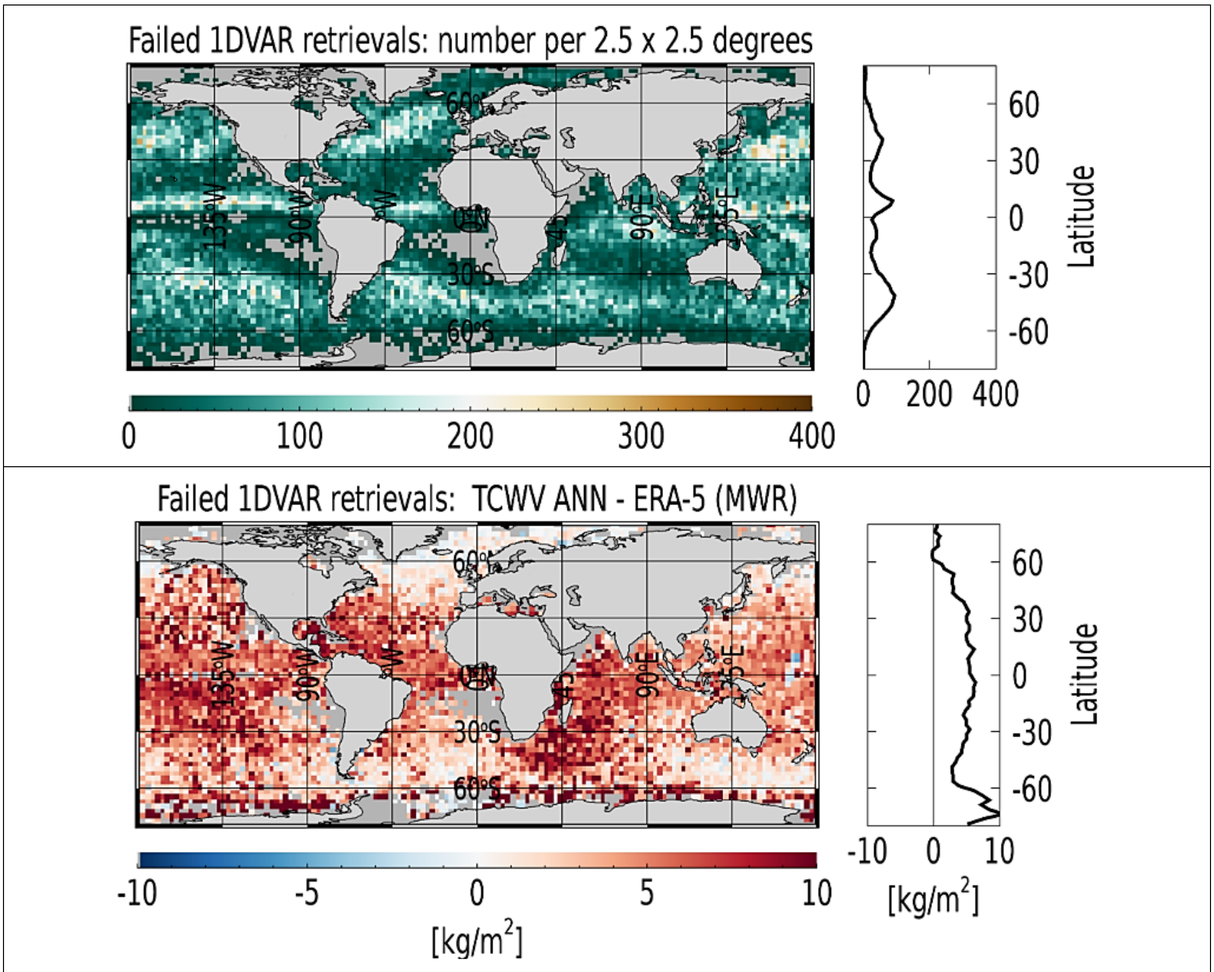

Figure 10: The upper panel shows the spatial distribution of those failed 1D-VAR retrievals where the ANN reports a result. The lower panel shows the difference between the ANN-derived TCWV and collocated ERA-5 TCWV for these cases.

An objective validation of the ANN retrievals for extreme conditions is not possible due to the lack of suitable reference measurements. However, it is unlikely that the ERA TCWV values are wrong by such a large margin. We therefore assume that the ANN produces biased TCWV retrievals in the presence of thick, precipitating clouds (where the 1D-VAR will frequently not converge and hence provide no retrievals).

