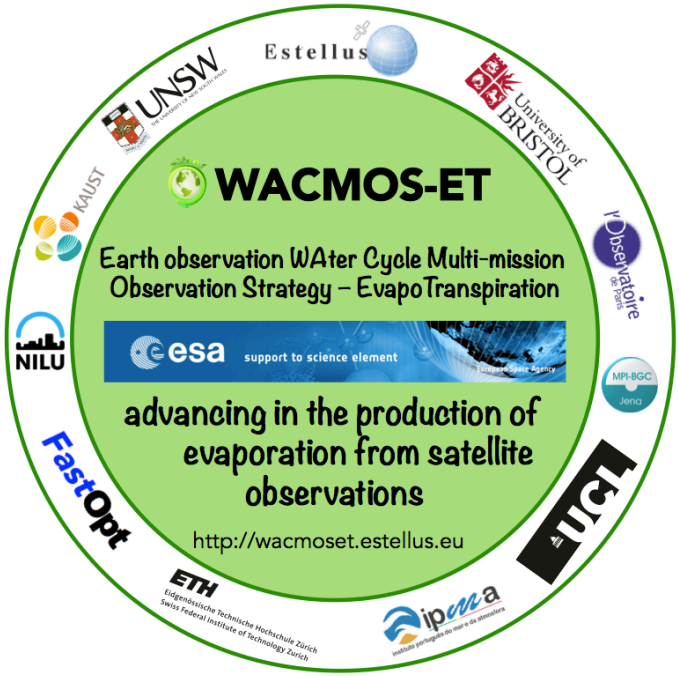



**WP2210-2220**

# Design of the Albedo/fAPAR/LAI Products

**Prepared by:**

Jan-Peter Muller

*Imaging Group, Mullard Space Science Laboratory, Dept. of Space and Climate Physics, University College London, Holmbury St Mary, RH5 6NT, UK*

Philip Lewis, Mat Disney

*NCEO/Remote Sensing Group, Dept. of Geography, University College London, Gower Street, London WC1E 6BT, UK*

8 May 2013

## Contents



# 1 Land surface radiation and vegetation parameters

## 1.1 Requirements

For the generation of regional and global ET, land surface albedo is required for snow and snow-free surfaces as well as vegetation parameters such as fAPAR and LAI, where vegetated land cover is present. An additional core requirement would seem to be the ability to calculate projected canopy cover at arbitrary solar zenith angles to enable consistency with the interpretation of thermal/passive microwave observations. A more subtle expression of this might be to enable the calculation of sunlit and shaded proportions of canopy and soil, but that does not appear to be a core requirement in this case.

## 1.2 Derivation of Albedo: Summary of Algorithm

Earth surface albedo is generally split into two spectral components: visible (VIS) and near infrared (NIR). The main reason for this is that visible albedo is of primary importance in consideration of photosynthesis in vegetation canopies: the radiation absorbed in this region over vegetation canopies, $1 - A_{VIS}$, is in essence partitioned between that absorbed by the canopy (*fAPAR* – the faction of absorbed *PAR*, where *PAR* is so-called 'photosynthetically-active radiation', a misnomer since it is the vegetation that is active in this sense, not the radiation) and that absorbed by soil or non-photosynthetically active components of the canopy. In this argument, we assume *PAR* and the visible waveband to be equivalent. Then we can note that $fAPAR \leq 1 - A_{VIS}$, i.e. $1 - A_{VIS}$ provides an upper bound estimate of $fAPAR$. Another reason for splitting the shortwave albedo into (at least two) spectral components is that the proportion of solar radiation in these wavebands varies with atmospheric conditions (the main factors being aerosols at shorter wavelengths and water vapour at longer wavelengths). This comment gives an immediate insight into some issues that arise in the estimation of albedo: as it depends of atmospheric state (at the very least the proportion of downwelling radiation in VIS and NIR bands) it is not an intrinsic property of the Earth surface. Instead, we can say that it is a function of some intrinsic characterisation of the surface and the illumination conditions (spectral and angular). It is for this reason that we separate 'optimal parameter estimation' from 'albedo estimation' in consideration of the

problem. In attempting to arrive at a useful *intrinsic* surface product *related to* albedo, GCOS (2004) specify 'black-sky albedo' (directional-hemispherical reflectance) as the product required for climate change purposes. This is essentially equivalent to the albedo in the absence of diffuse illumination. It is still however a function, rather than a fixed parameter, as it can vary significantly with solar zenith angle. It is therefore usually computed for a specific time (such as local solar noon) to provide a consistent framework. An obvious issue to arise from that though is that since the sun angle at local solar noon changes throughout the year, this normalized 'black-sky albedo' would apparently change over time, even if the surface underwent no change. Further, 'black-sky albedo' at some particular solar zenith obviously does not *directly* allow the calculation of diurnal shortwave energy budgets and so is of limited value in many applications. The fuller (functional) description of 'black-sky albedo' should then generally be preferred.

A more flexible description of albedo can be provided by a data product that provides estimates of intrinsic surface properties that, with an appropriate radiative transfer model, allow the estimation of spectral directional reflectance (the spectral BRDF). This term, the spectral BRDF, $BRDF_\lambda$ is the fundamental description of surface reflectance, being the ratio of reflected spectral radiance ($Wm^{-2}sr^{-1}nm^{-1}$) exiting around a direction vector $\Omega$ (relative to a surface normal vector) to the spectral irradiance ($Wm^{-2}nm^{-1}$) incident on the surface from direction $\Omega'$ at some wavelength $\lambda$. More practically, we define the spectral BRF, the bidirectional reflectance factor $R_\lambda(\Omega,\Omega')$ (unitless) at some wavelength $\lambda$ or $R_\Lambda(\Omega,\Omega')$ (unitless) the integral of $R_\lambda(\Omega,\Omega')$ over a waveband $\Lambda$, these being the ratio of the BRDF to that of a perfect Lambertian reflectors under the same (directional) illumination conditions. We recall of course that we are primarily interested here only in the (spectral and) angular integral of this term, the 'black-sky albedo':

$$\bar{R}_\Lambda(\Omega_i) = \frac{1}{\pi}\int_{\Lambda_{min}}^{\Lambda_{max}} d\lambda \int_0^{2\pi} d\phi_v \int_0^1 R_\lambda(\Omega_v,\Omega_i)\mu_v d\mu_v$$

A full description of albedo $A_\Lambda(\Omega_s)$ at a solar geometry $\Omega_s$ then can be given by (Roman et al., 2010):

$$A_\Lambda(\Omega_s) = \frac{(1-D_{0\Lambda})\bar{R}_\Lambda(\Omega_s) + D_{0\Lambda}\bar{\bar{R}}'_\Lambda}{1 - D_{0\Lambda}\bar{\bar{\rho}}_\Lambda\left(\bar{\bar{R}}_\Lambda - \bar{\bar{R}}'_\Lambda\right)}$$

where $D_{0\Lambda}$ is the proportion of diffuse illumination (formally, that assuming a totally absorbing lower boundary), $\bar{\bar{\rho}}_\Lambda$ is the atmospheric 'albedo' (meaning the bihemispherical integral of reflectance of the atmosphere over the surface), $\bar{\bar{R}}_\Lambda$ is the bihemispherical integral of surface reflectance:

$$\bar{\bar{R}}_\Lambda = \frac{1}{\pi}\int_{\Lambda_{min}}^{\Lambda_{max}} d\lambda \int_0^{2\pi} d\phi_i \int_0^1 \bar{R}_\lambda(\Omega_i)\mu_i d\mu_i$$

that is sometimes known as the 'white sky albedo' (the integral under perfect isotropic conditions) and $\overline{\overline{R'}}_\Lambda$ is a normalised sky radiance distribution-weighted integral of the 'black sky albedo':

$$\overline{\overline{R'}}_\Lambda = \frac{1}{\pi}\int_{\Lambda_{min}}^{\Lambda_{max}} d\lambda \int_0^{2\pi} d\phi_i \int_0^1 \overline{R}_\lambda(\Omega_i) N_{sky\lambda}(\Omega_i)\mu_i d\mu_i$$

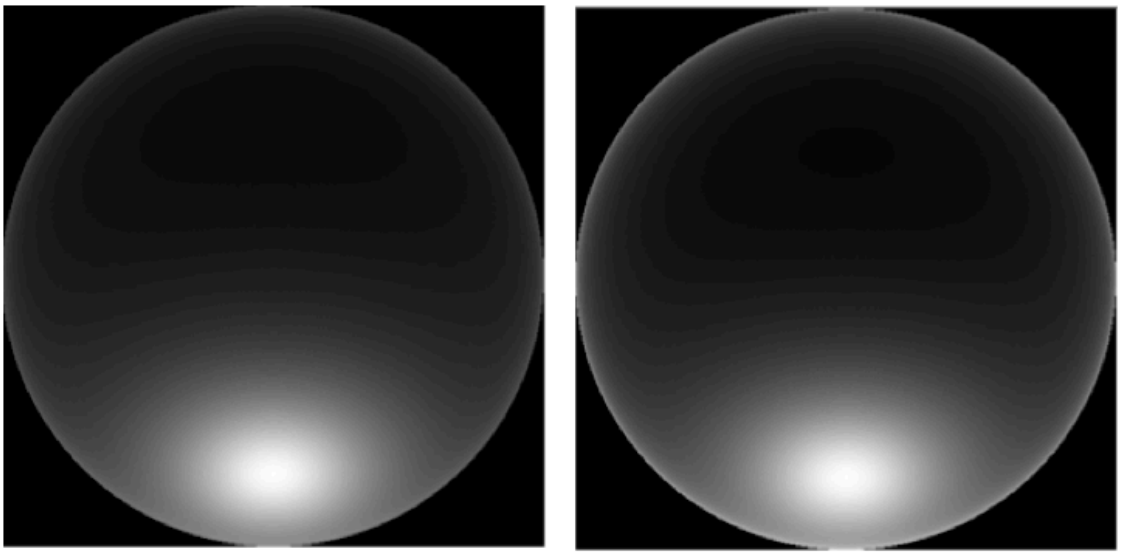
where

$$N_{sky\lambda}(\Omega_i) = \frac{L_{0\lambda}(\Omega_i)}{\eta_\lambda \downarrow(\mu_s)}$$

