Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

Remote Sensing of Environment 100 (2006) 95 – 113

Remote Sensing of Environment

www.elsevier.com/locate/rse

# Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images

Xiangming Xiao [a,*], Stephen Boles [a], Steve Frolking [a], Changsheng Li [a], Jagadeesh Y. Babu [a,b], William Salas [c], Berrien Moore III [a]

[a] *Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Durham, NH 03824, USA*
[b] *Central Rice Research Institute, Cuttack, 753006, Orissa, India*
[c] *Applied Geosolutions, LLC, Durham, NH 03824, USA*

Received 27 April 2005; received in revised form 12 August 2005; accepted 1 October 2005

**Abstract**

In this paper, we developed a new geospatial database of paddy rice agriculture for 13 countries in South and Southeast Asia. These countries have ~30% of the world population and ~2/3 of the total rice land area in the world. We used 8-day composite images (500-m spatial resolution) in 2002 from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the NASA EOS Terra satellite. Paddy rice fields are characterized by an initial period of flooding and transplanting, during which period a mixture of surface water and rice seedlings exists. We applied a paddy rice mapping algorithm that uses a time series of MODIS-derived vegetation indices to identify the initial period of flooding and transplanting in paddy rice fields, based on the increased surface moisture. The resultant MODIS-derived paddy rice map was compared to national agricultural statistical data at national and subnational levels. Area estimates of paddy rice were highly correlated at the national level and positively correlated at the subnational levels, although the agreement at the national level was much stronger. Discrepancies in rice area between the MODIS-derived and statistical datasets in some countries can be largely attributed to: (1) the statistical dataset is a sown area estimate (includes multiple cropping practices); (2) failure of the 500-m resolution MODIS-based algorithm in identifying small patches of paddy rice fields, primarily in areas where topography restricts field sizes; and (3) contamination by cloud. While further testing is needed, these results demonstrate the potential of the MODIS-based algorithm to generate updated datasets of paddy rice agriculture on a timely basis. The resultant geospatial database on the area and spatial distribution of paddy rice is useful for irrigation, food security, and trace gas emission estimates in those countries.


*Keywords:* Enhanced vegetation index; Land surface water index

## 1. Introduction

Paddy rice fields account for over 11% of global cropland area (Maclean et al., 2002). The major rice-producing countries of Asia account for over half of the world's population and rice represents over 35% of their daily caloric intake (FAO, 2004). Monitoring and mapping of paddy rice agriculture in a timely and efficient manner is very important for agricultural and environmental sustainability, food and water security, and greenhouse gas emissions. Because paddy rice is grown on flooded soils (irrigated and rainfed), water resource management is a major concern. Irrigation for agriculture accounts for over 80% of the fresh water withdrawals in our Asian study area, with several of the countries reporting over 95% of fresh water used for irrigation (FAOSTAT, 2001). These high levels of irrigation also raise concerns about maintenance and contamination of the water supply. Also, seasonally flooded rice paddies are a significant source of the greenhouse gas methane (Denier Van Der Gon, 2000; Li et al., 2002; Neue & Boonjawat, 1998), contributing over 10% of the total methane flux to the atmosphere (Prather & Ehhalt, 2001), which may have substantial impacts on atmospheric chemistry and climate. Field studies have shown that water management can have a significant influence on total methane emissions during a cropping season (Wassmann et al., 2000; Sass et al., 1999), so paddy water management has become a target scenario for greenhouse gas mitigations (Wassmann et al., 2000; Li et al., 2005).

* Corresponding author. Tel.: +603 862 3818; fax: +603 862 0188.
*E-mail address:* xiangming.xiao@unh.edu (X. Xiao).

0034-4257/$ - see front matter 
doi:10.1016/j.rse.2005.10.004

Several global datasets of paddy rice were developed in the late 1980s and early 1990s (Aselman & Crutzen, 1989; Matthews et al., 1991; Olson, 1992; Wilson & Henderson-Sellers, 1992) and were used in analyses of climate and trace gas emissions; most of these datasets have a coarse spatial resolution (0.5° to 5°). More recently (Leff et al., 2004), provided a global rice map at a spatial resolution of five arc minutes as part of their global cropland product. An Asian rice dataset was generated using agricultural statistical data from the 1970s (Huke, 1982) and updated with agricultural census data in the early 1990s (Huke & Huke, 1997). This updated Asia rice database was used to estimate methane emissions in Asia (Knox et al., 2000; Matthews et al., 2000). It is important to note that all of the above-mentioned rice datasets were wholly or partially based on agricultural statistical data. These statistical data sources cannot meet the needs of science and policy researchers that require updated geospatial databases of paddy rice agriculture at improved spatial and temporal resolutions.

In Asia, rice is grown over a large spatial domain (60° of latitude and 80° of longitude) and a wide range of landscape types. Such a large area contains a wide variety of climatic conditions, ranging from the temperate zones in the north to the tropical equatorial zones in the south. With such a large variation in landscapes and climates in the rice-growing region of Asia, a large number of unique paddy farming methods have also evolved, based on farming type (irrigated, rainfed, deepwater), crop management (single crop, multi-crop), and seasonality (wet season, dry season). This variation in potential rice farming scenarios makes the generation of a timely and spatially explicit paddy rice dataset a challenging task. Optical satellite remote sensing provides a viable means to improve geospatial datasets of paddy rice fields, and a number of earlier studies have used Landsat or AVHRR images to generate local to regional-scale estimates of paddy rice fields (Fang, 1998; Fang et al., 1998; Okamoto & Fukuhara, 1996; Okamoto & Kawashima, 1999; Tennakoon et al., 1992; Van Niel et al., 2003). Most of those previous satellite-based rice analyses used image classification procedures and required abundant local knowledge (e.g., crop calendars) of rice paddy fields. A rice mapping method that is both timely and requires less prior knowledge of local farming management would be a tremendous asset for large-scale mapping of paddy rice fields.

Recently, we developed an approach that takes advantage of a new generation of optical sensors such as VEGETATION (VGT; Xiao et al., 2002b) and the Moderate Resolution Imaging Spectroradiometer (MODIS; Xiao et al., 2005). This approach is based on a unique physical feature of paddy rice fields—rice is grown on flooded soils and paddy fields are a mixture of open water and green rice plants during the early part of the growing season. An algorithm was developed to identify and track those image pixels that experienced flooding and rice transplanting over time. Unlike other satellite-based classification algorithms that primarily use the Normalized Difference Vegetation Index (NDVI; Eq. (1)), our temporal profile analysis algorithm combines vegetation indices that are sensitive to the development of canopy (e.g., leaf area index, chlorophyll) and vegetation indices that are sensitive to changes in the land surface water content. We have applied this algorithm to map paddy rice fields in central China at a local scale using VGT data (Xiao et al., 2002b) and at a regional scale (13 provinces in China) using 8-day MODIS composite data (Xiao et al., 2005).

In this study, we use this algorithm to map paddy rice fields in 13 countries of South and Southeast Asia. Our objective is to generate an updated geospatial database of paddy rice at 500-m spatial resolution, using 8-day MODIS composites in 2002. The resultant geospatial database could be used to support various studies of land-use and land-cover change, methane emission estimations, and food and water security in Asia.

## 2. Brief description of the study area

The study area encompasses 13 countries in South and Southeast Asia, ranging from 68°E to 142°E and 10°S to 35°N (Fig. 1). The region contains a variety of climate zones, including tropical and subtropical areas in the southeast, temperate areas in northern India and Nepal, and dry areas in western India. In the areas where rice growth is limited by precipitation or temperature, there is usually one rice crop per year, as in most of the dry and temperate zones. However, in many of the tropical regions, two rice crops per year are common and, in some areas (such as the Mekong Delta in Vietnam), three crops per year are grown. Seasonal patterns of precipitation are driven by the monsoon climate system that dominates over the Indian subcontinent and Southeast Asia. The monsoons are seasonal winds that bring torrential rains in the summer (May/June to September/October) and sunny and dry weather in the winter.

The 13 countries in our study area are home to almost 1.8 billion people (Table 1), almost 30% of the global population. Rice is a highly important product in this part of the world, where much of the population is still employed by the agriculture sector. Rice represents a significant portion of total cropland area and the amount of daily caloric intake (Table 1). With approximately 1,000,000 km$^2$ of area sown to paddy rice (Table 1), our study area represents almost two-thirds of the world's total area sown to rice (1.53 million km$^2$ in 2004).

## 3. Data and methods

### 3.1. MODIS image data

The MODIS sensor has 36 spectral bands, 7 of which are designed for the study of vegetation and land surfaces: blue (459–479 nm), green (545–565 nm), red (620–670 nm), near infrared (NIR$_1$: 841–875 nm, NIR$_2$: 1230–1250 nm), and shortwave infrared (SWIR$_1$: 1628–1652 nm, SWIR$_2$: 2105–2155 nm). Daily global imagery is provided at spatial resolutions of 250-m (red and NIR$_1$) and 500-m (blue, green, NIR$_2$, SWIR$_1$, SWIR$_2$). The MODIS Land Science Team provides a suite of standard MODIS data products to users, including the 8-day composite MODIS Surface Reflectance

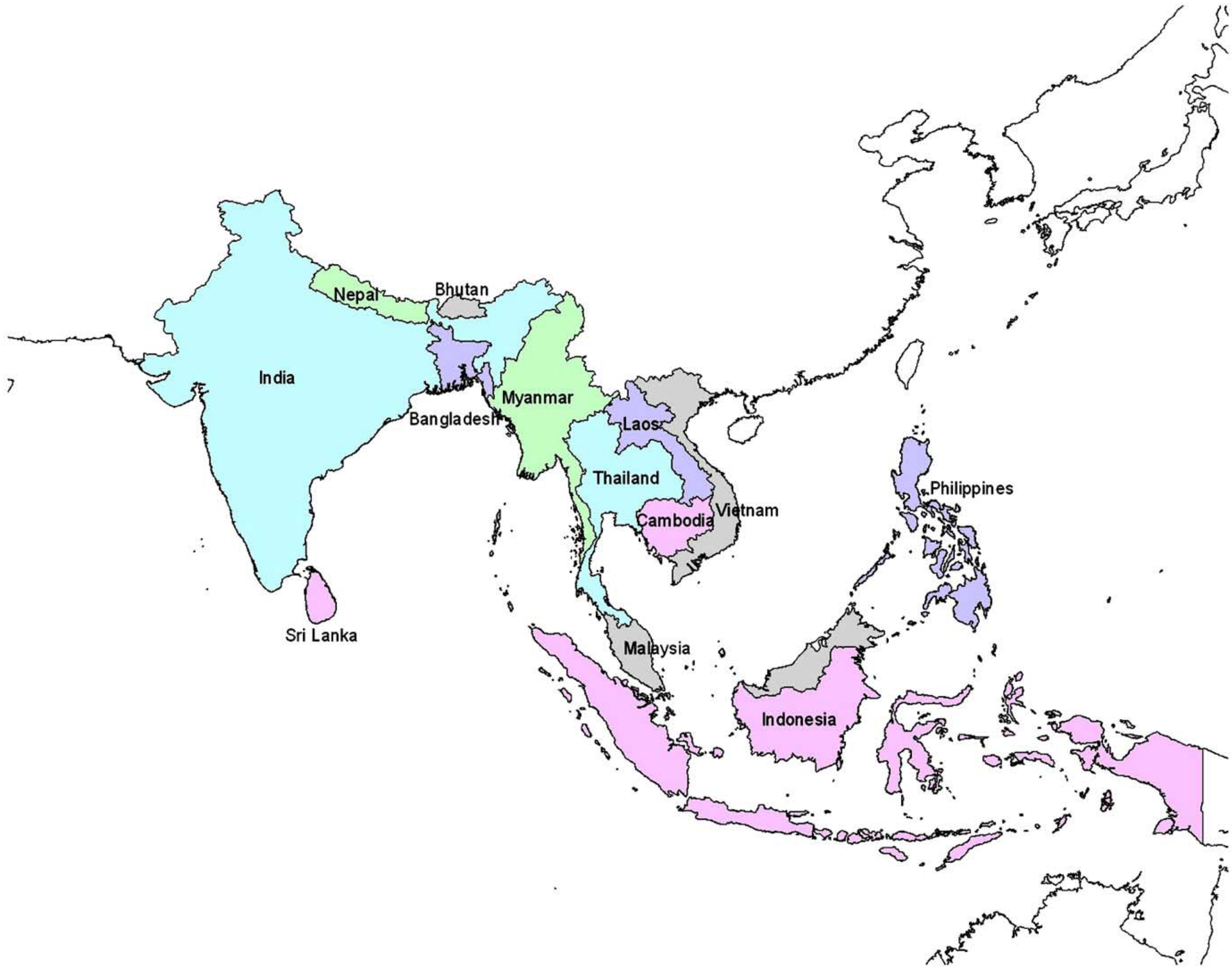


Fig. 1. Spatial extent and location of the 13 countries in South and Southeast Asia.

