International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

Data-model fusion for improving LAI mapping: a case study over China's land mass

Mei Huang [a b], Jing M. Chen [a] & Feng Deng [a]

[a] Department of Geography, University of Toronto, Toronto, Ontario, M5S 3G3, Canada

[b] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China

Available online: 29 Jul 2011

To cite this article: Mei Huang, Jing M. Chen & Feng Deng (2011): Data-model fusion for improving LAI mapping: a case study over China's land mass, International Journal of Remote Sensing, DOI:10.1080/01431161.2010.520347

To link to this article: http://dx.doi.org/10.1080/01431161.2010.520347

International Journal of Remote Sensing
iFirst, 2011, 1–18

Taylor & Francis
Taylor & Francis Group

Data-model fusion for improving LAI mapping: a case study over China's land mass

MEI HUANG*†‡, JING M. CHEN† and FENG DENG†

†Department of Geography, University of Toronto, Toronto, Ontario, M5S 3G3, Canada
‡Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

(*Received 7 October 2008; in final form 21 July 2010*)

A simple data-model fusion method is developed to improve leaf area index (LAI) mapping using satellite data. The objective is to overcome two issues with satellite-derived LAI maps: (1) optical remote sensing data are often seriously affected by the atmosphere due to clouds, and in some areas no reliable data are obtained in the whole growing season, and (2) seasonal variations in conifer LAI derived from satellite data are often distorted by the seasonal variations in leaf greenness (pigments), the background vegetation and snow cover, etc., and the derived LAI reflects the overall greenness rather than the actual forest leaf area present in a pixel. These shortcomings of satellite measurements can be greatly alleviated when an ecological model is used to simulate the LAI in the absence of reliable remote sensing data and to estimate the seasonal variation of LAI according to ecological principles. The usefulness of this fusion method is demonstrated through improving a China-wide LAI map series in 10-day intervals at 1 km resolution using Satellite Pour l'Observation de la Terre (SPOT) VEGETATION (VGT) data.

1. Introduction

Leaf area index (LAI) is one of the most important parameters in ecological studies as the leaf area is the basis for radiation absorption, energy and mass transfer to the atmosphere, carbon dioxide (CO_2) absorption for photosynthesis, etc. LAI is defined as one-half the total green leaf area (all sided) per unit ground surface area (Chen and Black 1992). Many previous studies have demonstrated the important effects of LAI on the exchanges of energy, momentum and mass between the atmosphere and terrestrial ecosystems (Sellers *et al*. 1997, Bondeau *et al*. 1999, Buermann *et al.* 2001, Foley *et al.* 2003). The study of Xue *et al.* (1996) on the impact of vegetation on US summer weather prediction suggested that accurate simulation of LAI seasonal variations can improve the accuracy of temperature and precipitation prediction. Alan *et al.* (2004) indicated that the seasonal variation of LAI for deciduous forests is the main factor determining net ecosystem productivity. Several studies have shown that the uncertainty in LAI is one of the main errors in the estimation of net primary productivity (NPP) (Willams and Rastetter 1999, Churkina *et al.* 2003), and without accurate LAI

*Corresponding author. Email: huangm@igsnrr.ac.cn

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2011 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431161.2010.520347

seasonal trajectory, it would be impossible to model plant response to global change (Eweret 2004).

LAI can be measured on the ground by a number of direct and indirect methods (Gower *et al.* 1999), but these methods are time-consuming and strenuous. Moreover, it is quite difficult to use ground-based methods alone for studies across large temporal and spatial scales. Remote sensing techniques are a way to estimate LAI time series for large areas. However, satellite-derived LAI products are inevitably affected by atmospheric conditions, including clouds, aerosols, water vapour, ozone, etc. Although much of the effect of these conditions can be removed using real-time or near real-time atmospheric observations made by the same sensor and other sensors, the remaining effect can sometimes be very large (Vermote *et al.* 2002, Chen *et al.* 2006). Variations in other factors, such as satellite view angle and solar illumination angle, can also cause errors in retrieved LAI. The seasonal variations in the greenness of leaves, that is, the chlorophyll content, and the soil background covers (understorey and snow cover in forests) incur additional challenges in deriving the seasonal variation in LAI. Without additional information on leaf greenness and background, the retrieved LAI seasonal variation is considerably distorted, especially for conifer forests (Tan *et al.* 2005, Yang *et al.* 2006, Pisek *et al.* 2007). Because of the decrease in leaf chlorophyll content in the winter as well as snow cover on the background, LAI of conifer forests retrieved through optical remote sensing often shows very small values, while in reality, it is only slightly smaller than the summer peak value and plays an important role in radiation absorption and energy balance.

On the other hand, LAI can also be generated by ecosystem models driven by environmental factors that affect plant growth and therefore LAI. On the assumption that resource availability constrains the distribution of plants, Prentice *et al.* (1992), Neilson (1995), Woodward *et al.* (1995), Kergoat (1998) and Luo *et al.* (2002) simulated LAI at regional and global scales. The atmosphere–vegetation interaction model version 2 (AVIM2) used in this study to generate LAI is different from the above models. It is based on biophysical and biogeochemical processes at the land surface. LAI, which is the internal variable of the model, can vary with growth, and its variation would also feed back to the mass and energy balance calculations for the canopy. In AVIM2, the allocation of photosynthetic assimilates to various biomass components and the estimation of plant phenology often have considerable errors, and therefore the parameterization for these processes needs independent data for validation. Albeit having the inevitable atmospheric effects, satellite data can be used to derive the maximum LAI in the growing season more reliably than models for all cover types. They can also be used to derive the phenology (e.g. leaf-on and leaf-off dates) more reliably than models. It is therefore logical to develop a methodology that can make use of the advantages of both satellite data and models while avoiding their shortcomings.

The objective of this study is to develop a methodology to fuse model-generated LAI with satellite-derived LAI to improve LAI mapping. The methodology is tested using an existing satellite LAI map over China's land mass.

2. Materials and methods

2.1 *Data source and data pre-processing*

The satellite-derived LAI data set from 2003 used in this study was described in Chen *et al.* (2006) and Deng *et al.* (2006). It is derived based on 10-day synthesis Satellite Pour l'Observation de la Terre (SPOT) VEGETATION (VGT) reflectance images

Table 1. The class names and codes of the vegetation cover map and the number of detected cloudy pixels as a percentage of the total number of China pixels.

