A VISION FOR A LAND OBSERVING SYSTEM

P. Lewis, J. Gomez-Dans, M. Disney

Department of Geography UCL and NCEO, Gower St., London WC1E 6BT, UK, Email: p.lewis@ucl.ac.uk

ABSTRACT

In this paper, we argue that the exploitation of EO land surface data for modelling and monitoring would be greatly facilitated by the routine generation of inter-operable low-level surface bidirectional reflectance factor (BRF) products. We consider evidence from a range of ESA, NASA and other products and studies as well as underlying research to outline the features such a processing system might have, and to define initial research priorities.

1. INTRODUCTION

Optical remote sensing of the terrestrial land surface provides society with vital information on the state and dynamics of the part of our planet that we live in and depend on. The range of applications is wide: from physical measures relating to energy exchanges such as albedo or the fraction of shortwave radiation absorbed by vegetation, through monitoring of land cover, urban areas, ecosystem dynamics and crop yields, to monitoring the occurrence and spread of pest outbreaks and wildfire, to name but a few.

Sensing and platform technologies are rapidly developing and there are increasing opportunities for local-scale monitoring from in-situ or UAV-mounted sensors (and of course there is a long tradition of monitoring from aircraft platforms), but the mainstay of global, region and national monitoring is data from Earth Observation satellite sensors. Part of the importance of these EO datasets comes from the ever-increasing legacy collections that we now have, such as data from the Landsat programme going back over 40 years or AVHRR for nearly as long. This provides opportunities not only for exploring and monitoring the current state of the planet but also for looking back at trends and impacts over recent decades. There is recognition of the importance of this legacy for climate and climate impact monitoring, and of the need to develop robust methods to extract the information we need from these and other data [1]. This has led to international efforts such as that of GTOS [2] that have specified the requirements for such 'Essential Climate Variables' (EVCs) and programmes such as the ESA CCI initiative [3] to implement and use the most appropriate algorithms to meet the scientific requirements.

Although the emphasis has rightly been placed on generating the physical measures required (land cover, albedo, fire disturbance etc. for the land surface), we note that a high proportion of these make use of global optical EO data of various sorts and the majority of these need measures of surface directional spectral reflectance.

Historical data are interesting and important but there is also a need to develop systems for monitoring the current state of the Earth land surface. Monitoring in this sense ranges from providing (near) 'real time' data for operational weather forecasting for meteorological agencies or information on fire or illegal logging activity, through to the provision of regular updates of land cover or other land surface properties. This is recognised in efforts such as the EU Copernicus programme [4] that aims to deliver environmental information services (one of which is land monitoring, another, climate change) for the benefit of EU citizens [5] and wider society. This involves the launch and maintenance of a suite of Sentinel satellites for the Copernicus Space Component as well as using data from other 'Contributing Missions'. Although optical EO data cannot in itself serve all such needs, it again makes a crucial contribution.

In summary, we need surface spectral directional reflectance from optical EO data for a range of roles involved in monitoring the land surface. Such data comes from many different sensors, on many platforms. The traditional approach to processing and interpreting such data has been to develop product suites from individual sensors (e.g. the MODIS land products [6]), but it has been recognised that there could be great benefits from improved coordination. This has led to the idea of seeing the combined set of satellite resources as a 'virtual constellation', one part of which is the Land Surface Imaging (LSI) Portal [7] that provides an interface to sensor information and data portals for a range of (optical and other) sensor data in a partially coordinated manner. Many of these datasets (in practice at present, mainly US data) are freely and easily available to users. ESA data are not directly included in this, but the revised ESA data policy [8] and free online access to many datasets [9] shows improvement in this area. US experience, for example within the MODIS and Landsat programmes, has clearly shown the benefits of easy, freely available data. The decision in 2008 to allow Landsat images to be downloaded free of charge has very significantly extended the use of these data and enabled the development of new approaches to processing and using such data: there are now more Landsat data downloads in one day than in a year when the data were sold [10]. This change in data policy has altered the way researchers use such data, practically

*Proc. 'ESA Living Planet Symposium 2013', Edinburgh, UK
9–13 September 2013 (ESA SP-722, December 2013)*

allowing for the first time practical multi-temporal approaches to imaging at moderate resolution. One advantage of such methods is that we are not limited to using the occasional 'cloud-free' scene, but can instead make use of all surface observations when we get them.

