# Land Surface Temperature Measurement from Space: AATSR Algorithm Theoretical Basis Document

Dr Fred Prata

CSIRO Atmospheric Research
Aspendale, Australia.

14 January, 2002.

# Contents

**Abstract**

The theoretical basis for a land surface temperature (LST) product using the split-window channels of the Advanced Along-Track Scanning Radiometer (AATSR) is given. Several restrictions imposed by the near-real time processor have affected the choice of algorithm. These are:

(1) the algorithm must not utilise the visible channels of the AATSR,
(2) the algorithm must be computationally fast, must not occupy large amounts of computer memory and must not utilise large amounts of disk space,
(3) a method for cloud detection over the land must be implemented, and
(4) a protocol for validation of the product must exist.

The algorithm proposed uses pixel-by-pixel top-of-the-atmosphere cloud-free, calibrated and navigated day and night brightness temperatures from the 11 and 12 $\mu$m AATSR channels to produce a global LST product. Additional seasonally-dependent land cover classification, fractional vegetation and precipitable water data are required for the operation of the algorithm. The algorithm is fast and can be used globally by employing regression coefficients in a look-up table that is updatable.

This document describes the mathematical basis for the algorithm, the ancillary data-sets required, exception handling, quality control flags and the validation strategy.

# 1 Introduction

This document provides an overview of, and rationale for, the definition of a multi-channel Land Surface Temperature (LST) data product from the AATSR. Land surface temperature (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 (about 1K) are obtained at night, when differential surface heating is absent. Since surface spectral emissivities are not known to sufficient accuracy and not measured over large parts of the land surface, an LST algorithm that has no explicit dependence on surface emissivity is required. Angular effects on directional temperature and surface emissivity are also difficult to include, so the LST algorithm must rely on the nadir AATSR split-window radiances. Further research will investigate the incorporation of the forward view AATSR data into the LST algorithm.

# 2 Algorithm overview

Over the past 5 years CSIRO has been actively developing (and validating) satellite-based algorithms (e.g. using ATSR and AVHRR data) for estimating LST. These algorithms are 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. The algorithms have been subjected to a thorough validation using a unique network of ground-truth sites across Australia. The basic algorithm may be stated as:

$$LST = a_o + b_o T_{11} + c_o T_{12}. \tag{1}$$

where $a_o$, $b_o$ and $c_o$ are coefficients that depend on atmospheric water vapour, viewing angle and land surface emissivity. 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. Moreover, these effects can be very complex (e. g. heterogeneity of mixed land covers) and suggest that accounting for them accurately 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—see section 3.7.

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 just 13 landcover classes or biomes, covering desert conditions to pine forest. To include the effects of mixed landcovers and allow for seasonal vegetation growth, coefficients are linearly

combined and a time-dependency is included. Special coefficients are used for snow and ice covered surfaces. 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, rather than by validating sets of input variables and parameters.

# 3 Algorithm description

## 3.1 Theoretical Description

### 3.1.1 Physics of the problem

The problem of determining LST from the split-window nadir channels of the AATSR is similar in concept to the problem of determining SST. The so-called split-window method, which utilizes 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 solve the problem of determining LST 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. Global surface spectral emissivity information is currently unavailable.

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 regression coefficients have physical meaning and physical constraints can be utilised to ensure their validity.

The temperature that is retrieved using the algorithm is a *radiative* 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 AATSR LST product definition.

### 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, 1994). 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 AATSR land surface temperature product (LST) is a 1x1 km$^2$, pixel by pixel quantity using only the nadir split-window (11 and 12 $\mu$m) channels of the AATSR. The product utilises the cloud-free top-of-the-atmosphere 11 and 12 $\mu$m brightness temperatures and ancillary information to correct for water vapour absorption and spectral emissivity effects. 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 provided in equation (9); following the treatment of the radiative transfer problem at the land-atmosphere interface and utilising several reasonable assumptions.

The starting point is the 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,

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

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

$$I_\nu^{atmos} = \int_o^\infty B_\nu[T(p)] \frac{\partial \tau}{\partial z}(z, \infty) dz. \tag{4}$$

where:

- $I_\nu$ is the radiance at the radiometer,
- $I_\nu^{surface}$ is the surface leaving radiance,
- $I_\nu^{atmos}$ is the radiance from the atmosphere,
- $\tau$ is the atmospheric transmittance,
- $\nu$ is wavenumber,
- $z$ is height,
- $F$ is the filter response function of the radiometer,
- s is a unit vector defining the view direction,

- $\mathrm{s}'$ is a unit vector defining the sun's direction,
- $T_s$ is the surface temperature,
- $\epsilon_\nu$ is the surface emissivity,
- $B_\nu$ is the Planck function,
- $\rho_\nu$ is the surface reflectance,
- $I_\nu^{sky}$ is the downwelling sky radiance.

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

$$\int_{\Omega^-} \mathrm{n}\cdot\mathrm{s}\epsilon_\nu(\mathrm{s})\mathrm{d}\mathbf{\Omega} = \int_{\mathbf{\Omega}^+} \mathrm{n}\cdot\mathrm{s}\left\{\mathbf{1} - \frac{\mathbf{1}}{\pi}\int_{\mathbf{\Omega}^-} \mathrm{n}\cdot\mathrm{s}'\rho_\nu(\mathrm{s},\mathrm{s}')\mathrm{d}\mathbf{\Omega}'\right\}\mathrm{d}\mathbf{\Omega}. \tag{5}$$

We assume that the surface is Lambertian. Then $\epsilon_\nu$ and $\rho_\nu$ are independent of direction,

$$\epsilon_\nu = 1 - \rho_\nu \tag{6}$$

The flux density of sky radiation is:

$$F_\nu^{sky} = \int_o^{2\pi}\int_0^{\pi/2} I_\nu^{sky}\cos\theta\sin\theta d\theta d\phi, \tag{7}$$

where $\theta$ is the satellite zenith view angle, and $\phi$ is the satellite azimuth view angle.

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

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

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

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

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 ($\approx 1$ $\mu$m width) with relatively smooth filter response functions, the variation of the Planck function with wavenumber is small. Thus an integration of the various quantities ($I_\nu$, $\epsilon_\nu$, $L_\nu^{sky}$, etc.) over the filter function is appropriate.

The definition is only valid under the assumptions outlined above. Under most circumstances we expect the assumptions to remain valid. Under some conditions they will

be violated, but we expect this to be a weak violation 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 are taking follows Prata (1994) and shows that the surface temperature may be written as a regression relation involving the brightness temperatures in the 11 and 12 $\mu$m channels. The relation takes account of atmospheric absorption (water vapour) and spectral emissivity effects.

## 3.2 Split-window approximation

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

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

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,

$$\tau_{\nu}(z,\infty) = exp\left\{ -\int_{z}^{\infty} k_{\nu}(z')w(z')sec\theta dz' \right\}. \tag{11}$$

where:

- $k_{\nu}$ 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. AATSR 11 and 12 $\mu$m channels):

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

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

Linearise around $\nu_{11}$:

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

Manipulate:

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

where:

$$\gamma = \frac{1-\tau_{11}}{\tau_{11}-\tau_{12}},$$

$$\Delta\epsilon = \epsilon_{11} - \epsilon_{12},$$

$$\alpha = -\frac{(1-\tau_{11})\tau_{12}(1-\epsilon_{12})I_{12} - (1-\tau_{12})\tau_{11}(1-\epsilon_{11})I_{11}}{\epsilon_{12}\tau_{12}(1-\tau_{11}) - \epsilon_{11}\tau_{11}(1-\tau_{12})}.$$

$I'_{11}$ 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:

$$\Delta\epsilon = 0,$$

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

(ii) Emissivity $\approx 1$ (e.g. sea surface):

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

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

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

After some manipulation,

$$LST = a_o + b_o T_{11} + c_o T_{12}, \tag{18}$$

$$a_o = \alpha \left(\frac{\partial B}{\partial T}\right)^{-1}_{\overline{T}},$$

$$b_o = \frac{1+\gamma}{\epsilon_{11}} \left[\frac{1}{1+\gamma\tau_{12}\Delta\epsilon/\epsilon_{11}}\right],$$

$$c_o = -\frac{\gamma}{\epsilon_{12}} \left[ \frac{1}{1 + (1+\gamma)\tau_{11}\Delta\epsilon/\epsilon_{12}} \right].$$

This mathematical development shows that under the assumptions it is possible to relate the brightness temperatures in the 11 and 12 $\mu$m channels linearly to the land surface temperature. We now consider some further approximations and also treat the case where the channels have the same bandpass but view the surface from two different zenith angles.

