SENTINEL-3 OPTICAL PRODUCTS AND ALGORITHM DEFINITION

Land Surface Temperature

DOCUMENT REF: S3-L2-SD-03-T03-ULNILU-ATBD_L2LST

DELIVERABLE REF: SD-03-T

VERSION: 2.3

CREATION DATE: 2009-08-30

LAST MODIFIED: 2012-10-10

DOCUMENT SIGNATURE TABLE

	NAME	FUNCTION	INSTITUTION	SIGNATURE AND DATE
PREPARED	J. Remedios	Consultant	U. Leicester	10-10-2012
VERIFIED	S. Emsley	Project Manager	ARGANS	10-10-2012
APPROVED				

DOCUMENT DIFFUSION LIST

ORGANISATION	NAME
ESA	Philippe Goryl, Alessandra Buongiorno and Carla Santella
EUMETSAT	Vincent Fournier-Sicre and Vincenzo Santacesaria
Consortium Partners	ARGANS, ACRI-ST, RAL, Brockmann-Consult, Elsag-Datamat

DOCUMENT CHANGE RECORD

VERSION	DATE	AUTHOR	DESCRIPTION
1.0	30/08/2009	J. Remedios	PDR submission
2.0	07/04/2010	J. Remedios	CDR submission
2.1	29/04/2010	J. Remedios	CDR RID update
2.2	02/07/2010	J. Remedios	Delta-CDR
2.3	10/10/2012	J. Remedios	FR submission

Table of Contents

List of Figures

1 INTRODUCTION

1.1 Acronyms,Abbreviations and Symbols

ATBD	Algorithm Theoretical Basis Document
AATSR	Advanced Along-Track Scanning Radiometer
ATSR	Along-Track Scanning Radiometer
BT	Brightness Temperature
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CYCLOPES	Carbon cYcle and Change in Land Observational Products from an Ensemble of Satellites
ECMWF	European Centre for Medium-Range Weather Forecasting
ERA-40	ECMWF Re-analysis 40 year project
FV	Fractional Vegetation
HITRAN	Hi-resolution Transmission database
IGBP	International Global Biosphere Program
LST	Land Surface Temperature
MODIS	Moderate Resolution Imaging Spectroradiometer
MODTRAN	Moderate spectral resolution atmospheric transmittance algorithm
NDVI	Normalised Difference Vegetation Index
NRT	Near real-time
RFM	Oxford Reference Forward Model
SLSTR	Sea and Land Surface Temperature Radiometer
SST	Sea Surface Temperature
TOA	Top of Atmosphere

1.1.1 Symbols

- I_v is the top-of-atmosphere radiance at the radiometer,
- $I_v^{surface}$ is the surface leaving radiance,
- I_v^{atmos} is the top-of-atmosphere radiance from the atmosphere,
- τ is the atmospheric transmittance,
- v is wavenumber,
- z is height and p is pressure
- F is the filter response function of the radiometer,
- s is a unit vector defining the view direction,

- $\mathbf{s}'$ is a unit vector defining the sun's direction,
- T_s is the surface temperature,
- ϵ_ν is the surface emissivity,
- B_ν is the Planck function,
- ϱ_ν is the surface reflectance,
- I_ν^{sky} is the downwelling sky radiance.
- T_{11} and T_{12} are the brightness temperatures of the measurements channels and 11 and 12 microns respectively; δT_{11} *and* δT_{12} are the radiometric noise in each channel
- θ is the satellite zenith view angle, and ϕ is the satellite azimuth view angle
- $T(z)$ is the atmospheric temperature profile,
- z is height,
- $\tau(z, z')$ is the transmittance profile between two heights.
- $a_{f,i,pw}$, $b_{f,i}$, and $c_{f,i}$ re classes of coefficients, applied to BTs; a_0, b_0, and c_0 are the theoretical equivalents
- $n = \frac{1}{\cos(\theta/m)}$,
- m is a variable parameter controlling the dependence on view angle,
- f corresponds to fractional vegetation.
- i corresponds to vegetation type (biome),
- pw corresponds to precipitable water (in cm).
- The error, δe_{rand}, is the random error.

1.2 Purpose and Scope

This document presents the Algorithm Theoretical Basis Document (ATBD) for the retrieval of Land Surface Temperature (LST) from thermal radiances measured by the Sea and Land Surface Temperature Radiometer (SLSTR). LST is an extremely important parameter that controls the exchange of longwave radiation between the surface and atmosphere. Because of the extreme heterogeneity of most natural land surfaces, LST is a difficult parameter to estimate and to validate. Algorithms for deriving land surface temperature using split-window radiances are sufficiently advanced that accuracies of 1-3 K are possible. Better accuracies (closer to 1K) are obtained at night, when differential surface heating is absent although cooling differentials can occur.

There are a number of factors that fundamentally influence the derivation of LST including temperature variations with angle, sub-pixel inhomogeneities in temperature and cover, surface spectral emissivity at the channel wavelengths, atmospheric temperature and humidity variations, clouds and large aerosol particles such as dust. Cloud flagging, which tends also to remove larger concentrations of aerosol dust, is usually performed independently of LST retrieval, and so is considered separately (see section 3.5.4). The chief drivers for LST algorithms are therefore surface spectral emissivity and atmospheric correction combined with angular effects on temperature and emissivity. The angular effects on directional temperature and emissivity have proved difficult to estimate and are the subject of continuing investigations. Hence, the LST algorithm at this stage must rely on nadir split-window radiance only.

There is an increasing amount of work on surface spectral emissivities. However, the absolute values of surface spectral emissivity and variations with wavelength are not known to sufficient accuracy and not yet verified over large parts of the land surface. Therefore, the LST algorithm must be restricted to the two closely adjacent thermal channels at 11 and 12 μm, and the algorithm should not maintain an explicit dependence on surface emissivity. Instead, dependences on emissivity are indirectly incorporated via statistical regression coefficients calculated off-line and translated through biome/fractional vegetation maps for application to a particular SLSTR pixel; fortunately the radiative transfer equations are sufficiently linear for such approaches to be realistic. These coefficients also account for variations in atmospheric temperature and humidity.

In summary, the LST algorithm specifiable at this stage for SLSTR should be a nadir-only algorithm operating on two split-window thermal channels at 11 and 12 μm, and employing a coefficient approach to derive LST where the coefficients are derived by regression.

1.3 Algorithm Identification

The algorithm consists of:

- A split-window LST algorithm which is based on a regression coefficient approach.

- The LST algorithm is based on three *classes* of coefficients applied to the nadir-views of the 11 and 12 µm channels. The three classes of coefficients are $a_{f,i,pw}$, $b_{f,i}$, and $c_{f,i}$ corresponding to the basic functionality of weighting with respect to brightness temperature.

- There are two detectors providing the 1 km view in interleaved fashion, potentially with different spectral characteristics. The characteristics below are for each level 1b pixel of each detector. It is expected that level 1b data are obtained independently for the output from each detector, so that the LST can be analysed separately for each. In this case, each detector has its own *base retrieval coefficients* in principle. Base retrieval coefficients are the coefficients that are actually supplied in an auxiliary data file as coefficient values.

- The retrieval equation will be applied to each level 1b pixel of each detector. LST should be derived for orphaned pixels which should be identified as such. Cosmetic pixels should also be identified and LST values retrieved for them.

- LST values should be retrieved for all level 1b pixels whether cloud flagged or not. The output file should contain the cloud flags so that the user can decide how to treat the data.

- For each biome, six base retrieval coefficients must be supplied (two per class of coefficient), corresponding to the fully vegetated and bare soil states. The retrieval coefficients will be supplied for both day and night (making 12 retrieval coefficients in all) for each biome. Note again that there are 12 retrieval coefficients for each detector element. There are also two constant parameters, per biome, *d* and *m*, whose values will be supplied as auxiliary parameters which are the same for both detector elements. The *d* and *m* should be supplied in auxiliary files as 2-d arrays with maximum dimension set to the maximum number of biomes and values for both day and night for each biome. In addition, a column for utility flag may be needed (next point).

- There is also a further constraint that when T_{11} is less than T_{12}, the *n* factor should be set equal to one (this can also be achieved by specifying the parameter *m* for this case). In case of further algorithm updates, it would be wisest to incorporate a utility flag (yes/no) as input to determine whether this process should be used. The flag could be incorporated into the auxiliary file with *d* and *m* as an additional column to the two-dimensional arrays.

- Base coefficients should be provided on a monthly basis with the *option* (enabled in the first instance) for interpolation of LST derived from the two nearest sets of monthly base coefficients.

The coefficients will be supplied in one file with a time stamp common to other monthly data sets (centre of the month in the first instance). If annual coefficients are desired, then the monthly data sets will simply be repeated with the same coefficients. Annual average coefficients are a minimum.

- If no interpolation is performed on a monthly basis, then all base coefficients will be changed at the monthly date boundary.

- Day and night base coefficients are selected based on solar zenith angle. The solar zenith angle switching point should be specified in a file but should be set to 90° in the first instance. There will be no interpolation between day and night base coefficients.

- In principle (first approach), the base coefficients should be interpolated first and then the LST derived using the interpolated base coefficients and the interpolated auxiliary data. This should ensure that the correct auxiliary data are also written to the output file; auxiliary data per pixel in the output file must be consistent with final output LST per pixel. A second approach is to calculate two LSTs and then interpolate the LSTs, however in this case great care must be taken also with the auxiliary data to ensure consistency. Therefore the first approach is recommended.

- The base retrieval coefficients must be generated by regression of simulated brightness temperatures against a suitable high specification line-by-line forward model with appropriate ranges of parameters for surface temperature, surface spectral emissivity, atmospheric profiles of temperature and humidity, and also contaminant gases. [These aspects except for water vapour are entirely implicit in the algorithm implemented in the operational processor]

- The derivation of base retrieval coefficients requires an accurately known spectral response function for each SLSTR detector element from the instrument specification/calibration and each of the two thermal channels.

- For the algorithm, the sensor view angle (from nadir) at the surface is required. For day/night specification, error characterisation and verification, the solar zenith angle is required.

- Auxiliary data must be supplied which will specify, for a given spatial location: biome, fractional vegetation and precipitable water. The number of biomes is currently those of Globcover, i.e. 22, but for flexibility in likely processor updates, should be assumed to be a maximum of 50. All files will be filled (no missing values).

- Auxiliary data and auxiliary files are not necessarily static and the processing must have the flexibility to incorporate updated information in the future. Therefore auxiliary information for each input parameter would be stored in separate file, but all information on the input parameter is stored together in one file. Tiled and packed forms of files are acceptable.

- For SLSTR processing, after testing it is recommended that (1/120)° resolution be employed currently for the biome auxiliary data file but with the capacity to expand this as required, e.g. to double the resolution; (1/120)° corresponds to approximately 1 km at the equator.

- The biome map is expected to be static with a slow update period e.g. annually or less often. Annual would be the most desirable but will depend on availability of Globcover-type data.

- The appropriate biome for analysis for a given pixel should be selected as the biome which is the nearest neighbour.

- Modules to provide both un-gridding of the SLSTR pixel map (level 1b) to original measurement space (level 1a) and re-gridding to 1 x 1 km^2 space should be designed and made accessible as part of the operational processor design. It is understood that level 1b to level 1a referencing is provided. Re-gridding back to level 1b would be a standalone tool and should be made available.

- For the application of the biome map, it will be essential to have accurate cross-referencing of the sensor and the co-ordinates of the biome map.

- The baseline fractional vegetation (FV) fields for the LST algorithm will be climatology with relatively high time resolution, consisting typically of files spaced at, for example, ten day intervals. Spatial resolution will be close to that of the biome data. The data will be time stamped. To allow for off-line re-processing or for NRT generation of the auxiliary data, the file specifications should include year date. The FV fields will be close to but not necessarily exactly the same as the spatial resolution of the biome map. A nearest neighbour approach can be employed to find the relevant FV fields for the LST. FV fields will be interpolated to the current time of the measurement.

- As an additional check, the FV field should be limited to an upper value of one in the code.

- Water vapour data should be obtainable from two sources: (a) A monthly climatology file of water vapour for each year (resolution of no greater than 0.l°); (b) total column water vapour from meteorological analyses. In the processing there should be an option to switch between these two sources of water vapour (the option to use meteorological data should be enabled in the first instance). The monthly climatology file shall be the baseline for the algorithm. Water vapour fields will be interpolated to the current time of the measurement. If meteorological fields are not available, then the climatology will be used. For water vapour, interpolation between analysis fields and forecast fields is acceptable. In near real-time operations, if the forecast field is available then climatology will have to be used. In off-line processing, analysis fields should be used and not forecast fields.

- The water vapour climatology will provided as filled fields with twelve monthly fields in one annual file (which can be updated from year to year if required e.g. for re-processing). The water vapour climatology should be interpolated to the required point in time, latitude and longitude.

