# DUE GLOBVAPOUR

## Software Development Plan

Issue 1 Revision 0

16 April 2010

Project nr: ESRIN/AO/1-6090/09/I-OL

Project Coordinator: Marc Schröder
Deutscher Wetterdienst
marc.schroeder@dwd.de

# Document Change Record

| Document, Version | Date | Changes | Originator |
|---|---|---|---|
| SDP, v1.0 | 2010.01.25 | Original version | T. Steenbergen,<br>M. Stengel,<br>M. Schröder,<br>D. Loyola,<br>R. Lindstrot |
| | | | |
| | | | |
| | | | |

# Table of Contents

# 1 Introduction

## 1.1 Purpose

This document provides requirements and guidelines for the software development needed throughout the project. The architecture of the various processors and processing environment and the EO Data Processing System (EODAPS) for GlobVapour is described.

## 1.2 Definitions, acronyms and abbreviations

| | |
|---|---|
| AATSR | Advanced Along Track Scanning Radiometer |
| AR | Acceptance Review |
| ATSR | Along-Track Scanning Radiometer |
| CDR | Critical Design Review |
| CEOS | Committee on Earth Observation Satellites |
| EODAPS-GV | EO Data Processing System for GlobVapour |
| FCC | Formal Configuration Control |
| FCDR | Fundamental Climate Data Record |
| GCWVP | GOME-type Climate Water Vapour Processor |
| GSICS | Global Space-based Inter-Calibration System |
| INSPIRE | Infrastructure for Spatial Information in Europe |
| MOMO | Matrix Operator Model |
| OE | Optimal Estimation |
| QR | Qualification Review |
| RTM | Radiative Transfer Model |
| SSM/I | Special Sensor Microwave/Imager |
| TCWV | Total Column Water Vapour |
| UCWVP | UV/VIS Climate Water Vapour Processor |
| UKMO | United Kingdom Meteorological Office |
| UPAS | Universal Processor for UV/VIS Atmospheric Spectrometers |
| WACMOS | Water Cycle Multimission Observation Strategy |

## 1.3 Applicable Documents

[AD-1] ESRIN Statement of Work; EOEP-DUEP-EOPS-SW-09-0003, issue 1 revision 1, 13 May 2009.

[AD-2] European Cooperation for Space Standardization: Space Engineering Software; ECSS-E-ST-40C, Part 1B, 6 March 2009; available from http://www.ecss.nl.

[AD-3] DUE GLOBVAPOUR Proposal; issue 1 revision 3, 9 July 2009.

[AD-4] DUE GLOBVAPOUR Clarification Note; issue 1, revision 1, 29 October 2009.

[AD-5] CEOS QA4EO guide for product quality indicators; QA4EO-CEOS-GEN-DQK-001, N. Fox, version 2.0, 13 September 2008; available from http://calvalportal.ceos.org/CalValPortal/welcome.do.

[AD-6] DUE GLOBVAPOUR Technical Specification Document; issue 1, revision 0, 10 March 2010

[AD-7] INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119; Drafting Team Metadata and European Commission Joint Research Centre; Version 1.1; 18 Feb 2009; available from http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/metadata/MD_IR_and_ISO_20090218.pdf.

## 1.4 Reference Documents

[R-1] Andersson, A., C. Klepp, K. Fennig, S. Bakan, H. Grassl, and J. Schulz, 2009: The HOAPS climatology: Essential water cycle components over global oceans derived from satellite data. Submitted to J. Appl. Met. Clim.

[R-2] Bennartz, R. and J. Fischer, 2001: Retrieval of columnar water vapour over land from back-scattered solar radiation using the Medium Resolution Imaging Spectrometer (MERIS). Remote Sensing of Environment, 78, 271-280.

[R-3] Deblonde, D., 2001: Variational retrieval using SSM/I and SSM/T-2 brightness temperatures in clear and cloudy situations. J. Atm. Oc. Tech., 18(4), 559-576

[R-4] Fell, F., and J. Fischer, 2001: 'Numerical simulation of the light field in the atmosphere-ocean system using the matrix-operator method', J. Quant. Spectrosc. Radiat. Transfer, 69, 351-388.

[R-5] Fennig, K. 2001: Interkalibration verschiedener SSM/I Mikrowellenradiometer im Hinblick auf eine gemeinsame Nutzung für eine fernerkundete Klimatologie (Homogenization of different SSM/I microwave radiometers for a combined satellite-derived climatology). Diploma Thesis, Met. Institute of the University of Hamburg, Hamburg.

[R-6] Li, Z. L., L. Jia, Z. Su, Z. Wan, R. Zhang, 2003: A new approach for retrieving precipitable water from ATSR2 split-window channel data over land area. Int. J.

Remote Sensing, 24, 5059-5117.

[R-7] Lindau, R. and M. Schröder, 2010, Objective analysis (Kriging) for water vapour products. Algorithm Theoretical Basis Document, Reference Number: SAF/CM/DWD/ATBD/KRIGING, Issue 1.0, 28 January 2010.

[R-8] Lindenbergh, R., Keshin, M., Van der Marel, H., and Hanssen, R., 2008: High resolution spatio-temporal water vapour mapping using GPS and MERIS observations. Int. J. Rem. Sens., 29(8), 2393-2409.

[R-9] Lindstrot, R., R. Preusker, and J. Fischer, 2010: 'The retrieval of cloud-top pressure of multilayered clouds using combined measurements of MERIS and AATSR'. Accepted by J. Appl. Meteor. Climat.

