Improvements of the MODIS terrestrial gross and net primary production global data set

Maosheng Zhao\textsuperscript{a,}\textsuperscript{*,} Faith Ann Heinsch\textsuperscript{a}, Ramakrishna R. Nemani\textsuperscript{b}, Steven W. Running\textsuperscript{a}

\textsuperscript{a}Numerical Terradynamic Simulation Group, Dept. of Ecosystem and Conservation Sciences, University of Montana, Missoula, MT 59812, USA
\textsuperscript{b}NASA Ames Research Center, Moffett Field, CA 94035, USA

Received 15 April 2004; received in revised form 12 December 2004; accepted 14 December 2004

Abstract

MODIS primary production products (MOD17) are the first regular, near-real-time data sets for repeated monitoring of vegetation primary production on vegetated land at 1-km resolution at an 8-day interval. But both the inconsistent spatial resolution between the gridded meteorological data and MODIS pixels, and the cloud-contaminated MODIS FPAR/LAI (MOD15A2) retrievals can introduce considerable errors to Collection4 primary production (denoted as \textit{C4 MOD17}) results. Here, we aim to rectify these problems through reprocessing key inputs to MODIS primary vegetation productivity algorithm, resulting in improved Collection5 MOD17 (here denoted as \textit{C5 MOD17}) estimates. This was accomplished by spatial interpolation of the coarse resolution meteorological data input and with temporal filling of cloud-contaminated MOD15A2 data. Furthermore, we modified the Biome Parameter Look-Up Table (BPLUT) based on recent synthesized NPP data and some observed GPP derived from some flux tower measurements to keep up with the improvements in upstream inputs. Because MOD17 is one of the down-stream MODIS land products, the performance of the algorithm can be largely influenced by the uncertainties from upstream inputs, such as land cover, FPAR/LAI, the meteorological data, and algorithm itself. MODIS GPP fits well with GPP derived from some flux tower measurements to keep up with the improvements in upstream inputs. Because MOD17 is one of the down-stream MODIS land products, the performance of the algorithm can be largely influenced by the uncertainties from upstream inputs, such as land cover, FPAR/LAI, the meteorological data, and algorithm itself. MODIS GPP fits well with GPP derived from some flux tower measurements to keep up with the improvements in upstream inputs. Because MOD17 is one of the down-stream MODIS land products, the performance of the algorithm can be largely influenced by the uncertainties from upstream inputs, such as land cover, FPAR/LAI, the meteorological data, and algorithm itself. MODIS GPP fits well with GPP derived from some flux tower measurements to keep up with the improvements in upstream inputs.

1. Introduction

Vegetation primary production is vital to human society not only because it provides essential materials, such as food, fiber and wood, but also because it creates environments suitable for human inhabitation. Primary production has received more attention in recent years because it is directly related to the global carbon cycle. The atmospheric concentration of CO$_2$ has increased by 31% since 1750 (IPCC, 2001) due to human activities, and the increased CO$_2$ in atmosphere can alter the energy balance of land surface resulting in changing climate system (Knutson et al., 2000; Scott et al., 2000). The terrestrial biosphere can sequester significant amounts of atmospheric CO$_2$ (Fan et al., 1998; Wofsy et al., 1993), and this process is largely influenced by increased atmospheric CO$_2$ and changing climate (Nemani et al., 2002, 2003; Schimel et al., 2000). The magnitude and the cause of C uptake are still uncertain (Barford et al., 2001), and understanding the carbon cycle on regional and global scales requires spatial and temporal monitoring of earth surface processes (Running et al., 1999). The MODerate Resolution Imaging Spectroradiometer (MODIS) is one of the primary global monitoring sensors on NASA Earth Observing System.
The current version of MODIS primary productivity product is C4 MOD17 and there are nearly 4 years of data. However, some shortcomings exist in the C4 MOD17 products. First, the C4 MOD17 operational process fails to account for the mismatching spatial resolution between a 1-km MODIS pixel and the corresponding 1°×1.25° meteorological data from the Data Assimilation Office (DAO). Secondly, the C4 MOD17 process produces GPP and NPP regardless of errors caused by contaminated or missing 8-day FPAR/LAI (MOD15A2) mainly due to cloud cover or sensor malfunction. These contaminated or missing MOD15A2 can introduce considerable error to 8-day MODIS GPP and therefore to annual NPP and GPP. Thirdly, the C4 MOD17 Biome Parameter Look-Up Table (BPLUT) was primarily developed at MODIS launch, and it was tested using different upstream inputs from those operationally being used (Running et al., 2000), which leads to different results. Finally, C4 MOD17 contains a meaningless annual quality assessment control (QC), in that a constant fill value is being used across all pixels because there were insufficient data at launch to establish meaningful annual QC values.