Product (MOD09A1). Each 8-day composite includes estimates of surface reflectance of the seven spectral bands at 500-m spatial resolution. In the production of MOD09A1, atmospheric corrections for gases, thin cirrus clouds, and aerosols are implemented (Vermote & Vermeulen, 1999). MOD09A1 composites are generated in a multi-step process that first eliminates pixels with a low score or low observational coverage, and then selects an observation with the minimum blue-band value during the 8-day period (http://modis-land.gsfc.nasa.gov/MOD09/MOD09ProductInfo/

Table 1
A comparison of agricultural (FAO, 2004), nutritional (FAO, 2002), and population (FAO, 2003) statistics for 13 countries in Southeast and South Asia

| Country | Cropland sown area (km$^2$) | Paddy rice sown area (km$^2$) | Percent of cropland that is paddy rice | Rice production (×000 Mt) | % of caloric intake from rice | Population (×000) |
|---|---|---|---|---|---|---|
| Bangladesh | 145,661 | 110,000 | 76 | 37,910 | 74 | 146,736 |
| Bhutan | 1014 | 200 | 20 | 45 | 21 | 2257 |
| Cambodia | 27,446 | 23,000 | 84 | 4710 | 69 | 14,144 |
| India | 1,906,335 | 425,000 | 22 | 124,410 | 33 | 1,065,462 |
| Indonesia | 321,707 | 117,527 | 37 | 53,100 | 50 | 219,883 |
| Laos | 11,382 | 8200 | 72 | 2700 | 64 | 5657 |
| Malaysia | 62,832 | 6700 | 11 | 2184 | 25 | 24,425 |
| Myanmar | 141,912 | 60,000 | 42 | 23,000 | 68 | 49,485 |
| Nepal | 44,023 | 15,500 | 35 | 4300 | 38 | 25,164 |
| Philippines | 126,8161 | 40,000 | 32 | 14,200 | 43 | 79,999 |
| Sri Lanka | 19,566 | 7555 | 39 | 2510 | 37 | 19,065 |
| Thailand | 177,610 | 98,000 | 55 | 25,200 | 42 | 62,833 |
| Vietnam | 130,302 | 74,000 | 57 | 35,500 | 65 | 81,377 |
| Total | 3,116,606 | 985,682 | 32 | 329,769 | 42.4 | 1,796,487 |

MOD09_L3_8-day.htm). The composites still have reflectance variations associated with the bidirectional reflectance distribution function. MOD09A1 also includes quality control flags to account for various image artifacts (e.g., clouds, cloud shadow). Standard MODIS products are organized in a tile system with the Sinusoidal projection; each tile covers an area of 1200 × 1200 km (approximately 10° latitude × 10° longitude at equator). In this study, we acquired 23 tiles of MOD09A1 data for 2002 (forty-six 8-day composites per year) from the USGS EROS Data Center (http://edc.usgs.gov/).

### 3.2. Algorithms for identifying inundation and paddy rice field

A unique physical feature of paddy rice fields is that rice plants are grown on flooded soils. Temporal dynamics of paddy rice fields can be characterized by three main periods: (1) the flooding and rice transplanting period; (2) the growing period (vegetative growth, reproductive, and ripening stages); and (3) the fallow period after harvest (Le Toan et al., 1997). During the flooding and rice transplanting period, the land surface is a mixture of surface water and green rice plants, with water depths usually between 2 and 15 cm. About 50 to 60 days after transplanting rice plant, canopies cover most of the surface area. At the end of the growth period prior to harvesting, there is a decrease of leaf and stem moisture content and a decrease of the number of leaves. Individual farmers have different flooding and rice transplanting schedules for their paddy rice fields, which poses a great challenge for remote sensing analyses at large spatial scales.

To identify the changes in the mixture of surface water and green vegetation in paddy rice fields over time requires spectral bands or vegetation indices that are sensitive to both water and vegetation. For each MOD09A1 composite, we calculate Normalized Difference Vegetation Index (NDVI; Eq. (1)), Land Surface Water Index (LSWI; Eq. (2)), and Enhanced Vegetation Index (EVI; Eq. (3)), using surface reflectance values from the blue, red, NIR (841–875 nm), and SWIR (1628–1652 nm) bands:

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{nir}} + \rho_{\mathrm{red}}} \quad (1)$$

$$\mathrm{LSWI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{swir}}}{\rho_{\mathrm{nir}} + \rho_{\mathrm{swir}}} \quad (2)$$

$$\mathrm{EVI} = 2.5 \times \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{nir}} + 6 \times \rho_{\mathrm{red}} - 7.5 \times \rho_{\mathrm{blue}} + 1} \quad (3)$$

NDVI is closely correlated to the leaf area index (LAI) of paddy rice fields (Xiao et al., 2002c). The blue band is sensitive to atmospheric conditions and is used for atmospheric correction. EVI directly adjusts the reflectance in the red band as a function of the reflectance in the blue band, and it accounts for residual atmospheric contamination and variable soil and canopy background reflectance (Huete et al., 2002, 1997). The SWIR spectral band is sensitive to leaf water and soil moisture, and is used to develop improved vegetation indices that are sensitive to equivalent water thickness (EWT, g $H_2O/m^2$), including LSWI (Maki et al., 2004; Xiao et al., 2002a,b).

We have developed an algorithm to identify paddy rice fields through a temporal profile analysis of LSWI, NDVI, and EVI (Xiao et al., 2002b, 2005). The algorithm focuses on the period from flooding/transplanting through rapid plant growth in the early part of the growing season to the point where a full canopy exists. Our hypothesis is that a temporary inversion of the vegetation indices, where LSWI either approaches or is higher than NDVI or EVI values, may signal flooding in paddy rice fields. To slightly relax the simple threshold assumption (LSWI > NDVI) used in the earlier study with 1-km VGT images (Xiao et al., 2002b), for 500-m MODIS images, we used the following thresholds for identifying a flooded pixel: LSWI + 0.05 ≥ EVI or LSWI + 0.05 ≥ NDVI (Xiao et al., 2005). After a pixel was identified as a "flooding and transplanting" pixel, a procedure was implemented to determine whether rice growth occurs in that pixel, using the assumption that the EVI value of a true rice pixel reaches half of the maximum EVI value (in that crop cycle) within five 8-day composites (40 days) following the date of flooding and transplanting. Rice crops grow rapidly after transplanting and LAI usually reaches its peak within 2 months (Xiao et al., 2002c).

This algorithm has proven successful in detecting paddy rice fields in a variety of climate regimes and types of farm water management at various spatial scales within China (Xiao et al., 2002b, 2005). In this study, we will apply the algorithm to an even larger spatial domain, where climate and agricultural practices differ from China. As a test of our algorithm's ability to detect rice in different environments outside of China, we took advantage of field validation data provided by colleagues at the International Water Management Institute (Thenkabail et al., 2005; http://www.iwmidsp.org/iwmi/info/main.asp). By using geographic points that they validated as 90% or more rice area within 90-$m^2$ sampling units, we were able to confidently extract time series from rice ecosystems to test our algorithm. Three common rice management regimes were sampled in different parts of India to test our algorithm, including a single-rice crop in Bihar state (Fig. 2a), a double-rice crop in Karnataka state (Fig. 2b), and a double crop (single-rice + other crop) regime in Andhra Pradesh state (Fig. 2c). In all three instances, our algorithm identified the periods of flooding and transplanting at the onset of the rice-growing season.

### 3.3. Regional implementation of the paddy rice mapping algorithm

The implementation of our MODIS paddy rice detection algorithm at the continental scale is a challenging task and requires careful consideration of many factors that could potentially affect the seasonal dynamics of vegetation indices, including snow cover, clouds, water bodies, and other vegetated land-cover types. We have developed a procedure for regional implementation of the algorithm by generating various masks for clouds, snow cover, water bodies, and evergreen vegetation in an effort to minimize their potential impacts (Fig. 3).

The cloud cover mask is generated through two steps. The MOD09A1 file includes quality control flags for clouds. We

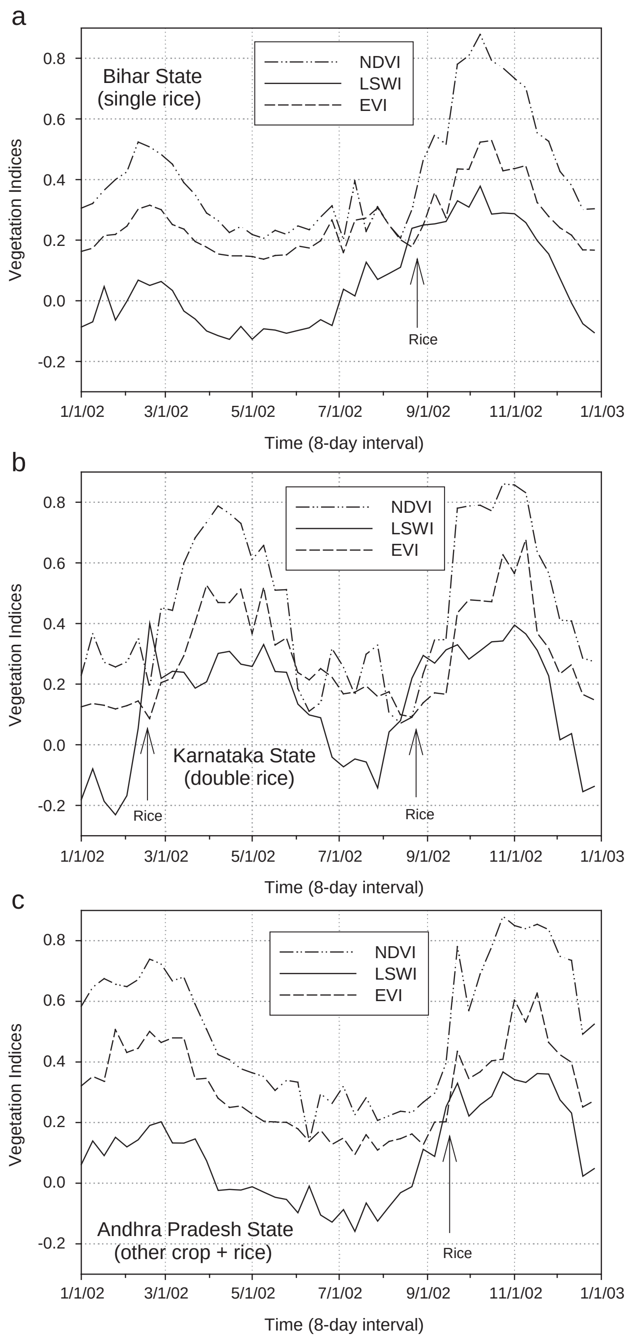


Fig. 2. The seasonal dynamics of the Normalized Difference Vegetation Index (NDVI), the Enhanced Vegetation Index (EVI), and the Land Surface Water Index (LSWI) at selected sites in: (a) a single-rice crop in Bihar, India (24.693°N, 84.499°E), (b) a double-rice crop in Karnataka, India (14.383°N, 75.755°E), (c) a single-rice+other crop in Andhra Pradesh, India (16.249°N, 80.49°E). Arrows define the approximate start of rice growth for each pixel. Note that the Andhra Pradesh site is a double cropping system; however, the non-rice crop (January–April) does not exhibit the flooding signal at the onset of the growing season.

extracted the information on clouds and generated masks of cloud cover for all time periods of each MODIS tile. It was noticed that a number of pixels had a high blue band reflectance but were not labeled as clouds in the MOD09A1 cloud quality flag. These pixels tended to have high LSWI relative to NDVI and EVI, potentially resulting in false identification of paddy rice areas. An additional restriction was then applied, whereas pixels with a blue reflectance of $\geq 0.2$ were also masked as cloudy pixels. For each MODIS tile, 46 cloud cover maps were generated; all cloud observations were excluded from further analyses. To determine the potential influence of clouds on the performance of our MODIS algorithm, the percent of all land pixels that were contaminated by clouds was calculated. There is an increase of cloud cover during the peak of the monsoon season (June–August), when contamination levels were approximately 40%. During the remainder of the year, cloud contamination levels fluctuated between 10% and 30%.

