# Mapping rice areas of South Asia using MODIS multitemporal data

Murali Krishna Gumma
Andrew Nelson
Prasad S. Thenkabail
Amrendra N. Singh

# Mapping rice areas of South Asia using MODIS multitemporal data

**Murali Krishna Gumma,**[a,][*] **Andrew Nelson,**[a] **Prasad S. Thenkabail,**[b] **and Amrendra N. Singh**[a]

[a] International Rice Research Institute, Social Sciences Division, DAPO Box 7777, Metro Manila, 7777 Philippines
m.gumma@cgiar.org; muraligk5@gmail.com

[b] Southwest Geographic Science Center, U.S. Geological Survey, Flagstaff, Arizona 86001

**Abstract.** Our goal is to map the rice areas of six South Asian countries using moderate-resolution imaging spectroradiometer (MODIS) time-series data for the time period 2000 to 2001. South Asia accounts for almost 40% of the world's harvested rice area and is also home to 74% of the population that lives on less than $2.00 a day. The population of the region is growing faster than its ability to produce rice. Thus, accurate and timely assessment of where and how rice is cultivated is important to craft food security and poverty alleviation strategies. We used a time series of eight-day, 500-m spatial resolution composite images from the MODIS sensor to produce rice maps and rice characteristics (e.g., intensity of cropping, cropping calendar) taking data for the years 2000 to 2001 and by adopting a suite of methods that include spectral matching techniques, decision trees, and ideal temporal profile data banks to rapidly identify and classify rice areas over large spatial extents. These methods are used in conjunction with ancillary spatial data sets (e.g., elevation, precipitation), national statistics, and maps, and a large volume of field-plot data. The resulting rice maps and statistics are compared against a subset of independent field-plot points and the best available subnational statistics on rice areas for the main crop growing season (kharif season). A fuzzy classification accuracy assessment for the 2000 to 2001 rice-map product, based on field-plot data, demonstrated accuracies from 67% to 100% for individual rice classes, with an overall accuracy of 80% for all classes. Most of the mixing was within rice classes. The derived physical rice area was highly correlated with the subnational statistics with $R^2$ values of 97% at the district level and 99% at the state level for 2000 to 2001. These results suggest that the methods, approaches, algorithms, and data sets we used are ideal for rapid, accurate, and large-scale mapping of paddy rice as well as for generating their statistics over large areas. © *2011 Society of Photo-Optical Instrumentation Engineers (SPIE)*. [DOI: 10.1117/1.3619838]

## 1 Introduction

South Asia (Fig. 1), which comprises Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka, has been described as the "food basket" and "food bowl" of Asia. Agriculture in the region provides employment and livelihoods for tens of millions of rural families directly or indirectly. The region accounts for almost 40% of the world's harvested rice area[1] and almost 25% of the world's population (circa 1.5 billion people).[1–4] Almost 74% of the population lives on less than $2.00 a day,[2] and rice provides around 30% of the calorific requirements of the
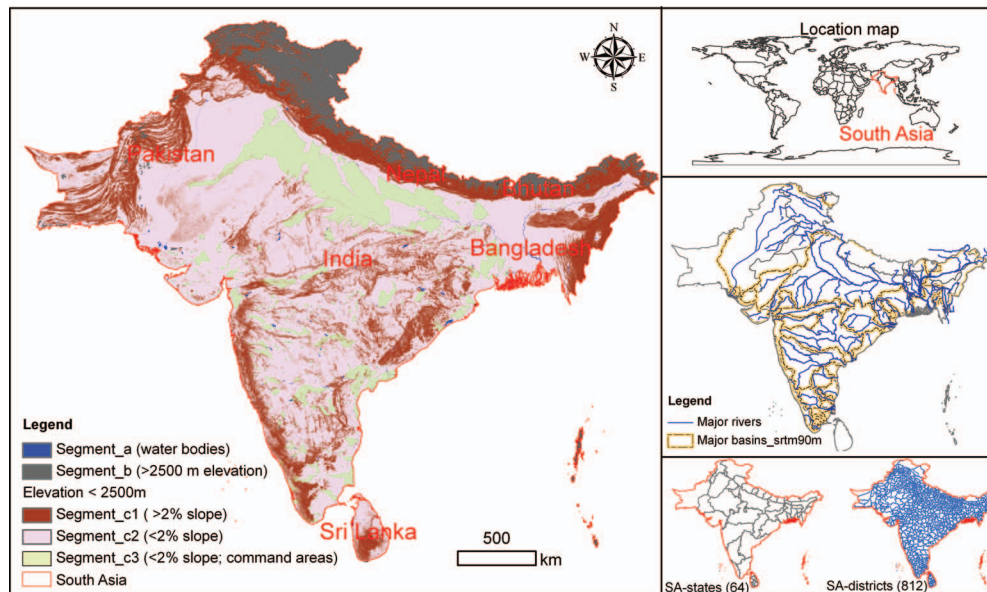
---

*Corresponding author.

**Fig. 1** Study area showing South Asia (Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka) and its segments. The South Asia study area showing classification segments, location map, major rivers with river basins, and administrative units. An overwhelming proportion of rice is grown in segment C3 (shown in green), indicating the high value of segmenting the image before classification and class identification, thus making the class identification process easier and accuracy higher.

population.[3] The population of the region is projected to grow faster than its ability to produce sufficient rice to meet the demand,[4] making South Asia food insecure in coming decades. Thus, rice-based agricultural systems are a vital component of any strategy to ensure food security and alleviate poverty in the region. Rice agriculture, especially on this vast spatial scale, also has important implications for natural resource management.[5] The cultivation of paddy rice on flooded soils requires large quantities of fresh water and thus has implications for water security and water quality,[6] while waterlogged soils are one of the largest sources of methane gas emissions.[7,8] Mapping and monitoring rice cultivation will provide important information to planners, decision makers, and scientists on where exactly rice is cultivated, its intensities, and changes over space and time. For example, substantial portions of rice croplands are lost to urbanization, and biofuel plantations in recent times. Exact mapping of rice crop areas will lead to more accurate assessments of their water use, crop productivity, and water productivity. Exact mapping of rice areas will also help assess methane emissions far more accurately. The interlinkage of rice-crop mapping to all these crucial agrological, ecological, and climatic factors will have a substantial influence on food security assessment and planning.

Compilations of regional and subnational statistical data on rice areas and coarse-scale rice maps in the 1980s and 1990s (Refs. 9–11) were followed in the 2000s by the automated mapping of rice areas using medium spatial resolution remote sensing[6,12–14] as well as more sophisticated combinations of various data sources.[5,15] Agricultural statistical data are invaluable for understanding historical trends in rice agriculture, but the data are rarely available in a timely manner and may not have sufficient spatial resolution, whereas automated remote-sensing techniques can rapidly generate up-to-date maps of paddy rice on a large scale, but they are difficult to validate on such a scale. Experiments by Sun et al.[14] in China have demonstrated that a single rice mapping algorithm with little field-plot and no crop calendar input may not be appropriate for the accurate identification of rice areas across different rice agroecologies. Rice is cultivated across a wide range of climatic conditions under many different crop-management

**Table 1** Country area and arable agricultural area in South Asia for 2000-01. Note that 41% of the geographic area in South Asia is arable.

| Country | Area ('000 ha) | Arable land, 2000 ('000 ha) |
|---------|---------------|------------------------------|
| Bangladesh | 14,804 | 8,084 |
| Bhutan | 4,365 | 130 |
| India | 345,623 | 162,717 |
| Nepal | 16,210 | 2,354 |
| Pakistan | 89,167 | 21,292 |
| Sri Lanka | 6,453 | 915 |
| Total | 476,622 | 195,492 |

Source: World rice statistics
http://beta.irri.org/index.php/

techniques, and some degree of prior knowledge of the variation in these factors is essential to produce an accurate map of rice-growing areas.

In this study, we present a method to identify and classify rice-growing areas in South Asia (Fig. 1, Table 1) using a suite of methods, including spectral matching techniques (SMTs), decision-tree algorithms, and an ideal temporal profile for rice classes. [Note: We deliberately use the term "temporal profile" throughout this paper for seasonal patterns of normalized difference vegetation index (NDVI). This is because we use 100's of bands of data in a single megafile data cube, which is akin to 100's of bands of hyperspectral data. The megafile data cube is described in Sec. 2.3.3, also see Refs. 16–19]. We use a time series of eight-day, 500-m spatial resolution, seven-band reflectance data (Table 2) composite images for 2000 to 2001 from the moderate-resolution imaging spectroradiometer (MODIS) sensor as well as ancillary spatial data sets and field-plot observations to identify and classify rice areas over a large spatial extent. We used data from 2000 to 2001 because this is the most recent "good" year for rice production in South Asia,[1] that is, there were no region-wide severe droughts or floods and no acute pest or disease outbreaks. This also matches the information collected from field observations, and detailed agricultural census data were available for this period for all six countries. As such, this represents a contemporary picture of rice cultivation in South Asia under good conditions. We demonstrate the accuracy of the approach in terms of classification accuracy for identifying different rice agroecologies and by comparing the MODIS-derived physical rice area against published national statistics.

**Table 2** MODIS data sets (seven bands): MODIS Terra seven-band, 500-m reflectance data characteristics used in this study. (Source: adapted from Ref. 16.)

| | MOD09A1 product | | | |
|---------|---------------------|----------------------|-----------------|----------------------------------------------------|
| MODIS bands | Band width (nm$^3$) | Band center (nm$^3$) | Visible range | Potential application |
| b3 | 459–479 | 470 | Blue | Soil/vegetation differences |
| b4 | 545–565 | 555 | Green | Green vegetation |
| b1 | 620–670 | 648 | Red | Absolute land cover transformation, vegetation chlorophyll |
| b2 | 841–876 | 858 | NIR1 | Cloud amount, vegetation land cover transformation |
| b5 | 1230–1250 | 1240 | NIR2 | Leaf/canopy differences |
| b6 | 1628–1652 | 1640 | SWIR1 | Snow/cloud differences |
| b7 | 2105–2155 | 2130 | SWIR2 | Cloud properties, land properties |

## 2 Materials

### 2.1 Study Area

South Asia extends between $5°38'39''$ and $36°54'38''$ latitude and $61°05'04''$ and $97°14'12''$ longitude, with a land mass of 477 million hectares (Fig. 1, Table 1). The region lies within six agroecological zones: humid tropics, subhumid tropics, semiarid tropics, semiarid, subtropics, and arid.[20] Rice is grown within all these zones but is concentrated in the semiarid tropics and subhumid tropics. South Asia contains almost 200 million hectares of arable land area. Water supply is either from irrigation (∼80 million hectares, where water is supplied via canals, water tanks, or groundwater) or rainfall (∼120 million hectares).[21] Figure 1 shows the major drainage basins in South Asia and the command areas where irrigation is present in India.

Rice cropping occurs in three seasons across the region: the main kharif or aman during the monsoon (June to mid-December), the rabi or boro in the postmonsoon dry season (mid-December to March), and a summer-season aus crop (April and May). Rice is harvested on some 60 million hectares across all three seasons (Table 1). Irrigated areas include double cropping of rice and other grains, single cropping of sugarcane, chilli, cotton, fodder grass, and some areas of light irrigation of corn, sorghum, and sunflower. Rain-fed crops include grains (sorghum, millet), pulses (red gram, green gram, and chickpea), and oilseeds (sunflower and groundnut).