Code	Class name	Cloudy percentage (%)
1	Tree cover, broadleaved, evergreen	3.23
2	Tree cover, broadleaved, deciduous, closed	3.39
3	Tree cover, broadleaved, deciduous, open	0.00
4	Tree cover, needle-leaved, evergreen	4.61
5	Tree cover, needle-leaved, deciduous	0.48
6	Tree cover, mixed leaf type	0.04
7	Tree cover, regularly flooded, fresh water	0.00
8	Tree cover, regularly flooded, saline water	0.00
9	Mosaic: tree cover/other natural vegetation	0.45
10	Tree cover, burnt	0.00
11	Shrub cover, closed-open, evergreen	1.81
12	Shrub cover, closed-open, deciduous	0.02
13	Herbaceous cover, closed-open	2.55
14	Sparse herbaceous or sparse shrub cover	0.40
15	Regularly flooded shrub and/or herbaceous cover	0.12
16	Cultivated and managed areas	0.00
17	Mosaic: cropland/tree cover/other natural vegetation	0.00
18	Mosaic: cropland/shrub and/or grass cover	0.00
19	Bare areas	0.00
20	Water bodies	0.00
21	Snow and ice	0.00
22	Artificial surfaces and associated areas	0.00
23	No data	

at 1 km resolution. The LAI algorithm was developed using the geometrical optical model (Four-Scale) to consider the bidirectional reflectance distribution function without a prior angular normalization procedure (Chen and Leblanc 1997). A variety of spatial data sets for 2003, including climate, soil texture and vegetation, are needed to drive AVIM2. The daily total precipitation, daily mean temperature, wind speed, relative humidity and cloud cover of 720 weather stations in China were interpolated to $0.1^\circ \times 0.1^\circ$ grids using the ANUSPLINE technique (Hutchinson 1989). The soil texture map used in this study is a combination of three regional maps of similar scales (1:14 000 000, 1:4 000 000 and 1:5 000 000) (Zhang *et al.* 2004). The soil texture was divided into 12 grades: gravel, sand, coarse sand, fine sand, silty sand, sandy silt, silt, silty clay, silty loam, loam, loamy clay and clay. The vegetation map for this study is the global land cover 2000 (GLC 2000) which was downloaded from the website http://www-gem.jrc.it/glc2000/objectivesGLC2000.htm. The vegetation cover was divided into 22 types (table 1).

2.2 *Methodology*

In order to take the advantages of both satellite-derived and model-generated LAI, the following steps were taken: (1) development of methods for cloud detection, (2) use of the satellite-derived LAI in cloud-free pixels to determine the leaf-on and leaf-off dates of each plant functional type (PFT) in the year and then estimation of the corresponding accumulated thermal energy (degree-days) as inputs to AVIM2,

(3) running AVIM2 to get 10-day averaged LAI and (4) fusion of model-generated and satellite-derived LAI.

2.2.1 Cloudy pixel detection method. The satellite-derived LAI series used in this study were screened to reduce the atmospheric effects using a locally adjusted cubic-spline capping method (LACC; Chen *et al.* 2006). The LACC method finds the optimum LAI seasonal trajectory for each pixel by replacing abnormally low LAI values in the trajectory with cubic-spline capping values. The capping curvature in different times of the growing season is automatically adjusted according to the curvature of a preliminary curve fitting. In order for LACC to perform reliably, 20 data points in a growing season are required. However, in areas with persistent cloud cover, there are not enough data points in a 10-day interval to meet this minimum requirement, and the seasonal variation pattern remains irregular after performing the LACC algorithm. The first step of this study was to find these residual cloud-affected pixels which were not successfully screened by LACC. A cloudiness index (CI) was used for this purpose. It is defined as the aggregate of the local maximum LAI minus local minimum LAI in a three-date moving window:

$$\mathrm{CI} = \sum_{i=1}^{n} (L_{\mathrm{max}i} - L_{\mathrm{min}i}), \tag{1}$$

where n is the total number of local minimum LAI points determined with the moving window during the growing season from 1 May to 31 October. Within the moving window of three consecutive dates in 10-day intervals, the local minimum LAI ($L_{\mathrm{min}i}$) is found when the ith 10-day LAI value (LAI_i) is smaller than both the $(i-1)$th 10-day LAI value (LAI_{i-1}) and $(i+1)$th 10-day LAI value (LAI_{i+1}), and the local maximum LAI ($L_{\mathrm{max}i}$) is found when LAI_i is greater than both LAI_{i-1} and LAI_{i+1}. The accumulated difference between the consecutive local minimum and maximum LAI values over the growing season signifies the cloudiness. Figure 1 shows an example of calculating CI for a deciduous forest in north-eastern China. There are three local maximum LAI points and two local minimum LAI points in this LAI seasonal trajectory. In this case CI is given as:

$$\mathrm{CI} = (L_{\mathrm{max}1} - L_{\mathrm{min}1}) + (L_{\mathrm{max}2} - L_{\mathrm{min}2}). \tag{2}$$

This is an extreme example. In fact, most of the detected number of local minimum LAI in all the study pixels are one, only a few are greater than two.

If there is no local minimum LAI point detected in an LAI curve then CI is zero. Table 2 shows different thresholds of CI for differentiating cloudy pixels and the corresponding detected percentages of cloudy pixels of total China pixels. Any pixels having CI greater than the threshold are considered to be cloudy. If the strictest threshold (0.0) is used to detect the percentage of cloudy pixels of total China pixels, then 19.67% of pixels are judged as cloudy. Table 2 shows the detected cloudy-pixel percentage increases from 2.7% to 19.67% as the threshold decreases from 2.0 to 0.0. The stricter the threshold, the higher the detected cloudy-pixel percentage. We checked the LAI seasonal trajectories pixel by pixel for different CI values and found that for the threshold 0.0, some slight declines of LAI were judged as cloudy effects, and these LAI drops may have been caused by other reasons (such as drought, pests, etc.). To

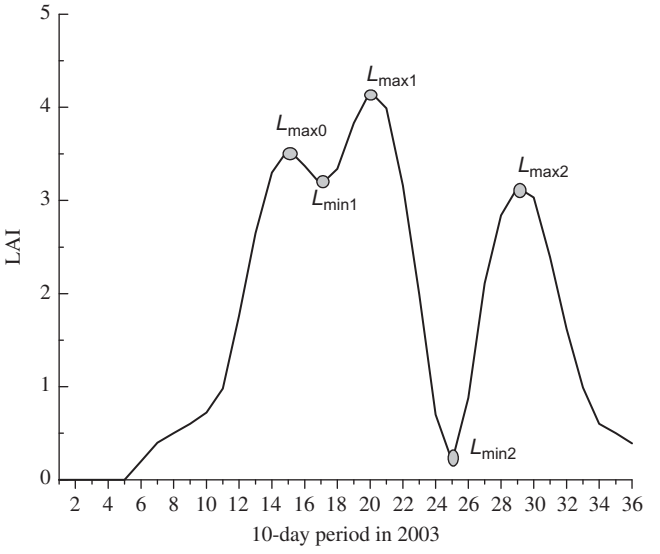

Figure 1. An example of calculating the CI value for a broadleaved deciduous forest LAI at 43.89° N, 129.83° E.

Table 2. Different thresholds of CI and the detected percentage of cloudy pixels of total China pixels.

CI	Cloudy pixels (%)	CI	Cloudy pixels (%)
0.00	19.7	0.1	16.1
0.02	18.6	0.2	14.3
0.04	17.8	0.4	11.9
0.06	17.1	1.0	7.0
0.08	16.6	2.0	2.7

avoid misjudgement, the threshold 0.06 was chosen. It is a moderately strict standard. The unreliable pixel percentage is 17.1% for the chosen threshold. Table 1 shows the constituents of these cloudy pixels. The most affected vegetation cover is evergreen needle-leaved forest (4.61%), followed by deciduous broadleaved forest (3.39%), evergreen broadleaved forest (3.23%), herbaceous cover (2.55%) and evergreen shrub cover (1.81%). For pixels in cropland and managed areas, the CI is set to 0.0, since it is hard to differentiate between erratic LAI variations caused by atmospheric effects and those caused by human activity. The affected percentages for other vegetation covers are less than 0.5.

