Chin. Geogra. Sci. 2016 Vol. 26 No. 1 pp. 22–34
doi: 10.1007/s11769-015-0789-3

Springer Science Press
www.springerlink.com/content/1002-0063

# A Simple Method to Extract Tropical Monsoon Forests Using NDVI Based on MODIS Data: A Case Study in South Asia and Peninsula Southeast Asia

LIN Sen[1, 2], LIU Ronggao[1]

(1. *State Key Laboratory of Resources and Environmental Information System*, *Institute of Geographic Sciences and Natural Resources Research*, *Chinese Academy of Sciences*, *Beijing* 100101, *China*; 2. *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*)

**Abstract:** Distribution of monsoon forests is important for the research of carbon and water cycles in the tropical regions. In this paper, a simple approach is proposed to map monsoon forests using the Normalized Difference Vegetation Index (NDVI) derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) data. Owing to the high contrast of greenness between wet season and dry season, the monsoon forest can be easily discriminated from other forests by combining the maximum and minimum annual NDVI. The MODIS-based monsoon forest maps ($MOD_{MF}$) from 2000 to 2009 are derived and evaluated using the ground-truth dataset. The $MOD_{MF}$ achieves an average producer accuracy of 80.0% and the Kappa statistic of 0.719. The variability of $MOD_{MF}$ among different years is compared with that calculated from MODIS land cover products (MCD12Q1). The results show that the coefficient of variation of total monsoon forest area in $MOD_{MF}$ is 7.3%, which is far lower than that in MCD12Q1 with 24.3%. Moreover, the pixels in $MOD_{MF}$ which can be identified for 7 to 9 times between 2001 and 2009 account for 53.1%, while only 7.9% of MCD12Q1 pixels have this frequency. Additionally, the monsoon forest areas estimated in $MOD_{MF}$, Global Land Cover 2000 (GLC2000), MCD12Q1 and University of Maryland (UMD) products are compared with the statistical dataset at national level, which reveals that $MOD_{MF}$ has the highest $R^2$ of 0.95 and the lowest RMSE of 14 014 $km^2$. This algorithm is simple but reliable for mapping the monsoon forests without complex classification techniques.

**Keywords:** monsoon forest; Moderate Resolution Imaging Spectroradiometer (MODIS); Normalized Difference Vegetation Index (NDVI) amplitude; threshold; classification

**Citation:** Lin Sen, Liu Ronggao, 2016. A simple method to extract tropical monsoon forests using NDVI based on MODIS data: a case study in South Asia and Peninsula Southeast Asia. *Chinese Geographical Science*, 26(1): 22–34. doi: 10.1007/s11769-015-0789-3

## 1 Introduction

Monsoon forests, also called tropical seasonal forests, are forests that have a dry season during which the trees partly or totally shed their leaves (Eamus, 1999). They distribute widely in South Asia and Peninsula Southeast Asia, Sub-Saharan Africa, South America and Central America (Miles *et al*., 2006). The monsoon forests, which cover over 40% of tropical forests area (Murphy and Lugo, 1986; FAO, 2001), not only sustain huge population by providing fuel and wood productions (Trejo and Dirzo, 2000; Steininger *et al*., 2001), but also play a significant role in controlling the concentration of $CO_2$ in atmosphere (Eamus, 1999; Kalacska *et al*., 2005). The high variation of monsoon forests between wet season and dry season has an important impact on the photosynthetic activity and transpiration (Malhi *et al*., 2002; Vourlitis *et al*., 2002), and contributes to sea-

Received date: 2015-03-24; accepted date: 2015-05-21
Foundation item: Under the auspices of National Natural Science Foundation of China (No. 41171285), Research and Development Special Fund for Public Welfare Industry (Meteorology) of China (No. GYHY201106014)
Corresponding author: LIU Ronggao. E-mail: liurg@igsnrr.ac.cn

sonal changes in tropical net primary productivity and water cycle. Meanwhile, the increasing extreme climate and weather events, such as floods and droughts, can change the length of forest growing season and affect the energy exchanges and biogeochemical cycles between forests and atmospheric layer. Therefore, the monsoon forests are also sensitive to global changes (Yoshifuji *et al*., 2006). An accurate description of the spatial distribution of tropical monsoon forests is the first step to many researches on global carbon cycle (Foley *et al*., 1996) and ecological processes (Peel *et al*., 2001). As most trees in monsoon forests shed leaves due to the lack of water in dry season, the forests shift seasonally from carbon sinks to carbon sources (Valentini *et al*., 2000; Vourlitis *et al*., 2004). At the same time, water deficit limits plant transpiration process during dry season and less water will be consumed from groundwater resources (Tanaka *et al*., 2008; Bohlman, 2010). Therefore, the deciduousness probably plays a significant role in carbon balance and water-use efficiency of tropical ecosystems (Tang *et al*., 2014).

Remote sensing technology has been applied to map regional land cover since 1980s (Tucker *et al*., 1985; Townshend *et al*., 1987). Up to now, many land cover products have been produced using satellite data, including National Oceanic and Atmospheric Administration-Advanced Very High Resolution Radiometer (NOAA-AVHRR) (Hansen *et al*., 2000; Loveland *et al*., 2000), TERRA/AQUA-Moderate Resolution Imaging Spectroradiometer (MODIS) (Friedl *et al*., 2002; 2010), Systeme Probatoire d′Observation de la Terre (SPOT)-VEGETATION (Bartholome and Belward, 2005), ENVISAT-Medium Resolution Imaging Spectrometer (MERIS) (Bicheron *et al*., 2008; Bontemps *et al*., 2011) and Landsat-Thematic Mapper (TM)/Enhanced Thematic Mapper Plus (ETM+) (Gong *et al*., 2013). In these products, monsoon forests were classified as deciduous broadleaf forests (DBF) with temperate deciduous forests. Nevertheless, the characteristic of monsoon forests is different from that of temperate deciduous forests. Defoliation process of the latter is associated with both rainfall and temperature, while deciduousness of monsoon forests is mainly determined by rainfall as the temperature in tropics is above the forest growth threshold over the whole year (Wright and Cornejo, 1990; Eamus, 1999). Given that this difference could lead to different phenology, growth patterns or forest behaviors between monsoon forests and temperate deciduous forests, and it might not be appropriate to cluster two forest types as a single type, which could partly account for the relatively low accuracy of DBF in global classifications (Herold *et al*., 2008). In recent years, there have been several attempts to extract monsoon forests as a specific land-cover type. They were mostly developed based on the unique feature of monsoon forests that the forests have high contrast of greenness between wet season and dry season. For instance, Krishnaswamy *et al*. (2004) used two NDVI change vectors corresponding to NDVI amplitude, mean NDVI and coefficient of variation as classification tree inputs to extract monsoon forests in Southern India. Htun *et al*. (2011) also demonstrated the usefulness of NDVI amplitude in classifying monsoon forests in Myanmar. Both studies used the tree classifier to generate the classification rules and proved NDVI amplitude to be effective in boosting the classification accuracy of monsoon forests. However, both studies selected single satellite imagery to represent the dry season or the wet season. As is known, the monsoon climate influence different regions during different periods and thus occurrences of minimum NDVI in forests are spatially different, which could not be captured in a snap-shot. It is difficult and time-consuming to manually choose specific imagery to represent wet season or dry season in large-area classification for different regions. In addition, the dry season deciduousness of monsoon forests can be short or long (Ito *et al*., 2008). Using single imagery in classification might omit the information of some forests with short leaf-fall periods. To improve the accuracy, most classification algorithms employ multi-temporal features from time series data as their inputs (Homer *et al*., 2004; Friedl *et al*., 2010; Jia *et al*., 2014; Zhu and Woodcock, 2014), especially in the vegetation classification, vegetation growth information could be well depicted by time-series profiles of vegetation indexes and various vegetation types could be detected based on their unique profile characteristics (Lenney *et al*., 1996; Xiao *et al*., 2005; Brown *et al*., 2013).

