Supplement Report to De3.32 “Generic Validation Strategy for CDRs/ECVs”

Supplement Report to Deliverable De3.32

Issue 1.0

**CORE-CLIMAX**

**COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS**

**FP7 Grant agreement No 313085**

# 1. Copernicus

## 1.1. Global Land vegetation products (e.g. fAPAR)

**Reference documents**

<http://land.copernicus.eu/global/products/fapar>

Validation report: <http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_VR_FAPAR_I1.00.pdf>

ATBD: <http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_ATBD_FAPAR_I1.00.pdf>

Product user manual: <http://land.copernicus.eu/global/sites/default/files/products/GIO-GL1_PUM_FAPAR_I1.00.pdf>

**Validation Strategy**

The consortium is responsible for the self-assessment of the products. This includes the (1) technical and service quality, (2) the scientific product quality per variable and (3) a cross cutting analysis to verify the consistency across variables. The independent assessment is implemented by a separate entity. The independent experts review all documents on a 6-monthly basis, and have the opportunity to perform validation activities themselves, independently from those executed by the Global Land Service. Evolution actions can be defined after each review process when necessary.

**The scientific product quality** is split in two parts. The first is an exhaustive evaluation of new products before the operational production. This validation has to be approved by the independent assessment team before production. Second part is the quality monitoring, which is a continuous and automated evaluation to check if the operational product quality is stable over time. This process is repeated every 6 months. The continuous quality monitoring is basically a lighten validation procedure compared to the exhaustive quality assessment analysis. It focuses on the metrics and criteria that can be easily automated so that the process becomes less and less time consuming.

**The independent assessment** is implemented by a separate entity (under a different contract, Lot II) providing a high quality and independent reviewing process to all aspects, both from engineering and scientific points of view of the operational activities. This entity, led by Spacebel N.V., also manages a technical user group to provide feedback at technical level on the use of the Global Land Service. The independent experts review all documents on a 6-monthly basis, and have the opportunity to perform validation activities themselves, independently from those executed by the Global Land Service.

**Ground Data**

In-situ data processed according with CEOS/WGCV LPV guidelines accounting for up-scaling effects was used. The “state of the art” direct validation approach (Morisette et al., 2006) consists in using high spatial resolution imagery to scale the ground measurements up to moderate resolution pixel. For this, a “transfer function” between high spatial resolution surface reflectances and biophysical measurements is established, and then is applied to an appropriate extent of the high resolution image. The resulting high spatial resolution map is finally aggregated to the larger pixel size (typically over 3x3 km) for comparison with moderate resolution products.

Most of the FAPAR maps are coming from the VALERI project. Additional data coming from SAFARI-2000 (S2K) Kalahari campaign (Privette et al., 2004) and the SMOS validation activities in the Valencia Anchor Station (Utiel, Spain) were used.

The description of ground data and processing for S2K can be found at g2-BP-RP-BP22-1 and for SMOS at g2-BP-RP-BP22 -5 & - 6. <http://web.vgt.vito.be/documents/BioPar/g2-BP-RP-BP022-ValidationReport-HRProducts-I1.10.pdf>

The ground data were first evaluated to select the most reliable values. Most of the sites are located in Europe and Africa, and very few sites in Asia and Oceania. The reference maps belong to the land cover type ‘forest’, ‘cultivated’ and ‘grass’ (including herbaceous, shrubs and savanna cover types).

The uncertainties of the ground data reference maps was also assessed.

**Validation process**

(1) Spatial consistency analysis

1. *Spatial distribution of retrievals*: Difference maps with global products (mean bias error, B) and root mean square error maps (RMSE) were analysed over the period. Global maps of overall consistency of the GEOV1 products with existing global products were derived based on the relative root mean square error (RRMSE) along the period (1 image per month). An optimal consistency is achieved if the RRMSE is lower than 10% for FAPAR where differences in product definition/assumptions are higher.
2. *Spatial continuity*: The missing values or pixels flagged as invalid over land were quantified. Global maps of percentage of missing values over the period were produced at the original projection and temporal resolution. The information of missing values coming from the quality flag was also analysed. The percentage of the missing values as a function of the latitude and period of the year was discussed.
3. *Magnitude of retrievals*: The value of vegetation variables over desert areas and densest forests were analysed in search of possible bias in the products. d. *Regional assessment*: The spatial consistency between products was evaluated over several regions (20ºx20º). Difference maps and metrics were analysed.

(2) Global statistical analysis: The statistical analysis per aggregated land cover types was performed over the BELMANIP-2 network of sites that was designed to represent globally the variability of land surface types (Baret et al., 2006). The land surface type is defined here using 8 generic classes derived from the GLOBCOVER classification, namely: Evergreen Broadleaf Forest (EBF), Broadleaf Deciduous Forest (BDF), Needle leaf Forest (NLF), Mosaic (M), Herbaceous (H), Shrublands (S), Sparse and Bare areas (SBA).

1. *Histograms*: Statistical distributions of retrievals were computed over main aggregated land cover types. Histograms were computed during the whole considered period considering all observations.
2. *Scatter-plots* between different products over all sites and the whole period at a monthly frequency were displayed to analyse the overall performance with existing products. The temporal evolution of the statistics was also analysed in order to find seasonal trends.
3. *Statistical assessment per biome*: Metrics (R2, B, RMSE, RRMSE) among different products were computed over the main aggregated land cover types.

(3) Temporal consistency analysis

1. *Temporal variations:* Seasonal variations and Inter-annual variations were analysed displaying temporal profiles over BELMANIP-2 sites plus over additional sites where ground reference maps were available. The products were evaluated over 3x3 pixels at the original temporal resolution. Information of ground maps was also displayed as reference even if the year was not coincident (only for natural vegetated areas). The precision of the products (i.e., inter-annual consistency) was qualitatively investigated over natural areas. Attention was paid to the temporal profiles of the GEOV1FF - filtered & filled products – to know if the interpolated value is reliable or not.
2. *Temporal realism*: The realism of the temporal variation was assessed by comparison with multi-temporal ground truth data available at Mongu site (Kalahari, Zambia). It is expected to acquire additional multi-temporal ground data to increase the representativeness of this exercise.
3. *Temporal smoothness*: The smoothness of products was evaluated by taking three consecutive observations and computing the absolute value of the difference delta between the centre P(dn+1) and the corresponding linear interpolation between the two extremes P(dn) and P(dn+2) as follows
4. *Error bar reliability*: The reliability of the GEOV1 error bar was analysed considering on one hand the uncertainty introduced by the different remote sensing product estimations and, on the other hand, according to the ground reference maps.

(4) Direct Validation:

1. Accuracy: Quantitative uncertainties of GEOV1 and reference products were computed against ground reference data representative of an area of 3x3 km and accounting for up-scaling effects. The scattering is carried out by comparing the field data with the closest product date. In-situ values for natural vegetation sites measured outside the period 2003-2004 are compared with satellite product estimates for the year in which the agreement was better.

(5) Cross cutting evaluation using the LDAS model.

1. Metrics derived from the passive monitoring of FAPAR: Spearman correlation, z-score, RMSD, SDD, mean bias.

**Error Characterization**

Table 1 Description of the quality flag provided for the versions 1 of LAI, FAPAR, FCOVER.

|  |  |  |
| --- | --- | --- |
| **Quality flag**  | **Nature** | **Bits**  |
| Bit 1: Land/Sea  | Land | 0 |
| Bit 2: Snow status  | Clear | 0 |
| Bit 3: Suspect  | Suspect or No suspect | 0 or 1 |
| Bit 4: Aerosol status  | Pure or mixte | 0 or 1 |
| Bit 5: Aerosol source  | Modis or climate | 0 or 1 |
| Bit 6: Input status  | OK | 0 |
| Bit 7: LAI status  | OK | 0 |
| Bit 8: FAPAR status  | OK | 0 |
| Bit 9: FCOVER status  | OK | 0 |
| Bit 10: B2 saturation status  | OK | 0 |
| Bit 11: B3 saturation status  | OK | 0 |
| Bit 12: Filtering status  | OK | 0 |
| Bit 13: Gap filling status  | OK | 0 |

Product uncertainties associated with fAPAR are derived from the theoretical performances evaluation: for each case in the training data base, the RMSE values are computed for the cases in the input feature space that are within the uncertainty domain. The sun zenith angles selected have also to be within ±5° around the case considered.

**Global Land Cross-cutting validation**

The direct validation of land CDRs is not easy, as in situ observations are limited in space and time. Therefore, indirect validation has a key role to play. It consists in comparing the products with similar pre-existing products derived from satellite observations or from land surface model (LSM) simulations. State-of-the-art LSMs are able to represent the diurnal cycle of the surface fluxes together with the seasonal, inter-annual and decadal variability of the vegetation biomass. They are able to diagnose a number of land ECVs such as LAI, fAPAR, surface albedo, land surface temperature (LST), and soil moisture. The most advanced indirect validation technique consists in integrating satellite-derived ECV products into a LSM using a data assimilation scheme. The LSM and the data assimilation scheme (e.g. an Extended Kalman Filter or an Ensemble Kalman filter) are embedded into a modelling platform forming a Land Data Assimilation System (LDAS). The reanalysis provided by the LDAS accounts for the synergies of the various upstream products and provides statistics which can be used to monitor the quality of the assimilated observations. The current version of the LDAS used for the cross-cutting validation in the Copernicus Global Land Service is operated by MTF and assimilates SPOT-VGT LAI and ASCAT surface soil moisture (SSM) products over France (8km x 8km). A passive monitoring of albedo, FAPAR and LST is performed (i.e., the simulated values are compared with the satellite products). The system is used to monitor the quality of upstream products. The LDAS generates statistics whose trends can be analyzed in order to detect possible drifts in the quality of the products: (1) for LAI and SSM, metrics derived from the active monitoring (i.e. assimilation) such as innovations (observations vs. model forecast), residuals (observations vs. analysis), and increments (analysis vs. model forecast) ; (2) for albedo, LST, and FAPAR, metrics derived from the passive monitoring such as the Pearson correlation coefficient, z-score, RMSD, SDD, mean bias.

