



## **An Earth Observation Land Data Assimilation System (EO-LDAS)**

Jose Gomez-Dans (1), Philip Lewis (1), Tristan Quaife (2), Tomas Kaminski (3), and Jon Styles (4)

(1) Department of Geography, UCL, and National Centre for Earth Observation, Gower St., London, WC1E 6BT, United Kingdom (j.gomez-dans@ucl.ac.uk), (2) Dept. of Meteorology and National Centre for Earth Observation, University of Reading, United Kingdom, (3) FastOpt, Hamburg, Germany, (4) Assimila, United Kingdom

In order to monitor the land surface, EO data provides the means of achieving global coverage in a timely fashion. Different sensors orbit the Earth acquiring data at different times and with different spectral and spatial properties. Blending all these observations presents a considerable challenge. Purely statistical methods based on machine learning techniques require accurate and extensive ground truth for "training" models. The complexities of the processes that take place in the scene result in limited usefulness of these models outside their training region or period. Models that describe the physical processes that give rise to the measurements, based on radiative transfer theory, offer a more robust way of interpreting the recorded data and relating it to surface properties such as leaf area index, chlorophyll concentration, etc. Unfortunately, the information content in the signals is rarely sufficient to unambiguously determine the many parameters that are required in typical radiative transfer models. To improve on this, the use of prior information is required. Typically, this information is given as parameter ranges, or maybe even distributions, which can have a positive effect in the so-called "inverse problem". Data assimilation techniques allow one to use models of the land surface as priors, to constrain the inverse problem. These models can be very useful in improving the ability of inverting the observations, as the models can give very valuable information on the dynamics of some parameters, like LAI. However, some parameters that have a strong bearing on the observations (some pigments, leaf angle distributions...) have no analogues in typical DGVMs.

In this work, we introduce and demonstrate the use of weak constraint 4DVAR data assimilation to the problem of inverting optical RT models. We demonstrate that the use of this technique results in important gains in parameter uncertainty reduction for a typical satellite mission, including parameter correlations. The proposed framework also results in seamless inference of land surface parameters even when no observations are available due to cloudiness or orbital configuration. Finally, the use of physically based models allows the combined use of different sensors with different spectral or angular characteristics. As no process models are available for typical RT model parameters, we use regularisation concepts to implement a first or second order difference model, although other models (e.g. DGVMs) could be used for some parameters as mentioned above. These other models can be incorporated by means of linearised versions.

The prototype code (available from <<http://www.eoldas.info>>) is demonstrated for time series data, both in a synthetic experiment, but also with real data. We show that we can retrieve many parameters per time step using data from the upcoming Sentinel-2 or MODIS data. The reduction in uncertainty in the estimates is also noteworthy, a factor of 2 over the typical inversions is usually achieved. Strong parameter correlations are observed, indicating that parameters compensate for each other. We also apply the concept spatially, again using regularisation models. This approach allows the combination of observations with different spatial resolutions. The reconstructed surface parameter maps are in very good agreement with the truth. Finally, we discuss some issues in the methodology, such as the choice of hyperparameters, and practical avenues to speed up the inversion procedure.