Snow cover has large surface reflectance values in the visible spectral bands and could potentially affect vegetation index values, particularly LSWI and EVI. To minimize the potential impact of those observations with snow cover in the winter and spring, we used the snow cover algorithms developed for the MODIS snow product (Hall et al., 1995, 2002) to generate snow cover masks. Normalized Difference Snow Index (NDSI; Eq. (4)) was first calculated for each 8-day composite, using surface reflectance values from the green and NIR bands, and then thresholds (NDSI > 0.40 and NIR > 0.11) were applied to identify snow-covered pixels. Those 8-day observations identified as snow during the year were excluded from identification of flooding and rice transplanting.

$$\mathrm{NDSI} = \frac{\rho_{\mathrm{green}} - \rho_{\mathrm{nir}}}{\rho_{\mathrm{green}} + \rho_{\mathrm{nir}}} \tag{4}$$

There is also a need to separate persistent water bodies from seasonally flooded pixels (e.g., paddy rice). We first analyzed

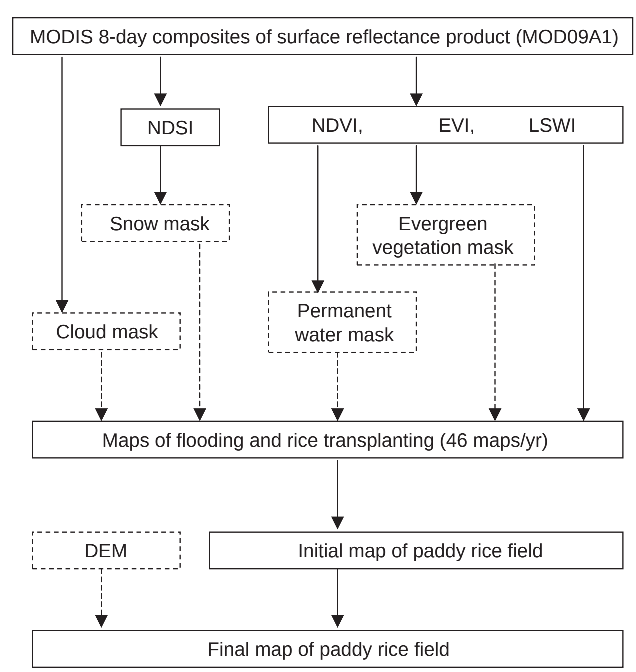


Fig. 3. A schematic diagram illustrating the algorithm for large-scale mapping of flooding and paddy rice from MODIS 8-day surface reflectance images at 500-m spatial resolution. One year of 8-day MODIS surface reflectance data (a total of 46 composites) are used as input.

temporal profiles of NDVI and LSWI, and assumed a pixel is covered by water if NDVI<0.10 and NDVI<LSWI, and generated a file that counts the number of 8-day periods within a year that a pixel is classified as water. Second, we assumed a pixel to be a persistent water body if it was identified as water in 10 or more 8-day composite periods in the year. Since the flooding/transplanting period is temporary, flooded rice pixels are expected to have fewer than 10 composite periods classified as water. Those pixels identified as persistent water bodies were not included in the identification of flooding and rice transplanting.

A mask of evergreen vegetation was generated using a two-test procedure that employs both NDVI and LSWI time series data. This step was necessary to remove the influence of vast regions of moist tropical and mangrove forests in the equatorial zones. If certain climatic conditions (e.g., flood) occur, these forests can have the tendency to exhibit some similar temporal characteristics as paddy rice fields, such as the temporary increase in LSWI. For the first test, evergreen forest areas tend to have consistently high NDVI values throughout the year, while rice pixels tend to have high NDVI values only in a few 8-day periods, mostly prior to harvesting. Evergreen forest areas were identified as those pixels having NDVI values of $\geq$0.7 over at least twenty 8-day composites during the year. Since the NDVI forest restriction is a cumulative count, we used a gap-filled product that corrects NDVI values in the time series where clouds were present. The second test, using LSWI, was designed to identify evergreen shrublands and woodlands. Croplands usually have some periods of time with exposed soils (after-harvest or land preparation) when LSWI values are very low. After examining the seasonal dynamics of LSWI for various vegetation types in the study area, we found that natural evergreen vegetation rarely has a LSWI value of <0.10. We assign those pixels with no LSWI values of <0.10 during the year to be natural evergreen vegetation. For each MODIS tile, one mask of natural evergreen vegetation was generated, and those pixels were excluded in the identification of flooding and paddy rice fields.

The most intensive areas of rice agriculture in the study area occur in the valleys and deltas of some of the great rivers, such as the Mekong (Vietnam, Cambodia, Laos, Thailand), Ganges (India, Bangladesh), and Ayeyarwady (Myanmar). However, there is some significant topography over much of the study area that could pose challenges to the implementation of a rice detection algorithm. Using the GTOPO30 digital elevation model (Global 30-Arc Second Elevation Dataset; http://edc.usgs.gov/products/elevation/gtopo30.html), it was determined that less than 0.5% of the rice area (from national rice statistics datasets) was situated at elevations greater than 2000 m in our study area. Therefore, we generated an elevation mask and used it to exclude those areas above 2000 m or with a slope greater than 2°.

We computed vegetation indices, masks, flooding maps, and rice maps for all the individual MODIS tiles that cover the study area. We then used an administrative boundary map of the study area to generate summaries of paddy rice area at national, provincial and subnational levels. Our subnational administrative base map was derived from two sources: (a) International Rice Research Institute (IRRI; www.irri.org) maps were used for administrative unit boundaries within the countries and (b) ArcWorld (ESRI, 1992) maps were used for international boundaries due to minor misalignments between national boundaries of individual IRRI maps. For each individual country, administrative boundary maps (IRRI) were merged with the respective ArcWorld national boundary map. Sliver polygons (a result of imperfect boundary matches between the datasets) were removed. Individual country maps were joined to form a complete subnational map of the study area comprising of 1586 polygons.

### 3.4. Ancillary data for evaluation of MODIS-based paddy rice map

Accuracy assessment of moderate-resolution (500-m to 1-km) land-cover products is a challenging task, as these maps can overestimate or underestimate areas of land-cover types due to the fragmentation and subpixel proportion of individual land-cover types. Because of budget constraints and human resource limitations, we were not able to conduct extensive field surveys for collection of site-specific data. As an alternative approach to field surveys, we assembled national agriculture statistical data (hereafter referred to as NAS) to compare to the satellite-derived rice estimates. Rice area statistics were obtained at the subnational level (analogous to a state, a county or district), and represent the total cropland area sown to rice, which double or triple count land areas that are multi-cropped with rice. NAS data for most countries (all except India, Bangladesh, Malaysia, Myanmar) were obtained from the website of the 'Regional Data Exchange System on Food and Agricultural Statistics in Asia and Pacific countries' (http://www.faorap-apcas.org/), a project funded by the Government of Japan and executed by the FAO. The data were supplied by the agricultural statistical offices of the various countries (Table 2) and, in most cases, contains data from the same year (2002) as our remote sensing analysis. NAS data obtained through the Regional Data Exchange System generally represents the total area sown to rice and does not subdivide the rice crop into farm management types (irrigated, upland, etc.). NAS data for India was obtained from the website of the Ministry of Agriculture's Directorate of Rice Development (http://dacnet.nic.in/rice/). India data was obtained at the district level for the year 2000 and represents the total area sown to rice. NAS data for Malaysia was obtained from the website of the Department of Agriculture (http://agrolink.moa.my/doa/BI/Statistics/jadual_perangkaan.html). Malaysia data was obtained at the state level for the year 1999 and represents the total area sown to rice in each of the two main rice seasons. NAS data for Bangladesh and Myanmar was obtained from the Huke and Huke (1997) dataset, a summary of rice sown area derived from agriculture census statistics from the early 1990s. Huke and Huke (1997) subdivide rice sown area into four main categories: irrigated, rainfed lowland, upland, and a deepwater.

Table 2
A summary of the national agricultural statistical datasets, including source agencies and dates of data used in this study (http://www.faorap-apcas.org/)

| Country | Date | Source agency |
|---|---|---|
| Bangladesh | 1993 | Huke and Huke (1997) |
| Bhutan | 2002 | Ministry of Agriculture |
| Cambodia | 2002 | Ministry of Agriculture, Forestry, and Fisheries |
| India | 2000 | Department of Agriculture and Cooperation; Ministry of Agriculture (http://dacnet.nic.in/rice/) |
| Indonesia | 2002 | BPS (Badan Pusat Statistik)—Statistics Indonesia |
| Laos | 2002 | Ministry of Agriculture and Forestry, Statistics Division |
| Malaysia | 1999 | Department of Agriculture (http://agrolink.moa.my/doa/BI/Statistics/jadual_perangkaan.html) |
| Myanmar | 1993 | Huke and Huke (1997) |
| Nepal | 2000 | Central Bureau of Statistics |
| Philippines | 2002 | Bureau of Agricultural Statistics; Department of Agriculture |
| Sri Lanka | 2002 | Department of Census and Statistics |
| Thailand | 2002 | Office of Agricultural Economics; Ministry of Agriculture and Cooperatives |
| Vietnam | 2002 | General Statistics Office |

IRRI classifies rice ecosystems into four categories: irrigated, rainfed lowland, upland, and deepwater (Maclean et al., 2002). Irrigated and rainfed lowland rice is grown in fields with small levees or dikes. Although irrigated rice accounts for only half of the world's rice land, more than 75% of the world's rice production comes from irrigated rice due to multi-cropping and improved technology (Maclean et al., 2002). Rainfed lowland rice, which accounts for 34% of the world's rice land, is flooded for at least part of the cropping season to water depths that exceed 100 cm for no more than 10 consecutive days (Maclean et al., 2002). Upland rice fields are generally not flooded, and dry soil preparation and direct seeding are common. It is important to note that the agriculture census data usually report the total area of rice cultivation and do not separate these four categories of rice fields. Our paddy rice algorithm is designed to identify those fields that could hold flooded/irrigated water for a period of a few weeks, and it is likely that the algorithm would fail to identify large portions of upland and deepwater rice. Therefore, the comparison between our MODIS-based map and the national rice statistical data needs to exclude both upland and deepwater rice, if the data are available.

We evaluated the MODIS-derived rice map in three ways: (1) spatial distribution of paddy rice, (2) national level comparison, and (3) subnational level comparison. We divided the 13 countries into four general geographical groups to conduct national and subnational level comparisons: (1) India; (2) Nepal, Bangladesh, Bhutan, Sri Lanka; (3) Myanmar, Thailand, Vietnam, Laos, Cambodia; and (4) Indonesia, Philippines, and Malaysia. These four regional groups were chosen because of general similarities of climate, landscape, and cropping systems among the countries in each group.

## 4. Results

### 4.1. Spatial distribution of paddy rice agriculture in South and Southeast Asia from MODIS-derived rice map

Fig. 4 shows the spatial distribution of paddy rice fields in 2002 across South and Southeast Asia at 500-m spatial resolution (hereafter referred to as $MOD_{rice}$). Paddy rice fields occur extensively, and are largely concentrated in the valleys and deltas of the major rivers in the region, such as the Mekong and Ganges river basins. To better facilitate the comparison between $MOD_{rice}$ and national agricultural statistics (NAS), we aggregated the 500-m $MOD_{rice}$ product to a subnational level polygon map. There are 1586 subnational administrative units in the study area, and we calculated percentage of rice area over the total land area for each of these units, normalizing comparisons between large and small districts. The spatial pattern of paddy rice from $MOD_{rice}$ (Fig. 5) is in general agreement with that of NAS (Fig. 6), but there are significant regional differences. These regional differences are outlined in more detail in Sections 4.2–4.5.

The $MOD_{rice}$ map estimates a total area of 766,810 km$^2$ of paddy rice fields in the 13 countries, which is about 78% of the total sown rice area (986,080 km$^2$) from the NAS dataset (Table 3). At the national level, 9 out of 13 countries have larger NAS estimates than $MOD_{rice}$ estimates. A simple linear regression model for rice area estimates of 13 countries between the $MOD_{rice}$ and NAS datasets has an $r^2=0.97$ and a root mean squared error (RMSE) of 31,072 km$^2$. The total paddy rice area from $MOD_{rice}$ is about 80% of the total sown acreage estimate (957,320 km$^2$) from the Huke dataset (Huke & Huke, 1997). The Huke dataset provides additional information on deepwater and upland rice (Table 3). When we exclude the areas of upland and deepwater rice, the Huke dataset estimates a rice area of 837,000 km$^2$ (sum of irrigated and rainfed rice), which is about 8% higher than the $MOD_{rice}$ estimate. At the country level, 9 out of 13 countries have larger Huke (irrigated plus rainfed rice) estimates than $MOD_{rice}$ estimates (Table 3). A simple linear regression model of rice area estimates of 13 countries between the $MOD_{rice}$ map and the Huke dataset (irrigated plus rainfed rice) has an $r^2=0.98$ and RMSE of 15,766 km$^2$. Discrepancies between the $MOD_{rice}$ and NAS datasets in some countries can be largely attributed to: (1) the NAS dataset is a sown area total that includes multiple cropping practices in a year and (2) failure of the 500-m resolution MODIS-based algorithm in identifying small patches of paddy rice fields, primarily in areas where topography poses restrictions to field sizes. These issues are further examined in the Discussion section.

