Title: ATBD Water Vapour

Doc. no: MAPP-ATBD-WV

Issue: 1

Revision: 1

Date: 1999-02-03

internal Distribution

Name	Quantity

external Distribution

Name	Quantity

Change Record

Issue	Revision	Date	Description	Change pages

Table of Contents

1. INTRODUCTION

Columnar water vapour content will be one of the central atmospheric application of the Medium Resolution Imaging Spectrometer (MERIS) onboard ENVISAT. Within the framework of the German MERIS processor level 2 water vapour retrievals will operationally be derived over land, clouds, and water surfaces. To ensure a high accuracy of the individual retrievals, the water vapour retrievals make use of the high resolution land surface albedo as well as of MERIS level 2 cloud products, such as cloud optical depth and cloud top pressure.
The retrieval itself will be performed by means of a neural network. Besides its capability to properly ft the entire range of variability of retrieved water vapour contents, the proposed neural net is also capable to internally derive estimates of the quality of the input data, which are useful estimates of the accuracy of the algorithm.

1.1 Algorithm Identification

Columnar water vapour path over land, water, and clouds.

2. ALGORITHM OVERVIEW

The water vapour retrieval algorithms dedicated for MERIS are based on the work of Fischer (1989), Bartsch, (1996) and Bartsch *et al.* (1997). The general algorithm approach is to relate the columnar water vapour content to the ratio of MERIS channels 14 and 15, located at 890 nm and 900 nm, respectively.
The retrieval will be performed using a neural network which will be trained with results of the below described radiative transfer model MOMO. Although conventional regression approaches for model inversion have successfully been applied to the inverse problem, the use of neural networks in this context will lead to significant improvements in the quality of the retrieval. This is especially the case for the low- and high end of the observed water vapour contents, where regression type algorithms typically lead to biases, because of there overall minimisation. It further allows to derive estimates about the quality of each retrieval, which allows to accept (or reject) the derived retrieval as being consistent (or inconsistent) with the neural network's training data. In contrast to simple regression techniques no threshold techniques have to be applied and pixels which e. g. have undergone an erroneous cloud classification in the pre-processing will be rejected from the neural network itself as inconsistent with the expected range of data. Details of the applied method are given in section 3.

3. ALGORITHM DESCRIPTION

3.1 Theoretical Description

3.1.1 Physics of the problem

Within the solar spectral range the ρστ-band around 935 nm exhibits strong absorption by water vapour. The transmission for two model atmospheres within the spectral range relevant for MERIS channel 15 is shown in Figure 1. The strength and broadness of individual lines is controlled by total water vapour amount, pressure, and temperature variations. Variations in atmospheric transmission directly translate to variations in the measured radiance. Over land surfaces the magnitude of the measured signal is mainly governed by variations of surface reflectance. Independence of surface albedo is gained relying on radiance ratios between MERIS channel 15 and the almost absorption-free reference channel 14. The assumption is, that scattering and absorption properties of other atmospheric constituents and the surface albedo do not differ significantly for these two channels. This presumption is crucial for the successful retrieval of water vapour. It is shown that for