and $L_{0\lambda}(\Omega_i)$ is the downwelling diffusely transmitted radiance at the bottom of the atmosphere for a totally absorbing lower boundary ($Wm^{-2}sr^{-1}nm^{-1}$) at the bottom of the atmosphere and $\eta_\lambda \downarrow (\mu_s)$ the integral of $L_{0\lambda}(\Omega_i)$ over the sky illumination hemisphere. We can note that in this description of albedo, the denominator term (that accounts for multiple scattering between the surface and the atmosphere) is not directly a function of the surface reflectance but rather the difference between the bihemispherical integral of reflectance and the $N_{sky}$-weighted integral, so would be zero under isotropic illumination. This residual multiple scattering term arises because both the downwelling and upwelling components of reflectance contain multiple scattering, so the albedo, the ratio of these two, nearly cancels out this component. It is typical therefore to ignore this term and to approximate albedo as:

$$A_\Lambda(\vartheta_s) = (1 - D_{0\Lambda})\overline{R}_\Lambda(\Omega_s) + D_{0\Lambda}\overline{\overline{R'}}_\Lambda$$

Assuming the sky radiance distribution is not too far from isotropic, this is quite often further simplified by approximating the $N_{sky}$-weighted integral by the 'white sky albedo' (the bihemispherical integral of reflectance). This can introduce some error at high solar zenith angles, for highly anisotropic surfaces and for turbid atmospheres but is the expression most commonly used to calculate 'blue sky' albedo, i.e. that under particular illumination conditions.



(a) $N_{sky}$ 650 nm (scaled 0.000 to 6.71) (a) $N_{sky}$ 859 nm (scaled 0.000 to 6.33)

**Figure 1. Illustrative Nsky distributions for a solar zenith angle of 45 degrees at two wavelengths.**

Whilst there are several approaches that can be used to describe the BRF to relate sample satellite observations to albedo, one of the most practical is linear kernel-driven BRDF models (Roujean et al., 1992; Wanner et al., 1997). Whilst there are a family of such kernel models, the core idea is to separate the surface scattering terms into a volumetric scattering component, a 'geometric' term describing shadowing and shadow-hiding by surface protrusions (roughness elements) and an isotropic term (that generally is assumed to mimic the multiple scattered component of the surface radiation).

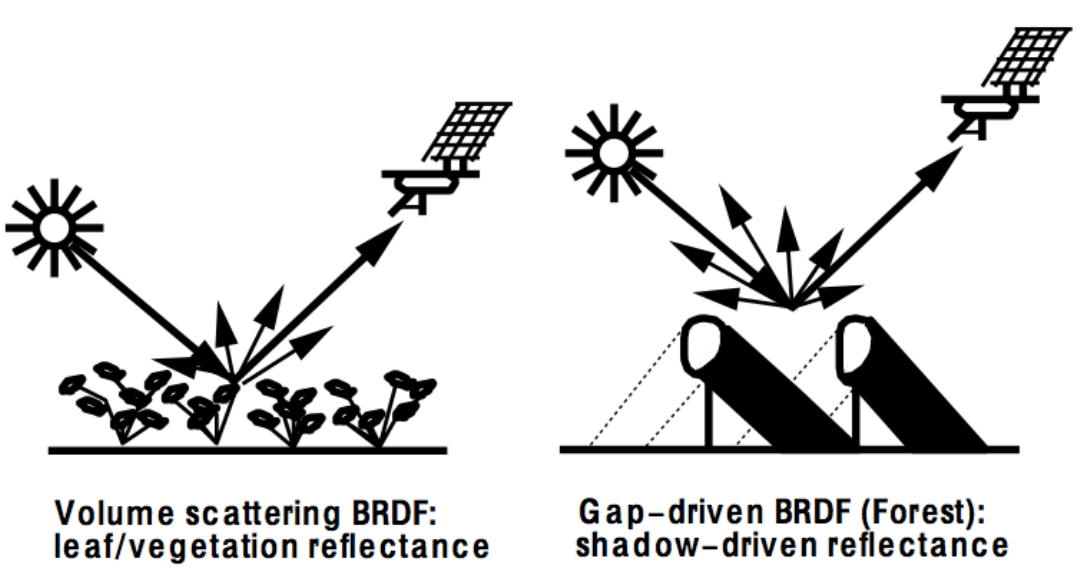



**Figure 2. Assumed scattering components in kernel-driven models (Source: MODIS BRDF/Albedo product ATBD)**

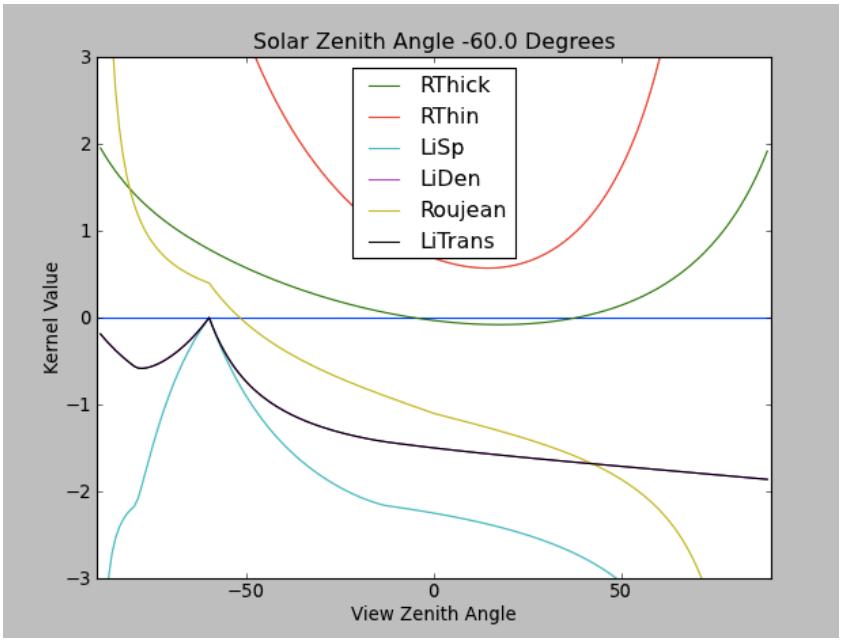



**Figure 3. The main family members of the linear kernel-driven models at a solar zenith angle of 60 degrees**

A linear model of this family can be written as

$$R_\lambda(\Omega,\Omega') = f_{iso\lambda} + f_{vol\lambda}k_{vol}(\Omega,\Omega') + f_{geo\lambda}k_{geo}(\Omega,\Omega')$$

where the model parameters $f_{x\lambda}$, $x = \{iso,vol,geo\}$ are defined as a function of wavelength. The angular kernels $k_x(\Omega,\Omega')$ are functions of viewing and illumination angle only as illustrated above.

These kernel models are considered as 'semi-empirical' in that they start from physically-based principles and are then combined in an empirical manner. Two main versions of the volumetric kernels exist: RossThick and RossThin, being based on slightly different approximations in the linearisation process (optically thick, and optically thin media, respectively). The 'theoretical' basis for these models is a solution for first order scattering from a medium of infinitesimal scatterers (leaves) over a soil background. The leaves are assumed uniformly distributed, with a spherical angular distribution, and of equal reflectance and transmittance $s/2$. The soil is assumed Lambertian, of reflectance $\rho_0$. These assumptions give (Wanner et al., 1997):

$$R_{vol} = \frac{4s}{3\pi}\frac{(\pi/2-\xi)\cos\xi + \sin\xi}{\mu_i + \mu_v}\left(1 - e^{-X}\right) + \rho_0 e^{-X}$$

with

$$X = \frac{LAI}{2}\left(\frac{\mu_i + \mu_v}{\mu_i\mu_v}\right)$$

where $\cos\xi = \cos\vartheta_i \cos\vartheta_v + \sin\vartheta_i \cos\vartheta_v \cos\phi$ is the cosine of the phase angle (for relative azimuth $\phi$). All angles are in radians. If *LAI* is assumed high, $e^{-X} \approx e^{-LAIB}$ where *B* is assumed constant, equal to an average of the zenith angle function $\langle(\sec\vartheta_i + \sec\vartheta_v)/2\rangle$. This step is a linearisation of the model so it can be expressed as:

$$R_{vol} = c_{1\lambda} k_{vol}(\Omega,\Omega') + c_{2\lambda}$$

with

$$\begin{aligned} k_{vol}(\Omega,\Omega') &= \frac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\mu_i + \mu_v} - \frac{\pi}{4} \\ c_{1\lambda} &= \frac{4s}{3\pi}\left(1 - e^{-LAIB}\right) \\ c_{2\lambda} &= \frac{s}{3} + \left(\rho_0 - \frac{s}{3}\right)e^{-LAIB} \end{aligned}$$

Note that the angular kernel $k_{vol}(\Omega,\Omega')$ is designed to equal zero at nadir viewing and illumination (following (Roujean et al., 1992)). Implementation of this kernel is aided by taking absolute values of view zenith angles whilst adding $\pi$ to the relative azimuth where the view zenith was negative. Relative azimuth angles of zero, with the same viewing and illumination zenith angles, define the hot spot direction.