### 3.2.1 Approximate split-window form

Approximations:

(i) Sky radiances in each channel nearly equal,
(ii) Spectral emissivity dependence small ($\Delta\epsilon < 0.05$).

Consider the $b$ coefficient:

$$b_o = \frac{1+\gamma}{\delta}, \tag{19}$$

$$\delta = \epsilon_{11} + \gamma\tau_{12}\Delta\epsilon,$$

As $\Delta\epsilon \to 0$,

$$\hat{b} \approx \frac{1+\gamma}{\epsilon_{11}}. \tag{20}$$

Coefficient $c$:

$$\hat{c} \approx -\frac{\gamma}{\epsilon_{11}}. \tag{21}$$

Off-set coefficient $a$:

$$\hat{a} = \left[ \frac{1-\delta}{\delta} \right] L^{sky} \left( \frac{\partial B}{\partial T} \right)^{-1}_{\overline{T}}. \tag{22}$$

### 3.2.2 How good is this approximate form ?

Typical values: $\gamma \approx 2.40$, $\tau_{11} \approx 0.85$, $\tau_{12} \approx 0.79$, $\Delta\epsilon \approx 0.01$, $\epsilon_{11} \approx 0.97$, $\delta \approx 0.989$, and

$$L^{sky} \left( \frac{\partial B}{\partial T} \right)^{-1} |_{\overline{T}} \approx 40K.$$

e.g. With $T_{11}$=20.0°C $T_{12}$=19.0°C, the accurate value is: LST$_o$=23.1 °C and the approximate value is $L\hat{S}T$=24.3 °C. This shows that spectral emissivity effects are important for accurate LST determination.

## 3.3 Dual-angle algorithm

The dual angle algorithm follows directly from the above by considering each channel viewing the surface at zenith angle $\theta_i$ and noting that the transmittance is,

$$\tau_{\theta_i} \approx 1 - k_i w \sec\theta_i$$

The Planck radiance from the surface is then,

$$B_\nu[T_s] = \left[\frac{1+\gamma_\theta}{\epsilon_{\theta_1} + \gamma_\theta[1 - k_\nu w \sec\theta_2]\Delta\epsilon}\right] I_{\theta_1} - \left[\frac{\gamma_\theta}{\epsilon_{\theta_2} + (1+\gamma_\theta)[1 - k_\nu w \sec\theta_1]\Delta\epsilon}\right] I_{\theta_2} \tag{23}$$

where the subscripts $1$ and $2$ refer to two (different) satellite zenith view angles,

$$\gamma_\theta = \frac{1}{\cos\theta_1/\cos\theta_2 - 1}$$

is an angle-dependent parameter analogous to the atmospheric parameter $\gamma$ used earlier and

$$\Delta\epsilon_\theta = \epsilon_{\theta_1} - \epsilon_{\theta_2}.$$

$$T_s = \frac{T_{\theta_1}}{\epsilon} + \frac{\gamma_\theta}{\epsilon}(T_{\theta_1} - T_{\theta_2}) + \frac{1-\epsilon}{\epsilon}\left[B_\nu[T_{\theta_1}]\left(\frac{\partial B}{\partial T}\right)^{-1}_{T_{\theta_1}}\right]. \tag{24}$$

The dual angle algorithm has some significant advantages over the split-window form. However, there are also some disadvantages which include the problem of determining clear fields-of-view for both views, surface homogeneity and non-simultaneity of the two views. This last consideration is significant over the land only during the daytime part of the orbit. In the light of these problems only the split-window channels will be considered for the operational LST product.

## 3.4 Global LST Algorithm: Regression relation

The basic form of the algorithm is (Prata, 1993, 1994):

$$LST = a_o + b_o T_{11} + c_o T_{12}. \tag{25}$$

- $a_o$, $b_o$, and $c_o$ are 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 brightness temperatures (top-of-the atmosphere) in the AATSR 11 $\mu$m and 12 $\mu$m channels.
- $\theta$ 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,

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

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

The algorithm has been developed with all temperatures in units of degrees Celsius and the algorithm returns the land surface temperature in units of degrees Celsius.

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 in the simulation data-set for desert regions, particularly at night and in the morning shortly after sunrise.

The coefficients are:

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

$$b_{f,i} = fb_{v,i} + (1 - f)b_{s,i},$$

$$c_{f,i} = fc_{v,i} + (1 - f)c_{s,i},$$

$$0 \leq f \leq 1.$$

Values for the coefficients have been determined using simulation data-sets. Updated values for all parameters will be supplied nearer launch time and from time to time after launch using the quality control and validation procedures described later. The parameters $d$ and $m$ were empirically determined using the radiative transfer simulations for regions where some validation data were available. Since both of these parameters only weakly affect the retrieval of LST, their precise values are not needed and again these are left as 'tuning' parameters.

The fractional vegetation cover is given by $f$. 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. At the commencement of the AATSR commissioning, tables of coefficients will be available for immediate use. These tables will initially consist of regression coefficients for 13 biomes (landcover classes), and one class representing lakes, viz.

- Broadleaf evergreen trees,
- Broadleaf deciduous trees,

- Broadleaf and needleleaf trees,
- Needleleaf- evergreen trees,
- Needleleaf-deciduous trees,
- Broadleaf trees with groundcover,
- Groundcover (grassland, Uardry cal/val site),
- Broadleaf shrubs with groundcover (tropical savanna, Thangoo cal/val site),
- Broadleaf shrubs with bare soil,
- Dwarf trees, shrubs with groundcover,
- Bare soil (desert, Amburla cal/val site),
- Broadleaf-deciduous trees with winter wheat,
- Perennial land ice.
- Permanent lakes.

These biomes were determined from Dorman and Sellers (1989) and a seasonal variation (monthly) and fractional cover map have been developed from the data, which is tabulated at 1°x1° resolution. The precipitable water data is based on the NVAP climatology at 2.5°x2.5° resolution and monthly intervals. These data will be supplied as a baseline climatology; other sources of water vapour information (e. g. from a microwave instrument or infrared data) could be used if they become available in a timely manner.

The initial set of coefficients have been determined via a modelling and simulation approach utilising previously derived coefficients and AVHRR/ATSR and validation data-sets at several continental cal/val sites in Australia (Prata, 1994; Prata and Cechet, 1999). For ice covered surfaces, coefficients derived by Key and Haeflager (1992) and Stroeve and Stefan (1998) based on AVHRR data have been utilised.

From time to time these coefficients will be updated and accessed via a dedicated web page. The updating of the coefficients will be done when necessary after completion of detailed analyses of the performance of the algorithm in the three target biomes −grassland, bare soil/desert and tropical savanna.