The objectives of this paper are: (1) to give a brief introduction to the MOD17 products and discuss uncertainties in them, (2) to describe the methods used to resolve known problems in C4 MOD17, and (3) to show some (Collection5) C5 MOD17 results from 2001 to 2003 as well as some validation results by comparison with GPP derived from 12 flux towers, Ecosystem Model–Data Intercomparison (EMDI) NPP data set, Multivariate ENSO Index (MEI), and global CO2 growth rate.

2. The MOD17 products

2.1. Brief introduction

The MOD17 algorithm provides the first operational, near-real-time calculation of global GPP and NPP products from EOS MODIS sensor. It has two subproducts: (1) MOD17A2, storing 8-day composite GPP, net photosynthesis (PsnNet) and corresponding QC, and (2) MOD17A3, which contains annual NPP and QC. These products are saved as formatted HDF EOS files (http://hdfeos.gsfc.nasa.gov) in a two-dimensional array with 1200 columns and 1200 rows in a Sinusoidal projection. In order to reduce data volume, the data are stored as signed or unsigned 2-byte integer data. Users need to be cautious about the data type, scale factor (gain), offset and units so that the data are restored correctly. These data are freely available to the public from the Numerical Terradynamic Simulation Group (NTSG) (http://www.ntsg.umt.edu) or the EROS Data Center Distributed Active Archive Center (EDC DAAC).

Users should note that the composite MOD17A2 is an 8-day summation of GPP and PsnNet, and annual GPP and NPP for MOD17A3 are annual summations of two variables. Additionally, PsnNet is defined as,

$$\text{PsnNet} = \text{GPP} - R_{ml} - R_{mr}$$

where $R_{ml}$ and $R_{mr}$ are maintenance respiration by leaves and fine roots, respectively. Growth respiration is not taken into account in 8-day PsnNet. Annual NPP is expressed as,

$$\text{NPP} = \sum_{i=1}^{365} \text{PsnNet} - (R_{mo} + R_{g})$$

where $R_{mo}$ is maintenance respiration by all other living parts except leaves and fine roots (e.g., livewood), and $R_{g}$ is growth respiration.

The QC field in MOD17A2 is inherited from MOD15A2, and each pixel’s QC value denotes the sensor and cloud
conditions, as well as the algorithm used to derive corresponding composite FPAR/LAI (Myneni et al., 2002). The MOD17A3 QC field is defined in Section 3.5.

2.2. Uncertainties from inputs and algorithm

There are three sources of MOD17 inputs. For each pixel, biome type information is derived from MODIS land cover products (MOD12Q1); daily meteorological data are derived from the DAO data set; and FPAR and LAI are obtained from MOD15A2. The uncertainties in MOD12Q1, DAO, MOD15A2, and the algorithm itself would all influence MOD17 results.

First, MOD12Q1 accuracies are falling in the range of 70–80%, and most “mistakes” are between similar classes (Strahler et al., 2002). For a pixel with misclassified land cover, a misuse of parameters from MOD17 BPLUT will occur, resulting in less reliable MOD17 results. Another problem is that the current 1-km MODIS global land cover classification unit may be too general for local application. Croplands, for example, are very diverse, yet the same set of parameters is applied indiscriminately to cropland everywhere, introducing large uncertainties for some crops in some regions.

Secondly, DAO is an assimilated meteorological data set not observed data. As a result, it may contain systematic errors in some regions. Uncertainties in meteorological data are largely responsible for the “unrealistic” negative NPP in some small regions. For these pixels located in harsh environments, overestimated temperature alone, for example, can be enough to produce negative NPP, because higher temperature results in higher respiration and lower GPP due to the higher Vapor Pressure Deficit (VPD) calculated by overestimated temperature. If respiration is greater than GPP, negative NPP will be produced. Because the MOD17 algorithm is very sensitive to meteorological inputs, a more detailed discussion of this aspect can be found elsewhere (Zhao et al., 2005).

For MOD15A2, the pixel-by-pixel comparison with the ground measurements has a poor correlation and retrieved LAI tends to be overestimated under most conditions (Wang et al., 2004). Comparison at the patch level can significantly improve the results, but retrieved LAI still tends to be higher (Wang et al., 2004). In the MOD17 algorithm, FPAR is directly related to assimilation and LAI is only related to respiration; an overestimated LAI from MOD15A2 may result in an underestimated NPP even if FPAR is relatively accurate. Although the temporal filling of unreliable FPAR/LAI greatly improves the accuracy of inputs, as discussed in Section 3.2 below, the filled values are artificial and contain uncertainties.