On top of free and easy access to the raw image data, another reason for the success (certainly in terms of science use and impact) of such data has been the provision of (easily and freely available) derived low-level datasets and tools. We consider low-level processing to include the treatment of: cloud and cloud shadow detection, ortho-rectification (topographic and sensor-related projection and sampling issues); basic feature detection (e.g. water bodies, snow, burned areas); atmospheric effects; and the BRDF effects. A fundamental requirement for the physical interpretation of EO signals is of course that the sensor is well calibrated. The role of such processing is to prepare the data for further analysis by accounting for (typically, 'correcting') extraneous geometrical or radiometric effects and to allow access to the part of the signal that relates to land surface properties. Basic feature detection (a form of classification) can be important for targeting analysis at particular land surface elements (vegetation, for example) that may require different tools for interpretation. Song et al. [11] note that not all of these steps may be necessary or desirable for some applications, but when multi-temporal approaches are used (e.g. for change detection), they are more important. A danger with any modification of the measured signal can be that it can introduce artefacts that may cause misinterpretation [12], but for the vast majority of land surface monitoring scenarios (where both space and time are important dimensions) we suggest that placing all observations into the space of a regular sampling grid has distinct advantages. When some form of compositing algorithm is used, account should be taken of the true correspondence between the assumed pixel area and the sample grid cell. For the MODIS product suite, this is defined by a term *obscov* that gives the weighted proportion of the measurement covered by a grid pixel [13].

2. A LAND OBSERVING SYSTEM

2.1 An integrated Land Observing System

We believe that all of the core technologies are now in place to develop an integrated Land Observing System (iLOS) and that such a system could play a key role in enabling the improved routine use of EO data and helping to fulfil the requirements of programmes such as Copernicus, the ECVs and related monitoring efforts.

What then do we mean by an iLOS? The fundamental requirement we see for an iLOS is to provide (easily and freely available) *gridded information on the best estimate of surface reflectance over a given spatial and temporal support at all wavelengths over the solar spectrum, along with associated per pixel uncertainty estimates* (the concept could of course be expanded to other wavelength domains). We see this as the fundamental information that is required to generate land surface ECVs and other EO-based information services. Currently, significant efforts are spent in developing 'tailor made' datasets of this sort (e.g. within individual projects or countries), efforts that we believe would be better spent on an integrated effort.

What uses would an iLOS have? First, it would provide the input data for higher-level algorithms for products such as LAI, land cover classification etc. Second, it would be of intrinsic value in directly providing an estimate of the (spectral and directional) shortwave energy interactions at the land surface, for example for energy budget studies, for sensor simulation studies, to aid the interpretation of atmospheric observations (by conditioning the lower boundary interactions). It would also provide an expectation of surface reflectance to aid the detection of non-land surface artefacts (clouds, cloud shadows etc.), provide data to aid in sensor calibration, as well as allow the simulation of data from proposed or forthcoming sensors. An advantage of an integrated system would mean that validation efforts might be more easily coordinated and targeted. As an integrated product, we intend that it should be capable of ingesting data from *any* optical EO instrument that sees that land surface, although the definition of the processing grid will condition the range of data spatial resolutions that would be usefully combined (which would seem to imply that a hierarchy of gridded representations would be most appropriate).

It can be argued that diversity in methods and products is good for the community, but in practice there is little real variation in the robust treatment of low-level processing (often, rather arbitrary, historical or political choices about which models or datasets to use).