This study aimed to propose a simple approach using NDVI amplitude to extract monsoon forests, which could be more applicable and efficient in large-area classifications. In this method, the NDVI amplitude was calculated by multi-temporal features of maximum and minimum annual NDVI, which were composited using the whole year MODIS observations. In this way, the composites can get the real NDVI amplitude and eliminate the effect of temporal difference of deciduousness in large-scale classification. Specifically, this algorithm will not employ any classifiers, which could simplify the process of classification. Moreover, South Asia and Peninsular Southeast Asia was selected as the case study area for a great amount of monsoon forests distributing in this region.

The data employed and the procedures of algorithm are illustrated in section 2 of the paper. The extracted maps and the analysis of the results will be shown in section 3. The limitation and potential improvements of the algorithm are discussed in section 4. Finally, the conclusions are given in section 5.

## 2 Materials and Methods

### 2.1 Study area

The study area (5.6°–35.6°N, 68.0°–109.5°E) covers 10 countries in South Asia and Peninsula Southeast Asia, including India, Nepal, Bhutan, Bangladesh, Sri Lanka, Myanmar, Laos, Thailand, Cambodia, Vietnam (Fig. 1).

The climate of this area is influenced by the prevailing South Asian Monsoon. The summer monsoonal wind picks up much moisture from the Indian Ocean and brings a humid rainy season. The summer rain accounts for more than 80% of the whole annual rainfall in this region. The winter monsoon is not as powerful as the summer monsoon and carries little moisture. The dry season, usually with precipitation less than 100 mm, could last for several months.

According to the Forest Resources Assessment 2010 (FAO, 2010), forests cover 32% of the land area in this region and most of the forests are tropical and subtropical species. As monsoon forests are more close to human habitations with great economic values, this area have suffered from deforestation. Although the declining rate of forest area has slowed down in 2000s, the extent of forests still shows negative trends and uncontrolled deforestation keeps an alarming rate in most countries. Monsoon forests in this region occupy a significant part of total forest area and represent the most typical monsoon forests around the world (Ruangpanit, 1995).

### 2.2 MODIS data

The satellite data used in this study is MODIS land surface reflectance product (MOD09A1), which contains 7 bands at 500 m resolution. The data are provided in an 8-day gridded product using the Sinusoidal projection. MOD09A1 has been corrected for atmospheric gases and aerosols (Vermote *et al.*, 2002) and includes quality control flags to show the various image states. Standard MODIS land products are organized in a tile system and each tile covers an area of 1200 km × 1200 km. In this research, 16 tiles of MOD09A1 data for 2000–2009 were used.

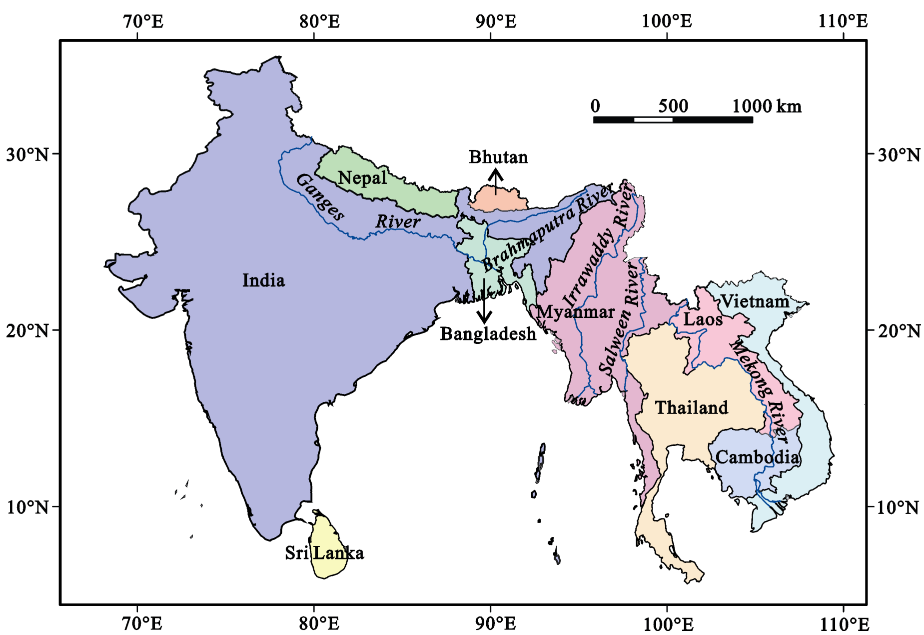


**Fig. 1** Spatial distribution of 10 countries in South Asia and Peninsula Southeast Asia

For each 8-day composite, a cloud mask proposed by Liu and Liu (2013) was employed to remove the residual clouds. Then NDVI was derived from Band 1 (red: 620–670 nm) and Band 2 (near infrared, NIR: 841–875 nm) of MOD09A1 by Equation (1):

$$NDVI_{\text{MODIS}} = \frac{Band2 - Band1}{Band2 + Band1} \tag{1}$$

Three NDVI-based metrics were prepared, including annual maximum NDVI (referred to as MaxNDVI), annual minimum NDVI (referred to as MinNDVI) and annual amplitude of NDVI (referred to as NDVI amplitude), which is the difference between MaxNDVI and MinNDVI.

(1) Compositing annual maximum NDVI

The MaxNDVI is defined as the annual maximum NDVI value, which represents the greenest status of vegetation. MaxNDVI was composited by selection of the three largest NDVI which calculated from the MOD09A1 data. And then the three largest NDVI were averaged to the mean value as the maximum NDVI in this whole year.

(2) Compositing annual minimum NDVI

MinNDVI is defined as the annual minimum NDVI value, which represent the brownest status of vegetation. MinNDVI was composited from the whole year MOD09A1 data with selection of the minimum of a spectral index named as the Brown Vegetation Index (BVI). The residual contamination in MOD09A1 makes it fail to directly select the real minimum NDVI (Delbart *et al*., 2005), while the Brown Vegetation Index (BVI) could avoid cloud/snow contamination and the lower BVI corresponds well to the lower contamination-free NDVI observations. Therefore, BVI is taken as a proxy to composite the minimum vegetation NDVI data. BVI was derived from band 4 (green: 545–565 nm) and band 7 (short-wave infrared, SWIR: 2105–2155 nm) of MOD09A1 by Equation (2):

$$BVI_{\text{MODIS}} = \frac{Band4 - Band7}{Band4 + Band7} \tag{2}$$

### 2.3 Algorithm for extracting monsoon forests

The algorithm is based on the fact that NDVI amplitude of monsoon forests is larger than that of tropical evergreen forests (Fig. 2). NDVI amplitude reflects the difference of two extreme greenness conditions of forests between wet season and dry season, which are respecttively represented by the annual maximum and minimum NDVI. In some regions, the annual maximum NDVI values of monsoon forests are even as high as those of evergreen forests, but the annual minimum NDVI values of monsoon forests are much lower.