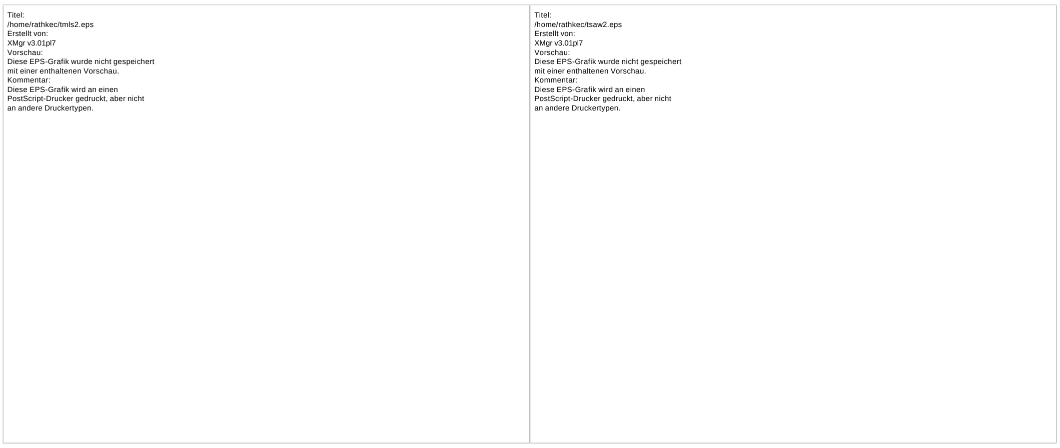

Figure 3.1: Atmospheric transmission between 895 nm and 905 nm for two standard atmospheres calculated form the HITRAN-96 database.

the majority of measurements the presumption is fulfilled to a high accuracy.

Outside the sun glitter water surfaces show a significantly lower surface reflectivity at higher wavelengths. The measured signal over water surfaces is mainly governed by aerosol scattering. Even though the overall structure of the ocean algorithm described by Equation (1) holds, the underlying physics differs significantly from the land algorithm. The effects of variable aerosol optical depth and of variations in aerosol and water vapour vertical profiles are the most significant factors affecting the retrieval accuracy. This has been accounted for by introducing a dependence of the regression coefficients on the aerosol optical depth. Despite this correction, the retrieval accuracy of the ocean algorithm, outside the sun glitter, is still inferior to that over land surfaces.

The sun glitter region has to be treated separately within the retrieval algorithm for water surfaces. Within the sun glitter region, the measured signal is comparable to that above land surfaces and hence land algorithms may be applied. The accuracy of retrievals over sun glitter is expected to be comparable to land surfaces.

Clouds are bright targets with small spectral variation in reflectivity. Water vapour retrieval results under cloudy conditions are therefore expected to be feasible. Variations in penetration depth of the incoming solar radiation lead to uncertainties in the assignment of the measured water vapour content to an altitude range. The penetration depth is mainly governed by the ratio between cloud optical depth and cloud geometrical thickness. The issues associated with variable penetration depth are closely related to cloud top pressure retrieval and are discussed in more detail in the cloud top pressure ATBD. Another factor, limiting the retrieval accuracy, is the comparably low water vapour content available above high clouds which generally leads to an increasing relative errors with increasing cloud top height.

3.1.2 Mathematical Description of the Algorithm

The retrieval algorithm will be based on a neural network which is schematically shown in Figure 2. The main difference between the proposed approach and classical approaches is that the neural net consists of 75 output neurons n_i, each associated with water vapour path between 1 kg/m^2 and 75 kg/m^2. The neural network is trained in a way that for a water vapour content of e. g. 10.4 kg/m^2 the 10th output neuron gives a value of 0.6

and the 11th neuron of 0.6, since the result is closer to the 10th neuron than to the 11th. All other neurons are trained for output values of zero.

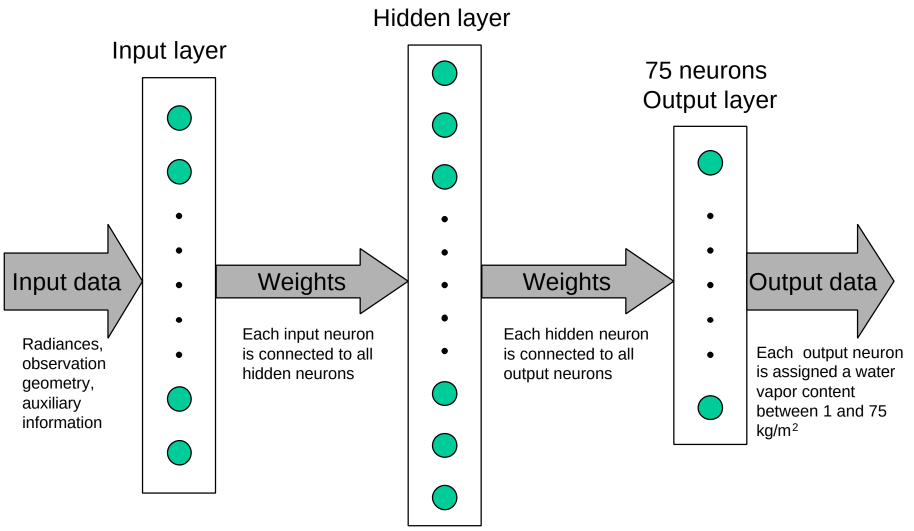

Figure 2: Conceptual view of the proposed neural network for water vapour retrieval**.**