2.2 BRDF/Albedo products

A good starting point for consideration of what an iLOS might look like is the NASA MODIS MCD43 BRDF/albedo and MOD09/MYD09 surface reflectance set of products [14, 15] at a base 250/500m spatial resolution, the former at an 8-day sampling and the latter daily (with potentially multiple samples per day). These products have been regularly generated over the MODIS era (2000 onwards) and have undergone algorithmic development and reprocessing to address any issues with the products and to make use of enhanced processing capabilities. It is seen (and has on the whole been funded as) as a long-term, strategic project to deliver high quality data that users require. It has also been backed up by the development of formal validation strategies (within CEOS-WGCV) and datasets, allowing users to know what confidence they might put in the products. MCD43 has currently achieved stage 3, and MOD09 stage 2 validation [16]. The products will soon be reprocessed for collection 6 (C6). MOD09 provides gridded daily surface reflectance products from the NASA Terra and Aqua platforms, along with associated data layers describing basic scene features (cloud, cloud shadow, snow etc.), geometric information such as *obscov* and a host of QA information. It does not at present supply information on per pixel uncertainty. MOD09 provides the input data for the generation of MCD43. Estimation of the BRDF model parameters (that then allow the estimation of BRF at any angular configuration for which the models are valid, practically, up to around 70º zenith) is achieved by combining observations over a 16-day window (in 8 day steps in C5) and solving using linear least squares methods. If the sampling is insufficient for solving for the model parameters (3 per waveband), a backup algorithm is invoked that assumes prior knowledge of the BRDF shape. Some outlier detection is performed during the processing. One advantage of the product is that the estimates every 16 days are independent, so uncertainties should be uncorrelated. In practice this is likely swamped by correlation in the angular sampling regimes for each processing window (e.g. the solar zenith angle will be similar for neighbouring windows). The product flags whether the full or backup algorithms are used, providing a form of QA measure (but no uncertainty) and whether the input data were mostly over snow or snow-free conditions (as this can of course significantly affect the reflectance and the appearance or disappearance of snow within a compositing window can degrade the quality of the product.

For a generic processing system, we should attempt to make use of generic radiative transfer (RT) models rather than imposing particular assumptions on the solution that may conflict with downstream applications. Fundamental technological breakthroughs that have allowed this for BRDF/albedo products occurred in the early-mid 1990s [17-19] with the advent of the MRPV and linear kernel-driven BRDF models. Various flavours of these models form the basis of all current operational products, and they have mostly been shown to provide very similar results in terms of fitting to observations and extrapolation. Comparisons between different broadband albedo products show very strong correspondence [20-21] showing that the discrepancy is mostly in the range +/- 0.02 and bias of around 0.01. This is quite remarkable given the differences in spectral characteristics, scale of observations and angular sampling regimes of the different sensors, and is one of the reasons that we have some confidence to propose combined reflectance products in an iLOS. One feature of these 'semi-empirical' BRDF models however is that they have no inherent 'spectral' components, so a different set of model parameters are required for each waveband to be processed.

There currently exist then, various *sensor-based* global products for BRDF and albedo (e.g. NASA MCD43, NASA MISR, NASA MAIAC using MODIS data [22], EUMETSAT MSG Land SAF [23]). An exception, ESA GlobAlbedo [24] is designed to work with data from any moderate resolution optical sensor but produces *only* broadband albedo estimates (not spectral BRDF). All of these products use similar RT models and similar forms of input data to produce gridded descriptions of land surface reflectance mapped at core near-native spatial resolutions in the spectral space of the sensing instruments (e.g. for 7 MODIS land wavebands for MCD43). They also generally produce integrated broadband (visible, near infrared and total shortwave) estimates that have been seen to mostly agree within reasonable tolerance. It is a difficult task to tease apart what causes the remaining differences, but they are likely caused by model performance issues (e.g. in extrapolation), varying skill in atmospheric correction, and sampling differences (and different uncertainties) arising from 'narrow to broadband' conversion (i.e. the estimation of broadband albedo from narrow band spectral samples) and the angular configurations of sensor samples. Taberner et al. [21] comment that there is value in maintaining independent data product streams as they can be compared and serve to benchmark each other. Since most current products (other than GlobAlbedo and the MSG albedo product) do not produce per pixel uncertainty estimates that can provide assessment of the propagation of uncertainties throughout the processing chain, we can currently only compare mean estimates and cannot readily combine the products. This is of course further complicated by the use of different BRDF models.

The MCD43 product was original designed to take observations from both MODIS and MISR instruments on the NASA Terra platform [25], but that has not yet been achieved, MISR instead producing its own albedo product using different models and assumptions, although [26] and [27] showed how a combined product (using a consistent model) offers improvements in sampling.

An interesting feature of the MSG Land SAF product [23] is its use a temporal weighting function to combine observations rather than compositing over a fixed time window with constant weighting as in MCD43. This is conceptually similar to using a Kalman filter for the parameter estimation [28]. Another feature is the propagation of uncertainty in the approach, though the treatment of atmospheric correction uncertainty is simplistic. It produces BRF and albedo estimates, along with associated uncertainty, in the sensor bandpass channels (0.6-, 0.8- and 1.6 mm) in the MSG grid near real time. The concepts of temporal weighting and uncertainty propagation have also been used in the GlobAlbedo [24], but in that case, a fuller propagation of uncertainty in the treatment of atmospheric effects and gridding effects is implemented, including correlations in uncertainty between wavebands.