### 4.2. Paddy rice agriculture in India

Rice-growing areas in India are primarily in the eastern coastal regions and the two great river basins in the northern part of the country, the Ganges and the Brahmaputra (Fig. 4a). Of the 44.6 million ha of harvested rice area in 2000, about 46% were irrigated, 28% were rainfed lowland, 12% were

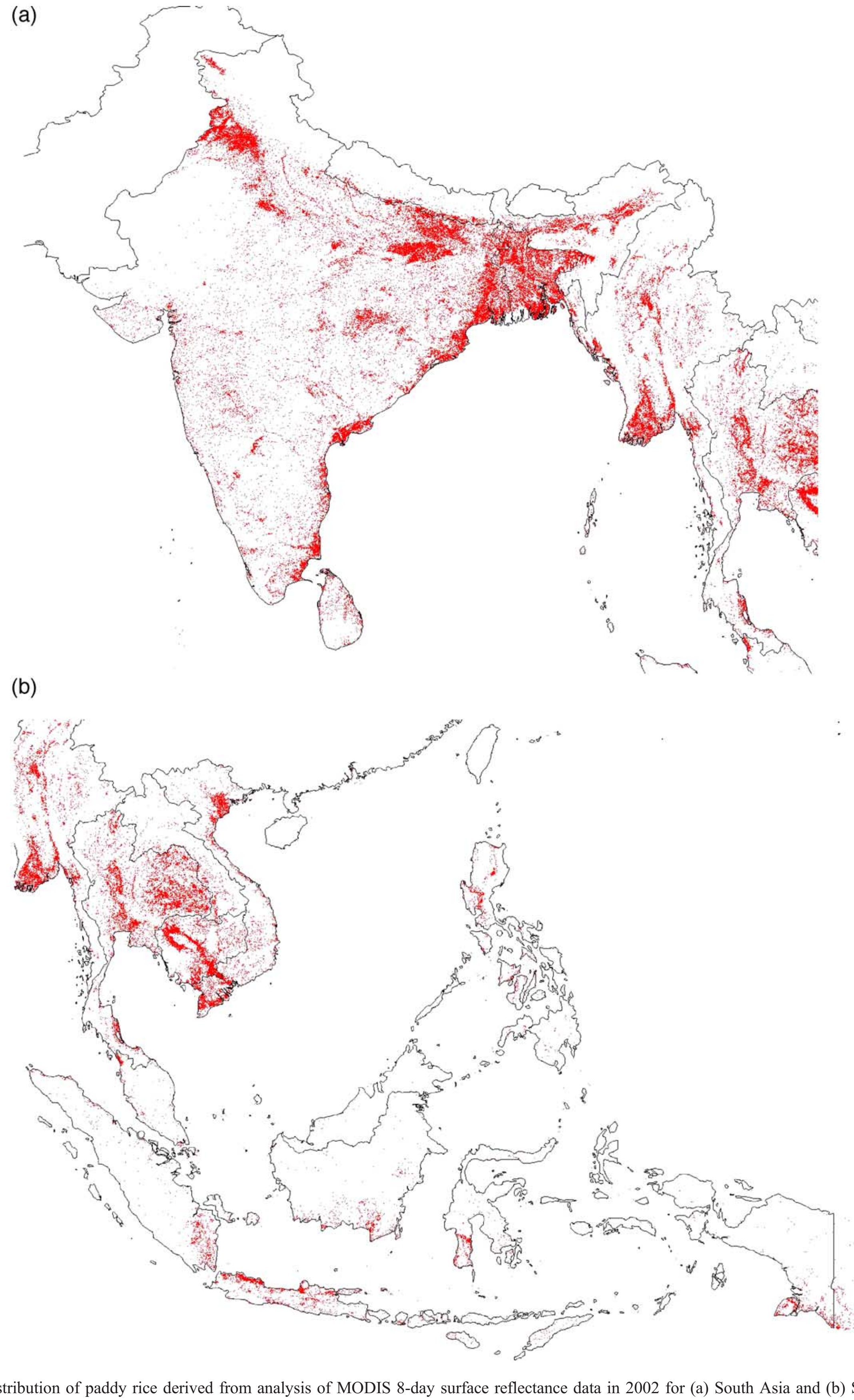


Fig. 4. Spatial distribution of paddy rice derived from analysis of MODIS 8-day surface reflectance data in 2002 for (a) South Asia and (b) Southeast Asia. The resultant paddy rice map has a spatial resolution of 500 m.

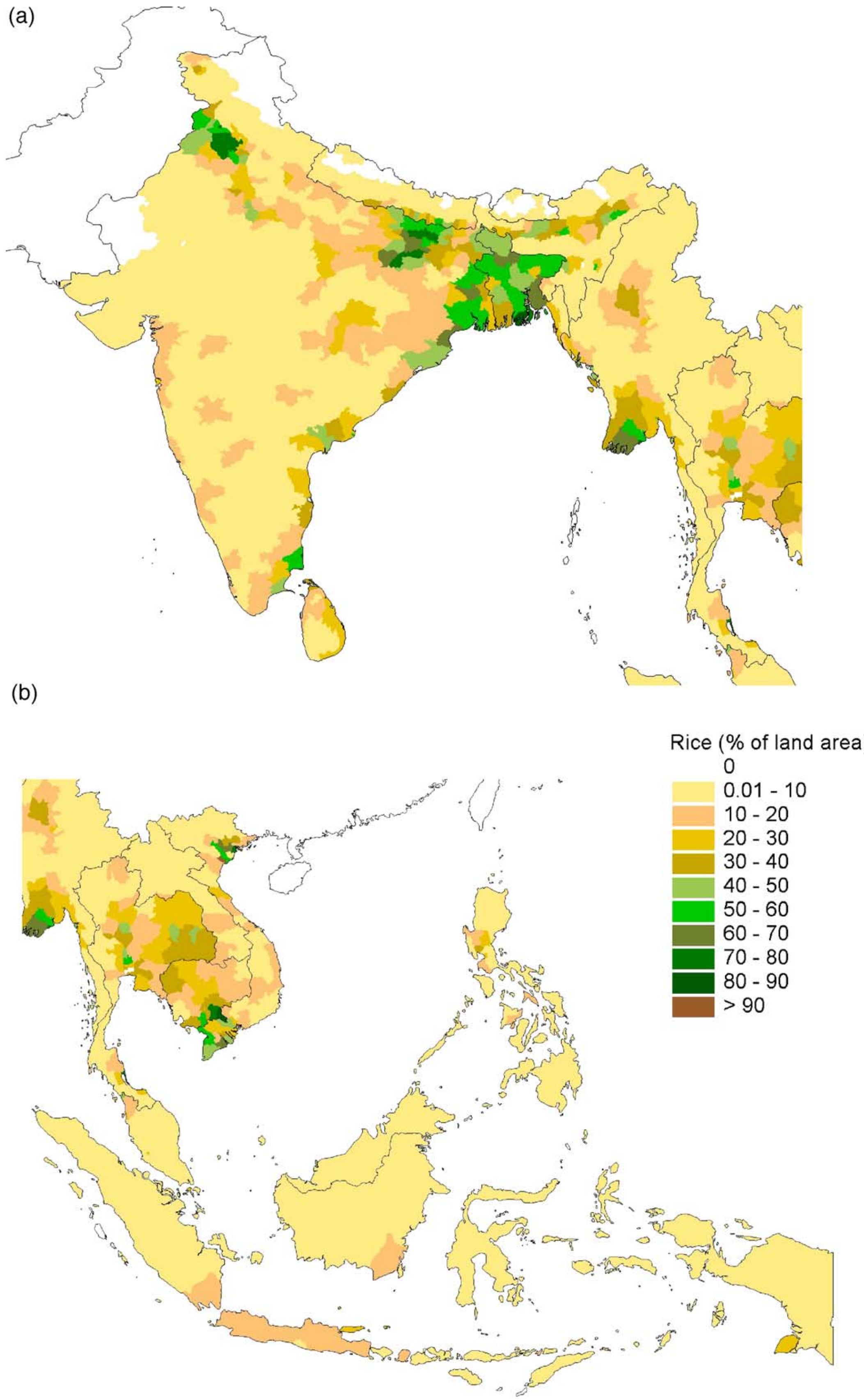


Fig. 5. Spatial distribution of paddy rice area at the district level in 2002, as aggregated from the MODIS-derived paddy rice map at 500 m (see Fig. 4) by administrative unit (district level). Rice area is displayed as the percentage of the district land area dedicated to paddy rice in (a) South Asia and (b) Southeast Asia.

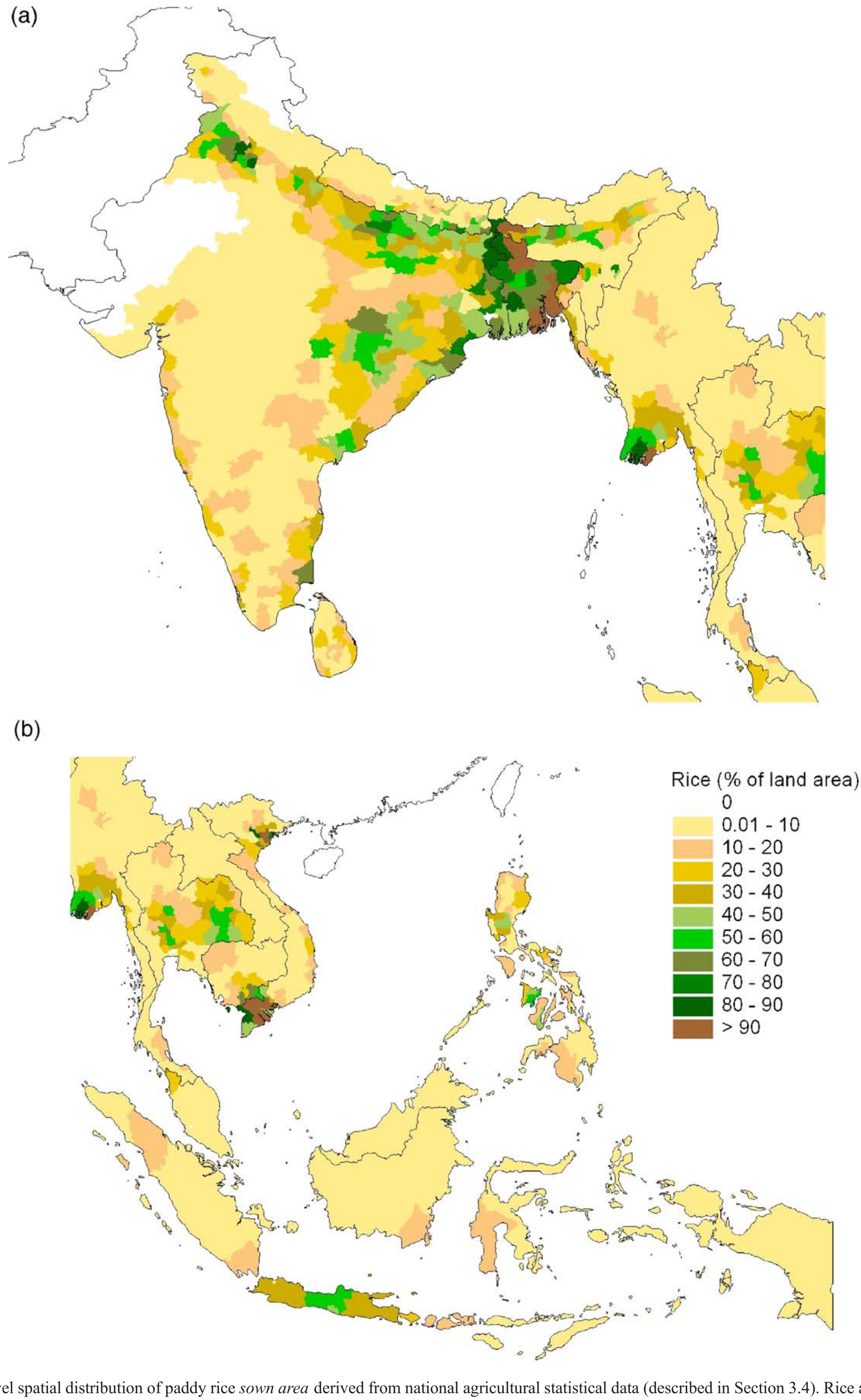


Fig. 6. District-level spatial distribution of paddy rice *sown area* derived from national agricultural statistical data (described in Section 3.4). Rice area is displayed as the percent of the district land area dedicated to paddy rice in (a) South Asia and (b) Southeast Asia.