(Breon et al., 2002; Maignan et al., 2004) criticised the Ross kernel models because although they are created to mimic first order scattering, they ignore enhanced reflectance in the backscatter (hot spot direction). They suggest a 'hot spot' model for these kernels, which for RossThick[1] gives:

$$k'_{vol}(\Omega,\Omega') = \frac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\mu_i + \mu_v}\left(1 + \frac{1}{1 + \xi/\xi_0}\right)$$

where the parameter $\xi_0$ controls the angular width of the hot spot feature. For low phase angles $(\xi \approx 0)$, the enhancement factor is equal to 2. At $\xi = \xi_0$ the enhancement factor is 1.5, and for $\xi >> \xi_0$ it tends to unity. There are two main arguments for the inclusion of this modification: (i) the hot spot is a real feature of first order volumetric scattering that is otherwise ignored; (ii) Observations from POLDER over a wide range of cover types (Breon et al., 2002) suggest that the parameter $\xi_0$ might reasonably be fixed to a value equating to $1.5^o$. The arguments for not including this term are: (i) most of the sensors we are dealing with very seldom have sampling near to the hot spot; (ii) any missing hot spot effects will likely be compensated for by the geometric kernel (see below); (iii) multiple scattering is assumed isotropic in these models (see below) but any failings in this assumption are likely to be compensated for in the volumetric kernel. Including a strong hot spot effect based on considerations of first order scattering might therefore over-emphasise the feature, especially when poorly constrained by sampling; (iv) The hot spot feature at the angular widths of interest here has little impact on angular integrals (although it does have an impact on angular normalisation); (v) the MODIS processing chain does not use this modification, so if we

[1] Note that the normalisation term is left out here.

include it here we will not be working with the same kernel set as that sensor. In fact, none of these reasons for not using this enhancement are very strong, but the latter, the fact that the kernel parameters would not be consistent with MODIS, has a practical impact for the optimal estimation framework used here, so we will not include this term.

At surface directional reflectances (SDRs) usually consist of BRFs (Bi-directional reflectance Factors) can only be retrieved when not obscured by clouds. For areas of persistent cloud cover, such as tropical forests or areas where it is extremely difficult to differentiate between clouds and underlying bright reflectances such as permanent snow and ice, on the basis of reflectance and/or emissive properties alone, these BRF cloudy scenes usually appear as gaps in the final BRDF/albedo record. To address this, (Moody et al., 2007; Moody et al., 2008) developed an empirical method for filling-in gaps for the MODIS product using vegetation phenology. However, to the present day, the resultant interpolated values remain unvalidated.

For MISR, individual orbital strips of only 380km swath-width contain only few areas where there are retrievals due to the need to simultaneously retrieve aerosols at 17.6km (due to frequent cloud cover) and land surface BRDFs at 1.1km. Compositing over a 16-day time period (the repeat of the EOS-Terra satellite) is another possible solution but MISR adopted a monthly binning at 0.5 x 0.5 degrees using the statistical methods proposed by (Braverman and Girolamo, 2002) which appears to contain sufficient retrievals for most parts of the Earth's land surface apart from the aforementioned tropical forests and snow/ice regions. MISR surface BRDF/albedo retrievals (Diner et al., 2005a) include estimates of uncertainties for BHR (so-called white-sky retrievals). Unfortunately, there has been only very limited validation to date of MISR albedos with validation mainly focused on inter-comparisons with MODIS (Pinty et al., 2011b; Sun et al., 2006; Taberner et al., 2010). European sensors are mostly not designed for obtaining global BRDF due to their narrow swath-width for medium resolution sensors (512km for AATSR, 1100km for MERIS) or their low resolution (≈6.67km) for capturing variability associated with the terrestrial land surface (Maignan et al., 2004). Fortunately, the SPOT-VEGETATION sensor fills this gap with 2250km swath-width and daily coverage at 1km albeit with only limited broader spectral band coverage

It is known that the information content of the European sensors that are considered here (MERIS and VEGETATION) have significantly less information content (in their ability to constrain estimates of albedo) than the data used in some other products, notably the MODIS BRDF/albedo product. This is mainly because of spectral sampling issues (particularly for MERIS beyond 1μm), angular sampling (particularly MERIS), and viewing opportunities (comparing those of VEGETATION with two MODIS sensors). The best scientific dataset then would most likely be obtained from a combination of data from all of these sensors, including MODIS, provided there were no significant calibration biases between the instruments. This is however not feasible within the WACMOS-ET project, and there is a further desire to generate a product that is not dependent on non-European data (e.g. to safeguard future production). Therefore non-European data is only used in the generation of a one-off ancillary datasets for this product. In particular, MODIS data is used to provide a prior estimate of model parameters $F_p$ and associated uncertainty information $C_p^{-1}$.

There are many significant advantages to using a 'prior' in the processing. In particular, it can provide a constraint which allows a solution to be calculated, even when sampling is weak. This is important as it obviates the need for a 'backup algorithm': there is instead a smooth, statistically-based blending of information from the prior estimate and the new observations, the quality of which can be assessed through a consistent uncertainty measure.

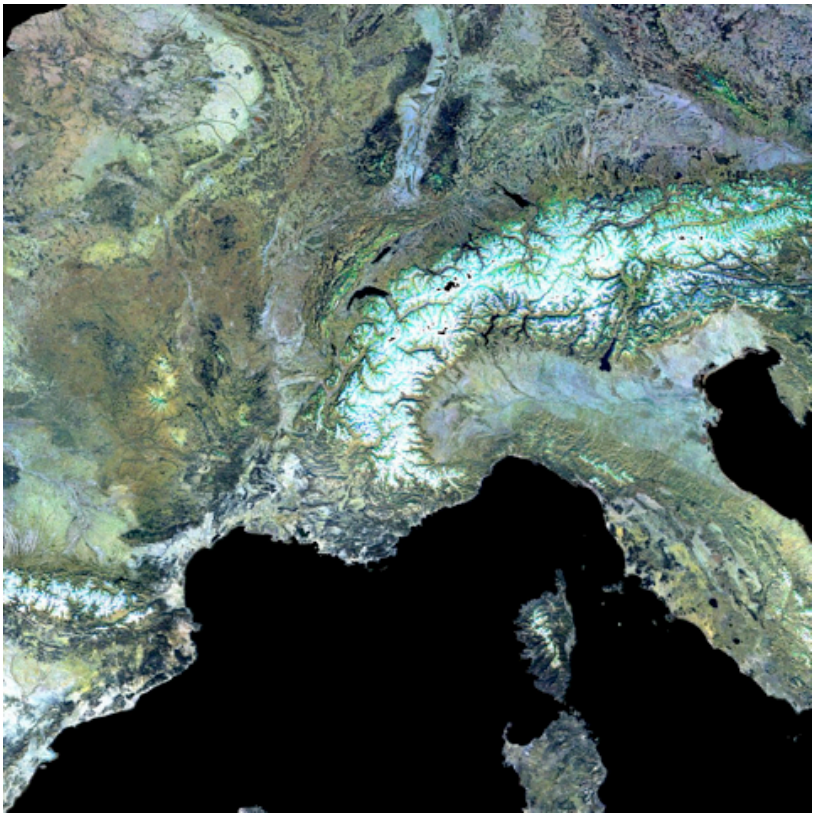


**Figure 4. MODIS climatology VIS 'snow free' bihemispherical reflectance for DOY (009,041,105 on RGB) scaled [0.0160:0.128]**

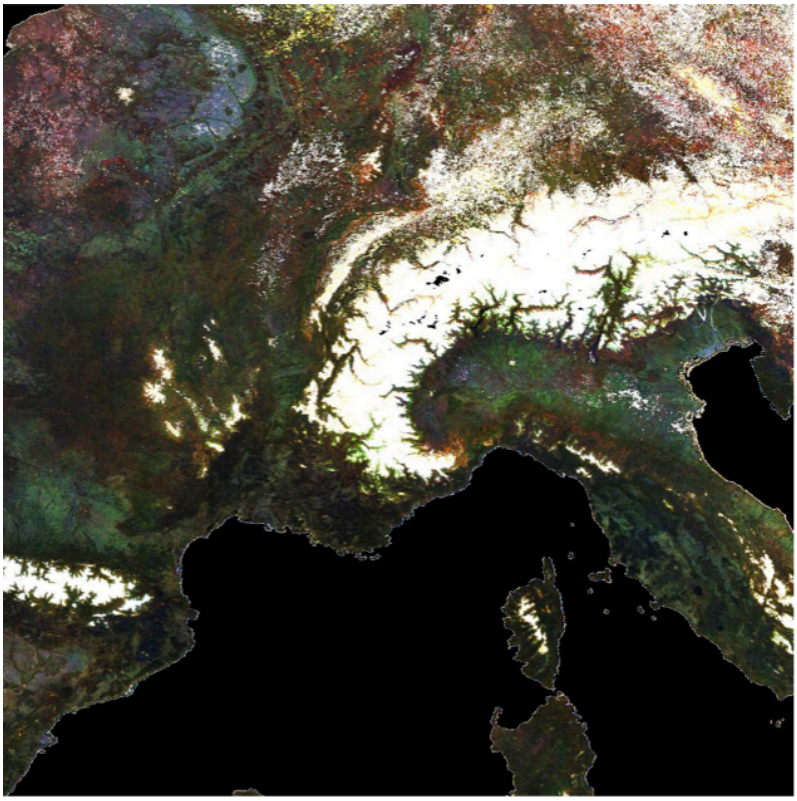


**Figure 5. MODIS climatology uncertainty VIS 'snow free' bihemispherical reflectance for DOY (009,041,105 on RGB) scaled [0.000:0.065]**

In the case of a small number of observations (when the estimate is strongly reliant on the prior estimate) the uncertainty should be relatively high, i.e. we should only use an external prior as this as a relatively weak constraint. A strong prior constraint would mean that the resulting product would be very dependent on the quality of the prior as a good estimate of model parameters, and therefore sensitive to any errors or biases they contain. It also makes it sensitive to the particular way in which the MODIS product is calculated (e.g. 16 day moving window).