[R-10] NWP-SAF User's Guide, 2001: Standalone 1D-var scheme for the SSM/I, SSMIS and AMSU. G. Deblonde, NWPSAF-MO-UD-001 Version 1.0, 22 August 2001.

[R-11] Rodgers, C. D., 1976: Retrieval of atmospheric temperature and composition from remote measurement of thermal radiation. Rev. Geophys. Space Phys., 14, 609-624.

[R-12] Noyes, E., G. Corlett, J. Remedios, X. Kong, and D. Llewellyn-Jones, 2007: An accuracy assessment of AATSR LST data using empirical and theoretical methods. Proc. 'Envisat Symposium 2007', Montreux, Switzerland, 23–27 April 2007 (ESA SP-636, July 2007).

[R-13] O'Carroll, A.G., J.R. Eyre, and R.W. Saunders, 2008: Three-Way Error Analysis between AATSR, AMSR-E, and In Situ Sea Surface Temperature Observations. J. Atmos. Oceanic Tech., 25, 1197–1207.

[R-14] Pougatchev, N., 2008: Validation of atmospheric sounders by correlative measurements. Appl. Opt., 47, 4739-4748.

[R-15] Pougatchev, N. et al., 2009: IASI temperature and water vapour retrievals – error assessment and validation. Atmos. CHem. Phys. Discuss., 9, 7972 – 7989.

[R-16] Stoffelen, A., 1998: Error modelling and calibration; towards the true surface wind speed. J. Geophys. Res., 103, 7755-7766.

[R-17] 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.

## 1.5 Structure of the document

Section 2 gives an overview of the complete processing system and environment, as well as the system architecture. Section 3 provides a description of the scientific processors to be developed and implemented. In Section 4 the quality assurance aspects of the development and testing activities are discussed, including configuration control and problem reporting. Section 5 deals with the validation tools to be developed. Finally, conclusions are stated in the last section.

# 2 System development

## 2.1 Description of the processing environment

The EO Data Processing System for GlobVapour (EODAPS-GV) is the software that is capable of transforming input data from multi satellite instruments at different processing levels, into higher level output products.

The processing within GlobVapour comprises of two distinct elements:

1) The scientific processors: Each scientific processor is a piece of software that takes one or more input data, runs a specific scientific algorithm on these products (e.g. an atmospheric correction) and generates one or more output products. The primary objectives are the appropriateness of the scientific algorithm and its correct and effective implementation.

2) The processing system: This connects all scientific processors that are required to generate the final products, and takes care of running the chain of processors on all input data. It is composed of a data management system (DMS) and a processing management system (PMS).

These two elements should be decoupled for maximum reuse. The processors should be configurable via external configuration information (typically stored in config files), invoked with a processing request (that specifies input and output data for a specific processing job) and should ideally provide status information during execution. The processing should terminate with a report on the completed job. The processing system should support these interfaces: it should maintain a configuration control database for the processors, prepare the processing requests, monitor the processing and register the result of the completed processing in product and processing meta data. It should finally take care of providing input data and submitting the output data to the next processing job.

In summary, the interface between scientific processors and processing system should be

- Processor configuration
- Processing request
- Processing monitoring
- Input/output data management

As long as these items are provided, the actual implementation language of a processor does not impact the processing system.

The processing system is developed during the following project phases:

- Development of the prototype processing environment (WP 390)
- Development of the stand-alone processing environment (WP 470)
- Update of the final processing system delivery at ESRIN (WP 520)

### *2.1.1 System Overview*

The EODAPS-GV requires

- management of all input, auxiliary, intermediate and output products,
- correct execution of the scientific processing steps, and
- effective usage of the available hardware resources.

The EODAPS-GV will be a highly modular and desynchronised system. In particular, the scientific processing modules (for example cloud screening, retrieval of water vapour, retrieval of surface properties, sensor merging, …) will be stand alone executables that will be called by the processing management system. However, the modularisation goes much further in order to achieve the requested high performance of the system. The specific hardware on which the system runs will be described in configuration tables. This makes the system flexible for deployment on different platforms. This enables processing to be performed in the different processing centres and hence to distribute the processing load to multiple partners in a project. Also a future transfer of the system to a different hardware platform and service provider is possible by defining the proper configuration.

Figure 1 shows a schematic overview of the elements of the processing system. Its hardware is composed of the processing nodes, the storage devices and the communication channels (ports). The system has to deal with input data, auxiliary files (e.g. DEM), intermediate data generated during processing but required after verification of the results and configuration information. The processing chains determine how the input data is converted into output data; this involves one or more scientific processors (see chapter 3). A processing chain is considered as a recipe and its ingredients are the processors, input and aux data and processing parameters (configuration information). The processing system has to execute these recipes and take care of availability of input data, storage of intermediate and output data, success or failure of the processing and effective management of available computing nodes and storage capacity. How this is done is described in the following section (2.1.2).

Special attention is to be paid to the quality control of the input data. The data completeness (gaps, overlap) of the satellite data must be checked, and implicit quality flags must be evaluated. Also the auxiliary data must be checked prior to processing instigation.