The actual output value of the retrieval is than derived from:

$$w = \frac{\sum_{i=1}^{75} o(i) \cdot i}{\sum_{i=1}^{75} o(i)} \tag{1}$$

where w is the water vapour content in kg/m^2, i is the respective number of the output neuron, which is at the same time the water vapour content associated with the neuron, and o(i) the neuron's respective activation. Since the network is trained to reproduce an integral output of 1 which is spread over only two adjacent neurons, the following values can be understood as measures of the overall capacity of the neural net to handle a given set of input values.

$$E_1 = \sum_{i=1}^{75} o(i) \tag{2}$$

$$E_2 = \sqrt{\sum_{i=1}^{75} \left[o(i) - \frac{E_1}{75} \right]^2} \tag{3}$$

The first error estimate E_1 describes the actual accuracy with which the network is able to identify the input data. If the input data is perfectly recognised, E_1 equals 1, if the input data falls apart from the manifold covered by the training dataset, E_1 will significantly deviate from unity. The second error estimate E_2 is the standard deviation of the output activities and is a measure of the certainty with which the neural network retrieves the water vapour from given input data. Note, that in this case a large standard deviation is associated with a more certain estimate, while a small standard deviation translates to a less certain estimate. This is because in an ideal case the activation should only be spread over the two neurons adjacent to the actual water vapour content. If the estimate is more uncertain, then the network produces activities over a larger number of neurons and E_2 decreases.

Although still relying on the data used for the network training, the error estimates E_1 and E_2 are superior to a fixed error estimate derived from the training data alone, since not only the theoretical error from the inversion procedure is estimated, rather all possible deviations of the actually measured data from those used for training are include. Sources of these deviations may be three-dimensional radiative transfer effects, wrong cloud classifications and other sources of geophysical noise, which can only partly be covered during algorithm development. A positive side-effect during the evaluation stage of the method is that in the unlikely case of severe processing or algorithm development errors, these would easily be identified via deviations of true from the expected error estimates.

3.2 Practical Considerations

3.2.1 Numerical computation considerations

TBD

3.2.2 Calibration and Validation

The overall calibration and validation strategy is visualised in Figure 3. Three validation steps are considered necessary to first generate a physical meaningful algorithm and second to correct for systematic deviations of the algorithm results from the actual water vapour content.

Step one includes all required simulation work to implement an algorithm which is consistent with the forward radiative transfer model. Step two embodies a first application to satellite data. The Modular Optical