Table 3
National-level rice area estimates (× 000 ha) derived from three data sources: Huke and Huke (1997), national agricultural statistics (NAS, see Section 3.4), and the $MOD_{rice}$ algorithm (this study)

| Country | Huke rice category (× 000 ha) | | | | | | NAS (× 000 ha) | Paddy intensity | $MOD_{rice}$ (× 000 ha) | $MOD_{rice}$ × paddy intensity (× 000 ha) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Upland | Deepwater | Irrigated paddy | Rainfed paddy | Total (columns 2–5) | Total (irrigated + rainfed) | | | | |
| Bangladesh | 698 | 1221 | 2617 | 6144 | 10,680 | 8761 | 11,000 | 1.75[a] | 6322 | 11,064 |
| Bhutan | 4 | 0 | 5 | 17 | 26 | 22 | 19 | 1.0[b] | 4 | 4 |
| Cambodia | 24 | 152 | 305 | 1418 | 1900 | 1723 | 1995 | 1.24[c] | 4242 | 5260 |
| India | 5060 | 1364 | 19,660 | 16,432 | 42,516 | 36,092 | 43,278 | 1.10[d] | 34,447 | 37,892 |
| Indonesia | 1209 | 2 | 5926 | 3878 | 11,015 | 9804 | 11,521 | 1.40[e] | 6740 | 9436 |
| Laos | 219 | 0 | 44 | 348 | 611 | 392 | 738 | 1.12[f] | 989 | 1111 |
| Malaysia | 80 | 0 | 438 | 150 | 668 | 588 | 609 | 1.43[g] | 489 | 699 |
| Myanmar | 214 | 362 | 3198 | 2511 | 6285 | 5709 | 6488 | 1.43[h] | 6724 | 9615 |
| Nepal | 68 | 118 | 730 | 572 | 1488 | 1302 | 1560 | 1.07[i] | 811 | 868 |
| Philippines | 165 | 0 | 2204 | 1252 | 3621 | 3456 | 4046 | 1.70[j] | 1484 | 2523 |
| Sri Lanka | 0 | 0 | 628 | 239 | 867 | 867 | 820 | 1.46[k] | 783 | 1143 |
| Thailand | 203 | 342 | 939 | 8160 | 9644 | 9099 | 9105 | 1.15[l] | 9306 | 10702 |
| Vietnam | 322 | 177 | 3260 | 2614 | 6373 | 5874 | 7504 | 1.83[a] | 4265 | 7805 |
| Total | 8265 | 3737 | 39,995 | 43,735 | 95,732 | 83,730 | 98,683 | 1.28 | 76,606 | 98,118 |

Sources for paddy intensity statistics are footnoted.
[a] Maclean et al. (2002).
[b] No statistics available. An intensity value of 1 was assigned based on its geographic location.
[c] http://www.cardi.org.kh/Library/AgStats.htm.
[d] Frolking and Babu (submitted for publication).
[e] http://www.indonesiaphoto.com/content/view/148/45/.
[f] http://www.faorap-apcas.org/lao/busdirectory/search_results.asp.
[g] http://www.fao.org/ag/agl/swlwpnr/reports/y_ta/z_my/my.htm#s125.
[h] http://www.fao.org/docrep/003/x0736m/rep2/myanmar.htm.
[i] http://www.riceweb.org/countries/nepal.htm (site visited in 2004).
[j] http://www.fao.org/ag/agl/swlwpnr/reports/y_ta/z_ph/ph.htm#s126.
[k] http://www.faorap-apcas.org/srilanka/busdirectory/search_results.asp.
[l] http://oae.go.th/statistic/yearbook/2001-02/indexe.html.

upland, and 14% were deepwater (Maclean et al., 2002). Depending on the location, Indian rice is grown in the *kharif* (summer, wet) or *rabi* (winter, dry) seasons, or both. The majority of rice is grown during the kharif season, which is a combination of both rainfed and irrigated. Rice grown in the rabi season is primarily irrigated.

At the state level (31 states in India), there are similar spatial patterns of rice fields between the $MOD_{rice}$ (Fig. 5a) and NAS (Fig. 6a) datasets. In many states, the NAS data had greater fractional areas of rice due to multiple-cropping, such as in east–central India (Figs. 5a and 6a). The four states that have the largest amount of upland rice (Orissa, Madhya Pradesh, West Bengal, and Uttar Pradesh) also tend to have relatively large differences between the $MOD_{rice}$ and NAS rice datasets (Table 4). At the state level, the $MOD_{rice}$ and NAS datasets are highly correlated, with an $r^2$ value of 0.90 and an RMSE of 8093 $km^2$ (Fig. 7a). These results suggest that the MODIS algorithm is capable of identifying the majority of irrigated and rainfed lowland rice fields at the state level, but may miss large portions of upland rice fields.

There are 457 administrative districts in India. At the district level, there is a positive correlation between the $MOD_{rice}$ and NAS datasets ($r^2 = 0.47$, RMSE = 809 $km^2$; Fig. 7b), but is lower than the agreement at the state level. A similar pattern of decreased agreement from the provincial level to the county level also occurred in a previous study in southern China (Xiao et al., 2005). The impacts based on the spatial resolution of MODIS, such as problems detecting small fields, tend to be more pronounced at the district level.

### *4.3. Paddy rice agriculture in Nepal, Bangladesh, and Sri Lanka*

Agriculture in Nepal occurs on a thin strip of plains in the southern portion of the country and the vast majority of rice is either irrigated or rainfed (Table 3). At the national level, the total area of paddy rice fields from $MOD_{rice}$ was just over half of the total rice area from the NAS dataset. At the district level, the overall agreement between the $MOD_{rice}$ and NAS datasets was positive ($r^2 = 0.48$, RMSE = 198 $km^2$; Fig. 8a), but a number of districts with under 200 $km^2$ of rice area (NAS) had little or no rice detected by the MODIS algorithm. It is possible that these are areas where the increased complexity of topography restricts the size of rice fields that can occur, with much of the rice growing on terraced slopes.

In Bangladesh, nearly 50% of the cropland is double cropped and 13% is triple cropped (Maclean et al., 2002). As a result, much of the country has areas where the fraction of sown rice is over 90% of the land area (Fig. 6a). Rice ecosystems in Bangladesh are dominated by rainfed (over 50% of the rice area) and irrigated, although significant amounts of upland and deepwater rice still exist. The national rice area from $MOD_{rice}$ is substantially lower than the total sown area of rice fields from NAS (Table 3). However, much of this

Table 4
Rice area (×000 ha) estimates for India states derived from three data sources: Huke and Huke (1997), national agricultural statistics (NAS; see Section 3.4), and the $MOD_{rice}$ algorithm

| State | Huke rice category (×000 ha) | | | | | | NAS (×000 ha) | Paddy intensity | $MOD_{rice}$ (×000 ha) | $MOD_{rice}$ × paddy intensity (×000 ha) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Upland | Deepwater | Irrigated paddy | Rainfed paddy | Total (columns 2–5) | Total (irrigated+rainfed) | | | | |
| Andhra Pradesh | 105 | 42 | 3859 | 0 | 4006 | 3859 | 3828 | 1.26 | 2853 | 3582 |
| Assam | 544 | 272 | 530 | 1144 | 2490 | 1674 | 2503 | 1.10 | 2134 | 2339 |
| Bihar | 510 | 457 | 1954 | 2473 | 5393 | 4427 | 4987 | 1 | 5216 | 5216 |
| Gujarat | 0 | 0 | 215 | 315 | 531 | 531 | 608 | 1 | 867 | 867 |
| Jammu and Kashmir | 0 | 0 | 266 | 0 | 266 | 266 | 271 | 1 | 276 | 276 |
| Karnataka | 110 | 0 | 852 | 204 | 1166 | 1056 | 1354 | 1.20 | 1134 | 1363 |
| Kerala | 30 | 0 | 256 | 273 | 559 | 529 | 463 | 1.14 | 242 | 277 |
| Madhya Pradesh | 840 | 0 | 994 | 3228 | 5062 | 4222 | 5298 | 1 | 3429 | 3429 |
| Maharashtra | 351 | 0 | 331 | 900 | 1581 | 1231 | 1525 | 1.01 | 2152 | 2183 |
| Orissa | 853 | 67 | 1556 | 1928 | 4404 | 3484 | 4495 | 1.09 | 2624 | 2862 |
| Rajasthan | 0 | 0 | 120 | 0 | 120 | 120 | 152 | 1 | 725 | 725 |
| Tamil Nadu | 20 | 23 | 1830 | 0 | 1873 | 1830 | 2156 | 1.12 | 2113 | 2366 |
| Uttar Pradesh | 549 | 218 | 2570 | 2277 | 5615 | 4847 | 5604 | 1 | 3126 | 3126 |
| West Bengal | 840 | 253 | 1251 | 3469 | 5813 | 4720 | 5866 | 1.25 | 3497 | 4379 |
| Himachal Pradesh | 0 | 0 | 85 | 0 | 85 | 85 | 83 | 1 | 45 | 45 |
| Haryana | 0 | 0 | 667 | 0 | 667 | 667 | 905 | 1 | 830 | 830 |
| Punjab | 0 | 0 | 2024 | 0 | 2024 | 2024 | 2234 | 1 | 2579 | 2579 |
| Other States | 309 | 32 | 300 | 220 | 861 | 520 | 945 | 1.07 | 605 | 647 |
| TOTAL | 5060 | 1364 | 19,660 | 16,432 | 42,516 | 36,092 | 43,278 | 1.10 | 34,447 | 38,034 |

State-level paddy intensity values are from Frolking and Babu (submitted for publication).

discrepancy is likely to be attributed to the double or triple rice cropping over a significant portion of the cropland in Bangladesh. The spatial distribution of the rice is similar between these two datasets (Figs. 5a and 6a) and the subnational (region) agreement in area estimates is good ($r^2$=0.70, RMSE=2021 km$^2$; Fig. 8b).

In Sri Lanka, multiple rice cropping occurs in parts of the country and all of the rice is either irrigated or rainfed (Table 3). At the national level, the rice area estimates from $MOD_{rice}$ and NAS are quite close (Table 3). However, at the district level, the correlation in rice area estimates between these two datasets (Fig. 8c) is the weakest of all of the 13 countries analyzed in this study ($r^2$=0.31, RMSE=293 km$^2$), likely due to smaller amounts of scattered rice areas. In most districts of Sri Lanka, the fraction of the land area that is sown or planted to rice is under 30% (Figs. 5a and 6a).

### *4.4. Paddy rice agriculture in Myanmar, Thailand, Vietnam, Laos, and Cambodia*

In Myanmar, rice cultivation occurs throughout much of the northern part of the country, but the majority of the rice production occurs in the delta areas of the Ayeyarwady and Sittoung rivers (Fig. 4a). Of the total rice area, rainfed rice accounts for 52%, irrigated rice is 18%, deepwater rice is 24%, and upland rice is 6% (Maclean et al., 2002). At the national level, the difference in rice area estimates between the $MOD_{rice}$ and NAS datasets are within 2360 km$^2$, approximately 3.5% of the total paddy rice area from $MOD_{rice}$ (Table 3). At the district level, the spatial distribution of rice fields derived from the MODIS algorithm and NAS are very similar, except for a slight overestimation by $MOD_{rice}$ in the northern interior areas (Figs. 5a and 6a). At the district level, the correlation between these two datasets is also good ($r^2$=0.74, RMSE=679 km$^2$; Fig. 9a).

While rice is distributed over much of Thailand (Fig. 4b), nearly half of the rice land is located in the northeast interior region, where the majority of the rice fields are rainfed. At a national level, the difference in rice area estimates between the $MOD_{rice}$ and NAS datasets are within 2010 km$^2$, about 2.2% of the total paddy rice area from the $MOD_{rice}$ dataset (Table 3). At the subnational province level, the spatial distribution of rice derived from $MOD_{rice}$ and NAS are very similar, except for a slight underestimation by the MODIS algorithm in the northeast rainfed region (Figs. 5b and 6b). At the province level, the correlation between these two datasets is also high ($r^2$=0.87, RMSE=481 km$^2$; Fig. 9b).

In Vietnam, much of the rice cultivation is concentrated in two river deltas, the Mekong (over half of the country's rice area) and the Red (Fig. 4b). The rice sown area in 2000 was about 7.7 million ha and the cropping intensity (ratio of sown area to land area for a given crop) was about 183% (Maclean et al., 2002), the highest in the world. The high cropping intensity is largely due to the triple rice crops that are common in much of the Mekong Delta area. Over 92% of the total rice area in Vietnam is either irrigated or rainfed (Table 3). At a national level, the paddy rice area from $MOD_{rice}$ is substantially lower than the total rice area reported in the NAS dataset (Table 3) due to the high cropping intensity. At the subnational province level, the correlation in rice area estimates between the $MOD_{rice}$ and NAS datasets is positive ($r^2$=0.42, RMSE=1183 km$^2$; Fig. 9c), but is weaker than in Thailand or Myanmar.

In Laos, over 35% of the rice crop is upland (Table 3), the largest percentage of any country in our study. At the national level, the rice area from $MOD_{rice}$ is substantially higher than the rice area from the NAS dataset (Table 3). However, at the

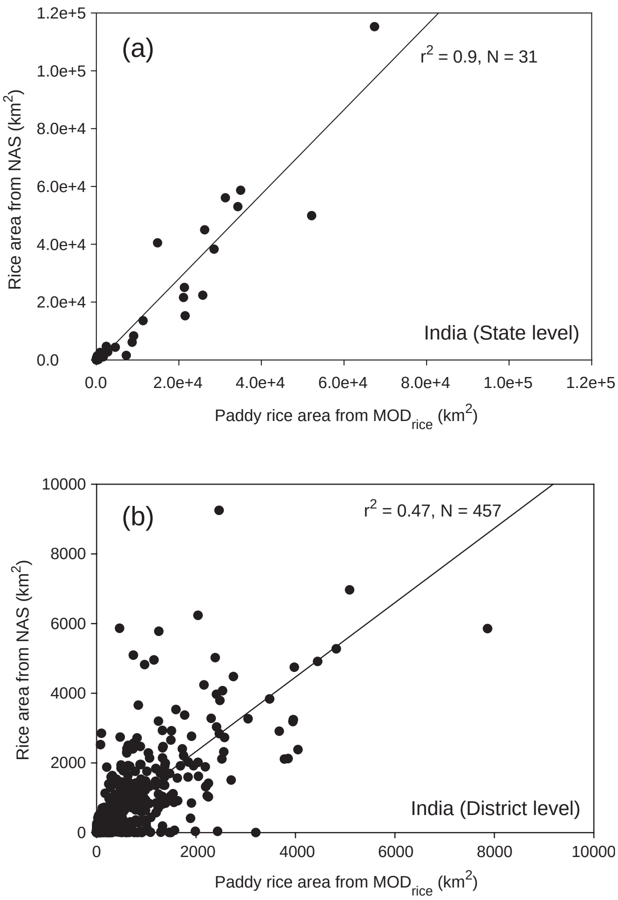


Fig. 7. A comparison of rice area in India between the MODIS rice algorithm ($MOD_{rice}$) and national agriculture statistics (see Section 3.4) at (a) state level and (b) district level.

subnational province level, these two datasets were well correlated ($r^2$=0.79, RMSE=457 km$^2$; Fig. 9d), indicating that the spatial distribution of rice was similar even though the $MOD_{rice}$ estimates were higher.