## 1.2. MACC II Atmospheric Composition

### Aerosols

**MACC-II**

The Copernicus **MACC-II** - Monitoring Atmospheric Composition and Climate - Interim Implementation - is the current pre-operational Copernicus Atmosphere Service.  So far MACC-II has only experimental aerosol propducts (ATSR-DV, SYNAER, SEVIRI NRT), not CDRs (<http://www.gmes-atmosphere.eu/catalogue/#ProductListPlace:sp=IjkyIg==@0@h1ds5nvCW7duIxDmdC8czB9cHoY>=). Thus there validation is not reviewed here.

### Carbon dioxide

Carbon dioxide products (column-averaged mixing ratios and tropospheric columns / stratospheric profiles of CO2) of ESA GHG-CCI are validated using essentially the same methodology as for the methane products. Therefore, the information regarding validation as described in section 2.1.2.2 is directly applicable as-is, even though the retrieval algorithms and product contents are different.

### Water vapour

<EUMETSAT has the lead, FMI will support if needed>

## 1.3. MyOcean II Baltic Sea Regional NRT SAR Based Sea Ice

**Reference documents**

Product Webpage: <http://www.myocean.eu/web/69-myocean-interactive-catalogue.php/?option=com_csw&view=details&product_id=SEAICE_BAL_SEAICE_L4_NRT_OBSERVATIONS_011_004>

User Manual: <http://catalogue.myocean.eu.org/static/resources/myocean/pum/MYO2-SIW-PUM-011-004-and-011-v2.2.pdf>

Quality Information Document: <http://catalogue.myocean.eu.org/static/resources/myocean/quid/MYO2-OSI-QUID-011-ALL-V1.5.pdf>

Validation Report: <http://myocean.met.no/SIW-TAC/doc/myo-wp14-siw-fmi-bal-seaice_hr-obs-validation_latest.pdf>

**General validation strategy**

Validation is done with direct comparison with reference data. Comparisons are resulted as distribution of differences between datasets, scatter plots and statistics on contingency tables.

**Generation of ECV**

Sea ice thickness CDR is based on the sea ice thickness history from the most recent IC available before the SAR. After SAR data is preprocessed for some corrections, an ice thickness value is assigned to each SAR segment. This thickness value is dependent on the ice thickness range adopted from the IC and on the mean SAR segment backscattering such that higher SAR backscattering corresponds to thicker ice. Finally the open water areas are updated according to the SAR open water classification.

Ice drift is produced after receiving two SAR images over a same area in the Baltic Sea with a time gap of less than three days between the SAR images. Each received SAR image is studied after it has been received and if it has common areas with an earlier SAR image less than three days older, the product is computed. The product is based on computing phase correlation of pairwise data windows sampled from the two images in two resolutions.

Ice concentration CDR is based on the digitized FMI ice charts produced by ice analysts.

**Validation data sets**

(1) Ground data

Icebreakers produce ice thickness drill measurements. Icebreakers produced 47 measurements until 27 March 2013.

Reference data used for ice concentration validation is based on SSMIS data and produced with ASI algorithm by University of Bremen (BR).

The SAR based short and long ice drift vectors are compared with the motion of the GPS buoys deployed on the ice in the Gulf of Bothnia and Gulf of Finland. Buoys collected location data from January 1 to March 17, 2013. The buoys indicate their location based on their GPS coordinates once an hour.

**Validation process**

(1) Comparison with in-situ

For ice thickness, each measurement from an icebreaker was associated to a pixel of the ice thickness map that represents the mean ice thickness. Some icebreaker ice thickness measurements had minimum and maximum ice thickness instead of single value, in that case mean ice thickness was used in the comparison. Statistical analysis is done by reporting scatter plots and distribution of the difference values. Regression is used to approximate the scattering as a line.

The comparison for sea concentration was performed by inspecting the distributions of the classified concentration subtractions and by confusion matrices of the classified concentrations. In both cases 3x3 homogeneous windows were used to eliminate the possible errors due to the re-rectification of the BR-maps. Statistical analysis is done by reporting distribution of difference values.

The ice drift monitoring period was from Jan 1st to Mar 19th 2013. During the monitoring period totally 117 ice drift estimates were used in the comparison, i.e. the two buoy locations was within the overlapping area of the two adjacent SAR images used in the drift estimation for the acquisition times of the SAR images. The ice drift estimates in comparison were divided into two categories: 68 of these were classified to short drift category (buoy motion less than 500m) and 49 measurements to long drift category. The quality given by the SAR algorithm for the short drift data the quality varied from 40% to 80% in both short drift and long drift categories. For the short drift data category only the motion magnitude was evaluated but for the long drift data both magnitude and direction were estimated. The direction couldn’t be evaluated in the short drift category because in short drift estimates the SAR registration errors can cause large relative errors and so defining the direction can become ambiguous due to the quantification of the direction. For the long drift data both the magnitude and direction were combined. Statistical analysis is done by reporting the accuracy of estimated vectors and scatter plots.

**Error characterization**

Products are given with additional quality flags in datasets, e.g. “1 to 5” for ice concentration and percentage for ice drift.

## 1.4. GMES in-situ Coordination

GISC was an FP7 funded project running from January 2010 till December 2013.

Currently, the project website is already closed down and information is hardly found. However, some Deliverables are available through the [internet](http://gisc.ew.eea.europa.eu/deliverables):

[Deliverable 2.1](http://gisc.ew.eea.europa.eu/deliverables/d2.1.pdf) provides a report on in-situ data requirements for a multiplicity of ECVs.

[Deliverable 2.2](http://gisc.ew.eea.europa.eu/deliverables/d2.2-d2.4.pdf) analyzed in situ requirements and [Deliverable 2.4](http://gisc.ew.eea.europa.eu/deliverables/d2.2-d2.4.pdf) reported criteria to determine priorities for support.

In April 2013 a GISC workshop ‘Monitoring Matters’ took place in Copenhagen ([workshop report](http://gisc.ew.eea.europa.eu/deliverables/monitoring-matters-workshop-report.pdf)).

**Action:** Retrieve all relevant information and documents (final report) of the GISC project. Identify current state of GISC. (Contact EEA focal point)

The objectives of the project were i) to enable data providers or network of data providers to provide the required in-situ data for GMES/Copernicus, ii) to identify priorities and requirements of in-situ data for GMES/Copernicus core services and iii) to integrate in-situ assets and networks into long-term sustainable frameworks for GMES/Copernicus services.

One task of GISC was the identification and classification of the in situ data required by the GMES/Copernicus services.

According to a [background note](http://ec.europa.eu/research/participants/portal/doc/call/fp7/fp7-space-2012-1/31663-gisc_-_r_d_requirements_of_in-situ_networks_en.pdf), GISC identified the following observation networks central to GMES/Copernicus (*to be updated*):

* EuroArgo (marine)
* EuroSites (marine)
* FerryBox (marine)
* SeaDataNet (marine)
* EARLINET-ASOS (atmosphere)
* EUSAAR (atmosphere)
* IAGOS (atmosphere)
* ICOS (atmosphere)
* E-SURFMAR (marine)

# 2. EUMETSAT SAFs

## 2.1. O3M

### Ozone

The first ozone data producer discussed is the O3M SAF (Satellite Application Facility on Ozone and Atmospheric Chemistry Monitoring) project of EUMETSAT.

**Reference documents**

O3M SAF documents available at http://o3msaf.fmi.fi/documents.html

O3M SAF Validation Report for O3M-41 and O3M-42, doc. no. SAF/O3M/AUTH/VRR/O3, issue 1 / 0

Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Tropospheric Ozone, NO2, Tropospheric NO2, BrO, SO2, H2O, HCHO, OClO and Cloud Properties. O3M SAF, Doc.No DLR/GOME-2/ATBD/01, rev. 2/H

**Generation of ECV**

The total ozone columns are retrieved from GOME-2 instruments by the GOME Data Processor (GDP) algorithm. The current version of the algorithm in operational use is 4.7. The algorithm is based on on Differential Optical Absorption Spectroscopy (DOAS), where gas-specific narrow band absorption features are detected in UV/Visible radiances. A least-squares fitting solves the gas density in the observed slant column, which is converted to a vertical column density following the computation of an appropriate Air Mass Factor (AMF) for that gas column (O3M SAF ATBD for total column products).

**General validation strategy**

The Near Real Time Total Ozone (NRT-O3) and Offline Total Ozone (OTO-O3) products are validated through direct comparisons of the satellite retrievals and ground-based measurements of total ozone (using Dobson and Brewer spectrophotometers). The validation is carried out by assessing the retrieval difference statistics on a global, latitudinal band and hemispherical basis rather than for singular validation sites (O3M SAF Validation Report for O3M-41 and O3M-42).

**Validation data sets**

The validation data of the O3M SAF ozone products is provided by the World Ozone and UV Data Centre (WOUDC) in Canada. The data is quality-monitored at the monitoring station and also by WOUDC. Daily mean Dobson/Brewer measurement data is selected as the satellite data comparison reference. 37 sites using Brewer and 21 sites using Dobson spectrophotometers are included in the O3M SAF product validation (during the latest validation period, December 2012 – April 2013).

The measurement uncertainty in the Dobson data is estimated to be ±1 - 2%, with systematic errors of up to 4% (Bernhard et al., 2005). Brewer measurement uncertainty is given at about 1%. Uncertainty in the reference data is expected to be larger in low Sun elevation (<15°) conditions.

**Validation process**

1. The satellite retrievals are matched against ground observations by finding the closest satellite product pixel to the ground measurement location within a 150 km radius of the ground site for each day in the validation period.
2. The daily mean validation results from the various sites are further collated into 10° latitude band averages as well as Northern and Southern hemispheric averages. Retrieval errors are also analyzed as functions of Solar Zenith Angle (SZA), cloud cover and cloud top height.
3. The retrievals from the two different GOME2 instruments (2A and 2B) are also intercompared, and compared to OMI-TOMS retrievals of total ozone.

**Error characterization**

The total column ozone product files contain a quality flag for each ground pixel relating to the retrieval validity range and estimated error in the slant column density. The content of this quality flag is shown in Figure 1.

Table 2: Quality flags in the O3MSAF products.



Ozone products have also been produced by the ESA Ozone\_CCI-project, which we shall discuss next.

### Ozone Precursor

The O3MSAF project processes global total and stratospheric NO2 data products from GOME-2 measurements.

**Reference documents**

O3M SAF documents available at http://o3msaf.fmi.fi/documents.html

O3M SAF ORR VALIDATION REPORT for NTO/OTO, products O3M-50, -51, -52, and -53. version 1.1, 30/06/2013.

Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Tropospheric Ozone, NO2, Tropospheric NO2, BrO, SO2, H2O, HCHO, OClO and Cloud Properties. version 2/H, 21/05/2013.

Boersma, K. F., Eskes, H. J. and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res., 109(D4), doi:10.1029/2003JD003962, 2004.

Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.: Tropospheric NO2 vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008–2011) and satellite validation, Atmos. Chem. Phys., 13, 1547-1567, doi:10.5194/acp-13-1547-2013, 2013.

**Generation of ECV**

The NO2 column retrievals are based on the GOME-2 backscattered Earthshine radiances and solar irradiance spectra. The retrieval algorithm calculates and corrects for the effects of clouds and air mass and separates the stratospheric and tropospheric NO2 columns (ATBD for GOME-2 Total Columns).

**General validation strategy**

The validation includes, but is not limited to, comparisons with correlative in situ measurements of total NO2 data. The production chain is also validated for the critical individual components, as described in the O3MSAF NO2 validation report (NTO/NO2 & OTO/NO2 Validation Report):

**Validation data sets**

The main reference data in stratospheric NO2 validation are the twilight measurements from DOAS UV/visible zenith-looking spectrometers. The data are made available by the NDACC network. The in situ reference data for stratospheric NO2 is estimated to have an uncertainty on the order of 5% (NTO/NO2 & OTO/NO2 Validation Report).

For the tropospheric NO2, MAXDOAS instruments are used as the reference (e.g. Brinksma et al. 2008, Ma et al., 2013). The MAXDOAS instruments are UV/visible spectrometers with multiangular observation capabilities for scattered light. Viewing directions of 30° or 45° elevation are typically used to retrieved the tropospheric NO2 column.

**Validation process**

The validation process is described quite clearly in the NTO/NO2 & OTO/NO2 Validation Report:

“The end-to-end validation approach adopted […] consists in:

(a) an assessment of the quality of GOME-2 DOAS analysis results, by confrontation of GDP 4.7 retrievals performed respectively on GOME-2A and GOME-2B spectra, both on an orbit-to-orbit base and time-series comparisons, and by confrontation of GDP 4.7 retrievals with GOME-2 retrievals of the scientific data processor TEMIS

(b) an assessment of the geophysical validity of total/stratospheric column measurements by comparison with stratospheric column measurements provided by zenith-sky DOAS UV-visible spectrometers affiliated with the Network for the Detection of Atmospheric Composition Change (NDACC); and

(c) an assessment of the validity of the GOME-2B tropospheric NO2 column data, with respect to MAX-DOAS observations performed by BIRA-IASB at OHP (France) and Beijing (P.R. China), and comparing them to GOME-2A results”

The validation report specifically lists challenges in direct comparison of nadir-observed NO2 columns from GOME-2 measurements at daytime with twilight reference observations from the NDACC spectrometers. Specifically, the main challenges are (NTO/NO2 & OTO/NO2 Validation Report):

“(1) its marked diurnal cycle, which can generate dramatic differences between NO2 column measurements acquired at different hours of the day,

(2) the wide range and variability of the NO2 burden in the troposphere, to which the nadir-viewing geometry of GOME-2 is one order of magnitude more sensitive than the zenith-sky viewing geometry of NDACC spectrometers, and

(3) the presence of spatial structures with significant horizontal gradients, requiring more accurate satellite/ground collocation criteria than classical geographical windows – e.g. a radius of 300 km around the station”

The validation process features corrections and data selection procedures to overcome these challenges. For specific information, please see the NTO/NO2 & OTO/NO2 Validation Report.

**Error Characterization**

The NO2 ATBD provides the following error budget estimation table (Table 3):

Table 6: Error budget table for the O3MSAF NO2 product



The NO2 products contain per-pixel quality flag information about the retrieval conditions (Table 4) as well as a quality flag describing the quality of the total column retrieval (Table 2).

Table 7: NO2 Tropospheric column retrieval condition quality flags



### other long live trace gas

The O3M SAF project’s products include offline products of bromine oxide (BrO), sulfur dioxide (SO2), and formaldehyde (HCHO). A demonstrational product for chlorine dioxide (OClO) exists but is not assessed here due to its non-operational status.

**Reference documents**

O3M SAF documents available at http://o3msaf.fmi.fi/documents.html

Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Tropospheric Ozone, NO2, Tropospheric NO2, BrO, SO2, H2O, HCHO, OClO and Cloud Properties. O3M SAF, Doc.No DLR/GOME-2/ATBD/01, rev. 2/H

O3M SAF ORR VALIDATION REPORT, O3M-58, issue/rev. 1 / 2

O3M SAF ORR VALIDATION REPORT, O3M-55, O3M-56, issue/rev. 1 / 1

O3M SAF ORR VALIDATION REPORT, O3M-82, issue/rev. 1 / 1

Yang, K., N. Krotkov, A. Krueger, S. Carn, P. K. Bhartia, and P. Levelt (2007), Retrieval of Large Volcanic SO2 columns from the Aura Ozone Monitoring Instrument (OMI): Comparisons and limitations , JGR 112, D24S43, doi:10.1029/2007JD008825

Clarisse, L., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Ngadi, Y., and Coheur, P.-F.: Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI), Atmos. Meas. Tech., 5, 581-594, 2012.

**Generation of ECV**

The trace gas column retrievals are based on the GOME-2 backscattered Earthshine radiances and solar irradiance spectra. The algorithm is based on Differential Optical Absorption Spectroscopy (DOAS), where gas-specific narrow band absorption features are detected in UV/Visible radiances. A least-squares fitting solves the gas density in the observed slant column, which is converted to a vertical column density following the computation of an appropriate Air Mass Factor (AMF) for that gas column. Note that the algorithm, called GOME Data Processor (GDP), calculates all the trace gases mentioned here as well as O3, NO2 and H2O column products during a single processing run.

The retrievals of the trace gas columns have gas-specific parts, adjusting e.g. the fitting wavelength window of the retrieval or adding specific offset corrections. Full details on the algorithm(s) are available in the O3M SAF ATBD for GOME-2 Total Column Products.

**General validation strategy**

The BrO column products are validated using the following methods (OTO/BRO Validation Report, 2013):

1. validation against in situ reference observations
2. intercomparison of different satellite-based BrO retrievals

The BrO validation was noted to be limited in scope because quality-assured reference data is scarce. Also, the satellite-based retrieval intercomparisons were hampered either by a mismatch in temporal coverage (SCIAMACHY vs. GOME-2) or instrument degradation of the comparison reference (GOME-2A vs. GOME-2B).

The SO2 validation is challenged by the large variability in column densities (1 – 1000 DU), reduced measurement sensitivity close to the surface, and the lack of in situ observations for volcanic eruptions with large SO2 columns. As a result, the current approach is to intercompare GOME-2B SO2 products with their counterparts from the GOME-2A, OMI, and IASI instruments. The product team notes that this does “pose a number of problems because the different satellite sensors have different overpass time, swath, spatial resolution and measurement sensitivity to SO2” (O3M SAF ORR Validation Report for O3M-56 and O3M-55).

The validation strategy for HCHO products is based on “a step-by-step validation of each subproduct in the HCHO production algorithm” (O3M SAF ORR Validation Report for O3M-58). In practice this means that the GOME-2B intermediate HCHO products are compared with corresponding retrievals from a reference algorithm for each step in the production process. Validation by comparisons to in situ reference observations are planned, using MAX-DOAS spectrometer measurements at a pilot site in Xianghe, China.

**Validation data sets**

For the (latest) BrO validation, the following data sets were used in comparison/validation:

* 5 years of averaged SCIAMACHY data (2007-2011) to assess the latitudinal and temporal features of GOME-2B retrievals for coherence.
* GOME-2A retrievals, although instrument degradation limits comparability even after post-processing corrections.
* Ground-based total BrO columns from a measurement site in Harestua, Norway, covering more than a decade.

For the SO2 validation, the following data sets were used in comparison/validation:

* GOME-2A retrievals, although instrument degradation limits comparability even after post-processing corrections.
* GOME-2B data processed with a reference algorithm (BIRA-IASB).
* SO2 columns for the volcanic cases from GOME-2A, OMI (Aura) and IASI (MetOp-A). The OMI data are from the Linear Fit algorithm (Yang et. al, 2007) and are filtered for the row anomaly pixels; the IASI algorithm is described in Clarisse et al. (2012).

For the HCHO validation, the following data sets were used in comparison/validation:

* GOME-2B and GOME-2A data processed with two versions of a reference algorithm (BIRA-IASB v07 and v13).

**Validation process**

The BrO validation process consists of a general verification of the coherence and feasibility of observed spatial and temporal column features against 5 years of SCIAMACHY retrievals, an intercomparison of GOME-2A and -2B retrievals both regionally and over specific overpass sites, and finally a direct comparison of total GOME-2A and -2B BrO columns with ground-based BrO measurements from a zenith-pointing spectrometer at sunrise (O3M SAF ORR Validation Report for O3M-82).

The SO2 validation consists of a direct comparison of the various intermediate products of the SO2 retrieval, much like in the HCHO validation process (see below). The reference algorithm (BIRA-IASB) intermediate outputs are compared with the operational GDP4.7 algorithm intermediate outputs, both using GOME-2B data, and the differences are analyzed (O3M SAF ORR Validation Report for O3M-55 and O3M-56).

For the total (vertical) SO2 columns, direct comparisons to GOME-2A, OMI and IASI are utilized. In the direct comparison between different instruments, a caveat is given regarding the need for accurate a priori knowledge of SO2 plume height in order for the retrievals to be considered comparable.

The validation process for HCHO consists of consequential verification of the HCHO sub-products against the BIRA-IASB reference algorithm (O3M SAF ORR Validation Report for O3M-58). The evaluated sub-products are:

* slant columns (SCD)
* normalized slant columns (ΔSCD) (see [ATBD] or De Smedt et al., 2008 for more details about the reference sector correction)
* air mass factors without cloud correction (AMF clear)
* air mass factors with independent pixel cloud correction (AMF)
* total vertical columns (VCD) (the bulk of the formaldehyde column lies in the lower troposphere, the contribution from the stratosphere is negligible).

