DLR

Deutsches Zentrum
für Luft und Raumfahrt e.V.

GOME

Level 0 to 1 Algorithms Description

TECHNICAL NOTE

ER–TN–DLR–GO–0022
Iss./Rev. 6/A
September 30, 2006

Deutsches Zentrum für Luft und Raumfahrt e.V. – DLR
Institut für Methodik der Fernerkundung – IMF
Oberpfaffenhofen, Germany

prepared by B. Aberle, DLR/IMF, _____ Date _____

W. Balzer, DLR/IMF, _____ Date _____

A. von Bargen, DLR/IMF, _____ Date _____

E. Hegels, (former DLR/IMF), _____ Date _____

D. Loyola, DLR/IMF, _____ Date _____

S. Slijkhuis, DLR/IMF, _____ Date _____

R. Spurr, DLR (now R.T. Solutions) _____ Date _____

Distribution List

Available to the Public via Internet

Document Change Log

Rev.	Date	Section	Description of Change	approved by	Sign
1/A	15.9.93	all	Completely new, only polarisation issues are covered		
1/B	15.11.93	1.3.2, 2.5, 3.5, 4	Update of the polarisation issues after comments by A. Hahne, K. Chance and R. Spurr		
2/A	18.2.94	all	Update of the polarisation issues after comments by G. Nuetzel D. Diebel and R. Spurr, inclusion of other algorithms and rename of the document (see Introduction)		
2/B	6.4.94	all	Minor updates after the GSAG meetings in March and inclusion of the error calculation section (provisional)		
3/A	15.7.95	all	Update of all algorithm descriptions according to the actual implementation and inclusion of new algorithms		
4/A	9.8.96	all	Update of all algorithms reflecting the 1st operational version		
5/A	19.12.00	all	Update of all algorithms reflecting the GDP Level 0 to 1 version 2.0		
5/B	10.04.02	all	Appendix B added		
6/A	30.09.06	all	Update of algorithms reflecting GDP level 0–to–1 version 3.0		

Table of Contents

0 Preface

0.1 Preface to Fourth Issue

The first issue of this document was released in September 1993; the second issue was written to include updated polarisation correction algorithm description and error calculations. The third issue (July 1995) updated the algorithm descriptions in the light of the first operational trials following the successful launch of the ERS–2 satellite in April 1995. After the instrument commissioning phase, the data processing validation campaign took place in the period October 1995 to March 1996, followed by a three–month period of analysis and trouble–shooting. In July 1996 the GDP 0–1 system reached a plateau, with the first public release of level 2 products scheduled for August 1996, and the first public release of Level 1 data planned for January 1997 following further improvements in key data and 0–1 algorithms..

At the final meeting of the ”Tiger Team” validation group at DLR on June 24/25 1996, a high priority recommendation was made to update GDP documentation. The 4th Issue of the Level 0 to 1 Algorithms Description Technical Note is one of the main outcomes of this recommendation. The basic algorithm descriptions in the original issues 1–3 are still relevant. In chapter 2, the algorithm texts have been left alone, and in order to avoid excessive re–writes of large amounts of text, extra ***comments pertinent to the current issue have been included in short additional subsections*** where appropriate.

0.2 Preface to Fifth Issue

After 5 year successful operation of GOME and its data processing, several algorithm updates have been performed inbetween. Thus, this issue will take care of the new implemented features and corrections of the algorithms described in the fourth issue. Additionaly, the GDP Update Report, which describes the changes of the algorithm initiated for the GDP level 0–to–1 version 2.0, is a further source of detailed information of the implemented changes. In order to keep track with the former description (issue 4), that description has been extended in some sections by adding a subsection which describes briefly the changes introduced with issue 5. The sections about the cross–correlation, the degradation of the BSDF–function, and the degradation of the measurement channels are totally new, since they cover new algorithmical ideas. The link to these sections is clarified at the appropriate subsections. Thus, we recommend to study the subsections which describe the additions of the 5th issue in all sections, especially for spectral and radiometric calibration.

0.3 Preface to Sixth Issue

After 11 years of GOME measurements, with the instrument still in good shape, ESA has decided on the first complete Level 0–to–1 processing of the entire GOME data set. The main driver for this reprocessing has been the gaps in GDP solar calibration spectra over long time periods, caused by pointing issues on the ERS–2 platform. The new GDP contains improved limit checking on solar spectra, improved line selection for better spectral calibration stability, and algorithmic changes described in this document. The precision error on the solar spectrum has been reformulated, and several options have been updated in the GDP extraction software, following the CHEOPS–GOME study. This comprises improved Peltier crosstalk correction in Band 1a, seasonal correction on BSDF, improvement in Polarisation Correction, and a new degradation correction of Earthshine Reflectivity. As for the Fifth Issue, we maintained the older algorithm descriptions and added the new ones to the end of each chapter. A more detailed description of the new extractor options has been added to the Appendix C.

1 Introduction

1.1 Purpose and Scope

The German Remote Sensing Data Center (DFD) plays a major role in the design, implementation and operation of the GOME Data Processor (GDP). GDP is the off–line ground segment for the GOME instrument on the ERS–2 satellite. It incorporates a Level 0 to 1 processing chain, the complete GOME data archive, the DOAS O_3 total column retrieval process, and an image processing chain for the generation of higher level products.

During the Level 0 to 1 processing GOME data is converted into ”calibrated radiances” by applying calibration algorithms and calibration parameters. Many calibration parameters are established on a regular basis from in–flight observations of the calibration lamp, the internal LED and the sun and under dark conditions. In addition, data from pre–flight instrument calibration and characterisation (e.g. the polarisation characteristics of the optical chain) is required. The Calibration and Characterisation Subcommittee of the GOME Scientific Advisory Group delivered a reference document for the entire calibration activities (”Scientific Requirements for the Calibration and Characterization of the Global Ozone Monitoring Experiment” [R3]) and the GOME project at ESTEC prepared a document about the required in–flight calibration activities (”GOME In–flight Calibration and Characterisation Plan” [R4]) which were taken into consideration in the present technical note.

The following auxiliary data (which are *not* part of the raw GOME data) are required in GDP level 0 to 1 processing:

- orbital information (e.g. state vectors from ESOC) to generate the geolocation data,
- time correlation file to correct the time in the raw data (from ESOC),
- pre–flight calibration and characterization data (from TPD and ESTEC).

Structure and usage of the orbital information and the time correlation file is well known from ERS–1 ground segment. The orbital information is used by a satellite orbit propagator programme (originally this was VENI, but replaced recently by the ESTEC ERS–2 satellite propagator) to calculate the geolocation information during scanning and the azimuth and zenith angles on the sun diffuser during sun calibration measurements. The usage of the time correlation file is described in section 2.1.

The initial scope of the present document was the evaluation of the results of the breadboard model (BBM) calibration and characterisation in view of the requirements for the Level 0 to 1 processing. These results were provided and initially evaluated by TPD/TNO (Delft, Netherlands) under contract of the GOME project at ESTEC. During the results review meeting at TPD in September 1993 the GDP project received the action item to closely review the presented information and documentation which was provided prior to that meeting.

After the first versions of the present document it became clear that the contents *are more or less the algorithm description for the Level 0 to 1 Processing.* Therefore, it was decided to include also algorithms which do not require pre–flight data and to change the name of the document from ”GOME Pre–Flight Calibration Data Requirements” (Issue 1) to ”GOME Level 0 to 1 Algorithms Description” (Issue 2 and beyond).

During the pre–flight calibration, some changes to the initial understanding of the behaviour of the GOME instrument occurred (cross–talk during dark measurements, etc.). These generated some alterations to the structure of the algorithms and the actual implementation of the software. Issue 3 of the present document reflects the status of the implemented algorithms immediately prior to launch.

Issue 4 reports on the status of the *first operational system*, wherein additional changes suggested during the validation campaign have been implemented (for example, the change of orbit propagator or the correction for Peltier noise.

1.2 Definitions, Abbreviations and Acronyms

A list of about all abbreviations and acronyms which are used throughout the SRDs for the GDP is given below:

BBM	Breadboard Model
BSDF	Bi–directional Scattering Distribution Function
CU	Calibration Unit
DFD	Deutsches Fernerkundungsdatenzentrum
DLR	Deutsche Forschungsanstalt für Luft– und Raumfahrt e.V.
DOAS	Differential Optical Absorption Spectroscopy
ERS	European Remote Sensing Satellite
ESA	European Space Agency
ESOC	European Space Operation Centre
ESTEC	European Space Center of Technology
ESRIN	European Space Research Institute
FM	Flight Model
FPA	Focal Plane Assembly
FPN	Fixed Pattern Noise
FSM	Flight Spare Model
GDP	GOME Data Processor
GOME	Global Ozone Monitoring Experiment
HK	House Keeping (Data)
LED	Light Emitting Diode
IFE	Institute of Environmental Physics and Remote Sensing at the University of Bremen
IMF	Institut für Methodik der Fernerkundung
MMCC	Mission Management and Control Center
PMD	Polarisation Measurement Device
PPG	Pixel–to–Pixel Gain
SAO	Smithsonian Astrophysical Observatory
SRD	Software Requirement Document
SRON	Space Research Organization Netherlands
TPD/TNO	Technisch Physische Dienst
UV	Ultra–Violet
VENI	Visibility and Ephemeris Investigations for Satellite Orbit Analysis

1.3 Documents

1.3.1 Applicable Documents

The following applicable documents provide background material:

[R1] ”The Interim GOME Science Report” September 1993

[R2] ”Report of the GOME Scientific Advisory Group, Data and Algorithm Subcommittee” February 1992

[R3] ”Report of the GOME Scientific Advisory Group, Calibration and Characterization Subcommittee, Scientific Requirements for the Calibration and Characterization of the Global Ozone Monitoring Experiment” November 1992

[R4] GOME In–flight Calibration and Characterisation Plan, ER–PL–ESA–GO–0303, Issue 1/–, August 93

1.3.2 Reference Documents

The following documentation was taken into account for the descriptions in the present document:

[A1] GOME Requirements Specification, ER–RS–ESA–GO–0001, Issue 2/–, 15.12.1993

[A2] GOME Calibration Facility Design Review Data Package, Issue 1, 18.11.1992

[A3] GOME BBM Calibration Results Review I Data Package, Issue 1, 7.9.1993

[A4] GOME BBM Calibration Results Review II Data Package, Issue 1, 9.3.1994

[A5] GOME FM Calibration Results Review Data Package, Issue 1, 14.11.94

[A6] GOME FS Calibration Results Review Data Package, Issue 1, 10.1.95

[A7] System Requirements Document of the GOME Data Processor, ER–SR–DLR–GO–0020, Issue 1, 15.7.93

[A8] Functional Software Requirements of the GOME Data Processor (Level 1), ER–SR–DLR–GO–0008, Issue 1, 15.7.93

[A9] A simple estimate of polarisation relevant to GOME and SCIAMACHY, Piet Stammes (KNMI), 4.10.1993

[A10] Data Reduction from Experimental Histograms, W.R. Falk (University of Manitoba, Winnipeg, Canada), Nuclear Instruments and Methods in Physics Research 220 (1984) 473–478

[A11] GOME Dark Signal Characterization, ER–TN–ESA–GO–0473, C.Caspar, 06.11.1995

[A12] Problems in using diode arrays for open path DOAS measurements of atmospheric species, Jochen Stutz and Ulrich Platt, Institut für Umweltphysik, Universität Heidelberg

[A13] PtCrNe lamp lines,TPD line list TPD–ERS–GO–TN–91 & TPD–ERS–GO–MIR–2 and ESA Study FTS Measurements by Murray

[A14] Moon Study

[A15] Serco, Tracegas Study Final Progress Report, ESA Contract 10728/94/NL/CN, 2.3.95

[A16] Journal of Research of the National Institute of Standards and Technology, January–February 1992, Volume 97, Number 1

[A17] CRC Handbook of Chemistry and Physics, 69th Edition, 1988–1989, CRC Press Inc., Boca Ratona, Florida

[A18] Numerical Recipes in C, William H. Press et. al., Cambridge University Press, 1988

[A19] C. Caspar and K. Chance, GOME wavelength calibration using solar and atmospheric spectra. Proceedings of the Third ERS Symposium on Space, 1997

[A20] C. Tanzi, pers. communication, 1999

[A21] A. von Bargen and S. Slijkhuis, Wavelength Calibration of GOME Spectra Utilising Cross–correlation, Proceedings of the European Symposium on Atmospheric Measurements, ESA WPP–161, pp709–713, Noordwijk, The Netherlands, 1999

[A22] J. Tonry and M. Davis, A Survey of Galaxy Redshifts: I. Data Reduction Techniques, Astron. J., 84(10), 1511–1525, 1979

[A23] GDP Update Report, ER–TN–DLR–GOM–0043, Issue 1/A, August 1999

[A24] R.L. Kurucz, I. Furenlid, J. Brault, and L. testerman, Solar Flux Atlas from 296 to 1300 nm, National Solar Observatory, Sunspot, New Mexico, 240pp., 1984

[A25] Technical Report, TPD TPD–ERS–GO–MIR–14, Issue 2, 03.11.1994

[A26] GDAQI – Final Report, TN–GDAQI–003SR/2000, SRON, September 2000

1.4 Overview

The present technical note includes the following sections:

- Level 0 to 1 Algorithms
  The following algorithms are described:
  - Conversion to Engineering Units
  - Correction for Leakage Current
  - Correction for FPA (Peltier) cross–talk noise
  - Correction for Pixel–to–Pixel Gain
  - Correction for Straylight
  - Spectral Calibration
  - Radiometric Calibration
  - Polarisation Correction
  - Degradation Correction
  - Quality Flagging
  - Error Calculation
  - Other Algorithms and Procedures
- Summary of Pre–flight Calibration Data Requirements
  This section lists the pre–flight data requirements.
- Appendix
  Describes some of the mathematical derivations used in the polarisation correction algorithm and of the cross–correlation option implemented in the Level 0–1 extractor software.

1.5 Other Algorithms and Procedures

1.5.1 Moon Calibration

GOME can observe the moon whenever it crosses the sensor's field of view. These moon observations are a potential supplement to the radiometric calibration, and could be used to characterise degradations of the light path (e.g. contaminations of the scan mirror). Without going into the details, there are a lot of problems which have to be solved before it might be possible to obtain any significant information from moon observations [R4]. The complete matter was subject of a special moon calibration study [A14]. It is expected that a reasonable algorithm could not be developed before real moon observations were available. The moon measurements are stored in the level 1 product.

2 Level 0 to 1 Algorithms

To understand the environment of the Level 0 to 1 algorithms figure 1 shows the functional model diagram of the Level 0 to 1 Processing, including the main functions and the internal and external data stores. This figure was taken from the Software Requirements Document of the GOME Data Processor [A8]. It provides a convenient overview of the level 0 to 1 algorithms, and the original conception for this model diagram is still valid.

The conversion to engineering units is part of the *Split_Orbit* function. The following six boxes shaded gray in the diagram will be covered in detail in six corresponding sections in this chapter. There are two additional sections dealing with algorithms not mentioned in the diagram (FPA noise and Straylight corrections). The penultimate section deals with error propagation through level 0 to

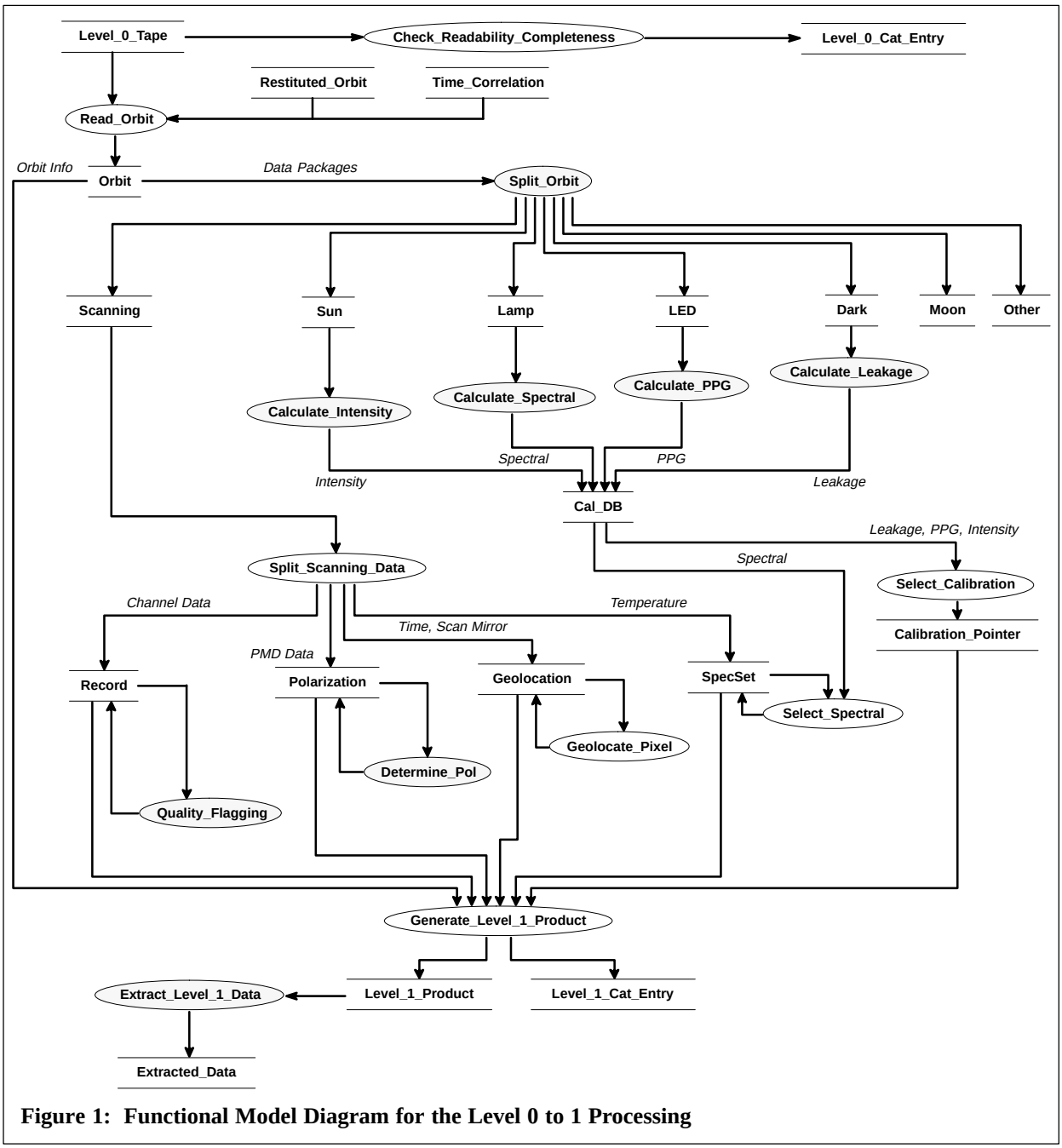

Figure 1: Functional Model Diagram for the Level 0 to 1 Processing

1 processing, and the final section mentions briefly other algorithms not included as part of the operational level 0 to 1 processing chain.

All level 0 to 1 algorithms (except one) fall naturally into 2 stages – (a) the *derivation* of appropriate calibration parameters from GOME raw data and ancillary information, and (b) the *application* of those parameters to generate the desired calibrated radiances at level 1. The application of calibration parameters is the main task of the *Extract_Level_1_Data* function (also shaded gray in Figure 1.) The exception to this rule is the Straylight Correction; this is a single step implementation and is performed after the application of the PPG correction algorithm.

The separation between the derivation and application of calibration parameters is required for the operational control flow during the Level 0 to 1 Processing. For convenience of exposition in the present document, this distinction is not kept in the descriptions that follow.

2.1 Conversion to Engineering Units

The GOME data packages include a variety of data with different formats and meanings. Besides the detector array and PMD signals (which are expressed as binary units until the application of the radiometric calibration), there are the following categories of information:

- digital information
  these values are used as they appear; no further transformations are required. Example – frame counters.
- enumerations
  these data consist of either one bit or two bits which have a certain semantic meaning; no further transformations are required. Example – mode switches for detector array temperatures.
- digitised analog measurements
  these values are converted by the use of a polynomial (see below); e.g. temperatures of the optical bench.
- time information
  the time is given by a 32–bit wrap–around counter; the time correlation files received from MMCC are required to transform this counter into a UTC time (see below).

For the digitised analog information the following formula applies:

$$y = a_3^* x^3 + a_2^* x^2 + a_1^* x + a_0^* \quad (1)$$

where x is the digital input from the data package, y the analog result value and a_i^* (i = 0..3) are the polynomial coefficients for a certain value of the data package.

The time correlation file is given at the ascending node of the corresponding orbit and consists of the following information:

- Orbit number;
- UTC_{days} (the number of days since the 1.1.1950);
- UTC_{msecs} (the milliseconds of the day);
- SBT_{TC} (the satellite binary counter of the above UTC time);
- $SBTP$ (the satellite binary counter period which is the number of nanoseconds yielding one tick of the satellite binary counter)

The UTC time of the data package is calculated as follows:

$$UTC_{DP} = UTC_{TC} + \frac{(SBT_{DP} - SBT_{TC}) \cdot SBTP \cdot 10^{-6}}{86.400.000} \quad (2)$$

where UTC_{DP} and UTC_{TC} are the UTC times in days of the data package and the time correlation file respectively. SBT_{DP} and SBT_{TC} are satellite binary counters of the data package and the time correlation file respectively. $SBTP$ is the satellite binary counter period which is given in nanoseconds. The factor 10^{-6} is required to transform nanoseconds to milliseconds and 86.400.000 is the number of milliseconds in one day. UTC_{TC} is calculated as follows from the entries in the time correlation file:

$$UTC_{TC} = UTC_{days} + \frac{UTC_{msecs}}{86.400.000} \quad (3)$$

Remark: The subtraction in formula (2) takes into account the wrap–around nature of the satellite binary time counter in the case when SBT_{DP} is close to zero and SBT_{TC} is close to (2^{32}–1).

2.2 Correction for Leakage Current and Determination of Noise

The detectors used for the 4 channels of the GOME sensor are random access linear photo–diode arrays. One characteristic of these devices is a certain amount of leakage current. This is produced by thermal leakage and it is expected that this current will depend on the orbital position of the satellite and also the time into the mission (sensor degradation). Therefore it is necessary to continuously monitor the leakage current and associated noise, and this is done by means of periodically taken dark–side measurements. From these, one can correct the charge readouts of the detector pixels for any leakage current contribution.

The PMD detectors are non–integrating devices, and therefore do not have a leakage current. Nevertheless those detectors must be corrected for their zero offset and the noise level from dark–side measurements must be monitored.

The following definitions are relevant:

$S_i^{measured,\,k}$	arbitrary measured signal of detector pixel *i* with an integration time pattern *k* for complete detector arrays [BU]
S_i	signal of detector pixel *i* corrected for leakage current [BU]
I_i^{DC}	leakage current of detector pixel *i* [BU/s]
S_i^{FPN}	fixed pattern noise of detector pixel *i* [BU]
t	integration time [s]
$S_i^{Dark,\,k}$	signal of detector pixel *i* taken under dark conditions with an integration time pattern *k* for complete detector arrays [BU]
$\overline{S_i^{Dark,\,k}}$	mean signal of detector pixel *i* with an integration time pattern *k* for complete detector arrays [BU]

In theory, the signal of a detector pixel consists of the following components:

$$S_i^{measured} = S_i + I_i^{DC} \cdot t + S_i^{FPN} \tag{4}$$

where $S_i^{measured}$ is the measured signal, and the other quantities are defined above. Due to cross–talk generation when the different channels are not integrating with the same readout times, it is not possible to determine values of I_i^{DC} and S_i^{FPN} valid for all integration times. Instead, one must measure under dark conditions with the same integration time patterns as those employed for scanning and other calibration measurements. This requires at least 10 (the more the better) consecutive dark measurements for each integration time pattern; from these a mean value array of leakage charge signals can be calculated:

$$\overline{S_i^{Dark,\,k}} = \frac{1}{n}\sum_{j=1}^{n}\left(S_i^{Dark,\,k}\right)_j \tag{5}$$

This means that the leakage current correction is simply the selection of the right leakage charge array for integration time pattern *k*, and the subtraction of this array from the main measurement:

$$S_i = S_i^{measured,\,k} - \overline{S_i^{Dark,\,k}} \tag{6}$$

Because of the low temperature (–38°C) of the detector arrays, the leakage current itself is very small (about 1 BU/s). Therefore, the change of the leakage current $I_{i,\,DC}$ under illumination is so small that there is no need for a second correction, as was originally proposed in the literature [A12].

The noise on the signals of the detector pixel readouts is also expected to be very small and a constant over all detector pixels. All dark measurements(with the same integration time) from one orbit are

used to calculate standard deviations for each detector pixel; the mean value of these standard deviations will yield the mean noise value for the detector pixel readouts.

For the PMDs, mean values of the dark measurements over the same orbit period are calculated to yield mean zero offsets. Mean noise values on the PMDs are calculated similarly to those of the detector arrays.

The influence of the orbital variation on the leakage current and the detector noise (including the effect this variation has on the non–illuminated pixels) will be the subject matter of a study during the commissioning phase. This study may lead to an algorithm to correct the leakage current and noise value with respect to the orbital position.

The following diagram (figure 2) identifies the sequence of steps required for the correction of leakage current.

