

RETRIEVALS OF CANOPY BIOPHYSICAL VARIABLES USING MULTI-TEMPORAL REMOTE SENSING DATA

Zhiqiang Xiao ^{a, b, *}, Jindi Wang ^{a, b}, Tiegang Tong ^c

^a State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing Applications, 100875, China

^b Research Center for Remote Sensing and GIS, School of Geography, Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing Normal University, 100875, China - zhqxiao@bnu.edu.cn

^c School of Info-Physics and Geomatics Engineering, Central South University, Changsha, 410083, China

KEY WORDS: Remote Sensing, Retrieval, Leaf area index, Multi-temporal

ABSTRACT:

The integrated application of multi-source and multi-temporal remote sensing data is the trend of remote sensing application research, and it is also the practical need to solve the inversion problem of remote sensing. In this paper, a method is developed to retrieve canopy biophysical variables using multi-temporal remote sensing data. The inherent change rules of biophysical variables are introduced into the retrieval methods by coupling the radiative transfer model with land process model to simulate time series surface reflectances. A cost function is constructed to compare the reflectances simulated by the coupled model with time series reflectances measured by sensors and the canopy biophysical variables with the available prior information. And an optimization method is used to minimize the cost function by adjusting the values of input canopy biophysical variables such as the temporal behaviour of the reflectances simulated reaches the best agreement with the multi-temporal reflectances measured. Retrieval of leaf area index from MODIS surface reflectance data (MOD09) at the Bondville site was performed to validate this method. The experimental results shows that the use of multi-temporal remote sensing data can significantly improve estimation of canopy biophysical variables.

1. INTRODUCTION

Remote sensing data have been widely used to estimate the canopy biophysical variables which are applied to large area water and carbon cycle simulation, climatic modelling and global change research. Therefore, it is very important to precisely estimate these variables from remote sensing data at the regional or global scale. Currently, there are many methods to estimate biophysical variables from remote sensing data (Liang, 2004; Weiss, 1999). And they can be roughly divided into following classes: by statistical relationship between LAI and spectral vegetation indices, by physical model inversion and by other nonparametric methods. These methods have their own advantage and disadvantage. Since the model inversion methods is physically based and can adjust to a wide range of situation, radiative transfer models are more and more used in the inverse mode to estimate the canopy biophysical variables (Kuusk, 1991; Jacquemoud, 1993).

It is well known that the inverse problem is by nature an ill-posed problem mainly because of the not unique solution and the measurement and model uncertainties (Combal, 2003). Currently, there are two sorts of ways to solve the ill-posed problem. One is dependent on some kinds of hypothetical condition, which may reduce the accuracy of remote sensing data products retrieved by radiative transfer model, and the other is to use the prior information of biophysical variables in model inversion.

With the development of remote sensing technique, many new types of sensors have been developed, and there are large numbers of remote sensing data with different temporal, spatial

and spectral resolution obtained from space-borne or air-borne sensors. The integrated application of multi-source, multi-temporal remote sensing data is the trend of remote sensing application research, and it is also the practical need to solve the inversion problem of remote sensing. Noted that most of the biophysical variables, such as LAI, are time-dependent and possess inherent change rules along with time which are often represented by process models such as crop growth models, it is an important way for us to introduce the inherent change rules of biophysical variables into the retrieval methods to add the amount of information needed to retrieve biophysical variables and improve the accuracy of remote sensing products. In this paper, we will present our research on the issue of integrating multi-temporal remote sensing data to retrieve biophysical variables.

Radiative transfer model are coupled with land process model to simulate time series surface reflectances. And a cost function is constructed, according to the posterior probability formula defined by Tarantola, to compare the reflectances simulated by the coupled model with time series reflectances measured by sensors. And an optimization method is used to minimize the cost function by adjusting the values of input canopy biophysical variables such as the temporal behaviour of the reflectances simulated reaches the best agreement with the multi-temporal reflectances measured. By introducing the physical constraints from the land surface model, the method can integrate multi-temporal remote sensing data to retrieve biophysical variables.

* Corresponding author.

