	GlobAlbedo Technical Specifications **Authors:** J-P Muller, P. Lewis, P. North, C. Brockmann, U. Krämer, J. Fischer

TITLE: Technical Specifications

Document Number: GlobAlbedo_TS_D02_V1_2 Date 29/05/10

Distribution:

ESA	O. Leonard	X
	S. Pinnock	X
University College London	J-P. Muller	X
	T. Kennedy	X
	P. Lewis	X
Swansea	P. North	X
	A. Heckel	X
Free University of Berlin	J. Fisher	X
	L. Guanter	X
Brockmann Consult	C. Brockmann	X
	N. Fomferra	X
	U. Krämer	X

Author:	Prof. J-P Muller	Date:	29.05.10
Manager/Project Office		Date:	9/6/10
PA:		Date:	9TH JUNE 2010

CHANGE RECORD

ISSUE	DATE	PAGES CHANGED	COMMENTS
1.1	11/01/10	All new	First Issue
1.2	29/05/10	All	Updates to include PM1 RIDs

CONTENTS

1 Introduction

This document relates to Task 1 of the ITT (RD-1).

"*As a response to the Requirements Baseline, the Contractor shall write a Technical Specification specifying the following:*

- *Content of the GlobAlbedo products, including a detailed specification of quality flags and uncertainties.*
- *For each instrument, justification and specification of the algorithms to be prototyped for atmospheric correction, snow and cloud detection, BRF modelling, angular and spectral integration, time and space compositing, and product merging.*
- *Approach to implementing uniform retrieval methods to each of the instruments, intercalibration and merging.*
- *Software requirements specification.*
- *Approach to minimising dependency on non-European data sets.*
- *Proposed product formats and delivery mechanisms*
- *Planned test data set for algorithm prototyping.*
- *Detailed plan for validation of the albedo products, including a specification of the in-situ and ancillary data to be used.*
- *Limitations of the algorithms and validation techniques."*

The following sections provide the necessary inputs to meet this requirement.

2 Applicable and Reference documents

2.1.1 Applicable documents

AD1 European Cooperation for Space Standardization: Space Engineering Software, ECSS-E-ST-40C (6 March 2009), available from http://www.ecss.nl

2.1.2 Reference Documents

RD1 EOEP-DUEP-EOPS-SW-09-0001 SoW Statement of Work for DUE-GlobAlbedo Project, Version 1.0, April 2009

Baret, F., O. Hagolle, B. Geiger, P. Bicheron, B. Miras, M. Huc, B. Berthelot, F. Niño, M. Weiss, O. Samain, J.-L. Roujean and M. Leroy, (2007), LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. Remote Sensing of Environment, 110, 275-286.

Berthelot, B. and Dedieu, G.,(2000), Operational method to correct VEGETATION satellite measurements from atmospheric effects. Proceedings of IGARSS 2000. IEEE 2000 International. 01/02/200002/2000; 2:831-833 vol.2. DOI: 10.1109/IGARSS.2000.861718.

Birks et al, (2007), Improvements to the AATSR IPF relating to Land Surface Temperature Retrieval and Cloud Clearing over Land, RUTHERFORD APPLETON LABORATORY

Cao, C., S. Uprety, J. Xiong, A. Wu, P. Jing, D. Smith, G. Chander, N. Fox, and S. Ungar, 2010: Establishing the Antarctic Dome C Community Reference Standard Site towards Consistent Measurements from Earth Observation Satellites. Canadian Journal of Remote Sensing, in press.

Gobron, N., B. Pinty, F. Mélin, M. Taberner, M. M. Verstraete, M. Robustelli and J.-L. Widlowski (2007), Evaluation of the MERIS/ENVISAT FAPAR product. Advances in Space Research, 39, 105-115.

Gomez-Chova, Luis, G. Camps-Valls, J. Calpe, J. Munoz, and J. Moreno, 2009, MERIS/AATSR Synergy Algorithms for Cloud Screening, Aerosol Retrieval, and Atmospheric Correction; ESRIN Contract No. 21090/07/I-LG

Govaerts, Y., Wagner, S. , Lattanzio, A. and Watts, P. (2010), Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 1. Theory. J. Geophys. Res., 115, D02203,doi:10.1029/2009JD011779.

Greuell, W., J. Oerlemans (2004), Narrowband-to-broadband albedo conversion for glacier ice and snow: equations based on modeling and ranges of validity of the equations. Remote Sensing of Environment, 89, 95-105.

Grey, W.M.F., North., P.R.J., Los, S.O., and Mitchell, R.M. (2006a), Aerosol optical depth and land surface reflectance from multi-angle AATSR measurements: Global validation and inter-sensor comparisons. IEEE Transactions on Geoscience and Remote Sensing, 44(8): 2184 – 2197.

Grey, W.M.F, North., P.R.J. and Los, S.O, (2006b), Computationally efficient method for retrieving aerosol optical depth from ATSR-2 and AATSR data, App. Optics, 45(12), 2786-2795.

Guanter, L.; L. Gomez-Chova, J. Moreno, (2008), Coupled retrieval of aerosol optical thickness, columnar water vapor and surface reflectance from ENVISAT/MERIS data over land. Remote Sensing of Environment, 112, 2898-2913.

Hess, M., P. Koepke, and I. Schult, (1998). Optical Properties of Aerosols and Clouds: The Software Package OPAC. Bull. Amer. Meteor. Soc., 79, 831–844.

Hess, M., P. Koepke, and I. Schult, (1998). Optical Properties of Aerosols and Clouds: The Software Package OPAC. Bull. Amer. Meteor. Soc., 79, 831–844.

Hu, B., Lucht, W. and Strahler, A. H. (1999), The interrelationship of atmospheric correction of reflectances and surface BRDF Retrieval: a sensitivity study. IEEE Trans. Geosci. Rem. Sens. 37(2), 724-738.

Levy, R.C., Remer, L., Mattoo, S., Vermote, E. and Kaufman, Y.J. (2007), Second generation algorithm for retrieving aerosol properties over land from MODIS spectral reflectance. J. Geophys. Res., 112, D13211,doi:10.1029/2006JD007811.

Liang, S. (2000), Narrowband to broadband conversions of land surface albedo. I Algorithms. Remote Sensing of Environment, 76, 213-238

Lissens, Gil, Pieter Kempeneers and Freddy Fierens, 2000, Development of a cloud, snow and cloud shadow mask for VEGETATION imagery. Contract no. SP67008 Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC) in the framework of the SPOT4 - VEGETATION preparatory programme. http://www.spot-vegetation.com/vegetationprogramme/ Pages/vgtprep/vgt2000/lissenscloud.pdf

Lyapustin, A., Y. Wang, J. Martonchik, J.L. Privette, B. Holben, I. Slutsker, A. Sinyuk, and A. Smirnov, (2006), Local analysis of MISR surface BRF and albedo over GSFC and Mongu AERONET sites, IEEE Trans. Geosci. Rem. Sens., 44, 1707-1718.

Maisongrande, P. et al. (2000), A new algorithmic concept for atmospheric correction

of surface reflectances delivered by the vegetation system. In proceedings, VEGETATION – 2000, Lake Maggiore - Italy, 3-6 April 2000.

North, P.R.J. (2002a), Estimation of fAPAR, LAI and vegetation fractional cover from ATSR-2 imagery, Remote Sensing of Environment 80:114-121.

North, P.R.J. (2002b). Estimation of aerosol opacity and land surface bidirectional reflectance from ATSR-2 dual-angle imagery: operational method and validation. Journal of Geophysical Research 107 (D12), DOI: 10.1029/2000JD000207: 1-11.

North, P. R. J., Plummer, S.E., Briggs, S.A. and Settle, J.J. (1999), Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery. IEEE Trans. Geosci. Rem. Sens. 37, 526–537.

Remer, L. Y. Kaufman, D. Tanre, S. Mattoo, D. Chu, J. Martins, R.-R. Li, C. Ichoku, R. Levy, R. Kleidman, T. Eck, E. Vermote, and B. Holben, (2005), The MODIS Aerosol Algorithm, Products and Validation, Journal of the Atmospheric Sciences, 62, 947-973.

Rodgers, C.D. (2000), Inverse methods for atmospheric sounding: Theory and practice', World Scientific Publishng Ltd.

Roman, M. O., C. B. Schaaf, C. E. Woodcock, A. H. Strahler, X. Yang, R. H. Braswell, P. S. Curtis, K. J. Davis, D. Dragoni, M. L. Goulden, L. Gu, D. Y. Hollinger, T. E. Kolb, T. P. Meyers, J. W. Munger, J. L. Privette, A. D. Richardson, T. B. Wilson, and S. C. Wofsy, 2009: The MODIS (Collection V005) BRDF/albedo product: Assessment of spatial representativeness over forested landscapes. Remote Sensing of Environment, 113, 2476-2498.

Santer, R. et al, 1997, MERIS ATBD Pixel identification, http://envisat.esa.int/instruments/meris/pdf/atbd_2_17.pdf

Santer, R., D. Ramon, J. Vidot, E. Dilligeard, (2007), A surface reflectance model for aerosol remote sensing over land, International Journal of Remote Sensing, 28(3-4), 737-760.

von Hoyningen-Huene, W. et al., (2006), Simultaneous determination of aerosol and surface characteristics from top-of-atmosphere reflectance using MERIS on board of Envisat, Advances in Space Research, 37, 2172-2177.

3 Acronyms and Abbreviations

3.1 Acronyms

Item	Definition
AD	Applicable document
AERONET	Aerosol Robotic Network
AR	Acceptance Review
ARM	US Atmospheric Radiation Measurement program
ATBD	Algorithm theoretical basis document
ATD	Acceptance Test Document
ATSR	Along-track scanning radiometer (ESA)
AATSR	Advanced ATSR (ESA)

BHR	Bi-hemispherical reflectance
BRDF	Bi-directional reflectance distribution function
BRF	Bi-directional reflectance factor
BSRN	Baseline surface radiation network
CDR	Critical Design Review
COTS	Commercial off-the-shelf (software packages)
CTIV	Centre de Traitement des Images VEGETATION
DGVM	Dynamic global vegetation model
DUE	Data User Element of the ESA Earth Observation Envelope Programme
ECSS	European Cooperation for Space Standardization
EO	Earth observation
FAPAR	Fraction of absorbed photosynthetically active radiation
FTP	File Transfer Protocol
GCM	General circulation model
GSA	Geostationary Meteorological Albedo
HDF	Hierarchical Data Format
HDRF	Hemispherical-directional reflectance factor
IPCC	Intergovernmental Panel on Climate Change
ITT	Invitation to tender
KO	Project kick-off
LAI	Leaf area index
LTO	Linear Tape-Open (magnetic media standard)
MERIS	Medium resolution imaging spectrometer (ESA)
MISR	Multiangle imaging Spectroradiometer (NASA)
MODIS	Moderate resolution imaging spectroradiometer (NASA)
NetCDF	Network Common Data Format
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical weather prediction
OLCI	Ocean and land colour instrument
PDF	Probability density function
PDR	Preliminary Design Review
PM	Progress meeting
PNG	Portable Network Graphics
PUG	Product User Guide
QA	Quality assessment
QR	Qualification Review
RB	Requirements Baseline document
RD	Reference document
RID	Review item discrepancy
SIN	Sinusoidal Grid
SLSTR	Sea and land surface temperature radiometer
SoW	Statement of Work
SUM	Software User Manual
SURFRAD	Surface Radiation Network
SVP	Software Validation Plan
TOA	Top of atmosphere

TR	Technical requirement
TS	Technical Specification document
UCM	User consultation meeting
URL	Uniform resource locator

3.2 Abbreviations

4 Description of GlobAlbedo Processing chain

4.1 Content of the GlobAlbedo products, including a detailed specification of quality flags and uncertainties

4.1.1 Surface Directional Reflectance (SDR)

To be computed per pixel at all wavebands for each instrument (ATSR-2, AATSR, MERIS and SPOT-VGT), with the exception of channels 11 and 15 for MERIS due to atmospheric absorption. In addition we retrieve aerosol optical depth (AOD) on a 10km grid, and record the aerosol model used, and uncertainty in AOD.