Finally, weaknesses in the MOD17 algorithm may lead to uncertainties in GPP/NPP. For example, there is still little known about the actual value of some parameters in the BPLUT, such as fine root maintenance respiration base and biomass ratio of fine root to leaf.

All three upstream inputs and the algorithm itself can introduce uncertainties to MOD17. In some regions, or during some seasons, these uncertainties may be large.

3. Improvements to C4 MOD17

3.1. Spatial interpolation of DAO

Because of the near-real-time and computer efficiency requirements of global MODIS products, daily meteorological data input must be provided in real time with a relatively coarse spatial resolution. MOD17 uses daily DAO as meteorological inputs. The core assimilation system for the DAO data is the Goddard EOS Data Assimilation System (GEOS-DAS), which uses General Circulation Model (GCM) outputs, boundary conditions (sea surface temperature, terrain, etc.) and surface observations to form a regular gridded meteorological data set (Atlas & Luches, 2000). The DAO version currently being used is GEOS402 and the original data set is global daily data with 1.0°×1.25° spatial resolution and a 3-h interval. From this, the MOD17 algorithm derives daily minimum temperature (Tmin), average temperature (Tavg), daytime average temperature (Tday), daily actual vapor pressure (AVP), daytime average VPD and total shortwave radiation (SWrad).

In the C4 MOD17 algorithm, each 1-km pixel falling into the same 1.00°×1.25° DAO grid cell will inherit the same meteorological data, creating a noticeable DAO footprint (Fig. 1a,c). Such treatment, on global or regional scale, may be acceptable, but at the local scale, especially for the terrain with topographical variation or located at relatively abruptly climatic gradient zones, it may cause large inaccuracies.

To enhance the meteorological inputs, we have interpolated coarse resolution DAO data down to the 1-km MODIS pixel level. The four DAO cells surrounding a given 1-km MODIS pixel are used in the interpolation algorithm (Fig. 2). The use of four DAO cells per 1-km MODIS pixel will not slow down the computational efficiency of MOD17 data stream, and reasonably assumes more minimal elevation variation within four DAO cells than any great number of DAO cells.

Although there are many formulae for non-linear spatial interpolation, for simplicity, we chose a cosine function, which constrains the result between 1 and 0 if the input variable is between 0 and π/2. As the function did not effectively eliminate DAO cell boundary lines in a MOD17 image, we tried to use its second, third and fourth power to increase the weighting value of the nearest DAO cell. Finally we found its fourth power form successfully removed DAO footprint even in region with abrupt climatic gradients:

\[
D_i = \cos^4\left(\left(\frac{\pi}{2}\right)^*(d_i/d_{\text{max}})\right) \quad i = 1, 2, 3, 4
\]
where \( D_i \) is the non-linear distance between the 1-km MODIS pixel and any one of the four surrounding DAO cells, \( d_i \) is the great-circle distance between the 1-km pixel and the same DAO cell, and \( d_{\text{max}} \) is the great-circle distance between the two farthest DAO cells of the four being used (Fig. 2). This ensures that \( D_i = 1 \) when \( d_i = 0 \), and \( D_i = 0 \) when \( d_i = d_{\text{max}} \). Based on the non-linear distance \( (D_i) \), the weighted value \( W_i \) can be expressed as

\[
W_i = \frac{D_i}{\sum_{i=1}^{4} D_i} \quad (4)
\]

and, for a given pixel, the corresponding smoothed variable, \( V_i \) (i.e., interpolated \( \text{Tmin}, \text{Tavg}, \text{Tday}, \text{AVP}, \text{SWrad} \)) is

\[
V = \sum_{i=1}^{4} \left( W_i \cdot V_i \right) \quad (5)
\]

VPD is derived from the interpolated \( \text{Tday} \) and \( \text{AVP} \).

Theoretically, this DAO spatial interpolation improves the accuracy of meteorological data for each 1-km pixel because it removes these abrupt changes from one side of a DAO boundary to the other, as used by C4 MOD17. Fig. 1 shows how this method makes embedded DAO cell effects disappear from C5 MOD17 image. The degree to which this interpolated DAO will improve the accuracy of meteorological inputs, however, is largely dependent on the accuracy of DAO data and local environmental conditions, elevation and weather patterns. To explore this question, we use observed daily weather data from the World Meteorological Organization (WMO) daily surface.
observation network (>5000 weather stations, Fig. 3) to compare changes in root mean squared error (RMSE) and correlation (COR) between the original and enhanced DAO data for 2000–2003.