### 2.1.1 MODIS surface reflectance data

The MODIS eight-day composite surface reflectance product from the Terra platform (MOD09A1) is ideal for monitoring vegetation at a continental scale.[16] The seven bands of reflectance data (Table 2) at a resolution of 15 arc s (500 m), coupled with a high-repeat frequency, can capture the seasonal variations in vegetation vigor, soil, and vegetation moisture, and surface water that characterize the key stages of rice cultivation.[16] The reflectance data undergo several preprocessing steps, including algorithms for atmospheric correction. Furthermore, the rate of observation coverage, the viewing angle, cloud or cloud shadow coverage, and aerosol loading are all assessed on a pixel-by-pixel basis to ensure that each pixel contains the best observations during that eight-day period. MOD09A1 also includes two quality assessment data sets at the pixel and band level, which are vital for user postprocessing to identify and remove areas of persistent cloud and snow cover. The MOD09A1 (Version 005) data are available in a tile system, in which each tile covers $10 \times 10$ deg ($1111.2 \times 1111.2$ km at the equator). We downloaded 12 tiles, for every eight days, from Ref. 22 covering South Asia for all dates between June 2000 and May 2001, which includes the rice crops in three seasons (kharif, rabi, and aus) for every 12-month period.

### 2.2 MODIS Data Preprocessing

### 2.2.1 Blue-band minimum-reflectivity threshold for clouds

South Asia is subject to the influences of the oscillating subtropical convergence zone, which include monsoonal activity from June to September (kharif season). It is during this part of the year that there is a significant change in vegetation cover, rapid changes in the dynamics of vegetation, and changes in biomass accumulation. It is also a period when cloud cover is more frequent. To retain the highest possible number of pixels in the time series, the following approach was adopted:[16] (a) retain all images with <5% cloud cover and (b) apply a cloud-masking algorithm in order to eliminate areas of cloud cover and retain the rest of the image in an unchanged form.[16,22] The cloud-masking algorithm applies a threshold to band 3, a simple algorithm for band 3 cloud removal in ERDAS-ER Mapper (ERDAS, 2010) was: for band 3 (if $b_3 \geq 18\%$, then null else $b_3$) (see Table 2) in the MODIS imagery, where, if the reflectance value is >18%, then this is considered as null (this cutoff value was arrived at by selecting

several samples over cloud patches throughout the study area). For other bands, where band 3 is more than 18%, then the values in all bands are replaced with a null value. For a more detailed description of cloud-removal algorithms for MODIS, refer to Ref. 16.

### 2.2.2 Normalized-difference vegetation index and monthly maximum value composite

An NDVI was generated using surface reflectance values of red and NIR bands in Eq. (1). Monthly, three to four eight-day composites were available, for a total of 45 eight-day composites. Monthly maximum value composites (MVCs) for June 2000 through May 2001 were created using the eight-day images in order to minimize cloud effects during the monsoon season in Eq. (2). The monthly MVCs were stacked into a 12-band megafile data cube (MFDC).

$$NDVI = \frac{\lambda_{NIR} - \lambda_{red}}{\lambda_{NIR} + \lambda_{red}}, \tag{1}$$

$$NDVI_{MVC_i} = Max(NDVI_{i1}, NDVI_{i2}, NDVI_{i3}, NDVI_{i4}), \tag{2}$$

where $MVC_i$ is the monthly maximum value composite of the $i$'th month and $i_1$, $i_2$, $i_3$, and $i_4$ are every eight days' data in a month.

### 2.2.3 Megafile data cube data-set generation

A megafile data cube[17] was composed for June 2000 through May 2001. The MFDC consisted of 327 bands: 12 bands of MVC NDVI images (one MVC image per month) and a further 315 bands, composed from all seven bands in the eight-day composites. This MFDC was then used to generate the ideal and class temporal profile (NDVI pattern) as follows.

### 2.3 Field-Plot Data for Identifying, Labeling, and Assessing the Accuracy of MODIS-Based Rice Maps

Field-plot data were collected during October 11–26, 2003, and August 30–September 28, 2005. Interviews with farmers and agricultural extension officers were conducted to determine cropping types and conditions, including the collection of historical data as far back as 2000 to 2001. A total of 1004 locations covering the major crop-land areas (Fig. 2) were chosen based on the knowledge of local agricultural extension officers to ensure that adequate samples of rice as well as other crops were gathered. The local experts also provided information on crop calendars, cropping intensity (single or double crop), and percentage canopy cover for these locations. Of the 1004 data points, 75% (751) were used for class identification and calculating rice fractions (within this, 149 rice points were used for ideal temporal profile generation) and the rest of the points (253) were used for the following accuracy assessment.

The precise locations of the sample sites were recorded using a Garmin handheld global positioning system unit. Ideally, 50 or more samples should be collected per land use class,[23] but, due to limited resources, the sample size varied from 15 to 25 for each major crop-land use/land cover (LULC) class.

At each of the 1004 locations, the following data were recorded for 2000 to 2001 based on interviews with local agricultural extension officers and/or farmers:

1. Land-use parameters, including farm land, range lands, and open lands, which relate to levels I, II, and III in the Thenkabail approach for land use/land cover classifications.[17,18] As an example, level I includes detailed information such as scale (large-/small-scale agriculture) and intensity (single crop, double crop, and continuous crop) with crop
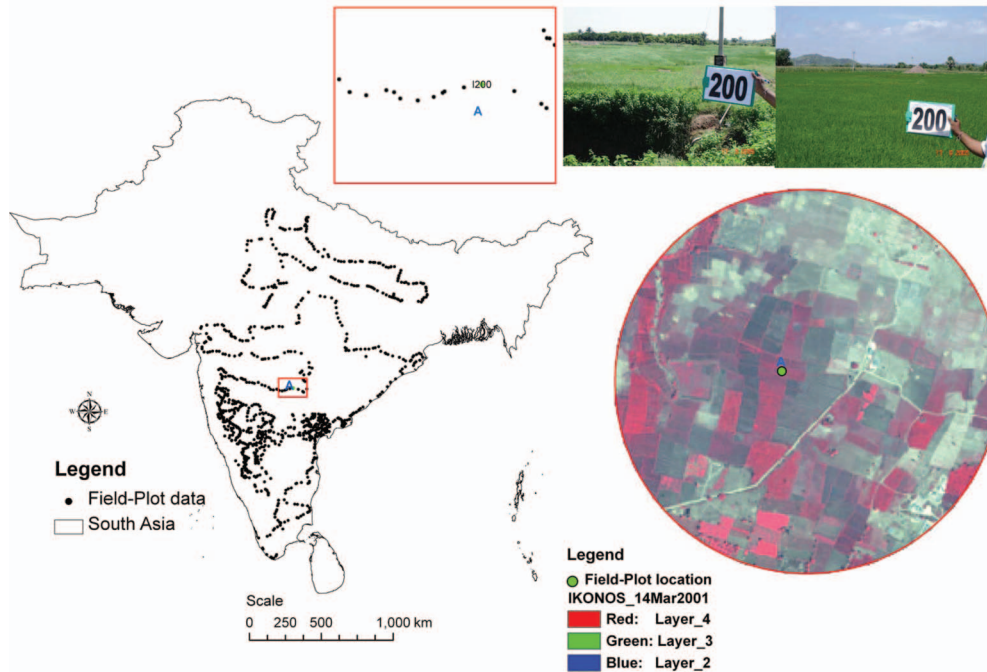
**Fig. 2** Field-plot data-point locations in South Asia. There are 1004 field-plot locations where crop type, cropping intensity, watering source (irrigated versus rainfed), and a number of other parameters (e.g., digital photos, land cover distribution) were also collected.

      type, irrigation type (groundwater, surface water, and conjunctive), and watering method (irrigated, rainfed, and other LULC)

2. Land-cover types (percent cover): trees, shrubs, grasses, built-up area, water, fallow lands, weeds, different crops, sand, rock, and fallow farms
3. Crops cultivated in the kharif, rabi, and summer seasons
4. Cropping calendars for kharif, rabi, and summer seasons
5. Irrigation sources (groundwater, surface water, and tank) at each location.

These class labels were assigned in the field, but additional class information was incorporated later on the basis of visual interpretation of the LandSat GEOCOVER 2000 data set.[24] Most of the Google Earth images were very high resolution (submeter to 4 m; from IKONOS, Quickbird satellites, and Landsat images) and were acquired in recent years. These data again relate to the target mapping for the 2000 to 2001 season.

## 2.4 Ancillary Spatial and Statistical Data

Several ancillary (or secondary) spatial data sets were used to segment the MODIS data into various components in order to augment the classification procedure and increase the accuracy in identifying rice classes.

### 2.4.1 Elevation and slope constraints for rice cultivation

Rice is rarely grown at elevations above 2500 m; thus, the digital elevation model (DEM) was used to remove these areas from the analysis. Slope is also a limiting factor, and although rice can be cultivated on steep slopes by means of terracing, this is not a common system in South Asia (except in Nepal and Bhutan). A series of thresholds were used to classify the terrain into

flood plains and lower valley slopes with ≤2% gradient as one segment and >2% gradient as another segment.

The Shuttle Radar Topography Mission (SRTM) obtained elevation data on a near-global scale to generate the most complete high-resolution digital topographic database of the Earth at a resolution of 3 arc s.[25–28] The SRTM tiles were downloaded from Ref. 29 and mosaicked into a regional DEM. A slope map was computed from this DEM, and both elevation and slope were used as constraint maps to determine where rice could possibly be cultivated.

### 2.4.2 Irrigation command areas to segment them from the rest of the areas

Irrigated command areas are a clear indication of where surface (canal)-irrigated crops are cultivated. We obtained irrigated command area maps from India and used them to further split the ≤2% slope gradient segment into irrigated and nonirrigated areas (Fig. 3). Crop lands within these command areas are almost always irrigated (with the possible exception of more remote areas during drought periods). Naturally, there are always some LULC classes within the command area segment that are not crop land, but the command area segment will still aid classification accuracy with respect to irrigated rice areas.

### 2.4.3 Water body segment derived from Landsat Geo cover and elevation data

One straightforward land-cover class is permanent surface-water bodies, and it is desirable to identify these prior to further analysis and classification. Water bodies were extracted from Landsat Geocover data for the year 2000 and SRTM DEM. Water bodies accounted for 1,879,587 hectares in South Asia.

### 2.4.4 Very high-resolution imagery via Google Earth® for class identification, labeling, and accuracy assessment

Google Earth provides very high-resolution images down to submeter resolution with free access, which is valuable for the visual interpretation of land cover, especially to ascertain whether a class is irrigated or rainfed crop land. Google Earth data[30] were also used to identify the presence of any irrigation structures (e.g., canals, irrigation channels, open wells). Most of the very high-resolution imagery (VHRI: <5 m; e.g., IKONOS, Quickbird, Geoeye) in South Asia that is available in Google Earth was acquired between 2000 and 2010. The rice crop has unique features, such as small bunds, along with very well-defined irrigation structures (e.g., canals, irrigation channels, open wells). These are clearly visible in Google Earth VHRI. This study was conducted for 2000 to 2001, and Google Earth VHRI was available mostly between 2003 and 2010. However, the irrigation structures and other features (e.g., rice bunds) do not change over time and space (except in highly insignificant very small fragments, which hardly exceed 1% over such a large geographic area in such short time spans) in South Asia. Hence, it is completely valid to use VHRI in identifying and labeling irrigated rice versus rain-fed rice. Furthermore, it should be noted that the Google Earth VHRI was used along with other distinct data sets described in various subsections of Sec. 2 and provides supplemental supporting information. A combination of these measures was used in the protocol of class identification and labeling. Using the same logic, VHRI was used in accuracy assessments[17] using visually derived analyst information from a large number of spatially well-distributed points, which is akin to ground-truth data.