Temporal processing for the large datasets involved in these products can be computationally expensive, but there can be much value in having an expectation of BRF. For instance: MOD09 processing in C5 makes use of the previous time step estimate of BRF from MCD43 in estimating BRF shape to deal with surface atmosphere radiative coupling; Roy et al [29] use a previous estimate of BRF to detect sudden signal changes for burned area detection; because of the weighting scheme of the MSG Land SAF product and the GlobAlbedo product, information from previous (and in GlobAlbedo, future) time steps are passed through to constrain the BRDF model parameter estimates as a *prior* constraint. Another interesting product in this context then is MAIAC [22] which uses its accumulated estimate of BRF to detect outliers (residual clouds, cloud shadows etc. that are not flagged in initial pixel identification). This allows for improved 'cleaning' of the dataset (to ensure that pixels are representative of the land surface reflectance), as well as an estimate of the BRF for atmospheric coupling and refined atmospheric correction. There is a not insignificant computational cost to such iterative processing, but the quality of the result seems very high.

As noted, most BRDF/albedo products are generated using data from individual sensor types. This can greatly simplify the processing chain, especially when semi-empirical BRDF models are used as all samples are in a consistent set of wavebands and we have only to solve for the model coefficients. If we wish to combine data from sensors with different wavebands using these models, we have two main choices [30], either: (i) develop some set of spectral basis functions to apply to each of the model parameters and solve for the combined spectral-directional model parameters; or (ii) apply some (e.g. linear) transformations to map from one spectral band to an estimate in another. Both of these approaches have been considered for providing broadband albedo estimates from narrowband spectral samples, though typically the latter is used. To develop such transforms, RT simulations are generally run over a range of conditions to derive effective mapping functions. Since the accuracy of any transformation will depend strongly on particular instrument spectral sampling coefficients, the uncertainty in this transform should be propagated through, as in GlobAlbedo.

The approach taken for GlobAlbedo builds on the work of [30] in this regard. In processing streams for e.g. MCD43 or the MSG product, the BRDF model parameters are first solved for in the spectral sampling space of the instrument. Linear transformations are then applied to convert the model parameters to broadband equivalents. Lewis notes [24] that provided a linear transformation is used, and provided the BRDF models are themselves *linear* (as in the class of kernel-driven BRDF models), and provided *only* broadband (i.e. not narrow band spectral) albedo estimates are required then the order of processing can be changed and we can use the linear transformations to map the input *reflectance* samples to broadband equivalents. So, a (relatively) simple processing chain is enabled within which atmospheric correction and flagging of clouds etc. is carried out in the spectral space of each sensor, but the surface spectral reflectance is transformed to broadband values on a grid, along with associated uncertainty from these processes. At the end of this 'data preparation' step *for whichever sensor we use*, (MERIS and SPOT VEGETATION for GlobAlbedo) we have a product that is very similar in form to MOD09, except that the 'wavebands' represented are broadband values (and we have associated uncertainty).

In the ESA ADAM project [31] this idea is taken further and a set of spectral basis functions developed (from spectral databases) to map from MODIS waveband sampling to full spectrum representations. This means that from estimates of the BRDF model parameters in the 7 MODIS bands, we can estimate the spectral directional reflectance at any angle of wavelength (along with associated uncertainty if uncertainty in the base product is provided). These basis functions also contain the information required to map from any other spectral sampling to MODIS sampling, although in practice it is preferable to develop specific functions for mapping from individual sensors to/from MODIS spectral space (or directly to the basis function space). This is a very interesting idea, and one that in many ways provides the final technology required to develop an iLOS. It can be seen to follow the excellent thesis of Samain [30] and illuminates how to implement what should be one of the key features of an iLOS (full spectral and angular coverage). In practice, it provides the information needed to use methods similar to that of GlobAlbedo where we transform all input datasets to a common spectral representation, but not now just for broadband coverage. Indeed, this was proposed as a refinement for ADAM. The choice of the MODIS spectral representation as the core spectral sampling is a little arbitrary, but pragmatic and convenient. In that context, the current MCD43 of MAIAC products (with uncertainty) could be to all intents and purposes, a contribution to an iLOS that we can use predict the spectral directional reflectance at any configuration. There are many advantages offered by using linear models throughout the processing chain: it simplifies the propagation of uncertainty, and allows for linear transformations e.g. between different waveband at will (e.g. taking a iLOS MCD43 to predict reflectance in MERIS bands), as well as the many others [32].