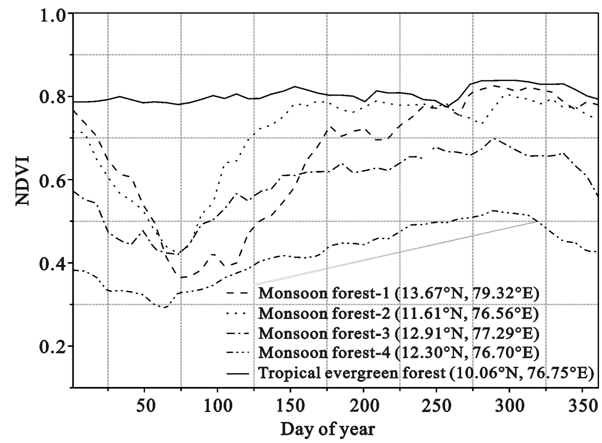


**Fig. 2** MODIS-derived Normalized Difference Vegetation Index (NDVI) seasonal profiles in 2000 for tropical evergreen forests and monsoon forests at 5 different plots in the study area

The method to extract monsoon forests from the whole forests includes two steps. Firstly, the non-forests land covers pixels, including grasslands, croplands and shrublands, are excluded using the mean NDVI value. Secondly, the classification threshold of NDVI amplitude between monsoon forests and other forests is determined by visual observation.

## 3 Results

### 3.1 Spatial distribution of monsoon forests for 2000 to 2009

The extracted maps of monsoon forests (referred to as $MOD_{MF}$) for 2000–2009 are depicted in Fig. 3. As they are shown, monsoon forests distribute widely but unevenly in all countries in South Asia and Peninsula Southeast Asia. Large areas of monsoon forests concentrate in several major river regions, for examples, Ganges River in India, the confluence of the Ganges River and the Brahmaputra River in Bangladesh, Irrawaddy River and Salween River in Myanmar, Mekong River across Myanmar, Laos, Thailand, Cambodia and Vietnam. The characteristics of the spatial distributions imply the transitional attributes of monsoon forests. They tend to occur at the edges of evergreen forests, such as in Peninsula Southeast Asia and southwestern India. There also exist some monsoon forests in the valleys, where display a more heterogeneous pattern, such as in eastern India (Fig. 3).

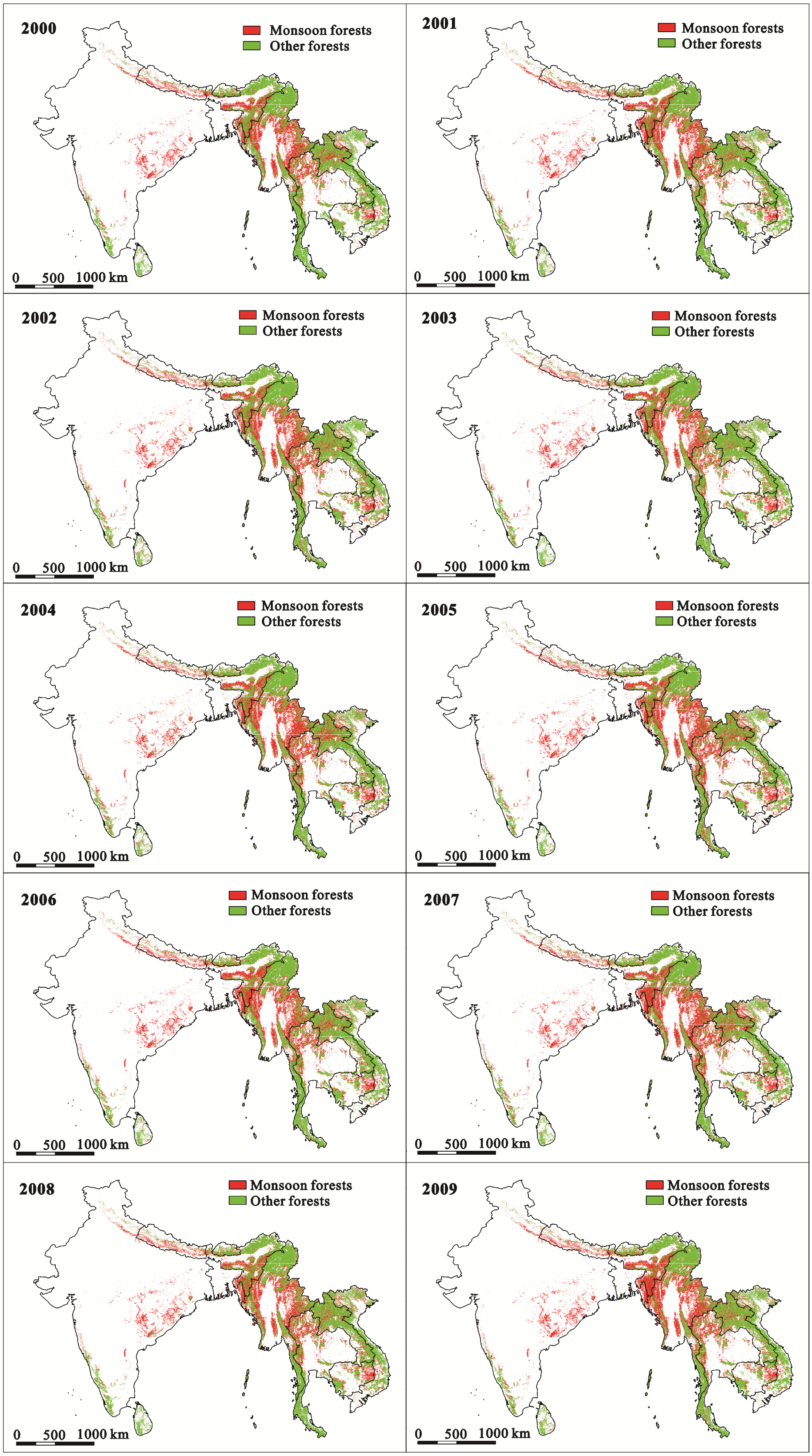


**Fig. 3** Spatial patterns of MODIS-derived monsoon forests ($MOD_{MF}$) in South Asia and Peninsula Southeast Asia between 2000 and 2009

In terms of subregional distributions, South Asia monsoon forests mainly distribute in eastern and western Ghats and northeastern India. Lots of monsoon forests also distribute in the northern and northeastern India where many rivers flow and some plains extend. In Bangladesh and Sri Lanka Island, there exist a small amount of monsoon forests. As for Peninsula Southeast Asian countries, Myanmar keeps 54% of monsoon forests in this area and other countries each accounts for parts. Several important south-north direction rivers, as mentioned above, cut the ranges as patches, which shapes the pattern of most monsoon forests region as the similar south-north direction. Overall, the area of monsoon forests in Peninsula Southeast Asian countries is 362 794 km$^2$ in 2009, which is far larger than that in South Asia with 222 333 km$^2$.

### 3.2 Classification threshold sensitivity analysis

The composites of MaxNDVI, MinNDVI and NDVI amplitude for 2001 are displayed in Fig. 4. The classification threshold of NDVI amplitude for extracting monsoon forests is between 0.17 and 0.21, which is obtained by direct visual observation from the composites of NDVI amplitude between 2000 and 2009. The final threshold is determined as 0.19, that is to say, a pixel in the forest mask would be detected as the monsoon forest if its NDVI amplitude is larger than 0.19, otherwise, it would be labeled as other forests.

The visual observation method to decide the threshold is based on the assumption that the derived monsoon forest area is not too sensitive to the classification threshold. Table 1 shows the variation of the extracted pixels in 2000 for the threshold ranging from 0.17 to 0.21. It is seen that the bias of 0.19 within range of ±0.01 NDVI amplitude units is between 5% and 6%.

### 3.3 Comparison of classification stability

The stability, here, means the consistency of derived monsoon forests among different years. Two statistical comparisons were performed to compare the stability of the long-term $MOD_{MF}$ with that of MODIS land cover product (MCD12Q1) since there is an overlapped period during 2001and 2009 between the two datasets.