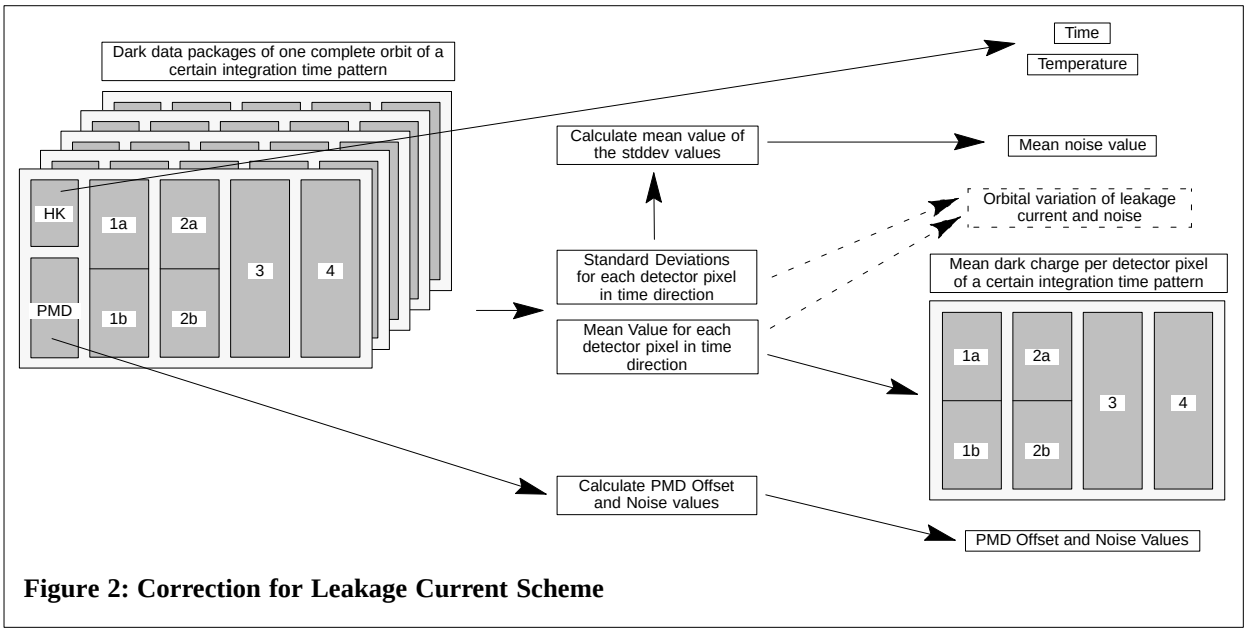

Figure 2: Correction for Leakage Current Scheme

2.2.1 Algorithm Update (Issue 4)

Following the introduction of the *co–adding mode* for scanning measurements, the definition of an ”integration pattern” has to include not only the integration time of each band and a flag indicating the integration completion, but also an additional flag indicating whether the co–adding mode was active or not. Measurements made using the co–adding mode may have a different leakage signal than measurements taken without co–adding.

2.3 Correction for FPA Noise

This Algorithm is based on the investigation of C. Caspar [A11]. An additional source of noise on the reticon signal is correlated to the voltage controlling the Peltier coolers on the Focal Plane Assembly (FPA).

This noise is correlated to the integration time. It can be determined by multiplying the Peltier cooler control signal by a scaling factor, which has to be included in the pre–flight data. For each integration time, a different scaling factor is required. A correction is only necessary for integration times of 6s or longer (channel 1 measurements only).

2.3.1 Summary of Algorithm Steps

The following steps are implemented in the correction for FPA noise :

- Collect all Peltier output signals of the orbit.
- Remove long–time trends in the Peltier output signals by applying a high pass filter.
- Calculate an average value of the Peltier outputs for the current integration period. Since Peltier results are available every 1.5s, then for an integration period of 12s, a mean value of 8 Peltier outputs can be constructed.
- Calculate the FPA noise by multiplying the mean Peltier output by the (pre–flight) scaling factor specified for this integration time. Figure 3 shows typical FPA noise data for one orbit.
- Subtract the FPA noise from the signals of the entire band to be corrected.

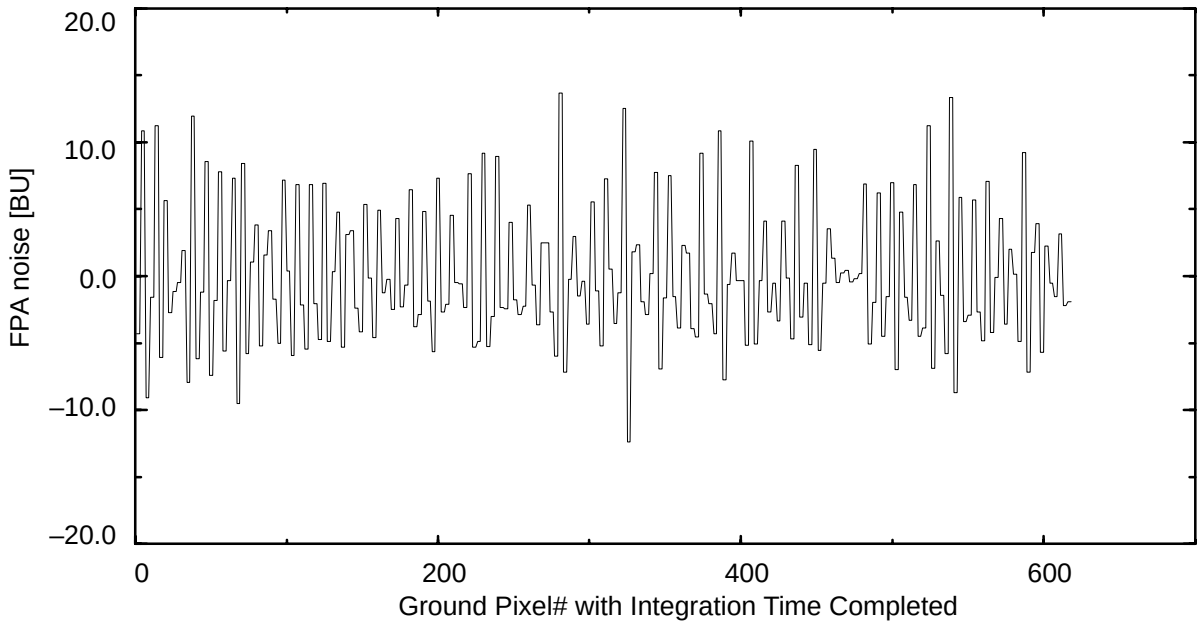

Figure 3: Typical FPA Noise of one Orbit

2.3.2 Algorithm Update (Issue 6)

It appears that after removal of the Peltier crosstalk as described above, a residual offset remains which is too large for e.g. O3 profile retrieval. A correction of this residual offset, using the signal in the ”straylight 1a” band, is now implemented. The application of FPA correction is done in two steps: first the correction as described in section 2.3.1, followed (optionally) by straylight correction, and then correction for residual offset. For details, see section C.1 .

2.4 Correction for the Pixel–to–Pixel Gain

One of the characteristics of a diode detector array is that the individual detector pixels have slightly different sensitivities. To correct for this feature, the determination of pixel–to–pixel gain (PPG) correction factors is required. Figure 4 shows typical PPG data (dotted line fine structure); the solid line is the smoothed PPG data (see below for explanation).

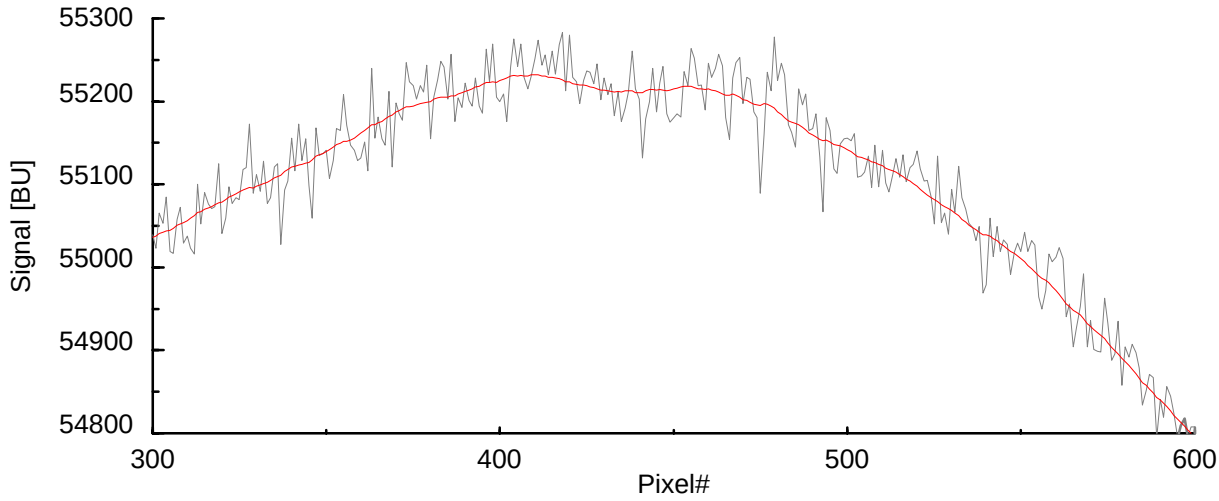

Figure 4: Typical Pixel–to–Pixel Gain Data

From the GOME BBM calibration and characterisation at TPD, two results have emerged which are important for the PPG correction:

1. the variation of the sensitivity from one pixel to the other is around 0.2% (see figure 5). This means that the size of the absolute error is not very large. Nevertheless, when looking for weak absorption features the relative error (between individual pixels) could be of great importance.

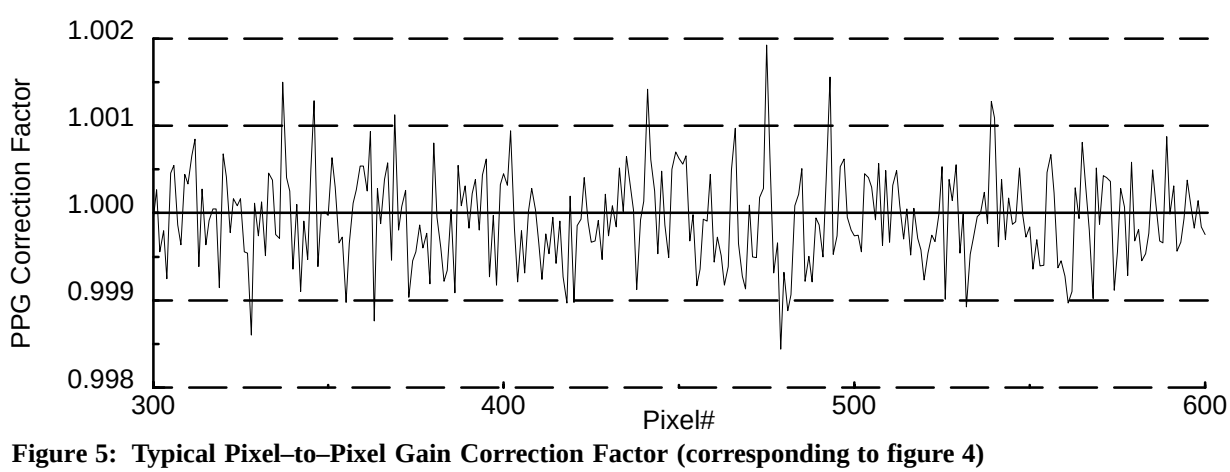

Figure 5: Typical Pixel–to–Pixel Gain Correction Factor (corresponding to figure 4)

2. Owing to reflections at lens mounting components of the BBM, the smooth illumination (3rd order polynomial) from the LED on the detector array was disturbed.

The conclusion from these observations was that the calculation and application of the PPG correction (leaving aside the coarse evaluation of defective or obstructed pixels) appeared to be unreasonable for this instrument, yet it was clear that an algorithm for PPG correction was required. The situation was improved considerably with the advent of constructive changes to the flight models (FM

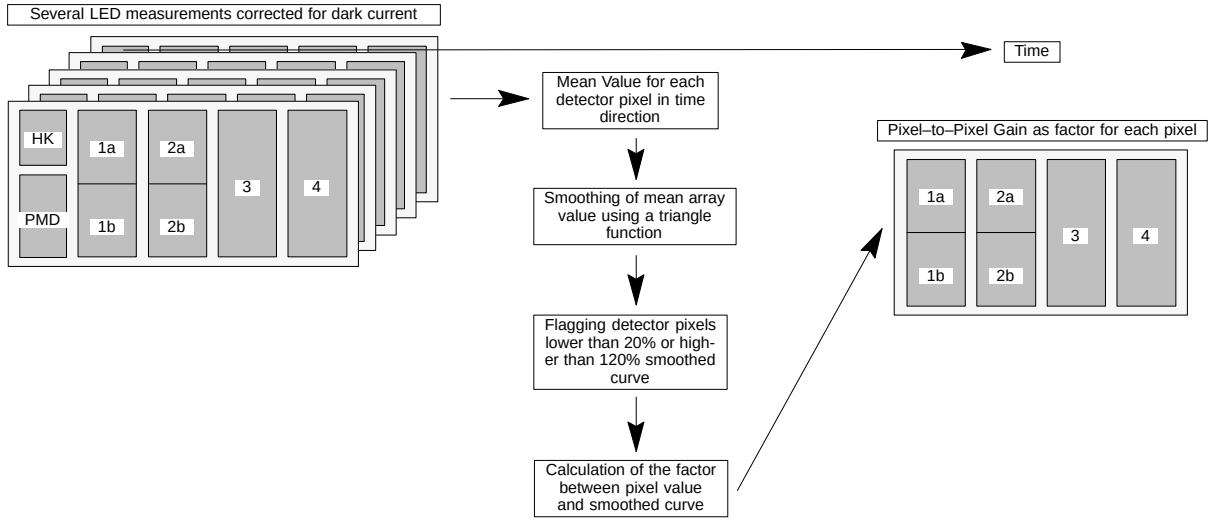

Figure 6: Correction for Pixel–to–Pixel Gain Scheme

and FSM) to reduce these reflections. Now, instead of fitting a third order polynomial to the LED measurement as was initially proposed, a more sophisticated smoothing algorithm has been set up to assess the pixel–to–pixel gain correction factors.

The following definitions are needed:

S_i	measured scanning or sun calibration signal value of detector pixel i [BU]
S_i^{corr}	corrected scanning or sun calibration signal value of detector pixel i [BU]
$\overline{S_i^{LED}}$	mean value of several measured LED signal values of detector pixel i [BU]
S_i^{smooth}	smoothed signal value of detector pixel i [BU]
c_i	pixel–to–pixel gain correction factor of detector pixel i [–]
n	smoothing value (width of triangle filtering window) [–]

The PPG correction factor is defined as follows:

$$c_i = \frac{S_i^{smooth}}{\overline{S_i^{LED}}} \tag{7}$$

where $\overline{S_i^{LED}}$ is the mean value of the signal values of several consecutive LED measurements and S_i^{smooth} the smoothed curve through this averaged measurements. The smoothed curve is calculated by means of:

$$S_i^{smooth} = \frac{\sum_{k=-n}^{n} \frac{n-|k|}{n} \cdot \overline{S_{i+k}^{LED}}}{\sum_{k=-n}^{n} \frac{n-|k|}{n}} \tag{8}$$

using a triangle filtering window of width n. The application of the PPG correction is then simply:

$$S_i^{corr} = S_i \cdot c_i \tag{9}$$

2.5 Correction for Straylight

2.5.1 Introduction

After the first calibration and characterisation measurements of the GOME BBM at TPD, it became obvious that straylight was a major issue; something which not only has to be characterised, but corrected for during the in–flight calibration exercise. Specifically in channel 1 and 2 the signal readouts are spoiled by a non–negligible amount of straylight. The following main sources of straylight in GOME were identified:

- a uniform (or very slowly changing) quantity of straylight over the detector pixels induced by diffuse reflections within the FPA;
- ghost straylight signals induced by reflections from the surfaces of the detector arrays and the lenses of the channel telescope. Two types of ghosts were detected during BBM calibration and characterisation (see [A3]):
  - symmetrical ghosts, which are signals mirrored at the middle of the detector array (close to pixel 512);
  - asymmetrical ghosts, which are signals mirrored at some arbitrary detector pixels.
- out–of–band straylight on the PMDs (induced by radiation outside the wavelength range of the array detectors).

The uniform straylight may be measured in the flight model (FM) by reserving specific detector pixel regions for straylight measurements (one region on the short wavelength end of the array detector spectrum area, the other on the long wavelength end). This has been done for channel 1 and 2. These detector pixels were not available for the BBM pre–flight calibration; they were assigned 'after the fact'. For channel 3 and 4, the short wavelength end of the detector arrays receive no signal higher than noise level. In addition, a relative level of uniform straylight w.r.t. the averaged signal level was determined during pre–flight FM calibration.

The ghost characteristics of the BBM have been evaluated thoroughly [A4]. The 3 characteristics of a ghost are (1) position of the centre pixel in each channel about which ghost signals are mirrored; (2) the efficiency (peak–to–peak ratio of ghost to parent) of the mirrored images; and (3) a "focussing factor" (necessary because mirrored images are out of focus with the detector arrays, and parent images are not).

A number of modifications to the FM hardware and software were designed to reduce the unacceptable straylight levels noted in the BBM characterisation. As a result, the straylight behaviour of GOME FM and FSM has improved in several ways:

- asymmetric ghosts have disappeared (tilt changes to the gratings);
- Uniform and ghost straylight levels were reduced (use of anti–reflection coatings);
- relative straylight at the short wavelength side of channel 2 was reduced (change of the channel separation between channel 1 and 2);
- removal of important light leakage sources (improvement of internal baffling).

Despite these improvements, a correction algorithm for straylight is still required, at least for channels 1 and 2 of GOME FM and FSM.

A completely new discovery during the GOME FM calibration (see [A5]) was the detection of out–of–band straylight on the PMDs. All three PMDs receive light above 800 nm, and one would expect this effect to be worst for PMD 1 (which has the lowest signal level in general). However, in orbit it was noticed that PMD 3 measured more signal than expected, while PMD 1 showed negligible straylight effect. PMD 3 is especially sensitive to wavelengths above 790 nm, where the channel 4

spectral region ends. Because there are no means in GOME to measure light above 790 nm, the correction factor for the PMD signals have to be estimated in some fashion from the expected spectral distribution of light collected by GOME in orbit. This is not constant, and so the straylight ratio for light above 790 nm is also not constant. Furthermore the wavelength ranges beyond 790 nm for which the PMDs are sensitive are not known.

Using sun nadir and PMD measurements, plus that part of the polarisation correction algorithm which calculates the fractional polarisation values for the PMD regions (see section 2.9.2.1 on page 36), the straylight correction factor may be set so that the fractional polarisation values yield 0.5 (expected value for unpolarised sunlight).

2.5.2 Algorithm Description

The following steps are applied for the straylight correction of scanning and sun calibration data packets (after subtraction of leakage current and PPG corrections):

- normalisation of all signals by division with the integration time (yields signal fluxes);
- uniform straylight contributions
  - an averaged signal flux for all detector arrays is calculated;
  - the averaged signal fluxes per detector array are multiplied with the relative uniform straylight level values from pre–flight calibration.
- ghost contributions
  - mirroring of the measured signal flux around the pixel centres for each ghost;
  - multiplication of the mirrored signal fluxes with the efficiencies of the corresponding ghost ;
  - smoothing of these mirrored and reduced signal fluxes with the corresponding focussing factors.
- array detector straylight correction
  - addition of the uniform and all ghost straylight contributions for each channel;
  - subtraction of the resulting absolute straylight flux multiplied by the integration time from the measured signal for each channel;
- PMD straylight correction
  - multiplication of all PMD signals with their corresponding correction factors.

A possible approach to improve straylight correction factors for the PMD signals will be:

- By means of a literature sun reference spectrum, the radiation response information of the PMDs and the correction factors for the sun measurements, the cut–of–band wavelength regions of the PMDs may be determined.
- By means of using expected spectral distributions for several earthshine situations (ocean, ice, desert, etc.) with respect to the spectral distribution of the literature sun reference spectrum as used in previous step, the calculated correction factors may be modified according to the different distribution of light between the sun reference and the earthshine spectra. This would yield a set of correction factors for each defined earthshine situation.

2.6 Spectral Calibration

The objective of the spectral calibration is to assign a wavelength to each individual GOME detector pixel during the flight of the sensor. For this calibration, two steps are required:

1. *calculation* of the spectral calibration parameters from the calibration lamp measurements; this task again is separated into two steps:
   a. determination of the pixel number centre of the spectral lines and
   b. fitting of a fourth order polynomial through these pixel number/wavelength pairs;
2. *application* of these parameters to scanning and calibration data packets.

The following algorithm description shows how this is done. For the spectral calibration no pre–flight calibration data is needed, but some auxiliary information is useful – the estimated pixel positions and their precise wavelength values of the PtCrNe–lamp lines in the wavelength range of GOME. It is possible to determine this information even after launch from the corresponding literature available on this subject (PtCrNe lamp lines[A13], NIST atlas [A16], CRC Handbook [A17]).

2.6.1 Notations

The following definitions are used:

S_i	signal of detector pixel i [BU/s]
N	sum of all signals which make up the spectral line [BU/s]
n_i	number of detector pixel i [–]
n_1 or n_2	certain numbers of a detector pixel [–]
$\overline{n}$	centroid of a spectral line as a broken number of detector pixels [–]
λ_i	wavelength of detector pixel i [nm]
Var	variance of a spectral line as a broken number of detector pixels [–]
σ	standard deviation of a spectral line as a broken number of detector pixels [–]
$Skew$	skewness of a spectral line [–]
$FWHM$	full width half maximum of a spectral line as a broken number of detector pixels [–]
a_i	coefficients of the spectral calibration polynomial for i = 0,...,4 [1/nm]

2.6.2 Algorithm Description

Figure 7 shows a typical line which might have been measured with GOME. One can regard such a data set as a statistical distribution, and use statistical methods to solve the pixel number positioning of the spectral lines, as discussed in [A10].

In this application, the size N of the statistical sample is given by:

$$N = \sum_{i=n_1}^{n_2} S_i \tag{10}$$

where S_i is the measured signal on detector pixel i and n_1 and n_2 are the first and last detector pixels which cover the spectral line. One can then calculate the moments of the statistical signal distribution, and these yield some characteristic information about the spectral line. They are:

1. the mean value

$$\overline{n} = \frac{1}{N} \cdot \sum_{i=n_1}^{n_2} i \cdot S_i \tag{11}$$

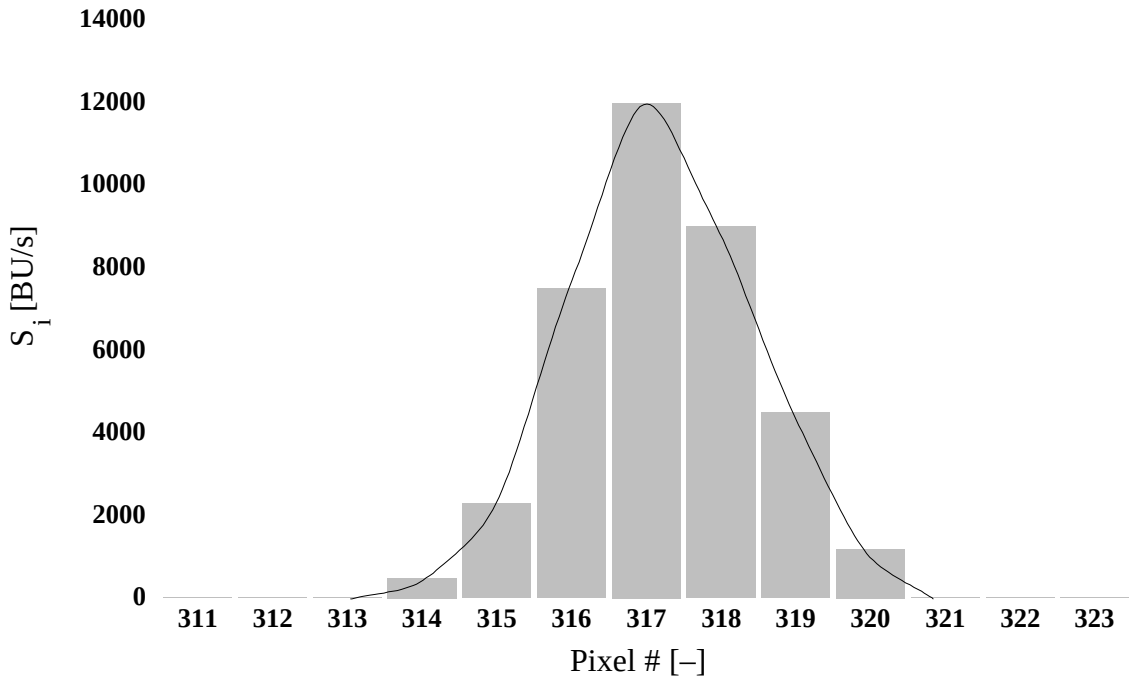

Figure 7: Typical PtCrNe lamp line

2. the variance and standard deviation

$$Var = \frac{1}{N-1} \cdot \sum_{i=n_1}^{n_2} (i - \bar{n})^2 \cdot S_i \tag{12}$$

$$\sigma = \sqrt{Var} \tag{13}$$

3. the skewness

$$Skew = \frac{1}{N} \cdot \sum_{i=n_1}^{n_2} \left(\frac{i - \bar{n}}{\sigma} \right)^3 \cdot S_i \tag{14}$$

For the positioning of spectral lines, the above statistical properties can be interpreted as follows:

1. The mean value gives the centroid of the spectral line, provided the other statistical properties fulfill certain prerequisites. The mean value is therefore the desired pixel centre result.