As an additional note, one can see from Figure 10 that near the ice edge, a few failed retrievals show negative biases. There is also a spot around 10° N/170 °E where negative biases exist. As pointed out in Section 5.2.1, these cases accumulate to about only 0.03% of the dataset.

For future operational 1D-VAR applications, the use of collocated NWP TCWV for altimeter correction should be considered. Those collocated TCWV values could potentially be forced to be on average bias-free with respect to the overall 1D-VAR retrieval by subtracting the overall TCWV difference between 1D-VAR and NWP.

5.3 A posteriori uncertainties of 1D-VAR TCWV and WTC retrievals

The 1D-VAR retrieval provides physically based *a posteriori* uncertainty estimates for all retrieved parameters at the level of individual retrievals. For the successfully retrieved TCWV values, mean and associated *a posteriori* uncertainty amount to 24.5±1.46 kg/m^2 corresponding to an average relative uncertainty of ca. 6 %. For WTC, the corresponding values are 14.7±0.9 cm. We note that these uncertainties are largely driven by different types of noise, including instrument noise *[Bennartz et al., 2017]*. Random uncorrelated instrument noise contributes significantly to these uncertainties, so that averaging over a set of observations will reduce the uncertainty further.

Beyond these top-level numbers, it is instructive to evaluate the absolute and relative a-posteriori uncertainty as function of TCWV itself as shown in Figure 11. One can see that while the absolute uncertainty increases with increasing TCWV, the relative uncertainty approaches values around 4% to 5% for higher TCWV values. Only for low TCWV does the relative uncertainty increase significantly, while the absolute uncertainty approaches 1 kg/m^2.

Relative uncertainties for WTC are virtually identical to those of TCWV.