- All auxiliary data used in the calculation of LST per pixel will be written to the output file per pixel; retrieval coefficients are not auxiliary data in this sense. The total water vapour column per pixel will be written in the output file.

- There should be two numerical error fields available in the output level 2 file: (a) the LST random error inferred from radiometric noise (calculated similarly to SST; (b) the total error which is the radiometric noise error combined with the absolute uncertainty loaded from an auxiliary look-up table. If only error field is provided, then this should be the total error.

- The random error should be calculated analytically.

- The absolute uncertainty (systematic) errors will be calculated through three independent variables: biome, latitude (temperature; see below), and total column water vapour. Each error field will be a 1-d array with an integer giving the actual dimension of the each array (start, increment and number of values also specified in the file). The maximum dimensions expected for each array are 50 (maximum number of biomes), 18 (latitudes), and 20 (water vapour).

- The output level 2 file should also contain the following error values and quality flags: radiometric error; total uncertainty which is a combination of radiometric error and look-up table errors; an aerosol quality flag (see below).

- A validation flag should be set through an external file as a value per biome. This should have the values: 2: fully validated; 1: some verification information available; 0: not validated.

- In *future evolution* of the algorithms, it is expected that the following algorithms might be implemented and so design of the processor should include now the functionality which would enable these to be implemented later.

  - *Dynamic fractional vegetation*. In this approach, a two step process will be used for fractional vegetation: (a) for day time retrievals, an operational SLSTR fractional vegetation will be used using coincident views of the visible channels; (b) for night time retrievals, a 'live' or dynamic climatology will be used with values being held in an auxiliary data file; (c) The climatological fractional vegetation map should be provided at the same spatial resolution as the auxiliary biome data file. The values in the file will be non-integer and fall between 0 and 1; (d) the auxiliary data file will need to be updatable on a weekly basis with a process running in parallel which takes daytime L2 NDVI values and combines with the previously available climatology to produce a new climatology for use with nighttime data.

    To provide useful dynamic auxiliary information to the user and to allow for the dynamic fractional vegetation upgrade in the future, NDVI should be calculated and provided in the output file.

- ➢ *Use of aerosol information to flag LST data*.

  An aerosol flag should be implemented in the quality flags of the level 2 LST output files.

  Provision should be made in the LST processor to make use of external aerosol information (e.g. from other SLSTR information, synergy information, or from ECMWF model aerosol fields if available). These data would need to be processed in an aerosol flag module, filtering the data to that required for thermal infra-red LST retrievals, i.e., dust particles. Alternatively, a direct aerosol (dust) flagging method using SLSTR brightness temperatures could be implemented in the LST processor.

- ➢ *Pre-processor for cloud*. The operational processor should maintain provision for a pre-processing cloud module in the level 2 processor for LST which would identify clouds more rigorously and specifically for the thermal infra-red channels. This algorithm might either be a more specific form of the level 1 cloud clearing or else a Bayesian-type approach for LST as for SST.

- ➢ *Improved specification of errors*: In the future, the latitude dependence is likely to be replaced by a temperature–related term which is more physical. However this has not been defined as yet. Error dependences on topography (digital elevation difference from sea level), view angle and solar zenith angle will be investigated. In addition, a cloud error parameterisation would be useful but cloud information is not yet available in suitable form.

- The cloud flag in the level 1b data will be used initially for cloud removal for the LST processor. This will be the baseline processor. However, the operational processor should include provision for a pre-processing cloud module in the level 2 processor for LST which would identify clouds more rigorously and specifically for the thermal infra-red channels.

In summary:

The baseline option shall consist of a split window algorithm with three classes of coefficients and a total of 12 retrieval coefficients for each biome (day/night coefficients; two per class of coefficient). The coefficients will be supplied for each month (annually by setting all monthly fields to the same values), and applied through the use of auxiliary data including a biome map (infrequently updated; typically annual); climatological fractional vegetation; water vapour climatology with option to use meteorological analyses such as ECMWF. The cloud flagging will be taken from the level 1b cloud detection scheme.

Future updates will include dynamic vegetation, most likely utilising NDVI from co-located SLSTR measurements in the daytime and a dynamic nighttime climatology; improved use of ECMWF water vapour data; a dedicated cloud pre-processing module; an aerosol (dust) flag technique; improved specification of errors.

2 ALGORITHM OVERVIEW

2.1 Objectives

The objective of this algorithm is to provide a measurement and its associated uncertainty of LST from the top-of-atmosphere (TOA) brightness temperatures (BT) measured by the SLSTR. The algorithm is an evolution of the LST algorithm developed for the Advanced Along track Radiometer (AATSR) by CSIRO (Prata 2002). The algorithm is soundly based on radiative transfer theory as applied to the exchange of radiation between the surface and atmosphere. The effects of land surface emissivity are *implicitly* taken into account in these algorithms via biome and fractional vegetation. The basic algorithm may be stated as:

$$LST = a_0 + b_0 T_{11} + c_0 T_{12} \qquad (1)$$

where a_0, b_0, and c_0 are classes of coefficients, applied to brightness temperatures (BTs), that depend on atmospheric water vapour, satellite viewing angle and land surface emissivity. The T_{11} and T_{12} are the BTs at 11 μm and 12 μm respectively.

Hence the basic algorithm is a split-window (two channel) approach.

The essence of the algorithm is the recognition that over the land, both atmospheric water vapour effects and surface emissivity effects (spectral and angular) play important roles in modifying the amount of radiation reaching the satellite-borne radiometer. The basic physical sensitivities, particularly those of atmospheric correction for clear sky trace gases, can be accounted for by the coefficients provided that these are derived for an appropriate range of conditions. Some effects can be very complex (e. g. heterogeneity of mixed land covers) and suggest that accounting for them accurately, in the sense of approaching SST retrieval quality (better than 0.5 K), can only be done in a few special cases, when surface and atmospheric properties are very well known. Such cases do exist, and can be used to test algorithms and for validation.

Accounting for all of the complex effects introduced by surface heterogeneity, shadowing, terrain variability (e.g. height variations) and atmospheric variability may be achieved under special experimental conditions, but will not be possible for global conditions. Consequently the approach adopted is to determine robust regression coefficients that can be used for classes of landcover conditions, atmospheric water vapour loadings and seasons. These classes are quite broad - in the present case there are 22 landcover classes (Globcover) or biomes, covering desert conditions to pine forest. To include the effects of seasonal vegetation growth, coefficients are linearly combined and a time-dependency of both coefficients and auxiliary data can be included. Special coefficients are used for snow and ice covered surfaces, as well as lakes.

The advantages of this approach are:

1. Regression-based algorithms are fast and easy to implement on a computer.
2. The regression coefficients can be implemented as a lookup table and be updated in a routine manner.
3. Validation of the algorithms can be performed for a subset of surfaces by comparing LSTs directly, as well as validating sets of input variables and parameters.

There are two detectors providing the 1 km view in interleaved fashion, potentially with different spectral characteristics. It is expected that level 1b data are obtained independently for the output from each detector, so that the LST can be analysed separately for each. In this case, each detector has its *own retrieval coefficients* due to potentially differing spectral characteristics of each which must be known.

The retrieval equation will be applied to each level 1b pixel of each detector. LST should be derived for orphaned pixels which should be identified as such. Cosmetic pixels should also be identified and LST values retrieved for them.

Modules to provide both un-gridding of the SLSTR pixel map to original measurement space and re-gridding to 1 x 1 km^2 space should be designed and made accessible as part of the operational processor design.

LST values should be retrieved for all level 1b pixels whether cloud flagged or not because cloud detection schemes over land are limited in accuracy (see also section 3.5.4). The output file should contain the cloud flags so that the user can decide how to treat the data.

3 ALGORITHM OVERVIEW

3.1 Theoretical Description

3.1.1 Physics of the problem

The problem of determining LST from the split-window nadir channels of the SLSTR is similar in concept to the problem of determining SST. The so-called split-window method, which utilises the radiances reaching the sensor in two channels whose band centres are close in wavelength, can be used for both the SST and LST problems. The method provides an estimate of the surface temperature from two brightness temperature measurements and assumes that the linearity of the relationship results from linearisation of the Planck function (generally a good assumption), and linearity of the variation of atmospheric transmittance with column water vapour amount (sometimes a poor approximation). In the case of the ocean, because the emissivity is high and varies little, the surface and atmosphere are effectively decoupled and can be treated almost independently. For the land, where the emissivity can be much lower than the ocean and where emissivity varies significantly with surface cover and type, the surface and atmosphere must be treated as a coupled system. There are two approaches to solving the problem of LST determination using the split-window channels. The first assumes that the effects due to the land and atmosphere can be decoupled and the method is then to separate out the surface effects (emissivity) from the atmospheric effects (water vapour). The second approach is to accept that the surface and atmosphere are coupled, solve the problem without taking explicit account of either emissivity or water vapour, but to allow for their effects simultaneously. The difficulty of the first approach is that an estimate of the emissivity must be provided or retrieved and validated.

The approach used here is the second approach which is outlined mathematically in the following section. Having established that there is a linear relation between the surface leaving radiance and the two split-window radiances for the land, the problem is reduced to one of multiple, linear regression. The retrieval coefficients, derived by regression, have physical meaning and physical constraints can be utilised to ensure their validity.

The temperature that is retrieved using the algorithm is a radiative surface temperature; it is appropriate for use as the temperature corresponding to the radiative flux density from the surface (i.e. Stefan-Boltzmann law). When used in modelling studies care must be taken to ensure that the model output temperature corresponds to the SLSTR LST product definition.

The definition of LST from SLSTR is the effective radiometric temperature of the Earth's surface "skin" in the instrument field of view. "Skin" temperature here refers to the temperature of the top surface in bare soil conditions and to the effective emitting temperature of vegetation "canopies" as determined from a view of the top of a canopy.

3.1.2 Mathematical description

The mathematical development of the problem of determining LST from a satellite radiometer with split-window channels follows closely that of Price (1984), McMillin and Crosby (1986) and Prata (1993, 1994a,b). These papers show that under certain assumptions, it is possible to formulate the surface leaving radiance in terms of a linear combination of radiances reaching the satellite sensor in two channels close in their respective central wavebands.

The proposed SLSTR land surface temperature product (LST) is a gridded 1x1 km^2, pixel by pixel quantity using only the nadir split-window (11 and 12 µm) channels of the SLSTR. The product utilises the cloud-free top-of-the-atmosphere 11 and 12 µm brightness temperatures and ancillary information to correct for water vapour absorption and spectral emissivity effects; conventionally input radiances are gridded to 1 x 1 km^2 prior to retrieval (but see section regarding the use of un-gridded radiance data). The product is generated using a regression relation and look-up tables that accommodate global and seasonal variations in the main perturbing influences. The mathematical basis for the formulation is provided here.

The mathematical definition of the LST is given in equation (9), preceded by a treatment of the radiative transfer problem at the land-atmosphere interface and utilising several reasonable assumptions.

The starting point for any LST algorithm is a consideration of the thermal radiative transfer equation for monochromatic radiation emitted and reflected from a surface that is assumed homogenous, and received by a spaceborne radiometer. The homogeneous area is defined by the angular field-of-view of the radiometer. The radiance received at the satellite-borne radiometer may be written,

$$I_\nu(s) = \int_\nu \mathbf{F}_\nu\{\tau_\nu(\mathbf{s})\mathbf{I}_\nu^{\mathbf{surface}}(\mathbf{s}) + \mathbf{I}_\nu^{\mathbf{atmos}}(\mathbf{s})\}\mathrm{d}\nu, \qquad (2)$$

$$I_\nu^{surface} = \epsilon_\nu B_\nu[T_s] + \frac{1}{\pi}\int_{\Omega-} \mathrm{n}\cdot\mathrm{s}\,\varrho_\nu(\mathrm{s},\mathrm{s}')\mathbf{I}_\nu^{\mathbf{sky}}\,\mathrm{d}\mathbf{\Omega}, \qquad (3)$$

$$I_\nu^{atmos} = \int_0^\infty B_\nu[T(p)]\frac{\partial\tau}{\partial z}(z,\infty)\mathrm{d}z. \qquad (4)$$

where:

- I_ν is the radiance at the radiometer,
- $I_\nu^{surface}$ is the surface leaving radiance,
- I_ν^{atmos} is the radiance from the atmosphere,
- τ is the atmospheric transmittance,

- ν is wavenumber,
- z is height and p is pressure
- F is the filter response function of the radiometer,
- s is a unit vector defining the view direction,
- s' is a unit vector defining the sun's direction,
- T_s is the surface temperature,
- ϵ_ν is the surface emissivity,
- B_ν is the Planck function,
- ϱ_ν is the surface reflectance,
- I_ν^{sky} is the downwelling sky radiance.