### 2.4.5 Rice area statistics from national sources for accuracy assessments

A comparison data source (ideally, complete ground-truth information) is required in order to assess the skill of the rice-mapping approach. The 253 points that were retained from the ground-truth survey can be used to assess how well the classification performed for those areas
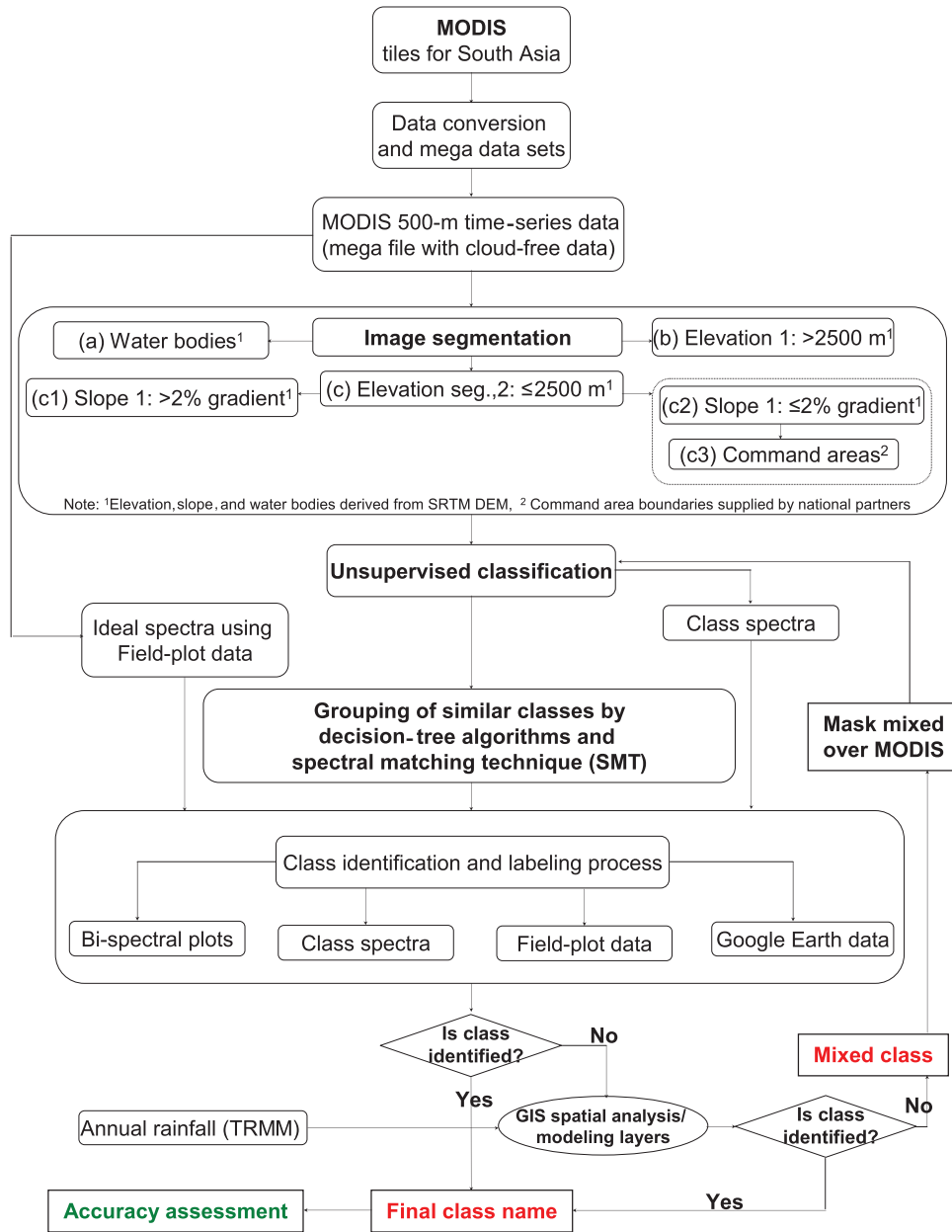
**Fig. 3** Methodology overview. Overview of the methodology for mapping rice areas using the eight-day MODIS MOD09A1 data and other ancillary spatial data. There were six image segments of a MFDC.

but does not indicate how well the classification performed in other areas. This is common to all remote-sensing approaches that rely on limited point survey data sources. An alternative or companion approach is to compare the rice area estimates from the map to area estimates from another source, such as agricultural census statistics. This is a very different approach, which—unlike the ground-truth points—is comprehensive in spatial coverage but restricted to the spatial detail in the census data. It also says nothing about the spatial distribution of rice within the census reporting units or about rice classification accuracy.

In this case, the 2000 to 2001 kharif-season rice area statistics were obtained at the national and subnational levels (analogous to a state, county, or district) from the six countries for comparison against the MODIS-derived rice areas.

Data for India were obtained from the website of the Ministry of Agriculture's Directorate of Rice Development.[31] Data for Bangladesh, Nepal, Pakistan, and Bhutan were obtained from the national statistical departments. In total, rice area statistics were collected for 812 spatial units across South Asia. The veracity of agricultural census data is often questioned, but, without access to region-wide high-resolution rice maps (should they even exist), there is no other realistic alternative source of independent information.

### 2.4.6 TRMM rainfall data

Monthly rainfall data for 2000 were downloaded from the Tropical Rainfall Measuring Mission (TRMM) with 0.25-deg spatial resolution to separate mixed irrigated and rain-fed classes based on rainfall.[32] The main purpose of the rainfall data was to separate tank-irrigated rice areas from rain-fed rice areas; however, some high-rainfall zones had a rice signature similar to that of irrigated rice areas (see Sec. 4.5).

## 3 Methods

### 3.1 Overview of Methods

An overview of the methods is shown in Fig. 3, and each step is described in detail in this section.

The process starts with the conversion of the MODIS data into GeoTIFF format, the removal of clouds as documented in Ref. 16, and the generation of MFDCs (see Ref. 17 for a detailed description), in which the multispectral time-series data are combined into a single-file data cube involving numerous data layers (see also Sec. 4.1.3). The MODIS MFDC is divided into distinct segments (see Secs. 3.3.1–3.3.3) to allow for easier class spectrum separation and class identification.

### 3.2 Ideal Temporal Profile (Normalized-Difference Vegetation Index Curves) Creation

The rice and nonrice fields were located exactly on the South Asia image (Fig. 4). The 149 rice points represent distinct categories of rice (e.g., double crop, single crop). The sample sizes of each category are indicated within the brackets of class legend in Fig. 5. The temporal profile signatures (e.g., NDVI, reflectivity) gathered for the 149 rice fields were grouped according to their unique categories and a typical ideal temporal profile (Fig. 5) of that particular category of rice crop established. When there were classes within classes (e.g., when an irrigated-rice-double crop has distinctly different signatures because of, say, cropping calendars), then we have two ideal temporal profiles for this class. This resulted in an ideal temporal profile of 12 distinct rice classes (Fig. 5) aggregated from the 149 sample locations.

### 3.3 Class Temporal Profile (Normalized-Difference Vegetation Index Curves) Generation

The MFDC was divided into five distinct segments (see Fig. 3) based on elevation, slope, water bodies, and command areas. The first set of segments was (a) water-body segment, (b) elevation segment 1: >2500 m, and (c) elevation segment 2: ≤2500 m. Segment c was further divided into three subsegments, which were c1 slope segment 1: >2% gradient, c2 slope segment 2: ≤2% gradient, and c3 irrigation command area. The main aim of the segmentation process was
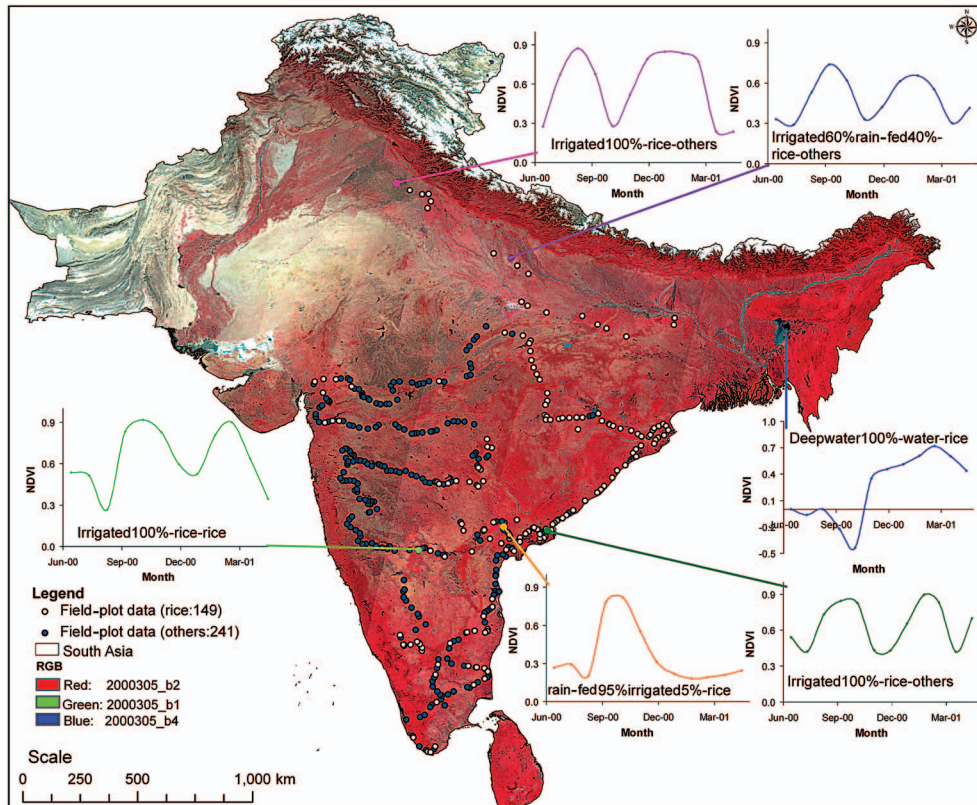
**Fig. 4** Phenology of the rice crop. Typical phenologies of six types of rice crops (e.g., double crop, single crop) of South Asia illustrated using NDVI signatures of MODIS time series. Top image is a false color-composite image of South Asia based on MODIS bands 1, 2, and 4 from October 31, 2000.

to focus more on the segments having higher amounts of crop-land classes such as the Central Board of Irrigation and Power (CBIP) command areas (e.g., India), valley bottoms, deltas, and river banks.

Unsupervised classification using the ISOCLASS cluster classification[17,18] (ISODATA in ERDAS Imagine 9.2™)[33] was applied to each of the MFDC segments. The classification was set at a maximum of 100 iterations and a convergence threshold of 0.99. In this study, the number of classes varied from 20 to 100 based on area spreading and complexity of segment. Segment 1 (elevation >2500 m) was classified into 20 classes because of lower complexity, and most of the areas are forested, snow-/ice-covered, or barren. Slope segments 1 and 2 were classified into 50 classes each, whereas the irrigation command area segment, which covered a larger and more complex area (with a mix of irrigated crops, irrigated dry crops, and rain-fed crops), contained 100 initial classes, making a total of 220 classes across all segments.