For all data sets, the variance-covariance matrices calculated as a result of SDR and necessary for the subsequent calculation of uncertainties should not be recorded as a product due to high storage requirements (n^2, where n is no. of bands). Instead, codes will be provided as part of the processing chain to calculate these terms as a function of the imaging geometry and atmospheric parameters.

Note. These data will be created as part of the processing, but it is not possible to commit to storing and delivering them as products unless additional project resources are made available (ESA).

ATSR-2 L2 SDR Reflectance Product

ID: ATSR2_SDR_*i*

Name: ATSR-2 L2 Surface Directional Reflectance in channel *i*, with *i* = 1 to 8, where 1-4 denote nadir reflectance, and 5-8 denote along-track reflectance

Standard Name: *TBD*

Description: The product provides atmospherically corrected surface reflectances for 4 wavebands and 2 view directions. Inputs are 8 channels from ATSR-2 - all solar reflective channels channels i.e 550nm, 660nm, 870nm and 1610nm. The grid of the product is L1b ATSR-2 reference grid providing the ortho-geolocation information for each pixel. The resolution is ~1km (nadir) and ~2km (along-track).

Units: *Dimensionless*

Uncertainty measure: Code will be provide to estimate error covariance matrix for all SDRs retrieved, taking as input the uncertainty in AOD, the SDR and the solar/view geometry. Uncertainty is stored only as 1 s.d. of error in retrieval of AOD.The covariance should not be recorded as a product due to high storage requirements (n^2, where n is no. of bands).

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Selected L1b Flags (TBD)

Time period: 1995-2002 [NB Further processing of overlap period 2002-2007 possible as an option. Note also there is a data gap in ATSR-2 for 1/1/96 – 1/7/96]

Size: TBD

AATSR L2 SDR Reflectance Product

ID: AATSR_SDR_*i*

Name: AATSR L2 Surface Directional Reflectance in channel *i*, with *i* = 1 to 8, where 1-4 denote nadir reflectance, and 5-8 denote along-track reflectance.

Standard Name: ***TBD***

Description: The product provides atmospherically corrected surface reflectances for 4 wavebands and 2 view directions. Inputs are 8 channels from ATSR-2 - all solar reflective channels channels i.e 550nm, 660nm, 870nm and 1610nm. The grid of the product is L1b ATSR-2 reference grid providing the ortho-geolocation information for each pixel. The resolution is ~1km (nadir) and ~2km (along-track).

Units: *Dimensionless*

Uncertainty measure: Code will be provide to estimate error covariance matrix for all SDRs retrieved, taking as input the uncertainty in AOD, the SDR and the solar/view geometry. Uncertainty is stored only as 1 s.d. of error in retrieval of AOD.

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Selected L1b Flags (TBD)

Time period: 2002-2010

Size: TBD

MERIS L2 SDR Reflectance Product

ID: MERIS_SDR_*i*

Name: MERIS L2 Surface Directional Reflectance in channel *i*, with *i* = 1 to 13, i.e. all MERIS channels excluding O_2 absorbing channel (761.4 nm) and water vapour channel (900nm).

Standard Name: ***TBD***

Description: The product provides atmospherically corrected surface reflectances for 13 wavebands. Inputs are all 15 channels from MERIS, which lie between 412.5 nm and 900 nm. The grid of the product is MERIS RR reference grid providing the ortho-geolocation information for each pixel. The resolution is ~1.2km.

Units: *Dimensionless*

Uncertainty measure: Code will be provide to estimate error covariance matrix for all SDRs retrieved, taking as input the uncertainty in AOD, the SDR and the solar/view geometry. Uncertainty is stored only as 1 s.d. of error in retrieval of AOD.

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Selected L1b Flags (TBD)

Time period: 2002-2010

Size: TBD

SPOT-VGT1 L2 SDR Reflectance Product

ID: VGT1_SDR_*i*

Name: SPOT-VGT1 L2 Surface Directional Reflectance in channel *i*, with *i* = 1 to 4, ie all SPOT_VGT channels.

Standard Name: ***TBD***

Description: The product provides atmospherically corrected surface reflectances for 4 wavebands. Inputs are all 4 channels from SPOT-VGT (450nm, 640nm, 830nm and 1660nm).
The grid of the product is SPOT-VGT1 reference grid providing the ortho-geolocation information for each pixel. The resolution is ~1.15km.

Units: *Dimensionless*

Uncertainty measure: Code will be provide to estimate error covariance matrix for all SDRs retrieved, taking as input the uncertainty in AOD, the SDR and the solar/view geometry. Uncertainty is stored only as 1 s.d. of error in retrieval of AOD.

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Selected L1b Flags (TBD)

Time period: 1998 - 2002

SPOT-VGT2 L2 SDR Reflectance Product

ID: VGT2_SDR_*i*

Name: SPOT-VGT L2 Surface Directional Reflectance in channel *i*, with *i* = 1 to 4, ie all SPOT_VGT channels.

Standard Name: ***TBD***

Description: The product provides atmospherically corrected surface reflectances for 4 wavebands. Inputs are all 4 channels from SPOT-VGT2 (450nm, 640nm, 830nm and 1660nm). The grid of the product is SPOT-VGT2 reference grid providing the ortho-geolocation information for each pixel. The resolution is ~1.15km.

Units: *Dimensionless*

Uncertainty measure: Code will be provide to estimate error covariance matrix for all SDRs retrieved, taking as input the uncertainty in AOD, the SDR and the solar/view geometry. Uncertainty is stored only as 1 s.d. of error in retrieval of AOD.

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Selected L1b Flags (TBD)

Time period: 2002 – 2010

Broadband SDR

To be computed per pixel on the production grid for 3 broad wavebands (visible, shortwave infrared, total shortwave). Data will be produced from the instrument-specific narrow band SDRs, including a full variance/covariance structure (3x3 matrix, 6 unique entries as matrix is symmetric). Each dataset will contain information on the sensor used in production of the broadband SDR (BBSDR).

ID: BROADBAND_SDR_*i*

Name: Broadband Surface Directional Reflectance and uncertainty for a particular instrument / day / view angle (where multiple available). Contains 14 channels of information. In channel *i*, with *i* = 1 to 3 (visible, shortwave infrared, total shortwave) we have broadband reflectance. In channel *i*, with *i* = 4 to 6 (visible, shortwave infrared, total shortwave) standard error in estimate. In channel *i*, with *i* = 7 to 9 (visible-shortwave infrared, visible-total shortwave, shortwave infrared-total shortwave) correlation coefficients. In channel *i*, with *i* =10 to 13 viewing and illumination angles (radians) at ground relative to geoid normal (view zenith; view azimuth (clockwise from N), solar zenith, solar azimuth (clockwise from N)). Sensor information to be coded in metadata. Status flags in channel 14 (including cloud and snow).

Standard Name: *TBD*

Description: The product provides broad waveband atmospherically corrected surface reflectances and associated uncertainty, in the product (1 km) grid. Inputs are data from ATSR-2, AATSR, MERIS or VEGETATION.

Units: *Dimensionless other than angles (radians)*

Status flags: No-data (= AC not performed), Invalid (= Uncertainty below threshold), Quantised cloud probability flags. Quantised snow probability flags. Selected L1b Flags (TBD)

Time period: 1995-2010

Size: TBD

4.1.1 Broadband BRDF kernels

To be computed per pixel on the production grid for 3 linear model kernels for broad wavebands (visible, shortwave infrared, total shortwave) (i.e 9 channels of information, plus associated uncertainty). Data will be produced by fitting the kernel models (per waveband) to broadband SDRs over a weighted temporal window, on an 8-day rolling window (defined to be the same as MODIS) and associated prior estimate. Output includes a full variance/covariance structure (9x9 matrix, 45 unique entries as matrix is symmetric) and a term indicating the relative contribution of the prior values.

ID: BROADBAND_Kernels_*i*

Name: Broadband kernel model values and uncertainty for a particular 8-day period. Contains 55 channels of information. In channel *i*, with *i* = 1 to 9, kernel weights (isotropic, volumetric, geometric) for each broad band (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 10 to 18 standard error in kernel weights (isotropic, volumetric, geometric) for each broad band (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 19 to 54, 36 combinations of correlation coefficients between each kernel weight and waveband. Status flags in channel 55.

Standard Name: *TBD*

Description: The product provides broad waveband kernel model weights and associated uncertainty, in the product (1 km) grid. Inputs are broadband SDRs from all available sensors over the time window and prior estimate of parameters.

Units: *Dimensionless*

Status flags: Fill value (= no data, even in prior), QA category (broad interpretation of uncertainty). Quantised prior weighting (unity=no information from data in time period, close to 0=most information comes from observations)

Time period: 1995-2010

Size: TBD

4.1.2 Broadband albedo

To be computed per pixel on the production grid from the 3 linear model kernels for 3 broad wavebands (visible, shortwave infrared, total shortwave), an estimate of bihemispherical integral of reflectance ('white sky albedo') per waveband, directional-hemispherical integral of reflectance ('black sky albedo') at local solar noon [OR, directional-hemispherical integral of reflectance ('black sky albedo') integrated over solar path for time period]. The white sky albedo is simply the isotropic kernel parameter value. The other terms are functions of the 3 kernel weights per waveband and angular integrals of the kernels. Data are defined over a weighted temporal window, on an 8-day rolling window (define the same as MODIS). Output includes a full variance/covariance structure (6x6 matrix, 21 unique entries as matrix is symmetric).

ID: BROADBAND_Albedo_i

Name: Broadband albedo values and uncertainty for a particular 8-day period. Contains 29 channels of information. In channel *i*, with *i* = 1 to 3, broad band white sky albedo (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 4 to 6, broad band black sky albedo (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 7 to 9, standard error in white sky albedo for each broad band (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 10 to 12, standard error in black sky albedo for each broad band (visible, shortwave infrared, total shortwave). In channel *i*, with *i* = 13 to 27, 15 combinations of correlation coefficients between each albedo and waveband. Status flags in channel 28. Sun angle associated with black sky albedo in channel 19.

Standard Name: *TBD*

Description: The product provides broad waveband albedo values and associated uncertainty, in the product (1 km) grid. Inputs are broadband kernel weights over the time window and prior estimate of parameters.

Units: *Dimensionless*

Status flags: Fill value (= no data, even in prior), QA category (broad interpretation of uncertainty). Quantised prior weighting (unity=no information from data in time period, close to 0=most information comes from observations)

Time period: 1995-2010

Size: TBD

4.2 Justification and specification for each input instrument of the pre-processing algorithms

This section describes the algorithms to be prototyped for atmospheric correction, snow and cloud detection, BRF modelling, angular and spectral integration, time and space compositing, and product merging.

4.2.1 Unified Pixel Classification

The current operational detection of clouds in SPOT VGT data relies on spectral threshold tests using the reflectance in the blue and SWIR bands (Lissens et al, 2000). Different threshold combinations are used to identify cloudy and clear pixels. A pixel which does not pass either test is declared uncertain. A snow mask is calculated using spectral threshold tests on the red and MIR channels combined with 3 spectral slope tests which exploit the lower scattering of snow in MIR and SWIR bands compared with clouds. A cloud shadow is finally added based on an estimation of a potential cloud shadow and a test on the NDVI of concerned pixels.

Current MERIS cloud screening uses spectral thresholds on shortwave bands, complemented by spectral slope tests in order to recover bright land surface and snow (Santer, 1997). In the current reprocessing of MERIS these cloud and snow tests are significantly changed and improved (Brockmann and Santer, in preparation) by adding tests on the height of the scattering surface (based on the oxygen absorption measurements in MERIS band 11), and new tests for snow and ice detection using the MERIS Differential Snow Index (MDSI), based on the ratio of bands 13 (865nm) and 14 (885nm).

