DATA ASSIMILATION OF SENTINEL-2 OBSERVATIONS: PRELIMINARY RESULTS FROM EO-LDAS AND OUTLOOK

Lewis, P.[(1)], Gomez-Dans, J.[(1)], Kaminski, T.[(2)]; Settle, J.[(3)], Quaife, T.[(3)], Gobron, N.[(4)], Styles, J.[(5)], Berger, M.[(6)]

[(1)] *UCL and NCEO, Gower St., London, WC1E 6BT, UK, Email:* {p.lewis|j.gomez-dans}@ucl.ac.uk
[(2)] *FastOpt, Lerchenstr. 28a, D-20767 Hamburg, Germany, Email:* Thomas.Kaminski@fastopt.com
[(3)] *University of Reading and NCEO, Reading RG6 6AL, UK, Email:* jjs@mail.nerc-essc.ac.uk
[(4)] *European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, Global Environment Monitoring Unit, TP 272, via Enrico Fermi 2749, I-21027 Ispra (VA), Italy, Email:* nadine.gobron@jrc.ec.europa.edu
[(5)] *Assimila Ltd., 1 Earley Gate, Reading RG6 6AT, UK, Email:* jon.styles@assimila.eu
[(6)] *ESA ESRIN, Science Strategy, Coordination and Planning Office (EOP-SA), Via Galileo Galilei, Casella Postale 64, 00044 Frascati (RM), Italy, Email:* Michael.Berger@esa.int

ABSTRACT

Attractive properties of Sentinel-2 MSI for monitoring vegetation dynamics include its reasonable spatial resolution, spectral sampling and revisit capabilities. But the large number of factors that affect vegetation reflectance spectra and the impact of clouds mean that the likely uncertainties when monitoring crops and forest resources are likely still relative high. One way of improving on this is to use data assimilation with optical radiative transfer models as observation operators. This allows for fuller use of the measured signals (rather than e.g. vegetation indices using only a few channels), for better use of multi-temporal data and ultimately reduced uncertainties in vegetation and soil state. In this paper, we discuss state estimation approaches and data assimilation in particular. We present an Earth Observation Data Assimilation System (EO-LDAS) which is an implementation of such ideas. EO-LDAS, developed under ESA funding, uses the semi-discrete radiative transfer model as observation operator and a temporal regularisation constraint as dynamic model and allows state variables to be solved at a daily time step with associated uncertainty. As well as temporal interpolation, the system provides a reduction in uncertainty of around 2 over a scenario MSI data alone. The outlook for the application of such methods is also discussed.

1. INTRODUCTION

1.1. The Remote Sensing Problem

The 'remote sensing problem' involves the estimation of information from remote, (generally) radiometric measurements. There have been many approaches taken to this, but at heart, it is an optimal estimation problem. We can interpret this as some representation of *state* that we wish to infer from our measurements. This state can be viewed in a Bayesian context [1] as some joint probability density function (PDF), which might usefully in many circumstances be simplified to a multivariate Gaussian distribution that we can describe with a mean vector x_{post} and a variance/covariance matrix C_{post}. We can also describe our observations in this same way, e.g. for assumed Gaussian distributions as a mean vector y with associated uncertainty C_{obs}. The remote sensing problem then is to provide an estimate of x_{post} and C_{post} given y and C_{obs}. To solve this, we will need an operator to map between these spaces, which we can phrase as an *observation operator* $\hat{y} = H(x)$, where $H(x)$ provides an estimate of y for given x, $\hat{y}$. There will typically be some uncertainty associated with this mapping, which, if assumed Gaussian we can represent by C_H. Often in remote sensing we have seen the remote sensing problem as simply trying to find and apply the inverse of $H(x)$, $\hat{x} = H^{-1}(y)$ [2], but in a more rigorously approach we recognise that this needs to take account of the various uncertainties.

In this paper, we consider the particular problem of trying to estimate the state variables of a vegetation canopy (leaf area index (LAI), leaf chlorophyll concentration etc.) from optical multispectral measurements such as those available from the Sentinel-2 MSI sensor [3]. In such a case $H(x)$ could be a radiative transfer model that predicts top of canopy or top of atmosphere radiance measurements [4]. Alternatively, it could simply be a mapping via a vegetation index [5].