Class temporal profiles were generated using ISOCLASS *k*-means classification of each of the segment MFDCs. Each segment was classified into 100 classes. The signature file of the 100 classes was plotted, and similar classes were then grouped [e.g., Fig. 6(b) illustrates a few of the class temporal profiles]. The process of matching class temporal profile [e.g., Fig. 6(b)] with ideal temporal profile [e.g., Fig. 6(a)] is illustrated in Figs. 6(c) and 6(d). The process is repeated for other groups of classes, until all classes are resolved. The class temporal profiles that do not have a match with any ideal temporal profile are resolved using other approaches (e.g., Secs. 3.4 and 3.5). The quantitative approaches of SMTs are described in detail by Ref. 19. The process is repeated for classes from all segments.
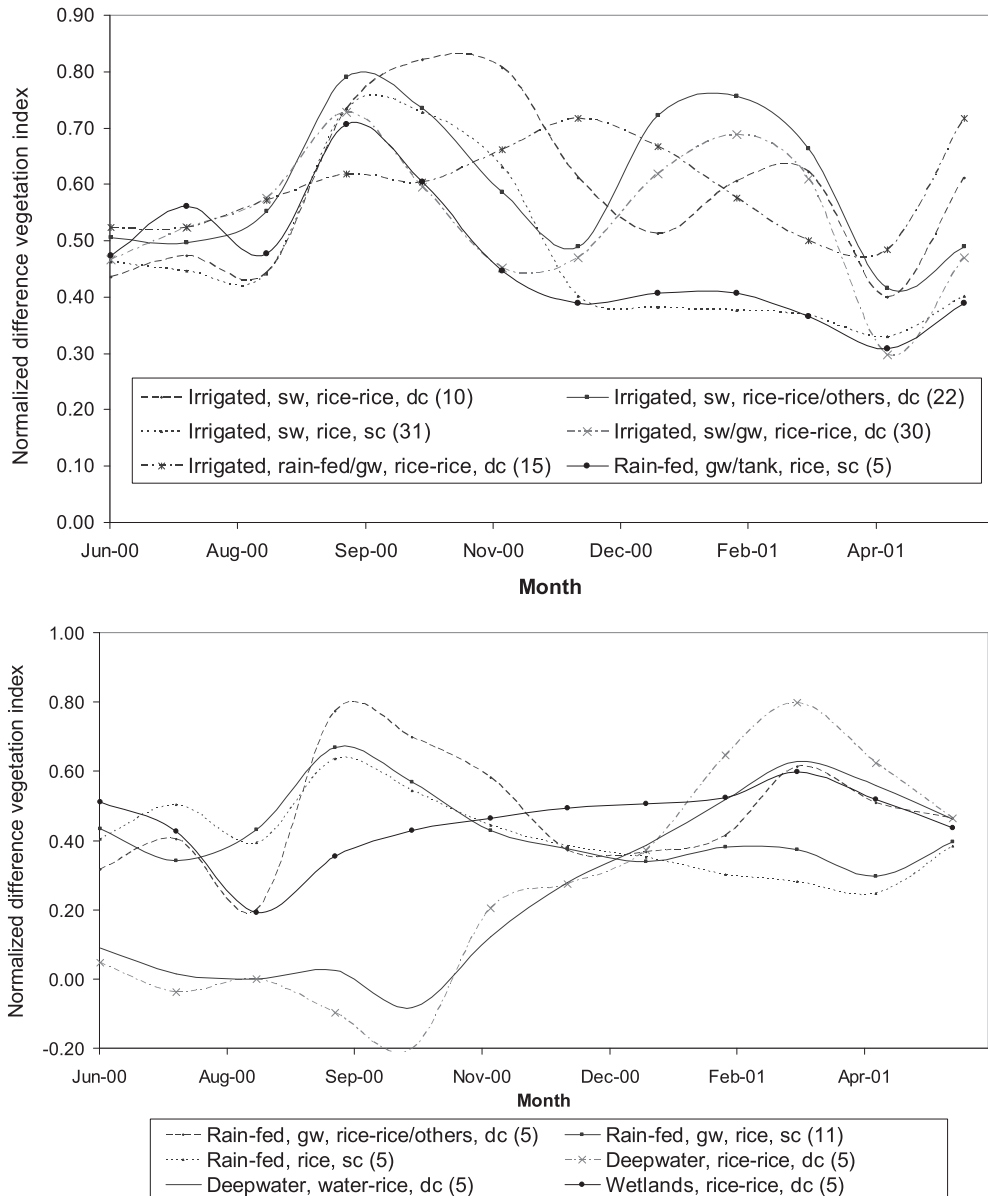
**Fig. 5** Ideal temporal signatures (NDVI curves) for the 12 rice classes. The ideal temporal profiles of 12 rice classes were generated using MODIS time-series data. Note: Field-plot sample size is shown in brackets ("sw" is surface water; "gw" is groundwater, "sc" is single crop, and "dc" is double crop).

## 3.4 Decision-Tree Algorithms

Several techniques were used simultaneously to group the classes in each segment. The first reduction in classes used a decision tree[34] based on the temporal NDVI MVC data. The decision tree is based on NDVI thresholds at different stages in the season that define the vegetation growth cycle, and these algorithms help to identify similar classes. The dates and threshold values were derived from the ideal temporal profile. The decision tree for segment c3 is shown in Fig. 7. This first stage reduced the total number of classes from 220 to 127.
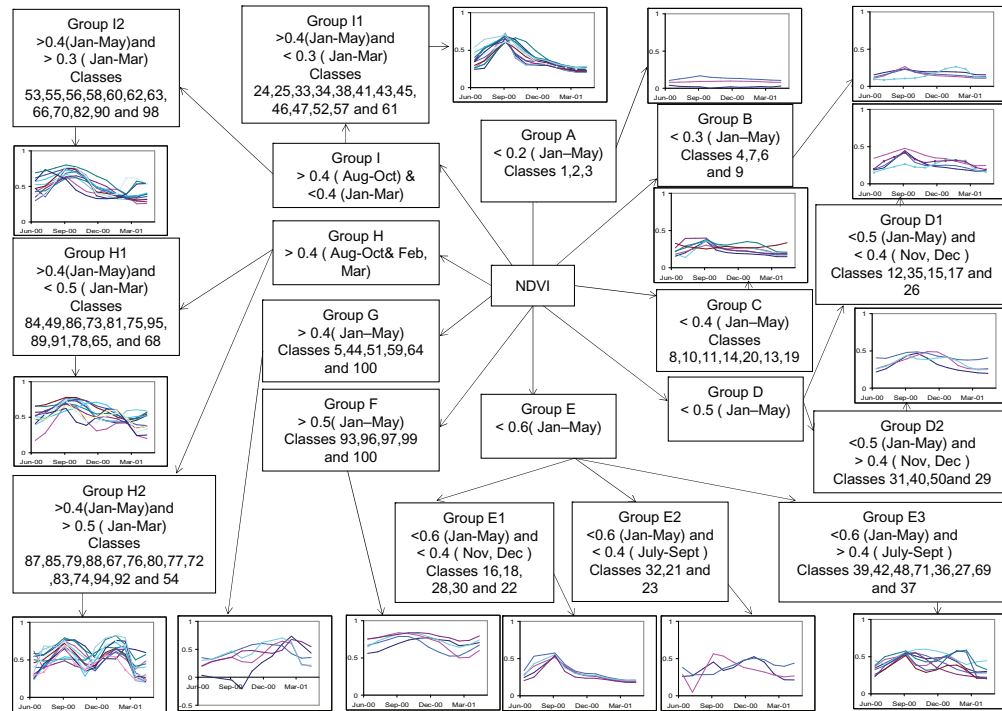
**Fig. 6** Decision-tree algorithm. The classes were also grouped and identified using the NDVI-based decision-tree algorithms. Note: Illustrated here for segment c3 classes. Each plot is NDVI (y) versus time (x).

The second and subsequent groupings used spectral matching techniques (SMT),[17,18] such as those used for hyperspectral analysis of minerals (e.g., Refs. [19], [34], and [35]). Time-series data, such as the eight-day MODIS NDVI data, are similar to such hyperspectral data with the 327 bands stacking a single instance of a hyperspectral image (or monthly NDVI values). These similarities imply that the SMTs applied for hyperspectral image analysis also have potential for application in identifying agricultural land-use classes from multiband time-series satellite imagery.

The temporal profile signatures of the classes within each segment were further grouped based on temporal profile similarity and compared against the ideal temporal profile in order to identify and label the grouped classes (Fig. [6]). Additionally, we used the field-plot data points, imagery from Google Earth and GeoCover, as well as national statistical data as auxiliary data in the class labeling. On the basis of one or more of these data, we clearly established that irrigated crop lands have distinct phenological cycles compared to rain-fed crop lands. For example, irrigated crops are often planted later than rain-fed crop lands, have a longer phenological cycle, and often have a following crop. All these factors are monitored using the MODIS time series. However, it must be noted that the magnitude of NDVI is not an indicator in distinguishing between rain-fed and irrigated crops, because we found that, in some cases, rain-fed crops actually have a higher magnitude of NDVI than irrigated crops. This may happen, for example, in flood-irrigated crops, for which background water can substantially absorb near-infrared reflectivity, thus reducing NDVI. Two main properties of class signatures allow the separation of groundwater and surface-water irrigation: annual average NDVI, which is a function of the irrigated fraction, and timing of the onset of greenness, which is a function of the timing of water availability for vegetation. Annual NDVI in both continuous and double-irrigated systems exceeds annual NDVI in groundwater systems, reflecting the higher
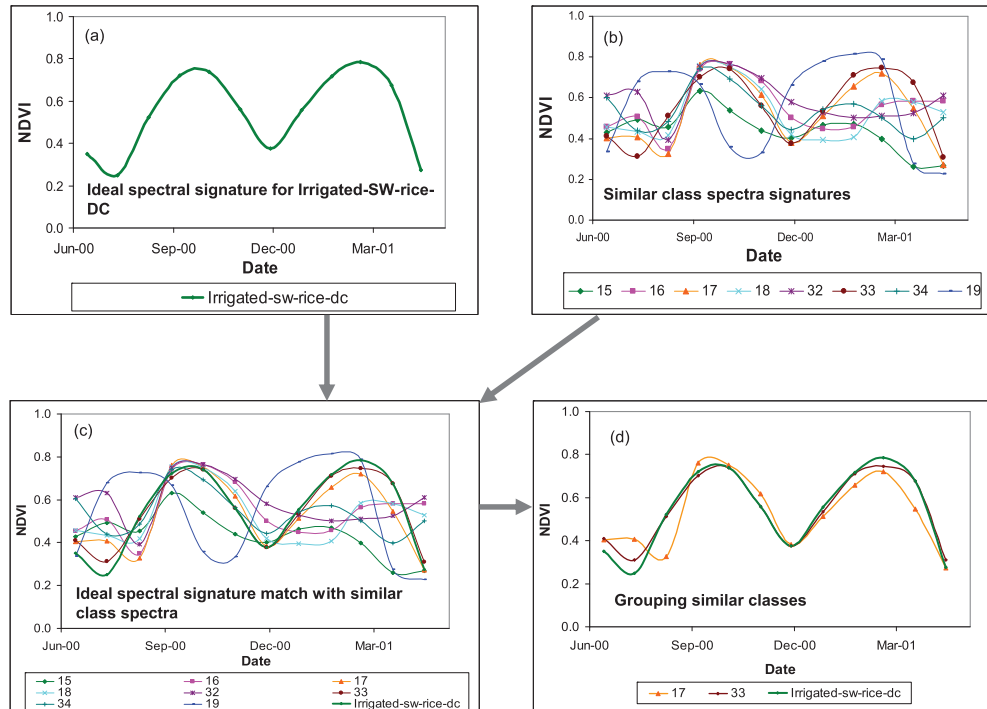
**Fig. 7** Spectral matching technique (SMT). In SMTs, the class temporal profiles (NDVI curves) are matched with ideal temporal profiles (quantitatively based on temporal profile similarity values) in order to group and identify classes as illustrated for a rice class in this figure. (a) Ideal temporal profile illustrated for "irrigated-surface-water-rice-double crop"; (b) some of the class temporal profile signatures that are similar, (c) ideal temporal profile signature [Fig. 7(a)] matched with class temporal profiles [Fig. 7(b)], and (d) the ideal temporal profile [Fig. 7(a), in deep green] matches with class temporal profiles of classes 17 and 33 perfectly. Then one can label classes 17 and 33 to be the same as the ideal temporal profile ("irrigated-surface-water-rice-double crop"). This is a qualitative illustration of SMTs. For quantitative methods, see Ref. 29.

irrigated fraction in areas irrigated with surface water. An example of spectral matching is shown in Fig. 6 for the irrigated-surface-water-rice-double crop.