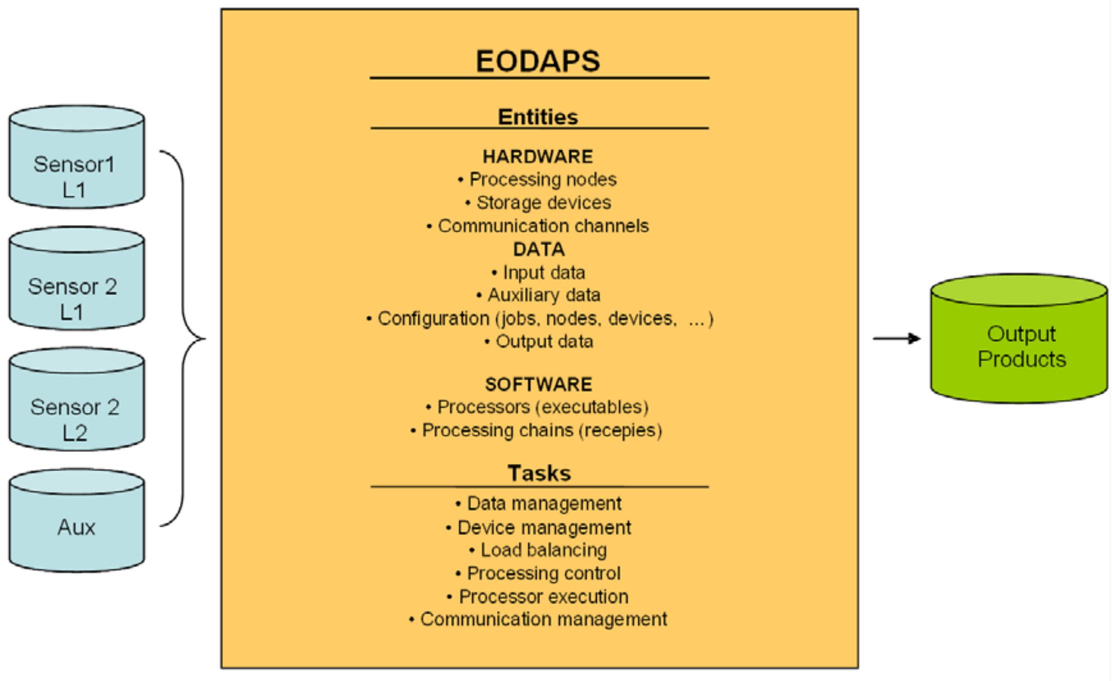


**Figure 1: Schematic overview of the components of the Earth Observation Data Processing System (EODAPS-GV).**

### *2.1.2 High Level Architecture Elements*

The processing system will be highly decoupled and desynchronised. This ensures failure tolerance and most efficient usage of resources. The following system is based on the existing MERCI and WAQSS systems at BC combined with the experience of the NASA Ocean Biology Processing Group, that has developed a concept for managing and processing a large volume of data in an extremely short period of time.

A data storage allocation is handled by a device manager. The device manager distributes the input data during loading round robin to the available devices. Data that are requested during processing will be delivered by the device manager. Through this mechanism the concurrent I/O requests will be equally distributed among all available disks and hence minimizing the latency time.

All executables are encapsulated in a standardised recipe wrapper class. The rest of the processing system does not need to know anything about the internal of the processing and calls the class via its standard interface. The recipe has a defined syntax for input and auxiliary data request, processing parameters (via XML configuration) and output.

The dispatcher is responsible for assigning processing jobs to processing nodes. The dispatcher is passive; it maintains a list of jobs and its status (ready for execution, under execution, completed) and processing node request jobs from the dispatcher, when they are free for the next job. The dispatcher is a core element of the architecture and includes features such as dynamic job priority and node monitoring.

Buffering is a concept that is associated with local storage associated with each computing node. The nodes never interact directly with the final RAID storage but write their output to a temporary local storage and send a notification to the device manager. When this is free it picks up the file and transfers it into the final location.

The token manager is responsible for the communication in the processing system. Each interaction in the system takes place on a dedicated port that is assigned by the token manager.

This system will be realised in a Service Oriented Architecture (SOA), where all communication is handles via standardised XML protocol.

# 3 Processor development and test products

The scientific processors will be stand-alone executables which will be invoked via a processing request, generated by the EODAPS-GV. Table 3-1 shows a list of expected processors and the institution responsible for the development, testing and maintenance of the processor.

**Table 3-1: List of expected GlobVapour processors to be integrated into EODAPS-GV.**

| Processor | Responsible Institution |
|---|---|
| GOME/SCIAMACHY/GOME-2 | DLR |
| MERIS | FUB |
| SSM/I (MWR) | DWD |
| MERIS-SSM/I Blend | DWD |
| ATSR/AATSR | FUB |
| IASI[1] | DWD |
| SEVIRI | DWD |
| Multi IASI-SEVIRI | DWD |

[1] The responsibility for the IASI processor depends on the results of the IASI retrieval assessment.

## 3.1 Development of GOME/SCIAMACHY/GOME-2 processing system

Figure 2 shows a schematic representation of the processing flow for the generation of the GOME/SCIAMACHY/GOME-2 climate dataset. Inputs are level 1 products from the corresponding sensors (the input for each sensor has a satellite specific format). The UPAS (Universal Processor for UV/VIS Atmospheric Spectrometers) system retrieves water vapour from each single satellite and generates a HDF5 product. These HDF5 products are the input to the UV/VIS Climate Water Vapour Processor (UCWVP) that generates the final NetCDF product. An external configuration file will be used to control/update algorithm and program settings in a flexible manner.

The components in yellow are already available at DLR (UPAS is a pre-existing software system owned by DLR for the operational retrieval of trace gas species from UV/VIS measurements). The components in blue will be developed as part of the GlobVapour project. The UCWVP software will be integrated into EODAPS-GV.