If the surface is in thermodynamic equilibrium with the atmosphere, then according to Kirchhoff's law:

$$\int_{\Omega-} \mathrm{n}\cdot\mathrm{s}\epsilon_\nu(\mathrm{s})\,\mathrm{d}\Omega = \int_{\Omega+} \mathrm{n}\cdot\mathrm{s}\left\{1-\frac{1}{\pi}\int_{\Omega-}\mathrm{n}\cdot\mathrm{s}'\varrho_\nu(\mathrm{s},\mathrm{s}')\mathrm{d}\Omega'\right\}\mathrm{d}\Omega, \qquad (5)$$

We assume that the surface is Lambertian. Then ϵ_ν and ϱ_ν are independent of direction,

$$\epsilon_\nu = 1-\rho_\nu \qquad (6)$$

The flux density of sky radiation is:

$$F_\nu^{sky} = \int_0^{2\pi}\int_0^{\pi/2} I_\nu^{sky}\cos\theta\sin\theta\,\mathrm{d}\theta\mathrm{d}\phi, \qquad (7)$$

where θ is the satellite zenith view angle, and ϕ is the satellite azimuth view angle.

$$I_\nu^{surface} = \epsilon_\nu B_\nu[T_s] + \{1-\epsilon_\nu\}L_\nu^{sky}, \qquad (8)$$

$$L_\nu^{sky} = \frac{F_\nu^{sky}}{\pi}.$$

This leads to the definition of surface temperature as sensed by a space-borne infrared radiometer:

$$T_s = B_\nu^{-1}\left\{\frac{I_\nu^{surface} - (1-\epsilon_\nu)L_\nu^{sky}}{\epsilon_\nu}\right\}. \qquad (9)$$

This definition has the attribute that T_s is directly measurable from space (e.g. the AATSR), is valid at any scale, and for a homogeneous surface it is equivalent to the thermodynamic temperature.

The definition is only strictly true for monochromatic radiation. For sufficiently narrow channels (≈ 1 µm width) with relatively smooth filter response functions, the variation of the Planck function with wavenumber is small. Thus an integration of the various quantities (I_ν, ϵ_ν, L_ν^{sky}, etc.) over the filter function is appropriate.

The definition is only strictly valid under the assumptions outlined above. Under most circumstances we expect the assumptions to remain valid and violation are weak so that the definition (and hence derivation of the surface temperature) is approximately correct.

Determination of the quantities in (9) can be done by various means. The approach we have taken follows Prata (1994a,b) and shows that the surface temperature may be written as a regression relation involving the brightness temperatures in the 11 and 12 µm channels. The relation takes account of atmospheric absorption (water vapour) and spectral emissivity effects.

3.1.3 Emissivity

It is well-known that variations in surface properties cause variations in the emission of radiation from natural surfaces and this is at the root of some complications in LST retrievals.

One major source of variation is due to the structural properties of the surface and this affects the efficiency of emission and reflection of thermal radiation from the surface. There are substantial variations in surface emissivity over the global. The lowest values occur in sandy regions where the emissivity may be as low as 0.92 at 11 µm (Sutherland, 1979). Over highly vegetated surfaces (e.g. closed-canopy trees) the emissivity is known to be spectrally uniform and high (ϵ_{11} > 0.98, e.g. Salsbury and D'Aria, 1992). Within a particular surface type the variation of emissivity is not well known, but measurements suggest it is small ≈ ±0.01, except when structural changes occur as in senescent vegetation. Thus the greatest concern for deriving LSTs is the variation between surface types rather than the variation within surface types. The scheme for accounting for emissivity variations between surface types relies on a surrogate measure of the surface structure; in this case we have used fractional vegetation cover and vegetation type. Snyder et al. (1998) have suggested using a classification based emissivity system for MODIS LST products. Their system uses 17 IGBP 'static' land cover classes. Also of concern is the directional variation of emissivity. Generally, the variation is strongest with view angles greater than 50° or so. Little is known of the variation with azimuth angle.

While it is important to note the role that emissivity plays in determining the emission and reflection of thermal radiation from the land surface, it must be stressed that sufficient field measurements of emissivity at scales appropriate to the AATSR pixel size are not absolutely accurate, although increasingly available. Thus while it is possible to retrieve an emissivity from thermal satellite measurements, its validation is problematic. Moreover, none of the emissivity schemes proposed can claim accuracies better than ±0.02. It is likely that the retrieval errors and biases will be re-mapped from atmospheric transmittance errors, since the radiative transfer problem shows that the surface emissivity and atmospheric transmittance always appear as a product. Separating their effects accurately suggests that the atmospheric transmittance must be known at least to the same accuracy.

These factors should be borne in mind in considering the derivation of the LST algorithm itself in the following sections.

3.2 Split-window approximation

By utilising the mean value theorem it can be shown that (McMillin and Crosby, 1984):

$$\overline{I}_{\nu}^{atmos} = \frac{1}{1-\tau_{\nu}} \int_{0}^{\infty} B_{\nu}[T(z)] \frac{\partial \tau(z,\infty)}{\partial z} \, \mathrm{d}z, \qquad (10)$$

where:

- $T(z)$ is the atmospheric temperature profile,
- z is height,
- $\tau(z,z')$ is the transmittance profile between two heights.

The transmittance may be written,

$$\tau(z,\infty) = \exp\left\{-\int_{z}^{\infty} k_{\nu}(z') w(z') \sec\theta \, \mathrm{d}z'\right\}. \qquad (11)$$

where:

- k_{ν} is the absorption coefficient,
- $w(z)$ is the vertical profile of the absorber amount.

This leads directly to the ‘so-called’ split-window formulation.

Consider two wavelengths (e.g. SLSTR 11 and 12 µm channels and introduce appropriate subscripts):

$$I_{11} = [\epsilon_{11}B_{11}(T_s) + (1-\epsilon_{11})L_{11}]\tau_{11} + (1-\tau_{11})\overline{I}_{11}^{atmos}. \qquad (12)$$

$$I_{12} = [\epsilon_{12}B_{12}(T_s) + (1-\epsilon_{12})L_{12}]\tau_{12} + (1-\tau_{12})\overline{I}_{12}^{atmos}. \qquad (13)$$

Linearise around ν_{11}:

$$B(\nu,T) = B(\nu_{11},T) + \left(\frac{\partial B}{\partial \nu}\right)_{\nu_{11}} (\nu - \nu_{11}),$$

Manipulate:

$$B\ (T_s) = \left[\frac{1+\gamma}{\epsilon_{11} + \gamma\tau_{12}\Delta\epsilon}\right] I_{11} - \left[\frac{\gamma}{\epsilon_{12} + (1+\gamma)\tau_{11}\Delta\epsilon}\right] I'_{11} + \alpha, \qquad (14)$$

where:

- $\gamma = \dfrac{1-\tau_{11}}{\tau_{11}-\tau_{12}}$,
- $\Delta\epsilon = \epsilon_{11} - \epsilon_{12}$,
- $\alpha = -\dfrac{(1-\tau_{11})\tau_{12}(1-\epsilon_{12})I_{12} - (1-\tau_{12})\tau_{11}(1-\epsilon_{11})I_{11}}{\epsilon_{12}\tau_{12}(1-\tau_{11}) - \epsilon_{11}\tau_{11}(1-\tau_{12})}$.
- I'_{11} is the radiance at $\nu = \nu_{11}$ that yields a temperature equal to T_{12}. Thus,
- $I'_{11} = B_{11}[T_{12}]$.

Some special cases are worth considering:

(i) No spectral emissivity dependence:

$\Delta\epsilon = 0$

$$B_s = \frac{1+\gamma}{\epsilon} I_{11} - \frac{\gamma}{\varepsilon} I'_{11} + \alpha. \qquad (15)$$

(ii) Emissivity ≈ 1 (e.g. sea surface):

$$B_s = (1+\gamma)I_{11} - \gamma I'_{11} \qquad (16)$$

By linearising the Planck function about a mean atmospheric temperature, the algorithm can be formulated in terms of brightness temperatures.

$$B(\nu,T) = B(\nu,\overline{T}) + \left(\frac{\partial B}{\partial T}\right)_{\overline{T}}(T-\overline{T}). \qquad (17)$$

After some manipulation,

$$LST = a_0 + b_0 T_{11} + c_0 T_{12} \qquad (18)$$

- $a_0 = \alpha\left(\frac{\partial B}{\partial T}\right)_{\overline{T}}^{-1},$
- $b_0 = \frac{1+\gamma}{\epsilon_{11}}\left[\frac{1}{1+\gamma\tau_{12}\,\Delta\epsilon/\epsilon_{11}}\right],$
- $c_0 = \frac{\gamma}{\epsilon_{12}}\left[\frac{1}{1+(1+\gamma)\tau_{11}\,\Delta\epsilon/\epsilon_{12}}\right].$

This mathematical development shows that under the assumptions it is possible to relate the brightness temperatures in the 11 and 12 µm channels linearly to the land surface temperature.

Although ε_{11} and ε_{12} are non-unity for land surface emissivities, from a radiative transfer point-of-view, they are sufficiently close to one for the approach to be appropriate.

3.3 Global LST Algorithm: Regression relation

The basic form of the algorithm, in clear-sky, is (Prata, 1993, 1994a,b):

$$LST = a_0 + b_0 T_{11} + c_0 T_{12} \qquad (25)$$

- a_0, b_0, and c_0 are classes of coefficients that depend on land cover type, vegetation fraction, season, time of day (day or night), precipitable water, and satellite zenith view angle.
- T_{11} and T_{12} are the measurements in units of brightness temperatures (top-of-the atmosphere) in the SLSTR 11 µm and 12 µm channels; units are degrees Kelvin.

- θ is the satellite zenith view angle.

Recognising that the essential approximations lead to a linear form for the dependence of surface temperature on top-of-the-atmosphere brightness temperatures, we introduce some weak non-linearity by permitting the temperature difference to vary with a power n. Then we write the algorithm as,

$$LST = a_{f,i,pw} + b_{f,i}(T_{11} - T_{12})^n + (b_{f,i} + c_{f,i})T_{12}, \qquad (26)$$

- $n = \frac{1}{\cos(\theta/m)}$,
- m is a variable parameter controlling the dependence on view angle,
- f corresponds to fractional vegetation.
- i corresponds to vegetation type (biome),
- pw corresponds to precipitable water (in cm).

The algorithm has been developed with all temperatures in units of degrees Kelvin and the algorithm returns the land surface temperature in units of degrees Kelvin. It has been applied operationally in then AATSR processing albeit with limited auxiliary data fields, particularly in terms of biome and surface classification. The requirements for SLSTR in terms of auxiliary data are specified in the following sections.

The parameter n can be set equal to unity ($m \to \infty$) and the algorithm relaxes to the form derived earlier. The purpose of introducing the parameter here is to permit some tuning of the algorithm based on the analysis of the validation data.

For the case where $T_{11} - T_{12} \leq 0$, the parameter n must be set equal to unity. Such cases do occur in practice and have been identified previously for desert regions, particularly at night and in the morning shortly after sunrise. This constraint can also be achieved by specifying the parameter *m* for this case. In case of further algorithm updates, it would be wisest to incorporate a flag (yes/no) as input to determine whether this process should be used. The flag could be incorporated into the auxiliary file with *d* and *m* as an additional column to the two-dimensional arrays (see below).

In the regression relation proposed, the emissivity does not appear explicitly. Since the derivation is based on the mathematical formulation given above, in principle, it will be possible to determine the quantities ϵ_v from (9). However, there is no current intention at present to determine the emissivities as part of the LST product, nor as a separate product.

Of course the emissivity factors must be present implicitly and this becomes apparent when we write down formulae for the classes of regression coefficients in the LST algorithm.

The three classes of coefficients are:

- $a_{f,i,pw} = d[\sec\theta - 1]pw + f a_{v,i} + (1-f)a_{s,i}$,
- $b_{f,i} = f b_{v,i} + (1-f)b_{s,i}$,
- $c_{f,i} = f c_{v,i} + (1-f)c_{s,i}$,

$(0 \leq f \leq 1)$

Hence there is a set of coefficients which are applied to each pixel type depending on biome *i*, fractional vegetation cover *f* and *pw* but also depend on the values at 100% vegetation (*v*) and bare soil (*s*). The $a_{v,i}$, $a_{s,i}$, $b_{v,i}$, $b_{s,i}$ and $c_{v,i}$, $c_{s,i}$ are defined as the base retrieval coefficients in this ATBD. The parameter, *d*, is a variable controlling the water vapour dependence on viewing angle and is independent of the variable surface and atmosphere parameters.

Values for the coefficients will be determined using simulation data-sets in which BTs have been calculated for a range of LSTs, appropriate surface and atmosphere conditions (section 3.4). The parameters d and m will be empirically determined using the radiative transfer simulations for regions, most likely where some validation data were available. The parameters, *d* and *m*, can in principle depend on biome. Hence the *d* and *m* should be supplied in auxiliary files as 2-d arrays with maximum dimension set to the maximum number of biomes and values for both day and night for each biome. In addition, a column for utility flag may be needed; the overall file containing d, m and utility flags should be 2-dimensional.