3. A PROTOTYPE iLOS

3.2 Overview

We now attempt to demonstrate some of the features that an iLOS provides and to develop a prototype system. In doing so, we aim to take the best features of the various BRDF/Albedo products developed. At present, we limit the design to a single pass of processing, though obviously refinements (e.g. to atmospheric correction) could be made by iteration as in MAIAC.

In essence, our proposal for a prototype iLOS involves a first-pass pixel identification and atmospheric correction with propagation of uncertainty using codes such as those developed in GlobAlbedo and ADAM (or MAIAC/MCD43). During this processing, a gridded product is generated in the MODIS grid and spectral space (at 10 km for ADAM processing). We then have a set of synthetic MODIS inputs, as well as actual MODIS (MOD09) observations.

We keep track of three versions of land surface state, as in GlobAlbedo processing: snow-free, snow and actual (snow or snow-free) conditions, represented by an initial estimate of the BRDF model parameters for such conditions, for the first day that we wish to ingest data (doy 001, 2005 in the example processing here). We have developed these initial estimates from MCD43 climatologies to demonstrate principles here, and then refined the estimate by reversing the order of data in the first month (to mimic boundary conditions has we processed further years of data).

In the next stage, we apply a Kalman Filter with a zero-order process model to these data. For each day of processing, we read each dataset for that day (in time order of acquisition) and refine the pixel identification by detecting outliers as unexpected values. We then apply the Kalman filter to the remaining samples and update out estimate of state (BRDF parameters). When we reach the end of the time series, we process the data in reverse order and combine the results to achieve processing with a Kalman smoother. This provides the posterior estimate of state and its uncertainty, on a 500 m grid for (in this case) daily time steps.

3.2 Spectral mapping

Although it requires some further investigation, we suppose (as in the ADAM project [31]) for the present that a spectral representation in MODIS waveband space is sufficient to permit full spectrum mapping (at 1 nm intervals). This greatly simplifies processing chains, as the core requirement then is to derive a representation of BRDF model parameters in these wavebands. This is clearly trivial for MODIS data. For other sensors, we develop linear models for example from SPOT VGT and MERIS:

$$R_{MODIS} = R_{0VGT} + M_{VGT}R_{VGT}$$

$$R_{MODIS} = R_{0MERIS} + M_{MERIS}R_{MERIS}$$

Examples of these are shown (developed in the ADAM project) in Fig. 1, developed using a range of spectral databases for MERIS. Fig. 1 also shows the normalised eigenvalues associated with these functions (along with cumulative values). We can see that functions 9 to 15 in Fig. 1 have rather high (positive and negative) magnitude but only contribute a small amount to the variance in the spectra. It is probably worthwhile then applying some cutoff to this set of functions rather than using the full set. We also see that the main role of these higher order basis functions is in predicting the MODIS reflectance at wavelengths longer than those sampled by MERIS, which will in any case have a significantly higher uncertainty than the prediction of reflectance in the visible and near infrared bands. Similar information is available for mapping from the SPOT VGT wavebands to MODIS. In this case, we are attempting to map from 4 input wavebands to the 7 MODIS bands and at least 3 basis functions will be required for this.

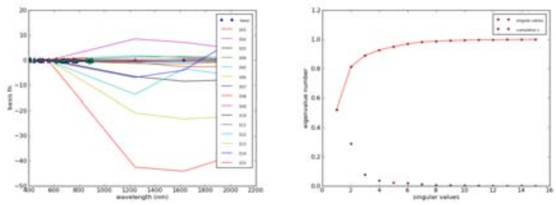

Figure 1. Left: MERIS to MODIS basis functions; Right: Normalised Eigenvalues associated MERIS basis functions

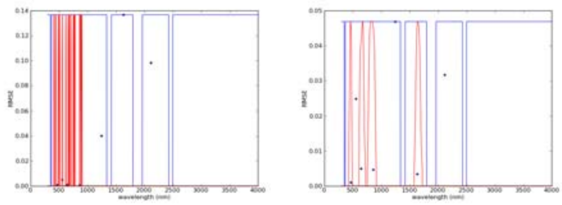

Figure 2. RMSE over independent test spectra (Left:MERIS, Right: VGT)

