# 8. Data Processing and Products

The processing of SCIAMACHY measurement data has to serve the requirements of different user communities. Therefore, products are generated and provided at various levels, both from a temporal and content point of view. In addition, product generation may occur in two environments. Whereas the operational data processing follows the specific guidelines and rules of the ENVISAT PDS architecture and organisation, scientific product generation is based on science group algorithms which have not been developed within the framework of the PDS. This twofold approach, when fully implemented, has the advantage that standard products up to level 2, with well-defined requirements, can be generated in a configuration controlled but configurable processor regime. It relieves users from providing significant resources for basic processing steps, e.g. calibrating the measurement data. On the other hand, scientific products can be based on most recent, often advanced algorithms since the stringent requirements of the operational ground segment need not to be fulfilled. Also new geophysical parameters which are not part of the operational chains will first be introduced as scientific products demonstrating the potential of SCIAMACHY's measurement data. Later these algorithms may be translated into the operational environment thus providing continuous global data sets of additional geophysical parameters. It is this feedback between scientific and operational products that ensures gaining an optimum knowledge level in a mission lasting for many years.

The baseline for each higher level product is the level 1b product containing orbitwise geo-referenced measurements, including their calibration and instrument monitoring data. Level 2 products are specifically dedicated to geophysical parameters including column densities and stratospheric profiles of atmospheric constituents as well as information about clouds and aerosols. Since the retrieval uses the corresponding level 1b product, also level 2 products are orbit driven. Each day of SCIAMACHY operations generates 14 or 15 level 1b and level 2 products. With level 2, ESA's responsibility for data generation, dissemination and archiving ends. This does not exclude however, as outlined above, experienced science users from processing their own scientific products of the same levels. For levels higher than 2, known as *Value Added* products, scientific applications begin to play a role. It is up to the individual user how to specify and process VA products. In the VA environment, geophysical parameters are often gridded on a global scale, i.e. these products are well suited to provide the interested public with SCIAMACHY's view of the Earth's atmosphere.

## 8.1 Operational Processing Chains

Operational processing of SCIAMACHY measurement data occurs in the Payload Data Segment as part of the ENVISAT mission. The objective of the PDS is twofold. It comprises the processing of measurements of all ENVISAT sensors to dedicated operational data products and the archiving and dissemination of operational data products. The operational data processing covers:

- facilities to serve the need of generation, archiving, and dissemination of data products including the interface to the users, e.g. PACs and the archiving and dissemination process,
- software to generate the data products, e.g. data processors and calibration facilities which run within the processing centres,
- data products on level 0, level 1b, and level 2,
- monitoring and quality control of the operational production chain.

Fig. 8-1 sketches the two main processing chains, near-realtime and offline, together with the required flow of auxiliary data. Each level 1b product is processed from a level 0 product which contains all measurement data for one complete orbit in the case of consolidated products or the measurement data between two consecutive data dumps when dealing with the NRT process. Additional information stems from auxiliary files such as on-ground calibration data and initialisation data for processing control. The level 0-1b chain is completed by an extra loop for the generation of most recent valid calibration data within the Instrument Engineering and Calibration Facility (IECF) and the associated software SciCal. The IECF contains, besides algorithms for calibration, a database comprising all generated calibration data stored in Annotation Data Sets (ADS). From this ADS database the Global Annotation Data Sets (GADS) are formed within the IECF as input for level 0-1b processing. They are based on the most recent calibration data available in the IECF database. Their content changes with time depending on the frequency of calibration & monitoring mission scenarios. Some calibration data updates occur on a

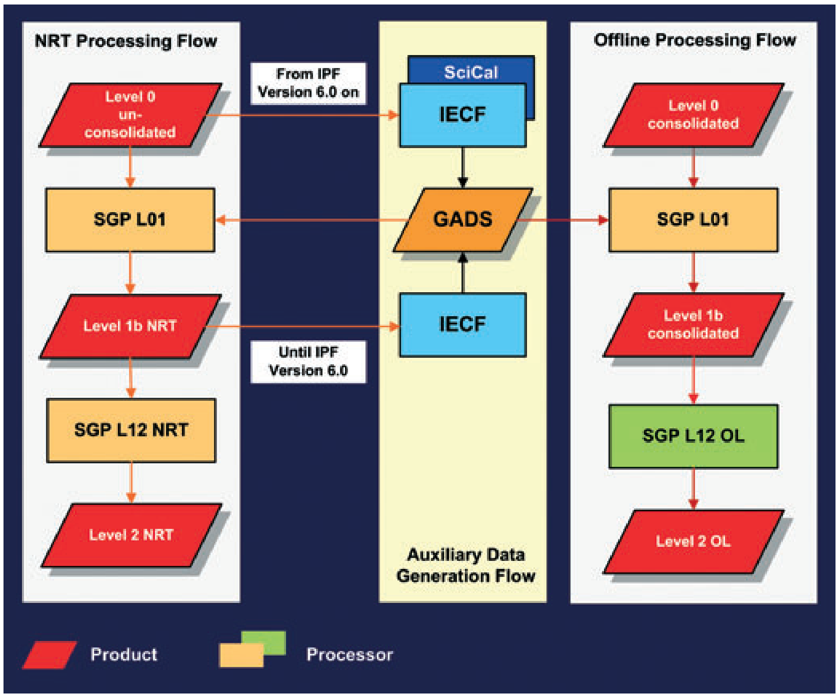


Fig. 8-1: General processing flow in NRT (left) and OL (right). Both chains receive the calibration information from the same auxiliary flow. (graphics: DLR-IMF)

daily basis, others on weekly, monthly or even longer time intervals. Until recently the generation of GADS was based on level 1b products and will be changed to level 0 with processor version 6.0 in 2006. Once the calibrated level 1b product exists, the corresponding level 2 product is generated by applying retrieval algorithms.