In trying to tackle this problem, we often find that there is insufficient *information* in the observations to strongly affect our estimate of some or all of the state vector. Alternatively we might find that there are many versions of the state vector that are capable of reproducing the observations. We can say then that the remote sensing problem is often *ill conditioned*. There have been various responses to this. One of the most

commonly used in mapping vegetation properties has been to attempt to transform the observations (e.g. through a vegetation index (VI)) to maximise the sensitivity of the transformed variable to a single element of x (e.g. LAI) [5]. In practice, many such transformations (such as the Normalised Difference Vegetation Index) are essentially measurements of the depth or relative depth of an absorption feature and are typically useful in obtaining first order estimates of the state variables of interest. Despite their popularity, VIs suffer from a number of known failings [6].

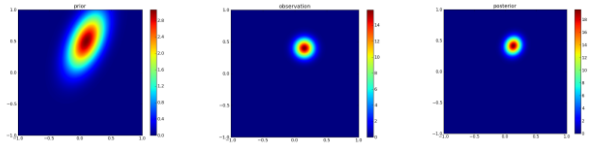
1.2. Optimal estimation

As indicated above, an alternative approach is to treat the operator $H(x)$ as a radiative transfer (RT) model based on an understanding of the physics of radiation scattering, but since such models may typically have more than ten parameters we tend to hit the problem of ill conditioning if we attempt to solve for all of the state vector. The pragmatic response to this has been to assume that some of the elements in the state vector are known and to solve for a limited subset. An example of this is the MODIS LAI/fAPAR product [7] that uses a land cover-based LUT from RT modelling. All state variables other than LAI are assumed fixed (i.e. known) in the modelling, although their values (and assumptions about vegetation structural arrangement) vary with land cover class. In a Bayesian context we can call this a form of *a priori* constraint: the algorithm developers have a belief that (non target) state variables such as leaf reflectance and transmittance realistically only vary within certain limits, so they choose to fix these to some average of their expectation. Properly, there should be some uncertainty associated with these fixed values. In that case, we can take this assumed prior knowledge as a vector x_{prior} with associated (Gaussian here) uncertainty C_{prior}. We can illustrate the simplest case of this form of constraint by assuming a toy example where the observation operator to be an Identity operator $I(x)$ so that $\hat{y} = x$. We suppose a two-dimensional state vector with a prior estimate of state represented by $x_{prior} = (0.1, 0.5)$ and $C_{prior} = \big((0.04, 0.03), (0.03, 0.09)\big)$ and an observation $y = (0.15, 0.40)$ with uncertainty $C_{obs} = \big((0.10, 0.00), (0.00, 0.01)\big)$. This is illustrated in figure 1, panels (a) and (b). It is trivial to show that in this case the combined PDF is obtained by:

$$\begin{aligned} x_{post} &= \left(C_{prior}^{-1} + C_{obs}^{-1}\right)^{-1}\left(C_{prior}^{-1}x_{prior} + C_{obs}^{-1}y\right) \\ C_{post} &= \left(C_{prior}^{-1} + C_{obs}^{-1}\right)^{-1} \end{aligned} \quad (1)$$

Eq. 1 is of course just textbook statistics: if we interpret x_{prior} as simply another (independent) observation of x, then we see immediately that this expresses that we combine two observations by taking an uncertainty-weighted mean. In the most trivial case where the uncertainties are the same, the best estimate of x is the mean and the posterior variance/covariance becomes reduced by a factor of $1/\sqrt{2}$. We present this here however *because* it should be a useful and rather intuitive starting point for most readers and to use it introduce a framework for solving the remote sensing problem by combining PDFs which is Bayes theorem:

$$P(b|a) = P(b)P(a|b)/P(a)$$



(a) Prior estimate (b) Observation (c) Posterior estimate

Figure 1. Illustration of combining PDFs.

We can now state the result of our state estimation (our solution to the remote sensing problem) as a *conditional* probability of state b given the information a, $P(b|a)$. The importance of this is perhaps better illustrated by writing $P(b|a) \propto P(b)P(a|b)$ since the term $P(a)$ above simply acts to normalise the PDF integral. This then is the theorem that tells us to combine PDFs, we multiply them.

The Gaussian PDF is an exponential of a term $J(x)$, $\exp\big(-J(x)\big)$ where e.g. for x_{prior} and y:

$$J\big(x_{prior}\big) = \frac{1}{2}\big(x_{prior} - x\big)^T C_{prior}^{-1}\big(x_{prior} - x\big)$$
$$J(y) = \frac{1}{2}\big(y - H(x)\big)^T (C_{obs}^{-1} + C_H^{-1})\big(y - H(x)\big)$$

Where $.^T$ denotes the transpose operation. So Bayes theorem tells us that to combine Gaussian PDFs we multiply exponentials, giving the *posterior* as a Gaussian distribution that is the sum of the $J(x)$ terms, so developing the above example and writing:

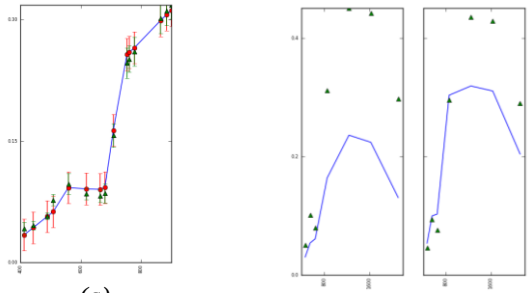
$$J\big(x_{post}\big) = J\big(x_{prior}\big) + J(y)$$

and the maximum likelihood estimate of x_{post} is given by the minimum of $J\big(x_{post}\big)$. We recognise $J\big(x_{post}\big)$ as a *cost function* and note that Bayes theorem tells us that it is simply a sum of cost functions providing a set of constraints. To solve for the minimum of this, we need to find where the derivative $J'\big(x_{post}\big)$ (the Jacobian), which is the sum of the derivatives of the individual cost functions, is zero. The estimate of the posterior uncertainty is given by the curvature of the cost function at its minimum, which is the second order derivative

of $J(x_{post})$, the Hessian, $J''(x_{post})$, which is the sum of the Hessians for the individual cost functions.

To illustrate this, we take the observation operator to be the canopy RT model of [8] with a spectral function for leaf reflectance and transmittance that is an approximation to the PROSPECT model [9] [10] with the soil spectrum defined by the basis functions of [11] as described in [12]. This model of canopy state has 13 parameters for each location, covering LAI, soil biochemistry etc., although only 7 or 8 of these may be accessible depending on spectral and angular sampling from typical Earth Observation (EO) instruments. We will deal with Gaussian distributions of error here, so apply exponential transformations to appropriate variables (LAI and leaf biochemistry concentrations) to approximately linearise the sensitivity. For observations, we take a MERIS top of canopy spectrum over a field site in Germany. We apply a constraint using x_{prior}, but with large uncertainty. This helps to condition the solution, but will not strongly influence it. We then solve for the minimum of the combined cost function for observations and prior for 7 of the model parameters. Table 1. shows the prior and posterior model state vectors, along with the associated standard deviations. For a full description of the model parameters and units see [12], but we can recognise for instance that the posterior estimate of LAI is $-2\ln(0.72) = 0.66$ here and that 95% confidence intervals on this would give an upper limit of $-2\ln(0.72 - 0.04\text{x}1.96) = 0.89$ and a lower one of $-2\ln(0.72 + 0.04\text{x}1.96) = 0.45$ which equates to 2/3 of the signal and seems quite a large uncertainty. However, if we equate this to an effective standard deviation in LAI of 0.11 this would compare very favourably to errors in current LAI products if this were a true estimate of error [13] so we might consider this a reasonable result: even without fixing the non target state vector elements we are able to provide a viable estimate of all (sensitive) parameters from the 'inversion' of an RT model with only weak prior constraints.

Fig. 2a. shows the reflectance data used here in the 15 MERIS bands and we can see that this 'viable' result is obtained by the model fitting well to the observations.. If we then try to use this estimate of state to predict what we would see from another sensor (MODIS here) (Fig. 2b) at a different set of view zenith and azimuth angles (around 40° view zenith here, rather than the near nadir MERIS observation) we see that the predictive power of this model and state vector set is rather poor. It is unsurprising that this is the case outside of the wavelength range of the MERIS data, where leaf and soil water content (not solved from the MERIS data). This poor performance might be due to errors in the model assumptions or scale differences, but this is unlikely as the observations are of a large wheat field that ought to reasonably correspond to the model assumptions. The most likely reason then is simply the large uncertainties in the state vector estimate, compounded by correlation effects in the parameters.



(a) *(b)*

Figure 2. (a) MERIS observation (ToC reflectance) as a function of wavelength (nm) in red/blue and posterior estimate of reflectance (green); (b) Prediction of MODIS reflectance (at two different view angles) for the same day/location from the MERIS-estimated state vector (observations in blue here, predictions in green).

	x_{prior}	σ_{prior}	x_{post}	σ_{post}
$e^{\frac{-LAI}{2}}$	0.95	1.00	0.72	0.04
$e^{\frac{C_{ab}}{100}}$	0.10	1.00	0.78	0.04
Scen.	0.00	1.00	0.00	0.06
$e^{-50C_{dm}}$	0.30	1.00	0.48	0.04
N	0.90	2.50	1.75	0.28
s_1	0.01	4.00	1.78	0.49
s_2	0.01	5.00	1.39	1.02

Table 1. Prior and posterior model state from MERIS

2. DATA ASSIMILATION

2.1 Context

Data assimilation (DA) is a set of statistical and computational methods that enable the optimal merging of *models* and *data*. The statistical basis for it is that presented above (although we have limited most discussion here to Gaussian distributions). In essence, it allows us to use multiple constraints to solve the problem. In general, a DA system will contain: (i) a background (or prior) constraint, involving a PDF from climatology or previous DA runs; (ii) a process model, predicting linkages between the state vector elements in space and/or time; (iii) observational constraints.

