Retrieving spectral and biophysical parameters of land vegetation by the Earth Observation Land Data Assimilation System

Dissertation

(kumulativ)

zur Erlangung des akademischen Grades doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt dem Rat der Chemisch-Geowissenchaftlichen Fakulat der Friedeich-Schiller-Universitat Jena

von Chernetskiy Maxim

geboren am 14.04.1978 in USSR

Gutachter:

1. Prof Dr. Christiane Schmullius, Institut für Geographie, Friedrich-Schiller-Universität
2. PD. Dr. Angela Lausch, Helmholtz-Zentrum für Umweltforschung GmbH – UFZ

(3. …….)

Tag der Verteidigung: ..21.09.2018……..

List of manuscripts

Appended papers

Chernetskiy, M., Gobron, N., Gomez-Dans, J., Morgan, O., Disney, M., Lewis, P., Schmullius, C. 2018. Simulating arbitrary hyperspectral bandsets from multispectral observations via a generic Earth Observation-Land Data Assimilation System (EO-LDAS). *Advances in Space Research* (In Press, Available Online, DOI: 10.1016/j.asr.2018.07.015).

Chernetskiy, M., Gomez-Dans, J., Gobron, N., Morgan, O., Lewis, P., Truckenbrodt, S. and Schmullius, C., 2017. Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing, 9(7), 656.

Chernetskiy, M., Gobron, N., Gómez-Dans, J., Lewis, P. and Schmullius, C.C., 2017. Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site. In Earth Observation for Land and Emergency Monitoring. John Wiley & Sons, Ltd, 117–139.

Additional publications

Shevyrnogov, A., **Chernetskiy, M.**, Vysotskaya, G. Multiyear trends of Normalized Difference Vegetation Index and temperature in the south of Krasnoyarsk Krai. Izvestiya Atmospheric and Oceanic Physics, 49, 9 (2013), 1047–1056.

Chernetskiy, M., Pasko, I., Shevyrnogov, A., Slyusar, N., Khodyayev, A. A study of forest vegetation dynamics in the south of the KrasnoyarskiiKrai in spring. Advances in Space Research, 48 (2011) 819–825.

Sentinel-3 for science workshop, Venice (Italy), June 2-5, 2015
Poster presentation: *Retrieval of biophysical canopy parameters on Sentinel-3 validation test sites using the Earth Observation Land Data Assimilation System and multiangular information of MISR* **Chernetskiy, M.**, Gobron, N., Morgan, O., Gomez-Dans, J., Lewis, P., Gitelson, A., Schmullius, C.

The 4th International Symposium on Recent Advances in Quantitative Remote Sensing (RAQRS), Valencia (Spain), September 22-26, 2014
Poster presentation and Proceedings: *Effectiveness of MISR multiangular observations in the new generation of Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P.

40th COSPAR Scientific Assembly, Moscow (Russia), August 2 – 10, 2014
Oral presentation: *Reconstruction of hyperspectral CHRIS/PROBA signal by the Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P.

GIONET Summer School, Frascati (Italy), June 30 - July 4, 2014

Poster presentation: *Earth Observation Land Data Assimilation System (EO-LDAS) regularization constraints over Barrax site*
Chernetskiy, M.

Sentinel-2 for science workshop, Frascati (Italy), May 20 - 22, 2014
Poster presentation and Proceedings: *Validation of effectiveness of MISR multiangular data in the Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P., Schmullius, C.

ESA Living Planet Symposium, Edinburgh (UK), September 9 - 13, 2013
Oral presentation and Proceedings: *Validation of the Earth Observation Land Data Assimilation System by the field data of ESA SPARC field campaign*
Chernetskiy, M., Gomez-Dans, J., Lewis, P., Schmullius, C.

ESA EO Summer School 2012 on Earth system monitoring and modelling, Frascati (Italy), July 30 – August 10, 2012
Poster presentation: *Multi-scale remote sensing synergy for land process studies: increasing the revisit frequency of high resolution sensors.*
Chernetskiy, M.

39th COSPAR Scientific Assembly, Mysore (India), July 14 - 22, 2012
Oral presentation: *Multi-scale remote sensing synergy for land process studies: blending of high and low resolution imagery*
Chernetskiy, M.

The 18th Biennial Conference of International Society for Ecological Modelling
In Procedia Environmental Sciences 13 (2012) 194 – 201: *Boreal forests contribution to global seasonal dynamic of carbon dioxide in the atmosphere*
Bartsev, S., Degermendzhi, A., Ivanova, Y., Shchemel, A., **Tchernetsky M.**

Symbols and abbreviations

List of symbols

Symbol	Unit	Description
C		Covariance matrix
H(χ)		Observational operator
J		Cost function
J_{model}		Dynamic model term of cost function
J_{obs}		Observational term of cost function
J_{prior}		Prior term of cost function
Θ		View zenith angle
$Θ_0$		Sun zenith angle
$μ_0$		Cosine of sun zenith angle
ϕ		View azimuth angle
$ϕ_0$		Sun azimuth angle

χ	State vector
χ_p	Prior state vector
ξ	Clumping (structural) factor

List of abbreviations

Abbreviation	Description
BRDF	Bidirectional Reflectance Distribution Function
BRF	Bi-directional Reflectance Factor
C_{ab}	Chlorophyll a+b concentration
ECV	Essential Climate Variable
EO	Earth Observation
EO-LDAS	Earth Observation Land Data Assimilation System
ETM+	Landsat Enhanced Thematic Mapper Plus
FAPAR	Fraction of Photosynthetically Active Radiation
GCOS	Global Climate Observing System
GEOSS	Global Earth Observation System of Systems
GPE	Gaussian Process Emulator
LAD	Leaf Angle Distribution
LAI	Leaf Area Index
LHS	Latin hypercube sampling
LUT	Look Up Table
MERIS	Medium Resolution Imaging Spectrometer
MISR	Multiangle Imaging Spectroradiometer
MODIS	Moderate Resolution Imaging Spectro-radiometer
NIR	Near Infrared
OLI	Landsat Operational Land Imager
PAR	Photosynthetically Active Radiation
PDF	Probability Density Function
POLDER	Polarization and Directionality of the Earth's Reflectances
RMSE	Root Mean Square Error
RTM	Radiative Transfer Model
SAIL	Scattering by Arbitrarily Inclined Leaves
SPOT	Satellite Pour l'Observation de la Terre
TIP	Two Stream Inversion Package
VI	Vegetation Index

Summary

In this thesis, a new methodology for retrieval of land spectral and biophysical parameters from optical remote sensing data has been designed and used. The result of the work was a physically based methodology for FAPAR and LAI retrievals, simulation of hyper-spectral information and estimation of associated uncertainties. The parameters derived by presented methodology are self-consistent, i.e., the same physical (RT) model assumptions are made to retrieve LAI, FAPAR and simulated spectral information.

The presented methodology is based on the generic Earth Observation-Land Data Assimilation System (EO-LDAS) which is a system that interprets spectral observations and provides an optimal estimate of the land surface state. It allows to combine observations from various sensors with different spatial and spectral resolution and repetition frequencies. EO-LDAS is based on variational Data Assimilation (DA) and exploit physically-based RTM to transform from state to observation space.

In the course of the work the following answers to the *research questions* were found:

1. *How to retrieve continuous FAPAR and its associated uncertainties if number of available observations are low.*

EO-LDAS can be used for daily estimation of FAPAR and associated uncertainties without any in-situ information and when the number of available observations is low. The results were in line with the field measurements with r^2 varying from 0.84 to 0.92 and RMSE from 0.11 to 0.16. This was the highest rate among compared products (JRC-TIP, MERIS FR and MODIS MCD15). This outcome demonstrated that EO-LDAS is able to estimate absorbed fluxes relying only on multi-spectral information. The use of temporal regularisation provides a controlled way of interpolating over data-free periods and the ability to track fast phenological changes phenology. This was especially important in the case of e.g. MISR observations, which were temporally sparse. Temporal regularisation was able to capture the leading and ending edges of the FAPAR curve in all of the fields and for most of the years. Furthermore, since EO-LDAS generated results for each day-of-year, it showed ability to predict FAPAR between available satellite observations: this may be a new way for performing space product validation in absence of ground-based measurements.

2. *How to retrieve consistent time series of biophysical parameters using EO-LDAS?*

Results showed that EO-LDAS can be used for retrieval of different parameters providing consistency between them. Thus, in this study daily LAI, leaf chlorophyll content and fraction of senescent leaves were estimated. Estimation of LAI showed high correlation with ground truth measurements but with some underestimation for high LAI values. Retrieved time series of leaf chlorophyll content and fraction of senescent leaves showed a realistic temporal trajectory which was consistent with dynamics of LAI and FAPAR.

3. *How does inclusion of static or dynamic prior information to EO-LDAS influence retrieval of LAI and FAPAR from CHRIS/PROBA, MISR and Landsat data?*

It was shown, that using MISR information, EO-LDAS temporal regularization and generic dynamic prior, it was possible to stabilize results of the retrieval and to obtain better results than MERIS FAPAR or JRC-TIP MISR. In addition, inclusion of generic static and dynamic prior information, decreases posterior uncertainties and can increase accuracies compared to in-situ data.

4. *What effect has multi-angular information of MISR and CHRIS/PROBA on retrieval of biophysical parameters over agricultural fields?*

It was found, that sequential increase of MISR multi-angular observations improves the RMSE of results based on synthetic data. Calculation of the cross entropy did not show, that multi-angular observations always provide more information than the same number of mono-angular nadir observations. This suggested that - at least in case of this experimental setup - the number of observations was of greater significance than different viewing angle positions.

5. *What are the major mechanisms controlling spectral variations in the visible and near infrared regions over agricultural fields? Can these mechanisms be parameterized by multispectral information of EO sensors?*

The results showed that proper estimation of LAI and soil parameters were sufficient to simulate a hyper-spectral signal between 400 and 1000 nm with acceptable precision: best RMSE is equal to 0.03 for real data and less than 0.008 for synthetic data. This implies that in case of the given experimental set-up, LAI and soil parameters are the major mechanisms controlling spectral variations in the visible and near infrared regions.

6. *Is EO-LDAS able to simulate a hyper-spectral signal for a range of crops relying only on multispectral bands in visible and NIR region?*

Results showed, that the methodology can be used to simulate arbitrary band sets in the visible and NIR regions, which are of importance to any application which requires combining new and existing streams of EO data in the optical domain, particularly inter-calibration of EO satellites in order to get continuous time series of surface reflectance, across programmes and sensors of different designs.

7. *How does the combination of different optical sensors in EO-LDAS influence the retrieval of biophysical parameters?*

Merging information of ETM+ and multi-angular data of MISR can increase accuracies and decrease a-posteriori uncertainties of hyper-spectral data simulations. These results demonstrate a new possibility to combine different satellite constellation sensors.

The presented work concentrates on the retrieval land parameters from optical sensor data using the data assimilation framework. In general presented methodology shows good performance however, there are many directions for further development. Currently the system exploits Gaussian Process (GP) emulators (Gómez-Dans et al. 2016) which significantly improve speed of the processing but it is still too slow for global applications with spatial

resolution more than 1 km. One way to improve processing speed is to use Discrete Cosine Transform (DCT) when we can represent remote sensing data as a sum of cosine functions with different frequencies which in theory allows to perform fast inversions. Another way to improve speed of the processing is using linearization of a RT model's signal. This is possible because first and second derivatives are the side products of the GP emulation.

Currently EO-LDAS uses 1D RT models, such as PROSAIL and semidiscrete, which are practical but these type of models significantly simplify canopy structure. However, exploitation of GP emulators open a way for emulation of 3D RT models such as Drat (Lewis, 1999) or Raytran (Govaerts and Verstraete, 1998) which can be very important for retrieval of parameters from high resolution sensors such as Sentinel-2 MSI. Because in case of higher resolution it is not always possible to represent canopy as a turbid medium.

All published results of EO-LDAS concern only top-of-canopy level which is easier to implement. The reason is that only three models have to be coupled: canopy (semidiscrete, SAIL), leaf (PROSPECT) and soil (Price, Hapke) (Verhoef, 1984, Gobron, 1997, Pinty et al. 1989). However this leads to incorporation of uncertainties from atmospheric correction which are not always easily trackable. The solution for this is implementing combination of atmospheric (6S), canopy, leaf and soil models (Vermote, 1997). The first experiments for coupling are vegetation and atmospheric models in EO-LDAS are not published but showed promising results.

EO-LDAS uses "today as tomorrow" process model for regularization of the inversion process. This assumption works well in many cases of natural development of vegetation however it is not realistic in case of abrupt changes such as fires, clear cuts, etc. One way of taking into account abrupt changes is to use edge preserving regularization which was implemented for inversion of linear kernel models for fire monitoring purposes. However, due to speed issues was not used for EO-LDAS.

The other way of avoiding unrealistic other smoothing is to use ecological process models for some state parameters such as LAI or chlorophyll.

An important part of EO-LDAS is an observational operator which currently is a RT model in optical domain. However, potentially EO-LDAS can be used for merging information of different electromagnetic domains such as thermal infrared and passive and active microwave. The main complication for ingestion SAR data into EO-LDAS scheme is lack of stable and mature physical models. The other complication is difference in input parameters of such models. This means potential doubling in number of parameters which in turn will lead to increasing of "ill-posedness" of the problem. An additional issue is difference in assumptions on LAI. However, despite of all these complications the problem of combining SAR and optical data is solvable by using newly developed models and including additional constraints.

Zusammenfassung

In dieser Arbeit wurde eine neue Methode zum Abruf von spektralen und biophysikalischen Parametern aus optischen Fernerkundungsdaten entwickelt und angewendet. Das Ergebnis der Arbeit war eine physikalisch basierte Methodik für FAPAR- und LAI-Abrufen, die Simulation von Hyperspektralinformationen und die Schätzung von Unsicherheiten, verbunden mit den abgerufenen Parametern. Die von der vorgestellten Methodik abgeleiteten Parameter sind selbstkonsistent, d.h. dieselben physikalischen (RT) Modellannahmen wurden gemacht, um LAI-, FAPAR- und simulierte Spektralinformationen abzurufen**.**

Im Laufe der Arbeit wurden folgende Antworten auf die *Forschungsfragen* gefunden:

1. *Wie können die kontinuierliche FAPAR und die damit verbundenen Unsicherheiten abgerufen werden, wenn die Anzahl der verfügbaren Beobachtungen gering ist.*

EO-LDAS kann zur täglichen Schätzung von FAPAR und damit verbundenen Unsicherheiten ohne In-situ-Information und wenn die Anzahl der verfügbaren Beobachtungen gering ist, verwendet werden. Die Ergebnisse waren in Übereinstimmung mit den Feldmessungen mit r^2 variierend zwischen 0,84-0,92 und RMSE von 0,11 bis 0,16. Dies war die höchste Rate unter den verglichenen Produkten (JRC-TIP, MERIS FR und MODIS MCD15). Dies zeigte, dass EO-LDAS in der Lage war, die absorbierten Strömungen abzuschätzen, stützend sich nur auf multispektrale Information. Die Verwendung der temporale Regularisierung, bot eine kontrollierte Methode der Interpolation über freie Daten Perioden an und die Fähigkeit, die schnellen Änderungen der Phänologie zu verfolgen. Dies war besonders wichtig bei z. B. MISR-Beobachtungen, die temporal spärlich waren. Die Temporale Regularisierung war in der Lage, die Führende**-** und Endkanten der FAPAR-Kurve in allen Bereichen und für die meisten Jahre zu erfassen. Weiterhin, weil EO-LDAS Ergebnisse für jeden Tag des Jahres generierte, zeigte es die Fähigkeit, FAPAR zwischen verfügbaren Satellitenbeobachtungen zu prognostizieren: Dies könnte eine neue Methode sein, die Validierung der Weltraumprodukten in Abwesenheit von bodenbasierter Messungen zu ermöglichen.

2. *Wie können konsistente Zeitreihen von biophysikalischen Parametern mit EO-LDAS abgerufen werden?*

Die Ergebnisse zeigten, dass EO-LDAS für den Abruf verschiedener Parameter verwendet werden kann, um die Konsistenz zwischen ihnen zu gewährleisten. So wurden in dieser Studie tägliche LAI, Blattchlorophyllgehalt und Anteil der seneszenten Blätter geschätzt. Die Schätzung von LAI zeigte eine hohe Korrelation mit der Messung der Bodenwahrheit, jedoch mit einer zu hohen Schätzung für hohe LAI-Werte. Wiedergefundene Zeitreihen des Blattchlorophyllgehaltes und des Anteils seneszenten Blätter, zeigten eine realistische zeitliche Flugbahn, die mit der Dynamik von LAI und FAPAR übereinstimmte.

3. *Wie beeinflusst der Einschluss von statischen oder dynamischen Vorabinformationen in EO-LDAS die Abfrage von LAI und FAPAR aus CHRIS / PROBA-, MISR- und Landsat-Daten?*

Es wurde gezeigt, dass mithilfe von MISR-Informationen, EO-LDAS-Regularisierung und generischer dynamischer Priorisierung, die die Vegetationsstruktur und die optischen Eigenschaften der Blätter detaillieren, war möglich, die Ergebnisse des Abrufs zu stabilisieren und bessere Ergebnisse als MERIS FAPAR oder JRC-TIP MISR zu erzielen. Darüber hinaus wurde gezeigt, dass die Aufnahme generischer, statischer und dynamischer Vorinformationen, verringert die posterioren Unsicherheiten und erhöht die Genauigkeit des Vergleichs mit In-situ-Daten.

4. *Welchen Effekt haben Multi-Angular-Informationen von MISR und CHRIS / PROBA auf den Abruf der biophysikalischer Parametern über landwirtschaftliche Felder?*

Es wurde festgestellt, dass eine erhöhte Anzahl von Mehrwinkelbeobachtungen (von MISR bereitgestellt) half nur zur Verbesserung der Genauigkeit (RMSE) im Fall der synthetischen Simulationen und nur dann für Cab. Die Berechnung der Kreuzentropie für steigende Zahlen der MISR-Kameras zeigte nicht, dass Mehrwinkelbeobachtungen immer mehr Informationen lieferten als die gleiche Anzahl von monogonalen Nadirbeobachtungen (Fig. 12). Dies deutet darauf hin, dass zumindest im Falle dieses Versuchsaufbaus, die Anzahl der Beobachtungen von größerer Bedeutung war als die verschiedenen Blickwinkelpositionen.

5. *Was sind die Hauptmechanismen, die Spektralvariationen im sichtbaren und nahen Infrarotbereich über landwirtschaftlichen Feldern steuern? Können diese Mechanismen durch multispektrale Informationen von EO-Sensoren parametrisiert werden?*

Die Ergebnisse zeigten, dass eine korrekte Abschätzung der LAI- und Bodenparameter für die Simulation des hyperspektralen Signals zwischen 400 und 1000 nm mit akzeptabler Genauigkeit von: der besten RMSE-Wert 0,03 für reale Daten und weniger als 0,008 für synthetische Daten ausreichte. Dies implizierte, dass LAI und Bodenparameter die Hauptmechanismen waren, die Spektralvariationen im sichtbaren und nahen Infrarotbereichen kontrollieren.

6. *Ist EO-LDAS in der Lage, hyperspektrales Signal für eine Reihe von Pflanzen zu simulieren, die nur auf multispektralen Bändern im sichtbaren und VNIR-Bereich beruhen?*

Die Ergebnisse haben gezeigt, dass das Verfahren verwendet werden kann, um willkürliche Bandensätze in sichtbaren und NIR-Regionen zu simulieren, die für jede Anwendung wichtig sind, die erfordern neue und existierende Ströme neuer EO-Daten in der optischen Domäne zu kombinieren, insbesondere Interkalibrierung von EO-Satelliten, um kontinuierliche Zeitreihen der Oberflächenreflexion zu erhalten, über Programme und Sensoren verschiedener Designs.

7. *Wie beeinflusst die Kombination verschiedener optischer Sensoren in EO-LDAS den Abruf biophysikalischer Parameter?*

Zusammenführung der Beobachtungen von multispektralen Sensoren mit unterschiedlichen spektralen und Winkeleigenschaften, wurde durch Kombinierung von MISR- und Landsat ETM+ Daten demonstriert, um das hyperspektrale CHRIS / PROBA-Signal über eine landwirtschaftliche Teststelle zu simulieren.

Es wurde gezeigt, dass das Hinzufügen von Informationen von ETM + zu Multi-Angular Daten von MISR, kann die Genauigkeit erhöhen und die spätere Unsicherheiten von Hyperspektraldatensimulationen verringern. Diese Ergebnisse haben eine Methode zur Kombination verschiedener Sensorenkonstellation der Satelliten demonstriert.

Die vorliegende Arbeit konzentriert sich auf die Parameter des Retrieval Lands aus optischen Sensordaten unter Verwendung eines Datenassimilations-Frameworks. Im Allgemeinen zeigt die vorgestellte Methodik eine gute Leistung. Es gibt jedoch viele Richtungen für die weitere Entwicklung. Gegenwärtig nutzt das System Gaussian Process (GP) -Emulatoren (Gómez-Dans et al. 2016), die die Geschwindigkeit der Verarbeitung erheblich verbessern, aber es ist immer noch zu langsam für globale Anwendungen mit räumlichen Auflösungen von mehr als 1 km. Eine Möglichkeit, die Verarbeitungsgeschwindigkeit zu verbessern, ist die Verwendung der diskreten Kosinustransformation (DCT), wenn Fernerkundungsdaten als Summe von Kosinusfunktionen mit unterschiedlichen Frequenzen dargestellt werden können, was theoretisch schnelle Inversionen ermöglicht. Eine andere Möglichkeit, die Verarbeitungsgeschwindigkeit zu verbessern, ist die Linearisierung eines RT-Modellsignals. Dies ist möglich, weil erste und zweite Ableitungen Nebenprodukte der GP-Emulation sind.

Derzeit verwendet EO-LDAS 1D RT-Modelle wie PROSAIL und Semidiskrete, die praktisch sind, aber diese Art von Modellen vereinfacht die Struktur der Überdachung erheblich. Die Verwendung von GP-Emulatoren eröffnet jedoch einen Weg zur Emulation von 3D-RT-Modellen wie Drat (Lewis, 1999) oder Raytran (Govaerts und Verstraete, 1998), was sehr wichtig sein kann für die Abfrage von Parametern von hochauflösenden Sensoren wie Sentinel-2 MSI. Denn bei höherer Auflösung ist es nicht immer möglich, den Schirm als trübes Medium darzustellen.

Alle veröffentlichten Ergebnisse von EO-LDAS betreffen nur die oberste Ebene, die einfacher zu implementieren ist. Der Grund ist, dass nur drei Modelle gekoppelt werden müssen: Baldachin (Semidiscrete, SAIL), Blatt (PROSPECT) und Boden (Price, Hapke) (Verhoef, 1984, Gobron, 1997, Pinty et al. 1989). Dies führt jedoch zur Aufnahme von Unsicherheiten aus der atmosphärischen Korrektur, die nicht immer leicht verfolgt werden können. Die Lösung hierfür ist eine Kombination von Atmosphären- (6S), Überdachungs-, Blatt- und Bodenmodellen (Vermote, 1997). Erste Experimente für eine solche Kopplung sind Vegetations- und Atmosphärenmodelle in EO-LDAS - nicht veröffentlicht, aber vielversprechend.

EO-LDAS verwendet ein "heute wie morgen" Prozessmodell zur Regularisierung des Inversionsprozesses. Diese Annahme funktioniert in vielen Fällen der natürlichen Entwicklung der Vegetation gut, ist aber bei abrupten Veränderungen wie Bränden, klaren Einschnitten usw. nicht realistisch. Eine Möglichkeit, plötzliche Veränderungen zu berücksichtigen, ist die Verwendung der kantenerhaltenden Regularisierung, die implementiert wurde zur Inversion von linearen Kernel-Modellen für Brandmeldezwecke. Aufgrund von Geschwindigkeitsproblemen wurde jedoch nicht für EO-LDAS verwendet.

Eine weitere Möglichkeit, eine unrealistische Glättung zu vermeiden, ist die Verwendung ökologischer Prozessmodelle für einige Zustandsparameter wie LAI oder Chlorophyll.

Ein wichtiger Teil von EO-LDAS ist ein Beobachtungsoperator, der derzeit ein RT-Modell im optischen Bereich ist. Potenziell kann EO-LDAS zum Zusammenführen von Informationen verschiedener elektromagnetischer Domänen, wie z. B. thermisches Infrarot und passive und aktive Mikrowellen, verwendet werden. Die Hauptkomplikation bei der Aufnahme von SAR-Daten in das EO-LDAS-Schema ist das Fehlen stabiler und ausgereifter physikalischer Modelle. Eine weitere Komplikation ist der Unterschied der Eingangsparameter solcher Modelle. Dies bedeutet eine potentielle Verdoppelung der Anzahl der Parameter, was wiederum zu einer Zunahme der "Unbeliebtheit" des Problems führt. Ein zusätzliches Problem ist die unterschiedliche Annahme von LAI. Trotz all dieser Komplikationen ist das Problem der Kombination von SAR und optischen Daten lösbar, indem neu entwickelte Modelle mit zusätzlichen Einschränkungen verwendet werden.

Contents

Chapter 1: Introduction

1.1. State of the art

Vegetation is an essential component of land ecosystems and plays a key role in regulating life of Earth's biosphere as a producer of complex organic compounds using energy from the sun. Biospherical balance exists during millions of years and instability of Earth's system can cause drastic effects. Nowadays human induced climate change has been recognized by authorities (IPCC 2014). Therefore additional efforts are introduced in monitoring and studying of changing biosphere, climate and biodiversity (GCP 2003; DIVERSITAS 2002). Earth Observation (EO) is a crucial part of such efforts because it provides consistent information about state and dynamic of vegetation over a range of geographical areas with relatively high temporal frequency. In 2005 the Global Earth Observation System of Systems (GEOSS) was initiated to develop an international system for providing spatially and temporally consistent observations (GEO 2005). Copernicus program is the European Union contribution to global Earth monitoring and it is coordinated by European Commission and implemented together with the Member States, the European Space Agency (ESA), and others European institutions.

Eventually knowledge of biophysical parameters helps answering on such questions as what are the main limiting factors of crops growth, how to manage crops more effectively, how climate change will influence future land use and how it will influence future economic situation, how humankind influences global carbon cycle, what are the main sources and sinks of carbon dioxide, what are the main scenarios of biosphere development due to climate change and many others (Stehfest et al. 2007; Noormets 2009).

Quantitative knowledge about state and dynamic of vegetation on global scale can be expressed in a form of such biophysical parameters as Leaf Area Index (LAI), Leave chlorophyll content or Fraction of Photosynthetically Active Radiation (FAPAR). One of the programs which regulates the needs for global satellite products is the Global Climate Observing System (GCOS) (GCOS-200 2016). GCOS recommends a set of Essential Climate Variables (ECV) as a priority of retrieval for the international space agencies. There are 18 recommended terrestrial ECVs and among them are LAI and FAPAR. These two variables are relevant for crop production models and global carbon cycle models.

Accurate and continuous monitoring of land canopy biophysical parameters such as LAI and FAPAR is required for understanding the growth of vegetation on the global scale (Noormets 2009). These parameters provide essential inputs to bio-geochemical cycles modelling, climate modelling, agricultural irrigation management, forecasting of crop production, forest mapping and management and in many other fields (Pinty et al. 2007). Nowadays information for monitoring the state and changes of terrestrial surfaces is provided by a number of optical remote sensing sensors such as Moderate Resolution Imaging Spectro-radiometer (MODIS), Satellite Pour l'Observation de la Terre (SPOT), Landsat Operational Land Imager (OLI), Sentinel-2 Multispectral Instrument (MSI), etc. Optical sensors measure reflected sun signal (i.e. the spectral reflectance) in the range between 400 and 1100 nm. This information is used to retrieve different state variable products such as LAI, Chlorophyll a+b

concentration (C_{ab}) and radiative fluxes like surface albedo and FAPAR (Myneni et al. 2002; Dash et al. 2010; Gobron et al. 1997; Bunnik 1978).

Terrestrial vegetation variables cannot be directly measured from space. They only can be inferred by the spectral information observed by a space sensor. There are two main approaches to infer land variables: one is based on empirical parametric relationships using vegetation indices (VI) and another is physically-based. The first approach exploits statistical relationships between biophysical parameters and satellite spectral measurements. Such method is relatively easy for implementation in the case of comprehensive field measurements. However empirical relationships can be used only for a specific land cover with the same sensor at identical season. They often cannot be used for physically sound interpretation and VIs do not provide information about uncertainty of estimation.

The second way is to obtain parameters by physically-based radiative transfer models (RTM) inversion. Spectral information, measured by space satellite sensors, is an output of RTM which model spectral response of vegetation by given set of parameters. In this case it is considered that estimation of vegetation properties is possible only if we can understand how solar radiation interacts with vegetation and environmental factors. This interaction is described by RTM. Usually typical inputs of a canopy RTM are biophysical parameters such as LAI, height, leaf angle distribution (LAD), etc. The output is reflected spectrum. Biophysical properties of the land surface can be obtained by using methods of inverse modelling (Tarantola 2005). This approach requires knowledge of spectral bidirectional reflectance factor (BRF) values as input and/or multi-angular information. We can define BRF as reflectance which is normalized by the reflectance of a Lambertian surface with identical illumination and geometrical properties (Martonchik et. al. 2000).

In most of the cases the problem of RTM inversion is known as ill-posed, meaning that there can be infinite number of solutions which fit observed information equally well (Kimes et. al., 2000). In order to solve it, various techniques or assumptions can be made such as use a-priori information, fixing few parameters or use a time period of BRFs to increase the number of observations.

During last decades researchers demonstrated having increased interest in developing physically based retrieval operational algorithms for vegetation parameters. Pinty et al. (2007) offered the Two Stream Inversion Package (TIP) to retrieve state canopy parameters, such LAI, and then ratio of fluxes, such as FAPAR. TIP works with surface albedo as input information. Surface albedo is defined as fraction of short-wave radiation reflected by the land's surface (Martonchik et. al. 2000). The inversion scheme of TIP is based on a generic formulation of the inversion problem i.e. the optimal solution can be found by minimization of a cost function J(x). More information about definition of inversion problem can be found for example in Tarantola (2005). In the case of TIP, J(x) is a sum of observational term $J_{obs}(x)$ and prior term $J_{prior}(x)$, where $J_{obs}(x)$ is the mismatch between measured and modelled surface albedos; $J_{prior}(x)$ is the deviation of the state x from prior values in form of the Probability Density Function (PDF). TIP models surface albedos by the two-stream model (Pinty et al. 2006). It exploits adjoint code of the two-stream model and hessian code of J(x). It provides information about posterior uncertainties as posterior covariance matrix. TIP doesn't use all information content of observations because it utilizes only two spectral bands in a form of broad band albedo in red and near infrared (NIR) spectral ranges that are mainly used in climate model. From one hand using only two bands limits amount of available information but from the other hand it reduces number of free parameters. The only

constrain which is used in TIP is the prior term of the cost function. So far TIP is routinely processed at JRC for retrieving of LAI, FAPAR using MODIS albedo and has been used by Disney et al. (2016) .

Other example of a physically based method is the JRC FAPAR algorithm (Gobron et al. 2000; Gobron et al. 2002). The method makes sensor-specific formulae which estimate FAPAR values. The main idea of the algorithm is to optimize coefficients of polynomial functions to make output values first as close as possible to normalized top of canopy values and second to FAPAR. The method uses Top of Atmosphere (TOA) data in the blue, red and near-infrared bands and optimization procedure corrects for atmospheric absorption and scattering effects, soil influence and illumination/observation geometry. It models FAPAR values using combination of the semi-discrete radiative transfer model (Gobron et al. 1997); 6S atmospheric RT model (Vermote et al. 1997); Rahman-Pinty-Verstraete (RPV) model (Rahman et al. 1993); and the Price soil spectra (Price 1990). Advantage of the method is that TOA data are used and no complicated step of atmospheric correction is required. It works much faster than most other RTM based methods. However this algorithm can estimate only FAPAR; methods of regularization cannot be incorporated; it has to be recalibrated for each sensor; and this formula cannot take into account all range of possible geophysical, biological and illumination conditions.

Knyazikhin et al. (1998) offer a retrieval algorithm which allows synergy use of MODIS and MISR data and thereby employ 7 spectral bands and revisiting frequency of MODIS and multi-angularity of MISR. The algorithm is based on Look Up Tables (LUT) made a canopy RT model. The algorithm relates total satellite derived reflectance to spectral properties of leafs. I.e. it establishes relations between a model and a sensor which exclude soil contribution. This works as a constraint because with exclusion of soil we get much less spectral variations. Thus we can consider it as introducing of additional information what reduces number of possible solutions. Also authors propose using only 6 biomes from MODIS land cover classification to constrain the task even more. Each biome applies some restrictions to specific variations of LAI and some over properties. For example in the case of grasses, LUT is calculated by 1D model but in the case of forests 3D model is used. Disadvantage of such approach is that the algorithm depends on accuracy of land cover classification. The optimization procedure is based on generic formulation of inverse problem. I.e. on minimization of cost function which consists of one or several terms where each term is a difference between modelled and observed signal. Despite of physical correctness the algorithm does not produce information about posterior uncertainty required in various ecological models.

None of the cited above methods however cannot restore gaps in data because it does not use spatial or temporal regularization, i.e. assumption of smooth change in space or in time. In more general words we can say what spatial or temporal regularization is using of prior knowledge of spatial or temporal distribution of the signal as constrains for inversion of a RT model.

Many efforts were made in order to find possibilities to regularize the inversion problem and to estimate posterior uncertainties. Quaife and Lewis (2010) imposed conditions of temporal smoothness to constrain linear BRDF model. Laurent et al. (2013) estimated LAI and C_{ab} from hyperspectral data of the Airborne Prism Experiment. These authors inverted the TOA data using the coupled SLC-MODTRAN4 canopy-atmosphere model. The regularization was done by model coupling, using a-priori information and object-level spatial constraints. Dubovik et al. (2011) used expectation of smoothness for spatial-temporal regularization to

estimate some aerosol properties. In the work of Mousivand et al. (2015) multi-temporal, multi-sensor constraints were used to retrieve LAI, Cab and fCover.

Over land surface, Lewis et al. (2012a) proposed the Earth Observation Land Data Assimilation System (EO-LDAS) which is based on weak constrain data assimilation (DA), meaning that a model is not perfect and has a model error which is regulated by a so-called regularization parameter (Zupanski 1997). The system implements mechanisms for constraining input remote sensing information in order to get optimal estimate of vegetation parameters. In EO-LDAS, one can constraint the optimal solution by prior information, time or/and spatial regularization. Lewis et al. (2012b) used EO-LDAS temporal regularization to obtain some state parameters by synthetic Sentinel-2 data. Lewis et al. (2012a) estimated LAI over the Gebesee field in Germany using EO-LDAS temporal regularization.

Despite of a number of efforts listed above there were no examples of prior, multi-angular, multi-sensor and temporal constraints used together in order to improve retrieval of biophysical parameters.

1.2. Motivation and relevance

Despite of recent achievements, the problem of biophysical parameters retrieval by inversion of physical models against satellite multispectral information has still large uncertainties. Information provided by satellite sensors usually not constrains RTM inversion enough which is resulted in too many possible solutions. Procedure of decreasing number of solutions by introducing some additional information (constraining) is called regularization. One way of introducing information is a priori knowledge about vegetation state. However in practice there is no additional information than satellite sensor signal. More ever most amount of space optical information has many gaps in time and in space due to variability of clear sky/cloudy conditions. Possible regularizations are temporal, spatial or combining together data of various sensors.

Partially these gaps can be filled by combining data of various EO sensors due to their different revisiting frequency. Therefore the problem of combination of different sensors is more and more important for operational use as well as for reanalysis of archived information. Technical difficulty is how to combine sensors with difference in spectral bands, spatial resolutions and view/sun geometries in a reliable way. In addition combining together various sensors can add some spectral information for such variables as chlorophyll content or leaf water content.

EO-LDAS can fill gaps in time series of satellite information in time and in space providing reliable output uncertainties. Also it can combine various sensors providing additional multi-sensor constraints. Previously EO-LDAS has been validated against synthetic Sentinel-2 data and real MODIS data using time regularization (Lewis et al., 2012b). However in this study only mono-angle systems were considered and only LAI field measurements were compared against retrieved values. Despite of its promising results the EO-LDAS scheme requires much more validation efforts.

One of the advantages of EO-LDAS is that it allows including different pieces of information for yielding *a posteriori* estimation. Therefore it is relatively easy to include multi-angular and/or multi-sensor information. Despite of the fact that the knowledge of multi-angular spectral observations has been recognized as required (Gobron et al. 2000a, Liu et al. 2004; Widlowski et al. 2004) there is no clear evidence how multi-angular or multi-sensor data can improve retrieval of parameters in EO-LDAS scheme.

When state biophysical parameters are retrieved RTM can simulate any optical sensor using these state parameters as an inputs for forward modelling. This allows simulating hyperspectral information which is quite valuable and rare over the globe. In addition it creates a basis for validation of RTM-inversion results in the case when no *in-situ* data are available. However it requires some validation efforts.

Most of widely used RTM do not allow direct estimation of FAPAR values. However it can be calculated indirectly with the closure of energy balance equation. Many existing FAPAR products are based on physically based models but most of them do not use time and/or space regularization techniques. Potentially estimation of FAPAR with using EO-LDAS regularization capabilities allows making new advanced FAPAR product.

1.3. Research Objectives

The main objective is to develop a physically based and automated approach for retrieving LAI and FAPAR and simulated hyper-spectral information over a wide variety of satellite sensors, observational/illumination conditions and agricultural crops.

In order to fulfil this objective, answers to the following research questions have to be found:

1. How to retrieve continuous FAPAR and its associated uncertainties if number of available observations is low.
2. How to retrieve consistent time series of biophysical parameters using EO-LDAS?
3. How inclusion of static or dynamic prior information to EO-LDAS influences retrieval of LAI and FAPAR from CHRIS/PROBA, MISR and Landsat data?
4. What effect has multi-angular information of MISR and CHRIS/PROBA on retrieval of biophysical parameters over agricultural fields?
5. What are the major mechanisms controlling spectral variations in the visible and near infrared regions over agricultural fields? Can these mechanisms be parameterized by multispectral information of EO sensors?
6. Is EO-LDAS able to simulate a hyper-spectral signal for a range of crops relying only on multispectral bands in visible and NIR region?
7. How does the combination of different optical sensors in EO-LDAS influence retrieval of biophysical parameters?

1.4. Structure of the thesis

The introductory chapter of the thesis is followed by the chapter 2 which gives description of methods and materials used. The first section of the chapter describes optical EO sensors which were used in the study. The theoretical description of LAI and FAPAR are then presented. The next section discusses methods and limitations of radiative transfer modeling which can be used to retrieve these parameters. Section 2.4 provides explanation of multi-angular information and why it can be useful for constraining of RTM inversion. The last section of the chapter describes how we can combine different optical sensors, multi-angular information and RTM for retrieval of biophysical parameters by using data assimilation.

Chapter 3 summarizes main results and discussion of each article of this thesis. The following chapter summarizes the main achievements.

The last three annexes provide the three independent peer-reviewed articles:1) investigation into the influence of prior constraints and space regularization on the retrieval of Leaf Area Index and Chlorophyll a+b content using EO-LDAS; 2) the use of MISR multi-angular information for the reconstruction of hyperspectral signals; and 3) the retrieval of FAPAR time series with using temporal regularization and MISR data.

The papers have been published in the Remote Sensing journal, Advances in Space Research journal and in Earth Observation for Land and Emergency Monitoring book. All papers went through a peer-review process.

Chapter 2: Methods and materials

2.1. Optical Sensors

Optical remote sensing sensors can only record the radiation in visible part of the electromagnetic spectrum, reflective infrared and thermal infrared. Table1 and Figure 1 provide an overview of the main optical sensors used for vegetation monitoring.

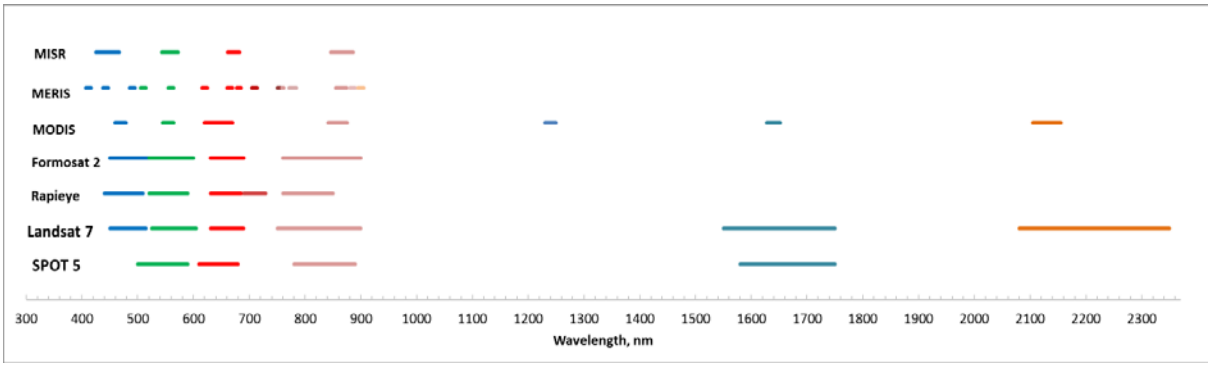


Figure 1: Wavebands of the some sensors in optical domain.

One of the requirements of the accurate data assimilation process is the availability of large amount of observations. It is better if the time series of observations are available. Nowadays this type of requirements can provide such sensors as MODIS, MERIS and MISR. MODIS and MERIS are multispectral, wide field of view sensors. They provide traditional type of remote sensing information. On the other hand, Multi-angle imaging Spectro Radiometer (MISR) can provide multi-angular data for each of 4 spectral bands (Figure 2).

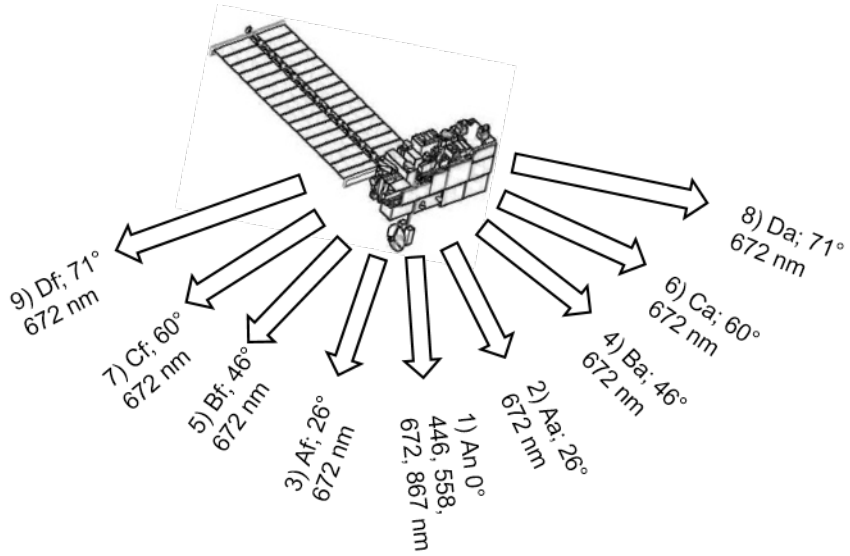


Figure 2: The cameras of the MISR instrument

MISR has 9 view angles per band. This means that it has nine cameras pointed at fixed angles. One camera looks to the nadir and others image Earth at 26.1, 45.6, 60.0, and 70.5 degrees forward and afterward. This configuration allows obtaining characteristics of the surface anisotropy in a quite precise way. In addition it can give more free parameters for RTM inversion. I.e. for example MODIS has 7 bands but MISR has 36 bands. 7 bands of MODIS have 500m resolution but MISR bands have 275m. The only disadvantage of MISR that MODIS and MERIS have more spectral information.

Table 1: Parameters of the sensors in optical domain

	SPOT-5	Landsat-7	Rapideye	Formosat-2	MODIS	MERIS	MISR
Spectral Bands (nm)	500-590 610-680 780-890 1580-1750	0.45 -0.515 0.525 0.605 0.63-0.69 0.75-0.90 1.55-1.75 10.4-12.5 2.09-2.35	440-510 520-590 630-685 690-730 760-850	0.45-0.52 0.52-0.60 0.63-0.69 0.76-0.90	620-670 841-876 459-479 54-565 1230-1250 1628-1652 2105-2155	407-417 437-447 485-495 505-515 555-565 615-625 660-670 677-685 704-714 750-757 758-762 771-785 855-875 880-890 895-905	425-467 543-773 661-683 846-886
Spatial resolution (m)	10	30	5	8	500	300	275
Revisit time (days)	2-3	16	1 (order)	1 (order)	1/2	3	9
Altitude (km)	822	705	630	888	705	785	705
Swath width (m)	60	185	77	24	2330	1150	360

2.2. State Parameters (LAI, FAPAR)

2.2.1. LAI

"Leaf area index (LAI) is a dimensionless variable and is defined as the total one-sided area of photosynthetic tissue per unit ground surface area" (Watson 1947). LAI is an important parameter because many models in agriculture, ecology, climate and other require information about the amount of leaves presented in a given Earth surface (Monteith and Unsworth 1990).

The typical way to infer LAI by the satellite derived spectral information is by using radiative transfer models. 1-D RTM simplified representation of vegetation by assuming a random distribution of leaves which have discrete or infinitely small size and specific leaf angle distribution (LAD). Usually this representation does not contain any description of canopy structure. Indeed, low and middle spatial resolution satellite data integrate various properties of surface in one pixel. In this case 1D representation of canopy is meaningful. However, true LAI depends on structural properties of canopy. The effective LAI was introduced to satisfy 1-D representation of vegetation canopy (Pinty et al. 2004). It can be defined as

$$\hat{LAI}(\mu_0) = < LAI > \xi(\mu_0)$$

where $LAI(\mu_0)$ is effective LAI, <LAI> the domain averaged LAI of the layer, ξ the clumping (structural) factor and μ_0 the cosine of sun zenith angle.

Among other parameters LAI is a typical input to 1-D RTM but it is the most influential parameter. This makes priority in validation of LAI compared to other model inputs. As one example, Figure 3 shows sensitivity of the PROSAIL model (Jacquemoud et al. 2009) outputs to changes of different parameters (leaf structure, chlorophyll, coratenoids, brown pigment, water thickness, dry matter, LAI and hot spot). One can see that LAI provides the highest variability on the whole range of wavelengths from 400 to 2500 nm.

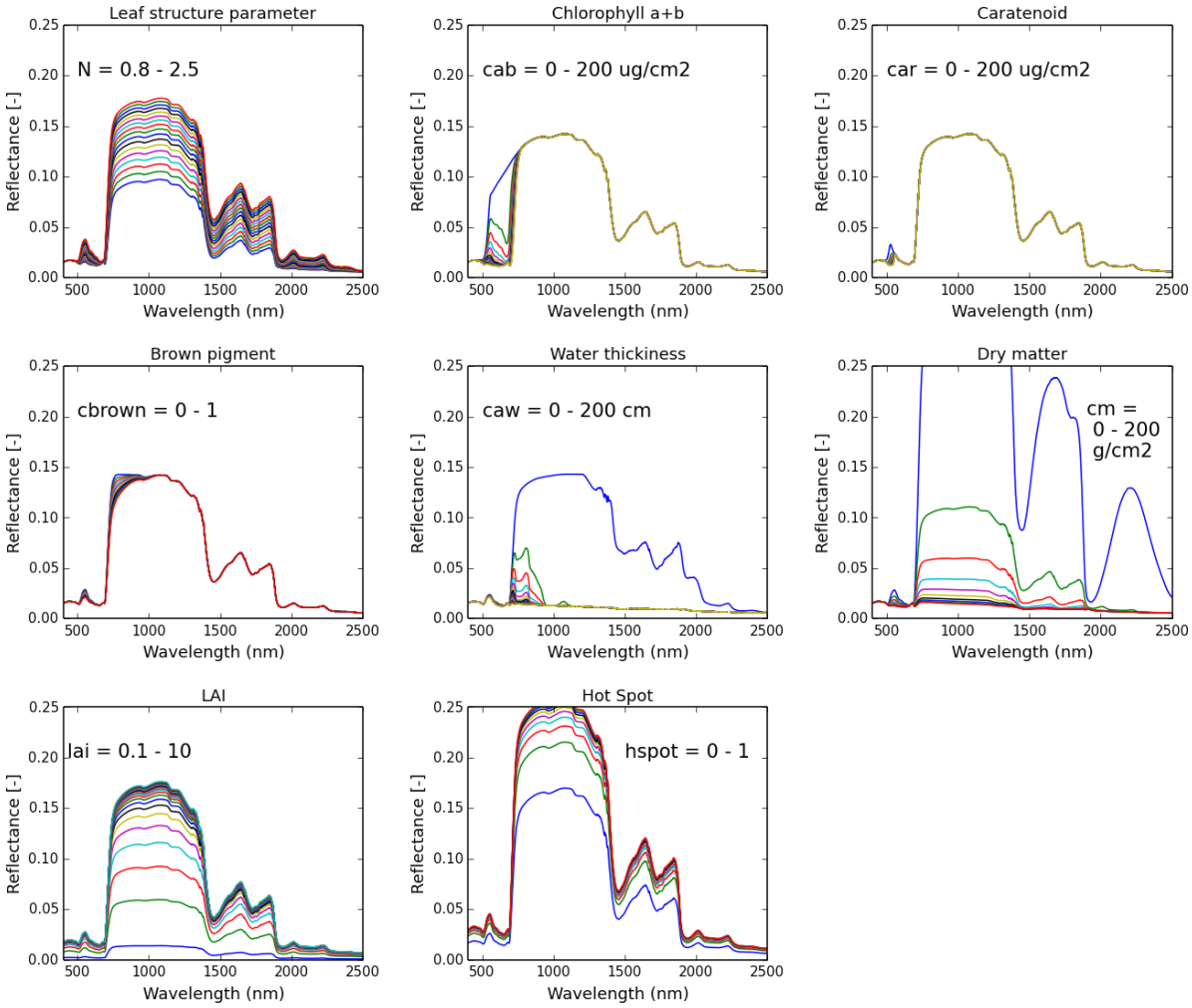


Figure 3: Sensitivity of the PROSAIL model

Currently the following global satellite LAI products are available: NASA-LAI; GLOBCARBON; LANDSAF; JRC-TIP and Copernicus Global Land Service LAI products (Knyazikhin et al. 1998; Plummer 2007; Roujean and Lacaze 2002; Pinty et al. 2007).

2.2.2. FAPAR

Very important parameter for energy balance of ecosystems is the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR). FAPAR is the part of photosynthetically active radiation (PAR) in the range 400-700 nm which is absorbed by vegetation tissue.

FAPAR is a main parameter controlling photosynthetic activity of vegetation. It can serve as an indicator of land carbon sink rate, presence of live plants, desertification and productivity of agricultural fields (Gobron and Verstraete 2009). Also it is important because it provides information about intensity of land carbon cycle. Therefore it helps in understanding greenhouse gas dynamics. Global biosphere models require some measurements of CO2 fluxes which are essential for data assimilation. However spatial distribution of in-situ CO2 measurements is seldom and therefore does not provide adequate accuracy. At the same time long time series of FAPAR provide good spatial-temporal coverage at global scale.

FAPAR can be defined as “green” which means that PAR is only absorbed by lived leafs or as “total” FAPAR which includes PAR absorbed by truck, branches or dead leafs.

Nowadays there are several available FAPAR products: Joint Research Centre (JRC) MERIS-FAPAR; NASA MODIS-FAPAR; NASA MISR-FPAR; JRC-TIP; GLOBCARBON; LANDSAF and Copernicus Global Land Service FAPAR products (Gobron et al. 2000; Knyazikhin et al. 1998; Knyazikhin et al. 1998; Pinty et al. 2007; Plummer et al. 2006; Roujean and Breon 1995). Extensive overview of all available FAPAR/LAI methods and products can be found in https://lpvs.gsfc.nasa.gov/producers2.php?topic=Fpar

The main disadvantages of both LAI and FAPAR retrievals are gaps in satellite acquisitions, underdetermined RTM problem and difference between various retrieval schemes. In recommendations of the Global Terrestrial Observing System (GTOS) it is recommended to focus on reliability and accuracy of LAI and FAPAR to improve ingestion by data assimilation systems (Gobron and Verstraete, M. 2009). This means that it is important to improve accuracy of current products and to provide information about associated uncertainties which is essential for the data assimilation process.

