VALIDATION OF THE EARTH OBSERVATION LAND DATA ASSIMILATION SYSTEM BY THE FIELD DATA OF ESA SPARC FIELD CAMPAIGN

ChernetskiyMaxim [1], Gomez-Dans Jose[2], Lewis Philip [2]

[1] Friedrich Schiller University, Institute of Geography, Department of Earth Observation, Jena, Germany, Email:maxim.chernetskiy@uni-jena.de
[2] NCEO & UCL, London, United Kingdom, Email: j.gomez-dans@ucl.ac.uk

ABSTRACT

The Earth Observation Land Data Assimilation System (EO-LDAS) project is uses the weak constraint variational data assimilation (DA) technique for the estimation of land surface parameters and their uncertainties by the remote sensing data. The main goal of the project is to make full use of different sources of optical sensors data, to provide improved estimation of structural and biophysical parameters of land surface. Therefore a software tool – the EO-LDAS prototype – was developed.

Within the frame of this work, the possibilities of EO-LDAS have been demonstrated for MERIS/Envisat and CHRIS/Proba data acquired during ESA SPARC 2004 field campaign over an agricultural test-site near Barrax (Spain).

We have used a regularization approach and conditions of spatial smoothness in order to better constrain the problem. The EO-LDAS prototype has been used to implement the weak constrain data assimilation (DA) system, to estimate leaf area index (LAI) and Chlorophyll (a + b) concentration as well as their uncertainties.

1. INTRODUCTION

A large number of sensors orbits the earth and provides observations at a variety of spatial and temporal resolutions and spectral windows. These observations require an interpretation, to extract useful information for the understanding and monitoring of the land surface. Typically, various interpretation approaches were applied on data of individual sensors to generate information on land surface parameters. With the advent of the Sentinel era, this approach is counter-productive.

Modern data assimilation techniques allow the use of physically based radiative transfer models to get a robust estimate of the state of the land surface (as well as detailed uncertainty information) conditional on all available observations.

In this contribution, we explore the recently published EO-LDAS (Earth Observation Land Data Assimilation System) as a way to combine optical data with different spatial resolutions and spectral characteristics[1]. The EO-LDAS framework provides a 4DVAR assimilation scheme with a weak constraint, resulting in an estimate of the state that is constrained by both the observations (using a suitable radiative transfer model) and prior knowledge of the state values or their spatial/temporal evolution. While vegetation models could be used to predict the temporal trajectory of some states such as leaf area index (LAI), for other components of the state that are required to model the observed signals, these models do not exist. Similarly, no spatial models are available.

In the previous publication, it was found that the EO-LDAS prototype is able to simultaneously estimate a state vector of over 2000 elements of biophysical characteristics in the synthetic experiment. It was demonstrated the reduction in uncertainties of the estimation of parameters in a temporal sense. Further it was noted that it is possible to extend EO-LDAS to spatial constraints but it wasn't explored at those time[1].

In order to solve the problem the spatial regularization approach was proposed. Validation of the proposed approach has been done with data from the ESA Barrax test site.

2. THE EO-LDAS SCHEME

EO-LDAS was developed to solve the problem of estimation of the state of the land surface by using all given observations as well as all other sources of information. Each source of information is represented by a probability density function (PDF). A result of combination of these PDF is *a posteriori* PDF - the solution of the problem. Thus, posterior probability is determined by the minimization of the cost function in the form[1]:

$$J(x) = J_{space}(x) + J_{obs}(x) + J_{prior}(x) \quad (1)$$

Where *a priori* information constraint:

$$J_{prior}(x) = -\frac{1}{2}\left(x - x_p\right)^T C_p{}^{-1}(x - x_p) \quad (2)$$

an observational constraint:

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

$$J_{obs}(x) = -\frac{1}{2}\left(x - H(x)\right)^T C_{obs}{}^{-1}\left(x - H(x)\right) \quad (3)$$

where $H(x)$– observational operator. In this case a semidiscrete model for the scattering of light by vegetation [2]. C_p, C_{obs} – covariance matrices which describe the uncertainties in the prior information and in the observations.