2. According to [A10] the full width half maximum (*FWHM*) of the line may be calculated as follows from the standard deviation σ:

$$FWHM = \sqrt{8 \cdot ln(2)} \cdot \sigma \tag{15}$$

3. While the mean and the standard deviation are dimensional quantities (that is, they have the same units as the measurement variables), the skewness (as defined here) is a non–dimensional number that characterises only the shape of the distribution, in our case the degree of asymmetry around the mean position of the spectral line. A positive (negative) value of skewness signifies a distribution with an asymmetry tail extending out towards more positive (negative) values. For formula (14) to be meaningful we need to have some idea of the standard deviation of this skewness statistic as an estimator of the skewness of the underlying distribution. For the

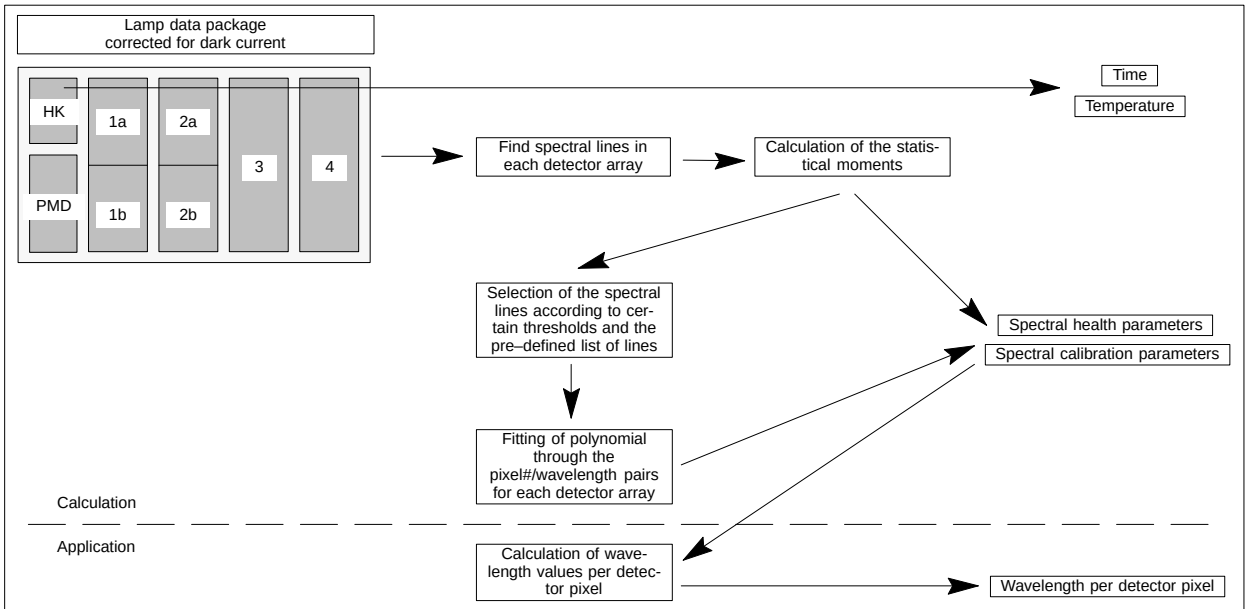

Figure 8: Spectral Calibration Algorithm

idealised case of a normal distribution, the standard deviation of formula (14) is approximately $\sqrt{6/N}$. In practice, one should regard the skewness as significant only when it reaches values several times this number (see NO TAG on page 474).

Once these statistics have been computed, one must then select lines (from a known pre–compiled list) for the fitting of a low–order polynomial. Selected lines should meet the following statistical criteria:

- the signal of the centre pixel shall not be below a certain minimum (well above the noise level);
- the *FWHM* shall not be below a certain value, in order to fulfill the Nyquist criteria for the digital recording of analog signals;
- the skewness shall not be larger than a certain value (spectral lines are reasonably symmetric).

Extensive analysis of lamp spectra measured with GOME during the pre–flight measurements has been carried out to establish reasonable values for these criteria; the requirements should be as strict as possible, given that at least seven spectral lines per channel are needed for the polynomial fitting.

For each line, one has a mean pixel number and a known wavelength. For each channel, a fourth order polynomial is fitted to the line–selected pixel–number/wavelength pairs (using the Singular Value Decomposition Algorithm, as described in NO TAG). The fitted coefficients of this polynomial are the spectral calibration parameters for one detector array.

The *application* of the spectral calibration is simply the evaluation of the following formula for each detector pixel of each array:

$$\lambda_i = a_4\, n_i^4 + a_3\, n_i^3 + a_2\, n_i^2 + a_1\, n_i + a_0 \qquad (16)$$

where the a_i are the polynomial coefficients for the specific detector array.

2.6.3 Summary of Algorithm Steps

The following list (see also figure 8) summarises the steps in level 0 to 1 spectral calibration algorithm. [Reference is made to the above formulae].

- Searching for possible spectral lines by shifting a window (9, 7, 5 and 3 detector pixel wide) over the detector pixel readings where the centre pixel has the highest reading and the surrounding pixels are decreasing to the outside of this window;
- Determination of the statistical moments as described in formulae (11), (13) and (14);
- Selection of those lines which fulfill certain thresholds for the signal level of the centre pixel, the FWHM and the skewness, e.g. $S_{centre} > 50$, $FWHM > 1.5$ and $Skew < 0.6$;
- Selection of lines that are defined in the pre–defined list of candidate lines for the spectral calibration (KeyData) out of above selection list;
- fitting of a fourth order polynomial through those pixel number/wavelength pairs;
- calculation of a wavelength value for each detector pixel using the fitted coefficients of the fourth order polynomial.

2.6.4 Algorithm Update (Issue 4)

The statistical criteria used to select lamp lines for fitting have the following operational defaults at the time of writing:

- the normalised signal of the centre pixel should not be less than 50 BU/s for channel 1, and 300 BU/s for the other channels;
- the standard deviation should not be lower than 0.6;
- the *FWHM* should not be less than 1.5;
- the skewness should not exceed 0.6.

2.6.5 Algorithm Update (Issue 5)

Within the framework of algorithm update to version 2.0 (see also the GDP Update Report [A23]), a re–examintation of the calibration data, especially the application of polynomials has been performed. For a brief description, see the GDP Update Report [A23]. Thus, the calibration polynomials mentioned above in Eqn.(16) are

- modified for detector channel # 1 and # 2 to a third–order polynomial instead a fourth order one,
- unchanged for detector channel # 3 and # 4.

Additionally, the wavelength calibration has been modified by the application of the results of cross–correlation algorithm, which is described in the following section. The modification is implemented in the following manner:

- Calculate the polynomials for wavelength calibration as before,
- Calculate once off–line correction coefficients for the polynomial coefficients determined in the first step by application of cross–correlation to the wavelength calibration. The cross–correlation calibration is described in the following section.
- Add the correction factors on the polynomial coefficients and apply these corrected coefficients according Eqn. to compute the wavelength scale.

2.7 Wavelength calibration by Cross–Correlation

A possible failure of the onboard calibration spectral lamp requires an additional wavelength calibration method. Caspar and Chance initiated to apply this method to the wavelength calibration of GOME spectra. The goal of the method is to calibrate a GOME spectrum using a well–calibrated reference spectrum (template). It is actually possible to determine shifts and squeezes of the wavelength scale by cross correlation. However, the dominant influence originates from the shift and that will be determined solely. The implementation of this algorithm is described briefly. A more detailed discussion can be found in [A19], [A21], and [A23].

2.7.1 Notation

The following definitions are used:

c	Cross–correlation function
g, G	GOME sunshine or earthshine measurement spectrum and its Fourier transform
t, T	Template or reference spectrum and its Fourier transform
λ, λ', $\lambda*$	Wavelength, convolution wavelength, and finally determined wavelength
k	Wavenumber corresponding to wavelength
N	Number of data points
σ_g, σ_g	RMS of the GOME and template spectrum, respectively
ε	Mean error of cross–correlation
h	Peak height maximum of cross–correlation function
σ_a	Root mean square of cross–correlation function
B	half maximum point of Fourier transform of cross correlation
a_i	Coefficients of the spectral calibration polynomial for i = 0,...,4 [1/nm]

2.7.2 Principle of Cross–correlation

The principle of cross–correlation used for astrophysical applications can be found in a publication of Tonry and Davis [A22]. In the following, a function g and G is related to a GOME spectrum, while a function t or T is related to the reference spectrum, The definitions are valid both for GOME solar and earthshine measurements.

The discrete cross–correlation $c(\lambda)$ is defined by:

$$c(\lambda) = g \times t(\lambda) = \frac{1}{N\sigma_g\sigma_t} \sum_{\lambda'} g(\lambda')\, t(\lambda' - \lambda), \tag{17}$$

where the number of data points is denoted by N. Taking the discrete Fourier–transform of the cross–correlation leads to:

$$C(k) = \sum_{\lambda} c(\lambda)\, \exp(-2\pi i k\lambda/N), \tag{18}$$

and the Fourier–transform of the cross–correlation writes:

$$C(k) = \frac{1}{N\sigma_g\sigma_t}\, G(k)\, T^*(k), \tag{19}$$

where σ_i is the *rms* of the GOME (i=g) and the reference (i=t) spectrum:

$$\sigma_i^2 = \frac{1}{N} \sum_{\lambda} g(\lambda)^2 \qquad i = g, t. \tag{20}$$

Note that the complex conjugate of *T(k)* has been taken in Eqn.(19).

As a result, the cross–correlation function contains the overlap of information of the correlated spectra. Simply spoken, the overlap can be separated in abscissa and ordinate contributions, e.g. wavelength and intensity. Thus, from the correlated patterns of the spectra the deviation in wavelength between both spectra is found by the shifted position of the maximum peak of the cross–correlation function.

In the special case of a cross–correlation of a function with itself (auto–correlation), the function is symmetrically centered around $\lambda=0$. The cross–correlation is not symmetric in general, but can be approximately expressed in a symmetric and asymmetric contribution. This fact can be used to derive the mean errors of the cross–correlation function and the shift within. For a detailed derivation see in [A22]. Finally, the mean error of the cross–correlation function is given by:

$$\epsilon = \frac{1}{4}\frac{N}{2B}\frac{1}{1+r} \tag{21}$$

where the parameter r is defined by:

$$r = \frac{h}{\sqrt{2}\,\sigma_a}. \tag{22}$$

B is the highest wavenumber, where the Fourier–transform of *c(λ)* has a distinct amplitude, e.g. the half–maximum point. For a more detailed explanation of *B*, see in [A22].

Note that the method of cross–correlation is actually given by the Eqs.(19) and (21), only.

2.7.3 Application of Cross–correlation to GOME Spectra

A Fraunhofer spectrum recorded and compiled by Kurucz et al. [A24] with a spectral resolution of 0.01 nm and an accuracy of 0.001 nm has been selected as a reference spectrum. In order to apply the cross–correlation to calibrate the GOME spectrum, several preliminary steps have to be performed:

- Since the resolution spectrum differs from that of the GOME spectrum, the reference spectrum has to be convolved. A Gaussian function with a half width at 1/e calculated from the slit function is appropriate.
- Different wavelength grids of both the GOME spectrum and the reference spectrum have to be taken into account. Therefore, one spectrum must be resampled so that both spectra are available on the same wavelength grid. Here, the wavelength grid of the Fraunhofer spectrum is taken as reference grid.
- Usually, both spectra contain different continuum information which must be corrected for. In order to remove these structures, a linear function or a polynomial of order 2 and higher can be used by means of least squares fitting.
- In order to perform the Fourier–transformation, both spectra need to be apodised.
- The signal–to–noise ratio and the mean error can be reduced by Fourier–filtering.

After the preparing steps and application of the cross–correlation, a peak finding routine which retrieves the peak maximum height position, e.g. the wavelength shift, has to be applied. The parameter *B* has to be determined from the unfiltered Fourier–transform of the cross–correlation function. Subsequently, the mean error can be determined. Usually, *B* is set to a constant value. Here, it is taken from a preparing, representative test run which was performed off–line.

Since the wavelength shift differs within each channel for different parts of the channel, the above described procedure is performed for several (15–20) wavelength ranges (windows) equally sized. It turned out that a window size of around 50 detector pixels is the most appropriate size in the sense of a minimum mean error. The retrieved values of all windows (of one channel) are combined to a polynomical function by means of weighted least squares fitting. The wavelength calibration for each channel is finally given by these polynomial functions coadded to the calibration polynomials which are retrieved by the method described in the previous section.

2.7.4 Algorithm Steps

The algorithm are divided into steps which are performed *off–line* and *on–line* during the data processing. The cross–correlation is applied off–line to provide polynomial coefficients as an additional correction to the polynomial coefficients of the existing wavelength calibration which is described in the previous section. The *off–line* computation which has been performed once, is characterised by following steps:

- Definition of small windows within each measuremnt channel for the application of cross–correlation,
- Application of preparational steps:
  - Convolution of reference spectra appropriate to the spectral resolution of the GOME spectrum,
  - Resampling of convolved reference spectrum to the wavelength grid of GOME spectrum,
  - Removal of continuum information,
  - Apodisation of the spectra as preparation for Fourier–transform, and
  - Application of Fourier–filtering.
- Retrieval of wavelength shift for each defined window by utilisation of cross–correlation, and
- Computation of polynomial coefficients for each measurement channel from the wavelength shift results by weighted means of least squares fitting.

The computed polynomial coefficients Δa_i are stored and read into the data processing unit. Their values are added to the values of the coefficient set retrieved by spectral calibration based on spectral lines of the onboard lamp, so that the wavelength λ_j, as defined in Eqn.(16), is modified to

$$\lambda_j^* = \lambda_j + \sum_{i=0}^{4} \Delta a_i \cdot \lambda_j^{\,i} \tag{23}$$

where a_i are the coefficient set retrieved from spectral lamp information for the spectral calibration.

2.7.5 Algorithm update (Issue 5B)

Due to the up–coming uncertainty of the wavelength calibration by hardware failures in the years 2001 and 2002, it was decided to implement the cross–correlation calibration algorithm into the Level 1 extraction software. All required specifications are given in part B of the appendix.

- The algorithm steps, as described in section 2.7.4, remain unchanged in the Level 0–1 processing and are added to the Level 0–1 extractor.
- The cross–correlation calibration as described in section 2.7.2 has been implemented as option into the Level 1 extraction software.

2.8 Radiometric Calibration

The objective of the radiometric calibration is to transform the 16–bit binary units (BU) of the detector pixel readouts into calibrated radiances (photons s^{-1} m^{-2} nm^{-1} sr^{-1}). There are two steps necessary to accomplish this task:

1. an intensity calibration using the GOME sun measurements taken via the diffuser in the calibration unit;
2. the application of a polarisation correction to the solar and scanning data signals.

This section gives a description of the first step; of particular importance is the use of parameters determined during the pre–flight calibration activities performed at TPD/TNO. The polarisation correction algorithm (PCA) is described in the next section.

2.8.1 Notation

The following definitions are used:

λ_i or λ_j	wavelength [nm] of detector pixel *i* or *j*; the indices *i* or *j* always identify individual detector pixels; whenever a detector pixel is mentioned in the following text, this assumes that the appropriate wavelength calibration parameters are known and the corresponding wavelengths assigned.
$S_{i(readout)}$	signal of detector pixel *i* [BU]
t	integration time of a detector pixel [s]
S_i	normalised signal of detector pixel *i* [BU/s]
$(S_{Sun,i})_{pol}$	normalised signal of detector pixel *i* during a sun calibration measurement before application of the polarisation correction [BU/s]
$S_{Sun,i}$	normalised signal of detector pixel *i* during a sun calibration measurement [BU/s]
$S_{Earth,j}$	normalised signal of detector pixel *j* during a scanning measurement [BU/s]
$(S_{Earth,i})_{pol}$	normalised signal of detector pixel *i* during a scanning measurement before application of the polarisation correction [BU/s]
$S_{Earth,i}$	normalised signal of detector pixel *i* during a scanning measurement [BU/s]
$I_{Sun}(\lambda_i)$	irradiance transmitted from the sun at wavelength λ_i [photons s^{-1} m^{-2} nm^{-1}]
$I_{Earth}(\lambda_i)$	radiance back–scattered from the earth at wavelength λ_i [photons s^{-1} m^{-2} nm^{-1} sr^{-1}]
σ	scan mirror angle [degrees]
$H^*(\lambda_i,\sigma)$	radiance sensitivity function at wavelength λ_i and scan mirror angle σ [(BU s^{-1})/(photons s^{-1} m^{-2} nm^{-1} sr^{-1})]. All such quantities derived from the pre–flight calibration and characterisation activity (at ESTEC and TPD) are *known* inputs, and are marked with an asterisk.
$BSDF^*_{CU}(\lambda_i,Az, El)$	bi–directional scattering distribution function of the diffuser at wavelength λ_i and at specified azimuth and elevation angles [sr^{-1}]
$BSDF_{Earth}(\lambda_i)$	absorption spectrum of the earth's atmosphere at wavelength λ_i [sr^{-1}]
c_i	polarisation correction factor for the scanning measurements [–]
$c^*_{CU}(\lambda_i)$	polarisation correction factor (known) for the sun calibration measurements at wavelength λ_i [–]
$Az(a)$	$BSDF^*_{CU}$ dependency from the azimuth *a*
$El(\varepsilon)$	$BSDF^*_{CU}$ dependency from the elevation angle ε
$w(\lambda)$	$BSDF^*_{CU}$ dependency from wavelength λ.

2.8.2 Algorithm Description

Figure 9 shows the schematic illumination configuration for the intensity calibration. During the sun calibration measurement the scan mirror of GOME is illuminated with sunlight via the diffuser

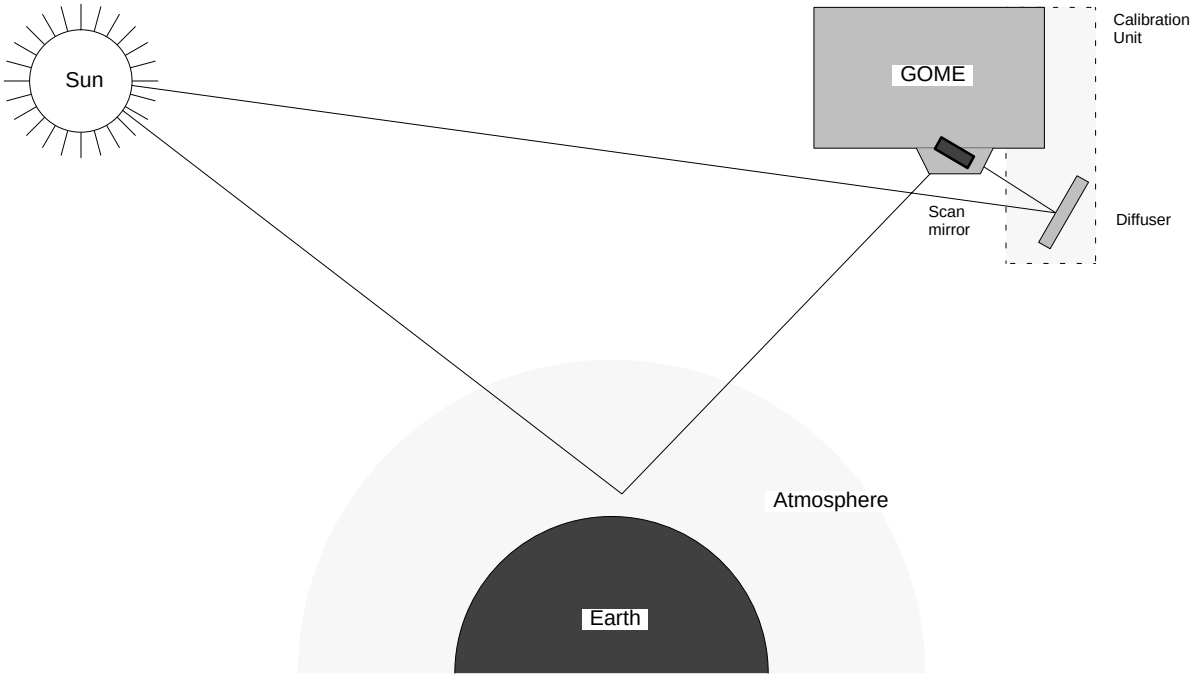

Figure 9: Schematic Illumination Configuration for the Intensity Calibration

and during the normal scanning operation the back–scattered light of the earth's atmosphere is entering GOME directly.

For each detector pixel, the signal readout (corrected for leakage current, pixel–to–pixel gain and straylight) is divided by the integration time to yield a normalised unit of binary units per second:

$$S_i = \frac{S_{i(readout)}}{t} \left[\frac{BU}{s}\right] \tag{24}$$

For signals of a *solar* calibration measurement, the following calibration formula is valid (includes the physical units of each term):

$$S_{Sun,i} = I_{Sun}(\lambda_i) \cdot H^*(\lambda_i, \sigma) \cdot BSDF^*_{CU}(\lambda_i, Az, El) \tag{25}$$

$$\left[\frac{\mathrm{BU}}{\mathrm{s}}\right] = \left[\frac{\mathrm{photons}}{\mathrm{s\ m^2\ nm}}\right] \cdot \left[\frac{\mathrm{BU\ s^{-1}}}{\mathrm{photons\ s^{-1}\ m^{-2}\ nm^{-1}\ sr^{-1}}}\right] \cdot \left[\mathrm{sr^{-1}}\right]$$

where $S_{Sun,i}$ is the measured signal on the detector arrays, $I_{Sun}(\lambda_i)$ the irradiance transmitted from the sun at the wavelength λ_i, $H^*(\lambda_i,\sigma)$ the radiance sensitivity function at the wavelength λ_i and scan mirror angle σ, and $BSDF^*_{CU}(\lambda_i,Az,El)$ the bi–directional scattering distribution function of the diffuser in the calibration unit, which is dependent on the wavelength, the azimuth and the elevation of the sunlight on the diffuser. For sun calibration measurements σ will be taken as nadir, because the scan mirror dependence is included in the BSDF.

For signals of a *scanning* measurement the corresponding formula is:

$$S_{Earth,j} = I_{Sun}(\lambda_j) \cdot H^*(\lambda_j, \sigma) \cdot BSDF_{Earth}(\lambda_j) \tag{26}$$

where $S_{Earth,j}$ is the measured signal on the detector arrays and $BSDF_{Earth}(\lambda_j)$ the wavelength dependent absorption spectrum of the earth's atmosphere. Indices (i, j) in formulae (25) and (26) are

not the same because the measurements are taken at different orbital positions, for which temperatures at the pre–disperse prism will (in general) not be the same, and consequently the two sets of spectral calibration parameters will also be unequal.

This implies an interpolation of one measured spectrum to the other. One orbit will produce several nadir scanning spectral calibration parameter sets, but only one solar spectral calibration parameter set. Thus scanning measurements of one orbit are interpolated to the wavelength grid of the corresponding sun calibration measurement; it is convenient to work with one wavelength grid per orbit. One can re–cast (19) on the same grid as (18):

$$S_{Earth,i} = I_{Sun}(\lambda_i) \cdot H^*(\lambda_i, \sigma) \cdot BSDF_{Earth}(\lambda_i) \tag{27}$$

Division of formulae (25) and (27) yields the atmosphere's absorption spectrum which is so important for the geophysical retrieval of trace gas constituents:

$$BSDF_{Earth}(\lambda_i) = S_{Earth,i} \cdot \frac{BSDF^*_{CU}(\lambda_i, Az, El)}{S_{Sun,i}} \tag{28}$$

For the last factor (BSDF divided by sun signals), a mean value is computed from several sun observations; this value is the 'mean sun response'. There will be 30 seconds of sun calibration measurements during one orbit out of the 14 orbits per day. To ensure that the sun is completely in the field of view of the diffuser, the first and last 6 seconds of these measurements are ignored. This leaves 3 measurements in channels 1a to 2a (short–wave region) and 12 measurements in channels 2b to 4 (longer–wave region) for the computation of this mean GOME sun spectrum.

Besides the atmosphere's absorption spectrum, the irradiance of the sunlight is part of the Level 1 Extracted Data; it is calculated as follows:

$$I_{Sun}(\lambda_i) = \frac{S_{Sun,i}}{H^*(\lambda_i, \sigma) \cdot BSDF^*_{CU}(\lambda_i, Az, El)} \tag{29}$$

The radiances of the back–scattered light from the earth's atmosphere will not be given directly in the Level 1 Data product; this is largely due to limitations on the accuracies of the pre–flight calibration parameters. In general, the radiance/irradiance calibration yields an accuracy of $\pm$5% for the values of $H^*(\lambda_i)$, whereas the calibration of the diffuser yields an accuracy of $\pm$1% for the values of $BSDF^*_{CU}(\lambda_i,Az,El)$. The calibrated radiances of the earth–shine measurements can of course be recovered quickly in the level 1 extraction function; the formula is:

$$I_{Earth}(\lambda_i) = I_{Sun}(\lambda_i) \cdot BSDF_{Earth}(\lambda_i) \tag{30}$$

So far, we have assumed that the radiances $I_{Sun}(\lambda_i) \cdot BSDF^*_{CU}(\lambda_i,Az,El)$ and $I_{Earth}(\lambda_i)$ at the entrance slit of GOME are unpolarised. However, both the atmosphere and the mirror system in the calibration unit polarise the incident solar light. Thus, correction factors must be applied to the signals of the scanning and the sun measurements before the above formulae (28) and (29) can be used. We may write:

$$S_{Earth,i} = \left(S_{Earth,i}\right)_{pol} \cdot c_i \tag{31}$$

$$S_{Sun,i} = \left(S_{Sun,i}\right)_{pol} \cdot c^*_{CU}(\lambda_i) \tag{32}$$

where the atmospheric correction factor c_i is calculated according to the algorithm description in the next section, and the diffuser correction factor $c^*_{CU}(\lambda_i)$ is an input from pre–flight calibration phase.