Rice in Cambodia is concentrated in the lowlands surrounding lake Tonle Sap and the lower reaches of the Mekong River in the southern part of the country (Fig. 4b). The majority of rice in this poor country is rainfed (Table 3). At the national level, the paddy rice area estimate from $MOD_{rice}$ is substantially higher than the total rice area reported in NAS (Table 3); possible reasons for this large discrepancy are raised in the Discussion section. At the subnational province level, the correlation in area estimates between the $MOD_{rice}$ and NAS datasets is positive ($r^2$=0.44, RMSE=1664 km$^2$; Fig. 9e).

### *4.5. Paddy rice agriculture in Malaysia, Philippines, and Indonesia*

Most of Malaysia's rice cultivation occurs in the northwest corner of the peninsular section, close to the Thailand border (Fig. 4b), and almost 90% of it is irrigated or rainfed (Table 2). At the national level, the difference in area estimates between the $MOD_{rice}$ and NAS datasets was relatively small, especially if the upland component in the Huke dataset is excluded (Table 3). At the subnational state level, these two datasets were well correlated ($r^2$=0.71, RMSE=350 km$^2$; Fig. 10a) and had similar spatial distribution patterns of paddy rice fields (Figs. 5b and 6b).

In the Philippines, over 95% of all rice is either irrigated or rainfed (Table 3), with the remainder being upland. At the national level, the paddy rice area estimate from $MOD_{rice}$ is less than half the total rice area of the NAS dataset (Table 3), likely due to a combination of multi-cropping and topographic restraints on field sizes. At the subnational region level, these two datasets were moderately correlated ($r^2$=0.60, RMSE=

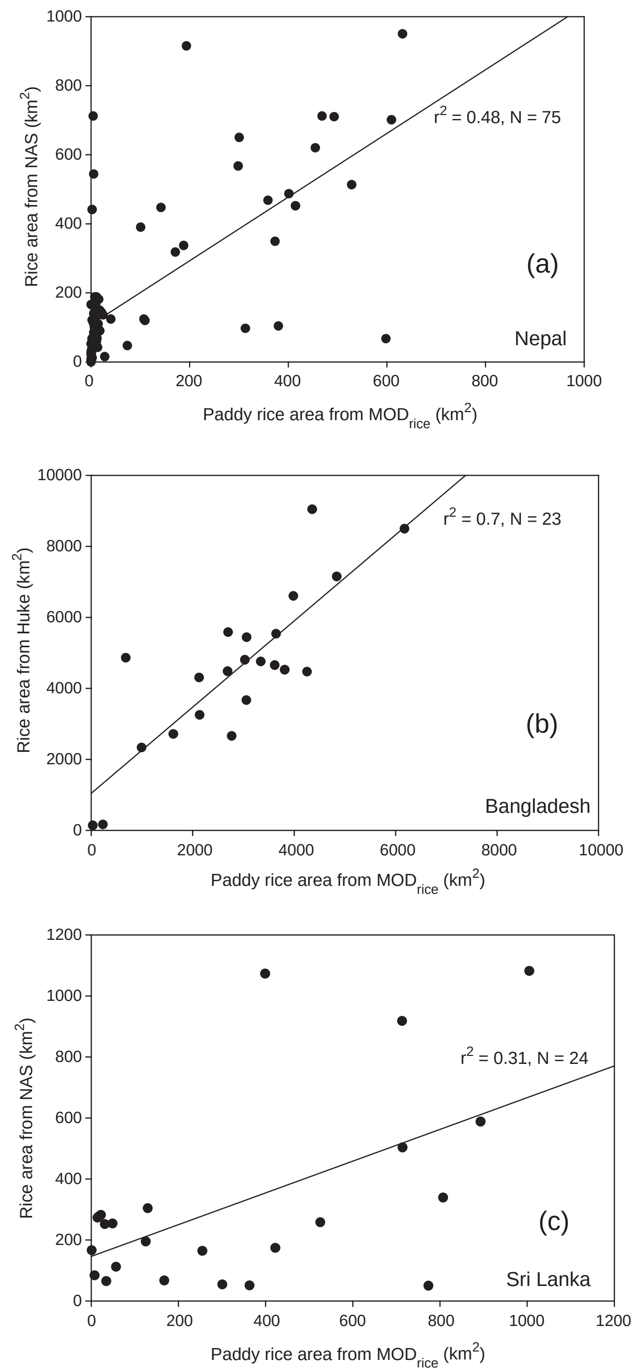


Fig. 8. A subnational comparison of rice area between the MODIS rice algorithm ($MOD_{rice}$) and national agriculture statistics (see Section 3.4) in (a) Nepal, (b) Bangladesh, and (c) Sri Lanka.

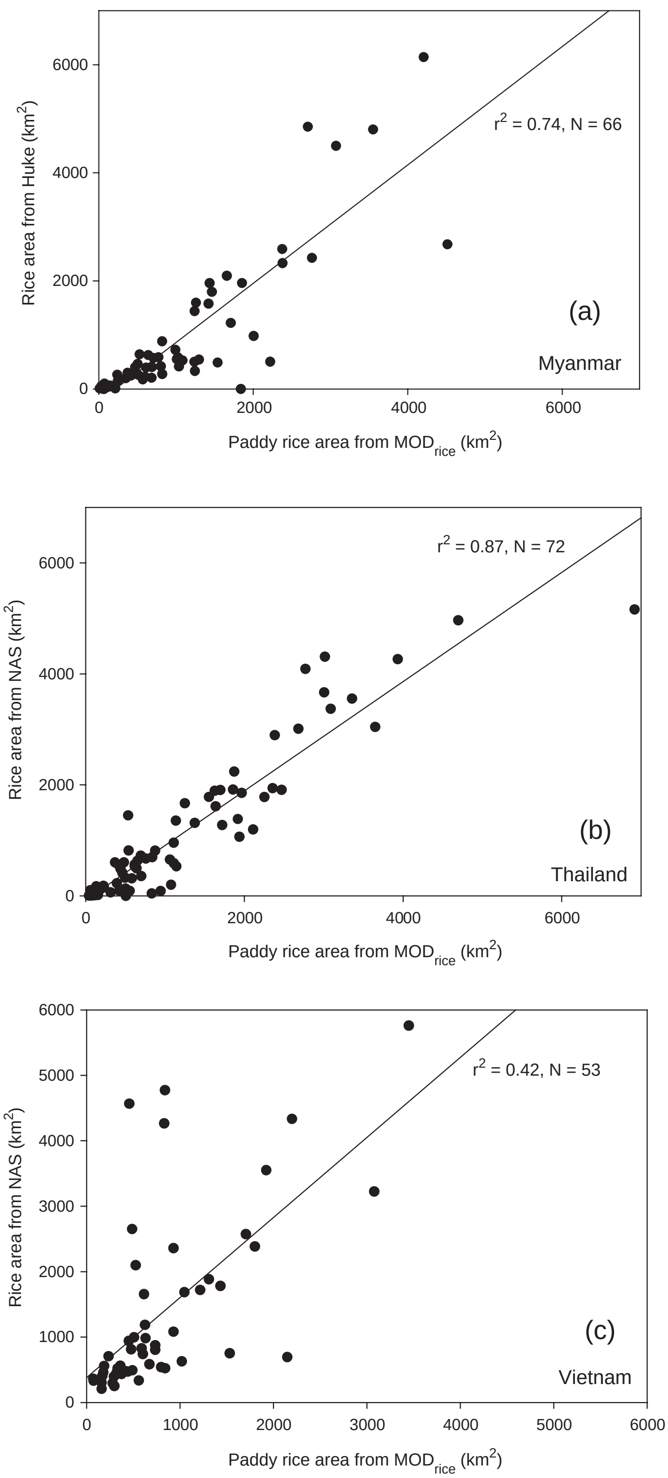


Fig. 9. A subnational comparison of rice area between the MODIS rice algorithm ($MOD_{rice}$) and national agriculture statistics (see Section 3.4) in (a) Myanmar, (b) Thailand, (c) Vietnam, (d) Laos, and (e) Cambodia.

534 $km^2$; Fig. 10b). $MOD_{rice}$ had much lower fractional rice areas in several of the central islands than the NAS dataset (Figs. 5b and 6b).

In Indonesia, irrigated and rainfed rice account for almost 90% of the total rice area, with the remaining 11% being upland (Maclean et al., 2002). While each of Indonesia's five main islands has some areas of intense rice production, heavily populated Java is the most productive rice area (Fig. 4b). At the national level, the paddy rice area estimate from $MOD_{rice}$ is lower than the total rice area estimate from the NAS dataset (Table 3). This discrepancy is likely a result of high cropping intensities and greater sown area totals in the NAS dataset, as much of the irrigated and rainfed areas are double cropped in Indonesia (Maclean et al., 2002). At the subnational province

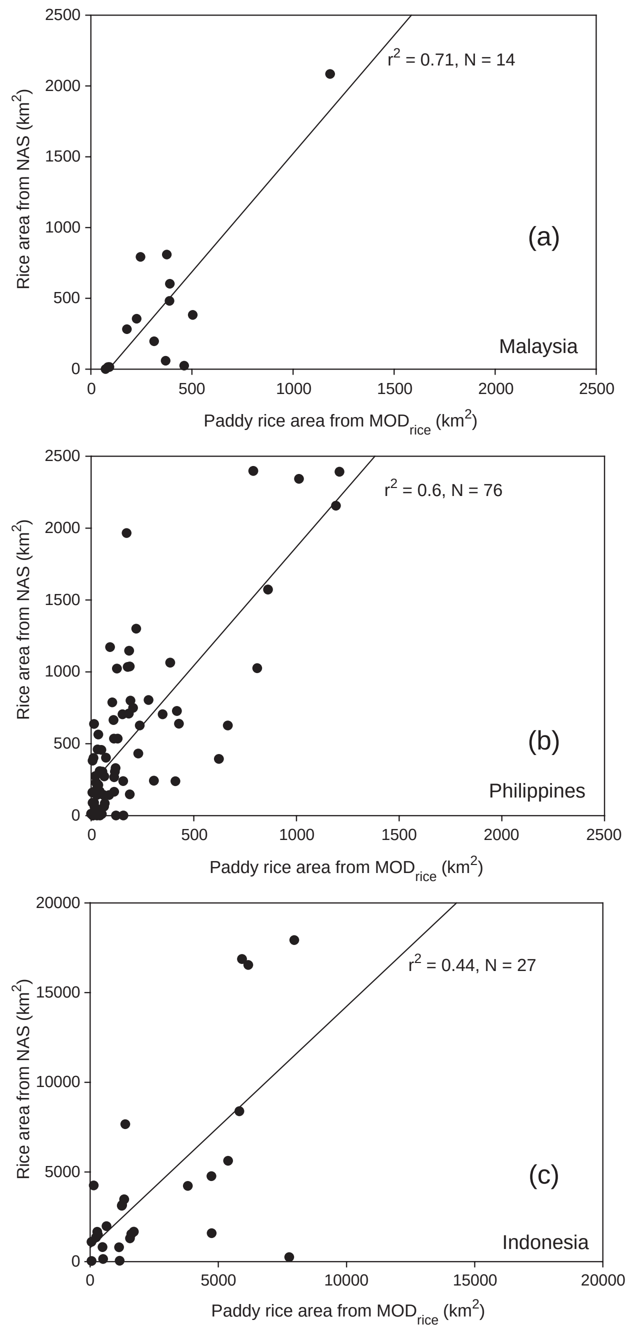


Fig. 10. A subnational comparison of rice area between the MODIS rice algorithm ($MOD_{rice}$) and national agriculture statistics (see Section 3.4) in (a) Malaysia, (b) Philippines, and (c) Indonesia.

level, these two datasets were moderately correlated ($r^2$=0.44, RMSE=4201 km$^2$; Fig. 10c). Spatial distribution patterns of paddy rice from the $MOD_{rice}$ and NAS datasets are similar, although $MOD_{rice}$ has considerably lower fractional rice areas on primarily double-cropped Java (Figs. 5b and 6b).