Early examples of DA include those used to improve short-term weather predictions from meteorological models [14]. Here observations are used to improve predictions of the state of the atmosphere, this being represented by a large number of interconnected cells in a 3D grid. In such applications, a 'strong constraint' DA is often used within which the physics of the

meteorological model (which is the process model in this case) are assumed to be without error, with the uncertainty in the estimate of the atmospheric state coming from errors in the initial conditions. By using DA to update the estimate of the initial conditions, the forecast from that state is improved. The *a priori* estimate for the state for the next run of the assimilation can be provided for example from the *a posteriori* result of the previous run. Within the field of meteorology, along with areas of remote sensing such as atmospheric sounding where DA has been used for some time [15] there has been a development of a set of techniques for DA that researchers are starting to apply to a wider range of problems. In contrast to the strong constraint approach mentioned above, a *weak constraint* method considers uncertainty in the process model. Other examples of DA include: the work of [16] who used a strong constraint DA to estimate parameters of vegetation process model; [17][18] who coupled an LAI phenology model to a RT observation operator to estimate LAI; or various applications in remote sensing of hydrology (e.g. [19]). We can identify two main approaches to DA: (i) variational methods and (ii) sequential methods.

2.2 Variational methods

In variational methods, numerical algorithms such as iterative gradient descent methods are used to find the minimum of the combined cost function. In a sense, this is similar to numerical methods used for many years in trying to 'invert' optical canopy reflectance models [2], but significant differences are: (i) computer code for the Jacobian is often used, involving tangent linear or adjoint codes that can be developed using automatic differentiation tools such as TAF [20] [21] or TAPENADE [22] and which allow for more efficient solutions to the optimisation, e.g. using L-BFGS-B [23]; (ii) the use of a prior estimate in the DA deals with many of the problems of ill posedness (see e.g. [24], [25]); (iii) the use of a process model provides additional constraints on expected state vector behaviour in space/time which better constrains the solution. In a variational scheme, an estimate of all elements of the state vector will be sought at the same time, although, as in the meteorological example mentioned above, they can be applied sequentially to subsequent (e.g. time) windows. As a result, they can provide posterior uncertainty information relating all elements in the state vector. These methods can be used to solve very large scale problems. They are easy to define, but to be efficient generally require that code for the Jacobian is available. A drawback is that they can be difficult to use unless Gaussian statistics are assumed. A schematic of a strong constraint variational DA is illustrated in fig. 2. From an initial estimate of the state the process model $Q_t(x)$ is applied to produce a proposed state at time t, x_t. This is transformed into the space of the observations by $H(x_t)$ as $\hat{y}$, and a cost function J_{obs} built from the difference between this and the observations y_{obs}, using the uncertainty in the observations (and potentially the observation operator). An additional constraint is developed as J_{prior} that is the cost of the state departing from our estimate of what the state ought to be, x_{prior} relative to the uncertainty in the prior. The derivatives J'_{obs} and J'_{prior} are generally calculated at the same time and the combined cost function and its derivative fed into an algorithm such as L-BFGS-B to attempt to find the cost function minimum and provide the *a posteriori* estimate of the state x_{post}. At that point, the Hessian is calculated and the posterior uncertainty estimated. The 'strong constraint' term refers to the process model that we assume can map the initial conditions x_{post} to the time/space of the observations. Although uncertainty in the observation operator can be incorporated, it is often not known.