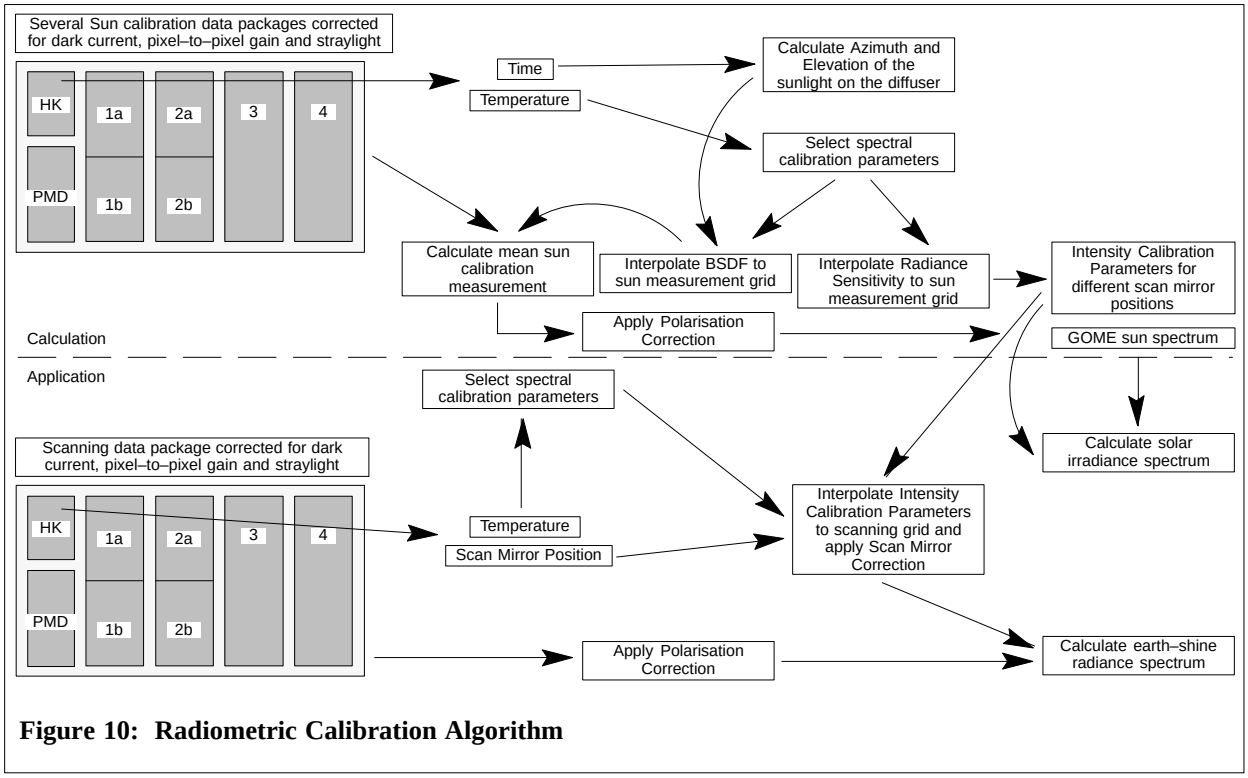

Figure 10: Radiometric Calibration Algorithm

2.8.3 Summary of Algorithm Steps

The following list (and the diagram in figure 10) summarises the steps involved in the radiometric calibration algorithm. [Refer to above equations where necessary].

- Calculate the azimuth and elevation angles of the sunlight at the diffuser, using an ephemeris programme (originally VENI, now ERS propagator);
- Calculate the BSDF of the diffuser for the wavelength grid of the sun calibration measurement and calculate its values for the azimuth and elevation angles (the functional dependence of the BSDF on these variables is known from the pre–flight diffuser characterisation);
- Interpolate the instrument response function to the wavelength grid of the sun calibration measurement. This yields the intensity calibration parameters to be included in the Level 1 Data product;
- Calculate the mean sun calibration measurement;
- Apply the polarisation correction factor for the calibration unit to the mean sun calibration measurement. This yields the GOME sun spectrum to be included in the Level 1 Data product. [Note – the calibration unit of GOME was not designed to polarise incoming light, and pre–flight measurements showed insignificant residual polarisation; the correction factor in (32) is taken as unity in the operational algorithm];
- Calculate the solar irradiance spectrum (29).
- Interpolate Intensity Calibration Parameters to the scanning measurement and apply the scan mirror position correction;
- Apply the polarisation correction to the scanning measurement (see next section) (31);
- Calculate the earth–shine radiance spectrum (30);

2.8.4 Algorithm Update (Issue 4)

Recent changes to the algorithm are as follows (these are the current operational defaults). First, define $M_{Sun}(\lambda_i)$, the mean sun reference spectrum at wavelength λ_i [BU s^{-1} sr^{-1}]. This is the mean value of all sun measurements divided by their integration times and appropriate BSDF entries :

$$M_{Sun}(\lambda_i) = = \frac{1}{(n)} \cdot \sum_{i=0}^{n} \frac{S_{Sun,i}}{BSDF^{*}_{CU}(\lambda_i, Az, El)} \tag{33}$$

where n is the number of sun measurements. Note that the azimuth and elevation at the diffuser are different for each measurement.

In the revised algorithm, the radiance/irradiance results are now given by:

$$I_{Earth}(\lambda_i) = \frac{S_{Earth,i}}{H^{*}(\lambda_i, \sigma)} \tag{34}$$

for the calibrated radiance of the earth–shine measurement, and

$$I_{Sun}(\lambda_i) = \frac{M_{Sun}(\lambda_i)}{H^{*}(\lambda_i, \sigma)} \tag{35}$$

for the solar irradiance spectrum.

The summary of the algorithm steps now reads:

- Calculate the azimuth and elevation angles of the sunlight on the diffuser using the ERS orbit propagator for each sun spectrum. Check if the angles are inside given limits;
- Calculate the BSDF of the diffuser for the wavelength grid of each sun calibration measurement and calculate its values for the azimuth and elevation angles (as before);
- Divide each sun spectrum by BSDF and integration time;
- Calculate the mean sun reference spectrum (33) to be included in the Level 1 Data product;
- Interpolate the instrument response function to the wavelength grid of the sun calibration measurement and apply the scan mirror position correction for the effective scan mirror position of the sun measurements;
- Calculate the solar irradiance spectrum (35);
- Interpolate Intensity Calibration Parameters to the scanning measurement and apply the scan mirror position correction;
- Calculate the calibrated earth–shine radiance spectrum (34);

2.8.5 Algorithm Update (Issue 5)

BSDF

The $BSDF^{*}_{CU}(\lambda_i,Az,El)$ is implemented using following function:

$$BSDF^{*}_{CU}(\lambda_i, Az, El) = BSDF_0 \cdot Az(\alpha) \cdot El(\varepsilon) \cdot w(\lambda_i) \tag{36}$$

Here, the function $Az(a)$ describes the dependency from the azimuth a, $El(\varepsilon)$ from the elevation angle ε, and $w(\lambda)$ from the wavelength λ. These functions and the constant $BSDF_0$ were derived by TPD [A25]:

$$Az(\alpha) = 1 - c_a \cdot \alpha^2 \tag{37}$$

$$El(\epsilon) = 1 + c_e \cdot \varepsilon \tag{38}$$

$$w(\lambda) = \sum_{i=0}^{7} c_i \cdot \left(\frac{\lambda - \lambda_0}{\lambda_0} \right)^i \tag{39}$$

The constants c_a, c_e and c_i (with λ_0 = 500 nm) have been determined by TPD during pre–flight calibration [A25].

An improved algorithm with new parameters for azimuth correction was developed by DLR in the framework of the ESA/ESRIN–project GOME Support. The correction can be applied as an option by the extractor software of GDP Level 0–to–1 using

$$Az(\alpha) = 1 + c_{a1} \cdot \alpha - c_{a2} \cdot \alpha^2 \tag{40}$$

instead of eqn. (37). More details are given in [A23] and [A26].

Albedo

The extractor software of GDP Level 0–to–1 provides the calculation of albedo instead radiance, if the option *–a* is selected.Usually, the calculation of albedo ϱ is performed by division of earthshine radiance through sun irradiance spectrum which are both given in eqns.(34) and (35), respectively:

$$\rho = \frac{I_{Earth}(\lambda_i)}{I_{Sun}(\lambda_i)} \tag{41}$$

Unitl now, for the implementation of the albedo computation it was assumed that for both spectra, earthshine and sunshine, the scan mirror position has been equal during the measurement. Thus, the the same value for the radiance sensitivity function H^* has been taken. In [A23] , initiated by C. Tanzi [A20], it is outlined that the assumption of an equal mirror position for both, earthshine and sun, spectra does not hold. Thus, the albedo computation has to be corrected by taking into account the scan mirror dependency of the radiance sensitivity function H^*:

$$\rho = \frac{I_{Earth}(\lambda_i)}{I_{Sun}(\lambda_i)} = \frac{S_{Earth} \cdot H^*_{Sun}(\lambda_i, \sigma)}{M_{Sun}(\lambda_i) \cdot H^*_{Earth}(\lambda_i, \sigma)} \tag{42}$$

The albedo is now calculated according to this equation.

2.8.6 Algorithm Update (Issue 6)

BSDF

The optional BSDF correction in the extractor software of GDP Level 0–to–1, as described under Algorithm Update Issue 5, now has been replaced by a scheme developed by DLR in the framework of the ESA/ESRIN project GOME–CHEOPS. This scheme uses lookup tables, instead of an analytical parametrisation. For details, see section C.2 . The lookup tables can be found on the same external degradation file as used before (see also section 2.10.3).

2.9 Polarisation Correction

Since the intensity part of the radiometric calibration uses unpolarised solar light[1] , the *polarisation correction* has to transform the measured signal (with fractional polarisation p) to an unpolarised light signal (for which p=0.5). The following algorithm description shows how this is done using the parameters determined during the pre–flight calibration activities performed by TPD/TNO.

2.9.1 Definitions and Preliminary Results

The following definitions are used in the text:

λ_i or λ_j	wavelength [nm] of detector pixel i or j; the indices i or j always identify individual detector pixels; whenever a detector pixel is mentioned in the following text, it is assumed that the wavelength corresponding to the pixel is known for the current set of wavelength calibration parameters.
I	total intensity of the incoming light [W m^{-2} nm^{-1} sr^{-1}]
I_p	intensity of the parallel polarised light [W m^{-2} nm^{-1} sr^{-1}]
$I_{p,i}$	intensity of the parallel polarised light at wavelength λ_i [W m^{-2} nm^{-1} sr^{-1}]
I_s	intensity of the perpendicular polarised light [W m^{-2} nm^{-1} sr^{-1}]
$I_{s,i}$	intensity of the perpendicular polarised light at wavelength λ_i [W m^{-2} nm^{-1} sr^{-1}]
I_{PMD}	total intensity of the light reaching a PMD [W m^{-2} nm^{-1} sr^{-1}]
$I_{p,PMD}$	intensity of the parallel polarised light reaching a PMD [W m^{-2} nm^{-1} sr^{-1}]
p	fractional polarisation value [–]
p_i	fractional polarisation value at wavelength λ_i [–]
p_{21}	fractional polarisation value for the ”Array 2/PMD 1”–combination [–]; other combinations are named accordingly
$n(p)$	weighted mean pixel number of a certain fractional polarisation value [–]
$S_i^{readout}$	signal of detector pixel i [BU]
t	integration time of a detector pixel [s]
S_i or S_j	normalised signal of detector pixel i or j [BU/s]
$S_{corr,i}$	normalised signal of detector pixel i for p=0.5 [BU/s]
$S_{PMDj,\,i}^{readout}$	signal of a PMD [BU]
S_{PMD}	normalised signal of a PMD [BU/s]
m	number of data packages corresponding to a normalised PMD signal [–]
$S_{s,i}$	signal of detector pixel i taken during during a measurement for a completely polarised beam in the perpendicular (s) position [BU/s]
$S_{p,i}$	the same as above for the (p) parallel position [BU/s]
$a_{p,i}$ or $b_{p,j}$	parallel polarisation sensitivity of detector pixel i or j [–]
$a_{s,i}$ or $b_{s,j}$	perpendicular polarisation sensitivity of detector pixel i or j [–]
$d_{p,PMD}$	parallel polarisation sensitivity of a PMD [–]
$d_{p,i}$	parallel polarisation sensitivity of a ”virtual” PMD pixel [–]
$\eta_{nadir,i}^{*}$	polarisation sensitivity ratio of detector pixel i [–]; these values are given for the nadir position of the scan mirror (all such quantities determined during the pre–flight calibration and characterisation activity are regarded as input information, and are annotated accordingly with an asterisk)

1 This is true after the polarising effect of the mirrors in the calibration unit has been taken into account (see formula (32)). However, it has been shown that this effect is very small for the GOME CU, and it has been neglected in this version of the polarisation correction algorithm.

σ	scan mirror angle [degrees]
$\chi^*(\lambda_i,\alpha)$	polarisation sensitivity correction factor for wavelength λ_i and scan mirror angle position α [–]
η_i	polarisation sensitivity ratio of detector pixel i [–]; it is assumed that the appropriate scan mirror angle correction factor has been applied
$\eta_{a,i}$ or $\eta_{b,j}$	polarisation sensitivity ratio of detector pixel i or j [–] in channels a and b; this notation is necessary to distinguish η^* in an overlapping region
ξ^*_i	PMD "virtual" polarisation sensitivity ratio of detector pixel i [–]; these values are given for each PMD
$\varkappa^*(\lambda_i,T)$	PMD "virtual" polarisation sensitivity ratio correction factor for the wavelength λ_i and the temperature T measured at the PMDs [–]; not used in the first version of the polarisation correction algorithm
$R_{u,i}$	signal ratio in an overlapping region taken for unpolarised light [–]; these may be derived from in–flight solar measurements or from pre–flight observations with an unpolarised white light source.
$R_{p,i}$	signal ratio in an over–lapping region taken for polarised light [–]
c_i	polarisation correction factor for detector pixel i [–]

The following preliminary results are required:

- The intensity of the incoming light is given by:

$$I = I_p + I_s = p \cdot I + (1–p) \cdot I \tag{43}$$

where I is the total intensity, I_p and I_s are the intensities of the parallel and perpendicular polarised light entering GOME, and p is the fractional polarisation.

- The measured signal of the detector pixels can be written:

$$S_i = a_{p,i} \cdot I_{p,i} + a_{s,i} \cdot I_{s,i} = a_{p,i} \cdot p_i \cdot I_i + a_{s,i} \cdot (1–p_i) \cdot I_i \tag{44}$$

where S_i is the measured signal, p_i the fractional polarisation value of a certain detector pixel and $a_{p,i}$ and $a_{s,i}$ are the wavelength–dependent polarisation sensitivities of the corresponding polarisation directions (p=parallel, s=perpendicular). For the PMDs, the measured signal is given by:

$$S_{PMD} = d_{p,PMD} \cdot I_{p,PMD} = d_{p,PMD} \cdot p \cdot I_{PMD} \tag{45}$$

where S_{PMD} is the measured signal of a certain PMD and $d_{p,PMD}$ is the parallel polarisation sensitivity of the PMD. It is a design feature of the instrument that the perpendicular sensitivity of the PMDs is practically zero.

- Since the absolute radiance of the calibration light source is not known, it is not possible to perform an absolute measurement to determine the parallel and perpendicular polarisation sensitivities independently. A relative measurement between the parallel and perpendicular position was performed using a completely polarised beam. This yields the following quantity (polarisation sensitivity) for each detector pixel:

$$\eta_i^* = \frac{S_{s,i}}{S_{p,i}} = \frac{\left(a_{s,i} \cdot I_{s,i} + a_{p,i} \cdot I_{p,i}\right)_{s–position}}{\left(a_{p,i} \cdot I_{p,i} + a_{s,i} \cdot I_{s,i}\right)_{p–position}} = \frac{a_{s,i} \cdot I_{s,i}}{a_{p,i} \cdot I_{p,i}} = \frac{a_{s,i}}{a_{p,i}} \tag{46}$$

[During the measurement of the light beam in parallel position, the intensity in the perpendicular direction is zero; similarly in perpendicular position, the intensity in the parallel direction is

zero. Moreover, $I_{p,i}$ and $I_{s,i}$ are equal in magnitude, because they are generated by the same beam in different positions].

The primary measurement of η is for the nadir position of the scan mirror during the pre–flight calibration ($\eta^*_{nadir,i}$). However, η is also dependent on the scan mirror angle, and a correction factor for $\eta^*_{nadir,i}$ is applied as follows:

$$\eta_i = \eta^*_{nadir,i} \cdot \chi^*(\lambda_i, \alpha) \tag{47}$$

where $\chi^*(\lambda_i, \alpha)$ is the correction factor for the scan mirror angle α at the wavelength λ_i.

- Measurements through the PMDs using a narrow–filter (”monochromatic”) unpolarised light beam yield the polarisation sensitivity ξ^* per ”virtual” pixel of the PMDs. The nadir scan mirror position was used in these pre–flight measurements. The measurements yield the following PMD ratios:

$$\frac{d_{p,i} \cdot I_{p,i}}{a_{p,i} \cdot I_{p,i} + a_{s,i} \cdot I_{s,i}} = \frac{d_{p,i}}{a_{p,i} + a_{s,i}} = \frac{d_{p,i}}{a_{p,i}(1 + \eta^*_{nadir,i})} \tag{48}$$

where $d_{p,i}$ is the polarisation sensitivity of the PMD and $I_{p,i}$ is equal to $I_{s,i}$, because of the unpolarised light source. In fact, the key data supplied by TPD is the quantity:

$$\xi^*_i = \frac{d_{p,i}}{a_{p,i}} \tag{49}$$

With these definitions, formula (29) can be written:

$$S_{PMD} = \sum^{i} d_{p,i} \cdot I_{p,i} = \sum^{i} d_{p,i} \cdot p \cdot I_i \tag{50}$$

If $I_{p,i}$ was constant over the wavelength range of the corresponding PMD, the following formulae would hold:

$$I_{p,PMD} = \sum^{i} I_{p,i} \qquad \text{and} \qquad d_{p,PMD} = \sum^{i} d_{p,i} \tag{51}$$

The PMD ratios per pixel were measured in ambient air pressure and at a certain temperature T_0. These PMD ratios may also depend on the temperature measured at the PMDs (this changes somewhat in space). A correction factor for ξ^* can be applied as follows:

$$\xi^*_i = \xi^*_{(Air,T_0),i} \cdot \kappa^*(\lambda_i, T) \tag{52}$$

where $\varkappa^*(\lambda_i, T)$ is the correction factor for PMD temperature T at wavelength λ_i. As it has not proved possible to generate these factors in the pre–flight stage, they will be be ignored (set to unity) in the algorithm.

- For each detector pixel, the signal readout is divided by the integration time t to yield a normalised unit for the readouts of binary units per second:

$$S_i = \frac{S^{readout}_i}{t} \left[\frac{BU}{s}\right] \tag{53}$$

PMDs are summed and averaged as follows:

$$S_{PMD} = \frac{1}{1.5 \cdot m} \cdot \sum_{i=1}^{m} \sum_{j=1}^{16} S^{readout}_{PMDj,i} \left[\frac{BU}{s}\right] \tag{54}$$

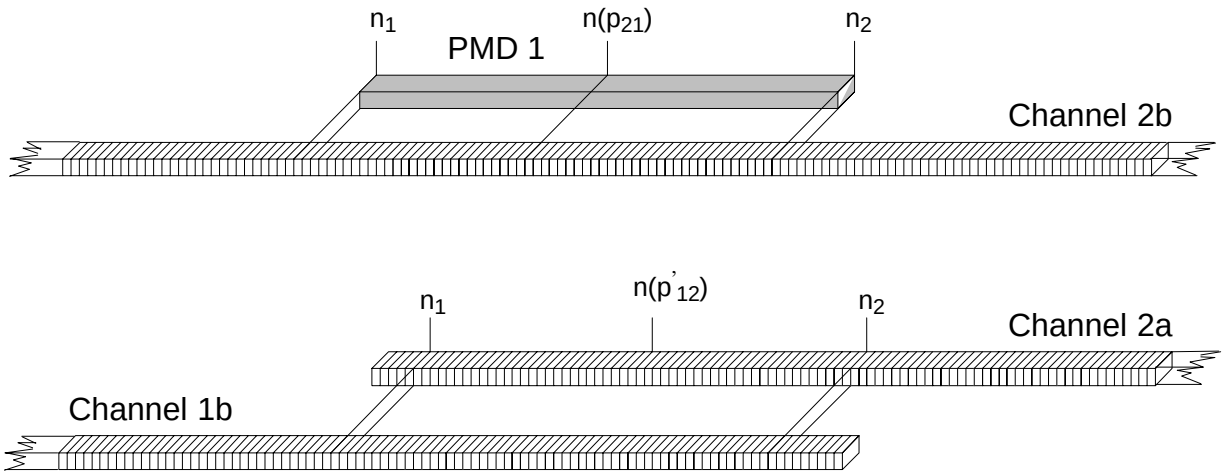

Figure 11: Configuration of the polarisation sensitive Detectors (schematic)

where *m* is the integration time of the channel divided by 1.5 seconds[2] , and $S^{readout}_{PMDj,\,i}$ are the signals of the PMD during this time. *j* runs over the 16 PMD measurements which are performed during one PMD readout (1.5 seconds).

Measurements carried out during the FM calibration have shown that all three PMDs are sensitive to light above 790 nm. Early in–flight solar data showed that straylight appears to be worst in PMD 3 (≈.1% PMD 1, ≈2% PMD 2,≈13% PMD 3). This straylight should be accounted for in the comparison of averaged PMD signals with their ”virtual pixel” equivalents (equation (36)). [see below, and also the discussion in section 2.7.5].

- For the GOME spectral range, there are six wavelength regions which are covered by two detectors having different polarisation sensitivities – these are (a) the three Polarisation Measurement Devices (PMDs) which overlap the wavelength ranges of channels 2, 3 and 4, and (b) the three overlapping regions of the channels themselves. [See figure 11]. *The chief purpose of the (first version of the) Polarisation Correction Algorithm (PCA) is to calculate 6 values of fractional polarisation for the 3 overlapping and 3 PMD wavelength regions. A seventh point on the polarisation curve is supplied from a model calculation.* Well–defined procedures have been established to calculate these fractional polarisation values, and these procedures are described in section 2.7.2.

- Fractional polarisation must be specified at all GOME wavelengths, and the *application* of the Polarisation Correction Algorithm is concerned with (a) the interpolation of *p* from the seven determinable values, and (b) the subsequent generation of correction factors to convert the polarised signals to unpolarised values. The PCA application is dealt with in section 2.7.3, and a summary of the algorithm steps is given in section 2.7.4.

- There are many difficulties involved in this algorithm. These include the pixel–to–pixel variations in fractional polarisation likely to be produced in regions of strong atmospheric absorption such as the O_2 A–band, the straylight problem in PMD 3, and the loss of the second channel overlap region. These points are discussed in section 2.7.5, which also includes a blueprint for a second version of the PCA.

2 e.g. integration time 30 sec. means *m*=20

2.9.2 Calculation of Fractional Polarisation Values

Sub–section 2.7.2.1 deals with the generation of p–values from the PMD–Array overlaps, while 2.7.2.2 describes the generation of p–values from the channel overlap regions. Section 2.7.2.3 deals with the ”seventh point” model calculation of *p* in the Ultra–Violet part of the spectrum.

2.9.2.1 PMD–Array Overlap

The following description concentrates on the evaluation of one fractional polarisation value for a given PMD region, for example with PMD 1 and the corresponding over–lapping channel regions. The other PMDs are treated similarly.

Using formulae (44) we may write the pixel intensity as follows:

$$I_i = \frac{S_i}{a_{p,i} \cdot p + a_{s,i}.(1-p)} \tag{55}$$

We now assume that p is a constant for a given PMD wavelength range. Substituting (41) in the PMD ”virtual pixel” sum (equation (36)) we get:

$$S_{PMD} = \sum^{i} d_{p,i} \cdot p \cdot \frac{S_i}{a_{p,i} \cdot p + a_{s,i}.(1-p)} \tag{56}$$

$$\therefore \; S_{PMD} = \sum^{i} \xi_i^* \cdot p \cdot \frac{S_i}{\eta_i + p(1-\eta_i)} \tag{57}$$

using the definitions (32) and (35) of ξ and η respectively [it is assumed that η is corrected for the nadir–viewing scan mirror angle position Ω]. Equation (43) is the basic result for this p–value; a numerical solution must be applied. There are three such equations (for the 3 PMDs), and care must be taken to include all detector pixel signals for which ξ^*_i is non–zero.

A corresponding result can be written for the solar measurements T_{PMD} and T_i. Since it is known that solar extraterrestrial light is unpolarised (p=1/2), we have:

$$T_{PMD} = \sum^{i} \xi_i^* \cdot \frac{T_i}{\eta_i^{SUN} + 1} \tag{58}$$

where η_i^{SUN} denotes the polarisation sensitivity corrected for the solar position of the scan mirror. There are no unknowns here, so assuming no PMD straylight (beyond 790 nm), *equation (44) for solar measurements should be identically true*. The difference between the two sides of (44) is then a measure of the PMD straylight for an unpolarised light source. In fact, we may define a relative correction factor for these sun measurements :

$$Q_{SUN} = \frac{\left(T_{PMD} - T_{VS}\right)}{T_{PMD}} \tag{59}$$

where T_{VS} is the virtual sum on the right hand side of (44).