The AATSR cloud screening is also based primarily on threshold tests (Birks, 2007). The AATSR gross cloud test flags as cloudy those pixels whose brightness temperature in the 12 micron channel falls below a specified threshold. The small-scale spatial coherence test works by calculating the standard deviation of the 11 micron brightness temperature in a 3 x 3 group of pixels and comparing it with a threshold. If the standard deviation exceeds the threshold, the pixels in the group are flagged as cloudy. The Visible Channel Cloud Test can only be used in the daytime. The NDVI (*Normalized Differential Vegetation Index*) is defined as NDVI = (R87 – R67)/(R87 + R67), where R87 and R67 are the calibrated reflectances in the 0.87 and 0.67 micron channels respectively. Two indices are defined involving the 0.55 micron channel reflectance R55: NDI2 = (R67 – R55)/(R67 + R55). The method uses two of these indices, NDVI and NDI2, to define a two-dimensional classification space. In this space, pixels of different surface types form clusters, and by identifying into which cluster a pixel falls, the surface type at the pixel can be determined. The Snow Index (NDSI) test based uses the bands centred at 0.555 and 1.640 microns respectively, NDSI = (R55 – R16)/(R55 + R16), where R16 is the calibrated reflectivity of the 1.6 micron channel.

Both MERIS and AATSR cloud screening are not optimal because of missing spectral information in each of the single instrument (SWIR and TIR bands in MERIS, O2 and water vapour bands in AATRS). In the framework of the MERIS – AATSR synergy project an algorithm has been developed that combines the data from both instruments (Gomez-Chova, 2009). A thorough analysis has been undertaken on the information content w.r.t to cloud detection in both instruments, and a set of 19 features has been identified as optimal with respect to the number of features (which should be kept low) and information content. These features include the spectral reflectances of the two instruments, and a number of band combinations. The cloud screening algorithm is a combination of feature extraction and supervised classification and spectral unmixing. The training vectors for the supervised classification have been obtained from a database of radiative transfer calculations. The results of the algorithm are a binary cloud mask resulting from the feature tests and a cloud abundance values (between 0 and 1) from the unmixing. These two values can be used by subsequent algorithms to decide if a pixel can be processed, or a final logic is applied to conclude on the pixel status.

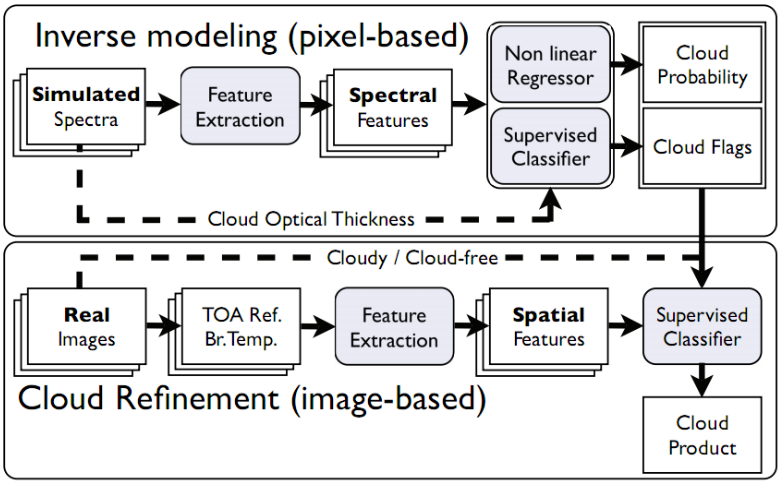

Figure 4-1: Overview of synergistic cloud classification algorithm

A critical step for the synergistic use of MERIS and AATSR is the collocation of the products. Due to the high spatial and temporal dynamic of clouds, misalignment of the two data sets would impact the cloud retrieval. Figure 4-2 shows the steps included in the preprocessing of the data of the two instruments.

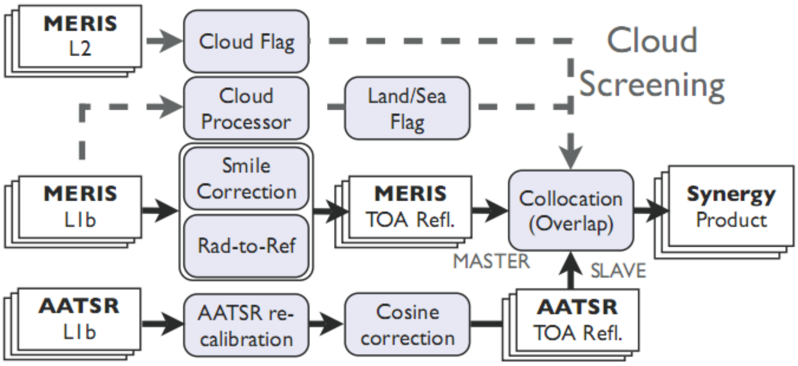

Figure 4-2: Flowchart of the MERIS and AATSR synergy preprocessing module

4.3 Approach to implementing uniform retrieval methods

4.3.1 Algorithm trade-off

Neither of the single instrument algorithms discussed in the previous section is considered of sufficient quality for the purpose of the GlobAlbedo project. The MERIS and AATSR cloud screening algorithms have been criticised during the MERIS-AATSR user workshops, and also the SPOT VGT algorithm does not screen out sufficiently the doubtful cloudy pixels. The reason is that the global Level 2 algorithms cannot be too severe in order to permit analysis of single Level 2 products. On the contrary, for a Level 3 product such as the Albedo, a clear sky conservative (i.e. severe cloud screening) approach is required. Even a small number of undetected clouds can significantly impact the final albedo product. This has been demonstrated in the MERIS Land Albedo and the Globcover projects.

The synergistic method of Gomez-Chova is a significant improvement. It exploits optimally all features available in both instruments and combines these in a non-linear, self trained mathematical way. The comparison shown in the MERIS-AATSR-Synergy Project demonstrates the improvements compared to the standard MERIS and AATSR algorithms. However, for the Globalbedo a synergistic use of MERIS and AATSR is not possible because of (1) the limitation of AATSR to the centre of the MERIS swath, thus not providing data for half of the MERIS swath and (2) the overall design of the processing architecture which does not foresee multi-sensor processing at this stage. The latter point could of course be resolved, however, point (1) still remains valid.

In conclusion the Globalbedo pixel classification will be an unique method, adapting the principles and mathematical implementation of the Gomez-Chova approach, but tailored to the features provided by the three instruments, each treated separately.

4.3.2 Globalbedo Pixel Classification

4.3.2.1 Pixel types and flags

The pixel classification results in identification (=flag) of

- Land/water
- Cloudy pixel
- Snow/ice pixel
- Cloud shadow
- Cloud edge
- Exceptional observation

Clear sky: All pixels not flagged as cloudy or snow/ice are treated as clear sky and shall be submitted to further processing.

Snow/ice pixel should also be included in further processing but may deserve special processing (if not, they can simply be treated as clear sky).

Cloud shadow pixels are clear sky pixels but special treatment should be applied, taking the reduced incident sun light into account.

Cloud edges are geometrically identified by horizontally extending a cloud; these pixels can be excluded from further processing, for safety reasons.

Exceptional observations are those measurements among a set of pixels collected for a given location and time window (binning cell), which spectrum differs significantly from the mean spectrum of all measurements within the binning cell.

4.3.2.2 Processing Logic

The processing sequence includes the following steps:

1) During SDR Processing (Level 2 processing)

4.3.2.2.1 Calculation of features.

For each of the instruments the feature set will be composed of the spectral band and complemented by the following quantities:

- Brightness = average spectral reflectance
- Whiteness = average spectral slope
- Snow detection using a spectral slope test that exploits the different absorption of snow and clouds in NIR or SWIR part. The mathematics will be unique for all instruments, but the selected bands and thresholds will be instrument specific.
- Additional features will be the apparent height of the scattering surface (MERIS only)

4.3.2.2.2 Clustering of the features.

This will be either done by an unsupervised classification, e.g. Entropy-Maximisation (EM) Clustering, which is an instrument independent step, followed by an automated labelling (instrument specific), or by a supervised classification, e.g. using a Support Vector Machine (SVM) or Artificial neural network technique (ANN). These will depend on the chosen features and are instrument specific steps.

The results of this step are the flags “cloudy pixel” and “snow/ice pixel”.

4.3.2.2.3 Land/water reclassification

Using geo-physical data instead of a static map as included in the Level 1b product, the purpose of this classification is to discriminate land from water pixels in cases where the Level1b *a priori* classification leads to ambiguities which may occur from:

- geo-location error;
- land /ocean atlas error: uncharted land or water, etc.;
- transient emerged land: tidal flats, etc.

The reclassification is limited to a priori land pixels, i.e. those pixels flagged as land in the Level 1b product and not identified as cloudy in the previous step. The purpose is to confirm the classification as land or to reclassify it as water (lake).

The tests are performed on reflectances corrected for gaseous absorption.

First, a comparison of the reflectance at 665 nm with a threshold coming from a LUT is performed. This LUT contains the darkest possible reflectance of a land surface (dark vegetation). If the tested reflectance is lower than this threshold, it is an indication that the measured surface is water:

*test_1 = rho_gascorr(band6) < alpha * LandIdentification_LUT(viewing geometry)*

Alpha is a tuning parameter, which is used to add a “margin” to the LUT values. Initially, alpha should be equals 0.8.

A second test is made which compares the reflectance at 665 nm with the reflectance at 865 nm; it is an indication of water if the reflectance at 665 nm is greater than the reflectance at 865 nm. A tuning parameter *beta* is multiplied with the reflectance at 865nm as a margin for safety reasons. Initially, this parameter should be set to 1.

*test_2 = (beta * rho_gascorr(band13) < rho_gascorr(band6))*

The pixel is classified as water if both tests *test_1 and test_2* are true.

The LUTs used in the tests are the same as used in the Land/water test of the MERIS ground segment processor. They are publicly available from ESA.

The results of this step are the flags “L2_water” and “L2_land”.

4.3.2.2.4 Cloud edge processing

In the cloud edge processing, each pixel in an n*n pixel area surrounding a cloud pixel will be flagged as cloud edge. For MERIS RR data n=1 is chosen.

The result of this step is the “cloud_edge” flag.

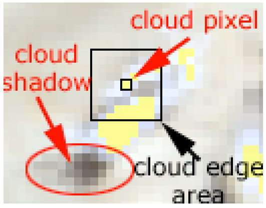

Figure 4-3: Cloud shadows and edges

4.3.2.2.5 Cloud shadow identification

To detect dark cloud shadow pixels -whose spectra are polluted by the shadow- the position of the shadow will be computed from cloud top height and the sun – pixel geometry. The following Figure 4 makes the assumption that for these computations on small regional scales, the earth might be considered as flat. Also the figure is only valid for the case in which sun, cloud and satellite are on the same West-East line. Not shown in the figure but important to be taken into account is the surface elevation when calculating the cloud position and the cloud shadow.