**Error Characterization**

The O3M SAF trace gas product files contain both general quality flags, as well as quality flags specific for each trace gas product and algorithm. Generally speaking, all of the trace gas products contain pixel-specific retrieval error estimates provided by the GDP algorithm. The general processing quality flags are shown in Table 2 (section 2.1.2.1). Additional trace gas-specific quality flags are as given in Tables 4 and 5:

Table 4: Specific quality flags in SO2 data products of O3M SAF



Table 5: Specific quality flags in HCHO data products of O3M SAF



The BrO products have no reported specific quality flags.

### AI

**Reference documents**

<http://o3msaf.fmi.fi/docs/vr/Validation_Report_ARS_AAI_Jun_2013.pdf>

<http://o3msaf.fmi.fi/docs/atbd/Algorithm_Theoretical_Basis_Document_ARS_Jun_2013.pdf>

<http://o3msaf.fmi.fi/docs/pum/Product_User_Manual_ARS_Sep_2013.pdf>

**Generation of ECV**

The Absorbing Aerosol Index separates the spectral contrast at two ultraviolet (UV) wavelengths caused by aerosol scattering and absorption from that of other effects, including molecular Rayleigh scattering, surface reflection, and gaseous absorption. AAI from the Main Science Channels is derived from reflectances measured by GOME-2 at 340 and 380 nm using the residue method.

**General validation strategy**

The verification consists of a comparison of daily GOME-2 AAI maps with similar maps from the OMI-TOMS AAI, a direct intercomparison of the AAI from SCIAMACHY and GOME-2, and verification of the statistical properties of the GOME-2 data.

**Validation data sets**

SCIAMACHY is the ideal instrument to compare with, for a number of reasons:

* SCIAMACHY is, like GOME-2, in a descending orbit with an almost identical orbital period.
* The equator passing times of SCIAMACHY (10:00 LT) and GOME-2 (09:30 LT) are only 30 minutes apart.
* The residue wavelength pair is completely identical.
* The viewing geometry for collocated orbits is identical. The solar geometry is different by only 30 minutes in time, resulting in a maximum difference between the SCIAMACHY and GOME-2 solar zenith angles of only 7 degrees, leading to residue differences of no more than 0.2 index points.

The errors caused by the overpass time difference of 30 minutes is small compared to the error caused

by the fact that the SCIAMACHY and GOME-2 footprints do not perfectly overlap. This can explain errors up to 0.5 index points for individual measurements. However, many measurements are used to analyse each day of data.

The GOME-2 AAI data set is also compared with OMI-TOMS AAI values. Assuming that the calibration of both instruments is perfect and that the ozone column is known (or the climatological column is close enough to the real column), one has to consider that:

* The difference in overpass time: at 13:45 LT, OMI sees the atmosphere when the convection has been more developed than GOME-2 (at 09:30 LT). Aerosol can be transported to higher levels, leading to an increased residue value for OMI.
* The viewing geometry of the two instruments is different. OMI also has a smaller SZA than GOME-2, due to its overpass time closer to local noon. This means a difference of 0.1–0.2 in AAI.
* The viewing angles of OMI are larger at the extremes of the swaths (also leading to a difference of approximately 0.1–0.2 points in AAI). Finally, the relative azimuth is very different, leading to differences of about 0.1–1.0 index points.
* The wavelength pair used in the retrieval of the two instruments is different: 340/380 nm for GOME-2 and 331/360 nm for OMI-TOMS. As mentioned above, this can lead to a 25% difference in the AAI. As a consequence, a GOME-2 AAI of 2 index points may actually correspond to a OMI-TOMS AAI of 1.5 index points.
* The spatial resolution of the two instruments is different. The pixel size of GOME-2 is about 9 times larger than the OMI pixel size at nadir. Thus the local AAI peaks in GOME-2 will be smaller due to averaging effects.

**Validation process**

Because the AAI is calculated from the difference in reflectance at two UV wavelengths, it is not possible to directly relate it to a single aerosol quantity. The AAI is not a physical quantity but a unitless index. It is therefore hard to do a quantitative verification, or an intercomparison between GOME-2 and other instruments.

For a given day, all SCIAMACHY and GOME-2 AAI orbits that are available are gathered. For each SCIAMACHY orbit the equator passing point (EPP) is determined. A GOME-2 orbit that has a more or less identical EPP is searched for. In positive case the data are matched. The slight difference in equator passing time and solar zenith angle (SZA) is taken into account. For all SCIAMACHY forward scan pixels between 70N and 70S that have a SZA below 80 corresponding GOME-2 forward scan pixels are looked for. Their residues are recorded and the mean is taken if more than one are found. The result we will be the “collocated GOME-2 AAI”.

The analysis described in the previous section was performed on the entire GOME-2 AAI reprocessed data set that was available, which covers most of the years 2007 and 2008. While processing the data set, the number of orbits for which was recorded with successful link between the SCIAMACHY data and GOME-2 data. When a SCIAMACHY monitoring orbit was encountered (narrow swath, nadir static, et cetera), then this orbit was skipped altogether. When a GOME-2 narrow swath or nadir static orbit was encountered, this orbit was not skipped, but it was recorded that a narrow swath/nadir static orbit was used that day. It was decided to remove days with more than two of such orbits from the analysis. Because of the slightly different orbital periods of SCIAMACHY and GOME-2, in general the orbit tracks of the two instruments do not overlap. Every nine days, however, the situation occurs that they do overlap, for a relatively short period of 2 days. For each day for which we could find SCIAMACHY and GOME-2 orbits with overlapping orbit tracks, we recorded the number of these orbits. Slope and intercept of the linear fit to the collocated AAI’s, performed in the way described before, were also recorded.

*I. Spatial consistency analysis*

The SCIAMACHY and GOME-2 footprints do not overlap completely, so there will always be a spatial collocation mismatch between the SCIAMACHY and GOME-2 footprints.

*2. Global statistical analysis:*

GOME-2 AAI was compared to the SCIAMACHY AAI data globally over whole time period. Scatter plots of the daily retrievals against the reference data were made and residues of linear regression to this data were determined. Seasonal variation was analysed as well. Clouds and aerosol loading was paid attention to.

*3. Temporal consistency analysis*

Analysing the global mean residue is a simple and robust validation technique for the AAI. The daily global mean residue, is defined as the average of all healthy residue measurements on a day located between 60N and 60S and having solar zenith angles below 85 degrees.

**Error Characterization**

The error sources are given in the table below:

Error Source Error Range Remarks

Radiometric Calibration 2 pt PPF 3.9.0 to 4.0.0, eastern side

Wavelength Pair 25 %

Solar Zenith Angle 0.1 - 0.2 pt difference between east and west

Viewing Zenith Angle 0.1 - 1.0 pt crosstrack variation

Ozone Column <= 1.0 pt for differences up to 100 DU

In the Product Requirements Document (PRD) [Hovila et al.,2009], the accuracy requirements for the AAI product are given as follows. The “threshold level” is set to 1.0 index points, the “target level” is set to 0.5 index points, and the “optimal level” is set to 0.2 index points.

Quality flags

Quality flags are indicators for the correctness of both the input and the retrieved AAI. There are two quality flag groups: QualityInput and QualityProcessing. In the input flags the types of failures are set: missing data (geometry-wise or spectral-wise), or out of range/invalid values. In the input quality flags there is a flag for sun glint. The sun glint causes a spurious AAI signal. This sun glint flag in the QualityInput comes directly from the Level-1b, and is on a scan level basis (i.e.: valid for the whole scan instead for the subpixel.

## 2.2. LSA-SAF Snow Cover

**Reference documents**

Product Webpage: <http://landsaf.meteo.pt/algorithms.jsp;jsessionid=F05A2D5B0715F079381912AFCF3E07BA?seltab=4&starttab=4>

User Manual: <http://landsaf.meteo.pt/GetDocument.do?id=280>

ATBD: <http://landsaf.meteo.pt/GetDocument.do?id=279>

Validation Report: <http://landsaf.meteo.pt/GetDocument.do?id=289>

Publications:

Siljamo, Niilo, Otto Hyvärinen, 2011: New Geostationary Satellite–Based Snow-Cover Algorithm. *J. Appl. Meteor. Climatol.*, **50**, 1275–1290.

**General validation strategy**

Validation methods covers visual inspections of products, inter-comparison with reference satellite derived data and comparison with in-situ data but in-situ comparison is not done yet. For the validation, not only the product itself with reference data but also the old version of the product is validated with same method and data and compared to the new version.

**Generation of ECV**

The MSG snow cover (SC) retrieval is based on multispectral threshold technique applied to each pixel of the image. In the algorithm MSG/SEVIRI radiance and brightness temperatures of several channels are used together with land surface temperature (LST) and solar and satellite angles to classify each pixel of the land areas. Production is made daily.

**Validation data sets**

(1) NOAA/NESDIS IMS Data:

The NOAA/NESDIS IMS analyses are available as gridded data in American Standard Code for Information Interchange (ASCII) format on the Internet. The higher resolution version of 4 km (614436144 grid) was used in this study. IMS products are disseminated in a polar stereographic projection; they were reprojected to same projection as the LSA SAF SC. This projection is not an area-preserving projection, and pixels correspond to areas of different sizes. However, as it is the projection in which the LSASAF SC is disseminated, it was the natural projection for the comparison.

(2) MODIS Images

MODIS RGB images are used for a few cases for the visual inspection.

(3) Data from the old version

Old version of the LSA-SAF product is also used in the validation for comparison of the versions.

**Validation process**

(1) Visual inspection of products:

For subjective evaluation, false-color RGB combinations provide a useful tool. The image is constructed from three grayscale images of satellite channels, each with different characteristics, and the colors can be given a physical interpretation. Different false-color RGB combinations can be constructed that emphasize different phenomena.

The RGB composite of SEVIRI images at 1200 UTC26 January 2007 and LSA SAF SC (version 2) and IMS products for the same day are assessed. Different locations on the maps are compared and results are verbally reported.

With MSG images, for a few cases MODIS images are used.

(2) Inter-comparison with reference satellite data:

Because of clouds or inadequate solar illumination, not all pixels can be classified by LSA SAF SC, and the number of pixels classified varied from day to day. No attempt was made to mitigate this; LSA SAF SC was compared to IMS only for pixels classified by LSA SAF SC. The data of this study can be thought of as a three dimensional grid consisting of two-dimensional maps and time. Each data point has one of the four possible values are listed. First, the data points were merged for all dimensions, resulting in one contingency table. Second, the values on each map were merged, resulting in one contingency table for each map; the results can be shown as time series.