As a result of the smoothing process, the percent of WMO stations with increased RMSE and increased COR are, on average, 29% (i.e., 71% stations have reduced RMSE) and 81%, respectively, relative to the original data as both compared with the observations (Fig. 4). For most of WMO stations, spatial interpolation reduced RMSE and increased COR, suggesting that the non-linear spatial interpolation considerably improves DAO inputs.

3.2. Temporal filling of unreliable MOD15A2

Despite the fact that the MOD15A2 product is an 8-day composite based on the maximum value of FPAR and corresponding LAI, there are still cloud-contaminated FPAR/LAI for some pixels during some periods (Myneni et al., 2002). In most of these cases, input FPAR/LAI tend to be underestimated due to cloud effects, and this will subsequently introduce considerable error to 8-day GPP/PsNet as shown in Fig. 5, and thereafter to annual GPP/NPP fields. The MOD15A2 product has two QC fields denoting cloud state, snow/ice presence, and the algorithm employed, called quality assessment fields. According to the MOD15A2 quality assessment scheme provided by Myneni et al. (2002), FPAR/LAI values retrieved by the main algorithm (i.e., Radiation Transfer process, denoted as RT) are most reliable, and those retrieved by the back-up algorithm (i.e., the empirical relationship between FPAR/LAI and NDVI) are less reliable because the back-up algorithm is employed mostly when cloud cover, strong atmospheric effects, or snow/ice are detected. In the case of snow cover, however, NDVI will be low simply because a large part of incoming solar radiation is reflected, and FPAR, therefore, would have a low value regardless of the algorithm used.

![Fig. 3. Distribution of WMO surface observation stations (n>5000) from 2000 to 2004 for evaluation of spatial interpolation.](image)

![Fig. 4. Percent of WMO stations with increased RMSE and COR by spatial interpolation in comparison to that by non-spatial interpolation from 2000 to 2004 years.](image)
To separate reliable and unreliable FPAR/LAI from the annual time series data array, for a given MODIS pixel, we have developed criteria based on the above information and our own analysis. For LAI, those periods without a cloud flag and derived by RT algorithm would be reliable; FPAR is slightly more complex. If no snow/ice is detected, FPAR derived by RT and without cloud flag would be considered reliable in the same way as employed for LAI. However, if

Fig. 5. Two examples on how temporal filling unreliable 8-day FPAR and LAI, and thereafter improved 8-day GPP and PsnNet for one MODIS 1-km pixel located in Amazon (AMZN) basin (lat=–5.0, lon=–65.0), and another located in Montana, USA (MTUS) (lat=46.754, lon=–113.829), respectively. The mark ‘B’ above some 8-day accepted FPAR in Montana figure denotes the FPAR was derived from MOD15 back-up algorithm under snow/ice cover but no cloud cover. MODIS land cover for AMZN is evergreen broadleaf forest (EBF), and evergreen needle forest (ENF) for MTUS. In 2002, for AMZN, the improved annual GPP and NPP are 2759 g C/m²/year and 914 g C/m²/year, respectively, in comparison to corresponding original annual GPP and NPP of 2252 g C/m²/year and 871 g C/m²/year. Similarly, for MTSU, the improved annual GPP and NPP are 695 g C/m²/year and 290 g C/m²/year against original annual GPP of 655 g C/m²/year and NPP of 294 g C/m²/year.
snow/ice are detected, only these FPAR without cloud flags would be chosen as reliable, regardless of whether the RT or back-up algorithm was used.

In addition to error introduced by contaminated FPAR/LAI, errors arise from missing periods of MODIS data due to malfunction of the MODIS sensor. In June 2001, for example, there are 2 or 3 periods of missing FPAR/LAI values, depending on the MODIS tile location. Given the contaminated or missing FPAR/LAI, it is necessary to reconstruct FPAR/LAI profile to enhance inputs to MOD17.