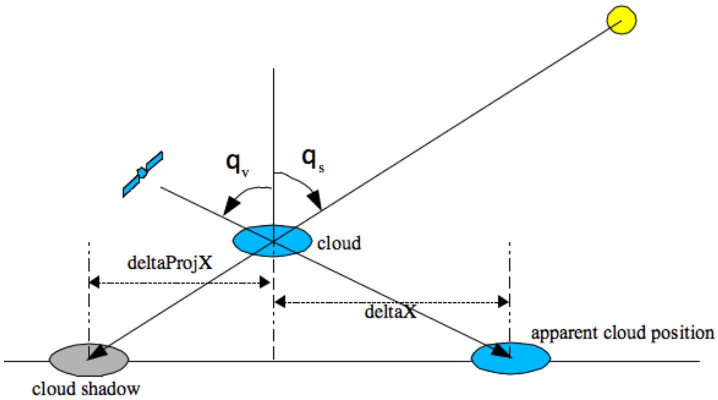

Figure 4-4: geometrical conditions for the cloud shadow calculation

The first is to calculate the real cloud position. When a pixel is detected as cloudy, one can retrieve the cloud position according to:

- the apparent position *lat_app* and *lon_app* of the cloud
- the cloud altitude *alt_cloud* assumed constant or determined in a previous step
- the viewing angles q_v and q_s
- the altitude of the apparently cloudy pixel retrieved from the DEM
- *R_earth* the Earth radius

We have:

$$\Delta x = -(alt_{cloud} - dem(lat_{app}, lon_{app})) \times \tan\theta_c \sin\phi_v$$
$$\Delta y = -(alt_{cloud} - dem(lat_{app}, lon_{app})) \times \tan\theta_v \cos\phi_v$$

We then calculate the cloud position by adding the apparent one to the correction one:

$$Lat_{Cloud} = lat_{app} - \frac{\Delta y}{R_{earth}} \frac{180}{\pi}$$
$$Lon_{Cloud} = lon_{app} - \frac{\Delta x}{R_{earth}\cos(lat_{app})} \frac{180}{\pi}$$

Once the cloud is correctly located, its shadow is computed with the DEM taking into account the intersection between the sun-cloud direction and the surface. This procedure is found iteratively:

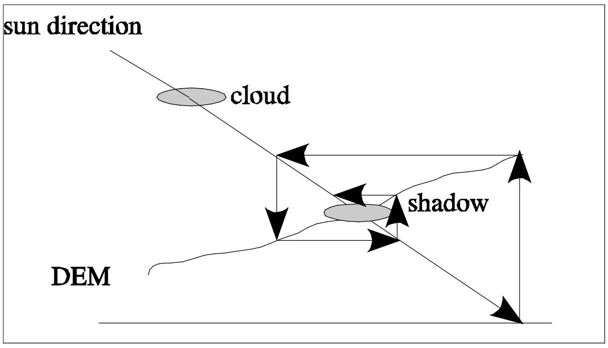

Figure 4-5: Iterative procedure to find the real cloud position

The algorithm (see Pseudo code below) could be resumed as follows with sza and saa representing respectively θ_q and ϕ_s, *InterpDEM* an interpolation method of the DEM to determine the altitude at any position and values of *MaxIterNb* and *distMin* having to be adjusted.

Pseudo Code:

```
i=0
Lat(i) = Lat_cloud
Lon(i) = Lon_cloud
Dem(i) = InterpDEM (Lat_cloud, Lon_cloud)

Do {
        i = i+1
        deltaProjX(i) = - (alt_cloud – dem(i) )* tan (sza) * sin (saa)
        deltaProjY(i) = - (alt_cloud – dem(i) )* tan (sza) * cos (saa)
        distLon = -deltaProjX(i) / R_earth * 180/pi
        distLat = -deltaProjY(i) / R_earth * 180/pi
        lat(i) = Lat_cloud + distLat
        lon(i) = Lon_cloud + distLon
        dem(i) = InterpDEM(lat(i), lon(i))

}       while ( (i< MaxIterNb) and
            (dist ( (lat(i), lon(i), (lat(i-1), lon(i-1)) ) > distMin) and dem(i) < alt_cloud)

            If (dem(i) > alt_cloud) then
                    No projection
            Else
                    Conversion from lon(i), lat(i) to pixel position (i_clshadow, j_clshadow)

set cloud_shadow (i_clshadow, j_clshadow)
```

The result of this step is the flag “cloud_shadow”.

4.3.2.3 During the Broadband reflectance conversion/albedo retrieval (Level 3 processing)

4.3.2.3.1 Temporal consistency check

During the temporal aggregation of the data we will include a statistical test, applied to each binning cell:

- all clear sky observation will be collected
- mean and standard deviation will be calculated.
- If sufficient data (>n) are available, pixel with reflectances greater than a factor f * standard deviation will be removed from the sample.

The parameters n and f will be tuned during algorithm development. If sufficient data are available, this process can be iterated.

The result of this test is the flag “exceptional_measurement”.

4.3.3 Atmospheric correction

Objectives

The satellite datasets will be processed from top-of-atmosphere (TOA) observations to obtain surface directional reflectance. The initial satellite measurements are strongly affected by molecular and aerosol scattering, and absorption by ozone and water vapour. The correction of these effects is necessary prior to estimation of surface albedo.

Atmospheric correction is performed in two stages. In the first step, the atmospheric properties are determined at the time of satellite overpass, which includes the estimation of aerosol loading and model from the data. Secondly, a radiative transfer model of the atmosphere is inverted to estimate surface reflectance, accounting for the atmospheric scattering and absorption. The parameters required to model aerosol radiative effects are aerosol optical depth (AOD) for a given reference wavelength, and aerosol model describing spectral dependence of AOD, single scattering albedo, and phase function. The atmospheric correction process thus also provides information on the magnitude of the downwelling diffuse flux at the surface, by providing AOD and model at time of acquisition.

The atmospheric correction process here is designed according to the following requirements:

- To retrieve surface reflectance for cloud-free imagery
- Provide a per-pixel error estimate on the surface reflectance, related to instrument error
- Avoid the use of a priori estimates of surface reflectance since this may bias the subsequent albedo retrieval
- Derive aerosol information directly from data, to provide estimates simultaneous with observations, and allow application to regions and periods where other observations are not available.
- Provide sufficient diagnostic variables to allow validation of the method
- Provide a single atmospheric model for the retrieval, consistent across the different platforms

Retrieval of aerosol properties

The algorithm for aerosol retrieval is based on the optimal estimation technique, which has been widely applied to atmospheric retrievals (Rodgers 2000; Govaerts et al., 2010). The algorithm has two stages: (1) Given a set of satellite TOA radiances, and an initial guess of aerosol parameters (namely AOD and aerosol model), we estimate the corresponding set of surface reflectances. (2) Application of the constraint set to this set of reflectances results in an error metric, where a low value of this metric should correspond to a set of surface reflectances (and hence aerosol parameters) which is realistic. Step (1) is repeated with a refined set of aerosol parameters until convergence at an optimal solution. The curvature of the error surface gives a measure of confidence in the retrieved AOD, which can be related to corresponding uncertainty in derived SDRs. The main elements therefore are (i) design of an efficient and accurate look-up table (LUT) scheme for deriving surface reflectance for known aerosol parameters, and (ii) formulation of constraints on the land surface reflectance suitable for inversion.

LUT surface reflectance estimation

For computational efficiency we employ a set of pre-calculated LUTs which map top-of - atmosphere reflectance to surface reflectance for a range of surface and atmospheric conditions. These are defined for each instrument, and allow rapid and accurate estimation of surface directional reflectance given TOA observations. The lookup tables of atmospheric properties for the algorithm are generated using the forward simulations of

MOMO atmospheric radiative transfer code model developed at FUB. For MERIS, we avoid bands 11 and 15 affected by strong oxygen and water vapor absorptions. Correction for water vapour and ozone are applied prior to further corrections. Coupling effects, such as multiple land-atmosphere scattering are implicitly represented in the LUTs. An two-stage iteration of the inversion allows compensation for non-Lambertian surface scattering effects, to a theoretical 1-2% relative error (Hu et al., 1999). The retrieval of aerosol properties will be made at a more coarse grid (10km) than the per-pixel retrievals, to allow computational efficiency and selection of reliable AOD inversions.

Instrument-specific constraints

(A)ATSR(2)

The team at SU have led development of a method for simultaneous estimation of aerosol opacity and surface reflectance for (A)ATSR imagery which will be applied for GlobAlbedo retrievals. The method is documented in the papers (North et al., 1999; North 2002; Grey et al., 2006a,b. The basis of the constraint is the angular variation of the land surface reflectance, giving a corresponding error. For AATSR, the ratio of surface reflectances at the nadir and forward viewing angles is well correlated across wavebands, and the variation in anisotropy may be modelled simply (North et al., 1999). This avoids the need for assumptions on absolute surface brightness or spectral properties. Our method differs from early approaches by using a more sophisticated surface model to account for some spectral variation of the angular shape owing to the variation of the diffuse fraction of light with wavelength

MERIS

When viewing from a single direction, such as MERIS, we must rely on the spectral signature to distinguish aerosol from ground scattering. For a given set of surface reflectances derived by assuming a certain atmospheric profile, this may be expressed as an error based on the fit of the retrieved surface reflectance to the assumed target reflectance. It is central to the method for albedo that no bias is introduced by assumptions of absolute brightness, and so the constraint used will be based on a parameterised model of the shape of spectrum rather than assuming absolute values.

This constraint allows estimation of the atmospheric aerosol by optimal estimation as before. Where a target of approximately known reflectance can be identified, such as dense vegetation or a body of water, aerosol optical depth at the target location may be estimated on the basis of known correlation of ground reflectance at different wavelengths Remer et al (2005). Where a large number of wavebands are available it is possible to represent the target reflectance as a linear mixture of an idealized vegetation and soil spectrum. A number of variations on such methods have been used successfully for aerosol retrieval with MERIS (von Hoyningen-Huene, et al., 2006; Guanter et al., 2008; Santer et al., 2007). However, accurate application is limited to regions where such targets are available at the appropriate spatial resolution (i.e. oceans and dark dense vegetation), so we must employ interpolation of the aerosol field to derive values at image points suitable for atmospheric correction. Recent results suggest improvement of this method is possible using calibration of the spectral relationship over a range of representative land covers, corresponding to selected AERONET sites (Levy et al., 2007) allowing correction for view-angle effects on surface spectra and generalisation to brighter surfaces.

SPOT-VGT

Relative to MERIS there are a small number of bands (4), but spanning a wide spectral range. The wide swath offers more frequent samples, but results in acquisition of data at greater atmospheric path lengths and consequently increased atmospheric scattering. Since no operational algorithm currently exists for correction of SPOT-VGT based on aerosol derived from the data, a method will be developed based on the optimal estimation framework adapted for this instrument. The blue channel in particular is highly sensitive to aerosol scattering, giving information on aerosol content. The basis of the algorithm will be to develop a constraint based on spectral correlation of surface reflectance, similar to that successfully applied to MODIS (Levy et al., 2007), but accounting for some variability of the ratio with VGT view angle and density, as outlined in (Maisongrande et al., 2000). The correlation will be suitable primarily for dark vegetation targets, where good correlations exist between the blue, red and 1.6µm channels. A second step will then interpolate to give continuous field suitable for per-pixel atmospheric correction. The constraint developed by analysis of corrected(A)ATSR & MERIS images with similar wavelengths, and will be tested and optimised by analysis of SPOT VGT over AERONET sites.

Ancillary data requirements

Surface pressure, ozone and water vapour will be estimated from best available ancillary data (ECWMF fields). For aerosol scattering, the Angstrom coefficient, single scattering albedo and phase function are derived from mixtures of 6 components aerosol types from refractive index and Junge size distribution. These are specified by the Global Aerosol Data Set (GADS), and derived from the software package OPAC (Optical Properties of Aerosols and Clouds) (Hess et al., 1998). To avoid spatial artifacts caused by changes in aerosol model used throughout small regions, the AOD will be the principal property derived from satellite observations. Climatologically-defined aerosol models will be used to constrain the remaining aerosol parameters.

Uncertainty estimates

A per-pixel estimate of uncertainty in the SDRs will made considering the errors due to the correction algorithm, ancillary data and instrument noise and calibration accuracy. The main sources of error in the SDR to be quantified are:

- Aerosol retrieval error: the optimal estimation algorithm returns an error estimate based on curvature of the error surface around the minimum. This error in AOD will be related to spectral surface reflectance by the RTM model sensitivity study.

- Instrument calibration errors: a separate study will characterise this error for each instrument

- Radiative transfer error: this is due to approximations in the RTM model, LUT and the representativeness of the aerosol model

- Interpolation error: error resulting in derivation of local AOD by interpolation from points distant to the retrieved reflectance. Error will be based on autocorrelation of AOD field

The error due to residual undetected clouds will be reported as a positive bias. Experience suggests the optimal estimation framework can act as a further cloud detection stage, since sub-pixel cloud contamination leads to poor model fit. The ancillary information produced by the retrieval will include quantitative uncertainty in the retrieved AOD, and code to relate this to produce the surface reflectance covariance matrix for each retrieval.