Separate regression coefficients will be supplied for bare and vegetated surfaces for each of the landcover classes i. The coefficients will also be supplied separately for daytime and nighttime. There should be an option for interpolation of coefficients between months which should be implemented in the first instance. There will be no interpolation between day and night coefficients. Rather the coefficients should be switched according to a specifiable solar zenith angle threshold.

Note that the coefficient route for the algorithm allows and requires auxiliary data which can be updated at any stage without altering the essential processing algorithm. Indeed for *f*, and potentially *pw*, this is essential.

Atmospheric correction

The combination of the split-window equation above and the derivation of coefficients through radiative transfer calculations *intrinsically* provide an effective atmosphere correction in clear sky conditions. The water vapour dependence is explicit in the *pw* term across the swath but is otherwise expected to be adequately covered through the statistical regression fit of the base retrieval coefficients. The most important thing is that the BT simulation process described in Section 3.4, and specifically the radiative transfer model, is capable of computing accurately the atmospheric transmission. Since the atmosphere can also change considerably from month to month, it is recommended that the coefficients be implemented on a monthly basis.

The dependence of thermal infra-red channels on atmospheric water vapour is the most important clear sky atmospheric dependence. In principle, water vapour information from any appropriate satellite sensor would be useful. However, the height-dependent sensitivity functions (or weighting functions as they are known) should be matched for best corrections and in reality the input atmosphere information should have good vertical sensitivity. Hence thermal infra-red instruments tend to be the best source if they provide vertical resolution and also lower troposphere sensitivity appropriate to window channels as for SLSTR.

Summary of the LST algorithm:

- The LST algorithm is based on three classes of coefficients applied to the nadir-views of the 11 and 12 μm channels (split-window approach).
- For each biome, six base retrieval coefficients must be supplied (two per class of coefficient). The retrieval coefficients will be supplied for both day and night (making 12 retrieval coefficients in all). A complete set of base retrieval coefficients will be supplied per detector owing to the differing spectral responses of the detectors.
- Base coefficients should be provided on a monthly basis with the *option* (enabled in the first instance) for interpolation of LST derived from the two nearest sets of monthly base coefficients. The base retrieval coefficients will be supplied in one file with a time stamp common to other monthly data sets (centre of the month in the first instance). If annual coefficients are desired, then the monthly data sets will simply be repeated with the same coefficients. Annual average coefficients are a minimum.
- There are two constant parameters, *d* and *m*, whose values will be supplied as auxiliary parameters which are the same for both detector elements.
- For the algorithm, the sensor view angle (from nadir) at the surface is required. For day/night specification (e.g. switching of coefficients), error characterisation and verification, the solar zenith angle is required.

- If no interpolation is performed on a monthly basis, then all base coefficients will be changed at the monthly date boundary.

- There will be no interpolation between day and night coefficients.

- The solar zenith angle switching point should be specified in a file but should be set to 90° in the first instance.

- In principle (first approach), the base coefficients should be interpolated first and then the LST derived using the interpolated base coefficients and the interpolated auxiliary data. This should ensure that the correct auxiliary data are also written to the output file; auxiliary data per pixel in the output file must be consistent with final output LST per pixel. A second approach is to calculate two LSTs and then interpolate the LSTs, however in this case great care must be taken also with the auxiliary data to ensure consistency. Therefore the first approach is recommended.

- Auxiliary data must be supplied which will specify, for a given spatial location: biome, fractional vegetation and precipitable water. The number of biomes is currently those of Globcover, i.e. 22, but for flexibility in likely processor updates, should be assumed to be a maximum of 50. All files will be filled (no missing values).

- Auxiliary data and auxiliary files are not necessarily static and the processing must have the flexibility to incorporate updated information in the future. Therefore auxiliary information for each input parameter would be stored in separate file, but all information on the input parameter is stored together in one file. Tiled and packed forms of files are acceptable.

3.4 Generation of retrieval coefficients

3.4.1 Introduction

This section describes the process for generating the coefficients for the retrieval of LST form SLSTR.As described in the previous section, retrievals coefficients are the essential and fundamental component of the LST retrieval algorithm. The base retrieval coefficients must be generated by regression of simulated BTs against a suitable high specification line-by-line forward model with appropriate ranges of parameters for land surface temperature, surface spectral emissivity, atmospheric profiles of temperature and humidity, and also contaminant gases. These aspects except for water vapour are entirely implicit in the algorithm implemented in the operational processor but are made explicit in the coefficient generation.

The simulations require:

- An accurate forward model of the radiative transfer

- Specified ranges of land surface temperature for each biome
- A range of expected spectral emissivity and fractional vegetation for each biome
- A full range of atmospheric profiles, particularly atmospheric water vapour.
- Instrument parameters, most importantly accurate spectral response functions for the channels at 11 and 12 μm.

The simulated brightness temperatures (BTs) are then regressed against the known input land surface temperatures to determine the base retrieval coefficients. The computations are performed for each biome with states including fully vegetated and bare soil conditions.

3.4.2 Specification of the forward model

3.4.2.1 Introduction

The forward model required for the simulations is quite sophisticated and must accurately simulate atmospheric radiative transfer.

A forward model for LST retrievals can in fact serve a number of different purposes. It can be:

1. used directly in the retrievals in level 2 processing (on-line direct fitting),
2. applied to the off-line derivation of retrieval coefficients which are subsequently applied to derive the LST in level 2 processing.
3. employed in error analyses and sensitivity tests for the LST product.

The primary description of the forward model for this LST ATBD is, however, concerned with the second task since the current specification of the LST retrieval processor is based on a coefficient approach. Brief reference will also be made to the third task, where appropriate, although the error analysis itself is described in a separate document.

As long as the retrieval approaches remain consistent, the ATBD for the LST will be similar in its description of the forward model to the ATBD for Sea Surface Temperature (SST). This is because the channels employed for LST are a sub-set of those for SST. Forward views are not used for LST because of research required to understand the angular dependence of spectral emissivity and angular anisotropy of temperature. The current LST retrieval is therefore equivalent to the dual-channel, nadir retrieval from the 11 and 12 μm channels.

The differences in radiative transfer between the LST and SST analyses lie in the details of the surface variation and impacts on thermal emissivity of the heterogeneous land surface, the greater range of surface-atmosphere temperature differences, and the expected strong variability of lower tropospheric trace gases over land, particularly water vapour. In addition, aerosol variations over land may present additional difficulties in accurately representing the full range of land surface temperature radiance variations.

In order to clarify the requirements for this model, it is necessary to return to the description of radiative transfer in Section 3.1.2, and to then describe a suitable radiative transfer model and its functionality. This is followed by a specification of the input data required for the radiative transfer calculations. Finally, the specification of the forward model and processing concept are presented.

3.4.2.2 Radiative transfer to space

The signal detected by the thermal infra-red channels of a radiometer such as SLSTR is intrinsically determined by the total number of photons falling on a detector within the spectral passband (the spectral filter) of each channel of the instrument. As a result of calibration, through views of blackbodies and space, these raw photon counts are converted to calibrated radiance or more usually, in this case, to brightness temperature. For the purposes of this ATBD for LST, it is sufficient to consider the measurement, **y**, to be a brightness temperature (BT) although uncertainties arising from any preceding steps (counts to BT) need to be taken into account in any error analysis.

The upwelling thermal radiance at the top of the atmosphere is primarily determined by Planck function from the surface according to its temperature, modified by absorption and re-emission processes in the atmosphere. A full treatment shows that the radiance measured by the satellite, L^{sat}, can be written as three terms, and using L to denote radiance in brightness temperatures:

$$L^{\mathrm{sat}} = L^{\mathrm{ground}} + L^{\mathrm{atm}} + L^{\mathrm{atm\text{-}reflected}} \quad (27)$$

where L^{ground}, L^{atm} and $L^{\mathrm{atm\text{-}reflected}}$ are respectively the upwelling radiance emitted by the ground and absorbed in the atmosphere; the radiation emitted and absorbed in the atmosphere, and the down-welling radiance emitted by the atmosphere that is reflected by the ground towards the sensor. For a radiometer, the right-hand terms can be expressed more specifically as (see Susskind et al. (1984) and Dash et al. (2002)):

$$L_i^{ground} = \int_{\lambda_1}^{\lambda_2} f_i(\lambda)\epsilon(\lambda)B(\lambda, T_s)\tau(\lambda)d\lambda \quad (1.2)$$

$$L_i^{atm} = \int_{\lambda_1}^{\lambda_2}\int_{p=0}^{p_s} f_i(\lambda)B(\lambda, T_p)\frac{d\tau}{dp}d\lambda dp \quad (1.3)$$

$$L_i^{atm\_reflected} = \int_{\lambda_1}^{\lambda_2}\int_{\theta=0}^{\frac{\pi}{2}}\int_{\phi=0}^{2\pi} f_i(\lambda)(1-\epsilon(\lambda))L^{\downarrow}(\lambda,\theta,\phi)\tau(\lambda)\sin 2\theta d\lambda d\theta d\phi \quad (1.4)$$

where i is the radiometer channel, $f_i(\lambda)$ is the normalised channel response function (or spectral filter function), θ is the zenith angle, φ is the azimuth angle, λ is wavelength, p is pressure in the atmosphere, ps is the pressure at the Earth's surface, Tp is the mean temperature of air at pressure level p, τ(λ) is the spectral atmospheric transmissivity, and ε(λ) is the surface spectral emissivity.

Although these equations look complicated, in a clear sky atmosphere the important geophysical parameters in addition to surface temperature are the surface emissivity, the atmosphere transmission at the wavelengths of interest and the atmosphere temperature. In effect it the knowledge of these parameters and their separation from the land surface temperature response which allows the surface temperature to be retrieved from radiometer measurements and at the same time also limits the accuracy with which this can be done.

Fortunately, the radiative transfer processes at thermal wavelengths are on the whole well understood at least in the clear sky case. The surface emissivity depends on the nature of the surface and can be more complicated for canopy structures. However, the largest variations in thermal emissivity are for bare soil and it is the classification of land cover and the numerical emissivity values that are more uncertain rather than the theoretical use of emissivity values. There is an angular dependence to emissivity but it is less clear that it matters compared to angular variations of temperature. Again, the theory is not a problem but knowledge of the actual emissivity may cause some error in "nadir" views across the swath which results in larger angle views.

The transmission through the atmosphere essentially depends on the concentrations of the gases in the atmosphere and its spectroscopy. The relevant spectral parameters are usually tabulated in world standard databases, except for the water vapour continuum which must be calculated using consistent information from the database. The continuum effect on radiative transfer is well understood and is conceptually simple to implement in radiative terms although its calculation is considerably more complex. There are radiative aspects such as non-local thermodynamic equilibrium, line mixing and non-standard lineshapes which matter generally in thermal radiative transfer and are still the subject of research, but effects are usually small in the thermal infra-red channels which are considered here (but see below).

For clear sky radiative transfer, we do not need to consider scattering but this can change where there are sufficient particles with diameters of the order of the wavelengths of interest (typically 10 microns). In non-volcanic cases, the particles that induce scattering are desert dust aerosols and thin cirrus; thicker cirrus and directly injected volcanic ash generally result in scenes with poor atmospheric transmission which are likely to be flagged as cloud. Again, for the purposes of calculations, the first order radiative transfer is well understood although scattering codes must be included in the forward model. The uncertainties tend to lie primarily in uncertainties in the refractive index and in non-spherical particle shapes.

3.4.2.3 Radiative Transfer Model Functionality

The radiative transfer equations described in the previous section must be implemented in a numerical code which accurately implements the equations described above. For both the generation of retrieval coefficients and for error analyses, the ideal code should be a so-called line-by-line model. Line-by-line models calculate the monochromatic transmission of the atmosphere at high spectral resolution, typically finer than the widths of the individual spectral lines which matter. Although the line shapes tend to be Lorentzian in the troposphere, the best accuracy is achieved by using a Voigt routine, which provides a convolution of the Lorentz and Doppler line shapes.

Line-by-line models are standard and a number exist, but the most tested and utilised within European Space Agency studies is the Oxford Reference Forward Model (or RFM). It can perform radiance or transmittance calculations over a wide range of wavelengths at spectral resolutions as high as 0.0004 cm^{-1}. In fact, the computation is performed at a fine mesh resolution of typically $5x10^{-4}$ at wavenumbers of less than 1-2 cm^{-1} from the line centre; the line wings of overlapping lines are taken into account by a wide mesh calculation at wavenumbers up to 25 cm^{-1} from the line centre. The user then specifies then output spectral resolution.