2.3. Retrieval of Parameters by Radiative Transfer Modelling

2.3.1. Land Vegetation Radiative Transfer Models

Description of transfer of sun radiation through vegetation requires a vegetation canopy model coupled with a leaf spectra model and a soil model. Canopy level models can represent vegetation medium as simplified 1-D structure or more realistic 3-D. 1-D models treat vegetation as randomly and uniform distributed particles (leaves) which have prescribed leaf angle distribution (LAD). Examples of canopy 1-D RT models, among others, are Scattering by Arbitrarily Inclined Leaves (SAIL) (Verhoef 1984), semi-discrete model (Gobron et al. 1997), 2-stream based model (Pinty et al. 2006; Sellers 1985) and A two-layer Canopy Reflectance Model (ACRM) (Kuusk 2001). More complex 3-D models describe canopy structure in 3 dimensional space (Ni et al. 1999; Govaerts 1996; Lewis, 1999). These models are very computationally expensive and have too many input parameters. Usually they are not appropriate for operational inversion of remote sensing derived reflectance.

One of most widely used leaf model is PROSPECT which represents a leaf as several transparent absorbing plates with rough surfaces (Jacquemoud et al. 2009; Jacquemoud and Baret 1990).

One example of a soil model is the Price's model which represents soil reflectance as a linear combination of 4 base functions (Price 1990). Other way to represent soil is using a modified Hapke bidirectional reflectance model (Pinty et al. 1989; Liang and Townshend 1996).

2.3.2. Inversion

The task of an inversion technique is to infer structural and biophysical parameters of canopy by known satellite spectral BRFs. In other words inversion of a canopy RTM means adjusting model parameters until model output are more or less equal to satellite measurements.

The following main factors influence remote sensing of vegetation by its reflectance (Goel 1988):

b. The source of radiation - the sun defined by its spectral intensity $I(\lambda)$ ($Wsr^{-1}\mu m^{-1}$), zenith θ_s and azimuth ψ_s angles.
c. Atmosphere – characterized by different absorptions and concentrations, aerosols particles, water vapour, ozone.
d. Vegetation canopy, characterized by a set of parameters. They include optical parameters (reflectance and transmittance) and structural (geometrical shapes and positions) of vegetation components (leaves, stems, etc.), planting geometry, and environmental parameters like temperature, relative humidity, wind speed, and precipitation. In general these parameters have wavelength, temporal, and spatial dependencies.
e. Ground or Soil - reflectance and absorption, surface roughness, surface texture, bulk density and moisture profile.
f. Detector - spectral sensitivity, aperture, calibration and position (view zenith angle θ_0 and view azimuth angle ψ_0).

Due to these factors process of remote sensing of vegetation can be described as (Goel 1988):

$$Ri = f(a_i, b_i, c_i, d_i, e_i) \qquad (1)$$

R_i – set of attributes of the radiation received by the detector.

f – function which produce set of R_i by radiative transfer processes.

There are two problems concerning relationship (1):

Forward problem – defining of R by given (a_i, b_i, c_i, d_i, e_i). I.e. $\{R_i\} = f(a_i, b_i, c_i, d_i, e_i)$

Inverse problem – defining of c_i by given R and (a_i, b_i, d_i, e_i). I.e. $\{c_i\} = g(R_i, a_i, b_i, d_i, e_i)$. In general remote sensing inverse problem can be seen as an optimization problem. Typical optimization problem is finding minimum of a function. For the EO problem this is a function of distribution of residuals. In this case residuals mean the vector of difference between multiple runs of a model and EO measurements.

Model parameters can be free or fixed. Free - adjustable parameters. If there is p free parameters then at least n=p independent equations in the form (2) are required to evaluate them (Kimes et al. 2000):

$$R_1 = f(b_i, c_i, d_i, a_i, 1, e_i, 1),$$

$$R_2 = f(b_i, c_i, d_i, a_i, 2, e_i, 2),$$

$$R_n = f(b_i, c_i, d_i, a_i, n, e_i, n). \quad (2)$$

If $n<p$, the system is underdetermined and more independent equations (for example reflectance samples) are required for a unique solution. This is the common situation for multi-spectral remote sensing. Typical canopy RTM should have at least 13 input parameters in order to comprehensively describe interaction of solar radiance and vegetation. On the other hand many multi-spectral sensors have only 4-7 optical spectral bands.

If $n=p$, there is normally no unique solution due to model and measurement inaccuracies.

If $n>p$, there is no single solution because the system is over determined. In this case, however, a statistically "best" solution can be determined as is most often done in practice (Kimes et al. 2000).

This means that a number of free parameters has to be smaller than number of spectral bands of a sensor. For example in Verhoef and Bach (2003) canopy RTM inversion is the part of data assimilation process. In this research model GeoSAIL (Huemmrich, 2001) was used for canopy level and PROSPECT for leaf level. Number of free parameters was three (LAI, fraction brown leaves and soil moisture), at the same time a number of used Landsat TM bands is six.

There are many approaches to solve the problem of the model inversion. (Kimes et al. 2000) highlights traditional approaches, the neural network and look-up-table methods.

In the traditional approach, iterative techniques are used to solve the inverse problem. During model inversion, the optimization algorithm iteratively adjusts the free parameters until the modelled reflectance closely corresponds to the measured reflectance. There is three most commonly used traditional minimizations routines: the downhill simplex method, the conjugate direction set method and the quasi-Newton method (Kimes et al. 2000). Kimes et al. (2000) conclude that the traditional inversion methods are very computationally expensive and are not appropriate for using on a per pixel basis for regional and global studies. The look-up table and neural network are more computationally efficient and can be applied on a per pixel basis. These methods can be applied to the most sophisticated RTM without any simplifications. Because look-up table or the training process (in case of neural networks) are generated before the actual operational application of the method, the computation time to generate this prior information is not so important. Consequently, the neural network and look-up table methods can be applied to any RTM. One more advantage of these methods is that they do not require any initial guesses to model variables as do the traditional inversion methods. Poor choices in initial guesses can lead to a great risk of terminating at local minima (Kimes et al. 2000).

The disadvantages of look up tables are that they are sensitive to distribution of the input variables and to the step size of the variables values. It could leads to a non-representative result.

The other way of doing minimization is Gaussian Process emulator (GPE) (Gómez-Dans 2016, Rasmussen and Williams 2006). It can be used as a statistical non-linear approximation of a computer model which is referred as a simulator. This means that we can approximate RTM by GPE using a set of training runs. Then we can use GPE as a representation of RTM and it works significantly faster than a real RTM. The important assumption of GPE is that variation of the output (reflectance) is changing relatively smoothly with changing of the inputs (state parameters). GPEs have been using in computer model calibrations and sensitivity analysis due to their effectiveness (Oakley and O'Hagan 2004; Bayarri et al. 2007). A GPE is required representative set of RTM samples. It can be done using one of the multidimensional sampling methods. For example Latin-hypercube sampling (LHS), which guarantees that the generated samples represent the real model variability with a relatively small number of model runs (McKay et al. 1979).

On figure 4 is given an example of comparison GPE-generated and RT-modelled data.

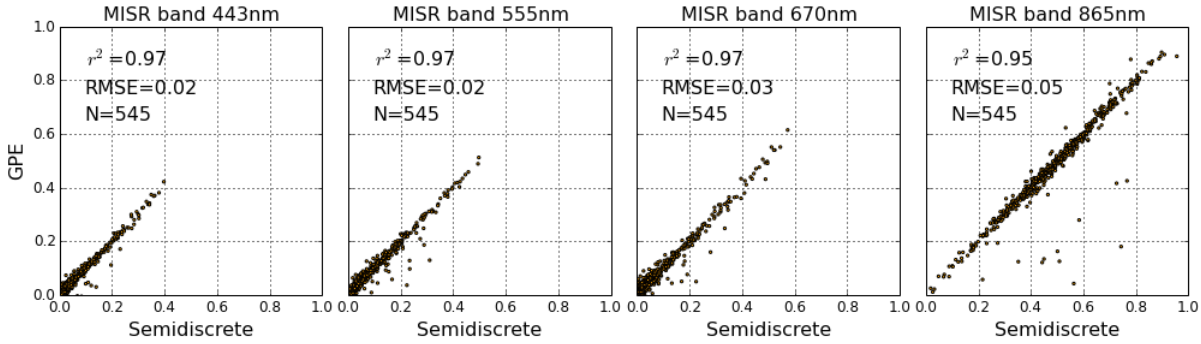


Figure 4: Comparison of the GPE and semidiscrete spectra over 800 samples for MISR bands with central wavelengths 443, 555, 670 and 865 nm.

In general canopy radiative transfer model inversion has ill-posed nature because usually a number of free parameters are larger than available information, such as spectral reflectance in different sensor's bands. The consequence of ill-posed inverse problem is that almost identical spectra could be produced by quite different combinations of parameters.

The quality of the inversion could be improved by increasing the dimensionality of observations of by constraining of the free variables. Hyperspectral measurements can significantly increase dimensionality but such data are not always available. In addition, increasing spectral dimensionality not always increases a number of relevant information. The best way to improve model inversion is to introduce a priori information on varaiables of interest. Such kind of information may be collected from literature, field measurements, other sensors, or previous experiments (Dorigo et al. 2007). Usually *a priori* information is used for fixation of variables or for defining the range of their possible change.

2.4. Multi-angular remote sensing

Availability of multi-angular satellite information since a decade brought new opportunities to understanding of radiation transfer processes in atmosphere and vegetation (Pinty et al. 2002). The main advantage of multi-angular data is to improve accuracy of retrieved state parameters due to additional constraints to solutions of radiative transfer problem. The nature of these constraints lies in better description of reflectance anisotropy.

As an example right (left) hand side panel of figure 5 shows BRF as a function of view zenith and azimuth angles (view zenith angle in one azimuth plane) at a given sun zenith equal to 28°.

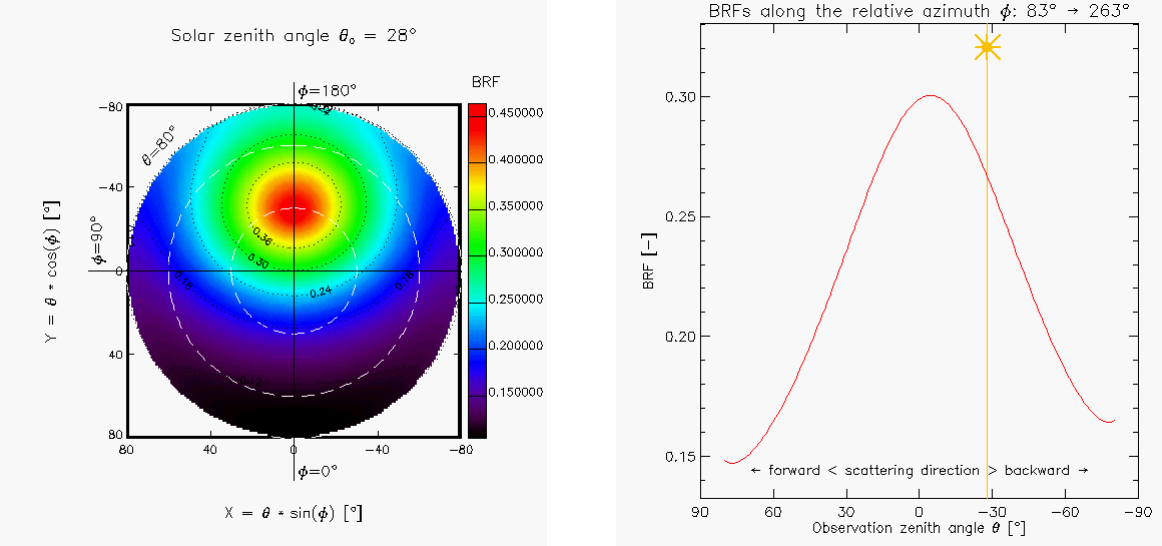


Figure 5: AnisView (Vogt and Verstraete, 2016) representation of BRF. Right hand side panel: BRF as a function of view zenith and azimuth angles. Left hand side panel side: BRFs in a given azimuth plane equal to 83°.

One can see what a single view angle doesn't provide all available information. So with reasonable increase of view angles we use more information and potentially can decrease uncertainty in solution of a radiative transfer model inversion.

Two instruments provide global multi-angular measurements the Polarization and Directionality of Earth Reflectances (POLDER) (Deschamps et al. 1994) and the Multi-angle Imaging Spectroradiometer (MISR) (Diner et al. 1998). This latter has nine view angles and a spatial resolution from 275 m to 1.1 km. Several studies proofed that MISR multi-angular information can improve retrieval of some land parameters such as the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) and the Leaf Area Index (LAI). For instance the study of Knyazikhin et al. (1998) describes the algorithm for synergistic retrieval of LAI and FPAR from Moderate Resolution Imaging Spectroradiometer (MODIS) and MISR measurements. This algorithm takes into account soil anisotropy, non-Lambertian surface and canopy 3D effects. Gobron et al. (2000) demonstrates that using multiple observational angles from MISR reduce the number of solutions when inverting a canopy radiative transfer model. The studies of Pinty et al. (2002), Gobron et al. (2002) and Widlowski et al. (2004) proofed that MISR can also provide knowledge on structure and heterogeneity of vegetation. High potential of MISR for retrieval of forest stands parameters by coupling canopy and atmosphere RT models are demonstrated in Laurent et al. (2011). More recently Pisek et al. (2013) showed that it is possible to derive vegetation clumping index from 275 m MISR data.

2.5. Data Assimilation

The main task of the system is the minimization of a cost function J_{post} which is the sum of three following cost functions:

$$J_{post} = J_{prior} + J_{obs} + J_{model} \quad (3)$$

where J_{prior} provided an priori knowledge of state variables:

$$J_{prior} = \frac{-1}{2}(x - x_p)^T C_p^{-1}(x - x_p) \quad (4)$$

in which C_p is the covariance matrix which describes uncertainty of the prior state, x is a vector of state variables and x_p the prior estimates.

The function of a prior constraint is to correct the cost function J_{post} by the prior term J_{prior}. This constraint is controlled by the prior state vector x_p and our belief to this state, i.e. the uncertainty of the prior model state C_p.

The second term J_{obs} corresponds to the observations cost function:

$$J_{obs} = \frac{-1}{2}(x - H(x))^T C_0^{-1}(x - H(x)) \quad (5)$$

where H(x) corresponds to the radiative transfer model for the scattering of light by vegetation and C_o the covariance matrix describing the uncertainty in the observations.

Finally the dynamic model cost function J_{model}:

$$J_{model} = \frac{\gamma^2}{2} x^T (D^T D) x \quad (6)$$

where γ = 1/δ is the regularization parameter which represents the model error and controls the smoothness of output. D is the differential operator in the form of a matrix:

$$D = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & -1 & 0 \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix} \quad (7)$$

The described scheme allows estimation of posterior uncertainties of the data. Uncertainties can be obtained by considering curvature at the minimum of the state space. I.e. the more steep area at minimum the lower uncertainties. Technically it can be done by inversion of the Hessian matrix. The result of the inversion is covariance matrix which in turn can be transform to correlation matrix (Figure 6).

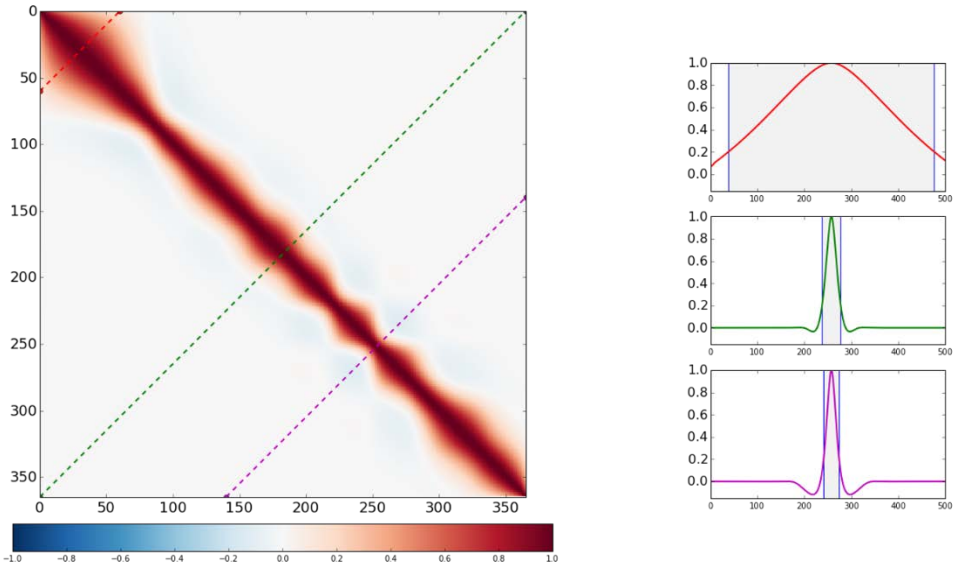


Figure 6: Left panel: the posterior correlation matrix which shows correlation between days of year. Higher correlation outside main diagonal generally means more interpolation because of temporal gaps for these dates. Right panel: Profiles which show how much information we get from other days. Correlation = 1 is a day we want to retrieve.

Time regularization can be applied when time series of remote sensing data are available. It is assumed that the development of biophysical variables can be described by a dynamical model. Time regularization is already proofed as a powerful tool for extraction of biophysical information from time series of remote sensing reflectance data (Houborg et al. 2007; Lauvernet et al. 2008; Quaife and Lewis 2010) and has been already used in EO-LDAS with Sentinel-2 like data in (Lewis et al. 2012b).

In EO-LDAS level of time regularization is controlled by regularization parameter γ which has to be estimated. Figure 7 shows how different values of γ affect the retrieval.

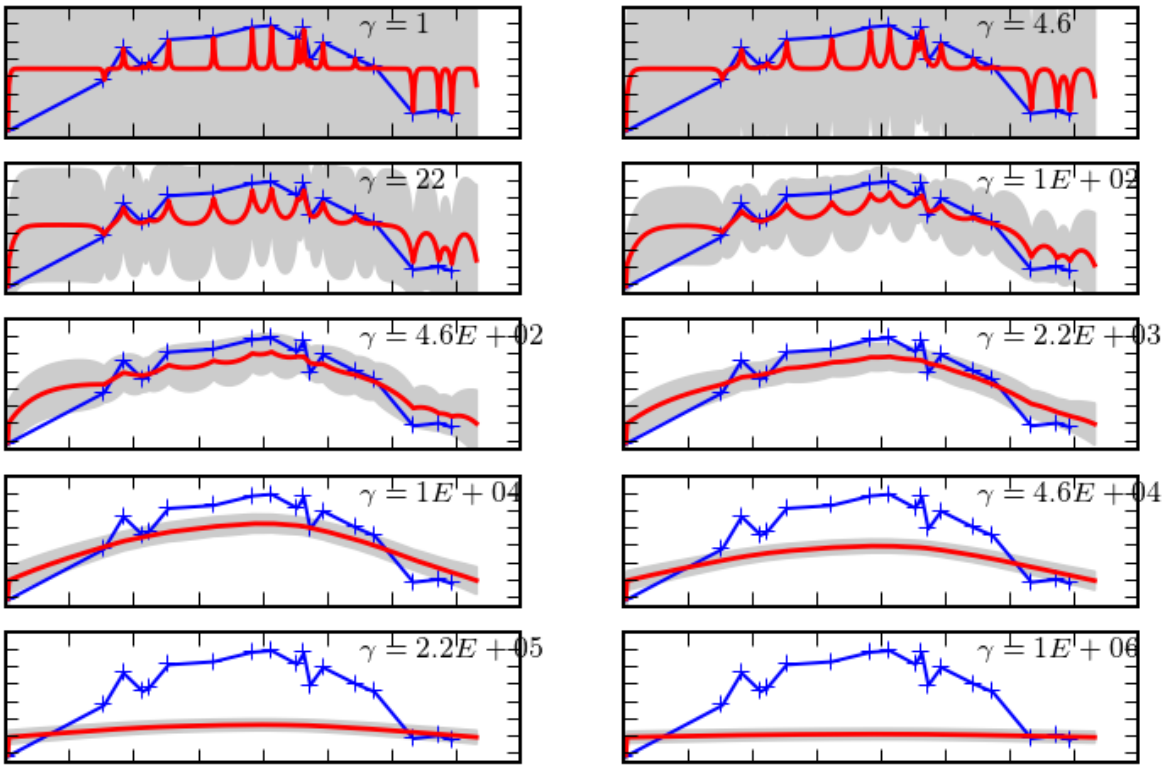


Figure 7: An example of the regularization of time series on synthetic data.

Spatial regularization is a new way in constraining of RT canopy model inversion. The main idea of spatial regularization is to use statistics of surrounded pixels in addition to a spectral signature. There are several studies which demonstrated increasing performance of the inversion after applying such spatial constrains (Atzberger and Richter 2012; Atzberger 2004; Wang et al. 2008; Laurent et al. 2013).

Chapter 3: Results and Discussion

3.1. Regularization constraints in EO-LDAS

The main aim here was to overview the EO-LDAS regularization methods for improving retrieval of canopy biophysical parameters and accuracy of hyper-spectral simulations. The work was done over the Barrax agricultural area in Spain. Different constraints were tested, including prior information on the canopy parameters, temporal regularization, spatial regularization and multi-sensor information (Chernetskiy et. al. 2015b).

The study was divided into four parts by the considered methods: 1) *a priori* knowledge of canopy state parameters within a specific range, 2) temporal regularization by a dynamic vegetation model of zero order, 3) spatial regularization using an assumption about spatial dependency between the pixels, and 4) constraints by using multi-sensor information. The first three methods were used for retrieving the biophysical parameters. While combination of multiple sensors measurements, such as MODIS 500 m, Landsat ETM+ and MISR full resolution, aimed at simulating hyper-spectral signatures of CHRIS/PROBA. The corresponding surface reflectance over the Barrax site was used. Field measurements needed for validation purpose are part of the ESA-SPARC campaign which was carried out over Barrax site in Spain in 2004 (Gandia et al., 2004). As a result it was found that all considered possibilities potentially can be used for constraining different problems in EO-LDAS: 1) the not strict prior constraint can decrease uncertainty of estimated LAI and Chlorophyll content values on pixel by pixel basis up to 100%, 2) time regularization decreases uncertainty of LAI values to up 100%, 3) spatial constraint for CHRIS/PROBA and MERIS sensors can reduce the uncertainty from 30% to 50% and 4) multi-sensor information in the form of combinations ETM+ and MODIS / ETM+ and MISR can constrain the problem of simulating hyper-spectral signatures of CHRIS/PROBA.

3.2. Reconstruction of CHRIS/Proba hyperspectral signal

This study presented a validation of multi-sensor and multi-angular constraints of EOLDAS in a form of reconstruction of CHRIS/PROBA hyper-spectral information over the Barrax test site (Chernetskiy et. al. 2015a).

The validation has been carried out for real and synthetic data sets: synthetic MISR, synthetic MISR and ETM+, real MISR, real MISR and ETM+. Each dataset contains 18 CHRIS/PROBA hyper-spectral measurements. These measurements correspond to following crops: Alfalfa, Sugar Beet, Corn, Garlic, Onion, Potato, Sunflower and Vineyard. In all four cases numbers of MISR cameras were increased one by one from nadir camera (An) to cameras up to 70 deg. It was found that in all considering cases posterior uncertainties were decreased by increasing the number of MISR cameras. The results based on the real MISR data had lowest accuracy due to low spatial resolution as its signal was mixed between crops and bare soil therefore it had high discrepancy with CHRIS/PROBA data. However adding spectral information of ETM+ always increased accuracy and decreased uncertainties relative

to both MISR and ETM+ and MISR only results. This was valid for synthetic as well as for real datasets.

This study showed that hyper-spectral signal between 400 and 1000 nm can be reconstructed based on a limited number of canopy parameters. Thus, the only parameters which were retrieved with reasonable accuracy were LAI and soil coefficients. However, we were still able to reconstruct CHRIS/Proba signal with root-mean-square error (RMSE) of 0.03 for more dense vegetation cover and RMSE=0.005 for less dense vegetation.

The results demonstrated that EO-LDAS can be used for simulation of hyper-spectral data when hyper-spectral information was not available. In addition it can be used for validation of biophysical parameters retrieval if hyper-spectral information was available but field measurements were not available. In addition it was shown a way of combination (only in spectral domain) of mixed coarse spatial resolution data and high resolution data.

3.3. Retrieval of Fraction of Photosynthetically Active Radiation (FAPAR)

This work concentrated on the retrieval of FAPAR time series over an agricultural test site (Chernetskiy et. al. 2015c). The retrieval of FAPAR using MISR full resolution (275 meters) over a FluxNet site with EO-LDAS was considered. The retrieval was done with temporal regularization using seven MISR cameras. Results were compared against ground-based data MERIS Full Resolution, MODIS MCD15 and Joint Research Centre Two-stream Inversion Package (JRC-TIP) products for years 2001-2008.

Among the compared products (JRC-TIP, MERIS FR and MODIS MCD15) the EO-LDAS results had both the highest linear correlation (r^2>0.8) and the lowest root mean square error (0.14). The EO-LDAS LAI retrievals showed a high correlation with ground-truth measurements, but some underestimation for high LAI values as a result of saturation. In this study, we also estimated the daily leaf chlorophyll content and fraction of senescent leaves. Although retrieved values of both parameters were not compared with any in situ measurements, they showed a realistic temporal trajectory. These results suggested that the EO-LDAS approach was successful in retrieving both FAPAR and other land surface parameters. A large part of this success was based on the use of a dynamic regularisation model that counteracted the poor temporal sampling from the MISR instrument. This demonstrated that at least for these particular sites, EO-LDAS was able to estimate absorbed fluxes and other land surface parameters relying on multi-spectral information, only. A main factor of the good results was based on a dynamic regularisation model that softened the effect of the sparse temporal sampling from the MISR instrument. EO-LDAS generated results for each day of year; it showed ability to predict FAPAR between available satellite observations: this could be a new way for making space products validation in absence of ground-based measurements.

Conclusions

In this thesis, a new methodology for retrieval of land spectral and biophysical parameters from optical remote sensing data has been designed and used. The result of the work was a physically based methodology for FAPAR and LAI retrievals, simulation of hyper-spectral information and estimation of associated uncertainties. The parameters derived by presented methodology are self-consistent, i.e., the same physical (RT) model assumptions are made to retrieve LAI, FAPAR and simulated spectral information.

In the course of the work the following answers to the *research questions* were found:

1. *How to retrieve continuous FAPAR and its associated uncertainties if number of available observations are low.*

EO-LDAS can be used for daily estimation of FAPAR and associated uncertainties without any in-situ information and when the number of available observations is low. The results were in line with the field measurements with r^2 varying from 0.84 to 0.92 and RMSE from 0.11 to 0.16. This was the highest rate among compared products (JRC-TIP, MERIS FR and MODIS MCD15). This outcome demonstrated that EO-LDAS is able to estimate absorbed fluxes relying only on multi-spectral information. The use of temporal regularisation provides a controlled way of interpolating over data-free periods and the ability to track fast phenological changes phenology. This was especially important in the case of e.g. MISR observations, which were temporally sparse. Temporal regularisation was able to capture the leading and ending edges of the FAPAR curve in all of the fields and for most of the years. Furthermore, since EO-LDAS generated results for each day-of-year, it showed ability to predict FAPAR between available satellite observations: this may be a new way for performing space product validation in absence of ground-based measurements.

2. *How to retrieve consistent time series of biophysical parameters using EO-LDAS?*

Results showed that EO-LDAS can be used for retrieval of different parameters providing consistency between them. Thus, in this study daily LAI, leaf chlorophyll content and fraction of senescent leaves were estimated. Estimation of LAI showed high correlation with ground truth measurements but with some underestimation for high LAI values. Retrieved time series of leaf chlorophyll content and fraction of senescent leaves showed a realistic temporal trajectory which was consistent with dynamics of LAI and FAPAR.

3. *How does inclusion of static or dynamic prior information to EO-LDAS influence retrieval of LAI and FAPAR from CHRIS/PROBA, MISR and Landsat data?*

It was shown, that using MISR information, EO-LDAS temporal regularization and generic dynamic prior, it was possible to stabilize results of the retrieval and to obtain better results

than MERIS FAPAR or JRC-TIP MISR. In addition, inclusion of generic static and dynamic prior information, decreases posterior uncertainties and can increase accuracies compared to in-situ data.

4. *What effect has multi-angular information of MISR and CHRIS/PROBA on retrieval of biophysical parameters over agricultural fields?*

It was found, that sequential increase of MISR multi-angular observations improves the RMSE of results based on synthetic data. Calculation of the cross entropy did not show, that multi-angular observations always provide more information than the same number of mono-angular nadir observations. This suggested that - at least in case of this experimental setup - the number of observations was of greater significance than different viewing angle positions.

5. *What are the major mechanisms controlling spectral variations in the visible and near infrared regions over agricultural fields? Can these mechanisms be parameterized by multispectral information of EO sensors?*

The results showed that proper estimation of LAI and soil parameters were sufficient to simulate a hyper-spectral signal between 400 and 1000 nm with acceptable precision: best RMSE is equal to 0.03 for real data and less than 0.008 for synthetic data. This implies that in case of the given experimental set-up, LAI and soil parameters are the major mechanisms controlling spectral variations in the visible and near infrared regions.

6. *Is EO-LDAS able to simulate a hyper-spectral signal for a range of crops relying only on multispectral bands in visible and NIR region?*

Results showed, that the methodology can be used to simulate arbitrary band sets in the visible and NIR regions, which are of importance to any application which requires combining new and existing streams of EO data in the optical domain, particularly inter-calibration of EO satellites in order to get continuous time series of surface reflectance, across programmes and sensors of different designs.

7. *How does the combination of different optical sensors in EO-LDAS influence the retrieval of biophysical parameters?*

Merging information of ETM+ and multi-angular data of MISR can increase accuracies and decrease a-posteriori uncertainties of hyper-spectral data simulations. These results demonstrate a new possibility to combine different satellite constellation sensors.

Outlook

The presented work concentrates on the retrieval land parameters from optical sensor data using the data assimilation framework. In general presented methodology shows good performance however, there are many directions for further development. Currently the system exploits Gaussian Process (GP) emulators (Gómez-Dans et al. 2016) which significantly improve speed of the processing but it is still too slow for global applications with spatial resolution more than 1 km. One way to improve processing speed is to use Discrete Cosine Transform (DCT) when we can represent remote sensing data as a sum of cosine functions with different frequencies which in theory allows to perform fast inversions. Another way to improve speed of the processing is using linearization of a RT model's signal. This is possible because first and second derivatives are the side products of the GP emulation.

Currently EO-LDAS uses 1D RT models, such as PROSAIL and semidiscrete, which are practical but these type of models significantly simplify canopy structure. However, exploitation of GP emulators open a way for emulation of 3D RT models such as Drat (Lewis, 1999) or Raytran (Govaerts and Verstraete, 1998) which can be very important for retrieval of parameters from high resolution sensors such as Sentinel-2 MSI. Because in case of higher resolution it is not always possible to represent canopy as a turbid medium.

All published results of EO-LDAS concern only top-of-canopy level which is easier to implement. The reason is that only three models have to be coupled: canopy (semidiscrete, SAIL), leaf (PROSPECT) and soil (Price, Hapke) (Verhoef, 1984, Gobron, 1997, Pinty et al. 1989). However this leads to incorporation of uncertainties from atmospheric correction which are not always easily trackable. The solution for this is implementing combination of atmospheric (6S), canopy, leaf and soil models (Vermote, 1997). The first experiments for coupling are vegetation and atmospheric models in EO-LDAS are not published but showed promising results.

EO-LDAS uses "today as tomorrow" process model for regularization of the inversion process. This assumption works well in many cases of natural development of vegetation however it is not realistic in case of abrupt changes such as fires, clear cuts, etc. One way of taking into account abrupt changes is to use edge preserving regularization which was implemented for inversion of linear kernel models for fire monitoring purposes. However, due to speed issues was not used for EO-LDAS.

The other way of avoiding unrealistic other smoothing is to use ecological process models for some state parameters such as LAI or chlorophyll.

An important part of EO-LDAS is an observational operator which currently is a RT model in optical domain. However, potentially EO-LDAS can be used for merging information of different electromagnetic domains such as thermal infrared and passive and active microwave. The main complication for ingestion SAR data into EO-LDAS scheme is lack of stable and mature physical models. The other complication is difference in input parameters of such models. This means potential doubling in number of parameters which in turn will lead to increasing of "ill-posedness" of the problem. An additional issue is difference in

assumptions on LAI. However, despite of all these complications the problem of combining SAR and optical data is solvable by using newly developed models and including additional constraints.

Acknowledgments

This work would never be possible without extraordinary and kind support of prof. Schmullius. This work would never be possible without help and patience of Dr. Gobron. I would like to say many thanks to them. This work would also not be possible without many years work of prof. Lewis and Dr. Gómez-Dans. I would like to thank Dr. Gómez-Dans for his help, support and critical view on my work. I would like to thank Dr. Disney for his support. I would like to thank prof. Heiko Balzter for making the GIONET project possible. I would like to say many thanks to Martyna Stelmaszczuk-Górska for her kind support in so many things. I would like to thank Sina Truckenbrodt for her help and friendship.

I would like to acknowledge financial support of the project GIONET, funded by the European Commission, Marie Curie Programme Initial Training Network, Grant Agreement Number PITN-GA-2010-264509; the EU H20:20 BACI project (Project No. 640176); ESA-Fire-CCI project. I would like to acknowledge the European Space Agency and the National Aeronautics and Space Administration for providing data for this work. I also wish to gratefully acknowledge Anatoly Gitelson and Andrew Suyker for the ground-based FAPAR data, as well as Marco Zuehlke and Brockmann's Consult for providing the MERIS FR results. The JRC-authors thank Linda Hunt and Michel Verstraete for their technical support on MISR HR processing chain.

References

Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. Remote Sensing of Environment, 93(1-2), 53–67.

Atzberger, C. and Richter, K., 2012. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. Remote Sensing of Environment, 120, 208–218.

Bayarri, M.J., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.H. and Tu, J., 2007. A framework for the validation of computer models. Technometrics, 49(2), 138–154.

Bunnik, N.J.J., 1978. The multispectral reflectance of shortwave radiation by agricultural crops in relation with their morphological and optical properties, Wageningen, Netherland.

Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P., Schmullius, C. 2015. Simulations of CHRIS/PROBA spectra with Earth Observation Land Data Assimilation System using MISR data. *Advances in Space Research* (in preparation).

Chernetskiy, M., Gobron, N., Gómez-Dans, J., Lewis, P. and Schmullius, C.C., 2017. Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site. In Earth Observation for Land and Emergency Monitoring. John Wiley & Sons, Ltd, 117–139.

Chernetskiy, M., Gomez-Dans, J., Gobron, N., Morgan, O., Lewis, P., Truckenbrodt, S. and Schmullius, C., 2017. Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing, 9(7), 656.

Dash, J., Jeganathan, C. and Atkinson, P.M., 2010. The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India. Remote Sensing of Environment, 114(7), 1388–1402.

Deschamps, P.-Y., Breon, F.-M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J.-C. andSeze, G., 1994. The POLDER mission: instrument characteristics and scientific objectives. Geoscience and Remote Sensing, IEEE Transactions on, 32(3), 598–615.

Diner, D.J., Beckert, J.C., Reilly, T.H., Bruegge, C.J., Conel, J.E., Kahn, R.A., Martonchik, J. V, Ackerman, T.P., Davies, R., Gerstl, S.A.W., Gordon, H.R., Muller, J.-P., Myneni, R.B., Sellers, P.J., Pinty, B. and Verstraete, M.M., 1998. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. Geoscience and Remote Sensing, IEEE Transactions on, 36(4), 1072–1087.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Voβbeck, M., Lewis, P., Pinty, B. A new global FAPAR and LAI dataset derived from optimal albedo estimates: Comparison with MODIS products. Remote Sensing 2016, 8.

DIVERSITAS, 2002. DIVERSITAS SCIENCE PLAN; DIVERSITAS Report No. 1, Paris.

Dorigo, W., Baret, F., Richter, R., Ruecker, G., Schaepman, M. and Mueller, M., 2007. Retrieving canopy variables by radiative transfer model inversion an automated regional approach for imaging spectrometer data. In 5th EARSeL Workshop on Imaging Spectroscopy. Brughes, Belgium: EARSeL.

Dubovik, O., Herman, M., Holdak, a., Lapyonok, T., Tanré, D., Deuzé, J.L., Ducos, F., Sinyuk, a. and Lopatin, a., 2011. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. Atmospheric Measurement Techniques, 4(5), 975–1018.

Gandia, S., Fernández, G., García, J.C. and Moreno, J., 2004. Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign, In Proc. 2nd CHRIS/Proba Workshop, ESA/ESRIN, Frascati, Italy, 2004.

GCP, 2003. The Global Carbon Project, Science Framework and Implementation. Earth System Science Partnership (IGBP, IHDP, WCRP, DIVERSITAS) Report No. 1; Global Carbon Project Report No. 1, Canberra.

GEO, 2005. Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan,

Gobron, N., Pinty, B., Verstraete, M.M. and Govaerts, Y., 1997. A semidiscrete model for the scattering of light by vegetation. Journal of Geophysical Research, 102, 9431–9446.

Gobron, N., Pinty, B., Verstraete, M.M., Martonchik, J. V., Knyazikhin, Y. and Diner, D.J., 2000. Potential of multiangular spectral measurements to characterize land surfaces: Conceptual approach and exploratory application. Journal of Geophysical Research, 105(D13), 17539.

Gobron, N., Pinty, B., Verstraete, M.M., Widlowski, J., Diner, D.J. and Member, A., 2002. Uniqueness of Multiangular Measurements — Part II : Joint Retrieval of Vegetation Structure and Photosynthetic Activity From MISR. , 40(7), 1574–1592.

Gobron, N., Pinty, B., Verstraete, M.M. and Widlowski, J., 2000. Advanced Vegetation Indices Optimized for Up-coming Sensors: Design, Perfomance, and Applications. , 38(6), 2489–2505.

Gobron, N., Pinty, B., Verstraete, M.M. and Widlowski, J.-L., 2000. Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications. Geoscience and Remote Sensing, IEEE Transactions on, 38(6), 2489–2505.

Gobron, N., Pinty, B. and Verstraete, M.M., 1997. Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. IEEE Transactions on Geoscience and Remote sensing, 35(6), 1438–1445.

Gobron, N. and Verstraete, M.M., 2009. Leaf Area Index. Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables, 11, 12.

Gobron, N. and Verstraete, M., M., 2009. Fraction of Absorbed Photosynthetically Active Radiation (FAPAR). Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables., 10, 14.

Goel, N.S., 1988. Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. Remote Sensing Reviews, 4(March 2012), 1–212.

Gómez-Dans, J.L.; Lewis, P.E.; Disney, M., 2016. Efficient Emulation of Radiative Transfer Codes Using Gaussian Processes and Application to Land Surface Parameter Inferences. Remote Sens. 8, 119.

Govaerts, Y.M., 1996. A model of light scattering three-dimensional plant canopies: a Monte Carlo ray tracing approach, Ispra, Italy, 1996.

Houborg, R., Soegaard, H. and Boegh, E., 2007. Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. Remote Sensing of Environment, 106(1), 39–58.

Huemmrich, K. F., 2001. The GeoSail model: A simple addition to the SAIL model to describe discontinuous canopy reflectance. Remote Sensing of Environment, 75, 423−431.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jacquemoud, S. and Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment, 34(2), 75–91.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C. and Ustin, S.L., 2009. PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment, 113, S56–S66.

Kimes, D.S., Knyazikhin, Y., Privette, J.L., Abuelgasim, A.A. and Gao, F., 2000. Inversion methods for physically based models. Remote Sensing Reviews, 18, 381–439.

Knyazikhin, Y., Martonchik, J. V., Myneni, R.B., Diner, D.J. and Running, S.W., 1998. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. Journal of Geophysical Research, 103(D24), 32257.

Knyazikhin, Y., Martonchik, J.V., Diner, D.J., Myneni, R.B., Verstraete, M., Pinty, B., Gobron, N. and N., 1998. Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data. Journal of Geophysical Research, 103, 32239– 32256.

Kuusk, A., 2001. A two-layer canopy reflectance model. Journal of Quantitative Spectroscopy & Radiative Transfer, 71, 1–9.

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W. and Schaepman, M.E., 2011. Inversion of a coupled canopy–atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. Remote Sensing of Environment, 115(10), 2603–2612.

Laurent, V.C.E., Verhoef, W., Damm, A., Schaepman, M.E. and Clevers, J.G.P.W., 2013. A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from APEX at-sensor radiance data. Remote Sensing of Environment, 139, 6–17.

Lauvernet, C., Baret, F., Hascoët, L., Buis, S. and Le Dimet, F.-X., 2008. Multitemporal-patch ensemble inversion of coupled surface–atmosphere radiative transfer models for land surface characterization. Remote Sensing of Environment, 112(3), 851–861.

Lewis, P. 1999 Three-dimensional plant modelling for remote sensing simulation studies using the botanical plant modelling system Agron. Agric. Environment, 19 (1999), pp. 185-210

Lewis, P., Gomez-Dans, J., Kaminski, T., Settle, J., Quaife, T., Gobron, N., Styles, J. and Berger, M., 2012. Data assimilation of Sentinel-2 observations: preliminary results from EO-LDAS and outlook. In Proc. "First Sentinel-2 Preparatory Symposium", Frascati, Italy 23–27 April 2012, ESA SP-707, July 2012.

Lewis, P., Gómez-Dans, J., Kaminski, T., Settle, J., Quaife, T., Gobron, N., Styles, J. and Berger, M., 2012. An Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing of Environment, 120, 219–235.

Liang, S. and Townshend, J.R.G., 1996. A modified hapke model for soil bidirectional reflectance. Remote Sensing of Environment, 55(1), 1–10.

Liu, X., Kafatos, M., Gomez, R.B. and Wolf, H., 2004. Multi-angular satellite remote sensing and forest inventory data for carbon stock and sink capacity in the Eastern United States forest ecosystems. In ISPRS Congress Istanbul 2004. 1–6.

Martonchik, J. V., C. J. Bruegge, and A. Strahler, 2000. A Review of Reflectance Nomenclature Used in Remote Sensing, Remote Sensing Reviews, 19, 9-20

McKay, M.D., Beckman, R.J. and Conover, W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics, 21(2), 239–245.

Monteith, J.L. and Unsworth, M.H., 1990. Principles of Environmental Physics, Edward Arnold, London, UK.

Mousivand, A., Menenti, M., Gorte, B. and Verhoef, W., 2015. Multi-temporal, multi-sensor retrieval of terrestrial vegetation properties from spectral–directional radiometric data. Remote Sensing of Environment, 158, 311–330.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.., Lotsch, A., Friedl, M., Morisette, J.., Votava, P., Nemani, R..and Running, S.., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sensing of Environment, 83(1-2), 214–231.

Ni, W., Xiaowen, L., Woodcock, C.E., Caetano, M.R. and Strahler, A.H., 1999. An analytical hybrid GORT model for bidirectional reflectance over discontinuous plant canopies. IEEE Transactions on Geoscience and Remote Sensing, 37, 987–999.

Noormets, A., 2009. Phenology of Ecosystem Processes: Applications in Global Change Research, Springer.

Oakley, J.E. and O'Hagan, A., 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society: Series B, 66(3), 751–769.

Pinty, B. , Verstraete, M. M. and Dickinson, R. E. 1989. A physical model for predicting bidirectional reflectances over bare soil, Remote Sensing of Environment, Volume 27, Issue 3, Pages 273-288,Pinty, B., Gobron, N., Widlowski, J.-L., Lavergne, T. and Verstraete, M.M., 2004. Synergy between 1-D and 3-D radiation transfer models to retrieve vegetation canopy properties from remote sensing data. Journal of Geophysical Research: Atmospheres, 109(D21), D21205.

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

Pinty, B., Lavergne, T., Voßbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M.M. and Widlowski, J.-L., 2007. Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products. Journal of Geophysical Research, 112(D10), D10116.

Pinty, B., Widlowski, J., Gobron, N., Verstraete, M.M., Diner, D.J. and Member, A., 2002. Uniqueness of Multiangular Measurements — Part I : An Indicator of Subpixel Surface Heterogeneity From MISR. , 40(7), 1560–1573.

Plummer, S., Arino, O., Simon, W. and Steffen, W., 2006. Establishing an Earth observation product service for the terrestrial carbon community: the GLOBCARBON initiative. Mitigation and Adaptation Strategies for Global Change, 11, 97–111.

Plummer, S., 2007. An update on the GLOBCARBON initiative: multi-sensor estimation of global biophysical products for global terrestrial carbon studies. In Proceedings of Envisat Symposium 2007. Montreux (Switzerland). ESA Communication Production Office (Publ.), Noordwijk.

Price, J.C., 1990. On the information content of soil reflectance spectra. Remote Sensing of Environment, 113–121.

Quaife, T. and Lewis, P., 2010. Temporal Constraints on Linear BRDF Model Parameters. IEEE Transactions on Geoscience and Remote Sensing, 48(5), 2445–2450.

Rahman, H., Pinty, B. and Verstraete, M.M., 1993. Coupled Surface-Atmosphere Reflectance (CSAR) Model 2. Semiempirical Surface Model Usable With NOAA Advanced Very High Resolution Radiometer Data. Journal of Geophysical Research, 98(D11), 20,791–20,801.

Rasmussen, C.E.R. and Williams, C.K.I., 2006. Gaussian processes for machine learning, The MIT Press.

Roujean, J.-L. andBreon, F.-M., 1995. Estimating {PAR} absorbed by vegetation from bidirectional reflectance measurements. Remote Sensing of Environment, 51(3), 375–384.

Roujean, J.L. and Lacaze, R., 2002. Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation. Journal of Geophysical Research, 107D, 10129–10145.

Sellers, P.J., 1985. Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing, 6(8), 1335–1372.

Stehfest, E., Heistermann, M., Priess, J.A., Ojima, D.S. and Alcamo, J., 2007. Simulation of global crop production with the ecosystem model DayCent. Ecological Modelling, 209(2-4), 203–219.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM.

Verhoef, W. and Bach, H., 2003. Remote sensing data assimilation using coupled radiative transfer models. Physics and Chemistry of the Earth, Parts A/B/C, 28(1-3), 3–13.

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sensing of Environment, 16, 125– 141.

Vermote, E., Tanre, D., Deuze, J.L., Herman, M. and Morcrette, J.J., 1997. Second Simulation of the Satellite Signal in the Solar Spectrum (6S): An Overview. IEEE Transactions on Geoscience and Remote sensing, 35(3), 675–686.

Vogt P.; Verstraete, M.M.; 2016. AnisView. Available at the following web site: http://forest.jrc.ec.europa.eu/download/software/anisview

Wang, Y., Yang, C. and Li, X., 2008. Regularizing kernel-based BRDF model inversion method for ill-posed land surface parameter retrieval using smoothness constraint. Journal of Geophysical Research, 113(D13), D13101.

Watson, D.J., 1947. Comparative physiological studies in the growth of field crops. I: Variation in net assimilation rate and leaf area between species and varieties, and within and between years. Ann. Bot., 11, 41–76.

Widlowski, J.-L., Pinty, B., Gobron, N., Verstraete, M.M., Diner, D.J. and Davis, A.B., 2004. Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies. Climatic Change, 67, 403–415.

Zupanski, D., 1997. A general weak constraint applicable to operational 4DVAR data assimilation systems. Monthly Weather Review, 125, 2274–2292.

7

Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site

M. Chernetskiy[1], N. Gobron[2], J. Gómez-Dans[3], P. Lewis[3] and C.C. Schmullius[1]

[1] *Friedrich-Schiller-University, Institute of Geography, Department for Earth Observation, Jena, Germany*
[2] *EC Joint Research Centre, Ispra (VA), Italy*
[3] *University College London (UCL), London, UK*

7.1 Background

Vegetation plays a key role in regulating the life of earth's biosphere. Accurate and continuous monitoring of land canopy biophysical parameters such as Leaf Area Index (LAI) is required for understanding the growth of vegetation [1]. These parameters in turn also provide essential input to bio-geochemical cycle modelling, climate modelling, agricultural irrigation management, forecasting of crop production, forest mapping and management and many other fields [2]. Nowadays information for monitoring the state and changes of terrestrial surfaces is provided by a number of optical remote sensing sensors such as Moderate Resolution Imaging Spectroradiometer (MODIS), Satellite Pour l'Observation de la Terre (SPOT), Landsat, etc. The spectral reflectances are used to retrieve different state variable products such as LAI, Chlorophyll concentration (Cab) and radiative fluxes like surface albedo and fraction of photosynthetically active radiation (fAPAR) [3–7].

There are two main methods to obtain these land variables: one is based on empirical parametric relationship with vegetation indices (VI) and another is physically based. The first one uses statistical relationships between biophysical parameters and spectral measurements. Such a method is relatively easy for implementation in the case of comprehensive field measurements. However empirical relationships can be used only with the same configuration of observations, i.e. for a single sensor and type of land cover. The second way is physically based as a canopy radiative transfer model, used through inversion technique. Extraction of vegetation properties is possible only if we can understand how solar radiation interacts with vegetation and environmental factors. This interaction is described by RTM. Bio-geophysical properties of the earth surface can be obtained by inversion of RTM. This approach requires several spectral bidirectional reflectance factor (BRF) values as inputs and/or the knowledge of multi-angular information. The task of an inversion technique is to infer canopy parameters by known spectral BRFs, i.e. input variables of a radiative transfer model adjusted to best explain measured fields of reflectance. Examples of canopy 1-D RT models, among

Earth Observation for Land and Emergency Monitoring, First Edition. Edited by Heiko Balzter.
© 2017 John Wiley & Sons Ltd. Published 2017 by John Wiley & Sons Ltd.

others, are Scattering by Arbitrarily Inclined Leaves (SAIL) [8], semi-discrete model [5], 2-stream based models [9,10] and A two-layer Canopy Reflectance Model (ACRM) [11]. More complex 3-D models describe canopy structure in 3-dimensional space [12,13]. These models are very expensive and have too many input parameters. Usually they are not appropriate for operational inversion of remote sensing derived reflectance.

One problem of RT modelling inversion is that the number of unknown parameters is usually higher than the number of input data, especially for mono-angular sensors [14,15]. This means that a number of free parameters (inputs of a model) has to be smaller than the number of spectral bands of a sensor. This leads to the so-called ill-posed problem, i.e. the number of possible solutions is too high and so there is no a unique solution. In order to solve this problem, various techniques or assumptions can be made such as fixing a few parameters or using a time period of BRFs to increase the number of observations. All these methods characterized by introducing some additional information can be called methods of regularization [16].

The quality of the inversion may be improved by increasing the dimensionality of observations by constraining or in other words regularizing the free variables. Hyperspectral measurements can significantly increase dimensionality, but such kinds of data are not always available. In addition, increasing spectral dimensionality does not always increase the amount of relevant information. One of the ways to improve inversion of variables is to introduce a priori information. Such information may be collected from literature, field measurements, other sensors, or previous experiments [17]. Usually a priori information is used for the fixation of variables or for defining the range of their possible change. Other methods of regularization use additional information from temporal or spatial development of reflectance fields [18,19].

Recently many efforts have been made in order to find possibilities for regularization of the inversion problem as well as for uncertainties estimation [20–22]. Over land surface, Lewis *et al.* (2012) [23] proposed the Earth Observation Land Data Assimilation System (EO-LDAS) which is based on weak constraint data assimilation (DA), meaning that a model is not perfect and has model error, which is regulated by a so-called regularization parameter [24]. The system implements mechanisms for constraining input remote sensing information in order to obtain the optimal estimate of vegetation parameters of the canopy surface. In EO-LDAS, one can constrain the optimal solution by prior information, time or/and spatial regularization.

The main task of the system is the minimization of a cost function J_{post}, which is the sum of the three following cost functions:

$$J_{\text{post}} = J_{\text{prior}} + J_{\text{obs}} + J_{\text{model}} \tag{7.1}$$

where J_{prior} provided a priori knowledge of state variables:

$$J_{\text{prior}} = -\frac{1}{2}\left(x - x_{\text{p}}\right)^{T} C_{\text{p}}^{-1}\left(x - x_{\text{p}}\right) \tag{7.2}$$

in which C_p is the covariance matrix which describes uncertainty of the prior state, x is a vector of state variables and x_p the prior estimates.

The function of a prior constraint is to correct the cost function J_{post} by the prior term J_{prior}. This constraint is controlled by the prior state vector x_p and our belief to this state, i.e. the uncertainty of the prior model state C_p.

The second term J_{obs} corresponds to the observations cost function:

$$J_{obs} = -\frac{1}{2}\left(x - H(x)\right)^T C_0^{-1}\left(x - H(x)\right) \tag{7.3}$$

where H(x) corresponds to the radiative transfer model for the scattering of light by vegetation and C_o the covariance matrix describing the uncertainty in the observations.

Finally the dynamic model cost function J_{model}:

$$J_{model} = -\frac{\gamma^2}{2} x^T \left(D^T D\right) x \tag{7.4}$$

where $\gamma = 1/\delta$ is the regularization parameter which represents the model error and controls the smoothness of output. D is the differential operator in the form of a matrix:

$$D = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \dots & -1 & 0 \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix} \tag{7.5}$$

Time regularization can be applied when time series of remote sensing data are available. It is assumed that the development of biophysical variables can be described by a dynamic model. Time regularization is already proven as a powerful tool for extraction of biophysical information from time series of remote sensing reflectance data [20,25,26] and has been already used in EO-LDAS with Sentinel-2 like data in [23].