4.3.4 Narrow-to-Broadband SDR conversion

A linear model will be used for the spectral integration of the SDR products derived by the atmospheric correction process. It will generate broadband directional reflectance by the linear combination of the SDR at the spectral channels of each sensor. This linear modeling enables the decoupling of the spectral and angular integrals to be performed for the conversion from SDRs to broadband albedo so that the angular and spectral integrals can be permuted. In this way, the amount of data to be inverted during the BRF model inversion to be performed for the angular integration of reflectance is reduced to the three spectral windows of the final albedo products.

Surface broadband albedos are not only measures of the surface properties, but also of the atmospheric conditions through the downward irradiance field. Therefore, the narrow-to-broadband conversion process depends on parameters such as the input reflectance, the surface elevation, the solar zenith angle and the aerosol parameters. Whether all or part of these parameters should be inputs to the narrow-to-broadband conversion or they will be assumed in the error budget of the narrow-to-broadband SDR conversion process will be analysed during the study.

The methodology which will be followed for this analysis and for the derivation of the spectral conversion coefficients is based on that of Liang (2000). It consists in the generation of a large data base of simulated high spectral resolution albedos by means of the combination of a large spectral library of surface reflectance patterns with different atmospheric conditions simulated with radiative transfer calculations for a wide range of atmospheric parameters. The resulting high resolution albedos are integrated in the broadband windows and convolved to the instruments spectral response functions. The conversion coefficients are derived by the inversion of the broadband albedos against the ones simulated at the instruments channels. The spectral albedos in this approach will be SDRs in our case. Radiative transfer calculations will be performed with the MOMO radiative transfer code, which allows consistency with the other radiative transfer calculations to be performed along the GlobAlbedo processing chain. Concerning the input spectral reflectance library, the Johns Hopkins University (JHU) Spectral Library provides a wide set of surface types and the appropriate spectral range and sampling. It contains a large number of spectra of e.g. vegetation, soils, minerals, manmade materials, ice and snow in the solar spectral range. However, the need for a second data set of snow and ice reflectance spectra could be preferred due to the relatively different spectral reflectance of these surfaces with respect to the rest of land covers (Greuell and Oerlemans, 2004).

4.3.5 Broadband BRF model inversion

The purpose of the **Broadband BRF model inversion framework** is to provide an optimal estimate of a set of parameters $f_+(T)$ for some discrete time sample T (in *days* from some baseline) over some defined area A that enable albedo to be estimated for that area and time under varying illumination conditions for wavebands B. The requirement for this comes from the need to translate sample observations of spectral directional reflectance obtained from satellite data into spectral and directional integrals of reflectance that can be used to calculate albedo. This framework achieves angular, temporal and 'product' merging (i.e. uses data from multiple satellite sensors), achieving an optimal combination through propagation of uncertainties in the various inputs. An underlying assumption is that there are multiple observations of BRF available over time in a common set of wavebands (the required broadband wavebands here) over the same viewing area A. Since the assumption is likely to be violated in combining multiple observations from different sensors, this component of spatial compositing should be treated as a source of uncertainty on the measurements as observations of A. Since the satellite observations do not directly provide estimates of BRF in the required wavebands, there will be a source of uncertainty in estimating these from any particular observation set. Since all observations cannot be expected to be achieved at time T, uncertainty arising from temporal compositing must also be treated.

This framework takes as input:

(i) a set of N estimates (indexed i), directly derived from observations, of broadband BRF data that we denote $\rho(t_i,\Omega_i,\Omega_i')$ for observation times t_i, viewing vectors Ω_i, illumination vectors Ω_i' for a set of wavebands B, where $B=\{VIS,SWIR,SW\}$ represents wavebands in the visible (PAR), shortwave infrared and total shortwave infrared respectively, with associated uncertainty matrices $C_{obs}(t_i)$;

(ii) a prior estimate of model parameters $f_-(T)$ with associated uncertainty matrix $C_-(T)$;

(iii) an expectation of temporal smoothness in reflectance for each waveband, expressed as expected variance associated with reflectance (not including BRDF effects) denoted $C_{time}(T-t_i)$.

The primary outputs of the framework are then:

(iv) estimates of model parameters at time T, $f_+(T)$;

(v) uncertainty matrices $C_+(T)$ expressing the uncertainty of $f_+(T)$.

The input broadband BRF data are broadband estimates derived from narrow band SDR data for the (A)ATSR, MERIS, and VEGETATION sensors. The strategy is to convert all spectral reflectance data (for a given sensor, location, time and viewing/illumination geometry) into broadband estimates of reflectance through the standard linear modelling method described above. Linear kernel-driven BRDF models are then applied to the total set of measurements over the relevant time compositing window. When such modelling is performed for a single sensor (e.g. MODIS) the modelling is generally done the other way around, i.e. the linear kernel models are fitted to observations in the 'raw' narrow bands of the sensor. The resultant kernel weights (see below) are then combined into broad band estimates. Since both models are linear, it makes no difference which way around these operations take place. When merging data from sensors with different wavebands however, there is a distinct advantage to doing it the way described here as: (i) processing costs are reduced; (ii) more robust estimates of the (broadband) model parameters will be produced all samples from all sensors can be directly combined. Thus, the 'sensor merging' takes place prior to the BRF modelling. The potential downsides to this approach are: (i) there are no narrow band estimates of BRF produced; (ii) there are no 'sensor specific' albedo products produced. In some ways the first of these is a shame, as there have been many 'side' advantages to having such data from MODIS. However, MODIS is specifically a BRDF/albedo product, whereas the only focus here is albedo. In any case, the angular sampling from MERIS and (A)ATSR are rather poor, and it is not straightforward to provide BRDF modelling from these sensors alone: some 'merging' with other data must at some point be considered. The second 'disadvantage' is probably actually an advantage. Provided there are no significant cross- or absolute calibration issues, EO science should in any case move towards such synthesis products, rather than simply generating more and more sets of more or less well-constrained datasets from individual sensors. To obtain the 'optimal' estimate of albedo does not need any particular sensor: at some points angular sampling issues will be of overriding importance, at others spectral sampling. Often high temporal (cloud free) observations will be the most important issues. The 'system of systems' here can make use of observations from any sensor at an appropriate spatial resolution.

The optimal estimate of the model parameters is achieved via a minimisation of a measure of departures between expected and observed quantities weighted by uncertainties in each. It is standard to use an L2 norm for this:

$$\begin{aligned} e^2 \;=\; & \sum_{i=0}^{N-1} \left[\rho(t_i,\Omega_i,\Omega'_i) - \hat{\rho}(T,\Omega_i,\Omega'_i)\right]^T \left[C_{obs}(t_i) + C_{time}(T-t_i)\right]^{-1} \left[\rho(t_i,\Omega_i,\Omega'_i) - \hat{\rho}(T,\Omega_i,\Omega'_i)\right] + \\ & \left[f'_+(T) - f'_-(T)\right]^T C_-^{-1}(T) \left[f'_+(T) - f'_-(T)\right] \end{aligned} \tag{1}$$

where $\hat{\rho}(T,\Omega_i,\Omega'_i)$ is a model estimate of $\rho(T,\Omega_i,\Omega'_i)$, ^{T} denotes the transpose operation and $^{-1}$ denotes a matrix inverse.

$\rho(t_i,\Omega_i,\Omega_i')$ is a vector of dimensions 3 x 1 for the 3 wavebands for sample *i*:

$$\rho(t_i,\Omega_i,\Omega_i') = \begin{bmatrix} \rho_{VIS}(t_0,\Omega_0,\Omega_0') \\ \rho_{SWIR}(t_0,\Omega_0,\Omega_0') \\ \rho_{SW}(t_0,\Omega_0,\Omega_0') \end{bmatrix}.$$

The observation variance/covariance matrix for time t_i, of dimensions 3 x 3 is:

$$C_{obs}(t_i) = \begin{bmatrix} \sigma^2_{VIS,VIS}(t_i) & \sigma^2_{VIS,SWIR}(t_i) & \sigma^2_{VIS,SW}(t_i) \\ \sigma^2_{VIS,SWIR}(t_i) & \sigma^2_{SWIR,SWIR}(t_i) & \sigma^2_{SWIR,SW}(t_i) \\ \sigma^2_{VIS,SW}(t_i) & \sigma^2_{SWIR,SW}(t_i) & \sigma^2_{SW,SW}(t_i) \end{bmatrix}$$

where $\sigma^2_{B1,B2}(t_i)$ is the covariance for sample i between bands $B1$ and $B2$.

The model we use here is linear in nature, has 3 parameters for each waveband and can be stated:

$$\hat{\rho}(T,\Omega_i,\Omega_i') = f_+(T)k(\Omega_i,\Omega_i') \qquad (2)$$

where $f_+(T)$ is a 3x3 vector:

$$f_+(T) = \begin{bmatrix} f_{+0,VIS} & f_{+1,VIS} & f_{+2,VIS} \\ f_{+0,SWIR} & f_{+1,SWIR} & f_{+2,SWIR} \\ f_{+0,SW} & f_{+1,SW} & f_{+2,SW} \end{bmatrix}$$

$k(\Omega_i,\Omega_i')$ is a 3x1 vector:

$$k(\Omega_i,\Omega_i') = \begin{bmatrix} 1 \\ k_1(\Omega_i,\Omega_i') \\ k_2(\Omega_i,\Omega_i') \end{bmatrix}$$

For the second term on the RHS of equation 1 $C_{-}^{-1}(T)$ should ideally consider covariances both between the model parameters and spectral covariances. Thus, we state the model parameter vector as:

$$f'_{+}(T) = \left[f_{+0,VIS} \quad f_{+1,VIS} \quad f_{+2,VIS} \quad f_{+0,SWIR} \quad f_{+1,SWIR} \quad f_{+2,SWIR} \quad f_{+0,SW} \quad f_{+1,SW} \quad f_{+2,SW} \right]^{T}$$

with $C_{-}(T)$ a symmetric matrix:

$$C_{-}(T) = \begin{bmatrix} \sigma^{2}_{0,VIS,0,VIS} & \sigma^{2}_{0,VIS,1,VIS} & \cdots & \sigma^{2}_{0,VIS,2,SW} \\ \sigma^{2}_{0,VIS,1,VIS} & \sigma^{2}_{1,VIS,1,VIS} & \cdots & \sigma^{2}_{1,VIS,2,SW} \\ & & \ddots & \vdots \\ & & & \sigma^{2}_{2,SW,2,SW} \end{bmatrix}$$

The particular set of kernel models used here are those of the kernel-driven semiempirical BRDF model 'Ambrals', used in the MODIS and SEVIRI products. In the above notation then, k1 is RossThick-LiSparse kernel functions for characterizing volume and k2 surface scattering (Wanner et al., 1995; 1997; Lucht et al., 2000; Schaaf et al., 2002). A reciprocal version of the LiSparse kernel is employed, as in related products. The two main reasons for the choice of model are: (ii) the team has significant experience with these models and confidence in their use (relative to other choices available); (ii) the product is then directly consistent with most other such global/hemispherical products (i.e. MODIS and SEVIRI), allowing comparison of model parameters rather than just albedo and allowing an initial prior estimate to be derived from 10 years of MODIS observations.

The framework assumes that uncertainties arising from spatial compositing and waveband interpolation/extrapolation are provided via $C_{obs}(t_i)$, along with uncertainties arising from any geolocation error, uncertainty in prediction of BRF from satellite radiances, satellite calibration. It also assumes that there is no bias over the set of N 'observations' of $\rho_B(t_i, \Omega_i, \Omega'_i)$. Practically, if N is reasonably large and errors assumed random, spatial compositing (especially on top of temporal compositing) results in a form of regularisation of the data and probably not too great an issue, except in areas of high spatial frequency where some of the higher frequencies will be dampened. Explicit treatment of spatial compositing issues could be achieved if an expectation of the local spatial statistics (at time T) were available, but this would greatly complicate the processing for this global product, so it is likely better to incorporate some simple estimate of additional uncertainty arising from the compositing in $C_{obs}(t_i)$.