The importance of the high spectral resolution of the RFM has been tested for the thermal channels considered here (Noyes, 2006). For example, comparison of RFM simulations with those from the MODerate spectral resolution atmospheric TRANSmittance algorithm and computer model (MODTRAN - see Kniezys et al. (1996) for details), showed that the spectral resolution of a radiative transfer model is critical in terms of the accuracy of simulated AATSR TOA BTs. For example, by using a spectral resolution of 1 cm^{-1} which is the highest mode available using MODTRAN, the error on the simulated AATSR TOA BTs was estimated to be of the order of tenths of a K (e.g. >0.4 K for the nadir 11 μm channel for a tropical climatology). Spectral resolutions of 0.01 cm^{-1} were found to be much more reasonable as a compromise between accuracy and computational time (errors of 0.01 to 0.02 K in BT terms).

In order to correctly model the atmospheric transmittance, the radiative transfer model needs to have as inputs vertical profiles of atmospheric temperature, pressure and trace gas mixing ratios. The radiative transfer calculation also requires accurate spectral parameters which characterise the absorptions of the atmospheric trace gases which affect the radiance in the spectral passbands of the channels of interest. These spectral parameters include line parameters for the main trace gases, absorption cross-sections for heavier molecules, and continuum parameters for the main atmospheric continua (in the SLSTR passbands for LST, these are water vapour continua and the nitrogen continuum).

It also requires spectral emissivities for the wavelengths and ranges of surfaces of interest. In principle, a complex model of emissivities is not required since the forward model is not being used to directly fit to the radiances. For retrieval coefficient generations, the model only needs to have some mechanism for user input of spectral emissivities for a given land type.

3.4.2.4 Requirements for the forward model.

The requirements for the forward model for land surface temperature, assuming the need to model only the 11 and 12 μm channels, are the following:

- should be capable of calculating output radiances at spectral resolutions of better than 0.01 cm^{-1}.
- Should utilise fine grid calculations at high spectral resolution
- Should use well-characterised and efficient Voigt calculation routines
- Should have the capability to use up-to-date HITRAN spectral databases
- Should utilise continuum calculations as appropriate, including water vapour continuum calculated consistently with the appropriate HITRAN database.
- Should provide a flexible interface to use vertical profiles of atmospheric temperature, pressure and trace gas mixing ratios
- Should provide for a separate input for surface temperature and emissivity.

3.4.3 Auxiliary parameters and input data sets for coefficient generation

The preceding description of the radiative transfer model makes it clear that the specification of auxiliary parameters and input data sets is a key aspect for the generation of retrieval coefficients. For auxiliary parameters, the primary requirement is for very good information on instrument parameters and spectroscopic databases. For input data sets, the requirement for retrieval coefficients is that the surface data and atmosphere profiles cover a realistic and well sampled range of conditions. Implicit in this but worth highlighting for LST, is that the surface-air temperature difference is particularly important. Fractional vegetation is also an important consideration in so far as including varying emissivities for a given biome are concerned.

For the error analysis and sensitivity tests, the data sets above are less crucial but are nonetheless of great utility in performing retrieval algorithm tests. The main error analysis is likely to be based on a set of standard conditions with estimates of uncertainty or variability in the input data sets. For these, the atmosphere data sets will be drawn from ECMWF analyses and the Reference atmosphere for MIPAS standard atmospheres (RAMStan; Remedios et al, 2007).

The error analysis results themselves are described in a separate document in preparation (S3-L2-SLSTR-ULNILU-AEAD-0001_L2LST_v1.0_100831.doc)

3.4.3.1 Instrument parameters

Instrument parameters required for the forward model are clear. The primary instrument information for the forward model is the spectral passband of the two thermal channels, expected to be centred at 11 and 12 µm. It is particularly crucial to know the low wings of these spectral responses to better than 10^{-4} of the peak response.

It is also important to have a knowledge of the field-of-view of the instrument in each detector channel although the requirement for this is less stringent here as it is required for error analyses but not for the operational processor itself. It will become more important if the retrieval process incorporates knowledge of vegetation at small scales. Knowledge of the field-of-view is important for users and hence it should be supplied in any case as a stand-alone function/tool.

3.4.3.2 Surface properties

The primary surface parameter required as input to forward model calculations is the LST itself. Globally, there are several sources of LST including ECMWF data, AATSR LST data and MODIS data. Since temperature varies considerably, the most appropriate sets of observed temperatures are those recorded close to the likely overpass times of SLSTR. Currently no such data sets exist and appropriate data sets will have to be developed to really constrain the ranges required for the retrieval coefficients.

Thermal emissivity is at the root of the radiative transfer calculation for LST. Typically, the emissivity varies between 0.9 and 1.0 but for bare soils and deserts, the emissivity may reach lower values. For the current retrieval scheme, these emissivity variations are important and need to be captured by the range of radiative transfer calculations for retrieval coefficients. In order to focus the problem, and also to be consistent with the retrieval processor, the range of emissivities can be restricted by generating retrieval coefficients as a function of biome and fractional vegetation. Therefore, the problem is reduced to utilisation of typical emissivities for each biome under consideration.

At the current time, the biome classification will be of the order of 22 but can be expanded with 50 recommended as a suitable limit. For each biome type, typical emissivities must be assigned from spectral databases of surface properties such as that in the ASTER spectral data base or compilations of derived surface emissivity.

The range and distribution of fractional vegetation indices can be deduced from satellite data sets of Normalised difference vegetation indices (NDVI), for example. An appropriate data set is described in section 6.3.

The biome classification and range of fractional vegetation will also influence the range of surface-air temperature differences (see next section).

3.4.3.3 Atmosphere variables

The key atmosphere variables are pressure and temperature, and trace gases, of which the most important is water vapour. It is convenient to draw the pressure, temperature and water vapour together (henceforth referred to as the met variables) as they can be taken from common sources e.g. radiosondes, meteorological analyses and satellite data sets. As noted for SST, the range and consistency of sampling of the atmosphere can be a significant consideration. In the case of LST, this range and consistency has to be established per biome.

For the atmosphere system, because of meteorological services, the analyses such as ECMWF probably provide a reasonable estimate of these meteorological parameters and will be used for the current LST system. These should provide the best method of obtaining a statistically complete set of profiles for the LST retrieval coefficient process. Data sets exist which randomly sample the range of water vapour profiles required and can be selected at the time of generation of the retrieval coefficients (but must be documented with the delivery of the retrieval coefficients).

The other main atmospheric trace gases in the thermal channels are carbon dioxide, nitric acid, chlorofluorocarbons (CFCs) and ozone. All of these trace gases have relatively small effects on the final temperature accuracies in terms of the expected performance of the LST retrievals. The seasonal and latitudinal variations will be taken from the Reference atmospheres for MIPAS initial guess climatology (RAMIgclim; Remedios et al, 2007). The RAMIgclim contains reasonable vertical profile climatologies for these gases with realistic vertical profiles for a range of latitudes; data sets are based on satellite observations, in situ data sets and model observations.

Overall, it is recommended that the forward model calculations should include all trace gases contributing brightness temperatures of greater than 0.1 K to the overall signal levels in each channel.

3.5 Operational processor auxiliary data and cloud/aerosol flagging

Auxiliary data must be supplied which will specify, for a given spatial location: biome, fractional vegetation and precipitable water. All files will be filled (no missing values).

Auxiliary data and auxiliary files are not necessarily static and the processing must have the flexibility to incorporate updated information in the future. Therefore auxiliary information for each input parameter would be stored in separate file, but all information on the input parameter is stored together in one file. Tiled and packed forms of files are acceptable.

3.5.1 Operational processor auxiliary data: Biome

The biome map for SLSTR is envisaged to be an auxiliary data file containing a high spatial resolution map of vegetation type. Examination of BT images for the AATSR instrument clearly shows a strong relationship to features in land cover maps at high spatial resolutions. Typically high spatial resolution maps are available at up to (1/360)° resolution in global maps; corresponds to approximately 300 m at the equator.

For SLSTR processing, after testing it is recommended that (1/120)° resolution be employed currently for the auxiliary data file but with the capacity to expand this as required, e.g. to double the resolution; corresponds to approximately 1 km at the equator.

The number of biomes is currently those of Globcover, i.e. 22, but for flexibility in likely processor updates, should be assumed to be a maximum of 50.

The biome map is expected to be static with a slow update period e.g. annually or less often. Annual would be the most desirable but will depend on availability of Globcover-type data

The number of biomes in the biome map should be matched by the number of types of coefficients which are to be applied in the LST processor. The values in the file will be integers.

The appropriate biome for analysis for a given pixel should be selected as the biome which is the nearest neighbour. Studies indicate that the number of mixed pixels without a majority biome is only a few %.

For the application of the biome map, it will be essential to have accurate cross-referencing of the sensor and the co-ordinates of the biome map.

All information on the input parameter is stored together in one file. Tiled and packed forms of files are acceptable.

3.5.2 Operational processor auxiliary data: Vegetation fraction

For some of the biomes it is recognised that the temporal variation of the vegetation fraction will be important.

The exact source of fractional vegetation is the study of ongoing research, but two methods of estimating fractional vegetation from visible satellite observations are being evaluated:

1. A Normalised Difference Vegetation Index (NDVI)-based method (see for example Carlson & Ripley, 1997), and

2. The Scaled Difference Vegetation Index (SDVI)-method.

Both methods make use of the differing visible properties of soil and vegetation at red and near infrared (NIR) wavelengths. As SLSTR has channels at 660 nm and 870 nm, both these methods could potentially be used for operational estimation of fractional vegetation. However, existing research and application is best suited to implementation of an NDVI-based climatology but at high temporal frequency.

In this ATBD, the vegetation climatology is specified as a fractional vegetation (FV) index with maximum value of 1 (corresponding to 100% vegetation) and 0 (corresponding to 0% vegetation). Updates should be at least twice per month and the current temporal spacing is every 10 days.

In the future, a two step process should in any case be used but this has not been tested operationally. Allowance for this in the definition of the processor would be helpful in allowing this potential update path to the processor to be maintained. Briefly,

- For day time retrievals, an operational SLSTR fractional vegetation will be used using coincident views of the visible channels to generate NDVI.
- For night time retrievals, a 'live' or dynamic climatology will be used with values being held in an auxiliary data file.
- From an operational point-of-view, the climatological fractional vegetation map should be provided at the same spatial resolution as the auxiliary biome data file. The values in the file will be non-integer and fall between 0 and 1. As an additional check, the FV field should be limited to an upper value of one in the code.

The night time 'live' climatology will be constructed from existing fractional vegetation climatologies, and will be updated on a weekly or monthly basis using day time NDVI measurements from the previous month.

Hence the auxiliary data file would need to be updatable on a weekly basis with a process running in parallel which takes daytime L2 NDVI values and combines with the previously available climatology to produce a new climatology for use with nighttime data.

Although it is possible to use other sensors for the generation of a dynamic FV nighttime climatology, particularly if coverage is greater or complementary to that of SLSTR, the specification and provision of such data become more complicated within an operational system. Therefore it is expected that SLSTR visible observations would be the foundation for a new FV system since co-located daytime data would be available. A synergy system utilising OLCI data would clearly be worthy of investigation also.

Therefore to summarise, the current baseline recommendation is for climatology with relatively high time resolution, consisting typically of files spaced at, for example, ten day intervals. Spatial resolution will be close to that of the biome data. The data will be time stamped. To allow for off-line re-processing or for NRT generation of the auxiliary data, the file specifications should include year date. The FV fields will be close to but not necessarily exactly the same as the spatial resolution of the biome map. A nearest neighbour approach can be employed to find the relevant FV fields for the LST. FV fields will be interpolated to the current time of the measurement. The files will be filled (no missing values).

Auxiliary data and auxiliary files are not necessarily static and the processing must have the flexibility to incorporate updated information in the future. Therefore auxiliary information for fractional vegetation would be stored together in one file but so that the fractional vegetation information could be updated on the order of once per week basis in the future. Tiled and packed forms of files are acceptable.

A potential update path has been identified for the fractional vegetation processing involving use of co-located daytime observations of NDVI

3.5.3 Operational processor auxiliary data: Water Vapour

It should be clear from section 3.3 that a large part of the atmospheric correction problem is provided by the split window formulation and by the use of retrieval coefficients derived by regressions against simulated data. The use of water vapour auxiliary data is to account better for swath angle dependences of the atmospheric correction factor which is a smaller effect.

Water vapour data should be obtainable from two sources:

1) Total column water vapour from meteorological analyses. This is already planned to be available in the processing.

2) A monthly climatology file of water vapour for each year. The water vapour climatology is likely to be at a resolution of no greater than 0.l°

In the processing there should be an option to switch between these two sources of water vapour (the option to use meteorological data should be enabled in the first instance). The monthly climatology file shall be the baseline for the algorithm. Water vapour fields will be interpolated to the current time of the measurement. If meteorological fields are not available, then the climatology will be used. For water vapour, interpolation between analysis fields and forecast fields is acceptable. In near real-time operations, if the forecast field is available then climatology will have to be used. In off-line processing, analysis fields should be used and not forecast fields.