Spatial regularization is a relatively new branch in constraining RT canopy model inversion. The main idea of spatial regularization is to use statistics of surrounded pixels in addition to a spectral signature. There are several studies which have demonstrated increased performance of the inversion after applying such spatial constraints [18,19,21,27].

As can be seen, the main purpose of EO-LDAS is the retrieval of biophysical land variables. However, once the state is known after inverting some observations, the system can be used to forward model and predict other observations. In this case availability of hyperspectral measurements such as CHRIS/PROBA data is highly desirable, because it allows validation of a full spectrum in a certain range of wavebands.

Another technique which can be employed in EO-LDAS is using multi-angular information. Conventional remote sensing methods use only one view zenith angle. Meanwhile a multi-angular sensor uses several view angles. This allows better characterization of the structure of vegetation. Knowledge of multi-angular spectral observations has been recognized as essential [28,29]. Since multi-angularity introduces additional information it can be considered as a regularization technique.

The main aim of this contribution is to give an overview of the possibilities offered by the EO-LDAS tool using separately prior information, time regularization and spatial regularization. We also propose the use of multi-angular information in addition to spectral signatures of low spatial/spectral resolution sensors in order to simulate the spectral signatures of higher spatial/spectral resolution sensors. These four exercises will be conducted using space data over the test site Barrax in Spain.

7.2 Methods and Data

7.2.1 Site Description

An agriculture area near Barrax in Spain, was chosen as a test site due to the availability of both field measurements of the European Space Agency (ESA) SPectrabARrax Campaign(SPARC) campaign 2004 (Figure 7.1) [30] and a series of satellite acquisitions (see Table 7.1).

Barrax is a widely studied area, where many field campaigns have taken place over the years 2003, 2004, 2005 and 2009. These campaigns have mostly dealt with the retrieval

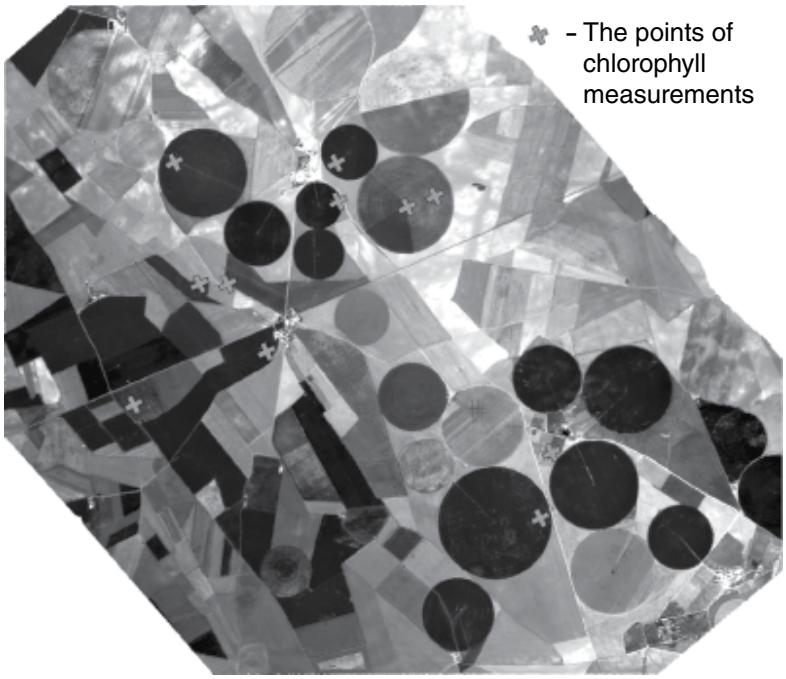


Figure 7.1 INTA-AHS 80 Airborne Hyperspectral Scanner (AHS) scene of the Barrax test site for 15.07.2004, 2 m spatial resolution. Measurement points are marked as crosses. Imagery from the SPARC 2004 database.

Table 7.1 Summary of data characteristics.

Sensor	Spatial resolution	Spectral bands	Number of View Zenith Angels (VZA)	Date of acquisition
CHRIS/PROBA	25 m	62 bands from 400 to 1100 nm	5	2004.07.16
ETM+/Landsat	30 m	483, 565, 660, 825, 1650 and 2220 nm	1	2004.07.18
MODIS/Terra	500 m	483, 565, 660, 825, 1240, 1650 and 2220 nm	1	2004.07.16
MISR/Terra	275 m	446, 558, 672 and 867 nm	9	2004.07.16
MERIS/Envisat	300 m	412, 442, 490, 510, 560, 620, 665, 681, 708, 753, 762, 779, 865, 885 and 900 nm	1	2004.07.17

of biophysical parameters, such as LAI, chlorophyll concentration, leaf equivalent water thickness, leaf dry matter and soil spectral measurements. The various works carried out during the campaign mostly related to the retrieval of biophysical parameters by space optical measurements. These latter data have been collected during the campaigns and include two high-resolution hyperspectral sensors CHRIS/PROBA and INTA-AHS.

7.2.2 Remote Sensing Data

The Compact High Resolution Imaging Spectrometer (CHRIS) is an experimental instrument on board the Proba-1 satellite. It can be configured to provide from 19 to 63 spectral bands in range (400–1050 nm) with spatial resolution from 17 to 34 m. CHRIS/PROBA was constructed for collection of Bidirectional Reflectance Distribution Function (BRDF); therefore it has five view zenith angles from nadir to 55° [31].

One scene of 25 m spatial resolution CHRIS/PROBA was acquired on 16 July 2004, with 62 spectral bands in the range from 400 to 1100 nm at five different viewing zenith angles. The closest nadir geometrical properties are: 8.4°, 283.6°, 20.8°, 325.5° – view zenith angle (VZA), view azimuth angle (VAA), sun zenith angle (SZA) and sun azimuth angle (SAA) respectively.

The Multi-angle Imaging Spectro-Radiometer (MISR) is the operational multi-angular optical sensor, which acquires information globally. We used the MISR top of canopy reflectance at 275 m on 16 July 2004. Only bands with 275 m resolution were used: the four nadir spectral bands and the red bands over eight other cameras.

Now not operational, the MEdium-spectral Resolution Imaging Spectrometer (MERIS) had 15 spectral bands from 390 to 1040 nm and spatial resolution 300 m. We used MERIS data acquired on 17 July 2004.

Additionally we used the surface reflectance from MODIS (MOD09GA) on 16 July 2004 at 500 m spatial resolution in seven spectral bands [32].

The last source of data was acquired on 18 July 2004 by the Landsat ETM+ with spatial resolution of 30 m including six spectral bands. The atmospheric correction was done by the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software [33]. We used the standard LEDAPS routine. All required input information was taken by LEDAPS from Landsat metadata; ancillary data of (Total Ozone Mapping Spectrometer) TOMS and National Centres for Environmental Prediction (NCEP) Reanalysis.

7.2.3 EO-LDAS

The EO-LDAS [23] is based on weak constraint data assimilation, including a radiative transfer canopy model [5], a leaf spectra model [34] and a spectral soil model [35].

In the following section, we will use prior information (see Table 7.2) [36]. Some values are given in logarithmic scale because EO-LDAS linearizes these parameters by logarithmic transformation.

The only fixed parameter of the canopy model is the leaf angle distribution (LAD) as we will use EO-LDAS over crops fields: it is either fixed as planophile for alfalfa, potato, sugar beet, sunflower and vineyard or to erectophile for corn, garlic and onion.

Reference data are taken from the SPARC field measurements: they represent different crop species and among them both Leaf Area Index and chlorophyll a + b content

Table 7.2 A priori knowledge of state variables in EO-LDAS.

Parameter	Lower limit (non log)	Upper limit (non log)	Prior (non log)
LAI, the single sided leaf area per unit ground area, xlai (log. transform. $e^{(-xlai/2)}$)	0.01(9.2)	0.99 (0.02)	0.05 (5.99)
The canopy height, m, xhc	1.0	5	0.1
The leaf radius/dimension, m, rpl	0.001	0.1	0.01
The concentration of chlorophyll a + b, μg/cm^2, xkab (log. transform. $e^{(-xkab/100)}$)	0.1 (230.26)	0.99 (1.0)	0.1 (230.26)
The proportion of senescent material, scen	0	1	0.001
Equivalent leaf water, cm, xkw (log. transform. $e^{(-xkw*50)}$)	0.01 (0.092)	0.99 (0.0002)	0.99 (0.0002)
Dry matter, μg/cm^2, xkm (log. transform. $e^{(-xkm*100)}$)	0.3 (0.012)	0.99 (0.0001)	0.35 (0.01)
The number of leaf layers, xleafn	1.0	2.5	1.5
Soil PC1 (soil brightness), xs1	0.05	0.4	0.05
Soil PC2 (soil wetness), xs2	–0.1	0.1	0.005
Leaf angle distribution, lad	n/a	n/a	fixed

have been measured. The ground-based data is available for 18 fields of different crop types. These fields were masked manually. This allowed excluding mixed pixels at the field borders.

7.2.4 Observational Operator

Observational operator in EO-LDAS is the combination of the 1-D semi-discrete model for the scattering of light by vegetation; a model of leaf optical properties spectra (PROSPECT) and a spectral soil model. The advantage of the semi-discrete model is that it takes into account the true size of the leaves, which allows the "hot spot" effect of parameterization. This effect has quite a big influence on reflectance measured by a satellite sensor. Conventional 1-D models based only on turbid medium assumption are not able to directly take into account this effect. In the semi-discrete model, the two first orders of scattering are calculated by taking into account the size of leaves in three-dimensional space. Multiple scattering is calculated using a turbid medium assumption. In the semi-discrete model the bidirectional reflectance factor (BRF) is given by Equation [7.6]:

$$\rho\left(\Omega_0,\Omega\right)=\rho^0\left(z_0,\Omega_0,\Omega\right)+\rho^1\left(z_1,\Omega_0,\Omega\right)+\rho^M\left(z_0,\mu_0,\mu\right) \tag{7.6}$$

where ρ^0 is uncollided term; ρ^1 is first order of scattering; ρ^M is multiple scattering.

PROSPECT model is based on the idea of representing a leaf as a number of absorbing plates which have rough surfaces [34]. The model has two groups of parameters, which are leaf structure parameter and several parameters representing leaf biochemical content.

According to Ref. [35] soil reflectance can be obtained by using the following equation:

$$x = \sum_{i=1}^{K} S_i \varphi_i \tag{7.7}$$

Where x is reflectance; K is the number of terms – two for this study; S is soil principal components; φ is the Price Empirical Orthogonal Functions (EOF).

In this study EOF are represented by: φ_1 – mean wet soil spectrum, φ2 – the mean difference between the mean dry and mean wet spectra [37].

7.3 Results

7.3.1 A Prior Constraint in EO-LDAS

The goal of this section is to study the performance of the EO-LDAS *prior constraint* for the retrieval of state variables values. The performance is estimated by comparing the case without and with prior knowledge against ground-based measurements.

In this exercise, the number of CHRIS/PROBA bands is reduced to 17 because of high correlation between hyper-spectral bands [38]; therefore the number of parameters is lower than the number of inputs value.

EO-LDAS without and with prior information is applied over 18 sites where ground-based measurements are available. Inversion was done independently for the three CHRIS/Proba cameras with the fly-by zenith angles 0, +36 and +55 degrees using 17 spectral values. The only fixed canopy parameter is the leaf angle distribution (LAD) which depends on the crop type. Figures 7.2 and 7.3 show the results using logarithmic scale respectively for the chlorophyll content, $e^{-Chl/100}$ and LAI,$e^{-LAI/2}$. Top and bottom panels show the results without prior and with prior by comparing retrieval values with in-situ ones (y-axis), respectively. Dotted symbols show the finding values and the error bar the uncertainties associated to the retrieval. These uncertainties are expressed as the standard deviation (σ).

In the case when EO-LDAS without a priori knowledge runs, the Pearson correlation coefficient r^2 ranges from [0.51–0.74] and [0.61–0.72] for LAI and Cab, respectively. The estimated probability of rejecting or accepting the null hypothesis (two sided p value) for all considered cases is less than 0.05. This means that chance of coincidence is low, the correlations are statistically significant and we can use obtained r^2 values. Root mean square error (RMSE) ranges from 15% to 25%. The change of RMSE between data with prior information and data without prior information is small and can be considered as insignificant. However, the uncertainties associated with the results without prior are very high (top panels of Figures 7.2 and 7.3). This means that the probability density functions (PDF) of estimated biophysical parameters having such a broad range of values have to be used with caution.

When prior constraint are used, the range of r^2 values increases to [0.7–0.8] and to [0.6–0.9] for LAI (bottom panel of Figure 7.2) for Cab (bottom panel of Figure 7.3), respectively. It can be seen that additional information content in the form of prior constraint (Table 7.1) is the reason of decreasing of uncertainties by 90–100%. Prior uncertainty for LAI and Cab in logarithmic scale was equal to 1: the uncertainties for

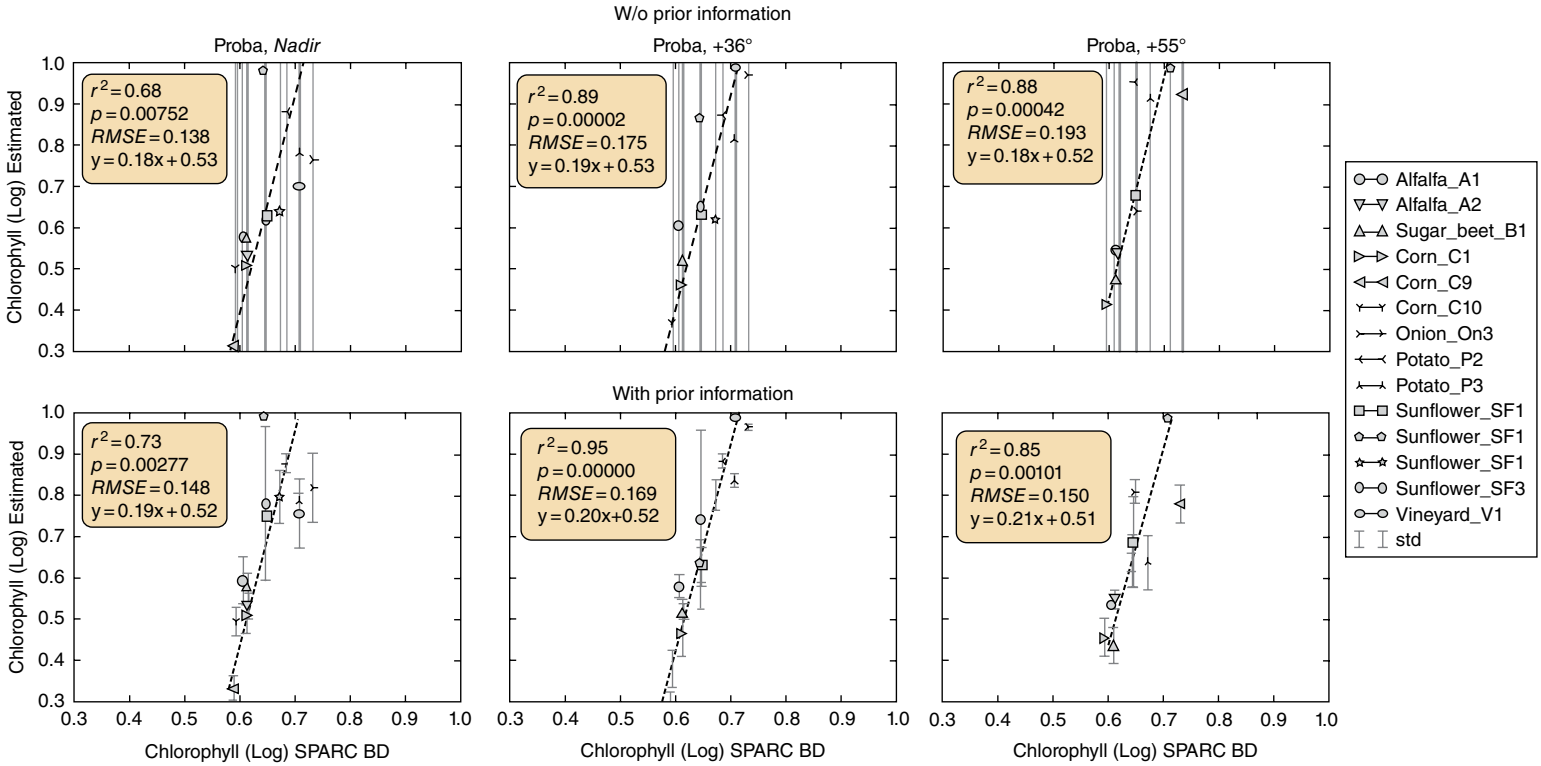


Figure 7.2 Comparison between Chlorophyll a + b content retrieved with EO-LDAS using CHRIS-PROBA data (y-axis) with (bottom panel) and without (top panel) prior information and in-situ measurements (x-axis) over different crop types of the Barrax area. Values are in logarithmic scale.

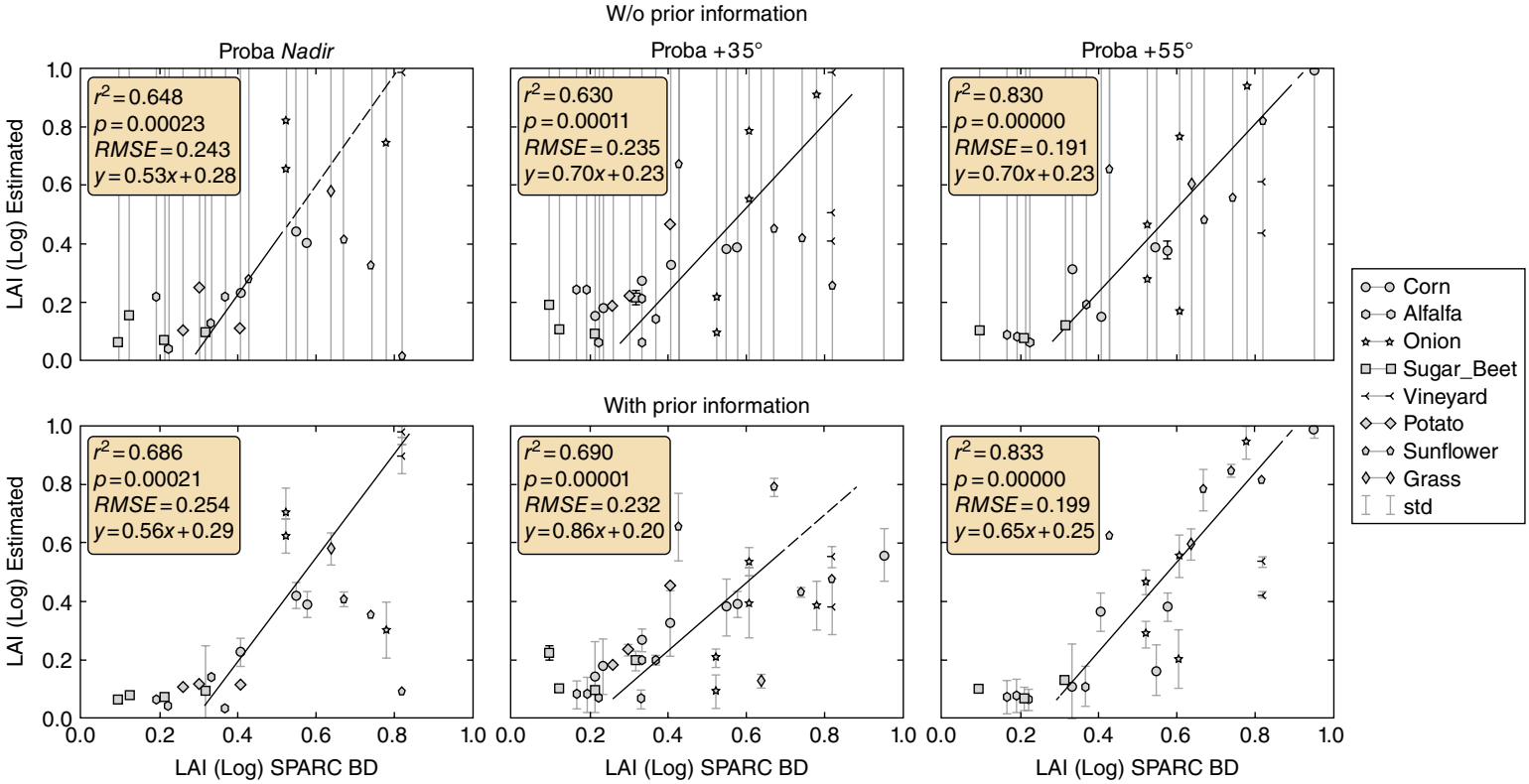


Figure 7.3 Comparison between LAI values retrieved with EO-LDAS using CHRIS-PROBA data (y-axis) with and w/o prior information and in-situ measurements (x-axis) over different crop types of the Barrax area. Values are in logarithmic scale.

retrieved LAI and Cab are less than 0.1. At the same time, the correlation values increase: this can be explained by a better convergence of the cost function. The only exception is the chlorophyll content for camera +55°. However the difference in 0.01 can be considered as an error during optimization process.

7.3.2 Time Regularization with MISR Data

The following exercise explores the time regularization using one year of MISR surface reflectance over the Alfalfa A1 site (39°4'55"N, 2°8'8"W), one of the Barrax fields. The nadir camera (An) of four spectral bands was used. The run was performed for retrieving 10 model parameters (LAD value was fixed to be planophile).

Time regularization imposes temporal smoothness constraint on the surface reflectance data. This constraint is controlled by the regularization parameter γ. However finding optimal γ value can be a challenging and computationally expensive task. One of the possible solutions for obtaining the best γ is cross-validation by removing part of observations and used them for verification of restored temporal dynamic. In this exercise every second observation was removed from 19 observations over 2004.

Figure 7.4 shows a set of solutions for different values of γ.

One can see that when the value of γ is bigger or equal to 200, the dynamic curve of the solution starts to look more stabilized with uncertainties decreased by more than 100% and at $\gamma = 14000$ becomes a flat line with uncertainties close to zero. In order to find the best γ, each set of retrieved parameters values were used in a forward model run to simulate the four MISR bands at the nadir view. The difference between observed surface reflectance data (not used in the optimization) and the simulated one is used as a verification tool. The minimum difference between retrieved reflectance and actual MISR reflectance corresponds to $\gamma = 900$ (Figures 7.4 and 7.5). An important note is that EO-LDAS dynamical "model" doesn't have any information about changing of LAI with time. However with increasing of γ, the LAI curve starts looking like LAI measurements, which include the start of the growing season and the end of the growing season. It demonstrates the ability of EO-LDAS time regularization to constrain the solution.

7.3.3 Spatial Regularization with CHRIS/PROBA Data

Regularization with CHRIS/Proba Only

The spatial constraint is applied with only four CHRIS/Proba spectral bands at 452, 553, 683 and 890 nm. We fix the LAD at planophile or erectophile according to crop type in order to retrieve 10 state variables.

As an example we use the data around the cornfield C9 site (39°4′53″N, 2°7′18″) results of retrieved LAI and chlorophyll content (Cab) values with no space constraint and without prior knowledge are displayed at the top panel of Figure 7.6 and Figure 7.7, respectively. Both retrieved and associated are reported. The ground-based measurement values of LAI are in the range of 2.92 to 3.1 whereas Cab is about 52.94 mg/cm^2. The estimations of LAI are from 3 to 4.5 and Cab in the range of 60–78 mg/cm^2. The uncertainties are displayed in the right column of the top panel of Figures 7.6 and 7.7: These values are quite high for the whole area of the field and correspond to 300% and 400% on average.

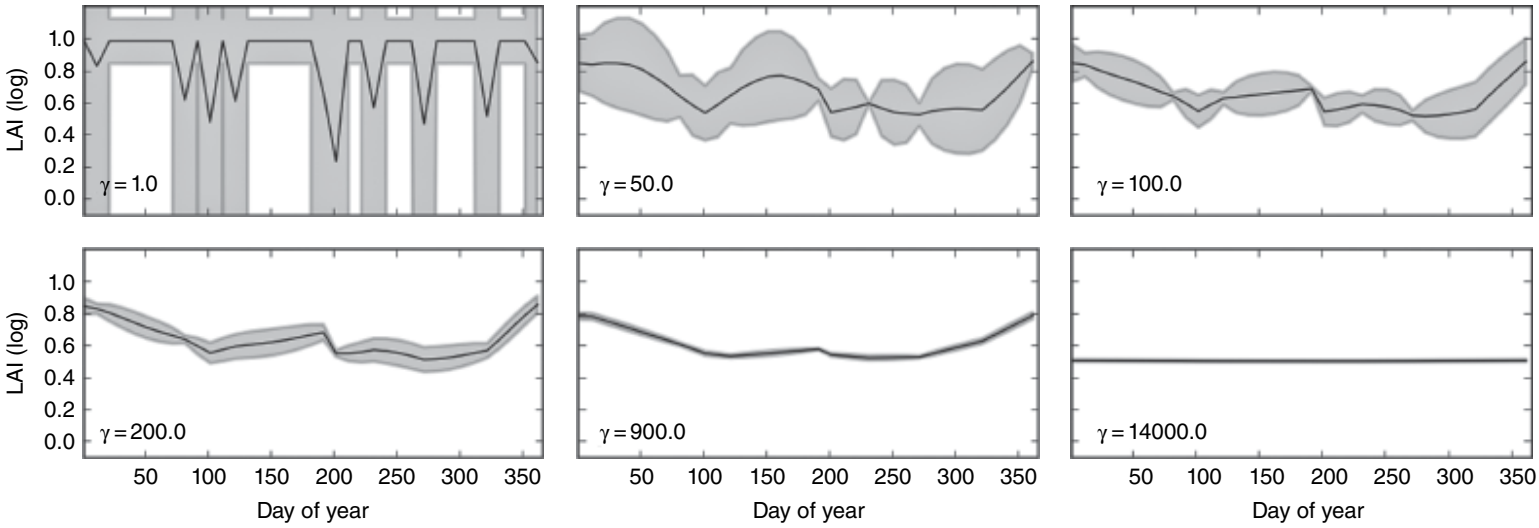


Figure 7.4 Time series of LAI obtained by time regularization of the MISR data of Alfalfa A1 field (39°4′55″N, 2°8′8″W) over 2004 with six different values of γ.

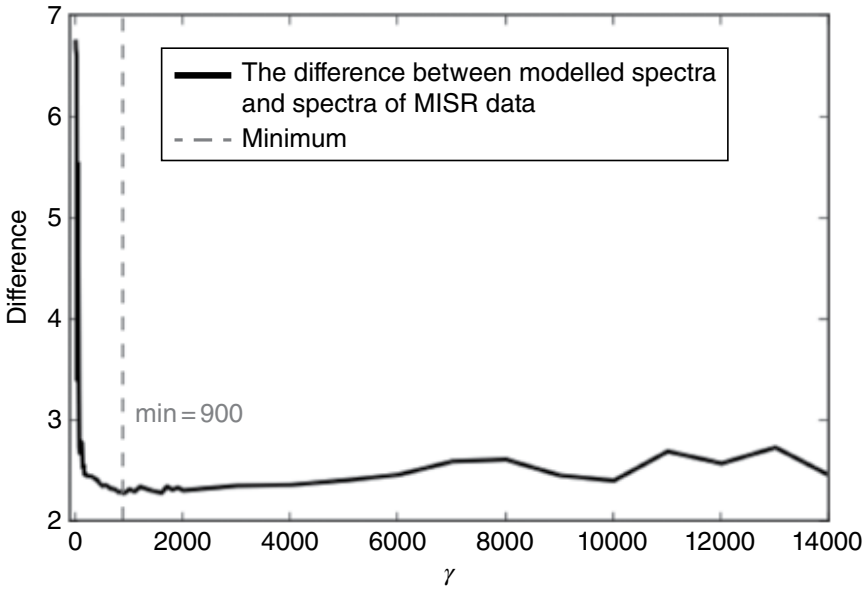


Figure 7.5 Estimation of optimal γ. The difference between time series modelled spectra and spectra of MISR data, which were taken out for cross-validation.

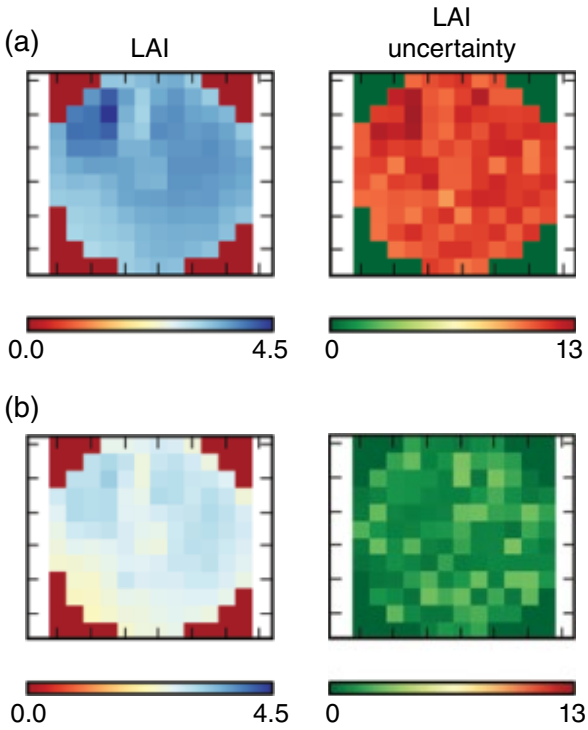


Figure 7.6 Estimation of LAI, CHRIS/PROBA (using 4 bands) over the Corn Field C9 (39°4′53″N, 2°7′18″W). (a) – without priors and spatial regularization (J = Jobs), (b) – with priors and spatial regularization ($J = J_{obs} + J_{prior} + J_{model}$).

When we apply the space constraint we obtain values of LAI in the range of 2.5 to 3.5, and Cab at about 50 to 60 mg/cm^2 (see bottom panels of the Figures 7.6 and 7.7). It means that the spatial regularization and the prior information provide better solutions, close to the range of the ground-based values. In addition, the uncertainties values decrease significantly to 50% and 30% as illustrated in the bottom panels of Figures 7.6 and 7.7.

Figure 7.8 shows that after using EO-LDAS retrieval with prior information and spatial regularization, uncertainties are reduced for all the state parameters from about 60% to 100%.

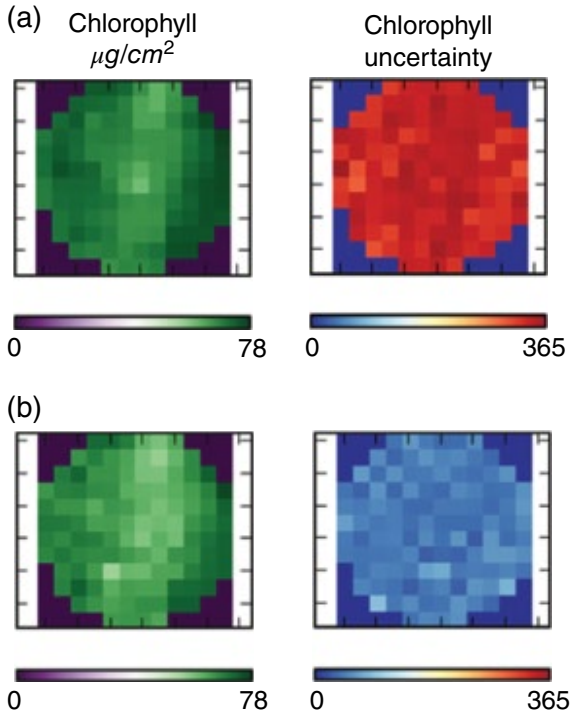


Figure 7.7 Estimation of Chlorophyll a + b content, CHRIS/PROBA (4 bands). Corn, Field C9. (a) – without priors and spatial regularization ($J = J_{obs}$), (b) – with priors and spatial regularization ($J = J_{obs} + J_{prior} + J_{model}$).

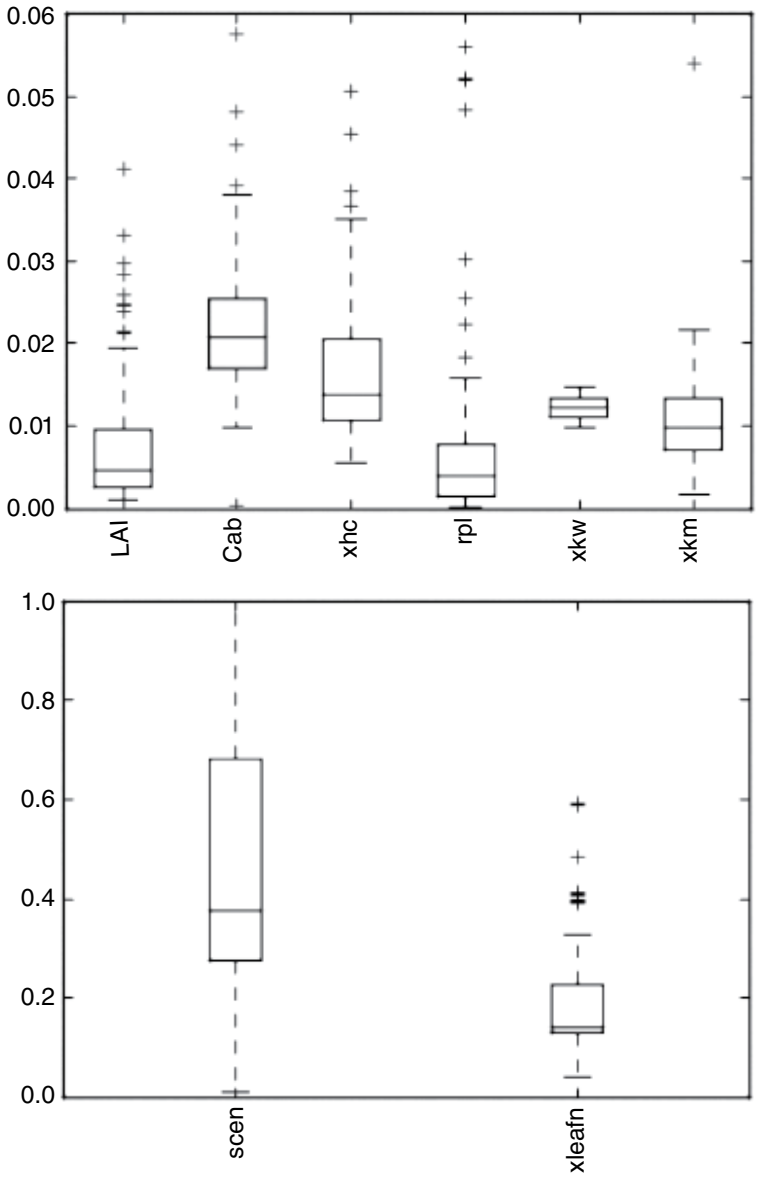


Figure 7.8 Reduction of the uncertainties for eight retrieved canopy parameters: standard deviation with prior and model divided by without prior and model. Parameters: LAI, Cab, canopy height (xhc), leaf radius (rpl), senescent material (scen), leaf water (xkw), dry matter (xkm) and number of leaf layers (xleafn).

Spatial Regularization with CHRIS/Proba and MERIS Data

To test the EO-LDAS efficiency, an additional exercise is run with two optical sensors with different spatial and spectral resolution separately and together. We first use CHRIS/Proba data with only two spectral bands, i.e. the red and near-infrared ones and secondly 15 bands of MERIS/Envisat data. This exercise is performed over the sunflower field (39°4′51″N, 2°6′50″W) with true values of LAI in the range 0.4–0.8 and true values of Cab in the range 43–44 mg/cm^2. The only fixed parameter is LAD, which is set to erectophile. The priors are the same as in the previous example. On the one hand, MERIS does not have enough spatial information for spatial constraining because of the relatively small Barrax field, i.e. size of a MERIS pixel is 300 m, which is comparable to the size of a Barrax field (~300 m). It means no spatial regularization can be made in this case. As a result MERIS-based solution has an overestimation of LAI (2.8–2.9) and underestimation of Cab (5–6 mg/cm^2) (see middle panels of Figures 7.9 and 7.10).

On the other hand, two bands of the high-resolution sensor cannot have enough information for the inversion once 10 model parameters have to be retrieved, despite the use of spatial regularization (upper panels of Figures 7.9 and 7.10). The results show an overestimation of LAI values (0.9–1.22) and an underestimation of Cab ones (9–14 mg/cm^2). In both cases, values of chlorophyll content are quite far away from the ground-based estimates (upper and middle panels of Figure 7.10). However, after solving the problem by following Equation 7.8, i.e. with datasets of the two sensors together,

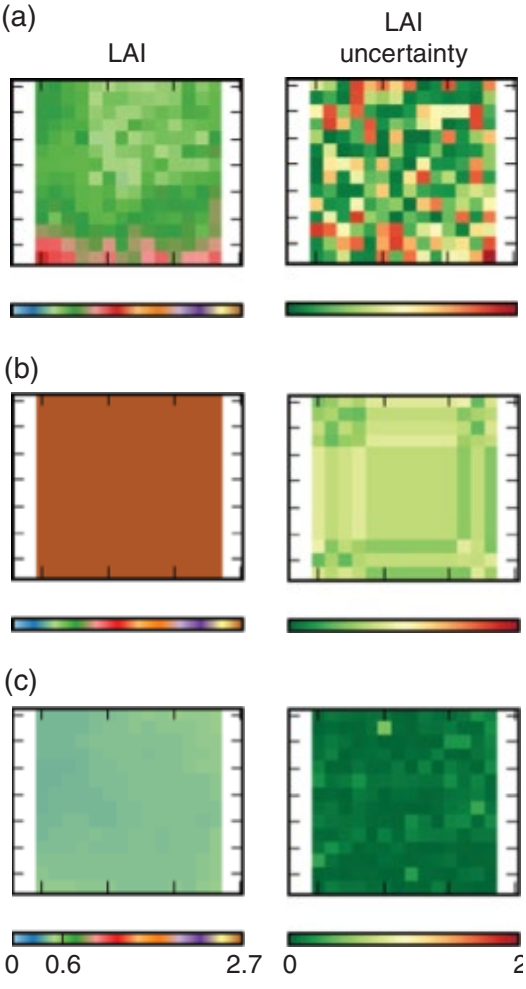


Figure 7.9 Estimation of LAI. Sunflower, Field SF1 (39°4′51″N, 2°6′50″W). (a) – CHRIS/Proba, (b) – MERIS, (c) – CHRIS/Proba + MERIS.

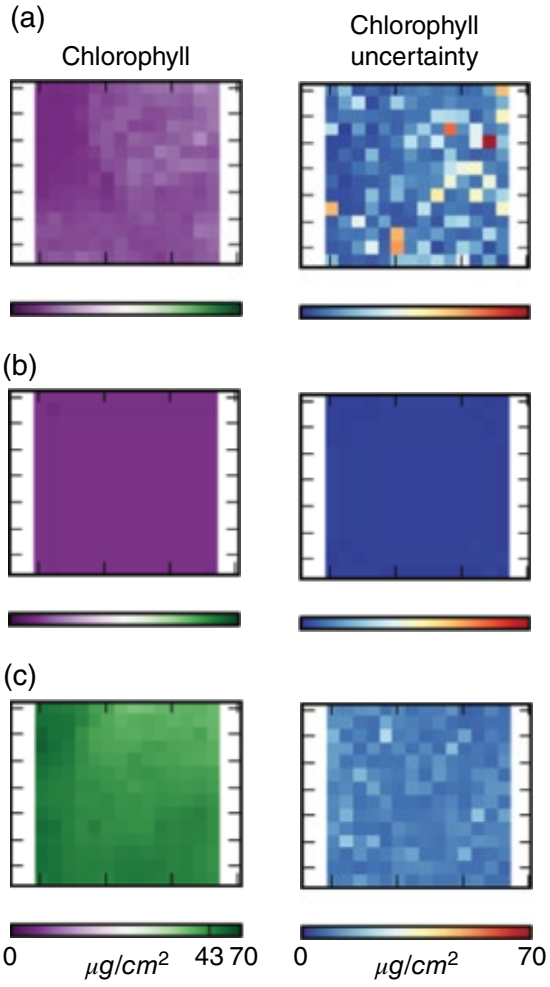


Figure 7.10 Estimation of Chlorophyll content. Sunflower, Field SF1. (a) – CHRIS/Proba, (b) – MERIS, (c) – CHRIS/Proba + MERIS.

values of chlorophyll content are shifted to the range of the estimate values (Bottom panel of Figure 7.10).

$$J_{\text{posterior}}(\mathrm{x}) = J_{\text{space\_model}}(\mathrm{x}) + J_{\text{obs\_lores}}(\mathrm{x}) + J_{\text{obs\_hires}}(\mathrm{x}) + J_{\text{prior}}(\mathrm{x}) \tag{7.8}$$

Thus, using both sensors separately does not retrieve good ground-based estimate values, but when combining them together, the optimal solution is found.

7.3.4 Combination of Low and Medium Resolution Sensors

In this section spectral signatures obtained by EO-LDAS using low and medium resolution sensors, such as ETM+/Landsat, MODIS and MISR, are compared with actual CHRIS/PROBA spectral signatures. Simulated reflectance is calculated for the same geometry of the "nadir PROBA" image for 16.07.2004.

Combination of high and low spatial resolution imagery allows the use of high spatial details of a fine resolution sensor and high temporal frequency of a low-resolution sensor. In addition low-resolution surface reflectance inputs increase information content and may improve quality of filling the time/space gaps in the terms of accuracy and uncertainty.

In the case of Barrax coarse resolution data, such as those of MODIS and MISR, the signal corresponds to mixed vegetation with bare soil. It means that the spectral profile of coarse resolution pixels could mainly correspond to the soil component. The diameter

of Barrax fields, i.e. 300 m on average, is comparable to one MODIS or MISR spatial resolution pixel. So if a ground-based point is measured at the border of a field, the spectral profile of a coarse resolution pixel, which corresponds to this ground-based point, can contain mainly soil component.

Considering relative spectral homogeneity of the Barrax fields, a procedure is implemented when higher resolution pixel spectral measurements are compared with the corresponding coarse resolution ones which have the same geographical coordinates and its eight neighbouring pixels. We consider that a coarse resolution pixel fits a corresponding ground-based measurement point if the difference between its spectral bands values and a higher resolution pixel one is minimal. Here we ignore spectral differences of corresponding bands of ETM+, MODIS and MISR. So this procedure is looking for a best match between high-resolution pixel and 9 low resolution pixels. This procedure helps to combine high and low resolution data. In addition, it can correct some errors of geo-referencing.

EO-LDAS takes into account all available space data and its uncertainties: Therefore, the assumption of this exercise is that higher resolution data is more trusted than coarse resolution. Due to this assumption, standard deviation (SD) for all ETM+ bands was set to 0.01. SD for all MODIS and MISR bands was set to 0.02. Therefore, Landsat data are trusted more.

ETM+ and MODIS

Figure 7.11 illustrates the modelled spectra over Alfalfa A1 (39°5'83"N, 2°8'15"W), after applying EO-LDAS using coarse resolution sensors together with CHRIS/PROBA ones. The input bands of multispectral sensor are over-plotted with symbols.

The left panel of Figure 7.11 corresponds to the solution, which was found when only ETM+ data are used. This solution is depicted as a solid grey spectral curve with corresponding uncertainties. The CHRIS/PROBA spectrum is shown as a solid black line. Triangle and dotted symbols correspond to MODIS and ETM+ bands data, respectively. The middle panel shows the result when only MODIS is used in EO-LDAS. The right panel illustrates the solution when ETM+ and MODIS data are combined. The main difference between ETM+ and MODIS solutions (left and middle panels) is that the latter has lower values in the NIR region, which can be explained by lower resolution of MODIS sensor (500 m), which is the reason for a mixture of signals by vegetation and bare soil. Usually reflectance in the NIR region is greater for vegetation than for inorganic materials. The best agreement between modelled spectra with the reference one is found in the latter case when both MODIS and ETM+ are used. We can observe that the red-edge region (680–730 nm) was well retrieved even if the multi-spectral data have no band in this spectral region. In this example, we can see that ETM+ data pull the solution to higher values in the NIR band whereas MODIS data provide better correspondence in red-edge region.

ETM+ and MISR

The MISR data were used in the EO-LDAS by increasing number of the view zenith angles (VZA) from one to nine. The first VZA was the nadir and then we increase the number of VZA one by one from angle at nadir to angles up to 70° (i.e. D cameras). Figure 7.12 shows that MISR-alone solution is similar to MODIS-alone solution (middle panel of Figure 7.11). However it provides slightly better fit in the NIR region because MISR has better spatial resolution (275 m versus 500 m). The results of the solution based on the combination of MISR and ETM+ show that with the increase of number of MISR cameras impact strongly on the results (Figure 7.13). From one to

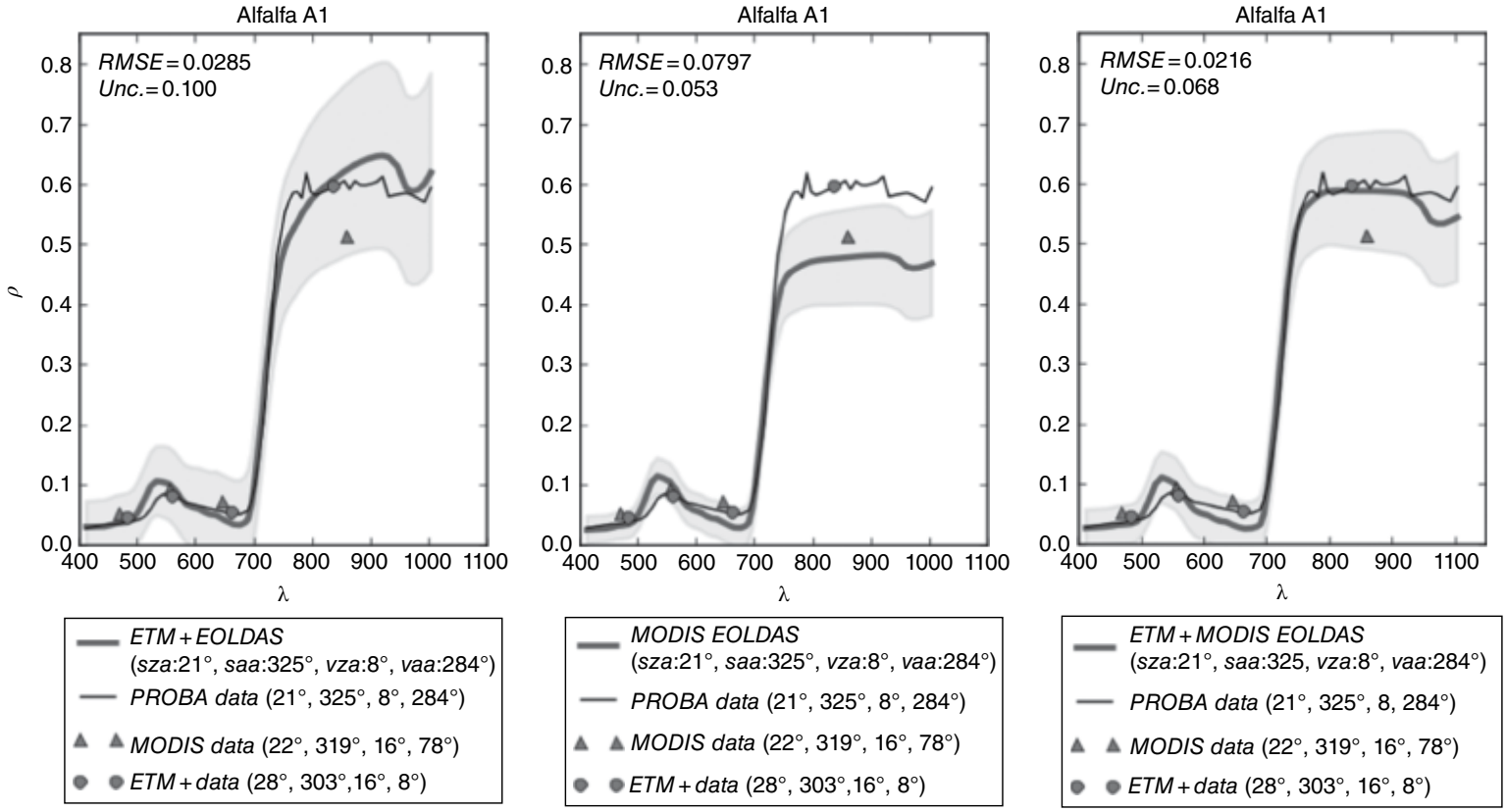


Figure 7.11 Spectral signatures obtained by EO-LDAS with the use of ETM+, MODIS and combination of ETM+/MODIS data over Alfalfa A1 site (39°5′83″N, 2°8′15″W). CHRIS/PROBA view/sun geometry: (sza: 21°, saa: 325°, vza: 8°, vaa: 284°); MODIS: (22°, 319°, 16°, 78°); ETM+: (28°, 303°, 16°, 8°); EO-LDAS reported solutions have the same geometry as CHRIS/PROBA. RMSE values which estimate agreement between Proba and modelled spectra are provided. Unc. values are mean uncertainty of modelled spectra over all bands.

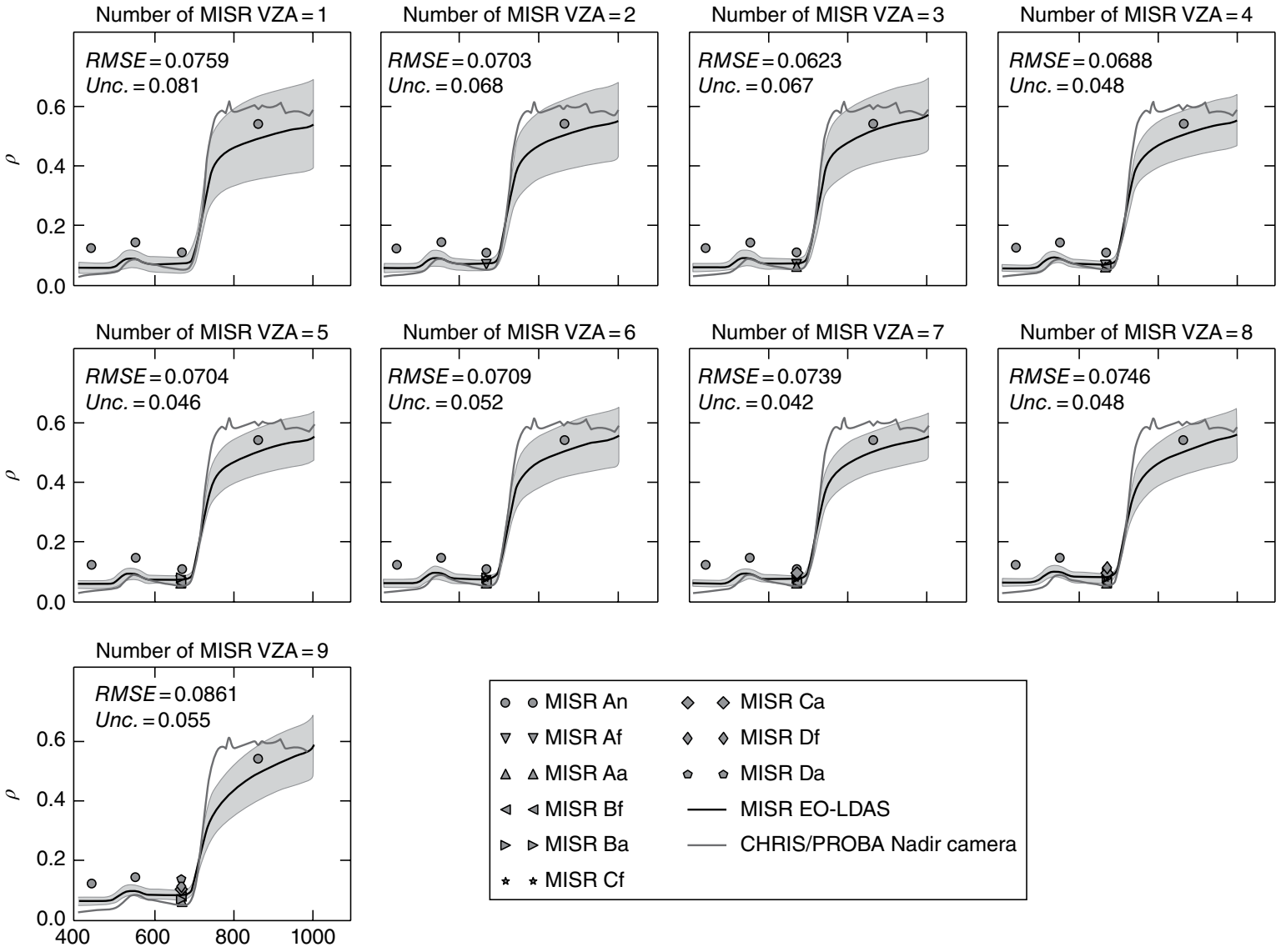


Figure 7.12 Spectral signatures obtained by EO-LDAS with the data of MISR for field Alfalfa A1 (39°5′83″N, 2°8′15″W) over the Barrax. MISR view/sun geometry (sza: 22°, saa: 319°, vza: –70°–70°, vaa: 6°– 345°); CHRIS/PROBA: (sza: 21°, saa: 325°, vza: 8°, vaa: 284°). RMSE values which estimate agreement between Proba and modelled spectra are provided. Unc. values are mean uncertainty of modelled spectra over all bands.