## 3.5 Vegetation fraction and temporal variation

For some of the biomes it has been recognised that the temporal variation of the vegetation fraction may be important. To provide a realistic estimate of this variation in the climatological mean sense, two data-sets have been used. These are the Dorman and Sellars (1989) monthly vegetation fraction estimates for the 13 biomes given above (biome 14−permanent lakes−assumes no vegetation, but in an improvement the fraction could be used to signify changes in lake area which would then allow inclusion of the ephemeral inland lakes of the arid zones), and an estimate of 'greenness' derived from the ISLSCP release of the global normalised difference vegetation index (NDVI) global composite at 1°x1° spatial resolution and monthly temporal resolution. The two data-sets are combined

in the following manner:
First, for ech biome the Dorman and Sellars vegetation fraction is obtained. This fraction lies between 0 and 1 by definition. Then for each 0.5°x0.5° cell, the ISLSCP NDVI is retrieved and multiplied by the D&S fraction. The maximum NDVI over the year (at a monthly time interval) for each cell is used to normalise the result. Finally, to account for differences in seasons between the northern and southern hemisphere months, the D&S vegetation fractions are used as if month 1 is January in the northern hemisphere and July in the southern hemisphere, and month 6 is June in the northern hemisphere, but December in the southern hemisphere. Table 1 below makes this coding more explicit.

Table 1: Indexing used to construct the normalised vegetation fractions for northern (NH) and southern (SH) hemispheres, based on the Dorman and Sellars (1989) phenology.

| Month | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NH | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| SH | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 |

## 3.6 Emissivity

It is well-known that variations in surface properties cause variations in the emission of radiation from natural surfaces. 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 $\mu$m (Sutherland, 1979). Over highly vegetated surfaces (e.g. closed-canopy trees) the emissivity is known to be spectrally uniform and high ($\epsilon_{11} > 0.98$, 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 $\approx \pm 0.01$, 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, but the NDVI could have been used or perhaps a scene based measure using visible data from the AATSR. 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 non-existent. 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 $\pm 0.02$. It is likely that the retrieval errors and biases are simply re-mapped from transmittance errors, since the radiative transfer problem shows that the surface emissivity and atmospheric transmittance always appear as

a product and separating their effects accurately suggests that the transmittance must be known at least to the same accuracy.

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 $\epsilon_\nu$ from (9). There is no intention at present to determine the emissivities as part of the LST product, nor as a separate product.

## 3.7 Practical Considerations

### 3.7.1 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. Interpolation of the regression coefficients will also be required.

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

- Nadir 11 $\mu$m brightness temperature,
- Nadir 12 $\mu$m brightness temperature,
- Geographic latitude of pixel,
- Geographic longitude of pixel,
- Nadir zenith view angle at pixel,
- Nadir azimuth view angle at pixel,
- Cloud flag,
- Land/sea flag,
- Mask value,
- Time at pixel,
- Date,
- Ancillary data consisting of precipitable water, vegetation type and vegetation fraction.

## 3.8 Calibration and Validation

The AATSR brightness temperatures are assumed to be calibrated in-flight and no instrumental or other anomalous effects will be corrected for. Validation of the brightness temperatures will not be done in a systematic manner. It will be assumed that validation measures put into place as part of the sea surface temperature validation protocol will apply to the LST product. Some limited opportunities exist for validation of the at surface brightness temperatures because radiometers will be deployed over some land covers and lake surfaces as part of the land AATSR validation program.

The major part of the validation exercise will involve routine collection and monitoring of land surface temperatures at three principal cal/val sites situated in Australia. These sites will supply continuous LST estimates from in situ sensors and radiometers. Some of the measurements will be made at off-nadir angles, allowing some validation of the forward view radiances. In general the validation data will be collected, quality controlled and archived by CSIRO. Within a few hours of the AATSR overpass validation data (with some quality control) will be released via a dedicated web page. These snapshot data will allow immediate monitoring of the LST product and permit a quick assessment of the performance of the algorithm over three landcovers: grassland, bare soil (desert) and acacia woodland (tropical savanna).

Further aspects of the calibration and validation program for the AATSR thermal products can be found in the AATSR calibration and validation (cal/val) documentation described by Parkes et al. (1998) and available on the world wide web at:

`http://www.le.ac.uk/physics/research/eos/aatsr/val1.html.`

### 3.8.1 LST validation protocol

The ideal method for obtaining an independent validation data-set is to use two multichannel radiometers with AATSR bandpasses from an aircraft and from the ground. The measurements must be made as close as possible to the overpass time and ancillary data relating to atmospheric profiles and emissivity measurements must be collected. The target validation site must be uniform at scales of 100's m to several kilometres. The sky must be clear during the measurement period and the aerosol content of the atmosphere must be low. The measurement protocol would consist of the following:

**[a-1]** The airborne and ground-based instruments must be calibrated to an accuracy of $\pm 0.1$ K and be traceable to a NIST blackbody.

**[b-1]** Vertical profiles of temperature and moisture must be obtained within $\pm 10$ minutes of the radiometer measurements and within $\pm$ 10 km of the target area.

**[c-1]** Directional sky radiance measurements using a ground-based duplicate radiometer at the target should be made at the time of the overpass. Sufficient directional measurements should be made to permit an accurate determination of the flux density of spectral sky radiance.

**[d-1]** Sun photometer measurements of at least the aerosol optical depth are required at the target area using the same temporal and spatial considerations as the profile data.

**[e-1]** Directional spectral ($8-14$ $\mu$m) emissivity measurements must be made at the target with appropriate spatial sampling and within a few days of the overpass, provided no significant changes have occurred at the target surface during the period of measurements (e.g. no rainfall, fires etc.). The emissivity measurements should be made as close as possible to the local time of overpass.

**[f-1]** For a vegetated target, measurements of the density and height of the vegetation might prove useful. The fractional vegetation cover must be estimated and emissivity measurements of the components must be made.

**[g-1]** Daytime and nighttime validation data will be required.

**[h-1]** An all-sky camera should be used at the target to provide an objective measure of cloudiness at the time of the overpass during the daytime. At night an upward viewing pyrgeometer should be used.

It is unlikely that this ideal measurement protocol will be met on many occasions. Thus at least three alternate strategies will be used to provide sufficient validation data to give confidence in the LST product. These strategies will consist of:
(i) gathering large amounts of less accurate near-simultaneous temperature measurements at well-chosen cal/val sites,
(ii) using numerical model temperature simulations from high resolution mesoscale models, and
(iii) comparison of the AATSR LST product with MODIS, AVHRR and GLI LST products.

The first strategy has been implemented for a limited set of sites in Australia, which cover three of the major land biomes. The data from these sites consist of contact temperature measurements and some radiometric measurements. The contact temperatures have been shown (Prata, 1994) to be representative of the satellite measurement, but biases will exist under certain atmospheric and surface conditions. The radiometer measurements also show a strong correlation with the satellite data, but these measurements must be emissivity corrected (this also involves a measure of the sky radiance) prior to comparison with the AATSR LST product. An intermediate comparison of the AATSR 'at surface' (atmospherically corrected) brightness temperatures may be possible with the uncorrected surface-based radiometer measurements. Currently there are 3 radiometers operating at the tropical site (Thangoo), 4 radiometers at the grassland site (Uardry) and 1 radiometer at the arid, bare-soil site (Amburla). The validation data-set from the Australian sites will consist of:

**[a-2a]** 30-minute time-series of 2-minute temperature averages and standard deviations, centred on the time of the overpass. The temperatures will consist of spatially averaged contact measurements over an area of approximately 1 x 1 $km^2$. The number of sensors used in the spatial average will vary.

**[a-2b]** 30-minute time series of radiometer measurements (average and standard deviations) appropriate for the nadir and forward views of the AATSR.

**[b-2]** Radiosonde profiles of temperature and moisture made at the nearest upper-air station to the site. In the case of Uardry, the station is Wagga Wagga about 200 km east of the site and the sondes are launched between 08:30 LT and 09:30 LT, daily. In the case of Amburla, the station is at Alice Springs about 100 km east of the site and the sondes are launched at 09:00 LT, daily. In the case of Thangoo, the station is Broome about 30 km north of the site and sondes are launched at 07:00 to 08:00 LT and 19:00 to 20:00 LT, sometimes twice per day.