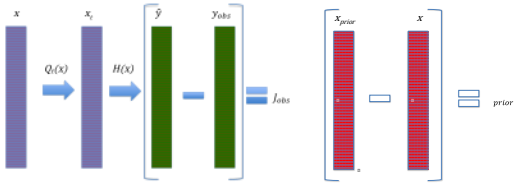


Figure 3. Strong constraint variational DA

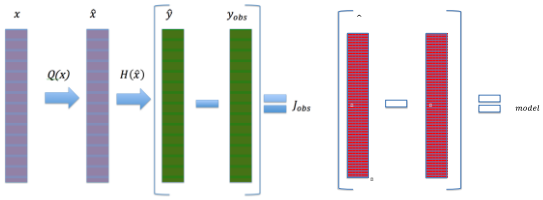


Figure 4. Weak constraint variational DA

In a weak constrain system uncertainty in the process model is explicitly treated. This can be phrased as an additional constraint J_{model}. An example of this (fig. 3) would be to consider that x now contains a representation of state at all times/locations (rather than just the initial conditions) so that the process model now maps from subsets of the state vector to other subset (e.g. from the part of the state at a particular time/space to another time/space). We now treat this prediction from the model as an estimate of state $\hat{x}$ with associated uncertainty C_{model} and form the cost function J_{model} between this and estimate and the *a posteriori* estimate that we wish to solve for. We can of course use the prior (background) constraint as previously, though we note that in this example it applies to the whole of the state vector at all (sampled) times/space. We can illustrate this with the simplest form of process model, a *zero order* process model, in which the value of state at a neighbouring sample (in space/time) is modelled the current state, i.e. $Q(x) = x$ with associated uncertainty. This can be phrased using a (linear) differential operator

D so we have the expectation $Dx = 0$ with uncertainty C_{model}. In the simplest form of this we can assume that the uncertainty is constant over all time/space, so $C_{model} = \gamma^{-2} I$ where I is the identity operator and γ is the inverse of the standard deviation in the model, which is of course a form of *tolerance* to the credibility of this model, a *smoothness* term on the variation of state, or equally the inverse of our expectation of how much we expect the state to change (in a mean squared sense) from one time/space location to another. In this case:

$$J_{model} = \frac{\gamma^2}{2} x^T (D^T D) x$$

The derivatives of this are straightforwardly calculated as it is a linear model: $J'_{model} = \gamma^2 (D^T D) x$ and $J''_{model} = \gamma^2 (D^T D)$. The matrix operator D has ones along the leading diagonal and -1 for the neighbouring state (scaled by the distance to the neighbour). There are various options for how this is treated at the boundaries of the time/space domain, such as periodicity, reflectivity etc. [26]. The impact of such a process model is to *regularise*, i.e. smooth the state, the degree of smoothness being controlled by γ. Since the degree of smoothness desired is often not known, this must typically be estimated by cross validation approaches or by making other assumptions. [27] use such a regularisation approach to solve for linear kernel-driven BRDF parameters from MODIS data, matching the observational residuals $(y - \hat{y})$ with the expected uncertainty in y.

2.3 Sequential methods

These include methods such as the Kalman filter and its variants [28] that operate at individual time /space step. In essence, they attempt to solve the same problem as the variational schemes but do so by breaking the problem up into small steps that converge to the same solution (or an approximation to it). They can be more flexible in dealing with non-linear effects and non-Gaussian distributions. When applied in multiple directions (in space/time) these methods are often known as *smoothers*, whereas if they are used in a single direction (e.g. predicting forward in time) they are known as *filters*. [29] used ensemble Kalman filters to assimilate satellite reflectance data into a vegetation process model to improve carbon flux estimates. [30] used similar techniques to assimilate snow observations into a process model. [31] used the related (but more flexible) method of particle filtering to assimilate microwave temperature data into a soil moisture dynamics model. One advantage of these methods is that adjoint code is not generally required. Of note also are Markov Chain Monte Carlo (MCMC) methods. These allow for the numerical solution of the posterior, using variations of the Metropolis algorithm [32]. The main benefit is their flexibility in combining different distributions (the data and/or prior need not be Gaussian, for example) and that no assumptions on the nature of the posterior are made (it is perfectly plausible to explore multimodal posterior distributions, for example). The main drawback is that MCMC algorithms in general take a long time to explore the solution space, therefore requiring many realisations of the dynamic model and observation operator. Additionally, convergence needs to be carefully monitored. In [33], MCMC methods are used to invert MODIS surface reflectance data using a canopy reflectance model.

A useful review article on the use of DA in estimating surface biogeophysical parameters is given by [34].

3. EO-LDAS

3.1 The EO-LDAS prototype

EO-LDAS is an ESA STSE-funded project to build a prototype Earth Observation Data Assimilation System that has recently been completed. The project is described in more detail in [35] and a tutorial for the use of the prototype software given in [36]. The software is soon to be released as a python package. Part of the motivation of the project was to build such a prototype tool to allow potential users to gain experience with using DA with EO data. The prototype software includes an interface to the top of canopy RT model described above (with associated adjoint code) which allow experiments such as the above MERIS and MODIS example to be conducted. There is also an interface to the 6s atmospheric code for ingesting top of atmosphere radiance, although this is slow as no adjoint is currently available. In addition, codes for linear kernel-driven BRDF models are included.