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

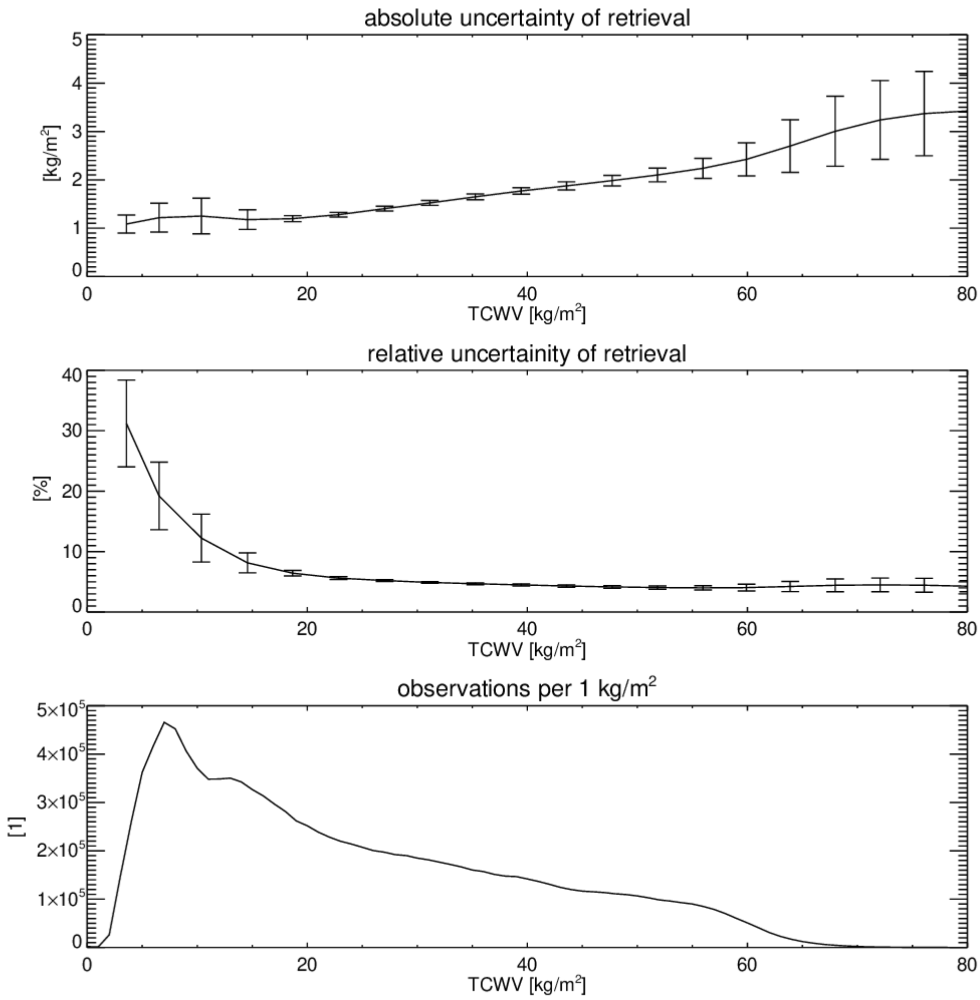

Figure 11: Absolute (upper panel) and relative (middle panel) *a posteriori* uncertainty of 1D-VAR TCWV retrieval as function of TCWV. The lower panel shows the amount of data for bins of 1 kg/m^2 TCWV. The TCWV uncertainty values [in kg/m^2] can be transferred with sufficient accuracy to WTC uncertainties [in cm] by multiplication with a factor of 0.61. A TCWV uncertainty of 1 kg/m^2 therefore yields a WTC uncertainty of approximately 0.61 cm.

5.4 Validation of TCWV and WTC retrieval at Level 2

Here we present an evaluation of the 1D-VAR TCWV and WTC retrievals at Level 2 against ERA-5 reanalysis for the full, bias-corrected AMTROC dataset (14 million observations). We note that the comparison against reanalysis is not a full validation. Such a validation is performed in Section 5.5 against the RSS climatology for TCWV and using cross-over analysis for WTC (Section 5.6).

The results are summarized in Table 5, where "1D-VAR (NO NWP)" refers to an application of 1D-VAR with input to the parameter vector ***b*** *[AD-AMTROC-ATBD]* coming from climatology instead of NWP. Input to ***b*** comprises the water-vapour averaged mean inverse atmospheric temperature T_m *[AD-AMTROC-ATBD]*, the atmospheric surface pressure (PSFC), and the sea surface temperature (SST). The observed degradation of "1D-VAR (NO NWP)" versus the full 1D-VAR highlights the importance of NWP estimates for the provision of input to the parameter vector ***b***.

Table 5: Comparison of biases and RMS errors for different retrievals of TCWV and WTC with respect to collocated ERA-5 at Level-2.

Retrieval scheme	TCWV bias	TCWV RMSE (bias corrected)	WTC bias	WTC RMSE (bias corrected)
	kg/m²	kg/m²	cm	cm
1D-VAR	0.01	2.63	0.00	1.62
1D-VAR (NO NWP)	-0.29	2.69	0.17	1.66
ANN	0.48	2.76	-0.33	1.69

Compared against collocated ERA-5, the full 1D-VAR performs best, closely followed by 1D-VAR (NO NWP) and ANN. Figure 12 evaluates results against altimetry-derived wind speed. Once can see that 1D-VAR has a lower dependency on wind speed than does ANN.

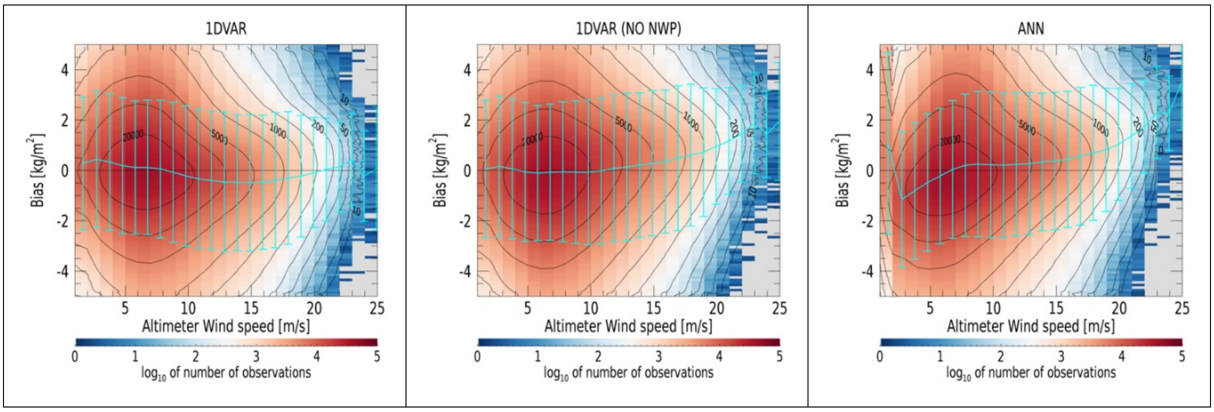

Figure 12: Difference between TCWV retrievals and ERA-5 as function of altimeter-derived surface wind speed for the entire 14 million retrievals. The cyan curves give the average bias and standard deviation as function of wind speed.