The derived monsoon forests areas in $MOD_{MF}$ and MCD12Q1 are assembled in Table 2. The coefficient of variance (CV) was used to compare the variation of monsoon forests areas estimated in $MOD_{MF}$ with that of

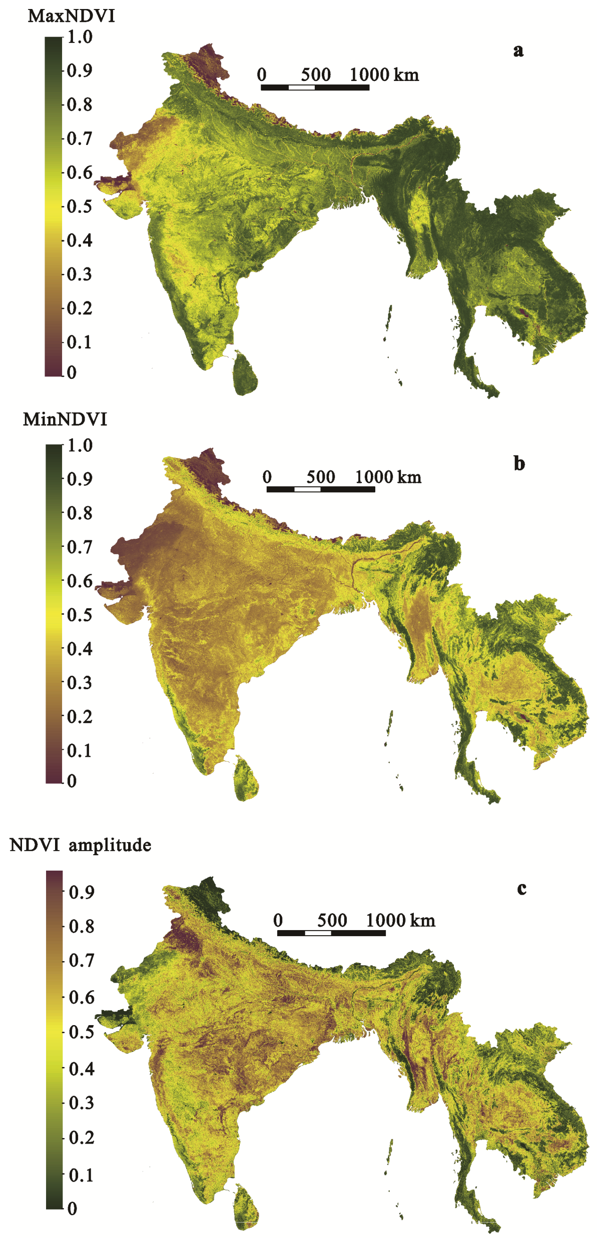


**Fig. 4** Composites of annual maximum NDVI value (MaxNDVI) (a), annual minimum NDVI value (MinNDVI) (b) and NDVI amplitude (c) using MOD09A1 product in 2001

**Table 1** Numbers of derived monsoon forest pixels in 2000 using different NDVI amplitude thresholds ranging from 0.17 to 0.21

| NDVI amplitude threshold | 0.17 | 0.18 | 0.19 | 0.20 | 0.21 |
|---|---|---|---|---|---|
| Number of Pixels | 2 672 554 | 2 532 840 | 2 384 024 | 2 266 162 | 2 139 388 |

MCD12Q1. The CV is also known as relative standard deviation and defined as the radio of the standard deviation (SD) to the mean. It is better than SD to compare

**Table 2** Monsoon forests areas derived from satellite-based datasets and a collected national statistical monsoon forests areas (referred to as $NSA_{MF}$) for 10 countries (km$^2$)

| Country | UMD | GLC2000 | MCD12Q1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1993 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
| Nepal | 723 | 24 572 | 5183 | 2733 | 1741 | 1660 | 2073 | 2429 | 2798 | 3681 | 2333 |
| India | 44 512 | 480 788 | 62 867 | 43 446 | 41 997 | 40 057 | 43 397 | 45 324 | 53 662 | 4672 | 36 060 |
| Sri Lanka | 1269 | 14 758 | 2332 | 1639 | 963 | 1163 | 1243 | 2301 | 1513 | 1133 | 302 |
| Bangladesh | 220 | 336 | 4437 | 3074 | 2396 | 2468 | 3221 | 4502 | 4170 | 3946 | 2938 |
| Bhutan | 5726 | 215 | 1312 | 601 | 527 | 506 | 593 | 600 | 726 | 886 | 832 |
| Cambodia | 2326 | 31 104 | 4173 | 1669 | 877 | 600 | 966 | 1397 | 1288 | 1400 | 1099 |
| Laos | 32 103 | 6074 | 2581 | 1451 | 865 | 661 | 1805 | 4383 | 4456 | 3594 | 2461 |
| Myanmar | 49 004 | 95 415 | 45 076 | 29 166 | 22 808 | 19 026 | 34 210 | 42 593 | 49 558 | 41 259 | 33 223 |
| Thailand | 8697 | 51 431 | 10 258 | 5574 | 3218 | 1973 | 4832 | 9063 | 11 482 | 11 208 | 8703 |
| Vietnam | 13 072 | 5956 | 4437 | 2134 | 1347 | 1881 | 3049 | 5640 | 4644 | 5230 | 3110 |
| Total | 157 652 | 710 649 | 142 655 | 91 487 | 76 739 | 69 996 | 95 389 | 118 232 | 134 297 | 118 810 | 91062 |

| Country | $MOD_{MF}$ | | | | | | | | | | $NSA_{MF}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | |
| Nepal | 26 283 | 26 539 | 23 170 | 22 679 | 22 883 | 24 807 | 26 810 | 23 192 | 25 372 | 26 611 | 26 696[a] |
| India | 150 260 | 161 735 | 141 173 | 148 807 | 150 037 | 150 408 | 162 738 | 160 161 | 140 504 | 174 895 | 162 170[b] |
| Sri Lanka | 1511 | 2572 | 3307 | 1218 | 2983 | 1952 | 2001 | 1769 | 1245 | 2290 | 2439[c] |
| Bangladesh | 8824 | 12 172 | 9717 | 10 829 | 11 111 | 11 628 | 12 504 | 12 810 | 11 629 | 13 026 | 1125[d] |
| Bhutan | 4863 | 6027 | 3843 | 4049 | 3117 | 4247 | 5158 | 3615 | 4415 | 5511 | 4400[e] |
| Cambodia | 19 820 | 25 835 | 28 921 | 28 516 | 30 076 | 33 271 | 27 983 | 31 594 | 29 647 | 28 161 | 43 012[c] |
| Laos | 33 545 | 31 057 | 32 974 | 21 327 | 43 426 | 50 307 | 37 766 | 49 743 | 32 197 | 33 188 | 25 810[b] |
| Myanmar | 174 961 | 176 261 | 175 366 | 163 654 | 189 871 | 190 496 | 164 694 | 197 657 | 165 055 | 194 986 | 151 465[c] |
| Thailand | 68 337 | 69 696 | 84 210 | 62 721 | 86 469 | 99 436 | 74 223 | 91 746 | 74 440 | 80 476 | 82 859[c] |
| Vietnam | 23 337 | 25 964 | 27 834 | 25 879 | 35 470 | 43 261 | 29 533 | 40 564 | 33 541 | 25 983 | 19 510[b] |
| Total | 511 741 | 537 858 | 530 515 | 489 679 | 575 443 | 609 813 | 543 410 | 612 851 | 518 045 | 585 127 | 519 486 |

Notes: although the MODIS pixel resolution is nominally 500 m, the actual pixel size is 463.312714 m. $MOD_{MF}$ is monsoon forest product produced in this study; MCD12Q1 is MODIS land cover product; GLC2000 is Global Land Cover 2000 product; UMD is University of Maryland land cover product; $NSA_{MF}$ is a collected national statistical monsoon forests areas. Data sources for national statistical areas of monsoon forests: a, The value was estimated from the Sal forests (a dominant local deciduous tree species) cover from Lillesø *et al*. (2005) and forest area from FAO (2010); b, Collins *et al*. (1991); c, FAO (1998); d, Bangladesh Forest Department (http://www.bforest.gov.bd/index.php/forest-category/tropical-moist-deciduous-forests); e, Bhutan Trust Fund for Environment Conservation (http://www.bhutantrustfund.bt/biodiversity)

datasets with different means. As the Table 3 shows, CV of total monsoon forest areas in $MOD_{MF}$ is 7.3%, which is far lower than that in MCD12Q1 of 24.3%. Mean of CVs in MCD12Q1 is 41%, nearly triple that in $MOD_{MF}$ of 14.5%. At the national level, CVs of monsoon forests in $MOD_{MF}$ were lower than those in MCD12Q1 for all countries.