## Operational Processors Timescales

Separate near-realtime and offline chains exist because users request geophysical parameters on short and long timescales. NRT applications include forecasting so that the time elapsed between sensing and product dissemination must be kept to the order of a few hours. NRT dissemination usually occurs within 3 hours of acquisition. The products in OL mode can be delayed by up to several weeks. This leads to certain deviations between NRT and OL products. Not only geolocation is affected because of the non-availability of restituted orbit and attitude information shortly after sensing but also the method of how in-flight calibration measurements are reflected in the product. NRT products may rely on GADS which correspond to the previous day while the OL product, due to its late generation, can use the most recent GADS from the same orbit. In this case both the level 1b and level 2 content may vary slightly between NRT and OL data. A further significant difference depends on the start/stop times of the product. NRT processing is always initiated by unconsolidated level 0 data. OL products however are based on consolidated level 0 input. Thus only the OL level 1b and level 2 products allow a straightforward comparison with measurement execution.

In order to cope with the different requirements of NRT and OL processing the operational processors have been set up in the following way: For 0-1b processing only one single processor, SGP L01, exists. It serves both the NRT and the OL chain. The next step uses two processors, a NRT (SGP L12 NRT) and an OL (SGP L12 OL) processor. ESA managed the development and implementation of the SGP L12 NRT version. It has been integrated at the NRT processing sites in Kiruna and ESRIN to form together with SGP L01 the Instrument Processing Facility (IPF). Hosting site of the SGP L12 OL processor is the D-PAC. While the NRT chain is completely handled by the PDHS facilities, the OL chain is a combined effort of the LRAC and the D-PAC.

Basic processing related information is provided via the file names of the products. A filename includes the sequence:

- product ID: *SCI_NL__0P*, *SCI_NL__1P* or *SCI_NL__2P* for level 0, 1b or 2 data
- processing status flag: *N* for NRT products, letters between *N* and *V* for consolidated products
- originator ID: *PDK*, *PDE*, *LRA*, *D-P* for PDHS-K, PDHS-E, LRAC or D-PAC
- start date: year, month and day of measurement start
- start time: hours, minutes and seconds of measurement start in UTC
- duration of product coverage in seconds
- ENVISAT mission phase
- ENVISAT cycle number
- relative orbit number within cycle
- absolute orbit number at product start
- file counter
- file extension: *N1* corresponds to ENVISAT

which permit an unambiguous allocation of the file to a particular measurement orbit.

Information about SCIAMACHY's operational data products is provided via ESA's websites for the

ENVISAT mission. The detailed product descriptions can be found under:

http://envisat.esa.int/instruments/sciamachy/data-app/

and product status, including disclaimers, under

http://envisat.esa.int/dataproducts/availability/

All operational data products are subject to quality monitoring under the responsibility of the PDCC. The goal is to screen all products at the time of generation in order to identify anomalies or deviations from expected results. Quality monitoring includes content and consistency checks, e.g. formal correctness of the product or parameter limits. In the case of detected anomalies, data shall be flagged to initiate further actions. The PDCC executes quality monitoring activities on various timescales ranging from daily to multi-monthly.

## 8.2 Operational Level 0-1b Processing

SCIAMACHY level 1 data products comprise geolocated and calibrated radiances of the scientific measurements, as well as measurements for calibration and instrument monitoring. The algorithms used in operational level 0 to 1 processing are primarily driven by the scientific needs to convert measured signals into calibrated radiances (see chapter 5). However, constraints imposed by instrument operation, and in particular constraints imposed by the operational data processing environment, may force use of different strategies as would be employed for a ground-based instrument. The wish to obtain a level 1 data product which is not excessively large and complicated imposes an additional constraint. The principal processing cycle is outlined in fig. 8-2. It starts with instrument level 0 data and ends with the level 1b product and the level 1c extraction. A major constraint imposed by the ENVISAT PDS architecture rules is that there may be only one output product per processing chain, in this case the level 1b product. As a consequence, the level 1b product must not only hold processed science data but also calibration measurements and instrument monitoring data, as well as newly calculated calibration parameters. The latter are collected in the ADS (see above) in contrast to the measurement data which can be found in the Measurement Data Sets (MDS). In order to keep the size of the product as small as possible, the individual dark signal calibration measurements are discarded after calculation of the dark signal calibration parameters. For each dark state only the average signal of all measurements together with the deviation derived from the averaging is kept. All other calibration and monitoring measurements are retained, albeit in an unprocessed form.

In the operational processing from level 0 to 1b, all necessary calibration constants for each science measurement are processed from the input calibration data, ground-