**[c-2]** 30-minute time-series of upward and downward longwave and shortwave flux density measurements.

**[d-2]** 30-minute time-series of aerosol optical depths at 0.55, 0.67 and 0.87 $\mu$m obtained from multi-filter rotating shadowband radiometers (MFRSR) located at each site.

**[e-2]** Library spectral emissivity profile measurements made at each site. Only one data set per site is envisaged, but this may be expanded to several seasonal data-sets as resources permit.

**[f-2]** Standard surface meteorological data (air temperature, wind-speed, humidity and surface pressure) at each site as 30-minute time-series centred on the overpass time.

The data-set required from the mesoscale model will depend on the exact details of the model employed. A minimum data-set might comprise:

**[a-3]** Surface temperature. This must be as close to the surface as possible without being below the ground. The lowest vertical level in the atmosphere will not be appropriate. Inversion of the Stefan-Boltzmann flux density to get a radiative temperature would be an acceptable value to use.

**[b-3]** Vertical profiles of temperature and moisture.

**[c-3]** Description of surface character, e.g. longwave or spectral emissivity (if possible), albedo, soil moisture status.

**[d-3]** Model estimate of degree of cloudiness.

**[e-3]** Description of model grid size, integration step size and vertical resolution.

Comparisons with products from other satellites can only be done when the measurements are nearly simultaneous. The kind of information required for this validation will include:

**[a-4a]** 11 and 12 $\mu$m top-of-the-atmosphere brightness temperatures at the site.

**[a-4b]** Pixel-by-pixel LST at a minimum of 5 x 5 pixels centred on the pixel closest to the site.

**[a-4c]** Pixel-by-pixel spectral emissivity retrievals (if available) at a minimum of 5 x 5 pixels centred on the pixel closest to the site.

**[b-4]** Satellite date and time of acquisition of collocated pixel.

**[c-4]** Satellite view zenith and azimuth angles of collocated pixel; for daytime data, sun zenith and azimuth angles.

**[d-4]** Satellite measure of cloudiness. This could be a cloud flag, a confidence flag or for daytime data reflectance values in the shortwave channels of the sensor being used.

**[e-4]** Description of the theoretical basis for the LST product.

### 3.8.2 Quality Control and Diagnostics

Quality control will be done in 5 main ways:

1. Quality control of the AATSR brightness temperatures using the on-board calibration subsystem.

2. Routine monitoring of the brightness temperatures over the Australian cal/val sites.

3. Systematic checks of the LST product against the ground temperatures measured at the grassland site (Uardry, New South Wales), the bare soil/desert site (Amburla, Northern Territory), and the tropical savanna site (Thangoo, Western Australia).

4. Systematic checks of the LST product against the output of a regional high resolution GCM (e. g. CSIROs DARLAM). These checks will be used to diagnose gross deficiencies in the LST algorithm, where errors can be noted on physical grounds.

5. Intercomparison of the AATSR LST product with MODIS, ADEOS-II, AVHRR-3 and ASTER LST products. Such comparisons will necessarily be performed sometime after data collection and cannot be relied upon to quality control the near-real time products.

### 3.8.3 Exception Handling

All thermal data products must carry a quality control flag, which will indicate the confidence level of the associated product. This flag could be contained in a single byte or conform to some other format, but it must be present. Two types of exceptions are expected for the AATSR thermal data products. These are:

- Data validity flags. Cloudy data, corrupt data and missing data must be flagged appropriately to terminate further processing. This implies that a quality data flag must be accessed by the code providing the AATSR thermal data products.

- Data mask flags. The AATSR thermal data product is a global product. However, there are known regions of the globe where a thermal data product will be of doubtful value. Mountainous regions, inland lakes and seas, and coastal regions will be masked to indicate that the product is of suspect quality. These exceptions can be handled by either terminating processing or providing a thermal product with a poor quality flag.

### 3.8.4 Output Products

There will be three AATSR thermal land products. A pixel-by-pixel land surface temperature (LSTPbP) An NxN pixel average land surface temperature (LSTNxN). A 0.5 degree by 0.5 degree average land surface temperature (ALST) Each product will have a quality control flag to indicate the confidence level of the product. The LSTPbP product assumes that the pixel is cloud free and the input brightness temperatures are valid. The LSTNxN product provides the average of nxn cloud-free pixels where, n <N. In practice this product can be generated from LSTPbP. (The value of N is to be determined, but a value of N=3 to N=9 would be practical). The final product, ALST, is a 50 km x 50 km

spatially averaged product, similar in concept to the ATSR-2 ASST product. It can be formed directly from cloud-free LSTPbP products.

### 3.9 Error Budget Estimates

We define the Noise Amplification Factor (NAF, see Pearce *et al.*, 1989) as

$$NAF = \sqrt{b_o^2 + c_o^2} \tag{27}$$

The error due to instrumental noise in the LST depends on the NE$\Delta$T of the 11 $\mu$m and 12 $\mu$m brightness temperatures. At 300 K these errors are expected to be less than 0.1 K. Assuming that they are the same in both channels and equal to $\delta$T, then the error due to instrumental effects alone will be $\sqrt{2}\cdot$NAF$\cdot\delta T$ (Kelvin).

Another source of error is due to incorrect specification of the ancillary information required to generate the precipitable water correction. This error is unknown at present, but we anticipate that it will not be significant for this algorithm because the explicit dependence on pw is small.

Errors due to undetected or subpixel cloud will probably be significant and worse at night. For warm clouds at night it is highly likely that they will be undetected. This source of error is very difficult to quantify. Experience with ATSR-2 data suggests that at night the forward view radiances can be used to check on the temperature uniformity of the pixel. During the day it is not possible to use the forward radiances in this way because the 2 minute interval causes temperature changes that may be indistinguishable from subpixel clouds.

The in situ validation data will be the major source of information for quantifying the magnitude of the error in the LST estimate. Care will be taken to ensure coincidence and simultaneity of the validation and satellite data. Current accuracies for determining LSTs from the AVHRR-2, ATSR-2 and GMS-5 thermal channels suggests errors of 1-2 K at night and 2-4 K during the day. Consequently the target accuracies for the LSTPbP product will be: 1.0 K for night time products and 2.5 K for daytime products.

## 4 Assumptions and limitations

The following assumptions have been made:

1. The nadir AATSR brightness temperatures have been flagged as cloud-free.
2. The nadir AATSR brightness temperatures have been flagged as valid data.
3. Time, date, zenith view angle and azimuth view angle are available for each pixel.
4. Geographic location of each pixel is available.
5. 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. 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.

# Appendix I: Exact number and size of LUTs.

The following are the LUTs required:

## 4.1 LST Coefficients

3 REAL*4 arrays, each of size 2x2x14x4 bytes. The arrays could be defined in FORTRAN77 using this code fragment:

REAL*4 A(2,2,14),B(2,2,14),C(2,2,14)

The first index corresponds to vegetation or soil. The second corresponds to day or night. The third corresponds to 13 biomes and 1 surface type representing permanent lakes.

## 4.2 Fractional vegetation cover

A 0.5 x 0.5 degree latitude/longitude vegetation fraction scaled INTEGER*2 array of size 720 x 360 x 2 x 12 bytes. The temporal variation of surface characteristics is divided into 12 monthly time periods.

The array could be defined in FORTRAN77 using this code fragment:

INTEGER*2 FRAC(720,360,12)

## 4.3 Vegetation type cover map

A 0.5 x 0.5 degree latitude/longitude resolution vegetation cover INTGEGER*! array of size 720 x 360 bytes.

The array could be defined in FORTRAN77 using this code fragment:

INTEGER*1 VEG(720,360)

This array contains values in the range 1-13, corresponding to the 13 biomes. The current release does not include type 14−permanent lakes, but this land type will be included in the next release. A value of 0 is used for ocean. The proposed global land cover classes and fractional vegetation cover maps are shown in Figs. 1 and 2.