## 5. Discussion

In this study, we used a temporal profile analysis of MODIS-derived vegetation indices to identify and map paddy rice over 13 countries in South and Southeast Asia. While there are several factors that can affect rice mapping using our MODIS method (sensor temporal and spatial resolution, cloud cover, snow, seasonally inundated wetlands), at the outset of our study, we were most concerned about the impact of frequent cloud cover in subtropical and tropical Asia, where paddy rice fields are widely distributed (Thenkabail et al., 2005). An innovative feature of our paddy rice algorithm is that rice paddies are identified from a relatively short time period of image data during the flooding and early growth period. One benefit of this approach is that one does not need cloud-free observations throughout the entire year or crop cycle for image classification purposes; one can obtain reasonable results as long as cloud-free observations occur within the short period of flooding and rice transplanting. Farmers generally select sunny days for rice transplanting, because continuous cloudy/rainy days could reduce the growth of young seedlings, and excessive water level in the fields could potentially result in die-back of seedlings. Cloud contamination affects between 35% and 45% of all land pixels during the height of the monsoon (June–August), and between 10% and 30% of land pixels the rest of the year. While these levels of cloud contamination introduce some degree of underestimation in the MODIS-derived rice areas, the results of this study suggest that our algorithm to a large degree overcomes the obstacle associated with frequent cloud cover occurrence in moist tropical Asia. Our paddy rice mapping algorithm conducts image classification pixel by pixel and is different from conventional image classification algorithms that were built upon spatial pattern recognition (Friedl et al., 2002; Loveland et al., 2000; Thenkabail et al., 2005; Xiao et al., 2002a). The latter approach is likely to have large error and uncertainty if one uses two moderate-resolution maps generated from algorithms based on spatial pattern recognition to infer land-cover and land-use changes. The mean and standard deviation of spectral clusters change when different years of satellite images are used, and interpretation of land-cover classes become less objective and more difficult. The temporal profile analysis of individual pixels is readily applicable to different years for quantifying changes in crop calendars and multiple cropping rotations, and thus has a potential for substantially reducing the error and uncertainty in quantifying land-cover conversion and land-use intensification.

Although the spatial distribution of paddy rice from $MOD_{rice}$ agrees reasonably with the spatial pattern of rice agriculture from the agricultural census data (NAS), there are significant regional differences among the 13 countries. Four factors may to various degrees contribute to the discrepancies between the $MOD_{rice}$ and NAS rice area estimates. First, both the NAS and Huke datasets used in this study are sown area statistics, including double- to triple cropping of paddy rice in a year, which leads to double or multiple counting of the area of paddy rice fields. The paddy intensity (ratio of sown paddy area to paddy land area) is a statistic that can be used to provide a direct comparison between $MOD_{rice}$ and NAS rice area estimates. If the paddy intensity is used as a multiplier, an estimate of sown area can be derived from $MOD_{rice}$ totals (Tables 3 and 4). Using this method, discrepancies between

MODIS- and NAS-derived rice areas of several countries (Vietnam, Bangladesh, India, Indonesia, and Philippines) were greatly reduced. While the discrepancies in a few countries (Myanmar, Cambodia, and Thailand) were increased using this method of comparison, the overall discrepancy for our Asian study area was reduced from over 22,000 km$^2$ (without the paddy intensity multiplier) to just 565 km$^2$ (Table 3). Second, when using MODIS data at 500-m spatial resolution, the algorithm could fail to identify paddy rice fields in regions with complex topographic relief and/or locations where paddy rice fields are much smaller than the MODIS pixels can resolve satisfactorily, which leads to underestimation of the area of paddy rice fields. In an earlier study (Xiao et al., 2005), a similar pattern was observed as the MODIS rice algorithm underestimated rice areas in the hilly provinces of southern China. To explore this phenomenon further, we summarized the total rice area of the two datasets (MOD$_{rice}$, NAS) according to elevation (Fig. 11). It is evident that the MODIS algorithm more consistently approximates the NAS dataset at lower elevations (<300m), with MOD$_{rice}$ totals within 20% of NAS totals in each elevation range that was analyzed. Almost 80% of all rice paddies occur at elevations less than 300 meters (Fig. 11); these areas tend to be large, flat river mouths and valleys with a relatively homogenous land-cover consisting primarily of rice paddies. At higher elevations, agreements between MOD$_{rice}$ and NAS are much less consistent (Fig. 11), although the level of consistency does not seem to decrease with increasing elevation. Third, pixels contaminated by clouds represent approximately 10–40% of all land pixels, depending on the time of year. As these pixels are not included in the MODIS algorithm, this introduces a potential source of underestimation in MOD$_{rice}$. Fourth, the NAS dataset has its own uncertainty in the reporting process due to a number of social and economic issues (including politics, tax), particularly at subnational levels. For instance, agricultural census statistics in China are routinely under-reported (Frolking et al., 2002), a practice that could potentially occur in other nations also.

To decrease the discrepancies one would require fine resolution satellite images (e.g., Landsat) to be acquired at several times during the year. In previous studies that used moderate resolution images for land-cover classification, Landsat images were often used for accuracy evaluation (Boles et al., 2004; Eva et al., 2004; Latifovic et al., 2004; Loveland et al., 2000). In an earlier study for mapping paddy rice in southern China, we used a Landsat-based land-cover map to evaluate the MODIS-based algorithm and resultant map of paddy rice fields (Xiao et al., 2005). In this study, we do not have sufficient resources to generate Landsat-derived land-cover maps for the 13 countries, and we expect that further evaluation may be carried out after the MODIS-derived paddy rice field dataset is released to the public for comments. In the next paragraph, we present a case study in Cambodia to illustrate the role of Landsat images in accuracy evaluation.

The largest percentage discrepancy in rice area estimates between the MOD$_{rice}$ and NAS datasets occurs in Cambodia (Table 3). The paddy rice area from MOD$_{rice}$ is substantially higher than rice area reported in the NAS dataset and is largely concentrated around the Tonle Sap Basin in the northwest part of Cambodia (Figs. 4–6). The basin covers an area of about 80,000 km$^2$ and total agricultural land area is about 15,787 km$^2$ (Wright et al., 2004). The Tonle Sap Basin is one of the regions in Cambodia that has a large area of rice fields and high rice productivity. The wet (monsoon) season usually occurs between May and October and the dry season occurs between November and April. Correspondingly, Cambodia has a wet-season rice crop and a dry-season rice crop. For the wet-season rice crop, the seedling-bed process starts in late May through July when the first rains of the monsoon begin to inundate and soften the land. Rice seedlings are transplanted from late June through September. Rice harvesting usually occurs in late November to December. The dry-season rice crop is usually planted in November in some areas that have trapped or retained part of the monsoon rains or irrigation infrastructure, and the cropping cycle (planting to harvest) is about 3 months. The area of dry-season rice crop is generally small, because of lack of irrigation infrastructure. We compared the MOD$_{rice}$ map with a Landsat ETM+ image that covers a large part of the Tonle Sap Basin (Fig. 12). The ETM+ image on January 11,

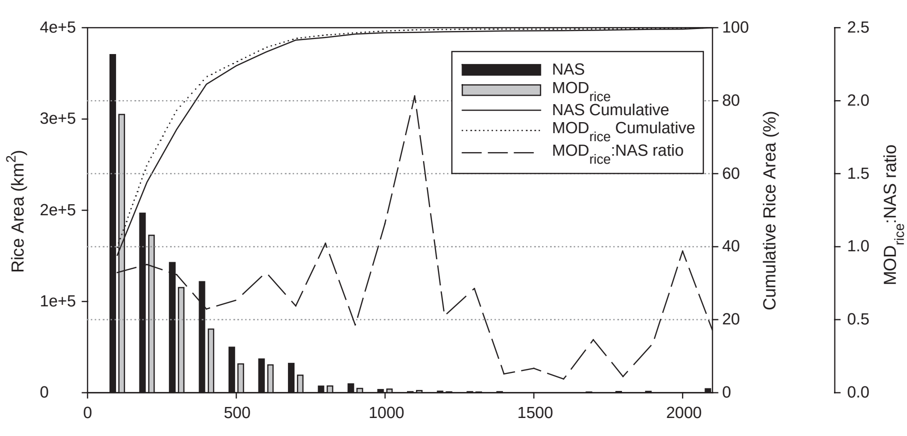


Fig. 11. The relationship between rice area and elevation (GTOPO30 digital elevation model; see Methods) for the study area. The solid and dotted lines represent the cumulative percentage of national agricultural statistic (NAS) and MOD$_{rice}$ rice areas, respectively. The dash line represents the ratio of MODIS-derived rice area to NAS rice area.

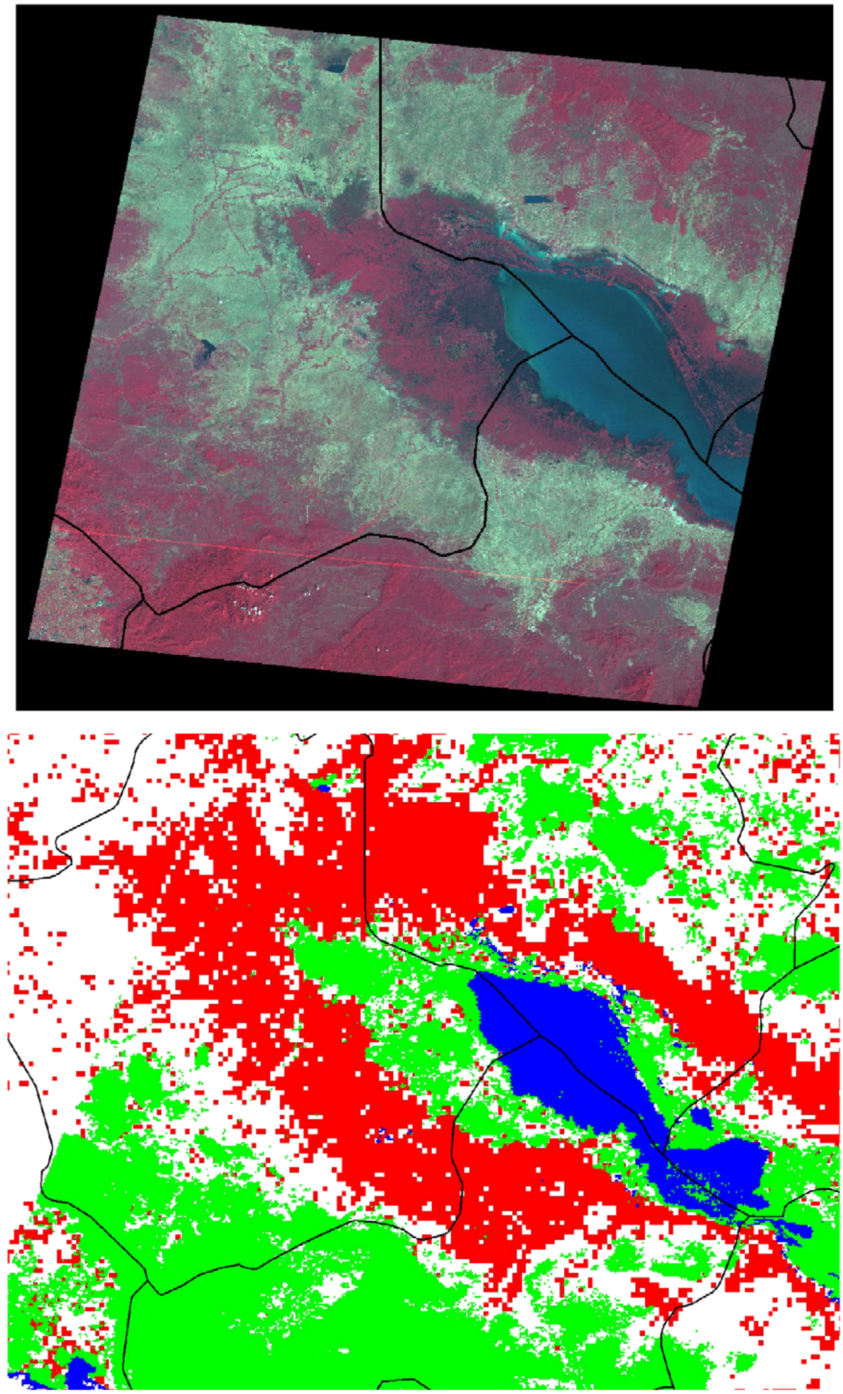

Fig. 12. A comparison between the $MOD_{rice}$ map and a Landsat ETM+ image on January 11, 2002 in the Tonle Sap Basin, Cambodia. The provincial boundary map is overlaid to aid in the comparison of the two maps. The upper panel is a false color composite of band 4-3-2 of Landsat ETM+ image. In the lower panel, paddy rice = red color, permanent water body = blue color, and forest mask = green color.