There are several methods widely used to reconstruct remotely sensed vegetation index time profiles, such as Fourier wave adjustment (Sellers et al., 1994), polynomial fitting (Karnieli et al., 2002), and piecewise logistic function fitting (Zhang et al., 2003). Unlike AVHRR, however, MODIS Level 3 land products have quality assessment fields as mentioned above, so there is no need to find and replace contaminated MODIS data by BISE. Additionally, we assume reliable FPAR/LAI are accurate enough for further use. The proposed mathematical fitting methods listed above, on the other hand, would eventually change reliable values. Hence, we use a simple linearly interpolation to fill unreliable or missing data based on reliable FPAR/LAI. The process entails two steps (see Fig. 5): (1) if the first (or last) 8-day FPAR/LAI is unreliable or missing, it will be replaced by the closest reliable 8-day value. This step ensures that the second step can be performed; (2) other unreliable FPAR/LAI will be replaced by linear interpolation of the nearest reliable value prior to it and the closest reliable value after it. If there are no reliable FPAR/LAI during the entire year, the annual maximum FPAR and corresponding LAI will be chosen from unreliable periods in current year, and they will be used as a constant value across the entire year. This exceptional case occurs for only a few pixels in tropical evergreen forest due to extreme cloud cover, and in tropical barren margin areas probably because of sparsely vegetated land cover or dust atmospheric conditions.

Fig. 5 illustrates how this temporal filling approach is applied to a MODIS pixel in the Amazon region where higher frequency and persistence of cloud cover exists, and another pixel in Montana, USA where severe snow and cloud cover usually occurs in winter, and low frequency of cloud cover appears in summer. For Montana site, before Julian day of 137 in 2002, due to cloud or severe snow/ice cover, all 8-day unreliable LAIs derived by MOD15 back-up algorithm are filled by reliable LAI value (1.1 m²/m²) for 8-day 137. But for FPAR, before Julian day of 137, some 8-day FPARs, derived by MOD15 back-up algorithm under snow cover but no cloud cover (with a mark ‘B’ above the value in Fig. 5), are accepted under our criterion described above. These FPARs (around 20%) tend to be lower corresponding to the LAI of 1.1 m²/m² according to Beer’s law (Monsi & Saeki, 1953) and light extinction coefficient of approximately 0.51 for evergreen needle forests (White et al., 2000), simply because of lower NDVI for snow/ice covered surface and higher LAI hidden beneath the snow/ice. This phenomenon did not exist in later 2002, which leads to the discrepancy in Montana FPAR between year begin (21%) and year end (65%). As depicted in Fig. 5, contaminated FPAR/LAI were greatly improved as the result of the filling process. Some unusual 8-day periods remain, which have lower FPAR/LAI with good QC or have higher FPAR/LAI with bad QC. In spite of this, QC fields from MOD15A2 are the only available source to distinguish between reliable and unreliable FPAR/LAI estimates. Improved FPAR/LAI inputs lead to enhanced MOD17 GPP and PsnNet. Under most conditions, 8-day composite GPP increases because the temporal filling process generally acts to increase FPAR. Changes in 8-day PsnNet, however, depend on the changes in both FPAR and LAI because improved FPAR/LAI lead to increases in not only GPP but also respiration (Heinsch et al., 2003). In addition, it is evident that the degree of the improvement is largely dependent on the region and season. Regions or seasons with higher cloud cover will have greater enhancements. As shown in Fig. 5, the site in the Amazon, for example, in 2002, had improved annual GPP and NPP of 2759 g C/m²/year (if no units mentioned, all annual GPP and NPP units are g C/m²/year hereafter) and 914, respectively, in comparison to the original GPP of 2252 and NPP of 871. The site with a lower frequency of cloud cover during the growing season will have fewer or almost no improvements, such as the site located in Montana, had improved GPP of 695 and NPP of 290 in comparison to the original GPP of 655 and NPP of 294.

Unfortunately, this temporal filling of FPAR/LAI can only be performed if an entire year of 8-day FPAR/LAI values is available. This limits the application of Collection5 MOD17A2 product to retrospective analysis. For MOD17A3 (i.e., annual GPP/NPP), however, it is applicable. For this reason, at the beginning of each calendar year, after all of the previous year’s 8-day MOD15A2 are available, we reprocess the previous year’s 8-day MOD17A2 based on the above two methods to improve the MOD17A2 product, and simultaneously produce an improved MOD17A3.

3.3 Combination of improved DAO and MOD15A2 data

Fig. 1 illustrates how the above improvements of meteorological data and FPAR/LAI inputs work together to enhance MOD17A3 NPP. The 4 subfigures are based on 4 schemes, (a) no spatial interpolation of DAO and no temporal filling of FPAR/LAI (Fig. 1a), (b) spatial interpolation of DAO with no temporal filling of FPAR/LAI (Fig. 1b), (c) no spatial interpolation of DAO with temporal filling of FPAR/LAI (Fig. 1c), and (d) spatial interpolation of DAO with temporal filling of FPAR/LAI.
Spatially interpolating DAO eliminates the unrealistic DAO footprint on the image and enhances MOD17 accuracy. Temporally filling unreliable FPAR/LAI enhances MOD15A2 inputs, and consequently improves MOD17, leading to increases in NPP, especially for these pixels located in regions of more frequent cloud cover. Annual MOD17A3 GPP and 8-day MOD17A2 GPP and PsnNet are similarly affected. These two approaches enhance MOD17 in different ways but produce more reliable MODIS primary production products.