The last term is a spatial smoothness constraint:

$$J_{space}(x) = \gamma_x \cdot \left\|\Delta_x x_{i,j}\right\|^2 + \gamma_y \cdot \left\|\Delta_y x_{i,j}\right\|^2 \quad (4)$$

where $\Delta_{x,y}$ is smoothness operators for x and y direction of smoothing. Spatial smoothness is weighted by the corresponding regularization parameter γ or in other words the model error. This parameter is determines to which extent we trust to the expectation of smoothness. Smoothness constraint is carried out per parameter. This means that for instance there is an expectation of LAI varying spatially smoothly independent of the smooth variation of chlorophyll content.

In the case of multi-resolution processing eq. 1becomes

$$J(x) = J_{space}(x) + J_{hires}(x) + J_{lores}(x) + J_{prior}(x) \quad (5)$$

where $J_{hires}(x)$, $J_{lores}(x)$ are observational constraints for high resolution and low resolution sensors respectively.

3. REGION OF INTEREST

This study demonstrates the spatial capabilities of EO-LDAS. As such, we think that concentrating efforts on a couple of agricultural fields will be a reasonable choice of sites to work on. Barrax is a widely-studied area, where many field campaigns have taken place over the years [3]–[6]. These campaigns have mostly dealt with retrieval of biophysical parameters from optical data, and as such, it is common for measurements of LAI, Chlorophyll concentration, leaf equivalent water thickness, leaf dry matter and soil spectral measurements to be available. There is also a large amount of EO data that has been collected by many sensors, and the area is typically cloud free in summer. This wealth of data allows one to test the impact of cloudiness or sparse temporal sampling by simply applying retrieval algorithms to a part of the dataset and validating on the remaining observations. For all this, Barrax is a very good site to try to understand spatial capabilities of EO-LDAS.

4. DATA

In this study, we used data of PROBA\CHRIS and Envisat\MERIS. CHRIS – Compact High Resolution Imaging Spectrometer has 62 spectral bands from 400 nm to 1050 and a spatial resolution of 36 m. The images are acquired in sets of 5 zenith angles: +/-55, +/-36 and nadir. In this work we used 3 sets: nadir, $+36^0$, $+55^0$. Due to high correlation between hyperspectral bands, the number of information is much lower than the number of bands. Because of this fact the number of bands was reduced to 17 according to [7].

MERIS – medium spectral resolution, imaging spectrometer has 15 bands and a spatial resolution of 300 m.

The prior information used in this study had very high standard deviations (tab. 1). Hence, the problem is not constrained to any specific values of the parameters. The reason is that we modeled the situation when very low amount of information about the land surface is available. However, the prior information can decrease chances to be trapped to a local minimum because it changes the parameter space[7].

Table 1. The prior information

Parameter	**Mean; standard deviation**
LAI, the single sided leaf area per unit ground area	5.99; +6.0/-9.2
The canopy height, m	0.1; +/- 1
Leaf radius/ dimension, m	0.01; +/- 1
The concentration of chlorophyll a+b, µg/cm^2	230; +239.8/-773.7
The proportion of senescent material	0.001; +/- 1
Equivalent leaf water, cm	0.0002; +0.01/-0.23
Dry matter, µg/cm^2	0.01; +0.01/-0.1
The number of leaf layers	1.5; +/- 1
Soil PC1 (soil brightness)	0.05; +/- 1
Soil wetness	0.005; +/-1

5. RESULTS

The results of pixel by pixel inversion with the prior information showed that despite of independent processes of inversion for each pixel we can see spatial structure of the considered area for all biophysical parameters. Fig. 1A-B demonstrates an example for LAI and chlorophyll. In addition the values of the parameters are in range of the true values. It can be seen that the canopy RTM is not able to retrieve chlorophyll content for areas of the bare soil (Fig. 1).