2.2.2 The atmosphere–vegetation interaction model (AVIM2). The AVIM2 used in this study to generate LAI every 10 days is a process-based model which was first developed by Ji (1995). This model consists of a plant-growth module, a soil–vegetation–atmosphere transfer (SVAT) scheme, and a soil-carbon and nitrogen-dynamics module. The detailed description of the plant-growth model and the SVAT scheme can be found in Ji (1995), Lu and Ji (2006) and Dan and Ji (2007). The soil-carbon and nitrogen-dynamics module is fully described by Huang *et al.* (2007). Here

we describe only the processes which are closely related to the simulation of LAI. In this model, the forest biomass is divided into three parts: foliage, root and the remaining part. For grassland, it is only separated into foliage and root, and its allocation coefficient for foliage is the same as for forest.

Vegetation LAI is calculated by:

$$\mathrm{LAI} = M_{\mathrm{f}}(\mathrm{SLA}), \tag{3}$$

where M_{f} is the foliage biomass, and SLA is the specific leaf area (Schulze *et al.* 1994). The biomass of a tissue is determined by photosynthesis, respiration, allocation of dry matter to the tissue, phenological phase and its mortality rate, that is,

$$\frac{\mathrm{d}M_i}{\mathrm{d}t} = \eta_i(A - R) - D_i, \tag{4}$$

where M_i is the biomass of tissue i. i can be f, r or s, representing foliage, root and the remaining part, respectively. A is the photosynthetic rate, and R is the sum of maintenance and growth respiration. η_i denotes the allocation coefficient of tissue i, and D_i refers to its mortality rate. The following allocation schemes were used in this study: (1) during the early stage of the growing season, more assimilates are allocated to foliage and thus the vegetation can absorb more CO_2 more quickly, and (2) as LAI increases with time, the portions allocated to other parts increase gradually to maintain the nutrient and water balance of the vegetation (Farrar 1992, Ji 1995, Dickinson *et al.* 1998, Lu and Ji 2002). On the basis of these assumptions, the allocation coefficients for foliage (η_{f}), root (η_{r}) and the other parts (η_{s}) are written as:

$$\begin{cases} \eta_{\mathrm{f}} = (1 - a_1)\exp(-b_1(\mathrm{LAI})/(\mathrm{LAI})_{\max}) \\ \eta_{\mathrm{s}} = a_2(1 - \eta_{\mathrm{f}}) \\ \eta_{\mathrm{r}} = (1 - a_2)(1 - \eta_{\mathrm{f}}) \end{cases}, \tag{5}$$

where b_1 is an empirical constant, and a_1 and a_2 are parameters determined by the phenological phase; $(\mathrm{LAI})_{\max}$ is the maximum LAI.

In this study, the phenological phase was assumed to be controlled by the accumulated thermal energy as measured by the degree-days above a threshold air temperature. The complete phenological cycle was determined using both satellite and meteorological data in the following steps: (1) the onset and the offset dates for deciduous vegetation were found from the satellite-generated LAI seasonal trajectory in cloudless pixels near a pixel of interest, (2) these dates were correlated with the accumulated degree-days of the daily mean air temperature above 5°C and (3) the phenology of a cloudy pixel in question was determined using the measure of degree-days. For evergreen coniferous trees, the relative magnitude of seasonal LAI variation is determined by the lifespan of the leaves. The leaf lifespan for coniferous forest was set as 6.17 years (Reich *et al.* 1999). In AVIM2, a minimum temperature of 0°C was set for photosynthesis for all PFTs.

2.2.3 Strategies for data-model fusion. For evergreen forest, the new LAI value after fusion in the ith 10-day period ($\mathrm{LAI}_{\mathrm{N}i}$) is given as:

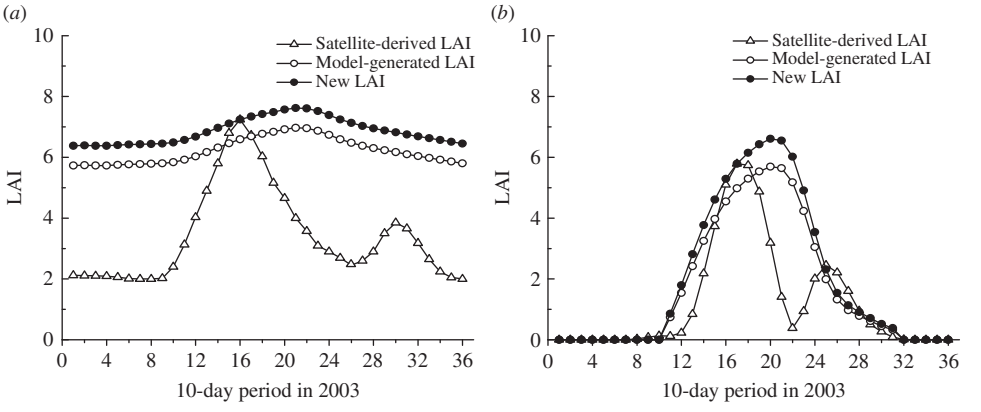

Figure 2. Examples of data-model strategy (*a*) for evergreen forest; (*b*) for deciduous forest, shrubland and grassland.

$$\mathrm{LAI}_{\mathrm{N}i} = (\mathrm{LAI})_{\mathrm{A}i} + ((\mathrm{LAI})_{\mathrm{Rmax}} - (\mathrm{LAI})_{\mathrm{Amax}}), \tag{6}$$

where $\mathrm{LAI}_{\mathrm{A}i}$ is the ith 10-day LAI generated by AVIM2, and $\mathrm{LAI}_{\mathrm{Rmax}}$ and $\mathrm{LAI}_{\mathrm{Amax}}$ represent the satellite-derived and model-generated maximum LAI, respectively. Figure 2(*a*) shows an example of this method. This data-model fusion strategy was developed to minimize the errors in satellite LAI data in the seasonal trajectory. Coniferous forests have only small LAI seasonal variations in reality, depending on the needle longevity. However, in remote sensing images, the LAI of coniferous forests often shows large seasonal variations due to the variation in leaf pigments and the variation in the background (understorey, moss, snow, etc.). In particular, during the winter months, when the background is covered by snow and leaves are less green, the LAI retrieved from remote sensing is often close to zero, while in reality a large LAI is still present and can greatly affect radiation absorption and energy balance. If LAI derived from remote sensing in this way is not corrected, it becomes unreliable as input to land-surface schemes, such as those used in numerical weather forecast and climate simulations, which require accurate simulation of surface energy balance. Our fusion strategy was to rely on model simulation of the seasonal trajectory while using remote sensing data to determine the maximum LAI in the season. The fusion process basically shifts the modelled LAI curve by the difference in the maximum LAI between the model and the satellite values.