Equation (43) for the earthshine measurements can be re–cast with a straylight correction term:

$$\therefore \; S_{PMD} = \sum^{i} \xi_i^* \cdot p \cdot \frac{S_i}{\eta_i + p(1-\eta_i)} + Q_{EARTH} \tag{60}$$

Unfortunately Q_{EARTH} cannot be equated to the (known) solar straylight correction Q_{SUN}, because the polarisation state of the atmosphere in the straylight region is not known. However, assuming the polarisation for a given PMD takes the same (constant) value in the straylight region, we may use the following rough estimate.

$$Q_{EARTH} = 2 \cdot p \cdot \frac{T_{PMD}}{S_{PMD}} \cdot Q_{SUN} \tag{61}$$

This estimate is most appropriate for PMD 3, which has wavelength range adjoining the straylight region beyond 790 nm. As noted above, first estimates from GOME solar measurements show that Q_{SUN} is roughly 13% for PMD 3, 2% for PMD 2 and 0.1% for PMD 1. In the operational baseline, (47) has been used in the p–value determination from PMD 3 and PMD 2. See also the discussion in section 2.9.5 below.

2.9.2.2 Array–Array Overlap

This sub–section concentrates on the evaluation of one fractional polarisation value for a channel overlapping region, e.g. between channel 2 and channel 3. Here, *a* and *b* refer to separate channels.

First, we require the ratio of *unpolarised* light signals (p = 0.5). In this case:

$$\textit{Channel a}: \quad T_i = a_{p,i} \cdot I_{p,i} + a_{s,i} \cdot I_{s,i} = 0.5 \cdot (a_{p,i} + a_{s,i}) \cdot I_i \tag{62}$$

$$\textit{Channel b}: \quad T_j = b_{p,j} \cdot I_{p,j} + b_{s,j} \cdot I_{s,j} = 0.5 \cdot (b_{p,j} + b_{s,j}) \cdot I_j \tag{63}$$

$$\therefore R_{u,i} = \frac{T_i}{T_j} = \frac{a_{p,i} + a_{s,i}}{b_{p,j} + b_{s,j}} = \frac{a_{p,i}}{b_{p,j}} \cdot \frac{1 + \eta^{\Omega}_{a,i}}{1 + \eta^{\Omega}_{b,i}} \tag{64}$$

where T_i is the signal of detector pixel *i* of the first array and T_j is the *interpolated signal of a virtual detector pixel j of the second array.* [One pixel grid in the overlap region is chosen, and signals not defined on that particular grid must be interpolated (here λ_i is equal to λ_j)]. $R_{u,i}$ is a ratio of these two signals. $\eta_{a,i}$ and $\eta_{b,i}$ are corrected for the appropriate scan mirror angle position Ω.. We may invert (50) to obtain the ratio of p–polarised sensitivities:

$$\frac{a_{p,i}}{b_{p,j}} = \frac{1 + \eta^{\Omega}_{b,j}}{1 + \eta^{\Omega}_{a,i}} \cdot R_{u,i} \tag{65}$$

For the ratio of two *polarised signals* S_i and S_j with fractional polarisation p, the corresponding version of equation (50) is:

$$R_{p,i} = \frac{S_i}{S_j} = \frac{a_{p,i}}{b_{p,j}} \cdot \frac{\eta^{\Theta}_{a,i} + p\,(1 - \eta^{\Theta}_{a,i})}{\eta^{\Theta}_{b,i} + p\,(1 - \eta^{\Theta}_{b,i})} \tag{66}$$

where the appropriate scan mirror angle position is indicated by Θ (not necessarily the same as Ω). Substituting the ratio (a_p/b_p) from (51) we get:

$$R_{p,i} = \frac{1 + \eta^{\Omega}_{b,j}}{1 + \eta^{\Omega}_{a,i}} \cdot R_{u,i} \cdot \frac{\eta^{\Theta}_{a,i} + p\,(1 - \eta^{\Theta}_{a,i})}{\eta^{\Theta}_{b,i} + p\,(1 - \eta^{\Theta}_{b,i})} \tag{67}$$

which can be inverted to give a value of p:

$$p = \frac{\left\{X_i \cdot \eta^{\Theta}_{b,j} \cdot (1 + \eta^{\Omega}_{a,i}) - \eta^{\Theta}_{a,i} \cdot (1 + \eta^{\Omega}_{b,j})\right\}}{\left\{(1 + \eta^{\Omega}_{b,j}) \cdot (1 - \eta^{\Theta}_{a,i}) - X_i \cdot (1 + \eta^{\Omega}_{a,i}) \cdot (1 - \eta^{\Theta}_{b,j})\right\}} \quad (68)$$

where $X_i = \frac{R_{p,i}}{R_{u,i}}$ (69)

Equation (54) is in theory valid for each pixel in the channel overlap region. In practice, p is calculated for a number of values in the neighbourhood of the "cross–over points" (equal signals from the two arrays). Signals must not be too small (introduction of noise – a minimum of 100 BU is current default), and this will limit the size of the final choice of grid points to use in the overlap region. Values of p calculated from (51) are then averaged for the final fractional polarisation answer. Note that unpolarised and polarised signals may have different wavelength registrations – interpolation must be made to *one choice of pixel grid*. Note also the different scan mirror dependencies in η.

It is possible to use a *pre–flight source* $R^*_{u,i}$ for the ratio of unpolarised light signals (for example, measurements taken with an FEL lamp, with scan mirror in nadir position). However, it is considered better to use in–flight solar measurements to derive $R_{u,i}$. In this case, signals in equation (50) are solar signals, the scan mirror angle Ω is that for the sun position; any polarisation dependency of the Calibration Unit is neglected.

2.9.2.3 The "Seventh" Point

For the UV part of the spectrum a fractional polarisation value may be derived from a model calculation. Below 300 nm, polarisation is dominated by Rayleigh scattering, and a simple calculation based on a Rayleigh single–scattering model (see [A9]) will return *a constant value* depending only on the illumination geometry. This is a useful adjunct to the p–values derived from GOME measurements, since the last available measurement point is the first overlap (around 312.5 nm) – the wavelength region 300–325 nm contains the strongest variation of atmospheric polarisation, and any additional pointers to the polarisation in this region are useful.

The *degree of polarisation* due to single scattering by molecules is

$$P_{ss} = \frac{1 - \cos^2\Theta}{1 + \Delta + \cos^2\Theta} \quad (70)$$

where Δ is an anisotropy constant related to the Rayleigh depolarisation ratio ($\Delta = 0.0574$ for depolarisation ratio 0.0295), and Θ is the scattering angle at a representative scatter height in the atmosphere ($\approx$30 km for wavelengths less than 300 nm). Θ is given by :

$$\cos\Theta = -\cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\phi - \phi_0) \quad (71)$$

where θ, θ_0 and ϕ–ϕ_0 are the *local geometrical angles at the scattering height* (line–of–sight zenith, solar zenith and relative azimuth respectively). Note that these angles must be derived from the geolocation information – the conversion from satellite height to scatter height and the derivation of local angles follows standard results from spherical trigonometry.

Also required is the angle of the polarisation plane (with respect to the local meridian plane). This is given by

$$\chi_{ss} = \begin{cases} 180 - \beta & (\text{if } \sin\beta \geq 0) \\ -\beta & (\text{if } \sin\beta < 0) \end{cases} \quad (72)$$

where:

$$\sin\beta = \frac{\cos\theta_0 + \cos\theta\cos\Theta}{\sin\theta\sin\Theta} \tag{73}$$

When $\theta = 0$, the local meridian plane is not defined (direct nadir view), and χ_{ss} is then taken as the angle between the scanning plane and the perpendicular to the plane of incident sunlight. For more details of this special case, see [A9].

The relation between degree of polarisation P and fractional polarisation p_7 is given by:

$$1 - p_7 = \frac{(1 + P_{ss}\cos 2\chi_{ss})}{2} \tag{74}$$

We follow here the convention used in the calibration of GOME (p = fractional polarisation parallel to GOME optics); the original seventh point formulation in [A9] specified p for the perpendicular optics direction – hence the (1–p) in (57). Although the degree of polarisation P_{ss} has greater physical meaning, it is more convenient in the calculation of the polarisation correction factor (see section 2.7.4) to use the fractional polarisation p. Model studies have shown that equation (53) for P_{ss} is valid for λ less than 300 nm, while the plane of polarisation angle χ remains very close to χ_{ss} over the whole GOME range.

Once the geolocation is known, p_7 can be calculated readily. Orbit simulations show that the value of p follows the solar zenith curve quite closely, and there are some variations across the swath (east versus west pixels) – this is especially true for the large swath width option (960 km).

Geolocation is only specified every 1.5 seconds, and care must be taken to calculate accurately the geometrical angles for the short integration time (0.375 and 0.75 seconds) pixels. In the original specification, the short readout geometrical angles were evaluated by interpolation; following the change to the ERS propagator, it is now possible to generate these angles using precise ephemeris modules. However, the advent of co–adding has made this point redundant.

2.9.2.4 Wavelength of Fractional Polarisation Values

For the channel overlap regions, there is little problem with the wavelength definition. The representative wavelengths are the mid–points of those small windows defined round the cross–over points and used to average (49) for the required p–value. This gives us a well defined point $[\lambda, p]$ on the polarisation curve.

The situation is not so simple for the PMD points. Here one must assign a representative wavelength for the single p–value assumed constant over the PMD wavelength range. Rather than placing this point at the middle of the range, it is situated at some pixel centre of gravity found by weighting the array detector signals over this range. The formula suggested for this weighting is:

$$n(p) = \frac{1}{(n_2 - n_1 + 1)} \cdot \sum_{j=n_1}^{n_2} \frac{S_j}{\overline{S}} \cdot j \tag{75}$$

where $\overline{S}$ is the mean value of the array detector signals from n_1 to n_2 and S_j is the signal of each individual pixel.

2.9.3 Application of the Polarisation Correction Algorithm

The first two subsections herein describe the interpolation of p–values to wavelengths away from the 6 or 7 values determined from the algorithm above. Section 2.7.3.3 gives the final calculation of correction factors.

2.9.3.1 Interpolation of Fractional Polarisation Values

The polarisation Correction Algorithm (PCA) of GOME Level 0 to 1 Processing, as described above, generates a small array of fractional polarisation values at certain well–spaced wavelengths across the GOME spectral range. Up to 7 points can be specified (see figure 12). Though interpolation with fractional polarisation p is equivalent to interpolation with degree of polarisation P, we shall work with p throughout.

Three points come from the comparison of integrated nadir signals with broad–band PMD signals (points p_2, p_4 and p_6 in Figure 12; corresponding wavelengths λ_2, λ_4 and λ_6 are approximately 370 nm, 500 nm and 700 nm, depending on signal values). Three more points come from the channel overlap regions (points p_1, p_3 and p_5, corresponding to wavelengths 312.5, 400, and 600 nm approximately). The last ("seventh") point (labelled $[\lambda_0, p_0]$) comes from the Rayleigh single–scatter model simulation of polarisation in the Ultra Violet.

We describe the *interpolation scheme for generating values of polarisation at all GOME wavelengths*, based on the 7 points described above. This "polarisation Interpolation" module forms part of the Application of level 0–1 Processing for the generation of Extracted Level 1 data.

In figure 12, the polarisation curve is divided into 5 regimes. For the 5 possible measurement–derived points between λ_2 to λ_6 (regime IV), a suitable interpolation routine is sufficient (Akima interpolation has been selected) – this applies for PMDs 1, 2 and 3, and for overlaps 2 and 3. Interpolation routines are not reliable for extrapolation beyond λ_6 to the end of the GOME range. Instead, a constant has been assumed for regime V, with value equal to p_6.

For the critical regions in the UV (regimes I, II and III in Figure 12), a *parameterisation of the polarisation curve* is assumed (analytic functions described by a small number of parameters which must be fitted according to the available information in that region). The parameterisation is described as follows:

Simulations of polarisation show that up to wavelength λ_0 ($\approx$300 nm), the degree of polarisation $P(\lambda)$ is virtually a constant, so we set fractional polarisation equal to the *single–scatter 7th point value p_0*. $P(\lambda)$ then changes rapidly between λ_0 and a value $\overline{\lambda} \sim 325$ nm, then shows smooth behaviour

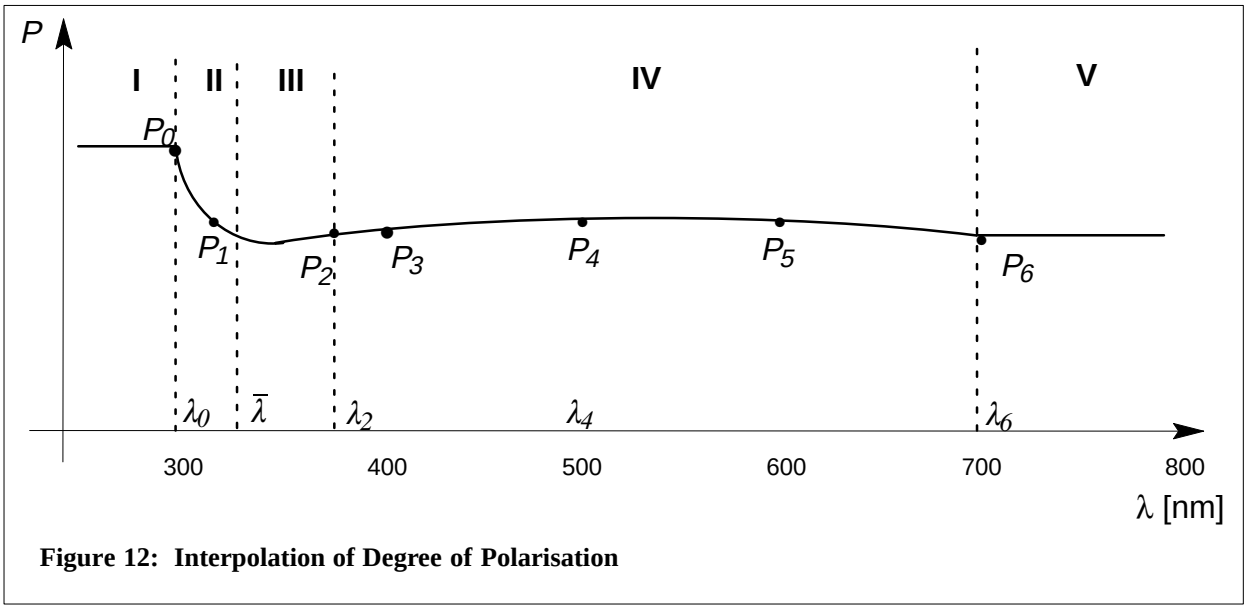

Figure 12: Interpolation of Degree of Polarisation

into the visible (*ca.* 325–400 nm). Thereafter, P(λ) is a smooth slowly–varying function of wavelength.

The critical region is λ_0 to $\bar{\lambda}$ (300–325 nm; regime II). An analytic function is required that mirrors the shape of simulated polarisations for all atmospheric scenarios. This function must have a turning point at the 7th point λ_0, with parabolic fall–off behaviour near this point. The function must also show exponential tail behaviour as λ tends towards $\bar{\lambda}$. The *Generalized Distribution Function* (GDF) :

$$F_1(\lambda) = \bar{p} + \frac{w_0\, e^{-(\lambda-\lambda_0)\beta}}{\left[1 + e^{-(\lambda-\lambda_0)\beta}\right]^2} \tag{76}$$

fulfils this role. Here $\{\bar{p}, w_0, \beta\}$ are parameters that characterise the GDF; they must be found to fit the given interpolation points.

Beyond $\bar{\lambda}$ and out to the first interpolation point λ_2, the polarisation curve shows smooth variation with λ. p in this regime III ($\bar{\lambda}$ to λ_2) will be approximated by a polynomial of low order :

$$F_2(\lambda) = \sum_{k=0}^{k=2} A_k \cdot \left(1 - \frac{\lambda}{\bar{\lambda}}\right)^k \tag{77}$$

Parameters $\{A_k\}$ again must be chosen to fit interpolation points and to retain continuity of the interpolation curve.

Figure 13 illustrates in more detail the parameterisation scheme proposed for wavelengths up to the first PMD value λ_2. With 3 parameters for the GDF, and 3 for the polynomial, we require 6 equations for the solution. Two equations come from the values p_0 and p_1 at λ_0 and λ_1 respectively. A smoothness assumption is also assumed (function and gradient continuity). Thus 2 more equations come from the continuity of p and its gradient at $\lambda=\bar{\lambda}$, and two more from continuity at $\lambda=\lambda_2$. [The gradient at $\lambda=\lambda_2$ equals that computed from the Akima interpolation at λ_2]. The gradient at $\lambda=\lambda_0$ is automatically continuous from the nature of the GDF function $F_1(\lambda)$ which has a turning point there. Mathematics of the parameter derivations are given in the Appendix A on page 62.

When the first overlap point p_1 is absent, we drop the polynomial parameterisation, and try to link the GDF curve directly with the first interpolation point (λ_2, p_2). The seventh point value p_0 and the 2 continuity conditions at $\lambda=\lambda_2$ provide enough information to calculate the 3 GDF parameters. However, when the gradient $\lambda=\lambda_2$ at is positive, the GDF curve cannot be linked to the interpolation (GDF curve always has a negative gradient over range of application). In this special case (which mainly occurs for values of p close to 0.5), a parabola is fitted between $\lambda=\bar{\lambda}$ (325 nm) and $\lambda=\lambda_2$, with a minimum halfway between these wavelengths.

The first version of the polarisation application had separate curves for short and long integration times – the first overlap point (λ_1, p_1) was only present for long integration times. The situation has now changed as a result of FM calibration; (λ_1, p_1) is available at short integration times, and there is no need to consider polarisation for long integration times.

First Operational Baseline (25 July 1995)
The first data from GOME in space have shown that the most reliable measurement points for p come from the PMDs. Overlap 2 is absent, and overlaps 1 and 3 show too many variations at present to be reliable sources for the corresponding p values; further investigation is required. The 7th point calculation is also reliable, provided care is taken over geolocation values for short and long integration times. The baseline then is to perform Akima interpolation over the 3 PMD values, and to join

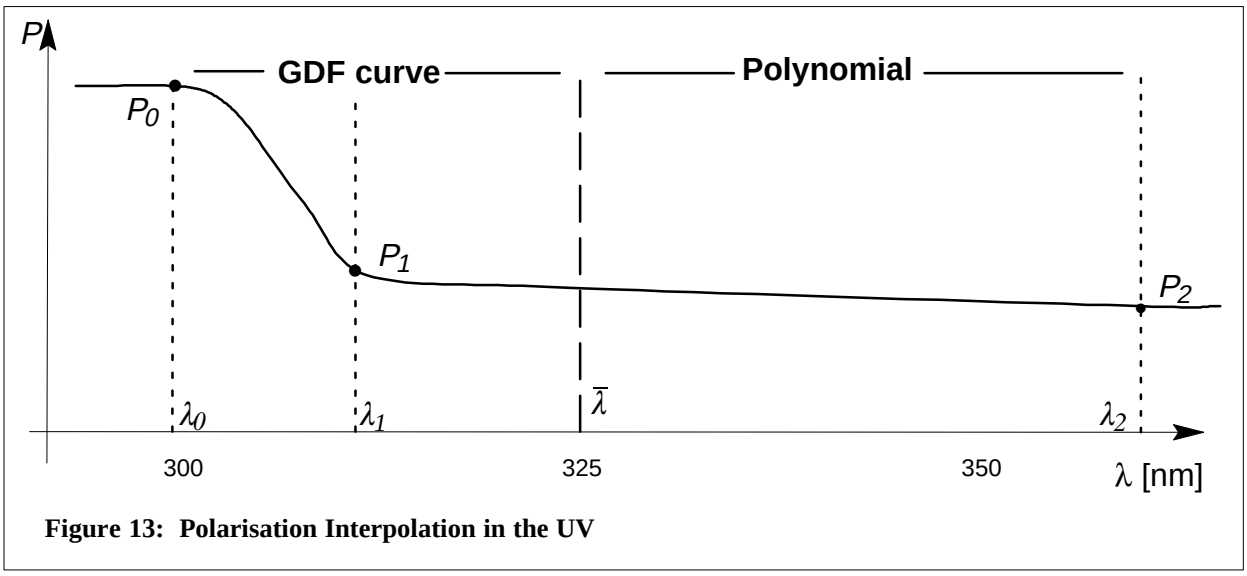

Figure 13: Polarisation Interpolation in the UV

the interpolation curve smoothly to a GDF function in the UV to 300 nm. Fractional polarisation will be constant for wavelengths out to 800 nm beyond the third PMD point.

2.9.3.2 Implementation of the Polarisation Interpolation Module

The ”polarisation interpolation” algorithm as described above is implemented as a self–contained module in the Application of polarisation Correction sub–algorithm in Level 0–1 processing. The inputs are the 7 fractional polarisation values and their wavelengths [λ_i, p_i]. [The polarisation angle χ is also passed, but is no longer used in the application]. Also inputted are the wavelengths for each band (labelled according to pixel number 1 to 1024). The result is simply the interpolated values of fractional polarisation on the given wavelength grid.

When a p–value occurs outside the range [0,1] due to corrupted data or some other factor, it is assigned a fixed value equal to the parameter constant INVALID_FRAC_POL. If only one of the interpolation points [λ_i, p_i], i = 1, .. 6 is valid, then all values of polarisation from $\lambda=\bar{\lambda}$ onwards are constant equal to the single correct p–value (regimes II, IV and V) – interpolation using the GDF curve is still possible over regime II provided the 7th point is present. If two values of the above set are present, the interpolation in regime IV is a straight line – regime III is then a continuation of this line. Akima interpolation is valid for 3 or more such interpolation points. If all 6 interpolation points are absent, the module returns a serious error status.

If the 7th point is invalid, then no analytic curves can be calculated. In this case, Akima interpolation is performed over the available valid points, with straight line extrapolation beyond the first interpolation point (beyond last point, assume constant out to 800 nm). This situation is unlikely to arise with in–flight processing (point [λ_0, p_0] requires only geolocation information for its generation, and should therefore be present for every ground pixel). However, it is useful to retain this contingency, as the polarisation correction algorithm will require testing with GOME pre–flight FM measurements (no 7th point).

2.9.3.3 Calculation of the Correction Factor

Once the polarisation interpolation is complete, it is straightforward to compute the *polarisation correction*. The conversion of polarised signals to ”corrected” unpolarised signals proceeds as follows. For unpolarised signals (p = 1/2) we may write:

$$S_i^{unpol} = \frac{1}{2} \cdot I_i \cdot \left\{ a_{p,i} + a_{s,i} \right\} \qquad (64)$$

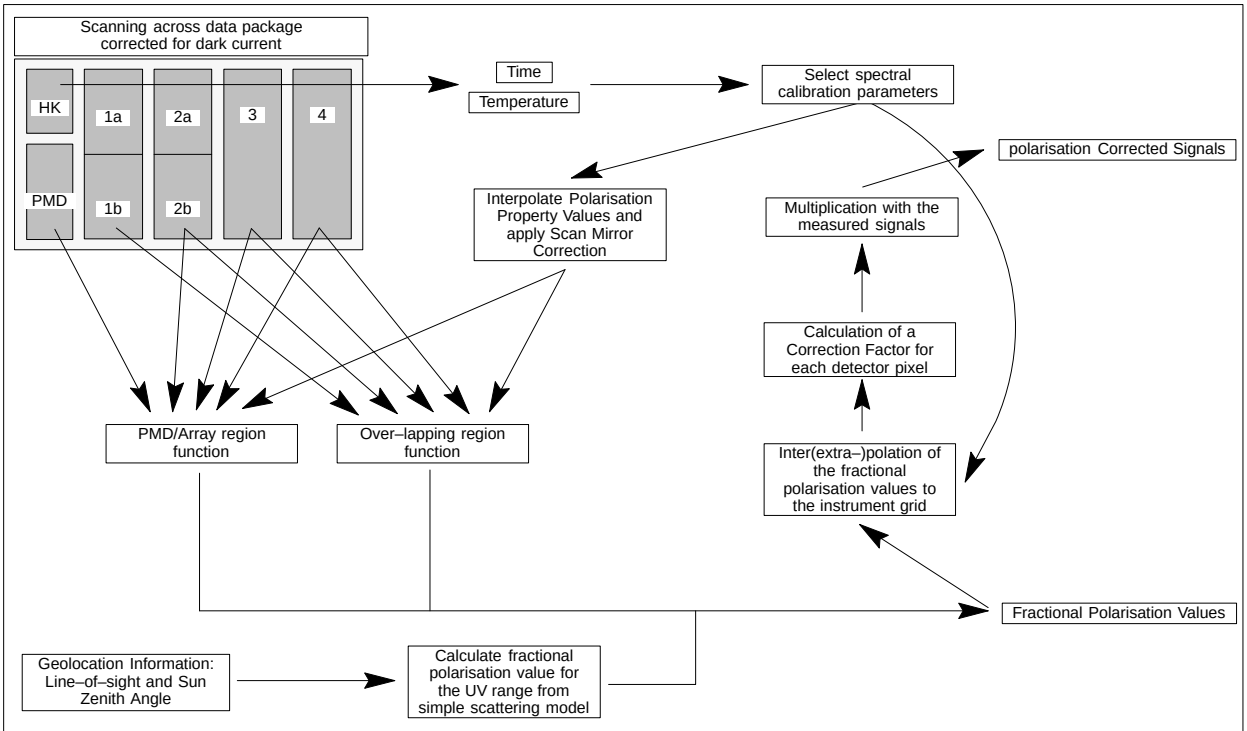

Figure 14: Determination and Application of Polarisation Correction

For polarised detector pixel elements:

$$S_i^{pol} = I_i \cdot \left\{ a_{p,i}\, p_i \; + \; a_{s,i}\,(1–p_i) \right\} \tag{65}$$

where I_i is the total intensity and $a_{s,i} / a_{p,i} = \eta_i$, the (pre–flight measured) pixel polarisation sensitivity, corrected for scan mirror position.