That said, there are now more than 14 years of global MODIS BRDF/albedo data, at 500 m spatial resolution, every 8 days. There is clearly a lot to learn from these data, and one way we have done this is to have developed for the first decade a spatially and temporally complete prior parameter estimate from them.

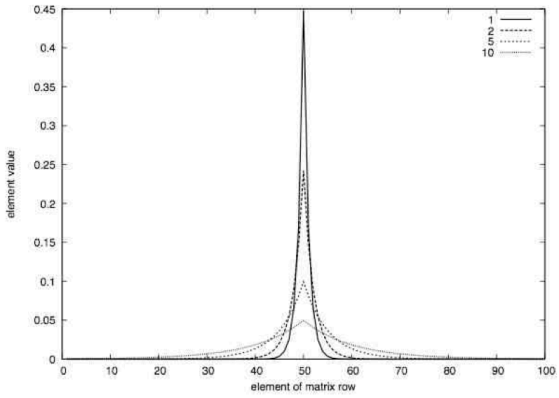



**Figure 6. Temporal weighting functions**

A further constraint that can be applied is a temporal weighting on observation uncertainty that mimics a form of temporal regularisation and gets around the idea of fixed (e.g. 16 day) sample windows used in many products. This means that samples from e.g. 3 months prior to the target time period have an (albeit small) influence of the albedo. For consistency with e.g. 8 day sampling, the width of the smoothing filter can be set to have a value of 0.5 at +/- 8 days, i.e. the contribution of samples 8 days distant to the target time period are half of any samples on the actual day. A 'peaky' weighting function (a Laplace distribution function $e^{-|t-t_0|/11.54}$ is used to mimic a first-order difference constraint. So at 16 days distant to the target, the weighting is 0.25, at 32 days it is 0.0625, and at 96 days (around 3 months) it is 0.00024.

For many satellite products, only an overall product quality and local QA information are provided. Recently, reasonably in depth validation has been performed of the MODIS broadband albedo product based on accounting for the heterogeneity of the tower albedometer sites for a subset of 53 sites within FLUXNET (Cescatti et al., 2012) based on previous work (Salomon et al., 2006). However, it is more appropriate to provide users with a per-pixel estimate of uncertainty. This is difficult to achieve for bias, but straightforward in an optimal estimation framework to propagate uncertainty from the atmospheric correction through to the final product, particularly if linear models are used.

Within the ESA DUE GlobAlbedo project (Muller et al., 2011, 2012b; Lewis et al., 2011) an optimal estimation approach has been developed for combining a background field of snow and snow-free BRDF derived from a decade of MODIS BRDF measurements together with angular sampling from VEGETATION and spectral sampling from MERIS and the regularisation constraint described above.

The optimal estimation processing chain is shown schematically in Figure 9. Satellite images, comprising DN data in a spatial array in the sensor swath coordinate system, along with ancillary information (on the atmospheric state and related information such as surface pressure/altitude) are ingested into a pre-processing step. All sensor information is treated the same way here, although minor details such as the specific LUTs and thresholds may vary between sensors.

The output of the pre-processing stage is a set of BBDR data in three spectral channels (VIS, NIR, SW) on a SIN grid (see below), along with associated uncertainty and ancillary information (primarily, weighted linear kernel integrals here), as well as pixel classification information (the primary items of interest being snow/no snow and land flags).

Sets of these data over some time window are then fed into an optimal estimation framework, along with a **prior** estimate of the model parameters. This produces a set of kernel model parameter estimates (with associated uncertainty) from which albedo may be estimated under any atmospheric conditions by the determination of appropriately weighted integrals. The model parameters are fed into a post-processing step, which summarises the information into black sky albedo (directional-hemispherical integral reflectance) (at local solar noon) and white sky albedo (bi-hemispherical integral of reflectance).

## 1.3 fAPAR

During the post-processing phase, fAPAR is estimated from the white sky albedo data and other required datasets are passed through to the output product, although the fAPAR product currently generated does not fully take into account the uncertainties in the albedo. In fact, in GlobAlbedo a simple LUT is applied that maps bi-hemispherical reflectance in the visible and near infrared bands to fAPAR. This is seen as a 'test' product in GlobAlbedo as it has not been a major focus of development. It is of value in that it is simple and traceable to publications, but should be interpreted with caution.

Pinty et al. (2006) present a two-stream model that models diffuse fluxes in a layered medium over a Lambertian lower boundary. The parameters of this model are: LAI, leaf single scattering albedo, leaf scattering asymmetry, soil reflectance. So, per waveband, there are three spectral terms and one structural. A further parameter, zeta, can be used to account for clumping, but is ignored in this present discussion, meaning that LAI is interpreted as effective LAI. In various papers following this (e.g. Pinty et al., 2010a,b; Pinty et al. 2011), Pinty and colleagues apply a parameter estimation package (JRC-TIP) that provides an estimate of the model state variables (7 terms if two wavebands are used) for input bihemispherical reflectance in visible and near infrared broad wavebands.

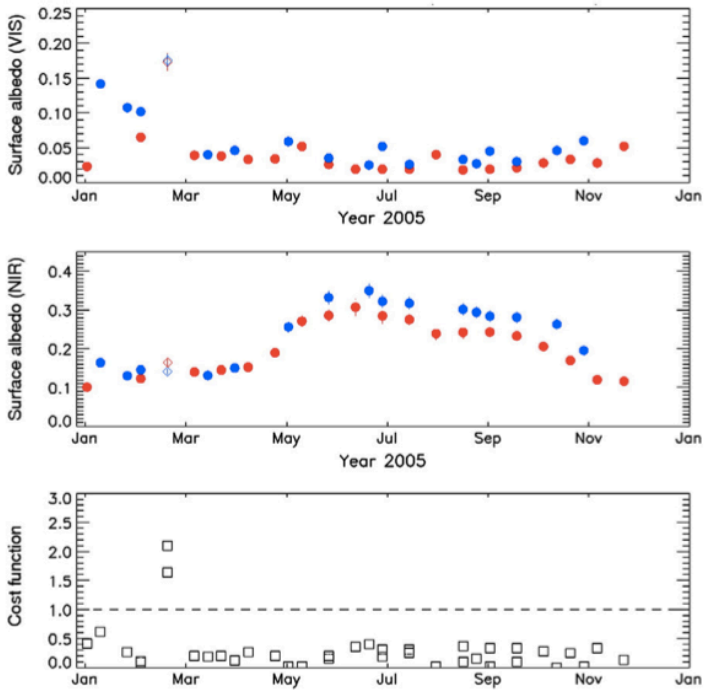



**Fig. 2.** Time series of the broadband BHR–white sky–surface albedo at 0.01° spatial resolution in the visible (top panel) and near-infrared (middle panel) domain over the site of Hainich (51° 05′ 05″ N; 10° 26′ 41″ E). MODIS (MISR) derived albedo values are featured in red (blue). High cost function values (bottom panel) are identified with an open diamond symbol. The vertical bars display the one-sigma posterior uncertainty values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Figure 7. Figure 2 from Pinty et al. 2011 illustrating input albedo data for a forest site.**

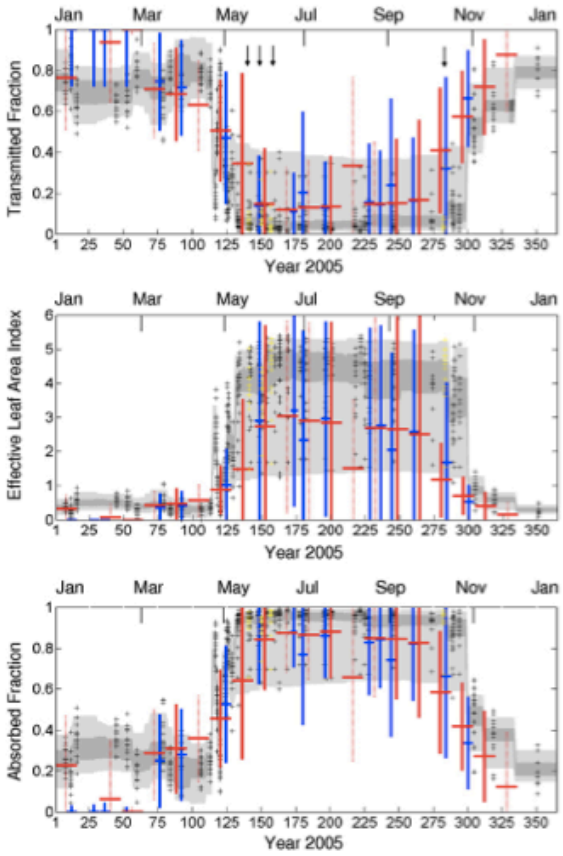



**Fig. 3.** Top panel: Time series of the domain-averaged fractions of diffuse transmission in the visible spectral domain measured over the Hainich site. The grey shaded zones indicate the range (light grey) and interquartile (dark grey) range estimated both from the individual sampled points over all seven years (black crosses). The points available in 2005 at dates indicated by the downward arrows are shown with yellow crosses. The JRC-TIP derived estimates from MODIS (MISR) are shown in red (blue). The full (dashed) vertical bars indicate the one-sigma posterior uncertainty associated with the best (good) quality MODIS input albedos. Middle and bottom panels: same as above except for the effective LAI and the fraction of absorbed flux in the vegetation layer, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Figure 8. Figure 3 from Pinty et al. 2011 illustrating derived quantities for a forest site.**

| **Variable** | $\mathbf{X}_{prior}$ | $\sigma_{\mathbf{X}_{\mathbf{prior}}}$ |
|---|---|---|
| LAI | 1.5000 | 5.0 |
| $\omega_l(\lambda_1)$ | 0.1700 | 0.1200 |
| $d_l(\lambda_1)$ | 1.0000 | 0.7000 |
| $r_g(\lambda_1)$ | 0.1000 [1] and 0.50 [2] | 0.0959 [1] and 0.346 [2] |
| $\omega_l(\lambda_2)$ | 0.7000 | 0.1500 |
| $d_l(\lambda_2)$ | 2.0000 | 1.5000 |
| $r_g(\lambda_2)$ | 0.1800 [1] and 0.350 [2] | 0.2000 [1] and 0.25 [2] |

[1(2)] Values adopted for the bare soil (snow) case with a correlation factor of 0.8862 (0.8670) set in $\mathbf{C}_{X_{prior}}$.