Summarising, at Level 2, the results of 1D-VAR and ANN are very close to each other with the 1D-VAR performing slightly better when compared against reanalysis. In addition, 1D-VAR shows a reduced dependency on wind speed than does ANN. If NWP input for SST, PSFC, and TM is replaced by climatology, 1DVAR and ANN show comparable retrieval performance.

5.5 Validation of TCWV at Level 3

Here we compare MWR-retrieved TCWV with three gridded TCWV climatologies (RSS, ERA-5 and ERA-Interim). From MWR, we include 1D-VAR and ANN retrievals as well as an additional dataset that collocates ERA-5 data with MWR observations and re-grids these collocations to the same grid resolution (2.5°×2.5°) as the other data. This last dataset is of interest as the comparison with the actual ERA-5 climatology allows to isolate the effects of the MWR sampling. Global comparison statistics are shown in Table 6. One can see that all datasets are in agreement within a small margin.

At the monthly mean time scale, the differences between ANN and 1D-VAR are smaller than the differences between different climatologies. ANN and 1D-VAR are therefore considered identical in terms of their TCWV retrieval performances when compared against monthly mean climatologies.

Table 6: Comparison of biases and RMS errors at Level-3 for different retrievals of TCWV with respect to different climatologies. All values are given in units of kg/m².

	RSS		ERA-5		ERA-Interim	
	Bias	RMSE (bias corrected)	Bias	RMSE (bias corrected)	Bias	RMSE (bias corrected)
1D-VAR	-0.79	0.77	0.06	0.98	-0.01	1.13
ANN	-0.33	0.82	0.53	0.94	0.46	1.08
ERA-5 re-gridded	-0.94	0.81	-0.09	0.76	-0.16	0.90

Figure 13 and Figure 14 show regional results in terms of biases against different climatologies. Again, the results between ANN and 1D-VAR are very similar.

TCWV retrievals from 1D-VAR and ANN are very close to each other. Independent validation against TCWV climatologies does not provide any indication that one is better than the other.