To evaluate the stability of the algorithm, another approach is also proposed to count the frequency of one pixel labeled as monsoon forest during 2001 and 2009. This method is based on the assumption that most forest land covers are relatively static and stable; that is, if one pixel was labeled as the monsoon forest by a classification algorithm, it should be detected during a long time period. The more pixels with high frequency there are in a product, the more robust the classification algorithm is. In this comparison, all the pixels labeled as monsoon forests during 2001 and 2009 were selected to record their frequency of labeling as the monsoon forest in $MOD_{MF}$ and MCD12Q1. The frequency count was divided into low interval (1–3), middle interval (4–6) and high interval (7–9). Under the assumption mentioned above, the more pixels with frequency in high interval, and the more stable the algorithm is. As shown in Fig. 5, the proportion of high frequency of $MOD_{MF}$ is about 53.1%, which is far higher than the proportion in MCD12Q1 of 7.9%. The percentage of low frequency of MCD12Q1 reaches 74.9% and it means that almost three fourths of monsoon pixels were labeled discontinuously between 2001 and 2009 in MCD12Q1. By contrast, the percentage of low interval in $MOD_{MF}$ is as low as 29.6%.

**Table 3** Coefficient of variance of monsoon forest areas during 2001 and 2009 for $MOD_{MF}$ and MCD12Q1

| Country | $MOD_{MF}$ | | | MCD12Q1 | | |
|---|---|---|---|---|---|---|
| | Mean ($km^2$) | SD ($km^2$) | CV (%) | Mean ($km^2$) | SD ($km^2$) | CV (%) |
| Nepal | 24 835 | 1621 | 6.5 | 2737 | 1103 | 40.3 |
| India | 154 072 | 10 134 | 6.6 | 45 920 | 7963 | 17.3 |
| Sri Lanka | 2085 | 670 | 32.1 | 1399 | 642 | 45.9 |
| Bangladesh | 11 425 | 1282 | 11.2 | 3461 | 820 | 23.7 |
| Bhutan | 4485 | 856 | 19.1 | 731 | 254 | 34.8 |
| Cambodia | 28 382 | 3450 | 12.2 | 1497 | 1054 | 70.4 |
| Laos | 36 553 | 8521 | 23.3 | 2473 | 1424 | 57.6 |
| Myanmar | 179 300 | 12 348 | 6.9 | 35 213 | 10 317 | 29.3 |
| Thailand | 79 175 | 10 834 | 13.7 | 7368 | 3547 | 48.1 |
| Vietnam | 31 137 | 6448 | 20.7 | 3497 | 1551 | 44.3 |
| Total | 551 448 | 40 103 | 7.3 | 104 296 | 25 325 | 24.3 |
| Mean of CVs | – | – | 14.5 | – | – | 41.0 |

Notes: mean is average derived monsoon forests areas; SD is standard deviation; CV stands for coefficient of variance; '–' means no data; $MOD_{MF}$ is monsoon forest product produced in this study; MCD12Q1 is MODIS land cover product

### 3.4 Comparison of monsoon forest areas at national level

Three satellite-based global land cover datasets, MODIS land cover product (MCD12Q1), Global Land Cover 2000 (GLC2000) map and University of Maryland (UMD) land cover dataset, were used to compare the derived monsoon forest areas. In tropical Asia, the DBF type in these global products could be considered equivalent to tropical monsoon forests. The estimated areas of monsoon forests in GLC2000 and UMD are also shown in Table 2. A 100% stacked comparison was performed to evaluate the consistency in each country among the four products (Fig. 6). The UMD derived low percent of monsoon forests except in Bhutan, Laos and Vietnam. Similarly, MCD12Q1 also accounts for quite low percent in most countries. It is seen that MCD12Q1 and UMD mapped far lower monsoon forest areas than GLC2000 and $MOD_{MF}$ did, which is also shown in Table 2. The main reason is that the definition of forests in International Geosphere-Biosphere Program (IGBP) classification system which is employed by MCD12Q1 and UMD products is different from that in Land Cover Classification System (LCCS) which is employed by GLC2000. The former restricts 60% forest cover as the low limit in the definition of forests, whereas the latter only requires 15%. Percent of GLC2000 varies greatly

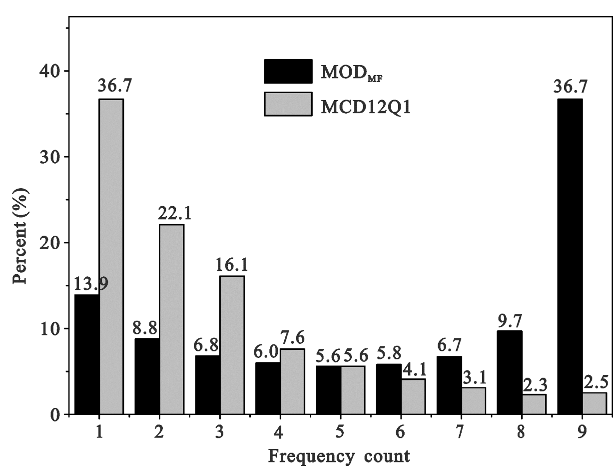


**Fig. 5** Percentage of every frequency count for comparing the stability of long-term derived monsoon forests between monsoon forest product produced in this study ($MOD_{MF}$) and MODIS land cover product (MCD12Q1)

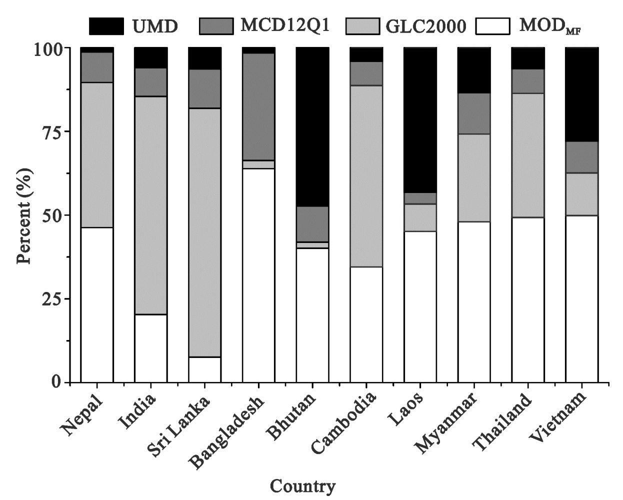


**Fig. 6** Monsoon forests areas percent stacked comparison among MODIS land cover product (MCD12Q1) in 2001, Global Land Cover 2000 product (GLC2000), University of Maryland land cover product (UMD) and monsoon forest product produced in this study ($MOD_{MF}$) in 2000 in 10 countries

among the countries, with some exceeding more than half and others accounting for only a little (Fig. 6). The percent areas of $MOD_{MF}$ are relatively even compared with other three datasets, except in Bangladesh, whose monsoon forests area appears to be overestimated. Among the four datasets, it appears poor agreement on the extracted monsoon forests areas in all countries.