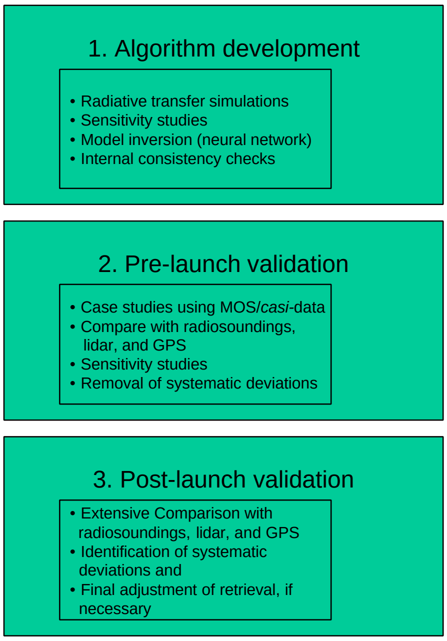

Figure 3: Validation strategy for MERIS water vapour retrieval algorithm**.**

Spectrometer (MOS) onboard the Indian satellite Indsat allows for a pre-launch calibration/validation of the algorithms developed for MERIS. The slightly different channel settings for the MOS water vapour channels (at 865 nm/945 nm) require a re-calibration of the MERIS algorithms to MOS.
In a third step, the post-launch calibration/validation phase, the total water vapour amounts retrieved from MERIS measurements will be compared with water vapour products of the Global Telecommunication System (GTS) radiosonde network. The use of ground-based upward-looking microwave radiometer as well as sun photometer measurements for further validation is strongly recommended.
Additional validation aircraft campaigns, including multi-spectral radiance measurements and active water vapour DIAL (DIfferential Absorption Lidar) measurements are necessary to establish the technique for the retrieval of the total water vapour content from future MERIS measurements. Such campaigns should be performed before launch of ENVISAT and at least during a commission phase.

3.2.3 Quality Control and Diagnostics

Quality control and diagnostics are considered a main component of the algorithm development. Instead of simply thresholding input and output values of the algorithm, emphasis is given to the internal error estimates of the neural network, as outlined in section 3.1. These estimates can be seen as cost-functions allowing to assess the match-up between the data used for algorithm development and the observed data.

The exact ranges of validity of the error estimates are to be investigated from pre-launch calibration/validation efforts and may also be adjust during the post-launch calibration/validation period.

3.2.4 Exception Handling

TBD

3.2.5 Output Product

The output product will consist of water vapour estimates over land, water, and clouds as well as of the above described error estimates.

4. ERROR BUDGET ESTIMATES

For a general estimate of the errors associated with different sources of geophysical and instrument noise, such as land surface albedo and albedo slope, aerosol content, observation geometry, see Fischer and Bennartz (1997). While most of the error sources discussed in Fischer and Bennartz (1997) are independent on the technique used for model inversion (neural network or regression), considerable improvements due to the neural network are mainly expected for the cases of extremely low or high water vapour content, since extreme values are in general not well represented in regression-type algorithms.

Details TBD

5. ASSUMPTIONS AND LIMITATIONS

6. REFERENCES

Bartsch, B., 1996: Fernerkundung des Wasserdampfgehaltes der Atmosphäre über Land aus rückgestreuter Sonnenstrahlung. Berichte aus dem Zentrum für Meeres- und Klimaforschung, No. 21, 11 pages, Hamburg, Germany.

Bartsch, B. and J. Fischer, 1997: Passive remote sensing of columnar water vapour content over land surfaces. MPI-report No.234, Hamburg, March 1997. ISBN 0937-1060.

Bennartz, R. and J. Fischer, 1999: A modified k-distribution approach applied to narrow band water vapour and oxygen estimates in the near infrared. Submitted to Journal of Quantitative Spectroscopy and Radiative Transfer.

Fischer, J., 'High Resolution Spectroscopy for Remote Sensing of Physical Cloud Properties and Water Vapour', In: Current Problems in Atmospheric Radiation, Ed. Lenoble and Geleyn, Deepak Publishing, 151-156, 1988.

Fischer, J. and R. Bennartz 1997: Retrieval of total water vapour content from MERIS measurements. ESA reference number: PO-TN-MEL-GS-005, published by ESA-ESTEC, Noordwijk, The Netherlands.

Rothman, L. S. *et al.*, 'The HITRAN Molecular Database: Editions of 1991 and 1992', J. Quant. Spectrosc. Radiat. Transfer, **48**, No. 5/6, pp. 469 - 507, 1992.

7. MAPP DATA PRODUCT SUMMARY SHEET

Product name:	Water vapour
Product code:	
Product Level:	2
Description of the Product:	Columnar water vapour content over land, clouds, and water
Product Parameters:	
Coverage	
Packaging:	
Units:	kg/m^2
Range:	MERIS Level 2
Sampling:	
Resolution:	
Accuracy:	
Geo-location:	
Format:	
Appended data:	
Frequency of generation:	Overpass
Size of product:	1 Value + 1 error estimate per pixel
Additional information:	
Identification of bands used in algorithm	14,15
Assumption on MERIS input data	
Identification of ancillary and auxiliary data	Surface pressure, albedo map, cloud top pressure (MERIS level 2), cloud optical depth (level 2)
Assumptions on ancillary and auxiliary data	