## 3.5 Resolving Mixed Classes

A rigorous class-identification and labeling process (Sec. 3.4) helped in identifying, grouping, and labeling many classes. However, some complex classes remained unresolved as mixed, because the purity of these classes could not be adequately validated using field-plot data and/or very high-resolution imagery, and/or other means. The mixed classes have more than one class in them. When classes continued to be mixed, in spite of the various methods and techniques discussed in previous sections, we adopted GIS spatial modeling approaches to resolve the classes. This involved taking mixed classes and applying spatial modeling techniques, such as overlay, matrix, and recode,[33] based on the theory of map algebra and Boolean logic.[36–38] The spatial data layers used included elevation zones and rainfall zones. Any one or a combination of these data layers usually helped in separating the mixed classes. For example, when some other agricultural areas mixed with rice classes, we ran a matrix with elevation classes as one axis of the matrix and a mixed class as the other axis to obtain classes at different elevation gradients. In such a scenario, we separate the mixed class into various elevation zones and rice may be one of these classes. Similarly, we can use elevation with precipitation as one of the axes of the matrix. All these different possibilities were attempted until the mixed class got resolved.

Furthermore, a rule-based decision-tree algorithm was used to help split the mixed classes. One sample application used NDVI variations in specific months in different parts of the same class to separate the classes. Whenever spatial modeling and decision-tree algorithms were unsuccessful, the mixed classes were masked out and the MFDC covering this area was used to reclassify the image into a number of classes. This was followed by a repetition of the entire class-identification and labeling process.

In spite of the rigorous class-identification process described in these sections, there were still some mixed classes. Typically, the unresolved classes were then split up into 5–10 or more subclasses (depending on the extent of area and complexity), reclassified using the masked area of the megafile data cube and the class-identification and labeling process (as described previously), and repeated.

### 3.6 Actual or Subpixel Area Calculations

The composite MODIS pixels cover an area of 21.5 hectares, which is larger than many agricultural fields in the study area. Thus, many pixels contain more than one land-cover class (Table 3). Any rice classification derived from such data will provide only the full pixel area (FPA), whereas the actual rice area per pixel can be obtained only by computing subpixel areas (SPAs) [Eq. (3)],[17,19,39] which are defined as

$$SPA_n = FPA_n \times RAF_n \qquad (3)$$

where $SPA_n$ is the subpixel area of class $n$, $FPA_n$ is the FPA of class $n$, and $RAF_n$ is the rice area fraction of class $n$ as derived from the field-plot observation data. The RAFs of each class were computed based on a large (30–60) sample size of points that are spatially well distributed in a class. The RAFs were derived by a combination of ground observations and very high-resolution ($\leq 5$ m) data. The SPA of each class (last column in Table 3) is computed by multiplying the FPA of that class with the RAF of the class. Later, the SPAs of all classes are summed to obtain the actual rice areas from all the classes (Table 3).

## 4 Accuracy Assessments

### 4.1 Classification Assessment Based on Field-Plot Data

A fuzzy accuracy assessment (Table 4) was performed using 25% (251 data points out of 1004) of the field-plot data to derive a robust understanding of the accuracies of the data sets used in this study. The field-plot data were based on an extensive field campaign conducted throughout India during kharif and rabi seasons by International Water Management Institute researchers, and they consisted of 1004 points. About 30% of the points were collected during the rabi season.

Fuzzy accuracy assessment provides realistic class accuracies for which land cover is heterogeneous and pixel sizes exceed the size of uniform land cover units (see Refs. 16, 17, 39, and 40). For this study, we had assigned 3 × 3 cells of MODIS pixels around each of the field-plot points to one of six categories: absolutely correct (100% correct), largely correct (75% or more correct), correct (50% or more correct), incorrect (50% or more incorrect), mostly incorrect (75% or more incorrect), and absolutely incorrect (100% incorrect). Class areas were tabulated for a 3 × 3-pixel (nine-pixel) window around each field-plot point.

On the basis of the theoretical description given by Ref. 23, Eqs. (4)–(6) were used to estimate accuracies and errors. Field-plot information was used to determine robust accuracies, using Eqs. (4)–(6),

$$\text{Accuracy of rice area classes} = A_{\text{ra}} = 100 \times \frac{\text{RFPCRA}}{\text{TRFP}}, \qquad (4)$$

**Table 3** Rice classes (irrigated and rain-fed) mapped in this study. The table shows FPA, rice-area fraction (RAF), and SPA or actual area. SPA = FPA × RAF. Note: 12-class map in Fig. 8, signatures in Fig. 9.

| Class code | Rice classification (water application in percent) | Sample size | Full pixel area | Rice fraction | Net areas ('000 ha) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Trees | Shrubs | Grass | Built-up area | Water | Weeds | Other crops | Rice[a] |
| RCL01 | 01. Irrigated 100%—Rice/Rice | 36 | 4,109 | 90.5 | 57 | 23 | 49 | 17 | 11 | 1 | 232 | 3,718 |
| RCL02 | 02. Irrigated 100%—Rice/Rice or Rice/Other | 27 | 8,999 | 90.3 | 96 | 27 | 113 | 38 | 245 | 10 | 344 | 8,127 |
| RCL03 | 03. Irrigated 100%—Rice | 40 | 11,234 | 92.0 | 81 | 32 | 80 | 8 | 147 | 6 | 543 | 10,338 |
| RCL04 | 04. Irrigated ≥60% / Rainfed ≤40%—Rice/Rice or Other | 51 | 4,061 | 87.6 | 63 | 24 | 61 | 2 | 102 | 8 | 244 | 3,559 |
| RCL05 | 05. Irrigated ≤30% / Rainfed ≥70%—Rice/Rice or Rice/Other | 18 | 7,264 | 88.4 | 58 | 35 | 77 | 24 | 34 | 0 | 617 | 6,419 |
| RCL06 | 06. Upland ≥80% / Rainfed ≤10% / Irrigated ≤10%—Rice | 6 | 5,326 | 91.0 | 47 | 4 | 65 | 4 | 24 | 27 | 312 | 4,844 |
| RCL07 | 07. Rainfed ≥60% / Irrigated ≤40%—Rice/Rice | 2 | 2,015 | 93.3 | 31 | 3 | 50 | 1 | 0 | 0 | 50 | 1,879 |
| RCL08 | 08. Rainfed ≥90% / Irrigated ≤10%—Rice | 11 | 5,774 | 87.6 | 32 | 12 | 152 | 1 | 165 | 0 | 354 | 5,057 |
| RCL09 | 09. Rainfed 100%—Rice | 9 | 6,655 | 87.1 | 136 | 19 | 192 | 1 | 6 | 0 | 505 | 5,797 |
| RCL10 | 10. Deepwater 100%—Rice/Rice | 10 | 599 | 100.0 | 0 | 0 | 0 | 0 | 599 | 0 | 0 | 599 |
| RCL11 | 11. Deepwater 100%—Water/Rice | 10 | 450 | 100.0 | 0 | 0 | 0 | 0 | 450 | 0 | 0 | 450 |
| RCL12 | 12. Wetlands 100%—Rice/Rice | 10 | 337 | 100.0 | 0 | 0 | 0 | 0 | 337 | 0 | 0 | 337 |

[a]The last column is actual (subpixel) area obtained by multiplying FPA × RAF.

**Table 4** Fuzzy accuracy assessment from field-plot data. Values in the table indicate the percent of field-plot windows in each class with a given correctness percentage.

| Rice class number and class name | Sample size | Absolutely correct (100%) | Mostly correct (76–99%) | Partly correct (51–75%) | Partly incorrect (26–50%) | Mostly incorrect (1–26%) | Absolutely incorrect (0%) |
|---|---|---|---|---|---|---|---|
| 01. Irrigated 100%—Rice/Rice | 60 | 73.8 | 7.3 | 3.0 | 5.7 | 5.7 | 4.5 |
| 02. Irrigated 100%—Rice/Rice or Rice/Other | 34 | 68.7 | 5.1 | 1.3 | 11.5 | 11.5 | 1.9 |
| 03. Irrigated 100%—Rice | 61 | 64.9 | 0.3 | 2.3 | 15.6 | 15.6 | 1.2 |
| 04. Irrigated ≥60% / Rainfed ≤40%—Rice/Rice or Other | 4 | 71.8 | 8.3 | 0.0 | 9.9 | 9.9 | 0.0 |
| 05. Irrigated ≤30% / Rainfed ≥70%—Rice/Rice or Rice/Other | 38 | 69.1 | 2.0 | 0.0 | 11.2 | 11.2 | 6.5 |
| 06. Upland ≥80% / Rainfed ≤10% / Irrigated ≤10%—Rice | 34 | 62.1 | 7.9 | 0.0 | 15.0 | 15.0 | 0.1 |
| 07. Rainfed ≥60% / Irrigated ≤40%—Rice/Rice | 3 | 71.6 | 0.0 | 0.0 | 11.6 | 11.6 | 5.3 |
| 08. Rainfed ≥90% / Irrigated ≤10%—Rice | 9 | 73.8 | 0.0 | 0.0 | 10.1 | 10.1 | 6.0 |
| 09. Rainfed 100%—Rice | 10 | 66.7 | 0.0 | 0.0 | 13.9 | 13.9 | 5.4 |
| 10. Deepwater 100%—Rice/Rice | 5[a] | 100.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 11. Deepwater 100%—Water/Rice | 5[a] | 100.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 12. Wetlands 100%—Rice/Rice | 5[a] | 100.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Total for Sample Size; Average for other columns | 268 | 76.9 | 2.6 | 0.6 | 8.7 | 8.7 | 2.6 |

[a]Extracted from other secondary sources and high-resolution images.

$$\text{Errors of commission for the rice area class} = E_c = 100 \times \frac{\text{NRFPRA}}{\text{TNRFP}}, \qquad (5)$$

$$\text{Errors of omission for the rice area class} = E_o = 100 \times \frac{\text{RFPNRA}}{\text{TRFP}}, \qquad (6)$$

where RFPCRA = rice field plots classified as rice areas (number), TRFP = total rice field plots (number), NRFPRA = nonrice field-plot points classified as rice area (number), TNRFP = total nonrice field plots (number), and RFPNRA = rice field plots classified as nonrice areas (number).

## 4.2 Accuracy Assessment Based on Correlations between Census-Derived Rice Areas versus MODIS-Derived Rice Areas

The field-plot data points were used to compute fractional rice areas for each rice class, that is, the proportion of the "rice" MODIS pixel that was planted to rice, and this in turn was multiplied by the number of MODIS rice pixels per subnational unit to derive a physical rice area estimate to compare to published rice area statistics. This combined the strengths of both "accuracy-assessment" approaches. We performed the comparison at the regional, national, state, and district levels.

## 5 Results and Discussion

In this section, we focus on the resulting rice classification, vegetation phenology of various rice classes, the derivation of the rice area fractions per rice class, the classification accuracy assessment based on field-plot data, and a comparison between MODIS rice area estimates, subnational statistics, and other published area estimates.