### 3.4. Recalibration of the BPLUT

The C4 MOD17 BPLUT was parameterized using a global simulation of the general ecosystem model, BIOME-BGC, and calibrated to different meteorological data set and FPAR/LAI prior to launch (Running et al., 2000). The MOD17 primary production is sensitive to meteorological data (Zhao et al., 2005) and differences in FPAR/LAI data set inputs (Nemani et al., 2003). Furthermore, both DAO and MOD15A2 have been improved since launch, resulting in consistent and stable data sets. As a result, the BPLUT needed recalibrated with simulation of global GPP/NPP using complete 3-year MOD15A2 and DAO data as inputs.

Rather than directly using field measurements of NPP to calibrate the model, synthesized NPP data for different biomes (Olson et al., 2001; Roy et al., 2001; Clark et al., 2001) were used as baseline to recalibrate the BPLUT, because field NPP data were measured using various methods, in different years, stands and environments. Furthermore, most of these measurements were limited to aboveground NPP, on small scale, and on good stands with higher-than-average production than the average (Jarvis et al., 2001). Synthesized NPP data are inferred from these existing field data and current knowledge of carbon allocation patterns for different NPP components (Clark et al., 2001; Gower et al., 1997, 2001; Olson et al., 2001). Direct use of unevenly distributed field data to calibrate ecological parameters of regional or global model, therefore, is not as valuable as the use of synthesized NPP.

The recalibration process involves four steps. First, to speed the process, we build a 0.5° global land cover by choosing the dominant land cover type within each 0.5° region based on 1-km MOD12Q1. For FPAR/LAI, we perform temporal filling of unreliable periods to enhance the contaminated values, and then we create a global improved 0.5° 8-day FPAR/LAI data set from 1-km enhanced MOD15A2 by averaging these 1-km pixels with the same land cover as the dominant land cover. Second, we run global MOD17 at 0.5° based on this 0.5° MODIS land cover and 8-day MOD15A2, produce histograms and calculate mean annual GPP/NPP for the different land cover types. Third, we compared global mean MODIS GPP for the different land cover types with observed mean GPP data from 12 flux tower sites in 2001 (Heinsch et al., 2005; Turner et al., 2003a,b). Finally, we compared global mean NPP with a recent NPP summary for different biomes (Clark et al., 2001; Roy et al., 2001), and EMDI NPP data (Olson et al., 2001). For a given biome, if the differences between modeled and observed GPP and synthesized NPP are significant, parameters in BPLUT for the corresponding biome will be changed, repeating the last three steps until the differences are negligible. Compared with C4 BPLUT, Specific Leaf Area (SLA) and maximum VPD were changed for most biomes. Adjustment of VPD control effectively corrected GPP when compared to that derived from eddy flux observations. Altering the SLA can easily make MODIS NPP comparable to synthesized NPP data. The modified VPD and SLA values are within a typical range for a given land cover type (Hoffmann & Franco, 2003; Poorter, 2001; White et al., 2000). The C5 BPLUT can be found in MOD17 User’s Guide (Heinsch et al., 2003).

### 3.5. Addition of annual GPP and meaningful annual QC

We have added an annual GPP field to C5 MOD17A3 to free users from the labor required in getting annual GPP by summation of all 46 8-day GPP values from MOD17A2.

The 8-day MOD17A2 QC field is inherited from MOD15A2 in the same period. For annual MOD17A3 products, however, there was not enough data to define annual QC at launch and a constant value (33) was used across all vegetated pixels. For C5 MOD17, we have created a more meaningful annual GPP/NPP QC, expressed as

\[
QC = \left( \frac{NU\text{g}}{\text{TOTAL}\text{g}} \right) \times 100
\]

where \(NU\text{g}\) is the number of days during growing season with unreliable or missing MODIS LAI inputs, and \(TOTAL\text{g}\) is total number of days in the growing season. The growing season is defined as all days with \(T\text{min}\) above \(-8^\circ\)C, which is also used as the minimum temperature control on photosynthesis for all biomes in the BPLUT. Although there is respiration when \(T\text{min}\) below \(-8^\circ\)C and LAI is above 0, its magnitude is negligible due to both low LAI and low temperature during non-growing season days. Since the annual QC definition is limited to data in the growing season, the number of unreliable LAI values caused by snow cover will contribute much less to the QC than those caused by cloud cover. Therefore, the annual MOD17A3 QC mainly reveals how many growing days (%) use artificially filled FPAR/LAI due to cloud cover to calculate 8-day GPP and annual GPP/NPP. Users can infer the reliability of annual GPP/NPP based on the corresponding annual QC field.