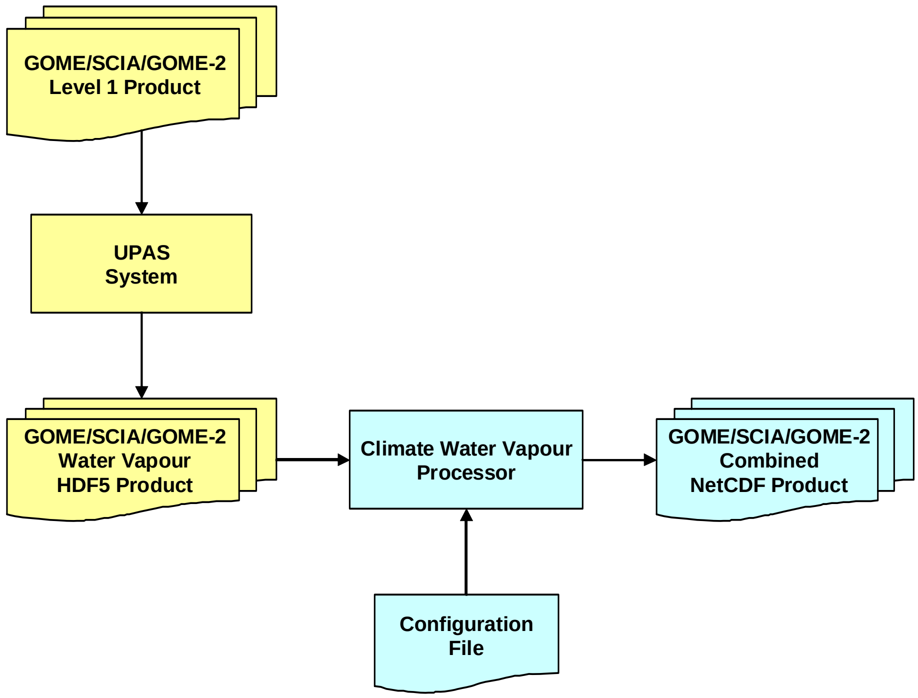


**Figure 2: GOME/SCIAMACHY/GOME-2 processing flow chart.**

*GOME-type Climate Water Vapour Processor (GCWVP)*

Input to the GCWVP are the L2 water vapour products from GOME, SCIAMACHY and GOME-2 in HDF5 format. The output is the combined dataset in NetCDF format. Settings for the processor and algorithm parameters like latitudinal/temporal bias corrections are read from an external configuration file.

The harmonisation of the total column water vapour (TCWV) time series from the various GOME-type instruments will be done following the approach specified in GlobVapour Technical Specification Document (TSD) from February 2010. The corresponding algorithms are integrated into the GCWVP. The usage of the Kriging software package from CM-SAF will be investigated as an alternative option.

## 3.2 Development of MERIS-SSM/I processing system

Figure 3 shows the flow chart for the combined SSM/I-MERIS product. It comprises of the elements of the preparation of level-1 data, the water vapour retrieval schemes, a quality monitoring module, a bias correction module between SSM/I and MERIS water vapour estimates, a Kriging module to be used for the same sensor merging of SSM/I estimates and a blending or combination module to bring SSM/I over ocean and MERIS over land onto one geographical distribution.

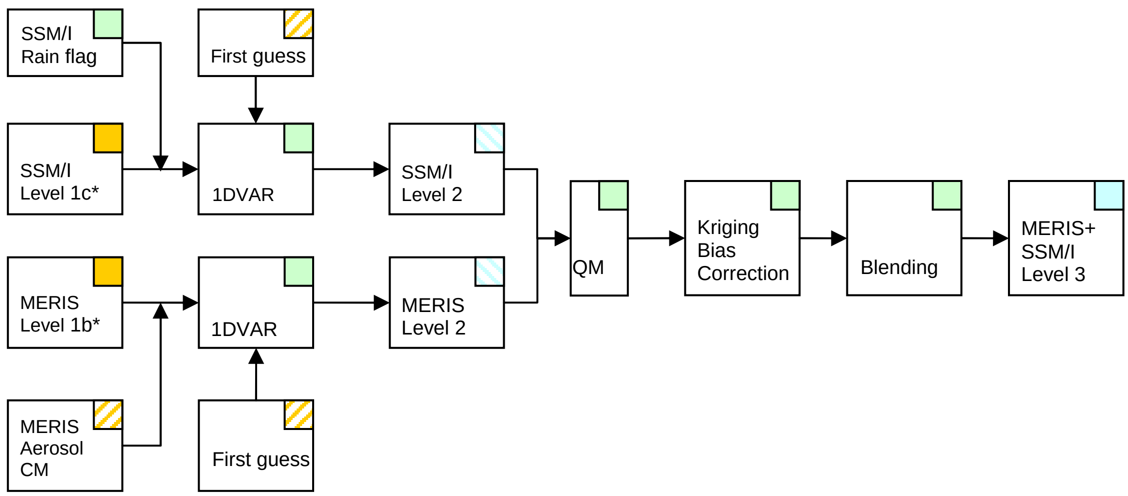


**Figure 3: Flow chart for the merged MERIS-SSM/I processor. Input data is marked orange (orange shaded: higher level input data), products are marked blue (blue shaded: instantaneous level 2 products) and software development is marked green. (QM: quality monitoring, CM: cloud mask).**

*Level 1 data preparation*

A major step between the prototype and the test products is the need to transform the radiance observations from each instrument into a homogenised or even inter-calibrated time series, that is, producing FCDRs from SSM/I and MERIS observations.