The hypothesis was that spatial regularization and even not strict prior information can self-contain the problem in such a way that the values of the results will be shifted to the range of the true values. In order to show it the problem was solved for several Barrax fields.

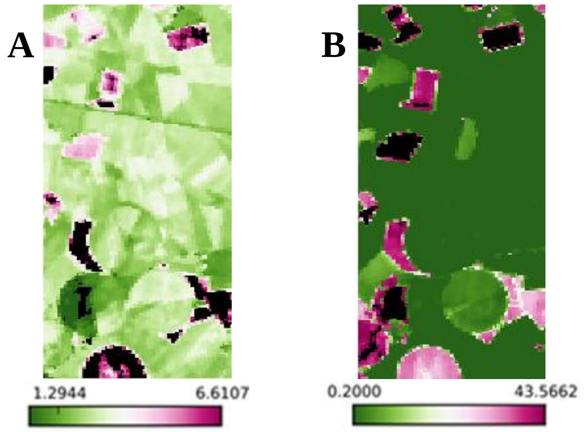

Figure 1. Results of pixel by pixel inversion. LAI (A) and chlorophyll content (B). 62 bands of CHRIS/Proba on16/07/2004 with prior information.

In the first case, EO-LDAS inversion without spatial regularization and without prior information was applied for each pixel of a field where in situ data were available. Inversion was done independently for three CHRIS/Proba cameras. LAI $r^2 = 0.51 - 0.74$, $p = 0.000001 - 0.008$;Chl r^2=0.64-0.72, p=0.003-0.03. However, the uncertainties were huge +/- 10 for LAI and +/- 400 for chlorophyll. That was to be expected because we tried to estimate 12 parameters by quite highly correlated spectral bands using only one set of view angles. In fig. 2, the results for the chlorophyll content estimation are shown in logarithmic scale i.e.$e^{-Chl/100}$.

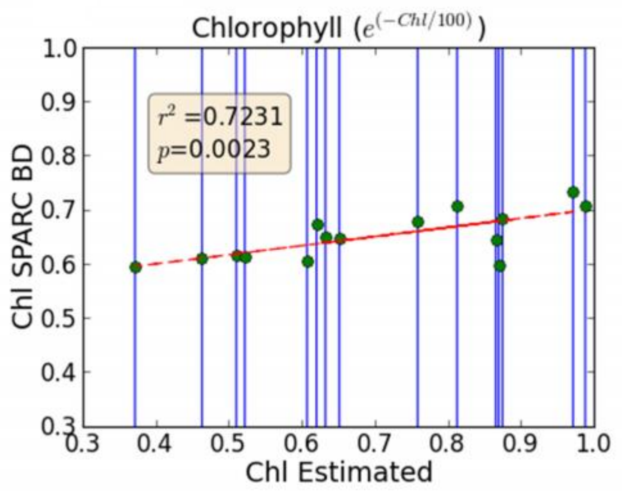

Figure 2. Validation of the Chlorophyll a+b content retrieval. $J=J_{obs}$ - w/o prior information. Logarithmic scale. Blue lines – std. dev. CHRIS/Proba 36^0.

In the second case the inversion with spatial regularization and with the priors was applied. In this case LAI r^2 = 0.71 - 0.8, p=0.000001-0.0004, Chl r^2 = 0.62-0.86, p=0.000001-0.3. The LAI uncertainties = +/-4, Chl uncertainties = +/-50. There is significant decreasing in uncertainties and increasing of the estimation accuracy. Fig. 3 shows the results for CHRIS/Proba camera $+36^0$.