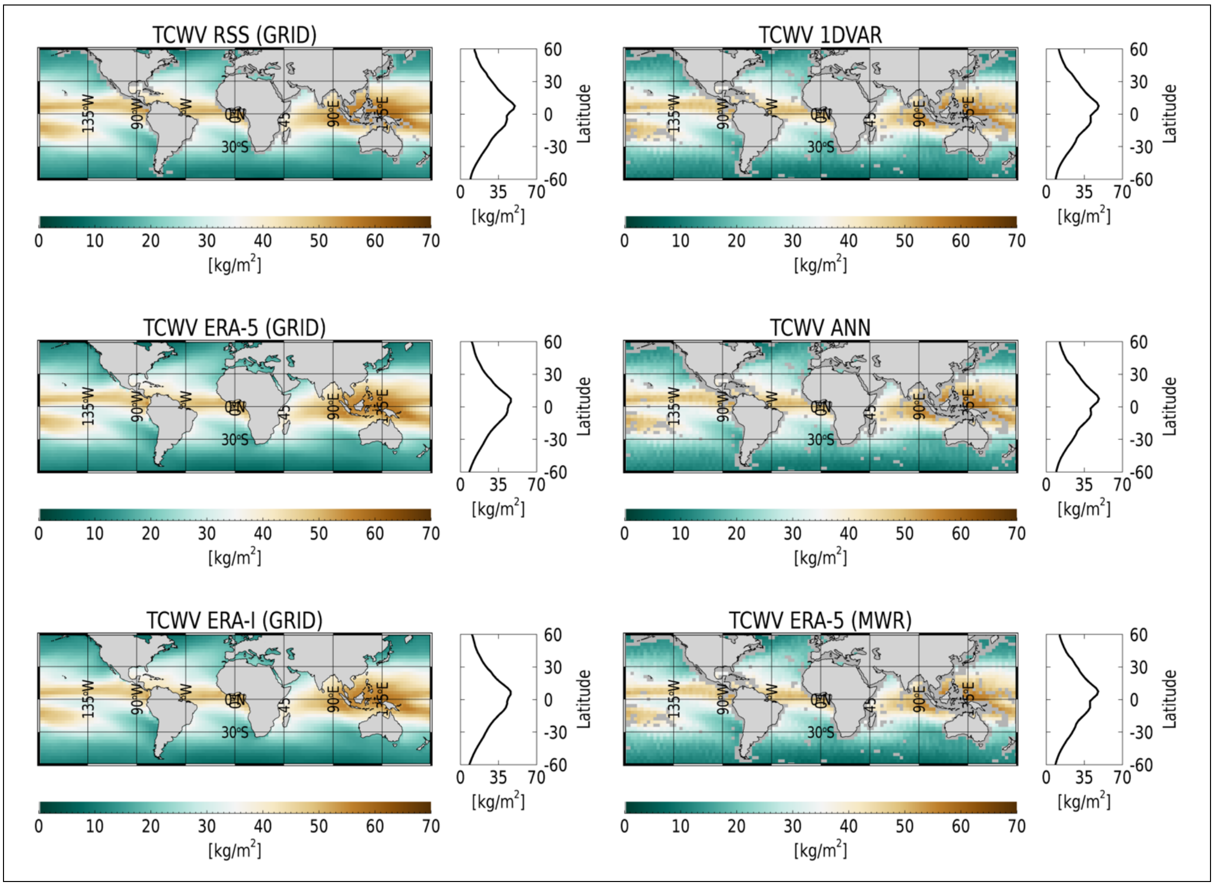

Figure 13: Long-term mean TCWV from three different gridded products (left: RSS, ERA-5, and ERA-Interim) and from MWR (right: 1D-VAR, ANN, and ERA-5 collocated with MWR and re-gridded).

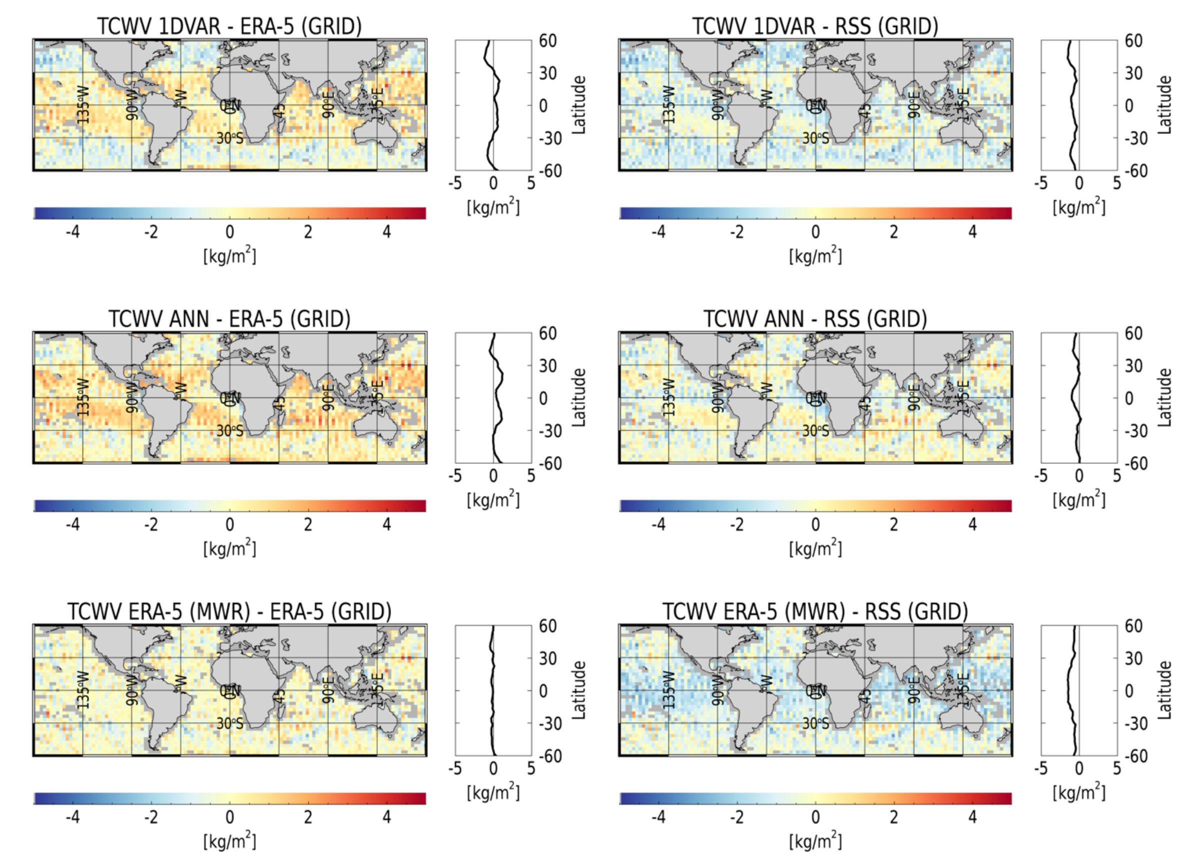

Figure 14: Difference between three different MWR-derived retrievals and ERA-5 (right) and RSS (right).