If the wavelength λ_i is known for pixel i, then $p_i = p(\lambda_i)$ is known from the polarisation interpolation module. $\eta_i = \eta(\lambda_i)$ is interpolated from the pre–flight data on a given wavelength grid, then adjusted to the appropriate scan mirror angle position. The correction factor for pixel i is given by :

$$c_i \;\equiv\; \frac{S_i^{unpol}}{S_i^{pol}} \;=\; \frac{1}{2} \cdot \frac{1 + \eta_i}{p_i \cdot (1–\eta_i) + \eta_i} \tag{66}$$

This completes the Level 0–1 polarisation correction application.

2.9.4 Summary of Algorithm Steps

The following list and the diagram in figure 14 summarises the steps in the polarisation correction algorithm.

- Interpolate the polarisation property values (η^*_{nadir} and ξ^*) to the instrument wavelength grid specified by the current set of spectral calibration parameters;
- Apply the scan mirror correction factor to η^*_{nadir} to yield η (36);
- [Apply the PMD temperature correction factor to $\xi^*_{(Air,T0)}$ to yield ξ^* for the following calculation (52) ; NO LONGER REQUIRED]

- Calculate the normalised signals in BU/s for the detector arrays and the PMDs (53) (54);
- Find the fractional polarisation values for the PMD detectors (43);
- Find the fractional polarisation values for the three over–lapping regions (54);
- Calculate the fractional polarisation value for the UV part of the spectrum using a Rayleigh single–scattering model (60);
- Calculate a fractional polarisation value for each detector pixel using the interpolation algorithm;
- Calculate a correction factor for each detector pixel (66);
- Apply the correction factor to the signals.

2.9.5 Discussion and Open Issues

2.9.5.1 Problems of the Algorithm

(a) Availability of polarisation points

Results from the FM calibration have indicated that the *p–value from second channel overlap point cannot be retrieved*. This point is absent, and it was this consideration which led to the generation of a software patch to include the *first* overlap point at every short integration time. See note above on First Operational Baseline.

(b) PMD Straylight

The problem with PMD response beyond 790 nm is potentially serious. While it is possible to characterise the PMD straylight for the *unpolarised* solar light by a comparison between the solar PMD signals and their virtual sum equivalents, such a procedure is not permissible for the earthshine spectra, because of the *unknown state of atmospheric polarisation* in the straylight region.

The low value for PMD 1 is in agreement with straylight tests carried in the TV chamber at Galileo. These results for unpolarised sunlight are in contrast to the ”Brewster” tests carried out with light at constant polarisation. There, the PMD 1 straylight was 20–25% – which translates into a 10–15% error in retrieved p–value. Straylight errors for PMD 2 and PMD 3 were found to be an order of magnitude smaller. These discrepancies need to be sorted out, but the GOME in–flight values look to be the right ones.

(c) Algorithm Deficiencies

The algorithm presented here represents a compromise to extract the maximum information from what is a very sparse set of observational points. The use of model data (the seventh point) in conjunction with measurement–derived p–values is typical of the compromise.

Apart from the neglect of PMD straylight, there is a further problem with the PMD formulation. It was assumed that p–values are constant across a given PMD range. While this assumption is (in general) not too bad for PMDs 2 and 3, where the polarisation curves are slowly varying and flat (in general!), the situation for PMD 1 is otherwise – here the polarisation shows marked non–linear behaviour and often large changes over the interval 300–400 nm. The approximation with a constant weighted to a particular pixel is probably unwarranted. This constant value may be grossly in error, and the resulting parameterisation scheme used in the application may be based on false inputs – such an error is difficult to quantify at this stage.

Though the polarisation curve is in general smooth across the GOME range, there are some detailed fine structures imposed on the general form. In particular, these include small structures due to O_3 absorption in the Hartley–Huggins bands, and a somewhat larger deviation in polarisation over the O_2 A–band. Recent modelling has shown strong perturbations in polarisation are possible in the visible (linked to surface albedo spectral dependence). It is important to realise that, with only 7 or less polarisation points covering the range 240–800 nm, it is really only possible to reproduce the overall shape of the polarisation within reasonable error bounds – extra information is required to fix the small–scale features. The first object of the present interpolation scheme is to fix as accurately as possible the *broad shape* of the polarisation interpolation curve. The status of polarisation parameterisation is summarised in the next section.

2.9.5.2 Model Comparisons for the parameterisation of P(λ)

Simulations of the Degree of polarisation P(λ) were carried out over the range 290–350 nm at a resolution of 0.2 nm, for a number of viewing geometries and a selection of model atmospheres. The simulated data was smoothed to remove the small–scale O_3 Huggins structures before comparison with analytic expressions. Polynomial and GDF functions were used as described above, and suitable parameters were found by minimizing the root–mean–square error between the analytic approximation and the smoothed model data over the range 290–350 nm. [$\bar{\lambda}$ was set at 325 nm]. The resulting analytic interpolation curves were found to be less than 2.5% different from the model data equivalents for all wavelengths and for all cases studied.

To obtain the best comparisons with model data, the λ_0 value was allowed to vary according to the value of P_0 (for example, if $P_0 \sim 0.80$, then $\lambda_0 \sim 302.0$ nm; if $P_0 \sim 0.20$, then $\lambda_0 \sim 298.0$ nm). The point to note here is that λ_0 depends (through P_0) only on the geometry, not on the atmospheric scenario or GOME measurement values . A pre–determined set of values $\{\lambda_0^*, P_0^*\}$ can be generated from model calculations, and then the true value λ_0 corresponding to P_0 found from interpolation over $\{P_0^*\}$.

Examination of model data has shown that the wavelength at which the gradient of polarisation is steepest depends only on the scattering geometry (or equivalently on the value of p_0), and is independent of the atmospheric composition. This means we can pre–calculate a data set $\{\lambda_m^*, P_0^*\}$ of maximum–gradient wavelengths λ_m^* against model simulations P_0^*, then interpolate to find the actual value of λ_m corresponding to a given 7^{th} point value P_0.

Although small data sets $\{\lambda_0^*, P_0^*\}$ and $\{\lambda_m^*, P_0^*\}$ were constructed from model simulations with the intention to improve the Level 0–1 polarisation algorithm, it was decided for simplicity to stick to an interpolation scheme requiring no additional data. Modelling studies are not sufficiently advanced yet to generate reliable look–up tables to be used in the application of polarisation. Again, it is possible to improve the parameterisation in the UV range by artificially superimposing Huggins structures on the analytic curves to match the accurately calculated model data; at present, this second order improvement has not been implemented.

Despite these cavils, the overall shape of the polarisation curve in the UV is very well approximated by a combination of polynomial and GDF functions.

2.9.5.3 Blueprint for a second version of PCA

The following recommendations are made for a more sophisticated version of the Polarisation Correction Algorithm (Mark 2). The requirements are for a more consistent ”cleaner” scheme that properly addresses the PMD straylight issues, and the overall parameterisation.

- The *entire polarisation curve should be parameterised as a series* o*f piecewise–continuous analytic functions*. If possible the polarisation gradient with wavelength should also be piece-

wise continuous. This means that PMD equations will solve for parameters rather than constant p–values.

- *The PMD straylight contributions should be included in the PMD equation*s. Some extra information is required here – the PMD quantum efficiency and some model output of spectra in the near infra–red beyond 790 nm (where water vapour absorption dominates the earthshine spectrum).
- Further study is required for the examination of the parameterisation scheme (using radiative transfer model with polarisation). In particular the second order structures need to be properly characterised.
- The known PMD straylight for *solar* signals must be used to determine the instrument response of each PMD, and this information is then used in the *earthshine* straylight computation.
- In this scheme, *application of the polarisation correction has no interpolation*, due to the parameterisation assumption (this includes parameterisation of the PMD straylight region).
- Determination of p–values from channel overlaps is not changed.

2.9.6 Algorithm Update (Issue 4)

- Averaging PMD signals

Because of the definition of the polarisation sensitivity ξ (PMD–pixel–ratio) the PMD values do not need to be normalised by integration time. For the p calculation using PMD signals, we need only the mean value of all PMD readouts during the integration period of the corresponding channel. Since signals of several channels are used in equation (43), it is a prerequisite that all such channel integration times are equal.

PMDs are summed and averaged as follows:

$$S_{PMD} = \frac{1}{16 \cdot m} \cdot \sum_{i=1}^{m} \sum_{j=1}^{16} S_{PMDj,\,i}^{readout} \quad [BU] \tag{67}$$

where m is the integration time of the channel divided by 1.5 seconds[3] and $S_{PMDj,\,i}^{readout}$ are the signals of the PMD during this time. j runs over the 16 PMD measurements which are performed during one PMD readout (1.5 seconds).

- The assignation of wavelengths for the 3 p–values derived from PMD signals has been improved to take into account the polarisation sensitivity ξ into the weighting formula (61). A revised version of this equation is:

$$n(p) = \frac{1}{(4096)} \cdot \sum_{j=0}^{4095} \frac{S\xi^*_j}{\overline{S\xi^*_j}} \cdot j \tag{68}$$

where $\overline{S\xi}$ is the mean value of the array detector signals multiplied by the ξ, S_j is the signal of each individual pixel, ξ_j the polarisation sensitivity of each individual pixel. Since all four GOME channels together have a grand total of 4096 detector pixels, the loop is extended to cover this number of indices.

3 e.g. integration time 30 sec. means m=20

- Determination of Fractional Polarisation Values derived using overlapping signals

Fractional Polarisation Values are calculated for all wavelengths in the overlapping region. The resulting curve is then checked for flatness near the cross–over intersection; if the curve is smooth the p value at the exact intersection wavelength is taken as a valid algorithm result, otherwise p is set to INVALID_FRAC_POL.

- Calculation of the ”Seventh Point” for integration times other than 1.5s

The ”Seventh point” p–value is a function only of the viewing geometry – azimuth and zenith angles of solar and line–of–sight paths. These angles depend on the integration time of the detector bands. In GDP, geolocation information is calculated for every 1.5 s ground pixel. For longer integration times (multiples of 1.5 seconds), it is possible to calculate a mean value for p_7 by averaging the 1.5–second values. p_7 values for shorter integration times can be computed readily since the geolocation information for smaller pixels can be calculated using the ERS ephemeris programme. For long readout pixels, line–of sight geometry changes substantially over the pixel scene, and it makes more sense to average 1–5–second values rather than calculating p_7 from scratch using an unrepresentative line–of–sight geometry.

2.9.7 Algorithm Update (Issue 5)

Computing the fractional polarisation value *p* of PMD #1 for a long observiation time, a clear deviation from the expected theoretical value of 0.5 is observed for PMD channel # 1. Since the PMD channel # 1 has not been before corrected for degradation, the degradation correction as implemented for PMD channel # 2 and # 3 is now also applied to PMD channel # 1. The correction is given by Eqn..(61) in section 2.9.2.1.

2.9.8 Algorithm Update (Issue 6)

The wavelength interpolation of the polarisation in the UV, using the GDF curve (section 2.9.3.1), has been updated. For details, see section C.3 . Note that this interpolation incorporates a different calculation of the wavelength of the seventh point. This is calculated only in the wavelength interpolation routine, which is implemented in the GDP extraction software. The value for the seventh point on the binary Level 1 product has not been changed.

2.10 Correction of Array Detector Degradation

2.10.1 Notation

The following definitions are used:

λ	wavelength [nm].
t	elapsed time since start of GOME measurements
$I_{Sun}(\lambda,t)$	irradiance transmitted from the sun dependent on time and wavelength
$P_{Deg}(\lambda,t)$	degradation function dependent on time and wavelength
$C_{SED}(t)$	Sun–Earth distance intensity correction
$Residual(\lambda,t)$	remaining structure of degradation algorithm
$a_i(t)$	polynomial degradation parameters for solar spectra $I_{Sun}(\lambda,t)$
b_{ij}	set of degradation parameters characterizing time and wavelength dependent behavior of degradation

2.10.2 Algorithm Description

In all GOME wavelength regions there is degradation caused by hard UV light and cosmic radiation. Coatings, optical parts and the detectors are damaged by this. Additionally changes in instrument sensitivity are due to material deposited on the surfaces or out gassing of coatings.

Following algorithm for degradation correction based on solar measurements was developed by DLR in the framework of the ESA/ESRIN–project GDAQI. The principle idea to monitor the changes is the use of ratios of all solar spectra with a reference measurement of the early GOME lifetime (t_0) as given in eqn. (78). This is explained in detail in [A26]. The dominant part of the degradation influence can be found by study of the solar irradiance of the Mean Sun Measurements, because most of the light path via the critical optical parts is identical for all measuring modes.

Please note: Since the light path for Sun and Earthshine measurements are different (the sun uses the diffuser and is picked up under a different scan mirror angle), also Earth–shine measurements should be analysed though this is not straightforward, since atmospheric parameters are variable one can only look statistically at large data sets. Unfortunately, this is particularly important since the ratio of Earth–shine to solar spectra is the quantity used in atmospheric tracegas retrievals.

$$\frac{I_{Sun}(\lambda, t)}{I_{Sun}(\lambda, t_0)} = P_{Deg}(\lambda, t) \cdot C_{SED}(t) \cdot Residual(\lambda, t) \tag{78}$$

where $P_{Deg}(\lambda,t)$ is the degradation function dependent on time and wavelength, $C_{SED}(t)$ is the Sun–Earth distance intensity correction due to its seasonal variation and $Residual(\lambda,t)$ is the remaining structure.

A broadband Etalon structure is present in all detector channels, together with a change of the dichroic structures for the detector channels 3 and 4. These effects are not taken into account and are found as remaining structures in the residuals. A two–step approach has been chosen to describe the degradation phenomenon: First, each radiance ratio has been approximated by a wavelength dependent polynomial function. Then, the coefficients $a_i(t)$ of these polynomials have been taken to retrieve a time dependent polynomial function with coefficients b_{ij} . Thus, for the degradation function $P_{Deg}(\lambda,t)$ following expressions have been obtained:

$$P_{Deg}(\lambda, t) = \sum_{i=0}^{n} a_i(t) \cdot (\lambda - \lambda_0)^i \tag{79}$$

$$a_i(t) = \sum_{j=0}^{m} b_{ij} \cdot (t - t_0)^j \qquad (80)$$

The coefficients b_{ij} and central wavelengths λ_0 are used by by the extractor software of GDP Level 0–to–1 (read from a file) to correct the degradation.

2.10.3 Algorithm Update (Issue 5)

This algorithm is valid for GDP level 0 to 1 algorithm from version 2.0 and higher and can be applied optionally by the extractor software of GDP Level 0–to–1 to all measurement modes. The file providing the degradation parameters can be found at http://auc.dfd.dlr.de/GOME. It is intended to update this file on a regular basis.

2.10.4 Algorithm Update (Issue 6)

Degradation in the light path difference between the Sun measurements and the Earth–shine measurements has now been derived. A correction may be applied in the extractor software of GDP Level 0–to–1. For details, see section C.4. The data for this correction are taken from the same external degradation parameter file given in section 2.10.3.

2.11 Quality Flagging

From sun calibration and scanning measurements the following quality assessment calculations are performed, and corresponding flags are set for these measurements.

- Identification of dead pixels by inspection of the associated pixel–to–pixel gain correction set; the location of the dead pixel are identified by a ”0” entry in the PPG correction array.
- Identification of ”hot pixels” by a statistical analysis in both wavelength and time of the signal data; after the initial identification, a hot pixel will remain so assigned until a certain time interval has passed (tbd 1). In addition to the time–occurrence hot pixel flag, the spatial location of hot pixels is recorded in a variable data structure in the data product. [Not applicable to sun and scanning data packets].
- Identification of saturation by searching for threshold failures in the signal data.This check can not be done when the instrument is in averaging mode.
- Calculation of a mean value and the standard deviation of the signal data for each detector array and comparison of these values with a set of thresholds (to be elaborated during the commissioning phase) (tbd 2).
- Identification of the possibility of sun–glint by comparison of the line–of–sight and solar zenith angles from the geolocation record, and reference to a land/sea mask. [Applies only to scanning data packets].
- Check of the spectral calibration parameters by comparison of wavelength registration – against one or more Fraunhofer lines in the range of each detector array. If the wavelength difference is larger than a certain value but smaller than a certain maximum threshold , the flag will set (tbd 3).
- Calculation of an additional leakage current and noise correction factor to be multiplied with the leakage current parameter set and the noise terms in the Level 1 Data product. This should deal with the orbital variation of these calibration parameters; this check will use the non–illuminated detector pixels of channel 1 whenever the asynchronous interface of the data handling unit (DDHU) to the ICU of ATSR transmits this data; if the data is not transmitted a factor derived from a global map of expected photon fluctuations will be used (acquired during the commissioning phase) (tbd 4).

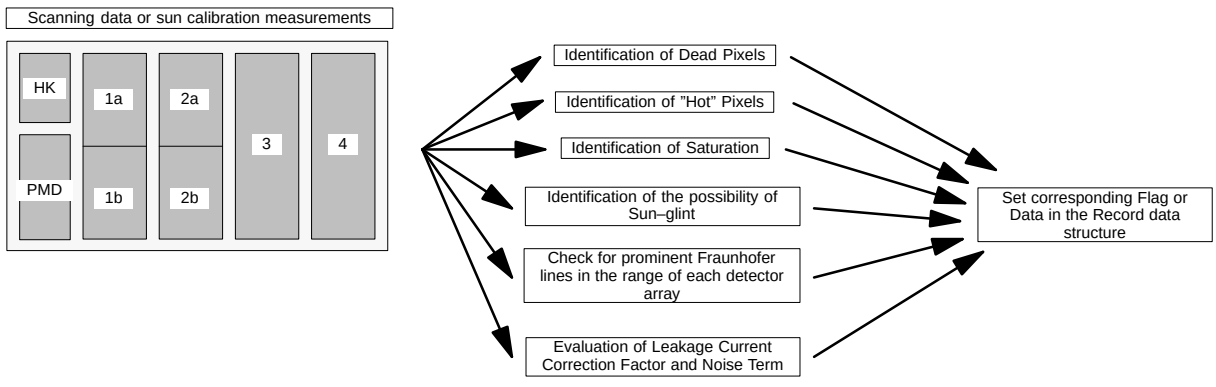

Figure 15: Quality Assessment and Flagging Scheme

2.12 Determination of Errors

Two major types of error which are determined during the Level 0 to 1 Processing, namely:

1. Wavelength errors – the uncertainty on the spectral calibration parameters; and

2. Radiometric errors – the relative/absolute errors on the radiance and irradiance measurements.

The wavelength errors are determined from the polynomial fitting of the spectral calibration parameters from lamp line centroid–pixel/wavelength pairs; these fitting errors are archived in the in–flight calibration database. Whenever these parameters are required for a specific product file, they are copied from this data base into the Level 1 Data products. Radiometric errors are calculated for each individual detector pixel and written to the output during the extraction of Level 1 Data from Level 1 Data products.

2.12.1 Prerequisites

The following *sources of errors* can be distinguished for GOME measurements:

- *Coarse Errors*

  These errors are severe deviations from expected results. They occur for internally or externally caused instrument failures or wrong handling of the instrument. There is no precise rule or mathematical description for such errors. Although careful and comprehensive measurement procedures should eliminate most coarse errors, chance events can occur to produce unavoidable gross errors. It is then necessary to identify the measurement as 'wrong', and either reject it or try to correct it with an appropriate model. A "hot" detector pixel is an example of a coarse error source.

- *Systematic Errors*

  Systematic errors are expected deviations from the true values of a measurement that may be described with a physical model. It is the major objective of every calibration activity to take these physical laws into account and correct for such kind of measurement errors. There are three types of systematic errors:

  – *Static Errors* due to a known constant bias or a known functional dependence of the measurements with respect to the physical parameter in question (e.g. linear function of the temperature sensors);
  – *Drift Errors* caused by long–term changes of parameters (e.g. warming of the optical bench during the lifetime of the sensor);
  – *Periodic Errors* are similar to drift errors, but the changes of certain parameters are periodic (e.g. temperature change during orbit; etalon structures)

  The application of physical rules to estimate systematic errors requires the ancillary measurement of relevant physical values (e.g. temperature) and/or the use of appropriate models. As these other values and models are also not free of errors, and the propagation of any such additional errors must be accounted for. One can say that the in–flight calibration activities of the GOME sensor are mainly the correction for systematic errors – this must include the propagation of errors on the correction factors throughout the calibration algorithms.

- *Random Errors*

  Random errors on measurements are caused by known physical effects, but by their nature are unpredictable in size and direction. Normally these errors are the so–called noise; this should be as low as possible compared to the measurement signal. During the design phase, the signal to

noise level (S/N) is one of the most important performance criteria by which the quality of the instrument can be properly assessed.

Random errors can only be analysed by statistical methods. Single measurements cannot be corrected for random errors; a sample of measurements (sample size ideally > 10) taken under the same conditions is needed to compute mean values, standard deviations and other statistics. statistical algorithms like mean value, standard deviations, etc. If such samples are not available, noise levels should be characterised during a calibration exercise in order to assign *error bars* to individual measurements.

It is important to distinguish between the terms accuracy and precision:

- The *accuracy* of a measurement is a measure of how close its result is to the true value.
- The *precision* of a measurement is a measure of how well its result has been determined, without reference to its agreement with the true value. Precision is therefore a measure of the reproducibility of the result.

Assessment of Random Errors

If x is the true value and a the result of one measurement of x, the true error ϵ^{true} of this measurement is defined to be:

$$\epsilon^{true} = a - x \tag{81}$$

If x is unknown, so also is ϵ^{true}. From a statistical sample of similar measurements, an approximate value μ for x and its probability boundaries $\pm\sigma$ can be derived. If these measurements are given by $a_1, a_2, a_3, ..., a_n$ the statistical parameters are:

$$\mu = \frac{1}{n} \cdot \sum_{i=1}^{n} a_i \tag{82}$$

$$\sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (a_i - \mu)^2} \tag{83}$$

With the approximate value μ in place of the unknown true value x, the approximate error ε_i of each individual measurement a_i is:

$$\epsilon_i = a_i - \mu \tag{84}$$

The estimated error ε_μ of the mean value μ of a measurement series is given as follows:

$$\epsilon_\mu = \frac{1}{n} \cdot \sqrt{\sum_{i=1}^{n} \epsilon_i^2} = \frac{\sigma}{\sqrt{n}} \tag{85}$$

In most cases, the measurement value is assumed to be Gaussian Normal Distributed, with the probability density function ϕ characterised by the parameters μ and σ:

$$\phi(a) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(a-\mu)^2}{2\sigma^2}} \tag{86}$$

Thus the probability P that the measurement a of a parameter x will lie between certain boundaries $(\pm\Delta)$ is given by:

$$P = \int_{\mu-\Delta}^{\mu+\Delta} \phi(a)\, da \tag{87}$$

Some typical and useful values of P are given in the following table:

Boundary Δ	Probability P
0.67σ	0.500
1.00σ	0.683
1.96σ	0.950
2.00σ	0.954
2.58σ	0.990
3.00σ	0.997

In what follows, error calculations will use 2σ as error boundary to express the relative errors. Approximated errors used for the error propagation calculation are 1σ values.

Error Propagation

GOME data is converted into ”calibrated radiances” by applying calibration algorithms and calibration parameters. Many calibration parameters are established on a regular basis from in–flight observations of the calibration lamp, the internal LED and the sun and under dark conditions. In addition, data from pre–flight instrument calibration and characterisation (e.g. the polarisation characteristics of the optical chain) is required. All of these calibration parameters have errors. The size and type of errors on the pre–flight calibration parameters is described in [A5]. The assessment of errors on the in–flight calibration parameters is described in the following sections.

During the calculation and application of the calibration parameters, a number of formulae are used to determine intermediate and the final results. In the GDP level 0 to 1 processing chain, the propagation of the errors through these formulae to the end results are done by means of the Gaussian Error Propagation Law. If $f(x_1, x_2, x_3, ..., x_n)$ is a function of the values $x_1, x_2, x_3, ..., x_n$ and $\varepsilon_1, \varepsilon_2, \varepsilon_3, ..., \varepsilon_n$ are the corresponding approximated errors of these values, the error on the function result is defined as follows:

$$\epsilon = \sqrt{\left(\frac{\partial f(x_1,\ ...,\ x_n)}{\partial x_1}\right)^2 \cdot \epsilon_1^2 + ... + \left(\frac{\partial f(x_1,\ ...,\ x_n)}{\partial x_n}\right)^2 \cdot \epsilon_n^2} \tag{88}$$

For addition, multiplication and division of two parameters, this has the following effect:

$$f(x_1, x_2) = x_1 + x_2 \qquad \epsilon = \sqrt{\epsilon_1^2 + \epsilon_2^2} \tag{89}$$

$$f(x_1, x_2) = x_1 \cdot x_2 \qquad \epsilon = \sqrt{x_2^2\,\epsilon_1^2 + x_1^2\,\epsilon_2^2} \tag{90}$$

$$f(x_1, x_2) = \frac{x_1}{x_2} \qquad \epsilon = \sqrt{\left(\frac{1}{x_2}\right)^2 \epsilon_1^2 + \left(\frac{-x_1}{x_2^2}\right)^2 \epsilon_2^2} \tag{91}$$

2.12.2 Wavelength Uncertainty

In the calculation of the spectral calibration parameters, the ’broken’ pixel number of a selected lamp line is determined as a statistical signal mean value, and the spectral calibration coefficients

come from the linear fitting of a polynomial to a selected number of pixel number/wavelength pairs; the formulae are repeated here for convenience.

$$\bar{n} = \frac{1}{N} \cdot \sum_{i=n_1}^{n_2} i \cdot S_i \tag{92}$$

$$\lambda_i = a_4 n_i^4 + a_3 n_i^3 + a_2 n_i^2 + a_1 n_i + a_0 \tag{93}$$

Figure 16 shows an array of lamp line measurements. Some 20 lines are visible here, though not all have been selected for fitting; some of the non–showing lines have been fitted. The plotted points represent the deviation (in pixel number, for each line) of the fitted wavelength compared with the lamp line wavelength of the centroid.