Although full access to the code is given, the most straightforward way of operating the code is through configuration files. In these, the user can read in observational data and set up a series of constraints for the DA. These would typically include: a background constraint by specifying *a priori* estimates of the model state with associated uncertainty; one or more observational constraints (one is set up for each sensor being used to facilitate sensor-specific configuration); and a (process) model constraint. The only process models implemented in the prototype so far are regularisation methods (N^{th} order difference constraints such as the zero order process model mentioned above) although the user can code or interface to their own models. Examples of this might include process models of vegetation dynamics, e.g. to interface to carbon flux calculations as in [16] or [29]. However, the *only* overlap between parameters of most state-of-the-art

vegetation dynamics models and those driving optical observation operators is LAI (or more precisely, the foliar carbon pool, which can be related to LAI), so even if we were to include such biogeochemistry models (as they are currently used) they would provide a constraint to only one of the terms linking to the observations. We would still then need to have some model (or make some further assumptions) regarding the other (7+) terms controlling reflectance to achieve the DA. For this reason, we have concentrated efforts at this stage on trying to derive a generic approach that can be used to estimate all state variables that affect the observations we are using. [29] discuss some of the other issues that must be considered when linking e.g. biogeochemical models with EO data in a DA system, including the need for a consistency in the assumptions made about the structural arrangement of vegetation in all models used. EO-LDAS implements a variational method. If the regularisation model constraint is used, then this is effectively a weak constraint DA system. Whilst this does not cover all of the approaches that might be used in DA it is a very useful starting point, particularly for users wishing to gain familiarity with the concepts. This is partly because it is straightforward to incorporate new constraints by adding new cost function terms as described above. The only real drawbacks are: (i) is that for it to be efficient, adjoint codes are needed if a model is non-linear; (ii) Gaussian statistics are assumed. This second restriction is somewhat mitigated against by allowing transformation functions to be defined for state variables (e.g. the exponentials given above), so a wide class of distributions can be used, provided they can be transformed by a continuous differentiable operator to a Gaussian. The (regularisation) process model in EO-LDAS is defined so that it can be applied in both space and time, although we have only explored the temporal aspect of this in any detail to date.

3.2 Application to MODIS data

Although it can do other things, the EO-LDAS software then is readily set up to ingest top of canopy spectral directional reflectance data, e.g. for a given location over some time period and apply prior and (regularisation) model constraints to provide an interpretation of the data in terms of the biophysical parameters driving the observation operator. Although a full 'validation' of such a system is hard to achieve, within the EO-LDAS project a comparison of LAI estimated using such an approach was performed, driven by MODIS (500m) observations. The field data were collected and processed by project partners from FSU Jena over large agricultural crop fields in Germany.

Fig. 5 shows envelopes of likely MODIS reflectance assuming the field LAI data to be true, using parameter ranges for the other observation operator variables taken from assumed prior distributions. We can see that there is generally quite a large range of reflectance values that could result from a given LAI. This is a useful first check before performing a DA exercise, as if observations are outside the envelopes (as with a few samples in bands 1 and 3) the DA *would not be able* to solve for the measured LAI values if these observations were matched closely.

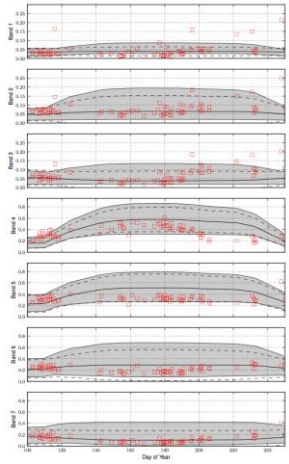


Figure 5. Panels of reflectance in MODIS wavebands as a function of day of year 2010 over the Gebesee site for a field of winter wheat. The grey envelopes show the domain of reflectance using field measured LAI and the a priori distributions of other parameters. The red marks indicate observed MODIS reflectance. The black solid line indicates the envelope mean.

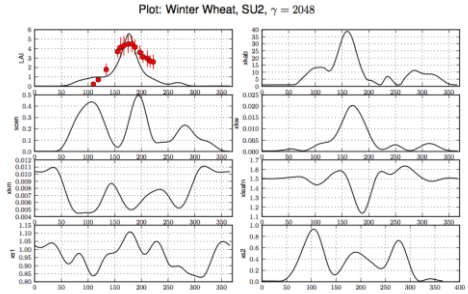


Figure 6. State estimates resulting from the EO-LDAS DA using MODIS data. The state for 8 parameters is sampled daily over a year, giving a state vector of around 3000 elements. Field measured LAI in red.