For deciduous forest, shrubland and grassland, the new LAI after fusion is given as:

$$\mathrm{LAI}_{\mathrm{N}i} = \mathrm{LAI}_{\mathrm{A}i} \frac{(\mathrm{LAI})_{\mathrm{Rmax}}}{(\mathrm{LAI})_{\mathrm{Amax}}}. \tag{7}$$

An example of this method is shown in figure 2(*b*). This fusion is similar to that for conifers, that is, combining the modelled seasonal trajectory with the satellite-derived maximum LAI value. The mathematical formulation (equation (7)) is different from equation (5) in order to maintain the zero LAI values before and after the growing season.

In both cases, the maximum LAI from the satellite data can occur at different times of the seasons from the modelled value. We matched the satellite-based maximum value with the modelled value at the same date, so that the fused LAI can sometimes have a slightly larger LAI value than that derived from satellite data.

In this study, the satellite-derived LAI values for croplands and managed areas were not altered by the fusion process, as their seasonal variation patterns are much more affected by human activity than meteorology, and our modelled seasonal trajectories become less reliable than satellite data. There is a supplementary step for grassland. If LAI_{Rmax} is 50% less than LAI_{Amax}, then the new LAI is given as:

$$LAI_{Ni} = (LAI)_{Ai}. \tag{8}$$

The reason for this is that the satellite-derived LAI values for grassland have not been validated in China, whereas the model-generated LAI values agree well with the field observations. The satellite LAI algorithm is sensitive to the soil background optical properties, which are difficult to acquire for large areas. The comparisons between the grassland measurements and AVIM2 simulations are shown in figure 3. The field-observed LAI values were obtained from an Inner Mongolian grassland ecosystem observation station located at 43.63° N, 116.70° E (Du *et al.* 2001).

In this study, the resolution of satellite-derived LAI is 1 km and that of modelled LAI is 0.1°. As the modelled shape of the LAI seasonal trajectory does not vary greatly in space and the maximum LAI is determined at 1 km resolution, the fused LAI effectively remains at 1 km resolution. For detected cloudy pixels, where the satellite maximum LAI was also unreliable, the following protocols were followed: if a 1 km pixel had the same vegetation type as the coarse pixel, then its LAI was taken as the coarse pixel LAI, and otherwise its LAI was taken as that of an adjacent coarse pixel which had the same vegetation type. It seldom happened that there were no reliable

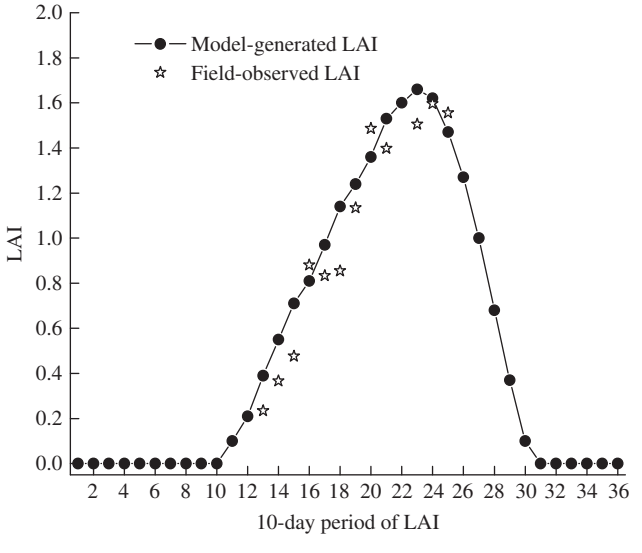

Figure 3. Comparison of model-generated LAI seasonal trajectory with the field observations (field data source: Du *et al.* 2001).

LAI values for a large region. When this was the case, we used a modelled LAI value for the same PFT in the nearest distance.

3. Results

3.1 *Cloudy pixel detection results*

The calculated CI values are shown in figure 4(*a*). Any pixel with CI value greater than 0.06 was detected as a cloudy pixel. These are the pixels that have frequent cloud cover in the year, for which the seasonal trajectory could not be reliably reconstructed using the LACC method (Chen *et al.* 2006), and their LAI values were replaced using the data-model fusion method. Figure 4(*b*) shows the distribution of annual mean cloud amount, averaged over the period of 1961–1990. The cloud amount (0–10 tenths of sky cover) data were from ground meteorological station observations and were interpolated to 0.1° resolution. The detected geographical locations of cloud-affected pixels in 2003 correspond very well with the mean cloud amount. They are mainly distributed in southern and north-eastern China, where the annual mean cloud amount values are greater than five. The examples of LAI with low and high CI values for deciduous and evergreen forests are shown in Figure 4(*c*) and (*d*). It is obvious that the LAI curve with CI around the threshold of 0.06 shows a smooth seasonal cycle, whereas the LAI curves with higher CI values show abrupt seasonal variations.

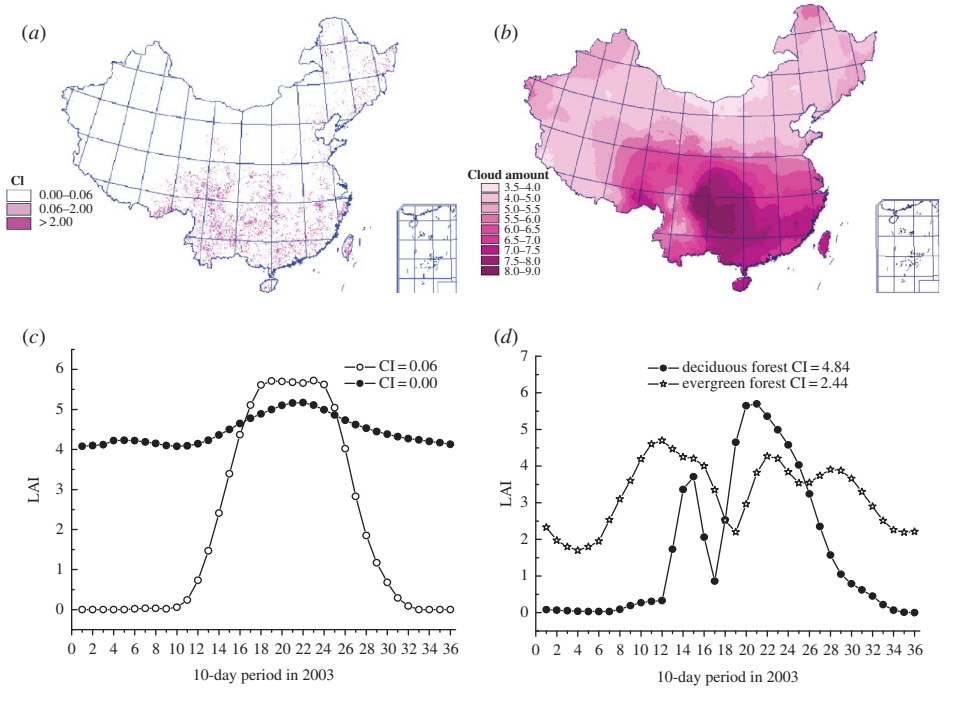

Figure 4. (*a*) The calculated cloudiness index map; (*b*) the averaged annual mean cloud amounts; (*c*) examples of CI equal to the threshold 0.06 for deciduous forest at 44.12° N, 126.75° E, and CI equal to 0.00 for evergreen forest at 25.00° N, 100.28° E; (*d*) examples of high CI for deciduous forest at 44.94° N, 129.60° E and for evergreen forest at 30.00° N, 102.21° E .