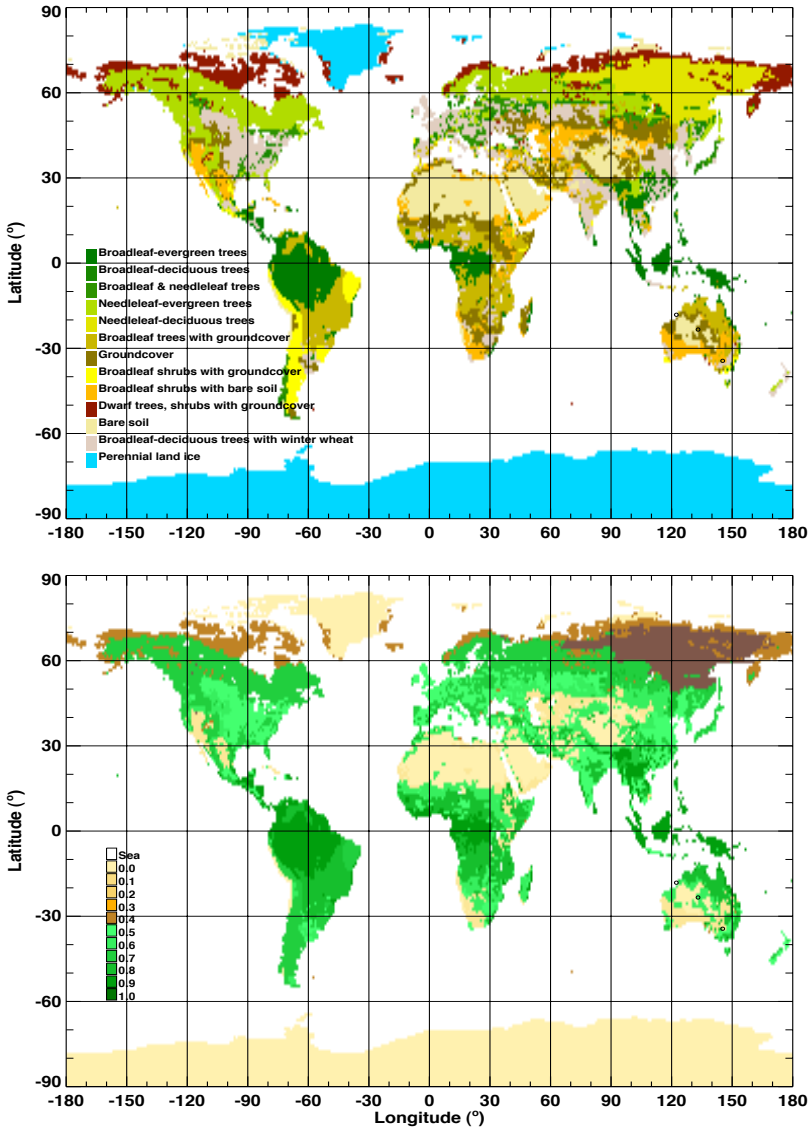


Figure 1. Global land cover classes and fractional vegetation cover.

As explained previously the temporal variation of fractional vegetation cover is achieved by using the ISLSCP 'greenness index' coupled with the Dorman and Sellars (1989) fractions by cover type. An example of the time-varying fractional cover is shown in Figure 2. In this map the fractional vegetation cover varies within each 0.5° x 0.5° cell from 0 to 1, for each cover class. For perennial land ice and bare surfaces the fractional vegetation cover should always be 0. Note that the data used cover only a two year period 1987-1988 and therefore may not necessarily represent the true climatology of any particular region.

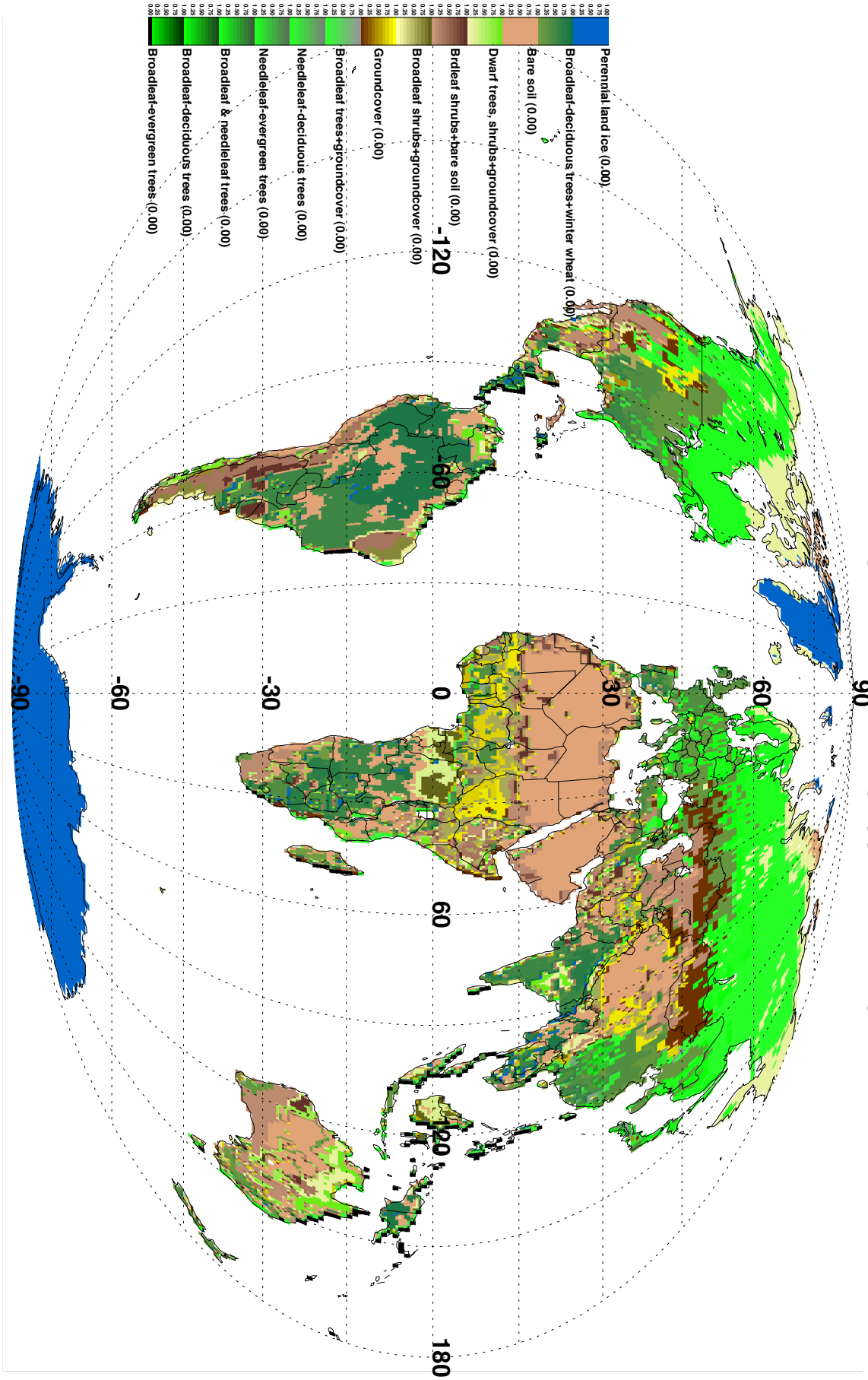


Figure 3. Global land vegetation fractions by type for January.

## 4.4 Precipitable water data

A 0.5 x 0.5 degree latitude/longitude resolution precipitable water INTEGER*2 array of size 720 x 360 x 12 x 2 bytes. Currently these data are derived from the 8-year NVAP data-set as monthly mean values.