Third, the values in the time dimension were merged, producing a map where each pixel has a contingency table of its own. This makes it possible to assess the spatial performance of the algorithm. It is reasonable to suppose that snow behaves differently over different terrain types. The effect of land cover was investigated subjectively first by examining the maps showing the distribution of measures and then objectively, in quantitative fashion. Using a verification measure (e.g., PC or HSS), it was calculated whether certain areas (e.g., needle-leaved evergreen forest) were more probable to have values higher or lower than the median value. The median, instead of the mean, was used, because the distribution of values is far from Gaussian.

Results are merged in one contingency table and statistical scores are calculated. Proportion correct, Heidke skill score, critical success index, false alarm rate values, probability of detection and bias are used.

Additionally, spatial variation of the results is considered and shown on the maps for European area.

**Error characterization**

Additional pixel by pixel quality information is supplied in metadata. (e.g. quality flags)

## 2.3. H-SAF H10 (SN-OBS-1) Snow Mask

**Reference documents**

Product Webpage: <http://hsaf.meteoam.it/description-sn-obs-1.php>

User Manual: <http://hsaf.meteoam.it/documents/PUM/SAF_HSAF_PUM-10_1_1.pdf>

ATBD: <http://hsaf.meteoam.it/documents/ATDD/SAF_HSAF_ATBD-10_1_1.pdf>

Validation Report: [http](file:///D%3A%5CGDrive%5CCORE_CLIMAX%5Cwp3%5Chttp)[://hsaf.meteoam.it/documents/PVR/SAF\_HSAF\_PVR-10\_1\_1.pdf](http://hsaf.meteoam.it/documents/PVR/SAF_HSAF_PVR-10_1_1.pdf)

Publications and other:

Surer, S. and Akyurek, Z., 2012. Evaluating the utility of the EUMETSAT HSAF snow recognition product over mountainous areas of eastern Turkey. *Hydrological Sciences Journal*, 57 (8), 1–11.

**General validation strategy**

For H-SAF products, a common methodology is used. The common validation methodology is based on ground data comparison to produce large statistic (multi‐categorical), and case study analysis. Both components (large statistic and case study analysis) are considered complementary in assessing the accuracy of the implemented algorithms. The main steps of the validation procedure are:

1. Observation data containing snow cover measurements have to be gathered.

2. Satellite product needs to be acquired.

3. Both observation and satellite data series need to be checked for consistency.

4. Comparison between the observation data and the product has to be performed.

5. Results of the comparison need to be presented.

**Generation of ECV**

Product SN‐OBS‐1 (Snow detection (snow mask) by VIS/IR radiometry) is based on multi‐channel analysis of the SEVIRI instrument onboard Meteosat satellites.

**Validation data sets**

(1) Ground data

For the validation of the H10 the data come from eight countries (Belgium, Bulgaria, Finland, Germany, Italy, Poland, Slovakia, and Turkey) with different climatology and orography have been used. The ground data inventory shows that synoptic, automatic, handmade snow measurements are used in the SPVG.

The validation is based on measurements at ground stations (SYNOP and other lower level posts) made on a daily basis at 0600 UTC. Metadata concerning the method and instrument used for snow measurement as well as accuracy and frequency are included. From the data collected by ground network, a subset containing snow cover depth (SD) for the reference season (1.10.2009 – 31.09.2010) is extracted and a local database is created. The data is stored in plain text. Each file contains the data from all reporting stations for one day of the reference season.

Detailed information of ground data for each country can be found in the validation report.

(2) Mountain mask

The mountain mask that has been developed by METU to be used in HSAF Project depends on two main features of topography. One is the slope and the other one is the elevation of the terrain. For the mountain mask development GTOPO Digital Elevation Model (DEM) which has 1 km of spatial distribution has been used.

(3) MODIS Data

MOD10A2 Product quicklook images are used visual inspections/comparisons for case studies.

(4) MSG RGB Images

RGB Images from Meteosat-9 satellite are used for visual inspections in case studies.

(5) NWC SAF Cloud Type

Data is used for the case study in Poland.

**Validation process**

(1) Gathering observation data

Observation data is in plain text format as described in the previous section. As the first step assigning the columns of the files is explained in the validation report.

(2) Obtaining satellite data (product)

How to get the product data and required data conversions are explained in the validation report.

(3) Data consistency check

To guarantee high quality of the validation it is advised to check if both the observation data and the satellite product are available for all days of the reference season. It is recommended to make a 'sanity check' both on the satellite product downloaded from the ftp server and the observational data – a quick look on the filename format and file modification dates can prevent making validation on wrong (e.g. old version of the product) or incomplete datasets (e.g. missing observation data).

(4) Comparison between the observation data and the product (with error matrix)

To compare the satellite product with observation data, the measurement from the station that is the nearest to the satellite pixel is used. From satellite product, only pixels with code 0 (snow) and 85 (ground) are taken into consideration. Cloudy and data‐missing pixels are discarded from comparison.

It has been stipulated to treat the measurement as snow occurrence if the snow depth parameter value is equal or greater than 2 cm. Error matrices are produced according to this rule. With the matrices, probability of detection, false alarm ratio, probability of false detection, accuracy, critical success index and Heidke skill score values are calculated.

(5) Results

Results are reported for each different country for both orography and also for all the data together in large tables. Not only the statistical values, hits, false alarms misses and negative correctives are also reported.

(6) Case studies

Multiple case studies for different days and locations are used for validation. These studies include visual inspections/comparisons, quantitative comparisons with error matrices with statistical results and also assessment of correlation with other types of data records. (e.g. temperature)

**Error characterization**

Additional quality control information is supplied in a different file with the product. QC Flags shows data retrieval information, i.e. no data from satellite or location is dark.

## 2.4. H-SAF H12 (SN-OBS-3) Effective Snow Cover

**Reference documents**

Product Webpage: <http://hsaf.meteoam.it/description-sn-obs-3.php>

User Manual: <http://hsaf.meteoam.it/documents/PUM/SAF_HSAF_PUM-12_1_0.pdf>

ATBD: <http://hsaf.meteoam.it/documents/ATDD/SAF_HSAF_ATBD-12_1_0.pdf>

Validation Report: <http://hsaf.meteoam.it/documents/PVR/SAF_HSAF_PVR-12_1.2.pdf>

**General validation strategy**

For H-SAF products, a common methodology is used. The common validation methodology is based on ground data comparison to produce large statistic (multi‐categorical), and case study analysis. Both components (large statistic and case study analysis) are considered complementary in assessing the accuracy of the implemented algorithms. The main steps of the validation procedure are:

1. Observation data containing e-codes or snow course data with visual estimates of snow covered area with values from 0 to 100 have to have to be gathered.

2. Satellite products need to be acquired.

3. Both observation and satellite data series need to be checked for consistency.

4. Comparison between the observation data and the product has to be performed.

5. Results of the comparison need to be presented.

**Generation of ECV**

Product SN-OBS-3 (*Effective snow cover by VIS/IR radiometry*) is based on multi-channel analysis of the AVHRR instrument onboard NOAA and MetOp satellites.

**Validation data sets**

(1) Ground Data

For the validation of the H12 the data come from two countries (Finland and Turkey) with different climatology and orography have been used. The ground data inventory shows that synoptic, automatic, handmade snow measurements are used in the SPVG.

The validation is based on measurements at ground stations (SYNOP and other lower level posts) made on a daily basis at 0600 UTC. Metadata concerning the method and instrument used for snow measurement as well as accuracy and frequency are included. From the data collected by ground network, a subset containing snow cover depth (SD) for the reference season (1.10.2009 – 31.09.2010) is extracted and a local database is created. The data is stored in plain text. Each file contains the data from all reporting stations for one day of the reference season.

A subset in the collected ground data is e-codes. E-code values are assigned with the visual inspection in the observation area. These e-codes are classified as snow fraction for the validation.

Detailed information of ground data for each country can be found in the validation report.

(2) Mountain mask

The mountain mask that has been developed by METU to be used in HSAF Project depends on two main features of topography. One is the slope and the other one is the elevation of the terrain. For the mountain mask development GTOPO Digital Elevation Model (DEM) which has 1 km of spatial distribution has been used.

(4) MSG SEVIRI RGB Composite Images

Data is used for visual inspections in the case studies in Finland.

(3) METOP/AVHRR RGB Composite Images

Data is used for visual inspections in the case studies in Turkey.

(4) MODIS MOD10A1 Data

Data is used for comparison with the product in a quantitative way in case studies in Turkey.

**Validation process**

(1) Gathering observation data

Observation data is in plain text format as described in the previous section. As the first step assigning the columns of the files is explained in the validation report.

(2) Obtaining satellite data (product)

How to get the product data and required data conversions are explained in the validation report.

(3) Data consistency check

To guarantee high quality of the validation it is advised to check if both the observation data and the satellite product are available for all days of the reference season. It is recommended to make a 'sanity check' both on the satellite product downloaded from the ftp server and the observational data – a quick look on the filename format and file modification dates can prevent making validation on wrong (e.g. old version of the product) or incomplete datasets (e.g. missing observation data).

(4) Comparison between the observation data and the product (with error matrix)

The nearest e-code extracted for the pixel is used in the validation. The snow cover fraction (FSCA) and the e-code values are classified into four classes as explained in the report. Then a multi-category contingency table with four classes is used for statistical calculations.

(5) Results

Results are reported for each country separately and together. Statistical values reported are RMSE and accuracy (user’s accuracy, producer’s accuracy and overall accuracy). Scatter plots are also presented.

(6) Case studies

Multiple case studies for different days and locations are used for validation. These studies include visual inspections/comparisons, quantitative comparisons error matrices with statistical results.

**Error characterization**

Additional quality control information is supplied in a different file with the product. QC dataset gives the number of observations of the surface, i.e. non-cloudy day and during daytime.