The current SSM/I processing at CM-SAF already applies a homogenisation to the radiance observations. The statistical approach utilises SSM/I on F11 as reference and is described in Fennig (2001) and Andersson et al. (2009). NOAA-STAR determined inter-calibration coefficients using Simultaneous Conical Overpasses (SCOs) among SSM/Is on different satellites. Also, at Colorado State University the SSM/I Level 1C development aims at inter-calibrated radiance records from SSM/I. It is intended to use one those approaches for GlobVapour. For the software development that means a potential implementation of existing software to create the SSM/I FCDR. If the SSM/I FCDR used within the CM-SAF will be utilized no further software needs to be developed and only the existing level-1 data will be utilized in the retrieval scheme. A decision on this is expected at the point where the production of prototype products is finished.

For MERIS only one instrument record is handled in GlobVapour so nothing of what is described above will be needed for MERIS.

*SSM/I 1DVAR module*

The core of the 1D-VAR scheme used for SSM/I is based on developments at ECMWF by Phalippou (1996), which were initially focussing on SSMIS microwave imager/sounder and AMSU. It was further extended to be a stand-alone scheme applicable to SSM/I; SSMIS and AMSU (Deblonde, 2001). The actual utilisation of the scheme is given in NWP-SAF User’s Guide (2001).

The 1D-Var scheme retrieves humidity, near-surface wind speed, and cloud liquid water. It also provides an additional mode in which the total water content is retrieved instead of water vapour and liquid water separately. When retrieving the total water content a separation module is applied subsequently for dividing the result into a vapour and liquid part. As a first step, the mode the scheme is run in has to be chosen.

The 1D-Var scheme can only be applied to rain free pixels. To filter out rainy pixels a SSM/I rain flag module is applied that is based on a brightness temperature threshold for identification and discrimination of areas with precipitation. Several schemes from the literature can be implemented but nothing has to be developed.

This 1D-Var scheme will be implemented at DWD, which itself as well as corresponding interfaces will have to be adjusted. The main task here will be the development of an interface software that provides first guess data to the 1D-Var. This contains the collocation of the SSM/I observations to ERA-Interim forecast data, extracting spatially interpolated profiles, and converting these to the input format as used by the 1D-Var.

If the used background error covariance matrix as included in the 1D-Var turns out to be insufficient, a tool may have to be developed that analyses ERA–Interim forecast quality and estimates the corresponding error covariances.

*MERIS 1D-Var module*

The 1D-VAR module for MERIS will be a further development of the existing water vapour retrieval algorithm implemented in the MERIS ground segment (Bennartz and Fischer, 2001). It is based on the differential absorption technique (DOAS), exploiting the MERIS measurements of bands 14 (885 nm) and 15 (900 nm) near and within the $\rho\sigma\tau$-water vapour absorption around 935 nm. In order to provide proper error propagation, an optimal estimation algorithm (Rodgers et al, 1976) will be developed in the GlobVapour framework, allowing for a pixel-by-pixel accuracy estimation. The optimization is based on a gradient descent, where the measurement error (co-variances) as well as the forward model errors and the background knowledge are integrated into a cost function, assigning the lowest cost to the solution with the highest probability, following the Bayes' theorem. First guess and background information will be provided by ERA-Interim data.

For the optimal estimation approach, a very fast forward operator for the simulation of the water vapour absorption is needed, since standard RTMs lack the computational efficiency to be used in iterative fitting techniques in the VIS/NIR spectral region. Both a single scattering model and an artificial neural network will be tested with respect to their accuracy and speed, as both methods are fast and allow for an analytic calculation of the Jacobian matrices. The usage of an artificial neural network for the efficient forward modelling within an optimal estimation algorithm has for example been successfully demonstrated for MERIS cloud retrievals by Lindstrot et al. (2010). However, both forward approximations have to be carefully validated against full RTM calculations using the FUB radiative transfer code MOMO (Fell and Fischer, 2001).

In case the planned forward approximations shape up as insufficient, the fall back procedure is the standard MERIS L2 retrieval, which is based on a multi-dimensional nonlinear regression. The regression coefficients were retrieved using the results of a large number of RTM simulations, covering the naturally occurring variability of surface properties, aerosol amounts and properties, temperature and water vapour profiles and amounts. In case the standard MERIS L2 algorithm is used, the pixel based error analysis will be based on look-up tables from stratified error analyses.

*Quality monitoring module*

The TCWV above land surfaces, derived from MERIS, and above sea surfaces, derived from SSM/I will be thoroughly quality checked. These checks will be performed at L2 stage, separately for each instrument. The initial test will screen out retrieval values exceeding thresholds, e.g. at 5 and 95 kg/m². Additionally, both TCWV products from MERIS will be checked toward inconsistencies in the spatial distribution of the product. This is performed on the spatial variability of the TCWV field, converted to the virtual field at sea level in order to eliminate the influence of the topography.

Further quality indication for both 1D-Var systems will be provided by the OE cost function, revealing some information about the quality of the measurement fit and the deviation of the found solution from the background knowledge. Furthermore, the deviation of the derived water vapour from the ERA-Interim background knowledge will reveal further information about the retrieval quality. Software is needed to incorporate these numbers into the level-2 output.

On top of that, the MERIS L2 data archive will be analysed with respect to the temporal variability of the derived water vapour fields. A quality flag will be raised in case the deviation of the derived water vapour from the values derived during previous orbits exceeds a threshold of two standard deviations.

*Kriging for same sensor merging (SSM/I)*

For both data products (SSM/I and MERIS) an objective analysis method (Kriging) will be applied to achieve an interpolation with the possibility to also provide uncertainty information for the merged products. We will employ the approach as described in Lindau and Schulz (2004) which is the version currently used by CM-SAF operations. The Kriging will be adjusted to account for the retrieval errors when merging and to provide also corresponding grid box representative measures of uncertainties, which are a combination of the initial pixel-based retrieval errors and the sampling errors.