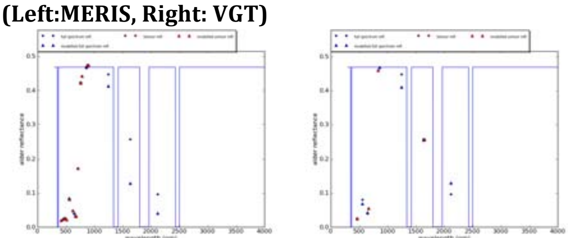

Figure 3. Reconstruction of Alder spectrum (Left:MERIS, Right: VGT)

Fig 2. shows the RMSE achieved in the mapping using a test spectral database for MERIS (left) and VGT (right) to MODIS mapping. Unsurprisingly, the errors at the unsampled (longer wavelength) bands are relatively high, but provided we characterise this correctly and propagate the uncertainties, we should still achieve a consistent product. The case of VGT is in many ways simpler. Although the error is generally higher, it is relatively consistent across MODIS wavebands, due to the fuller spectral sampling offered by this instrument. Fig. 3 shows example reconstructed spectra in MODIS wavebands, given these mapping functions and spectral samples in MERIS (left) and VGT (right) wavebands. We see that even in the case of MERIS, the correlations inherent in reflectance spectra allow a reasonable mapping to MODIS wavebands, with larger errors in the VGT and MERIS mappings corresponding to where we see higher overall errors in Fig. 2.

3.3 Outlier detection

Although first-pass pixel identification (identification of snow, cloud etc.) can be quite successful, it is never completely so and we must expect the masked surface reflectance dataset to contain contamination. This is likely to be better for some instruments than others (e.g. it is generally easier to detect cloud if a thermal band is available). Whilst the modelling of temporal parameter development with a Kalman filter/smoother will be robust to outliers to some extent, it is preferable to remove them from consideration. Further, improved pixel identification can lead to improved atmospheric correction in a multi-pass system such as MAIAC. In the prototype, we apply a simple but seemingly robust approach in which we compare each pixel initially identified as clear with a simulation of the image (at the correct viewing and illumination angle configuration) and calculate a form of Z-score between the predicted and observed values (relative to the uncertainty in the measurement and the prediction). We then examine samples for which the Z-score is over a given threshold (3.0 here) to make a decision to reject them as snow or snow-free samples. The use of different state estimates for snow and snow free conditions adds to the processing time, but seems to allow much better filtering of outliers (as well as interesting by-products).

3.4 Kalman Filter/Smoother

We initialise the processing with a prior estimate of state for the first day we consider (001, 2005 here), and then apply the Kalman Filter to update the state estimate for pixels that are flagged as uncontaminated land surface observations. We also ingest pixels identified as clear and snow free into the update of the snow-free state, and similarly for the snow state. We then step on to the next day with our zero-order process model, inflating the uncertainty by our expectation of change (i.e. departure from this model) in this time period (a standard deviation of 0.001 here). This process is then iterated, moving over each day until the end of the time series. For each time period we have observations, we write out a state representation file (in netCDF format) on the forward pass through the data. We then implement a return pass to achieve a Kalman smoother where we load the state file and in essence combine this with the state estimate from the reverse sweep. In both forward and reverse sweeps, we keep careful track of pixels identified as outliers, with the intention of studying these for persistent effects that may be indicative of fire activity or other causes of dramatic change to the land surface state.

4. RESULTS

Figure 4. State estimates over Scotland, shown as BHR RGB true colour composites for snow-free (top), snow (middle), and actual (bottom) conditions. DOY 1 2005.

We have performed an initial implementation of the algorithm, and initial testing (using only MODIS Terra and Aqua data, for the year 2005) over Scotland. This is a challenging environment to do land surface monitoring, due to high cloud cover and complex snow impacts. This is illustrated in Fig. 5, which shows the cumulative (obscov-weighted) number of samples from MODIS available over the area by the end of February 2005. This is very small over mountainous areas, but we can see in Fig. 6 that we have captured the dynamics of the snow cover well with this.

