LEAF AREA INDEX ESTIMATION FROM MODIS DATA USING THE ENSEMBLE KALMAN SMOOTHER METHOD

Huaan Jin, Jindi Wang, Zhiqiang Xiao, Zhuo Fu

1.State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications, CAS, Beijing, China, 100875
2. Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing Normal University, Beijing, China, 100875
3. School of Geography and Remote Sensing Science, Beijing Normal University, Beijing, China, 100875
Email: wangjd@bnu.edu.cn, jinhuaan2008@hotmail.com

1. INTRODUCTION

Leaf area index (LAI) is an important structural parameter of vegetation canopy and is also one of the basis variables driving the models used in regional and global biogeochemical, ecological and meteorological applications. Although several crop growth models can simulate the continuous variation of LAI, the prediction errors caused by the uncertainties of model and parameters would accumulate gradually so that simulation may diverge from the true state. Data assimilation algorithms have been introduced to estimate time series LAI from remote sensing observations by integrating canopy reflectance model and crop growth model. It permits a reduction in the uncertainties in model predictions. From a purely algorithmic perspective, all current data assimilation algorithms can be grouped into two categories: variational assimilation and sequential assimilation [1]. The variational approach generally requires an adjoint model, which is difficult to derive from a land surface model. The sequential method, on the other hand, has attracted increased attention in recent years due to its flexibility and versatility in the assimilation of remotely sensed data into the process model. There is a great deal of research in sequential data assimilation to employ Ensemble Kalman filter (EnKF). However, there has been little work so far on applying Ensemble Kalman smoother (EnKS) for estimating LAI from remotely sensed surface reflectance. This paper aims to estimate LAI using EnKS by coupling a radiative transfer model and an empirical LAI temporal profile model. The algorithm is tested using MODIS data at some sites which are from several existing observation networks (Fluxnet, Bigfoot, etc.) and validated using ground LAI measurements. The results are compared with those from the ENKF. The advantages and disadvantages of each method are discussed.

2. METHODOLOGY AND DATA

In general, the data assimilation scheme is composed of a data assimilation algorithm, a model operator, an observation operator, and data sets.

2.1 Data assimilation algorithms

Data assimilation algorithm is used to integrate simulation and observation, which utilizes observation information to update the state variables produced by model operator. In this study, the EnKS [2] and EnKF [3] are adopted as assimilation algorithms.

Define the matrix holding an n-dimensional model state vector $\varphi_i \in \Re^n$, $A = (\varphi_1, \varphi_2, ..., \varphi_N) \in \Re^{n \times N}$, where N is the number of ensemble and n is the size of the model state vector. The ensemble mean is stored in each column of a matrix $\overline{A} \in \Re^{n \times N}$, and the ensemble perturbation matrix $A' \in \Re^{n \times N}$ can be defined as $A' = A - \overline{A}$. Then, at observation time t, the standard ENKF analysis equation is

$$A^a(t) = A(t) + A'(t) A'^T(t) H^T \left(HA'(t) A'^T(t) H^T + R\right)^{-1} \left(D(t) - HA(t)\right) \quad (1)$$

where $H \in \Re^{m \times n}$ (with m being the number of measurements) is observation operator relating the model state to the observations, $R \in \Re^{m \times m}$ represents the measurement error covariance matrix, and $D(t) \in \Re^{m \times N}$ is the perturbed observations ensemble matrix. The superscript 'a' denotes the analysis state and the superscript 'T' indicates a matrix transpose.

The EnKS uses the observation at time t to update the state at previous estimation times t' ($t > t'$) too. Similar to the analysis equation (1), the analysis for a prior time t' can be written as

$$A^a(t') = A(t') + A'(t') A'^T(t) H^T \left(HA'(t) A'^T(t) H^T + R\right)^{-1} \left(D(t) - HA(t)\right) \quad (2)$$

The equation (2) is updated repetitively every time a new set of measurements are introduced at future times t [4]. The EnKS can be considered as an extension of the EnKF since the EnKF solution is used as the first guess for the analysis, which is propagated backward in time by using the ensemble covariances [2].