A software tool will be developed to determine the spatial TCWV correlation functions. This tool will analyse an exemplary dataset, e.g., one month of data. Based on the results the correlation functions will be defined. A likely approach will be an exponential decrease of the spatial correlation function. The actual correlation length will then be derived for each month employing the correlation function.

Furthermore, an analysis of the variance is necessary for both uncertainty estimation and normalisation. The implementation of the systematic difference of the input data streams within the Kriging approach will also be analysed.

The Kriging module applied to SSM/I data over ocean and MERIS data over land brings both data sources onto regular grids with spatial resolution described in the TSD [AD-6].

*SSM/I-MERIS TCWV bias correction module*

A module will be developed in order to eliminate systematic differences between the MERIS and SSM/I TCWV retrievals. This tool will contain firstly, a modified version of the MERIS TCWV algorithm to enable its application to ocean cases with sun glint, and secondly, a comparison tool to assess possible systematic deviations in the TCWV products of MERIS and SSM/I. This includes a collocation module that collocates the MERIS sunglint pixels with the nearest SSM/I observation in time and space applying limits that appear useful from the above describes correlation analysis. This information will then be used to tune the MERIS data to avoid significant biases in the final Level 3 product.

*Combination module*

The combination of SSM/I and MERIS will be performed at level-3. A combination module will be developed that merges the TCWV products of SSM/I and MERIS. Within this module the final global coverage will be provided by SSM/I TCWV grid points over ocean together with the complementing MERIS grid points over land surfaces. For coastal areas the module will select grid points from MERIS.

Despite a lower accuracy of the MERIS retrievals over coastal pixel compared to land surfaces, these retrievals are expected to be qualitatively better than SSM/I retrieval in these regions. The chosen spatial resolutions of these data of $(0.5°)^2$ for SSM/I and $(0.05°)^2$ support an optimal and complete intersection between the two products.

## 3.3 Development of ATSR/AATSR processing system

Since the prototype of the (A)ATSR(2) TCWV will be developed as a BEAM plug-in, it will be already in a production environment. The scientific community using (A)ATSR(2) for sea surface temperature measurements made extraordinary effort to ensure that these are homogenised and inter-calibrated. Hence this does not have to be accounted for in GlobVapour.

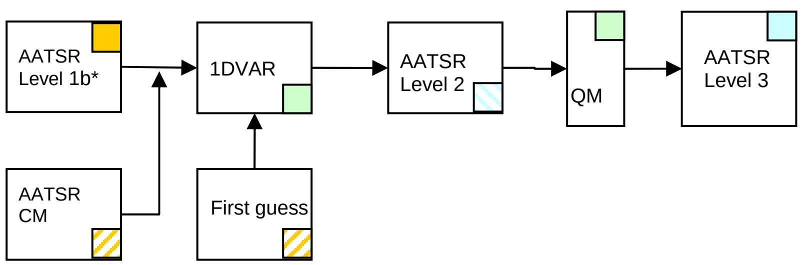


**Figure 4: Flow chart for the merged ATSR/AATSR processor. Input data is marked orange (orange shaded: higher level input data), products are marked blue (blue shaded: instantaneous level 2 products) and software development is marked green. (QM: quality monitoring, CM: cloud mask).**

*AATSR 1DVAR module*

An (A)ATSR(-2) retrieval scheme will be developed for the estimation of the TCWV from the thermal channels at 11 and 12 µm. Both spectral bands are used to derive the sea surface (SST) and land surface temperature (LST). The accuracy of the derived surface temperature is quite sensitive to TCWV, whereby an overall increase in bias is observed with both increasing LST and water vapour loading (Noyes et al., 2007). The impact of water vapour on the atmospheric radiation is different at both spectral bands. To account for the water vapour influence in the retrieval of SST and LST, most of the algorithms are based on the ‘split-window technique’. Following Li et al. (2003), the TCWV can be estimated from AATSR nadir measurements by

$$wv = 13.73 - 13.622\, \tau_{12} / \tau_{11}$$

where $\tau_i = \tau_{11}$ and $\tau_j = \tau_{12}$ are the transmittance for the bands at 11 and 12 µm and *wv* is the atmospheric water vapour content.

Additional information about the TCWV will be provided by the 3.7 µm channel of (A)ATSR(-2), showing a rather strong sensitivity to water vapour as compared to the influences of other gases. It will be mainly used during nighttime observations, since the correction of the solar contribution in daylight conditions is difficult.

In a preliminary simulation study, the optimal strategy for the algorithm development will be assessed. The selected algorithm will be realised in a full optimal estimation framework using ERA-Interim data as source of first guess background information.

*Quality Monitoring module*

The (A)ATSR(-2) quality monitoring module will be similar to that of MERIS. It will thus be based on the analysis of the retrieval cost function, the spatial and temporal homogeneity of the retrieved water vapour fields and the deviation from the ERA-Interim data.

## 3.4 Development of IASI-SEVIRI processing system

Figure 5 shows the flow chart for the combined IASI-SEVIRI product. It comprises of 1D-Var water vapour retrieval schemes, a quality monitoring module, a bias correction module between IASI and SEVIRI water vapour estimates, and a Kriging module to be used for the sensor merging of IASI and SEVIRI estimates.

The prototype software will be adjusted to the output formats from the IASI processing system. After adaptation to lessons learned from validation, the software will be transferred into a stand alone program.