**Table 1. Prior constraints used by Pinty et al.,**

By using a set of prior constraints on the state variables (items (1) in table 1), a solution to these 7 terms can be derived for given VIS and NIR white sky albedo then, for a given uncertainty. The cost function involved is a weighting between the prior cost and the observation cost, this weighting being essentially defined by the relative uncertainty of the

observations and the priors. Once these state variables are defined, derived radiative quantities can also be output, such as the transmitted or absorbed proportion of radiation for each waveband. In this way, fAPAR, interpreted as the radiation absorbed in the canopy in the VIS band, can be estimated.

## 2 Derivation of LAI and fAPAR: summary of algorithm

We do not propose implementing a new development of LAI/fAPAR retrieval, and will rely on the published method based on the Two-Stream Inversion Package (TIP) developed explicitly for driving these parameters for large-scale gridded Earth System models (ESMs). TIP has been developed by JRC jointly in conjunction with project partners FastOpt (http://www.fastopt.com). The development and testing publications are listed in http://www.fastopt.com/products/tip/tip.html. Below we provide a brief summary of this method.

Here, LAI is defined as the one-sided leaf area per unit ground area; fAPAR is the fraction of absorbed photosynthetically active radiation (PAR), defined as the 400-700 nm region.

For regional-to-global applications, EO-derived LAI/fAPAR parameter retrieval has two key requirements:

I. Rapid enough to be applied to global datasets (i.e. requirements for simple, robust retrievals with 'sufficient' physical detail);
II. Consistency between derived parameters and the representation of these parameters in downstream ESM applications.

### 2.1 Two-stream Inversion Package (TIP)

#### 2.1.1 Two-stream model

The TIP is a simplified inversion scheme explicitly developed for rapid, robust LAI/fAPAR parameter retrieval from observations of albedo, particularly at large spatial scales. In particular, the underlying radiation transfer (RT) model approach was designed to make use of so-called 'effective' parameters i.e. parameters that provide the correct radiation properties within the model, given the model assumptions, but parameters that would not be observed if they were (or could be) 'measured' in some way at the same scale as the model is representing them (Pinty et al., 2004; 2006). Large-scale ESMs almost always depend on a two-stream approximation to represent radiation fluxes at the surface; this in turn depends on a simplified 1-D approximation of surface canopy structural properties to be used in place of a more realistic 3-D representation. The approach of Pinty et al. (2006) is to accept this simplification, and use it as the basis of parameter retrieval, as long as the resulting parameters are then recognised as being the 'effective' 1-D equivalents required to provide a correct, consistent RT representation solution for what is an inherently 3-D canopy system.

The approach proposed by Pinty et al. (2006) uses a standard two-stream approach allows decomposition of the directional hemispherical reflectance (DHR, black sky, collimated albedo) field emerging from the top of a canopy layer at depth $z_{toc}$ into three terms:

$$R^{total}_{coupled}(z_{toc},\mu_0) = R^{Coll}_{veg}(z_{toc},\mu_0) + R^{UnColl}_{bgd}(z_{toc},\mu_0) + R^{Coll}_{bgd}(z_{toc},\mu_0)$$

**Equation 1**

where $\mu_0$ is the solar zenith angle. The terms on the RHS represent, respectively: radiation that has interacted with the vegetation canopy elements only; that interacting with the soil only i.e. travelling both down and upwards through gaps in the canopy; and that interacting multiple times with both soil and background. The major advantage of this approach (over a more detailed 3-D approach for example) is that it can be solved analytically. The only additional requirement is the introduction of one additional parameter, a structural term representing the effects of spatial variation of leaf density. The use of effective variables guarantees the correct simulation of the scattered, transmitted and absorbed radiant fluxes. Pinty et al. (2006) derive values of the effective variables by inverting a 1-D radiant flux model against fluxes generated by a 3-D model. This guarantees accurate simulations of the three radiant fluxes when using, in direct mode, these effective values. In addition it does not require explicit consideration of other canopy elements (woody material) as these are considered implicitly in the diffuse (multiple scattered) fluxes. The resulting approach has been shown to agree very well with 3-D model simulations across a wide range of structural configurations (Pinty et al., 2006).

### 2.1.2 Two-stream model inversion

Inversion of the two-stream model is described in Pinty et al. (2007). In brief, the process is a standard inversion via minimization of a cost function representing an a priori optimal set of model state parameters, $\boldsymbol{X}$, that minimizes the difference between a set of observations $\boldsymbol{d}$ of the fluxes predicted by the two-stream radiative transfer model, $M(\boldsymbol{X})$. The parameters $\boldsymbol{X}$ required by the model are LAI, leaf single-scattering albedo $\omega_l(\lambda_{1,2})$, background albedo $r_g(\lambda_{1,2})$ and $d_l(\lambda_{1,2})$ for the visible and near-infrared spectral domains respectively. By solving for the scattered and transmitted radiant fluxes, the absorbed component (fAPAR), or $A_{veg}^{Coll}(z_{bgd}, \mu_0)$ is given by considering energy closure from Equation 1 in the case of a black (totally absorbing) background i.e. $A_{veg}^{Coll} = 1 - \left(R_{veg}^{Coll} + T_{veg}^{Coll}\right)$, where $T_{veg}^{Coll}$ is simply the component transmitted by the vegetation through multiple interactions.

The model inversion problem is simplified by assumptions of observation and model parameter probability distribution as Gaussian, as well as a locally-linear approximation of $M(\boldsymbol{X})$. In this case, the posterior distribution of model parameters can be represented as

$$P(\boldsymbol{X}) \approx exp\left(-\frac{1}{2}\left(\boldsymbol{X}-\boldsymbol{X}_{post}\right)^T \boldsymbol{C}_{\boldsymbol{X}_{post}}^{-1}\left(\boldsymbol{X}-\boldsymbol{X}_{post}\right)\right)$$

**Equation 2**

where $^T$ represents the transpose operator; $\boldsymbol{X}_{post}$ and $\boldsymbol{C}_{\boldsymbol{X}_{post}}$ are the mean and covariance matrix of the maximum likelihood estimator of $P(\boldsymbol{X})$; $\boldsymbol{C}_{\boldsymbol{X}_{post}}^{-1}$ is the inverse of $\boldsymbol{C}_{\boldsymbol{X}_{post}}$. The $\boldsymbol{X}_{post}$ minimises a cost function $J(\boldsymbol{X})$ as

$$J(\boldsymbol{X}) = -\frac{1}{2}\left[(M(\boldsymbol{X})-\boldsymbol{d})^T \boldsymbol{C}_{\boldsymbol{d}}^{-1}(M(\boldsymbol{X})-\boldsymbol{d}) + \left(\boldsymbol{X}-\boldsymbol{X}_{prior}\right)^T \boldsymbol{C}_{\boldsymbol{X}_{prior}}^{-1}\left(\boldsymbol{X}-\boldsymbol{X}_{prior}\right)\right]$$

**Equation 3**

where the first term on the right-hand side quantifies the mismatch between the model predicted values and the observations, and the second term represents the model constraint represented by the a priori estimate of model parameters, $\boldsymbol{X}$. In practice Equation 3 is solved

using gradient descent via automated differentiation methods to calculate the adjoint of $J(\boldsymbol{X})$, as developed by FastOpt.

Pinty et al. (2007) demonstrate the utility of the TIP in retrieving LAI and fAPAR from Moderate Resolution Imaging Spectroradiometer (MODIS) and Multiangle Imaging Spectroradiometer (MISR) broadband surface albedo products. More recently, Pinty et al. (2011, b) have demonstrated the performance of the TIP in deriving surface parameters on global scales from MODIS albedo data. In particular, they show the benefits of the TIP scheme for the generation of physically consistent information about the density and absorbing properties of the vegetation layer together with the brightness of the background underneath. LAI values follow expected trajectories of seasonal and spatial distributions. The estimated uncertainties on LAI are strongly correlated with the LAI itself and to a lesser extent to the brightness of the underlying background. In addition, the TIP allows control over correlations between values of state variables of radiation transfer in the visible and near-infrared spectral ranges that are typically implicit in more approximate retrieval schemes based on empirical spectral indices.

In this project, the LAI and fAPAR values will be derived in the same way as outlined above, using the GLOBALBEDO observations in the visible and NIR.