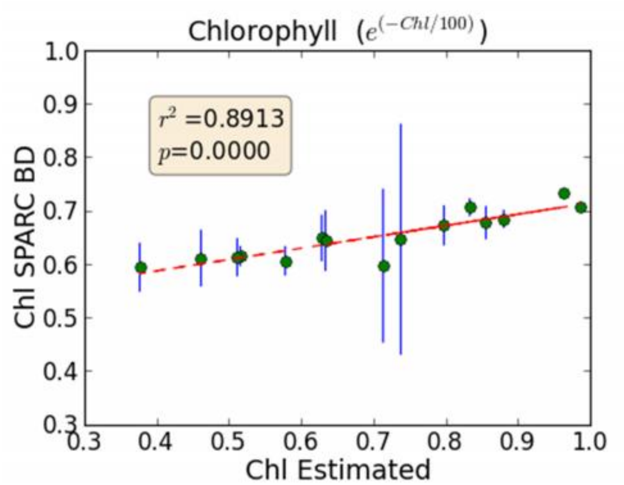

Figure 3. Validation of the Chlorophyll a+b content retrieval. $J=J_{obs}+J_{prior}$ – with prior information. Logarithmic scale. Blue lines – std. dev. CHRIS/Proba $+36^0$.

The above results show that EO-LDAS is able to estimate some of the biophysical parameters on pixel-by-pixel basis. The prior information can help to decrease the number of possible inversion solutions. However, the uncertainties are still quite big and in the case of availability of only single observation, we cannot apply constraining by a dynamical model. A possible solution is the use of the spatial constraints (eq. 4).

For the spatial constraint problem due to speed limitations of the prototype software the number of the CHRIS/Proba bands were decreased to 4 – 452, 553, 683 and 890 nm

In fig. 4A and 5A an example of the spatial solution without constraining by space is given for the cornfield C9. For this field we have 2 points of the field measurements. The in situ values of LAI are in the range of 2.92 to 3.1 and the chlorophyll (Chl) content is about 52.94 mg/cm^2. The estimation of the parameters for this case: LAI = 3-4.5, Chl = 60- 78 mg/cm^2.The large values of the uncertainties can be seen for the whole area of the field.

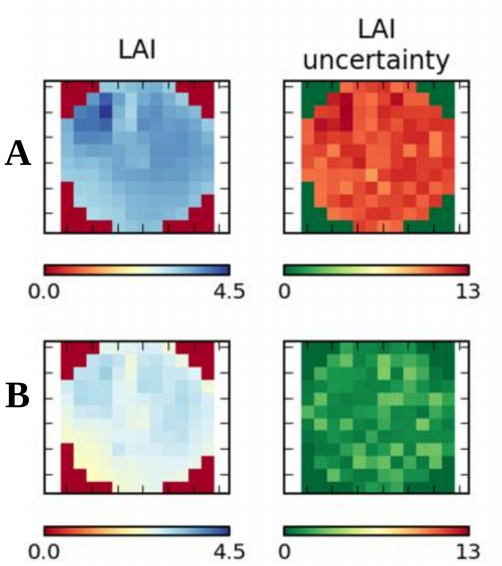

Figure 4. Estimation of LAI, CHRIS/PROBA (4bands). Corn, Field C9. A - without priors and spatial regularization (J=Jobs), B – with priors and spatial regularization ($J=J_{obs}+J_{prior}+J_{model}$).

In the second case i.e. with constraining by space LAI = 2.5-3.5, Chl=50- 60 mg/cm^2. I.e. solution after applying spatial regularization and the prior information was shifted to the range of the true values. The values of uncertainties of the second-case estimation decreased significantly (fig. 4B, 5B).