2.2 Model operator and observation operator

In this study, the LAI climatology is derived from LAI averages of multi-year MODIS LAI products (MOD15A2) with best quality (quality control values less than 32) for specific vegetation type. Owing to the profile of LAI averages fluctuates sometimes, especially in the growing season, the adaptive Savitzky-Golay filtering [5] is then used to derive the LAI envelope. The following dynamic model is constructed based on the LAI envelope to provide the forecast of LAI.

$$LAI_t = F_t \bullet LAI_{t-1} \quad (3)$$

and

$$F_t = 1 + \frac{1}{LAI_t^{env} + \varepsilon} \times \frac{dLAI_t^{env}}{dt} \quad (4)$$

where LAI_t^{env} is the temporal profile from the LAI envelope, and $\varepsilon = 10^{-3}$ is to prevent a null denominator.

The observation operator is used to build the relationship between simulated state variables and observations. In our scheme, the simulated variable is LAI and the observations are the MODIS surface bi-directional reflectance data. The two-layer canopy reflectance model ACRM [6] is selected as the observation operator, which is applied to convert the simulated LAI into canopy reflectance.

Besides LAI, the reflectance simulated by ACRM model shows a high sensitivity to changes in some other parameters (defined as 'free parameters'), such as weight of the first/second Price function r_{s1}, r_{s2}, leaf structure parameter N, etc. During the process of assimilation, these free parameters are evolved by a stationary model as follow:

$$X_n = F \bullet X_{n-1} \tag{5}$$

where X denotes the state variable, the operator F is equal to the identity matrix. It implies that these free parameters remain constant when no observations are available.

2.3 Data sets

The data sets used in this paper include remote sensing data and field measurement data. MODIS LAI data product MOD15A2 and MODIS surface reflectance data product MOD09A1 are used in this study.

The selected sites include Bondville, Konza, Mead, etc. The Bondville site (40.006° N, 88.292° W) is an agricultural site in the mid-western part of the United States. The field was continuous no-till with alternating years of soybean and maize crops. The Konza site (39.089° N, 96.571° W) is located in the Flint Hills region of northeastern Kansas. Vegetation at the site was tallgrass prairie. In addition to grassland, there are areas of gallery forest and some croplands in the northern part of the study site. The Mead site (41.180° N, 96.440° W) is located at the University of Nebraska Agricultural Research and Development Center near Mead, Nebraska. The field in this site was under no-till with alternating years of maize and soybean crops since 2001.

Except LAI, field data include some environmental variables, such as vegetation type, soil type, leaf wetness, soil moisture, etc. Some sites provide LAI reference map based on the relationship between field LAI data and remote sensing information from ETM+ sensor.

3. RESULTS AND DISCUSSION

This research results demonstrate the feasibility of implementing the Ensemble Kalman smoother in the time series LAI estimation. Although the CPU requirements for the EnKS are similar to those needed for the EnKF, the EnKS is capable of providing an improvement over the EnKF estimate. Using the EnKF solution as a first guess, the EnKS estimate eliminates the discontinuities or abrupt spikes normally obtained with EnKF, and generates a smoothed LAI estimation in time series. In the future, we hope to carry out more extensive validation for other land cover types and plan to extend the EnKS algorithm at regional scale.

4. REFERENCES

[1] O. Talagrand, “Assimilation of observations, an introduction”, *Journal of the Meteorological Society of Japan*, vol. 75, pp. 191-209, 1997.

[2] G. Evensen and P. J. van Leeuwen, “An ensemble Kalman smoother for nonlinear dynamics”, *Monthly Weather Review*, vol. 128, pp. 1852-1867, 2000.

[3] G. Evensen, “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics”, *Journal of Geophysical Research*, vol. 99, no. C5, pp. 10143-10162, 1994.

[4] G. Evensen, “The Ensemble Kalman Filter: theoretical formulation and practical implementation”, Ocean Dynamics, vol. 53, pp. 343-367, 2003.

[5] A. Savitzky and M. J. E. Golay, “Smoothing and differentiation of data by simplified least squares procedures”, *Analytical Chemistry*, vol. 36, no. 8, pp. 1627-1639, 1964.

[6] A. Kuusk, “A two-layer canopy reflectance model”, *Journal of Quantitative Spectroscopy & Radiative Transfer*, vol. 71, pp. 1-9, 2001.