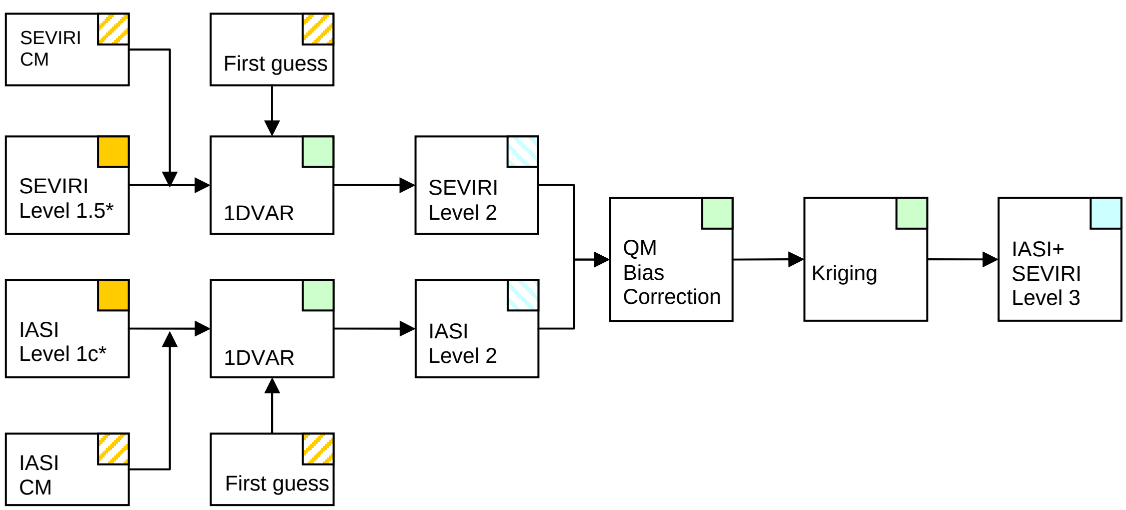


**Figure 5: Flow chart for the merged IASI-SEVIRI processor. Input data is marked orange (orange shaded: higher level input data), products are marked blue (blue shaded: instantaneous level 2 products) and software development is marked green. (QM: quality monitoring, CM: cloud mask).**

The IASI time series determines the length of the merged product. Though relatively short, the temporal stability of SEVIRI and IASI observations should be assessed. Here, results from GSICS and the bias monitoring of SEVIRI radiances from CM-SAF can be considered, if inconsistencies occur.

*IASI 1D-Var module*

The IASI 1D-Var retrieval scheme, which will be selected after the IASI retrieval assessment, will be built into an appropriate framework. The input and output interfaces will be modified so as to read collocated ERA-Interim forecast profiles and provide the output in NetCDF format when necessary. Optional extensions to this scheme, such as the inclusion of detailed land surface emissivity data and the optimisation of the computational cost, will be implemented if necessary and feasible during the course of this project.

*SEVIRI 1D-Var module*

The SEVIRI 1D-Var module is based on an existing 1D-Var retrieval scheme developed in the framework of the NWC-SAF. The core of the used scheme in GlobVapour is identical to the SEVIRI 1D-Var used in WACMOS, which is another ESA project investigating the water cycle and soil moisture, and will build on the framework it is currently implemented into within the WACMOS project. Due to the extend application period in GlobVapour compared to WACMOS, which emphasises a computational efficiency of the retrieval, further extensions to the scheme will possibly be necessary.

*Quality Monitoring and Bias Correction module*

The SEVIRI and IASI data undergo the QM and Bias Correction module from WACMOS. Some QM properties of this module are: a threshold test rejecting all retrieval with exceeding pre-defined limiters and an optional outlier check. Furthermore, a bias correction is performed to support an optimal Kriging performance. Here, the biases between the SEVIRI and IASI retrievals are investigated, while the latter serves as trusted reference. Subsequently, the determined biases of the SEVIRI product are removed.

*Kriging module*

The Kriging module for the IASI-SEVIRI processing system is, similarly as for the merged SSM/I-MERIS product, based on an objective analysis method, which will be applied to achieve an interpolation with the possibility to also provide uncertainty information for the merged products. We will employ the results of the WACMOS project. Within WACMOS the work of Lindenbergh et al. (2008) who merged MERIS TCWV observations with hourly ground based GPS observations of a single day, is applied and combined with the currently used CM-SAF operational Kriging approach described in Lindau and Schröder (2010).

The first step of the sensor merging procedure is the determination of the horizontal and temporal correlation functions. The correlation functions will be analysed on an exemplary data set, e.g., one month of data. Based on the results the correlation functions will be defined. A likely approach will be an exponential decrease of the spatial correlation function and a spherical function for the temporal correlation. Furthermore, analysis of the variance is necessary for both uncertainty estimation and normalisation. The implementation of the systematic difference of the input data streams within the Kriging approach will also be analysed.

# 4 Software Quality Assurance

## 4.1 Standards and conventions

The GlobVapour products will follow a service-oriented approach adopting common standards on metadata and data models as described by INSPIRE [AD-7] and the GEOSS Architecture and Data Committee. Following these rules will ensure global connectivity and interoperability, also providing appropriate links to the WMO Information System (WIS). The INSPIRE-compliant infra-structure will enable considerable data-sharing efficiencies to be realised between GlobVapour and the various GMES services requiring access to climate-related data. As an example the NetCDF Climate Forecast convention will be used as a standard for all GlobVapour end products. The product quality indicators will follow the CEOS QA4EO standard [AD-5].

For software related system requirements analysis, the software management process, software design and implementation engineering, software validation and software delivery and acceptance process the ECSS-E-ST-40C standard [AD-2] will be applied. The EODAPS-GV processing environment will be fully compatible with the ECSS-E-ST-40C standard.