To validate this method, we used double logistic model to describe the temporal LAI profile of agriculture crops, and Kuusk model are coupled with the empirical statistical model to simulate the surface reflectance. The experimental results show that the use of multi-temporal remote sensing data can significantly improve estimation of canopy biophysical variables.

2. DATA

Retrieval of leaf area index from MODIS surface reflectance data (MOD09) at the Bondville site was performed to validate this method. The Bondville site, located at (40.0061000, -88.291867), is an agricultural site in the Mid-western part of the United States, near Champaign, Illinois. The site is part of the network of eddy covariance flux towers associated with AmeriFlux and the network of Core Validation Sites associated with the MODIS Land Team. It was established in 1996, with the long-term goal of obtaining the necessary in situ information to test and improve the representation of land-surface processes in soil-vegetative-atmosphere transfer (SVAT) models. The field was continuous no-till with alternating years of soybean and maize crops (Meyers, 2004). In 2001, the crop was maize with the maximum leaf area of 4.38 and an associated height of 2.4m. And there are time series of field measurements of LAI, which can be used to compare with the retrieved LAI.

To test the new methods, the input data include multi-year MODIS LAI product (MOD15A2) and time series of MODIS reflectance product (MOD09A1) in 2001. All these products are from Collection 4. And a 49km² region around the tower or field site is extracted, so there are 7×7 subsets from MODIS LAI product with the spatial resolution of 1km, 14×14 subsets from MODIS reflectance product with the spatial resolution of 500m.

3. METHOD

3.1 Retrieval method using multi-temporal data

The problem of parameter estimation from remote sensing data is underestimated, and Tarantola gives the theory to resolve it (Tarantola, 2004). Let \mathbf{M} be the model space, and \mathbf{D} the data space. Tarantola defined the posterior probability density in the space of $\mathbf{D} \times \mathbf{M}$ as follows.

$$\sigma(\mathbf{d}_{i,j}, \mathbf{m}_{i,j}) = k \frac{\rho(\mathbf{d}_{i,j}, \mathbf{m}_{i,j}) \Theta(\mathbf{d}_{i,j}, \mathbf{m}_{i,j})}{\mu(\mathbf{d}_{i,j}, \mathbf{m}_{i,j})} \quad (1)$$

where k is a normalization constant, $\mathbf{m}_{i,j}$ and $\mathbf{d}_{i,j}$ are vectors in model space and data space respectively, $\rho(\mathbf{d}_{i,j}, \mathbf{m}_{i,j})$ is the prior probability density in the space of $\mathbf{D} \times \mathbf{M}$, which represents the prior information of observation variables and model parameters, $\Theta(\mathbf{d}_{i,j}, \mathbf{m}_{i,j})$ is the theoretical probability density which constructs the physical correlations between the observation variables and model parameters, and $\Theta(\mathbf{d}_{i,j}, \mathbf{m}_{i,j}) = \theta(\mathbf{d}_{i,j} | \mathbf{m}_{i,j}) \mu(\mathbf{m}_{i,j})$ given the model parameter $\mathbf{m}_{i,j}$, $\mu(\mathbf{d}_{i,j}, \mathbf{m}_{i,j})$ is the homogeneous probability density of observation variables and model parameters. Then, the

posteriori information in the model space is given by the marginal probability density

$$\sigma_{\mathbf{M}}(\mathbf{m}_{i,j}) = \int_{\mathbf{D}} d\mathbf{d}_{i,j} \sigma(\mathbf{d}_{i,j}, \mathbf{m}_{i,j}) \quad (2)$$

Suppose the prior information of observation variables and model parameters is independent, equation (2) is then

$$\sigma_{\mathbf{M}}(\mathbf{m}_{i,j}) = k \rho_{\mathbf{M}}(\mathbf{m}_{i,j}) \int_{\mathbf{D}} d\mathbf{d}_{i,j} \frac{\rho_{\mathbf{D}}(\mathbf{d}_{i,j}) \theta(\mathbf{d}_{i,j} | \mathbf{m}_{i,j})}{\mu_{\mathbf{D}}(\mathbf{d}_{i,j})} \quad (3)$$