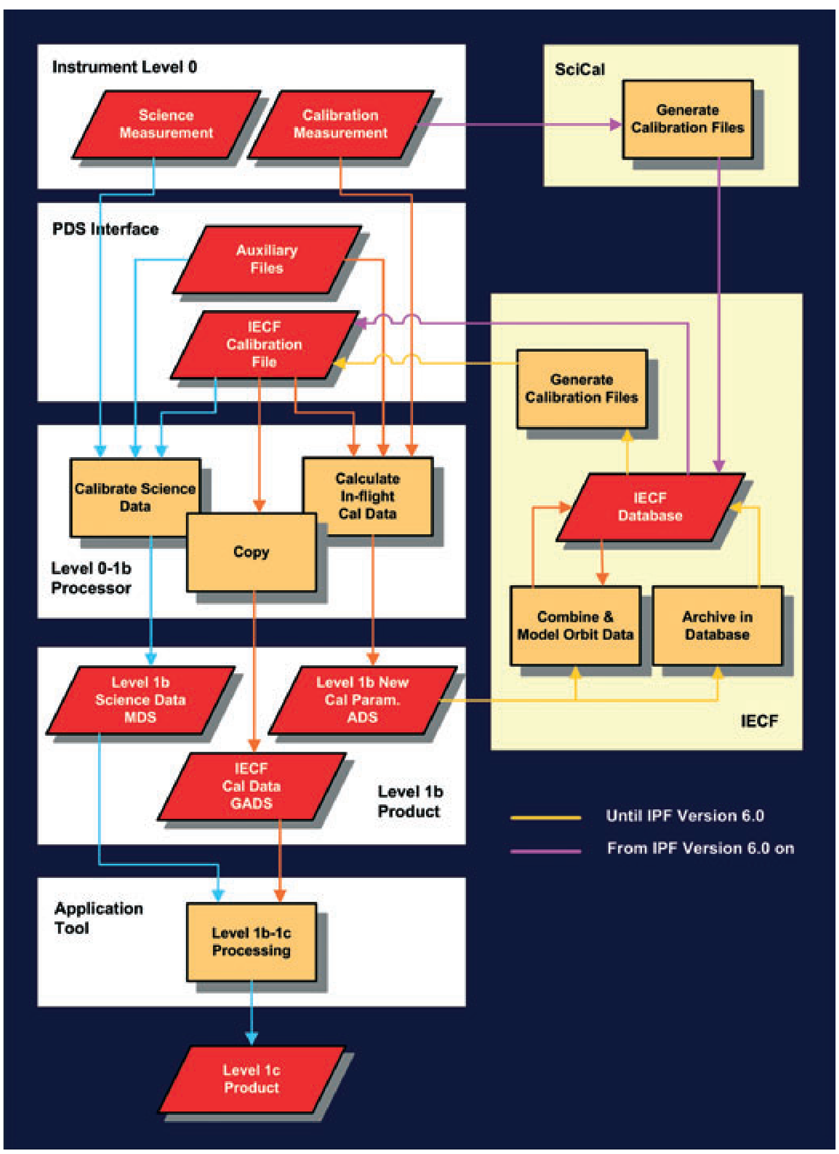


Fig. 8-2: Sketch of operational data processing from level 0 to level 1b and level 1c. (graphics: DLR-IMF)

based and in-flight as well. The level 1b data product contains the raw detector signals of the science measurements plus these calibration constants, mainly coded as single byte integers. In addition to measurement specific calibration constants, e.g. stray light or atmospheric polarisation determination for each measurement, lookup tables (LUT) are generated for globally applicable calibration constants, e.g. instrument polarisation sensitivity as function of scan angle.

The IECF generates selected calibration data based on level 0 products using the software SciCal. SciCal is also able to perform processing tasks such as modelling calibration data as function of orbit phase. These calibration data are copied into the calibration database of the IECF. In addition, the IECF is the tool to extract data from the calibration database and present them in the form of auxiliary calibration data products to the level 0-1b processor.

Level 1c data contain the fully calibrated measurements. The calibration application tool SciaL1c, available to the users, inflates the operational level 1b product into a level 1c product. For the user's convenience certain calibrations can be optionally omitted at extraction, or a subset of data can be filtered out. Level 1c products are not operationally generated by the PDS, they have to be generated by the users.

Fig. 8-3 provides an overview of the processing steps for level 0 to 1b and level 1b to 1c calibration of scientific measurements (nadir, limb and solar occultation), as well as the corresponding output of level 1b and level 1c data products in a flow diagram. The individual calibration related steps reflect the calibration characteristics as described in chapter 5 and specified by *Slijkhuis (2004)*.

## Global Calibration Constants

Globally valid calibration constants for the radiance and polarisation sensitivity of the instrument are taken from the calibration *Key Data* measured on-ground. For the determination of atmospheric polarisation in level 0-1b processing, the combined data from PMD and detector channels are required. Once the atmospheric polarisation is determined, the application of the polarisation correction in level 1b-1c only needs the polarisation sensitivity of the detector channels. This is provided for the level 1b product as a lookup table. A similar approach is pursued for the radiance sensitivity. A complication arises from the fact that the PMD channels are sampled independently and at a different rate from the channel detectors. Using the time stamps of both types of measurements, the PMD channels are re-sampled on a time grid synchronised with the detector readout, at twice the BCPS frequency (the BCPS corresponds to the maximum detector data packet generation rate of 62.5 msec). For channels 1-5 the time of detector readout depends on the pixel number, as each pixel is read out sequentially. The synchronisation is approximately performed for the part of the detector corresponding to the mean wavelength of each PMD passband.

## Memory Effect

All channel detector readouts have to be corrected for the Memory effect (channels 1-5) or non-linearity (channels 6-8). The Memory effect correction requires the value of the previous readout. If this is not available, for example for co-added measurements, an estimate has to be made. In case of co-adding, the value of each readout before co-adding is estimated on the basis of the PMD intensity which is sampled at a higher rate than the detector. Non-linearity correction is implemented similarly for co-added signals. A larger uncertainty in Memory effect correction exists for limb measurements: At each new tangent height a reset readout of the detectors with an integration time of 31.25 msec is performed. This reset readout is discarded on-board, i.e. it is not available in the data. Similarly, the Memory effect for the very first readout in a state can only be estimated, because before the first readout the detector picks up an unknown signal as the mirrors are driven to their start position. For both cases an estimation of the Memory effect is done based on the measured signal and state setup parameters.