The water vapour climatology will provided as filled fields with twelve monthly fields in one annual file (which can be updated from year to year if required e.g. for re-processing). The water vapour climatology should be interpolated to the required point in time, latitude and longitude.

3.5.4 Clouds and aerosols

Cloud particles and aerosol particles may both affect thermal infra-red properties but the magnitudes of effects depend very much on optical thickness factors including refractive index and size of particles. All cloud particles, i.e. water droplets and ice particles, will have some infra-red effect but where clouds become thick then the signal at the satellite tends to drop dramatically, except for low warm clouds for example, due to cloud top temperatures being lower than surface temperatures. In contrast, it is typically only aerosol particles that affect thermal infra-red radiances significantly: desert dust, volcanic ash and sulphate aerosol particles. Where such aerosol events significantly reduce atmospheric transmission, then such data are likely to be removed by cloud clearing algorithms.

The major effect on LST is the effect of clouds and therefore the efficacy of cloud clearing algorithms over land is the major issue. A detailed discussion, and solution, is out of scope of this ATBD. However, there are two main points that can usefully be made. Firstly, cloud flagging is undertaken at level 1b by a conventional cloud detection algorithm that has been implemented operationally for ATSR-like sensors such as SLSTR. LST values should be retrieved for all level 1b pixels whether cloud flagged or not. The output file should contain the cloud flags so that the user can decide how to treat the data. Secondly, advanced cloud schemes, such as Bayesian approaches, are being actively pursued at the moment. The gain from advancing cloud detection schemes means that it is prudent to allow for such an update path.

Hence, the baseline for the operational processor is to use the level 1b cloud flags to flag the output level 2 files but to analyse the data for all pixels. The operational processor should maintain provision for a pre-processing cloud module in the level 2 processor for LST which would identify clouds more rigorously and specifically for the thermal infra-red channels. This algorithm might either be a more specific form of the level 1 cloud clearing or else a Bayesian-type approach as for LST.

For aerosol flagging or for aerosol correction, good information is required on aerosol which may have significant effects in the thermal infra-red for which there are two main solutions: 1) co-located aerosol observations from visible sensors with good aerosol type information or observations combining visible and thermal aerosol channels; 2) operational analyses/models with good aerosol optical depth data and aerosol-type information. Algorithms for thermal infra-red retrievals of aerosol parameters globally are not as well developed as visible wavelengths retrievals. There are also considerable uncertainties in visible aerosol products for desert dust, for example. Another approach is to develop an infra-red aerosol index or use the difference between 11 and 12 μm BTs. Over land, these methods have not been extensively tested and are not ready for operational implementation.

Hence at the current time, the best approach that can be adopted for aerosol is that of the cloud case which is to analyse all level 1b pixels whether affected by cloud or not. An aerosol flag should be included in the level 2 output file. If possible, this should be filled in future by an ECMWF-type aerosol flagging system or data from a synergy product for aerosol.

3.6 Practical considerations

3.6.1 Numerical computation considerations

The algorithm requires brightness temperatures, viewing angle information and location data on a pixel-by- pixel basis. The most frequent numerical operation will be multiplication. The mostly costly operation in terms of computing time will be the identification of the biome type, fractional vegetation and precipitable water. Interpolation of the regression coefficients will also be required.

3.6.2 Input data fields

The generation of the LST product will require the following data to be available:

- Nadir 11 μm brightness temperature,

- Nadir 12 µm brightness temperature,
- Geographic latitude of pixel,
- Geographic longitude of pixel,
- Elevation of pixel
- Nadir zenith view angle at pixel,
- Nadir azimuth view angle at pixel,
- Cloud flag,
- Aerosol (thermal infra-red) flag if possible
- Land/sea flag,
- Mask value,
- Time at pixel,
- Date,
- Ancillary data consisting of vegetation type (biome), vegetation fraction and precipitable water.

3.6.3 Output data fields

The generation of the LST product will require the following output fields:

- LST
- Random LST error
- Total LST error
- Orphaned/cosmetic fill information
- Time
- Latitude
- Longitude
- Elevation of pixel
- Satellite viewing: elevation angle

- Satellite viewing: azimuth angle
- Solar zenith angle (zero for nighttime)
- Solar azimuth angle (zero for nighttime)
- Confidence flag
- Biome
- Fractional Vegetation
- Precipitable water total column
- NDVI
- Cloud flag
- Aerosol flag
- Validation flag

4 VALIDATION

The validation of the SLSTR algorithm, which is identical in form to the operational AATSR algorithm, has been performed utilising brightness temperatures simulated with the AATSR filter functions. The validation itself is performed by comparing LST retrieved from simulated brightness temperatures with input surface temperatures for each biome and day/night combination. The mean differences between input surface temperature and retrieved LST are less than or equal to approximately 0.3 K with standard deviations being less than or equal to approximately 1.0 K depending on biome and day/night selection.

Across track performance of the algorithm is validated by comparing LST retrievals with respect to changes in precipitable water: retrievals applying nadir coefficients and simulated edge-of-swath brightness temperatures are compared with retrievals applying edge-of-swath coefficients and simulated edge-of-swath brightness temperatures for iterations of the *d* and *m* parameters when applied to the retrievals with nadir coefficients. The iteration with the lowest RMSE is set in the auxiliary data file. Mean differences are less than or equal to 0.032 K with standard deviations less than or equal to 0.081 K depending on biome.

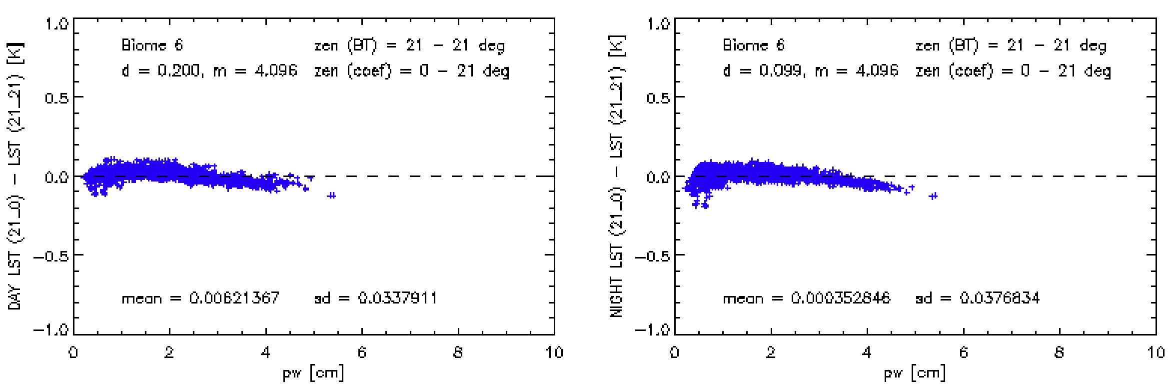

Figure 1: *Simulated errors for SLSTR within the AATSR regime of spectral response function sensitivities and range of "nadir" view angles. Scatterplots with respect to changes in precipitable water of simulated LST difference (nadir coefficients and simulated edge-of-swath BTs minus edge-of-swath coefficients and simulated edge-of-swath BTs) for biome class 6.*

These results are applicable to SLSTR assuming that the SLSTR spectral response functions are not significantly different to those of AATSR, particularly in terms of sensitivities of top of atmosphere BTs to atmospheric temperatures and water vapour. However, they are not applicable outside the range of "nadir" swath angles observed by AATSR, i.e. 21.6° (SLSTR "nadir" views extend to greater than 50°).

5 UNCERTAINTY BUDGET

Both random and pseudo-systematic errors contribute to the total error budget for LST but the latter dominate by far. Pseudo-systematic errors are errors which are not constant biases but have structure in space (and time). Typical errors include both uncertainties due to retrieval uncertainties (e.g., the ability of the regression coefficients to fit the data) and uncertainties in the values of auxiliary data. These structural differences can vary as a function of biome, latitude (a proxy for temperature), view angle and solar zenith angle.

The approach taken here is to calculate the random error from radiometric noise (calculated similarly to radiometric noise). The total error is then a combination of this random error with pseudo-systematic errors.

In principle, the pseudo-systematic errors are dependent on biome, fractional vegetation, latitude (temperature; see below), water vapour, topography (digital elevation difference from sea level), view angle and solar zenith angle. However, at the current time, no existing LST algorithm explicitly accounts for these errors and for their inter-dependence. Therefore, instead the approach is to concentrate on the three main sources of quantifiable error: biome, latitude (temperature; see below), and total column water vapour. The error fields will be loaded from an auxiliary file. Each error field will be a 1-d array with an integer giving the actual dimension of the each array (start, increment and number of values also specified in the file). The maximum dimensions expected for each array are 50 (maximum number of biomes), 18 (latitudes), and 20 (water vapour).

The pseudo-systematic errors can be calculated through simulations of the sensitivity of the LST to variations in the parameters using the forward modelling approach described for generation of retrieval coefficients. In addition, the standard deviations of the results with respect to those expected from the fits of the regression coefficients also provide a measure of the errors.

The random error, δe_{rand}, can be calaculated analytically. Differentiating first with respect to T_{11} (equation 28) and then with respect to T_{12} (equation 29), one can express the total random error in the following manner:

$$db_{f,i} = n \times b_{f,i}(T_{11} - T_{12})^{n-1} \qquad (28)$$

$$dc_{f,i} = -n \times b_{f,i}(T_{11} - T_{12})^{n-1} + b_{f,i} + c_{f,i} \ , \qquad (29)$$

$$\delta e_{rand} = \left((db_{f,i} \times \delta T_{11})^2 + (dc_{f,i} \times \delta T_{12})^2 \right)^{0.5} \qquad (30)$$

The δT_{11} *and* δT_{12} are the radiometric noise in each relevant channel.

The total error is the root-sum-square of this error squared with the total pseudo-systematic error squared. The random and total error should be written to the level 2 file for each pixel.

In the future, the latitude dependence is likely to be replaced by a temperature–related term which is more physical. However this has not been defined as yet. Error dependences on topography (digital elevation difference from sea level), view angle and solar zenith angle will be investigated. In addition, a cloud error parameterisation would be useful but cloud information is not yet available in suitable form.

6 ASSUMPTIONS AND LIMITATIONS

The following assumptions have been made:

1. The nadir SLSTR brightness temperatures have been flagged as cloud-free.
2. The nadir SLSTR brightness temperatures have been flagged as valid data.
3. Level 1b pixels are provided independently for each detector.
4. Time, date, zenith view angle and azimuth view angle are available for each pixel.
5. Geographic location of each pixel is available.
6. Ancillary data are available (precipitable water, vegetation type and fractional vegetation cover).

Quality control of the data product can be done at only three of the thirteen biomes although this could expand with time. Guarantees of accuracy and error estimates can only be provided for these biomes. The quality of the thermal product is critically dependent upon the quality of the brightness temperatures (e.g., calibration and noise characteristics), and dependent upon the ability to screen the data for clouds.

A validation flag should be set through an external file as a value per biome. This should have the values:

2: fully validated,

1: some verification information available

0: not validated.

7 FUTURE EVOLUTION

- In *future evolution* of the algorithms, it is expected that the following algorithms might be implemented and so design of the processor should include now the functionality which would enable these to be implemented later.

  - *Dynamic fractional vegetation*. In this approach, a two step process will be used for fractional vegetation: (a) for day time retrievals, an operational SLSTR fractional vegetation will be used using coincident views of the visible channels; (b) for night time retrievals, a 'live' or dynamic climatology will be used with values being held in an auxiliary data file; (c) The climatological fractional vegetation map should be provided at the same spatial resolution as the auxiliary biome data file. The values in the file will be non-integer and fall between 0 and 1; (d) the auxiliary data file will need to be updatable on a weekly basis with a process running in parallel which takes daytime L2 NDVI values and combines with the previously available climatology to produce a new climatology for use with nighttime data.

    To provide useful dynamic auxiliary information to the user and to allow for the dynamic fractional vegetation upgrade in the future, NDVI should be calculated and provided in the output file.

  - *Use of aerosol information to flag LST data*.

    An aerosol flag should be implemented in the quality flags of the level 2 LST output files.

    Provision should be made in the LST processor to make use of external aerosol information (e.g. from other SLSTR information, synergy information, or from ECMWF model aerosol fields if available). These data would need to be processed in an aerosol flag module, filtering the data to that required for thermal infra-red LST retrievals, i.e., dust particles. Alternatively, a direct aerosol (dust) flagging method using SLSTR brightness temperatures could be implemented in the LST processor.

  - *Pre-processor for cloud*. The operational processor should maintain provision for a pre-processing cloud module in the level 2 processor for LST which would identify clouds more rigorously and specifically for the thermal infra-red channels. This algorithm might either be a more specific form of the level 1 cloud clearing or else a Bayesian-type approach as for LST.