5.6 Crossover analysis of WTC

Crossover (X-Over) analysis is a metric particularly well adapted to the assessment of altimeter corrections applied to sea surface height (SSH) observations *[LeTraon et al., 1994]*: a new correction brings improvement to SSH retrievals when it minimizes the variance of the SSH differences between ascending and descending passes separated by less than 10 days.

The computation of the X-Over locations and of the SSH at these locations has been assessed against the statistics provided by the MPC cyclic report. Note that the SSH is computed at any X-Over location by a simple linear interpolation of the two closest measurements.

Unfortunately, the first MPC cyclic report including results obtained with the same reprocessing baseline than the one used for the present study (IPF-SM-2 06.12, S3A Processing Baseline 2.27) covers cycle 017 *[S3MPC.CLS.PR.06-017]*. We have compared those results with the results of the current study obtained for cycle 015, the last complete cycle of the reprocessing dataset.

In the MPC ocean validation report for cycle 017, the bias of SSH at X-Over is reported to have a mean of +0.22 cm and a standard deviation of 3.59 cm. With the current approach for cycle 015, the SSH bias has a mean of +0.24 cm and a standard deviation of 3.56 cm.

Apart from the difference of two months (resulting from the temporal difference between cycles 15 and 17), note that the statistics of the MPC report are computed only for water depths >1000 m (similar to discarding MWR observations less than 100 km offshore applied herein) and for ocean variability (computed over several years) below 0.2 cm (not taken into account in this study). Despite these differences, the statistics are very close. They are also consistent when compared to the temporal monitoring of the SSH bias and standard deviation displayed in the MPC annual report which includes results obtained with the same processing baseline.

To assess the performance of the 1D-VAR retrieval, a reference amongst the operational products has to be chosen. An X-Over analysis is applied to compare 3-i NN and 5-i NN WTC solutions from the operational S3 products. The statistics show a small degradation of the 5-i NN compared to the 3-i NN (+0.7 cm^2, global mean over a 4°×4° gridded map established over the whole period), mainly located in the mid-latitudes (see Figure 15). No further explanation on the sources of the degradation can be provided without a detailed analysis of the 5-i NN algorithm. Considering the good performances of the 3-i NN solution, the impact of using climatological maps of SST and γ_{800} instead of collocated outputs from ECMWF analysis would be a first lead to investigate.

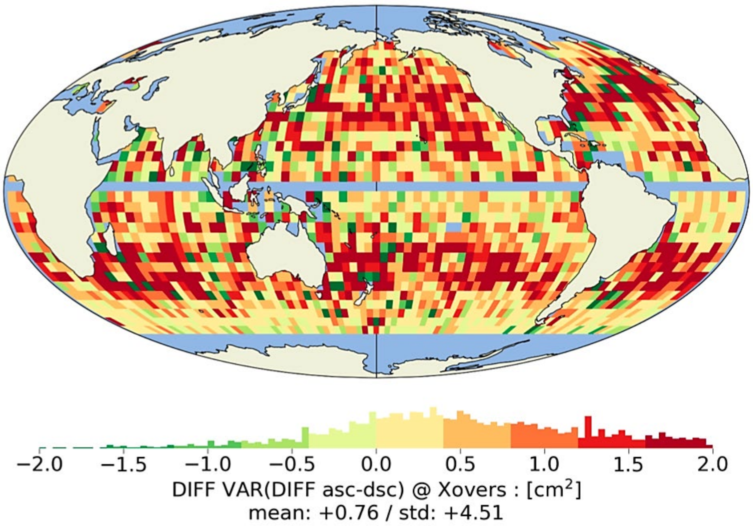

Figure 15: Geographical distribution of the difference of variance of SSH differences at X-Overs, comparing 3-i NN and 5-i NN operational products over the whole reprocessing period. The difference is (5-i NN – 3-i NN). The 5-i NN degrades the overall performances with respect to 3-i NN at mid-latitudes.

As shown in Figure 17 (light blue), the small degradation of the 5-i NN with respect to the 3-i NN is constant throughout the whole period of reprocessing. Those results are in contradiction of the MPC annual report that shows similar performances between 3-i NN and 5-i NN WTC retrievals over the first two years. Nevertheless, no direct comparison between the two retrievals is provided, only independent comparison of each solution to the model WTC. The differences between the two results could be explained by differences in the computation of the location of X-Overs and of the SSH at those locations or in the specific filtering applied in the MPC reports (low oceanic variability). Considering those results, the 3-i NN solution is taken as a reference for the comparison to the 1D-VAR retrieval.