To better understand how well the $MOD_{MF}$ algorithm performed compared with other three remote sensing products, a set of national statistical areas of monsoon forests (hereafter referred to as $NSA_{MF}$) was collected (Table 2). At the national level, the derived areas of monsoon forests in the four products were compared with $NSA_{MF}$. In many countries, especially the countries where monsoon forests widely distribute, such as in India and Myanmar, there are the minimum differences between the $MOD_{MF}$ and $NSA_{MF}$ datasets. The result of Bangladesh seems to have relatively large differences between the $MOD_{MF}$ and $NSA_{MF}$ datasets, which was also mentioned in the 100% stacked comparison. Simple linear regressions of monsoon forests areas estimates of 10 countries between satellite products and $NSA_{MF}$ were performed (Fig. 7). They show that areas of the $MOD_{MF}$ and $NSA_{MF}$ are strongly correlated, with the highest $R^2$ value of 0.95 compared with 0.88 of MCD12, 0.61 of GLC2000 and 0.67 of UMD. Moreover, the regression between $MOD_{MF}$ and $NSA_{MF}$ has the lowest RMSE of 14 014 km$^2$, which is far lower than 56 969 km$^2$ of MCD12, 108 800 km$^2$ of GLC2000 and 59 867 km$^2$ of UMD.

### 3.5 Evaluation of map accuracy

To objectively evaluate the results, accuracies of the $MOD_{MF}$ maps were assessed using a dataset of 100 ground-truth forest observations, which were collected in all 10 countries in this study (Fig. 8). These sites were sampled corresponding to the proportion of forest area in each country according to FAO reports (FAO, 2010). Most of the observation sites in field were selected in the vicinity of national parks, reserved forests, wildlife sanctuaries, or the most reprehensive forests in a country. One ground-truth site was identified as the monsoon forest (MF) or non-MF by digging the local information from three sources: 1) the research papers relating monsoon forests in the study area; 2) Google Earth, a

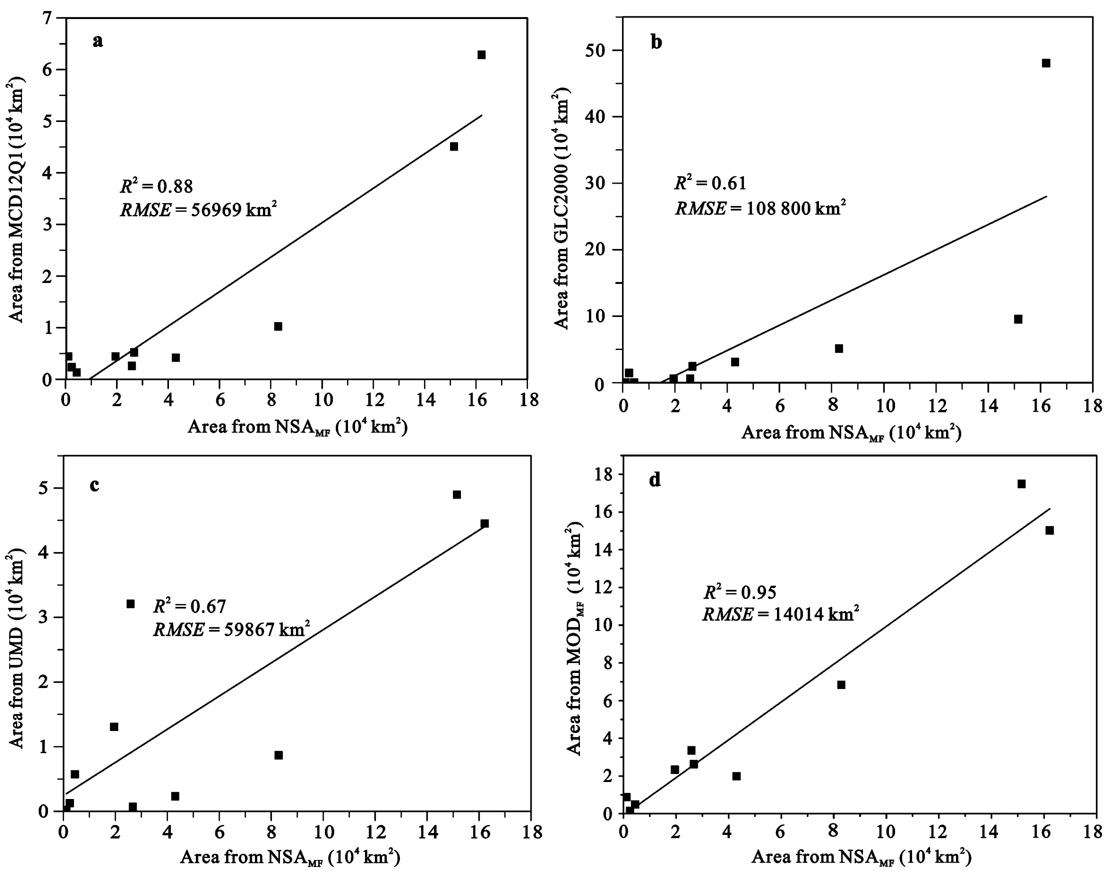


**Fig. 7** Linear Regressions of monsoon forests areas between satellite-based datasets and $NSA_{MF}$ (a collected national statistical monsoon forests areas). (a) MCD12Q1 in 2001; (b) GLC2000; (c) UMD; (d) $MOD_{MF}$ in 2000. Meanings of all abbreviations see Fig. 6

fashionable virtual globe that assembles high resolution images and seas of photographs taken by volunteers; 3) descriptions from local websites. The precise coordinate information of a ground-truth site was obtained from 1) or 2) and the decision of its land cover type was mostly made based on both descriptions from 1) and 3). Once labeled, the 100 sites were compared with the corresponding pixels in the $MOD_{MF}$ map at the same location.

The confusion matrix for this study is reported in Table 4. From the viewpoint of a map producer, the accuracy of monsoon forests is 80.0% and the accuracy for non-MF is 99.1%. At the standpoint of a map user, this new monsoon forests map achieved a high accuracy of 99.4%. The user's accuracy of non-MF is only 70.9% because of the misclassified monsoon forests. For a transitional ecosystem, the imagery pixels in transition zones, especially at the border of two land covers, are easy to be misclassified as other types. The overall accuracy of $MOD_{MF}$ maps is 86.3%.

In order to provide more information on classification errors, Cohen's Kappa coefficient, which is thought to be a more reliable measure than simple percent judgment by taking into account the chance agreement, is calculated (Table 4). The Kappa coefficient is 71.9% for this classification.

## 4 Discussion

In most satellite-based classifications, the monsoon forest is not classified as an individual land-cover type but mixed up with temperate deciduous forests in common DBF legend, however, two forest types have different mechanisms of leaf falling and thus forest behaviors differentiate greatly. It is necessary to separate monsoon forests from other forests.