## 5.1 Rice Map and Area Statistics

Altogether, 12 rice classes were identified and labeled (Fig. 8). The final class name or label (Fig. 8, Table 3) is based on the predominance of a particular rice class (e.g., single- or double-season rice), and the dominant water source (e.g., irrigated or rain-fed). For example, the name for class 1 is "01. Irrigated 100 percent—rice-rice (meaning first-season rice followed by a second rice crop)." This means that rice class is dominated by rice cultivation in both the kharif and rabi seasons, and the area is predominantly irrigated from surface-water sources. This class occurs in the major command areas such as the Ganges, Indus, and Krishna. Similarly, class 9 is labeled "09. Rain-fed 100 percent—rice" because this is an intensely cropped rice class, but heavily dependent on seasonal rains. This class is predominantly found in heavy-rainfall areas such as those in Sri Lanka and some parts of eastern and southern India. Similarly, class 2 is labeled "02. Irrigated 100 percent—rice-rice or rice-other crop," found predominantly in the Indo-Ganges basin command areas and Godavari delta. Class 3 is called "03. Irrigated 100 percent—rice," occurring primarily across Orissa, Chhattisgarh, and fragmented areas in Uttar Pradesh in India and across Pakistan. Classes 4 and 5, "irrigated/rainfed rice areas," predominate in Tamil Nadu and Bangladesh. Class 6 is "06.Upland ≥80 percent/rain-fed ≤10 percent/irrigated ≤10 percent—rice" areas, which are found primarily in Madhya Pradesh, Maharashtra, and fragmented locations across the study areas. Classes 10 and 11 ("Deepwater 100 percent rice-rice" and "Deepwater 100 percent water-rice") and Class 12 ("Wetlands 100 percent rice-rice") are predominant in Bangladesh and fragmented areas along the coastal regions (Fig. 8, Table 3). The spectral separability in the temporal NDVI signatures for each of the rice classes is shown in Fig. 9.
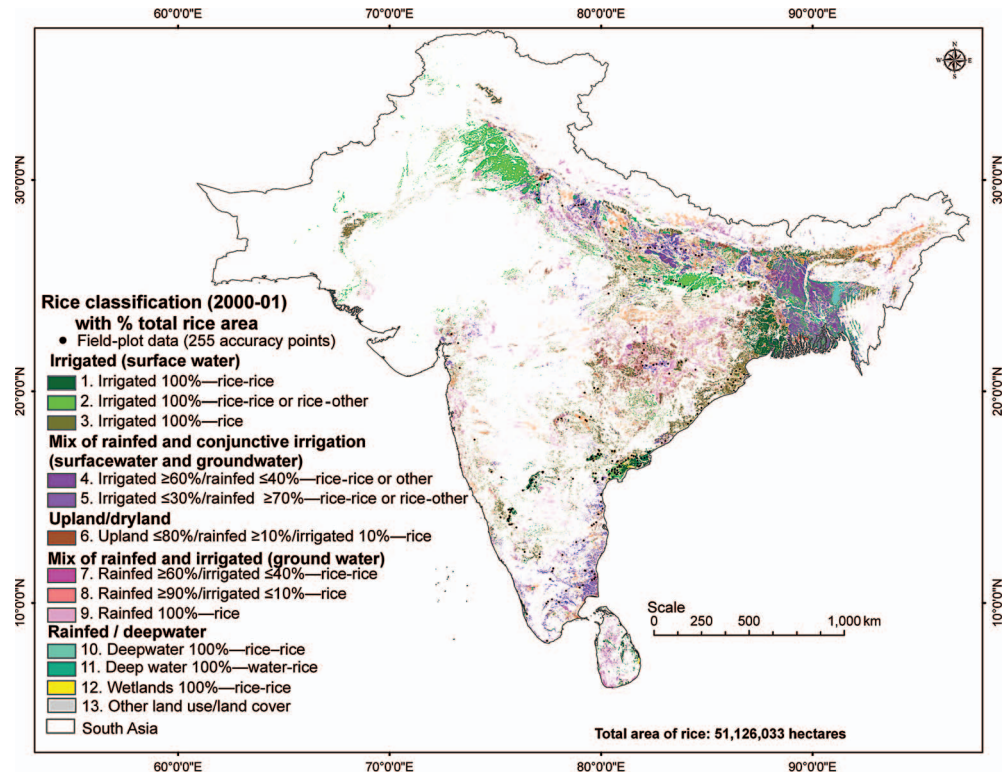
**Fig. 8** The 12 rice classes of six South Asian countries mapped based on MODIS 500-m seven-band time-series data for 2000 to 2001. Note: Area statistics of these classes in Table 3, signatures in Fig. 9.

## 5.2 Vegetation Phenology of Various Rice Classes

Rice-crop phenology was studied using NDVI time-series plots (Fig. 9). These NDVI time-series profiles provided information on (Fig. 8): (*i*) cropping intensities (e.g., single or double crop); (*ii*) crop calendar (i.e., when a crop begins and when it is harvested); and (*iii*) crop health and vigor (indicated by magnitude of NDVI). Each rice class (Fig. 8) has a distinctly different phenology depicted by the NDVI magnitude and/or seasonality (Fig. 9).

The NDVI time series also allows the separation of rain-fed rice from irrigated rice based on factors such as when a crop calendar begins and the magnitude of NDVI. For example, class 2 (Fig. 9) shows a kharif crop beginning around June 20, NDVI peaking around August 15, and the crop harvested by the end of October. The rabi crop begins around November 15, NDVI peaks around February 15, and all crops are harvested by April 15. Around October 15 (Fig. 9), this class 2 has the lowest NDVI and a uniquely high NDVI (compared to all other classes) around December 15. Such distinctive features indicate a unique class with a firm set of characteristics that define that class. This can be said of all classes depicted in Fig. 9.

## 5.3 Rice Area Fractions

Each rice class has several different land cover classes (Fig. 9, Table 3); however, all 12 rice classes have more than a 90% rice-area fraction (RAF) (Table 3). Thus, all 12 rice classes are relatively pure rice classes. However, when calculating actual rice areas (or subpixel areas), we should use RAFs. For example, in class 2 (02. Irrigated 100 percent—rice/rice or rice/other),
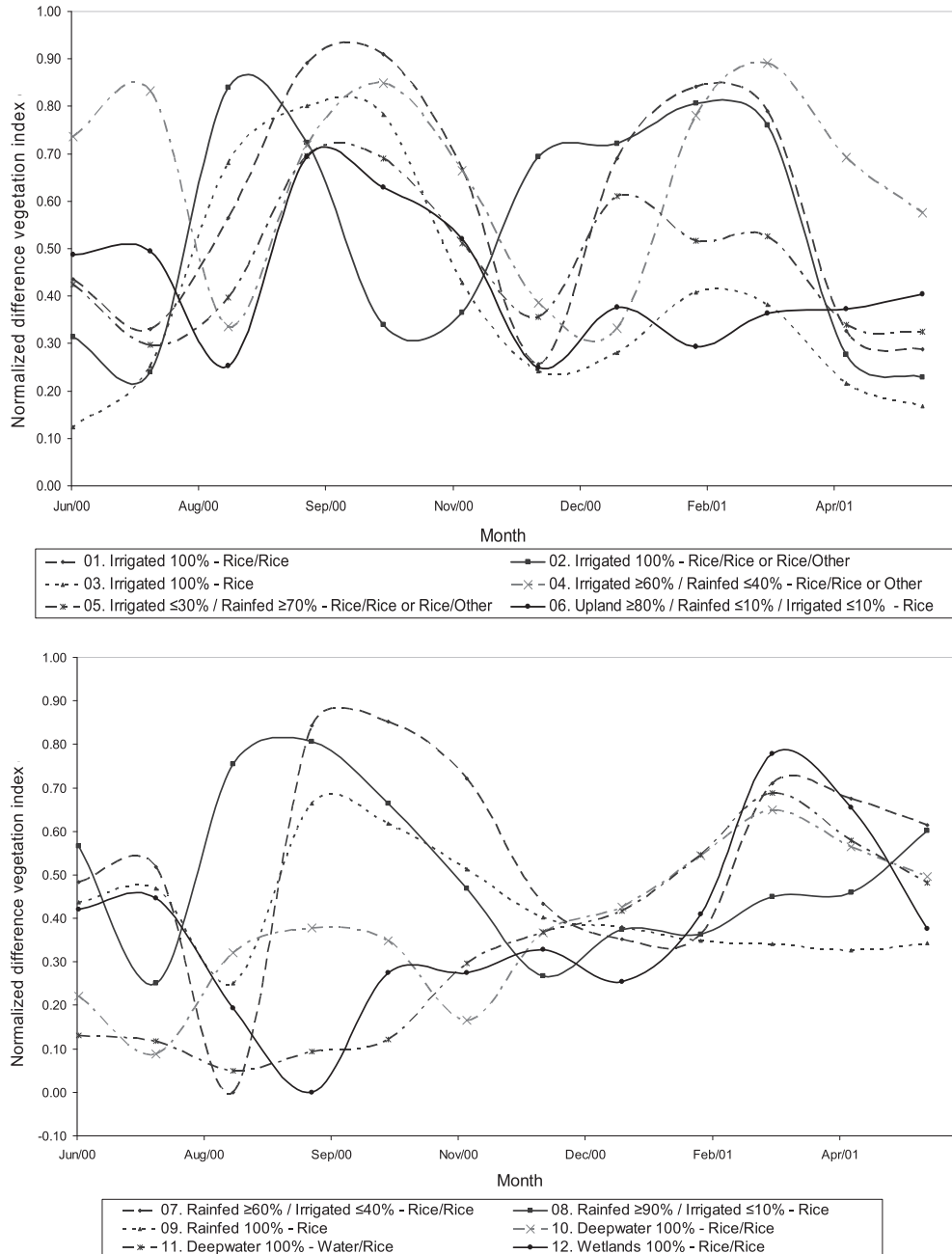
**Fig. 9** The temporal mean MODIS 500-m NDVI signatures of the 12 rice classes derived using data for 2000 to 2001. Note: The 12-class map is in Fig. 8 and area statistics in Table 3.

rice areas (90.3%; Table 3) dominate, but there are other land-cover types, including 1.1% trees, 0.3% shrubs, 1.3% grass, 0.4% built-up area, 2.7% water, and 3.8% other crops (Table 3).

Therefore, the actual areas or SPAs of class 2 in Table 3 = 8,999,360 hectares of FPAs × (93.3/100 of RAFs) = 8,127,089 hectares. Using the same approach, the actual areas (SPAs) of the other 11 rice classes are calculated (Table 3).

## 5.4 Accuracy Assessment

The classification accuracies obtained from the 268 independent field-plot observation data points (15 extracted from other secondary sources and high-resolution images) are summarized in Table 4. The fuzzy classification accuracy varied between 67% and 100% across 12 classes, with an overall accuracy of 80%. However, it must be noted that most rice classes intermix among themselves. Thus, the uncertainty of ∼20% is due to the intermixing among the various rice classes. Thus, rice versus no-rice class accuracy will be very high. The irrigated classes generally have higher classification accuracies than the rain-fed or mixed irrigated/rain-fed classes (Table 4).

## 5.5 Accuracy Assessment Based on Comparison to Subnational Statistics and Other Published Rice-Area Estimates

Table 5 compares the summed RAFs across all classes against the published rice statistics by the six countries. The comparison at the district level (812 spatial units) was performed across South Asia, but, for reasons of space, we report only the tabulated areas at the state level (54 spatial units), and the district-level comparison is shown in Fig. 10. Figure 10 shows the relationship between the MODIS area summarized at district and state levels and the subnational rice areas at the same level of spatial detail. The level of agreement between the MODIS area estimates and the published statistics is very good, 97% at the district level [Fig. 10(a)] and 99% at the state level [Fig. 10(b)]. This excellent match between MODIS-derived rice areas and national and subnational statistics clearly illustrates the high accuracy with which MODIS data have been classified.

These rice maps and statistics produced in this study were a clear advancement over the early rice maps[9–11] in which remote-sensing data were not used. Those early maps were produced by putting together maps and statistics of varying scales and accuracies obtained from different countries and regions using distinctly different approaches to mapping.[9–11] Later rice maps[6,12–14] used remote sensing and were vast improvements over earlier[9–11] maps. However, these later efforts used a single methodological approach (e.g., vegetation indices as in Ref. 6) that left considerable uncertainties.[5,15] As a result, this study used not only multitemporal remote sensing from MODIS but also a suite of methods to decrease uncertainties and achieve a high degree of accuracy in spatial maps and/or statistics derived from them.