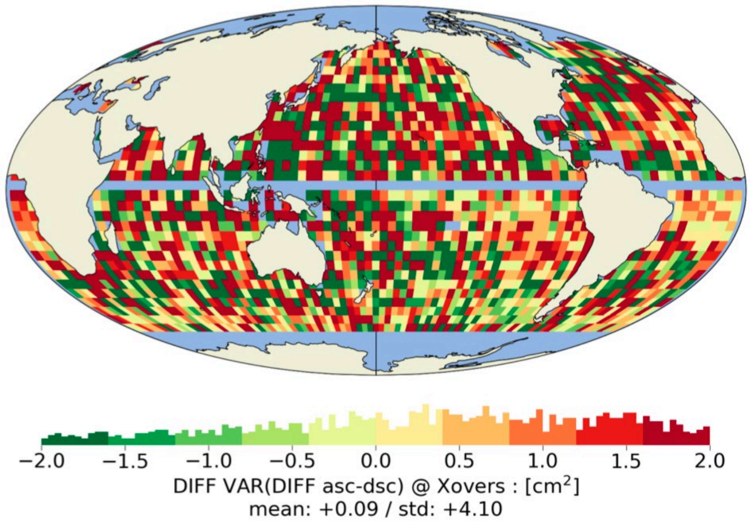

Figure 16: Geographical distribution of the difference of variance of SSH differences at X-Overs, comparing 3-i NN and 1DVAR retrievals over the whole period of reprocessing. The difference is (3-i NN – 1D-VAR). The performances are very similar with no clear geographical pattern and a large dispersion.

The geographical distribution of the X-Over analysis (Figure 16) between 3-i NN and 1D-VAR WTC shows a noisy figure without any clear pattern. The global statistics indicate similar performances between the two (1D-VAR with minor degradation of +0.1 cm^2) but with a large dispersion.

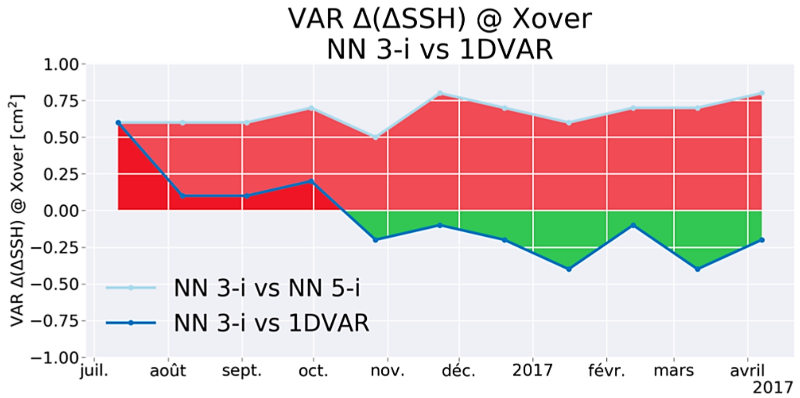

Figure 17: Temporal monitoring of the difference of variance of SSH differences at X-Overs, comparing 3-i NN vs. 5-i NN operational products (light blue) and 3-i NN *vs.* 1D-VAR retrievals (dark blue). The degradation of the 5-i NN against the 3-i NN retrieval is small but constant over time (ca. +0.6 cm^2). The performance of the 1D-VAR solution compared to 3-i NN oscillates between small degradation and small improvement over the whole period, resulting in a very similar overall performance.

The temporal monitoring of the difference of variance (see Figure 17) shows that the performance of the 1D-VAR oscillates between degradation (at the beginning of the period) and improvement (over the second part of the period). Considering those results, the X-Over analysis suggests similar performances between the 1D-VAR and the 3-i NN operational products. The second part of the period analysed during this study would even conclude to a potential slight improvement of the 1D-VAR compared to 3-i NN.

A longer period of analysis would allow to consolidate those results and to improve the physical assessment of the differences between the ANN and the 1DVAR approaches.

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

6 Recommendations

The following recommendation emerge from this validation report:

(1) The AMTROC 1D-VAR retrieval performs similarly well as do the operational ANN-based algorithms. It should be considered as a candidate algorithm for future operational retrievals of TCWV and WTC.