## 3 Globalbedo processing

More details concerning the pre-processing stage are illustrated in Figure 10. Here we see that there are five main processes involved:

1. pixel identification
2. aerosol concentration estimation
3. atmospheric correction
4. spectral integration of reflectance
5. binning/gridding

The first of these, pixel identification, involves taking the raw DN value and attempting to assign a classification to the pixel (e.g. cloud). This is covered in detail in the document (Lewis et al., 2011) and outlined briefly below. The second stage involves attempting to estimate the aerosol optical thickness from the data and supplementary datasets. This is detailed in (Lewis et al., 2011) and described briefly below. The final stages of processing involve applying the atmospheric correction to the data identified as clear of cloud in the classifier to give spectral directional reflectance factors (Lambertian equivalent). These are then converted into estimates of broadband directional reflectance, and finally each pixel of the input swath data is assigned to processing grid cell (SIN projection grid). These stages of pre-processing are dealt with in detail in the document (Lewis et al., 2011). All information on the optimal estimation framework and post-processing to albedo (and related information) are given in (Lewis et al., 2011).

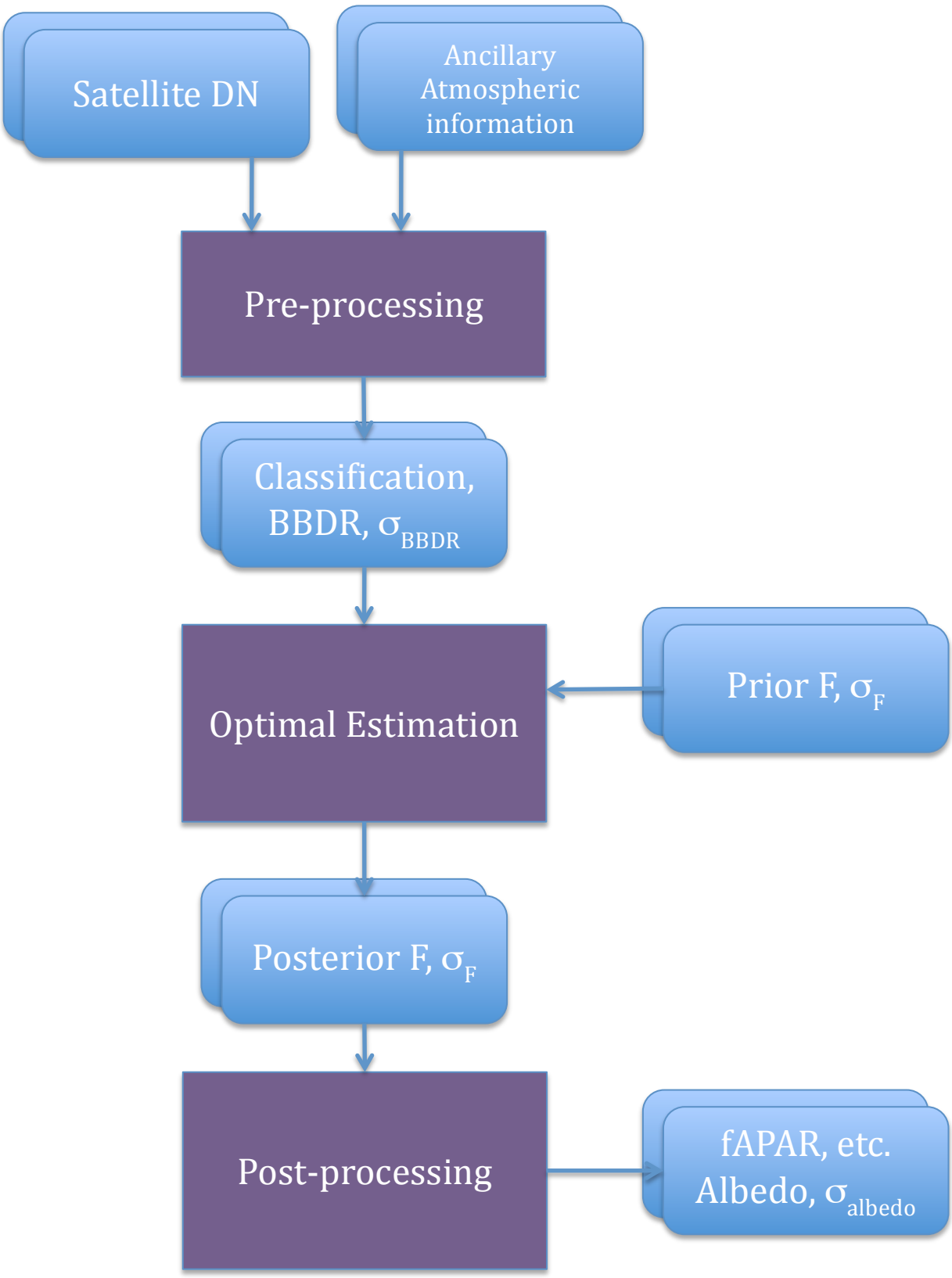



**Figure 9. GlobAlbedo Overall processing chain**

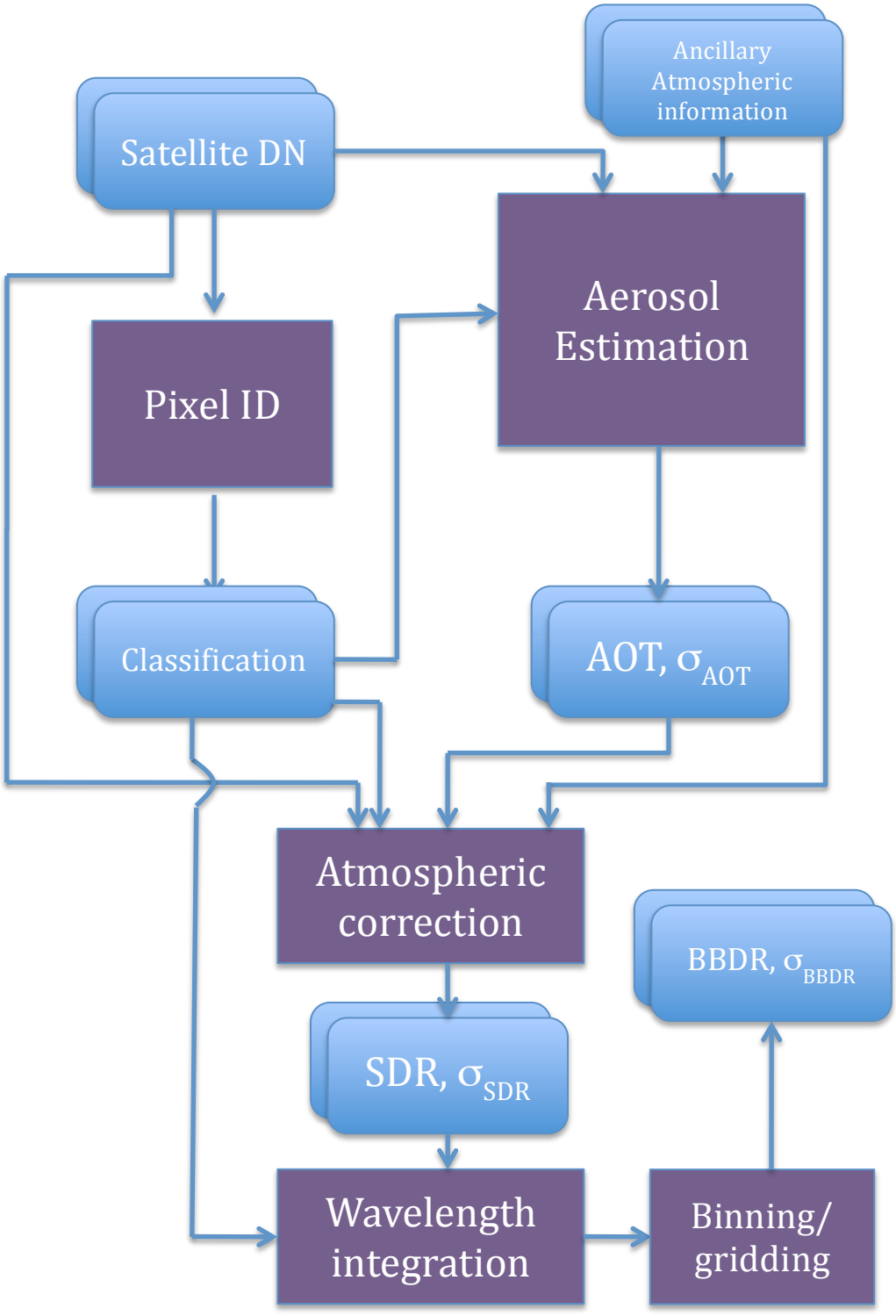



**Figure 10. GlobAlbedo Pre-processing Steps**

This results in a 1km product containing DHR (black sky) and BHR (white-sky) albedos in 3 broadband regions (0.3-0.7µm, 0.7-3µm and 0.3-3µm) along with uncertainty estimates for each pixel, an estimate of the contribution of the MODIS prior through relative entropy and the number of samples. An example of the resultant BHR, uncertainty (expressed as the coefficient of variation of standard deviation divided by mean) and number of samples is given in Figure 11. Clearly, interim products of (snow and snow-free) kernel parameters are also produced.