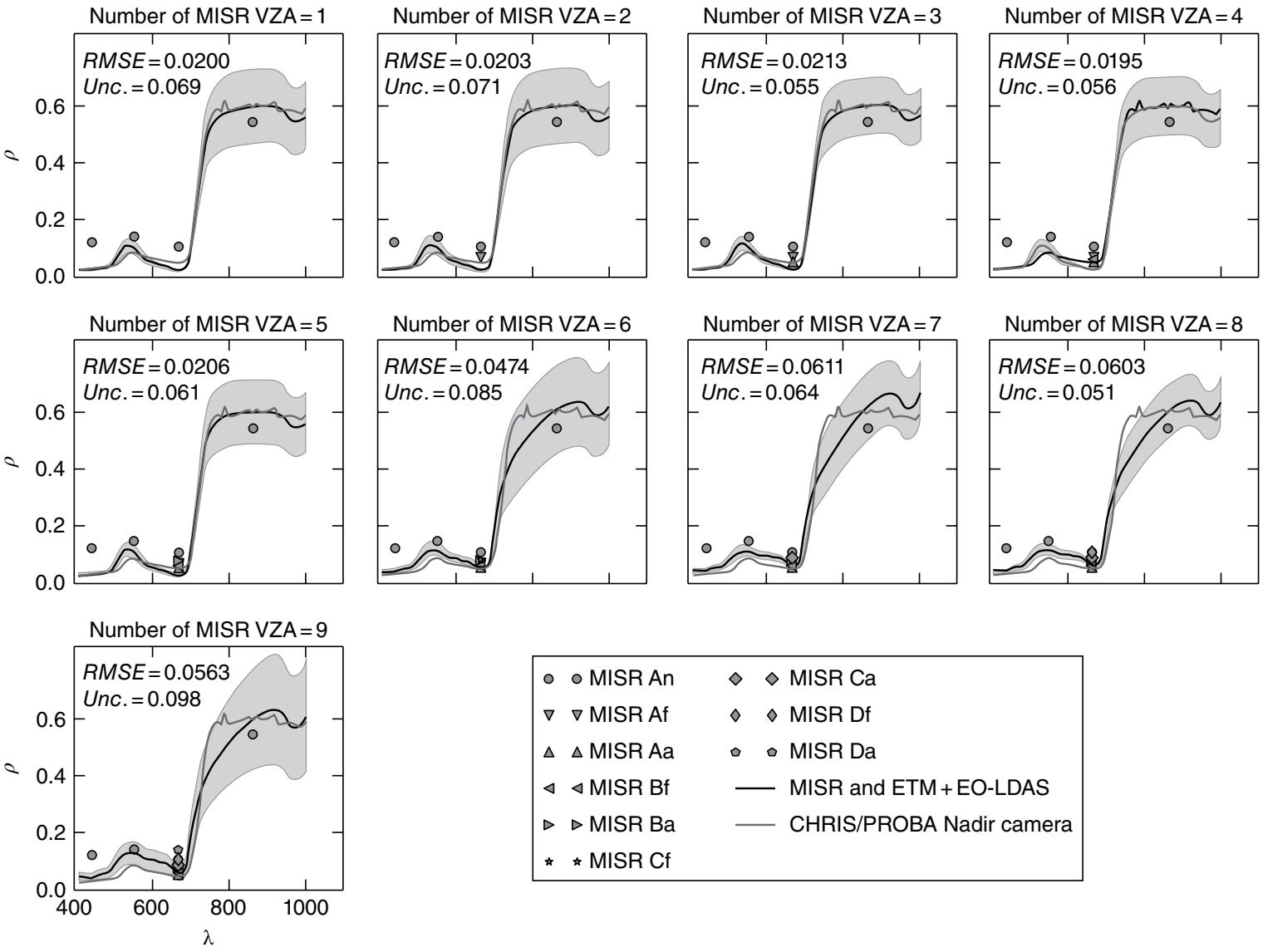


Figure 7.13 Spectral signatures obtained by EO-LDAS with the data of combination of ETM+/MISR for field Alfalfa A1 (39°5′83″N, 2°8′15″W) on the Barrax test site. MISR view/sun geometry: (sza: 22°, saa: 319°, vza: –70°–70°, vaa: 6°– 345°); CHRIS/PROBA: (sza: 21°, saa: 325°, vza: 8°, vaa: 284°); ETM+: (28°, 303°, 16°, 8°). RMSE values which estimate agreement between Proba and modelled spectra are provided. Unc. values are mean uncertainty of modelled spectra over all bands.

five MISR cameras help to improve the results, providing even better fit in the terms of RMSE and lower uncertainties (Figure 7.13) than MODIS and ETM+ solution (right panel of Figure 7.11). The reason is that additional cameras give more information in the red region. However increasing the number of view angles from six to nine has a rather negative effect, i.e. MISR starts to "pull" solution from ETM+. This is because by increasing the amount of MISR view angles, we increase the number of MISR-conditioned constraints. These constraints contain additional information about surface, which supplements ETM+ information, but it is still mixed low-resolution pixel.

7.3.5 Discussion

This contribution overviewed the EO-LDAS functionalities over the Barrax agriculture area in Spain. Various constraints were experimented upon: prior information on the canopy parameters, time regularization, spatial regularization and multi-sensor information.

The three first exercises were used for retrieving the biophysical information in terms of state variables such as LAI and Chlorophyll content. The last experiment was performed to simulate hyper-spectral signatures over the canopy surface.

The results of using prior information show that EO-LDAS is able to estimate leaf area index and chlorophyll concentration on a pixel-by-pixel basis: this helps to decrease the number of possible solutions and significantly decrease the output uncertainty and increase correlation between estimated and ground truth data of LAI and chlorophyll. Correlation ranges from 0.63 to 0.95 for all cases. Average increase of the correlation values when prior information is used is 0.05. RMSE varies from 0.138 to 0.254 without significant change between solutions with prior information and without prior information. However, the uncertainties decrease from 90% to 99% for all considered cases.

Feasibility of the EO-LDAS time regularization was already proven using synthetic time series of Sentinel-2 and actual MODIS time-series [39]. In this work we used four spectral bands of MISR at 275 m and the LAI estimation demonstrates a reduction of uncertainties up to 100%, which means a decrease from mean value 42 to 0.015.

The spatial regularization was explored over small homogeneous fields. However, for the purpose of validation we had to assume that each individual field was homogeneous and one–two ground measurements represent a whole field. In this exercise, we did not take into account the edge problem, but in the case of data that are more realistic, we managed it using clustering, before estimation of the parameters, as proposed in Ref. [21]. Another complexity was the estimation of the regularization parameter γ. We solved this problem by using cross-validation as was proposed in Ref. [23]. Despite the small size of the studied area, the spatial regularization was shown to be efficient by decreasing the uncertainties of retrieved estimates by 70–96%.

The goal of the last section was to demonstrate that we could use the EO-LDAS to simulate hyper-spectral signatures of vegetation. We used Landsat/ETM+, Terra/MODIS and Terra/MISR data. It was shown that combining spectral information of MISR or MODIS to ETM+ can improve results by decreasing the difference with CHRIS/PROBA spectral signature.

7.4 Conclusions

The results of this study show that there is a lot of further work to be done for EO-LDAS applications. One issue is the knowledge of input data uncertainties, especially at the top of the canopy level. There is a requirement to investigate how to propagate uncertainties through the pre-processing chain: i.e. calibration, georeferencing, atmospheric correction, dependency on spatial resolution, etc. This will provide a proper balance between the cost function terms: models, radiometric information and prior information and can significantly improve results, i.e. it can improve the balance of cost function terms and this balance strongly depends on the input uncertainties definition. Further improvements in spatial regularization are related to investigation of clustering schemes, which should help in defining proper homogeneous regions. Also it is essential to make new efforts in understanding how multi-angular information can improve the results of EO-LDAS solutions. Synergy of optical RT inversion with Synthetic Aperture Radar (SAR) information on the basis of physical modelling and data assimilation can provide a lot of new possibilities in whole land remote sensing. Another interesting possibility is the coupling of atmosphere and canopy RT models together in EO-LDAS.

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 the European Space Agency for providing data access by the "EO-Support" system (projects ID 13803 and 13931) and Olivier Morgan (JRC) for his technical support.

References

1 A. Noormets, *Phenology of Ecosystem Processes: Applications in Global Change Research*. Springer, (2009).
2 B. Pinty *et al.*, Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products, *J. Geophys. Res.*, **112**, no. D10, p. D10116, (2007).
3 R.B. Myneni *et al.*, Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, **83**, no. 1–2, 214–231, (2002).
4 J. Dash *et al.*, The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India, *Remote Sens. Environ.*, **114**, no. 7, 1388–1402, (2010).
5 N. Gobron *et al.*, A semidiscrete model for the scattering of light by vegetation, *J. Geophys. Res.*, **102**, 9431–9446, (1997).
6 N. Gobron *et al.*, Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data, *IEEE Trans. Geosci. Remote Sens.*, **35**, no. 6, 1438–1445, (1997).

7 N.J.J. Bunnik, *The multispectral reflectance of shortwave radiation by agricultural crops in relation with their morphological and optical properties*, Wageningen, Netherlands, (1978).

8 W. Verhoef, Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model, *Remote Sens. Environ.*, **16**, 125–141, (1984).

9 B. Pinty *et al.*, Simplifying the Interaction of Land Surfaces with Radiation for Relating Remote Sensing Products to Climate Models, *J. Geophys. Res. – Atmos.*, **111**, no. 2, (2006).

10 P.J. Sellers, Canopy reflectance, photosynthesis and transpiration, *Int. J. Remote Sens.*, **6**, no. 8, 1335–1372, (1985).

11 A. Kuusk, A two-layer canopy reflectance model, *J. Quant. Spectrosc. Radiat. Transf.*, **71**, 1–9, (2001).

12 W. Ni *et al.*, An analytical hybrid GORT model for bidirectional reflectance over discontinuous plant canopies, *IEEE Trans. Geosci. Remote Sens.*, **37**, 987–999, (1999).

13 Y.M. Govaerts, *A model of light scattering three-dimensional plant canopies: a Monte Carlo ray tracing approach*, Ispra, Italy, 1996.

14 R. Myneni and J. Ross, *Photon-Vegetation Interactions: Applications in Optical Remote Sensing and Plant Ecology*. Springer Verlag, New York, (1991).

15 D.S. Kimes *et al.*, Inversion methods for physically based models, *Remote Sens. Rev.*, **18**, pp. 381–439, (2000).

16 A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-Posed Problems*. New York: Winston, (1977).

17 W. Dorigo *et al.*, Retrieving canopy variables by radiative transfer model inversion an automated regional approach for imaging spectrometer data, in *5th EARSeL Workshop on Imaging Spectroscopy*, (2007).

18 C. Atzberger and K. Richter, Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery, *Remote Sens. Environ.*, **120**, 208–218, (2012).

19 Y. Wang *et al.*, Regularizing kernel-based BRDF model inversion method for ill-posed land surface parameter retrieval using smoothness constraint, *J. Geophys. Res.*, **113**, no. D13, p. D13101, (2008).

20 T. Quaife and P. Lewis, Temporal Constraints on Linear BRDF Model Parameters, *IEEE Trans. Geosci. Remote Sens.*, **48**, no. 5, 2445–2450, (2010).

21 V.C.E. Laurent *et al.*, A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from APEX at-sensor radiance data, *Remote Sens. Environ.*, **139**, 6–17, (2013).

22 O. Dubovik *et al.*, Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations, *Atmos. Meas. Tech.*, **4**, no. 5, 975–1018, (2011).

23 P. Lewis *et al.*, An Earth Observation Land Data Assimilation System (EO-LDAS), *Remote Sens. Environ.*, **120**, 219–235, (2012).

24 D. Zupanski, A general weak constraint applicable to operational 4DVAR data assimilation systems, *Mon. Weather Rev.*, **125**, 2274–2292, (1997).

25 R. Houborg *et al.*, Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data, *Remote Sens. Environ.*, **106**, no. 1, 39–58, (2007).

26 C. Lauvernet *et al.*, Multitemporal-patch ensemble inversion of coupled surface–atmosphere radiative transfer models for land surface characterization, *Remote Sens. Environ.*, **112**, no. 3, 851–861, (2008).

27 C. Atzberger, Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models, *Remote Sens. Environ.*, **93**, no. 1–2, 53–67, (2004).

28 X. Liu *et al.*, Multi-angular satellite remote sensing and forest inventory data for carbon stock and sink capacity in the Eastern United States forest ecosystems, in *ISPRS Congress Istanbul 2004*, pp. 1–6, (2004).

29 J.-L. Widlowski *et al.*, Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies, *Clim. Change*, **67**, 403–415, (2004).

30 S. Gandia *et al.*, Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign, in *Proc. 2nd CHIRS/Proba Workshop, ESA/ESRIN, Frascati, Italy*, no. July, (2004).

31 M.A. Cutter and L.S. Johns, *CHRIS Data format* (Product Document). p. 31, (2005).

32 E.F. Vermote and A. Vermeulen, *Atmospheric correction algorithm: spectral reflectances (MOD09).* Algorithm Technical Background Document, University of Maryland, Department of Geography, pp. 1–107, (1999).

33 J.G. Masek *et al.*, *LEDAPS Landsat Calibration, Reflectance, Atmospheric Correction Preprocessing Code. Model product.* Available on-line http://daac.ornl.gov. from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. 10.3334/ORNLDAAC/1080, (2012).

34 S. Jacquemoud and F. Baret, PROSPECT: A model of leaf optical properties spectra, *Remote Sens. Environ.*, **34**, no. 2, 75–91, (1990).

35 J.C. Price, On the information content of soil reflectance spectra, *Remote Sens. Environ.*, **II**, pp. 113–121, (1990).

36 M. Chernetskiy *et al.*, Validation of the Earth Observation Land Data Assimilation System by the field data of ESA SPARC field campaign, in *Proc. ESA Living Planet Symposium 2013, Edinburgh, UK 9–13 September 2013 (ESA SP-722, December 2013)*, vol. **2013**, September, pp. 1–5, (2013).

37 J.L. Gomez-Dans *et al.*, EO-LDAS Validation Report, (2012).

38 F. Vuolo *et al.*, 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).

39 P. Lewis *et al.*, Data assimilation of Sentinel-2 observations: preliminary results from EO-LDAS and outlook, in *Proc. First Sentinel-2 Preparatory Symposium, Frascati, Italy 23–27 April 2012, ESA SP-707, July 2012*, no. July, (2012).

remote sensing

MDPI

Article

Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS)

Maxim Chernetskiy [1,*], Jose Gómez-Dans [1,2,*], Nadine Gobron [3], Olivier Morgan [3], Philip Lewis [1,2], Sina Truckenbrodt [4] and Christiane Schmullius [4]

1 Department of Geography, University College London, Gower Street, London WC1E 6BT, UK; p.lewis@ucl.ac.uk
2 National Centre for Earth Observation (NCEO), Department of Geography, University College London, Gower Street, London WC1E 6BT, UK
3 Directorate D—Sustainable Resources, Knowledge for Sustainable Development and Food Security Unit, European Commission Joint Research Centre, TP 122, Via Enrico Fermi, 2749, 21027 Ispra, Italy; nadine.gobron@jrc.ec.europa.eu (N.G.); olivier.morgan@jrc.ec.europa.eu (O.M.)
4 Institute of Geography, Department for Earth Observation, Friedrich Schiller University, Grietgasse 6, 07743 Jena, Germany; sina.truckenbrodt@uni-jena.de (S.T.); c.schmullius@uni-jena.de (C.S.)
* Correspondence: m.chernetskiy@ucl.ac.uk (M.C.); j.gomez-dans@ucl.ac.uk (J.G.-D.)

Academic Editors: Jose Moreno and Prasad S. Thenkabail
Received: 24 March 2017; Accepted: 22 June 2017; Published: 27 June 2017

Abstract: The Fraction of Absorbed Photosynthetically-Active Radiation (FAPAR) is an important parameter in climate and carbon cycle studies. In this paper, we use the Earth Observation Land Data Assimilation System (EO-LDAS) framework to retrieve FAPAR from observations of directional surface reflectance measurements from the Multi-angle Imaging SpectroRadiometer(MISR) instrument. The procedure works by interpreting the reflectance data via the semi-discrete Radiative Transfer (RT) model, supported by a prior parameter distribution and a dynamic regularisation model and resulting in an inference of land surface parameters, such as effective Leaf Area Index (LAI), leaf chlorophyll concentration and fraction of senescent leaves, with full uncertainty quantification. The method is demonstrated over three agricultural FLUXNET sites, and the EO-LDAS results are compared with eight years of in situ measurements of FAPAR and LAI, resulting in a total of 24 site years. We additionally compare three other widely-used EO FAPAR products, namely the MEdium Resolution Imaging Spectrometer (MERIS) Full Resolution, the MISR High Resolution (HR) Joint Research Centre Two-stream Inversion Package (JRC-TIP) and MODIS MCD15 FAPAR products. The EO-LDAS MISR FAPAR retrievals show a high correlation with the ground measurements ($r^2 > 0.8$), as well as the lowest average $RMSE$ (0.14), in line with the MODIS product. As the EO-LDAS solution is effectively interpolated, if only measurements that are coincident with MISR observations are considered, the correlation increases ($r^2 > 0.85$); the $RMSE$ is lower by 4–5%; and the bias is 2% and 7%. The EO-LDAS MISR LAI estimates show a strong correlation with ground-based LAI (average $r^2 = 0.76$), but an underestimate of LAI for optically-thick canopies due to saturation (average $RMSE = 2.23$). These results suggest that the EO-LDAS approach is successful in retrieving both FAPAR and other land surface parameters. A large part of this success is based on the use of a dynamic regularisation model that counteracts the poor temporal sampling from the MISR instrument.

Keywords: biophysical parameters; inverse problems; FAPAR; leaf area index; radiative transfer; vegetation

1. Introduction

The Fraction of Absorbed Photosynthetically-Active Radiation (FAPAR) is recognised as an essential climate variable, and it plays an important role in biosphere and climate modelling [1]. FAPAR is defined as incident solar radiation in the range 400–700 nm that is absorbed by the photosynthetic tissue of canopy [2] and, thus, is an important control on the photosynthetic activity of vegetation. FAPAR has been widely used for monitoring drought, biodiversity, land degradation, phenology, CO_2 emission studies and Dynamic Global Vegetation Models (DGVM) [3–8]. Although we consider FAPAR to be a land surface parameter (e.g., only related to the land), the amount of direct and diffuse radiation affects its value [9,10].

In order to infer the state of the land surface, the inversion of physically-based models that describe the interaction of incoming radiation with the soil-leaf-canopy medium, typically based on radiative transfer (RT) theory, are generally used [11,12]. The main benefits of using physically-based RT models is their ability to cope with different sensor properties (angular and spectral sampling characteristics, etc.) and that they are more generic than empirical approaches, as they incorporate basic physical laws (e.g., energy conservation) that are universally applicable, and should result in a more robust interpretation of the measurements. The retrieval of land surface parameters using RT models is complicated by the problem being 'ill-posed' [13]. A well-posed problem is one that has a solution; the solution is unique and changes continuously with changing input. A problem that does not hold these conditions is an ill-posed problem [14]. In the context of EO, it often means that an infinite number of land surface parametrisations results in equally likely predictions of the observations. Here, we can see "inputs" as the inputs to RT model, i.e., state variables (LAI, chlorophyll, leaf water content, etc.). One of the possible solutions for improving the situation is using a priori knowledge, such as physically-realistic parameter distributions and/or constraints on parameter smoothness (e.g., in time, space) [15,16]. In practice, ill-posedness means that retrieved parameters have very large uncertainties. Prior information restricts the possible space of potential solutions. This strategy is deployed by the Joint Research Centre Two-Stream Inversion Package (JRC-TIP) product [17]. Other sources of uncertainty in the retrievals arise from sparse observations, e.g., due to cloudiness or orbital and sensor design characteristics. These problems call for a credible and traceable uncertainty quantification framework that allows users to understand shortcomings in the inverted data.

A final comment on FAPAR products is that its magnitude is closely related to other biophysical properties, such as Leaf Area Index (LAI), leaf optical properties, single scattering albedo, etc. These parameters are often derived independently from the same original datasets in RT model inversion schemes, making a number of assumptions on, e.g., canopy structure, leaf optical properties, etc., that might result in inconsistencies between derived products' datasets.

In order to meet the requirements described above, this paper explores the use of the Earth Observation Land Data Assimilation System (EO-LDAS), a general purpose Data Assimilation (DA) framework, to invert a time series of surface directional reflectance observations from the Moderate-resolution Imaging Spectroradiometer (MODIS) and Multi-angle Imaging SpectroRadiometer (MISR) sensors to infer land surface parameters. We use these parameters (and associated uncertainties) to provide a consistent estimation of FAPAR. The results are compared with other products available over the same sites. Two recent papers provide an approachable and non-specialist overview of this area [18,19].

EO-LDAS is a system that allows interpreting spectral observations to provide an optimal quantitative estimate of the Earth surface state. It permits the combination of observations from different sensors despite differences in spatial and spectral resolution and acquisition frequencies. EO-LDAS is based on variational DA and uses physically-based RT modes to map from state (LAI, leaf and soil optical properties, for example) to observation space (in this case, surface directional reflectance). EO-LDAS essentially allows a flexible description of both the 'fit to the observations' using RT models and the prior information, either as parameter distributions, or temporal, or spatial regularisation constraints.

Previous EO-LDAS results have been validated using synthetic Sentinel-2 data [16,20]. In [21], emulators (fast surrogate approximations to computationally-expensive physical RT models) are demonstrated within the EO-LDAS framework, showing that the addition of a simple regularisation dynamic model results in improved retrievals in a synthetic example that combines observations from Sentinel-2/MSI, Sentinel-3/SLSTR and Proba-V observations. In all of these studies, the authors found that adding temporal regularisation as an additional prior constraint resulted in a significant reduction of uncertainty in the estimates of the inferred land surface parameters. Here, we analyse the results of EO-LDAS temporal regularization with MISR observations by comparing them against ground-based FAPAR estimated over an agricultural test site [22] and against MEdium Resolution Imaging Spectrometer (MERIS) FAPAR at 300 m, MISR High Resolution (HR) JRC-TIP and MODIS MOD15 [23] products for 2001–2008.

Multi-angular remote sensing of the land surface can help to reveal structural properties of the vegetation and thus to improve characterisation of the vegetation cover [24]. Three widely-used instruments provide multi-angular information at the global scale: Polarization and Directionality of Earth Reflectance (POLDER) [25], Sea and Land Surface Temperature Radiometer (SLSTR) [26] and MISR [27]. MISR has nine cameras pointed to directions from $-70°$–$70°$, four spectral bands from blue to near-infrared and spatial resolution at 275 m for the nadir camera and red band and at 1.1 km for others. A number of studies has demonstrated that exploiting multi-angular information of MISR can improve retrieval of LAI and FAPAR [28–31]. One physically-based approach for deriving FAPAR from 275 m is the JRC-TIP approach that uses MISR resolution data [32,33] to invert a radiative transfer model. This package is based on a two-stream model [34] and uses prior information to constrain the RT model inversion. In addition, this application provides information about theoretical uncertainties for both output state parameters and output fluxes. JRC-TIP output was tested against independent parameter estimates over a range of different areas [17,35].

The next sections describe the test site, EO data and FAPAR retrieval algorithms with their respective definition and assumptions. Following this, we present a short summary of the theoretical basis of EO-LDAS and FAPAR estimation. We then present the retrieved values from EO-LDAS and proceed to compare them with ground measurements. We also show and discuss comparisons with other products and discuss these comparisons. We finally draw some conclusions.

2. Materials

2.1. Test Site Description and Ground-Based Data Collection

The study area is comprised of the three agricultural FLUXNET sites, US-Ne1 (41.165°N, 96.477°W), US-Ne2 (41.1649°N, 96.470°W) and US-Ne3 (41.1797°N, 96.4730°W) (http://fluxnet.ornl.gov) located at the Lincoln Agricultural Research and Development Center near Mead (NE, USA) (Figure 1). Each field has an area of approximately 65 ha. The US-Ne1 site was cultivated with maize from 2001–2008. The sites US-Ne2 and US-Ne3 were cultivated with maize in 2001, 2003, 2005 and 2007 and with soy bean in 2002, 2004, 2006 and 2008 [22,36]. Fields US-Ne1 and US-Ne2 are irrigated, and US-Ne3 is rainfed. The growing season is approximately from May–October.

In situ measurements of FAPAR were carried out from 2001–2008 between June and September/October with an interval of 2–6 days. These measurements were performed with a Li-Cor quantum sensor (LI-COR Inc., Lincoln, NE, USA) by detecting incoming photosynthetically active radiation (PAR_{inc}), PAR reflected by the canopy and soil (PAR_{out}), PAR transmitted through the canopy (PAR_{transm}) and PAR reflected by the soil (PAR_{soil}) (Equation (1)) [36].

$$FAPAR_{total} = \frac{PAR_{inc} - PAR_{out} - PAR_{transm} + PAR_{soil}}{PAR_{inc}} \quad (1)$$

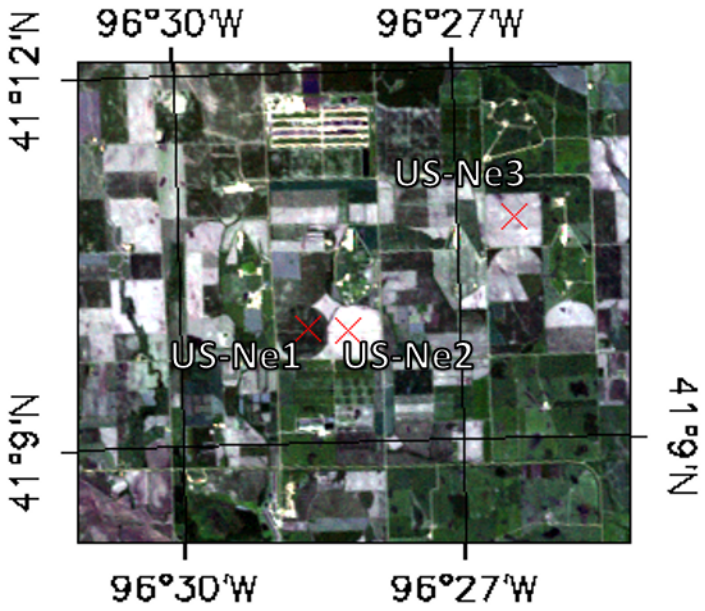


Figure 1. Map of the test site (Landsat). Red crosses are the points of ground measurements.

Hourly values of radiation measured throughout a day were integrated in order to get daily values when PAR_{inc} was >1 µmol m^{-2} s^{-1} [22]. A detailed description of the measurements protocols can be found in [22,36]. In summary, the authors obtain the absorption PAR for only green elements, i.e., $FAPAR_{green}$, with the following correction for which LAI_{green} and LAI_{total} have been measured through a destructive determination technique:

$$FAPAR_{green} = FAPAR_{total} \frac{LAI_{green}}{LAI_{total}} \quad (2)$$

It should be noted that this method does not take into account senescent leaves, which can lead to some underestimation of FAPAR at the end of the growing season.

One important technical issue concerns the different spatial scales of the ground and satellite-derived estimates, as well as the geo-location differences between in situ and MISR observations. Another problem is the way that in situ green FAPAR has been calculated (see Equation (2)): the common assumption of a linear correlation between LAI and FAPAR will inevitably introduce some errors.

2.2. Remote Sensing Data and Products

2.2.1. MISR Observations

The EO-LDAS inversions use the MISR full resolution surface reflectance as inputs, i.e., at 275 m, using seven cameras acquiring in the nadir direction (camera: An), and (60°, 46° and 26° afterwards (cameras: Ca, Ba, Aa) and forward (cameras: Cf, Bf, Af). The acquired data have been pre-processed with the sharpening method of [33]. MISR has four spectral bands with central wavelengths at 446 nm, 558 nm, 672 nm and 867 nm [37]. The test area is acquired by MISR on the paths P27, P28 and P29. In every acquisition, the pixels that are closest to the positions of the flux towers were selected as being representative of the associated fields. This means that over the entire time period, the ground location of the selected pixels may change slightly. In order to maintain the native MISR pixel resolution, we did not re-sample pixels. Spatial standard deviation of surface Bidirectional Reflectance Factors (BRFs) in the 3 × 3 pixel area for eight years does not exceed 0.04, except for four dates (not shown here).

The surface BRF uncertainties were assumed Gaussian, with a zero mean and with a standard deviation of 0.05 units of reflectance for all bands. The uncertainty between bands was assumed uncorrelated.

2.2.2. The JRC-TIP MISR Product

The JRC-TIP approach generates surface biophysical products using broadband surface albedo in visible and near-infrared domains by the minimisation of a cost function $J(\vec{x})$, derived from the log posterior distribution:

$$J(\vec{x}) = -\frac{1}{2}\left[\left[\vec{d} - H_t(\vec{x}^t)\right]^\top \mathbf{C}_o^{-1}\left[\vec{d} - H_t(\vec{x}^t)\right] + \left[\vec{x}^t - \vec{x}_p^t\right]^\top \mathbf{C}_p^{-1}\left[\vec{x}^t - \vec{x}_p^t\right]\right] \quad (3)$$

where $\vec{d}$ is the MISR white sky albedo in visible and near-infrared broadbands; $\vec{x}^t$ is the state vector, which is comprised of the inputs of the two stream model $H_t(\vec{x}^t)$ [35]. The state parameters are: the effective Leaf Area Index (LAI); the background albedos (BA) $r_g(\lambda)$; the Single Scattering Albedo of leaf (SSA):

$$\omega_l(\lambda) = r_l(\lambda) + t_l(\lambda) \quad (4)$$

and the backward/forward scattering efficiency $r_l(\lambda)/t_l(\lambda)$ in each broadband. The ratio $r_l(\lambda)/t_l(\lambda) < 1$ ($r_l(\lambda)/t_l(\lambda) > 1$), then forward (backward) scattering is predominant; $\vec{x}_p^t$ is the state vector of prior information; $\mathbf{C}_o$ is the covariance matrix of observations, which is responsible for the uncertainty in the observations; $\mathbf{C}_p$ is the covariance of prior information. An efficient minimisation of Equation (3) is achieved by exploiting the adjoint code (the adjoint provides an efficient estimate of the gradient of a function and can be obtained using automatic differentiation tools [38]) of the two stream model [35], which allows the use of gradient descent methods. These methods exploit the availability of the gradient of the cost function provided by the adjoint to propose the search direction of the optimisation. Equation (3) can be seen as the sum of two terms: $J_{obs}(\vec{x})$, which is responsible for the observations, and $J_{prior}(\vec{x})$, which is the prior term:

$$J(x) = J_{obs}(\vec{x}) + J_{prior}(\vec{x}). \quad (5)$$

The surface reflectance values used in EO-LDAS and the surface albedo values used in the JRC-TIP correspond to the same pixel with a resolution of 275 m and have been derived with the same pre-processing chain used in the sharpening method and subsequent atmospheric corrections [33]. The JRC-TIP assumes that the Leaf Angle Distribution (LAD) is spherical and that the a priori leaf spectra values correspond either to 'polychrome' (standard) or 'green' leaf. The polychrome leaf is vaguer about the leaf single scattering albedo, whereas the green leaf has a very tight prior distribution of single scattering albedo that is broadly consistent with a healthy green leaf [17]. We will include this latter assumption, as it is consistent with choices in the MERIS algorithm [39,40]. The JRC-TIP FAPAR corresponds to the post-processing absorbed fluxes in the visible broadband, i.e., in the PAR domain. However, we note that for the JRC-TIP, the absorption corresponds to the white-sky value, i.e., absorption under diffuse radiation.

2.2.3. The JRC MERIS FAPAR Product

The MERIS FAPAR Full Resolution products come from the operational ESA products. They do not yet contain associated uncertainties or/and updated cloud masking, which is planned for the fourth reprocessing [41,42]. In this work, we use a version with uncertainties. The design of the MERIS FAPAR retrieval is based on a two-step procedure, where the spectral radiances measured in the red and near-infrared bands are first rectified in order to ensure their decontamination from atmospheric and angular effects. The outputs are then converted into FAPAR by using a function that has been defined by fitting it to pairings of input rectified reflectances and output FAPAR [40]. This retrieval method assumes that the leaves are alive and photosynthesising, hence the name green FAPAR. We assume that $\omega_l(\lambda)$ and the leaf single scattering albedo are fixed to a value representing standard

leaves using standard biochemical leaf properties [41,42]. In this study, we average MERIS values over 3×3 pixels around the central pixel in order to minimise the impact of remaining clouds and cloud shadow contamination [41,42].

2.2.4. The MODIS FAPAR Product

Finally, we also use the MODIS FAPAR product (Product MCD15A2H, Collection 6 [23]). The MODIS FAPAR product is produced at 500-m spatial resolution, which is a coarser resolution than the MISR (275 m) or MERIS (300 m) products, which can potentially lead to inconsistencies due to the spatial heterogeneity. The MODIS FAPAR product uses a look-up table approach to invert the reflectances acquired by the TERRA and AQUA sensors over a particular time window in the red and near-infrared bands [28,43]. A radiative transfer model is used to populate the Lookup Table (LUT), with some assumptions on the optical properties of the leaves and soil used to constrain the inversion.

3. Methods

3.1. The EO-LDAS Approach

The EO-LDAS scheme [16] is a generic land Data Assimilation (DA) system, which uses a set of observational operators together with prior information and a process model to provide an inference on the state of the land surface that is a consistent interpretation of the observations, prior information and dynamic model. The inference on the different land surface parameters is quantified as a full-probability density function (pdf), which encodes the uncertainty in the state. EO-LDAS is implemented as a variational system, where the dynamic model is implemented as a weak constraint [44]. In this work, we have used the `eoldas_ng` Python implementation available from [45].

A priori information reduces the volume of the solution space [15]. The inclusion of a priori information derives directly from a Bayesian understanding of the inverse problem. In this sense, the a priori information is a probability density function (pdf) that describes the expected distribution of the state. Typically, normal or uniform distribution (the latter just indicating parameter boundaries) have been used [16,46]. In this contribution, the prior distribution is Gaussian as required by the variational framework used and is made fairly uninformative (e.g., with a large variance) in order to test how the system generalises.

A particularly useful form of a priori information exploits the often smooth nature of the temporal or spatial evolution of the land surface state. These so-called "regularisation" methods [47–54] assume temporal and/or spatial correlation as part of the prior distribution, resulting in a much reduced uncertainty [16,21]. In a similar vein, there are DA methods that exploit predictions of the land surface state from a dynamic vegetation model (typically a function of LAI, FAPAR) [55]. A main disadvantage in the dynamic model approach is the lack of suitable models of the temporal and/or spatial evolution of many of the variables that have a direct control on the observations (e.g., equivalent leaf water or leaf chlorophyll content). Regularisation in this sense is the application of a zero-order model on the evolution of the land surface parameters [16,21,47].

The fundamental task of EO-LDAS is to infer the land surface state by minimising a cost function made up of three terms:

$$J(\vec{x}) = J_{obs}(\vec{x}) + J_{prior}(\vec{x}) + J_{model}(\vec{x}), \quad (6)$$

where $J_{obs}(\vec{x})$ is the observational constraint (or the fit to the data component); $J_{prior}(\vec{x})$ is the prior constraint, which includes the departure of the state from its prior normal distribution, and $J_{model}(\vec{x})$ is the dynamic model constraint, which penalises trajectories of the land surface that depart from those given by a dynamic model. The elements of $\vec{x}$ are shown in Table 2. In all three cases, the statistics are assumed to be Gaussian. Equation (6) is in effect the logarithm of the posterior of the inverse problem EO-LDAS tries to solve. Under the assumption of Gaussian statistics and weak non-linearities, the minimum of $J(\vec{x})$ coincides with the Maximum A Posteriori (MAP)

of the inferred state, and the uncertainty can be calculated as the inverse of the Hessian matrix at the MAP point [16,56]. As in [16,21], we use the parameter transformations in [57] (shown in Table 2) to provide a quasi-linearisation of the model.

The posterior uncertainties are determined as the main diagonal of the covariance matrix, which is an inversion of the Hessian matrix. i.e., uncertainties are estimated as level of steepness of curvature of decision space. The more data we have, the more chances to have a single deep minimum and correspondingly lower uncertainties. If we have less data or data that cannot be described by a model, the minimum becomes flatter, and the local minimum appears. The inverse of the Hessian provides a good approximation of covariance matrix for not too non-linear models [56].

3.2. Fit to Observations

The term responsible for fitting to the observations is J_{obs}, in effect, the log-likelihood:

$$J_{obs}(\vec{x}) = -\frac{1}{2}(\vec{R} - H(\vec{x}))^{\top}\mathbf{C}_o^{-1}(\vec{R} - H(\vec{x})) \tag{7}$$

where $H(\vec{x})$ corresponds to an observation operator and $\mathbf{C}_o$ is the covariance matrix describing the uncertainty in the observations. The observational operator $H(\vec{x})$ is implemented as the coupling of the semi-discrete canopy RT model of [58], the spectral leaf optical properties RT model (PROSPECT) of [59] and an adapted version of the spectral soil model of Price [60]. Note that this $H(\vec{x})$ term is equivalent to the first term in the right-hand side of Equation (3), with the only difference here being that we operate with reflectance measurements, and the JRC-TIP uses broadband albedos. The observation operator is also consequently different.

The observational operator consists of three main terms:

$$H(\vec{x}) = R_0 + R_1 + R_m \tag{8}$$

where R_0 corresponds to the BRF due to zero order scattering or in other words absence of scattering, only two transmissions through the canopy and reflection by soil background; R_1 and R_m are the BRFs due to first order and multiple scattering, respectively.

In EO-LDAS the soil background reflectance is implemented as:

$$R_0 = s_1\phi_1 + s_2\phi_2 \tag{9}$$

where ϕ_1 and ϕ_2 are Price's basis functions [60], weighted by two scalars, to be inferred. This model assumes a Lambertian soil. In order to infer the values of the Price functions, we have taken MODIS observations daily surface reflectance observations between 2009 and 2015 and selected observations between Days of Year 1 and 60, where NDVI was less than 0.25, assuming that these two conditions would result in observed bare soil. We fitted regularised linear kernel models [47] to calculate nadir illumination, nadir viewing reflectances and fitted the first two Price spectral basis functions over the seven MODIS bands using a standard least squares approach. The resulting values (see Table 2) were then used as a description of the soil throughout.

We assume independence of observations for different bands and cameras. However, due to the atmospheric correction and sharpening procedures, the independence assumption between cameras and bands might not strictly hold. The magnitude of these potential correlations is unknown, so we ignore them here, i.e., $\mathbf{C}_o$ is a diagonal matrix with per band variances on the main diagonal.

3.3. The Prior

The prior constraint is written as:

$$J_{prior}(\vec{x}) = -\frac{1}{2}(\vec{x} - \vec{x}_p)^\top \mathbf{C}_p^{-1}(\vec{x} - \vec{x}_p) \tag{10}$$

where $\mathbf{C}_p$ is the covariance matrix which describes uncertainty of the prior state; $\vec{x}_p$ is the vector of prior means, again as it is for the JRC-TIP (Equation (3)). Equation (10) states that in EO-LDAS, the prior distribution must always be Gaussian.

In the case of LAI, we propose a simple model of temporal development based on a double logistic function [61], which can be appropriate for vegetation exhibiting a clear (single, annual) phenology in the Northern Hemisphere:

$$LAI(t) = wP + (mP - wP) \times (\frac{1}{1 + \exp(-mS(t - S))} + \frac{1}{1 + \exp(-mA(t - A))} - 1) \tag{11}$$

where wP and mP are the expected minimum background and maximum values of LAI throughout the year; S is the spring 'green up' date (increasing point of inflection); A is the autumn 'senescence' or 'brown down' date (decreasing point of inflection); mS (mA) is related to the rate of increase (decrease) at the point of inflection S.

Equation (11) provides an estimate of the mean LAI, and for the associated standard deviation, we have assumed a Gaussian temporal distribution:

$$SD = \frac{G}{max(G)} + b \tag{12}$$

where b is the background value. SD is the LAI standard deviation associated with G, the Gaussian normal distribution:

$$G = \frac{1}{\sqrt{2\pi}\delta} e^{\frac{-0.5(x-\mu)^2}{\delta^2}} \tag{13}$$

where μ is the mean of G and σ the standard deviation of G. We have used the same functional shape for LAI, leaf chlorophyll content and leaf equivalent water thickness. The values used in these functions are summarized in Table 1. These values come from fitting the double logistic function against AVHRR LAI over 20 years [62], and we thus obtain mP for LAI and S, A, mS and mA for LAI, chlorophyll and senescence. wP and mP are chosen as reasonable upper/lower bounds.

Table 1. Summary of the double logistic model and Gaussian distribution parameters.

	S	***A***	***mS***	***mA***	***wP***	***mP***	***μ***	***σ***	***b* (log)**
LAI	175	245	0.04	0.05	0.15	4.22	200	30	0.25
Leaf chlorophyll content	175	245	0.04	0.05	1	90	200	40	0.05
Proportion of senescence material	175	245	0.04	0.05	0.001	0.7	200	70	0.05

Having this trajectory of standard deviation assumes that the model in Equation (11) is suitable for winter (low uncertainty (but also low, near-zero absolute values of vegetation cover), but the model is chosen to be uninformative in summer (very high uncertainty) (Figure 2). This is due to the fact that there are few observations during winter due to clouds, snow, etc.

We use this model to provide a prior mean and variance for LAI, leaf chlorophyll content and senescence. The remaining components of the state vector are assumed known and set to the values indicated in Table 2. The purpose of the model is to approximately predict seasonal development in the case where there are no satellite data for a long period, which is typically the case at the start and end of the year. Due to the large variance introduced by Equation (12) over the vegetation period,

the prior only has an influence at the beginning and end of year, being otherwise uninformative when the vegetation is active.

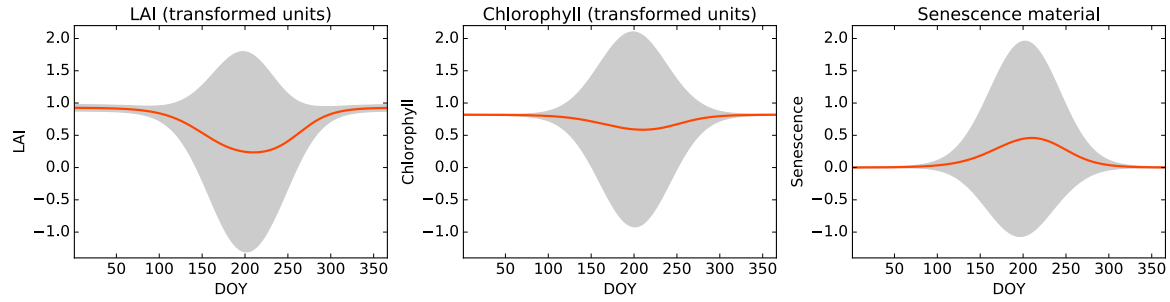


Figure 2. Prior information for LAI (transformed units), leaf chlorophyll content (transformed units) and proportion of senescent material. The shaded region represents the uncertainty range in each case. Parameter transformations are shown in Table 2.

3.4. Temporal Regularisation

The dynamic model provides a prediction of the temporal evolution of the land surface state. We make the simplest possible assumption, i.e., that over time, the land surface parameter does not change. However, clearly, this model has an error. In other words, we assume that the difference between the state between two consecutive time steps is Gaussian, with mean zero and a particular variance. The rationale behind this very simple model is that solutions where a large high frequency component is present are penalised, resulting in a smooth temporal evolution of parameters, as expected for the parameters of interest here and in the scales that are being considered. This constraint can be written as:

$$J_{model}(\vec{x}) = \frac{\gamma^2}{2}\vec{x}^{\top}(\mathbf{D}^{\top}\mathbf{D})^{\top}\vec{x} \tag{14}$$

where γ is the regularization parameter, which represents in this case the inverse of the model error variance and controls the smoothness of retrievals. $\mathbf{D}$ is the differential operator of the first order. Note that this constraint is also prior information (as stated above, it imposes a Gaussian distribution on first differences of parameters). We applied temporal regularisation to all state parameters listed as "dynamic" in the Table 2. The uncertainty associated with each parameter (i.e., γ in Equation (14)) was estimated by cross-validation. Each parameter was given a different value of γ, and the optimal value of γ changed from year to year, reflecting the different number of observations available every year.

In this study, the Leaf Angle Distribution (LAD) is prescribed as a spherical distribution. Some testing of other distributions suggest that this was not a major influence, resulting in differences in retrieved mean a posteriori FAPAR by 3–5%.

To sum up, in this work, we have solved for LAI, leaf chlorophyll content and senescent material assuming these parameters evolve with time and using the regularisation assumption described above. All other parameters were assumed known and prescribed to values given in Table 2. The other prescribed spectral parameters have little or no effect in the spectral range of the considered MISR observations [63]. The inferences on the three state components were done every five days.

Table 2. Summary of the state parameters.

Name	Symbol	Units	Default or Prior Value	Lower Limit	Upper Limit	Prior STD (Transf. Units)	Transform
Leaf Area Index (LAI)	LAI	$(m^2 \cdot m^2)$	Dynamic	0.02	8.4	Dynamic	$e^{(-\frac{LAI}{2})}$
Canopy height	xh	(m)	1	0.05	10	1	-
Leaf radius	xr	(m)	0.1	0.01	0.1	1	-
Chlorophyll a,b	C_{ab}	$(mg \cdot cm^{-2})$	Dynamic	20	51	Dynamic	$e^{(-\frac{C_{ab}}{100})}$
Proportion of senescent material	C_{sen}	na	Dynamic	0.001	1	Dynamic	-
Leaf water	C_w	(cm^{-1})	0.0001	0.00002	0.092	1	$e^{(-50 \cdot C_w)}$
Dry matter	C_{dm}	$(g \cdot cm^{-2})$	0.00005	0.00001	0.012	1	$e^{(-100 \cdot C_{dm})}$
Leaf layers	N	na	1.9	1	5	1	-
Soil PC1	S_1	na	1.22	0.5	2	1	-
Soil PC2	S_2	na	1.32	-1	1.5	1	-
Leaf angle distribution	LAD	na	Spherical (Uniform)				-

3.5. Gaussian Process Emulators

A limitation of variational DA methods is that they require multiple evaluations of the cost function and its associated gradient within a gradient descent minimisation scheme. The use of advanced RT models in such schemes is therefore potentially computationally costly, further compounded by the fact that the gradient needs to be evaluated either numerically by finite differences [52], or using an adjoint of the RT model, which still takes significant time to evaluate. To overcome these limitations, we propose the use of Gaussian Process (GP) emulators [21,64,65]. An emulator provides a prediction of the RT model output with respect to the input parameters. To do this, the emulator is trained with a limited set of RT model input/output pairs to be able to produce this mapping. GPs are fast, cope well with non-linear RT models and produce an estimation of uncertainty of the model output prediction that can be included in the DA scheme [21] (although in this case the uncertainty is very small compared to the observational error, so it is ignored). Additionally, GP emulators can be used to provide an estimate of the emulated model gradient. The emulation of various leaf, canopy and atmospheric RT models (semi-discrete, PROSAILand 6S) is demonstrated in [21], showing speed-up by a factor of 40,000 or more. In this contribution, we have used the Python implementation of the emulators provided in [66].

3.6. FAPAR

After the inference of land state parameters and associated uncertainties, the energy balance function of the semi-discrete model can be used for computing FAPAR. Note that this provides the absorption under direct illumination and requires as inputs the estimated LAI, height of canopy and leaf diameter values. In addition, the spectral parameters, i.e., soil albedo (rg), leaf reflectance (rl) and leaf transmittance (tl) from the PROSPECT model and the background albedo from the Price soil model, have to be converted into the Photosynthetically Active Radiation (PAR) region. The sun zenith angle has been computed for 12:00 local times over the three sites using [67].

4. Results

Figures 3–5 show the evolution of LAI, leaf chlorophyll content, proportion of senescent material and FAPAR retrieved by EO-LDAS over the US-Ne1, US-Ne2 and US-Ne3 sites, respectively, between 2001 and 2008.

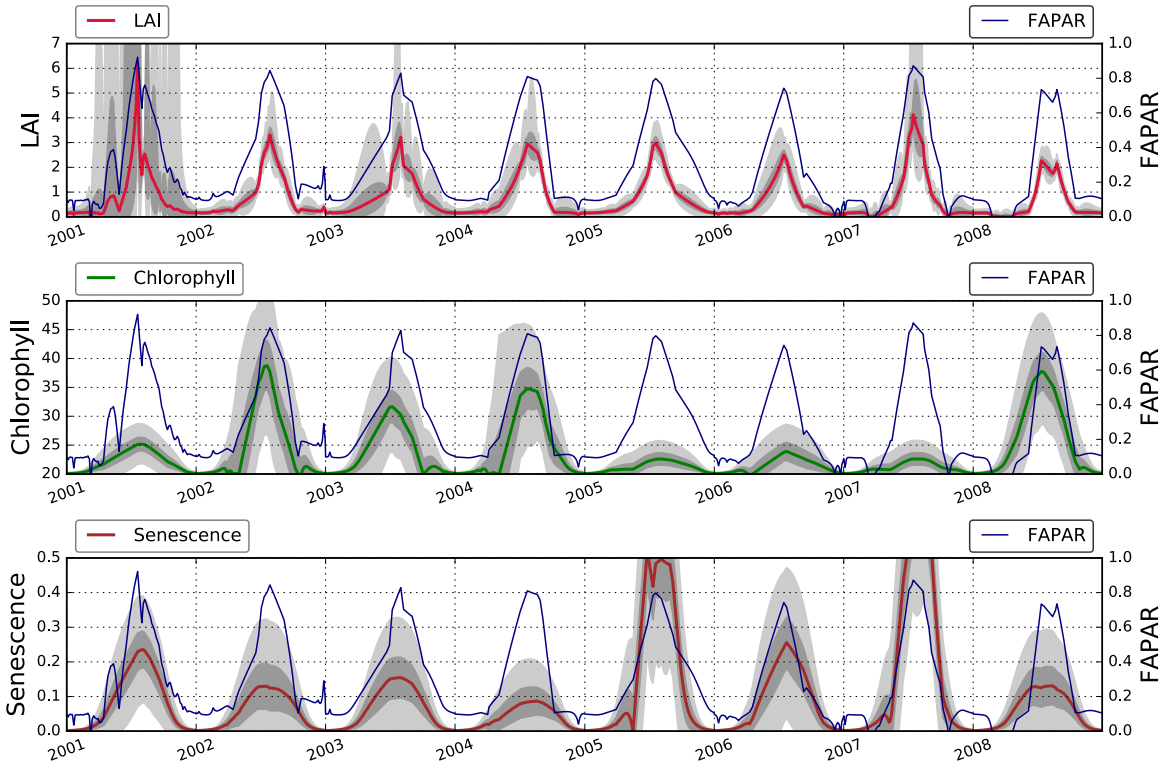


Figure 3. State parameters on US-Ne1 (2001–2008) retrieved from the Multi-angle Imaging SpectroRadiometer(MISR) data with Earth Observation Land Data Assimilation System (EO-LDAS). Dark and light shaded areas correspond to 75% and 95% credible intervals, respectively.