The array could be defined in FORTRAN77 using this code fragment:

INTEGER*2 PW(720,360,12)

These are the 4 LUTs required to implement the AATSR LST algorithm. The array sizes depend on the spatial resolution of the ancillary data. Should it be deemed necessary to use higher resolution data (e.g. 1 km x 1 km) then the array sizes will increase accordingly. Global annual mean NVAP precipitable water and its standard deviation are shown in Fig. 3.

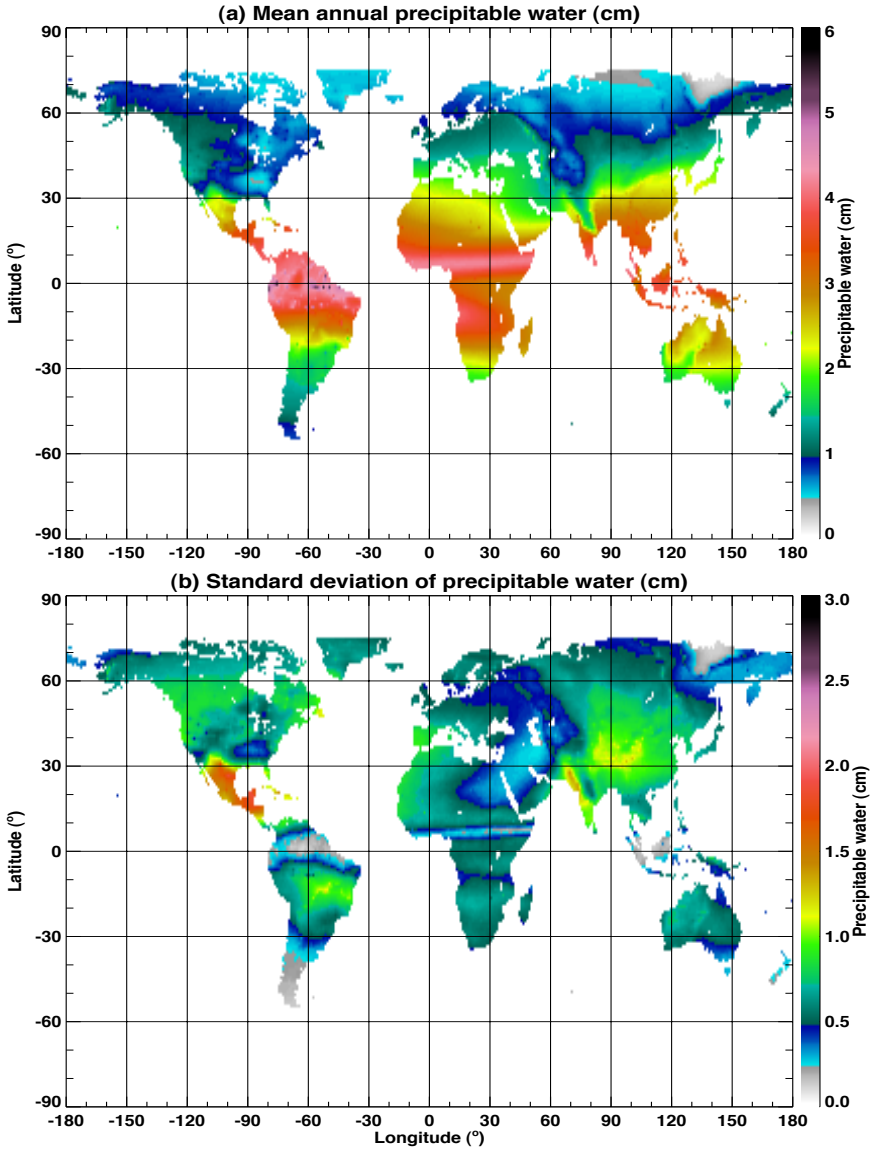


Figure 3(a). Annual mean precipitable water (in cm) derived from the NVAP analysis. (b) Standard deviation of annual mean.

## 4.5 Initial LUT generation and updating

- Now until launch: Freeze the LUTs at some point before launch and after full testing of the processor. Initial LUTs will be provided by CSIRO.
- After launch to 1st 6 months: Review performance and evaluate need to update LUTs based on validation data. During this period the LUTs may be updated quite frequently and a web-based method will be employed.
- After 1st year: Once a full-year of data (capturing the annual cycle in LST) has been acquired a second review should be conducted.
- After the first full year of acquisition we should be in a position to stabilise the LUTs.

# Appendix II: Flags

## 4.6 Data masking

There are three possible 'problem' areas envisaged. These are:

- Land/sea boundaries or coastal areas. The land/sea mask could be used to mask these areas. Daytime AATSR data could also be used by setting reflectance thresholds, but this may need some tuning because of clouds, sea fogs and scene brightness changes.
- Mountainous regions. These are easily masked at high resolution using an altitude threshold based on existing topographic data-sets.
- Lakes, rivers, flooded areas. These areas are more dynamic than the last two categories and it is not clear how to mask them out. Initially they should be included and a set of coefficients be developed for estimating t heir surface temperature. Maybe an SST algorithm would be adequate. Large lakes within the 0.5 x 0.5 degree vegetation fraction map - simply set a value of say 1.1 as being water on land. Smaller lakes and rivers would not be captured by this method and it is likely that either a global river/lake mask would need to be established or these locations are found using a scene-based algorithm (e. g. reflectance ratio and threshold).

## 4.7 Confidence flags

### 4.7.1 Topographic Variance Flag (TVF)

A topographic variance flag will be used to give the user an indication of the confidence of the LST retrieval for the particular grid square. This is a 2-bit flag (0 to 3) giving the size of the variation of the topography over a 0.5 degree by 0.5 degree grid square. It has been generated from a 30 arc second digital elevation model (DEM) and degraded to the 0.5 degree nominal grid size. The value of the flag has the following meaning:

| Flag value | Meaning |
|---|---|
| 0 | Extremely flat ground (very high confidence) |
| 1 | Some topographic variation (good confidence) |
| 2 | Significant topographic variation (low confidence) |
| 3 | Extreme topographic variation (No confidence) |

The TVF can be stored in 720 x 360 / 4 bytes. However, it has been implemented here by storing the flag in the lowest 2-bits of 720 x 360 8-bit words (bytes). The TVF is illustrated in Fig. 4.