Further assume that model parameters, observation variables and the a priori information on the model parameters are Gaussian, then we can get the cost function shown in equation (4) that has been widely used in the parameter retrieval from remote sensing data.

$$S(\mathbf{m}_{i,j}) = \frac{1}{2} \left(g(\mathbf{m}_{i,j}) - \mathbf{d}_{i,j}^{obs} \right)^T C_{\mathbf{D}}^{-1} \left(g(\mathbf{m}_{i,j}) - \mathbf{d}_{i,j}^{obs} \right) + \frac{1}{2} \left(\mathbf{m}_{i,j} - \mathbf{m}_{i,j}^{prior} \right)^T C_{\mathbf{M}}^{-1} \left(\mathbf{m}_{i,j} - \mathbf{m}_{i,j}^{prior} \right) \quad (4)$$

where $g(\mathbf{m}_{i,j})$ is the forward model, $C_{\mathbf{D}}$ is the covariance matrix representing the measurement uncertainties and model uncertainties, $C_{\mathbf{M}}$ is the covariance matrix representing the uncertainties of a priori information on the model parameters. The retrieval of canopy biophysical variables from remote sensing data is to minimize the equation (4) to find the model parameters $\mathbf{m}_{i,j}$, which possess the maximum a posterior probability. However, the equation is constructed just using the individual pixel measurement and a priori information on the model parameters.

Noted that most of the biophysical variables, such as LAI, are time-dependent and possess inherent change rules along with time which are often represented by process models such as crop growth models, we made an attempt to retrieve canopy biophysical variables using the multi-temporal remote sensing data by introducing the inherent change rules of biophysical variables into the retrieval methods.

The observational information of pixel

(i, j) at times $t+1$ and t can be integrated to estimate canopy biophysical variables of pixel (i, j) at time t . Thus, the data space is extended as $\mathbf{D} = \mathbf{D}_{i,j}^t \times \mathbf{D}_{i,j}^{t+1}$, in which each vector is $\mathbf{d} = (\mathbf{d}_{i,j}^t, \mathbf{d}_{i,j}^{t+1})$. Therefore, the posterior probability density in the model space can be expressed as follow.

$$\sigma_{\mathbf{M}}(\mathbf{m}_{i,j}^t) = k \rho_{\mathbf{M}}(\mathbf{m}_{i,j}^t) \int_{\mathbf{D}} d\mathbf{d} \frac{\rho_{\mathbf{D}}(\mathbf{d}) \theta(\mathbf{d} | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}}(\mathbf{d})} \quad (5)$$

Assume the observational data of pixel (i, j) at times $t+1$ and t are independent each other. Then, equation (5) can be extended as

$$\sigma_M(\mathbf{m}_{i,j}^t) = k \rho_M(\mathbf{m}_{i,j}^t) \int_{\mathbf{D}_{i,j}^t} d\mathbf{d}_{i,j}^t \frac{\rho_{\mathbf{D}_{i,j}^t}(\mathbf{d}_{i,j}^t) \theta(\mathbf{d}_{i,j}^t | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}_{i,j}^t}(\mathbf{d}_{i,j}^t)} \quad (6)$$

$$\int_{\mathbf{D}_{i,j}^{t+1}} d\mathbf{d}_{i,j}^{t+1} \frac{\rho_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1}) \theta(\mathbf{d}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1})}$$

where the first integral term, similar to the integral term in equation (3), represents the parameter retrieval constraint from the observational data at time t , and the second integral terms represent parameter retrieval constraints from the observational data at time $t+1$.