3.2 *Comparison of satellite-derived and modelled LAI in cloudless pixels*

The modelled maximum LAI values averaged for the individual PFTs (table 1) were compared with those of the satellite-derived LAI (figure 5), which shows that the two sets of values generally agree very well. For evergreen shrub and herbaceous cover, the averaged modelled LAI values are slightly greater than satellite-derived LAI values. This close agreement suggests that the data-model fusion would not introduce systematic bias to the satellite-derived LAI while removing the errors due to the atmospheric effects and the seasonal variation.

3.3 *Comparison of LAI maps before and after data-model fusion*

The LAI maps before and after data-model fusion in winter and summer are visually compared in figure 6. Before the fusion, the LAI map in January (figure 6(*a*)) has maximum LAI values that are smaller than 2.0, which are obviously incorrect for coniferous forests. After the fusion (figure 6(*b*)), the LAI values of many pixels are increased to around 5.0. These pixels are mostly in subtropical forests and in middle- and high-latitude coniferous forests. After the fusion, the forest LAI values in August are also increased in cloudy pixels (figure 6(*d*) in comparison with figure 6(*c*)). Before the fusion, grassland LAI values are between 0.0 and 1.0 in Inner Mongolia and the Tibetan plateau (figure 6(*c*)), and these values are increased to 1.0–2.0 after the fusion (figure 6(*d*)).

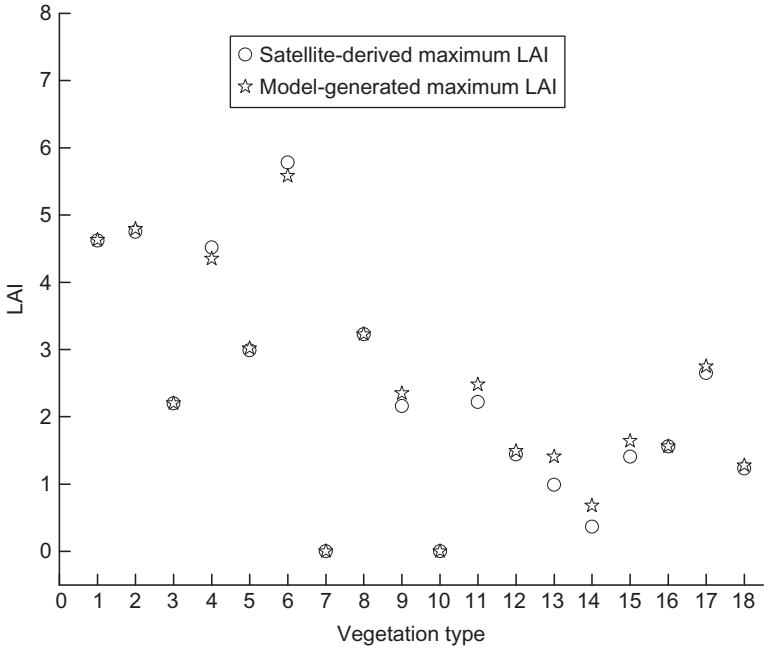

Figure 5. Comparison of the maximum LAI values, averaged for individual PFTs.

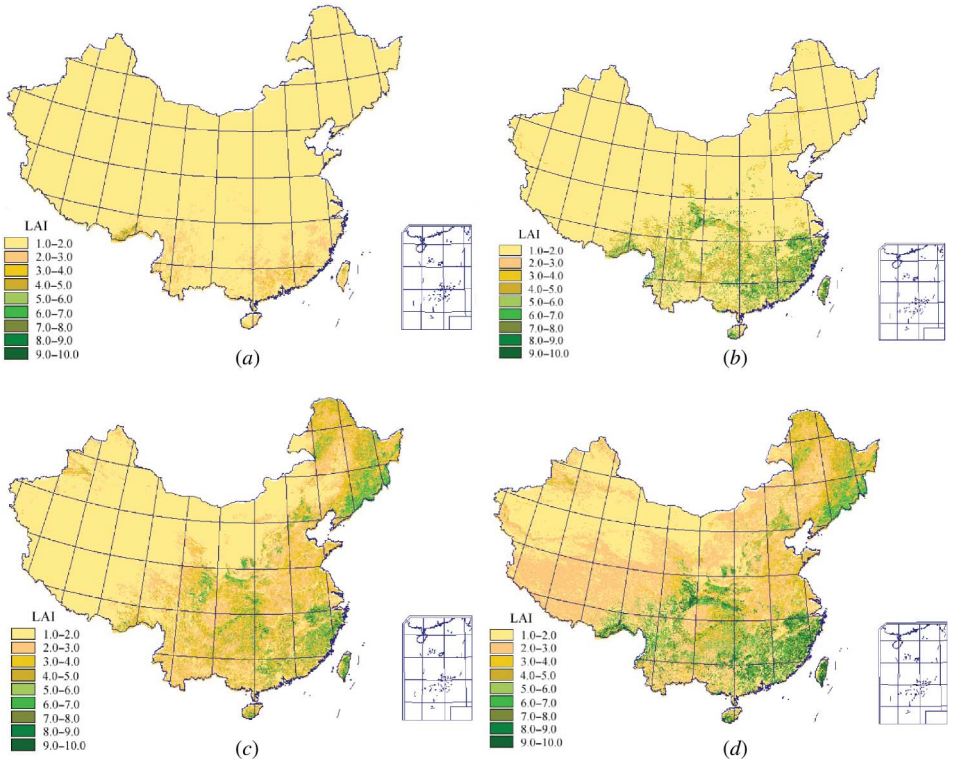

Figure 6. LAI distributions in the first 10-day period of January, (*a*) before and (*b*) after the data-model fusion; LAI distributions in the first 10-day period of August, (*c*) before and (*d*) after the data-model fusion.

3.4 *Comparison of China-wide seasonal maximum LAI with an existing global data set*

In order to compare the LAI values in China with other available data, the PFTs were merged into forest, grassland and shrubland, and the China-wide mean maximum LAI values in the growing season for the three cover-type groups were compared with an existing global data set (table 3). The global mean maximum LAI data were compiled by Asner *et al.* (2003), who collected more than 1000 published estimates and removed the statistical outliers using inter-quartile range analysis.

Table 3 shows that after the data-model fusion, the mean maximum LAI values over China's land mass for forest, grassland and shrubland are all slightly lower than their

Table 3. Comparison of China-wide mean maximum LAI values after fusion with global mean estimates (the data in the brackets are global estimates).

Vegetation type	Mean	Standard deviation
Forest	4.24 (4.36)	1.46 (2.00)
Grassland	1.51 (1.70)	1.04 (1.20)
Shrub	1.93 (2.10)	1.19 (1.60)

corresponding global mean measurements. The maximum forest LAI is 4.24 ± 1.46 in China, and the global mean measurement is 4.36 ± 2.0. The maximum LAI values of grassland and shrubland in China are 1.59 ± 1.04 and 1.93 ± 1.19 respectively, compared to the corresponding global values of 1.7 ± 1.2 and 2.1 ± 1.6. Since 55% of the LAI values referenced by Asner *et al*. (2003) were from the USA and Japan, the differences between these two data sets are reasonable.