The estimated deviation of the wavelength for each detector array is:

$$\epsilon_\lambda = \frac{\frac{1}{k} \cdot \sum_k \epsilon_i}{\sqrt{k-f}}$$

where k is the number of lines used in the fit, f is the degree of freedom of the fitted polynomial and:

$$\epsilon_i = |\lambda_{lit} - \lambda_i| \quad \text{(literature wavelength minus calculated wavelength)} \tag{94}$$

If ε_λ is divided by the linear coefficient of the spectral calibration parameters (a_1 in (93)) the error is be given in detector pixels.

2.12.3 Radiometric Errors

Unfortunately, the assessment of the radiometric errors is much more complicated. Several corrections are applied to the raw signal values using pre–flight and in–flight calibration parameters; all such processes will propagate parameter errors to the final result. The following errors on the radiometric signals are determined for the Level 0 to 1 Processing.

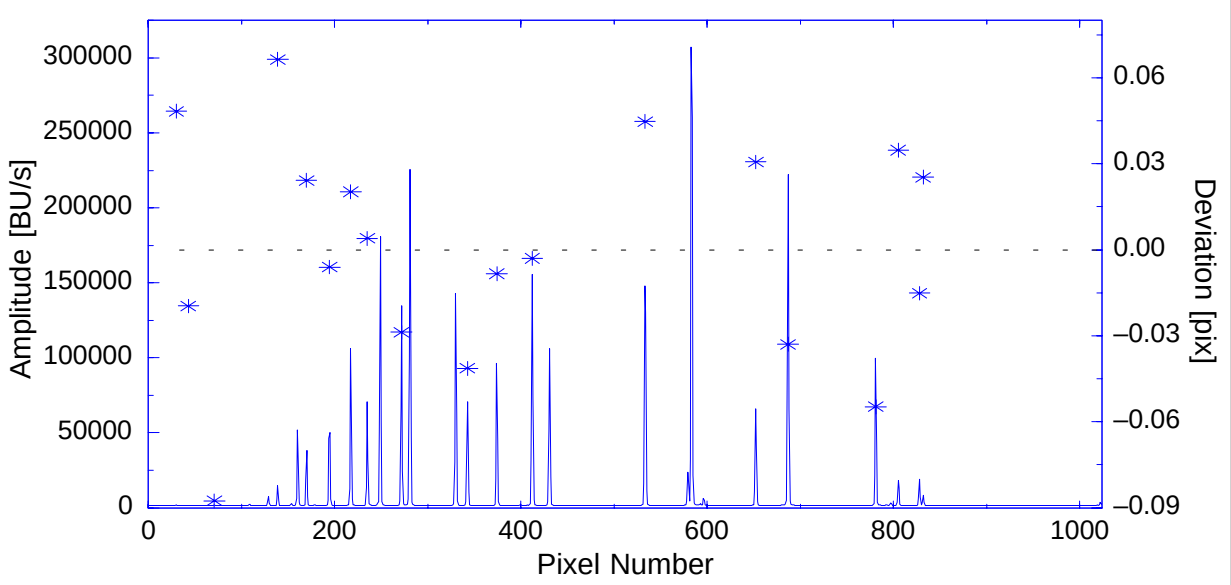

Figure 16: Typical Wavelength Uncertainty Diagram

- calculation of leakage current noise (includes electronic/readout noise and digitisation effects)
- calculation of photon shot noise induced by incident light
- estimation of the error on the fractional polarisation values
- calculation of the error on the polarisation correction factor

The first two items comprise the precision of the radiance measurements done with GOME, while the second two error sources contribute to the accuracy of the measurements. In addition, the following error sources, taken from the error budget analysis of TPD [A5] for the pre–flight calibration parameters, are also required for the accuracy calculation of GOME measurements:

- error on the BSDF
- error on the radiance sensitivity
- error on the polarisation sensitivity

2.12.3.1 Precision of the Radiometric Error

In the calculation of the leakage current, a mean value of standard deviations σ_L for the dark signals of all detector pixels is determined. If the individual detector pixel signal is denoted by *S*, then a relative measurement noise ε_R per pixel can be calculated ::

$$\epsilon_R = \sqrt{S_e + \epsilon_L^2 + \epsilon_D^2} \tag{95}$$

This includes the electronic readout noise (ε_L), shot noise and the digitisation effect (ε_D):

$$S_e = S \cdot N_e \quad \text{(shot noise term)} \tag{96}$$

$$\epsilon_L = \sigma_L \cdot N_e \tag{97}$$

$$\epsilon_D = 0.5 \cdot N_e \tag{98}$$

$$\mathrm{N_e} = 937\,[\text{electrons/BU}] \tag{99}$$

Linearity is assumed over the dynamic ranges of the detector arrays. Signal–to–noise values and relative measurement errors are then:

$$(S/N) = \frac{S_e}{\epsilon_R} \tag{100}$$

$$E_{S,\,rel} = \frac{\epsilon_R}{S_e} \tag{101}$$

2.12.3.2 Accuracy of the Solar Irradiance

The solar irradiance is given as follows:

$$I_{Sun}(\lambda_i) = \frac{S_{Sun,i}}{H^*(\lambda_i, \sigma) \cdot BSDF^*_{CU}(\lambda_i, Az, El)} \tag{102}$$

where $S_{Sun,i}$ is the signal of detector pixel *i* during a sun calibration measurement and $H^*(\lambda_i, \sigma)$ and $BSDF^*_{CU}(\lambda_i, Az, El)$ are the pre–flight calibration parameters for the radiance sensitivity and the bi–

directional scattering distribution function (BSDF) respectively. From the Gaussian error propagation equation, the error on $I_{Sun}(\lambda_i)$ is:

$$\epsilon_I = \sqrt{\left(\frac{1}{H \cdot BSDF}\right)^2 \cdot \epsilon_{S_S}^2 + \left(\frac{-S_S}{H^2 \cdot BSDF}\right)^2 \cdot \epsilon_H^2 + \left(\frac{-S_S}{H \cdot BSDF^2}\right)^2 \cdot \epsilon_B^2} \qquad (103)$$

where ϵ_{S_S}, ϵ_B and ϵ_H are the errors on the sun measurement signal, the BSDF and the radiance sensitivity respectively. This is the error for one sun measurement; however several solar spectra (3 in bands 1a/1b/2a, 12 in bands 2b/3/4) are recorded during a 30 second sun calibration period, and the reduced mean error on the solar irradiance is then:

$$\epsilon_{Iges} = \frac{\epsilon_I}{\sqrt{n}} \qquad (104)$$

where *n* is the number of measurements. This assumes that every sun spectrum in the sample has the same standard deviation (not quite the case!).

2.12.3.3 Error on Fractional Polarisation Values

Despite the complexity of the algorithms required for the determination of fractional polarisation values, explicit error calculations for the 3 overlap values and the 3 PMD values can be performed using the general Gaussian Error Propagation Law (88).

(a) *Errors on fractional polarisation in Channel Overlap Regions*

The basic formula for channel overlap fractional polarisation value is $p = (\alpha / \beta)$, where

$$\alpha = \left\{X_i \cdot \eta_{b,j}^{\Theta} \cdot (1 + \eta_{a,i}^{\Omega}) - \eta_{a,i}^{\Theta} \cdot (1 + \eta_{b,j}^{\Omega})\right\} \cdot \qquad (105)$$

$$\beta = \left\{(1 + \eta_{b,j}^{\Omega}) \cdot (1 - \eta_{a,j}^{\Theta}) - X_i \cdot (1 + \eta_{a,i}^{\Omega}) \cdot (1 - \eta_{b,j}^{\Theta})\right\} \cdot \qquad (106)$$

where $X_i = \frac{R_{p,i}}{R_{u,i}}$; $R_{u,i} = \frac{T_i}{T_j}$ *and* $R_{p,i} = \frac{S_i}{S_j}$ are signal ratios.

Here, the two channel polarisation sensitivities $\eta_{a,i}$ and $\eta_{b,j}$ are corrected for the appropriate scan mirror angle positions (Ω for sun position, Θ for earth nadir viewing scan position).

Errors on signal values S_i S_j etc. are given by the equations in section 10.3.1 above, and require only knowledge of the signal (photon shot noise) and the electron readout noise. Formulae (88) and (89) require the interpolation of all variables onto a common wavelength grid, and ***this also applies to their errors***. In what follows, it is assumed that any errors due to such interpolation have been accounted for. Absolute errors are assumed throughout. There are 5 dependent variables, and the error propagation law is :

$$\epsilon_p = \sqrt{\left(\frac{\partial p}{\partial \eta_a^{\Omega}}\right)^2 \epsilon_{\eta_a^{\Omega}}^2 + \left[\frac{\partial p}{\partial \eta_b^{\Omega}}\right]^2 \epsilon_{\eta_b^{\Omega}}^2 + \left(\frac{\partial p}{\partial \eta_a^{\Theta}}\right)^2 \epsilon_{\eta_a^{\Theta}}^2 + \left[\frac{\partial p}{\partial \eta_b^{\Theta}}\right]^2 \epsilon_{\eta_b^{\Theta}}^2 + \left(\frac{\partial p}{\partial X_i}\right)^2 \epsilon_{X_i}^2} \qquad (107)$$

The required partial derivatives are :

$$\frac{\partial \alpha}{\partial X} = (1 + \eta_a^{\Omega}) \cdot \eta_b^{\Theta} \qquad \frac{\partial \beta}{\partial X} = -(1 + \eta_a^{\Omega}) \cdot (1 - \eta_b^{\Theta}) \qquad (108)$$

$$\frac{\partial\alpha}{\partial\eta_a^\Omega} = X \cdot \eta_b^\Theta \qquad \frac{\partial\beta}{\partial\eta_a^\Omega} = -X \cdot (1-\eta_b^\Theta) \tag{109}$$

$$\frac{\partial\alpha}{\partial\eta_b^\Omega} = -\eta_a^\Theta \qquad \frac{\partial\beta}{\partial\eta_b^\Omega} = (1-\eta_a^\Theta) \tag{110}$$

$$\frac{\partial\alpha}{\partial\eta_a^\Theta} = -(1 + \eta_b^\Omega) \qquad \frac{\partial\beta}{\partial\eta_a^\Theta} = -(1 + \eta_b^\Omega) \tag{111}$$

$$\frac{\partial\alpha}{\partial\eta_b^\Theta} = X \cdot (1 + \eta_a^\Omega) \qquad \frac{\partial\beta}{\partial\eta_b^\Theta} = X \cdot (1 + \eta_a^\Omega) \tag{112}$$

$$\text{and} \quad \nabla p = \frac{(\nabla\alpha - p \cdot \nabla\beta)}{\beta} \tag{113}$$

Note that the error on X is readily computed from the definition of X, and the knowledge of signal errors :

$$\frac{\epsilon_{X_i}}{X_i} = \sqrt{\left(\frac{\epsilon_{S_i}}{S_i}\right)^2 + \left(\frac{\epsilon_{S_j}}{S_j}\right)^2 + \left(\frac{\epsilon_{T_i}}{T_i}\right)^2 + \left(\frac{\epsilon_{T_j}}{T_j}\right)^2} \tag{114}$$

(a) *Errors on fractional polarisation value from PMD Overlaps*

The basic formula for the solution of p is

$$\therefore \; S_{PMD} = \sum_{i=n_1}^{n_2} \xi_i^* \cdot p \cdot \frac{S_i}{p + (1-p) \cdot \eta_i} + Q_{PMD} \tag{115}$$

for average PMD signal S_{PMD}, detector signals S_i in the virtual sum, and the virtual pixel polarisation sensitivity ξ^*. In what follows, we assume the out–of–band stray–light correction factor Q_{PMD} is zero. The polarisation sensitivities ξ_i and η_i are interpolated to the wavelength grid of the signals S_i, and it is assumed once again that interpolation errors have been accounted for in ε_ξ and ε_η.

The first version of the polarisation correction algorithm assumes that p takes a constant value for the PMD wavelength range. The solution of (98) for p is implicit; a numerical root–finding technique is required. However, partial derivatives of p with respect to the dependent variables $\{S_{PMD}, S_i, \xi_i, \eta_i\}$ can be calculated explicitly, and the error ε_p is then readily computed from the error propagation law. The result is :

$$\epsilon_p = Y \cdot \sqrt{\epsilon_{S_{PMD}}^2 + \sum_{j=n_1}^{n_2} \left\{ \frac{\epsilon_{U_j}^2 + \beta^2 \cdot U_j^2 \cdot \epsilon_{\eta_j}^2}{\left(1 + \beta \cdot \eta_j\right)^2} \right\}} \tag{116}$$

$$\text{where } Y = \frac{p^2}{\sum_{j=n_1}^{n_2} \left\{ \frac{U_j \cdot \eta_j}{\left(1+\beta\cdot\eta_j\right)^2} \right\}} \qquad \text{and} \quad \beta = \frac{(1-p)}{p} \tag{117}$$

$$\text{also} \quad U_j = \xi_j \cdot S_j \ , \quad \text{with} \ \frac{\epsilon_{U_j}}{U_j} = \sqrt{\left(\frac{\epsilon_{\xi_j}}{\xi_j}\right)^2 + \left(\frac{\epsilon_{S_j}}{S_j}\right)^2} \tag{118}$$

The wavelength corresponding to the PMD–derived fractional polarisation is assigned according to the *signal–weighted pixel centroid* :

$$\bar{n} = \frac{1}{(n_2 - n_1 + 1)} \cdot \sum_{j=n_1}^{n_2} \frac{S_j}{\bar{S}} \cdot j \ , \quad \text{where} \ \ \bar{S} = \frac{1}{(n_2 - n_1 + 1)} \cdot \sum_{j=n_1}^{n_2} S_j \tag{119}$$

This depends only on signal values. The error on this pixel centroid is readily calculated from the error propagation law :

$$\epsilon_{\bar{n}} = \frac{1}{\bar{S} \cdot (n_2 - n_1 + 1)} \cdot \sqrt{\sum_{j=n_1}^{n_2} \cdot \epsilon_{S_j}^2 \, (j^2 + \bar{n}^2)} \tag{120}$$

This pixel centroid error readily translates to a wavelength uncertainty upon application of the wavelength calibration parameters and the usual 4th order dispersion polynomial.

(c) *Error on "seventh point" polarisation value*

The value of fractional polarisation at and below 300 nm is taken from a simple model formula. The error on this p–value will contain a systematic component due to assumptions intrinsic to the model, plus random errors due to variations in the input parameters (especially geolocation variables).

The model assumes polarisation is entirely caused by molecular scattering (Rayleigh single scatter). Comparisons with more sophisticated models show that the polarisation is accurate to better than 1% computed with the simple model. In what follows, we shall drop this systematic algorithm error. The degree of polarisation in this model is given by

$$P_{ss} = \frac{1 - \cos^2\Theta}{1 + \Delta + \cos^2\Theta} \tag{121}$$

where Δ is an anisotropy constant related to the Rayleigh depolarisation ratio ϱ [$\Delta = 2\varrho / (1-\varrho)$], and Θ is the scattering angle at a representative scatter height in the atmosphere ($\approx$30 km for wavelengths less than 300 nm). Assuming the error propagation law, we get:

$$\epsilon_{P_{ss}} = \frac{\sin\Theta}{\left(1 + \Delta + \cos^2\Theta\right)^2} \cdot \sqrt{4\,(\Delta + 2)^2 \cos^2\Theta \, \epsilon_\Theta^2 + \sin^2\Theta \, \epsilon_\Delta^2} \tag{122}$$

Errors on the scattering angle may be caused by mispointing at the satellite. One can specify the geolocation errors and follow these through the spherical geometry calculation to the final error ε_Θ, but for now we will assume ε_Θ = 0.5%. There are conflicting values of the Rayleigh depolarisation ratio ϱ in the literature; we shall take $\varrho = 0.0295$, with relative error 10%. This implies $\Delta = 0.0574$, with relative ε_Δ =10.3%). Fortunately, P_{ss} is not very sensitive to Δ.

The "seventh point" fractional polarisation p_7 is given by:

$$p_7 = \frac{(1 + P_{ss} \cos 2\chi)}{2} \ , \quad \chi \text{ defines the orientation of the polarisation plane}$$

χ depends solely on the viewing geometry so its error is not independent of that for P_{ss}. Rather than work through the complete geolocation algebra, we shall simply assume that $\varepsilon_{p7} = (1/2)\varepsilon_{Pss} \cos 2\chi$.

(d) *Interpolation errors in the polarisation correction application*

For wavelengths less than or equal to 300 nm, p is constant ($p = p_7$), and the error is ε_{p7}as given above. For that wavelength range over which polynomial interpolation is performed, the error on p is also found by interpolation through the ”perturbed” data set $\{\lambda_i, p_i+\varepsilon(p_i)\}$; there are up to 5 points (PMDs 1,2 and 3; Overlaps 2 and 3).

For the region between 300 nm and the first PMD point, a parameterisation scheme has been employed to establish intermediate values of p. One can also carry out the parameterisation using ”error perturbed” values for the determined points (in this case, the 7th point, overlap 1 and PMD1). One then computes the difference between the two parameterisations for the error. There is also a systematic error due to the parameterisation assumption – in the range 300–340 nm in particular, there are detailed small–scale features on the p–curve corresponding to O_3 Huggins signatures. The parameterisation can only reproduce the broad–scale behaviour of the polarisation in this range, but initial comparisons of parameterised polarisations against high–resolution accurate model output has shown that parameterised p–values are accurate to within 2% over this range. We shall therefore take this value as an additional error to be placed on the polarisation.

(e) *Remarks*

It has already been noted that there are serious inaccuracies for the p–determination from the first PMD1 (high out of band straylight, invalid assumption of constant p). While these problems are present in the other PMDs, they are an order of magnitude more serious for PMD1. Uncertainties in PMD1 p–determination due to these factors are large enough that they constitute a ***gross error***.

2.12.3.4 Error on Polarisation Correction Factors

The polarisation correction factor is a function of the polarisation sensitivity η and the fractional polarisation value p, as given in (123). Following the general Gaussian Error Propagation Law (88) the error on the polarisation correction factor may be given as in (126).

$$c(\eta, p) = \frac{1}{2} \cdot \frac{1 + \eta}{p \cdot (1-\eta) + \eta} \tag{123}$$

$$\frac{\partial c(\eta, p)}{\partial \eta} = \frac{1}{2} \cdot \frac{2p-1}{(p\eta-p-\eta)^2} \tag{124}$$

$$\frac{\partial c(\eta, p)}{\partial p} = \frac{1}{2} \cdot \frac{(1 + \eta) \cdot (\eta-1)}{(p\eta-p-\eta)^2} \tag{125}$$

$$\epsilon_c = \sqrt{\left[\frac{1}{2} \cdot \frac{2p-1}{(p\eta-p-\eta)^2}\right]^2 \cdot \epsilon_\eta^2 + \left[\frac{1}{2} \cdot \frac{(1 + \eta) \cdot (\eta-1)}{(p\eta-p-\eta)^2}\right]^2 \cdot \epsilon_p^2} \tag{126}$$

where ϵ_η and ϵ_p are the errors on the polarisation sensitivity and the fractional polarisation value, respectively.

2.12.3.5 Accuracy of the Earthshine Radiance

The earthshine radiance is given as follows:

$$I_{Earth}(\lambda_i) = \frac{S_{Earth,i} \cdot c_i}{H^*(\lambda_i, \sigma)} \tag{127}$$

where $S_{Earth,i}$ is the signal of detector pixel *i* during a scanning measurement, c_i the polarisation correction factor and $H^*(\lambda_i, \sigma)$ the pre–flight calibration parameter for the radiance sensitivity. Using (90) and (91) the error on $I_{Sun}(\lambda_i)$ may be given as follows:

$$\epsilon_I = \sqrt{\left(\frac{c}{H}\right)^2 \cdot \epsilon_{S_E}^2 + \left(\frac{S_E}{H}\right)^2 \cdot \epsilon_c^2 + \left(\frac{-S_E \cdot c}{H^2}\right)^2 \cdot \epsilon_H^2} \tag{128}$$

where ϵ_{S_E}, ϵ_c and ϵ_H are the errors on the scanning signal, the polarisation correction factor and the radiance sensitivity, respectively.

2.12.3.6 Accuracy of the Atmospheric Attenuation

The absolute atmospheric attenuation is given as follows:

$$BSDF_{Earth}(\lambda_i) = S_{Earth,i} \cdot c_i \cdot \frac{BSDF^*_{CU}(\lambda_i, Az, El)}{S_{Sun,i}} \tag{129}$$

where $S_{Earth,i}$ is the signal of detector pixel *i* during a scanning measurement and $S_{Sun,i}$ is the signal of the same detector pixel during a sun calibration measurement, c_i is the polarisation correction factor for the scanning measurement and $BSDF^*_{CU}(\lambda_i, Az, El)$ is the pre–flight calibration parameter for the bi–directional scattering distribution function (BSDF). Using (90) and (91) the error on $BSDF_{Earth}(\lambda_i)$ may be given as follows:

$$\epsilon_B = \sqrt{\left(\frac{c \cdot BSDF}{S_S}\right)^2 \cdot \epsilon_{S_E}^2 + \left(\frac{-S_E \cdot c \cdot BSDF}{S_S^2}\right)^2 \cdot \epsilon_{S_S}^2 + \left(\frac{S_E \cdot c}{S_S}\right)^2 \cdot \epsilon_B^2 + \left(\frac{S_E \cdot BSDF}{S_S}\right)^2 \cdot \epsilon_c^2} \tag{130}$$

where ϵ_{S_E},ϵ_{S_S}, ϵ_B and ϵ_c are the errors on the earthshine signal, the sun calibration signal, the BSDF and the polarisation correction factor respectively.

2.12.3.7 Algorithm Update (Issue 6)

Precision of the solar irradiance

The precision error on the solar irradiance (in the Level 1b product called 'relative error') used to be calculated according to (103), (104). This formula incorporates a precision error from the on–ground calibration keydata, which is not appropriate as the BSDF in the on–ground keydata is analytically specified by a polynomial function. Multiplication of detector counts by a polynomial function does not introduce a precision error.

The new formulation is (for definition of variables see section 2.12.3.1 and section 2.12.3.2):

$$\epsilon_{Iges} = \frac{\sqrt{\sum_n \left(\frac{\epsilon_R}{Ne}\right)^2}}{\sum_n S} + \epsilon_{fixed} \tag{131}$$

where n denotes index over measurement in the solar calibration sequence. The term ϵ_{fixed} accounts for noise introduced by dark signal subtraction and interpolation noise from regridding the radiance keydata to the actual wavelength grid. It is set to an estimated value of 3.e–4 .