  - *Improved specification of errors*: The ideal is to provide two error fields, both random and systematic (as outlined above); this should replace total error if this is the only error specified. In the future, the latitude dependence is likely to be replaced by a temperature–related term which is more physical. However this has not been defined as yet. Error dependences on topography (digital elevation difference from sea level), view angle and solar zenith angle will be investigated. In addition, a cloud error parameterisation would be useful but cloud information is not yet available in suitable form.

- The cloud flag in the level 1b data will be used initially for cloud removal for the LST processor. This will be the baseline processor. However, the operational processor should include provision for a pre-processing cloud module in the level 2 processor for LST which would identify clouds more rigorously and specifically for the thermal infra-red channels.

REFERENCES

Carlson, T. N. and Ripley, D. A., 1997, On the relation between NDVI, fractional vegetation cover, and leaf area index, Remote Sensing of Environment, 62, 241-252.

Dorman, J. L., and P. J. Sellers, 1989, A global climatology of albedo, roughness length, and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB), J. Appl. Meteorol., 28, 833–855.

Kniezys, F. X., Abreu, L. W., Anderson, G. P., Chetwynd, J. H., Shettle, E. P., Berk, A., Bernstein, L. S., Robertson, D. C., Rothman, P. A. L. S., Selby, J. E. A., Gallery, W. O., & Clough, S. A, 1996, The MODTRAN 2/3 report and LOWTRAN 7 model, Technical Report 4, Phillips Lab., Geophys. Directorate, PL/GPOS, Hanscom AFB, MA 01731, Contract F19628-91-C-0132.

McMillin, L. M. and D. S. Crosby, 1984, Theory and validation of the multiple window sea surface temperature technique, J. Geophys. Res., 89, 3655–3661.

Noyes, E.J., 2006, An investigation into the accuracy of surface temperature retrievals from the AATSR, PhD. Thesis, University of Leicester.

Prata, A. J., 1993, Land surface temperatures derived from the AVHRR and the ATSR, 1, Theory, J. Geophys. Res., 98(D9), 16,689–16,702.

Prata, A. J., 1994a, Land surface temperatures derived from the AVHRR and the ATSR, 2, Experimental results and validation of AVHRR algorithms, J. Geophys. Res., 99(D6), 13,025–13,058.

Prata, A. J., 1994b, Land surface temperature determination from satellites, *Advances in space research*, 14, 15-26.

Prata, A. J., and R. P. Cechet 1999, An assessment of the accuracy of land surface temperatures from the Japanese GMS-5 VISSR, Rem. Sensing Environ., 67:1–14. Price, J. C., 1984, Land surface temperature measurements from the split-window channels of the NOAA- 7/AVHRR, J. Geophys. Res., 89, 7231–7237.

Prata, A., 2002, 'Land surface temperature measurement from space: AATSR algorithm theoretical basis document', *Contract Report to ESA, CSIRO Atmospheric Research, Aspendale, Victoria, Australia.*

Price, J. C., 1984, Land surface temperature measurements from the split-window channels of the NOAA- 7/AVHRR, *J. Geophys. Res.*, **89**, 7231–7237.

Remedios, J.J., R. J. Leigh, A. M. Waterfall, D. P. Moore, H. Sembhi, I. Parkes, J. Greenhough, M. P. Chipperfield and D. Hauglustaine, 2007, MIPAS Reference Atmospheres and comparisons to v4.61/v4.62 MIPAS level 2 geophysical data sets, Atmospheric Chemistry and Physics Discussions.

Salsbury, J. W., and D. M. D'Aria, 1992, Emissivity of terrestrial materials in the 8 – 14 µm atmospheric window, Rem. Sensing Environ., 42, 83–106.

Snyder, W. C., Wan, Z., Zhang, Y., and Y.-Z. Feng, 1998, Classification-based emissivity for land surface temperature measurement from space, Int. J. Remote Sensing, 19(4), 2753–2774.

Sutherland, R. A., 1979, Broadband and spectral emissivities of natural sands and vegetation (2–20 µm), J. Atmos. Oceanic Technol., 3, 199–202.

8 Appendix I: Typical number and size of LUTs

8.1 LST Coefficients

An example set of LST coefficients is given below based on AASTR LST coefficients

As shown in equation (26), the LST retrieval is based on the brightness temperatures T11 and T12, view angle and on the coefficients a,b,c dependent on fractional vegetation index f, precipitation water pw (a only), and the vegetation types i:

$$LST = a_{f,i,pw} + b_{f,i}(T_{11} - T_{12})^{p(\theta)} + (b_{f,i} + c_{f,i})T_{12}$$

The coefficients are split in vegetation (index v) and soil (index s) part for each landclass. The table below show all coefficients of the Dorman and Sellers (D&S) biomes for day and night condition (only for permanent water, biome 14) :

D&S biome no	night	a_v	a_s	b_v	b_s	c_v	c_s
1	0	0.6907	6.0951	3.8129	4.5637	-2.8455	-3.3617
1	1	0.6907	6.0951	3.8129	4.5637	-2.8456	-3.3617
2	0	-0.5393	4.6301	3.6472	4.3652	-2.7218	-3.2155
2	1	-0.5393	4.6301	3.6472	4.3652	-2.7218	-3.2155
3	0	-0.6885	4.8786	3.6472	4.3652	-2.7218	-3.2155
3	1	-0.6885	4.8786	3.6472	4.3652	-2.7218	-3.2155
4	0	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909
4	1	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909
5	0	0.7804	1.491	3.2721	3.8117	-2.3374	-2.7233
5	1	0.7804	1.491	3.2721	3.8117	-2.3374	-2.7233
6	0	0.9089	0.0348	3.3511	3.9038	-2.389	-2.7891
6	1	0.9089	0.0348	3.3511	3.9038	-2.389	-2.7891
7	0	0.7994	0.7994	3.5088	3.5088	-2.5065	-2.5065
7	1	0.7994	0.7994	3.5088	3.5088	-2.5065	-2.5065
8	0	1.5662	0.7833	3.1384	3.656	-2.2419	-2.6121
8	1	1.5662	0.7833	3.1384	3.656	-2.2419	-2.6121
9	0	0.8965	0.8965	3.4867	3.4867	-2.4908	-2.4908
9	1	0.8965	0.8965	3.4867	3.4867	-2.4908	-2.4908
10	0	1.0817	1.0817	3.3039	3.3039	-2.2955	-2.2955
10	1	1.0817	1.0817	3.3039	3.3039	-2.2955	-2.2955
11	0	0.7075	0.7041	3.7832	3.7832	-2.7868	-2.7868
11	1	0.7075	0.7041	3.7832	3.7832	-2.7868	-2.7868
12	0	0.881	0.881	3.4106	3.4106	-2.4133	-2.4133
12	1	0.881	0.881	3.4106	3.4106	-2.4133	-2.4133
13	0	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909
13	1	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909

14	0	-0.0005	-0.0005	2.4225	2.4225	-1.4344	-1.4344
14	1	-0.3658	-0.3658	2.3823	2.3823	-1.3556	-1.3556

Based on these coefficients of the D&S biomes, coefficients for the Globcover (GC) biomes have been cestimated. For every GC landclass, the global distribution of D&S landclasses has been investigated. Several GC landclasses are allocated quite unambiguously to a D&S landclass, e.g. GC40 to D&S 1. In this case the coefficients of the D&S biome are adopted for the according GC landclass. In cases of broad distribution of D&S landclasses to one GC landclass, its coefficient consist of the mean of the 2 most occurring D&S landclasses, e.g. the coefficients for GC130 (shrubland) are calculated of the mean of those of D&S 8 and 9. The coefficient of GC landclasses representing regularly flooding of cropland or natural vegetation (GC11, 160, 170 and 180) are composed of the mean of the most occurring D&S landclass and D&S14 (water). Thus, the coefficients of these GC landclasses (and naturally of GC210 – water) differ between day and night conditions.

GC biome no	night	a_v	a_s	b_v	b_s	c_v	c_s	GC coefficients composed by D&S landclasses
11	0	0.4402	0.4402	2.9165	2.9165	-1.9238	-1.9238	(12 + 14)/2
	1	0.2576	0.2576	2.8964	2.8964	-1.8844	-1.8844	
14	0	0.881	0.881	3.4106	3.4106	-2.4133	-2.4133	12
	1	0.881	0.881	3.4106	3.4106	-2.4133	-2.4133	
20	0	0.8903	0.5989	3.3907	3.575	-2.4052	-2.5385	(2x12 + 6)/3
	1	0.8903	0.5989	3.3907	3.575	-2.4052	-2.5385	
30	0	0.8949	0.4579	3.3808	3.6572	-2.4011	-2.6012	(12 + 6)/2
	1	0.8949	0.4579	3.3808	3.6572	-2.4011	-2.6012	
40	0	0.6907	6.0951	3.8129	4.5637	-2.8455	-3.3617	1
	1	0.6907	6.0951	3.8129	4.5637	-2.8455	-3.3617	
50	0	-0.6139	4.7543	3.6472	4.3652	-2.7218	-3.2155	(2 + 3)/2
	1	-0.6139	4.7543	3.6472	4.3652	-2.7218	-3.2155	
60	0	0.9089	0.0348	3.3511	3.9038	-2.389	-2.7891	6
	1	0.9089	0.0348	3.3511	3.9038	-2.389	-2.7891	
70	0	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909	4
	1	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909	
90	0	0.9302	1.2855	3.2846	3.5544	-2.3141	-2.5071	(4 + 5)/2
	1	0.9302	1.2855	3.2846	3.5544	-2.3141	-2.5071	
100	0	0.5403	2.2634	3.4138	3.6532	-2.4345	-2.5991	(2x4 + 3)/3
	1	0.4905	2.3462	3.4138	3.6532	-2.4345	-2.5991	
110	0	0.8541	0.4171	3.4299	3.7063	-2.4477	-2.6478	(6 + 7)/2
	1	0.8541	0.4171	3.4299	3.7063	-2.4477	-2.6478	
120	0	0.931	1.2863	3.288	3.5578	-2.3164	-2.5093	(5 + 10)/2
	1	0.931	1.2863	3.288	3.5578	-2.3164	-2.5093	
130	0	1.2313	0.8399	3.3125	3.5713	-2.3663	-2.5514	(8 + 9)/2
	1	1.2313	0.8399	3.3125	3.5713	-2.3663	-2.5514	
140	0	0.7993	0.7993	3.5088	3.5088	-2.5065	-2.5065	7
	1	0.7993	0.7993	3.5088	3.5088	-2.5065	-2.5065	

150	0	1.0199	1.0199	3.3648	3.3648	-2.3606	-2.3606	(2x10 + 9)/3
	1	1.0199	1.0199	3.3648	3.3648	-2.3606	-2.3606	
160	0	0.345	3.0472	3.1177	3.4931	-2.1399	-2.398	(1 + 14)/2
	1	0.1624	2.8646	3.0976	3.473	-2.1006	-2.3586	
170	0	0.4542	0.0171	2.8867	3.1631	-1.9117	-2.1117	(6 + 14)/2
	1	0.2715	-0.1655	2.8667	3.143	-1.8723	-2.0723	
180	0	0.5406	0.5406	2.8632	2.8632	-1.8649	-1.8649	(10 + 14)/2
	1	0.3579	0.3579	2.8431	2.8431	-1.8255	-1.8255	
190	0	0.7942	0.7925	3.5969	3.5969	-2.6	-2.6	(11 + 12)/2
	1	0.7942	0.7925	3.5969	3.5969	-2.6	-2.6	
200	0	0.7075	0.7041	3.7832	3.7832	-2.7868	-2.7868	11
	1	0.7075	0.7041	3.7832	3.7832	-2.7868	-2.7868	
210	0	-0.0005	-0.0005	2.4224	2.4224	-1.4344	-1.4344	14
	1	-0.3658	-0.3658	2.3822	2.3822	-1.3556	-1.3556	
220	0	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909	13
	1	1.0801	1.0801	3.2972	3.2972	-2.2909	-2.2909	

8.2 Vegetation type cover or biome map

A specific fractional vegetation auxiliary file will be delivered and separately documented.

An example is illustrated here. The biome map of Dorman and Sellers [1989] (abb: D&S) with a resolution of 0.5° (Fig 1) previously used for the LST retrieval is replaced by Globcover biome (Fig 2) with its origin resolution of 1/360° (≈300m). Because of the 1km resolution of the AATSR satellite, LST jumps at the edge of 0.5° bins may occur when the biome changes (Fig 3 left). By using Globcover for the LST retrieval, there are no jumps.