## 6 Conclusions

This study demonstrated a suite of methods, approaches, and algorithms to accurately map a rice "footprint" and classify rice areas using the temporal profile and temporal characteristics of the MODIS eight-day, 500-m, seven-band surface-reflectance data across South Asia leading to three main accomplishments. First, a baseline rice map of South Asia with 12 classes (Fig. 8) with their statistics (Table 3) and signatures (Fig. 9) was produced based on data for the years 2000 to 2001. Second, a fuzzy classification accuracy showed that the 12 rice classes were mapped with absolute accuracy ranging from 67% to 100% for individual classes and an overall classification accuracy of 80% for all 12 classes. Almost all the intermixing was between rice classes. The overall classification accuracy when all 12 rice classes were pooled into a single class of rice crop was ∼95%. Furthermore, the accuracy was also determined by correlating the MODIS-derived rice areas with subnational statistics obtained from the six South Asian countries. For this, the $R^2$ values were 97% at the district level and 99% at the state level for 2000 to 2001. Thus, we can state that all the rice classes can be mapped with an accuracy of ∼95% and the 12 individual rice classes can be mapped with accuracy of 67–100%. Third, the results suggest that the methods and approaches including area calculations, data sets, and algorithms used in this study are ideal for rapid, accurate, and large-scale mapping of paddy rice as well as generating their statistics at the national and subnational levels of South Asia.

**Table 5** Comparison of MODIS rice area estimates with subnational statistics, by state.

| Country | State | MODIS rice area ('000 ha) | Subnational statistcs ('000 ha) |
|---|---|---|---|
| Bangladesh | Barisal | 680.1 | 728.4 |
| Bangladesh | Chittagong | 921.9 | 1,138.9 |
| Bangladesh | Dhaka | 1543.5 | 1,562.2 |
| Bangladesh | Khulna | 864.2 | 871.9 |
| Bangladesh | Rajshahi | 1863.7 | 1,917.8 |
| Bangladesh | Sylhet | 665.8 | 568.4 |
| **Bangladesh** | | **6539.3** | **6,787.6** |
| **Bhutan** | | **20.9** | **19.1** |
| India | Andhra Pradesh | 3564.7 | 3,956.3 |
| India | Arunachal Pradesh | 119.1 | 108.2 |
| India | Assam | 2113.7 | 2,646.2 |
| India | Bihar | 3270.3 | 3,656.2 |
| India | Chandigarh | 1.7 | - NA - |
| India | Chhattisgarh | 3540.7 | 3,794.1 |
| India | Dadra and Nagar Haveli | 11.9 | - NA - |
| India | Daman and Diu | 0.7 | - NA - |
| India | Delhi | 6.7 | - NA - |
| India | Goa | 52.7 | 68.0 |
| India | Gujarat | 681.8 | 634.6 |
| India | Haryana | 983.2 | 1,049.0 |
| India | Himachal Pradesh | 90.8 | 79.1 |
| India | Jammu and Kashmir | 272.8 | 244.1 |
| India | Jharkhand | 1408.7 | 1,463.7 |
| India | Karnataka | 1238.6 | 1,481.4 |
| India | Kerala | 333.3 | 346.6 |
| India | Lakshadweep | 0.6 | - NA - |
| India | Madhya Pradesh | 1674.2 | 1,699.0 |
| India | Maharashtra | 1389.4 | 1,511.5 |
| India | Manipur | 180.4 | 212.4 |
| India | Meghalaya | 102.8 | 106.8 |
| India | Mizoram | 44.6 | 67.2 |
| India | Nagaland | 130.0 | 150.5 |
| India | Orissa | 3854.3 | 4,437.3 |
| India | Puducherry | 18.6 | 28.0 |
| India | Punjab | 2567.1 | 2,611.0 |
| India | Rajasthan | 155.9 | 166.1 |
| India | Sikkim | 0.4 | - NA - |
| India | Tamil Nadu | 2125.5 | 2,113.3 |
| India | Tripura | 100.1 | 241.2 |
| India | Uttar Pradesh | 5512.2 | 5,817.5 |
| India | Uttaranchal | 227.7 | 285.5 |
| India | West Bengal | 4540.9 | 5,041.4 |
| **India** | | **40315.8** | **44,016.1** |
| Nepal | Central | 476.1 | 531.8 |
| Nepal | East | 471.5 | 505.5 |
| Nepal | Far-Western | 162.3 | 138.9 |
| Nepal | Mid-Western | 152.2 | 157.7 |
| Nepal | West | 291.7 | 309.6 |
| **Nepal** | | **1553.9** | **1,643.4** |
| Pakistan | Baluchistan | 101.9 | 124.3 |
| Pakistan | Khyber | 12.3 | 50.4 |
| Pakistan | Kashmir | 51.2 | - NA - |
| Pakistan | Punjab_Pak | 1213.9 | 1,217.7 |
| Pakistan | Sind | 585.3 | 613.4 |
| Pakistan | Tribal areas | 10.9 | 11.2 |
| **Pakistan** | | **1975.5** | **2,017.0** |
| Sri Lanka | Central | 41.2 | 43.4 |
| Sri Lanka | Eastern | 131.5 | 117.8 |
| Sri Lanka | North Central | 149.3 | 138.5 |
| Sri Lanka | North Western | 95.3 | 102.4 |
| Sri Lanka | Northern | 55.5 | 38.2 |
| Sri Lanka | Sabaragamuwa | 37.1 | 37.9 |
| Sri Lanka | Southern | 102.9 | 87.3 |
| Sri Lanka | Uva | 52.5 | 33.9 |
| Sri Lanka | Western | 55.3 | 57.1 |
| **Sri Lanka** | | **720.7** | **656.5** |
| **South Asia** | | **51,126.0** | **55,139.6** |
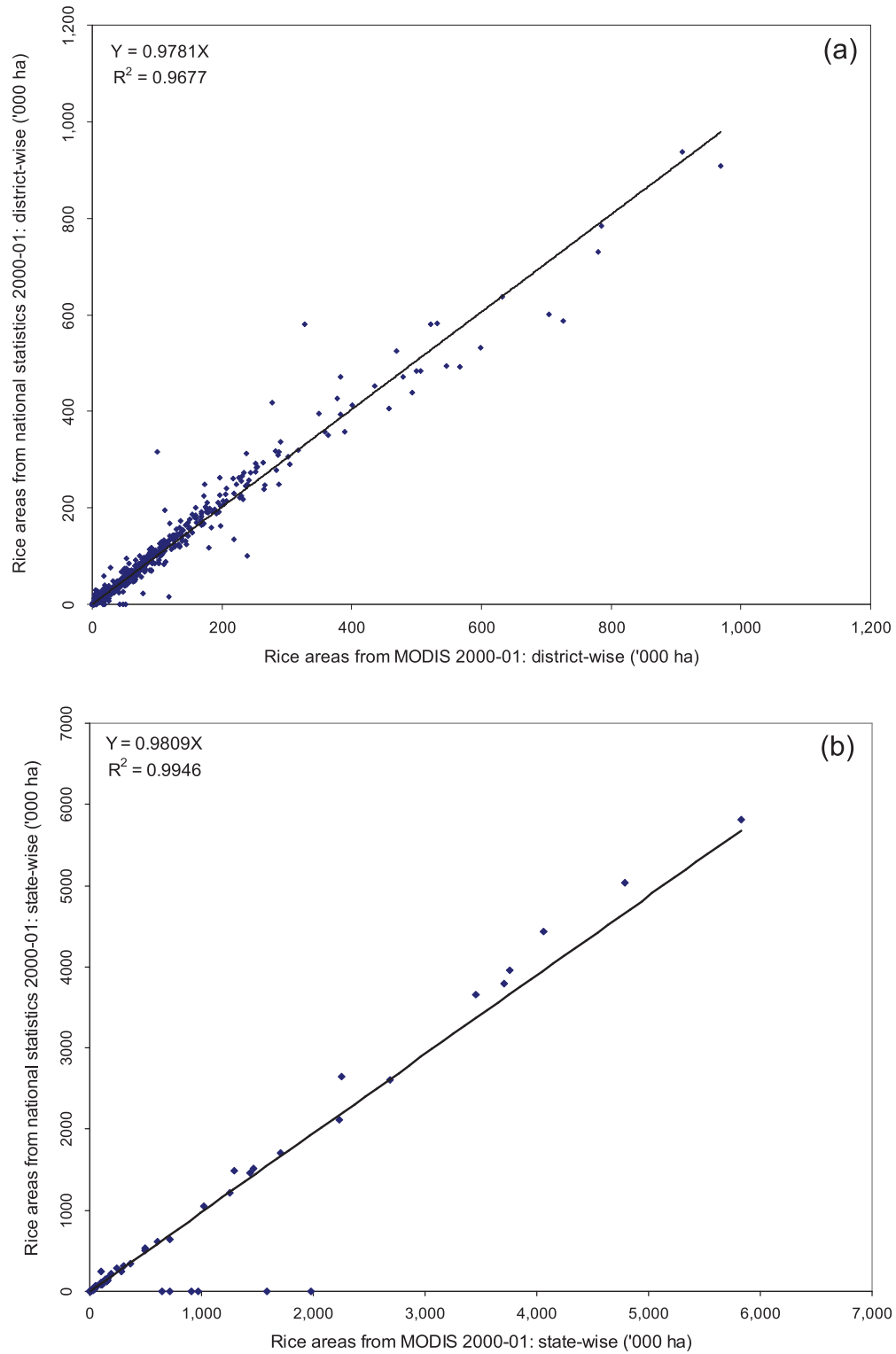
**Fig. 10** Rice areas of South Asia derived using MODIS 500 m compared to agricultural census data for 2000 to 2001: (a) district-wise and (b) state-wise (below). (Administrative boundaries are shown in Fig. 1.)

Mapping rice areas is the first step in characterizing important rice-growing environments of South Asia. Detailed and up-to-date maps and statistics such as these are important inputs for assessing the impact of abiotic stresses, such as droughts and floods, which regularly affect the region and are predicted to increase in frequency and intensity in a changing climate scenario. The study used MODIS data from one good year and demonstrated that it is possible to produce highly accurate maps and statistics of rice maps using a suite of methods. The need to test the rice maps and statistics using approaches and methods advocated in this study using data for non-normal years has been recognized.[41]

## Acknowledgments

## References

1. USDA, United States Department of Agriculture. Foreign Agricultural Service. Production, Supply and Distribution Online. www.fas.usda.gov/psdonline/psdAvailability.aspx (accessed June 21, 2010).
2. World Bank. PovcalNet, http://iresearch.worldbank.org/PovcalNet/povcalNet.html (accessed June 25, 2010).
3. FAO. FAOSTAT, http://faostat.fao.org/ (accessed June 21, 2010).
4. M. K. Papademetriou, "Rice production in the Asia-Pacific region: issues and perspectives," in *Bridging the Rice Yield Gap in the Asia-Pacific Region*, M. K. Papademetriou, F. J. Dent and E. M. Herath, Eds., Bangkok, Food and Agriculture Organization of the United Nations, p. 220 (2000).
5. S. Frolking, J. Qiu, S. Boles, X. Xiao, J. Liu, Y. Zhuang, C. Li, and X. Qin, "Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China," *Global Biogeochem. Cycles* **16**, 3801–3810 (2002).
6. X. Xiao, S. Boles, S. Frolking, C. Li, J. Y. Babu, W. Salas, and B. Moore III, "Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images," *Remote Sens. Environ.* **100**, 95–113 (2006).
7. H. Denier Van Der Gon, "Changes in CH4 emission from rice fields from 1960s to 1990s: 1. impacts of modern rice technology," *Global Biogeochem. Cycles* **1**, 61–72 (2000).
8. R. Wassmann, R. S. Lantin, H. U. Neue, L. V. Buendia, T. M. Corton, and Y. Lu, "Characterization of methane emissions from rice fields in Asia. III. mitigation options and future research needs," *Nutrient Cycling Agroecosyst.* **58**, 23–36 (2000).
9. R. E. Huke, *Rice Area by Type of Culture: South, Southeast, and East Asia*, International Rice Research Institute, Los Baños, Laguna, Philippines (1982).
10. R. E. Huke and E. H. Huke, *Rice Area by Type of Culture: South, Southeast, and East Asia, A Revised and Updated Data Base*, International Rice Research Institute, Los Baños, Laguna, Philippines (1997).
11. I. Aselman and P. J. Crutzen, "Global distribution of natural freshwater wetlands and rice paddies, their net primary productivity, seasonality and possible methane emissions," *J. Atmos. Chem.* **8**, 307–358 (1989).
12. X. Xiao, S. Boles, J. Liu, D. Zhuang, S. Frolking, C. Li, W. Salas, and B. Moore III, "Mapping paddy rice agriculture in southern China using multi-temporal MODIS images," *Remote Sens. Environ.* **95**, 480–492 (2005).