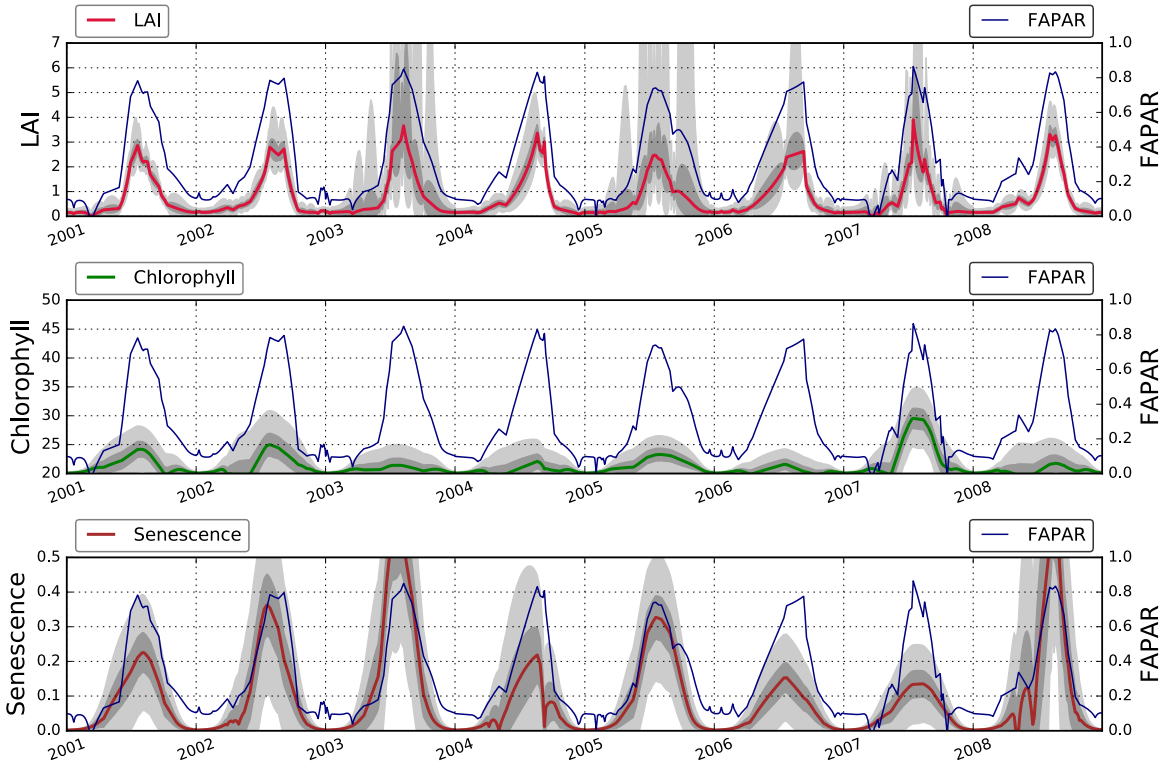


Figure 4. State parameters on US-Ne2 (2001–2008) retrieved from the MISR data with EO-LDAS. Dark and light shaded areas correspond to 75% and 95% credible intervals, respectively.

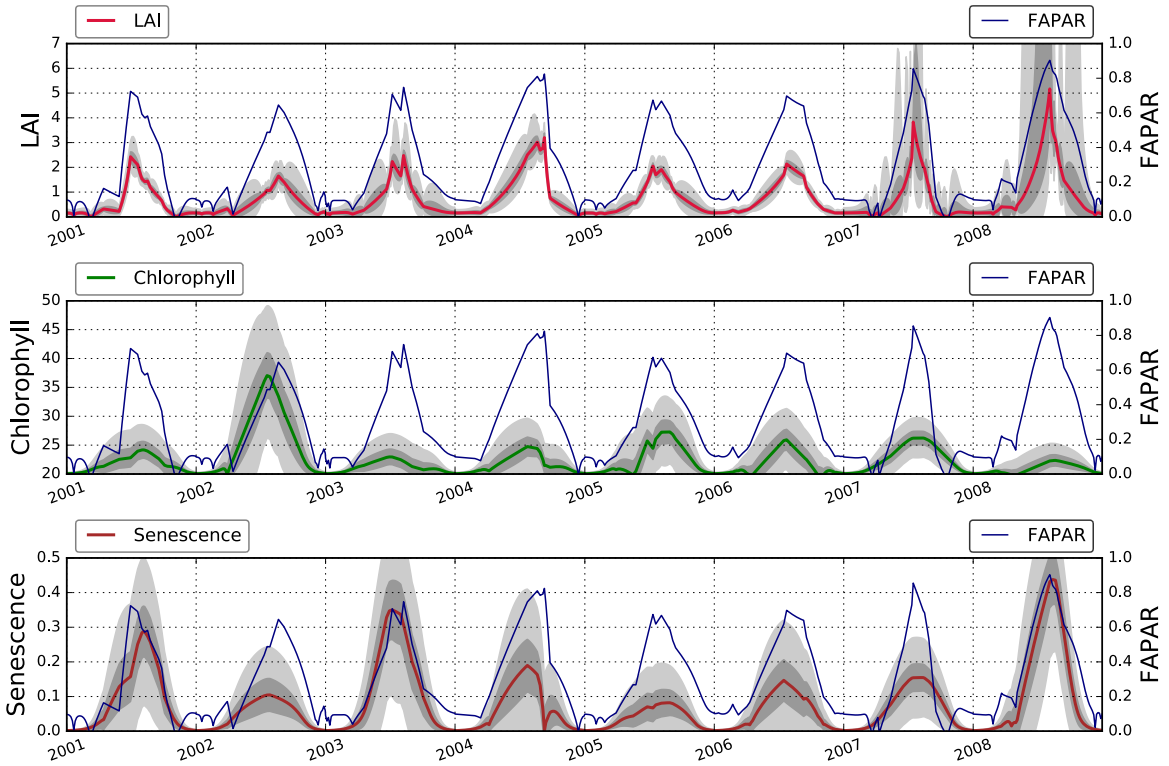


Figure 5. State parameters on US-Ne3 (2001–2008) retrieved from the MISR data with EO-LDAS. Dark and light shaded areas correspond to 75% and 95% credible intervals, respectively.

Figure 3 shows that the temporal trajectories of LAI, FAPAR, leaf chlorophyll content and the fraction of senescent leaves all broadly follow an annual pattern, with peaks in the summer months (around July-September). For LAI, the growing season is characterised by high uncertainties, a consequence of the saturation of reflectance for high LAI values and the small contribution of the prior term in this period (characterised by a large uncertainty). The temporal trajectories of the leaf pigments (chlorophyll and senescence) are characterised by large uncertainties, but the posterior mean shows a clear annual cycle, with chlorophyll leading senescence, as expected. This is a significant observation, as the temporal dynamics are not fixed by the prior term (the same functional form is used for all three parameters). The value of the proportion of senescent material does not go above 0.1, whereas at the end of the growing cycle one would expect there to be no senescent leaves. Large uncertainties in the parameters are due to complex interactions in how the parameters interact (with LAI and optical properties compensating each other). It is important to note that the dynamic model used in the `eoldas_ng` inversion results in temporal continuous inferences, even though no observations might be available on a particular date. The results from the other two fields (Figures 4 and 5) show a very similar behaviour, with both sites showing clear seasonalities for all retrieved parameters.

The parameters retrieved and shown in Figure 3 were then combined with the parameters in Table 2 and used to run the semi-discrete model and predict FAPAR. The MISR time series of EO-LDAS derived FAPAR are compared to ground-based measurements over US-Ne1, US-Ne2 and US-Ne3 sites (see Figure 6). The retrieved FAPAR tends to track the ground observations, with most of the ground observations being within the uncertainty bounds during the peak vegetation period. The inferences tend to overestimate the start and end of the growing season, and to slightly underestimate at peak LAI. The reason for the underestimate at the start of the growing season is that there is a paucity of observations at this time, and the retrieval is governed by the dynamic model interpolating between the observation-rich high LAI period and the prior-driven period with no vegetation at the beginning of the year. Towards the end of the growing season, the dynamic

model sometimes fails to track the fast changes in FAPAR, again due to poor observation availability. Uncertainties are much larger when observations are not available and dynamic model restores the data. This is especially noticeable when temporal gaps are larger than one month. For example at the beginning and end of year 2002 (Ne-1), beginning of year 2003 (Ne-1), end of year 2004 (Ne-3), etc. In Figure 7 we show the retrieved FAPAR, uncertainties as well as ground measurements for US-Ne1 for 2002 only. It is clear that there are no observations available between the beginning of May and mid-July, and so the regularisation results in an overestimation over that period. Note, however, that the paucity of observations results in a noticeable increase in uncertainty, which results in the ground measurements actually being within the 95% credible interval.

Figure 8 shows linear correlations between the retrieved FAPAR and the ground measurements for the three fields and all days for which ground measurements are available (top row) and also only for the dates where there are ground observations and coincident MISR overpasses (bottom row). It is clear from the top row in Figure 8 that the retrievals for low FAPAR are overestimated, and that this is caused by the interpolation provided by the dynamic model. When only points with observations are considered, the correlation increases, and the bias, slope and intercept all decrease, suggesting that the inversion works well where observations are available and the quality of the inferences drops as one moves away from the observations, with the dynamic model being too simple to track the changes in the rates of the process (as shown in Figure 7. In summary, we see that the correlation between retrieved and in situ FAPAR is high ($r^2 > 0.8$ for all dates, increasing to $r^2 > 0.85$ for retrievals coincidental with ground measurements) and the average $RMSE$ is 0.14 (only at days of satellite acquisitions). The slope of the retrieved FAPAR is consistent with an overestimation of FAPAR at the start and end of the growing season, and a slight underestimate in summer. When MISR observations are coincident with ground measurements, the slope becomes closer to the 1:1 line and the bias in the linear model tends to vanish.

In Figure 9, we show the results of comparing ground-measured and retrieved LAI. The comparisons show that there is a strong correlation between retrieved and in situ LAI, but an important underestimation. One has to recall that any retrieval LAI from space or ground-based data correspond to an effective value which depends on the RT used during the retrieval or protocol in the case of in situ [68]. In general, retrieved effective LAI is close to the in situ measurements when LAI is low ($\leq$2). After that, the retrieved LAI is lower than the in situ measurements. As the canopy becomes optically thicker, the sensitivity of the MISR observations decreases. This is expected as the value is close to the theoretical limit of retrieval of LAI as described in [69]. This is accompanied by an increase in uncertainty in the retrieved LAI estimate (see Figure 3). The summary statistics show that there is ample room for improvement: the correlations are 0.86, 0.80 and 0.68 (for US-Ne1, US-Ne2 and US-Ne3, respectively), slopes are 1.79, 1.86 and 1.55 (same order), and intercepts are -0.26, -0.36 and 0.28. $RMSE$ values are 1.92, 2.03 and 2.16.

Although we have stated that the impact of the prior term should only be to tighten the retrieval in the period with no vegetation and few observations, this has not been demonstrated. In Figure 10, we show the comparison of retrieved FAPAR for the days of coincident satellite overpasses, where the prior is just set to a constant mean and uninformative (i.e., large) variance. The results are virtually the same as those shown in Figure 8, which demonstrates that the effect of the prior is minimal in the retrievals, as expected.

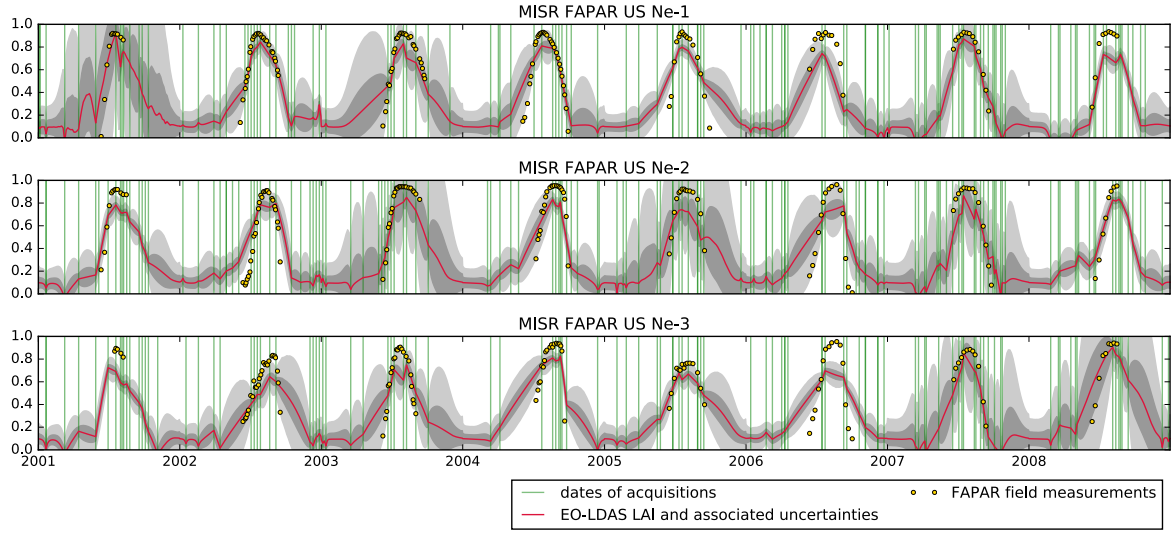


Figure 6. Comparison of FAPAR MISR time series between field measurements (yellow dots) and EO-LDAS predictions (red lines) over US-Ne1, US-Ne2 and US-Ne3 sites. Green lines indicate dates of MISR surface available data. Light and dark shaded areas correspond to uncertainties 95% and 75% credible interval, respectively.

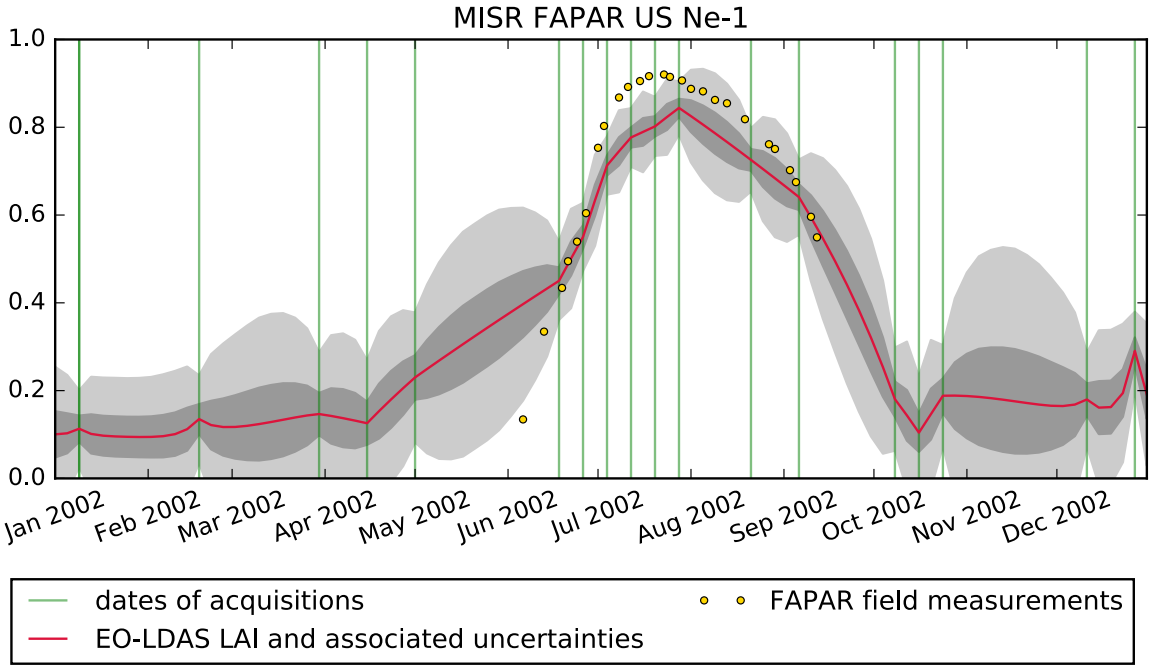


Figure 7. Comparison of FAPAR MISR time series between field measurements (yellow dots) and EO-LDAS predictions (red lines) over the US-Ne1 site for year 2002. Green lines indicate dates of MISR surface available data. Light and dark shaded areas correspond to uncertainties 95% and 75% credible interval, respectively.

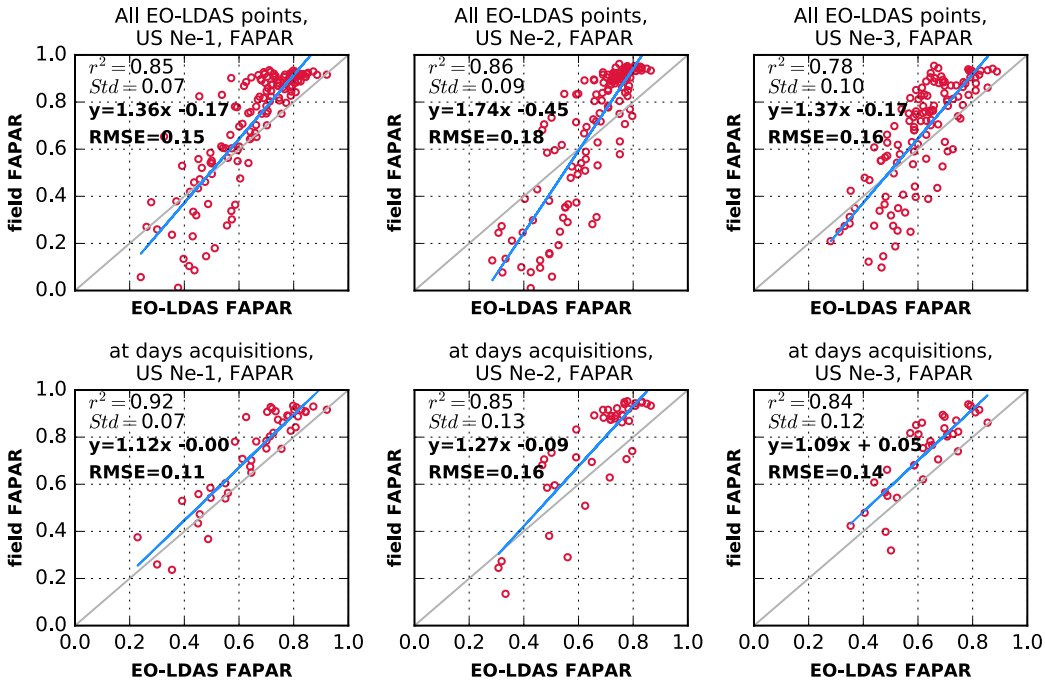


Figure 8. Scatter-plots of all available data pairs of in situ measured and MISR EO-LDAS derived FAPAR and field FAPAR for the investigated time period from 2001–2008. (Top panels) all available EO-LDAS data points; (bottom panels) only at days of satellite acquisitions. Blue lines were derived from least square regression, while 1:1 lines are depicted in grey.

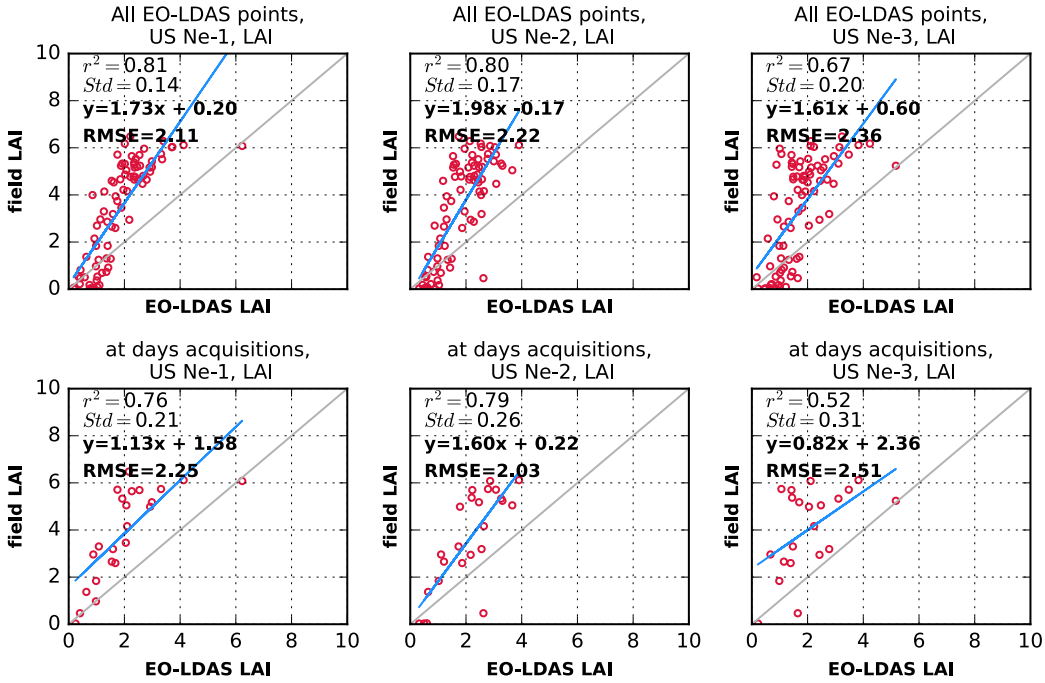


Figure 9. Scatter-plots of ground-based estimated and MISR EO-LDAS derived effective LAI. (Top panels) all available EO-LDAS data points; (bottom panels) only at days of satellite acquisitions. Blue lines were derived from least square regression, while 1:1 lines are depicted in grey.

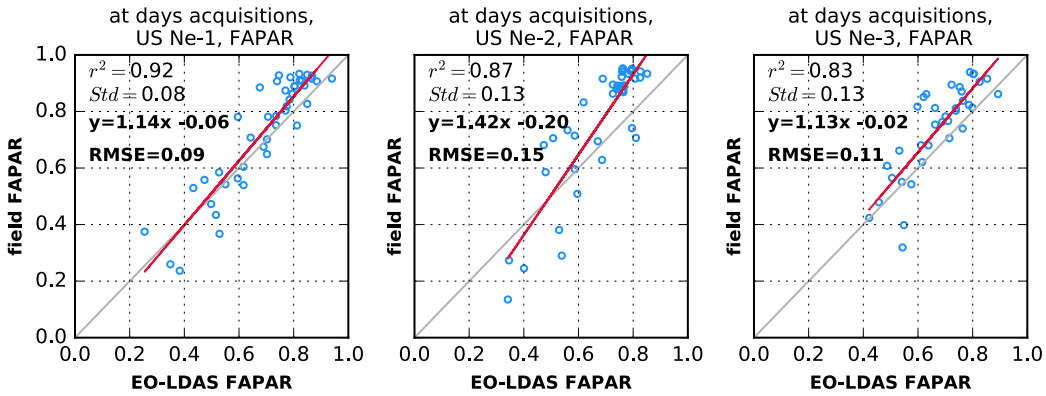


Figure 10. Scatter-plots of data pairs of in situ measured and MISR EO-LDAS derived FAPAR without dynamical prior for the investigated time period from 2001 to 2008. Only data from days of satellite acquisitions are shown. Red lines were derived from least square regression, while 1:1 lines are depicted in grey.

4.1. The JRC-TIP Results

The results of FAPAR from the JRC-TIP product is shown in Figure 11. We note that this product is only able to provide an estimate of FAPAR when satellite observations are available. Furthermore, we note that for a number of years, the JRC-TIP FAPAR results in a more peaky growing season, which often peaks earlier than the ground measurements.

Figure 12 shows the results from the JRC-TIP product assuming green leaves (in effect, a tighter prior on the leaf single scattering albedo). The results show a clear underestimate of FAPAR compared to the in situ measurements.

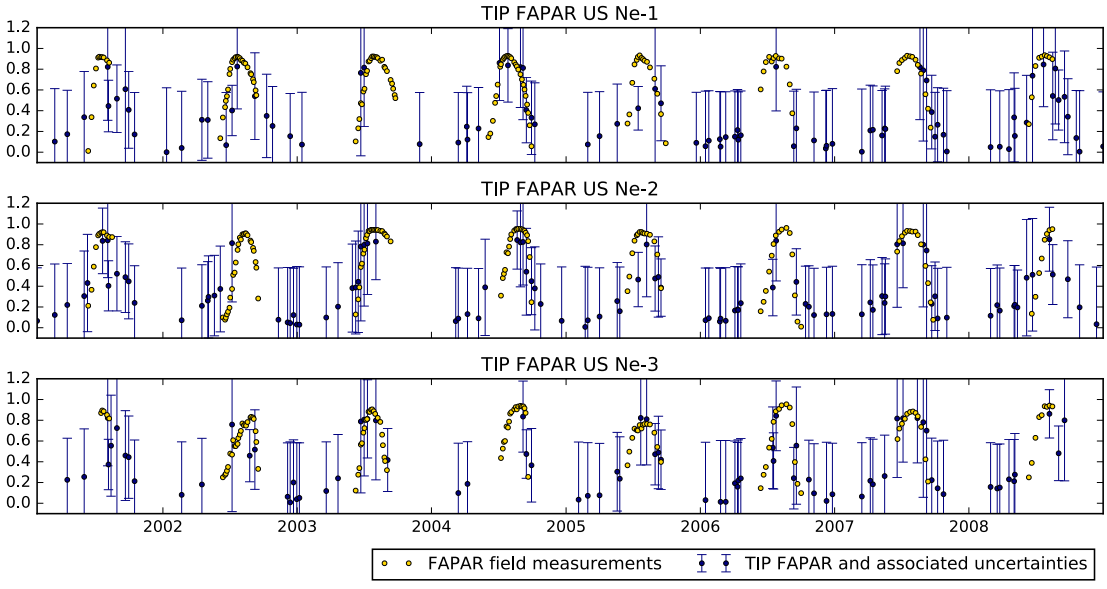


Figure 11. Time series of Joint Research Centre Two-stream Inversion Package (JRC-TIP) FAPAR from 2001–2008. Error bars correspond to 95% credible intervals.

We have also carried out a comparison between ground LAI and retrieved LAI, but we know that the retrieved LAI is effective [34,70]. The RT model used in this retrieval has indeed been developed for climate modelling assimilation of space land products; therefore it is based on a two-stream theory and retrieval values from surface broadband albedo. In [34] is recalled the need for a correction of a structure

factor at the pixel resolution associated with the heterogeneous nature of the canopy volume. In general, the correlation coefficient is low (0.21–0.64), and there is a large bias (>1.3), as well as a scatter of around 2. The large bias is due to the JRC-TIP saturating at around three, as explained in [17].

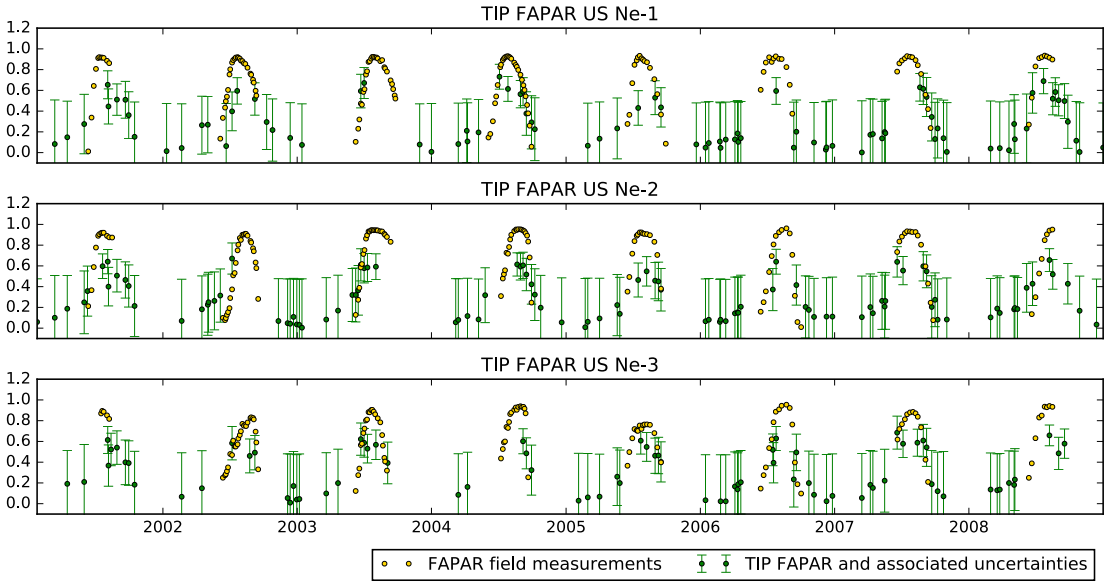


Figure 12. Time series of JRC-TIP Green FAPAR from 2001–2008. Error bars correspond to 95% credible intervals.

4.2. The JRC MERIS Product

Figure 13 shows a comparison of the JRC MERIS FAPAR product and the in situ measurements. In general, the MERIS product provides an accurate description of the FAPAR annual trajectory, although with some underestimation of FAPAR at the peak of the growing season, as well as some overestimates of the very low FAPAR values at the beginning and end of the growing season.

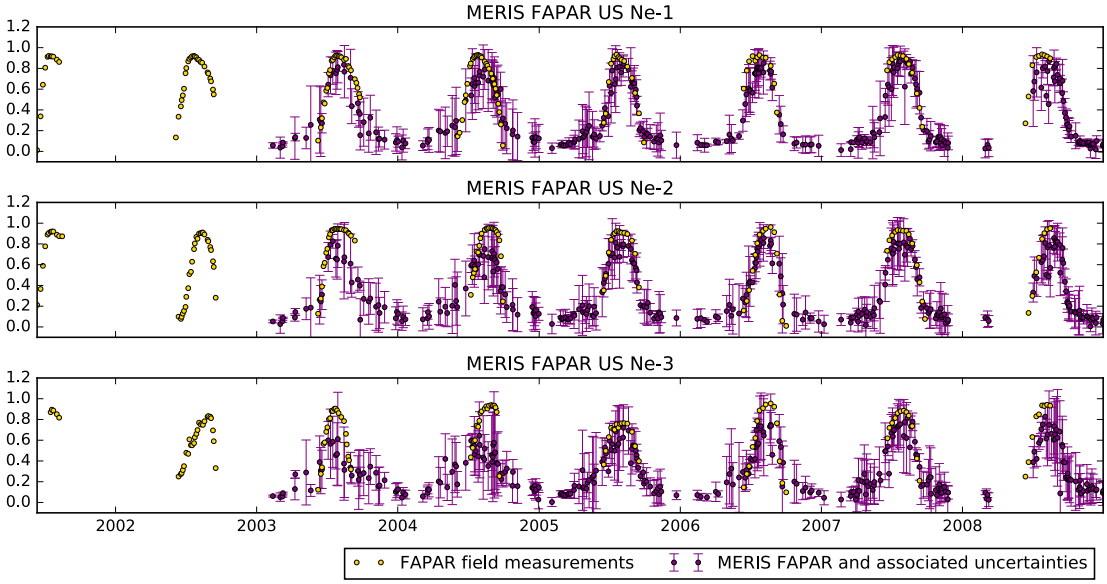


Figure 13. Time series of JRC MERIS FAPAR from 2002–2008. Error bars correspond to 95% credible intervals.

4.3. The MODIS FAPAR/LAI Product

Figure 14 shows FAPAR from the standard MODIS product MCD15A2H Collection 6. The results are generally good, the fine temporal sampling of using the two MODIS sensors resulting in a good coverage of the annual period. The estimates, however, tend to undershoot the peak FAPAR value consistently, and to overestimate the leading and trailing edges of the growing season.

The MODIS LAI product has a similar behaviour to the EO-LDAS LAI retrievals: with reasonably high correlations 0.84, 0.71 and 0.60 for US-Ne1, US-Ne2 and US-Ne3, respectively, but, again, resulting in an underestimate, particularly when the LAI is high.

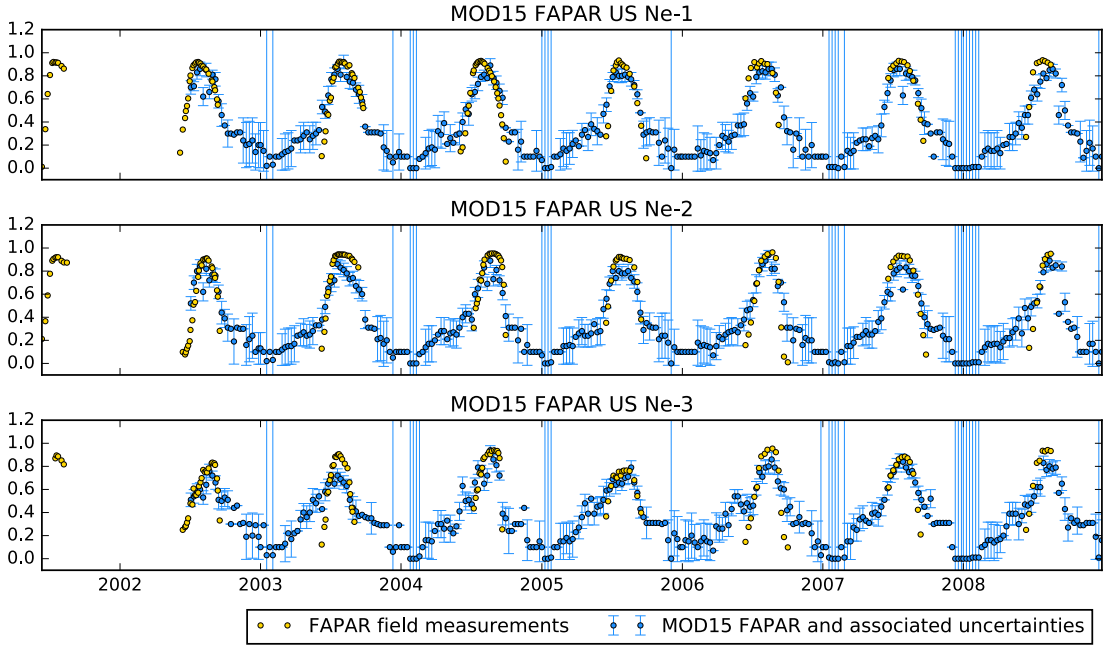


Figure 14. Time series (2002–2008) of MODIS MCD15A2H Collection 6. Error bars correspond to 95% credible intervals.

5. Discussion

Comparisons of FAPAR retrievals for all products with ground-based data are shown in Figure 15 and Tables 3–5. The EO-LDAS results have the highest linear correlation among all the compared products, as well as the lowest root mean square error. There is a positive slope, in line with the comments made in the last section about underestimates of FAPAR during the start and end of the vegetation period. If only dates where MISR observations are available are taken into account (remember that of all these products, EO-LDAS is the only one that will produce estimates for the entire time series), then the correlation increases, the slope approaches unity, and the bias disappears. These are remarkable results since MISR has far fewer observations per year (around 18) than MODIS (around 200) or MERIS (around 50), and yet the retrieved EO-LDAS MISR FAPAR performance is better than the other products (Table 6). The small number of observations and the very different temporal sampling of the MISR data affect the values of the regularisation parameter γ retrieved by cross validation. In years with sparse observations, this can affect the impact of the dynamic model on the retrieved parameter trajectories. For example for US Ne-1, in 2001 (Figure 6), the regularisation parameter was found to be 3.6, in contrast to regularisation values around two orders of magnitude higher. This means that the character of the solution is likely to be less smooth and have higher associated uncertainties [48].

The poorest results in this comparison are with the JRC-TIP with green leaves, followed closely by the JRC-TIP with polychromatic leaves. The MERIS and MODIS products are comparable, with MERIS

showing a slight negative bias. In general, all the tested products struggle to reach the highest in situ measured FAPAR, with a significant underestimate of FAPAR when in situ FAPAR is over 0.8. For EO-LDAS, this limitation can be traced back to the inability of the system to reach high values of LAI, a problem which is also expected with the JRC-TIP, where LAI is basically unconstrained whereas the optical properties of the leaves and soil are heavily constrained by priors. The other common striking feature of the comparisons is how the retrievals result in an overestimation of FAPAR when the in situ FAPAR is low (less than 0.3). This is a common feature of all products, but it is perhaps more exaggerated with the MODIS observations, which produce an overestimation of FAPAR where in situ values are below 0.6. In the case of the EO-LDAS retrievals, the effect is quite strong when all the in situ observations are considered, but the effect is attenuated when only the dates that have coincident observations are considered. By referring back to Figure 6, we explain this behaviour by pointing out that the low FAPAR values are usually in periods of very sparse MISR data (this is common to all years/sites), so that the effect is that of the system interpolating between two widely separated observations.

Table 3. Summary statistics derived from the comparison of FAPAR field measurements (2001–2008) collected on the US Ne-1 site with all corresponding EO-LDAS derived values (EO-LDAS all), EO-LDAS derived values at days of satellite acquisitions (EO-LDAS Obs.) and values obtained from JRC-TIP, JRC-TIP green, JRC MERIS and MODIS MCD15A2H products.

Stat. Param	EO-LDAS All	EO-LDAS Obs.	TIP	TIP Gr.	MERIS	MCD15
r^2	0.85	0.92	0.41	0.45	0.83	0.80
σ	0.07	0.07	0.13	0.18	0.08	0.09
slope	1.36	1.12	0.38	0.58	1.06	1.31
intercept	−0.17	−0.00	0.44	0.38	0.06	−0.20
RMSE	0.15	0.11	0.26	0.28	0.16	0.14

Table 4. Summary statistics derived from the comparison of FAPAR field measurements (2001–2008) collected on the US Ne-2 site with all corresponding EO-LDAS derived values (EO-LDAS all), EO-LDAS derived values at days of satellite acquisitions (EO-LDAS Obs.) and values obtained from JRC-TIP, JRC-TIP green, JRC MERIS and MODIS MCD15A2H products.

Stat. Param	EO-LDAS All	EO-LDAS Obs.	TIP	TIP Gr.	MERIS	MCD15
r^2	0.86	0.85	0.59	0.64	0.83	0.86
σ	0.09	0.13	0.17	0.27	0.09	0.09
slope	1.74	1.27	0.86	1.60	1.25	1.64
intercept	−0.45	−0.09	0.14	−0.13	−0.03	−0.42
RMSE	0.18	0.16	0.24	0.29	0.19	0.16

Table 5. Summary statistics derived from the comparison of FAPAR field measurements (2001–2008) collected on the US Ne-3 site with all corresponding EO-LDAS derived values (EO-LDAS all), EO-LDAS derived values at days of satellite acquisitions (EO-LDAS Obs.), and values obtained from JRC-TIP, JRC-TIP green, JRC MERIS and MODIS MCD15A2H products.

Stat. Param	EO-LDAS All	EO-LDAS Obs.	TIP	TIP Gr.	MERIS	MCD15
r^2	0.78	0.84	0.28	0.21	0.59	0.80
σ	0.10	0.12	0.17	0.31	0.14	0.10
slope	1.37	1.09	0.30	0.42	0.90	1.41
intercept	−0.17	0.05	0.48	0.46	0.23	−0.24
RMSE	0.16	0.14	0.23	0.26	0.26	0.14

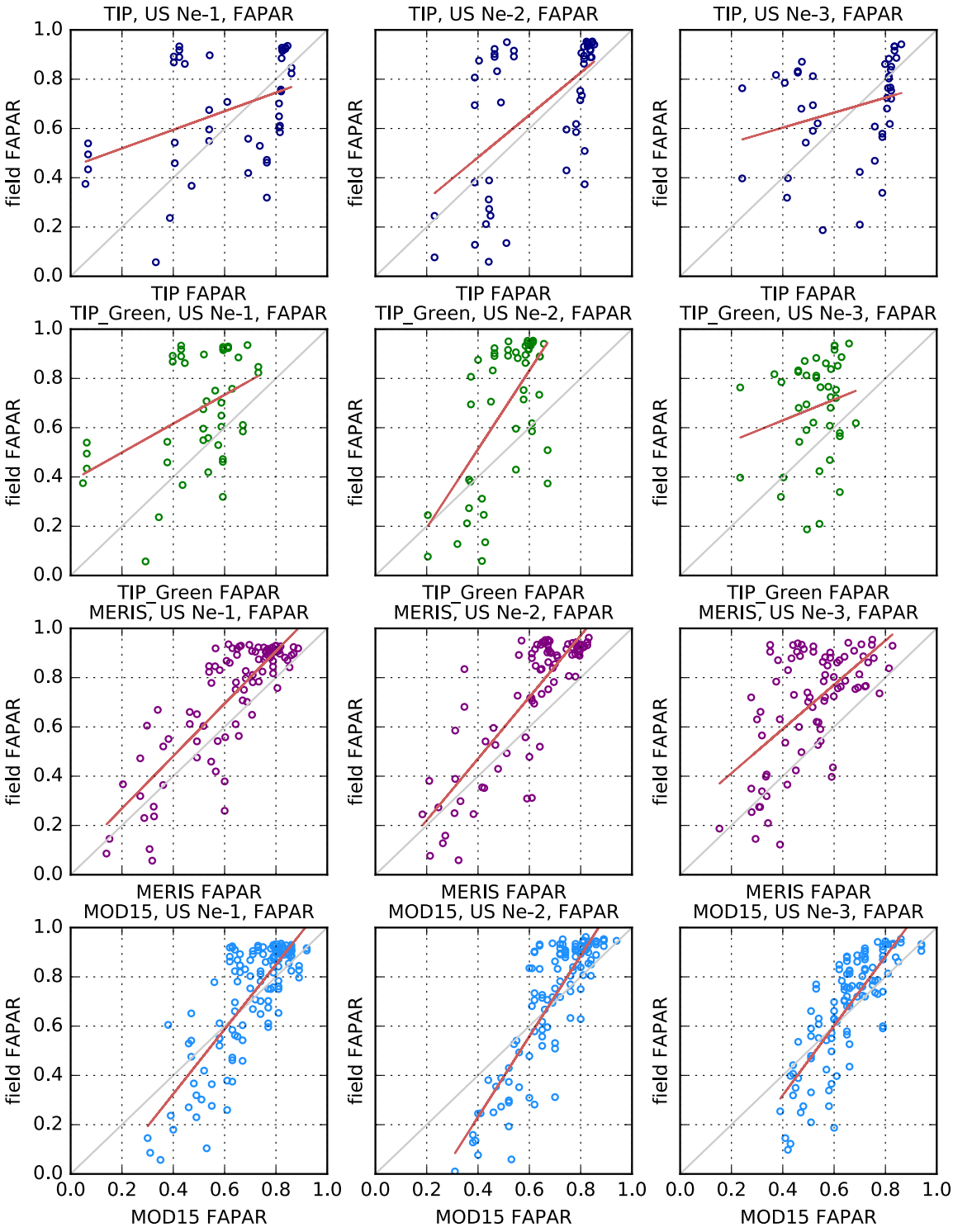


Figure 15. Relationship between field measurements and JRC-TIP FAPAR for standard leaf (top panels), JRC-TIP Green FAPAR (second row of panels), 3 × 3 MERIS FR (third row) and MODIS MCD15A2H (bottom panels), the period between 2001 and 2008. Red lines were derived from least square regression, while 1:1 lines are depicted in grey.

Table 6. Average number of observations over all available years.

Field	MISR	MODIS	MERIS
US-Ne1	16 ± 4	207 ± 37	50 ± 9
US-Ne2	18 ± 3	203 ± 34	50 ± 9
US-Ne3	14 ± 3	122 ± 18	51 ± 10
Years	2001–2008	2001–2008	2003–2008

An important feature of the EO-LDAS approach is the use of a dynamic model for the retrievals. As the uncertainty in the dynamic model (γ in Equation (14)) is constant, the trajectory of the state will tend to change not as rapidly as the observations on the ground, and this can lead to solutions that are too smooth in periods where the dynamics of the state are very fast. In our EO-LDAS results, the regularisation results in an increase in uncertainty, that although the MAP estimate overshoots the ground measurements, most of the measurements are within the uncertainty boundaries (except for a couple of cases of the earliest and latest stage of the growing season). Having more observations over those highly dynamic periods would alleviate this problem. In observation-sparse periods (e.g., 2001 for LAI in US-Ne1, Figure 3), observations produce a discontinuity in parameter trajectories, which is noticeable in the MAP solution, but small in terms of the total uncertainty. This artefact could be caused by locally defective convergence of the optimisation. In the case of abrupt changes, different techniques would need to be used, such as those presented in [71].

A further observation is that in the case at hand, we are using surface directional reflectance. The non-vegetative period is characterised by a rough soil surface with crop residue. The simple soil model is unable to properly model this period, for example being unable to, model BRDF effects present in the observations. The only freedom allowed by the model in this period is to modify canopy parameters to account for this, which is what causes the departure from zero of LAI in Figure 3 and the oscillations of C_{ab} in the non-vegetative period. We can demonstrate this by running the following experiment: by fixing the leaf optical properties to sensible values and setting the LAI to be the (temporally interpolated) in situ LAI, we can then predict the observations. We can see the predictions and observations in the red and near-infra-red bands for 2002 and the US-Ne1 site in Figure 16. It is clear that the model fits the growing season reasonably well, but struggles with the bare soil period, indicating that here the RT model has problems replicating the data. Extending the RT model to have proper treatment of BRDF (for example, adding a Walthall [72] or Hapke [73,74] soil model) would alleviate this problem. Another approach would be to consider in the uncertainty budget in Equation (7) the model uncertainty. Under the assumption that the model uncertainty is (i) independent of the measurement uncertainty and (ii) normally distributed, we have that the observational constraint is given by:

$$J_{obs}(\vec{x}) = -\frac{1}{2}(\vec{R} - H(\vec{x}))^{\top} \left[\mathbf{C}_o^{-1} + \mathbf{C}_m^{-1}\right] (\vec{R} - H(\vec{x})), \tag{15}$$

a case that is readily implemented in `eoldas_ng`, but where the model uncertainty encoded in $\mathbf{C}_m$ might not be straightforward to assess. The effect of the prior introduced in this study goes somewhat to soften this problem, but leads to a dampening the contribution of $J_{obs}(\vec{x})$ to the minimisation.

It is instructive to compare retrievals from the JRC-TIP product and EO-LDAS. Both approaches share the philosophy of calculating FAPAR by running a RT model with a parametrisation of the land surface state derived from inverting observations. In Figure 17, it is clear that there is a strong correlation between the LAI value retrieved from both approaches (taking into account that the retrieved effective LAI), but it is also apparent that the Single Scattering Albedo (SSA) estimates from both approaches is anti-correlated. In Figure 18, we can see that the JRC-TIP estimation for the soil background albedo in the visible changes throughout the year, whereas in the EO-LDAS case we have just fixed a particular soil spectral model. LAI (middle panel in Figure 18) shows some agreement,

with EO-LDAS providing a more realistic trajectory due to the interpolation. The uncertainties from both products have similar trends with low values when LAI is low and high values in summer. The temporal evolution of the SSA in interesting: the JRC-TIP solution barely shifts from the prior, whereas the EO-LDAS version has a clear seasonality as a consequence of leaf chlorophyll content and fraction of senescent leaves changing throughout the growing period. In this case, the use of spectrally-resolved observations results in a richer interpretation of the EO data, rather than working from an spectrally integrated broad band, as is the case with the inputs to the JRC-TIP product.

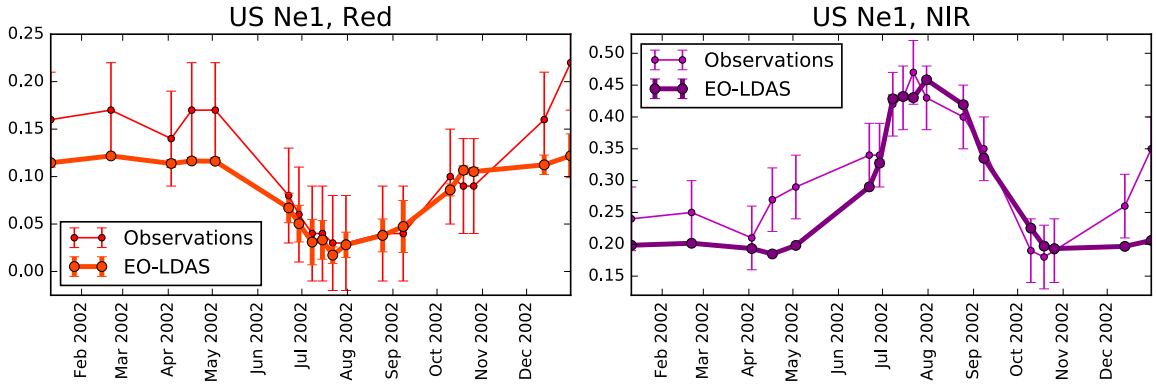


Figure 16. Comparison of reflectance from MISR observations (An camera, US Ne1 site, 2002) and reflectance obtained from a forward run of the model with LAI fixed to the ground reference data.

The relatively good performance of the MERIS FAPAR algorithm is probably caused by its simplicity. Rather than targeting inference of a full set of land surface parameters, an equation is used to map from top of atmosphere reflectance to FAPAR, making strong assumptions on leaf single scattering albedo and soil background. Going through the EO-LDAS or JRC-TIP approaches needs the inference of a larger set of parameters, which necessarily results in larger uncertainties. However, a limitation of the MERIS algorithm is that it only produces an estimation where observations are present, and would need to be extended to cover different sensors simultaneously to produce a consistent FAPAR value. The method also does not estimate the underlying land surface parameters (e.g., LAI) which, if required for a particular application, would need to be obtained from a different product, potentially introducing inconsistencies e.g., due to different choices in the underlying RT model used for inversion.

The results from comparing the products that provide LAI estimates with ground LAI measurements show a common trend: all products underestimate LAI substantially. For the JRC-TIP, we note that this is expected as the LAI is effective [34], so this is not an entirely fair comparison. For the EO-LDAS MISR and MODIS products, we see an important underestimate of LAI when the ground value is high and effective LAI reach the saturation limits.. The different spatial scales of the satellite and ground measurements can result in very different LAI values [70,75,76]. It is interesting to note that the uncertainties in the retrieval of LAI with EO-LDAS MISR show a strong asymmetry (e.g., the uncertainty region above the posterior mean is much larger than the uncertainty region below the posterior mean) for high LAI (see Figures 3 and 4), which is a statement of the limited sensitivity of the observations to high LAI.

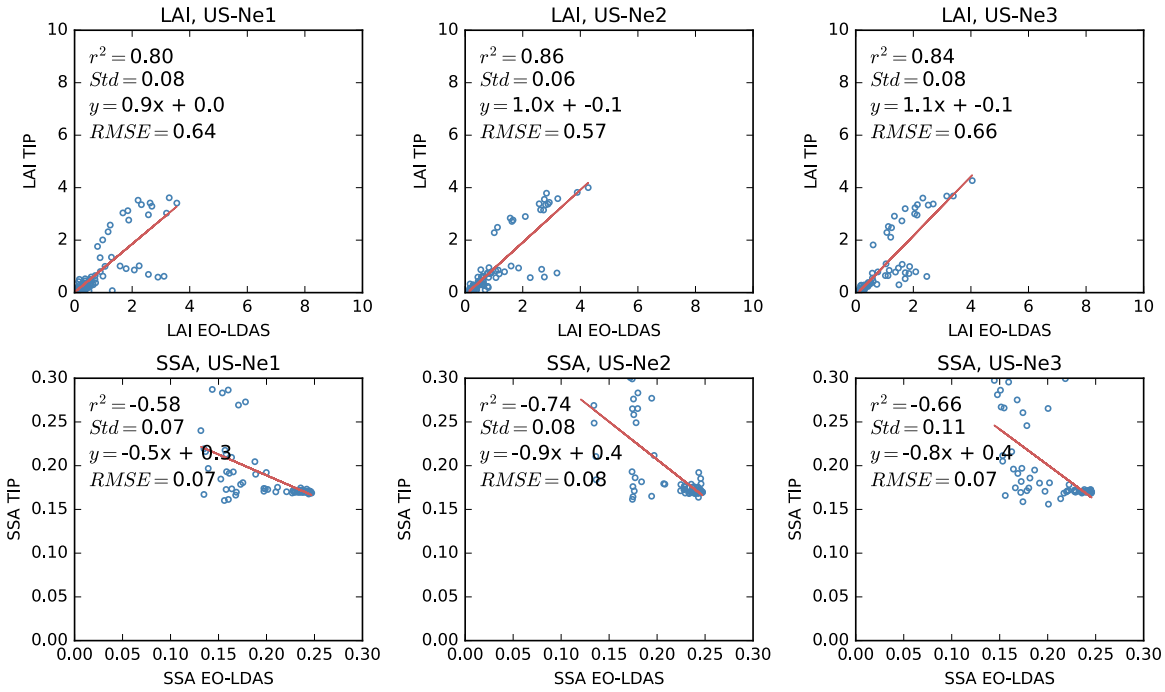


Figure 17. MISR: Comparison between EO-LDAS and JRC-TIP Leaf Area Index (LAI; upper panels) and leaf Single Scattering Albedo (SSA; lower panels) in the case of polychrome leaf assumption for the JRC-TIP. Red lines were derived from least square regression.