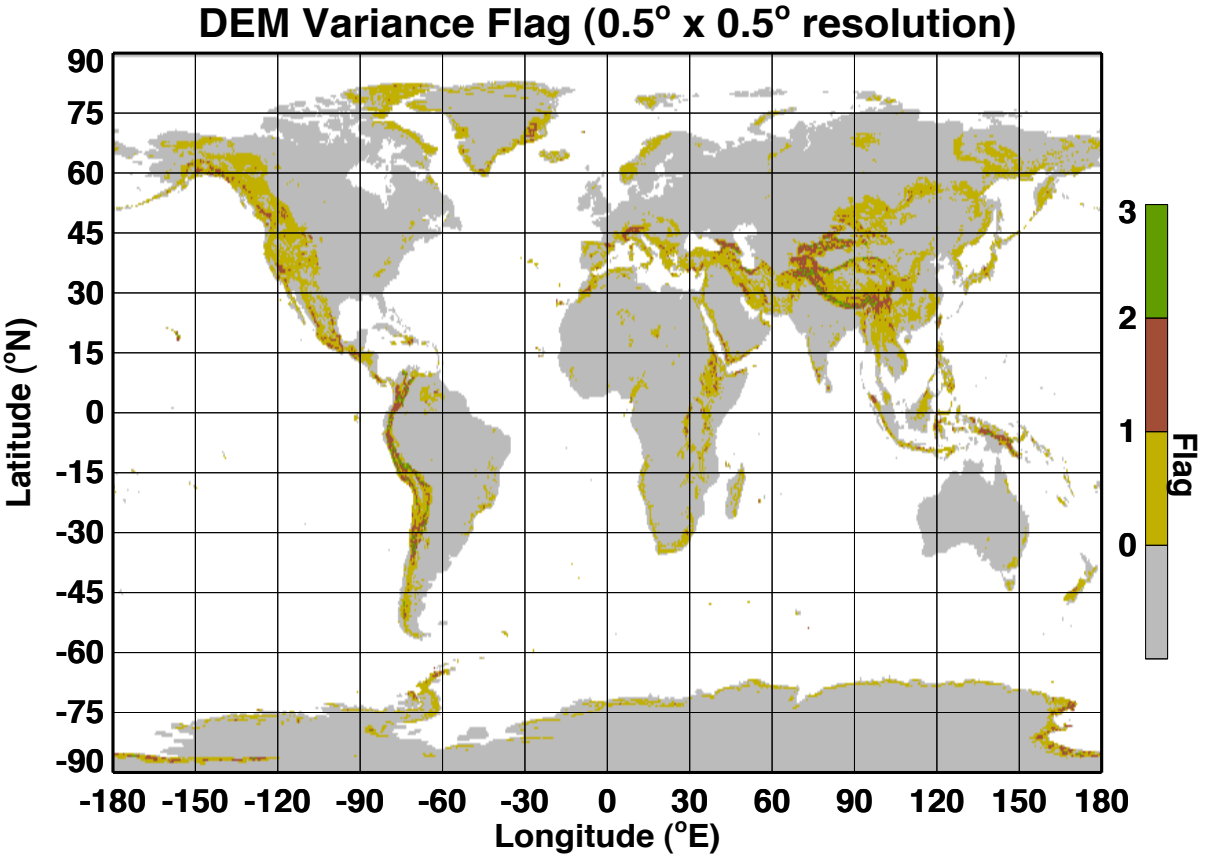


Figure 4. Global topographic variance flag.

### 4.7.2 Validation data flag

Currently no data confidence flag based on validation is planned, however this will be researched.

## 4.8 Quality flags

The main criterion here would be the cloudiness of the pixel. This means good cloud detection algorithms over land should be in place. Plans are being made to investigate simple and effective cloud detection algorithms over land. A temperature saturation flag will be used to indicate pixels over land that have reached or exceeded the maximum number of bits used to represent the digital count.

# Appendix III: Regression coefficients

The *initial* regression coefficients to be used with equation (9) are given in Table 1. These will be updated once AATSR data are processed and analysed.

Table 2: Regression Coefficients for the AATSR Global LST Algorithm.

| Land Cover Class | $a_{v,i}$ | $a_{s,i}$ | $b_{v,i}$ | $b_{s,i}$ | $c_{v,i}$ | $c_{s,i}$ |
|---|---|---|---|---|---|---|
| 1 | +0.6006 | +5.3001 | +3.3156 | +3.9684 | −2.4744 | −2.9232 |
| 2 | −1.3994 | +3.3001 | +3.3156 | +3.9684 | −2.4744 | −2.9232 |
| 3 | −1.3532 | +3.7078 | +3.3156 | +3.9684 | −2.4744 | −2.9232 |
| 4 | +0.7880 | +0.7880 | +3.3305 | +3.3305 | −2.3140 | −2.3140 |
| 5 | −0.0198 | +0.6980 | +3.3051 | +3.8502 | −2.3610 | −2.7508 |
| 6 | +0.1027 | −0.7219 | +3.1614 | +3.6828 | −2.2583 | −2.6312 |
| 7 | −0.1957 | −0.1957 | +3.4232 | +3.4232 | −2.4454 | −2.4454 |
| 8 | +0.8329 | +0.0801 | +3.0177 | +3.5154 | −2.1557 | −2.5116 |
| 9 | −0.1957 | −0.1957 | +3.3526 | +3.3526 | −2.3950 | −2.3950 |
| 10 | +0.7880 | +0.7880 | +3.3305 | +3.3305 | −2.3140 | −2.3140 |
| 11 | +0.4847 | +0.4847 | +3.6377 | +3.6377 | −2.6796 | −2.6796 |
| 12 | +0.2689 | +0.2689 | +3.2637 | +3.2637 | −2.3094 | −2.3094 |
| 13 | +0.7880 | +0.7880 | +3.3305 | +3.3305 | −2.3140 | −2.3140 |
| 14 day | −0.0005 | −0.0005 | +2.4225 | +2.4225 | −1.4344 | −1.4344 |
| 14 night | −0.3658 | −0.3658 | +2.3823 | +2.3823 | −1.3556 | −1.3556 |

# Appendix IV: Lake surface temperatures

Regression coefficients for lake surface temperature have been developed for the ATSR-2 by analysis of a set of *in situ* radiometric measurements made on Lake Tahoe during 1999. The data, their reduction and analysis and the regression results are reported in Hook *et al.* (*submitted*, 2002). The form of the algorithm is the classical expression:

$$LkST = a_o + b_o T_{11} + c_o T_{12}$$

, where the temperatures are expressed in degrees Celsius, and $LkST$ is the lake surface temperature (skin) also in degrees Celsius. Although this is not exactly the same as the nonlinear expression used for LST, the code is easily modified to include separate coefficients for lakes by setting $n = 1$ in (26) and either setting $d = 0$ or defaulting the precipitable water to zero for land type 14 (permanent lakes). The anticipated accuracy of the algorithm is $\pm$0.3 °C.

Coefficients are supplied for day and night separately and for consistency with the LST algorithm, the coefficients are repeated for the fraction of vegetation and bare ground. A side benefit of retaining this formulation is that in a future release, the fractional covers can be used to represent lake area change, where $f$ could signify the amount of standing water in the cell and $1 - f$ the amount of dry lake bed. This would be most useful for the semi-arid ephemeral lakes of Australia, Africa and USA.

# Appendix V: An Implementation of the CLAVR Cloud Detection Algorithm for ATSR-2/AATSR

The CLAVR cloud detection algorithm has been developed by Stowe *et al.* (1999) to sense cloud cover using the AVHRR GAC (4 km) data. The algorithm has been used on a global basis and offers the possibility of a good general purpose and relatively efficient operational cloud detection scheme. The algorithm uses the combination of a number of spatial and spectral tests to determine if a pixel in an AVHRR scene is cloud free. The general methodology is, however, applicable to a range of other sensors.

## Method

The algorithm uses a sequence of multi-spectral contrast, spectral and spatial signature threshold tests to classify the AVHRR scene into clear, cloudy and mixed categories. Figure 1 shows that there are really four groups of tests comprising the CLAVR algorithm with each group being specific to a combination of day/night and land/ocean.

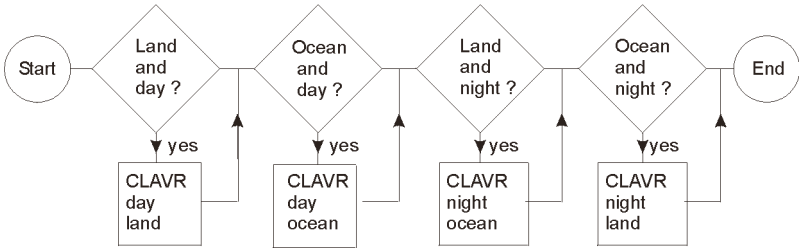


Figure 5. CLAVR algorithm structure.