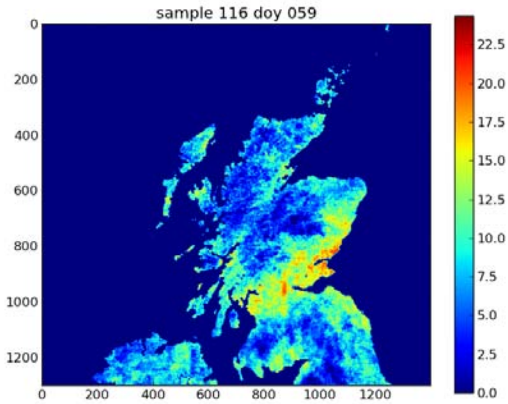

Figure 5. Cumulative weighted number of MODIS (Terra and Aqua) samples between 1st Jan and 28th Feb 2005 (total, snow and clear).

Figure 6. BHR estimates over Scotland, as RGB true colour composite for actual conditions: DOY 59

Figs. 4 and 6 illustrate the state datasets on the 500 m MODIS grid and MODIS wavebands. We have not as yet produced any validation of the product, but a visual inspection confirms the data to be of high quality. In particular, the appearance and disappearance of snow is consistent in these data over Scotland with historical weather reports and other dynamics as expected.

5. DISCUSSION

The algorithm we propose and prototype here is in many ways, simple and elegant. It takes heterogeneous input datasets to/from a homogeneous representation (as in GlobAlbedo) to enable it to function with any optical input datasets. Of vital importance to this step is the estimation of uncertainty inherent in the observations (sensor calibration uncertainty and gridding effects), in the treatment of atmospheric effects, and in estimation of reflectance at MODIS wavebands (or some other basis set) and the application of data assimilation (DA) methods through the use of the Kalman Filter/smoother. Whilst this is one of the simplest DA methods available, it is the most relevant to use in this context where all models are linear. As demonstrated previously with MAIAC and further shown here, outlier detection can be simply and robustly applied once an expectation of surface state is available. We have not yet investigated other potential improvements likely to be provided by this expectation, e.g. improved atmospheric correction etc., but this is likely to occur with this form of processing as with MAIAC.

Although we suppose the core processing grid to (conveniently) be at 500 m resolution here (to match MODIS datasets), it is very likely that the estimates of surface state obtained will enable the improved conditioning (e.g. atmospheric and BRDF correction) of higher resolution datasets by providing an expectation of coarser scale reflectance in a similar manner to many recent studies using MODIS and Landsat TM data. It is also likely to have significant impact when processing coarser resolution datasets, particularly for improved methods of (sub pixel) cloud detection.

Whilst there is clearly further work required to fully test this and develop an operational algorithm, the ideas behind it are based on methods used in the processing of existing land surface products, so development is likely to be rapid.

This form of a processing system should be robust and generic in providing routine estimates of surface BRF (with by-products, albedo, normalised BRF, snow-free BRF, change events). By using generic linear models throughout, practical algorithms can be developed that are in keeping with the desire for not over-constraining the BRF estimates. The approach provides a route for the integration of data from heterogeneous (in wavebands and spatial scale of observations) data sources, and does so within a Bayesian framework that tracks uncertainties throughout.

Research priorities we identify include: implement, test and compare demonstrator products, including the testing some options (e.g. change detection); get community by-in to the idea and agree the broad approach and requirements; define conditions to be met for merging data streams (learning from existing operational DA; and further investigate spectral mapping issues; and demonstrate and test spatial scaling concepts.