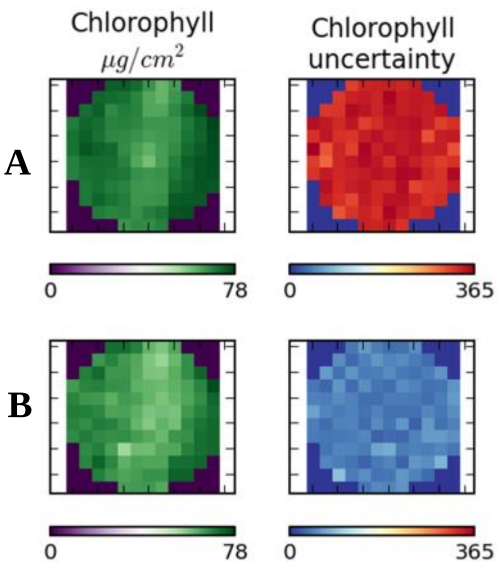

Figure 5. Estimation of the Chlorophyll a+b content, CHRIS/PROBA (4bands). Corn, Field C9. A - without priors and spatial regularization ($J=J_{obs}$), B – with priors and spatial regularization ($J=J_{obs}+J_{prior}+J_{model}$).

The next step is to model the situation when two sensors with different resolution are available. The first sensor has high spatial resolution but only a few bands and the second one has low resolution but more spectral bands. For this constellation, we use CHRIS/Proba with a reduced number of bands (near infrared (NIR) and Red) and MERIS/Envisat with 15 bands. On the one hand, MERIS does not have enough spatial information for spatial constraining because of the relatively small Barrax fields. I.e. there are no homogeneous pixels, which can constrain each other (fig. 6B, 7B).

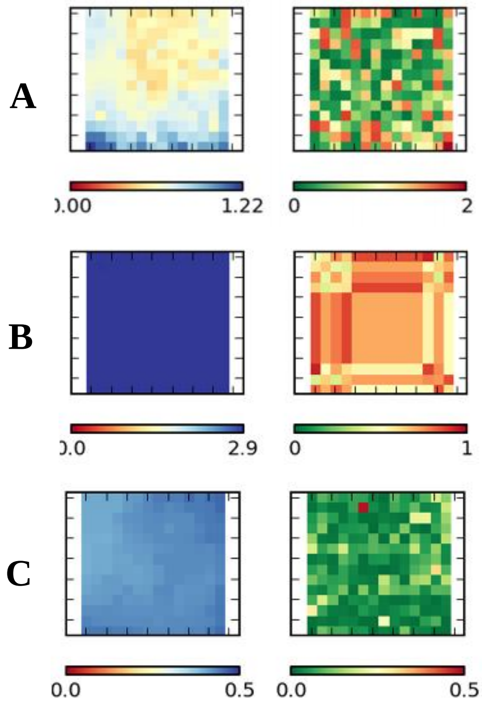

Figure 6. Estimation of LAI. Sunflower, Field SF1. A – CHRIS/Proba, B – MERIS, C - CHRIS/Proba + MERIS.

On the other hand, two bands of the high-resolution sensor do not have enough information content for the model inversion (fig. 6A, 7A). In both cases, values of chlorophyll content are quite far away from the truth (fig. 6A-B, 7A-B). However, after solving the problem in a way of eq. 5 values of chlorophyll were shifted to the range of the true values (fig. 6C, 7C). I.e. both sensors did not have true values in the estimation but after combining them together optimal decision was found.