3 Summary of Pre–flight Calibration KeyData Requirements

The following is a list of parameters and calibration results (KeyData) which are required for the algorithms of the Level 0 to 1 Processing:

- Confidence limit of Detector Values
- Relative Error Budget on KeyData Functions
- Uniform Straylight Level
- Straylight Ghost Characteristics
- Out of Band Straylight PMD's Correction
- Spectral Calibration Lines
- Scale Factors for FPA Correction
- Number of FPA Filter Coefficients Used
- Filter Coefficients for FPA Correction
- BSDF Bi–directional Scattering Distribution Function Coefficients
- Radiance Response
- Radiance Scan Mirror Dependency Factor Angles
- Radiance Scan Mirror Dependency Factor for Channel 1 to 4
- Radiance Overlap Correction Function f2 Temperatures
- Radiance Overlap Correction Function f2 for Channel 1 to 4
- Eta_Nadir Polarization Sensitivity Ratio (η)
- Chi Polarization Sensitivity Correction Factor Angles (χ)
- Chi Polarization Sensitivity Correction Factor for Channel 1 to 4 (χ)
- Ksi Polarization Sensitivity Ratio for PMD 1 to 3 (ξ)

Appendix

A Parameter Derivation for Polarisation curves

When the first overlap is present, the equations corresponding to the continuity and point conditions are (see section 2.7.3 for definitions) :

$$F_1(\lambda_0) = p_0 \qquad \text{(7th point value at } \lambda = \lambda_0) \tag{132}$$

$$F_1(\lambda_1) = p_1 \qquad \text{(first overlap value at } \lambda = \lambda_1) \tag{133}$$

$$F_2(\lambda_2) = p_2 \qquad \text{(continuity at } \lambda = \lambda_2) \tag{134}$$

$$\left.\frac{dF_2(\lambda)}{d\lambda}\right|_{\lambda=\lambda_2} = m_2 \qquad \text{(continuity of gradient at } \lambda = \lambda_2) \tag{135}$$

$$F_1(\bar{\lambda}) = F_2(\bar{\lambda}) \qquad \text{(continuity at } \lambda = \bar{\lambda}) \tag{136}$$

$$\left.\frac{dF_1(\lambda)}{d\lambda}\right|_{\lambda=\bar{\lambda}} = \left.\frac{dF_2(\lambda)}{d\lambda}\right|_{\lambda=\bar{\lambda}} \quad \text{(continuity of gradient at } \lambda = \bar{\lambda}) \tag{137}$$

Using the polynomial (parabola) and GDF functions, these 6 conditions yield respectively :

$$p_0 = \bar{p} + \frac{w_0}{4} \tag{138}$$

$$p_1 = \bar{p} + \frac{w_0 f}{(1+f)^2} \tag{139}$$

$$A_1 + A_2 \cdot x + A_3 \cdot x^2 = p_2 \tag{140}$$

$$-\left(A_2 + 2 \cdot x \cdot A_3\right) = \bar{\lambda} \cdot m_2 \tag{141}$$

$$\bar{p} + \frac{w_0 g}{(1+g)^2} = A_1 \tag{142}$$

$$A_2 = \frac{\bar{\lambda} \cdot w_0 \cdot g \cdot \beta \cdot (1-g)}{(1+g)^3} \tag{143}$$

where

$$f = exp\left[-\beta\left(\lambda_1 - \lambda_0\right)\right] \tag{144}$$

$$g = exp\left[-\beta\left(\bar{\lambda} - \lambda_0\right)\right] \tag{145}$$

$$x = 1 - \frac{\lambda_2}{\bar{\lambda}} \tag{146}$$

Eliminate A_3 from (140) and (141):

$$2 \cdot A_1 + x \cdot A_2 = \bar{\lambda} \cdot m_2 + 2 \cdot p_2 \tag{147}$$

Now use (121) and (122) for A_1 and A_2 in this:

$$2 \cdot p_2 + \bar{\lambda} \cdot m_2 = 2 \cdot \bar{p} + \frac{2\, w_0\, g}{(1+g)^2} + \frac{\bar{\lambda}\, w_0\, x\, g\, \beta\, (1-g)}{(1+g)^3} \tag{148}$$

From equations (138) and (139):

$$\bar{p} = p_0 - \frac{w_0}{4} \tag{149}$$

$$p_0 - p_1 = w_0 \left\{ \frac{1}{4} - \frac{f}{(1+f)^2} \right\} \tag{150}$$

We can eliminate w_0 and $\bar{p}$ from these last three equations to get :

$$\frac{p_0 - p_2 - \frac{\bar{\lambda} m_2}{2}}{(p_0 - p_1)} = \frac{\frac{(1-g)^2}{(1+g)^2} - \frac{2\bar{\lambda}\, x\, g\, \beta\, (1-g)}{(1+g)^3}}{\frac{(1-f)^2}{(1+f)^2}} \tag{151}$$

The unknown here is parameter β, since both f and g are functions of β through (13) and (14). This equation can be solved numerically using a standard root–finding technique. Newton–Raphson iteration was found to be somewhat unreliable for this root, and a safe Bisection root–finding routine was implemented from Numerical Recipes. With β thus calculated, the other GDF parameters w_0 and $\bar{p}$ are found easily from (149) and (19) above.

For the special case whereby the first overlap is absent, the GDF curve will be matched directly to the interpolation curve. *This is the situation that has been adopted in the first operational baseline*. Now there are only 3 unknowns, and the relevant defining conditions are :

$$F_1(\lambda_0) = p_0 \qquad \text{(7th point value at } \lambda = \lambda_0\text{)} \tag{152}$$

$$F_1(\lambda_2) = p_2 \qquad \text{(continuity at } \lambda = \lambda_2\text{)} \tag{153}$$

$$\left. \frac{dF_2 1(\lambda)}{d\lambda} \right|_{\lambda=\lambda_2} = m_2 \qquad \text{(continuity of gradient at } \lambda = \lambda_2\text{)} \tag{154}$$

which yield (notation as above):

$$p_0 = \bar{p} + \frac{w_0}{4} \tag{155}$$

$$p_2 = \bar{p} + \frac{w_0\, h}{(1+h)^2} \tag{156}$$

$$-m_2 = \frac{w_0 \cdot h \cdot \beta \cdot (1-h)}{(1+h)^3} \tag{157}$$

where now

$$h = exp\left[-\beta\, (\lambda_2 - \lambda_0)\right] \tag{158}$$

This system yields the following equation for β, which again must be solved numerically by the Bisection technique:

$$(1 + h) \cdot (1-h) + h \cdot \beta \cdot K = 0 \quad (159)$$

where K is a known value, given by :

$$K = \frac{(p_0 - p_2)}{m_2} \quad (160)$$

Equation (28) has no solution if K > 0. In this case, we re–introduce the intermediate point $\overline{\lambda}$, and stipulate a parabola between $\overline{\lambda}$ and λ_2. With the 7th point, 2 continuity equations at $\overline{\lambda}$ and 2 more continuity equations at λ_2, we require one more condition. This is achieved by allowing the parabola to have a turning point at a wavelength halfway between $\overline{\lambda}$ and λ_2. The equations can be written down as above, and the solution for β found again by numerical root–finding. In this way, a solution is always found for the parameters, provided the 7th point is present. If p_0 is not valid, there is no GDF parameterisation.

B Cross Correlation for spectral calibration

The cross–correlation algorithms as applied for wavelength calibration is described within this document in section 2.7. For completeness, we will give here the input parameters for the execution of the cross–correlation algorithm.

The window segments used in each measurement channel for the application of cross–correlation are as follows:

Window #	Channel #1 Start pixel #	Channel #2 Start pixel #	Channel #3 Start pixel #	Channel #4 Start pixel #
Number of pixels	44	44	50	50
1	81	43	135	151
2	84	70	150	157
3	180	199	152	234
4	183	202	155	255
5	195	226	220	350
6	294	253	355	352
7	297	394	412	378
8	300	397	436	425
9	408	502	439	597
10	411	505	472	599
11	531	523	591	755
12	534	526	599	758
13	–	703	658	779
14	–	706	660	782
15	–	730	765	902

Window #	Channel #1 Start pixel #	Channel #2 Start pixel #	Channel #3 Start pixel #	Channel #4 Start pixel #
16	–	757	768	907
17	–	–	824	–
18	–	–	827	–
19	–	–	915	–
20	–	–	918	–

Apodisation:

- Shape: sinusoidal
- Apodisation fraction: 0.25

Continuum correction:

- Polynomial function
- Degree: 2

Error estimation parameter:

- Half maximum of Fourier–transform (parameterisation): 10

Overall fitting polynomial for calibration correction:

- Channel #1 and channel #2: linear
- Channel #3 and channel #4: quadratical

C Algorithm updates for the extraction software (Doc. Issue 6)

C.1 Apply Residual Offset Correction

This correction is based on the CHEOPS-GOME study of signal background [C8]. The results of this study are obtained by using a method where the correction is made after the correction for straylight.

One may argue that, from a point of principle, it might be better to apply this correction *before* the calculation of straylight, since straylight correction uses the amount of photons in the channel, which may be obtained from the signal corrected for dark signal and corrected for Peltier offset. Note, that this could only be implemented in an iterative approach, because the calculation of the amount of residual offset can only be done after straylight is removed. As the straylight correction is based on the total intensity in the channel, and as the number of counts in the Peltier correction is small compared to the total intensity, the error made in the straylight correction by not applying the residual offset correction first is expected to be negligible. Therefore, we stick to the procedure described in [C8] and implement the correction *after* the correction for straylight.

The offset is only corrected for long integration times, which limits it to channel 1a. Here we use the signal of the “straylight band 1a” just before the beginning of the nominal band 1a.

1. The algorithm uses the first 20 pixels of the straylight band 1a, starting at pixel 206 in channel 1 (we count from 0 in this description).
2. These are sorted w.r.t. signal intensity.
3. The intensity of the *k*-th sorted pixel is used to further correct the measurements.
   We take $k = 9$ (i.e. the 10th sorted pixel when counting from 1)

The lowest intensity is not used because of noise, higher intensities are not used because they may be affected by cosmic rays or by straylight.

We then get for the signal S of pixel i in channel 1a:

$$S_{i\_1a}(\text{corrected}) = S_{i\_1a} - S_{k\_\text{stray}}(\text{sorted}) \tag{161}$$

C.2 Correct seasonal variation in BSDF

This correction is based on the CHEOPS-GOME study of seasonal variability of the diffuser BSDF [C7]. The BSDF is described as polynomial function of solar azimuth angle on the diffuser.

Two approaches are presented in the study:

1. using a smoothed BSDF which is a 3rd order polynomial fit in wavelength; this is provided for all channels
2. a spectrally resolved BSDF, which can be provided for channels 1,3,4 but not for channel 2; this is because the frequent changes in etalon structure in channel 2 destroys the information in the analysis [C7]. As it may be that the spectral structure correlates with the detector etalon (although this is not obvious from the data) there may also be a real limitation for channel 2 here.

The CHEOPS-GOME study on diffuser BSDF [C7] proposes to fit the azimuth dependence with a polynomial. However, the data volume concerned is not a problem for the GDP extraction software. We can obtain a higher accuracy by using a look-up table (LUT) for a number of azimuth angles.

For the smoothed BSDF, the wavelength dependence for each azimuth angle is calculated as an N^{th} (default: 3^{rd}) order polynomial; the coefficients of this polynomial are in the LUT with dimensions (M, N+1) where the number of rows, M, is the number of azimuth angles.

For an azimuth angle $\alpha_{azi}(m)$, the smoothed BSDF as function of measurement wavelength λ_i is for each pixel i is given by:

$$BSDF(\lambda_i, \alpha_{azi}(m)) = \sum_{j=0}^{N} LUT_{mj} \cdot (\lambda_i - \lambda_c)^j \quad (162)$$

where λ_c is a 'central wavelength' in the channel (to keep the numerical values of LUT_{mj} in a smaller range of magnitude).

For an azimuth angle $\alpha_{azi}(m) \le \alpha_{azi} \le \alpha_{azi}(m+1)$, the BSDF is calculated for table indices (m) and (m+1) and then linearly interpolated to α_{azi}.

For the spectrally resolved BSDF, a LUT is provided with dimensions (N_pix, M) where the number of columns, M, is the number of azimuth angles and N_pix is the number of pixels in a channel. The data are presented together with a grid of N_pix wavelengths λ_j. The processing is here:

1. calculate for the applicable azimuth angle the BSDF on the wavelength grid :

$$BSDF(\lambda_j, \alpha_{azi}(m)) = LUT_{jm} \quad (163)$$

2. interpolate $BSDF(\lambda_j)$ to the wavelength grid of the measurement using Akima interpolation

For an azimuth angle $\alpha_{azi}(m) \le \alpha_{azi} \le \alpha_{azi}(m+1)$, the BSDF is calculated for table indices (m) and (m+1) and then linearly interpolated to α_{azi}.

As the analysis of the spectrally resolved BSDF has been limited to wavelength windows, one must accept a jump in level 1 radiances at the window edges after correction. The current windows are: 250 - 312 nm, 315-393 nm, 415-590 nm, 620-790 nm; perhaps this may be slightly enlarged in the final data set. Note, that the current windows cover all operationally retrieved trace gas species in the GDP. Also the ozone profile retrieval wavelengths used in the CHEOPS project are covered, with a possible exception in the 312-315 nm range (t.b.d.).

Finally, it is noted that the spectrally resolved (or 'unsmoothed') BSDF should be used with care, especially in channels 3 and 4. In these channels there is a strong dependence of high-frequency structures (probably from interference in the dichroic filter) on solar azimuth angle on the diffuser. Especially after the problems with the ERS-2 gyros after 2001, this angle as given on the Level 1 product has some uncertainties. Furthermore, the features are only defined w.r.t. azimuth angle 0, not in an absolute radiometric sense.

C.3 Improvements of the polarisation correction algorithm

The PMDs measure polarisation for wavelengths $\lambda \ge 380$ nm. For shorter wavelengths, a theoretical model is required. Up to a wavelength $\lambda_{SS} \approx 300$ nm, the polarisation can be taken as constant and equal to the value for single scattering [C3]. In the GDP, the gap between these wavelength regions is bridged using a parameterisation given by the 'Generalised Distribution Function' (GDF):

$$F_1(\lambda) = \bar{P}(\bar{\lambda}) + \frac{w_0 \cdot \exp(-[\lambda - \lambda_{SS}] \cdot \beta)}{\{1 + \exp(-[\lambda - \lambda_{SS}] \cdot \beta)\}^2} \tag{164}$$

Here $\{\bar{P}, w0, \beta\}$ are parameters that characterise the GDF; they must be found to fit the given interpolation points. In the previous GDP extraction software, λ_{SS} was taken as constant (300 nm) and the GDF was calculated up to a wavelength $\bar{\lambda} = 325\,\text{nm}$. The value $\bar{P}(\bar{\lambda})$ was calculated in such a way that the GDF is continuous in gradient with a spline interpolation (Akima interpolation [C2]) through the PMD polarisation values.

The previous GDP formulation has the drawback that $\bar{P}(\bar{\lambda})$ depends strongly on information from PMD-2 and PMD-3. However, the polarisation in the visible is physically decoupled from the polarisation in the UV and hence information is used which is not applicable. In the SCIAMACHY ATBD [C1], the GDF is initially connected to λ_A, the effective wavelength of PMD-1. Using the observation that the steepest gradient of the GDF is always near a fixed wavelength λ_m, an analytical solution to the GDF parameters is possible [C1]. These are sequentially calculated using:

1. $\beta = (\ln(2 + \sqrt{3}))/(\lambda_m - \lambda_{SS})$
2. $\bar{P}(\bar{\lambda}) = (P_A - P_0 \cdot g_A(\beta))/(1 - g_A(\beta))$
3. $w_0 = 4 \cdot (P_0 - \bar{P}(\bar{\lambda}))$

   where P_0 is the theoretical polarisation at λ_{SS} (in this document also named ‘7th point‘), P_A is the polarisation of PMD-1, and $g_A(\beta)$ is an auxiliary function defined by:

$$g_A(\beta) = 4 \cdot \frac{\exp(-[\lambda_A - \lambda_{SS}] \cdot \beta)}{\{1 + \exp(-[\lambda_A - \lambda_{SS}] \cdot \beta)\}^2} \tag{165}$$

An improved algorithm has been proposed by Schutgens and Stammes [C5]. Here, λ_{SS} and λ_m are not fixed values, but they are parameterised as function of airmass, ground albedo and ozone content:

$$\lambda_{SS} = a_0 + \sum_{i=1}^{2} a_i/M^i + b_i \cdot \left(\frac{VCD}{VCD_0} - 1\right)^i \tag{166}$$

$$\lambda_m = \sum_{i=0}^{2}\sum_{j=0}^{2} c_{ij}/M^i \cdot A^j + \sum_{i=1}^{2} d_i \cdot \left(\frac{VCD}{VCD_0} - 1\right)^i \tag{167}$$

where M is the airmass, A is the surface albedo, and VCD is the vertical ozone column [DU]; VCD_0 = 345.8 DU.

The parameters have been fitted from model calculations as:

```
a = [ 308.68, -29.10, 11.46 ]
b = [     0.,   7.58, -4.26 ]
    [316.43, -41.89, 29.49 ]
c = [  0.33,  -0.06,  0.66 ]
    [ -1.11,   0.56, -3.46 ]
```

where index *i* runs over column and index *j* over row

```
d = [0., 7.20, -4.08 ]
```

In the GomeCal algorithm from KNMI [C6], this parameterisation has been implemented using the Fortuyn-Kelder ozone climatology [C4] to calculate *VCD*, whereas the albedo is calculated by comparing the measured GOME reflectance at 380 nm with a pre-calculated lookup table. We have followed this implementation here.

An improvement to GomeCal may be made in the airmass calculation. In Ref. [C5], [C6] the formula for the plane-parallel case is used. This limits the calculation to solar zenith angles SZA < 85°. We use a formula for spherical geometry from H. Eskes (private communication):

$$M = \frac{1}{\cos(VZA)} + \frac{\sqrt{\cos(SZA)^2 + (h/R)^2 + 2(h/R)} - \cos(SZA)}{h/R} \tag{168}$$

where *VZA* is the nadir viewing angle, *h* is the top of atmosphere height, and *R* is the Earth radius. This formula gives physical results for the airmass even for SZA > 90°, although the numerical values then become strongly dependent on the assumed top of atmosphere height. We will use $h = 60$ and $R = 6300$.

The parameters *a-d* in the equations above have been calculated for SZA < 75° and 220 < *VCD* < 440. We find that using our airmass calculation, these parameters still yield plausible results for SZA < 95° and 100 < *VCD* < 600 DU. There are thus no practical limitations in applying the parameters in the GDP software.

Reference [C5] provides parameterisations of λ_{SS} and λ_m, but does not mention how the value of $\bar{P}(\bar{\lambda})$ has to be obtained, or how the GDF has to be connected to the PMD polarisation points. In the KNMI algorithm, the GDF parameters are calculated as above, i.e. the GDF is connected to the point (λ_A ,P_A). The parameterisation is then used between λ_{SS} and $\bar{\lambda} = 325$ and Akima interpolation is used from $\bar{\lambda} = 325$ onwards.

The GDF is a slowly decreasing function for wavelengths above ~320 nm. Using it up to 325 nm therefore prescribes a rather flat function, and a continuous connection in gradient forces also the Akima interpolation to be pretty flat until the polarisation point from PMD-1. This behaviour is not always supported by simulations of polarisation, especially above clear surfaces with low albedo (high polarisation). Here polarisation shows a local minimum near 320 nm (depending on airmass) and a local maximum near 400-450 nm.

We prefer to use the method from the SCIAMACHY ATBD [C1]. Here, the GDF parameterisation is only used between λ_{SS} and a wavelength ($\lambda_0 + \Delta\lambda_{GDF}$) which is chosen such that it falls still in the strongly descending part of the GDF. Continuation in gradient will force it to follow the GDF further downwards as would the parameterisation do, but it gives the algorithm a bit more freedom to shape the curve according to the gradient obtained from PMD measurements around λ_A. In that sense it is a bit of an intermediate between the previous GDP algorithm and the prescription of a spectrally flat function. A difference with the previous GDP algorithm is that the GDF parameters are not *determined* by all PMD measurements: they are determined only by the theoretical polarisation and the measurement of PMD-1, but the polarisation curve can smoothly shape itself between ($\lambda_0 + \Delta\lambda_{GDF}$) and λ_A using the gradient between PMD-1 and PMD-2. To keep close to the GomeCal approach, we will use a value of $\Delta\lambda_{GDF} = 25$ nm; this yields on average a connecting point near 325 nm.

A limitation needs to be implemented against using unphysical polarisation values from PMD-1. Unphysical are polarisation fractions where the Stokes fraction *Q/I* of the theoretical point is smaller than, or has a different sign as, the one from PMD-1. These are given by the conditions:

$$|P_0 - 0{,}5| < |P_A - 0{,}5| \tag{169}$$

$$(P_0 - 0{,}5) \cdot (P_A - 0{,}5) < 0 \tag{170}$$

Analysis of GOME PMD-1 data has shown that, on average, the degree of polarisation at PMD-1 is half that of the theoretical single scatter value.

In case of unphysical PMD-1 values, replace P_A by this statistical value:

$$P_A \rightarrow 0{,}5 \cdot (1 - \Delta + 2\Delta \cdot P_0) \tag{171}$$

with $\Delta = 0{,}5$

Finally, as clarification, a note on the definition of the polarisation coordinate frame.The PMDs measure intensity polarised parallel to the spectrometer's slit. If we define the plane of polarisation to be coincident with the local meridian plane (through viewing direction and through nadir; this coincides with the scanning plane - see Figure 17) then the PMDs measures the $-Q$ Stokes Intensity. In the GOME calibration algorithm, the polarisation is described using the fraction of light polarised parallel to the spectrometer's slit. This polarisation parameter is denoted p and is a simple transformation of the Stokes fraction Q/I given by: $p = 0{,}5 \cdot (1 - Q/I)$

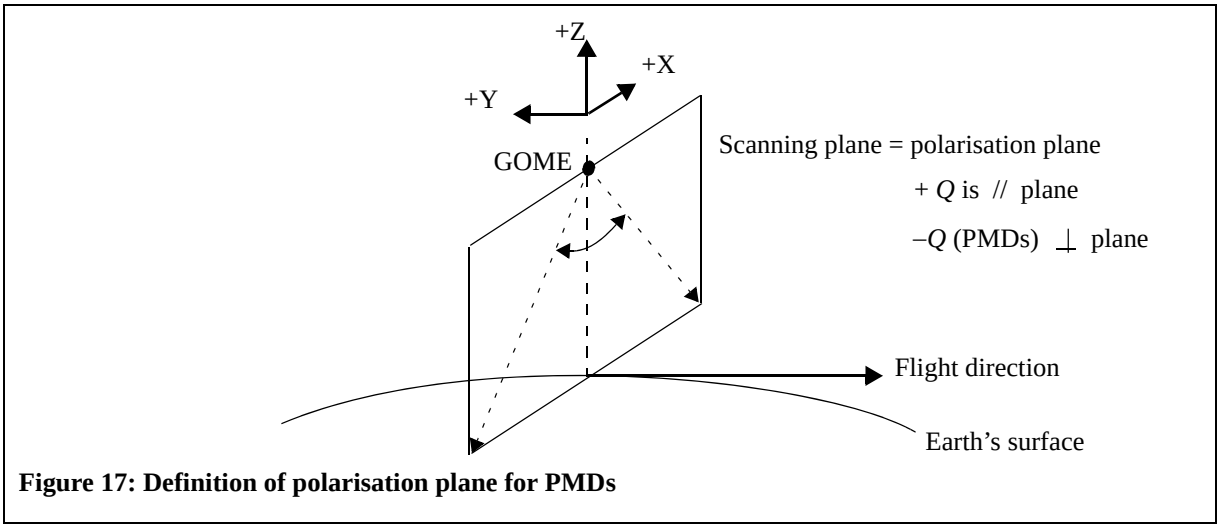

Figure 17: Definition of polarisation plane for PMDs

C.4 Correct degradation in reflectivity

The post-processing algorithm used by KNMI [C6] corrects the reflectivity (sun-normalised radiance) on the level 1 data on two levels:

1. A correction is made on the on-ground calibration of the BSDF (wavelength dependent correction, fixed in time)
2. A correction is made for time-dependent degradation of the BSDF (on the ratio of Radiance to Irradiance)

For the correction we will use the results from the CHEOPS-GOME degradation study, which has been performed by SRON [C9]. Currently implemented is an option -F in the GDP extraction software, which implicitly makes the correction 1) above. It is foreseen to add an option (probably with flag -E) which only corrects for the degradation w.r.t. the begin of the in-orbit situation.

The correction is only available for channels 1 and 2. The degradation is given as a LUT for each wavelength per channel (using a fixed wavelength scale), for a number (~70) of dates. The first index in the LUT describes the wavelength dependence, the second index refers to the sequence of dates (given as Julian Date).

Let the LUT have dimensions (N_pix, M_day). Application of this LUT is as follows (the value of the degradation spectrum calculated from the LUT - after an interpolation - is denoted by *D*).

1. On the LUT's wavelength grid we get $D(\lambda_j, t) = LUT(\lambda_j, t)$, where for measurement day t between t_m and t_{m+1}, a linear interpolation is used between the degradations $D(\lambda_i, t_m)$ and $D(\lambda_i, t_{m+1})$.

2. The level 1 earthshine radiances (not solar irradiance) for each day t are divided by the corresponding degradation $D(\lambda_i, t)$., after (Akima) interpolation of D to the wavelength grid of the measurement.

Note, that currently there is an option in the GDP extractor to apply degradation correction factors, derived from irradiance measurements, on both, radiances and irradiances. When this irradiance degradation is then applied on the reflectivity degradation, we get a radiance degradation. This is why *D* is applied on the radiance - although for calculation of the reflectivity one might also apply it on the BSDF instead.

C.5 References C

[C1] SCIAMACHY Level 0 to 1c Processing - Algorithm Theoretical Basis Document, ENV-ATB-DLR-SCIA-0041, Issue 2, DLR, 14.12.2001

[C2] Hiroshi Akima, "A new method of interpolation and smooth curve fitting based on local procedures", J. ACM, Vol. 17(4), 1970, 589-602

[C3] Piet Stammes, "The seventh point polarisation algorithm, Internal Report" (GOME and SCIAMACHY Data & Algorithm group), KNMI De Bilt, 1994

[C4] Fortuin, J.P.F. and Kelder, H.M."An ozone climatology based on ozonezonde and satellite measurements," J. Geophys. Res. 103, p.31,709, 1998

[C5] N.A.J. Schutgens, P. Stammes, "Parametrisation of earth's polarisation spectrum from 290 to 330 nm", J.Quant.Spectr.Rad.Transf. 75, p.239, 2002

[C6] J.H.G.M. van Geffen, "Documentation of the software packeage GomeCal (Version 1.0)", KNMI Technical Report TR-255, 2003

[C7] S. Slijkhuis, "Study on Seasonal Effects on the ERS-2/GOME Diffuser BSDF", DLR report CH-TN-DLR-GO-0001, Issue 1, 9.5.2004

[C8] S. Slijkhuis, "Study on Correction of Signal Background on the ERS-2/GOME Level 1 data", DLR report CH-TN-DLR-GO-0002, Issue 1, 3.9.2004

[C9] J.M. Krijger, I Aben, J. Landgraf, "GOME: scan-mirror degradation correction", ESA Atm.Sci.Conf. , May 2006, in press