Assume that model parameters, observation variables and the a priori information on the model parameters are Gaussian, then, the first integral term in equation (6) can be written analogously as follow.

$$T(\mathbf{m}_{i,j}^t) = \int_{\mathbf{D}_{i,j}^t} d\mathbf{d}_{i,j}^t \frac{\rho_{\mathbf{D}_{i,j}^t}(\mathbf{d}_{i,j}^t) \theta(\mathbf{d}_{i,j}^t | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}_{i,j}^t}(\mathbf{d}_{i,j}^t)} \quad (7)$$

$$= \alpha \exp\left(-\frac{1}{2} \left(g(\mathbf{m}_{i,j}^t) - \mathbf{d}_{i,j}^{t,obs}\right)^T C_D^{-1} \left(g(\mathbf{m}_{i,j}^t) - \mathbf{d}_{i,j}^{t,obs}\right)\right)$$

And the second integral term is

$$R(\mathbf{m}_{i,j}^t) = \int_{\mathbf{D}_{i,j}^{t+1}} d\mathbf{d}_{i,j}^{t+1} \frac{\rho_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1}) \theta(\mathbf{d}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1})} \quad (8)$$

$$= \int_M d\mathbf{m}_{i,j}^{t+1} \left\{ \int_{\mathbf{D}_{i,j}^{t+1}} d\mathbf{d}_{i,j}^{t+1} \frac{\rho_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1}) \theta(\mathbf{d}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)}{\mu_{\mathbf{D}_{i,j}^{t+1}}(\mathbf{d}_{i,j}^{t+1})} \right\} \theta(\mathbf{m}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)$$

$$= \int_M d\mathbf{m}_{i,j}^{t+1} T(\mathbf{m}_{i,j}^{t+1}) \theta(\mathbf{m}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)$$

where $\theta(\mathbf{m}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t)$ is a transition probability that the parameter is $\mathbf{m}_{i,j}^{t+1}$ at time $t+1$ given the parameter was $\mathbf{m}_{i,j}^t$ at time t and is related to process models which describe the relationship between $\mathbf{m}_{i,j}^{t+1}$ and $\mathbf{m}_{i,j}^t$.

Then, equation (6) can be written as follow.

$$\sigma_M(\mathbf{m}_{i,j}^t) = k \rho_M(\mathbf{m}_{i,j}^t) T(\mathbf{m}_{i,j}^t) R(\mathbf{m}_{i,j}^t) \quad (9)$$

Apply logarithm to both sides of the equation (9), then

$$COST(\mathbf{m}_{i,j}^t) = -\frac{1}{2} \left(\mathbf{m}_{i,j}^t - \mathbf{m}_{i,j}^{t,prior}\right)^T C_D^{-1} \left(\mathbf{m}_{i,j}^t - \mathbf{m}_{i,j}^{t,prior}\right) - \frac{1}{2} \left(g(\mathbf{m}_{i,j}^t) - \mathbf{d}_{i,j}^{t,obs}\right)^T C_D^{-1} \left(g(\mathbf{m}_{i,j}^t) - \mathbf{d}_{i,j}^{t,obs}\right) + \ln \int_M d\mathbf{m}_{i,j}^{t+1} T(\mathbf{m}_{i,j}^{t+1}) \theta(\mathbf{m}_{i,j}^{t+1} | \mathbf{m}_{i,j}^t) \quad (10)$$

Equation (10) is the cost function to retrieve the canopy biophysical variables at pixel (i, j) using the observational information at time t and time $t+1$. Similar method can be applied to derive the cost function to retrieve the canopy

biophysical variables at time t using the observational information at times $t-N, \dots, t-1, t, t+1, \dots, t+N$ together.

3.2 Radiative transfer model

Radiative transfer models describe the relationship between canopy characteristics and reflectance, and many of them have been developed to obtain land surface biophysical parameters (Kuusk, 1994; Jacquemoud, 1992; Liang, 1993). In our parameter retrieval, the Markov chain reflectance model (MCRM) developed by Kuusk (Kuusk, 1995; Kuusk, 2001) is chosen as the forward model to simulate the canopy reflectance. This model incorporates the Markov properties of stand geometry into an analytical multispectral canopy reflectance model, which makes the model more flexible and more applicable. The MCRM can calculate the angular distribution of the canopy reflectance for a given solar direction from 400 to 2500 nm. The inputs of the forward MCRM are summarized in Table 1.