3.5 *Comparison with two Landsat ETM images*

To investigate the accuracy of individual pixel LAI values, comparisons were made between the improved LAI images after fusion and those in the corresponding Enhanced Thematic Mapper (ETM) images in China. One ETM scene is located in Liping County in the south-west of China, and the other in the Changbaishan Nature Reserve in north-east China. The LAI values for the two images were calculated from atmospherically corrected Landsat ETM$^+$ reflectance using algorithms based on field observations. The resolutions for the two images are 30 m (Wang *et al.* 2007, Zheng *et al.* 2007). Since the spatial resolution is 1 km for the VGT LAI image and 30 m for the ETM LAI images, both coarse- and fine-resolution images were resampled to 3 km for comparison purposes. The final comparison was made at 3 km rather than 1 km, in order to reduce the effect of pixel misregistration of the coarse resolution image, following the methodology of Chen *et al.* (2002). For the Changbaishan area, with a large LAI dynamic range due to the altitude variation, the VGT LAI is well-correlated with the ETM LAI (figure 7(*a*)), and the coefficient of determination (R^2) value reaches 0.8 with the root mean square error (RMSE) equal to 0.36. For the Liping area, with a small LAI dynamic range, the correlation is not as good, with a smaller R^2 value of 0.36, but the RMSE is reduced to 0.30 (figure 7(*b*)). The *p*-values for the two regressions are both less then 0.0001.

The comparison of mean LAI values of improved LAI with those of ETM LAI for the Changbaishan and Liping images is shown in table 4. For the image of Changbaishan, the dominant vegetation cover is deciduous broadleaved forest, deciduous needle-leaved forest, mixed forest and herbaceous cover. The improved LAI

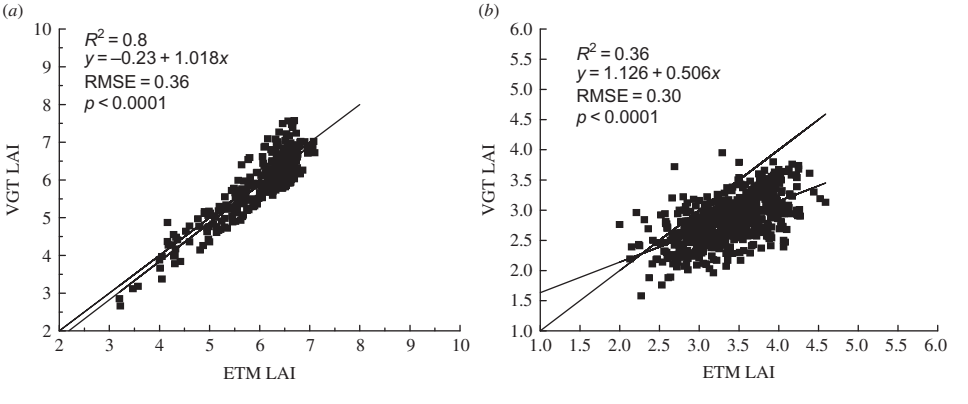

Figure 7. Comparisons of LAI values after fusion with those derived from ETM for: (*a*) Changbaishan; (*b*) Liping.
Notes: ETM, Enhanced Thematic Mapper; LAI, leaf area index; VGT, VEGETATION.

Table 4. Comparison of LAI mean values after fusion with those from ETM of Changbaishan and Liping (the data in the brackets are standard deviations).

Changbaishan			Liping		
Vegetation code	VGT LAI	ETM LAI	Vegetation code	VGT LAI	ETM LAI
2	6.11 (0.56)	6.26 (0.45)	1	2.71 (0.52)	3.33 (0.57)
5	6.6 (0.35)	6.56 (0.23)	4	2.67 (0.45)	3.32 (0.61)
6	4.49 (0.58)	4.75 (0.55)	11	2.57 (0.51)	3.15 (0.68)
13	4.73 (0.98)	4.94 (0.86)	13	2.21 (0.62)	2.71 (0.63)

Notes: The meaning of the vegetation code can be found in table 1. ETM, Enhanced Thematic Mapper; LAI, leaf area index; VGT, VEGETATION.

mean values for the above vegetation cover types are 6.11 ± 0.56, 6.6 ± 0.35, 4.49 ± 0.58 and 4.73 ± 0.98, respectively, which are very close to the ETM LAI mean values of 6.26 ± 0.45, 6.56 ± 0.23, 4.75 ± 0.55 and 4.94 ± 0.86. For the image of Liping, the improved LAI mean values are 2.71 ± 0.52, 2.67 ± 0.45, 2.57 ± 0.51 and 2.21 ± 0.62 for evergreen broadleaved forest, evergreen needle-leaved forest, evergreen shrub and herbaceous cover respectively, which are all lower than the ETM LAI mean values of 3.33 ± 0.57, 3.32 ± 0.61, 3.15 ± 0.68 and 2.71 ± 0.63. The larger systematic error at the Liping site may be due to the following reasons: (1) the landscape within the Liping county in southern China is heterogeneous due to intensive land management and small-scale topographical variations, and most pixels at 1 km resolution are mixed, causing errors in scaling from 30 m to 1 km resolution; and (2) the influence of topographical variations on remote sensing signals at Liping county would be considerable and is not corrected on the ETM and VGT images used in this study. These errors are much less at the Changbaishan Nature Reserve, where the landscape is more homogeneous and the topography varies at a larger scale. Additional work is still needed to reduce the scaling error and to consider the topographical effects.

3.6 *Comparison with field observations*

Although the direct comparison of satellite-generated LAI values at a coarse resolution with field observations in small plots is often not meaningful because of the mismatch in spatial resolution, we can still get some information from such comparisons when a large data set is used. The field LAI observations used for comparison for this purpose, obtained from the website http://www.daac.ornl.gov, were 794 plots with mature or nearly mature stand ages originally collected by Luo (1996) and Luo *et al.* (2002) from the ecological research plots in the Chinese literature and inventory plots. In the comparison, if there are three or more plots located at a forest pixel then the averaged ground observation is compared with the satellite data, and plots located within a non-forest pixel or less than three plots located within one pixel are excluded. Figure 8 shows that the correlations between field-observed LAI and satellite-derived LAI are improved after the data-model fusion. The R^2 increased from 0.14 to 0.23 and the RMSE decreased from 1.34 to 1.24 after data-model fusion. The p-values for the two regressions are both less then 0.0001. The reason for this is that many of the forest LAI measurements were obtained in the detected cloudy pixels in southern

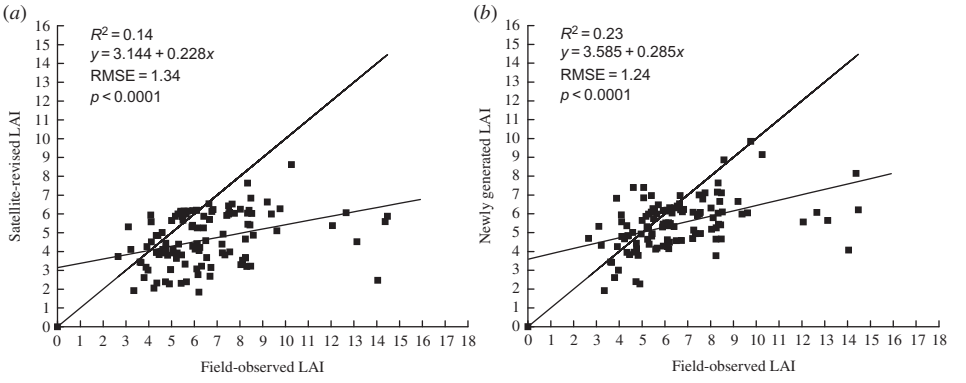

Figure 8. Comparison of satellite-based LAI values with field observations (Luo 1996, Luo *et al*. 2002), (*a*) before data-model fusion; (*b*) after fusion.