The annual QC image in Fig. 6c shows the average annual QC from 2001 to 2003 and clearly demonstrates that tropical forest areas, western Europe, East Asian monsoon regions and the Pacific Northwest have higher values (less reliable annual GPP/NPP). Relatively dry
regions with biomes such as shrublands, grasslands or savannas have lower QC values and, therefore, more reliable annual GPP/NPP.

4. Results and validation of GPP and NPP

4.1. Results from 2001 to 2003

The 3-year (2001–2003) average annual global 1-km improved GPP, NPP and QC images are shown in Fig. 6. As expected, MODIS GPP and NPP have high values in areas covered by forests and woody savannas, especially in these tropical regions. Low NPP occurs in areas dominated by adverse environments, such as high latitudes with short growing seasons constrained by low temperatures, and dry areas with limited water availability. Global mean total GPP is 109.29 Pg C/year and NPP is 56.02 Pg C/year, ignoring barren land cover as defined by MOD12Q1. The annual QC image reflects the percent of filled FPAR/LAI during the growing season as discussed above.

The mean and standard deviation of GPP and NPP for different land cover types from the three-year averaged global 1-km data set are shown in Fig. 7, and the corresponding mean values and the ratio of NPP to GPP are listed in Table 1. Generally, annual GPP is about twice of NPP.
4.2. Validation

Validation of global 1-km GPP/NPP product is problematic because there are limited available field data in comparison to global coverage data sets. Ideally, the testing sites should cover as many as biomes types and climate regimes as possible. Eddy flux towers offer invaluable opportunities to validate process-based ecosystem models and satellite data because they measure carbon, water and energy exchange on a long-term and continuous basis (Baldocchi et al., 2001; Running et al., 1999). GPP can be derived from eddy flux measurements (Falge et al., 2002), and is being used to validate MODIS GPP (Heinsch et al., 2005; Turner et al., 2003b). At present, FLUXNET, a global network with over 250 towers, is operationally providing ground information for validating MODIS land products (http://www.daac.ornl.gov/FLUXNET/fluxnet.html), and 8-day MODIS GPP. Additionally, 7×7-km subsets of MODIS land surface products for the network are provided to researchers for validation efforts. These ongoing validation activities at surface flux site are vital for evaluation of model performance and refinements of MOD17 such as recalibrating BPLUT as mentioned above.

The use of eddy covariance flux tower data allows us to evaluate the MOD17 algorithm at the local scale, although care must be taken to ensure that participating towers are representative of the larger landscape surrounding them. Participants include members of the Ameriflux and Fluxnet communities. The towers are located across North America from the middle latitudes to the polar region and represent several land cover types, including forests, shrublands and grasslands. Direct comparison of MODIS annual GPP (MOD17A3) with observations for 37 site-years has resulted in a higher correlation and lower bias ($r^2=0.6993$, relative error=19%, unpublished data) than MODIS annual GPP calculated using tower meteorology ($r^2=0.595$, relative error=−2%). Overall, the average relative error of the difference between DAO and tower meteorology based GPP results is 27% (±45%), indicating that the DAO meteorology plays an important role in the accuracy of the GPP algorithm, and that this role is site-specific. These results suggest that the MODIS GPP compares favorably with observations, and the accuracy of MODIS primary production is very sensitive to meteorological inputs at the local level. A detailed comparison of MODIS GPP and flux tower data can be found elsewhere (Heinsch et al., 2005).

NPP is validated using the EMDI data set (Fig. 8), which was assembled from extensive worldwide NPP data.