The algorithm developed in this study to map monsoon forests based on NDVI amplitude has two advantages. First, the composite of a whole year observations is applied to get the annual maximum and minimum NDVI, which could better show the annual variation amplitude of NDVI than single imagery and solve the

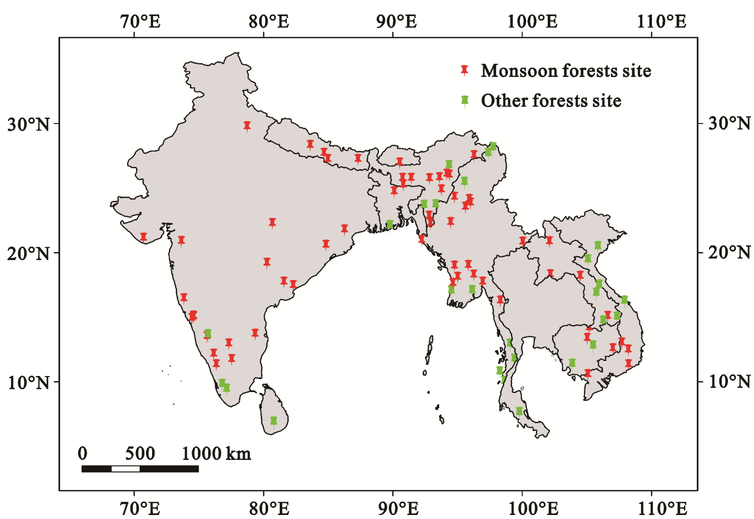


**Fig. 8** Locations of collected ground-truth sites

**Table 4** Confusion matrix for this study

| | | Classified class | | | Producer's accuracy (%) |
|---|---|---|---|---|---|
| | | MF | Non-MF | Total | |
| Observed classes | MF | 53.6 | 13.4 | 67.0 | 80.0 |
| | Non-MF | 0.3 | 32.7 | 33.0 | 99.1 |
| | Total | 53.9 | 46.1 | 100.0 | |
| User's accuracy (%) | | 99.4 | 70.9 | | |
| Overall accuracy = 86.3% | | | | | |
| Kappa coefficient = 71.9% | | | | | |

Note: MF stands for monsoon forests

difficulties of different deciduous periods in different regions in large-area classification. The composite is an assemble of absolute amplitude of NDVI for every pixel and it makes the classification avoid the step in single image classification to select different images to represent wet season and dry season for different regions. Second, as the analysis showed that the derived results were not sensitive to the classification threshold, this algorithm did not require any classifiers but used the threshold of NDVI amplitude, which simplifies the classification procedure.

The results have shown that the algorithm did well in balancing the simplicity and the accuracy. Although there exists variation in derived monsoon forests among different years, both the coefficient of variance and the frequency-based comparison demonstrate that the stability of the presented algorithm in long-time mapping is better than that of MCD12Q1 in the study region. The linear regression comparisons show that $MOD_{MF}$ has the highest $R^2$ value and the lowest RMSE, which suggests $MOD_{MF}$ has closer forest areas in most countries to the national statistical data than MCD12Q1, GLC2000 and UMD did. The $MOD_{MF}$ achieved a high producer′s accuracy of 80.0% for mapping monsoon forests and the Kappa coefficient of 71.9% also indicates a good agreement between $MOD_{MF}$ and the ground-truth sites (Landis and Koch, 1977; Fleiss *et al*., 2003). These results indicate that the algorithm has the potential to be employed in other tropical regions. However, the thresholds produced in this study for monsoon forests are just reference values, because the thresholds could be different in other regions. In small-area classification, the thresholds can be manually modified to maximally separate monsoon forests from other forests. Moreover, if more classification features were employed as the inputs, such as MaxNDVI and MinNDVI, the accuracy of mapping monsoon forests could be improved to some extent.

The results suggest that the algorithm is capable of identifying monsoon forests. However, there still remain some factors that affect the results. The first factor is the variation of inter-annual NDVI. The threshold of NDVI amplitude was direct determined by visual observation but NDVI-based metrics of forests could be different year by year. In this paper, the classification threshold in 2000 was used for the classification. This might lead to some errors and oscillations in the final products for 2001 to 2009. In Table 2, it shows that the monsoon forest area for every country varies year by year, and the variations in some counties even account for up to 30% of the average areas (Table 3). The second factor is that the other land covers might disturb the extraction results. Although the mean of NDVI were used to exclude non-forest information, the confusion between monsoon forests and mixed woodland and shrublands does exists (Herold *et al*., 2008). Especially at the border of two land covers, they might be misclassified as each other and this could reduce the accuracy of $MOD_{MF}$ maps. Thirdly, as the high producer′s accuracy of non-MF contributes to the overall accuracy, the overall accuracy is less valuable to evaluate the classification than the producer′s accuracy for monsoon forests in this research.

Two recommendations can be made for improving in further researches: 1) adding some other features as classification inputs which might be related with the monsoon forests, they could further enhance the accuracy of classification; 2) assessing the classification accuracy by employing more diversified reference data, not only relying on the sampling point validation.

## 5 Conclusions

In this paper, a simple algorithm was proposed to map monsoon forests using MODIS data. This algorithm is based on the high greenness variation of different seasons in monsoon forests. The NDVI amplitude was used as the classification input and the classification threshold for monsoon forests and other forests was decided by direct visual observation. In this way, monsoon forest maps ($MOD_{MF}$) for 2000–2009 were produced. The accuracies assessed by a ground-truth dataset and the Kappa coefficient both proved the results to be good. Comparisons between $MOD_{MF}$, MCD12Q1, UMD, GLC2000 and the national statistical areas reveals that $MOD_{MF}$ are highly correlated to the statistical data with the highest $R^2$ value and the lowest RMSE. Moreover, the stability comparison demonstrates that the long-term maps extracted by this algorithm are more stable in delineating the spatial pattern of monsoon forests than MCD12Q1.

The proposed algorithm selected the classification feature of NDVI amplitude based on the understanding of the unique characteristic of monsoon forests, so the classification rules was not necessary to be created using any classifier. This algorithm could simplify the process of classification and is suitable to those classifications without complicated techniques. At the same time, the results have shown that the presented method could balance well between accuracy and efficiency and it has the potential to map monsoon forests in other tropical regions.

## References

Bartholome E, Belward A S, 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26(9): 1959–1977. doi: 10.1080/01431160412331291297

Bicheron P, Defourny P, Brockmann C *et al*., 2008. *Globcover*: *Products Description and Validation Report*. Toulouse: MEDIAS France.

Bohlman S A, 2010. Landscape patterns and environmental controls of deciduousness in forests of central Panama. *Global Ecology and Biogeography*, 19(3): 376–385. doi: 10.1111/j.1466-8238.2009.00518.x

Bontemps S, Defourny P, Bogaert E V *et al*., 2011. *GLOBCOVER 2009*: *Products Description and Validation Report*. European Space Agency and Université catholique de Louvain.

Brown J C, Kastens J H, Coutinho A C *et al*., 2013. Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. *Remote Sensing of Environment*, 130(4): 39–50. doi: 10.1016/j.rse.2012.11.009

Collins N M, Sayer J A, Whitmore T C, 1991. *The Conservation Atlas of Tropical Forests*: *Asia and the Pacific*. London: Macmillan Press Ltd.

Delbart N, Kergoat L, Le Toan T *et al*., 2005. Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment*, 97(1): 26–38. doi: 10.1016/j.rse.2005.03.011

Eamus D, 1999. Ecophysiological traits of deciduous and evergreen woody species in the seasonally dry tropics. *Trends in Ecology & Evolution*, 14(1): 11–16. doi: 10.1016/S0169-5347(98)01532-8

FAO (Food and Agriculture Organization of the United Nations), 1998. *Asia-Pacific Forestry Sector Outlook Study Working Paper Series*: *Country Report*. Roma.

FAO (Food and Agriculture Organization of the United Nations), 2001. *Global Forest Resources Assessment 2000*: *Main Report*. Roma.