REFERENCES

1. GCOS, (2013) Essential Climate Variables, http://www.wmo.int/pages/prog/gcos/index.php?name =EssentialClimateVariables.
2. GTOS (2013) Terrestrial ECVs, http:// www.fao.org/gtos/topcECV.html
3. ESA (2013) Climate Change Initiative, http://www.esa-cci.org
4. European Commission, (2013) Copernicus: the European Earth Observation Programme, http://copernicus.eu/
5. European Commission, (2013) Press release: Copernicus: new name for European Earth Observation Programme, http://europa.eu/rapid/ press-release_IP-12-1345_en.htm
6. NASA (2013) MODIS Land, http://modis-land.gsfc.nasa.gov/
7. CEOS (2013) Land Surface Imaging Portal, http://eros.usgs.gov/ceos/lsip.shtml
8. ESA (2013) Revised ESA Earth Observation Data Policy, https://earth.esa.int/web/guest/-/revised-esa-earth-observation-data-policy-7098
9. ESA (2013) Earthnet Online: Data Access, https://earth.esa.int/web/guest/data-access/online-archives (see also https://ladsweb.nascom.nasa.gov/ MERIS).
10. Wulder, M.A., et al. (2012), Opening the archive: How free data has enabled the science and monitoring promise of Landsat, *Remote Sensing of Environment*, (122), 2-10.
11. Song, C., et al. (2000), Classification and Change Detection Using Landsat TM data: When and How to Correct Atmospheric Effects?, *Remote Sensing of Environment*, 75, 230-244.
12. Tan, B., et al. (2006), The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions, *Remote Sensing of Environment*, 105(2), 98-114.
13. Wolfe, R. E., et al.(1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Trans. Geosci. Rem. Sens.*, 36, 1324–1338.
14. Schaaf, C. B., et al., First Operational BRDF, Albedo and Nadir Reflectance Products from MODIS, *Remote Sens. Environ.*, 83, 135-148, 2002.
15. Vermote, E. F., et al. (2002), Atmospheric correction of MODIS data in the visible to middle infrared: first results, Rem. Sens. Environ., 83 (1-2), 97-111.
16. NASA (2013) MODIS Land Team Validation http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD43, http://landval.gsfc.nasa.gov/ ProductStatus.php ?ProductID=MOD09
17. Roujean, J., Leroy, M., and Deschamps, P., A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data., *Journal of Geophysical Research*, *97*, 20 455–20 468, 1992.
18. Engelsen, O. et al. (1996) Parametric bidirectional reflectance factor models: evaluation, improvements and applications, Tech. Rep. EUR 16426, European Commision, Ispra, Italy, 1996.
19. Wanner, W. et al. (1995) On the derivation of kernels for kernel-driven models of bidirectional reflectance. *J. Geophys. Res.*, *100(D10)* 21077–21089.
20. Carrer, D., et al. (2010), Comparing operational MSG/SEVIRI land surface albedo products from Land SAF with ground measurements and MODIS, *IEEE Trans. Geosci. Rem. Sens.*, 48(4), 1714-1728.
21. Tabener, M., et al. (2010), Comparison of MISR and MODIS land surface albedos: Methodology, *J. Geophys. Res.*, 115(D05101) doi:10.1029/2009JD012665.
22. Lyapustin, A. et al. (2012). Improved cloud and snow screening in MAIAC aerosol retrievals using spectral and spatial analysis. Atmos. Meas. Tech., 5(4), 843-850.
23. Geiger B, et al. (2007), Land Surface Albedo derived on a daily basis from Meteosat Second Generation Observations, *IEEE Trans. Geosci. Rem. Sens.*, 46, 3841-3856, 2008.
24. Lewis, P. et al. (2012), The ESA globAlbedo project: Algorithm, IGARSS, pp.5745-5748
25. Wanner, W. et al. (1997) Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm, *J. Geophys. Res*, 102, D14,17143-17161.
26. Jin, Y. et al. (2002), Improving MODIS Surface BRDF/Albedo Retrieval with MISR Multi-angle Observations, *IEEE Trans. Geosci. Remote Sens.*, 40, 1593-1604, 2002.
27. Lucht, W. and Lewis, P. (2000) Theoretical noise sensitivity of BRDF and albedo retrieval from the EOS-MODIS and MISR sensors with respect to angular sampling. *International Journal of Remote Sensing* 21(1) 81-89.
28. Samain, O., et al. (2008) Use of a Kalman filter for the retrieval of surface BRDF coefficients with a time-evolving model based on the ECOCLIMAP land cover classification, *Remote Sens. Environ.*, 112(4), 1337–1346.
29. Roy, D. et al. (2001) Burned Area Mapping Using Multi-Temporal Moderate Spatial Resolution Data - a Bi-Directional Reflectance Model-Based Expectation Approach, Rem. Sens. Environ., 83, 263-286.
30. Samain, O., (2004), Fusion multi-capteurs de données satellitaires optiques pour la détermination de variables biophysiques de surface, Ph.D. dissertation, Université Paul Sabatier, Toulouse, France, 2004.
31. Muller, J.P., et al. (2013) A Surface Reflectance Database for ESA's Earth Observation Missions (ADAM), *This symposium*.
32. Lewis, P., (1995), The Utility of Linear Kernel-Driven BRDF Models in Global BRDF and Albedo Studies. Proc IGARSS'95, Firenze, Italy.Volume 2, 10-14 July 1995 Page(s):1186 - 1188 vol.2.