For the purposes of the land-surface temperature product the cloud ocean elements can be omitted. The cloud tests used for both the AVHRR GAC version and the ATSR-2 version comprise a set of gross reflectance and thermal threshold tests, a reflectance ratio test, an 11 12 $\mu$m threshold test, and a spatial uniformity test in both reflectance and thermal. The ATSR-2/AATSR tests use the 0.67, 0.87, 3.7, 11 and 12 $\mu$m channels as for the AVHRR channels 1 to 5. The 0.55 and the 1.6 mm channels are not used. To accommodate the expected higher variance in the spatial uniformity tests due to the higher resolution of the pixels the spatial uniformity thresholds have been increased. Other test thresholds have not be adjusted. Tests like the land and ocean 11 12 $\mu$m test that use theoretically derived parameters may need adjustment to suit the exact ATSR channel wavelengths. However, it does not appear that there should be any large changes. The CLAVR cirrus cloud test based on 3.7 and 11 $\mu$m also uses and AVHRR channel 1 count value. This has been omitted. Stowe *et al.* suggest that the algorithm has difficulties in winter high latitudes, over some deserts and mountainous regions and over areas of ocean specular reflection. Stowe et al. further suggest that CLAVR tends to underestimate cloud amount when it is large and over estimate when small.

## Implementation

The CLAVR-1 algorithm as described by Stowe *et al.* (1999) comprises four component algorithms for detection cloud over combinations of ocean and land in daylight and at

night. Many of the basic tests are common to each component. Table summarises the tests applied.

Table 3: Summary of cloud tests applied to ATSR-2/AATSR.

| Test | Channels | Day | | Night | |
|---|---|---|---|---|---|
| | | Land | Ocean | Land | Ocean |
| Reflectance Gross Cloud | 0.67 $\mu$m threshold | yes | yes | no | no |
| Reflectance Uniformity | 0.67 $\mu$m spatial | yes | yes | no | no |
| Reflectance Ratio | 0.87/0.67 $\mu$m | yes | yes | no | no |
| Channel 3 (3.7 $\mu$m) albedo | 3.7, 11 and 12 $\mu$m | yes | yes | no | no |
| Thermal uniformity | 11 $\mu$m | yes | yes | yes | yes |
| Four minus Five (11  12 $\mu$m) | 11 and 12 $\mu$m | yes | yes | yes | yes |
| Thermal Gross Cloud | 11 $\mu$m | yes | yes | yes | yes |
| Channel 3 (3.7 $\mu$m) Restoral | 3.7, 11 and 12 $\mu$m | yes | yes | no | no |
| Thermal Uniformity | 11 $\mu$m | yes | yes | yes | yes |
| Uniform Low Stratus | 3.7, 11 and 12 $\mu$m | | | yes | yes |
| Cirrus | 3.7 and 12 $\mu$m | | | yes | yes |

The inputs required for the operation of the procedure are listed in table 3.

Table 4: Summary of inputs required by the cloud detection process.

| Required inputs | Comments |
|---|---|
| 0.67 and 0.87 $\mu$m reflectance | normalised by cosine solar zenith |
| 3.7 $\mu$m reflectance | Derived from 3.7, 11 and 12 $\mu$m |
| Brightness temperature 3.7 T$\mu$m | |
| Brightness temperature 11 $\mu$m | |
| Brightness temperature 12 $\mu$m | |
| gamma | Sun glint mask (not required for land) |
| Solar Zenith angle | |
| Latitude | Latitude grid for data |
| Day/night mask | Derived from solar zenith angle |
| Land/sea mask | RAL Chris Mutlow or vegetation map |
| Desert mask | For areas of very high albedo |
| Coastline mask | For spatial coherence tests |

In addition to the standard CLAVR algorithm it is recommended that the spatial coherence tests be masked out near land/sea borders.

# Appendix VI: Summary of data files

The following table summarises the various data-sets required for the LST product.

Table 5: Summary of data files used for the LST data product.

| Data file | Description | Units | Dimensions | File size (bytes) | Type |
|---|---|---|---|---|---|
| LSTCoefs.txt | Regression coefficients | N/A | 3 x 14 x 2 x 2 | 2,286 | ASCII |
| PW.climate | Precipitable water | mm x 100 | 720 x 360 x 12 | 6,220,800 | BINARY I*2 |
| *VegFrac.txt*[1] | *Vegetation fraction* | *N/A* | *14 x 12* | *1,744* | *ASCII* |
| Biome.dat | Landtype (biome) | N/A | 720 x 360 | 259,200 | BINARY I*1 |
| Greenness.dat | Vegetation fraction variation with time | x 1000 | 720 x 360 x12 | 6,220,800 | BINARY I*2 |
| TVF.dat | Topographic variance flag | N/A | 720 x 360 | 259,200 | BINARY I*1 |

[1]This file is not required.

Explanations:

- The LST coefficients file contains 3 coefficient values for 14 biomes (no valid data for oceans), 2 surface classes (bare and vegetated), and 2 times (day or night). For biomes 1 to 13 the day and night values are identical. For biome 14 (lakes) the day and night values are different and the coefficients are identical for bare and vegetated classes.
- The biome class (landtype) file contains values of 0 to 13. 0 (zero) is used to signify ocean. Currently biome 14 is not defined in the file.
- The VegFrac.txt file contains the Dorman and Sellars vegetation fractions for each month. **This file is not required in the implementaton of the algorithm because it has already been used to generate the time varyng fractions in the file Greennesas.dat. However, it is included here for completeness.**
- The agreed spatial resolution employed is 0.5° x 0.5°. However it should be noted that the basic resolution of the input ancillary data is no better than 1° x 1° in most cases. Data are provided as 720 longitudes by 360 latitudes and generally are valid for all grid cells. Thus the vegetation fraction is valid over the ocean and has a value of 0. The ocean landtype is 0 which is not a valid land biome. In all cases longitudes start at $-180$ (the cell is defined as -180 to -179.5) and latitudes start at -90 (the cell is -90 to -89.5).
- Precipitable water is in units of mm scaled by a factor of 100. In the algorithm the units employed are centimetres.
- Vegetation fractions are provided as scaled integers in the files (a factor of 1000 is used for the scaling). In the algorithm the fraction is defined in the range 0 to 1.
- Topographic variance is a flag that has been calculated from the variance of the topography at 30 arc-second resolution in each half-degree grid cell. The original specification had this as a 4-bit flag (with values ranging from 0 to 6). The flag has been reduced to 2-bit with valid range of 0–3.

# References


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.

Hook, S. J., Prata, A. J., Alley, R. E., Abtahi, A., Richards, R. C., Schladow, G. S. and Palmarsson, S. O., 2002, Retrieval of Lake Bulk-and Skin-Temperatures using Along Track Scanning Radiometer (ATSR) Data: A Case Study using Lake Tahoe, CA. *Submitted to J. Atmos. Oceanic. Technol.*

Key, J., and M. Haeflager, 1992, Arctic ice surface temperature retrieval from AVHRR thermal channels, J. Geophys. Res., 97(D5), 5,884–5,893.

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.

Parkes, I. M., M. D. Steven, D. Llewelyn-Jones, C. T. Mutlow, C. J. Donlon, J. Foot, F. Prata, I. Grant and T. Nightingale, 1998, AATSR Validation Plan, PO-PL-GAD-AT-005, Version 2.2, Earth Observation Centre, University of Leicester, UK.

Pearce, A.,F. A. J. Prata, and C. R. Manning, 1989, Comparison of NOAA/AVHRR-2 sea surface temperatures with surface measurements in coastal waters, Int. J. Remote Sens., 10(1), 37–52.

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., 1994, 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., 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.

Salsbury, J. W., and D. M. D'Aria, 1992, Emissivity of terrestrial materials in the $8-14$ $\mu$m 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.

Stroeve, J., and K. Steffen, 1998, Variability of AVHRR-derived clear-sky surface temperature over the Greenland ice sheet, J. Appl. Meteorol., 37, 23–31.

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

Stowe, L.L., Davis, P.A., McClain, E.P., 1999, Scientific Basis and Initial Evaluation of the CLAVR-1 Global Clear/Cloud Classification Algorithm for the Advanced Very High Resolution Radiometer. *J. Atmospheric and Oceanic Technology*, **16**, 656-681.