## PPG & Etalon

The correction for PPG/Etalon requires input from in-flight calibration measurements. Changes in PPG and Etalon are expected to be generally slow. These are determined by regular comparison of the WLS measurement to a reference WLS spectrum from the on-ground calibration. A jump in Etalon structure may occur after an instrument switch-off. In such a case instrument operation may plan a WLS measurement just when regular science measurements are resumed thus obtaining information on a possible Etalon jump.

## Spectral Calibration and Dark Signal

Dark signal and spectral calibration depend on instrument temperature, and therefore on the position in the orbit (characterised by the 'orbit phase' on the level 1b product). Modelling over the orbit occurs in the IECF where information from several orbits may be grouped together, contrary to level 0-1b processing which always operates on just one orbit, i.e. the

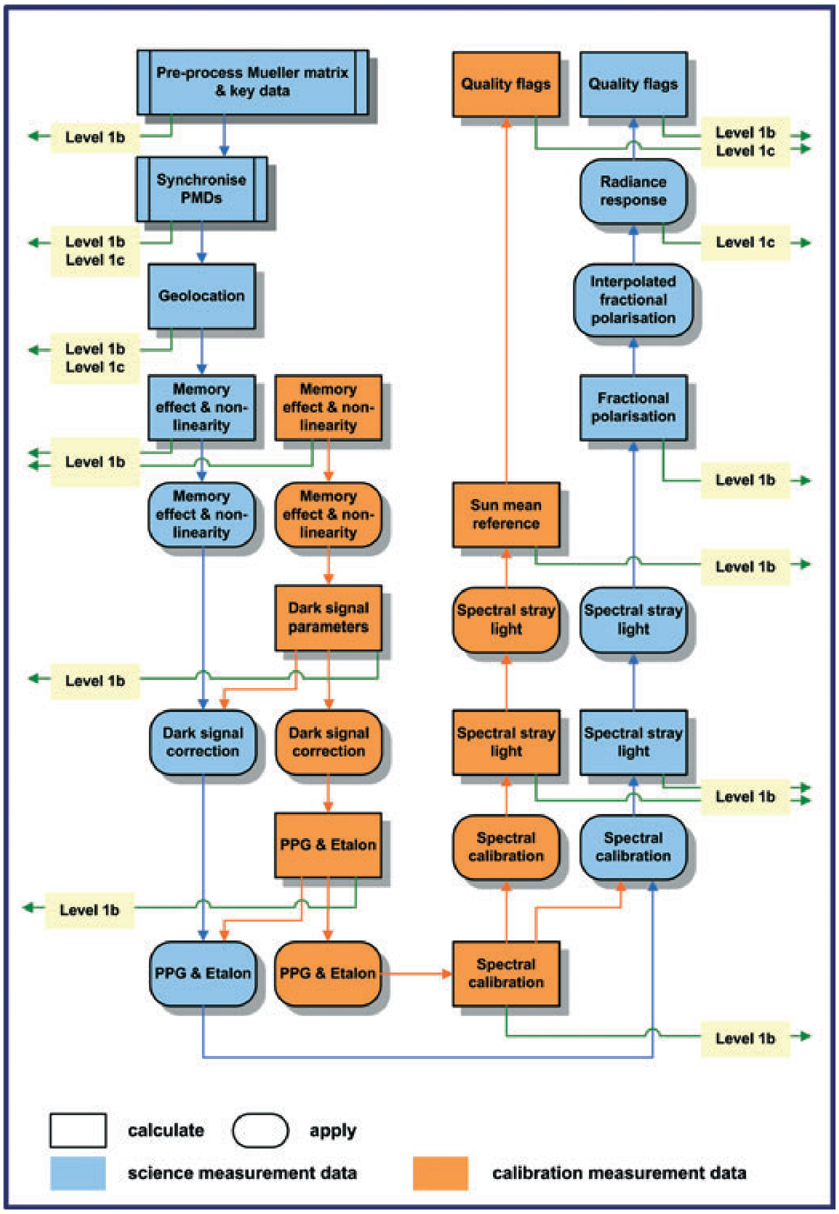


Fig. 8-3: Science and calibration measurements data flow. Algorithms applied to scientific measurements are in blue boxes, while calibration measurement algorithms are highlighted in orange boxes. Also indicated is the output into the level 1b or level 1c intermediate product. (graphics: DLR-IMF)

one provided in the level 0 file. In the original calibration concept it was anticipated that the drift in calibration would be slow, typically on a seasonal scale. While this remains true for the spectral calibration and for the dark signal in channels 1-6, the situation proved different for channels 7 and 8. In channel 7 and particularly in channel 8, a significant contribution to the dark signal originates in thermal radiation of the instrument itself. The unexpected layer of ice on the detector influences both the transmission of the source spectra and the thermal background (spectrally unresolved). As the thickness of ice may vary rapidly – noticeable on timescales of one orbit – weekly or even daily dark signal calibrations are no longer sufficient. For this reason, dark measurements have been implemented on the eclipse side of each orbit early in the routine operations phase. The level 0-1b processing has – since processor version 6.0 – for channels 6-8 the option to use these dark measurements from the same orbit. Usually, no calibration measurements are used which lie in the South Atlantic Anomaly (SAA). This is to reduce contamination from particle radiation which shows itself as spikes on the detector signal. For orbital dark measurements this rule would imply that for many orbits per day the dark signal is not updated. Therefore, measurements from the SAA are incorporated into the orbital dark calibration for channels 6-8, even though this may imply a small degradation in quality. A 'hot pixel detection algorithm' which uses for each detector pixel a sequence of measurements to detect and reject spikes in the data will filter out large deviations. The application of the dark signal uses in most cases the calibrated dark signal described above. Only for limb states there is a dedicated dark measurement performed at a tangent height of 250 km (early in the mission this was at 150 km until it became obvious that spatial stray light is still detectable at this altitude). Optionally the processor can be configured to use this dark measurement of the state itself.