Fig. 6 shows the state vector resulting from the DA for a value of γ derived from generalised cross validation. The estimated LAI broadly matches that in the field data, but is clearly too 'peaky', probably as a result of over-smoothing here. Of more interest perhaps is the variation in the other parameters. Leaf Chlorophyll and water broadly matches the LAI trajectory, and there is an apparent decrease in leaf N (leaf structural complexity) with increasing LAI. The term xs2 broadly

corresponds to soil moisture in this experiment, which over the green vegetation period broadly matches the leaf water pattern. The leaf senescence term is high around day 100 which may be due to residual stubble in the field, but more interestingly has a strong peak soon after the time of maximum LAI.

3.2 Synthetic experiment for Sentinel-2 MSI

A set of experiments using EO-LDAS with simulated Sentinel-2 MSI data is described in detail in [12]. Temporal trajectories for the biophysical parameters are modelled (shown as a dotted green line in the figures below) and we attempt to retrieve these using synthetic MSI data. This is first attempted using MSI data for each observational day alone (akin to the use of MERIS data shown in Fig. 2), with results for a cloudy scenario with 50% of samples missing over the year shown in Fig. 7. When the zero-order process model is applied, we reduce the uncertainties by a factor of around 2 and also provide a continuous estimation of state. Although estimates of the biophysical parameters can be derived from the MSI data, the uncertainty (error bars show 95% C.I.) is high. The impact of DA even with this simple regularisation model is to dramatically improve the estimates whilst mostly keeping the true value within the C.I.

4. OUTLOOK

One of the major goals of terrestrial remote sensing has been to quantify the properties (state) of vegetation canopies. We have had access to RT tools to understand signals and interpret data for decades, but fast and simple VI methods still tend to dominate the field. As we move into an era where we are more concerned with estimating both state and uncertainty, we must start to recognise the impact of the assumptions made in interpretation as part of such error budgets. This requires more sophisticated statistical frameworks than have often been used in the past. This might be seen as a burden to producers of EO products, but rather affords great opportunities for *explicitly* applying multiple constraints on our interpretations. The information we estimate from EO is regularly and increasingly used to drive models (e.g. biogeochemistry models), but these models themselves can often provide information to help constrain estimates from EO. In many circumstances then it makes sense to combine the (process) modelling with interpretation of low level EO data such as ToA radiance or ToC reflectance. This can provide better consistency, more easily track uncertainty, allow constraint from model expectations, and ultimately better test and drive the models. This sort of task *requires* DA. We must also recognise that some terms that affect the EO data do not overlap with concepts in our process models, so we must pay attention to empirical approaches such as regularisation that can fit into the DA framework when we lack models or understanding of process.

EO-LDAS as a tool, is a start at exploring how we should be making better use of DA concepts in monitoring the land surface (and vegetation in particular) from EO. It can be applied practically to state estimation from (one or more) existing sensors, and we can have also demonstrated its use in exploring how we might improve mapping from future sensors such as Sentinel-2 MSI. The process models within the tool are limited at present, but there is plenty of scope for expanding this. Also, whilst interesting concepts such as spatial DA, which could form the basis for a multi-scale DA system are implemented, they have not been fully explored.

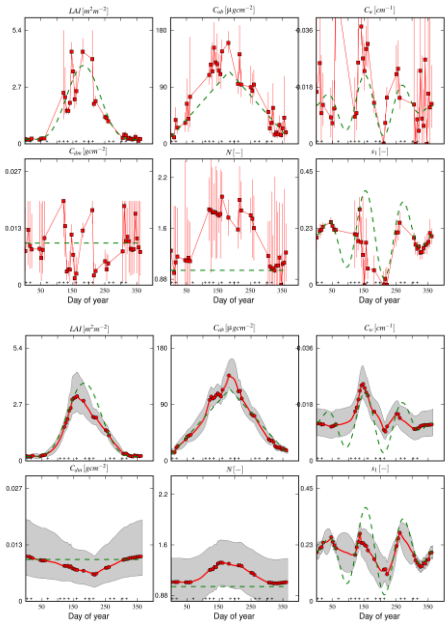


Figure 7. State vector estimate from EOLDAS for 6 biophysical parameters as a function of time derived from synthetic Sentinel-2 MSI data with EO-LDAS. (a) single date solutions; (b) regularised solutions.