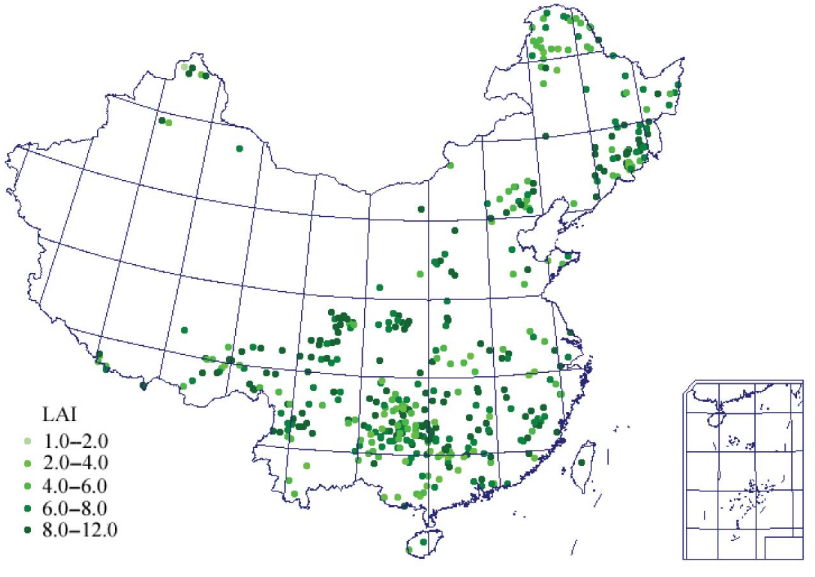

Figure 9. The locations of field observation sites with their LAI values.

China (figure 9), and LAI values in these cloudy pixels were increased after data-model fusion.

The mean values of field-observed LAI, the old satellite-based LAI and the improved LAI are shown in table 5. It shows that after fusion, the mean LAI values for deciduous and evergreen forests are closer to the field measurements than before.

4. Conclusions

The data-model fusion method used in this study is shown to be effective in improving the quality of the satellite-derived LAI through identifying the atmospherically affected data points and reconstructing the seasonal trajectory of LAI. In cloudless

Table 5. Comparison of the mean values of field-observed LAI with those of satellite-based LAI before and after improvement (the data in the brackets are standard deviations).

Vegetation code	Observation	Satellite-based LAI	Improved LAI
1	7.21 (2.99)	4.08 (1.25)	6.26 (1.02)
2	6.01 (1.61)	5.39 (0.98)	5.43 (1.04)
4	6.86 (2.44)	4.56 (1.44)	5.48 (1.35)
5	5.46 (1.44)	4.44 (1.50)	4.58 (1.54)

Note: The meaning of the vegetation code can be found in table 1.

pixels, the satellite-derived maximum LAI is relatively reliable in comparison with our ecological model AVIM2 used for LAI simulation, and it is the basis for determining the spatial LAI distribution. For deciduous cover types, the leaf-on and leaf-off dates observed by satellite are also reliable and can be used to calibrate AVIM2. However, the seasonal variations of coniferous forests observed by satellite sensors are often not reliable due to the variation in leaf pigment contents and the background vegetation or snow cover, while the seasonal variation of conifer LAI can be well-modelled. The model AVIM2, driven by environmental variables (soil and meteorology), has the advantage of stability and reasonability. Its results are not influenced by atmospheric effects, and thus can be used to reconstruct the LAI seasonal trajectory in cloudy pixels and for conifer pixels. It is therefore logical to use a model to improve satellite-derived LAI data through data-model fusion. It is demonstrated in this study that this fusion method considerably improved an existing LAI map series in comparison with several independent LAI data sets.

In this study, the data-model fusion method was not suitable for pixels of cropland and managed areas because (1) these pixels are highly influenced by human activity and there is insufficient information to judge if the abrupt reductions in the LAI seasonal trajectory were caused by atmospheric effects or human activity (such as harvests in double or multiple cropping systems), and (2) croplands were taken as grasslands in the AVIM2 simulation, which ignores the influence of human activity such as irrigation and fertilization. These shortcomings may be overcome in future studies when information on agricultural activity that affects LAI seasonal variations is available for large areas.

Acknowledgements

This research was supported by the Natural Science and Engineering Research Council of Canada Project 'Hyperspectral and multi-angle remote sensing of terrestrial ecosystems', the Chinese Academy of Sciences International Partnership Project 'Human activities and ecosystem changes' (CXTD-Z2005-1) and the Chinese Academy of Sciences Important Project 'Coupling vegetation dynamical model with land surface bio-geochemical process' (KZCX2-YW-219).

References

Alan, G.B., Black, T.A., Hogg, E.H., Kljun, N., Morgenstern, K. and Nesic, Z., 2004, Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. *Agricultural and Forest Meteorology*, **126**, pp. 237–255.

Asner, G.P., Scurlock, J.M.O. and Hicke, J.A., 2003, Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, **12**, pp. 191–205.

Bondeau, A., Kicklighter, D.W., Kaduk, J. and The Participants of The Potsdam Npp Model Intercomparison, 1999, Comparing global models of terrestrial net primary productivity (NPP): importance of vegetation structure on seasonal NPP estimates. *Global Change Biology*, **5**, pp. 35–45.

Buermann, W., Dong, J., Zeng, X., Myneni, R.B. and Dickinson, R.E., 2001, Evaluation of the utility of satellite-based vegetation leaf area index data for climate simulations. *Journal of Climate*, **14**, pp. 3536–3550.

Chen, J.M. and Black, T.A., 1992, Define leaf area index for non-flat leaves. *Plant Cell Environment*, **15**, pp. 421–429.

Chen, J.M., Deng, F. and Chen, M., 2006, Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 2230–2238.

Chen, J.M. and Leblanc, S.G., 1997, A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, pp. 1316–1337.

Chen, J.M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S.G., White, H.P., Hall, R.J., Peddle, D.R., King, D.J., Trofymow, J.A., Swift, E., Van Der Sanden, J. and Pellikka, P.K.E., 2002, Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment*, **80**, pp. 165–184.