## Spectral Stray Light and Polarisation

The corrections for spectral stray light and for atmospheric polarisation do not require in-orbit calibration measurements but are derived from each science measurement using calibration constants measured on-ground. Contrary to the situation on-ground, both corrections have to cope with the clustered readout of the detectors in flight. The stray light at one detector pixel is the sum of stray light originating from wavelengths corresponding to all other detector pixels in a channel, sometimes even of a neighbouring channel. If only part of the channel, i.e. one or more clusters is read out, the intensity at other wavelengths is not known and must be estimated. A similar situation occurs for the polarisation calculation where PMD signals, which are integrated over a broad wavelength band, are compared to channel detector signals. For level 0-1b processing, the stray light and polarisation calculation algorithms are applied after the spectrum covering the complete SCIAMACHY wavelength range has been estimated. Such a spectrum is calculated for each integration time in the state. After each longest IT, a complete spectrum has been measured. For shorter ITs, intensity from missing clusters is estimated by assuming that the spectrum for shorter ITs may only have a different intensity but retains the spectral shape of the longest integration time measurement, i.e. a complete spectrum for each IT is estimated by scaling the complete spectrum at the longest IT to the shorter ITs. On the level 1b product this is reflected by polarisation values $q$, $u$ which are given for all different ground pixel sizes defined for the state. For calibration of a ground pixel which is larger than the smallest ground pixel in the state it is more accurate to use the polarisation values derived for this size than to co-add fully calibrated clusters of smaller size. The latter carry larger estimated errors of stray light and polarisation. Polarisation values $q$ are calculated for 6 (broad) PMD bands. This calculation needs the corresponding value of $u$, which is derived iteratively based on the assumption that the ratio $u/q$ is constant and depends only on the polarisation angle. For a full description of related theories the reader is referred to chapter 5.5 which outlines in detail SCIAMACHY's polarisation concept.

### Sun Mean Reference Spectrum

For the purpose of trace gas retrieval in level 1-2 processing, the level 0-1b processing generates for each state of solar observations a Sun Mean Reference (SMR) spectrum. Although only diffuser spectra are to be used for level 1-2 processing, a SMR is generated for monitoring purposes from all solar states, including solar occultation states – in the latter case only spectra above a tangent height of 105 km are used. Absolute irradiance calibration on-ground was only performed for the ESM diffuser mode. The irradiance calibration of all other solar observation modes is based on a rough estimate which leads to an incorrect absolute irradiance level and may leave residual instrument polarisation features. Because of this, and because test trace gas retrievals showed good results with 'uncalibrated' ASM diffuser spectra, two SMRs are generated for each solar state: one with all calibrations applied and one without radiometric calibration. As the sun is a source of unpolarised radiation, solar measurements can be used to perform a self-calibration of the PMD channels which appears on the data products as the so-called *out-of-band signal*.

More details on the level 0-1b processing can be found in the relevant Algorithm Technical Basis Document (ATBD, *Slijkhuis 2004*) which will be updated after every major processor change.

## 8.3 Operational Level 1b-2 Processing

The goal of the level 1b-2 data processing is to provide geophysical parameters as column densities and profiles from atmospheric constituents as well as cloud and aerosol parameters. These data are given in the MDS of the level 2 product in combination with geolocation and additional auxiliary information (state geolocation, quality flagging, etc.) in the appropriate ADSs.

As outlined above, the operational level 1b-2 data processing occurs in NRT and OL mode using two different types of processors, SGP L12 NRT and SGP L12 OL. Since NRT processing has to fulfil the requirement to provide level 2 products within a certain time span, these products cannot be based on the most accurate level 1b product because of the extra calibration loop via the IECF. In order to ensure a sufficient NRT product performance, the level 2 NRT product palette has been focused on nadir observation geometry for the UV-VIS-NIR and SWIR spectral region. In contrast, the level 1b-2 OL product includes, in addition, retrievals from limb observation geometry. The operational processing of occultation measurements is currently not intended but the existing system can be extended to cope with it at a later stage. Restrictions in the current calibration status, partly caused by the ice layer in channels 7 and 8, prevent the PDS from retrieving geophysical parameters in the SWIR range. Only species specified in the UV-VIS-NIR are presently accessible. Nevertheless, algorithms and processors permanently evolve so that the product suite will be subject to changes. Table 8-1 summarises the planned level 2 product content.

The general processing chain for the level 1b-2 data processors is outlined in fig. 8-4 (nadir) and 8-5 (limb). It presents the functional flow. The software implementation is completely different between the SGP L12 NRT and the SGP L12 OL. While the SGP L12 NRT is a single node data processor, the SGP L12 OL is implemented within a multi-processor Linux cluster system. Besides the general application steps, for example level 1b input and auxiliary file input, the applications and data in green boxes are present for both processors. All other algorithms are

| Spectral Range | Nadir (NRT & OL) | | Limb (OL) |
|---|---|---|---|
| **UV-VIS-NIR** | $O_3$ (vertical column)[1)] | HCHO (slant column) | $O_3$ (profile) |
| | $NO_2$ (vertical column)[1)] | cloud cover | $NO_2$ (profile) |
| | BrO (slant column) | cloud optical thickness | BrO (profile) |
| | $SO_2$ (slant column) | cloud top height | |
| | OClO (slant column) | AAI | |
| | $H_2O$ (vertical column) | | |
| **SWIR** | CO (vertical column) | $CH_4$ (vertical column) | |
| | $CO_2$ (vertical column) | | |