# 3. ESA CCI Projects

## 3.1. ACCI (ESA)

**Reference documents**

<http://www.esa-aerosol-cci.org/?q=webfm_send/226>

<http://www.esa-aerosol-cci.org/?q=webfm_send/477>

**Generation of ECV**

The algorithm AATSR Dual View (ADV) uses both the forward and the nadir view provided by AATSR to eliminate the effects of land surface reflectance on the radiation received at the top of the atmosphere (TOA). Over ocean the surface reflectance is modelled and both views are used independently to retrieval aerosol properties. The aerosol properties retrieved are the aerosol optical depths at the wavelengths in the visible to near-infra-read, and the Ångström exponent describing the wavelengths dependence of the AOD. This is achieved by fitting the reflectance spectra at TOA computed using a forward model, the radiative transfer model DAK (Double Adding KNMI), to the measured those derived directly from the AATSR radiance measurments. The forward model uses a number of aerosol models for which the optical properties are computed using a Mie code. Two aerosol components, usually a fine and a coarse mode component, are mixed and the mixing ratio is varied to obtain the optimum wavelengths dependence using a least-square routine. The mixing ratio of these two, pre-selected, aerosol components is provided as an output as well.

**General validation strategy**

The validation activities are two-fold: 1) evaluations against quality-controlled reference data with known accuracy, as provided by ground-based sun-photometry, and 2) inter-comparison to existing and established products from satellite remote sensing. Aerosol\_CCI data products were evaluated by four partners not involved in the EO data production. The diversity in evaluation methods is complementary and redundancy in evaluation of the same product with different methods will help to avoid (method-related) evaluation biases.

The analysis of aerosol\_CCI datasets focused on validating aerosol optical depth (AOD) – primarily at the reference wavelength 550 nm - and to a lesser extent derived parameters such as Ångström coefficient (ANG), or fine and coarse mode AOD. The assessment of pixel level uncertainty provided as part of the products has reached a first assessment level. Inter-comparison to other satellite remote sensing datasets as reference was mainly done for comparing to state of the art. .

**Validation data sets**

LOA maintains a part of AERONET. It used its local ICARE remote sensing data-base (http://www.icare.univ-lille1.fr/archive/), where many satellite remote sensing products (e.g. POLDER, MODIS) are stored. ICARE also hosted the storage facility for downloading Aerosol\_cci retrieval results for easy access by the validation partners. LOA evaluations built on existing and new LOA/ICARE tools, which were developed for evaluation of the POLDER, AATSR, MERIS sensor data. LOA focused on validation of level2 products also taking quality flags into account.

Met-No hosts the AeroCom data-base and web site (http://aerocom.met.no/data.html). Here data are accessible from multiple model simulations, including simulations for the reference year 2008. In addition, a wide variety of reference data exists, including data by sun-/sky-photometry and data from satellite instruments such as MODIS and MISR. Existing evaluation tools include visual images and standard statistics (bias, correlation and RMS). MetNo focused on level3 (daily gridded) product evaluation versus daily mean AERONET AOD. AeroCom tools were used to produce monthly mean maps for visual inspection of coverage and features. Further refinements of the AeroCom tools were developed to allow for consistent across retrieval comparisons having different coverage and filtering of stations.

MPI-M has been involved by supplying reference data into the AeroCom data-base. Most importantly AERONET sun-/sky-photometer data were prepared and are regularly updated. In addition, a new quantitative scoring method has been developed to serve as diagnostic tool in comparative analyses to

identify (on a regional and seasonal basis) performance strength and limitations and monitor overall progress over time. This scoring was applied to quantitatively assess temporal and spatial correlations with reference AERONET data.

NILU hosts the GAW data and its World Data Centre for Aerosols (http://ebas.nilu.no/). NILU focused on level2 validation and in particular assessment of AOD uncertainties in the data products. Furthermore, NILU was responsible for the validation of the stratospheric products from GOMOS using lidar (CALIOP and NDACC ground-based) datasets.

**Validation process**

*Level2 validation*

The analysis compared level2 datasets (AOD550 and Angstrom coefficient) for each month to AERONET measurements. Comparisons were performed separately over land and ocean with AERONET sites and satellite pixels within ± 30 min and ± 20 km (LOA/ICARE) or ± 50 km (NILU). Where available, quality flags in the products were taken into account to select best pixels. The validation considered AOD / Ångström coefficient scatter plots and difference histograms. Statistical measures evaluated were Pearson correlation coefficient, linear fit parameters, standard deviations (from linear fit and from AOD/ Ångström coefficient difference histogram), average difference, and numbers of AERONET sites and satellite pixels used. Criteria used for ranking were correlation coefficient, standard deviation and number of satellite pixels.

ed for ranking were correlation coefficient, standard deviation and bias.

*Level3 validation*

This evaluation of daily gridded products works with the nearest satellite pixel value on a 1x1 degree grid corresponding to daily mean AERONET values excluding mountain sites. The evaluation with the AEROCOM tools provides bias maps, histograms, scatter plots, time series graphs, zonal mean comparisons, and score tables. Any user can produce the same analysis for all retrievals at the MetNo AeroCom web interface. This analysis included all pixels regardless of quality flags. The ORAC land / sea mask was used for all retrievals to differentiate land, coast and sea cases.

*Comparison to other datasets*

The monthly mean maps of three AATSR Aerosol\_CCI algorithms (ADV, ORAC, SU) were compared to reference datasets (MISR, MODIS Terra) to derive total regional error.

*Scoring of spatial and temporal correlations (MPI)*

This approach relies on matches of satellite 1x1 degree gridded data with AERONET data measurements within a 30 min window around 10:00 local time (ENVISAT equator crossing) and within 50 km to the AERONET station. The temporal correlation scoring requires at least 10 data points at each location over the 4 months of data provided – this excludes retrievals like SYNAER with lower coverage from the analysis. The scoring aggregates regional spatial scores and regional temporal scores to total regional scores and then to total global scores. Sufficient numbers of AERONET sites are only available for Europe, North America and Asia. This analysis aims to tackle the combination of ‘sufficient’ coverage and of ‘sufficient’ accuracy.

*Validation of stratospheric extinction profiles (NILU)*

A preliminary evaluation of the gridded L3 stratospheric aerosol AOD (column) and extinction at 500 nm derived from the GOMOS AERGOM product version v2.1 was conducted. As reference CALIPSO Level 2 profile data (5-km product, CAL\_LID\_L2\_05kmAPro-Prov-V3-01) were used. Clouds are generally removed from the CALIPSO aerosol data. Stratospheric AOD values over 0.2 are excluded, which means a rough cloud filtering of the CALIPSO data. The individual L2 overpass data were averaged to the 2.5° x 10° latitude- longitude grid.

*Use of synthetic case studies for validation (UBremen)*

As additional means of validation the use of synthetic case studies was tested, whereby blind tests of simulated retrieval cases were made with the participating algorithms. However, it turned out, that the small number of test cases cannot provide comprehensive assessments of the retrieval algorithms in a wide range of conditions (atmosphere, surface, aerosol).

*Evaluation of first 10 year time series from AATSR (MetNo)*

A first time series from AATSR (2002 – 2012) was produced and evaluated within an option to the project. It was compared to AERONET and MODIS data.

**Error Characterization**

An analysis was made to assess the uncertainties contained as pixel-level auxiliary data in the product files (NILU).

## 3.2. GlobSnow Snow Extent

**Reference documents**

Product Webpage: <http://www.globsnow.info/index.php?page=Snow_Extent>

User Manual: [http](file:///D%3A%5CGDrive%5CCORE_CLIMAX%5Cwp3%5Chttp)[://www.globsnow.info/se/GlobSnow\_SE\_product\_readme\_v1.2.pdf](http://www.globsnow.info/se/GlobSnow_SE_product_readme_v1.2.pdf)

Snow Extent Algorithm Development Document 2: (Internal validation report) <http://www.globsnow.info/swe/Aux_data/Deliverables_GS2/GS2_DEL_11_WP2.2_v1_r2_final.pdf>

Publications:

Sari Metsämäki, Olli-Pekka Mattila, Jouni Pulliainen, Kirsikka Niemi, Kari Luojus, Kristin Böttcher, An optical reflectance model-based method for fractional snow cover mapping applicable to continental scale, Remote Sensing of Environment, Volume 123, August 2012, Pages 508-521

**General validation strategy**

The work on Fractional Snow Cover (FSC) evaluation of the Full Production Set (FPS) Version 1.2 has focussed on:

1. Inter‐comparison of GlobSnow SE and MODIS Fractional Snow Cover products for the northern hemisphere in order to identify systematic differences between the two products

2. Inter‐comparison of high‐resolution optical sensors (Landsat) for different environments and climate zones in order to study the capabilities and limitations of the GlobSnow algorithms

3. Visual inspection to determine anomalies (errors or deviations from expected results).

Additional validation work is reported in journal papers. This validation covers both in-situ and satellite derived inter-comparison and comparison of the validation for the product and the reference satellite derived product.

**Generation of ECV**

The GlobSnow SE processing system applies optical measurements in the visual‐to -thermal part of the electromagnetic spectrum acquired by the ERS‐2 sensor ATSR‐2 and the Envisat sensor AATSR. Clouds are detected by a cloud‐cover retrieval algorithm (SCDA) and masked out. Large water bodies (oceans, lakes and rivers) and glaciers are also masked out. The snow cover information is retrieved by two algorithms, one for high‐mountain areas of steep topography above the tree line (NLR) and another for forested and open areas (SCAmod). The domains of the algorithms are determined by the thematic masks, and the retrieval results are merged.

**Validation data sets**

(1) Snow course measurements:

Snow course data is retrieved from the network of SYKE. It consist of ~150 courses which are monthly visited. These courses have 40-80 locations to observe and each location is observed for its area of 25m radius to calculate the snow fraction. Snow depth is also measured. Land covers types are also considered and assigned to each location.

(2) Weather station measurements:

Weather station data is retrieved from the network of FMI; consisting of ~250 stations. Snow depth and coverage measurements are made on daily basis. Snow coverage here has the e-code description of WMO.

(3) Landsat-TM/ETM+ Satellite data (snow fraction):

A set of 13 Landsat-ETM+ scenes from years 2003 in different parts of Europe was selected for comparisons. These represent different landscapes (tundra, steppe, boreal forest) and different stages of snow melt (FSC ranges from nearly zero to 100%). From all four DDS years, only 2003 was chosen due the sensor deterioration after that year. The narrow swath width of the AATSR limited the number of scenes to only 13 (tens of otherwise feasible TM-scenes were outside the AATSR acquisition area).

(4) ENVEO Satellite Data:

For the validation activities various dates in 2003 were selected, where Landsat 7 ETM+ images were coincidently acquired with AATSR scenes. For some of these dates also AVHRR data are available. For these dates, snow products derived by different algorithms and from different sensors are compared.