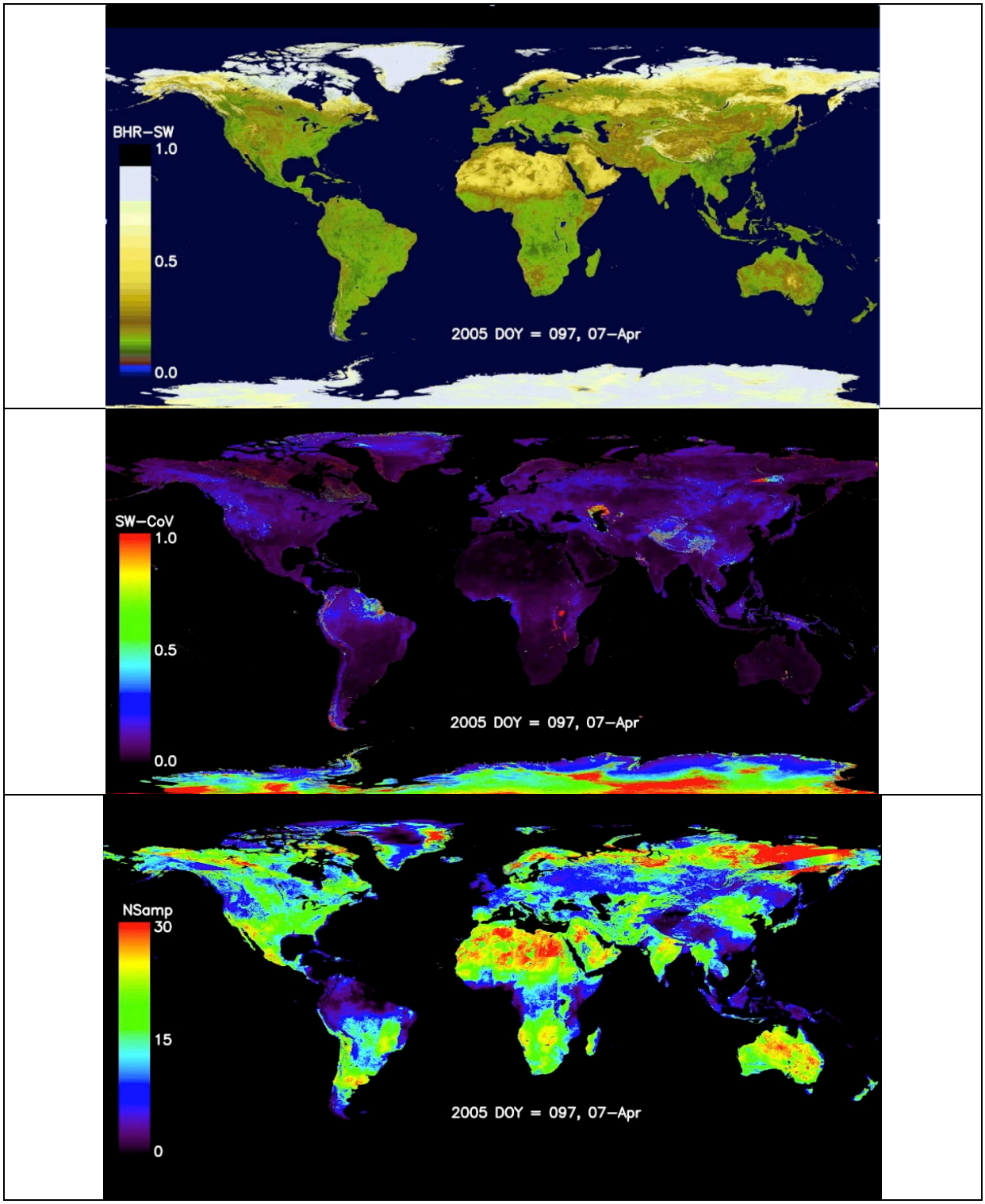



**Figure 11. GlobAlbedo example products for global 0.05º fields of (uppermost) BHR-SW for DoY 97 (7-April-2005); (middle) Coefficient of Variation for the same time period; (c) Weighted number of samples for the same time period.**

Geostationary satellites, such as METEOSAT also provide broadband shortwave albedos, (Govaerts et al., 2008) albeit on a random day within a 10-day time period and only for the

region from ±55º of latitude with restricted longitudinal cover and at low resolution (≈5km at sub-satellite point). We plan to use the aforementioned scheme to calculate an effective LAI. Soil background albedo may also be calculated using the approach described originally by (Houldcroft et al., 2009) and more recently by (Pinty et al., 2011a).

To meet the requirements for global products (10-25km grids) and regional products for Europe, Australia, US and Africa (1km grids), we plan to re-use the BRDF products as well as, if necessary the Broadband Directional Reflectances (BBDRs) processed during GlobAlbedo to obviate the need to repeat this processing. Each year's processing (from BBDRs to BRD/albedos) takes around 2-3 weeks of continuous CPU utilisation on a 224-core/14-blade system.

FAPAR and LAI products will be acquired from the GEOLAND product and compared against those derived from GlobAlbedo broadband albedo using the LUT approach proposed by (Pinty et al., 2011a). The most significant issue in the GlobAlbedo retrievals is the cloud pixel identification but little further can be achieved with the GlobAlbedo datasets as both radiance-based, multispectral and temporal sampling have already been applied to eliminate virtually all cloudy pixels. The poorest retrievals in GlobAlbedo occur over permanent snow/ice, such as Greenland and Antarctica. These poor retrievals are due to distinguishing cloud from snow/ice. The inadequacy of the BRDF for snow/ice compared to their excellent performance for vegetation & deserts appear to have little impact for albedo retrievals, at least for MODIS (Stroeve et al., 2005). However, these areas are not included in the regional datasets and become less of an issue at the 10-25km global resolution. The next poorest retrievals occur over regions with recent snowfall with snow below trees or melting snow on upper storey canopies. However, the need for temporal compositing means that it is very unlikely that further improvements can be made here either. An exploration will be performed of the BBDR datasets from VEGETATION and MERIS to see if further progress can be made as well as through the inversion process to try to flag these poor retrievals better. A more detailed analysis will also be performed of the accuracy of the internal uncertainty metrics employed to ensure that weighting factors have been properly applied in the pre-processing chain.

The full BRDF model parameters will be provided for each of the three wavebands (0.3-0.7, 0.7-3 and 0.3-3μm) along with the full uncertainty matrix. This will allows downstream BHR and DHR albedo products to be generated for data assimilation for any given input solar angle. More complete uncertainty information will be provided in the product for use by FastOpt in their TIP retrievals. If regional or global aerosol information is provided on a daily basis, the actual (so-called) Blue-Sky albedo can also be calculated. We will employ outputs from LUT described by (Pinty et al., 2007). Critical to this will be an assessment of the impact of snow, especially ephemeral snow under forest given that snow appears when there are cloudy conditions and depending on how much snow-melt takes place will define how much of the background snow is visible in comparison with upper canopy snow. Each BBDR indicates whether a pixel is snow-free or not and the temporal changes in this are currently not utilised.

The current MODIS priors employed in the GlobAlbedo project was based on the first 10 years of MODIS BRDF data every 8 days (from Collection 5). Work is underway in WACMOS-ET to re-calculate this MODIS prior every 8 days with the full 14 years of MODIS BRDF data now available.

The output of the WACMOS-ET product will consist of the following:

1. BroadBand BRDF parameters (isotropic, volumetric, geometric-optical) with associated uncertainty matrix for each parameter and each waveband region. This will permit a BHR or DHR albedo to be calculated for any solar angle and provided that the Direct-to-Diffuse ratio is measured (only available for limited FLUXNET or 2 BSRN sites in Europe or AERONET sites worldwide), a blue-sky albedo can be calculated for every solar angle for every view angle specified
2. FAPAR and effective LAI based on the LUT defined by radiative transfer simulations from FastOpt including estimated uncertainties for each quantity

Extensive validation has been performed of blue sky albedo within GlobAlbedo from FLUXNET tower measurements (Muller et al., 2012a) as well as for BHR between GlobAlbedo, MODIS, MISR and METEOSAT First Generation (MSA) (loc.cit.). An example of such an evaluation is shown for Fort Peck (a US NOAA SURFRAD site) in Figure 12. Note here the differences between the tower measurements and satellite derivations for a small footprint size (hence seeing more of the snow beneath the trees cf km pixel) in the winter months. Recent work (Schaaf & Wang, private communication, 2013) have shown that if sampling takes place on a daily basis that such large differences disappear as the ephemeral nature of the snow is much better represented. Unfortunately the time-scale of the MODIS Collection 6 processing (loc.cit.) does not meet the timescales of the WACMOS-ET project so it will NOT be possible to generate a new daily prior and a subsequent daily albedo, fAPAR & LAI field.

In WACMOS-ET, we plan to evaluate the temporal consistency of albedo measurements using the FLUXNET towers based on the BRDF parameters. For FAPAR and LAI, inter-comparisons will be made with the GlobAlbedo product and that generated from GEOLAND2 using the triple collocation technique (Stoffelen, 1998) (Muller et al., 2012a) to gain a better understanding of the biases in the different approaches.