[1)] = constitute METEO products available for the meteorological community

Table 8-1: Operational SCIAMACHY level 2 product suite. The SWIR products are strongly affected by the icing in channels 7 and 8, and operational processing within the PDS needs major revisions to cope with the icing.

only part of the OL processing chain. Since scientific progress in retrieval techniques continues, an outline of level 1b-2 algorithms can only provide a snapshot. The reader is referred to the ATBD (*Spurr 2000* with updates) which will be modified after every major processor change. Particularly experience gained from scientific product generation will be reflected in the operational chain stepwise. Assuming that scientific algorithms are usually more evolved than the operational ones due to development timescales determined by the rules for PDS processor s/w updates, it is a valid scientific requirement to regenerate operational level 2 products whenever significant algorithm know-how can be transferred from the scientific to the operational environment. Even beyond the in-orbit mission lifetime, an ambitious atmospheric science project like SCIAMACHY requires continuing retrieval algorithm research and reprocessing.

### Column Densities in the UV-VIS-NIR for Nadir Observations

NRT and OL data processing used the DOAS approach for the retrieval of slant column densities from launch onwards. The DOAS concept, together with the derivation of vertical column densities is described in chapter 7.2. Vertical column densities are deduced for the main species $O_3$ and $NO_2$, for all other trace gases listed in table 8-1 currently only slant columns are provided. The current operational DOAS algorithm is based on the approach originally implemented for the GOME data processor (GDP, *Thomas and Spurr 1999*). Since DOAS algorithm development undergoes permanent improvements – currently GDP 4.0 is operational (*Spurr et al. 2004*) – also SCIAMACHY's processor SGP L12 OL follows these changes. For $O_3$ and $NO_2$ the GDP 4.0 implementation yields vertical columns which are derived in an iterative way by taking into account radiative transfer calculations for the airmass factor and the DOAS slant column density determination for each iteration step. It replaces the early approach where a distinction between AMF extraction, slant column density derivation, and the vertical column density calculation was made.

### Cloud and Aerosol Parameters in the UV-VIS-NIR for Nadir Observations

The cloud parameter *cloud fraction* is determined by the threshold algorithm OCRA *(Loyola 1998)*. PMD measurements identified to be free of clouds from the red, green, and blue PMD are stored as RGB composite. The distance between the white point in the RGB reflectance space, which is dedicated to the fully cloudy case, and the actually determined PMD reflectance is taken to determine the cloud fraction. Heuristically derived scaling factors allow the appropriate representation of cloud fraction between zero (cloud-free) and one (fully cloudy). In the initial processor versions cloud top height or cloud top pressure was extracted from the ISCCP (International Satellite Cloud Climatology Project) database but it is replace from version 3.0 onwards by an algorithm based on evaluation of the $O_2$ A-band in the visible spectral region (SACURA, *Rozanov and Kokhanovsky 2004*). This yields improved cloud top height and cloud optical thickness parameters since in the former implementation only one value for optical thickness serves as fixed input.

The aerosol parameter AAI relies on the ratio of the reflectances between a spectral band which is not covered by spectral absorptions and a spectral band