(2) Regardless of which retrieval is used (1D-VAR or ANN), a thorough bias-correction of the underlying brightness temperatures is crucial to the success of any retrieval:

- *a) Such bias correction should be documented and implemented in a way transparent to the user.*
- *b) It needs to be performed not only with respect to the instrument but also with respect to the retrieval system used.*
- *c) It must be re-derived if the calibration of the MWR instrument changes.*

(2) Beyond brightness temperatures, a set of auxiliary parameters is needed to perform high-quality retrievals (again regardless of which retrieval is used).

- *a) These auxiliary parameters include most prominently surface wind speed or some other measure of the sea surface roughness, but also sea surface temperature and information on the atmospheric temperature profile.*
- *b) The auxiliary parameters should ideally be derived from concomitant measurements (e.g. wind speed from altimeter), or from temporally and spatially collocated NWP fields.*
- *c) The use of climatological data for those auxiliary parameters leads to a degradation of the retrieval results in cases where the climatology diverges from the actual situation and is therefore discouraged.*

(3) Precipitation screening is deemed crucial for the improvement of WTC and TCWV estimates under precipitating conditions.

- *a) Typical two-channel MWRs, such as the ones on the Sentinel-3 satellites, provide only limited information on precipitation.*
- *b) The 1D-VAR is slightly more conservative than the ANN in flagging precipitation-affected retrievals and hence performs slightly better in this regard.*
- *c) The addition of one or more higher-frequency channels (> 80 GHz) to future MWRs would be beneficial as higher frequency observations will allow for a better screening of precipitation-affected observations.*

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

7 References

Bennartz, R., Höschen, H., Picard, B., Schröder, M., Stengel, M., Sus, O., Bojkov, B., Casadio, S., Diedrich, H., Eliasson, S., Fell, F., Fischer, J., Hollmann, R., Preusker, R., and Willén, U. (2017): An intercalibrated dataset of total column water vapour and wet tropospheric correction based on MWR on board ERS-1, ERS-2, and Envisat, *Atmospheric Measurement Techniques*, 10, 1387-1402, 10.5194/amt-10-1387-2017.

Dee, D. P., et al. (2011): The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553-597, 10.1002/qj.828.

Hocking, J., Rayer, P., Saunders, R., Matricardi, M., Geer, A., and Brunel, P. (2011): RTTOV v10 Users Guide, EUMETSAT Satellite Application Facility on Numerical Weather Prediction.

LeTraon, P.-Y., Stum, J., Dorandeu, J., Gaspar, P., and Vincent, P. (1994): Statistical analysis of Topex and Poseidon data. *Journal of Geophysical Research*. 992. 24619-24631. 10.1029/94JC01110.

Obligis, E., Eymard, L., Ngan, T., Labroue, S., and Féménias, P. (2006): First three years of the Microwave Radiometer Aboard ENVISAT: In-flight calibration, Processing and Validation of the geophysical products. *Journal of Atmospheric and Oceanic Technology*. 10.1175/JTECH1878.1.

Obligis, E., Rahmani, A., Eymard, L., Labroue, S., and Bronner, E. (2009): An Improved Retrieval Algorithm for Water Vapor Retrieval: Application to the Envisat Microwave Radiometer. *IEEE Transactions on Geoscience and Remote Sensing*, 47. 3057 - 3064. 10.1109/TGRS.2009.2020433.

Saunders, R., Rayer, P., Brunel, P., von Engeln, A., Bormann, N., Strow, L., Hannon, S., Heilliette, S., Liu, X., Miskolczi, F., Han, Y., Masiello, G., Moncet, J. L., Uymin, G., Sherlock, V., and Turner, D. S. (2007): A comparison of radiative transfer models for simulating Atmospheric Infrared Sounder (AIRS) radiances, *Journal of Geophysical Research-Atmospheres*, 112, D01s90 10.1029/2006jd007088, 2007.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C. (2018): An update on the RTTOV fast radiative transfer model (currently at version 12), *Geoscientific Model Development*, 11, 2717-2737, 10.5194/gmd-11-2717-2018.

[S3MPC-STM-APR] S3MPC STM Annual Performance Report - Year 2018, S3MPC.CLS.APR.004, 28/02/2019, 1rev0.

[S3MPC.CLS.PR.06-017] S3-A Ocean Validation Cyclic Performance Report, Cycle 017, S3MPC.CLS.PR.06-017, 23/05/2017, 1rev0.

[S3MPC-STM-PEP] STM Product Evolution for Processing Baseline 2.24, S3MPC.CLS.REP.030, 16/01/2018, 1rev0.

Weng, F., and Yang, H. (2016): Validation of ATMS Calibration Accuracy Using Suomi NPP Pitch Manoeuvre Observations, *Remote Sensing*, 8, 10.3390/rs8040332.

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