**Validation process**

(1) Qualitative comparison:

A visual comparison by experts is a very valuable tool to gain first impression of the overall performance of the snow, but also of the cloud masking algorithms. Even though this step involves a certain amount of subjectivity, obvious systematic errors can already be detected and attributed at this stage. Therefore, every team was asked to provide a first visual check of the satellite product and reference data before initiating automated validation procedures.

(2) Binary metrics: (Error matrix)

Measures based on contingency table statistics are used for the accuracy assessment of binary snow products, generated as described in Section 2.6: POD (probability of detection), FAR (false alarm rate), HR (hit rate), KSS (Kuiper’s skill score).

For the validation of the AATSR SCAmod and AVHRR SCAmod datasets based on model snow data, binary information of snow cover had to be derived from all data. Therefor the remotely sensed daily fractional snow cover (DFSC in %) had to be transformed to a snow map containing snow or no-snow information. A threshold of ≥ 50% snow cover was used to define a FSC pixel as fully covered by snow. The validation was carried out for the years 2003, 2004 and 2006.

For the reference data, the model snow data, three different thresholds have been set to convert snow depth in cm to binary snow cover. This validation was carried out separately for ≥ 4 cm, 15 cm and 40 cm thresholds do define a pixel as 100 % snow covered.

Additionally a land cover mask (forest class and all other classes) was applied to the data in order to compare the results with the validation for Finnish forest (SYKE). We used ESA GlobCover 2009 masks to calculate the validation for forested areas and all other classes. On a daily basis the re-classified snow covered SCAmod pixel was compared with the binary snow information from the spatially and temporarily corresponding pixel of the model data (reference data). In a first step, no mask was applied. In the second and third steps, the tasks for forested and non-forested were included. Contingency matrices for each day with coinciding data elements were designed to calculate a variety of statistical indices and scores. Eventually the daily indices and scores were averaged for the three years individually and for the entire period.

(3) Fractional metrics:

Comparison of estimated FSC and the reference FSC can be conducted using the original continuous values or, alternatively, the values can be categorized into certain classes and then these classes are compared. With the former approach, the differences between estimated and independent reference FSC (N cases) are analyzed using root-mean-squared-error (RMSE).

(4) Intersatellite comparison:

The comparison between the two satellite SE products (AATSR SCAmod and SPARC AVHRR) is carried out in two steps. In general, the comparison is limited to the European Alps and only pixels classified as „clear-sky“ in both products were used as different cloud masks are applied. Fractional Snow cover area percentage (FSC) [0-100%]) is defined as the percentage area of a certain reference area (i.e., the Alpine region) covered by snow at a specific point in time. Following analyses are done.

Regression analysis of the respective snow fractions detected for the same day during the winter season and ablation phase: The results are expected to provide indication of the general agreement and potential systematic biases of the two products as well as information on the dependence of agreement on the actual SCA. As measures for comparison, R2 coefficient of determination, RMSE and bias values are used.

Analyses in amount of SCA differences (SCAmod AATSR – SPARC AVHRR) over the course of one year: This analysis is supposed to reveal the time/phase when the highest differences occur between the SCA detection of the two products.

Results are displayed separately for the major land cover classes as the two algorithm performances are expected to vary across different land surface types. Furthermore, the analysis is carried out for all of the four validation years (2003, 2004, 2006 and 2010).

**Error characterization**

Additional quality flags are used in metadata. Bit flags in thematic layer include information like low solar elevation, critical forest density, sensing algorithm etc. Explanation of these is seen in Table x:



Table x: Explanation of bit flags

## 3.3. Methane

Methane products are produced in the ESA Climate Change Initiative project on Greenhouse Gases (GHG-CCI).

**Reference documents**

GHG-CCI documents available from http://www.esa-ghg-cci.org/?q=node/95

Algorithm selection report (ASR) of GHG-CCI, final version.

Product validation and Intercomparison report (PVIR) of GHG-CCI, version 2.0.

B. Dils, M. Buchwitz, M. Reuter, O. Schneising, H. Boesch, R. Parker, S. Guerlet, I. Aben, T. Blumenstock, J. P. Burrows, A. Butz, N. M. Deutscher, C. Frankenberg, F. Hase, O. P. Hasekamp, J. Heymann, M. De Mazière, J. Notholt, R. Sussmann, T. Warneke, D. Griffith, V. Sherlock, D. Wunch :The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSOFTS/ GOSAT CO2 and CH4 retrieval algorithm products with measurements from the TCCON network, submitted to AMTD, 2013

**Generation of ECV**

In the GHG-CCI, methane products are produced using a set of varying algorithms selected after a round robin evaluation of candidates. Algorithms are separated by instrument (SCIAMACHY or GOSAT) and sensitivity to near-surface methane mixing ratio, i.e. the capacity to identify methane sources and sinks. For details on the algorithms, see the algorithm selection report (ASR) of GHG-CCI.

**General validation strategy**

The methane products in GHG-CCI have primarily been validated through temporally and spatially collocated comparisons against reference in situ observations from the Total Carbon Column Observing Network (TCCON). The general validation methodology in the latest available validation and intercomparison report (PVIR) is stated to closely follow Dils et al. (2013).

**Validation data sets**

The primary validation dataset for methane products in GHG-CCI is the Total Carbon Column Observing Network (TCCON). 13 TCCON sites provide reference data for the validation, the longest reference data record being ~8.5 years.

**Validation process**

The satellite retrievals and in situ observations are collocated spatially and temporally, with matching boundaries of 500 km and +/- 2 hours, respectively. The satellite retrieval data are corrected before the comparison for differing a-priori profiles, although the effect of this correction is stated to be minor in the PVIR of GHG-CCI. In situ observations within the +/- 2 hour satellite retrieval window are averaged to obtain a unique satellite-in situ data pair.

The validation metrics used are:

* Bias (average satellite-in situ difference)
* Scatter (standard deviation of said difference)
* Relative Accuracy (standard deviation of overall biases at each station)
* Seasonal Relative Accuracy (as relative accuracy, but restricted to seasonal periods JFM, AMJ, JAS, OND)

Validation results derived from less than 10 data points or having a standard error exceeding the threshold Relative Accuracy requirement of 10 ppb are excluded from all compared algorithms’ results.

**Error Characterization**

Error characterization of the GHG-CCI methane products varies according to algorithm. The IMAP-DOAS algorithm Product User Guide (PUG) describes a custom data filter that all users are advised to use instead of the quality flag included in the product data files. SCIAMACHY detector degradation since 2005 for some detector pixels is noted and data is filtered to remove poor-quality retrievals.

Data files produced by the WFM-DOAS algorithm contain a quality flag whose composition is described in the algorithm ATBD. Quality flagging conditions were changed in all products since 2005 to better account for the SCIAMACHY detector degradation.

Both IMAP-DOAS and WFM-DOAS data products also contain a retrieval uncertainty for each data pixel. The retrieval uncertainty is derived from either the least-squares fitting in WFM-DOAS, or the (optimal estimation) forward model calculation in IMAP-DOAS.

The algorithms used for GOSAT-based methane retrievals (OCPR and SRFP) also function on similar basis, providing pixel-specific quality flags.

## 3.4. Ozone

**Reference documents**

Ozone\_CCI documents available at http://www.esa-ozone-cci.org/?q=documents

Ozone\_CCI Product Specification Document (PSD), ref: Ozone\_cci\_ PSD\_3.1, version 3.0

Ozone\_CCI Product Validation and Selection Report (PVASR), ref: Ozone\_cci\_PVASR, issue 1.5

Brühl, C., S. R. Drayson, J. M. Russell III, P. J. Crutzen, J. McInerney, P. N. Purcell, H. Claude, H. Gernand, T. McGee, I. McDermid, and M. R. Gunson: HALOE Ozone Channel Validation, J. Geophys. Res., Vol. 101, No. D6, pp. 10,217-10,240, April 1996.

Nazaryan, H., M.P. McCormick, and J.M. Russell III: New studies of SAGE II and HALOE ozone profile and long-term change comparisons, J.Geophys. Res., 110 (D9), doi: 10.1029/2004JD005425, 2005.

Wang, H. J., D.M. Cunnold, L.W. Thomason, J.M. Zawodny, and G.E. Bodeker, "Assessment of SAGE version 6.1 ozone data quality", J.Geophys. Res., 107, doi:10.1029/2002JD002418, 2002.

**Generation of ECV**

Ozone\_CCI products are organized along three main product lines: Total ozone from nadir UV backscatter sensors, ozone profiles from nadir UV backscatter sensors, and ozone profiles from limb and occultation sensors. The sensors used are GOME/ERS-2, GOME-2/METOP, OMI/AURA, SCIAMACHY/ENVISAT, MIPAS/ENVISAT, GOMOS/ENVISAT, and OSIRIS. The aim of the Ozone\_CCI is to create common state-of-the-art algorithms for the various sensors for the retrieval of total ozone and ozone profiles. The baseline algorithm for total ozone calculation is the GOME-type direct-fitting retrieval scheme (GODFIT). This algorithm is based on a least-squares fitting including direct multi-spectral radiative transfer simulation of radiances and retrieval parameters Jacobians.

For the ozone profile products, two existing scientific algorithms developed at KNMI (OPERA) and at RAL, and applicable to all sensors, will be intercompared and characterized in a Round Robin exercise. Results of this evaluation will support the determination of the new Ozone-cci nadir profile algorithm baseline, which will be a combined RAL/OPERA algorithm. More specific information is available in the Ozone\_cci Product Specification Document (PSD).

**General validation strategy**

The total ozone products were validated with direct comparisons against total ozone column in situ data mainly from Dobson and Brewer UV spectrophotometers. The nadir ozone profile products are validated against ozone sonde and lidar observations for the part of the atmosphere over 30-35 km height.

For the limb/occultation ozone profiles, more emphasis is placed on intercomparison of satellite products, in particular against HALOE v19 (Brühl et al., 1996, Nazaryan et al., 2005) and SAGE II v6.2 (Wang et al., 2002, Nazaryan et al., 2005), although in situ validation against sonde and lidar observations is included.

**Validation data sets**

The total ozone column in situ data was extracted from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) archives. 32 Brewer and 47 Dobson stations were considered, although not all stations were used as reference against every data product made during the Round Robin exercise. In situ data is limited up to 80° SZA for Brewers MK-III and MK-IV and 70-75° of SZA for Dobsons and other Brewers. The in situ data is required to feature “a bi-weekly sampling of the time series over at least five years” (Ozone-CCI PVASR), meaning that sites with long data gaps will likely not be considered as reference.