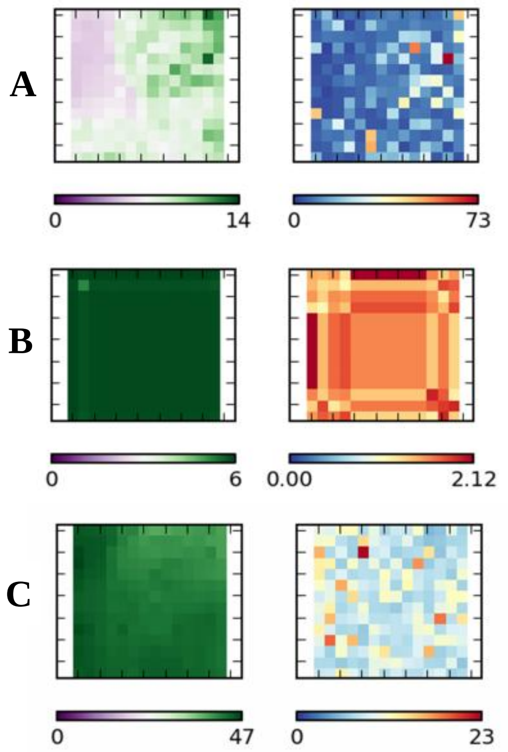

Figure 7. Estimation of the Chlorophyll content. Sunflower, Field SF1. A – CHRIS/Proba, B – MERIS, C - CHRIS/Proba + MERIS.

6. DISCUSSION AND CONCLUSIONS

The main goal of the paper was the exploration of spatial constraining possibilities of EO-LDAS. In the frame of this work, small homogeneous fields were used. The computational costs of the current EO-LDAS implementation do not allow obtaining results for bigger fields. However, we assumed that each field was homogeneous and it is possible to extend values of one-two measurements to a whole field. In this research, we did not take into account the edge problem but in the case of data that are more realistic, we should manage it or do a kind of clusterization before estimation of the parameters.

Another complication of this study is the estimation of γ. The best way to do it is the cross-validation. However, it is not possible to do now due to speed limitations.

Despite of some simplicity of the input data the spatial regularization has shown significant decreasing of the uncertainties in the estimated data as well as improvement in the accuracy.

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support of this project through GIONET, funded by the European Commission, Marie Curie Programme Initial Training Network, Grant Agreement number PITN-GA-2010-264509. We also acknowledge European Space Agency for providing data access by the "EO-Support" system (projects ID 13803 and 13931).

REFERENCES

[1] P. Lewis, J. Gómez-Dans, T. Kaminski, J. Settle, T. Quaife, N. Gobron, J. Styles, and M. Berger, “An Earth Observation Land Data Assimilation System (EO-LDAS),” *Remote Sens. Environ.*, vol. 120, pp. 219–235, May 2012.

[2] N. Gobron, B. Pinty, M. M. Verstraete, and Y. Govaerts, “A semidiscrete model for the scattering of light by vegetation,” *J. Geophys. Res.*, vol. 102, pp. 9431–9446, 1997.

[3] S. Gandia, G. Fernández, J. C. García, and J. Moreno, “Retrieval of vegetation biophysical variables from CHRIS / PROBA data in the SPARC campaign,” in *Proc. 2nd CHIRS/Proba Workshop, ESA/ESRIN, Frascati, Italy*, 2004, no. July.

[4] M. Gonzalez-Sanpedro, T. Le Toan, J. Moreno, L. Kergoat, and E. Rubio, “Seasonal variations of leaf area index of agricultural fields retrieved from landsat data,” *Remote Sens. Environ.*, vol. 112, no. 3, pp. 810–824, 2008.

[5] A. Verger, B. Martinez, F. Camacho-de Coca, and F. Garcia-Haro, “Accuracy assessment of fraction of vegetation cover and leaf area index estimates from pragmatic methods in a cropland area,” *Int. J. Remote Sens.*, vol. 30, no. 10, pp. 2685–2704, 2009.

[6] J. Delegido, C. Vergara, J. Verrelst, S. Gandia, and J. Moreno, “Remote estimation of crop chlorophyll content by means of high-spectral-resolution re ectance techniques,” *Agron. J.*, vol. 103, no. 6, pp. 1834–1842, 2011.

[7] F. Vuolo, L. Dini, and G. D’Urso, “Assessment of LAI retrieval accuracy by inverting a RT model and a simple empirical model with multiangular and hyperspectral CHRIS/PROBA data from SPARC,” in *Proc. of the 3rd ESA CHRIS/Proba Workshop, 21–23 March, ESRIN, Frascati, Italy, (ESA SP-593, June 2005)*, 2005.