Temporal uncertainty can be similarly readily taken into account via some quantification of temporal smoothness in an uncertainty framework. The algorithm used for MODIS processing currently assumes that all samples within some time window W (16 days) can be treated equally as observations at time T. This is itself an implicit form of regularisation (i.e. temporal smoothing) with a top hat filter in the temporal domain. Assuming this function to apply to each model parameter (i.e. locally constant in time) is equivalent to applying this filter directly to the observation set. If more sophisticated temporal models are assumed to apply to the model parameters (e.g. local smoothness) these do not generally translate to simple temporal filters that can be applied as weightings on the observations. The main advantage of such weightings would be a greatly simplified processing system, as parameter smoothness would need to be implemented via some (more) complex data assimilation framework. However, the general 'impact' of such assumptions *can* be achieved with temporal weighting of the observations (as demonstrated in the SEVIRI albedo product). This then requires some expectation of temporal smoothness in reflectance for each waveband, as noted above. This is perhaps slightly more complex to strictly define and therefore to calculate (e.g. as it should not include BRDF effects as these are explicitly modelled) but a simple method is proposed in this work to achieve estimates of the required uncertainties. With this in mind, the time variance/covariance matrix can be given:

$$C_{time}(T-t_i) = \begin{bmatrix} \sigma_{VIS}^2(T-t_i) & 0 & 0 \\ 0 & \sigma_{SWIR}^2(T-t_i) & 0 \\ 0 & 0 & \sigma_{SW}^2(T-t_i) \end{bmatrix}$$

where:

$$\sigma_B^2(T-t_i) = \exp\left(\frac{(T-t_i)^2}{s_B^2}\right)$$

and s_B^2 is the expectation of change in reflectance in band *B.* Thus 'temporal' compositing is achieved by applying the (constrained) optimisation (minimisation of error in model and prior fit – a weighting of the priors and the observations) over a given time period (e.g. 3 months, but to be determined experimentally from analysis of MODIS data) with the temporal function above used to weight observations over that period relative to the 'focal point' of the temporal window (the centre here).

To use these data in any modelling, we must assume that data filtering routines such as snow and cloud clearing (including cloud shadow) are *mostly* effective in their job. Some tolerance to errors can be built in to the framework, but this will ultimately rely on identifying such errors as outliers as would be problematic unless the majority of identifications are correct in any particular dataset. Estimation of albedo relies not only on the fidelity of the input data, but also on the quantity and angular sampling of the samples. Thus, ideally, as well as QA flags being used to identify snow, cloud/cloud shadow, probabilities should also be assigned to these detections.

4.4 Approach to implementing uniform retrieval methods

Uniform retrieval methods for each of the instruments, intercalibration and merging.

4.4.1 Pixel Classification

The Gomez-Chova approach applied for Globalbedo requires the following processing steps:

1) Calculation of features. For each of the instruments the feature set will be composed of the spectral band and complemented by the following quantities:

- Brightness = average spectral reflectance
- Whiteness = average spectral slope
- Snow detection using a spectral slope test that exploits the different absorption of snow and clouds in NIR or SWIR part. The mathematics will be unique for all instruments, but the selected bands and thresholds will be instrument specific.
- Additional features will be the apparent height of the scattering surface (MERIS only)

2) Clustering of the features. This will be either done by an unsupervised classification, e.g. Entropy-Maximisation (EM) Clustering, which is an instrument independent step, followed by an automated labelling (instrument specific), or by a supervised classification, e.g. using a Support Vector Machine (SVM) or Artificial neural network technique (ANN). These will depend on the chosen features and are instrument specific steps.

In addition to the Gomez-Chova approach will be a temporal consistency check:

During the temporal aggregation of the data we will include a statistical test: for each binning cell all cloud free observation will be collected, and mean and standard deviation will be calculated. If sufficient data (>n) are available, pixel with reflectances greater than the factor * standard deviation will be removed from the sample. The parameters n and f will be tuned during algorithm development. If sufficient data are available, this process can be iterated.

4.4.2 Atmospheric correction

For atmospheric correction to SDRs, we use a common framework based on a consistent LUT set defined by the radiative transfer code (MOMO) and set of aerosol models. The AOD is retrieved at time of overpass using iterative retrieval of AOD using the same framework, but with instrument-specific constraints. The method will be coded within the BEAM environment to provide consistent format of outputs.

4.4.3 Sensor intercalibration

Cao et al. (2010) have recently described the CEOS-GEO sponsored QA4EO protocols and results for several sensors for sensor inter-calibration over a high altitude polar site, Antarctic Dome C site (75°06'S, 123°21'E, elevation 3.2 km) for December 2008-January 2009. This site has been initially chosen for Simultaneous Nadir Overpasses (SNO) for inter-comparison of polar orbiting satellite sensors such as NOAA/AVHRR and NASA/MODIS. The site is covered with snow and atmospheric aerosol effects are minor. The site includes a NASA AERONET station which records Aerosol Optical Depth (AOD) and ground-based Ozone measurements. Observations from seven satellite instruments involving four space agencies were compared, including OrbView-2/SeaWiFS, Terra/Aqua/MODIS, EO-1/Hyperion, MetOp/AVHRR, Envisat/MERIS/AATSR, and Landsat 7/ETM+. Recently, a team in the UK led by S. Mackin, including Dr Nigel Fox (Chair, CEOS-WGCV-IVOS panel) and a joint student, Dale Potts with Prof. J-P Muller have extended this analysis to include all DMC satellites and sensors, AWiFS as well as MERIS, and AATSR. The GlobAlbedo team at UCL-MSSL plan to continue this analysis to include SPOT-VEGETATION, Terra-MODIS, Terra-MISR and POLDER/PARASOL.

4.5 Software requirements specification

The objective of this chapter is to specify the software requirements of the GlobAlbedo Processors and Processing System. The general conception of the workflow outlined in the Statement of Work is being translated into a series of specifications describing the system to be implemented.

Therefore, the GlobAlbedo system is analysed and decomposed into a number of functional groups. The requirements for these are described in the following. Each requirement is translated into a feature, which is described as follows:

Feature Group Title

Requirement description, either a compiled summary or directly taken from SoW

Realisation: Brief description of how the requirement is planned to be implemented

Priority: Essential | Expected | Desired | Optional	**Effort: Months | Weeks | Days | Hours**	**Risk: Dangerous | 3-Risks | 2-Risks | 1-Risk | Safe**
Description of priority rating	Description of effort rating	Description of risk rating

The GlobAlbedo Processing System shall be capable to produce SDR products from MERIS, (A)ATSR(-2) and VEGETATION input products. These SDR products are merged to BB albedo products in a subsequent step (see ***Figure 4-6***).

In order to meet the non-functional requirements, the Processing System needs to support distributed processing, with a common processing management. It must be able to handle large data volumes to fulfill the critical processing performance requirements, i.e. to process all 15 years of satellite input data in one month.

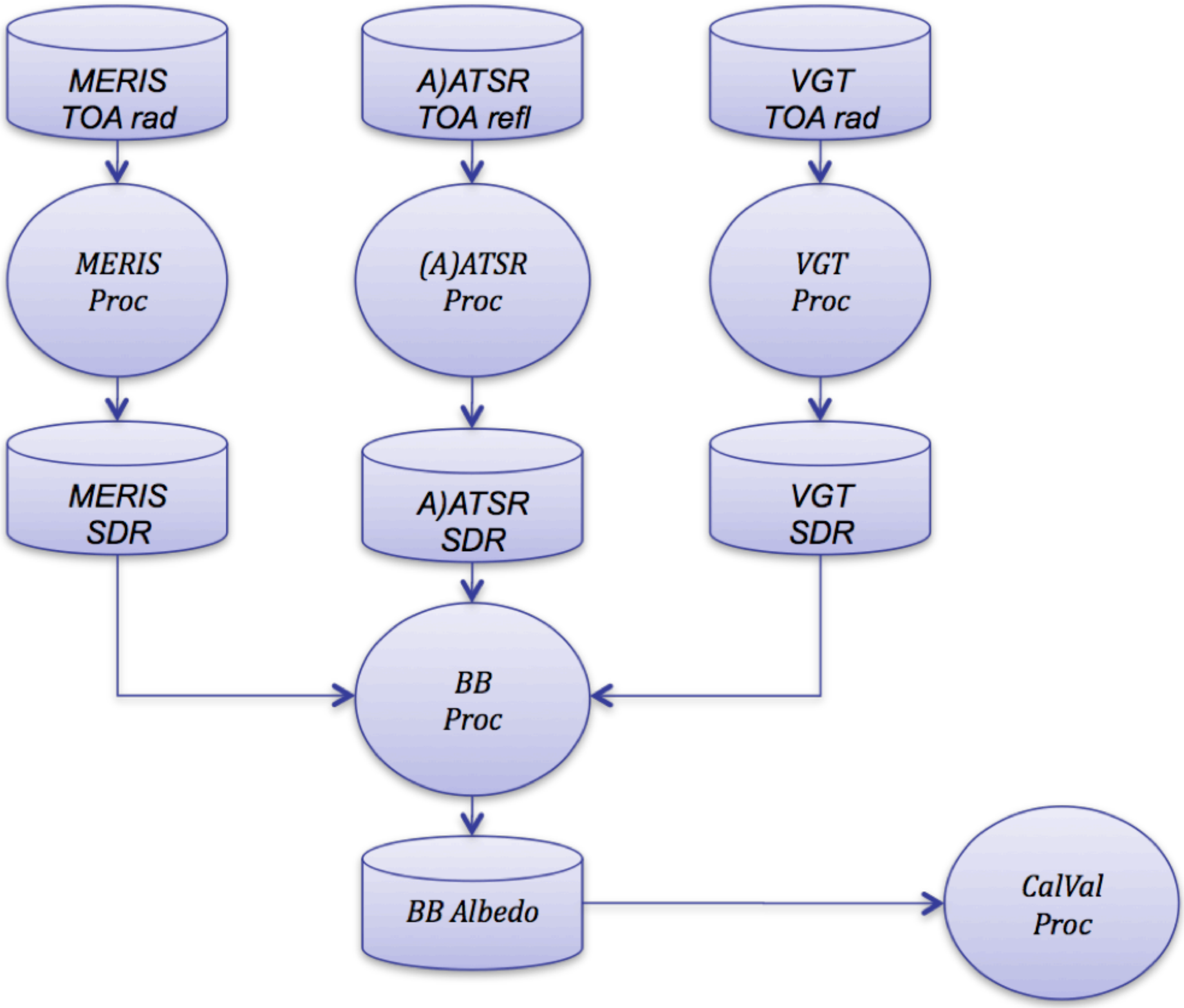

Figure 4-6: Processing Overview

The Software Requirements will be divided into the following sections:

Section	**Feature Group Title**
General Processor Requirements	SR-GP
Specific Processor Requirements: MERIS SDR	SR-MER-SDR
Specific Processor Requirements: (A)ATSR(-2) SDR	SR-ATS-SDR
Specific Processor Requirements: VEGETATION SDR	SR-VGT-SDR
Specific Processor Requirements: BB-albedo	SR-BBA
Processing System Requirements: Software	SR-SYS-SW
Processing System Requirements: Hardware	SR-SYS-HW
Non-functional Requirements	RQ-NF

4.5.1 General Processor Requirements

SR-GP-001 *The SDR processors shall operate separately for each available sensor.*

SR-GP-002 *The SDR processors shall write SDR output products in BEAM DIMAP format.*

SR-GP-003 *The SDR processors shall be able to write SDR output products in user-friendly EXPORT formats. This requirement might be implemented as a converter tool.*

Realisation: The SDR processors will be designed by implementing the breadboard prototypes as BEAM processors implementing the Graph Processing Framework (GPF) interface. The export functionality might be transferred into a dedicated conversion tool.