References

[1] Enting, I.G., 2002. Inverse problems in atmospheric constituent transport. Cambridge University Press.

[2] Kimes, D.S., et al. 2000. Inversion methods for physically-based models. Rem. Sens. Rev. 18, 381–439.

[3] Berger, M., et al., 2012, ESA's sentinel missions in support of Earth system science, Rem. Sens. Environ.. 120, pp. 84-90.

[4] Goel, N.S., (1988), Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. Rem. Sens. Rev. 4, 1–212.

[5] Gobron, N., et al. 2002. Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications. IEEE Trans. Geosci. Rem. Sens., 38, 2489–2505.
[6] Baret, F. and Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Rem. Sens. Environ., 35, 161–173
[7] Knyazikhin, Y., et al. (1998) Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. J. Geophys. Res., 103:32,257-32,276.
[8] Gobron, N., et al. 1997. A semidiscrete model for the scattering of light by vegetation. J. Geophys. Res., 102, 9431–9446.
[9] Féret, J.B., et al. 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments, Rem. Sens. Environ 112, 3030–3043.
[10] Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Rem. Sens. Environ., 34, 75–91.
[11] Price, J.C., 1990. On the information content of soil reflectance spectra. Rem. Sens. Environ., 33, 113–121.
[12] Lewis, P. et al., 2012, An Earth Observation Land Data Assimilation System (EO-LDAS), Rem. Sens. Environ., 120, 219-235.
[13] Garrigues et al., 2008, Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, J. Geophys. Res., 113, G02028, doi: 0.1029/2007JG000635
[14] Ghil, M. and Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography. Adv. Geophys 33, 141–266.
[15] Rodgers, C.D., 2000. Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific Publishing Company.
[16] Knorr, W. et al. 2010. Carbon cycle data assimilation with a generic phenology model. J. Geophys. Res., 115, G04017.
[17] Xiao, Z., et al. 2009. A temporally integrated inversion method for estimating leaf area index from MODIS data. IEEE Trans. Geosci. Rem. Sens., 47, 2536–2545.
[18] Xiao, Z., et al. 2011. Real-time retrieval of Leaf Area Index from MODIS time series data. Rem. Sens. Environ., 115, 97-106.
[19] McLaughlin, D., 2002. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. Adv. Wat. Res., 25, 1275–1286.
[20] Giering, R. and Kaminski, T., 1998. Recipes for adjoint code construction. ACM Trans. Math. Soft. (TOMS) 24, 437–474.
[21] Lavergne, T., et al., 2007. Application to MISR land products of an RPV model inversion package using adjoint and Hessian codes. Rem. Sens. Environ., 107, 362–375.
[22] Qin, J., et al. 2007. A weak-constraint-based data assimilation scheme for estimating surface turbulent fluxes. IEEE Geosci. and Rem. Sens. Lett., 4, 649–653.
[23] Zhu, C., Byrd, R.H., Lu, P., Nocedal, J., 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans. Math. Soft. (TOMS) 23, 550–560.
[24] Pinty, B., et al. (2007) Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products, J. Geophys. Res., 112, D10116, doi:10.1029/2006JD008105.
[25] Clerici, M., et al. 2010. Consolidating the two-stream inversion package (JRC-TIP) to retrieve land surface parameters from albedo products. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 3, 286–295.
[26] Hansen, P.C. et al.. (2006) Deblurring images: matrices, spectra and filtering, Society for Industrial and Applied Mathematics (SIAM).
[27] Quaife, T., Lewis, P., 2010. Temporal Constraints on Linear BRDF Model Parameters. IEEE Trans. Geosci. Rem. Sens., 48, 2445–2450.
[28] Evensen, G., 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean dynamics 53, 343–367.
[29] Quaife, T., et al. 2008. Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. Rem. Sens. Environ., 112, 1347-1364.
[30] Slater, A.G. and Clark, M.P., 2009. Snow data assimilation via an ensemble Kalman filter. J. Hydromet., 7, 478-492.
[31] Qin, J., et al. 2009. Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. J. Geophys. Res. 114, D15103.
[32] Hastings, W.K., 1970, Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika 57 (1): 97–109. doi:10.1093/biomet/57.1.97
[33] Zhang, Q., et al., 2005, Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model, Remote Sensing of Environment, 99(3), pp 357-371.
[34] Liang, S., 2007, Recent developments in estimating land surface biogeophysical variables from optical remote sensing, Progress in Physical Geography (31) pp. 501-516
[35] http://www.assimila.eu/eoldas/
[36] http://www2.geog.ucl.ac.uk/~plewis/eoldas