Churkina, G., Tenhunen, J., Thornton, P., Falge, E.M., Elbers, J.A., Erhard, M., Grünwald, T., Kowalski, A.S., Rannik, Ü. and Sprinz, D., 2003, Analysing the ecosystem carbon dynamics of four European coniferous forests using a biogeochemistry model. *Ecosystems*, **6**, pp. 168–184.

Dan, L. and Ji, J., 2007, The surface energy, water, carbon flux and their intercorrelated seasonality in a global climate-vegetation coupled model. *Tellus B*, **59**, pp. 425–438.

Deng, F., Chen, J.M., Plummer, S., Chen, M. and Pisek, J., 2006, Algorithm for global leaf area index retrieval using satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 2219–2229.

Dickinson, R.E., Shaikh, M., Bryant, R. and Graumlich, L., 1998, Interactive canopies for a climate model, *Journal of Climate*, **11**, pp. 2823–2836.

Du, Z., Yang, Z. and Cui, X., 2001, A comparative study of leaf area index of five plant communities in a typical steppe region of Inner Mongolia. *Grassland of China*, **23**, pp. 13–18. (In Chinese.)

Ewert, F., 2004, Modelling plant responses to elevated CO_2: how important is leaf area index? *Annals of Botany*, **93**, pp. 619–627.

Farrar, J.F., 1992, The whole plant: carbon partitioning during development. In *Carbon partitioning within and between organisms*, C.J. Pollock and A.J. Gordon (Eds.), pp. 163–179 (Oxford: Bios Scientific Publishers Limited).

Foley, J.A., Costa, M.H., Delire, C., Ramankutty, N. and Snyder, P., 2003, Green surprise? How terrestrial ecosystems could affect earth's climate. *The Ecological Society of America*, **1**, pp. 38–44.

Gower, S.T., Kucharik, C.J. and Norman, J.M., 1999, Direct and indirect estimation of leaf area index, f_{APAR}, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, **70**, pp. 29–51.

Huang, M., Ji, J., Li, K., Liu, Y., Yang, F. and Tao, B., 2007, The ecosystem carbon accumulation after conversion of grasslands to pine plantations in subtropical red soil of south China. *Tellus B*, **59**, pp. 439–448.

Hutchinson, M.F., 1989, *A new objective method for spatial interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean*

solar radiation, temperature, precipitation and windrun, pp. 95–104. Division of Water Resources Technical Memorandum 89. (Clayton South, Australia: Commonwealth Scientific and Industrial Research Organisation (CSIRO)).

Ji, J., 1995, A climate-vegetation interaction model simulating the physical and biological processes at the surface. *Journal of Biogeography*, **22**, pp. 445–451.

Kergoat, L., 1998, A model for hydrological equilibrium of leaf area index on a global scale. *Journal of Hydrology*, **212–213**, pp. 268–286.

Lu, J. and Ji, J., 2002, A simulation study of atmosphere-vegetation interactions over the Tibetan Plateau. Part I: Physical fluxes and parameters. *Chinese Journal of Atmospheric Sciences*, **26**, pp. 111–126. (In Chinese.)

Lu, J. and Ji, J., 2006, A simulation and mechanism analysis of long-term variations at land surface over arid/semi-arid area in north China. *Journal of Geophysical Research*, **111**, pp. 1–19.

Luo, T.X., 1996, Patterns of net primary productivity for Chinese major forest types and their mathematical models. PhD thesis, Chinese Academy of Sciences.

Luo, T.X., Neilson, R.P., Tian, H., Vörösmarty, C.J., Zhu, H. and Liu, S., 2002, A model for seasonality and distribution of leaf area index of forests and its application to China. *Journal of Vegetation Science*, **13**, pp. 817–830.

Neilson, R.P., 1995, A model for prediction of continental-scale vegetation distribution and water balance. *Ecological Applications*, **5**, pp. 362–385.

Pisek, J., Chen, J.M. and Deng, F., 2007, Canada-wide validation of a new global leaf area index data set from SPOT-4 VEGETATION data. *Canadian Journal of Remote Sensing*, **33**, pp. 1–16.

Prentice, C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A. and Solomon, A.M., 1992, A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, pp. 117–134.

Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C. and Bowman, W.D., 1999, Generality of leaf trait relationships: a test across six biomes. *Ecology*, **80**, pp. 1955–1969.

Schulze, E., Kelliher, F.M., Korner, C., Lloyd, J. and Leuning, R., 1994, Relationships among maximum stomatal conductance, ecosystem surface conductance, carbon assimilation rate and plant nitrogen nutrition: a global ecology scaling exercise. *Annual Review of Ecology and Systematics*, **25**, pp. 629–660.

Sellers, P.J., Hall, F.G., Kelly, R.D., Black, A., Baldocchi, D., Berry, J., Ryan, M., Ranson, K.J., Crill, P.M., Lettenmaier, D.P., Margolis, H., Cihlar, J., Newcomer, J., Fitzjarrald, D., Jarvis, P.G., Gower, S.T., Halliwell, D., Williams, D., Goodison, B., Wickland, D.E. and Guertin, F.E., 1997, BOREAS in 1997: experiment overview, scientific results and future directions. *Journal of Geophysics Research*, **102**, pp. 28 731–28 770.

Tan, B., Hu, J., Huang, D., Yang, W., Zhang, P., Shabanov, N.V., Knyazikhin, Y., Nemani, R.R. and Myneni, R.B., 2005, Assessment of the broadleaf crops' leaf area index product from the terra MODIS Instrument. *Agricultural and Forest Meteorology*, **135**, pp. 124–134.

Vermote, E.F., Saleous, N.E. and Justice, C.O., 2002, Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment*, **83**, pp. 97–111.

Wang, P., Sun, R., Hu, J., Zhu, Q., Hou, Y., Li, L. and Chen, J.M., 2007, Measurements and simulation of forest leaf area index and net primary productivity in northern China. *Journal of Environmental Management*, **85**, pp. 607–615.

Williams, M. and Rastetter, E.B., 1999, Vegetation characteristics and primary productivity along an arctic transect: implications for scaling-up. *Journal of Ecology*, **87**, pp. 885–898.

Woodward, F.I., Smith, T.M. and Emanuel, W.R., 1995, A global land productivity and phytogeography model. *Global Biogeochemical Cycles*, **9**, pp. 471–490.

Xue, Y., Fennessy, M.J. and Sellers, P.J., 1996, Impact of vegetation properties on US summer weather prediction. *Journal of Geophysical Research*, **101**, pp. 7419–7430.

Yang, W., Shabanov, N.V., Huang, D., Wang, W., Dickinson, R.E., Nemani, R.R., Knyazikhin, Y. and Myneni, R.B., 2006, Analysis of leaf area index products from a combination of MODIS Terra and Aqua data. *Remote Sensing of Environment*, **104**, pp. 297–312.

Zhang, S., Peng, G. and Huang, M., 2004, The feature extraction and data fusion of regional soil textures based on GIS techniques. *Climatic and Environmental Research.* **9**, pp. 65–79. (in Chinese.)

Zheng, G., Chen, J.M., Tian, Q.J., Ju, W.M. and Xia, X.Q., 2007, Combining remote sensing imagery and forest age inventory for biomass mapping. *Journal of Environmental Management*, **85**, pp. 616–623.