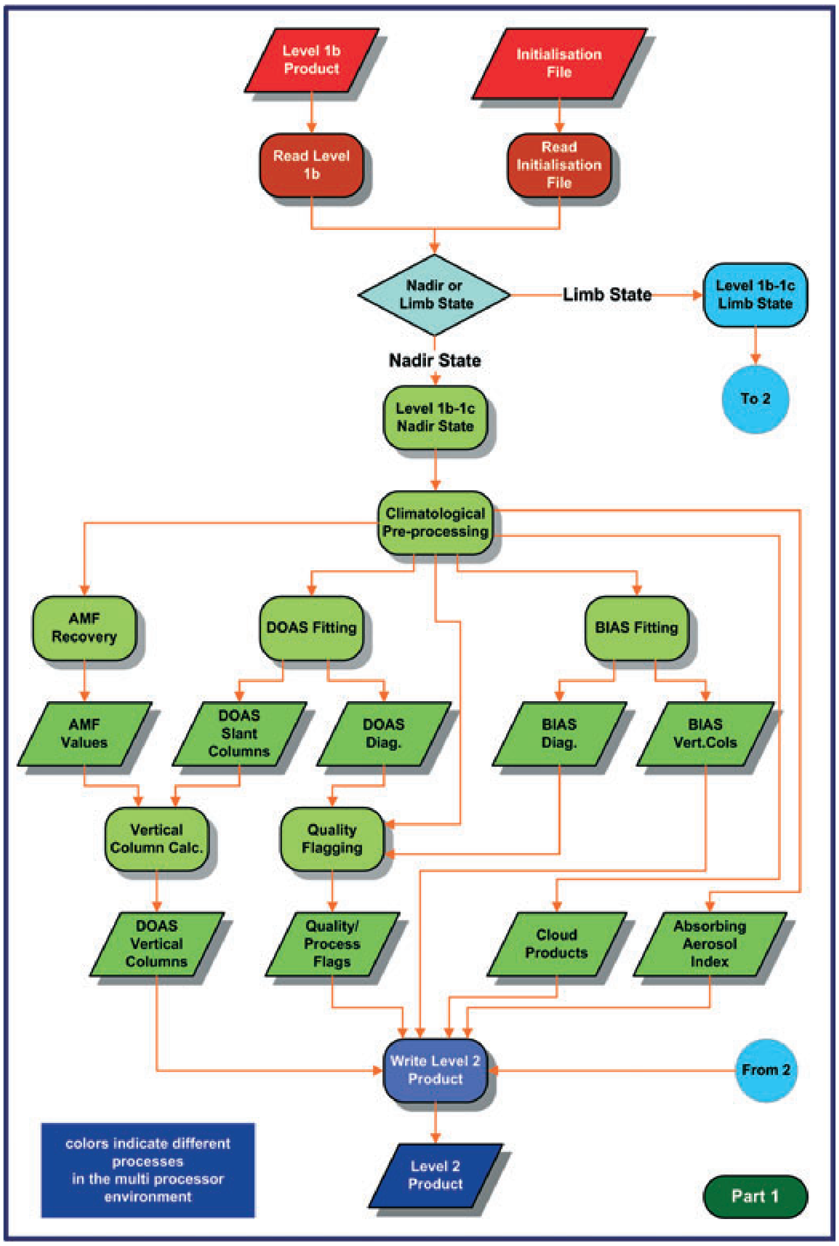


Fig. 8-4: Level 1b-2 data processing flow diagram. Limb algorithms are continued in fig. 8-5. (graphics: DLR-IMF)

with ozone absorptions. From the logarithmic difference, which is similar to the ratio of the reflectances, the AAI is computed. It indicates the presence of absorbing aerosols and depends strongly on an accurate calibration of the reflectance.

## Column Density in SWIR for Nadir Observations

In the SWIR range the cross-sections of line absorbers strongly depend on pressure, temperature and wavelength. Therefore, retrieval algorithms from the UV-VIS-NIR cannot be applied here and new algorithms, taking into account the characteristics of the SWIR absorption spectroscopy, are under development. With the ice on channels 7 & 8, where most of the trace gases were planned to be retrieved operationally, a major revision of the processor baseline was necessary and is still ongoing. Nevertheless, first results from scientific retrievals already demonstrate that meaningful data can be derived from the SWIR measurements. The know-how from the scientific retrievals will be transferred into the operational processing in the future.

## Profile Retrieval from Limb Observations

The original approach – until version 2.5 – for the retrieval of trace gas profiles from SCIAMACHY limb measurements represents an implementation of Optimal Estimation (OE). OE follows the original algorithm of *Rodgers (2000)* for the retrieval and a single scattering radiative transfer model improved by a multiple scattering correction LUT for the forward model. The retrieval is directly applied to intensities. These are constructed from the ratio of the radiance of the particular tangent height against the radiance of a reference height at 43 km for each tangent layer. $O_3$ profile retrieval is performed in the Huggins band while $NO_2$ profile information comes from the visible spectral region taking the ozone profile as auxiliary quantity into account.

Currently the limb retrieval baseline is under revision making it less dependent on strong spectral variations. The new algorithm (version 3.0), the Iterative Regularized Gauss-Newton Method *(Doicu et al. 2002)*, includes sophisticated techniques to select the iteration and regularisation parameters. A polynomial is fitted to radiances to remove broad spectral features and ratioed radiances are constructed for each tangent height. Finally the logarithm of the ratios determines the basis for the retrieval *(Doicu 2005)*. Fig. 8-5 depicts the s/w architecture of both limb retrieval algorithms with two preparational steps and the retrieval of the profile information. The function of regularisation matrix set-up does not appear in the OE case.