Priority: Essential	**Effort:** Months	**Risk:** Safe
One major processing pillar	DIMAP is the standard BEAM format and therefore it does not require additional software development. However, the export functionalities need to be implemented. The developments shall take place in KO+1 through KO+9.	Well-known technology

SR-GP-004 *The processors shall internally use a MODIS Aggregation Product (MODAGG) type of grid using the MODIS SIN.*

SR-GP-005 *The processors shall use a pyramidal data structure in order to provide multi-resolution products*

SR-GP-006 *The processors shall provide both command-line and VISAT GUI as user interfaces.*

4.5.2 Specific Processor Requirements: MERIS SDR

SR-MER-SDR-001 *The MERIS SDR processor breadboards and prototype shall be implemented as a BEAM processor.*

SR-MER-SDR-002 *The MERIS SDR processor shall be designed to implement the BEAM GPF interface.*

SR-MER-SDR-003 *The MERIS SDR processor shall be easily configurable through an XML file passed as a parameter.*

SR-MER-SDR-004 *The MERIS SDR processor shall be able to ingest MERIS L1b products in N1 format and BEAM DIMAP format.*

SR-MER-SDR-005 *The MERIS SDR processor shall not depend on auxiliary data when using MERIS AATSR synergy heritage. The fallback solution using AlbedoMap heritage depends on LARS aerosol auxiliary data for aerosol correction.*

Realisation: The MERIS SDR processor will be designed using MERIS-AATSR synergy processor heritage. The breadboard prototype shall be designed as a BEAM processor implementing the GPF interface.

Priority: Essential	**Effort: Months**	**Risk: Safe**
One major processing pillar	Iteration between scientists and BC shall result in a stable version. Backup version available from Albedomap/GlobCover. The developments shall take place in KO+1 through KO+9.	Well-known technology

4.5.3 Specific Processor Requirements: (A)ATSR(-2) SDR

SR-ATS-SDR-001 *The (A)ATSR(-2) SDR processor breadboards and prototype shall be implemented as a BEAM processor.*

SR-ATS-SDR-002 *The (A)ATSR(-2) SDR processor shall be designed to implement the BEAM GPF interface.*

SR-ATS-SDR-003 *The (A)ATSR(-2) SDR processor shall be easily configurable through an XML file passed as a parameter.*

SR-ATS-SDR-004 *The (A)ATSR(-2) SDR processor shall be able to ingest ATSR-2 and AATSR L1b products in generic N1 format and BEAM DIMAP format.*

Realisation: The (A)ATSR(-2) SDR processor will be designed using MERIS-AATSR synergy processor heritage. The breadboard prototype shall be designed as a BEAM processor implementing the GPF interface.

Priority: Essential	**Effort: Months**	**Risk: Safe**
One major processing pillar	Iteration between scientists and BC shall result in a stable version. Initial version available from Synergy project. The developments shall take place in KO+1 through KO+9.	Well-known technology

4.5.4 Specific Processor Requirements: VEGETATION SDR

SR-VGT-SDR-001 *The VEGETATION SDR processor breadboards and prototype shall be implemented as a BEAM processor.*

SR-VGT-SDR-002 *The VEGETATION SDR processor shall be designed to implement the BEAM GPF interface.*

SR-VGT-SDR-003 *The VEGETATION SDR processor shall be easily configurable through an XML file passed as a parameter.*

SR-VEG-SDR-004 *The VEGETATION SDR processor shall be able to ingest SPOT VEGETATION input products in HDF format and BEAM DIMAP format.*

Realisation: The VEGETATION SDR processor will be designed using analogue technology like the MERIS and (A)ATSR(-2) processors. The breadboard prototype shall be designed as a BEAM processor implementing the GPF interface.

Priority: Essential	**Effort: Months**	**Risk: Safe**
One major processing pillar	Iteration between scientists and BC shall result in a stable version. No initial version available yet. The developments shall take place in KO+1 through KO+9.	Well-known technology

4.5.5 Specific Processor Requirements: BB-albedo

SR-BBA-001 *The GlobAlbedo Broad Band Albedo processor breadboard and prototype shall be implemented as a BEAM L3 processor.*

SR-BBA-002 *The GlobAlbedo Broad Band Albedo processor shall be designed to implement the BEAM GPF interface.*

SR-BBA-003 *The GlobAlbedo Broad Band Albedo processor shall be easily configurable through an XML file passed as a parameter.*

SR-BBA-004 *The GlobAlbedo Broad Band Albedo processor shall be able to ingest L2 output products from SDR processing in BEAM DIMAP format.*

SR-BBA-005 *The GlobAlbedo Broad Band Albedo processor shall be capable to take multiple SDR products as input.*

SR-BBA-006 *The GlobAlbedo Broad Band Albedo processor shall work on a time window specified in the configuration file.*

Realisation: The GlobAlbedo Broad Band albedo processor will be designed as a BEAM L3 processor implementing the GPF interface. Because there is no possibility to re-use the MERIS AlbedoMap processor, it has to be implemented from scratch.

Priority: Essential	**Effort: Months**	**Risk: Safe**
One major processing pillar	Iteration between scientists and BC shall result in a stable version. No backup version available. The developments shall take place in KO+1 through KO+9.	Well-known technology

4.5.6 Processing System Requirements: Software

SR-SYS-SW-001 *The Processing System Software shall run under a Linux operating system. The preferred distributions are RedHat Server and Ubuntu Server.*

SR-SYS-SW-002 *The Processing System Software shall avoid I/O overhead between processing steps*

SR-SYS-SW-003 *The Processing System Software shall be capable to exploit multi-core CPU architectures.*

SR-SYS-SW-004 *The Processing System Software shall translate a job into a number of tasks running in parallel on a cluster of machines.*

SR-SYS-SW-005 *The Processing System Software shall force local data I/O to avoid network I/O.*

SR-SYS-SW-006 *To provide load balancing, the Processing System Software shall be capable to assign tasks to idle or to newly added computing nodes.*

SR-SYS-SW-007 *The Processing System Software shall provide means to monitor performance and to generate processing statistics.*

SR-SYS-SW-007 *The Processing System Software shall integrate all GlobAlbedo processors.*

Realisation: The Processing System Software will be developed by BC and installed onto the Demonstration Hardware (see §4.5.7). It will be used to optimise for maximum performance and scalability. Possibly the Apache Hadoop technology will be used.

Priority: Essential	**Effort:** Months	**Risk:** 1-Risk
Major processing component	The development and refinement of the Processing System shall take place in KO+9 through KO+16.	Only short-term expertise available with the development of this type of processing system software. However, the experience with the Apache Hadoop technology is very encouraging so that the non-functional performance requirements can be met by the system.

4.5.7 Processing System Requirements: Hardware

SR-SYS-HW-001 *The Processing System Hardware shall provide as much CPU cores as needed to fulfill the non-functional performance requirements described in §4.5.8.*

SR-SYS-HW-002 *The Processing System Hardware shall be easily scalable: it should be possible to add both CPU/RAM and disk storage (= new servers) with minimum effort.*

SR-SYS-HW-003 *The Processing System Hardware shall be fault tolerant in terms of recovery from HW failures. Broken hardware shall be replacable wiithout the need to reconfigure the whole system.*

Realisation: Processing System Demonstration Hardware will be purchased and made available by BC. Together with the Processing System Software it will be used to optimise for maximum processing performance and scalability. The Demonstration System should be understood as a fully functional Processing System that is able to perform all GlobAlbedo processing steps from end to end but with limited performance. To meet the performance requirements of the project, the requested GlobAlbedo processing will be performed on larger systems. The specifications of the target hardware will be derived from the specifications of the demonstration hardware.

Priority: Essential	**Effort:** Weeks	**Risk:** Safe
Major processing component	The installation and set-up of the machine is expected to take not longer than 2 days.	Well-known tasks.

4.5.8 Non-functional Requirements

RQ-NF-001 *The GlobAlbedo Processing System shall be capable to process the specified input satellite data (partially 15 years of acquisitions) in one month.*

Realisation: From the performance of the Processing System Demonstration Hardware and Software it might be deduced what the scaling factor will be to meet the performance requirements of the project.

Priority: Essential	**Effort: N/A**	**Risk: 1-Risk**
Major performance requirement	Prerequisite.	Resource-Demanding.

4.6 Approach to minimising dependency on non-European data sets.

No non-European data is planned to be employed in the generation of SDRs or the production of full retrievals. However, rather than use proxy methods such as LUT via land cover or develop a model Earth with similar phenology as the real one to complete gaps, a different approach is being taken within this project as described below. This approach relies on the concept of determining incremental "information" updates from a base dataset of a decade of measurements.

It is very likely that the information content of the sensors proposed for use in GlobAlbedo have significantly less information content (in their ability to constrain estimates of albedo) that the data used in some other products, notably the MODIS BRDF/albedo product. This is mainly because of spectral sampling issues (particularly for MERIS), angular sampling (particularly MERIS and (A)ATSR/ATSR-2), and viewing opportunities (comparing those of VEGETATION with 2 MODIS sensors). The best scientific dataset then would most likely be obtained from a combination of data from all of these sensors, including MODIS. This is however not feasible within this project, and there is a further desire to generate a product that is not *dependent* on non-European data (e.g. to safeguard future production). It is therefore proposed here that non-European data will only be used here in the generation of one-off ancillary datasets for this product. In particular, MODIS data will be used to provide a prior estimate of model parameters $\underline{f}'(T)$ and associated uncertainty information $\underline{C}^{-1}(T)$, as well as initial estimates of the temporal uncertainty terms s_B^2. These ancillary products will be targeted at a particular year of production (2005). After this point (in processing both forward and backwards in time), posterior estimates derived via equation 1 will be used, so the dependency will be reduced with time and the independence of the datasets increase.

To understand the dependence of the product on the MODIS 'prior' the relative contribution of the prior (related to the average innovation of the observations over the prior) will be quantified and output into the product. For the year 2005, this will refer directly to the influence of the MODIS prior on the GlobAlbedo product. For processing of this initial 'test' year, a version of the product will also be generated for selected areas that makes no use of the MODIS product, so that a complete justification to use the MODIS prior can be made. For subsequent years of processing the innovation measure will refer to a lesser influence of MODIS as it is gradually 'watered down'.

There are many significant advantages to using a 'prior' in the processing here. The reasons for using MODIS for this is that this exists, is accessible, and has 10 years of data from which this product will benefit. The use of a prior obviates the need for a backup algorithm. The longer, weighted moving window will make use of whatever observations happen to occur during the (e.g. 3 month) observation time period in a manner analogous to a smoother/regularisation scheme. If the observations are particularly sparse in time, even this will be a relatively poor estimate of the surface parameters. The prior estimate (that resulting from the previous year) allows for the transfer of the 'shape' of the land cover dynamics from one year to the next, as well as the regulariser working locally in time. The only time this may cause problems is if there is e.g. land cover change from one year to the next, so the prior is a poor estimate. There are two approaches that could be taken to that: (i) if there is a large discrepancy between the observations and the prior, then simply down weight the prior further; or (ii) approach the 'prior' from both the year ahead and the previous year in a 2-way sweep of the processing and consider outlier detection as part of the process. The second option is not really feasible within the processing limits of this task, but might be a worthwhile future option (in fact, this might as well be considered as a 2-way sweep of a regularisation filter, giving a smoother). Clearly, the degree of influence of the prior is essentially assigned by the weighting given to it, relative to uncertainties (via quality, number and angular and spectral sampling) of the observations. This relative weighting (controlled largely via the additional uncertainty assigned to the posterior in moving forward one year and forming the new prior) should be set at a level that will produce an albedo product of the required accuracy. It is not possible to know this without some (significant) experimentation, but the aim is to set the lowest possible weighting (highest uncertainty) on the prior that will allow the observations and down-weighted prior produce the required accuracy.

4.7 Proposed product formats and delivery mechanisms

The Globalbedo products are stored and distributed in NetCDF format files.

4.7.1 File format

According to the UniData NetCDF page: "NetCDF (Network Common Data Format) is an interface for array-oriented data access and a library that provides an implementation of the interface. The NetCDF library also defines a machine-independent format for representing scientific data. Together, the interface, library, and format support the creation, access, and sharing of scientific data. NetCDF readers allow to extract information from the Globalbedo products. Such readers are included in BEAM but also commercial packages such as IDL, Matlab, etc.

4.7.2 Meta Data

A convention shall be used to define the metadata describing each variable in NetCDF format, so that it is self-contained. The recommended convention to be used for Globalbedo products is the NetCDF Climate and Forecast (CF) Metadata Convention. This convention already defines a certain number of variables especially in the field of climate and forecast. However, probably this convention may not include all the definitions of variables related to Globalbedo products.