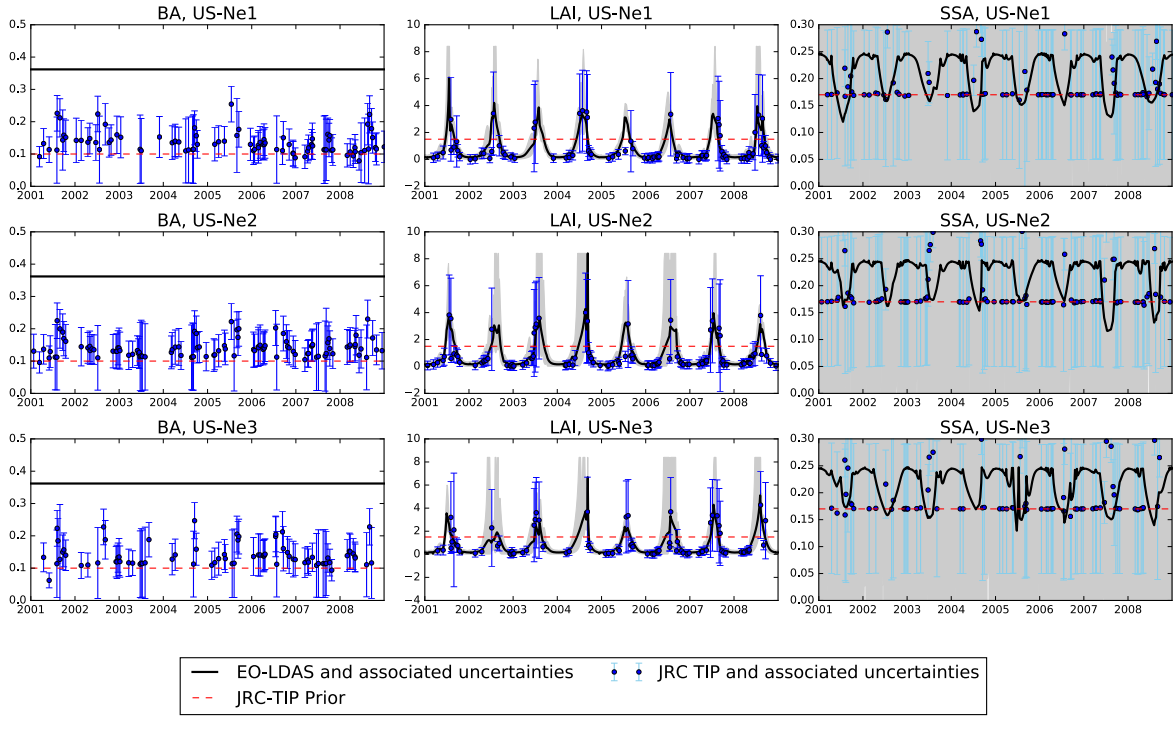


Figure 18. Times series of state parameters derived from MISR observations: Background Albedo (BA; left column), Leaf Area Index (LAI; central column) and Single Scattering Albedo (SSA; right column) for EO-LDAS retrieval (black line) and JRC-TIP (blue line).

6. Concluding Remarks

The main focus of this study is the estimation of daily FAPAR values by using EO-LDAS on MISR data. The evaluation of results against JRC-TIP, MERIS FR, MODIS MCD15 and in situ measured 'green' FAPAR was carried out for single pixels covering three agricultural FLUXNET sites over eight years. We have compared results from a number of different approaches and sensors: the EO-LDAS and JRC-TIP approaches both use MISR data (surface directional reflectance for the former and broadbands bi-hemispherical reflectance for the latter). These two methods rely on the inversion of an RT model, auxiliated by prior parameter distributions, and in the case of EO-LDAS, a dynamic regularisation model that allows the inference even at times where no satellite observations are present. The JRC MERIS FAPAR product uses MERIS top of the atmosphere reflectance and a polynomial that maps these measurements to FAPAR, and the MODIS product uses a Lookup Table (LUT) that maps red and near-infrared surface reflectance to FAPAR. It is important to point out that the difference in temporal sampling between the three instruments: MISR has a much lower number of observations than MERIS, and MERIS has a much lower number of observations than MODIS.

We have compared the different products with in situ ground measurements, and these indicate that FAPAR is retrieved with reasonable accuracy for three products: EO-LDAS ($r^2 > 0.8$, $RMSE < 0.18$ in units of FAPAR), MERIS ($r^2 > 0.6$, $RMSE < 0.2$) and MODIS ($r^2 > 0.8$, $RMSE < 0.16$). The JRC-TIP results show a poor performance, with r^2 values between 0.3 and 0.6 and $RMSE$ in excess of 0.26. If only dates with MISR overpasses are considered for EO-LDAS, the estimates from EO-LDAS improve to an r^2 between 0.85 and 0.94, an $RMSE < 0.14$ and with a bias that is between 2 and 7%.

All products have problems tracking the high FAPAR peak of the growing season, resulting in all products underestimating FAPAR at this point by 10–20%. Additionally, all products tend to overestimate FAPAR for the flanks of the growing season.

Three products (EO-LDAS, JRC-TIP and MODIS) also retrieve LAI (or effective LAI). The results all indicate that high LAI is underestimated. We propose that two processes are having an effect here: first, as the canopy becomes optically thicker, underestimation of LAI is expected [17], and secondly, the comparison of coarse resolution observations with point measurements introduces the effects of sub-pixel landscape heterogeneity [75,76]. In the literature [77,78], the use of empirical methods that have been trained with ground observations of the same area limits the generality of the methods for global applications. Additionally, no simultaneous inferences on FAPAR are presented in either of these two references. For the three coarse-resolution products that we considered in this study, comparisons with in situ LAI result in ($r^2 > 0.7$, $RMSE < 2.2$) for EO-LDAS, ($r^2 > 0.6, RMSE < 2.4$) for MODIS and r^2 between 0.4 and 0.6, $RMSE < 2.3$ for the JRC-TIP (but again, note that LAI for the JRC-TIP is effective). The RMSE between all products and the ground observations is of the order of two units of LAI. In this study, we have used point measurements of LAI as a comparison. The recommended best practice is to use these measurements to provide a spatially explicit map, i.e., the minimum size of a validation site has to be compatible with resolution of satellite data [79,80].

Additionally, the MISR EO-LDAS approach also estimates the daily evolution of leaf chlorophyll content and fraction of senescent leaves. Both of these parameters show a credible temporal trajectory (although they have not been compared with any in situ measurements), with a clear arch for both when LAI is high, but showing leaf chlorophyll dropping off earlier than senescence and senescence growing later than leaf chlorophyll. This is a very encouraging result, as no timing information on these two parameters as provided in the priors pdf used, and only observations from four spectral bands in the visible-NIR region were available.

The MISR EO-LDAS approach is characterised by large uncertainties in both land surface parameters (LAI, C_{ab}, C_{brown}), as well as fluxes (FAPAR). This is inevitable due to the poor temporal sampling. We note that high uncertainty in LAI during periods of high LAI is to be expected, and we note that this uncertainty (see Figure 3) is asymmetric: high above the MAP estimate and low below, encoding the inability of the RT model of interpreting the reflectance as a clear high value of LAI, but quite certain that it is larger than e.g., two. This is in marked contrast to the uncertainties

in LAI for the JRC-TIP product (see Figure 18). The advantage of EO-LDAS is obtained by using the quasi-linearising transformations in Table 2.

In this paper, we deliberately only compare EO-derived estimates of FAPAR from approaches that rely on the inversion of RT models (and hence, ought to show some level of consistency and physical accuracy). However, we also note that there is a long history of exploiting the relationship between empirical vegetation indices (such as NDVI) and FAPAR. Recent work by [22] has used ground spectral measurements, as well as MODIS 250-m NDVI data to revisit these relationships for the sites that are considered in this study. In [22], different regressions relating NDVI to FAPAR are shown to produce a very large spread of possible values of FAPAR. The authors also derive a set of equations that can be applied to the site, but note that they need to be split by crop type (maize and soya) and by development (vegetative and reproductive) stage. The resulting correlations are similar to those reported in the present study, whereas other non-specific formulations of the mapping between NDVI and FAPAR show a large scatter, indicating that significant effort is required for local calibration to provide useful results. In [22], additional understanding of the crop type and phenology are required for optimal performance. Using the EO-LDAS approach, we have only used very generic priors detailing vegetation structure and leaf optical properties, although of course, more informative priors could be used, if they were available. As part of the process, estimates of widely-used land surface parameters have also been retrieved, along with well-quantified uncertainty. This is a particular feature of the EO-LDAS approach, which provides an important benefit over other methods. In particular, LAI has been retrieved with an error that is in the range of other typical global LAI products for this site, and more generally [81]. The EO-LDAS approach also exploits temporal regularisation to provide a consistently gap-filled estimate of parameters even when no observations are available. The use of radiative transfer models within this regularisation framework allows for simple inclusion of observations from other sensors, as these are interpreted and assimilated in terms of common descriptions of the land surface. These are important advantages of the EO-LDAS approach presented here, and they have been proven to be useful: having a much poorer temporal sampling regime, results from the EO-LDAS MISR approach are in line with the JRC MERIS and MODIS FAPAR products. LAI retrievals are also in line with the other products, indicating that the approach based on interpretation of the observations using physical models succeeds in providing consistent estimates of FAPAR and other land surface parameters.

An additional important consideration is that the EO-LDAS retrieved parameters are self-consistent, i.e., the same physical (RT) model assumptions are made to retrieve e.g., both LAI and FAPAR. This is important if such parameter estimates are then used to drive models [8]. Inconsistencies across suites of parameters can result in hard-to-trace model deviations and uncertainties in model predictions. Parameter consistency is hard, if not impossible, to ensure when using locally-calibrated empirical relationships where different parts of the state vector (e.g., LAI, leaf chlorophyll concentration or FAPAR) are derived separately. Further, the incorporation of priors makes explicit the assumption that both the chosen physical model and sets of priors are adequate for the task at hand. The use of Bayesian approaches with physical models necessitates clear statements about observational uncertainty [82,83]. These ought to be provided as the standard with EO-derived surface reflectance products, but rarely are in practice due to the limitations of retrieval processes. Finally, although in this study we have ignored the uncertainty associated with using an RT model, this can be readily introduced in the EO-LDAS formalism, if information on the properties of this error were available. Unacknowledged model error can result in biases in the solution.

We demonstrate that at least for the particular sites shown here, EO-LDAS is able to estimate absorbed fluxes, other biophysical parameters and associated uncertainties relying on multi-angular surface reflectance observations. The EO-LDAS scheme also allows for a simple combination of other available observations, thus opening the door to multi-sensor estimates of fluxes and/or biophysical parameters. The estimation of uncertainties, as well as the retrieval of a complete set of ground

biophysical parameters is also an important tool in providing data that allows us to learn more about the land surface.

Acknowledgments: We acknowledge financial support of the project GIONET, funded by the European Commission, Marie Curie Programme Initial Training Network, Grant Agreement Number PITN-GA-2010-264509; the EU H20:20 BACI project (Project No. 640176); the EU H20:20 MULTIPLY project (Project No. 687320); and the BMWIproject EO-LDAS-App (Grant No. 50EE1307). The authors are grateful for European Space Agency support under Contract 4000112388/14/I-NB for funding the development of the underlying software that made this work possible. We also wish to gratefully acknowledge Anatoly Gitelson and Andrew Suyker for the ground-based FAPAR data, as well as Marco Zuehlke and Brockmann's Consult for providing the MERIS FR results. The JRCauthors thank Linda Hunt and Michel Verstraete for their technical support on MISR HR processing chain.

Author Contributions: Maxim Chernetskiy designed the experiments, performed the experiments and wrote the paper. Jose Gómez-Dans contributed to writing the paper, experiment design and development of `eoldas_ng`. Nadine Gobron designed the experiments and contributed to writing the paper. Olivier Morgan prepared the data and contributed to writing the paper. Philip Lewis contributed to the development of `eoldas_ng`. Sina Truckenbrodt contributed to writing the paper. Christiane Schmullius supervised the related PhD work as part of the GIONET project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Global Terrestrial Observing System (GTOS). *Terrestrial Essential Climate Variables for Climate Change Assessment, Mitigation and Adaptation*; Technical Report; Food and Agriculture Organization (FAO): Rome, Italy, 2008.
2. Gower, S.T.; Kucharik, C.J.; Norman, J.M. Direct and Indirect Estimation of Leaf Area Index, fAPAR, and Net Primary Production of Terrestrial Ecosystems. *Remote Sens. Environ.* **1999**, *70*, 29–51.
3. Gobron, N.; Pinty, B.; Mélin, F.; Taberner, M.; Verstraete, M.M.; Belward, A.; Lavergne, T.; Widlowski, J. The state of vegetation in Europe following the 2003 drought. *Int. J. Remote Sens.* **2005**, *26*, 2013–2020.
4. Senna, M.; Costa, M.; Shimabukuro, Y. Fraction of photosynthetically active radiation absorbed by Amazon tropical forest: A comparison of field measurements, modeling, and remote sensing. *J. Geophys. Res. Biogeosci.* **2005**, *110*, G01008.
5. Verstraete, M.M.; Gobron, N.; Aussedat, O.; Robustelli, M.; Pinty, B.; Widlowski, J.L.; Taberner, M. An automatic procedure to identify key vegetation phenology events using the JRC-FAPAR products. *Adv. Space Res.* **2008**, *41*, 1773–1783.
6. Coops, N.; Wulder, M.; Duro, D.; Han, T.; Berry, S. The development of a Canadian dynamic habitat index using multi-temporal satellite estimates of canopy light absorbance. *Ecol. Indic.* **2008**, *8*, 754–766.
7. Kaminski, T.; Rayner, P.J.; Voßbeck, M.; Scholze, M.; Koffi, E. Observing the continental-scale carbon balance: Assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design. *Atmos. Chem. Phys.* **2012**, *12*, 7867–7879.
8. Loew, A.; van Bodegom, P.M.; Widlowski, J.L.; Otto, J.; Quaife, T.; Pinty, B.; Raddatz, T. Do we (need to) care about canopy radiation schemes in DGVMs? Caveats and potential impacts. *Biogeosciences* **2014**, *11*, 1873–1897.
9. Weiss, M.; Baret, F.; Garrigues, S.; Lacaze, R. LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: Validation and comparison with MODIS collection 4 products. *Remote Sens. Environ.* **2007**, *110*, 317–331.
10. Román, M.O.; Schaaf, C.B.; Lewis, P.; Gao, F.; Anderson, G.P.; Privette, J.L.; Strahler, A.H.; Woodcock, C.E.; Barnsley, M. Assessing the coupling between surface albedo derived from MODIS and the fraction of diffuse skylight over spatially-characterized landscapes. *Remote Sens. Environ.* **2010**, *114*, 738–760.
11. D'Odorico, P.; Gonsamo, A.; Pinty, B.; Gobron, N.; Coops, N.; Mendez, E.; Schaepman, M.E. Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. *Remote Sens. Environ.* **2014**, *142*, 141–154.
12. Global Climate Observing System (GCOS). The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC. Available online: https://www.wmo.int/pages/prog/gcos/Publications/gcos-82_2AR.pdf (accessed on 26 June 2017).

13. Kimes, D.; Knyazikhin, Y.; Privette, J.; Abuelgasim, A.; Gao, F. Inversion methods for physically based models. *Remote Sens. Rev.* **2000**, *18*, 381–439.
14. Hadamard, J. Sur les problèmes aux dérivées partielles et leur signification physique. *Princet. Univ. Bull.* **1902**, *13*, 49–52.
15. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2002**, *84*, 1–15.
16. Lewis, P.; Gómez-Dans, J.; Kaminski, T.; Settle, J.; Quaife, T.; Gobron, N.; Styles, J.; Berger, M. An Earth Observation Land Data Assimilation System (EO-LDAS). *Remote Sens. Environ.* **2012**, *120*, 219–235.
17. Pinty, B.; Jung, M.; Kaminski, T.; Lavergne, T.; Mund, M.; Plummer, S.; Thomas, E.; Widlowski, J.L. Evaluation of the JRC-TIP 0.01° products over a mid-latitude deciduous forest site. *Remote Sens. Environ.* **2011**, *115*, 3567–3581.
18. Kaminski, T.; Mathieu, P.P. Reviews and syntheses: Flying the satellite into your model: On the role of observation operators in constraining models of the Earth system and the carbon cycle. *Biogeosciences* **2017**, *14*, 2343–2357.
19. Kaminski, T.; Pinty, B.; Voßbeck, M.; Lopatka, M.; Gobron, N.; Robustelli, M. Consistent retrieval of land surface radiation products from EO, including traceable uncertainty estimates. *Biogeosciences* **2017**, *14*, 2527–2541.
20. Lewis, P.; Gomez-Dans, J.; Kaminski, T.; Settle, J.; Quaife, T.; Gobron, N.; Styles, J.; Berger, M. Data assimilation of Sentinel-2 observations: Preliminary results from EO-LDAS and outlook. In Proceedings of the First Sentinel-2 Preparatory Symposium, Frascati, Italy 23–27 April 2012; Volume ESA SP-707.
21. Gómez-Dans, J.L.; Lewis, P.E.; Disney, M. Efficient Emulation of Radiative Transfer Codes Using Gaussian Processes and Application to Land Surface Parameter Inferences. *Remote Sens.* **2016**, *8*, 119.
22. Gitelson, A.A.; Peng, Y.; Huemmrich, K.F. Relationship between fraction of radiation absorbed by photosynthesizing maize and soybean canopies and NDVI from remotely sensed data taken at close range and from MODIS 250 m resolution data. *Remote Sens. Environ.* **2014**, *147*, 108–120.
23. Myneni, R.; Knyazikhin, Y.; Park, T. MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500 m SIN Grid V006. NASA EOSDIS Land Processes DAAC. 2015. Available online: https://doi.org/10.5067/MODIS/MCD15A2H.006 (acccessed on 26 June 2017).
24. Pinty, B.; Widlowski, J.L.; Gobron, N.; Verstraete, M.M.; Diner, D.J.; Member, A. Uniqueness of Multiangular Measurements—Part I: An Indicator of Subpixel Surface Heterogeneity From MISR. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1560–1573.
25. Deschamps, P.Y.; Breon, F.M.; Leroy, M.; Podaire, A.; Bricaud, A.; Buriez, J.C.; Seze, G. The POLDER mission: Instrument characteristics and scientific objectives. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 598–615.
26. SLSTR Instrument. Available online: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-slstr/instrument/description (accessed on 20 June 2017).
27. Diner, D.J.; Beckert, J.C.; Reilly, T.H.; Bruegge, C.J.; Conel, J.E.; Kahn, R.A.; Martonchik, J.V.; Ackerman, T.P.; Davies, R.; Gerstl, S.A.W.; et al. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1072–1087.
28. Knyazikhin, Y.; Martonchik, J.V.; Myneni, R.B.; Diner, D.J.; Running, S.W. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *J. Geophys. Res.* **1998**, *103*, 32257.
29. Gobron, N.; Pinty, B.; Verstraete, M.M.; Martonchik, J.V.; Knyazikhin, Y.; Diner, D.J. Potential of multiangular spectral measurements to characterize land surfaces: Conceptual approach and exploratory application. *J. Geophys. Res.* **2000**, *105*, 17539.
30. Widlowski, J.L.; Pinty, B.; Gobron, N.; Verstraete, M.M.; Diner, D.J.; Davis, A.B. Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies. *Clim. Chang.* **2004**, *67*, 403–415.
31. Laurent, V.C.; Verhoef, W.; Clevers, J.G.; Schaepman, M.E. Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. *Remote Sens. Environ.* **2011**, *115*, 2603–2612.

32. Pinty, B.; Lavergne, T.; Voßbeck, M.; Kaminski, T.; Aussedat, O.; Giering, R.; Gobron, N.; Taberner, M.; Verstraete, M.M.; Widlowski, J.L. Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products. *J. Geophys. Res.* **2007**, *112*, D10116.
33. Verstraete, M.M.; Member, S.; Hunt, L.A.; Scholes, R.J.; Clerici, M.; Pinty, B.; Nelson, D.L. Generating 275-m Resolution Land Surface Products From the Multi-Angle Imaging SpectroRadiometer Data. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3980–3990.
34. Pinty, B.; Lavergne, T.; Dickinson, R.E.; Widlowski, J.L.; Gobron, N.; Verstraete, M.M. Simplifying the Interaction of Land Surfaces with Radiation for Relating Remote Sensing Products to Climate Models. *J. Geophys. Res. Atmos.* **2006**, *111*, D02116.
35. Pinty, B.; Clerici, M.; Andredakis, I.; Kaminski, T.; Taberner, M.; Verstraete, M.M.; Gobron, N.; Plummer, S.; Widlowski, J.L. Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers. *J. Geophys. Res. Atmos.* **2011**, *116*, D09106.
36. Viña, A.; Gitelson, A.A. New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops. *Geophys. Res. Lett.* **2005**, *32*, L17403.
37. MISR Web Site. Available online: http://www-misr.jpl.nasa.gov/ (accessed on 27 January 2017).
38. Giering, R.; Kaminski, T. Recipes for adjoint code construction. *ACM Trans. Math. Softw.* **1998**, *24*, 437–474.
39. Gobron, N.; Pinty, B.; Verstraete, M.; Widlowski, J.L. Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2489–2505.
40. Gobron, N.; Pinty, B.; Verstraete, M.; Govaerts, Y. The MERIS Global Vegetation Index (MGVI): Description and preliminary application. *Int. J. Remote Sens.* **1999**, *20*, 1917–1927.
41. Gobron, N.; Pinty, B.; Aussedat, O.; Taberner, M.; Faber, O.; Mélin, F.; Lavergne, T.; Robustelli, M.; Snoeij, P. Uncertainty estimates for the FAPAR operational products derived from MERIS—Impact of top-of-atmosphere radiance uncertainties and validation with field data. *Remote Sens. Environ.* **2008**, *112*, 1871–1883.
42. Gobron, N. Uncertainties assessment for MERIS/OLCI FAPAR. In Proceeding of the 2015 ESA Sentinel-3 for Science Workshop, Venice, Italy, 2–5 June 2015.
43. Yang, W.; Tan, B.; Huang, D.; Rautiainen, M.; Shabanov, N.V.; Wang, Y.; Privette, J.L.; Huemmrich, K.F.; Fensholt, R.; Sandholt, I.; et al. MODIS leaf area index products: From validation to algorithm improvement. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1885–1898.
44. Zupanski, D. A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Weather Rev.* **1997**, *125*, 2274–2292.
45. Gómez-Dans, J. The `eoldas_ng` Python Library. Available online: http://github.com/jgomezdans/eoldas_ng (accessed on 19 March 2017).
46. Verhoef, W.; Bach, H. Coupled soil—Leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sens. Environ.* **2007**, *109*, 166–182.
47. Quaife, T.; Lewis, P. Temporal Constraints on Linear BRDF Model Parameters. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2445–2450.
48. Gomez-Dans, J.L.; Lewis, P. EOLDAS Users' Documentation. Available online: http://jgomezdans.github.io/eoldas_release/EOLDAS_UsersGuide.pdf (accessed on 26 June 2017).
49. Atzberger, C.; Richter, K. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. *Remote Sens. Environ.* **2012**, *120*, 208–218.
50. Atzberger, C. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sens. Environ.* **2004**, *93*, 53–67.
51. Wang, Y.; Yang, C.; Li, X. Regularizing kernel-based BRDF model inversion method for ill-posed land surface parameter retrieval using smoothness constraint. *J. Geophys. Res.* **2008**, *113*, D13101.
52. Laurent, V.C.; Verhoef, W.; Damm, A.; Schaepman, M.E.; Clevers, J.G. A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from APEX at-sensor radiance data. *Remote Sens. Environ.* **2013**, *139*, 6–17.
53. Chernetskiy, M.; Gomez-Dans, J.; Lewis, P. Validation of the Earth Observation Land Data Assimilation System by the field data of ESA SPARC field campaign. In Proceedings of the 2013 ESA Living Planet Symposium, Edinburgh, UK, 9–13 September 2013; Volume ESA SP-722, pp. 1–5.

54. Mousivand, A.; Menenti, M.; Gorte, B.; Verhoef, W. Multi-temporal, multi-sensor retrieval of terrestrial vegetation properties from spectral–directional radiometric data. *Remote Sens. Environ.* **2015**, *158*, 311–330.
55. Quaife, T.; Lewis, P.; De Kauwe, M.; Williams, M.; Law, B.; Disney, M.; Bowyer, P. Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sens. Environ.* **2008**, *112*, 1347–1364.
56. Thacker, W.C. The role of the Hessian matrix in fitting models to measurements. *Geophys. Res.* **1989**, *94*, 6177–6196.
57. Weiss, M.; Baret, F.; Myneni, R.; Pragnère, A.; Knyazikhin, Y. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie* **2000**, *20*, 3–22.
58. Gobron, N.; Pinty, B.; Verstraete, M.M.; Govaerts, Y. A semidiscrete model for the scattering of light by vegetation. *J. Geophys. Res.* **1997**, *102*, 9431–9446.
59. Jacquemoud, S.; Baret, F. PROSPECT: A model of leaf optical properties spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91.
60. Price, J. On the information content of soil reflectance spectra. *Remote Sens. Environ.* **1990**, *33*, 113–121.
61. Beck, P.S.A.; Atzberger, C.; Høgda, K.A.; Johansen, B.; Skidmore, A.K. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sens. Environ.* **2006**, *100*, 321–334.
62. Shunlin, L.; Zhiqiang, X. Global Land Surface Products: Leaf Area Index Product Data Collection (1985–2010). 2012. Available online: http://glcf.umd.edu/data/lai/ (accessed on 26 June 2017).
63. Zarco-Tejada, P.J.; Rueda, C.; Ustin, S. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* **2003**, *85*, 109–124.
64. Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.; Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—A review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290.
65. Camps-Valls, G.; Verrelst, J.; Munoz-Mari, J.; Laparra, V.; Mateo-Jimenez, F.; Gomez-Dans, J. A Survey on Gaussian Processes for Earth-Observation Data Analysis: A Comprehensive Investigation. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 58–78.
66. Gómez-Dans, J.; Lewis, P. `gp_emulator`: A Python Library for Emulating Radiative Transfer Codes. 2016. Available online: http://jgomezdans.github.io/gp_emulator/ (accessed on 26 June 2017).
67. Earth System Research Laboratory: Solar Calculation Details. Available online: http://www.esrl.noaa.gov/gmd/grad/solcalc/calcdetails.html (accessed on 30 January 2017).
68. Weiss, M.; Baret, F.; Smith, G.J.; Jonckheere, I.; Coppin, P. Review of methods for in situ leaf area index (LAI) determination Part II. Estimation of LAI, errors and sampling. *Agric. For. Meteorol.* **2004**, *121*, 37–53.
69. Gobron, N.; Pinty, B.; Verstraete, M.M. Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 1438–1445.
70. Disney, M.; Muller, J.P.; Kharbouche, S.; Kaminski, T. A New Global fAPAR and LAI Dataset Derived from Optimal Albedo Estimates: Comparison with MODIS Products. *Remote Sens.* **2016**, *8*, 275.
71. Gomez-Dans, J.; Lewis, P.; Disney, M.; Roy, D.; Quaife, T.; Wooster, M. Edge-Preserving Data Assimilation for Fire Monitoring Using Optical Data. Available online: https://ftp.space.dtu.dk/pub/Ioana/papers/s252_gome.pdf (accessed on 26 June 2017).
72. Walthall, C.L.; Norman, J.M.; Welles, J.M.; Campbell, G.; Blad, B.L. Simple equation to approximate the bidirectional reflectance from vegetative canopies and bare soil surfaces. *Appl. Opt.* **1985**, *24*, 383.
73. Hapke, B. Bidirectional reflectance spectroscopy: 1. Theory. *J. Geophys. Res.* **1981**, *86*, 3039–3054.
74. Liang, S.; Townshend, J.R. A modified hapke model for soil bidirectional reflectance. *Remote Sens. Environ.* **1996**, *55*, 1–10.
75. Liang, S. Numerical experiments on the spatial scaling of land surface albedo and leaf area index. *Remote Sens. Rev.* **2000**, *19*, 225–242.
76. Pfeifer, M.; Disney, M.; Quaife, T.; Marchant, R. Terrestrial ecosystems from space: A review of earth observation products for macroecology applications. *Glob. Ecol. Biogeogr.* **2012**, *21*, 603–624.
77. Kira, O.; Nguy-Robertson, A.L.; Arkebauer, T.J.; Linker, R.; Gitelson, A.A. Toward Generic Models for Green LAI Estimation in Maize and Soybean: Satellite Observations. *Remote Sens.* **2017**, *9*, 318.

78. Nguy-Robertson, A.L.; Gitelson, A.A. Algorithms for estimating green leaf area index in C3 and C4 crops for MODIS, Landsat TM/ETM+, MERIS, Sentinel MSI/OLCI, and Venμs sensors. *Remote Sens. Lett.* **2015**, *6*, 360–369.
79. Morisette, J.T.; Baret, F.; Privette, J.L.; Myneni, R.B.; Nickeson, J.E.; Garrigues, S.; Shabanov, N.V.; Weiss, M.; Fernandes, R.A.; Leblanc, S.G.; et al. Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1804–1814.
80. Baret, F.; Weiss, M.; Allard, D.; Garrigue, S.; Leroy, M.; Jeanjean, H.; Fernandes, R.; Myneni, R.; Privette, J.; Morisette, J.; et al. VALERI: A network of sites and a methodology for the validation of medium spatial resolution land satellite products. *Remote Sens. Environ.* **2005**, *76*, 36–39.
81. Garrigues, S.; Lacaze, R.; Baret, F.; Morisette, J.T.; Weiss, M.; Nickeson, J.E.; Fernandes, R.; Plummer, S.; Shabanov, N.V.; Myneni, R.B.; et al. Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. *J. Geophys. Res.* **2008**, *113*, G02028.
82. Povey, A.C.; Grainger, R.G. Known and unknown unknowns: uncertainty estimation in satellite remote sensing. *Atmos. Meas. Tech.* **2015**, *8*, 4699–4718.
83. Pinnington, E.M.; Casella, E.; Dance, S.L.; Lawless, A.S.; Morison, J.I.L.; Nichols, N.K.; Wilkinson, M.; Quaife, T.L. Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using Four-dimensional Variational data assimilation. *Agric. For. Meteorol.* **2016**, *228–229*, 299–314.

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Elsevier Editorial System(tm) for Advances in Space Research

Manuscript Draft

Manuscript Number: ASR-D-18-00123R1

Title: Simulating arbitrary hyperspectral bandsets from multispectral observations via a generic Earth Observation-Land Data Assimilation System (EO-LDAS)

Article Type: ES - Earth Sciences

Keywords: MISR; Landsat; CHRIS/PROBA; EO-LDAS; Semi-discrete radiative transfer model; Barrax.

Corresponding Author: Mr. Maxim Chernetskiy,

Corresponding Author's Institution: University College London (UCL)

First Author: Maxim Chernetskiy

Order of Authors: Maxim Chernetskiy; Nadine Gobron, Dr.; Jose Gomez-Dans, Dr.; Olivier Morgan; Mathias Disney, Dr.; Philip Lewis, Prof. Dr.; Christiane Schmullius, Prof. Dr.

Abstract: This paper presents results of using multi-sensor and multi-angular constraints in the generic Earth Observation-Land Data Assimilation System (EO-LDAS) for reproducing arbitrary bandsets of hyperspectral reflectance at the top-of-canopy (TOC) level by merging observations from multispectral sensors with different spectral characteristics. This is demonstrated by combining Multi-angle Imaging Spectroradiometer (MISR) and Landsat Enhanced Thematic Mapper Plus (ETM+) data to simulate the Compact High Resolution Imaging Spectrometer CHRIS/PROBA hyperspectral signal over an agricultural test site, in Barrax, Spain. However, the method can be more generally applied to any combination of spectral data, providing a tool for merging EO data to any arbitrary hyperspectral bandset.
Comparisons are presented using both synthetic and observed MISR and Landsat data, and retrieving surface biophysical properties. We find that when using simulated MISR and Landsat data, the CHRIS/PROBA hyperspectral signal is reproduced with RMSE 0.0001 - 0.04. LAI is retrieved with r2 from 0.97 to 0.99 and RMSE of from 0.21 to 0.38. The results based on observed MISR and Landsat data have lower performances, with RMSE for the reproduced CHRIS/PROBA hyperspectral signal varying from 0.007 to 0.2. LAI is retrievedwith r2 from 0.7 to 0.9 and RMSE from 0.7 to 1.4. We found that for the data considered here the main spectral variations in the visible and near infrared regions can be described by a limited number of parameters (3-4) that can be estimated from multispectral information. Results show that the method can be used to simulate arbitrary bandsets, which will be of importance to any application which requires combining new and existing streams of new EO data in the optical domain, particularly intercalibration of EO satellites in order to get continuous time series of surface reflectance, across programmes and sensors of different designs.

Manuscript
Click here to view linked References

Simulating arbitrary hyperspectral bandsets from multispectral observations via a generic Earth Observation-Land Data Assimilation System (EO-LDAS)

M. Chernetskiy[1], N. Gobron[2], J. Gomez-Dans[1,3] , O. Morgan[2], M. Disney[1,3], P. Lewis[1,3] and C. Schmullius[4]

[1] *University College London (UCL), United Kingdom*

[2] *European Commission Joint Research Centre, Ispra (VA), Italy*

[3] *NERC National Centre for Earth Observation (NCEO), United Kingdom*

[4] *Friedrich Schiller University, Institute of Geography - Department for Earth Observation, Jena, Germany*

Correspondence to: M. Chernetskiy: *m.chernetskiy@ucl.ac.uk*

Abstract

This paper presents results of using multi-sensor and multi-angular constraints in the generic Earth Observation-Land Data Assimilation System (EO-LDAS) for reproducing arbitrary bandsets of hyperspectral reflectance at the top-of-canopy (TOC) level by merging observations from multispectral sensors with different spectral characteristics. This is demonstrated by combining Multi-angle Imaging Spectroradiometer (MISR) and Landsat Enhanced Thematic Mapper Plus (ETM+) data to simulate the Compact High Resolution Imaging Spectrometer CHRIS/PROBA hyperspectral signal over an agricultural test site, in Barrax, Spain. However, the method can be more generally applied to any combination of spectral data, providing a tool for merging EO data to any arbitrary hyperspectral bandset.

Comparisons are presented using both synthetic and observed MISR and Landsat data, and retrieving surface biophysical properties. We find that when using simulated MISR and Landsat data, the CHRIS/PROBA hyperspectral signal is reproduced with RMSE 0.0001 - 0.04. LAI is retrieved with r^2 from 0.97 to 0.99 and RMSE of from 0.21 to 0.38. The results based on observed MISR and Landsat data have lower performances, with RMSE for the reproduced CHRIS/PROBA hyperspectral signal varying from 0.007 to 0.2. LAI is

retrievedwith r^2 from 0.7 to 0.9 and RMSE from 0.7 to 1.4. We found that for the data considered here the main spectral variations in the visible and near infrared regions can be described by a limited number of parameters (3-4) that can be estimated from multispectral information. Results show that the method can be used to simulate arbitrary bandsets, which will be of importance to any application which requires combining new and existing streams of new EO data in the optical domain, particularly intercalibration of EO satellites in order to get continuous time series of surface reflectance, across programmes and sensors of different designs.

*Keywords:*MISR; Landsat; CHRIS/PROBA; EO-LDAS; Semi-discrete radiative transfer model; Barrax.

1 Introduction

An understanding of surface reflectance over the solar reflective domain (with wavelengths from 400 to 2500 nm) is important in order to monitor the land surface with spaceborne passive optical sensors. Typically, these sensors acquire data in a limited set of bands (e.g. multispectral sensors, such as Landsat (6 bands), Moderate Resolution Imaging Spectroradiometer (MODIS) (7 bands) or Sentinel-2/MSI (12 bands). By contrast, the Compact High Resolution Imaging Spectrometer (CHRIS) instrument on board Proba-1 (Barnsley et al., 2004) collects data with a higher spectral resolution (62 bands of 1.3-12 nm width), typically referred to as a hyperspectral sensor. Although hyperspectral data are presented as airborne sensors measurements for decades, they are not so common on the space borne platforms. We note that a number of space missions are expected to be launched in the next few years with the remit of acquiring hyperspectral data. Among these missions are the Environmental Mapping and Analysis Program (EnMAP), Hyperspectral PRecursor of the Application Mission (PRISMA) and Hyperspectral Infrared Imager (HyspIRI) (Guanter et. al. 2015; Candela et. al. 2016; Lee et. al. 2016).

Hyperspectral data are routinely collected from airborne sensors (e.g. Asner et al. 2016), on the ground and have been used for validation and calibration of space borne sensors (Gupta et al., 1998, Baccini et al., 2007, Hay et al., 1997, 2001). Hyperspectral sensors are also being mounted on automated acquisition platforms in flux towers (Porcar-Castell et al. 2015). There has been an increased interest in hyperspectral observations as a way to characterise leaf traits such as specific leaf area or leaf nitrogen content (Roelofsen 2014, Musavi 2015)

A typical application of Earth Observation (EO) data in the optical domain over the land surface is the retrieval of biophysical parameters, often carried out through the inversion of a radiative transfer (RT) model. The resulting derived parameters are usually 'validated' by comparisons with ground based measurements (Baret et al, 2006) and/or comparisons among different products (Disney et al. 2016). These validation methods have a number of shortcomings, as detailed for leaf area index (LAI) in e.g. (Disney et al. 2016), which mentions that a number of incompatible assumptions can be made when gathering "ground truth" data and the retrieval scheme chosen (or among different products). The issue of the scale of the measurements is also important (Pfeifer et al. 2012; Widlowski et al. 2005). Validation of other parameters that describe leaf optical properties is also fraught with complications due to similar reasons. An additional, independent, test of an inversion scheme is that the results obtained from the inversion ought to allow one to predict observations from a different sensor, with arbitrary angular and spectral properties. In this respect, hyperspectral sensors present a spectrally comprehensive dataset to compare against.

Data assimilation (DA) schemes, such as the Earth Observation Land Data Assimilation System (EO-LDAS) of Lewis et al. (2012) and Gomez-Dans et al (2016) produce inferences of land surface parameters based on EO data combined with a number of *a priori* additional constraints. The EO-LDAS approach maps land surface parameters (such as LAI, or leaf and soil optical properties) to surface directional reflectance by means of an RT model scheme. Thus, if the land surface parameters are known, the RT model can be used to predict observations from another sensor, with different acquisition geometries, spectral characteristics, etc. Additionally, the ability to produce a complete time series of parameters allows the prediction of observations when no other sensor data is available.

An example of this approach is provided in Verhoef and Bach (2003), where estimates of LAI, fraction of brown leaves and soil moisture derived from inverting Landsat data, were used to forward model HyMap imaging spectrometer observations. Verhoef and Bach (2007) concentrate on simulation of Top Of Atmosphere (TOA) multi-angular hyperspectral signal by coupling soil-leaf-canopy and atmosphere RT models. In order to validate the RT model results, they simulate hyperspectral signal of CHRIS/PROBA at the bottom-of-atmosphere (BOA) level for bare soil, maize, dense and sparse forest. Comparison of results between real and simulated CHRIS/Proba measurements show RMSE from 0.011 to 0.027.

The inverse problem is known to be ill-posed (Kimes et al. 2000), in practice meaning that

there may be infinite solutions that fit the observations equally well. One way around this is to use prior constraints (Combal et al. 2003), complemented by regularisation approaches (Lewis et al. 2012) or by models of the parameter evolution (Koetz et al. 2005; Quaife et al. 2007; Gomez-Dans et al. 2016). Adding more (independent) observations is also an obvious way to add more constraints to the problem. In all these cases, the original set of observations are being complemented by extra information that restricts the solution space.

Several studies have shown that multi-angular information can improve retrieval of land parameters such as LAI. For instance, Knyazikhin et al. (1998) describe the algorithm for synergistic retrieval of LAI and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) from MODIS and MISR measurements. Gobron et al. (2002) demonstrate that using multiple observational angles from MISR reduces the number of solutions when inverting a canopy RT model. Other studies have demonstrated that MISR can also provide information on structure and heterogeneity of vegetation (Pinty et al. (2002), Gobron et al. (2002) and Widlowski et al. (2004)).

There are currently in excess of 100 EO sensors acquiring data over the land surface. Products that combine observations from different sensors are still relatively rare, as the different spectral, spatial, angular and temporal characteristics of the data, as well as artefacts introduced by parts of the processing, result in a challenging problem. The identification of so-called essential climate variables (ECVs) (Hollman et al 2013, Bojinksy et al 2014) is providing a push towards datasets of scientific parameters that use observations from all available satellites, resulting in consistent, uncertainty-quantified, long term records.

In this paper, we aim to advance the development of multi-sensor products by demonstrating a method for using data from a sensor with relatively few spectral bands can be used to predict data from a hyperspectral sensor, via a DA approach. To achieve this we use data from the MISR (Diner et al., 1998) and Landsat ETM+ sensors to predict bottom-of-atmosphere (BOA) hyperspectral reflectance from the CHRIS/PROBA sensor.

The main assumption we make here is that the major spectral variations in the visible-near infrared region (400 to 1000 nm, broadly) are controlled by a small number of mechanisms that can be parameterised by a few magnitudes (basis functions) provided by multispectral information. For example Price (1994; 1998) showed that vegetation reflectance exhibits a strong degree of redundancy i.e. much of the variation is concentrated in a few key features.

Price (1990) examined the reflectance spectra of 500 different soils at field scale, and showed via principal component analysis that 99.6% of the variance could be explained by just 5 principal components. The main reason of this redundancy is the broad spectral representation of a limited number of underlying physical parameters. More recently, Verrelstet. al. (2016) carried out a sensitivity analysis of the PROSAIL model (Jacquemoud et. al. 2009) and reported that the main driving factors are LAI, Leaf Angular Distribution (LAD) and soil coefficients, which contribute up to 90% of whole spectral variability. Another example of this issue of redundancy is given by Mousivandet. al. (2015) who found that the main driving factors for the Soil-Leaf-Canopy (SLC) model (Verhoef and Bach 2003) and CHRIS-Proba are crown coverage, LAI, leaf inclination distribution function, soil moisture, chlorophyll content of green leaves and fraction of brown leaves. Under this assumption, these few parameters may be retrieved from a limited spectral sampling, and then used to forward model observations from other sensors with different spectral properties. This in turn allows sensors with different spectral properties to be compared directly, or more importantly, merged in a common spectral framework.

In the first part of the study, we test the assumptions above with a theoretical experiment, in which a synthetic observational scenario is used to assess the limitations of the proposed approach without considering complicating factors such as residual effects from atmospheric correction, gridding or sensor calibration. In the second part we apply the same approach to the real multispectral data of MISR and Landsat, both separately and then combined.

2 Methods

Fig. 1 shows the steps of the study. In the “Synthetic data” section, we generate multispectral Landsat and MISR data, then we use the EO-LDAS approach to invert the semi-discrete RT model of Gobron et al. (1997) and retrieve the state parameters. Then we use forward model to simulate hyperspectral and multi-angular observations from CHRIS-PROBA. In the “Real data” section, we apply the same steps to the real multispectral data. For “Comparison”, we use field measurements collected during the SPectrabARrax Campaign (SPARC) 2004 (Gandia et al., 2004).

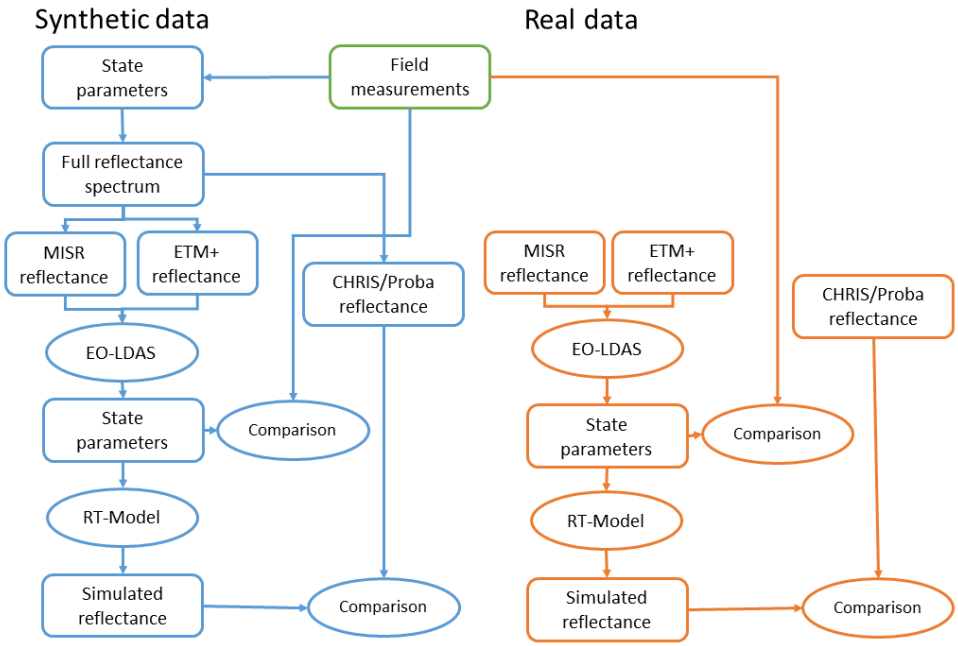


Fig. 1. Simulating synthetic and real CHRIS/PROBA spectra

2.1 Study site

The test site of this study is an agricultural site located near Barrax in Spain (fig. 2).

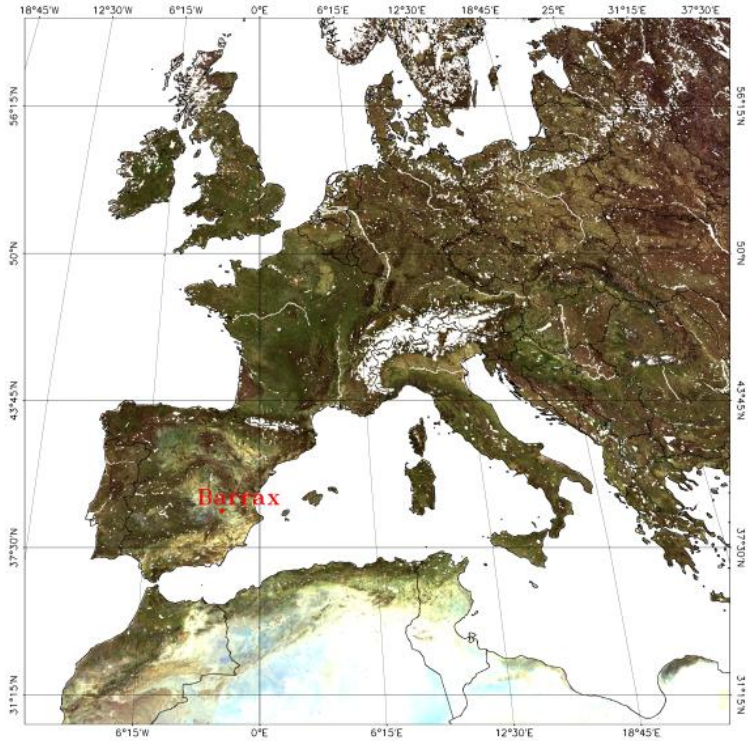


Fig. 2. The Barrax test site on the satellite map of Europe.

Barrax is an European Space Agency (ESA) test site where many field campaigns have been

carried out over the years 2003, 2004, 2005 and 2009. The activities of the SPARC 2004 (Gandia et al., 2004) ground-based campaign included the retrieval of biophysical parameters, such as LAI, chlorophyll concentration, leaf equivalent water thickness and leaf dry matter. Ground spectral measurements were taken with an Analytical Spectral Devices (ASD) FieldSpec Pro FR Spectroradiometer (ASD inc.), including bare soil and vegetation of some fields. The crop types across the site include alfalfa, sugar beet, corn, garlic, onion, potato, sunflower and vineyard, with a total of 18 fields being measured. Biophysical parameters which are relevant for this work are presented in Table 1.

Table 1: Parameters from the SPARC database.

field number	field code	field code	LAI (m2/m2)	Cab (ug/cm2)	Leaf water (Cw) (g/m2)	Dry matter (Cdm) (g/m2)
1	Alfalfa	A1	3.73	50.17	36.0	144.6
2	Alfalfa	A2	3.36	48.7	41.55	119.3
3	Sugar Beet	B1	4.48	49.05	64.5	424.1
4	Corn	C1	1.69	49.33	76.2	148.2
5	Corn	C9	2.92	52.94	76.0	148.0
6	Corn	C10	2.57	52.16	76.0	148.0
7	Garlic	G1	0.63	51.71	100.4	654.8
8	Garlic	G1	0.63	48.22	100.0	655.0
9	Garlic	G1	0.63	43.6	100.0	655.0
10	Garlic	G1	0.63	39.02	100.0	655.0
11	Onion	On3	0.95	31.26	71.3	499.4
12	Potato	P2	3.96	37.98	40.7	270.8
13	Potato	P3	4.03	34.85	41.0	271.0
14	Sunflower	SF1	0.5	43.35	87.5	368.6
15	Sunflower	SF1	0.5	44.1	87.0	369.0
16	Sunflower	SF1	0.5	39.76	87.0	369.0
17	Sunflower	SF3	0.66	43.71	69.6	404.9
18	Vineyard	V1	2.51	34.58	91.0	202.3

2.2 Remote sensing data

Data from the CHRIS/Proba sensor have been collected contemporary with the SPARC ground measurements campaign. In mode 1, the data provide 62 spectral bands in the visible and near-infrared region, from 411 to 997 nm with a spatial resolution at nadir of 34 m and five view zenith angles: +/- 55°, +/-36° and 0° (nominally). Spectral resolution varies from 1.3 nm at the 410 nm band to 12 nm at the 1050 nm band (Barnsley et al. 2004). Fig. 3 illustrates the illumination (stars coloured in orange) and viewing geometry (green dots) of CHRIS/Proba data acquired over the Barrax site on 16th July 2004.

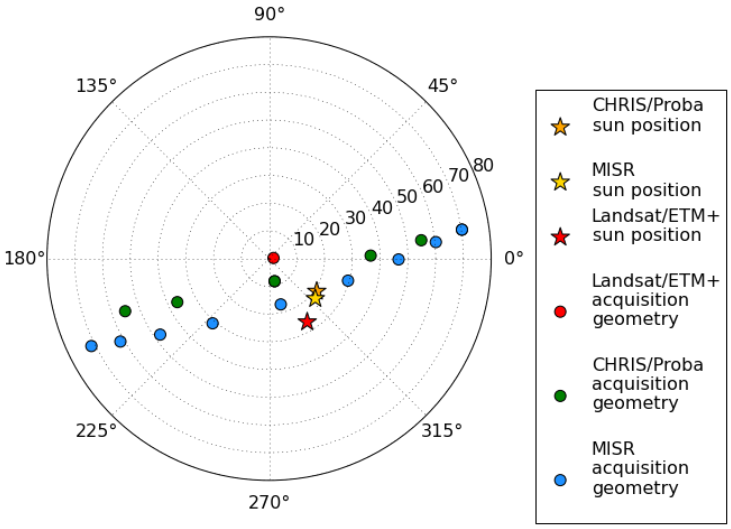


Fig. 3. Acquisition geometry in polar coordinates of the MISR, Landsat ETM+ and CHRIS/Proba data.

MISR is currently the only operational multi-angular optical sensor that obtains information globally and simultaneously at multiple (nine) view angles. We use the MISR top-of-canopy (TOC) reflectance at 275 m spatial resolution produced by the sharpening method of Verstraete et al. (2012) for the 16th July 2004. The data comprise four spectral bands at 446 nm, 558 nm, 672 nm and 867 nm with nine cameras at nadir view, +/-26.1°, +/-45.6°, +/-60.0°, and +/-70.5°. Acquisition geometry for all cameras is also represented in Fig. 3 in blue dots. MISR per band uncertainties for real data are specified as 0.05 for all four bands. (Chernetskiy et al 2017).

The Landsat ETM+ data were acquired on the 18th July 2004 with 30 m spatial resolution and 6 spectral bands. The viewing and illumination geometry are also displayed in Fig. 3. Atmospheric correction was performed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software (Masek et al., 2012). The default LEDAPS routine was used. All input information required for atmospheric correction was taken from Landsat

metadata; ancillary data of (Total Ozone Mapping Spectrometer) TOMS and National Centres for Environmental Prediction (NCEP) Reanalysis. Fig. 4 shows true colour composites of the three types of satellite observations. We assume that a Landsat pixel corresponds to centre of MISR pixel where we have highest response of Point Spread Function (PSF). Landsat per band uncertainties for real data are assumed to be normal with zero mean and standard deviations of0.0202, 0.0246, 0.0250, 0.0214, 0.0395 and 0.0272 for bands 1-5, 7 respectively (Maiersperger et al 2013). Real per band uncertainties for MISR and Landsat are not provided and their estimation is a complicated task, which is outside of this study.

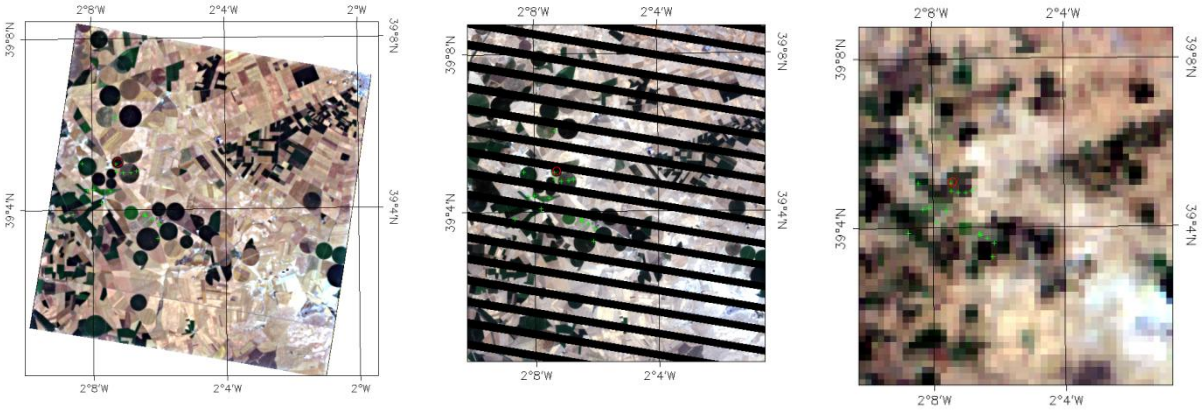


Fig. 4. True colour Images of CHRIS/Proba (left hand side panel), Landsat ETM+ (middle panel) and MISR at 275 m (right hand side panel). Green cross indicate the 18 fields listed in Table 1.