2002 clearly shows the spatial distribution of wetland (evergreen forests) surrounding Tonle Sap and areas of harvested cropland. The paddy rice fields identified in $MOD_{rice}$ are in general spatial agreement with the areas of harvested cropland in the ETM+ image. The wetland areas immediately surrounding the open water are likely to be flooded during the wet season. These flooded forest areas are not identified as paddy rice fields in $MOD_{rice}$, which indicates that our paddy rice mapping algorithm successfully eliminates the potential errors introduced by seasonally flooded wetlands (see Methods and Fig. 3). We also examined time series of MODIS vegetation indices (NDVI, EVI, and LSWI) for a number of rice pixels in the basin. Time series of vegetation indices from two pixels that are identified as paddy rice (Fig. 13) show that these pixels have the unique spectral signature we identify as paddy rice fields. Note that, because of budget limitations, we were not able to conduct field work in the Tonle Sap Basin to evaluate our paddy rice map. The time series of MODIS

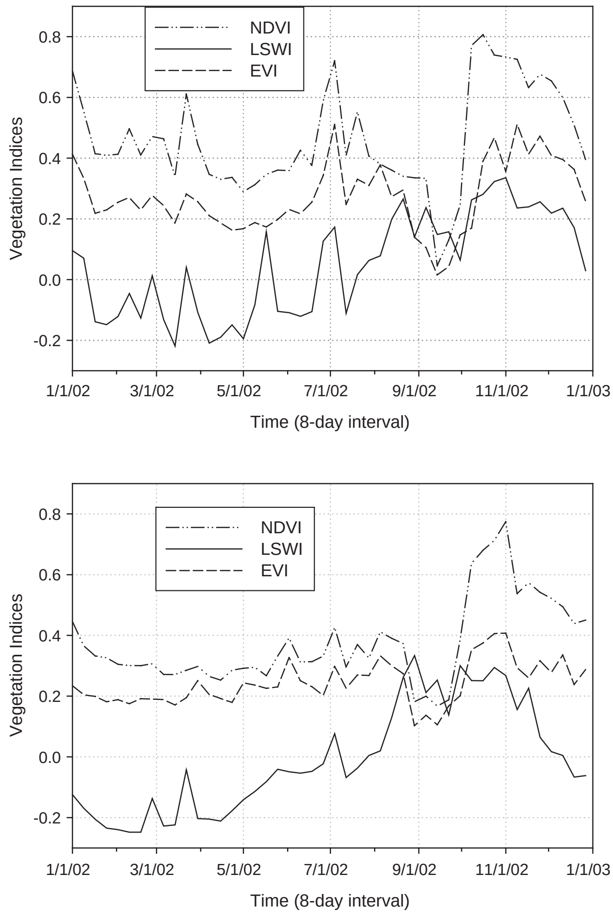


Fig. 13. The seasonal dynamics of three vegetation indices (EVI, LSWI, and NDVI) in 2002 for two MODIS pixels in the Tonle Sap Basin, Cambodia (top = 13.464°N, 103.098°E; bottom = 13.569°N, 103.628°E). Arrows define the approximate start of rice growth for each pixel.

vegetation indices and visual interpretation of Landsat ETM+ images suggest that this paddy rice algorithm does identify irrigated pixels in the basin. The large discrepancy in paddy rice area between the $MOD_{rice}$ and the NAS might be attributed to either under-reporting of the NAS data or other types of irrigated croplands.

## 6. Summary

This study represents our continuing efforts towards mapping individual crops by studying unique spectral features of individual crop systems. We have developed a database of paddy rice agriculture in monsoon South and Southeast Asia at 500-m spatial resolution, which is to our knowledge the finest-resolution database of paddy rice at such a large spatial domain. This is made possible by the availability of water-sensitive shortwave infrared bands from a new generation of optical sensors (MODIS and VGT) that enable us to progress beyond previous mapping algorithms that are primarily dependent on NDVI as the spectral input. There are certain sources of error that are inherent to optical sensors, such as cloud contamination, topographic effects, and resolution limitations (both spatial and temporal). However, in general, the output of the MODIS rice algorithm was similar to datasets derived from census statistics, both in terms of spatial distribution and area totals. Floods and drought events associated with the monsoon climate system can substantially affect the timing and spatial distribution of paddy rice agriculture in Asia. In the future, we intend to apply this algorithm to multi-year MODIS data to examine its potential for quantifying inter-annual variations of paddy rice fields due to extreme climate events and/or human-driven land-use changes. Other future efforts may include global application of this algorithm to provide an updated global dataset of paddy rice, and exploring the temporal profile analysis approach for its ability to map other crops (e.g., cotton) that have a period of significant irrigation at the start of the plant-growing season.

## Acknowledgement

We thank Prasad S. Thenkabail for providing information on individual field sites he has collected of paddy rice fields in the Ganges, Indus, and Krishna river basins in India. We thank three anonymous reviewers for their comments and suggestions on the earlier version of the manuscript. This study is supported by grants from the NASA Terrestrial Ecology Program (NAG5-12838), the NASA Earth Observing System Interdisciplinary Science Program (NAG5-10135), and the NASA Land Cover and Land Use Change Program (NAG5-11160).

## References

Aselman, I., & Crutzen, P. J. (1989). Global distribution of natural freshwater wetlands and rice paddies, their net primary productivity, seasonality and possible methane emissions. *Journal of Atmospheric Chemistry*, *8*, 307–358.

Boles, S., Xiao, X., Liu, J., Zhang, Q., Munkhtuya, S., Chen, S., et al. (2004). Land cover characterization of temperate East Asia using multi-temporal VEGETATION sensor data. *Remote Sensing of Environment*, *90*, 477–489.

Denier Van Der Gon, H. (2000). Changes in CH4 emission from rice fields from 1960s to 1990s: 1. Impacts of modern rice technology. *Global Biogeochemical Cycles*, *1*, 61–72.

Eva, H., Belward, A., De Miranda, E., Di Bella, C., Gond, V., Huber, O., et al. (2004). A land cover map of South America. *Global Change Biology*, *10*, 731–744.

Fang, H. (1998). Rice crop area estimation of an administrative division in China using remote sensing data. *International Journal of Remote Sensing*, *17*, 3411–3419.

Fang, H., Wu, B., Liu, H., & Xuan, H. (1998). Using NOAA AVHRR and Landsat TM to estimate rice area year-by-year. *International Journal of Remote Sensing*, *3*, 521–525.

FAOSTAT. (2001). Statistical Database of the Food and Agricultural Organization of the United Nations.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*, 287–302.

Frolking, S., Qiu, J., Boles, S., Xiao, X., Liu, J., Zhuang, Y., et al. (2002). Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. *Global Biogeochemical Cycles*, *16*(4). doi:10.1029/2001GB001425.

Frolking, S., Babu Y. J., (submitted for publication). New district-level maps of rice cropping in India: A foundation for scientific input into policy assessment. Agricultural Systems.

Hall, D. K., Riggs, G. A., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, *54*, 127–140.

Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., & Bayr, K. J. (2002). MODIS snow-cover products. *Remote Sensing of Environment*, *83*, 181–194.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Huete, A. R., Liu, H. Q., Batchily, K., & vanLeeuwen, W. (1997). A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, *59*, 440–451.

Huke, R. E., Huke, E. H. (1997). *Rice area by type of culture: South, Southeast, and East Asia, A revised and updated data base.* International Rice Research Institute, Los Banos, Laguna, Philippines.

Knox, J. W., Matthews, R. B., & Wassmann, R. (2000). Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia: III. Databases. *Nutrient Cycling in Agroecosystems*, *58*, 179–199.

Latifovic, R., Zhu, Z. L., Cihlar, J., Giri, C., & Olthof, I. (2004). Land cover mapping of North and Central America—Global Land Cover 2000. *Remote Sensing of Environment*, *89*, 116–127.

Leff, B., Ramankutty, N., & Foley, J. A. (2004). Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, *18*. doi:10.1029/2003GB002108.

Le Toan, T., Ribbes, F., Wang, L., Floury, N., Ding, K., Kong, J., et al. (1997). Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. *IEEE Transactions on Geoscience and Remote Sensing*, *1*, 41–56.

Li, C. S., Qui, J.J., Frolking, S., Xiao, X.M., Salas, W., Moore, B., et al. (2002). Reduced methane emissions from large-scale changes in water management of China's rice paddies during 1980–2000. *Geophysical Research Letters*, *29*, (art. no.-1972).

Li, C. S., Frolking, S., Xiao, X. M., Moore, B., Boles, S., Qiu, J. J., et al. (2005). Modeling impacts of farming management alternatives on CO2, CH4, and N2O emissions: A case study for water management of rice agriculture in China. *Global Biogeochemical Cycles*, *19*(3).

Loveland, T. R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., 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*, 1303–1330.

Maclean, J. L., Dawe, D. C., Hardy, B., & Hettel, G. P. (2002). *Rice almanac: Source book for the most important economic activity on earth* (3rd edn.). CABI Publishing.

Maki, M., Ishiahra, M., & Tamura, M. (2004). Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data. *Remote Sensing of Environment*, *90*, 441–450.

Matthews, E., Fung, I., & Lerner, J. (1991). Methane emission from rice cultivation: Geographic and seasonal distribution of cultivated areas and emissions. *Global Biogeochemical Cycles*, *5*, 3–24.

Matthews, R. B., Wassmann, R., Knox, J. W., & Buendia, L. V. (2000). Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia: IV. Upscaling to national levels. *Nutrient Cycling in Agroecosystems*, *58*, 201–217.

Neue, H., & Boonjawat, J. (1998). Methane emissions from rice fields. In J. Galloway, & J. Melillo (Eds.), *Asian change in the context of global climate change* (pp. 187–209). Cambridge: Cambridge University Press.

Okamoto, K., & Fukuhara, M. (1996). Estimation of paddy rice field area using the area ratio of categories in each pixel of Landsat TM. *International Journal of Remote Sensing*, *9*, 1735–1749.

Okamoto, K., & Kawashima, H. (1999). Estimating of rice-planted area in the tropical zone using a combination of optical and microwave satellite sensor data. *International Journal of Remote Sensing*, *5*, 1045–1048.

Olson, J. S. (1992). World Ecosystems (WE1.4): Digital Raster Data on a 10-minute geographic $1080 \times 2160$ grid. In J. J. Kineman & M. A. Ochrenschall (Eds.), In Global Ecosystems Database version 1.0: Disc A, Documentation manual. US. Department of Commerce/National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, Colorado.

Prather, M., & Ehhalt, D. (2001). Atmospheric chemistry and greenhouse gases. In J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, & C. A. Johnson (Eds.), *Climate change 2001: The scientific basis* (pp. 239–287). Cambridge, U.K.: Cambridge University Press.

Sass, R. L., Fisher, F. M., Ding, A., & Huang, Y. (1999). Exchange of methane from rice fields: National, regional, and global budgets. *Journal of Geophysical Research – Atmospheres*, *104*(D21), 26943–26951.

Tennakoon, S. B., Murty, V. V. N., & Etumnoh, A. (1992). Estimation of cropped area and grain yield of rice using remote sensing data. *International Journal of Remote Sensing*, *13*, 427–439.

Thenkabail, P., Schull, M., & Turral, H. (2005). Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment*, *95*, 317–341.

Van Niel, T. G., McVicar, T. R., Fang, H., & Liang, S. (2003). Calculating environmental moisture for per-field discrimination of rice crops. *International Journal of Remote Sensing*, *24*, 885–890.

Vermote, E. F., & Vermeulen, A. (1999). Atmospheric correction algorithm: Spectral reflectance (MOD09). *MODIS algorithm technical background document, version 4.0*. University of Maryland, Department of Geography.

Wassmann, R., Lantin, R. S., Neue, H. U., Buendia, L. V., Corton, T. M., & Lu, Y. (2000). Characterization of methane emissions from rice fields in Asia. III. Mitigation options and future research needs. *Nutrient Cycling in Agroecosystems*, *58*(1–3), 23–36.

Wilson, M. F. & Henderson-Sellers, A. (1992). A global archive of land cover and soils data for use in general circulation models. In: J. J. Kineman & M. A. Ochrenschall (Eds.), Global Ecosystems Database version 1.0: Disc A, Documentation Manual. US. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, Colorado.

Wright, G., Moffatt, D., & Wager, J. (2004). *TA4212-CAM establishment of the Tonle Sap Basin management organization: Tonle Sap Basin profile*. Asian Development Bank and Cambodia National Mekong Committee.

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., et al. (2002a). Landscape-scale characterization of cropland in China using Vegetation and Landsat TM images. *International Journal of Remote Sensing*, *23*, 3579–3594.

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., et al. (2002b). Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. *International Journal of Remote Sensing*, *23*, 3009–3022.

Xiao, X., He, L., Salas, W., Li, C., Moore, B., Zhao, R., et al. (2002c). Quantitative relationships between field-measured leaf area index and vegetation index derived from VEGETATION images for paddy rice fields. *International Journal of Remote Sensing*, *23*, 3595–3604.

Xiao, X. M., Boles, S., Liu, J. Y., Zhuang, D. F., Frolking, S., Li, C. S., et al. (2005). Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, *95*(4), 480–492.