Top down modular design will be used for the development of the processing system. The software shall be designed to provide maximum flexibility for modification and expansion. System and path variables, as well as shared parameters for processing control, scaling and calibration, shall be written in configuration files. Software component design notes and test results will be documented by the software lead engineers.

All source code will be developed in a consistent manner and will contain headers that include the file name, purpose, version number, change history, and list of procedures in the file. The data processing functions of the software modules will include a data verification step so as to prevent uncontrolled behaviour of the software due to inconsistent or corrupt data. The software modules will only exit in a controlled manner. Clear and consistent log messages shall be written throughout the processing.

## 4.2 Configuration Management

### *4.2.1 Configuration Control*

All software items of the entire EODAPS-GV system, including any configuration files, batch files, command files, data files, or other software files needed for the installation and operation of the system, shall be under formal configuration control (FCC), using one standard COTS tool. The following guidelines will be applied:

- Version numbering shall be maintained for all components.
- Separate version numbers shall be used for the scientific processors and processing system modules.
- Version numbers reflecting major and minor software status shall be listed in the source code headers.
- The version numbers of all software components shall be listed in an engineering configuration control report.

- The version numbers shall be reflected in all test reports and review documents.

It has been decided that major software submissions of the modules and the processing system will be under central version control at Brockmann Consult. A closed source policy of the module code is foreseen. The test data will not necessarily be involved. The version control database will be fully accessible by all project responsible persons.

### *4.2.2 Problem Reporting and Correction*

Software problems will be managed as part of FCC, and reported to the software QA responsible. Software problems identified during integration and testing will be documented using the software problem report form.

Corrective actions will be applied in response to reported software problems in a timely manner. The resolution of software problems as well as the implementation of change requests will be tracked using a software change control mechanism.

The software problem reporting and bug tracking will also be performed centrally at Brockmann Consult, with complete visibility to the project responsible persons.

## 4.3 Verification

### *4.3.1 Design Verification*

The aim of software verification is to ensure that adequate specifications exist for any software development activity and that the outputs of the activities are correct and consistent with the specifications. The verification plan and verification reports for each software component are documented in the Software Verification Plan (SVP) which is part of the Design Justification File (DJF).

Adequate testing of the software will be performed in order to ensure that all necessary software requirements are achieved. Verification tests are carried out on three different levels: Unit-level tests, integration tests and acceptance tests.

### *4.3.2 Implementation Verification*

A code walkthrough will be held at an early stage of each development phase (being the development of the prototype and of the stand-alone processing environment).

The source code will be tested during code implementation for each development phase, in unit test configuration on respectively prototype and stand-alone processor hardware. Unit-level tests are defined and carried out by developers. Unit-level tests aim at asserting the proper implementation of individual software units: a procedure or function in a procedural design or a class in an object-oriented language. For the Java programming language JUnit framework is used. Cobertura is used for giving an overview of the lines of source code that are effectively covered by unit level tests. GUI components are designed so that any logic is separated from the view, which implies that the logic can be tested without knowing the view.

The results of the unit testing will be documented per software module in Unit Test Reports (no formal deliverables).

### 4.3.3 Integration Testing

The source code of the (prototype and stand-alone) processing environment will be tested in system test configuration prior to delivery for integration during each development phase.

Integration tests of the complete system, i.e. retrieval scheme modules integrated into the EODAPS-GV processing environment, will be based on the Requirements Baseline Document (RBD). The test plan and conditions are to be defined by testers in consultation with developers.

Integration tests, which are supported by the QF test tool, will be conducted by testers, and the results will be documented.

### 4.3.4 Acceptance Testing

Acceptance tests are defined by the ATP and ensure that the software has achieved all necessary requirements. Acceptance testing of the final system delivery shall not commence until integration testing of the processing environment has been successfully executed and formally passed the QR.

The acceptance tests will be conducted by BC. The final acceptance test requires supervision by ESA, and the results will be written in an Acceptance Review Report.

## 4.4 Delivery and Deployment

The executable software will be prepared, including configuration files, batch files, command files, data files, or other software files needed to install and operate the software on its target computer(s). A list of such items together with all software items as part of the full system delivery, along with their configuration controlled version numbers, will be summarised.

A Product User Guide will also be prepared and delivered.

# 5 Validation tools

The development of the validation tools will take place in the frame of WP 230 (01.04.2010-30.11.2010).

At first, existing software to load and read validation data will be collected and missing software for loading and reading will be developed. The validation data will be quality controlled and depending on the outcome, not considered for validation and for the DDS.

For the collocation of the Diagnostic Dataset (DDS) with GlobVapour products existing collocation software will be collected and adapted or developed if needed.

Software for the validation of the products will be gathered and developed, that is, software for the determination of parameters from a common metric for validation. In order to inter-compare the GlobVapour products a triple collocation software will be developed following Stoffelen (1998 )and O'Carroll, (2008).

For the IASI assessment it is proposed to follow the approach of Pougatchev (2008) and Pougatchev et al. (2009). It is further planned to try to adapt this method to the comparisons of total column water vapour estimates to ground-based observations and to comparisons among the products.

The validation tools will be under configuration control and will be tested prior to validation with input data subsets of the DDS.

# 6 Conclusions

Requirements and guidelines for the software development needed throughout the project has been provided. The architecture of the various processors and processing environment as well as the EO Data Processing System (EODAPS) for GlobVapour are described. Associated standards for software development and data production were introduced. Finally, the software verification procedure is outlined.