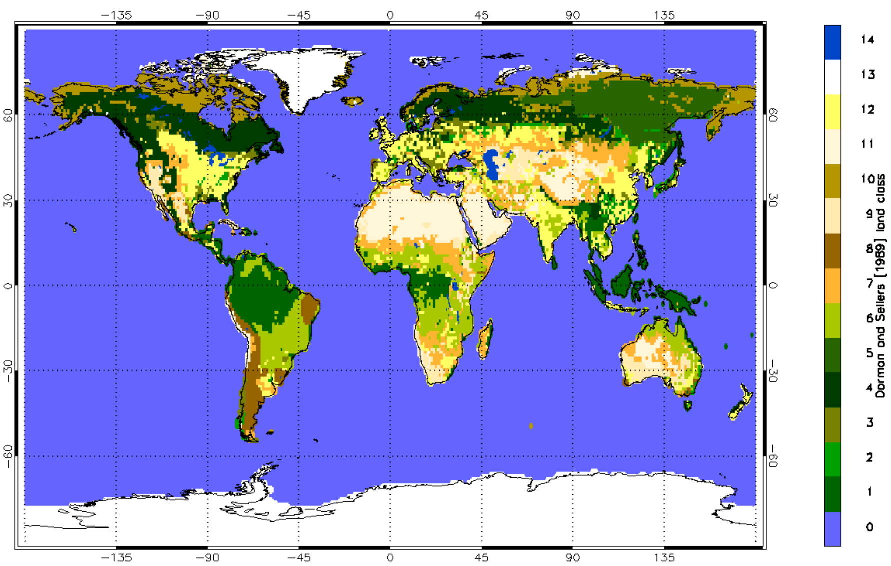

Figure 2*: Global auxiliary biome map for the previous AATSR LST retrieval from based on Dorman and Sellers [1989]*

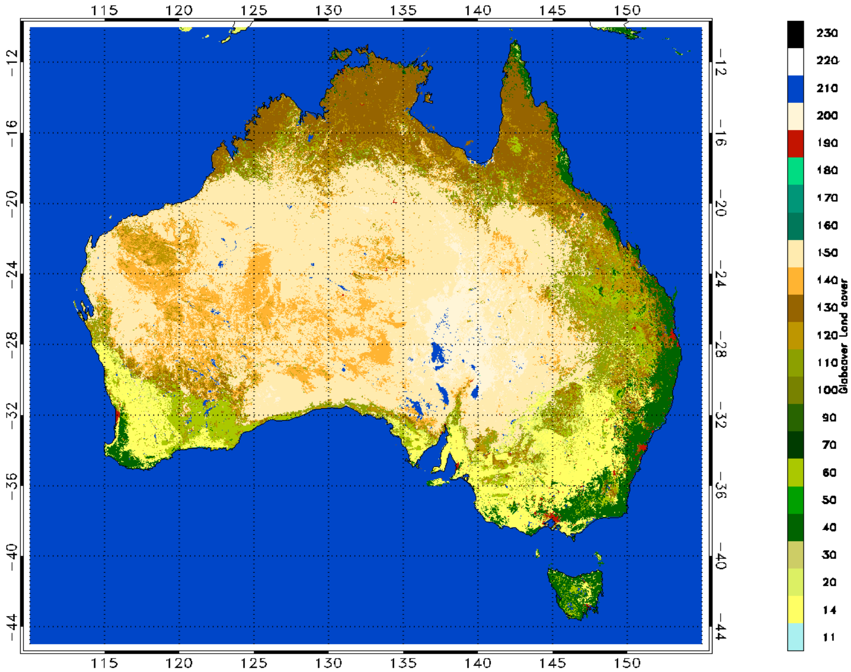

***Figure 3**: Globcover biome map for Australia as the new biome map for the LST retrieval*

Typical differences due to using higher resolution land cover data are shown in the following plots:

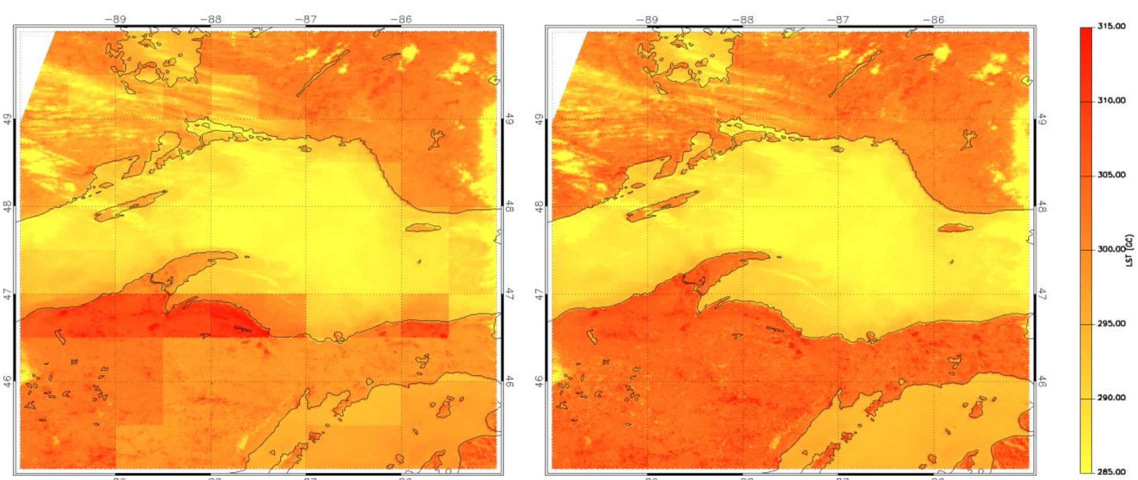

***Figure 4**: left: LST over Lake Michigan based on D&S biome during daytime from AATSR orbit no. 484 for 3rd August 2008; right: Same like left, but based on Globcover biome.*

The Biome description for Globcover is:

Biome no. GC	**Biome description**
GC 11	Post-flooding or irrigated croplands
GC 14	Rainfed croplands
GC 20	Mosaic Cropland (50-70%) / Vegetation (grassland, shrubland, forest) (20-50%)
GC 30	Mosaic Vegetation (grassland, shrubland, forest) (50-70%) / Cropland (20-50%)
GC 40	Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest (>5m)
GC 50	Closed (>40%) broadleaved deciduous forest (>5m)
GC 60	Open (15-40%) broadleaved deciduous forest (>5m)
GC 70	Closed (>40%) needleleaved evergreen forest (>5m)
GC 90	Open (15-40%) needleleaved deciduous or evergreen forest (>5m)
GC 100	Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)
GC 110	Mosaic Forest/Shrubland (50-70%) / Grassland (20-50%)
GC 120	Mosaic Grassland (50-70%) / Forest/Shrubland (20-50%)
GC 130	Closed to open (>15%) shrubland (<5m)
GC 140	Closed to open (>15%) grassland
GC 150	Sparse (>15%) vegetation (woody vegetation, shrubs, grassland)
GC 160	Closed (>40%) broadleaved forest regularly flooded – Fresh water
GC 170	Closed (>40%) broadleaved semi-deciduous and/or evergreen forest regularly flooded – Saline water
GC 180	Closed to open (>15%) vegetation (grassland, shrubland, woody vegetation) on regularly flooded or waterlogged soil – Fresh, brackish or saline water
GC 190	Artificial surfaces and associated areas (urban areas >50%)
GC 200	Bare areas
GC 210	Water bodies
GC 220	Permanent snow and ice
GC 230	No data (burnt areas, clouds,…)

8.3 Fractional vegetation cover

A specific fractional vegetation auxiliary file will be delivered and separately documented.

As an illustration of typical data, the current fractional vegetation cover file is described. This has been produced using data from the CYCLOPES (Carbon cYcle and Change in Land Observational Products from an Ensemble of Satellites) project. The time resolution is 10 days resulting into 36 fractional vegetation data per year. All fractional vegetation data from 1999 – 2003 have been used to calculate a climatology. The data domain extends between the values 0 (0.0 – no vegetation) and 250 (1.0 – full vegetation). However, there are also values of 255 for cases of corrupt data, but also over cloud, snow and sea flags. The data produced has a spatial resolution of r~1 km (1/112°).

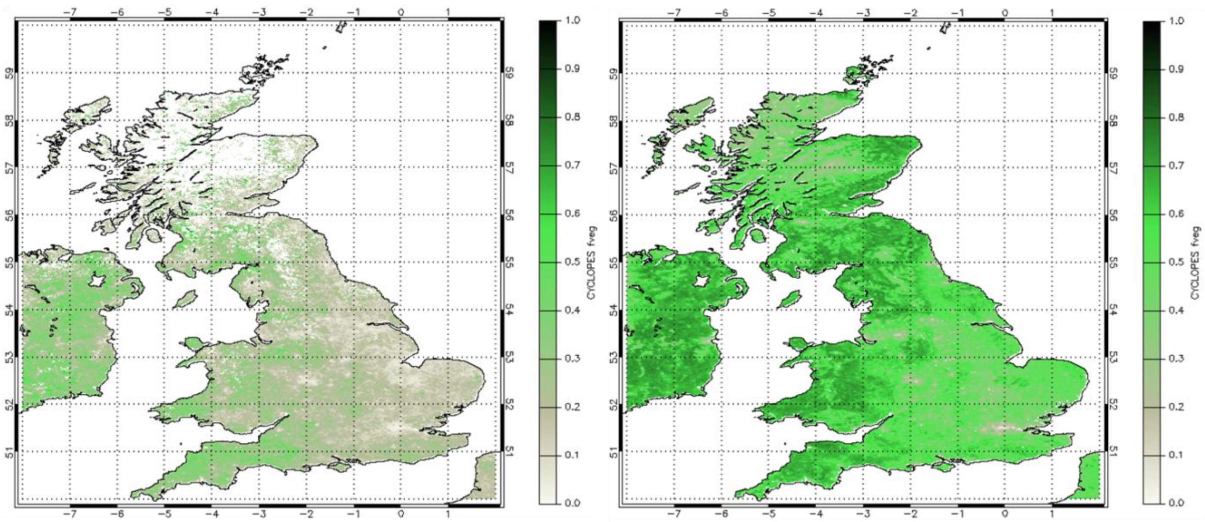

Figure 5*: Fractional vegetation cover climatology from CYCLOPES over Great Britain for 11th – 20th January (left) and 11th – 20th July (right) within the time range of 1999 – 2003.*

8.4 Precipitable water data

A specific precipitable auxiliary file will be delivered and separately documented.

An example is illustrated here. A climatology has been developed from the ECMWF Re-analysis 40 year (ERA-40) project data. The ERA-40 year data-set was selected and a new climatology derived for the last 5 years, 1997–2001. Figure 5 shows the January monthly mean PW (in units of kg m^{-2}) for the 5-year ERA-40 climatology for each of the years, 1997–2001. The final climatology will be a monthly average of these data sets. The climatology is output at four synoptic times: 00, 06, 12 and 18 UTC.

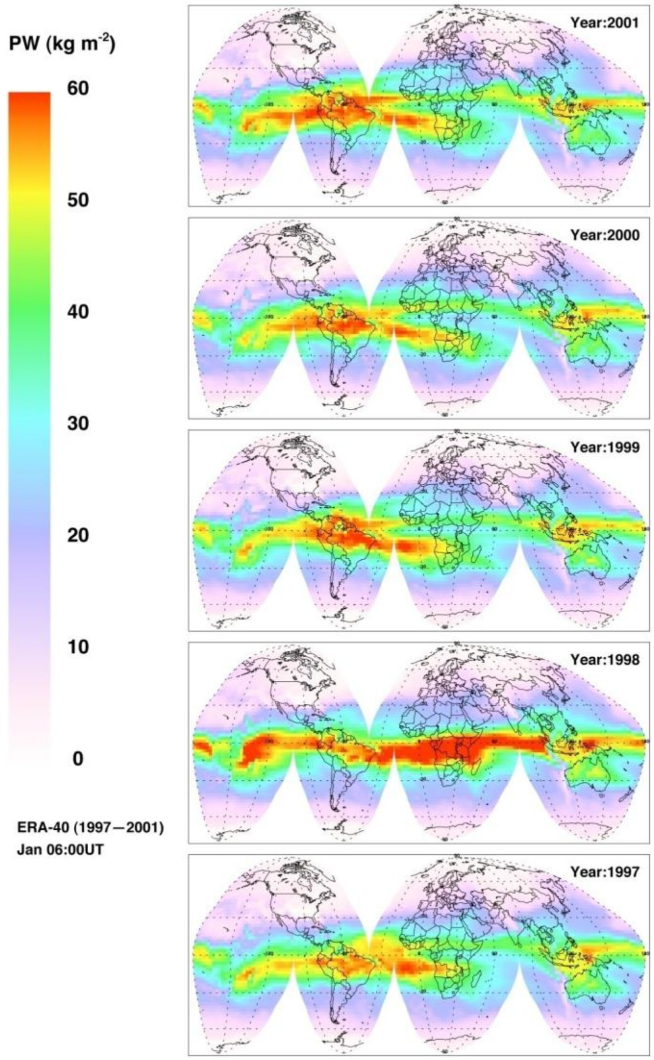

Figure 6: *ERA-40 January 06UT mean precipitable water for each year 1997–2001*.