Parameters	Value range	Unit
Solar zenith angle	0~90	Degree
Relative azimuth angle	0~180	Degree
Viewing zenith angle	0~90	Degree
Angstrom turbidity factor	0.1~0.5	
Ratio of leaf dimension and canopy height	0.02~0.4	
*Markov parameter	0.4~1.0	
Factor for refraction index	0.7~1.2	
Eccentricity of the leaf angle distribution	0.0~4.5	Degree
Mean leaf angle of the elliptical LAD	0.0~90.0	Degree
Leaf specific weight	100	g/m ²
Chlorophyll AB content	0.3~0.8	% of SLW
Leaf water content	100~200	% of SLW
Leaf dry matter content	95~100	% of SLW
Leaf structure parameter	1.0~3.0	
*Weight of the first Price function	0.05~0.4	
*Weight of the second Price function	-0.1~0.1	
Weight of the third Price function	-0.05~0.05	
Weight of the forth Price function	-0.05~0.05	
*Leaf area index	0~10	m ² /m ²

* free parameters

Table 1. The parameters needed to run the MCRM

3.3 Process model

Fisher (Fisher, 1994(a); Fisher, 1994(b)) used an empirical statistical model to describe the temporal NDVI profile of agriculture crops. The model is a double logistic function to describe the NDVI profile. In our study, the double logistic model, shown in formula (11), is used to describe the seasonal LAI trajectory.

$$LAI(\theta; t) = vb + \frac{av}{1 + \exp(-c(t-p))} - \frac{av + vb - ve}{1 + \exp(-d(t-q))} \quad (11)$$

where t is the time variable representing the day of the year, and January 1 is set zero, av is related to the asymptotic value of LAI, c and d denote the slopes at the first and second inflection points, p and q are the date of these two points, and vb and ve are the LAI values at the beginning and the end of the growing season

The averages of MODIS LAI within the site area are fit to determine the parameters of double logistic function which is used as the process model in our method to retrieve LAI using multi-temporal remote sensing data. Figure 2 shows the averages of MODIS LAI at the Bondville site within a 49km² region around the tower or field site. On days 169 and 177, there are no LAI values over the region due to instrument problems. And the fitted double logistic model is also shown in Figure 1. Obviously, the double logistic function can effectively describe the LAI profiles for these vegetation types.

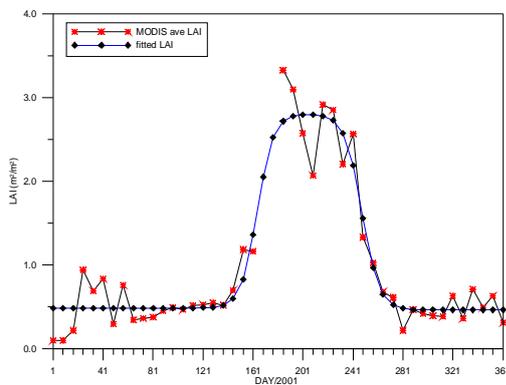


Figure 1. The averages of MODIS LAI at the Bondville site together with the fitted double logistic model

4. EXPERIMENTAL RESULTS

In order to test the above algorithm, the MODIS surface reflectance data (MOD09) at the Bondville site are used to retrieve LAI. The results are also compared with LAI retrieved the basic method which just uses the individual pixel measurement. Figure 2 demonstrates the retrieved LAI time series for crops. The LAI time series retrieved by the basic method are shown in Figure 2(a). And Figure 2(b), 2(c) and 2(d) demonstrate the LAI time series retrieved by the new method by integrating three, five and seven continuous MODIS surface reflectance data respectively. It is clear that the LAI values at this flux site have markedly underestimated the field measurements in the crop growing season. And there are fluctuations, especially in the crop growing season, because it is difficult to acquire cloud-free image due to the high amount of moisture content in the atmosphere during the growing season. By comparison, the temporally integrated inversion method can remove noise shown as abrupt rises or drops, especially when more MODIS surface reflectance data are integrated. Moreover, the accuracy of the LAI by the new method has been significantly improved over the LAI retrieved by the basic method compared to the field measured LAI data.