13. W. Takeuchi and Y. Yasuoka, "Sub-pixel mapping of rice paddy fields over Asia using MODIS time series," in *Remote Sensing of Global Croplands for Food Security*, P. Thenkabail, C. Biradar, and J. G. Lyon, Eds., CRC Press, Boca Raton (2009).

14. H. Sun, J. Huang, A. R. Huete, D. Peng, and F. Zhang, "Mapping paddy rice with multi-date moderate-resolution imaging spectroradiometer (MODIS) data in China," *J. Zhejiang Univ. Sci. A* **10**, 1509–1522 (2009).

15. S. Frolking, J. B. Yeluripati, and E. Douglas, "New district-level maps of rice cropping in India: a foundation for scientific input into policy assessment," *Field Crops Res.* **98**, 164–177 (2006).

16. P. S. Thenkabail, M. Schull, and H. Turral, "Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data," *Remote Sens. Environ.* **95**, 24 (2005).

17. P. S. Thenkabail, C. Biradar, P. Noojipady, V. Dheeravath, Y. Li, M. Velpuri, M. Gumma, G. Reddy, H. Turral, X. Cai, J. Vithanage, M. Schull, and R. Dutta, "Global irrigated area map (GIAM) for the end of the last millennium derived from remote sensing," *Int. J. Remote Sens.* **30**(14), 3679–3733 (2009).

18. P. S. Thenkabail, G. Lyon, H. Turral, and C. Biradar, "Remote sensing of global croplands for food security," CRC Press, Boca Raton, p. 556 (2009).

19. P. S. Thenkabail, P. Gangadhara Rao, T. Biggs, M. Krishna, and H. Turral, "Spectral matching techniques to determine historical land use/land cover (LULC) and irrigated areas using time-series AVHRR pathfinder datasets in the Krishna River Basin, India," *Photogramm. Eng. Remote Sens.* **73**(9), 1029–1040 (2007).

20. H. Choice, "Agroecological zones," http://harvestchoice.org/production/biophysical/agroecology. (accessed January 5, 2010).

21. A. Mukherji, T. Facon, J. Burke, C. de Fraiture, J.-M. Faurès, B. Füleki, M. Giordano, D. Molden, and T. Shah, "Revitalizing Asia's irrigation: to sustainably meet tomorrow's food needs," International Water Management Institute; Food and Agriculture Organization of the United Nations, Colombo, Sri Lanka: Rome, Italy (2009).

22. National Aeronautics and Space Administration (NASA). Moderate Resolution Imaging Spectrometer (MODIS). Available online: http://modis.gsfc.nasa.gov/data/dataprod/index.php (accessed on January 21, 2010).

23. R. G. Congalton and K. Green, "*Assessing the accuracy of remotely sensed data: principles and practices*," Lewis, New York (1999).

24. C. Tucker, D. Grant, and J. Dykstra, "NASA's global orthorectified Landsat data set," *Photogramm. Eng. Remote Sens.* **70**, 313–322 (2005), http://zulu.ssc.nasa.gov/mrsid.

25. B. Rabus, M. Eineder, A. Roth, and R. Bamler, "The Shuttle Radar Topography Mission—a new class of digital elevation models acquired by spaceborne radar," *Photogramm. Eng. Remote Sens.* **57**(2), 241–262 (2003).

26. T. Farr and M. Kobrick, "Shuttle Radar Topography Mission produces a wealth of data, *EOS Trans*.," *AGU* **81**, 583–585 (2000).

27. T. Farr, P. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf, "*The Shuttle Radar Topography Mission*," *Rev. Geophys.* **45**, RG2004 (2007).

28. E. Rodriguez, C. Morris, J. Belz, E. Chapin, J. Martin, W. Daffer, and S. Hensley, "An assessment of the SRTM topographic products," Tech. Rep. No. JPL D-31639, Jet Propulsion Laboratory, Pasadena (2005).

29. U.S. Geological Survey (USGS). Shuttle Radar Topography Mission (SRTM) "Finished" 3-arc second SRTM Format Documentation. Available online: http://gcmd.nasa.gov/records/GCMD_DMA_DTED.html (accessed on November 6, 2006).

30. Google Earth. Available online: http://www.google.com/earth/index.html (accessed on October 17, 2010).

31. Ministry of Agriculture and National Informatics Centre (NIC). Available online: http://dacnet.nic.in/rice (accessed on February 9, 2010).
32. Tropical Rainfall Measuring Mission (TRMM), rainfall data. Available online: http://trmm.gsfc.nasa.gov (accessed on October 17, 2011).
33. ERDAS, *ERDAS Field Guide*, Vol. 1 (October 2007).
34. R. S. DeFries, M. Hansen, J. R. G. Townshend, and R. Sohlberg, "Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers," *Int. J. Remote Sens.* **19**, 3141–3168 (1998).
35. S. Homayouni and M. Roux, "Material mapping from hyperspectral images using spectral matching in urban area," presented at IEEE Workshop in Honour of Prof. Landgrebe, Washington, DC (October 2003).
36. R. Tomlinson, *Thinking about Geographic Information Systems Planning for Managers*, ESRI Press (2003).
37. C. Tomlin, *Geographic Information Systems and Cartographic Modeling*, Prentice-Hall, Englewood Cliffs, NJ (1990).
38. D. Peuquet and D. Marble, *Introductory Readings in Geographic Information Systems*, Taylor and Francis, New York (1990).
39. M. K. Gumma, P. S. Thenkabail, I. V. Murali Krishna, N. M. Velpuri, T. P. Gangadhara Rao, V. Dheeravath, C. M. Biradar, S. A. Nalam, and A. Gaur, "Changes in agricultural cropland areas between a water-surplus year and water-deficit year impacting food security determined using MODIS 250 m time-series data and spectral matching techniques in the Krishna River Basin (India)," *Int. J. Remote Sens.* **32**(12), 3679–3733 (2011).
40. S. Gopal and C. Woodcock, "Theory and methods for accuracy assessment of thematic maps using fuzzy sets," *Photogramm. Eng. Remote Sens.* **60**, 181–188 (1994).
41. M. K. Gumma, D. Gauchan, A. Nelson, S. Pandey, and A. Rala, "Temporal changes in rice-growing area and their impact on livelihood over a decade: a case study of Nepal," *Agric. Ecosyst. Environ.* **142**(3-4), 382–392 (2011).

**Murali Krishna Gumma** is a postdoctoral fellow at the International Rice Research Institute, Los Baños, Philippines. He is the author of more than 30 peer-reviewed scientific papers and book chapters in two books. Over the last 10 years, he has been working on a wide range of natural resource and agricultural issues within the CGIAR. His current research interests include global-level crop mapping, land remote sensing identifying the best sites in inland valley wetland rice cultivation in Africa, land use and land cover change, cropping intensity, climate change, and ecology of infectious diseases.

**Andrew Nelson** is a scientist and geographer in the Social Sciences Division of the International Rice Research Institute (IRRI), Los Baños, Philippines. He received his BEng from the University of Nottingham, his MSc from the University of Leicester, and his PhD in geography from the University of Leeds, all in the UK. Over the last 14 years, he has conducted spatial research on a wide range of natural resource and agricultural issues within the CGIAR, the World Bank, FAO, UNEP, and the European Commission. He is now head of the GIS/RS group at IRRI, responsible for the spatial characterization and mapping of rice-growing areas in Asia.

**Prasad S. Thenkabail** is a research geographer at the U.S. Geological Survey. His roles include being a lead researcher for a number of projects (e.g., hyperspectral remote sensing, irrigated crop-land water productivity in California, global crop lands and their water use), coordinator (2010 to present) of the Committee for Earth Observation Systems (CEOS) Agriculture Societal Beneficial Area (SBA), and a Science Advisor (2010 to present) to the Land Surface Imaging Constellation for CEOS, GEO, and GEOSS. He co-leads an IEEE "Water for the World" project. He is also an adjunct professor, Department of Soil, Water, and Environmental Science (SWES), University of Arizona (USA). He has 25+ years' experience working as a well-recognized international expert in remote sensing and geographic information systems (RS/GIS) and their application to agriculture, wetlands, natural resource management, water resources, forests,

sustainable development, and environmental studies. His work experience spans over 25+ countries spread across West and Central Africa (Republic of Benin, Burkina Faso, Cameroon, Central African Republic, Côte d'Ivoire, Gambia, Ghana, Mali, Nigeria, Senegal, and Togo), southern Africa (Mozambique, South Africa), South Asia (Bangladesh, India, Myanmar, Nepal, and Sri Lanka), Southeast Asia (Cambodia), the Middle East (Israel, Syria), East Asia (China), Central Asia (Uzbekistan), North America (the United States), South America (Brazil), and the Pacific (Japan). He has led the remote-sensing programs at the International Water Management Institue, 2003 to 2008; International Center for Integrated Mountain Development, 1995 to 1997; and International Institute of Tropical Agriculture, 1992 to 1995. He was a remote-sensing scientist at the Center for Earth Observation, Yale University (1997 to 2003), and the National Remote Sensing Agency of Indian Space Research Organization during 1986 to 1988. He got his PhD from the Ohio State University in 1992 and has his MS in hydraulics and water resources engineering and BS in civil engineering from Mysore University, India. He served as the chief architect of numerous international remote-sensing science and applications projects and has trained and built capacity in remote sensing and GIS at all the international institutes where he worked. In 2008, Prasad and co-authors were the Second Place Recipients of the 2008 John I. Davidson ASPRS President's Award for practical papers (for their paper on Spectral Matching Techniques used in mapping global irrigated areas). He won the 1994 Autometric Award of the American Society of Photogrammetric Engineering and Remote Sensing (ASPRS) for superior publications in remote sensing. In 1997 he was awarded the Special Achievement in GIS by ESRI in San Diego. He was on a three-member scientific advisory board of Rapideye (a German Satellite Company during its design phase), helping them design the best wavebands for studying agriculture. Rapideye is now a constellation of 5 satellites orbiting the Earth. He is a member of the Landsat Science team and is on the editorial board of *Remote Sensing of Environment* and *The Journal of Remote Sensing*. He is the editor-in-chief of two pioneering books: *Remote Sensing of Global Croplands for Food Security* (Taylor and Francis, 2009) and *Hyperspectral Remote Sensing of Vegetation* (Taylor and Francis, expected publication, October, 2011).

**Amrendra N. Singh** is a consultant in the Social Sciences Division in charge of variety tracking at the International Rice Research Institute. His current research interests include salinity research, flood-related problems, changes in cropping pattern, and characterization of drought-prone areas in eastern India. He is the author of more than 60 peer-reviewed scientific papers and several book chapters. He has 35+ years' experience working as a well-recognized international expert in remote sensing and geographic information systems (RS/GIS) and their application to agriculture and natural resource management.