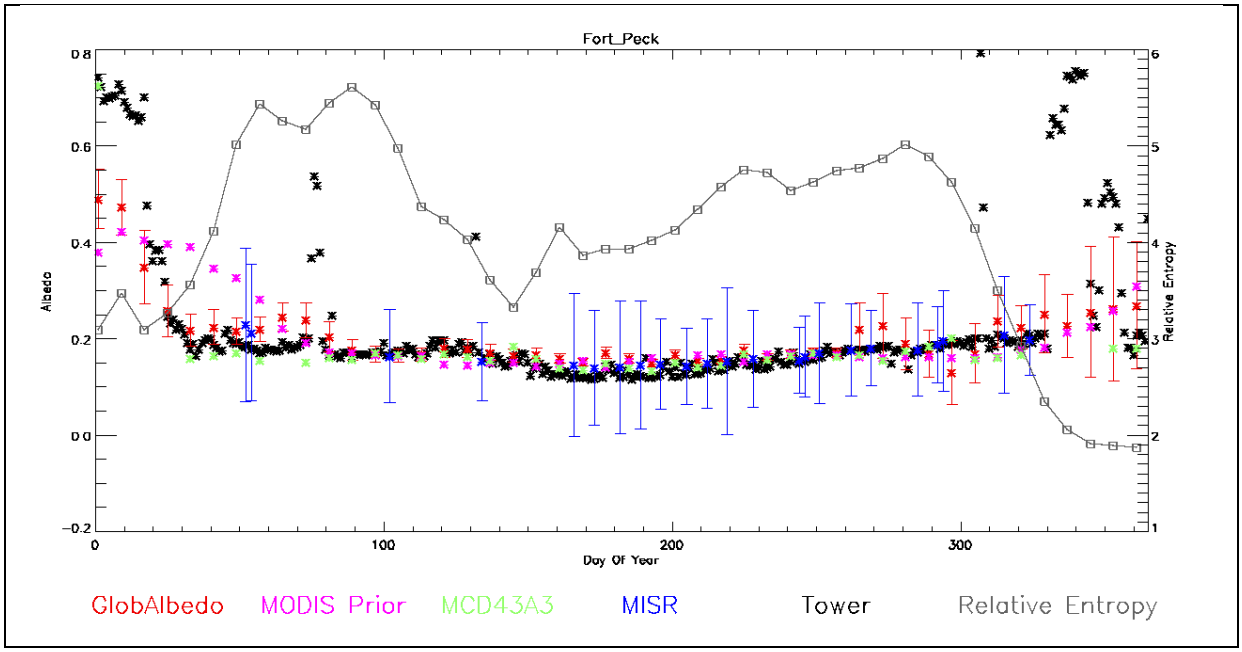



**Figure 12. Temporal plot of Blue-Sky Albedo from GlobAlbedo, MODIS priors, MC43A3 (MODIS Collection 5 at 500m), MISR and SURFRAD tower as well as the relative entropy showing little impact in the GlobAlbedo product from the MODIS priors. Note also the large discrepancy between satellite and tower measurements for ephemeral snow during winter and autumn months.**

**References cited**


Braverman, A., Girolamo, L.D., 2002. MISR global data products: a new approach. Ieee T Geosci Remote 40, 1626-1636.

Breon, F.M., Maignan, F., Leroy, M., Grant, I., 2002. Analysis of hot spot directional signatures measured from space. J. Geophys. Res.-Atmos. 107, art. no.-4282.

Cescatti, A., Marcolla, B., Vannan, S.K.S., Pan, J.Y., Roman, M.O., Yang, X., Ciais, P., Cook, R.B., Law, B.E., Matteucci, G., Migliavacca, M., Moors, E., Richardson, A.D., Seufert, G., Schaaf, C.B., 2012. Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network. Remote Sens. Environ. 121, 323-334.

Diner, D., Martonchik, J., Kahn, R., Pinty, B., Gobron, N., Nelson, D., Holben, B., 2005a. Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land. Remote Sens. Environ. 94, 155-171.

Diner, D.J., Braswell, B.H., Davies, R., Gobron, N., Hu, J.N., Jin, Y.F., Kahn, R.A., Knyazikhin, Y., Loeb, N., Muller, J.P., Nolin, A.W., Pinty, B., Schaaf, C.B., Seiz, G., Stroeve, J., 2005b. The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces, Remote Sens. Environ., pp. 495-518.

Diner, D.J., Martonchik, J.V., Borel, C., Gerstl, S.A.W., Gordon, H.R., Myneni, R., Pinty, B., Verstraete, M.M., 1999. Level 2 Surface Retrieval Algoithm Theoretical Basis. JPL.

GCOS. 2004: Implementation plan for the Global observing system for Climate in support of the UnFCCC. Report GCOS – 92 (WMO/TD No. 1219). 136p

Govaerts, Y.M., Lattanzio, A., Taberner, M., Pinty, B., 2008. Generating global surface albedo products from multiple geostationary satellites. Remote Sens. Environ. 112, 2804-2816.

Houldcroft, C.J., Grey, W.M.F., Barnsley, M., Taylor, C.M., Los, S.O., North, P.R.J., 2009. New vegetation albedo parameters and global fields of soil background albedo derived from MODIS for use in a climate model. J. Hydrometeorol. 10, 183-198.

Lewis, P., Brockmann, C., Danne, O., Fischer, J., Guanter, L., Heckel, A., Krueger, O., López, G., Muller, J.-P., North, P.R., Preusker, R., 2011. GlobAlbedo Algorithm Theoretical Basis Document: Version 3.0. UCL, http://www.globalbedo.org, p. 295pp.

Maignan, F., Breon, F.M., Lacaze, R., 2004. Bidirectional reflectance of Earth targets: Evaluation of analytical models using a large set of spaceborne measurements with emphasis on the Hot Spot. Remote Sens. Environ. 90, 210-220.

Moody, E.G., King, M.D., Schaaf, C.B., Hall, D.K., Platnick, S., 2007. Northern Hemisphere five-year average (2000-2004) spectral albedos of surfaces in the presence of snow: Statistics computed from Terra MODIS land products. Remote Sens. Environ. 111, 337-345.

Moody, E.G., King, M.D., Schaaf, C.B., Platnick, S., 2008. MODIS-Derived Spatially Complete Surface Albedo Products: Spatial and Temporal Pixel Distribution and Zonal Averages. Journal of Applied Meteorology and Climatology 47, 2879-2894.

Muller, J.-P., Lopez, G., Shane, N.S., Danne, O., Brockmann, C., Krämer, U., Zühlke, M., Heckel, A., North, P.R., Domench, C., Guanter, L., Fischer, J., Wang, Z., Schaapman-Strub, G., Cescatti, A., 2012a. GlobAlbedo Test Product Validation Report, http://www.globalbedo.org/docs/GlobAlbedo_TPVR_V2_2.pdf. UCL, UCL-MSSL, p. 92pp.

Muller, J.-P., López, G., Watson, G., Shane, N.S., Tom Kennedy, T.E., Lewis, P., Fischer, J., Guanter, L., Domench, C., Preusker, R., North, P.R., Heckel, A., Danne, O., Krämer,

U., Zühlke, Z., Brockmann, C., 2012b. The ESA GlobAlbedo Project for mapping the Earth's land surface albedo for 15 Years from European Sensors., IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2012. IEEE, Munich, Germany, 22-27.7.12.

Pinty B., Lavergne T., Dickinson R. E., Widlowski J.-L., Gobron N. and Verstraete M. M. (2006) Simplifying the Interaction of Land Surfaces with Radiation for Relating Remote Sensing Products to Climate Models, Journal of Geophysical Research Atmospheres, 111(2).

Pinty, B., Jung, M., Kaminski, T., Lavergne, T., Mund, M., Plummer, S., Thomas, E., Widlowski, J.L., 2011a. Evaluation of the JRC-TIP 0.01° products over a mid-latitude deciduous forest site. Remote Sens. Environ. 115, 3567-3581.

Pinty, B., Andredakis, I., Clerici, M., Kaminski, T., Taberner, M., Verstraete, M. M., et al. (2010a). Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP) Part I: effective Leaf Area Index, Vegetation and Soil properties. Journal of Geophysical Research, 116.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., et al. (2010b). Exploiting the MODIS albedos with the Two-stream Inversion Package JRC-TIP) Part II: Fractions of transmitted and absorbed fluxes in the vegetation and soil layers. Journal of Geophysical Research, 116.

Pinty, B., Lavergne, T., Vossbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M.M., Widlowski, J.-L., 2007. Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products. J Geophys Res-Atmos 112, D10116.

Pinty, B., Taberner, M., Haemmerle, V.R., Paradise, S.R., Vermote, E., Verstraete, M.M., Gobron, N., Widlowski, J.-L., 2011b. Global-Scale Comparison of MISR and MODIS Land Surface Albedos. J Climate 24, 732-749.

Roman, Miguel O. , Crystal B. Schaaf, Philip Lewis, Gail P. Anderson, Feng Gao, Jeffrey L. Privette, Alan H. Strahler, Curtis E. Woodcock, and Mike Barnsley, (2010) Assessing the Coupling between Surface Albedo derived from MODIS and the Fraction of Diffuse Skylight over Spatially-Characterized Landscapes, Remote Sensing of Environment 114(4),738-760

Roujean, J.L., Leroy, M., Deschamps, P.Y., 1992. A Bidirectional Reflectance Model of the Earth’s Surface for the Correction of Remote Sensing Data. J. Geophys. Res. 97, 20,455-420,468

Salomon, J.G., Schaaf, C.B., Strahler, A.H., Gao, F., Jin, Y., 2006. Validation of the MODIS Bidirectional Reflectance Distribution Function and Albedo Retrievals Using Combined Observations From the Aqua and Terra Platforms. IEEE Trans. Geosci. Remote Sensing 44, 1555-1565.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X.W., Tsang, T., Strugnell, N.C., Zhang, X.Y., Jin, Y.F., Muller, J.P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B.X., Liang, S.L., Privette, J.L., Roy, D., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135-148.

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. Journal of Geophysical Research 103, 7755-7766.

Stroeve, J., Box, J., Gao, F., Liang, S., Nolin, A., Schaaf, C., 2005. Accuracy assessment of the MODIS 16-day albedo product for snow: comparisons with Greenland in situ measurements. Remote Sens. Environ. 94, 46-60.

Sun, W., Loeb, N., Davies, R., Loukachine, K., Miller, W.F., 2006. Comparison of MISR and CERES top-of-atmosphere albedo. Geophys. Res. Lett. 33.

Taberner, M., Pinty, B., Govaerts, Y., Liang, S., Verstraete, M.M., Gobron, N., Widlowski, J., 2010. Comparison of MISR and MODIS land surface albedos: Methodology. J Geophys Res-Earth 115, D05101.
Wanner, W., Strahler, A.H., Hu, B., Lewis, P., Muller, J.P., Li, X., Schaaf, C.L.B., Barnsley, M.J., 1997. Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm. J. Geophys. Res.-Atmos. 102, 17143-17161.