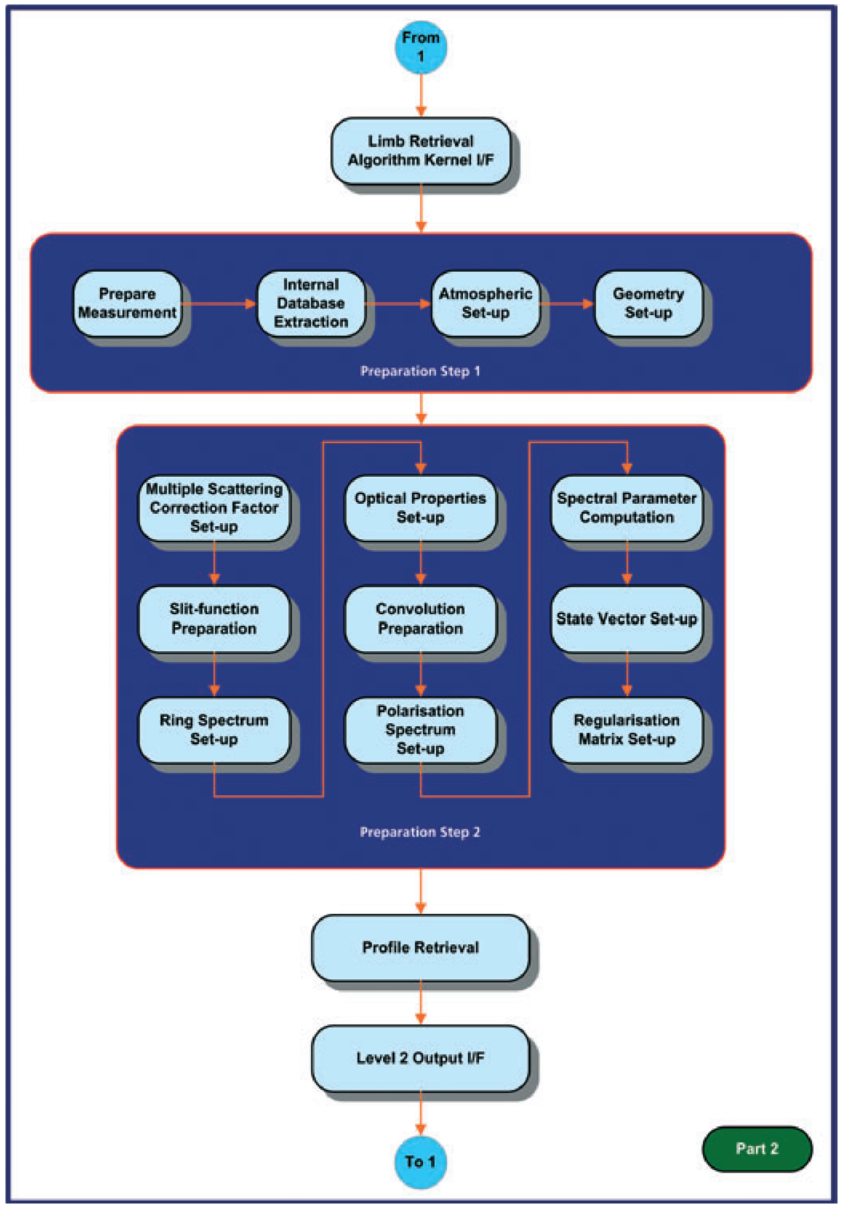


Fig. 8-5: Level 1b-2 data processing flow diagram for limb retrieval. At the end of the limb processing, the data flow returns to fig. 8-4 for product output. (graphics: DLR-IMF)

## 8.4 Scientific and Value Added Products

### Scientific Products

As explained above, scientific products are a vital service of the mission. Although scientific products usually refer to level 2, scientific level 1b data is also generated from operational level 1b or level 0 data by applying calibration steps differently to what is done in the operational ground segment. Since scientific product generation does not necessarily follow the stringent requirements of the operational chains – which does not preclude them from being as precise as possible and even permitting the utilisation of the most accurate and evolved algorithms – they may be implemented in various ways. Therefore it is not possible to provide here a detailed description of scientific processor architectures as for the operational ones. What can be given is a summary of the products available and a guide where to find additional information.

Table 8-2 presents a listing of geophysical parameters available as scientific products. Details can be found on SCIAVALIG's website under:

http://www.sciamachy.org/validation/products/

with succeeding species or parameter dependent links.

### Value Added Products

The VA products are generated on a routine basis for various applications. Their focus lies in an enhancement of data utilisation by users which are not directly connected to satellite retrieval and thus not trained in their exploitation. Requirements for value adding originate from very different user groups. Whereas scientific applications – if not based on level 2 data – usually need data sets on regular grids with complete information on pixel-wise data quality, applications for governmental monitoring duties, public services or education programs ask for synoptic, i.e. for one point in time, maps on different scales without any gaps and for long time series of continuous observations. In some cases, near-realtime data provision is also requested to initialise forecast services.

A wide set of requirements on atmosphere-related monitoring and forecasting are targeted within the ESA GMES Element project PROMOTE where atmospheric satellite observations are exploited. Topics under observation are distribution and trends of the stratospheric ozone layer, UV irradiance at the surface, near surface air quality and greenhouse gases. Application fields range from global and European protocol monitoring (Montreal, Kyoto, CLRTAP – Convention on Long-range Transboundary Air

| ***Geophysical Parameters*** | | | |
|---|---|---|---|
| $O_3$ | vertical column, profile<br>stratosphere and mesosphere | OClO | slant column, profile |
| $NO_2$ | slant/vertical/tropospheric column, profile | $H_2O$ | vertical column |
| BrO | slant/vertical column, profile | CO | vertical column |
| $SO_2$ | slant/vertical column | $CH_4$ | vertical column |
| HCHO | vertical column | $CO_2$ | vertical column |
| $T_{mesop.}$ | mesopause temperature | | |
| Clouds | cover, pressure, top height, optical thickness, liquid water path, phase index,<br>Top-of-Atmosphere (TOA) reflectance, droplet effective radius, PSC index, NLC index | | |
| Aerosol | AAI, AOT | | |
| UV | index, dose | | |

Table 8-2: Scientific SCIAMACHY level 2 product suite (due to the nature of scientific products,the table is evolving and does not claim to be complete)

Pollution) to public information services such as skin exposure sunburn time warning, street level air quality forecasts or flight route warning of dust events. Further currently developed applications include services to support planning and management of solar power plants (ESA market development project ENVISOLAR – Environmental Information Services for Solar Energy Industries) or to improve the quality of satellite land images by providing an automatic atmospheric correction with actual information on the state of the atmosphere (e.g. ClearView processor of DLR-DFD). In all these utilisation fields, quality information is essential to improve confidence in the VA products. In the future the use of value adding will most likely increase as new requirements call for integration of all observations – satellite, aircraft, balloons, ground – together with models to extract the best possible and comprehensive knowledge on the state of the atmosphere. For more details the reader is referred to the IGACO (Integrated Global Atmospheric Chemistry Observations Theme) strategy. One key tool for achieving this goal is the use of data assimilation techniques which provide a mathematically consistent technology to integrate measurements and models together with their statistical errors (see chapter 7.7).

Value added products are based on the operational products generated within the PDS. In most cases value adding transforms level 2 products from their satellite projection and the time of the satellite overpasses to integrated datasets (level 3 and 4) which contain a geophysical parameter as observed by one instrument. However, also synergetic value adding is conducted by either combining level 2 products of several sensors into one comprehensive level 3 and 4 data set or by exploiting level 1 products (top-of-atmosphere radiances or spectra) from different sensors to retrieve a level 2 product with new information. Averaging level 2 data over a certain period and grid cell generates level 3 products yielding daily, weekly, monthly or annual mean data sets. Level 4 products are based on the combination of satellite observations (level 2) and atmospheric models such as e.g. meteorological circulation and chemistry.