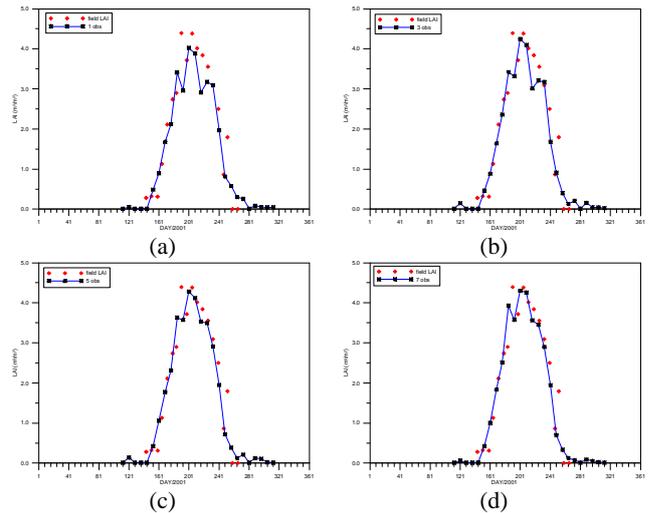


Figure 2. Retrieved LAI time series using multi-temporal remote sensing data

5. CONCLUSION

A method to retrieve LAI using multi-temporal remote sensing data was designed to produce spatially and temporally continuous LAI products with relatively higher quality. The algorithm integrates the inherent change rules of biophysical variables into the retrieval methods to improve the temporal consistency of the retrieved LAI by coupling the radiative transfer model with the empirical statistical model. Results as described in this paper have shown that the new algorithm is able to produce more continuous LAI product, and the validation of the retrieved LAI against the field measurements shows that the use of multi-temporal remote sensing data can significantly improve the accuracy of the parameter retrieval.

REFERENCES

- Combal, B., Baret, F., Weiss, M. et al, 2003. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment*, 84(1), pp. 1-15.
- Fisher A., 1994(a). A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote sensing of Environment*, 48, pp. 220-230.
- Fisher A., 1994(b). A simple model for the temporal variations of NDVI at regional scale over agricultural countries: validation with ground radiometric measurements. *International Journal of Remote sensing*, 15, pp. 1421-1446.
- Jacquemoud S. and Baret F., 1990. PROSPECT: a model of leaf optical properties spectra. *Remote sensing of Environment*, 34, pp. 75-91.
- Jacquemoud S. and Baret F., 1993. Estimating vegetation biophysical parameters by inversion of a reflectance model on high spectral resolution data. In: *Crop structure and light microclimate: Characterization and Applications*, Paris, France: INRA, pp. 339-350.

- Kuusik A., 1991. The inversion of the Nilson-kuusk canopy reflectance model, a test case. In: *Proceeding of the 11th International Geoscience and Remote Sensing Symposium*, Helsinki, Finland, pp.1547-1550.
- Kuusik A., 1994. A multispectral canopy reflectance model. *Remote Sensing of Environment*, 50, pp. 75-82.
- Kuusik A., 1995. A Markov chain model of canopy reflectance. *Agricultural and Forest Meteorology*, 76, pp. 221-236.
- Kuusik A., 2001. A two-layer canopy, reflectance model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 71, pp. 1-9.
- Liang S., Strahler A. H., 1993. An analytic BRDF model of canopy radiative transfer and its inversion. *IEEE Transactions on Geoscience and Remote Sensing*, 31, pp. 1081-1092.
- Liang S., 2004. *Quantitative Remote Sensing of Land Surfaces*. New York: John Wiley and Sons, Inc.
- Meyers T. P., Hollinger S. E., 2004. An assessment of storage terms in the surface energy balance of maize and soybean. *Agricultural and Forest Meteorology*, 125, pp. 105-115.
- Tarantola A., 2004. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia.
- Weiss M. and Baret F., 1999. Evaluation of canopy biophysical variable retrieval performances from the accumulation of large swath satellite data. *Remote Sensing of Environment*, 70, pp. 293-306.

ACKNOWLEDGEMENTS

This research was supported by Chinese 973 Program under grant 2007CB714407, the National Natural Science Foundation of China under grant 40571107, 40701102.