Globalbedo intermediate BB SDR as well primary albedo products are defined on a sinusoidal grid (SIN), described in detail below. The final user products will be remapped to a Plate-Carrée grid as well as the MODIS SIN grid.

4.7.3 SIN Grid

The average size of the grid bins is equal to 1/120°, leading to 21600 rows in latitude, i.e. 10800 latitude rows in each hemisphere (the equatorial line is located between two rows of bins). This discretisation corresponds to roughly 1 km. Just below and above the equatorial line, the rows have 43,200 bins. This number decreases regularly from the equator to the poles where the last latitude row have only 3 bins. For each row, the left side of the first bin is always aligned with longitude -180° while the right side of the last bin is aligned with longitude +180°, covering all the latitude row area. The number of bins per row is always an integer number, computed in order to have the bin cell size as close as possible to 1 km, so that the effective longitudinal bin size may vary from one row to the next one. Applying this simple rules from South to North pole leads to a total of 594 042 200 bins.

Only grid cells with data are stored, i.e. empty grid cells will not inflate the product size. Globalbedo product are defined over land surfaces only, which makes roughly 1/3 or 200 000 000 bins.

Pseudo-code is available that can be used to describe the sinusoidal grid, i.e. to link the index of each bin (n) to its geolocation in longitude and latitude (bin_lon[n], bin_lat[n]).

4.7.4 Product overview

The files shown in *Table 4-1* will be generated:

Product	Grid	Approximate Size*	Number of Products	Persistence
MERIS BB SDR	ISIN	400MB * No of bands	1 / day **	Temporary
(A)ATSR BB SDR				
VGT BB SDR				
Primary BB Albedo	ISIN	400MB * No of bands	1 / <integrating period>	Archive
User defined BB albedo	Plate-Caree	Depends on user definition; max 1.7GB per band for full earth in 1km.	Depending on user requests	Temporary

Table 4-1: Summary of GlobAlbedo products

* For this estimation it is assumed that all data will be stored as 16 bit integers with scaling factors.

** compiling all 14 orbits / day

4.7.5 Online Product availability

GlobAlbedo BBA products will be available through 3 different mechanisms:

- A web-GIS for display of individual Web Map Services (WMS) and (probably) FLASH animations of products which is OGC-compliant using the Open Source University of Minnesota Map Server (UMMS) and the open source OpenLayers GUI. This will be based on the CEOS-WGISS ICEDS GEO data portal and its recent update in DEMqis. This will include the ability to generate still and animated sequences for display in Google Earth
- The full MERCI system for extraction of product subsets by geographic area and time duration including single pixels, averaged sets of pixels over small windows from both GlobAlbedo and contemporaneous global albedo datasets

- A version of the NASA GIOVANNI system for inter-comparison of global, regional, local, point and transect datasets.

4.8 Planned test data set for algorithm prototyping.

In [RD-1], the year 2005 was recommended and currently, subject to there not being any issues with data input, it is envisaged that 2005 will be employed. Initially testing will take place using two months (tbd) taken from this test year.

4.9 Detailed Validation plan for GlobAlbedo products

Validation of the albedo products, including a specification of the in-situ and ancillary data to be used.

SDRs and atmospheric correction

The algorithm operation will be evaluated by comparing retrieved AOD values with AERONET values. An estimate of error in surface reflectance based on AOD error will be made. In addition, the method of Lyapustin et al. (2006) will be followed. In this approach the sun-photometer data themselves are used for atmospheric correction at locations in close proximity to the AERONET sites. In principal this gives us a more accurate atmospheric profile that we can obtain using the retrieved satellite inversion estimates. A set of surfaces reflectances using the AERONET dataset and another set of surface reflectance calculated from the satellite inversions can be compared.

During the ESA sponsored MERIS AlbedoMap project, a number of novel methods were developed for the validation of spectral and broadband albedo products. It is planned to re-use these methods within GlobAlbedo albeit on a much larger and more systematic scale. As previously mentioned, we plan to use a QA4EO approach to systematise the validation approach.

The FAPAR and NDVI products will be derived from the SDR data for each data set. NDVI is calculated as a simple ratio between red and near-infrared bands, so no dedicated validation beyond the SDR validation is foreseen. In the case of FAPAR, direct comparison with the MERIS MGVI/FAPAR (Gobron et al., 2007) and VEGETATION FAPAR products (Baret et al., 2007) will be carried out.

4.9.1 Evaluation of the Albedo products

In AlbedoMap, the products were evaluated by a number of simple, but highly effective methods. These included false colour and natural colour composite display which identified insufficient removal of aerosols. Display of time sequences as animations which identified snow and cloud issues. Display of bit-map fields of QC bits, specifically full inversions vs magnitude inversions vs cosmetic fill and their time sequence help to understand the role of clouds. Analysis of the number of pixels by year coming from different sources over all land pixels helped identified issues concerned with insufficient looks due to cloud cover which were not helped through the use of combined Aqua+Terra BRDFs. These only helped with increasing the number of full inversions cf magnitude inversions. Production of natural and false colour composites quicklook browse images. In the case of GlobAlbedo, as we will only be producing broadband products, some of these visualization techniques are not feasible. However, use of pseudo colour display of VIS-

BBA, NIR-BBA, BBA as well as false colour composites of these three are likely to be very insightful.

4.9.2 Inter-comparison with European data products

SU and UCL have previously processed VEGETATION into BBA products and these will be compared to GlobAlbedo results, both on a single pixel level for specific field sites, in zonal mean displays and in animated side-by-side animations to identify locations and reasons for differences. The MERIS-BBAs will be similarly compared against those from GlobAlbedo for the 4.5 years of coincidences, again both for individual pixels, global differences and zonal variations.

4.9.3 Inter-comparison with NASA data products

MODIS have completed production of Collection 5 at the end of the MERIS AlbedoMap project but too late to be employed for inter-comparison. Given that MODIS has an additional spectral band at 2.2 µm, simulation studies will be carried out to assess what impact, if any, this extra band will have on the resultant accuracy of the BBA. MODIS BBAs will be compared against those from GlobAlbedo after being degraded to 1km and transformation into the same co-ordinate system and map projection. An extension to BEAM will be written to allow ingest of MISR Level 2 BBAs using the JPL MISR toolkit. This will enable instantaneous MISR BBA to be directly compared to the aggregated BBA planned to be produced within GlobAlbedo. This inter-comparison will be targeted at specific sites where anomalies have been detected in the MODIS-GlobAlbedo. An inter-comparison will also be carried out at the 0.5 degree level using the MISR Level 3 monthly, seasonal and annual products to determine whether there are any systematic differences in BBA.

4.9.4 Inter-comparison with CNES data products

Previous attempts to read POLDER and PARASOL products were unsuccessful due to vagaries of IDL and the map projection system employed by these CNES products. Since the end of the ESA AlbedoMap project, these problems have been overcome. For lower resolution GlobAlbedo products, inter-comparisons will be performed using the methods previously described.

4.9.5 Inter-comparison with METEOSAT data products

A variety of historical and current METEOSAT BBA products are in production or archived [R38-40]. Previous intercomparisons showed that MODIS and MISR had similar BBAs but specific meteorological conditions could lead to large differences (loc. cit.). Further investigations will be carried out to determine if the same situation applies with GlobAlbedo products.

4.9.6 Inter-comparison with *in situ* data products

In situ albedometer data which has been screened and filtered by colleagues at BU will be provided to the project for use in intercomparison with the satellite-derived values. These will include BSRN measurements and those over flux tower sites (FLUXNET) recently described by Roman et al (2009). The principal difficulty here will be to ensure that the *in* *situ* data are representative of a large enough area. Roman et al (2009) provided a proposal for a standard uniform method which is applicable not just to forested sites but is being applied by colleagues of C. Schaaf (BU) to BSRN sites to determine which ones are representative of the landscape and which are not. We will take the result of these assessments to select albedometer data from BSRN and FLUXNET sites which are spatially representative at 1km. BU have also filtered observations for cloud-free and fully cloud-covered which will be flagged in their records so that appropriate adjustments can be made to these albedometer measurements to make them compatible with that derived in GlobAlbedo products.

G. Schapmann-Strub has offered to provide the GlobAlbedo project with albedometer measurements from a forest campaign being held in Siberia during the summer of 2010. These measurements will be similarly screened for clouds. Dr Schapmann-Strub will also investigate if there is suitable albedometer which might be made available to the GlobAlbedo project which has been collected in Russia over the last 2 decades which would be screened in a similar manner to that from BU.

Given the spatial representativeness of these albedometer measurements made after applying the Roman et al. (2009) methods it should be possible to directly compare these albedometer measurements with GlobAlbedo products transformed into the same blue sky estimates employed by the ESA AlbedoMap project and by MODIS.

4.9.7 Use of the MERCI system for storing the GlobAlbedo products (and possibly the SDRs in future)

The MERCI (MULTIMISSION EARTH OBSERVATION CATALOGUE AND INVENTORY) system is a software-based platform to manage and process Earth Observation data regardless of the data volume and regardless of the storage media – even and specifically online storage is supported, either partly or for the complete archive of data to be managed.

MERCI is used internally by Brockmann Consult to manage an online archive of many years of MERIS data, and is installed in various international organisations such as ESA. MERCI provides all functions for data acquisition, registration, quality control, product query, child product generation and dissemination. MERCI comes with a Web-Interface to make the EO data accessible from any place in the world. MERCI is currently supporting MERIS, A(A)TSR and MODIS data. Support to other data is under development which will include the GlobAlbedo products.

MERCI capabilities include the following:

- Product registration: EO data coming from any source media (CD, DVD, Satellite Link) are registered in a product database and organised in a file repository.
- Quick-Look generation: During product registration, quick-look and thumbnail images are generated and made available through the Web Interface
- Statistics: A large number of statistical information is generated during registration: blank line, numbers of flagged pixels, averages of band values, …

- Automated Quality Control: Tests can be defined on the statistics and product metadata to perform quality control of products during registration. A final “yes/no” is calculated and the test details are available through the Web.
- Product query: A powerful query interface includes the statistical parameters and an interactive map for searching relevant products from the archive
- Child product generation: Working on the quick-look image, a subset of a full orbit product can be selected and generated on-the-fly
- Pre-Defined sites of interest: Fixed areas of interest can be defined which are automatically extracted from full orbit products.

MERCI installation has been accomplished at ESA where MERCI is used to distribute reprocessed (A)ATSR products. The installation at the UK-PAC, Farnborough, consists of one web-server computer and four compute nodes that access 70 TB NAS storage.

This system hosts the complete archive of the ATSR-1, ATSR-2 and AATSR sensor products from 1991 up to today.

The albedo products from GlobAlbedo, MODIS, MISR and the NASA-NOAA LTDR for AVHRR is plnned to be stored in this system alongside new products to be generated from the ENVISAT-DDS receiving station at UCL-MSSL.

4.10 Limitations of the algorithms and validation techniques.

The chief limitations of algorithms for atmospheric correction are:

- Over snow and ice surfaces, algorithms are currently untested
- For VGT, algorithms have yet to be developed and accuracy is currently unknown
- For MERIS and VGT, performance is only robust over dark targets; however the effects of errors in AOD on SDR is lower over desert targets, as surface reflectance is nearer the ‘critical point’ where TOA reflectance is insensitive to AOD.
- For many pixels we will rely on interpolated AOD; at points distant from the retrieved AODs error is likely to increase, but will be modelled

Limitations, validation of atmospheric correction

Direct validation of SDR at spatial scales of 1km is difficult due to required simultaneity of measurement, correspondence of view angle and instrument calibration

accuracy. Validation of the atmospheric correction will therefore focus on AOD, which can be measured by AERONET.

4.11 Delivery of products

A complete set of the GlobAlbedo products shall be delivered to ESA and to each User Group member on suitable media.

4.12 Traceability to Baseline Requirements

The traceability between the BR and the TS are shown in GlobAlbedo tractability matrix (accompanying excel sheet).