Figure 4 shows that the image of Landsat-7 ETM+ was acquired with Scan Line Corrector (SLC) turned off. This means that we have vertical 'stripes' where data are lost. However apart from some loss of data the SLC-off imagery has the same radiometric and geometric quality as imagery collected before SLC failure. We can see that most of the field measurements are within regions where data are available. The two fields where data are lost are Alfalfa (A2) and Onion (On3) are excluded from the analysis of the real Landsat data.

2.3 Methods

EO-LDAS (Lewis et al., 2012) is a generic variational data assimilation system designed to infer land surface properties (and associated uncertainties) from heterogeneous combinations of observations (interpreted through observation operators, typically RT models) and an array of prior constraints. The DA system (Gomez-Dans and Lewis 2012) defines a cost function, which is then minimised by a gradient descent approach. This cost function describes the model fit to observations, as well as penalties from departing from prior values:

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

where x is the state vector: each element contains the value of a land surface parameter. In this study, x contains ten parameters, eight vegetation and two soil parameters. Various parameter transformations are used to quasi-linearise the RT model (Weiss et al. 2000). $J_{obs}(x)$ is the fit to the observations and $J_{prior}(x)$ is the *a priori* parameter distribution.

The observational operator embedded in $J_{obs}(x)$ provides a mapping from land surface parameters to directional surface reflectance using the semi-discrete radiative transfer canopy model of Gobron et al. (1997), coupled to the leaf optical PROpertiesSPECTra (PROSPECT) of Jacquemoud and Baret (1990) and the spectral soil model of Price (1990) for the albedo of background soil. We use modification of the PROSPECT-5 (Lewis et al 2008, Féret et al 2008). The soil model is described as:

$$R_0 = s_1\varphi_1 + s_2\varphi_2 \qquad (2)$$

where R_0 is soil background reflectance, φ_1 and φ_2 are basis functions (principle components) of Price, s_1 and s_2 are soil model coefficients (Price 1990).

The prior component $J_{prior}(x)$ contains the prior parameter distribution. The prior probability distribution function (pdf) in EO-LDAS is Gaussian (so defined by a mean vector and a covariance matrix) and it encodes the belief on the distribution of the parameter values before the observations are taken into account. In this study, we assume the prior distribution to have means and standard deviations given by the values shown in Table 2. There are 8 state parameters for the canopy and leaf model and two for the soil model. Note that the Leaf Angle Distribution (LAD) is fixed to one of the five Bunnik archetype expressions (Bunnik 1978). Prior information on parameters is taken from (Lewis at al. 2012). LAI and height are set to values which are reasonable for an agricultural field. The prior covariance matrix is assumed to be diagonal.

Table 2: The bounds of state variables and prior information.

Parameter	Transformation	Lower limit	Upper limit	Prior mean	Prior sd.
LAI [m^2/m^2]	$e^{(-x/2)}$	0.08	0.95	0.37	1.0
Canopy height (xh) [m]	-	0.1	5.0	1	3
The leaf radius (xr) [m]	-	0.001	0.3	0.15	0.15
Concentration of chlorophyll a+b (C_{ab}), [µg/cm^2]	$e^{(-x/100)}$	0.5	0.75	0.6	0.4
Proportion of senescent material	-	0.001	1.0	0.5	0.5

(scen)					
Equivalent leaf water (C_w) [cm]	$e^{(-x*50)}$	0.03	0.999	0.6	0.5
Dry matter (C_{dm}), [µg/cm^2]	$e^{(-x*100)}$	0.13	0.99	0.37	0.7
Number of leaf layers (N)	-	1	2.5	1.5	1.5
Soil PC1 (s_1)	-	-20.0	-1.0	-12.21	1.33
Soil PC2 (s_2)	-	-1.0	2.0	0.32	2.14
Leaf angle distribution (LAD)	-	Planophile - 1 Erectophile - 2 Plagiophile - 3 Extremophile - 4 Uniform - 5			

The aim of the exponential transformation of some parameters from Table 2 is to quasi-linearize the RT model, so that the assumption that the posterior PDF is Gaussian is better met.

The assumptions in EO-LDAS result in the posterior pdf of x being given by a multivariate normal. The mean vector is given by the minimum of $J(x)$ (effectively, minus the log posterior), and the posterior uncertainty is given by the inverse of the Hessian at the minimum, using a linear approximation of the Hessian (Lewis et al. 2012).

One important issue is to find an initial estimate of the state parameters because it determines starting values – the first guess for an inversion process. A good first guess reduces the chance of being trapped in local minima of the decision space and typically results in faster convergence of the gradient descent procedure. A sampling design of 10 starting points over parameter space was used. We assume a uniform distribution and use a Latin Hypercube Sampling (LHS) scheme (Sacks et al. 1989). The optimisation that resulted in the lowest cost was chosen. The main advantage of using a LHS over other methods (e.g. random sampling) is that it will produce a space-filling sampling design for a given number of samples and dimensions.

3 Results

3.1 Soil model

Many fields of the Barrax region have relatively low LAI in range from 0.5 to 1.0, meaning the soil reflectance will have a strong influence on the overall reflectance signal and hence that incorrect soil description can be a source of uncertainties (Table 1). However, availability of the ASD soil measurements gives the possibility of finding more precise φ_1 and φ_2

functions (Eq 2). We apply Principal Component Analysis (PCA) to 62 ASD SPARC spectral measurements of bare soil and find that first two principal components (PC) explain more than 97 % of all variability (Fig 5).

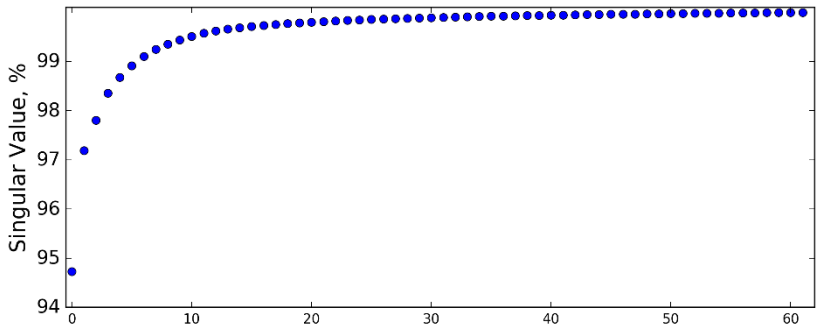


Fig. 5. Results of PCA analysis of 62 ASD measurements.

We use these PCs as φ_1 and φ_2 vectors of the Price soil model. Then we invert this soil model against available ASD measurements and find s_1 and s_2 coefficients. For this we solve a linear system:

$$\bar{s} = (AC^{-1}A^T)^{-1}AC^{-1}\bar{\rho} \qquad (3)$$

where ;

A is a matrix of phi functions:

$$A = [[\varphi_1(t_1,\lambda_1),\dots,\varphi_1(t_1,\lambda_{2101}),\dots,\varphi_1(t_n,\lambda_{2101})],[\varphi_2(t_1,\lambda_1),\dots,\varphi_2(t_1,\lambda_{2101}),\dots,\varphi_2(t_n,\lambda_{2101})]] \quad (4)$$

C is a covariance matrix; ρ is a vector of reflectance:

$$\rho = [\rho(t_1\lambda_1),\dots,\rho(t_1\lambda_{2101}),\dots,\rho(t_n\lambda_1),\dots,\rho(t_n\lambda_{2101})] \quad (5)$$

We find that μ_{s1}=-12.21, σ_{s1}=1.33 and μ_{s2}=0.32, σ_{s2}=2.14. The soil model defined by the resulting coefficients is shown in Fig 6. We can see that in general the model follows the ASD measurements except some for some small differences in the NIR region, but still remaining well within the range of uncertainties. This model is a more precise description of the Barrax soil and allows for the use of the resulting μ and σ values in the EO-LDAS prior term.

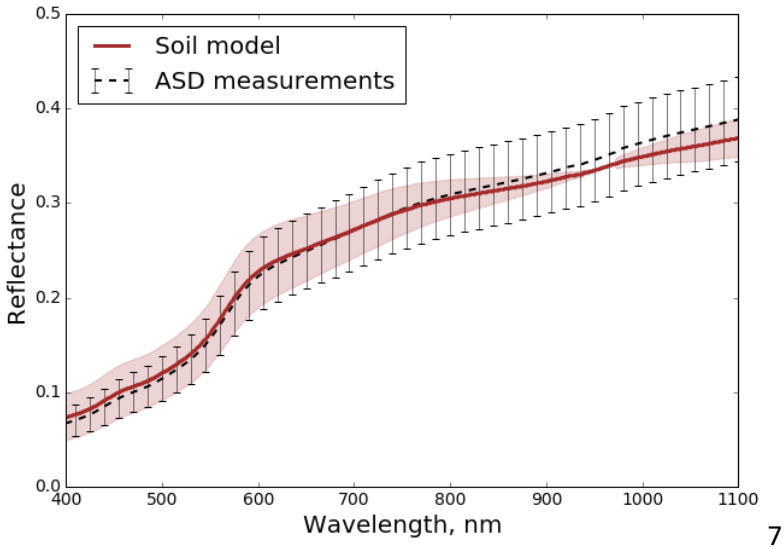


Fig. 6. Soil model with associated uncertainties in 1σ (brown line). Mean and standard deviation (1σ) of bare soil ASD measurements (dotted line).

3.2 Reproducing real spectral information

Here, we test the potential of the semi-discrete RT model to reproduce spectral characteristics of the Barrax vegetation. Values of LAI, C_{ab}, C_w and C_{dm} are known from the SPARC database (Table 1), s_1 and s_2 were found in the previous section. This means that we know six parameters, which according to, e.g. Mousivandet. al. 2015, are most influential. The unknown values of xh, xr, scen and N can be found by inversion of available CHRIS/Proba data. In order to simplify and speed-up the process we generate a look-up-table (LUT) with 4000 entries by means of LHS filled with the values of xh, xr, scen and N. The resulting spectra obtained by minimization of the LUT are shown in figures 7-8. In order to check ability of the model to reproduce the bidirectional reflectance distribution function (BRDF) in this particular situation, we modelled reflectance for all five CHRIS/Proba 'fly-by' positions. Table 3 shows RMSE in range from 0.004 to 0.061 for all 'fly-by' positions. We can see that slightly higher minimum RMSE (0.009-0.01) is for +/-55° positions but highest maximum RMSE (0.061) is for nadir view. Fig.7 shows that the best results are for the fields with lowest LAI, demonstrating the accuracy of the chosen soil model. The highest RMSE is 0.061 for the potato field P2 which is responsible for highest RMSE for all viewing directions (Table 3). So maximum values of RMSE mostly due to discrepancies in some fields which can not be parameterised with given model assumptions (e.g continuous canopy) and spatial resolution. If we accept the correctness of the field measurements these discrepancies can potentially be explained by structural effects at the canopy (clumping) and/or leaf level, that cannot be

simulated by the available model assumptions. This example also demonstrates that the soil model correctly represents background signal and can reproduce BRDF. Results of this section demonstrate that with given ground truth measurements, it is possible to find a set of parameters that allow reproducing hyperspectral data over the Barrax fields in all CHRIS/Proba viewing directions.

Table 3: Minimum and maximum RMSE for fit between real and modelled CHRIS/Proba. LAI, C_{ab}, C_w and C_{dm} are fixed to the ground truth values.

'fly-by' position	0°	-36°	+36°	+55°	-55°
min RMSE	0.004	0.005	0.005	0.010	0.009
max RMSE	0.061	0.048	0.055	0.059	0.053
mean RMSE	0.022	0.023	0.024	0.024	0.027
std. RMSE	0.015	0.015	0.013	0.013	0.014

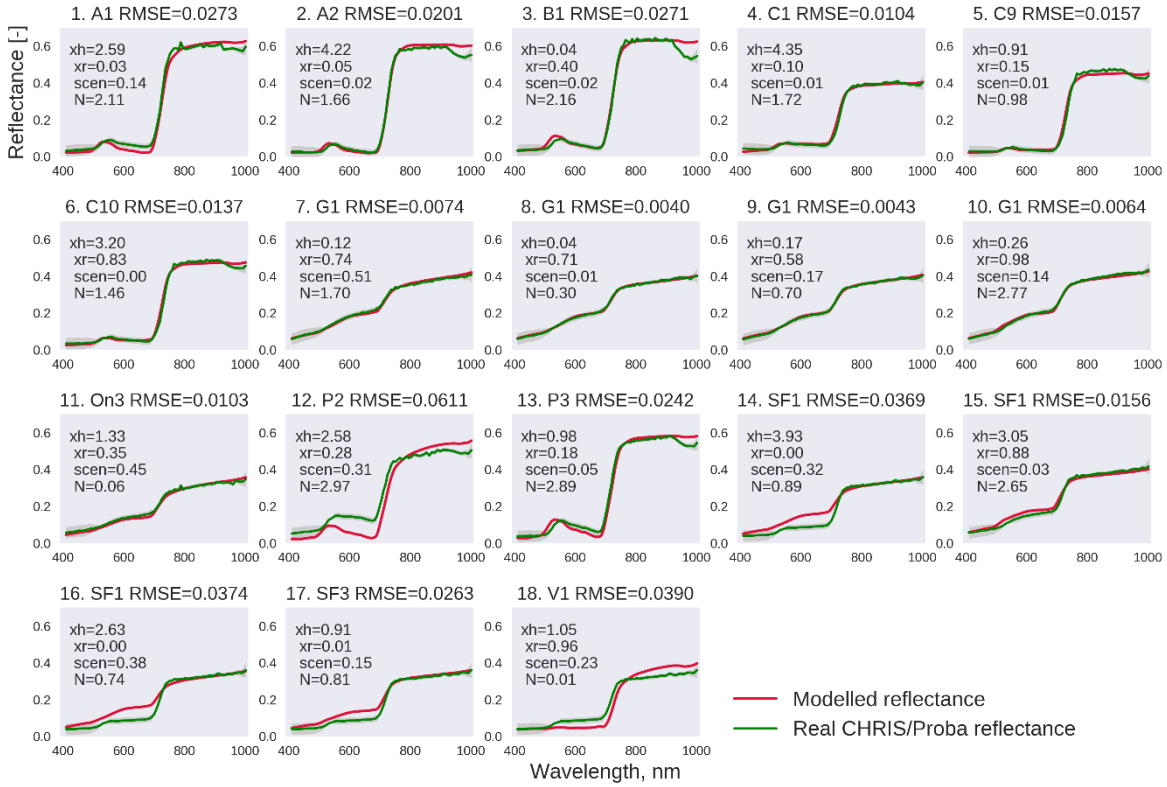


Fig. 7. Real and modelled CHRIS/Proba spectra for the nadir view. LAI, C_{ab}, C_w and $C_{dm}$are fixed to the field measured values. Each panel corresponds to each field.

Figures 7 and 8 show that some fields exhibit some difference between real and modelled spectra in green and NIR regions. The most likely reason of this is discontinuous canopies which can not be detected by given spatial resolution and model assumptions. This exercise potentially can be improved by including field measurements of some parameters which are unavailable in this study such as N and by using different model assumption such as clumping, etc.

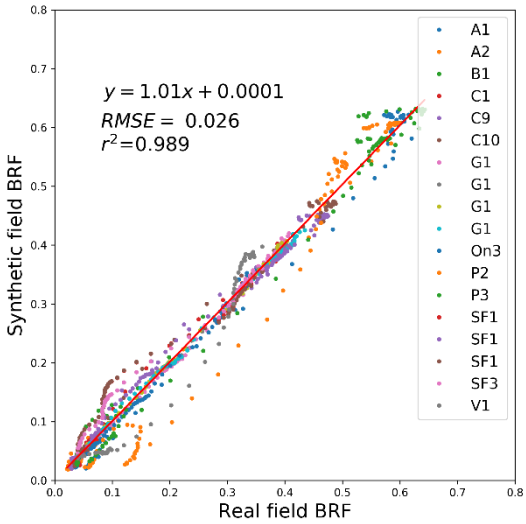


Fig. 8. Scatterplot of real and modelled CHRIS/Proba spectra (400-2500 nm) for the nadir view. LAI, C_{ab}, C_w and $C_{dm}$are fixed to the field measured values. Each field has its own colour code.

3.3 Synthetic data

3.3.1 Generation of the synthetic datasets

Synthetic data represent an 'idealised' test case and allow checking potential of the EO-LDAS parameter retrievals and subsequent reproduction of hyperspectral information.

We simulate the Bidirectional Reflectance Factors (BRFs) of CHRIS/Proba, MISR and Landsat (ETM+) over the 18 SPARC field measurement points. The synthetic data are based on the SPARC database (ESA, 2004) with given information on LAI, C_{ab}, C_w and C_{dm} (Table 1). All other parameters were taken from Lewis et al. (2012) (Table 2). Since each MISR 275x275 m^2 pixel contains approximately 81 Landsat pixels we generate 18 9x9 pixel fields with 10 layers of parameters each and add correlated Gaussian noise but keeping the central pixel with information from Tables 1 and 2. Then each field was averaged to a single value and the number of spectral bands was reduced to those of CHRIS, Landsat and MISR. Finally, we added random noise, which we define as 0.1 times the uncertainty of the real data i.e. [0.00202, 0.00246, 0.00250, 0.00214, 0.00395 and 0.00272] (section 2.2). These numbers are reasonable because they give some noise but keep it low because the aim of the synthetic experiment is to test the best possible (idealised) conditions for retrieval. Increasing of noise can lead to decreasing sensitivity to some of parameters and we will not see potential possibility of retrieval. Input per band uncertainties depend on random noise and are in the range from 0.0002 to 0.002 for Landsat; and from 0.001 to 0.04 for MISR. The resulting synthetic data are similar but not necessarily identical to real spectral information because we

have five parameters, which are not available from the SPARC database (Table 1).

Table 4: Parameters for synthetic data

Field N	field	Canopy height (xh)	leaf radius (xr)	Proportion of senescent material (scen)	Number of leaf layers (N)	Leaf Angle Distribution (LAD)
1	A1	5	0.01	0	1	1
2	A2	5	0.01	0	1	1
3	B1	5	0.01	0	1	5
4	C1	5	0.01	0	1	2
5	C9	5	0.01	0	1	2
6	C10	5	0.01	0	1	2
7	G1	5	0.01	0	1	2
8	G1	5	0.01	0	1	2
9	G1	5	0.01	0	1	2
10	G1	5	0.01	0	1	2
11	On3	5	0.01	0	1	2
12	P2	5	0.01	0	1	5
13	P3	5	0.01	0	1	5
14	SF1	5	0.01	0	1	1
15	SF1	5	0.01	0	1	1
16	SF1	5	0.01	0	1	1
17	SF3	5	0.01	0	1	1
18	V1	5	0.01	0	1	1

3.3.2 Retrieval of biophysical parameters

In this work, we simulate a hyperspectral signal by forward run of the RT model with input parameters retrieved by the model inversion. So it is implied that in order to restore the hyperspectral signal we have to precisely retrieve one or several state parameters.

We solve the problem only for LAI, Cab, Cw and soil coefficients. All other parameters are fixed to values from table 2. Some of the fixed parameters such as N have influence on leaf

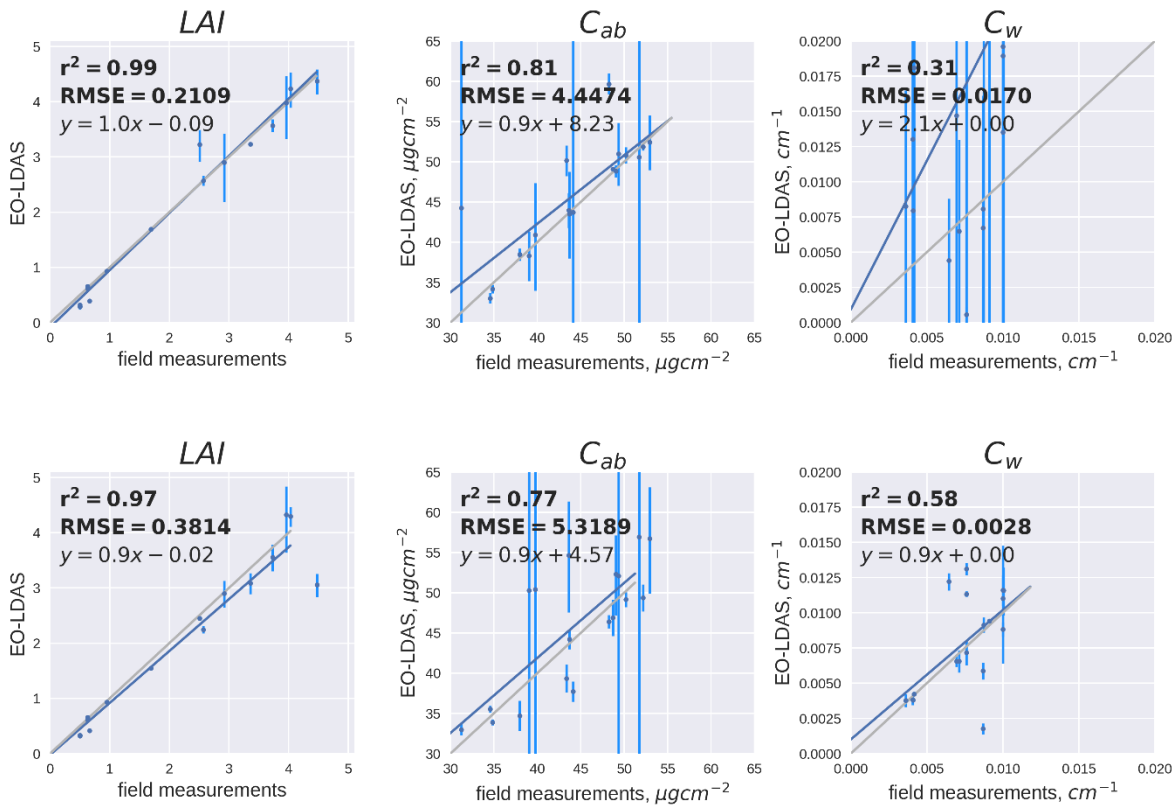
spectra from 400 to 2500. However, this contribution has been ignored.

One of the goals of this work is to test the influence of the MISR cameras combinations on the retrieval of parameters and on the subsequent simulation of hyperspectral signal. We do the retrievals by increasing the number of cameras from one to nine i.e. An camera on the first step, An-Af on the second, An-Af-Aa on the third, etc. This means that at first, we have the nadir-only camera, and then the range of View Zenith Angles (VZAs) is increasing with increasing number of cameras. We assume that bands and cameras are independent and treat each camera as additional J_{obs} term in equation 1.

Fig. 9 displays the results of retrieval of LAI, C_{ab} and C_w by inversion of MISR and Landsat data separately and then MISR+Landsat together. LAI is retrieved with r^2 from 0.97 to 0.99 and RMSE from 0.21 to 0.38 for all considered cases.

The results from LAI retrievals using Landsat only data are similar to the MISR only results. The RMSE in the retrievals essentially spans the entire dynamic range of the field values.

Combining both sensors for LAI results in an inference of LAI that, compared with LAI measurements, is similar to the MISR only case.



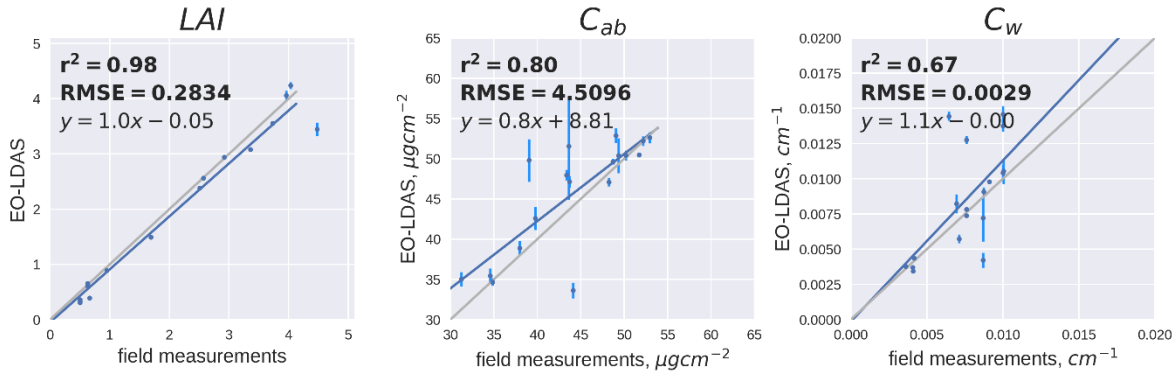


Fig. 9. Comparison of field measurements and EO-LDAS retrievals of LAI, Cab and Cw with synthetic data. MISR only (top panels), Landsat only (middle panels) and MISR+Landsat (bottom panels). MISR nadir camera (An).

Retrieval of C_{ab} is similar for all sensors with lowest r^2 and highest RMSE in case of Landsat-only results which is reflected in the highest RMSE and uncertainties. Results of C_w show worst performance for MISR-only, because MISR does not have sensitivity to water content, as spectral bands sensitive to water absorption are not available. Best C_w results are achieved for Landsat-only and the combination of MISR and Landsat, with nearly the same RMSE and better r^2 for the latter. This might be due to the system reacting to a better inference of LAI, which in turn limits the possible values of $C_{w,}$ resulting in a better estimate of a parameter to which one of the sensors has no sensitivity, as explained above.

Uncertainties are high for C_w in MISR-only case and for C_{ab} in Landsat-only case. However, with the combination of two sensors they are lower for all three parameters.

Results of the retrieval of parameters demonstrate precise retrieval of LAI in all three cases and reasonable retrieval of C_{ab} and C_w in case of combination of two sensors. These results were acquired using nadir-only MISR-camera. Next results show what happens if the number of the cameras is increased.

Fig.10 shows the performance of the retrieval with increasing numbers of MISR cameras. We can see that in case of LAI and C_w there is no distinct linear trend with increasing number of cameras. The exception is nadir-only LAI where RMSE is higher. Lower RMSE for C_w in case of Landsat and MISR+Landsat expected because of sensitivity of Landsat to this parameter. The parameter which is affected by increasing of number of cameras is C_{ab}. We can see decreasing of RMSE for retrieval where from four to nine MISR cameras were used.

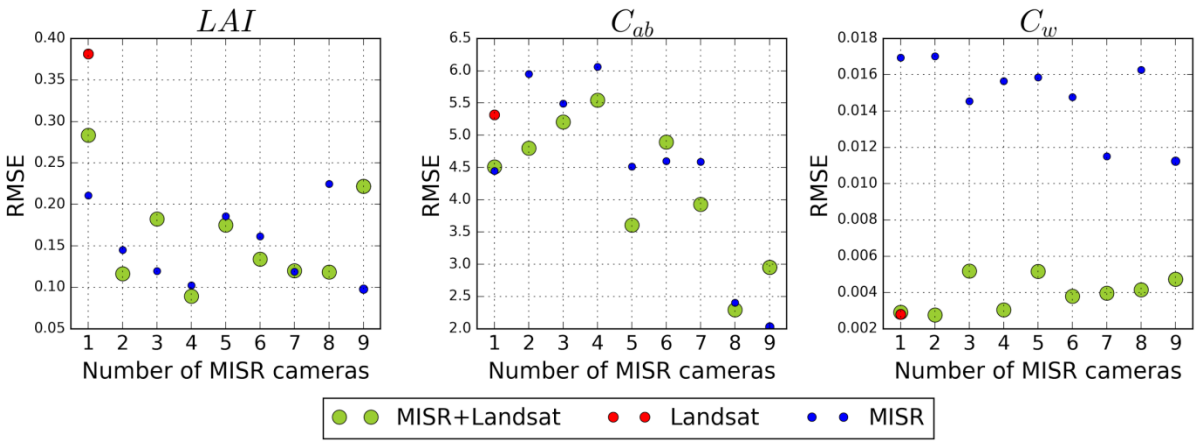


Fig. 10. RMSE of LAI (left hand side panel), Cab (middle panel) and Cw (right hand side panel) for increasing number of MISR cameras.

However, Fig 11 shows decreasing uncertainties with increasing number of cameras for LAI and Cab. Also it shows that combination of sensors have lower uncertainty in case of all three parameters.

LAI and Cab retrievals show that the mean *a posteriori* does not change significantly, but that as more MISR cameras are considered, the uncertainty decreases. This is an automatic consequence of the set-up of the problem, and assuming uncertainty in the bands is independent.

These results are not unexpected: the increase of more observations from MISR cameras results in a shrinking of the posterior uncertainty, as expected. This is further enhanced by adding the extra Landsat data. Generally speaking, the Landsat inversions are closer to the true values, again as expected due to the lower uncertainty in the measurements.

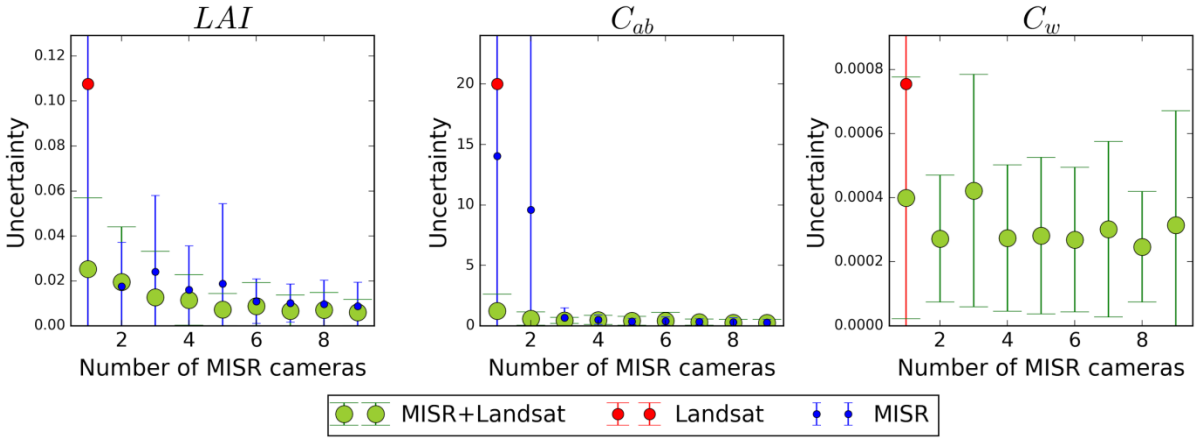


Fig. 11. Averaged uncertainty of LAI (left hand side panel), Cab (middle panel) and Cw (right hand side panel) for increasing number of MISR cameras.

A similar pattern is also present with leaf chlorophyll concentration, although in this case, the

a posteriori mean is less stable. This is expected: the assumption of statistical independence in the observations produces a reduction of the posterior variance in both cases. The sensitivity of the data to LAI is higher than to chlorophyll, as well as its distribution having a higher dynamic range, which accounts for the higher stability of the LAI solution relative to the Cab inversion.

3.3.3 Cross entropy

One of the results of the section 3.3.2 is that with increasing number of MISR cameras posterior uncertainties are decreasing. The possible reason for this is increasing amount of information. In this section, we make an experiment to find out what is the main reason for decreasing of uncertainties: number of viewing directions or number of observations.

We generate seven synthetic nadir MISR observations with different random noise. For the same state parameters (Table 2) we generate 7 multiangular observations. Then we estimate the state parameters and the cross entropy is calculated as

$$H_e = \frac{1}{2} ln(det(C_{prior} H_{hess})) (6)$$

where H_e - cross entropy, C_{prior} - prior covariance, H_{hess} - Hessian matrix.

Hessian is a matrix of partial second derivatives, calculated at the minimum of the solution space or maximum *a posteriori* (MAP) point. The Hessian matrix describes the radius of curvature around MAP point, and in the case of Gaussian distributions, it is equivalent to the inverse *a posteriori* covariance matrix, and as such encodes the *a poteriori* uncertainty (e.g. variance/covariance structure) in the retrieved parameters.

Under the assumption of a Gaussian *a priori* pdf for the parameters, Cprior encodes the *a priori* covariance between parameters. In this contribution, we assume Cprior to be diagonal, and thus only define *a priori* variances and not covariances. The cross-entropy He quantifies the "information gain" of the system going from prior to posterior distribution, or in other words, it quantifies how the system interprets the evidence from the observations in terms of parameters.

Fig. 12 demonstrates that with increasing number of observations, the amount of information is increasing. However, it is nearly the same for both nadir-only data and multiangular data until the third camera, and then fluctuates from fourth to seventh cameras.

Note that in nadir only case in the sequence of adding of cameras the cross entropy is increasing but values of the retrieved parameters remain the same.

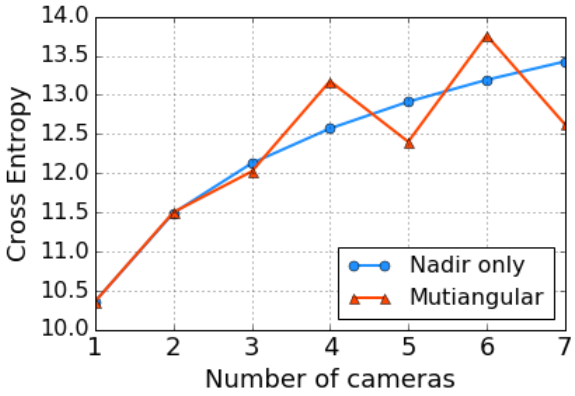


Fig. 12. Changing cross entropy with increasing number of the MISR cameras: 7 nadir only cameras, 7 multiangular observations from An to C MISR cameras.

This example shows that, at least in case of this angular/spectral configuration, increasing number of observations has nearly the same effect as increasing number of different viewing directions. This is in agreement with results of the previous section where we can see that increasing number of MISR cameras leads to a decrease of posterior parameter uncertainties. As above, this is a consequence of the assumptions of independent bands. In practice, this is an important issue, as when EO data has been atmospherically corrected, it is likely that strong correlations in the uncertainty appear in different bands.

3.3.4 Simulation of the CHRIS/PROBA signal

This section investigates possibility to restore hyperspectral signal based on the results from section 3.3.2.

CHRIS/Proba reflectance is retrieved by forward run of the RT model for each combination of the MISR cameras and for each CHRIS/Proba 'fly-by' positions.

Fig. 13 shows one example of direct comparison between synthetic CHRIS/PROBA and retrieved reflectance spectra for Alfalfa crop field for the nadir camera (An) when using only An camera (left hand side panel) and nine cameras (right side panel). We can see that there is no distinct RMSE trend with increasing number of cameras which is in line with results obtained for retrieval of the state parameters and cross-entropy.

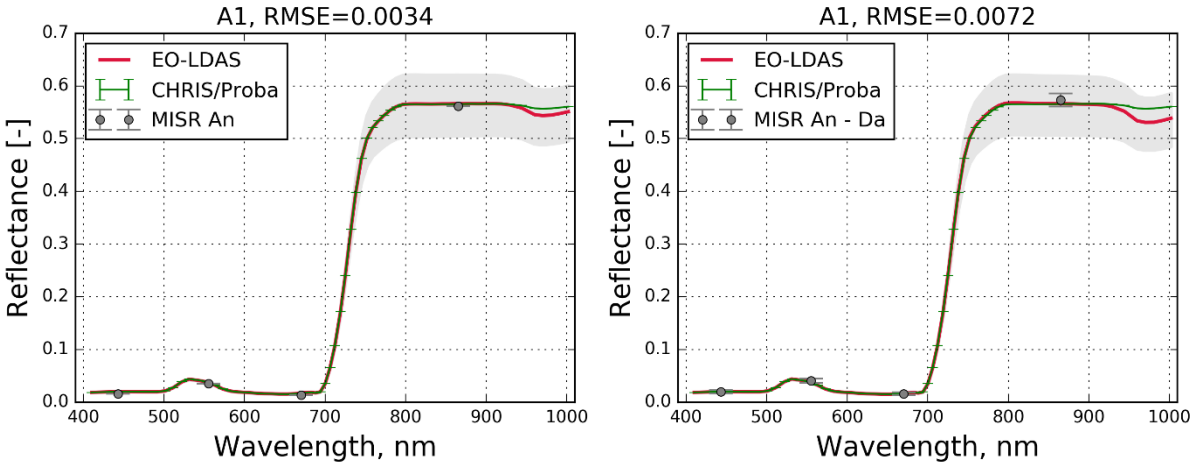


Fig. 13. Comparison of CHRIS synthetic (green line) for nadir view and MISR-only retrieved reflectance spectra (red line) over the Alfalfa field. Left hand side panel is for An (nadir) camera only, right hand side panel for nine cameras. Grey surface indicates the posterior uncertainties in 95% credible interval.

Fig 14 and 15 show that with given set of retrieved parameters, synthetic CHRIS/PROBA spectra can be simulated very close to unity with both Landsat and MISR+Landsat.

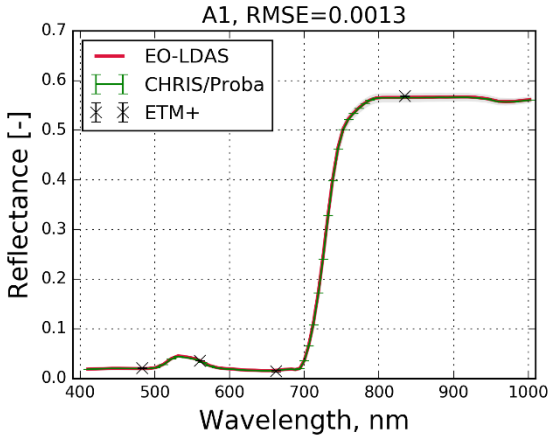


Fig. 14. Comparison of CHRIS synthetic (green line) and Landsat-only retrieved reflectance spectra (red line) over the Alfalfa field. Grey surface indicates the posterior uncertainties in 95% credible interval.

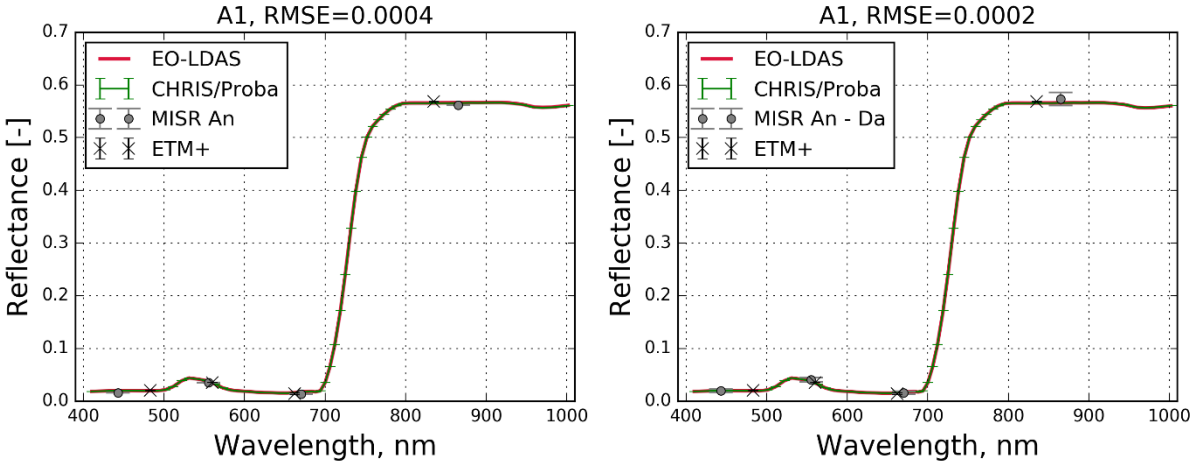


Fig. 15. Comparison of CHRIS synthetic (green line) for nadir view and MISR+Landsat retrieved reflectance

spectra (red line) over the Alfalfa field. Left hand side panel is for An (nadir) camera only, right hand side panel for nine cameras. Grey surface indicates the posterior uncertainties in 95% credible interval.

Fig. 16 summarizes the comparison of the 'true' and retrieved hyperspectral data over the 18 fields with synthetic MISR data only. Here we show the nine cameras case because one of the results of section 3.3.2 is that using nine cameras provides lower uncertainties. In the left-hand column, the results of the simulation of CHRIS/Proba camera VZA 8.4° are shown in terms of RMSE. The right panel corresponds to VZA -55.2°. The lines correspond to the 18 Barrax fields. We show result only for two CHRIS/Proba 'fly-by' positions because no big difference was found for other three directions. Note that the Z scale is logarithmic.

Fig. 16 shows low squared difference for all fields with slightly higher values in NIR region for fields 1-4, 6, 9-11 and 18. However taking into account that highest squared difference varies from 10^{-3} to 10^{-4} the error is very low. RMSE is in the range from 0.0002 to 0.04.

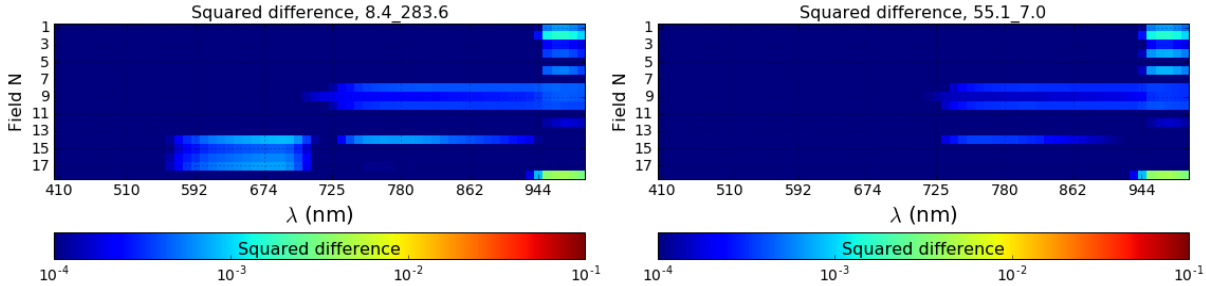


Fig. 16. Squared difference between synthetic CHRIS-Proba and MISR-only (nine cameras) retrieval over 18 fields for CHRIS-Proba camera VZA 8.4° (left hand side panel) and VZA -55.2° (right hand side panel).

Fig 17 displays the Landsat only case. Landsat/ETM+ has three additional bands in the SWIR region which are outside the CHRIS/Proba spectral range but due to better constraining of retrieval of parameters there is a noticeable improvement for fields 7-9 and 15-18. RMSE varies in the range from 0.001 to 0.01.

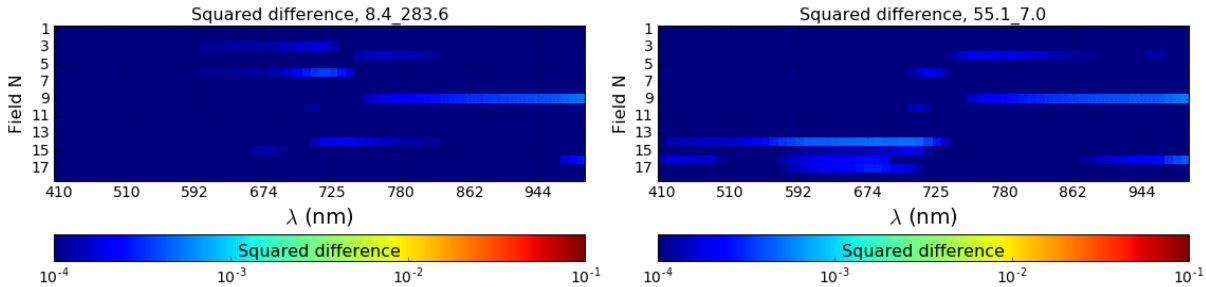


Fig. 17. Squared difference between synthetic CHRIS-Proba and Landsat-only retrieval over 18 fields for CHRIS-Proba camera VZA 8.4° (left hand side panel) and VZA -55.2° (right hand side panel).

The MISR+Landsat case and has slightly lower RMSE range for all simulated CHRIS/Proba

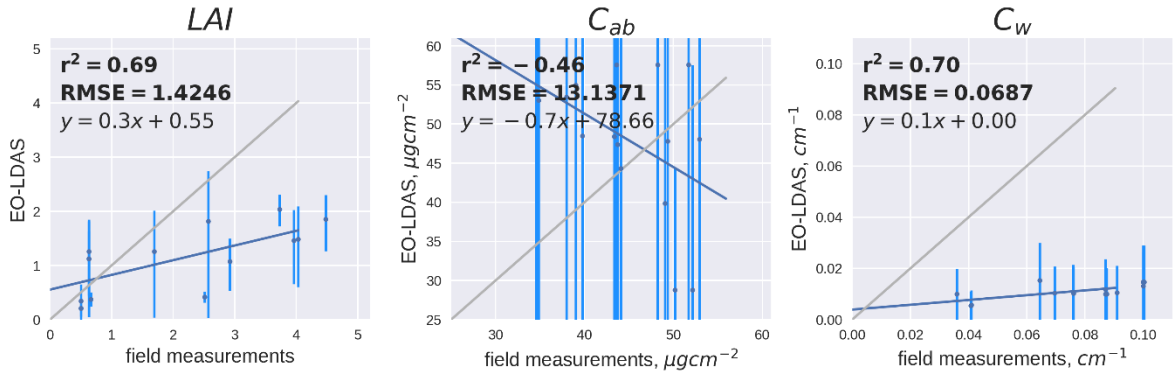
cameras: from 0.0001 to 0.01. However, in general this result is very similar to Landsat-only. This tells that Landsat data dominates the retrieval process and that even if spectral information content of Landsat is not enough for precise retrieval of all parameters (Fig. 7-9), it is enough for precise simulation of a hyperspectral sensor.

3.4 Real Data

Real data experiment in general follows the same steps as synthetic experiment. I.e. we use CHRIS/PROBA as a benchmark for simulated hyperspectral data which arebased on forward run of the model using parameters obtained by MISR and Landsat data inversion. Inversion is doneby MISR-only data, Landsat-only and with merged MISR+Landsat. We use soil model described in section 3.1 as a prior.

3.4.1 Retrieval of biophysical parameters

Fig 18 shows that retrieval of LAI based on the MISR-only information underestimates higher LAI values, resulting in bias, lower r^2 (0.69) and high RMSE (1.42). In case of Landsat and MISR+Landsat LAI retrieval has better results with slightly lower RMSE and uncertainties for the blended sensors. Retrieval of C_{ab} and C_w does not have meaningful accuracy and mostly show results of the prior minimization. We can see it because many of the C_{ab} values for Landsat and MISR+Landsat cases equal to 29 µg/cm^2.



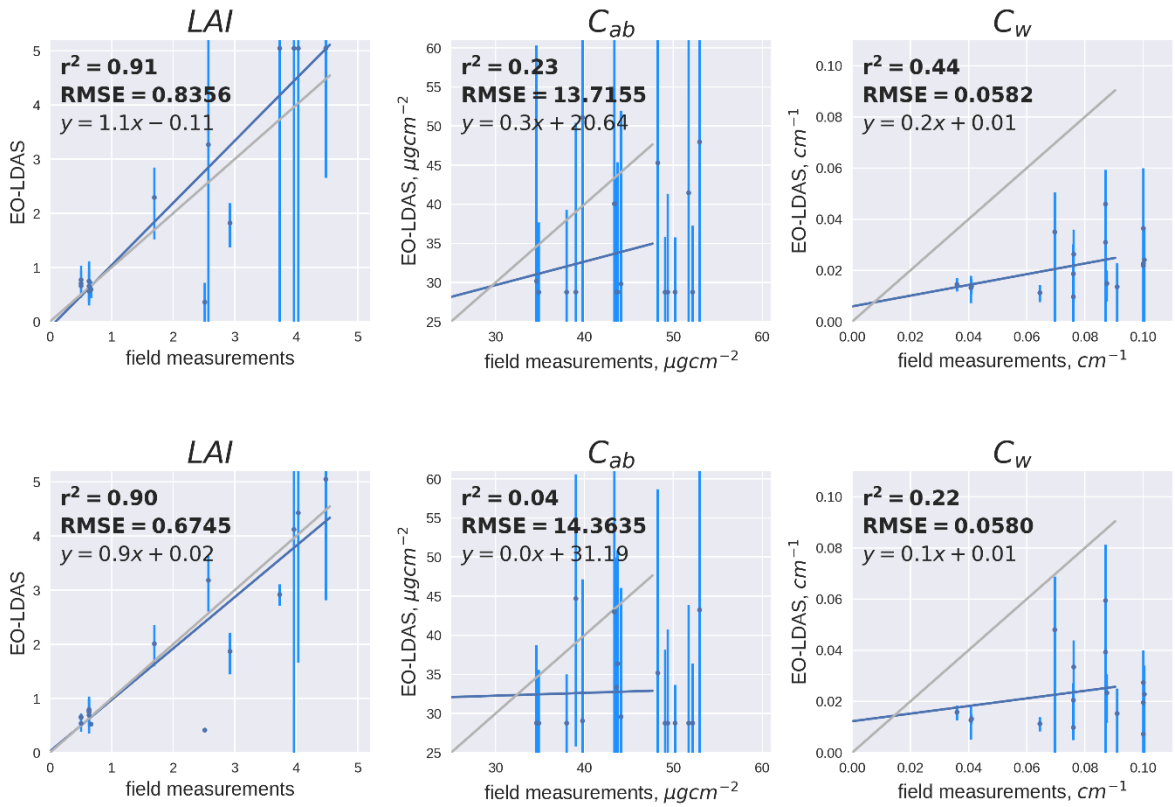


Fig. 18. Comparison of field measurements and EO-LDAS retrievals of LAI, C_{ab} and C_w with real data. MISR only (top panels), Landsat only (middle panels) and MISR+Landsat (bottom panels). MISR nadir camera (An).

Fig. 18 shows that only retrieval of LAI is acceptable because for C_{ab} and C_w r^2 is low and RMSE is high. When we compare the LAI retrieved from Landsat data alone and then MISR+Landsat data, with the "true" LAI the correlation is high and the slope is close to unity. The RMSE (Landsat, Landsat+MISR) is around 0.8 LAI units or lower, which is in line with a number of global products (Garrigues et. al. 2008). In Landsat-onlycase, the results of LAI show that the retrieved LAI tends to overestimate the true LAI for high LAI values, which is explained by the saturation of reflectance for high LAI, and thus is a reduced sensitivity effect. The retrieval of C_{ab} and C_w, suggesting a low sensitivity of the real observations to these parameters. Discrepancy between results of this section and synthetic data section is due to idealized synthetic datasets which are aimed to show potential retrieval of parameters.

3.4.2 Simulation of the CHRIS/Proba signal

Results of the section 3.4.1 show that among three retrieved parameters (LAI, C_w and C_{ab}) only the retrieved LAI error is less than 20%. This means that with given data we know 'correct' values only for three from ten parameters (LAI, s_1 and s_2). The goal of this section is to test whether it is possible to simulate CHRIS/Proba signal if the only known parameters are

LAI and the soil coefficients. If so, this would demonstrate the power and flexibility of the reconstruction process.

Fig. 19illustrates simulation of hyperspectral information using MISR-only data over the Alfalfa A1 field. RMSE between actual and reconstructed CHRIS/Proba spectra ranges from 0.0920 to 0.0985. There is no distinct trend of RMSE/uncertainty with varying the number of MISR cameras.

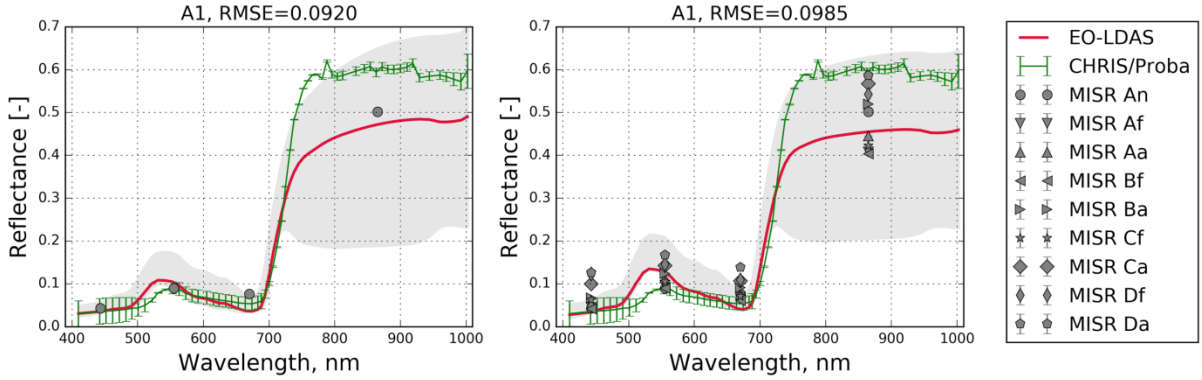


Fig. 19. Comparison of CHRIS actual (green line) for NADIR view and retrieved reflectance spectra (red line) over the Alfalfa field. Retrieval was done using actual MISR data only by increasing the number of cameras from one to nine. Left hand side panel is for An (nadir) camera only, right hand side panel for nine cameras. Grey surface indicates the posterior uncertainties.