<table>
<thead>
<tr>
<th>Landcover Types</th>
<th>GPP (g C/m²/year)</th>
<th>NPP (g C/m²/year)</th>
<th>Ratio (NPP/GPP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evergreen needle forests</td>
<td>818</td>
<td>2699</td>
<td>0.54</td>
</tr>
<tr>
<td>Evergreen broadleaf forests</td>
<td>441</td>
<td>1224</td>
<td>0.45</td>
</tr>
<tr>
<td>Deciduous needle forests</td>
<td>703</td>
<td>1366</td>
<td>0.43</td>
</tr>
<tr>
<td>Deciduous broadleaf forests</td>
<td>301</td>
<td>482</td>
<td>0.35</td>
</tr>
<tr>
<td>Mixed forests</td>
<td>1125</td>
<td>524</td>
<td>0.47</td>
</tr>
<tr>
<td>Closed shrublands</td>
<td>868</td>
<td>405</td>
<td>0.47</td>
</tr>
<tr>
<td>Open shrublands</td>
<td>336</td>
<td>1250</td>
<td>0.63</td>
</tr>
<tr>
<td>Woody savannas</td>
<td>212</td>
<td>705</td>
<td>0.56</td>
</tr>
<tr>
<td>Savannas</td>
<td>1121</td>
<td>627</td>
<td>0.56</td>
</tr>
<tr>
<td>Grasslands</td>
<td>396</td>
<td>259</td>
<td>0.65</td>
</tr>
<tr>
<td>Croplands</td>
<td>721</td>
<td>420</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Fig. 7. Three-year (2001–2003) mean and standard deviation of annual GPP, NPP for all vegetated land cover types delineated using MODIS land cover (full name and values for different land cover types are given in Table 1).

Fig. 8. Comparison of MODIS mean annual NPP with the EMDI NPP data set.
MODIS NPP agrees well with the EMDI data, despite the fact that this comparison contains large uncertainties, mainly associated with the discrepancy in land cover between MODIS and EMDI. The scatter at high NPP corresponds mainly to tropical forests, and these discrepancies may arise from two sources: (1) large uncertainties in DAO data inputs, because of increased errors from DAO data in tropical regions (Zhao et al., 2005), and (2) larger uncertainties in the EMDI tropical NPP data set, because field data not only are very scarce in this region, but also contain more uncertainties due to complex environmental characteristics in tropical forests (Clark et al., 2001).

Using the MOD17 algorithm and a 19-year monthly AVHRR data set, Nemani et al. (2003) found that, statistically, global terrestrial NPP is inversely related to atmospheric CO2 growth rate, and both are strongly associated with the ENSO cycle. To test this new generation of global MODIS primary production products, we compared global total GPP and NPP with the annual CO2 growth rate. 2001 was a La Nina year and 2002 and 2003 were weak El Nino years (http://www.cdc.noaa.gov/ENSO/enso.mei_index.html). Although 3-year short-term results limit our analysis, there are some interesting findings. First, MODIS global total GPP values (110.76 Pg C/year in 2001, 107.82 Pg C/year in 2002 and 107.50 Pg C in 2003) are comparable to the recent studies not only in magnitude but also in interannual variability. Large uncertainties may exist in tropical forests due to relatively large uncertainties from cloud contaminated FPAR/LAI, and meteorological inputs. The improvements of upstream input products and advances in knowledge of vegetation primary production will enable refinement of MOD17 to continue for many years. However, we feel these GPP and NPP data sets are now sufficiently mature to be used in a wide variety of applications, particularly where regular, spatially referenced measures of vegetation activity are desired (Running et al., 2004).

Acknowledgements

The work is funded by NASA Earth Science Enterprise MODIS contract NAS 5-31368. Daily surface observations from WMO stations were provided by National Climate Data Center. EMDI NPP data set was provided by Oak Ridge National Laboratory Distributed Active Archive Center. We thank the researchers who worked at AmeriFlux towers; and the addition of annual GPP and meaningful annual QC make MOD17A3 more convenient for the user community.

Validation campaigns are an important component of the product, because they can evaluate the performance of the model at different temporal and spatial scales. These activities help us understand the strengths and weaknesses of model and lead to refinements of the model in the future. These validations are not limited to MOD17, because the accuracies of the inputs have a great impact on MOD17. These campaigns need to test important upstream inputs, including land cover, FPAR/LAI and meteorological inputs to explore how uncertainties from inputs propagate to MODIS GPP and NPP. The 3-year MODIS GPP/NPP results compare favorably to observed GPP and NPP, and the global GPP/NPP are comparable to the recent studies not only in magnitude but also in interannual variability. Large uncertainties may exist in tropical forests due to relatively large uncertainties from cloud contaminated FPAR/LAI, and meteorological inputs.

The improvements of upstream input products and advances in knowledge of vegetation primary production will enable refinement of MOD17 to continue for many years. However, we feel these GPP and NPP data sets are now sufficiently mature to be used in a wide variety of applications, particularly where regular, spatially referenced measures of vegetation activity are desired (Running et al., 2004).

References