For the ozone profile products, validation data sets are collected from the Data Host Facility (DHF) of the Network for the Detection of Atmospheric Composition Change (NDACC) and from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC).

**Validation process**

Total ozone validation is carried out by comparison of in situ to satellite-retrieved data over individual validation sites as well as using latitudinal zones (tropics, middle latitudes and polar areas). Base level of comparison is the daily mean, although monthly mean accuracies are used in reporting.

**Error Characterization**

Both the total ozone and ozone profile products feature pixel-specific quality flags for the level2. Total ozone quality flags report the processing quality level (retrieval successful / no algorithm convergence reached / bad retrieval values / no retrieval possible). Level3 total ozone product reports a grid cell-specific error associated to the total ozone column (Dobson units).

Nadir ozone level2 products feature detailed quality flags according to Table 2. Associated retrieval error is also reported.

Table 3: Ozone\_CCI nadir ozone profile product quality flags

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset name | Data type | Unit | Description |
| QualityInput | Int arr,rank 2,size 32 | N/A | Quality flags for the input data. 0 = false, 1 = true0: Non-nominal level 1 due to instrument degradation;DEGRADED\_INST\_MDR in Level1b)1: Non-nominal level 1 due to processing degradation;DEGRADED\_PROC\_MDR in Level1b2: Groundpixel is in SAA; F\_SSA in Level1b/PCD\_BASIC3: Sunfile of date missing: older sunfile used4: Meteoforecast file missing: climatological meteo dataused5: Meteoforecast data missing: climatological meteo dataused6: Meteoforecast data invalid7: Earthshine radiance data missing8: Earthshine radiance data invalid9: Solar irradiance data missing10: Solar irradiance data invalid11: Measurement data invalid12: Auxiliary data invalid13: Absorbing Aerosol Index data invalid14: Failure in setup of the Forward Model Input15: Failure in State vector definition setup16: Sunglint flag17: Cloud fraction forced to zero18: Cloud Pressure Adjusted to Surface Pressure19: Other error20 – 31: reserved for future use |
| QualityProcessing | Int arr,rank 2,size 32 | N/A | Quality flags for processing. 0 = false, 1 = true, -999 = NoRetrieval done, -1 = value not initialized / not used.0: Overall Convergence was reached (indicatessuccessful retrieval)1: Convergence reached on Cost2: Convergence reached on State3: Convergence not reached after maximum number ofiterations4: Out of bound retrieval values5: Too high values for Chi Square6: No retrieval done! (due to incorrect inputs or otherreasons).7 – 31: reserved for future use |

## 3.5. CCI Soil Moisture

A round robin exercise was organized to benchmark several retrieval algorithms for both active and passive microwave observations (ASCAT and AMSR-E, respectively). The triple-(multiple-) collocation method was applied to surface soil moisture[[1]](#footnote-1). Initially, this technique was applied to wind products over the ocean[[2]](#footnote-2).

# 4. In-situ Networks Climate Monitoring

The table below gives an overview on selected in-situ station networks and gridded datasets, which are produced, hosted or contributed by European institutions or initiatives, pointing to corresponding documentation on quality control, gridding methodology, validation approach and uncertainties. The GCOS Reference Upper Air Network (GRUAN) can be regarded as a role model for in-situ Cal/Val, as this international network has been designed by researchers with particular focus on calibration and monitoring of long-term stability for climate applications.

|  |
| --- |
| Station databases / in-situ networks |
| [ECA&D](http://eca.knmi.nl/)  (KNMI)ECA dataset contains series of daily observations at meteorological stations throughout Europe.Algorithm Theoretical Basis Document ([ATBD](http://eca.knmi.nl/documents/atbd.pdf)) 🡪 provides information on quality control, homogeneity tests etc. |
| Baseline Surface Radiation Network ([BSRN](http://www.bsrn.awi.de/en/data/data_retrieval_via_pangaea/))[GCOS-174](http://www.wmo.int/pages/prog/gcos/Publications/gcos-174.pdf), 2013, Baseline Surface Radiation Network (BSRN) - Update of the Technical Plan for BSRN Data Management.🡪 report describing quality control etc. |
| GSN (GCOS Surface Network) / GUAN (GCOS Upper Air Network)[GCOS-144](http://www.wmo.int/pages/prog/gcos/Publications/GCOS-144_en.pdf), Guide to the GCOS Surface Network (GSN) and GCOS Upper-Air Network (GUAN) (2010 Update of GCOS-73) |
| GRUAN[GCOS-171](http://www.wmo.int/pages/prog/gcos/Publications/gcos-171.pdf), The GCOS Upper-Air Reference Network (GRUAN) GUIDE[GCOS-170](http://www.wmo.int/pages/prog/gcos/Publications/gcos-170.pdf), The GCOS Upper-Air Reference Network (GRUAN) MANUAL |

|  |
| --- |
| Gridded datasets (derived by station measurements) |
| [E-OBS](http://eca.knmi.nl/download/ensembles/ensembles.php) (KNMI)Precipitation, Temperature ([Haylock et al. 2008](http://eca.knmi.nl/download/ensembles/Haylock_et_al_2008.pdf))Sea level pressure ([van den Besselaar et al. 2011](http://onlinelibrary.wiley.com/doi/10.1029/2010JD015468/abstract)) |
| MetOffice Hadley Centre observation datasets [CRUTEM4](http://www.metoffice.gov.uk/hadobs/crutem4) - Gridded land monthly temperatures ([Jones et al. 2012](http://www.metoffice.gov.uk/hadobs/crutem4/CRUTEM4_accepted.pdf))[HadCRUT4](http://www.metoffice.gov.uk/hadobs/hadcrut4) - Gridded monthly temperatures ([Morice et al. 2012](http://www.metoffice.gov.uk/hadobs/hadcrut4/HadCRUT4_accepted.pdf))[HadSLP2](http://www.metoffice.gov.uk/hadobs/hadslp2) - Monthly gridded sea-level pressures ([Allan & Ansell 2006](http://www.metoffice.gov.uk/hadobs/hadslp2/allan_ansell.pdf)) |
| Climate Research Unit (CRU)[CRU TS3.10](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__ACTIVITY_fe67d66a-5b02-11e0-88c9-00e081470265) gridded time series dataset ([Harris et al. 2013](http://onlinelibrary.wiley.com/doi/10.1002/joc.3711/abstract)) |
| Global Precipitation Climatology Centre (GPCC, DWD)GPCC Full Data Reanalysis Version 6 ([Becker et al. 2013](http://www.earth-syst-sci-data.net/5/71/2013/essd-5-71-2013.html)) |

An overview on communities and projects dealing with issues of standardization, quality control, calibration, validation and homogeneity is provided in the following tables:

|  |
| --- |
| WMO bodies |
| *Commission for Instruments and Methods of Observation (CIMO)*🡪 promotes and facilitates international standardization and compatibility of instruments and methods of observationsOPAG Standardization and Intercomparisons* Expert Team on Standardization
* Expert Team on New In-Situ Technologies
* Expert Team on Instrument Intercomparisons
 |
| *Commission for Climatology (CCl)*OPACE 1 Climate Data Management * Expert Team on Climate Data Base Management Systems
* Expert Team on Data Rescue
* Task Team on Observational Standards and Practices

Other focus groups:Expert Group on Quality Management for Climatology |
| *Commission for Basic Systems (CBS)*OPAG on Integrated Observing Systems (OPAG-IOS)Expert Team on Surface-Based Observing Systems (ET-SBO)OPAG on Information Systems and Services (OPAG-ISS)Inter-Programme Expert Team on Metadata and Data Representation Development (IPET-MDRD) |
| *Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology (JCOMM)*DBCP (Data Buoy Cooperation Panel) |

|  |
| --- |
| COST Actions |
| [S0601](http://www.cost.eu/domains_actions/essem/Actions/ES0601): “Advances in homogenisation methods of climate series: an integrated approach HOME” (May 2007-November 2011)Website: <http://www.homogenisation.org/v_02_15/>Special Issue of [Időjárás 2013, 117, 1-158.](http://www.met.hu/en/omsz/kiadvanyok/idojaras/index.php?id=82)Venema et al. 2012: Benchmarking homogenization algorithms for monthly data, Clim. Past, 8, 89-115, doi:[10.5194/cp-8-89-2012](http://www.clim-past.net/8/89/2012/cp-8-89-2012.html) |
| [719](http://www.cost.eu/domains_actions/essem/Actions/719): “The use of geographic information systems in Climatology and Meteorology” (February 2001-July 2006)*Publications:*OE. Tveito, M. Wegehenkel, F. van der Wel, H. Dobesch (2008): [The Use of Geographic Information Systems in Climatology and Meteorology](http://www.cost.eu/media/publications/07-37-Spatial-Interpolation-for-Climate-Data)H. Dobesch, P. Dumolard, I. Dyras (2007): [Spatial Interpolation for Climate Data](http://www.cost.eu/media/publications/07-37-Spatial-Interpolation-for-Climate-Data)  |

|  |
| --- |
| Projects/organizations on metadata/traceability of measurements |
| [CHARMe](http://charme.org.uk/) (Characterisation of Metadata to enable high-quality climate applications and services) |
| [MeteoMet](http://www.meteomet.org/index.php?option=com_content&task=view&id=83&Itemid=77) (Metrology for pressure, temperature, humidity and airspeed in the Atmosphere) |
| [BIPM](http://www.bipm.org/) (Bureau International des Poids et Mesures)🡪 Guide to the Expression of Uncertainty in Measurement ([GUM](http://www.bipm.org/en/publications/guides/gum.html))🡪 WMO-BIPM Workshop (2010) on Measurement Challenges for Global Observation Systems for Climate Change Monitoring: Traceability, Stability and Uncertainty [[Report](http://www.bipm.org/utils/common/pdf/rapportBIPM/2010/08.pdf)] |

1. Zwieback, S., K. Scipal, W. Dorigo, and W. Wagner: Structural and statistical properties of the collocation technique for error characterization. Nonlin. Processes Geophys., 19, 69–80, 2012. Available on: www.nonlin-processes-geophys.net/19/69/2012/ [↑](#footnote-ref-1)
2. Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, J. Geophys. Res., 103, 7755–7766, doi:10.1029/97jc03180, 1998 [↑](#footnote-ref-2)