Unlike the synthetic simulation there is a relatively high discrepancy between simulated and real spectra in the NIR region. The reason for this maybe the much lower spatial resolution of MISR that results in spectrally mixed pixels. This probably means that the procedure with equation (2) cannot completely separate the spectra. However, this was expected since equation (2) is only a first guess assumption.

Fig. 20 shows results for Landsat-only case. RMSE is lower mainly due to better correspondence in NIR region. This can be explained by higher spatial resolution and better spectral coverage. However, some discrepancy in NIR and red regions is still noticeable.

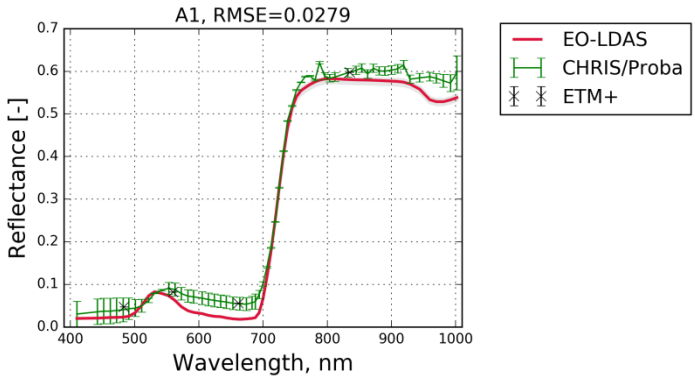


Fig. 20. Comparison of CHRIS actual (green line) for NADIR view and retrieved reflectance spectra (red line)

over the Alfalfa field. Retrieval was done using actual Landsat-only data. Grey surface indicates the posterior uncertainties.

The next figure illustrates combination of real Landsat and MISR data. Fig. 21 shows that when a combination of the two sensors is used, MISR pulls NIR part of the curve down resulting in lower RMSE. The more cameras we add the more deviation from the actual CHRIS/Proba. However, we can notice better correspondence in visible part. The possible reason for higher error in NIR part is unknown real per band uncertainties.

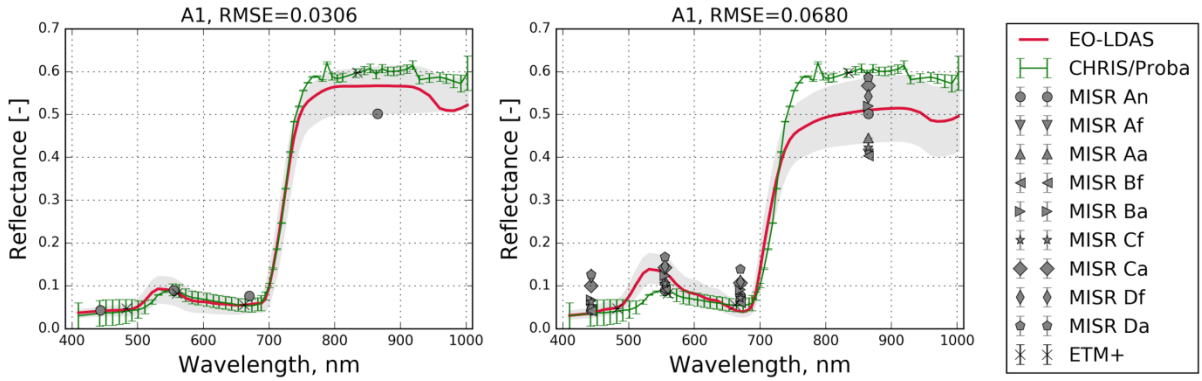


Fig. 21. Comparison of CHRIS actual (green line) for NADIR view and retrieved reflectance spectra (red line) over the Alfalfa field. Retrieval was done using actual MISR data and Landsat by increasing the number of cameras from one to nine. Left panel is for An (nadir) camera only, right panel for nine cameras. Grey surface indicates the posterior uncertainties.

Fig. 22 summarizes results based on the real MISR-only data over 18 fields. Results for fields with higher LAI (Table 1) exhibit higher squared difference especially in the NIR region. Meanwhile RMSE for fields with lower LAI has lower values. Low squared difference is especially noticeable for the garlic fields 7-9. This result is expected since garlic fields have LAI of 0.63 leading to a very strong signal from soil, which is well described (section 3.1). Difference between simulations of different CHRIS/Proba cameras is minimal except higher difference is red region for the garlic fields. RMSE is in the range from 0.007 to 0.2 where the lowest RMSE belongs to field seven (garlic).

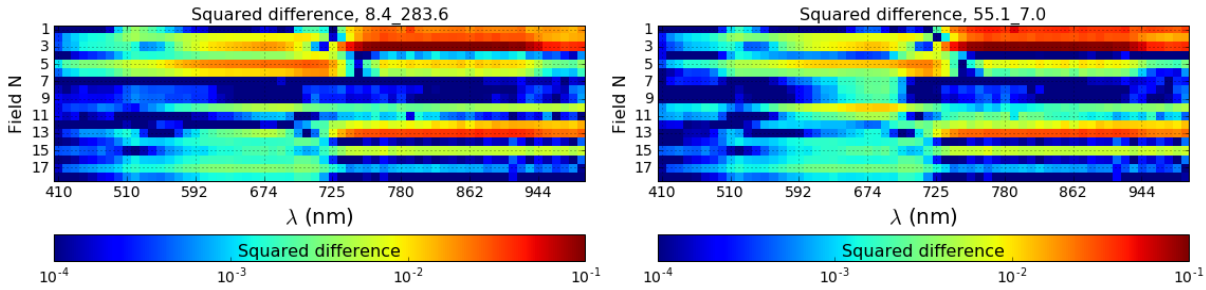


Fig. 22. Squared difference between real CHRIS-Proba and MISR-only (nine cameras) retrieval over 18 fields for CHRIS-Proba camera VZA 8.4° (left hand side panel) and VZA -55.2° (right hand side).

Fig. 23 displays results of hyperspectral simulations using Landsat-only data. We can see that

in general, values of squared difference are lower for all fields, corresponding to the difference between Fig. 19 and Fig. 20. However, the difference is slightly higher in NIR region in case of low LAI (fields 7-9). RMSE is in the range from 0.01 to 0.1.

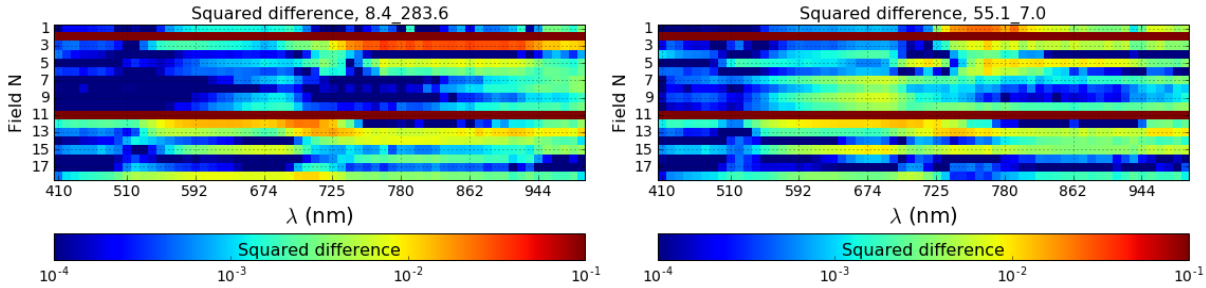


Fig. 23. Squared difference between real CHRIS-Proba and Landsat-only retrieval over 18 fields for CHRIS-Proba camera VZA 8.4° (left hand side panel) and VZA -55.2° (right hand side).

Fig 24 shows squared difference for MISR+Landsat. These results correspond to Fig 21 i.e. small difference from the Landsat-only results. However, unlike the difference between Fig. 20 and Fig. 21 there is no higher difference in NIR region except for field A1. RMSE is in the range from 0.02 to 0.1. This tells that in blending the two sensors and using nine MISR cameras, Landsat plays the leading role. However, with adding more cameras results can deviate from Landsat-only decisions according to the results shown on Fig 21. One other thing to note is that for both Fig 23 and Fig 24 the differences for CHRIS/Proba VZA 55.2° have slightly higher values in red region for garlic fields.

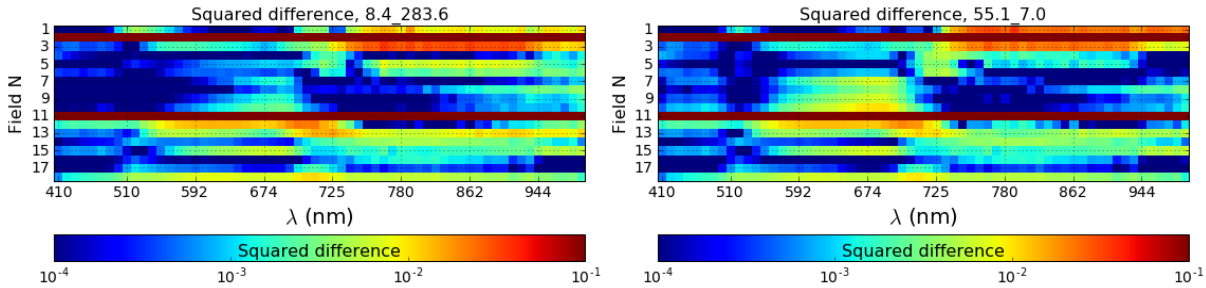


Fig. 24. Squared difference between real CHRIS-Proba and MISR+Landsat (nine cameras) retrieval over 18 fields for CHRIS-Proba camera VZA 8.4° (left hand side panel) and VZA -55.2° (right hand side).

Note that the fields 2 and 11 on Fig 23 and 24 have no data because they correspond to SLC-OFF pixels.

3.5 On uncertainty estimation

Realistic uncertainty is one of the key factors to combining observations from different sensors. We can do an experiment assuming that uncertainties of the each MISR band is nine

times higher than Landsat ones. This setup makes a balance between two sensors when we use all nine MISR cameras. Results are displayed on Fig 25. We can see that if we have only one MISR camera the result is almost the same as for Landsat-only case (Fig 20) but with five cameras RMSE is lower than on both Fig 20 and Fig 21. Due to the fact that uncertainties of each MISR band are much lower than Landsat ones, adding more than one MISR cameras does not pull simulated Proba-curve down in NIR region like on Fig 25.

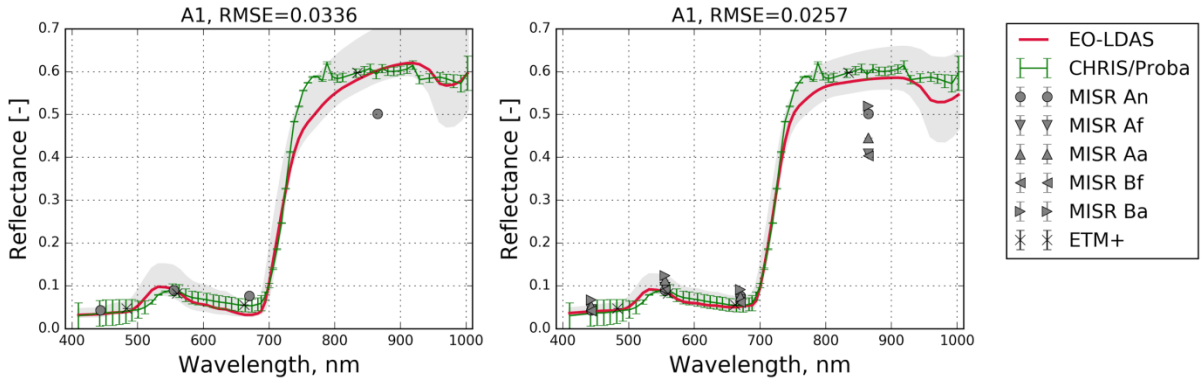


Fig. 25. Comparison of CHRIS real (green line) for nadir view and MISR+Landsat retrieved reflectance spectra (red line) over the Alfalfa field. Left hand side panel is for An (nadir) camera only, right hand side panel for nine cameras. Grey surface indicates the posterior uncertainties in 95% credible interval.

3.6 An experiment with fixed LAI

We can make an experiment when we fix LAI to the ground truth value and s_1/s_2 to their prior values. Then we use LHS to generate 100 random LUTs with seven unknown parameters with 7000 entries in each table. We then find best fit to the real CHRIS/Proba data for each of these tables. Results show low RMSE in the range between 0.004 and 0.02. Here we show an example of Alfalfa and garlic fields (Fig 26 and Table 5 and 6). Tables 5-6 show that a broad variation in land surface parameters still results in a prediction very much in agreement with the observations. We can see it as a depiction of ill-posedness, where there is no unique solution, but rather a wide range of values fit the observations equally well. However, taking into account that LAI and soil parameters are fixed we can say that this is in line with Mousivand et al. 2015 and shows that in case of this data the number of the main controlling factors are even less and they are LAI and s_1/s_2.

Fig 26 displays spectra with minimum and maximum RMSE for the Alfalfa and Garlic fields. We can see that even in the case of maximum RMSE (0.018), the resulting simulated spectral curve is quite similar to real CHRIS/Proba data.

Table 5: min and max values for fit with fixed LAI and soil model for the Alfalfa field

	xhc (m)	rpl (m)	xkab ()	scen	xkw	xkm	xleafn
min value	0.011	0.038	25.486	0.003	0.000	0.001	0.619
max value	4.843	0.999	118.157	0.303	0.055	0.024	2.999

Table 6: min and max values for fit with fixed LAI and soil model for the Garlic field

	xhc (m)	rpl (m)	xkab ()	scen	xkw	xkm	xleafn
min value	0.000	0.101	21.724	0.003	0.000	0.000	0.038
max value	2.131	0.999	84.602	0.683	0.059	0.020	2.946

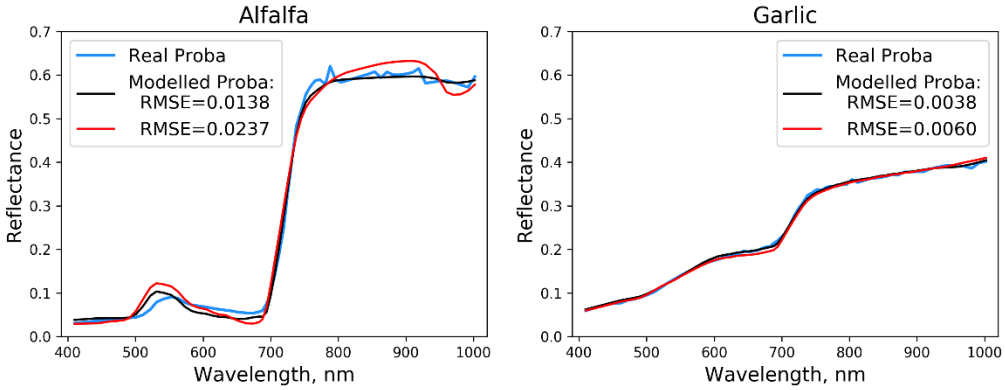


Fig. 26. Simulation of CHRIS/Proba spectra with fixed LAI and s1/s2 to the 'true' values. Left hand side panel - Alfalfa A1 field, right hand side panel - Garlic G1 field. Black lines represent spectra with minimum RMSE, red lines - maximum RMSE.

4 Discussion and Conclusion

We propose a way of using the EO-LDAS data assimilation system to reproduce hyperspectral information (effectively arbitrary band combinations) with multi-sensor and multi-angular constraints, using a restricted number of known biophysical parameters derived from by multi-spectral sensors. We have demonstrated the performance of EO-LDAS on MISR and Landsat sensors but the method can be extended to other optical multispectral sensors. We have shown that EO-LDAS provides an optimal way of combining spectral and angular observations for a two-sensor combination. It was found that increased amount of multiangular observations (provided by MISR) only helps to improve accuracy (RMSE) in the case of the synthetic simulations, and only then for C_{ab}. Calculation of the cross entropy for increasing numbers of the MISR cameras does not show that multiangular observations always provide more information than the same number of mono-angular nadir observations

(Fig. 12). This suggests that, at least in case of this experimental setup, the number of observations is of greater significance than different viewing angle positions.

One important note is that per-band uncertainties for both MISR and Landsat surface directional reflectances are not provided. Estimation of these uncertainties will require propagation of the sensor thermal noise, as well as atmospheric composition, through the atmospheric correction component, not a trivial task, and outside the scope of this study.

However, it should also be kept in mind that these are not "real" uncertainties, but is an approximation to the linear part of the Hessian, with higher order terms ignored (Lewis et al, 2012). Additionally, it is assumed that the posterior is Gaussian, and that there is no model error. It can be seen as estimation of steepness of curvature in the decision space, which measures how many possible solutions we have at the end of optimization process. It is not always related to "real" uncertainty, which is very hard to estimate. For example, the process of optimization can be trapped in local minima and in this case posterior uncertainty reflects the situation around this local minimum. However, in many cases posterior uncertainty does reflect the estimation of accuracy, as shown for example by Thacker (1989).

From the above, it is clear that an accurate description of the statistical properties of the inputs (here, surface directional reflectance), is required to produce an accurate estimate of the posterior distribution. It is also important to consider potential biases between observations from different sensors (although the biases can be considered part of the observational uncertainty).

The example from section 3.5 is an attempt to show that realistic prior uncertainties is an essential factor in merging different sources of information and potentially better results can be achieved if EO data providers can include information about uncertainties to reflectance products.

The real data results suggest that proper estimation of LAI and soil parameters could be enough for simulation of hyperspectral signal between 400 and 1000 nm with RMSE<=0.03. This result is in line with Verrelstet. al. (2016) and Mousivandet. al. (2015), who both show that LAI has a stronger influence in all bands than most of other parameters. Indeed, the example from the section 3.6 supports the hypothesis that most of variations of hyperspectral data in visible and infrared ranges are controlled by just a few mechanisms. In case of the data of this work it is mainly LAI and soil parameters.

These comparisons show that EO-LDAS can be used for simulation of hyperspectral data by observations (retrievals) from a multispectral sensor or by a combination of multispectral sensors. In addition, results show that accuracy of retrieved LAI and accuracy of simulated hyperspectral information are consistent. Finally, results demonstrate consistent and physically based technique for merging information from two or more multispectral sensors.

The simulation of one sensor signal by another sensor, which we have demonstrate here, is likely to be useful for the intercalibration of small EO satellites. 92 small satellites were launched in 2013, more than 140 in 2014 and more than 500 expected in 2015-2019 (Messier, 2015). Many of these satellites are designed for EO and can provide quite high spatial resolution data from 10 to 30 m. Large numbers of these satellites can provide both high revisiting rate and fine resolution. However, to combine information from all these sensors they have to be properly intercalibrated. The method we present here is a possibility to intercalibrate small EO satellites in order to get continuous time series of surface reflectance, across programmes and sensors of different designs.

An important issue, which we do not consider here, is how to propagate *a priori* uncertainties through the pre-processing chain: i.e. calibration, georeferencing, atmospheric correction, dependency on spatial resolution etc. So further efforts are required in this area because this is likely to provide a better balance between the cost function terms: models, radiometric information and prior information and can significantly improve results.

Acknowledgements

Authors gratefully acknowledges financial support from the BACI project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640176. We gratefully acknowledge the GIONET project, funded by the European Commission, Marie Curie Programme Initial Training Network, Grant Agreement number PITN-GA-2010- 264509. We also acknowledge MULTIPLY project (Project No. 687320). We acknowledge 40th COSPAR organizing committee for travel financial support; U.S. Geological Survey (USGS) and National Center for Earth Resource Observations and Science (EROS) for providing Landsat data through the Earth Explorer system (http://earthexplorer.usgs.gov/). JRC authors thank Linda Hunt and Michel Verstraete for their technical support on MISR HR processing chain. This research would be impossible without work of JPL/NASA who created and launched into space MISR instrument.

J Gomez-Dans acknowledges support from the NERC NCEO, European Space Agency support under Contract 4000112388/14/I-NB, and the European Union's Horizon 2020 research and innovation programme under grant agreement No 687320 for the EU H2020

MULTIPLY project.

M. Disney acknowledges support from the NERC NCEO, NERC Standard Grants NE/N00373X/1 and NE/ P011780/1 and the European Union's Horizon 2020 research and innovation programme under grant agreement No 640176 for the EU H2020 BACI project.

References

Asner, G. P., Townsend, P., Martin, R. E. and Chadwick, K. D. Forest Biophysical and Biochemical Properties from Hyperspectral and LiDAR Remote Sensing. In *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* (pp. 429–448), 2016.

Baccini, A., Friedl, M.A., Woodcock, C.E. and Zhu, Z. Scaling Field Data to Calibrate and Validate Moderate Spatial Resolution Remote Sensing Models, Photogrammetric Engineering & Remote Sensing 73(8) p. 945 - 954, 2007.

Baret, F., Morissette, J. T., Fernandes, R. A., Champeaux, J. L., Myneni, R. B., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S. and Nickeson, J. E.: Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP, IEEE Trans. Geosci. Remote Sens., 44(7), 1794–1803, 2006.

Barnsley, M. J., Settle, J. J., Cutter, M., Lobb, D. and Teston, F. The PROBA/CHRIS mission: A low-cost smallsat for hyperspectral multiangle observations of the earth surface and atmosphere, IEEE Transactions on Geoscience and Remote Sensing, 42(7), 1512−1520, 2004.

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A. and Zemp, M.: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. Bull. Amer. Meteor. Soc., 95, 1431–1443, 2014.

Bunnik, N.J.J. The multispectral reflectance of shortwave radiation of agricultural crops in relation with their morphological and optical properties, technical report, MededelingenLandbouwhogeschool, Wageningen, Netherlands, 1978.

Candela, L., Formaro, R., Guarini, R., Loizzo, R., Longo, F. and Varacalli, G. The PRISMA Mission, Proceedings of the IEEE IGARSS (International Geoscience and Remote Sensing Symposium) Conference, Beijing, China, July 10-15, 2016.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes,

R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y. and Yang, W. Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, Journal of Geophysical Research: Biogeosciences, 113, G2, 2156-2202, 2008.

Chernetskiy, M., Gomez-Dans, J., Gobron, N., Morgan, O., Lewis, P., Truckenbrodt, S. and Schmullius, C. Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing, 9(7), 656, 2017.

Combal, B., Baret, F, Weiss, M., Trubuil, A., Macé D, Pragnère A, et al. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. Remote Sens Environ, 84(1), 1–15, 2003.

Diner, D.J., Beckert, J.C., Reilly, T.H., Bruegge, C.J., Conel, J.E., Kahn, R.A., Martonchik, J. V, Ackerman, T.P., Davies, R., Gerstl, S.A.W., Gordon, H.R., Muller, J.-P., Myneni, R.B., Sellers, P.J., Pinty, B. and Verstraete, M.M. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, Geoscience and Remote Sensing, IEEE Transactions on, 36(4), 1072–1087, 1998.

Disney M, Muller J-P, Kharbouche S., Kaminski T., Voßbeck M., Lewis P., et al. A New Global fAPAR and LAI Dataset Derived from Optimal Albedo Estimates: Comparison with MODIS Products. Remote Sensing, 8, 275, 2016.

ESA: SPARC Data Acquisition Report., Contract no: 18307/04/NL/FF, University of Valencia [Online], Available: http://earth.esa.int/campaigns/, 2004.

Feret, J.B., Francois, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P., et al. (2008) PROSPECT-4 and 5: advances in the leaf optical properties model separating photosynthetic pigments. Remote Sensing of Environment. 112(6): 3030-3043, 2008.

Gandia, S., Fernández, G., García, J.C. and Moreno, J. Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign, In Proc. 2nd CHRIS/Proba Workshop, ESA/ESRIN, Frascati, Italy, 2004.

Gobron, N., Pinty, B., Verstraete, M.M. and Govaerts, Y. A semi-discrete model for the scattering of light by vegetation, Journal of Geophysical Research, 102, 9431–9446, 1997

Gobron, N., Pinty, B., Verstraete, M.M., Martonchik, J. V., Knyazikhin, Y. and Diner, D.J. Potential of multiangular spectral measurements to characterize land surfaces: Conceptual

approach and exploratory application, Journal of Geophysical Research, 105(D13), 17539, 2000.

Gomez-Dans, J. and Lewis, P. Earth Observation Land Data Assimilation System (EO-LDAS), Available online: https://github.com/jgomezdans/eoldas, 2012.

Gomez-Dans, J., Lewis, P., & Disney, M.: Efficient Emulation of Radiative Transfer Codes Using Gaussian Processes and Application to Land Surface Parameter Inferences. Remote Sensing, 8(2), 119,2016Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M., Mühle, H., Müller, R., Habermeyer, M., Ohndorf, A., Hill, J., Buddenbaum, H., Hostert, P., van der Linden, S., Leitão, P., Rabe, A., Doerffer, R., Krasemann, H., Xi, H., Mauser, W., Hank, T., Locherer, M., Rast, M., Staenz, K., Sang, B.: The EnMAPSpaceborne Imaging Spectroscopy Mission for Earth Observation. Remote Sens. 7(7), 8830-8857, 2015.

Gupta, R. K., Prasad, S., and P.V., K. R. Evaluation of spatial upscaling algorithms for different land cover types. Adv. Space Res., 22(5), 625–628, 1998

Hay, G. J., Niemann, K. and Goodenough, D. G. Spatial Thresholds, Image-Objects, and Upscaling: A Multiscale Evaluation. Remote Sensing of Environment, 62, 1–19,1997.

Hay, G. J., Marceau, D. J. and Dub, P. A multiscale framework for landscape analysis: Object-specific analysis and upscaling. Landscape Ecology, 16, 471–490, 2001.

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., and Wagner, W.: The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. Bull. Amer. Meteor. Soc., 94, 1541–1552, 2013.

Jacquemoud, S. and Baret, F. PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment, 34(2), 75–91, 1990.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., Ustin, S. L.: PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, *113*, S56–S66, 2009.

Kimes D.S., Knyazikhin Y., Privette J.L., Abuelgasim A.A. and Gao F. Inversion methods for

physically-based models. Remote Sens Rev, 18, 381–439, 2000.

Knyazikhin, Y., Martonchik, J. V., Myneni, R.B., Diner, D.J. and Running, S.W. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data, Journal of Geophysical Research, 103(D24), 32257, 1998.

Koetz B., Baret F., Poilvé H. and Hill J. Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics. Remote Sens Environ, 95, 115–124, 2005.

Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D. J., Middleton, E. M.: An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities, Remote Sensing of Environment, 167, 6-19, 2015.

Lewis, P., Gómez-Dans, J., Kaminski, T., Settle, J., Quaife, T., Gobron, N., Styles, J. and Berger, M.: An Earth Observation Land Data Assimilation System (EO-LDAS), Remote Sensing of Environment, 120, 219–235, 2012.

Maiersperger, T.K., Scaramuzza, P.L., Leigh, L., Shrestha, S., Gallo, K.P., Jenkerson, C.B. and Dwyer, J.L.: Characterizing LEDAPS surface reflectance products by comparisons with AERONET, field spectrometer, and MODIS data. Remote Sensing of Environment, 136, 1–13, 2013.

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J. and Lim, T.K.: LEDAPS Landsat Calibration, Reflectance, Atmospheric Correction Preprocessing Code. Model product. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. http://dx.doi.or, 2012

Messier, D. Euroconsult Sees Large Market for Smallsats,*Parabolic Arc*. Retrieved 8 March 2015.

Mousivand, A., Menenti, M., Gorte, B., Verhoef, W.: Multi-temporal, multi-sensor retrieval of terrestrial vegetation properties from spectral–directional radiometric data, Remote Sensing of Environment, 158, 311-330, 2015.

Musavi, T., Migliavacca, M., van de Weg, M. J., Kattge, J., Wohlfahrt, G., van Bodegom, P.

M., Reichstein, M., Bahn, M., Carrara, A., Domingues, T. F., Gavazzi, M., Gianelle, D., Gimeno, C., Granier, A., Gruening, C., Havránková, K., Herbst, M., Hrynkiw, C., Kalhori, A., Kaminski, T., Klumpp, K., Kolari, P., Longdoz, B., Minerbi, S., Montagnani, L., Moors, E., Oechel, W. C., Reich, P. B., Rohatyn, S., Rossi, A., Rotenberg, E., Varlagin, A., Wilkinson, M., Wirth, C. and Mahecha, M. D.: Potential and limitations of inferring ecosystem photosynthetic capacity from leaf functional traits.EcolEvol, 6, 7352–7366, 2016.

Pinty, B., Widlowski, J., Gobron, N., Verstraete, M.M., Diner, D.J. and Member, A.: Uniqueness of Multiangular Measurements — Part I : An Indicator of Subpixel Surface Heterogeneity From MISR, IEEE Transactions on Geoscience and Remote Sensing, 40(7), 1560–1573. 2002.

Porcar-Castell A, Mac Arthur A, Rossini M, Eklundh L, Pacheco-Labrador J, Anderson K, et al.: EUROSPEC: at the interface between remote-sensing and ecosystem CO_2 flux measurements in Europe, Biogeosciences, 12, 6103–6124. 2015.

Pfeifer M, Disney M, Quaife T, Marchant R.: Terrestrial ecosystems from space: a review of earth observation products for macroecology applications. Glob EcolBiogeogr, 21, 603–624 2012.

Price, J. C.: On the information content of soil reflectance spectra. Remote Sensing of Environment, 113–121, 1990.

Price, J. C.: Band selection procedure for multispectral scanners. Applied Optics, 33(15), 3281–8, 1994.

Price, J. C.: An Approach for Analysis of Reflectance Spectra. Remote Sensing of Environment, 64(3), 316–330, 1998.

Roelofsen, H. D., van Bodegom, P. M., Kooistra, L., & Witte, J.-P. M.: Predicting leaf traits of herbaceous species from their spectral characteristics. Ecology and Evolution, 4(6), 706–719, 2014.

Quaife T, Lewis P, Disney MI, De Kauwe M, Williams M, Law B, et al.: Assimilating reflectance data into an ecosystem model to improve estimates of terrestrial carbon flux, 2007.

Thacker, W.C.: The role of the Hessian matrix in fitting models to measurements, Geophys.

Res. 94(C5), 6177-6196, 1989.

Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P.: Design and analysis of computer experiments. Stat. Sci. 4, 409–423, 1989

Verhoef, W. and Bach, H.: Coupled soil–leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data, Remote Sensing of Environment, 109(2), 166–182, 2007.

Verhoef, W. and Bach, H.: Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models, Remote Sensing of Environment, 87(1), 23–41. 2003.

Verrelst, J., Sabater, N., Rivera, J. P., Muñoz-Marí, J., Vicent, J., Camps-Valls, G., & Moreno, J.: Emulation of Leaf, Canopy and Atmosphere Radiative Transfer Models for Fast Global Sensitivity Analysis. *Remote Sensing*, *8*(8), 673, 2016.

Verstraete, M.M., Hunt, L.A., Scholes, R.J., Clerici, M., Pinty, B. and Nelson, D.L.: Generating 275-m Resolution Land Surface Products From the Multi-Angle Imaging SpectroRadiometer Data, IEEE Transactions on Geoscience and Remote Sensing, 50(10), 3980–3990, 2012.

Weiss, M., Baret, F., Myneni, R. B., Pragnère, A. and Knyazikhin, Y.: Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data, Agronomie, 20(1), 3-22, 2000.

Widlowski, J.-L., Pinty, B., Gobron, N., Verstraete, M.M., Diner, D.J. and Davis, A.B.: Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies, Climatic Change, 67, 403–415, 2004.

Widlowski JL, Pinty B, Lavergne T, Verstraete MM, Gobron N.: Using 1-D Models to Interpret the Reflectance Anisotropy of 3-D Canopy Targets: Issues and Caveats. IEEE Trans Geosci Remote Sens, 43, 2008–2017, 2005.

Account Info Help

Book: Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site

Author: M. Chernetskiy, N. Gobron, J. Gómez-Dans, et al

Publisher: John Wiley and Sons

Date: Mar 25, 2017

Copyright © 2017, John Wiley and Sons

Logged in as:
Maxim Chernetskiy
UCL
Account #:
3001290417

LOGOUT

Order Completed

Thank you for your order.

This Agreement between Maxim Chernetskiy ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your order details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License number	Reference confirmation email for license number
License date	May, 25 2018
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Wiley Books
Licensed Content Title	Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site
Licensed Content Author	M. Chernetskiy, N. Gobron, J. Gómez-Dans, et al
Licensed Content Date	Mar 25, 2017
Licensed Content Pages	23
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Text extract
Number of Pages	23
Will you be translating?	No
Title of your thesis / dissertation	Retrieving spectral and biophysical parameters of land vegetation by Earth Observation Land Data Assimilation System
Expected completion date	Sep 2018
Expected size (number of pages)	150
Requestor Location	UCL Flat 21 London, NW9 5QD United Kingdom Attn: UCL
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing address	UCL Flat 21 London, United Kingdom NW9 5QD Attn: UCL
Total	0.00 GBP

CLOSE WINDOW

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

[Remote Sensing] Manuscript ID: remotesensing-189179; doi:10.3390/rs9070656. Paper has been published.

MDPI – Website Editor <website@mdpi.com>

Tue 27/06/2017 15:09

To: Chernetskiy, Maxim <m.chernetskiy@ucl.ac.uk>; Gomez-Dans, Jose <j.gomez-dans@ucl.ac.uk>; nadine.gobron@jrc.ec.europa.eu <nadine.gobron@jrc.ec.europa.eu>; olivier.morgan@jrc.ec.europa.eu <olivier.morgan@jrc.ec.europa.eu>; Lewis, Philip <p.lewis@ucl.ac.uk>; sina.truckenbrodt@uni-jena.de <sina.truckenbrodt@uni-jena.de>;

Cc: billing@mdpi.com <billing@mdpi.com>; remotesensing@mdpi.com <remotesensing@mdpi.com>; website@mdpi.com <website@mdpi.com>; erika.zhao@mdpi.com <erika.zhao@mdpi.com>;

Dear Authors,

We are pleased to inform you that your article "Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS)" has been published in Remote Sensing and is available online:

Abstract: http://www.mdpi.com/2072-4292/9/7/656/
PDF Version: http://www.mdpi.com/2072-4292/9/7/656/pdf
Manuscript: http://www.mdpi.com/2072-4292/9/7/656/manuscript (available to authors after login)

Please take a moment to check that everything is correct. You can reply to this email if there is a problem. Note that at this stage we will not accept further changes to the manuscript text.

This link can be forwarded to your colleagues, and you can share the paper on various social networks by clicking the links on the article webpage. Alternatively, MDPI can post an announcement of your article on our Twitter channel (https://twitter.com/MDPIOpenAccess), please provide the doi number of the manuscript and a text of up to 117 characters with spaces to socialmedia@mdpi.com. Our service Scitations.net will automatically notify authors cited in your article.

We would be happy to keep you updated about new issue releases of Remote Sensing. Please enter your e-mail address in the box at http://www.mdpi.com/journal/remotesensing/toc-alert/ to receive notifications.

To order high quality reprints of your article in quantities of 25-1000, visit: http://www.mdpi.com/2072-4292/9/7/656/reprints

We support the multidisciplinary preprint platform /Preprints/, which permanently archives full text documents and datasets of working papers in all subject areas. Posting on the platform is entirely free of charge, and full details can be viewed at http://www.preprints.org.

We are dedicated to providing an outstanding publishing service, and we invite you to complete our author satisfaction survey https://www.surveymonkey.com/r/mdpiAuthorFeedback. The survey contains 16 short questions and will only take a couple of minutes to complete.

Thank you for choosing Remote Sensing to publish your work, we look forward to receiving further contributions from your research group in the future.

Kind regards,

Website Cn
Website Editor

--
MDPI AG
Postfach, CH - 4020 Basel, Switzerland
Office: St. Alban-Anlage 66, 4052 Basel, Switzerland
Tel. +41 61 683 77 34
Fax: +41 61 302 89 18
E-mail: website@mdpi.com
http://www.mdpi.com/

Your accepted manuscript [JASR_13844] is now available online

Elsevier - Article Status <Article_Status@elsevier.com>

Mon 23/07/2018 23:35

To:Chernetskiy, Maxim <m.chernetskiy@ucl.ac.uk>; mchernetskiy@gmail.com <mchernetskiy@gmail.com>;

Article title: Simulating arbitrary hyperspectral bandsets from multispectral observations via a generic Earth Observation-Land Data Assimilation System (EO-LDAS)
Article reference: JASR13844
Journal title: Advances in Space Research
Corresponding author: Mr. M. Chernetskiy
First author: Mr. M. Chernetskiy
Accepted manuscript available online: 23-JUL-2018
DOI information: 10.1016/j.asr.2018.07.015

Dear Mr. Chernetskiy,

We are pleased to inform you that your accepted manuscript (unformatted and unedited PDF) is now available online at:

https://doi.org/10.1016/j.asr.2018.07.015

You might like to bookmark this permanent URL to your article. Please note access to the full text of this article will depend on your personal or institutional entitlements.
This version of your article has already been made available at this early stage to provide the fastest access to your article. It is not intended to be the final version of
your article. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note changes to the
article should not be requested at this stage.

Your article can already be cited using the year of online availability and the DOI as follows: Author(s), Article Title, Journal (Year), DOI.

Once the full bibliographic details (including volume and page numbering) for citation purposes are available, you will be alerted by e-mail.

To track the status of your article throughout the publication process, please use our article tracking service:

https://authors.elsevier.com/tracking/article/details.do?aid=13844&jid=JASR&surname=Chernetskiy

Kind regards,
Elsevier Author Support

HAVE QUESTIONS OR NEED ASSISTANCE?
For further assistance, please visit our Customer Support site where you search for solutions on a range of topics and find answers for frequently asked questions. You can also talk to our customer support team by hone 24 hours a day from Monday-Friday and 24/7 by live chat and email.

Get started at > http://service.elsevier.com

© 2016 Elsevier Ltd | Privacy Policy http://www.elsevier.com/privacypolicy
Elsevier Limited, The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, United Kingdom, Registration No. 1982084. This e-mail has been sent to you from Elsevier Ltd. To ensure delivery to your inbox (not bulk or junk folders), please add article_status@elsevier.com to your address book or safe senders list.

[T-14-20180404]

Maxim Chernetskiy

Tel: +447754588496
E-mail: m.chernetskiy@ucl.ac.uk

SUMMARY

- Working experience in University College London (UCL) and Friedrich Schiller University (Germany).
- Experience as a part of international, inter-sectoral EU (Marie Curie) project; H2020 Towards a Biosphere Atmosphere Change Index (BACI) project; and ESA Fires CCI project.
- PhD in remote sensing submitted, with significant data assimilation and radiative transfer modelling components
- Strong programming skills (Python (numpy, matplotlib, gdal), C/C++, IDL)
- Familiar with C#, R, ERDAS, Visual Basic, MS Office automation, relational databases, SQL
- Previous use of various satellite sensors (Landsat/TM/ETM+, MISR, CHRIS/Proba, MODIS, AVHRR, MERIS) in a range of applications (e.g., retrieval land parameters using canopy radiative transfer modelling; time series analyses; land cover/land use change mapping; detection and monitoring of phytoplankton blooms in lakes)
- Skills in ecology and statistical analyses
- Experience in leading various environmental science and engineering research projects, related to remote sensing in particular
- Proven success in preparing research reports, publication in the scientific literature and conference presentations
- Proficient spoken and written English, beginner German, fluent Russian (mother tongue)

EDUCATION

PhD

Department of Earth Observation
Friedrich Schiller University Institute for Geography (Jena, Germany)
Topic: Multi-scale Remote Sensing Synergy for Land Process Studies
Supervisor: Prof. Dr. Christiane Schmullius, Dr. Nadine Gobron
Submitted, expected: September, 2018

Postgraduate student

Institute of Biophysics
Siberian Branch of Russian Academy of Science, (Krasnoyarsk, Russia)
2001 - 2004

Engineer (equivalent to MEng)

Specialization: Information systems; Ecology, natural resource management and scientific studies
Krasnoyarsk State Technical University (Krasnoyarsk, Russia)
Obtained: 2001

PROFESSIONAL EXPERIENCE

Research Associate
University College London, Department of Geography
ESA Fire CCI (Climate Change Initiative) project.
From November 2017

Research Associate
University College London, Department of Geography
Biosphere Atmosphere Change Index (BACI) project (H2020).
From December 2015

Marie Curie Early Stage Researcher
Copernicus (formerly known as GMES) Initial Operations – Network for Earth Observation Research Training (GIONET) project (Grant Agreement Number PITN-GA-2010-264509)
Department of Earth Observation
Friedrich Schiller University Institute for Geography (Jena, Germany)
December 2011 - December 2014

Visiting teacher
Transforming UK Agriculture through new technology. Harpenden. United Kingdom
4-5 June, 2018

Visiting Researcher
Department of Geography
University College London (London, UK)
5 - 8 October, 2012; 18 - 28 February, 2013; 5 - 8 September, 2013; 30 March - 11 April, 2014

Visiting Researcher
European Commission - Joint Research Centre Institute for Environment and Sustainability (Ispra, Italy)
24 May - 7 June, 2013; 1 - 7 December, 2013; 6 - 14 October, 2014

Lead Mathematician (Engineer Programmer)
Institute of Biophysics
Siberian Branch of Russian Academy of Science, (Krasnoyarsk, Russia)
August 2004 - November 2011

Software Engineer
Krasnoyarsk state administration (Krasnoyarsk , Russia)
October 2001 - November 2002

PUBLICATIONS

Chernetskiy, M., Gomez-Dans, J., Gobron, N., Morgan, O., Truckenbrodt, S., Lewis, P., Schmullius, C. 2017. Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing, 9(7), 656.

Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P., Schmullius, C. 2018. Simulations of CHRIS/PROBA spectra with Earth Observation Land Data Assimilation System using MISR data. *Advances in Space Research* (Accepted).

Chernetskiy, M., Gobron, N., Gómez-Dans, J., Lewis, P. and Schmullius, C.C., 2017. Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site. In Earth Observation for Land and Emergency Monitoring. John Wiley & Sons, Ltd, 117–139.

Shevyrnogov, A., **Chernetskiy, M.**, Vysotskaya, G. Multiyear trends of Normalized Difference Vegetation Index and temperature in the south of Krasnoyarsk Krai. Izvestiya Atmospheric and Oceanic Physics, 49, 9 (2013), 1047–1056.

Chernetskiy, M., Pasko, I., Shevyrnogov, A., Slyusar, N., Khodyayev, A. A study of forest vegetation dynamics in the south of the Krasnoyarskii Krai in spring. Advances in Space Research, 48 (2011) 819–825.

Chernetskiy, M., Shevyrnogov, A., Shevnina, S., Vysotskaya, G., Sidko, A. Investigations of the Krasnoyarsk Reservoir waters based on the multispectral satellite data. Advances in Space Research, 43(2), 15 January 2009, 206-213

Rogozin, D. Y., Zykov, V. V., **Chernetsky, M. Y.**, Degermendzhy, A. G., Gulati, R. D. Effect of winter conditions on distributions of anoxic phototrophic bacteria in two meromictic lakes in Siberia, Russia. Aquatic Ecology, 43 (3), (2009), 661-672

Shevyrnogov, A., Vysotskaya, G., Sukhinin, A., Frolikova, O., **Tchemetsky, M.** Results of analysis of human impact on environment using the time series of vegetation satellite images around large industrial centers. Advances in zSpace Research, 41(1), 2008, 36-40.

Chernetskiy, M., Disney, M., Gomez-Dans, J., Urban, M., Lewis, P., Schmullius, C. 2018. Temporal regularisation: Merging optical, thermal and SAR for the BACI project. Remote Sensing. (in preparation)

Chernetskiy, M., Disney, M., Gomez-Dans, S., Lewis, P. 2018. Influence of surface reflectance uncertainties on change detection. Remote Sensing. (in preparation)

CONFERENCE PRESENTATIONS AND PROCEEDINGS

PV2018 Conference - Adding value and preserving data, Rutherford Appleton Laboratory, Harwell Space Cluster (UK), May 15-17, 2018.
Poster presentation: *Processing surface state vector by temporal regularization of optical, thermal and SAR data*
Chernetskiy M., Disney M., Delgado A., Nagini M., Urban, M. and Schmullius, C.

IGARSS 2018, Valencia, Spain, July 22-27
Title: Uncertainty for Burnt Area Products
Brennan, J., Gomez-Dans, J., Lewis, P., **Chernetskiy, M.,** Heil, A.

IGARSS 2018, Valencia, Spain, July 22-27
Title: Uncertainty Characterisation & Validation within ESA Fire-CCI
Brennan, J., Lewis, P., Gomez-Dans, J., **Chernetskiy, M.,** Chuvieco, M., Lizundia, J., Campagnolo, M., Pereira, J., Oom, D.

Sentinel-3 for science workshop, Venice (Italy), June 2-5, 2015
Poster presentation: *Retrieval of biophysical canopy parameters on Sentinel-3 validation test sites using the Earth Observation Land Data Assimilation System and multiangular information of MISR*
Chernetskiy, M., Gobron, N., Morgan, O., Gomez-Dans, J., Lewis, P., Gitelson, A., Schmullius, C.

The 4th International Symposium on Recent Advances in Quantitative Remote Sensing (RAQRS), Valencia (Spain), September 22-26, 2014
Poster presentation and Proceedings: *Effectiveness of MISR multiangular observations in the new generation of Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P.

40th COSPAR Scientific Assembly, Moscow (Russia), August 2 – 10, 2014
Oral presentation: *Reconstruction of hyperspectral CHRIS/PROBA signal by the Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P.

GIONET Summer School, Frascati (Italy), June 30 - July 4, 2014
Poster presentation: *Earth Observation Land Data Assimilation System (EO-LDAS) regularization constraints over Barrax site*
Chernetskiy, M.

Sentinel-2 for science workshop, Frascati (Italy), May 20 - 22, 2014
Poster presentation and Proceedings: *Validation of effectiveness of MISR multiangular data in the Earth Observation Land Data Assimilation System (EO-LDAS)*
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Lewis, P., Schmullius, C.

ESA Living Planet Symposium, Edinburgh (UK), September 9 - 13, 2013
Oral presentation and Proceedings: *Validation of the Earth Observation Land Data Assimilation System by the field data of ESA SPARC field campaign*
Chernetskiy, M., Gomez-Dans, J., Lewis, P., Schmullius, C.

ESA EO Summer School 2012 on Earth system monitoring and modelling, Frascati (Italy), July 30 – August 10, 2012
Poster presentation: *Multi-scale remote sensing synergy for land process studies: increasing the revisit frequency of high resolution sensors.*
Chernetskiy, M.

39th COSPAR Scientific Assembly, Mysore (India), July 14 - 22, 2012
Oral presentation: *Multi-scale remote sensing synergy for land process studies: blending of high and low resolution imagery*
Chernetskiy, M.

The 18th Biennial Conference of International Society for Ecological Modelling
In Procedia Environmental Sciences 13 (2012) 194 – 201: *Boreal forests contribution to global seasonal dynamic of carbon dioxide in the atmosphere*
Bartsev, S., Degermendzhi, A., Ivanova, Y., Shchemel, A., **Tchernetsky M.**

POTENTIAL REFERENCES

Dr. Mathias (Mat) Disney
Department of Geography, University College London
Tel.: +44 (0)20 7679 0592; Email: mathias.disney@ucl.ac.uk

Prof. Dr. Christiane Schmullius
Department of Earth Observation, Friedrich Schiller University Institute for Geography
Tel.: +49 (0) 3641 94 88 80; Email: c.schmullius@uni-jena.de

Dr. Nadine Gobron, Project Manager
European Commission Joint Research Centre Institute for Environment and Sustainability
Email: nadine.gobron@jrc.ec.europa.eu

D.Sc. Prof. Anatoli Shevyrnogov, Vice-Director
Institute of Biophysics Siberian Branch of the Russian Academy of Sciences
Tel.: +7 (3912) 433 566; Email: ecoinf@ibp.ru; ap_42@mail.ru

Erklärung zu den Eigenanteilen des Promovenden/der Promovendin sowie der weiteren Doktoranden/Doktorandinnen als Koautoren an den Publikationen und Zweitpublikationsrechten bei einer kumulativen Dissertation (in die kumulative Dissertation einbinden)

Falls eine Publikation schon in einer kumulativen Promotion Verwendung gefunden hat, bitter Art und Umfang der Verwendung aufführen.

Publikation (Vollständiges Zitat): **Chernetskiy, M.**, Gobron, N., Gómez-Dans, J., Lewis, P. and Schmullius, C.C., 2017. Earth Observation Land Data Assimilation System (EO-LDAS) Regularization Constraints over Barrax Site. In Earth Observation for Land and Emergency Monitoring. John Wiley & Sons, Ltd, 117–139.					
Beteiligt an (Zutreffendes ankreuzen)					
	Maxim Chernetskiy	Nadine Gobron	Jose Gómez-Dans	Philip Lewis	Christiane Schmullius
Konzeption des Forschungsansatzes	X	X	X	X	X
Planung der Untersuchung	X	X			
Datenerhebung	X				
Datenanalyse und interpretation	X				
Schreiben des Manuskript	X				
Vorschlag Anrechnung Publikationsäquivalente	1.0	n.a	n.a	n.a	n.a

Publikation (Vollständiges Zitat): **Chernetskiy, M.**, Gomez-Dans, J., Gobron, N., Morgan, O., Lewis, P., Truckenbrodt, S. and Schmullius, C., 2017. Estimation of FAPAR over Croplands Using MISR Data and the Earth Observation Land Data Assimilation System (EO-LDAS). Remote Sensing, 9(7), 656.							
Beteiligt an (Zutreffendes ankreuzen)							
	Maxim Chernetskiy	Jose Gómez-Dans	Nadine Gobron	Olivier Morgan	Philip Lewis	Sina Truckenbrodt	Christiane Schmullius
Konzeption des Forschungsansatzes	X		X		X	X	X
Planung der Untersuchung	X						
Datenerhebung	X			X			
Datenanalyse und interpretation	X						
Schreiben des Manuskript	X	X	X				
Vorschlag Anrechnung Publikationsäquivalente	1.0	n.a	n.a	n.a	n.a	n.a	n.a

Publikation (Vollständiges Zitat):
Chernetskiy, M., Gobron, N., Gomez-Dans, J., Morgan, O., Disney, M., Lewis, P., Schmullius, C. 2018. Simulating arbitrary hyperspectral bandsets from multispectral observations via a generic Earth Observation-Land Data Assimilation System (EO-LDAS). *Advances in Space Research* (submitted).

Beteiligt an (Zutreffendes ankreuzen)							
	Maxim Chernetskiy	Nadine Gobron	Jose Gómez-Dans	Olivier Morgan	Mathias Disney	Philip Lewis	Christiane Schmullius
Konzeption des Forschungsansatzes	X	X				X	X
Planung der Untersuchung	X						
Datenerhebung	X			X			
Datenanalyse und interpretation	X						
Schreiben des Manuskript	X	X	X		X		
Vorschlag Anrechnung Publikationsäquivalente	1.0	n.a	n.a	n.a	n.a	n.a	n.a

Erklärung zu den Eigenanteilen des Promovenden/der Promovendin sowie der weiteren Doktoranden/Doktorandinnen als Koautoren an den Publikationen und Zweitpublikationsrechten bei einer kumulativen Dissertation

Für alle in dieser kumulativen Dissertation verwendeten Manuskripte liegen die notwendigen Genehmigungen der Verlage für die Zweitpublikation vor.

Die Co-Autoren der in dieser kumulativen Dissertation verwendeten Manuskripte sind sowohl über die Nutzung, als auch über die oben angegebenen Eigenanteile informiert und stimmen dem zu.

Die Anteile des Promovenden / der Promovendin sowie der weiteren Doktoranden / Doktorandinnen als Koautoren an den Publikationen und Zweitpublikationsrechten bei einer kumulativen Dissertation in der Anlage aufgeführt.

Name des Promovenden/der Promovendin	Datum	Ort	Unterschrift
Maxim Chernetskiy	08.02.2019	London	

Ich bin mit der Abfassung der Dissertation als publikationsbasiert, d.h. kumulativ, einverstanden und bestätige die vorstehenden Angaben. Eine entsprechend begründete Befürwortung mit Angabe des wissenschaftlichen Anteils des Doktoranden/der Doktorandin an den verwendeten Publikationen werde ich parallel an den Rat der Fakultät der Chemisch-Geowissenschaftlichen Fakultät richten.

Name Erstbetreuer(in)	Datum	Ort	Unterschrift

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe.

Ort, Datum

London 08.02.2019

Unterschrift des Verfassers/der Verfasserin