FAO (Food and Agriculture Organization of the United Nations), 2010. *Global Forest Resources Assessment 2010*: *Main Report*. Roma.

Fleiss J L, Levin B, Paik M C, 2003. *Statistical Methods for Rates and Proportions*. New Jersey: John Wiley & Sons Inc.

Foley J A, Prentice I C, Ramankutty N *et al*., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles*, 10(4): 603–628. doi: 10.1029/96gb02692

Friedl M A, McIver D K, Hodges J C F *et al*., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1–2): 287–302. doi: 10.1016/S0034-4257(02)00078-0

Friedl M A, Sulla-Menashe D, Tan B *et al*., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1): 168–182. doi: 10.1016/j.rse.2009.08.016

Gong P, Wang J, Yu L *et al*., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34(7): 2607–2654. doi: 10.1080/01431161.2012.748992

Hansen M C, De Fries R S, Townshend J R G *et al*., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6): 1331–1364. doi: 10.1080/014311600210209

Herold M, Mayaux P, Woodcock C E *et al*., 2008. Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, 112(5): 2538–2556. doi: 10.1016/j.rse.2007.11.013

Homer C, Huang C Q, Yang L M *et al*., 2004. Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70(7): 829–840. doi: 10.14358/PERS.70.7.829

Htun N Z, Mizoue N, Yoshida S, 2011. Classifying tropical deciduous vegetation: a comparison of multiple approaches in Popa Mountain Park, Myanmar. *International Journal of Remote Sensing*, 32(24): 8935–8948. doi: 10.1080/01431161.2010.531779

Ito E, Araki M, Tith B *et al*., 2008. Leaf-shedding phenology in lowland tropical seasonal forests of Cambodia as estimated from NOAA satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10): 2867–2871. doi: 10.1109/Tgrs.2008.919820

Jia K, Liang S, Wei X *et al*., 2014. Land cover classification of Landsat data with phenological features extracted from time series MODIS NDVI data. *Remote Sensing*, 6(11): 11518–11532. doi: 10.3390/rs61111518

Kalacska M E R, Sanchez-Azofeifa G A, Calvo-Alvarado J C *et al*., 2005. Effects of season and successional stage on leaf area index and spectral vegetation indices in three mesoamerican tropical dry forests. *Biotropica*, 37(4): 486–496. doi: 10.1111/j.1744-7429.2005.00067.x

Krishnaswamy J, Kiran M C, Ganeshaiah K N, 2004. Tree model based eco-climatic vegetation classification and fuzzy mapping in diverse tropical deciduous ecosystems using multi-season NDVI. *International Journal of Remote Sensing*, 25(6): 1185–1205. doi: 10.1080/0143116031000149989

Landis J R, Koch G G, 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1): 159–174. doi: 10.2307/2529310

Lenney M P, Woodcock C E, Collins J B *et al*., 1996. The status of agricultural lands in Egypt: the use of multitemporal NDVI features derived from Landsat TM. *Remote Sensing of Environment*, 56(95): 8–20. doi: 10.1016/0034-4257(95)00152-2

Lillesø J P B, Shrestha T B, Dhakal L P *et al*., 2005. *The Map of Potential Vegetation of Nepal—A Forestry/Agro-Ecological/Biodiversity Classification System*. Copenhagen: Danish Centre for Forest Landscape & Planning Kvl.

Liu R G, Liu Y, 2013. Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sensing of Environment*, 133(12): 21–37. doi: 10.1016/j.rse.2013.01.019

Loveland T R, Reed B C, Brown J F *et al*., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21(6–7): 1303–1330. doi: 10.1080/014311600210191

Malhi Y, Pegoraro E, Nobre AD *et al*., 2002. Energy and water dynamics of a central Amazonian rain forest. *Journal of Geophysical Research*: *Atmospheres*, 107(D20): LBA 45-1–LBA 45-17. doi: 10.1029/2001jd000623

Miles L, Newton A C, DeFries R S *et al*., 2006. A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, 33(3): 491–505. doi: 10.1111/j.1365-2699.2005.01424.x

Murphy P G, Lugo A E, 1986. Ecology of tropical dry forest. *Annual Review of Ecology and Systematics*, 17(1): 67–88. doi: 10.1146/annurev.es.17.110186.000435

Peel M C, McMahon T A, Finlayson B L *et al*., 2001. Identification and explanation of continental differences in the variability of annual runoff. *Journal of Hydrology*, 250(1): 224–240. doi: 10.1016/S0022-1694(01)00438-3

Ruangpanit N, 1995. Tropical seasonal forests in monsoon Asia: with emphasis on continental southeast Asia. *Vegetatio*, 121(1–2): 31–40. doi: 10.1007/Bf00044670

Steininger M K, Tucker C J, Ersts P *et al*., 2001. Clearance and fragmentation of tropical deciduous forest in the Tierras Bajas, Santa Cruz, Bolivia. *Conservation Biology*, 15(4): 856–866. doi: 10.1046/j.1523-1739.2001.015004856.x

Tanaka N, Kume T, Yoshifuji N *et al*., 2008. A review of evapotranspiration estimates from tropical forests in Thailand and adjacent regions. *Agricultural and Forest Meteorology*, 148(5): 807–819. doi: 10.1016/j.agrformet.2008.01.011

Tang X, Li H, Desai A R *et al*., 2014. How is water-use efficiency of terrestrial ecosystems distributed and changing on Earth? *Scientific Reports*, 4: 7483–7483. doi: 10.1038/srep07483

Townshend J R G, Justice C O, Kalb V, 1987. Characterization and classification of south American land cover types using satellite data. *International Journal of Remote Sensing*, 8(8): 1189–1207. doi: 10.1080/01431168708954764

Trejo I, Dirzo R, 2000. Deforestation of seasonally dry tropical forest: a national and local analysis in Mexico. *Biological Conservation*, 94(2): 133–142. doi: 10.1016/S0006-3207(99)00188-3

Tucker C J, Townshend J R G, Goff T E, 1985. African land-cover classification using satellite data. *Science*, 227(4685): 369–375. doi: 10.1126/science.227.4685.369

Valentini R, Matteucci G, Dolman A J *et al*., 2000. Respiration as the main determinant of carbon balance in European forests. *Nature*, 404(6780): 861–865. doi: 10.1038/35009084

Vermote E F, El Saleous N Z, Justice C O, 2002. Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment*, 83(1–2): 97–111. doi: 10.1016/S0034-4257(02)00089-5

Vourlitis G L, Priante N, Hayashi M M S *et al*., 2002. Seasonal variations in the evapotranspiration of a transitional tropical forest of Mato Grosso, Brazil. *Water Resources Research*, 38(6): 30-1–30-11. doi: 10.1029/2000wr000122

Vourlitis G L, Priante N, Hayashi M M S *et al*., 2004. Effects of meteorological variations on the $CO_2$ exchange of a Brazilian transitional tropical forest. *Ecological Applications*, 14(4): S89–S100. doi: 10.1890/01-6005

Wright S J, Cornejo F H, 1990. Seasonal drought and leaf fall in a tropical forest. *Ecology*, 71(3): 1165–1175. doi: 10.2307/1937384

Xiao X M, Boles S, Liu J Y *et al*., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, 95(4): 480–492. doi: 10.1016/j.rse.2004.12.009

Yoshifuji N, Kumagai T O, Tanaka K *et al*., 2006. Inter-annual variation in growing season length of a tropical seasonal forest in northern Thailand. *Forest Ecology and Management*, 229(1–3): 333–339. doi: 10.1016/j.foreco.2006.04.013

Zhu Z, Woodcock C E, 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144(1): 152–171. doi: 10.1016/j.rse.2014.01.011